LINEARIZATION, SENSITIVITY AND BACKWARD PERTURBATION ANALYSIS OF MULTIPARAMETER EIGENVALUE PROBLEMS

by

ARNAB GHOSH

DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
GUWAHATI-781039, INDIA
August 2017

LINEARIZATION, SENSITIVITY AND BACKWARD PERTURBATION ANALYSIS OF MULTIPARAMETER EIGENVALUE PROBLEMS

A thesis submitted
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

by

Arnab Ghosh

(Roll No. 09612307)

Department of Mathematics
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
August 2017

To

My Parents and Mou

Certificate

It is certified that the work contained in this thesis entitled **"Linearization, sensitivity and backward perturbation analysis of multiparameter eigenvalue problems"** by **Arnab Ghosh,** a student of 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.

August 2017

Professor Rafikul Alam
Department of Mathematics
Indian Institute of Technology Guwahati

DECLARATION

It is certified that the work contained in this thesis entitled **“Linearization, sensitivity and backward perturbation analysis of multiparameter eigenvalue problems”** has done by me, under the supervision of **Prof. Rafikul Alam**, Professor, Department of Mathematics, Indian Institute of Technology Guwahati for the award of the degree of Doctor of Philosophy and this work has not been submitted elsewhere for a degree.

August 2017

Arnab Ghosh
Roll No. 09612307
Department of Mathematics
Indian Institute of Technology Guwahati

Acknowledgement

First and foremost, I would like to express my sincerest gratitude to my thesis supervisor Prof. Rafikul Alam. He has been very supportive, inspiring me all along the way, and has shown extraordinary patience with me. I am very grateful to him for introducing me Numerical Linear Algebra, for all his contributions of time, and ideas to make my Ph.D experience productive and stimulating. He has also provided insightful discussions about the research and during this period I have been able to learn many things with his help. Without his guidance and persistent help this dissertation would not have been possible.

I want to convey my sincere thanks to the doctoral committee of members Prof. Rajen Kumar Sinha, Dr. Shreemayee Bora and Dr. Anjan Kumar Chakrabarty for their advice and valuable suggestions on this thesis and giving their precious time to evaluate the progress of my work. I am very grateful to Prof. M. Guru Prem Prasad, Dean of Academic Affairs, IIT Guwahati and Prof. Swaroop Nandan Bora, Head of the Department, Department of Mathematics, IIT Guwahati for their support and encouragement for carrying the research.

I sincerely acknowledge Indian Institute of Technology Guwahati for facilities provided to me during my research work. I am most grateful to Ministry of Human Resource and Development, Government of India, for providing me financial assistance for the completion of my thesis work. I thank our technical superintendent Mr. Shantanu Majumdar and Mr. Pranpratim Borgohain and office staff members Mr. Sridhar Samal, Mr. Phatik Kumar and Mr. Saurav Choudhury of the department, for their assistance in various ways during my research period.

I am also thankful to Dr. Sk. Safique Ahmad, Dr. Bibhas Adhikari and Dr. Namita Behera for their encouragement during the research work. I also thank Mr. Ranjan Kumar Das and Mr. Jibrail Ali for their help during this period.

I take this opportunity to convey my gratitude to Mr. Atanu Bera, Dr. Sandipan Gupta, Dr. Sudipta Purkait, Dr. Ahin Ray, Mr. Sounak Naha, Dr. Debopam Chakraborty, Dr. Imran Hossain, Mr. Gundappa Saha, Mr. Anabil Gayen, Mr. Pradip Das, Mr. Indrajit Das and Mr. Siddhanta Roy for all the unforgettable memories which I will cherish forever. I also thank Kaushik, Murali, Himadri, Manideepa di, Biswajit da, Santu, Mandar da, Kalyan, Barun, Abhishek, Anirban, Diganta, Arghya and many others with whom I have shared some of the best moments of my life at IITG.

I pay my sincere regards to my grandmother Late Malati Ghosh and my uncle Late Nilratan

Ghosh. It is not possible for me to adequately express my heartfelt gratitude towards my parents Mr. Gopal Chandra Ghosh and Mrs. Madhuri Ghosh for everything they have done for me from the very first day of my life. I thank my eleder brother Mr. Biplab Ghosh and sister in law Mrs. Tanushree Ghosh. I also convey my gratitude to my father in law Mr. Haradhan Das, my mother in law Mrs. Anita Das and brother in law Mr. Prasenjit Das for their continuous support and love.

I cannot thank my brother Mr. Omkar Majumder enough for his support, encouragement and love in the most difficult situations of my life. I convey my gratitude to uncle Mr. Sanjoy Majumder and aunty Mrs. Ramala Majumder for their warmth love.

I have no words to express my love and gratitude towards my beloved Mou who has been part and parcel of my life since April, 2007. She stood like a pillar beside me in every failure and has been the driving force of my life. I have no adjective to express her countless sacrifices for me during this work. Hopefully we will walk together with more spirit, love and integrity for the rest of our lives by the grace of almighty.

Abstract

We develop a general framework for the sensitivity and backward perturbation analysis of linear and polynomial MEPs. For a general norm on the space of MEPs, we define the condition number $\mathrm{cond}(\lambda, \mathbb{W})$ of a simple eigenvalue λ of an MEP $\mathbb{W}$ and derive three equivalent representations of $\mathrm{cond}(\lambda, \mathbb{W})$. We also analyze holomorphic perturbation of a simple eigenvalue of $\mathbb{W}$. Further, we define the backward error $\eta(\lambda, \mathbb{W})$ of an approximate eigenvalue λ of $\mathbb{W}$. We determine $\eta(\lambda, \mathbb{W})$ and construct an optimal perturbation $\Delta\mathbb{W}$ such that $\lambda \in \sigma(\mathbb{W} + \Delta\mathbb{W})$ and $|||\Delta\mathbb{W}||| = \eta(\lambda, \mathbb{W})$. We also define the backward error $\eta(\lambda, x, \mathbb{W})$ of an approximate eigenpair $(\lambda, x := x_1 \otimes \cdots \otimes x_m) \in \mathbb{C}^m \times (\mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m})$. We determine $\eta(\lambda, x, \mathbb{W})$ and construct an optimal perturbation $\Delta\mathbb{W}$ such that $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ and $|||\Delta\mathbb{W}||| = \eta(\lambda, x, \mathbb{W})$.

We define and analyze two vector spaces, namely, the right ansatz space and the left ansatz space of potential linearizations of a two-parameter matrix polynomial of the form $P(\lambda, \mu) := \sum_{i=0}^{k} \sum_{j=0}^{k-i} \lambda^i \mu^j A_{ij}$. We also analyze conditions under which a pencil in the ansatz spaces is a linearization of $P(\lambda, \mu)$.

We consider structured linear MEPs and define structured backward error $\eta^{\mathbb{S}}(\lambda, x, \mathbb{W})$ of an approximate eigenpair $(\lambda, x := x_1 \otimes \cdots \otimes x_m)$. We determine $\eta^{\mathbb{S}}(\lambda, x, \mathbb{W})$ and construct an optimal structured perturbation $\Delta\mathbb{W}$ such that $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ and $|||\Delta\mathbb{W}||| = \eta^{\mathbb{S}}(\lambda, x, \mathbb{W})$.

Contents

1 Introduction 1

1.1 Preliminaries 6

1.1.1 Vector space of linearizations of matrix polynomial 10

1.1.2 Nonsingular MEP 12

2 Sensitivity and Backward Perturbation Analysis of Nonhomogeneous MEP 17

2.1 Introduction 17

2.2 Nonhomogeneous MEP 18

2.2.1 Norms on $\mathbb{MEP}(n_1, \cdots, n_m)$ 22

2.3 Sensitivity analysis of simple eigenvalues 25

2.4 Holomorphic perturbations 33

2.5 Backward perturbation analysis 36

3 Polynomial MEP **41**

3.1 Introduction 41

3.2 Polynomial MEP 42

3.2.1 Norms on $\mathbb{PMEP}(k; n_1, n_2)$ 46

3.3 Sensitivity analysis of a simple eigenvalue 47

3.4 Holomorphic perturbations 52

3.5 Backward perturbation analysis 54

4 Sensitivity and Backward Perturbation Analysis of Homogeneous MEP 59

4.1 Introduction 59

4.2 Homogeneous MEP . . . 59
4.3 Sensitivity analysis of simple eigenvalues . . . 68
4.4 Holomorphic perturbations . . . 75
4.5 Backward perturbation analysis . . . 77

5 Vector Space of Linearizations of a two-parameter PEP 83
5.1 Introduction . . . 83
5.2 Ansatz spaces for two-parameter polynomials . . . 85
  5.2.1 Right and left ansatz spaces . . . 92
  5.2.2 Linearizations in $\mathbb{L}_1(P)$ . . . 107
  5.2.3 Double ansatz space . . . 114
5.3 Linearizations of a two-parameter PEP . . . 115

6 Structured Backward Perturbation Analysis of MEP 117
6.1 Introduction . . . 117
6.2 Structured MEPs . . . 118
6.3 Frobenious norm structured backward error . . . 124
6.4 Spectral norm structured backward error . . . 144

7 Summary 157

List of Symbols

$\mathbb{N}$	the set of the natural numbers
$\mathbb{Z}$	the set of the integers
$\mathbb{R}$	the field of the real numbers
$\mathbb{C}$	the field of the complex numbers
$\mathbb{C}^{m\times n}$	the space of complex matrices of size $m \times n$
e_k	the vector with the kth component equal to 1 and all the other entries are zero.
A^T	the transpose of a matrix $A \in \mathbb{C}^{m\times n}$
A^{-1}	the inverse of a matrix $A \in \mathbb{C}^{n\times n}$
A^*	complex conjugate transpose of a matrix $A \in \mathbb{C}^{m\times n}$
$\mathrm{rank}(A)$	rank of matrix A
I_n	the identity matrix of order n
$A \otimes B$	Kronecker product of matrix A and B
$A^{\mathcal{B}}$	block transpose of matrix A
$\mathrm{adj}(A)$	adjugate of matrix $A \in \mathbb{C}^{n\times n}$
$\det(A)$	the determinant of a matrix $A \in \mathbb{C}^{n\times n}$
$\mathrm{tr}(A)$	the trace of a matrix $A \in \mathbb{C}^{n\times n}$
$\sigma(A)$	the spectrum of $A \in \mathbb{C}^{n\times n}$
$\sigma_{\max}(A)$	the largest nonzero singular value of $A \in \mathbb{C}^{m\times n}$

$\sigma_{\min}(A)$	the smallest nonzero singular value of $A \in \mathbb{C}^{m\times n}$
$\|x\|_p := (\sum\limits_{i=1}^{n} \|x_i\|^p)^{1/p}$	the Hölder p-norm on $\mathbb{C}^n$
$\|A\|_2 := \max_{\|x\|_2=1} \|Ax\|_2$	the spectral norm $\|.\|_2$ on $\mathbb{C}^{n\times n}$
$\|A\|_F := (\mathrm{tr}(A^*A))^{1/2}$	the Frobenius norm $\|.\|_F$ on $\mathbb{C}^{n\times n}$
$\langle A, B\rangle$	inner product of matrix A and B
$\|A\|_*$	dual norm of matrix norm $\|A\|$
$Df(a)$	the derivative of f at a
$\partial\|a\|$	subdifferential of $x \mapsto \|x\|$ at a

1

Introduction

Multiparameter eigenvalue problems (MEPs) arise in many applications and have been analyzed extensively over the years. A finite dimensional linear nonhomogeneous MEP is defined as follows. Let $\mathbb{W} : \mathbb{C}^m \longrightarrow \prod_{j=1}^m \mathbb{C}^{n_j \times n_j}$ be an affine map given by

$$\mathbb{W}(z) := (W_1(z), \ldots, W_m(z)), \text{ where } W_i(z) := A_i + \sum_{j=1}^{m} z_j B_{ij} \tag{1.1}$$

and $A_i, B_{ij} \in \mathbb{C}^{n_i \times n_i}$ for $i, j = 1, \cdots, m$. We say that $\mathbb{W}$ is *regular* if $\prod_{j=1}^m \det(W_j(z)) \neq 0$ for some $z \in \mathbb{C}^m$. For $x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ and $z \in \mathbb{C}^m$, we define

$$\mathbb{W}(z)x := (W_1(z)x_1, \ldots, W_m(z)x_m) \text{ and } x^*\mathbb{W}(z) := (x_1^* W_1(z), \ldots, x_m^* W_m(z)).$$

Suppose that $\mathbb{W}$ is regular. Let $\lambda \in \mathbb{C}^m$ and $x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ be a nonzero decomposable tensor. Then λ is said to be an *eigenvalue* of $\mathbb{W}$ and x is said to be a *right eigenvector* of $\mathbb{W}$ corresponding to λ if $\mathbb{W}(\lambda)x = 0$. Similarly, a nonzero decomposable tensor $y := y_1 \otimes \cdots \otimes y_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ is said to be a *left eigenvector* of $\mathbb{W}$ corresponding to λ if $y^*\mathbb{W}(\lambda) = 0$. The *multiparameter eigenvalue problem* is to solve the problems

$$\mathbb{W}(\lambda)x = 0 \text{ and } y^*\mathbb{W}(\lambda) = 0$$

for $\lambda \in \mathbb{C}^m$ and nonzero decomposable tensors x and y in $\mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$. We refer to the tuple $\mathbb{W}$ as a nonhomogeneous MEP.

The origin of multiparameter eigenvalue problems can be traced back to mathematical physics (for details, see [11, 13]). For example, the vibration of a fixed membrane is governed by the differential equation $\Delta u + k^2 u = 0$ in $\mathbb{R}^2$. For a rectangular membrane, the seperation of variables technique leads to two parameter (Sturm-Liouville) eigenvalue problem [52]

$$\left.\begin{aligned} y_1''(x_1) + \lambda_1 y_1(x_1) = 0 \\ y_2''(x_2) + \lambda_2 y_2(x_2) = 0. \end{aligned}\right\} \tag{1.2}$$

On the other hand, for a circular membrane, the angular and radial equation in polar coordinates form leads to the MEP [52]

$$\left.\begin{aligned} &\phi'' + \lambda_1\phi = 0 \\ &r^{-1}(rR')' + (\lambda_2 - \lambda_1 r^-2)R = 0. \end{aligned}\right\} \tag{1.3}$$

More generally, the separation of variables technique for solving a partial differential equation leads to multiparameter Sturm-Liouville eigenvalue problems with various boundary conditions. For example, considering the m intervals $[a_r, b_r]$ for $r = 1, \cdots, m$, we have the following multiparameter Sturm-Liouville eigenvalue problem in m parameters

$$\left.\begin{aligned} &y_r''(x_r) - q_r(x_r)y_r(x_r) + \textstyle\sum_{s=1}^m \lambda_s p_{rs}(x_r)y_r(x_r) = 0, \quad x_r \in [a_r, b_r] \\ &y_r(a_r)\cos\alpha_r = y_r'(a_r)\sin\alpha_r, \quad 0 \le \alpha_r < \pi \\ &y_r(b_r)\cos\beta_r = y_r'(b_r)\sin\beta_r, \quad 0 < \beta_r \le \pi, \end{aligned}\right\} \tag{1.4}$$

where $p_{rs}(x_r), q_r(x_r)$ are real valued continuous functions on $[a_r, b_r]$ for all $r, s = 1, \cdots, m$. A scalar tuple $(\lambda_1, \cdots, \lambda_m) \in \mathbb{C}^m$ for which a non-trivial solutions exist in (1.4) is an eigenvalue of the MEP (1.4).

Multiparameter spectral theory has applications in various fields and is an important area of research. R. D. Carmichael indicated the vast area awaiting investigation and suggested a possible line of attack in [20, 21, 22]. F.V.Atkinson [12] revived the theory back in 1960s. Given a linear MEP $\mathbb{W}$, he defined operator determinants Δ_i associated with $\mathbb{W}$ for $i = 0, \cdots, m$. A linear homogeneous MEP $\mathbb{W}$ is said to be nonsingular if $\Delta := \sum_{i=1}^m \alpha_i\Delta_i$ is invertible for some scalars $\alpha_0, \cdots, \alpha_m$. On the other hand, a linear nonhomogeneous MEP $\mathbb{W}$ is said to be nonsingular if Δ_0 is invertible. Atkinson [12] proved that a homogeneous MEP $\mathbb{W}$ is equivalent to the m coupled generalized eigenproblems $\Delta_i z = \Delta z$ for $i = 1, \cdots, m$. On the other hand, a nonhomogeneous MEP $\mathbb{W}$ is equivalent to the m coupled generalized eigenproblems

$\Delta_i z = \Delta_0 z \quad$ for $i = 1, \cdots, m$. Based on Atkinson's approach multiparameter spectral theory has been studied extensively in literature [14, 19, 28, 29, 37, 38, 39, 46]. Root subspaces associated with nonderogatory eigenvalues of a finite dimensional MEP is analyzed in [28, 37, 46]. Also a version of the Cayley-Hamilton theorem for MEP is proved in [39]. Structured MEPs such as hermitian, and left and right definite MEPs are analyzed in [12, 13, 14, 15, 16, 17, 18, 19, 52].

The development of numerical methods for computing eigenvalues and eigenvectors of MEP is an active area of research. A Jacobi-Davidson type method for the two-parameter eigenvalue problem is discussed in [30, 32]. Harmonic and refined extraction methods for an MEP is studied in [33]. Also several numerical methods are discussed in [43, 47] for a polynomial MEP.

The sensitivity and backward error analysis plays in important role in analyzing the accuracy of computed eigenvalues and eigenvectors. The condition number of an eigenvalue measures its sensitivity to small perturbations in the coefficient matrices. The condition number of an eigenvalue and backward error of an approximate eigenpair of an MEP are discussed in [31] for a specific choice of weighted norm on the space of MEPs. For $A \in \mathbb{C}^{m\times m}$ and $\theta := (\theta_1, \cdots, \theta_m)^T \in \mathbb{C}^m$, the θ-weighted norm is defined as

$$||A||_\theta := \max\{||Az||_2 : z = (z_1, \cdots, z_m)^T \in \mathbb{C}^m \text{ and } |z_i| = |\theta_i| \text{ for } i = 1, \cdots, m\}.$$

Then for a simple eigenvalue λ of $\mathbb{W}$ and an associated eigenvector $x = x_1 \otimes \cdots \otimes x_m$, the condition number of λ is defined as

$$\begin{aligned} \kappa(\lambda, \mathbb{W}) \;&:=\; \limsup_{\epsilon\to 0}\Bigg\{\frac{||\Delta\lambda||_2}{\epsilon} : \left((A_i + \Delta A_i) + \sum_{j=1}^{m}(\lambda_j + \Delta\lambda_j)(B_{ij} + \Delta B_{ij})\right)(x_i + \Delta x_i) = 0, \\ &\qquad ||\Delta A_i||_2 \le \epsilon ||E_i||_2 \text{ and } ||\Delta B_{ij}||_2 \le \epsilon ||F_{ij}||_2 \text{ for } i, j = 1, \cdots, m\Bigg\}. \end{aligned}$$

It is shown in [31] that $\kappa(\lambda, \mathbb{W}) = ||B_0^{-1}||_\theta$, where $\theta_i := ||E_i||_2 + \sum_{j=1}^m |\lambda_j|\; ||F_{ij}||_2$ for $i = 1, \cdots, m$ and $B_0 := [y_i^* B_{ij} x_i]_{m\times m}$.

On the other hand, considering (λ, x) with $\lambda \in \mathbb{C}^m$ and $x = x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ as an approximate eigenpair of $\mathbb{W}$, the backward error of (λ, x) is defined as

$$\eta(x,\lambda) := \min\{\epsilon : \mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0, ||\Delta A_i||_2 \leq \epsilon ||E_i||_2 \text{ and } ||\Delta B_{ij}||_2 \leq \epsilon ||F_{ij}||_2 \text{ for } i,j = 1,\cdots,m\}.$$

It is shown in [31] that $\eta(x,\lambda) = \max_{i=1,\cdots,m} \frac{||r_i||_2}{\widetilde{\theta}_i}$, where $r_i := W_i(\lambda)x_i$ and $\widetilde{\theta}_i := ||E_i||_2 + \sum_{j=1}^{m} |\lambda_j|\ ||F_{ij}||_2$ for $i = 1,\cdots,m$. Minimizing $\eta(\lambda, x)$ over $\|x_i\|_2 = 1, i = 1,\ldots,m$, yields backward error of λ.

Linearization is a standard method for solving polynomial eigenvalue problems. A process that transforms a polynomial eigenvalue problem to an equivalent generalized eigenvalue problem of larger size is called a linearization. Linearizations for quadratic two-parameter eigenvalue problems are considered in [34, 44] and a linearization for a two-parameter matrix polynomial of degree k is considered in [44]. A vector space of potential linearizations of a quadratic two-parameter eigenvalue problem is developed and analyzed in [1].

Our main aim in this thesis is two-fold. First, to develop a general framework for the sensitivity and backward perturbation analysis of linear and polynomial MEPs. Second, to develop and analyze two vector spaces of potential linearizations of two parameter polynomial eigenvalue problems. The thesis is organized as follows.

The remainder of Chapter 1 is devoted to basic results for ready reference. In Chapter 2, we undertake the sensitivity and backward perturbation analysis of nonhomogeneous MEPs. We develop a general framework for the spectral analysis of nonhomogeneous MEPs. For a general norm on the space of MEPs, we define the condition number $\text{cond}(\lambda, \mathbb{W})$ of a simple eigenvalue λ of an MEP $\mathbb{W}$ and derive three equivalent representations of $\text{cond}(\lambda, \mathbb{W})$. We also analyze holomorphic perturbation of a simple eigenvalue of $\mathbb{W}$. Further, we define the backward error of an approximate eigenvalue λ of $\mathbb{W}$ by

$$\eta(\lambda, \mathbb{W}) := \inf\{\|\Delta\mathbb{W}\| : \lambda \in \sigma(\mathbb{W} + \Delta\mathbb{W})\},$$

where $\sigma(\mathbb{W})$ denotes the spectrum of $\mathbb{W}$. We determine $\eta(\lambda, \mathbb{W})$ and construct an optimal perturbation $\Delta\mathbb{W}$ such that $\lambda \in \sigma(\mathbb{W} + \Delta\mathbb{W})$ and $\|\Delta\mathbb{W}\| = \eta(\lambda, \mathbb{W})$. Furthermore, we define the backward error of an approximate eigenpair $(\lambda, x := x_1 \otimes \cdots \otimes x_m) \in$

$\mathbb{C}^m \times (\mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m})$ by

$$\eta(\lambda, x, \mathbb{W}) := \inf\{|\!|\!|\Delta\mathbb{W}|\!|\!| : \mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0\}.$$

We determine $\eta(\lambda, x, \mathbb{W})$ and construct an optimal perturbation $\Delta\mathbb{W}$ such that $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ and $|\!|\!|\Delta\mathbb{W}|\!|\!| = \eta(\lambda, x, \mathbb{W})$.

Chapter 3 is devoted to the sensitivity and backward perturbation analysis of two-parameter polynomial eigenvalue problems (PEPs). We extend the results in Chapter 2 to the case of two parameter PEPs.

In Chapter 4, we develop a general framework for spectral analysis of linear homogeneous MEPs. We introduce condition number $\mathrm{cond}(\lambda, \mathbb{W})$ of a simple eigenvalue λ of a homogeneous MEP $\mathbb{W}$ and derive three equivalent representations of $\mathrm{cond}(\lambda, \mathbb{W})$. We also determine backward errors of approximate eigenvalues and approximate eigenpairs of $\mathbb{W}$ and construct optimal perturbations.

In Chapter 5, we analyze linearizations of two parameter polynomial eigenvalue problems of degree k. We consider a two-parameter matrix polynomial of the form

$$P(\lambda, \mu) := \sum_{i=0}^{k} \sum_{j=0}^{k-i} \lambda^i \mu^j A_{ij}, \tag{1.5}$$

where $A_{ij} \in \mathbb{C}^{n\times n}$ for all $i = 0, \cdots, k$ and $j = 0, \cdots, k-i$. Then an $ln \times ln$ two-parameter matrix pencil $L(\lambda, \mu) := \mathbb{A} + \lambda\mathbb{X} + \mu\mathbb{Y}$ is said to be a linearization of $P(\lambda, \mu)$ if there exist unimodular matrix polynomials $E(\lambda, \mu)$ and $F(\lambda, \mu)$ (i.e., determinants of $E(\lambda, \mu)$ and $F(\lambda, \mu)$ are non-zero constants) such that

$$F(\lambda, \mu)L(\lambda, \mu)E(\lambda, \mu) = \begin{bmatrix} P(\lambda, \mu) & 0 \\ 0 & I_{(l-1)n} \end{bmatrix}.$$

We define and analyze two vector spaces, namely, the right ansatz space and the left ansatz space of potential linearizations of $P(\lambda, \mu)$ and analyze conditions under which a pencil in the ansatz spaces is a linearization of $P(\lambda, \mu)$.

Finally, in Chapter 6, we consider structured linear MEPs and analyze structured backward errors and structured backward perturbations. We denote the space of structured MEPs by $\mathbb{S}$ and define structured backward error of an approximate eigenpair

$(\lambda, x := x_1 \otimes \cdots \otimes x_m)$ by

$$\eta^{\mathbb{S}}(\lambda, x, \mathbb{W}) := \inf\{\|\Delta\mathbb{W}\| : \Delta\mathbb{W} \in \mathbb{S} \text{ and } \mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0\}.$$

We determine $\eta^{\mathbb{S}}(\lambda, x, \mathbb{W})$ and construct an optimal structured perturbation $\Delta\mathbb{W}$ such that $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ and $\|\Delta\mathbb{W}\| = \eta^{\mathbb{S}}(\lambda, x, \mathbb{W})$.

1.1 Preliminaries

Throughout this thesis, we use $\mathbb{C}^n$ to denote the vector space of column vectors $(x_1, \cdots, x_n)^T$, where $x_i \in \mathbb{C}$ and $\mathbb{C}^{m\times n}$ to denote the vector space of m-by-n matrices with entries from $\mathbb{C}$. We denote the $n \times n$ identity matrix by I_n and the $n \times n$ zero matrix by 0_n.

Kernel and range of a matrix. Let $A \in \mathbb{C}^{m\times n}$. Then the kernel of A is defined by $\ker(A) := \{x \in \mathbb{C}^n : Ax = 0\}$. We denote the dimension of the kernel of A by $\dim\ker(A)$. The range of A is defined by $\text{range}(A) := \{Ax : x \in \mathbb{C}^n\}$. The dimension of $\text{range}(A)$ is called the rank of A and is denoted by $\text{rank}(A)$.

Unitary matrix. Let $A = [a_{ij}] \in \mathbb{C}^{m\times n}$. Then $A^T = [a_{ji}] \in \mathbb{C}^{n\times m}$ denotes the transpose of A and $A^* = [\overline{a_{ji}}] \in \mathbb{C}^{n\times m}$ denotes the conjugate transpose of A. A matrix $A \in \mathbb{C}^{n\times n}$ is said to be unitary if $AA^* = A^*A = I_n$.

Spectrum of a matrix. Let $A \in \mathbb{C}^{n\times n}$. Then $\lambda \in \mathbb{C}$ is said to be an eigenvalue of A if $\det(A - \lambda I_n) = 0$. The set of all eigenvalues of A is called the spectrum of A and is denoted by $\sigma(A)$.

Let $\lambda \in \sigma(A)$. Then $x \in \mathbb{C}^n$ and $y \in \mathbb{C}^n$ are said to be right and left eigenvectors of A corresponding to λ if $Ax = 0$ and $y^*A = 0$, respectively. The geometric multiplicity of λ is given by $\dim\ker(A - \lambda I_n)$ and the algebraic multiplicity of λ is its multiplicity as a root of the characteristic polynomial $p(z) = \det(A - zI_n)$. Note that the algebraic multiplicity of an eigenvalue is always greater than or equal to the geometric multiplicity of that eigenvalue.

Singular value decomposition (SVD). Let $A \in \mathbb{C}^{m\times n}$ be such that $\text{rank}(A) = r$. Then the SVD of A is given by $A = U\Sigma V^*$, where $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ are unitary and $\Sigma = \text{diag}(\sigma_1, \cdots, \sigma_r, 0, \cdots, 0) \in \mathbb{C}^{m\times n}$ is a diagonal matrix with $\sigma_1 \geq \cdots \geq \sigma_r > 0$, where $\sigma_1, \sigma_2, \cdots, \sigma_r$ are the nonzero singular values of A. We denote the smallest nonzero singular value of a matrix A by $\sigma_{\min}(A)$.

Adjugate of a matrix. Let $A \in \mathbb{C}^{n\times n}$. Then the $(i,j)^{\text{th}}$ minor of A is the determinant of the $(n-1)\times(n-1)$ matrix that results from deleting i^{th} row and j^{th} column of A. The $(i,j)^{\text{th}}$ minor of A is denoted by M_{ij}. Then the matrix $[(-1)^{i+j}M_{ij}]$ is called the cofactor matrix of A and is denoted by C. The adjugate matrix of A is denoted by $\operatorname{adj}(A)$ and is defined by the transpose of the cofactor matrix of A, i.e., $\operatorname{adj}(A) = C^T$.

It is well known that

$$A \cdot \operatorname{adj}(A) = \operatorname{adj}(A) \cdot A = \det(A) \cdot I_n.$$

Proposition 1.1.1. *[10] Let $A \in \mathbb{C}^{n\times n}$ and $\operatorname{rank}(A) = n-1$. Then $\operatorname{adj}(A) = vu^*$ for some nonzero vectors u and v such that $Av = 0$ and $u^*A = 0$.*

Trace of a matrix. The trace of an square matrix A is defined by the sum of the diagonal entries of A and is denoted by $\operatorname{tr}(A)$. Suppose that $A = [a_{ij}] \in \mathbb{C}^{n\times n}$. Then $\operatorname{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii}$.

Jacobi's formula. In subsequent development, we use the Jacobi's formula which expresses the derivative of the determinant of a matrix.

Proposition 1.1.2. *[35] The determinant map $\det : \mathbb{C}^{n\times n} \to \mathbb{C}, A \mapsto \det(A)$ is differentiable and*

$$d\det(A) = \operatorname{tr}(\operatorname{adj}(A)\cdot dA), \tag{1.6}$$

where dA denotes the differential of A.

Kronecker product of matrices. The Kronecker product of matrices is defined as follows.

Definition 1.1.1. *[49] Let $A = [a_{ij}] \in \mathbb{C}^{m\times n}, B \in \mathbb{C}^{p\times q}$. Then the Kronecker product (tensor product) of A and B is defined by*

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

Some properties of the kronecker product are as follows [49]:

(i) Let $A \in \mathbb{C}^{m\times n}, B \in \mathbb{C}^{r\times s}, C \in \mathbb{C}^{n\times p}$ and $D \in \mathbb{C}^{s\times t}$. Then

$$(A \otimes B)(C \otimes D) = AC \otimes BD \in \mathbb{C}^{mr\times pt}.$$

(ii) For all A and B, $(A \otimes B)^T = A^T \otimes B^T$ and $(A \otimes B)^* = A^* \otimes B^*$.

(iii) If $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{p\times p}$ are nonsingular then $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.

Norm. Norm plays an important role in perturbation analysis. We briefly consider vector and matrix norms to be used to define general norms on the space of MEPs.

Definition 1.1.2. (Vector norm) *A function* $||\cdot|| : \mathbb{C}^n \to \mathbb{R}$ *is said to be a norm (vector norm) on* $\mathbb{C}^n$ *if it satisfies the following conditions:*

(i) $||x|| = 0 \Leftrightarrow x = 0.$

(ii) $||\alpha x|| = |\alpha|\ ||x||$ *for* $\alpha \in \mathbb{C}$ *and* $x \in \mathbb{C}^n$.

(iii) $||x + y|| \leq ||x|| + ||y||$ *for* $x, y \in \mathbb{C}^n$.

The p-norm $||\cdot||_p$ on $\mathbb{C}^n$ is given by

$$||x||_p := \begin{cases} (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}} & 1 \leq p < \infty \\ \max\limits_{i=1,\cdots,n} |x_i| & p = \infty \end{cases}$$

for $x = (x_1, \cdots, x_n)^T \in \mathbb{C}^n$.

Definition 1.1.3. (Monotone norm) *A norm* $||\cdot||$ *on* $\mathbb{C}^n$ *is said to be a monotone norm if* $|x| \leq |y|$ *implies that* $||x|| \leq ||y||$, *where* $|x| \leq |y|$ *means* $|x_j| \leq |y_j|$ *for* $j = 1, \cdots, n$.

It is easy to see that the p-norm $||\cdot||_p$ is a monotone norm on $\mathbb{C}^n$.

Next we consider matrix norm on $\mathbb{C}^{n\times n}$. Let $||\cdot||$ be vector norm on $\mathbb{C}^n$. Define $||\cdot|| : \mathbb{C}^{n\times n} \to \mathbb{R}$ by $||A|| := \sup\,\{||Ax|| : x \in \mathbb{C}^n \text{ and } ||x|| = 1\}$. Then $||\cdot||$ is a norm on $\mathbb{C}^{n\times n}$ and we refer to it as the induced operator norm or the subordinate matrix norm. The subordinate matrix norm induced by the 2-norm $||\cdot||_2$ on $\mathbb{C}^n$ is denoted by $||A||_2$. We refer to $||\cdot||_2$ as the the spectral norm on $\mathbb{C}^{n\times n}$. The Frobenius norm on $\mathbb{C}^{n\times n}$ is defined as $||A||_F := \mathrm{tr}\,(A^*A)^{\frac{1}{2}}$. The spectral and Frobenius norms satisfy the followings:

(i) $||Ux|| = ||x||$ for an unitary matrix U.

(ii) $||UAV^*|| = ||A||$ for unitary matrices U and V.

(iii) $||A \otimes B|| = ||A||\ ||B||$ for $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{m\times m}$.

The following result shows that the spectral norm of $\text{adj}(A)$ is the product of the singular values of A.

Lemma 1.1.1. *[48] Let A be a $n \times n$ matrix. Then $||\text{adj}(A)||_2 = \prod_{i=1}^{n-1} \sigma_i(A)$, where $\sigma_1(A) \geq \cdots \geq \sigma_n(A)$ are the singular values of A.*

Inner Product. Let V be a finite dimensional vector space over a field $\mathbf{F}$. Then $\langle \cdot, \cdot \rangle : V \times V \to \mathbf{F}$ is said to be an inner product if the following conditions hold.

(i) $\langle v, v\rangle \geq 0$ for all $v \in V$.

(ii) $\langle v, v\rangle = 0$ if and only if $v = 0$.

(iii) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

(iv) $\langle av, w\rangle = a\langle v, w\rangle$ for all $v, w \in V$ and $a \in \mathbf{F}$.

(v) $\langle v, w\rangle = \overline{\langle w, v\rangle}$ for all $v, w \in V$.

Note that $\langle x, y\rangle = y^*x$ for all $x, y \in \mathbb{C}^n$ defines an inner product on $\mathbb{C}^n$. It is also easy to see that an inner product on $\mathbb{C}^{n\times n}$ is given by $\langle A, B\rangle = \text{tr}(B^*A)$ for all $A, B \in \mathbb{C}^{n\times n}$.

Dual Norm. Let H be a finite dimensional inner product space with an inner product $\langle \cdot, \cdot \rangle$. Let $|| \cdot ||$ be a norm on H. Then

$$||h||_* := \sup\{|\langle h, g\rangle : ||g|| = 1\}$$

defines a norm on H and the norm $|| \cdot ||_*$ is called the dual norm of $|| \cdot ||$ relative to the inner product $\langle \cdot, \cdot \rangle$. Note that $|\langle x, y\rangle| \leq ||x||\ ||y||_*$ for all $x, y \in H$. It is easy to see that the dual of the Hölder's p-norm $|| \cdot ||_p$ is the Hölder's q-norm $|| \cdot ||_q$, where $p^{-1} + q^{-1} = 1$.

Subdifferential. The subdifferential of the map $y \mapsto ||y||$ at x is given by

$$\partial||x|| := \{y \in H : \langle x, y\rangle = ||x|| \text{ and } ||y||_* = 1\}.$$

In case of the p-norm $|| \cdot ||_p$ on $\mathbb{C}^n$, we have the following result.

Proposition 1.1.3. *[8] Let $x = (x_1, \cdots, x_n)^T \in \mathbb{C}^n$ be nonzero. Consider the p-norm $||\cdot||_p$ on $\mathbb{C}^n$ for $1 \leq p \leq \infty$. For $z \in \mathbb{C}$, define $\text{sign}(z) := \frac{\bar{z}}{|z|}$ if $z \neq 0$ and $\text{sign}(0) := 0$. Then we have the followings.*

(a) For $1 < p < \infty$, we have

$$\partial||x||_p = \{y = (y_1, \cdots, y_n)^T \in \mathbb{C}^n : y_i := \text{sign}(x_i)|x_i|^{p-1}||x||_p^{1-p}\}.$$

(b) We have

$$\partial||x||_1 = \{y \in \mathbb{C}^n : y_i := \text{sign}(x_i) \text{ if } x_i \neq 0 \text{ else } |y_i| \leq 1 \text{ is arbitrary}\}.$$

(c) If $||x||_\infty = |x_i|$ for some $i \in \{1, \cdots, n\}$ and $|x_j| < |x_i|$ for all $j \neq i$ then

$$\partial||x||_\infty = \{y \in \mathbb{C}^n : y_i = \text{sign}(x_i) \text{ and } y_j = 0, \text{ for } j \neq i\}.$$

If $||x||_\infty = |x_{i_1}| = \cdots = |x_{i_m}|$ for some $i_1, \cdots, i_m \in \{1, \cdots, n\}$ and $|x_j| < |x_{i_k}|$ for $j \notin \{i_1, \cdots, i_m\}$, then

$$\partial||x||_\infty = \{y \in \mathbb{C}^n : y_r = \frac{\text{sign}(x_r)}{m} \text{ for } r = i_1, \cdots, i_m \text{ and } y_r = 0, \text{ otherwise}\}.$$

1.1.1 Vector space of linearizations of matrix polynomial

Now we briefly review the vector spaces of linearizations of a matrix polynomial of degree k.

Definition 1.1.4. *[27] Let $P(\lambda)$ be a matrix polynomial of degree k. Then a linear matrix polynomial $L(\lambda) = \lambda X + Y$ with $X, Y \in \mathbb{C}^{kn \times kn}$ is said to be a linearization of P if there exist $kn \times kn$ unimodular matrix polynomials $E(\lambda)$ and $F(\lambda)$ (i.e., $\det(E(\lambda))$ and $\det(F(\lambda))$ are nonzero constants, independent of λ) such that*

$$E(\lambda)L(\lambda)F(\lambda) = \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{(k-1)n} \end{bmatrix}.$$

Let $P(\lambda) := \sum_{j=0}^{k} \lambda^j A_j$. Then the matrix pencils

$$C_1(\lambda) := \lambda \left[\begin{array}{c|ccc} A_k & & & \\ \hline & I_n & & \\ & & \ddots & \\ & & & I_n \end{array}\right] + \left[\begin{array}{ccc|c} A_{k-1} & A_{k-2} & \cdots & A_0 \\ \hline -I_n & & & \\ & \ddots & & \\ & & -I_n & \end{array}\right]$$

and

$$C_2(\lambda) := \lambda \left[\begin{array}{c|ccc} A_k & & & \\ \hline & I_n & & \\ & & \ddots & \\ & & & I_n \end{array}\right] + \left[\begin{array}{c|ccc} A_{k-1} & -I_n & & \\ A_{k-2} & & \ddots & \\ \vdots & & & -I_n \\ \hline A_0 & & & \end{array}\right],$$

are known as the first and second companion form of $P(\lambda)$, respectively. Generalizing the companion forms, the following two classes of vector spaces of linearizations were developed in [42]:

$$\begin{aligned} \mathbb{L}_1(P) &:= \{L(\lambda) : L(\lambda) \cdot (\Lambda \otimes I_n) = v \otimes P(\lambda), v \in \mathbb{C}^k\}, \\ \mathbb{L}_2(P) &:= \{L(\lambda) : (\Lambda^T \otimes I_n) \cdot L(\lambda) = w^T \otimes P(\lambda), w \in \mathbb{C}^k\}, \end{aligned}$$

where $L(\lambda) \in \mathbb{C}^{kn \times kn}$ linear matrix polynomial and $\Lambda := (\lambda^{k-1}, \lambda^{k-2}, \cdots, \lambda, 1)^T$. $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ are called the right and left ansatz space, respectively. The vector v is called the right ansatz vector for $L(\lambda) \in \mathbb{L}_1(P)$ and the vector w is called the left ansatz vector for $L(\lambda) \in \mathbb{L}_2(P)$.

Theorem 1.1.1. *[42]*(Eigenvector recovery property) *Let $P(\lambda)$ be a regular matrix polynomial of degree k and $L \in \mathbb{L}_1(P)$ with non-zero right ansatz vector v. Then $x \in \mathbb{C}^n$ is a right eigenvector of P corresponding to the finite eigenvalue $\lambda \in \mathbb{C}$ if and only if $\Lambda \otimes x$ is an eigenvector of L corresponding to the eigenvalue λ. In fact, every right eigenvector of L with finite eigenvalue λ is of the form $\Lambda \otimes x$ for some eigenvector x of P corresponding to λ.*

Similarly, if $L \in \mathbb{L}_2(P)$ with nonzero left ansatz vector w then $y \in \mathbb{C}^n$ is a left eigenvector of P corresponding to the finite eigenvalue $\lambda \in \mathbb{C}$ if and only if $\overline{\Lambda} \otimes y$ is a left eigenvector of L corresponding to the eigenvalue λ. In fact, every left eigenvector *of L with finite eigenvalue λ is of the form $\overline{\Lambda} \otimes y$ for some left eigenvector y of P corresponding to λ.*

The right and left ansatz space provide a wide class of matrix pencils corresponding to a given matrix polynomial. It is shown in [42] that all the matrix pencils in the right and left ansatz space are not linearizations of the given matrix polynomial. The following result gives the linearization condition for $L \in \mathbb{L}_1(P)$ corresponding to the nonzero right ansatz vector αe_1, where $0 \neq \alpha \in \mathbb{C}$.

Theorem 1.1.2. *[42] Let $P(\lambda) = \sum_{j=0}^{k} \lambda^j A_j$ be a matrix polynomial of degree k. Suppose that $L \in \mathbb{L}_1(P)$ with nonzero right ansatz vector $v = \alpha e_1$. Then*

$$L(\lambda) = \lambda \left[\begin{array}{c|c} \alpha A_k & X \\ \hline 0 & -Z \end{array}\right] + \left[\begin{array}{c|c} Y & \alpha A_0 \\ \hline Z & 0 \end{array}\right],$$

where $X, Y \in \mathbb{C}^{n \times (k-1)n}$ and $Z \in \mathbb{C}^{(k-1)n \times (k-1)n}$. If Z is nonsingular then $L(\lambda)$ is a linearization of $P(\lambda)$.

Procedure to determine the linearization condition for a pencil in $\mathbb{L}_1(P)$[42]

(1) Let $P(\lambda)$ be a matrix polynomial of degree k and $L(\lambda) = \lambda X + Y \in \mathbb{L}_1(P)$ corresponding to the nonzero right ansatz vector $v \in \mathbb{C}^k$.

(2) Select a nonsingular matrix M such that $Mv = \alpha e_1$, where $0 \neq \alpha \in \mathbb{C}$.

(3) Consider $\widetilde{L}(\lambda) := (M \otimes I_n)L(\lambda)$. Then

$$\widetilde{L}(\lambda) = \lambda \left[\begin{array}{c|c} \widetilde{X}_{11} & \widetilde{X}_{12} \\ \hline 0 & -Z \end{array}\right] + \left[\begin{array}{c|c} \widetilde{Y}_{11} & \widetilde{Y}_{12} \\ \hline 0 & 0 \end{array}\right].$$

(4) Extract $\det(Z) \neq 0$, the linearization condition for $L(\lambda)$.

1.1.2 Nonsingular MEP

Now we review some results on MEP.

Let $\mathbb{W} : \mathbb{C}^{m+1} \longrightarrow \prod_{j=1}^{m} \mathbb{C}^{n_j \times n_j}$ be given by

$$\mathbb{W}(z) := (W_1(z), \ldots, W_m(z)), \text{ where } W_i(z) := z_0 A_i + \sum_{j=1}^{m} z_j B_{ij} \tag{1.7}$$

and $A_i, B_{ij} \in \mathbb{C}^{n_i \times n_i}$ for $i, j = 1, \cdots, m$. We refer to $\mathbb{W}$ as a homogeneous MEP. Note that, $\mathbb{W}$ is *regular* if $\prod_{j=1}^{m} \det(W_j(z)) \neq 0$ for some $z \in \mathbb{C}^m$. Suppose that $\mathbb{W}$ is regular. For $x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ and $z \in \mathbb{C}^m$, we define

$$\mathbb{W}(z)x := (W_1(z)x_1, \ldots, W_m(z)x_m) \;\text{ and }\; x^*\mathbb{W}(z) := (x_1^*W_1(z), \ldots, x_m^*W_m(z)).$$

Then $\lambda \in \mathbb{C}^{m+1} \setminus \{0\}$ is called an *eigenvalue* of $\mathbb{W}$ if $\det(W_i(\lambda)) = 0$ for all $i = 1, \cdots, m$. And a nonzero $x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1 \otimes \cdots \otimes \mathbb{C}^{n_m}}$ is called a *right eigenvector* of $\mathbb{W}$ corresponding to λ if $\mathbb{W}(\lambda)x = 0$. Similarly, a nonzero $y := y_1 \otimes \cdots \otimes y_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ is called a *left eigenvector* of $\mathbb{W}$ corresponding to λ if $y^*\mathbb{W}(\lambda) = 0$.

Let $B_{ij}^{\dagger} : \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m} \to \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ be the map induced by B_{ij} and given by

$$B_{ij}^{\dagger}(x_1 \otimes \cdots \otimes x_m) = x_1 \otimes \cdots \otimes B_{ij}x_i \otimes \cdots \otimes x_m$$

for $i, j = 1, \cdots, m$. Similarly define the maps $A_i^{\dagger} : \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m} \to \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ given by

$$A_i^{\dagger}(x_1 \otimes \cdots \otimes x_m) = x_1 \otimes \cdots \otimes A_ix_i \otimes \cdots \otimes x_m, \tag{1.8}$$

for $i = 1, \cdots, m$. Define the operator determinants $\Delta_i : \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m} \to \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ by

$$\Delta_0 := \det\left(\begin{bmatrix} B_{11}^{\dagger} & B_{12}^{\dagger} & \cdots & B_{1m}^{\dagger} \\ B_{21}^{\dagger} & B_{22}^{\dagger} & \cdots & B_{2m}^{\dagger} \\ \vdots & \vdots & & \vdots \\ B_{m1}^{\dagger} & B_{m2}^{\dagger} & \cdots & B_{mm}^{\dagger} \end{bmatrix}\right) \tag{1.9}$$

and

$$\Delta_i = \det\left(\begin{bmatrix} B_{11}^{\dagger} & \cdots & B_{1,i-1}^{\dagger} & A_1^{\dagger} & B_{1i}^{\dagger} & \cdots & B_{1m}^{\dagger} \\ B_{21}^{\dagger} & \cdots & B_{2,i-1}^{\dagger} & A_2^{\dagger} & B_{2i}^{\dagger} & \cdots & B_{2m}^{\dagger} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ B_{m1}^{\dagger} & \cdots & B_{m,i-1}^{\dagger} & A_m^{\dagger} & B_{mi}^{\dagger} & \cdots & B_{mm}^{\dagger} \end{bmatrix}\right) \tag{1.10}$$

for $i = 1, \cdots, m$. For a decomposable tensor $x = x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$

$$\begin{aligned} \Delta_0 x &= \sum_{\sigma \in \mathbf{S}_m} (-1)^{\mathrm{sgn}(\sigma)} B_{1\sigma(1)} x_1 \otimes \cdots \otimes B_{m\sigma(m)} x_m \\ &= \det \left(\begin{bmatrix} B_{11}x_1 & B_{12}x_1 & \cdots & B_{1m}x_1 \\ B_{21}x_2 & B_{22}x_2 & \cdots & B_{2m}x_2 \\ \vdots & \vdots & & \vdots \\ B_{m1}x_m & B_{m2}x_m & \cdots & B_{mm}x_m \end{bmatrix} \right), \end{aligned} \tag{1.11}$$

where $\mathbf{S}_m$ is the set of all permutations of the set $\{1, 2, \cdots, m\}$ and $\mathrm{sgn}(\sigma)$ is the sign of a permutation $\sigma \in \Pi_m$. These operator determinants satisfy the following.

Lemma 1.1.2. *[12] The operators Δ_0 and Δ_i given in (1.9) and (1.10) have the following properties.*

(i) If two columns are interchanged, the sign of the determinant is reversed.

(ii) If two columns are identical, the determinant vanishes.

(iii) The value of the determinant is unchanged if a scalar multiple of one column is added to another column.

Atkinson [12] used the operator determinants to define nonsingular MEP.

Definition 1.1.5. *[12] Let $\mathbb{W}$ be as given in (1.7). Then $\mathbb{W}$ is said to be nonsingular if there exists scalars $\mu_0, \cdots, \mu_m$ such that*

$$\Delta := \sum_{s=0}^{m} \mu_s \Delta_s = \det \left(\begin{bmatrix} \mu_0 & \mu_1 & \cdots & \mu_m \\ A_1^\dagger & B_{11}^\dagger & \cdots & B_{1m}^\dagger \\ \vdots & \vdots & & \vdots \\ A_m^\dagger & B_{m1}^\dagger & \cdots & B_{mm}^\dagger \end{bmatrix} \right)$$

defines an isomorphism.

Let $\mathbb{W}$ be nonsingular. Define $S_i := \Delta^{-1}\Delta_i$ for $i = 1, \cdots, m$. It is shown in [12] that these operators are commutative, i.e., $S_i S_j = S_j S_i$. The commuting tuple of operators $\mathbf{S} = (S_1, \cdots, S_m)$ is said to be the associated system of the homogeneous MEP $\mathbb{W}$. Next we recall the definition of the spectrum of a commuting tuple of operators.

Definition 1.1.6. *[37] Let* $\mathbf{T} = (\mathbf{T}_1, \cdots, \mathbf{T}_m)$ *be a commuting tuple of operators. Then* $\lambda = (\lambda_1, \cdots, \lambda_m)^T \in \mathbb{C}^m$ *is said to be an eigenvalue of* $\mathbf{T}$ *if* $\bigcap \ker(\lambda_i I - \mathbf{T}_i) \neq \{0\}$*. The set of all eigenvalues of* $\mathbf{T}$ *is said to be the spectrum of* $\mathbf{T}$ *and is denoted by* $\sigma(\mathbf{T})$*.*

Atkinson[12] proved that the spectrum of a nonsingular homogeneous MEP is same as the spectrum of the corresponding associated system of the MEP. In fact, a nonsingular homogeneous MEP is equivalent to the following system of generalized eigenvalue problem

$$\Delta_i x = \lambda_i \Delta x \quad i = 0, 1, \cdots, m.$$

Theorem 1.1.3. *[12] Let* $\mathbb{W}$ *be as given in (1.7). Let* $\lambda = (\lambda_0, \cdots, \lambda_m) \in \sigma(\mathbb{W})$ *and* $\mathbf{S} = (S_1, \cdots, S_m)$ *be the associated system of* $\mathbb{W}$*. Then the eigensubspace associated to* λ *is same as* $\bigcap_{i=0}^{m} \ker(S_i - \lambda_i I)$*. Furthermore,* $\sigma(W) = \sigma(T)$ *and*

$$\ker(\mathbf{S} - \lambda I) = \ker(W_1(\lambda)) \otimes \cdots \otimes \ker(W_m(\lambda)). \tag{1.12}$$

Consider a nonhomogeneous MEP $\mathbb{W}$ as given in (1.1). Then $\mathbb{W}$ is said to be nonsingular if the corresponding operator determinant Δ_0 given in (1.9) is invertible. A nonsingular nonhomogeneous MEP is equivalent to the system of generalized eigenvalue problem

$$\Delta_i x = \lambda_i \Delta_0 x \quad i = 1, \cdots, m.$$

In this case the operators $T_i := \Delta_0^{-1} \Delta_i$ for $i = 1, \cdots, m$, are commutative, i.e., $T_i T_j = T_j T_i$. The commuting tuple $\mathbf{T} = (T_1, \cdots, T_m)$ is said to be the associated system of the nonsingular nonhoogeneous MEP $\mathbb{W}$.

Definition 1.1.7. *[31] Let* $\mathbb{W}$ *as given in (1.1) be nonsingular and* $\mathbf{T} = (T_1, \cdots, T_m)$ *be the corresponding associated system. Let* $\lambda = (\lambda_1, \cdots, \lambda_m) \in \sigma(\mathbb{W})$*. Then the algebraic multiplicity of* λ *is given by*

$$m(\lambda) = \dim \left(\bigcap_{\substack{j_1 + \cdots + j_m = N \\ j_1, \cdots, j_m \geq 0}} \ker[(T_1 - \lambda_1 I)^{j_1} \cdots (T_m - \lambda_m I)^{j_m}] \right), \tag{1.13}$$

where $N = \prod_{i=1}^{m} n_i$*. The geometric multiplicity of* λ *is given by*

$$g(\lambda) = \dim \left(\bigcap_{i=1}^{m} \ker(T_i - \lambda_i I) \right). \tag{1.14}$$

An eigenvalue $\lambda \in \sigma(\mathbb{W})$ *is said to be algebraically simple if* $m(\lambda) = 1$ *and is geometrically simple if* $g(\lambda) = 1$.

Let $\lambda \in \sigma(\mathbb{W})$ and $y = y_1 \otimes \cdots \otimes y_m, x = x_1 \otimes \cdots \otimes x_m$ be left and right eigenvectors of $\mathbb{W}$ corresponding to λ, respectively. Define

$$B_0 = \begin{bmatrix} y_1^* B_{11} x_1 & \cdots & y_1^* B_{1m} x_1 \\ & \vdots & \\ y_m^* B_{m1} x_m & \cdots & y_m^* B_{mm} x_m \end{bmatrix}, \tag{1.15}$$

Then the following result holds.

Corollary 1.1.4. *[37] Let* $\mathbb{W}$ *as given in (1.1) be nonsingular and* $\mathbf{T} = (T_1, \cdots, T_m)$ *be the associated system of* $\mathbb{W}$. *Let* $\lambda \in \sigma(\mathbb{W})$ *be geometrically simple and* $y = y_1 \otimes \cdots \otimes y_m, x = x_1 \otimes \cdots \otimes x_m$ *be left and right eigenvectors of* $\mathbb{W}$ *corresponding to* λ, *respectively. Then*

$$\dim \ker(\lambda I - T)^2 = 1 + \dim \ker B_0, \tag{1.16}$$

where B_0 *is given in (1.15) and*

$$\ker(\lambda I - T)^2 = \bigcap_{i=1}^{m} \bigcap_{j=1}^{m} \ker[(\lambda_i I - T_i)(\lambda_j I - T_j)].$$

2

Sensitivity and Backward Perturbation Analysis of Nonhomogeneous MEP

2.1 Introduction

Nonhomogeneous MEPs arise in various applications, see [23, 32, 41, 45]. Various numerical methods have been proposed in the literature for solving a linear nonsingular nonhomogeneous MEP, for details, see [32, 33]. In this chapter we undertake sensitivity and backward perturbation analysis for a regular linear nonhomogeneous MEP.

First we develop a framework for sensitivity and backward perturbation analysis of a regular nonhomogeneous MEP. For a general norm on the space of linear MEPs, we define the condition number of an eigenvalue λ of a regular nonhomogeneous MEP $\mathbb{W}$ by $\mathrm{cond}(\lambda, \mathbb{W}) := \limsup\limits_{|||\Delta\mathbb{W}|||\to 0} \frac{\mathrm{dist}(\lambda,\sigma(\mathbb{W}+\Delta\mathbb{W}))}{|||\Delta\mathbb{W}|||}$. For a simple eigenvalue $\lambda_{\mathbb{W}}$ of $\mathbb{W}$, we show that

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \; ||(1, \lambda_{\mathbb{W}}^*)||_{V,*} \max_i ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_*.$$

Further we derive $\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W})$ explicitly for various choice of norms.

Given $\lambda \in \mathbb{C}^m$ and $0 \neq x = x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$, we define the backward error of (λ, x) as an approximate eigenpair of $\mathbb{W}$ by $\eta(\lambda, x, \mathbb{W}) = \inf\{|||\Delta\mathbb{W}||| : \mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0\}$ and the backward error of λ as an approximate eigenvalue of $\mathbb{W}$ by $\eta(\lambda, \mathbb{W}) = \inf\{|||\Delta\mathbb{W}||| : \lambda \in \sigma(\mathbb{W} + \Delta\mathbb{W})\}$ and determine $\eta(\lambda, x, \mathbb{W})$ and $\eta(\lambda, \mathbb{W})$.

2.2 Nonhomogeneous MEP

Let $\mathbb{W} : \mathbb{C}^m \longrightarrow \prod_{j=1}^m \mathbb{C}^{n_j \times n_j}$ be given by

$$\mathbb{W}(z) := (W_1(z), \cdots, W_m(z)), \text{ where } W_i(z) := A_i + \sum_{j=1}^m z_j B_{ij}, \tag{2.1}$$

$z := (z_1, \cdots, z_m)^T \in \mathbb{C}^m$ and $A_i, B_{ij} \in \mathbb{C}^{n_i \times n_i}$ for all $i, j = 1, \cdots, m$. We refer to $\mathbb{W}$ as an MEP. We denote the space of all nonhomogeneous MEPs by $\mathbb{MEP}(n_1, \cdots, n_m)$. For $x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ and $z \in \mathbb{C}^m$, we define

$$\mathbb{W}(z)x := (W_1(z)x_1, \ldots, W_m(z)x_m) \text{ and } x^*\mathbb{W}(z) := (x_1^* W_1(z), \ldots, x_m^* W_m(z)).$$

Definition 2.2.1. (Regular MEP) *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be given by* $\mathbb{W}(z) = (W_1(z), \cdots, W_m(z))$. *Then* $\mathbb{W}$ *is said to be regular if* $\prod_{j=1}^m \det(W_j(z)) \neq 0$ *for some* $z \in \mathbb{C}^m$.

Throughout the chapter, we consider regular MEPs. Next we define

$$\mathbf{det} : \prod_{j=1}^m \mathbb{C}^{n_j \times n_j} \to \mathbb{C}^m, A \mapsto (\det(A_1), \cdots, \det(A_m))^T, \tag{2.2}$$

where $A = (A_1, \cdots, A_m) \in \prod_{j=1}^m \mathbb{C}^{n_j \times n_j}$ and $\det(A_i)$ denotes the determinant of A_i for $i = 1, \cdots, m$. Note that $\mathbb{W}(z) \in \prod_{j=1}^m \mathbb{C}^{n_j \times n_j}$ and thus

$$\mathbf{det}(\mathbb{W}(z)) = (\det(W_1(z)), \cdots, \det(W_m(z)))^T.$$

Now we use **det** function to define the spectrum of a regular nonhomogeneous MEP.

Definition 2.2.2. (Spectrum) *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular. Then* $\lambda \in \mathbb{C}^m$ *is said to be an eigenvalue of* $\mathbb{W}$ *if* $\mathbf{det}(\mathbb{W}(\lambda)) = 0$. *The set of all eigenvalues of* $\mathbb{W}$ *is called the spectrum of* $\mathbb{W}$ *and is denoted by* $\sigma(\mathbb{W})$. *Let* $\lambda \in \sigma(\mathbb{W})$. *Then* $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ *is called a right eigenvector of* $\mathbb{W}$ *corresponding to* λ *if* $\mathbb{W}(\lambda)x = 0$. *Similarly* $0 \neq y := y_1 \otimes \cdots \otimes y_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ *is called a left eigenvector of* $\mathbb{W}$ *corresponding to* λ *if* $y^*\mathbb{W}(\lambda) = 0$.

Let $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ be given by (2.1) and $\lambda \in \sigma(\mathbb{W})$. Let $x = x_1 \otimes \cdots \otimes x_m$ and $y = y_1 \otimes \cdots \otimes y_m$ be right and left eigenvectors of $\mathbb{W}$ corresponding to λ, respectively. Define

$$J_{\mathbb{W}}(\lambda; y, x) := \begin{bmatrix} y_1^* B_{11} x_1 & \cdots & y_1^* B_{1m} x_1 \\ & \vdots & \\ y_m^* B_{m1} x_m & \cdots & y_m^* B_{mm} x_m \end{bmatrix}. \tag{2.3}$$

Next we define geometric multiplicity of an eigenvalue of a regular nonhomogeneous MEP.

Definition 2.2.3. (Geometric multiplicity of an eigenvalue) *Let $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ given by $\mathbb{W}(z) = (W_1(z), \cdots, W_m(z))$ be regular and $\lambda \in \sigma(\mathbb{W})$. Then the geometric multiplicity of λ is defined by $g_{\mathbb{W}}(\lambda) = \prod_{i=1}^m \dim \ker(W_i(\lambda))$. If $g_{\mathbb{W}}(\lambda) = 1$, then λ is said to be geometrically simple.*

Given $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$, consider the differentiable function $\mathbb{C}^m \to \mathbb{C}^m$ given by $z \mapsto \mathbf{det}(\mathbb{W}(z))$ for $z \in \mathbb{C}^m$. To define an algebraically simple eigenvalue of $\mathbb{W}$, we use the Jacobian matrix of this function at z which we denote by $J_{\mathbb{W}}(z)$.

Proposition 2.2.1. *Let $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ given by (2.1) be regular. Then the Jacobian matrix $J_{\mathbb{W}}(z)$ of the map $\mathbb{C}^m \to \mathbb{C}^m, z \mapsto \mathbf{det}(\mathbb{W}(z))$, is given by*

$$J_{\mathbb{W}}(z) = \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(z))B_{11}) & \cdots & \mathrm{tr}(\mathrm{adj}(W_1(z))B_{1m}) \\ \vdots & & \vdots \\ \mathrm{tr}(\mathrm{adj}(W_m(z))B_{m1}) & \cdots & \mathrm{tr}(\mathrm{adj}(W_m(z))B_{mm}) \end{bmatrix}. \tag{2.4}$$

In particular, if $\lambda \in \sigma(\mathbb{W})$ is geometrically simple then there exist left and right eigenvectors $u := u_1 \otimes \cdots \otimes u_m$ and $v := v_1 \otimes \cdots \otimes v_m$ of $\mathbb{W}$ corresponding to λ, respectively, such that

$$J_{\mathbb{W}}(\lambda) = \begin{bmatrix} u_1^* B_{11} v_1 & \cdots & u_1^* B_{1m} v_1 \\ \vdots & & \vdots \\ u_m^* B_{m1} v_m & \cdots & u_m^* B_{mm} v_m \end{bmatrix} = J_{\mathbb{W}}(\lambda; u, v). \tag{2.5}$$

Proof. Note that $J_{\mathbb{W}}(z) = \left[\frac{\partial}{\partial z_j}(\det(W_i(z)))\right]_{m \times m}$. By Jacobi's formula stated in Proposition 1.1.2, we have $\frac{\partial}{\partial z_j}\left(\det(W_i(z))\right) = \mathrm{tr}\left(\mathrm{adj}(W_i(z)) \cdot \frac{\partial}{\partial z_j}(W_i(z))\right) = \mathrm{tr}\left(\mathrm{adj}(W_i(z))B_{ij}\right)$, which yeilds (2.4).

If $\lambda \in \sigma(\mathbb{W})$ is geometrically simple then it follows from Definition 2.2.3 that $\mathrm{rank}(W_i(\lambda)) = n_i - 1$ for all $i = 1, \cdots, m$. So by Proposition 1.1.1 there exist nonzero vectors $u_i, v_i \in \mathbb{C}^{n_i}$ such that $\mathrm{adj}(W_i(\lambda)) = v_i u_i^*$, where $W_i(\lambda)v_i = 0$ and $u_i^* W_i(\lambda) = 0$ for all $i = 1, \cdots, m$. Thus we have $\mathrm{tr}(\mathrm{adj}(W_i(\lambda))B_{ij}) = \mathrm{tr}(v_i u_i^* B_{ij}) = u_i^* B_{ij} v_i$, which proves (2.5). □

Now we use $J_{\mathbb{W}}(\lambda)$ to define an algebraically simple eigenvalue $\lambda \in \sigma(\mathbb{W})$.

Definition 2.2.4. (Algebraically simple eigenvalue) *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular and* $\lambda \in \sigma(\mathbb{W})$. *Then* λ *is said to be algebraically simple if* $\mathrm{rank}(J_{\mathbb{W}}(\lambda)) = m$. *If* $\mathrm{rank}(J_{\mathbb{W}}(\lambda)) < m$ *then* λ *is said to be a multiple eigenvalue of* $\mathbb{W}$.

We refer to algebraically simple eigenvalue as the simple eigenvalue. The following result shows that a simple eigenvalue is also geometrically simple.

Proposition 2.2.2. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *given by (2.1) be regular and* $\lambda \in \sigma(\mathbb{W})$ *be simple. Then* λ *is geometrically simple.*

Proof. Let $\lambda \in \sigma(\mathbb{W})$ be simple. If possible, suppose that λ is not geometrically simple, i.e., $\dim(\ker(W_i(\lambda))) > 1$ for some $i \in \{1, \cdots, m\}$. Then $\mathrm{adj}(W_i(\lambda)) = 0$. It follows from (2.4) that the i^{th} row of $J_{\mathbb{W}}(\lambda)$ is a zero row. So $\mathrm{rank}(J_{\mathbb{W}}(\lambda)) < m$ which contradicts that λ is simple. Hence $\dim(\ker(W_i(\lambda))) = 1$ for all $i = 1, \cdots, m$ and $g_{\mathbb{W}}(\lambda) = \prod_{i=1}^{m} \dim(\ker(W_i(\lambda))) = 1$, i.e., λ is geometrically simple. □

The next result characterizes multiple eigenvalues.

Theorem 2.2.1. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *given by (2.1) be regular and* $\lambda \in \sigma(\mathbb{W})$. *Then* λ *is a multiple eigenvalue if and only if there exists left and right eigenvectors* y *and* x *of* $\mathbb{W}$ *corresponding to* λ, *respectively, such that* $J_{\mathbb{W}}(\lambda; y, x)$ *is singular.*

Proof. Let $\lambda \in \sigma(\mathbb{W})$ be multiple. Then by Definition 2.2.4, $J_{\mathbb{W}}(\lambda)$ is singular.

If λ is geometrically simple, i.e., $\dim(\ker(W_i(\lambda))) = 1$ for all $i = 1, \cdots, m$, then by Proposition 2.2.1, there exist left and right eigenvectors u and v of $\mathbb{W}$ corresponding to λ, respectively, such that $J_{\mathbb{W}}(\lambda) = J_{\mathbb{W}}(\lambda; u, v)$. Thus $J_{\mathbb{W}}(\lambda; u, v)$ is singular as $J_{\mathbb{W}}(\lambda)$ is singular.

If λ is not geometrically simple, then $\dim(\ker(W_k(\lambda))) > 1$ for some $k \in \{1, \cdots, m\}$. Let $y := y_1 \otimes \cdots \otimes y_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ be a left eigenvector of $\mathbb{W}$ corresponding to λ. Consider $x_i \in \mathbb{C}^{n_i}$ such that $W_i(\lambda)x_i = 0$ for all $i = 1, \cdots, k-1, k+1, \cdots, m$. Define the linear functional $f : \ker(W_k(\lambda)) \to \mathbb{C}$ given by $f(z_k) = \det(J_{\mathbb{W}}(\lambda; y, z))$, where $z = x_1 \otimes \cdots \otimes x_{k-1} \otimes z_k \otimes x_{k+1} \otimes \cdots \otimes x_m$ is a right eigenvector of $\mathbb{W}$ corresponding to λ. Now by the Rank - Nullity Theorem, we have $\dim \ker(f) = \dim(\ker(W_k(\lambda))) - 1 > 0$. So there exists $0 \neq x_k \in \ker(W_k(\lambda))$ such that $f(x_k) = \det(J_{\mathbb{W}}(\lambda; y, x)) = 0$, where $0 \neq x = x_1 \otimes \cdots \otimes x_{k-1} \otimes x_k \otimes x_{k+1} \otimes \cdots x_m$. Note that x is a right eigenvector of $\mathbb{W}$ corresponding to λ such that $J_{\mathbb{W}}(\lambda; y, x)$ is singular.

Conversely, suppose that there are left and right eigenvectors y and x of $\mathbb{W}$ corresponding to λ, respectively, such that $J_{\mathbb{W}}(\lambda; y, x)$ is singular.

If λ is not geometrically simple, then $\dim(\ker(W_k(\lambda))) > 1$ for some $k \in \{1, \cdots, m\}$. Then $\mathrm{adj}(W_k(\lambda)) = 0$ and thus it follows from (2.4) that the k^{th} row of $J_{\mathbb{W}}(\lambda)$ is a zero row. So $J_{\mathbb{W}}(\lambda)$ is singular and λ is an multiple eigenvalue.

If λ is geometrically simple then by Proposition 2.2.1, there exist left and right eigenvectors $u = u_1 \otimes \cdots \otimes u_m$ and $v = v_1 \otimes \cdots \otimes v_m$ of $\mathbb{W}$ corresponding to λ, respectively, such that $J_{\mathbb{W}}(\lambda) = J_{\mathbb{W}}(\lambda; u, v)$. Since $\dim(\ker(W_i(\lambda))) = 1$ for all $i = 1, \cdots, m$, there exists nonzero scalars α_i, β_i such that $u_i = \alpha_i y_i$ and $v_i = \beta_i x_i$ for all $i = 1, \cdots, m$. Then $u_i^* B_{ij} v_i = (\alpha_i y_i)^* B_{ij} (\beta_i x_i) = \overline{\alpha_i} \beta_i y_i^* B_{ij} x_i$ for all $i, j = 1, \cdots, m$. So we have

$$J_{\mathbb{W}}(\lambda; u, v) = \begin{bmatrix} \overline{\alpha_1}\beta_1 y_1^* B_{11} x_1 & \cdots & \overline{\alpha_1}\beta_1 y_1^* B_{1m} x_1 \\ \vdots & & \vdots \\ \overline{\alpha_m}\beta_m y_m^* B_{m1} x_m & \cdots & \overline{\alpha_m}\beta_m y_m^* B_{mm} x_m \end{bmatrix} = \begin{bmatrix} \overline{\alpha_1}\beta_1 & & \\ & \ddots & \\ & & \overline{\alpha_m}\beta_m \end{bmatrix} J_{\mathbb{W}}(\lambda; y, x).$$

Therefore $J_{\mathbb{W}}(\lambda; y, x)$ is singular implies that $J_{\mathbb{W}}(\lambda; u, v)$ is singular and thus $J_{\mathbb{W}}(\lambda)$ is singular. Hence λ is a multiple eigenvalue. □

The next result shows that Definition 1.1.7 and Definition 2.2.4 are equivalent.

Theorem 2.2.2. *Let $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ be a regular nonsingular nonhomogeneous MEP and $\mathbf{T} = (T_1, \cdots, T_m)$ be its associated system. Let $\lambda \in \sigma(\mathbb{W})$. Then $m(\lambda) = 1$ if and only if* $\mathrm{rank}(J_{\mathbb{W}}(\lambda)) = m$.

Proof. Suppose that $m(\lambda) = 1$. Then it follows from Definition 1.1.7 that $\dim\ker(\mathbf{T} - \lambda I)^2 = 1$. By Corollary 1.1.4 we have $\dim\ker(J_{\mathbb{W}}(\lambda)) = 1 - \dim\ker(\mathbf{T} - \lambda I)^2 = 0$, i.e., $\mathrm{rank}(J_{\mathbb{W}}(\lambda)) = m$.

Conversely suppose that $\mathrm{rank}(J_{\mathbb{W}}(\lambda)) = m$. Then by Proposition 2.2.2, $\dim\ker(W_i(\lambda)) = 1$ for all $i = 1, \cdots, m$ and it follows from (1.12) that $\dim\ker(\mathbf{T}-\lambda I) = \prod_{i=1}^{m} \dim\ker(W_i(\lambda)) = 1$. Again by Corollary 1.1.4, we have $\dim\ker(\mathbf{T} - \lambda I)^2 = 1 - \dim\ker(J_{\mathbb{W}}(\lambda)) = 1 = \dim\ker(\mathbf{T} - \lambda I)$ and thus $m(\lambda) = 1$. □

2.2.1 Norms on $\mathbb{MEP}(n_1, \cdots, n_m)$

Let $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ be as given in (2.1). Consider a monotone norm $||\cdot||_V$ on $\mathbb{C}^{m+1}$ and a matrix norm $||\cdot||$ on $\mathbb{C}^{n_i \times n_i}$ for all $i = 1, \cdots, m$. Define $||W_i||_V := ||(||A_i||, ||B_{i1}||, \cdots, ||B_{im}||)||_V$ for $i = 1, \cdots, m$ and

$$|||\mathbb{W}|||_V := ||(||W_1||_V, \cdots, ||W_m||_V)||_V. \tag{2.6}$$

It is easy to see that $|||\cdot|||_V$ is a norm on $\mathbb{MEP}(n_1, \cdots, n_m)$. In particular replacing the monotone norm $||\cdot||_V$ by the usual Hölder p-norm $||\cdot||_p$ on $\mathbb{C}^{m+1}$ for $1 \le p \le \infty$, we define $|||\cdot|||_p$ on $\mathbb{MEP}(n_1, \cdots, n_m)$ by $|||\mathbb{W}|||_p = ||(||W_1||_p, \cdots, ||W_m||_p)||_p$, where $||W_i||_p = ||(||A_i||, ||B_{i1}||, \cdots, ||B_{im}||)||_p$ for $i = 1, \cdots, m$. We refer to $|||\cdot|||_p$ as the Hölder p-norm on $\mathbb{MEP}(n_1, \cdots, n_m)$.

Next we define weighted norms on $\mathbb{MEP}(n_1, \cdots, n_m)$. Consider $\mathbb{R}^m_+ := \{(x_1, \cdots, x_m)^T \in \mathbb{R}^m : x_i > 0\}$ and define the map $\mathbb{R}^m_+ \times \mathbb{MEP}(n_1, \cdots, n_m) \to \mathbb{MEP}(n_1, \cdots, n_m), (w, \mathbb{W}) \mapsto w \odot \mathbb{W}$ by $(w \odot \mathbb{W})(z) := (w_1 W_1(z), \cdots, w_m W_m(z))$, where $\mathbb{W}(z) = (W_1(z), \cdots, W_m(z))$ and $w = (w_1, \cdots, w_m)^T \in \mathbb{R}^m_+$.

Given a norm $|||\cdot|||$ on $\mathbb{MEP}(n_1, \cdots, n_m)$ and a weight $w \in \mathbb{R}^m_+$, we define a weighted norm $|||\cdot|||_w$ by $|||\mathbb{W}|||_w := |||w \odot \mathbb{W}|||$. For the norm $|||\cdot|||_V$ given in (2.6)

$$|||\mathbb{W}|||_{w,V} := ||(w_1||W_1||_V, \cdots, w_m||W_m||_V)||_V$$

defines a weighted norm on $\mathbb{MEP}(n_1, \cdots, n_m)$. The weighted Hölder p-norm is given by $|||\mathbb{W}|||_{w,p} := ||(w_1||W_1||_p, \cdots, w_m||W_m||_p)||_p$, where $1 \le p \le \infty$.

Inner Product. Now we define an inner product on $\mathbb{MEP}(n_1, \cdots, n_m)$. Let $\mathbb{X}, \mathbb{Y} \in \mathbb{MEP}(n_1, \cdots, n_m)$ be such that $\mathbb{X}(z) = (X_1(z), \cdots, X_m(z))$ and $\mathbb{Y}(z) = (Y_1(z), \cdots, Y_m(z))$, where $X_i(z) := \widehat{A}_i + \sum_{j=1}^m z_j \widehat{B}_{ij}$ and $Y_i(z) := \widetilde{A}_i + \sum_{j=1}^m z_j \widetilde{B}_{ij}$ for all $i = 1, \cdots, m$. Consider the usual inner product $\langle \cdot, \cdot \rangle$ on the space of matrices, i.e., $\langle A, B \rangle = \mathrm{tr}(B^*A)$, for all $A, B \in \mathbb{C}^{m\times m}$. Define $\langle X_i, Y_i \rangle := \langle \widehat{A}_i, \widetilde{A}_i \rangle + \sum_{j=1}^m \langle \widehat{B}_{ij}, \widetilde{B}_{ij} \rangle$ for all $i = 1, \cdots, m$. Then

$$\langle \mathbb{X}, \mathbb{Y} \rangle := \langle X_1, Y_1 \rangle + \cdots + \langle X_m, Y_m \rangle = \sum_{i=1}^m \left(\langle \widehat{A}_i, \widetilde{A}_i \rangle + \sum_{j=1}^m \langle \widehat{B}_{ij}, \widetilde{B}_{ij} \rangle \right) \tag{2.7}$$

defines an inner product on $\mathbb{MEP}(n_1, \cdots, n_m)$.

