THE CONDITIONING AND BACKWARD ERROR OF THE MACKEY PENCIL

ASHLEY MAILLOUX AND LUCAS MEDINA
ADVISOR: MARIBEL BUENO

Abstract. The standard way of solving the PEP (polynomial eigenvalue problem) $P(\lambda)x = 0$, where $P(\lambda)$ is the matrix polynomial $A_k\lambda^k + \cdots + A_1\lambda + A_0$ and $A_0, \ldots, A_k$ are $n \times n$ matrices, is by using linearizations (degree 1 matrix polynomials that share the eigenstructure of $P(\lambda)$). There are efficient algorithms, like the QZ method, used to solve the PEP for the linearizations. When choosing a linearization, it is convenient that it has similar conditioning and backward error to $P(\lambda)$ while also preserving its structural properties like symmetries in the coefficients. In our work, we study the Mackey pencil $T(\lambda)$, which is block-symmetric. In the literature, the pencils called D_1 and D_k belonging to the vector space $\mathbb{DL}(P)$, were shown to be optimal in regards to the conditioning and backward error associated with an eigenpair (δ, x), when $|\delta| \geq 1$ and $|\delta| \leq 1$, respectively. Moreover, D_1 and D_k are linearizations only if A_0 and A_k are nonsingular, respectively. Given a simple, finite, nonzero eigenvalue of $P(\lambda)$, we provide explicit expressions for the condition number and backward error of an eigenpair (δ, x) for both $P(\lambda)$ and $T(\lambda)$ as well as bounds on the quotients of these numbers. Unlike D_1 and D_k, the numerical behavior of the Mackey pencil does not depend on the modulus of δ and does not require A_k or A_0 to be nonsingular for $T(\lambda)$ to be a linearization of $P(\lambda)$. The only drawback is that this pencil is only available for k odd. We will provide results from numerical experiments.

1. Introduction

Let $P(\lambda)$ be the $n \times n$ matrix polynomial of degree k

$$P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i, \qquad A_i \in \mathbb{C}^{nxn}.,$$

In this paper, we will be assuming that P is regular, that is the $det(P(\lambda)) \neq 0$. If x and y are nonzero vectors such that they are a solution to the polynomial eigenvalue problem (PEP), $P(\lambda)x = 0$ and $y * P(\lambda) = 0$, then we will say that x and y are a right and left eigenvector, respectively. The scalar λ is called a (finite) eigenvalue of $P(\lambda)$.

The standard method of solving the PEP is to use a matrix pencil, a matrix polynomial of degree 1 that is a linearization of $P(\lambda)$. We say that $L(\lambda) = \lambda L_1 + L_0$ is a linearization of $P(\lambda)$ if there exists two unimodular matrices, $F(\lambda)$ and $G(\lambda)$ such that

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

When the degree of $P(\lambda)$ is larger than 1, the PEP becomes difficult to solve numerically. Choosing a linearization of P proves to be an easier way to solve the PEP since there are efficient algorithms for the generalized eigenvalue problem. The new problem that we are solving is

$$L(\lambda)z = 0,$$

where z is a right eigenvector of L.

When $P(\lambda)$ is structured (Hermitian, Skew-Hermitian, Palindromic, ...), it is convenient to use linearizations that preserve the structural properties of P. For example, if P is hermitian, meaning that $A_i = A_i^*$ for $i = 0 : k$, we want a linearization $L(\lambda) = \lambda L_1 + L_0$ such that $L_i^* = L_i$, $i = 0 : 1$. Note that A^* denotes the complex conjugate transpose of the matrix A.

Also, when applying an algorithm to solve the eigenvalue problem for $L(\lambda)$ we would like the linearization to be comparable to $P(\lambda)$ in regards to other properties, such as conditioning and backward error. When choosing a linearization, it is not desirable to choose one that has worse conditioing than $P(\lambda)$. Also, the standard algorithm to compute the eigenvalue problem for L, that is the QZ method, produces small backward errors. Thus, we want to choose a linearization whose backward error is approximately the same as that of P. Linearizations with these properties are considered optimal for our purposes.

In 2006, a group of mathematicians created a vector space which they named $\mathbb{DL}(P)$. The standard basis for this vector space contains k pencils, denoted by $D_1, \ldots D_k$. Extensive research has been done on the pencils in this space regarding conditioning and backward error. In [4], it was shown that there are two optimal linearizations (in regards to conditioning and backward error) in $\mathbb{DL}(P)$: D_1 and D_k. D_1 and D_k are optimal when $|\delta| \geq 1$ and $|\delta| \leq 1$, respectively and D_1 and D_k are linearizations of P only if A_0 and A_K are nonsingular, respectively.

The goal of our project is to find a block-symmetric pencil such that it is a linearization that behaves comparably to $P(\lambda)$ in regards to conditioning and backward error and that does not require nonsingularity. Furthermore, we worked on finding a pencil that behaved optimally regardless of the eigenvalues of any $P(\lambda)$.

Through research, we were able to find such a pencil, but only for matrix polynomials of odd degrees. We constructed bounds on the backward error for this pencil and present the proof of these bounds in our paper. Also, we conducted numerical experiments examining the conditioning and backward error of our pencil compared to other pencils.

The paper is structured as follows: In Section 2, we present relevant background information with examples to motivate our basic concepts. In Section 3, we discuss important pencils in literature. In Section 4, we describe the scaling technique that we used in our numerical experiments and in section 5, we present our numerical experiments and end our paper with our conclusions and future work.

2. Background

We will begin with defining commonly used terms throughout our paper and these concepts will motivate our results that are mentioned in later sections. This information has been drawn from several publications, all of which are referenced accordingly.

It is important to recall that the 2-norm of a matrix is the maximum of the norm of Ax where x ranges over the unit sphere. Also note that in our paper, $\otimes$ refers to the Kronecker product.

2.1. **Forward Error.** When using any algorithm to solve the PEP, it will produce some forward error. The relative forward error is

$$\frac{|\hat{y} - y|}{|y|}$$

where $\hat{y}$ is the computed value for y.

It happens to be that for our problem, the forward error is bounded above by the conditioning and the backward error of the problem. Specifically,

$$\text{forward error } \leq \text{condition number } \times \text{ backward error.}$$

2.2. **Conditioning.** To think of conditioning in a general sense, let us consider a polynomial of degree 2.

For example, let us examine

$$\begin{aligned} x^2 - 2x + 1 &= 0 \\ (x-1)^2 &= 0 \\ x &= 1, \text{(double root)} \end{aligned}$$

When using a computer to compute solutions, there will be errors due to rounding and truncation errors. So next we examine what happens when one of the coefficients change slightly. Let us change the constant term to .9999. The problem that we are now solving for is $x^2 - 2x + .9999 = 0$

Thus,

$$\begin{aligned} x^2 - 2x + .9999 &= 0 \\ (x - .99)(x - 1.01) &= 0 \\ x &= .99, 1.01 \end{aligned}$$

Notice that a small change in the constant term, exactly a change of 1.0×10^{-4} resulted in a change in the roots by1.0×10^{-2}. So a slight change in the input caused a much larger change in the output.

We can classify the way a problem is conditioned in one of two ways. A problem is said to be well-conditioned or ill-conditioned. A problem is well-conditioned if a small perturbation in the input results in a small change in the output. Ill-conditioned problems are those in which a small perturbation in the input results in a large change in the output. Thus, it is easy to see that the conditioning of a problem deals with the senstivity of the problem, not the algorithm. The words small and large are obviously subjective, but the value of the condition number can be classified as either depending on the problem at hand. For certain problems, a condition number could be said to be well-conditioned, but not for others.

