Robustness and Perturbations of Minimal Bases

Paul Van Dooren[*]   and   Froilán M. Dopico[†]

December 9, 2016

Abstract

Polynomial minimal bases of rational vector subspaces are a classical concept that plays an important role in control theory, linear systems theory, and coding theory. It is a common practice to arrange the vectors of any minimal basis as the rows of a polynomial matrix and to call such matrix simply a minimal basis. Very recently, minimal bases, as well as the closely related pairs of dual minimal bases, have been applied to a number of problems that include the solution of general inverse eigenstructure problems for polynomial matrices, the development of new classes of linearizations and ℓ-ifications of polynomial matrices, and backward error analyses of complete polynomial eigenstructure problems solved via a wide class of strong linearizations. These new applications have revealed that although the algebraic properties of minimal bases are rather well understood, their robustness and the behavior of the corresponding dual minimal bases under perturbations have not yet been explored in the literature, as far as we know. Therefore, the main purpose of this paper is to study in detail when a minimal basis $M(\lambda)$ is robust under perturbations, i.e., when all the polynomial matrices in a neighborhood of $M(\lambda)$ are minimal bases, and, in this case, how perturbations of $M(\lambda)$ change its dual minimal bases. In order to study such problems, a new characterization of whether or not a polynomial matrix is a minimal basis in terms of a finite number of rank conditions is introduced and, based on it, we prove that polynomial matrices are generically minimal bases with very specific properties. In addition, some applications of the results of this paper are discussed.

Key words. backward error analysis, dual minimal bases, genericity, linearizations, ℓ-ifications, minimal bases, minimal indices, perturbations, polynomial matrices, robustness, Sylvester matrices

AMS subject classification. 15A54, 15A60, 15B05, 65F15, 65F35, 93B18

1   Introduction

Minimal bases, formed by vectors with polynomial entries, of a rational vector subspace were made popular in standard references of control theory and linear systems theory as those of Wolovich [23], Forney [13], and Kailath [17], although all three of them cite earlier work for some theoretical developments on the so-called *minimal polynomial bases*. For instance, one can read in [17, p. 460] the following sentence: *I.C. Gohberg pointed out to the author that the significance of minimal bases was perhaps first realized by J. Plemelj in 1908 and then substantially developed in 1943 by N.I. Mushkelishvili and N.P. Vekua.* This means that this paper deals with an almost 110 years old classical mathematical notion. However, the discovery of the importance of this concept in applications had to wait until the 1970s, when the contributions of authors such as Wolovich,

Forney, Kailath, and others, provided computational schemes for constructing a minimal basis from an arbitrary polynomial basis, and showed the key role that this notion plays in multivariable linear systems. These systems could be modeled by rational matrices, polynomial matrices, or linearized state-space models, and had tremendous potential for solving analysis and design problems in control theory as well as in coding theory. The reader is referred to [13, 17] and the references therein for more information on minimal bases and their applications, and also to the brief revision included in the Section 2 of this paper. Moreover, many papers have been published after [17] on the computation and applications of minimal bases and some of them can be found in the references included in [3].

Very recently, minimal bases, and the closely related notion of pairs of dual minimal bases, have been applied to the solution of some problems that have attracted the attention of many researchers in the last fifteen years. For instance, minimal bases have been used (1) in the solution of inverse complete eigenstructure problems for polynomial matrices (see [7, 8] and the references therein), (2) in the development of new classes of linearizations and ℓ-ifications of polynomial matrices [5, 9, 10, 11, 12, 18, 22], which has allowed to recognize that many important linearizations commonly used in the literature are constructed via dual minimal bases (including the classical Frobenius companion forms), (3) in the explicit construction of linearizations of rational matrices [1], and (4) in the backward error analysis of complete polynomial eigenvalue problems solved via the so-called "block Kronecker linearizations" of polynomial matrices [10, Section 6], which include the interesting class of Fiedler linearizations (see [5, 10] for references on this class of linearizations), but *do not include most of the linearizations and ℓ-ifications that can be constructed from minimal bases* [5, 9, 10, 11, 12, 18, 22]. See also [19] for additional references on the role played by dual minimal bases in backward error analyses.

The backward error analysis in [10, Section 6] uses heavily the following two key ideas: (1) that the very particular minimal bases of degree one which are blocks of the block Kronecker linearizations are robust under perturbations, in the sense that all the polynomial matrices in a neighborhood of these minimal bases are also minimal bases with similar properties; and (2) that if these particular minimal bases are perturbed by a certain magnitude, then a particular choice of their dual minimal bases changes essentially by the same magnitude. These results have made clear that for extending the backward error analysis in [10, Section 6] to much more general contexts that include many other classes of linearizations and also ℓ-ifications of polynomial matrices, it is necessary to study in depth the robustness of general minimal bases and the behaviour of their dual minimal bases under perturbations, questions that have not been considered so far in the literature.

To solve the robustness and perturbation problems for minimal bases discussed in the previous paragraph is the main goal of this paper, which is achieved in Sections 6 and 7. We emphasize that the solution of these problems is based on a number of additional results that, in our opinion, are by themselves interesting contributions to the theory and applications of minimal bases. For instance, a new characterization of minimal bases in terms of Sylvester matrices is presented in Section 3, and we prove in Section 5 that polynomial matrices are generically minimal bases with very specific properties that are encoded in the concept of polynomial matrices with full-Sylvester-rank, which are studied in Section 4. In Section 8, the standard rank characterization of minimal bases is connected to the results in this paper. This work is completed with a discussion of some preliminary applications in Section 9 and the conclusions are presented in Section 10.

2 Preliminaries

The results in Sections 2, 3, and 4 of this paper hold for an arbitrary field $\mathbb{F}$, while in the remaining sections $\mathbb{F}$ is the field of real numbers $\mathbb{R}$ or of complex numbers $\mathbb{C}$, which will be simultaneously denoted by $\mathbb{F}$. We adopt standard notations in the area: $\mathbb{F}[\lambda]$ denotes the ring of polynomials in the variable λ with coefficients in $\mathbb{F}$ and $\mathbb{F}(\lambda)$ denotes the field of fractions of $\mathbb{F}[\lambda]$, also known as the field of rational functions over $\mathbb{F}$. Vectors with entries in $\mathbb{F}[\lambda]$ are called polynomial vectors. In addition, $\mathbb{F}[\lambda]^{m \times n}$ stands for the set of $m \times n$ polynomial matrices, and $\mathbb{F}(\lambda)^{m \times n}$ for the set of

$m \times n$ rational matrices. The *degree* of a polynomial vector, $v(\lambda)$, or matrix, $P(\lambda)$, is the highest degree of all its entries and is denoted by $\deg(v)$ or $\deg(P)$. The rigorous definition of genericity we adopt requires to work inside the vector space over $\mathbb{F}$ of $m \times n$ polynomial matrices with degree at most d, which is denoted by $\mathbb{F}[\lambda]_d^{m \times n}$. Finally, $\overline{\mathbb{F}}$ denotes the algebraic closure of $\mathbb{F}$, I_n the $n \times n$ identity matrix, and $0_{m \times n}$ the $m \times n$ zero matrix, where the sizes will be omitted when they are clear from the context.

Polynomial matrices with size $m \times n$ and degree at most d are represented in this paper in the monomial basis as $P(\lambda) = \sum_{i=0}^d P_i \lambda^i$, where $P_i \in \mathbb{F}^{m \times n}$. If $P_d \neq 0$, then the degree of $P(\lambda)$ is precisely d, that is $\deg(P) = d$. The rank of $P(\lambda)$ (sometimes called "normal rank") is just the rank of $P(\lambda)$ considered as a matrix over the field $\mathbb{F}(\lambda)$, and is denoted by $\mathrm{rank}(P)$. The finite eigenvalues of $P(\lambda)$ are the roots of its invariant polynomials, and associated to each such eigenvalue are elementary divisors of $P(\lambda)$. These and the rest of concepts used in this paper can be found in the classical books [14, 15, 17], as well as in the summaries included in [6, Sect. 2] and in [7, Sect. 2], which are recommendable since are brief and follow exactly the conventions in this paper.

Since "minimal basis" is the key concept of this paper, we revise its definition, characterization, main properties, and related notions. It is well known that every rational vector subspace $\mathcal{V}$, i.e., every subspace $\mathcal{V} \subseteq \mathbb{F}(\lambda)^n$ over the field $\mathbb{F}(\lambda)$, has bases consisting entirely of polynomial vectors. Among them some are minimal in the following sense introduced by Forney [13]: a *minimal basis* of $\mathcal{V}$ is a basis of $\mathcal{V}$ consisting of polynomial vectors whose sum of degrees is minimal among all bases of $\mathcal{V}$ consisting of polynomial vectors. The fundamental property [13, 17] of such bases is that the ordered list of degrees of the polynomial vectors in any minimal basis of $\mathcal{V}$ is always the same. Therefore, these degrees are an intrinsic property of the subspace $\mathcal{V}$ and are called the *minimal indices* of $\mathcal{V}$. This discussion immediately leads us to the definition of the minimal bases and indices of a polynomial matrix. An $m \times n$ polynomial matrix $P(\lambda)$ with rank r smaller than m and/or n has non-trivial left and/or right rational null-spaces, respectively, over the field $\mathbb{F}(\lambda)$:

$$\mathcal{N}_\ell(P) := \left\{ y(\lambda)^T \in \mathbb{F}(\lambda)^{1 \times m} : y(\lambda)^T P(\lambda) \equiv 0^T \right\},$$
$$\mathcal{N}_r(P) := \left\{ x(\lambda) \in \mathbb{F}(\lambda)^{n \times 1} : P(\lambda) x(\lambda) \equiv 0 \right\}.$$

Polynomial matrices with non-trivial left and/or right null-spaces are called *singular* polynomial matrices. If the rational subspace $\mathcal{N}_\ell(P)$ is non-trivial, it has minimal bases and minimal indices, which are called the *left minimal bases and indices* of $P(\lambda)$. Analogously, the *right minimal bases and indices* of $P(\lambda)$ are those of $\mathcal{N}_r(P)$, whenever this subspace is non-trivial.

The definition of minimal basis given above cannot be easily handled in practice. Therefore, [13, p. 495] includes five equivalent characterizations of minimal bases. Among those, we present in Theorem 2.2 only the one we believe is the most useful in practice, since relies on the ranks of constant matrices. The statement of Theorem 2.2 requires to introduce first Definition 2.1. For brevity, we use the expression "column (resp., row) degrees" of a polynomial matrix to mean the degrees of its column (resp., row) vectors.

Definition 2.1. *Let $d'_1, \ldots, d'_n$ be the column degrees of $N(\lambda) \in \mathbb{F}[\lambda]^{m \times n}$. The highest-column-degree coefficient matrix of $N(\lambda)$, denoted by N_{hc}, is the $m \times n$ constant matrix whose jth column is the vector coefficient of $\lambda^{d'_j}$ in the jth column of $N(\lambda)$. The polynomial matrix $N(\lambda)$ is said to be* column reduced *if N_{hc} has full column rank.*

Similarly, let $d_1, \ldots, d_m$ be the row degrees of $M(\lambda) \in \mathbb{F}[\lambda]^{m \times n}$. The highest-row-degree coefficient matrix of $M(\lambda)$, denoted by M_{hr}, is the $m \times n$ constant matrix whose jth row is the vector coefficient of λ^{d_j} in the jth row of $M(\lambda)$. The polynomial matrix $M(\lambda)$ is said to be row reduced *if M_{hr} has full row rank.*

Theorem 2.2 provides the announced characterization of minimal bases proved in [13].

Theorem 2.2. *The columns (resp., rows) of a polynomial matrix $N(\lambda)$ over a field $\mathbb{F}$ are a minimal basis of the subspace they span if and only if $N(\lambda_0)$ has full column (resp., row) rank for all $\lambda_0 \in \overline{\mathbb{F}}$, and $N(\lambda)$ is column (resp., row) reduced.*

Remark 2.3. In this paper we follow the convention in [13] and often say, for brevity, that a $p \times q$ polynomial matrix $N(\lambda)$ is a *minimal basis* if the columns (when $q < p$) or rows (when $p < q$) of $N(\lambda)$ form a minimal basis of the rational subspace they span.

Remark 2.4. If $M(\lambda) \in \mathbb{F}[\lambda]^{m \times k}$ is a row reduced polynomial matrix, then $M(\lambda)$ has full row (normal) rank. This can be seen as follows: let $d_i, i = 1, \ldots, m$, be the row degrees of $M(\lambda)$ and let $S(\lambda)$ be an $m \times m$ submatrix of $M(\lambda)$ such that the corresponding submatrix of M_{hr}, denoted by S_{hr}, is nonsingular. Then, $\det S(\lambda) = \lambda^{d_1 + \cdots + d_m} \det S_{hr} + $ (lower degree terms in λ) $\neq 0$. This means that the rows of $M(\lambda)$ form a basis of the rational subspace $\mathcal{V}$ they span (which is the row space of $M(\lambda)$) and, by definition of minimal basis, the sum of its row degrees $\sum_{i=1}^{m} d_i$ is an upper bound for the sum of the minimal indices $\sum_{i=1}^{m} \eta_i$ of $\mathcal{V}$. Moreover, equality of the sums implies that $M(\lambda)$ is a minimal basis of $\mathcal{V}$ and that the ordered lists $\{d_i\}$ and $\{\eta_i\}$ are also equal, by the uniqueness of the minimal indices.

Next, we introduce the concept of *dual minimal bases*, which has played a key role in a number of recent applications [7, 9, 10, 18, 22].

Definition 2.5. *Polynomial matrices $M(\lambda) \in \mathbb{F}[\lambda]^{m \times k}$ and $N(\lambda) \in \mathbb{F}[\lambda]^{n \times k}$ with full row ranks are said to be* dual minimal bases *if they are minimal bases satisfying $m + n = k$ and $M(\lambda) N(\lambda)^T = 0$.*

The name "dual minimal bases" seems to be very recent in the literature, because [7] is the first reference where is used, to our knowledge. However, the origins of such concept can be found in [13, Section 6]. According to [13, Section 6], dual minimal bases span rational vector subspaces of $\mathbb{F}(\lambda)^k$ that are dual to each other. In the language of null-spaces of matrix polynomials, observe that $M(\lambda)$ is a minimal basis of $\mathcal{N}_\ell(N(\lambda)^T)$ and that $N(\lambda)^T$ is a minimal basis of $\mathcal{N}_r(M(\lambda))$. As a consequence, the right minimal indices of $M(\lambda)$ are the row degrees of $N(\lambda)$ and the left minimal indices of $N(\lambda)^T$ are the row degrees of $M(\lambda)$. Note that dual minimal bases have been defined with more columns than rows, as in the classical reference [13], although, one can also use matrices with more rows than columns in the definition. It follows from this discussion that for every minimal basis, there exists a minimal basis that is dual to it. In addition, every minimal basis is a minimal basis of some matrix polynomial.

The next theorem is crucial for the rest of this paper. Its first part was proven in [13], while the second (converse) part has been proven very recently in [7].

