Advances in Operator Theory (2023) 8:65
https://doi.org/10.1007/s43036-023-00292-8

Tusi Mathematical Research Group

ORIGINAL PAPER

Check for updates

Hoffman–Wielandt type inequality for block companion matrices of certain matrix polynomials

Pallavi Basavaraju[1] · Shrinath Hadimani[1] · Sachindranath Jayaraman[1]

Received: 20 March 2023 / Accepted: 24 August 2023 / Published online: 11 September 2023
© Tusi Mathematical Research Group (TMRG) 2023

Abstract

Matrix polynomials with unitary or doubly stochastic coefficients form the subject matter of this manuscript. We prove that if $P(\lambda)$ is a quadratic matrix polynomial whose coefficients are either unitary matrices or doubly stochastic matrices, then under certain conditions on these coefficients, the corresponding block companion matrix C is diagonalizable. Consequently, if $Q(\lambda)$ is another quadratic matrix polynomial with corresponding block companion matrix D, then a Hoffman–Wielandt type inequality holds for the block companion matrices C and D.

Keywords Matrix polynomials with unitary or doubly stochastic coefficients · Eigenvalue bounds for matrix polynomials with doubly stochastic coefficients · Diagonalizability of block companion matrix · Hoffman–Wielandt type inequality

Mathematics Subject Classification 15A42 · 47A56 · 15A20 · 15A22

1 Introduction

We work either over the field $\mathbb{C}$ of complex numbers or over the field $\mathbb{R}$ of real numbers. The vector space of $n \times n$ matrices over $\mathbb{C}$ (resp., $\mathbb{R}$) is denoted by $M_n(\mathbb{C})$ (resp., $M_n(\mathbb{R})$). The notations $\|\cdot\|_2$ and $\|\cdot\|_F$ will denote, respectively, the spectral norm and the Frobenius norm of a square matrix. An $n \times n$ matrix polynomial of degree m is

Communicated by Dragan S. Djordjevic.

✉ Sachindranath Jayaraman
sachindranathj@iisertvm.ac.in; sachindranathj@gmail.com

Pallavi Basavaraju
pallavipoorna20@iisertvm.ac.in

Shrinath Hadimani
srinathsh3320@iisertvm.ac.in

[1] School of Mathematics, Indian Institute of Science Education and Research Thiruvananthapuram, Maruthamala P.O., Vithura, Thiruvananthapuram, Kerala 695 551, India

a mapping P from $\mathbb{C}$ to $M_n(\mathbb{C})$, defined by $P(\lambda) = \sum_{i=0}^{m} A_i\lambda^i$, where $A_i \in M_n(\mathbb{C})$ and $A_m \neq 0$. Matrix polynomials arise in several areas of applied mathematics. A good source of reference on matrix polynomials is the monograph by Gohberg et al. [8].

An $n \times n$ matrix polynomial $P(\lambda)$ is said to be regular if $\det(P(\lambda))$ is not identically zero. For a regular matrix polynomial $P(\lambda)$, the polynomial eigenvalue problem (PEP) seeks to find a scalar λ_0 and a nonzero vector v such that $P(\lambda_0)v = 0$. Equivalently, for a given regular matrix polynomial $P(\lambda)$, $\lambda_0 \in \mathbb{C}$ is an eigenvalue, if $\det P(\lambda_0) = 0$. The nonzero vector $v \in \mathbb{C}^n$ satisfying the equation $P(\lambda_0)v = 0$ is called an eigenvector of $P(\lambda)$ corresponding to an eigenvalue λ_0. Moreover, $\lambda_0 = 0$ is an eigenvalue of $P(\lambda)$ if and only if A_0 is singular. We say ∞ is an eigenvalue of $P(\lambda)$, when 0 is an eigenvalue of the reverse matrix polynomial $\widehat{P}(\lambda) := \lambda^m P(\frac{1}{\lambda}) = A_0\lambda^m + A_1\lambda^{m-1} + \cdots + A_{m-1}\lambda + A_m$. Notice that for an $n \times n$ matrix polynomial of degree $m \geq 1$, there are at most mn number of eigenvalues.

Given an $n \times n$ matrix polynomial $P(\lambda)$ of degree m with nonsingular leading coefficient, one can introduce a monic matrix polynomial corresponding to $P(\lambda)$ as follows: For $i = 0, \ldots, m-1$, let $U_i = A_m^{-1}A_i$. Define a monic matrix polynomial $P_U(\lambda) := I\lambda^m + U_{m-1}\lambda^{m-1} + \cdots + U_1\lambda + U_0$ so that $P(\lambda) = A_m P_U(\lambda)$. The matrix polynomials P and P_U have the same eigenvalues. Moreover, the eigenvalues of P_U are the same as that of the eigenvalues of the corresponding $mn \times mn$ block companion matrix $C := \begin{bmatrix} 0 & I & 0 & \cdots & 0 \\ 0 & 0 & I & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I \\ -U_0 & -U_1 & -U_2 & \cdots & -U_{m-1} \end{bmatrix}$ (see [9] for details). A nonzero vector $v \in \mathbb{C}^n$ is an eigenvector of $P(\lambda)$ corresponding to an eigenvalue λ_0 if and only if the vector $\begin{bmatrix} v \\ \lambda_0 v \\ \vdots \\ \lambda_0^{m-1} v \end{bmatrix} \in \mathbb{C}^{mn}$ is an eigenvector of C corresponding to the eigenvalue λ_0. Note that any eigenvector of C is of this form.

