LINEARIZATIONS AND DISTANCE PROBLEMS ASSOCIATED WITH MATRIX POLYNOMIALS

Ph.D. Thesis

Biswajit Das

DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
GUWAHATI-781039, ASSAM, INDIA

January 2020

LINEARIZATIONS AND DISTANCE PROBLEMS ASSOCIATED WITH MATRIX POLYNOMIALS

A Thesis Submitted

in Partial Fulfillment of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

by

Biswajit Das

(Roll Number: 136123016)

Under the Supervision of

Prof. Shreemayee Bora

to the

DEPARTMENT OF MATHEMATICS

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

GUWAHATI-781039, ASSAM, INDIA

January 2020

Contents

DECLARATION

I do hereby declare that the work contained in this thesis entitled **"Linearizations and Distance Problems associated with Matrix Polynomials"** has been done by me, a student in the Department of Mathematics, Indian Institute of Technology Guwahati under the guidance of **Prof. Shreemayee Bora** for the award of the degree of Doctor of Philosophy and that this work has not been submitted elsewhere for a degree.

January 2020

Guwahati

Biswajit Das
Roll No. 136123016
Department of Mathematics
Indian Institute of Technology Guwahati

CERTIFICATE

It is certified that the work contained in the thesis entitled **"Linearizations and Distance Problems associated with Matrix Polynomials"** by **Biswajit Das**, a student in the Department of Mathematics, Indian Institute of Technology Guwahati, for the award of the degree of Doctor of Philosophy has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

January 2020

Guwahati

Dr. Shreemayee Bora
Professor
Department of Mathematics
Indian Institute of Technology Guwahati

Dedicated to my Mother

Late Jyotsna Das

Acknowledgement

First and foremost, I would like to convey my sincere thanks and humble gratitude to my Supervisor Prof. Shreemayee Bora, Department of Mathematics, IIT Guwahati, who was kind enough to me to be my supervisor for this research study. I would like to thank her for the excellent guidance, patience, availability, and understanding. Her wholehearted cooperation and support inspired and enabled me to carry out and complete this research. Her erudition and rich experience shaped the course of the study and sharpened its outcome.

My sincere gratitude goes out to my collaborators Dr. Ashish Kothyari and Prof. Madhu N. Belur of IIT Bombay. I convey my thankfulness and gratitude to the doctoral committee members Prof. Rafikul Alam, Dr. Sriparna Bandopadhyay, and Prof. Rajen Kumar Sinha for their encouragement, insightful comments, and valuable suggestions. I greatly appreciate the help and cooperation of all the faculty members of the Department of Mathematics, IIT Guwahati. I also express my heartfelt gratitude to all the teachers who motivated and helped me in some way or another during my school and college days.

A very special gratitude goes to our former HODs Prof. Bhaba Kumar Sarma, Prof. Swaroop Nandan Bora, and Prof. N. Selvaraju and current HOD Prof. M. Guru Prem Prasad for their support and generous care. I am also thankful to the Department of Mathematics IIT Guwahati for providing me the necessary facilities for my research work.

I would also like to thank the staff members of the Department of Mathematics, IIT Guwahati, for their assistance in all technical and official matters.

I take this opportunity to sincerely acknowledge the IIT Guwahati and the MHRD for the financial assistance and facilities provided to me during my research work.

I am indebted to all my friends and juniors who made my stay joyful in IIT Guwahati and my seniors for a lot of help and encouragement. Among all my friends Devanand, Sonjoy, Ranjan, Madhu, Ashish and Debasish deserve a special mention for the friendship and warmth they extended to me during these years.

Finally, I express my very profound gratitude to my parents Mr. Jadu Das and Late Jyotsna Das, sister Sangeeta Das and especially my wife Aioshi Bhattacharjee

for their unfailing support and for always believing in me and encouraging me to follow my dreams. Without them, I would not have had the courage to embark on this journey.

My gratitude and appreciation goes out to each and everyone who contributed to this research endeavour in one way or the other and made this task possible in its present fruitful form.

Last but not the least, I would thank God for giving me the strength, ability, and opportunity to undertake this research study and to complete it.

January 2020

Guwahati

(Biswajit Das)

Abstract

Four different aspects of the study of matrix polynomials are considered in this thesis. Firstly, the vector space based study of linearizations of square regular matrix polynomials undertaken by Mackey et. al. (*Vector spaces of linearizations for matrix polynomials, SIAM J. Matrix Anal. Appl., 28(4):971–1004, 2006*) is extended to the case of the rectangular matrix polynomials. Secondly, it is established that there is a wide choice of linearizations of rectangular matrix polynomials arising from the proposed vector spaces that can be used to solve the associated complete eigenvalue problem for the matrix polynomial in a globally backward stable manner. Thirdly, the distance from a given square matrix polynomial to a nearest matrix polynomial with an elementary divisor of the form $(\lambda - \lambda_0)^j, j \geqslant r$, for a given complex number λ_0 and an integer $r > 1$, is considered. Fourthly, the distance from a given square matrix polynomial to a nearest normal rank deficient matrix polynomial with a given lower bound on the rank deficiency is considered. Special emphasis is on the distance to a nearest singular (or non-regular) matrix polynomial. Both the distance problems are considered in their most general forms and are long standing open questions. In each case, the matrix polynomials of interest are characterized in terms of the suitable rank deficiency of certain block Toeplitz matrices. Based on these characterizations, optimizations for computing the said distances are formulated and upper and lower bounds are derived in various norm settings. In particular, the distance to a nearest singular matrix polynomial is established as the reciprocal of certain μ-value problem and a numerical strategy to compute it based on the information about the minimal indices of a nearest singular matrix polynomial is proposed. The distances are computed via numerical software and compared with several upper and lower bounds in numerical experiments. Results are shown to compare favourably with those in the literature.

List of Figures

List of Tables

1 Introduction

Eigenvalue problems associated with matrix polynomials $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$, where $A_i, i = 0, \ldots, k$, are $m \times n$ matrices occur in a wide range of applications like vibration analysis of machines, buildings and vehicles, in control theory and linear systems theory and as approximate solutions of other nonlinear eigenvalue problems [32, 44, 69, 72, 78]. This thesis considers four problems associated with matrix polynomials.The first problem is concerned with the study of linearizations of rectangular matrix polynomials in a vector space framework. The second problem is concerned with the global backward stability of the solution of the eigenvalue problem associated with rectangular matrix polynomials via linearizations arising from a vector space setting. The third problem is the distance from a given square matrix polynomial to a nearest regular matrix polynomial with a specified elementary divisor. The fourth problem is about the distance from an $n \times n$ regular matrix polynomial to a nearest matrix polynomial of normal rank at most $r(\leqslant n-1)$ with special emphasis on the distance to a nearest singular matrix polynomial.

When the polynomial is square and regular, i.e., $\det(P(\lambda)) \not\equiv 0$, the associated polynomial eigenvalue problem consists of finding the finite and infinite eigenvalues and corresponding eigenvectors. However when the polynomial is singular, i.e., when it is either non-square or $\det(P(\lambda)) \equiv 0$, then the eigenvalue problem is said to be a complete eigenvalue problem as in addition to finite and infinite eigenvalues and corresponding elementary divisors, the minimal bases and indices corresponding to the left and right null spaces of the matrix polynomial also have to be computed. The most common approach for solving the eigenvalue problem associated with a matrix polynomial $P(\lambda)$ is to linearize it by converting the problem into an equivalent problem associated with a larger matrix pencil of the form $L(\lambda) = \lambda X + Y$ called a linearization of $P(\lambda)$, and solving the eigenvalue problem for $L(\lambda)$ by using standard algorithms like the QZ algorithm [33] when $L(\lambda)$ is regular, or the staircase algorithm [74] when $L(\lambda)$ is singular. The solution for $P(\lambda)$ is then recovered from that of its linearization. The linearization is called a strong linearization if reversing the order of its coefficient matrices results in a linearization of the polynomial obtained by reversing the order of the coefficient matrices of $P(\lambda)$. Note that it is possible to compute eigenvalues of $P(\lambda)$ at ∞ if any and associated elementary divisors only from strong linearizations of $P(\lambda)$.

The most commonly used forms of linearizations for solving polynomial eigenvalue problems associated with $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ are the first and second Frobenius companion forms. One of the first systematic studies of linearizations to be undertaken is [61] which introduced vector spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ of matrix pencils for a given $n \times n$ regular matrix polynomial $P(\lambda)$ as sources of linearizations of $P(\lambda)$. This work gave a whole new direction to research in the theory of linearizations due to the special properties of these vector spaces. For instance, it was shown that constructing pencils in these spaces is very simple and almost all the pencils are strong linearizations of $P(\lambda)$ from which the eigenvalues and corresponding eigenvectors can be easily recovered. It was shown in [14] that even when $P(\lambda)$ is square but singular, almost every pencil in $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ is a linearization of $P(\lambda)$ from which the solution of the complete eigenvalue problem for $P(\lambda)$ can be easily recovered. The problem of extending $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ to study linearizations of rectangular matrix polynomials in a vector space setting by defining vector spaces that coincide with $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ when the polynomial is square was left open in [14]. This problem is solved in Chapter 2 of this thesis, by introducing the concept of a generalized linearization or g-linearization of a rectangular matrix polynomial and its strong version. It is shown that with minor modifications, the spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ may be associated with a rectangular matrix polynomial and they have the same properties with respect to g-linearizations that the original spaces have with respect to linearizations of $P(\lambda)$. In particular, the complete eigenvalue problem associated with the matrix polynomial can be solved by using almost every matrix pencil from these spaces. Further, almost every pencil in these spaces can be 'trimmed' to form many smaller pencils that are strong linearizations of the matrix polynomial which readily solve the complete eigenvalue problem for the matrix polynomial. These linearizations are easier to construct and are often smaller than the Fiedler linearizations introduced in [17].

The literature on linearizations of matrix polynomials has expanded very rapidly in recent times (see for example [9, 11, 15, 16, 22, 60] and references therein). Choices of linearizations that guarantee the computation of an eigenvalue and a corresponding eigenvector of the polynomial in a backward stable manner have been identified in [40]. This may not be useful in practice as different linearizations may have to be used to compute different eigenvalue-eigenvector pairs for the same matrix polynomial. From the point of view of computation, a desirable property of linearizations is that the algorithm for solving the complete eigenvalue problem via such a linearization is globally backward stable. Global backward stability implies that the solution of the complete eigenvalue problem associated with $P(\lambda)$ via a linearization, say $L(\lambda)$, is the exact solution of the problem for $P(\lambda) + \Delta P(\lambda)$ where $|||\Delta P|||/|||P|||$ is of the order of unit roundoff $\mathbf{u}$ with respect to a norm $|||\cdot|||$ on the matrix polynomials. Moreover, the rules for extracting the left and right minimal indices of $P(\lambda)$ from $L(\lambda)$ remain the same for $P(\lambda) + \Delta P(\lambda)$ with respect to the class of linearization to which $L(\lambda)$ belongs. Global backward stability analysis of algorithms that use the Frobenius companion linearizations has been undertaken in [75]. More recently this has been extended to the block Kronecker linearizations in [22] which identifies optimal choices of block Kronecker linearizations that ensure global backward stability. In Chapter 3 the analysis in [22] is extended to linearizations of $P(\lambda)$ extracted from the g-linearizations in Chapter 2. Our analysis shows that there is a wider choice of linearizations beyond the ones identified in [22] that can be used to solve the complete eigenvalue problem for $P(\lambda)$ in a globally backward stable manner.

The two distance problems considered in this thesis are motivated by the following example. Consider a linear time invariant dynamical system

$$E\dot{x}(t) = Ax(t) + Bu(t), \quad y(t) = Cx(t) + Du(t),$$

where $x(t)$ is the state, $u(t)$ is the control, $y(t)$ is the output, A and E are $n \times n$ matrices and B, C and D are $n \times m$ and $p \times n$ and $p \times m$ matrices respectively. If $\lambda A - E$ is a regular matrix pencil with an elementary divisor $\lambda^j, j \geqslant 2$, then it is well known [13, 55] that the system may not be solvable for all initial states unless the controller $u(t)$ is sufficiently smooth. Again if the matrix pencil $\lambda A - E$ is singular (or non-regular), then the transfer function $G(s) = C(A - sE)^{-1}B$ does not exist and the system may not be (uniquely) solvable for all initial states even if the controller is smooth enough. If the pencil $\lambda A - E$ is regular and does not have $\lambda^j, j \geqslant 2$, as an elementary divisor then the above system is associated with the regular index at most one pencil $\lambda E - A$. Even if $\lambda E - A$ is a regular index at most one pencil but is close to a pencil without these properties, then the perturbations induced by modelling and other approximation errors may induce the system to exhibit unstable behaviour. Therefore, given a square matrix polynomial $P(\lambda)$, the distance to a nearest regular matrix polynomial with an elementary divisor of the form $(\lambda - \lambda_0)^j, j \geqslant r$, for a given scalar λ_0 and integer $r \geqslant 2$ is considered in Chapter 4 of this thesis. In Chapter 5 the distance from an $n \times n$ regular matrix polynomial $P(\lambda)$ to a nearest matrix polynomial of normal rank at most r, where $0 < r \leqslant n-1$, is considered with special emphasis on the case $r = n-1$ which is the distance to a nearest singular matrix polynomial. Both problems pose theoretically intriguing questions of very significant practical importance [12, 13, 79]. In fact finding a practical numerical method for computing the second distance is a long standing open question [12]. It is observed that certain block Toeplitz matrices play an important role in the computation of both the distances resulting in a strong connection between them.

To solve the first distance problem we show that it is enough to find the distance to a nearest matrix polynomial having λ_0 as an eigenvalue of algebraic multiplicity at least r. Unlike most of the work in the literature [50, 51, 53, 66], it is considered for a $P(\lambda)$ that is either regular or singular with perturbations affecting all the coefficient matrices. In fact by using elementary perturbation theory it is shown that if the matrix polynomial $P(\lambda)$ is singular, then it is arbitrarily close to a regular matrix polynomial with the desired property. This solves the distance problem for singular matrix polynomials. For regular matrix polynomials, the distance problem is shown to be equivalent to finding a smallest structure preserving perturbation such that certain block Toeplitz matrices formed by $P(\lambda_0)$ and its derivatives becomes suitably rank deficient. From this it follows that if λ_0 is not an eigenvalue of $P(\lambda)$, then solving the distance problem is equivalent to computing a generalized version of a structured singular value or μ-value [47]. It is well known that the μ-value computation is an NP-hard problem [7]. Therefore this problem is also likely to be NP-hard. The distance is formulated as an optimization and bounds are derived from the characterizations. Some special cases for which the solutions has a closed form expression is presented. Computed values of the distance obtained via Broyden Fletcher Goldfarb Shanno (BFGS) [58] and MATLAB's `globalsearch` [73] algorithms are compared with upper and lower bounds which appear to be tight in many instances.

The second distance problem is shown to be equivalent to computing a smallest structure preserving perturbation such that certain convolution matrices of the polynomial (that are block Toeplitz) become suitably rank deficient, thus leading to a characterization of the distance. Additionally by considering the relationship of the Jordan chain structure of a matrix polynomial at zero with its rank, the distance to a nearest singular matrix polynomial is shown to be also equivalent to the rank deficiency of another type of block Toeplitz matrix that appears in the first distance problem when $\lambda_0 = 0$. This leads to a new characterization that links the two distance problems under consideration. The characterizations imply that finding the distance to singularity may be an NP-hard problem as it is the reciprocal of a certain μ-value. Upper and lower bounds as well as information about the minimal indices of nearest singular matrix polynomials are derived from these characterizations. Based on this information, a numerical strategy to solve the optimizations that computes the distance to singularity is devised and implemented via BFGS and MATLAB's `globalsearch` algorithms. Numerical experiments show that the bounds are tight in many cases and the computed distances compare favourably with values obtained in the literature, thus demonstrating the efficacy of the strategy.

This thesis is organized as follows. Chapter 1 contains notations and basic concepts that are used in the subsequent chapters. These include fundamental notions associated with matrix polynomials and basic results associated with certain block Toeplitz matrices that play an important role throughout the thesis. Chapter 2 is about an effective vector space framework that generalizes the study of linearizations via the spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ in [61] to rectangular matrix polynomials. Chapter 3 undertakes a global backward stability analysis of the solution of the eigenvalue problem associated with rectangular matrix polynomials via the linearizations that arise in Chapter 2 and identifies a wide class of linearizations for which global backward stability is assured. Given a square matrix polynomial $P(\lambda)$, $\lambda_0 \in \mathbb{C}$ and $r > 0$, Chapter 4 studies the distance from $P(\lambda)$ to a nearest regular matrix polynomial having an elementary divisor of the form $(\lambda - \lambda_0)^j, j \geqslant r$. Chapter 5 studies the distance from a given $n \times n$ matrix polynomial $P(\lambda)$ to a nearest matrix polynomial of normal rank at most r where $0 < r \leqslant n - 1$ is given. Special emphasis is on the case $r = n - 1$ which corresponds to the distance to a nearest singular matrix polynomial.

1.1 Notations and terminology

The following notations have been used throughout the thesis.

- $\mathbb{R}$ and $\mathbb{C}$ are the sets of real and complex numbers, respectively.
- $\mathbb{R}^n$ and $\mathbb{C}^n$ are the sets of real and complex vectors of length n, respectively.

- $\mathbb{C}^{m\times n}$ is the set of complex matrices of size $m \times n$.
- $\mathbb{C}[\lambda]$ is the ring of polynomial with variable λ and coefficients in $\mathbb{C}$.
- $\mathbb{C}(\lambda)$ is the field of rational functions with coefficients in $\mathbb{C}$.
- $\mathbb{C}(\lambda)^n$ is the vector space of n-tuples of rational functions over the field $\mathbb{C}(\lambda)$.
- $\mathbb{C}[\lambda]^{m\times n}$ is the set of all $m \times n$ matrices with entries in $\mathbb{C}[\lambda]$.
- $P(\lambda)$ denotes the matrix polynomial $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ of size $m \times n$ of grade k, where A_i for $i = 0, 1, \ldots, k$ are real or complex matrices. The degree of $P(\lambda)$ denoted by deg P is the largest integer d such that $A_d \neq 0$. Degree one matrix polynomial is called matrix pencil.
- $\Delta P(\lambda)$ denotes the matrix polynomial $\Delta P(\lambda) = \sum_{i=0}^{k} \lambda^i \Delta A_i$ of size $m \times n$, where ΔA_i for $i = 0, 1, \ldots, k$ are real or complex matrices.
- $\det(A)$ denotes the determinant of the matrix A.
- A regular matrix polynomial $P(\lambda)$ is a square matrix polynomial for which $\det(P(\lambda_0)) \neq 0$ for some $\lambda_0 \in \mathbb{C}$.
- Normal rank of a matrix polynomial $P(\lambda)$ is the rank of $P(\lambda)$ considered as a matrix with entries in $\mathbb{C}(\lambda)$ and is denoted by $\operatorname{nrank} P$. It is equivalent to saying $\operatorname{nrank} P = \max\limits_{\lambda_0\in\mathbb{C}} \operatorname{rank}(P(\lambda_0))$.
- $\operatorname{rev}_k P(\lambda)$ denotes the k-reversal of $P(\lambda)$ defined by $\operatorname{rev}_k P(\lambda) := \sum_{i=0}^{k} \lambda^i A_{k-i}$.
- $P^*(\lambda)$ denotes the adjoint of $P(\lambda)$ defined by $P^*(\lambda) = \sum_{i=0}^{k} \lambda^i A_i^*$.
- $(P(\lambda))^T$ denotes the transpose of $P(\lambda)$ defined by $(P(\lambda))^T = \sum_{i=0}^{k} \lambda^i A_i^T$.
- For two matrix polynomials $P(\lambda)$ and $Q(\lambda)$ for which the product $P(\lambda)Q(\lambda)$ and the sum $P(\lambda) + Q(\lambda)$ are defined, for notational convenience we set $PQ(\lambda) := P(\lambda)Q(\lambda)$ and $(P + Q)(\lambda) := P(\lambda) + Q(\lambda)$.
- $e_j \in \mathbb{C}^n$ denotes the j-th standard unit vector of $\mathbb{C}^n$.
- A^T, $\bar{A}$ and A^* denote the transpose, complex conjugate and complex conjugate transpose of the matrix A, respectively.
- A^{-1} denotes the inverse of A.

- $A^\dagger$ denotes the Moore-Penrose pseudoinverse of matrix A.

- I denotes the identity matrix and I_n denotes the identity matrix of size n, whenever the dimension needs to be emphasized.

- When 0 denotes the zero matrix, then 0^T denotes its transpose. Also 0_n denotes a column of n zeros and $0_{m\times n}$ denotes the $m\times n$ zero matrix, whenever dimensions need to be emphasized.

- $\|\cdot\|_2$ and $\|\cdot\|_F$ denote the spectral norm and Frobenius norm respectively, of a vector or a matrix.

- A matrix A is called Hermitian if A satisfies $A^* = A$.

- A matrix polynomial $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ is called Hermitian if all of its coefficient matrices A_i are Hermitian.

- $\lambda_{\max}(H)$ denotes the largest eigenvalue of the Hermitian matrix H.

- $\sigma_{\min}(A)$ and $\sigma_{\max}(A)$ denotes the smallest and largest singular value of the matrix A, respectively.

- $\sigma_k(A)$ denotes the k^{th} largest singular value of the matrix A.

- Given two matrices $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{p\times q}$, $A \otimes B$ denotes the Kronecker product of A and B defined by

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \in \mathbb{C}^{mp\times nq}, \quad \text{where } A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}.$$

1.2 Preliminaries

1.2.1 Terminology and definitions related to matrix polynomials

In this section we discuss about the terms related to matrix polynomials.

Smith form of a matrix polynomial. For a matrix polynomial of grade k of the form $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$, $A_i \in \mathbb{C}^{m\times n}$, there exist two matrix polynomials $E(\lambda)$ and

$F(\lambda)$ of size $m \times m$ and $n \times n$ respectively with nonzero determinants independent of λ such that $P(\lambda) = E(\lambda)D(\lambda)F(\lambda)$, where

$$D(\lambda) = \begin{bmatrix} d_1(\lambda) & & & & & \\ & \ddots & & & & \\ & & d_t(\lambda) & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} \in \mathbb{C}[\lambda]^{m\times n}$$

is a diagonal matrix with monic scalar polynomials $d_i(\lambda)$ such that $d_{i-1}(\lambda)$ divides $d_i(\lambda)$. This is called the Smith form of $P(\lambda)$.

Important concepts associated with $P(\lambda)$ may be defined via its Smith form. The nonzero diagonal elements $d_1(\lambda), \ldots, d_t(\lambda)$ are called the invariant polynomials of $P(\lambda)$. The number of such invariant polynomials is the normal rank of $P(\lambda)$. The polynomial $P(\lambda)$ is said to be regular if $m = n$ and its normal rank is equal to its size n. Else it is said to be a non-regular or singular matrix polynomial. Any $\lambda_0 \in \mathbb{C}$ is a finite eigenvalue of $P(\lambda)$ if $\operatorname{rank} P(\lambda_0) < \operatorname{nrank} P$. We say that $P(\lambda)$ as a grade k matrix polynomial has an infinite eigenvalue if $\operatorname{rev}_k P(\lambda)$ has zero as an eigenvalue. If $\lambda_0 \in \mathbb{C}$ is an eigenvalue of $P(\lambda)$ then each $d_i(\lambda)$ can be written as

$$d_i(\lambda) = (\lambda - \lambda_0)^{c_i} \hat{d}_i(\lambda), \hat{d}_i(\lambda_0) \neq 0.$$

The factors $(\lambda - \lambda_0)^{c_i}, 1 \leqslant i \leqslant t$ are called the elementary divisors of $P(\lambda)$ associated with λ_0 and $c_1, \ldots, c_t$ are called the partial multiplicities of λ_0. The algebraic and geometric multiplicity of λ_0 as an eigenvalue of $P(\lambda)$ is the sum of all the partial multiplicities of λ_0 and the number of nonzero partial multiplicities of λ_0 respectively. Elementary divisors and partial multiplicities associated with an infinite eigenvalue of $P(\lambda)$ considered as grade k matrix polynomial are the elementary divisors and partial multiplicities associated with zero eigenvalue of $\operatorname{rev}_k P(\lambda)$ respectively. An eigenvalue λ_0 is said to be defective if its algebraic multiplicity is strictly more than its geometric multiplicity. Otherwise it is said to be semisimple. Clearly, if the matrix polynomial $P(\lambda)$ is regular, then the eigenvalues of $P(\lambda)$ are the roots of $\det(P(\lambda))$ with algebraic multiplicity equal to the multiplicity as a root of the polynomial $\det(P(\lambda))$.

Having an elementary divisor $(\lambda - \lambda_0)^r$ is also equivalent to the existence of vectors $x_0, \ldots, x_{r-1} \in \mathbb{C}^n$, $x_0 \neq 0$, satisfying the equations

$$\sum_{i=0}^{p} \frac{1}{i!} P^i(\lambda_0) x_{p-i} = 0, p = 0, \ldots, r-1$$

where $P^i(\lambda_0)$ denotes the i-th derivative of $P(\lambda)$ at λ_0 with respect to λ. The set of vectors $\{x_0, \ldots, x_{r-1}\}$ are said to form a Jordan chain of length r of $P(\lambda)$ corresponding to λ_0. The vector x_0 is called an eigenvector of $P(\lambda)$ corresponding to the eigenvalue λ_0 and $x_1, \ldots, x_{r-1}$ are generalized eigenvectors of $P(\lambda)$ corresponding to λ_0.

The following subspaces associated with $P(\lambda)$ are important.

Definition 1.2.1. *The right and left null spaces of an $m \times n$ matrix polynomial $P(\lambda)$, denoted by $N_r(P)$ and $N_l(P)$ respectively are defined as follows.*

$$N_r(P) := \{x(\lambda) \in \mathbb{C}(\lambda)^n : P(\lambda)x(\lambda) \equiv 0\},$$
$$N_l(P) := \{y(\lambda) \in \mathbb{C}(\lambda)^m : y(\lambda)^T P(\lambda) \equiv 0\}.$$

A vector polynomial is a vector whose entries are polynomials. For any subspace of $\mathbb{C}(\lambda)^n$, it is always possible to find a basis consisting entirely of vector polynomials. The degree of a vector polynomial is the greatest degree of its components, and the order of a polynomial basis is defined as the sum of the degrees of its vectors. Also any subspace of $\mathbb{C}(\lambda)^n$ has a polynomial basis of least order among all such bases and the ordered list of degrees of the vector polynomials in any such basis is always the same [27]. A minimal basis of the subspace is therefore defined as any polynomial basis of least order among all such bases and the minimal indices of the subspace are the ordered list of degrees of the vector polynomials in such a basis. In particular we have the following definitions.

Definition 1.2.2. *For a given $m \times n$ matrix polynomial $P(\lambda)$, a left minimal basis is a minimal basis of $N_l(P)$ and a right minimal basis is a minimal basis of $N_r(P)$.*

Definition 1.2.3. *For a given $m \times n$ matrix polynomial $P(\lambda)$, let $\{x_1(\lambda), \ldots, x_p(\lambda)\}$ be a right minimal basis and $\{y_1(\lambda), \ldots, y_q(\lambda)\}$ be a left minimal basis such that*

$$\deg x_1 \leqslant \cdots \leqslant \deg x_p \text{ and } \deg y_1 \leqslant \cdots \leqslant \deg y_q.$$

Setting $\epsilon_i = \deg x_i, i = 1, \ldots, p$, and $\eta_j = \deg y_j$, $j = 1, \ldots, q$, the right and left minimal indices of $P(\lambda)$ are defined as $\epsilon_1 \leqslant \cdots \leqslant \epsilon_p$, and $\eta_1 \leqslant \cdots \leqslant \eta_q$ respectively.

When $P(\lambda)$ is a square regular matrix polynomial, the associated polynomial eigenvalue problem (PEP) consists of finding the finite and infinite eigenvalues and corresponding eigenvectors. When $P(\lambda)$ is not regular in addition to the eigenvalues and corresponding elementary divisors, the left and right minimal bases and indices also have to be calculated. This is called the complete eigenvalue problem (CEP) associated with $P(\lambda)$. If $\deg P = 1$, then the polynomial eigenvalue problem is also called a generalized eigenvalue problem.

1.2.2 Canonical forms associated with matrices and matrix pencils

Various canonical forms associated with matrices and matrix pencils will be used throughout the thesis.

Jordan canonical form for matrices. For a matrix A of size $n \times n$ there exists an $n \times n$ invertible matrix P, such that PAP^{-1} is a block diagonal with each block is of the form $J_q(\alpha) := \begin{bmatrix} \alpha & 1 & & \\ & \alpha & \ddots & \\ & & \ddots & 1 \\ & & & \alpha \end{bmatrix} \in \mathbb{C}^{q\times q}$, called a Jordan block associated with the eigenvalue α.

Weierstrass canonical form for regular matrix pencils. For a regular matrix pencil $L(\lambda)$ of size $n\times n$, there exist two $n\times n$ invertible matrices P and Q such that $PL(\lambda)Q$ is block diagonal with each diagonal block being of the form $\lambda I_q - J_q(\alpha)$ or $\lambda J_q(0) - I_q$ where $J_q(\alpha)$ is a $q \times q$ Jordan block associated with some $\alpha \in \mathbb{C}$. The blocks $\lambda I_q - J_q(\alpha)$ and $\lambda J_q(0) - I_q$ correspond to a finite eigenvalue α and an infinite eigenvalue respectively. The size of the largest block of type $\lambda J_q(0) - I_q$ in the Weierstrass from is the index of the pencil $L(\lambda)$.

Kronecker canonical form for matrix pencils. For a matrix pencil $L(\lambda)$ of size $m \times n$ there exist two invertible matrices P and Q of size $m \times m$ and $n \times n$ respectively such that $PL(\lambda)Q$ is a block diagonal with each diagonal block being of one of the following forms:

$$\lambda I_q - J_q(\alpha),\ \lambda J_q(0) - I_q,\ \lambda G_q - F_q,\ \lambda G_q^T - F_q^T.$$

The first two blocks are as in the above mentioned Weierstrass form and

$$F_q := \begin{bmatrix} 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} \in \mathbb{C}^{q\times(q+1)} \text{ and } G_q := \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \end{bmatrix} \in \mathbb{C}^{q\times(q+1)}.$$

The blocks $\lambda I_q - J_q(\alpha)$ and $\lambda J_q(0) - I_q$ correspond to a finite eigenvalue α and an infinite eigenvalue respectively. The blocks $\lambda G_q - F_q$ and $\lambda G_q^T - F_q^T$ known as right and left singular bolck respectively, appear only when $L(\lambda)$ is singular and correspond to a right minimal index and a left minimal index respectively. Also clearly $L(\lambda)$ has an elementary divisor of the form $(\lambda - \lambda_0)^j$ associated with an eigenvalue λ_0 if and only if $\lambda I_j - J_j(\lambda_0)$ is a block of the Kronecker canonical form.

Canonical form of Hermitian matrix pencils under congruence. In [57] it was established that every Hermitian pencil $L(\lambda)$ is congruent to a Hermitian matrix pencil of the form

$$
\begin{aligned}
&0_{u\times u} \oplus \left(\lambda \begin{bmatrix} 0 & 0 & F_{\epsilon_1} \\ 0 & 0 & 0 \\ F_{\epsilon_1} & 0 & 0 \end{bmatrix} - G_{2\epsilon_1+1}\right) \\
&\oplus \cdots \oplus \left(\lambda \begin{bmatrix} 0 & 0 & F_{\epsilon_p} \\ 0 & 0 & 0 \\ F_{\epsilon_p} & 0 & 0 \end{bmatrix} - G_{2\epsilon_p+1}\right) \\
&\oplus \delta_1(\lambda G_{k_1} - F_{k_1}) \oplus \cdots \oplus \delta_r(\lambda G_{k_r} - F_{k_r}) \\
&\oplus \eta_1((\lambda - \alpha_1)F_{l_1} - G_{l_1}) \oplus \cdots \oplus \eta_q((\lambda - \alpha_q)F_{l_q} - G_{l_q}) \\
&\oplus \left(\begin{bmatrix} 0 & (\lambda - \beta_1)F_{m_1} \\ (\lambda - \bar{\beta}_1)F_{m_1} & 0 \end{bmatrix} - \begin{bmatrix} 0 & G_{m_1} \\ G_{m_1} & 0 \end{bmatrix}\right) \\
&\oplus \cdots \oplus \left(\begin{bmatrix} 0 & (\lambda - \beta_s)F_{m_s} \\ (\lambda - \bar{\beta}_s)F_{m_s} & 0 \end{bmatrix} - \begin{bmatrix} 0 & G_{m_s} \\ G_{m_s} & 0 \end{bmatrix}\right).
\end{aligned}
\tag{1.2.1}
$$

Here, $\epsilon_1 \leq \cdots \leq \epsilon_p$ and $k_1 \leq \cdots \leq k_r$ are positive integers, α_j are real numbers, β_j are complex nonreal numbers, $\delta_1, \delta_2, \ldots, \delta_r, \eta_1, \eta_2, \ldots, \eta_q$ are signs each equal to $+1$ or -1 called the signs of $L(\lambda)$, and F_m and G_m are the $m \times m$ matrices

$$
F_m = \begin{bmatrix} 0 & \cdots & \cdots & 0 & 1 \\ \vdots & & & 1 & 0 \\ \vdots & & & & \vdots \\ 0 & 1 & & & \vdots \\ 1 & 0 & \cdots & \cdots & 0 \end{bmatrix}, \quad G_m = \begin{bmatrix} F_{m-1} & 0 \\ 0 & 0 \end{bmatrix}
$$

with $F_1 = [1]$.

The form (1.2.1) is uniquely determined by $L(\lambda)$ up to a combination of the following permutations: a permutation of the blocks

$$\lambda \begin{bmatrix} 0 & 0 & F_{\epsilon_j} \\ 0 & 0 & 0 \\ F_{\epsilon_j} & 0 & 0 \end{bmatrix} - G_{2\epsilon_j+1}, \quad j = 1, 2, \ldots, p;$$

a permutation of the blocks

$$\delta_j(\lambda G_{k_j} - F_{k_j}), \quad j = 1, 2, \ldots, r;$$

a permutation of the blocks

$$\eta_j((\lambda - \alpha_j)F_{l_j} - G_{l_j}), \quad j = 1, 2, \ldots, q;$$

and a permutation of the blocks

$$\begin{bmatrix} 0 & (\lambda - \beta_j)F_{m_j} \\ (\lambda - \bar{\beta}_j)F_{m_j} & 0 \end{bmatrix} - \begin{bmatrix} 0 & G_{m_j} \\ G_{m_j} & 0 \end{bmatrix}, \quad j = 1, 2, \ldots, s;$$

with possible replacement of β_j by $\bar{\beta}_j$ within each such block.

1.2.3 Linearization of matrix polynomials

The most widely used approach for solving polynomial eigenvalue problems is linearization. Linearization is the process of converting the polynomial eigenvalue problem into a equivalent generalized eigenvalue problem associated with a larger matrix pencil.

Definition 1.2.4 (Linearization). *An $(m+s)\times(n+s)$ matrix pencil $L(\lambda) = \lambda X + Y$, is a linearization of an $m \times n$ matrix polynomial $P(\lambda)$ of grade k if there exist two unimodular matrix polynomials $E(\lambda) \in \mathbb{C}[\lambda]^{(m+s)\times(m+s)}$ and $F(\lambda) \in \mathbb{C}[\lambda]^{(n+s)\times(n+s)}$ for some positive integer s such that*

$$E(\lambda)L(\lambda)F(\lambda) = \begin{bmatrix} P(\lambda) & \\ & I_s \end{bmatrix}.$$

For example the first and second Frobenius companion forms $C_1(\lambda)$ and $C_2(\lambda)$ given by

$$C_1(\lambda) := \lambda \begin{bmatrix} A_k & & & \\ & I_n & & \\ & & \ddots & \\ & & & I_n \end{bmatrix} + \begin{bmatrix} A_{k-1} & A_{k-2} & \cdots & A_0 \\ -I_n & & & 0 \\ & \ddots & & \\ & & -I_n & 0 \end{bmatrix} \tag{1.2.2}$$

$$C_2(\lambda) := \lambda \begin{bmatrix} A_k & & & \\ & I_m & & \\ & & \ddots & \\ & & & I_m \end{bmatrix} + \begin{bmatrix} A_{k-1} & -I_m & & \\ A_{k-2} & & \ddots & \\ \vdots & & & -I_m \\ A_0 & 0 & & 0 \end{bmatrix} \tag{1.2.3}$$

are linearizations of $P(\lambda)$ with $s = (k-1)n$ and $s = (k-1)m$ respectively. The unimodular matrix polynomials $E(\lambda)$ and $F(\lambda)$ that appear in the relation

$$E(\lambda)C_1(\lambda)F(\lambda) = \begin{bmatrix} P(\lambda) & \\ & I_{(k-1)n} \end{bmatrix},$$

are

$$F(\lambda) = \begin{bmatrix} \lambda^{k-1}I_n & \cdots & \lambda I_n & I_n \\ \vdots & \iddots & I_n & \\ \lambda I_n & \iddots & & \\ I_n & & & \end{bmatrix} \text{ and } E(\lambda) = \begin{bmatrix} I_m & E_{k-1}(\lambda) & E_{k-2}(\lambda) & \cdots & E_1(\lambda) \\ & & & & -I_n \\ & & & \iddots & \\ & & -I_n & & \\ 0 & -I_n & & & \end{bmatrix}$$

where $E_k(\lambda) = A_k$ and $E_{i-1}(\lambda) = A_{i-1} + \lambda E_i(\lambda), i = k, \ldots, 2.$

And the unimodular matrix polynomials $\hat{F}(\lambda)$ and $\hat{E}(\lambda)$ that appear in the relation

$$\hat{F}(\lambda)C_2(\lambda)\hat{E}(\lambda) = \begin{bmatrix} P(\lambda) & \\ & I_{(k-1)m} \end{bmatrix},$$

are $\hat{F}(\lambda) = \begin{bmatrix} \lambda^{k-1}I_m & \cdots & \lambda I_m & I_m \\ \vdots & \iddots & I_m & \\ \lambda I_m & \iddots & & \\ I_m & & & \end{bmatrix}$ and $\hat{E}(\lambda) = \begin{bmatrix} I_n & & & & 0 \\ E_{k-1}(\lambda) & & & & -I_m \\ E_{k-2}(\lambda) & & & -I_m & \\ \vdots & & \iddots & & \\ E_1(\lambda) & -I_m & & & \end{bmatrix}$

where $E_k(\lambda) = A_k$ and $E_{i-1}(\lambda) = A_{i-1} + \lambda E_i(\lambda), i = k, \dots, 2$. It is clear that a matrix polynomial and its linearization have the same finite eigenvalues and corresponding elementary divisors (for details, see [32]). However, if the same is to be guaranteed for the eigenvalue at infinity also, then the linearization has to be a strong linearization of $P(\lambda)$.

Definition 1.2.5 (Strong Linearization). *A linearization $L(\lambda) = \lambda X + Y$ of a matrix polynomial $P(\lambda)$ is called a strong linearization of $P(\lambda)$ considered as a matrix polynomial of grade k if $\operatorname{rev}_1 L(\lambda)$ is a linearization of $\operatorname{rev}_k P(\lambda)$.*

1.3 Some block Toeplitz matrices associated with matrix polynomials

Block Toeplitz matrices associated with $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ of the form

$$C_j(P) = \underbrace{\begin{bmatrix} A_0 & & & \\ A_1 & A_0 & & \\ \vdots & A_1 & \ddots & \\ A_k & \vdots & \ddots & A_0 \\ & A_k & & A_1 \\ & & \ddots & \vdots \\ & & & A_k \end{bmatrix}}_{j+1 \text{ block columns}} \quad \text{for } j = 0, 1, \dots. \tag{1.3.1}$$

and the form

$$\bar{T}_j(P) = \begin{bmatrix} A_0 & & \\ \vdots & \ddots & \\ A_{j-1} & \cdots & A_0 \end{bmatrix} \quad \text{for } j = 1, 2, \dots. \tag{1.3.2}$$

play an important role in study of matrix polynomials. The matrices $C_j(P)$ are called convolution or Sylvester matrices associated with $P(\lambda)$ of level j. Both classes of matrices have been used to understand the algebraic structure of matrix polynomials. The convolution matrices $C_j(Q)$ and $C_j(Q^T)$ associated with a matrix polynomial $Q(\lambda)$ and its transpose $Q(\lambda)^T$ are used to analyse and compute its left and right minimal bases and indices when $Q(\lambda)$ is singular whereas the matrices

$\bar{T}_j(Q)$ and $\bar{T}_j(Q^T)$ are used to analyse the left and right Jordan chain structure of $Q(\lambda)$ at zero [2, 4, 29, 34, 45, 46, 76].

The following lemma which states some important and useful properties of the convolution matrices can be easily proved.

Lemma 1.3.1. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ and $Q(\lambda) = \sum_{i=0}^{l} \lambda^i B_i$ be matrix polynomials of grade k and l respectively and $C_j(P)$ and $C_j(Q)$ for $j = 0, 1, 2, \ldots,$ be the convolution matrices associated with $P(\lambda)$ and $Q(\lambda)$ respectively.*

(a) If $P(\lambda)$ and $Q(\lambda)$ are of same size and grade then $C_j(P+Q) = C_j(P)+C_j(Q)$ for all j.

(b) $\|C_j(P)\|_F = \sqrt{j+1}\|P\|_F$ for all j.

(c) If the product $P(\lambda)Q(\lambda)$ is defined, then considering it as a grade $k+l$ matrix polynomial, we have $C_0(PQ) = C_l(P)C_0(Q)$.

In Chapter 3 we require block Toeplitz matrices which are of a slightly different form than the one already defined in (1.3.1). Let $\widehat{C}_j(P)$ denote the j^{th} convolution matrix associated with $\text{rev}_k P(\lambda)$. Then

$$\widehat{C}_j(P) = \underbrace{\begin{bmatrix} A_k & & & \\ A_{k-1} & A_k & & \\ \vdots & A_{k-1} & \ddots & \\ A_0 & \vdots & \ddots & A_k \\ & A_0 & & A_{k-1} \\ & & \ddots & \vdots \\ & & & A_0 \end{bmatrix}}_{j+1 \text{ block columns}} \quad \text{for } j = 0, 1, 2, \ldots. \tag{1.3.3}$$

Note that Lemma 1.3.1 is also true for the matrices $\widehat{C}_j(P)$ associated with $P(\lambda)$. The next result is about the singular values of $\widehat{C}_j(G)$ of some level j associated with a right singular block G. A proof of this result is available in [22]. Our proof was made independently with different arguments.

Lemma 1.3.2. *Suppose* $H_{k-1}(\lambda) = \begin{bmatrix} -1 & \lambda & & & \\ & -1 & \lambda & & \\ & & \ddots & \ddots & \\ & & & -1 & \lambda \end{bmatrix} \in \mathbb{C}^{(k-1)\times k}$ *for any positive integer* $k > 1$. *Then*

$$\sigma_{\min}(\widehat{C}_{k-1}(H_{k-1})) = \sigma_{\min}(\widehat{C}_{k-2}(H_{k-1})) = 2\sin\left(\frac{\pi}{4k-2}\right).$$

Proof. The result is obvious for $k = 2$. Therefore we assume that $k > 2$. For simplicity, we denote $\widehat{C}_j(H_{k-1})$ by $\widehat{C}_j$ for $j = k-1, k-2$. To complete the proof it is enough to show that both the matrices $\widehat{C}_{k-1}$ and $\widehat{C}_{k-2}$ are full rank and the smallest nonzero eigenvalues of $S_{k-2} := \widehat{C}_{k-2}^*\widehat{C}_{k-2}$ and $S_{k-1} := \widehat{C}_{k-1}^*\widehat{C}_{k-1}$ are both equal to $2 + 2\cos\left(\frac{2(k-1)\pi}{2k-1}\right)$. This is because in such a case, the smallest singular values of $\widehat{C}_{k-1}$ and $\widehat{C}_{k-2}$ will both be equal to

$$\sqrt{2 + 2\cos\left(\frac{2(k-1)\pi}{2k-1}\right)} = \sqrt{\left(2\sin\left(\frac{\pi}{4k-2}\right)\right)^2} = 2\sin\left(\frac{\pi}{4k-2}\right).$$

For $j \geqslant 1$, let

$$D_j = \begin{bmatrix} 1 & & & & \\ & 2 & & & \\ & & \ddots & & \\ & & & 2 & \\ & & & & 1 \end{bmatrix}_{j\times j}, L_j = \begin{bmatrix} 0 & & & \\ -1 & 0 & & \\ & \ddots & \ddots & \\ & & -1 & 0 \end{bmatrix}_{j\times j} \text{ and,}$$

$$T_j = D_j + L_j + L_j^T + e_je_j^T, \hat{T}_j = D_j + L_j + L_j^T + e_1e_1^T,$$

where e_j is the j-th column of I_j. Now a simple multiplication shows that

$$S_{k-2} = \underbrace{\begin{bmatrix} D_k & L_k^T & & & & \\ L_k & D_k & L_k^T & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & L_k & D_k & L_k^T \\ & & & & L_k & D_k \end{bmatrix}}_{(k-1) \text{ block columns.}} \text{ and } S_{k-1} = \underbrace{\begin{bmatrix} D_k & L_k^T & & & & \\ L_k & D_k & L_k^T & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & L_k & D_k & L_k^T \\ & & & & L_k & D_k \end{bmatrix}}_{k \text{ block columns.}}.$$

To find the smallest nonzero eigenvalue of S_{k-2}, consider the permutation matrix

$$\tilde{\mathcal{P}} = \begin{bmatrix} \tilde{\mathcal{P}}_1 & \tilde{\mathcal{P}}_2 & \cdots & \tilde{\mathcal{P}}_{k+1} \end{bmatrix} \in \mathbb{C}^{k(k-1)\times k(k-1)}$$

where

$$\tilde{\mathcal{P}}_i = \begin{cases} \begin{bmatrix} \tilde{e}_i & \tilde{e}_{i+(k+1)} & \cdots & \tilde{e}_{i+(k-2)(k+1)} \end{bmatrix}, & \text{for } i \leqslant 2, \\ \begin{bmatrix} \tilde{e}_i & \tilde{e}_{i+(k+1)} & \cdots & \tilde{e}_{i+(k-3)(k+1)} \end{bmatrix}, & \text{for } 3 \leqslant i \leqslant k+1. \end{cases}$$

Here $\tilde{e}_j$ is the j-th column of $I_{k(k-1)}$. Then $\tilde{\mathcal{P}}^T S_{k-2} \tilde{\mathcal{P}}$ is a block diagonal matrix with $(k+1)$ blocks where T_{k-1} is the first block, $\hat{T}_{k-1}$ is the second block, and $\begin{bmatrix} \hat{T}_{k-j+1} & \\ & T_{j-3} \end{bmatrix}$, $j = 3, 4, \ldots, (k+1)$, is the j-th block, with T_0 and $\hat{T}_0$ being empty matrices.

Clearly the first two blocks have the same eigenvalues and the sub-blocks of all other blocks are principal submatrices of the first or second block. Hence by using the inclusion principal for eigenvalues of Hermitian matrices [43, Theorem 4.3.15] we can say that, the smallest eigenvalue of S_{k-2} is the smallest eigenvalue of any one of the first two blocks, in particular of the second block $\hat{T}_{k-1}$.

To find the smallest nonzero eigenvalue of S_{k-1}, consider the permutation matrix

$$\hat{\mathcal{P}} = \begin{bmatrix} \hat{\mathcal{P}}_1 & \hat{\mathcal{P}}_2 & \cdots & \hat{\mathcal{P}}_{k+1} \end{bmatrix} \in \mathbb{C}^{k^2\times k^2}$$

where

$$\hat{\mathcal{P}}_i = \begin{cases} \begin{bmatrix} \hat{e}_i & \hat{e}_{i+(k+1)} & \cdots & \hat{e}_{i+(k-1)(k+1)} \end{bmatrix}, & \text{for } i = 1, \\ \begin{bmatrix} \hat{e}_i & \hat{e}_{i+(k+1)} & \cdots & \hat{e}_{i+(k-2)(k+1)} \end{bmatrix}, & \text{for } 2 \leqslant i \leqslant k+1. \end{cases}$$

Here $\hat{e}_j$ is the j-th column of I_{k^2}. Then $\hat{\mathcal{P}}^T S_{k-1} \hat{\mathcal{P}}$ is a block diagonal matrix of $(k+1)$ blocks where the first block is $D_k + L_k + L_k^T$, the second block is $\hat{T}_{k-1}$, and the j-th block is $\begin{bmatrix} \hat{T}_{k-j+1} & \\ & T_{j-2} \end{bmatrix}$ for $j = 3, 4, \ldots, (k+1)$ with $\hat{T}_0$ being the empty matrix.

Clearly 0 is an eigenvalue of the first block and the second block can be obtained by removing the first row and first column of the first block. Again the sub-blocks of all other blocks are either submatrices of the second block $\hat{T}_{k-1}$ or of T_{k-1}. Since T_{k-1} and $\hat{T}_{k-1}$ have the same eigenvalues, hence by using the inclusion principle for eigenvalues of Hermitian matrices we can say that, the smallest eigenvalue of S_{k-1} is

0 and the second smallest eigenvalue is the smallest eigenvalue of the second block $\hat{T}_{k-1}$.

From [59, 81] the smallest eigenvalue of $\hat{T}_{k-1}$ is $2 + 2\cos\left(\frac{2(k-1)\pi}{2k-1}\right)$. Hence $2 + 2\cos\left(\frac{2(k-1)\pi}{2k-1}\right)$ is the smallest nonzero eigenvalue of both the matrices S_{k-1} and S_{k-2}. This completes the proof. $\square$

1.4 Problem definition

The following problems will be considered in this thesis.

Problem 1. Given a rectangular matrix polynomial $P(\lambda)$, find vector spaces that are a rich source of strong linearizations of $P(\lambda)$ with the following features:

(a) The vector spaces have properties akin to the spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ proposed in [61] for the study of regular matrix polynomials and extended in [14] to square singular matrix polynomials, and coincide with them when $P(\lambda)$ is square.

(b) The minimal bases and indices of $P(\lambda)$ can be easily recovered from those of the linearizations of $P(\lambda)$ arising from $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$.

Problem 2. Let $L(\lambda)$ be a strong linearization of a rectangular matrix polynomial $P(\lambda)$ arising from the vector space framework proposed as a solution of Problem 1. The computed solution of the complete eigenvalue problem for $L(\lambda)$ via a backward stable algorithm is the exact solution of the problem for some perturbed pencil $L(\lambda) + \Delta L(\lambda)$ where $\|\Delta L\|/\|L\|$ is of the order of unit roundoff $\mathbf{u}$. Investigate the following:

(a) Is $L(\lambda)+\Delta L(\lambda)$ a strong linearization of some polynomial $P(\lambda)+\Delta P(\lambda)$ where $\|\Delta P\|/\|P\| \approxeq O(\mathbf{u})$?

(b) Are the recovery rules for extracting minimal indices of $P(\lambda) + \Delta P(\lambda)$ from those of $L(\lambda)+\Delta L(\lambda)$ the same as the corresponding rules for $P(\lambda)$ from $L(\lambda)$?

(c) What choice of linearizations arising from the proposed vector spaces satisfy (a) and (b)?

Problem 3. Given a square matrix polynomial $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$, $\lambda_0 \in \mathbb{C}$ and $r > 1$, find the distance to a nearest matrix polynomial having $(\lambda - \lambda_0)^j, j \geqslant r$, as an elementary divisor, with respect to the norms $\|P\|_F := \|[A_0 \cdots A_k]\|_F$, $\|P\|_2 := \|[A_0 \cdots A_k]\|_2$ and $\|P\|_{2,\infty} := \max_{0\leq i\leq k} \|A_i\|_2$.

Problem 4. Given an $n \times n$ matrix polynomial $P(\lambda)$, find the distance to a nearest matrix polynomial of normal rank at most r, $(r < n)$, with respect to the norms $|||\cdot|||_F$ and $|||\cdot|||_2$.

1.5 Minimal norm mapping

Several perturbations of minimal norm that are constructed in the thesis depend on the following theorem.

Theorem 1.5.1. *[71] Let $A \in \mathbb{C}^{p\times m}$, $B \in \mathbb{C}^{n\times q}$, and $C \in \mathbb{C}^{p\times q}$ be given. Define*

$$\mathcal{S} = \{E \in \mathbb{C}^{m\times n} : AEB = C\}.$$

Then $\mathcal{S} \neq \emptyset$ if and only if A, B, C satisfy

$$AA^{\dagger}CB^{\dagger}B = C.$$

If $\mathcal{S} \neq \emptyset$, any $E \in \mathcal{S}$ can be expressed by

$$E = A^{\dagger}CB^{\dagger} + Z - A^{\dagger}AZBB^{\dagger},$$

where $Z \in \mathbb{C}^{m\times n}$. Moreover, there is a unique matrix $\hat{E} \in \mathcal{S}$ expressed by

$$\hat{E} = A^{\dagger}CB^{\dagger}$$

satisfying $\|\hat{E}\|_s = \inf\{\|E\|_s : E \in \mathcal{S}\}$ where $s = 2$ or F.

1.6 Structured singular value

A perturbation class S is a nonempty subset of $\mathbb{C}^{p\times q}$.

Definition 1.6.1 (μ-value). *[65, 47] Let $S \subset \mathbb{C}^{p\times q}$ be a perturbation class and let $\|\cdot\|$ be a norm on $\mathbb{C}^{p\times q}$. The μ-value of $M \in \mathbb{C}^{q\times p}$ with respect to S and $\|\cdot\|$ is*

$$\mu_{S,\|\cdot\|}(M) := (\inf\{\|\Delta\| : \Delta \in S, \operatorname{rank}(I - \Delta M) \leqslant p - 1\})^{-1}. \tag{1.6.1}$$

If there is no such $\Delta \in S$, then $\mu_{S,\|\cdot\|}(M) = 0$.

The generalized μ-value is now defined as follows.

Definition 1.6.2 (Generalized μ-value). *Let $S \subset \mathbb{C}^{p\times q}$ be a perturbation class and let $\|\cdot\|$ be a norm on $\mathbb{C}^{p\times q}$. The generalized μ-value of $M \in \mathbb{C}^{q\times p}$ with respect to S and $\|\cdot\|$ is defined as*

$$\mu^r_{S,\|\cdot\|}(M) := (\inf\{\|\Delta\| : \Delta \in S, \operatorname{rank}(I - \Delta M) \leqslant p - r\})^{-1}. \tag{1.6.2}$$

If there is no such $\Delta \in S$, then $\mu^r_{S,\|\cdot\|}(M) = 0$.

The above definition was first introduced in [39]. The following lemma, the proof of which is evident from elementary properties of singular values, will be used in the subsequent chapters.

Lemma 1.6.3. *Let $M \in \mathbb{C}^{p\times q}$ and $N \in \mathbb{C}^{q\times p}$. Then for $i = 1, \ldots, \min\{p, q\}$,*

$$\inf\{\|M\|_2 : \mathrm{nullity}(I - MN) \geqslant i\} = (\sigma_i(N))^{-1}.$$

2

Vector spaces of generalized linearizations

Given an $n \times n$ regular matrix polynomial $P(\lambda)$ of degree k, one of the first systematic studies of linearizations to be undertaken was [61] which introduced the following vector spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ of matrix pencils as sources of linearizations of $P(\lambda)$.

$$\mathbb{L}_1(P) := \{L(\lambda) = \lambda X + Y : L(\lambda)(\Lambda_k(\lambda) \otimes I_n) = v \otimes P(\lambda), v \in \mathbb{C}^k\}, \tag{2.0.1}$$

$$\mathbb{L}_2(P) := \{L(\lambda) = \lambda X + Y : (\Lambda_k(\lambda)^T \otimes I_n)L(\lambda) = w^T \otimes P(\lambda), w \in \mathbb{C}^k\} \tag{2.0.2}$$

$$\text{where } \Lambda_k(\lambda) := [\lambda^{k-1} \cdots \lambda \;\; 1]^T. \tag{2.0.3}$$

The defining identities in (2.0.1) and (2.0.2) are called the right and left ansatz equations respectively and the corresponding vectors v in (2.0.1) and the vector w in (2.0.2) are called right and left ansatz vectors. The first and second companion linearizations given by (1.2.2) and (1.2.3) belong to $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ respectively. This work gave a whole new direction to research in the theory of linearizations due to the special properties of these vector spaces. For instance, it was shown that constructing pencils in these spaces corresponding to a given ansatz vector is very simple and almost all the resulting pencils are strong linearizations of $P(\lambda)$ from which the eigenvalues and corresponding eigenvectors can be easily recovered. In particular in [24, 42, 60], it was shown that if $P(\lambda)$ has some special structure like, Hermitian, symmetric, $\star$-alternating and $\star$-palindromic, then there exist subspaces of $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ with the property that almost every pencil of the subspace is a structure preserving linearization of $P(\lambda)$ from which both finite and infinite eigenvalues of $P(\lambda)$ and corresponding eigenvectors can be easily recovered.

It was shown in [14] that even when $P(\lambda)$ is square but singular, almost every pencil in $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ is a linearization of $P(\lambda)$ from which the solution of the complete eigenvalue problem for $P(\lambda)$ can be easily recovered. The vector space setting for constructing linearizations has since been extended to cover other polynomial bases [25] and inspired further work that throws fresh light on these spaces [64]. Since then the literature on linearizations has expanded rapidly and various linearizations that take different factors into consideration have been proposed. Important choices of linearizations not covered by $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ are the Fiedler pencils and their generalizations [3, 9, 11, 15, 80] which are also sources of linearizations for non-square matrix polynomials [17]. Systematic studies of linearizations that cover both square and non-square linearizations are relatively recent in the literature. For example, [22] introduced the framework of block minimal bases pencils as potential linearizations of rectangular matrix polynomials with focus on particular subclasses like the block Kronecker pencils. These ideas were further extended in [10]. Inspired by [61], the recent work [26] considers linearizations of rectangular matrix polynomials in a vector space setting. Referred to as block Kronecker ansatz spaces, these vector spaces contain block Kronecker linearizations as well as Fiedler linearizations and their extensions modulo permutations and share some of the important properties that $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ have when $P(\lambda)$ is square. However, the block Kronecker ansatz spaces do not become $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ when $P(\lambda)$ is square.

In this chapter, we generalize the spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ to the case when $P(\lambda)$ is not square by forming vector spaces of matrix pencils that have some of the key features of $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ and coincide with them when $P(\lambda)$ is square. For this purpose we define generalized linearizations of matrix polynomials (which we refer to in short as g-linearizations) and their strong versions and show that the proposed vector spaces have all the properties with respect to being g-linearizations that $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ are shown to possess with respect to being linearizations of square singular polynomials in [14]. In particular the matrix pencils in these spaces can be easily constructed from the coefficient matrices of $P(\lambda)$ in a manner very similar to the ones in $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ when $P(\lambda)$ is square. We show that the solution of the complete eigenvalue problem for $P(\lambda)$ can be easily recovered from that of almost every pencil in these spaces. Finally we show that almost every g-linearization in these spaces may be 'trimmed' to produce strong linearizations of $P(\lambda)$.

2.1 Generalized linearizations of matrix polynomials

Definition 2.1.1 (g-Linearization)**.** *An $km \times kn$ matrix pencil $L(\lambda) = \lambda X + Y$ is called a g-linearization of an $m \times n$ matrix polynomial $P(\lambda)$ of grade k if there exist* *two unimodular matrices* $E(\lambda) \in \mathbb{C}[\lambda]^{mk\times mk}$ *and* $F(\lambda) \in \mathbb{C}[\lambda]^{nk\times nk}$ *such that*

$$E(\lambda)L(\lambda)F(\lambda) = \begin{bmatrix} P(\lambda) & \\ & I_{k-1}\otimes I_{m,n} \end{bmatrix}$$

Here $I_{m,n} = \begin{bmatrix} I_n \\ 0_{(m-n)\times n} \end{bmatrix}$ *if* $m > n$ *and* $I_{m,n} = \begin{bmatrix} I_m & 0_{m\times(n-m)} \end{bmatrix}$ *if* $m < n$. *Otherwise* $I_{m,n} = I_m = I_n$.

Definition 2.1.2 (Strong g-Linearization). *A matrix pencil* $L(\lambda) = \lambda X + Y$ *with* $X, Y \in \mathbb{C}^{mk\times nk}$ *is a strong g-linearization of an* $m \times n$ *matrix polynomial* $P(\lambda)$ *of grade* k *if* $L(\lambda)$ *is a g-linearization of* $P(\lambda)$ *and* $\mathrm{rev}_1 L(\lambda)$ *is a g-linearization of* $\mathrm{rev}_k P(\lambda)$.

From the above definition, it is clear that every linearization of a square matrix polynomial is also a generalized linearization, which justifies our choice for the term. Also, evidently a matrix polynomial has the same finite eigenvalues and elementary divisors as its g-linearization and the same finite and infinite eigenvalues and elementary divisors as its strong g-linearization. Therefore, to establish that the solution of a complete eigenvalue problem for a rectangular matrix polynomial can be obtained from a given strong g-linearization, it is enough to show that the minimal bases and indices of the polynomial can be easily recovered from the g-linearization.

2.1.1 The vector spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$

To extend the work in [61] to non-square matrix polynomials, we propose the following vector spaces associated with a matrix polynomial $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i \in \mathbb{C}[\lambda]^{m\times n}$ of grade k, which we continue to denote by $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ for ease of notation.

$$\mathbb{L}_1(P) := \{L(\lambda) = \lambda X + Y : L(\lambda)(\Lambda_k(\lambda) \otimes I_n) = v \otimes P(\lambda), v \in \mathbb{C}^k\}, \tag{2.1.1}$$

$$\mathbb{L}_2(P) := \{L(\lambda) = \lambda X + Y : (\Lambda_k(\lambda)^T \otimes I_m)L(\lambda) = w^T \otimes P(\lambda), w \in \mathbb{C}^k\}, \tag{2.1.2}$$

where $\Lambda_k(\lambda)$ is as in (2.0.3).

Following [61] we will refer to the vector v (w) in the identity (2.1.1), ((2.1.2)) satisfied by $L(\lambda) \in \mathbb{L}_1(P)$, ($L(\lambda) \in \mathbb{L}_2(P)$) as the right (left) ansatz vector corresponding to $L(\lambda)$. The sets $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ are not empty as $C_1^g(\lambda) := \lambda X_1 + Y_1 \in \mathbb{L}_1(P)$ with right ansatz vector $v = e_1 \in \mathbb{C}^k$, where $X_1 = \begin{bmatrix} A_k & & & \\ & I_{m,n} & & \\ & & \ddots & \\ & & & I_{m,n} \end{bmatrix}$, $Y_1 = \begin{bmatrix} A_{k-1} & A_{k-2} & \cdots & A_0 \\ -I_{m,n} & & & 0 \\ & \ddots & & \\ & & -I_{m,n} & 0 \end{bmatrix}$

and $C_2^g(\lambda) := \lambda X_2 + Y_2 \in \mathbb{L}_2(P)$ with left ansatz vector $w = e_1 \in \mathbb{C}^k$,

where $X_2 = \begin{bmatrix} A_k & & & \\ & I_{m,n} & & \\ & & \ddots & \\ & & & I_{m,n} \end{bmatrix}$, $Y_2 = \begin{bmatrix} A_{k-1} & -I_{m,n} & & \\ A_{k-2} & & \ddots & \\ & & & -I_{m,n} \\ A_0 & 0 & & 0 \end{bmatrix}$.

As Theorem 2.1.7 and Theorem 2.1.8 show, if $m \geqslant n$ then $C_1^g(\lambda)$ is a strong g-linearization of $P(\lambda)$ and if $m \leqslant n$ then $C_2^g(\lambda)$ is a strong g-linearization of $P(\lambda)$.

For any matrix polynomial $P(\lambda)$, clearly $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ are vector spaces over $\mathbb{C}$. In this section we find some important properties of these vector spaces. The results show that if the $m \times n$ matrix polynomial $P(\lambda)$ is tall, i.e., $m \geqslant n$, then the properties of $\mathbb{L}_1(P)$ with respect to g-linearizations are very similar to those of the corresponding space for square matrix polynomials considered in [61] and [14] with respect to linearizations. The same is true of $\mathbb{L}_2(P)$ when $P(\lambda)$ is broad, i.e., $m \leqslant n$.

For the case $m = n$, the matrix pencils in $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ were originally characterized in [61] by introducing special operations on block matrices called column shifted sums and row shifted sums respectively. We state these definitions with the aim of showing that the same characterizations also hold when $m \neq n$.

Definition 2.1.3 (Column and row shifted sums). *Let X and Y be block matrices*

$$X = \begin{bmatrix} X_{11} & \cdots & X_{1k} \\ X_{21} & \cdots & X_{2k} \\ \vdots & \ddots & \vdots \\ X_{k1} & \cdots & X_{kk} \end{bmatrix}, Y = \begin{bmatrix} Y_{11} & \cdots & Y_{1k} \\ Y_{21} & \cdots & Y_{2k} \\ \vdots & \ddots & \vdots \\ Y_{k1} & \cdots & Y_{kk} \end{bmatrix}$$

with blocks $X_{ij}, Y_{ij} \in \mathbb{C}^{m\times n}$ *then the operations*

$$X \boxplus\!\!\!\rightarrow Y := \begin{bmatrix} X_{11} & \cdots & X_{1k} & 0 \\ X_{21} & \cdots & X_{2k} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ X_{k1} & \cdots & X_{kk} & 0 \end{bmatrix} + \begin{bmatrix} 0 & Y_{11} & \cdots & Y_{1k} \\ 0 & Y_{21} & \cdots & Y_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & Y_{k1} & \cdots & Y_{kk} \end{bmatrix}, \textit{and}$$

$$X \underset{\downarrow}{\boxplus} Y := \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ X_{k1} & X_{k2} & \cdots & X_{kk} \\ 0 & 0 & \cdots & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & \cdots & 0 \\ Y_{11} & Y_{12} & \cdots & Y_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{k1} & Y_{k2} & \cdots & Y_{kk} \end{bmatrix},$$

where the zero blocks are also of size $m \times n$ *are referred to as the column shifted sum and the row shifted sum of* X *and* Y *respectively.*

The above definition immediately gives the following lemma, the proof of which is obvious.