Theorem 2.6. *Let $M(\lambda) \in \mathbb{F}[\lambda]^{m \times (m+n)}$ and $N(\lambda) \in \mathbb{F}[\lambda]^{n \times (m+n)}$ be dual minimal bases with row degrees $(\eta_1, \ldots, \eta_m)$ and $(\varepsilon_1, \ldots, \varepsilon_n)$, respectively. Then*

$$\sum_{i=1}^{m} \eta_i = \sum_{j=1}^{n} \varepsilon_j. \tag{2.1}$$

Conversely, given any two lists of nonnegative integers $(\eta_1, \ldots, \eta_m)$ and $(\varepsilon_1, \ldots, \varepsilon_n)$ satisfying (2.1), there exists a pair of dual minimal bases $M(\lambda) \in \mathbb{F}[\lambda]^{m \times (m+n)}$ and $N(\lambda) \in \mathbb{F}[\lambda]^{n \times (m+n)}$ with precisely these row degrees, respectively.

3   A finite number of rank conditions for minimal bases

The characterization of a minimal basis given in Theorem 2.2 does not seem tractable at first sight from a numerical point of view, since it requires a rank test over all $\lambda_0 \in \overline{\mathbb{F}}$. In this section we show that this can be reduced to a finite number of rank tests on matrices that can be described from the dimensions and coefficients of $M(\lambda)$. A crucial role will be played here by the so-called Sylvester matrices [2, 4] of a polynomial matrix $M(\lambda)$ of degree at most d,

$$M(\lambda) := M_0 + M_1 \lambda + \cdots + M_d \lambda^d, \quad M_i \in \mathbb{F}^{m \times (m+n)}, \tag{3.1}$$

defined for $k = 1, 2, \ldots$ as

$$
S_k := \begin{bmatrix}
M_0 & & & \\
M_1 & M_0 & & \\
\vdots & M_1 & \ddots & \\
M_d & & \ddots & M_0 \\
0 & M_d & & M_1 \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & M_d
\end{bmatrix}, \quad S_k \in \mathbb{F}^{(k+d)m \times k(m+n)}, \tag{3.2}
$$

$$
\underbrace{}_{k \text{ blocks}}
$$

where S_k has k block columns. In order to avoid trivialities, we assume throughout the paper that $m > 0, n > 0$, and $d > 0$. In certain results involving Sylvester matrices of several polynomial matrices, we will use the notation $S_k(M)$ to indicate that the Sylvester matrices are associated to the polynomial matrix $M(\lambda)$.

The ranks of the matrices S_k in (3.2) are fundamental in this paper and the following two simple lemmas are used very often.

Lemma 3.1. *Let S_k for a given index $k > 1$ have full column rank $r_k := k(m + n)$. Then all submatrices S_ℓ with $1 \leq \ell < k$ also have full column rank $r_\ell := \ell(m + n)$.*

Proof. Each matrix S_ℓ with $1 \leq \ell < k$, appropriately padded with zeros, forms the first ℓ block columns of S_k, from which the rank condition trivially follows. $\square$

Lemma 3.2. *Let S_k with a given index $k > 0$ have full row rank $r_k := (k+d)m$. Then all embedding matrices S_ℓ with $k < \ell$ also have full row rank $r_\ell := (\ell + d)m$.*

Proof. We first prove that the result holds for $\ell = k + 1$. Since S_k has full row rank r_k, its bottom block row $\begin{bmatrix} 0 & | & M_d \end{bmatrix}$ has full row rank m and so does the matrix M_d. Since the matrix S_{k+1} can be partitioned in a block triangular form

$$
S_{k+1} = \left[\begin{array}{c|c} S_k & X \\ \hline 0 & M_d \end{array} \right],
$$

where both S_k and M_d have full row rank, S_{k+1} has also full row rank. The result for general ℓ larger than k, then easily follows by induction. $\square$

In some of the results in this section, we will assume that $M(\lambda) \in \mathbb{F}[\lambda]^{m \times (m+n)}$ has (full) normal rank m so that it has a (right) null-space of dimension n (over the field of rational functions). The connection of this null-space $\mathcal{N}_r(M)$ with the Sylvester matrices in (3.2) comes from the fact that any $N(\lambda) \in \mathbb{F}[\lambda]^{n \times (m+n)}$ such that the columns of $N(\lambda)^T$ belong to $\mathcal{N}_r(M)$ (i.e. $M(\lambda)N(\lambda)^T = 0$) with expansion

$$
N(\lambda) := N_0 + N_1 \lambda + \cdots + N_s \lambda^s, \quad N_i \in \mathbb{F}^{n \times (m+n)}, \tag{3.3}
$$

will satisfy the equations

$$
S_k \begin{bmatrix} N_0^T \\ \vdots \\ N_s^T \\ 0 \end{bmatrix} = 0 \quad \text{for all} \quad k \geq s + 1.
$$

In particular, one can choose $N(\lambda)^T$ to be a minimal basis of $\mathcal{N}_r(M)$ and then to derive certain rank conditions from this. The following theorem is proven in [4] using this type of arguments.

Theorem 3.3. *Let $M(\lambda) \in \mathbb{F}[\lambda]^{m \times (m+n)}$ be a polynomial matrix of full row rank m and let S_k be its Sylvester matrices for $k = 1, 2, \ldots$. Let r_k be the rank of S_k, n_k be the right nullity of S_k, and α_k be the number of right minimal indices ε_i of $M(\lambda)$ equal to k. Then these magnitudes obey the following recursive relations*

$$\alpha_k = (n_{k+1} - n_k) - (n_k - n_{k-1}) = (r_k - r_{k-1}) - (r_{k+1} - r_k), \quad k = 1, 2, \ldots, \tag{3.4}$$

initialized with $r_0 = n_0 = 0$ and $\alpha_0 = m + n - r_1 = n_1$.

The second equality in (3.4) between ranks and nullities easily follows from the identity

$$r_k + n_k = k(m + n), \text{ for } k \geq 1, \tag{3.5}$$

and an intuitive proof of the first identity follows from the recursive definition of the null-spaces of S_k, for $k = 1, 2, \ldots$. For $k = 1$ the theorem says that the matrix

$$S_1 := \begin{bmatrix} M_0 \\ M_1 \\ \vdots \\ M_d \end{bmatrix} \in \mathbb{F}^{(d+1)m \times (m+n)}$$

has a right null space of dimension n_1 equal to α_0, the number of right minimal indices ε_i of $M(\lambda)$ equal to 0. Indeed, every index $\varepsilon_i = 0$ corresponds to a right null vector of degree 0 (i.e. constant) of S_1 and hence also of $M(\lambda)$. For $k = 2$, the matrix

$$S_2 := \begin{bmatrix} M_0 & 0 \\ M_1 & M_0 \\ \vdots & M_1 \\ M_d & \vdots \\ 0 & M_d \end{bmatrix} \in \mathbb{F}^{(d+2)m \times 2(m+n)}$$

has a right null space of dimension n_2 equal to $\alpha_1 + 2\alpha_0$, because S_2 contains S_1 (padded with zeros) twice as a submatrix and hence the null-space of S_1 will contribute also twice to the null-space of S_2. The "additional" α_1 linearly independent vectors in the null-space of S_2 correspond to the "true" vectors of degree 1 in $\mathcal{N}_r(M)$, i.e., those vectors in the null-space of S_2 which cannot be obtained as linear combinations of vectors in the null-space of S_1 padded with zeros. It then follows that $n_2 - n_1 = \alpha_0 + \alpha_1$. One uses the same arguments to show recursively that

$$n_{k+1} - n_k = \sum_{i=0}^{k} \alpha_i, \tag{3.6}$$

which is equivalent to the first equality in (3.4) together with the initializations $n_0 = 0$, $n_1 = \alpha_0$. We refer to [4] and the references therein for a more detailed proof. The following corollary, also given in [4], is then easily derived from this.

Corollary 3.4. *Let $M(\lambda) \in \mathbb{F}[\lambda]^{m \times (m+n)}$ be a polynomial matrix of full row rank m and let S_k be its Sylvester matrices for $k = 1, 2, \ldots$. Let r_k be the rank of S_k and n_k be the right nullity of S_k. If d' is the smallest index k for which*

$$n_{k+1} - n_k = n, \quad \text{or equivalently} \quad r_{k+1} - r_k = m, \tag{3.7}$$

then d' is the maximum right minimal index of $M(\lambda)$ or, equivalently, the maximum column degree of any minimal basis of $\mathcal{N}_r(M)$. Moreover, for all k larger than d', the equalities (3.7) still hold.

6

Proof. As soon as the α_i, the number of right minimal indices of $M(\lambda)$ equal to i, add up to n, we have found a complete polynomial basis of the rational right null-space of $M(\lambda)$. There can be no further linearly independent right null vectors of $M(\lambda)$ or nonzero α_i since the sum $\sum_{i=0}^{k} \alpha_i = n$ is the total number of right minimal indices of $M(\lambda)$. The corresponding largest index k with $\alpha_k \neq 0$ is therefore the largest right minimal index d' of $M(\lambda)$. Since all $\alpha_k = 0$ for k larger than d', equation (3.6) guarantees that the equality (3.7) continues to hold. The equivalent condition $r_{k+1} - r_k = m$ follows from (3.5). $\qquad\square$

We can also express the sum of the right minimal indices ε_j of $M(\lambda)$ as a function of the nullities n_k.

Corollary 3.5. *Let $M(\lambda) \in \mathbb{F}[\lambda]^{m \times (m+n)}$ be a polynomial matrix of full row rank m and let S_k be its Sylvester matrices for $k = 1, 2, \ldots$. Let n_k and r_k be the right nullity and the rank of S_k, respectively, α_k be the number of right minimal indices ε_i of $M(\lambda)$ equal to k, and d' be the maximum right minimal index of $M(\lambda)$. Then*

$$\sum_{j=1}^{n} \varepsilon_j = \sum_{k=1}^{d'} k\alpha_k = \sum_{k=1}^{d'} k(n_{k-1} - 2n_k + n_{k+1}) = n \cdot d' - n_{d'} = r_{d'} - m \cdot d'. \tag{3.8}$$

Proof. The first identity follows from the definition of α_k which is the number of right minimal indices ε_j equal to k. The second identity follows from Theorem 3.3. The third identity follows from the relations of Theorem 3.3, which written as follows

$$n_0 = 0, \quad
\begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{d'} \end{bmatrix}
=
\begin{bmatrix}
1 & & & & \\
-2 & 1 & & & \\
1 & -2 & 1 & & \\
 & \ddots & \ddots & \ddots & \\
 & & 1 & -2 & 1
\end{bmatrix}
\begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_{d'} \\ n_{d'+1} \end{bmatrix}, \tag{3.9}$$

allow us to see immediately that $\sum_{k=1}^{d'} k\alpha_k = d' \cdot n_{d'+1} - (d'+1)n_{d'}$. This equality combined with the identity in Corollary 3.4, i.e.,

$$n_{d'+1} = n_{d'} + n,$$

gives $\sum_{k=1}^{d'} k\alpha_k = n \cdot d' - n_{d'}$. The fourth identity follows from (3.5) applied to $k = d'$. $\qquad\square$

Remark 3.6. Notice that equation (3.9) in the proof of Corollary 3.5 establishes a one to one correspondence between the index sets $\{n_i, i = 1, \ldots, d' + 1\}$ and $\{\alpha_i, i = 0, \ldots, d'\}$, as long as all the indices are non-negative and $\sum_{i=0}^{d'} \alpha_i = n$ holds. This last identity is clearly equivalent to $n_{d'+1} - n_{d'} = n$.

We are now ready to formulate our necessary and sufficient finite number of rank conditions on constant matrices derived from $M(\lambda)$ to verify that it is a minimal basis. The conditions are given in the next theorem, where we note that $d' \leq \sum_{i=1}^{m} d_i$, as a consequence of Corollary 3.4 and Theorem 2.6, and so the number of rank conditions to be checked is indeed finite.

Theorem 3.7. *Let $M(\lambda) \in \mathbb{F}[\lambda]^{m \times (m+n)}$ be a polynomial matrix, let $d_i, i = 1, \ldots, m$, be its row degrees, and M_{hr} be its highest-row-degree coefficient matrix. Let S_k be the Sylvester matrices of $M(\lambda)$ for $k = 1, 2, \ldots$, and let r_k and n_k be the rank and the right nullity of S_k, respectively. Let d' be the smallest index k for which $n_{k+1} = n_k + n$, or equivalently, $r_{k+1} = r_k + m$. Then $M(\lambda)$ is a minimal basis if and only if the following rank conditions are satisfied*

$$\operatorname{rank} M_{hr} = m \quad \text{and} \quad r_{d'} - m \cdot d' = \sum_{i=1}^{m} d_i. \tag{3.10}$$

Proof. Let us assume first that the conditions (3.10) hold. The condition rank $M_{hr} = m$ says that $M(\lambda)$ is row reduced, which implies that $M(\lambda)$ has normal rank equal to its number of rows, m, (recall Remark 2.4) and that it has a right null space of dimension n and a set of n right minimal indices $\varepsilon_j, j = 1, \ldots, n$. These right minimal indices can then be computed via the recurrences of Theorem 3.3 and condition $r_{d'} - m \cdot d' = \sum_{i=1}^m d_i$ implies (according to Corollaries 3.5 and 3.4) that

$$\sum_{i=1}^m d_i = \sum_{j=1}^n \varepsilon_j. \tag{3.11}$$

Since $M(\lambda)$ has normal rank m, its rows form a polynomial basis of the row space of $M(\lambda)$ whose degree sum equals $\sum_{i=1}^m d_i$. If $\eta_1, \ldots, \eta_m$ are the minimal indices of the row space of $M(\lambda)$, then Theorem 2.6 combined with (3.11) imply $\sum_{i=1}^m d_i = \sum_{j=1}^n \varepsilon_j = \sum_{j=1}^m \eta_j$, which in turn implies that the rows of $M(\lambda)$ must form a minimal basis, by definition of minimal basis.

The reverse implication follows immediately: assume that $M(\lambda)$ is a minimal basis. Then rank $M_{hr} = m$ by Theorem 2.2. In addition, Corollaries 3.4 and 3.5 and Theorem 2.6 imply $r_{d'} - m \cdot d' = \sum_{i=1}^m d_i$. $\qquad\square$

Theorem 3.7 is considerably simplified under the additional generic assumption that the leading coefficient of $M(\lambda)$ has full rank. Corollaries 3.8 and 3.9 are two results in that direction.

Corollary 3.8. *Let $M(\lambda) = M_0 + M_1\lambda + \cdots + M_d\lambda^d \in \mathbb{F}[\lambda]^{m \times (m+n)}$ be a polynomial matrix such that* rank $M_d = m$. *Let S_k be the Sylvester matrices of $M(\lambda)$ for $k = 1, 2, \ldots$, let r_k be the rank of S_k, and let d' be the smallest index k such that $r_{k+1} = r_k + m$. Then, $M(\lambda)$ is a minimal basis if and only if $r_{d'} = m(d + d')$.*

Proof. Note that the assumptions of Corollary 3.8 imply that $M_{hr} = M_d$ and that the row degrees $d_1, d_2, \ldots, d_m$ of $M(\lambda)$ are all equal to d and, so, $\sum_{i=1}^m d_i = m \cdot d$. In this scenario, if $M(\lambda)$ is a minimal basis, then it satisfies the second equality in (3.10), which implies $r_{d'} = m(d + d')$. Conversely, if one assumes that $r_{d'} = m(d+d')$ holds, then the second equality in (3.10) is satisfied, while the first one is guaranteed by rank $M_d = m$. Therefore $M(\lambda)$ is a minimal basis. $\qquad\square$

Corollary 3.9. *Let $M(\lambda) = M_0 + M_1\lambda + \cdots + M_d\lambda^d \in \mathbb{F}[\lambda]^{m \times (m+n)}$ be a polynomial matrix and let S_k be the Sylvester matrices of $M(\lambda)$ for $k = 1, 2, \ldots$. Then, $M(\lambda)$ is a minimal basis with* rank $M_d = m$ *if and only if there exists an index k such that S_k has full row rank. In this case, if d' is the smallest index k for which S_k has full row rank, then d' is the largest right minimal index of $M(\lambda)$.*

Proof. If $M(\lambda)$ is a minimal basis with rank $M_d = m$, then Corollary 3.8 implies $r_{d'} = m(d + d')$, which is the number of rows of $S_{d'}$. Then $S_{d'}$ has full row rank.