Dual Norm. The dual norm $|\!|\!|\cdot|\!|\!|_*$ of a norm $|\!|\!|\cdot|\!|\!|$ on $\mathbb{MEP}(n_1, \cdots, n_m)$ relative to the inner product $\langle \cdot, \cdot \rangle$ in (2.7) is given by $|\!|\!|\mathbb{W}|\!|\!|_* = \sup\{|\langle \mathbb{X}, \mathbb{W} \rangle| : |\!|\!|\mathbb{X}|\!|\!| = 1\}$. Note that $|\langle \mathbb{X}, \mathbb{Y} \rangle| \leq |\!|\!|\mathbb{X}|\!|\!| |\!|\!|\mathbb{Y}|\!|\!|_*$ for all $\mathbb{X}, \mathbb{Y} \in \mathbb{MEP}(n_1, \cdots, n_m)$.

Remark 2.2.1. ***(Compatibility of norms)*** *Let $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ be given by (2.1). Consider a monotone norm $||\cdot||_V$ on $\mathbb{C}^{m+1}$ and a matrix norm $||\cdot||$ on $\mathbb{C}^{n_i\times n_i}$. Then for $i = 1, \cdots, m$, we have,*

$$\begin{aligned} ||W_i(\lambda)|| &= ||A_i + \sum_{j=1}^m \lambda_j B_{ij}|| \leq ||A_i|| + \sum_{j=1}^m |\lambda_j|\ ||B_{ij}|| \\ &= \langle (||A_i||, ||B_{i1}||, \cdots, ||B_{im}||), (1, |\lambda_1|, \cdots, |\lambda_m|) \rangle \\ &\leq ||(||A_i||, ||B_{i1}||, \cdots, ||B_{im}||)||_V\ ||(1, |\lambda_1|, \cdots, |\lambda_m|)||_{V,*} \\ &= ||W_i||_V ||(1, \lambda^*)||_{V,*}, \end{aligned} \tag{2.8}$$

where $||\cdot||_{V,*}$ is the dual norm of the monotone norm $||\cdot||_V$.

Finally we have the following result.

Proposition 2.2.3. *Let $\mathbb{MEP}(n_1, \cdots, n_m)$ be equipped with the $|\!|\!|\cdot|\!|\!|_V$ norm and $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ be given by (2.1). Then $|\!|\!|\mathbb{W}|\!|\!|_{V,*} = ||(||W_1||_{V,*}, \cdots, ||W_m||_{V,*})||_{V,*}$ and $||W_i||_{V,*} = ||(||A_i||_*, ||B_{i1}||_*, \cdots, ||B_{im}||_*)||_{V,*}$ for $i = 1, \cdots, m$, where $||\cdot||_*$ is the dual norm of the matrix norm $||\cdot||$.*

Proof. First note that

$$\begin{aligned} |\langle \mathbb{X}, \mathbb{W} \rangle| &= |\langle X_1, W_1 \rangle + \cdots \langle X_m, W_m \rangle| \leq |\langle X_1, W_1 \rangle| + \cdots + |\langle X_m, W_m \rangle| \\ &\leq ||X_1||_V ||W_1||_{V,*} + \cdots + ||X_m||_V ||W_m||_{V,*} \\ &= \langle (||X_1||_V, \cdots, ||X_m||_V), (||W_1||_{V,*}, \cdots, ||W_m||_{V,*}) \rangle \\ &\leq ||(||X_1||_V, \cdots, ||X_m||_V)||_V\ ||(||W_1||_{V,*}, \cdots, ||W_m||_{V,*})||_{V,*} \\ &= |\!|\!|\mathbb{X}|\!|\!|_V\ ||(||W_1||_{V,*}, \cdots, ||W_m||_{V,*})||_{V,*} \end{aligned}$$

for all $\mathbb{X} \in \mathbb{MEP}(n_1, \cdots, n_m)$. Then $\|\mathbb{W}\|_{V,*} \leq ||(||W_1||_{V,*}, \cdots, ||W_m||_{V,*})||_{V,*}$.

Let $Y_i \in \partial ||W_i||_{V,*}$ for all $i = 1, \cdots, m$ and $(t_1, \cdots, t_m) \in \partial\, ||(||W_1||_{V,*}, \cdots, ||W_m||_{V,*})||_{V,*}$. Then $||Y_i||_V = 1$ and $\langle Y_i, W_i\rangle = ||W_i||_{V,*}$, for all $i = 1, \cdots, m$. Also $||(t_1, \cdots, t_m)||_V = 1$ and $\langle (t_1, \cdots, t_m), (||W_1||_{V,*}, \cdots, ||W_m||_{V,*})\rangle = ||(||W_1||_{V,*}, \cdots, ||W_m||_{V,*})||_{V,*}$.

Consider $\mathbb{Y} \in \mathbb{MEP}(n_1, \cdots, n_m)$ such that $\mathbb{Y}(z) := (t_1 Y_1(z), \cdots, t_m Y_m(z))$ for all $z \in \mathbb{C}^m$. Then $\|\mathbb{Y}\|_V = ||(|t_1|\ ||Y_1||_V, \cdots, |t_m|\ ||Y_m||_V)||_V = ||(t_1, \cdots, t_m)||_V = 1$ and

$$\begin{aligned} \langle \mathbb{Y}, \mathbb{W}\rangle &= t_1\langle Y_1, W_1\rangle + \cdots + t_m\langle Y_m, W_m\rangle = t_1||W_1||_{V,*} + \cdots + t_m||W_m||_{V,*} \\ &= \langle (t_1, \cdots, t_m), (||W_1||_{V,*}, \cdots, ||W_m||_{V,*})\rangle = ||(||W_1||_{V,*}, \cdots, ||W_m||_{V,*})||_{V,*}. \end{aligned}$$

Hence $\|\mathbb{W}\|_{V,*} = ||(||W_1||_{V,*}, \cdots, ||W_m||_{V,*})||_{V,*}$.

Now we derive $||W_i||_{V,*}$. Let $X_i(z) := \widehat{A}_i + \sum_{j=1}^m z_j \widehat{B}_{ij}$. Then we have

$$\begin{aligned} |\langle X_i, W_i\rangle| &= \left|\langle \widehat{A}_i, A_i\rangle + \sum_{j=1}^m \langle \widehat{B}_{ij}, B_{ij}\rangle\right| \leq |\langle \widehat{A}_i, A_i\rangle| + \sum_{j=1}^m |\langle \widehat{B}_{ii}, B_{ij}\rangle| \\ &\leq ||\widehat{A}_i||\ ||A_i||_* + \sum_{j=1}^m ||\widehat{B}_{ij}||\ ||B_{ij}||_* \\ &= \langle (||\widehat{A}_i||, ||\widehat{B}_{i1}||, \cdots, ||\widehat{B}_{im}||), (||A_i||_*, ||B_{i1}||_*, \cdots, ||B_{im}||_*)\rangle \\ &\leq ||(||\widehat{A}_i||, ||\widehat{B}_{i1}||, \cdots, ||\widehat{B}_{im}||)||_V ||(||A_i||_*, ||B_{i1}||_*, \cdots, ||B_{im}||_*)||_{V,*} \\ &= ||X_i||_V ||(||A_i||_*, ||B_{i1}||_*, \cdots, ||B_{im}||_*)||_{V,*}. \end{aligned}$$

So $||W_i||_{V,*} \leq ||(||A_i||_*, ||B_{i1}||_*, \cdots, ||B_{im}||_*)||_{V,*}$.

To show the equality, let $\widetilde{A}_i \in \partial ||A_i||_*, \widetilde{B}_{ij} \in \partial ||B_{ij}||_*$ for all $j = 1, \cdots, m$ and $(u_0, u_1, \cdots, u_m)^T \in \partial ||(||A_i||_*, ||B_{i1}||_*, \cdots, ||B_{im}||_*)||_{V,*}$. Define $C_i := u_0 \widetilde{A}_i$ and $D_{ij} := u_j \widetilde{B}_{ij}$ for all $j = 1, \cdots, m$ and consider $\widetilde{X}_i(z) := C_i + \sum_{j=1}^m z_j D_{ij}$. Then $||\widetilde{X}_i||_V = ||(|u_0|\ ||\widetilde{A}_i||, |u_1|\ ||\widetilde{B}_{i1}||, \cdots, |u_m|\ ||\widetilde{B}_{im}||)||_V = 1$ and

$$\begin{aligned} \langle \widetilde{X}_i, W_i\rangle = \langle u_0 \widetilde{A}_i, A_i\rangle + \sum_{j=1}^m \langle u_j \widetilde{B}_{ij}, B_{ij}\rangle &= u_0||A_i||_* + u_1||B_{i1}||_* + \cdots + u_m||B_{im}||_* \\ &= ||(||A_i||_*, ||B_{i1}||_*, \cdots, ||B_{im}||_*)||_{V,*}. \end{aligned}$$

Therefore $||W_i||_{V,*} = ||(||A_i||_*, ||B_{i1}||_*, \cdots, ||B_{im}||_*)||_{V,*}$. This completes the proof. ☐

2.3 Sensitivity analysis of simple eigenvalues

Now we undertake the sensitivity analysis of a simple eigenvalue of a regular nonhomogeneous MEP. First we define the condition number of an eigenvalue of a regular nonhomogeneous MEP as follows.

Definition 2.3.1. (Condition number) *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular and* $\lambda \in \sigma(\mathbb{W})$. *Then we define the* ***condition number*** *of* λ *by*

$$\mathrm{cond}(\lambda, \mathbb{W}) := \limsup_{\|\Delta\mathbb{W}\| \to 0} \frac{\mathrm{dist}(\lambda, \sigma(\mathbb{W} + \Delta\mathbb{W}))}{\|\Delta\mathbb{W}\|},$$

where $\Delta\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *and* $\mathrm{dist}(\lambda, \sigma(\mathbb{W} + \Delta\mathbb{W})) := \min\{||\lambda - \mu||_2 : \mu \in \sigma(\mathbb{W} + \Delta\mathbb{W})\}$.

For a simple eigenvalue $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$, we prove the existence of an open neighborhood $\mathrm{nbd}(\mathbb{W})$ containing $\mathbb{W}$ and a smooth map $\lambda : \mathrm{nbd}(\mathbb{W}) \to \mathbb{C}^m$ such that $\lambda(\mathbb{W}) = \lambda_{\mathbb{W}}$ and $\lambda(\mathbb{X}) \in \sigma(\mathbb{X})$ is algebraically simple for all $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$. We also derive the derivative of this map in the following result.

Theorem 2.3.1. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular and* $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ *be a simple eigenvalue of* $\mathbb{W}$. *Then there is an open neighborhood* $\mathrm{nbd}(\mathbb{W}) \subset \mathbb{MEP}(n_1, \cdots, n_m)$ *containing* $\mathbb{W}$ *and a smooth function* $\lambda : \mathrm{nbd}(\mathbb{W}) \to \mathbb{C}^m$ *such that* $\lambda(\mathbb{W}) = \lambda_{\mathbb{W}}$ *and* $\lambda(\mathbb{X})$ *is a simple eigenvalue of* $\mathbb{X}$ *for all* $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$. *Further, for* $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$ *such that* $\mathbb{X}(z) = (X_1(z), \cdots, X_m(z))$, *the derivative* $D\lambda(\mathbb{X}) : \mathbb{MEP}(n_1, \cdots, n_m) \to \mathbb{C}^m$ *is given by*

$$D\lambda(\mathbb{X})\mathbb{H} = -(J_{\mathbb{X}}(\lambda(\mathbb{X})))^{-1} \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(X_1(\lambda(\mathbb{X})))H_1(\lambda(\mathbb{X}))) \\ \vdots \\ \mathrm{tr}(\mathrm{adj}(X_m(\lambda(\mathbb{X})))H_m(\lambda(\mathbb{X}))) \end{bmatrix} \tag{2.9}$$

for all $\mathbb{H} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *such that* $\mathbb{H}(z) = (H_1(z), \cdots, H_m(z))$.

Proof. Define $\phi : \mathbb{MEP}(n_1, \cdots, n_m) \times \mathbb{C}^m \to \mathbb{C}^m, (\mathbb{Y}, z) \mapsto \mathbf{det}(\mathbb{Y}(z))$ and consider $\mathbb{V}(\phi) := \{(\mathbb{Y}, z) \in \mathbb{MEP}(n_1, \cdots, n_m) \times \mathbb{C}^m : \phi(\mathbb{Y}, z) = 0\}$. It is obvious that $(\mathbb{W}, \lambda_{\mathbb{W}}) \in \mathbb{V}(\phi)$ and $\partial_\lambda \phi(\mathbb{W}, \lambda_{\mathbb{W}}) = J_{\mathbb{W}}(\lambda_{\mathbb{W}})$ is nonsingular by Theorem 2.2.1. Hence by the Implicit function theorem there is an open neighborhood $\mathrm{nbd}(\mathbb{W})$ containing $\mathbb{W}$ and a smooth function $\lambda : \mathrm{nbd}(\mathbb{W}) \to \mathbb{C}^m$ such that $\lambda(\mathbb{W}) = \lambda_{\mathbb{W}}$ and $\mathbb{V}(\phi) \cap (\mathrm{nbd}(\mathbb{W}) \times \mathbb{C}^m) = \{(\mathbb{X}, \lambda(\mathbb{X})) : \mathbb{X} \in \mathrm{nbd}(\mathbb{W})\}$ is the graph of the function $\mathbb{X} \mapsto \lambda(\mathbb{X})$ for all $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$. So $\phi(\mathbb{X}, \lambda(\mathbb{X})) = 0$, i.e.,$\lambda(\mathbb{X}) \in \sigma(\mathbb{X})$, for all $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$. In order to show that $\lambda(\mathbb{X})$ is simple for all $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$, if possible let $\lambda(\widehat{\mathbb{X}})$ is an multiple eigenvalue of $\widehat{\mathbb{X}}$, for some $\widehat{\mathbb{X}} \in \mathrm{nbd}(\mathbb{W})$. Then the eigenvalue $\lambda_{\mathbb{W}}$ and an eigenvalue $\mu_{\mathbb{W}} \neq \lambda_{\mathbb{W}}$ of the MEP $\mathbb{W}$ must move and coalesce at $\lambda(\widehat{\mathbb{X}})$ when $\mathbb{X}$ varies from $\mathbb{W}$ to $\widehat{\mathbb{X}}$. But the intersection of the eigenvalue paths $(\mathbb{X}, \lambda(\mathbb{X}))$ with $\lambda(\mathbb{W}) = \lambda_{\mathbb{W}}$ and $(\mathbb{X}, \mu(\mathbb{X}))$ with $\mu(\mathbb{W}) = \mu_{\mathbb{W}}$ at $(\widehat{\mathbb{X}}, \lambda(\widehat{\mathbb{X}}))$ would contradict that $\mathbb{V}(\phi) \cap (\mathrm{nbd}(\mathbb{W}) \times \mathbb{C}^m) = \{(\mathbb{X}, \lambda(\mathbb{X})) : \mathbb{X} \in \mathrm{nbd}(\mathbb{W})\}$ is the graph of the map $\mathbb{X} \mapsto \lambda(\mathbb{X})$ for all $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$. Hence $\lambda(\mathbb{X}) \in \sigma(\mathbb{X})$ is simple for all $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$.

For the remaining part, differentiating $\phi(\mathbb{X}, \lambda(\mathbb{X})) = 0$ with respect to $\mathbb{X}$ we have

$$\partial_{\lambda}\phi(\mathbb{X}, \lambda(\mathbb{X}))D\lambda(\mathbb{X})\mathbb{H} + \partial_{\mathbb{X}}\phi(\mathbb{X}, \lambda(\mathbb{X}))\mathbb{H} = 0 \tag{2.10}$$

for all $\mathbb{H} \in \mathbb{MEP}(n_1, \cdots, n_m)$. Note that $\partial_{\lambda}\phi(\mathbb{X}, \lambda(\mathbb{X})) = J_{\mathbb{X}}(\lambda(\mathbb{X}))$ is nonsingular and $\partial_{\mathbb{X}}\phi(\mathbb{X}, \lambda(\mathbb{X})) : \mathbb{MEP}(n_1, \cdots, n_m) \to \mathbb{C}^m$ is given by

$$\partial_{\mathbb{X}}\phi(\mathbb{X}, \lambda(\mathbb{X}))\mathbb{H} = \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(X_1(\lambda(\mathbb{X})))H_1(\lambda(\mathbb{X}))), & \cdots, & \mathrm{tr}(\mathrm{adj}(X_m(\lambda(\mathbb{X})))H_m(\lambda(\mathbb{X}))) \end{bmatrix}^T.$$

Therefore by (2.10), we have

$$D\lambda(\mathbb{X})\mathbb{H} = -(J_{\mathbb{X}}(\lambda(\mathbb{X})))^{-1} \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(X_1(\lambda(\mathbb{X})))H_1(\lambda(\mathbb{X}))) \\ \vdots \\ \mathrm{tr}(\mathrm{adj}(X_m(\lambda(\mathbb{X})))H_m(\lambda(\mathbb{X}))) \end{bmatrix},$$

which proves (2.9). □

The following result is immediate from Definition 2.3.1.

Proposition 2.3.1. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular and* $\lambda \in \sigma(\mathbb{W})$ *be simple. Then* $\mathrm{cond}(\lambda, \mathbb{W}) = ||D\lambda(\mathbb{W})||_V$, *where* $||\cdot||_V$ *is the matrix norm induced by the monotone vector norm* $||\cdot||_V$.

Theorem 2.3.2. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *given by (2.1) be regular and* $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ *be a simple eigenvalue. Let* $x := x_1 \otimes \cdots \otimes x_m$ *and* $y := y_1 \otimes \cdots \otimes y_m$ *be right and left* *eigenvector of* $\mathbb{W}$ *corresponding to* $\lambda_{\mathbb{W}}$*, respectively. Then there is an open neighborhood* $\mathrm{nbd}(\mathbb{W}) \subset \mathbb{MEP}(n_1, \cdots, n_m)$ *containing* $\mathbb{W}$ *and a smooth function* $\lambda : \mathrm{nbd}(\mathbb{W}) \to \mathbb{C}^m$ *such that* $\lambda(\mathbb{W}) = \lambda_{\mathbb{W}}$ *and* $\lambda(\mathbb{X})$ *is a simple eigenvalue of* $\mathbb{X}$ *for all* $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$*. Further, the derivative* $D\lambda(\mathbb{W}) : \mathbb{MEP}(n_1, \cdots, n_m) \to \mathbb{C}^m$ *is given by*

$$D\lambda(\mathbb{W})\mathbb{H} = -(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1} \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(\lambda_{\mathbb{W}}))H_1(\lambda_{\mathbb{W}})) \\ \vdots \\ \mathrm{tr}(\mathrm{adj}(W_m(\lambda_{\mathbb{W}}))H_m(\lambda_{\mathbb{W}})) \end{bmatrix} \tag{2.11}$$

$$= -(J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1} \begin{bmatrix} y_1^* H_1(\lambda_{\mathbb{W}}) x_1 \\ \vdots \\ y_m^* H_m(\lambda_{\mathbb{W}}) x_m \end{bmatrix} \tag{2.12}$$

for all $\mathbb{H} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *such that* $\mathbb{H}(z) = (H_1(z), \cdots, H_m(z))$*, where* $J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x)$ *is given in (2.3).*

a) The condition number $\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W})$ *is given by*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \; ||(1, \lambda_{\mathbb{W}}^*)||_{V,*} \max_i ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_*. \tag{2.13}$$

In particular, for the spectral or Frobenius norm on $\mathbb{C}^{n_i \times n_i}$*, we have*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \; ||(1, \lambda_{\mathbb{W}}^*)||_{V,*} \max_i \prod_{k=1}^{n_i-1} \sigma_k(W_i(\lambda_{\mathbb{W}})), \tag{2.14}$$

where $\sigma_1(W_i(\lambda_{\mathbb{W}})) \geq \cdots \geq \sigma_{n_i}(W_i(\lambda_{\mathbb{W}}))$ *are the singular values of* $W_i(\lambda_{\mathbb{W}})$ *for* $i = 1, \cdots, m$*.*

b) Also we have

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_V \; ||(1, \lambda_{\mathbb{W}}^*)||_{V,*} \max_i ||y_i x_i^*||_*. \tag{2.15}$$

For a subordinate matrix norm $|| \cdot ||$ *on* $\mathbb{C}^{n_i \times n_i}$*, we have*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_V \; ||(1, \lambda_{\mathbb{W}}^*)||_{V,*} \max_i ||y_i||_* ||x_i||. \tag{2.16}$$

In particular, for the spectral or Frobenius norm on $\mathbb{C}^{n_i \times n_i}$*, we have*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_V \; ||(1, \lambda_{\mathbb{W}}^*)||_{V,*} \max_i ||y_i||_2 ||x_i||_2. \tag{2.17}$$

Proof. By Theorem 2.3.1 the derivative $D\lambda(\mathbb{W}) : \mathbb{MEP}(n_1, \cdots, n_m) \to \mathbb{C}^m$ of λ at $\mathbb{W}$ is given by

$$D\lambda(\mathbb{W})\mathbb{H} = -(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1} \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(\lambda_{\mathbb{W}}))H_1(\lambda_{\mathbb{W}})) \\ \vdots \\ \mathrm{tr}(\mathrm{adj}(W_m(\lambda_{\mathbb{W}}))H_m(\lambda_{\mathbb{W}})) \end{bmatrix},$$

which yields (2.11).

Since $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ is simple, so by Proposition 2.2.2, it follows that $\lambda_{\mathbb{W}}$ is geometrically simple. Then by Proposition 2.2.1, there exist left and right eigenvectors $u = u_1 \otimes \cdots \otimes u_m$ and $v = v_1 \otimes \cdots \otimes v_m$ of $\mathbb{W}$ corresponding to $\lambda_{\mathbb{W}}$, respectively, such that $\mathrm{adj}(W_i(\lambda_{\mathbb{W}})) = u_i v_i^*$ for all $i = 1, \cdots, m$ and $J_{\mathbb{W}}(\lambda_{\mathbb{W}}) = J_{\mathbb{W}}(\lambda_{\mathbb{W}}; u, v)$. Also there exist nonzero scalars α_i, β_i such that $u_i = \alpha_i y_i$ and $v_i = \beta_i x_i$ for all $i = 1, \cdots, m$. Consequently, we have

$$J_{\mathbb{W}}(\lambda_{\mathbb{W}}) = J_{\mathbb{W}}(\lambda_{\mathbb{W}}; u, v) = \begin{bmatrix} \overline{\alpha_1}\beta_1 & & \\ & \ddots & \\ & & \overline{\alpha_m}\beta_m \end{bmatrix} J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x).$$

So

$$(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1} = (J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1} \begin{bmatrix} \overline{\alpha_1}\beta_1 & & \\ & \ddots & \\ & & \overline{\alpha_m}\beta_m \end{bmatrix}^{-1}.$$

For $i = 1, \cdots, m$, we have

$$\begin{aligned} \mathrm{tr}(\mathrm{adj}(W_i(\lambda_{\mathbb{W}}))H_i(\lambda_{\mathbb{W}})) &= \langle H_i(\lambda_{\mathbb{W}}), \mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^* \rangle = \langle H_i(\lambda_{\mathbb{W}}), u_i v_i^* \rangle \\ &= u_i^* H_i(\lambda_{\mathbb{W}}) v_i = (\alpha_i y_i)^* H_i(\lambda_{\mathbb{W}})(\beta_i x_i) = \overline{\alpha_i}\beta_i y_i^* H_i(\lambda_{\mathbb{W}}) x_i. \end{aligned}$$

Consequently,

$$\begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(\lambda_{\mathbb{W}}))H_1(\lambda_{\mathbb{W}})) \\ \vdots \\ \mathrm{tr}(\mathrm{adj}(W_m(\lambda_{\mathbb{W}}))H_m(\lambda_{\mathbb{W}})) \end{bmatrix} = \begin{bmatrix} \overline{\alpha_1}\beta_1 & & \\ & \ddots & \\ & & \overline{\alpha_m}\beta_m \end{bmatrix} \begin{bmatrix} y_1^* H_1(\lambda_{\mathbb{W}}) x_1 \\ \vdots \\ y_m^* H_m(\lambda_{\mathbb{W}}) x_m \end{bmatrix}.$$

Hence by (2.11), we have

$$D\lambda(\mathbb{W})\mathbb{H} = -(J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1} \begin{bmatrix} y_1^* H_1(\lambda_{\mathbb{W}}) x_1 \\ \vdots \\ y_m^* H_m(\lambda_{\mathbb{W}}) x_m \end{bmatrix},$$

which proves (2.12).

Since $D\lambda(\mathbb{W}) : \mathbb{MEP}(n_1, \cdots, n_m) \to \mathbb{C}^m$ is linear, so $||D\lambda(\mathbb{W})||_V = \sup\limits_{|||\mathbb{H}|||_V=1} ||D\lambda(\mathbb{W})\mathbb{H}||_V$. From (2.11) we have

$$||D\lambda(\mathbb{W})\mathbb{H}||_V \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \left|\left| \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(\lambda_{\mathbb{W}}))H_1(\lambda_{\mathbb{W}})) \\ \vdots \\ \mathrm{tr}(\mathrm{adj}(W_m(\lambda_{\mathbb{W}}))H_m(\lambda_{\mathbb{W}})) \end{bmatrix} \right|\right|_V . \tag{2.18}$$

It is easy to see that

$$\begin{aligned} |\mathrm{tr}(\mathrm{adj}(W_i(\lambda_{\mathbb{W}}))H_i(\lambda_{\mathbb{W}}))| = |\langle H_i(\lambda_{\mathbb{W}}), (\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*\rangle| &\leq ||H_i(\lambda_{\mathbb{W}})|| \; ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_* \\ &\leq ||H_i(\lambda_{\mathbb{W}})|| \; \max_i ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_* \end{aligned}$$

for all $i = 1, \cdots, m$. By (2.8) we have $||H_i(\lambda_{\mathbb{W}})|| \leq ||(1, \lambda^*_{\mathbb{W}})||_{V,*} \, ||H_i||_V$ for $i = 1, \cdots, m$. This shows that $|\mathrm{tr}(\mathrm{adj}(W_i(\lambda_{\mathbb{W}}))H_i(\lambda_{\mathbb{W}}))| \leq ||(1, \lambda^*_{\mathbb{W}})||_{V,*} \max\limits_i ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_* \, ||H_i||_V$ for all $i = 1, \cdots, m$. Since $|| \cdot ||_V$ is a monotone norm, we have

$$\left|\left| \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(\lambda_{\mathbb{W}}))H_1(\lambda_{\mathbb{W}})) \\ \vdots \\ \mathrm{tr}(\mathrm{adj}(W_m(\lambda_{\mathbb{W}}))H_m(\lambda_{\mathbb{W}})) \end{bmatrix} \right|\right|_V \leq ||(1, \lambda^*_{\mathbb{W}})||_{V,*} \max_i ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_* \, |||\mathbb{H}|||_V .$$

Now by (2.18) we have

$$||D\lambda(\mathbb{W})\mathbb{H}||_V \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V ||(1, \lambda^*_{\mathbb{W}})||_{V,*} \max_i ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_* \, |||\mathbb{H}|||_V .$$

Therefore

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) = ||D\lambda(\mathbb{W})||_V \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \; ||(1, \lambda^*_{\mathbb{W}})||_{V,*} \max_i ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_*,$$

which proves (2.13). By Lemma 1.1.1, we have $||\mathrm{adj}(W_i(\lambda))||_2 = \prod_{k=1}^{n_i-1} \sigma_k(W_i(\lambda))$ for all $i = 1, \cdots, m$ and thus (2.14) follows. Similarly, by (2.12), we have

$$||D\lambda(\mathbb{W})\mathbb{H}||_V \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_V ||(1, \lambda^*_{\mathbb{W}})||_{V,*} \max_i ||y_i x_i^*||_* \, |||\mathbb{H}|||_V ,$$

which proves (2.15). Also note that for a subordinate matrix norm $||y_i x_i^*||_* = ||y_i||_* ||x_i||$ and, in particular, for the spectral or Frobenius norm, $||y_i x_i^*||_2 = ||y_i||_2 ||x_i||_2$ for all $i = 1, \cdots, m$. Hence (2.16) and (2.17) follow. □

For Hölder p-norm on $\mathbb{MEP}(n_1, \cdots, n_m)$, we have the following result.

Corollary 2.3.3. *Let $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ given by (2.1) be regular and $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ be a simple eigenvalue. Let $x := x_1 \otimes \cdots \otimes x_m$ and $y := y_1 \otimes \cdots \otimes y_m$ be right and left eigenvector of $\mathbb{W}$ corresponding to $\lambda_{\mathbb{W}}$, respectively.*

a) Then we have

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_p \; ||(1, \lambda_{\mathbb{W}}^*)||_q \; \max_i ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_*,$$

where $p^{-1} + q^{-1} = 1$.

In particular, for the spectral or Frobenius norm on $\mathbb{C}^{n_i \times n_i}$, we have

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_p \; ||(1, \lambda_{\mathbb{W}}^*)||_q \; \max_i \prod_{k=1}^{n_i - 1} \sigma_k(W_i(\lambda_{\mathbb{W}})),$$

where $\sigma_1(W_i(\lambda_{\mathbb{W}})) \geq \cdots \geq \sigma_{n_i}(W_i(\lambda_{\mathbb{W}}))$ are singular values of $W_i(\lambda_{\mathbb{W}})$ for $i = 1, \cdots, m$.

b) Also we have

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_p \; ||(1, \lambda_{\mathbb{W}}^*)||_q \; \max_i ||y_i x_i^*||_*.$$

For a subordinate matrix norm $|| \cdot ||$ on $\mathbb{C}^{n_i \times n_i}$, we have

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_p \; ||(1, \lambda_{\mathbb{W}}^*)||_q \; \max_i ||y_i||_* ||x_i||.$$

In particular, for the spectral or Frobenius norm on $\mathbb{C}^{n_i \times n_i}$, we have

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_p \; ||(1, \lambda_{\mathbb{W}}^*)||_q \; \max_i ||y_i||_2 ||x_i||_2.$$

Remark 2.3.1. *When $m = 1$, the regular nonhomogeneous MEP reduces to a regular matrix pencil $\mathbb{W}(z) := A + zB$, where $z \in \mathbb{C}$ and $A, B \in \mathbb{C}^{n_1 \times n_1}$. Then for a simple eigenvalue $\lambda_{\mathbb{W}}$ of the regular matrix pencil $\mathbb{W}$, we have $J_{\mathbb{W}}(\lambda_{\mathbb{W}}) = \mathrm{tr}(\mathrm{adj}(W(\lambda_{\mathbb{W}}))B) \neq 0$ and $(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1} = \frac{1}{\mathrm{tr}(\mathrm{adj}(W(\lambda_{\mathbb{W}}))B)}$. So it follows from (2.13) that*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq \frac{||(1, \lambda_{\mathbb{W}})||_{V,*} \; ||(\mathrm{adj}(\mathbb{W}(\lambda_{\mathbb{W}})))^*||_*}{|\mathrm{tr}(\mathrm{adj}(\mathbb{W}(\lambda_{\mathbb{W}}))B)|}. \tag{2.19}$$

It is shown in [7] that the equality holds in (2.19). We, therefore, consider

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) = ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \ ||(1, \lambda_{\mathbb{W}}^*)||_{V,*} \max_j ||(\mathrm{adj}(W_j(\lambda_{\mathbb{W}})))^*||_*$$

to be the condition number of a simple eigenvalue $\lambda_{\mathbb{W}}$ of a regular nonhomogeneous MEP $\mathbb{W}$. Equivalently, if $x := x_1 \otimes \cdots \otimes x_m$ and $y := y_1 \otimes \cdots \otimes y_m$ are right and left eigenvectors of $\mathbb{W}$ corresponding to $\lambda_{\mathbb{W}}$, respectively, then we have

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) = ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_V \ ||(1, \lambda_{\mathbb{W}}^*)||_{V,*} \max_i ||y_i x_i^*||_*.$$

When $\mathbb{MEP}(n_1, \cdots, n_m)$ is equipped with the weighted norm with nonzero positive weights, we have the following result.

Theorem 2.3.4. *Let $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ given by (2.1) be regular and $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ be simple. Consider the weighted norm $|||\cdot|||_{w,V}$ on $\mathbb{MEP}(n_1, \cdots, n_m)$ with nonzero positive weights $w_i := ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_*$ for $i = 1, \cdots, m$. Then $\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) = ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \ ||(1, \lambda_{\mathbb{W}}^*)||_{V,*}$. In particular, for the weighted Hölder p-norm $|||\cdot|||_{w,p}$ on $\mathbb{MEP}(n_1, \cdots, n_m)$ with nonzero positive weights $w_i := ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_*$ for $i = 1, \cdots, m$, we have $\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) = ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_p \ ||(1, \lambda_{\mathbb{W}}^*)||_q$, where $p^{-1} + q^{-1} = 1$.*

Proof. By (2.8), we have

$$|\mathrm{tr}(\mathrm{adj}(W_i(\lambda_{\mathbb{W}}))H_i(\lambda_{\mathbb{W}}))| \leq ||(1, \lambda_{\mathbb{W}}^*)||_{V,*}||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_*||H_i||_V = w_i||(1, \lambda_{\mathbb{W}}^*)||_{V,*}||H_i||_V.$$

Thus we have $||D\lambda(\mathbb{W})\mathbb{H}||_V \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \ ||(1, \lambda_{\mathbb{W}}^*)||_{V,*} \ |||\mathbb{H}|||_{w,V}$.

To prove the equality, consider $\mathbb{Y} \in \mathbb{MEP}(n_1, \cdots, n_m)$ such that $\mathbb{Y}(z) = (Y_1(z), \cdots, Y_m(z))$ and $||Y_i||_V = 1$, $\mathrm{tr}(\mathrm{adj}(W_i(\lambda_{\mathbb{W}}))Y_i(\lambda_{\mathbb{W}})) = \langle Y_i, (1, \lambda_{\mathbb{W}}^*) \otimes \mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^* \rangle = w_i||(1, \lambda_{\mathbb{W}}^*)||_{V,*}$, where $(1, \lambda_{\mathbb{W}}^*) \otimes \mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^* : \mathbb{C}^m \to \mathbb{C}^{n_i \times n_i}$ is given by $((1, \lambda_{\mathbb{W}}^*) \otimes \mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^*)(z) = (1 + \lambda_{\mathbb{W}}^* z)\mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^*$ for $i = 1, \cdots, m$. Let $u = (u_1, \cdots, u_m)^T \in \mathbb{C}^m$ be such that $||u||_V = 1$ and $||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}u||_V = ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V$. Construct $\mathbb{X} \in \mathbb{MEP}(n_1, \cdots, n_m)$ such that $\mathbb{X}(z) := \left(\frac{u_1}{w_1}Y_1(z), \cdots, \frac{u_m}{w_m}Y_m(z)\right)$ for all $z \in \mathbb{C}^m$. Then we have

$$|||\mathbb{X}|||_{w,V} = \left|\left|\left(\frac{w_1|u_1|}{w_1} \ ||Y_1||_V, \cdots, \frac{w_m|u_m|}{w_m} \ ||Y_m||_V\right)\right|\right|_V = ||u||_V = 1$$

and by (2.11), we have

$$\begin{aligned}
||D\lambda(\mathbb{W})\mathbb{X}||_V &= \left\| -(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1} \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(\lambda_{\mathbb{W}}))\frac{u_1}{w_1}Y_1(z)) \\ \vdots \\ \mathrm{tr}(\mathrm{adj}(W_m(\lambda_{\mathbb{W}}))\frac{u_m}{w_m}Y_m(z)) \end{bmatrix} \right\|_V \\
&= \left\| -(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1} \begin{bmatrix} \frac{u_1}{w_1}\cdot w_1 \cdot ||(1,\lambda_{\mathbb{W}}^*)||_{V,*} \\ \vdots \\ \frac{u_m}{w_m}\cdot w_m \cdot ||(1,\lambda_{\mathbb{W}}^*)||_{V,*} \end{bmatrix} \right\|_V \\
&= ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}u||_V \; ||(1,\lambda_{\mathbb{W}}^*)||_{V,*} = ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \; ||(1,\lambda_{\mathbb{W}}^*)||_{V,*}.
\end{aligned}$$

So $\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) = ||D\lambda(\mathbb{W})||_V = ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \; ||(1,\lambda_{\mathbb{W}}^*)||_{V,*}$. This completes the proof. □

Next we define fast perturbation $\Delta\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ for a simple eigenvalue of a regular nonhomogeneous MEP.

Definition 2.3.2. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular and* $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ *be simple. Then* $\Delta\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *is said to be a fast perturbation for* $\lambda_{\mathbb{W}}$ *if* $||D\lambda(\mathbb{W})\Delta\mathbb{W}||_V = ||D\lambda(\mathbb{W})|| |||\Delta\mathbb{W}|||_V$.

When $\mathbb{MEP}(n_1, \cdots, n_m)$ is equipped with a weighted norm $|||\cdot|||_{w,V}$, we can construct a fast perturbation for a simple eigenvalue $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ as follows.

Proposition 2.3.2. (Fast Perturbation) *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular and* $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ *be simple. Let* $\mathbb{MEP}(n_1, \cdots, n_m)$ *be equipped with the weighted norm* $|||\cdot|||_{w,V}$ *with nonzero positive weights* $w_j := ||(\mathrm{adj}(W_j(\lambda_{\mathbb{W}})))^*||_*$ *for* $j = 1, \cdots, m$. *Let* $y = (y_1, \cdots, y_m)^T \in \mathbb{C}^m$ *be such that* $||y||_V = 1$ *and* $||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}\, y||_V = ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V$. *Let* $\mathbb{Y} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be such that* $\mathbb{Y}(z) := (Y_1(z), \cdots, Y_m(z))$ *and* $||Y_i||_V = 1$, $\mathrm{tr}(\mathrm{adj}(W_i(\lambda_{\mathbb{W}}))Y_i(\lambda_{\mathbb{W}})) = \langle Y_i, (1,\lambda_{\mathbb{W}}^*)\otimes\mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^*\rangle = w_i||(1,\lambda_{\mathbb{W}}^*)||_{V,*}$, *where* $(1,\lambda_{\mathbb{W}}^*)\otimes \mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^* : \mathbb{C}^m \to \mathbb{C}^{n_i\times n_i}$ *is given by* $((1,\lambda_{\mathbb{W}}^*)\otimes\mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^*)(z) = (1+\lambda_{\mathbb{W}}^* z)\mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^*$ *for* $i = 1, \cdots, m$. *Define* $\Delta\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *by* $\Delta\mathbb{W}(z) := \left(\frac{y_1}{w_1}Y_1(z), \cdots, \frac{y_m}{w_m}Y_m(z)\right)$. *Then* $\Delta\mathbb{W}$ *is a fast perturbation for* $\lambda_{\mathbb{W}}$, *i.e.,* $||D\lambda(\mathbb{W})\Delta\mathbb{W}||_V = ||D\lambda(\mathbb{W})|| |||\Delta\mathbb{W}|||_V$.

Proof. First note that $\|\Delta\mathbb{W}\|_{w,V} = \left|\left|\left(w_1\frac{|y_1|}{w_1}, \cdots, w_m\frac{|y_m|}{w_m}\right)\right|\right|_V = ||y||_V = 1$. Again it follows from (2.11) that

$$\begin{aligned} D\lambda(\mathbb{W})\Delta\mathbb{W} &= -(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}\begin{bmatrix} \mathrm{tr}\left(\mathrm{adj}(W_1(\lambda_{\mathbb{W}}))\frac{y_1 Y_1(\lambda_{\mathbb{W}})}{w_1}\right) \\ \vdots \\ \mathrm{tr}\left(\mathrm{adj}(W_m(\lambda_{\mathbb{W}}))\frac{y_m Y_m(\lambda_{\mathbb{W}})}{w_m}\right) \end{bmatrix} \\ &= -(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}\begin{bmatrix} ||(1,\lambda_{\mathbb{W}}^*)||_{V,*}\; y_1 \\ \vdots \\ ||(1,\lambda_{\mathbb{W}}^*)||_{V,*}\; y_m \end{bmatrix} = -||(1,\lambda_{\mathbb{W}}^*)||_{V,*}(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}y. \end{aligned}$$

Therefore $||D\lambda(\mathbb{W})\Delta\mathbb{W}||_V = ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V\; ||(1,\lambda_{\mathbb{W}}^*)||_{V,*} = ||D\lambda(\mathbb{W})||\; \|\Delta\mathbb{W}\|_{w,V}$. Hence by Definition 2.3.2, $\Delta\mathbb{W}$ is a fast perturbation for $\lambda_{\mathbb{W}}$. ☐

2.4 Holomorphic perturbations

Next we discuss holomorphic evolution of a simple eigenvalue of a regular nonhomogeneous MEP. Let $\mathbb{W} : \mathbb{C}^p \to \mathbb{MEP}(n_1, \cdots, n_m)$ be a holomorphic map. Then $\mathbb{W}(t) \in \mathbb{MEP}(n_1, \cdots, n_m)$ for $t \in \mathbb{C}^p$. We write $\mathbb{W}(t)(z)$ as $\mathbb{W}(t,z) := (W_1(t,z), \cdots, W_m(t,z))$. Then $\mathbb{W}(t,z)$ is said to be regular if for each $t \in \mathbb{C}^p$ there is a $z \in \mathbb{C}^m$ such that $\prod_{i=1}^m \det(W_i(t,z)) \neq 0$. Define $\mathcal{F}(t,z) := \mathbf{det}(\mathbb{W}(t,z)) = (\det(W_1(t,z)), \cdots, \det(W_m(t,z)))^T$. Recall that, $\lambda \in \mathbb{C}^m$ is said to be an eigenvalue of $\mathbb{W}(t,z)$ if $\mathcal{F}(t,\lambda) = 0$. We refer to $x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ and $y_1 \otimes \cdots \otimes y_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ as a right and left eigenvectors of $\mathbb{W}(t,z)$ corresponding to an eigenvalue λ, respectively, if $W_i(t,\lambda)x_i = 0 = y_i^* W_i(t,\lambda)$ for all $i = 1, \cdots, m$. An eigenvalue λ of $\mathbb{W}(t,z)$ is said to be geometrically simple if $\prod_{j=1}^m \dim\ker(W_j(t,\lambda)) = 1$. Further to define algebraically simple eigenvalue, consider the differentiable function $\mathbb{C}^m \to \mathbb{C}^m, z \mapsto \mathcal{F}(t,z)$ for $t \in \mathbb{C}^p$. An eigenvalue λ of $\mathbb{W}(t,z)$ is said to be simple if the Jacobian matrix $J_{\mathbb{W}(t,\lambda)}(\lambda) := \left[\mathrm{tr}(\mathrm{adj}(W_i(t,\lambda))\partial_{\lambda_j}(W_i(t,\lambda)))\right]_{m\times m}$ is nonsingular, where $\partial_{\lambda_j}(W_i(t,\lambda))$ is the partial derivative of $W_i(t,\lambda)$ with respect to λ_j. For left and right eigenvectors $y := y_1 \otimes \cdots \otimes y_m$ and $x := x_1 \otimes \cdots \otimes x_m$ of $\mathbb{W}(t,z)$ corresponding to λ, respectively, define $J_{\mathbb{W}(t,\lambda)}(t,\lambda; y, x) := \left[y_i^* \partial_{\lambda_j}(W_i(t,\lambda))x_i\right]_{m\times m}$.

Proposition 2.4.1. *Let* $\mathbb{W} : \mathbb{C}^p \to \mathbb{MEP}(n_1, \cdots, n_m)$ *be a holomorphic map such that* $\mathbb{W}(t,z)$ *is regular for all* $t \in \mathbb{C}^p$*. Let* λ *be a geometrically simple eigenvalue of* $\mathbb{W}(t,z)$*. Then there exist left and right eigenvectors* $u := u_1 \otimes \cdots \otimes u_m$ *and* $v := v_1 \otimes \cdots \otimes v_m$ *corresponding to* λ*, respectively, such that* $J_{\mathbb{W}(t,\lambda)}(\lambda) = J_{\mathbb{W}(t,\lambda)}(t, \lambda; u, v)$*.*

Proof. Since λ is geometrically simple so $\mathrm{rank}(W_i(t,\lambda)) = n_i - 1$. Then by Proposition 1.1.1 there exist nonzero vectors $u_i, v_i \in \mathbb{C}^{n_i}$ such that $\mathrm{adj}(W_i(\lambda)) = u_i v_i^*$, where $W_i(\lambda)v_i = 0$ and $u_i^* W_i(\lambda) = 0$ for all $i = 1, \cdots, m$. The rest of the proof is similar to that of Proposition 2.2.1. □

Next we have the following result.

Theorem 2.4.1. *Let* $\mathbb{W} : \mathbb{C}^p \to \mathbb{MEP}(n_1, \cdots, n_m)$ *be a holomorphic map such that* $\mathbb{W}(t,z)$ *is regular for all* $t \in \mathbb{C}^p$*. Let* $\widehat{\lambda} \in \mathbb{C}^m$ *be a simple eigenvalue of* $\mathbb{W}(\widehat{t}, z)$ *for some* $\widehat{t} \in \mathbb{C}^p$*. Let* $x := x_1 \otimes \cdots \otimes x_m$ *and* $y := y_1 \otimes \cdots \otimes y_m$ *be right and left eigenvector of* $\mathbb{W}(\widehat{t}, z)$ *corresponding to the eigenvalue* $\widehat{\lambda}$*, respectively. Then there is an open neighborhood* $\mathrm{nbd}(\widehat{t}) \subset \mathbb{C}^p$ *containing* $\widehat{t}$ *and a holomorphic function* $\lambda : \mathrm{nbd}(\widehat{t}) \to \mathbb{C}^m$ *such that* $\lambda(\widehat{t}) = \widehat{\lambda}$ *and* $\lambda(t)$ *is a simple eigenvalue of* $\mathbb{W}(t,z)$ *for all* $t \in \mathrm{nbd}(\widehat{t})$*. Further for all* $t \in \mathrm{nbd}(\widehat{t})$*, we have* $\lambda(t+h) = \lambda(t) + D\lambda(t)h + \mathcal{O}(||h||_2^2)$*, for sufficiently small* $||h||_2$*, where* $D\lambda(t) : \mathbb{C}^p \to \mathbb{C}^m$ *is given by*

$$D\lambda(t)h = -\left(J_{\mathbb{W}(t,\lambda(t))}(\lambda(t))\right)^{-1} \begin{bmatrix} \sum_{j=1}^{p} \mathrm{tr}\left(\mathrm{adj}(W_1(t,\lambda(t)))\partial_{t_j}(W_1(t,\lambda(t)))\right) h_j \\ \vdots \\ \sum_{j=1}^{p} \mathrm{tr}\left(\mathrm{adj}(W_m(t,\lambda(t)))\partial_{t_j}(W_m(t,\lambda(t)))\right) h_j \end{bmatrix} \tag{2.20}$$

for all $h = (h_1, \cdots, h_p)^T \in \mathbb{C}^p$*,* $\partial_{t_j}(W_i(t,\lambda(t))$ *is the partial derivative of* $W_i(t,\lambda(t))$ *with respect to* t_j*. In particular, we have*

$$\frac{\partial \lambda(\widehat{t})}{\partial t_j} = -\left(J_{\mathbb{W}(\widehat{t},\widehat{\lambda})}(\widehat{\lambda})\right)^{-1} \begin{bmatrix} \mathrm{tr}\left(\mathrm{adj}(W_1(\widehat{t},\widehat{\lambda}))\partial_{t_1}(W_1(\widehat{t},\widehat{\lambda}))\right) \\ \vdots \\ \mathrm{tr}\left(\mathrm{adj}(W_m(\widehat{t},\widehat{\lambda}))\partial_{t_m}(W_1(\widehat{t},\widehat{\lambda}))\right) \end{bmatrix} \tag{2.21}$$

and

$$\frac{\partial\lambda(\widehat{t})}{\partial t_j} = -\left(J_{\mathbb{W}(t,\lambda)}\left(\widehat{t},\widehat{\lambda};y,x\right)\right)^{-1}\begin{bmatrix} y_1^*\partial_{t_1}(W_1(\widehat{t},\widehat{\lambda}))x_1 \\ \vdots \\ y_m^*\partial_{t_m}(W_1(\widehat{t},\widehat{\lambda}))x_m \end{bmatrix}. \tag{2.22}$$

Proof. Define $\phi : \mathbb{C}^p \times \mathbb{C}^m \to \mathbb{C}^m, (t,z) \mapsto \mathcal{F}(t,z)$. Note that $\phi(\widehat{t},\widehat{\lambda}) = 0$ and $\partial_\lambda(\phi(\widehat{t},\widehat{\lambda})) = J_{\mathbb{W}(\widehat{t},\widehat{\lambda})}(\widehat{\lambda})$ is nonsingular since $\widehat{\lambda}$ is a simple eigenvalue of $\mathbb{W}(\widehat{t},z)$. Hence by Implicit Function Theorem there exist an open neighborhood $\text{nbd}(\widehat{t}) \subset \mathbb{C}^p$ containing $\widehat{t}$ and a holomorphic function $\lambda : \text{nbd}(\widehat{t}) \to \mathbb{C}^m$ such that $\lambda(\widehat{t}) = \widehat{\lambda}$ and $\phi(t,\lambda(t)) = 0$, for all $t \in \text{nbd}(\widehat{t})$. Now by the similar arguments as those in the proof of Theorem 2.3.1 it follows that $\lambda(t)$ is a simple eigenvalue of $\mathbb{W}(t,z)$, for all $t \in \text{nbd}(\widehat{t})$. For the remaining part, differentiating $\phi(\lambda(t),t) = 0$ with respect to t we have

$$\partial_\lambda(\phi(\lambda(t),t))D\lambda(t)h + \partial_t(\phi(\lambda(t),t))h = 0 \tag{2.23}$$

for all $h \in \mathbb{C}^p$. Note that $\partial_\lambda(\phi(\lambda(t),t)) = J_{\mathbb{W}(t,\lambda(t))}(\lambda(t))$ is nonsingular and $\partial_t(\phi(\lambda(t),t)) : \mathbb{C}^p \to \mathbb{C}^m$ is given by

$$\begin{aligned} \partial_t(\phi(\lambda(t),t))h &= \begin{bmatrix} \text{tr}(\text{adj}(W_1(t,\lambda(t)))DW_1(t,\lambda(t))h) \\ \vdots \\ \text{tr}(\text{adj}(W_m(t,\lambda(t)))DW_m(t,\lambda(t))h) \end{bmatrix} \\ &= \begin{bmatrix} \sum_{j=1}^p \text{tr}\left(\text{adj}(W_1(t,\lambda(t)))\partial_{t_j}(W_1(t,\lambda(t)))\right)h_j \\ \vdots \\ \sum_{j=1}^p \text{tr}\left(\text{adj}(W_m(t,\lambda(t)))\partial_{t_j}(W_m(t,\lambda(t)))\right)h_j \end{bmatrix}, \end{aligned}$$

since $DW_i(t,\lambda(t))h = \sum_{j=1}^p \partial_{t_j}(W_i(t,\lambda(t)))h_j$ for all $i = 1,\cdots,m$. By (2.23), we have

$$\begin{aligned} D\lambda(t)h &= (\partial_\lambda(\phi(\lambda(t),t)))^{-1}\partial_t(\phi(\lambda(t),t))h \\ &= -\left(J_{\mathbb{W}(t,\lambda(t))}(\lambda(t))\right)^{-1}\begin{bmatrix} \sum_{j=1}^p \text{tr}\left(\text{adj}(W_1(t,\lambda(t)))\partial_{t_j}(W_1(t,\lambda(t)))\right)h_j \\ \vdots \\ \sum_{j=1}^p \text{tr}\left(\text{adj}(W_m(t,\lambda(t)))\partial_{t_j}(W_m(t,\lambda(t)))\right)h_j \end{bmatrix}, \end{aligned}$$

which proves (2.20). Finally (2.21) and (2.22) follow from (2.20). □

2.5 Backward perturbation analysis

In this section, we analyze backward errors of approximate eigenelements of an MEP $\mathbb{W}$.

Definition 2.5.1. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular. Let* $\lambda \in \mathbb{C}^m$ *and* $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$*. Then the backward error of* (λ, x) *as an approximate eigenpair of* $\mathbb{W}$ *is defined by*

$$\eta(\lambda, x, \mathbb{W}) = \inf \{ \|\!|\Delta\mathbb{W}|\!\|_V : \mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0 \}$$

and the backward error of λ *as an approximate eigenvalue of* $\mathbb{W}$ *is defined by*

$$\eta(\lambda, \mathbb{W}) = \inf \{ \|\!|\Delta\mathbb{W}|\!\|_V : \lambda \in \sigma(\mathbb{W} + \Delta\mathbb{W}) \}.$$

First we prove the following result.

Proposition 2.5.1. *Let* $W_k : \mathbb{C}^m \to \mathbb{C}^{n_k \times n_k}$ *be given by* $W_k(z) := A_k + \sum_{j=1}^m z_j B_{kj}$*, where* $A_k, B_{kj} \in \mathbb{C}^{n_k \times n_k}$ *for all* $j = 1, \cdots, m$*. Let* $\lambda \in \mathbb{C}^m$ *and* $0 \neq x_k \in \mathbb{C}^{n_k}$*. Define*

$$\eta(\lambda, x_k, W_k) := \inf \{ ||\Delta W_k||_V : W_k(\lambda)x_k + \Delta W_k(\lambda)x_k = 0 \}. \tag{2.24}$$

Then for a subordinate matrix norm on $\mathbb{C}^{n_k \times n_k}$*, we have* $\eta(\lambda, x_k, W_k) = \frac{1}{||(1,\lambda^*)||_{V,*}} \frac{||r_k||}{||x_k||}$*, where* $r_k = W_k(\lambda)x_k$*.*

Let $y_k \in \partial ||x_k||$ *and* $t := (t_0, \cdots, t_m)^T \in \partial ||(1, \lambda^*)||_{V,*}$*. Define*

$$\Delta A_k := -\frac{\overline{t_0}}{||(1, \lambda^*)||_{V,*}} \frac{r_k y_k^*}{||x_k||}, \quad \Delta B_{kj} := -\frac{\overline{t_j}}{||(1, \lambda^*)||_{V,*}} \frac{r_k y_k^*}{||x_k||}$$

for $j = 1, \cdots, m$ *and consider* $\Delta W_k(z) := \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$*. Then* $W_k(\lambda)x_k + \Delta W_k(\lambda)x_k = 0$ *and* $||\Delta W_k||_V = \eta(\lambda, x_k, W_k)$*.*

Proof. Let ΔX_k be such that $W_k(\lambda)x_k + \Delta X_k(\lambda)x_k = 0$. Then $||r_k|| = ||\Delta X_k(\lambda)x_k|| \leq ||\Delta X_k(\lambda)||\ ||x_k||$. By (2.8), we have $||\Delta X_k(\lambda)|| \leq ||\Delta X_k||_V\ ||(1, \lambda^*)||_{V,*}$. Thus $||r_k|| \leq ||\Delta X_k||_V\ ||(1, \lambda)||_{V,*}||x_k||$, i.e., $||\Delta X_k||_V \geq \frac{1}{||(1,\lambda^*)||_{V,*}} \frac{||r_k||}{||x_k||}$. Therefore $\eta(\lambda, x_k, W_k) \geq \frac{1}{||(1,\lambda^*)||_{V,*}} \frac{||r_k||}{||x_k||}$.

Again we have $\Delta W_k(\lambda)x_k = -\frac{\overline{t_0}+\sum_{j=1}^m \lambda_j \overline{t_j}}{||(1,\lambda^*)||_{V,*}} \frac{r_k y_k^* x_k}{||x_k||} = -r_k$, i.e., $W_k(\lambda)x_k + \Delta W_k(\lambda)x_k = 0$. Since $||r_k y_k^*|| = ||r_k||\ ||y_k||_* = ||r_k||$, so we have $||\Delta A_k|| = \frac{|t_0|}{||(1,\lambda^*)||_{V,*}} \frac{||r_k||}{||x_k||}$ and $||\Delta B_{kj}|| = \frac{|t_j|}{||(1,\lambda^*)||_{V,*}} \frac{||r_k||}{||x_k||}$ for all $j = 1, \cdots, m$. Therefore

$$\begin{aligned} ||\Delta W_k||_V &= \left|\left|\left(\frac{|t_0|}{||(1,\lambda^*)||_{V,*}} \frac{||r_k||}{||x_k||}, \frac{|t_1|}{||(1,\lambda^*)||_{V,*}} \frac{||r_k||}{||x_k||}, \cdots, \frac{|t_m|}{||(1,\lambda^*)||_{V,*}} \frac{||r_k||}{||x_k||} \right)\right|\right|_V \\ &= \frac{||t||_V}{||(1,\lambda^*)||_{V,*}} \frac{||r_k||}{||x_k||} = \frac{1}{||(1,\lambda^*)||_{V,*}} \frac{||r_k||}{||x_k||}. \end{aligned}$$

So $\eta(\lambda, x_k, W_k) = \frac{1}{||(1,\lambda^*)||_{V,*}} \frac{||r_k||}{||x_k||}$. This completes the proof. □

Now we use the above result to derive $\eta(\lambda, x, \mathbb{W})$.

Theorem 2.5.1. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular,* $\lambda \in \mathbb{C}^m$ *and* $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$*. Consider* $r_k := W_k(\lambda)x_k$ *for all* $k = 1, \cdots, m$*. Then for a subordinate matrix norm on* $\mathbb{C}^{n_k \times n_k}$*, we have*

$$\eta(\lambda, x, \mathbb{W}) = ||(\eta(\lambda, x_1, W_1), \cdots, \eta(\lambda, x_m, W_m))||_V = \frac{1}{||(1,\lambda^*)||_{V,*}} \left|\left|\left(\frac{||r_1||}{||x_1||}, \cdots, \frac{||r_m||}{||x_m||} \right)\right|\right|_V, \tag{2.25}$$

$\eta(\lambda, x_k, W_k)$ *is as defined in (2.24) for all* $k = 1, \cdots, m$*. In particular, for the Hölder p-norm on* $\mathbb{MEP}(n_1, \cdots, n_m)$*, we have* $\eta(\lambda, x, \mathbb{W}) = \frac{1}{||(1,\lambda^*)||_q} \left|\left|\left(\frac{||r_1||}{||x_1||}, \cdots, \frac{||r_m||}{||x_m||} \right)\right|\right|_p$*, where* $p^{-1} + q^{-1} = 1$*.*

Let $y_k \in \partial||x_k||$ *and* $t := (t_0, \cdots, t_m)^T \in \partial||(1,\lambda^*)||_{V,*}$*. Define*

$$\Delta A_k := -\frac{\overline{t_0}}{||(1,\lambda^*)||_{V,*}} \frac{r_k y_k^*}{||x_k||}, \ \Delta B_{kj} := -\frac{\overline{t_j}}{||(1,\lambda^*)||_{V,*}} \frac{r_k y_k^*}{||x_k||}$$

for $j = 1, \cdots, m$*. Consider* $\Delta W_k(z) := \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$ *for* $k = 1, \cdots, m$ *and* $\Delta\mathbb{W}(z) := (\Delta W_1(z), \cdots, \Delta W_m(z))$*. Then* $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ *and* $\|\Delta\mathbb{W}\|_V = \eta(\lambda, x, \mathbb{W})$*.*

Proof. Let $\Delta\mathbb{X} \in \mathbb{MEP}(n_1, \cdots, n_m)$ be such that $\Delta\mathbb{X}(z) = (\Delta X_1(z), \cdots, \Delta X_m(z))$ for $z \in \mathbb{C}^m$ and $W_k(\lambda)x_k + \Delta X_k(\lambda)x_k = 0$ for all $k = 1, \cdots, m$. Then it is easy to see that $||\Delta X_k||_V \geq \eta(\lambda, x_k, W_k)$, where $\eta(\lambda, x_k, W_k)$ is as defined in (2.24) for all $k = 1, \cdots, m$. Since $||\cdot||_V$ is a monotone norm, we have $\|\Delta\mathbb{X}\|_V \geq ||(\eta(\lambda, x_1, W_1), \cdots, \eta(\lambda, x_m, W_m))||_V$ and thus $\eta(\lambda, x, \mathbb{W}) \geq ||(\eta(\lambda, x_1, W_1), \cdots, \eta(\lambda, x_m, W_m))||_V$.

Again it follows from Proposition 2.5.1 that $W_k(\lambda)x_k + \Delta W_k(\lambda)x_k = 0$ and $||\Delta W_k||_V = \eta(\lambda, x_k, W_k)$ for all $k = 1, \cdots, m$. Then $|||\Delta \mathbb{W}|||_V = ||(\eta(\lambda, x_1, W_1), \cdots, \eta(\lambda, x_m, W_m))||_V$. So $\eta(\lambda, x, \mathbb{W}) = ||(\eta(\lambda, x_1, W_1), \cdots, \eta(\lambda, x_m, W_m))||_V$. Therefore (2.25) follows by Proposition 2.5.1. $\square$

Next we derive $\eta(\lambda, \mathbb{W})$.

Proposition 2.5.2. *Let* $W_k : \mathbb{C}^m \to \mathbb{C}^{n_k \times n_k}$ *be given by* $W_k(z) := A_k + \sum_{j=1}^m z_j B_{kj}$, *where* $A_k, B_{kj} \in \mathbb{C}^{n_k \times n_k}$ *for all* $j = 1, \cdots, m$. *Let* $\lambda \in \mathbb{C}^m$ *and define*

$$\eta(\lambda, W_k) = \inf \{||\Delta W_k||_V : \mathrm{rank}(W_k(\lambda) + \Delta W_k(\lambda)) < n_k\}. \tag{2.26}$$

Then for a subordinate matrix norm on $\mathbb{C}^{n_k \times n_k}$, *we have* $\eta(\lambda, W_k) = \frac{r_k}{||(1,\lambda^*)||_{V,*}}$, *where* $r_k = \min_{||x_k||=1} \{||W_k(\lambda)x_k|| : x_k \in \mathbb{C}^{n_k}\}$.

Let $x_k \in \mathbb{C}^{n_k}$ *be such that* $||x_k|| = 1$ *and* $r_k = ||W_k(\lambda)x_k||$. *Let* $y_k \in \partial||x_k||$ *and* $t := (t_0, \cdots, t_m)^T \in \partial||(1, \lambda^*)||_{V,*}$. *Define*

$$\Delta A_k := -\frac{\overline{t_0}}{||(1, \lambda^*)||_{V,*}} W_k(\lambda)x_k y_k^*, \ \Delta B_{kj} := -\frac{\overline{t_j}}{||(1, \lambda^*)||_{V,*}} W_k(\lambda)x_k y_k^*$$

for $j = 1, \cdots, m$ *and consider* $\Delta W_k(z) := \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$. *Then* $W_k(\lambda)x_k + \Delta W_k(\lambda)x_k = 0$ *and* $||\Delta W_k||_V = \eta(\lambda, W_k)$.

Proof. Let ΔX_k be such that $\mathrm{rank}(W_k(\lambda) + \Delta X_k(\lambda)) < n_k$. Then there exists $\widehat{x}_k \in \mathbb{C}^{n_k}$ such that $||\widehat{x}_k|| = 1$ and $W_k(\lambda)\widehat{x}_k + \Delta X_k(\lambda)\widehat{x}_k = 0$. So $||W_k(\lambda)\widehat{x}_k|| = ||\Delta X_k(\lambda)\widehat{x}_k|| \leq ||\Delta X_k(\lambda)|| \leq ||\Delta X_k||_V ||(1, \lambda^*)||_{V,*}$. This shows that $||\Delta X_k||_V \geq \frac{||W_k(\lambda)\widehat{x}_k||}{||(1,\lambda^*)||_{V,*}}$ and thus $\eta(\lambda, W_k) \geq \frac{r_k}{||(1,\lambda^*)||_{V,*}}$.

Next note that $\Delta W_k(\lambda)x_k = -\frac{\overline{t_0} + \sum_{j=1}^m \lambda_j \overline{t_j}}{||(1,\lambda^*)||_{V,*}} W_k(\lambda)x_k y_k^* x_k = -W_k(\lambda)x_k$, i.e., $(W_k(\lambda) + \Delta W_k(\lambda))x_k = 0$ and thus $\mathrm{rank}(W_k(\lambda) + \Delta W_k(\lambda)) < n_k$. Also we have $||W_k(\lambda)x_k y_k^*|| = ||W_k(\lambda)x_k|| \, ||y_k||_* = ||W_k(\lambda)x_k|| = r_k$. Consequently,

$$||\Delta A_k|| = \frac{|t_0| \ r_k}{||(1, \lambda^*)||_{V,*}} \text{ and } ||\Delta B_{kj}|| = \frac{|t_j| \ r_k}{||(1, \lambda^*)||_{V,*}}.$$

Then

$$\begin{aligned} ||\Delta W_k||_V &= \left|\left|\left(\frac{|t_0| \ r_k}{||(1, \lambda^*)||_{V,*}}, \frac{|t_1| \ r_k}{||(1, \lambda^*)||_{V,*}}, \cdots, \frac{|t_m| \ r_k}{||(1, \lambda^*)||_{V,*}}\right)\right|\right|_V \\ &= \frac{||t||_V \ r_k}{||(1, \lambda^*)||_{V,*}} = \frac{r_k}{||(1, \lambda^*)||_{V,*}}. \end{aligned}$$

Hence $\eta(\lambda, W_k) = \frac{r_k}{||(1,\lambda^*)||_{V,*}}$. This completes the proof. $\square$

We now derive $\eta(\lambda, \mathbb{W})$ considering a subordinate matrix norm on $\mathbb{C}^{n_i \times n_i}$.

Theorem 2.5.2. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular and* $\lambda \in \mathbb{C}^m$. *Consider* $r_k := \min_{||x_k||=1} \{||W_k(\lambda)x_k|| : x_k \in \mathbb{C}^{n_k}\}$ *for* $k = 1, \cdots, m$. *Then for a subordinate matrix norm on* $\mathbb{C}^{n_k \times n_k}$, *we have*

$$\eta(\lambda, \mathbb{W}) = ||(\eta(\lambda, W_1), \cdots, \eta(\lambda, W_m))||_V = \frac{1}{||(1, \lambda^*)||_{V,*}} \; ||(r_1, \cdots, r_m)||_V,$$

where $\eta(\lambda, W_k)$ *is as defined in (2.26) for all* $k = 1, \cdots, m$. *In particular, for the Hölder* p*-norm on* $\mathbb{MEP}(n_1, \cdots, n_m)$, $\eta(\lambda, \mathbb{W}) = \frac{1}{||(1,\lambda^*)||_q} \; ||(r_1, \cdots, r_m)||_p$, *where* $p^{-1} + q^{-1} = 1$.

Let $x_k \in \mathbb{C}^{n_k}$ *be such that* $||x_k|| = 1$ *and* $r_k = ||W_k(\lambda)x_k||$. *Let* $y_k \in \partial||x_k||$ *and* $t := (t_0, \cdots, t_m)^T \in \partial||(1, \lambda^*)||_{V,*}$. *Define*

$$\Delta A_k := -\frac{\overline{t_0}}{||(1, \lambda^*)||_{V,*}} W_k(\lambda)x_k y_k^*, \;\; \Delta B_{kj} := -\frac{\overline{t_j}}{||(1, \lambda^*)||_{V,*}} W_k(\lambda)x_k y_k^*$$

for $j = 1, \cdots, m$. *Consider* $\Delta W_k(z) := \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$ *for* $k = 1, \cdots, m$ *and* $\Delta\mathbb{W}(z) = (\Delta W_1(z), \cdots, \Delta W_m(z))$. *Then* $\lambda \in \sigma(\mathbb{W} + \Delta\mathbb{W})$ *and* $\|\Delta\mathbb{W}\|_V = \eta(\lambda, \mathbb{W})$.

Proof. Let $\Delta\mathbb{X} \in \mathbb{MEP}(n_1, \cdots, n_m)$ be such that $\Delta\mathbb{X}(z) = (\Delta X_1(z), \cdots, \Delta X_m(z))$ and $\operatorname{rank}(W_k + \Delta X_k) < n_k$ for all $k = 1, \cdots, m$. Then there exists $0 \neq \widehat{x}_k \in \mathbb{C}^{n_k}$ such that $W_k(\lambda)\widehat{x}_k + \Delta X_k(\lambda)\widehat{x}_k = 0$ for all $k = 1, \cdots, m$. So $\lambda \in \sigma(\mathbb{W} + \Delta\mathbb{X})$. It is easy to see that $||\Delta X_k||_V \geq \eta(\lambda, W_k)$ for all $k = 1, \cdots, m$. Then $\|\Delta\mathbb{X}\|_V \geq ||(\eta(\lambda, W_1), \cdots, \eta(\lambda, W_m))||_V$ as $|| \cdot ||_V$ is a monotone norm and thus $\eta(\lambda, \mathbb{W}) \geq ||(\eta(\lambda, W_1), \cdots, \eta(\lambda, W_m))||_V$.

By Proposition 2.5.2, we have $W_k(\lambda)x_k + \Delta W_k(\lambda)x_k = 0$ and $||\Delta W_k||_V = \frac{r_k}{||(1,\lambda^*)||_{V,*}}$ for all $k = 1, \cdots, m$. So $\lambda \in \sigma(\mathbb{W} + \Delta\mathbb{W})$ and $\|\Delta\mathbb{W}\|_V = ||(\eta(\lambda, W_1), \cdots, \eta(\lambda, W_m))||_V$. Thus $\eta(\lambda, \mathbb{W}) = ||(\eta(\lambda, W_1), \cdots, \eta(\lambda, W_m))||_V$. Hence by Proposition 2.5.2, we have $\eta(\lambda, \mathbb{W}) = \frac{1}{||(1,\lambda^*)||_{V,*}} \; ||(r_1, \cdots, r_m)||_V$. □

For the spectral or Frobenious norm on $\mathbb{C}^{n_i \times n_i}$, we have the following result.

Theorem 2.5.3. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular and* $\lambda \in \mathbb{C}^m$. *Then for the spectral or Frobenious norm on* $\mathbb{C}^{n_i \times n_i}$, *we have*

$$\eta(\lambda, \mathbb{W}) = \frac{1}{||(1, \lambda^*)||_{V,*}} \; ||(\sigma_{\min}(W_1(\lambda)), \cdots, \sigma_{\min}(W_m(\lambda)))||_V \,, \tag{2.27}$$

where $\sigma_{\min}(W_k(\lambda))$ *is the smallest singular value of* $W_k(\lambda)$ *for all* $k = 1, \cdots, m$. *In particular, for the Hölder* p*-norm on* $\mathbb{MEP}(n_1, \cdots, n_m)$ *we have*

$$\eta(\lambda, \mathbb{W}) = \frac{1}{||(1, \lambda^*)||_q} \; ||(\sigma_{\min}(W_1(\lambda)), \cdots, \sigma_{\min}(W_m(\lambda)))||_p \,,$$

where $p^{-1} + q^{-1} = 1$.

Let $t = (t_0, \cdots, t_m)^T \in \partial||(1, \lambda^*)||_{V,*}$. *Consider the* SVD $W_k(\lambda) = U_k \Sigma_k V_k$ *and set* $u_k := U_k(:, n_k)$ *and* $v_k := V_k(:, n_k)$ *for* $k = 1, \cdots, m$.

Define

$$\Delta A_k := -\frac{\overline{t_0}}{||(1, \lambda^*)||_{V,*}} \; \sigma_{\min}(W_k(\lambda)) u_k v_k^* \text{ and } \Delta B_{kj} := -\frac{\overline{t_j}}{||(1, \lambda^*)||_{V,*}} \; \sigma_{\min}(W_k(\lambda)) u_k v_k^*$$

for all $k, j = 1, \cdots, m$. *Consider* $\Delta W_k(z) := \Delta A_k + \sum_{j=1}^{m} z_j \Delta B_{kj}$ *for* $k = 1, \cdots, m$ *and* $\Delta\mathbb{W}(z) := (\Delta W_1(z), \cdots, \Delta W_m(z))$. *Then* $\lambda \in \sigma(\mathbb{W} + \Delta\mathbb{W})$ *and* $\|\Delta\mathbb{W}\|_V = \eta(\lambda, \mathbb{W})$.

Proof. In this case, $\min\limits_{||x_k||=1} ||W_k(\lambda)x_k||_2 = \sigma_{\min}(W_k(\lambda)) = ||W_k(\lambda)v_k||_2$ for all $k = 1, \cdots, m$ and thus (2.27) follows by Theorem 2.5.2. Further, note that

$$\Delta W_k(\lambda) v_k = -\frac{\overline{t_0} + \sum_{j=1}^{m} \lambda_j \overline{t_j}}{||(1, \lambda^*)||_{V,*}} \; \sigma_{\min}(W_k(\lambda)) u_k = -\sigma_{\min}(W_k(\lambda)) u_k = -W_k(\lambda) v_k,$$

i.e., $(W_k(\lambda) + \Delta W_k(\lambda))v_k = 0$ and thus $\text{rank}(W_k(\lambda) + \Delta W_k(\lambda)) < n_k$. Since $||u_k v_k^*||_2 = ||u_k v_k^*||_F = ||u_k||_2 ||v_k||_2 = 1$, so we have

$$\begin{aligned} ||\Delta A_k||_2 = ||\Delta A_k||_F &= \frac{|t_0|}{||(1, \lambda^*)||_{V,*}} \sigma_{\min}(W_k(\lambda)), \\ ||\Delta B_{kj}||_2 = ||\Delta B_{kj}||_F &= \frac{|t_j|}{||(1, \lambda^*)||_{V,*}} \sigma_{\min}(W_k(\lambda)) \end{aligned}$$

for all $j = 1, \cdots, m$. Then $||\Delta W_k||_V = \frac{||t||_V}{||(1,\lambda^*)||_{V,*}} \; \sigma_{\min}(W_k(\lambda)) = \frac{1}{||(1,\lambda^*)||_{V,*}} \; \sigma_{\min}(W_k(\lambda))$. Hence $\|\Delta\mathbb{W}\|_V = \frac{1}{||(1,\lambda^*)||_{V,*}} \; ||(\sigma_{\min}(W_1(\lambda)), \cdots, \sigma_{\min}(W_m(\lambda)))||_V = \eta(\lambda, \mathbb{W})$. This completes the proof. ☐

3

Polynomial MEP

3.1 Introduction

Polynomial MEP arises in many applications [36, 43]. For example, the delay differential equation (DDE) $\sum_{k=0}^{m} B_k \dot{x}(t-\tau k) = \sum_{k=0}^{m} A_k x(t-\tau k)$ with m delays leads to the following two-parameter polynomial eigenvalue problem:

$$\begin{aligned} A_0 x &= \lambda \sum_{k=0}^{m} \mu^k B_k x - \sum_{k=1}^{m} \mu^k A_k x \\ -A_m^* y &= \lambda \sum_{k=0}^{m} B_{m-k}^* y + \sum_{k=1}^{m} \mu^k A_{m-k}^* y. \end{aligned}$$

Numerical methods to solve a nonhomogeneous polynomial MEP have been proposed in [47]. In this chapter, we undertake sensitivity and backward perturbation analysis of a two-parameter polynomial eigenvalue problem of degree k. Let $\mathbb{W}$ be a two-parameter polynomial eigenvalue problem of degree k. We develop a framework for sensitivity and backward perturbation analysis of $\mathbb{W}$. For a simple eigenvalue λ, we define the condition number $\mathrm{cond}(\lambda, \mathbb{W})$ by

$$\mathrm{cond}(\lambda, \mathbb{W}) := \limsup_{\|\Delta \mathbb{W}\| \to 0} \frac{\mathrm{dist}(\lambda, \sigma(\mathbb{W} + \Delta \mathbb{W}))}{\|\Delta \mathbb{W}\|}$$

and derive various expressions for $\mathrm{cond}(\lambda, \mathbb{W})$. Given $\lambda \in \mathbb{C}^2$ and $0 \neq x := x_1 \otimes x_2$, we determine the backward error $\eta(\lambda, x, \mathbb{W})$ of (λ, x) as an approximate eigenpair of $\mathbb{W}$. We also derive an optimal perturbation $\Delta \mathbb{W}$ such that $\mathbb{W}(\lambda)x + \Delta \mathbb{W}(\lambda)x = 0$ and $\|\Delta \mathbb{W}\|_V = \eta(\lambda, x, \mathbb{W})$. Further we compute similar results for an approximate eigenpair λ of $\mathbb{W}$.