For this specific problem, we will provide the roots for any change in the constant term in order to classify the problem in regards to conditioning.

$$x^2 - 2x + (1 + \Delta) = 0$$

Using the quadratic formula,

$$\begin{aligned}\delta + \Delta\delta &= \frac{2 \pm \sqrt{4 - 4(1+\Delta)}}{2} \\ &= \frac{2 \pm \sqrt{-4\Delta}}{2} \\ &= 1 \pm \sqrt{-\Delta}\end{aligned}$$

Thus,

$$\lim_{\Delta \to 0} \frac{\frac{|\delta_x|}{|x|}}{\frac{|\delta|}{|a|}} = \frac{\frac{|\sqrt{\Delta}|}{|\delta|}}{\frac{|\Delta|}{|a|}}$$

$$\begin{aligned}&= \frac{|a|}{|\delta|} \cdot \frac{\sqrt{\Delta}}{|\Delta|} \\ &= \infty\end{aligned}$$

It easily shown that a small change in input results in a drastic change in the output. Thus, we can classify the example above as being an ill-conditioned problem.

For our problem, the explicit expression for the condition number is

$$\kappa_P(\lambda) = \frac{\left(\sum_{n=0}^{k} |\lambda|^i \left\|A_i\right\|_2\right) \left\|y\right\|_2 \left\|x\right\|_2}{|\lambda||y^* P'(\lambda)x|},$$

where x and y are right and left eigenvectors and λ is a simple, finite, non-zero eigenvalue. P' is the entry-wise derivative of P.

2.3. **Backward Error.** Unlike conditioing, the backward error deals with an error in the algorithm itself. When the algorithm computes the solution, it is not the exact solution for the problem given. The backward error looks to find the exact problem that yields the computed solution. If we have a problem, $f(a) = b$, the algorithm produces the solution $\hat{b}$. We must find the exact problem that yields $\hat{b}$. That is,

$$f(a + \Delta a) = \hat{b}.$$

In a more formal setting, the normwise backward error [1] of an approximate eigenpair$(\hat{x}, \hat{\lambda})$ of $P(\lambda)$, where $\hat{\lambda}$ is finite is defined as

$$\eta_P(\hat{x}, \hat{\lambda}) = \min\{\epsilon : (P(\hat{\lambda}) + \Delta P(\hat{\lambda}))\hat{x} = 0, \left\|\Delta A_i\right\|_2 \le \epsilon \left\|A_i\right\|_2, i = 0 : m\}$$

where $\Delta P(\hat{\lambda}) = \sum_{i=0}^{m} \hat{\lambda}^i \Delta A_i$.

For our research, we use the explicit formula for backward error that was constructed by Tisseur[1] as

$$\eta_P(\hat{x}, \hat{\lambda}) = \frac{\left\|P(\hat{\lambda})\hat{x}\right\|_2}{\left(\sum_{i=0}^{m} |\hat{\lambda}^i| \left\|A_i\right\|_2\right) \left\|\hat{x}\right\|_2}$$

where $(\hat{x}, \hat{\lambda})$ is a computed eigenpair for P.

3. Noteworthy Pencils in Literature

3.1. The Vector Space $\mathbb{DL}(P)$.

Recall the vector space $\mathbb{DL}(P)$ that we mentioned briefly in the introduction. There are two vector spaces, $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ in which almost all of the pencils belonging to these vector spaces are linearizations for the matrix polynomial $P(\lambda)$. For linearizations that are in $\mathbb{L}_1(P)$, it has been shown that it easy to recover the right eigenvectors of the matrix polynomial from the right eigenvectors of the linearization. Similarly, the left eigenvectors for $P(\lambda)$ can be easily recovered from the left eigenvectors of a linearization in $\mathbb{L}_2(P)$. Mackey, Mackey, Mehl, and Mehrmann [4] constructed $\mathbb{DL}(P)$ from the intersection of these vector spaces.

Definition 3.1. For any matrix polynomial $P(\lambda)$ of size $n \times n$, the double ansatz space of P is

$$\mathbb{DL}(P) := \mathbb{L}_1(P) \cap \mathbb{L}_2(P),$$

which is the set of $nk \times nk$ pencils $L(\lambda)$ that simultaneously satisfy

$$L(\lambda) \cdot (\Lambda \otimes I) = v \otimes P(\lambda) \textit{ for some } v \in \mathbb{F}^k$$

$$(\Lambda^T \otimes I) \cdot L(\lambda) = w^T \otimes P(\lambda) \textit{ for some } w \in \mathbb{F}^k,$$

which are referred to as a "right" and "left ansatz," respectively.

Recall that almost all of the pencils in this vector space are linearizations for $P(\lambda)$ and that the standard basis for this vector space contains k pencils, denoted by $D_1, \dots D_k$.

The optimal linearizations in regards to conditioning and backward error in $\mathbb{DL}(P)$ are D_1 and D_k. D_1 and D_k are optimal when $|\delta| \geq 1$ and $|\delta| \leq 1$, respectively.

The bounds for the conditioning of the pencils D_1 and D_k provided below were calculated by Bueno, Dopico, and Furtado [2].

Theorem 1. *Let $P(\lambda)$ be a regular matrix polynomial of degree k with $A_0 \neq 0$. Assume that δ is a simple, finite, nonzero eigenvalue of $P(\lambda)$. Let $\rho = \frac{\max_{i=0:k}\{\|A_i\|_2\}}{\min\{\|A_k\|_2, \|A_0\|_2\}}$ and consider the natural weights for $D_1(\lambda, P)$ and $P(\lambda)$.*

- *If A_0 is nonsingular and $|\delta| \geq 1$, then*
$$\frac{1}{k\rho} \leq \frac{\kappa_{D_1}(\delta)}{\kappa_P(\delta)} \leq 2k^2\rho.$$
- *If A_k is nonsingular and $|\delta| \leq 1$, then*
$$\frac{1}{k\rho} \leq \frac{\kappa_{D_k}(\delta)}{\kappa_P(\delta)} \leq 2k^2\rho.$$

The backward error bounds were constructed by Higham, Li, and Tisseur for the homogenous form of $P(\lambda)$. The homogenous form allows for λ to be infinite or zero. In Section 4, we will define a scaling technique that we used that preserves the conditioning and backward error. If λ were to be infinite or 0, the scaling would no longer preserve the conditioning or backward error. Thus, in our research we only looked at finite, simple, non-zero eigenvalues.

The following two theorems are for upper bounds of the backward error for any $L \in \mathbb{DL}(P)$. We only were able to look at the right eigenvectors for P and L, so computations in the future must be done for the left eigenvectors as well.

Theorem 2. *Let $P(\lambda)$ be a matrix polynomial and let $L(\lambda) \in \mathbb{DL}(P)$ be a linearization for P with ansatz vector v. Let z be a right eigenvector of L and let $x = (v^T \otimes I_n)z$. Let η_P and η_L be the backward error for the matrix polynomial and $L \in \mathbb{DL}(P)$. The ratio $\frac{\eta_P}{\eta_L}$ is bounded above such that*

$$\frac{\eta_P}{\eta_L} \leq kr^{1/2}\rho\frac{|\lambda|+1}{|\lambda|^k+1}\frac{\|z\|_2}{\|x\|_2}$$

where $\rho = \frac{\max\{\|A_i\|_2\}}{\min\{\|A_k\|_2, \|A_0\|_2\}}$.