Conversely, if there exists an index k such that S_k has full row rank, then M_d has full row rank, because the last block row of S_k is $[0 \mid M_d]$, which also implies that $M(\lambda)$ has full row normal rank. Let k_0 be the smallest index k such that S_k has full row rank and denote by r_k the rank of any Sylvester matrix S_k. Then, according to Lemma 3.2, S_{k_0+1} also has full row rank and their ranks satisfy

$$r_{k_0+1} - r_{k_0} = m. \tag{3.12}$$

However, $r_{k_0-1} < (k_0 - 1 + d)m$, because S_{k_0-1} has not full row rank. Therefore, $r_{k_0} - r_{k_0-1} > m$ and, so, $r_{k+1} - r_k > m$ for all $k \leq k_0 - 1$, since Theorem 3.3 implies $r_j - r_{j-1} \geq r_{j+1} - r_j$ for all $j \geq 1$ because $\alpha_j \geq 0$. Therefore, k_0 is the smallest index k such that $r_{k+1} = r_k + m$, that is, $k_0 = d'$ in Corollary 3.8 and $r_{d'} = m(d' + d)$, since $S_{k_0} = S_{d'}$ has full row rank. So, Corollary 3.8 implies that $M(\lambda)$ is a minimal basis. In addition, observe that also $k_0 = d'$ in Corollary 3.4 and, so, $k_0 = d'$ is the largest right minimal index of $M(\lambda)$. $\qquad\square$

We illustrate Theorems 3.3 and 3.7, and Corollaries 3.4, 3.8, and 3.9 with three simple examples. With the purpose that the reader can easily check that these results yield the right outcome, in our three examples Theorem 2.2 also allows us to see very easily whether $M(\lambda)$ is a minimal basis or not. In addition, we display a matrix $N(\lambda)$ such that the columns of $N(\lambda)^T$ are a minimal basis of $\mathcal{N}_r(M)$ (and, so, $N(\lambda)$ is a minimal basis dual to $M(\lambda)$ whenever $M(\lambda)$ is a minimal basis), which can again be easily checked by Theorem 2.2.

Example 3.10. Let $M(\lambda) \in \mathbb{F}[\lambda]^{6\times 8}$ and $N(\lambda) \in \mathbb{F}[\lambda]^{2\times 8}$ be given:

$$M(\lambda) = \begin{bmatrix} -I_2 & \lambda I_2 & & \\ & -I_2 & \lambda I_2 & \\ & & -I_2 & \lambda I_2 \end{bmatrix}, \quad N(\lambda) = \begin{bmatrix} \lambda^3 I_2 & \lambda^2 I_2 & \lambda I_2 & I_2 \end{bmatrix}.$$

Then clearly $M(\lambda)N(\lambda)^T = 0$. The ranks r_k and nullities n_k of the Sylvester matrices S_k of $M(\lambda)$ are:

$$r_1 = 8, r_2 = 16, r_3 = 24, r_4 = 30; \quad n_1 = 0, n_2 = 0, n_3 = 0, n_4 = 2.$$

Then, clearly $d' = 3$ and (3.4) gives $[\alpha_0, \alpha_1, \alpha_2, \alpha_3] = [0, 0, 0, 2]$ (which agrees with the row degrees of $N(\lambda)$). The condition (3.10) becomes rank $M_{hr} = 6$ and $r_3 - 3 \cdot m = 24 - 3 \cdot 6 = 6$, which is indeed equal to $d_1 + d_2 + d_3 + d_4 + d_5 + d_6 = 6$. Therefore, $M(\lambda)$ is a minimal basis with right minimal indices $\{3, 3\}$. Observe that for proving just that $M(\lambda)$ is a minimal basis (not for getting that its right minimal indices are $\{3, 3\}$), one can use simply Corollary 3.9 because the leading coefficient of $M(\lambda)$ has clearly full rank. In this case the numbers of rows of S_k for $k = 1, 2, 3$ are $12, 18, 24$, respectively, which implies that S_3 has full row rank and that the largest right minimal index of $M(\lambda)$ is 3.

The second example corresponds to a polynomial matrix that is not a minimal basis and has three different right minimal indices.

Example 3.11. Let $M(\lambda) \in \mathbb{F}[\lambda]^{4\times 7}$ and $N(\lambda) \in \mathbb{F}[\lambda]^{3\times 7}$ be given:

$$M(\lambda) = \left[\begin{array}{cc|cc|cc|c} \lambda & 0 & & & & & \\ \hline & & -1 & \lambda & & & \\ \hline & & & & -1 & \lambda & \\ \hline & & & & & -1 & \lambda \end{array}\right], \quad N(\lambda) = \left[\begin{array}{cc|cc|ccc} 0 & 1 & & & & & \\ \hline & & \lambda & 1 & & & \\ \hline & & & & \lambda^2 & \lambda & 1 \end{array}\right].$$

Then clearly $M(\lambda)N(\lambda)^T = 0$. The ranks r_k and nullities n_k of the Sylvester matrices S_k of $M(\lambda)$ are:

$$r_1 = 6, r_2 = 11, r_3 = 15; \quad n_1 = 1, n_2 = 3, n_3 = 6.$$

Then clearly $d' = 2$ and (3.4) gives $[\alpha_0, \alpha_1, \alpha_2] = [1, 1, 1]$ (which agrees with the row degrees of $N(\lambda)$). The condition (3.10) becomes rank $M_{hr} = 4$ and $r_2 - 2 \cdot m = 11 - 2 \cdot 4 = 3$, which is not equal to $d_1 + d_2 + d_3 + d_4 = 1 + 1 + 1 + 1 = 4$. Therefore $M(\lambda)$ is not a minimal basis. Observe that for proving just that $M(\lambda)$ is not a minimal basis (not for getting that its right minimal indices are $\{0, 1, 2\}$), one can use simply Corollary 3.8 because the leading coefficient of $M(\lambda)$ has clearly full rank.

In the previous two examples $M(\lambda)$ has degree 1 and all its row degrees equal. The polynomial matrix $M(\lambda)$ in the next example does not satisfy any of these two properties.

Example 3.12. Let $M(\lambda) \in \mathbb{F}[\lambda]^{6\times 8}$ and $N(\lambda) \in \mathbb{F}[\lambda]^{2\times 8}$ be given:

$$M(\lambda) = \begin{bmatrix} -I_2 & \lambda I_2 & & \\ & -I_2 & \lambda I_2 & \\ & & -I_2 & \lambda^2 I_2 \end{bmatrix}, \quad N(\lambda) = \begin{bmatrix} \lambda^4 I_2 & \lambda^3 I_2 & \lambda^2 I_2 & I_2 \end{bmatrix}.$$

Then clearly $M(\lambda)N(\lambda)^T = 0$. The ranks r_k and nullities n_k of the Sylvester matrices S_k of $M(\lambda)$ are:

$$r_1 = 8, r_2 = 16, r_3 = 24, r_4 = 32, r_5 = 38; \quad n_1 = 0, n_2 = 0, n_3 = 0, n_4 = 0, n_5 = 2\,.$$

Then clearly $d' = 4$ and (3.4) gives $[\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4] = [0, 0, 0, 0, 2]$ (which agrees with the row degrees of $N(\lambda)$). The condition (3.10) becomes rank $M_{hr} = 6$ and $r_4 - 4 \cdot m = 32 - 4 \cdot 6 = 8$, which is equal to $d_1 + d_2 + d_3 + d_4 + d_5 + d_6 = 8$. Therefore, $M(\lambda)$ is a minimal basis. In this case the leading coefficient of $M(\lambda)$ (the one corresponding to degree 2) has not full row rank and Corollaries 3.8 and 3.9 cannot be used.

We emphasize that in the examples above the rank conditions are completely in terms of the coefficient matrices of $M(\lambda)$ via its Sylvester matrices, and that the minimal bases $N(\lambda)$ are displayed only for the purpose of comparison.

4 Full-Sylvester-rank polynomial matrices and their properties

In this section we characterize the polynomial matrices of size $m \times (m + n)$ and degree at most d all of whose Sylvester matrices S_k defined in (3.2) have full rank, either full column rank when S_k has more rows than columns or full row rank otherwise. We advance that such matrices are always minimal bases and that satisfy other additional properties. The *ceiling* function of a real number x is often used in the rest of this paper and is denoted by $\lceil x \rceil$. Recall that $\lceil x \rceil$ is the smallest integer that is larger than or equal to x.

The following definition will allow us to refer to the property of interest in this section in a concise way.

Definition 4.1. *Let $M(\lambda) \in \mathbb{F}[\lambda]^{m \times (m+n)}$ be a polynomial matrix of degree at most d, let S_k for $k = 1, 2, \ldots$ be its Sylvester matrices, and let r_k be the rank of S_k. The polynomial matrix $M(\lambda)$ is said to have full-Sylvester-rank if all the matrices S_k have full rank, i.e., if $r_k = \min\{(k+d)m\,,\, k(m+n)\}$ for $k = 1, 2, \ldots$.*

It is necessary and sufficient to check at most two ranks for determining whether a polynomial matrix has full-Sylvester-rank or not, as a consequence of Lemmas 3.1 and 3.2. This is stated in Lemma 4.2.

Lemma 4.2. *Let $M(\lambda) \in \mathbb{F}[\lambda]^{m \times (m+n)}$ be a polynomial matrix of degree at most d, let S_k for $k = 1, 2, \ldots$ be its Sylvester matrices, and let*

$$k' := \left\lceil \frac{md}{n} \right\rceil \quad and \quad nk' = md + t, \quad where \ 0 \le t < n. \tag{4.1}$$

Then the following statements hold.

(a) *k' is the smallest index k for which the number of columns of S_k is larger than or equal to the number of rows of S_k.*

(b) *If $k' > 1$ and $t > 0$, then $M(\lambda)$ has full-Sylvester-rank if and only if $S_{k'-1}$ has full column rank and $S_{k'}$ has full row rank.*

(c) *If $k' = 1$ or $t = 0$, then $M(\lambda)$ has full-Sylvester-rank if and only if $S_{k'}$ has full row rank.*

Proof. Part (a) follows from the size of S_k displayed in (3.2), because $(k + d)m \le k(m + n)$ is equivalent to $dm/n \le k$, and this is equivalent to $k' \le k$. Part (b) follows from Lemmas 3.1 and 3.2, together with the fact that if $t > 0$ then $S_{k'}$ has strictly more columns than rows, while $S_{k'-1}$ has strictly less columns than rows. Note that $S_{k'-1}$ is defined since $k' > 1$. In part (c) we have two scenarios. If $t = 0$, then $S_{k'}$ is a square matrix and the result follows again from Lemmas 3.1 and 3.2 for any value of k'. If $k' = 1$ and $t > 0$, the same argument as in part (b) proves the result, with the only difference that in this case $S_{k'-1}$ is not defined. $\qquad\square$

Note that Corollary 3.9 immediately implies that polynomial matrices with full-Sylvester-rank are minimal bases whose leading coefficient has full rank. So, they have full row normal rank and, since the ranks r_k of their Sylvester matrices are given by $r_k = \min\{(k+d)m, k(m+n)\}$, their right minimal indices are fixed by the recurrence in Theorem 3.3. This leads to a characterization of full-Sylvester-rank matrices in terms of their *complete eigenstructure*, i.e., their finite and infinite elementary divisors[1] and their left and right minimal indices. This characterization together with other properties are presented in Theorem 4.3.

Theorem 4.3. *Let $M(\lambda) = M_0 + M_1\lambda + \cdots + M_d\lambda^d \in \mathbb{F}[\lambda]^{m \times (m+n)}$ be a polynomial matrix of degree at most d, let α_k be the number of right minimal indices of $M(\lambda)$ equal to k, and let k' and t be defined as in (4.1). Then the following statements hold.*

(a) *$M(\lambda)$ has full-Sylvester-rank if and only if the complete eigenstructure of $M(\lambda)$ consists only of the following right minimal indices*

$$\alpha_{k'-1} = t, \quad \alpha_{k'} = n - t, \quad and \quad \alpha_j = 0 \ for \ j \notin \{k'-1, k'\}. \tag{4.2}$$

(b) *If $M(\lambda)$ has full-Sylvester-rank, then $M(\lambda)$ is a minimal basis with $\operatorname{rank} M_d = m$, i.e., with all its row degrees equal to d.*

(c) *If $M(\lambda)$ has full-Sylvester-rank, then the degree of any minimal basis dual to $M(\lambda)$ is equal to k'. That is, with the notation of Corollaries 3.4 and 3.5, $d' = k'$ holds for full-Sylvester-rank matrices.*

Proof. Proof of (a). We assume first that $M(\lambda)$ has full-Sylvester-rank. Then for all $k \geq k'$ the Sylvester matrix S_k of $M(\lambda)$ has full row rank, since S_k has full rank and does not have more rows than columns. Therefore, Corollary 3.9 implies that $M(\lambda)$ is a minimal basis and that M_d has full row rank. These properties imply in turn that the complete eigenstructure of $M(\lambda)$ consists only of n right minimal indices, since $M(\lambda)$ has not finite elementary divisors as a consequence of Theorem 2.2, $M(\lambda)$ has not infinite elementary divisors as a consequence of rank $M_d = m$, $M(\lambda)$ has not left minimal indices because has full row normal rank, and the number of right minimal indices of $M(\lambda)$ is $\dim \mathcal{N}_r(M) = m + n - \operatorname{rank}(M) = n$. It remains to determine the n right minimal indices of $M(\lambda)$. To this purpose note that the full-Sylvester-rank property together with Lemma 4.2-(a) imply that S_k has full column rank for all $1 \leq k < k'$ and has full row rank for all $k \geq k'$, which in terms of the rank, r_k, and right nullity, n_k, of S_k is equivalent to

$$\begin{aligned} r_k &= k(m+n), & n_k &= 0, & \text{for } 1 \leq k < k', \quad \text{and} \\ r_k &= (k+d)m, & n_k &= nk - dm, & \text{for } k \geq k'. \end{aligned} \tag{4.3}$$

The relations (4.3) can be expressed also in terms of k' and t in (4.1) as follows

$$\begin{aligned} r_k &= k(m+n), & n_k &= 0, & \text{for } 1 \leq k < k', \quad \text{and} \\ r_k &= (k+d)m, & n_k &= n(k-k') + t, & \text{for } k \geq k'. \end{aligned} \tag{4.4}$$