An easy consequence of the finite dimensional spectral theorem (see for instance Theorem 2.5.5, [10]) is that if A and B are two $n \times n$ commuting normal matrices with eigenvalues $\lambda_1, \ldots, \lambda_n$ and $\mu_1, \ldots, \mu_n$, respectively, then

$$\sum_{i=1}^{n} |\lambda_i - \mu_i|^2 = \|A - B\|_F^2. \tag{1.1}$$

What happens when the matrices do not commute has given rise to several interesting questions. One of the classical and well known inequality in matrix analysis is the Hoffman–Wielandt inequality which says that if A and B are two $n \times n$ normal matrices with eigenvalues $\lambda_1, \ldots, \lambda_n$ and $\mu_1, \ldots, \mu_n$, respectively given in some order, then there exists a permutation π of $\{1, \ldots, n\}$ such that

$$\sum_{i=1}^{n} |\lambda_i - \mu_{\pi(i)}|^2 \leq \|A - B\|_F^2 \tag{1.2}$$

(see for instance Theorem 6.3.5, [10]). Various generalizations of the Hoffman–Wielandt inequality by allowing one or both matrices to be non-normal have been studied in the literature (see for instance [1, 2, 5, 6]) and the references cited therein. The most general form assumes that one matrix is diagonalizable while the other remains arbitrary. In [11], the authors provide a concise overview of these generalizations. As in [11], we refer to this generalized form as the Hoffman–Wielandt type inequality, which is stated below:

Theorem 1.1 [11, Theorem 4] *Let A be a diagonalizable matrix of order n and B be an arbitrary matrix of order n, with eigenvalues $\alpha_1, \alpha_2, \ldots, \alpha_n$ and $\beta_1, \beta_2, \ldots, \beta_n$, respectively. Let X be a nonsingular matrix whose columns are eigenvectors of A. Then, there exists a permutation π of the indices $1, 2, \ldots, n$ such that*

$$\sum_{i=1}^{n} |\alpha_i - \beta_{\pi(i)}|^2 \leq \|X\|_2^2 \|X^{-1}\|_2^2 \|A - B\|_F^2. \tag{1.3}$$

The number $\|X\|_2 \|X^{-1}\|_2$ is the spectral condition number of X and is usually denoted by $\kappa(X)$. For a nonsingular matrix $X \in M_n(\mathbb{C})$, $\kappa(X) = \frac{\sigma_{\max}}{\sigma_{\min}}$, the ratio of the largest and the smallest singular values of X.

The purpose of this manuscript is to investigate inequalities (1.2) and (1.3) for block companion matrices of matrix polynomials. Examples illustrating that the inequality (1.2) does not hold in general for block companion matrices of matrix polynomials, even when the coefficients are normal or unitary matrices is presented first (see Remark 2.1). This motivates us in identifying specific classes of matrix polynomials for which the Hoffman–Wielandt type inequality (1.3) holds for the corresponding block companion matrices. As one may observe, inequality (1.3) demands that at least one of the block companion matrices of such matrix polynomials is diagonalizable. With this aim, we prove in Theorem 2.2 that the block companion matrix of linear matrix polynomials whose coefficients are either unitary or diagonal or positive semidefinite matrices is diagonalizable. Theorem 2.2 makes it pertinent to look at similar classes of matrix polynomials of higher degree whose block companion matrices are diagonalizable. The main results of the paper in this connection are Theorems 2.4 and 2.12, which concern quadratic matrix polynomials with commuting unitaries and 2×2 doubly stochastic matrix coefficients, respectively. We also point out in Theorem 2.10 that the block companion matrix of a linear matrix polynomial with 2×2 doubly stochastic matrices are diagonalizable. We establish the Hoffman–Wielandt type inequality for block companion matrices of these classes of matrix polynomials in Theorems 2.6, 2.7 and 2.14. Examples to demonstrate that the block companion matrix is not diagonalizable in general for other classes of matrix polynomials—linear, quadratic and cubic, as well as matrix polynomials not satisfying the assumptions of

Theorems 2.4 and 2.12—(see Remarks 2.3, 2.5, 2.11 and 2.13) are pointed out. The final outcome of the paper concerns estimating the spectral condition number of the matrix X that appears in each of Theorems 2.6, 2.7, and 2.14. We actually prove that bound on $\kappa(X)$ in Theorem 2.6 is independent of matrix polynomials and in Theorem 2.7, we give $\kappa(X)$ in terms of the eigenvalues of the leading coefficient of the matrix polynomial. To the best of our knowledge, the results presented in this manuscript seems new in proving the Hoffman–Wielandt type inequality for certain class of matrix polynomials. The only other source we are aware of in this direction is [12]. In connection to the results in the literature, one can see that if the matrix polynomial is monic and linear with normal or diagonalizable matrix coefficients, then our results reduce to the classical Hoffman–Wielandt and the Hoffman–Wielandt type inequality, respectively, for matrices.

The manuscript is organized as follows. Section 1 is introductory and contains a brief introduction to matrix polynomials that are needed for this manuscript. The main results are presented in Sect. 2, which is further divided into subsections for ease of reading. Each of these subsections is self-explanatory. The Hoffman–Wielandt type inequality is derived for the corresponding block companion matrices of matrix polynomials (with appropriate assumptions) in each of these subsections. In Sect. 2.3, the spectral condition number of a matrix X, which appears in the Hoffman–Wielandt type inequality is estimated. Wherever possible, necessary remarks and examples are provided to justify the assumptions made. Most of the examples were computationally verified using Matlab and SageMath.

Assumption Throughout this manuscript, we only work with matrix polynomials whose leading coefficient is nonsingular.

2 Main results

This section contains the main results of this paper. We first make a remark which gives justification for considering the Hoffman–Wielandt type inequality as against the inequality (1.2) in the context of matrix polynomials.

Remark 2.1 1. As mentioned in the introduction, the inequality (1.2) for block companion matrices of matrix polynomials whose coefficients are normal matrices fails to hold in general. Consider $P(\lambda) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}\lambda + \begin{bmatrix} 2 & 2 \\ 2 & -14 \end{bmatrix}$ and $Q(\lambda) = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{5}{4} \end{bmatrix}\lambda + \begin{bmatrix} 2 & 5 \\ 5 & -\frac{30}{4} \end{bmatrix}$ with the corresponding block companion matrices $C = \begin{bmatrix} -1 & -1 \\ 1 & -7 \end{bmatrix}$ and $D = \begin{bmatrix} -2 & -5 \\ 4 & -6 \end{bmatrix}$, respectively. The eigenvalues of C and D are $\lambda_1 = -4 - 2\sqrt{2}$, $\lambda_2 = -4 + 2\sqrt{2}$ and $\mu_1 = -4 + 4i$, $\mu_2 = -4 - 4i$, respectively. Note that for any permutation π on $\{1, 2\}$, $\sum_{i=1}^{2} |\lambda_i - \mu_{\pi(i)}|^2 = 48$, whereas $\|C - D\|_F^2 = 27$. Therefore, the inequality (1.2) fails to hold in this case. However, one can easily prove that the inequality (1.2) holds for the block companion matrix of linear matrix polynomials whose coefficients are unitary matrices.