Proof. Since $L(\lambda)z = v \otimes P(\lambda)x$, it follows that

$$\|v\|_2 \|P(\lambda)x\|_2 = \|v \otimes P(\lambda)x\|_2 = \|L(\lambda)z\|_2 \leq \|L(\lambda)\|_2 \|z\|_2$$

Thus,

$$\|P(\lambda)x\|_2 \leq \frac{\|L(\lambda)\|_2 \|z\|_2}{\|v\|_2}.$$

By multiplying both denominators by $(\sum_{i=0}^k |\lambda|^i \|A_i\|_2) \|x\|_2$, it follows that

$$\frac{\|P(\lambda)x\|_2}{(\sum_{i=0}^k |\lambda|^i \|A_i\|_2) \|x\|_2} \leq \frac{\|L(\lambda)z\|_2}{(\sum_{i=0}^k |\lambda|^i \|A_i\|_2) \|x\|_2 \|v\|_2}.$$

$$\begin{aligned}\frac{\eta_P}{\eta_L} &= \frac{\|P(\lambda)x\|_2}{(\sum_{i=0}^k |\lambda|^i \|A_i\|_2) \|x\|_2}\frac{(|\lambda| \|L_1\|_2 + \|L_0\|_2) \|z\|_2}{\|L(\lambda)z\|_2} \\ &\leq \frac{(|\lambda| \|L_1\|_2 + \|L_0\|_2) \|z\|_2}{(\sum_{i=0}^k |\lambda|^i \|A_i\|_2) \|x\|_2 \|v\|_2} \\ &\leq \frac{(|\lambda|+1)\max\{\|L_1\|_2, \|L_0\|_2\} \|z\|_2}{(\sum_{i=0}^k |\lambda|^i \|A_i\|_2) \|x\|_2 \|v\|_2}.\end{aligned}$$

where the last inequality follows because

$$|\lambda| \|L_1\| + \|L_0\| \leq |\lambda| \max\{\|L_1\|_2, \|L_0\|_2\} + \max\{\|L_1\|_2, \|L_0\|_2\} = (|\lambda|+1)\max\{\|L_1\|_2, \|L_0\|_2\}.$$

In [2, theorem 3.20] shows us

$$\max\{\|L_1\|_2, \|L_0\|_2\} \leq kr^{1/2} \max \|A_i\|_2 \|v\|_2$$

$$\frac{\eta_P(x,\lambda)}{\eta_L(z,\lambda)} \leq \frac{(|\lambda|+1)\max\{\|L_1\|_2, \|L_0\|_2\}\|z\|_2}{(\sum_{i=0}^k |\lambda|^i \|A_i\|_2)\|x\|_2 \|v\|_2}$$
$$\leq \frac{kr^{1/2}(|\lambda|+1)\max\{\|A_i\|_2\}\|z\|_2}{(\sum_{i=0}^k |\lambda|^i \|A_i\|_2)\|x\|_2}$$

It follows that

$$\sum_{i=0}^k |\lambda|^i \|A_i\|_2 \geq |\lambda|^k \|A_k\|_2 + \|A_0\|_2 . \geq (|\lambda|^k+1)\min\{\|A_k\|_2, \|A_0\|_2\}$$

Then, a calculation shows

$$\frac{\eta_P(x,\lambda)}{\eta_L(z,\lambda)} \leq \frac{kr^{1/2}(|\lambda|+1)\max\{\|A_i\|_2\}\|z\|_2}{(\sum_{i=0}^k |\lambda|^i \|A_i\|_2)\|x\|_2}$$
$$\leq kr^{1/2}\left(\frac{|\lambda|+1}{|\lambda|^k+1}\right)\frac{\max\{\|A_i\|_2\}}{\min\{\|A_0\|_2, \|A_k\|_2\}}\frac{\|z\|_2}{\|x\|_2}$$
$$= kr^{1/2}\rho\left(\frac{|\lambda|+1}{|\lambda|^k+1}\right)\frac{\|z\|_2}{\|x\|_2}$$

□

Notice that the upper bound for the backward error depends on the modulus of the eigenvalue. This requires the calculation of the eigenvalues, but furthermore, if the eigenvalue is large or small, the bounds could become extremely large or virtually 0. The next bound we provide is an alternative bound that does not depend on the modulus of the eigenvalue.

Theorem 3. *Let $P(\lambda)$ be a matrix polynomial and let $L(\lambda) \in \mathbb{DL}(P)$ be a linearization for P with ansatz vector v. Let z be a right eigenvector of L and let $x = (v^T \otimes I_n)z$. Let η_P and η_L be the backward error for the matrix polynomial and $L \in \mathbb{DL}(P)$. The ratio $\frac{\eta_P}{\eta_L}$ is bounded above such that*

$$\frac{\eta_P}{\eta_L} \leq k^{3/2} r^{1/2} \rho \frac{\|z\|_2}{\|x\|_2}.$$

Proof. Note that

$$\|(P(\lambda)x\|_2 = \left\|P(\lambda)(v^T \otimes I_n)z\right\|_2$$
$$= \left\|(v^T \otimes P(\lambda))z\right\|_2$$
$$= \left\|(\Lambda^T \otimes I_n)L(\lambda)z\right\|_2$$
$$\leq \left\|\Lambda^T \otimes I_n\right\|_2 \|L(\lambda)z\|_2$$

$$
\begin{aligned}
\frac{\eta_P}{\eta_L} &= \frac{\|P(\lambda)x\|}{\left(\sum_{i=0}^{k}|\lambda|^i\|A_i\|_2\right)\|x\|_2}\frac{\left(|\lambda|\,\|L_1\|_2+\|L_0\|_2\right)\|z\|_2}{\|L(\lambda)z\|_2} \\
&\leq \frac{\|\Lambda(\lambda)\otimes I_n\|_2\,\|L(\lambda)z\|_2\left(|\lambda|\,\|L_1\|_2+\|L_0\|_2\right)\|z\|_2}{\left(\sum_{i=0}^{k}|\lambda|^i\|A_i\|_2\right)\|x\|_2\,\|L(\lambda)z\|_2}.
\end{aligned}
$$

From the techniques used in the proof of the previous theorem,

$$
\begin{aligned}
\frac{\eta_P}{\eta_L} &\leq \frac{\left\|\Lambda^T(\lambda)\otimes I_n\right\|_2(|\lambda|+1)\max\{\|L_1\|_2,\|L_0\|_2\}\|z\|_2}{\left(\sum_{i=0}^{k}|\lambda|^i\|A_i\|_2\right)\|x\|_2} \\
&\leq \frac{\left\|\Lambda^T(\lambda)\otimes I_n\right\|_2(|\lambda|+1)kr^{1/2}\max\{\|A_i\|_2\}\|z\|_2}{(|\lambda|^k+1)\min\{\|A_k\|_2,\|A_0\|_2\}\|x\|_2}
\end{aligned}
$$

By Lemma A.1 in [5]

$$
\frac{\left\|\Lambda^T(\lambda)\otimes I_n\right\|_2(|\lambda|+1)}{|\lambda|^k+1}\leq\sqrt{k}.
$$

Thus,

$$
\begin{aligned}
\frac{\eta_P}{\eta_L} &\leq \frac{k^{3/2}r^{1/2}\max\{\|A_i\|_2\}}{\min\{\|A_k\|_2,\|A_0\|_2\}}\frac{\|z\|_2}{\|x\|_2} \\
&= k^{3/2}r^{1/2}\rho\frac{\|z\|_2}{\|x\|_2}
\end{aligned}
$$

where $\rho = \frac{\max\{\|A_i\|_2\}}{\min\{\|A_k\|_2,\|A_0\|_2\}}$. □

Even though these bounds work for any linearization in $\mathbb{DL}(P)$, the two optimal linearizations are D_1 and D_k. Unfortunately, D_1 and D_k are linearizations of $P(\lambda)$ only if A_0 **and** A_k are nonsingular, respectively. Another disadvantage of these two pencils is that the bounds for the backward error depend on the modulus of the eigenvalue. These pencils are optimal and linearizations, but only under the conditions mentioned.