Lemma 2.1.4. *Let* $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ *be an* $m \times n$ *matrix polynomial of grade* k *and let* $L(\lambda) = \lambda X + Y$ *be an* $km \times kn$ *pencil. Then for* $v, w \in \mathbb{C}^k$,

$$(\lambda X + Y)(\Lambda_k(\lambda) \otimes I_n) = v \otimes P(\lambda) \Leftrightarrow X \boxplus\!\!\!\rightarrow Y = v \otimes \begin{bmatrix} A_k & A_{k-1} & \cdots & A_0 \end{bmatrix}$$

$$(\Lambda_k(\lambda)^T \otimes I_m)(\lambda X + Y) = w^T \otimes P(\lambda) \Leftrightarrow X \underset{\downarrow}{\boxplus} Y = w^T \otimes \begin{bmatrix} A_k^T & A_{k-1}^T & \cdots & A_0^T \end{bmatrix}^T.$$

Thus we have an immediate characterization of the spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ in the next theorem the proof of which is omitted as it follows by arguing exactly as in the proof of [61, Theorem 3.5].

Theorem 2.1.5. *Let* $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ *be an* $m \times n$ *matrix polynomial of grade* k *and* $v, w \in \mathbb{C}^k$. *Then the pencils in* $\mathbb{L}_1(P)$ *with right ansatz vector* v *are* $L(\lambda) = \lambda X + Y$ *such that* $X = \begin{bmatrix} v \otimes A_k & -W \end{bmatrix}$ *and* $Y = \begin{bmatrix} W + v \otimes \begin{bmatrix} A_{k-1} & \cdots & A_1 \end{bmatrix} & v \otimes A_0 \end{bmatrix}$ *with* $W \in \mathbb{C}^{km \times (k-1)n}$ *chosen arbitrarily.*

Similarly, the pencils in $\mathbb{L}_2(P)$ *with left ansatz vector* w *are of the form* $L(\lambda) = \lambda X + Y$ *such that* $X = \begin{bmatrix} w^T \otimes A_k \\ -\hat{W} \end{bmatrix}$ *and* $Y = \begin{bmatrix} \hat{W} + w^T \otimes \begin{bmatrix} A_{k-1}^T & \cdots & A_1^T \end{bmatrix}^T \\ w^T \otimes A_0 \end{bmatrix}$ *with* $\hat{W} \in \mathbb{C}^{(k-1)m \times kn}$ *chosen arbitrarily.*

It is clear from Theorem 2.1.5 that the vector spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ are completely determined by the pairs (v, W) and $(w, \hat{W})$ respectively, where $v, w \in \mathbb{C}^k$, $W \in \mathbb{C}^{km\times(k-1)n}$ and $\hat{W} \in \mathbb{C}^{(k-1)m\times kn}$. Hence the dimensions of the vector spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ over $\mathbb{C}$ are both equal to $k(k-1)mn + k$. The following immediate corollary of Theorem 2.1.5 shows that in particular matrix pencils in $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ with corresponding ansatz vector $\alpha e_1 \in \mathbb{C}^k$ for some non zero scalar α are easy to construct from the coefficient matrices of $P(\lambda)$.

Corollary 2.1.6. *Suppose $L(\lambda) = \lambda X + Y \in \mathbb{L}_1(P)$ with right ansatz vector $v = \alpha e_1$ for $\alpha \neq 0$. Then $X = \left[\begin{array}{c|c} \alpha A_k & X_{12} \\ \hline & -Z \end{array}\right]$ and $Y = \left[\begin{array}{c|c} Y_{11} & \alpha A_0 \\ \hline Z & \end{array}\right]$ for some $Z \in \mathbb{C}^{(k-1)m\times(k-1)n}$ where X_{12}, Y_{11} are $m \times (k-1)n$ matrices that satisfy $X_{12} + Y_{11} = \alpha \begin{bmatrix} A_{k-1} & \cdots & A_1 \end{bmatrix}$.*

Similarly, if $L(\lambda) = \lambda X + Y \in \mathbb{L}_2(P)$ has left ansatz vector $w = \alpha e_1$ for $\alpha \neq 0$, then $X = \left[\begin{array}{c|c} \alpha A_k & \\ \hline \hat{X}_{12} & -\hat{Z} \end{array}\right]$ and $Y = \left[\begin{array}{c|c} \hat{Y}_{11} & \hat{Z} \\ \hline \alpha A_0 & \end{array}\right]$ for some $\hat{Z} \in \mathbb{C}^{(k-1)m\times(k-1)n}$ where $\hat{X}_{12}, \hat{Y}_{11}$ are $(k-1)m \times n$ matrices that satisfy $\hat{X}_{12} + \hat{Y}_{11} = \alpha \begin{bmatrix} A_{k-1} \\ \vdots \\ A_1 \end{bmatrix}$.

Given an $m \times n$ matrix polynomial $P(\lambda)$, it is easy to see that

$$L(\lambda) \in \mathbb{L}_2(P) \Leftrightarrow L(\lambda)^T \in \mathbb{L}_1(P^T). \tag{2.1.3}$$

Therefore, the results in the rest of the chapter for $\mathbb{L}_1(P)$ where $P(\lambda)$ is of size $m \times n$ with $m \geqslant n$, give rise to corresponding results for $\mathbb{L}_2(P)$ when $m \leqslant n$ with appropriate modifications. We provide proofs only for the statements concerning $\mathbb{L}_1(P)$ as the corresponding statements for $\mathbb{L}_2(P)$ follow either by using the correspondence (2.1.3) or by similar independent arguments. The first among these is an analog of [14, Theorem 4.1], that gives a sufficient condition for a pencil in $\mathbb{L}_1(P)$ ($\mathbb{L}_2(P)$) to be a strong g-linearization of $P(\lambda)$ when $m \geqslant n$ ($m \leqslant n$).

Theorem 2.1.7. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial. If $m \geqslant n$ and $L(\lambda) \in \mathbb{L}_1(P)$ with right ansatz vector $v \in \mathbb{C}^k \setminus \{0\}$, then for any nonsingular $M \in \mathbb{C}^{k\times k}$ such that $Mv = \alpha e_1$ for some $\alpha \neq 0$, the pencil $(M \otimes I_m)L(\lambda)$ satisfies*

$$(M \otimes I_m)L(\lambda) = \lambda \left[\begin{array}{c|c} \alpha A_k & X_{12} \\ \hline & -Z \end{array}\right] + \left[\begin{array}{c|c} Y_{11} & \alpha A_0 \\ \hline Z & \end{array}\right] \tag{2.1.4}$$

where $X_{12}, Y_{11} \in \mathbb{C}^{m\times(k-1)n}$ *satisfy* $X_{12} + Y_{11} = \alpha \begin{bmatrix} A_{k-1} & \cdots & A_1 \end{bmatrix}$ *and* Z *is a* $(k-1)m \times (k-1)n$ *matrix. If* Z *is of full rank, i.e.,* $\operatorname{rank} Z = (k-1)n$*, then* $L(\lambda) \in \mathbb{L}_1(P)$ *is a strong g-linearization of* $P(\lambda)$*.*

Proof. We first prove the theorem for the case that $v = \alpha e_1$ for some $\alpha \neq 0$. Then

$$L(\lambda) = \lambda \left[\begin{array}{c|c} \alpha A_k & X_{12} \\ \hline & -Z \end{array}\right] + \left[\begin{array}{c|c} Y_{11} & \alpha A_0 \\ \hline Z & \end{array}\right] =: \lambda X + Y,$$

where $X_{12}, Y_{11} \in \mathbb{C}^{m\times(k-1)n}$ satisfy $X_{12} + Y_{11} = \alpha \begin{bmatrix} A_{k-1} & \cdots & A_1 \end{bmatrix}$ and Z is a $(k-1)m \times (k-1)n$ matrix.
Partitioning Z as $Z = \begin{bmatrix} Z_1 & Z_2 & \cdots & Z_{k-1} \end{bmatrix}$ where $Z_i \in \mathbb{C}^{(k-1)m\times n}$, and setting

$$G(\lambda) = \begin{bmatrix} 1 & 0 & \cdots & \lambda^{k-1} \\ & \ddots & & \vdots \\ & & 1 & \lambda \\ & & & 1 \end{bmatrix} \otimes I_n,$$

we have,

$$\begin{aligned} L(\lambda)G(\lambda) &= \begin{bmatrix} * & * & \cdots & * & * \\ Z_1 & (Z_2 - \lambda Z_1) & \cdots & (Z_{k-1} - \lambda Z_{k-2}) & -\lambda Z_{k-1} \end{bmatrix} G(\lambda) \\ &= \begin{bmatrix} * & * & \cdots & * & \alpha P(\lambda) \\ Z_1 & (Z_2 - \lambda Z_1) & \cdots & (Z_{k-1} - \lambda Z_{k-2}) & 0 \end{bmatrix}. \end{aligned}$$

Now,

$$L(\lambda)G(\lambda) \begin{bmatrix} I_n & \lambda I_n & & & \\ & I_n & & & \\ & & I_n & & \\ & & & \ddots & \\ & & & & I_n \end{bmatrix} \begin{bmatrix} I_n & & & & \\ & I_n & \lambda I_n & & \\ & & I_n & & \\ & & & \ddots & \\ & & & & I_n \end{bmatrix} \cdots \begin{bmatrix} I_n & & & & \\ & \ddots & & & \\ & & I_n & \lambda I_n & \\ & & & I_n & \\ & & & & I_n \end{bmatrix}$$

$$= \begin{bmatrix} * & * & \cdots & * & \alpha P(\lambda) \\ Z_1 & Z_2 & \cdots & (Z_{k-1} - \lambda Z_{k-2}) & 0 \end{bmatrix} \begin{bmatrix} I_n & & & & \\ & I_n & \lambda I_n & & \\ & & I_n & & \\ & & & \ddots & \\ & & & & I_n \end{bmatrix} \cdots \begin{bmatrix} I_n & & & & \\ & \ddots & & & \\ & & I_n & \lambda I_n & \\ & & & I_n & \\ & & & & I_n \end{bmatrix}$$

$$= \begin{bmatrix} * & * & \cdots & * & \alpha P(\lambda) \\ Z_1 & Z_2 & \cdots & Z_{k-1} & 0 \end{bmatrix}$$
$$= \left[\begin{array}{c|c} * & \alpha P(\lambda) \\ \hline Z & \end{array}\right].$$

Therefore there exists a unimodular matrix $F(\lambda)$ such that

$$L(\lambda)F(\lambda) = \left[\begin{array}{c|c} P(\lambda) & W(\lambda) \\ \hline & Z \end{array}\right] \text{ for some } W(\lambda) \in \mathbb{C}[\lambda]^{m\times(k-1)n}. \tag{2.1.5}$$

If Z is of full rank, then $Z^\dagger Z = I_{(k-1)n}$. Therefore,

$$\left[\begin{array}{c|c} I_m & -W(\lambda)Z^\dagger \\ \hline & I_{(k-1)m} \end{array}\right] L(\lambda)F(\lambda) = \begin{bmatrix} P(\lambda) & \\ & Z \end{bmatrix}.$$

As rank $Z = (k-1)n = \operatorname{rank}(I_{k-1} \otimes I_{m,n})$, there exist $E \in \mathbb{C}^{(k-1)m\times(k-1)m}$ and $F \in \mathbb{C}^{(k-1)n\times(k-1)n}$ that are invertible such that $Z = E(I_{k-1} \otimes I_{m,n})F$. Therefore,

$$\left[\begin{array}{c|c} I_m & -W(\lambda)Z^\dagger \\ \hline & I_{(k-1)m} \end{array}\right] L(\lambda)F(\lambda) = \begin{bmatrix} I_m & \\ & E \end{bmatrix} \begin{bmatrix} P(\lambda) & \\ & I_{k-1} \otimes I_{m,n} \end{bmatrix} \begin{bmatrix} I_n & \\ & F \end{bmatrix},$$

and this implies that $L(\lambda)$ is a g-linearization of $P(\lambda)$. To show that $L(\lambda)$ is a strong g-linearization of $P(\lambda)$, notice that

$$\lambda^{k-1}\Lambda_k(1/\lambda) = \begin{bmatrix} 1 & \lambda & \cdots & \lambda^{k-1} \end{bmatrix}^T = R_k\Lambda_k(\lambda), \text{ where } R_k = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix}_{k\times k}.$$

Also $\operatorname{rev}_1 L(\lambda)(R_k\Lambda_k(\lambda)\otimes I_n) = \alpha e_1 \otimes \operatorname{rev}_k P(\lambda)$, as $L(\lambda) \in \mathbb{L}_1(P)$ with corresponding right ansatz vector αe_1. Therefore, $\underbrace{\operatorname{rev}_1 L(\lambda)(R_k \otimes I_n)}_{=:\tilde{L}(\lambda)}(\Lambda_k(\lambda)\otimes I_n) = \alpha e_1 \otimes \operatorname{rev}_k P(\lambda)$, and we have $\tilde{L}(\lambda) = \lambda\tilde{X} + \tilde{Y} \in \mathbb{L}_1(\operatorname{rev}_k P)$, where

$$\tilde{X} = Y(R_k \otimes I_n) = \left[\begin{array}{c|c} \alpha A_0 & \tilde{X}_{12} \\ \hline & -\tilde{Z} \end{array}\right] \text{ and } \tilde{Y} = X(R_k \otimes I_n) = \left[\begin{array}{c|c} \tilde{Y}_{11} & \alpha A_k \\ \hline \tilde{Z} & \end{array}\right]$$

with $\tilde{Z} = -Z(R_{k-1} \otimes I_n)$. Clearly $\tilde{Z}$ is of full rank if Z is of full rank. Hence $\tilde{L}(\lambda)$ is a g-linearization of $\operatorname{rev}_k P(\lambda)$ and consequently $\operatorname{rev}_1 L(\lambda)$ is a g-linearization of $\operatorname{rev}_k P(\lambda)$, this completes the proof for the case that $v = \alpha e_1$.

Now let $L(\lambda) \in \mathbb{L}_1(P)$ with corresponding right ansatz vector $v \in \mathbb{C}^k \setminus \{0\}$. Then $(M \otimes I_m)L(\lambda) \in \mathbb{L}_1(P)$ with right ansatz vector αe_1 and equation (2.1.4) holds in view of Corollary 2.1.6. From (2.1.4) it follows that $L(\lambda)$ is a strong g-linearization of $P(\lambda)$ if and only if

$$\hat{L}(\lambda) := \lambda \left[\begin{array}{c|c} \alpha A_k & X_{12} \\ \hline & -Z \end{array}\right] + \left[\begin{array}{c|c} Y_{11} & \alpha A_0 \\ \hline Z & \end{array}\right] \in \mathbb{L}_1(P)$$

with corresponding right ansatz vector αe_1, is a strong g-linearization of $P(\lambda)$. But we have $\operatorname{rank} Z = (k-1)n$. Therefore by the first part of the proof, it follows that $\hat{L}(\lambda)$ is a strong g-linearization of $P(\lambda)$ and this completes the proof. $\square$

The corresponding theorem for $\mathbb{L}_2(P)$ is as follows.

Theorem 2.1.8. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial. If $m \leqslant n$ and $L(\lambda) \in \mathbb{L}_2(P)$ with left ansatz vector $w \in \mathbb{C}^k \setminus \{0\}$, then for any nonsingular $\hat{M} \in \mathbb{C}^{k \times k}$ such that $\hat{M}w = \alpha e_1$ for some $\alpha \neq 0$, the pencil $L(\lambda)(\hat{M}^T \otimes I_n)$ satisfies*

$$L(\lambda)(\hat{M}^T \otimes I_n) = \lambda \left[\begin{array}{c|c} \alpha A_k & \\ \hline \hat{X}_{12} & -\hat{Z} \end{array}\right] + \left[\begin{array}{c|c} \hat{Y}_{11} & \hat{Z} \\ \hline \alpha A_0 & \end{array}\right] \tag{2.1.6}$$

with $\hat{Z} \in \mathbb{C}^{(k-1)m \times (k-1)n}$. If $\hat{Z}$ is of full rank, i.e., $\operatorname{rank} \hat{Z} = (k-1)m$, then $L(\lambda) \in \mathbb{L}_2(P)$ is a strong g-linearization of $P(\lambda)$.

It was proved in [61, Theorem 4.1 and Theorem 4.3] that if $P(\lambda)$ is a square regular polynomial, then $L(\lambda) \in \mathbb{L}_1(P)$ is a strong linearization of $P(\lambda)$ if and only if the matrix in the position of the block labelled Z in (2.1.4) is nonsingular. However as shown in [14, Example 2], the same is not a necessary condition for $L(\lambda)$ to be a strong linearization of $P(\lambda)$ if it is square but not regular. The following simple modification of that example shows that if $P(\lambda)$ is an $m \times n$ matrix polynomial with $m \geqslant n$, then $L(\lambda) \in \mathbb{L}_1(P)$ with corresponding nonzero right ansatz vector $v \in \mathbb{C}^k$, can be a strong g-linearization of $P(\lambda)$ even if the matrix labelled Z in (2.1.4) is rank deficient.

Example 2.1.9. Let $P(\lambda) = \lambda^2 A_2$ where $A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$. Then

$$L(\lambda) = \lambda \begin{bmatrix} A_2 & -\hat{X} \\ 0 & -Z \end{bmatrix} + \begin{bmatrix} \hat{X} & 0 \\ Z & 0 \end{bmatrix} \text{ where } \hat{X} = \begin{bmatrix} 0 & 0 \\ 0 & -1 \\ 0 & 0 \end{bmatrix}, \text{ and } Z = \begin{bmatrix} -1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

belongs to $\mathbb{L}_1(P)$ with right ansatz vector e_1. Although $\operatorname{rank} Z = 1$, interchanging the second and fifth rows of $L(\lambda)$ gives $C_1^g(\lambda)$ which is a strong g-linearization of $P(\lambda)$.

Since the matrix in the block labelled Z in the reduction (2.1.4) of $L(\lambda) \in \mathbb{L}_1(P)$ plays an important role in determining whether $L(\lambda)$ is a g-linearization of $P(\lambda)$, we refer to it as the Z-matrix of $L(\lambda)$ with respect to the pair (M, α) as it may vary depending on the choice of the nonsingular matrix M satisfying $Mv = \alpha e_1$. Therefore it is important to know whether its rank can change with change in the choice of M. The next theorem shows that this does not happen, i.e., the rank of the Z-matrix in a given $L(\lambda) \in \mathbb{L}_1(P)$ remains invariant of the choice of M. The proof of the theorem is omitted as it follows by arguing exactly as in the proof of [14, Lemma 4.2].

Theorem 2.1.10. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial with $m \geqslant n$ and $L(\lambda) = \lambda X + Y \in \mathbb{L}_1(P)$ with right ansatz vector $v \neq 0$. Suppose that $M_1, M_2 \in \mathbb{C}^{k\times k}$ are two nonsingular matrices such that $M_1 v = \alpha_1 e_1$ and $M_2 v = \alpha_2 e_1$ for some $\alpha_1 \neq 0$, and $\alpha_2 \neq 0$. If $Z_1, Z_2 \in \mathbb{C}^{(k-1)m\times(k-1)n}$ are the matrices in the block labelled Z in (2.1.4) corresponding to the pairs (M_1, α_1) and (M_2, α_2) respectively, then* $\operatorname{rank} Z_1 = \operatorname{rank} Z_2$.

In a similar way it can also be shown that if the $m \times n$ matrix polynomial $P(\lambda)$ satisfies $m \leqslant n$, the rank of the matrix labelled $\hat{Z}$ in the reduction (2.1.6) is independent of the choice of the nonsingular matrix $\hat{M}$. The above result allows us to make the following definition.

Definition 2.1.11. *For any $m \times n$ matrix polynomial $P(\lambda)$ with $m \geqslant n$ ($m \leqslant n$), the Z-rank of $L(\lambda) \in \mathbb{L}_1(P)$ ($L(\lambda) \in \mathbb{L}_2(P)$) is the rank of any matrix appearing in the block labelled Z ($\hat{Z}$) under any reduction of $L(\lambda)$ of the form (2.1.4) ((2.1.6)). If Z ($\hat{Z}$) in (2.1.4) ((2.1.6)) is of full rank, then we say that $L(\lambda) \in \mathbb{L}_1(P)$ ($L(\lambda) \in \mathbb{L}_2(P)$) has full Z-rank.*

The final result of this section shows that for a given $m \times n$ matrix polynomial $P(\lambda)$ with $m \geqslant n$, almost every pencil in $\mathbb{L}_1(P)$ is a g-linearization of $P(\lambda)$. By the term *almost every* we mean *all but a closed, nowhere dense set of measure zero in* $\mathbb{L}_1(P)$. Before going to the theorem let us introduce the term algebraic set.

Definition 2.1.12. *An algebraic subset $\mathcal{A}$ of a vector space V is the set of common zeros of finite number of polynomials in p variables where p is the dimension of V and coefficients are from the underlying field.*

As we know proper algebraic sets are closed, nowhere dense and have measure zero, hence it is sufficient to show that the pencils in $\mathbb{L}_1(P)$ which are not strong g-linearization of $P(\lambda)$ belong to a proper algebraic set in $\mathbb{L}_1(P)$.

Theorem 2.1.13. (Genericity of g-linearizations in $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$) *For any $m \times n$ matrix polynomial $P(\lambda)$ of grade k with $m \geqslant n$, ($m \leqslant n$,) almost every pencil in $\mathbb{L}_1(P)$ ($\mathbb{L}_2(P)$) is a strong g-linearization of $P(\lambda)$.*

Proof. Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial with $m \geqslant n$. The set of pencils in $\mathbb{L}_1(P)$ with right ansatz vector v consists of all $L(\lambda) = \lambda X + Y$ such that

$$X = \begin{bmatrix} v \otimes A_k & -W \end{bmatrix} \text{ and } Y = \begin{bmatrix} W + v \otimes \begin{bmatrix} A_{k-1} & \cdots & A_1 \end{bmatrix} & v \otimes A_0 \end{bmatrix}$$

with $W \in \mathbb{C}^{km \times (k-1)n}$ chosen arbitrarily. For a parametrization of $\mathbb{L}_1(P)$ we define the isomorphism

$$\Gamma : \mathbb{L}_1(P) \to \mathbb{C}^k \times \mathbb{C}^{km \times (k-1)n}$$
$$\lambda X + Y \mapsto (v, W).$$

Suppose $L(\lambda) = \lambda X + Y \in \mathbb{L}_1(P)$ with right ansatz vector $v = \begin{bmatrix} v_1 & \cdots & v_k \end{bmatrix}^T$. Then for

$$M = \left[\begin{array}{c|c} 1 & 0 \\ \hline -v_2 & \\ \vdots & v_1 I_{k-1} \\ -v_k & \end{array}\right],$$

$Mv = v_1 e_1$ and M is nonsingular if $v_1 \neq 0$. Now $(M \otimes I_m)L(\lambda) = \lambda(M \otimes I_m)X + (M \otimes I_m)Y$ where

$$(M \otimes I_m)X = \left[\begin{array}{c|c} Mv \otimes A_k & -(M \otimes I_m)W \end{array}\right] = \left[\begin{array}{c|c} v_1 e_1 \otimes A_k & -(M \otimes I_m)W \end{array}\right] := \left[\begin{array}{c|c} v_1 A_k & * \\ \hline & \tilde{Z} \end{array}\right].$$

Clearly $-\tilde{Z}$ is a Z-matrix of $L(\lambda)$ with respect to (M, v_1). Then

$$\mathcal{P}(v, W) := v_1 \sum \det(\text{ minor of } \tilde{Z} \text{ of order } (k-1)n)$$

is a polynomial in the $k + k(k-1)mn$ entries of v and W. The pair corresponding to $C_1^g(\lambda)$ has $v = e_1$ and $W = \left[\begin{array}{c} 0 \\ \hline -I_{k-1} \otimes I_{m,n} \end{array}\right]$. Hence, $\tilde{Z} = I_{k-1} \otimes I_{m,n}$ and thus

$\mathcal{P}(v, W) \neq 0$ for $C_1^g(\lambda)$. Therefore the zero set of $\mathcal{P}(v, W)$ corresponds to a proper algebraic subset of $\mathbb{L}_1(P)$. Clearly any pair (v, W) such that $\mathcal{P}(v, W) \neq 0$ has $v_1 \neq 0$ and any one of the minors of $\tilde{Z}$ of order $(k-1)n$ has nonzero determinant. So the corresponding $L(\lambda)$ will have full Z-rank and hence is a strong g-linearization of $P(\lambda)$. $\square$

An important difference between linearizations of regular and singular square matrix polynomials $P(\lambda)$ in the space $\mathbb{L}_1(P)$ is that while every linearization of $P(\lambda)$ in $\mathbb{L}_1(P)$ is also a strong linearization of $P(\lambda)$ when $P(\lambda)$ is a regular matrix polynomial [61, Theorem 4.3], the same is not true if $P(\lambda)$ is singular [14, Example 3]. The following example shows that the same also holds for g-linearizations of rectangular matrix polynomials, i.e., there exist rectangular matrix polynomials $P(\lambda)$ with g-linearizations in $\mathbb{L}_1(P)$ that are not strong g-linearizations.

Example 2.1.14. Let $P(\lambda) = \begin{bmatrix} \lambda^2 & \lambda \\ \lambda & 1 \\ 0 & 0 \end{bmatrix}$. Then

$$L(\lambda) = \lambda \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \in \mathbb{L}_1(P)$$

is a g-linearization of $P(\lambda)$ as $E(\lambda)L(\lambda)F(\lambda) = \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{3,2} \end{bmatrix}$ for

$$E(\lambda) = \begin{bmatrix} 0 & 0 & 1 & \lambda & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \text{ and } F(\lambda) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & -\lambda & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \end{bmatrix}.$$

But $L(\lambda)$ is not a strong g-linearization of $P(\lambda)$ as infinity is a eigenvalue of $L(\lambda)$ but not of $P(\lambda)$.

2.2 Recovery of minimal bases and indices in $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$

In this section we show the process of extraction of left and right minimal bases and indices of an $m \times n$ polynomial $P(\lambda)$ from that of a g-linearization in $\mathbb{L}_1(P)$ or $\mathbb{L}_2(P)$. In particular we show that these extractions are possible from g-linearizations of $P(\lambda)$ in $\mathbb{L}_1(P)$ with full Z-rank if $m \geqslant n$ and those of $P(\lambda)$ in $\mathbb{L}_2(P)$ if $m \leqslant n$. It is easy to see that if $m \geqslant n$ and $L(\lambda) \in \mathbb{L}_1(P)$ is of full Z-rank, then

$$\dim\, N_r(L) = \dim\, N_r(P) \text{ and } \dim\, N_l(L) = \dim\, N_l(P) + (k-1)(m-n).$$

On the other hand if $m \leqslant n$, and $L(\lambda) \in \mathbb{L}_2(P)$ has full Z-rank, then the above equalities hold when the positions of the right and left null spaces are interchanged for both $P(\lambda)$ and $L(\lambda)$. Therefore the process of extracting the right (left) minimal bases and indices of $P(\lambda) \in \mathbb{C}[\lambda]^{m\times n}$ from those of a g-linearization of $P(\lambda)$ in $\mathbb{L}_1(P)$ ($\mathbb{L}_2(P)$) is identical to the extraction of the same quantities from a linearization of a square singular polynomial in the respective spaces (as established in [14]). However, showing that the left (right) minimal bases and indices of $P(\lambda)$ can also be extracted from those of $L(\lambda) \in \mathbb{L}_1(P)$ ($L(\lambda) \in \mathbb{L}_2(P)$) with full Z-rank requires more work.

2.2.1 Recovery of right (left) minimal bases and indices in $\mathbb{L}_1(P)$ ($\mathbb{L}_2(P)$)

Given an $m \times n$ matrix polynomial $P(\lambda)$ with $m \geqslant n$, the following lemma provides an isomorphism between $N_r(P)$ and $N_r(L)$ that enables extraction of the right minimal bases and indices of $P(\lambda)$ from those of $L(\lambda) \in \mathbb{L}_1(P)$.

Lemma 2.2.1. *Let $P(\lambda)$ be an $m \times n$ matrix polynomial of grade k with $m \geqslant n$. Also let $L(\lambda) \in \mathbb{L}_1(P)$ with a right ansatz vector $v \neq 0$, and $x(\lambda) \in \mathbb{C}(\lambda)^n$. Then $\Lambda_k(\lambda) \otimes x(\lambda) \in N_r(L)$ if and only if $x(\lambda) \in N_r(P)$. Moreover, if $L(\lambda)$ is a g-linearization of $P(\lambda)$, then the mapping*

$$R_\Lambda : N_r(P) \to N_r(L)$$
$$x(\lambda) \mapsto \Lambda_k(\lambda) \otimes x(\lambda).$$

is a linear isomorphism between the $\mathbb{C}(\lambda)$-vector spaces $N_r(P)$ and $N_r(L)$. Furthermore, $x(\lambda) \in N_r(P)$ is a vector polynomial if and only if $\Lambda_k(\lambda) \otimes x(\lambda) \in N_r(L)$ is a vector polynomial.

We skip the proof as it follows by arguing exactly as in the proof of [14, Lemma 5.1]. Now the following theorem whose proof is immediate shows that the right minimal bases and indices of $P(\lambda) \in \mathbb{C}[\lambda]^{m\times n}$, $m \geqslant n$, have a very simple relationship with those of a g-linearization $L(\lambda) \in \mathbb{L}_1(P)$ and can be easily extracted from the latter.

Theorem 2.2.2. *Let $P(\lambda)$ be an $m \times n$ matrix polynomial of grade k with $m \geqslant n$ and* $\operatorname{nrank} P = r$. *Also let $L(\lambda) \in \mathbb{L}_1(P)$ be a g-linearization of $P(\lambda)$.*

1. *The right minimal indices of $P(\lambda)$ are $\epsilon_1 \leqslant \epsilon_2 \leqslant \cdots \leqslant \epsilon_{n-r}$ if and only if the right minimal indices of $L(\lambda)$ are $(k-1)+\epsilon_1 \leqslant (k-1)+\epsilon_2 \leqslant \cdots \leqslant (k-1)+\epsilon_{n-r}$.*

2. *Any right minimal basis of $L(\lambda)$ is of the form $\{\Lambda_k(\lambda) \otimes x_1(\lambda), \ldots, \Lambda_k(\lambda) \otimes x_{n-r}(\lambda)\}$ where $\{x_1(\lambda), \ldots, x_{n-r}(\lambda)\}$ is a right minimal basis of $P(\lambda)$.*

Similarly, if $P(\lambda)$ is an $m \times n$ matrix polynomial with $m \leqslant n$, then the mapping

$$R_\Lambda : N_l(P) \to N_l(L), \quad y(\lambda) \mapsto \Lambda_k(\lambda) \otimes y(\lambda),$$

is an isomorphism between $N_l(P)$ and $N_l(L)$ that also induces a bijection between vector polynomials in $N_l(P)$ and $N_l(L)$. This results in the following counterpart of Theorem 2.2.2 for extraction of the left minimal bases and indices of $P(\lambda)$ from those of any g-linearization $L(\lambda) \in \mathbb{L}_2(P)$.

Theorem 2.2.3. *Let $P(\lambda)$ be an $m \times n$ matrix polynomial of grade k with $m \leqslant n$ and* $\operatorname{nrank} P = r$. *Also let $L(\lambda) \in \mathbb{L}_2(P)$ be a g-linearization of $P(\lambda)$.*

1. *The left minimal indices of $P(\lambda)$ are $\eta_1 \leqslant \eta_2 \leqslant \cdots \leqslant \eta_{m-r}$ if and only if the left minimal indices of $L(\lambda)$ are $(k-1)+\eta_1 \leqslant (k-1)+\eta_2 \leqslant \cdots \leqslant (k-1)+\eta_{m-r}$.*

2. *Any left minimal basis of $L(\lambda)$ is of the form $\{\Lambda_k(\lambda) \otimes y_1(\lambda), \ldots, \Lambda_k(\lambda) \otimes y_{m-r}(\lambda)\}$ where $\{y_1(\lambda), \ldots, y_{m-r}(\lambda)\}$ is a left minimal basis of $P(\lambda)$.*

2.2.2 Recovery of left (right) minimal bases and indices in $\mathbb{L}_1(P)$ ($\mathbb{L}_2(P)$)

In this section we show that the left minimal bases and indices of $P(\lambda) \in \mathbb{C}[\lambda]^{m\times n}$ with $m \geqslant n$, can be extracted from the g-linearizations in $\mathbb{L}_1(P)$ that are of full Z-rank. The following lemmas will be very useful for establishing Theorem 2.2.6 which is the main result of this section.

Lemma 2.2.4. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial of grade k with $m \geqslant n$. Suppose $L(\lambda) \in \mathbb{L}_1(P)$ has full Z-rank and right ansatz vector $v \neq 0$. Then the mapping*

$$\begin{aligned}\mathcal{L}_v : N_l(L) &\to N_l(P)\\ y(\lambda) &\mapsto (v^T \otimes I_m) y(\lambda)\end{aligned}$$

is a linear map from the vector space $N_l(L)$ onto the vector space $N_l(P)$ over $\mathbb{C}(\lambda)$. Furthermore it is an onto map from the vector polynomials in $N_l(L)$ to the vector polynomials in $N_l(P)$ with the property that if $q(\lambda) \in N_l(P)$ is a vector polynomial of degree δ, then there exists a vector polynomial $y(\lambda) \in N_l(L)$ of degree δ such that $\mathcal{L}_v(y(\lambda)) = q(\lambda)$.

Proof. Let $y(\lambda) \in N_l(L)$. Since $L(\lambda) \in \mathbb{L}_1(P)$,

$$y(\lambda)^T L(\lambda)(\Lambda_k(\lambda) \otimes I_n) = 0 \Rightarrow y(\lambda)^T (v \otimes P(\lambda)) = 0 \Rightarrow y(\lambda)^T (v \otimes I_m) P(\lambda) = 0.$$

Therefore $(v^T \otimes I_m) y(\lambda) \in N_l(P)$ and this shows that $\mathcal{L}_v$ is well defined and clearly linear.

Let $q(\lambda) \in N_l(P)$ and

$$S_k(\lambda) = \begin{bmatrix} 1 & \lambda & \cdots & \lambda^{k-2} \\ & \ddots & \ddots & \\ & & \ddots & \lambda \\ & & & 1 \\ & & & 0 \end{bmatrix}_{k \times (k-1)},$$

$$q_1(\lambda) = \begin{bmatrix} q(\lambda) \\ \tilde{q}(\lambda) \end{bmatrix} \text{ and } y(\lambda) = (M^T \otimes I_m) q_1(\lambda) \tag{2.2.1}$$

where

$$\tilde{q}(\lambda)^T = -q(\lambda)^T \underbrace{\left(\lambda \begin{bmatrix} A_k & X_{12} \end{bmatrix} + \begin{bmatrix} Y_{11} & A_0 \end{bmatrix}\right) (S_k(\lambda) \otimes I_n) Z^{\dagger}}_{=:C}, \tag{2.2.2}$$

M being a nonsingular matrix such that $Mv = e_1$. Clearly $\mathcal{L}_v(y(\lambda)) = q(\lambda)$ and $\mathcal{L}_v$ is onto if $y(\lambda) \in N_l(L)$. This holds if and only if $q_1(\lambda) \in N_l(\hat{L})$, where $\hat{L}(\lambda) =$

$(M \otimes I_m)L(\lambda)$. Also as $L(\lambda) \in \mathbb{L}_1(P)$ corresponds to right ansatz vector v and $Mv = e_1$, therefore $\hat{L}(\lambda) \in \mathbb{L}_1(P)$ corresponds to right ansatz vector e_1. So,

$$\hat{L}(\lambda) = \lambda \left[\begin{array}{c|c} A_k & X_{12} \\ \hline & -Z \end{array}\right] + \left[\begin{array}{c|c} Y_{11} & A_0 \\ \hline Z & \end{array}\right]$$

where Z has full rank and $X_{12}+Y_{11} = \begin{bmatrix} A_{k-1} & A_{k-2} & \cdots & A_1 \end{bmatrix}$. This implies that,

$$X_{12} := \begin{bmatrix} X_1 & X_2 & \cdots & X_{k-1} \end{bmatrix} \text{ and } Y_{11} := \begin{bmatrix} Y_1 & Y_2 & \cdots & Y_{k-1} \end{bmatrix}$$

where $X_i, Y_i \in \mathbb{C}^{m\times n}$ satisfy $X_i + Y_i = A_{k-i}$ for $i = 1, 2, \ldots k-1$. Now from (2.2.1) and (2.2.2),

$$\begin{aligned} q_1(\lambda)^T \hat{L}(\lambda) &= [q(\lambda)^T \quad \tilde{q}(\lambda)^T]\hat{L}(\lambda) \\ &= q(\lambda)^T [I_m \quad -C]\left(\lambda \left[\begin{array}{c|c} A_k & X_{12} \\ \hline & -Z \end{array}\right] + \left[\begin{array}{c|c} Y_{11} & A_0 \\ \hline Z & \end{array}\right]\right) \\ &= q(\lambda)^T \left(\lambda[A_k \quad X_{12} + \hat{C}] + [Y_{11} - \hat{C} \quad A_0]\right) \end{aligned} \tag{2.2.3}$$

where

$$\begin{aligned} \hat{C} &= \left(\lambda \begin{bmatrix} A_k & X_1 & X_2 & \cdots & X_{k-1} \end{bmatrix} + \begin{bmatrix} Y_1 & Y_2 & \cdots & Y_{k-1} & A_0 \end{bmatrix}\right)(S_k \otimes I_n) \\ &= \left(\begin{bmatrix} \lambda A_k + Y_1 & \lambda X_1 + Y_2 & \lambda X_2 + Y_3 & \cdots & \lambda X_{k-2} + Y_{k-1} & \lambda X_{k-1} + A_0 \end{bmatrix}\right)(S_k \otimes I_n) \\ &= [\ \lambda A_k + Y_1 \quad \lambda^2 A_k + \lambda(Y_1 + X_1) + Y_2 \quad \lambda^3 A_k + \lambda^2(Y_1 + X_1) + \lambda(Y_2 + X_2) + Y_3 \\ &\quad \cdots \quad \cdots \quad \lambda^{k-1} A_k + \lambda^{k-2}(Y_1 + X_1) + \cdots + \lambda(Y_{k-2} + X_{k-2}) + Y_{k-1}\ ] \\ &= [\ \lambda A_k \quad \lambda^2 A_k + \lambda A_{k-1} \quad \cdots \quad \lambda^{k-1} A_k + \lambda^{k-2} A_{k-1} + \cdots + \lambda A_2\ ] + Y_{11} \end{aligned} \tag{2.2.4}$$

the 2-nd last equality being due to the fact that $X_i + Y_i = A_{k-i}, i = 1, 2, \ldots, k-1$. Therefore,

$$\begin{aligned} \lambda(X_{12} + \hat{C}) &= \lambda(X_{12} + Y_{11}) + \lambda[\lambda A_k \ \cdots \ \lambda^{k-1} A_k + \lambda^{k-2} A_{k-1} + \cdots + \lambda A_2] \\ &= [\ \lambda^2 A_k + \lambda A_{k-1} \quad \lambda^3 A_k + \lambda^2 A_{k-1} + \lambda A_{k-2} \quad \cdots \quad P(\lambda) - A_0\ ]. \end{aligned} \tag{2.2.5}$$

Using (2.2.5) and (2.2.4) in (2.2.3),

$$\begin{aligned} q_1(\lambda)^T \hat{L}(\lambda) = q(\lambda)^T & \left(\begin{bmatrix} \lambda A_k & \lambda^2 A_k + \lambda A_{k-1} & \lambda^3 A_k + \lambda^2 A_{k-1} + \lambda A_{k-2} & \cdots & P(\lambda) - A_0 \end{bmatrix}\right. \\ & \left. + \begin{bmatrix} -\lambda A_k & -(\lambda^2 A_k + \lambda A_{k-1}) & -(\lambda^3 A_k + \lambda^2 A_{k-1} + \lambda A_{k-2}) & \cdots & A_0 \end{bmatrix}\right) \end{aligned}$$

$$= q(\lambda)^T \begin{bmatrix} 0 & 0 & \cdots & 0 & P(\lambda) \end{bmatrix}$$
$$= 0.$$

Therefore $q_1(\lambda) \in N_l(\hat{L})$ and hence $\mathcal{L}_v$ is an onto linear map from the vector space $N_l(L)$ to the vector space $N_l(P)$ over $\mathbb{C}(\lambda)$. Now clearly, if $y(\lambda) \in N_l(L)$ is a vector polynomial, then so is $\mathcal{L}_v(y(\lambda))$. Conversely, if $q(\lambda) \in N_l(P)$ is a vector polynomial, then from (2.2.1) and (2.2.2) it follows that $\tilde{q}(\lambda)$, $q_1(\lambda)$ and $y(\lambda)$ are also all vector polynomials. Since $\mathcal{L}_v(y(\lambda)) = q(\lambda)$ and $y(\lambda) \in N_l(L)$, it follows that $\mathcal{L}_v$ maps the vector polynomials in $N_l(L)$ onto the vector polynomials in $N_l(P)$. To complete the proof we show that if the degree of $q(\lambda)$ is δ, then $y(\lambda)$ can be chosen so that it has degree δ.

Let $q(\lambda) \in N_l(P)$ and $y(\lambda) \in N_l(L)$ be vector polynomials such that $\mathcal{L}_v(y(\lambda)) = q(\lambda)$. Let $\deg q = \delta$ and suppose $\deg y = \hat{\delta} > \delta$. Let $q_1(\lambda) = (M^{-T} \otimes I_m)y(\lambda)$. Then $\deg q_1 = \hat{\delta}$ and $q(\lambda) = \mathcal{L}_v(y(\lambda)) = (v^T \otimes I_m)(M^T \otimes I_m)q_1(\lambda) = (e_1^T \otimes I_m)q_1(\lambda)$. This implies that $q_1(\lambda) = \begin{bmatrix} q(\lambda) \\ \tilde{q}(\lambda) \end{bmatrix}$ where $\deg \tilde{q} = \hat{\delta}$. Hence

$$q_1(\lambda) = \sum_{i=\delta+1}^{\hat{\delta}} \lambda^i \begin{bmatrix} 0 \\ t_i \end{bmatrix} + \underbrace{\begin{bmatrix} q(\lambda) \\ \sum_{i=0}^{\delta} \lambda^i t_i \end{bmatrix}}_{=:\hat{q}(\lambda)}$$

where $t_i \in \mathbb{C}^{(k-1)m}$, $i = 0, \ldots, \hat{\delta}$, with $t_{\hat{\delta}} \neq 0$. Clearly $\deg \hat{q} = \delta$ and

$$q(\lambda) = \mathcal{L}_v(y(\lambda)) = (e_1^T \otimes I_m)q_1(\lambda) = (e_1^T \otimes I_m)\hat{q}(\lambda) = \mathcal{L}_v(\underbrace{(M^T \otimes I_m)\hat{q}(\lambda)}_{=:\eta(\lambda)}).$$

So we have $\deg \eta = \delta$. Finally, we show that $\eta(\lambda) \in N_l(L)$. As $y(\lambda) \in N_l(L)$, therefore, $q_1(\lambda) \in N_l(\hat{L})$ and

$$q_1(\lambda)^T \hat{L}(\lambda) = 0$$

$$\Rightarrow \left(\begin{bmatrix} 0_m \\ \sum_{i=\delta+1}^{\hat{\delta}} \lambda^i t_i \end{bmatrix}^T + \hat{q}(\lambda)^T \right) \left(\lambda \left[\begin{array}{c|c} A_k & X_{12} \\ \hline & -Z \end{array}\right] + \left[\begin{array}{c|c} Y_{11} & A_0 \\ \hline Z & \end{array}\right] \right) = 0$$

$$\Rightarrow \lambda \begin{bmatrix} 0_m \\ \sum_{i=\delta+1}^{\hat{\delta}} \lambda^i t_i \end{bmatrix}^T \left[\begin{array}{c|c} A_k & X_{12} \\ \hline & -Z \end{array}\right] + \begin{bmatrix} 0_m \\ \sum_{i=\delta+1}^{\hat{\delta}} \lambda^i t_i \end{bmatrix}^T \left[\begin{array}{c|c} Y_{11} & A_0 \\ \hline Z & \end{array}\right] + \hat{q}(\lambda)^T \hat{L}(\lambda) = 0$$

$$\Rightarrow \begin{bmatrix} 0_n^T & -\sum_{i=\delta+1}^{\hat{\delta}} \lambda^{i+1} t_i^T Z \end{bmatrix} + \begin{bmatrix} \sum_{i=\delta+1}^{\hat{\delta}} \lambda^i t_i^T Z & 0_n^T \end{bmatrix} + \hat{q}(\lambda)^T \hat{L}(\lambda) = 0.$$

Therefore,

$$\lambda^{\delta+1}\left[t_{\delta+1}^T Z \quad 0_n^T\right] + \lambda^{\delta+2}\left(\left[t_{\delta+2}^T Z \quad 0_n^T\right] + \left[0_n^T \quad -t_{\delta+1}^T Z\right]\right) + \cdots$$
$$\cdots + \lambda^{\hat{\delta}}\left(\left[t_{\hat{\delta}}^T Z \quad 0_n^T\right] + \left[0_n^T \quad -t_{\hat{\delta}-1}^T Z\right]\right) + \lambda^{\hat{\delta}+1}\left[0_n^T \quad -t_{\hat{\delta}}^T Z\right] + \hat{q}(\lambda)^T \hat{L}(\lambda) = 0. \tag{2.2.6}$$

Since the degree of $\hat{q}(\lambda)^T\hat{L}(\lambda)$ is at most $\delta+1$, equating the coefficients of $\lambda^{\delta+2},\ldots,\lambda^{\hat{\delta}+1}$ in (2.2.6) to 0, we have $t_i^T Z = 0$ for $i = \delta+1,\ldots,\hat{\delta}$. Therefore (2.2.6) implies that $\hat{q}(\lambda)^T\hat{L}(\lambda) = 0$. Now as

$$\eta(\lambda)^T L(\lambda) = \hat{q}(\lambda)^T (M \otimes I_m) L(\lambda) = \hat{q}(\lambda)^T \hat{L}(\lambda) = 0,$$

this completes the proof. □

Lemma 2.2.5. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial of grade k with $m \geqslant n$ and let $L(\lambda) \in \mathbb{L}_1(P)$ corresponding to a nonzero right ansatz vector $v \in \mathbb{C}^k$, be of full Z-rank. Let $r = \operatorname{nrank} P$, $c = (k-1)(m-n)$ and $\mathcal{L}_v(y(\lambda)) = (v^T \otimes I_m)y(\lambda)$ for all $y(\lambda) \in N_l(L)$. Then there exists a minimal basis of $N_l(L)$ of the form*

$$\{y_1(\lambda),\ldots,y_{m-r}(\lambda), u_{m-r+1},\ldots,u_{m-r+c}\},$$

where $\{u_{m-r+1},\ldots,u_{m-r+c}\} \subset \mathbb{C}^{km}$ is a basis of the null space of the linear map $\mathcal{L}_v$ denoted by $N(\mathcal{L}_v)$.

Proof. Under the given hypothesis we have, $\dim(N_l(L)) = \dim(N_l(P)) + (k-1)(m-n)$. Since we have $r = \operatorname{nrank} P$, assume that $\{v_1(\lambda),\ldots,v_{m-r}(\lambda)\}$ is a minimal basis of $N_l(P)$ with $\deg v_j = \delta_j$ for $j = 1,\ldots,m-r$. By Lemma 2.2.4, there exist linearly independent vectors $y_1(\lambda),\ldots,y_{m-r}(\lambda) \in N_l(L)$ such that

$$\mathcal{L}_v(y_j(\lambda)) = v_j(\lambda), \text{ and } \deg y_j = \delta_j.$$

Also from Lemma 2.2.4 we have, $\operatorname{rank}\mathcal{L}_v = \dim N_l(P)$, and hence $\dim(N(\mathcal{L}_v)) = c$. Let $M \in \mathbb{C}^{k\times k}$ be nonsingular such that $Mv = e_1$. Then as $L(\lambda)$ has full Z-rank,

$$\hat{L}(\lambda) := (M \otimes I_m)L(\lambda) = \lambda \left[\begin{array}{c|c} A_k & X_{12} \\ \hline & -Z \end{array}\right] + \left[\begin{array}{c|c} Y_{11} & A_0 \\ \hline Z & \end{array}\right],$$

where X_{12}, Y_{11} and Z are as given in Theorem 2.1.7 such that $\operatorname{rank} Z = (k-1)n$. We show that $N(\mathcal{L}_v)$ has a basis consisting of vectors of the form $(M^T \otimes I_m)\begin{bmatrix} 0 \\ w \end{bmatrix}$

such that $w^T Z = 0$. Since $\operatorname{rank} Z = (k-1)n$, there exist c linearly independent vectors $w_1, \ldots, w_c \in \mathbb{C}^{(k-1)m}$ such that $w_i^T Z = 0$ for all $i = 1, \ldots, c$. So

$$\beta_w := \left\{ (M^T \otimes I_m) \begin{bmatrix} 0 \\ w_1 \end{bmatrix}, \ldots, (M^T \otimes I_m) \begin{bmatrix} 0 \\ w_c \end{bmatrix} \right\}$$

is a linearly independent subset of $N(\mathcal{L}_v)$ as

$$\left((M^T \otimes I_m) \begin{bmatrix} 0 \\ w_i \end{bmatrix} \right)^T L(\lambda) = \begin{bmatrix} 0 \\ w_i \end{bmatrix}^T (M \otimes I_m) L(\lambda) = \begin{bmatrix} 0 \\ w_i \end{bmatrix}^T \hat{L}(\lambda) = w_i^T Z = 0$$

and $(v^T \otimes I_m)(M^T \otimes I_m) \begin{bmatrix} 0 \\ w_i \end{bmatrix} = (e_1^T \otimes I_m) \begin{bmatrix} 0 \\ w_i \end{bmatrix} = 0$

for all $i = 1, \ldots, c$. Clearly β_w is also a basis of $N(\mathcal{L}_v)$ as it has c linearly independent vectors. Let

$$\beta = \{y_1(\lambda), \ldots, y_{m-r}(\lambda), u_{m-r+1}, \ldots, u_{m-r+c}\}$$

where $u_{m-r+j} := (M^T \otimes I_m) \begin{bmatrix} 0 \\ w_j \end{bmatrix}$, for $j = 1, \ldots, c$. Then β has $\dim(N_l(L))$ vectors as $\operatorname{rank} \mathcal{L}_v = \dim(N_l(P)) = m - r$. Therefore β is a basis of $N_l(L)$ if it is a linearly independent set. Suppose there exist $a_1(\lambda), \ldots, a_{m-r+c}(\lambda) \in \mathbb{C}(\lambda)$ such that

$$a_1(\lambda) y_1(\lambda) + \cdots + a_{m-r}(\lambda) y_{m-r}(\lambda) + a_{m-r+1}(\lambda) u_{m-r+1} + \cdots + a_{m-r+c}(\lambda) u_{m-r+c} = 0. \tag{2.2.7}$$

Then,

$$\mathcal{L}_v(a_1(\lambda) y_1(\lambda) + \cdots + a_{m-r}(\lambda) y_{m-r}(\lambda)) = 0 \Rightarrow a_1(\lambda) v_1(\lambda) + \cdots + a_{m-r}(\lambda) v_{m-r}(\lambda) = 0.$$

This gives $a_j(\lambda) = 0$ for $j = 1, \ldots, m-r$ as $\{v_1(\lambda), \ldots, v_{m-r}(\lambda)\}$ is a basis of $N_l(P)$. So by (2.2.7) we have $a_{m-r+1}(\lambda) u_{m-r+1} + \cdots + a_{m-r+c}(\lambda) u_{m-r+c} = 0$. As β_w is a basis of $N(\mathcal{L}_v)$, this implies that $a_j(\lambda) = 0$ for $j = m-r+1, \ldots, m-r+c$. Hence β is a basis of $N_l(L)$. Suppose that it is not a minimal basis of $N_l(L)$. Since the sum of the degrees of the polynomials in β is $\sum_{j=1}^{m-r} \delta_j$, there exists a minimal basis $\hat{\beta} := \{\hat{y}_1(\lambda), \ldots, \hat{y}_{m-r+c}(\lambda)\}$ of $N_l(L)$ such that $\sum_{i=1}^{m-r+c} \deg \hat{y}_i < \sum_{i=1}^{m-r} \delta_i$. Then $\mathcal{L}_v(\hat{\beta})$ is a spanning set in $N_l(P)$. Let

$$\{\mathcal{L}_v(\hat{y}_{i_1}(\lambda)), \ldots, \mathcal{L}_v(\hat{y}_{i_{m-r}}(\lambda))\} \subset \mathcal{L}_v(\hat{\beta})$$

be a basis of $N_l(P)$. Then

$$\sum_{j=1}^{m-r} \deg \mathcal{L}_v(\hat{y}_{i_j}) < \sum_{j=1}^{m-r} \delta_j = \sum_{j=1}^{m-r} \deg v_j.$$

But this contradicts the assumption that $\{v_1(\lambda), \ldots, v_{m-r}(\lambda)\}$ is a minimal basis of $N_l(P)$. Hence the proof. $\square$

The following theorem now shows how the left minimal bases and indices of an $m \times n$ matrix polynomial $P(\lambda)$ with $m \geqslant n$ can be extracted from those of $L(\lambda) \in \mathbb{L}_1(P)$ with full Z-rank.

Theorem 2.2.6. *Let $P(\lambda)$ be an $m \times n$ matrix polynomial of grade k with $m \geqslant n$, and let $L(\lambda) \in \mathbb{L}_1(P)$ corresponding to nonzero right ansatz vector $v \in \mathbb{C}^k$ be of full Z-rank. Let $r = \operatorname{nrank} P$, $c = (k-1)(m-n)$ and $\mathcal{L}_v(y(\lambda)) = (v^T \otimes I_m)y(\lambda)$ for all $y(\lambda) \in N_l(L)$. If $\beta = \{y_1(\lambda), \ldots, y_{m-r}(\lambda), u_{m-r+1}, \ldots, u_{m-r+c}\}$ is a minimal basis of $N_l(L)$ satisfying the properties of Lemma 2.2.5, then $\{\mathcal{L}_v(y_1(\lambda)), \ldots, \mathcal{L}_v(y_{m-r}(\lambda))\}$ is a minimal basis of $N_l(P)$. Moreover if $\eta_1 \geqslant \eta_2 \geqslant \cdots \geqslant \eta_{m-r} \geqslant \underbrace{0 = \cdots = 0}_{c \text{ zeros}}$ are the left minimal indices of $L(\lambda)$ then $\eta_1 \geqslant \eta_2 \geqslant \cdots \geqslant \eta_{m-r}$ are the left minimal indices of $P(\lambda)$.*

Proof. Since $\{u_{m-r+1}, \ldots, u_{m-r+c}\}$ forms a basis of $N(\mathcal{L}_v)$, $\mathcal{L}_v(u_i) = 0$ for $i = m-r+1, \ldots, m-r+c$. Let $z_i(\lambda) = \mathcal{L}_v(y_i(\lambda))$ for $i = 1, \ldots, m-r$. Clearly $\deg y_i = \deg z_i$ for all $i = 1, \ldots, m-r$ and $\{z_1(\lambda), \ldots, z_{m-r}(\lambda)\}$ is a basis of $N_l(P)$. If it is not a minimal basis of $N_l(P)$, then there exists a basis $\{\hat{z}_1(\lambda), \ldots, \hat{z}_{m-r}(\lambda)\}$ of $N_l(P)$ such that $\deg \hat{z}_{j_0} < \deg z_{j_0}$ for some $1 < j_0 < m-r$. Consequently, by Lemma 2.2.4 and Lemma 2.2.5, there exists a basis

$$\hat{\beta} := \{\hat{y}_1(\lambda), \ldots, \hat{y}_{m-r}(\lambda), u_{m-r+1}, \ldots, u_{m-r+c}\}$$

of $N_l(L)$ such that $\mathcal{L}_v(\hat{y}_j(\lambda)) = \hat{z}_j(\lambda)$ and $\deg \hat{y}_j = \deg \hat{z}_j$ for $j = 1, \ldots, m-r$. The sum of the degrees of the vector polynomials in $\hat{\beta}$ are clearly lower than that of the ones in β as, $\deg \hat{z}_{j_0} < \deg z_{j_0}$. But this contradicts the fact that β is a minimal basis of $N_l(L)$. Hence the proof follows. $\square$

Remark 2.2.7. *A minimal basis of $N_l(P)$ may be extracted from a basis of $N_l(L)$ that satisfies the assumptions of Lemma 2.2.5. We outline the steps for constructing such a basis from* any *given minimal basis of $N_l(L)$.*

1. *Let* $\{y_1(\lambda), \ldots, y_t(\lambda), y_{t+1}(\lambda), \ldots, y_{m-r+c}(\lambda)\}$ *be a minimal basis of* $N_l(L)$ *with* $\deg y_i = d_i$ *such that* $d_1 \geqslant d_2 \geqslant \cdots \geqslant d_{m-r+c}$. *Without loss of generality we may assume that there exists* $1 \leqslant t \leqslant m-r$ *such that* $d_i > 0$ *for* $i = 1, \ldots, t$ *and* $d_i = 0$ *for* $i = t+1, \ldots, m-r+c$.

2. *Consider* $u_{m-r+j} = (M^T \otimes I_m)\begin{bmatrix} 0 \\ v_j \end{bmatrix}$, *for* $j = 1, \ldots, c$ *such that* $\{v_1, \ldots, v_c\}$ *forms a basis of left null space of* Z. *Such a basis may be obtained by choosing* $v_1, \ldots, v_c$ *to be the complex conjugate of the last* c *columns of the matrix* Q *of a* QR *decomposition of* Z. *Then* u_{m-r+j} *belongs to* $N_l(L)$ *and* $\mathcal{L}_v(u_{m-r+j}) = 0$ *for* $j = 1, \ldots, c$.

3. *Clearly* $\beta = \{y_1(\lambda), \ldots, y_t(\lambda), u_{m-r+1}, \ldots, u_{m-r+c}\}$ *is linearly independent set. Check if* $y_{t+1}(\lambda)$ *is in* $\mathrm{span}(\beta)$, *and include it in* β *otherwise. Repeat the process for* $i = t+2, \ldots, m-r+c$ *with respect to the updated* β *after each step.*

Remark 2.2.8. *There are situations when left minimal bases of* $L(\lambda) \in \mathbb{L}_1(P)$ *will generically satisfy the assumptions in Lemma 2.2.5. For example if* $n \leqslant m \leqslant (k+1)n$, *then generically,* $\mathcal{A} := [A_0 \cdots A_k]$ *has full row rank and consequently, none of the left minimal indices of* $P(\lambda)$ *are zero. Consequently, there does not exist any vector polynomial of degree zero in a left minimal basis of* $L(\lambda)$ *that does not belong to* $N(\mathcal{L}_v)$. *This implies that any such basis where the elements are arranged in the decreasing order of their respective degrees, must satisfy the assumptions of Lemma 2.2.5.*

The preceding results imply that if $P(\lambda) \in \mathbb{C}[\lambda]^{m\times n}$ with $m \leqslant n$, then the right minimal bases and indices of $P(\lambda)$ can be extracted from those of $L(\lambda) \in \mathbb{L}_2(P)$ of full Z-rank. In particular we have in this case the following counterpart of Lemma 2.2.4 which can either be proved by arguing as in the proof of Lemma 2.2.4 or by using the relation (2.1.3).

Lemma 2.2.9. *Let* $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ *be an* $m \times n$ *matrix polynomial of grade* k *with* $m \leqslant n$. *Suppose* $L(\lambda) \in \mathbb{L}_2(P)$ *has full* Z*-rank and nonzero left ansatz vector* $w \in \mathbb{C}^k$. *Then the mapping*

$$\begin{aligned} \mathcal{L}_w : N_r(L) &\to N_r(P) \\ x(\lambda) &\mapsto (w^T \otimes I_n)x(\lambda) \end{aligned}$$

is a linear map from the vector space $N_r(L)$ onto the vector space $N_r(P)$ over $\mathbb{C}(\lambda)$. Furthermore it is an onto map from the vector polynomials in $N_r(L)$ to the vector polynomials in $N_r(P)$ with the property that if $q(\lambda) \in N_r(P)$ is a vector polynomial of degree δ, then there exists a vector polynomial $x(\lambda) \in N_r(L)$ of degree δ such that $\mathcal{L}_w(x(\lambda)) = q(\lambda)$.

Therefore, by arguing as in the proof of Lemma 2.2.5, there exists a minimal basis of $N_r(L)$ of the form

$$\{x_1(\lambda), \ldots, x_{n-r}(\lambda), u_{n-r+1}, \ldots, u_{n-r+c}\},$$

where $c = (n-m)(k-1)$ and $\{u_{n-r+1}, \ldots, u_{n-r+c}\} \subset \mathbb{C}^{kn}$ is a basis of $N(\mathcal{L}_w)$. This leads to the following theorem for extracting the right minimal indices and bases of $P(\lambda)$ from those of $L(\lambda) \in \mathbb{L}_2(P)$ with full Z-rank.

Theorem 2.2.10. *Let $P(\lambda)$ be an $m \times n$ matrix polynomial of grade k with $m \leqslant n$, and $L(\lambda) \in \mathbb{L}_2(P)$ with corresponding nonzero left ansatz vector $w \in \mathbb{C}^k$ be of full Z-rank. Let $r = \operatorname{nrank} P$, $c = (k-1)(n-m)$ and $\mathcal{L}_w$ as defined in Lemma 2.2.9. If $\{x_1(\lambda), \ldots, x_{n-r}(\lambda), u_{n-r+1}, \ldots, u_{n-r+c}\}$ is a minimal basis of $N_r(L)$ where $\{u_{n-r+1}, \ldots, u_{n-r+c}\} \subset \mathbb{C}^{kn}$ is a basis of $N(\mathcal{L}_w)$, then $\{\mathcal{L}_w(x_1(\lambda)), \ldots, \mathcal{L}_w(x_{n-r}(\lambda))\}$ is a minimal basis of $N_r(P)$. Moreover if $\epsilon_1 \geqslant \epsilon_2 \geqslant \cdots \geqslant \epsilon_{n-r} \geqslant \underbrace{0 = \cdots = 0}_{c \text{ zeros}}$ are the right minimal indices of $L(\lambda)$ then $\epsilon_1 \geqslant \epsilon_2 \geqslant \cdots \geqslant \epsilon_{n-r}$ are the right minimal indices of $P(\lambda)$.*

Remark 2.2.11. *Given any $m \times n$ matrix polynomial $P(\lambda)$ of grade k with $m \geqslant n$, a pencil $L(\lambda)$ belongs to $\mathbb{L}_1(P)$ with right ansatz vector $v \in \mathbb{C}^k \setminus \{0\}$ if and only if $L(\lambda)$ is of the form $L(\lambda) = \begin{bmatrix} v \otimes I_m & \Omega \end{bmatrix} C_1(\lambda)$ where $\Omega \in \mathbb{C}^{km \times (k-1)n}$ such that*

$$\begin{bmatrix} v \otimes I_m & \Omega \end{bmatrix} = (M^{-1} \otimes I_m) \left[\begin{array}{c|c} \alpha I_m & X_{12} \\ \hline & -Z \end{array}\right].$$

Here $M \in \mathbb{C}^{k \times k}$, such that $Mv = \alpha e_1$, and X_{12} and Z are the same as those in (2.1.4). If the pencil $L(\lambda)$ has full Z-rank then it is a strong g-linearization of the matrix polynomial $P(\lambda)$ and consequently the matrix $\begin{bmatrix} v \otimes I_m & \Omega \end{bmatrix}$ is also of full rank. Though we have already established the recovery rules for minimal bases and minimal indices of $P(\lambda)$ from those of $L(\lambda)$ with full Z-rank, this observation gives us one more way to see that the right minimal indices of $P(\lambda)$ are $(k-1)$ less than those of $L(\lambda)$ and the left minimal indices of $P(\lambda)$ are the same as those of $L(\lambda)$

after eliminating $(k-1)(m-n)$ many zeros. The above conclusion comes from the fact that the right minimal indices of $P(\lambda)$ are $(k-1)$ less than those of $C_1(\lambda)$ and the left minimal indices of $P(\lambda)$ are the same as those of $C_1(\lambda)$ and the presence of the matrix $\begin{bmatrix} v \otimes I_m & \Omega \end{bmatrix}$ contributes $(k-1)(m-n)$ additional degree-zero-vectors to any minimal basis of $N_l(L(\lambda))$.

Similarly, if $m \leqslant n$, then any pencil $L(\lambda)$ belongs to $\mathbb{L}_2(P)$ with nonzero left ansatz vector w, if and only if $L(\lambda)$ is of the form $L(\lambda) = C_2(\lambda)\begin{bmatrix} w^T \otimes I_n \\ \Omega \end{bmatrix}$ for some $(k-1)m \times kn$ matrix Ω. Therefore, by arguing as above, it follows from this relation that if $L(\lambda)$ has full Z-rank, then the left minimal indices of $P(\lambda)$ are $(k-1)$ less than those of $L(\lambda)$ and the right minimal indices of $P(\lambda)$ are the same as those of $L(\lambda)$ after eliminating $(k-1)(n-m)$ many zeros.