2. For higher degree matrix polynomials the inequality (1.2) fails to hold in general even when the coefficients are unitary matrices. For example, let $P(\lambda) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\lambda^2 + \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}\lambda + \begin{bmatrix} \frac{4}{\sqrt{41}} & \frac{5}{\sqrt{41}} \\ \frac{5}{\sqrt{41}} & -\frac{4}{\sqrt{41}} \end{bmatrix}$ and $Q(\lambda) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\lambda^2 + \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}\lambda + \begin{bmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix}$. If C and D are the corresponding block companion matrices of these matrix polynomials, then $\|C - D\|_F^2 = 4$. However, on computing the eigenvalues of C and D, one observes that for any permutation π on $\{1, 2, 3, 4\}$, $\sum_{i=1}^{4} |\lambda_i - \mu_{\pi(i)}|^2 \geq 4.5102 > 4$, where $\{\lambda_i\}$ and $\{\mu_j\}$ are eigenvalues of C and D, respectively. Thus, the inequality (1.2) fails to hold in this case.

It, therefore, makes it pertinent to consider matrix polynomials whose block companion matrices satisfy inequality (1.3). Further, in inequality (1.3), one of the matrices should be diagonalizable. Therefore, in the following subsections, we find classes of matrix polynomials whose block companion matrix is diagonalizable and prove inequality (1.3) for such matrix polynomials.

2.1 Hoffman–Wielandt type inequality for block companion matrices of matrix polynomials with unitary coefficients

We begin with the following theorem which gives the classes of linear matrix polynomials whose block companion matrix is diagonalizable.

Theorem 2.2 *Let $P(\lambda) = A_1\lambda + A_0$ be a linear matrix polynomial. Then, the corresponding block companion matrix C of $P(\lambda)$ is diagonalizable*

(1) *when the coefficients are unitary matrices.*
(2) *when the coefficients are diagonal matrices.*
(3) *when the coefficients are positive (semi)definite matrices.*

Proof In (1) and (2) the block companion matrix is unitary and diagonal, respectively. Hence, it is diagonalizable. In (3), the corresponding block companion matrix is $C = -A_1^{-1}A_0$. Note that $A_1^{-1}A_0 = A_1^{-1/2}\left(A_1^{-1/2}A_0A_1^{-1/2}\right)A_1^{1/2}$. One can verify that $A_1^{-1/2}A_0A_1^{-1/2}$ is a Hermitian matrix and is, therefore, diagonalizable. It now follows that $C = -A_1^{-1}A_0$, being similar to $-A_1^{-1/2}A_0A_1^{-1/2}$, is diagonalizable. □

Remark 2.3 It is not hard to construct linear matrix polynomials whose coefficients are either normal or upper (lower) triangular such that the corresponding block companion matrix C is not diagonalizable. If $P(\lambda)$ is a quadratic matrix polynomial whose coefficients are either (a) diagonal matrices (b) normal matrices (c) upper (lower) triangular matrices or (d) positive (semi)definite, then again, the corresponding block companion matrix need not be diagonalizable. The following matrix polynomial serves as an example for all of the above cases: $P(\lambda) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\lambda^2 + \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\lambda + \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$.

The above remark also implies that if a matrix polynomial has diagonal coefficients, it does not necessarily mean that the corresponding block companion matrix is diagonalizable. Theorem 2.2 and the remark that follows suggest that for deg ≥ 2, matrix polynomials with unitary coefficients might form a good candidate as far as diagonalizability of the block companion matrix is concerned. We discuss this below.

Given a matrix polynomial $P(\lambda) = I\lambda^m + A_{m-1}\lambda^{m-1} + \cdots + A_1\lambda + A_0$, let us consider the following matrix equation $X^m + A_{m-1}X^{m-1} + \cdots + A_1X + A_0 = 0$ with $X \in M_n(\mathbb{C})$. An $n \times n$ matrix X satisfying this equation is called a right solvent of the equation. We just call a right solvent X a solution here. There are at most $\binom{nm}{n}$ solutions to this equation [7]. Solutions $X_1, X_2, \ldots, X_m$ of the above matrix equation are said to be independent if the block Vandermonde matrix, $V := \begin{bmatrix} I & I & \cdots & I \\ X_1 & X_2 & \cdots & X_m \\ \vdots & \vdots & \ddots & \vdots \\ X_1^{m-1} & X_2^{m-1} & \cdots & X_m^{m-1} \end{bmatrix}$ is invertible. In such a case, the corresponding block companion matrix C of $P(\lambda)$ is similar to the block diagonal matrix $\begin{bmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & X_m \end{bmatrix}$ through the block Vandermonde matrix V. Details of these may be found in [4]. Thus, if each of these solutions happen to be diagonalizable, then the block companion matrix C is also diagonalizable. We shall use this in proving Theorem 2.4. Let us observe that it suffices to consider monic matrix polynomials while dealing with commuting unitary coefficients. For, if $P(\lambda) = V_2\lambda^2 + V_1\lambda + V_0$, where the V_i's are commuting unitary matrices, then we can consider the corresponding monic matrix polynomial $P_U(\lambda) = I\lambda^2 + U_1\lambda + U_0$ and observe that the coefficients of P_U are also commuting unitary matrices. With these observations, we have the following theorem.