3.2. **Mackey Pencil.** As seen in Section 3.1, the pencils D_1 and D_k were shown to be optimal in terms of conditioning and backward error, but both depended on the modulus of the eigenvalue. Finding a pencil that numerically behaves like D_1 and D_k for all eigenvalues would prove extremely helpful, as this pencil could replace the other two. This section of the paper focuses on one specific block-symmetric linearization of $P(\lambda)$ which we will refer to as the Mackey pencil.

Definition 3.2. The Mackey pencil $T(\lambda)$, is a linearization of $P(\lambda)$ of the following form, depending on k.

If k is even, $T(\lambda)$ is

$$\lambda \begin{bmatrix} 0 & I & & & \\ I & A_{k-1} & & & \\ & & \ddots & & \\ & & & 0 & I \\ & & & I & A_1 \end{bmatrix} - \begin{bmatrix} A_k^{-1} & & & & & & \\ & -A_{k-2} & I & & & & \\ & I & 0 & \ddots & & & \\ & & & & -A_2 & I & \\ & & & & I & 0 & \\ & & & & & & -A_0 \end{bmatrix}.$$

If k is odd, $T(\lambda)$ is

$$\lambda \begin{bmatrix} A_k & & & & & \\ & 0 & I & & & \\ & I & A_{k-2} & & & \\ & & & \ddots & & \\ & & & & 0 & I \\ & & & & I & A_1 \end{bmatrix} - \begin{bmatrix} -A_{k-1} & I & & & & \\ I & 0 & & & & \\ & & \ddots & & & \\ & & & -A_2 & I & \\ & & & I & 0 & \\ & & & & & -A_0 \end{bmatrix}.$$

Unlike pencils in $\mathbb{DL}(P)$, when k is odd A_0 and A_k do not need to be nonsingular. When k is even, we see that A_k must be nonsingular, which could be problematic. In this section of the paper we will theoretically focus on when k is odd due to the good theoretical behavior of k odd. Later on, we will show numerical experiments for when k is even, explaining why we did not further explore the even case theoretically.

The bounds for the condition number[2] were constructed by Bueno, Dopico, and Furtado for generalized Fiedler pencils.

Theorem 4. *Let $P(\lambda)$ be a regular matrix polynomial of odd degree k with $A_0 \neq 0$. Assume that δ is a simple, finite, nonzero eigenvalue of $P(\lambda)$. Let $T(\lambda)$ be the Mackey pencil of $P(\lambda)$. Then,*

$$\frac{1}{k} \leq \frac{\kappa_T(\delta)}{\kappa_P(\delta)} \leq 2k^4 \rho_1,$$

where $\rho_1 = \frac{\max_{i=0:k}\{\|A_i\|_2, 1\}^3}{\min\{\|A_k\|_2, \|A_0\|_2\}}$ *and* $\rho_2 = \frac{\min\{\max\{\|A_k\|_2,1\},\max\{\|A_0\|_2,1\}\}}{\max\{\|A_i\|_2\}}$.

Definition 3.3. For a computed eigenpair, the backward error for the matrix polynomial, $P(\lambda)$ and for the Mackey pencil, $T(\lambda)$ are given by

$$\eta_P(\hat{x}, \hat{\delta}) = \frac{\left\| P(\hat{\delta})\hat{x} \right\|_2}{(\sum_{i=0}^m |\hat{\delta}^i| \, \|A_i\|_2) \, \|\hat{x}\|_2}$$

where $(\hat{x}, \hat{\delta})$ is a computed eigenpair for P. and

$$\eta_T(\hat{z}, \hat{\delta}) = \frac{\left\| T(\hat{\delta})\hat{z} \right\|_2}{(|\hat{\delta}| \, \|T_1\|_2 + \|T_0\|_2)}.$$

where $(\hat{z}, \hat{\delta})$ is a computed eigenpair for L.

It would be convenient to find bounds on the backward error for a random matrix polynomial $P(\lambda)$ that does not depend on the modulus of the eigenvalue. Thus, the backward error for any matrix polynomial$P(\lambda)$ would be easy to compute.

Theorem 5. *Let $P(\lambda)$ be a regular matrix polynomial of odd degree k with $A_0 \neq 0$. Assume that $\hat{\delta}$ is a simple, nonzero, finite computed eigenvalue of $P(\lambda)$ and $\hat{\delta}$ is a computed eigenvalue of L. Let $T(\lambda)$ be the Mackey pencil of $P(\lambda)$. Then, the backward error for a computed eigenpair is bounded by*

$$\frac{1}{k} \leq \frac{\eta_P(\hat{x}, \hat{\delta})}{\eta_T(\hat{z}, \hat{\delta})} \leq \sqrt{4k^5 + 2k^2}\ \rho_1'$$

where $\rho_1' = \frac{\max\{\|A_i\|_2, 1\}^2}{\min\{\|A_k\|_2, \|A_0\|_2\}}$.

Proof. Recall that

$$\frac{\eta_P}{\eta_T} = \frac{\|P(\lambda)x\|_2}{\left(\sum_{i=0}^{k} |\lambda|^i \|A_i\|_2\right) \|x\|_2} \cdot \frac{\left(|\lambda| \|T_1\|_2 + \|T_0\|_2\right) \|z\|_2}{\|T(\lambda)z\|_2}.$$

From [2,Lemma 7.1] we know that

$$T(\lambda)\Delta = e_k \otimes P(\lambda)$$

and therefore that

$$T(\lambda)\Delta x = e_k \otimes P(\lambda)x.$$

From Theorem 7.2 in [2], a vector z is a right eigenvector of $T(\lambda)$ associated with λ if and only if $z = \Delta x$. Thus,

$$T(\lambda)z = e_k \otimes P(\lambda)x$$

which implies

$$\|T(\lambda)z\|_2 = \|P(\lambda)x\|_2 .$$

After simplifying,

$$\frac{\eta_P}{\eta_T} = \frac{\left(|\lambda| \|T_1\|_2 + \|T_0\|_2\right) \|z\|_2}{\left(\sum_{i=0}^{k} |\lambda|^i \|A_i\|_2\right) \|x\|_2}.$$

From [2, 7.7] states

$$|\lambda| \|T_1\|_2 + \|T_0\|_2 \leq (|\lambda| + 1)k \max\{\|A_i\|_2, 1\}.$$

Thus,

$$\frac{\eta_P}{\eta_T} \leq \frac{(|\lambda| + 1)k \max_{i=0:k}\{\|A_i\|_2, 1\} \|z\|_2}{\left(\sum_{i=0}^{k} |\lambda|^i \|A_i\|_2\right) \|x\|_2}.$$

As mentioned in the previous proof,

$$(\sum_{i=0}^{k} |\lambda|^i \|A_i\|_2) \geq |\lambda| \|A_k\|_2 + \|A_0\|_2 .$$