Finally, from Theorem 3.3 we have that $\alpha_0 = n_1$ and $\alpha_k = n_{k+1} - 2n_k + n_{k-1}$ for $k \geq 1$, which combined with (4.4) yields

$$\begin{aligned} \alpha_k &= 0, & \text{for } 1 \leq k \leq k' - 2, \\ \alpha_{k'-1} &= n_{k'} = t, \\ \alpha_{k'} &= n_{k'+1} - 2n_{k'} = n - t, \\ \alpha_k &= 0, & \text{for } k > k', \end{aligned} \tag{4.5}$$

where the last line follows immediately from the fact that the n right minimal indices of $M(\lambda)$ have been already determined in the previous lines of (4.5). Observe that according to (4.1) $k' \geq 1$

---

[1]In this paper the infinite elementary divisors of a polynomial matrix $M(\lambda)$ with degree at most d are the elementary divisors associated to the eigenvalue 0 of the *reversal polynomial matrix* $\operatorname{rev}_d M(\lambda) := \lambda^d M(1/\lambda)$ [6].

and that in the limit case $k' = 1$, i.e., $md \le n$, we get from (4.5) $\alpha_{k'-1} = \alpha_0 = n_1 = t$, which is consistent with the initialization in Theorem 3.3. In this limit case the first lines in (4.3), (4.4), and (4.5) are not present. The following table illustrates the nullities, ranks, and α_k numbers for $k = k' - 2, k' - 1, k', k' + 1, k' + 2$:

k	$k' - 2$	$k' - 1$	k'	$k' + 1$	$k' + 2$
n_k	0	0	t	$t + n$	$t + 2n$
r_k	$(k'-2)(m+n)$	$(k'-1)(m+n)$	$(k'+d)m$	$(k'+d+1)m$	$(k'+d+2)m$
α_k	0	t	$n - t$	0	0

Next we prove the sufficiency in part (a). Assume that the complete eigenstructure of $M(\lambda)$ consists only of the n right minimal indices described in (4.2). Since the number of right minimal indices is precisely n, $M(\lambda)$ has full normal rank, i.e., $\mathrm{rank}(M) = m$. Then, we can apply Theorem 3.3 to $M(\lambda)$. Note that (3.4) implies (3.6) and, so, the right nullities n_k of the Sylvester matrices are determined from the α_k numbers as follows

$$n_1 = \alpha_0, \qquad n_{k+1} = n_k + \sum_{i=0}^{k} \alpha_i \quad \text{for } k \ge 1. \tag{4.6}$$

This recursion combined with (4.2) yields

$$
\begin{aligned}
n_k &= 0, & \text{for } k < k', \\
n_k &= n(k - k') + t, & \text{for } k \ge k',
\end{aligned}
$$

which according to (4.4) implies that every Sylvester matrix S_k of $M(\lambda)$ has full rank. This completes the proof of part (a).

Part (b) has been already proved at the beginning of the proof of part (a). Part (c) follows from part (a) and the fact that the row degrees of any minimal basis dual to $M(\lambda)$ are precisely the right minimal indices of $M(\lambda)$. $\qquad\square$

Theorem 4.3-(a) allows us to state in Theorem 4.4 another necessary and sufficient condition for a polynomial matrix to have full-Sylvester-rank in terms only of its right minimal indices or, equivalently, the degrees of their dual minimal bases.

Theorem 4.4. *Let $M(\lambda) = M_0 + M_1\lambda + \cdots + M_d\lambda^d \in \mathbb{F}[\lambda]^{m \times (m+n)}$ be a polynomial matrix of degree at most d, let α_k be the number of right minimal indices of $M(\lambda)$ equal to k, and let k' and t be defined as in (4.1). Then, $M(\lambda)$ has full-Sylvester-rank if and only if the right minimal indices of $M(\lambda)$ are*

$$\alpha_{k'-1} = t, \quad \alpha_{k'} = n - t, \quad \text{and} \quad \alpha_j = 0 \ \text{for } j \notin \{k'-1, k'\}. \tag{4.7}$$

Proof. Theorem 4.3-(a) implies that if $M(\lambda)$ has full-Sylvester-rank, then its right minimal indices are those in (4.7), which proves the necessity. The sufficiency is proved as follows. If the right minimal indices of $M(\lambda)$ are those in (4.7), then $\dim \mathcal{N}_r(M) = n$ and $\mathrm{rank}(M) = m$. Therefore, $M(\lambda)$ has not left minimal indices. In addition, the Index Sum Theorem [6, Theorem 6.5] applied to the polynomial matrix $M(\lambda)$ of grade d implies that $M(\lambda)$ has neither finite nor infinite elementary divisors, since the sum of the right minimal indices of $M(\lambda)$ is

$$(k' - 1)\,\alpha_{k'-1} + k'\,\alpha_{k'} = (k' - 1)\,t + k'\,(n - t) = -t + k'\,n = md = \mathrm{rank}(M)\,\mathrm{grade}(M)\,.$$

Therefore, the complete eigenstructure of $M(\lambda)$ consists only of the right minimal indices in (4.7) and Theorem 4.3-(a) implies that $M(\lambda)$ has full-Sylvester-rank. $\qquad\square$

Remark 4.5. Observe that Theorem 4.3-(b) states that full-Sylvester-rank matrices are minimal bases with full rank leading matrix coefficient M_d. We emphasize that the converse result is not true: a minimal basis $C(\lambda) \in \mathbb{F}[\lambda]^{m \times (m+n)}$ with degree at most d and leading coefficient C_d of full

rank has not necessarily full-Sylvester-rank. This follows immediately from Theorem 2.6 because there exist dual minimal bases $C(\lambda) \in \mathbb{F}[\lambda]^{m \times (m+n)}$ and $D(\lambda) \in \mathbb{F}[\lambda]^{n \times (m+n)}$ with the row degrees of $C(\lambda)$ all equal to d (equivalently C_d has full rank) and with the row degrees of $D(\lambda)$ having arbitrary values whose sum is md. Since the row degrees of $D(\lambda)$ are the right minimal indices of $C(\lambda)$, they can be different than those in (4.7) and so $C(\lambda)$ has not full-Sylvester-Rank.

5 Genericity of full-Sylvester-rank matrices and consequences

It is well-known that $p \times q$ constant matrices have *generically*, i.e., typically or "almost always" if the entries are considered as random variables, full rank equal to $\min\{p, q\}$. Therefore, it is natural to expect that *generically* all of the Sylvester matrices of a polynomial matrix $M(\lambda) \in \mathbb{F}[\lambda]^{m \times (m+n)}$ have full rank. In other words, it is natural to expect that *generically* a polynomial matrix $M(\lambda) \in \mathbb{F}[\lambda]^{m \times (m+n)}$ has full-Sylvester-rank. However, this has to be rigorously proved, since Sylvester matrices are highly structured matrices containing many zero entries and with block columns intimately related each other. The development of such rigorous proof and the analysis of some interesting consequences of this result are the goals of this section. To this purpose, we need to define the precise meaning of *genericity*, which in this work is essentially the standard notion in Algebraic Geometry.

We define genericity inside the vector space $\mathbb{F}[\lambda]_d^{m \times (m+n)}$ of polynomial matrices of size $m \times (m+n)$ and degree at most d, where in this section, and in the rest of the paper, $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. One motivation for considering $\mathbb{F}[\lambda]_d^{m \times (m+n)}$ as our "ambient" space comes from the applications to backward error analyses of polynomial eigenproblems solved via linearizations or ℓ-ifications (see [6, 9] for the definition of ℓ-ification) that we have in mind for the results in this paper. In practice, backward errors are considered arbitrary perturbations, since the only information available on them is their magnitude, which do not increase the degree of the polynomial matrix [10, 21].

The first step in the definition of genericity is to identify $\mathbb{F}[\lambda]_d^{m \times (m+n)}$ with $\mathbb{R}^{(d+1)m(m+n)}$ when $\mathbb{F} = \mathbb{R}$, or with $\mathbb{R}^{2(d+1)m(m+n)}$ when $\mathbb{F} = \mathbb{C}$. If $\mathbb{F} = \mathbb{R}$ such identification can be made, for instance, by mapping each polynomial matrix $M(\lambda) = M_0 + M_1\lambda + \cdots + M_d\lambda^d \in \mathbb{F}[\lambda]_d^{m \times (m+n)}$ into $\mathrm{vec}([M_0\ M_1\ \cdots\ M_d]) \in \mathbb{R}^{(d+1)m(m+n)}$, where $\mathrm{vec}(\cdot)$ is the standard vectorization operator defined for instance in [16, Chapter 4]. If $\mathbb{F} = \mathbb{C}$, one considers the entrywise real and imaginary parts of each matrix coefficient M_i, denoted by $\mathrm{Re}(M_i)$ and $\mathrm{Im}(M_i)$, respectively, and the identification is made by mapping $M(\lambda)$ into $\mathrm{vec}([\mathrm{Re}(M_0)\ \mathrm{Im}(M_0)\ \cdots\ \mathrm{Re}(M_d)\ \mathrm{Im}(M_d)]) \in \mathbb{R}^{2(d+1)m(m+n)}$. Next, we recall that an *algebraic set* in $\mathbb{R}^p$ is the set of common zeros of a finite number of multivariable polynomials with p variables and coefficients in $\mathbb{R}$, and that an algebraic set is *proper* if it is not the whole set $\mathbb{R}^p$. With these concepts at hand, the standard definition of genericity of Algebraic Geometry is as follows: a *generic set of $\mathbb{R}^p$ is a subset of $\mathbb{R}^p$ whose complement is contained in a proper algebraic set*. This definition extends obviously to the corresponding one of *generic set of $\mathbb{F}[\lambda]_d^{m \times (m+n)}$*, with $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, through the bijective "vec" mappings discussed above, since these mappings allow us to define algebraic sets of $\mathbb{F}[\lambda]_d^{m \times (m+n)}$. In the sequel, expressions as "generically the polynomial matrices in $\mathbb{F}[\lambda]_d^{m \times (m+n)}$ have the property $\mathcal{P}$" have the precise meaning of "the polynomial matrices of $\mathbb{F}[\lambda]_d^{m \times (m+n)}$ that satisfy property $\mathcal{P}$ are a generic set of $\mathbb{F}[\lambda]_d^{m \times (m+n)}$".

Now, we are in the position of stating and proving the main result of this section.

Theorem 5.1. *Let* $\mathrm{Syl}[\lambda]_d^{m \times (m+n)} \subset \mathbb{F}[\lambda]_d^{m \times (m+n)}$ *be the set of polynomial matrices of size $m \times (m+n)$, degree at most d, and with full-Sylvester-rank. Then the complement of* $\mathrm{Syl}[\lambda]_d^{m \times (m+n)}$ *is a proper algebraic set of* $\mathbb{F}[\lambda]_d^{m \times (m+n)}$ *and, so,* $\mathrm{Syl}[\lambda]_d^{m \times (m+n)}$ *is a generic set of* $\mathbb{F}[\lambda]_d^{m \times (m+n)}$.

Proof. Let k' and t be defined as in (4.1). We will prove the theorem in the case $k' > 1$ and $t > 0$. The proofs in the cases $k' = 1$ or $t = 0$ are similar and are omitted. Taking into account Lemma

4.2-(b), we have that

$$\mathrm{Syl}[\lambda]_d^{m\times(m+n)} = \{M(\lambda) \in \mathbb{F}[\lambda]_d^{m\times(m+n)} : \det(S_{k'-1}^* S_{k'-1}) \cdot \det(S_{k'} S_{k'}^*) \neq 0\},$$

where $S_{k'-1}$ and $S_{k'}$ are Sylvester matrices of $M(\lambda)$. The complement of $\mathrm{Syl}[\lambda]_d^{m\times(m+n)}$ relative to $\mathbb{F}[\lambda]_d^{m\times(m+n)}$ is

$$\left(\mathrm{Syl}[\lambda]_d^{m\times(m+n)}\right)^c = \{M(\lambda) \in \mathbb{F}[\lambda]_d^{m\times(m+n)} : \det(S_{k'-1}^* S_{k'-1}) \cdot \det(S_{k'} S_{k'}^*) = 0\},$$

which is obviously an algebraic set. More precisely: if $\mathbb{F} = \mathbb{R}$, $\left(\mathrm{Syl}[\lambda]_d^{m\times(m+n)}\right)^c$ is the set of zeros of one multivariable polynomial in the entries of the matrix coefficients M_i, $i = 0, 1, \ldots, d$, of $M(\lambda)$, and if $\mathbb{F} = \mathbb{C}$, the real and imaginary parts of $\det(S_{k'-1}^* S_{k'-1}) \cdot \det(S_{k'} S_{k'}^*) = 0$ are equivalent to two multivariable polynomial equations in the real and imaginary parts of the entries of the matrix coefficients M_i, $i = 0, 1, \ldots, d$, of $M(\lambda)$. It only remains to prove that $\left(\mathrm{Syl}[\lambda]_d^{m\times(m+n)}\right)^c$ is proper, i.e., that there is at least one polynomial matrix $M(\lambda) \in \mathbb{F}[\lambda]_d^{m\times(m+n)}$ such that $M(\lambda) \notin \left(\mathrm{Syl}[\lambda]_d^{m\times(m+n)}\right)^c$. This follows immediately from Theorem 2.6, which guarantees the existence of dual minimal bases $M(\lambda) \in \mathbb{F}[\lambda]^{m\times(m+n)}$ and $N(\lambda) \in \mathbb{F}[\lambda]^{n\times(m+n)}$ with all the row degrees of $M(\lambda)$ equal to d (so $M(\lambda) \in \mathbb{F}[\lambda]_d^{m\times(m+n)}$), and with t row degrees of $N(\lambda)$ equal to $k' - 1$ and the other $n - t$ equal to k', because the row degrees of each of these two matrices sum up md. Therefore, the right minimal indices of $M(\lambda)$, which are the row degrees on $N(\lambda)$, are the ones in (4.7) and, by Theorem 4.4, $M(\lambda) \in \mathrm{Syl}[\lambda]_d^{m\times(m+n)}$ and $M(\lambda) \notin \left(\mathrm{Syl}[\lambda]_d^{m\times(m+n)}\right)^c$. $\qquad\square$

We have just proved that generically the polynomial matrices in $\mathbb{F}[\lambda]_d^{m\times(m+n)}$ have full-Sylvester-rank. This can be combined with Theorem 4.3 to prove that generically the polynomial matrices in $\mathbb{F}[\lambda]_d^{m\times(m+n)}$ satisfy other interesting properties, in particular, the property of being minimal bases whose degree-d matrix coefficient has full row rank. These properties are stated in the following corollary, whose simple proof is omitted.

Corollary 5.2. *Let k' and t be defined as in (4.1). Then, the following subsets of $\mathbb{F}[\lambda]_d^{m\times(m+n)}$ are generic in $\mathbb{F}[\lambda]_d^{m\times(m+n)}$:*

(a) *The set of $m \times (m + n)$ polynomial matrices of degree at most d and whose complete eigenstructure consists only of the following right minimal indices*

$$\alpha_{k'-1} = t, \quad \alpha_{k'} = n - t, \quad and \quad \alpha_j = 0 \ for \ j \notin \{k' - 1, k'\},$$

where α_j denotes the number of right minimal indices equal to j.