Theorem 2.4 *Let $P(\lambda) = I\lambda^2 + U_1\lambda + U_0$ be an $n \times n$ matrix polynomial where the coefficients U_0 and U_1 are commuting unitary matrices. Then, the corresponding block companion matrix C of $P(\lambda)$ is diagonalizable.*

Proof The matrices U_0 and U_1 being commuting unitary matrices, there exists an $n \times n$ unitary matrix W such that $WU_1W^* = D_1$ and $WU_0W^* = D_0$ where, D_1, D_0 are diagonal matrices whose diagonal entries are the eigenvalues of U_1 and U_0, respectively. Let $D_1 = \mathrm{diag}(a_{11}, a_{22}, \ldots, a_{nn})$ and $D_0 = \mathrm{diag}(b_{11}, b_{22}, \ldots, b_{nn})$. Let $Q(\lambda) := WP(\lambda)W^* = I\lambda^2 + D_1\lambda + D_0$. The corresponding block companion matrix of $Q(\lambda)$ is $D = \begin{bmatrix} 0 & I \\ -D_0 & -D_1 \end{bmatrix}$. Note that C and D are similar through the unitary matrix $U := W \oplus W$. It, therefore, suffices to prove that the matrix D is diagonalizable. Consider $Q(\lambda) = I\lambda^2 + D_1\lambda + D_0 = \begin{bmatrix} \lambda^2 + a_{11}\lambda + b_{11} & 0 & \cdots & 0 \\ 0 & \lambda^2 + a_{22}\lambda + b_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda^2 + a_{nn}\lambda + b_{nn} \end{bmatrix}$. Let $f_{ii}(\lambda) := \lambda^2 + a_{ii}\lambda + b_{ii}$, $1 \leq i \leq n$. It thus follows that

$$Q(\lambda) = \begin{bmatrix} f_{11}(\lambda) & 0 & \cdots & 0 \\ 0 & f_{22}(\lambda) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f_{nn}(\lambda) \end{bmatrix}.$$ Observe that for each $i,\ 1 \leq i \leq n$, the polynomial $f_{ii}(\lambda)$ has two distinct roots. For otherwise, $f_{ii}(\lambda) = (\lambda - \lambda_0)^2$ for some λ_0, so that $\lambda^2 + a_{ii}\lambda + b_{ii} = \lambda^2 - 2\lambda_0\lambda + \lambda_0^2$. Comparing the coefficients we get, $a_{ii} = -2\lambda_0$ and $b_{ii} = \lambda_0^2$. Taking the modulus we have, $|\lambda_0| = \frac{1}{2}$ and $|\lambda_0| = 1$, which is not possible. This proves the assertion that $f_{ii}(\lambda)$ has two distinct roots. Let $\lambda_i \neq \mu_i$ be two distinct roots of $f_{ii}(\lambda)$ for $1 \leq i \leq n$. Let us now define $X_1 := \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ and $X_2 := \text{diag}(\mu_1, \mu_2, \ldots, \mu_n)$. It is easy to verify that X_1 and X_2 satisfy the quadratic matrix equation $X^2 + D_1 X + D_0 = 0$. Moreover, $\det(X_1 - X_2) = \prod_{i=1}^{n}(\lambda_i - \mu_i) \neq 0$, as $\lambda_i \neq \mu_i$ for all i. Therefore, the block Vandermonde matrix $V = \begin{bmatrix} I & I \\ X_1 & X_2 \end{bmatrix}$ is invertible; in other words, X_1 and X_2 are two independent solutions to the matrix equation $X^2 + D_1 X + D_0 = 0$. It now follows that the block companion matrix $D = \begin{bmatrix} 0 & I \\ -D_0 & -D_1 \end{bmatrix}$ is similar to the block diagonal matrix $\widetilde{D} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix}$ through the block Vandermonde matrix V. Since both X_1 and X_2 are diagonal, $\widetilde{D}$ is diagonal. Hence, D is diagonalizable. □

Few remarks are in order.

Remark 2.5 1. Theorem 2.4 not only proves diagonalizability of the block companion matrix D, but also explicitly gives the invertible block Vandermonde matrix V through which diagonalization happens. Note that the block companion matrix C is then diagonalizable through the matrix $X = UV$, where U is the unitary matrix from the above theorem.

2. Theorem 2.4 does not hold good if the unitary matrices U_0 and U_1 do not commute. For example, when $n = 2$ consider the matrix polynomial $P(\lambda) = I\lambda^2 + U_1\lambda + U_0$ where, $U_0 = \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix}$ and $U_1 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$. Note that U_1 and U_0 are unitary matrices that do not commute. One can easily check that the corresponding block companion matrix C of $P(\lambda)$ has two distinct eigenvalues 1 and -1, each of algebraic multiplicity 2. However, a simple computation reveals that the geometric multiplicity of both these eigenvalues is 1. Hence, C is not diagonalizable.

   For $n = 3$, consider $P(\lambda) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\lambda^2 + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\lambda + \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$ with non-commuting unitary coefficients. The corresponding block companion matrix C of $P(\lambda)$ is not diagonalizable in this case too. Note that one can extend this example and see that the block companion matrix C is not diagonalizable for any size $n \geq 4$ as well.

3. If the degree of matrix polynomial is greater than 2, the corresponding block companion matrix need not be diagonalizable, even if all coefficients are commuting unitary matrices. For example consider $P(\lambda) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \lambda^3 + \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \lambda^2 + \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \lambda + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. The corresponding block companion matrix C of $P(\lambda)$ is not diagonalizable. The argument is the same as in the previous remark.

The above remark justifies the need to consider only quadratic matrix polynomials with commuting unitary coefficients. We now deduce the Hoffman–Wielandt type inequality for quadratic matrix polynomials with unitary coefficients.

Theorem 2.6 *Let P and Q be quadratic matrix polynomials of the same size, where P satisfies the conditions of Theorem 2.4. If C and D are the corresponding block companion matrices, then there exists a permutation π of the indices $1, 2, \ldots, 2n$ such that*

$$\sum_{i=1}^{2n} |\alpha_i - \beta_{\pi(i)}|^2 \leq \|X\|_2^2 \|X^{-1}\|_2^2 \|C - D\|_F^2,$$

where $\{\alpha_i\}$ and $\{\beta_i\}$ are the eigenvalues of C and D, respectively, and X is a nonsingular matrix whose columns are the eigenvectors of C.

Proof The assumptions on P ensure that the matrix C is diagonalizable. The result now follows from Theorem 1.1. □

The Hoffman–Wielandt type inequality for linear matrix polynomials is stated below. We skip the proof as it is similar to the above theorem.

Theorem 2.7 *Let P and Q be linear matrix polynomials of the same size, where P satisfies any of the conditions of Theorem 2.2. If C and D are the corresponding block companion matrices, then there exists a permutation π of the indices $1, 2, \ldots, n$ such that*

$$\sum_{i=1}^{n} |\alpha_i - \beta_{\pi(i)}|^2 \leq \|X\|_2^2 \|X^{-1}\|_2^2 \|C - D\|_F^2,$$

where $\{\alpha_i\}$ and $\{\beta_i\}$ are the eigenvalues of C and D, respectively, and X is a nonsingular matrix whose columns are the eigenvectors of C.