Thus,

$$|\lambda|^k \|A_k\|_2 + \|A_0\|_2 \geq (|\lambda|^k + 1) \min\{\|A_k\|_2 , \|A_0\|_2\}.$$

It follows that

$$\begin{aligned} \frac{\eta_P}{\eta_T} &\leq \frac{(|\lambda| + 1)k \max_{i=0:k}\{\|A_i\|_2 , 1\} \|z\|_2}{(|\lambda|^k + 1) \min\{\|A_k\|_2 , \|A_0\|_2\} \|x\|_2} \\ &\leq \frac{(|\lambda| + 1)}{(|\lambda|^k + 1)} \cdot k \cdot \frac{\max\{\|A_i\|_2 , 1\}}{\min\{\|A_k\|_2 , \|A_0\|_2\}} \cdot \frac{\|\Delta x\|_2}{\|x\|_2}. \end{aligned}$$

Bueno, Dopico, and Furtado [2, 7.3] provide a proof that will aid in the further construction of the backward error bounds.

$$\begin{aligned} \|\Delta x\|_2^2 &= \sum_{r=0}^{\frac{k-1}{2}} |\lambda|^{2r} \|x_1\|_2^2 + \sum_{r=1}^{\frac{k-1}{2}} |\lambda|^{2r} \|P_{k-2r}(\lambda)x_1\|_2^2 \\ &\leq \left[\sum_{r=0}^{\frac{k-1}{2}} |\lambda|^{2r} \|x_1\|_2^2 + \sum_{r=1}^{\frac{k-1}{2}} |\lambda|^{2r} \|P_{k-2r}(\lambda)\|_2^2 \right] \|x_1\|_2^2 . \end{aligned}$$

Taking the square root of both sides,

$$\begin{aligned} \|\Delta x_1\|_2 &= \sqrt{\sum_{r=0}^{\frac{k-1}{2}} |\lambda|^{2r} \|x_1\|_2^2 + \sum_{r=1}^{\frac{k-1}{2}} |\lambda|^{2r} \|P_{k-2r}(\lambda)x_1\|_2^2} \\ &\leq \sqrt{\left[\sum_{r=0}^{\frac{k-1}{2}} |\lambda|^{2r} \|x_1\|_2^2 + \sum_{r=1}^{\frac{k-1}{2}} |\lambda|^{2r} \|P_{k-2r}(\lambda)\|_2^2 \right]} \|x_1\|_2 \\ \frac{\|\Delta x_1\|_2}{\|x_1\|_2} &= \sqrt{\left[\sum_{r=0}^{\frac{k-1}{2}} |\lambda|^{2r} \|x_1\|_2^2 + \sum_{r=1}^{\frac{k-1}{2}} |\lambda|^{2r} \|P_{k-2r}(\lambda)\|_2^2 \right]}. \end{aligned}$$

It follows that

$$\begin{aligned} \|\lambda A_k + A_{k-1}\|_2 &\leq |\lambda + 1| \max\{\|A_i\|_2\} \\ \left\|\lambda^{k-2} A_k + \ldots + A_2\right\|_2 &\leq |\lambda^{k-2} + \ldots + 1| \max\{\|A_i\|_2\}, \end{aligned}$$

Thus,

$$\frac{\|\Delta x\|_2}{\|x\|_2} = \sqrt{\sum_{r=0}^{\frac{k-1}{2}} |\lambda|^{2r} + \sum_{r=1}^{\frac{k-1}{2}} \left(\max_{i=0:k}\{\|A_i\|_2\} \sum_{s=0}^{k-2r} |\lambda|^s \right)^2}.$$

Using a similar process from [2, Lemma 5.3 (pp.21)],

$$\begin{aligned}
\left(\sum_{s=r}^{k-r} |\lambda|^s\right)^2 &= \left(\sum_{s=0}^{k-2r} |\lambda|^{s+r}\right)^2 \\
&= |\lambda|^{2r} \left(\sum_{s=0}^{k-2r} |\lambda|^s\right)^2 \\
&\leq |\lambda|^{2r}(k-2r+1) \sum_{j=0}^{2(k-2r)} |\lambda|^s \\
&= (k-2r+1) \sum_{s=0}^{2(k-2r)} |\lambda|^{s+2r} \\
&= (k-2r+1) \sum_{s=2r}^{2k-2r} |\lambda|^s.
\end{aligned}$$

Then,

$$\begin{aligned}
\frac{\|\Delta x\|_2}{\|x\|_2} &\leq \sqrt{\left[\sum_{r=0}^{\frac{k-1}{2}} |\lambda|^{2r} + \sum_{r=1}^{\frac{k-1}{2}} \left((k-2r+1) \sum_{s=2r}^{2k-2r} |\lambda|^s \right)\right] \max\{\|A_i\|_2, 1\}^2} \\
&\leq \sqrt{\left[\sum_{r=0}^{\frac{k-1}{2}} |\lambda|^{2r} + \sum_{r=1}^{\frac{k-1}{2}} \left((k-2r+1) \sum_{s=2r}^{2k-2r} |\lambda|^s \right)\right]} \max\{\|A_i\|_2, 1\}.
\end{aligned}$$

Calculations show that

$$\frac{(|\lambda|+1)}{(|\lambda|^k+1)} \frac{\|\Delta x\|_2}{\|x\|_2} \leq \frac{(|\lambda|+1)}{(|\lambda|^k+1)} \sqrt{\left[\sum_{r=0}^{\frac{k-1}{2}} |\lambda|^{2r} + \sum_{r=1}^{\frac{k-1}{2}} \left((k-2r+1) \sum_{s=2r}^{2k-2r} |\lambda|^s \right)\right]} \max\{\|A_i\|_2, 1\}.$$

Notice that

$$\begin{aligned}
\sum_{r=0}^{\frac{k-1}{2}} |\lambda|^{2r} &\leq \sum_{r=0}^{\frac{k-1}{2}} 1 \\
&= \sqrt{\frac{k+1}{2} + \frac{(k-1)^2}{2}(2k-3)} \\
&\leq \frac{(|\lambda|+1)}{(|\lambda|^k+1)}\sqrt{\frac{k+1}{2} + \frac{(k-1)^2}{2}(2k-3)} \\
&\leq 2\sqrt{\frac{k+1}{2} + \frac{(k-1)^2}{2}(2k-3)} \\
&= \sqrt{2(k+1) + 2(k-1)^2(2k-3)} \\
&\leq \sqrt{4k^3+2}
\end{aligned}$$

The $\sqrt{4k^3}$ provided real roots, so we had to increase by 2 so that there were no real roots.