2.3 Linearizations arising from g-linearizations

Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial of grade k. In this section we show that although the pencils in $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ are generically g-linearizations of $P(\lambda)$, they can give rise to smaller pencils that are linearizations of $P(\lambda)$ from which the left and right minimal bases and indices of $P(\lambda)$ may be easily extracted. In the following we first describe the process of extracting these smaller pencils from g-linearizations of $P(\lambda)$ of full Z-rank in $\mathbb{L}_1(P)$ when $m \geqslant n$.

Let $L(\lambda) \in \mathbb{L}_1(P)$ with nonzero right ansatz vector $v \in \mathbb{C}^k$ be of full Z-rank. Let $M \in \mathbb{C}^{k\times k}$ be a nonsingular matrix such that $Mv = \alpha e_1$. From (2.1.5),

$$(M \otimes I_m)L(\lambda) = \left[\begin{array}{c|c} P(\lambda) & W(\lambda) \\ \hline & Z \end{array}\right] (F(\lambda))^{-1}.$$

Let $Z = Q\begin{bmatrix} \tilde{R} \\ 0 \end{bmatrix}$ be a QR decomposition of Z where Q is a $(k-1)m \times (k-1)m$ unitary matrix and $\tilde{R}$ is a $(k-1)n \times (k-1)n$ nonsingular and upper triangular matrix. Then we have,

$$\begin{bmatrix} I_m & \\ & Q^* \end{bmatrix} (M \otimes I_m)L(\lambda) = \left[\begin{array}{c|c} P(\lambda) & W(\lambda) \\ \hline & \tilde{R} \\ & 0 \end{array}\right] (F(\lambda))^{-1}. \qquad (2.3.1)$$

Let $Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}$, be a partition of Q such that $Z = Q_1\tilde{R}$ is the condensed QR decomposition of Z. Then recalling that $c = (k-1)(m-n)$, the last c rows of the matrix $\begin{bmatrix} I_m & \\ & Q^* \end{bmatrix} (M \otimes I_m)L(\lambda)$ form the matrix $\begin{bmatrix} 0 & Q_2^* \end{bmatrix} (M \otimes I_m)L(\lambda)$. But $\begin{bmatrix} 0 & Q_2^* \end{bmatrix} (M \otimes I_m)L(\lambda) = 0$ as the last c rows of the matrix on the RHS of (2.3.1) are zero. Consider $D \in \mathbb{C}^{m+(k-1)n \times km}$ such that

$$\begin{bmatrix} D \\ \begin{bmatrix} 0 & Q_2^* \end{bmatrix} (M \otimes I_m) \end{bmatrix} \tag{2.3.2}$$

is nonsingular. Then $\begin{bmatrix} D \\ \begin{bmatrix} 0 & Q_2^* \end{bmatrix} (M \otimes I_m) \end{bmatrix} L(\lambda) = \begin{bmatrix} DL(\lambda) \\ 0 \end{bmatrix}$. We set,

$$L_t(\lambda) = DL(\lambda). \tag{2.3.3}$$

Given a g-linearization $L(\lambda) \in \mathbb{L}_1(P)$ of full Z-rank, the above process of extracting the pencil $L_t(\lambda)$ from $L(\lambda)$ clearly depends not only on $L(\lambda)$ but also on the choice of the nonsingular matrix $M \in \mathbb{C}^{k\times k}$ satisfying $Mv = \alpha e_1$ and the matrix $D \in \mathbb{C}^{m+(k-1)n\times km}$ such that the matrix in (2.3.2) is nonsingular. For ease of expression, we will refer to $L_t(\lambda)$ as *the trimmed version of* $L(\lambda) \in \mathbb{L}_1(P)$, *with respect to* M *and* D, the sizes of the matrices M and D being evident from the context.

Clearly, for a given choice of nonsingular $M \in \mathbb{C}^{k\times k}$, there are infinitely many choices of D in (2.3.3). One possible choice is to set D to be the first $m + (k-1)n$ rows of the matrix $\begin{bmatrix} I_m & \\ & Q^* \end{bmatrix} (M \otimes I_m)$. Then the corresponding linearization is

$$\hat{L}_t(\lambda) := \lambda \left[\begin{array}{c|c} \alpha A_k & X_{12} \\ \hline & -\tilde{R} \end{array}\right] + \left[\begin{array}{c|c} Y_{11} & \alpha A_0 \\ \hline \tilde{R} & \end{array}\right]. \tag{2.3.4}$$

If $L(\lambda) = C_1^g(\lambda)$, then such a choice of D results in $L_t(\lambda)$ being the first Frobenius companion linearization $C_1(\lambda)$. Every other linearization $L_t(\lambda)$ that is not of the form (2.3.4) is strictly equivalent to some $\hat{L}_t(\lambda)$ as

$$\begin{aligned} L_t(\lambda) &= DL(\lambda) \\ &= D(M^{-1} \otimes I_m) \left\{ \lambda \left[\begin{array}{c|c} \alpha A_k & X_{12} \\ \hline & -Z \end{array}\right] + \left[\begin{array}{c|c} Y_{11} & \alpha A_0 \\ \hline Z & \end{array}\right] \right\} \end{aligned}$$

$$
= D(M^{-1}\otimes I_m)\begin{bmatrix} I_m & \\ & Q \end{bmatrix}\left\{\lambda\left[\begin{array}{c|c} \alpha A_k & X_{12} \\ \hline & -\tilde{R} \\ & 0 \end{array}\right] + \left[\begin{array}{c|c} Y_{11} & \alpha A_0 \\ \hline \tilde{R} & \\ 0 & \end{array}\right]\right\}
$$

$$
= \underbrace{D(M^{-1}\otimes I_m)\begin{bmatrix} I_m & \\ & Q_1 \end{bmatrix}}_{=E_1}\left\{\lambda\left[\begin{array}{c|c} \alpha A_k & X_{12} \\ \hline & -\tilde{R} \end{array}\right] + \left[\begin{array}{c|c} Y_{11} & \alpha A_0 \\ \hline \tilde{R} & \end{array}\right]\right\}
$$

Therefore,

$$
L_t(\lambda) = \underbrace{DE_1}_{=:\tilde{D}}\hat{L}_t(\lambda). \tag{2.3.5}
$$

where clearly, $\tilde{D} \in \mathbb{C}^{m+(k-1)n\times m+(k-1)n}$ is nonsingular as it satisfies

$$
\left[\begin{array}{c|c} \tilde{D} & * \\ \hline & I_{(k-1)(m-n)} \end{array}\right] = \begin{bmatrix} D \\ \begin{bmatrix} 0 & Q_2^* \end{bmatrix}(M\otimes I_m) \end{bmatrix}(M^{-1}\otimes I_m)\begin{bmatrix} I_m & \\ & Q \end{bmatrix}.
$$

Remark 2.3.1. *The QR decomposition of the Z-matrix of the g-linearization $L(\lambda)$ that has been used to extract the pencils $L_t(\lambda)$ from $L(\lambda)$ can easily be replaced by any other decomposition like the rank revealing QR decomposition or the SVD of Z without affecting the results and the analysis concerning these pencils. Therefore, the upper triangular structure of the block $\tilde{R}$ in (2.3.4) is not essential.*

2.3.1 Trimming a g-linearization results in a strong linearization

We now show that trimming a g-linearization of full Z-rank results in a strong linearization of $P(\lambda)$ from which the left and right minimal bases and indices of $P(\lambda)$ can easily be recovered. In doing so, we establish the connection between the resulting pencils with some of the important classes of linearizations for rectangular matrix polynomials that have been recently introduced in the literature. We begin with the block minimal bases pencils introduced in [22].

Definition 2.3.2. *A block minimal bases pencil is a pencil of the form*

$$
\left[\begin{array}{c|c} A(\lambda) & \hat{B}(\lambda)^T \\ \hline B(\lambda) & \end{array}\right],\ \left[\begin{array}{c} A(\lambda) \\ \hline B(\lambda) \end{array}\right]\ \text{or}\ \left[\begin{array}{c|c} A(\lambda) & \hat{B}(\lambda)^T \end{array}\right]
$$

where $A(\lambda)$, $B(\lambda)$ and $\hat{B}(\lambda)$ are themselves pencils such that the rows of $B(\lambda)$ and $\hat{B}(\lambda)$ form minimal bases of the rational subspaces spanned by them.

We will need a few important concepts and results related to block minimal bases pencils from [22]. For convenience, following [22], we refer to a matrix polynomial whose rows form a minimal basis of the rational subspace spanned by them as a minimal basis. Such a minimal basis can be associated with a dual minimal basis defined as follows.

Definition 2.3.3. *Two minimal bases $B(\lambda) \in \mathbb{C}[\lambda]^{n_1 \times n}$ and $C(\lambda) \in \mathbb{C}[\lambda]^{n_2 \times n}$ are called (a pair of) dual minimal bases if $n_1 + n_2 = n$ and $B(\lambda)C(\lambda)^T = 0$ holds.*

For example, the matrix polynomials

$$H_j(\lambda) = \begin{bmatrix} -1 & \lambda & & & \\ & -1 & \lambda & & \\ & & \ddots & \ddots & \\ & & & -1 & \lambda \end{bmatrix}_{j \times (j+1)} . \tag{2.3.6}$$

and $\Lambda_{j+1}(\lambda)^T$ given by (2.0.3) are dual minimal bases. For most practical purposes, we will need the following special kind of block minimal bases from [22].

Definition 2.3.4. *A block minimal bases pencil*

$$L(\lambda) = \left[\begin{array}{c|c} A(\lambda) & \hat{B}(\lambda)^T \\ \hline B(\lambda) & \end{array}\right] \tag{2.3.7}$$

is called a strong block minimal bases pencil if it has the following additional properties:

(a) The row degrees of $B(\lambda)$ and $\hat{B}(\lambda)$ are all equal to one.

(b) The row degrees of any minimal basis dual to $B(\lambda)$ are all equal.

(c) The row degrees of any minimal basis dual to $\hat{B}(\lambda)$ are all equal.

We will adopt the convention that if the block $B(\lambda)$ ($\hat{B}(\lambda)$) is absent, then the corresponding dual minimal basis is an identity matrix of the same size as the number of columns (rows) of $A(\lambda)$. The following theorem about block minimal bases pencils which is a combination of [22, Theorems 3.3 and 3.6] will be important for the results in this section and the next one.

Theorem 2.3.5. *Let $L(\lambda)$ be a block minimal bases pencil given by (2.3.7) and $C(\lambda)$ and $\hat{C}(\lambda)$ be the dual minimal bases of $B(\lambda)$ and $\hat{B}(\lambda)$ respectively. Then $L(\lambda)$ is a linearization of the matrix polynomial*

$$Q(\lambda) = \hat{C}(\lambda)A(\lambda)C(\lambda)^T. \tag{2.3.8}$$

Moreover, if $L(\lambda)$ is a strong block minimal bases pencil, then the following hold.

(a) If $Q(\lambda)$ considered as a polynomial of grade $1 + \deg C + \deg \hat{C}$, then $L(\lambda)$ is a strong linearization of $Q(\lambda)$.

(b) If $0 \leqslant \epsilon_1 \leqslant \epsilon_2 \leqslant \cdots \leqslant \epsilon_p$ are the right minimal indices of $Q(\lambda)$, then

$$\epsilon_1 + \deg\, C \leqslant \epsilon_2 + \deg\, C \leqslant \cdots \leqslant \epsilon_p + \deg\, C,$$

are the right minimal indices of $L(\lambda)$.

(c) If $0 \leqslant \eta_1 \leqslant \eta_2 \leqslant \cdots \leqslant \eta_q$ are the left minimal indices of $Q(\lambda)$, then

$$\eta_1 + \deg\, \hat{C} \leqslant \eta_2 + \deg\, \hat{C} \leqslant \cdots \leqslant \eta_q + \deg\, \hat{C}$$

are the left minimal indices of $L(\lambda)$.

Given a strong block minimal bases pencil, the above result shows the construction of a polynomial from the pencil such that the pencil is a strong linearization of the polynomial and lays out the recovery rules for extracting left and right minimal indices of the polynomial from those of the pencil. However in practice, we are generally more interested in the reverse process, i.e., given a matrix polynomial $Q(\lambda)$ of grade k, we are interested in constructing a strong linearization from which the left and right minimal indices of the polynomial can be easily extracted. It was shown in [22] that this can be easily achieved by the so called block Kronecker pencils that are a special class of strong block minimal bases pencils. For these we have

$$\hat{B}(\lambda) = H_\epsilon(\lambda) \otimes I_m \text{ and } B(\lambda) = H_\eta(\lambda) \otimes I_n$$

with $\epsilon + \eta + 1 = k$, and $H_j(\lambda)$ given by (2.3.6). The conditions on the block $A(\lambda)$ under which the block Kronecker pencils become strong linearizations of a given polynomial $P(\lambda)$ are given in [22, Theorem 4.4]. Now we have the main result of this section.

Theorem 2.3.6. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial of grade k with $m \geqslant n$ and $r = \operatorname{nrank} P$. Let $L(\lambda) \in \mathbb{L}_1(P)$ with right ansatz vector $v \in \mathbb{C}^k \setminus \{0\}$ be of full Z-rank. Let $L_t(\lambda)$ be the pencil obtained by trimming $L(\lambda) \in \mathbb{L}_1(P)$, with respect to M and D. Then $L_t(\lambda)$ is a strong linearization of $P(\lambda)$ such that the following hold.*

(a) Every minimal basis of $N_r(L_t)$ is of the form $\{\Lambda_k(\lambda) \otimes x_1(\lambda), \ldots, \Lambda_k(\lambda) \otimes x_{n-r}(\lambda)\}$ where $\{x_1(\lambda), \ldots, x_{n-r}(\lambda)\}$ is a minimal basis of $N_r(P)$.

(b) The right minimal indices of $L_t(\lambda)$ are those of $P(\lambda)$ shifted by $k-1$.

(c) Every minimal basis of $N_l(P)$ is of the form $\{\mathcal{L}_v(D^T y_1(\lambda)), \ldots, \mathcal{L}_v(D^T y_{m-r}(\lambda))\}$ where $\{y_1(\lambda), \ldots, y_{m-r}(\lambda)\}$ is a minimal basis of $N_l(L_t)$ and $\mathcal{L}_v$ is as defined in Lemma 2.2.4.

(d) The left minimal indices of $P(\lambda)$ are equal to those of $L_t(\lambda)$.

Proof. From (2.3.4) and (2.3.5),

$$L_t(\lambda) := \tilde{D} \left[\begin{array}{c|c} I_m & \\ \hline & -\tilde{R} \end{array}\right] \underbrace{\left(\lambda \left[\begin{array}{c|c} \alpha A_k & X_{12} \\ \hline & I_{(k-1)n} \end{array}\right] + \left[\begin{array}{c|c} Y_{11} & \alpha A_0 \\ \hline -I_{(k-1)n} & \end{array}\right] \right)}_{=:K(\lambda)} \tag{2.3.9}$$

Clearly $K(\lambda)$ is in the block Kronecker form $\left[\begin{array}{c} A(\lambda) \\ \hline B(\lambda) \end{array}\right]$ with

$$A(\lambda) := \lambda \left[\begin{array}{cc} \alpha A_k & X_{12} \end{array}\right] + \left[\begin{array}{cc} Y_{11} & \alpha A_0 \end{array}\right] \text{ and } B(\lambda) := H_{k-1}(\lambda) \otimes I_n.$$

Now $A(\lambda)C(\lambda)^T = P(\lambda)$, where $C(\lambda) = \frac{1}{\alpha}(\Lambda_k(\lambda) \otimes I_n)^T$ is a dual of $B(\lambda)$. As the block $\hat{B}(\lambda)$ is absent in $K(\lambda)$ (and consequently, $\hat{C}(\lambda) = I_m$,) by Theorem 2.3.5, $K(\lambda)$ is a strong linearization of $P(\lambda)$ such that $P(\lambda)$ and $K(\lambda)$ have the same left minimal indices and the right minimal indices of $K(\lambda)$ are those of $P(\lambda)$ shifted by $k-1$. The relation (2.3.9), shows that the same is true of each pencil $L_t(\lambda)$ obtained by trimming a strong g-linearization in $\mathbb{L}_1(P)$ and this proves (b) and (d).

The process of obtaining $L_t(\lambda)$ from $L(\lambda)$ implies that $N_r(L) = N_r(L_t)$. Therefore the proof of (a) follows from Theorem 2.2.2. To prove (c) we consider the map

$$\mathcal{L} : N_l(L_t) \to N_l(P)$$
$$y(\lambda) \mapsto \mathcal{L}_v(D^T y(\lambda))$$

where $\mathcal{L}_v$ is as defined in Lemma 2.2.4. If $y(\lambda) \in N_l(L_t)$ then $D^T y(\lambda) \in N_l(L)$ and hence $(v^T \otimes I_m)D^T y(\lambda) = \mathcal{L}_v(D^T y(\lambda)) \in N_l(P)$. Therefore $\mathcal{L}$ is well defined. Also from the definition of $\mathcal{L}$ it is clear that it is a linear map from $N_l(L_t)$ to $N_l(P)$. We first show that $\mathcal{L}$ is bijective. Let Z be the Z-matrix of $(M \otimes I_n)L(\lambda)$ and Q_2 be the last $c = (k-1)(m-n)$ columns of the unitary matrix Q of a QR decomposition of Z. Now $y(\lambda) \in N(\mathcal{L})$ if and only if $D^T y(\lambda) \in N(\mathcal{L}_v)$. As noted in Remark 2.2.7, $\left\{ (M^T \otimes I_m) \begin{bmatrix} 0 \\ q_1 \end{bmatrix}, \ldots, (M^T \otimes I_m) \begin{bmatrix} 0 \\ q_c \end{bmatrix} \right\}$ where $c = (k-1)(m-n)$ and $\{\bar{q}_1, \ldots, \bar{q}_c\}$ are the columns of Q_2 is a basis of $N(\mathcal{L}_v)$. Therefore there exists a nonzero $a \in \mathbb{C}^c$ such that

$$D^T y(\lambda) = (M^T \otimes I_m) \begin{bmatrix} I_m & \\ & \overline{Q}_2 \end{bmatrix} \begin{bmatrix} 0 \\ a \end{bmatrix} = (M^T \otimes I_m) \begin{bmatrix} 0 \\ \overline{Q}_2 a \end{bmatrix}.$$

This implies that $\begin{bmatrix} y(\lambda) \\ -a \end{bmatrix}^T \begin{bmatrix} D \\ \begin{bmatrix} 0 & Q_2^* \end{bmatrix} (M \otimes I_m) \end{bmatrix} = 0$, and as $\begin{bmatrix} D \\ \begin{bmatrix} 0 & Q_2^* \end{bmatrix} (M \otimes I_m) \end{bmatrix}$ is nonsingular, we have $y(\lambda) = 0$. Therefore $\mathcal{L}$ is a one to one linear map. Since $N_l(L_t)$ and $N_l(P)$ are of the same dimension, it follows that $\mathcal{L}$ is a bijective linear map.

Now we will show for any vector polynomial $p(\lambda) \in N_l(P)$ of degree δ we can find a polynomial vector $z(\lambda) \in N_l(L_t)$ of degree δ such that $\mathcal{L}(z(\lambda)) = p(\lambda)$. By Lemma 2.2.4 there exists $\hat{z}(\lambda) \in N_l(L)$ such that $\mathcal{L}_v(\hat{z}(\lambda)) = p(\lambda)$ and $\deg \hat{z} = \deg p$. Set $z(\lambda)$ to be the vector formed by the first $m + (k-1)n$ entries of the vector polynomial $\hat{y}(\lambda)$ which satisfies $\begin{bmatrix} D \\ \begin{bmatrix} 0 & Q_2^* \end{bmatrix} (M \otimes I_m) \end{bmatrix}^T \hat{y}(\lambda) = \hat{z}(\lambda)$. Then $z(\lambda) \in N_l(L_t)$ as,

$$z(\lambda)^T L_t(\lambda) = \hat{y}(\lambda)^T \begin{bmatrix} L_t(\lambda) \\ 0_c \end{bmatrix} = \hat{y}(\lambda)^T \begin{bmatrix} D \\ \begin{bmatrix} 0 & Q_2^* \end{bmatrix} (M \otimes I_m) \end{bmatrix} L(\lambda) = \hat{z}(\lambda)^T L(\lambda) = 0.$$

Now as $(v^T \otimes I_m)(M^T \otimes I_m) \begin{bmatrix} 0_{m \times c} \\ \overline{Q}_2 \end{bmatrix} = (\alpha e_1^T \otimes I_m) \begin{bmatrix} 0_{m \times c} \\ \overline{Q}_2 \end{bmatrix} = 0$, we have

$$\mathcal{L}(z(\lambda)) = (v^T \otimes I_m)D^T z(\lambda) = (v^T \otimes I_m) \begin{bmatrix} D \\ \begin{bmatrix} 0 & Q_2^* \end{bmatrix} (M \otimes I_m) \end{bmatrix}^T \hat{y}(\lambda) = \mathcal{L}_v(\hat{z}(\lambda)) = p(\lambda).$$

Also it is clear that $z(\lambda)$ and $p(\lambda)$ have the same degree as

$$\deg\, z \geqslant \deg\, p = \deg\, \hat{z} = \deg\, \hat{y} \geqslant \deg\, z.$$

Now the proof of part (c) follows by arguing as in the proof of Theorem 2.2.6. $\square$

Remark 2.3.7. *It can be proved that the pencils $L_t(\lambda)$ are strong linearizations of $P(\lambda)$ without establishing their connection with block Kronecker pencils. However, we prefer to give this connection to highlight their position in the current literature of strong linearizations of rectangular matrix polynomials. In fact it is also clear from (2.3.9) that among the linearizations formed by trimming g-linearizations in $\mathbb{L}_1(P)$, only the pencils $\hat{L}_t(\lambda)$ of the form (2.3.4) belong to the block Kronecker ansatz spaces $\mathbb{G}_1(P)$ introduced in [26].*

Remark 2.3.8. *It may be possible to derive recovery rules for the left and right minimal bases of $P(\lambda)$ from those of $L_t(\lambda)$ using the results given in [22]. However, as the pencils $K(\lambda)$ in (2.3.9) to which the pencils $L_t(\lambda)$ are strictly equivalent are special types of block Kronecker pencils, we prefer to prove these parts directly by using the notions and techniques previously introduced in this chapter.*

In a similar way, if $m \leqslant n$, pencils in $\mathbb{L}_2(P)$ of full Z-rank can provide strong linearizations of $P(\lambda)$. In particular if $L(\lambda) \in \mathbb{L}_2(P)$ with nonzero left ansatz vector $w \in \mathbb{C}^k$ has full Z-rank, then for any nonsingular matrix $\hat{M} \in \mathbb{C}^{k\times k}$ such that $\hat{M}w = \alpha e_1$ for some $\alpha \neq 0$,

$$L(\lambda)(\hat{M}^T \otimes I_n) = \lambda \left[\begin{array}{c|c} \alpha A_k & \\ \hline \hat{X}_{12} & -\hat{Z} \end{array}\right] + \left[\begin{array}{c|c} \hat{Y}_{11} & \hat{Z} \\ \hline \alpha A_0 & \end{array}\right],$$

where $\hat{Z} \in \mathbb{C}^{(k-1)m\times(k-1)n}$ with rank $\hat{Z} = (k-1)m$. If $\hat{Z}^* = Q\begin{bmatrix} \hat{R} \\ 0 \end{bmatrix}$ is a QR decomposition of $\hat{Z}^*$ and Q_2 is the matrix formed by the last $c = (k-1)(n-m)$ columns of Q, then it is easy to see that

$$L(\lambda)(\hat{M}^T \otimes I_n)\begin{bmatrix} 0 \\ Q_2 \end{bmatrix} = 0.$$

For any choice of $\hat{D} \in \mathbb{C}^{kn\times(n+(k-1)m)}$, such that the $kn \times kn$ matrix

$$\left[\begin{array}{cc} \hat{D} & (\hat{M}^T \otimes I_n)\begin{bmatrix} 0 & Q_2^T \end{bmatrix}^T \end{array}\right] \tag{2.3.10}$$

is nonsingular, we get the pencils $L(\lambda)\hat{D}$. We refer to them as *the pencils formed by trimming* $L(\lambda) \in \mathbb{L}_2(P)$ *with respect to* $\hat{M}$ *and* $\hat{D}$. For instance, the second companion linearization $C_2(\lambda)$ arises from $C_2^g(\lambda) \in \mathbb{L}_2(P)$, with respect to $\hat{M} = I_k$ and $\hat{D} = \begin{bmatrix} I_n & \\ & I_{k-1} \otimes I_{n,m} \end{bmatrix}$.

By arguing as in the proof of Theorem 2.3.6, these pencils can be shown to be strictly equivalent to block Kronecker linearizations $\begin{bmatrix} A(\lambda) & H_{k-1}(\lambda)^T \otimes I_m \end{bmatrix}$ of $P(\lambda)$ from which the left and right minimal bases and indices of $P(\lambda)$ may be easily extracted. In fact we have the following theorem.

Theorem 2.3.9. *Let* $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ *be an* $m \times n$ *matrix polynomial of grade* k *with* $m \leqslant n$ *and* $r = \operatorname{nrank} P$. *Let* $L(\lambda) \in \mathbb{L}_2(P)$ *with left ansatz vector* $w \in \mathbb{C}^k \setminus \{0\}$ *be of full Z-rank. Let* $L_t(\lambda)$ *be the pencil formed by trimming* $L(\lambda) \in \mathbb{L}_2(P)$ *with respect to* $\hat{M}$ *and* $\hat{D}$. *Then* $L_t(\lambda)$ *is a strong linearization of* $P(\lambda)$ *and the following hold.*

(a) Every minimal basis of $N_l(L_t)$ *is of the form* $\{\Lambda_k(\lambda) \otimes y_1(\lambda), \ldots, \Lambda_k(\lambda) \otimes y_{m-r}(\lambda)\}$ *where* $\{y_1(\lambda), \ldots, y_{m-r}(\lambda)\}$ *is a minimal basis of* $N_l(P)$.

(b) The left minimal indices of $L_t(\lambda)$ *are those of* $P(\lambda)$ *shifted by* $k-1$.

(c) Every minimal basis of $N_r(P)$ *is of the form* $\{\mathcal{L}_w(\hat{D}x_1(\lambda)), \ldots, \mathcal{L}_w(\hat{D}x_{n-r}(\lambda))\}$ *where* $\{x_1(\lambda), \ldots, x_{n-r}(\lambda)\}$ *is a minimal basis of* $N_r(L_t)$ *and* $\mathcal{L}_w$ *is as defined in Lemma 2.2.9.*

(d) The right minimal indices of $P(\lambda)$ *are equal to those of* $L_t(\lambda)$.

Remark 2.3.10. *For* $L(\lambda) \in \mathbb{L}_2(P)$ *of full Z-rank with left ansatz vector* $w \in \mathbb{C}^k \setminus \{0\}$ *and a given choice of* $\hat{M} \in \mathbb{C}^{k \times k}$ *such that* $\hat{M}w = \alpha e_1$, *if the matrix* $\hat{D}$ *in (2.3.10) is chosen to be the first* $n + (k-1)m$ *columns of* $(\hat{M}^T \otimes I_n) \begin{bmatrix} I_n & \\ & Q \end{bmatrix}$, *then in fact, the resulting pencil belongs to the block Kronecker ansatz space* $\mathbb{G}_k(P)$ *introduced in [26]. Also every other pencil formed by trimming* $L(\lambda)$ *with respect to* $\hat{M}$ *and some other choice of* $\hat{D}$ *is strictly equivalent to such a pencil but does not belong to* $\mathbb{G}_k(P)$.

The following example shows that linearizations $L_t(\lambda)$ arising from the pencils of full Z-rank in $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ are not subclasses of the class of block minimal bases linearizations.

Example 2.3.11. Consider $P(\lambda) = \lambda^2 A_2 + \lambda A_1 + A_0$ where

$$A_2 = \begin{bmatrix} 1 & 2 \\ 2 & 5 \\ 4 & 9 \end{bmatrix},\ A_1 = \begin{bmatrix} 3 & 4 \\ 9 & 2 \\ 15 & 10 \end{bmatrix},\ \text{and } A_0 = \begin{bmatrix} 1 & 7 \\ 2 & 5 \\ 4 & 19 \end{bmatrix}.$$

Then $L(\lambda) = \lambda \hat{X} + \hat{Y} \in \mathbb{L}_1(P)$ where

$$\hat{X} = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 2 & 5 & 0 & 0 \\ 4 & 9 & 0 & 0 \end{bmatrix},\ \hat{Y} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 3 & 4 & 1 & 7 \\ 9 & 2 & 2 & 5 \\ 15 & 10 & 4 & 19 \end{bmatrix},$$

with corresponding right ansatz vector $v = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Now $\underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{=:M} v = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, and

$$(M \otimes I_3) L(\lambda) = \lambda X + Y,$$

where

$$X = (M \otimes I_3)\hat{X} = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 5 & 0 & 0 \\ 4 & 9 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix},\ \text{and } Y = (M \otimes I_3)\hat{Y} = \begin{bmatrix} 3 & 4 & 1 & 7 \\ 9 & 12 & 2 & 5 \\ 15 & 10 & 4 & 19 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Clearly, $Z = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ has full rank with QR decomposition $Z = I_3 Z$. Hence,

$$\begin{bmatrix} 0 & Q_2^* \end{bmatrix} (M^T \otimes I_3) = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \text{ and } \left[\begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 & -1 & 1 \\ \hline 0 & 0 & 1 & 0 & 0 & 0 \end{array}\right]$$

is nonsingular. So,

$$L_t(\lambda) := \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 & -1 & 1 \end{bmatrix} L(\lambda) = \lambda \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 2 & 0 & 0 \\ 2 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 3 & 4 & 1 & 7 \\ 9 & 2 & 2 & 5 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

is a strong linearization of $P(\lambda)$. Evidently it is not a block minimal bases linearization.

Remark 2.3.12. *It is clear that if $m \geqslant n$ the size $(m + (k-1)n) \times kn$ of $L_t(\lambda)$ is the same as that of the first Frobenius companion linearization $C_1(\lambda)$ of $P(\lambda)$. On the other hand, if $m \leqslant n$, then $L_t(\lambda)$ is of size $km \times (n + (k-1)m)$ which is the same as that of the second Frobenius companion linearization $C_2(\lambda)$. Since $C_1(\lambda)$ and $C_2(\lambda)$ are the smallest among all possible Fiedler and block Kronecker linearizations of $P(\lambda)$, therefore, the size of $L_t(\lambda)$ is less than or equal to that of all such linearizations for rectangular matrix polynomials.*

Remark 2.3.13. *Although the pencils $L_t(\lambda)$ are extracted from g-linearizations in $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$, in practice, it is not necessary to form them by trimming g-linearizations. For example, given a matrix polynomial $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$, of grade k and $m \geqslant n$, we can directly build these linearizations by using the fact that they are of the form*

$$L_t(\lambda) := \tilde{D} \left(\lambda \left[\begin{array}{c|c} \alpha A_k & X_{12} \\ \hline & -\tilde{R} \end{array}\right] + \left[\begin{array}{c|c} Y_{11} & \alpha A_0 \\ \hline \tilde{R} & \end{array}\right] \right)$$

where $\alpha \neq 0$, $X_{12} + Y_{11} = \alpha \begin{bmatrix} A_{k-1} & \cdots & A_1 \end{bmatrix}$ *and* $\tilde{D}$ *and* $\tilde{R}$ *are nonsingular matrices of size* $m + (k-1)n$ *and* $(k-1)n$ *respectively. The process of trimming g-linearizations to form linearizations of this type can be seen as a means to connect the g-linearizations of* $P(\lambda)$ *with linearizations.*

Conclusion. In this chapter we have extended the spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ introduced in [61] to the case of the rectangular matrix polynomials. Though the pencils belonging to these spaces are not linearizations of the rectangular matrix polynomial $P(\lambda)$, we can still solve the complete eigenvalue problem for $P(\lambda)$ via most of these pencils. We have also shown that almost every pencil from these spaces gives rise to many strong linearizations of $P(\lambda)$ which are of the same size as that of the smallest Frobenius and block Kronecker linearizations of $P(\lambda)$. Moreover the complete eigenvalue problem for $P(\lambda)$ may be easily solved via these linearizations.

3

Global backward error analysis of solutions via linearizations arising from g-linearizations

In this chapter we carry out a global backward error analysis of the process of solving the complete eigenvalue problem associated with a rectangular matrix polynomial $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i \in \mathbb{C}[\lambda]^{m\times n}$ of grade k by using linearizations that arise from a g-linearization in the vector spaces $\mathbb{L}_1(P)$ or $\mathbb{L}_2(P)$ introduced in Chapter 2. The analysis is an extension of the one in [22] for block Kronecker linearizations. Any solution of such a problem involves finding the finite and infinite eigenvalues and associated elementary divisors as well as the left and right minimal bases and indices of $P(\lambda)$. Typically this is done by initially finding the said quantities for some choice of strong linearization via very effective backward stable methods like the *staircase algorithm* proposed in [74] and further developed in [19, 20]. The backward stability of such algorithms guarantees that any computed solution of the eigenvalue problem corresponding to a linearization say, $L(\lambda)$ of $P(\lambda)$, is the exact solution of the problem for a pencil $L(\lambda) + \Delta L(\lambda)$ where $|||\Delta L|||/|||L||| = O(\mathbf{u})$ with respect to some norm $|||\cdot|||$, $\mathbf{u}$ being unit roundoff. Unit roundoff $\mathbf{u}$ is the largest relative rounding or representation error that can occur when a normalized number, say x, is rounded or represented by the nearest floating point number in a given finite precision environment under a well defined set of rounding or representation rules. Thus it is given by

$$\mathbf{u} = \sup_{l \le |x| \le u} \left| \frac{\mathrm{fl}(x) - x}{x} \right|,$$

where $\mathrm{fl}(x)$ is the rounded version of x or the representation of x in the system, and l and u are the smallest and largest positive normalized numbers respectively that can be stored in the system. The value of $\mathbf{u}$ depends both on the finite precision system and the rules of rounding. The values of $\mathbf{u}$ are approximately 6×10^{-8} and 10^{-16} for IEEE single and double precision respectively and these are the finite precision systems most widely used in computations. The solution of the complete eigenvalue problem for $P(\lambda)$ is then computed from the solution for $L(\lambda) + \Delta L(\lambda)$ by applying the same recovery rules to $L(\lambda) + \Delta L(\lambda)$ that would have been applied to the solution for $L(\lambda)$ if it were available. Following [22], the process is said to be globally backward stable if it is the exact solution of the complete eigenvalue problem for $P(\lambda) + \Delta P(\lambda)$ with the following conditions being met.

(a) If $P(\lambda)$ is of grade k, the perturbed pencil $L(\lambda)+\Delta L(\lambda)$ a strong linearization of $P(\lambda) + \Delta P(\lambda)$ of grade k such that $\|\Delta P\|/\|P\| = O(\mathbf{u})$.

(b) The rules for extracting the minimal indices of $P(\lambda)$ from those of $L(\lambda)$ remain the same when replaced by $P(\lambda) + \Delta P(\lambda)$ and $L(\lambda) + \Delta L(\lambda)$ respectively.

The analysis in [22], showed that (a) and (b) are satisfied for optimal choices of block Kronecker linearizations of $P(\lambda)$ with respect to the norm $\|\cdot\|_F$. In particular, it was shown that there exists a constant $C_{L,P}$ depending on $L(\lambda)$ and $P(\lambda)$ such that

$$\frac{\|\Delta P\|_F}{\|P\|_F} \leqslant C_{L,P}\frac{\|\Delta L\|_F}{\|L\|_F},$$

where, $C_{L,P} \approxeq k^3\sqrt{m+n}$ under certain conditions that are satisfied by the appropriate choice of block Kronecker linearizations and scaling of $P(\lambda)$.

We establish that the same analysis can be extended to solutions obtained via linearizations $L_t(\lambda)$ of $P(\lambda) \in \mathbb{C}[\lambda]^{m\times n}$ that arise from g-linearizations in $\mathbb{L}_1(P)$ when $m \geqslant n$. Similar arguments can easily complete the corresponding analysis for the case $m \leqslant n$ with respect to linearizations that arise from g-linearizations in $\mathbb{L}_2(P)$.

Our choice of norm $\|\cdot\|_F$ on $\mathbb{C}[\lambda]^{m\times n}$ considered as a vector space over $\mathbb{C}$ is not submultiplicative. The following lemma from [22] which bounds the Frobenius norm of the product of two matrix polynomials will therefore be useful in the analysis. The proof given here is made independently. For notational convenience we set $\Lambda_{k,p}(\lambda) := (\Lambda_k(\lambda) \otimes I_p)$, where $\Lambda_k(\lambda)$ is as in (2.0.3).

Lemma 3.0.1. *Let $P(\lambda) = \sum_{i=0}^{d_1} \lambda^i A_i$ and $Q(\lambda) = \sum_{i=0}^{d_2} \lambda^i B_i$ be two matrix polynomials such that all the products below are defined. Then the following inequalities hold.*

(a) $\|PQ\|_F \leqslant \min\{\sqrt{d_1+1}, \sqrt{d_2+1}\}\|P\|_F\|Q\|_F$.

(b) $\|P\Lambda_{k,p}\|_F \leqslant \min\{\sqrt{d_1+1}, \sqrt{k}\}\|P\|_F$.

Proof. To prove part (a), we write $PQ(\lambda)$ as $\sum_{i=0}^{d_1} \lambda^i A_i Q(\lambda)$. Then,

$$\|PQ\|_F \leqslant \sum_{i=0}^{d_1} \|\lambda^i A_i Q(\lambda)\|_F = \sum_{i=0}^{d_1} \|A_i Q(\lambda)\|_F.$$

As $\|A_i Q(\lambda)\|_F \leqslant \|A_i\|_F \|Q\|_F$, $\|PQ\|_F \leqslant \sqrt{d_1+1}\|P\|_F\|Q\|_F$. Again by writing $PQ(\lambda)$ as $\sum_{i=0}^{d_2} \lambda^i P(\lambda) B_i$ we have $\|PQ\|_F \leqslant \sqrt{d_2+1}\|P\|_F\|Q\|_F$. Hence

$$\|PQ\|_F \leqslant \min\{\sqrt{d_1+1}, \sqrt{d_2+1}\}\|P\|_F\|Q\|_F.$$

To prove part (b), partition $A_i, i = 0, \ldots, d_1$ as $A_i = \begin{bmatrix} A_{i,k-1} & \cdots & A_{i,0} \end{bmatrix}$ where $A_{i,j}, i = 0, \ldots, d_1, j = 0, \ldots, k-1$ has p columns. Now

$$\|P\Lambda_{k,p}\|_F^2 = \left\|\sum_{i=0}^{d_1} \lambda^i A_i \Lambda_{k,p}\right\|_F^2 = \left\|\sum_{i=0}^{d_1}\left(\sum_{j=0}^{k-1} \lambda^{i+j} A_{i,j}\right)\right\|_F^2 \leqslant (d_1+1)\sum_{i=0}^{d_1}\left\|\sum_{j=0}^{k-1} \lambda^{i+j} A_{i,j}\right\|_F^2$$

$$= (d_1+1)\sum_{i=0}^{d_1}\sum_{j=0}^{k-1} \|A_{i,j}\|_F^2 = (d_1+1)\sum_{i=0}^{d_1} \|A_i\|_F^2 = (d_1+1)\|P\|_F^2.$$

Again partition $P(\lambda)$ as $P(\lambda) = \begin{bmatrix} P_{k-1}(\lambda) & \cdots & P_0(\lambda) \end{bmatrix}$ where $P_i(\lambda), i = 0, \ldots, k-1$ has p columns. Then

$$\|P\Lambda_{k,p}\|_F^2 = \left\|\sum_{i=0}^{k-1} \lambda^i P_i(\lambda)\right\|_F^2 = \left\|\sum_{i=0}^{k-1} \lambda^i \sum_{j=0}^{d_1} \lambda^j A_{j,i}\right\|_F^2 = \left\|\sum_{i=0}^{k-1}\sum_{j=0}^{d_1} \lambda^{i+j} A_{j,i}\right\|_F^2$$

$$\leqslant k\sum_{i=0}^{k-1}\left\|\sum_{j=0}^{d_1} \lambda^{i+j} A_{j,i}\right\|_F^2 = k\sum_{i=0}^{k-1}\sum_{j=0}^{d_1} \|A_{j,i}\|_F^2 = k\sum_{i=0}^{k-1} \|P_i\|_F^2 = k\|P\|_F^2.$$

Hence $\|P\Lambda_{k,p}\|_F \leqslant \min\{\sqrt{d_1+1}, \sqrt{k}\}\|P\|_F$. □

3.1 Global backward error analysis of solutions using $L_t(\lambda)$

Initially we analyse the global backward stability of the process of computing a solution of the complete eigenvalue problem for $P(\lambda)$ via linearizations $\hat{L}_t(\lambda)$ of the form (2.3.4). Later on we extend this analysis to the case where any linearization $L_t(\lambda)$ arising from a g-linearization in $\mathbb{L}_1(P)$ is used.

Since the matrix $\tilde{R} \in \mathbb{C}^{(k-1)n \times (k-1)n}$ of $\hat{L}_t(\lambda)$ given by (2.3.4) is upper triangular and nonsingular, $\hat{L}_t(\lambda)$ is a strong block minimal bases pencil of the form

$$\left[\begin{array}{c} A(\lambda) \\ \hline B(\lambda) \end{array}\right]$$

where,

$$A(\lambda) = \lambda \left[\begin{array}{cc} \alpha A_k & X_{12} \end{array}\right] + \left[\begin{array}{cc} Y_{11} & \alpha A_0 \end{array}\right] \tag{3.1.1}$$

with $\alpha \neq 0$ and $X_{12} + Y_{11} = \alpha \left[\begin{array}{ccc} A_{k-1} & \cdots & A_1 \end{array}\right]$, and

$$B(\lambda) = \lambda \left[\begin{array}{cc} 0_{(k-1)n\times n} & -\tilde{R} \end{array}\right] + \left[\begin{array}{cc} \tilde{R} & 0_{(k-1)n\times n} \end{array}\right]. \tag{3.1.2}$$

Note that $\Lambda_{k,n}(\lambda)^T$ is a dual minimal basis of $B(\lambda)$.

Any computed solution of the complete eigenvalue problem associated with $\hat{L}_t(\lambda)$ is an exact solution of a perturbed pencil $\hat{L}_t(\lambda) + \Delta\hat{L}_t(\lambda)$ where

$$\Delta\hat{L}_t(\lambda) = \left[\begin{array}{c} \Delta A(\lambda) \\ \hline \Delta B(\lambda) \end{array}\right]$$

with $\Delta A(\lambda) \in \mathbb{C}[\lambda]^{m\times kn}$ and $\Delta B(\lambda) \in \mathbb{C}[\lambda]^{(k-1)n\times kn}$ so that

$$\hat{L}_t(\lambda) + \Delta\hat{L}_t(\lambda) = \left[\begin{array}{c} A(\lambda) + \Delta A(\lambda) \\ \hline B(\lambda) + \Delta B(\lambda) \end{array}\right].$$

The global backward stability analysis of the process of computing a solution of the complete eigenvalue problem for $P(\lambda)$ via linearizations $\hat{L}_t(\lambda)$ of the form (2.3.4) can also be performed by applying the results in [22]. These may then be used to extend the analysis to linearizations $L_t(\lambda)$ arising from g-linearizations. However, from the point of view of obtaining a larger set of linearizations that guarantee global backward stability, we also undertake a more detailed analysis that results in a different bound on $\|\Delta L_t\|_F$ and a different expression for the constant $C_{L,P}$. We first perform the analysis by applying the results in [22] in Section 3.1.1 and then undertake a more detailed analysis in Section 3.1.2. Finally we provide arguments to support the claim that the detailed analysis provides a larger set of linearizations that ensure global backward stability.

3.1.1 Global backward error analysis: A brief approach

As a consequence of [22, Theorem 5.22], the computation of the complete eigenvalue problem for $P(\lambda)$ can be shown to be globally backward stable for suitable choices of linearizations of the form (2.3.4). This result from [22] is stated below in a form that is relevant to our analysis.

Theorem 3.1.1. *Let* $P(\lambda) = \Sigma_{i=0}^{k} \lambda^i A_i$ *be of grade* k *and* $L(\lambda) = \left[\begin{array}{c} M(\lambda) \\ \hline H_{k-1}(\lambda) \otimes I_n \end{array}\right]$ *be a block Kronecker pencil such that* $P(\lambda) = M(\lambda)\Lambda_{k,n}(\lambda)$ *where* $H_{k-1}(\lambda)$ *is given by (2.3.6). If* $\Delta L(\lambda)$ *is a pencil satisfying* $\|\Delta L\|_F < \frac{1}{2k^{3/2}}$*, and is of the same size as* $L(\lambda)$*, then* $L(\lambda) + \Delta L(\lambda)$ *is a strong linearization of* $P(\lambda) + \Delta P(\lambda)$ *of grade* k *such that*

$$\frac{\|\Delta P\|_F}{\|P\|_F} \leqslant 2k \frac{\|L\|_F}{\|P\|_F}(1 + \|M\|_F)\frac{\|\Delta L\|_F}{\|L\|_F}.$$

Also the right minimal indices of $L(\lambda)+\Delta L(\lambda)$ *are those of* $P(\lambda)+\Delta P(\lambda)$ *shifted by* $k-1$ *and the left minimal indices of* $L(\lambda) + \Delta L(\lambda)$ *are the same as those of* $P(\lambda) + \Delta P(\lambda)$*. This coincides with the corresponding relationship between the minimal indices of* $L(\lambda)$ *and* $P(\lambda)$*.*

This can be used to prove the following result for linearizations of the form (2.3.4).

Corollary 3.1.2. *Let* $\hat{L}_t(\lambda)$ *be any linearization of* $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i \in \mathbb{C}[\lambda]^{m\times n}$ *of grade* k *with* $m \geqslant n$*, of the form (2.3.4). Let* $A(\lambda)$ *and* $B(\lambda)$ *be the blocks of* $\hat{L}_t(\lambda)$ *as specified by (3.1.1) and (3.1.2) respectively and* $\tilde{R}$ *be the nonsingular upper triangular matrix appearing in the block* $B(\lambda)$*. If* $\Delta\hat{L}_t(\lambda)$ *is any pencil of the same size as* $\hat{L}_t(\lambda)$ *such that*

$$\|\Delta \hat{L}_t\|_F < \frac{1}{2k^{3/2} \max\{1, 1/\sigma_{\min}(\tilde{R})\}}$$

then $\hat{L}_t(\lambda) + \Delta\hat{L}_t(\lambda)$ *is a strong linearization of a matrix polynomial* $P(\lambda) + \Delta P(\lambda)$ *of grade* k *and*

$$\frac{\|\Delta P\|_F}{\|P\|_F} \leqslant C^b_{\hat{L}_t,P} \frac{\|\Delta \hat{L}_t\|_F}{\|\hat{L}_t\|_F},$$

where

$$C^b_{\hat{L}_t,P} = 2k \frac{\|\hat{L}_t\|_F}{|\alpha| \|P\|_F}(1 + \|A\|_F) \max\{1, 1/\sigma_{\min}(\tilde{R})\}. \tag{3.1.3}$$

Also the right minimal indices of $\hat{L}_t(\lambda) + \Delta\hat{L}_t(\lambda)$ *are those of* $P(\lambda) + \Delta P(\lambda)$ *shifted by* $k-1$ *and the left minimal indices of* $\hat{L}_t(\lambda)+\Delta\hat{L}_t(\lambda)$ *are the same as those* *of $P(\lambda) + \Delta P(\lambda)$. This coincides with the corresponding relationship between the minimal indices of $\hat{L}_t(\lambda)$ and $P(\lambda)$.*

Proof. Observe that,

$$\left[\begin{array}{c|c} I_m & \\ \hline & -\tilde{R}^{-1} \end{array}\right] \hat{L}_t(\lambda) = \left[\begin{array}{c} A(\lambda) \\ \hline H_{k-1}(\lambda) \otimes I_n \end{array}\right]$$

where $A(\lambda)$ and $H_{k-1}(\lambda)$ are given by (3.1.1) and (2.3.6) respectively, is a block Kronecker pencil. Therefore

$$\left[\begin{array}{c|c} I_m & \\ \hline & -\tilde{R}^{-1} \end{array}\right] (\hat{L}_t(\lambda) + \Delta\hat{L}_t(\lambda)) = \left[\begin{array}{c} A(\lambda) \\ \hline H_{k-1}(\lambda) \otimes I_n \end{array}\right] + \underbrace{\left[\begin{array}{c|c} I_m & \\ \hline & -\tilde{R}^{-1} \end{array}\right] \Delta\hat{L}_t(\lambda)}_{:=\Delta\mathcal{L}_t(\lambda)}.$$

Now as $\|\Delta\hat{L}_t\|_F < \frac{1}{2k^{3/2}\max\{1, 1/\sigma_{\min}(\tilde{R})\}}$, we have,

$$\|\Delta\mathcal{L}_t\|_F \leqslant \left\| \left[\begin{array}{c|c} I_m & \\ \hline & -\tilde{R}^{-1} \end{array}\right] \right\|_2 \|\Delta\hat{L}_t\|_F \leqslant \max\{1, 1/\sigma_{\min}(\tilde{R})\}\|\Delta\hat{L}_t\|_F < \frac{1}{2k^{3/2}}.$$

As $A(\lambda)\Lambda_{k,n}(\lambda) = \alpha P(\lambda)$, by Theorem 3.1.1, $\left[\begin{array}{c} A(\lambda) \\ \hline H_{k-1}(\lambda) \otimes I_n \end{array}\right] + \Delta\mathcal{L}_t(\lambda)$ is a strong linearization of $\alpha(P(\lambda) + \Delta P(\lambda))$ where

$$\frac{\|\Delta P\|_F}{\|P\|_F} \leqslant 2k \frac{\|\hat{L}_t\|_F}{|\alpha| \|P\|_F} (1 + \|A\|_F) \frac{\|\Delta\mathcal{L}_t\|_F}{\|\hat{L}_t\|_F}.$$

This implies $\hat{L}_t(\lambda) + \Delta\hat{L}_t(\lambda)$ is a strong linearization of $P(\lambda) + \Delta P(\lambda)$ where

$$\frac{\|\Delta P\|_F}{\|P\|_F} \leqslant 2k \frac{\|\hat{L}_t\|_F}{|\alpha| \|P\|_F} (1 + \|A\|_F) \max\{1, 1/\sigma_{\min}(\tilde{R})\} \frac{\|\Delta\hat{L}_t\|_F}{\|\hat{L}_t\|_F} = C^b_{\hat{L}_t,P} \frac{\|\Delta\hat{L}_t\|_F}{\|\hat{L}_t\|_F}.$$

Also the right minimal indices of $\hat{L}_t(\lambda) + \Delta\hat{L}_t(\lambda)$ are those of $P(\lambda) + \Delta P(\lambda)$ shifted by $k-1$ and the left minimal indices of $\hat{L}_t(\lambda) + \Delta\hat{L}_t(\lambda)$ are the same as those of left minimal indices of $P(\lambda) + \Delta P(\lambda)$. □

Now we extend the above result to solutions of the complete eigenvalue problem for $P(\lambda)$ obtained via any linearization $L_t(\lambda)$ arising from a g-linearization in $\mathbb{L}_1(P)$. As noted in Section 2.3, any such linearization $L_t(\lambda)$ is strictly equivalent to a linearization $\hat{L}_t(\lambda)$ of the form (2.3.4). Using this fact and the above result we have the following corollary.

Corollary 3.1.3. *Let $L_t(\lambda)$ be any linearization of $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i \in \mathbb{C}[\lambda]^{m\times n}$ of grade k with $m \geqslant n$, arising from a g-linearization in $\mathbb{L}_1(P)$. Let $\hat{L}_t(\lambda) = \tilde{D}^{-1}L_t(\lambda)$ where $\tilde{D}$ is as given in (2.3.5). Then $\hat{L}_t(\lambda)$ is of the form (2.3.4). Let $A(\lambda)$ and $B(\lambda)$ be the blocks of $\hat{L}_t(\lambda)$ as specified by (3.1.1) and (3.1.2) respectively and $\tilde{R}$ be the nonsingular upper triangular matrix appearing in the block $B(\lambda)$. If $\Delta L_t(\lambda)$ is any pencil with the same size as $L_t(\lambda)$ such that*

$$\|\Delta L_t\|_F < \frac{\sigma_{\min}(\tilde{D})}{2k^{3/2}\max\{1, 1/\sigma_{\min}(\tilde{R})\}} \tag{3.1.4}$$

then $L_t(\lambda) + \Delta L_t(\lambda)$ is a strong linearization of a matrix polynomial $P(\lambda) + \Delta P(\lambda)$ of grade k and

$$\frac{\|\Delta P\|_F}{\|P\|_F} \leqslant C^b_{L_t,P}\frac{\|\Delta L_t\|_F}{\|L_t\|_F},$$

where

$$C^b_{L_t,P} = 2k\frac{\kappa_2(\tilde{D})}{|\alpha|}\frac{\|\hat{L}_t\|_F}{\|P\|_F}(1 + \|A\|_F)\max\{1, 1/\sigma_{\min}(\tilde{R})\}. \tag{3.1.5}$$

Also the right minimal indices of $L_t(\lambda) + \Delta L_t(\lambda)$ are those of $P(\lambda) + \Delta P(\lambda)$ shifted by $k-1$ and the left minimal indices of $L_t(\lambda)+\Delta L_t(\lambda)$ are the same as those of $P(\lambda) + \Delta P(\lambda)$. This coincides with the corresponding relationship between the minimal indices of $L_t(\lambda)$ and $P(\lambda)$.

Proof. As $\hat{L}_t(\lambda) = \tilde{D}^{-1}L_t(\lambda)$ is of the form (2.3.4), using (3.1.4) and the fact that

$$L_t(\lambda) + \Delta L_t(\lambda) = \tilde{D}(\hat{L}_t(\lambda) + \tilde{D}^{-1}\Delta L_t(\lambda)),$$

we have,

$$\|\tilde{D}^{-1}\Delta L_t\|_F < \frac{1}{2k^{3/2}\max\{1, 1/\sigma_{\min}(\tilde{R})\}}.$$

Therefore by Corollary 3.1.2, $\hat{L}_t(\lambda) + \tilde{D}^{-1}\Delta L_t(\lambda)$ is a strong block minimal bases linearization of some polynomial $P(\lambda) + \Delta P(\lambda)$ of grade k such that

$$\frac{\|\Delta P\|_F}{\|P\|_F} \leqslant \frac{2k}{|\alpha|}\frac{\|\hat{L}_t\|_F}{\|P\|_F}(1 + \|A\|_F)\max\{1, 1/\sigma_{\min}(\tilde{R})\}\frac{\|\tilde{D}^{-1}\Delta L_t\|_F}{\|\hat{L}_t\|_F}.$$

Using the relations, $\|\tilde{D}^{-1}\Delta L_t\|_F \leqslant \|\tilde{D}^{-1}\|_2\|\Delta L_t\|_F$ and $\|L_t\|_F \leqslant \|\tilde{D}\|_2\|\hat{L}_t\|_F$, we get,

$$\begin{aligned}\frac{\|\Delta P\|_F}{\|P\|_F} &\leqslant \frac{2k}{|\alpha|}\frac{\|\hat{L}_t\|_F}{\|P\|_F}(1 + \|A\|_F)\max\{1, 1/\sigma_{\min}(\tilde{R})\}\frac{\|\tilde{D}^{-1}\|_2\|\Delta L_t\|_F\|\tilde{D}\|_2}{\|L_t\|_F}\\ &= 2k\frac{\kappa_2(\tilde{D})}{|\alpha|}\frac{\|\hat{L}_t\|_F}{\|P\|_F}(1 + \|A\|_F)\max\{1, 1/\sigma_{\min}(\tilde{R})\}\frac{\|\Delta L_t\|_F}{\|L_t\|_F}.\end{aligned}$$

As $L_t(\lambda)+\Delta L_t(\lambda) = \tilde{D}(\hat{L}_t(\lambda)+\tilde{D}^{-1}\Delta L_t(\lambda))$, the pencils $L_t(\lambda)+\Delta L_t(\lambda)$ and $\hat{L}_t(\lambda)+\tilde{D}^{-1}\Delta L_t(\lambda)$ are strictly equivalent. Hence $L_t(\lambda)+\Delta L_t(\lambda)$ is a strong linearization of the polynomial $P(\lambda)+\Delta P(\lambda)$ and the rules for recovering the minimal indices of $P(\lambda)+\Delta P(\lambda)$ from those of $L_t(\lambda)+\Delta L_t(\lambda)$ are the same as the ones applied to $\hat{L}_t(\lambda)+\tilde{D}^{-1}\Delta L_t(\lambda)$. Therefore it follows from Corollary 3.1.2, that the right minimal indices of $L_t(\lambda)+\Delta L_t(\lambda)$ are those of $P(\lambda)+\Delta P(\lambda)$ shifted by $k-1$ and left minimal indices of $L_t(\lambda)+\Delta L_t(\lambda)$ are same as those of $P(\lambda)+\Delta P(\lambda)$. □

Clearly, the solution of the complete eigenvalue problem for $P(\lambda)$ via linearizations $L_t(\lambda)$ is a globally backward stable process if the constant $C^b_{L_t,P}$ given by (3.1.5) is moderate. For instance, one such choice is any $L_t(\lambda)$ for which $\kappa_2(\tilde{D}) \approxeq 1$, $\|A\|_F \approxeq |\alpha|\|P\|_F \approxeq 1$ and $\kappa_2(\tilde{R}) \approxeq \sigma_{\min}(\tilde{R}) \approxeq 1$.

3.1.2 Global backward error analysis: A detailed approach

Our initial aim is to show that for small enough $\|\Delta\hat{L}_t\|_F$, $\hat{L}_t(\lambda)+\Delta\hat{L}_t(\lambda)$ is a strong block minimal bases linearization of $P(\lambda)+\Delta P(\lambda)$ of grade k such that $\|\Delta P\|_F/\|P\|_F$ is bounded above by a small multiple of $\|\Delta\hat{L}_t\|_F/\|\hat{L}_t\|_F$.

We establish an upper bound on $\|\Delta B\|_F$ such that $\hat{L}_t(\lambda)+\Delta\hat{L}_t(\lambda)$ is a strong block minimal bases pencil given by Definition 2.3.4. For this it is enough to show that the following conditions are satisfied.

Condition (A) $B(\lambda)+\Delta B(\lambda)$ is a minimal basis with all row degrees equal to one.

Condition (B) There exists a matrix polynomial $\Delta D(\lambda) \in \mathbb{C}[\lambda]^{kn\times n}$ of grade $k-1$ such that the polynomial $\Lambda_{k,n}(\lambda)^T+\Delta D(\lambda)^T$ is a dual minimal basis of $B(\lambda)+\Delta B(\lambda)$ with all row degrees equal to $k-1$.

Following the strategy in [22], we will use the concept of convolution matrices associated with a matrix polynomial. Recall that $\widehat{C}_j(P)$ denotes the j^{th} convolution matrix associated with $\mathrm{rev}_k P(\lambda)$ as in (1.3.3).

The next theorem from [22] for convolution matrices will be useful to show that for sufficiently small $\|\Delta B\|_F$, $B(\lambda)+\Delta B(\lambda)$ can be a minimal basis with all row degrees equal to 1.

Theorem 3.1.4. *For any positive integer l, let $N(\lambda) = A+\lambda B \in \mathbb{C}[\lambda]^{ln\times(l+1)n}$ and $\widehat{C}_j(N)$ for $j=0,1,\ldots,$ be the sequence of matrices as in (1.3.3) of $N(\lambda)$. Then $N(\lambda)$ is a minimal basis with all its row degrees equal to 1 and all the row degrees of any dual minimal basis equal to l, if and only if $\widehat{C}_{l-1}(N) \in \mathbb{C}^{l(l+1)n\times l(l+1)n}$ is nonsingular and $\widehat{C}_l(N) \in \mathbb{C}^{l(l+2)n\times(l+1)^2 n}$ has full row rank.*

Recall that for any matrix M, $\sigma_{\min}(M)$ denotes the smallest singular value of M. Observing that $B(\lambda) = -\tilde{R}(H_{k-1}(\lambda) \otimes I_n)$ where $H_{k-1}(\lambda)$ is given by (2.3.6), the next lemma will be useful in establishing a bound on $\|\Delta B\|_F$ that achieves the desired objectives.

Lemma 3.1.5. *For $k \geqslant 2$, let $\tau(\lambda) = H_{k-1}(\lambda) \otimes I_n$, where $H_{k-1}(\lambda)$ is as in (2.3.6). Then,*

$$\sigma_{\min}(\widehat{C}_{k-2}(\tau)) = \sigma_{\min}(\widehat{C}_{k-1}(\tau)) = 2\sin\left(\frac{\pi}{4k-2}\right) \geqslant \frac{3}{2k}.$$

Proof. Since $\widehat{C}_j(\tau) = \widehat{C}_j(H_{k-1}) \otimes I_n$ for $j = k-1, k-2$, the first two equalities follow from Lemma 1.3.2 and the last inequality follows from the fact that $\sin(x) \geqslant \frac{3x}{\pi}$ for $0 \leqslant x \leqslant \frac{\pi}{6}$. □

The following result bounds $\|\Delta B\|_F$ such that $B(\lambda) + \Delta B(\lambda)$ is a minimal basis with all row degrees equal to 1.

Theorem 3.1.6. *Let $B(\lambda)$ be as in (3.1.2), and $\Delta B(\lambda) \in \mathbb{C}[\lambda]^{(k-1)n \times kn}$ be any pencil such that*

$$\|\Delta B\|_F < \frac{3\sigma_{\min}(\tilde{R})}{2k^{3/2}}. \tag{3.1.6}$$

Then $B(\lambda) + \Delta B(\lambda)$ is a minimal basis with all its row degrees equal to 1 and all row degrees of any minimal basis dual to it equal to $k-1$.

Proof. In view of Theorem 3.1.4, the proof follows by establishing that $\widehat{C}_{k-2}(B+\Delta B)$ is nonsingular and $\widehat{C}_{k-1}(B+\Delta B)$ has full row rank. For $j = k-1$ or $k-2$,

$$\sigma_{\min}(\widehat{C}_j(B)) = \sigma_{\min}\left(\begin{bmatrix} -\tilde{R} & & \\ & \ddots & \\ & & -\tilde{R} \end{bmatrix} \widehat{C}_j(\tau)\right) \geqslant \sigma_{\min}(\tilde{R})\sigma_{\min}(\widehat{C}_j(\tau)),$$

where $\tau(\lambda)$ is as in Lemma 3.1.5. Now by Lemma 3.1.5,

$$\sigma_{\min}(\widehat{C}_j(B)) \geqslant 2\sigma_{\min}(\tilde{R})\sin\left(\frac{\pi}{4k-2}\right) \geqslant \sigma_{\min}(\tilde{R})\frac{3}{2k}. \tag{3.1.7}$$

Since $\tilde{R}$ is nonsingular, $\widehat{C}_{k-2}(B)$ is nonsingular and $\widehat{C}_{k-1}(B)$ has full row rank. By Lemma 1.3.1(a),

$$\widehat{C}_j(B+\Delta B) = \widehat{C}_j(B) + \widehat{C}_j(\Delta B)$$

for $j = k-1$ and $k-2$. Again for $j = k-1$ and $k-2$,

$$\|\widehat{C}_j(\Delta B)\|_F = \sqrt{j+1}\|\Delta B\|_F < \sqrt{j+1}\frac{3\sigma_{\min}(\tilde{R})}{2k^{3/2}} < \sqrt{k}\frac{3\sigma_{\min}(\tilde{R})}{2k^{3/2}} \leqslant \sigma_{\min}(\widehat{C}_j(B)),$$

where the equality follows from Lemma 1.3.1(b), the first and last inequality is due to equations (3.1.6) and (3.1.7) respectively. As $\|\widehat{C}_j(\Delta B)\|_F < \sigma_{\min}(\widehat{C}_j(B))$, for $j = k-1$ and $k-2$, $\widehat{C}_{k-2}(B+\Delta B)$ is nonsingular and $\widehat{C}_{k-1}(B+\Delta B)$ has full row rank. $\square$

To establish the required upper bound on $\|\Delta B\|_F$ such that both **Condition (A)** and **Condition (B)** are fulfilled, we use [22, Corollary 5.16] which is stated below as a lemma.

Lemma 3.1.7. *Let $\Delta D(\lambda) \in \mathbb{C}^{kn\times n}$ be a matrix polynomial of grade $k-1$ and $\|\Delta D\|_F < \frac{1}{\sqrt{2}}$. Then $\Lambda_{k,n}(\lambda)^T + \Delta D(\lambda)^T$ is a minimal basis with all the row degrees equal to $k-1$ and with all row degrees of any minimal basis dual to it equal to 1.*

The following result gives the desired upper bound on $\|\Delta B\|_F$.

Theorem 3.1.8. *Let $B(\lambda)$ as in (3.1.2) and $\Delta B(\lambda) \in \mathbb{C}[\lambda]^{(k-1)n\times kn}$ be any pencil such that*

$$\|\Delta B\|_F < \frac{\sigma_{\min}(\tilde{R})}{2k^{3/2}}. \tag{3.1.8}$$

Then there exists a matrix polynomial $\Delta D(\lambda) \in \mathbb{C}[\lambda]^{kn\times n}$ of grade $k-1$ such that

(a) $B(\lambda) + \Delta B(\lambda)$ and $\Lambda_{k,n}(\lambda)^T + \Delta D(\lambda)^T$ are dual minimal bases, with all the row degrees equal to 1 and $k-1$ respectively, and

(b) $\|\Delta D\|_F \leqslant \frac{k\sqrt{2}}{\sigma_{\min}(\tilde{R})}\|\Delta B\|_F < \frac{1}{\sqrt{2k}}$.