2.2 Hoffman–Wielandt type inequality for block companion matrices of matrix polynomials with doubly stochastic coefficients

We now consider matrix polynomials whose coefficients are doubly stochastic matrices. Recall that a nonnegative square matrix is doubly stochastic if all the row and column sums are 1. A classical result of Birkhoff says that any doubly stochastic matrix is a convex combination of permutation matrices (see Section 8.7 of [10]). Since permutation matrices are unitary, it is natural to ask if the Hoffman–Wielandt type inequality holds for matrix polynomials with doubly stochastic coefficients. We explore this question in this section.

We begin this section with eigenvalue bounds for matrix polynomials whose coefficients are doubly stochastic matrices, with some added assumptions. We state this below and skip the proof as the proof technique is essentially the same as in [3] with the observation that the spectral norm of a doubly stochastic matrix is 1 and the inverse of a permutation matrix is again a permutation matrix. We illustrate via examples the validity of these assumptions.

Theorem 2.8 1. *Let $P(\lambda) = A_m\lambda^m + A_{m-1}\lambda^{m-1} + \cdots + A_1\lambda + A_0$, where A_m, A_0 are $n \times n$ permutation matrices and $A_{m-1}, \ldots, A_1$ are $n \times n$ doubly stochastic matrices. If λ_0 is an eigenvalue of $P(\lambda)$, then $\frac{1}{2} < |\lambda_0| < 2$.*

2. *Let $\mathcal{D} = \left\{ P(\lambda) = \sum_{i=0}^{m} A_i\lambda^i \,:\, A_i$'s are $n \times n$ doubly stochastic matrices, A_m, A_0 are permutation matrices and $m, n \in \mathbb{N} \right\}$ and let $S_{\mathcal{D}} = \{|\lambda_0| : \lambda_0$ is an eigenvalue of $P(\lambda) \in \mathcal{D}\}$. Then $\inf S_{\mathcal{D}} = \frac{1}{2}$ and $\sup S_{\mathcal{D}} = 2$.*

Few remarks are in order.

Remark 2.9 If the leading coefficient or the constant term (or both) is a doubly stochastic matrix, but not a permutation matrix, then the eigenvalues may not necessarily lie in the region $\frac{1}{2} < |\lambda| < 2$. The following examples illustrate this. Let $P(\lambda) = I\lambda^2 + \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{1}{4} \end{bmatrix}\lambda + \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$. One of the eigenvalue is $\frac{3-\sqrt{57}}{12} = -0.3792$, which is less than $\frac{1}{2}$ in absolute value. Similarly if $P(\lambda) = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}\lambda^2 + \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{1}{4} \end{bmatrix}\lambda + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. The eigenvalues of $P(\lambda)$ are $\frac{-1 \pm i\sqrt{3}}{2}$ and $\frac{-3 \pm \sqrt{57}}{4}$ and the absolute value of $\frac{-3-\sqrt{57}}{4}$ is $2.637 > 2$.

We are now in a position to derive diagonalizability of the block companion matrix of a matrix polynomial with doubly stochastic coefficients. We begin with some easy observations.

Theorem 2.10 *Let $P(\lambda) = A_1\lambda + A_0$ be a linear matrix polynomial whose coefficients are 2×2 doubly stochastic matrices. Then, the corresponding block companion matrix is diagonalizable.*

Proof Writing A_0 and A_1 as $A_0 = \begin{bmatrix} b & 1-b \\ 1-b & b \end{bmatrix}$, $A_1 = \begin{bmatrix} a & 1-a \\ 1-a & a \end{bmatrix}$ where, $0 \le a, b \le 1$, we observe that $C = -\begin{bmatrix} \frac{a+b-1}{2a-1} & \frac{a-b}{2a-1} \\ \frac{a-b}{2a-1} & \frac{a+b-1}{2a-1} \end{bmatrix}$, a real symmetric matrix. Hence C is diagonalizable. □

Remark 2.11 Let $P(\lambda) = A_1\lambda + A_0$ be an $n \times n$ matrix polynomial with $n \ge 3$. If one of A_1 or A_0 is a doubly stochastic matrix which is not a permutation matrix, then C

need not be diagonalizable. For example consider $P(\lambda) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \lambda + \begin{bmatrix} \frac{1}{8} & \frac{1}{2} & \frac{3}{8} \\ \frac{1}{4} & \frac{3}{8} & \frac{3}{8} \\ \frac{5}{8} & \frac{1}{8} & \frac{1}{4} \end{bmatrix}$.
We can easily check that C is not diagonalizable.

Let us now prove that the block companion matrix of a quadratic matrix polynomial with 2×2 doubly stochastic coefficients is diagonalizable. We make use of Theorem 2.8 to prove this.

Theorem 2.12 *Let $P(\lambda) = A_2\lambda^2 + A_1\lambda + A_0$ where, A_2, A_0 are 2×2 permutation matrices and A_1 is a 2×2 doubly stochastic matrix. Then, the corresponding block companion matrix C of $P(\lambda)$ is diagonalizable.*