Therefore,

$$\frac{\eta_P}{\eta_T} \leq \sqrt{4k^5+2k^2}\rho_1'$$

To find a lower bound [3], note that

$$\eta_P(\hat{x},\hat{\delta}) = \min\{\epsilon : [P(\hat{\delta}) + \Delta P(\hat{\delta})]\hat{x} = 0, \|\Delta A_i\|_2 \leq \epsilon \|A_i\|_2\}.$$

Let $\{\Delta A_i\}$ be an optimal perturbation of $\{A_i\}$.

$$T(\lambda) = \lambda X_1 + Y_1$$

$$T(\lambda) = \lambda \begin{bmatrix} A_k & & & & & \\ & 0 & I & & & \\ & I & A_{k-2} & & & \\ & & & \ddots & & \\ & & & & 0 & I \\ & & & & I & A_1 \end{bmatrix} - \begin{bmatrix} -A_{k-1} & I & & & & & & \\ I & 0 & & & & & & \\ & & -A_{k-3} & I & & & & \\ & & I & 0 & & & & \\ & & & & \ddots & & & \\ & & & & & -A_2 & I & \\ & & & & & I & 0 & \\ & & & & & & & -A_0 \end{bmatrix}$$

$$\Delta T(\lambda) = \lambda \Delta X_1 + \Delta Y_1$$

$$\Delta T(\lambda) = \lambda \begin{bmatrix} \Delta A_k & & & & & & \\ & 0 & & & & & \\ & & \Delta A_{k-2} & & & & \\ & & & 0 & & & \\ & & & & \Delta A_{k-4} & & \\ & & & & & \ddots & \\ & & & & & & \Delta A_1 \end{bmatrix} - \begin{bmatrix} \Delta A_{k-1} & & & & & \\ & 0 & & & & \\ & & \Delta A_{k-3} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & \Delta A_0 \end{bmatrix}$$

Using the well known property from Section 5 in [2] that for any $l \times n$ block-matrix $B = B_{i,j}$,

$$\max_{i,j} \|B_{ij}\|_2 \leq \|B\|_2 \leq \sqrt{ln} \max_{i,j} \|B_{ij}\|_2 ,$$

it follows that

$$\begin{aligned} \|\Delta X_1\|_2 &\leq \sqrt{k^2} \max\{\|\Delta A_k\|_2 , \dots \|\Delta A_1\|_2\} \\ &\leq k \cdot \eta_P \max\{\|\Delta A_k\|_2 , \dots \|\Delta A_1\|_2\} \\ &\leq k \cdot \eta_P \|X_1\|_2 . \end{aligned}$$

Similarly,

$$\begin{aligned} \|\Delta Y_1\|_2 &\leq \sqrt{k^2} \max\{\|\Delta A_{k-1}\|_2 , \dots \|\Delta A_0\|_2\} \\ &\leq k \cdot \eta_P \max\{\|\Delta A_{k-1}\|_2 , \dots \|\Delta A_0\|_2\} \\ &\leq k \cdot \eta_P \|Y_1\|_2 . \end{aligned}$$

Therefore,

$$\eta_T \leq k \cdot \eta_P.$$

This results in the bounds of the backward error,

$$\frac{1}{k} \leq \frac{\eta_P}{\eta_T} \leq \sqrt{4k^5 + 2k^2}\ \rho_1'$$

where $\rho_1' = \frac{\max_{i=0:k}\{\|A_i\|_2, 1\}^2}{\min\{\|A_k\|_2, \|A_0\|_2\}}$. □

Unlike D_1 and D_k, we notice that the bounds for the backward error do not depend on the modulus of the eigenvalue. Thus, theoretically the bounds for the backward error are the same for a matrix polynomial, regardless of the eigenvalue.

3.3. **The First Companion Form C_1 of P.** In Section 5 we will be comparing the condition number and backward error of the Mackey pencil, $T(\lambda)$ with other pencils that are linearizations in literature. One of these pencils is C_1, the first companion form of P. The pencil C_1 is a commonly used pencil for unstructured matrix polynomials. By comparing T with C_1, which is a pencil of reference for general matrix polynomials P, we can show that since T behaves similarly to C_1, we can use T for non-structured polynomials as well.

Definition 3.4. We define C_1 for matrix polynomial P of degree k in the following way:

$$C_1(\lambda) = M^P_{-k}\lambda + M^P_{k-1,k-2,\ldots,1,0}$$

C_1 has the following structure for its coefficients:

$$C_1(\lambda) = \lambda \begin{bmatrix} A_k & & & & & \\ & I & 0 & & & \\ & 0 & I & & & \\ & & & \ddots & & \\ & & & & & 0 \\ & & & & 0 & I \end{bmatrix} - \begin{bmatrix} -A_{k-1} & -A_{k-2} & & \cdots & & -A_0 \\ I & 0 & & & & 0 \\ 0 & I & 0 & & & \\ \vdots & 0 & \ddots & \ddots & & \vdots \\ & \vdots & & & & \\ 0 & 0 & & & I & 0 \end{bmatrix}$$

Theorem 6. *Let P be a matrix polynomial and C_1 be its first companion form. Let $\hat{z}$ be an approximate right eigenvector for C_1 corresponding to the approximate eigenvalue $\hat{\delta}$. Then for $z_i = z((i-1):n+1:in)$, $i = 1:k$, we have:*

$$\frac{1}{\sqrt{k}} \leq \frac{\eta_P(\hat{z}_i, \hat{\delta})}{\eta_{C_1}(\hat{z}, \hat{\delta})} \leq k^{5/2} \frac{\|\hat{z}_2\|}{\|\hat{z}_i\|_2} \rho_1'.$$

- *If $|\hat{\delta}| \leq 1$, then $i = k$.*
- *If $|\hat{\delta}| \geq 1$, then $i = 1$.*

This theorem and its proof can be found in [1, Theorem 3.6].

4. Scaling P to Improve Condition Number and Backward Error Ratios

Since the original PEP $P(\lambda)x = 0$ is being replaced with the problem $L(\lambda)\hat{x} = 0$, it is favorable that the sensitivity of the new problem is not worse than that of the original problem. This way, it can be ensured that computed solutions from the linearization are close to the computed solutions for the original matrix polynomial. Thus, it is desirable that $\kappa_L \leq a \cdot \kappa_P$, where $a \approx 1$. If this inequality holds, then it is clear that the condition number for the linearization is not larger, meaning worse than the condition number for

P. The condition number ratio $\frac{\kappa_L(\delta)}{\kappa_P(\delta)}$ is therefore ideal when it is close to 1, as this shows that the condition numbers for the linearizations are not worse than those for the matrix polynomials.

It is well-known that backward errors using the QZ-algorithm are small for linearizations of P. Thus it is also desirable that $\eta_P \leq a \cdot \eta_L$, where $a \approx 1$. If this inequality holds, then it is clear that the backward error for P is no worse than the backward error for L, which is already known to be small. The backward error ratio $\frac{\eta_P}{\eta_L}$ is therefore ideal when it is close to 1. This demonstrates that the backward errors for the matrix polynomial are not worse than those for the linearization.

Our goal is to find a linearization of P which producescondition number and backward error ratios that are close to 1. Well-known pencils in literature D_1, D_k, and C_1 manage to fulfill this goal, but each of these pencils have disadvantages, which make these pencils useless under certain situations.

Recall from earlier in the previous section that the pencils D_1 and D_k are only linearizations for P when A_0 and A_k are nonsingular,respectively, meaning that these pencils only work well for very specific P. The optimality of the condition number and backward error for both pencils depends on the modulus of the eigenvalue. Ideally, we would like a single pencil which can be used regardless of the modulus of the eigenvalues. C_1 is a useful linearization for unstructured P, but it is not useful for P with structural properties such as being Hermitian.

The Mackey Pencil $T(\lambda)$ has signficant advantages over D_1, D_k, and C_1 when k is odd. In this case, T does not require that A_0 or A_k are nonsingular. For matrix polynomials where the coefficient norms are close together, T produces ratios close to 1 regardless of the modulus of the eigenvalues. Not only is T useful for structured P, but numerical experiments show that it is comparable to C_1 for unstructured P. This means that T can be applied more generally than D_1, D_k and C_1.