3.2 Polynomial MEP

Consider $\mathbb{N}_0 := \mathbb{N} \bigcup \{0\}$, where $\mathbb{N}$ denotes the set of natural numbers. For $\alpha := (\alpha_1, \cdots, \alpha_m) \in \mathbb{N}_0^m$, define $|\alpha| := \alpha_1 + \cdots + \alpha_m$. For $z = (z_1, \cdots, z_m)^T \in \mathbb{C}^m$, define $z^\alpha := \prod_{j=1}^m z_j^{\alpha_j} \in \mathbb{C}$. Let $\mathbb{W} : \mathbb{C}^m \to \prod_{j=1}^m \mathbb{C}^{n_j \times n_j}$ be given by

$$\mathbb{W}(z) = (W_1(z), \cdots, W_m(z)), \text{ where } W_i(z) := \sum_{|\alpha| \leq k} A_\alpha^{(i)} z^\alpha, \tag{3.1}$$

$z \in \mathbb{C}^m$, $\alpha \in \mathbb{N}_0^m$, $A_\alpha^{(i)} \in \mathbb{C}^{n_i \times n_i}$ for all $i = 1, \cdots, m$ and $k > 1$ is a positive integer. Then we refer to $\mathbb{W}$ as a polynomial MEP of degree k. We denote the space of polynomial MEPs of degree k by $\mathbb{PMEP}(k; n_1, \cdots, n_m)$. In particular, when $m = 2$, $\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$ is given by

$$\mathbb{W}(z) = (W_1(z), W_2(z)), \text{ where } W_i(z) := \sum_{m=0}^{k} \sum_{n=0}^{k-m} A_{mn}^{(i)} z_1^m z_2^n, \tag{3.2}$$

$z = (z_1, z_2) \in \mathbb{C}^2$ and $A_{mn}^{(i)} \in \mathbb{C}^{n_i \times n_i}$ for $i = 1, 2, m = 0, \cdots, k, n = 0, \cdots, k - m$. We refer to $\mathbb{W}$ as a two-parameter polynomial eigenvalue problem.

Given $\lambda := (\lambda_1, \lambda_2) \in \mathbb{C}^2$, define

$$\Lambda_j(\lambda_1, \lambda_2) := \begin{bmatrix} \lambda_1^j \\ \lambda^{j-1}\lambda_2 \\ \vdots \\ \lambda_1 \lambda_2^{j-1} \\ \lambda_2^j \end{bmatrix} \in \mathbb{C}^{j+1} \text{ for } j = 0, \cdots, k \text{ and } \Lambda := \begin{bmatrix} \Lambda_k(\lambda_1, \lambda_2) \\ \vdots \\ \Lambda_0(\lambda_1, \lambda_2) \end{bmatrix} \in \mathbb{C}^{\frac{(k+1)(k+2)}{2}}. \tag{3.3}$$

In particular when $k = 2$, we have $\Lambda = (\lambda_1^2, \lambda_1\lambda_2, \lambda_2^2, \lambda_1, \lambda_2, 1)^T$.

Definition 3.2.1. *Let* $\mathbb{W} \in \mathbb{PMEP}(k; n_1, \cdots, n_m)$ *be given by* $\mathbb{W}(z) = (W_1(z), \cdots, W_m(z))$. *Then* $\mathbb{W}$ *is said to be regular if* $\prod_{j=1}^m \det(W_j(z)) \neq 0$ *for some* $z \in \mathbb{C}^m$.

Throughout this chapter we consider regular polynomial MEPs. For $x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ and $z \in \mathbb{C}^m$, we define

$$\mathbb{W}(z)x := (W_1(z)x_1, \ldots, W_m(z)x_m) \text{ and } x^*\mathbb{W}(z) := (x_1^* W_1(z), \ldots, x_m^* W_m(z)).$$

Theorem 3.2.1. (BézoutsTheorem[24]) *Let* $f(x,y) = g(x,y) = 0$ *be a system of two polynomial equations in two unknowns. If it has only finitely many common complex zeros* $(x,y) \in \mathbb{C} \times \mathbb{C}$*, then the number of those zeros is at most* $\text{degree}(f) \cdot \text{degree}(g)$*.*

Next we define the spectrum of a regular polynomial MEP. Recall that the **det** function is defined in (2.2).

Definition 3.2.2. (Spectrum) *Let* $\mathbb{W} \in \mathbb{PMEP}(k; n_1, \cdots, n_m)$ *be regular. Then* $\lambda \in \mathbb{C}^m$ *is said to be an eigenvalue of* $\mathbb{W}$ *if* $\mathbf{det}(\mathbb{W}(\lambda)) = 0$*. The set of all eigenvalues of* $\mathbb{W}$ *is called the spectrum of* $\mathbb{W}$ *and is denoted by* $\sigma(\mathbb{W})$*. Let* $\lambda \in \sigma(\mathbb{W})$*. Then* $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ *is called a right eigenvector of* $\mathbb{W}$ *corresponding to* λ *if* $\mathbb{W}(\lambda)x = 0$*. Similarly* $0 \neq y := y_1 \otimes \cdots \otimes y_2 \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ *is called a left eigenvector of* $\mathbb{W}$ *corresponding to* λ *if* $y^*\mathbb{W}(\lambda) = 0$*.*

Next we define geometric multiplicity of an eigenvalue of a regular polynomial MEP of degree k.

Definition 3.2.3. (Geometric multiplicity) *Let* $\mathbb{W} \in \mathbb{PMEP}(k; n_1, \cdots, n_m)$ *given by* $\mathbb{W}(z) = (W_1(z), \cdots, W_m(z))$ *be regular and* $\lambda \in \sigma(\mathbb{W})$*. Then the geometric multiplicity of* λ *is defined by* $g_{\mathbb{W}}(\lambda) = \prod_{i=1}^{m} \dim \ker(W_i(\lambda))$*. If* $g_{\mathbb{W}}(\lambda) = 1$ *then* λ *is said to be geometrically simple.*

Let $\mathbb{W} : \mathbb{C}^m \to \prod_{j=1}^{m} \mathbb{C}^{n_j \times n_j}$ given by $\mathbb{W}(z) = (W_1(z), \cdots, W_m(z))$ be differentiable, where $W_i : \mathbb{C}^m \to \mathbb{C}^{n_i \times n_i}$ is a nonlinear function for $i = 1, \cdots, m$. Then $\lambda \in \mathbb{C}^m$ is said to be an eigenvalue of $\mathbb{W}$ if $\mathbf{det}(\mathbb{W}(\lambda)) = 0$. The set of all eigenvalues of $\mathbb{W}$ is called the spectrum of $\mathbb{W}$ and is denoted by $\sigma(\mathbb{W})$. Then $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ is called a right eigenvector of $\mathbb{W}$ corresponding to λ if $\mathbb{W}(\lambda)x = 0$. Similarly $0 \neq y := y_1 \otimes \cdots \otimes y_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ is called a left eigenvector of $\mathbb{W}$ corresponding to λ if $y^*\mathbb{W}(\lambda) = 0$. An eigenvalue $\lambda \in \sigma(\mathbb{W})$ is said to be geometrically simple if $\prod_{i=1}^{m} \dim \ker(W_i(\lambda)) = 1$. Let $\lambda \in \sigma(\mathbb{W})$ and $x = x_1 \otimes \cdots \otimes x_m$ and $y = y_1 \otimes \cdots \otimes y_m$ be right and left eigenvector of $\mathbb{W}$ corresponding to the eigenvalue λ, respectively. Then we define

$$J_{\mathbb{W}}(\lambda; y, x) := \begin{bmatrix} y_1^* \frac{\partial}{\partial \lambda_1}(W_1(\lambda))x_1 & \cdots & y_1^* \frac{\partial}{\partial \lambda_m}(W_1(\lambda))x_1 \\ & \vdots & \\ y_m^* \frac{\partial}{\partial \lambda_1}(W_m(\lambda))x_m & \cdots & y_m^* \frac{\partial}{\partial \lambda_m}(W_m(\lambda))x_m \end{bmatrix}. \tag{3.4}$$

We denote the Jacobian of the map $\mathbb{C}^m \to \mathbb{C}^m, z \mapsto \mathbf{det}(\mathbb{W}(z))$ by $J_{\mathbb{W}}(z)$. Then we have the following result.

Proposition 3.2.1. *Let* $\mathbb{W} : \mathbb{C}^m \to \prod_{j=1}^m \mathbb{C}^{n_j \times n_j}$ *given by* $\mathbb{W}(z) = (W_1(z), \cdots, W_m(z))$ *be differentiable, where* $W_i : \mathbb{C}^m \to \mathbb{C}^{n_i \times n_i}$ *is a nonlinear function for* $i = 1, \cdots, m$. *Then*

$$J_{\mathbb{W}}(z) = \begin{bmatrix} \mathrm{tr}\left(\mathrm{adj}(W_1(z))\frac{\partial}{\partial z_1}(W_1(z))\right) & \cdots & \mathrm{tr}\left(\mathrm{adj}(W_1(z))\frac{\partial}{\partial z_m}(W_1(z))\right) \\ & \vdots & \\ \mathrm{tr}\left(\mathrm{adj}(W_m(z))\frac{\partial}{\partial z_1}(W_m(z))\right) & \cdots & \mathrm{tr}\left(\mathrm{adj}(W_m(z))\frac{\partial}{\partial z_m}(W_m(z))\right) \end{bmatrix}. \tag{3.5}$$

In particular, if $\lambda = (\lambda_1, \cdots, \lambda_m)^T \in \sigma(\mathbb{W})$ *is geometrically simple then there exist left and right eigenvectors* $u := u_1 \otimes \cdots \otimes u_m$ *and* $v := v_1 \otimes \cdots \otimes v_m$ *of* $\mathbb{W}$ *corresponding to* λ, *respectively, such that*

$$J_{\mathbb{W}}(\lambda) = \begin{bmatrix} u_1^* \frac{\partial}{\partial \lambda_1}(W_1(\lambda))v_1 & \cdots & u_1^* \frac{\partial}{\partial \lambda_m}(W_1(\lambda))v_1 \\ & \vdots & \\ u_m^* \frac{\partial}{\partial \lambda_1}(W_m(\lambda))v_m & \cdots & u_m^* \frac{\partial}{\partial \lambda_m}(W_m(\lambda))v_m \end{bmatrix} = J_{\mathbb{W}}(\lambda; u, v). \tag{3.6}$$

Proof. Note that $J_{\mathbb{W}}(z) = \left[\frac{\partial}{\partial z_j}(\det(W_i(z)))\right]_{m \times m}$. By Jacobi's formula given in Proposition 1.1.2, it follows that $\frac{\partial}{\partial z_j}(\det(W_i(z))) = \mathrm{tr}\left(\mathrm{adj}(W_i(z))\frac{\partial}{\partial z_j}(W_i(z))\right)$ for all $i, j = 1, \cdots, m$. This proves (3.5).

If $\lambda \in \sigma(\mathbb{W})$ is geometrically simple then $\mathrm{rank}(W_i(\lambda)) = n_i - 1$ for all $i = 1, \cdots, m$. So by Proposition 1.1.1, there exist nonzero vectors $u_i, v_i \in \mathbb{C}^{n_i}$ such that $\mathrm{adj}(W_i(\lambda)) = v_i u_i^*$, where $W_i(\lambda)v_i = 0$ and $u_i^* W_i(\lambda) = 0$ for all $i = 1, \cdots, m$. Hence the desired result follows. ☐

Definition 3.2.4. (Algebraically simple eigenvalue) *Let* $\mathbb{W} \in \mathbb{PMEP}(k; n_1, \cdots, n_m)$ *be regular and* $\lambda \in \sigma(\mathbb{W})$. *Then* λ *is said to be algebraically simple if* $\mathrm{rank}(J_{\mathbb{W}}(\lambda)) = m$. *If* $\mathrm{rank}(J_{\mathbb{W}}(\lambda)) < m$ *then* λ *is said to be a multiple eigenvalue of* $\mathbb{W}$.

Throughout this chapter we refer to algebraically simple eigenvalue as simple eigenvalue.

Proposition 3.2.2. *Let* $\mathbb{W} \in \mathbb{PMEP}(k; n_1, \cdots, n_m)$ *be regular and* $\lambda \in \sigma(\mathbb{W})$ *be simple. Then* λ *is geometrically simple.*

Proof. If $\dim \ker(W_i(\lambda)) > 1$ for some $i \in \{1, \cdots, m\}$ then the i^{th} row of $J_{\mathbb{W}}(\lambda)$ is a zero row as we have seen in the proof of Proposition 2.2.2. But this is not possible since λ is simple. Hence $\dim \ker(W_i(\lambda)) = 1$ for all $i = 1, \cdots, m$ and $g_{\mathbb{W}}(\lambda) = \prod_{i=1}^{m} \dim \ker(W_i(\lambda)) = 1$, i.e., λ is geometrically simple. ☐

The following result characterizes a multiple eigenvalue.

Theorem 3.2.2. *Let* $\mathbb{W} \in \mathbb{PMEP}(k; n_1, \cdots, n_m)$ *be regular and* $\lambda \in \sigma(\mathbb{W})$*. Then* λ *is multiple if and only if there exist left and right eigenvectors* y *and* x *of* $\mathbb{W}$ *corresponding to* λ*, respectively, such that* $J_{\mathbb{W}}(\lambda; y, x)$ *is singular.*

Proof. Let $\lambda \in \sigma(\mathbb{W})$ be multiple. Then by similar arguments as those in Theorem 2.2.1, it follows that there exist left and right eigenvectors y and x of $\mathbb{W}$ corresponding to λ, respectively, such that $J_{\mathbb{W}}(\lambda; y, x)$ is singular.

Conversely, suppose that there exist left and right eigenvectors y and x of $\mathbb{W}$ corresponding to λ, respectively, such that $J_{\mathbb{W}}(\lambda; y, x)$ is singular. If λ is not geometrically simple then $\dim \ker(W_i(\lambda)) > 1$ for some $k \in \{1, \cdots, m\}$ and $\operatorname{adj}(W_k(\lambda)) = 0$. Thus it follows from (3.5) that the k^{th} row of $J_{\mathbb{W}}(\lambda)$ is a zero row and $\operatorname{rank}(J_{\mathbb{W}}(\lambda)) < m$. Therefore λ is multiple.

If λ is geometrically simple then by Proposition 3.2.1, there exist left and right eigenvectors u and v of $\mathbb{W}$ corresponding to λ, respectively, such that $J_{\mathbb{W}}(\lambda) = J_{\mathbb{W}}(\lambda; u, v)$. By similar arguments as those in the proof of Theorem 2.2.1, it follows that there exist nonzero scalars α_i, β_i for all $i = 1, \cdots, m$ such that

$$J_{\mathbb{W}}(\lambda; u, v) = \begin{bmatrix} \overline{\alpha}_1\beta_1 & & \\ & \ddots & \\ & & \overline{\alpha}_m\beta_m \end{bmatrix} J_{\mathbb{W}}(\lambda; y, x).$$

Hence the desired conclusion follows. ☐

3.2.1 Norms on $\mathbb{PMEP}(k; n_1, n_2)$

For perturbation analysis of a regular two-parameter polynomial eigenvalue problem of degree k, we define norms, inner product and dual norms on $\mathbb{PMEP}(k; n_1, n_2)$ as follows.

Norm. Let $\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$ be as given in (3.2). Consider a matrix norm $||\cdot||$ on $\mathbb{C}^{n_i \times n_i}$ and a monotone vector norm $||\cdot||_V$ on $\mathbb{C}^d$. Define $||W_i||_V := ||(M_k^{(i)}, \cdots, M_0^{(i)})||_V$, where $M_j^{(i)} = (||A_{j,0}^{(i)}||, ||A_{j-1,1}^{(i)}||, \cdots, ||A_{0,j}^{(i)}||) \in \mathbb{C}^{j+1}$ for $i = 1, 2$ and $j = 0, \cdots, k$. Then

$$|||\mathbb{W}|||_V = ||(||W_1||_V, ||W_2||_V)||_V. \tag{3.7}$$

defines a norm on $\mathbb{PMEP}(k; n_1, n_2)$. In particular, replacing the monotone norm $||\cdot||_V$ by the usual Hölder p-norm $||\cdot||_p$ for $1 \leq p \leq \infty$, we define $|||\cdot|||_p$ on $\mathbb{PMEP}(k; n_1, n_2)$ by $|||\mathbb{W}|||_p := ||(||W_1||_p, ||W_2||_p)||_p$, where $||W_i||_p = ||(M_k^{(i)}, \cdots, M_0^{(i)})||_p$ for $i = 1, 2$. We refer to $|||\cdot|||_p$ as the Hölder p-norm on $\mathbb{PMEP}(k; n_1, n_2)$.

Next we define weighted norms on $\mathbb{PMEP}(k; n_1, n_2)$. Consider $\mathbb{R}_+^2 := \{(x_1, x_2)^T \in \mathbb{R}^2 : x_i > 0\}$ and define the map $\mathbb{R}_+^2 \times \mathbb{PMEP}(k; n_1, n_2) \to \mathbb{PMEP}(k; n_1, n_2), (w, \mathbb{W}) \mapsto w \odot \mathbb{W}$, by $(w \odot \mathbb{W})(z) := (w_1 W_1(z), w_2 W_2(z))$, where $\mathbb{W}(z) = (W_1(z), W_2(z))$ and $w = (w_1, w_2)^T \in \mathbb{R}_+^2$. Given a norm $|||\cdot|||$ on $\mathbb{PMEP}(k; n_1, n_2)$, we define the weighted norm $|||\cdot|||_w$ by $|||\mathbb{W}|||_w := |||w \odot \mathbb{W}|||$. For the norm $|||\cdot|||_V$ given in (3.7), $|||\mathbb{W}|||_{w,V} := ||(w_1||W_1||_V, w_2||W_2||_V)||_V$ defines a weighted norm on $\mathbb{PMEP}(k; n_1, n_2)$. The weighted Hölder p-norm on $\mathbb{PMEP}(k; n_1, n_2)$ is given by $|||\mathbb{W}|||_{w,p} := ||(w_1||W_1||_p, w_2||W_2||_p)||_p$ for $1 \leq p \leq \infty$.

Inner Product. Next we define inner product on $\mathbb{PMEP}(k; n_1, n_2)$. Let $\mathbb{X}, \mathbb{Y} \in \mathbb{PMEP}(k; n_1, n_2)$ be such that $\mathbb{X}(z) = (X_1(z), X_2(z))$ and $\mathbb{Y}(z) = (Y_1(z), Y_2(z))$, where $X_i(z) = \sum_{m=0}^{k} \sum_{n=0}^{k-m} A_{mn}^{(i)} z_1^m z_2^n$ and $Y_i(z) = \sum_{m=0}^{k} \sum_{n=0}^{k-m} B_{mn}^{(i)} z_1^m z_2^n$ for $i = 1, 2$. Consider the usual inner product $\langle \cdot, \cdot \rangle$ on the space of matrices, i.e., $\langle A, B \rangle := \mathrm{tr}(B^* A)$. Define $\langle X_i, Y_i \rangle := \sum_{m=0}^{k} \sum_{n=0}^{k-m} \left\langle A_{mn}^{(i)}, B_{mn}^{(i)} \right\rangle$ for $i = 1, 2$. Then

$$\langle \mathbb{X}, \mathbb{Y} \rangle := \langle X_1, Y_1 \rangle + \langle X_2, Y_2 \rangle = \sum_{i=1}^{2} \sum_{m=0}^{k} \sum_{n=0}^{k-m} \langle A_{mn}^{(i)}, B_{mn}^{(i)} \rangle \tag{3.8}$$

defines an inner product on $\mathbb{PMEP}(k; n_1, n_2)$.

Dual Norm. Given a norm $|||\cdot|||$ on $\mathbb{PMEP}(k; n_1, n_2)$, we define the dual norm $|||\cdot|||_*$

relative to the inner product $\langle\cdot,\cdot\rangle$ given in (3.8) by $\|\mathbb{W}\|_* = \sup\{|\langle\mathbb{W},\mathbb{X}\rangle| : \|\mathbb{X}\| = 1\}$. It follows that $|\langle\mathbb{X},\mathbb{Y}\rangle| \le \|\mathbb{X}\|\|\mathbb{Y}\|_*$ for all $\mathbb{X},\mathbb{Y} \in \mathbb{PMEP}(k; n_1, n_2)$.

Remark 3.2.1. ***(Compatibility of norms)*** *Let $\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$ be as given in (3.2). Consider a monotone vector norm $||\cdot||_V$ and a matrix norm $||\cdot||$ on $\mathbb{C}^{n_i\times n_i}$. Then*

$$||W_i(\lambda)|| = \left|\left|\sum_{m=0}^{k}\sum_{n=0}^{k-m} A_{mn}^{(i)}\lambda_1^m\lambda_2^n\right|\right| \;\le\; \sum_{m=0}^{k}\sum_{n=0}^{k-m} ||A_{mn}^{(i)}||\;|\lambda_1^m\lambda_2^n| = \left|\left\langle (M_k^{(i)},\cdots,M_0^{(i)}),\Lambda\right\rangle\right|$$
$$\le\; ||(M_k^{(i)},\cdots,M_0^{(i)})||_V||\Lambda||_{V,*} = ||W_i||_V||\Lambda||_{V,*}, \quad (3.9)$$

where Λ is as given in (3.3).

3.3 Sensitivity analysis of a simple eigenvalue

Now we analyze the sensitivity of a simple eigenvalue of a regular polynomial MEP.

Definition 3.3.1. (Condition number) *Let $\mathbb{W} \in \mathbb{PMEP}(k; n_1, \cdots, n_m)$ be regular and $\lambda \in \sigma(\mathbb{W})$. Then we define the* ***condition number*** *of λ by*

$$\mathrm{cond}(\lambda,\mathbb{W}) := \limsup_{\|\Delta\mathbb{W}\|\to 0} \frac{\mathrm{dist}(\lambda,\sigma(\mathbb{W}+\Delta\mathbb{W}))}{\|\Delta\mathbb{W}\|},$$

where $\Delta\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$ and $\mathrm{dist}(\lambda,\sigma(\mathbb{W}+\Delta\mathbb{W})) := \min\{||\lambda-\mu||_2 : \mu \in \sigma(\mathbb{W}+\Delta\mathbb{W})\}$.

We show that given a simple eigenvalue $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$, there exists an open neighborhood $\mathrm{nbd}(\mathbb{W})$ containing $\mathbb{W}$ and a smooth function $\lambda : \mathrm{nbd}(\mathbb{W}) \to \mathbb{C}^m$ such that $\lambda(\mathbb{W}) = \lambda_{\mathbb{W}}$ and $\lambda(\mathbb{X}) \in \sigma(\mathbb{X})$ is a simple eigenvalue for all $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$.

Theorem 3.3.1. *Let $\mathbb{W} \in \mathbb{PMEP}(k; n_1, \cdots, n_m)$ be regular and $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ be a simple eigenvalue. Then there is an open neighborhood $\mathrm{nbd}(\mathbb{W}) \subset \mathbb{PMEP}(k; n_1, \cdots, n_m)$ and a smooth function $\lambda : \mathrm{nbd}(\mathbb{W}) \to \mathbb{C}^m$ such that $\lambda(\mathbb{W}) = \lambda_{\mathbb{W}}$ and $\lambda(\mathbb{X})$ is a simple eigenvalue of $\mathbb{X}$ for all $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$. Further, for $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$ such that $\mathbb{X}(z) = (X_1(z), \cdots, X_m(z))$, the derivative $D\lambda(\mathbb{X}) : \mathbb{PMEP}(k; n_1, \cdots, n_m) \to \mathbb{C}^m$ of λ at $\mathbb{X}$ is given by*

$$D\lambda(\mathbb{X})\mathbb{H} = -(J_{\mathbb{X}}(\lambda(\mathbb{X})))^{-1}\begin{bmatrix} \mathrm{tr}(\mathrm{adj}(X_1(\lambda(\mathbb{X})))H_1(\lambda(\mathbb{X}))) \\ \vdots \\ \mathrm{tr}(\mathrm{adj}(X_m(\lambda(\mathbb{X})))H_m(\lambda(\mathbb{X}))) \end{bmatrix} \quad (3.10)$$

for all $\mathbb{H} \in \mathbb{PMEP}(k; n_1, \cdots, n_m)$ *such that* $\mathbb{H}(z) = (H_1(z), \cdots, H_2(z))$.

Proof. Define $\phi : \mathbb{PMEP}(k; n_1, \cdots, n_m) \times \mathbb{C}^m \to \mathbb{C}^m$ given by $\phi(\mathbb{Y}, z) = \mathbf{det}(\mathbb{Y}(z))$. Then by the similar arguments as those in the proof of Theorem 2.3.1, it follows that there is an open neighborhood $\mathrm{nbd}(\mathbb{W}) \subset \mathbb{PMEP}(k; n_1, \cdots, n_m)$ and a smooth function $\lambda : \mathrm{nbd}(\mathbb{W}) \to \mathbb{C}^m$ such that $\lambda(\mathbb{W}) = \lambda_{\mathbb{W}}$ and $\lambda(\mathbb{X})$ is a simple eigenvalue of $\mathbb{X}$ for all $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$. Finally differentiating $\phi(\mathbb{X}, \lambda(\mathbb{X})) = 0$ with respect to $\mathbb{X}$, we obtain (3.10) as we have done in the proof of Theorem 2.3.1. □

Now we determine $\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W})$ when $\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$ and $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ is simple.

Theorem 3.3.2. *Let* $\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$ *be regular and* $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ *be simple. Let* $x := x_1 \otimes x_2$ *and* $y := y_1 \otimes y_2$ *be right and left eigenvectors of* $\mathbb{W}$ *corresponding to* $\lambda_{\mathbb{W}}$*, respectively. Then there is an open neighborhood* $\mathrm{nbd}(\mathbb{W}) \subset \mathbb{PMEP}(k; n_1, n_2)$ *containing* $\mathbb{W}$ *and a smooth function* $\lambda : \mathrm{nbd}(\mathbb{W}) \to \mathbb{C}^2$ *such that* $\lambda(\mathbb{W}) = \lambda_{\mathbb{W}}$ *and* $\lambda(\mathbb{X})$ *is a simple eigenvalue of* $\mathbb{X}$ *for all* $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$*. Further, the derivative* $D\lambda(\mathbb{W}) : \mathbb{PMEP}(k; n_1, n_2) \to \mathbb{C}^2$ *is given by*

$$D\lambda(\mathbb{W})\mathbb{H} = -(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1} \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(\lambda_{\mathbb{W}}))H_1(\lambda_{\mathbb{W}})) \\ \mathrm{tr}(\mathrm{adj}(W_2(\lambda_{\mathbb{W}}))H_2(\lambda_{\mathbb{W}})) \end{bmatrix} \tag{3.11}$$

$$= -(J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1} \begin{bmatrix} y_1^* H_1(\lambda_{\mathbb{W}}) x_1 \\ y_2^* H_2(\lambda_{\mathbb{W}}) x_2 \end{bmatrix} \tag{3.12}$$

for all $\mathbb{H} \in \mathbb{PMEP}(k; n_1, n_2)$ *such that* $\mathbb{H}(z) = (H_1(z), H_2(z))$.

a) The condition number $\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W})$ *is given by*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \; ||\Lambda_{\mathbb{W}}||_{V,*} \max_i ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_*,$$

where $\Lambda_{\mathbb{W}}$ *be the column vector as defined in (3.3) corresponding to* $\lambda_{\mathbb{W}}$.

In particular, for the spectral or Frobenius norm on $\mathbb{C}^{n_i \times n_i}$*, we have*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \; ||\Lambda_{\mathbb{W}}||_{V,*} \max_i \prod_{k=1}^{n_i - 1} \sigma_k(W_i(\lambda_{\mathbb{W}})),$$

where $\sigma_1(W_i(\lambda_{\mathbb{W}})) \geq \cdots \geq \sigma_{n_i}(W_i(\lambda_{\mathbb{W}}))$ *are the singular values of* $W_i(\lambda_{\mathbb{W}})$ *for* $i = 1, 2$.

b) Also we have

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_V \; ||\Lambda_{\mathbb{W}}||_{V,*} \; \max_i ||y_i x_i^*||_*.$$

For a subordinate matrix norm $||\cdot||$ *on* $\mathbb{C}^{n_i \times n_i}$*, we have*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_V \; ||\Lambda_{\mathbb{W}}||_{V,*} \; \max_i ||y_i||_* ||x_i||.$$

In particular, for the spectral or Frobenius matrix norm on $\mathbb{C}^{n_i \times n_i}$*, we have*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_V \; ||\Lambda_{\mathbb{W}}||_{V,*} \; \max_i ||y_i||_2 ||x_i||_2.$$

Proof. By Theorem 3.3.1, it follows that there is an open neighborhood $\mathrm{nbd}(\mathbb{W}) \subset \mathbb{PMEP}(k; n_1, n_2)$ containing $\mathbb{W}$ and a smooth function $\lambda : \mathrm{nbd}(\mathbb{W}) \to \mathbb{C}^2$ such that $\lambda(\mathbb{W}) = \lambda_{\mathbb{W}}$ and $\lambda(\mathbb{X})$ is a simple eigenvalue of $\mathbb{X}$ for all $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$. Further, (3.11) follows from (3.10). By Theorem 2.3.2, it follows that there exist nonzero scalar $\alpha_1, \alpha_2, \beta_1, \beta_2$ such that

$$J_{\mathbb{W}}(\lambda_{\mathbb{W}}) = \begin{bmatrix} \overline{\alpha}_1 \beta_1 & \\ & \overline{\alpha}_2 \beta_2 \end{bmatrix} J_{\mathbb{W}}(\lambda; y, x)$$

and $\mathrm{tr}(\mathrm{adj}(W_i(\lambda_{\mathbb{W}}))H_i(\lambda_{\mathbb{W}})) = \overline{\alpha}_i \beta_i y_i^* H_i(\lambda_{\mathbb{W}}) x_i$ for $i = 1, 2$. Then by (3.11), we obtain (3.12). The rest of the proof is similar to that of Theorem 2.3.2. □

When $\mathbb{PMEP}(k; n_1, n_2)$ is equipped with the Hölder p-norm $|||\cdot|||_p$ for $1 \leq p \leq \infty$, we have the following result.

Corollary 3.3.3. *Let* $\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$ *be regular and* $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ *be a simple eigenvalue of* $\mathbb{W}$*. Let* $x := x_1 \otimes x_2$ *and* $y := y_1 \otimes y_2$ *be right and left eigenvectors of* $\mathbb{W}$ *corresponding to* $\lambda_{\mathbb{W}}$*, respectively.*

a) Then we have

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_p \; ||\Lambda_{\mathbb{W}}||_q \; \max_i ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_*,$$

where $p^{-1} + q^{-1} = 1$*. In particular, for the spectral or Frobenius norm on* $\mathbb{C}^{n_i \times n_i}$*, we have*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_p \; ||\Lambda_{\mathbb{W}}||_q \; \max_i \prod_{k=1}^{n_i - 1} \sigma_k(W_i(\lambda_{\mathbb{W}})),$$

where $\sigma_1(W_i(\lambda_{\mathbb{W}})) \geq \cdots \geq \sigma_{n_i}(W_i(\lambda_{\mathbb{W}}))$ *are the singular values of* $W_i(\lambda_{\mathbb{W}})$ *for* $i = 1, 2.$

b) Also we have

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_p \; ||\Lambda_{\mathbb{W}}||_q \; \max_i ||y_i x_i^*||_*. \tag{3.13}$$

For a subordinate matrix norm $||\cdot||$ *on* $\mathbb{C}^{n_i \times n_i}$*, we have*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_p \; ||\Lambda_{\mathbb{W}}||_q \; \max_i ||y_i||_* ||x_i||. \tag{3.14}$$

In particular, for the spectral or Frobenius matrix norm on $\mathbb{C}^{n_i \times n_i}$*, we have*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_p \; ||\Lambda_{\mathbb{W}}||_q \; \max_i ||y_i||_2 ||x_i||_2. \tag{3.15}$$

Remark 3.3.1. *The regular two-parameter nonhomogeneous polynomial MEP of degree* k *reduces to a regular matrix polynomial* $\mathbb{W}(z) := \sum_{j=0}^{k} z^j A_j, z \in \mathbb{C}, A_j \in \mathbb{C}^{n_1 \times n_1}$*, of degree* k *when* $i = 1$*. In this case, we have* $\Lambda_{\mathbb{W}} = (\lambda_{\mathbb{W}}^k, \lambda_{\mathbb{W}}^{k-1}, \cdots, \lambda_{\mathbb{W}}, 1)^T$ *and*

$$J_{\mathbb{W}}(\lambda_{\mathbb{W}}) = \mathrm{tr}\left(\mathrm{adj}(\mathbb{W}(\lambda_{\mathbb{W}}))\frac{\partial}{\partial \lambda}(\mathbb{W}(\lambda_{\mathbb{W}}))\right) \neq 0 \Rightarrow (J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1} = \frac{1}{\mathrm{tr}\left(\mathrm{adj}(\mathbb{W}(\lambda_{\mathbb{W}}))\frac{\partial}{\partial \lambda}(\mathbb{W}(\lambda_{\mathbb{W}}))\right)}.$$

It follows from Theorem 3.3.2 that

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq \frac{||\Lambda_{\mathbb{W}}||_{V,*}}{\left|\mathrm{tr}(\mathrm{adj}(\mathbb{W}(\lambda_{\mathbb{W}})) \cdot \frac{\partial}{\partial \lambda}(\mathbb{W}(\lambda_{\mathbb{W}})))\right|} ||(\mathrm{adj}(\mathbb{W}(\lambda_{\mathbb{W}})))^*||_*. \tag{3.16}$$

It is shown in [7] that the equality holds in (3.16). So we consider

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) = ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \; ||\Lambda_{\mathbb{W}}||_{V,*} \; \max_i ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_*$$

to be the condition number of a simple eigenvalue $\lambda_{\mathbb{W}}$ *of a regular polynomial MEP* $\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$*. Further if* $x := x_1 \otimes x_2$ *and* $y := y_1 \otimes y_2$ *be right and left eigenvectors of* $\mathbb{W}$ *corresponding to* $\lambda_{\mathbb{W}}$*, respectively, then we have*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) = ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_V \; ||\Lambda_{\mathbb{W}}||_{V,*} \; \max_i ||y_i x_i^*||_*.$$

When $\mathbb{PMEP}(k; n_1, n_2)$ is equipped with the weighted norm $|||\cdot|||_{w,V}$ with nonzero positive weights, we have the following result.

Theorem 3.3.4. *Let* $\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$ *be regular and* $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ *be simple. Consider the weighted norm* $|||\cdot|||_{w,V}$ *on* $\mathbb{PMEP}(k; n_1, n_2)$ *with nonzero positive weights* $w_i := ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_*$ *for* $i = 1, 2$. *Then* $\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) = ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \; ||\Lambda_{\mathbb{W}}||_{V,*}$. *In particular, for the weighted Hölder p-norm* $|||\cdot|||_{w,p}$ *on* $\mathbb{PMEP}(k; n_1, n_2)$ *with nonzero positive weights* $w_i := ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_*$, *we have* $\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) = ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_p \; ||\Lambda_{\mathbb{W}}||_q$, *where* $p^{-1} + q^{-1} = 1$.

Proof. By (3.9), we have

$$|\mathrm{tr}(\mathrm{adj}(W_i(\lambda_{\mathbb{W}}))H_i(\lambda_{\mathbb{W}}))| \leq ||H_i(\lambda_{\mathbb{W}})|| \; ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_* \leq w_i ||\Lambda_{\mathbb{W}}||_{V,*} ||H_i||_V$$

for $i = 1, 2$. Then by (3.11), we have $||D\lambda(\mathbb{W})\mathbb{H}||_V \leq ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V ||\Lambda_{\mathbb{W}}||_{V,*} |||\mathbb{H}|||_{w,V}$. Now to show the equality, consider $\mathbb{Y} \in \mathbb{PMEP}(k; n_1, n_2)$ such that $\mathbb{Y}(z) = (Y_1(z), Y_2(z))$, $||Y_i||_V = 1$ and $\langle Y_i, \Lambda_{\mathbb{W}}^* \otimes \mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^* \rangle = w_i ||\Lambda_{\mathbb{W}}^*||_{V,*}$, where $\Lambda_{\mathbb{W}}^* \otimes \mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^* : \mathbb{C}^2 \to \mathbb{C}^{n_i \times n_i}$ is given by $(\Lambda_{\mathbb{W}}^* \otimes \mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^*)(z) = (\Lambda_{\mathbb{W}}^* \Lambda_z)\mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^*$ for $i = 1, 2$ and $\Lambda_{\mathbb{W}}, \Lambda_z \in \mathbb{C}^{\frac{(k+1)(k+2)}{2}}$ are given by (3.3) corresponding to $\lambda_{\mathbb{W}}, z$ respectively. Let $u = (u_1, u_2)^T \in \mathbb{C}^2$ be such that $||u||_V = 1$ and $||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}u||_V = ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V$. Construct $\mathbb{X} \in \mathbb{MEP}(n_1, \cdots, n_m)$ such that $\mathbb{X}(z) := \left(\frac{u_1}{w_1}Y_1(z), \frac{u_2}{w_2}Y_2(z)\right)$. By similar arguments as those in the proof of Theorem 2.3.4, it follows that $|||\mathbb{X}|||_{w,V} = 1$ and $||D\lambda(\mathbb{W})\mathbb{X}||_V = ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \; ||\Lambda_{\mathbb{W}}^*||_{V,*}$. Hence $||D\lambda(\mathbb{W})||_V = ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \; ||\Lambda_{\mathbb{W}}||_{V,*}$ and the desired result follows. ☐

The fast perturbation for a simple eigenvalue is defined as follows.

Definition 3.3.2. *Let* $\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$ *be regular and* $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ *be simple. Then* $\Delta\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$ *is said to be a fast perturbation for* $\lambda_{\mathbb{W}}$ *if* $||D\lambda(\mathbb{W})\Delta\mathbb{W}||_V = ||D\lambda(\mathbb{W})|| |||\Delta\mathbb{W}|||_V$.

Considering a weighted norm $|||\cdot|||_{w,V}$ on $\mathbb{PMEP}(k; n_1, n_2)$, we construct a fast perturbation for a simple eigenvalue $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$.

Proposition 3.3.1. (Fast Perturbation) *Let* $\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$ *be regular and* $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ *be simple. Consider a weighted norm* $|||\cdot|||_{w,V}$ *on* $\mathbb{PMEP}(k; n_1, n_2)$ *with nonzero positive weights* $w_i := ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_*$ *for* $i = 1, 2$. *Let* $y = (y_1, y_2)^T \in \mathbb{C}^2$ *be such that*

$||y||_V = 1$ *and* $||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}\; y||_V = ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V$*. Let* $\mathbb{Y} \in \mathbb{PMEP}(k; n_1, n_2)$ *be such that* $\mathbb{Y}(z) := (Y_1(z), Y_2(z))$*,* $||Y_i||_V = 1$ *and* $\langle Y_i, \Lambda_{\mathbb{W}}^* \otimes \mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^*\rangle = w_i||\Lambda_{\mathbb{W}}||_{V,*}$ *where* $\Lambda_{\mathbb{W}}^* \otimes \mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^* : \mathbb{C}^2 \to \mathbb{C}^{n_i \times n_i}$ *is given by* $(\Lambda_{\mathbb{W}}^* \otimes \mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^*)(z) = (\Lambda_{\mathbb{W}}^*\Lambda_z)\mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^*$ *for* $i = 1, 2$ *and* $\Lambda_{\mathbb{W}}, \Lambda_z \in \mathbb{C}^{\frac{(k+1)(k+2)}{2}}$ *are given by (3.3) corresponding to* $\lambda_{\mathbb{W}}, z$ *respectively. Define* $\Delta\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$ *by* $\Delta\mathbb{W}(z) := \left(\frac{y_1}{w_1}Y_1(z), \frac{y_2}{w_2}Y_2(z)\right)$*. Then* $\Delta\mathbb{W}$ *is a fast perturbation for* $\lambda_{\mathbb{W}}$*, i.e.,* $||D\lambda(\mathbb{W})\Delta\mathbb{W}||_V = ||D\lambda(\mathbb{W})|| |||\Delta\mathbb{W}|||_V$*.*

Proof. First note that $|||\Delta\mathbb{W}|||_{w,V} = \left|\left|\left(w_1\frac{|y_1|}{w_1}||Y_1||_V, w_2\frac{|y_2|}{w_2}||Y_2||_V\right)\right|\right|_V = ||y||_V = 1$. Also by (3.11), we have $D\lambda(\mathbb{W})\Delta\mathbb{W} = -(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}\begin{bmatrix} ||\Lambda_{\mathbb{W}}||_{V,*}y_1 \\ ||\Lambda_{\mathbb{W}}||_{V,*}y_2 \end{bmatrix} = -||\Lambda_{\mathbb{W}}||_{V,*}(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}y$. So $||D\lambda(\mathbb{W})\Delta\mathbb{W}||_V = ||\Lambda_{\mathbb{W}}||_{V,*}||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}y||_V = ||(J_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V||\Lambda_{\mathbb{W}}||_{V,*}$. □

3.4 Holomorphic perturbations

Next we consider the effect of holomorphic perturbations on a simple eigenvalue of a regular nonhomogeneous polynomial MEP. Let $\mathbb{W} : \mathbb{C}^p \to \mathbb{PMEP}(k; n_1, \cdots, n_m)$ be a holomorphic map. Then $\mathbb{W}(t) \in \mathbb{PMEP}(k; n_1, \cdots, n_m)$ for $t \in \mathbb{C}^p$. We write $\mathbb{W}(t)(z)$ as $\mathbb{W}(t, z) := (W_1(t, z), \cdots, W_m(t, z))$. Then $\mathbb{W}(t, z)$ is said to be regular if for each $t \in \mathbb{C}^p$ there is a $z \in \mathbb{C}^m$ such that $\prod_{i=1}^m \det(W_i(t, z)) \neq 0$. Define $\mathcal{F}(t, z) := (\det(W_1(t, z)), \cdots, \det(W_m(t, z)))^T$. Recall that $\lambda \in \mathbb{C}^m$ is said to be an eigenvalue of $\mathbb{W}(t, z)$ if $\mathcal{F}(t, \lambda) = 0$. We refer to $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ and $0 \neq y := y_1 \otimes \cdots \otimes y_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ as right and left eigenvectors of $\mathbb{W}(t, z)$ corresponding to an eigenvalue λ, respectively, if $\mathbb{W}(t, \lambda)x = 0 = y^*\mathbb{W}(t, \lambda)$. An eigenvalue λ of $\mathbb{W}(t, z)$ is said to be geometrically simple if $\prod_{j=1}^m \dim \ker(W_j(t, \lambda)) = 1$. An eigenvalue λ of $\mathbb{W}(t, z)$ is said to be simple if the Jacobian matrix $J_{\mathbb{W}(t,\lambda)}(\lambda)$ of the function $\mathbb{C}^m \to \mathbb{C}^m, z \mapsto \mathcal{F}(t, z)$ at λ, given by $J_{\mathbb{W}(t,\lambda)}(\lambda) := \left[\mathrm{tr}(\mathrm{adj}(W_i(t, \lambda))\partial_{\lambda_j}(W_1(t, \lambda(t))))\right]_{m\times m}$ is nonsingular, where $\partial_{\lambda_j}(W_i(t, \lambda))$ is the partial derivative of $W_i(t, \lambda)$ with respect to λ_j for all $i, j = 1, \cdots, m$. For the left and right eigenvectors $y := y_1 \otimes \cdots \otimes y_m$ and $x := x_1 \otimes \cdots \otimes x_m$ of $\mathbb{W}(t, z)$ corresponding to λ, respectively, define $J_{\mathbb{W}(t,\lambda)}(t, \lambda; y, x) := \left[y_i^*\partial_{\lambda_j}(W_i(t, \lambda))x_i\right]_{m\times m}$.

Proposition 3.4.1. *Let* $\mathbb{W} : \mathbb{C}^p \to \mathbb{PMEP}(k; n_1, \cdots, n_m)$ *be holomorphic such that* $\mathbb{W}(t, z) \in \mathbb{PMEP}(k; n_1, \cdots, n_m)$ *is regular for* $t \in \mathbb{C}^p$*. Let* λ *be a geometrically simple* *eigenvalue of* $\mathbb{W}(t,z)$*. Then there exist left and right eigenvectors* $u := u_1\otimes\cdots\otimes u_m$ *and* $v := v_1\otimes\cdots\otimes v_m$ *corresponding to* λ*, respectively, such that* $J_{\mathbb{W}(t,\lambda)}(\lambda) = J_{\mathbb{W}(t,\lambda)}(t,\lambda;u,v)$*.*

Proof. Since λ is geometrically simple so $\operatorname{rank}(W_i(t,\lambda)) = n_i - 1$. Then by Proposition 1.1.1, there exist nonzero vectors $u_i, v_i \in \mathbb{C}^{n_i}$ such that $\operatorname{adj}(W_i(\lambda)) = v_i u_i^*$, where $W_i(\lambda)v_i = 0$ and $u_i^* W_i(\lambda) = 0$ for all $i = 1,\cdots,m$. Then

$$\operatorname{tr}(\operatorname{adj}(W_i(t,\lambda))\partial_{\lambda_j}(W_1(t,\lambda(t)))) = y_i^*\partial_{\lambda_j}(W_i(t,\lambda))x_i$$

for $i,j = 1,\cdots,m$, which gives the desired result. $\square$

Next we have the following result.

Theorem 3.4.1. *Let* $\mathbb{W} : \mathbb{C}^p \to \mathbb{PMEP}(k;n_1,\cdots,n_m)$ *be holomorphic such that* $\mathbb{W}(t,z)$ *is regular for* $t \in \mathbb{C}^p$*. Let* $\widehat{\lambda} \in \mathbb{C}^m$ *be a simple eigenvalue of* $\mathbb{W}(\widehat{t},z)$*. Let* $x_1\otimes\cdots\otimes x_m$ *and* $y_1\otimes\cdots\otimes y_m$ *be right and left eigenvectors of* $\mathbb{W}(\widehat{t},z)$ *corresponding to the eigenvalue* $\widehat{\lambda}$*, respectively. Then there is an open neighborhood* $\operatorname{nbd}(\widehat{t}) \subset \mathbb{C}^p$ *containing* $\widehat{t}$ *and a holomorphic function* $\lambda : \operatorname{nbd}(\widehat{t}) \to \mathbb{C}^m$ *such that* $\lambda(\widehat{t}) = \widehat{\lambda}$ *and* $\lambda(t)$ *is a simple eigenvalue of* $\mathbb{W}(t,z)$ *for all* $t \in \operatorname{nbd}(\widehat{t})$*. Further for* $t \in \operatorname{nbd}(\widehat{t})$*, we have* $\lambda(t+h) = \lambda(t) + D\lambda(t)h + \mathcal{O}(||h||_2^2)$ *for sufficiently small* $||h||_2$ *and the derivative* $D\lambda(t) : \mathbb{C}^p \to \mathbb{C}^m$ *is given by*

$$D\lambda(t)h = -\left(J_{\mathbb{W}(t,\lambda(t))}(\lambda(t))\right)^{-1}\begin{bmatrix} \sum_{j=1}^{p}\operatorname{tr}\left(\operatorname{adj}(W_1(t,\lambda(t)))\partial_{t_j}(W_1(t,\lambda(t)))\right)h_j \\ \vdots \\ \sum_{j=1}^{p}\operatorname{tr}\left(\operatorname{adj}(W_m(t,\lambda(t)))\partial_{t_j}(W_m(t,\lambda(t)))\right)h_j \end{bmatrix} \tag{3.17}$$

for all $h = (h_1,\cdots,h_p)^T \in \mathbb{C}^p$*, where* $\partial_{t_j}(W_i(t,\lambda(t))$ *is the partial derivative of* $W_i(t,\lambda(t))$ *with respect to* t_j *for* $i = 1,\cdots,m$ *and* $j = 1,\cdots,p$*. In particular, we have*

$$\frac{\partial\lambda(\widehat{t})}{\partial t_j} = -\left(J_{\mathbb{W}(\widehat{t},\widehat{\lambda})}(\widehat{\lambda})\right)^{-1}\begin{bmatrix} \operatorname{tr}\left(\operatorname{adj}(W_1(\widehat{t},\widehat{\lambda}))\partial_{t_j}(W_1(\widehat{t},\widehat{\lambda}))\right) \\ \vdots \\ \operatorname{tr}\left(\operatorname{adj}(W_m(\widehat{t},\widehat{\lambda}))\partial_{t_j}(W_1(\widehat{t},\widehat{\lambda}))\right) \end{bmatrix} \tag{3.18}$$

$$= -\left(J_{\mathbb{W}(\widehat{t},\widehat{\lambda})}(\widehat{t},\widehat{\lambda};y,x)\right)^{-1}\begin{bmatrix} y_1^*\partial_{t_j}(W_1(\widehat{t},\widehat{\lambda}))x_1 \\ \vdots \\ y_m^*\partial_{t_j}(W_m(\widehat{t},\widehat{\lambda}))x_m \end{bmatrix}. \tag{3.19}$$

Proof. Define $\phi : \mathbb{C}^p \times \mathbb{C}^m \to \mathbb{C}^m, (t, z) \mapsto \mathbf{det}(\mathbb{W}(t, z))$. Then by the similar arguments as those given in the proof of Theorem 2.4.1, it follows that there is an open neighborhood $\text{nbd}(\widehat{t}) \subset \mathbb{C}^p$ containing $\widehat{t}$ and a holomorphic function $\lambda : \text{nbd}(\widehat{t}) \to \mathbb{C}^m$ such that $\lambda(\widehat{t}) = \widehat{\lambda}$ and $\lambda(t)$ is a simple eigenvalue of $\mathbb{W}(t, z)$ for all $t \in \text{nbd}(\widehat{t})$. Further, differentiating $\phi(\lambda(t), t) = 0$ with respect to t, we obtain (3.17). It is easy to see that (3.18) follows from (3.17).

Since $\widehat{\lambda} \in \mathbb{C}^m$ is a simple eigenvalue of $\mathbb{W}(\widehat{t}, z)$, so $\widehat{\lambda}$ is geometrically simple, i.e., $\dim \ker(W_j(\widehat{t}, \widehat{\lambda})) = 1$ for all $j = 1, \cdots, m$. Then there exist left and right eigenvectors $u := u_1 \otimes \cdots \otimes u_m$ and $v := v_1 \otimes \cdots \otimes v_m$ corresponding to $\widehat{\lambda}$, respectively, such that $\text{adj}(W_j(\widehat{t}, \widehat{\lambda})) = v_j u_j^*$ and $J_{\mathbb{W}(\widehat{t},\widehat{\lambda})}(\widehat{\lambda}) = J_{\mathbb{W}(\widehat{t},\widehat{\lambda})}(\widehat{t}, \widehat{\lambda}; u, v) = [u_i^* \partial_{\lambda_j}(W_i(\widehat{t}, \widehat{\lambda}))v_i]_{m\times m}$. Further there exist nonzero scalars α_j, β_j such that $u_j = \alpha_j y_j$ and $v_j = \beta_j x_j$. Now it is easy to see that $u_i^* \partial_{\lambda_j}(W_i(\widehat{t}, \widehat{\lambda}))v_i = \overline{\alpha}_i \beta_i y_i^* \partial_{\lambda_j}(W_i(\widehat{t}, \widehat{\lambda}))x_i$ and thus

$$J_{\mathbb{W}(\widehat{t},\widehat{\lambda})}(\widehat{t}, \widehat{\lambda}; u, v) = \begin{bmatrix} \overline{\alpha}_1\beta_1 & & \\ & \ddots & \\ & & \overline{\alpha}_m\beta_m \end{bmatrix} J_{\mathbb{W}(\widehat{t},\widehat{\lambda})}(\widehat{t}, \widehat{\lambda}; y, x).$$

Also we have $\text{tr}\left(\text{adj}(W_i(\widehat{t}, \widehat{\lambda}))\partial_{t_j}(W_1(\widehat{t}, \widehat{\lambda}))\right) = u_i^* \partial_{t_j}(W_m(\widehat{t}, \widehat{\lambda}))v_i = \overline{\alpha}_i \beta_i y_i^* \partial_{t_j}(W_m(\widehat{t}, \widehat{\lambda}))x_i$ for all $i, j = 1, \cdots, m$. Hence by (3.18) we have

$$\frac{\partial\lambda(\widehat{t})}{\partial t_j} = -\left(J_{\mathbb{W}(\widehat{t},\widehat{\lambda})}(\widehat{t}, \widehat{\lambda}; y, x)\right)^{-1} \begin{bmatrix} y_1^* \partial_{t_j}(W_1(\widehat{t}, \widehat{\lambda}))x_1 \\ \vdots \\ y_m^* \partial_{t_j}(W_m(\widehat{t}, \widehat{\lambda}))x_m \end{bmatrix},$$

which yields (3.19). $\square$

3.5 Backward perturbation analysis

Now we undertake backward perturbation analysis of approximate eigenelements of a regular two-parameter polynomial eigenvalue problem of degree k.

Definition 3.5.1. *Let* $\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$ *be regular. Let* $\lambda \in \mathbb{C}^2$ *and* $0 \neq x := x_1 \otimes x_2 \in \mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2}$*. Then the backward error of* (λ, x) *as an approximate eigenpair of* $\mathbb{W}$ *is defined by* $\eta(\lambda, x, \mathbb{W}) := \inf \{\|\Delta\mathbb{W}\|_V : \mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0\}$*. The backward* *error of* λ *as an approximate eigenvalue of* $\mathbb{W}$ *is defined by* $\eta(\lambda, \mathbb{W}) := \inf\{\|\Delta\mathbb{W}\|_V : \lambda \in \sigma(\mathbb{W} + \Delta\mathbb{W})\}$.

First we prove the following result.

Proposition 3.5.1. *Let* $W_i : \mathbb{C}^2 \to \mathbb{C}^{n_i \times n_i}$ *be given by* $W_i(z) := \sum_{m=0}^{k}\sum_{n=0}^{k-m} A_{mn}^{(i)} z_1^m z_2^n$, *where* $z = (z_1, z_2)^T \in \mathbb{C}^2$, $A_{mn}^{(i)} \in \mathbb{C}^{n_i \times n_i}$. *Let* $\lambda = (\lambda_1, \lambda_2)^T \in \mathbb{C}^2$ *and* $0 \neq x_i \in \mathbb{C}^{n_i}$. *Define* $\eta(\lambda, x_i, W_i) := \inf\{||\Delta W_i||_V : W_i(\lambda)x_i + \Delta W_i(\lambda)x_i = 0\}$. *Then for a subordinate matrix norm on* $\mathbb{C}^{n_i\times n_i}$, *we have* $\eta(\lambda, x_i, W_i) = \frac{1}{||\Lambda||_{V,*}}\frac{||r_i||}{||x_i||}$, *where* $r_i = W_i(\lambda)x_i$.

Let $t := (t_{k,0}, t_{k-1,1}, \cdots, t_{0,k}, t_{k-1,0}, t_{k-2,1}, \cdots, t_{0,k-1}, \cdots, t_{1,0}, t_{0,1}, t_{0,0})^T \in \mathbb{C}^{\frac{(k+1)(k+2)}{2}}$ *be such that* $t \in \partial||\Lambda||_{V,*}$ *and* $y_i \in \partial||x_i||$. *Define* $\Delta A_{mn}^{(i)} := -\frac{\overline{t_{m,n}}}{||\Lambda||_{V,*}}\frac{r_i y_i^*}{||x_i||}$ *for* $m = 0, \cdots, k$ *and* $n = 0, \cdots, k-m$. *Consider* $\Delta W_i(z) := \sum_{m=0}^{k}\sum_{n=0}^{k-m}\Delta A_{mn}^{(i)} z_1^m z_2^n$. *Then* $W_i(\lambda)x_i + \Delta W_i(\lambda)x_i = 0$ *and* $||\Delta W_i||_V = \eta(\lambda, x_i, W_i)$.

Proof. By similar arguments as those in the proof of Proposition 2.5.1, we have $\eta(\lambda, x_i, W_i) \geq \frac{1}{||\Lambda||_{V,*}}\frac{||r_i||}{||x_i||}$. Again note that, $\Delta W_i(\lambda)x_i = -\sum_{m=0}^{k}\sum_{n=0}^{k-m}\frac{\overline{t_{m,n}}\;\lambda_1^m\lambda_2^n}{||\Lambda||_{V,*}}\frac{r_i y_i^* x}{||x_i||} = -\frac{\langle\Lambda, t\rangle}{||\Lambda||_{V,*}}r_i = -r_i$. So $W_i(\lambda)x_i + \Delta W_i(\lambda)x_i = 0$. Also we have $||\Delta A_{mn}^{(i)}|| = \frac{|t_{m,n}|}{||\Lambda||_{V,*}}\frac{||r_i||}{||x_i||}$ and thus $||\Delta W_i||_V = \frac{||t||_V}{||\Lambda||_{V,*}}\frac{||r_i||}{||x_i||} = \frac{1}{||\Lambda||_{V,*}}\frac{||r_i||}{||x_i||}$. Therefore $\eta(\lambda, x_i, W_i) = \frac{1}{||\Lambda||_{V,*}}\frac{||r_i||}{||x_i||}$. □

Next we use the above result to derive $\eta(\lambda, x, \mathbb{W})$.

Theorem 3.5.1. *Let* $\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$ *be regular,* $\lambda = (\lambda_1, \lambda_2)^T \in \mathbb{C}^2$ *and* $0 \neq x := x_1 \otimes x_2 \in \mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2}$. *Consider* $r_i = W_i(\lambda)x_i$ *for* $i = 1, 2$. *Then for a subordinate matrix norm on* $\mathbb{C}^{n_i\times n_i}$, *we have*

$$\eta(\lambda, x, \mathbb{W}) = ||(\eta(\lambda, x_1, W_1), \eta(\lambda, x_2, W_2))||_V = \frac{1}{||\Lambda||_{V,*}}\left|\left|\left(\frac{||r_1||}{||x_1||}, \frac{||r_2||}{||x_2||}\right)\right|\right|_V,$$

where $\eta(\lambda, x_i, W_i)$ *is as defined in Proposition 3.5.1. In particular, for the Hölder p-norm on* $\mathbb{PMEP}(k; n_1, n_2)$, *we have* $\eta(\lambda, x, \mathbb{W}) = \frac{1}{||\Lambda||_q}\left|\left|\left(\frac{||r_1||}{||x_1||}, \frac{||r_2||}{||x_2||}\right)\right|\right|_p$, *where* $p^{-1} + q^{-1} = 1$.

Let $t := (t_{k,0}, t_{k-1,1}, \cdots, t_{0,k}, t_{k-1,0}, t_{k-2,1}, \cdots, t_{0,k-1}, \cdots, t_{1,0}, t_{0,1}, t_{0,0})^T \in \mathbb{C}^{\frac{(k+1)(k+2)}{2}}$ *be such that* $t \in \partial||\Lambda||_{V,*}$ *and* $y_i \in \partial||x_i||$ *for* $i = 1, 2$. *Define* $\Delta A_{mn}^{(i)} := -\frac{\overline{t_{m,n}}}{||\Lambda||_{V,*}}\frac{r_i y_i^*}{||x_i||}$ *for* $m = 0, \cdots, k, n = 0, \cdots, k-m$ *and* $\Delta W_i(z) := \sum_{m=0}^{k}\sum_{n=0}^{k-m}\Delta A_{mn}^{(i)} z_1^m z_2^n$. *Consider* $\Delta\mathbb{W}(z) := (\Delta W_1(z), \Delta W_2(z))$. *Then* $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ *and* $\|\Delta\mathbb{W}\|_V = \eta(\lambda, x, \mathbb{W})$.

Proof. It follows from the similar arguments as those in the proof of Theorem 2.5.1 that $\eta(\lambda, x, \mathbb{W}) \geq ||(\eta(\lambda, x_1, W_1), \eta(\lambda, x_2, W_2))||_V$. By Proposition 3.5.1, we have $W_i(\lambda)x_i +$

$\Delta W_i(\lambda)x_i = 0$ and $||\Delta W_i||_V = \eta(\lambda, x_i, W_i)$. So $|||\Delta\mathbb{W}|||_V = ||(\eta(\lambda, x_1, W_1), \eta(\lambda, x_2, W_2))||_V$. Hence by Proposition 3.5.1, we have $\eta(\lambda, x, \mathbb{W}) = ||(\eta(\lambda, x_1, W_1), \eta(\lambda, x_2, W_2))||_V = \frac{1}{||\Lambda||_{V,*}} \left|\left| \left(\frac{||r_1||}{||x_1||}, \frac{||r_2||}{||x_2||} \right) \right|\right|_V$. $\square$

Next we derive backward error of an approximate eigenvalue $\lambda \in \mathbb{C}^m$ of $\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$.

Proposition 3.5.2. *Let* $W_i : \mathbb{C}^2 \to \mathbb{C}^{n_i \times n_i}$ *be given by* $\sum_{m=0}^{k} \sum_{n=0}^{k-m} A_{mn}^{(i)} z_1^m z_2^n$, *where* $z = (z_1, z_2)^T \in \mathbb{C}^2$, $A_{mn}^{(i)} \in \mathbb{C}^{n_i \times n_i}$. *Let* $\lambda = (\lambda_1, \lambda_2)^T \in \mathbb{C}^2$ *and define* $\eta(\lambda, W_i) := \inf \{||\Delta W_i||_V : \text{rank}(W_i(\lambda) + \Delta W_i(\lambda)) < n_i\}$. *Then for a subordinate matrix norm on* $\mathbb{C}^{n_i \times n_i}$, *we have* $\eta(\lambda, W_i) = \frac{r_i}{||\Lambda||_{V,*}}$, *where* $r_i = \min\limits_{||x_i||=1} \{||W_i(\lambda)x_i|| : x_i \in \mathbb{C}^{n_i}\}$.

Let $x_i \in \mathbb{C}^{n_i}$ *be such that* $||x_i|| = 1$ *and* $r_i = ||W_i(\lambda)x_i||$. *Let* $y_i \in \partial||x_i||$ *and* $t := (t_{k,0}, t_{k-1,1}, \cdots, t_{0,k}, t_{k-1,0}, t_{k-2,1}, \cdots, t_{0,k-1}, \cdots, t_{1,0}, t_{0,1}, t_{0,0})^T \in \mathbb{C}^{\frac{(k+1)(k+2)}{2}}$ *be such that* $t \in \partial||\Lambda||_{V,*}$. *Define* $\Delta A_{mn}^{(i)} := -\frac{\overline{t_{m,n}}}{||\Lambda||_{V,*}} \frac{r_i y_i^*}{||x_i||}$ *for* $m = 0, \cdots, k$ *and* $n = 0, \cdots, k-m$. *Consider* $\Delta W_i(z) := \sum_{m=0}^{k} \sum_{n=0}^{k-m} \Delta A_{mn}^{(i)} z_1^m z_2^n$. *Then* $\text{rank}(W_i(\lambda) + \Delta W_i(\lambda)) < n_i$ *and* $||\Delta W_i||_V = \eta(\lambda, W_i)$.

Proof. By similar arguments as those in the proof of Proposition 2.5.2, we have $\eta(\lambda, W_i) \geq \frac{r_i}{||\Lambda||_{V,*}}$. It is also easy to see that $W_i(\lambda)x_i + \Delta W_i(\lambda)x_i = 0$, i.e., $\text{rank}(W_i(\lambda) + \Delta W_i(\lambda)) < n_i$ and $||\Delta A_{mn}^{(i)}|| = \frac{|t_{m,n}|}{||\Lambda||_{V,*}} \frac{||r_i||}{||x_i||}$. Consequently, we have $||\Delta W_i||_V = \frac{r_i}{||\Lambda||_{V,*}}$ and $\eta(\lambda, W_i) = \frac{r_i}{||\Lambda||_{V,*}}$. $\square$

Now we derive $\eta(\lambda, \mathbb{W})$ considering subordinate matrix norm on $\mathbb{C}^{n_i \times n_i}$.

Theorem 3.5.2. *Let* $\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$ *be regular and* $\lambda = (\lambda_1, \lambda_2)^T \in \mathbb{C}^2$. *Consider* $r_i := \min\limits_{||x_i||=1} \{||W_i(\lambda)x_i|| : x_i \in \mathbb{C}^{n_i}\}$ *for* $i = 1, 2$. *Then for a subordinate matrix norm on* $\mathbb{C}^{n_k \times n_k}$, *we have*

$$\eta(\lambda, \mathbb{W}) = ||(\eta(\lambda, W_1), \eta(\lambda, W_2))||_V = \frac{1}{||\Lambda||_{V,*}} ||(r_1, r_2)||_V,$$

where $\eta(\lambda, W_i)$ *is as defined in Proposition 3.5.2. In particular, for the Hölder p-norm on* $\mathbb{MEP}(n_1, \cdots, n_m)$, $\eta(\lambda, \mathbb{W}) = \frac{1}{||\Lambda||_q} ||(r_1, r_2)||_p$, *where* $p^{-1} + q^{-1} = 1$.

Let $x_i \in \mathbb{C}^{n_i}$ *be such that* $||x_i|| = 1$ *and* $r_i = ||W_i(\lambda)x_i||$ *for* $i = 1, 2$. *Let* $y_i \in \partial||x_i||$ *and* $t := (t_{k,0}, t_{k-1,1}, \cdots, t_{0,k}, t_{k-1,0}, t_{k-2,1}, \cdots, t_{0,k-1}, \cdots, t_{1,0}, t_{0,1}, t_{0,0})^T \in \mathbb{C}^{\frac{(k+1)(k+2)}{2}}$ *be* *such that* $t \in \partial||\Lambda||_{V,*}$. *Define* $\Delta A_{mn}^{(i)} := -\frac{\overline{t_{m,n}}}{||\Lambda||_{V,*}} \frac{r_i y_i^*}{||x_i||}$ *for* $m = 0, \cdots, k, n = 0, \cdots, k-m$ *and* $\Delta W_i(z) := \sum_{m=0}^{k} \sum_{n=0}^{k-m} \Delta A_{mn}^{(i)} z_1^m z_2^n$. *Consider* $\Delta \mathbb{W}(z) = (\Delta W_1(z), \Delta W_2(z))$. *Then* $\lambda \in \sigma(\mathbb{W} + \Delta \mathbb{W})$ *and* $\|\Delta \mathbb{W}\|_V = \eta(\lambda, \mathbb{W})$.

Proof. By similar arguments as those in the proof of Theorem 2.5.2, it follows that $\eta(\lambda, \mathbb{W}) \geq ||(\eta(\lambda, W_1), \eta(\lambda, W_2))||_V$. Again by Proposition 3.5.2, we have $\text{rank}(W_i(\lambda) + \Delta W_i(\lambda)) < n_i$ for $i = 1, 2$, i.e., $\lambda \in \sigma(\mathbb{W} + \Delta \mathbb{W})$ and $||\Delta W_i||_V = \eta(\lambda, W_i)$. So $\|\Delta \mathbb{W}\|_V = ||(\eta(\lambda, W_1), \eta(\lambda, W_2))||_V$. Hence the desired result follows. ☐

For the spectral or Frobenious norm on $\mathbb{C}^{n_i \times n_i}$, we have the following result.

Theorem 3.5.3. *Let* $\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$ *be regular and* $\lambda \in \mathbb{C}^2$. *Then*

$$\eta(\lambda, \mathbb{W}) = \frac{1}{||\Lambda||_{V,*}} ||(\sigma_{\min}(W_1(\lambda)), \sigma_{\min}(W_2(\lambda)))||_V,$$

where $\sigma_{\min}(W_j(\lambda))$ *denotes the smallest singular value of* $W_j(\lambda)$ *for all* $j = 1, 2$. *In particular, for the Hölder p-norm on* $\mathbb{MEP}(n_1, \cdots, n_m)$ *we have*

$$\eta(\lambda, \mathbb{W}) = \frac{1}{||\Lambda||_q} ||(\sigma_{\min}(W_1(\lambda)), \sigma_{\min}(W_2(\lambda)))||_p.$$

Let $t := (t_{k,0}, t_{k-1,1}, \cdots, t_{0,k}, t_{k-1,0}, t_{k-2,1}, \cdots, t_{0,k-1}, \cdots, t_{1,0}, t_{0,1}, t_{0,0})^T \in \mathbb{C}^{\frac{(k+1)(k+2)}{2}}$ *be such that* $t \in \partial||\Lambda||_{V,*}$. *Consider the* SVD $W_i(\lambda) = U_i \Sigma_i V_i$ *and set* $u_i := U_i(:, n_i), v_i := V_i(:, n_i)$ *for* $i = 1, 2$. *Define* $\Delta A_{mn}^{(j)} := -\frac{\overline{t_{m,n}}}{||\Lambda||_{V,*}} \sigma_{\min}(W_j(\lambda)) u_j v_j^*$ *for* $m = 0, \cdots, k, n = 0, \cdots, k-m$ *and* $\Delta W_i(z) := \sum_{m=0}^{k} \sum_{n=0}^{k-m} \Delta A_{mn}^{(i)} z_1^m z_2^n$. *Consider* $\Delta \mathbb{W}(z) := (\Delta W_1(z), \Delta W_2(z))$. *Then* $\lambda \in \sigma(\mathbb{W} + \Delta \mathbb{W})$ *and* $\|\Delta \mathbb{W}\|_V = \eta(\lambda, \mathbb{W})$.

Proof. Proof is similar to Theorem 2.5.3. ☐

4

Sensitivity and Backward Perturbation Analysis of Homogeneous MEP

4.1 Introduction

In this chapter we develop a framework for sensitivity and backward perturbation analysis of a regular homogeneous MEP $\mathbb{W}$. More precisely, we define condition number $\mathrm{cond}(\lambda, \mathbb{W})$ of a simple eigenvalue λ of $\mathbb{W}$ and derive various representations of $\mathrm{cond}(\lambda, \mathbb{W})$. We also define the backward error $\eta(\lambda, \mathbb{W})$ of λ as an approximate eigenvalue of $\mathbb{W}$ and determine $\eta(\lambda, \mathbb{W})$ as well as an optimal perturbation $\Delta\mathbb{W}$ such that $\lambda \in \sigma(\mathbb{W} + \Delta\mathbb{W})$ and $|||\Delta\mathbb{W}||| = \eta(\lambda, \mathbb{W})$. We also derive similar results for an approximate eigenpair $(\lambda, x = x_1 \otimes \cdots \otimes x_m)$ of $\mathbb{W}$.

4.2 Homogeneous MEP

Let $\mathbb{W} : \mathbb{C}^{m+1} \longrightarrow \prod_{j=1}^{m} \mathbb{C}^{n_j \times n_j}$ be a linear map given by

$$\mathbb{W}(z) := (W_1(z), \cdots, W_m(z)), \text{ where } W_i(z) = z_0 A_i + \sum_{j=1}^{m} z_j B_{ij}, \tag{4.1}$$

$z := (z_0, \cdots, z_m)^T \in \mathbb{C}^{m+1}$ and $A_i, B_{ij} \in \mathbb{C}^{n_i \times n_i}$ for all $i, j = 1, \cdots, m$. We refer to $\mathbb{W}$ as homogeneous MEP. We denote the space of all homogeneous MEPs by $\mathbb{MEP}(n_1, \cdots, n_m)$. For $x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ and $z \in \mathbb{C}^m$, we define

$$\mathbb{W}(z)x := (W_1(z)x_1, \ldots, W_m(z)x_m) \text{ and } x^*\mathbb{W}(z) := (x_1^* W_1(z), \ldots, x_m^* W_m(z)).$$

Definition 4.2.1. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be given by* $\mathbb{W}(z) = (W_1(z), \cdots, W_m(z))$*. Then* $\mathbb{W}$ *is said to be regular if* $\prod_{j=1}^{m} \det(W_j(z)) \neq 0$ *for some* $z \in \mathbb{C}^{m+1}$*.*

Throughout the chapter, we consider regular homogeneous MEPs. Next we define the spectrum of a regular homogeneous MEP. Recall that the **det** function is defined in (2.2).

Definition 4.2.2. (Spectrum) *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular. Then* $\lambda \in \mathbb{C}^{m+1} \backslash \{0\}$ *is said to be an eigenvalue of* $\mathbb{W}$ *if* $\mathbf{det}(\mathbb{W}(\lambda)) = 0$*. The set of all eigenvalues of* $\mathbb{W}$ *is called the spectrum of* $\mathbb{W}$ *and is denoted by* $\sigma(\mathbb{W})$*. Let* $\lambda \in \sigma(\mathbb{W})$*. Then* $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ *is called a right eigenvector of* $\mathbb{W}$ *corresponding to* λ *if* $\mathbb{W}(\lambda)x = 0$*. Similarly* $0 \neq y := y_1 \otimes \cdots \otimes y_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ *is called a left eigenvector of* $\mathbb{W}$ *corresponding to* λ *if* $y^*\mathbb{W}(\lambda) = 0$*.*

Let $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ be a regular homogeneous MEP as given in (4.1) and $\lambda \in \sigma(\mathbb{W})$. Let $x = x_1 \otimes \cdots \otimes x_m$ and $y = y_1 \otimes \cdots \otimes y_m$ be right and left eigenvector of $\mathbb{W}$ corresponding to λ, respectively. Then we define

$$J_{\mathbb{W}}(\lambda; y, x) := \begin{bmatrix} y_1^* A_1 x_1 & y_1^* B_{11} x_1 & \cdots & y_1^* B_{1m} x_1 \\ \vdots & \vdots & & \vdots \\ y_m^* A_m x_m & y_m^* B_{m1} x_m & \cdots & y_m^* B_{mm} x_m \end{bmatrix}. \tag{4.2}$$

Next we define geometric multiplicity of an eigenvalue of an MEP.

Definition 4.2.3. (Geometric multiplicity) *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *given by* $\mathbb{W}(z) = (W_1(z), \cdots, W_m(z))$ *be regular and* $\lambda \in \sigma(\mathbb{W})$*. Then the geometric multiplicity of* λ *is defined by* $g_{\mathbb{W}}(\lambda) = \prod_{i=1}^{m} \dim \ker(W_i(\lambda))$*. If* $g_{\mathbb{W}}(\lambda) = 1$*, then* λ *is said to be geometrically simple.*

Let $J_{\mathbb{W}}(z)$ denote the Jacobian of the differentiable function $\mathbb{C}^{m+1} \to \mathbb{C}^m, z \mapsto \mathbf{det}(\mathbb{W}(z))$, at z. Note that $J_{\mathbb{W}}(z)$ is a rectangular $m \times (m+1)$ matrix.

Proposition 4.2.1. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *given by (4.1) be regular. Then we* *have*

$$J_{\mathbb{W}}(z) = \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(z))A_1) & \mathrm{tr}(\mathrm{adj}(W_1(z))B_{11}) & \cdots & \mathrm{tr}(\mathrm{adj}(W_1(z))B_{1m}) \\ \vdots & \vdots & & \vdots \\ \mathrm{tr}(\mathrm{adj}(W_m(z))A_m) & \mathrm{tr}(\mathrm{adj}(W_m(z))B_{m1}) & \cdots & \mathrm{tr}(\mathrm{adj}(W_m(z))B_{mm}) \end{bmatrix} \tag{4.3}$$

for all $z \in \mathbb{C}^{m+1}$. In particular, if $\lambda \in \sigma(\mathbb{W})$ is geometrically simple then there exist left and right eigenvectors $\widehat{u} := \widehat{u}_1 \otimes \cdots \otimes \widehat{u}_m$ and $\widehat{v} := \widehat{v}_1 \otimes \cdots \otimes \widehat{v}_m$ of $\mathbb{W}$ corresponding to λ, respectively, such that

$$J_{\mathbb{W}}(\lambda) = \begin{bmatrix} \widehat{u}_1^* A_1 \widehat{v}_1 & \widehat{u}_1^* B_{11} \widehat{v}_1 & \cdots & \widehat{u}_1^* B_{1m} \widehat{v}_1 \\ \vdots & \vdots & & \vdots \\ \widehat{u}_m^* A_m \widehat{v}_m & \widehat{u}_m^* B_{m1} \widehat{v}_m & \cdots & \widehat{u}_m^* B_{mm} \widehat{v}_m \end{bmatrix} = J_{\mathbb{W}}(\lambda; \widehat{u}, \widehat{v}). \tag{4.4}$$

Proof. Note that $J_{\mathbb{W}}(z) = \left[\frac{\partial}{\partial z_j}(\det(W_i(z)))\right]_{m\times(m+1)}$. By Jacobi's formula stated in Proposition 1.1.2, it follows that

$$\frac{\partial}{\partial z_j}(\det(W_i(z))) = \begin{cases} \mathrm{tr}\,(\mathrm{adj}(W_i(z))A_i) & \text{for } j = 0 \\ \mathrm{tr}\,(\mathrm{adj}(W_i(z))B_{ij}) & \text{for } j = 1, \cdots, m, \end{cases}$$

which yeilds (4.3).