(b) *The set of $m \times (m + n)$ polynomial matrices of degree at most d that are minimal bases with degree-d matrix coefficient of full row rank.*

(c) *The set of $m \times (m + n)$ polynomial matrices of degree at most d that are minimal bases and such that their dual minimal bases have degree equal to k'.*

Observe that the set defined in Corollary 5.2-(a) is precisely $\mathrm{Syl}[\lambda]_d^{m\times(m+n)}$, while $\mathrm{Syl}[\lambda]_d^{m\times(m+n)}$ is strictly included in the sets $\mathcal{A}$ and $\mathcal{B}$ defined in parts (b) and (c) of Corollary 5.2, respectively, as a consequence of the discussion in Remark 4.5. Therefore the complements of these sets satisfy $\mathcal{A}^c \subset \left(\mathrm{Syl}[\lambda]_d^{m\times(m+n)}\right)^c$ and $\mathcal{B}^c \subset \left(\mathrm{Syl}[\lambda]_d^{m\times(m+n)}\right)^c$, respectively.

6 Robustness of minimal bases and of full-Sylvester-rank matrices

This section has three goals: first, to characterize when a minimal basis $M(\lambda) \in \mathbb{F}[\lambda]_d^{m \times (m+n)}$ is robust under perturbations in the sense that all the polynomial matrices in a neighborhood of $M(\lambda)$ are also minimal bases; second, to estimate the size of such neighborhood; and, third, to prove that full-Sylvester-rank matrices are robust and to estimate the sizes of the corresponding neighborhoods of robustness.

Throughout the rest of the paper the singular values of a constant matrix $A \in \mathbb{F}^{p \times q}$ are denoted by $\sigma_1(A) \geq \sigma_2(A) \geq \cdots \geq \sigma_{\min\{p,q\}}(A)$ and the Sylvester matrices of any $P(\lambda) \in \mathbb{F}[\lambda]_d^{m \times (m+n)}$ are denoted by $S_k(P)$ for $k = 1, 2, \ldots$. In order to study the questions of interest in this section, we define a norm in $\mathbb{F}[\lambda]_d^{m \times (m+n)}$ as follows: the norm of any $P(\lambda) \in \mathbb{F}[\lambda]_d^{m \times (m+n)}$ is $\|S_1(P)\|_2$, where $\|A\|_2 = \sigma_1(A)$ is the standard spectral norm of the matrix A [20]. This norm induces the distance $\|S_1(P) - S_1(\widetilde{P})\|_2 = \|S_1(P - \widetilde{P})\|_2$ between any two polynomial matrices $P(\lambda), \widetilde{P}(\lambda) \in \mathbb{F}[\lambda]_d^{m \times (m+n)}$.

In this section, as well as in the rest of this manuscript, we will use very often Lemma 6.2, whose proof relies on Lemma 6.1.

Lemma 6.1. *Let* $A = \begin{bmatrix} A_1 & | & A_2 & | & \cdots & | & A_k \end{bmatrix}$ *where* $A_i \in \mathbb{F}^{m \times n_i}, i = 1, \ldots, k$. *Then*

$$\max_i \sigma_1(A_i) \leq \sigma_1(A) \leq \sqrt{\sum_{i=1}^k \sigma_1^2(A_i)} \leq \sqrt{k} \cdot \max_i \sigma_1(A_i).$$

Proof. This is a simple consequence of [16, Corollary 3.1.3] and [16, Problem 22 in p. 217]. $\square$

Lemma 6.2. *Let* $P(\lambda) \in \mathbb{F}[\lambda]_d^{m \times (m+n)}$. *Then the following inequalities hold for the Sylvester matrices of* $P(\lambda)$:
$$\|S_1(P)\|_2 \leq \|S_k(P)\|_2 \leq \sqrt{k} \cdot \|S_1(P)\|_2.$$

Proof. This is a direct consequence of Lemma 6.1 and the structure of $S_k(P)$. $\square$

Next theorem proves that a minimal basis of degree at most d is robust inside $\mathbb{F}[\lambda]_d^{m \times (m+n)}$ if and only if its row degrees are all equal to d or, equivalently, if and only if its degree d matrix coefficient has full rank.

Theorem 6.3. *Let* $M(\lambda) = M_0 + M_1 \lambda + \cdots + M_d \lambda^d \in \mathbb{F}[\lambda]_d^{m \times (m+n)}$ *be a minimal basis. Then the following statements hold:*

(a) *If* $\operatorname{rank} M_d < m$, *then for all* $\epsilon > 0$ *there exists a polynomial matrix* $\widetilde{M}(\lambda) \in \mathbb{F}[\lambda]_d^{m \times (m+n)}$ *that is not a minimal basis and satisfies* $\|S_1(M) - S_1(\widetilde{M})\|_2 < \epsilon$. *That is, as close as we want to* $M(\lambda)$ *there are polynomial matrices that are not minimal bases.*

(b) *If* $\operatorname{rank} M_d = m$, *then there exists an index* k *such that* $S_k(M)$ *has full row rank and every polynomial matrix* $\widetilde{M}(\lambda) = \widetilde{M}_0 + \widetilde{M}_1 \lambda + \cdots + \widetilde{M}_d \lambda^d \in \mathbb{F}[\lambda]_d^{m \times (m+n)}$ *that satisfies*

$$\|S_1(M) - S_1(\widetilde{M})\|_2 < \frac{\sigma_{(k+d)m}(S_k(M))}{\sqrt{k}} \tag{6.1}$$

is a minimal basis with $\operatorname{rank} \widetilde{M}_d = m$. *That is, all the polynomial matrices sufficiently close to* $M(\lambda)$ *are minimal bases with full rank leading coefficient.*

Proof. Proof of (a). Note that M_d has at least one zero row, because otherwise all the row degrees of $M(\lambda)$ would be equal to d, then M_d would be the highest-row-degree coefficient matrix of $M(\lambda)$ (recall Definition 2.1), and $\operatorname{rank} M_d = m$ by Theorem 2.2, which is a contradiction.

Assume first $m > 1$, then either $M_d \neq 0$ or $M_d = 0$ (in this latter case the degree of $M(\lambda)$ is strictly less than d). In the case $M_d \neq 0$, M_d has at least one zero row and at least one nonzero row

$w_k \neq 0$. A polynomial matrix $\widetilde{M}(\lambda)$ as in the statement can be constructed as follows: $\widetilde{M}(\lambda)$ is equal to $M(\lambda)$ except that one of the zero rows of M_d is replaced by the row vector $(0.5\epsilon/\|w_k\|_2)\, w_k$. This implies that the highest-row-degree coefficient matrix $\widetilde{M}_{hr}$ of $\widetilde{M}(\lambda)$ has two linearly dependent rows, so $\widetilde{M}(\lambda)$ is not a minimal basis by Theorem 2.2, and that $\|S_1(M) - S_1(\widetilde{M})\|_2 = 0.5\epsilon < \epsilon$. In the case $M_d = 0$, a polynomial $\widetilde{M}(\lambda)$ as in the statement can be constructed as follows: $\widetilde{M}(\lambda)$ is equal to $M(\lambda)$ except that two zero rows of M_d are both replaced by the same arbitrary vector v_ϵ with norm $\|v_\epsilon\|_2 < \epsilon/2$. Then, $\widetilde{M}_{hr}$ of $\widetilde{M}(\lambda)$ has two rows equal to v_ϵ, so $\widetilde{M}(\lambda)$ is not a minimal bases by Theorem 2.2, and $\|S_1(M) - S_1(\widetilde{M})\|_2 = \|[v_\epsilon \,;\, v_\epsilon]\|_2 < \epsilon$.

The limiting case $m = 1$, i.e., when $M(\lambda)$ has only one row and has degree smaller than d, requires a different proof. Note that in this case rank $M_d < m = 1$ is equivalent to $M_d = 0$. Let us express $M(\lambda) = [m_1(\lambda) \cdots m_{1+n}(\lambda)]$ in terms of its entries. From Theorem 2.2, we know that $M(\lambda_0) \neq 0$ for any $\lambda_0 \in \overline{\mathbb{F}}$. Given $\epsilon > 0$, define a number λ_ϵ whose modulus is large enough to satisfy $|m_i(\lambda_\epsilon)/\lambda_\epsilon^d| < \epsilon/\sqrt{1+n}$ for $i = 1, \ldots, 1+n$. Note that such number exists because all of the scalar polynomials $m_i(\lambda)$ have degree strictly smaller than d. In this case, we construct $\widetilde{M}(\lambda)$ as follows

$$\widetilde{M}(\lambda) = M(\lambda) - \lambda^d \left[\frac{m_1(\lambda_\epsilon)}{\lambda_\epsilon^d} \cdots \frac{m_{1+n}(\lambda_\epsilon)}{\lambda_\epsilon^d} \right],$$

which satisfies $\widetilde{M}(\lambda_\epsilon) = 0$, so $\widetilde{M}(\lambda)$ is not a minimal basis, and $\|S_1(M) - S_1(\widetilde{M})\|_2 < \epsilon$. This completes the proof of part (a).

Proof of (b). Corollary 3.9 guarantees the existence of a Sylvester matrix $S_k(M)$ with full row rank or, equivalently, with minimal singular value $\sigma_{(k+d)m}(S_k(M)) > 0$. Condition (6.1) and Lemma 6.2 imply

$$
\begin{aligned}
\|S_k(M) - S_k(\widetilde{M})\|_2 &= \|S_k(M - \widetilde{M})\|_2 \\
&\leq \sqrt{k}\, \|S_1(M - \widetilde{M})\|_2 = \sqrt{k}\, \|S_1(M) - S_1(\widetilde{M})\|_2 \\
&< \sigma_{(k+d)m}(S_k(M)),
\end{aligned}
$$

which in turns implies that $S_k(\widetilde{M})$ has full row rank by Weyl's perturbation theorem for singular values [20]. Then, part (b) follows from applying Corollary 3.9 to $\widetilde{M}(\lambda)$. $\qquad\square$

Remark 6.4. The natural choice of k in Theorem 6.3-(b) is the smallest index k_0 for which $S_k(M)$ has full row rank, since in this way the denominator of the right hand side of (6.1) is the smallest possible one, which favors a larger estimation of the neighbourhood of robustness. However, note that the numerator plays a nontrivial role and another larger k might be a better choice.

As a consequence of the proof of Theorem 6.3-(a) when $M_d = 0$, it is obvious that minimal bases are never robust under perturbations that increase their degrees. This simple fact is stated for completeness in the next corollary.

Corollary 6.5. *Let $M(\lambda) \in \mathbb{F}[\lambda]^{p \times q}$ be any minimal basis with $p < q$ and let d be any integer such that $d > \deg(M)$. Then, there exist polynomial matrices in $\mathbb{F}[\lambda]_d^{p \times q}$ which are not minimal bases and are as close as we want to $M(\lambda)$.*

We have proved in Section 5 that generically the polynomial matrices in $\mathbb{F}[\lambda]_d^{m \times (m+n)}$ have full-Sylvester-rank, which implies that generically the polynomials in $\mathbb{F}[\lambda]_d^{m \times (m+n)}$ are minimal bases with all their row degrees equal to d *and with the row degrees of the minimal bases dual to them fully determined by* (4.7). It is not surprising, due to their genericity, that full-Sylvester-rank polynomial matrices are robust. This is established in Theorem 6.6 together with estimates of the sizes of the robustness neighbourhoods. We advance that Theorem 6.6 plays a key role in the perturbation theory of dual minimal bases developed in Section 7, which is based on the mentioned idea that the row degrees of the minimal bases dual to full-Sylvester-rank matrices (or, equivalently, the right minimal indices of full-Sylvester-rank matrices) are completely fixed (recall Theorems 4.3-(a) and

16

4.4) and, thus, remain constant in a robustness neighborhood of any full-Sylvester-rank matrix. This invariance property is lost in any neighborhood of any minimal basis without full-Sylvester-rank, which implies that perturbation results for dual minimal bases are no longer possible in that scenario, unless perturbations with particular properties are considered.

Theorem 6.6. *Let $M(\lambda) \in \mathbb{F}[\lambda]_d^{m \times (m+n)}$ be a polynomial matrix with full-Sylvester-rank and let k' and t be defined as in (4.1). Then the following statements hold:*

(a) *If $k' > 1$ and $t > 0$, then every $\widetilde{M}(\lambda) \in \mathbb{F}[\lambda]_d^{m \times (m+n)}$ such that*

$$\|S_1(M) - S_1(\widetilde{M})\|_2 < \min \left\{ \frac{\sigma_{(k'-1)(m+n)}(S_{k'-1}(M))}{\sqrt{k'-1}}, \frac{\sigma_{(k'+d)m}(S_{k'}(M))}{\sqrt{k'}} \right\}$$

has full-Sylvester-rank.

(b) *If $k' = 1$ or $t = 0$, then every $\widetilde{M}(\lambda) \in \mathbb{F}[\lambda]_d^{m \times (m+n)}$ such that*

$$\|S_1(M) - S_1(\widetilde{M})\|_2 < \frac{\sigma_{(k'+d)m}(S_{k'}(M))}{\sqrt{k'}}$$

has full-Sylvester-rank.

Proof. Proof of (a). Recall that, from Lemma 4.2-(b), $S_{k'-1}(M)$ and $S_{k'}(M)$ have full column and full row rank, respectively, which implies that the smallest singular values of these matrices satisfy $\sigma_{(k'-1)(m+n)}(S_{k'-1}(M)) > 0$ and $\sigma_{(k'+d)m}(S_{k'}(M)) > 0$. The assumption in part (a) and Lemma 6.2 imply

$$\|S_{k'-1}(M) - S_{k'-1}(\widetilde{M})\|_2 = \|S_{k'-1}(M - \widetilde{M})\|_2 \leq \sqrt{k'-1}\,\|S_1(M - \widetilde{M})\|_2$$
$$= \sqrt{k'-1}\,\|S_1(M) - S_1(\widetilde{M})\|_2 < \sigma_{(k'-1)(m+n)}(S_{k'-1}(M))$$

and

$$\|S_{k'}(M) - S_{k'}(\widetilde{M})\|_2 \leq \sqrt{k'}\,\|S_1(M) - S_1(\widetilde{M})\|_2 < \sigma_{(k'+d)m}(S_{k'}(M)).$$

Then, Weyl's perturbation theorem for singular values [20] implies that $S_{k'-1}(\widetilde{M})$ and $S_{k'}(\widetilde{M})$ have full column and full row rank, respectively, and Lemma 4.2-(b) applied to $\widetilde{M}(\lambda)$ implies that $\widetilde{M}(\lambda)$ has full Sylvester rank.

The proof of part (b) is based on Lemma 4.2-(c) and is analogous to that of part (a). Thus, it is omitted. □

Example 6.7. Let us illustrate Theorem 6.6 for the pencil in Example 3.10. In this example, $m = 6$, $n = 2$, $d = 1$, $k' = 3$, and $t = 0$. Moreover, $S_3(M)$ has size 24×24 and we saw in Example 3.10 that $r_3 = \text{rank}\,S_3(M) = 24$. Thus, Lemma 4.2-(c) guarantees that $M(\lambda)$ has full-Sylvester-rank and we can apply Theorem 6.6-(b) to determine that the neighbourhood of full-Sylvester-rank-robustness is in this example

$$\|S_1(M) - S_1(\widetilde{M})\|_2 < \frac{\sigma_{(k'+d)m}(S_{k'}(M))}{\sqrt{k'}} = \frac{\sigma_{24}(S_3(M))}{\sqrt{3}} \approx 0.2569.$$

Since $\|S_1(M)\|_2 = \sqrt{2}$, this is a rather large robustness neighborhood.