Recall that the bounds for $\frac{\kappa_T(\delta)}{\kappa_P(\delta)}$ are as follows:

$$\frac{1}{k} \leq \frac{\kappa_T(\delta)}{\kappa_P(\delta)} \leq 2k^4\rho_1.$$

The backward error ratio bounds for $\frac{\eta_P}{\eta_T}$ are

$$\frac{1}{k} \leq \frac{\eta_P}{\eta_T} \leq \sqrt{4k^5 + 2k^2}\ \rho_1'.$$

Finally, let us recall the backward error bounds for $\frac{\eta_P}{\eta_{C_1}}$ are

$$\frac{1}{k^{1/2}} \leq \frac{\eta_P(z_i,\delta)}{\eta_{C_1}(z,\delta)} \leq k^{5/2}\frac{\|z\|_2}{\|z_i\|_2}\rho_1',$$

where

$$\rho_1 = \frac{\max_{i=0:k}\{\|A_i\|_2, 1\}^3}{\min\{\|A_k\|_2, \|A_0\|_2\}},\ \rho_2 = \frac{\min\{\max\{\|A_k\|_2, 1\}, \max\{\|A_0\|_2, 1\}\}}{\max\{\|A_i\|_2\}},\ \rho_1' = \frac{\max_{i=0:k}\{\|A_i\|_2, 1\}^2}{\min\{\|A_k\|_2, \|A_0\|_2\}}.$$

Since the ideal ratios for the condition number and the backward error are close to 1, it would be useful for the theoretical bounds to be closer to 1. It is clear that by minimizing constants ρ_1, ρ_1' and ρ_2, the theoretical bounds for condition number and backward error will tighten. Notice that these constants depend on the norms of the matrix coefficients of P. Therefore, scaling the matrix polynomial P into another matrix polynomial $\tilde{P}$ in such a way that the eigenstructure of $\tilde{P}$ is the same as that of P will tighten these theoretical bounds and will produce better conditioning and backward error results for $T(\lambda)$.

The chosen scaling is a simple one which depends only on the maximal norm of the coefficients, A_i of P. We define the scaling in the following way:

$$\tilde{P}(\lambda) = \frac{1}{\beta}P(\lambda), \text{ where } \beta = \max_{i=0:k}\{\|A_i\|_2\}.$$

Using this scaling, $\tilde{P}$ has the same eigenstructure as P and it is easy to show that $\tilde{P}$ also has the same condition number and backward error as P. Each coefficient of $\tilde{P}$ has norm no more than 1. Therefore, constructing a Mackey pencil with $\tilde{P}$ will produce better results than constructing a Mackey pencil with P, since the theoretical bounds for the condition number ratios and the backward error ratios will improve and tighten.

5. Numerical Experiments

In this section we present our numerical experiments which support the theory addressed in this report. Experiments were programmed and were run on MATLAB_R2015a. Condition numbers and backward errors were compared between the matrix polynomial P and linearizations D_1, D_k, the first companion form C_1 of P, and the Mackey pencil $T(\lambda)$. Random selections for P were examined as well as selected matrix polynomials from applied problems in literature. Eigenvalues, eigenvectors, and condition numbers were computed exactly up to 40 digits using MATLAB's symbolic toolbox alongside its VPA (Variable Precision Arithmetic). Graphs were produced in order to show the behavior of each linearization with respect to P. Scaling was then applied to P in order to improve the numerical behavior of the Mackey pencil $T(\lambda)$. Graphs were remade for the scaled P to demonstrate the improved behavior of $T(\lambda)$. Numerical experiments were done for both k odd and k even to show the desirable behavior of $T(\lambda)$ when k is odd and the undesirable behavior of $T(\lambda)$ when k is even.

First, we examine the case where k is odd and P is a random matrix polynomial with coefficient norms that are close together. The random matrix polynomials P were generated by choosing random matrix coefficients with entries between $[-10, 10]$. These random

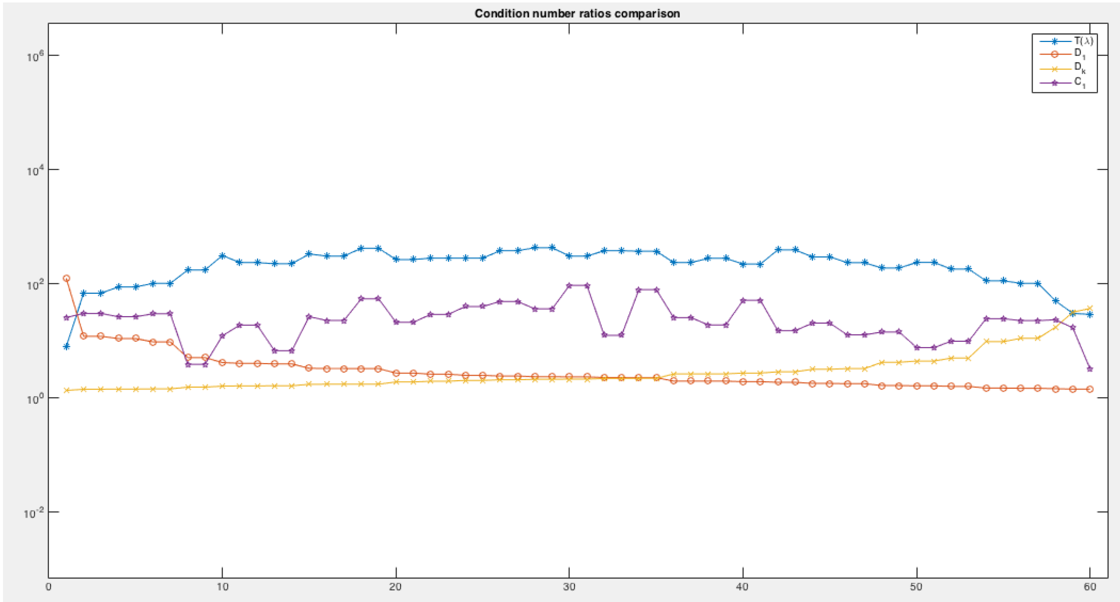

Figure 1. Condition number comparison for unscaled random matrix polynomial P of degree $k = 3$ with coefficients 20×20.

coefficients were created with the MATLAB built-in function "randn", which creates invertible matrices with probability 1. The condition numbers were compared between L and a random matrix polynomial P. The linearizations considered are: D_1 and D_k from the standard basis of $\mathbb{DL}(P)$, the first companion form C_1 of P, and the Mackey pencil $T(\lambda)$. In the figures for randomly generated P with coefficient norms that are close together, the x-axis has the indices $1, 2, ..., nk$ for the non-zero, simple eigenvalues for L and P, which are sorted in increasing order by magnitude (1 is the first eigenvalue in this order, and so on). On the y-axis are the computed $\frac{\kappa_L(\delta)}{\kappa_P(\delta)}$ for linearization L and matrix polynomial P.

Figure 1 shows these ratios for a random matrix polynomial P, where $k = 3$ and $n = 20$. Note that in Figure 1, the behavior of D_1 and D_k follow exactly what was proven by N. Higham, S. Mackey, R. Li and F. Tisseur in [1] and [4]. For $|\delta| \leq 1$ the optimal linearization is D_k and for $|\delta| \geq 1$ the optimal linearization is D_1. In this figure, P is unscaled. The behavior of $T(\lambda)$ in this case is far from ideal, with ratios mostly close to 10^2 and the same undesired behavior is seen with C_1. Creating a new graphic with P scaled yields remarkably better results for $T(\lambda)$ and C_1 which can be seen in Figure 2. We see that after scaling, the ratios for C_1 are much closer to 1. However the ratios for $T(\lambda)$ are even closer to 1 and optimality does not depend on $|\delta|$. It is worth noting that while the scaling improved the behavior of $T(\lambda)$, it had no effect on the condition number ratios for D_1 and D_k and that in this particular case, $T(\lambda)$ outperforms C_1 after scaling.