If $\lambda \in \sigma(\mathbb{W})$ is geometrically simple then $\mathrm{rank}(W_i(\lambda)) = n_i - 1$ for all $i = 1, \cdots, m$. Thus by Proposition 1.1.1 there exist nonzero vectors $u_i, v_i \in \mathbb{C}^{n_i}$ such that $\mathrm{adj}(W_i(\lambda)) = v_i u_i^*$, where $W_i(\lambda)v_i = 0$ and $u_i^* W_i(\lambda) = 0$ for all $i = 1, \cdots, m$. Hence the desired result in (4.4) follows. $\square$

Remark 4.2.1. *Let $\mathbb{W} : \mathbb{C}^{m+1} \to \prod_{j=1}^{m} \mathbb{C}^{n_j \times n_j}$ given by $\mathbb{W}(z) = (W_1(z), \cdots, W_m(z))$ be differentiable. Then the Jacobian matrix of the map $\mathbb{C}^{m+1} \to \mathbb{C}^m, z \mapsto \mathbf{det}(\mathbb{W}(z))$ is given by*

$$J_{\mathbb{W}}(z) = \begin{bmatrix} \mathrm{tr}\left(\mathrm{adj}(W_1(z)) \cdot \frac{\partial}{\partial z_0}(W_1(z))\right) & \cdots & \mathrm{tr}\left(\mathrm{adj}(W_1(z)) \cdot \frac{\partial}{\partial z_m}(W_1(z))\right) \\ \vdots & & \vdots \\ \mathrm{tr}\left(\mathrm{adj}(W_m(z)) \cdot \frac{\partial}{\partial z_0}(W_m(z))\right) & \cdots & \mathrm{tr}\left(\mathrm{adj}(W_m(z)) \cdot \frac{\partial}{\partial z_m}(W_m(z))\right) \end{bmatrix}$$

for all $z = (z_0, z_1, \cdots, z_m)^T \in \mathbb{C}^{m+1}$.

Now we define an algebraically simple eigenvalue of a regular homogeneous MEP.

Definition 4.2.4. (Algebraic simplicity) *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular and* $\lambda \in \sigma(\mathbb{W})$. *Then* λ *is said to be algebraically simple if* $\operatorname{rank}(J_{\mathbb{W}}(\lambda)) = m$. *If* $\operatorname{rank}(J_{\mathbb{W}}(\lambda)) < m$ *then* λ *is said to be a multiple eigenvalue of* $\mathbb{W}$.

Throughout the chapter, we refer to algebraically simple eigenvalues as simple eigenvalues. As expected, a simple eigenvalue is geometrically simple.

Proposition 4.2.2. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *given by (4.1) be regular and* $\lambda \in \sigma(\mathbb{W})$ *be simple. Then* λ *is geometrically simple.*

Proof. If $\dim(\ker(W_i(\lambda))) > 1$ for some $i \in \{1, \cdots, m\}$ then $\operatorname{adj}(W_i(\lambda)) = 0$. Thus it follows from (4.3) that the i^{th} row of $J_{\mathbb{W}}(\lambda)$ is a zero row which contradicts that λ is simple. Hence $\dim(\ker(W_i(\lambda))) = 1$ for all $i = 1, \cdots, m$ and $g_{\mathbb{W}}(\lambda) = \prod_{i=1}^{m} \dim(\ker(W_i(\lambda))) = 1$, i.e., λ is geometrically simple. □

Given $\lambda \in \mathbb{C}^{m+1}$, define $G_\lambda : \mathbb{C}^{m+1} \to \mathbb{C}^{m+1}$ by

$$G_\lambda(x) = \begin{bmatrix} \mathbf{det}(\mathbb{W}(x)) \\ \langle x - \lambda, \lambda \rangle \end{bmatrix}, \quad x \in \mathbb{C}^{m+1}, \tag{4.5}$$

where $\langle \cdot, \cdot \rangle$ is the usual inner product on $\mathbb{C}^{m+1}$. It is easy to see that G_λ is a differentiable function. We derive the derivative $DG_\lambda(x)$ of $G_\lambda(x)$ with respect to x in the following result.

Proposition 4.2.3. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *given by (4.1) be regular and* $\lambda \in \mathbb{C}^{m+1}$. *Consider the differentiable function* G_λ *given in (4.5). Then*

$$DG_\lambda(x) = \begin{bmatrix} J_{\mathbb{W}}(x) \\ \lambda^* \end{bmatrix} \tag{4.6}$$

for all $x \in \mathbb{C}^{m+1}$.

Proof. It follows from Proposition 4.2.1 that $\left[\frac{\partial}{\partial x_j}(\det(W_i(x)))\right]_{m \times (m+1)} = J_{\mathbb{W}}(x)$. Again $\frac{\partial}{\partial x_j}(\langle x - \lambda, \lambda \rangle) = \overline{\lambda_j}$. Thus (4.6) follows. □

It follows from (4.5) that $\lambda \in \sigma(\mathbb{W}) \Leftrightarrow G_\lambda(\lambda) = 0$. We now show that $\lambda \in \sigma(\mathbb{W})$ is simple $\Leftrightarrow$ λ is a simple zero of $G_\lambda(z)$. Note that by simple zero of $G_\lambda(z)$ we mean a solution of $G_\lambda(z) = 0$ at which the derivative $DG_\lambda(z)$ is nonsingular.

Theorem 4.2.1. *Let $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ given by (4.1) be regular and $\lambda \in \sigma(\mathbb{W})$. Then λ is simple if and only if λ is a simple zero of G_λ, i.e., if and only if $G_\lambda(\lambda) = 0$ and* $\mathrm{rank}(DG_\lambda(\lambda)) = m + 1$.

Proof. Let λ be a simple eigenvalue of $\mathbb{W}$, i.e., $\mathrm{rank}(J_{\mathbb{W}}(\lambda)) = m$. If possible, suppose that λ is not a simple zero of G_λ, i.e., $\mathrm{rank}(DG_\lambda(\lambda)) < m + 1$. Then there exists $0 \neq \alpha \in \mathbb{C}^m$ such that $\lambda^* = \alpha^T J_{\mathbb{W}}(\lambda)$. Since $\lambda \in \sigma(\mathbb{W})$, it is easy to see that

$$\begin{aligned} J_{\mathbb{W}}(\lambda)\lambda &= \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(z))A_1) & \mathrm{tr}(\mathrm{adj}(W_1(z))B_{11}) & \cdots & \mathrm{tr}(\mathrm{adj}(W_1(z))B_{1m}) \\ \vdots & \vdots & & \vdots \\ \mathrm{tr}(\mathrm{adj}(W_m(z))A_m) & \mathrm{tr}(\mathrm{adj}(W_m(z))B_{m1}) & \cdots & \mathrm{tr}(\mathrm{adj}(W_m(z))B_{mm}) \end{bmatrix} \begin{bmatrix} \lambda_0 \\ \vdots \\ \lambda_m \end{bmatrix} \\ &= \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(\lambda)) \cdot W_1(\lambda)) \\ \vdots \\ \mathrm{tr}(\mathrm{adj}(W_m(\lambda)) \cdot W_m(\lambda)) \end{bmatrix} = \begin{bmatrix} \mathrm{tr}(\det(W_1(\lambda))I_{n_1}) \\ \vdots \\ \mathrm{tr}(\det(W_m(\lambda))I_{n_m}) \end{bmatrix} = 0. \end{aligned}$$

Then $\lambda^*\lambda = \alpha^T J_{\mathbb{W}}(\lambda)\lambda = 0$, i.e., $||\lambda||_2^2 = 0$, which is not possible since $\lambda \in \sigma(\mathbb{W})$. Thus $\mathrm{rank}(DG_\lambda(\lambda)) = m + 1$ and λ is a simple zero of G_λ.

Conversely let λ be a simple zero of G_λ, i.e., $\mathrm{rank}(DG_\lambda(\lambda)) = m + 1$. Since the rows of $J_{\mathbb{W}}(\lambda)$ are the first m rows of $DG_\lambda(\lambda)$ so it follows that $\mathrm{rank}\,(J_{\mathbb{W}}(\lambda))) = m$. Hence λ is a simple eigenvalue of $\mathbb{W}$. ☐

Next we define normal derivative and normal Jacobian matrix of the function $x \mapsto \mathbf{det}(\mathbb{W}(x))$ at some nonzero $z \in \mathbb{C}^{m+1}$.

Definition 4.2.5. *Let $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ be regular and $z \in \mathbb{C}^{m+1}$ be nonzero. Consider the hyperplane $\{z\}^\perp := \{y \in \mathbb{C}^{m+1} : \langle z, y\rangle = 0\}$ in $\mathbb{C}^{m+1}$. Then we define $D_{z^\perp}\mathbf{det}(\mathbb{W}(z)) : \{z\}^\perp \to \mathbb{C}^m$ by $u \mapsto J_{\mathbb{W}}(z)u$ for all $u \in \{z\}^\perp$ and we refer to $D_{z^\perp}\mathbf{det}(\mathbb{W}(z))$ as the normal derivative of the function $x \mapsto \mathbf{det}(\mathbb{W}(x))$ at z.*

Let $U := \begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}$ be a basis of $\{z\}^\perp$. Then $N_{\mathbb{W}}(z) := J_{\mathbb{W}}(z) \cdot U$ denotes the matrix representation of $D_{z^\perp}\mathbf{det}(\mathbb{W}(z))$ with respect to the basis U of $\{z\}^\perp$ and the *standard basis of* $\mathbb{C}^m$*. We refer to* $N_{\mathbb{W}}(z)$ *as the normal Jacobian matrix of the function* $x \mapsto \mathbf{det}(\mathbb{W}(x))$ *at* z.

Let $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ be regular and $\lambda \in \sigma(\mathbb{W})$. Let $x = x_1 \otimes \cdots \otimes x_m$ and $y = y_1 \otimes \cdots \otimes y_m$ be right and left eigenvector of $\mathbb{W}$ corresponding to λ, respectively. Let $U := \begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}$ be a basis of $\{\lambda\}^{\perp}$.Then we define

$$N_{\mathbb{W}}(\lambda; y, x) := \begin{bmatrix} y_1^* W_1(u_1) x_1 & \cdots & y_1^* W_1(u_m) x_1 \\ & \vdots & \\ y_m^* W_m(u_1) x_m & \cdots & y_m^* W_m(u_m) x_m \end{bmatrix}. \tag{4.7}$$

Proposition 4.2.4. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular and* $\lambda \in \mathbb{C}^{m+1}$ *be nonzero. Let* $U := \begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}$ *be a basis of* $\{\lambda\}^{\perp}$*. Then*

$$N_{\mathbb{W}}(\lambda) = \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(\lambda)) \cdot W_1(u_1)) & \cdots & \mathrm{tr}(\mathrm{adj}(W_1(\lambda)) \cdot W_1(u_m)) \\ \vdots & & \vdots \\ \mathrm{tr}(\mathrm{adj}(W_m(\lambda)) \cdot W_m(u_1)) & \cdots & \mathrm{tr}(\mathrm{adj}(W_m(\lambda)) \cdot W_m(u_m)) \end{bmatrix}. \tag{4.8}$$

In particular, if $\lambda \in \sigma(\mathbb{W})$ *is geometrically simple then there exist left and right eigenvectors* $\widehat{u} := \widehat{u}_1 \otimes \cdots \otimes \widehat{u}_m$ *and* $\widehat{v} := \widehat{v}_1 \otimes \cdots \otimes \widehat{v}_m$ *of* $\mathbb{W}$ *corresponding to* λ*, respectively, such that*

$$N_{\mathbb{W}}(\lambda) = \begin{bmatrix} \widehat{u}_1^* W_1(u_1) \widehat{v}_1 & \cdots & \widehat{u}_1^* W_1(u_m) \widehat{v}_1 \\ \vdots & & \vdots \\ \widehat{u}_m^* W_m(u_1) \widehat{v}_m & \cdots & \widehat{u}_m^* W_m(u_m) \widehat{v}_m \end{bmatrix} = N_{\mathbb{W}}(\lambda; \widehat{u}, \widehat{v}). \tag{4.9}$$

Proof. We have $N_{\mathbb{W}}(\lambda) = J_{\mathbb{W}}(\lambda) \cdot U = \begin{bmatrix} J_{\mathbb{W}}(\lambda) u_1 & \cdots & J_{\mathbb{W}}(\lambda) u_m \end{bmatrix}$. Now

$$\begin{aligned} J_{\mathbb{W}}(\lambda) u_i &= \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(\lambda)) A_1) & \mathrm{tr}(\mathrm{adj}(W_1(\lambda)) B_{11}) & \cdots & \mathrm{tr}(\mathrm{adj}(W_1(\lambda)) B_{1m}) \\ \vdots & \vdots & & \vdots \\ \mathrm{tr}(\mathrm{adj}(W_m(\lambda)) A_m) & \mathrm{tr}(\mathrm{adj}(W_m(\lambda)) B_{m1}) & \cdots & \mathrm{tr}(\mathrm{adj}(W_m(\lambda)) B_{mm}) \end{bmatrix} \cdot u_i \\ &= \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(\lambda)) \cdot W_1(u_i)) \\ \vdots \\ \mathrm{tr}(\mathrm{adj}(W_m(\lambda)) \cdot W_m(u_i)) \end{bmatrix} \quad \text{for } i = 1, \cdots, m. \end{aligned}$$

This proves (4.8).

If $\lambda \in \sigma(\mathbb{W})$ is geometrically simple then it follows from Definition 4.2.3 that $\operatorname{rank}(W_i(\lambda)) = n_i - 1$ for all $i = 1, \cdots, m$. So by Proposition 1.1.1, there exist nonzero vectors $\widehat{u}_i, \widehat{v}_i \in \mathbb{C}^{n_i}$ such that $\operatorname{adj}(W_i(\lambda)) = \widehat{v}_i\widehat{u}_i^*$, where $W_i(\lambda)\widehat{v}_i = 0$ and $\widehat{u}_i^* W_i(\lambda) = 0$ for all $i = 1, \cdots, m$. Thus we have $\operatorname{tr}(\operatorname{adj}(W_i(\lambda))W_i(u_j)) = \operatorname{tr}(\widehat{v}_i\widehat{u}_i^* W_i(u_j)) = \widehat{u}_i^* W_i(u_j)\widehat{v}_i$, which yields (4.11). □

Consider a special case when $\lambda := (\lambda_0, \cdots, \lambda_m)^T \in \mathbb{C}^{m+1}$ is such that $\lambda_0 \neq 0$. Define $u_j := \overline{\lambda_j}e_0 - \overline{\lambda_0}e_j$ for $j = 1, \cdots, m$. Then $U = \begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}$ is a basis of $\{\lambda\}^\perp$ and as an immediate consequence of Proposition 4.2.4, we have the following result.

Corollary 4.2.2. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *given by (4.1) be regular and* $\lambda := (\lambda_0, \cdots, \lambda_m)^T \in \mathbb{C}^{m+1}$ *be such that* $\lambda_0 \neq 0$*. Consider the basis* $U = \begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}$ *of* $\{\lambda\}^\perp$*, where* $u_j := \overline{\lambda_j}e_0 - \overline{\lambda_0}e_j$ *for* $j = 1, \cdots, m$*. Then*

$$N_{\mathbb{W}}(\lambda) = \begin{bmatrix} \operatorname{tr}(\operatorname{adj}(W_1(\lambda)) \cdot (\overline{\lambda_1}A_1 - \overline{\lambda_0}B_{11})) & \cdots & \operatorname{tr}(\operatorname{adj}(W_1(\lambda)) \cdot (\overline{\lambda_m}A_1 - \overline{\lambda_0}B_{1m})) \\ \vdots & & \vdots \\ \operatorname{tr}(\operatorname{adj}(W_m(\lambda)) \cdot (\overline{\lambda_1}A_m - \overline{\lambda_0}B_{m1})) & \cdots & \operatorname{tr}(\operatorname{adj}(W_m(\lambda)) \cdot (\overline{\lambda_m}A_m - \overline{\lambda_0}B_{mm})) \end{bmatrix}. \tag{4.10}$$

In particular, if $\lambda \in \sigma(\mathbb{W})$ *is geometrically simple then there exist left and right eigenvectors* $\widehat{u} := \widehat{u}_1 \otimes \cdots \otimes \widehat{u}_m$ *and* $\widehat{v} := \widehat{v}_1 \otimes \cdots \otimes \widehat{v}_m$ *of* $\mathbb{W}$ *corresponding to* λ*, respectively, such that*

$$N_{\mathbb{W}}(\lambda) = \begin{bmatrix} \widehat{u}_1^*(\overline{\lambda_1}A_1 - \overline{\lambda_0}B_{11})\widehat{v}_1 & \cdots & \widehat{u}_1^*(\overline{\lambda_m}A_1 - \overline{\lambda_0}B_{1m})\widehat{v}_1 \\ \vdots & & \vdots \\ \widehat{u}_m^*(\overline{\lambda_1}A_m - \overline{\lambda_0}B_{m1})\widehat{v}_m & \cdots & \widehat{u}_m^*(\overline{\lambda_m}A_m - \overline{\lambda_0}B_{mm})\widehat{v}_m \end{bmatrix}. \tag{4.11}$$

Proof. Note that $W_i(u_j) = \overline{\lambda_j}A_i - \overline{\lambda_0}B_{ij}$ for all $i, j = 1, \cdots, m$. Hence the proof follows from Proposition 4.2.4. □

Next we characterize a simple eigenvalue $\lambda \in \sigma(\mathbb{W})$.

Theorem 4.2.3. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular and* $\lambda \in \sigma(\mathbb{W})$*. Then* λ *is simple if and only if* $N_{\mathbb{W}}(\lambda)$ *is nonsingular.*

Proof. Let $U := \begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}$ be a basis of $\{\lambda\}^\perp$.

($\Rightarrow$) Let $\lambda \in \sigma(\mathbb{W})$ be simple. Suppose that $N_{\mathbb{W}}(\lambda)\alpha = 0$ for some $\alpha \in \mathbb{C}^m$. By (4.6) we have

$$DG_\lambda(\lambda)U\alpha = \begin{bmatrix} J_{\mathbb{W}}(\lambda)U\alpha \\ \lambda^* U\alpha \end{bmatrix} = \begin{bmatrix} N_{\mathbb{W}}(\lambda)\alpha \\ 0 \end{bmatrix} = 0.$$

Now by Theorem 4.2.1, $DG_\lambda(\lambda)$ is nonsingular since λ is simple. So $DG_\lambda(\lambda)U\alpha = 0$ implies that $U\alpha = 0 \Rightarrow \alpha = 0$. Hence $N_{\mathbb{W}}(\lambda)$ is nonsingular.

($\Leftarrow$) Let $N_{\mathbb{W}}(\lambda)$ be nonsingular. Suppose that $DG_\lambda(\lambda)z = 0$ for some $z \in \mathbb{C}^{m+1}$. Note that there exist $\beta \in \mathbb{C}$ and $\gamma \in \mathbb{C}^m$ such that $z = U\gamma + \beta\lambda$. Since $J_{\mathbb{W}}(\lambda)\lambda = 0$ as $\lambda \in \sigma(\mathbb{W})$ it follows from (4.6) that

$$DG_\lambda(\lambda)z = \begin{bmatrix} J_{\mathbb{W}}(\lambda)(U\gamma + \beta\lambda) \\ \lambda^*(U\gamma + \beta\lambda) \end{bmatrix} = \begin{bmatrix} J_{\mathbb{W}}(\lambda)U\gamma \\ \beta||\lambda||_2^2 \end{bmatrix} = \begin{bmatrix} N_{\mathbb{W}}(\lambda)\gamma \\ \beta||\lambda||_2^2 \end{bmatrix} = 0.$$

Since $\lambda \neq 0$, $\beta||\lambda||_2^2 = 0$ implies that $\beta = 0$. Again nonsingularity of $N_{\mathbb{W}}(\lambda)$ implies that $\gamma = 0$. So $z = 0$ and thus $DG_\lambda(\lambda)$ is nonsingular. Therefore by Theorem 4.2.1, it follows that λ is simple. ☐

The following result also characterizes simple eigenvalues.

Theorem 4.2.4. *Let $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ be regular and $\lambda \in \mathbb{C}^{m+1}$ be nonzero. Let $U := \begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}$ be a basis of $\{\lambda\}^\perp$. Then the following conditions are equivalent:*

(i) λ is simple.

(ii) λ is a simple zero of the function G_λ which is defined in (4.5).

(iii) $N_{\mathbb{W}}(\lambda)$ is nonsingular.

(iv) $N_{\mathbb{W}}(\lambda; y, x)$ is nonsingular for all left and right eigenvectors y, x of $\mathbb{W}$ corresponding to λ, respectively.

Proof. Note that (i) $\Rightarrow$ (ii) follows by Theorem 4.2.1 and (ii) $\Rightarrow$ (iii) follows by Theorem 4.2.3.

Next we show (iii) $\Rightarrow$ (iv). Suppose that $N_{\mathbb{W}}(\lambda)$ is nonsingular. If $\dim \ker(W_k(\lambda))) > 1$ for some $k \in \{1, \cdots, m\}$, then $\operatorname{adj}(W_k(\lambda)) = 0$. Then it follows from (4.8) that the

k^{th} row of $N_{\mathbb{W}}(\lambda)$ is a zero row which contradicts that $N_{\mathbb{W}}(\lambda)$ is nonsingular. Therefore $\dim\ker(W_i(\lambda))) = 1$ for all $i = 1, \cdots, m$, i.e., λ is geometrically simple. Now by Proposition 4.2.4, there exist left and right eigenvectors $\widehat{u}$ and $\widehat{v}$ of $\mathbb{W}$ corresponding to λ, respectively, such that $N_{\mathbb{W}}(\lambda) = N_{\mathbb{W}}(\lambda; \widehat{u}, \widehat{v})$. Note that $N_{\mathbb{W}}(\lambda; \widehat{u}, \widehat{v})$ is nonsingular as $N_{\mathbb{W}}(\lambda)$ is nonsingular. Since $\dim\ker(W_i(\lambda))) = 1$ for all $i = 1, \cdots, m$, there exists nonzero scalars α_i and β_i such that $\widehat{u}_i = \alpha_i y_i$ and $\widehat{v}_i = \beta_i x_i$ for all $i = 1, \cdots, m$. Then $\widehat{u}_i^* W_i(u_j)\widehat{v}_i = (\alpha_i y_i)^* W_i(u_j)(\beta_i x_i) = \overline{\alpha_i}\beta_i y_i^* W_i(u_j) x_i$. So we have

$$\begin{aligned} N_{\mathbb{W}}(\lambda; \widehat{u}, \widehat{v}) &= \begin{bmatrix} \overline{\alpha_1}\beta_1 y_1^* W_1(u_1) x_1 & \cdots & \overline{\alpha_1}\beta_1 y_1^* W_1(u_m) x_1 \\ \vdots & & \vdots \\ \overline{\alpha_m}\beta_m y_m^* W_m(u_1) x_m & \cdots & \overline{\alpha_m}\beta_m y_m^* W_m(u_m) x_m \end{bmatrix} \\ &= \begin{bmatrix} \overline{\alpha_1}\beta_1 & & \\ & \ddots & \\ & & \overline{\alpha_m}\beta_m \end{bmatrix} N_{\mathbb{W}}(\lambda; y, x). \end{aligned}$$

Therefore $N_{\mathbb{W}}(\lambda; y, x)$ is nonsingular since $N_{\mathbb{W}}(\lambda; \widehat{u}, \widehat{v})$ is nonsingular.

Finally we prove (iv) $\Rightarrow$ (i). Suppose that $N_{\mathbb{W}}(\lambda; y, x)$ is nonsingular for all left and right eigenvectors y and x of $\mathbb{W}$ corresponding to λ, respectively. If possible, suppose that $\dim\ker(W_k(\lambda)) > 1$ for some $k \in \{1, \cdots, m\}$. Let $y := y_1 \otimes \cdots \otimes y_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ be a left eigenvector of $\mathbb{W}$ corresponding to λ. Consider $x_i \in \mathbb{C}^{n_i}$ such that $W_i(\lambda)x_i = 0$ for all $i = 1, \cdots, k-1, k+1, \cdots, m$. Define the linear functional $g : \ker(W_k(\lambda)) \to \mathbb{C}$ given by $g(z_k) = \det(N_{\mathbb{W}}(\lambda; y, z))$, where $z = x_1 \otimes \cdots \otimes x_{k-1} \otimes z_k \otimes x_{k+1} \otimes \cdots \otimes x_m$ is a right eigenvector of $\mathbb{W}$ corresponding to λ. Now by the Rank - Nullity Theorem, we have $\dim\ker(g) = \dim(\ker(W_k(\lambda))) - 1 > 0$. So there exists $x_k \in \ker(W_k(\lambda))$ such that $g(x_k) = \det(N_{\mathbb{W}}(\lambda; y, x)) = 0$, where $x = x_1 \otimes \cdots \otimes x_{k-1} \otimes x_k \otimes x_{k+1} \otimes \cdots \otimes x_m$ is a right eigenvector of $\mathbb{W}$ corresponding to λ. This contradicts the assumption that $N_{\mathbb{W}}(\lambda; y, x)$ is nonsingular for all left and right eigenvectors y and x of $\mathbb{W}$ corresponding to λ, respectively. Therefore $\dim\ker(W_i(\lambda)) = 1$ for all $i = 1, \cdots, m$, i.e., λ is geometrically simple. Again by Proposition 4.2.4, there exist left and right eigenvectors $\widehat{u}$ and $\widehat{v}$ of $\mathbb{W}$ corresponding to λ, respectively, such that $N_{\mathbb{W}}(\lambda) = N_{\mathbb{W}}(\lambda; \widehat{u}, \widehat{v})$. So $N_{\mathbb{W}}(\lambda)$ is nonsingular and by Theorem 4.2.3, λ is simple. $\square$

4.3 Sensitivity analysis of simple eigenvalues

In this section we undertake sensitivity analysis of an algebraically simple eigenvalue of a regular homogeneous MEP and derive condition number of a simple eigenvalue. First we define the condition number of an eigenvalue of a regular homogeneous MEP.

Definition 4.3.1. (Condition number) *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular and* $\lambda \in \sigma(\mathbb{W})$. *Then we define the* ***condition number*** *of* λ *by*

$$\mathrm{cond}(\lambda, \mathbb{W}) := \limsup_{|||\Delta\mathbb{W}||| \to 0} \frac{\mathrm{dist}(\lambda, \sigma(\mathbb{W} + \Delta\mathbb{W}))}{|||\Delta\mathbb{W}|||},$$

where $\Delta\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *and* $\mathrm{dist}(\lambda, \sigma(\mathbb{W} + \Delta\mathbb{W})) := \min\{||\lambda - \mu||_2 : \mu \in \sigma(\mathbb{W} + \Delta\mathbb{W})\}$.

Let $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ be as given in (4.1). Recall that given a monotone norm $|| \cdot ||_V$ on $\mathbb{C}^{m+1}$ and a matrix norm $|| \cdot ||$ on $\mathbb{C}^{n_i \times n_i}$ for $i = 1, \cdots, m$, the norm $|||\cdot|||_V$ on $\mathbb{MEP}(n_1, \cdots, n_m)$ is given by $|||\mathbb{W}|||_V = ||(||W_1||_V, \cdots, ||W_m||_V)||_V$, where $||W_i||_V = ||(||A_i||, ||B_{i1}||, \cdots, ||B_{im}||)||_V$ for $i = 1, \cdots, m$. We define Hölder p-norm $|||\cdot|||_p$ on $\mathbb{MEP}(n_1, \cdots, n_m)$ by $|||\mathbb{W}|||_p = ||(||W_1||_p, \cdots, ||W_m||_p)||_p$, where $||W_i||_p$ is given by $||W_i||_p = ||(||A_i||, ||B_{i1}||, \cdots, ||B_{im}||)||_p$ for $1 \leq p \leq \infty$.

Theorem 4.3.1. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *given by (4.1) be regular and* $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ *be simple. Let* $y := y_1 \otimes \cdots \otimes y_m$ *and* $x := x_1 \otimes \cdots \otimes x_m$ *be left and right eigenvectors of* $\mathbb{W}$ *corresponding to* $\lambda_{\mathbb{W}}$, *respectively. Let* $U = \begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}$ *be a basis of* $\{\lambda_{\mathbb{W}}\}^{\perp}$ *and* $N_{\mathbb{W}}(\lambda_{\mathbb{W}})$ *be the normal Jacobian matrix of the function* $z \mapsto \mathbf{det}(\mathbb{W}(z))$ *at* $\lambda_{\mathbb{W}}$. *Then there is an open neighborhood* $\mathrm{nbd}(\mathbb{W}) \subset \mathbb{MEP}(n_1, \cdots, n_m)$ *containing* $\mathbb{W}$ *and a smooth function* $\lambda : \mathrm{nbd}(\mathbb{W}) \to \mathbb{C}^{m+1}$ *such that* $\lambda(\mathbb{W}) = \lambda_{\mathbb{W}}$ *and* $\lambda(\mathbb{X})$ *is a simple eigenvalue of* $\mathbb{X}$ *for all* $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$. *Further the derivative* $D\lambda(\mathbb{W}) : \mathbb{MEP}(n_1, \cdots, n_m) \to \mathbb{C}^{m+1}$ *of* λ

at $\mathbb{W}$ *is given by*

$$D\lambda(\mathbb{W})\mathbb{H} = -U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}\begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(\lambda_{\mathbb{W}}))H_1(\lambda_{\mathbb{W}})) \\ \vdots \\ \mathrm{tr}(\mathrm{adj}(W_m(\lambda_{\mathbb{W}}))H_m(\lambda_{\mathbb{W}})) \end{bmatrix} \tag{4.12}$$

$$= -U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}\begin{bmatrix} y_1^* H_1(\lambda_{\mathbb{W}})x_1 \\ \vdots \\ y_m^* H_m(\lambda_{\mathbb{W}})x_m \end{bmatrix} \tag{4.13}$$

for all $\mathbb{H} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *such that* $\mathbb{H}(z) = (H_1(z), \cdots, H_m(z))$.

a) *The condition number* $\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W})$ *is given by*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \; ||\lambda_{\mathbb{W}}^*||_{V,*} \max_i ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_*.$$

In particular, for the spectral or Frobenius norm on $\mathbb{C}^{n_i \times n_i}$, *we have*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \; ||\lambda_{\mathbb{W}}^*||_{V,*} \max_i \prod_{k=1}^{n_i-1} \sigma_k(W_i(\lambda_{\mathbb{W}})),$$

where $\sigma_1(W_i(\lambda_{\mathbb{W}})) \geq \cdots \geq \sigma_{n_i}(W_i(\lambda_{\mathbb{W}}))$ *are the singular values of* $W_i(\lambda_{\mathbb{W}})$ *for* $i = 1, \cdots, m$.

b) *Also we have*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_V \; ||\lambda_{\mathbb{W}}^*||_{V,*} \max_i ||y_i x_i^*||_*.$$

For a subordinate matrix norm $||\cdot||$ *on* $\mathbb{C}^{n_i \times n_i}$, *we have,*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_V \; ||\lambda_{\mathbb{W}}^*||_{V,*} \max_i ||y_i||_* ||x_i||.$$

In particular, for the spectral or Frobenius matrix norm on $\mathbb{C}^{n_i \times n_i}$, *we have,*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_V \; ||\lambda_{\mathbb{W}}^*||_{V,*} \max_i ||y_i||_2 ||x_i||_2.$$

Proof. Define $\phi : \mathbb{MEP}(n_1, \cdots, n_m) \times \mathbb{C}^{m+1} \to \mathbb{C}^{m+1}$ by $\phi(\mathbb{Y}, z) = \begin{bmatrix} \mathbf{det}(\mathbb{Y}(z)) \\ \langle z - \lambda_{\mathbb{W}}, \lambda_{\mathbb{W}} \rangle \end{bmatrix}$. Consider $\mathbb{V}(\phi) := \{(\mathbb{Y}, z) \in \mathbb{MEP}(n_1, \cdots, n_m) \times \mathbb{C}^{m+1} : \phi(\mathbb{Y}, z) = 0\}$. It is obvious that

$(\mathbb{W}, \lambda_{\mathbb{W}}) \in \mathbb{V}(\phi)$ and $\partial_\lambda \phi(\mathbb{W}, \lambda_{\mathbb{W}}) = DG_{\lambda_{\mathbb{W}}}(\lambda_{\mathbb{W}})$ which is nonsingular by Theorem 4.2.1 since $\lambda_{\mathbb{W}}$ is simple. Hence by the Implicit function theorem there is an open neighborhood $\mathrm{nbd}(\mathbb{W})$ containing $\mathbb{W}$ and a smooth function $\lambda : \mathrm{nbd}(\mathbb{W}) \to \mathbb{C}^{m+1}$ such that $\lambda(\mathbb{W}) = \lambda_{\mathbb{W}}$ and $\mathbb{V}(\phi) \cap (\mathrm{nbd}(\mathbb{W}) \times \mathbb{C}^{m+1}) = \{(\mathbb{X}, \lambda(\mathbb{X})) : \mathbb{X} \in \mathrm{nbd}(\mathbb{W})\}$ is the graph of the function $\mathbb{X} \mapsto \lambda(\mathbb{X})$ for all $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$. So $\phi(\mathbb{X}, \lambda(\mathbb{X})) = 0$ for all $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$, i.e., $\lambda(\mathbb{X}) \in \sigma(\mathbb{X})$ for all $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$. By similar arguments as given in Theorem 2.3.1, it follows that $\lambda(\mathbb{X}) \in \sigma(\mathbb{X})$ is simple for all $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$.

Now differentiating $\phi(\mathbb{X}, \lambda(\mathbb{X})) = 0$ with respect to $\mathbb{X}$ at $\mathbb{W}$ we have

$$\partial_\lambda \phi(\lambda_{\mathbb{W}}, \mathbb{W}) D\lambda(\mathbb{W})\mathbb{H} + \partial_{\mathbb{X}} \phi(\lambda_{\mathbb{W}}, \mathbb{W})\mathbb{H} = 0 \tag{4.14}$$

for all $\mathbb{H} \in \mathbb{MEP}(n_1, \cdots, n_m)$. Note that $\partial_\lambda \phi(\lambda_{\mathbb{W}}, \mathbb{W}) = DG_{\lambda_{\mathbb{W}}}(\lambda_{\mathbb{W}})$ and $\partial_{\mathbb{X}} \phi(\lambda_{\mathbb{W}}, \mathbb{W}) : \mathbb{MEP}(n_1, \cdots, n_m) \to \mathbb{C}^{m+1}$ is given by

$$\partial_{\mathbb{X}} \phi(\lambda_{\mathbb{W}}, \mathbb{W})\mathbb{H} = \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(\lambda_{\mathbb{W}}))H_1(\lambda(\lambda_{\mathbb{W}}))), & \cdots, & \mathrm{tr}(\mathrm{adj}(W_m(\lambda_{\mathbb{W}}))H_m(\lambda_{\mathbb{W}})), & 0 \end{bmatrix}^T.$$

Therefore by (4.14), we have

$$\begin{aligned} D\lambda(\mathbb{W})\mathbb{H} &= -(DG_{\lambda_{\mathbb{W}}}(\lambda_{\mathbb{W}}))^{-1} \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(\lambda_{\mathbb{W}}))H_1(\lambda_{\mathbb{W}})) \\ \vdots \\ \mathrm{tr}(\mathrm{adj}(X_m(\lambda_{\mathbb{W}}))H_m(\lambda_{\mathbb{W}})) \\ 0 \end{bmatrix} \\ &= -(DG_{\lambda_{\mathbb{W}}}(\lambda_{\mathbb{W}}))^{-1}(:, 1:m) \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(\lambda_{\mathbb{W}}))H_1(\lambda_{\mathbb{W}})) \\ \vdots \\ \mathrm{tr}(\mathrm{adj}(X_m(\lambda_{\mathbb{W}}))H_m(\lambda_{\mathbb{W}})) \end{bmatrix}. \end{aligned} \tag{4.15}$$

Note that $B = \begin{bmatrix} U & \lambda_{\mathbb{W}} \end{bmatrix}$ is a basis of $\mathbb{C}^{m+1}$. By (4.6) we have

$$(DG_{\lambda_{\mathbb{W}}}(\lambda_{\mathbb{W}})) \cdot B = \begin{bmatrix} J_{\mathbb{W}}(\lambda_{\mathbb{W}}) \\ \lambda_{\mathbb{W}}^* \end{bmatrix} \begin{bmatrix} U & \lambda_{\mathbb{W}} \end{bmatrix} = \begin{bmatrix} N_{\mathbb{W}}(\lambda_{\mathbb{W}}) & 0 \\ 0 & ||\lambda_{\mathbb{W}}||_2^2 \end{bmatrix}.$$

Consequently, we have

$$\begin{aligned} B^{-1}(DG_{\lambda_{\mathbb{W}}}(\lambda_{\mathbb{W}}))^{-1} &= \begin{bmatrix} (N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1} & 0 \\ 0 & \frac{1}{\|\lambda_{\mathbb{W}}\|_2^2} \end{bmatrix} \\ \Rightarrow \quad (DG_{\lambda_{\mathbb{W}}}(\lambda_{\mathbb{W}}))^{-1} &= \begin{bmatrix} U & \lambda_{\mathbb{W}} \end{bmatrix} \begin{bmatrix} (N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1} & 0 \\ 0 & \frac{1}{\|\lambda_{\mathbb{W}}\|_2^2} \end{bmatrix} \\ \Rightarrow \quad (DG_{\lambda_{\mathbb{W}}}(\lambda_{\mathbb{W}}))^{-1}(:, 1:m) &= U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}. \end{aligned}$$

Hence by (4.15) we have

$$D\lambda(\mathbb{W})\mathbb{H} = -U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1} \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(\lambda_{\mathbb{W}}))H_1(\lambda_{\mathbb{W}})) \\ \vdots \\ \mathrm{tr}(\mathrm{adj}(W_m(\lambda_{\mathbb{W}}))H_m(\lambda_{\mathbb{W}})) \end{bmatrix},$$

which proves (4.12).

Since $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ is simple, by Proposition 4.2.2, it follows that $\lambda_{\mathbb{W}}$ is geometrically simple. Then by Proposition 4.2.4, there exists left and right eigenvectors $\widehat{u} = \widehat{u}_1 \otimes \cdots \otimes \widehat{u}_m$ and $\widehat{v} = \widehat{v}_1 \otimes \cdots \otimes \widehat{v}_m$ of $\mathbb{W}$ corresponding to $\lambda_{\mathbb{W}}$, respectively, such that $\mathrm{adj}(W_i(\lambda_{\mathbb{W}})) = \widehat{v}_i\widehat{u}_i^*$ for all $i = 1, \cdots, m$ and $N_{\mathbb{W}}(\lambda_{\mathbb{W}}) = N_{\mathbb{W}}(\lambda_{\mathbb{W}}; \widehat{u}, \widehat{v})$. Also there exist nonzero scalars α_i, β_i such that $\widehat{u}_i = \alpha_i y_i$ and $\widehat{v}_i = \beta_i x_i i$ for $i = 1, \cdots, m$. As we have seen in the proof of Theorem 4.2.4, we have

$$N_{\mathbb{W}}(\lambda_{\mathbb{W}}) = N_{\mathbb{W}}(\lambda_{\mathbb{W}}; \widehat{u}, \widehat{v}) = \begin{bmatrix} \overline{\alpha_1}\beta_1 & & \\ & \ddots & \\ & & \overline{\alpha_m}\beta_m \end{bmatrix} N_{\mathbb{W}}(\lambda; y, x)$$

which gives

$$(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1} = (N_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1} \begin{bmatrix} \overline{\alpha_1}\beta_1 & & \\ & \ddots & \\ & & \overline{\alpha_m}\beta_m \end{bmatrix}^{-1}.$$

Again note that

$$\begin{aligned} \mathrm{tr}(\mathrm{adj}(W_i(\lambda_{\mathbb{W}}))H_i(\lambda_{\mathbb{W}})) &= \langle H_i(\lambda_{\mathbb{W}}), \mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^* \rangle = \langle H_i(\lambda_{\mathbb{W}}), u_i v_i^* \rangle \\ &= u_i^* H_i(\lambda_{\mathbb{W}}) v_i = (\alpha_i y_i)^* H_i(\lambda_{\mathbb{W}})(\beta_i x_i) = \overline{\alpha_i}\beta_i y_i^* H_i(\lambda_{\mathbb{W}}) x_i \end{aligned}$$

for all $i = 1, \cdots, m$. Thus we have

$$\begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(\lambda_{\mathbb{W}}))H_1(\lambda_{\mathbb{W}})) \\ \vdots \\ \mathrm{tr}(\mathrm{adj}(W_m(\lambda_{\mathbb{W}}))H_m(\lambda_{\mathbb{W}})) \end{bmatrix} = \begin{bmatrix} \overline{\alpha_1}\beta_1 & & \\ & \ddots & \\ & & \overline{\alpha_m}\beta_m \end{bmatrix} \begin{bmatrix} y_1^* H_1(\lambda_{\mathbb{W}})x_1 \\ \vdots \\ y_m^* H_m(\lambda_{\mathbb{W}})x_m \end{bmatrix}.$$

Therefore by (4.12), we have

$$D\lambda(\mathbb{W})\mathbb{H} = -U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1} \begin{bmatrix} y_1^* H_1(\lambda_{\mathbb{W}})x_1 \\ \vdots \\ y_m^* H_m(\lambda_{\mathbb{W}})x_m \end{bmatrix},$$

which proves (4.13). The rest of the proof is similar to that of Theorem 2.3.2. □

When $\mathbb{MEP}(n_1, \cdots, n_m)$ is equipped with the Hölder p-norm $|||\cdot|||_p$, we have the following result.

Corollary 4.3.2. *Let $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ be regular and $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ be a simple eigenvalue of $\mathbb{W}$. Let $x := x_1 \otimes \cdots \otimes x_m$ and $y := y_1 \otimes \cdots \otimes y_m$ be right and left eigenvectors of $\mathbb{W}$ corresponding to $\lambda_{\mathbb{W}}$, respectively.*

a) Then

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_p \; ||\lambda_{\mathbb{W}}^*||_q \; \max_i ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_*,$$

where $p^{-1} + q^{-1} = 1$. In particular, for the spectral or Frobenius norm on $\mathbb{C}^{n_i \times n_i}$, we have

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_p \; ||\lambda_{\mathbb{W}}^*||_q \; \max_i \prod_{k=1}^{n_i - 1} \sigma_k(W_i(\lambda_{\mathbb{W}})),$$

where $\sigma_1(W_i(\lambda_{\mathbb{W}})) \geq \cdots \geq \sigma_{n_i}(W_i(\lambda_{\mathbb{W}}))$ are singular values of $W_i(\lambda_{\mathbb{W}})$ for $i = 1, \cdots, m$.

b) Also we have

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_p \; ||\lambda_{\mathbb{W}}^*||_q \; \max_i ||y_i x_i^*||_*.$$

For a subordinate matrix norm $|| \cdot ||$ on $\mathbb{C}^{n_i \times n_i}$, we have

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_p \; ||\lambda_{\mathbb{W}}^*||_q \; \max_i ||y_i||_* ||x_i||.$$

In particular, for the spectral or Frobenius matrix norm on $\mathbb{C}^{n_i \times n_i}$, *we have*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq ||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_p \; ||\lambda_{\mathbb{W}}^*||_q \max_i ||y_i||_2 ||x_i||_2.$$

Remark 4.3.1. *When* $m = 1$, *the regular homogeneous MEP reduces to a regular homogenous matrix pencil* $\mathbb{W}(z) := z_1 A + z_2 B$, *where* $z = (z_1, z_2)^T \in \mathbb{C}^2$ *and* $A, B \in \mathbb{C}^{n_1 \times n_1}$. *For a simple eigenvalue* $\lambda_{\mathbb{W}} = (\lambda_1, \lambda_2)^T$ *of* $\mathbb{W}$, *consider* $u = \left(\frac{\overline{\lambda_2}}{||(\overline{\lambda_2}, -\overline{\lambda_1})||_V}, \frac{-\overline{\lambda_1}}{||(\overline{\lambda_2}, -\overline{\lambda_1})||_V} \right)^T$. *Then* $U = [u]$ *is a basis of* $\{\lambda_{\mathbb{W}}\}^{\perp}$. *Note that*

$$N_{\mathbb{W}}(\lambda_{\mathbb{W}}) = \mathrm{tr}(\mathrm{adj}(W(\lambda_{\mathbb{W}}))\mathbb{W}(u)) \neq 0 \text{ and } (N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1} = \frac{1}{\mathrm{tr}(\mathrm{adj}(W(\lambda_{\mathbb{W}}))\mathbb{W}(u))}.$$

Then $||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V = \frac{1}{|\mathrm{tr}(\mathrm{adj}(W(\lambda_{\mathbb{W}}))\mathbb{W}(u))|} \; ||u||_V = \frac{1}{|\mathrm{tr}(\mathrm{adj}(W(\lambda_{\mathbb{W}}))\mathbb{W}(u))|}$. *So it follows from Theorem 4.3.1 that*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) \leq \frac{||\lambda_{\mathbb{W}}||_{V,*} \; ||(\mathrm{adj}(\mathbb{W}(\lambda_{\mathbb{W}})))^*||_*}{|\mathrm{tr}(\mathrm{adj}(\mathbb{W}(\lambda_{\mathbb{W}}))\mathbb{W}(u))|}. \tag{4.16}$$

It is shown in [7] that the equality holds in (4.16). We, therefore, consider

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) = ||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \; ||\lambda_{\mathbb{W}}^*||_{V,*} \max_i ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_*$$

to be the condition number of a simple eigenvalue $\lambda_{\mathbb{W}}$ *of a regular nonhomogeneous MEP* $\mathbb{W}$. *Equivalently if* $x := x_1 \otimes \cdots \otimes x_m$ *and* $y := y_1 \otimes \cdots \otimes y_m$ *are right and left eigenvectors of* $\mathbb{W}$ *corresponding to* $\lambda_{\mathbb{W}}$, *respectively, then we have*

$$\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) = ||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}; y, x))^{-1}||_V \; ||\lambda_{\mathbb{W}}^*||_{V,*} \max_i ||y_i x_i^*||_*.$$

Considering the weighted norm $|||\cdot|||_{w,V}$ on $\mathbb{MEP}(n_1, \cdots, n_m)$, we have the following result.

Theorem 4.3.3. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular and* $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ *be simple. Consider the weighted norm* $|||\cdot|||_{w,V}$ *on* $\mathbb{MEP}(n_1, \cdots, n_m)$ *with nonzero positive weights* $w_i := ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_*$ *for* $i = 1, \cdots, m$. *Then* $\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) = ||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \; ||\lambda_{\mathbb{W}}^*||_{V,*}$.

In particular, for the weighted Hölder p-norm $|||\cdot|||_{w,p}$ *on* $\mathbb{MEP}(n_1, \cdots, n_m)$ *with nonzero positive weights* $w_i := ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_*$ *for* $i = 1, \cdots, m$, *we have* $\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) = ||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_p \; ||\lambda_{\mathbb{W}}^*||_q$, *where* $p^{-1} + q^{-1} = 1$.

Proof. By similar arguments as those in the proof of Theorem 2.3.4, it follows from (4.12) that $||D\lambda(\mathbb{W})\mathbb{H}||_V \leq ||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \; ||\lambda_{\mathbb{W}}^*||_{V,*} \; |||\mathbb{H}|||_{w,V}$. To prove the equality, consider $\mathbb{Y} \in \mathbb{MEP}(n_1, \cdots, n_m)$ such that $\mathbb{Y}(z) = (Y_1(z), \cdots, Y_m(z))$, $||Y_i||_V = 1$ and

$$\langle Y_i, \lambda_{\mathbb{W}}^* \otimes \mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^* \rangle = \mathrm{tr}(\mathrm{adj}(W_i(\lambda_{\mathbb{W}}))Y_i(\lambda_{\mathbb{W}})) = w_i||\lambda_{\mathbb{W}}^*||_{V,*},$$

where $\lambda_{\mathbb{W}}^* \otimes \mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^* : \mathbb{C}^{m+1} \to \mathbb{C}^{n_i \times n_i}$ is given by $(\lambda_{\mathbb{W}}^* \otimes \mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^*)(z) = (\lambda_{\mathbb{W}}^* z)\mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^*$ for $i = 1, \cdots, m$. Let $u = (u_1, \cdots, u_m)^T \in \mathbb{C}^m$ be such that $||u||_V = 1$ and $||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}u||_V = ||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V$. Construct $\mathbb{X} \in \mathbb{MEP}(n_1, \cdots, n_m)$ such that $\mathbb{X}(z) := \left(\frac{u_1}{w_1}Y_1(z), \cdots, \frac{u_m}{w_m}Y_m(z)\right)$. Then it is easy to see that $|||\mathbb{X}|||_{w,V} = 1$. Also by (4.12), we have $D\lambda(\mathbb{W})\mathbb{X} = ||\lambda_{\mathbb{W}}^*||_{V,*} \; U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}u$ and $||D\lambda(\mathbb{W})\mathbb{X}||_V = ||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V \; ||\lambda_{\mathbb{W}}^*||_{V,*}$. Hence $\mathrm{cond}(\lambda_{\mathbb{W}}, \mathbb{W}) = ||D\lambda(\mathbb{W})||_V = ||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V||\lambda_{\mathbb{W}}^*||_{V,*}$. This completes the proof. ☐

Next we define fast perturbation for a regular homogeneous MEP.

Definition 4.3.2. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular and* $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ *be simple. Then* $\Delta\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *is said to be a fast perturbation for* $\lambda_{\mathbb{W}}$ *if* $||D\lambda(\mathbb{W})\Delta\mathbb{W}||_V = ||D\lambda(\mathbb{W})|| \, |||\Delta\mathbb{W}|||_V$.

Now we construct a fast perturbation for a simple eigenvalue of a homogeneous MEP considering a weighted norm on $\mathbb{MEP}(n_1, \cdots, n_m)$.

Proposition 4.3.1. (Fast Perturbation) *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular and* $\lambda_{\mathbb{W}} \in \sigma(\mathbb{W})$ *be simple. Consider the weighted norm* $|||\cdot|||_{w,V}$ *on* $\mathbb{MEP}(n_1, \cdots, n_m)$ *with nonzero positive weights* $w_i = ||(\mathrm{adj}(W_i(\lambda_{\mathbb{W}})))^*||_*$ *for* $i = 1, \cdots, m$. *Let* $y = (y_1, \cdots, y_m)^T \in \mathbb{C}^m$ *be such that* $||y||_V = 1$ *and* $||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1} \; y||_V = ||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V$. *Let* $\mathbb{Y} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be such that* $\mathbb{Y}(z) := (Y_1(z), \cdots, Y_m(z))$, $||Y_i||_V = 1$ *and*

$$\mathrm{tr}(\mathrm{adj}(W_i(\lambda_{\mathbb{W}}))Y_i(\lambda_{\mathbb{W}})) = \langle Y_i, \lambda_{\mathbb{W}}^* \otimes \mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^* \rangle = w_i||\lambda_{\mathbb{W}}^*||_{V,*},$$

where $\lambda_{\mathbb{W}}^* \otimes \mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^* : \mathbb{C}^{m+1} \to \mathbb{C}^{n_i \times n_i}$ *is given by* $(\lambda_{\mathbb{W}}^* \otimes \mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^*)(z) = (\lambda_{\mathbb{W}}^* z)\mathrm{adj}(W_i(\lambda_{\mathbb{W}}))^*$ *for* $i = 1, \cdots, m$. *Consider* $\Delta\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *such that* $\Delta\mathbb{W}(z) := \left(\frac{y_1}{w_1}Y_1(z), \cdots, \frac{y_m}{w_m}Y_m(z)\right)$. *Then* $\Delta\mathbb{W}$ *is a fast perturbation for* $\lambda_{\mathbb{W}}$.

Proof. First note that $|||\Delta\mathbb{W}|||_{w,V} = \left|\left|\left(w_1\frac{|y_1|}{w_1}||Y_1||_V, \cdots, w_m\frac{|y_m|}{w_m}||Y_m||_V\right)\right|\right|_V = ||y||_V = 1.$ Also by (4.12), we have $D\lambda(\mathbb{W})\Delta\mathbb{W} = ||\lambda^*||_{V,*}U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}y$. Therefore $||D\lambda(\mathbb{W})\Delta\mathbb{W}||_V = ||\lambda^*||_{V,*}||U(N_{\mathbb{W}}(\lambda_{\mathbb{W}}))^{-1}||_V$. Hence the proof completes. ☐

4.4 Holomorphic perturbations

Let $\mathbb{W} : \mathbb{C}^p \to \mathbb{MEP}(n_1, \cdots, n_m)$ be holomorphic. Then $\mathbb{W}(t) \in \mathbb{MEP}(n_1, \cdots, n_m)$ for $t \in \mathbb{C}^p$. We write $\mathbb{W}(t)(z)$ as $\mathbb{W}(t,z) := (W_1(t,z), \cdots, W_m(t,z))$. Then $\mathbb{W}(t,z)$ is said to be regular if for each $t \in \mathbb{C}^p$ there is a $z \in \mathbb{C}^{m+1}$ such that $\prod_{i=1}^m \det(W_i(t,z)) \neq 0$. We refer to $\mathbb{W}(t,z)$ as a regular holomorphic MEP. Define $\mathcal{F} : \mathbb{C}^p \times \mathbb{C}^{m+1} \to \mathbb{C}^m$ by $\mathcal{F}(t,z) := (\det(W_1(t,z)), \cdots, \det(W_m(t,z)))^T$. Recall that a nonzero $\lambda \in \mathbb{C}^{m+1}$ is said to be an eigenvalue of $\mathbb{W}(t,z)$ if $\mathcal{F}(t,\lambda) = 0$. Then $x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ is said to be a right eigenvector of $\mathbb{W}(t,z)$ corresponding to λ if $W_i(t,\lambda)x_i = 0$ for all $i = 1, \cdots, m$. Similarly $y_1 \otimes \cdots \otimes y_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ is said to be a right eigenvector of $\mathbb{W}(t,z)$ corresponding to λ if $y_i^*W_i(t,\lambda) = 0$ for all $i = 1, \cdots, m$. An eigenvalue λ of $\mathbb{W}(t,z)$ is said to be geometrically simple if $\prod_{i=1}^m \dim\ker(W_j(t,\lambda)) = 1$. An eigenvalue λ of $\mathbb{W}(t,z)$ is said to be simple if the Jacobian matrix $J_{\mathbb{W}(t,\lambda)}(\lambda)$ of the map $\mathbb{C}^{m+1} \to \mathbb{C}^m, z \mapsto \mathcal{F}(t,z)$ given by $J_{\mathbb{W}(t,\lambda)}(\lambda) := \left[\mathrm{tr}(\mathrm{adj}(W_i(t,\lambda))\partial_{\lambda_j}(W_i(t,\lambda(t))))\right]_{m\times(m+1)}$ is a full-rank matrix.

Let $\lambda \in \mathbb{C}^{m+1}$ be an eigenvalue of $\mathbb{W}(t,z)$. Let $U := \begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}$ be a basis of $\{\lambda\}^\perp$. Define $N_{\mathbb{W}(t,\lambda)}(\lambda) := J_{\mathbb{W}(t,\lambda)}(\lambda)U$. We refer to $N_{\mathbb{W}(t,\lambda)}(\lambda)$ as the normal Jacobian matrix of the map $\{\lambda\}^\perp \to \mathbb{C}^m, u \mapsto J_{\mathbb{W}(t,\lambda)}(\lambda)U$. For left and right eigenvectors $y := y_1 \otimes \cdots \otimes y_m$ and $x := x_1 \otimes \cdots \otimes x_m$ corresponding to an eigenvalue λ, respectively, define $N_{\mathbb{W}(t,\lambda)}(t,\lambda; y, x) := [y_i^* W_i(t,u_j))x_i]_{m\times m}$.

Proposition 4.4.1. *Let* $\mathbb{W} : \mathbb{C}^p \to \mathbb{MEP}(n_1, \cdots, n_m)$ *be holomorphic such that* $\mathbb{W}(t,z)$ *is regular for all* $t \in \mathbb{C}^p$. *Let* λ *be a geometrically simple eigenvalue of* $\mathbb{W}(t,z)$. *Then there exist left and right eigenvectors* $\widehat{u} := \widehat{u}_1 \otimes \cdots \otimes \widehat{u}_m$ *and* $\widehat{v} := \widehat{v}_1 \otimes \cdots \otimes \widehat{v}_m$ *corresponding to* λ, *respectively, such that* $N_{\mathbb{W}(t,\lambda)}(\lambda) = N_{\mathbb{W}(t,\lambda)}(t,\lambda;\widehat{u},\widehat{v})$.

Proof. Since $\mathrm{rank}(W_i(t,\lambda)) = n_i - 1$ for all $i = 1, \cdots, m$, by Proposition 1.1.1 there exist nonzero vectors $\widehat{u}_i, \widehat{v}_i \in \mathbb{C}^{n_i}$ such that $\mathrm{adj}(W_i(t,\lambda)) = \widehat{v}_i\widehat{u}_i^*$, where $W_i(\lambda)\widehat{v}_i = 0$ and $\widehat{u}_i^*W_i(\lambda) = 0$ for all $i = 1, \cdots, m$. The rest of the proof is similar to that of Proposition 2.2.1. ☐

Next we have the following result.

Theorem 4.4.1. *Let* $\mathbb{W} : \mathbb{C}^p \to \mathbb{MEP}(n_1, \cdots, n_m)$ *be a holomorphic map such that* $\mathbb{W}(t, z)$ *is regular for all* $t \in \mathbb{C}^p$*. Let* $\widehat{\lambda}$ *be a simple eigenvalue of* $\mathbb{W}(\widehat{t}, z)$ *for some* $\widehat{t} \in \mathbb{C}^p$*. Let* $x := x_1 \otimes \cdots \otimes x_m$ *and* $y := y_1 \otimes \cdots \otimes y_m$ *be right and left eigenvector of* $\mathbb{W}(\widehat{t}, z)$ *corresponding to the eigenvalue* $\widehat{\lambda}$*, respectively. Then there is an open neighborhood* $\mathrm{nbd}(\widehat{t}) \subset \mathbb{C}^p$ *containing* $\widehat{t}$ *and a holomorphic function* $\lambda : \mathrm{nbd}(\widehat{t}) \to \mathbb{C}^{m+1}$ *such that* $\lambda(\widehat{t}) = \widehat{\lambda}$ *and* $\lambda(t)$ *is a simple eigenvalue of* $\mathbb{W}(t, z)$ *for all* $t \in \mathrm{nbd}(\widehat{t})$*. Further, let* $U(t)$ *be a basis of* $\{\lambda(t)\}^\perp$ *and* $N_{\mathbb{W}(t,\lambda(t))}(\lambda(t)) := J_{\mathbb{W}(t,\lambda)}(\lambda) \cdot U(t)$ *be the normal Jacobian matrix of the function* $z \mapsto \mathcal{F}(t, z)$ *at* $\lambda(t)$ *for all* $t \in \mathrm{nbd}(\widehat{t})$*. Then* $\lambda(t + h) = \lambda(t) + D\lambda(t)h + \mathcal{O}(||h||_2^2)$ *for sufficiently small* $||h||_2$*, where*

$$D\lambda(t)h = -U(t)(N_{\mathbb{W}(t,\lambda(t))}(\lambda(t)))^{-1} \begin{bmatrix} \sum_{j=1}^{p} \mathrm{tr}\left(\mathrm{adj}(W_1(t, \lambda(t)))\partial_{t_j}(W_1(t, \lambda(t)))\right) h_j \\ \vdots \\ \sum_{j=1}^{p} \mathrm{tr}\left(\mathrm{adj}(W_m(t, \lambda(t)))\partial_{t_j}(W_m(t, \lambda(t)))\right) h_j \end{bmatrix}, \tag{4.17}$$

for all $h = (h_1, \cdots, h_p)^T \in \mathbb{C}^p$*. In particular, we have*

$$\frac{\partial\lambda(\widehat{t})}{\partial t_j} = -U(\widehat{t})(N_{\mathbb{W}(\widehat{t},\widehat{\lambda})}(\widehat{\lambda}))^{-1} \begin{bmatrix} \mathrm{tr}\left(\mathrm{adj}(W_1(\widehat{t}, \widehat{\lambda}))\partial_{t_j}(W_1(\widehat{t}, \widehat{\lambda}))\right) \\ \vdots \\ \mathrm{tr}\left(\mathrm{adj}(W_m(\widehat{t}, \widehat{\lambda}))\partial_{t_j}(W_1(\widehat{t}, \widehat{\lambda}))\right) \end{bmatrix} \tag{4.18}$$

$$= -U(\widehat{t})(N_{\mathbb{W}(\widehat{t},\widehat{\lambda})}(\widehat{t}, \widehat{\lambda}; y, x))^{-1} \begin{bmatrix} y_1^* \partial_{t_j}(W_1(\widehat{t}, \widehat{\lambda}))x_1 \\ \vdots \\ y_m^* \partial_{t_j}(W_1(\widehat{t}, \widehat{\lambda}))x_m \end{bmatrix}. \tag{4.19}$$

Proof. Define $\phi : \mathbb{C}^p \times \mathbb{C}^{m+1} \to \mathbb{C}^{m+1}$ by $\phi(t, z) = \begin{bmatrix} \mathcal{F}(t, z) \\ \langle z - \widehat{\lambda}, \widehat{\lambda}\rangle \end{bmatrix}$ for $t \in \mathbb{C}^p$ and $z \in \mathbb{C}^{m+1}$. Note that $\phi(\widehat{t}, \widehat{\lambda}) = 0$ and $\partial_{\widehat{\lambda}}(\phi(\widehat{t}), \widehat{\lambda}) = DG_{\widehat{\lambda}}(\widehat{\lambda})$ is nonsingular by Theorem 4.2.1 since $\widehat{\lambda}$ is a simple eigenvalue of $\mathbb{W}(\widehat{t}, z)$. Hence by Implicit Function Theorem there is an open neighborhood $\mathrm{nbd}(\widehat{t})$ containing $\widehat{t}$ and a holomorphic function $\lambda : \mathrm{nbd}(\widehat{t}) \to \mathbb{C}^{m+1}$ such that $\lambda(\widehat{t}) = \widehat{\lambda}$ and $\phi(t, \lambda(t)) = 0$ for all $t \in \mathrm{nbd}(\widehat{t})$. By similar arguments as those in the proof of Theorem 4.3.1 it follows that $\lambda(t)$ is a simple eigenvalue of $\mathbb{W}(t, z)$ for all $t \in \mathrm{nbd}(\widehat{t})$. Now differentiating $\phi(\lambda(t), t) = 0$ with respect to t we have

$$\partial_\lambda(\phi(\lambda(t), t))D\lambda(t)h + \partial_t(\phi(\lambda(t), t))h = 0 \tag{4.20}$$

for all $h = (h_1, \cdots, h_p)^T \in \mathbb{C}^p$. Note that $\partial_\lambda(\phi(\lambda(t), t)) = DG_{\lambda(t)}(\lambda(t))$ is nonsingular and using $DW_i(t, \lambda(t))h = \sum_{j=1}^{p} \partial_{t_j}(W_i(t, \lambda(t)))h_j$, we have $\partial_t(\phi(\lambda(t), t)) : \mathbb{C}^p \to \mathbb{C}^{m+1}$ given by

$$\begin{aligned}
\partial_t(\phi(\lambda(t), t))h &= \begin{bmatrix} \mathrm{tr}(\mathrm{adj}(W_1(t, \lambda(t)))DW_1(t, \lambda(t))h) \\ \vdots \\ \mathrm{tr}(\mathrm{adj}(W_m(t, \lambda(t)))DW_m(t, \lambda(t))h) \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} \sum_{j=1}^{p} \mathrm{tr}\left(\mathrm{adj}(W_1(t, \lambda(t)))\partial_{t_j}(W_1(t, \lambda(t)))\right) h_j \\ \vdots \\ \sum_{j=1}^{p} \mathrm{tr}\left(\mathrm{adj}(W_m(t, \lambda(t)))\partial_{t_j}(W_m(t, \lambda(t)))\right) h_j \\ 0 \end{bmatrix}.
\end{aligned}$$

So by (4.20), we have

$$\begin{aligned}
D\lambda(t)h &= (DG_{\lambda(t)}(\lambda(t)))^{-1} \begin{bmatrix} \sum_{j=1}^{p} \mathrm{tr}\left(\mathrm{adj}(W_1(t, \lambda(t)))\partial_{t_j}(W_1(t, \lambda(t)))\right) h_j \\ \vdots \\ \sum_{j=1}^{p} \mathrm{tr}\left(\mathrm{adj}(W_m(t, \lambda(t)))\partial_{t_j}(W_m(t, \lambda(t)))\right) h_j \\ 0 \end{bmatrix} \\
&= (DG_{\lambda(t)}(\lambda(t)))^{-1}(:, 1:m) \begin{bmatrix} \sum_{j=1}^{p} \mathrm{tr}\left(\mathrm{adj}(W_1(t, \lambda(t)))\partial_{t_j}(W_1(t, \lambda(t)))\right) h_j \\ \vdots \\ \sum_{j=1}^{p} \mathrm{tr}\left(\mathrm{adj}(W_m(t, \lambda(t)))\partial_{t_j}(W_m(t, \lambda(t)))\right) h_j \end{bmatrix}.
\end{aligned}$$

By similar arguments as those in the proof of Theorem 4.3.1, we have $(DG_{\lambda(t)}(\lambda(t)))^{-1}(:, 1:m) = U(t)(N_{\mathbb{W}(t,\lambda(t))}(\lambda(t)))^{-1}$, which proves (4.17). Then (4.18) and (4.19) follow from (4.17). □

4.5 Backward perturbation analysis

We now derive backward error of an approximate eigenpair and an approximate eigenvalue of a regular homogeneous MEP.

Definition 4.5.1. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular. Let* $\lambda \in \mathbb{C}^{m+1} \setminus \{0\}$ *and* $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$*. Then the backward error of* (λ, x) *as an approximate eigenpair of* $\mathbb{W}$ *is defined by*

$$\eta(\lambda, x, \mathbb{W}) = \inf \, \{\|\Delta\mathbb{W}\|_V : \mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0\}$$

and the backward error of λ *as an approximate eigenvalue of* $\mathbb{W}$ *is defined by*

$$\eta(\lambda, \mathbb{W}) = \inf \, \{\|\Delta\mathbb{W}\|_V : \lambda \in \sigma(\mathbb{W} + \Delta\mathbb{W})\}.$$

First we prove the following result.

Proposition 4.5.1. *Let* $W_k : \mathbb{C}^{m+1} \to \mathbb{C}^{n_k \times n_k}$ *be given by* $W_k(z) := z_0 A_k + \sum_{j=1}^m z_j B_{kj}$*, where* $A_k, B_{kj} \in \mathbb{C}^{n_k \times n_k}$ *for all* $j = 1, \cdots, m$*. Let* $\lambda \in \mathbb{C}^{m+1} \setminus \{0\}$ *and* $0 \neq x_k \in \mathbb{C}^{n_k}$*. Define*

$$\eta(\lambda, x_k, W_k) = \inf \, \{||\Delta W_k||_V : W_k(\lambda)x_k + \Delta W_k(\lambda)x_k = 0\}. \tag{4.21}$$

Then for a subordinate matrix norm on $\mathbb{C}^{n_k \times n_k}$*, we have* $\eta(\lambda, x_k, W_k) = \frac{1}{||\lambda||_{V,*}} \frac{||r_k||}{||x_k||}$*, where* $r_k = W_k(\lambda)x_k$*.*

Let $y_k \in \partial||x_k||$ *and* $t := (t_0, \cdots, t_m)^T \in \partial||\lambda||_{V,*}$*. Define*

$$\Delta A_k := -\frac{\overline{t_0}}{||\lambda||_{V,*}} \frac{r_k y_k^*}{||x_k||}, \;\; \Delta B_{kj} := -\frac{\overline{t_j}}{||\lambda||_{V,*}} \frac{r_k y_k^*}{||x_k||}$$

for $j = 1, \cdots, m$ *and consider* $\Delta W_k(z) := z_0 \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$*. Then* $W_k(\lambda)x_k + \Delta W_k(\lambda)x_k = 0$ *and* $||\Delta W_k||_V = \eta(\lambda, x_k, W_k)$*.*

Proof. By similar arguments as those in the proof of Proposition 2.5.1, we have $\eta(\lambda, x_k, W_k) \geq \frac{1}{||\lambda||_{V,*}} \frac{||r_k||}{||x_k||}$. Note that $\Delta W_k(\lambda)x_k = -\frac{\sum_{j=1}^m \lambda_j \overline{t_j}}{||\lambda||_{V,*}} \frac{r_k y_k^* x_k}{||x_k||} = -r_k$, i.e., $W_k(\lambda)x_k + \Delta W_k(\lambda)x_k = 0$ and $||\Delta A_k|| = \frac{|t_0|}{||\lambda||_{V,*}} \frac{||r_k||}{||x_k||}$, $||\Delta B_{kj}|| = \frac{|t_j|}{||\lambda||_{V,*}} \frac{||r_k||}{||x_k||}$ for $j = 1, \cdots, m$. Now $||\Delta W_k||_V = \frac{||t||_V}{||\lambda||_{V,*}} \frac{||r_k||}{||x_k||} = \frac{1}{||\lambda||_{V,*}} \frac{||r_k||}{||x_k||}$ shows that $\eta(\lambda, x_k, W_k) = \frac{1}{||\lambda||_{V,*}} \frac{||r_k||}{||x_k||}$. ☐

Next we derive $\eta(\lambda, x, \mathbb{W})$ by considering a subordinate matrix norm on $\mathbb{C}^{n_i \times n_i}$.

Theorem 4.5.1. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular,* $\lambda \in \mathbb{C}^{m+1} \setminus \{0\}$ *and* $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$*. Consider* $r_k := W_k(\lambda)x_k$ *for all* $k = 1, \cdots, m$*. Then for a subordinate matrix norm on* $\mathbb{C}^{n_k \times n_k}$*, we have*

$$\eta(\lambda, x, \mathbb{W}) = ||(\eta(\lambda, x_1, W_1), \cdots, \eta(\lambda, x_m, W_m))||_V = \frac{1}{||\lambda||_{V,*}} \left|\left| \left(\frac{||r_1||}{||x_1||}, \cdots, \frac{||r_m||}{||x_m||} \right) \right|\right|_V, \tag{4.22}$$

where $\eta(\lambda, x_k, W_k)$ *is as defined in (4.21) for all* $k = 1, \cdots, m$. *In particular, for the Hölder p-norm on* $\mathbb{MEP}(n_1, \cdots, n_m)$, *we have* $\eta(\lambda, x, \mathbb{W}) = \frac{1}{||\lambda||_q} \left|\left| \left(\frac{||r_1||}{||x_1||}, \cdots, \frac{||r_m||}{||x_m||} \right) \right|\right|_p$, *where* $p^{-1} + q^{-1} = 1$.

Let $y_k \in \partial ||x_k||$ *and* $t := (t_0, \cdots, t_m)^T \in \partial ||\lambda||_{V,*}$. *Define*

$$\Delta A_k := -\frac{\overline{t_0}}{||\lambda||_{V,*}} \frac{r_k y_k^*}{||x_k||}, \ \Delta B_{kj} := -\frac{\overline{t_j}}{||\lambda||_{V,*}} \frac{r_k y_k^*}{||x_k||}$$

for $j = 1, \cdots, m$. *Consider* $\Delta W_k(z) := z_0 \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$ *for all* $k = 1, \cdots, m$ *and* $\Delta \mathbb{W}(z) := (\Delta W_1(z), \cdots, \Delta W_m(z))$. *Then* $\mathbb{W}(\lambda)x + \Delta \mathbb{W}(\lambda)x = 0$ *and* $\|\Delta \mathbb{W}\|_V = \eta(\lambda, x, \mathbb{W})$.

Proof. By similar arguments, as those in the proof of Theorem 2.5.1, we have $\eta(\lambda, x, \mathbb{W}) \geq ||(\eta(\lambda, x_1, W_1), \cdots, \eta(\lambda, x_m, W_m))||_V$. Again by Proposition 4.5.1, it follows that $W_k(\lambda)x_k + \Delta W_k(\lambda)x_k = 0$ and $||\Delta W_k||_V = \eta(\lambda, x_k, W_k)$ for all $k = 1, \cdots, m$. This shows that $\|\Delta \mathbb{W}\|_V = ||(\eta(\lambda, x_1, W_1), \cdots, \eta(\lambda, x_m, W_m))||_V$. Hence

$$\eta(\lambda, x, \mathbb{W}) = ||(\eta(\lambda, x_1, W_1), \cdots, \eta(\lambda, x_m, W_m))||_V$$

and the last equality in (4.22) follows from Proposition 4.5.1. ☐

Next we derive backward error corresponding to an approximate eigenvalue.

Proposition 4.5.2. *Let* $W_k : \mathbb{C}^{m+1} \to \mathbb{C}^{n_k \times n_k}$ *be given by* $W_k(z) := z_0 A_k + \sum_{j=1}^m z_j B_{kj}$, *where* $A_k, B_{kj} \in \mathbb{C}^{n_k \times n_k}$ *for all* $j = 1, \cdots, m$. *Let* $\lambda \in \mathbb{C}^{m+1} \setminus \{0\}$ *and define*

$$\eta(\lambda, W_k) = \inf \{||\Delta W_k||_V : \mathrm{rank}(W_k(\lambda) + \Delta W_k(\lambda)) < n_k\}. \tag{4.23}$$

Then for a subordinate matrix norm on $\mathbb{C}^{n_k \times n_k}$, *we have* $\eta(\lambda, W_k) = \frac{r_k}{||\lambda||_{V,*}}$, *where* $r_k = \min_{||x_k||=1} \{||W_k(\lambda)x_k|| : x_k \in \mathbb{C}^{n_k}\}$.

Let $x_k \in \mathbb{C}^{n_k}$ *be such that* $||x_k|| = 1$ *and* $r_k = ||W_k(\lambda)x_k||$. *Let* $y_k \in \partial ||x_k||$ *and* $t := (t_0, \cdots, t_m)^T \in \partial ||\lambda||_{V,*}$. *Define*

$$\Delta A_k := -\frac{\overline{t_0}}{||\lambda||_{V,*}} W_k(\lambda) x_k y_k^*, \ \Delta B_{kj} := -\frac{\overline{t_j}}{||\lambda||_{V,*}} W_k(\lambda) x_k y_k^*$$

for $j = 1, \cdots, m$ *and consider* $\Delta W_k(z) := z_0 \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$. *Then* $W_k(\lambda)x_k + \Delta W_k(\lambda)x_k = 0$ *and* $||\Delta W_k||_V = \eta(\lambda, W_k)$.

Proof. By Proposition 2.5.2, we have $\eta(\lambda, W_k) \geq \frac{r_k}{||\lambda||_{V,*}}$. It is also easy to see that $(W_k(\lambda) + \Delta W_k(\lambda))x_k = 0$, i.e, $\text{rank}(W_k(\lambda) + \Delta W_k(\lambda)) < n_k$ and $||\Delta A_k|| = \frac{|t_0|\ r_k}{||\lambda||_{V,*}}$, $||\Delta B_{kj}|| = \frac{|t_j|\ r_k}{||\lambda||_{V,*}}$ for $j = 1, \cdots, m$. Consequently, we have $||\Delta W_k||_V = \frac{r_k}{||\lambda^*||_{V,*}}$ and hence $\eta(\lambda, W_k) = \frac{r_k}{||\lambda^*||_{V,*}}$. ☐

Now we derive $\eta(\lambda, \mathbb{W})$.

Theorem 4.5.2. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular and* $\lambda \in \mathbb{C}^{m+1} \setminus \{0\}$. *Consider* $r_k := \min_{||x_k||=1} \{||W_k(\lambda)x_k|| : x_k \in \mathbb{C}^{n_k}\}$ *for* $k = 1, \cdots, m$. *Then for a subordinate matrix norm on* $\mathbb{C}^{n_k \times n_k}$, *we have*

$$\eta(\lambda, \mathbb{W}) = ||(\eta(\lambda, W_1), \cdots, \eta(\lambda, W_m))||_V = \frac{1}{||\lambda||_{V,*}} ||(r_1, \cdots, r_m)||_V,$$

where $\eta(\lambda, W_k)$ *is as defined in (4.23) for all* $k = 1, \cdots, m$. *In particular, for the Hölder p-norm on* $\mathbb{MEP}(n_1, \cdots, n_m)$, $\eta(\lambda, \mathbb{W}) = \frac{1}{||\lambda||_q} ||(r_1, \cdots, r_m)||_p$, *where* $p^{-1} + q^{-1} = 1$.

Let $x_k \in \mathbb{C}^{n_k}$ *be such that* $||x_k|| = 1$ *and* $r_k = ||W_k(\lambda)x_k||$. *Let* $y_k \in \partial||x_k||$ *and* $t := (t_0, \cdots, t_m)^T \in \partial||\lambda||_{V,*}$. *Define*

$$\Delta A_k := -\frac{\overline{t_0}}{||\lambda||_{V,*}} W_k(\lambda)x_k y_k^*, \ \Delta B_{kj} := -\frac{\overline{t_j}}{||\lambda||_{V,*}} W_k(\lambda)x_k y_k^*$$

for $j = 1, \cdots, m$. *Consider* $\Delta W_k(z) := z_0 \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$ *for* $k = 1, \cdots, m$ *and* $\Delta\mathbb{W}(z) = (\Delta W_1(z), \cdots, \Delta W_m(z))$. *Then* $W_k(\lambda)x_k + \Delta W_k(\lambda)x_k = 0$ *for all* $k = 1, \cdots, m$ *and* $\|\Delta\mathbb{W}\|_V = \eta(\lambda, \mathbb{W})$.