Proof The proof involves three cases.
Case 1: Suppose $A_2 = A_1 = A_0 = I$. Then, $P(\lambda)$ can be written as, $P(\lambda) = \begin{bmatrix} \lambda^2+\lambda+1 & 0 \\ 0 & \lambda^2+\lambda+1 \end{bmatrix}$. Therefore $\frac{-1+i\sqrt{3}}{2}$ and $\frac{-1-i\sqrt{3}}{2}$ are eigenvalues of $P(\lambda)$ and hence of C, of multiplicity 2 each. We can check that $[1, 0]^t$ and $[0, 1]^t$ are eigenvectors of $P(\lambda)$ corresponding to both the eigenvalues $\frac{-1+i\sqrt{3}}{2}$ and $\frac{-1-i\sqrt{3}}{2}$. Therefore, $u_1 = [1, 0, \frac{-1+i\sqrt{3}}{2}, 0]^t$, $u_2 = [0, 1, 0, \frac{-1+i\sqrt{3}}{2}]^t$, $u_3 = [1, 0, \frac{-1-i\sqrt{3}}{2}, 0]^t$ and $u_4 = [0, 1, 0, \frac{-1-i\sqrt{3}}{2}]^t$ are linearly independent eigenvectors of C. This proves diagonalizability of C.
Case 2: If $A_2 = A_1 = A_0 = I'$, where $I' = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Then, the monic matrix polynomial corresponding to $P(\lambda)$ is $P_U(\lambda) = I\lambda^2 + I\lambda + I$. Hence by Case 1, C is diagonalizable.
Case 3: Consider the corresponding monic matrix polynomial, $P_U(\lambda) = I\lambda^2 + B_1\lambda + B_0$ where, $B_1 = A_2^{-1}A_1$ is a doubly stochastic matrix and $B_0 = A_2^{-1}A_0$ is a permutation matrix. Let $B_1 = \begin{bmatrix} a & 1-a \\ 1-a & a \end{bmatrix}$ and $B_0 = \begin{bmatrix} b & 1-b \\ 1-b & b \end{bmatrix}$ where, $0 \le a, b \le 1$. Then, $P_U(\lambda) = \begin{bmatrix} \lambda^2 + a\lambda + b & (1-a)\lambda + (1-b) \\ (1-a)\lambda + (1-b) & \lambda^2 + a\lambda + b \end{bmatrix}$ and $\det P_U(\lambda) = (\lambda^2 + \lambda + 1)(\lambda^2 + (2a-1)\lambda + (2b-1))$. Note that $\lambda^2 + \lambda + 1 \neq \lambda^2 + (2a-1)\lambda + (2b-1)$. Otherwise $a = b = 1$ which then will imply that $A_2 = A_1 = A_0 = I$ or $A_2 = A_1 = A_0 = I'$. Moreover, since both $\lambda^2 + \lambda + 1$ and $\lambda^2 + (2a-1)\lambda + (2b-1)$ are real polynomials they do not have common roots. Now, we claim that $\lambda^2 + (2a-1)\lambda + (2b-1)$ has two distinct roots. Suppose there is only one root, say, λ_0. Then, we have $\lambda^2 + (2a-1)\lambda + (2b-1) = (\lambda - \lambda_0)^2 = \lambda^2 - 2\lambda_0\lambda + \lambda_0^2$. Comparing the coefficients, we get $2a - 1 = -2\lambda_0$. Since $0 \le a \le 1$, we have $-1 \le 2a - 1 \le 1$. This implies $2|\lambda_0| = |2a - 1| \le 1$. Therefore, $|\lambda_0| \le \frac{1}{2}$, a contradiction to Theorem 2.8. Thus $\lambda^2 + (2a-1)\lambda + (2b-1)$ has two distinct roots. Since $\lambda^2 + \lambda + 1$ also has two distinct roots, $P(\lambda)$ and hence C has four distinct eigenvalues. Hence, C is diagonalizable. □

The following remark justifies the assumptions made in the above theorem.

Remark 2.13 1. In the above theorem, if the leading coefficient or the constant term (or both) is a doubly stochastic matrix, but not a permutation matrix, then the corresponding block companion matrix need not be diagonalizable. For example, consider $P(\lambda) = \begin{bmatrix} \frac{11}{24} & \frac{13}{24} \\ \frac{13}{24} & \frac{11}{24} \end{bmatrix} \lambda^2 + \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{1}{4} \end{bmatrix} \lambda + \begin{bmatrix} \frac{1}{8} & \frac{7}{8} \\ \frac{7}{8} & \frac{1}{8} \end{bmatrix}$. We can check that the corresponding block companion matrix is not diagonalizable. One can also look at the matrix polynomial $P(\lambda) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \lambda^2 + \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \lambda + \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$.

2. If the size of a quadratic matrix polynomial $P(\lambda)$ is greater than 2 then the corresponding block companion matrix C need not be diagonalizable even when all the coefficients are permutation matrices (see the Example in Remark 2.5). Note that the coefficients of the matrix polynomial in that example are non-commuting permutation matrices. However, when the coefficients of $P(\lambda)$ are commuting permutation matrices the corresponding block companion matrix C is diagonalizable, as already proved in Theorem 2.4.
3. Let $P(\lambda) = A_2\lambda^2 + A_1\lambda + A_0$ be an $n \times n$ matrix polynomial with $n \geq 3$. If one of A_2, A_1, A_0 is a doubly stochastic matrix which is not permutation, then the corresponding block companion matrix C need not be diagonalizable even if the coefficients commute. For example, consider $P(\lambda) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \lambda^2 + \begin{bmatrix} \frac{5}{12} & \frac{5}{12} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \lambda + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. We can check that the corresponding block companion matrix of $P(\lambda)$ is not diagonalizable.
4. Let $P(\lambda)$ be a matrix polynomial of degree greater than 2. Then, the corresponding block companion matrix need not be diagonalizable even with coefficients of $P(\lambda)$ being commuting permutation matrices. For example consider $P(\lambda) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \lambda^3 + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \lambda^2 + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \lambda + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. The coefficients are commuting permutation matrices. However, the corresponding block companion matrix is non-diagonalizable.

We end this section by pointing out that the Hoffman–Wielandt type inequality for matrix polynomials with doubly stochastic coefficients can be derived as in the unitary case. For the sake of completeness, we state below only the quadratic polynomials version and skip the proof.

Theorem 2.14 *Let P and Q be quadratic matrix polynomials of the same size, where P satisfies conditions of Theorem 2.12. If C and D are the corresponding block companion matrices, then there exists a permutation π of the indices $1, \ldots, 4$ such that*

$$\sum_{i=1}^{4} |\alpha_i - \beta_{\pi(i)}|^2 \leq \|X\|_2^2 \|X^{-1}\|_2^2 \|C - D\|_F^2,$$

where $\{\alpha_i\}$ and $\{\beta_i\}$ are the eigenvalues of C and D, respectively, and X is a nonsingular matrix whose columns are the eigenvectors of C.