We finish this section with an improvement of a particular case of Theorem 6.6, which is motivated by the following discussion. The estimations in Theorem 6.6 of the sizes of the neighbourhoods of full-Sylvester-rank-robustness are based on the smallest singular values of $S_{k'-1}(M)$ and $S_{k'}(M)$ (or only $S_{k'}(M)$ when $k' = 1$ or $t = 0$), which leads to easily computable estimations. However, what really determines the robustness of the full-Sylvester-rank of $M(\lambda)$ is the smallest of the norms of the perturbations $\Delta M(\lambda)$ that make $S_{k'-1}(M + \Delta M)$ or $S_{k'}(M + \Delta M)$ rank deficient,

as a consequence of Lemma 4.2. The point is that this is not measured by the smallest singular values of $S_{k'-1}(M)$ and $S_{k'}(M)$, because these singular values measure the distance of $S_{k'-1}(M)$ and $S_{k'}(M)$ to rank deficient matrices under unstructured perturbations that do not preserve the zero structure of the Sylvester matrices $S_{k'-1}(M + \Delta M)$ and $S_{k'}(M + \Delta M)$. This implies that the estimates of the sizes of the robustness neighborhoods obtained in Theorem 6.6 are, in general, strictly smaller than the actual sizes of those neighborhoods, except in the case $k' = 1$, i.e., $md \leq n$, corresponding to very "flat" polynomial matrices, where the size of the robustness neighborhood is sharp because $S_1(M)$ has no special zero structure. This is stated in the next corollary.

Corollary 6.8. *Let $M(\lambda) \in \mathbb{F}[\lambda]_d^{m \times (m+n)}$ be a polynomial matrix such that $md \leq n$. Then the following statements hold:*

(a) *$M(\lambda)$ has full-Sylvester-rank if and only if $S_1(M(\lambda))$ has full row rank.*

(b) *Every $\widetilde{M}(\lambda) \in \mathbb{F}[\lambda]_d^{m \times (m+n)}$ such that $\|S_1(M) - S_1(\widetilde{M})\|_2 < \sigma_{(d+1)m}(S_1(M))$ has full-Sylvester-rank.*

(c) *There exists a polynomial matrix $\widetilde{M}(\lambda) \in \mathbb{F}[\lambda]_d^{m \times (m+n)}$ that has not full-Sylvester-rank and satisfies $\|S_1(M) - S_1(\widetilde{M})\|_2 = \sigma_{(d+1)m}(S_1(M))$.*

Proof. Parts (a) and (b) follows immediately from the fact that $k' = 1$ in (4.1), from Lemma 4.2 and Theorem 6.6-(b). The matrix polynomial in part (c) is extracted from any rank deficient matrix $A \in \mathbb{F}^{(d+1)m \times (m+n)}$ such that $\|S_1(M) - A\|_2 = \sigma_{(d+1)m}(S_1(M))$ with the identification $A = S_1(\widetilde{M})$. One such A is obtained from the singular value decomposition of $S_1(M)$ by replacing the smallest singular value by zero. $\square$

7 Perturbations of minimal bases dual to full-Sylvester-rank matrices

Some applications require to know how the minimal bases dual to a given minimal basis $M(\lambda)$ change when $M(\lambda)$ is perturbed. See Section 9 for more information about such applications. In fact, this problem has been solved in [10, Section 6.2] for certain particular minimal bases that are the off-diagonal blocks of block Kronecker linearizations of polynomial matrices [10] and has been applied to the backward error analysis of polynomial eigenvalue problems solved via block Kronecker linearizations. The minimal bases considered in [10, Section 6.2] are very particular instances of full-Sylvester-rank polynomial matrices that have all their row degrees equal to 1 and all the row degrees of their dual minimal bases equal. The purpose of this section is to solve the corresponding perturbation problem for any polynomial matrix $M(\lambda)$ that has full-Sylvester-rank. Therefore, the degrees of its dual minimal bases are not always all equal, although they are completely fixed according to (4.7) by the degree and size of $M(\lambda)$, and so remain constant in a robustness neighborhood of the full-Sylvester-rank property. We emphasize that only for the case of full-Sylvester-rank polynomial matrices such perturbation theory is possible, since in other cases tiny perturbations of $M(\lambda)$ may change dramatically the row degrees of the minimal bases of its dual space, even in the case that the perturbation of $M(\lambda)$ is still a minimal basis.

The main result in this section is Theorem 7.1, which deserves some comments before being stated. A key observation on our problem is that there are infinitely many minimal bases $N(\lambda)$ dual to a given full-Sylvester-rank polynomial matrix $M(\lambda)$ and a key feature of the solution we propose is that Theorem 7.1 provides the same perturbation bound for a *relative* change of all of them. However, the allowed norms of the perturbations of $M(\lambda)$ do depend on the considered dual basis $N(\lambda)$ in order to guarantee that the perturbed $\widetilde{N}(\lambda)$ found in the proof of Theorem 7.1 is indeed a minimal basis. In Theorem 7.1, the Frobenius norm [20] $\|S_k(P)\|_F$ of the Sylvester matrices of a polynomial matrix $P(\lambda)$ is used, in addition to the spectral or 2-norm used in Section 6. The reason of using the Frobenius norm is that the equality $\|S_k(P)\|_F = \|S_k(P^T)\|_F$ holds, while it

is not valid in general for the spectral norm, and this equality makes the proof simpler and the bounds cleaner.

Theorem 7.1. *Let $M(\lambda) \in \mathbb{F}[\lambda]_d^{m\times(m+n)}$ be a polynomial matrix with full-Sylvester-rank, let k' and t be defined as in (4.1), and let $N(\lambda) \in \mathbb{F}[\lambda]_{k'}^{n\times(m+n)}$ be a minimal basis dual to $M(\lambda)$ with highest-row-degree coefficient matrix $N_{hr} \in \mathbb{F}^{n\times(m+n)}$. Moreover, let us define the quantities $\theta_1(M)$ and $\theta_2(M)$ as follows:*

(a) *If $k' > 1$ and $t > 0$*

$$\theta_1(M) := \min\left\{\frac{\sigma_{(k'-1)(m+n)}(S_{k'-1}(M))}{\sqrt{k'-1}}, \frac{\sigma_{(k'+d)m}(S_{k'}(M))}{\sqrt{k'}}, \frac{\sigma_{(k'+1+d)m}(S_{k'+1}(M))}{\sqrt{k'+1}}\right\},$$

$$\theta_2(M) := \min\left\{\frac{\sigma_{(k'+d)m}(S_{k'}(M))}{\sqrt{k'}}, \frac{\sigma_{(k'+1+d)m}(S_{k'+1}(M))}{\sqrt{k'+1}}\right\};$$

(b) *If $k' = 1$ and $t > 0$,*

$$\theta_1(M) = \theta_2(M) := \min\left\{\sigma_{(d+1)m}(S_1(M)), \frac{\sigma_{(d+2)m}(S_2(M))}{\sqrt{2}}\right\};$$

(c) *If $t = 0$*

$$\theta_1(M) := \min\left\{\frac{\sigma_{(k'+d)m}(S_{k'}(M))}{\sqrt{k'}}, \frac{\sigma_{(k'+1+d)m}(S_{k'+1}(M))}{\sqrt{k'+1}}\right\},$$

$$\theta_2(M) := \frac{\sigma_{(k'+1+d)m}(S_{k'+1}(M))}{\sqrt{k'+1}}.$$

Then, every $\widetilde{M}(\lambda) \in \mathbb{F}[\lambda]_d^{m\times(m+n)}$ such that

$$\|S_1(M) - S_1(\widetilde{M})\|_2 < \frac{1}{2} \cdot \theta_1(M) \cdot \frac{\sigma_n(N_{hr})}{\|S_1(N)\|_F} \tag{7.1}$$

has full-Sylvester-rank and has a dual minimal basis $\widetilde{N}(\lambda) \in \mathbb{F}[\lambda]_{k'}^{n\times(m+n)}$ that satisfies

$$\frac{\|S_1(N) - S_1(\widetilde{N})\|_F}{\|S_1(N)\|_F} \leq \frac{2}{\theta_2(M)} \cdot \|S_1(M) - S_1(\widetilde{M})\|_2. \tag{7.2}$$

In addition, if $t = 0$, then all the row degrees of $\widetilde{N}(\lambda)$ and $N(\lambda)$ are equal to k'.

Proof. We only prove the most difficult case, which is (a), i.e., when $k' > 1$ and $t > 0$. The proof of the case (b) is essentially the same as that of (a), with the differences coming from the differences between parts (a) and (b) of Theorem 6.6. The proof of (c) is considerably simpler, as it is briefly explained at the end of this proof. In order to perform the proof of case (a), note first that N_{hr} is a submatrix of $S_1(N)$. Thus, we get

$$\frac{\sigma_n(N_{hr})}{\|S_1(N)\|_F} \leq \frac{\sigma_n(N_{hr})}{\|S_1(N)\|_2} \leq \frac{\sigma_n(N_{hr})}{\|N_{hr}\|_2} \leq 1. \tag{7.3}$$

This inequality and (7.1) imply that $\|S_1(M) - S_1(\widetilde{M})\|_2$ is bounded as in Theorem 6.6-(a). Therefore, any $\widetilde{M}(\lambda)$ that satisfies (7.1) has full-Sylvester-rank and every of its dual minimal bases $\widetilde{N}(\lambda)$ has t row degrees equal to $k' - 1$, and $n - t$ row degrees equal to k', as it also happens for the minimal basis $N(\lambda)$ dual to $M(\lambda)$ (recall Theorem 4.4). As a consequence we can write, modulo row permutations,

$$N(\lambda) = \begin{bmatrix} X(\lambda) \\ Y(\lambda) \end{bmatrix} \quad \text{and} \quad \widetilde{N}(\lambda) = \begin{bmatrix} X(\lambda) \\ Y(\lambda) \end{bmatrix} + \begin{bmatrix} \Delta X(\lambda) \\ \Delta Y(\lambda) \end{bmatrix}, \tag{7.4}$$

19

where $X(\lambda), X(\lambda) + \Delta X(\lambda) \in \mathbb{F}[\lambda]_{k'-1}^{t \times (m+n)}$ have both all their row degrees equal to $k' - 1$, and $Y(\lambda), Y(\lambda) + \Delta Y(\lambda) \in \mathbb{F}[\lambda]_{k'}^{(n-t) \times (m+n)}$ have both all their row degrees equal to k'. In the rest of the proof, among the infinitely many minimal bases $\widetilde{N}(\lambda)$ dual to $\widetilde{M}(\lambda)$, we will determine one that satisfies (7.2) by determining the corresponding $\Delta X(\lambda)$ and $\Delta Y(\lambda)$. To this purpose, note that the equation $\widetilde{M}(\lambda)\, \widetilde{N}(\lambda)^T = 0$ defining dual minimal bases is equivalent with the notation in (7.4) to the pair of equations

$$\widetilde{M}(\lambda)\,(X(\lambda)^T + \Delta X(\lambda)^T) = 0 \quad \text{and} \quad \widetilde{M}(\lambda)\,(Y(\lambda)^T + \Delta Y(\lambda)^T) = 0\,. \tag{7.5}$$

By introducing the notation $\widetilde{M}(\lambda) = M(\lambda) + \Delta M(\lambda)$ and by using

$$M(\lambda)N(\lambda)^T = [M(\lambda)X(\lambda)^T \; M(\lambda)Y(\lambda)^T] = 0,$$

the equations in (7.5) can be written in the following equivalent form

$$\widetilde{M}(\lambda)\,\Delta X(\lambda)^T = -\Delta M(\lambda)\,X(\lambda)^T \quad \text{and} \quad \widetilde{M}(\lambda)\,\Delta Y(\lambda)^T = -\Delta M(\lambda)\,Y(\lambda)^T\,,$$

which in terms of the corresponding Sylvester matrices are equivalent to the equations

$$S_{k'}(\widetilde{M})\, S_1(\Delta X^T) = -S_{k'}(\Delta M)\, S_1(X^T) \quad \text{and} \tag{7.6}$$

$$S_{k'+1}(\widetilde{M})\, S_1(\Delta Y^T) = -S_{k'+1}(\Delta M)\, S_1(Y^T) \tag{7.7}$$

for the unknowns $S_1(\Delta X^T)$ and $S_1(\Delta Y^T)$. Since $\widetilde{M}(\lambda)$ has full-Sylvester-rank, the Sylvester matrices $S_{k'}(\widetilde{M})$ and $S_{k'+1}(\widetilde{M})$ have both full row rank and the equations (7.6) and (7.7) are consistent. In addition, for each of them, its minimum Frobenius norm solution is given by

$$S_1(\Delta X^T) = -S_{k'}(\widetilde{M})^\dagger\, S_{k'}(\Delta M)\, S_1(X^T) \quad \text{and} \tag{7.8}$$

$$S_1(\Delta Y^T) = -S_{k'+1}(\widetilde{M})^\dagger\, S_{k'+1}(\Delta M)\, S_1(Y^T)\,, \tag{7.9}$$

where $(\cdot)^\dagger$ stands for the Moore-Penrose pseudoinverse of a matrix. From (7.8), Lemma 6.2, Weyl's perturbation theorem for singular values, and (7.1), we get

$$\|S_1(\Delta X^T)\|_F \le \frac{\|S_{k'}(\Delta M)\|_2}{\sigma_{(k'+d)m}(S_{k'}(\widetilde{M}))} \|S_1(X^T)\|_F \le \frac{\sqrt{k'}\,\|S_1(\Delta M)\|_2}{\sigma_{(k'+d)m}(S_{k'}(M)) - \|S_{k'}(\Delta M)\|_2} \|S_1(X^T)\|_F$$

$$\le \frac{\sqrt{k'}\,\|S_1(\Delta M)\|_2}{\sigma_{(k'+d)m}(S_{k'}(M)) - \sqrt{k'}\,\|S_1(\Delta M)\|_2} \|S_1(X^T)\|_F \le \frac{\sqrt{k'}\,\|S_1(\Delta M)\|_2}{\frac{1}{2}\,\sigma_{(k'+d)m}(S_{k'}(M))} \|S_1(X^T)\|_F$$

$$= \frac{2\,\|S_1(\Delta M)\|_2}{\frac{1}{\sqrt{k'}}\,\sigma_{(k'+d)m}(S_{k'}(M))} \|S_1(X^T)\|_F\,. \tag{7.10}$$

Analogously, we get from (7.9)

$$\|S_1(\Delta Y^T)\|_F \le \frac{2\,\|S_1(\Delta M)\|_2}{\frac{1}{\sqrt{k'+1}}\,\sigma_{(k'+1+d)m}(S_{k'+1}(M))} \|S_1(Y^T)\|_F\,. \tag{7.11}$$