Proof. By Theorem 2.5.2, we have $\eta(\lambda, \mathbb{W}) \geq ||(\eta(\lambda, W_1), \cdots, \eta(\lambda, W_m))||_V$. Further by Proposition 4.5.2, it follows that $W_k(\lambda)x_k + \Delta W_k(\lambda)x_k = 0$ and $||\Delta W_k||_V = \eta(\lambda, W_k)$ for all $k = 1, \cdots, m$. Hence $\|\Delta\mathbb{W}\|_V = ||(\eta(\lambda, W_1), \cdots, \eta(\lambda, W_m))||_V$. By Proposition 4.5.2, we have $\eta(\lambda, \mathbb{W}) = ||(\eta(\lambda, W_1), \cdots, \eta(\lambda, W_m))||_V = \frac{1}{||\lambda||_{V,*}} ||(r_1, \cdots, r_m)||_V$. ☐

For the spectral or Frobenious norm on $\mathbb{C}^{n_k \times n_k}$, we have the following result.

Theorem 4.5.3. *Let* $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *be regular and* $\lambda \in \mathbb{C}^{m+1} \setminus \{0\}$. *Then for the spectral or Frobenious norm on* $\mathbb{C}^{n_i \times n_i}$, *we have*

$$\eta(\lambda, \mathbb{W}) = \frac{1}{||\lambda||_{V,*}} \ ||(\sigma_{\min}(W_1(\lambda)), \cdots, \sigma_{\min}(W_m(\lambda)))||_V ,$$

where $\sigma_{\min}(W_k(\lambda))$ *is the smallest singular value of* $W_k(\lambda)$ *for all* $k = 1, \cdots, m$. *In particular, for the Hölder p-norm on* $\mathbb{MEP}(n_1, \cdots, n_m)$ *we have*

$$\eta(\lambda, \mathbb{W}) = \frac{1}{||\lambda||_q} \; ||(\sigma_{\min}(W_1(\lambda)), \cdots, \sigma_{\min}(W_m(\lambda)))||_p \,,$$

where $p^{-1} + q^{-1} = 1$.

Let $t = (t_0, \cdots, t_m)^T \in \partial||\lambda||_{V,*}$. *Consider the* SVD $W_k(\lambda) = U_k \Sigma_k V_k$ *and set* $u_k := U_k(:, n_k)$ *and* $v_k := V_k(:, n_k)$ *for* $k = 1, \cdots, m$. *Define*

$$\Delta A_k := -\frac{\overline{t_0}}{||\lambda||_{V,*}} \, \sigma_{\min}(W_k(\lambda)) u_k v_k^*, \; \Delta B_{kj} := -\frac{\overline{t_j}}{||\lambda||_{V,*}} \, \sigma_{\min}(W_k(\lambda)) u_k v_k^*$$

for all $k, j = 1, \cdots, m$. *Consider* $\Delta W_k(z) := z_0 \Delta A_k + \sum_{j=1}^{m} z_j \Delta B_{kj}$ *for* $k = 1, \cdots, m$ *and* $\Delta\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ *such that* $\Delta\mathbb{W}(z) := (\Delta W_1(z), \cdots, \Delta W_m(z))$. *Then* $\lambda \in \sigma(\mathbb{W} + \Delta\mathbb{W})$ *and* $\|\Delta\mathbb{W}\|_V = \eta(\lambda, \mathbb{W})$.

Proof. The proof is similar to that of Theorem 2.5.3. □


भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

5

Vector Space of Linearizations of a two-parameter PEP

5.1 Introduction

Linearization is a standard method for solving a polynomial eigenvalue problem (PEP). Linearization is a process which transforms a polynomial eigenvalue problem to a generalized eigenvalue problem of larger size.

Let $P(\lambda) = \sum_{j=0}^{k} A_j \lambda^j$, where $A_j \in \mathbb{C}^{n\times n}$ for $j = 0, \cdots, k$, be a matrix polynomial of degree k. Then the matrix pencils

$$\begin{aligned} C_1(\lambda) &:= \lambda \left[\begin{array}{c|ccc} A_k & & & \\ \hline & I_n & & \\ & & \ddots & \\ & & & I_n \end{array}\right] + \left[\begin{array}{ccc|c} A_{k-1} & A_{k-2} & \cdots & A_0 \\ \hline -I_n & & & \\ & \ddots & & \\ & & -I_n & \end{array}\right] \\ C_2(\lambda) &:= \lambda \left[\begin{array}{c|ccc} A_k & & & \\ \hline & I_n & & \\ & & \ddots & \\ & & & I_n \end{array}\right] + \left[\begin{array}{c|ccc} A_{k-1} & -I_n & & \\ A_{k-2} & & \ddots & \\ \vdots & & & -I_n \\ \hline A_0 & & & \end{array}\right], \end{aligned}$$

known as the first and second companion form of $P(\lambda)$, respectively, are linearization of $P(\lambda)$. Generalizing the companion forms, the following two classes of vector spaces of linearizations were characterized in [42]:

$$\begin{aligned}\mathbb{L}_1(P) &:= \{L(\lambda) : L(\lambda)\cdot(\Lambda\otimes I_n) = v\otimes P(\lambda), v\in\mathbb{C}^k\},\\ \mathbb{L}_2(P) &:= \{L(\lambda) : (\Lambda^T\otimes I_n)\cdot L(\lambda) = w^T\otimes P(\lambda), w\in\mathbb{C}^k\},\end{aligned}$$

where $\Lambda := (\lambda^{k-1}, \lambda^{k-2}, \cdots, \lambda, 1)^T$. For a two-parameter polynomial eigenvalue problem, a few linearizations are proposed in [34, 44]. Also a vector space of linearizations for a two-parameter quadratic eigenvalue problem is provided in [1]. In this chapter, our main objective is to construct linearizations of a two-parameter polynomial eigenvalue problem. Consider the two-parameter matrix polynomial $P(\lambda,\mu)$ given by

$$P(\lambda,\mu) := \sum_{i=0}^{k}\sum_{j=0}^{k-i}\lambda^i\mu^j A_{ij}, \quad \text{where } A_{ij}\in\mathbb{C}^{n\times n}.$$

Define

$$\begin{aligned}K_{ij}(\lambda,\mu) &:= A_{ij} \quad \text{for } i+j<k-1,\\ K_{ij}(\lambda,\mu) &:= A_{ij}+\lambda A_{i+1,j} \quad \text{for } i+j=k-1, i\neq 0,\\ K_{0,k-1}(\lambda,\mu) &:= A_{0,k-1}+\lambda A_{1,k-1}+\mu A_{0,k},\\ K_r &:= \begin{bmatrix} K_{r0} & K_{r-1,1} & \cdots & K_{1,r-1} & K_{0r}\end{bmatrix} \quad \text{for } r=0,\cdots,k,\\ T_r &:= \left[\begin{array}{ccc} \lambda I_n & & \\ & \ddots & \\ & & \lambda I_n \\ \hline & & \mu I_n \end{array}\right] \quad \text{for } r=1,\cdots,k.\end{aligned}$$

Then it is proved in [44] that

$$L(\lambda,\mu) := \begin{bmatrix} K_0 & K_1 & \cdots & K_k \\ T_1 & -I_{2n} & & \\ & \ddots & \ddots & \\ & & T_k & -I_{kn}\end{bmatrix}$$

is a linearization of $P(\lambda,\mu)$. For the two-parameter quadratic matrix polynomial $Q(\lambda,\mu) = \lambda^2 A_{20} + \lambda\mu A_{11} + \mu^2 A_{02} + \lambda A_{10} + \mu A_{01} + A_{00}$, the linearization proposed in [1] is given by

$$\lambda \begin{bmatrix} A_{20} & A_{11} & 0_n \\ 0_n & 0_n & 0_n \\ 0_n & 0_n & I_n \end{bmatrix} + \mu \begin{bmatrix} 0_n & A_{02} & 0_n \\ 0_n & 0_n & I_n \\ 0_n & 0_n & 0_n \end{bmatrix} + \begin{bmatrix} A_{10} & A_{01} & A_{00} \\ 0_n & -I_n & 0_n \\ -I_n & 0_n & 0_n \end{bmatrix} \tag{5.1}$$

and the linearization proposed in [44] is given by

$$\lambda \begin{bmatrix} 0_n & A_{20} & A_{11} \\ I_n & 0_n & 0_n \\ 0_n & 0_n & 0_n \end{bmatrix} + \mu \begin{bmatrix} 0_n & 0_n & A_{02} \\ 0_n & 0_n & 0_n \\ I_n & 0_n & 0_n \end{bmatrix} + \begin{bmatrix} A_{00} & A_{10} & A_{01} \\ 0_n & -I_n & 0_n \\ 0_n & 0_n & -I_n \end{bmatrix}. \tag{5.2}$$

In this chapter, we characterize the following two vector spaces of linear two-parameter matrix polynomials of the foem $L(\lambda,\mu) = \lambda A + \mu B + C$:

$$\begin{aligned} \mathbb{L}_1(P) &:= \left\{ L(\lambda,\mu) : L(\lambda,\mu)\cdot(\Lambda \otimes I_n) = v \otimes P(\lambda,\mu), v \in \mathbb{C}^{\frac{k(k+1)}{2}} \right\}, \\ \mathbb{L}_2(P) &:= \left\{ L(\lambda,\mu) : (\Lambda^T \otimes I_n)\cdot L(\lambda,\mu) = w^T \otimes P(\lambda,\mu), w \in \mathbb{C}^{\frac{k(k+1)}{2}} \right\}. \end{aligned}$$

We refer to $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ as the right ansatz space and left ansatz space, respectively. We derive the linearization conditions for an element in the right or left ansatz space.

5.2 Ansatz spaces for two-parameter polynomials

Consider a two-parameter matrix polynomial $P(\lambda,\mu)$ of degree k given by

$$P(\lambda,\mu) := \sum_{i=0}^{k}\sum_{j=0}^{k-i} \lambda^i \mu^j A_{ij} = \sum_{j=0}^{k} P_j(\lambda,\mu), \tag{5.3}$$

where $A_{ij} \in \mathbb{C}^{n\times n}$ and $P_j(\lambda,\mu)$ is a homogeneous matrix polynomial of degree j given by $P_j(\lambda,\mu) := \sum_{l=0}^{j} \lambda^{j-l}\mu^l A_{jl}$ for $j = 0,\cdots,k$. In particular when $k=2$, we consider a two-parameter quadratic matrix polynomial given by

$$Q(\lambda,\mu) = \lambda^2 A_{20} + \lambda\mu A_{11} + \mu^2 A_{02} + \lambda A_{10} + \mu A_{01} + A_{00}. \tag{5.4}$$

First we recall the definition of linearization.

Definition 5.2.1. *[44] A $ln \times ln$ two-parameter matrix pencil $L(\lambda,\mu) = \lambda X + \mu Y + Z$ is a linearization of an $n \times n$ two-parameter matrix polynomial $P(\lambda,\mu)$ if there exist* *unimodular matrix polynomials* $E(\lambda,\mu)$ *and* $F(\lambda,\mu)$ *(i.e.,determinants of* $E(\lambda,\mu)$ *and* $F(\lambda,\mu)$ *are non-zero constant independent of* λ *and* μ*) such that*

$$F(\lambda,\mu)L(\lambda,\mu)E(\lambda,\mu) = \begin{bmatrix} P(\lambda,\mu) & 0 \\ 0 & I_{(l-1)n} \end{bmatrix}.$$

Next we recall the definition of block-transpose of a block matrix.

Definition 5.2.2. *[42] Let* $A = \sum_{i,j}(E_{ij} \otimes B_{ij})$ *be a* $k \times l$ *block matrix with* $m \times n$ *blocks* B_{ij}*. Then the block-transpose of* A *is* $A^{\mathcal{B}} := \sum_{i,j}(E_{ij}^T \otimes B_{ij})$*, a* $l \times k$ *block matrix also with* $m \times n$ *blocks.*

For the rest of the chapter, we define $p := \frac{k(k+1)}{2}$ and $q := \frac{(k+1)(k+2)}{2}$.

Let $P(\lambda,\mu)$ be a two-parameter matrix polynomial of degree k as given in (5.3). Define the block matrices

$$\mathcal{S} := \begin{bmatrix} A_{k0} & A_{k-1,1} & \cdots & A_{1,k-1} \end{bmatrix} \tag{5.5}$$

$$\mathcal{T} := \begin{bmatrix} 0 & \cdots & 0 & A_{0k} \end{bmatrix}, \tag{5.6}$$

$$\mathcal{A}_j := \begin{bmatrix} A_{j0} & A_{j-1,1} & \cdots & A_{1,j-1} & A_{0j} \end{bmatrix}, \quad j = 0, \cdots, k \tag{5.7}$$

$$\mathcal{I}_j := \begin{bmatrix} I_j \otimes I_n \\ 0_{n \times jn} \end{bmatrix} \in \mathbb{C}^{(j+1)n \times jn}, \quad j = 1, \cdots, k-1 \tag{5.8}$$

$$D_j := \left[\begin{array}{c|c} 0_{j\times(j-1)} & 0_{j\times 1} \\ \hline 0_{1\times(j-1)} & 1 \end{array}\right] \in \mathbb{C}^{(j+1)\times j}, \text{ and } \mathcal{J}_j := D_j \otimes I_n \in \mathbb{C}^{(j+1)n \times jn} \tag{5.9}$$

for $j = 1, \cdots, k-1$. Also define

$$\Lambda_j(\lambda,\mu) := \begin{bmatrix} \lambda^j \\ \lambda^{j-1}\mu \\ \vdots \\ \lambda\mu^{j-1} \\ \mu^j \end{bmatrix} \in \mathbb{C}^{j+1}, \; j = 0, \cdots, k-1 \text{ and } \Lambda := \begin{bmatrix} \Lambda_{k-1}(\lambda,\mu) \\ \vdots \\ \Lambda_0(\lambda,\mu) \end{bmatrix} \in \mathbb{C}^p. \tag{5.10}$$

In particular, when $k = 2$, $\Lambda = (\lambda,\mu,1)^T \in \mathbb{C}^3$. Throughout the chapter, we use the notations given in (5.5)-(5.10).

It is easy to see that $(\lambda\mathcal{S}+\mu\mathcal{T})\cdot(\Lambda_{k-1}(\lambda,\mu)\otimes I_n) = P_k(\lambda,\mu)$ and $\mathcal{A}_j\cdot(\Lambda_j(\lambda,\mu)\otimes I_n) = P_j(\lambda,\mu)$ for $j = 0,\cdots,k-1$. Thus we can rewrite $P(\lambda,\mu)x = 0$ for some $0 \neq x \in \mathbb{C}^n$ as

$$(\lambda\mathcal{S}+\mu\mathcal{T})\cdot(\Lambda_{k-1}(\lambda,\mu)\otimes x) + \sum_{j=0}^{k-1}\mathcal{A}_j\cdot(\Lambda_j(\lambda,\mu)\otimes x) = 0$$

and consequently we have

$$\left(\lambda\underbrace{\left[\begin{array}{c|ccc} \mathcal{S} & & & \\ \hline & \mathcal{I}_{k-1} & & \\ & & \ddots & \\ & & & \mathcal{I}_1 \end{array}\right]}_{=\mathbb{S}} + \mu\underbrace{\left[\begin{array}{c|ccc} \mathcal{T} & & & \\ \hline & \mathcal{J}_{k-1} & & \\ & & \ddots & \\ & & & \mathcal{J}_1 \end{array}\right]}_{=\mathbb{T}} + \underbrace{\left[\begin{array}{c|ccc} \mathcal{A}_{k-1} & \cdots & \mathcal{A}_1 & \mathcal{A}_0 \\ \hline -I_{kn} & & & \\ & \ddots & & \\ & & -I_{2n} & \end{array}\right]}_{=\mathbb{A}}\right)$$

$$\cdot\left(\begin{bmatrix} \Lambda_{k-1}(\lambda,\mu) \\ \vdots \\ \Lambda_0(\lambda,\mu) \end{bmatrix} \otimes x\right) = 0. \tag{5.11}$$

Define a two-parameter matrix pencil

$$L_1(\lambda,\mu) := \lambda\mathbb{S} + \mu\mathbb{T} + \mathbb{A}, \tag{5.12}$$

where $\mathbb{S},\mathbb{T},\mathbb{A}$ are as defined in (5.11). Next we show that $L_1(\lambda,\mu)$ is a linearization of $P(\lambda,\mu)$.

Theorem 5.2.1. *Let $P(\lambda,\mu)$ be as given in (5.3) and $L_1(\lambda,\mu)$ be as defined in (5.12). Then $L_1(\lambda,\mu)$ is a linearization of $P(\lambda,\mu)$.*

Proof. Consider a $p \times p$ block matrix

$$E(\lambda,\mu) := \left[\begin{array}{c|ccccc} \Lambda_{k-1}(\lambda,\mu)\otimes I_n & -I_k\otimes I_n & -(\lambda\mathcal{I}_{k-1}+\mu\mathcal{J}_{k-1}) & & & \\ \vdots & & \ddots & \ddots & & \\ \Lambda_2(\lambda,\mu)\otimes I_n & & & & -I_3\otimes I_n & -(\lambda\mathcal{I}_2+\mu\mathcal{J}_2) \\ \Lambda_1(\lambda,\mu)\otimes I_n & & & & & -I_2\otimes I_n \\ \hline \Lambda_0(\lambda,\mu)\otimes I_n & & & & & \end{array}\right].$$

Since $(\lambda\mathcal{S}+\mu\mathcal{T})\cdot(\Lambda_{k-1}(\lambda,\mu)\otimes I_n)+\sum_{j=0}^{k-1}\mathcal{A}_j\cdot(\Lambda_j(\lambda,\mu)\otimes I_n) = \sum_{j=0}^{k} P_j(\lambda,\mu) = P(\lambda,\mu)$ and

$$(\lambda\mathcal{I}_j+\mu\mathcal{J}_j)\cdot(\Lambda_{j-1}\otimes I_n) = \begin{bmatrix} \lambda I_n & & \\ & \ddots & \\ & & \lambda I_n \\ & & \mu I_n \end{bmatrix} \begin{bmatrix} \lambda^{j-1} I_n \\ \lambda^{j-2}\mu I_n \\ \vdots \\ \lambda\mu^{j-2} I_n \\ \mu^{j-1} I_n \end{bmatrix} = \begin{bmatrix} \lambda^{j} I_n \\ \lambda^{j-1}\mu I_n \\ \vdots \\ \lambda\mu^{j-1} I_n \\ \mu^{j} I_n \end{bmatrix} = \Lambda_j(\lambda,\mu)\otimes I_n$$

for $j = 1,\cdots,k-1$, we have

$$L_1(\lambda,\mu)E(\lambda,\mu) = \left[\begin{array}{c|ccc} P(\lambda,\mu) & \mathcal{M}_k(\lambda,\mu) & \cdots & \mathcal{M}_2(\lambda,\mu) \\ \hline & I_k\otimes I_n & & \\ & & \ddots & \\ & & & I_2\otimes I_n \end{array}\right] \in \mathbb{C}^{pn\times pn},$$

where $\mathcal{M}_k(\lambda,\mu) = -(\lambda\mathcal{S}+\mu\mathcal{T}+\mathcal{A}_{k-1}), \mathcal{M}_{k-1}(\lambda,\mu) = -(\lambda\mathcal{S}+\mu\mathcal{T}+\mathcal{A}_{k-1})\cdot(\lambda\mathcal{I}_{k-1}+\mu\mathcal{J}_{k-1}) - \mathcal{A}_{k-2}$ and $\mathcal{M}_j(\lambda,\mu) = -\mathcal{A}_j(\lambda\mathcal{I}_j+\mu\mathcal{J}_j) - \mathcal{A}_{j-1}$ for $j = 2,\cdots,k-2$.

Now to complete the proof, construct $p\times p$ block matrix

$$F(\lambda,\mu) := \left[\begin{array}{c|ccc} I_n & -\mathcal{M}_k(\lambda,\mu) & \cdots & -\mathcal{M}_2(\lambda,\mu) \\ \hline & I_k\otimes I_n & & \\ & & \ddots & \\ & & & I_2\otimes I_n \end{array}\right]$$

and then we have

$$F(\lambda,\mu)L_1(\lambda,\mu)E(\lambda,\mu) = \left[\begin{array}{c|ccc} P(\lambda,\mu) & & & \\ \hline & I_k\otimes I_n & & \\ & & \ddots & \\ & & & I_2\otimes I_n \end{array}\right] = \left[\begin{array}{c|c} P(\lambda,\mu) & \\ \hline & I_{p-1}\otimes I_n \end{array}\right].$$

It follows that

$$\begin{aligned}\det(E(\lambda,\mu)) &= \det(\Lambda_0(\lambda,\mu)\otimes I_n)\\ &\quad\cdot\det\left(\begin{bmatrix} -I_k\otimes I_n & -(\lambda\mathcal{I}_{k-1}+\mu\mathcal{J}_{k-1}) & & \\ & \ddots & \ddots & \\ & & -I_3\otimes I_n & -(\lambda\mathcal{I}_2+\mu\mathcal{J}_2)\\ & & & -I_2\otimes I_n\end{bmatrix}\right)\\ &= \det(\mathrm{I_n})\cdot\prod_{j=2}^{k}\det(-I_j\otimes I_n)\end{aligned}$$

and

$$\det(F(\lambda,\mu)) = \det\left(\begin{bmatrix} I_n & -\mathcal{M}_k(\lambda,\mu) & \cdots & -\mathcal{M}_2(\lambda,\mu)\\ & I_k\otimes I_n & & \\ & & \ddots & \\ & & & I_2\otimes I_n\end{bmatrix}\right) = \prod_{j=1}^{k}\det(I_j\otimes I_n) = 1.$$

Therefore $E(\lambda,\mu)$ and $F(\lambda,\mu)$ are unimodular and hence $L_1(\lambda,\mu)$ is a linearization of $P(\lambda,\mu)$. □

We refer to $L_1(\lambda,\mu)$ as the *first standard form* of $P(\lambda,\mu)$. In particular, when $k=2$, the *first standard form* of the two-parameter quadratic matrix polynomial $Q(\lambda,\mu)$ given in (5.4) is given by

$$L_1(\lambda,\mu) = \lambda\begin{bmatrix} A_{20} & A_{11} & 0_n\\ 0_n & 0_n & I_n\\ 0_n & 0_n & 0_n\end{bmatrix} + \mu\begin{bmatrix} 0_n & A_{02} & 0_n\\ 0_n & 0_n & 0_n\\ 0_n & 0_n & I_n\end{bmatrix} + \begin{bmatrix} A_{10} & A_{01} & A_{00}\\ -I_n & 0_n & 0_n\\ 0_n & -I_n & 0_n\end{bmatrix}. \qquad (5.13)$$

Note that the first standard form given in (5.13) is permutationally equivalent to the linearizations (5.1) and (5.2).

Next consider the two-parameter matrix pencil

$$\begin{aligned} L_2(\lambda,\mu) &:= \lambda\mathbb{S}^{\mathcal{B}}+\mu\mathbb{T}^{\mathcal{B}}+\mathbb{A}^{\mathcal{B}}\\ &= \lambda\left[\begin{array}{c|ccc} \mathcal{S}^{\mathcal{B}} & & & \\ \hline & \mathcal{I}^{\mathcal{B}}_{k-1} & & \\ & & \ddots & \\ & & & \mathcal{I}^{\mathcal{B}}_1\end{array}\right] + \mu\left[\begin{array}{c|ccc} \mathcal{T}^{\mathcal{B}} & & & \\ \hline & \mathcal{J}^{\mathcal{B}}_{k-1} & & \\ & & \ddots & \\ & & & \mathcal{J}^{\mathcal{B}}_1\end{array}\right] \\ &\quad + \left[\begin{array}{c|ccc} \mathcal{A}^{\mathcal{B}}_{k-1} & -I_k\otimes I_n & & \\ \vdots & & \ddots & \\ \mathcal{A}_1^{\mathcal{B}} & & & -I_2\otimes I_n \\ \hline \mathcal{A}_0^{\mathcal{B}} & & & \end{array}\right], \end{aligned} \tag{5.14}$$

where $\mathcal{S}^{\mathcal{B}}, \mathcal{T}^{\mathcal{B}}, \mathcal{A}_j^{\mathcal{B}}, \mathcal{I}_j^{\mathcal{B}}, \mathcal{J}_j^{\mathcal{B}}$ denotes the block transpose of the block matrices $\mathcal{S}, \mathcal{T}, \mathcal{A}_j, \mathcal{I}_j, \mathcal{J}_j$ respectively and the zero blocks are block transposed in proper order. In the following result, we show that $L_2(\lambda,\mu)$ is also a linearization of $P(\lambda,\mu)$.

Theorem 5.2.2. *Let $P(\lambda,\mu)$ be as given in (5.3) and $L_2(\lambda,\mu)$ be as defined in (5.14). Then $L_2(\lambda,\mu)$ is a linearization of $P(\lambda,\mu)$*

Proof. Consider the $p\times p$ block matrix

$$E(\lambda,\mu) := \left[\begin{array}{cccc|c} \Lambda^T_{k-1}(\lambda,\mu)\otimes I_n & \cdots & \Lambda^T_2(\lambda,\mu)\otimes I_n & \Lambda^T_1(\lambda,\mu)\otimes I_n & \Lambda^T_0(\lambda,\mu)\otimes I_n \\ \hline -I_k\otimes I_n & & & & \\ -(\lambda\mathcal{I}^{\mathcal{B}}_{k-1}+\mu\mathcal{J}^{\mathcal{B}}_{k-1}) & \ddots & & & \\ & \ddots & -I_3\otimes I_n & & \\ & & -(\lambda\mathcal{I}^{\mathcal{B}}_2+\mu\mathcal{J}^{\mathcal{B}}_2) & -I_2\otimes I_n & \end{array}\right],$$

where $\Lambda_j^T(\lambda,\mu)$ denotes the transpose of the $\Lambda_j(\lambda,\mu)$, for all $j=0,\cdots,k-1$. Note that

$$(\Lambda^T_{k-1}(\lambda,\mu)\otimes I_n)\cdot(\lambda\mathcal{S}^{\mathcal{B}}+\mu\mathcal{T}^{\mathcal{B}}) = \begin{bmatrix} \lambda^{k-1}I_n & \lambda^{k-2}\mu I_n & \cdots & \mu^{k-1}I_n\end{bmatrix}\cdot\begin{bmatrix} \lambda A_{k0} \\ \lambda A_{k-1,1} \\ \vdots \\ \lambda A_{1,k-1} \\ \mu A_{0k}\end{bmatrix} = P_k(\lambda,\mu) \tag{5.15}$$

and

$$(\Lambda_j^T(\lambda,\mu)\otimes I_n)\cdot \mathcal{A}_j^{\mathcal{B}} = \begin{bmatrix} \lambda^{j-1}I_n & \lambda^{j-2}\mu I_n & \cdots & \mu^{j-1}I_n \end{bmatrix}\cdot \begin{bmatrix} A_{j0} \\ A_{j-1,1} \\ \vdots \\ A_{1,j-1} \\ A_{0j} \end{bmatrix} = P_j(\lambda,\mu), \quad (5.16)$$

for $j = 0, \cdots, k-1$. Also we have

$$\begin{aligned} (\Lambda_{j-1}^T(\lambda,\mu)\otimes I_n)\cdot(\lambda\mathcal{I}_j^{\mathcal{B}} + \mu\mathcal{J}_j^{\mathcal{B}}) &= \begin{bmatrix} \lambda^{j-1}I_n & \lambda^{j-2}\mu I_n & \cdots & \mu^{j-1}I_n \end{bmatrix} \left[\begin{array}{ccc|c} \lambda I_n & & & \\ & \ddots & & \\ & & \lambda I_n & \mu I_n \end{array}\right] \\ &= \begin{bmatrix} \lambda^{j}I_n & \lambda^{j-1}\mu I_n & \cdots & \mu^{j}I_n \end{bmatrix} = \Lambda_j^T(\lambda,\mu)\otimes I_n \qquad (5.17) \end{aligned}$$

for $j = 1, \cdots, k-1$. Then by (5.15)-(5.17) we have

$$E(\lambda,\mu)L_2(\lambda,\mu) = \left[\begin{array}{c|ccc} P(\lambda,\mu) & & & \\ \hline \widehat{\mathcal{M}}_k(\lambda,\mu) & I_k\otimes I_n & & \\ \vdots & & \ddots & \\ \widehat{\mathcal{M}}_2(\lambda,\mu) & & & I_2\otimes I_n \end{array}\right],$$

where $\widehat{\mathcal{M}}_k(\lambda,\mu) = -(\lambda\mathcal{S}^{\mathcal{B}} + \mu\mathcal{T}^{\mathcal{B}} + \mathcal{A}_{k-1}^{\mathcal{B}})$, $\widehat{\mathcal{M}}_{k-1}(\lambda,\mu) = -(\lambda\mathcal{I}_{k-1}^{\mathcal{B}} + \mu\mathcal{J}_{k-1}^{\mathcal{B}})\cdot(\lambda\mathcal{S}^{\mathcal{B}} + \mu\mathcal{T}^{\mathcal{B}} + \mathcal{A}_{k-1}^{\mathcal{B}})$ and $\widehat{\mathcal{M}}_j(\lambda,\mu) = -(\lambda\mathcal{I}_j^{\mathcal{B}} + \mu\mathcal{J}_j^{\mathcal{B}})\cdot\mathcal{A}_j^{\mathcal{B}} - \mathcal{A}_{j-1}^{\mathcal{B}}$ for $j = 2, \cdots, k-2$.

To conclude the proof, consider the $p \times p$ block matrix

$$F(\lambda,\mu) := \left[\begin{array}{c|ccc} I_n & & & \\ \hline -\widehat{\mathcal{M}}_k(\lambda,\mu) & I_k\otimes I_n & & \\ \vdots & & \ddots & \\ -\widehat{\mathcal{M}}_2(\lambda,\mu) & & & I_2\otimes I_n \end{array}\right].$$

Then we have

$$E(\lambda,\mu)L_2(\lambda,\mu)F(\lambda,\mu) = \left[\begin{array}{c|ccc} P(\lambda,\mu) & & & \\ \hline & I_k\otimes I_n & & \\ & & \ddots & \\ & & & I_2\otimes I_n \end{array}\right] = \left[\begin{array}{c|c} P(\lambda,\mu) & \\ \hline & I_{p-1}\otimes I_n \end{array}\right].$$

It is easy to see that $E(\lambda,\mu)$ and $F(\lambda,\mu)$ are unimodular. Hence $L_2(\lambda,\mu)$ is a linearization of $P(\lambda,\mu)$. $\square$

We refer to $L_2(\lambda,\mu)$ as the *second standard form* of $P(\lambda,\mu)$. The second standard form of two-parameter quadratic matrix polynomial $Q(\lambda,\mu)$ given in (5.4) is given by

$$L_2(\lambda,\mu) = \lambda \begin{bmatrix} A_{20} & 0 & 0 \\ A_{11} & 0 & 0 \\ 0 & I_n & 0 \end{bmatrix} + \mu \begin{bmatrix} 0 & 0 & 0 \\ A_{02} & 0 & 0 \\ 0 & 0 & I_n \end{bmatrix} + \begin{bmatrix} A_{10} & -I_n & 0 \\ A_{01} & 0 & -I_n \\ A_{00} & 0 & 0 \end{bmatrix}.$$

5.2.1 Right and left ansatz spaces

Let $P(\lambda,\mu)$ be a two-parameter matrix polynomial of degree k given in (5.3). Recall the first standard form $L_1(\lambda,\mu)$ given in (5.12). It is straightforward to see that $L_1(\lambda,\mu)\cdot(\Lambda\otimes I_n) = e_1\otimes P(\lambda,\mu)$, where $e_1\in\mathbb{C}^p$ is the first column of the identity matrix I_p. This motivates us to concentrate on two-parameter matrix pencils $L(\lambda,\mu)$ such that $L(\lambda,\mu)\cdot(\Lambda\otimes I_n) = v\otimes P(\lambda,\mu)$ for some $v\in\mathbb{C}^p$. Consider $\mathcal{V}_P := \{v\otimes P(\lambda,\mu) : v\in\mathbb{C}^p\}$ and define the space

$$\mathbb{L}_1(P) := \{L(\lambda,\mu) := \lambda X + \mu Y + Z : L(\lambda,\mu)\cdot(\Lambda\otimes I_n)\in\mathcal{V}_P\}.$$

Note that $\mathbb{L}_1(P)$ is nonempty as the first standard form $L_1(\lambda,\mu)\in\mathbb{L}_1(P)$. Also it is easy to check that $\mathbb{L}_1(P)$ is a vector space. We refer to the space $\mathbb{L}_1(P)$ as the *right ansatz space*. If $L(\lambda,\mu)\in\mathbb{L}_1(P)$ is such that $L(\lambda,\mu)\cdot(\Lambda\otimes I_n) = v\otimes P(\lambda,\mu)$ for some $v\in\mathbb{C}^p$ then we refer v as the *right ansatz vector*.

To characterize a two-parameter matrix pencil $L(\lambda,\mu)\in\mathbb{L}_1(P)$, we define a column shifted sum of block matrices. For the rest of this chapter, we follow the convention that $A(i,j)$ is the $(i,j)^{\text{th}}$ block of A whenever A is a block matrix with each block being an $n\times n$ matrix.

Definition 5.2.3. (Column shifted sum) *Let X, Y, Z be $p\times p$ block matrices with each* *blocks being an $n \times n$ matrix. Then we define the column shifted sum by*

$$\begin{aligned} & X \boxplus Y \boxplus Z \\ = & \begin{bmatrix} X(:,1:k) & 0 & X(:,k+1:2k-1) & 0 & \cdots & 0 & X(:,p-2:p-1) & 0 & X(:,p) & 0 & 0 \end{bmatrix} \\ + & \begin{bmatrix} 0 & Y(:,1:k) & 0 & Y(:,k+1:2k-1) & 0 & \cdots & 0 & Y(:,p-2:p-1) & 0 & Y(:,p) & 0 \end{bmatrix} \\ + & \begin{bmatrix} \underbrace{0 \cdots 0}_{(k+1) \text{ block columns}} & Z(:,1) & \cdots & Z(:,p) \end{bmatrix}, \end{aligned}$$

where 0 *denotes a zero block column with* $n \times n$ *zero matrices and '+' denotes usual matrix addition.*

For example, when $k = 2$, consider the 3×3 block matrices $X = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix}, Y = \begin{bmatrix} Y_1 & Y_2 & Y_3 \end{bmatrix}, Z = \begin{bmatrix} Z_1 & Z_2 & Z_3 \end{bmatrix}$, where X_j, Y_j, Z_j are block columns with each blocks being an $n \times n$ matrix. Then we have

$$\begin{aligned} X \boxplus Y \boxplus Z &= \begin{bmatrix} X(:,1:2) & 0 & X(:3) & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & Y(:,1:2) & 0 & Y(:,3) & 0 \end{bmatrix} \\ &+ \begin{bmatrix} 0 & 0 & 0 & Z(:,1) & Z(:,2) & Z(:,3) \end{bmatrix} \\ &= \begin{bmatrix} X_1 & X_2 + Y_1 & Y_2 & X_3 + Z_1 & Y_3 + Z_2 & Z_3 \end{bmatrix}, \end{aligned}$$

which is same as the "box addition" defined in [1].

Also, when $k = 3$, consider the 6×6 block matrices $X = \begin{bmatrix} X_1 & \cdots & X_6 \end{bmatrix}, Y = \begin{bmatrix} Y_1 & \cdots & Y_6 \end{bmatrix}, Z = \begin{bmatrix} Z_1 & \cdots & Z_6 \end{bmatrix}$, where X_j, Y_j, Z_j are block columns with each blocks being an $n \times n$ matrix. Then we have

$$\begin{aligned} X \boxplus Y \boxplus Z &= \begin{bmatrix} X(:,1:3) & 0 & X(:4:5) & 0 & X(:,6) & 0 & 0 \end{bmatrix} \\ &+ \begin{bmatrix} 0 & Y(:,1:3) & 0 & Y(:,4:5) & 0 & Y(:,6) & 0 \end{bmatrix} \\ &+ \begin{bmatrix} 0 & 0 & 0 & 0 & Z(:,1) & \cdots & Z(:,6) \end{bmatrix}. \end{aligned}$$

The column shifted sum of three $p \times p$ block matrices is a $p \times q$ block matrix which we characterize in the next result.

Lemma 5.2.1. *Let X, Y, Z be $p \times p$ block matrices with each block being an $n \times n$ matrix. Let S be a $p \times q$ block matrix with each blocks being an $n \times n$ matrix be such that $X \boxplus Y \boxplus Z = S$. Then*

$$S(:,1) = X(:,1), S(:,k+1) = Y(:,k), S(:,q) = Z(:,p),$$

$$S(:,j) = X(:,j) + Y(:,j-1) \quad \text{for } j = 2,\cdots,k,$$

$$S\left(:,\frac{j}{2}(2k-j+3)+1\right) = X\left(:,\frac{j}{2}(2k-j+1)+1\right) + Z\left(:,\frac{j-1}{2}(2k-j+2)+1\right)$$

for $j = 1,\cdots,k-1,$

$$S\left(:,\frac{j+1}{2}(2k-j+2)\right) = Y\left(:,\frac{j+1}{2}(2k-j)\right) + Z\left(:,\frac{j}{2}(2k-j+1)\right)$$

for $j = 1,\cdots,k-1,$

$$\begin{aligned} S\left(:,\frac{j}{2}(2k-j+3)+(r+1)\right) &= X\left(:,\frac{j}{2}(2k-j+1)+(r+1)\right) + Y\left(:,\frac{j}{2}(2k-j+1)+r\right) \\ &\quad + Z\left(:,\frac{j-1}{2}(2k-j+2)+(r+1)\right) \end{aligned}$$

for $j = 1,\cdots,k-2, r = 1,\cdots,k-j-1.$

Proof. The proof is computational and follows from Definition 5.2.3. ☐

Next we derive the column shifted sum of the coefficient matrices of the first standard form $L_1(\lambda,\mu)$.

Proposition 5.2.1. *Let* $P(\lambda,\mu)$ *be as given in (5.3). Consider the first standard form* $L_1(\lambda,\mu) := \lambda\mathbb{S} + \mu\mathbb{T} + \mathbb{A}$ *given in (5.12). Then* $\mathbb{S} \boxplus \mathbb{T} \boxplus \mathbb{A} = e_1 \otimes \begin{bmatrix} \mathcal{A}_k & \cdots & \mathcal{A}_0 \end{bmatrix}$.

Proof. Let $S = \mathbb{S} \boxplus \mathbb{T} \boxplus \mathbb{A}$. Now by Lemma 5.2.1, we have

$$S(:,1) = \mathbb{S}(:,1) = \begin{bmatrix} A_{k0} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, S(:,k+1) = \mathbb{T}(:,k) = \begin{bmatrix} A_{0k} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, S(:,q) = \mathbb{A}(:,p) = \begin{bmatrix} \mathcal{A}_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{5.18}$$

$$S(:,j) = \mathbb{S}(:,j) + \mathbb{T}(:,j-1) = \begin{bmatrix} A_{k-j+1,j-1} \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} A_{k-j+1,j-1} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{5.19}$$

for $j = 2, \cdots, k$,

$$\begin{aligned} S\left(:, \frac{j}{2}(2k-j+3)+1\right) &= \mathbb{S}\left(:, \frac{j}{2}(2k-j+1)+1\right) + \mathbb{A}\left(:, \frac{j-1}{2}(2k-j+2)+1\right) \\ &= \begin{bmatrix} 0 \\ I_n \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} A_{k-j,0} \\ -I_n \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} A_{k-j,0} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \end{aligned} \tag{5.20}$$

for $j = 1, \cdots, k-1$,

$$\begin{aligned} S\left(:, \frac{j+1}{2}(2k-j+2)\right) &= \mathbb{T}\left(:, \frac{j+1}{2}(2k-j)\right) + \mathbb{A}\left(:, \frac{j}{2}(2k-j+1)\right) \\ &= \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -I_n \end{bmatrix} + \begin{bmatrix} A_{0,k-j} \\ 0 \\ \vdots \\ 0 \\ I_n \end{bmatrix} = \begin{bmatrix} A_{0,k-j} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \end{aligned} \tag{5.21}$$

for $j = 1, \cdots, k-1$,

$$\begin{aligned} S\left(:, \frac{j}{2}(2k-j+3)+(r+1)\right) &= \mathbb{S}\left(:, \frac{j}{2}(2k-j+1)+(r+1)\right) + \mathbb{T}\left(:, \frac{j}{2}(2k-j+1)+r\right) \\ &\quad + \mathbb{A}\left(:, \frac{j-1}{2}(2k-j+2)+(r+1)\right) \\ &= \begin{bmatrix} 0 \\ e_{\frac{j-1}{2}(2k-j+2)+(r+1)} \otimes I_n \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \\ &\quad + \begin{bmatrix} A_{k-j-r,r} \\ -e_{\frac{j-1}{2}(2k-j+2)+(r+1)} \otimes I_n \end{bmatrix} = \begin{bmatrix} A_{k-j-r,r} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \end{aligned} \tag{5.22}$$

for $j = 1, \cdots, k-2, r = 1, \cdots, k-j-1$.

By (5.18)-(5.22), we have

$$S(:,1:k+1) = \begin{bmatrix} A_{k0} & A_{k-1,1} & \cdots & A_{1,k-1} & A_{0k} \\ 0 & 0 & \cdots & 0 & 0 \\ & & \vdots & & \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix} = \begin{bmatrix} \mathcal{A}_k \\ 0 \\ \vdots \\ 0 \end{bmatrix}, S(:,q) = \begin{bmatrix} \mathcal{A}_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$$\begin{aligned} S\left(:, \frac{j}{2}(2k-j+3)+1 : \frac{j+1}{2}(2k-j+2)\right) &= \begin{bmatrix} A_{k-j,0} & A_{k-j-1,1} & \cdots & A_{1,k-j-1} & A_{0,k-j} \\ 0 & 0 & \cdots & 0 & 0 \\ & & \vdots & & \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} \mathcal{A}_{k-j} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{for } j = 1, \cdots, k-1. \end{aligned}$$

Therefore, we have

$$\mathbb{S} \boxplus \mathbb{T} \boxplus \mathbb{A} = \begin{bmatrix} \mathcal{A}_k & \cdots & \mathcal{A}_0 \\ 0 & \cdots & 0 \\ & \vdots & \\ 0 & \cdots & 0 \end{bmatrix} = e_1 \otimes \begin{bmatrix} \mathcal{A}_k & \cdots & \mathcal{A}_0 \end{bmatrix}.$$

$\square$

In particular, for a two-parameter quadratic matrix polynomial, we have the following result which is proved in [1].

Corollary 5.2.3. *[1] Let $Q(\lambda,\mu)$ be as given in (5.4). Consider the first standard form $L_1(\lambda,\mu)$ given in (5.13). Then the column shifted sum of the coefficient matrices of $L_1(\lambda,\mu)$ is $e_1 \otimes \begin{bmatrix} A_{20} & A_{11} & A_{02} & A_{10} & A_{01} & A_{00} \end{bmatrix}$.*

Recall that the first standard form $L_1(\lambda,\mu) = \lambda\mathbb{S}+\mu\mathbb{T}+\mathbb{A} \in \mathbb{L}_1(P)$ with right ansatz vector e_1 and by Proposition 5.2.1, $\mathbb{S} \boxplus \mathbb{T} \boxplus \mathbb{A} = e_1 \otimes \begin{bmatrix} \mathcal{A}_k & \cdots & \mathcal{A}_0 \end{bmatrix}$. The next result characterizes $L(\lambda,\mu) \in \mathbb{L}_1(P)$ via column shifted sum.

Lemma 5.2.2. *Let $P(\lambda,\mu)$ be as given in (5.3). Let $L(\lambda,\mu) = \lambda X + \mu Y + Z$ be a two-parameter matrix pencil, where X, Y, Z are $p \times p$ block matrices with $n \times n$ blocks. Then $L(\lambda,\mu) \in \mathbb{L}_1(P)$ with right ansatz vector $v \in \mathbb{C}^p$ if and only if $X \boxplus Y \boxplus Z = v \otimes \begin{bmatrix} \mathcal{A}_k & \cdots & \mathcal{A}_0 \end{bmatrix}$.*

Proof. First note that $(L(\lambda,\mu)\cdot(\Lambda \otimes I_n))$

$$\begin{aligned}
= {} & \lambda^k X(:,1) + \mu^k Y(:,k) + Z(:,p) + \sum_{j=2}^{k} \lambda^{k-j+1}\mu^{j-1}\left(X(:,j) + Y(:,j-1)\right) \\
& + \sum_{j=1}^{k-1} \lambda^{k-j}\left(X\left(:,\frac{j}{2}(2k-j+1)+1\right) + Z\left(:,\frac{j-1}{2}(2k-j+2)+1\right)\right) \\
& + \sum_{j=1}^{k-1} \mu^{k-j}\left(Y\left(:,\frac{j+1}{2}(2k-j)\right) + Z\left(:,\frac{j}{2}(2k-j+1)\right)\right) \\
& + \sum_{j=1}^{k-2}\sum_{r=1}^{k-j-1} \lambda^{k-j-r}\mu^r\left(X\left(:,\frac{j}{2}(2k-j+1)+(r+1)\right) + Y\left(:,\frac{j}{2}(2k-j+1)+r\right)\right. \\
& \left. + Z\left(:,\frac{j-1}{2}(2k-j+2)+(r+1)\right)\right).
\end{aligned} \tag{5.23}$$

Now $L(\lambda,\mu) \in \mathbb{L}_1(P)$ with right ansatz vector $v \in \mathbb{C}^p$ if and only if

$$L(\lambda,\mu)\cdot(\Lambda \otimes I_n) = v \otimes P(\lambda,\mu) = \sum_{i=0}^{k}\sum_{j=0}^{k-i} \lambda^i\mu^j\left(v \otimes A_{ij}\right). \tag{5.24}$$

Then comparing the coefficient for the respective powers in (5.23) and (5.24) we have,

$$X(:,1) = v \otimes A_{k0},\, Y(:,k) = v \otimes A_{0k},\, Z(:,p) = v \otimes A_{00}, \tag{5.25}$$

$$X(:,j) + Y(:,j-1) = v \otimes A_{k-j+1,j-1} \quad \text{for } j = 2,\cdots,k, \tag{5.26}$$

$$X\left(:,\frac{j}{2}(2k-j+1)+1\right) + Z\left(:,\frac{j-1}{2}(2k-j+2)+1\right) = v \otimes A_{k-j,0} \tag{5.27}$$

for $j = 1,\cdots,k-1$,

$$Y\left(:,\frac{j+1}{2}(2k-j)\right) + Z\left(:,\frac{j}{2}(2k-j+1)\right) = v \otimes A_{0,k-j} \tag{5.28}$$

for $j = 1,\cdots,k-1$ and

$$\begin{aligned}
& X\left(:,\frac{j}{2}(2k-j+1)+(r+1)\right) + Y\left(:,\frac{j}{2}(2k-j+1)+r\right) \\
& + Z\left(:,\frac{j-1}{2}(2k-j+2)+(r+1)\right) = v \otimes A_{k-j-r,r}
\end{aligned} \tag{5.29}$$

for $j = 1, \cdots, k-2, r = 1, \cdots, k-j-1$. Now by Lemma 5.2.1 and (5.25)-(5.29) it follows that

$$(X \boxplus Y \boxplus Z)(:, 1:k+1) = \begin{bmatrix} v \otimes A_{k0} & v \otimes A_{k-1,1} & \cdots & v \otimes A_{1,k-1} & v \otimes A_{0k} \end{bmatrix} = v \otimes \mathcal{A}_k,$$
$$(X \boxplus Y \boxplus Z)\left(:, \frac{j}{2}(2k-j+3)+1 : \frac{j+1}{2}(2k-j+2)\right)$$
$$= \begin{bmatrix} v \otimes A_{k-j,0} & v \otimes A_{k-j-1,1} & \cdots & v \otimes A_{1,k-j-1} & v \otimes A_{0,k-j} \end{bmatrix} = v \otimes \mathcal{A}_{k-j}$$

for $j = 1, \cdots, k-1$, and $(X \boxplus Y \boxplus Z)(:, q) = v \otimes A_{00} = v \otimes \mathcal{A}_0$.

Hence $X \boxplus Y \boxplus Z = v \otimes \begin{bmatrix} \mathcal{A}_k & \cdots & \mathcal{A}_0 \end{bmatrix}$. This completes the proof. ☐

For a quadratic two-parameter matrix polynomial we have the following result which is proved in [1].

Corollary 5.2.4. *[1] Let $Q(\lambda, \mu)$ be as given in (5.4). Let $L(\lambda, \mu) = \lambda X + \mu Y + Z$ be a two-parameter matrix pencil such that $X = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix}, Y = \begin{bmatrix} Y_1 & Y_2 & Y_3 \end{bmatrix}, Z = \begin{bmatrix} Z_1 & Z_2 & Z_3 \end{bmatrix}$, where X_j, Y_j, Z_j are block columns with $n \times n$ matrices for $j = 1, 2, 3$. Then $L(\lambda, \mu) \in \mathbb{L}_1(Q)$ with right ansatz vector $v \in \mathbb{C}^3$ if and only if*

$$X \boxplus Y \boxplus Z = v \otimes \begin{bmatrix} A_{20} & A_{11} & A_{02} & A_{10} & A_{01} & A_{00} \end{bmatrix}.$$

Now using Lemma 5.2.2 we characterize a linear two-parameter matrix polynomial $L(\lambda, \mu) \in \mathbb{L}_1(P)$ with right ansatz vector $v \in \mathbb{C}^p$.

Theorem 5.2.5. *Let $P(\lambda, \mu)$ be as given in (5.3) and $L(\lambda, \mu) \in \mathbb{L}_1(P)$ with right ansatz vector $v \in \mathbb{C}^p$. Then*

$$\begin{aligned} L(\lambda, \mu) &= \lambda\left(\begin{bmatrix} v \otimes \mathcal{S} & 0 & \cdots & 0 \end{bmatrix} + X\right) + \mu\left(\begin{bmatrix} v \otimes \mathcal{T} & 0 & \cdots & 0 \end{bmatrix} + Y\right) \\ &\quad + \begin{bmatrix} v \otimes \mathcal{A}_{k-1} & \cdots & v \otimes \mathcal{A}_0 \end{bmatrix} + Z, \end{aligned}$$

where $\mathcal{S}, \mathcal{T}, \mathcal{A}_j$ are as defined in (5.5)-(5.7) and X, Y, Z are $p \times p$ block matrices with $n \times n$ blocks such that $X \boxplus Y \boxplus Z = 0$.

Proof. Define a linear map $\mathcal{M} : \mathbb{L}_1(P) \to \mathcal{V}_P$ given by $L(\lambda, \mu) \mapsto L(\lambda, \mu) \cdot (\Lambda \otimes I_n)$. We first show that $\mathcal{M}$ is surjective. Let $v \otimes P(\lambda, \mu) \in \mathcal{V}_P$ for some $v \in \mathbb{C}^p$. Consider a two-parameter matrix pencil

$$\widehat{L}(\lambda, \mu) := \lambda \begin{bmatrix} v \otimes \mathcal{S} & 0 & \cdots & 0 \end{bmatrix} + \mu \begin{bmatrix} v \otimes \mathcal{T} & 0 & \cdots & 0 \end{bmatrix} + \begin{bmatrix} v \otimes \mathcal{A}_{k-1} & \cdots & v \otimes \mathcal{A}_0 \end{bmatrix}.$$

Then

$$\begin{bmatrix} v\otimes\mathcal{S} & 0 & \cdots & 0 \end{bmatrix} \boxplus \begin{bmatrix} v\otimes\mathcal{T} & 0 & \cdots & 0 \end{bmatrix} \boxplus \begin{bmatrix} v\otimes\mathcal{A}_{k-1} & \cdots & v\otimes\mathcal{A}_0 \end{bmatrix} = v\otimes \begin{bmatrix} \mathcal{A}_k & \cdots & \mathcal{A}_0 \end{bmatrix}.$$

Thus by Lemma 5.2.2, it follows that $\widehat{L}(\lambda,\mu)\cdot(\Lambda\otimes I_n) = v\otimes P(\lambda,\mu)$. So $\mathcal{M}$ is a surjective map. Therefore the set of all $\mathcal{M}$-preimages of $v\otimes P(\lambda,\mu)$ is $\widehat{L}(\lambda,\mu)+\ker\mathcal{M}$. Note that $\ker\mathcal{M} = \{L(\lambda,\mu) = \lambda X+\mu Y+Z\in\mathbb{L}_1(P) : L(\lambda,\mu)\cdot(\Lambda\otimes I_n) = 0\}$. By Lemma 5.2.2 it follows that, if $(\lambda X+\mu Y+Z)\cdot(\Lambda\otimes I_n) = 0$ then $X\boxplus Y\boxplus Z = 0$. So the set of all $\mathcal{M}$-preimages of $v\otimes P(\lambda,\mu)$ is

$$\begin{aligned} L(\lambda,\mu) &= \lambda\left(\begin{bmatrix} v\otimes\mathcal{S} & 0 & \cdots & 0 \end{bmatrix} + X\right) + \mu\left(\begin{bmatrix} v\otimes\mathcal{T} & 0 & \cdots & 0 \end{bmatrix} + Y\right) \\ &\quad + \begin{bmatrix} v\otimes\mathcal{A}_{k-1} & \cdots & v\otimes\mathcal{A}_0 \end{bmatrix} + Z, \end{aligned}$$

where $X\boxplus Y\boxplus Z = 0$. This completes the proof. □

For $L(\lambda,\mu)\in\mathbb{L}_1(Q)$, we have the following result.

Corollary 5.2.6. *Let $Q(\lambda,\mu)$ be as given in (5.4) and $L(\lambda,\mu)\in\mathbb{L}_1(Q)$ with right ansatz vector $v\in\mathbb{C}^3$. Then*

$$\begin{aligned} L(\lambda,\mu) &= \lambda\begin{bmatrix} v\otimes A_{20} & v\otimes A_{11}-Y_1 & -Z_1 \end{bmatrix} + \mu\begin{bmatrix} Y_1 & v\otimes A_{02} & -Z_2 \end{bmatrix} \\ &\quad + \begin{bmatrix} v\otimes A_{10}+Z_1 & v\otimes A_{01}+Z_2 & v\otimes A_{00} \end{bmatrix}, \end{aligned}$$

where Y_1, Z_1, Z_2 are arbitrary block columns and each block being an $n\times n$ matrix.

Next we derive the dimension of the right ansatz space.

Proposition 5.2.2. *Let $P(\lambda,\mu)$ be as given in (5.3). Then*

$$\dim(\mathbb{L}_1(P)) = \frac{k(k+1)}{2}(k^2-1)n^2 + \frac{k(k+1)}{2}.$$

Proof. Recall the linear map $\mathcal{M} : \mathbb{L}_1(P)\to\mathcal{V}_P, L(\lambda,\mu)\mapsto L(\lambda,\mu)\cdot(\Lambda\otimes I_n)$ defined in the proof of Theorem 5.2.5. Since $\mathcal{M}$ is surjective, so $\dim(\mathbb{L}_1(P)) = \dim\ker\mathcal{M} + \dim\mathcal{V}_P$. Now $\dim\mathcal{V}_P = p = \frac{k(k+1)}{2}$. To obtain $\dim\ker\mathcal{M}$, first note that $\ker\mathcal{M} = \{L(\lambda,\mu) = \lambda X+\mu Y+Z\in\mathbb{L}_1(P) : X\boxplus Y\boxplus Z = 0\}$. By Proposition 5.2.1, if $L(\lambda,\mu) = \lambda X+\mu Y+Z\in\ker\mathcal{M}$ then we have

$$X(:,1) = Y(:,k) = Z(:,p) = 0, \tag{5.30}$$

$$X(:,j) = -Y(:,j-1) \quad \text{for } j = 2,\cdots,k, \tag{5.31}$$

$$X\left(:, \frac{j}{2}(2k-j+1)+1\right) = -Z\left(:, \frac{j-1}{2}(2k-j+2)+1\right) \tag{5.32}$$

for $j = 1,\cdots,k-1$,

$$\begin{aligned} X\left(:, \frac{j}{2}(2k-j+1)+(r+1)\right) &= -Y\left(:, \frac{j}{2}(2k-j+1)+r\right) \\ &\quad -Z\left(:, \frac{j-1}{2}(2k-j+2)+(r+1)\right) \end{aligned} \tag{5.33}$$

for $j = 1,\cdots,k-2, r = 1,\cdots,k-j-1$,

$$Y\left(:, \frac{j+1}{2}(2k-j)\right) = -Z\left(:, \frac{j}{2}(2k-j+1)\right) \quad \text{for } j = 1,\cdots,k-1, \tag{5.34}$$

$Y\left(:, \frac{j-1}{2}(2k-j+2)+1\right), \cdots, Y\left(:, \frac{j-1}{2}(2k-j+2)+(k-j)\right)$ are arbitrary block columns for $j = 1,\cdots,k-1$, and $Z(:,j)$ is arbitrary block column for $j = 1,\cdots,p-1$. So it follows from (5.30)-(5.34) that dim ker $\mathcal{M} = \frac{k(k+1)}{2}(k^2-1)n^2$. Therefore $\dim(\mathbb{L}_1(P)) = \frac{k(k+1)}{2}(k^2-1)n^2 + \frac{k(k+1)}{2}$. □

In particular when $k = 2$, it follows from Proposition 5.2.2 that $\dim(\mathbb{L}_1(Q)) = 9n^2+3$, which is proved in [1].

Let $x \in \ker(P(\lambda,\mu))$ and $L(\lambda,\mu) \in \mathbb{L}_1(P)$ with a nonzero right ansatz vector $v \in \mathbb{C}^p$. Then $L(\lambda,\mu)(\Lambda \otimes x) = L(\lambda,\mu)(\Lambda \otimes I_n)x = v \otimes P(\lambda,\mu)x = 0$. So $\Lambda \otimes x \in \ker(L(\lambda,\mu))$.

Proposition 5.2.3. *Let $P(\lambda,\mu)$ be as given in (5.3) and $L(\lambda,\mu) \in \mathbb{L}_1(P)$ with nonzero right ansatz vector $v \in \mathbb{C}^p$. Then $0 \neq x \in \ker(P(\lambda,\mu))$ if and only if $\Lambda \otimes x \in \ker(L(\lambda,\mu))$.*

Proof. Define $g : \ker(P(\lambda,\mu)) \to \ker(L(\lambda,\mu))$ given by $x \mapsto \Lambda \otimes x$. It is easy to see that g defines an isomorphism. □

Next consider the second standard form $L_2(\lambda,\mu)$ given in (5.14) and note that $(\Lambda^T \otimes I_n) \cdot L_2(\lambda,\mu) = e_1^T \otimes P(\lambda,\mu)$. This motivates us to analyze two-parameter matrix pencils $L(\lambda,\mu)$ such that $(\Lambda^T \otimes I_n) \cdot L(\lambda,\mu) = w^T \otimes P(\lambda,\mu)$ for $w \in \mathbb{C}^p$. Let $\mathcal{W}_P := \{w^T \otimes P(\lambda,\mu) : w \in \mathbb{C}^p\}$ and define the space

$$\mathbb{L}_2(P) := \{L(\lambda,\mu) := \lambda X + \mu Y + Z : (\Lambda^T \otimes I_n) \cdot L(\lambda,\mu) \in \mathcal{W}_P\}.$$

We call $\mathbb{L}_2(P)$ as the *left ansatz space*. If $L(\lambda,\mu) \in \mathbb{L}_2(P)$ satisfies $(\Lambda^T \otimes I_n) \cdot L(\lambda,\mu) = w^T \otimes P(\lambda,\mu)$ for some $w \in \mathbb{C}^p$ then we call w as the *left ansatz vector*. Left ansatz space is nonempty as the second standard form $L_2(\lambda,\mu) \in \mathbb{L}_2(P)$. Also it is easy to check that $\mathbb{L}_2(P)$ is a vector space. To characterize the left ansatz space, we define row shifted sum as follows.

Definition 5.2.4. (row shifted sum) *Let X, Y, Z be $p \times p$ block matrices with $n \times n$ matrices as blocks. Then we define the row shifted sum by*

$$X \boxtimes Y \boxtimes Z = \begin{bmatrix} X(1:k,:) \\ 0 \\ X(k+1:2k-1,:) \\ 0 \\ \vdots \\ 0 \\ X(p-2:p-1,:) \\ 0 \\ X(p,:) \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ Y(1:k,:) \\ 0 \\ Y(k+1:2k-1,:) \\ 0 \\ \vdots \\ 0 \\ Y(p-2:p-1,:) \\ 0 \\ Y(p,:) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ Z(1,:) \\ \vdots \\ Z(p,:) \end{bmatrix},$$

where $X(j,:), Y(j,:), Z(j,:)$ denote the j^{th} block row of X, Y, Z, respectively, 0 denotes a zero block row with $n \times n$ zero matrices and '+' denotes usual matrix addition.

For example, when $k = 2$, consider the 3×3 block matrices

$$X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}, Y = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix}, Z = \begin{bmatrix} Z_1 \\ Z_2 \\ Z_3 \end{bmatrix},$$

where X_j, Y_j, Z_j are block rows with each blocks being an $n \times n$ matrix. Then

$$X \boxtimes Y \boxtimes Z = \begin{bmatrix} X(1:2,:) \\ 0 \\ X(3,:) \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ Y(1:2,:) \\ 0 \\ Y(3,:) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ Z(1,:) \\ Z(2,:) \\ Z(3,:) \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 + Y_1 \\ Y_2 \\ X_3 + Z_1 \\ Y_3 + Z_2 \\ Z_3 \end{bmatrix}.$$

Also, when $k = 3$, consider the 6×6 block matrices

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_6 \end{bmatrix}, Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_6 \end{bmatrix}, Z = \begin{bmatrix} Z_1 \\ \vdots \\ Z_6 \end{bmatrix},$$

where X_j, Y_j, Z_j are block rows with each blocks being an $n \times n$ matrix. Then

$$X \boxtimes Y \boxtimes Z = \begin{bmatrix} X(1:3,:) \\ 0 \\ X(4:5,:) \\ 0 \\ X(6,:) \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ Y(1:3,:) \\ 0 \\ Y(4:5,:) \\ 0 \\ Y(6,:) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ Z(1,:) \\ \vdots \\ Z(6,:) \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 + Y_1 \\ X_3 + Y_2 \\ Y_3 \\ X_4 + Z_1 \\ X_5 + Y_4 + Z_2 \\ Y_5 + Z_3 \\ X_6 + Z_4 \\ Y_6 + Z_5 \\ Z_6 \end{bmatrix}.$$

The row shifted sum of three $p \times p$ block matrices is a $q \times p$ block matrix which we characterize in the next result.

Lemma 5.2.3. *Let X, Y, Z be $p \times p$ block matrices with $n \times n$ matrices. Let T be $q \times p$ block matrix with $n \times n$ matrices be such that $X \boxtimes Y \boxtimes Z = T$. Then*

$$T(1,:) = X(1,:), T(k+1,:) = Y(k,:), T(q,:) = Z(p,:),$$

$$T(j,:) = X(j,:) + Y(j-1,:) \qquad \text{for } j = 2, \cdots, k,$$

$$T\left(\frac{j}{2}(2k-j+3)+1,:\right) = X\left(\frac{j}{2}(2k-j+1)+1,:\right) + Z\left(\frac{j-1}{2}(2k-j+2)+1,:\right)$$

for $j = 1, \cdots, k-1,$

$$T\left(\frac{j+1}{2}(2k-j+2),:\right) = Y\left(\frac{j+1}{2}(2k-j),:\right) + Z\left(\frac{j}{2}(2k-j+1),:\right)$$

for $j = 1, \cdots, k-1,$

$$\begin{aligned} T\left(\frac{j}{2}(2k-j+3)+(r+1),:\right) &= X\left(\frac{j}{2}(2k-j+1)+(r+1),:\right) + Y\left(\frac{j}{2}(2k-j+1)+r,:\right) \\ &\quad + Z\left(\frac{j-1}{2}(2k-j+2)+(r+1),:\right) \end{aligned}$$

for $j = 1, \cdots, k-2, r = 1, \cdots, k-j-1.$

Proof. The proof is computational and follows from Definition 5.2.4. ☐

Next we derive the row shifted sum of the coefficient matrices of the second standard form given in (5.14).

Proposition 5.2.4. *Let* $P(\lambda, \mu)$ *be as given in (5.3). Consider the second standard from* $L_2(\lambda, \mu) = \lambda \mathbb{S}^{\mathcal{B}} + \mu \mathbb{T}^{\mathcal{B}} + \mathbb{A}^{\mathcal{B}}$ *given in (5.14). Then* $\mathbb{S}^{\mathcal{B}} \boxtimes \mathbb{T}^{\mathcal{B}} \boxtimes \mathbb{A}^{\mathcal{B}} = e_1^T \otimes \begin{bmatrix} \mathcal{A}_k^{\mathcal{B}} \\ \vdots \\ \mathcal{A}_0^{\mathcal{B}} \end{bmatrix}.$

Proof. Proof is similar to the proof of the Proposition 5.2.1. ☐

For a two-parameter quadratic matrix polynomial, we have the following result.

Corollary 5.2.7. *Let* $Q(\lambda, \mu)$ *be as given in (5.4). Let* $L_2(\lambda, \mu)$ *be the second standard form as given in (5.2). Then the row shifted sum of the coefficient matrices of* $L_2(\lambda, \mu)$ *is* $e_1^T \otimes \begin{bmatrix} A_{20} & A_{11} & A_{02} & A_{10} & A_{01} & A_{00} \end{bmatrix}^{\mathcal{B}}$.

Next we characterize a two-parameter matrix pencil $L(\lambda, \mu) \in \mathbb{L}_2(P)$ with left ansatz vector w via row shifted sum of the coefficient matrices of $L(\lambda, \mu)$.

Lemma 5.2.4. *Let* $P(\lambda, \mu)$ *be as given in (5.3). Let* $L(\lambda, \mu) = \lambda X + \mu Y + Z$ *be a two-parameter matrix pencil such that* X, Y, Z *are* $p \times p$ *block matrices with* $n \times n$

matrices as blocks. Then $L(\lambda,\mu) \in \mathbb{L}_2(P)$ *with left ansatz vector* $w \in \mathbb{C}^p$ *if and only if* $X \boxtimes Y \boxtimes Z = w^T \otimes \begin{bmatrix} \mathcal{A}_k^{\mathcal{B}} \\ \vdots \\ \mathcal{A}_0^{\mathcal{B}} \end{bmatrix}$.

Proof. Note that $(\Lambda^T \otimes I_n) \cdot (L(\lambda,\mu))$

$$\begin{aligned}
= \;& \lambda^k X(1:,) + \mu^k Y(k,:) + Z(p,:) + \sum_{j=2}^{k} \lambda^{k-j+1}\mu^{j-1}\left(X(j,:) + Y(j-1,:)\right) \\
&+ \sum_{j=1}^{k-1} \lambda^{k-j}\left(X\left(\frac{j}{2}(2k-j+1)+1,:\right) + Z\left(\frac{j-1}{2}(2k-j+2)+1,:\right)\right) \\
&+ \sum_{j=1}^{k-1} \mu^{k-j}\left(Y\left(\frac{j+1}{2}(2k-j),:\right) + Z\left(\frac{j}{2}(2k-j+1)\right)\right) \\
&+ \sum_{j=1}^{k-2}\sum_{r=1}^{k-j-1} \lambda^{k-j-r}\mu^{r}\left(X\left(\frac{j}{2}(2k-j+1)+(r+1),:\right) + Y\left(\frac{j}{2}(2k-j+1)+r,:\right)\right. \\
&\left. + Z\left(\frac{j-1}{2}(2k-j+2)+(r+1),:\right)\right). \qquad (5.35)
\end{aligned}$$

Now $L(\lambda,\mu) \in \mathbb{L}_2(P)$ with left ansatz vector $w \in \mathbb{C}^p$ if and only if

$$(\Lambda^T \otimes I_n) \cdot L(\lambda,\mu) = w^T \otimes P(\lambda,\mu) = \sum_{i=0}^{k}\sum_{j=0}^{k-i} \lambda^i \mu^j \left(w^T \otimes A_{ij}\right). \qquad (5.36)$$

Then comparing the coefficient for the respective powers in (5.35) and (5.36) we have,

$$X(1,:) = w^T \otimes A_{k0},\, Y(k,:) = w^T \otimes A_{0k},\, Z(p,:) = w^T \otimes A_{00}, \qquad (5.37)$$

$$X(j,:) + Y(j-1,:) = w^T \otimes A_{k-j+1,j-1} \quad \text{for } j = 2,\cdots,k, \qquad (5.38)$$

$$X\left(\frac{j}{2}(2k-j+1)+1,:\right) + Z\left(\frac{j-1}{2}(2k-j+2)+1,:\right) = w^T \otimes A_{k-j,0} \qquad (5.39)$$

for $j = 1,\cdots,k-1$,

$$Y\left(\frac{j+1}{2}(2k-j),:\right) + Z\left(\frac{j}{2}(2k-j+1),:\right) = w^T \otimes A_{0,k-j} \qquad (5.40)$$

for $j = 1,\cdots,k-1$,

$$\begin{aligned}
&X\left(\frac{j}{2}(2k-j+1)+(r+1),:\right) + Y\left(\frac{j}{2}(2k-j+1)+r,:\right) \\
&+Z\left(\frac{j-1}{2}(2k-j+2)+(r+1),:\right) = w^T \otimes A_{k-j-r,r} \qquad (5.41)
\end{aligned}$$

for $j = 1, \cdots, k-2, r = 1, \cdots, k-j-1$.

Now by Lemma 5.2.3 and (5.37)-(5.41) it follows that

$$(X \boxtimes Y \boxtimes Z)(1:k+1,:) = \begin{bmatrix} w^T \otimes A_{k0} \\ w^T \otimes A_{k-1,1} \\ \vdots \\ w^T \otimes A_{1,k-1} \\ w^T \otimes A_{0k} \end{bmatrix} = w^T \otimes \mathcal{A}_k^{\mathcal{B}},$$

$$(X \boxtimes Y \boxtimes Z)\left(\frac{j}{2}(2k-j+3)+1 : \frac{j+1}{2}(2k-j+2),:\right) = \begin{bmatrix} w^T \otimes A_{k-j,0} \\ w^T \otimes A_{k-j-1,1} \\ \vdots \\ w^T \otimes A_{1,k-j-1} \\ w^T \otimes A_{0,k-j} \end{bmatrix} = w^T \otimes \mathcal{A}_{k-j}^{\mathcal{B}}$$

for $j = 1, \cdots, k-1$, and $(X \boxtimes Y \boxtimes Z)(q,:) = w^T \otimes A_{00} = w^T \otimes \mathcal{A}_0^{\mathcal{B}}$.

Hence we have $X \boxtimes Y \boxtimes Z = w^T \otimes \begin{bmatrix} \mathcal{A}_k^{\mathcal{B}} \\ \vdots \\ \mathcal{A}_0^{\mathcal{B}} \end{bmatrix}$. □

In particular, for a quadratic two-parameter matrix polynomial we have the following result.

Corollary 5.2.8. *Let $Q(\lambda, \mu)$ be as given in (5.4). Let $L(\lambda, \mu) = \lambda X + \mu Y + Z$ be a two-parameter matrix pencil such that X, Y, Z are 3×3 block matrices with each blocks being $n \times n$ matrix. Then $L(\lambda, \mu) \in \mathbb{L}_2(Q)$ with left ansatz vector $w \in \mathbb{C}^3$ if and only if $X \boxtimes Y \boxtimes Z = w^T \otimes \begin{bmatrix} A_{20} & A_{11} & A_{02} & A_{10} & A_{01} & A_{00} \end{bmatrix}^{\mathcal{B}}$.*

Now we use Lemma 5.2.4 to characterize a two-parameter matrix pencil in $\mathbb{L}_2(P)$.

Theorem 5.2.9. *Let $P(\lambda, \mu)$ be as given in (5.3) and $L(\lambda, \mu) \in \mathbb{L}_2(P)$ with left ansatz vector $w \in \mathbb{C}^p$. Then*

$$L(\lambda, \mu) = \lambda \left(\begin{bmatrix} w^T \otimes \mathcal{S}^{\mathcal{B}} \\ 0 \\ \vdots \\ 0 \end{bmatrix} + X \right) + \mu \left(\begin{bmatrix} w^T \otimes \mathcal{T}^{\mathcal{B}} \\ 0 \\ \vdots \\ 0 \end{bmatrix} + Y \right) + \begin{bmatrix} w^T \otimes \mathcal{A}_{k-1}^{\mathcal{B}} \\ \vdots \\ w^T \otimes \mathcal{A}_0^{\mathcal{B}} \end{bmatrix} + Z, \tag{5.42}$$

where X, Y, Z are $p \times p$ block matrices with $n \times n$ blocks such that $X \boxtimes Y \boxtimes Z = 0$.

Proof. Define a linear map $\mathcal{M} : \mathbb{L}_2(P) \to \mathcal{W}_P$ by $L(\lambda, \mu) \mapsto (\Lambda^T \otimes I_n) \cdot L(\lambda, \mu)$. First to show that $\mathcal{M}$ is surjective, let $w^T \otimes P(\lambda, \mu) \in \mathcal{W}_P$ for some $w \in \mathbb{C}^p$. Consider the linear two-parameter matrix pencil

$$\widehat{L}(\lambda, \mu) := \lambda \begin{bmatrix} w^T \otimes \mathcal{S}^{\mathcal{B}} \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \mu \begin{bmatrix} w^T \otimes \mathcal{T}^{\mathcal{B}} \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} w^T \otimes \mathcal{A}^{\mathcal{B}}_{k-1} \\ \vdots \\ w^T \otimes \mathcal{A}^{\mathcal{B}}_0 \end{bmatrix}$$

Note that

$$\begin{bmatrix} w^T \otimes \mathcal{S}^{\mathcal{B}} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \boxtimes \begin{bmatrix} w^T \otimes \mathcal{T}^{\mathcal{B}} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \boxtimes \begin{bmatrix} w^T \otimes \mathcal{A}^{\mathcal{B}}_{k-1} \\ \vdots \\ w^T \otimes \mathcal{A}^{\mathcal{B}}_0 \end{bmatrix} = w^T \otimes \begin{bmatrix} \mathcal{A}^{\mathcal{B}}_k \\ \vdots \\ \mathcal{A}^{\mathcal{B}}_0 \end{bmatrix}.$$

and thus by Lemma 5.2.4 we have $(\Lambda^T \otimes I_n) \cdot \widehat{L}(\lambda, \mu) = w^T \otimes P(\lambda, \mu)$. Therefore $\mathcal{M}$ is a surjective map. Now the set of all $\mathcal{M}$-preimages of $w^T \otimes P(\lambda, \mu)$ is $\widehat{L}(\lambda, \mu) + \ker \mathcal{M}$. Note that $\ker \mathcal{M} = \{L(\lambda, \mu) = \lambda X + \mu Y + Z \in \mathbb{L}_2(P) : (\Lambda^T \otimes I_n) \cdot L(\lambda, \mu) = 0\}$. By Lemma 5.2.4 it follows that, if $(\Lambda^T \otimes I_n) \cdot (\lambda X + \mu Y + Z) = 0$ then $X \boxtimes Y \boxtimes Z = 0$. Hence (5.42) follows. $\square$

For a two-parameter quadratic matrix polynomial, we have the following result.

Corollary 5.2.10. *Let $Q(\lambda, \mu)$ be as given in (5.4) and $L(\lambda, \mu) \in \mathbb{L}_2(Q)$ with left ansatz vector $w \in \mathbb{C}^3$. Then*

$$L(\lambda, \mu) = \lambda \begin{bmatrix} w^T \otimes A_{20} \\ w^T \otimes A_{11} - Y_1 \\ -Z_1 \end{bmatrix} + \mu \begin{bmatrix} Y_1 \\ w^T \otimes A_{02} \\ -Z_2 \end{bmatrix} + \begin{bmatrix} w^T \otimes A_{10} + Z_1 \\ w^T \otimes A_{01} + Z_2 \\ w^T \otimes A_{00} \end{bmatrix},$$

where Y_1, Z_1, Z_2 are arbitrary block rows with $n \times n$ matrices as blocks.

Next we compute the dimension of $\mathbb{L}_2(P)$.

Proposition 5.2.5. *Let* $P(\lambda,\mu)$ *be as given in (5.3). Then*

$$\dim\, \mathbb{L}_2(P) = \frac{k(k+1)}{2}(k^2-1)n^2 + \frac{k(k+1)}{2}.$$

Proof. The proof is similar to that of Proposition 5.2.2. □

Note that the dimensions of the right and left ansatz space are same. In particular for $k=2$, we have $\dim(\mathbb{L}_1(\mathbb{Q})) = \dim(\mathbb{L}_2(\mathbb{Q})) = 9n^2+3$.