Proof. As

$$\|\Delta B\|_F < \frac{\sigma_{\min}(\tilde{R})}{2k^{3/2}} < \frac{3\sigma_{\min}(\tilde{R})}{2k^{3/2}},$$

therefore, by Theorem 3.1.6, $B(\lambda)+\Delta B(\lambda)$ is minimal basis with all its row degrees equal to 1 and all row degrees of any minimal basis dual to it equal to $k-1$. If possible, suppose $\Delta D(\lambda)$ is a matrix polynomial of grade $k-1$ satisfying

$$(B(\lambda) + \Delta B(\lambda))(\Lambda_{k,n}(\lambda) + \Delta D(\lambda)) = 0 \tag{3.1.9}$$

Since $B(\lambda)\Lambda_{k,n}(\lambda) = 0$, so (3.1.9) becomes

$$\begin{aligned}
&(B(\lambda) + \Delta B(\lambda))\Delta D(\lambda) = -\Delta B(\lambda)\Lambda_{k,n}(\lambda)\\
\Leftrightarrow &\widehat{C}_0((B+\Delta B)\Delta D) = -\widehat{C}_0(\Delta B\Lambda_{k,n})\\
\Leftrightarrow &\widehat{C}_{k-1}(B+\Delta B)C_0(\Delta D) = -\widehat{C}_0(\Delta B\Lambda_{k,n}).
\end{aligned} \tag{3.1.10}$$

where (3.1.10) follows from Lemma 1.3.1(c). As $\widehat{C}_{k-1}(B+\Delta B)$ has full row rank, we have

$$\widehat{C}_0(\Delta D) = -(\widehat{C}_{k-1}(B+\Delta B))^{\dagger}\widehat{C}_0(\Delta B\Lambda_{k,n}),$$

which solves (3.1.10). This gives the matrix polynomial $\Delta D(\lambda)$ of grade $k-1$ satisfying (3.1.9). To prove part (b) note that,

$$\|\Delta D\|_F = \|\widehat{C}_0(\Delta D)\|_F \leqslant \|(\widehat{C}_{k-1}(B+\Delta B))^{\dagger}\|_2\|\widehat{C}_0(\Delta B\Lambda_{k,n})\|_F. \tag{3.1.11}$$

As $\|(\widehat{C}_{k-1}(B+\Delta B))^{\dagger}\|_2 = \frac{1}{\sigma_{\min}(\widehat{C}_{k-1}(B+\Delta B))}$, using (3.1.7), (3.1.8) and standard results for perturbation of singular values, it follows that

$$\|(\widehat{C}_{k-1}(B+\Delta B))^{\dagger}\|_2 \leqslant \frac{1}{\frac{3\sigma_{\min}(\tilde{R})}{2k} - \sqrt{k}\|\Delta B\|_F} \leqslant \frac{k}{\sigma_{\min}(\tilde{R})},$$

$$\text{and } \|\widehat{C}_0(\Delta B\Lambda_{k,n})\|_F = \|\Delta B\Lambda_{k,n}\|_F \leqslant \sqrt{2}\|\Delta B\|_F < \frac{\sigma_{\min}(\tilde{R})}{\sqrt{2}k^{3/2}}.$$

Hence from (3.1.11), $\|\Delta D\|_F \leqslant \frac{\sqrt{2}k}{\sigma_{\min}(\tilde{R})}\|\Delta B\|_F < \frac{1}{\sqrt{2k}}$. Now as the matrix polynomial $\Delta D(\lambda)$ of grade $k-1$ satisfies (3.1.9) with $\|\Delta D\|_F < \frac{1}{\sqrt{2k}} < \frac{1}{\sqrt{2}}$, part (a) follows from Lemma 3.1.7. □

Next we have the main result which completes the global backward error analysis for solutions of the complete eigenvalue problem for $P(\lambda)$ obtained from the linearizations $\hat{L}_t(\lambda)$ given by (2.3.4).

Theorem 3.1.9. *Let $\hat{L}_t(\lambda)$ be any linearization of $P(\lambda) = \sum_{i=0}^{k}\lambda^i A_i \in \mathbb{C}[\lambda]^{m\times n}$ of grade k with $m \geqslant n$, of the form (2.3.4). Let $A(\lambda)$ and $B(\lambda)$ be the blocks of $\hat{L}_t(\lambda)$ as specified by (3.1.1) and (3.1.2) respectively and $\tilde{R} \in \mathbb{C}^{(k-1)n\times(k-1)n}$ be the nonsingular upper triangular matrix appearing in the block $B(\lambda)$. If $\Delta\hat{L}_t(\lambda)$ is any pencil of the same size as $\hat{L}_t(\lambda)$ such that*

$$\|\Delta\hat{L}_t\|_F < \frac{\sigma_{\min}(\tilde{R})}{2k^{3/2}},$$

then $\hat{L}_t(\lambda) + \Delta\hat{L}_t(\lambda)$ is a strong linearization of a matrix polynomial $P(\lambda) + \Delta P(\lambda)$ of grade k and

$$\frac{\|\Delta P\|_F}{\|P\|_F} \leqslant C^d_{\hat{L}_t,P}\frac{\|\Delta\hat{L}_t\|_F}{\|\hat{L}_t\|_F}$$

where $C^d_{\hat{L}_t,P} = \frac{1}{|\alpha|}\frac{\|\hat{L}_t\|_F}{\|P\|_F}\left(3 + 2k\frac{\|A\|_F}{\sigma_{\min}(\tilde{R})}\right)$.

The right minimal indices of $\hat{L}_t(\lambda) + \Delta\hat{L}_t(\lambda)$ are those of $P(\lambda) + \Delta P(\lambda)$ shifted by $k-1$ and left minimal indices of $\hat{L}_t(\lambda) + \Delta\hat{L}_t(\lambda)$ are the same as those of $P(\lambda) +$

$\Delta P(\lambda)$. *This coincides with the corresponding relationship between the minimal indices of* $\hat{L}_t(\lambda)$ *and* $P(\lambda)$.

Proof. Clearly, $\|\Delta \hat{L}_t\|_F < \frac{\sigma_{\min}(\tilde{R})}{2k^{3/2}} \Rightarrow \|\Delta B\|_F < \frac{\sigma_{\min}(\tilde{R})}{2k^{3/2}}$. By Theorem 3.1.8, there exists $\Delta D(\lambda)$ of grade $k-1$ such that $B(\lambda)+\Delta B(\lambda)$ and $\Lambda_{k,n}(\lambda)^T + \Delta D(\lambda)^T$ are dual minimal bases with all the row degrees 1 and $k-1$ respectively. Therefore $\hat{L}_t(\lambda)+\Delta\hat{L}_t(\lambda)$ is a strong block minimal bases pencil and Theorem 2.3.5 implies that $\hat{L}_t(\lambda)+\Delta\hat{L}_t(\lambda)$ is a strong block minimal bases linearization of

$$\frac{1}{\alpha}(A(\lambda)+\Delta A(\lambda))(\Lambda_{k,n}(\lambda)+\Delta D(\lambda)) =: P(\lambda)+\Delta P(\lambda)$$

of grade k. As $P(\lambda)=\frac{1}{\alpha}A(\lambda)(\Lambda_{k,n}(\lambda))$, we have,

$$\Delta P(\lambda) = \frac{1}{\alpha}\{(A(\lambda)+\Delta A(\lambda))\Delta D(\lambda) + \Delta A(\lambda)\Lambda_{k,n}(\lambda)\}.$$

Therefore, by applying Lemma 3.0.1 and Theorem 3.1.8 (b),

$$\begin{aligned}
\|\Delta P\|_F &\leqslant \frac{1}{|\alpha|}\{\|A(\Delta D)\|_F + \|(\Delta A)(\Delta D)\|_F + \|(\Delta A)\Lambda_{k,n}\|_F\} \\
&\leqslant \frac{1}{|\alpha|}\{\sqrt{2}\|A\|_F\|\Delta D\|_F + \sqrt{2}\|\Delta A\|_F\|(\Delta D)\|_F + \sqrt{2}\|\Delta A\|_F\} \\
&\leqslant \frac{1}{|\alpha|}\left\{\sqrt{2}\|A\|_F \frac{k\sqrt{2}}{\sigma_{\min}(\tilde{R})}\|\Delta B\|_F + 3\|\Delta A\|_F\right\} \\
&\leqslant \frac{1}{|\alpha|}\left(3+2k\frac{\|A\|_F}{\sigma_{\min}(\tilde{R})}\right)\|\Delta\hat{L}_t\|_F.
\end{aligned}$$

This implies that, $\frac{\|\Delta P\|_F}{\|P\|_F} \leqslant \frac{1}{|\alpha|}\frac{\|\hat{L}_t\|_F}{\|P\|_F}\left(3+2k\frac{\|A\|_F}{\sigma_{\min}(\tilde{R})}\right)\frac{\|\Delta\hat{L}_t\|_F}{\|\hat{L}_t\|_F}$. Also as the block $\hat{B}(\lambda)$ is absent in the linearization $\hat{L}_t(\lambda)+\Delta\hat{L}_t(\lambda)$, we have $\hat{C}(\lambda) = I_m$ in (2.3.8) and consequently by Theorem 2.3.5, the right minimal indices of $\hat{L}_t(\lambda)+\Delta\hat{L}_t(\lambda)$ are those of $P(\lambda)+\Delta P(\lambda)$ shifted by $k-1$ and left minimal indices of $\hat{L}_t(\lambda)+\Delta\hat{L}_t(\lambda)$ are the same as those of $P(\lambda)+\Delta P(\lambda)$. By Theorem 2.3.6, the shifting relations between the left and right minimal indices of $P(\lambda)+\Delta P(\lambda)$ and $\hat{L}_t(\lambda)+\Delta\hat{L}_t(\lambda)$ are exactly the same as those between $P(\lambda)$ and $\hat{L}_t(\lambda)$. □

Now we extend the above analysis to solutions of the complete eigenvalue problem for $P(\lambda)$ obtained via any linearization $L_t(\lambda)$ arising from a g-linearization in $\mathbb{L}_1(P)$. Using this fact that any such linearization $L_t(\lambda)$ is strictly equivalent to a linearization $\hat{L}_t(\lambda)$ of the form (2.3.4), and the results for $\hat{L}_t(\lambda)$, we have the following theorem.

Theorem 3.1.10. *Let $L_t(\lambda)$ be any linearization of $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i \in \mathbb{C}[\lambda]^{m\times n}$ of grade k with $m \geqslant n$, arising from a g-linearization in $\mathbb{L}_1(P)$. Let $\hat{L}_t(\lambda) = \tilde{D}^{-1}L_t(\lambda)$ where $\tilde{D}$ is as given in (2.3.5). Then $\hat{L}_t(\lambda)$ is of the form (2.3.4). Let $A(\lambda)$ and $B(\lambda)$ be the blocks of $\hat{L}_t(\lambda)$ as specified by (3.1.1) and (3.1.2) respectively and $\tilde{R}$ be the nonsingular upper triangular matrix appearing in the block $B(\lambda)$. If $\Delta L_t(\lambda)$ is any pencil of the same size as $L_t(\lambda)$ such that*

$$\|\Delta L_t\|_F < \frac{\sigma_{\min}(\tilde{R})\sigma_{\min}(\tilde{D})}{2k^{3/2}}, \tag{3.1.12}$$

then $L_t(\lambda) + \Delta L_t(\lambda)$ is a strong linearization of a matrix polynomial $P(\lambda) + \Delta P(\lambda)$ of grade k such that

$$\frac{\|\Delta P\|_F}{\|P\|_F} \leqslant C^d_{L_t,P} \frac{\|\Delta L_t\|_F}{\|L_t\|_F}, \tag{3.1.13}$$

where $C^d_{L_t,P} = \frac{\kappa_2(\tilde{D})}{|\alpha|}\frac{\|\hat{L}_t\|_F}{\|P\|_F}\left(3 + 2k\frac{\|A\|_F}{\sigma_{\min}(\tilde{R})}\right)$, $\kappa_2(\tilde{D})$ being the 2-norm condition number of $\tilde{D}$. The right minimal indices of $L_t(\lambda) + \Delta L_t(\lambda)$ are those of $P(\lambda) + \Delta P(\lambda)$ shifted by $k-1$ and left minimal indices of $L_t(\lambda) + \Delta L_t(\lambda)$ are same as those of $P(\lambda) + \Delta P(\lambda)$. This coincides with the corresponding relationship between the minimal indices of $L_t(\lambda)$ and $P(\lambda)$.

Proof. Evidently, $\hat{L}_t(\lambda) = \tilde{D}^{-1}L_t(\lambda)$ is of the form (2.3.4). Using (3.1.12) and the fact that

$$L_t(\lambda) + \Delta L_t(\lambda) = \tilde{D}(\hat{L}_t(\lambda) + \tilde{D}^{-1}\Delta L_t(\lambda)),$$

we have,

$$\|\tilde{D}^{-1}\Delta L_t\|_F < \frac{\sigma_{\min}(\tilde{R})}{2k^{3/2}}.$$

Therefore by Theorem 3.1.9, $\hat{L}_t(\lambda) + \tilde{D}^{-1}\Delta L_t(\lambda)$ is a strong block minimal bases linearization of some polynomial $P(\lambda) + \Delta P(\lambda)$ of grade k such that

$$\frac{\|\Delta P\|_F}{\|P\|_F} \leqslant \frac{1}{|\alpha|}\frac{\|\hat{L}_t\|_F}{\|P\|_F}\left(3 + 2k\frac{\|A\|_F}{\sigma_{\min}(\tilde{R})}\right)\frac{\|\tilde{D}^{-1}\Delta L_t\|_F}{\|\hat{L}_t\|_F}.$$

Using the relations, $\|\tilde{D}^{-1}\Delta L_t\|_F \leqslant \|\tilde{D}^{-1}\|_2\|\Delta L_t\|_F$ and $\|L_t\|_F \leqslant \|\tilde{D}\|_2\|\hat{L}_t\|_F$, we get,

$$\begin{aligned}\frac{\|\Delta P\|_F}{\|P\|_F} &\leqslant \frac{1}{|\alpha|}\frac{\|\hat{L}_t\|_F}{\|P\|_F}\left(3 + 2k\frac{\|A\|_F}{\sigma_{\min}(\tilde{R})}\right)\frac{\|\tilde{D}^{-1}\|_2\|\Delta L_t\|_F\|\tilde{D}\|_2}{\|L_t\|_F} \\ &= \frac{\kappa_2(\tilde{D})}{|\alpha|}\frac{\|\hat{L}_t\|_F}{\|P\|_F}\left(3 + 2k\frac{\|A\|_F}{\sigma_{\min}(\tilde{R})}\right)\frac{\|\Delta L_t\|_F}{\|L_t\|_F}.\end{aligned}$$

As $L_t(\lambda) + \Delta L_t(\lambda) = \tilde{D}(\hat{L}_t(\lambda) + \tilde{D}^{-1}\Delta L_t(\lambda))$, the pencils $L_t(\lambda) + \Delta L_t(\lambda)$ and $\hat{L}_t(\lambda) + \tilde{D}^{-1}\Delta L_t(\lambda)$ are strictly equivalent. Hence $L_t(\lambda) + \Delta L_t(\lambda)$ is a strong linearization of the polynomial $P(\lambda) + \Delta P(\lambda)$ and the rules for recovering the minimal indices of $P(\lambda) + \Delta P(\lambda)$ from those of $L_t(\lambda) + \Delta L_t(\lambda)$ are the same as the ones applied to $\hat{L}_t(\lambda) + \tilde{D}^{-1}\Delta L_t(\lambda)$. Therefore it follows from Theorem 3.1.9, that the right minimal indices of $L_t(\lambda) + \Delta L_t(\lambda)$ are those of $P(\lambda) + \Delta P(\lambda)$ shifted by $k-1$ and left minimal indices of $L_t(\lambda) + \Delta L_t(\lambda)$ are same as those of $P(\lambda) + \Delta P(\lambda)$. □

If the complete eigenvalue problem for $L_t(\lambda)$ is solved by using a backward stable algorithm, then $\frac{\|\Delta L_t\|_F}{\|L_t\|_F} = O(\mathbf{u})$. In such a situation (3.1.13) shows that the process of solving the complete eigenvalue problem for $P(\lambda)$ via linearizations $L_t(\lambda)$, is globally backward stable if $C^d_{L_t,P}$ is not very large. As $C^d_{L_t,P} = \kappa_2(\tilde{D})C^d_{\hat{L}_t,P}$, so a good choice of $L_t(\lambda)$ would be one for which $\kappa_2(\tilde{D}) \approxeq 1$ and $C^d_{\hat{L}_t,P}$ is not large for the corresponding pencil $\hat{L}_t(\lambda) = \tilde{D}^{-1}L_t(\lambda)$. To identify such linearizations, we first note that for the block $A(\lambda)$ of $\hat{L}_t(\lambda)$,

$$\alpha P(\lambda) = A(\lambda)\Lambda_{k,n}(\lambda) \Rightarrow |\alpha|\, \|P\|_F = \|A\Lambda_{k,n}\|_F \leqslant \sqrt{2}\|A\|_F.$$

This implies that

$$\frac{\|\hat{L}_t\|_F}{\|P\|_F} \geqslant \frac{\|A\|_F}{\|P\|_F} \geqslant \frac{|\alpha|}{\sqrt{2}}.$$

Now if $|\alpha|\, \|P\|_F \gg \sigma_{\min}(\tilde{R})$ then $\frac{\sqrt{2}\|A\|_F}{\sigma_{\min}(\tilde{R})} \gg 1$ and since $\frac{\|\hat{L}_t\|_F}{|\alpha|\|P\|_F} \geqslant \frac{1}{\sqrt{2}}$, so $C^d_{\hat{L}_t,P}$ is big. Again if $|\alpha|\, \|P\|_F \ll \sigma_{\min}(\tilde{R})$ then $\frac{\|\hat{L}_t\|_F}{|\alpha|\|P\|_F} > \frac{\|\tilde{R}\|_F}{|\alpha|\|P\|_F} \gg 1$ and once again $C^d_{\hat{L}_t,P}$ is big. So, a good choice of $L_t(\lambda)$ would be one for which $|\alpha|\, \|P\|_F \approxeq \sigma_{\min}(\tilde{R})$.

Besides, if $\|A\|_F \approxeq |\alpha|\, \|P\|_F \approxeq \sigma_{\min}(\tilde{R})$, and $\kappa_2(\tilde{R}) \approxeq 1$ then

$$C^d_{L_t,P} \approxeq (3+2k)\sqrt{1+2(k-1)n}$$

and then

$$\frac{\|\Delta P\|_F}{\|P\|_F} \lessapprox (3+2k)\sqrt{1+2(k-1)n}\frac{\|\Delta L_t\|_F}{\|L_t\|_F}.$$

In summary, by using linearizations $L_t(\lambda)$ satisfying

(i) $\kappa_2(\tilde{D}) \approxeq 1$ and $\kappa_2(\tilde{R}) \approxeq 1$ and

(ii) $\|A\|_F \approxeq |\alpha|\, \|P\|_F \approxeq \sigma_{\min}(\tilde{R})$,

we will have $\frac{\|\Delta P\|_F}{\|P\|_F} = O(\mathbf{u})$ if $\frac{\|\Delta L_t\|_F}{\|L_t\|_F} = O(\mathbf{u})$. So the complete eigenvalue problem for $P(\lambda)$ can be solved in a globally backward stable manner by using backward stable algorithms to solve the complete eigenvalue problem for such choices of $L_t(\lambda)$.

The optimal block Kronecker linearizations of the form $\begin{bmatrix} A(\lambda) \\ B(\lambda) \end{bmatrix}$ ensuring global backward stability that were identified in [22] are included in the above choices. In fact they are the ones for which $\tilde{D} = I_{m+(k-1)n}$, $|\alpha| = 1/\|P\|_F$, $\tilde{R} = I_{(k-1)n}$ and $\|X_{12}\|_F^2 + \|Y_{11}\|_F^2 \approxeq \frac{1}{\|P\|_F^2}\sum_{i=1}^{k-1}\|A_i\|_F^2$ in (2.3.9) and include the Frobenius companion form $C_1(\lambda)$. Our analysis shows that there exist many more choices of linearizations among the pencils $L_t(\lambda)$ with which the complete eigenvalue problem for $P(\lambda)$ can be solved in a globally backward stable manner.

3.2 Comparison of the brief and detailed analyses

Optimal linearizations for global backward stability which can be identified using Corollary 3.1.3 form a strict subset of the optimal linearizations found by our detailed analysis. To see this, suppose $L_t(\lambda)$ is a linearization of $P(\lambda)$ arising from a g-linearization such that $\sigma_{\min}(\tilde{R}) \leqslant 1$. If $\|\Delta L_t\|_F < \frac{\sigma_{\min}(\tilde{D})}{2k^{3/2}\max\{1, 1/\sigma_{\min}(\tilde{R})\}}$, then by Corollary 3.1.3 $L_t(\lambda) + \Delta L_t(\lambda)$ is a linearization of $P(\lambda) + \Delta P(\lambda)$ such that $\frac{\|\Delta P\|_F}{\|P\|_F} \leqslant C^b_{L_t,P}\frac{\|\Delta L_t\|_F}{\|L_t\|_F}$, where

$$\begin{aligned} C^b_{L_t,P} &= 2k\frac{\kappa_2(\tilde{D})}{|\alpha|}\frac{\|\hat{L}_t\|_F}{\|P\|_F}(1+\|A\|_F)\max\{1, 1/\sigma_{\min}(\tilde{R})\} \\ &= 2k\frac{\kappa_2(\tilde{D})}{|\alpha|}\frac{\|\hat{L}_t\|_F(1+\|A\|_F)}{\|P\|_F\sigma_{\min}(\tilde{R})}. \end{aligned} \tag{3.2.1}$$

As $\sigma_{\min}(\tilde{R}) \leqslant 1$, we also have $\|\Delta L_t\|_F < \frac{\sigma_{\min}(\tilde{R})\sigma_{\min}(\tilde{D})}{2k^{3/2}}$. Hence by Theorem 3.1.10 $L_t(\lambda) + \Delta L_t(\lambda)$ is a linearization of $P(\lambda) + \Delta P(\lambda)$ such that $\frac{\|\Delta P\|_F}{\|P\|_F} \leqslant C^d_{L_t,P}\frac{\|\Delta L_t\|_F}{\|L_t\|_F}$, where

$$C^d_{L_t,P} = \frac{\kappa_2(\tilde{D})}{|\alpha|}\frac{\|\hat{L}_t\|_F}{\|P\|_F}\left(3 + 2k\frac{\|A\|_F}{\sigma_{\min}(\tilde{R})}\right).$$

Now as $k \geqslant 2$, using (3.2.1) we have,

$$C^d_{L_t,P} < 2k\frac{\kappa_2(\tilde{D})}{|\alpha|}\frac{\|\hat{L}_t\|_F}{\|P\|_F}\left(1 + \frac{\|A\|_F}{\sigma_{\min}(\tilde{R})}\right) = \frac{C^b_{L_t,P}(\sigma_{\min}(\tilde{R}) + \|A\|_F)}{(1+\|A\|_F)} \leqslant C^b_{L_t,P}.$$

By similar arguments, we can easily show that $C^d_{L_t,P} < C^b_{L_t,P}$ when $\sigma_{\min}(\tilde{R}) \geqslant 1$. So if $C^b_{L_t,P}$ is a moderate constant then so is $C^d_{L_t,P}$ and therefore if $L_t(\lambda)$ can be used to solve the complete eigenvalue problem for $P(\lambda)$ in a globally backward stable manner on the basis of Corollary 3.1.3, then our detailed analysis also draws the same conclusion. Moreover, our detailed analysis provides a tighter bound on the backward error $\frac{\|\Delta P\|_F}{\|P\|_F}$ then the one obtained from Corollary 3.1.3.

Now suppose $L_t(\lambda)$ is a linearization of $P(\lambda)$ arising from g-linearization such that $\kappa_2(\tilde{D}) \approxeq 1, \kappa_2(\tilde{R}) \approxeq 1$ and $\|A\|_F \approxeq |\alpha|\, \|P\|_F \approxeq \sigma_{\min}(\tilde{R})$.

First consider the case that $\sigma_{\min}(\tilde{R}) \geqslant 1$. If the perturbation $\Delta L_t(\lambda)$ induced by the backward error analysis satisfies $\frac{\sigma_{\min}(\tilde{D})}{2k^{3/2}} \leqslant \|\Delta L_t\|_F < \frac{\sigma_{\min}(\tilde{R})\sigma_{\min}(\tilde{D})}{2k^{3/2}}$, then clearly our detailed analysis has an advantage. Suppose $\|\Delta L_t\|_F < \frac{\sigma_{\min}(\tilde{D})}{2k^{3/2}}$. Then $\|\Delta L_t\|_F < \frac{\sigma_{\min}(\tilde{R})\sigma_{\min}(\tilde{D})}{2k^{3/2}}$ and by Theorem 3.1.10, $\frac{\|\Delta P\|_F}{\|P\|_F} \leqslant C^d_{L_t,P}\frac{\|\Delta L_t\|_F}{\|L_t\|}$ where

$$C^d_{L_t,P} \approxeq (3+2k)\sqrt{1+2(k-1)n}.$$

As $\sigma_{\min}(\tilde{R}) \geqslant 1$, we have $\|\Delta L_t\|_F < \frac{\sigma_{\min}(\tilde{D})}{2k^{3/2}\max\{1, 1/\sigma_{\min}(\tilde{R})\}}$. Therefore by Corollary 3.1.3, $\frac{\|\Delta P\|_F}{\|P\|_F} \leqslant C^b_{L_t,P}\frac{\|\Delta L_t\|_F}{\|L_t\|_F}$, where

$$C^b_{L_t,P} = 2k\frac{\kappa_2(\tilde{D})}{|\alpha|}\frac{\sqrt{\|A\|_F^2 + 2\|\tilde{R}\|_F^2}}{\|P\|_F}(1+\|A\|_F)\max\{1, 1/\sigma_{\min}(\tilde{R})\} \approxeq 2k\sqrt{1+2(k-1)n}(1+\|A\|_F).$$

Now if $\tilde{R}$ is chosen in such a way that $\sigma_{\min}(\tilde{R}) \gg 1$ while at the same time ensuring that $\|\Delta L_t\|_F < \frac{\sigma_{\min}(\tilde{D})}{2k^{3/2}}$, then $C^b_{L_t,P}$ is not a moderate constant as $\|A\|_F \approxeq \sigma_{\min}(\tilde{R})$ is large.

Next consider the case that $\sigma_{\min}(\tilde{R}) \leqslant 1$. If $\|\Delta L_t\|_F < \frac{\sigma_{\min}(\tilde{R})\sigma_{\min}(\tilde{D})}{2k^{3/2}}$, then once again $\frac{\|\Delta P\|_F}{\|P\|_F} \leqslant C^d_{L_t,P}\frac{\|\Delta L_t\|_F}{\|L_t\|_F}$ where $C^d_{L_t,P} \approxeq (3+2k)\sqrt{1+2(k-1)n}$. Moreover, we also have $\|\Delta L_t\|_F < \frac{\sigma_{\min}(\tilde{D})}{2k^{3/2}\max\{1, 1/\sigma_{\min}(\tilde{R})\}}$. By Corollary 3.1.3, $\frac{\|\Delta P\|_F}{\|P\|_F} \leqslant C^b_{L_t,P}\frac{\|\Delta L_t\|_F}{\|L_t\|_F}$, where

$$C^b_{L_t,P} = 2k\frac{\kappa_2(\tilde{D})}{|\alpha|}\frac{\sqrt{\|A\|_F^2 + 2\|\tilde{R}\|_F^2}}{\|P\|_F}\frac{(1+\|A\|_F)}{\sigma_{\min}(\tilde{R})} \approxeq 2k\sqrt{1+2(k-1)n}\left(\frac{1}{\sigma_{\min}(\tilde{R})}+1\right).$$

Now $C^b_{L_t,P}$ is not a moderate constant if $\sigma_{\min}(\tilde{R}) \ll 1$.

Hence in both the cases $C^d_{L_t,P}$ is a moderate constant while $C^b_{L_t,P}$ is a large number. For such choices of linearizations $L_t(\lambda)$, it can be said on the basis of our detailed analysis that the complete eigenvalue problem for $P(\lambda)$ can be solved in a globally backward stable manner. But the same cannot be said about them on the basis of Corollary 3.1.3.

Conclusion. In this chapter we have carried out a global backward error analysis of the process of solving the complete eigenvalue problem for a rectangular matrix polynomial $P(\lambda)$ via linearizations extracted from strong g-linearizations in the spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ introduced in Chapter 2. It shows that these g-linearizations provide a wide choice of linearizations that can solve the complete eigenvalue problem for $P(\lambda)$ in a globally backward stable manner.

4 Nearest matrix polynomials with a specified elementary divisor

Given an matrix polynomial $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ of degree k where A_i, $i = 0, \ldots, k$ are $n \times n$ real or complex matrices, this chapter investigates the distance from $P(\lambda)$ to a nearest matrix polynomial with an elementary divisor $(\lambda - \lambda_0)^j, j \geqslant r$, for a given $\lambda_0 \in \mathbb{C}$ and integer $r \geqslant 2$.

Although the problem is considered only for finite values of λ_0, the analysis also covers the infinite case which is equivalent to the reversal polynomial $\text{rev}_k P(\lambda)$ having an elementary divisor $\lambda^j, j \geqslant r$. In particular in such cases the distance under consideration is important from the point of view of control theory for the following reasons. If $P(\lambda) = \lambda A_1 - A_0$ is a regular matrix pencil with index greater than one, then this is equivalent to the existence of a Jordan chain of length at least 2 at ∞ for $P(\lambda)$, or equivalently an elementary divisor $\lambda^j, j \geqslant 2$, for $\text{rev}_1 P(\lambda)$. In such a case the associated differential algebraic equation $A_1 \dot{x}(t) = A_0 x(t) + Bu(t)$, may not have any solution for certain choices of initial conditions unless the controller $u(t)$ is sufficiently smooth. In fact the larger the length of a Jordan chain at ∞, the greater are the smoothness requirements on $u(t)$. In particular, for dynamical systems arising from matrix pencils as above to be stable or asymptotically stable, it is necessary that the matrix pencil has index at most one. Moreover, for the stability of such systems it is necessary that the purely imaginary eigenvalues of $P(\lambda)$ are not associated with Jordan chains of length 2 or more. For more details see [13, 23, 79] and references therein.

It is well known that arbitrarily small perturbations to matrix pencils with λ_0 as an eigenvalue of algebraic multiplicity r can result in a matrix pencil having an elementary divisor $(\lambda - \lambda_0)^r$. In fact this result can also be extended to all matrix polynomials a proof of which is provided in Theorem 4.1.2. Due to this fact, the distance problem under consideration is also equivalent to finding the distance to a nearest matrix polynomial with an eigenvalue at λ_0 with algebraic multiplicity at least r. This problem has been considered in the literature in various forms. The distance to a nearest matrix polynomial with a prescribed multiple eigenvalue is considered in [66] and bounds on the distance are obtained under certain conditions. In [67] this work is extended to find the distance from a given matrix polynomial to a nearest matrix polynomial with a specified eigenvalue of algebraic multiplicity at least r and a Jordan chain of length at most k and an upper bound of the distance to a nearest matrix polynomial with a specified eigenvalue of algebraic multiplicity at least r. The latter is done by constructing a perturbation to the given matrix polynomial which has the desired feature. However, the construction is possible under certain conditions. The results are extended to matrix polynomials in [51] where a similar construction is made to find an upper bound on the distance to a nearest matrix polynomial with specified eigenvalues of desired multiplicities. The distance from an $n \times m$ matrix pencil $A + \lambda B$ with $n \geqslant m$, to a nearest matrix pencil having specified eigenvalues such that the sum of their multiplicities is at least r is considered in [53]. Under the assumption that rank $B \geqslant r$, and only A is perturbed, the distance is shown to be given by a certain singular value optimization under certain conditions. These ideas are extended in [50] to find the same distance from a square matrix polynomial that has no infinite eigenvalues. Under certain conditions similar to those in [53], a singular value optimization is shown to be equal to the distance when only the constant coefficient of the matrix polynomial is perturbed. A lower bound is found for the general case when all coefficient matrices are perturbed. The techniques are further extended to find the same distance for more general nonlinear eigenvalue problems in [49].

The analysis of the distance problem in this chapter has several key features. Firstly the stated distance is considered for a square matrix polynomial that is either regular or singular and perturbations are considered on all the coefficient matrices of the polynomial. Note that with the exception of [53] where a rectangular matrix pencil is considered, in all other works in the literature the matrix pencil or polynomial is assumed to be regular. However, [53] considers perturbations only to the constant coefficient matrix of the pencil. In fact by using elementary perturbation theoretic arguments it is established that if the matrix polynomial $P(\lambda)$ is singular, then it is arbitrarily close to a regular matrix polynomial with an elementary divisor $(\lambda - \lambda_0)^j, j \geqslant r$. This makes it possible to assume that the matrix polynomial $P(\lambda)$

is regular in the rest of the chapter. A necessary and sufficient condition is obtained for $P(\lambda)$ to have λ_0 as an eigenvalue of algebraic multiplicity at least r. Due to this it is possible to show that finding the stated distance is equivalent to finding a structure preserving perturbation such that the nullity of a certain block Toeplitz matrix is at least r. This leads to a lower bound on the distance and allows for several characterizations of the distance in terms of optimization problems. Under the mild assumption that λ_0 is not an eigenvalue of $P(\lambda)$, for different choices of norms it is established that computing the distance from $P(\lambda)$ to a nearest matrix polynomial with an elementary divisor $(\lambda-\lambda_0)^j, j \geqslant r$, is equivalent to computing a generalized version of a structured singular value or μ-value. As μ-value computation is an NP hard problem, these results are likely to throw light on the computational complexity of the distance problem. The characterization in terms of generalized μ-values also yields a lower bound on the distance. Alternatively, the distance is characterized by another optimization problem which is computed via Broyden Fletcher Goldfarb Shanno (BFGS) and MATLAB's `globalsearch` algorithms. This also results in an upper bound on the distance. Special cases for which the solution of the distance problem has a closed form expression are also discussed. Finally, computed values of the distance via BFGS and MATLAB's globalsearch are compared with upper and lower bounds.

The normwise distance of $P(\lambda)$ to the set of all matrix polynomials having a elementary divisor $(\lambda-\lambda_0)^j$ where $j \geqslant r$ will be considered with respect to the following norms.

$$\delta_s(P,\lambda_0,r) := \inf\left\{\|\!|\Delta P|\!\|_s | (P+\Delta P)(\lambda) \text{ has an elementary divisor } (\lambda-\lambda_0)^j, j \geqslant r\right\},$$

$$\delta_{2,\infty}(P,\lambda_0,r) := \inf\left\{\|\!|\Delta P|\!\|_{2,\infty} | (P+\Delta P)(\lambda) \text{ has an elementary divisor } (\lambda-\lambda_0)^j, j \geqslant r\right\},$$

where $\|\!|P|\!\|_s = \|[A_0 \cdots A_k]\|_s$, $s = 2, F$ and $\|\!|P|\!\|_{2,\infty} = \max_{i\in\{0,\ldots,k\}}\{\|A_i\|_2\}$. Also the matrix polynomial $\Delta P(\lambda) = \sum_{i=0}^k \lambda^i \Delta A_i$ is such that any of the coefficient matrices ΔA_i may be zero.

Due to the importance of the case $\lambda_0 = 0$ in practical applications and also because the results for this case involve expressions that are relatively simpler than the general case, in many instances, initially the results are obtained for this special case and then extended for other choices of λ_0. The following lemma will be useful for making these extensions.

Lemma 4.0.1. *Given any $n \times n$ matrix polynomial $Q(\lambda) = \sum_{i=0}^k \lambda^i B_i$, and $\lambda_0 \in \mathbb{C}$,*

$$\begin{bmatrix} Q(\lambda_0) & Q'(\lambda_0) & \cdots & \frac{1}{p!}Q^p(\lambda_0) \end{bmatrix} = \begin{bmatrix} B_0 & \cdots & B_k \end{bmatrix} M(\lambda_0; r)$$

where $p = \min\{r, k\}$ *and* $M(\lambda_0; r)$ *is a* $(k+1)n \times (p+1)n$ *matrix given by* $H(\lambda_0) \otimes I_n$, $H(\lambda_0)$ *being a* $(k+1) \times (p+1)$ *matrix with* (i, j) *entry equal to* $\frac{1}{(j-1)!} \frac{d^{j-1}\lambda^{i-1}}{d\lambda^{j-1}}\Big|_{\lambda=\lambda_0}$.

Proof. The proof follows from the fact that for each $i = 1, \ldots, p+1$, $\frac{Q^{(i-1)}(\lambda_0)}{(i-1)!}$ is given by the product of $\begin{bmatrix} B_0 & \cdots & B_k \end{bmatrix}$ with $H_i(\lambda_0) \otimes I_n$, $H_i(\lambda_0)$ being the i-th column of $H(\lambda_0)$. □

4.1 Polynomials for which the distance is zero

Given a matrix polynomial $P(\lambda)$, it is interesting to identify cases when the distance under consideration is zero. One such situation is obviously the case that $k = 1$ and λ_0 is an eigenvalue of $P(\lambda)$ of multiplicity at least r. The main result of this section is a proof of the fact that if λ_0 is an eigenvalue of $P(\lambda)$ of multiplicity at least r for any value of k, or $P(\lambda)$ is a singular matrix polynomial, then the distance under consideration is zero. The following theorem proves this for matrix pencils which is later generalized to matrix polynomials. For the sake of completeness, the result that the distance is zero if λ_0 is an eigenvalue of the pencil of multiplicity at least r is also included in the statement of the theorem. Also note that although the theorem is proved with respect to the norm $|||\cdot|||_F$, clearly it also holds for all other choices of norms.

Theorem 4.1.1. *For a given* $n \times n$ *matrix pencil* $L(\lambda) = A + \lambda E$ *and a positive integer* $r \leqslant n$, *if*

(a) $L(\lambda)$ *is regular and algebraic multiplicity of* λ_0 *as an eigenvalue of* $L(\lambda)$ *is greater than or equal to* r, *or*

(b) $L(\lambda)$ *is singular,*

then it is arbitrarily close to a regular pencil having an elementary divisor $(\lambda - \lambda_0)^j$ *where* $j \geqslant r$.

Proof.

(a) The proof of this part is obvious owing to the structure of the Kronecker canonical form of a pencil having λ_0 as an eigenvalue of algebraic multiplicity at least r.

(b) Let $L(\lambda)$ be a singular pencil and $\epsilon > 0$ be arbitrarily chosen. Without loss of generality it may be assumed that $L(\lambda)$ is in Kronecker canonical form, i.e.,

$$L(\lambda) = \left[\begin{array}{c|c|c} R_f(\lambda) & & \\ \hline & R_{inf}(\lambda) & \\ \hline & & S(\lambda) \end{array}\right]$$

where $R_f(\lambda)$ and $R_{inf}(\lambda)$ represents the regular part of $L(\lambda)$ corresponding to finite and infinite eigenvalues respectively and $S(\lambda)$ represents the singular part. The idea of the proof is to construct a pencil $\Delta L(\lambda)$ such that $\|\Delta L\|_F < \epsilon$ and $(L+\Delta L)(\lambda)$ is a regular matrix pencil with 0 as an eigenvalue of algebraic multiplicity n. Then by part (a), $(L+\Delta L)(\lambda)$ is arbitrarily close to having an elementary divisor λ^n where clearly $n \geqslant r$. The arguments are then extended to the case that $\lambda_0 \neq 0$.

Initially it is assumed that all three types of blocks are present in $L(\lambda)$. Let the sizes of $R_f(\lambda)$, $R_{inf}(\lambda)$ and $S(\lambda)$ be $n_1 \times n_1$, $n_2 \times n_2$ and $n_3 \times n_3$ respectively such that $n_1 + n_2 + n_3 = n$. Also choose any $\tilde{\epsilon} \in (0, \epsilon)$.

The block $R_f(\lambda)$ is bidiagonal with super-diagonal entries 0 or -1. Construct an $n_1 \times n_1$ pencil $\Delta R_f(\lambda)$ such that all the sub-diagonal entries are $\tilde{\epsilon}\lambda$ and all other entries are 0.

The block $R_{inf}(\lambda)$ is also bidiagonal with super-diagonal entries λ or 0. Construct an $n_2 \times n_2$ pencil $\Delta R_{inf}(\lambda)$ by replacing the super-diagonal entries λ and 0 of $R_{inf}(\lambda)$ by 0 and $\tilde{\epsilon}\lambda$ respectively and setting all other entries to 0.

The number of right and left singular blocks in the singular part $S(\lambda)$ are equal. Without loss of generality assume that right singular block and left singular blocks in $S(\lambda)$ appear alternatively so that $S(\lambda)$ can be considered block diagonal with square diagonal blocks formed by placing one right and one left singular block next to each other. Each such diagonal block of $S(\lambda)$ has exactly one row and one column independent of λ and all other rows and columns have exactly one entry as λ. Assuming that there are μ blocks in $S(\lambda)$, suppose that rows $i_1, \ldots, i_\mu$ and columns $j_1, \ldots, j_\mu$ of $S(\lambda)$ are independent of λ. Construct an $n_3 \times n_3$ block diagonal pencil $\Delta S(\lambda)$ with square blocks on the diagonal of the same size as the diagonal blocks of $S(\lambda)$ such that the entries $(i_2, j_2), \ldots, (i_\mu, j_\mu)$ are $\tilde{\epsilon}\lambda$ and all other entries are 0.

Set $\Delta L(\lambda) = \left[\begin{array}{c|c|c} \Delta R_f(\lambda) & \tilde{\epsilon}\lambda & \\ \hline & \Delta R_{inf}(\lambda) & \tilde{\epsilon}\lambda \\ \hline \tilde{\epsilon}\lambda & & \Delta S(\lambda) \end{array}\right]$, where $\tilde{\epsilon}\lambda$ has been placed in the positions $(1, n_1+1)$, $(n_1+n_2, n_1+n_2+j_1)$ and $(n_1+n_2+i_1, n_1)$ of $\Delta L(\lambda)$. Choose $\tilde{\epsilon}$ small enough so that $\|\Delta L\|_F < \epsilon$. Now,

$$(L+\Delta L)(\lambda) = \left[\begin{array}{c|c|c} \hat{R}_f(\lambda) & \tilde{\epsilon}\lambda & \\ \hline & \hat{R}_{inf}(\lambda) & \tilde{\epsilon}\lambda \\ \hline \tilde{\epsilon}\lambda & & \hat{S}(\lambda) \end{array}\right],$$

where

$$\hat{R}_f(\lambda) := R_f(\lambda) + \Delta R_f(\lambda) = \begin{bmatrix} \lambda - \lambda_1 & * & & \\ \tilde{\epsilon}\lambda & \ddots & \ddots & \\ & \ddots & \lambda - \lambda_{n_1-1} & * \\ & & \tilde{\epsilon}\lambda & \lambda - \lambda_{n_1} \end{bmatrix},$$

with $*$ representing either 0 or -1, and

$$\hat{R}_{inf}(\lambda) = R_{inf}(\lambda) + \Delta R_{inf}(\lambda) = \begin{bmatrix} -1 & \star & & \\ & \ddots & \ddots & \\ & & -1 & \star \\ & & & -1 \end{bmatrix},$$

with $\star$ representing λ or $\tilde{\epsilon}\lambda$ and $\hat{S}(\lambda) = S(\lambda) + \Delta S(\lambda)$. For $1 \leqslant \tilde{i}_1, \tilde{i}_2, \tilde{j}_1, \tilde{j}_2 \leqslant n$, let $F[\tilde{i}_1, \tilde{i}_2; \tilde{j}_1, \tilde{j}_2](\lambda)$ denote the determinant of the submatrix of $L(\lambda) + \Delta L(\lambda)$ obtained by deleting rows $\tilde{i}_1, \tilde{i}_2$ and columns $\tilde{j}_1, \tilde{j}_2$.

The determinant $\det(L(\lambda) + \Delta L(\lambda))$ is evaluated by first expanding along row $n_1 + n_2$ which has two nonzero entries, viz., $\tilde{\epsilon}\lambda$ at position $n_1 + n_2 + j_1$ and -1 at position $n_1 + n_2$ and then expanding along column n_1 which has at most three nonzero entries, viz., $\tilde{\epsilon}\lambda$ at position $n_1 + n_2 + i_1$, $\lambda - \lambda_{n_1}$ at position n_1 and -1 (if $* = -1$) at position $n_1 - 1$. Setting $\tilde{i}_1 = n_1 + n_2$, $\tilde{j}_1 = n_1 + n_2 + j_1$, $\tilde{i}_2 = n_1 + n_2 + i_1$, $\tilde{j}_2 = \tilde{i}_3 = n_1$, $\tilde{j}_3 = \tilde{i}_1$, and $\tilde{i}_4 = n_1 - 1$,

$$\begin{aligned}\det((L + \Delta L)(\lambda)) = &\pm (\tilde{\epsilon}\lambda)^2 F[\tilde{i}_1, \tilde{i}_2; \tilde{j}_1, \tilde{j}_2](\lambda) \pm \tilde{\epsilon}\lambda(\lambda - \lambda_{n_1}) F[\tilde{i}_1, \tilde{i}_3; \tilde{j}_1, \tilde{j}_2](\lambda) \\ &\pm \tilde{\epsilon} c \lambda F[\tilde{i}_1, \tilde{i}_4; \tilde{j}_1, \tilde{i}_3](\lambda) \pm \tilde{\epsilon}\lambda F[\tilde{i}_1, \tilde{i}_2; \tilde{j}_3, \tilde{j}_2](\lambda) \\ &\pm (\lambda - \lambda_{n_1}) F[\tilde{i}_1, \tilde{i}_3; \tilde{j}_3, \tilde{j}_2](\lambda) \pm c F[\tilde{i}_1, \tilde{i}_4; \tilde{j}_3, \tilde{j}_2](\lambda)\end{aligned}$$

where $c = 0$ or $c = -1$.

As $\hat{S}(\lambda)$ and its submatrices obtained by removing row i_1 or column j_1 are all singular pencils, the last five terms in the right hand side of the above equation are zero. Now $F[\tilde{i}_1, \tilde{i}_2; \tilde{j}_1, \tilde{j}_2](\lambda)$ is the product of the determinants of two matrices, viz., $\hat{S}(\lambda)$ with rows i_1 and j_1 removed and the submatrix of

$$\hat{R}(\lambda) := \left[\begin{array}{c|c} \hat{R}_f(\lambda) & \tilde{\epsilon}\lambda \\ \hline & \hat{R}_{inf}(\lambda) \end{array}\right]$$

obtained by removing column n_1 and the last row. Due to the manner of constructing $\hat{S}(\lambda)$, the determinant of $\hat{S}(\lambda)$ with rows i_1 and j_1 removed is given by $\pm\tilde{\epsilon}^{p_1}\lambda^{n_3 - 1}$ for some positive integer p_1. Let $G[\tilde{i}_1, \tilde{i}_2; \tilde{j}_1, \tilde{j}_2](\lambda)$ denote the determinant of $\hat{R}(\lambda)$ with rows $\tilde{i}_1, \tilde{i}_2$ and columns $\tilde{j}_1$ and $\tilde{j}_2$ removed. Then the determinant of submatrix of $\hat{R}(\lambda)$ with the last row and column n_1 removed is given by

$$\pm(\tilde{\epsilon}\lambda) G[n_1 + n_2, 1; n_1, n_1 + 1](\lambda) \pm G[n_1 + n_2, n_1 + 1; n_1, n_1 + 1](\lambda)$$

when expanded along column $n_1 + 1$. But $G[n_1 + n_2, n_1 + 1; n_1, n_1 + 1](\lambda) = 0$ as the submatrix obtained by eliminating the indicated rows of $\hat{R}(\lambda)$ is singular. Therefore,

$$\det((L + \Delta L)(\lambda)) = \pm(\tilde{\epsilon}\lambda)^2 \left(\tilde{\epsilon}^{p_1}\lambda^{n_3 - 1}\right) \left(\tilde{\epsilon}\lambda G[n_1 + n_2, 1; n_1, n_1 + 1](\lambda)\right), \qquad (4.1.1)$$

where

$$G[n_1+n_2,1;n_1,n_1+1](\lambda) = \det\left(\left[\begin{array}{ccccc|cccc} \tilde{\epsilon}\lambda & \lambda-\lambda_2 & * & & & & & & \\ & \ddots & \ddots & * & & & & & \\ & & \tilde{\epsilon}\lambda & \lambda-\lambda_{n_1-1} & & & & & \\ & & & \tilde{\epsilon}\lambda & & & & & \\ \hline & & & & & \star & & & \\ & & & & & -1 & \ddots & & \\ & & & & & & \ddots & \star & \\ & & & & & & & -1 & \star \end{array}\right]\right)$$

$$= \tilde{\epsilon}^{p_2}\lambda^{n_1+n_2-2}, \tag{4.1.2}$$

for some positive integer p_2. From (4.1.1) and (4.1.2),

$$\det((L+\Delta L)(\lambda)) = \pm\tilde{\epsilon}^{p_1+p_2+3}\lambda^{n_1+n_2+n_3} = \pm\tilde{\epsilon}^{p}\lambda^{n},$$

for $p = p_1+p_2+3$. This establishes that $L(\lambda)+\Delta L(\lambda)$ is a regular matrix pencil with 0 an eigenvalue of algebraic multiplicity n.

The cases where all three types of blocks $R_f(\lambda)$, $R_{inf}(\lambda)$ and $S(\lambda)$ are not present in $L(\lambda)$ are dealt with in the following manner.

- If the block $R_f(\lambda)$ is not present then $n_1 = 0$. We construct $\Delta L(\lambda)$ as

$$\Delta L(\lambda) = \left[\begin{array}{c|c} \Delta R_{inf}(\lambda) & \\ & \tilde{\epsilon}\lambda \\ \hline \tilde{\epsilon}\lambda & \\ & \Delta S(\lambda) \end{array}\right],$$

where the (n_2, n_2+j_1) and $(n_2+i_1, 1)$ entries of $\Delta L(\lambda)$ are $\tilde{\epsilon}\lambda$ and the construction of $\Delta R_{inf}(\lambda)$ and $\Delta S(\lambda)$ are the same as earlier. Now,

$$(L+\Delta L)(\lambda) = \left[\begin{array}{c|c} \hat{R}_{inf}(\lambda) & \\ & \tilde{\epsilon}\lambda \\ \hline \tilde{\epsilon}\lambda & \\ & \hat{S}(\lambda) \end{array}\right],$$

where

$$\hat{R}_{inf}(\lambda) = R_{inf}(\lambda) + \Delta R_{inf}(\lambda) = \begin{bmatrix} -1 & \star & & \\ & \ddots & \ddots & \\ & & -1 & \star \\ & & & -1 \end{bmatrix},$$

$\star$ being λ or $\tilde{\epsilon}\lambda$ and $\hat{S}(\lambda) = S(\lambda) + \Delta S(\lambda)$. For $1 \leqslant \tilde{i}_1, \tilde{i}_2, \tilde{j}_1, \tilde{j}_2 \leqslant n$, let $F[\tilde{i}_1, \tilde{i}_2; \tilde{j}_1, \tilde{j}_2](\lambda)$ denote the determinant of the submatrix of $L(\lambda) + \Delta L(\lambda)$ obtained by deleting rows $\tilde{i}_1, \tilde{i}_2$ and columns $\tilde{j}_1, \tilde{j}_2$. To find $\det((L+\Delta L)(\lambda))$ we first expand it along row n_2 which has two nonzero entries, viz., $\tilde{\epsilon}\lambda$ at position $n_2 + j_1$ and -1 at position n_2 and then expanding along column 1 which has 2 nonzero entries, viz., $\tilde{\epsilon}\lambda$ at position $n_2 + i_1$ and -1 at position 1. So,

$$\begin{aligned} \det((L+\Delta L)(\lambda)) =& \pm(\tilde{\epsilon}\lambda)^2 F[n_2, n_2+i_1; 1, n_2+j_1](\lambda) \pm \tilde{\epsilon}\lambda F[1, n_2; 1, n_2+j_1](\lambda) \\ & \pm \tilde{\epsilon}\lambda F[n_2, n_2+i_1; 1, n_2](\lambda) \pm F[1, n_2; 1, n_2](\lambda). \end{aligned}$$

As $\hat{S}(\lambda)$ and its submatrices obtained by removing row i_1 or column j_1 are all singular pencils, the last three terms in the right hand side of the above equation are zero. Now $F[n_2, n_2+i_1; 1, n_2+j_1](\lambda)$ is the product of the determinants of two matrices, viz., $\hat{S}(\lambda)$ with rows i_1 and j_1 removed and the matrix

$$\begin{bmatrix} \star & & & \\ -1 & \star & & \\ & \ddots & \ddots & \\ & & -1 & \star \end{bmatrix}$$

which has determinant $\tilde{\epsilon}^{p_2}\lambda^{n_2-1}$ for some non-negative integer p_2. Due to the manner of constructing $\hat{S}(\lambda)$, the determinant of $\hat{S}(\lambda)$ with rows i_1 and j_1 removed is given by $\pm\tilde{\epsilon}^{p_1}\lambda^{n_3-1}$ for some positive integer p_1. Hence

$$\det((L+\Delta L)(\lambda)) = \pm\tilde{\epsilon}^{p_1+p_2+2}\lambda^{n_2+n_3} = \pm\tilde{\epsilon}^{p}\lambda^{n},$$

for $p = p_1 + p_2 + 2$. This proves that $(L+\Delta L)(\lambda)$ is a regular matrix pencil with 0 as an eigenvalue of algebraic multiplicity n.

- If the block $R_{inf}(\lambda)$ is not present then $n_2 = 0$ and we construct $\Delta L(\lambda)$ as

$$\Delta L(\lambda) = \left[\begin{array}{c|c} \Delta R_f(\lambda) & \tilde{\epsilon}\lambda \\ \hline \tilde{\epsilon}\lambda & \Delta S(\lambda) \end{array}\right],$$

where the (n_1+i_1, n_1) and $(1, n_1+j_1)$ entries of $\Delta L(\lambda)$ are $\tilde{\epsilon}\lambda$. The construction of $\Delta R_f(\lambda)$ and $\Delta S(\lambda)$ are the same as above. Now,

$$(L + \Delta L)(\lambda) = \left[\begin{array}{c|c} \hat{R}_f(\lambda) & \tilde{\epsilon}\lambda \\ \hline \tilde{\epsilon}\lambda & \hat{S}(\lambda) \end{array}\right],$$

where

$$\hat{R}_f(\lambda) = R_f(\lambda) + \Delta R_f(\lambda) = \begin{bmatrix} \lambda - \lambda_1 & * & & \\ \tilde{\epsilon}\lambda & \ddots & \ddots & \\ & \ddots & \lambda - \lambda_{n_1-1} & * \\ & & \tilde{\epsilon}\lambda & \lambda - \lambda_{n_1} \end{bmatrix},$$

$*$ being 0 or -1 and $\hat{S}(\lambda) = S(\lambda) + \Delta S(\lambda)$. For $1 \leqslant \tilde{i}_1, \tilde{i}_2, \tilde{j}_1, \tilde{j}_2 \leqslant n$, let $F[\tilde{i}_1, \tilde{i}_2; \tilde{j}_1, \tilde{j}_2](\lambda)$ be the determinant of the submatrix of $L(\lambda)+\Delta L(\lambda)$ obtained by deleting rows $\tilde{i}_1, \tilde{i}_2$ and columns $\tilde{j}_1, \tilde{j}_2$. To find $\det((L + \Delta L)(\lambda))$ we first expand it along row 1 which has at most three nonzero entries, viz., $\tilde{\epsilon}\lambda$ at position $n_1 + j_1$ and -1 (if $* = -1$) at position 2 and $\lambda - \lambda_1$ at position 1 and then expand along column n_1 which has at most 3 nonzero entries, viz., $\tilde{\epsilon}\lambda$ at position $n_1 + i_1$, $\lambda - \lambda_{n_1}$ at position n_1 and -1 (if $* = -1$) at position $n_1 - 1$. So,

$$\begin{aligned} &\det((L + \Delta L)(\lambda)) \\ =&\pm(\tilde{\epsilon}\lambda)^2 F[1, n_1 + i_1; n_1, n_1 + j_1](\lambda) \pm \tilde{\epsilon}\lambda(\lambda - \lambda_{n_1})F[1, n_1; n_1, n_1 + j_1](\lambda) \\ &\pm c\tilde{\epsilon}\lambda F[1, n_1 - 1; n_1, n_1 + j_1](\lambda) \pm c\tilde{\epsilon}\lambda F[1, n_1 + i_1; 2, n_1](\lambda) \end{aligned}$$

$$\pm c(\lambda-\lambda_{n_1})F[1,n_1;2,n_1](\lambda) \pm c^2 F[1,n_1-1;2,n_1](\lambda)$$
$$\pm\tilde{\epsilon}\lambda(\lambda-\lambda_1)F[1,n_1+i_1;1,n_1](\lambda) \pm (\lambda-\lambda_{n_1})(\lambda-\lambda_1)F[1,n_1;1,n_1](\lambda)$$
$$\pm c(\lambda-\lambda_1)F[1,n_1-1;1,n_1](\lambda),$$

where $c = 0$ or $c = -1$.

As $\hat{S}(\lambda)$ and its submatrices obtained by removing row i_1 or column j_1 are all singular pencils, the last eight terms in the right hand side of the above equation are zero. Now $F[1, n_1 + i_1; n_1, n_1 + j_1](\lambda)$ is the product of the determinants of two matrices, viz., $\hat{S}(\lambda)$ with rows i_1 and j_1 removed and the matrix

$$\begin{bmatrix} \tilde{\epsilon}\lambda & \lambda-\lambda_2 & * & & \\ & \tilde{\epsilon}\lambda & \ddots & \ddots & \\ & & \ddots & \ddots & * \\ & & & \ddots & \lambda-\lambda_{n_1-1} \\ & & & & \tilde{\epsilon}\lambda \end{bmatrix}$$

which has determinant $\tilde{\epsilon}^{n_1-1}\lambda^{n_1-1}$. Due to the manner of constructing $\hat{S}(\lambda)$, the determinant of $\hat{S}(\lambda)$ with rows i_1 and j_1 removed is given by $\pm\tilde{\epsilon}^{p_1}\lambda^{n_3-1}$ for some positive integer p_1. Hence

$$\det((L+\Delta L)(\lambda)) = \pm\tilde{\epsilon}^{p_1+n_1+1}\lambda^{n_1+n_3} = \pm\tilde{\epsilon}^{p}\lambda^{n},$$

for $p = p_1 + n_1 + 1$. This proves that $(L + \Delta L)(\lambda)$ is a regular matrix pencil with 0 an eigenvalue of algebraic multiplicity n.

- If only the singular blocks occur in $L(\lambda)$ then $n_1 = n_2 = 0$ and we construct $\Delta L(\lambda)$ as a block diagonal matrix with blocks of the same size as $S(\lambda)$ such that the (i_1, j_1), ..., (i_μ, j_μ) entries are $\tilde{\epsilon}\lambda$ and all other entries are 0. Then clearly

  $$\det((L+\Delta L)(\lambda)) = \pm\tilde{\epsilon}^{\mu}\lambda^{n}.$$

  This proves that $(L + \Delta L)(\lambda)$ is a regular matrix pencil with 0 an eigenvalue of algebraic multiplicity n.

In each case the above arguments may be extended to show that by choosing $\tilde{\epsilon}$ small enough, $\|\Delta L\|_F < \epsilon$ and $(L + \Delta L)(\lambda)$ is a regular pencil whose determinant is a scalar multiple of λ^n.

Now suppose $\lambda_0 \neq 0$. The pencil $L(\lambda)$ may be written in the form

$$L(\lambda) = A + \lambda_0 E + (\lambda - \lambda_0)E.$$

Setting $\hat{A} = A + \lambda_0 E$ and $\hat{\lambda} = \lambda - \lambda_0$ and arguing as above, for a given $\epsilon > 0$ there exists $\Delta E \in \mathbb{C}^{n\times n}$ satisfying $\|\Delta E\|_F < \epsilon$ such that $\hat{A} + \hat{\lambda}(E + \Delta E)$ is a regular pencil and $\det(\hat{A} + \hat{\lambda}(E + \Delta E)) = \pm\tilde{\epsilon}^p \hat{\lambda}^n$ for some $\tilde{\epsilon} \in (0, \epsilon)$. This implies that

$$\det((A - \lambda_0 \Delta E) + \lambda(E + \Delta E)) = \pm\tilde{\epsilon}^p(\lambda - \lambda_0)^n.$$

Therefore for $\Delta L(\lambda) := -(\lambda_0 \Delta E) + \lambda \Delta E$, λ_0 is an eigenvalue of $(L + \Delta L)(\lambda)$ of algebraic multiplicity $n \geqslant r$. The proof now follows from the fact that $\tilde{\epsilon}$ may be chosen small enough so that $\|\Delta L\|_F = \sqrt{|\lambda_0|^2 + 1}\|\Delta E\|_F < \epsilon$. □

Note that Theorem 4.1.1(b) can be proved by using parts (iii), (iv) and (vi) of [28, Theorem III]. In particular, by repeated application of parts (iii) and/or (iv), it may be shown that there exists a singular pencil with only singular blocks in its Kronecker canonical form that is arbitrarily close to $L(\lambda)$. Part (vi) of the same theorem then implies that this singular pencil is arbitrarily close to a regular pencil with a Jordan chain of length n at λ_0. But here we give a proof using elementary perturbation techniques, which is a bit long but it is more accessible to a wider readership as it is not couched in the language of versal deformations.

The above result may be extended to matrix polynomials by considering the first companion linearization $C_1(\lambda)$ of $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ as in (1.2.2). It is an example of a block Kronecker linearization as introduced in [22] where it was shown that if $L(\lambda)$ is a block Kronecker linearization of $P(\lambda)$ and $\Delta L(\lambda)$ is a pencil of the same size as $L(\lambda)$ with $\|\Delta L\|_F < \epsilon$, for some sufficiently small $\epsilon > 0$, then $(L + \Delta L)(\lambda)$ is a strong linearization of $(P + \Delta P)(\lambda)$ such that $\|\Delta P\|_F < C\epsilon$ for some positive constant C. Due to this result the following theorem is an immediate consequence of Theorem 4.1.1.

Theorem 4.1.2. *For a given $n \times n$ matrix polynomial $P(\lambda)$ of degree k, and a positive integer $r \leqslant kn$, if*

(a) $P(\lambda)$ is regular and λ_0 is an eigenvalue of algebraic multiplicity greater than or equal to r, or

(b) $P(\lambda)$ is singular,

then $P(\lambda)$ is arbitrarily close to a regular matrix polynomial having an elementary divisor $(\lambda - \lambda_0)^j$ where $j \geqslant r$.

Note that Theorem 4.1.2 is also implied by the work done in [21] which uses the above mentioned result in [22] to construct the closure hierarchy graphs of matrix polynomials from those of their first companion linearizations.

In fact, by using the arguments in the proof of Theorem 4.1.1, it is clear that any $n \times n$ singular matrix polynomial $P(\lambda)$ of degree k is arbitrarily close to a regular matrix polynomial having an elementary divisor $(\lambda - \lambda_0)^{kn}$. In view of the above theorem, it is now possible to assume without loss of generality that the distances $\delta_F(P, \lambda_0, r)$, $\delta_2(P, \lambda_0, r)$ and $\delta_{2,\infty}(P, \lambda_0, r)$ are being computed for a regular matrix polynomial $P(\lambda)$ which does not have λ_0 as an eigenvalue of algebraic multiplicity r. This also has the effect of removing the uncertainty that was earlier associated with the situation that perturbations being made to the matrix polynomial for the desired objectives could result in a singular matrix polynomial.

4.2 A characterization via block Toeplitz matrices

One of the aims of this chapter is to show that computing the distance to a nearest matrix polynomial with an elementary divisor $(\lambda - \lambda_0)^j, j \geqslant r$, for appropriate choices of norms is equivalent to finding a structured singular value or generalized μ-value. The next result is an important step in this direction. Since the expression for the optimization is more aesthetic if r is replaced by $r + 1$, in the rest of the chapter the distance is considered in the form $\delta_F(P, \lambda_0, r + 1)$, $\delta_2(P, \lambda_0, r + 1)$, and $\delta_{2,\infty}(P, \lambda_0, r + 1)$. We introduce a set that will be frequently used. Consider

$$\Gamma := \{[\gamma_1 \cdots \gamma_r] : \gamma_i > 0, 1 \leqslant i \leqslant r\}. \tag{4.2.1}$$

For any $\gamma := \begin{bmatrix} \gamma_1 & \cdots & \gamma_r \end{bmatrix} \in \Gamma$ and $\alpha \in \mathbb{C}$, let $T_\gamma(Q, \alpha, r)$ be a function from the set of all $n \times n$ matrix polynomial $Q(\lambda) = \sum_{i=0}^{k} \lambda^i B_i$, to the set of $(r+1)n \times (r+1)n$ matrices defined by

$$T_\gamma(Q, \alpha, r) := \begin{bmatrix} Q(\alpha) & & & & \\ \gamma_1 Q'(\alpha) & Q(\alpha) & & & \\ \gamma_1\gamma_2 \frac{Q''(\alpha)}{2!} & \gamma_2 Q'(\alpha) & Q(\alpha) & & \\ \vdots & \vdots & \ddots & \ddots & \\ \left(\prod_{i=1}^{r}\gamma_i\right)\frac{Q^r(\alpha)}{r!} & \left(\prod_{i=2}^{r}\gamma_i\right)\frac{Q^{r-1}(\alpha)}{(r-1)!} & \cdots & \gamma_r Q'(\alpha) & Q(\alpha) \end{bmatrix}. \tag{4.2.2}$$

When $r = 1$, we consider $\gamma \in \Gamma$ where Γ is the set of all positive real numbers.