Observe that if $\Delta X(\lambda) = \sum_{i=0}^{k'-1} \Delta X_i\, \lambda^i$ and $\Delta Y(\lambda) = \sum_{i=0}^{k'} \Delta Y_i\, \lambda^i$, then

$$\|S_1(N^T) - S_1(\widetilde{N}^T)\|_F = \|S_1([\Delta X(\lambda)^T \; \Delta Y(\lambda)^T])\|_F = \left\| \begin{bmatrix} \Delta X_0^T & \Delta Y_0^T \\ \vdots & \vdots \\ \Delta X_{k'-1}^T & \Delta Y_{k'-1}^T \\ 0 & \Delta Y_{k'}^T \end{bmatrix} \right\|_F$$

$$= \sqrt{\|S_1(\Delta X^T)\|_F^2 + \|S_1(\Delta Y^T)\|_F^2}\,,$$

which combined with (7.10) and (7.11) yields

$$\|S_1(N^T) - S_1(\widetilde{N}^T)\|_F \leq \frac{2\,\|S_1(\Delta M)\|_2}{\theta_2(M)} \sqrt{\|S_1(X^T)\|_F^2 + \|S_1(Y^T)\|_F^2} = \frac{2\,\|S_1(\Delta M)\|_2}{\theta_2(M)}\,\|S_1(N^T)\|_F\,.$$

Since $\|S_k(P)\|_F = \|S_k(P^T)\|_F$ for any polynomial matrix, this inequality is (7.2). It remains to prove that $\widetilde{N}(\lambda)$ is a minimal basis, since we have only proved so far that $\widetilde{M}(\lambda)\,\widetilde{N}(\lambda)^T = 0$. In order to prove that $\widetilde{N}(\lambda)$ is a minimal basis, we will show that

$$Z := N_{hr} + \begin{bmatrix} \Delta X_{k'-1} \\ \Delta Y_{k'} \end{bmatrix} \tag{7.12}$$

has full row rank, because in this case (7.4) guarantees that Z is the highest-row-degree coefficient matrix of $\widetilde{N}(\lambda)$, corresponding to t row degrees equal to $k'-1$ and $n-t$ row degrees equal to k', and that $\widetilde{N}(\lambda)$ is row reduced. Thus $\widetilde{N}(\lambda)$ has full row normal rank (recall Remark 2.4) and since its row degrees are those of a minimal basis dual to $\widetilde{M}(\lambda)$ and it satisfies $\widetilde{M}(\lambda)\,\widetilde{N}(\lambda)^T = 0$, we get that $\widetilde{N}(\lambda)$ must indeed be a minimal basis dual to $\widetilde{M}(\lambda)$. The fact that Z in (7.12) has full row rank follows from (7.2) (which has been already proved) combined with the hypothesis (7.1) which imply

$$\|\begin{bmatrix} \Delta X_{k'-1} \\ \Delta Y_{k'} \end{bmatrix}\|_2 \leq \|\begin{bmatrix} \Delta X_{k'-1} \\ \Delta Y_{k'} \end{bmatrix}\|_F \leq \|S_1(\widetilde{N}) - S_1(N)\|_F < \frac{\theta_1(M)}{\theta_2(M)}\,\sigma_n(N_{hr}) \leq \sigma_n(N_{hr}).$$

Taking into account that N_{hr} has full row rank, this inequality implies that Z has also full row rank. The proof of the case (a) is completed.

Observe that in the case $t = 0$ considered in (c), the matrices $X(\lambda)$ and $\Delta X(\lambda)$ are not present in (7.4), i.e., the partition into two block rows corresponding to polynomial matrices of different degrees $k'-1$ and k' is no longer needed, since all the row degrees of $N(\lambda)$ and $\widetilde{N}(\lambda)$ are equal to k'. The proof is considerably simpler and follows from the one above for the case (a) by ignoring all equations involving $X(\lambda)$ and $\Delta X(\lambda)$. $\qquad\square$

It is interesting to emphasize with respect to Theorem 7.1 that if $t > 0$, then the minimal bases $N(\lambda)$ dual to full-Sylvester-rank matrices $M(\lambda) \in \mathbb{F}[\lambda]_d^{m\times(m+n)}$ are never robust under arbitrary perturbations in $\mathbb{F}[\lambda]_{k'}^{n\times(m+n)}$ since the row degrees of $N(\lambda)$ are not all equal (recall Theorem 6.3). However, Theorem 7.1 proves that such $N(\lambda)$ can behave smoothly if the perturbations are restricted to be in the set of minimal bases dual to the full-Sylvester-rank matrices satisfying (7.1). In addition, if $t = 0$ then we prove in Theorem 7.2 that any such $N(\lambda)$ has also full-Sylvester-rank and, therefore, $N(\lambda)$ admits a robustness result analogous to Theorem 6.6-(b) and a perturbation theorem similar to Theorem 7.1-(c) with respect to its dual minimal basis $M(\lambda)$.

Theorem 7.2. *Let $M(\lambda) \in \mathbb{F}[\lambda]_d^{m\times(m+n)}$ be a polynomial matrix with full-Sylvester-rank, let k' and t be defined as in (4.1), and let $N(\lambda) \in \mathbb{F}[\lambda]_{k'}^{n\times(m+n)}$ be a minimal basis dual to $M(\lambda)$. Then, $N(\lambda)$ has full-Sylvester-rank if and only if $t = 0$, i.e., if and only if all the row degrees on $N(\lambda)$ are equal.*

Proof. If $t > 0$, then the row degrees of $N(\lambda)$ take the two values $k'-1$ and k', as a consequence of Theorem 4.4, and $N(\lambda)$ has not full-Sylvester-rank according to Theorem 4.3-(b) (applied to $N(\lambda)$).

If $t = 0$, then the quantities in (4.1) corresponding to $N(\lambda)$ are $k'_N = \lceil nk'/m \rceil = d$ and $t'_N = 0$. In addition, $N(\lambda)$ has exactly m right minimal indices equal to d because these are the row degrees of $M(\lambda)$. Therefore, Theorem 4.4 implies that $N(\lambda)$ has full-Sylvester-rank. $\qquad\square$

We finish this section with another nice feature of the case $t = 0$. In this case, also the reversal polynomials of $M(\lambda)$ and $N(\lambda)$ have full-Sylvester-rank and, therefore, admit robustness and perturbation results as those in Theorem 6.6-(b) and Theorem 7.1-(c).

Theorem 7.3. *Let $M(\lambda) \in \mathbb{F}[\lambda]_d^{m \times (m+n)}$ be a polynomial matrix with full-Sylvester-rank, let k' and t be defined as in (4.1), let $N(\lambda) \in \mathbb{F}[\lambda]_{k'}^{n \times (m+n)}$ be a minimal basis dual to $M(\lambda)$, and let*

$$\mathrm{rev}_d M(\lambda) := \lambda^d M\left(\frac{1}{\lambda}\right) \quad and \quad \mathrm{rev}_{k'} N(\lambda) := \lambda^{k'} N\left(\frac{1}{\lambda}\right).$$

If $t = 0$, then $\mathrm{rev}_d M(\lambda)$ and $\mathrm{rev}_{k'} N(\lambda)$ are dual minimal bases and both have full-Sylvester-rank.

Proof. Since $t = 0$, all the row degrees of $N(\lambda)$ are equal to k' (recall Theorem 4.4). Then, [10, Theorem 2.7] implies that $\mathrm{rev}_d M(\lambda)$ and $\mathrm{rev}_{k'} N(\lambda)$ are dual minimal bases with all their row degrees equal, respectively, to d and k'. The fact that both have full-Sylvester-rank then follows from Theorem 4.4, since the right minimal indices of $\mathrm{rev}_d M(\lambda)$ (resp. $\mathrm{rev}_{k'} N(\lambda)$) are the row degrees of $\mathrm{rev}_{k'} N(\lambda)$ (resp. $\mathrm{rev}_d M(\lambda)$). $\qquad \square$

8 The classical rank conditions for a minimal basis revisited

In this section we return to the classical characterization of a minimal basis given in Theorem 2.2, originally presented in [13], in the case of minimal bases over $\mathbb{C}$ that are robust under arbitrary perturbations with the same or smaller degree, which according to Theorem 6.3 are those whose row degrees are all equal. These minimal bases include all polynomial matrices with full-Sylvester-rank and also other polynomial matrices, as it is explained in Remark 4.5. More precisely, our goal in this section is to prove that in the case of robust minimal bases all of the infinitely many constant matrices whose ranks are involved in Theorem 2.2 have minimum singular values bounded below by a common number determined by one of the Sylvester matrices of the considered minimal basis, and, so, none of such matrices is extremely closed to be rank deficient. This is proved in the following theorem.

Theorem 8.1. *Let $M(\lambda) = M_0 + M_1\lambda + \cdots + M_d\lambda^d \in \mathbb{C}[\lambda]^{m \times (m+n)}$ be a minimal basis with $\mathrm{rank}\, M_d = m$, let S_k be its Sylvester matrices for $k = 1, 2, \ldots$, and let d' be the largest right minimal index of $M(\lambda)$ or, equivalently, let d' be the degree of any minimal basis dual to $M(\lambda)$. Then*

$$\sigma_{(d+d')m}(S_{d'}) \leq \inf_{\lambda_0 \in \mathbb{C}} \sigma_m(M(\lambda_0)) \quad and \quad \sigma_{(d+d')m}(S_{d'}) \leq \sigma_m(M_d).$$

Proof. From Corollary 3.9, we obtain that $S_{d'}$ has full row rank. Therefore, its smallest singular value is larger than zero, i.e., $\sigma_{(d+d')m}(S_{d'}) > 0$. We use in this proof the well known fact [16, 20] that any matrix $A \in \mathbb{C}^{p \times q}$ with $p \leq q$ satisfies

$$\sigma_p(A) = \min_{0 \neq x \in \mathbb{C}^p} \frac{\|A^* x\|_2}{\|x\|_2} = \min_{0 \neq x \in \mathbb{C}^p} \frac{\|x^* A\|_2}{\|x\|_2}. \tag{8.1}$$

This result applied to $A = S_{d'}$ immediately implies that $\sigma_{(d+d')m}(S_{d'}) \leq \sigma_m(M_d)$, since the last block row of $S_{d'}$ is $[0 \ M_d]$ and one can choose in (8.1) vectors x with entries not corresponding to this last block row equal to zero. To prove the first inequality in Theorem 8.1 note that the polynomial matrix $M(\lambda)$ satisfies the following equality between polynomial matrices:

$$\begin{bmatrix} I & I\lambda & I\lambda^2 & \ldots & I\lambda^{d+d'-1} \end{bmatrix} \begin{bmatrix} M_0 & & & \\ M_1 & M_0 & & \\ \vdots & M_1 & \ddots & \\ M_d & & \ddots & M_0 \\ 0 & M_d & & M_1 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & M_d \end{bmatrix} = M(\lambda) \begin{bmatrix} I & I\lambda & I\lambda^2 & \ldots & I\lambda^{d'-1} \end{bmatrix},$$

$$\underbrace{}_{d' \text{ blocks}}$$

$$\tag{8.2}$$

where the identity matrices in the left hand side are equal to I_m and in the right hand side are equal to I_{m+n}. Theorem 2.2 implies that $\sigma_0 := \sigma_m(M(\lambda_0)) > 0$ for any $\lambda_0 \in \mathbb{C}$. Let $u_0 \in \mathbb{C}^m$ and $v_0 \in \mathbb{C}^{(m+n)}$ be left and right singular vectors of $M(\lambda_0)$ corresponding to σ_0, that is $\|u_0\|_2 = \|v_0\|_2 = 1$ and $u_0^* M(\lambda_0) = \sigma_0 v_0^*$. Then it follows from (8.2) that

$$
\left(\begin{bmatrix} 1 & \lambda_0 & \lambda_0^2 & \dots & \lambda_0^{d+d'-1} \end{bmatrix} \otimes u_0^* \right) \underbrace{\begin{bmatrix} M_0 & & & & \\ M_1 & M_0 & & & \\ \vdots & M_1 & \ddots & & \\ M_d & & \ddots & M_0 & \\ 0 & M_d & & M_1 & \\ \vdots & \ddots & \ddots & \vdots & \\ 0 & \dots & 0 & M_d \end{bmatrix}}_{d' \text{ blocks}} = \sigma_0 \left(\begin{bmatrix} 1 & \lambda_0 & \lambda_0^2 & \dots & \lambda_0^{d'-1} \end{bmatrix} \otimes v_0^* \right).
$$

From this last equality and (8.1) applied to $A = S_{d'}$, we get

$$
\sigma_{(d+d')m}(S_{d'}) \le \sigma_0 \sqrt{\frac{\sum_{i=1}^{d'} |\lambda_0|^{2(i-1)}}{\sum_{i=1}^{d+d'} |\lambda_0|^{2(i-1)}}} \le \sigma_0 = \sigma_m(M(\lambda_0)).
$$

Since this holds for all $\lambda_0 \in \mathbb{C}$, the result is proved. $\qquad\square$

Theorem 8.1 is consistent with the robustness result proved in Theorem 6.3-(b) that guarantees that all polynomial matrices in a neighborhood of a minimal basis $M(\lambda)$ with all its row degrees equal are also minimal bases, as long as we restrict to the space of polynomial matrices with degree less than or equal to $\deg(M)$. The point is that if $\inf_{\lambda_0 \in \mathbb{C}} \sigma_m(M(\lambda_0)) = 0$ (note that one can think this might happen although $\sigma_m(M(\mu)) > 0$ for each μ), then it would be possible to construct an $\widetilde{M}(\lambda)$ as close as we want to $M(\lambda)$ with $\sigma_m(\widetilde{M}(\mu)) = 0$ for some particular μ and such $\widetilde{M}(\lambda)$ would not be a minimal basis by Theorem 2.2.

9 Applications to backward error analyses

In this section, we discuss very briefly two concrete applications of some of the results in this paper and develop with detail a third application. The goals of this third application are to illustrate, first, how the results in this paper allow us to establish a framework for performing backward error analyses of polynomial eigenvalue problems solved by applying backward stable algorithms to strong linearizations or ℓ-ifications, and, second, that considerable nontrivial work is still needed to analyze the behaviour of particular linearizations or ℓ-ifications within this framework.

The first application is to extend the backward error analysis performed in [10, Section 6] for polynomial eigenvalue problems solved via the block Kronecker linearizations introduced in [10, Section 5] to the much more general class of strong block minimal bases (SBMB) linearizations of polynomial matrices introduced in [10, Section 3], whose construction is based on polynomial matrices with full-Sylvester-rank of degree 1 whose right minimal indices are all equal. This is a challenging problem with very interesting applications, since these SBMB linearizations include, among many others, strong linearizations of polynomial matrices expressed in non-monomial bases [18, 22], all Fiedler-like linearizations (modulo permutations) published so far in the literature [5], as well as some interesting vector spaces of potential strong linearizations of polynomial matrices [12].