Proposition 5.2.6. *Let* $P(\lambda,\mu)$ *be as given in (5.3). Then* $\mathbb{L}_2(P) = [\mathbb{L}_1(P^T)]^T$, *where* $[\mathbb{L}_1(P^T)]^T := \big\{(L(\lambda,\mu))^T : L(\lambda,\mu) \in \mathbb{L}_1(P^T)\big\}$.

Proof. Let $L(\lambda,\mu) \in \mathbb{L}_2(P)$ with left ansatz vector $w \in \mathbb{C}^p$, i.e., $(\Lambda^T \otimes I_n)\cdot L(\lambda,\mu) = w^T \otimes P(\lambda,\mu)$. Taking transpose on both sides we have $(L(\lambda,\mu))^T \cdot (\Lambda \otimes I_n) = w \otimes (P(\lambda,\mu))^T$ which implies that $(L(\lambda,\mu))^T \in \mathbb{L}_1(P^T)$ with right ansatz vector w. So $\mathbb{L}_2(P) \subseteq [\mathbb{L}_1(P^T)]^T$. The other inclusion, i.e., $[\mathbb{L}_1(P^T)]^T \subseteq \mathbb{L}_2(P)$ follows similarly. □

Let $y \in \ker(P(\lambda,\mu))$ and $L(\lambda,\mu) \in \mathbb{L}_2(P)$ with a nonzero right ansatz vector $w \in \mathbb{C}^p$. Then $(\overline{\Lambda}\otimes y)^* L(\lambda,\mu) = y^*(\Lambda^T \otimes I_n)L(\lambda,\mu) = y^*(w^T \otimes P(\lambda,\mu)) = w^T \otimes y^* P(\lambda,\mu) = 0$. So $\overline{\Lambda}\otimes y \in \ker(L(\lambda,\mu))$. In fact, it is easy to see that the map $\ker(P(\lambda)) \to \ker(L(\lambda)), y \mapsto \overline{\Lambda}\otimes y$ is an isomorphism.

Proposition 5.2.7. *Let* $P(\lambda,\mu)$ *be a two-parameter matrix polynomial of degree* k *as given in (5.3) and* $L(\lambda,\mu) \in \mathbb{L}_2(P)$ *with nonzero left ansatz vector* $w \in \mathbb{C}^p$. *Then* $0 \neq y \in \ker(P(\lambda,\mu))$ *if and only if* $\overline{\Lambda}\otimes y \in \ker(L(\lambda,\mu))$.

5.2.2 Linearizations in $\mathbb{L}_1(P)$

Let $P(\lambda,\mu)$ be a two-parameter matrix polynomial of degree k given by (5.3). We have constructed ansatz space of $P(\lambda,\mu)$. Now the question is whether $L(\lambda,\mu) \in \mathbb{L}_1(P)\bigcup\mathbb{L}_2(P)$ is a linearization of $P(\lambda,\mu)$ or not. It is straightforward to see that $L(\lambda,\mu) \in \mathbb{L}_1(P)$ with right ansatz vector 0 is not a linearization of $P(\lambda,\mu)$. So not all elements in $\mathbb{L}_1(P)$ or $\mathbb{L}_2(P)$ are linearizations of $P(\lambda,\mu)$. Thus it is necessary to derive conditions for $L(\lambda,\mu) \in \mathbb{L}_1(P)\bigcup\mathbb{L}_2(P)$.

First we derive a linearization condition for $L(\lambda,\mu) \in \mathbb{L}_1(P)$ with right ansatz vector $\alpha e_1 \in \mathbb{C}^p$, for some $0 \neq \alpha \in \mathbb{C}$.

Theorem 5.2.11. *Let $P(\lambda,\mu)$ be as given in (5.3). Let $L(\lambda,\mu) \in \mathbb{L}_1(P)$ with right ansatz vector $\alpha e_1 \in \mathbb{C}^p$, for some $0 \neq \alpha \in \mathbb{C}$, be given by*

$$\begin{aligned} L(\lambda,\mu) &= \lambda\left(\begin{bmatrix} \alpha e_1 \otimes \mathcal{S} & 0 & \cdots & 0 \end{bmatrix} + X\right) + \mu\left(\begin{bmatrix} \alpha e_1 \otimes \mathcal{T} & 0 & \cdots & 0 \end{bmatrix} + Y\right) \\ &\quad + \begin{bmatrix} \alpha e_1 \otimes \mathcal{A}_{k-1} & \cdots & \alpha e_1 \otimes \mathcal{A}_0 \end{bmatrix} + Z, \end{aligned}$$

where $X = [X_{ij}], Y = [Y_{ij}], Z = [Z_{ij}]$ are $p \times p$ block matrices with $n \times n$ matrices as blocks such that $X \boxplus Y \boxplus Z = 0$. Further, suppose that the following conditions hold:

(i) $Y_{i,\frac{j-1}{2}(2k-j+2)+r} = 0$, *for* $i = 2, \cdots, p, j = 1, \cdots, k-1, r = 1, \cdots, k-j$,

(ii) $\mathcal{Z} := \begin{bmatrix} Z_{21} & \cdots & Z_{2,p-1} \\ & \vdots & \\ Z_{p1} & \cdots & Z_{p,p-1} \end{bmatrix}$ *is nonsingular.*

Then $L(\lambda,\mu)$ is a linearization of $P(\lambda,\mu)$.

Proof. First note that $L(\lambda,\mu) = \widehat{L}(\lambda,\mu) + (\lambda X + \mu Y + Z)$, where

$$\begin{aligned} \widehat{L}(\lambda,\mu) &= \lambda\left(\begin{bmatrix} \alpha e_1 \otimes \mathcal{S} & 0 & \cdots & 0 \end{bmatrix}\right) + \mu\left(\begin{bmatrix} \alpha e_1 \otimes \mathcal{T} & 0 & \cdots & 0 \end{bmatrix}\right) \\ &\quad + \begin{bmatrix} \alpha e_1 \otimes \mathcal{A}_{k-1} & \cdots & \alpha e_1 \otimes \mathcal{A}_0 \end{bmatrix}. \end{aligned}$$

Define the $(k-i+1) \times (k-i-j+1)$ block matrix

$$R_{ij}(\lambda,\mu) := \left[\begin{array}{c|c} \lambda^j I_{(k-i-j)} \otimes I_n & \\ \hline & \Lambda_j(\lambda,\mu) \otimes I_n \end{array}\right]$$

for $i = 1, \cdots, k-2, j = 1, \cdots, k-i-1$ and consider the $p \times p$ block matrix

$$E(\lambda,\mu) := \left[\begin{array}{c|ccccc} \frac{1}{\alpha}\Lambda_{k-1}(\lambda,\mu) \otimes I_n & I_k \otimes I_n & R_{11}(\lambda,\mu) & R_{12}(\lambda,\mu) & \cdots & R_{1,k-2}(\lambda,\mu) \\ \frac{1}{\alpha}\Lambda_{k-2}(\lambda,\mu) \otimes I_n & & I_{k-1} \otimes I_n & R_{21}(\lambda,\mu) & \cdots & R_{2,k-3}(\lambda,\mu) \\ \vdots & & & \ddots & \ddots & \vdots \\ \frac{1}{\alpha}\Lambda_2(\lambda,\mu) \otimes I_n & & & & I_3 \otimes I_n & R_{k-2,1}(\lambda,\mu) \\ \frac{1}{\alpha}\Lambda_1(\lambda,\mu) \otimes I_n & & & & & I_2 \otimes I_n \\ \hline \frac{1}{\alpha}\Lambda_0(\lambda,\mu) \otimes I_n & & & & & \end{array}\right].$$

Now we have

$$\widehat{L}(\lambda,\mu)\cdot\begin{bmatrix} \frac{1}{\alpha}\Lambda_{k-1}(\lambda,\mu)\otimes I_n \\ \frac{1}{\alpha}\Lambda_{k-2}(\lambda,\mu)\otimes I_n \\ \vdots \\ \frac{1}{\alpha}\Lambda_{1}(\lambda,\mu)\otimes I_n \\ \frac{1}{\alpha}\Lambda_{0}(\lambda,\mu)\otimes I_n \end{bmatrix} = \frac{1}{\alpha}\widehat{L}(\lambda,\mu)\cdot(\Lambda(\lambda,\mu)\otimes I_n) = \frac{1}{\alpha}\cdot\alpha e_1\otimes P(\lambda,\mu) = e_1\otimes P(\lambda,\mu).$$

A little calculation shows that

$$\widehat{L}(\lambda,\mu)\cdot E(\lambda,\mu) = \left[\begin{array}{c|ccc} P(\lambda,\mu) & M_1(\lambda,\mu) & \cdots & M_{k-1}(\lambda,\mu) \\ \hline 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \\ 0 & 0 & \cdots & 0 \end{array}\right] \tag{5.43}$$

for some $M_j(\lambda,\mu)\in\mathbb{C}^{n\times(k+1-j)n}$ for $j=1,\cdots,k-1$. Since $X \boxplus Y \boxplus Z = 0$, by Lemma 5.2.2 we have

$$(\lambda X+\mu Y+Z)\cdot\begin{bmatrix} \frac{1}{\alpha}\Lambda_{k-1}(\lambda,\mu)\otimes I_n \\ \frac{1}{\alpha}\Lambda_{k-2}(\lambda,\mu)\otimes I_n \\ \vdots \\ \frac{1}{\alpha}\Lambda_{1}(\lambda,\mu)\otimes I_n \\ \frac{1}{\alpha}\Lambda_{0}(\lambda,\mu)\otimes I_n \end{bmatrix} = \frac{1}{\alpha}(\lambda X+\mu Y+Z)\cdot(\Lambda(\lambda,\mu)\otimes I_n) = 0.$$

By the condition given in (i) and using the fact that $X \boxplus Y \boxplus Z = 0$, we have

$$\lambda X+\mu Y+Z = \begin{bmatrix} * & * & \cdots & * \\ \widehat{Z}_1(\lambda,\mu) & \widehat{Z}_2(\lambda,\mu) & \cdots & \widehat{Z}_k(\lambda,\mu) \end{bmatrix},$$

where

$$\begin{aligned} \widehat{Z}_1(\lambda,\mu) &:= \begin{bmatrix} Z_{21} & \cdots & Z_{2k} \\ & \vdots & \\ Z_{p1} & \cdots & Z_{pk} \end{bmatrix}, \\ \widehat{Z}_2(\lambda,\mu) &:= -\begin{bmatrix} \lambda Z_{21} & \cdots & \lambda Z_{2,k-2} & \lambda Z_{2,k-1}+\mu Z_{2,k} \\ \vdots & & \vdots & \vdots \\ \lambda Z_{p1} & \cdots & \lambda Z_{p,k-2} & \lambda Z_{p,k-1}+\mu Z_{p,k} \end{bmatrix} + \begin{bmatrix} Z_{2,k+1} & \cdots & Z_{2,k+(k-1)} \\ & \vdots & \\ Z_{p,k+1} & \cdots & Z_{p,k+(k-1)} \end{bmatrix}, \end{aligned}$$

$$\begin{aligned}\widehat{Z}_3(\lambda,\mu) \;&:=\; -\begin{bmatrix} \lambda Z_{2,k+1} & \cdots & \lambda Z_{2,k+(k-3)} & \lambda Z_{2,k+(k-2)} + \mu Z_{2,k+(k-1)} \\ \vdots & & \vdots & \vdots \\ \lambda Z_{p,k+1} & \cdots & \lambda Z_{p,k+(k-3)} & \lambda Z_{p,k+(k-2)} + \mu Z_{p,k+(k-1)} \end{bmatrix} \\ &+ \begin{bmatrix} Z_{2,k+(k-1)+1} & \cdots & Z_{2,k+(k-1)+(k-3)} \\ & \vdots & \\ Z_{2,k+(k-1)+1} & \cdots & Z_{2,k+(k-1)+(k-3)} \end{bmatrix},\end{aligned}$$

so on and $\widehat{Z}_k(\lambda,\mu) := -\begin{bmatrix} \lambda Z_{2,p-2} + \mu Z_{2,p-1} \\ \vdots \\ \lambda Z_{p,p-2} + \mu Z_{p,p-1} \end{bmatrix}$. Note that $\widehat{Z}_1(\lambda,\mu)\cdot(I_k \otimes I_n) = \widehat{Z}_1(\lambda,\mu)$.

Further, we have

$$\begin{aligned}\widehat{Z}_1(\lambda,\mu)\cdot R_{11}(\lambda,\mu) \;&=\; \begin{bmatrix} Z_{21} & \cdots & Z_{2k} \\ & \vdots & \\ Z_{p1} & \cdots & Z_{pk} \end{bmatrix} \cdot \left[\begin{array}{c|c} \lambda I_{k-2}\otimes I_n & \\ \hline & \lambda I_n \\ & \mu I_n \end{array}\right] \\ &= \begin{bmatrix} \lambda Z_{21} & \cdots & \lambda Z_{2,k-2} & \lambda Z_{2,k-1} + \mu Z_{2,k} \\ \vdots & & \vdots & \vdots \\ \lambda Z_{p1} & \cdots & \lambda Z_{p,k-2} & \lambda Z_{p,k-1} + \mu Z_{p,k} \end{bmatrix}.\end{aligned}$$

So we have

$$\begin{aligned}&(\widehat{Z}_1(\lambda,\mu)\cdot R_{11}(\lambda,\mu)) + (\widehat{Z}_2(\lambda,\mu)\cdot(I_{k-1}\otimes I_n)) \\ = &\begin{bmatrix} \lambda Z_{21} & \cdots & \lambda Z_{2,k-2} & \lambda Z_{2,k-1} + \mu Z_{2,k} \\ \vdots & & \vdots & \vdots \\ \lambda Z_{p1} & \cdots & \lambda Z_{p,k-2} & \lambda Z_{p,k-1} + \mu Z_{p,k} \end{bmatrix} - \begin{bmatrix} \lambda Z_{21} & \cdots & \lambda Z_{2,k-2} & \lambda Z_{2,k-1} + \mu Z_{2,k} \\ \vdots & & \vdots & \vdots \\ \lambda Z_{p1} & \cdots & \lambda Z_{p,k-2} & \lambda Z_{p,k-1} + \mu Z_{p,k} \end{bmatrix} \\ &+ \begin{bmatrix} Z_{2,k+1} & \cdots & Z_{2,k+(k-1)} \\ & \vdots & \\ Z_{p,k+1} & \cdots & Z_{p,k+(k-1)} \end{bmatrix} = \begin{bmatrix} Z_{2,k+1} & \cdots & Z_{2,k+(k-1)} \\ & \vdots & \\ Z_{p,k+1} & \cdots & Z_{p,k+(k-1)} \end{bmatrix}.\end{aligned}$$

Continuing this way, we obtain that

$$(\lambda X + \mu Y + Z)\cdot E(\lambda,\mu) = \left[\begin{array}{c|ccc} 0 & \widehat{M}_1(\lambda,\mu) & \cdots & \widehat{M}_{k-1}(\lambda,\mu) \\ \hline 0 & Z_{21} & \cdots & Z_{2,p-1} \\ \vdots & & \vdots & \\ 0 & Z_{p1} & \cdots & Z_{p,p-1} \end{array}\right] \tag{5.44}$$

for some $\widehat{M_j}(\lambda,\mu) \in \mathbb{C}^{n\times(k+1-j)n}$ for $j = 1,\cdots,k-1$. Finally combining (5.43) and (5.44), we have

$$L(\lambda,\mu)\cdot E(\lambda,\mu) = \left[\begin{array}{c|c} P(\lambda,\mu) & \mathcal{M}(\lambda,\mu) \\ \hline & \mathcal{Z} \end{array}\right],$$

where

$$\mathcal{M}(\lambda,\mu) = \left[\begin{array}{ccc} M_1(\lambda,\mu)+\widehat{M}_1(\lambda,\mu) & \cdots & M_{k-1}(\lambda,\mu)+\widehat{M}_{k-1}(\lambda,\mu) \end{array}\right].$$

By the condition given in (ii), $\mathcal{Z}$ is nonsingular. Consider

$$F(\lambda,\mu) := \left[\begin{array}{c|c} I_n & -\mathcal{M}(\lambda,\mu)\mathcal{Z}^{-1} \\ \hline & \mathcal{Z}^{-1} \end{array}\right].$$

Then

$$\begin{aligned} F(\lambda,\mu)\cdot L(\lambda,\mu)\cdot E(\lambda,\mu) &= \left[\begin{array}{c|c} I_n & -\mathcal{M}(\lambda,\mu)\mathcal{Z}^{-1} \\ \hline & \mathcal{Z}^{-1} \end{array}\right]\cdot\left[\begin{array}{c|c} P(\lambda,\mu) & \mathcal{M}(\lambda,\mu) \\ \hline & \mathcal{Z} \end{array}\right] \\ &= \left[\begin{array}{c|c} P(\lambda,\mu) & \\ \hline & I_{(p-1)n} \end{array}\right]. \end{aligned}$$

It is easy to see that $\det(E(\lambda,\mu))$ and $\det(F(\lambda,\mu))$ are independent of λ,μ. Hence $L(\lambda,\mu)$ is a linearization of $P(\lambda,\mu)$. □

For a two-parameter quadratic matrix polynomial we have the following result.

Corollary 5.2.12. *Let $Q(\lambda,\mu)$ be as given in (5.4). Let $L(\lambda,\mu) \in \mathbb{L}_1(Q)$ with nonzero right ansatz vector $v = \alpha e_1$, for some $0 \neq \alpha \in \mathbb{C}$, be given by*

$$\begin{aligned} L(\lambda,\mu) &= \lambda\left[\begin{array}{ccc} \alpha e_1\otimes A_{20} & \alpha e_1\otimes A_{11}-Y_1 & -Z_1 \end{array}\right]+\mu\left[\begin{array}{ccc} Y_1 & \alpha e_1\otimes A_{02} & -Z_2 \end{array}\right] \\ &\quad+\left[\begin{array}{ccc} \alpha e_1\otimes A_{10}+Z_1 & \alpha e_1\otimes A_{01}+Z_2 & \alpha e_1\otimes A_{00} \end{array}\right], \end{aligned}$$

where

$$Y_1 = \begin{bmatrix} Y_{11} \\ Y_{21} \\ Y_{31} \end{bmatrix}, Z_1 = \begin{bmatrix} Z_{11} \\ Z_{21} \\ Z_{31} \end{bmatrix}, Z_2 = \begin{bmatrix} Z_{12} \\ Z_{22} \\ Z_{32} \end{bmatrix}$$

are 3×3 block matrices with $n\times n$ blocks such that *(i)* $Y_{21} = Y_{31} = 0$,

(ii) $\mathcal{Z} = \begin{bmatrix} Z_{21} & Z_{22} \\ Z_{31} & Z_{32} \end{bmatrix}$ *is nonsingular.*

Then $L(\lambda, \mu)$ *is a linearization of* $Q(\lambda, \mu)$.

Next we use Theorem 5.2.11 to determine linearization condition of $L(\lambda, \mu) \in \mathbb{L}_1(P)$ with right ansatz vector $0 \neq v \in \mathbb{C}^p$.

Procedure to determine the linearization condition for a two-parameter matrix pencil in $\mathbb{L}_1(P)$:

(1) Let $L(\lambda, \mu) \in \mathbb{L}_1(P)$ corresponding to a nonzero right ansatz vector $v \in \mathbb{C}^p$ be given by

$$\begin{aligned} L(\lambda, \mu) &= \lambda\left(\begin{bmatrix} v \otimes \mathcal{S} & 0 & \cdots & 0 \end{bmatrix} + X\right) + \mu\left(\begin{bmatrix} v \otimes \mathcal{T} & 0 & \cdots & 0 \end{bmatrix} + Y\right) \\ &\quad + \begin{bmatrix} v \otimes \mathcal{A}_{k-1} & \cdots & v \otimes \mathcal{A}_0 \end{bmatrix} + Z, \end{aligned}$$

where $X = [X_{ij}], Y = [Y_{ij}], Z = [Z_{ij}]$ are such that $X \boxplus Y \boxplus Z = 0$.

(2) Consider a nonsingular matrix $M = [m_{ij}] \in \mathbb{C}^{p \times p}$ such that $Mv = \alpha e_1 \in \mathbb{C}^p$, where $0 \neq \alpha \in \mathbb{C}$.

(3) Apply the corresponding block transformation $M \otimes I_n$ to $L(\lambda, \mu)$. Then we have

$$\begin{aligned} \widetilde{L}(\lambda, \mu) &= (M \otimes I_n) \cdot L(\lambda, \mu) \\ &= \lambda\left(\begin{bmatrix} \alpha e_1 \otimes \mathcal{S} & 0 & \cdots & 0 \end{bmatrix} + \widetilde{X}\right) + \mu\left(\begin{bmatrix} \alpha e_1 \otimes \mathcal{T} & 0 & \cdots & 0 \end{bmatrix} + \widetilde{Y}\right) \\ &\quad + \begin{bmatrix} \alpha e_1 \otimes \mathcal{A}_{k-1} & \cdots & \alpha e_1 \otimes \mathcal{A}_0 \end{bmatrix} + \widetilde{Z}, \end{aligned}$$

where $\widetilde{X} = [\widetilde{X}_{ij}], \widetilde{Y} = [\widetilde{Y}_{ij}], \widetilde{Z} = [\widetilde{Z}_{ij}]$ are given by $\widetilde{X}_{ij} = \sum_{l=1}^{p} m_{il} X_{lj}, \widetilde{Y}_{ij} = \sum_{l=1}^{p} m_{il} Y_{lj}, \widetilde{Z}_{ij} = \sum_{l=1}^{p} m_{il} Z_{lj}$ for all $i, j = 1, \cdots, p$. Note that $\widetilde{L}(\lambda, \mu) \in \mathbb{L}_1(P)$ corresponding to the nonzero right ansatz vector αe_1.

(4) (Linearization Condition)

(i) $\widetilde{Y}_{i, \frac{j-1}{2}(2k-j+2)+r} = 0$ for $i = 2, \cdots, p, j = 1, \cdots, k-1$ and $r = 1, \cdots, k-j$.

(ii) $\widetilde{\mathcal{Z}} := \begin{bmatrix} \widetilde{Z}_{21} & \cdots & \widetilde{Z}_{2,p-1} \\ & \vdots & \\ \widetilde{Z}_{p1} & \cdots & \widetilde{Z}_{p,p-1} \end{bmatrix}$ is nonsingular. $\square$

Example 5.1. Let $Q(\lambda, \mu)$ be a two-parameter quadratic matrix polynomial given in (5.4) and $L(\lambda, \mu) \in \mathbb{L}_1(Q)$ with right ansatz vector $v = (1, 2, 3)^T$. Then it follows from Corollary 5.2.10 that

$$\begin{aligned} L(\lambda,\mu) &= \lambda \begin{bmatrix} A_{20} & A_{11} - Y_{11} & -Z_{11} \\ 2A_{20} & 2A_{11} - Y_{21} & -Z_{21} \\ 3A_{20} & 3A_{11} - Y_{31} & -Z_{31} \end{bmatrix} + \mu \begin{bmatrix} Y_{11} & A_{02} & -Z_{12} \\ Y_{21} & 2A_{02} & -Z_{22} \\ Y_{31} & 3A_{02} & -Z_{32} \end{bmatrix} \\ &\quad + \begin{bmatrix} A_{10} + Z_{11} & A_{01} + Z_{12} & A_{00} \\ 2A_{10} + Z_{21} & 2A_{01} + Z_{22} & 2A_{00} \\ 3A_{10} + Z_{31} & 3A_{01} + Z_{32} & 3A_{00} \end{bmatrix}, \end{aligned}$$

where $Y_{ij}, Z_{ij} \in \mathbb{C}^{n \times n}$ are arbitrary matrices. Consider $M = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$ and note that $Mv = e_1$. Let $\widetilde{L}(\lambda, \mu) := (M \otimes I_n) \cdot L(\lambda, \mu)$. Then

$$(M \otimes I_n) \cdot \begin{bmatrix} Y_{11} \\ Y_{21} \\ Y_{31} \end{bmatrix} = \begin{bmatrix} I_n & 0 & 0 \\ -2I_n & I_n & 0 \\ -3I_n & 0 & I_n \end{bmatrix} \begin{bmatrix} Y_{11} \\ Y_{21} \\ Y_{31} \end{bmatrix} = \begin{bmatrix} Y_{11} \\ Y_{21} - 2Y_{11} \\ Y_{31} - 3Y_{11} \end{bmatrix}$$

and

$$(M \otimes I_n) \cdot \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \\ Z_{31} & Z_{32} \end{bmatrix} = \begin{bmatrix} I_n & 0 & 0 \\ -2I_n & I_n & 0 \\ -3I_n & 0 & I_n \end{bmatrix} \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \\ Z_{31} & Z_{32} \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} - 2Z_{11} & Z_{22} - 2Z_{12} \\ Z_{31} - 3Z_{11} & Z_{32} - 3Z_{12} \end{bmatrix}$$

Hence the linearization conditions are:

(i) $Y_{21} - 2Y_{11} = 0 = Y_{31} - 3Y_{11}$, i.e., $Y_{21} = 2Y_{11}$ and $Y_{31} = 3Y_{11}$,

(ii) $\begin{bmatrix} Z_{21} - 2Z_{11} & Z_{22} - 2Z_{12} \\ Z_{31} - 3Z_{11} & Z_{31} - 3Z_{12} \end{bmatrix}$ is nonsingular.

5.2.3 Double ansatz space

Let $Q(\lambda,\mu)$ be a two-parameter quadratic matrix polynomial given in (5.4). Then we define the double ansatz space by $\mathbb{DL}(Q) := \mathbb{L}_1(Q)\bigcap\mathbb{L}_2(Q)$. It is easy to see that $L(\lambda,\mu)\in\mathbb{DL}(Q)\Leftrightarrow L(\lambda,\mu)\cdot(\Lambda\otimes I_n) = v\otimes Q(\lambda,\mu)$ and $(\Lambda^T\otimes I_n)\cdot L(\lambda,\mu) = v^T\otimes I_n$ for some $v\in\mathbb{C}^3$. Lemma 5.2.2 and Lemma 5.2.4 can be used to characterize $\mathbb{DL}(Q)$. Consider

$$\begin{aligned} L(\lambda,\mu) \;&:=\; \lambda\begin{bmatrix} 0 & 0 & \alpha A_{20} \\ 0 & 0 & \alpha A_{11}+Z \\ \alpha A_{20} & -Z & \alpha A_{10}\end{bmatrix} + \mu\begin{bmatrix} 0 & 0 & -Z \\ 0 & 0 & \alpha A_{02} \\ \alpha A_{11}+Z & \alpha A_{02} & \alpha A_{01}\end{bmatrix} \\ &\quad + \begin{bmatrix} -\alpha A_{20} & Z & 0 \\ -\alpha A_{11}-Z & -\alpha A_{02} & 0 \\ 0 & 0 & \alpha A_{00}\end{bmatrix}, \end{aligned} \tag{5.45}$$

where $0\neq\alpha\in\mathbb{C}$ and $Z\in\mathbb{C}^{n\times n}$ is an arbitrary matrix. Then it is easy to see that $L(\lambda,\mu)\cdot(\Lambda\otimes I_n) = v\otimes Q(\lambda,\mu)$ and $(\Lambda^T\otimes I_n)\cdot L(\lambda,\mu) = v^T\otimes I_n$, where $v = (0,0,\alpha)^T\in\mathbb{C}^3$ and $\Lambda = (\lambda,\mu,1)^T\in\mathbb{C}^3$. So $L(\lambda,\mu)\in\mathbb{DL}(Q)$.

Theorem 5.2.13. *Let $Q(\lambda,\mu)$ be as given in (5.4) and $L(\lambda,\mu)$ be as given in (5.45). Then $L(\lambda,\mu)$ is a linearization of $Q(\lambda,\mu)$ if $\widehat{Z} = \begin{bmatrix} Z & -\alpha A_{20} \\ -\alpha A_{02} & -\alpha A_{11}-Z\end{bmatrix}$ is nonsingular.*

Proof. Consider $E(\lambda,\mu) = \begin{bmatrix} \frac{1}{\alpha}\lambda I_n & 0 & I_n \\ \frac{1}{\alpha}\mu I_n & I_n & 0 \\ \frac{1}{\alpha}I_n & 0 & 0\end{bmatrix}$ and note that

$$\begin{aligned} L(\lambda,\mu)\cdot E(\lambda,\mu) \;&=\; \begin{bmatrix} 0 & Z & -\alpha A_{20} \\ 0 & -\alpha A_{02} & -\alpha A_{11}-Z \\ Q(\lambda,\mu) & -\lambda Z+\mu\alpha A_{02} & \lambda\alpha A_{20}+\mu\alpha A_{11}+\mu Z\end{bmatrix} \\ &=\; \left[\begin{array}{c|c} 0 & \widehat{Z} \\ \hline Q(\lambda,\mu) & \mathcal{M}(\lambda,\mu)\end{array}\right], \end{aligned}$$

where $\mathcal{M}(\lambda,\mu) := \begin{bmatrix} -\lambda Z+\mu\alpha A_{02} & \lambda\alpha A_{20}+\mu\alpha A_{11}+\mu Z\end{bmatrix}$.

Next construct $F(\lambda,\mu) := \begin{bmatrix} -\mathcal{M}(\lambda,\mu)\widehat{Z}^{-1} & I_n \\ \widehat{Z}^{-1} & 0 \end{bmatrix}$. Then we have

$$F(\lambda,\mu)\cdot L(\lambda,\mu)\cdot E(\lambda,\mu) = \left[\begin{array}{c|c} -\mathcal{M}(\lambda,\mu)\widehat{Z}^{-1} & I_n \\ \hline \widehat{Z}^{-1} & 0 \end{array}\right]\left[\begin{array}{c|c} 0 & \widehat{Z} \\ \hline Q(\lambda,\mu) & \mathcal{M}(\lambda,\mu) \end{array}\right] = \begin{bmatrix} Q(\lambda,\mu) & 0 \\ 0 & I_{2n} \end{bmatrix}.$$

Note that $\det(E(\lambda,\mu))$ and $\det(F(\lambda,\mu))$ are independent of λ and μ and hence $L(\lambda,\mu)$ is a linearization of $Q(\lambda,\mu)$. □

We mention that $L(\lambda,\mu)$ given in (5.45) is the only pencil in $\mathbb{DL}(Q)$ which is a linearization of $Q(\lambda,\mu)$.

5.3 Linearizations of a two-parameter PEP

Let $\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$ be a two-parameter PEP of degree k given by

$$\mathbb{W}(\lambda,\mu) := (W_1(\lambda,\mu), W_2(\lambda,\mu)) \text{ and } W_i(\lambda,\mu) := \sum_{m=0}^{k}\sum_{n=0}^{k-m} A_{mn}^{(i)}\lambda^m\mu^n, \tag{5.46}$$

where $(\lambda,\mu) \in \mathbb{C}^2$, $A_{mn}^{(i)} \in \mathbb{C}^{n_i\times n_i}$ for $i = 1,2$, $m = 0,\cdots,k$ and $n = 0,\cdots,k-m$.

First we recall the definition of the linearization of a two-parameter polynomial MEP.

Definition 5.3.1. *[1] Let $\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$ be a two-parameter nonhomogeneous polynomial MEP of degree k such that $\mathbb{W}(\lambda,\mu) := (W_1(\lambda,\mu), W_2(\lambda,\mu))$ and $\mathbb{L} \in \mathbb{MEP}(pn_1, pn_2)$ be a two-parameter linear MEP such that $\mathbb{L}(\lambda,\mu) := (L_1(\lambda,\mu), L_2(\lambda,\mu))$. Then $\mathbb{L}$ is said to be a linearization of $\mathbb{W}$ if $L_i(\lambda,\mu)$ is a linearization of $W_i(\lambda,\mu)$ for $i = 1,2$.*

Let $\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$ be as given in (5.46). Let $L_i(\lambda,\mu)$ be the first standard form of $W_i(\lambda,\mu)$ for $i = 1,2$. Consider $\widehat{\mathbb{L}} \in \mathbb{MEP}(pn_1, pn_2)$ such that $\widehat{\mathbb{L}}(\lambda,\mu) := (L_1(\lambda,\mu), L_2(\lambda,\mu))$. Then $\widehat{\mathbb{L}}$ is a linearization of $\mathbb{W}$. We refer to the linearization $\widehat{\mathbb{L}}$ as the first standard form of $\mathbb{W}$ given in (5.46). We define the right ansatz space of $\mathbb{W}$ by

$$\begin{aligned} \mathcal{L}_1(\mathbb{W}) \;:=\; & \{\mathbb{L} \in \mathbb{MEP}(pn_1, pn_2) : \mathbb{L}(\lambda,\mu) := (L_1(\lambda,\mu), L_2(\lambda,\mu)) \text{ and} \\ & L_i(\lambda,\mu)\cdot(\Lambda \otimes I_{n_i}) = v_i \otimes W_i(\lambda,\mu),\; v_i \in \mathbb{C}^p \text{ for } i = 1,2\}, \end{aligned}$$

where Λ as given in (5.10). We refer to $\begin{bmatrix} v_1 & v_2 \end{bmatrix} \in \mathbb{C}^{p\times 2}$ as the right ansatz matrix. Note that, the first standard form $\widehat{\mathbb{L}} \in \mathcal{L}_1(\mathbb{W})$ corresponding to the right ansatz matrix $\begin{bmatrix} e_1 & e_1 \end{bmatrix}$.

Next consider $\widetilde{\mathbb{L}} \in \mathbb{MEP}(pn_1, pn_2)$ such that $\widetilde{\mathbb{L}}(\lambda,\mu) := (L_1(\lambda,\mu), L_2(\lambda,\mu))$, where $L_i(\lambda,\mu)$ is the second standard form of $W_i(\lambda,\mu)$ for $i = 1,2$. Then $\widetilde{\mathbb{L}}$ is also a linearization of $\mathbb{W}$. We refer to the linearization $\widetilde{\mathbb{L}}$ as the second standard form of $\mathbb{W}$. Now we define the left ansatz space of $\mathbb{W}$ by

$$\begin{aligned}\mathcal{L}_2(\mathbb{W}) \;\; := \;\; & \{\mathbb{L} \in \mathbb{MEP}(pn_1, pn_2) : \mathbb{L}(\lambda,\mu) := (L_1(\lambda,\mu), L_2(\lambda,\mu)) \text{ and} \\ & (\Lambda^T \otimes I_{n_i}) \cdot L_i(\lambda,\mu) = w_i^T \otimes W_i(\lambda,\mu),\; w_i \in \mathbb{C}^p \text{ for } i = 1,2\big\}.\end{aligned}$$

We refer to $\begin{bmatrix} w_1 & w_2 \end{bmatrix} \in \mathbb{C}^{p\times 2}$ as the left ansatz matrix. It is easy to see that $\widetilde{\mathbb{L}} \in \mathcal{L}_2(\mathbb{W})$ corresponding to the left ansatz matrix $\begin{bmatrix} e_1 & e_1 \end{bmatrix}$.

Theorem 5.3.1. (Eigenvector Recovery Property) *Let $\mathbb{W} \in \mathbb{PMEP}(k; n_1, n_2)$ be as given in (5.46).*

Let $\mathbb{L} \in \mathcal{L}_1(\mathbb{W})$ be a linearization of $\mathbb{W}$ with right ansatz matrix $\begin{bmatrix} v_1 & v_2 \end{bmatrix}$, where $0 \neq v_i \in \mathbb{C}^p$ for $i = 1,2$. Then $0 \neq x_1 \otimes x_2 \in \mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2}$ is a right eigenvector of $\mathbb{W}$ corresponding to an eigenvalue (λ,μ) if and only if $(\Lambda \otimes x_1) \otimes (\Lambda \otimes x_2)$ is a right eigenvector of $\mathbb{L}$ corresponding to an eigenvalue (λ,μ).

Let $\mathbb{L} \in \mathcal{L}_2(\mathbb{W})$ be a linearization of $\mathbb{W}$ with left ansatz matrix $\begin{bmatrix} w_1 & w_2 \end{bmatrix}$, where $0 \neq w_i \in \mathbb{C}^p$ for $i = 1,2$. Then $0 \neq y_1 \otimes y_2 \in \mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2}$ is a left eigenvector of $\mathbb{W}$ corresponding to an eigenvalue (λ,μ) if and only if $(\overline{\Lambda} \otimes y_1) \otimes (\overline{\Lambda} \otimes y_2)$ is a left eigenvector of $\mathbb{L}$ corresponding to an eigenvalue (λ,μ).

Proof. The proof follows from Proposition 5.2.3 and Proposition 5.2.7. ☐

6

Structured Backward Perturbation Analysis of MEP

6.1 Introduction

Consider a regular MEP $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ given by $\mathbb{W}(z) := (W_1(z), \cdots, W_m(z))$ with $W_k(z) := A_k + \sum_{j=1}^{m} z_j B_{kj}$ and $z = (z_1, \cdots, z_m)^T \in \mathbb{C}^m$, where $A_k, B_{kj} \in \mathbb{C}^{n_k \times n_k}$ for all $k, j = 1, \cdots, m$. Then $\mathbb{W}$ is said to be Hermitian if all the coefficient matrices of $\mathbb{W}$ are Hermitian. Hermitian and right definite MEPs have been analyzed in [12, 52]. Since structured MEPs arise in many applications, in this chapter, we undertake structured backward perturbation analysis of structured MEPs. Throughout the chapter, we denote the conjugate tranpose of a given matrix A by A^H.

We consider twelve special structured MEPs, namely T-symmetric, T-skew symmetric, T-even, T-odd, T-even alternating, T-odd alternating, H-Hermitian, H-skew Hermitian, H-even, H-odd, H-even alternating and H-odd alternating. These structures are defined in the next section. We denote the set of the structured MEPs by $\mathbb{S}$. Let $\mathbb{W} \in \mathbb{S}$, $\lambda \in \mathbb{C}^m$ and $0 \neq x = x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$. Then we define the structured backward error of (λ, x) as an approximate eigenpair of $\mathbb{W}$ by

$$\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) := \inf\{|||\Delta\mathbb{W}|||_p : \Delta\mathbb{W} \in \mathbb{S} \text{ and } \mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0\},$$

where $1 \leq p \leq \infty$. We show that there exists $\Delta\mathbb{W} \in \mathbb{S}$ such that $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$. Therefore $\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) < \infty$. For the Frobenius and spectral norm on $\mathbb{C}^{n_k \times n_k}$, we determine $\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W})$ and construct $\Delta\mathbb{W}$ such that $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ and $|||\Delta\mathbb{W}|||_p = \eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W})$.

6.2 Structured MEPs

Let $\mathbb{W} \in \mathbb{MEP}(n_1, \cdots, n_m)$ be given by

$$\mathbb{W}(z) := (W_1(z), \cdots, W_m(z)), \text{where } W_k(z) := A_k + \sum_{j=1}^{m} z_j B_{kj}, \tag{6.1}$$

$z = (z_1, \cdots, z_m)^T \in \mathbb{C}^m$ and $A_k, B_{kj} \in \mathbb{C}^{n_k \times n_k}$ for all $k, j = 1, \cdots, m$. Recall that $\mathbb{W}$ is said to be regular if $\prod_{j=1}^{m} \det(W_j(z)) \neq 0$ for some $z \in \mathbb{C}^m$. Also recall that, for $x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ and $z \in \mathbb{C}^m$, we have

$$\mathbb{W}(z)x := (W_1(z)x_1, \ldots, W_m(z)x_m) \;\text{ and }\; x^H\mathbb{W}(z) := (x_1^H W_1(z), \ldots, x_m^H W_m(z)).$$

Let $\mathbb{S}$ be a (real or complex) linear subspace of $\mathbb{MEP}(n_1, \cdots, n_m)$. We refer to the MEPs in $\mathbb{S}$ as the structured MEPs. We consider the following structures:

- **T-symmetric:** $(W_k(z))^T = W_k(z)$, i.e., $A_k^T = A_k$ and $B_{kj}^T = B_{kj}$ for all $k, j = 1, \cdots, m$.
- **T-skew-symmetric:** $(W_k(z))^T = -W_k(z)$, i.e., $A_k^T = -A_k$ and $B_{kj}^T = -B_{kj}$ for all $k, j = 1, \cdots, m$.
- **T-even:** $(W_k(z))^T = W_k(-z)$, i.e., $A_k^T = A_k$ and $B_{kj}^T = -B_{kj}$ for all $k, j = 1, \cdots, m$.
- **T-odd:** $(W_k(z))^T = -W_k(-z)$, i.e., $A_k^T = -A_k$ and $B_{kj}^T = B_{kj}$ for all $k, j = 1, \cdots, m$.
- **T-even alternating:** $A_k^T = A_k, B_{kj}^T = (-1)^j B_{kj}$ for all $k, j = 1, \cdots, m$.
- **T-odd alternating:** $A_k^T = -A_k, B_{kj}^T = (-1)^{j+1} B_{kj}$ for all $k, j = 1, \cdots, m$.
- **H-Hermitian:** $(W_k(z))^H = W_k(\overline{z})$, i.e., $A_k^H = A_k$ and $B_{kj}^H = B_{kj}$ for all $k, j = 1, \cdots, m$.
- **H-skew-Hermitian:** $(W_k(z))^H = -W_k(\overline{z})$, i.e., $A_k^H = -A_k$ and $B_{kj}^H = -B_{kj}$ for all $k, j = 1, \cdots, m$.

- H-**even:** $(W_k(z))^H = W_k(-\overline{z})$, i.e., $A_k^H = A_k$ and $B_{kj}^H = -B_{kj}$ for all $k, j = 1, \cdots, m$.

- H-**odd:** $(W_k(z))^H = -W_k(-\overline{z})$, i.e., $A_k^H = -A_k$ and $B_{kj}^H = B_{kj}$ for all $k, j = 1, \cdots, m$.

- H-**even alternating:** $A_k^H = A_k, B_{kj}^H = (-1)^j B_{kj}$ for all $k, j = 1, \cdots, m$.

- H-**odd alternating:** $A_k^H = -A_k, B_{kj}^H = (-1)^{j+1} B_{kj}$ for all $k, j = 1, \cdots, m$.

We consider

$$\begin{aligned}\mathbb{S} \;\in\; & \{T\text{-symmetric}, T\text{-skew-symmetric}, T\text{-even}, T\text{-odd}, T\text{-even alternating},\\ & T\text{-odd alternating}, H\text{-Hermitian}, H\text{-skew-Hermitian}, H\text{-even}, H\text{-odd},\\ & H\text{-even alternating}, H\text{-odd alternating}\}\,.\end{aligned}$$

Let $\mathbb{W} \in \mathbb{S}$ be regular. Let $\lambda \in \mathbb{C}^m, 0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ and $0 \neq y := y_1 \otimes \cdots \otimes y_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$. Then (λ, x, y) is an eigentriple of $\mathbb{W}$ if λ is an eigenvalue of $\mathbb{W}$ and x and y are right and left eigenvectors of $\mathbb{W}$ corresponding to λ respectively, i.e., $\mathbb{W}(\lambda)x = 0$ and $y^H\mathbb{W}(\lambda) = 0$.

Theorem 6.2.1. *Let $\mathbb{W} \in \mathbb{S}$ as given in (6.1) be regular. Then we have the followings:*

$\mathbb{S}$	***eigenvalue pairing***	***eigentriple***
T-symmetric/ T-skew-symmetric	λ	$(\lambda, x = x_1 \otimes \cdots \otimes x_m, \overline{x} = \overline{x}_1 \otimes \cdots \otimes \overline{x}_m)$
T-even/ T-odd	$(\lambda, -\lambda)$	$(\lambda, x = x_1 \otimes \cdots \otimes x_m, \overline{y} = \overline{y}_1 \otimes \cdots \otimes \overline{y}_m)$, $(-\lambda, y = y_1 \otimes \cdots \otimes y_m, \overline{x} = \overline{x}_1 \otimes \cdots \otimes \overline{x}_m)$
T-even alternating/ T-odd alternating	$(\lambda, \widehat{\lambda})$, *where* $\widehat{\lambda}_j = (-1)^j \lambda_j$	$(\lambda, x = x_1 \otimes \cdots \otimes x_m, \overline{y} = \overline{y_1} \otimes \cdots \otimes \overline{y_m})$ $(\widehat{\lambda}, y = y_1 \otimes \cdots \otimes y_m, \overline{x} = \overline{x_1} \otimes \cdots \otimes \overline{x_m})$
H-Hermitian / H-skew-Hermitian	$(\lambda, \overline{\lambda})$	$(\lambda, x = x_1 \otimes \cdots \otimes x_m, y = y_1 \otimes \cdots \otimes y_m)$ $(\overline{\lambda}, y = y_1 \otimes \cdots \otimes y_m, x = x_1 \otimes \cdots \otimes x_m)$
H-even/ H-odd	$(\lambda, -\overline{\lambda})$	$(\lambda, x = x_1 \otimes \cdots \otimes x_m, y = y_1 \otimes \cdots \otimes y_m)$ $(-\overline{\lambda}, y = y_1 \otimes \cdots \otimes y_m, x = x_1 \otimes \cdots \otimes x_m)$
H-even alternating/ H-odd alternating	$(\lambda, \widehat{\lambda})$, *where* $\widehat{\lambda}_j = (-1)^j \overline{\lambda_j}$	$(\lambda, x = x_1 \otimes \cdots \otimes x_m, y = y_1 \otimes \cdots \otimes y_m)$ $(\widehat{\lambda}, y = y_1 \otimes \cdots \otimes y_m, x = x_1 \otimes \cdots \otimes x_m)$

Proof. Let $\mathbb{W}$ be T-symmetric. Using $(W_k(\lambda))^T = W_k(\lambda)$, we obatin that if $W_k(\lambda)x_k = 0$ then $\overline{x}_k^H W_k(\lambda) = 0$ for all $k = 1, \cdots, m$. So $(\lambda, x = x_1 \otimes \cdots \otimes x_m, \overline{x} = \overline{x}_1 \otimes \cdots \otimes \overline{x}_m)$ is an eigentriple of $\mathbb{W}$. Proof is similar when $\mathbb{W}$ is T-skew symmetric.

Let $\mathbb{W}$ be T-even. BY $(W_k(\lambda))^T = W_k(-\lambda)$, it follows that if $W_k(\lambda)x_k = 0 = W_k(-\lambda)y_k$ then we have $\overline{x}_k^H W_k(-\lambda) = 0 = \overline{y}_k^H W_k(\lambda)$ for all $k = 1, \cdots, m$. This shows that $(\lambda, -\lambda)$ pairing of eigenvalues and $(\lambda, x = x_1 \otimes \cdots \otimes x_m, \overline{y} = \overline{y}_1 \otimes \cdots \otimes \overline{y}_m)$ and $(-\lambda, y = y_1 \otimes \cdots \otimes y_m, \overline{x} = \overline{x}_1 \otimes \cdots \otimes \overline{x}_m)$ are eigentriples of $\mathbb{W}$. Proof is similar when $\mathbb{W}$ is T-odd.

Let $\mathbb{W}$ be T-even alternating. It follows from the definition of T-even alternating MEP that $(W_k(\lambda))^T = A_k + \sum_{j=1}^{m}(-1)^j \lambda_j B_{kj} = W_k(\widehat{\lambda})$ for all $k = 1, \cdots, m$. Then it follows that if $W_k(\lambda)x_k = 0 = W_k(\widehat{\lambda})y_k$ then $\overline{x}_k^H W_k(\widehat{\lambda}) = 0 = \overline{y}_k^H W_k(\lambda) = 0$ for all $k = 1, \cdots, m$. Thus $(\lambda, \widehat{\lambda})$ pairing of eigenvalues and $(\lambda, x = x_1 \otimes \cdots \otimes x_m, \overline{y} = \overline{y}_1 \otimes \cdots \otimes \overline{y}_m)$ and $(\widehat{\lambda}, y = y_1 \otimes \cdots \otimes y_m, \overline{x} = \overline{x}_1 \otimes \cdots \otimes \overline{x}_m)$ are eigentriples of $\mathbb{W}$. Proof is similar when $\mathbb{W}$ is T-odd alternating.

Let $\mathbb{W}$ be H-Hermitian. In this case using $(W_k(\lambda))^H = W_k(\overline{\lambda})$, we observe that if $W_k(\lambda)x_k = 0 = W_k(\overline{\lambda})y_k = 0$ then $x_k^H W_k(\overline{\lambda}) = 0 = y_k^H W_k(\lambda)$ for all $k = 1, \cdots, m$. Thus $(\lambda, \overline{\lambda})$ pairing of eigenvalues and $(\lambda, x = x_1 \otimes \cdots \otimes x_m, y = y_1 \otimes \cdots \otimes y_m)$ and $(\overline{\lambda}, y = y_1 \otimes \cdots \otimes y_m, x = x_1 \otimes \cdots \otimes x_m)$ are eigentriples of $\mathbb{W}$. Proof is similar when $\mathbb{W}$ is H-skew-Hermitian.

Let $\mathbb{W}$ be H-even. Then by $(W_k(\lambda))^H = W_k(-\overline{\lambda})$, it follows that if $W_k(\lambda)x_k = 0 = W_k(-\overline{\lambda})y_k$ then $x_k^H W_k(-\overline{\lambda}) = 0 = y_k^H M_k(\lambda)$ for all $k = 1, \cdots, m$. So $(\lambda, -\overline{\lambda})$ are pairing of eigenvalues and $(\lambda, x = x_1 \otimes \cdots \otimes x_m, y = y_1 \otimes \cdots \otimes y_m)$ and $(-\overline{\lambda}, y = y_1 \otimes \cdots \otimes y_m, x = x_1 \otimes \cdots \otimes x_m)$ are eigentriples of $\mathbb{W}$. Proof is similar when $\mathbb{W}$ is H-odd.

Let $\mathbb{W}$ be H-even alternating. By definition $(W_k(\lambda))^H = A_k + \sum_{j=1}^{m}(-1)^j \overline{\lambda_j} B_{kj} = W_k(\widehat{\lambda})$ and it follows that if $W_k(\lambda)x_k = 0 = W_k(\widehat{\lambda})y_k$ then $x_k^H W_k(\widehat{\lambda}) = 0 = y_k^H W_k(\lambda)$ for all $k = 1, \cdots, m$. So $(\lambda, \widehat{\lambda})$ are pairing of eigenvalues and $(\lambda, x = x_1 \otimes \cdots \otimes x_m, y = y_1 \otimes \cdots \otimes y_m)$ and $(\widehat{\lambda}, y = y_1 \otimes \cdots \otimes y_m, x = x_1 \otimes \cdots \otimes x_m)$ are eigentriples of $\mathbb{W}$. Proof is similar when $\mathbb{W}$ is H-odd alternating. $\square$

Next we show that given $\mathbb{W} \in \mathbb{S}$, $\lambda \in \mathbb{C}^m$ and $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ there exists $\Delta\mathbb{W} \in \mathbb{S}$ such that (λ, x) is an eigenpair of $\mathbb{W} + \Delta\mathbb{W}$.

Theorem 6.2.2. *Let* $\mathbb{W} \in \mathbb{S}$ *given in (6.1) be regular. Let* $\lambda \in \mathbb{C}^m$, $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ *be such that* $x_k^H x_k = 1$ *and set* $r_k := -W_k(\lambda)x_k$ *for all* $k = 1, \cdots, m$. *Let* $y = (y_0, \cdots, y_m)^T \in \partial||(1,\lambda)||_q$, *where* $p^{-1} + q^{-1} = 1$. *Define*

$$\Delta A_k := \begin{cases} -\overline{x_k}x_k^T A_k x_k x_k^H + \frac{\overline{y_0}}{||(1,\lambda)||_q}[\overline{x_k}r_k^T + r_k x_k^H - 2(x_k^T r_k)\overline{x_k}x_k^H] & \text{if } A_k = A_k^T \\ \frac{\overline{y_0}}{||(1,\lambda)||_q}[r_k x_k^H - \overline{x_k}r_k^T] & \text{if } A_k = -A_k^T \end{cases}$$

$$\Delta B_{kj} := \begin{cases} -\overline{x_k}x_k^T B_{kj} x_k x_k^H + \frac{\overline{y_j}}{||(1,\lambda)||_q}[\overline{x_k}r_k^T + r_k x_k^H - 2(x_k^T r_k)\overline{x_k}x_k^H] & \text{if } B_{kj} = B_{kj}^T \\ \frac{\overline{y_j}}{||(1,\lambda)||_q}[r_k x_k^H - \overline{x_k}r_k^T] & \text{if } B_{kj} = -B_{kj}^T \end{cases}$$

and

$$\Delta A_k := \begin{cases} \begin{array}{l} \frac{1}{||(1,\lambda)||_q}[y_0 x_k r_k^H(I_{n_k} - x_k x_k^H) + \overline{y_0}(I_{n_k} - x_k x_k^H) r_k x_k^H] \\ -x_k x_k^H A_k x_k x_k^H \end{array} & \text{if } A_k = A_k^H \\ \begin{array}{l} -\frac{1}{||(1,\lambda)||_q}[y_0 x_k r_k^H(I_{n_k} - x_k x_k^H) - \overline{y_0}(I_{n_k} - x_k x_k^H) r_k x_k^H] \\ -x_k x_k^H A_k x_k x_k^H \end{array} & \text{if } A_k = -A_k^H, \end{cases}$$

$$\Delta B_{kj} := \begin{cases} \begin{array}{l} \frac{1}{||(1,\lambda)||_q}[y_j x_k r_k^H(I_{n_k} - x_k x_k^H) + \overline{y_j}(I_{n_k} - x_k x_k^H) r_k x_k^H] \\ -x_k x_k^H B_{kj} x_k x_k^H \end{array} & \text{if } B_{kj} = B_{kj}^H \\ \begin{array}{l} -\frac{1}{||(1,\lambda)||_q}[y_j x_k r_k^H(I_{n_k} - x_k x_k^H) - \overline{y_j}(I_{n_k} - x_k x_k^H) r_k x_k^H] \\ -x_k x_k^H B_{kj} x_k x_k^H \end{array} & \text{if } B_{kj} = -B_{kj}^H \end{cases}$$

for all $k, j = 1, \cdots, m$. *Consider* $\Delta W_k(z) := \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$ *for* $k = 1, \cdots, m$ *and* $\Delta\mathbb{W}(z) := (\Delta W_1(z), \cdots, \Delta W_m(z))$. *Then* $\Delta\mathbb{W} \in \mathbb{S}$ *and* $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$.

Proof. By the construction of ΔA_k and ΔB_{kj}, we have

$$(\Delta A_k)^T = \begin{cases} \Delta A_k & \text{if } A_k \text{ is symmetric} \\ -\Delta A_k & \text{if } A_k \text{ is skew symmetric} \end{cases}, \tag{6.2}$$

$$(\Delta B_{kj})^T = \begin{cases} \Delta B_{kj} & \text{if } B_{kj} \text{ is symmetric} \\ -\Delta B_{kj} & \text{if } B_{kj} \text{ is skew symmetric} \end{cases} \tag{6.3}$$

and

$$(\Delta A_k)^H = \begin{cases} \Delta A_k & \text{if } A_k \text{ is hermitian} \\ -\Delta A_k & \text{if } A_k \text{ is skew hermitian} \end{cases}, \tag{6.4}$$

$$(\Delta B_{kj})^H = \begin{cases} \Delta B_{kj} & \text{if } B_{kj} \text{ is hermitian} \\ -\Delta B_{kj} & \text{if } B_{kj} \text{ is skew hermitian} \end{cases} \tag{6.5}$$

for all $k, j = 1, \cdots, m$. By (6.2)-(6.5) it follows that if $\mathbb{W} \in \mathbb{S}$ then $\Delta\mathbb{W} \in \mathbb{S}$.

Suppose that $\mathbb{S} = \{T\text{-symmetric}\}$ and $\mathbb{W} \in \mathbb{S}$. Then it follows that

$$\begin{aligned} \Delta W_k(\lambda)x_k &= -\overline{x_k}x_k^T W_k(\lambda)x_k + \frac{\overline{y_0} + \sum_{j=1}^m \lambda_j \overline{y_j}}{||(1,\lambda)||_q}[r_k - (x_k^T r_k)\overline{x_k}] \\ &= (x_k^T r_k)\overline{x_k} + r_k - (x_k^T r_k)\overline{x_k} \quad (\text{since } y \in \partial||(1,\lambda)||_q) \\ &= r_k = -W_k(\lambda)x_k \quad \text{for all } k = 1, \cdots, m. \end{aligned}$$

So $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$. The proof is similar for T-skew symmetric MEPs.

Next suppose that $\mathbb{S} = \{T\text{-even}\}$ and $\mathbb{W} \in \mathbb{S}$. Note that $x_k^T B_{kj} x_k = 0$ since B_{kj} is skew symmetric for $k, j = 1, \cdots, m$ and thus $x_k^T r_k = -x_k^T A_k x_k$. Now we have,

$$\begin{aligned} \Delta W_k(\lambda)x_k &= -(x_k^T A_k x_k)\overline{x_k} + \frac{\overline{y_0} + \sum_{j=1}^m \lambda_j \overline{y_j}}{||(1,\lambda)||_q}[r_k - (x_k^T r_k)\overline{x_k}] \\ &= (x_k^T r_k)\overline{x_k} + r_k - (x_k^T r_k)\overline{x_k} = r_k = -W_k(\lambda)x_k \quad \text{for all } k = 1, \cdots, m. \end{aligned}$$

Then $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$. The proof is similar for T-odd MEPs.

Let $\mathbb{S} = \{T\text{-even alternating}\}$ and $\mathbb{W} \in \mathbb{S}$. Note that in this case $x_k^T B_{kj} x_k = 0$ when j is odd and thus it follows that $x_k^T r_k = -x_k^T A_k x_k - \sum_{j \text{ even}} x_k^T B_{kj} x_k$, for all $k = 1, \cdots, m$. Then we have

$$\begin{aligned} \Delta W_k(\lambda)x_k &= -(x_k^T A_k x_k + \sum_{j \text{ even}} x_k^T B_{kj} x_k)\overline{x_k} + \frac{\overline{y_0} + \sum_{j=1}^m \lambda_j \overline{y_j}}{||(1,\lambda)||_q}[r_k - (x_k^T r_k)\overline{x_k}] \\ &= (x_k^T r_k)\overline{x_k} + r_k - (x_k^T r_k)\overline{x_k} = r_k = -W_k(\lambda)x_k \quad \text{for all } k = 1, \cdots, m. \end{aligned}$$

Thus $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$. The proof is similar for T-odd alternating MEPs.

Let $\mathbb{S} = \{H\text{-Hermitian}\}$ and $\mathbb{W} \in \mathbb{S}$. In this case we have,

$$\begin{aligned} \Delta W_k(\lambda)x_k &= -x_k x_k^H (A_k + \sum_{j=1}^m \lambda_j B_{kj})x_k + \frac{y_0 + \sum_{j=1}^m \lambda_j y_j}{||(1,\lambda)||_q}(x_k r_k^H x_k - x_k r_k^H x_k) \\ &\quad + \frac{\overline{y_0} + \sum_{j=1}^m \lambda_j \overline{y_j}}{||(1,\lambda)||_q}(r_k - x_k x_k^H r_k) \\ &= x_k x_k^H r_k + r_k - x_k x_k^H r_k = r_k = -W_k(\lambda)x_k \quad \text{for all } k = 1, \cdots, m. \end{aligned}$$

So $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$. The proof is similar for H-skew Hermitian MEPs.

Let $\mathbb{S} = \{H\text{-even}\}$ and $\mathbb{W} \in \mathbb{S}$. Then we have,

$$\begin{aligned}
\Delta W_k(\lambda)x_k &= -x_k x_k^H(A_k + \sum_{j=1}^{m} \lambda_j B_{kj})x_k + \frac{y_0 - \sum_{j=1}^{m} \lambda_j y_j}{||(1,\lambda)||_q}(x_k r_k^H x_k - x_k r_k^H x_k) \\
&\quad + \frac{\overline{y_0} + \sum_{j=1}^{m} \lambda_j \overline{y_j}}{||(1,\lambda)||_q}(r_k - x_k x_k^H r_k) \\
&= x_k x_k^H r_k + r_k - x_k x_k^H r_k = r_k = -W_k(\lambda)x_k \;\; \text{for all } k = 1, \cdots, m.
\end{aligned}$$

Thus $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$. The proof is similar for H-odd MEPs.

Let $\mathbb{S} = \{H\text{-even alternating}\}$ and $\mathbb{W} \in \mathbb{S}$. Now we have,

$$\begin{aligned}
\Delta W_k(\lambda)x_k &= -x_k x_k^H(A_k + \sum_{j=1}^{m} \lambda_j B_{kj})x_k + \frac{\overline{y_0} + \sum_{j=1}^{m} \lambda_j \overline{y_j}}{||(1,\lambda)||_q}(r_k - x_k x_k^H r_k) \\
&\quad + \frac{y_0 - \sum\limits_{j \text{ odd}} \lambda_j y_j + \sum\limits_{j \text{ even}} \lambda_j y_j}{||(1,\lambda)||_q}(x_k r_k^H x_k - x_k r_k^H x_k) \\
&= x_k x_k^H r_k + r_k - x_k x_k^H r_k = r_k = -W_k(\lambda)x_k \;\; \text{for all } k = 1, \cdots, m
\end{aligned}$$

and $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$. The proof is similar for H-odd alternating MEPs. □

Next we define structured backward error.

Definition 6.2.1. *Let* $\mathbb{W} \in \mathbb{S}$ *be regular. Let* $\lambda \in \mathbb{C}^m$ *and* $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$*. Then the structured backward error of* (λ, x) *as an approximate eigenpair of* $\mathbb{W}$ *is defined by*

$$\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) := \inf\{\|\Delta\mathbb{W}\|_p : \Delta\mathbb{W} \in \mathbb{S} \text{ and } \mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0\},$$

where $1 \leq p \leq \infty$.

Note that by Theorem 6.2.2, it follows that $\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) < \infty$. We prove the following result regarding the smallest p-norm solutions of a system of linear equations which we will use in the sequel.

Lemma 6.2.1. *Let* $a, b_1, \cdots, b_m, c \in \mathbb{C}^n$. *Consider the following minimization problem*

$$\left.\begin{aligned}
&\text{minimize } \; ||(||a||_p, ||b_1||_p, \cdots, ||b_m||_p)||_p \\
&\text{subject to } \; a + \lambda_1 b_1 + \cdots + \lambda_m b_m = c,
\end{aligned}\right\} \tag{6.6}$$

where $\lambda = (\lambda_1, \cdots, \lambda_m)^T \in \mathbb{C}^m$. *Then the solutions of (6.6) are given by* $\widehat{a} = \frac{\overline{y_0}}{||(1,\lambda)||_q}\, c$ *and* $\widehat{b}_j = \frac{\overline{y_j}}{||(1,\lambda)||_q}\, c$ *for all* $j = 1, \cdots, m$, *where* $(y_0, \cdots, y_m)^T \in \partial||(1,\lambda)||_q$, $1 \leq p \leq \infty$ *and* $p^{-1} + q^{-1} = 1$.

Proof. First note that, $\widehat{a} + \sum_{j=1}^m \lambda_j \widehat{b}_j = \frac{\overline{y_0} + \sum_{j=1}^m \lambda_j \overline{y_j}}{||(1,\lambda)||_q}\, c = c$. Now

$$\left|\left|\left(||\widehat{a}||_p, \left|\left|\widehat{b}_1\right|\right|_p, \cdots, \left|\left|\widehat{b}_m\right|\right|_p\right)\right|\right|_p = \frac{||y||_p}{||(1,\lambda)||_q}\, ||c||_p = \frac{||c||_p}{||(1,\lambda)||_q}.$$

Again by Hölder's inequality, we have

$$||c||_p = \left|\left| a + \sum_{j=1}^m \lambda_j b_j \right|\right|_p \leq ||(||a||_p, ||b_1||_p, \cdots, ||b_m||_p)||_p\, ||(1,\lambda)||_q$$

and thus $||(||a||_p, ||b_1||_p, \cdots, ||b_m||_p)||_p \geq \frac{||c||_p}{||(1,\lambda)||_q} = \left|\left|\left(||\widehat{a}||_p, \left|\left|\widehat{b}_1\right|\right|_p, \cdots, \left|\left|\widehat{b}_m\right|\right|_p\right)\right|\right|_p$. This completes the proof. □

We refer to $\widehat{a}, \widehat{b}_1, \cdots \widehat{b}_m$ as the smallest p-norm solution of $a + \sum_{j=1}^m \lambda_j b_j = c$.

6.3 Frobenious norm structured backward error

In this section we derive structured backward errors of structured MEP by considering Frobenious norm on $\mathbb{C}^{n_k \times n_k}$ for all $k = 1, \cdots, m$.

Theorem 6.3.1. *Let* $\mathbb{S}$ *be the space of all* T*-symmetric MEPs. Let* $\mathbb{W} \in \mathbb{S}$ *be as given in (6.1),* $\lambda := (\lambda_1, \cdots, \lambda_m)^T \in \mathbb{C}^m$ *and* $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ *be such that* $x_k^H x_k = 1$. *Consider* $r_k := -W_k(\lambda) x_k$ *for all* $k = 1, \cdots, m$. *Then*

$$\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \frac{1}{||(1,\lambda)||_q} \left|\left|\left(\left(2||r_1||_2^2 - |x_1^T r_1|^2\right)^{\frac{1}{2}}, \cdots, \left(2||r_m||_2^2 - |x_m^T r_m|^2\right)^{\frac{1}{2}} \right)\right|\right|_p, \quad (6.7)$$

where $p^{-1} + q^{-1} = 1$.

Let $(y_0, \cdots, y_m)^T \in \partial||(1,\lambda)||_q$. *Define*

$$\Delta A_k = \frac{\overline{y_0}}{||(1,\lambda)||_q}\left[\overline{x_k} r_k^T + r_k x_k^H - (r_k^T x_k)\overline{x_k} x_k^H\right], \Delta B_{kj} = \frac{\overline{y_j}}{||(1,\lambda)||_q}\left[\overline{x_k} r_k^T + r_k x_k^H - (r_k^T x_k)\overline{x_k} x_k^H\right] \quad (6.8)$$

for all $k, j = 1, \cdots, m$. *Consider* $\Delta W_k(z) := \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$ *for* $k = 1, \cdots, m$ *and* $\Delta\mathbb{W}(z) := (\Delta W_1(z), \cdots, \Delta W_m(z))$. *Then* $\Delta\mathbb{W}$ *is* T*-symmetric,* $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ *and* $\|\Delta\mathbb{W}\|_p = \eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W})$.

Proof. It follows from Theorem 6.2.2 that there exists $\Delta\mathbb{W} \in \mathbb{S}$ such that $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$. Choose $Q^{(k)} \in \mathbb{C}^{n_k \times (n_k - 1)}$ such that $Q_k = \begin{bmatrix} x_k & Q^{(k)} \end{bmatrix}$ is unitary for all $k = 1, \cdots, m$. Define

$$\widetilde{\Delta A}_k := Q_k^T \; \Delta A_k \; Q_k = \begin{bmatrix} a_k & (\widehat{a}_k)^T \\ \widehat{a}_k & \widehat{A}_k \end{bmatrix}, \widetilde{\Delta B}_{kj} := Q_k^T \; \Delta B_{kj} \; Q_k = \begin{bmatrix} b_{kj} & (\widehat{b}_{kj})^T \\ \widehat{b}_{kj} & \widehat{B}_{kj} \end{bmatrix}, \tag{6.9}$$

where $a_k, b_{kj} \in \mathbb{C}$, $\widehat{a}_k, \widehat{b}_{kj} \in \mathbb{C}^{n_k - 1}$ and $\widehat{A}_k, \widehat{B}_{kj} \in \mathbb{C}^{(n_k-1)\times(n_k-1)}$ are symmetric matrices, for all $k, j = 1, \cdots, m$. Since Q_k is unitary and $Q_k^H x_k = e_1$, the first column of the identity matrix, we have

$$\Delta W_k(\lambda)x_k = r_k \Rightarrow \left(Q_k^T \Delta W_k(\lambda) Q_k\right) Q_k^H x_k = Q_k^T r_k \Rightarrow \left(\widetilde{\Delta A}_k + \sum_{j=1}^{m} \lambda_j \widetilde{\Delta B}_{kj}\right) e_1 = Q_k^T r_k. \tag{6.10}$$

By (6.9) we have

$$\begin{bmatrix} a_k + \sum_{j=1}^m \lambda_j b_{kj} \\ \widehat{a}_k + \sum_{j=1}^m \lambda_j \widehat{b}_{kj} \end{bmatrix} = \begin{bmatrix} x_k^T r_k \\ Q^{(k)^T} r_k \end{bmatrix} \quad \text{for all } k = 1, \cdots, m.$$

Let $(y_0, \cdots, y_m)^T \in \partial||(1,\lambda)||_q$. It follows from Lemma 6.2.1 that the smallest p-norm solution of $a_k + \sum_{j=1}^m \lambda_j b_{kj} = x_k^T r_k$ is given by $a_k = \frac{\overline{y_0}}{||(1,\lambda)||_q} x_k^T r_k, b_{kj} = \frac{\overline{y_j}}{||(1,\lambda)||_q} x_k^T r_k$ for $k, j = 1, \cdots, m$. The minimal p-norm solution of $\widehat{a}_k + \sum_{j=1}^m \lambda_j \widehat{b}_{kj} = Q^{(k)^T} r_k$ is given by $\widehat{a}_k = \frac{\overline{y_0}}{||(1,\lambda)||_q} Q^{(k)^T} r_k, \widehat{b}_{kj} = \frac{\overline{y_j}}{||(1,\lambda)||_q} Q^{(k)^T} r_k$ for $k, j = 1, \cdots, m$. Thus we have

$$\widetilde{\Delta A}_k = \begin{bmatrix} \frac{\overline{y_0}}{||(1,\lambda)||_q} x_k^T r_k & \frac{\overline{y_0}}{||(1,\lambda)||_q} (Q^{(k)^T} r_k)^T \\ \frac{\overline{y_0}}{||(1,\lambda)||_q} Q^{(k)^T} r_k & \widehat{A}_k \end{bmatrix} \tag{6.11}$$

$$\widetilde{\Delta B}_{kj} = \begin{bmatrix} \frac{\overline{y_j}}{||(1,\lambda)||_q} x_k^T r_k & \frac{\overline{y_j}}{||(1,\lambda)||_q} (Q^{(k)^T} r_k)^T \\ \frac{\overline{y_j}}{||(1,\lambda)||_q} Q^{(k)^T} r_k & \widehat{B}_{kj} \end{bmatrix}. \tag{6.12}$$

Note that $||\widetilde{\Delta A}_k||_F$ and $||\widetilde{\Delta B}_{kj}||_F$ are minimized when $\widehat{A}_k = \widehat{B}_{kj} = 0$. So setting $\widehat{A}_k = \widehat{B}_{kj} = 0$ we have

$$\begin{aligned} ||\Delta A_k||_F &= ||\widetilde{\Delta A}_k||_F = \frac{|y_0|}{||(1,\lambda)||_q}\left(|x_k^T r_k|^2 + 2||Q^{(k)^T} r_k||_2^2\right)^{\frac{1}{2}} \\ ||\Delta B_{kj}||_F &= ||\widetilde{\Delta B}_{kj}||_F = \frac{|y_j|}{||(1,\lambda)||_q}\left(|x_k^T r_k|^2 + 2||Q^{(k)^T} r_k||_2^2\right)^{\frac{1}{2}}. \end{aligned}$$

Since $Q_kQ_k^H = I_{n_k}$, we have

$$\begin{bmatrix} x_k & Q^{(k)} \end{bmatrix} \begin{bmatrix} x_k^H \\ Q^{(k)^H} \end{bmatrix} = I_{n_k} \Rightarrow Q^{(k)}Q^{(k)^H} = I_{n_k} - x_kx_k^H \Rightarrow \overline{Q^{(k)}}Q^{(k)^T} = I_{n_k} - \overline{x_k}x_k^T.$$

Consequently,

$$||Q^{(k)^T} r_k||_2^2 = ||\overline{Q^{(k)}}Q^{(k)^T} r_k||_2^2 = ||(I_{n_k} - \overline{x_k}x_k^T)r_k||_2^2 = ||r_k||_2^2 - |x_k^T r_k|^2. \tag{6.13}$$

Thus

$$||\Delta A_k||_F = \frac{|y_0|}{||(1,\lambda)||_q}\Big(2||r_k||_2^2 - |x_k^T r_k|^2\Big)^{\frac{1}{2}} \text{ and } ||\Delta B_{kj}||_F = \frac{|y_j|}{||(1,\lambda)||_q}\Big(2||r_k||_2^2 - |x_k^T r_k|^2\Big)^{\frac{1}{2}}.$$

Hence

$$||\Delta W_k||_p = \frac{||y||_p}{||(1,\lambda)||_q}(2||r_k||_2^2 - |x_k^T r_k|^2)^{\frac{1}{2}} = \frac{1}{||(1,\lambda)||_q}(2||r_k||_2^2 - |x_k^T r_k|^2)^{\frac{1}{2}}$$

and

$$\|\Delta \mathbb{W}\|_p = \frac{1}{||(1,\lambda)||_q}\ \Big|\Big|\Big((2||r_1||_2^2 - |x_1^T r_1|^2)^{\frac{1}{2}}, \cdots, (2||r_m||_2^2 - |x_m^T r_m|^2)^{\frac{1}{2}}\Big)\Big|\Big|_p,$$

which yields (6.7).

Using $\overline{Q^{(k)}}Q^{(k)^T} = I_{n_k} - \overline{x_k}x_k^T$, by (6.11) and (6.12), we have

$$\begin{aligned}
\Delta A_k &= \overline{Q_k}\widetilde{\Delta A_k}Q_k^H = \begin{bmatrix} \overline{x_k} & \overline{Q^{(k)}} \end{bmatrix} \begin{bmatrix} \frac{\overline{y_0}}{||(1,\lambda)||_q} x_k^T r_k & \frac{\overline{y_0}}{||(1,\lambda)||_q}(Q^{(k)^T} r_k)^T \\ \frac{\overline{y_0}}{||(1,\lambda)||_q} Q^{(k)^T} r_k & \widehat{A}_k \end{bmatrix} \begin{bmatrix} x_k^H \\ Q^{(k)^H} \end{bmatrix} \\
&= \frac{\overline{y_0}}{||(1,\lambda)||_q}\Big[\overline{x_k}x_k^T r_k x_k^H + \overline{Q^{(k)}}Q_k^T r_k x_k^H + \overline{x_k} r_k^T Q^{(k)}Q^{(k)^H}\Big] + \overline{Q^{(k)}}\widehat{A}_k Q^{(k)^H} \\
&= \frac{\overline{y_0}}{||(1,\lambda)||_q}\Big[\overline{x_k} r_k^T + r_k x_k^H - (x_k^T r_k)\overline{x_k}x_k^H\Big] + \overline{Q^{(k)}}\widehat{A}_k Q^{(k)^H} \qquad (6.14)
\end{aligned}$$

and

$$\begin{aligned}
\Delta B_{kj} &= \overline{Q_k}\widetilde{\Delta B_{kj}}Q_k^H = \begin{bmatrix} \overline{x_k} & \overline{Q^{(k)}} \end{bmatrix} \begin{bmatrix} \frac{\overline{y_j}}{||(1,\lambda)||_q} x_k^T r_k & \frac{\overline{y_j}}{||(1,\lambda)||_q}(Q^{(k)^T} r_k)^T \\ \frac{\overline{y_j}}{||(1,\lambda)||_q} & \widehat{B}_{kj} \end{bmatrix} \begin{bmatrix} x_k^H \\ Q^{(k)^H} \end{bmatrix} \\
&= \frac{\overline{y_j}}{||(1,\lambda)||_q}\Big[\overline{x_k}x_k^T r_k x_k^H + \overline{Q^{(k)}}Q^{(k)^T} r_k x_k^H + \overline{x_k} r_k^T Q^{(k)}Q^{(k)^H}\Big] + \overline{Q^{(k)}}\widehat{B}_{kj} Q^{(k)^H} \\
&= \frac{\overline{y_j}}{||(1,\lambda)||_q}\Big[\overline{x_k} r_k^T + r_k x_k^H - (x_k^T r_k)\overline{x_k}x_k^H\Big] + \overline{Q^{(k)}}\widehat{B}_{kj} Q^{(k)^H} \qquad (6.15)
\end{aligned}$$

for all $k, j = 1, \cdots, m$. Setting $\widehat{A}_k = \widehat{B}_{kj} = 0$ for all $k, j = 1, \cdots, m$ in (6.14) and (6.15), we obtain the desired ΔA_k and ΔB_{kj} given in (6.8). This completes the proof. □

Next we consider T-skew symmetric MEPs.