Theorem 4.2.1. *A scalar $\lambda_0 \in \mathbb{C}$ is an eigenvalue of a $n \times n$ matrix polynomial $Q(\lambda)$ of algebraic multiplicity at least $r+1$ if and only if the rank of $T_\gamma(Q, \lambda_0, r)$ as defined by (4.2.2) is at most $(r+1)(n-1)$.*

Proof. Let λ_0 be an eigenvalue of $Q(\lambda)$ of multiplicity $r+1$. Suppose there are p Jordan chains $\{x_{11}, x_{12}, \ldots, x_{1k_1}\}$, $\{x_{21}, x_{22}, \ldots, x_{2k_2}\}$, $\ldots$, $\{x_{p1}, x_{p2}, \ldots, x_{pk_p}\}$ of $Q(\lambda)$ corresponding to λ_0 where $x_{ij} \in \mathbb{C}^n$ for $i = 1, \ldots, p$, $j = 1, \ldots, k_i$, satisfying $\sum_{i=1}^{p} k_i \geqslant r+1$. The i^{th} Jordan chain satisfies the following equations for $i = 1, \ldots, p$.

$$
\begin{aligned}
&Q(\lambda_0)x_{i1} = 0\\
&Q(\lambda_0)x_{i2} + Q'(\lambda_0)x_{i1} = 0\\
&Q(\lambda_0)x_{i3} + Q'(\lambda_0)x_{i2} + \frac{Q''(\lambda_0)}{2!}x_{i1} = 0\\
&\vdots\\
&Q(\lambda_0)x_{ik_i} + Q'(\lambda_0)x_{i(k_i-1)} + \frac{Q''(\lambda_0)}{2!}x_{i(k_i-2)} + \cdots + \frac{Q^{k_i-1}(\lambda_0)}{(k_i-1)!}x_{i1} = 0.
\end{aligned}
$$

It may be assumed that $\{x_{11}, x_{21}, \ldots, x_{p1}\}$ is a linear independent set. Clearly the i^{th} Jordan chain contributes the following k_i vectors

$$
\begin{bmatrix} 0 \\ \vdots \\ 0 \\ \frac{x_{i1}}{\gamma_{(r-k_i+2)}\cdots\gamma_r} \\ \vdots \\ \frac{x_{i(k_i-3)}}{\gamma_{r-2}\gamma_{r-1}\gamma_r} \\ \frac{x_{i(k_i-2)}}{\gamma_{r-1}\gamma_r} \\ \frac{x_{i(k_i-1)}}{\gamma_r} \\ x_{ik_i} \end{bmatrix},
\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \frac{x_{i1}}{\gamma_{(r-k_i+3)}\cdots\gamma_r} \\ \vdots \\ \frac{x_{i(k_i-3)}}{\gamma_{r-1}\gamma_r} \\ \frac{x_{i(k_i-2)}}{\gamma_r} \\ x_{i(k_i-1)} \end{bmatrix},
\begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \frac{x_{i1}}{\gamma_{(r-k_1+4)}\cdots\gamma_r} \\ \vdots \\ \frac{x_{i(k_i-3)}}{\gamma_r} \\ x_{i(k_i-2)} \end{bmatrix}, \ldots,
\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ x_{i1} \end{bmatrix}
$$

of length$(r+1)n$ to the null space $N(T_\gamma(Q, \lambda_0, r))$ of $T_\gamma(Q, \lambda_0, r)$ for $i = 1, \ldots, p$. All the above vectors are linearly independent as $\{x_{11}, x_{21}, \ldots, x_{p1}\}$ are linearly independent. Hence the nullity of $T_\gamma(Q, \lambda_0, r)$ is at least $r+1$.

Conversely suppose that $\operatorname{rank}(T_\gamma(Q, \lambda_0, r)) \leqslant (r+1)(n-1)$ so that the nullity of $T_\gamma(Q, \lambda_0, r)$ is at least $r+1$. Let $\{x_1, x_2, \ldots, x_{r+1}\}$ be a linearly independent ordered list in $N(T_\gamma(Q,\lambda_0,r))$, where

$$x_j = \begin{bmatrix} x_{r+1,j}^T & \cdots & x_{i,j}^T & \cdots & x_{2,j}^T & x_{1,j}^T \end{bmatrix}^T$$

with $x_{i,j} \in \mathbb{C}^n$ for $i,j \in \{1,\ldots,r+1\}$. If $x_{r+1,j} \neq 0$ for some j, then $\{x_{r+1,j},\ldots,x_{1,j}\}$ will be a Jordan chain of $Q(\lambda)$ of length $r+1$ corresponding to λ_0 and the proof follows. Therefore without loss of generality it may be assumed that for each $j = 1,\ldots,r+1$, there exists $t_j, 0 < t_j < r+1$ such that $x_{ij} = 0$ for all $i = t_j + 1,\ldots,r+1$. Setting $t = \max_{1\leqslant j\leqslant r+1} t_j$ and $p = t - \min_{1\leqslant j\leqslant r+1} t_j$ and reordering the list if necessary, it may be assumed that the first $k_1 + \cdots + k_s$ vectors of the list satisfy $x_{ij} = 0$ for all $s = 1,\ldots,p+1$ and $i = t-s+2,\ldots,r+1$, so that $k_1 + \cdots k_{p+1} = r+1$. Note that k_j may be zero for some or all $j = 2,\ldots,p+1$. Consider $X = \begin{bmatrix} x_1 & x_2 & \cdots & x_{r+1} \end{bmatrix}$. Then in fact,

$$X = \begin{bmatrix}
0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
x_{t,1} & \cdots & x_{t,k_1} & 0 & \cdots & 0 & & 0 & \cdots & 0 \\
\cdots & \cdots & \cdots & x_{t-1,k_1+1} & \cdots & x_{t-1,k_1+k_2} & \cdots & \cdots & \cdots & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & x_{t-p,r+2-k_{p+1}} & \cdots & x_{t-p,r+1} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
x_{1,1} & \cdots & x_{1,k_1} & x_{1,k_1+1} & \cdots & x_{1,k_1+k_2} & \cdots & x_{1,r+2-k_{p+1}} & \cdots & x_{1,r+1}
\end{bmatrix}.$$

It is possible that some of the vectors $x_{t-s+1,1+\sum_{j=1}^{s-1} k_j},\ldots,x_{t-s+1,\sum_{j=1}^{s} k_j}$ in the consecutive columns $1+\sum_{j=1}^{s-1} k_j$ to $\sum_{j=1}^{s} k_j$ of X can be made zero for each $s = 1,\ldots,p+1$, via elementary column operations on X that affect only those columns. The columns of the transformed X will also be a linearly independent list in $N(T_\gamma(Q,\lambda_0,r))$. Assume without loss of generality that X has been formed after such transformations have already been made and the submatrices

$$\begin{bmatrix} x_{t-s+1,1+\sum_{j=1}^{s-1} k_j} & \cdots & x_{t-s+1,\sum_{j=1}^{s} k_j} \end{bmatrix}$$

of X have full column rank for each $s = 1,\ldots,p+1$.

Clearly the columns of X are linearly independent and belong to $N(T_\gamma(Q,\lambda_0,r))$. Also the first k_1 columns give rise to Jordan chains of $Q(\lambda)$ corresponding to λ_0 of length t, the next k_2 columns give rise to Jordan chains of of length $t-1$ and so on, with the last k_{p+1} columns forming Jordan chains of length $t-p$. Due to the structure of X, $\beta_s := \{x_{t-s+1,k_1+\cdots+k_{s-1}+1}, \ldots, x_{t-s+1,k_1+\cdots+k_s}\}$, $s = 1, \ldots, p+1$, are ordered lists of linearly independent vectors. Consider the list

$$\beta := \beta_{p+1} = \{x_{t-p,k_1+\cdots+k_p+1}, \ldots, x_{t-p,k_1+\cdots+k_{p+1}}\}.$$

If the first vector of the list β_p, does not belong to $\operatorname{span}(\beta)$, it is included in β. If it belongs to $\operatorname{span}(\beta)$ then it can be uniquely represented by a linear combination of the vectors of β and at least one of the scalar coefficients in the representation is nonzero. Replace one of the vectors from β whose associated coefficient in the linear combination is nonzero by the first vector of β_p. Now consider the second vector of the list β_p. If it does not belong to the span of the updated β, then include it in β. Otherwise it is a linear combination of the vectors of β with at least one of the scalar coefficients in the linear combination being non zero. As β_p is a linearly independent list, a vector associated with such a non zero scalar in the linear combination can be chosen from β_{p+1}. Update the set β by replacing this vector by the second vector from β_p. This process is continued for the rest of the vectors in β_p as well as those of $\beta_{p-1}, \ldots, \beta_1$. The final β clearly forms a linearly independent list of eigenvectors of $P(\lambda)$ corresponding to λ_0. Moreover the sums of the lengths of the Jordan chains associated with these eigenvectors is at least $tk_1 + \sum_{s=1}^{p}(t-s)(k_{s+1}-k_s)$. But

$$tk_1 + \sum_{s=1}^{p}(t-s)(k_{s+1}-k_s) = \sum_{s=1}^{p} k_s + (t-p)k_{p+1} \geqslant \sum_{s=1}^{p+1} k_s = r+1.$$

Hence λ_0 is an eigenvalue of $Q(\lambda)$ of algebraic multiplicity at least $r+1$ and the proof follows. □

Note that the part of Theorem 4.2.1, showing that *if* $\operatorname{rank}(T_\gamma(Q,\lambda_0,r)) \leqslant (r+1)(n-1)$ *then* λ_0 *is an eigenvalue of* $Q(\lambda)$ *of algebraic multiplicity at least* $r+1$ can be proved by using a result from [77], where it was shown that the algebraic multiplicity $a(Q,\lambda_0)$ of an eigenvalue λ_0 of an $m \times n$ full rank matrix polynomial $Q(\lambda) = \sum_{i=0}^{k} B_i(\lambda-\lambda_0)^i$ with $m \leqslant n$, is given by

$$a(Q,\lambda_0) = ml - \operatorname{rank} \mathcal{T}_{l-1}(Q). \tag{4.2.3}$$

Here $l \geqslant g$, g being the length of the largest Jordan chain corresponding to λ_0 and

$$\mathcal{T}_q(Q) = \begin{bmatrix} B_q & \cdots & B_0 \\ \vdots & \iddots & \\ B_0 & & \end{bmatrix}, \; q = 0, 1, \ldots.$$

A similar result is stated in [76] without restrictions on the rank and size of $Q(\lambda)$.

However the other part of Theorem 4.2.1 does not follow from (4.2.3) unless an additional assumption is made. If λ_0 is an eigenvalue of $Q(\lambda)$ of algebraic multiplicity greater than or equal to $r+1$, then proving that $\operatorname{rank}(T_\gamma(Q,\lambda_0,r)) \leqslant (r+1)(n-1)$, is equivalent to proving that $n(r+1) - \operatorname{rank}(\mathcal{T}_r(Q)) \geqslant r+1$. If the Jordan chains associated with λ_0 are not longer than $r+1$, then the proof is immediate due to (4.2.3). However this may not always hold as r can be any positive integer less than kn in Theorem 4.2.1. Therefore this part of the proof of Theorem 4.2.1 does not follow directly from the results in [76, 77] unless it is further assumed that the length of the Jordan chain of $Q(\lambda)$ associated with λ_0 is at most $r+1$.

Remark 4.2.2. *Theorem 4.2.1 is established in [66] for the particular case that $Q(\lambda)$ has an eigenvalue of multiplicity* 2*. Under the assumption that the leading coefficient matrix is of full rank, another characterization of a matrix polynomial $Q(\lambda)$ having a specified eigenvalue of multiplicity r is obtained in [50] via a different block Toeplitz matrix that involves $r(r+1)/2$ parameters.*

In view of part (a) of Theorem 4.1.2, the following corollary of Theorem 4.2.1 is immediate.

Corollary 4.2.3. *Given any $n\times n$ matrix polynomial $P(\lambda)$ consider the collection $\mathcal{S}(P,\lambda_0)$ of all $n\times n$ matrix polynomials $\Delta P(\lambda) = \sum_{i=0}^{k} \lambda^i \Delta A_i$ such that the block Toeplitz matrices $T_\gamma(P+\Delta P,\lambda_0,r)$ as defined in (4.2.2) with $\gamma = [1\cdots 1]$, have rank at most $(r+1)(n-1)$. For any choice of norm $|||\cdot|||$, the distance to a nearest matrix polynomial with an elementary divisor $(\lambda-\lambda_0)^j, j \geqslant r+1$, is given by* $\inf\{|||\Delta P||| : \Delta P(\lambda) \in \mathcal{S}(P,\lambda_0)\}$.

4.3 The distance as the reciprocal of a generalized μ value

Corollary 4.2.3 implies that for any given choice of norm, finding the distance from $P(\lambda)$ to a nearest matrix polynomial with an elementary divisor λ^{r+1} is equivalent to finding the smallest structure preserving perturbation to the block Toeplitz matrix

$$\begin{bmatrix} P(0) & & & & \\ P'(0) & P(0) & & & \\ \frac{1}{2!}P''(0) & P'(0) & P(0) & & \\ \vdots & \vdots & \ddots & \ddots & \\ \frac{1}{r!}P^{r}(0) & \frac{1}{(r-1)!}P^{r-1}(0) & \cdots & P'(0) & P(0) \end{bmatrix}$$

so that the rank of the perturbed matrix is at most $(r+1)(n-1)$. This fact is used in this section to show that if $\lambda_0 \in \mathbb{C}$ is not already an eigenvalue of $P(\lambda)$, then computing the distance from $P(\lambda)$ to a nearest matrix polynomial with the desired elementary divisor $(\lambda - \lambda_0)^j, j \geqslant r+1$, with respect to the norms $\|\!|\cdot|\!\|_F$, $\|\!|\cdot|\!\|_2$ and $\|\!|\cdot|\!\|_{2,\infty}$ is the reciprocal of a generalized notion of a μ-value as defined in (1.6.2).

Let

$$T(Q, \lambda_0, r) := T_\gamma(Q, \lambda_0, r) \text{ with } \gamma = [1 \cdots 1]. \tag{4.3.1}$$

The following lemma provides a useful factorization of $T(Q, \lambda_0, r)$. The proof of the lemma follows from direct multiplication of the stated factors and is therefore skipped.

Lemma 4.3.1. *For a given positive integer r and an $n \times n$ matrix polynomial $Q(\lambda)$ of degree k,*

$$T(Q, \lambda_0, r) = \left(I_{r+1} \otimes \begin{bmatrix} Q(\lambda_0) & Q'(\lambda_0) & \cdots & \frac{Q^{\min\{r,k\}}(\lambda_0)}{\min\{r,k\}!} \end{bmatrix} \right) E$$

where

$$E = \begin{cases} \begin{bmatrix} E_1^T & E_2^T & \cdots & E_{r+1}^T \end{bmatrix}^T \otimes I_n & \text{if } r \leqslant k \\ \begin{bmatrix} \tilde{E}_1^T & \tilde{E}_2^T & \cdots & \tilde{E}_{r+1}^T \end{bmatrix}^T \otimes I_n & \text{if } r > k \end{cases}$$

such that E_i, $i = 1, \ldots, r+1$ are the $(r+1) \times (r+1)$ matrices,

$$E_1 = \begin{bmatrix} 1 & & \\ & & \\ & & \\ & & \end{bmatrix}, E_2 = \begin{bmatrix} & 1 & \\ 1 & & \\ & & \\ & & \end{bmatrix}, \ldots, E_r = \begin{bmatrix} & & 1 & \\ & \iddots & & \\ 1 & & & \\ & & & \end{bmatrix}, E_{r+1} = \begin{bmatrix} & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & \end{bmatrix}$$

and $\tilde{E}_i \in \mathbb{C}^{(k+1)\times(r+1)}$ are the first $k+1$ rows of E_i.

The next theorem is the main result of this section.

Theorem 4.3.2. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $n \times n$ matrix polynomial of degree k, and $1 \leqslant r < kn$. For $\Delta A_i \in \mathbb{C}^{n\times n}, i = 0, \ldots, k$, let S_1, S_2, S_3 and S_4 be the perturbation classes of all perturbations of the type*

$$I_{r+1} \otimes \begin{bmatrix} \Delta A_0 & \Delta A_1 & \cdots & \Delta A_k \end{bmatrix},$$

$$I_{r+1} \otimes \begin{bmatrix} \Delta A_0 & \Delta A_1 & \cdots & \Delta A_{\min\{r,k\}} \end{bmatrix},$$

$$I_{r+1} \otimes \begin{bmatrix} \Delta A_0 & & \\ & \ddots & \\ & & \Delta A_k \end{bmatrix} \quad \textit{and}$$

$$I_{r+1} \otimes \begin{bmatrix} \Delta A_0 & & \\ & \ddots & \\ & & \Delta A_{\min\{r,k\}} \end{bmatrix}$$

respectively. For any $\lambda_0 \in \mathbb{C}$ which is not an eigenvalue of $P(\lambda)$, let $T(P, \lambda_0, r)$ be defined by (4.3.1) and E and $M(\lambda_0; r)$ be as given in Lemma 4.3.1 and Lemma 4.0.1 respectively. Then,

$$\delta_2(P, \lambda_0, r+1) = \begin{cases} \left[\mu^{r+1}_{S_1,\|\cdot\|_2}\left((I_{r+1} \otimes M(\lambda_0; r))E\ (T(P, \lambda_0, r))^{-1}\right)\right]^{-1} & \textit{if } \lambda_0 \neq 0, \\ \left[\mu^{r+1}_{S_2,\|\cdot\|_2}\left(E\ (T(P, 0, r))^{-1}\right)\right]^{-1} & \textit{otherwise,} \end{cases}$$

$$\delta_F(P, \lambda_0, r+1) = \begin{cases} \dfrac{\left[\mu^{r+1}_{S_1,\|\cdot\|_F}\left((I_{r+1} \otimes M(\lambda_0; r))E\ (T(P, \lambda_0, r))^{-1}\right)\right]^{-1}}{\sqrt{r+1}} & \textit{if } \lambda_0 \neq 0, \\ \dfrac{\left[\mu^{r+1}_{S_2,\|\cdot\|_F}\left(E\ (T(P, 0, r))^{-1}\right)\right]^{-1}}{\sqrt{r+1}} & \textit{otherwise,} \end{cases}$$

and

$$\delta_{2,\infty}(P, \lambda_0, r+1) = \begin{cases} \left[\mu^{r+1}_{S_3,\|\cdot\|_2}\left((I_{r+1} \otimes M(\lambda_0; r))E\ (T(P, \lambda_0, r))^{-1}\left(I_{r+1} \otimes \underbrace{\begin{bmatrix} I_n & \cdots & I_n \end{bmatrix}}_{k+1 \text{ block columns}}\right)\right)\right]^{-1} & \textit{if } \lambda_0 \neq 0, \\ \left[\mu^{r+1}_{S_4,\|\cdot\|_2}\left(E\ (T(P, 0, r))^{-1}\left(I_{r+1} \otimes \underbrace{\begin{bmatrix} I_n & \cdots & I_n \end{bmatrix}}_{\min\{r,k\}+1 \text{ block columns}}\right)\right)\right]^{-1} & \textit{otherwise.} \end{cases}$$

Proof. As $P(\lambda_0)$ is invertible, $-\Delta P(\lambda) \in S(P, \lambda_0)$ where $S(P, \lambda_0)$ is as in Corollary 4.2.3 if and only if

$$\text{nullity}(I - T(\Delta P, \lambda_0, r)T(P, \lambda_0, r)^{-1}) \geqslant r + 1. \tag{4.3.2}$$

Due to Lemma 4.3.1 the above relation may be written as,

$$\text{nullity}\left(I - \left(I_{r+1} \otimes \begin{bmatrix}\Delta P(\lambda_0) & \cdots & \frac{\Delta P^p(\lambda_0)}{p!}\end{bmatrix}\right) E\ (T(P, \lambda_0, r))^{-1}\right) \geqslant r + 1. \tag{4.3.3}$$

where $p = \min\{r, k\}$. By Lemma 4.0.1 this is equivalent to

$$\text{nullity}\left(I - (I_{r+1} \otimes [\Delta A_0 \cdots \Delta A_k] M(\lambda_0; r))\, E\ (T(P, \lambda_0, r))^{-1}\right) \geqslant r + 1,$$

which may also be written as

$$\text{nullity}\left(I - (I_{r+1} \otimes [\Delta A_0 \cdots \Delta A_k])\, (I_{r+1} \otimes M(\lambda_0; r))\, E\ (T(P, \lambda_0, r))^{-1}\right) \geqslant r + 1. \tag{4.3.4}$$

Then

$$\inf\{\|\Delta P\|_2 | \Delta P(\lambda) \in S(P, \lambda_0)\} = \left[\mu^{r+1}_{S_1, \|\cdot\|_2}\left((I_{r+1} \otimes M(\lambda_0; r))E\ (T(P, \lambda_0, r))^{-1}\right)\right]^{-1}.$$

Similarly,

$$\inf\{\|\Delta P\|_F | \Delta P(\lambda) \in S(P, \lambda_0)\} = \frac{\left[\mu^{r+1}_{S_1, \|\cdot\|_F}\left((I_{r+1} \otimes M(\lambda_0; r))E\ (T(P, \lambda_0, r))^{-1}\right)\right]^{-1}}{\sqrt{r+1}}.$$

Again (4.3.4) can be written as

$$\text{nullity}\left(I - \left(I_{r+1} \otimes \underbrace{\begin{bmatrix} I_n & \cdots & I_n\end{bmatrix}}_{k+1 \text{ block columns}}\right)\left(I_{r+1} \otimes \begin{bmatrix}\Delta A_0 & & \\ & \ddots & \\ & & \Delta A_k\end{bmatrix}\right)(I_{r+1} \otimes M(\lambda_0; r))\, E\ (T(P, \lambda_0, r))^{-1}\right) \geqslant r + 1.$$

This is equivalent to

$$\text{nullity}\left(I - \left(I_{r+1} \otimes \begin{bmatrix}\Delta A_0 & & \\ & \ddots & \\ & & \Delta A_k\end{bmatrix}\right)(I_{r+1} \otimes M(\lambda_0; r))\, E\ (T(P, \lambda_0, r))^{-1}\left(I_{r+1} \otimes \underbrace{\begin{bmatrix} I_n & \cdots & I_n\end{bmatrix}}_{k+1 \text{ block columns}}\right)\right) \geqslant r + 1. \tag{4.3.5}$$

Hence

$$\inf\{\|\Delta P\|_{2,\infty} | \Delta P(\lambda) \in S(P, \lambda_0)\} = \left[\mu^{r+1}_{S_3, \|\cdot\|_2}\left((I_{r+1} \otimes M(\lambda_0; r))E\ (T(P, \lambda_0, r))^{-1}\left(I_{r+1} \otimes \underbrace{\begin{bmatrix} I_n & \cdots & I_n\end{bmatrix}}_{k+1 \text{ block columns}}\right)\right)\right]^{-1}.$$

The proof for the case $\lambda_0 \neq 0$ now follows from Corollary 4.2.3.

If $\lambda_0 = 0$, then the equation (4.3.3) becomes

$$\text{nullity}\left(I - \left(I_{r+1} \otimes \begin{bmatrix} \Delta A_0 & \cdots & \Delta A_p \end{bmatrix}\right) E\,(T(P,0,r))^{-1}\right) \geqslant r+1. \tag{4.3.6}$$

Then setting $\Delta A_i = 0$ for $i = r+1, \ldots, k$ if $r < k$,

$$\inf\{\|\Delta P\|_2 | \Delta P(\lambda) \in S(P,0)\} = \left[\mu^{r+1}_{S_2,\|\cdot\|_2}\left(E\,(T(P,0,r))^{-1}\right)\right]^{-1} \text{ and}$$
$$\inf\{\|\Delta P\|_F | \Delta P(\lambda) \in S(P,0)\} = \frac{\left[\mu^{r+1}_{S_2,\|\cdot\|_F}\left(E\,(T(P,0,r))^{-1}\right)\right]^{-1}}{\sqrt{r+1}}.$$

Equation (4.3.6) can also be written as

$$\text{nullity}\left(I - \left(I_{r+1} \otimes \underbrace{\begin{bmatrix} I_n & \cdots & I_n \end{bmatrix}}_{p+1 \text{ block columns}}\right)\left(I_{r+1} \otimes \begin{bmatrix} \Delta A_0 & & \\ & \ddots & \\ & & \Delta A_p \end{bmatrix}\right) E\,(T(P,0,r))^{-1}\right) \geqslant r+1.$$

This is equivalent to

$$\text{nullity}\left(I - \left(I_{r+1} \otimes \begin{bmatrix} \Delta A_0 & & \\ & \ddots & \\ & & \Delta A_p \end{bmatrix}\right) E\,(T(P,0,r))^{-1}\left(I_{r+1} \otimes \underbrace{\begin{bmatrix} I_n & \cdots & I_n \end{bmatrix}}_{p+1 \text{ block columns}}\right)\right) \geqslant r+1. \tag{4.3.7}$$

Setting $\Delta A_i = 0$ for $i = r+1, \ldots, k$ if $r < k$,

$$\inf\{\|\Delta P\|_{2,\infty} | \Delta P(\lambda) \in S(P,0)\} = \left[\mu^{r+1}_{S_3,\|\cdot\|_2}\left(E\,(T(P,0,r))^{-1}\left(I_{r+1} \otimes \underbrace{\begin{bmatrix} I_n & \cdots & I_n \end{bmatrix}}_{p+1 \text{ block columns}}\right)\right)\right]^{-1}.$$

Hence the proof follows from Corollary 4.2.3. □

4.4 An alternative formulation of the distance as an optimization

An alternative formulation for the distance $\delta_s(P, \lambda_0, r+1)$ is obtained in this section for $s = 2$ or F.

Theorem 4.4.1. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $n \times n$ matrix polynomial of degree k. For a given integer r, such that $0 < r < kn$, consider the sets, Γ as in (4.2.1) and*

$$\mathbb{C}_0^{(r+1)n} := \{[x_0^T \cdots x_r^T]^T : x_i \in \mathbb{C}^n, i = 0, \ldots, r, x_0 \neq 0\}. \tag{4.4.1}$$

Now let $\mathbb{C}_\Gamma^{r,n}$ be the collection of all block Toeplitz like matrices X given by,

$$X = \begin{cases} \begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_r \\ & \gamma_1 x_0 & \gamma_2 x_1 & \cdots & \gamma_r x_{r-1} \\ & & \gamma_1\gamma_2 x_0 & \cdots & \gamma_{r-1}\gamma_r x_{r-2} \\ & & & \ddots & \vdots \\ & & & & \Big(\prod_{i=1}^{r} \gamma_i\Big) x_0 \end{bmatrix} & \text{if } r \leqslant k, \\ \begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_k & \cdots & x_r \\ & \gamma_1 x_0 & \gamma_2 x_1 & \cdots & \gamma_k x_{k-1} & \cdots & \gamma_r x_{r-1} \\ & & \gamma_1\gamma_2 x_0 & \cdots & \gamma_{k-1}\gamma_k x_{k-2} & \cdots & \gamma_{r-1}\gamma_r x_{r-2} \\ & & & \ddots & \vdots & \vdots & \vdots \\ & & & & \Big(\prod_{i=1}^{k} \gamma_i\Big) x_0 & \cdots & \Big(\prod_{i=r-k+1}^{r} \gamma_i\Big) x_{r-k} \end{bmatrix} & \text{otherwise,} \end{cases}$$

where $[x_0^T \cdots x_r^T]^T \in \mathbb{C}_0^{(r+1)n}$ and $[\gamma_1 \cdots \gamma_r] \in \Gamma$. Then for $s = 2$ or F,

$$\delta_s(P, 0, r+1) = \begin{cases} \inf_{X \in \mathbb{C}_\Gamma^{r,n}} \left\| \begin{bmatrix} A_0 & \cdots & A_r \end{bmatrix} XX^\dagger \right\|_s & \text{if } r \leqslant k, \\ \inf_{X \in \mathbb{C}_\Gamma^{r,n}} \left\| \begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix} XX^\dagger \right\|_s & \text{otherwise.} \end{cases} \tag{4.4.2}$$

and in general,

$$\delta_s(P, \lambda_0, r+1) = \begin{cases} \inf_{X \in \mathbb{C}_\Gamma^{r,n}} \left\| \begin{bmatrix} P(\lambda_0) & \cdots & \frac{1}{r!} P^r(\lambda_0) \end{bmatrix} X(M(\lambda_0; r)X)^\dagger \right\|_s & \text{if } r \leqslant k, \\ \inf_{X \in \mathbb{C}_\Gamma^{r,n}} \left\| \begin{bmatrix} P(\lambda_0) & \cdots & \frac{1}{k!} P^k(\lambda_0) \end{bmatrix} X(M(\lambda_0; r)X)^\dagger \right\|_s & \text{otherwise.} \end{cases} \tag{4.4.3}$$

where $M(\lambda_0; r)$ is as given in Lemma 4.0.1.

Proof. Initially consider the case that $r \leqslant k$. Let $\Delta P(\lambda) = \sum_{i=0}^{k} \lambda^i \Delta A_i$ be any $n \times n$ matrix polynomial of grade k such that $(P + \Delta P)(\lambda)$ is a regular matrix polynomial. Then $(P + \Delta P)(\lambda)$ has an elementary divisor $(\lambda - \lambda_0)^j$ where $j \geqslant r + 1$ if and only if there exists vectors $x_0, x_1, \ldots, x_r \in \mathbb{C}^n$ with $x_0 \neq 0$ and r positive scalars $\gamma_1, \ldots, \gamma_r$ such that

$$
\begin{aligned}
&(P+\Delta P)(\lambda_0)x_0 = 0\\
&(P+\Delta P)(\lambda_0)x_1 + \gamma_1(P+\Delta P)'(\lambda_0))x_0 = 0\\
&(P+\Delta P)(\lambda_0)x_2 + \gamma_2(P+\Delta P)'(\lambda_0)x_1 + \gamma_1\gamma_2\frac{(P+\Delta P)''(\lambda_0)}{2!}x_0 = 0\\
&\ldots\ldots\ldots\\
&(P+\Delta P)(\lambda_0)x_r + \gamma_r(P+\Delta P)'(\lambda_0)x_{r-1} + \cdots + \left(\prod_{i=1}^{r}\gamma_i\right)\frac{(P+\Delta P)^r(\lambda_0)}{r!}x_0 = 0.
\end{aligned}
$$

This is equivalent to

$$
\begin{bmatrix}
(P+\Delta P)(\lambda_0) & & & & \\
\gamma_1(P+\Delta P)'(\lambda_0) & (P+\Delta P)(\lambda_0) & & & \\
\gamma_1\gamma_2\frac{(P+\Delta P)''(\lambda_0)}{2!} & \gamma_2(P+\Delta P)'(\lambda_0) & (P+\Delta P)(\lambda_0) & & \\
\vdots & \vdots & & \ddots & \\
\left(\prod_{i=1}^{r}\gamma_i\right)\frac{(P+\Delta P)^r(\lambda_0)}{r!} & \left(\prod_{i=2}^{r}\gamma_i\right)\frac{(P+\Delta P)^{r-1}(\lambda_0)}{(r-1)!} & \cdots & \cdots & (P+\Delta P)(\lambda_0)
\end{bmatrix}
\begin{bmatrix} x_0\\ x_1\\ \vdots\\ x_{r-1}\\ x_r \end{bmatrix} = 0,
$$

which can be written in the form

$$
\begin{bmatrix}\Delta P(\lambda_0) & \Delta P'(\lambda_0) & \cdots & \frac{1}{r!}\Delta P^r(\lambda_0)\end{bmatrix} X = -\begin{bmatrix}P(\lambda_0) & P'(\lambda_0) & \cdots & \frac{1}{r!}P^r(\lambda_0)\end{bmatrix} X, \tag{4.4.4}
$$

where

$$
X = \begin{bmatrix}
x_0 & x_1 & x_2 & \cdots & x_r\\
 & \gamma_1 x_0 & \gamma_2 x_1 & \cdots & \gamma_r x_{r-1}\\
 & & \gamma_1\gamma_2 x_0 & \cdots & \gamma_{r-1}\gamma_r x_{r-1}\\
 & & & \ddots & \vdots\\
 & & & & \left(\prod_{i=1}^{r}\gamma_i\right)x_0
\end{bmatrix} \in \mathbb{C}_{\Gamma}^{r,n}.
$$

When $\lambda_0 = 0$, (4.4.4) takes the form

$$
\begin{bmatrix}\Delta A_0 & \cdots & \Delta A_r\end{bmatrix} X = -\begin{bmatrix}A_0 & \cdots & A_r\end{bmatrix} X. \tag{4.4.5}
$$

By Theorem 1.5.1, the minimum 2 or Frobenius norm solution of this equation is given by

$$
\begin{bmatrix}\Delta A_0 & \cdots & \Delta A_r\end{bmatrix} = -\begin{bmatrix}A_0 & \cdots & A_r\end{bmatrix} XX^{\dagger}. \tag{4.4.6}
$$

Setting $\Delta A_i = 0$ for $i = r+1,\ldots,k$ if $r < k$, $(P+\Delta P)(\lambda)$ has an elementary divisor $\lambda^j, j \geqslant r+1$. Therefore in this case, $\delta_s(P,0,r+1) = \inf_{X \in \mathbb{C}_\Gamma^{r,n}} \left\| \begin{bmatrix} A_0 & \cdots & A_r \end{bmatrix} XX^\dagger \right\|_s$, for $s = 2$ or F. When $\lambda_0 \neq 0$, Lemma 4.0.1 implies that

$$\begin{bmatrix} \Delta P(\lambda_0) & \Delta P'(\lambda_0) & \cdots & \frac{1}{r!}\Delta P^r(\lambda_0) \end{bmatrix} = \begin{bmatrix} \Delta A_0 & \cdots & \Delta A_k \end{bmatrix} M(\lambda_0; r).$$

Using this in (4.4.4), the minimum 2 or Frobenius norm solution of the resulting equation is given by

$$\begin{bmatrix} \Delta A_0 & \cdots & \Delta A_k \end{bmatrix} = -\begin{bmatrix} P(\lambda_0) & P'(\lambda_0) & \cdots & \frac{1}{r!}P^r(\lambda_0) \end{bmatrix} X(M(\lambda_0;r)X)^\dagger,$$

thus proving that if $r \leqslant k$ then for $s = 2$ or F,

$$\delta_s(P,\lambda_0,r+1) = \inf_{X \in \mathbb{C}_\Gamma^{r,n}} \left\| \begin{bmatrix} P(\lambda_0) & P'(\lambda_0) & \cdots & \frac{1}{r!}P^r(\lambda_0) \end{bmatrix} X(M(\lambda_0;r)X)^\dagger \right\|_s.$$

If $r > k$, then $(P+\Delta P)(\lambda)$ has an elementary divisor $(\lambda-\lambda_0)^j$ where $j \geqslant r+1$ if and only if there exists vectors $x_0, x_1, \ldots, x_r \in \mathbb{C}^n$ with $x_0 \neq 0$ and r positive scalars $\gamma_1, \ldots, \gamma_r$ such that

$$\begin{aligned}
&(P+\Delta P)(\lambda_0)x_0 = 0\\
&(P+\Delta P)(\lambda_0)x_1 + \gamma_1(P+\Delta P)'(\lambda_0)x_0 = 0\\
&(P+\Delta P)(\lambda_0)x_2 + \gamma_2(P+\Delta P)'(\lambda_0)x_1 + \gamma_1\gamma_2\frac{(P+\Delta P)''(\lambda_0)}{2!}x_0 = 0\\
&\ldots\ldots\ldots\\
&(P+\Delta P)(\lambda_0)x_k + \gamma_k(P+\Delta P)'(\lambda_0)x_{k-1} + \cdots + \left(\prod_{i=1}^{k}\gamma_i\right)\frac{(P+\Delta P)^k(\lambda_0)}{k!}x_0 = 0\\
&\ldots\ldots\ldots\\
&(P+\Delta P)(\lambda_0)x_r + \gamma_r(P+\Delta P)'(\lambda_0)x_{r-1} + \cdots + \left(\prod_{i=r-k+1}^{r}\gamma_i\right)\frac{(P+\Delta P)^k(\lambda_0)}{k!}x_{r-k} = 0.
\end{aligned}$$

This set of equations can be written in the form

$$\begin{bmatrix}
(P+\Delta P)(\lambda_0) & & & & \\
\gamma_1(P+\Delta P)'(\lambda_0) & (P+\Delta P)(\lambda_0) & & & \\
\vdots & \vdots & \ddots & & \\
\left(\prod_{i=1}^{k}\gamma_i\right)\frac{(P+\Delta P)^k(\lambda_0)}{k!} & & & (P+\Delta P)(\lambda_0) & \\
 & \ddots & & & \ddots \\
 & & \left(\prod_{i=r-k+1}^{r}\gamma_i\right)\frac{(P+\Delta P)^k(\lambda_0)}{k!} & \cdots & (P+\Delta P)(\lambda_0)
\end{bmatrix}
\begin{bmatrix} x_0\\ x_1\\ \vdots\\ x_k\\ \vdots\\ x_r \end{bmatrix} = 0.$$

This is equivalent to

$$\left[\Delta P(\lambda_0) \quad \Delta P'(\lambda_0) \quad \cdots \quad \tfrac{1}{k!}\Delta P^k(\lambda_0)\right] X = -\left[P(\lambda_0) \quad P'(\lambda_0) \quad \cdots \quad \tfrac{1}{k!}P^k(\lambda_0)\right] X,$$

where

$$X = \begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_k & \cdots & x_r \\ & \gamma_1 x_0 & \gamma_2 x_1 & \cdots & \gamma_k x_{k-1} & \cdots & \gamma_r x_{r-1} \\ & & \gamma_1\gamma_2 x_0 & \cdots & \gamma_{k-1}\gamma_k x_{k-2} & \cdots & \gamma_{r-1}\gamma_r x_{r-2} \\ & & & \ddots & \vdots & \vdots & \vdots \\ & & & & \left(\prod_{i=1}^{k}\gamma_i\right) x_0 & \cdots & \left(\prod_{i=r-k+1}^{r}\gamma_i\right) x_{r-k} \end{bmatrix} \in \mathbb{C}_{\Gamma}^{r,n}.$$

Therefore the proof follows by arguing as in the previous case. $\square$

Remark 4.4.2. *The parameters γ_i can all be taken to be 1 in the optimization that computes $\delta_s(P, \lambda_0, r+1)$ for $s = 2$ or F. As shall be seen in Section 4.6, this will also decrease the number of variables in the optimization. But these parameters play an important role when computing the upper bound for $\delta_s(P, \lambda_0, r+1)$. However, there is no particular advantage in choosing them to be nonzero real or complex numbers when deriving the upper bound.*

4.5 Lower bounds

The first lower bound on the distance $\delta_s(P, \lambda_0, r+1)$ where $s = 2$ or F is derived from Theorem 4.2.1.

Theorem 4.5.1. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $n \times n$ matrix polynomial of degree k and $r < kn$ be a positive integer. Let $\gamma := [\gamma_1 \cdots \gamma_r] \in \Gamma$, for Γ as in (4.2.1) and $T_\gamma(P, \lambda_0, r)$ be as in (4.2.2). For $s = 2$ or F, the distance $\delta_s(P, \lambda_0, r+1)$ to a nearest matrix polynomial having $(\lambda - \lambda_0)^j$ as an elementary divisor with $j \geqslant r+1$ satisfies*

$$\delta_s(P, \lambda_0, r+1) \geqslant \begin{cases} \sup_{\gamma\in\Gamma} \frac{f(\gamma)}{\sqrt{F_1(\gamma)}} & \text{if } r \leqslant k, \\ \sup_{\gamma\in\Gamma} \frac{f(\gamma)}{\sqrt{F_2(\gamma)}} & \text{otherwise,} \end{cases}$$

where, $f(\gamma) := \sigma_{(r+1)n-r}(T_\gamma(P, \lambda_0, r)),$

$$F_1(\gamma) := \|M(\lambda_0; r)\|_2^2 \max_{1\leqslant p\leqslant r} \left(1 + \sum_{t=1}^{p}\prod_{i=t}^{p} \gamma_{r+1-i}^2\right) \quad \textit{and}$$

$$F_2(\gamma) := \|M(\lambda_0;r)\|_2^2 \max\left\{ \max_{1\leqslant p\leqslant k}\left(1+\sum_{t=1}^{p}\prod_{i=t}^{p}\gamma_{r+1-i}^2\right), \max_{k+1\leqslant p\leqslant r}\left(1+\sum_{t=p-k+1}^{p}\prod_{i=t}^{p}\gamma_{r+1-i}^2\right)\right\},$$

$M(\lambda_0;r)$ being as defined in Lemma 4.0.1.

Proof. Consider the case that $r\leqslant k$. Let $\Delta P(\lambda)=\sum_{i=0}^{k}\lambda^i\Delta A_i$ be an $n\times n$ matrix polynomial such that $(P+\Delta P)(\lambda)$ has an elementary divisor $(\lambda-\lambda_0)^j$ with $j\geqslant r+1$. By Theorem 4.2.1, $\sigma_{(r+1)n-r}(T_\gamma(P+\Delta P,\lambda_0,r))=0$ where $T_\gamma(P+\Delta P,\lambda_0,r)$ is as defined in (4.2.2). By the perturbation theory for singular values,

$$\begin{aligned} f(\gamma) &= \left|\sigma_{(r+1)n-r}(T_\gamma(P+\Delta P,\lambda_0,r))-\sigma_{(r+1)n-r}(T_\gamma(P,\lambda_0,r))\right| \\ &\leqslant \|T_\gamma(\Delta P,\lambda_0,r)\|_2\,. \end{aligned}$$

Observe that

$$T_\gamma(\Delta P,\lambda_0,r)=\left(I_{r+1}\otimes\begin{bmatrix}\Delta P(\lambda_0) & \cdots & \frac{1}{r!}\Delta P^r(\lambda_0)\end{bmatrix}\right)\begin{bmatrix}\hat{E}_{\gamma,1}^T & \hat{E}_{\gamma,2}^T\cdots & \hat{E}_{\gamma,r+1}^T\end{bmatrix}^T$$

where $\hat{E}_{\gamma,j}=E_{\gamma,j}\otimes I_n$, $E_{\gamma,j}$, $j=1,\ldots,r+1$, being the $(r+1)\times(r+1)$ matrices

$$E_{\gamma,1}=\begin{bmatrix}1 & & & \\ & & & \\ & & & \\ & & & \end{bmatrix}, E_{\gamma,2}=\begin{bmatrix} & 1 & & \\ \gamma_1 & & & \\ & & & \\ & & & \end{bmatrix}, \ldots, E_{\gamma,r+1}=\begin{bmatrix} & & & 1\\ & & \gamma_r & \\ & \gamma_r\gamma_{r-1} & & \\ & \iddots & & \\ \prod_{i=1}^{r}\gamma_i & & & \end{bmatrix}.$$

Therefore,

$$f(\gamma)\leqslant\left\|\begin{bmatrix}\Delta P(\lambda_0) & \cdots & \frac{1}{r!}\Delta P^r(\lambda_0)\end{bmatrix}\right\|_2\left\|\begin{bmatrix}\hat{E}_{\gamma,1}^T & \hat{E}_{\gamma,2}^T & \cdots & \hat{E}_{\gamma,r+1}^T\end{bmatrix}^T\right\|_2.$$

Now,

$$\left\|\begin{bmatrix}\Delta P(\lambda_0) & \cdots & \frac{1}{r!}\Delta P^r(\lambda_0)\end{bmatrix}\right\|_2\leqslant\|\Delta P\|_2\|M(\lambda_0;r)\|_2,$$

where the last inequality holds due to Lemma 4.0.1. This shows for $s=2$ or F,

$$\frac{f(\gamma)}{\sqrt{F_1(\gamma)}}\leqslant\delta_s(P,\lambda_0,r+1),$$

and the proof follows by taking the supremum of the left hand side as γ varies over Γ. The lower bound for the case $r>k$ can be proved by similar arguments. □

A second lower bound of $\delta_s(P,\lambda_0,r+1)$ for $s=2$ or F and a lower bound of $\delta_{2,\infty}(P,\lambda_0,r+1)$ are obtained from Theorem 4.3.2.

Theorem 4.5.2. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $n \times n$ matrix polynomial of degree k, and $1 \leqslant r < kn$. For any $\lambda_0 \in \mathbb{C}$ which is not an eigenvalue of $P(\lambda)$, let $T(P, \lambda_0, r)$ be as defined by (4.3.1) and E and $M(\lambda_0; r)$ be as defined in Lemma 4.3.1 and Lemma 4.0.1 respectively. For a given positive integer t, let $\mathbb{S}^t$ be the collection of column vectors of length t with positive entries and*

$$W_{t,n}(a) := \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_t \end{bmatrix} \otimes I_n,$$

where $a = \begin{bmatrix} a_1 & \cdots & a_t \end{bmatrix}^T \in \mathbb{S}^t$. Then for $s = 2$ or F,

$$\delta_s(P, \lambda_0, r+1) \geqslant \sup_{a \in \mathbb{S}^{r+1}} (\sigma_{r+1}(B(\lambda_0, a, P)))^{-1},$$

where

$$B(\lambda_0, a, P) := W_{r+1,(k+1)n}(a)(I_{r+1} \otimes M(\lambda_0; r))E\ (T(P, \lambda_0, r))^{-1} W_{r+1,n}^{-1}(a),$$

and

$$\delta_{2,\infty}(P, \lambda_0, r+1) \geqslant \sup_{a \in \mathbb{S}^{(k+1)(r+1)}} (\sigma_{r+1}(B_{\inf}(\lambda_0, a, P)))^{-1},$$

where

$$B_{\inf}(\lambda_0, a, P) = W\,(I_{r+1} \otimes M(\lambda_0; r))\,E\ (T(P, \lambda_0, r))^{-1}\Big(I_{r+1} \otimes \underbrace{\begin{bmatrix} I_n & \cdots & I_n \end{bmatrix}}_{k+1 \text{ block columns}}\Big)W^{-1},$$

and $W := W_{(k+1)(r+1),n}(a)$.

If $\lambda_0 = 0$ then for $s = 2$, or F,

$$\delta_s(P, 0, r+1) \geqslant \sup_{a \in \mathbb{S}^{r+1}} (\sigma_{r+1}(\widehat{B}(0, a, P)))^{-1},$$

where $\widehat{B}(0, a, P) := W_{r+1,(\min\{r,k\}+1)n}(a)E\ (T(P, 0, r))^{-1} W_{r+1,n}^{-1}(a)$, and

$$\delta_{2,\infty}(P, 0, r+1) \geqslant \sup_{a \in \mathbb{S}^{(p+1)(r+1)}} (\sigma_{r+1}(\widehat{B}_{\inf}(0, a, P)))^{-1},$$

where

$$\widehat{B}_{\inf}(0, a, P) := \widehat{W}E\ (T(P, 0, r))^{-1}\Big(I_{r+1} \otimes \underbrace{\begin{bmatrix} I_n & \cdots & I_n \end{bmatrix}}_{p+1 \text{ block columns}}\Big)\widehat{W}^{-1},$$

and $\widehat{W} := W_{(r+1)(p+1),n}(a)$, p being the minimum of r and k.

Proof. Let $S(P,\lambda_0)$ be as in Corollary 4.2.3. If $-\Delta P(\lambda) \in S(P,\lambda_0)$, from (4.3.4) the nullity of

$$I - W_{r+1,n}^{-1}(a)\,(I_{r+1} \otimes [\Delta A_0 \cdots \Delta A_k])\, W_{r+1,(k+1)n}(a)(I_{r+1} \otimes M(\lambda_0; r))E\,(T(P,\lambda_0,r))^{-1}$$

is at least $r+1$. Therefore nullity $(I - (I_{r+1} \otimes [\Delta A_0 \cdots \Delta A_k])\, B(\lambda_0, a, P)) \geqslant r+1$. By Lemma 1.6.3, $\left\| \begin{bmatrix} \Delta A_0 & \cdots & \Delta A_k \end{bmatrix} \right\|_2 \geqslant (\sigma_{r+1}(B(\lambda_0, a, P)))^{-1}$, so that for $s=2$ or F, $\delta_s(P, \lambda_0, r+1) \geqslant \sup\limits_{a \in \mathbb{S}^{r+1}} (\sigma_{r+1}(B(\lambda_0, a, P)))^{-1}$. Again from (4.3.5) the nullity of

$$I - W^{-1}\left(I_{r+1} \otimes \begin{bmatrix} \Delta A_0 & & \\ & \ddots & \\ & & \Delta A_k \end{bmatrix} \right) W\,(I_{r+1} \otimes M(\lambda_0; r))\,E\,(T(P,\lambda_0,r))^{-1}\Big(I_{r+1} \otimes \underbrace{\begin{bmatrix} I_n & \cdots & I_n \end{bmatrix}}_{k+1 \text{ block columns}}\Big)$$

is at least $r+1$. This is equivalent to

$$\text{nullity}\left(I - \left(I_{r+1} \otimes \begin{bmatrix} \Delta A_0 & & \\ & \ddots & \\ & & \Delta A_k \end{bmatrix} \right) B_{\inf}(\lambda_0, a, P) \right) \geqslant r+1.$$

By Lemma 1.6.3

$$\left\| \begin{bmatrix} \Delta A_0 & & \\ & \ddots & \\ & & \Delta A_k \end{bmatrix} \right\|_2 \geqslant (\sigma_{r+1}(B_{\inf}(\lambda_0, a, P)))^{-1},$$

hence

$$\delta_{2,\infty}(P, \lambda_0, r+1) \geqslant \sup_{a \in \mathbb{S}^{(k+1)(r+1)}} (\sigma_{r+1}(B_{\inf}(\lambda_0, a, P)))^{-1}.$$

Similarly if $-\Delta P(\lambda) \in S(P, 0)$, from inequality (4.3.6) the nullity of

$$I - W_{r+1,n}^{-1}(a)\left(I_{r+1} \otimes \begin{bmatrix} \Delta A_0 & \cdots & \Delta A_p \end{bmatrix} \right) W_{r+1,(p+1)n}(a) E\,(T(P,0,r))^{-1}$$

is at least $r+1$, where $p = \min\{r,k\}$. This is equivalent to

$$\text{nullity}\left(I - \left(I_{r+1} \otimes \begin{bmatrix} \Delta A_0 & \cdots & \Delta A_p \end{bmatrix} \right) \widehat{B}(0, a, P) \right) \geqslant r+1.$$

By Lemma 1.6.3, $\left\| \begin{bmatrix} \Delta A_0 & \cdots & \Delta A_p \end{bmatrix} \right\|_2 \geqslant (\sigma_{r+1}(\widehat{B}(0,a,P)))^{-1}$. Therefore

$$\delta_s(P, 0, r+1) \geqslant \sup_{a \in \mathbb{S}^{r+1}} (\sigma_{r+1}(\widehat{B}(0,a,P)))^{-1} \text{ for } s = 2 \text{ or } F.$$

Again from (4.3.7), the nullity of

$$I-\widehat{W}^{-1}\left(I_{r+1}\otimes\begin{bmatrix}\Delta A_0 & & \\ & \ddots & \\ & & \Delta A_p\end{bmatrix}\right)\widehat{W}E\ (T(P,0,r))^{-1}\Big(I_{r+1}\otimes\underbrace{\begin{bmatrix}I_n & \cdots & I_n\end{bmatrix}}_{p+1\text{ block columns}}\Big)$$

is at least $r+1$. This is equivalent to

$$\text{nullity}\left(I-\left(I_{r+1}\otimes\begin{bmatrix}\Delta A_0 & & \\ & \ddots & \\ & & \Delta A_p\end{bmatrix}\right)\widehat{B}_{\inf}(0,a,P)\right)\geqslant r+1.$$

By Lemma 1.6.3, $\left\|\begin{bmatrix}\Delta A_0 & & \\ & \ddots & \\ & & \Delta A_p\end{bmatrix}\right\|_2 \geqslant (\sigma_{r+1}(\widehat{B}_{\inf}(0,a,P)))^{-1}$. Therefore

$$\delta_{2,\infty}(P,0,r+1)\geqslant\sup_{a\in\mathbb{S}^{(p+1)(r+1)}}(\sigma_{r+1}(\widehat{B}_{\inf}(0,a,P)))^{-1}.$$

This completes the proof. □

4.6 Upper bound

In this section an upper bound on the distance $\delta_s(P,\lambda_0,r+1)$ for $s=2$ or F that can be used in conjunction with the lower bound obtained in Theorem 4.5.1 is derived.

Theorem 4.6.1. *Let $P(\lambda)=\sum_{i=0}^{k}\lambda^iA_i$ be an $n\times n$ matrix polynomial of degree k and $r<kn$ be a positive integer. Let $\gamma:=[\gamma_1\ldots\gamma_r]\in\Gamma$, for Γ as in (4.2.1), and let $f(\gamma):=\sigma_{(r+1)n-r}(T_\gamma(P,\lambda_0,r))$ where $T_\gamma(P,\lambda_0,r)$ is as in (4.2.2). Suppose that $v(\gamma)=\begin{bmatrix}v_0^T & v_1^T & \cdots & v_r^T\end{bmatrix}^T$ and $u(\gamma)=\begin{bmatrix}u_0^T & u_1^T & \cdots & u_r^T\end{bmatrix}^T$ are the corresponding right and left singular vectors with $v_i,u_i\in\mathbb{C}^n$ for $i=0,1,\ldots,r$ dependent on γ. Also let $\Gamma_0\subset\Gamma$ be the collection of all $\gamma\in\Gamma$ with the property that the vector v_0 formed by the first n entries of a right singular vector $v(\gamma)$ associated with the singular value $f(\gamma)$ of $T_\gamma(P,\lambda_0,r)$ is nonzero. Then for $s=2$ or F,*

$$\delta_s(P,\lambda_0,r+1)\leqslant\begin{cases}\inf_{\gamma\in\Gamma_0}f(\gamma)\left\|U(\gamma)V(\gamma)^\dagger\right\|_s & \text{if } \lambda_0=0,\\ \inf_{\gamma\in\Gamma_0}f(\gamma)\left\|U(\gamma)(M(\lambda_0;r)V(\gamma))^\dagger\right\|_s & \text{otherwise,}\end{cases}\tag{4.6.1}$$

where $M(\lambda_0; r)$ is as defined in Lemma 4.0.1, $U(\gamma) = \begin{bmatrix} u_0 & \cdots & u_r \end{bmatrix}$,

$$V(\gamma) := \begin{cases} \begin{bmatrix} v_0 & v_1 & v_2 & \cdots & v_r \\ & \gamma_1 v_0 & \gamma_2 v_1 & \cdots & \gamma_r v_{r-1} \\ & & \gamma_1\gamma_2 v_0 & \cdots & \gamma_{r-1}\gamma_r v_{r-2} \\ & & & \ddots & \vdots \\ & & & & \left(\prod_{i=1}^{r} \gamma_i\right) v_0 \end{bmatrix} & \text{if } r \leqslant k, \\ \begin{bmatrix} v_0 & v_1 & v_2 & \cdots & v_k & \cdots & v_r \\ & \gamma_1 v_0 & \gamma_2 v_1 & \cdots & \gamma_k v_{k-1} & \cdots & \gamma_r v_{r-1} \\ & & \gamma_1\gamma_2 v_0 & \cdots & \gamma_{k-1}\gamma_k v_{k-2} & \cdots & \gamma_{r-1}\gamma_r v_{r-2} \\ & & & \ddots & \vdots & \vdots & \vdots \\ & & & & \left(\prod_{i=1}^{k} \gamma_i\right) v_0 & \cdots & \left(\prod_{i=r-k+1}^{r} \gamma_i\right) v_{r-k} \end{bmatrix} & \text{otherwise,} \end{cases}$$

and the infimum is taken to be ∞ if $\Gamma_0 = \emptyset$.

Proof. Initially we consider the case $\lambda_0 \neq 0$. From (4.4.3) it is clear that if $\gamma \in \Gamma_0$, then $V(\gamma)$ satisfies

$$\delta_s(P, \lambda_0, r+1) \leqslant \begin{cases} \left\| \begin{bmatrix} P(\lambda_0) & \cdots & \frac{1}{r!}P^r(\lambda_0) \end{bmatrix} V(\gamma)(M(\lambda_0; r)V(\gamma))^\dagger \right\|_s & \text{if } r \leqslant k, \\ \left\| \begin{bmatrix} P(\lambda_0) & \cdots & \frac{1}{k!}P^k(\lambda_0) \end{bmatrix} V(\gamma)(M(\lambda_0; r)V(\gamma))^\dagger \right\|_s & \text{otherwise} \end{cases}$$

for $s = 2$ or F. As $v(\gamma)$ is a right singular vector of $T_\gamma(P, \lambda_0, r)$ corresponding to $f(\gamma)$, it is clear that $\begin{bmatrix} P(\lambda_0) & \cdots & \frac{1}{p!}P^p(\lambda_0) \end{bmatrix} V(\gamma) = U(\gamma)$ where $p = \min\{r, k\}$. In either case, $\delta_s(P, \lambda_0, r+1) \leqslant f(\gamma) \left\| U(\gamma)(M(\lambda_0; r)V(\gamma))^\dagger \right\|_s$, and the proof follows by taking the infimum of the right hand side of the above inequality as γ varies over Γ_0. The result for $\lambda_0 = 0$ follows from (4.4.2) by similar arguments. □

Remark 4.6.2. *A matrix polynomial for which $\Gamma_0 = \emptyset$ has never been encountered in practice. Therefore it is conjectured that the upper bound in Theorem 4.6.1 is never ∞. In fact numerical experiments show that in many cases this upper bound is very close to the computed value of the distance.*

4.7 Distance to the set of matrix polynomials having a defective eigenvalue at zero

The quantities $\delta_s(P,0,2), s = 2, F,$ and $\delta_{2,\infty}(P,0,2)$ are the measure of the distance to a matrix polynomial nearest to $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$, having a defective eigenvalue at 0. In this case the problem becomes equivalent to finding a nearest matrix pencil to $\lambda A_1 + A_0$ in the chosen norm that has 0 as a defective eigenvalue. Note that this distance is of significant practical interest as when $P(\lambda)$ is a pencil then the physical significance of $\delta_F(P,0,2)$, $\delta_2(P,0,2)$ and $\delta_{2,\infty}(P,0,2)$ is that they are different measures of the distance of the autonomous system $A_0\dot{x} = -A_1 x$ to a nearest system which has an impulsive solution for some initial conditions. One can find the details in [52]. Here we give an alternative characterization of the distance $\delta_F(P,0,2)$. Under the assumptions that rank $A_0 < n$ and the ranks of the perturbation to each coefficient of $P(\lambda)$ are at-most one, we give a closed form solution of the distance with respect to the norm $|||\cdot|||_F$. Finally in the $|||\cdot|||_{2,\infty}$ norm setting we consider the case where the polynomial under consideration is Hermitian and only Hermitian perturbations are allowed.

Firstly, the upper bound for the distances in Theorem 4.6.1 is given by

$$\inf_{\gamma\in\Gamma_0} f(\gamma) \left\| \begin{bmatrix} u_0 & u_1 \end{bmatrix} \begin{bmatrix} v_0 & v_1 \\ & \gamma v_0 \end{bmatrix}^{\dagger} \right\|_s$$

for $s = 2$ or F where $\begin{bmatrix} v_0 \\ v_1 \end{bmatrix}$ and $\begin{bmatrix} u_0 \\ u_1 \end{bmatrix}$ are the right and left singular vectors of $\begin{bmatrix} P(0) & \\ \gamma P'(0) & P(0) \end{bmatrix}$ corresponding to its $(2n-1)^{\text{th}}$ singular value denoted by $f(\gamma)$. In this case, γ can be allowed to vary over all positive real numbers as the restriction $v_0 \neq 0$ can be removed. To see this, assume that $\gamma > 0$ is such that the corresponding vector $v_0 = 0$. Then clearly $u_0 = 0$ and $f(\gamma) \left\| \begin{bmatrix} 0 & u_1 \end{bmatrix} \begin{bmatrix} 0 & v_1 \\ 0 & 0 \end{bmatrix}^{\dagger} \right\|_s = f(\gamma)$. Let $\Delta P(\lambda) = \sum_{i=0}^{k} \lambda^i \Delta A_i$ where $\Delta A_0 = -f(\gamma) u_1 v_1^*$ and $\Delta A_i = 0$ for all $i = 1, \ldots, k$. Then

$$[\Delta A_0 \quad \Delta A_1] = -f(\gamma) \begin{bmatrix} 0 & u_1 \end{bmatrix} \begin{bmatrix} 0 & v_1 \\ 0 & 0 \end{bmatrix}^{\dagger} \text{ and } |||\Delta P|||_s = f(\gamma) \left\| \begin{bmatrix} 0 & u_1 \end{bmatrix} \begin{bmatrix} 0 & v_1 \\ 0 & 0 \end{bmatrix}^{\dagger} \right\|_s = f(\gamma)$$

for $s = 2$, or F and the relations

$$\begin{bmatrix} P(0) & \\ \gamma P'(0) & P(0) \end{bmatrix} \begin{bmatrix} 0 \\ v_1 \end{bmatrix} = f(\gamma) \begin{bmatrix} 0 \\ u_1 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & u_1^* \end{bmatrix} \begin{bmatrix} P(0) & \\ \gamma P'(0) & P(0) \end{bmatrix} = f(\gamma) \begin{bmatrix} 0 & v_1^* \end{bmatrix},$$

imply that $A_0 v_1 = f(\gamma) u_1, u_1^* A_0 = f(\gamma) v_1^*$ and $u_1^* A_1 = 0$. Therefore,

$$u_1^*(P + \Delta P)(0) = u_1^* A_0 - f(\gamma) v_1^* = 0, \ (P + \Delta P)(0) v_1 = A_0 v_1 - f(\gamma) u_1 = 0$$

and $u_1^*(P + \Delta P)'(0) v_1 = u_1^* A_1 v_1 = 0$. So unless $(P + \Delta P)(\lambda)$ is singular, 0 is a multiple eigenvalue of $(P + \Delta P)(\lambda)$. In either case the objective is achieved as the polynomial $(P + \Delta P)(\lambda)$ is arbitrarily close to having an elementary divisor λ^j, $j \geqslant 2$.

From Theorem 4.5.1, $\delta_s(P, 0, 2) \geqslant \sup\limits_{\gamma > 0} \dfrac{f(\gamma)}{\sqrt{1 + \gamma^2}}$ for $s = 2$ and $s = F$, where $f(\gamma)$ as in Theorem 4.5.1. Additionally the following lower bound holds for $\delta_F(P, 0, 2)$.

Theorem 4.7.1. *Let* $P(\lambda) = \sum_{i=0}^k \lambda^i A_i \in \mathbb{C}[\lambda]^{n \times n}$ *and* $T_\gamma(P, 0, 1) = \begin{bmatrix} A_0 & \\ \gamma A_1 & A_0 \end{bmatrix}$. *Then*

$$\sup_{\gamma > 0} \frac{\sigma_{2n-1}(T_\gamma(P, 0, 1))}{\max\{\sqrt{2}, \gamma\}} \leqslant \delta_F(P, 0, 2).$$

Proof. For any $\Delta A_0, \Delta A_1 \in \mathbb{C}^{n \times n}$, and $\gamma > 0$,

$$\left| \sigma_{2n-1} \left(\begin{bmatrix} A_0 + \Delta A_0 & \\ \gamma(A_1 + \Delta A_1) & A_0 + \Delta A_0 \end{bmatrix} \right) - \sigma_{2n-1} \left(\begin{bmatrix} A_0 & \\ \gamma A_1 & A_0 \end{bmatrix} \right) \right| \leqslant \left\| \begin{bmatrix} \Delta A_0 & \\ \gamma \Delta A_1 & \Delta A_0 \end{bmatrix} \right\|_2.$$

If the pencil $(A_0 + \Delta A_0) + \lambda(A_1 + \Delta A_1)$ has a multiple eigenvalue at zero, then by Theorem 4.2.1

$$\sigma_{2n-1} \left(\begin{bmatrix} A_0 + \Delta A_0 & \\ \gamma(A_1 + \Delta A_1) & A_0 + \Delta A_0 \end{bmatrix} \right) = 0$$

for any $\gamma > 0$. Therefore, for $\gamma > 0$,

$$\sigma_{2n-1}(T_\gamma(P, 0, 1)) \leqslant \left\| \begin{bmatrix} \Delta A_0 & \\ \gamma \Delta A_1 & \Delta A_0 \end{bmatrix} \right\|_2 \leqslant \sqrt{2 \|\Delta A_0\|_F^2 + \gamma^2 \|\Delta A_1\|_F^2} \leqslant \max\{\sqrt{2}, \gamma\} \sqrt{\|\Delta A_0\|_F^2 + \|\Delta A_1\|_F^2},$$

and hence, $\sup\limits_{\gamma > 0} \dfrac{\sigma_{2n-1}(T_\gamma(P, 0, 1))}{\max\{\sqrt{2}, \gamma\}} \leqslant \delta_F(P, 0, 2)$. This completes the proof. □

Since the distance to a nearest matrix polynomial with a defective eigenvalue at zero is equal to the distance to nearest matrix polynomial with a multiple eigenvalue at zero, we state here some results from literature [53, 66] concerning the distance to a nearest matrix polynomial with an eigenvalue of specified multiplicity. Though in [66] the authors have used a different norm to measure the said distance, we have restated the main result of [66] so that it holds in the norm setting chosen here.