A second application is related to the fact that the SBMB and block Kronecker linearizations of polynomial matrices recently introduced in [10] can be extended to SBMB and block Kronecker ℓ-ifications of polynomial matrices [11], which are constructed by using polynomial matrices with full-Sylvester-rank of degree ℓ and whose right minimal indices are all equal. The reader is referred

to [6, 9] for formal definitions of the concept of an ℓ-ification of a polynomial matrix. In this paper it is enough to know that an ℓ-ification of a polynomial matrix $P(\lambda)$ is another polynomial matrix of degree ℓ that has the same (finite and infinite) elementary divisors, the same number of left, and the same number of right minimal indices as $P(\lambda)$. The results in [11] pose naturally the question whether or not the backward error analysis for block Kronecker linearizations in [10, Section 6] can be extended to the classes of ℓ-ifications introduced in [11]. This seems to be a hard problem whose potential solution will be related for sure to the results in this paper.

The third application in this section is to perform a backward error analysis of polynomial eigenvalue problems solved by applying any backward stable algorithm to the strong ℓ-fications introduced in [9], which include also many interesting strong linearizations when $\ell = 1$. We develop this application with detail and for this purpose it is convenient to revise the most important results from [9]. Theorem 9.1 was proved in [9, Theorems 4.1 and 4.5]. The statement is essentially the one in [9] although the notation has been adapted to fit the one used in previous sections of this paper. Theorem 9.1 is valid in an arbitrary field, while the remaining results in this section are valid only when $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$.

Theorem 9.1. *Let*

$$L(\lambda) = \left[\begin{array}{c} K(\lambda) \\ M(\lambda) \end{array} \right] \in \mathbb{F}[\lambda]_{\ell}^{(p+m)\times(m+n)} \tag{9.1}$$

be a matrix polynomial of degree ℓ such that $M(\lambda) \in \mathbb{F}[\lambda]^{m\times(m+n)}$ is a minimal basis with all its row degrees equal to ℓ and such that the row degrees of any minimal basis dual to $M(\lambda)$ are all equal to $k' = m\ell/n$, i.e., all the right minimal indices of $M(\lambda)$ are equal. If $N(\lambda) \in \mathbb{F}[\lambda]_{k'}^{n\times(m+n)}$ is any minimal basis dual to $M(\lambda)$, then $L(\lambda)$ is a strong ℓ-ification of the polynomial matrix

$$P(\lambda) = K(\lambda)N(\lambda)^T \in \mathbb{F}[\lambda]^{p\times n} \, , \tag{9.2}$$

considered as a polynomial of degree less than or equal to $\ell + k'$. In addition, the right minimal indices of $L(\lambda)$ are obtained from the right minimal indices of $P(\lambda)$ by adding to each of them k', while the left minimal indices of $L(\lambda)$ are equal to those of $P(\lambda)$.

Observe that since $k' = m\ell/n$ must be an integer only some combinations of m, n, and ℓ are allowed in Theorem 9.1. Recall also that the value k' of the row degrees of $N(\lambda)$ is fixed to be precisely $k' = m\ell/n$ by Theorem 2.6 and the requirements that all the row degrees of $M(\lambda)$ are equal to ℓ and that all the row degrees of $N(\lambda)$ are equal. Theorem 9.1 establishes a wide framework for constructing strong ℓ-ifications of matrix polynomials. In particular, this framework includes the Frobenius-like strong ℓ-ifications introduced for first time in [6, Section 5.2]. Moreover, in [9, Section 4], it is explained how to determine $K(\lambda)$ from (9.2) when $P(\lambda)$ and $N(\lambda)$ are given and some interesting examples of strong ℓ-ifications are presented there.

It is interesting to emphasize that with the new results introduced in this paper *the polynomial matrix $M(\lambda)$ with the properties in Theorem 9.1 is just a matrix with full-Sylvester-rank with $t = 0$, i.e., with its right minimal indices all equal to k'*, which follows immediately from Theorem 4.4 and the fact that k' is an integer. This simple observation allows us to prove very easily the perturbation result for strong ℓ-ifications presented in Theorem 9.2 by using Theorem 7.1-(c). Recall in this context that $S_1(Q)$ denotes the first Sylvester matrix of the polynomial matrix $Q(\lambda)$ and that $\|S_1(Q)\|_F$ and $\|S_1(Q)\|_2$ denote the corresponding Frobenius and spectral norms, respectively.

Theorem 9.2. *Let*

$$L(\lambda) = \left[\begin{array}{c} K(\lambda) \\ M(\lambda) \end{array} \right] \in \mathbb{F}[\lambda]_{\ell}^{(p+m)\times(m+n)}$$

be a strong ℓ-ification of the polynomial matrix $P(\lambda) = K(\lambda)N(\lambda)^T \in \mathbb{F}[\lambda]^{p\times n}$, with the properties and notation stated in Theorem 9.1. If

$$\Delta L(\lambda) = \left[\begin{array}{c} \Delta K(\lambda) \\ \Delta M(\lambda) \end{array} \right] \in \mathbb{F}[\lambda]_{\ell}^{(p+m)\times(m+n)}$$

is another polynomial matrix of degree at most ℓ partitioned as $L(\lambda)$ and such that

$$\|S_1(\Delta M)\|_2 < \frac{1}{2} \cdot \theta_1(M) \cdot \frac{\sigma_n(N_{hr})}{\|S_1(N)\|_F}, \tag{9.3}$$

where $\theta_1(M)$ is defined as in Theorem 7.1-(c) with d replaced by ℓ and N_{hr} is the highest-row-degree coefficient matrix of $N(\lambda)$, then $L(\lambda) + \Delta L(\lambda)$ is a strong ℓ-ification of a polynomial matrix $P(\lambda) + \Delta P(\lambda)$ of degree less than or equal to $\ell + k'$ such that

$$\frac{\|S_1(\Delta P)\|_F}{\|S_1(P)\|_F} \leq \min\{\sqrt{k'+1}, \sqrt{\ell+1}\} \, C_{P,L} \, \frac{\|S_1(\Delta L)\|_F}{\|S_1(L)\|_F}, \tag{9.4}$$

with

$$C_{P,L} = \frac{\|S_1(L)\|_F}{\|S_1(P)\|_F} \|S_1(N)\|_F \left(1 + \frac{2\sqrt{k'+1}}{\sigma_{(k'+1+\ell)m}(S_{k'+1}(M))} \left(\|S_1(K)\|_F + \|S_1(\Delta K)\|_F\right)\right). \tag{9.5}$$

In addition, the right minimal indices of $L(\lambda) + \Delta L(\lambda)$ are obtained from the right minimal indices of $P(\lambda) + \Delta P(\lambda)$ by adding to each of them k', while the left minimal indices of $L(\lambda) + \Delta L(\lambda)$ are equal to those of $P(\lambda) + \Delta P(\lambda)$.

Proof. The condition (9.3) is just (7.1) in Theorem 7.1. Therefore, according to Theorem 7.1-(c,) the submatrix $M(\lambda) + \Delta M(\lambda) \in \mathbb{F}[\lambda]_\ell^{m \times (m+n)}$ of $L(\lambda) + \Delta L(\lambda)$ has full-Sylvester-rank, has parameter $t = 0$ (where t and k' are defined as in (4.1) with d replaced by ℓ), i.e., all the row degrees of any minimal basis dual to $M(\lambda) + \Delta M(\lambda)$ are equal to $k' = m\ell/n$, and there exists one of such dual minimal bases $N(\lambda) + \Delta N(\lambda)$ that satisfies (7.2). Thus, Theorem 9.1 applied to $L(\lambda) + \Delta L(\lambda)$ implies that $L(\lambda) + \Delta L(\lambda)$ is a strong ℓ-ification of the polynomial matrix

$$P(\lambda) + \Delta P(\lambda) = (K(\lambda) + \Delta K(\lambda)) (N(\lambda) + \Delta N(\lambda))^T \tag{9.6}$$

and that the relationship between the minimal indices of $L(\lambda) + \Delta L(\lambda)$ and $P(\lambda) + \Delta P(\lambda)$ is the one in the statement. From (9.6), we get

$$\Delta P(\lambda) = \Delta K(\lambda) \, N(\lambda)^T + K(\lambda) \, \Delta N(\lambda)^T + \Delta K(\lambda) \, \Delta N(\lambda)^T. \tag{9.7}$$

In the rest of the proof, we bound $\|S_1(\Delta P)\|_F$ for $\Delta P(\lambda)$ in (9.7) in order to get (9.4) and (9.5). To this purpose note that for any polynomial matrix $Q(\lambda)$ the norm $\|S_1(Q)\|_F$ coincides with the norm $\|Q(\lambda)\|_F$ introduced in [10, Definition 2.15]. Therefore, we can use the inequalities in [10, Lemma 2.16] which, together with (7.2) in the case (c) in Theorem 7.1 and the inequality $\|S_1(\Delta M)\|_2 \leq \|S_1(\Delta M)\|_F$, lead directly to (9.4) with $C_{P,L}$ replaced by

$$\widehat{C}_{P,L} = \frac{\|S_1(L)\|_F}{\|S_1(P)\|_F} \|S_1(N)\|_F \left(\frac{\|S_1(\Delta K)\|_F}{\|S_1(\Delta L)\|_F}\right.$$
$$\left. + \frac{2\sqrt{k'+1}}{\sigma_{(k'+1+\ell)m}(S_{k'+1}(M))} \frac{\|S_1(\Delta M)\|_F}{\|S_1(\Delta L)\|_F} \left(\|S_1(K)\|_F + \|S_1(\Delta K)\|_F\right)\right).$$

Finally, the trivial inequalities $\|S_1(\Delta K)\|_F / \|S_1(\Delta L)\|_F \leq 1$ and $\|S_1(\Delta M)\|_F / \|S_1(\Delta L)\|_F \leq 1$ prove the result. $\qquad \square$

Theorem 9.2 establishes a general framework to analyze when the solution of the complete polynomial eigenvalue problem for $P(\lambda)$ (including minimal indices) through the use of a backward stable algorithm applied to any of its strong ℓ-ifications described in Theorem 9.1 is backward stable from the polynomial point of view. Such backward stability will be guaranteed whenever $C_{P,L} \approx 1$, or is at least moderate. Clearly, a detailed analysis of $C_{P,L}$ is only possible for specific strong ℓ-ifications for which $K(\lambda)$, $M(\lambda)$, and $N(\lambda)$ are known. In particular, to determine lower bounds for $\sigma_{(k'+1+\ell)m}(S_{k'+1}(M))$ will be a necessary challenging task to study the backward stability of specific strong ℓ-ifications which is beyond the scope of this paper.

10   Conclusions

This paper introduces a new characterization of those polynomial matrices which are minimal bases in terms of the ranks of a finite number of the Sylvester matrices of the considered polynomial matrix. This characterization is applied to prove in a rigorous sense that polynomial matrices are generically minimal bases with very specific properties encoded in the concept of polynomial matrices with full-Sylvester-rank, which are carefully studied in this paper. In addition, the new characterization permits us to give a necessary and sufficient condition for a minimal basis to be robust under perturbations, which is just that the leading matrix coefficient has full rank, and to determine finite and simple estimations of the size of its robustness neighborhoods. Such results are particularly interesting in the case of full-Sylvester-rank polynomial matrices, since in this case they allow us to study in a very precise way how any minimal basis dual to a given minimal basis with full-Sylvester-rank changes when the given basis is perturbed. Finally some applications of the results of this paper are discussed with detail. We believe that they are just a very few among many other potential applications of this work.

References

[1] A. Amparan, F. M. Dopico, S. Marcaida, and I. Zaballa, Strong linearizations of rational matrices, submitted. Available as MIMS-eprint 2016.51, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, October 2016.

[2] B.D.O. Anderson and E.I. Jury, Generalized Bezoutian and Sylvester matrices in multivariable linear control, IEEE Trans. Aut. Contr., Vol.AC-21(4) (1976) 551-556.

[3] E. N. Antoniou, A. I. G. Vardulakis, and S. Vologiannidis, Numerical computation of minimal polynomial bases: A generalized resultant approach, Linear Algebra Appl. 405 (2005) 264–278.

[4] R.R. Bitmead, S.-Y. Kung, B.D.O. Anderson, and T. Kailath, Greatest common divisors via generalized Sylvester and Bezout matrices, IEEE Trans. Aut. Contr., Vol.AC-23(6) (1978) 1043-1047.

[5] M. I. Bueno, F. M. Dopico, J. Pérez, R. Saavedra, and B. Zykoski, A unified approach to Fiedler-like pencils via strong block minimal bases pencils, submitted. Available as arXiv preprint arXiv:1611.07170v1, 2016.

[6] F. De Terán, F. M. Dopico, and D. S. Mackey, Spectral equivalence of matrix polynomials and the index sum theorem, Linear Algebra Appl. 459 (2014) 264–333.

[7] F. De Terán, F. M. Dopico, D. S. Mackey, and P. Van Dooren, Polynomial zigzag matrices, dual minimal bases, and the realization of completely singular polynomials, Linear Algebra Appl. 488 (2016) 460–504.

[8] F. De Terán, F. M. Dopico, and P. Van Dooren, Matrix polynomials with completely prescribed eigenstructure, SIAM J. Matrix Anal. Appl. 36 (2015) 302–328.

[9] F. De Terán, F. M. Dopico, and P. Van Dooren, Constructing strong ℓ-ifications from dual minimal bases, Linear Algebra Appl. 495 (2016) 344–372.

[10] F. M. Dopico, P. W. Lawrence, J. Pérez, and P. Van Dooren, Block Kronecker linearizations of matrix polynomials and their backward errors, submitted. Available as MIMS EPrint 2016.34, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, May 2016.

[11] F. M. Dopico, J. Pérez, and P. Van Dooren, Block minimal bases ℓ-ifications of polynomial matrices, in preparation, 2016.

[12] H. Fassbender and P. Saltenberger, Block-Kronecker ansatz spaces for matrix polynomials, available as arXiv preprint arXiv:1610.05988.

[13] G. D. Forney, Minimal bases of rational vector spaces, with applications to multivariable linear systems, SIAM J. Control 13 (1975) 493–520.

[14] F. R. Gantmacher, The Theory of Matrices, Vol. I and II (transl.), Chelsea, New York, 1959.

[15] I. Gohberg, P. Lancaster, and L. Rodman, Matrix Polynomials, SIAM Publications, 2009. Originally published: Academic Press, New York, 1982.

[16] R.A. Horn and C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge 1991.

[17] T. Kailath, Linear Systems, Prentice Hall, Englewood Cliffs, NJ, 1980.

[18] P. W. Lawrence and J. Pérez, Constructing strong linearizations of matrix polynomials expressed in the Chebyshev bases, submitted. Available as MIMS EPrint 2016.12, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, 2016.

[19] P. W. Lawrence, M. Van Barel, and P. Van Dooren, Backward error analysis of polynomial eigenvalue problems solved by linearizations, SIAM J. Matrix Anal. Appl. 37 (2016) 123–144.

[20] G. W. Stewart and J. G. Sun, Matrix Perturbation Theory, Academic Press, Inc., Boston, MA, 1990.

[21] F. Tisseur, Backward error and condition of polynomial eigenvalue problems, Linear Algebra Appl. 309 (2000) 339–361.

[22] L. Robol, R. Vandebril, and P. Van Dooren, A framework for structured linearizations of matrix polynomials in various bases, to appear in SIAM J. Matrix Anal. Appl. Available as arXiv preprint arXiv:1603.05773, 2016.

[23] W. A. Wolovich, Linear Multivariable Systems, Applied Mathematical Sciences Series Vol.11, Springer-Verlag, New York, 1974.