2.3 Estimation of the spectral condition number

A diagonalizable matrix can be diagonalized through more than one matrix and the spectral condition number of these matrices may not be identical. We also know from Theorem 2.4 that diagonalization of the block companion matrix of matrix polynomials with commuting unitary coefficients is achieved through a particular block Vandermonde matrix V. On the other hand, in Theorem 2.12, diagonalization is achieved more directly. This gives us some hope in estimating the spectral condition number in Theorems 2.6 and 2.14. We set out to do this in this section.

2.3.1 Condition number of matrix *X* that appears in Theorem 2.6

We first prove that the spectral condition number of the block Vandermonde matrix V obtained in Theorem 2.4 is less than 2.

Theorem 2.15 *Let V be the block Vandermonde matrix obtained in Theorem 2.4. Then, $\kappa(V) < 2$.*

Proof By Theorem 2.4, we have $V = \begin{bmatrix} I & I \\ X_1 & X_2 \end{bmatrix}$ with $X_1 = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ and $X_2 = \operatorname{diag}(\mu_1 \ldots, \mu_n)$, where λ_i and μ_i are the distinct roots of $f_{ii}(\lambda) = \lambda^2 + a_{ii}\lambda + b_{ii}$ for $i = 1, \ldots, n$ and a_{ii}, b_{ii} are the eigenvalues of U_1, U_0, respectively. We thus have $|\lambda_i + \mu_i| = 1$ and $|\lambda_i \mu_i| = 1$ for $i = 1, \ldots, n$. It is then easy to show that $(\lambda_i - \mu_i)^2 = a_{ii}^2 - 4b_{ii}$. Thus, $|\lambda_i - \mu_i|^2 = |a_{ii}^2 - 4b_{ii}| \geq \| a_{ii}^2| - 4|b_{ii}\| = |1 - 4| = 3$. Using the parallelogram identity, we have $2(|\lambda_i|^2 + |\mu_i|^2) = |\lambda_i + \mu_i|^2 + |\lambda_i - \mu_i|^2 \geq 1 + 3 = 4$. This implies that $|\lambda_i|^2 + |\mu_i|^2 \geq 2$. Since $\big||\lambda_i| - |\mu_i|\big| \leq |\lambda_i + \mu_i| = 1$, we have $1 \geq (|\lambda_i| - |\mu_i|)^2 = |\lambda_i|^2 + |\mu_i|^2 - 2|\lambda_i||\mu_i|$. Therefore, $|\lambda_i|^2 + |\mu_i|^2 \leq 3$. We thus have proved the following:

$$2 \leq |\lambda_i|^2 + |\mu_i|^2 \leq 3, \quad i = 1, \ldots, n. \tag{2.1}$$

We know that $\kappa(V) = \frac{\sigma_{\max}}{\sigma_{\min}}$, where $\sigma_{\min}$ and $\sigma_{\max}$ are the smallest and the largest singular values of V, respectively. Since the singular values of V are the positive square roots of the eigenvalues of VV^*, we estimate bounds for these eigenvalues. Define $L := \begin{bmatrix} I & 0 \\ -(X_1 + X_2)(2I - I\lambda)^{-1} & I \end{bmatrix}$. Notice that L is a matrix with $\det(L) = 1$. Let us now compute the matrix $L(VV^* - I\lambda) = \begin{bmatrix} 2I - I\lambda & X_1^* + X_2^* \\ 0 & -(X_1 + X_2)(2I - I\lambda)^{-1}(X_1^* + X_2^*) + X_1X_1^* + X_2X_2^* - I\lambda \end{bmatrix}$. Since X_1 and X_2 are diagonal matrices, $\det(VV^* - I\lambda) = \det\big(L(VV^* - I\lambda)\big) = \det(2I - I\lambda)\big(- (X_1 + X_2)(2I - I\lambda)^{-1}(X_1^* + X_2^*) + (X_1X_1^* + X_2X_2^* - I\lambda)\big) = \det\big(I\lambda^2 - (2 + X_1X_1^* + X_2X_2^*)\lambda + (X_1X_1^* + X_2X_2^* - X_1X_2^* - X_2X_1^*)\big) = \prod_{i=1}^{n}\big(\lambda^2 - (2 + |\lambda_i|^2 + |\mu_i|^2)\lambda + (|\lambda_i|^2 + |\mu_i|^2 - \lambda_i\bar{\mu}_i - \mu_i\bar{\lambda}_i)\big)$. Thus, in order to compute the eigenvalues of VV^*, it suffices to determine the roots of the polynomials $\lambda^2 - (2 + |\lambda_i|^2 + |\mu_i|^2)\lambda + (|\lambda_i|^2 + |\mu_i|^2 - \lambda_i\bar{\mu}_i - \mu_i\bar{\lambda}_i)$. These are given by $\alpha_i =$

$\frac{(2+|\lambda_i|^2+|\mu_i|^2)-\sqrt{(2-(|\lambda_i|^2+|\mu_i|^2))^2+4}}{2}$ and $\beta_i = \frac{(2+|\lambda_i|^2+|\mu_i|^2)+\sqrt{(2-(|\lambda_i|^2+|\mu_i|^2))^2+4}}{2}$, for $i = 1, \ldots, n$. Note that $\alpha_i \leq \beta_i$ as the term inside the square root symbol is positive. We now claim that $1 \leq \alpha_i \leq \beta_i < 4$ for all $i = 1, \ldots, n$. Suppose on the contrary, $\alpha_i < 1$ for some $i = 1, \ldots, n$. Then we have,

$$\frac{(2+|\lambda_i|^2+|\mu_i|^2)-\sqrt{(2-(|\lambda_i|^2+|\mu_i|^2))^2+4}}{2} < 1$$
$$\implies (2+|\lambda_i|^2+|\mu_i|^2)-\sqrt{(2-(|\lambda_i|^2+|\mu_i|^2))^2+4} < 2$$
$$\implies (|\lambda_i|^2+|\mu_i|^2)^2 < (2-(|\lambda_i|^2+|\mu_i|^2))^2+4$$
$$\implies (|\lambda_i|^2+|\mu_i|^2)^2 < 4+(|\lambda_i|^2+|\mu_i|^2)^2-4(|\lambda_i|^2+|\mu_i|^2)+4$$
$$\implies |\lambda_i|^2+|\mu_i|^2 < 2,$$

a contradiction to the inequality (2.1). Therefore, we have $1 \leq \alpha_i$ for all $i = 1, \ldots, n$. Similarly, if $4 \leq \beta_i$ for some $i = 1, \ldots, n$, then we have $3 < |\lambda_i|^2 + |\mu_i|^2$ which is again a contradiction to the inequality (2.1). Therefore, $\beta_i < 4$ for all $i = 1, \ldots, n$. We thus have $1 \leq \alpha_i \leq \beta_i < 4$, which implies that $1 \leq \sigma_{\min}$ and $\sigma_{\max} < 2$. Therefore, $\kappa(V) = \frac{\sigma_{\max}}{\sigma_{\min}} < 2$. □