Proposition 4.7.2. *[66] Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i \in \mathbb{C}[\lambda]^{n\times n}$ with A_k nonsingular. For $\gamma > 0$ let $f(\gamma) = \sigma_{2n-1}\begin{bmatrix} A_0 & \\ \gamma A_1 & A_0 \end{bmatrix}$. For $f(\hat{\gamma}) := \max_{\gamma>0} f(\gamma)$ and $\hat{\alpha} := \max_{\gamma>0} \frac{f(\gamma)}{\left\|\begin{bmatrix} 1 & 0 \\ \gamma & 1 \end{bmatrix}\right\|_2}$, $\delta_F(P,0,2)$ and $\delta_2(P,0,2)$ satisfy*

$$\hat{\alpha} \leqslant \delta_F(P,0,2) \leqslant \; f(\hat{\gamma}) \; \left\| \begin{bmatrix} u_1(\hat{\gamma}) & u_2(\hat{\gamma}) \end{bmatrix} \begin{bmatrix} v_1(\hat{\gamma}) & v_2(\hat{\gamma}) \end{bmatrix}^{\dagger} \right\|_F ,$$
$$\hat{\alpha} \leqslant \delta_2(P,0,2) \leqslant f(\hat{\gamma}).$$

where $u(\hat{\gamma}) := \begin{bmatrix} u_1(\hat{\gamma}) \\ u_2(\hat{\gamma}) \end{bmatrix}$ and $v(\hat{\gamma}) := \begin{bmatrix} v_1(\hat{\gamma}) \\ v_2(\hat{\gamma}) \end{bmatrix}$ are respectively the left and right singular vectors of $\begin{bmatrix} A_0 & \\ \hat{\gamma} A_1 & A_0 \end{bmatrix}$ with respect to $f(\hat{\gamma})$ with $v_1(\hat{\gamma}), v_2(\hat{\gamma}), u_1(\hat{\gamma}), u_2(\hat{\gamma}) \in \mathbb{C}^n$.

Given $P(\lambda) = A_0 + \lambda A_1$, the nearest pencil with zero as a multiple eigenvalue when only A_0 is perturbed is established to be $\max_{\gamma>0} \sigma_{2n-1} \begin{bmatrix} A_0 & \\ \gamma A_1 & A_0 \end{bmatrix}$ in [53] under certain assumptions. The result in [53] is in a more general form, but here we restate it in a form best suited for our purpose.

Proposition 4.7.3. *[53] Let $A_0, A_1 \in \mathbb{C}^{n\times n}$ and $f(\hat{\gamma})$, $v(\hat{\gamma})$, $v_1(\hat{\gamma})$ and $v_2(\hat{\gamma})$ be as defined in Proposition 4.7.2. If $f(\hat{\gamma})$ is a simple singular value of $\begin{bmatrix} A_0 & \\ \hat{\gamma} A_1 & A_0 \end{bmatrix}$ and $\operatorname{rank} \begin{bmatrix} v_1(\hat{\gamma}) & v_2(\hat{\gamma}) \end{bmatrix} = 2$, then*

$$\inf\{\|\Delta A_0\|_2 \; \big| (A_0 + \Delta A_0) + \lambda A_1 \textit{ has a multiple eigenvalue at zero}\} = f(\hat{\gamma}).$$

Next we formulate the distances $\delta_F(P,0,2)$ and $\delta_{2,\infty}(P,0,2)$ as optimizations. The formulation for $\delta_F(P,0,2)$ is different from the one in Theorem 4.4.1. The key to the results is the fact that the distance to a nearest regular polynomial with zero as a defective eigenvalue is equal to the distance to a nearest regular matrix polynomial with a multiple eigenvalue at zero. From [6] it follows that a regular matrix polynomial $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ has a multiple eigenvalue at zero if and only if there exist nonzero vectors x and y such that $A_0 x = 0, y^* A_0 = 0$ and $y^* A_1 x = 0$. Therefore, the distance to a matrix polynomial having multiple eigenvalue at zero is equivalent to finding smallest possible perturbations ΔA_0 and ΔA_1 with respect to the chosen norm settings such that

$$(A_0 + \Delta A_0)x = 0, y^*(A_0 + \Delta A_0) = 0 \text{ and } y^*(A_1 + \Delta A_1)x = 0$$

for some nonzero vectors x and y. The following proposition, which is a combination of [70, Theorems 1.2.13 and 1.2.14] and Theorem 1.5.1, helps in finding such perturbations.

Proposition 4.7.4. *Let $E, A \in \mathbb{C}^{n\times n}$ and $x, y \in \mathbb{C}^n$ such that $\|x\|_2 = \|y\|_2 = 1$.*

(a) Let P_E be the collection of all $\Delta E \in \mathbb{C}^{n\times n}$ satisfying

$$\Delta E x = -Ex \text{ and } \Delta E^* y = -E^* y.$$

Then, any $\Delta E \in P_E$ is of the form

$$\Delta E := (I_n - yy^*)R(I_n - xx^*) - Exx^* - yy^*E(I_n - xx^*) \text{ for } R \in \mathbb{C}^{n\times n}.$$

(b) **Minimal Frobenius norm:**

$$\min_{\Delta E \in P_E} \|\Delta E\|_F = (\|Ex\|_2^2 + \|E^*y\|_2^2 - |y^*Ex|^2)^{\frac{1}{2}} \tag{4.7.1}$$

which is attained by $\hat{\Delta} E := -Exx^* - yy^*E(I_n - xx^*)$.

(c) **Minimal 2-norm:**

$$\min_{\Delta E \in P_E} \|\Delta E\|_2 = \max\{\|Ex\|_2, \|E^*y\|_2\}. \tag{4.7.2}$$

Case 1: $\|Ex\|_2 \geqslant \|E^*y\|_2$.
The minimal 2 norm is attained by

$$\hat{\Delta} E := y^*Ex\left(yx^* + \frac{((y^*Ex)y - Ex)\,((y^*Ex)x - E^*y)^*}{\|(y^*Ex)y - Ex\|_2^2}\right) - Exx^* - yy^*E$$

if Ex and y are linearly independent and by $\hat{\Delta} E := -yy^*E$, otherwise.

Case 2: $\|Ex\|_2 < \|E^*y\|_2$.
The minimal 2 *norm is attained by*

$$\hat{\Delta}E := y^*Ex\left(yx^* + \frac{((y^*Ex)y - Ex)\,((y^*Ex)x - E^*y)^*}{\|(y^*Ex)x - E^*y\|_2^2}\right) - Exx^* - yy^*E$$

if E^*y *and* x *are linearly independent and by* $\hat{\Delta}E := -Exx^*$, *otherwise.*

(d) Let P_A *be the collection of all* $\Delta A \in \mathbb{C}^{n\times n}$ *such that* $y^*\Delta Ax = -y^*Ax$. *Then any* $\Delta A \in P_A$ *is of the form*

$$\Delta_A = Z - yy^*Zxx^* - (y^*Ax)yx^* \text{ for } Z \in \mathbb{C}^{n\times n}.$$

(e) **Minimal Frobenius and 2-norm:**

$$\min_{\Delta A\in P_A} \|\Delta_A\|_F = \min_{\Delta A\in P_A} \|\Delta_A\|_2 = |y^*Ax| \tag{4.7.3}$$

and both minimal norms are attained by $\hat{\Delta}_A := -(y^*Ax)yx^*$.

We obtain the following characterization of $\delta_F(P,0,2)$ and $\delta_{2,\infty}(P,0,2)$ as a direct consequence of Proposition 4.7.4.

Theorem 4.7.5. *Suppose* $P(\lambda) = \sum_{i=1}^{k} \lambda^i A_i \in \mathbb{C}[\lambda]^{n\times n}$. *Then,*

$$\delta_F(P,0,2) = \inf_{\|x\|_2=\|y\|_2=1} \sqrt{\|A_0x\|_2^2 + \|A_0^*y\|_2^2 - |y^*A_0x|^2 + |y^*A_1x|^2} \tag{4.7.4}$$

and

$$\delta_{2,\infty}(P,0,2) = \inf_{\|x\|_2=\|y\|_2=1} \{\max\{\|A_0x\|_2, \|A_0^*y\|_2, |y^*A_1x|\}\}. \tag{4.7.5}$$

Remark 4.7.6. *From Theorem 4.7.5 it is clear that for a regular matrix polynomial, only the constant coefficient and the coefficient of* λ *need to be perturbed to achieve the objective of finding a nearest matrix polynomial having a multiple eigenvalue at zero. The optimal perturbations* ΔA_0 *and* ΔA_1 *that correspond to the minimizations in Theorem 4.7.5 can be explicitly constructed by using Proposition 4.7.4.*

4.7.1 Low rank perturbations

In this section we obtain a closed form expression for the distance from a given matrix polynomial $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$, where A_0 is singular to a nearest matrix polynomial having multiple zero eigenvalue via perturbations of rank at most one. We formulate the problem as

$$\inf\{\|\Delta P\|_F | (P+\Delta P)(\lambda) \text{ has a multiple eigenvalue at zero and } \operatorname{rank} \Delta A_i \leqslant 1\}.$$

Theorem 4.7.7. *Let* $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i \in \mathbb{C}[\lambda]^{n\times n}$ *with* $\operatorname{rank} A_0 = n-1$. *Suppose* $A_0 = U\Sigma V^*$ *is a Singular Value Decomposition (SVD) of* A_0 *and* $a_{i,j}$ *is the entry of* U^*A_1V *in the* i^{th} *row and* j^{th} *column. For*

$$X := \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_{n-1} & \\ a_{n,1} & \dots & \dots & a_{n,n-1} & a_{n,n} \end{bmatrix}, Y := \begin{bmatrix} \sigma_1 & & & & a_{1,n} \\ & \sigma_2 & & & a_{2,n} \\ & & \ddots & & \vdots \\ & & & \sigma_{n-1} & a_{n-1,n} \\ & & & & a_{n,n} \end{bmatrix},$$

where $\sigma_1 \geqslant \sigma_2 \geqslant \dots \geqslant \sigma_{n-1} > 0$ *are the singular values of* A_0, *we have*

$$\inf\{\|\Delta P\|_F | (P+\Delta P)(\lambda) \text{ has a multiple eigenvalue at } 0 \text{ and } \operatorname{rank} \Delta A_i \leqslant 1\} = \min\{\sigma_{\min}(X), \sigma_{\min}(Y)\}.$$

Proof. Since the Frobenius norm is unitarily invariant, it is enough to prove the theorem by replacing A_i by U^*A_iV, $i = 0, \dots, k$. Note that in such a case A_0 gets replaced by $\operatorname{diag}(\sigma_1, \dots, \sigma_{n-1}, 0)$.

For any $\Delta P(\lambda) = \sum_{i=0}^{k} \lambda^i \Delta A_i$, $(P+\Delta P)(\lambda)$ has a multiple eigenvalue at zero if and only if $(A_0 + \Delta A_0)x = 0, (A_0+\Delta A_0)^* y = 0$ and $y^*(A_1 + \Delta A_1)x = 0$ for some nonzero $x, y \in \mathbb{C}^n$.

If ΔA_0 has rank at most one then exactly one of the following two cases hold.
Case 1: $(A_0 + \Delta A_0)x = 0$, for some $x \in \mathbb{C}^n$ with $\|x\|_2 = 1$ and $\hat{y}^*(A_1 + \Delta A_1)x = 0$ for a corresponding $\hat{y} \in \mathbb{C}^n$ satisfying

$$\hat{y}^*(A_0 + \Delta A_0) = 0 \text{ and } \|\hat{y}\|_2 = 1.$$

Case 2: $y^*(A_0 + \Delta A_0) = 0$ for some $y \in \mathbb{C}^n$ with $\|y\|_2 = 1$ and $y^*(A_1 + \Delta A_1)\hat{x} = 0$ for a corresponding $\hat{x} \in \mathbb{C}^n$ satisfying

$$(A_0 + \Delta A_0)\hat{x} = 0 \text{ and } \|\hat{x}\|_2 = 1.$$

Proof for Case 1: Choices of $\Delta A_0^\bullet, \Delta A_1^\bullet \in \mathbb{C}^{n\times n}$ each with rank at most one satisfying $(A_0 + \Delta A_0^\bullet)x = 0$ and $\hat{y}^*(A_1 + \Delta A_1^\bullet)x = 0$ that are also minimal with respect to Frobenius norm are given by

$$\Delta A_0^\bullet := -A_0 x x^*, \text{ and } \Delta A_1^\bullet := -(\hat{y}^* A_1 x)\hat{y} x^*$$

respectively. Note that while the choice of $\Delta A_0^\bullet$ is obvious, that of $\Delta A_1^\bullet$ follows from part (e) of Proposition 4.7.4. Now

$$\hat{y}^*(A_0 + \Delta A_0^\bullet) = 0 \Rightarrow (I_n - xx^*)A_0^* \hat{y} = 0$$

and since $\|x\|_2 = 1$, either $A_0^*\hat{y} = 0$ or $A_0^*\hat{y} = \alpha x$ for some $\alpha \in \mathbb{C}\backslash\{0\}$. If $A_0^*\hat{y} = 0$, then $\hat{y} = \pm e_n$ and $\Delta A_1^\bullet = -(a_n x)e_n x^*$ where a_n is the last row of the matrix A_1. Therefore in this case

$$\|\Delta A_0^\bullet\|_F^2 + \|\Delta A_1^\bullet\|_F^2 = \|(A_0 x)x^*\|_F^2 + \|(a_n x)e_n x^*\|_F^2 = \|Xx\|_2^2 . \tag{4.7.6}$$

On the other hand, if $A_0^*\hat{y} = \alpha x$, then $x^* e_n = 0$ and $(A_0 + \Delta A_0^\bullet)e_n = 0$. Since we also have $(A_0 + \Delta A_0^\bullet)x = 0$, so $x^* e_n = 0$ implies that $\sqrt{\|\Delta A_0^\bullet\|_F^2 + \|\Delta A_1^\bullet\|_F^2} \geqslant \sigma_{n-1}$.

Proof for Case 2: Arguing as in the proof of Case 1, a choice of $\Delta A_0^\bullet, \Delta A_1^\bullet \in \mathbb{C}^{n\times n}$ with ranks at most one that satisfy $y^*(A_0 + \Delta A_0^\bullet) = 0$ and $y^*(A_1 + \Delta A_1^\bullet)\hat{x} = 0$ and are also minimum with respect to Frobenius norm are given by

$$\Delta A_0^\bullet := -yy^* A_0, \text{ and } \Delta A_1^\bullet := -(y^* A_1 \hat{x})y\hat{x}^*$$

respectively. Therefore,

$$(A_0 + \Delta A_0^\bullet)\hat{x} = 0 \Rightarrow (I_n - yy^*)A_0\hat{x} = 0.$$

and since $\|y\|_2 = 1$, either $A_0\hat{x} = 0$ or $A_0\hat{x} = \beta y$ for some $\beta \in \mathbb{C}\backslash\{0\}$. If $A_0\hat{x} = 0$, then $\hat{x} = \pm e_n$ and $\Delta A_1^\bullet = -(y^* a_n')y e_n^*$ where a_n' is the last column of the matrix A_1. Therefore,

$$\|\Delta A_0^\bullet\|_F^2 + \|\Delta A_1^\bullet\|_F^2 = \|(yy^*)A_0\|_F^2 + \|(y^* a_n')y e_n^*\|_F^2 = \|y^* Y\|_2^2 . \tag{4.7.7}$$

On the other hand, if $A_0\hat{x} = \beta y$, then $y^* e_n = 0$ and $e_n^*(A_0 + \Delta A_0^\bullet) = 0$. Since $y^*(A_0 + \Delta A_0^\bullet) = 0$, so $y^* e_n = 0$ implies that $\sqrt{\|\Delta A_0^\bullet\|_F^2 + \|\Delta A_1^\bullet\|_F^2} \geqslant \sigma_{n-1}$. Setting $\Delta P(\lambda) = \sum_{i=0}^k \lambda^i \Delta A_i^\bullet$ where $\Delta A_i^\bullet = 0$ for $i = 2, \ldots, k$, in either case it follows that

$$\begin{aligned} &\inf\{\|\Delta P\|_F | (P + \Delta P)(\lambda) \text{ has multiple eigenvalue at zero and rank}\, \Delta A_i \leqslant 1\} \\ &= \min\left\{\min_{\|x\|_2=1} \|Xx\|_2 , \min_{\|y\|_2=1} \|y^* Y\|_2 , \sigma_{n-1}\right\} \\ &= \min\{\sigma_{\min}(X), \sigma_{\min}(Y)\} \end{aligned}$$

where the last equality holds as σ_{n-1} is not strictly smaller than $\sigma_{\min}(X)$ due to the particular structure of X. This completes the proof. □

4.7.2 Hermitian perturbations

In certain cases, the physical structure of the system itself gives rise to certain structures (for example, Hermitian) in the coefficient matrices of the matrix polynomial. Many such examples have been studied in literature. For example, a typical control system, denoted as (E, A, B, C, D), that arises from the discretization of partial differential equations, model reduction, realization or system identification, is $E\dot{v} = Av + Bu,\ w = Cv + Du$, where E, A, B, C, D are real or complex matrices of size $n \times n, n \times n, n \times m, m \times n$, and $m \times m$ respectively. If $sE - A$ has no finite eigenvalues in the closed right half of the complex plane, then determining whether the system is passive, (i.e., whether it absorbs supply energy), is equivalent to the Hermitian pencil

$$s \begin{bmatrix} 0 & iE & 0 \\ -iE^* & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & A & B \\ A^* & 0 & C^* \\ B^* & C & D + D^* \end{bmatrix}$$

being regular, having only nonreal eigenvalues and infinity being an eigenvalue with equal algebraic and geometric multiplicities (see [37]). Motivated by this fact, in this section we consider the problem of finding the distance with respect to the norm $|||\cdot|||_{2,\infty}$ from a given $n \times n$ Hermitian matrix polynomial $P(\lambda)$ to a nearest matrix polynomial in the class, say $\mathcal{H}_n$, of $n \times n$ Hermitian matrix polynomials that are either singular or have an elementary divisor $\lambda^j, j \geqslant 2$. Recall that this is equivalent to finding the distance from $P(\lambda)$ to a nearest Hermitian matrix polynomial that is either singular or has ∞ as a defective eigenvalue. We denote and define the distance as

$$\delta^H_{2,\infty}(P, 0, 2) = \inf\{|||\Delta P|||_{2,\infty} : \Delta P^*(\lambda) = \Delta P(\lambda), (P + \Delta P)(\lambda) \in \mathcal{H}_n\}.$$

The following lemma characterizes the matrix polynomials in $\mathcal{H}_n$.

Lemma 4.7.8. *Let $Q(\lambda) := \sum_{i=0}^{k} \lambda^i B_i \in \mathcal{H}_n$. Then $B_0 x = 0$ and $x^* B_1 x = 0$ for some nonzero $x \in \mathbb{C}^n$. Conversely if there exists a nonzero $x \in \mathbb{C}^n$ satisfying $B_0 x = 0$ and $x^* B_1 x = 0$, then $Q(\lambda)$ belongs to the closure of $\mathcal{H}_n$ with respect to any norm on the $n \times n$ Hermitian matrix polynomials.*

Proof. Suppose $Q(\lambda) = \sum_{i=0}^{k} \lambda^i B_i \in \mathcal{H}_n$ is singular. By [41]

$$\mathcal{L}(\lambda) = \lambda \begin{bmatrix} & & & & B_k \\ & & & \iddots & B_{k-1} \\ & & \iddots & \iddots & \vdots \\ & \iddots & \iddots & & B_2 \\ B_k & B_{k-1} & \cdots & B_2 & B_1 \end{bmatrix} - \begin{bmatrix} & & & B_k & \\ & & \iddots & B_{k-1} & \\ & \iddots & \iddots & \vdots & \\ B_k & B_{k-1} & \cdots & B_2 & \\ & & & & -B_0 \end{bmatrix}$$

is a Hermitian linearization of $Q(\lambda)$ belonging to $\mathbb{L}_1(Q)$ (given by (2.0.1)) corresponding to the ansatz vector e_k. From the canonical form under congruence of Hermitian matrix pencils in (1.2.1), $\mathcal{L}(\lambda)$ has a vector polynomial $w(\lambda)$ in its right and left minimal bases satisfying $w(0) \neq 0$ and $w(0)^*\mathcal{L}'(0)w(0) = 0$. By [14], there exists vector polynomial $x(\lambda)$ in the left and right minimal bases of $Q(\lambda)$ such that $w(\lambda) = \Lambda_k(\lambda) \otimes x(\lambda)$. Clearly $w(0) \neq 0 \Rightarrow x(0) \neq 0$. As $\mathcal{L}(\lambda) \in \mathbb{L}_1(Q)$, $\mathcal{L}(\lambda)(\Lambda_k(\lambda) \otimes I_n) = e_k \otimes Q(\lambda)$. Differentiating both sides with respect to λ and multiplying on the left and right by $w(\lambda)^*$ and $1 \otimes x(\lambda)$ respectively,

$$\begin{aligned} &w(\lambda)^*\mathcal{L}'(\lambda)(\Lambda_k(\lambda) \otimes I_n)(1 \otimes x(\lambda)) = w(\lambda)^*(e_k \otimes Q'(\lambda)x(\lambda)) \\ \Rightarrow &w(\lambda)^*\mathcal{L}'(\lambda)w(\lambda) = \Lambda_k^T(\lambda)e_k(x(\lambda)^*Q'(\lambda)x(\lambda)) \end{aligned}$$

Since $\Lambda_k^T(\lambda)e_k = 1$, and the left hand side of the second equation is zero for $\lambda = 0$, we have $x^*B_1x = 0$ for $x = x(0) \neq 0$. Also clearly $B_0x = Q(0)x(0) = 0$.

Next let $Q(\lambda)$ be regular with an elementary divisor $\lambda^k, k \geqslant 2$. This is equivalent to the Hermitian pencil $\lambda B_1 + B_0$ having such an elementary divisor. Therefore the canonical form of $\lambda B_1 + B_0$ under congruence in (1.2.1), has an $m \times m$ block of the form $\lambda F_m + G_m$ where $m \geqslant 2$ and F_m and G_m are the $m \times m$ matrices

$$F_m = \begin{bmatrix} 0 & \cdots & \cdots & 0 & 1 \\ \vdots & & & 1 & 0 \\ \vdots & & & & \vdots \\ 0 & 1 & & & \vdots \\ 1 & 0 & \cdots & \cdots & 0 \end{bmatrix}, \; G_m = \begin{bmatrix} F_{m-1} & 0 \\ 0 & 0 \end{bmatrix}$$

Clearly, this implies that there exists $x \in \mathbb{C}^n \setminus \{0\}$ such that $B_0x = 0$ and $x^*B_1x = 0$.

Now suppose that there exists $x \in \mathbb{C}^n \setminus \{0\}$, such that $B_0x = 0$ and $x^*B_1x = 0$. Assume without loss of generality that $Q(\lambda)$ is regular and $\lambda^k, k \geqslant 2$, is not an elementary divisor of $Q(\lambda)$. This implies that the canonical form under congruence of the pencil $\lambda B_1 + B_0$ has a diagonal block $D_m := \lambda \mathrm{diag}(s_1 \cdots s_m)$ for some $m \geq 2$ where $s_i = \pm 1, i = 1, \ldots, m$ with $s_is_j = -1$ for some $i \neq j$. Consider a Hermitian perturbation to the canonical form such that for arbitrary $\epsilon > 0$, the block D_m changes to $D_m + \epsilon(e_i + e_j)(e_i + e_j)^*$ and the rest of the canonical form remains unchanged. Let $\lambda\Delta B_1 + \Delta B_0$ be the Hermitian perturbation to $\lambda B_1 + B_0$ that induces this change in the canonical form. Clearly 0 is a defective eigenvalue of $\lambda(B_1 + \Delta B_1) + (B_0 + \Delta B_0)$ and hence $\lambda^k, k \geq 2$ is an elementary divisor. Therefore

$(Q+\Delta Q)(\lambda) := \sum_{i=2}^{k} \lambda^i B_i + \lambda(B_1+\Delta B_1) + B_0 + \Delta B_0 \in \mathcal{H}_n$. By choosing $\epsilon > 0$ small enough, $(Q+\Delta Q)(\lambda)$ can be brought arbitrarily close to to $Q(\lambda)$. Therefore $Q(\lambda)$ belongs to the closure of $\mathcal{H}_n$ and this completes the proof. □

Theorem 4.7.9. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be a Hermitian matrix polynomial. Then*

$$\delta_{2,\infty}^{H}(P,0,2) = \inf_{\|x\|_2=1} \{\max\{\|A_0x\|_2\,, |x^*A_1x|\}\}\,.$$

Proof. Let $\Delta P(\lambda) = \sum_{i=0}^{k} \lambda^i \Delta A_i$ be a Hermitian matrix polynomial. By Lemma 4.7.8 finding $\delta_{2,\infty}^{H}(P,0,2)$ is equivalent to finding the smallest $|\!|\!|\Delta P|\!|\!|_{2,\infty}$ such that

$$(A_0+\Delta A_0)x = 0 \text{ and } x^*(A_1+\Delta A_1)x = 0 \tag{4.7.8}$$

for some $x \in \mathbb{C}^n$ with $\|x\|_2 = 1$. By [5, Theorem 2.6]

$$\min\{\|\Delta A_0\|_2\,|\Delta A_0 = \Delta A_0^*, \Delta A_0 x = -A_0x \text{ and } \|x\|_2 = 1\} = \|A_0x\|_2\,. \tag{4.7.9}$$

Moreover the norm is minimized by

$$\Delta A_0 = \frac{\|A_0x\|_2^2}{(x^*A_0x)^2 - \|A_0x\|_2^2}\left[A_0xx^* + xx^*A_0 - x^*A_0x\left(xx^* + \frac{A_0x(A_0x)^*}{\|A_0x\|_2^2}\right)\right] \tag{4.7.10}$$

if x and A_0x are linearly independent and by $\Delta A_0 = -A_0xx^*$ otherwise. Again by Theorem 1.5.1

$$\min\{\|\Delta A_1\|_2\,|\Delta A_1 = \Delta A_1^*, x^*\Delta A_1 x = -x^*A_1x \text{ and } \|x\|_2 = 1\} = |x^*A_1x| \tag{4.7.11}$$

where the minimal norm is attained by $-(x^*A_1x)xx^*$. The proof now follows from (4.7.9) and (4.7.11). □

If perturbations are allowed only on the coefficient A_0 of $P(\lambda)$ and the restricted distance is denoted by $\hat{\delta}_{2,\infty}^{H}(P,0,2)$, then the following characterization is obtained.

Theorem 4.7.10. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be a Hermitian matrix polynomial. Then*

$$\hat{\delta}_{2,\infty}^{H}(P,0,2) = \begin{cases} \infty, & \text{if } A_1 \text{ is definite,} \\ \sqrt{\lambda_{\min}(V^*A_0^*A_0V)}, & \text{if } A_1 \text{ is semidefinite,} \\ \sqrt{\sup_{t\in[t_1,\,t_2]} \lambda_{\min}(A_0^*A_0 + tA_1)}, & \text{if } A_1 \text{ is indefinite} \end{cases} \tag{4.7.12}$$

where V is an isometry whose columns form an orthonormal basis of the null space of A_1, $\lambda_{\min}(A_0^*A_0+tA_1)$ and $\lambda_{\min}(V^*A_0^*A_0V)$ are the smallest eigenvalues of $A_0^*A_0+tA_1$ and $V^*A_0^*A_0V$ respectively, $t_1 := \frac{\sigma_{\min}(A_0)^2-\sigma_{\max}(A_0)^2}{\lambda_{\max}(A_1)}$, and $t_2 := \frac{\sigma_{\min}(A_0)^2-\sigma_{\max}(A_0)^2}{\lambda_{\min}(A_1)}$, and $\lambda_{\max}(A_1)$ and $\lambda_{\min}(A_1)$ are the largest and smallest eigenvalues of A_1 respectively.

Proof. Along the lines of Theorem 4.7.9, we deduce that

$$\hat{\delta}^{H}_{2,\infty}(P,0,2) = \inf\{\|A_0x\|_2 \,|x \in \mathbb{C}^n, x^*A_1x = 0, \|x\|_2 = 1\}. \tag{4.7.13}$$

Now $\|A_0x\|_2^2 = x^*A_0^*A_0x$. Thus by [48, Theorem 6.2], we have,

$$\hat{\delta}^{H}_{2,\infty}(P,0,2) = \begin{cases} \infty, & \text{if } A_1 \text{ is definite,} \\ \sqrt{\lambda_{\min}(V^*A_0^*A_0V)}, & \text{if } A_1 \text{ is semidefinite,} \\ \sqrt{\sup\limits_{t\in[t_1,\,t_2]} \lambda_{\min}(A_0^*A_0 + tA_1)}, & \text{if } A_1 \text{ is indefinite.} \end{cases} \tag{4.7.14}$$

This completes the proof. □

4.8 Numerical Experiments

This section presents numerical experiments conducted to illustrate the upper and lower bounds on the distances and their values computed via the Broyden Fletcher Goldfarb Shanno (BFGS) algorithm and MATLAB's `globalsearch.m` algorithm from the formulation in Theorem 4.4.1.

Computing $\delta_F(P,\lambda_0,r+1)$ from the optimization in Theorem 4.4.1 via BFGS requires the gradient of the objective function

$$f(X) := \begin{cases} \left\|HXX^{\dagger}\right\|_F & \text{if } \lambda_0 = 0, \\ \left\|HX(M(\lambda_0;r)X)^{\dagger}\right\|_F & \text{otherwise,} \end{cases}$$

where X varies depending on whether $r \leqslant k$ or $r > k$, $H = \begin{bmatrix} P(\lambda_0) & \cdots & \frac{1}{p!}P^p(\lambda_0) \end{bmatrix}$ and $p = \min\{r,k\}$. We consider the case $\lambda_0 \neq 0$. By Lemma 4.0.1, $H = GM(\lambda_0;r)$, where $G = \begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix}$. Therefore

$$f(X) = \left\|G(M(\lambda_0;r)X)(M(\lambda_0;r)X)^{\dagger}\right\|_F.$$

Only real matrix polynomials are considered in the experiments. Since $M(\lambda_0;r)$ has full column rank, for any $X = X_0$ if there exists a neighborhood S of X_0 such that $\operatorname{rank} X_0 = \operatorname{rank} X$ for all $X \in S$ then $f(X)$ is differentiable at X_0. If we use any numerical scheme to find the infimum of $f(X)$, then generically at every step there exists a neighborhood S of X where every element of S is of full rank and consequently we can find gradient of $f(X)$ at those points. Additionally the matrix X involved in the objective function $f(X)$ has block Toeplitz structure which needs to be incorporated when finding the gradient of $f(X)$. For simplicity, the gradient is initially computed for the function $(f(X))^2$ without taking the structure of X into consideration with the changes due to the structure being incorporated later. Therefore the function under consideration is

$$g(X) := (f(X))^2 = \left\| G(M(\lambda_0; r)X)(M(\lambda_0; r)X)^\dagger \right\|_F^2 .$$

Considering $g(X)$ as a real valued function of the entries of X,

$$\nabla g(X)\Big|_{X=X_0} = \text{vec}\left(\frac{dg}{dX}\Big|_{X=X_0} \right).$$

Now, setting $Y = M(\lambda_0; r)X$,

$$dg = 2\langle GYY^\dagger, Gd(YY^\dagger)\rangle \text{ where } \langle A, B\rangle = \text{trace} A^T B.$$

Expanding the right hand side gives

$$dg = 2\langle G^T GYY^\dagger (Y^\dagger)^T, dY\rangle + 2\langle Y^T G^T GYY^\dagger, dY^\dagger\rangle \tag{4.8.1}$$

where,

$$dY^\dagger = (I - Y^\dagger Y)dY^T (Y^\dagger)^T (Y^\dagger) + (Y^\dagger)(Y^\dagger)^T dY^T (I - YY^\dagger) - (Y^\dagger)dY(Y^\dagger).$$

Therefore,

$$\begin{aligned} dg =& 2\langle G^T GYY^\dagger (Y^\dagger)^T, dY\rangle + 2\langle Y^T G^T GYY^\dagger, (I - Y^\dagger Y)dY^T (Y^\dagger)^T (Y^\dagger)\rangle \\ & + 2\langle Y^T G^T GYY^\dagger, (Y^\dagger)(Y^\dagger)^T dY^T (I - YY^\dagger)\rangle - 2\langle Y^T G^T GYY^\dagger, (Y^\dagger)dY(Y^\dagger)\rangle \\ =& 2\langle G^T GYY^\dagger (Y^\dagger)^T, dY\rangle + 2\langle (I - Y^\dagger Y)^T Y^T G^T GYY^\dagger (Y^\dagger)^T (Y^\dagger), dY^T\rangle \\ & + 2\langle (Y^\dagger)(Y^\dagger)^T Y^T G^T GYY^\dagger (I - YY^\dagger)^T, dY^T\rangle \\ & - 2\langle (Y^\dagger)^T Y^T G^T GYY^\dagger (Y^\dagger)^T, dY\rangle \\ =& 2\langle \mathcal{F}(G, Y), dY\rangle \end{aligned}$$

where

$$\begin{aligned} \mathcal{F}(G, Y) :=& G^T GYY^\dagger (Y^\dagger)^T + (Y^\dagger)^T Y^\dagger (Y^\dagger)^T Y^T G^T GY(I - Y^\dagger Y) \\ & + (I - YY^\dagger)(Y^\dagger)^T Y^T G^T GYY^\dagger (Y^\dagger)^T - (Y^\dagger)^T Y^T G^T GYY^\dagger (Y^\dagger)^T. \end{aligned}$$

As, $dY = M(\lambda_0; r)X$, $\frac{dg}{dX} = 2M(\lambda_0; r)^T \mathcal{F}(G, Y) =: \psi(X)$. Now at a fixed X_0, $\frac{dg}{dX}\Big|_{X=X_0} = \psi(X_0)$. Due to the structure of X, $\nabla g(X)\Big|_{X=X_0}$ is given by

$$\nabla g(X)\Big|_{X=X_0} = \begin{bmatrix} \sum_{i=1}^{r+1}\left(\begin{bmatrix} \psi(X_0)_{(i-1)n+1,i} \\ \vdots \\ \psi(X_0)_{in,i} \end{bmatrix}\right) \\ \sum_{i=1}^{r}\left(\begin{bmatrix} \psi(X_0)_{(i-1)n+1,i+1} \\ \vdots \\ \psi(X_0)_{in,i+1} \end{bmatrix}\right) \\ \vdots \\ \begin{bmatrix} \psi(X_0)_{1,r+1} \\ \vdots \\ \psi(X_0)_{n,r+1} \end{bmatrix} \end{bmatrix} \text{ if } r \leqslant k, \text{ and } \nabla g(X)\Big|_{X=X_0} = \begin{bmatrix} \sum_{i=1}^{k+1}\left(\begin{bmatrix} \psi(X_0)_{(i-1)n+1,i} \\ \vdots \\ \psi(X_0)_{in,i} \end{bmatrix}\right) \\ \vdots \\ \sum_{i=1}^{k+1}\left(\begin{bmatrix} \psi(X_0)_{(i-1)n+1,i+r-k} \\ \vdots \\ \psi(X_0)_{in,i+r-k} \end{bmatrix}\right) \\ \sum_{i=1}^{k}\left(\begin{bmatrix} \psi(X_0)_{(i-1)n+1,i+r-k+1} \\ \vdots \\ \psi(X_0)_{in,i+r-k+1} \end{bmatrix}\right) \\ \vdots \\ \begin{bmatrix} \psi(X_0)_{1,r+1} \\ \vdots \\ \psi(X_0)_{n,r+1} \end{bmatrix} \end{bmatrix} \text{ if } r > k.$$

The gradient of $(f(X))^2$ for the case $\lambda_0 = 0$ may be found on similar lines. Due to the difficulties in computing the gradient of the associated objective function, the optimization for $\delta_2(P, \lambda_0, r+1)$ in Theorem 4.4.1 is performed only via MATLAB's `globalsearch.m`. Also in each case, the optimizations involved in the lower and upper bounds are computed via `globalsearch.m` algorithm.

Example 4.8.1. Consider a 2×2 matrix polynomial of degree 3,

$$P(\lambda) = \begin{bmatrix} -0.1414 & -0.1490 \\ 1.1928 & 0.9702 \end{bmatrix} + \lambda \begin{bmatrix} 0.8837 & 0.9969 \\ 0.2190 & 0.0259 \end{bmatrix} + \lambda^2 \begin{bmatrix} 0.6346 & 0.9689 \\ 0.6252 & -0.0649 \end{bmatrix} + \lambda^3 \begin{bmatrix} -1.9867 & 1.2800 \\ 0.6097 & -0.1477 \end{bmatrix}.$$

Table 4.8.1 records the the values of the distance $\delta_F(P, 0, r)$ computed via BFGS and `globalsearch.m` algorithms using the formulation in Theorem 4.4.1 for various values of r together with lower bounds from Theorem 4.5.1 and Theorem 4.5.2 and the upper bound from Theorem 4.6.1. Table 4.8.2 records the same for $\delta_F(P, 1, r)$ as r varies from 2 to 6. Likewise, Table 4.8.3 and Table 4.8.4 records the corresponding quantities for the distances $\delta_2(P, 0, r)$ and $\delta_2(P, 1, r)$ respectively, except that in these cases the distance is computed only via the `globalsearch.m` algorithm. Tables 4.8.5 and 4.8.6 records the bounds for the distances $\delta_{2,\infty}(P, 0, r)$ and $\delta_{2,\infty}(P, 1, r)$ for all possible values of $r > 1$. In this case the lower bound is calculated using `globalsearch.m` on the corresponding formulation in Theorem 4.5.2. To obtain the upper bounds, the $|||\cdot|||_{2,\infty}$ norms of the perturbations attaining the upper bounds with respect to $|||\cdot|||_F$ and $|||\cdot|||_2$ in Theorem 4.6.1 are calculated. The $|||\cdot|||_{2,\infty}$

norms of the perturbations corresponding to the computed distances $\delta_F(P, \lambda_0, r)$ and $\delta_2(P, \lambda_0, r)$ where $\lambda_0 \in \{0, 1\}$ obtained from BFGS and/or `globalsearch.m` are also calculated. The least of the two sets of norms is then reported as the upper bound.

Distance measured	Lower bounds (Theorem 4.5.2)	Lower bounds (Theorem 4.5.1)	Estimated distance BFGS	Estimated distance `globalsearch`	Upper bound (Theorem 4.6.1)
$\delta_F(P, 0, 2)$	0.10797922	0.10683102	0.14992951	0.14992951	0.15049440
$\delta_F(P, 0, 3)$	0.17943541	0.17354340	0.27433442	0.27433442	0.27996519
$\delta_F(P, 0, 4)$	0.83444419	0.65889251	1.41424988	1.41424988	1.41894440
$\delta_F(P, 0, 5)$	0.90827444	0.75348431	1.46326471	1.46326471	1.47185479
$\delta_F(P, 0, 6)$	0.99263034	0.85789363	1.66359899	1.66359899	1.71794745

Table 4.8.1: Comparison of upper and lower bounds with the distance $\delta_F(P, 0, r)$ calculated by BFGS and `globalsearch.m` for Example 4.8.1.

Distance measured	Lower bounds (Theorem 4.5.2)	Lower bounds (Theorem 4.5.1)	Estimated distance BFGS	Estimated distance `globalsearch`	Upper bound (Theorem 4.6.1)
$\delta_F(P, 1, 2)$	1.35798224	0.70551994	1.35814780	1.35814780	1.39370758
$\delta_F(P, 1, 3)$	1.35690676	0.57675049	1.42078740	1.42078740	1.57015806
$\delta_F(P, 1, 4)$	1.35798160	0.56881053	1.42220397	1.42220397	1.65060167
$\delta_F(P, 1, 5)$	1.35689708	0.56908887	1.45865399	1.45865399	1.68613249
$\delta_F(P, 1, 6)$	1.35690633	0.56789237	1.46349849	1.46349849	1.57008146

Table 4.8.2: Comparison of upper and lower bounds with the distance $\delta_F(P, 1, r)$ calculated by BFGS and `globalsearch.m` for Example 4.8.1.

Example 4.8.2. Consider the matrix polynomial

$$P(\lambda) = \begin{bmatrix} 2.7694 & 0.7254 & -0.2050 \\ -1.3499 & -0.0631 & -0.1241 \\ 3.0349 & 0.7147 & 1.4897 \end{bmatrix} + \lambda \begin{bmatrix} 1.4090 & -1.2075 & 0.4889 \\ 1.4172 & 0.7172 & 1.0347 \\ 0.6715 & 1.6302 & 0.7269 \end{bmatrix} + \lambda^2 \begin{bmatrix} -0.3034 & 0.8884 & -0.8095 \\ 0.2939 & -1.1471 & -2.9443 \\ -0.7873 & -1.0689 & 1.4384 \end{bmatrix}.$$

Table 4.8.7 and Table 4.8.8 record the computed values of the distances $\delta_F(P, 0, r)$ and $\delta_F(P, -1, r)$ respectively obtained via BFGS and `globalsearch.m` algorithms

Distance measured	Lower bounds (Theorem 4.5.2)	Lower bounds (Theorem 4.5.1)	Estimated Distance `globalsearch`	Upper bound (Theorem 4.6.1)
$\delta_2(P,0,2)$	0.10797922	0.10683102	0.10797922	0.11368413
$\delta_2(P,0,3)$	0.17943541	0.17354340	0.19516022	0.21687613
$\delta_2(P,0,4)$	0.83444419	0.65889251	1.04384908	1.05026516
$\delta_2(P,0,5)$	0.90827444	0.75348431	1.13265966	1.17968107
$\delta_2(P,0,6)$	0.99263034	0.85789363	1.36385207	1.61899945

Table 4.8.3: Comparison of upper and lower bounds with the distance $\delta_2(P,0,r)$ calculated by `globalsearch.m` for Example 4.8.1.

Distance measured	Lower bound (Theorem 4.5.2)	Lower bound (Theorem 4.5.1)	Estimated Distance `globalsearch`	Upper bound (Theorem 4.6.1)
$\delta_2(P,1,2)$	1.35798224	0.70551994	1.35798224	1.35827634
$\delta_2(P,1,3)$	1.35690676	0.57675049	1.35805109	1.35813196
$\delta_2(P,1,4)$	1.35798160	0.56881053	1.35805159	1.37779220
$\delta_2(P,1,5)$	1.35689708	0.56908887	1.35805160	1.41164731
$\delta_2(P,1,6)$	1.35690633	0.56789237	1.35805609	1.43743061

Table 4.8.4: Comparison of upper and lower bounds with the distance $\delta_2(P,1,r)$ calculated by `globalsearch.m` for Example 4.8.1.

Distance measured	Lower bound (Theorem 4.5.2)	Upper bound
$\delta_{2,\infty}(P,0,2)$	0.10153149	0.10706018
$\delta_{2,\infty}(P,0,3)$	0.14545520	0.16793005
$\delta_{2,\infty}(P,0,4)$	0.55881158	0.84776583
$\delta_{2,\infty}(P,0,5)$	0.56734597	0.75348692
$\delta_{2,\infty}(P,0,6)$	0.62408329	1.17976723

Table 4.8.5: Upper and lower bounds of $\delta_{2,\infty}(P,0,r)$ for Example 4.8.1.

Distance measured	Lower bound (Theorem 4.5.2)	Upper bound
$\delta_{2,\infty}(P,1,2)$	0.67911046	0.69280941
$\delta_{2,\infty}(P,1,3)$	0.67826881	0.78460615
$\delta_{2,\infty}(P,1,4)$	0.67836252	0.78885169
$\delta_{2,\infty}(P,1,5)$	0.67837715	0.73067901
$\delta_{2,\infty}(P,1,6)$	0.67960385	0.85536125

Table 4.8.6: Upper and lower bounds of $\delta_{2,\infty}(P,1,r)$ for Example 4.8.1.

for all possible values of r together with the upper and lower bounds. The corresponding quantities for the distances $\delta_2(P,0,r)$ and $\delta_2(P,-1,r)$ are recorded in Table 4.8.9 and Table 4.8.10 respectively except that in these cases the computed value of the distance is obtained only via the `globalsearch.m` algorithm. Tables 4.8.11 and 4.8.12 records the bounds for the distances $\delta_{2,\infty}(P,0,r)$ and $\delta_{2,\infty}(P,-1,r)$ for all possible values of $r>1$. The strategies for computing them are the same as those in Example 4.8.1.

Distance	Lower bound		Estimated distance		Upper bound
measured	(Theorem 4.5.2)	(Theorem 4.5.1)	BFGS	`globalsearch`	(Theorem 4.6.1)
$\delta_F(P,0,2)$	0.25800277	0.25750097	0.25904415	0.25904415	0.26879600
$\delta_F(P,0,3)$	0.43621850	0.38556596	0.69617957	0.69617957	0.82200773
$\delta_F(P,0,4)$	0.88752500	0.83727454	1.84231345	1.84231345	2.04437686
$\delta_F(P,0,5)$	1.19949290	1.13421484	1.84468801	1.84468801	2.43953618
$\delta_F(P,0,6)$	1.28885600	1.07999296	2.60665217	2.60665217	2.73090090

Table 4.8.7: Comparison of upper and lower bounds with the distance $\delta_F(P,0,r)$ calculated by BFGS and `globalsearch.m` for Example 4.8.2.

Distance	Lower bound		Estimated distance		Upper bound
measured	(Theorem 4.5.2)	(Theorem 4.5.1)	BFGS	`globalsearch`	(Theorem 4.6.1)
$\delta_F(P,-1,2)$	0.99413714	0.49049043	1.14436402	1.14436402	1.14869786
$\delta_F(P,-1,3)$	1.23816383	0.57712979	2.22703947	2.22703947	2.37565159
$\delta_F(P,-1,4)$	1.33820455	0.56416354	2.33112163	2.33112163	2.51177974
$\delta_F(P,-1,5)$	1.36050277	0.59682624	2.44152499	2.44152500	2.70800270
$\delta_F(P,-1,6)$	1.46702487	0.61024547	2.62503371	2.66473662	2.90151393

Table 4.8.8: Comparison of upper and lower bounds with the distance $\delta_F(P,-1,r)$ calculated by BFGS and `globalsearch.m` for Example 4.8.2.

In almost every case the lower bound from Theorem 4.5.2 is better than the lower bound from Theorem 4.5.1. The perturbations $\Delta P(\lambda)$ constructed to find the upper bound in Theorem 4.6.1 may also be obtained by using nonzero singular values of $T_\gamma(P,\lambda_0,r)$ other than $f(\gamma)$ and a corresponding pair of left and right singular vectors. However the resulting upper bound obtained by taking the infimum of $|||\Delta P|||_s$, $s=2$ or F over all permissible γ does not seem to be an improvement over the one already obtained. For instance in Example 4.8.1, the matrix $T_\gamma(P,0,3)$

Distance measured	Lower bound (Theorem 4.5.2)	Lower bound (Theorem 4.5.1)	Estimated distance globalsearch	Upper bound (Theorem 4.6.1)
$\delta_2(P,0,2)$	0.25800277	0.25750097	0.25802614	0.25817920
$\delta_2(P,0,3)$	0.43621850	0.38556596	0.47215132	0.58937606
$\delta_2(P,0,4)$	0.88752500	0.83727454	1.11581046	1.57310992
$\delta_2(P,0,5)$	1.19949290	1.13421484	1.28762926	1.77192528
$\delta_2(P,0,6)$	1.28885600	1.07999296	1.80408754	2.39304209

Table 4.8.9: Comparison of upper and lower bounds with the distance $\delta_2(P,0,r)$ calculated by `globalsearch.m` for Example 4.8.2.

Distance measured	Lower bound (Theorem 4.5.2)	Lower bound (Theorem 4.5.1)	Estimated distance globalsearch	Upper bound (Theorem 4.6.1)
$\delta_2(P,-1,2)$	0.99413714	0.49049043	0.99413725	1.08915666
$\delta_2(P,-1,3)$	1.23816383	0.57712979	1.44791995	1.95311418
$\delta_2(P,-1,4)$	1.33820455	0.56416354	1.49553254	1.92175876
$\delta_2(P,-1,5)$	1.36050277	0.59682624	1.69980154	2.04689029
$\delta_2(P,-1,6)$	1.46702487	0.61024547	1.76475401	2.57442518

Table 4.8.10: Comparison of upper and lower bounds with the distance $\delta_2(P,-1,r)$ calculated by `globalsearch.m` for Example 4.8.2.

Distance measured	Lower bound (Theorem 4.5.2)	Upper bound
$\delta_{2,\infty}(P,0,2)$	0.25598106	0.25802614
$\delta_{2,\infty}(P,0,3)$	0.34805511	0.47215132
$\delta_{2,\infty}(P,0,4)$	0.67706047	1.02829835
$\delta_{2,\infty}(P,0,5)$	0.73910153	1.16000343
$\delta_{2,\infty}(P,0,6)$	0.86228680	1.55008462

Table 4.8.11: Upper and lower bounds of $\delta_{2,\infty}(P,0,r)$ for Example 4.8.2.

Distance measured	Lower bound (Theorem 4.5.2)	Upper bound
$\delta_{2,\infty}(P,-1,2)$	0.61030513	0.85449907
$\delta_{2,\infty}(P,-1,3)$	0.73192090	1.35676433
$\delta_{2,\infty}(P,-1,4)$	0.81667056	1.33897706
$\delta_{2,\infty}(P,-1,5)$	0.82859967	1.26177921
$\delta_{2,\infty}(P,-1,6)$	0.87090441	1.54433050

Table 4.8.12: Upper and lower bounds of $\delta_{2,\infty}(P,-1,r)$ for Example 4.8.2.

corresponding to the distance $\delta_2(P,0,4)$ is of size 8 and the upper bound from Theorem 4.6.1 reported in Table 4.8.3 is constructed by using $\sigma_5(T_\gamma(P,0,3))$ and it corresponding left and right singular vectors. If the same bound is constructed by considering the three smallest singular values $\sigma_6(T_\gamma(P,0,3))$, $\sigma_7(T_\gamma(P,0,3))$ and $\sigma_8(T_\gamma(P,0,3))$ and corresponding left and right singular vectors, then the values are 1.55784600, 1.65319413 and 2.42096365 respectively. Similar observations have been made by considering the other singular value of $T_\gamma(P,0,3)$.

Under certain assumptions, a singular value optimization characterization of the distance to a nearest matrix polynomial having an eigenvalue of specified algebraic multiplicity was obtained in [50, Corollary 3.4] with the restriction that only the constant coefficient matrix is perturbed. An upper bound of $\delta_2(P,\lambda_0,r)$ follows from this result. We computed it via `globalsearch.m` for Examples 4.8.1 and 4.8.2 and observed that the computed distance and upper bounds reported in the tables are better than this upper bound. For example the upper bounds from [50] for $\delta_2(P,0,2)$ and $\delta_2(P,1,3)$ in Example 4.8.1 are 0.20532034 and 8.59203088 respectively and those for $\delta_2(P,0,5)$ and $\delta_2(P,-1,4)$ in Example 4.8.2 are 2.1468713 and 4.67812467 respectively. In fact we find that the larger the value of r, the more the computed value of the upper bound from [50] than the ones derived in this work.

The next example provides a visual comparison of the various bounds of the distances $\delta_s(P,0,2)$ for $s=2$ or F.

Example 4.8.3. Consider the matrix pencil

$$P(\lambda)=\begin{bmatrix}1.8168 & & \\ & 0.8389 & \\ & & 0\end{bmatrix}+\lambda\begin{bmatrix}-1.7373 & -0.1861 & -0.1917\\ 0.8606 & -0.4003 & -0.0546\\ -0.7434 & -0.3821 & 0.1860\end{bmatrix}.$$

Table 4.8.13 gives the bounds on $\delta_s(P,0,2), s=2,F$. In this example the bounds in the literature are conservative with respect to the bounds in Theorem 4.5.1, Theorem 4.7.1 and Theorem 4.6.1. Further, the upper bounds in the literature are also conservative in comparison to the distance with respect to rank one perturbations on $P(\lambda)$ (Theorem 4.7.7). The value of $\delta_F(P,0,2)$ calculated via `globalsearch.m` and BFGS are all equal to the upper bound in Theorem 4.6.1 and slightly lower than the upper bound obtained from Theorem 4.7.7 and much lower than the upper bound from the literature. We present the variation of the bounds with respect to γ for this example. In the figures, the horizontal-axis covers an interval $(0,\, a]$ such that $\hat{\gamma}$ as specified in Proposition 4.7.2 belongs to this interval. The lower bound from Theorem 4.5.2 does not hold in this case as 0 is already an eigenvalue of $P(\lambda)$. Figure 4.8.1 shows the variation of the bounds (upper and lower) for $\delta_F(P,0,2)$. Similarly, Fig. 4.8.2 shows the variation of the upper and lower bounds for $\delta_2(P, 0, 2)$.

Distance measured	Lower Bound (Theorem 4.7.1)	Lower Bound (Theorem 4.5.1)	Lower Bound Papathanasiou et al. [66]	Upper Bound (Theorem 4.6.1)	Upper Bound Papathanasiou et al. [66]	Upper bound Kressner et al. [53]	Rank one Perturbations (Theorem 4.7.7)	Global-search	BFGS
$\delta_F(P, 0, 2)$	0.135	0.1122	0.0995	0.1567	0.31	-	0.1581	0.1567	0.1567
$\delta_2(P, 0, 2)$	-	0.1122	0.0995	0.1148	0.2195	0.2195	-	0.1137	0.1136

Table 4.8.13: Comparison of upper and lower bounds and the calculated value of $\delta_s(P, 0, 2)$ using BFGS and `globalsearch.m` with the upper and lower bounds available in literature for Example 4.8.3.

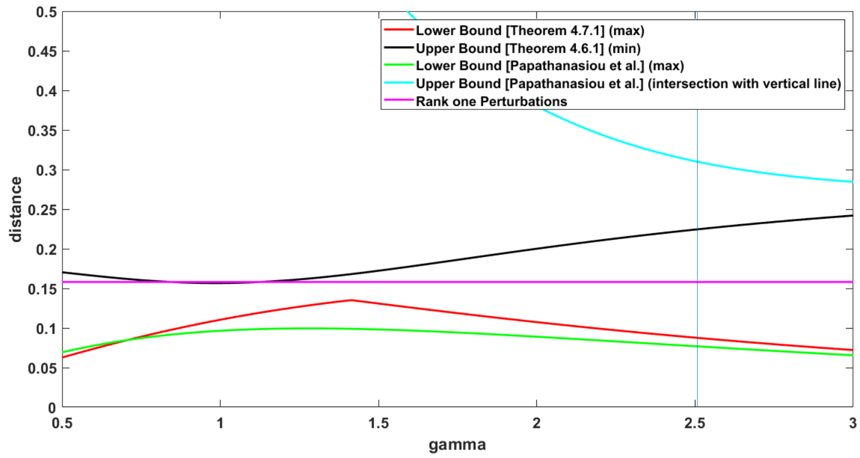

Figure 4.8.1: Variation of upper and lower bounds of $\delta_F(P, 0, 2)$ for Example 4.8.3.

Example 4.8.4. The following example is for the case when all the coefficient of the matrix polynomial are Hermitian and the distance is measured via Hermitian perturbations as in Theorem 4.7.10. Consider the matrix pencil

$$P(\lambda) = \begin{bmatrix} 1.8234 & & \\ & 0.2198 & \\ & & 0 \end{bmatrix} + \lambda \begin{bmatrix} -1.7435 & 1.5035 & 0.5159 \\ 1.5035 & -0.7858 & -1.2683 \\ 0.5159 & -1.2683 & 0.3103 \end{bmatrix}.$$

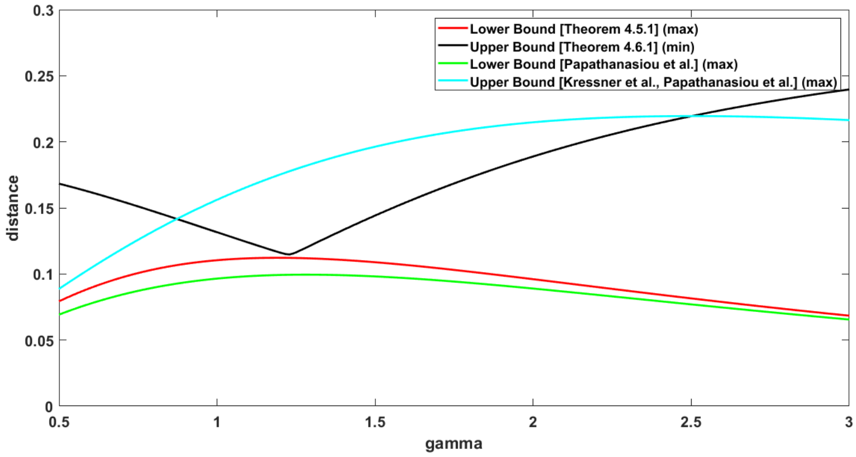

Figure 4.8.2: Variation of upper and lower bounds of $\delta_2(P, 0, 2)$ for Example 4.8.3.

Since A_1 is indefinite, by Theorem 4.7.10 the distance is given by

$$\sqrt{\sup_{t\in[t_1,t_2]} \lambda_{\min}(A_0^*A_0 + tA_1)},$$

where $t_1 = -2.7315$ and $t_2 = 1.0258$. By CVX algorithm [35, 36],

$$\hat{\delta}^H_{2,\infty}(P, 0, 2) = 0.02571369,$$

and the supremum is attained at $t_0 = -4.11354 \times 10^{-3}$. Now

$$x = \begin{bmatrix} -0.00085315 & 0.11677117 & 0.99315848 \end{bmatrix}^T$$

is an eigenvector corresponding to $\lambda_{\min}(A_0^*A_0 + t_0A_1)$ and x and A_0x are linearly independent. Hence by (4.7.10), the distance is attained by

$$\Delta A_0 = \begin{bmatrix} 8.43705 & 22.489 & 1563.71 \\ 22.489 & -3006.73 & -25489.6 \\ 1563.71 & -25489.6 & 2998.29 \end{bmatrix} \times 10^{-6}.$$

Conclusion. In this chapter we have characterized the matrix polynomials having an eigenvalue $\lambda_0 \in \mathbb{C}$ of algebraic multiplicity greater than or equal to a specified number. Using this characterization we have obtained upper and lower bounds and an optimization formulation of the distance to a nearest matrix polynomial having a specified elementary divisor in various norm settings. For some special cases we have found closed form solutions of the distance. We have also thrown light on the computational complexity of the distance calculation. Finally we have compared the computed distances with the bounds.

5 Nearest rank deficient matrix polynomials

A square regular matrix polynomial is nonsingular over the field of rational functions and in applications and theory, it plays a role similar to the square nonsingular scalar matrices which have many desirable properties. As already noted in Chapter 1 if E and A are $n \times n$ matrices such that $A - \lambda E$ is a regular matrix pencil, then the linear time invariant control system

$$E\dot{x}(t) = Ax(t) + Bu(t), \quad y(t) = Cx(t) + Du(t),$$

where B, C and D are $n \times m$, $p \times n$ and $p \times m$ respectively, is solvable for all choices of initial vectors x_0 and sufficiently smooth control $u(t)$. Moreover the solution $x(t)$ depends uniquely on the initial vector x_0 if the conditions are consistent and the transfer function $G(s) = C(A - sE)^{-1}B$ exists for all but finitely many choices of scalars s. The proximity of a nonsingular scalar matrix to a singular matrix is important as even if a square matrix is nonsingular but close to a singular matrix, then it is an ill conditioned matrix that poses several numerical challenges. Similarly the proximity of a regular matrix polynomial to a non-regular one is a theoretically intriguing question which is also of great importance in applications. For example if a linear time invariant system is associated with a regular matrix pencil that is close to a non-regular pencil, then the system will be unstable and modelling and other approximation errors may result in a system associated with a non-regular pencil (see for example [8, 13, 55, 56, 68] and references therein). Since matrix polynomials that are not regular are also referred to as singular matrix polynomials, this distance is also referred to as the distance to singularity (for matrix polynomials). However unlike its counterpart for the scalar matrices, the problem of finding the distance to singularity is not as well understood. In fact it is a long standing open problem [12].

The first work that provided significant insight into this problem for matrix pencils is [12] where the distance was formulated via two different optimizations and several upper and lower bounds were provided. Subsequently, [63] considered this distance for matrix pencils under the restriction that the perturbations that result in singular matrix pencils are low rank. In the more recent work [38], a differential equations approach has been taken for computing the distance to a nearest singular matrix pencil.

This chapter considers the general problem of finding the distance from a given matrix polynomial to a matrix polynomial with a specified upper bound on its normal rank with particular emphasis on the distance to singularity. Except for [30, 31], so far this general problem has not been considered in the literature. Moreover the algorithms in [30, 31] for computing the said distances are only with respect to the norm $|||\cdot|||_F$ under the assumption that the matrix polynomial is close to a singular matrix polynomial. In contrast the distances considered in this work are

$$\delta_s^r(P) := \inf\{|||\Delta P|||_s : \mathrm{nrank}(P + \Delta P) \leqslant r\}, s = 2, F; r < n,$$

without any assumptions on $P(\lambda)$.

As in Chapter 3 and Chapter 4, block Toeplitz matrices (1.3.1) and (1.3.2) associated with $P(\lambda)$ play an important role in computing these distances. It is established that an $n \times n$ matrix polynomial $P(\lambda)$ of degree k has normal rank at most r if and only if its level kr convolution matrix $C_{kr}(P)$ is sufficiently rank deficient. For the special case that $r = n - 1$, an additional characterization involving the suitable rank deficiency of the matrix $\bar{T}_{kn+1}(P)$ is obtained. Therefore the distance problem is equivalent to finding smallest structure preserving perturbations of these block Toeplitz matrices that cause the prescribed rank deficiencies to occur. These results are used to formulate the problems in terms of optimizations and to obtain various upper and lower bounds. In particular, the distance to singularity is formulated in terms of three different optimizations. An important consequence is the fact that computing these distances is equivalent to computing the reciprocal of a structured singular value or μ-value when $r = n-1$ and a generalized structured singular value in other cases. Therefore computing the distance to singularity may be an NP-hard problem. A key feature of this work is the additional insight obtained about the minimal indices of nearest singular matrix polynomials. This is helpful in interpreting the numerical results and formulating appropriate initialization strategies when performing the optimizations. Additionally, the specific connections of the optimizations that compute the distance to singularity with the optimizations that solve the problem of finding a nearest matrix polynomial with a Jordan chain of prescribed minimal length at zero are also revealed. Finally in the numerical experiments section optimizations via the Broyden Fletcher Goldfarb Shanno (BFGS) and Matlab's `globalsearch.m` algorithms are performed to compute the distances for examples that have already been considered in the literature. They are shown to compare favourably with the values obtained in literature. The upper and lower bounds for these examples are also seen to be tight in many cases.

5.1 Characterization of matrix polynomials of normal rank at most r

Let $Q(\lambda) = \sum_{i=0}^{k} \lambda^i B_i$ be any $n \times n$ matrix polynomial of grade k, where any of the coefficient matrices can be zero. Recall that $C_\varepsilon(Q)$ denotes the convolution matrix of $Q(\lambda)$ of level ε as in (1.3.1). If $Q(\lambda)$ is a singular matrix polynomial with normal rank at most r, then the vectors in the null spaces of $C_\varepsilon(Q), \varepsilon = 0, 1, \ldots$ determine the degrees of the vector polynomials in the right null space $N_r(Q)$. For example $N_r(Q)$ has a vector polynomial $\sum_{i=0}^{j} \lambda^i x_i$ of degree j if and only if $C_j(Q)x = 0$ for $x = [x_0^T, \cdots, x_j^T]^T$ where $x_i \in \mathbb{C}^n, i = 0, \ldots, j$, with $x_j \neq 0$. Similarly the degrees of the vector polynomials in $N_l(Q)$ are determined by the null spaces of the the convolution matrices $C_\varepsilon(Q^T), \varepsilon = 0, 1, \ldots$. If j is the smallest right (left) minimal index of $Q(\lambda)$ then $C_\varepsilon(Q)$ $(C_\varepsilon(Q^T))$ is nonsingular for $\varepsilon = 0, \ldots, j-1$ and $C_j(Q)x = 0$ $(C_j(Q^T)x = 0)$ for some $x = \begin{bmatrix} x_0^T & \cdots x_j^T \end{bmatrix}^T$ where $x_i \in \mathbb{C}^n, 0 \leqslant i \leqslant j$ with $x_0 \neq 0$ and $x_j \neq 0$.

Suppose that the normal rank of $Q(\lambda)$ is at least r for some positive integer r and consider $C_{kr}(Q) \in \mathbb{C}^{((kr+1)n+kn)\times((kr+1)n)}$. The main aim of this section is to characterize all $n \times n$ matrix polynomials of grade k and normal rank at most r by showing that they are equal to the set of $n \times n$ matrix polynomials of grade k with the property that rank $C_{kr}(Q) \leqslant (kr+1)r + kr$. We begin with some important concepts that are necessary to prove the main result.

Definition 5.1.1. *Given $y = \begin{bmatrix} y_0^T & y_1^T & \ldots & y_w^T \end{bmatrix}^T$, with $y_i \in \mathbb{C}^n, 0 \leqslant i \leqslant w$, if $y_{w-j} = 0$ for at least $j = 0, \ldots, i-1$, then the i^{th} shift y^i of y is defined to be the vector* $y^i = \begin{bmatrix} \underbrace{0^T \ \cdots \ 0^T}_{i \text{ vectors}} & y_0^T & y_1^T & \cdots & y_{w-i}^T \end{bmatrix}^T$.

Also, if p is the largest value of j such that $y_{w-j+1} = 0$ for all $j = 1, \ldots, p$, then we say that the maximum number of shifts allowed in y is p.