Theorem 6.3.2. *Let $\mathbb{S}$ be the space of all T-skew symmetric MEPs. Let $\mathbb{W} \in \mathbb{S}$ be as given in (6.1), $\lambda := (\lambda_1, \cdots, \lambda_m)^T \in \mathbb{C}^m$ and $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ be such that $x_k^H x_k = 1$. Consider $r_k := -W_k(\lambda)x_k$ for all $k = 1, \cdots, m$. Then*

$$\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \frac{\sqrt{2}}{||(1,\lambda)||_q} ||(||r_1||_2, \cdots, ||r_m||_2)||_p,$$

where $p^{-1} + q^{-1} = 1$. Further, for the T-skew symmetric MEP $\Delta\mathbb{W}$ given in Theorem 6.2.2, we have $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ and $|||\Delta\mathbb{W}|||_p = \eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W})$.

Proof. In this case, the coefficient matrices are all skew symmetric. So by similar arguments as given in the proof of Theorem 6.3.1, we have

$$\widetilde{\Delta A}_k := Q_k^T \, \Delta A_k \, Q_k = \begin{bmatrix} 0 & -(\widehat{a}_k)^T \\ \widehat{a}_k & \widehat{A}_k \end{bmatrix} \text{ and } \widetilde{\Delta B}_{kj} := Q_k^T \, \Delta B_{kj} \, Q_k = \begin{bmatrix} 0 & -(\widehat{b}_{kj})^T \\ \widehat{b}_{kj} & \widehat{B}_{kj} \end{bmatrix}, \tag{6.16}$$

where $\widehat{a}_k, \widehat{b}_{kj} \in \mathbb{C}^{n_k-1}$ and $\widehat{A}_k, \widehat{B}_{kj} \in \mathbb{C}^{(n_k-1)\times(n_k-1)}$ are skew-symmetric matrices for all $k, j = 1, \cdots, m$. Consequently, we have

$$\begin{bmatrix} 0 \\ \widehat{a}_k + \sum_{j=1}^m \lambda_j \widehat{b}_{kj} \end{bmatrix} = \begin{bmatrix} x_k^T r_k \\ Q^{(k)^T} r_k \end{bmatrix} \quad \text{for all } k = 1, \cdots, m.$$

Note that $x_k^T r_k = 0$ and the smallest p-norm solution of $\widehat{a}_k + \sum_{j=1}^m \lambda_j \widehat{b}_{kj} = Q^{(k)^T} r_k$ is given by $\widehat{a}_k = \frac{\overline{y_0}}{||(1,\lambda)||_q} Q^{(k)^T} r_k, \widehat{b}_{kj} = \frac{\overline{y_j}}{||(1,\lambda)||_q} Q^{(k)^T} r_k$ for $k, j = 1, \cdots, m$. Then

$$\widetilde{\Delta A}_k = \begin{bmatrix} 0 & -\frac{\overline{y_0}}{||(1,\lambda)||_q} (Q^{(k)^T} r_k)^T \\ \frac{\overline{y_0}}{||(1,\lambda)||_q} Q^{(k)^T} r_k & \widehat{A}_k \end{bmatrix}, \tag{6.17}$$

$$\widetilde{\Delta B}_{kj} = \begin{bmatrix} 0 & -\frac{\overline{y_j}}{||(1,\lambda)||_q} (Q^{(k)^T} r_k)^T \\ \frac{\overline{y_j}}{||(1,\lambda)||_q} Q^{(k)^T} r_k & \widehat{B}_{kj} \end{bmatrix}. \tag{6.18}$$

Setting $\widehat{A}_k = \widehat{B}_{kj} = 0$ and using $||Q^{(k)^T} r_k||_2^2 = ||r_k||_2^2 - |x_k^T r_k|^2 = ||r_k||_2^2$, we obtain $||\Delta A_k||_F = \frac{\sqrt{2}\,|y_0|}{||(1,\lambda)||_q} ||r_k||_2$ and $||\Delta B_{kj}||_F = \frac{\sqrt{2}\,|y_j|}{||(1,\lambda)||_q} ||r_k||_2$. So $||\Delta W_k||_p = \frac{\sqrt{2}\,||y||_p}{||(1,\lambda)||_q} ||r_k||_2 = \frac{\sqrt{2}}{||(1,\lambda)||_q} ||r_k||_2$ and $\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) = |||\Delta\mathbb{W}|||_p = \frac{\sqrt{2}}{||(1,\lambda)||_q} ||(||r_1||_2, \cdots, ||r_m||_2)||_p$.

Using $\overline{Q^{(k)}}Q^{(k)^T} = I_{n_k} - \overline{x_k}x_k^T$ and by further simplifications it follows from (6.17) and (6.18) that

$$\Delta A_k = \frac{\overline{y_0}}{||(1,\lambda)||_q}(r_k x_k^H - \overline{x_k} r_k^T) + \overline{Q^{(k)}}\widehat{A}_k Q^{(k)^H}, \Delta B_{kj} = \frac{\overline{y_j}}{||(1,\lambda)||_q}(r_k x_k^H - \overline{x_k} r_k^T) + \overline{Q^{(k)}}\widehat{B}_{kj} Q^{(k)^H}$$

for all $k, j = 1, \cdots, m$. Finally setting $\widehat{A}_k = \widehat{B}_{kj} = 0$, we obtain the T-skew symmetric MEP given in Theorem 6.2.2. $\square$

Next we derive the structured backward error of an approximate eigenpair of T-even MEPs.

Theorem 6.3.3. *Let $\mathbb{S}$ be the space of all T-even MEPs. Let $\mathbb{W} \in \mathbb{S}$ be as given in (6.1), $\lambda := (\lambda_1, \cdots, \lambda_m)^T \in \mathbb{C}^m$ and $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ be such that $x_k^H x_k = 1$. Consider $r_k := -W_k(\lambda)x_k$ for all $k = 1, \cdots, m$. Let $(y_0, \cdots, y_m)^T \in \partial||(1,\lambda)||_q$ for $p^{-1} + q^{-1} = 1$. Then $\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) = ||(c_1, \cdots, c_m)||_p$, where $c_k = ||(a_k, b_{k1}, \cdots, b_{km})||_p$ and*

$$a_k = \left(|x_k^T A_k x_k|^2 + \frac{2\ |y_0|^2}{||(1,\lambda)||_q^2}\left[||r_k||_2^2 - |x_k^T A_k x_k|^2\right]\right)^{\frac{1}{2}},\ b_{kj} = \frac{\sqrt{2}|y_j|}{||(1,\lambda)||_q}\left(||r_k||_2^2 - |x_k^T A_k x_k|^2\right)^{\frac{1}{2}}$$

for all $k, j = 1, \cdots, m$. In particular, for $p = 2$, we have

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \left(\sum_{k=1}^{m}\left(|x_k^T A_k x_k|^2 + \frac{2}{||(1,\lambda)||_2^2}(||r_k||_2^2 - |x_k^T A_k x_k|^2)\right)\right)^{\frac{1}{2}}.$$

Let $\Delta\mathbb{W} \in \mathbb{S}$ be as given in Theorem 6.2.2. Then $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ and $\|\Delta\mathbb{W}\|_p = \eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W})$.

Proof. Note that A_k is symmetric and B_{kj} is skew-symmetric for all $k, j = 1, \cdots, m$. Consider the unitary matrices $Q_k := \begin{bmatrix} x_k & Q^{(k)} \end{bmatrix}$, where $Q^{(k)} \in \mathbb{C}^{n_k \times (n_k - 1)}$ for all $k = 1, \cdots, m$. Define

$$\widetilde{\Delta A}_k := Q_k^T \Delta A_k Q_k = \begin{bmatrix} a_k & (\widehat{a}_k)^T \\ \widehat{a}_k & \widehat{A}_k \end{bmatrix},\ \widetilde{\Delta B}_{kj} := Q_k^T\ \Delta B_{kj}\ Q_k = \begin{bmatrix} 0 & -(\widehat{b}_{kj})^T \\ \widehat{b}_{kj} & \widehat{B}_{kj} \end{bmatrix},$$

where $a_k \in \mathbb{C}$, $\widehat{a}_k, \widehat{b}_{kj} \in \mathbb{C}^{n_k - 1}$, $\widehat{A}_k \in \mathbb{C}^{(n_k-1)\times(n_k-1)}$ is symmetric and $\widehat{B}_{kj} \in \mathbb{C}^{(n_k-1)\times(n_k-1)}$ are skew-symmetric matrices for all $k, j = 1, \cdots, m$. Now $(\widetilde{\Delta A}_k + \sum_{j=1}^{m} \lambda_j \widetilde{\Delta B}_{kj})e_1 =$

$Q_k^T r_k$ implies that $\begin{bmatrix} a_k \\ \widehat{a}_k + \sum_{j=1}^m \lambda_j \widehat{b}_{kj} \end{bmatrix} = \begin{bmatrix} x_k^T r_k \\ Q^{(k)^T} r_k \end{bmatrix}$ for all $k = 1, \cdots, m$. Note that $a_k = x_k^T r_k = -x_k^T A_k x_k$. By Lemma 6.2.1 the smallest p-norm solution of $\widehat{a}_k + \sum_{j=1}^m \lambda_j \widehat{b}_{kj} = Q^{(k)^T} r_k$ is given by $\widehat{a}_k = \frac{\overline{y_0}}{||(1,\lambda)||_q} Q^{(k)^T} r_k$, $\widehat{b}_{kj} = \frac{\overline{y_j}}{||(1,\lambda)||_q} Q^{(k)^T} r_k$. So we have

$$\widetilde{\Delta A}_k = \begin{bmatrix} -x_k^T A_k x_k & \frac{\overline{y_0}}{||(1,\lambda)||_q} (Q^{(k)^T} r_k)^T \\ \frac{\overline{y_0}}{||(1,\lambda)||_q} Q^{(k)^T} r_k & \widehat{A}_k \end{bmatrix} \tag{6.19}$$

$$\widetilde{\Delta B}_{kj} = \begin{bmatrix} 0 & -\frac{\overline{y_j}}{||(1,\lambda)||_q} (Q^{(k)^T} r_k)^T \\ \frac{\overline{y_j}}{||(1,\lambda)||_q} Q^{(k)^T} r_k & \widehat{B}_{kj} \end{bmatrix} \tag{6.20}$$

for $k, j = 1, \cdots, m$. Setting $\widehat{A}_k = \widehat{B}_{kj} = 0$ and noting that $||Q^{(k)^T r_k}||_2^2 = ||r_k||_2^2 - |x_k^T A_k x_k|^2$, we have $||\Delta A_k||_F = \left(|x_k^T A_k x_k|^2 + \frac{2\ |y_0|^2}{||(1,\lambda)||_q^2} \left[||r_k||_2^2 |x_k^T A_k x_k|^2 \right] \right)^{\frac{1}{2}} = a_k$ and $||\Delta B_{kj}||_F = \frac{\sqrt{2}|y_j|}{||(1,\lambda)||_q} \left(||r_k||_2^2 - |x_k^T A_k x_k|^2 \right)^{\frac{1}{2}} = b_{kj}$. Therefore $||\Delta W_k||_p = ||(a_k, b_{k1}, \cdots, b_{km})||_p = c_k$ and thus $\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) = |||\Delta \mathbb{W}|||_p = ||(c_1, \cdots, c_m)||_p$.

Using $\overline{Q^{(k)}} Q^{(k)^T} = I_{n_k} - \overline{x_k} x_k^T$ and simplifying the expressions, it follows from (6.19) and (6.20) that

$$\begin{aligned} \Delta A_k &= -\overline{x_k} x_k^T A_k x_k x_k^H + \frac{\overline{y_0}}{||(1,\lambda)||_q} \left[r_k x_k^H + \overline{x_k} r_k^T - 2(x_k^T A_k x_k) \overline{x_k} x_k^H \right] + \overline{Q^{(k)}} \widehat{A}_k Q^{(k)^H} \\ \Delta B_{kj} &= \frac{\overline{y_j}}{||(1,\lambda)||_q} \left(r_k x_k^H - \overline{x_k} r_k^T \right) + \overline{Q^{(k)}} \widehat{B}_{kj} Q^{(k)^H}. \end{aligned}$$

Setting $\widehat{A}_k = \widehat{B}_{kj} = 0$ we obtain the desired $\Delta \mathbb{W}$ given in Theorem 6.2.2. □

Now we consider the T-odd MEPs. In this case interchanging the role of symmetric and skew-symmetric matrices, the proof follows by similar arguments as those given in Theorem 6.3.3.

Theorem 6.3.4. *Let $\mathbb{S}$ be the space of all T-odd MEPs. Let $\mathbb{W} \in \mathbb{S}$ be as given in (6.1), $\lambda := (\lambda_1, \cdots, \lambda_m)^T \in \mathbb{C}^m$ and $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ be such that $x_k^H x_k = 1$. Consider $r_k := -W_k(\lambda) x_k$ for all $k = 1, \cdots, m$. Let $(y_0, \cdots, y_m)^T \in \partial ||(1,\lambda)||_q$ and $(z_1, \cdots, z_m)^T \in \partial ||\lambda||_q$, where $p^{-1} + q^{-1} = 1$. Then*

$$\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \begin{cases} ||(c_1, \cdots, c_m)||_p & \text{when } \lambda \neq 0 \\ \sqrt{2}\ ||(||r_1||_2, \cdots, ||r_m||_2)||_p & \text{when } \lambda = 0, \end{cases}$$

where $c_k = ||(a_k, b_{k1}, \cdots, b_{km})||_p$ *and*

$$a_k = \sqrt{2}\frac{|y_0|}{||(1,\lambda)||_q}\left(||r_k||_2^2 - |x_k^T r_k|^2\right)^{\frac{1}{2}},\ b_{kj} = \left(\frac{|z_j|^2}{||\lambda||_q^2}|x_k^T r_k|^2 + 2\frac{|y_j|^2}{||(1,\lambda)||_q^2}(||r_k||_2^2 - |x_k^T r_k|^2)\right)^{\frac{1}{2}}$$

for all $k, j = 1, \cdots, m$. *In particular, for* $p = 2$, *we have*

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \begin{cases} \left(\sum_{k=1}^m \left(\frac{1}{||\lambda||_2^2}|x_k^T r_k|^2 + \frac{2}{||(1,\lambda)||_2^2}(||r_k||_2^2 - |x_k^T r_k|^2)\right)\right)^{\frac{1}{2}} & \text{when } \lambda \neq 0 \\ \sqrt{2}\ ||(||r_1||_2, \cdots, ||r_m||_2)||_2 & \text{when } \lambda = 0. \end{cases} \tag{6.21}$$

Let $\Delta\mathbb{W} \in \mathbb{S}$ *be as given in Theorem 6.2.2. Then* $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ *and* $\|\Delta\mathbb{W}\|_p = \eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W})$.

Next we consider T-even alternating and T-odd alternating MEPs. For that purpose we define even coefficient projector $\Pi_{(e)} : \mathbb{C}^n \to \mathbb{C}^n$ as

$$z \mapsto \begin{cases} [0, z_2, \cdots, 0, z_{n-2}, 0, z_n]^T & \text{if } n \text{ is even} \\ [0, z_2, \cdots, 0, z_{n-3}, 0, z_{n-1}, 0]^T & \text{if } n \text{ is odd} \end{cases}$$

for all $z = (z_1, \cdots, z_n)^T \in \mathbb{C}^n$. Also the odd coefficient projector $\Pi_{(o)}$ is defined analogously and it satisfies $\Pi_{(o)}(z) = z - \Pi^{(e)}(z)$ for all $z \in \mathbb{C}^n$.

Theorem 6.3.5. *Let* $\mathbb{S}$ *be the space of all* T*-even alternating MEPs. Let* $\mathbb{W} \in \mathbb{S}$ *be as given in (6.1),* $\lambda := (\lambda_1, \cdots, \lambda_m)^T \in \mathbb{C}^m$ *and* $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ *be such that* $x_k^H x_k = 1$. *Consider* $r_k := -W_k(\lambda)x_k$ *for all* $k = 1, \cdots, m$. *Let* $(y_0, \cdots, y_m)^T \in \partial||(1,\lambda)||_q$ *and* $(z_0, \cdots, z_m)^T \in \partial||(1, \Pi_{(e)}(\lambda))||_q$, *where* $p^{-1} + q^{-1} = 1$. *Then* $\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) = ||(c_1, \cdots, c_m)||_p$, *where* $c_k = ||(a_k, b_{k1}, \cdots, b_{km})||_p$ *and*

$$\begin{aligned} a_k &= \left(\frac{|z_0|^2}{||(1, \Pi_{(e)}(\lambda))||_q^2}|x_k^T r_k|^2 + \frac{2|y_0|^2}{||(1,\lambda)||_q^2}(||r_k||_2^2 - |x_k^T r_k|^2)\right)^{\frac{1}{2}} \\ b_{kj} &= \begin{cases} \left(\frac{|z_j|^2}{||(1,\Pi_{(e)}(\lambda))||_q^2}|x_k^T r_k|^2 + \frac{2|y_j|^2}{||(1,\lambda)||_q^2}(||r_k||_2^2 - |x_k^T r_k|^2)\right)^{\frac{1}{2}} & \text{when } j \text{ is even} \\ \frac{\sqrt{2}|y_j|}{||(1,\lambda)||_q}(||r_k||_2^2 - |x_k^T r_k|^2)^{\frac{1}{2}} & \text{when } j \text{ is odd} \end{cases} \end{aligned}$$

for all $k, j = 1, \cdots, m$. *In particular, when* $p = 2$, *we have*

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \left(\sum_{k=1}^m \left(\frac{1}{||(1, \Pi_{(e)}(\lambda))||_2^2}|x_k^T r_k|^2 + \frac{2}{||(1,\lambda)||_2^2}(||r_k||_2^2 - |x_k^T r_k|^2)\right)\right)^{\frac{1}{2}}.$$

Define

$$\Delta A_k := \frac{\overline{y_0}}{||(1,\lambda)||_q}[r_k x_k^H + \overline{x_k} r_k^T - 2(x_k^T r_k)\overline{x_k} x_k^H] + \frac{\overline{z_0}}{||(1,\Pi_{(e)}(\lambda))||_q}(x_k^T r_k)\overline{x_k} x_k^H, \tag{6.22}$$

$$\Delta B_{kj} \;:=\; \begin{cases} \frac{\overline{y_j}}{||(1,\lambda)||_q}[r_k x_k^H + \overline{x_k} r_k^T - 2(x_k^T r_k)\overline{x_k} x_k^H] & \\ +\frac{\overline{z_j}}{||(1,\Pi_{(e)}(\lambda))||_q}(x_k^T r_k)\overline{x_k} x_k^H & \text{when } j \text{ is even} \\ \frac{\overline{y_j}}{||(1,\lambda)||_q}[r_k x_k^H - \overline{x_k} r_k^T] & \text{when } j \text{ is odd} \end{cases} \tag{6.23}$$

for all $k, j = 1, \cdots, m$ *and* $\Delta W_k(z) := \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$ *for* $k = 1, \cdots, m$. *Consider* $\Delta\mathbb{W}(z) := (\Delta W_1(z), \cdots, \Delta W_m(z))$. *Then* $\Delta\mathbb{W}$ *is* T*-even alternating,* $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ *and* $\|\Delta\mathbb{W}\|_p = \eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W})$.

Proof. In this case, A_k is symmetric, B_{kj} is symmetric when j is even and B_{kj} is skew-symmetric when j is odd for all $k, j = 1, \cdots, m$. Consider the unitary matrices $Q_k := \begin{bmatrix} x_k & Q^{(k)} \end{bmatrix}$, where $Q^{(k)} \in \mathbb{C}^{n_k \times (n_k - 1)}$ for all $k = 1, \cdots, m$. Define

$$\begin{aligned} \widetilde{\Delta A_k} \;&:=\; Q_k^T \, \Delta A_k \, Q_k = \begin{bmatrix} a_k & (\widehat{a}_k)^T \\ \widehat{a}_k & \widehat{A}_k \end{bmatrix}, \\ \widetilde{\Delta B_{kj}} \;&:=\; Q_k^T \, \Delta B_{kj} \, Q_k = \begin{cases} \begin{bmatrix} b_{kj} & (\widehat{b}_{kj})^T \\ \widehat{b}_{kj} & \widehat{B}_{kj} \end{bmatrix} & \text{when } j \text{ is even} \\ \begin{bmatrix} 0 & -(\widehat{b}_{kj})^T \\ \widehat{b}_{kj} & \widehat{B}_{kj} \end{bmatrix} & \text{when } j \text{ is odd,} \end{cases} \end{aligned}$$

where $a_k, b_{kj} \in \mathbb{C}$, $\widehat{a}_k, \widehat{b}_{kj} \in \mathbb{C}^{n_k - 1}$, $\widehat{A}_k \in \mathbb{C}^{(n_k-1)\times(n_k-1)}$ is symmetric and $\widehat{B}_{kj} \in \mathbb{C}^{(n_k-1)\times(n_k-1)}$ are such that $(\widehat{B}_{kj})^T = (-1)^j \widehat{B}_{kj}$ for all $k, j = 1, \cdots, m$. Since $(\widetilde{\Delta A_k} + \sum_{j=1}^m \lambda_j \widetilde{\Delta B_{kj}})e_1 = Q_k^T r_k$, we have $\begin{bmatrix} a_k + \sum\limits_{j \text{ even}} \lambda_j b_{kj} \\ \widehat{a}_k + \sum_{j=1}^m \lambda_j \widehat{b}_{kj} \end{bmatrix} = \begin{bmatrix} x_k^T r_k \\ Q^{(k)^T} r_k \end{bmatrix}$ for all $k = 1, \cdots, m$. By Lemma 6.2.1, the smallest p-norm solution for $a_k + \sum\limits_{j \text{ even}} \lambda_j b_{kj} = x_k^T r_k$ is $a_k = \frac{\overline{z_0}}{||(1,\Pi_{(e)}(\lambda))||_q} x_k^T r_k$, $b_{kj} = \frac{\overline{z_j}}{||(1,\Pi_{(e)}(\lambda))||_q} x_k^T r_k$ when j is even and $b_{kj} = 0$ otherwise. Also the smallest p-norm solution for $\widehat{a}_k + \sum_{j=1}^m \lambda_j \widehat{b}_{kj} = Q^{(k)^T} r_k$ is $\widehat{a}_k = \frac{\overline{y_0}}{||(1,\lambda)||_q} Q^{(k)^T} r_k$ and

$\widehat{b}_{kj} = \frac{\overline{y_j}}{||(1,\lambda)||_q} Q^{(k)^T} r_k$. Then we have

$$\widetilde{\Delta A}_k = \begin{bmatrix} \frac{\overline{z_0}}{||(1,\Pi_{(e)}(\lambda))||_q} x_k^T r_k & \frac{\overline{y_0}}{||(1,\lambda)||_q} (Q^{(k)^T} r_k)^T \\ \frac{\overline{y_0}}{||(1,\lambda)||_q} Q^{(k)^T} r_k & \widehat{A}_k \end{bmatrix}, \tag{6.24}$$

$$\widetilde{\Delta B}_{kj} = \begin{cases} \begin{bmatrix} \frac{\overline{z_j}}{||(1,\Pi_{(e)}(\lambda))||_q} x_k^T r_k & \frac{\overline{y_j}}{||(1,\lambda)||_q} (Q^{(k)^T} r_k)^T \\ \frac{\overline{y_j}}{||(1,\lambda)||_q} Q^{(k)^T} r_k & \widehat{B}_{kj} \end{bmatrix} & \text{when } j \text{ is even} \\ \begin{bmatrix} 0 & -\frac{\overline{y_j}}{||(1,\lambda)||_q} (Q^{(k)^T} r_k)^T \\ \frac{\overline{y_j}}{||(1,\lambda)||_q} Q^{(k)^T} r_k & \widehat{B}_{kj} \end{bmatrix} & \text{when } j \text{ is odd.} \end{cases} \tag{6.25}$$

Therefore setting $\widehat{A}_k = \widehat{B}_{kj} = 0$ and noting that $||Q^{(k)^T} r_k||_2^2 = ||r_k||_2^2 - |x_k^T r_k|^2$, we have $||\Delta A_k||_F = \left(\frac{|z_0|^2}{||(1,\Pi_{(e)}(\lambda))||_q^2} |x_k^T r_k|^2 + \frac{2|y_0|^2}{||(1,\lambda)||_q^2} (||r_k||_2^2 - |x_k^T r_k|^2) \right)^{\frac{1}{2}} = a_k$ and

$$\begin{aligned} ||\Delta B_{kj}||_F &= \begin{cases} \left(\frac{|z_j|^2}{||(1,\Pi_{(e)}(\lambda))||_q^2} |x_k^T r_k|^2 + \frac{2|y_j|^2}{||(1,\lambda)||_q^2} (||r_k||_2^2 - |x_k^T r_k|^2) \right)^{\frac{1}{2}} & \text{when } j \text{ is even} \\ \frac{\sqrt{2}|y_j|}{||(1,\lambda)||_q} (||r_k||_2^2 - |x_k^T r_k|^2)^{\frac{1}{2}} & \text{when } j \text{ is odd} \end{cases} \\ &= b_{kj} \end{aligned}$$

for all $k, j = 1, \cdots, m$. So we have $||\Delta W_k||_p = ||(a_k, b_{k1}, \cdots, b_{km})||_p = c_k$ and hence $\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) = |||\Delta \mathbb{W}|||_p = ||(c_1, \cdots, c_m)||_p$.

By (6.24) and (6.25), it follows that

$$\Delta A_k = \frac{\overline{z_0}}{||(1,\Pi_{(e)}(\lambda))||_q} (x_k^T r_k) \overline{x_k} x_k^H + \frac{\overline{y_0}}{||(1,\lambda)||_q} [r_k x_k^H + \overline{x_k} r_k^T - 2(x_k^T r_k) \overline{x_k} x_k^H] + \overline{Q^{(k)}} \widehat{A}_k Q^{(k)^H} \tag{6.26}$$

and

$$\Delta B_{kj} = \begin{cases} \frac{\overline{z_j}}{||(1,\Pi_{(e)}(\lambda))||_q} (x_k^T r_k) \overline{x_k} x_k^H + \overline{Q^{(k)}} \widehat{B}_{kj} Q^{(k)^H} & \\ + \frac{\overline{y_j}}{||(1,\lambda)||_q} [r_k x_k^H + \overline{x_k} r_k^T - 2(x_k^T r_k) \overline{x_k} x_k^H] & \text{when } j \text{ is even} \\ \frac{\overline{y_j}}{||(1,\lambda)||_q} [r_k x_k^H - \overline{x_k} r_k^T] + \overline{Q^{(k)}} \widehat{B}_{kj} Q^{(k)^H} & \text{when } j \text{ is odd} \end{cases} \tag{6.27}$$

for all $k, j = 1, \cdots, m$. Setting $\widehat{A}_k = \widehat{B}_{kj} = 0$ in (6.26) and (6.27) we obtain the desired perturbations given in (6.22) and (6.23) respectively. □

We next consider T-odd alternating MEPs.

Theorem 6.3.6. *Let* $\mathbb{S}$ *be the space of all* T*-odd alternating MEPs. Let* $\mathbb{W} \in \mathbb{S}$ *be as given in (6.1),* $\lambda := (\lambda_1, \cdots, \lambda_m)^T \in \mathbb{C}^m$ *and* $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ *be such that* $x_k^H x_k = 1$*. Consider* $r_k := -W_k(\lambda)x_k$ *for all* $k = 1, \cdots, m$*. Let* $y := (y_0, \cdots, y_m)^T \in \partial||(1,\lambda)||_q$ *and* $(z_1, \cdots, z_m)^T \in \partial||\Pi_{(o)}(\lambda)||_q$*, where* $p^{-1} + q^{-1} = 1$*. Then*

$$\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \begin{cases} ||(c_1, \cdots, c_m)||_p & \text{when } \Pi_{(o)}(\lambda) \neq 0 \\ \frac{\sqrt{2}\ ||\Pi_{(o)}(y)||_p}{||(1,\lambda)||_q} ||(||r_1||_2, \cdots, ||r_m||_2)||_p & \text{when } \Pi_{(o)}(\lambda) = 0, \end{cases}$$

where $c_k = ||(a_k, b_{k1}, \cdots, b_{km})||_p$, $a_k = \frac{\sqrt{2}|y_0|}{||(1,\lambda)||_q}(||r_k||_2^2 - |x_k^T r_k|^2)^{\frac{1}{2}}$ *and*

$$b_{kj} = \begin{cases} \frac{\sqrt{2}|y_j|}{||(1,\lambda)||_q}(||r_k||_2^2 - |x_k^T r_k|^2)^{\frac{1}{2}} & \text{when } j \text{ is even} \\ \left(\frac{|z_j|^2}{||\Pi_{(o)}(\lambda)||_q^2}|x_k^T r_k|^2 + \frac{2|y_j|^2}{||(1,\lambda)||_q^2}(||r_k||_2^2 - |x_k^T r_k|^2)\right)^{\frac{1}{2}} & \text{when } j \text{ is odd} \end{cases}$$

for all $k, j = 1, \cdots, m$*. In particular when* $p = 2$*, we have*

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \begin{cases} \left(\sum_{k=1}^m \left(\frac{1}{||\Pi_{(o)}(\lambda)||_2^2}|x_k^T r_k|^2 + \frac{2}{||(1,\lambda)||_2^2}(||r_k||_2^2 - |x_k^T r_k|)\right)\right)^{\frac{1}{2}} & \text{when } \Pi_{(o)}(\lambda) \neq 0 \\ \sqrt{2}\ ||(||r_1||_2, \cdots, ||r_m||_2)||_2 & \text{when } \Pi_{(o)}(\lambda) = 0. \end{cases}$$

Define $\Delta A_k := \frac{\overline{y_0}}{||(1,\lambda)||_q}[r_k x_k^H - \overline{x_k} r_k^T]$ *and*

$$\Delta B_{kj} = \begin{cases} \frac{\overline{y_j}}{||(1,\lambda)||_q}[r_k x_k^H - \overline{x_k} r_k^T] & \text{when } j \text{ is even} \\ \frac{\overline{y_j}}{||(1,\lambda)||_q}[r_k x_k^H + \overline{x_k} r_k^T - 2(x_k^T r_k)\overline{x_k} x_k^H] & \\ +\frac{\overline{z_j}}{||\Pi_{(o)}(\lambda)||_q}(x_k^T r_k)\overline{x_k} x_k^H & \text{when } j \text{ is odd} \end{cases}$$

for all $k, j = 1, \cdots, m$ *and* $\Delta W_k(z) := \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$ *for* $k = 1, \cdots, m$*. Consider* $\Delta\mathbb{W}(z) := (\Delta W_1(z), \cdots, \Delta W_m(z))$*. Then* $\Delta\mathbb{W}$ *is* T*-odd alternating,* $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ *and* $\|\Delta\mathbb{W}\|_p = \eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W})$*.*

Proof. Proof follows by the similar arguments as those given in the proof of Theorem 6.3.5. □

Now we derive structured backward errors in case of H-structures. Suppose that $w \in \mathbb{C}$ and $\mathbb{W} \in \mathbb{S}$. Define $w\mathbb{W}$ such that $w\mathbb{W}(z) := (wW_1(z), \cdots, wW_m(z))$ and consider $\mathbb{S}_w := \{w\mathbb{W} : w \in \mathbb{C} \text{ such that } |w| = 1 \text{ and } \mathbb{W} \in \mathbb{S}\}$. Then $\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \eta_p^{\mathbb{S}_w}(\lambda, x, w\mathbb{W})$.

Note that, for a H-Hermitian MEP, all the coefficient matrices are Hermitian whereas all the coefficient matrices are skew-Hermitian for a H-skew-Hermitian MEP. Now a matrix $A \in \mathbb{C}^{n\times n}$ is Hermitian if and only if iA is skew-Hermitian. Then note that

$$\begin{aligned}
&\mathbb{W} \text{ as given in (6.1) is } H\text{-Hermitian}\\
\Leftrightarrow\ & A_k, B_{kj} \text{ are Hermitian for all } k, j = 1, \cdots, m\\
\Leftrightarrow\ & iA_k, iB_{kj} \text{ are skew-Hermitian for all } k, j = 1, \cdots, m\\
\Leftrightarrow\ & i\mathbb{W} \text{ such that } i\mathbb{W}(z) := (iW_1(z), \cdots, iW_m(z)) \text{ is } H\text{-skew-Hermitian.}
\end{aligned}$$

Thus it follows that the maps H-Hermitian $\rightarrow$ H-skew-Hermitian given by $\mathbb{W} \mapsto i\mathbb{W}$ and H-skew-Hermitian $\rightarrow$ H-Hermitian given by $\mathbb{X} \mapsto i\mathbb{X}$ are isometric isomorphisms. Combining these facts, it follows that the structured backward error of (λ, x) as an approximate eigenpair of a skew-Hermitian MEP can be obtained from the structured backward error of (λ, x) as an approximate eigenpair of a Hermitian MEP and vice-versa. By similar arguments, further we can show that the structured backward error of (λ, x) as an approximate eigenpair of a H-odd MEP (H-odd alternating, respectively) can be obtained from the structured backward error of (λ, x) as an approximate eigenpair of a H-even MEP (H-even alternating, respectively) and vice-versa. Now we derive the structured backward error of (λ, x) as an approximate eigenpair of H-Hermitian, H-even and H-even alternating MEPs considering Hölder 2-norm $|||\cdot|||_2$ over $\mathbb{MEP}(n_1, \cdots, n_m)$. We consider the H-Hermitian MEPs in the following.

Theorem 6.3.7. *Let $\mathbb{S}$ be the space of all H-Hermitian MEPs and $\mathbb{W} \in \mathbb{S}$ be as given in (6.1). Let $\lambda := (\lambda_1, \cdots, \lambda_m)^T \in \mathbb{C}^m$ be such that $\lambda_j := \alpha_j + i\beta_j$ for all $j = 1, \cdots, m$, where $i = \sqrt{-1}$ and $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ be such that $x_k^H x_k = 1$. Consider $r_k := -W_k(\lambda)x_k$ for all $k = 1, \cdots, m$. Then*

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \begin{cases} \frac{1}{||(1,\lambda)||_2}\left(\sum_{k=1}^m \left(2||r_k||_2^2 - |x_k^H r_k|^2\right)\right)^{\frac{1}{2}} & \text{when } \lambda \in \mathbb{R}^m \\ \left(\sum_{k=1}^m \left(\sum_{j=0}^m |s_{kj}|^2 + \frac{2}{||(1,\lambda)||_2^2}(||r_k||_2^2 - |x_k^H r_k|^2)\right)\right)^{\frac{1}{2}} & \text{when } \lambda \in \mathbb{C}^m \setminus \mathbb{R}^m, \end{cases} \tag{6.28}$$

where

$$\begin{bmatrix} s_{k0} \\ s_{k1} \\ \vdots \\ s_{km} \end{bmatrix} = \begin{bmatrix} 1 & \alpha_1 & \cdots & \alpha_m \\ 0 & \beta_1 & \cdots & \beta_m \end{bmatrix}^{\dagger} \begin{bmatrix} \mathrm{Re}(x_k^H r_k) \\ \mathrm{Im}(x_k^H r_k) \end{bmatrix} \quad \text{for } k = 1, \cdots, m.$$

When $\lambda \in \mathbb{R}^m$*, define*

$$\begin{aligned} \Delta A_k &:= \tfrac{1}{\|(1,\lambda)\|_2^2}\left(r_k x_k^H + x_k r_k^H - (r_k^H x_k) x_k x_k^H\right) \\ \Delta B_{kj} &:= \tfrac{\lambda_j}{\|(1,\lambda)\|_2^2}\left(r_k x_k^H + x_k r_k^H - (r_k^H x_k) x_k x_k^H\right) \end{aligned}$$

and when $\lambda \in \mathbb{C}^m \setminus \mathbb{R}^m$*, define*

$$\begin{aligned} \Delta A_k &:= x_k s_{k0} x_k^H + \tfrac{1}{\|(1,\lambda)\|_2^2}\left[\left(I_{n_k} - x_k x_k^H\right) r_k x_k^H + x_k r_k^H \left(I_{n_k} - x_k x_k^H\right)\right] \\ \Delta B_{kj} &:= x_k s_{kj} x_k^H + \tfrac{1}{\|(1,\lambda)\|_2^2}\left[\,\overline{\lambda_j}\left(I_{n_k} - x_k x_k^H\right) r_k x_k^H + \lambda_j x_k r_k^H \left(I_{n_k} - x_k x_k^H\right)\right] \end{aligned}$$

for all $k, j = 1, \cdots, m$*. Consider* $\Delta W_k(z) := \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$ *for* $k = 1, \cdots, m$ *and* $\Delta\mathbb{W}(z) := (\Delta W_1(z), \cdots, \Delta W_m(z))$*. Then* $\Delta\mathbb{W}$ *is H-Hermitian,* $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ *and* $\|\Delta\mathbb{W}\|_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W})$*.*

Proof. By Theorem 6.2.2 it follows that there exists $\Delta\mathbb{W} \in \mathbb{S}$ such that $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$. Choose $Q^{(k)} \in \mathbb{C}^{n_k \times (n_k - 1)}$ such that $Q_k = \begin{bmatrix} x_k & Q^{(k)} \end{bmatrix}$ is unitary for all $k = 1, \cdots, m$. Define

$$\widetilde{\Delta A}_k := Q_k^H \Delta A_k Q_k = \begin{bmatrix} a_k & (\widehat{a}_k)^H \\ \widehat{a}_k & \widehat{A}_k \end{bmatrix}, \quad \widetilde{\Delta B}_{kj} := Q_k^H \Delta B_{kj} Q_k = \begin{bmatrix} b_{kj} & (\widehat{b}_{kj})^H \\ \widehat{b}_{kj} & \widehat{B}_{kj} \end{bmatrix},$$

where $a_k, b_{kj} \in \mathbb{R}$, $\widehat{a}_k, \widehat{b}_{kj} \in \mathbb{C}^{n_k - 1}$, $\widehat{A}_k, \widehat{B}_{kj} \in \mathbb{C}^{(n_k-1)\times(n_k-1)}$ are hermitian matrices, for all $k, j = 1, \cdots, m$. Since Q_k is unitary and $Q_k^H x_k = e_1$, we have $\left(Q_k^H \Delta W_k(\lambda) Q_k\right) Q_k^H x_k = Q_k^H r_k \Rightarrow \left(\widetilde{\Delta A}_k + \sum_{j=1}^m \lambda_j \widetilde{\Delta B}_{kj}\right) e_1 = Q_k^H r_k \Rightarrow \begin{bmatrix} a_k + \sum_{j=1}^m \lambda_j b_{kj} \\ \widehat{a}_k + \sum_{j=1}^m \lambda_j \widehat{b}_{kj} \end{bmatrix} = \begin{bmatrix} x_k^H r_k \\ {Q^{(k)}}^H r_k \end{bmatrix}$ for all $k = 1, \cdots, m$.

Case I: Suppose that $\lambda \in \mathbb{R}^m$. Then the minimum norm solution of $a_k + \sum_{j=1}^m \lambda_j b_{kj} = x_k^H r_k$ is given by $a_k = \frac{1}{\|(1,\lambda)\|_2^2} x_k^H r_k$, $b_{kj} = \frac{\lambda_j}{\|(1,\lambda)\|_2^2} x_k^H r_k$ and the minimum norm solution of $\widehat{a}_k + \sum_{j=1}^m \lambda_j \widehat{b}_{kj} = {Q^{(k)}}^H r_k$ is given by $\widehat{a}_k = \frac{1}{\|(1,\lambda)\|_2^2} {Q^{(k)}}^H r_k$, $\widehat{b}_{kj} = \frac{\lambda_j}{\|(1,\lambda)\|_2^2} {Q^{(k)}}^H r_k$ for

$k, j = 1, \cdots, m$. Therefore we have

$$\widetilde{\Delta A}_k = \begin{bmatrix} \frac{1}{||(1,\lambda)||_2^2} x_k^H r_k & \frac{1}{||(1,\lambda)||_2^2}(Q^{(k)^H} r_k)^H \\ \frac{1}{||(1,\lambda)||_2^2} Q^{(k)^H} r_k & \widehat{A}_k \end{bmatrix}, \; \widetilde{\Delta B}_{kj} = \begin{bmatrix} \frac{\lambda_j}{||(1,\lambda)||_2^2} x_k^H r_k & \frac{\lambda_j}{||(1,\lambda)||_2^2}(Q^{(k)^H} r_k)^H \\ \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2} Q^{(k)^H} r_k & \widehat{B}_{kj} \end{bmatrix} \tag{6.29}$$

for all $k, j = 1, \cdots, m$. To derive $||\widetilde{\Delta A}_k||_F$ and $||\widetilde{\Delta B}_{kj}||_F$, now we need to compute $||Q^{(k)^H} r_k||_2$. For that purpose note that $Q_k Q_k^H = I_{n_k} \Rightarrow Q^{(k)} Q^{(k)^H} = I_{n_k} - x_k x_k^H$ and thus we have $||Q^{(k)^H} r_k||_2^2 = ||Q_k Q_k^H r_k||_2^2 = ||(I_{n_k} - x_k x_k^H) r_k||_2^2 = ||r_k||_2^2 - |x_k^H r_k|^2$. Now setting $\widehat{A}_k = \widehat{B}_{kj} = 0$, we have $||\Delta A_k||_F = ||\widetilde{\Delta A}_k||_F = \frac{1}{||(1,\lambda)||_2^2}\left(2||r_k||_2^2 - |x_k^H r_k|^2\right)^{\frac{1}{2}}$ and $||\Delta B_{kj}||_F = ||\widetilde{\Delta B}_{kj}||_F = \frac{|\lambda_j|}{||(1,\lambda)||_2^2}\left(2||r_k||_2^2 - |x_k^H r_k|^2\right)^{\frac{1}{2}}$ for $k, j = 1, \cdots, m$. Therefore $||\Delta W_k||_2 = \frac{1}{||(1,\lambda)||_2}\left(2||r_k||_2^2 - |x_k^H r_k|^2\right)^{\frac{1}{2}}$ for all $k = 1, \cdots, m$ and $\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = |||\Delta \mathbb{W}|||_2 = \frac{1}{||(1,\lambda)||_2}\left(\sum_{k=1}^m \left(2||r_k||_2^2 - |x_k^H r_k|^2\right)\right)^{\frac{1}{2}}$.

Case II: Suppose that $\lambda \in \mathbb{C}^m \setminus \mathbb{R}^m$. By Lemma 6.2.1 it follows that the minimum norm solution of $\widehat{a_k} + \sum_{j=1}^m \lambda_j \widehat{b_{kj}} = Q^{(k)^H} r_k$ is given by $\widehat{a_k} = \frac{1}{||(1,\lambda)||_2} Q^{(k)^H} r_k$, $\widehat{b}_{kj} = \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2} Q^{(k)^H} r_k$. Now to obtain the real minimum norm solution of $a_k + \sum_{j=1}^m \lambda_j b_{kj} = x_k^H r_k$, we rewrite it as $\begin{bmatrix} 1 & \alpha_1 & \cdots & \alpha_m \\ 0 & \beta_1 & \cdots & \beta_m \end{bmatrix} \begin{bmatrix} a_k \\ b_{k1} \\ \vdots \\ b_{km} \end{bmatrix} = \begin{bmatrix} \mathrm{Re}(x_k^H r_k) \\ \mathrm{Im}(x_k^H r_k) \end{bmatrix}$ and then its minimum norm solution is $\begin{bmatrix} a_k \\ b_{k1} \\ \vdots \\ b_{km} \end{bmatrix} = \begin{bmatrix} 1 & \alpha_1 & \cdots & \alpha_m \\ 0 & \beta_1 & \cdots & \beta_m \end{bmatrix}^{\dagger} \begin{bmatrix} \mathrm{Re}(x_k^H r_k) \\ \mathrm{Im}(x_k^H r_k) \end{bmatrix}$. Therefore we have

$$\widetilde{\Delta A}_k = \begin{bmatrix} s_{k0} & \frac{1}{||(1,\lambda)||_2^2}(Q^{(k)^H} r_k)^H \\ \frac{1}{||(1,\lambda)||_2^2} Q^{(k)^H} r_k & \widehat{A}_k \end{bmatrix} \tag{6.30}$$

$$\widetilde{\Delta B}_{kj} = \begin{bmatrix} s_{kj} & \frac{\lambda_j}{||(1,\lambda)||_2^2}(Q^{(k)^H} r_k)^H \\ \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2} Q^{(k)^H} r_k & \widehat{B}_{kj} \end{bmatrix}, \tag{6.31}$$

for all $k, j = 1, \cdots, m$. Setting $\widehat{A}_k = \widehat{B}_{kj} = 0$, we have

$$\begin{aligned} ||\Delta A_k||_F &= \left(|s_{k0}|^2 + \frac{2}{||(1,\lambda)||_2^4}(||r_k||_2^2 - |x_k^H r_k|^2)\right)^{\frac{1}{2}} \\ ||\Delta B_{kj}||_F &= \left(|s_{kj}|^2 + \frac{2|\lambda_j|^2}{||(1,\lambda)||_2^4}(||r_k||_2^2 - |x_k^H r_k|^2)\right)^{\frac{1}{2}}, \end{aligned}$$

Therefore $||\Delta W_k||_2 = \left(\sum_{j=0}^m |s_{kj}|^2 + \frac{2}{||(1,\lambda)||_2^2}(||r_k||_2^2 - |x_k^H r_k|^2)\right)^{\frac{1}{2}}$ for all $k = 1, \cdots, m$ and finally $\eta_2^{\mathbb{S}}(\lambda, x_k, W_k) = \|\Delta\mathbb{W}\|_2 = \left(\sum_{k=1}^m \left(\sum_{j=0}^m |s_{kj}|^2 + \frac{2}{||(1,\lambda)||_2^2}(||r_k||_2^2 - |x_k^H r_k|^2)\right)\right)^{\frac{1}{2}}$. This proves (6.28).

When $\lambda \in \mathbb{R}^m$, using $Q^{(k)}Q^{(k)^H} = I_{n_k} - x_k x_k^H$ and by further simplifications it follows from (6.29) that

$$\begin{aligned} \Delta A_k &= \frac{1}{||(1,\lambda)||_2^2}\left(r_k x_k^H + x_k r_k^H - (r_k^H x_k) x_k x_k^H\right) + Q^{(k)}\widehat{B}_{kj}Q^{(k)^H} \\ \Delta B_{kj} &= \frac{\lambda_j}{||(1,\lambda)||_2^2}\left(r_k x_k^H + x_k r_k^H - (r_k^H x_k) x_k x_k^H\right) + Q^{(k)}\widehat{B}_{kj}Q^{(k)^H}. \end{aligned}$$

When $\lambda \in \mathbb{C}^m \setminus \mathbb{R}^m$, by (6.30) and (6.31), we have,

$$\begin{aligned} \Delta A_k &= x_k s_{k0} x_k^H + \frac{1}{||(1,\lambda)||_2^2}\left[\left(I_{n_k} - x_k x_k^H\right) r_k x_k^H + x_k r_k^H \left(I_{n_k} - x_k x_k^H\right)\right] + Q^{(k)}\widehat{A}_{kj}Q^{(k)^H} \\ \Delta B_{kj} &= x_k s_{kj} x_k^H + \frac{1}{||(1,\lambda)||_2}\left[\overline{\lambda_j}\left(I_{n_k} - x_k x_k^H\right) r_k x_k^H + \lambda_j x_k r_k^H \left(I_{n_k} - x_k x_k^H\right)\right] + Q^{(k)}\widehat{B}_{kj}Q^{(k)^H}. \end{aligned}$$

Setting $\widehat{A}_k = \widehat{B}_{kj} = 0$ we obtain the desired perturbation. □

Now we consider an H-even MEP.

Theorem 6.3.8. *Let $\mathbb{S}$ be the space of all H-even MEPs and $\mathbb{W} \in \mathbb{S}$ be as given in (6.1). Let $\lambda := (\lambda_1, \cdots, \lambda_m)^T \in \mathbb{C}^m$ be such that $\lambda_j := \alpha_j + i\beta_j$ for all $j = 1, \cdots, m$, where $i = \sqrt{-1}$ and $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ be such that $x_k^H x_k = 1$. Consider $r_k := -W_k(\lambda)x_k$ for all $k = 1, \cdots, m$. Then*

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \begin{cases} \frac{1}{||(1,\lambda)||_2}\left(\sum_{k=1}^m \left(2||r_k||_2^2 - |x_k^H r_k|^2\right)\right)^{\frac{1}{2}} & \text{when } \lambda \in i\mathbb{R}^m \\ \left(\sum_{k=1}^m \left(\sum_{j=0}^m |s_{kj}|^2 + \frac{2}{||(1,\lambda)||_2^2}(||r_k||_2^2 - |x_k^H r_k|^2)\right)\right)^{\frac{1}{2}} & \text{when } \lambda \in \mathbb{C}^m \setminus i\mathbb{R}^m, \end{cases} \tag{6.32}$$

where

$$\begin{bmatrix} s_{k0} \\ s_{k1} \\ \vdots \\ s_{km} \end{bmatrix} = \begin{bmatrix} 1 & -\beta_1 & \cdots & -\beta_m \\ 0 & \alpha_1 & \cdots & \alpha_m \end{bmatrix}^{\dagger} \begin{bmatrix} \mathrm{Re}(x_k^H r_k) \\ \mathrm{Im}(x_k^H r_k) \end{bmatrix}$$

for all $k = 1, \cdots, m$.

When $\lambda \in i\mathbb{R}^m$, *define*

$$\Delta A_k := \frac{1}{||(1,\lambda)||_2^2}\left(r_k x_k^H + x_k r_k^H - (r_k^H x_k) x_k x_k^H\right), \Delta B_{kj} := \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2}\left(r_k x_k^H + x_k r_k^H - (r_k^H x_k) x_k x_k^H\right)$$

and when $\lambda \in \mathbb{C}^m \setminus i\mathbb{R}^m$, *define*

$$\begin{aligned} \Delta A_k &:= s_{k0} x_k x_k^H + \tfrac{1}{||(1,\lambda)||_2^2}\left[\left(I_{n_k} - x_k x_k^H\right) r_k x_k^H + x_k r_k^H \left(I_{n_k} - x_k x_k^H\right)\right] \\ \Delta B_{kj} &:= i s_{kj} x_k x_k^H + \tfrac{1}{||(1,\lambda)||_2^2}\left[\overline{\lambda_j}\left(I_{n_k} - x_k x_k^H\right) r_k x_k^H - \lambda_j x_k r_k^H \left(I_{n_k} - x_k x_k^H\right)\right] \end{aligned}$$

for all $k, j = 1, \cdots, m$. *Consider* $\Delta W_k(z) := \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$ *for* $k = 1, \cdots, m$ *and* $\Delta\mathbb{W}(z) := (\Delta W_1(z), \cdots, \Delta W_m(z))$. *Then* $\Delta\mathbb{W}$ *is* H*-even,* $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ *and* $|||\Delta\mathbb{W}|||_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W})$.

Proof. Note that A_k is Hermitian and B_{kj} is skew-Hermitian for $k, j = 1, \cdots, m$. Choose $Q^{(k)} \in \mathbb{C}^{n_k \times (n_k - 1)}$ such that $Q_k = \begin{bmatrix} x_k & Q^{(k)} \end{bmatrix}$ is unitary for all $k = 1, \cdots, m$ and define

$$\widetilde{\Delta A_k} := Q_k^H \; \Delta A_k \; Q_k = \begin{bmatrix} a_k & \widehat{a_k}^H \\ \widehat{a_k} & \widehat{A_k} \end{bmatrix} \text{ and } \widetilde{\Delta B_{kj}} := Q_k^H \; \Delta B_{kj} \; Q_k = \begin{bmatrix} i b_{kj} & -\widehat{b_{kj}}^H \\ \widehat{b_{kj}} & \widehat{B_{kj}} \end{bmatrix},$$

where $a_k, b_{kj} \in \mathbb{R}$, $\widehat{a_k}, \widehat{b_{kj}} \in \mathbb{C}^{n_k - 1}$ and $\widehat{A_k}, \widehat{B_{kj}} \in \mathbb{C}^{(n_k-1)\times(n_k-1)}$ be such that $\widehat{A_k}^H = \widehat{A_k}$ and $\widehat{B_{kj}}^H = -\widehat{B_{kj}}$ for all $k, j = 1, \cdots, m$. Consequently, we have $\begin{bmatrix} a_k + \sum_{j=1}^m i\lambda_j b_{kj} \\ \widehat{a_k} + \sum_{j=1}^m \lambda_j \widehat{b_{kj}} \end{bmatrix} = \begin{bmatrix} x_k^H r_k \\ {Q^{(k)}}^H r_k \end{bmatrix}$ for all $k = 1, \cdots, m$. The minimal 2-norm solution of $\widehat{a_k} + \sum_{j=1}^m \lambda_j \widehat{b_{kj}} = {Q^{(k)}}^H r_k$ is given by $\widehat{a_k} = \frac{1}{||(1,\lambda)||_2^2} {Q^{(k)}}^H r_k$ and $\widehat{b_{kj}} = \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2} {Q^{(k)}}^H r_k$.

Case I: Suppose that $\lambda \in i\mathbb{R}^m$. In this case note that $\lambda_j = i\beta_j$ for all $j = 1, \cdots, m$. Then we can rewrite $a_k + \sum_{j=1}^m i\lambda_j b_{kj} = x_k^H r_k$ as $a_k - \sum_{j=1}^m \beta_j b_{kj} = x_k^H r_k$. Now by

Lemma 6.2.1, the minimum norm solution of $a_k - \sum_{j=1}^{m} \beta_j b_{kj} = x_k^H r_k$ is given by $a_k = \frac{1}{||(1,\lambda)||_2^2} x_k^H r_k$ and $b_{kj} = \frac{i\lambda_j}{||(1,\lambda)||_2^2} x_k^H r_k$. Then

$$\widetilde{\Delta A_k} = \begin{bmatrix} \frac{1}{||(1,\lambda)||_2^2} x_k^H r_k & \frac{1}{||(1,\lambda)||_2^2} (Q^{(k)^H} r_k)^H \\ \frac{1}{||(1,\lambda)||_2^2} Q^{(k)^H} r_k & \widehat{A_k} \end{bmatrix}, \ \widetilde{\Delta B_{kj}} = \begin{bmatrix} \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2} x_k^H r_k & -\frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2} (Q^{(k)^H} r_k)^H \\ \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2} Q^{(k)^H} r_k & \widehat{B_{kj}} \end{bmatrix} \tag{6.33}$$

for all $k, j = 1, \cdots, m$. Setting $\widehat{A_k} = \widehat{B_{kj}} = 0$ and noting that $||Q^{(k)^H} r_k||_2^2 = ||r_k||_2^2 - |x_k^H r_k|^2$ we have

$$||\Delta A_k||_F = \frac{1}{||(1,\lambda)||_2^2} \left(2||r_k||_2^2 - |x_k^H r_k|^2\right)^{\frac{1}{2}}, ||\Delta B_{kj}||_F = \frac{|\lambda_j|}{||(1,\lambda)||_2^2} \left(2||r_k||_2^2 - |x_k^H r_k|^2\right)^{\frac{1}{2}}.$$

So $||\Delta W_k||_2 = \frac{1}{||(1,\lambda)||_2} \left(2||r_k||_2^2 - |x_k^H r_k|^2\right)^{\frac{1}{2}}$ and thus

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \|\Delta \mathbb{W}\|_2 = \frac{1}{||(1,\lambda)||_2} \left(\sum_{k=1}^{m} \left(2||r_k||_2^2 - |x_k^H r_k|^2\right) \right)^{\frac{1}{2}}.$$

Case II: Suppose that $\lambda \in \mathbb{C}^m \setminus i\mathbb{R}^m$. To obtain the real minimal 2-norm solution of $a_k + \sum_{j=1}^{m} i\lambda_j b_{kj} = x_k^H r_k$, we can rewrite it as

$$\begin{bmatrix} 1 & -\beta_1 & \cdots & -\beta_m \\ 0 & \alpha_1 & \cdots & \alpha_m \end{bmatrix} \begin{bmatrix} a_k \\ b_{k1} \\ \vdots \\ b_{km} \end{bmatrix} = \begin{bmatrix} \mathrm{Re}(x_k^H r_k) \\ \mathrm{Im}(x_k^H r_k) \end{bmatrix}$$

and the minimum 2-norm solution is given by

$$\begin{bmatrix} a_k \\ b_{k1} \\ \vdots \\ b_{km} \end{bmatrix} = \begin{bmatrix} 1 & -\beta_1 & \cdots & -\beta_m \\ 0 & \alpha_1 & \cdots & \alpha_m \end{bmatrix}^{\dagger} \begin{bmatrix} \mathrm{Re}(x_k^H r_k) \\ \mathrm{Im}(x_k^H r_k) \end{bmatrix}.$$

So

$$\left.\begin{aligned} \widetilde{\Delta A_k} &= \begin{bmatrix} s_{k0} & \frac{1}{||(1,\lambda)||_2^2} (Q^{(k)^H} r_k)^H \\ \frac{1}{||(1,\lambda)||_2^2} Q^{(k)^H} r_k & \widehat{A_k} \end{bmatrix} \\ \widetilde{\Delta B_{kj}} &= \begin{bmatrix} i s_{kj} & -\frac{\lambda_j}{||(1,\lambda)||_2^2} (Q^{(k)^H} r_k)^H \\ \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2} Q^{(k)^H} r_k & \widehat{B_{kj}} \end{bmatrix} \end{aligned}\right\} \tag{6.34}$$

for all $k, j = 1, \cdots, m$. Then setting $\widehat{A_k} = \widehat{B_{kj}} = 0$ and noting that $||Q^{(k)^H} r_k||_2^2 = ||r_k||_2^2 - |x_k^H r_k|^2$ we have $||\Delta A_k||_F = \left(|s_{k0}|^2 + \frac{2}{||(1,\lambda)||_2^4}(||r_k||_2^2 - |x_k^H r_k|^2)\right)^{\frac{1}{2}}$ and $||\Delta B_{kj}||_F = \left(|s_{kj}|^2 + \frac{2|\lambda_j|^2}{||(1,\lambda)||_2^4}(||r_k||_2^2 - |x_k^H r_k|^2)\right)^{\frac{1}{2}}$ for all $k, j = 1, \cdots, m$. Therefore $||\Delta W_k||_2 = \left(\sum_{j=0}^m |s_{kj}|^2 + \frac{2}{||(1,\lambda)||_2^2}(||r_k||_2^2 - |x_k^H r_k|^2)\right)^{\frac{1}{2}}$ and

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \|\Delta \mathbb{W}\|_2 = \left(\sum_{k=1}^m \left(\sum_{j=0}^m |s_{kj}|^2 + \frac{2}{||(1,\lambda)||_2^2}(||r_k||_2^2 - |x_k^H r_k|^2)\right)\right)^{\frac{1}{2}},$$

which proves (6.32).

Simplifying the expressions, when $\lambda \in i\mathbb{R}^m$, by (6.33), it follows that

$$\begin{aligned} \Delta A_k &= \tfrac{1}{||(1,\lambda)||_2^2}\left[x_k r_k^H + r_k x_k^H - (x_k^H r_k) x_k x_k^H\right] + Q^{(k)} \widehat{A_k} Q^{(k)^H} \\ \Delta B_{kj} &= \tfrac{\overline{\lambda_j}}{||(1,\lambda)||_2^2}\left[r_k x_k^H - x_k r_k^H + (x_k^H r_k) x_k x_k^H\right] + Q^{(k)} \widehat{B_{kj}} Q^{(k)^H} \end{aligned}$$

and when $\lambda \in \mathbb{C}^m \setminus i\mathbb{R}^m$, by (6.34), we have

$$\begin{aligned} \Delta A_k &= x_k s_{k0} x_k^H + \tfrac{1}{||(1,\lambda)||_2^2}\left[\left(I_{n_k} - x_k x_k^H\right) r_k x_k^H + x_k r_k^H \left(I_{n_k} - x_k x_k^H\right)\right] + Q^{(k)} \widehat{A_k} Q^{(k)^H} \\ \Delta B_{kj} &= i x_k s_{kj} x_k^H + \tfrac{1}{||(1,\lambda)||_2^2}\left[\overline{\lambda_j}\left(I_{n_k} - x_k x_k^H\right) r_k x_k^H - \lambda_j x_k r_k^H \left(I_{n_k} - x_k x_k^H\right)\right] + Q^{(k)} \widehat{B_{kj}} Q^{(k)^H} \end{aligned}$$

for all $k, j = 1, \cdots, m$. Setting $\widehat{A_k} = \widehat{B_{kj}} = 0$ we obtain the desired perturbation. □

Finally we derive structured backward error of a H-even alternating MEP in the following result.

Theorem 6.3.9. *Let $\mathbb{S}$ be the space of all H-even alternating MEPs and $\mathbb{W} \in \mathbb{S}$ be as given in (6.1). Let $\lambda := (\lambda_1, \cdots, \lambda_m)^T \in \mathbb{C}^m$ be such that $\lambda_j := \alpha_j + i\beta_j$ for all $j = 1, \cdots, m$, where $i = \sqrt{-1}$ and $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ be such that $x_k^H x_k = 1$. Consider $r_k := -W_k(\lambda) x_k$ for all $k = 1, \cdots, m$. Then*

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \begin{cases} \frac{1}{||(1,\lambda)||_2}\left(\sum_{k=1}^m \left(2||r_k||_2^2 - |x_k^H r_k|^2\right)\right)^{\frac{1}{2}} & \text{when } \lambda \in i\mathbb{R}^m \\ \left(\sum_{k=1}^m \left(\sum_{j=0}^m |s_{kj}|^2 + \frac{2}{||(1,\lambda)||_2^2}(||r_k||_2^2 - |x_k^H r_k|^2)\right)\right)^{\frac{1}{2}} & \text{when } \lambda \in \mathbb{C}^m \setminus i\mathbb{R}^m, \end{cases} \tag{6.35}$$

where

$$\begin{bmatrix} s_{k0} \\ s_{k1} \\ \vdots \\ s_{km} \end{bmatrix} = \begin{bmatrix} 1 & \Pi_{(e)}(\alpha) - \Pi_{(o)}(\beta) \\ 0 & \Pi_{(o)}(\alpha) + \Pi_{(e)}(\beta) \end{bmatrix}^{\dagger} \begin{bmatrix} \mathrm{Re}(x_k^H r_k) \\ \mathrm{Im}(x_k^H r_k) \end{bmatrix} \quad \text{for } k = 1, \cdots, m$$

and $\alpha := (\alpha_1, \cdots, \alpha_m)^T, \beta := (\beta_1, \cdots, \beta_m)^T$.

When $\lambda \in i\mathbb{R}^m$, *define*

$$\begin{aligned} \Delta A_k &= \frac{1}{||(1,\lambda)||_2^2}\left[r_k x_k^H + x_k r_k^H - (r_k^H x_k) x_k x_k^H\right] \\ \Delta B_{kj} &= \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2}\left[r_k x_k^H + x_k r_k^H - (r_k^H x_k) x_k x_k^H\right] \end{aligned}$$

and when $\lambda \in \mathbb{C}^m \setminus i\mathbb{R}^m$, *define*

$$\Delta A_k = x_k s_{k0} x_k^H + \frac{1}{||(1,\lambda)||_2}\left[\left(I_{n_k} - x_k x_k^H\right) r_k x_k^H + x_k r_k^H \left(I_{n_k} - x_k x_k^H\right)\right]$$

$$\Delta B_{kj} = \begin{cases} \frac{1}{||(1,\lambda)||_2}\left[\overline{\lambda_j}\left(I_{n_k} - x_k x_k^H\right) r_k x_k^H + \lambda_j x_k r_k^H \left(I_{n_k} - x_k x_k^H\right)\right] & \\ +x_k s_{kj} x_k^H & \text{when } j \text{ is even} \\ \frac{1}{||(1,\lambda)||_2}\left[\overline{\lambda_j}\left(I_{n_k} - x_k x_k^H\right) r_k x_k^H - \lambda_j x_k r_k^H \left(I_{n_k} - x_k x_k^H\right)\right] & \\ +i x_k s_{kj} x_k^H & \text{when } j \text{ is odd} \end{cases}$$

for all $k, j = 1, \cdots, m$. *Consider* $\Delta W_k(z) := \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$ *for* $k = 1, \cdots, m$ *and* $\Delta\mathbb{W}(z) := (\Delta W_1(z), \cdots, \Delta W_m(z))$. *Then* $\Delta\mathbb{W}$ *is* H*-even alternating,* $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ *and* $\|\Delta\mathbb{W}\|_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W})$.

Proof. Consider $Q^{(k)} \in \mathbb{C}^{n_k \times (n_k - 1)}$ such that $Q_k = \begin{bmatrix} x_k & Q^{(k)} \end{bmatrix}$ is unitary for all $k = 1, \cdots, m$. Define

$$\widetilde{\Delta A_k} := \begin{bmatrix} a_k & \widehat{a_k}^H \\ \widehat{a_k} & \widehat{A_k} \end{bmatrix}, \ \widetilde{\Delta B_{kj}} := \begin{cases} \begin{bmatrix} b_{kj} & \widehat{b_{kj}}^H \\ \widehat{b_{kj}} & \widehat{B_{kj}} \end{bmatrix} & \text{when } j \text{ is even} \\ \begin{bmatrix} i b_{kj} & -\widehat{b_{kj}}^H \\ \widehat{b_{kj}} & \widehat{B_{kj}} \end{bmatrix} & \text{when } j \text{ is odd,} \end{cases}$$

where $a_k, b_{kj} \in \mathbb{R}, \widehat{a_k}, \widehat{b_{kj}} \in \mathbb{C}^{n_k - 1}$ and $\widehat{A_k}, \widehat{B_{kj}} \in \mathbb{C}^{(n_k - 1)\times(n_k - 1)}$ be such that $\widehat{A_k}^H = \widehat{A_k}$ and $\widehat{B_{kj}}^H = (-1)^j \widehat{B_{kj}}$ for all $k, j = 1, \cdots, m$. Consequently, we have

$$\begin{bmatrix} a_k + \sum\limits_{j \ even} \lambda_j b_{kj} + \sum\limits_{j \ odd} i\lambda_j b_{kj} \\ \widehat{a_k} + \sum_{j=1}^m \lambda_j \widehat{b_{kj}} \end{bmatrix} = \begin{bmatrix} x_k^H r_k \\ Q^{(k)^H} r_k \end{bmatrix}.$$

The minimum norm solution of $\widehat{a_k} + \sum_{j=1}^m \lambda_j \widehat{b_{kj}} = Q^{(k)^H} r_k$ is given by $\widehat{a_k} = \frac{1}{||(1,\lambda)||_2^2} Q^{(k)^H} r_k$ and $\widehat{b_{kj}} = \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2} Q^{(k)^H} r_k$ for $k, j = 1, \cdots, m$.

Case I: Suppose that $\lambda \in i\mathbb{R}^m$. In this case, the minimal 2-norm solution of $a_k + \sum\limits_{j\ even} \lambda_j b_{kj} + \sum\limits_{j\ odd} i\lambda_j b_{kj} = x_k^H r_k$ is given by $a_k = \frac{1}{||(1,\lambda)||_2^2} x_k^H r_k$, $b_{kj} = \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2} x_k^H r_k$ when j is even and $b_{kj} = -\frac{i\overline{\lambda_j}}{||(1,\lambda)||_2^2} x_k^H r_k$ when j is odd. Then

$$\widetilde{\Delta A_k} = \begin{bmatrix} \frac{1}{||(1,\lambda)||_2^2} x_k^H r_k & \frac{1}{||(1,\lambda)||_2^2} (Q^{(k)^H} r_k)^H \\ \frac{1}{||(1,\lambda)||_2^2} Q^{(k)^H} r_k & \widehat{A_k} \end{bmatrix} \tag{6.36}$$

$$\widetilde{\Delta B_{kj}} = \begin{cases} \begin{bmatrix} \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2} x_k^H r_k & \frac{\lambda_j}{||(1,\lambda)||_2^2} (Q^{(k)^H} r_k)^H \\ \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2} Q^{(k)^H} r_k & \widehat{B_{kj}} \end{bmatrix} & \text{when } j \text{ is even} \\ \begin{bmatrix} \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2} x_k^H r_k & -\frac{\lambda_j}{||(1,\lambda)||_2^2} (Q^{(k)^H} r_k)^H \\ \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2} Q^{(k)^H} r_k & \widehat{B_{kj}} \end{bmatrix} & \text{when } j \text{ is odd} \end{cases} \tag{6.37}$$

for all $k, j = 1, \cdots, m$. Setting $\widehat{A_k} = \widehat{B_{kj}} = 0$ and noting that $||Q^{(k)^H} r_k||_2^2 = ||r_k||_2^2 - |x_k^H r_k|^2$ we have

$$||\Delta A_k||_F = \frac{1}{||(1,\lambda)||_2^2} \left(2||r_k||_2^2 - |x_k^H r_k|^2\right)^{\frac{1}{2}} \text{ and } ||\Delta B_{kj}||_F = \frac{|\lambda_j|}{||(1,\lambda)||_2^2} \left(2||r_k||_2^2 - |x_k^H r_k|^2\right)^{\frac{1}{2}}.$$

Therefore $||\Delta W_k||_2 = \frac{1}{||(1,\lambda)||_2} \left(2||r_k||_2^2 - |x_k^H r_k|^2\right)^{\frac{1}{2}}$ and thus

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \|\Delta \mathbb{W}\|_2 = \frac{1}{||(1,\lambda)||_2} \left(\sum_{k=1}^{m} \left(2||r_k||_2^2 - |x_k^H r_k|^2\right) \right)^{\frac{1}{2}}.$$

Case II: Suppose that $\lambda \in \mathbb{C}^m \backslash i\mathbb{R}^m$. In this case, $a_k + \sum\limits_{j\ even} \lambda_j b_{kj} + \sum\limits_{j\ odd} i\lambda_j b_{kj} = x_k^H r_k$ can be written as

$$\begin{bmatrix} 1 & \Pi_{(e)}(\alpha) - \Pi_{(o)}(\beta) \\ 0 & \Pi_{(o)}(\alpha) + \Pi_{(e)}(\beta) \end{bmatrix} \begin{bmatrix} a_k \\ b_{k1} \\ \vdots \\ b_{km} \end{bmatrix} = \begin{bmatrix} \mathrm{Re}(x_k^H r_k) \\ \mathrm{Im}(x_k^H r_k) \end{bmatrix}$$

and thus the required minimum norm solution is given by

$$\begin{bmatrix} a_k \\ b_{k1} \\ \vdots \\ b_{km} \end{bmatrix} = \begin{bmatrix} 1 & \Pi_{(e)}(\alpha) - \Pi_{(o)}(\beta) \\ 0 & \Pi_{(o)}(\alpha) + \Pi_{(e)}(\beta) \end{bmatrix}^{\dagger} \begin{bmatrix} \mathrm{Re}(x_k^H r_k) \\ \mathrm{Im}(x_k^H r_k) \end{bmatrix}.$$

Therefore we have

$$\widetilde{\Delta A_k} = \begin{bmatrix} s_{k0} & \frac{1}{||(1,\lambda)||_2^2}({Q^{(k)}}^H r_k)^H \\ \frac{1}{||(1,\lambda)||_2^2}{Q^{(k)}}^H r_k & \widehat{A_k} \end{bmatrix}, \tag{6.38}$$

$$\widetilde{\Delta B_{kj}} = \begin{cases} \begin{bmatrix} s_{kj} & \frac{\lambda_j}{||(1,\lambda)||_2^2}({Q^{(k)}}^H r_k)^H \\ \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2}{Q^{(k)}}^H r_k & \widehat{B_{kj}} \end{bmatrix} & \text{when } j \text{ is even} \\ \begin{bmatrix} is_{kj} & -\frac{\lambda_j}{||(1,\lambda)||_2^2}({Q^{(k)}}^H r_k)^H \\ \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2}{Q^{(k)}}^H r_k & \widehat{B_{kj}} \end{bmatrix} & \text{when } j \text{ is odd} \end{cases} \tag{6.39}$$

for all $k, j = 1, \cdots, m$. Again setting $\widehat{A_k} = \widehat{B_{kj}} = 0$, it follows that $||\Delta A_k||_F = \left(|s_{k0}|^2 + \frac{2}{||(1,\lambda)||_2^4}(||r_k||_2^2 - |x_k^H r_k|^2)\right)^{\frac{1}{2}}$ and $||\Delta B_{kj}||_F = \left(|s_{kj}|^2 + \frac{2|\lambda_j|^2}{||(1,\lambda)||_2^4}(||r_k||_2^2 - |x_k^H r_k|^2)\right)^{\frac{1}{2}}$. So $||\Delta W_k||_2 = \left(\sum_{j=0}^m |s_{kj}|^2 + \frac{2}{||(1,\lambda)||_2^2}(||r_k||_2^2 - |x_k^H r_k|^2)\right)^{\frac{1}{2}}$ and hence

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \|\Delta\mathbb{W}\|_2 = \left(\sum_{k=1}^m \left(\sum_{j=0}^m |s_{kj}|^2 + \frac{2}{||(1,\lambda)||_2^2}(||r_k||_2^2 - |x_k^H r_k|^2)\right)\right)^{\frac{1}{2}}.$$

Therefore, we obtain (6.35).

When $\lambda \in i\mathbb{R}^m$, by further simplifications, it follows from (6.36) and (6.37) that

$$\begin{aligned} \Delta A_k &= \frac{1}{||(1,\lambda)||_2^2}\left[r_k x_k^H + x_k r_k^H - (r_k^H x_k) x_k x_k^H\right] + Q^{(k)}\widehat{A_k}{Q^{(k)}}^H \\ \Delta B_{kj} &= \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2}\left[r_k x_k^H + x_k r_k^H - (r_k^H x_k) x_k x_k^H\right] + Q^{(k)}\widehat{A_k}{Q^{(k)}}^H \end{aligned}$$

and when $\lambda \in \mathbb{C}^m \setminus i\mathbb{R}^m$, it follows from (6.38) and (6.39) that

$$\Delta A_k = x_k s_{k0} x_k^H + \frac{1}{||(1,\lambda)||_2}\left[\left(I_{n_k} - x_k x_k^H\right) r_k x_k^H + x_k r_k^H \left(I_{n_k} - x_k x_k^H\right)\right] + Q^{(k)}\widehat{A_k}{Q^{(k)}}^H,$$

$$\Delta B_{kj} = \begin{cases} \frac{1}{||(1,\lambda)||_2}\left[\overline{\lambda_j}\left(I_{n_k} - x_k x_k^H\right) r_k x_k^H + \lambda_j x_k r_k^H \left(I_{n_k} - x_k x_k^H\right)\right] & \\ + x_k s_{kj} x_k^H + Q^{(k)}\widehat{B_{kj}}{Q^{(k)}}^H & \text{when } j \text{ is even} \\ \frac{1}{||(1,\lambda)||_2}\left[\overline{\lambda_j}\left(I_{n_k} - x_k x_k^H\right) r_k x_k^H - \lambda_j x_k r_k^H \left(I_{n_k} - x_k x_k^H\right)\right] & \\ + i x_k s_{kj} x_k^H + Q^{(k)}\widehat{B_{kj}}{Q^{(k)}}^H & \text{when } j \text{ is odd.} \end{cases}$$

Hence setting $\widehat{A_k} = \widehat{B_{kj}} = 0$ we obtain the desired perturbation. □

6.4 Spectral norm structured backward error

In this section we derive structured backward errors by considering the Spectral norm on $\mathbb{C}^{n_k\times n_k}$ for all $k=1,\cdots,m$. The following result will play an important role in deriving the structured backward errors.

Theorem 6.4.1. (Davis-Kahan-Weinberger, [25]) *Let A,B,C be given matrices such that*

$$\left|\left|\begin{bmatrix} A \\ B \end{bmatrix}\right|\right|_2 = \mu = \left|\left|\begin{bmatrix} A & C \end{bmatrix}\right|\right|_2 .$$

Then there exists D such that

$$\left|\left|\begin{bmatrix} A & C \\ B & D \end{bmatrix}\right|\right|_2 = \mu.$$

Indeed, those D which have this property are exactly those of the form

$$D = -KA^HL + \mu(I-KK^H)^{\frac{1}{2}}Z(I-L^HL)^{\frac{1}{2}},$$

where $K^H := (\mu^2 I - A^HA)^{-\frac{1}{2}}B^H, L := (\mu^2 I - AA^H)^{-\frac{1}{2}}C$ and Z is an arbitrary contraction, that is, $||Z||_2 \leq 1$.

First we consider T-symmetric MEPs.