As mentioned in the Remark 2.5, the matrix which diagonalizes the block companion matrix C is $X = UV$, where U is a unitary matrix. Since the spectral condition number is unitarily invariant, we have $\kappa(X) = \kappa(V) < 2$. Thus, in Theorem 2.6, we have $\|X\|_2^2\|X^{-1}\|_2^2 < 4$.

2.3.2 Condition number of matrix *X* obtained in Theorem 2.7

In parts (1) and (2) of Theorem 2.2, the block companion matrices are unitary and diagonal matrices, respectively. Hence, both are diagonalizable through unitary matrices, whose spectral condition number is 1.

In part (3) of Theorem 2.2, the block companion matrix is $C = -A_1^{-1}A_0 = A_1^{-1/2}\left(-A_1^{-1/2}A_0A_1^{-1/2}\right)A_1^{1/2}$, where $-A_1^{-1/2}A_0A_1^{-1/2}$ is a Hermitian matrix, which is diagonalizable through a unitary matrix, say, U. Hence, $C = A_1^{-1/2}U^{-1}DUA_1^{1/2}$, where D is a diagonal matrix. Thus, C is diagonalizable through matrix $X = UA_1^{1/2}$, where $A_1^{1/2}$ is a Hermitian positive definite matrix. Thus, $\kappa(X) = \kappa(A_1^{1/2})$. Since $A_1^{1/2}$ is Hermitian positive definite, $\kappa(A_1^{1/2}) = \frac{\lambda_{\max}}{\lambda_{\min}}$, where $\lambda_{\max}$ and $\lambda_{\min}$ are, respectively, the maximum and the minimum eigenvalues of $A_1^{1/2}$. Therefore, $\kappa(X) = \frac{\lambda_{\max}}{\lambda_{\min}}$.

2.3.3 Condition number of matrix *X* obtained in Theorem 2.14

We again discuss two cases with reference to Theorem 2.12.

(1) If $A_2 = A_1 = A_0 = I$ or $A_2 = A_1 = A_0 = I'$ where I is the identity matrix and $I' = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, the coefficients are unitary matrices and this case reduces to the one discussed above. Thus, $\kappa(X) < 2$.

(2) In the general case, since the coefficients are of size 2×2 one can verify that $v_1 = \left[\frac{-1+i\sqrt{3}}{2}\ \frac{-1+i\sqrt{3}}{2}\ 1\ 1\right]^t$, $v_2 = \left[\frac{2i}{\sqrt{3}+i}\ \frac{2i}{\sqrt{3}+i}\ 1\ 1\right]^t$, $v_3 = \left[\frac{2}{(2a-1)+\sqrt{4a^2-4a-8b+5}}\ -\frac{2}{(2a-1)+\sqrt{4a^2-4a-8b+5}}\ -1\ 1\right]^t$ and $v_4 = \left[\frac{2}{(2a-1)-\sqrt{4a^2-4a-8b+5}}\ \frac{2}{(1-2a)+\sqrt{4a^2-4a-8b+5}}\ -1\ 1\right]^t$ are linearly independent eigenvectors of the block companion matrix C. We can thus choose $X = \left[v_1\ v_2\ v_3\ v_4\right]$.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s43036-023-00292-8.

Acknowledgements The authors are grateful to the referees for their valuable comments and suggestions. Pallavi Basavaraju and Shrinath Hadimani acknowledge the Council of Scientific and Industrial Research (CSIR) and the University Grants Commission (UGC), Government of India, for financial support through research fellowships.

Data availability Not applicable.

Declarations

Conflict of interest The authors declare that they have no competing interest.

References

1. Bhatia, R.: Perturbation Bounds for Matrix Eigenvalues, 2nd edn. SIAM, Philadelphia (2007)
2. Bhatia, R., Elsner, L.: The Hoffman–Wielandt inequality in infinite dimensions. Proc. Indian Acad. Sci. Math. Sci. **104**(3), 483–494 (1994)
3. Cameron, T.R.: Spectral bounds for matrix polynomials with unitary coefficients. Electron. J. Linear Algebra **30**, 585–591 (2015)
4. Connes, A., Schwarz, A.: Matrix Vieta theorem revisited. Lett. Math. Phys. **39**(4), 349–353 (1997)
5. Elsner, L.: A note on the Hoffman–Wielandt theorem. Linear Algebra Appl. **182**, 235–237 (1993)
6. Elsner, L., Friedland, S.: Singular values, doubly stochastic matrices and applications. Linear Algebra Appl. **220**, 161–169 (1995)
7. Fuchs, D., Schwarz, A.: Matrix Vieta theorem. In: Lie Groups and Lie Algebras: E. B. Dynkin's Seminar. American Mathematical Society. Transl., Ser. 2, Am. Math. Soc., Providence, RI **169**, pp. 15–22 (1995)
8. Gohberg, I., Lancaster, P., Rodman, L.: Matrix Polynomials, 2nd edn. SIAM, Philadelphia (2009)
9. Higham, N.J., Tisseur, F.: Bounds for eigenvalues of matrix polynomials. Linear Algebra Appl. **358**, 5–22 (2003)
10. Horn, R.A., Johnson, C.R.: Matrix Analysis, 2nd edn. Cambridge University Press, Cambridge (2012)
11. Ikramov, Kh.D., Nesterenko, Yu.R.: Theorems of the Hoffman–Wielandt type for coneigenvalues of complex matrices. Dokl. Math. **80**(1), 536–540 (2009)
12. Le, C.-T.: On Wielandt–Mirsky's conjecture for matrix polynomials. Bull. Korean Math. Soc. **56**(5), 1273–1283 (2019)

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.