Observe that for $y = \begin{bmatrix} y_0^T & y_1^T & \ldots & y_w^T \end{bmatrix}^T$, with $y_i \in \mathbb{C}^n$, if $y_{w-1} = y_w = 0$, then the first shift of y is $y^1 := \begin{bmatrix} 0^T & y_0^T & y_1^T & \dots & y_{w-1}^T \end{bmatrix}^T$, and the 2-nd shift y^2 is first shift of y^1. In general for $2 \leqslant i \leqslant w$, if the i^{th} shift y^i is defined, then it is the first shift of y^{i-1}. Also clearly for any nonzero y of the given form, the maximum number of shifts possible is w.

A key concept for proving the main result of this section is the index sum theorem established by De Téran et al. in [18, Theorem 6.5]. It states that the sum of the partial multiplicities of all finite and infinite eigenvalues and right and left minimal indices of any $n \times n$ matrix polynomial of grade k and normal rank r cannot exceed kr. Consequently, the degree of any element of a right minimal basis of $Q(\lambda)$ cannot exceed kr. This fact is important for proving the following lemmas and the main result given in Theorem 5.1.4.

Lemma 5.1.2. *Let $Q(\lambda)$ be an $n \times n$ matrix polynomial with* $\operatorname{nrank} Q = r - j$, $0 \leqslant j \leqslant r$, *of grade k. If $\{x_1(\lambda), \dots, x_{n-r+j}(\lambda)\}$ is a minimal basis of $N_r(Q)$ where $x_i(\lambda) = \sum_{t=0}^{kr} \lambda^t x_{i,t}$ for $i \in \{1, \dots, n-r+j\}$, then $y_i := \begin{bmatrix} x_{i,0}^T & x_{i,1}^T & \dots & x_{i,kr}^T \end{bmatrix}^T$ for $i = 1, \dots, n-r+j$ and all their possible shifts form a linearly independent set.*

Proof. Suppose l_i shifts are possible on y_i, $i = 1, \dots, n-r+j$. We have to show that the set $\mathcal{A} = \{y_1, y_1^1, \dots, y_1^{l_1}, y_2, y_2^1, \dots, y_2^{l_2}, \dots, y_{n-r+j}, y_{n-r+j}^1, \dots, y_{n-r+j}^{l_{n-r+j}}\}$ is linearly independent. If possible let the set $\mathcal{A}$ be linearly dependent. Then there exists $\alpha_{(i,t)} \in \mathbb{C}$ for $i = 0, \dots, n-r+j$ and $t = 0, 1, \dots, l_i$ not all zero such that

$$\alpha_{(1,0)} y_1 \dots + \alpha_{(1,l_1)} y_1^{l_1} + \dots + \alpha_{(n-r+j,0)} y_{n-r+j} + \dots + \alpha_{(n-r+j,l_{n-r+j})} y_{n-r+j}^{l_{n-r+j}} = 0. \quad (5.1.1)$$

Multiplying (5.1.1) from the left by the matrix $\begin{bmatrix} I_n & \lambda I_n & \dots & \lambda^{kr} I_n \end{bmatrix}$ we get,

$$\left(\sum_{i=0}^{l_1} \lambda^i \alpha_{(1,i)} \right) x_1(\lambda) + \dots + \left(\sum_{i=0}^{l_{n-r+j}} \lambda^i \alpha_{(n-r+j,i)} \right) x_{n-r+j}(\lambda) = 0$$

which implies linear dependency of $\{x_1(\lambda), \dots, x_{n-r+j}(\lambda)\}$, a contradiction to the fact that it is a minimal basis of $N_r(Q)$. Hence the proof. □

Lemma 5.1.3. *Let $Q(\lambda)$ be any $n \times n$ matrix polynomial of grade k. For any non negative integer m, at least one of the following alternatives hold for the convolution matrix $C_m(Q)$.*

(a) $C_m(Q)$ is of full rank.

(b) *There exists a spanning set of the null space of $C_m(Q)$ of the form*

$$\{y_1, y_1^1, \ldots, y_1^{t_1}, y_2, y_2^1, \ldots, y_2^{t_2}, \ldots, y_q, y_q^1, \ldots, y_q^{t_q}\}$$

where $y_i = \begin{bmatrix} y_{i,0}^T & y_{i,1}^T & \ldots & y_{i,m}^T \end{bmatrix}^T$ *and* $y_{i,j} \in \mathbb{C}^n, 0 \leqslant j \leqslant m$ *such that* $\{y_{1,0}, \ldots, y_{q,0}\}$ *is linearly independent.*

Proof. The proof is completed by showing that if statement (a) does not hold, then statement (b) holds. Suppose $C := C_m(Q)$ is rank deficient and let $N(C)$ denote its null space. For any vector $z \in N(C)$ we partition z as $z = \begin{bmatrix} z_0^T & z_1^T & \ldots & z_m^T \end{bmatrix}^T$ where $z_j \in \mathbb{C}^n \ \forall \ j = 0, \ldots, m$. We describe the process of finding a spanning set of $N(C)$ with the desired properties in a stepwise manner. If any of the steps is not possible, then it is skipped.

Step 1: If possible, choose a vector $y_1 = \begin{bmatrix} y_{1,0}^T & 0^T & \cdots & 0^T \end{bmatrix}^T \in N(C)$. Then all m shifts of y_1 belongs to $N(C)$. Consider $\mathcal{B} = \{y_1, y_1^1, \ldots, y_1^m\}$. Again if possible, choose a vector $y_2 = \begin{bmatrix} y_{2,0}^T & 0^T & \cdots & 0^T \end{bmatrix}^T \in N(C)$ such that $y_2 \notin \operatorname{span}(\mathcal{B})$. Then all m shifts of y_2 belongs to $N(C)$. Include all of them in $\mathcal{B}$. Continue this process till all the vectors of type $z = \begin{bmatrix} 0^T & \cdots & 0^T & z_i^T & 0^T & \cdots & 0^T \end{bmatrix}^T$ for some i are included in $\operatorname{span}(\mathcal{B})$. Assuming that s_1 repetitions are required to complete this process, after 1st step we have $\mathcal{B} = \{y_1, y_1^1, \ldots, y_1^m, \ldots, y_{s_1}, y_{s_1}^1, \ldots, y_{s_1}^m\}$ such that $\{y_{1,0}, \ldots, y_{s_1,0}\}$ is a linearly independent set.

Setting $S_p = \sum_{i=1}^p s_i$, suppose that after $(j-1)$th step we get

$$\mathcal{B} = \{y_1, \ldots, y_1^m, \ldots, y_{s_1}, \ldots, y_{s_1}^m, \ldots, y_{S_{j-2}+1}, \ldots, y_{S_{j-2}+1}^{m-j+2}, \ldots, y_{S_{j-1}}, \ldots, y_{S_{j-1}}^{m-j+2}\}$$

such that $\{y_{1,0}, \ldots, y_{s_1,0}, \ldots, y_{S_{j-2}+1,0}, \ldots, y_{S_{j-1},0}\}$ is linearly independent and $\mathcal{B}$ contains all the vectors of type

$$z = \begin{bmatrix} 0^T & \ldots & 0^T & z_i^T & \cdots & z_{i+j-2}^T & 0^T & \cdots & 0^T \end{bmatrix}^T$$

for some i.

Step j: If possible, choose a vector

$$y_{S_{j-1}+1} = \begin{bmatrix} y_{S_{j-1}+1,0}^T & \cdots & y_{S_{j-1}+1,j-1}^T & 0^T & \cdots & 0^T \end{bmatrix}^T \in N(C)$$

but $y_{S_{j-1}+1} \notin \operatorname{span}(\mathcal{B})$. Then all possible shifts of $y_{S_{j-1}+1}$ belongs to $N(C)$. Again, $\{y_{1,0}, \ldots, y_{s_1,0}, y_{S_{j-2}+1,0}, \ldots, y_{S_{j-1},0}, y_{S_{j-1}+1,0}\}$ is a linearly independent set as otherwise $y_{S_{j-1}+1,0} = c_1 y_{1,0} + \cdots + c_{S_{j-1}} y_{S_{j-1},0}$ for some $c_1, \ldots, c_{S_{j-1}} \in \mathbb{C}$ which implies

$$y_{S_{j-1}+1} - \sum_{i=1}^{S_{j-1}} c_i y_i = \begin{bmatrix} 0^T & \hat{y}_{S_{j-1}+1,1}^T & \cdots & \hat{y}_{S_{j-1}+1,j-1}^T & 0^T & \cdots & 0^T \end{bmatrix}^T \in \operatorname{span}(\mathcal{B}),$$

so $y_{S_{j-1}+1} \in \text{span}(\mathcal{B})$, which is a contradiction. Include all the possible shifts of $y_{S_{j-1}+1}$ to $\mathcal{B}$ and continue this process till all the vectors of type

$$z = \begin{bmatrix} 0^T & \cdots & 0^T & z_i^T & \cdots & z_{i+j-1}^T & 0^T & \cdots & 0^T \end{bmatrix}^T$$

for some i are included in $\mathcal{B}$. Suppose we have to repeat the process s_j times until the jth step is complete. Then we get a set

$$\mathcal{B} = \{y_1, \ldots, y_1^m, \ldots, y_{s_1}, \ldots, y_{s_1}^m, \ldots, y_{S_{j-1}+1}, \ldots, y_{S_{j-1}+1}^{m-j+1}, \ldots, y_{S_j}, \ldots, y_{S_j}^{m-j+1}\}$$

such that $\{y_{1,0}, \ldots, y_{s_1,0}, \ldots, y_{S_{j-1}+1,0}, \ldots, y_{S_j,0}\}$ are linearly independent.

Continue the process until $\text{span}(\mathcal{B}) = N(C)$. Suppose we achieve this after t steps and s_i is the number of repetitions required in the i^{th} step. Then we end up with a spanning set of $N(C)$ of the form $\{y_1, \ldots, y_1^m, \ldots, y_{s_1}, \ldots, y_{s_1}^m, \ldots, y_{S_t}, \ldots, y_{S_t}^{m-t+1}\}$ such that $\{y_{1,0}, \ldots, y_{s_1,0}, \ldots, y_{S_t,0}\}$ is a linearly independent set and this completes the proof. $\square$

Now we come to the main theorem of this section.

Theorem 5.1.4. *The normal rank of an $n \times n$ matrix polynomial $Q(\lambda)$ of grade k, is at most r if and only if*

$$\sigma_{(kr+1)(r+1)}(C_{kr}(Q)) = 0.$$

Proof. Let $\text{nrank}\, Q \leqslant r$. Then $\text{nullity}(Q(\lambda)) = n - r + j$ for $0 \leqslant j \leqslant r$. Let

$$\{x_1(\lambda), \ldots, x_{n-r+j}(\lambda)\} \text{ where } x_i(\lambda) = \sum_{t=0}^{kr} x_{i,t}\lambda^t,\ x_{i,t} \in \mathbb{C}^n$$

be a minimal basis of $N_r(Q)$ with $\deg x_i = d_i$ for $i = 1, \ldots, n-r+j$. Here we assume that $x_{i,t} = 0$ for all $t = d_i + 1, \ldots, kr$, $i = 1, \ldots, n - r + j$. Then at least $kr - d_i$ shifts are possible to $y_i = \begin{bmatrix} x_{i,0}^T & x_{i,1}^T & \cdots & x_{i,kr}^T \end{bmatrix}^T$ and from the index sum theorem $d_1 + \cdots + d_{n-r+j} \leqslant k(r - j)$. Set $C(Q) = C_{kr}(Q)$. Now from Lemma 5.1.2 it is clear that

$$\begin{aligned} \text{nullity}(C(Q)) &\geqslant (kr + 1 - d_1) + \cdots + (kr + 1 - d_{n-r+j}) \\ &= (kr + 1)(n - r + j) - (d_1 + \cdots + d_{n-r+j}) \\ &\geqslant (kr + 1)(n - r + j) - k(r - j) \\ &\geqslant (kr + 1)(n - r) - kr, \end{aligned}$$

which proves one part of the equivalence.

For proving the other part let $\sigma_{(kr+1)(r+1)}(C(Q)) = 0$. Then

$$\text{nullity}(C(Q)) = (kr+1)(n-r) - kr + j,$$

where $0 \leqslant j \leqslant kr(r+1) + r$. Let $N(C(Q))$ denote the null space of $C(Q)$. We partition any $y_i \in N(C(Q))$ as $y_i = \begin{bmatrix} y_{i,0}^T & \dots & y_{i,kr}^T \end{bmatrix}^T$, $y_{i,t} \in \mathbb{C}^n \ \forall \ t = 0, \dots, kr$. From the $(kr+1)(n-r)-kr+j$ linearly independent vectors in $N(C(Q))$ we construct at least $n-r$ linearly independent vector polynomials over the field $\mathbb{C}(\lambda)$ belonging to $N_r(Q)$. According to Lemma 5.1.3 there exists a spanning set of $N(C(Q))$, of the form

$$\mathcal{B} = \{y_1, y_1^1, \dots, y_1^{t_1}, y_2, y_2^1, \dots, y_2^{t_2}, \dots, y_q, y_q^1, \dots, y_q^{t_q}\}$$

such that $\{y_{1,0}, \dots, y_{q,0}\}$ is linearly independent. Now $q \geqslant n-r$ because if $q < n-r$ then

$$\dim(\text{span}(\mathcal{B})) \leqslant (kr+1)q \leqslant (kr+1)(n-r-1) < \dim(N(C(Q))),$$

leading to a contradiction. Thus we have a set of vectors $\{y_1, \dots, y_q\} \in N(C(Q))$ such that $\{y_{1,0}, \dots, y_{q,0}\}$ is linearly independent. Then $y_i(\lambda) := \sum_{j=0}^{kr} \lambda^j y_{i,j} \in N_r(Q)$ for each $i = 1, \dots, q$. To complete the proof we show that the set $\{y_1(\lambda), \dots, y_q(\lambda)\}$ is linearly independent over $\mathbb{C}(\lambda)$. If not, then

$$c_1(\lambda)y_1(\lambda) + \cdots + c_q(\lambda)y_q(\lambda) = 0 \tag{5.1.2}$$

where $c_i(\lambda)$ are polynomials not all zero and without loss of generality it may be assumed that $c_i(0) \neq 0$ for at least one $i, 1 \leqslant i \leqslant q$. Comparing the constant term on both sides of (5.1.2), $c_1(0)y_{1,0} + \cdots + c_q(0)y_{q,0} = 0$ thus contradicting the fact that $\{y_{1,0}, \dots, y_{q,0}\}$ is linearly independent set. Hence $\{y_1(\lambda), \dots, y_q(\lambda)\}$ is a linearly independent subset of $N_r(Q)$. As $q \geqslant n-r$, nullity$(Q(\lambda)) \geqslant q \geqslant n-r$ and the proof follows. □

5.2 Characterizations of singular matrix polynomials

As an immediate corollary of Theorem 5.1.4 we have the following characterization of a singular matrix polynomial.

Corollary 5.2.1. *The $n \times n$ matrix polynomial $Q(\lambda) = \sum_{i=0}^{k} \lambda^i B_i$ is singular if and only if the convolution matrix $C_{k(n-1)}(Q)$ is rank deficient.*

Remark 5.2.2. *The convolution matrix $C_j(Q)$ is rank deficient, then so is $C_{\hat{j}}(Q)$ where $\hat{j} > j$. Therefore Corollary 5.2.1 implies that $Q(\lambda)$ is a singular matrix polynomial if and only if any of the convolution matrices $C_j(Q), j = 0, \ldots, k(n-1)$, are rank deficient. Moreover, clearly the statement also holds if $C_j(Q)$ is replaced by $C_j(Q^T)$.*

An alternative characterization for a matrix polynomial $Q(\lambda)$ to be singular may be obtained via a suitable rank deficiency criterion of yet another block Toeplitz matrix $T(Q, 0, kn) := T_\gamma(Q, 0, kn)$ with $\gamma = [1 \cdots 1]$ associated with the matrix polynomial $Q(\lambda)$ where $T_\gamma(Q, 0, kn)$ is as defined in (4.2.2). Note that $\bar{T}_i(Q) = T(Q, 0, i-1)$ for $i = 1, 2, \ldots,$ where $\bar{T}_i(Q)$ is as defined in (1.3.2). The characterization is based on the following lemma that appears in the literature in various forms [2, 76]. Here it is restated in a form best suited for our purposes along with a short proof.

Lemma 5.2.3. *For a $n \times n$ matrix polynomial $Q(\lambda) = \sum_{i=0}^{k} \lambda^i B_i$ of grade k, let $\bar{r}_0 = 0$, $\bar{r}_1 = \operatorname{rank} \bar{T}_1(Q)$ and $\bar{r}_i = \operatorname{rank} \bar{T}_i(Q) - \operatorname{rank} \bar{T}_{i-1}(Q)$ for $i > 1$. Then*

1. *$(\bar{r}_i)$ is a nondecreasing sequence of non-negative integers.*

2. *$\bar{l}_i = \bar{r}_{i+1} - \bar{r}_i$ is the number of Jordan chains of length i of $Q(\lambda)$ corresponding to zero.*

3. *When $\bar{r}_i = \operatorname{nrank} Q$ for some i, all Jordan chains of $Q(\lambda)$ corresponding to zero are determined.*

Proof. Suppose there are p Jordan chains of $Q(\lambda)$ corresponding to zero of lengths $k_1 \leqslant k_2 \leqslant \cdots \leqslant k_p$. Let the j^{th} chain be denoted by $\{x_{j,0}, x_{j,2}, \ldots, x_{j,k_j-1}\}$ where $x_{j,i} \in \mathbb{C}^n, 0 \leqslant i \leqslant k_j - 1$ so that the set of eigenvectors corresponding to zero is $\{x_{1,0}, \ldots, x_{p,0}\}$. Again let $\operatorname{nrank} Q = r$ and $\{y_1(\lambda), \ldots, y_{n-r}(\lambda)\}$ be a minimal basis of right null space of $Q(\lambda)$ where each $y_j(\lambda) = \sum_{i=0}^{d_j} \lambda^i y_{j,i}$. Then clearly, $\{y_{10}, \ldots, y_{n-r,0}, x_{1,0}, \ldots, x_{p,0}\}$ is a linearly independent set. Therefore, the Jordan chains of length j corresponding to zero contribute j or i linearly independent vectors in the null space of $\bar{T}_i(Q)$ depending on $j < i$ or $j \geqslant i$ respectively and each $y_j(\lambda) \in N_r(Q)$ contributes i linearly independent vectors in the null space of $\bar{T}_i(Q)$. Let m_j and $\hat{m}_j$ represent the number of Jordan chains of length j and length at least j of $Q(\lambda)$ corresponding to zero respectively. Then

$$\operatorname{nullity} \bar{T}_i(Q) = m_1 + \cdots + (i-1)m_{i-1} + i\hat{m}_i + i(n-r).$$

Again,

$$\begin{aligned}\bar{r}_i &= \operatorname{rank}\bar{T}_i(Q) - \operatorname{rank}\bar{T}_{i-1}(Q)\\ &= n - (\operatorname{nullity}\bar{T}_i(Q) - \operatorname{nullity}\bar{T}_{i-1}(Q))\\ &= n - (\hat{m}_i + (n-r))\\ &= r - \hat{m}_i.\end{aligned}$$

Clearly from the above relation $(\bar{r}_i)$ is a nondecreasing sequence of non-negative integers.

Now $\bar{l}_i = \bar{r}_{i+1} - \bar{r}_i = r - \hat{m}_{i+1} - r + \hat{m}_i = m_i$, which shows that $\bar{l}_i$ is the number of Jordan chains of length i of $Q(\lambda)$ corresponding to zero.

Also clearly, as $\bar{r}_i = r - \hat{m}_i$, so $\bar{r}_i = r$ only when $\hat{m}_i = 0$ which implies that $Q(\lambda)$ cannot have any Jordan chain of length greater than or equal to i corresponding to zero if $\bar{r}_i = \operatorname{nrank} Q$. □

Theorem 5.2.4. *An $n \times n$ matrix polynomial $Q(\lambda) = \sum_{i=0}^{k} \lambda^i B_i$ of grade k is singular if and only if $\bar{T}_{kn+1}(Q)$ given by (1.3.2) has rank at most $(kn+1)(n-1)$.*

Proof. Let $\bar{T}_i(Q)$ be the block Toeplitz matrices as in (1.3.2). By Lemma 5.2.3, the number of Jordan chains of length i at 0 associated with $Q(\lambda)$ is $\bar{r}_{i+1} - \bar{r}_i$ where $\bar{r}_i = \operatorname{rank}\bar{T}_i(Q) - \operatorname{rank}\bar{T}_{i-1}(Q)$. Also as the longest such chain has length kn, $\bar{r}_j = \bar{r}_{j+1} = \operatorname{nrank} Q$ for $j \geqslant kn+1$. Therefore $Q(\lambda)$ is a singular matrix polynomial if and only if $\operatorname{rank}\bar{T}_{kn+1}(Q) - \operatorname{rank}\bar{T}_{kn}(Q) < n$. This is equivalent to the existence of a vector $\begin{bmatrix} x_0^T & \cdots & x_{kn}^T \end{bmatrix}^T$ in the null space of $\bar{T}_{kn+1}(Q)$ such that $x_i \in \mathbb{C}^n, i = 0, \ldots, kn$ and $x_0 \neq 0$. This implies that $\operatorname{rank}\bar{T}_{kn+1}(Q) \leqslant (kn+1)(n-1)$ as

$$\begin{bmatrix} 0 & \cdots & x_0^T \end{bmatrix}^T, \begin{bmatrix} 0 & \cdots & x_0^T & x_1^T \end{bmatrix}^T, \ldots, \begin{bmatrix} x_0^T & \cdots & x_{kn}^T \end{bmatrix}^T$$

are linearly independent vectors in the null space of $\bar{T}_{kn+1}(Q)$.

Conversely, let $\operatorname{rank}\bar{T}_{kn+1}(Q) < (kn+1)(n-1)$. Suppose if possible that $Q(\lambda)$ is a regular matrix polynomial. Then by arguing as in the proof of Theorem 4.2.1, 0 is an eigenvalue of $Q(\lambda)$ such that the sum of the lengths of all the Jordan chains corresponding to 0 is at least $kn+1$. Since this is impossible, it follows that $Q(\lambda)$ is a singular matrix polynomial. □

The above characterizations of singular matrix polynomials will be used to link the computation of the distance to singularity to structured and generalized structured singular values or μ values of certain matrices. Certain lower and upper bounds on the distance to singularity will also be derived from these characterizations.

5.3 Distance to nearest matrix polynomials of normal rank at most r

The characterization of matrix polynomials of normal rank at most r obtained in Theorem 5.1.4 is used to formulate the computation of $\delta_s^{(r)}(P)$ as an optimization problem in the following theorem.

Theorem 5.3.1. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $n \times n$ matrix polynomial of degree k and $r, 0 < r < n$ be a given integer. For $\phi := (kr+1)(n-r) - kr$, let $\mathcal{V}_\phi$ be the collection of all linearly independent subsets of $\mathbb{C}^{(kr+1)n}$ with ϕ elements. For any $\{z_1, \dots, z_\phi\} \in \mathcal{V}_\phi$, partition each z_i as $z_i = \begin{bmatrix} z_{i,0}^T & z_{i,1}^T & \cdots & z_{i,kr}^T \end{bmatrix}^T \in \mathbb{C}^{(kr+1)n}$, where $z_{i,j} \in \mathbb{C}^n$ for $j = 0, \dots, kr$, and define*

$$X(z_i) = \begin{bmatrix} z_{i,0} & z_{i,1} & \cdots & z_{i,k} & \cdots & \cdots & z_{i,kr} & & & \\ & z_{i,0} & z_{i,1} & \cdots & z_{i,k} & \cdots & \cdots & z_{i,kr} & & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & & z_{i,0} & z_{i,1} & \cdots & z_{i,k} & \cdots & \cdots & z_{i,kr} \end{bmatrix}.$$

Then for $s = 2$ or F

$$\delta_s^{(r)}(P) = \inf_{\{z_1,\dots,z_\phi\} \in \mathcal{V}_\phi} \left\| \begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix} \begin{bmatrix} X(z_1) \cdots X(z_\phi) \end{bmatrix} \begin{bmatrix} X(z_1) \cdots X(z_\phi) \end{bmatrix}^\dagger \right\|_s.$$

Proof. Let $\Delta P(\lambda)$ be any $n \times n$ matrix polynomial such that $\operatorname{nrank}(P + \Delta P) \leqslant r$. By Theorem 5.1.4, this is equivalent to $\operatorname{rank}(C_{kr}(P + \Delta P)) \leqslant (kr+1)r + kr$ and hence there are at least ϕ linearly independent vectors in the null space of $C_{kr}(P + \Delta P)$. Suppose $\{z_1, \dots, z_\phi\} \in \mathcal{V}_\phi$ is a subset of the null space of $C_{kr}(P + \Delta P)$ where each $z_i = \begin{bmatrix} z_{i,0}^T & z_{i,1}^T & \cdots & z_{i,kr}^T \end{bmatrix}^T$ with $z_{i,j} \in \mathbb{C}^n$ for $i = 1, \dots, \phi$, $j = 0, \dots, kr$. Then,

$$C_{kr}(P + \Delta P) \begin{bmatrix} z_1 & \cdots & z_\phi \end{bmatrix} = 0.$$

This is equivalent to

$$\begin{bmatrix} \Delta A_0 & \cdots & \Delta A_k \end{bmatrix} \begin{bmatrix} X(z_1) & \cdots & X(z_\phi) \end{bmatrix} = - \begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix} \begin{bmatrix} X(z_1) & \cdots & X(z_\phi) \end{bmatrix}.$$

By Theorem 1.5.1, a choice of $\begin{bmatrix} \Delta A_0 & \cdots & \Delta A_k \end{bmatrix}$ satisfying the above equation that is also minimal with respect to the 2 or Frobenius norm is given by

$$\begin{bmatrix} \Delta A_0 & \cdots & \Delta A_k \end{bmatrix} = - \begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix} \begin{bmatrix} X(z_1) & \cdots & X(z_\phi) \end{bmatrix} \begin{bmatrix} X(z_1) & \cdots & X(z_\phi) \end{bmatrix}^\dagger.$$

Hence

$$\delta_s^{(r)}(P) = \inf_{\{z_1,\ldots,z_\phi\}\in\mathcal{V}_\phi} \left\| \begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix} \begin{bmatrix} X(z_1)\cdots X(z_\phi) \end{bmatrix} \begin{bmatrix} X(z_1)\cdots X(z_\phi) \end{bmatrix}^{\dagger} \right\|_s,$$

for $s = 2$ or F. □

5.4 Distance to nearest singular matrix polynomials

Given a square regular matrix polynomial $P(\lambda)$, the distance $\delta_s^{(n-1)}(P)$ to a nearest singular matrix polynomial is an important special case of the distance to nearest matrix polynomials of normal rank at most r considered in the previous section. In this section we formulate the computation of $\delta_s^{(n-1)}(P)$ in terms of different optimization problems.

The characterization of singular matrix polynomials in Corollary 5.2.1 and the remark following it implies that for a given $\Delta P(\lambda) = \sum_{i=0}^{k} \lambda^i \Delta A_i$, $(P + \Delta P)(\lambda)$ is singular if either $C_j(P + \Delta P)$ or $C_j((P + \Delta P)^T)$ where $0 \leqslant j \leqslant k(n-1)$ is rank deficient for some j. For $j = 0, \ldots, k(n-1)$, and $s = 2, F$, let

$$\gamma_j^{(s)} = \inf\{\|\Delta P\|_s : C_j(P + \Delta P) \text{ is rank deficient }\}, \tag{5.4.1}$$

$$\eta_j^{(s)} = \inf\{\|\Delta P\|_s : C_j((P + \Delta P)^T) \text{ is rank deficient }\}. \tag{5.4.2}$$

Observe that $\gamma_j^{(s)}$ ($\eta_j^{(s)}$) computes the distance to a nearest singular matrix polynomial with a right (left) minimal index in the set $\{0, \ldots, j\}$. Also clearly

$$\gamma_0^{(s)} \geqslant \cdots \geqslant \gamma_{k(n-1)}^{(s)}, \text{ and } \eta_0^{(s)} \geqslant \cdots \geqslant \eta_{k(n-1)}^{(s)}.$$

We will refer to these as the γ-*sequence* and the η-*sequence* respectively. The following theorem gives the distance to a nearest singular matrix polynomial from $P(\lambda)$ and its connection with the right and left minimal indices of a nearest singular matrix polynomial.

Theorem 5.4.1. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $n \times n$ matrix polynomial of degree k. Suppose that its γ-sequence and η-sequence are given by (5.4.1) and (5.4.2) respectively with $s = 2$ or F. Let j_0 be the smallest index j such that $\gamma_{j_0}^{(s)} = \min_{0\leqslant j\leqslant k(n-1)} \gamma_j^{(s)}$ and i_0 be the smallest index i such that $\eta_{i_0}^{(s)} = \min_{0\leqslant i\leqslant k(n-1)} \eta_i^{(s)}$. Then the following hold.*

(a) If either j_0 or i_0 is equal to 0 or $k(n-1)$, then

$$\delta_s^{(n-1)}(P) = \min\{\gamma_0^{(s)}, \eta_0^{(s)}\}.$$

In particular,

$$\delta_F^{(n-1)}(P) = \min\left\{\sigma_{\min}\left(\begin{bmatrix} A_0^T & \cdots & A_k^T\end{bmatrix}^T\right), \sigma_{\min}\left(\begin{bmatrix} A_0 & \cdots & A_k\end{bmatrix}\right)\right\}.$$

(b) *There exists a matrix polynomial* $(P+\Delta P)(\lambda)$ *such that* $\|\Delta P\|_s = \delta_s^{(n-1)}(P)$ *with* j_0 (i_0) *as its least right (left) minimal index. Also the right (left) minimal indices of any other singular matrix polynomial* $Q(\lambda)$ *satisfying* $\|Q-P\|_s = \delta_s^{(n-1)}(P)$ *cannot be less than* j_0 (i_0).

(c) *The distance* $\delta_s^{(n-1)}(P) = \gamma_{j_0}^{(s)} = \eta_{i_0}^{(s)} = \gamma_{k(n-1)-i_0}^{(s)} = \eta_{k(n-1)-j_0}^{(s)}$.

(d) *The largest left (right) minimal index of any singular matrix polynomial* $(P+\Delta P)(\lambda) = \sum_{i=0}^{k} \lambda^i (A_i + \Delta A_i)$ *such that* $\|\Delta P\|_s = \delta_s^{(n-1)}(P)$ *is at most* $k(n-1) - j_0$ $(k(n-1) - i_0)$.

(e) *The distance* $\delta_s^{(n-1)}(P) = \min\left\{\gamma_q^{(s)}, \eta_q^{(s)}\right\}$, *where* $q := \left\lfloor \frac{k(n-1)}{2} \right\rfloor$. *Moreover,* $\delta_s^{(n-1)}(P) = \gamma_q^{(s)} = \eta_q^{(s)}$ *if and only if* $1 \leqslant i_0, j_0 \leqslant q$.

Proof. If $i_0 = 0$ or $j_0 = 0$ then part (a) clearly holds. Suppose that $j_0 = k(n-1)$. Then $\gamma_{k(n-1)-1}^{(s)} > \gamma_{k(n-1)}^{(s)} = \delta_F^{(n-1)}(P)$. Let $\Delta P(\lambda) = \sum_{i=0}^{k} \lambda^i \Delta A_i$ such that $\|\Delta P\|_s = \delta_s^{(n-1)}(P)$ and $C_{k(n-1)}(P+\Delta P)$ is rank deficient. For $x_i \in \mathbb{C}^n$, $i = 0, \ldots, k(n-1)$, let $x = \begin{bmatrix} x_0^T & \cdots & x_{k(n-1)}^T \end{bmatrix}^T \neq 0$ such that $C_{k(n-1)}(P+\Delta P)x = 0$. Then x_0 and $x_{k(n-1)}$ are non zero vectors as otherwise in the first case $\hat{x} = \begin{bmatrix} x_1^T & \cdots & x_{k(n-1)}^T \end{bmatrix}^T$ and in the second case $\hat{x} = \begin{bmatrix} x_0^T & \cdots & x_{k(n-1)-1}^T \end{bmatrix}^T$ are non zero vectors in the null space of $C_{k(n-1)-1}(P+\Delta P)$ implying that $j_0 = k(n-1)-1$ as $\gamma_{k(n-1)}^{(s)} = \|\Delta P\|_s \geqslant \gamma_{k(n-1)-1}^{(s)}$, which is impossible. Therefore $k(n-1)$ is a right minimal index of $(P+\Delta P)(\lambda)$ so that 0 is a left minimal index of $(P+\Delta P)(\lambda)$ and $\delta_s^{(n-1)}(P) = \|\Delta P\|_s = \eta_0^{(s)}$. The proof for the case that $i_0 = k(n-1)$ follows by replacing the convolution matrices in the preceding arguments by those with respect to $(P+\Delta P)(\lambda)^T$. In particular $\gamma_0^{(F)} = \sigma_{\min}\left(\begin{bmatrix} A_0^T & \cdots & A_k^T \end{bmatrix}^T\right)$ and $\eta_0^{(F)} = \sigma_{\min}\left(\begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix}\right)$. This completes the proof of part (a).

To prove part (b), let $\mathcal{S}$ be the collection of all singular matrix polynomials $(P+\Delta P)(\lambda) = \sum_{i=0}^{k} \lambda^i (A_i + \Delta A_i)$, such that $\|\Delta P\|_s = \delta_s^{(n-1)}(P)$. Clearly there exists a perturbation $\Delta P(\lambda)$ to $P(\lambda)$ such that $\|\Delta P\|_s = \gamma_{j_0}^{(s)}$. Then $(P+\Delta P)(\lambda)$ is singular and one of the numbers $\{0, \ldots, j_0\}$ is a right minimal index. Since $\gamma_{j_0}^{(s)} = \delta_s^{(n-1)}(P)$, $(P+\Delta P)(\lambda) \in \mathcal{S}$. If j_0 is not a right minimal index of $(P+\Delta P)(\lambda)$, then $C_j(P+\Delta P)$ is singular for some $j < j_0$ and $\gamma_j^{(s)} \leqslant \|\Delta P\|_s = \gamma_{j_0}^{(s)}$. But this contradicts the minimality of j_0. Suppose that there exists a matrix polynomial $(P+\widehat{\Delta P})(\lambda)$ in $\mathcal{S}$ such that its smallest right minimal index is $\hat{j}$ and $\hat{j} < j_0$. Then $C_{\hat{j}}(P+\widehat{\Delta P})x = 0$ for $x = \begin{bmatrix} x_0^T & \cdots & x_{\hat{j}}^T \end{bmatrix}^T$ where $x_i \in \mathbb{C}^n, i = 0, \ldots, \hat{j}$ and x_0 and $x_{\hat{j}}$ are nonzero. This is because otherwise $(P+\widehat{\Delta P})(\lambda)$ has a right minimal index that is strictly less then $\hat{j}$, which is impossible. Then $\gamma_{\hat{j}}^{(s)} \leqslant \|\widehat{\Delta P}\|_s = \gamma_{j_0}^{(s)}$ which again contradicts the minimality of j_0. This completes the proof of part (b).

From the definitions of j_0 and i_0 it is clear that $\delta_s^{(n-1)}(P) = \gamma_{j_0}^{(s)} = \eta_{i_0}^{(s)}$. Therefore to prove part (c), we establish that $\delta_s^{(n-1)}(P) = \gamma_{k(n-1)-i_0}^{(s)} = \eta_{k(n-1)-j_0}^{(s)}$. By part (b) there exist $\Delta P(\lambda) = \sum_{i=0}^{k} \lambda^i \Delta A_i$ such that $\|\Delta P\|_s = \delta_s^{(n-1)}(P)$ and $(P+\Delta P)(\lambda)$ has a right minimal index j_0. So $(P+\Delta P)(\lambda)$ has a vector polynomial of degree at most $k(n-1)-j_0$ in $N_l(P+\Delta P)$. Therefore $\|\Delta P\|_s \geqslant \eta_{k(n-1)-j_0}^{(s)}$. But this inequality cannot be strict as $\|\Delta P\|_s = \delta_s^{(n-1)}(P)$.

The proof of $\delta_s^{(n-1)}(P) = \gamma_{k(n-1)-i_0}^{(s)}$ follows from similar arguments due to part (b) which assures the existence of a matrix polynomial $\Delta P(\lambda) = \sum_{i=0}^{k} \lambda^i \Delta A_i$ such that $(P+\Delta P)(\lambda)$ has a left minimal index i_0 and $\|\Delta P\|_s = \delta_s^{(n-1)}(P)$. This completes the proof of part (c).

From parts (b) and (c) it is clear that any singular matrix polynomial $(P+\Delta P)(\lambda) = \sum_{i=0}^{k} \lambda^i (A_i + \Delta A_i)$ such that $\|\Delta P\|_s = \delta_s^{(n-1)}(P)$, has vector polynomials of degree at most $k(n-1)-j_0$ and $k(n-1)-i_0$ in $N_l(P+\Delta P)$ and $N_r(P+\Delta P)$ respectively. This proves part (d).

To prove part (e), observe that if anyone or both j_0 and i_0 is either 0 or $k(n-1)$, then the statement holds trivially due to part (a) and the fact that in such a case all the numbers in at least one of the γ-sequence and η-sequence are equal. Therefore without loss of generality it may be assumed that $1 \leqslant j_0, i_0 \leqslant k(n-1)-1$. Suppose $1 \leqslant j_0 \leqslant q$. As the γ-sequence is non increasing and $\delta_s^{(n-1)}(P) = \gamma_{j_0}^{(s)}$, it follows that $\delta_s^{(n-1)}(P) = \gamma_q^{(s)}$. If $j_0 > q$, then $k(n-1)-j_0 \leqslant q$ and by part (c), $\delta_s^{(n-1)}(P) = \eta_{k(n-1)-j_0}^{(s)} \geqslant \eta_q^{(s)}$. As this inequality cannot be strict, therefore in this case $\delta_s^{(n-1)}(P) = \eta_q^{(s)}$. If $1 \leqslant i_0, j_0 \leqslant q$, then as γ-sequence and η-sequence are non increasing, evidently $\delta_s^{(n-1)}(P) = \gamma_q^{(s)} = \eta_q^{(s)}$. Conversely, suppose $\delta_s^{(n-1)}(P) = \gamma_q^{(s)} = \eta_q^{(s)}$, and if possible suppose that at least one among i_0 and j_0 is greater than q. As $\delta_s^{(n-1)}(P) = \min\left\{\gamma_q^{(s)}, \eta_q^{(s)}\right\}$, only one of them, say i_0, is greater than q. Then clearly $\eta_q^{(s)} > \eta_{i_0}^{(s)} = \delta_s^{(n-1)}(P)$, which contradicts the assumption that $\delta_s^{(n-1)}(P) = \eta_q^{(s)}$. Similarly, if $j_0 > q$, then it is easy to see that $\gamma_q^{(s)} > \gamma_{j_0}^{(s)} = \delta_s^{(n-1)}(P)$, leading to the same contradiction. This completes the proves part (e). $\square$

Remark 5.4.2. *Observe that Theorem 5.4.1 holds for any choice of norm on the matrix polynomials. In [12] it was established that* $\delta_F^{(n-1)}(P) = \gamma_{k(n-1)}^{(F)} = \min_{0\leqslant j\leqslant k(n-1)} \gamma_j^{(F)}$ *when* $P(\lambda)$ *is a matrix pencil. Part (e) of Theorem 5.4.1 is a modified form of this result that also holds for the matrix polynomials. It may be utilized to formulate the computation of the distance to a nearest singular matrix polynomial as an optimization.*

The next result formulates the computation of $\delta_s^{(n-1)}(P)$ for $s = 2, F$ in terms of a variable projection least squares problem. For $s = F$, the formulation is based on part (e) of Theorem 5.4.1. Since a suitable variable projection least squares formulation for computing the η-sequence is not available for $s = 2$, the formulation for this case is based on Corollary 5.2.1.

Theorem 5.4.3. *Let* $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ *be an* $n \times n$ *matrix polynomial of degree* k*. For* $q \in \left\{k(n-1), \left\lfloor \frac{k(n-1)}{2} \right\rfloor\right\}$*, let* $\mathcal{S}_q$ *be the collection of all non zero vectors in* $\mathbb{C}^{(q+1)n}$*. Suppose that the vectors* $x \in \mathcal{S}_q$ *are partitioned as* $x = \begin{bmatrix} x_0^T & \cdots & x_q^T \end{bmatrix}^T \in \mathbb{C}^{(q+1)n}$ *where* $x_i \in \mathbb{C}^n, i = 0, \ldots, q$*. Let*

$$X_q = \begin{bmatrix} x_0 & x_1 & \cdots & x_k & \cdots & \cdots & x_q & & & \\ & x_0 & x_1 & \cdots & x_k & \cdots & \cdots & x_q & & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & & x_0 & x_1 & \cdots & x_k & \cdots & \cdots & x_q \end{bmatrix}.$$

Then for $q = \left\lfloor \frac{k(n-1)}{2} \right\rfloor$*,*

$$\delta_F^{(n-1)}(P) = \min\left\{ \inf_{x\in\mathcal{S}_q} \left\| \begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix} X_q X_q^\dagger \right\|_F, \inf_{x\in\mathcal{S}_q} \left\| \begin{bmatrix} A_0^T & \cdots & A_k^T \end{bmatrix} X_q X_q^\dagger \right\|_F \right\} \tag{5.4.3}$$

and for $q = k(n-1)$*,*

$$\delta_2^{(n-1)}(P) = \inf_{x\in\mathcal{S}_q} \left\| \begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix} X_q X_q^\dagger \right\|_2. \tag{5.4.4}$$

Proof. By Theorem 5.4.1(e), $\delta_F^{(n-1)}(P) = \min\left\{\gamma_q^{(F)}, \eta_q^{(F)}\right\}$, where $q = \left\lfloor \frac{k(n-1)}{2} \right\rfloor$, $\gamma_q^{(F)}$ and $\eta_q^{(F)}$ being as in (5.4.1) and (5.4.2) respectively. Let $\Delta P(\lambda) = \sum_{i=0}^{k} \lambda^i \Delta A_i$ such that $C_q(P + \Delta P)$ is rank deficient. Then $C_q(P + \Delta P)x = 0$ for some $x \in \mathcal{S}_q$. By arguing as in the proof of Theorem 5.3.1, this is equivalent to

$$\begin{bmatrix} \Delta A_0 & \cdots & \Delta A_k \end{bmatrix} X_q = -\begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix} X_q.$$

Therefore, by Theorem 1.5.1, a choice of $\begin{bmatrix} \Delta A_0 & \cdots & \Delta A_k \end{bmatrix}$ satisfying the above equation that is also minimal with respect to the Frobenius norm is given by

$$\begin{bmatrix} \Delta A_0 & \cdots & \Delta A_k \end{bmatrix} = -\begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix} X_q X_q^{\dagger}.$$

This implies that

$$\gamma_q^{(F)} = \inf_{x \in \mathcal{S}_q} \left\| \begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix} X_q X_q^{\dagger} \right\|_F.$$

Similarly, it follows that

$$\eta_q^{(F)} = \inf_{x \in \mathcal{S}_q} \left\| \begin{bmatrix} A_0^T & \cdots & A_k^T \end{bmatrix} X_q X_q^{\dagger} \right\|_F.$$

This establishes (5.4.3).

By Corollary 5.2.1, $\delta_2^{(n-1)}(P) = \inf\{\|\!|\Delta P\|\!|_2 : C_{k(n-1)}(P + \Delta P) \text{ is rank deficient}\}$. Therefore, (5.4.4) may be established via identical arguments. $\square$

The details concerning the strategy for computing the distance to singularity based on the above theorem are discussed in Section 5.7. For the Frobenius norm it is shown to be numerical advantageous over the computation based on the result in [12].

Alternatively by using Theorem 5.2.4, the distance to singularity may be formulated in terms of another optimization also involving a variable projection.

Theorem 5.4.4. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $n \times n$ matrix polynomial of degree k. Then for*

$$X = \begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_k & \cdots & x_{kn} \\ & x_0 & x_1 & \cdots & x_{k-1} & \cdots & x_{kn-1} \\ & & x_0 & \cdots & x_{k-2} & \cdots & x_{kn-2} \\ & & & \ddots & \vdots & \vdots & \vdots \\ & & & & x_0 & \cdots & x_{kn-k} \end{bmatrix}$$

where $x_i \in \mathbb{C}^n, i = 0, \ldots, kn,$ and $x_0 \neq 0$,

$$\delta_s^{(n-1)}(P) = \inf_{\substack{x_i \in \mathbb{C}^n \\ x_0 \neq 0}} \left\| \begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix} X X^{\dagger} \right\|_s, s = 2, F. \tag{5.4.5}$$

Proof. Let $(P+\Delta P)(\lambda)$ be singular where $\Delta P(\lambda)=\sum_{i=0}^{k}\lambda^i\Delta A_i$. Then by Theorem 5.2.4,

$$\bar{T}_{kn+1}(P+\Delta P)\begin{bmatrix} x_0\\ x_1\\ \vdots\\ x_{kn}\end{bmatrix}=\begin{bmatrix} A_0+\Delta A_0 & & & & \\ \vdots & \ddots & & & \\ A_k+\Delta A_k & & \ddots & & \\ & \ddots & & \ddots & \\ & & A_k+\Delta A_k & \cdots & A_0+\Delta A_0\end{bmatrix}\begin{bmatrix} x_0\\ x_1\\ \vdots\\ x_{kn}\end{bmatrix}=0$$

where $x_i\in\mathbb{C}^n, 0\leqslant i\leqslant kn, x_0\neq 0$. This can be written as

$$\begin{bmatrix} A_0+\Delta A_0 & \cdots & A_k+\Delta A_k\end{bmatrix}X=0,$$

$$\text{where } X=\begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_k & \cdots & x_{kn}\\ & x_0 & x_1 & \cdots & x_{k-1} & \cdots & x_{kn-1}\\ & & x_0 & \cdots & x_{k-2} & \cdots & x_{kn-2}\\ & & & \ddots & \vdots & \vdots & \vdots\\ & & & & x_0 & \cdots & x_{kn-k}\end{bmatrix} \text{ with } x_0\neq 0.$$

This implies $\begin{bmatrix}\Delta A_0 & \cdots & \Delta A_k\end{bmatrix}X=-\begin{bmatrix}A_0 & \cdots & A_k\end{bmatrix}X$ and by Theorem 1.5.1, a minimum 2 or Frobenius norm solution of this equation is given by

$$\begin{bmatrix}\Delta A_0 & \cdots & \Delta A_k\end{bmatrix}=-\begin{bmatrix}A_0 & \cdots & A_k\end{bmatrix}XX^{\dagger}.$$

Therefore

$$\delta_s^{(n-1)}(P)=\inf_{\substack{x_i\in\mathbb{C}^n\\ x_0\neq 0}}\left\|\begin{bmatrix}A_0 & \cdots & A_k\end{bmatrix}XX^{\dagger}\right\|_s, s=2 \text{ or } F.$$

This completes the proof. □

Remark 5.4.5. *In Theorem 4.4.1 it was shown that the distance with respect to the norms $\|\cdot\|_s, s=2,F$, from an $n\times n$ matrix polynomial $P(\lambda)=\sum_{i=0}^{k}\lambda^iA_i$ of degree k to a nearest matrix polynomial with a Jordan chain of length kn corresponding to zero is given by*

$$\inf_{\substack{x_i\in\mathbb{C}^n\\ x_0\neq 0}}\left\|\begin{bmatrix}A_0 & \cdots & A_k\end{bmatrix}\hat{X}\hat{X}^{\dagger}\right\|_s$$

where

$$\hat{X} = \begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_k & \cdots & x_{kn-1} \\ & x_0 & x_1 & \cdots & x_{k-1} & \cdots & x_{kn-2} \\ & & x_0 & \cdots & x_{k-2} & \cdots & x_{kn-3} \\ & & & \ddots & \vdots & \vdots & \vdots \\ & & & & x_0 & \ldots & x_{kn-k-1} \end{bmatrix}.$$

Observe that the matrix $\hat{X}$ is obtained by removing the last column of the matrix X involved in the projection $XX^\dagger$ in (5.4.5). This suggests an interesting relationship between the two distances.

5.5 The distances via structured singular values

In this section the results of Sections 5.1 and 5.2 are used to formulate the distance problems in terms of structured singular values or μ-values and their generalized versions. The following elementary result will be important for characterizing the distances in terms of μ-values.

Proposition 5.5.1. *Let M be an $m_1 \times n_1$ matrix such that $m_1 \geqslant n_1$. Consider $K = \begin{bmatrix} I_{m_1} & M \\ M^* & 0 \end{bmatrix}$. Then $\text{nullity}(M) \geqslant l$ if and only if $\text{nullity}(K) \geqslant l$.*

Also for the purpose of the characterization, with an $n \times n$ matrix polynomial $Q(\lambda) = \sum_{i=0}^{k} \lambda^i B_i$ of grade k we will frequently associate the matrix

$$\widehat{M}_{kr}(Q) = I_{kr+k+1} \otimes \begin{bmatrix} B_0 & \cdots & B_k \end{bmatrix}. \tag{5.5.1}$$

The following lemma that factorizes the convolution matrix $C_{kr}(Q)$ will be key to the characterization. The proof is skipped as it easily follows via direct multiplication of the stated factors.

Lemma 5.5.2. *For a given positive integer r and an $n \times n$ matrix polynomial $Q(\lambda) = \sum_{i=0}^{k} \lambda^i B_i$, the convolution matrix $C_{kr}(Q)$ given by (1.3.1) can be factored as*

$$C_{kr}(Q) = \widehat{M}_{kr}(Q)\widehat{Y},$$

where $\widehat{Y} = \begin{bmatrix} \hat{y}_1 & \hat{y}_2 & \cdots & \hat{y}_{(kr+1)} \end{bmatrix} \otimes I_n$ such that the entries in the $(i-1)k+i+j(k+2)$ position of $\hat{y}_i \in \mathbb{C}^{(kr+k+1)(k+1)}$ are one for each $j = 0, \ldots, k$ and all other entries are zero.

The next theorem shows that the distance $\delta_s^r(P)$ for $s = 2$ and $s = F$, is essentially the reciprocal of a generalized structured singular value.

Theorem 5.5.3. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $n \times n$ regular matrix polynomial of degree k. For any $n \times n$ matrix polynomial $\Delta P(\lambda) = \sum_{i=0}^{k} \lambda_i \Delta A_i$, let S be the perturbation class of all perturbations of type $\begin{bmatrix} & \widehat{M}_{kr}(\Delta P) \\ \widehat{M}_{kr}(\Delta P)^* & \end{bmatrix}$ where $\widehat{M}_{kr}(\Delta P)$ is as given in (5.5.1). Let $C_{kr}(P)$ and $\widehat{Y}$ be as defined in (1.3.1) and Lemma 5.5.2 respectively. Then,*

$$\delta_2^{(r)}(P) = \left[\mu_{S,\|\cdot\|_2}^{(kr+1)(n-r)-kr} \left(\begin{bmatrix} I & \\ & \widehat{Y} \end{bmatrix} \begin{bmatrix} I & C_{kr}(P) \\ C_{kr}(P)^* & \end{bmatrix}^{-1} \begin{bmatrix} I & \\ & \widehat{Y}^* \end{bmatrix} \right) \right]^{-1},$$

and

$$\delta_F^{(r)}(P) = \frac{\left[\mu_{S,\|\cdot\|_F}^{(kr+1)(n-r)-kr} \left(\begin{bmatrix} I & \\ & \widehat{Y} \end{bmatrix} \begin{bmatrix} I & C_{kr}(P) \\ C_{kr}(P)^* & \end{bmatrix}^{-1} \begin{bmatrix} I & \\ & \widehat{Y}^* \end{bmatrix} \right) \right]^{-1}}{\sqrt{2(kr+k+1)}}.$$

Proof. From Theorem 5.1.4, normal rank of $(P - \Delta P)(\lambda)$ is at most r for an $n \times n$ matrix polynomial $\Delta P(\lambda) = \sum_{i=0}^{k} \lambda_i \Delta A_i$ if and only if the nullity of $C_{kr}(P - \Delta P)$ is at least $(kr+1)(n-r) - kr$. Due to Proposition 5.5.1 this can be written as

$$\begin{aligned} &\text{nullity} \left(\begin{bmatrix} I & C_{kr}(P-\Delta P) \\ C_{kr}(P-\Delta P)^* & \end{bmatrix} \right) \geqslant (kr+1)(n-r) - kr \\ \Leftrightarrow &\text{nullity} \left(\begin{bmatrix} I & C_{kr}(P) \\ C_{kr}(P)^* & \end{bmatrix} - \begin{bmatrix} & C_{kr}(\Delta P) \\ C_{kr}(\Delta P)^* & \end{bmatrix} \right) \geqslant (kr+1)(n-r) - kr. \end{aligned}$$

Using the factorization of $C_{kr}(\Delta P)$ from Lemma 5.5.2,

$$\begin{aligned} &\text{nullity} \left(I - \begin{bmatrix} & \widehat{M}_{kr}(\Delta P)\widehat{Y} \\ \widehat{Y}^*\widehat{M}_{kr}(\Delta P)^* & \end{bmatrix} \begin{bmatrix} I & C_{kr}(P) \\ C_{kr}(P)^* & \end{bmatrix}^{-1} \right) \geqslant (kr+1)(n-r) - kr \\ \Leftrightarrow &\text{nullity} \left(I - \begin{bmatrix} I & \\ & \widehat{Y}^* \end{bmatrix} \begin{bmatrix} & \widehat{M}_{kr}(\Delta P) \\ \widehat{M}_{kr}(\Delta P)^* & \end{bmatrix} \begin{bmatrix} I & \\ & \widehat{Y} \end{bmatrix} \begin{bmatrix} I & C_{kr}(P) \\ C_{kr}(P)^* & \end{bmatrix}^{-1} \right) \geqslant (kr+1)(n-r) - kr \\ \Leftrightarrow &\text{nullity} \left(I - \begin{bmatrix} & \widehat{M}_{kr}(\Delta P) \\ \widehat{M}_{kr}(\Delta P)^* & \end{bmatrix} \begin{bmatrix} I & \\ & \widehat{Y} \end{bmatrix} \begin{bmatrix} I & C_{kr}(P) \\ C_{kr}(P)^* & \end{bmatrix}^{-1} \begin{bmatrix} I & \\ & \widehat{Y}^* \end{bmatrix} \right) \geqslant (kr+1)(n-r) - kr. \end{aligned} \tag{5.5.2}$$

Therefore,

$$\delta_2^{(r)}(P) = \left[\mu_{S,\|\cdot\|_2}^{(kr+1)(n-r)-kr}\left(\begin{bmatrix} I & \\ & \widehat{Y}\end{bmatrix}\begin{bmatrix} I & C_{kr}(P) \\ C_{kr}(P)^* & \end{bmatrix}^{-1}\begin{bmatrix} I & \\ & \widehat{Y}^*\end{bmatrix}\right)\right]^{-1}$$

and

$$\delta_F^{(r)}(P) = \frac{\left[\mu_{S,\|\cdot\|_F}^{(kr+1)(n-r)-kr}\left(\begin{bmatrix} I & \\ & \widehat{Y}\end{bmatrix}\begin{bmatrix} I & C_{kr}(P) \\ C_{kr}(P)^* & \end{bmatrix}^{-1}\begin{bmatrix} I & \\ & \widehat{Y}^*\end{bmatrix}\right)\right]^{-1}}{\sqrt{2(kr+k+1)}}$$

which completes the proof. □

By putting $r = n-1$ as an immediate corollary it follows that computing the distance to singularity $\delta_s^{(n-1)}(P)$, $s = 2, F$, is equivalent to computing the reciprocal of a structured singular value or μ-value.

Corollary 5.5.4. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $n \times n$ regular matrix polynomial of degree k. For any $n \times n$ matrix polynomial $\Delta P(\lambda) = \sum_{i=0}^{k} \lambda_i \Delta A_i$, let S be the perturbation class of all perturbations of type $\begin{bmatrix} & \widehat{M}_{kn-k}(\Delta P) \\ \widehat{M}_{kn-k}(\Delta P)^* & \end{bmatrix}$ where $\widehat{M}_{kn-k}(\Delta P)$ is as given in (5.5.1). Let $C_{k(n-1)}(P)$ and $\widehat{Y}$ be as defined in (1.3.1) and Lemma 5.5.2 respectively. Then,*

$$\delta_2^{(n-1)}(P) = \left[\mu_{S,\|\cdot\|_2}\left(\begin{bmatrix} I & \\ & \widehat{Y}\end{bmatrix}\begin{bmatrix} I & C_{kn-k}(P) \\ C_{kn-k}(P)^* & \end{bmatrix}^{-1}\begin{bmatrix} I & \\ & \widehat{Y}^*\end{bmatrix}\right)\right]^{-1},$$

and

$$\delta_F^{(n-1)}(P) = \frac{\left[\mu_{S,\|\cdot\|_F}\left(\begin{bmatrix} I & \\ & \widehat{Y}\end{bmatrix}\begin{bmatrix} I & C_{kn-k}(P) \\ C_{kn-k}(P)^* & \end{bmatrix}^{-1}\begin{bmatrix} I & \\ & \widehat{Y}^*\end{bmatrix}\right)\right]^{-1}}{\sqrt{2(kn+1)}}.$$

The next theorem provides an alternative characterization of the distance to singularity in terms of a generalized μ-value. It is based on Theorem 5.2.4 and leads to lower bounds of the distance to singularity in the next section. The following lemma which is a corollary of Lemma 4.3.1 will be used for the formulation.

Lemma 5.5.5. *For a given $n \times n$ matrix polynomial $Q(\lambda) = \sum_{i=0}^{k} \lambda^i B_i$ of grade k,*

$$\bar{T}_{kn+1}(Q) = \left(I_{kn+1} \otimes \begin{bmatrix} B_0 & \cdots & B_k \end{bmatrix}\right) E$$

where

$$E = \begin{bmatrix} \tilde{E}_1^T & \tilde{E}_2^T & \cdots & \tilde{E}_{kn+1}^T \end{bmatrix}^T \otimes I_n$$

$\tilde{E}_i \in \mathbb{C}^{(k+1)\times(kn+1)}$ being the first $k+1$ rows of the $kn+1 \times kn+1$ matrices E_i, $i = 1, \ldots, kn+1$ given by,

$$E_1 = \begin{bmatrix} 1 & & & \\ & & & \\ & & & \\ & & & \end{bmatrix}, E_2 = \begin{bmatrix} & 1 & & \\ 1 & & & \\ & & & \\ & & & \end{bmatrix}, \ldots, E_{kn} = \begin{bmatrix} & & & 1 & \\ & & \iddots & & \\ 1 & & & & \\ & & & & \end{bmatrix}, E_{kn+1} = \begin{bmatrix} & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & \end{bmatrix}.$$

Theorem 5.5.6. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $n \times n$ matrix polynomial of degree k. For $\Delta A_i \in \mathbb{C}^{n\times n}, i = 0, \ldots, k$, let $\hat{S}$ be the perturbation class of all perturbations of the type $\left(I_{kn+1} \otimes \begin{bmatrix} \Delta A_0 & \cdots & \Delta A_k \end{bmatrix}\right)$. Let A_0 be a non-singular matrix, $\bar{T}_{kn+1}(P)$ be as (1.3.2) and E be as given in Lemma 5.5.5. Then,*

$$\delta_2^{(n-1)}(P) = \left[\mu^{kn+1}_{\hat{S},\|\cdot\|_2}\left(E\, \bar{T}_{kn+1}(P)^{-1}\right)\right]^{-1},$$

and

$$\delta_F^{(n-1)}(P) = \frac{\left[\mu^{kn+1}_{\hat{S},\|\cdot\|_F}\left(E\, \bar{T}_{kn+1}(P)^{-1}\right)\right]^{-1}}{\sqrt{kn+1}}.$$

Proof. From Theorem 5.2.4, $(P - \Delta P)(\lambda)$ is singular for $\Delta P(\lambda) = \sum_{i=0}^{k} \lambda^i \Delta A_i$ if and only if $\text{nullity}(\bar{T}_{kn+1}(P - \Delta P)) \geqslant kn + 1$. As A_0 is invertible, $(P - \Delta P)(\lambda)$ is singular if and only if

$$\text{nullity}(I - \bar{T}_{kn+1}(\Delta P)\bar{T}_{kn+1}(P)^{-1}) \geqslant kn + 1.$$

Due to Lemma 5.5.5 the above relation may be written as,

$$\text{nullity}\left(I - \left(I_{kn+1} \otimes \begin{bmatrix} \Delta A_0 & \cdots & \Delta A_k \end{bmatrix}\right) E\, \bar{T}_{kn+1}(P)^{-1}\right) \geqslant kn + 1. \tag{5.5.3}$$

Therefore,

$$\delta_2^{(n-1)}(P) = \left[\mu^{kn+1}_{\hat{S},\|\cdot\|_2}\left(E\, \bar{T}_{kn+1}(P)^{-1}\right)\right]^{-1},$$

and

$$\delta_F^{(n-1)}(P) = \frac{\left[\mu^{kn+1}_{\hat{S},\|\cdot\|_F}\left(E\, \bar{T}_{kn+1}(P)^{-1}\right)\right]^{-1}}{\sqrt{kn+1}}$$

which completes the proof. □

5.6 Bounds on the distances

In this section we present several lower and upper bounds on the distances under consideration. Such bounds exist in the literature for the distance to singularity for matrix pencils. Some of the bounds in this section are generalizations of those for matrix polynomials. However the main results are the lower bounds arising from the formulation of the distances as structured singular values.

5.6.1 Lower bounds

The following lower bound is obtained from Theorem 5.5.3.

Theorem 5.6.1. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $n \times n$ regular matrix polynomial of degree k, $C_{kr}(P)$ and $C_{kr}(P^T)$ be the level kr convolution matrices of $P(\lambda)$ and $P(\lambda)^T$ respectively, and $\widehat{Y}$ be as defined in Lemma 5.5.2. Then*

$$\delta_F^{(r)}(P) \geqslant \max\{\beta_1, \beta_2\} \text{ and } \delta_2^{(r)}(P) \geqslant \beta_1,$$

where

$$\beta_1 := \sup_{a,b \in \mathbb{S}^{kr+k+1}} \left[\sigma_{(kr+1)(n-r)-kr} \left(\widetilde{W} \begin{bmatrix} & \widehat{Y} \\ I & \end{bmatrix} \begin{bmatrix} I & C_{kr}(P) \\ C_{kr}(P)^* & \end{bmatrix}^{-1} \begin{bmatrix} I & \\ & \widehat{Y}^* \end{bmatrix} \widehat{W} \right) \right]^{-1},$$

$$\beta_2 := \sup_{a,b \in \mathbb{S}^{kr+k+1}} \left[\sigma_{(kr+1)(n-r)-kr} \left(\widetilde{W} \begin{bmatrix} & \widehat{Y} \\ I & \end{bmatrix} \begin{bmatrix} I & C_{kr}(P^T) \\ C_{kr}(P^T)^* & \end{bmatrix}^{-1} \begin{bmatrix} I & \\ & \widehat{Y}^* \end{bmatrix} \widehat{W} \right) \right]^{-1},$$

$$\widetilde{W} := \begin{bmatrix} W_{kr+k+1,(k+1)n}(a) & \\ & W_{kr+k+1,n}(b) \end{bmatrix},$$

$$\widehat{W} := \begin{bmatrix} W^{-1}_{kr+k+1,n}(a) & \\ & W^{-1}_{kr+k+1,(k+1)n}(b) \end{bmatrix},$$

$W_{t,n}(a)$ and $\mathbb{S}^t$ being as in Theorem 4.5.2.

Proof. Initially we show that β_1 is a lower bound of $\delta_s^{(r)}(P)$ for $s = 2, F$. If the normal rank of $(P - \Delta P)(\lambda)$ is at most r for any matrix polynomial $\Delta P(\lambda)$ of grade k, then by (5.5.2)

$$\text{nullity}\left(I - \begin{bmatrix} \widehat{M}_{kr}(\Delta P) & \\ & \widehat{M}_{kr}(\Delta P)^* \end{bmatrix} \begin{bmatrix} & I \\ I & \end{bmatrix} \begin{bmatrix} I & \\ & \widehat{Y} \end{bmatrix} \begin{bmatrix} I & C_{kr}(P) \\ C_{kr}(P)^* & \end{bmatrix}^{-1} \begin{bmatrix} I & \\ & \widehat{Y}^* \end{bmatrix} \right) \geqslant (kr+1)(n-r) - kr$$

$$\Leftrightarrow \text{nullity}\left(I - \widehat{W}\begin{bmatrix}\widehat{M}_{kr}(\Delta P) & \\ & \widehat{M}_{kr}(\Delta P)^*\end{bmatrix}\widetilde{W}\begin{bmatrix} & \widehat{Y}\\ I & \end{bmatrix}\begin{bmatrix} I & C_{kr}(P)\\ C_{kr}(P)^* & \end{bmatrix}^{-1}\begin{bmatrix} I & \\ & \widehat{Y}^*\end{bmatrix}\right) \geqslant (kr+1)(n-r)-kr$$

where $\widehat{M}_{kr}(\Delta P)$ is as defined in (5.5.1), $\widetilde{W} = \begin{bmatrix} W_{kr+k+1,(k+1)n}(a) & \\ & W_{kr+k+1,n}(b)\end{bmatrix}$ and

$$\widehat{W} = \begin{bmatrix} W^{-1}_{kr+k+1,n}(a) & \\ & W^{-1}_{kr+k+1,(k+1)n}(b)\end{bmatrix}.$$

This implies

$$\text{nullity}\left(I - \begin{bmatrix}\widehat{M}_{kr}(\Delta P) & \\ & \widehat{M}_{kr}(\Delta P)^*\end{bmatrix}\widetilde{W}\begin{bmatrix} & \widehat{Y}\\ I & \end{bmatrix}\begin{bmatrix} I & C_{kr}(P)\\ C_{kr}(P)^* & \end{bmatrix}^{-1}\begin{bmatrix} I & \\ & \widehat{Y}^*\end{bmatrix}\widehat{W}\right) \geqslant (kr+1)(n-r)-kr.$$

Therefore by *Lemma* 1.6.3

$$\left\|\begin{bmatrix}\widehat{M}_{kr}(\Delta P) & \\ & \widehat{M}_{kr}(\Delta P)^*\end{bmatrix}\right\|_2 \geqslant \left[\sigma_{(kr+1)(n-r)-kr}\left(\widetilde{W}\begin{bmatrix} & \widehat{Y}\\ I & \end{bmatrix}\begin{bmatrix} I & C_{kr}(P)\\ C_{kr}(P)^* & \end{bmatrix}^{-1}\begin{bmatrix} I & \\ & \widehat{Y}^*\end{bmatrix}\widehat{W}\right)\right]^{-1},$$

and hence $\delta_s^{(r)}(P) \geqslant \beta_1$ for $s = 2$ or F. By replacing $P(\lambda)$ by its transpose in the above arguments and using the fact that $\delta_F^{(r)}(P^T) = \delta_F^{(r)}(P)$, it follows that $\delta_F^{(r)}(P) \geqslant \beta_2$. Hence the proof. $\square$

A lower bound of the distance to singularity $\delta_s^{(n-1)}(P)$ follows from the above result. An alternative lower bound may also be derived from Theorem 5.5.6.