Theorem 6.4.2. *Let $\mathbb{S}$ be the space of all T-symmetric MEPs and $\mathbb{W}\in\mathbb{S}$ be as given in (6.1). Let $\lambda=(\lambda_1,\cdots,\lambda_m)^T\in\mathbb{C}^m$ and $0\neq x := x_1\otimes\cdots\otimes x_m \in \mathbb{C}^{n_1}\otimes\cdots\otimes\mathbb{C}^{n_m}$ be such that $x_k^Hx_k=1$. Consider $r_k := -W_k(\lambda)x_k$ for all $k=1,\cdots,m$. Then $\eta_p^{\mathbb{S}}(\lambda,x,\mathbb{W}) = \frac{1}{||(1,\lambda)||_q}\,||(||r_1||_2,\cdots,||r_m||_2)||_p$, where $p^{-1}+q^{-1}=1$.*

Let $(y_0,\cdots,y_m)^T\in\partial||(1,\lambda)||_q$. When $||r_k||_2\neq|x_k^Tr_k|$ for some $k\in\{1,\cdots,m\}$ define

$$\begin{aligned}
\Delta A_k &:= \frac{\overline{y_0}}{||(1,\lambda)||_q}\left[\overline{x_k}r_k^T + r_kx_k^H - (r_k^Tx_k)\overline{x_k}x_k^H - \frac{\overline{x_k^Tr_k}(I_{n_k}-\overline{x_k}x_k^T)r_kr_k^T(I_{n_k}-x_kx_k^H)}{||r_k||_2^2-|x_k^Tr_k|^2}\right],\\
\Delta B_{kj} &:= \frac{\overline{y_j}}{||(1,\lambda)||_q}\left[\overline{x_k}r_k^T + r_kx_k^H - (r_k^Tx_k)\overline{x_k}x_k^H - \frac{\overline{x_k^Tr_k}(I_{n_k}-\overline{x_k}x_k^T)r_kr_k^T(I_{n_k}-x_kx_k^H)}{||r_k||_2^2-|x_k^Tr_k|^2}\right]
\end{aligned}$$

and when $||r_k||_2 = |x_k^Tr_k|$ for some $k\in\{1,\cdots,m\}$, define $\Delta A_k := \frac{\overline{y_0}}{||(1,\lambda)||_q}\overline{x_k}r_k^T = \frac{\overline{y_0}}{||(1,\lambda)||_q}r_kx_k^H$, $\Delta B_{kj} := \frac{\overline{y_j}}{||(1,\lambda)||_q}\overline{x_k}r_k^T = \frac{\overline{y_j}}{||(1,\lambda)||_q}r_kx_k^H$ for all $j=1,\cdots,m$. Consider

$\Delta W_k(z) := \Delta A_k + \sum_{j=1}^{m} z_j \Delta B_{kj}$ *for* $k = 1, \cdots, m$ *and* $\Delta \mathbb{W}(z) := (\Delta W_1(z), \cdots, \Delta W_m(z))$. *Then* $\Delta \mathbb{W}$ *is* T*-symmetric,* $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ *and* $\|\Delta \mathbb{W}\|_p = \eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W})$.

Proof. Recall from the proof of Theorem 6.3.1 that

$$\begin{aligned} \Delta A_k &= \overline{Q_k} \begin{bmatrix} \frac{\overline{y_0}}{||(1,\lambda)||_q} x_k^T r_k & \frac{\overline{y_0}}{||(1,\lambda)||_q} (Q^{(k)^T} r_k)^T \\ \frac{\overline{y_0}}{||(1,\lambda)||_q} Q^{(k)^T} r_k & \widehat{A_k} \end{bmatrix} Q_k^H, \\ \Delta B_{kj} &= \overline{Q_k} \begin{bmatrix} \frac{\overline{y_j}}{||(1,\lambda)||_q} x_k^T r_k & \frac{\overline{y_j}}{||(1,\lambda)||_q} (Q^{(k)^T} r_k)^T \\ \frac{\overline{y_j}}{||(1,\lambda)||_q} & \widehat{B_{kj}} \end{bmatrix} Q_k^H. \end{aligned}$$

satisfies that $W_k(\lambda)x_k + \Delta W_k(\lambda)x_k = 0$. When $||r_k||_2 \neq |x_k^T r_k|$, for $\mu_{\Delta A_k} := \frac{|y_0|}{||(1,\lambda)||_q} ||r_k||_2$ and $\mu_{\Delta B_{kj}} := \frac{|y_j|}{||(1,\lambda)||_q} ||r_k||_2$, by Theorem 6.4.1, we have

$$\widehat{A_k} = -\frac{\overline{y_0}\ \overline{x_k^T r_k} (Q^{(k)^T} r_k)(Q^{(k)^T} r_k)^T}{||(1,\lambda)||_q (||r_k||_2^2 - |x_k^T r_k|^2)} \text{ and } \widehat{B_{kj}} = -\frac{\overline{y_j}\ \overline{x_k^T r_k} (Q^{(k)^T} r_k)(Q^{(k)^T} r_k)^T}{||(1,\lambda)||_q (||r_k||_2^2 - |x_k^T r_k|^2)}. \tag{6.40}$$

When $||r_k||_2 = |x_k^T r_k|$ for some $k \in \{1, \cdots, m\}$ then note that $||Q^{(k)^T} r_k||_2 = (||r_k||_2^2 - |x_k^T r_k|^2)^{\frac{1}{2}} = 0$, i.e., $Q^{(k)^T} r_k = 0$. So we have

$$\Delta A_k = \overline{Q_k} \begin{bmatrix} \frac{\overline{y_0}}{||(1,\lambda)||_q} x_k^T r_k & 0 \\ 0 & \widehat{A_k} \end{bmatrix} Q_k^H, \ \Delta B_{kj} = \overline{Q_k} \begin{bmatrix} \frac{\overline{y_j}}{||(1,\lambda)||_q} x_k^T r_k & 0 \\ 0 & \widehat{B_{kj}} \end{bmatrix} Q_k^H.$$

For $\mu_{\Delta A_k} := \frac{|y_0|}{||(1,\lambda)||_q} |x_k^T r_k| = \frac{|y_0|}{||(1,\lambda)||_q} ||r_k||_2$ and $\mu_{\Delta B_{kj}} := \frac{|y_j|}{||(1,\lambda)||_q} ||r_k||_2$, by Theorem 6.4.1, we have $\widehat{A_k} = \widehat{B_{kj}} = 0$. So $||\Delta W_k||_p = \frac{||y||_p}{||(1,\lambda)||_q} ||r_k||_2 = \frac{1}{||(1,\lambda)||_q} ||r_k||_2$. Therefore $\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \|\Delta \mathbb{W}\|_p = \frac{1}{||(1,\lambda)||_q} \ ||(||r_1||_2, \cdots, ||r_m||_2)||_p$.

When $||r_k||_2 \neq |x_k^T r_k|$, substituting $\widehat{A_k}$ and $\widehat{B_{kj}}$ as given in (6.40) in (6.14) and (6.15), respectively, and when $||r_k||_2 = |x_k^T r_k|$, setting $\widehat{A_k} = \widehat{B_{kj}} = 0$, in (6.14) and (6.15) we obtain the desired perturbation. ☐

Next we consider T-skew symmetric MEPs.

Theorem 6.4.3. *Let* $\mathbb{S}$ *be the space of all* T*-skew symmetric MEPs and* $\mathbb{W} \in \mathbb{S}$ *be as given in (6.1). Let* $\lambda := (\lambda_1, \cdots, \lambda_m)^T \in \mathbb{C}^m$ *and* $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ *be such that* $x_k^H x_k = 1$. *Consider* $r_k := -W_k(\lambda)x_k$ *for all* $k = 1, \cdots, m$. *Then* $\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \frac{1}{||(1,\lambda)||_q} \ ||(||r_1||_2, \cdots, ||r_m||_2)||_p$*, where* $p^{-1} + q^{-1} = 1$.

Let $(y_0,\cdots,y_m)^T \in \partial||(1,\lambda)||_q$. *Define* $\Delta A_k = \frac{\overline{y_0}}{||(1,\lambda)||_q}\left(r_k x_k^H - \overline{x_k} r_k^T\right)$ *and* $\Delta B_{kj} = \frac{\overline{y_j}}{||(1,\lambda)||_q}\left(r_k x_k^H - \overline{x_k} r_k^T\right)$ *for all* $k,j=1,\cdots,m$. *Consider* $\Delta W_k(z) := \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$ *for* $k=1,\cdots,m$ *and* $\Delta\mathbb{W}(z) := (\Delta W_1(z),\cdots,\Delta W_m(z))$. *Then* $\Delta\mathbb{W}$ *is* T*-skew symmetric,* $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ *and* $\|\Delta\mathbb{W}\|_p = \eta_p^{\mathbb{S}}(\lambda,x,\mathbb{W})$.

Proof. Recall from the proof of Theorem 6.3.2 that in this case we have

$$\begin{aligned}
\Delta A_k &= \overline{Q_k}\begin{bmatrix} 0 & -\frac{\overline{y_0}}{||(1,\lambda)||_q}\left(Q^{(k)^T} r_k\right)^T \\ \frac{\overline{y_0}}{||(1,\lambda)||_q} Q^{(k)^T} r_k & \widehat{A_k} \end{bmatrix} Q_k^H, \\
\Delta B_{kj} &= \overline{Q_k}\begin{bmatrix} 0 & -\frac{\overline{y_j}}{||(1,\lambda)||_q}\left(Q^{(k)^T} r_k\right)^T \\ \frac{\overline{y_j}}{||(1,\lambda)||_q} Q^{(k)^T} r_k & \widehat{B_{kj}} \end{bmatrix} Q_k^H.
\end{aligned}$$

such that $W_k(\lambda)x_k + \Delta W_k(\lambda)x_k = 0$. Now for $\mu_{\Delta A_k} := \frac{|y_0|}{||(1,\lambda)||_q}||r_k||_2$ and $\mu_{\Delta B_{kj}} := \frac{|y_j|}{||(1,\lambda)||_q}||r_k||_2$ by Theorem 6.4.1, we have $\widehat{A_k} = \widehat{B_{kj}} = 0$. Then $\eta_p^{\mathbb{S}}(\lambda, x_k, W_k) = \frac{||y||_p}{||(1,\lambda)||_q}||r_k||_2 = \frac{1}{||(1,\lambda)||_q}||r_k||_2$ and thus we have $\eta_p^{\mathbb{S}}(\lambda,x,\mathbb{W}) = \frac{1}{||(1,\lambda)||_q}\ ||(||r_1||_2,\cdots,||r_m||_2)||_p$. Further simplifying ΔA_k and ΔB_{kj}, we have

$$\Delta A_k = \frac{\overline{y_0}}{||(1,\lambda)||_q}\left(r_k x_k^H - \overline{x_k} r_k^T\right) + \overline{Q^{(k)}}\widehat{A_k}Q^{(k)^H}, \Delta B_{kj} = \frac{\overline{y_j}}{||(1,\lambda)||_q}\left(r_k x_k^H - \overline{x_k} r_k^T\right) + \overline{Q^{(k)}}\widehat{B_{kj}}Q^{(k)^H}.$$

Setting $\widehat{A_k} = \widehat{B_{kj}} = 0$, we obtain the desired optimal perturbation. ☐

Now we derive $\eta_p^{\mathbb{S}}(\lambda,x,\mathbb{W})$ for a T-even MEP.

Theorem 6.4.4. *Let* $\mathbb{S}$ *be the space of all* T*-even MEPs and* $\mathbb{W} \in \mathbb{S}$ *be as given in (6.1). Let* $\lambda := (\lambda_1,\cdots,\lambda_m)^T \in \mathbb{C}^m$ *and* $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ *be such that* $x_k^H x_k = 1$. *Consider* $r_k := -W_k(\lambda)x_k$ *for all* $k = 1,\cdots,m$. *Let* $(y_0,\cdots,y_m)^T \in \partial||(1,\lambda)||_q$, *for* $p^{-1} + q^{-1} = 1$. *Then* $\eta_p^{\mathbb{S}}(\lambda,x,\mathbb{W}) = ||(c_1,\cdots,c_m)||_p$, *where* $c_k = ||(a_k, b_{k1},\cdots,b_{km})||_p$, $a_k = \left(|x_k^T A_k x_k|^2 + \frac{|y_0|^2}{||(1,\lambda)||_q^2}\left[||r_k||_2^2 - |x_k^T A_k x_k|^2\right]\right)^{\frac{1}{2}}$ *and* $b_{kj} = \frac{|y_j|}{||(1,\lambda)||_q}\left(||r_k||_2^2 - |x_k^T A_k x_k|^2\right)^{\frac{1}{2}}$ *for all* $k,j = 1,\cdots,m$. *In particular, when* $p = 2$, *we have* $\eta_2^{\mathbb{S}}(\lambda,x,\mathbb{W}) = \left(\sum_{k=1}^m \left(|x_k^T A_k x_k|^2 + \frac{1}{||(1,\lambda)||_2^2}\left(||r_k||_2^2 - |x_k^T A_k x_k|^2\right)\right)\right)^{\frac{1}{2}}$.

When $||r_k||_2 \neq |x_k^T A_k x_k|$, *define*

$$\begin{aligned} \Delta A_k &:= -\overline{x_k} x_k^T A_k x_k x_k^H + \frac{\overline{y_0}}{||(1,\lambda)||_q}\left[r_k x_k^H + \overline{x_k} r_k^T - 2(x_k^T A_k x_k)\overline{x_k} x_k^H\right] \\ &\quad - \frac{\overline{y_0}^2\ \overline{x_k^T A_k x_k}(I_{n_k} - \overline{x_k} x_k^T) r_k r_k^T (I_{n_k} - x_k x_k^H)}{|y_0|^2\left(||r_k||_2^2 - |x_k^T A_k x_k|^2\right)}, \\ \Delta B_{kj} &:= \frac{\overline{y_j}}{||(1,\lambda)||_q}\left(r_k x_k^H - \overline{x_k} r_k^T\right) \end{aligned}$$

and when $||r_k||_2 = |x_k^T A_k x_k|$, *define*

$$\begin{aligned} \Delta A_k &:= -\overline{x_k} x_k^T A_k x_k x_k^H + \frac{\overline{y_0}}{||(1,\lambda)||_q}\left[r_k x_k^H + \overline{x_k} r_k^T - 2(x_k^T A_k x_k)\overline{x_k} x_k^H\right], \\ \Delta B_{kj} &:= \frac{\overline{y_j}}{||(1,\lambda)||_q}\left(r_k x_k^H - \overline{x_k} r_k^T\right) \end{aligned}$$

for all $k, j = 1, \cdots, m$. *Consider* $\Delta W_k(z) := \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$ *for* $k = 1, \cdots, m$ *and* $\Delta\mathbb{W}(z) := (\Delta W_1(z), \cdots, \Delta W_m(z))$. *Then* $\Delta\mathbb{W}$ *is* T*-even,* $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ *and* $\|\Delta\mathbb{W}\|_p = \eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W})$.

Proof. It follows from the proof of Theorem 6.3.3 that we have

$$\begin{aligned} \Delta A_k &= \overline{Q_k}\begin{bmatrix} -x_k^T A_k x_k & \frac{\overline{y_0}}{||(1,\lambda)||_q}(Q^{(k)^T} r_k)^T \\ \frac{\overline{y_0}}{||(1,\lambda)||_q} Q^{(k)^T} r_k & \widehat{A_k} \end{bmatrix} Q_k^H, \\ \Delta B_{kj} &= \overline{Q_k}\begin{bmatrix} 0 & -\frac{\overline{y_j}}{||(1,\lambda)||_q}(Q^{(k)^T} r_k)^T \\ \frac{\overline{y_j}}{||(1,\lambda)||_q} Q^{(k)^T} r_k & \widehat{B_{kj}} \end{bmatrix} Q_k^H \end{aligned}$$

such that $W_k(\lambda)x_k + \Delta W_k(\lambda)x_k = 0$. When $||r_k||_2 \neq |x_k^T A_k x_k|$ for some $k \in \{1, \cdots, m\}$ for $\mu_{\Delta A_k} = a_k$ and $\mu_{\Delta B_{kj}} = b_{kj}$, by Theorem 6.4.1, we have

$$\widehat{A_k} = -\frac{\overline{y_0}^2\ \overline{x_k^T A_k x_k}(Q^{(k)^T} r_k)(Q^{(k)^T} r_k)^T}{|y_0|^2\left(||r_k||_2^2 - |x_k^T A_k x_k|^2\right)^{\frac{1}{2}}} \text{ and } \widehat{B_{kj}} = 0. \tag{6.41}$$

When $||r_k||_2 = |x_k^T A_k x_k|$ for some $k \in \{1, \cdots, m\}$, note that $Q^{(k)^T} r_k = 0$ and thus $\Delta A_k = \overline{Q_k}\begin{bmatrix} -x_k^T A_k x_k & 0 \\ 0 & \widehat{A_k} \end{bmatrix} Q_k^H$ and $\Delta B_{kj} = \overline{Q_k}\begin{bmatrix} 0 & 0 \\ 0 & \widehat{B_{kj}} \end{bmatrix} Q_k^H$. Again for $\mu_{\Delta A_k} = a_k$ and $\mu_{\Delta B_{kj}} = b_{kj}$, by Theorem 6.4.1, we have $\widehat{A_k} = \widehat{B_{kj}} = 0$. Then $||\Delta W_k||_2 =$

$||(a_k, b_{k1}, \cdots, b_{km})||_p = c_k$ and therefore $\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \|\Delta\mathbb{W}\|_p = ||(c_1, \cdots, c_m)||_p$. Further simplifying ΔA_k and ΔB_{kj}, we have

$$\begin{aligned} \Delta A_k &= -\overline{x_k} x_k^T A_k x_k x_k^H + \frac{\overline{y_0}}{||(1,\lambda)||_q} \left[r_k x_k^H + \overline{x_k} r_k^T - 2(x_k^T A_k x_k)\overline{x_k} x_k^H \right] + \overline{Q^{(k)}}\widehat{A_k} {Q^{(k)}}^H \\ \Delta B_{kj} &= \frac{\overline{y_j}}{||(1,\lambda)||_q} \left(r_k x_k^H - \overline{x_k} r_k^T \right) + \overline{Q^{(k)}}\widehat{B_{kj}} {Q^{(k)}}^H. \end{aligned}$$

When $||r_k||_2 \neq |x_k^T A_k x_k|$ for some $k \in \{1, \cdots, m\}$, setting $\widehat{A_k}$ and $\widehat{B_{kj}}$ as given in (6.41) and when $||r_k||_2 = |x_k^T A_k x_k|$ for some $k \in \{1, \cdots, m\}$, setting $\widehat{A_k} = \widehat{B_{kj}} = 0$, the desired perturbation follows. □

In case of a T-odd MEP, we have the following result.

Theorem 6.4.5. *Let* $\mathbb{S}$ *be the space of all* T*-odd MEPs and* $\mathbb{W} \in \mathbb{S}$ *be as given in (6.1). Let* $\lambda := (\lambda_1, \cdots, \lambda_m)^T \in \mathbb{C}^m$ *and* $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ *be such that* $x_k^H x_k = 1$. *Consider* $r_k := -W_k(\lambda)x_k$ *for all* $k = 1, \cdots, m$. *Let* $(y_0, \cdots, y_m)^T \in \partial||(1,\lambda)||_q$ *and* $(z_1, \cdots, z_m)^T \in \partial||\lambda||_q$, *for* $p^{-1} + q^{-1} = 1$. *Then*

$$\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \begin{cases} ||(c_1, \cdots, c_m)||_p & \text{when } \lambda \neq 0 \\ ||(||r_1||_2, \cdots, ||r_m||_2)||_p & \text{when } \lambda = 0, \end{cases}$$

where $c_k = ||(a_k, b_{k1}, \cdots, b_{km})||_p$ *and*

$$a_k = \frac{|y_0|}{||(1,\lambda)||_q}(||r_k||_2^2 - |x_k^T r_k|^2)^{\frac{1}{2}} \text{ and } b_{kj} = \left(\frac{|z_j|^2}{||\lambda||_q^2}|x_k^T r_k|^2 + \frac{|y_j|^2}{||(1,\lambda)||_q^2}(||r_k||_2^2 - |x_k^T r_k|^2) \right)^{\frac{1}{2}}$$

for all $k, j = 1, \cdots, m$. *In particular, when* $p = 2$, *we have*

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \begin{cases} \left(\sum_{k=1}^m \left(\frac{1}{||\lambda||_2^2}|x_k^T r_k|^2 + \frac{1}{||(1,\lambda)||_2^2}\left(||r_k||_2^2 - |x_k^T A_k x_k|^2 \right) \right) \right)^{\frac{1}{2}} & \text{when } \lambda \neq 0 \\ ||(||r_1||_2, \cdots, ||r_m||_2)||_2 & \text{when } \lambda = 0. \end{cases}$$

When $||r_k||_2 \neq |x_k^T r_k|$ *for some* $k \in \{1, \cdots, m\}$, *define*

$$\begin{aligned} \Delta A_k &:= \frac{\overline{y_0}}{||(1,\lambda)||_q}(r_k x_k^H - \overline{x_k} r_k^T) \\ \Delta B_{kj} &:= \frac{\overline{z_j}}{||\lambda||_q}(x_k^T r_k)\overline{x_k} x_k^H + \frac{\overline{y_j}}{||(1,\lambda)||_q}(r_k x_k^H + \overline{x_k} r_k^T - 2(x_k^T r_k)\overline{x_k} x_k^H) \\ &\quad - \frac{\overline{y_j}^2 \; z_j \; \overline{x_k^T r_k}(I_{n_k} - \overline{x_k} x_k^T) r_k r_k^T (I_{n_k} - x_k x_k^H)}{||\lambda||_q \; |y_j|^2 \; (||r_k||_2^2 - |x_k^T r_k|^2)} \end{aligned}$$

and when $||r_k||_2 = |x_k^T r_k|$ *for some* $k \in \{1, \cdots, m\}$*, define*

$$\begin{aligned} \Delta A_k &:= \frac{\overline{y_0}}{||(1,\lambda)||_q}(r_k x_k^H - \overline{x_k} r_k^T), \\ \Delta B_{kj} &:= \frac{\overline{z_j}}{||\lambda||_q}(x_k^T r_k)\overline{x_k} x_k^H + \frac{\overline{y_j}}{||(1,\lambda)||_q}(r_k x_k^H + \overline{x_k} r_k^T - 2(x_k^T r_k)\overline{x_k} x_k^H) \end{aligned}$$

for all $k, j = 1, \cdots, m$*. Consider* $\Delta W_k(z) := \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$ *for* $k = 1, \cdots, m$ *and* $\Delta \mathbb{W}(z) := (\Delta W_1(z), \cdots, \Delta W_m(z))$*. Then* $\Delta \mathbb{W}$ *is* T*-odd,* $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ *and* $\|\Delta \mathbb{W}\|_p = \eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W})$*.*

Now we consider T-even alternating MEPs.

Theorem 6.4.6. *Let* $\mathbb{S}$ *be the space of all* T*-even alternating MEPs and* $\mathbb{W} \in \mathbb{S}$ *be as given in (6.1). Let* $\lambda := (\lambda_1, \cdots, \lambda_m)^T \in \mathbb{C}^m$ *and* $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ *be such that* $x_k^H x_k = 1$*. Consider* $r_k := -W_k(\lambda)x_k$ *for all* $k = 1, \cdots, m$*. Let* $(y_0, \cdots, y_m)^T \in \partial||(1,\lambda)||_q$ *and* $(z_0, \cdots, z_m)^T \in \partial||(1, \Pi_{(e)}(\lambda)||_q$ *for* $p^{-1} + q^{-1} = 1$*. Then* $\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) = ||(c_1, \cdots, c_m)||_p$*, where* $c_k = ||(a_k, b_{k1}, \cdots, b_{km})||_p$ *and*

$$\begin{aligned} a_k &= \left(\frac{|z_0|^2}{||(1,\Lambda^{(e)}(\lambda))||_q^2}|x_k^T r_k|^2 + \frac{|y_0|^2}{||(1,\lambda)||_q^2}(||r_k||_2^2 - |x_k^T r_k|^2)\right)^{\frac{1}{2}} \\ b_{kj} &= \begin{cases} \left(\frac{|z_j|^2}{||(1,\Lambda^{(e)}(\lambda))||_q^2}|x_k^T r_k|^2 + \frac{|y_j|^2}{||(1,\lambda)||_q^2}(||r_k||_2^2 - |x_k^T r_k|^2)\right)^{\frac{1}{2}} & \text{when } j \text{ is even} \\ \frac{|y_j|}{||(1,\lambda)||_q}(||r_k||_2^2 - |x_k^T r_k|^2)^{\frac{1}{2}} & \text{when } j \text{ is odd} \end{cases} \end{aligned}$$

for all $k, j = 1, \cdots, m$*. In particular, for* $p = 2$*, we have*

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \left(\sum_{k=1}^m \left(\frac{1}{||(1,\Lambda^{(e)}(\lambda))||_2^2}|x_k^T r_k|^2 + \frac{1}{||(1,\lambda)||_2^2}\left(||r_k||_2^2 - |x_k^T r_k|^2\right)\right)\right)^{\frac{1}{2}}.$$

When $||r_k||_2 \neq |x_k^T r_k|$ *for some* $k \in \{1, \cdots, m\}$*, define*

$$\begin{aligned} \Delta A_k \;:=\; & \frac{\overline{z_0}}{||(1,\Lambda^{(e)}(\lambda))||_q}(x_k^T r_k)\overline{x_k} x_k^H + \frac{\overline{y_0}}{||(1,\lambda)||_q}[r_k x_k^H + \overline{x_k} r_k^T - 2(x_k^T r_k)\overline{x_k} x_k^H] \\ & - \frac{\overline{y_0}^2 z_0 \overline{x_k^T r_k}(I_{n_k} - \overline{x_k} x_k^T) r_k r_k^T (I_{n_k} - x_k x_k^H)}{|y_0|^2(||r_k||_2^2 - |x_k^T r_k|^2)||(1,\Lambda^{(e)}(\lambda))||_q}, \end{aligned}$$

$$\Delta B_{kj} := \begin{cases} \frac{\overline{z_j}}{||(1,\Lambda^{(e)}(\lambda))||_q}(x_k^T r_k)\overline{x_k} x_k^H + \frac{\overline{y_j}}{||(1,\lambda)||_q}[r_k x_k^H + \overline{x_k} r_k^T - 2(x_k^T r_k)\overline{x_k} x_k^H] & \\ -\frac{\overline{y_j}^2 z_j \overline{x_k^T r_k}}{|y_j|^2(||r_k||_2^2 - |x_k^T r_k|^2)||(1,\Lambda^{(e)}(\lambda))||_q}(I_{n_k} - \overline{x_k} x_k^T) r_k r_k^T (I_{n_k} - x_k x_k^H) & \text{when } j \text{ is even} \\ \frac{\overline{y_j}}{||(1,\lambda)||_q}[r_k x_k^H - \overline{x_k} r_k^T] & \text{when } j \text{ is odd} \end{cases}$$

and when $||r_k||_2 = |x_k^T r_k|$ *for some* $k \in \{1, \cdots, m\}$*, define*

$$\Delta A_k := \frac{\overline{z_0}}{||(1, \Lambda^{(e)}(\lambda))||_q}(x_k^T r_k)\overline{x_k}x_k^H + \frac{\overline{y_0}}{||(1,\lambda)||_q}[r_k x_k^H + \overline{x_k}r_k^T - 2(x_k^T r_k)\overline{x_k}x_k^H],$$

$$\Delta B_{kj} := \begin{cases} \frac{\overline{z_j}}{||(1,\Lambda^{(e)}(\lambda))||_q}(x_k^T r_k)\overline{x_k}x_k^H + \frac{\overline{y_j}}{||(1,\lambda)||_q}[r_k x_k^H + \overline{x_k}r_k^T - 2(x_k^T r_k)\overline{x_k}x_k^H] & \text{when } j \text{ is even} \\ \frac{\overline{y_j}}{||(1,\lambda)||_q}[r_k x_k^H - \overline{x_k}r_k^T] & \text{when } j \text{ is odd} \end{cases}$$

for all $k, j = 1, \cdots, m$*. Consider* $\Delta W_k(z) := \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$ *for* $k = 1, \cdots, m$ *and* $\Delta\mathbb{W}(z) := (\Delta W_1(z), \cdots, \Delta W_m(z))$*. Then* $\Delta\mathbb{W}$ *is* T*-even alternating,* $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ *and* $\|\Delta\mathbb{W}\|_p = \eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W})$*.*

Proof. Recall from the proof of Theorem 6.3.5 that

$$\Delta A_k = \overline{Q_k}\begin{bmatrix} \frac{\overline{z_0}}{||(1,\Pi_{(e)}(\lambda))||_q}x_k^T r_k & \frac{\overline{y_0}}{||(1,\lambda)||_q}(Q^{(k)^T} r_k)^T \\ \frac{\overline{y_0}}{||(1,\lambda)||_q}Q^{(k)^T} r_k & \widehat{A_k} \end{bmatrix} Q_k^H,$$

$$\Delta B_{kj} = \begin{cases} \overline{Q_k}\begin{bmatrix} \frac{\overline{z_j}}{||(1,\Pi_{(e)}(\lambda))||_q}x_k^T r_k & \frac{\overline{y_j}}{||(1,\lambda)||_q}(Q^{(k)^T} r_k)^T \\ \frac{\overline{y_j}}{||(1,\lambda)||_q}Q^{(k)^T} r_k & \widehat{B_{kj}} \end{bmatrix} Q_k^H & \text{when } j \text{ is even} \\ \overline{Q_k}\begin{bmatrix} 0 & -\frac{\overline{y_j}}{||(1,\lambda)||_q}(Q^{(k)^T} r_k)^T \\ \frac{\overline{y_j}}{||(1,\lambda)||_q}Q^{(k)^T} r_k & \widehat{B_{kj}} \end{bmatrix} Q_k^H & \text{when } j \text{ is odd} \end{cases}$$

for all $k, j = 1, \cdots, m$. When $||r_k||_2 \neq |x_k^T r_k|$ for some $k \in \{1, \cdots, m\}$, for $\mu_{\Delta A_k} = a_k$ and $\mu_{\Delta B_{kj}} = b_{kj}$, by Theorem 6.4.1, we have

$$\widehat{A_k} = -\frac{\overline{y_0}^2 z_0 \overline{x_k^T r_k}}{|y_0|^2(||r_k||_2^2 - |x_k^T r_k|^2)||(1, \Pi_{(e)}(\lambda))||_q}(Q^{(k)^T} r_k)(Q^{(k)^T} r_k)^T,$$

$$\widehat{B_{kj}} = \begin{cases} -\frac{\overline{y_j}^2 z_j \overline{x_k^T r_k}}{|y_j|^2(||r_k||_2^2 - |x_k^T r_k|^2)||(1,\Pi_{(e)}(\lambda))||_q}(Q^{(k)^T} r_k)(Q^{(k)^T} r_k)^T & \text{when } j \text{ is even} \\ 0 & \text{when } j \text{ is odd.} \end{cases}$$

When $||r_k||_2 = |x_k^T r_k|$ for some $k \in \{1, \cdots, m\}$ then noting that $Q^{(k)^T} r_k = 0$ we have $\Delta A_k = \overline{Q_k}\begin{bmatrix} \frac{\overline{z_0}}{||(1,\Pi_{(e)}(\lambda))||_q}x_k^T r_k & 0 \\ 0 & \widehat{A_k} \end{bmatrix} Q_k^H$, $\Delta B_{kj} = \overline{Q_k}\begin{bmatrix} \frac{\overline{z_j}}{||\lambda||_q}x_k^T r_k & 0 \\ 0 & \widehat{B_{kj}} \end{bmatrix} Q_k^H$ when j is even and $\Delta B_{kj} = \overline{Q_k}\begin{bmatrix} 0 & 0 \\ 0 & \widehat{B_{kj}} \end{bmatrix} Q_k^H$ when j is odd. Again for $\mu_{\Delta A_k} = a_k$ and $\mu_{\Delta B_{kj}} =$

b_{kj}, by Theorem 6.4.1, we have $\widehat{A_k} = \widehat{B_{kj}} = 0$. So $||\Delta W_k||_p = ||(a_k, b_{k1}, \cdots, b_{km})||_p = c_k$ and hence $\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \|\Delta \mathbb{W}\|_p = ||(c_1, \cdots, c_m)||_p$. Simplifying ΔA_k and ΔB_{kj}, the desired perturbation follows. $\square$

Similarly we have the following result for a T-odd alternating MEP.

Theorem 6.4.7. *Let $\mathbb{S}$ be the space of all T-odd alternating MEPs and $\mathbb{W} \in \mathbb{S}$ be as given in (6.1). Let $\lambda := (\lambda_1, \cdots, \lambda_m)^T \in \mathbb{C}^m$ and $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ be such that $x_k^H x_k = 1$. Consider $r_k := -W_k(\lambda) x_k$, for all $k = 1, \cdots, m$. Let $(y_0, \cdots, y_m)^T \in \partial ||(1, \lambda)||_q$ and $(z_1, \cdots, z_m)^T \in \partial ||\Pi_{(o)}(\lambda)||_q$ for $p^{-1} + q^{-1} = 1$. Then*

$$\eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \begin{cases} ||(c_1, \cdots, c_m)||_p & \text{when } \Pi_{(o)}(\lambda) \neq 0 \\ \frac{1}{||(1,\lambda)||_q} ||(||r_1||_2, \cdots, ||r_m||)||_p & \text{when } \Pi_{(o)}(\lambda) = 0, \end{cases}$$

where $c_k = ||(a_k, b_{k1}, \cdots, b_{km})||_p$, $a_k = \frac{|y_0|}{||(1,\lambda)||_q} (||r_k||_2^2 - |x_k^T r_k|^2)^{\frac{1}{2}}$ and

$$b_{kj} = \begin{cases} \frac{|y_j|}{||(1,\lambda)||_q} (||r_k||_2^2 - |x_k^T r_k|^2)^{\frac{1}{2}} & \text{when } j \text{ is even} \\ \left(\frac{|z_j|^2}{||\Lambda^{(o)}(\lambda)||_q^2} |x_k^T r_k|^2 + \frac{|y_j|^2}{||(1,\lambda)||_q^2} (||r_k||_2^2 - |x_k^T r_k|^2) \right)^{\frac{1}{2}} & \text{when } j \text{ is odd} \end{cases}$$

for all $k, j = 1, \cdots, m$. In particular, for $p = 2$, we have

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \begin{cases} \left(\sum_{k=1}^m \left(\frac{1}{||\Lambda^{(o)}(\lambda))||_2^2} |x_k^T r_k|^2 + \frac{1}{||(1,\lambda)||_2^2} \left(||r_k||_2^2 - |x_k^T r_k|^2 \right) \right) \right)^{\frac{1}{2}} & \text{when } \Pi_{(o)}(\lambda) \neq 0 \\ \frac{1}{||(1,\lambda)||_2} ||(||r_1||_2, \cdots, ||r_m||)||_2 & \text{when } \Pi_{(o)}(\lambda) = 0. \end{cases}$$

When $||r_k||_2 \neq |x_k^T r_k|$ for some $k \in \{1, \cdots, m\}$, define $\Delta A_k := \frac{\overline{y_0}}{||(1,\lambda)||_q} [r_k x_k^H - \overline{x_k} r_k^T]$,

$$\Delta B_{kj} := \begin{cases} \frac{\overline{y_j}}{||(1,\lambda)||_q} [r_k x_k^H - \overline{x_k} r_k^T] & \text{when } j \text{ is even} \\ \frac{\overline{z_j}}{||\Lambda^{(o)}(\lambda)||_q} (x_k^T r_k) \overline{x_k} x_k^H + \frac{\overline{y_j}}{||(1,\lambda)||_q} [r_k x_k^H + \overline{x_k} r_k^T - 2(x_k^T r_k) \overline{x_k} x_k^H] & \\ -\frac{\overline{y_j}^2 z_j \overline{x_k^T r_k}}{|y_j|^2 (||r_k||_2^2 - |x_k^T r_k|^2) ||\Lambda^{(e)}(\lambda)||_q} (I_{n_k} - \overline{x_k} x_k^T) r_k r_k^T (I_{n_k} - x_k x_k^H) & \text{when } j \text{ is odd} \end{cases}$$

and when $||r_k||_2 = |x_k^T r_k|$ for some $k \in \{1, \cdots, m\}$, define $\Delta A_k := \frac{\overline{y_0}}{||(1,\lambda)||_q} [r_k x_k^H - \overline{x_k} r_k^T]$,

$$\Delta B_{kj} := \begin{cases} \frac{\overline{y_j}}{||(1,\lambda)||_q} [r_k x_k^H - \overline{x_k} r_k^T] & \text{when } j \text{ is even} \\ \frac{\overline{z_j}}{||\Lambda^{(o)}(\lambda)||_q} (x_k^T r_k) \overline{x_k} x_k^H + \frac{\overline{y_j}}{||(1,\lambda)||_q} [r_k x_k^H + \overline{x_k} r_k^T - 2(x_k^T r_k) \overline{x_k} x_k^H] & \text{when } j \text{ is odd,} \end{cases}$$

for all $k, j = 1, \cdots, m$. Consider $\Delta W_k(z) := \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$ for $k = 1, \cdots, m$ and $\Delta \mathbb{W}(z) := (\Delta W_1(z), \cdots, \Delta W_m(z))$. Then $\Delta \mathbb{W}$ is T-odd alternating, $\mathbb{W}(\lambda) x + \Delta \mathbb{W}(\lambda) x = 0$ and $\|\Delta \mathbb{W}\|_p = \eta_p^{\mathbb{S}}(\lambda, x, \mathbb{W})$.

Now we derive the structured backward error for the H-structures by considering Hölder 2-norm $|||\cdot|||_2$ on $\mathbb{MEP}(n_1, \cdots, n_m)$.

Theorem 6.4.8. *Let $\mathbb{S}$ be the space of all H-Hermitian MEPs and $\mathbb{W} \in \mathbb{S}$ be as given in (6.1). Let $\lambda := (\lambda_1, \cdots, \lambda_m)^T \in \mathbb{C}^m$ be such that $\lambda_j := \alpha_j + i\beta_j$ for all $j = 1, \cdots, m$, where $i = \sqrt{-1}$ and $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ be such that $x_k^H x_k = 1$. Consider $r_k := -W_k(\lambda)x_k$ and $P_k = I_{n_k} - x_k x_k^H$ for all $k = 1, \cdots, m$. Then*

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \begin{cases} \frac{1}{||(1,\lambda)||_2} \, ||(||r_1||_2 \, , \cdots , ||r_m||_2)|| & \text{when } \lambda \in \mathbb{R}^m \\ \left(\sum_{k=1}^m \left(\sum_{j=0}^m |s_{kj}|^2 + \frac{1}{||(1,\lambda)||_2^2} \left(||r_k||_2^2 - |x_k^H r_k|^2\right)\right)\right)^{\frac{1}{2}} & \text{when } \lambda \in \mathbb{C}^m \setminus \mathbb{R}^m, \end{cases}$$

where s_{kj} is as defined in Theorem 6.3.7.

When $||r_k||_2 \neq |x_k^H r_k|$, for some $k \in \{1, \cdots, m\}$, define

$$\Delta A_k = \begin{cases} \frac{1}{||(1,\lambda)||_2^2} \left(r_k x_k^H + x_k r_k^H - (r_k^H x_k) x_k x_k^H\right) - \frac{x_k^H r_k P_k r_k r_k^H P_k}{||(1,\lambda)||_2^2 (||r_k||_2^2 - |x_k^H r_k|^2)} & \text{when } \lambda \in \mathbb{R}^m \\ s_{k0} x_k x_k^H + \frac{1}{||(1,\lambda)||_2^2} \left[P_k r_k x_k^H + x_k r_k^H P_k\right] - \frac{s_{k0} P_k r_k r_k^H P_k}{||r_k||_2^2 - |x_k^H r_k|^2} & \text{when } \lambda \in \mathbb{C}^m \setminus \mathbb{R}^m, \end{cases}$$

$$\Delta B_{kj} = \begin{cases} \frac{\lambda_j}{||(1,\lambda)||_2^2} \left(r_k x_k^H + x_k r_k^H - (r_k^H x_k) x_k x_k^H\right) - \frac{\lambda_j x_k^H r_k P_k r_k r_k^H P_k}{||(1,\lambda)||_2^2 (||r_k||_2^2 - |x_k^H r_k|^2)} & \text{when } \lambda \in \mathbb{R}^m \\ s_{kj} x_k x_k^H + \frac{1}{||(1,\lambda)||_2^2} \left[\overline{\lambda_j} P_k r_k x_k^H + \lambda_j x_k r_k^H P_k\right] - \frac{s_{kj} P_k r_k r_k^H P_k}{||r_k||_2^2 - |x_k^H r_k|^2} & \text{when } \lambda \in \mathbb{C}^m \setminus \mathbb{R}^m \end{cases}$$

and when $||r_k||_2 \neq |x_k^H r_k|$, for some $k \in \{1, \cdots, m\}$, define

$$\Delta A_k = \begin{cases} \frac{1}{||(1,\lambda)||_2^2} (x_k^H r_k) x_k x_k^H = \frac{1}{||(1,\lambda)||_2^2} r_k x_k^H & \text{when } \lambda \in \mathbb{R}^m \\ x_k s_{k0} x_k^H & \text{when } \lambda \in \mathbb{C}^m \setminus \mathbb{R}^m, \end{cases}$$

$$\Delta B_{kj} = \begin{cases} \frac{\lambda_j}{||(1,\lambda)||_2^2} (x_k^H r_k) x_k x_k^H = \frac{\lambda_j}{||(1,\lambda)||_2^2} r_k x_k^H & \text{when } \lambda \in \mathbb{R}^m \\ x_k s_{kj} x_k^H & \text{when } \lambda \in \mathbb{C}^m \setminus \mathbb{R}^m, \end{cases}$$

for all $k, j = 1, \cdots, m$. Consider $\Delta W_k(z) := \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$ for $k = 1, \cdots, m$ and $\Delta\mathbb{W}(z) := (\Delta W_1(z), \cdots, \Delta W_m(z))$. Then $\Delta\mathbb{W}$ is H-Hermitian, $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ and $|||\Delta\mathbb{W}|||_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W})$.

Proof. When $\lambda \in \mathbb{C}^m \setminus \mathbb{R}^m$, recall from Theorem 6.3.7 that

$$\begin{aligned} \Delta A_k &= Q_k \begin{bmatrix} s_{k0} & \frac{1}{||(1,\lambda)||_2} ({Q^{(k)}}^H r_k)^H \\ \frac{1}{||(1,\lambda)||_2} {Q^{(k)}}^H r_k & \widehat{A_k} \end{bmatrix} Q_k^H, \\ \Delta B_{kj} &= Q_k \begin{bmatrix} s_{kj} & \frac{\lambda_j}{||(1,\lambda)||_2} ({Q^{(k)}}^H r_k)^H \\ \frac{\overline{\lambda_j}}{||(1,\lambda)||_2} {Q^{(k)}}^H r_k & \widehat{B_{kj}} \end{bmatrix} Q_k^H \end{aligned}$$

for all $j = 1, \cdots, m$. So for $\mu_{\Delta A_k} := \left(|s_{k0}|^2 + \frac{1}{||(1,\lambda)||_2^2}\left(||r_k||_2^2 - |x_k^H r_k|^2\right)\right)^{\frac{1}{2}}$ and $\mu_{\Delta B_{kj}} := \left(|s_{kj}|^2 + \frac{|\lambda_j|^2}{||(1,\lambda)||_2^2}\left(||r_k||_2^2 - |x_k^H r_k|^2\right)\right)^{\frac{1}{2}}$ by Theorem 6.4.1, we have $\widehat{A_k} = -\frac{s_{k0}(Q^{(k)^H} r_k)(Q^{(k)^H} r_k)^H}{||r_k||_2^2 - |x_k^H r_k|^2}$ and $\widehat{B_{kj}} = -\frac{s_{kj}(Q^{(k)^H} r_k)(Q^{(k)^H} r_k)^H}{||r_k||_2^2 - |x_k^H r_k|^2}$. When $||r_k||_2 = |x_k^H r_k|$ for some $k \in \{1, \cdots, m\}$ then note that $||Q^{(k)^H} r_k||_2 = (||r_k||_2^2 - |x_k^H r_k|^2)^{\frac{1}{2}} = 0$, i.e., $Q^{(k)^H} r_k = 0$. So we have $\Delta A_k = Q_k \begin{bmatrix} s_{k0} & 0 \\ 0 & \widehat{A_k} \end{bmatrix} Q_k^H$ and $\Delta B_{kj} = Q_k \begin{bmatrix} s_{kj} & 0 \\ 0 & \widehat{B_{kj}} \end{bmatrix} Q_k^H$. For $\mu_{\Delta A_k} = |s_{k0}|$ and $\mu_{\Delta B_{kj}} = |s_{kj}|$, again by Theorem 6.4.1, we have $\widehat{A_k} = \widehat{B_{kj}} = 0$. Hence $||\Delta W_k||_2 = \left(\sum_{j=0}^m |s_{kj}|^2 + ||r_k||_2^2 - |x_k^H r_k|^2\right)^{\frac{1}{2}}$ and

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \|\Delta \mathbb{W}\|_2 = \left(\sum_{k=1}^m \left(\sum_{j=0}^m |s_{kj}|^2 + ||r_k||_2^2 - |x_k^H r_k|^2\right)\right)^{\frac{1}{2}}.$$

When $\lambda \in \mathbb{R}^m$, recall from Theorem 6.3.7 that

$$\begin{aligned} \Delta A_k &= Q_k \begin{bmatrix} \frac{1}{||(1,\lambda)||_2^2} x_k^H r_k & \frac{1}{||(1,\lambda)||_2^2}(Q^{(k)^H} r_k)^H \\ \frac{1}{||(1,\lambda)||_2^2} Q^{(k)^H} r_k & \widehat{A}_k \end{bmatrix} Q_k^H, \\ \Delta B_{kj} &= Q_k \begin{bmatrix} \frac{\lambda_j}{||(1,\lambda)||_2^2} x_k^H r_k & \frac{\lambda_j}{||(1,\lambda)||_2^2}(Q^{(k)^H} r_k)^H \\ \frac{\lambda_j}{||(1,\lambda)||_2^2} Q^{(k)^H} r_k & \widehat{B}_{kj} \end{bmatrix} Q_k^H \end{aligned}$$

for all $j = 1, \cdots, m$. When $||r_k||_2 \neq |x_k^H r_k|$ for $\mu_{\Delta A_k} = \frac{1}{||(1,\lambda)||_2^2}||r_k||_2$ and $\mu_{\Delta B_{kj}} = \frac{|\lambda_j|}{||(1,\lambda)||_2^2}||r_k||_2$ by Theorem 6.4.1, we have

$$\widehat{A_k} = -\frac{x_k^H r_k (Q^{(k)^H} r_k)(Q^{(k)^H} r_k)^H}{||r_k||_2^2 - |x_k^H r_k|^2} \text{ and } \widehat{B_{kj}} = -\frac{\lambda_j \; x_k^H r_k (Q^{(k)^H} r_k)(Q^{(k)^H} r_k)^H}{||r_k||_2^2 - |x_k^H r_k|^2}.$$

Finally when $||r_k||_2 = |x_k^H r_k|$ for some $k \in \{1, \cdots, m\}$ then $\Delta A_k = Q_k \begin{bmatrix} \frac{1}{||(1,\lambda)||_2^2} x_k^H r_k & 0 \\ 0 & \widehat{A_k} \end{bmatrix} Q_k^H$ and $\Delta B_{kj} = Q_k \begin{bmatrix} \frac{\lambda_j}{||(1,\lambda)||_2^2} x_k^H r_k & 0 \\ 0 & \widehat{B_{kj}} \end{bmatrix} Q_k^H$. For $\mu_{\Delta A_k} = \frac{1}{||(1,\lambda)||_2^2}||r_k||_2$ and $\mu_{\Delta B_{kj}} = \frac{|\lambda_j|}{||(1,\lambda)||_2^2}||r_k||_2$ by Theorem 6.4.1, we have $\widehat{A_k} = \widehat{B_{kj}} = 0$. Therefore $||\Delta W_k||_2 = \frac{1}{||(1,\lambda)||_2}||r_k||_2$ and and $\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \|\Delta \mathbb{W}\|_2 = \frac{1}{||(1,\lambda)||_2}||(||r_1||_2, \cdots, ||r_m||_2)||_2$. The desired optimal perturbations follow by simplifying ΔA_k and ΔB_{kj}. $\square$

Next we derive $\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W})$ for H-even MEP.

Theorem 6.4.9. *Let $\mathbb{S}$ be the space of all H-even MEPs and $\mathbb{W} \in \mathbb{S}$ be as given in (6.1). Let $\lambda := (\lambda_1, \cdots, \lambda_m)^T \in \mathbb{C}^m$ be such that $\lambda_j := \alpha_j + i\beta_j$,for all $j = 1, \cdots, m$, where $i = \sqrt{-1}$ and $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ be such that $x_k^H x_k = 1$. Consider $r_k := -W_k(\lambda)x_k$ and $P_k := I_{n_k} - x_k x_k^H$ for all $k = 1, \cdots, m$. Then*

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \begin{cases} \frac{1}{||(1,\lambda)||_2} ||(||r_1||_2, \cdots, ||r_m||_2)||_2 & \text{when } \lambda \in i\mathbb{R}^m \\ \left(\sum_{k=1}^m \left(\sum_{j=0}^m |s_{kj}|^2 + \frac{1}{||(1,\lambda)||_2^2}\left(||r_k||_2^2 - |x_k^H r_k|^2\right)\right)\right)^{\frac{1}{2}} & \text{when } \lambda \in \mathbb{C}^m \setminus i\mathbb{R}^m \end{cases}$$

where s_{kj} are as given in Theorem 6.3.8.

When $||r_k||_2 \neq |x_k^H r_k|$, for some $k \in \{1, \cdots, m\}$, define

$$\Delta A_k := \begin{cases} \frac{1}{||(1,\lambda)||_2^2}\left(r_k x_k^H + x_k r_k^H - (r_k^H x_k)x_k x_k^H\right) - \frac{x_k^H r_k P_k r_k r_k^H P_k}{||(1,\lambda)||_2^2\left(||r_k||_2^2 - |x_k^H r_k|^2\right)} & \lambda \in i\mathbb{R}^m \\ s_{k0} x_k x_k^H + \frac{1}{||(1,\lambda)||_2^2}\left[P_k r_k x_k^H + x_k r_k^H P_k\right] - \frac{s_{k0} P_k r_k r_k^H P_k}{||r_k||_2^2 - |x_k^H r_k|^2} & \lambda \in \mathbb{C}^m \setminus i\mathbb{R}^m, \end{cases}$$

$$\Delta B_{kj} := \begin{cases} \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2}\left(r_k x_k^H + x_k r_k^H - (r_k^H x_k)x_k x_k^H\right) + + \frac{i s_{kj} P_k r_k r_k^H P_k}{||r_k||_2^2 - |x_k^H r_k|^2} & \lambda \in i\mathbb{R}^m \\ i s_{kj} x_k x_k^H + \frac{1}{||(1,\lambda)||_2^2}\left[\overline{\lambda_j} P_k r_k x_k^H - \lambda_j x_k r_k^H P_k\right] + \frac{i s_{kj} P_k r_k r_k^H P_k}{||r_k||_2^2 - |x_k^H r_k|^2} & \lambda \in \mathbb{C}^m \setminus i\mathbb{R}^m \end{cases}$$

When $||r_k||_2 = |x_k^H r_k|$ for some $k \in \{1, \cdots, m\}$, define

$$\Delta A_k := \begin{cases} \frac{1}{||(1,\lambda)||_2^2} r_k x_k^H & \lambda \in i\mathbb{R}^m \\ s_{k0} x_k x_k^H & \lambda \in \mathbb{C}^m \setminus i\mathbb{R}^m, \end{cases} \qquad \Delta B_{kj} := \begin{cases} \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2} r_k x_k^H & \lambda \in i\mathbb{R}^m \\ i s_{kj} x_k x_k^H & \lambda \in \mathbb{C}^m \setminus i\mathbb{R}^m \end{cases}$$

for all $k, j = 1, \cdots, m$. Consider $\Delta W_k(z) := \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$ for $k = 1, \cdots, m$ and $\Delta\mathbb{W}(z) := (\Delta W_1(z), \cdots, \Delta W_m(z))$. Then $\Delta\mathbb{W}$ is H-even, $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ and $\|\Delta\mathbb{W}\|_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W})$.

Proof. When $\lambda \in i\mathbb{R}^m$, recall from the proof of Theorem 6.3.8 that

$$\begin{aligned} \Delta A_k &= Q_k \begin{bmatrix} \frac{1}{||(1,\lambda)||_2^2} x_k^H r_k & \frac{1}{||(1,\lambda)||_2^2}({Q^{(k)}}^H r_k)^H \\ \frac{1}{||(1,\lambda)||_2^2} {Q^{(k)}}^H r_k & \widehat{A_k} \end{bmatrix} Q_k^H, \\ \Delta B_{kj} &= Q_k \begin{bmatrix} \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2} x_k^H r_k & -\frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2}({Q^{(k)}}^H r_k)^H \\ \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2} {Q^{(k)}}^H r_k & \widehat{B_{kj}} \end{bmatrix} Q_k^H \end{aligned}$$

for all $j = 1, \cdots, m$. When $||r_k||_2 \neq |x_k^H r_k|$ for $\mu_{\Delta A_k} = \frac{1}{||(1,\lambda)||_2^2}||r_k||_2$ and $\mu_{\Delta B_{kj}} = \frac{|\lambda_j|}{||(1,\lambda)||_2^2}||r_k||_2$, by Theorem 6.4.1, we have $\widehat{A_k} = -\frac{\overline{x_k^H r_k}({Q^{(k)}}^H r_k)({Q^{(k)}}^H r_k)^H}{||(1,\lambda)||_2^2\left(||r_k||_2^2 - |x_k^H r_k|^2\right)}$ and $\widehat{B_{kj}} =$

$\frac{\lambda_j \overline{x_k^H r_k} (Q^{(k)^H} r_k)(Q^{(k)^H} r_k)^H}{||(1,\lambda)||_2^2 \left(||r_k||_2^2 - |x_k^H r_k|^2\right)}$. When $||r_k||_2 = |x_k^H r_k|$ for some $k \in \{1, \cdots, m\}$, we have $Q^{(k)^H} r_k = 0$. Hence

$$\Delta A_k = Q_k \begin{bmatrix} \frac{1}{||(1,\lambda)||_2^2} x_k^H r_k & 0 \\ 0 & \widehat{A_k} \end{bmatrix} Q_k^H \text{ and } \Delta B_{kj} = Q_k \begin{bmatrix} \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2} x_k^H r_k & 0 \\ 0 & \widehat{B_{kj}} \end{bmatrix} Q_k^H.$$

Again for $\mu_{\Delta A_k} = \frac{1}{||(1,\lambda)||_2^2} ||r_k||_2$ and $\mu_{\Delta B_{kj}} = \frac{|\lambda_j|}{||(1,\lambda)||_2^2} ||r_k||_2$ by Theorem 6.4.1, we have $\widehat{A_k} = \widehat{B_{kj}} = 0$. So $||\Delta W_k||_2 = \frac{1}{||(1,\lambda)||_2} ||r_k||_2$ and therefore $\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \|\Delta \mathbb{W}\|_2 = \frac{1}{||(1,\lambda)||_2} ||(||r_1||_2, \cdots, ||r_m||_2)||_2$.

When $\lambda \in \mathbb{C}^m \setminus i\mathbb{R}^m$, recall from Theorem 6.3.8 that

$$\begin{aligned} \Delta A_k &= Q_k \begin{bmatrix} s_{k0} & \frac{1}{||(1,\lambda)||_2^2} (Q^{(k)^H} r_k)^H \\ \frac{1}{||(1,\lambda)||_2^2} Q^{(k)^H} r_k & \widehat{A_k} \end{bmatrix} Q_k^H, \\ \Delta B_{kj} &= Q_k \begin{bmatrix} i s_{kj} & -\frac{\lambda_j}{||(1,\lambda)||_2^2} (Q^{(k)^H} r_k)^H \\ \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2} Q^{(k)^H} r_k & \widehat{B_{kj}} \end{bmatrix} Q_k^H, \end{aligned}$$

for all $j = 1, \cdots, m$. When $||r_k||_2 \neq |x_k^H r_k|$ for some $k \in \{1, \cdots, m\}$, for $\mu_{\Delta A_k} = \left(|s_{k0}|^2 + \frac{1}{||(1,\lambda)||_2^4} \left(||r_k||_2^2 - |x_k^H r_k|^2\right)\right)^{\frac{1}{2}}$ and $\mu_{\Delta B_{kj}} = \left(|s_{kj}|^2 + \frac{|\lambda_j|^2}{||(1,\lambda)||_2^4} \left(||r_k||_2^2 - |x_k^H r_k|^2\right)\right)^{\frac{1}{2}}$ by Theorem 6.4.1, we have $\widehat{A_k} = -\frac{s_{k0} (Q^{(k)^H} r_k)(Q^{(k)^H} r_k)^H}{||r_k||_2^2 - |x_k^H r_k|^2}$ and $\widehat{B_{kj}} = \frac{i s_{kj} (Q^{(k)^H} r_k)(Q^{(k)^H} r_k)^H}{||r_k||_2^2 - |x_k^H r_k|^2}$. When $||r_k||_2 = |x_k^H r_k|$ for some $k \in \{1, \cdots, m\}$ then we have $\Delta A_k = Q_k \begin{bmatrix} s_{k0} & 0 \\ 0 & \widehat{A_k} \end{bmatrix} Q_k^H$ and $\Delta B_{kj} = Q_k \begin{bmatrix} i s_{kj} & 0 \\ 0 & \widehat{B_{kj}} \end{bmatrix} Q_k^H$. For $\mu_{\Delta A_k} = |s_{k0}|$ and $\mu_{\Delta B_{kj}} = |s_{kj}|$, by Theorem 6.4.1, we have $\widehat{A_k} = \widehat{B_{kj}} = 0$. Therefore $||\Delta W_k||_2 = \left(\sum_{j=0}^m |s_{kj}|^2 + \frac{1}{||(1,\lambda)||_2^2} \left(||r_k||_2^2 - |x_k^H r_k|^2\right)\right)^{\frac{1}{2}}$ and thus $\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \|\Delta \mathbb{W}\|_2 = \left(\sum_{k=1}^m \left(\sum_{j=0}^m |s_{kj}|^2 + \frac{1}{||(1,\lambda)||_2^2} \left(||r_k||_2^2 - |x_k^H r_k|^2\right)\right)\right)^{\frac{1}{2}}$.

Simplifying $\widehat{A_k}$ and $\widehat{B_{kj}}$, we obtain the desired optimal perturbations. □

In a similar way, in case of a H-even alternating MEP, we have the following result.

Theorem 6.4.10. *Let $\mathbb{S}$ be the space of all H-even alternating MEPs and $\mathbb{W} \in \mathbb{S}$ be as given in (6.1). Let $\lambda := (\lambda_1, \cdots, \lambda_m)^T \in \mathbb{C}^m$ be such that $\lambda_j := \alpha_j + i\beta_j$ for all $j = 1, \cdots, m$, where $i = \sqrt{-1}$, $0 \neq x := x_1 \otimes \cdots \otimes x_m \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m}$ be such that*

$x_k^H x_k = 1$. *Consider* $r_k := -W_k(\lambda)x_k$ *and* $P_k := I_{n_k} - x_k x_k^H$ *for all* $k = 1, \cdots, m$. *Then*

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W}) = \begin{cases} \frac{1}{||(1,\lambda)||_2} ||(||r_1||_2, \cdots, ||r_m||_2)||_2 & \text{when } \lambda \in i\mathbb{R}^m \\ \left(\sum_{k=1}^m \left(\sum_{j=0}^m |s_{kj}|^2 + \frac{1}{||(1,\lambda)||_2^2}\left(||r_k||_2^2 - |x_k^H r_k|^2\right)\right)\right)^{\frac{1}{2}} & \text{when } \lambda \in \mathbb{C}^m \setminus i\mathbb{R}^m \end{cases}$$

where s_{kj} *are as given in Theorem 6.3.9.*

When $||r_k||_2 \neq |x_k^H r_k|$, *for some* $k \in \{1, \cdots, m\}$, *define*

$$\Delta A_k := \begin{cases} \frac{1}{||(1,\lambda)||_2^2}\left(r_k x_k^H + x_k r_k^H - (r_k^H x_k) x_k x_k^H\right) - \frac{x_k^H r_k P_k r_k r_k^H P_k}{||(1,\lambda)||_2^2\left(||r_k||_2^2 - |x_k^H r_k|^2\right)} & \text{when } \lambda \in i\mathbb{R}^m \\ s_{k0} x_k x_k^H + \frac{1}{||(1,\lambda)||_2^2}\left[P_k r_k x_k^H + x_k r_k^H P_k\right] - \frac{s_{k0} P_k r_k r_k^H P_k}{||r_k||_2^2 - |x_k^H r_k|^2} & \text{when } \lambda \in \mathbb{C}^m \setminus i\mathbb{R}^m \end{cases}$$

$$\Delta B_{kj} := \begin{cases} \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2}\left(r_k x_k^H + x_k r_k^H - (r_k^H x_k) x_k x_k^H\right) - \frac{\lambda_j x_k^H r_k P_k r_k r_k^H P_k}{||(1,\lambda)||_2^2\left(||r_k||_2^2 - |x_k^H r_k|^2\right)} & \text{when } \lambda \in i\mathbb{R}^m \\ s_{k0} x_k x_k^H + \frac{1}{||(1,\lambda)||_2^2}\left[P_k r_k x_k^H + x_k r_k^H P_k\right] - \frac{s_{kj} P_k r_k r_k^H P_k}{||r_k||_2^2 - |x_k^H r_k|^2} & \text{when } \lambda \in \mathbb{C}^m \setminus i\mathbb{R}^m \end{cases}$$

when j *is even and*

$$\Delta B_{kj} := \begin{cases} \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2}\left(r_k x_k^H + x_k r_k^H - (r_k^H x_k) x_k x_k^H\right) + \frac{\lambda_j x_k^H r_k P_k r_k r_k^H P_k}{||(1,\lambda)||_2^2\left(||r_k||_2^2 - |x_k^H r_k|^2\right)} & \text{when } \lambda \in i\mathbb{R}^m \\ i s_{kj} x_k x_k^H + \frac{1}{||(1,\lambda)||_2^2}\left[\overline{\lambda_j} P_k r_k x_k^H - \lambda_j x_k r_k^H P_k\right] + \frac{i s_{kj} P_k r_k r_k^H P_k}{||r_k||_2^2 - |x_k^H r_k|^2} & \text{when } \lambda \in \mathbb{C}^m \setminus i\mathbb{R}^m \end{cases}$$

When $||r_k||_2 = |x_k^H r_k|$ *for some* $k \in \{1, \cdots, m\}$ *define* $\Delta A_k := \frac{1}{||(1,\lambda)||_2^2} r_k x_k^H$, $\Delta B_{kj} := \frac{\overline{\lambda_j}}{||(1,\lambda)||_2^2} r_k x_k^H$ *when* $\lambda \in i\mathbb{R}^m$ *and* $\Delta A_k := s_{k0} x_k x_k^H$, $\Delta B_{kj} := \begin{cases} s_{kj} x_k x_k^H & \text{when } j \text{ is even} \\ i s_{kj} x_k x_k^H & \text{when } j \text{ is odd,} \end{cases}$ *when* $\lambda \in \mathbb{C}^m \setminus i\mathbb{R}^m$. *Consider* $\Delta W_k(z) := \Delta A_k + \sum_{j=1}^m z_j \Delta B_{kj}$ *for* $k = 1, \cdots, m$ *and* $\Delta\mathbb{W}(z) := (\Delta W_1(z), \cdots, \Delta W_m(z))$. *Then* $\Delta\mathbb{W}$ *is* H*-even alternating,* $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ *and* $\|\Delta\mathbb{W}\|_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathbb{W})$.

7

Summary

We have developed a general framework for the spectral analysis of nonhomogeneous MEPs. For a general norm on the space of MEPs, we have defined the condition number $\mathrm{cond}(\lambda, \mathbb{W})$ of a simple eigenvalue λ of an MEP $\mathbb{W}$ by

$$\mathrm{cond}(\lambda, \mathbb{W}) := \lim_{\|\Delta\mathbb{W}\|\to 0} \sup \frac{\mathrm{dist}(\lambda, \sigma(\mathbb{W}+\Delta\mathbb{W}))}{\|\Delta\mathbb{W}\|},$$

where $\Delta\mathbb{W}$ is an MEP and $\mathrm{dist}(\lambda, \sigma(\mathbb{W}+\Delta\mathbb{W})) := \min\{||\lambda-\mu||_2 : \mu \in \sigma(\mathbb{W}+\Delta\mathbb{W})\}$. Given a simple eigenvalue $\lambda_{\mathbb{W}}$ of $\mathbb{W}$, we have shown that there is an open neighborhood $\mathrm{nbd}(\mathbb{W})$ containing $\mathbb{W}$ and a smooth function λ such that $\lambda(\mathbb{W}) = \lambda_{\mathbb{W}}$ and $\lambda(\mathbb{X})$ is a simple eigenvalue of $\mathbb{X}$ for all $\mathbb{X} \in \mathrm{nbd}(\mathbb{W})$. We have determined the derivative $D\lambda(\mathbb{W})$ and used it to derive three equivalent representations of $\mathrm{cond}(\lambda, \mathbb{W})$. Also, we have constructed a fast perturbation for $\lambda_{\mathbb{W}}$ considering a weighted norm on the space of MEPs. We have defined the backward error of an approximate eigenvalue λ of $\mathbb{W}$ by

$$\eta(\lambda, \mathbb{W}) := \inf\{\|\Delta\mathbb{W}\| : \lambda \in \sigma(\mathbb{W}+\Delta\mathbb{W})\}.$$

We have derived $\eta(\lambda, \mathbb{W})$ and constructed an optimal perturbation $\Delta\mathbb{W}$ such that $\lambda \in \sigma(\mathbb{W}+\Delta\mathbb{W})$ and $\|\Delta\mathbb{W}\| = \eta(\lambda, \mathbb{W})$. Also, for an approximate eigenpair $(\lambda, x := x_1 \otimes \cdots \otimes x_m) \in \mathbb{C}^m \times (\mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_m})$, we have defined the backward error by $\eta(\lambda, x, \mathbb{W}) := \inf\{\|\Delta\mathbb{W}\| : \mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0\}$. We have derived $\eta(\lambda, x, \mathbb{W})$ and constructed optimal perturbation $\Delta\mathbb{W}$ for (λ, x) such that $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ and $\|\Delta\mathbb{W}\| = \eta(\lambda, x, \mathbb{W})$. We have extended the results to the case of two parameter PEPs. Further we have developed a general framework for spectral analysis of linear homogeneous MEPs. We have introduced condition number $\mathrm{cond}(\lambda, \mathbb{W})$ of a simple eigenvalue λ

of a homogeneous MEP $\mathbb{W}$ and derived $\mathrm{cond}(\lambda, \mathbb{W})$. We have also determined backward errors of approximate eigenvalues and approximate eigenpairs of $\mathbb{W}$ and constructed optimal perturbations.

We have defined and analyzed two vector spaces of potential linearizations of two-parameter polynomial of degree k. Considering a two-parameter matrix polynomial of the form

$$P(\lambda, \mu) := \sum_{i=0}^{k} \sum_{j=0}^{k-i} \lambda^i \mu^j A_{ij},$$

where $A_{ij} \in \mathbb{C}^{n \times n}$ for all $i = 0, \cdots, k$ and $j = 0, \cdots, k-i$, we have defined and analyzed two vector spaces, namely, the right ansatz space and the left ansatz space of potential linearizations of $P(\lambda, \mu)$. We have also analyzed conditions under which a pencil in the ansatz spaces is a linearization of $P(\lambda, \mu)$.

Finally, we have considered structured linear MEPs and analyzed structured backward errors and structured backward perturbations. We have denoted the space of structured MEPs by $\mathbb{S}$ and defined structured backward error of an approximate eigenpair $(\lambda, x := x_1 \otimes \cdots \otimes x_m)$ by

$$\eta^{\mathbb{S}}(\lambda, x, \mathbb{W}) := \inf\{\|\Delta\mathbb{W}\| : \Delta\mathbb{W} \in \mathbb{S} \text{ and } \mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0\}.$$

We have determined $\eta^{\mathbb{S}}(\lambda, x, \mathbb{W})$ and constructed an optimal structured perturbation $\Delta\mathbb{W}$ such that $\mathbb{W}(\lambda)x + \Delta\mathbb{W}(\lambda)x = 0$ and $\|\Delta\mathbb{W}\| = \eta^{\mathbb{S}}(\lambda, x, \mathbb{W})$.

Bibliography

[1] B. Adhikari, *Vector space of linearizations for the quadratic two-parameter matrix polynomial*, Linear Multilinear Algebra, 61(2013), pp.603-616.

[2] B. Adhikari, *Backward Perturbation and Sensitivity Analysis if Structured Polynomial Eigenvalue Problem*, PhD thesis, Department of Mathematics, Indian Institute of Technology Guwahati, India, 2009.

[3] B. Adhikari and R. Alam, *Structured backward errors and pseudospectra of structured matrix pencils*, SIAM J. Matrix Anal. Appl., 31(2009), pp.331-359.

[4] B. Adhikari and R. Alam, *On backward errors of structured polynomial eigenproblems solved by structure preserving linearizations*, Linear Algebra Appl., 434(2011), pp.1989-2017.

[5] B. Adhikari, R. Alam and D. Kressner, *Structured eigenvalue condition numbers and linearizations for matrix polynomials*, Linear Algebra Appl., 435(2011), pp.2193-2221.

[6] S. S. Ahmad, *Pseudospectra of matrix pencils and their applications in perturbation analysis of eigenvalues and eigendecompositions*, PhD thesis, Department of Mathematics, Indian Institute of Technology Guwahati, India, 2007.

[7] S.S. Ahmed and R. Alam, *Sensitivity analysis of nonlinear eigenproblems.*, preprint.

[8] S.S. Ahmed, R. Alam and R. Byers, *Pseudospectra, critical points and multiple eigenvalues of matrix pencils*, SIAM J. Matrix Anal. Appl., 31(2010), pp.1915-1933.

[9] S.S. Ahmed and R. Alam, *Pseudospectra, critical points and multiple eigenvalues of matrix polynomials*, Linear Algebra Appl., 430(2009), pp.1171-1195.

[10] R. Alam, *On the sensitivity analysis of eigenvalues*, Electron. J. Linear Algebra 29(2016), pp.223-235.

[11] F.V.Atkinson, *Multiparameter spectral theory*, Bull. Am. Math. Soc. 74(1968), pp.1-27.

[12] F.V.Atkinson, *Multiparameter Eigenvalue Problems*, Academic Press, New York, 1972.

[13] F.V.Atkinson and Angelo B. Mingarelli, *Multiparameter Eigenvalue Problems:Strum-Liouville Theory*, Academic Press, New York, 1972.

[14] T. Bhattacharyya, P. A. Binding and K. Seddighi, *Multiparameter sturm-liouville problems with eigenparameter dependent boundary conditions*, J. Math. Anal. Appl., 264(2001), pp.560-576.

[15] T. Bhattacharyya, T. Košir and B. Plestenjak, *Right definite multiparameter sturm-liouville problems with eigenparameter dependent boundary conditions*, Proc. Edinb. Math. Soc. (2), 45(2002), pp.565-578.

[16] P. Binding, *Left definite multiparameter eigenvalue problems*, Trans. Amer. Math. Soc., 272(1982), pp.475-486.

[17] P. Binding and P. J. Browne, *Two Parameter Eigenvalue Problems for Matrices*, Linear Algebra Appl., 113(1989), pp.139-157.

[18] P. J. Browne, *Abstract multiparameter theory,I*, J. Math. Anal. Appl., 60(1977), pp.259-273.

[19] P. J. Browne and B. D. Sleeman, *Regular multiparameter eigenvalue Problems with several parameters in the boundary conditions*, J. Math. Anal. Appl., 72(1979), pp.29-33.

[20] R. D. Carmichael, *Boundary value and expansion problems: algebraic basis of the theory*, Amer. J. Math., 43(1921), pp.69-101.

[21] R. D. Carmichael, *Boundary value and expansion problems: formulation of various transcendental problems*, Amer. J. Math., 43(1921), pp.232-270

[22] R. D. Carmichael *Boundary value and expansion problems: oscillatory, comparison and expansion problems*, Amer. J. Math., 44(1922), pp.129-152.

[23] N. Cottin, *Dynamic model updating - a multiparameter eigenvalue problem*, Mech. Systems Signal Process., 15(2001), pp.649-665.

[24] D. Cox, J. Little and D. Oshea, *Using algebraic geometry*, Springer Verlag, 1997.

[25] C. Davis, W. M. Kahan and H.F. Weinberger, *Norm-preserving dialations and their applications to optimal error bounds*, SIAM J. Numer. Anal., 19(1982), pp.445-469.

[26] J.-P. Dedieu and F. Tisseur, *Perturbation theory for homogeneous polynomial eigenvalue problems*, Linear Algebra Appl., 358 (2003), pp.7194.

[27] I. Gohberg, P. Lancaster and L. Rodman, *Matrix Polynomials*, Academic Press, Inc., New York, 1982.

[28] L. Grunenfelder and T. Košir, *An Algebraic approach to multiparameter spectral theory*, Trans. Amer. Math. Soc., 348(1996), pp.2983-2998.

[29] L. Grunenfelder and T. Košir, *Geometric aspects of multiparameter spectral theory*, Trans. Amer. Math. Soc., 350(1998), pp.2525-2546

[30] M. E. Hochstenbach and B. Plestenjak, *A JacobiDavidson type method for a right definite two-parameter eigenvalue problem*, SIAM J. Matrix Anal. Appl., 24(2002), pp.392-410.

[31] M.E. Hochstenbach and B. Plestenjak, *Backward error, condition number and pseudospectra for the multiparameter eigenvalue problem*, Linear Algebra Appl., 375(2003), pp.63-81

[32] M.E. Hochstenbach, T. Košir and B. Plestenjak, *A jacobi-davidson type method for the two-parameter eigenvalue problem* SIAM J. Matrix Anal. Appl., 26(2005), pp.477-497.

[33] M.E. Hochstenbach and B. Plestenjak, *Harmonic rayleigh ritz extraction for the multiparameter eigenvalue problem*, Electron. Trans. Numer. Anal., 29(2008), pp.81-96.

[34] M. E. Hochstenbach, A. Muhic and B. Plestenjak, *On linearizations of the quadratic two-parameter eigenvalue problems*, Linear Algebra Appl., 436(2012), pp.2725-2743.

[35] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, reprinted paperback version ed., 1995.

[36] E. Jarlebring and M. E. Hochstenbach, *Polynomial two-parameter eigenvalue problems and matrix pencil methods for stability of delay-differential equations*, Linear Algebra Appl., 431(2009), pp.369-380.

[37] T. Košir, *Finite dimensional multiparameter spectral theory: the nonderogatory case,* Linear Algebra Appl., 212/213(1994), pp.45-70.

[38] T. Košir, *Kronecker bases for linear matrix equations, with application to two-parameter eigenvalue problems*, Linear Algebra Appl., 249(1996), pp.259-288.

[39] T. Košir, *The cayley-hamilton theorem and inverse problems for multiparameter systems*, Linear Algebra Appl., 367(2003), pp.155-163.

[40] J. R. Kuttler and V. G. Sigillito, *Eigenvalues of the laplacian in two dimensions*, SIAM Rev., 26(1984), pp.163-193.

[41] B. C. Lesieutre, A. V. Mamishev, Y. Du, E. Keskiner, M. Zahn and G. C. Verghese, *Forward and inverse parameter estimation algorithms of interdigital dielectrometry sensors*, IEEE Trans. Dielectrecs Elect. Insul., 8(2001), pp.577-588.

[42] D. S. Mackey, N. Mackey, C. Mehl and V. Mehrmann, *Vector spaces of linearizations for matrix polynomials,* SIAM J. Matrix Anal. Appl., 28(2006), pp.971-1004.

[43] K. Meerbergen, C. Schröder and H. Voss, *A Jacobi-Davidson method for two-real-parameter nonlinear eigenvalue problems arising from delay-differential equations*, Numer. Lin. Alg. Appl., 20(2013), pp.852-868.

[44] A. Muhič and B. Plestenjak, *On the quadratic two-parameter eigenvalue problem and its linearization*, Linear Algebra Appl., 432 (2010), pp.2529-2542.

[45] M. R. Osborne, *Iterative procedures for solving finite-difference approximations to separable partial differential equation*, Comput. J., 6(1963), pp.93-99.

[46] B. Plestenjak, *A numerical algorithm for computing a basis for the root subspace at a nonderogatory eigenvalue of a multiparameter system,* Linear Algebra Appl., 285(1998), pp.257-276.

[47] B. Plestenjak, *Numerical methods for nonlinear two-parameter eigenvalue problems*, BIT Numer. Math., 56(2016), pp.241-262.

[48] R. A. Smith, *The condition numbers of the matrix eigenvalue problem*, Numer. Math., 10(1967), pp.232-240.

[49] W.H. Steeb, *Matrix calculus and Kronecker product with applications and C++ programs*, World Scientific Publishing Co., Inc., River Edge, NJ, 1997.

[50] F. Tisseur, *Backward error and condition of polynomial eigenvalue problems*, Linear Algebra Appl., 309(2000), pp.339361.

[51] F. Tisseur and K. Meerbergen, *The quadratic eigenvalue problem*, SIAM Rev., 43(2001), pp.235286.

[52] H. Volkmer, *Multiparameter Problems and Expansion Theorems*, Lecture Notes in Mathematics, Springer, Berlin, New York, 1988.