Theorem 5.6.2. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $n \times n$ matrix polynomial of degree k. Let A_0 be non-singular, $\bar{T}_{kn+1}(P)$ be as in (1.3.2) and E be as defined in Lemma 5.5.5. Then*

$$\delta_F^{(n-1)}(P) \geqslant \max\{\alpha_1, \alpha_2\} \text{ and } \delta_2^{(n-1)}(P) \geqslant \alpha_1,$$

where

$$\alpha_1 := \sup_{a\in\mathbb{S}^{kn+1}} \left[\sigma_{kn+1}(W_{kn+1,(k+1)n}(a)E\,\bar{T}_{kn+1}(P)^{-1}W^{-1}_{kn+1,n}(a))\right]^{-1},$$

$$\alpha_2 := \sup_{a\in\mathbb{S}^{kn+1}} \left[\sigma_{kn+1}(W_{kn+1,(k+1)n}(a)E\,\bar{T}_{kn+1}(P^T)^{-1}W^{-1}_{kn+1,n}(a))\right]^{-1},$$

and $W_{t,n}(a)$ and $\mathbb{S}^t$ are as defined in Theorem 4.5.2.

Proof. From (5.5.3), $(P-\Delta P)(\lambda)$ is singular if and only if the nullity of

$$I - W_{kn+1,n}^{-1}(a)\left(I_{kn+1}\otimes[\Delta A_0\cdots\Delta A_k]\right)W_{kn+1,(k+1)n}(a)E\,\bar{T}_{kn+1}(P)^{-1}$$

is at least $kn+1$. Therefore

$$\text{nullity}\left(I-(I_{kn+1}\otimes[\Delta A_0\cdots\Delta A_k])\,W_{kn+1,(k+1)n}(a)E\,\bar{T}_{kn+1}(P)^{-1}W_{kn+1,n}^{-1}(a)\right)\geqslant kn+1.$$

By Lemma 1.6.3,

$$\left\|\begin{bmatrix}\Delta A_0 & \cdots & \Delta A_k\end{bmatrix}\right\|_2 \geqslant \left[\sigma_{kn+1}(W_{kn+1,(k+1)n}(a)E\,\bar{T}_{kn+1}(P)^{-1}W_{kn+1,n}^{-1}(a))\right]^{-1}.$$

which implies, that $\delta_s^{(n-1)}(P)\geqslant\alpha_1$, for $s=2$ or F. Now the proof follows from the fact that $\delta_F^{(n-1)}(P^T)\geqslant\alpha_2$ and $\delta_F^{(n-1)}(P^T)=\delta_F^{(n-1)}(P)$. □

If the polynomial $(P+\Delta P)(\lambda)$ is singular, then $C_j(P+\Delta P)$ is rank deficient for some $j_0, 0\leqslant j_0\leqslant k(n-1)$ and

$$\sigma_{\min}(C_{j_0}(P))\leqslant\sqrt{j_0+1}\|\Delta P\|_s\Rightarrow\frac{\sigma_{\min}(C_{j_0}(P))}{\sqrt{j_0+1}}\leqslant\delta_2^{(n-1)}(P).$$

This gives the lower bound

$$\frac{\sigma_{\min}(C_{k(n-1)}(P))}{\sqrt{k(n-1)+1}}\leqslant\delta_2^{(n-1)}(P). \tag{5.6.1}$$

As $\delta_F^{(n-1)}(P)=\delta_F^{(n-1)}(P^T)$,

$$\frac{\sigma_{\min}(C_{k(n-1)}(P^T))}{\sqrt{k(n-1)+1}}\leqslant\delta_F^{(n-1)}(P). \tag{5.6.2}$$

Setting the left hand sides of (5.6.1) and (5.6.2) to be l_1 and l_2 respectively, it follows that

$$\max\{l_1,l_2\}\leqslant\delta_F^{(n-1)}(P). \tag{5.6.3}$$

The next lower bound is a generalization of the lower bounds in [12] for matrix pencils to matrix polynomials.

If $(P+\Delta P)(\lambda)$ is a singular matrix polynomial, then for every complex number λ_0, rank $(P+\Delta P)(\lambda_0)<n$. Every number in the extended complex plane $\mathbb{C}\cup\{\infty\}$ may be identified with ordered pairs of the form (α,β) where $\alpha\in\mathbb{R},\beta\in\mathbb{C}$ such that $\alpha^2+|\beta|^2=1$ and the polynomial $P(\lambda)$ may be written in homogeneous form as $P(\alpha,\beta)=\sum_{i=0}^k\beta^{k-i}\alpha^iA_i$. It is well known [1] that for each pair (α,β),

$$\min\{\|\Delta P\|_s:\text{rank}(P(\alpha,\beta)+\Delta P(\alpha,\beta))<n\}=\frac{\sigma_{\min}\left(\sum_{i=0}^k\beta^{k-i}\alpha^iA_i\right)}{\sqrt{\sum_{i=0}^k|\beta|^{2(k-i)}\alpha^{2i}}},\ s=2,F.$$

Therefore, $\frac{\sigma_{\min}\left(\sum_{i=0}^{k}\beta^{k-i}\alpha^i A_i\right)}{\sqrt{\sum_{i=0}^{k}|\beta|^{2(k-i)}\alpha^{2i}}} \leqslant \|\Delta P\|_s$, $s = 2, F$, which gives

$$\max_{|\beta|^2+\alpha^2=1} \frac{\sigma_{\min}\left(\sum_{i=0}^{k}\beta^{k-i}\alpha^i A_i\right)}{\sqrt{\sum_{i=0}^{k}|\beta|^{2(k-i)}\alpha^{2i}}} \leqslant \delta_s^{(n-1)}(P) \text{ for } s = 2 \text{ or } F. \tag{5.6.4}$$

A final lower bound follows from Theorem 5.2.4. Recall that $(P + \Delta P)(\lambda)$ is singular if and only if $\operatorname{rank}(\bar{T}_{kn+1}(P + \Delta P)) \leqslant (kn + 1)(n - 1)$. This is equivalent to $\sigma_{(kn+1)n-kn}(T_\gamma(P + \Delta P, 0, kn)) = 0$ where $T_\gamma(P + \Delta P, 0, kn)$ is as in (4.2.2).

Therefore $\sigma_{(kn+1)n-kn}(T_\gamma(P, 0, kn)) \leqslant \|T_\gamma(\Delta P, 0, kn)\|_2$, which implies

$$\frac{\sigma_{(kn+1)n-kn}(T_\gamma(P, 0, kn))}{F(\gamma)} \leqslant \|\Delta P\|_2 \tag{5.6.5}$$

where

$$F(\gamma) = \sqrt{\max\left\{\max_{1\leqslant p\leqslant k}\left(1 + \sum_{t=1}^{p}\prod_{i=t}^{p}\gamma^2_{(kn+1-i)}\right), \max_{k+1\leqslant p\leqslant kn}\left(1 + \sum_{t=p-k+1}^{p}\prod_{i=t}^{p}\gamma^2_{(kn+1-i)}\right)\right\}}.$$

Let Γ be as in (4.2.1) with $r = kn$. Taking the supremum of the left side of (5.6.5) over all $\gamma \in \Gamma$

$$\sup_{\gamma\in\Gamma} \frac{\sigma_{(kn+1)n-kn}(T_\gamma(P, 0, kn))}{F(\gamma)} \leqslant \delta_s^{(n-1)}(P) \text{ for } s = 2 \text{ or } F. \tag{5.6.6}$$

5.6.2 Upper bounds

The upper bounds of $\delta_s^{(n-1)}(P)$ for $s = 2, F$ extend the ones in [12] for the matrix pencils to the matrix polynomials. A trivial upper bound follows from the fact that $(P + \Delta P)(\lambda)$ is a singular matrix polynomial if its coefficient matrices have a common nonzero vector in their right (or left) null spaces. Therefore

$$\delta_2^{(n-1)}(P) \leqslant \delta_F^{(n-1)}(P) \leqslant \min\left\{\sigma_{\min}\left(\begin{bmatrix} A_0^T & \cdots & A_n^T \end{bmatrix}^T\right), \sigma_{\min}\left(\begin{bmatrix} A_0 & \cdots & A_n \end{bmatrix}\right)\right\}. \tag{5.6.7}$$

Another upper bound of $\delta_F^{(n-1)}(P)$ may be obtained by constructing a perturbation $\Delta P(\lambda) = \sum_{i=0}^{k}\lambda^i \Delta A_i$ to $P(\lambda) = \sum_{i=0}^{k}\lambda^i A_i$ such that the convolution matrices $C_j(P + \Delta P)$ and $C_j((P + \Delta P)^T)$ are rank deficient for $0 \leqslant j \leqslant k(n - 1)$. By arguing as in the proof of Theorem 5.4.3, $C_j(P + \Delta P)$ is rank deficient if

$\begin{bmatrix} \Delta A_0 & \cdots & \Delta A_k \end{bmatrix} = -\begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix} X_j X_j^\dagger$, where

$$X_j = \begin{bmatrix} x_0 & x_1 & \cdots & x_k & \cdots & \cdots & x_j & & & \\ & x_0 & x_1 & \cdots & x_k & \cdots & \cdots & x_j & & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & & x_0 & x_1 & \cdots & x_k & \cdots & \cdots & x_j \end{bmatrix}$$

with $x_i \in \mathbb{C}^n$ for $i = 0, \ldots, j$. For any choice of X_j

$$\delta_F^{(n-1)}(P) \leqslant \left\| \begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix} X_j X_j^\dagger \right\|_F \leqslant \left\| \begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix} X_j \right\|_F \left\| X_j^\dagger \right\|_2.$$

Let $x_i = v_i, i = 0, \ldots, j$ where $\begin{bmatrix} v_0^T & \cdots & v_j^T \end{bmatrix}^T$ is a right singular vector of $C_j(P)$ corresponding to $\sigma_{\min}(C_j(P))$. Then $C_j(P) \begin{bmatrix} v_0 \\ \vdots \\ v_j \end{bmatrix} = \sigma_{\min}(C_j(P)) \begin{bmatrix} u_0 \\ \vdots \\ u_j \end{bmatrix}$, where $\begin{bmatrix} u_0 \\ \vdots \\ u_j \end{bmatrix}$ is the corresponding left singular vector with $u_i \in \mathbb{C}^n$, $i = 1, \ldots, j$, such that $\sum_{i=1}^{j} \|u_i\|_2^2 = 1$. For $V_j := X_j$,

$$\left\| \begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix} V_j \right\|_F = \sigma_{\min}(C_j(P)).$$

Hence

$$\delta_F^{(n-1)}(P) \leqslant \min_{0 \leqslant j \leqslant k(n-1)} \frac{\sigma_{\min}(C_j(P))}{\sigma_{\min}(V_j)}. \tag{5.6.8}$$

Similarly replacing $C_j(P)$ by $C_j(P^T)$ in the above arguments gives,

$$\delta_F^{(n-1)}(P) \leqslant \min_{0 \leqslant j \leqslant k(n-1)} \frac{\sigma_{\min}(C_j(P^T))}{\sigma_{\min}(\tilde{V}_j)}. \tag{5.6.9}$$

Here $\tilde{V}_j = X_j$ with $x_i = \tilde{v}_i, i = 0, \ldots, j$, where $\begin{bmatrix} \tilde{v}_0^T & \cdots & \tilde{v}_j^T \end{bmatrix}^T$ is a right singular vector of $C_j(P^T)$ corresponding to $\sigma_{\min}(C_j(P^T))$. Setting $u1$ and $u2$ to be the right hand sides of (5.6.8) and (5.6.9) respectively, we have,

$$\delta_2^{(n-1)}(P) \leqslant \delta_F^{(n-1)}(P) \leqslant \min\{u1, u2\}. \tag{5.6.10}$$

5.7 Numerical Experiments

In this section we provide several numerical examples which compare the computed values of the distances with their upper and lower bounds. The optimization in

Theorem 5.3.1 is used to compute the distance from an $n \times n$ matrix polynomial to a nearest matrix polynomial with normal rank at most r when $r < n-1$. To compute the distance to a nearest singular matrix polynomial, the optimizations in Theorem 5.4.3 is used. The computations are done by using the Broyden Fletcher Goldfarb Shanno (BFGS) algorithm and MATLAB's `globalsearch.m`. A strategy for computing $\delta_s^{(r)}(P)$ when $r < n-1$ is explained in the following example from [31].

Example 5.7.1.

$$P(\lambda) = A_0 + \lambda A_1 + \lambda^2 A_2 + \lambda^3 A_3$$

where

$$A_0 = \begin{bmatrix} 0.09108776 & -0.05442464 & 0.3645006 & 0.01821543 \\ -0.1456436 & 0.03647524 & -0.07277662 & 0.07305016 \\ 0.05478714 & -0.05444916 & 0.437322 & 0.05478385 \\ -0.1274211 & 0.09124859 & -0.6556615 & -0.0544685 \end{bmatrix},$$

$$A_1 = \begin{bmatrix} 0.09116729 & 0.0000179769 & 0.2550857 & 0.05475106 \\ 0.0001156514 & 0.00001659159 & 0.09108906 & -0.05447104 \\ 0.05470823 & 0.03662426 & 0.1276959 & 0.03650378 \\ 0.05472202 & -0.1091389 & 0.1458359 & -0.09090507 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} 0.01833149 & 0.0366177 & 0.01824331 & 0.03660918 \\ 0.01837542 & -0.05442525 & 0 & 0.01832234 \\ 0.01841784 & 0.00003900436 & 0 & 0.01836515 \\ 0.01840752 & 0.00001508311 & 0.01839699 & 0.0365917 \end{bmatrix},$$

$$\text{and } A_3 = \begin{bmatrix} 0 & 0.01837967 & 0 & 0 \\ 0 & 0.01843603 & 0 & 0 \\ 0 & 0.01829203 & 0 & 0 \\ 0 & 0.01842778 & 0 & 0 \end{bmatrix}.$$

The aim is to compute the distance $\delta_s^{(2)}(P)$ to a nearest matrix polynomial with normal rank at most 2, where $s = 2$ or F. By Theorem 5.1.4 observe that any perturbed polynomial $(P + \Delta P)(\lambda)$ has normal rank at most 2 if and only if the null space of the convolution matrix $C_6(P + \Delta P)$ has at least 8 linearly independent vectors, say $z_1 \ldots, z_8$, each of length 28. By Theorem 5.3.1 any $\Delta P(\lambda) = \sum_{i=0}^{3} \lambda^i \Delta A_i$ satisfying this which is minimal with respect to $|||\cdot|||_s, s = 2, F$, may be taken to be of the form $\begin{bmatrix}\Delta A_0 & \cdots & \Delta A_3\end{bmatrix} = -\begin{bmatrix}A_0 & \cdots & A_3\end{bmatrix}\begin{bmatrix} X(z_1)\cdots X(z_8) \end{bmatrix}\begin{bmatrix} X(z_1)\cdots X(z_8) \end{bmatrix}^{\dagger}$. The matrix $\begin{bmatrix}A_0 & A_1 & A_2 & A_3\end{bmatrix}$ is of size 4×16 with rank 4, and $\begin{bmatrix}X(z_1)\cdots X(z_8)\end{bmatrix}$ is a 16×80 matrix of likely rank 16 for random choices of $z_1,\ldots,z_8$. Therefore for such choices, the computed value of the distance

$$\delta_s^{(2)}(P) = \inf_{\{z_1,\ldots,z_8\}\in\mathcal{V}_8}\left\|\begin{bmatrix}A_0 & A_1 & A_2 & A_3\end{bmatrix}\begin{bmatrix} X(z_1)\cdots X(z_8) \end{bmatrix}\begin{bmatrix} X(z_1)\cdots X(z_8) \end{bmatrix}^{\dagger}\right\|_s$$

for $s=2$ or F is unlikely to be less than $\sigma_{\min}\left(\begin{bmatrix}A_0 & A_1 & A_2 & A_3\end{bmatrix}\right)$ which is an upper bound of the distance from $P(\lambda)$ to a nearest singular matrix polynomial. To be able to find a value of the distance less than this, it is clear that the initial $z_1,\ldots,z_8$ have to be so chosen that the rank of $\begin{bmatrix}X(z_1)\cdots X(z_8)\end{bmatrix}$ is at most 12 and the computations performed in such a way that this rank restriction is fulfilled throughout the numerical process. We execute this by initially taking some of the vectors z_i with sufficiently many zero entries and the other vectors as their shifted versions (as defined in Definition 5.1.1) and then projecting the vectors obtained at the end of each iteration back to these forms. Following this strategy we obtain $\delta_F^{(2)}(P) = 7.7049\times 10^{-4}$ via BFGS and `globalsearch.m`, and $\delta_2^{(2)}(P) = 6.9394\times 10^{-4}$ via `globalsearch.m` for the choices

$$z_1 = \begin{bmatrix} z_{1,0} \\ z_{1,1} \\ z_{1,2} \\ 0_{16} \end{bmatrix} \text{ and } z_2 = \begin{bmatrix} z_{2,0} \\ z_{2,1} \\ z_{2,2} \\ 0_{16} \end{bmatrix}$$

where $z_{1,0}, z_{1,1}, z_{1,2}, z_{2,0}, z_{2,1}, z_{2,2} \in \mathbb{R}^4$ are random with the restriction that $\{z_{1,0}, z_{2,0}\}$ is linearly independent and the other 6 vectors $z_3,\ldots,z_8$ are the following shifted versions of z_1 and z_2:

$$z_1^i := \begin{bmatrix} 0_{4i} \\ z_{1,0} \\ z_{1,1} \\ z_{1,2} \\ 0_{(16-4i)} \end{bmatrix}, i = 1,2,3,4, \quad z_2^j := \begin{bmatrix} 0_{4j} \\ z_{2,0} \\ z_{2,1} \\ z_{2,2} \\ 0_{(16-4j)} \end{bmatrix}, j = 1,2.$$

Note that in [31], $\delta_F^{(2)}(P)$ was computed to be 7.844×10^{-4}. Evidently, this strategy may not yield the exact value of the distance but an upper bound. However in view of the choice of the $z_i, i = 1,\ldots,8$, if the list of right minimal indices of a nearest matrix polynomial of normal rank at most 2 contains $\{2, 2\}$ then it is likely to be a very good estimate of the distance.

The computation of the distance to a nearest singular matrix polynomial is based on Theorem 5.4.3. For computing $\delta_F^{(n-1)}(P)$ via (5.4.3), the choice of initial vectors $\left\{x_0, \ldots, x_{\lfloor \frac{k(n-1)}{2} \rfloor}\right\}$ is guided by Theorem 5.4.1 where it is established that there exists a nearest singular polynomial to $P(\lambda)$, say $(P + \Delta P)(\lambda)$, with a least right minimal index j_0 or a least left minimal index i_0, with properties as stated in the theorem. The aim of the computation is to find this singular matrix polynomial. Let $q = \left\lfloor \frac{k(n-1)}{2} \right\rfloor$. If $j_0 = 0$ or $i_0 = 0$ then by part(a) of Theorem 5.4.1, in the first case $\delta_F^{(n-1)}(P) = \sigma_{\min}\left(\begin{bmatrix} A_0^T & \cdots & A_k^T \end{bmatrix}^T\right)$ and in the second case, $\delta_F^{(n-1)}(P) = \sigma_{\min}\left(\begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix}\right)$. Therefore we focus on other cases where either $1 \leqslant j_0 \leqslant q$ or $1 \leqslant i_0 \leqslant q$. When $1 \leqslant j_0 \leqslant q$ ($1 \leqslant i_0 \leqslant q$) as j_0 (i_0) is the least right (left) minimal index of $(P + \Delta P)(\lambda)$, we have $C_q(P + \Delta P)x = 0$ ($C_q(P^T + \Delta P^T)x = 0$) for some $x = \begin{bmatrix} x_0^T & \cdots & x_q^T \end{bmatrix}^T$ such that $x_{j_0+1} = \cdots = x_q = 0$ but $x_0 \neq 0$ and $x_{j_0} \neq 0$. As j_0 and i_0 are not known in advance, the computations are performed by making q different choices of initial vectors given by the ordered sets $\{x_0^{(i)}, \ldots, x_q^{(i)}\}$, $i = 1, \ldots, q$. For $i = q$ the vectors are chosen randomly. For all other values of i, the i^{th} set is such that the first $i + 1$ vectors are random non zero vectors and the remaining vectors are zero. Also the corresponding computations $\inf_{x \in \mathcal{S}_q} \left\| \begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix} X_q X_q^\dagger \right\|_F$ and $\inf_{x \in \mathcal{S}_q} \left\| \begin{bmatrix} A_0^T & \cdots & A_k^T \end{bmatrix} X_q X_q^\dagger \right\|_F$ are performed in such a way that the updated versions of $\{x_{i+1}^{(i)}, \ldots x_q^{(i)}\}$ obtained after each iteration are put to 0. Let the computed values of $\inf_{x \in \mathcal{S}_q} \left\| \begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix} X_q X_q^\dagger \right\|_F$ and $\inf_{x \in \mathcal{S}_q} \left\| \begin{bmatrix} A_0^T & \cdots & A_k^T \end{bmatrix} X_q X_q^\dagger \right\|_F$ so obtained be g_i and e_i respectively. Observe that computing g_i is equivalent to finding $\inf_{x \in \mathcal{S}_i} \left\| \begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix} X_i X_i^\dagger \right\|_F$ and computing e_i is equivalent to finding $\inf_{x \in \mathcal{S}_i} \left\| \begin{bmatrix} A_0^T & \cdots & A_k^T \end{bmatrix} X_i X_i^\dagger \right\|_F$ without any restrictions on the vectors $x_0, \cdots x_i$ that form X_i. Setting $g_0 := \sigma_{\min}\left(\begin{bmatrix} A_0^T & \cdots & A_k^T \end{bmatrix}^T\right)$ and $e_0 := \sigma_{\min}\left(\begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix}\right)$, theoretically

$$g_i = \gamma_i^{(F)}, i = 0, \ldots, q \tag{5.7.1}$$

$$e_i = \eta_i^{(F)}, i = 0, \ldots, q. \tag{5.7.2}$$

In fact if the aim is to compute $\gamma_{k(n-1)}^{(F)}$ or $\eta_{k(n-1)}^{(F)}$, then larger sets of vectors

$$\{x_0, \ldots, x_{k(n-1)}\}$$

are initialized in an identical manner and g_i and e_i for $0 < i \leqslant k(n-1)$ are computed. Then the above equalities hold in theory for all values of i up to $k(n-1)$. Suppose that $S := \{j_0, \ldots, j_m\}$ is the ordered list of right minimal indices of a nearest singular matrix polynomial $(P + \Delta P)(\lambda)$. Then except for the convolution matrices $C_{j_0}(P + \Delta P), \ldots, C_{j_m}(P + \Delta P)$, the null space vectors in all the other convolution matrices $C_j(P + \Delta P)$, $j > j_0$, are linear combinations of null space vectors of lower level convolution matrices $C_l(P + \Delta P)$, $l \in S$ and their shifts. Consequently, for every j that is not a right minimal index of *any* of the nearest singular matrix polynomials, the infimum in $\inf_{x \in \mathcal{S}_j} \| \begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix} X_j X_j^{\dagger} \|_F$ is likely to be attained by vectors $\{x_0, \ldots, x_j\}$ which have a special relationship between them. Numerically, every iteration of the algorithm used to compute this infimum produces an X_j that has full rank so that in the final output the computed optimal vectors $\{x_0^c, \ldots, x_j^c\}$ satisfy $x_0^c \neq 0$ and $x_j^c \neq 0$. For random choices of initial vectors, the vectors $\{x_0^c, \ldots, x_j^c\}$ may not satisfy the special relationships that should hold to attain the infimum. An identical statement holds for the null spaces of the convolution matrices of $(P + \Delta P)(\lambda)^T$. Therefore, computationally the relations (5.7.1) and (5.7.2) may not hold and the computed values of g_i and e_i *do not necessarily* form a non-increasing sequence.

However, among all nearest singular matrix polynomials $(P + \Delta P)(\lambda)$, there is one with a least right (left) minimal index j_0 (i_0) that belongs to $\{0, \ldots, q\}$. In the first case the infimum corresponding to g_{j_0} and in the second case the infimum corresponding to e_{i_0} is likely to be attained for random choices of initial vectors. If $0 \leqslant j_0 \leqslant q$ and the value of g_{j_0} is optimally computed then,

$$\gamma_q^{(F)} = g_{j_0} = \min_{0 \leqslant i \leqslant q} g_i = \delta_F^{(n-1)}(P).$$

Likewise, if $0 \leqslant i_0 \leqslant q$ and the value of e_{i_0} is optimally computed then,

$$\eta_q^{(F)} = e_{i_0} = \min_{0 \leqslant i \leqslant q} e_i = \delta_F^{(n-1)}(P).$$

Therefore the computed value of $\delta_F^{(n-1)}(P)$ may be set as *the smallest among* $\min_{1 \leqslant i \leqslant q} g_i$, $\min_{1 \leqslant i \leqslant q} e_i$, $\sigma_{\min}\left(\begin{bmatrix} A_0^T & \cdots & A_k^T \end{bmatrix}^T\right)$ and $\sigma_{\min}\left(\begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix}\right)$. Note that if a nearest singular matrix polynomial also has right (left) minimal indices for some other values of i and the corresponding g_i (e_i) values are optimally computed, then it is possible that they are also good approximations of $\gamma_q^{(F)}$ ($\eta_q^{(F)}$) and hence of the distance. The approach is illustrated in the following example.

Example 5.7.2. Consider the 8×8 matrix pencil $\lambda A_1 + A_0$ arising from the model of a two dimensional, three link mobile manipulator from [54] which also appeared as [12, Example 14]. The matrices A_0 and A_1 are given by

$$A_0 = \begin{bmatrix} 0 & I_3 & 0 \\ -K_0 & -D_0 & F_0^T \\ F_0 & 0 & 0 \end{bmatrix} \text{ and } A_1 = \begin{bmatrix} -I_3 & 0 & 0 \\ 0 & -M_0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

where

$$M_0 = \begin{bmatrix} 18.7532 & -7.94493 & 7.94494 \\ -7.94493 & 31.8182 & -26.8182 \\ 7.94494 & -26.8182 & 26.8182 \end{bmatrix},$$

$$D_0 = \begin{bmatrix} -1.52143 & -1.55168 & 1.55168 \\ 3.22064 & 3.28467 & -3.28467 \\ -3.22064 & -3.28467 & 3.28467 \end{bmatrix},$$

$$K_0 = \begin{bmatrix} 67.4894 & 69.2393 & -69.2393 \\ 69.8124 & 1.68624 & -1.68617 \\ -69.8123 & -1.68617 & -68.2707 \end{bmatrix} \text{ and } F_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

In [54], the optimizations $\inf_{x \in \mathcal{S}_i} \left\| \begin{bmatrix} A_0 & A_1 \end{bmatrix} X_i X_i^\dagger \right\|_F$, $i = 1, \ldots, 7$, are computed for the above mentioned matrix pencil via the Structured Low Rank Approximation (SLRA) algorithm [62] and the computed γ-sequence for this example appears to exhibit properties contrary to the theoretical results. If the same sequence is computed via BFGS with random initializations, then we would actually be computing $g_i, i = 1, \ldots, 7$. The values are given in Table 5.7.1

g_0	g_1	g_2	g_3	g_4	g_5	g_6	g_7
0.0113	0.0135	0.0112	0.0161	0.0195	0.0253	0.0508	0.0712

Table 5.7.1: γ-sequence for Example 5.7.2.

The corresponding $e_i, i = 1, \ldots, 7$ are given in Table 5.7.2

Here $g_0 = \gamma_0^{(F)} = \sigma_{\min} \left\{ \begin{bmatrix} A_0^T & A_1^T \end{bmatrix}^T \right\}$ and $e_0 = \eta_0^{(F)} = \sigma_{\min} \left\{ \begin{bmatrix} A_0 & A_1 \end{bmatrix} \right\}$. As anticipated, both sets of values do not form a non-increasing sequence. However, in this case $\left\lfloor \frac{k(n-1)}{2} \right\rfloor = 3$, and the least of all the g_i and e_i values is $g_2 = e_4 = 0.0112$.

e_0	e_1	e_2	e_3	e_4	e_5	e_6	e_7
0.0494	0.0477	0.0203	0.0196	0.0112	0.0137	0.0113	0.0157

Table 5.7.2: η-sequence for Example 5.7.2.

Therefore, computationally $\delta_F^{(n-1)}(P) = 0.0112$. Also 2 is the least value of a right minimal index and 4 is the least value of a left minimal index of all the nearest singular matrix pencils so that $j_0 = 2$, and $i_0 = 4$. Note that computationally,

$$\min\{\gamma_3^{(F)}, \eta_3^{(F)}\} = \gamma_3^{(F)} = \min_{0\leqslant i\leqslant 3} g_i = g_2 = 0.0112.$$

Also note that $k(n-1) - j_0 = 7 - 2 = 5$ and computationally

$$\eta_5^{(F)} = \min_{0\leqslant i\leqslant 5} e_i = e_4 = 0.0112.$$

In this case, the pencil also seems to be very close to another singular matrix pencil having a right minimal index 0 and a left minimal index 6.

An identical strategy is used to compute $\delta_2^{(n-1)}(P)$ via (5.4.4). The only difference is that in this case $k(n-1)+1$ initial choices of $\{x_0^{(i)}, \ldots, x_{k(n-1)}^{(i)}\}$ are made in such a way that for $i = k(n-1)+1$ all the vectors $\{x_0^{(i)}, \ldots, x_{k(n-1)}^{(i)}\}$ are randomly chosen and for all smaller values of i, the first i vectors are chosen randomly with the rest being set to zero. Each of them is used to compute $\inf_{x\in\mathcal{S}_{k(n-1)}} \left\| \begin{bmatrix} A_0 & \cdots & A_k \end{bmatrix} X_{k(n-1)} X_{k(n-1)}^{\dagger} \right\|_F$. If $\hat{g}_i$ be the value computed in the i^{th} round, then $\gamma_{k(n-1)}^{(2)} = \min_{1\leqslant i\leqslant k(n-1)+1} \hat{g}_i$. A value of i for which the minimization is achieved is one for which either $i = j_0+1$ or one such that a nearest singular matrix polynomial has $i-1$ as a right minimal index. The corresponding $\hat{g}_i$ when optimally computed, is likely to be a good estimate of $\delta_2^{(n-1)}(P)$. Therefore the computed value of $\delta_2^{(n-1)}(P)$ is taken to be $\min_{1\leqslant i\leqslant k(n-1)+1} \hat{g}_i$.

Observe that if the computation of $\delta_F^{(n-1)}(P)$ is based on Corollary 5.2.1 which extends the result in [12] to matrix polynomials, then the strategy for initialization would have to be the same as the one for $\delta_2^{(n-1)}(P)$ and it will involve computing $\inf_{x\in\mathcal{S}_j} \left\| \begin{bmatrix} A_0^T & \cdots & A_k^T \end{bmatrix} X_j X_j^{\dagger} \right\|_F$, for $j = 0, \ldots, k(n-1)$ with no restrictions on the entries of x. Clearly as the sizes of the matrices X_j increase when $j > \left\lfloor \frac{k(n-1)}{2} \right\rfloor$, this is more expensive than the computation strategy based on part (e) of Theorem 5.4.1.

A few more examples are provided to illustrate the theoretical results and compare the computed values of the distances with the corresponding upper and lower bounds. In each case, in order to provide some information about the right and/or left minimal indices of the computed nearest singular matrix polynomial, the computed values of the γ-sequence and η-sequence for the $|||\cdot|||_F$ norm and the γ-sequence for the $|||\cdot|||_2$ norm are mentioned. Recall that for the $|||\cdot|||_F$ norm, we refer to the computed values of the γ-sequence and η-sequence as $g_j, 0 \leqslant j \leqslant k(n-1)$ and $e_j, 0 \leqslant j \leqslant k(n-1)$ respectively. Similarly, for norm $|||\cdot|||_2$, the computed values of the γ-sequence are denoted by $\hat{g}_j, 1 \leqslant j \leqslant k(n-1)+1$. In each case, the values of g_j and e_j that give the computed $\delta_F^{(n-1)}(P)$ and the value of $\hat{g}_j$ that gives the computed $\delta_2^{(n-1)}(P)$ are boldfaced.

Both BFGS and `globalsearch.m` are used in computations involving the $|||\cdot|||_F$ norm. It is worth noting that both algorithms give identical results in every case. The BFGS algorithm requires finding the gradient of the objective function which is done in a manner identical to the one in Section 4.8. Since the gradient computation is difficult for the objective functions involved in the computation of the γ-sequence for the $|||\cdot|||_2$ norm, these values are computed only via `globalsearch.m`.

Example 5.7.3. Consider the matrix polynomial

$$P(\lambda) = \begin{bmatrix} -0.1414 & -0.1490 \\ 1.1928 & .9702 \end{bmatrix} + \lambda \begin{bmatrix} 0.8837 & 0.9969 \\ 0.2190 & 0.0259 \end{bmatrix} + \lambda^2 \begin{bmatrix} 0.6346 & 0.9689 \\ 0.6252 & -0.0649 \end{bmatrix} + \lambda^3 \begin{bmatrix} -1.9867 & 1.2800 \\ 0.6097 & -0.1477 \end{bmatrix}.$$

Table 5.7.3 records the computed values of the γ-sequence and η-sequence with respect to the $|||\cdot|||_F$ norm obtained via the BFGS and `globalsearch.m` algorithms and the γ-sequence for $|||\cdot|||_2$ norm computed via `globalsearch.m`. In this case $k(n-1) = 3$ and $\left\lfloor \frac{k(n-1)}{2} \right\rfloor = 1$. The computed value of $\delta_F^{(n-1)}(P)$ is given by entries g_2 and e_1. Hence computationally $j_0 = 2$ and $i_0 = 1$. The computed value of $\delta_2^{(n-1)}(P)$ is given by $\hat{g}_3$ which also corresponds to $j_0 = 2$.

Table 5.7.4 records the various upper and lower bounds of the distance to a nearest singular matrix polynomial calculated via `globalsearch.m` along with the computed values of the distances.

It is worth noting that although

$$\delta_F^{(n-1)}(P) = \min\left\{\sigma_{\min}\left(\begin{bmatrix} A^T & B^T \end{bmatrix}^T\right), \sigma_{\min}\left(\begin{bmatrix} A & B \end{bmatrix}\right)\right\}$$

for $P(\lambda) = A + \lambda B$ of size 2×2 [12, Corollary 3], this example shows that may not hold for 2×2 matrix polynomials.

The next two examples appeared in both [31] and [38].

i	0	1	2	3
γ-sequence for $\|\|\cdot\|\|_F$ norm (g_i)	2.13568497	1.73451440	**1.67654181**	1.72284863
η-sequence for $\|\|\cdot\|\|_F$ norm (e_i)	1.72284863	**1.67654181**	1.73451440	2.13568497
γ-sequence for $\|\|\cdot\|\|_2$ norm ($\hat{g}_{i+1}$)	1.76225044	1.46091727	**1.39302305**	1.72284863

Table 5.7.3: γ-sequence and η-sequence for Example 5.7.3.

Distance measured	Lower bounds			Estimates of the distance		Upper bounds	
	Equation (5.6.4)	Equation (5.6.6)	Theorem 5.6.2	BFGS	`globalsearch`	Equation (5.6.7)	Equation (5.6.10)
$\delta_F^{(1)}(P)$	1.35805159	0.98501932	1.06017226	1.67654181	1.67654181	1.72284863	1.72284863
$\delta_2^{(1)}(P)$	1.35805159	0.98501932	1.06017226	-	1.39302305	1.72284863	1.72284863

Table 5.7.4: Comparison of various bounds on the distance $\delta_s^{(1)}(P)$ for Example 5.7.3.

Example 5.7.4. Consider the matrix pencil

$$P(\lambda) = \begin{bmatrix} 0 & 0.04 & 0.89 \\ 0.15 & -0.02 & 0 \\ 0.92 & 0.11 & 0.066 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Table 5.7.5 records the computed values of the γ-sequence and η-sequence for the $\|\cdot\|_F$ norm obtained via the BFGS and `globalsearch.m` algorithms and the γ-sequence for $\|\cdot\|_2$ norm computed via `globalsearch.m`. In this case $k(n-1) = 2$ and $\left\lfloor \frac{k(n-1)}{2} \right\rfloor = 1$. The computed value of $\delta_F^{(n-1)}(P)$ is given by entries g_1 and e_1. Hence computationally $j_0 = 1$ and $i_0 = 1$. The computed value of $\delta_2^{(n-1)}(P)$ is given by $\hat{g}_2$ which also corresponds to $j_0 = 1$.

Table 5.7.6 records the various upper and lower bounds of the distance to a nearest singular matrix polynomial computed via `globalsearch.m` along with the computed values of the distances. The computed values of $\delta_F^{(2)}(P)$ from [31] and [38] are also provided for reference.

i	0	1	2
γ-sequence for $\|\cdot\|_F$ norm (g_i)	0.90449335	**0.11554629**	0.88861007
η-sequence for $\|\cdot\|_F$ norm (e_i)	0.88861007	**0.11554629**	0.90449335
γ-sequence for $\|\cdot\|_2$ norm $(\hat{g}_{i+1})$	0.86581312	**0.09215174**	0.88861007

Table 5.7.5: γ-sequence and η-sequence for Example 5.7.4.

Distance measured	Lower bounds			Estimates of the distance				Upper bounds	
	Equation (5.6.4)	Equation (5.6.6)	Theorem 5.6.2	Guglielmi et al. [38]	Giesebrecht et al. [31]	BFGS	`globalsearch`	Equation (5.6.7)	Equation (5.6.10)
$\delta_F^{(2)}(P)$	0.09215174	0.08883492	0.09058250	0.1193	0.115585	0.11554629	0.11554629	0.88861	0.1427448
$\delta_2^{(2)}(P)$	0.09215174	0.08883492	0.09012456	-	-	-	0.09215174	0.88861	0.1427448

Table 5.7.6: Comparison of various bounds on the distance $\delta_s^{(1)}(P)$ for Example 5.7.4.

Example 5.7.5. Consider the pencil

$$P(\lambda) = \begin{bmatrix} -1.79 & 0.1 & -0.6 \\ 0.84 & -0.54 & 0.49 \\ -0.89 & 0.3 & 0.74 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Table 5.7.7 records the computed values of the γ-sequence and η-sequence with respect to $\|\cdot\|_F$ norm obtained via the BFGS and `globalsearch.m` algorithms and the γ-sequence for $\|\cdot\|_2$ norm computed via `globalsearch.m`. In this case $k(n-1) = 2$ and $\left\lfloor \frac{k(n-1)}{2} \right\rfloor = 1$. The computed value of $\delta_F^{(n-1)}(P)$ is given by entries g_2 and e_0. Hence computationally $j_0 = 2$ and $i_0 = 0$. The computed value of $\delta_2^{(n-1)}(P)$ is given by $\hat{g}_2$ which corresponds to $j_0 = 1$.

Table 5.7.8 records the various upper and lower bounds of the distance to a nearest singular matrix polynomial computed via `globalsearch.m`. Also the computed values of the distances are recorded here together with the computed values of $\delta_F^{(2)}(P)$ from [31] and [38] for reference.

The final example is a 3×3 matrix polynomial from [30].

Example 5.7.6. Consider the matrix polynomial

$$P(\lambda) = A_0 + \lambda A_1 + \lambda^2 A_2$$

i	0	1	2
γ-sequence for $\|\cdot\|_F$ norm (g_i)	1.07219625	0.94957852	**0.94356417**
η-sequence for $\|\cdot\|_F$ norm (e_i)	**0.94356417**	0.94957852	1.07219625
γ-sequence for $\|\cdot\|_2$ norm ($\hat{g}_{i+1}$)	0.88222447	**0.74394001**	0.94356417

Table 5.7.7: γ-sequence and η-sequence for Example 5.7.5.

Distance measured	Lower bounds			Estimates of the distance				Upper bounds	
	Equation (5.6.4)	Equation (5.6.6)	Theorem 5.6.2	Guglielmi et al. [38]	Giesebrecht et al. [31]	BFGS	`globalsearch`	Equation (5.6.7)	Equation (5.6.10)
$\delta_F^{(2)}(P)$	0.743524	0.636092	0.648559	0.9438619	0.94356416	0.94356417	0.94356417	0.94356417	0.94356417
$\delta_2^{(2)}(P)$	0.743524	0.636092	0.648559	-	-	-	0.74394001	0.94356417	0.94356417

Table 5.7.8: Comparison of various bounds on the distance $\delta_s^{(2)}(P)$ for Example 5.7.5.

where

$$A_0 = \begin{bmatrix} 0.0278 & 0.0563 & 0.1141 \\ -0.1758 & 0.327 & -0.173 \\ -0.056 & 0.0321 & -0.075 \end{bmatrix},\ A_1 = \begin{bmatrix} -0.2122 & 0.363 & -0.1385 \\ 0.18027 & -0.151 & 0.469 \\ -0.106 & 0.212 & -0.1514 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} -0.0376 & 0.107 & 0.293 \\ 0.003 & -0.14914 & -0.2859 \\ 0.0577 & 0.1455 & 0.231 \end{bmatrix}.$$

Table 5.7.9 records the computed values of the γ-sequence and η-sequence with respect to $\|\cdot\|_F$ norm obtained via the BFGS and `globalsearch.m` algorithms and the γ-sequence for $\|\cdot\|_2$ norm computed via `globalsearch.m`. In this case $k(n-1) = 4$ and $\left\lfloor \frac{k(n-1)}{2} \right\rfloor = 2$. The computed value of $\delta_F^{(n-1)}(P)$ is given by entries g_2 and e_2. Hence computationally $j_0 = 2$ and $i_0 = 2$. The computed value of $\delta_2^{(n-1)}(P)$ is given by $\hat{g}_3$ which also corresponds to $j_0 = 2$. Table 5.7.10 records the various upper and lower bounds of the distance to a nearest singular matrix polynomial computed via the `globalsearch.m` along with the computed values of the distances. The computed value of $\delta_F^{(2)}(P)$ from [30] is also provided for reference.

i	0	1	2	3	4
γ-sequence for $\lVert\cdot\rVert_F$ norm (g_i)	0.09097717	0.08198622	**0.02660446**	0.04708049	0.18100480
η-sequence for $\lVert\cdot\rVert_F$ norm (e_i)	0.18100480	0.04708049	**0.02660446**	0.08198622	0.09097717
γ-sequence for $\lVert\cdot\rVert_2$ norm ($\hat{g}_{i+1}$)	0.08623423	0.08054015	**0.01898730**	0.04047697	0.18100480

Table 5.7.9: γ-sequence and η-sequence for Example 5.7.6.

Distance measured	Lower bounds			Estimates of the distance			Upper bounds	
	Equation (5.6.4)	Equation (5.6.6)	Theorem 5.6.2	Giesebrecht et al. [30]	BFGS	`globalsearch`	Equation (5.6.7)	Equation (5.6.10)
$\delta_F^{(2)}(P)$	0.01666888	0.01650657	0.01825576	0.026604	0.02660446	0.02660446	0.09097717	0.08063977
$\delta_2^{(2)}(P)$	0.01666888	0.01650657	0.01671034	-	-	0.01898730	0.09097717	0.08063977

Table 5.7.10: Comparison of various bounds on the distance $\delta_s^{(2)}(P)$ for Example 5.7.6.

Conclusion. In this chapter we have characterized rank deficient matrix polynomials. Using this characterization we have obtained an optimization formulation of the distance to a nearest rank deficient matrix polynomial. We have also found an additional characterization of the singular matrix polynomials. The characterizations have led to upper and lower bounds and optimization formulations of the distance to singularity in various norm settings. The formulation via structured singular values shows that computing the distance to singularity may be an NP-hard problem. Insight into the left and right minimal indices of a nearest singular matrix polynomial is provided and a strategy for computing the distance to singularity is proposed. The computed distances and the bounds are illustrated via numerical examples. The results compare favourably with those in the literature.

Conclusion

Given an $m \times n$ rectangular matrix polynomial $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$, of grade k, in this thesis we have introduced the notion of a generalized linearization also referred to as a g-linearization of $P(\lambda)$. We have constructed vector spaces of rectangular matrix pencils such that almost every matrix pencil in the space provides solutions of the complete eigenvalue problem for $P(\lambda)$ with the property that the left and right minimal indices and bases of $P(\lambda)$ can be easily extracted from those of the pencil. These spaces become the vector spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ introduced in [61] whenever $P(\lambda)$ is square. They also have the same properties with respect to g-linearizations that the spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ have with respect to linearizations (as shown in [14]) when $P(\lambda)$ is square and singular. The results provide a direct extension of the theory of the vector spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ to the case of rectangular matrix polynomials. We have also shown a process of extracting many different strong linearizations from almost every pencil in $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$. We believe that our work complements the recent work in [26] which allows the study of linearizations of rectangular matrix pencils in a vector space setting by introducing the block Kronecker ansatz spaces. While [26] gives the relationship between the particular block Kronecker ansatz spaces $\mathbb{G}_1(P)$ and $\mathbb{G}_k(P)$ and the spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ respectively when $P(\lambda)$ is square and regular, our work extends the notion of the spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ to the rectangular case and shows the relationship between strong linearizations of $P(\lambda)$ extracted from the pencils in $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ and the linearizations in $\mathbb{G}_1(P)$ and $\mathbb{G}_k(P)$. A global backward error analysis of the process of solving the complete eigenvalue problem for $P(\lambda)$ via the linearizations that can be extracted from strong g-linearizations in $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ was also conducted. It showed that these g-linearizations provide a wide choice of linearizations that can solve the eigenvalue problem for $P(\lambda)$ in a globally backward stable manner.

Next we consider the problem of finding the distance to a nearest matrix polynomial with an elementary divisor of the form $(\lambda - \lambda_0)^j, j \geqslant r$, for a given square matrix polynomial $P(\lambda)$, $\lambda_0 \in \mathbb{C}$ and $r \geqslant 2$. The distance is shown to be zero for singular matrix polynomials. Square matrix polynomials with elementary divisors $(\lambda - \lambda_0)^j, j \geqslant r$, have been characterized in terms of rank deficiency of certain block Toeplitz matrices. Using the characterization, an optimization formulation of the distance has been obtained. This has been used to compute the distance via numerical software like BFGS and MATLAB's `globalsearch`. Upper and lower bounds have been derived and numerical experiments performed to compare them with the computed values of the distance show that they are quite tight in many cases. Also the characterization has been used to show that the distance is the reciprocal of a generalized notion of a μ-value. Since μ-value computation is an NP-hard problem, it is conjectured that the solution of the given distance problem is also NP-hard. The optimizations involved in the calculations are computationally quite expensive. But this is also the case with other optimizations proposed in the literature for computing similar distances. Also due to the nature of the optimizations, it is not clear that the values of the bounds from Theorem 4.5.1, Theorem 4.5.2 and Theorem 4.6.1 are the globally optimal values. However in many cases they are very close to the computed values of the distance. This leaves the question whether they may actually give the exact solution of the distance problem open for future research.

Finally in this thesis the distance from a given $n \times n$ regular matrix polynomial $P(\lambda)$ of degree k, to a nearest singular matrix polynomial of normal rank at most r for a given $r, (0 < r \leqslant n-1)$ has been considered with special emphasis on the distance to singularity, where $r = n - 1$. Square matrix polynomials with normal rank at most r have been characterized via a suitable rank deficiency criterion of the convolution matrix $C_{kr}(P)$. An additional characterization involving a different block Toeplitz matrix has been obtained for the case $r = n - 1$. The results have been used to set up the optimizations for computing nearest matrix polynomials of normal rank at most r in general, and nearest singular matrix polynomials in particular, under two different norm settings. They have also been used to formulate the distance as the reciprocal of a generalized structured singular value when $r < n-1$ and the reciprocal of a structured singular value or μ-value when $r = n-1$, thus establishing that the latter is likely to be an NP-hard problem. We conjecture that the computing the distance is also NP-hard when $r < n - 1$. The results also provide information about the possible left and right minimal indices of a nearest singular matrix pencil. At the same time they highlight the numerical challenges of computing the distance to nearest rank deficient matrices. They also provide a means of interpreting the numerical results and for devising a numerical strategy which we have used for calculating the distance to singularity via BFGS and MATLAB's `globalsearch` algorithm. The numerical examples show that the distances computed by this strategy compare favourably with those in the literature and the bounds are tight in many cases. Therefore although the computations involve variable projection least squares optimizations and it cannot be claimed that the exact distance has been calculated, it appears that the strategy is fairly efficient at least for problems that are not large.

Publications/Submissions from the Thesis

Publications:

1. The following publication is based on a part of Chapter 4 of this thesis:

   Ashish Kothyari, Biswajit Das, Shreemayee Bora, and Madhu N Belur. On the distance to singular descriptor dynamical systems with impulsive initial conditions. *IEEE Transactions on Automatic Control*, 64(3):1137 - 1149, 2019.

2. The following publication is based on Chapters 2 and 3 of this thesis:

   Biswajit Das and Shreemayee Bora. Vector spaces of generalized linearizations for rectangular matrix polynomials. *Electronic Journal of Linear Algebra*, 35:116-155, 2019.

3. The following manuscript based on Chapter 4 of this thesis has been accepted for publication:

   Biswajit Das and Shreemayee Bora. Nearest matrix polynomials with a specified elementary divisor. *SIAM Journal on Matrix Analysis and Applications*, 2020.

Submissions:

1. The following manuscript based on Chapter 5 of this thesis has been submitted for publication:

   Biswajit Das and Shreemayee Bora. Nearest rank deficient matrix polynomials.

Bibliography

[1] Sk Safique Ahmad and Rafikul Alam. Pseudospectra, critical points and multiple eigenvalues of matrix polynomials. *Linear Algebra and its Applications*, 430(4):1171–1195, 2009.

[2] JC Zúniga Anaya and Didier Henrion. An improved Toeplitz algorithm for polynomial matrix null-space computation. *Applied Mathematics and Computation*, 207(1):256–272, 2009.

[3] Efstathios N Antoniou and Stavros Vologiannidis. A new family of companion forms of polynomial matrices. *Electronic Journal of Linear Algebra*, 11(411):78–87, 2004.

[4] Daniel L Boley. The algebraic structure of pencils and block Toeplitz matrices. *Linear Algebra and its Applications*, 279(1-3):255–279, 1998.

[5] Shreemayee Bora, Michael Karow, Christian Mehl, and Punit Sharma. Structured eigenvalue backward errors of matrix pencils and polynomials with Hermitian and related structures. *SIAM Journal on Matrix Analysis and Applications*, 35(2):453–475, 2014.

[6] Lyonell Boulton, Peter Lancaster, and Panayiotis Psarrakos. On pseudospectra of matrix polynomials and their boundaries. *Mathematics of Computation*, 77(261):313–334, 2008.

[7] Richard P Braatz, Peter M Young, John C Doyle, and Manfred Morari. Computational complexity of μ calculation. *IEEE Transactions on Automatic Control*, 39(5):1000–1002, 1994.

[8] Kathryn Eleda Brenan, Stephen L Campbell, and Linda Ruth Petzold. *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*. SIAM Publications, Philadelphia, 1996.

[9] María I Bueno, Fernando De Terán, and Froilán M Dopico. Recovery of eigenvectors and minimal bases of matrix polynomials from generalized fiedler linearizations. *SIAM Journal on Matrix Analysis and Applications*, 32(2):463–483, 2011.

[10] María I Bueno, Froilán M Dopico, Javier Pérez, Rafael Saavedra, and Bradley Zykoski. A simplified approach to Fiedler-like pencils via block minimal bases pencils. *Linear Algebra and its Applications*, 547:45–104, 2018.

[11] María I Bueno and Susana Furtado. Palindromic linearizations of a matrix polynomial of odd degree obtained from Fiedler pencils with repetition. *Electronic Journal Linear Algebra*, 23:562–577, 2012.

[12] Ralph Byers, Chunyang He, and Volker Mehrmann. Where is the nearest non-regular pencil? *Linear Algebra and its Applications*, 285(1-3):81–105, 1998.

[13] Ralph Byers and Nancy K Nichols. On the stability radius of a generalized state-space system. *Linear Algebra and its Applications*, 188:113–134, 1993.

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

[15] Fernando De Terán, Froilán M Dopico, and D Steven Mackey. Fiedler companion linearizations and the recovery of minimal indices. *SIAM Journal on Matrix Analysis and Applications*, 31(4):2181–2204, 2010.

[16] Fernando De Terán, Froilán M Dopico, and D Steven Mackey. Palindromic companion forms for matrix polynomials of odd degree. *Journal of Computational and Applied Mathematics*, 236(6):1464–1480, 2011.

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

[18] Fernando De Terán, Froilán M Dopico, and D Steven Mackey. Spectral equivalence of matrix polynomials and the index sum theorem. *Linear Algebra and its Applications*, 459:264–333, 2014.

[19] James Demmel and Bo Kågström. The generalized Schur decomposition of an arbitrary pencil A–λB: robust software with error bounds and applications. Part I: theory and algorithms. *ACM Transactions on Mathematical Software (TOMS)*, 19(2):160–174, 1993.

[20] James Demmel and Bo Kågström. The generalized Schur decomposition of an arbitrary pencil A–λB: robust software with error bounds and applications. Part II: software and applications. *ACM Transactions on Mathematical Software (TOMS)*, 19(2):175–201, 1993.

[21] Andrii Dmytryshyn, Stefan Johansson, Bo Kågström, and Paul Van Dooren. Geometry of matrix polynomial spaces. *Foundations of Computational Mathematics*, 20:423–450, 2020.

[22] Froilán M Dopico, Piers W Lawrence, Javier Pérez, and Paul Van Dooren. Block Kronecker linearizations of matrix polynomials and their backward errors. *Numerische Mathematik*, 140(2):373–426, 2018.

[23] Nguyen Huu Du, Vu Hoang Linh, and Volker Mehrmann. Robust stability of differential-algebraic equations. In Achim Ilchmann and Timo Reis, editors, *Surveys in Differential-Algebraic Equations I*, pages 63–95. Springer, 2013.

[24] Heike Faßbender, D Steven Mackey, Niloufer Mackey, and Christian Schroeder. Structured polynomial eigenproblems related to time-delay systems. *Electronic Transactions on Numerical Analysis*, 31:306–330, 2008.

[25] Heike Faßbender and Philip Saltenberger. On vector spaces of linearizations for matrix polynomials in orthogonal bases. *Linear Algebra and its Applications*, 525:59–83, 2017.

[26] Heike Faßbender and Philip Saltenberger. Block Kronecker ansatz spaces for matrix polynomials. *Linear Algebra and its Applications*, 542:118–148, 2018.

[27] G David Forney, Jr. Minimal bases of rational vector spaces, with applications to multivariable linear systems. *SIAM Journal on Control*, 13(3):493–520, 1975.

[28] Vyacheslav Futorny, Tetiana Klymchuk, Vladimir V Sergeichuk, and Nadya Shvai. A constructive proof of Pokrzywa's theorem about perturbations of matrix pencils. *arXiv preprint arXiv:1907.03213*, 2019.

[29] Felix Ruvimovich Gantmacher. *The Theory of Matrices*, volume 2. Chelsea, New York, 1959.

[30] Mark Giesbrecht, Joseph Haraldson, and George Labahn. Computing the nearest rank-deficient matrix polynomial. In *Proceedings of the 2017 ACM on International Symposium on Symbolic and Algebraic Computation*, pages 181–188. ACM, 2017.

[31] Mark Giesbrecht, Joseph Haraldson, and George Labahn. Computing lower rank approximations of matrix polynomials. *Journal of Symbolic Computation*, 98:225–245, 2020.

[32] Israel Gohberg, Peter Lancaster, and Leiba Rodman. *Matrix Polynomials*. Academic Press, New York, 1982.

[33] Gene H Golub and Charles F Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, 2013.

[34] Juan M Gracia, Inmaculada de Hoyos, and Ion Zaballa. Perturbation of linear control systems. *Linear Algebra and its Applications*, 121:353–383, 1989.

[35] Michael Grant and Stephen Boyd. Graph implementations for nonsmooth convex programs. In Vincent D Blondel, Stephen P Boyd, and Hidenori Kimura, editors, *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, pages 95–110. Springer-Verlag Limited, 2008. `http://stanford.edu/~boyd/graph_dcp.html`.

[36] Michael Grant and Stephen Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. `http://cvxr.com/cvx`, March 2014.

[37] Stefano Grivet-Talocia. Passivity enforcement via perturbation of Hamiltonian matrices. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 51(9):1755–1769, 2004.

[38] Nicola Guglielmi, Christian Lubich, and Volker Mehrmann. On the nearest singular matrix pencil. *SIAM Journal on Matrix Analysis and Applications*, 38(3):776–806, 2017.

[39] George Halikias, G Galanis, Nicos Karcanias, and Efstathios Milonidis. Nearest common root of polynomials, approximate greatest common divisor and the structured singular value. *IMA Journal of Mathematical Control and Information*, 30(4):423–442, 2012.

[40] Nicholas J Higham, Ren-Cang Li, and Françoise Tisseur. Backward error of polynomial eigenproblems solved by linearization. *SIAM Journal on Matrix Analysis and Applications*, 29(4):1218–1241, 2007.

[41] Nicholas J Higham, D Steven Mackey, Niloufer Mackey, and Françoise Tisseur. Symmetric linearizations for matrix polynomials. *SIAM Journal on Matrix Analysis and Applications*, 29(1):143–159, 2006.

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

[43] Roger A Horn and Charles R Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

[44] Thomas Kailath. *Linear Systems*. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1980.

[45] Nicos Karcanias and Grigoris Kalogeropoulos. On the Segré, Weyr characteristics of right (left) regular matrix pencils. *International Journal of Control*, 44(4):991–1015, 1986.

[46] Nicos Karcanias and Grigoris Kalogeropoulos. Right, left characteristic sequences and column, row minimal indices of a singular pencil. *International Journal of Control*, 47(4):937–946, 1988.

[47] Michael Karow. *Geometry of Spectral Value Sets*. PhD thesis, University of Bremen, 2003.

[48] Michael Karow. μ-values and spectral value sets for linear perturbation classes defined by a scalar product. *SIAM Journal on Matrix Analysis and Applications*, 32(3):845–865, 2011.

[49] Michael Karow, Daniel Kressner, and Emre Mengi. Nonlinear eigenvalue problems with specified eigenvalues. *SIAM Journal on Matrix Analysis and Applications*, 35(3):819–834, 2014.

[50] Michael Karow and Emre Mengi. Matrix polynomials with specified eigenvalues. *Linear Algebra and its Applications*, 466:457–482, 2015.

[51] Esmaeil Kokabifar, Panayiotis Psarrakos, and Ghasem Barid Loghmani. On the distance from a matrix polynomial to matrix polynomials with some prescribed eigenvalues. *Linear Algebra and its Applications*, 544:158–185, 2018.

[52] Ashish Kothyari, Biswajit Das, Shreemayee Bora, and Madhu N Belur. On the distance to singular descriptor dynamical systems with impulsive initial conditions. *IEEE Transactions on Automatic Control*, 64(3):1137–1149, 2019.

[53] Daniel Kressner, Emre Mengi, Ivica Nakić, and Ninoslav Truhar. Generalized eigenvalue problems with specified eigenvalues. *IMA Journal of Numerical Analysis*, 34(2):480–501, 2014.

[54] Daniel Kressner and Matthias Voigt. Distance problems for linear dynamical systems. In Peter benner, Matthias Bollhöfer, Daniel Kressner, Christian Mehl, and Tatjana Styket, editors, *Numerical Algebra, Matrix Theory, Differential-Algebraic Equations and Control Theory*, pages 559–583. Springer, 2015.

[55] Peter Kunkel and Volker Mehrmann. *Differential-Algebraic Equations: Analysis and Numerical Solution*. EMS Publising House, Zürich, Switzerland, 2006.

[56] René Lamour, Roswitha März, and Caren Tischendorf. *Differential-Algebraic Equations: A Projector Based Analysis*. Differential-Algebraic Equations Forum, Springer-Verlag, 2013.

[57] Peter Lancaster and Leiba Rodman. Canonical forms for Hermitian matrix pairs under strict equivalence and congruence. *SIAM Review*, 47(3):407–443, 2005.

[58] Adrian S Lewis and Michael L Overton. Nonsmooth optimization via quasi-Newton methods. *Mathematical Programming*, 141(1-2):135–163, 2013.

[59] Laszlo Losonczi. Eigenvalues and eigenvectors of some tridiagonal matrices. *Acta Mathematica Hungarica*, 60(3-4):309–322, 1992.

[60] D Steven Mackey, Niloufer Mackey, Christian Mehl, and Volker Mehrmann. Structured polynomial eigenvalue problems: Good vibrations from good linearizations. *SIAM Journal on Matrix Analysis and Applications*, 28(4):1029–1051, 2006.

[61] D Steven Mackey, Niloufer Mackey, Christian Mehl, and Volker Mehrmann. Vector spaces of linearizations for matrix polynomials. *SIAM Journal on Matrix Analysis and Applications*, 28(4):971–1004, 2006.

[62] Ivan Markovsky. *Low Rank Approximation: Algorithms, Implementation, Applications*. Berlin, Germany, Springer-Verlag, 2012.

[63] Christian Mehl, Volker Mehrmann, and Michal Wojtylak. On the distance to singularity via low rank perturbations. *Operators and Matrices*, 9(4):733–772, 2015.

[64] Yuji Nakatsukasa, Vanni Noferini, and Alex Townsend. Vector spaces of linearizations for matrix polynomials: a bivariate polynomial approach. *SIAM Journal on Matrix Analysis and Applications*, 38(1):1–29, 2017.

[65] Andrew Packard and John Doyle. The complex structured singular value. *Automatica*, 29(1):71–109, 1993.

[66] Nikolaos Papathanasiou and Panayiotis Psarrakos. The distance from a matrix polynomial to matrix polynomials with a prescribed multiple eigenvalue. *Linear Algebra and its Applications*, 429(7):1453–1477, 2008.

[67] Panayiotis J Psarrakos. Distance bounds for prescribed multiple eigenvalues of matrix polynomials. *Linear Algebra and its Applications*, 436(11):4107–4119, 2012.

[68] Ricardo Riaza. *Differential-Algebraic Systems: Analytical Aspects and Circuit Applications*. World Scientific Publishing Co. Pte. Ltd., Singapore, 2008.

[69] Howard Harry Rosenbrock. *State-space and Multivariable Theory*. Nelson, 1970.

[70] Punit Sharma. *Eigenvalue Backward Errors of Polynomial Eigenvalue Problems under Structure Preserving Perturbations*. PhD thesis, Indian Institute of Technology Guwahati, 2015.

[71] Ji-Guang Sun. Backward perturbation analysis of certain characteristic subspaces. *Numerische Mathematik*, 65(1):357–382, 1993.

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

[73] Zsolt Ugray, Leon Lasdon, John Plummer, Fred Glover, James Kelly, and Rafael Martí. Scatter search and local nlp solvers: A multistart framework for global optimization. *INFORMS Journal on Computing*, 19(3):328–340, 2007.

[74] Paul Van Dooren. The computation of Kronecker's canonical form of a singular pencil. *Linear Algebra and its Applications*, 27:103–140, 1979.

[75] Paul Van Dooren and Patrick Dewilde. The eigenstructure of an arbitrary polynomial matrix: computational aspects. *Linear Algebra and its Applications*, 50:545–579, 1983.

[76] Paul Van Dooren, Patrick Dewilde, and Joos Vandewalle. On the determination of the Smith-Macmillan form of a rational matrix from its Laurent expansion. *IEEE Transactions on Circuits and Systems*, 26(3):180–189, 1979.

[77] Joos Vandewalle and Patrick Dewilde. On the determination of the order and the degree of a zero of a rational matrix. *IEEE Transactions on Automatic Control*, 19(5):608–609, 1974.

[78] Antonis I G Vardulakis. *Linear Multivariable Control: Algebraic Analysis and Synthesis Methods*. J. Wiley, 1991.

[79] Andras Varga. On stabilization methods of descriptor systems. *Systems & Control Letters*, 24(2):133–138, 1995.

[80] Stavros Vologiannidis and Efstathios N Antoniou. A permuted factors approach for the linearization of polynomial matrices. *Mathematics of Control, Signals, and Systems*, 22(4):317–342, 2011.

[81] Wen-Chyuan Yueh. Eigenvalues of several tridiagonal matrices. *Applied Mathematics e-Notes*, 5:66–74, 2005.