Figure 3 and Figure 4 demonstrate the improvement of the backward error after scaling the same P. Again, we notice that the backward error ratios for D_1 and D_k behave as

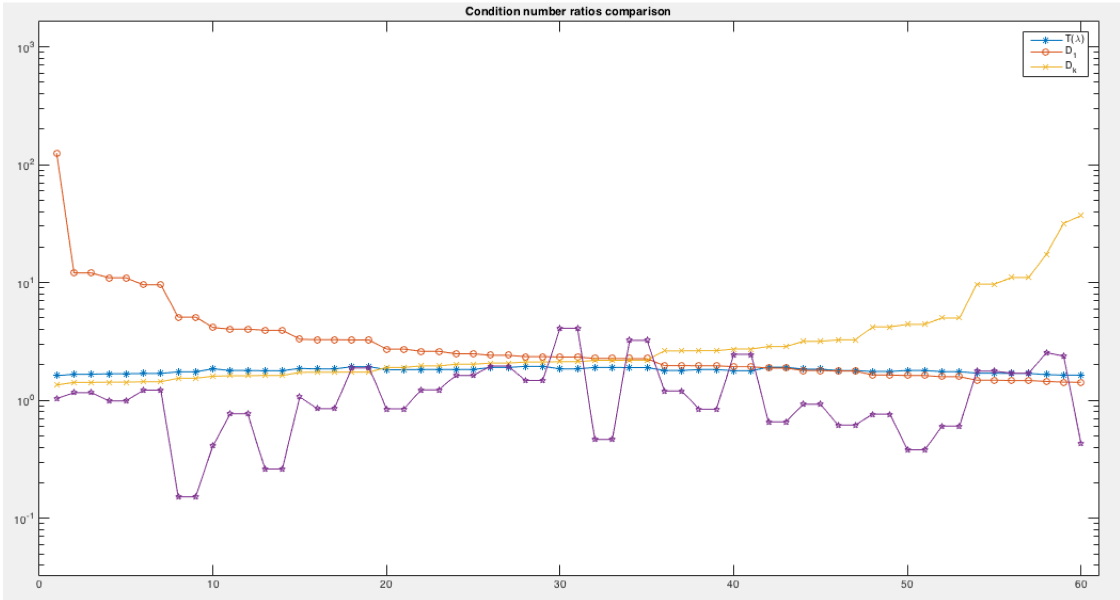

FIGURE 2. Condition number comparison after scaling the matrix polynomial P in Figure 1.

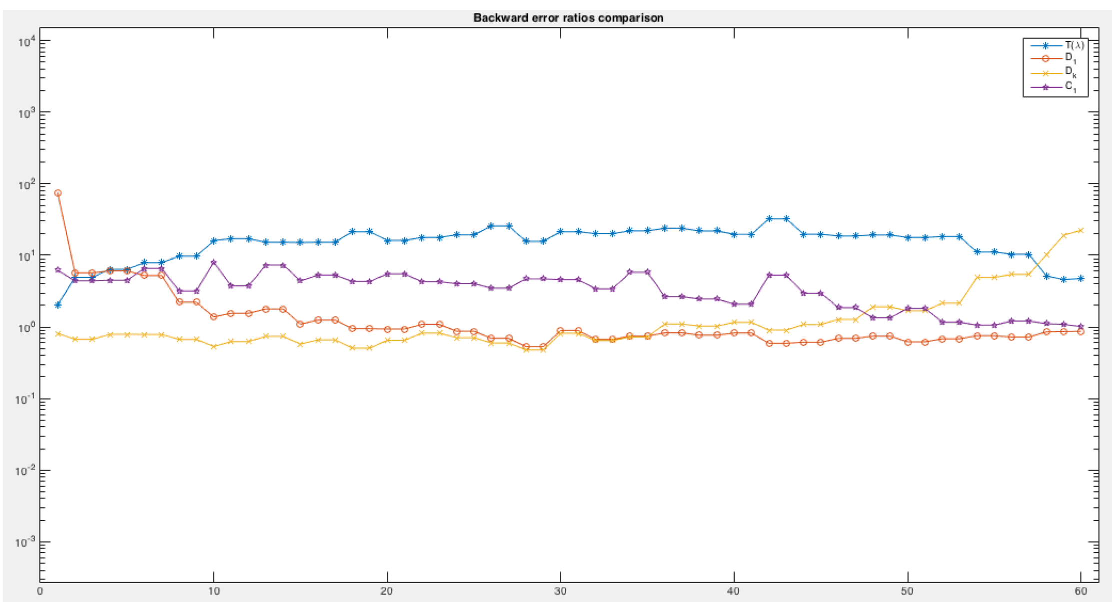

FIGURE 3. Backward error comparison for unscaled P.

what was proven theoretically. The ratios for T and C_1 are not ideal before scaling, but become much closer to 1 after being constructed from $\tilde{P}$. All data pertaining to Figures 1-4 can be seen in Table 1.

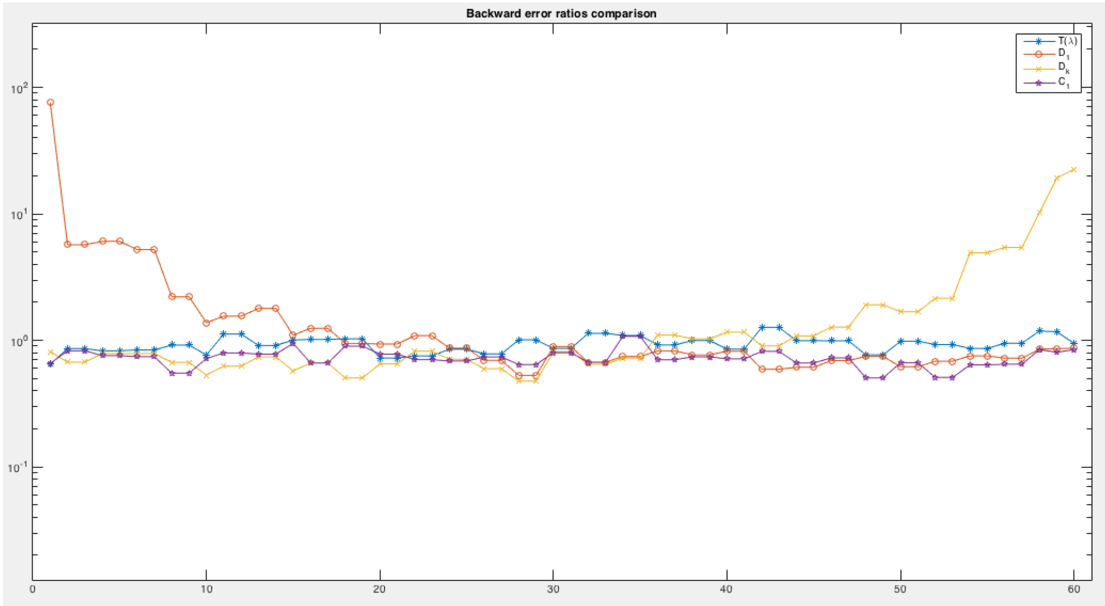

FIGURE 4. Backward error comparison after scaling the matrix polynomial P.

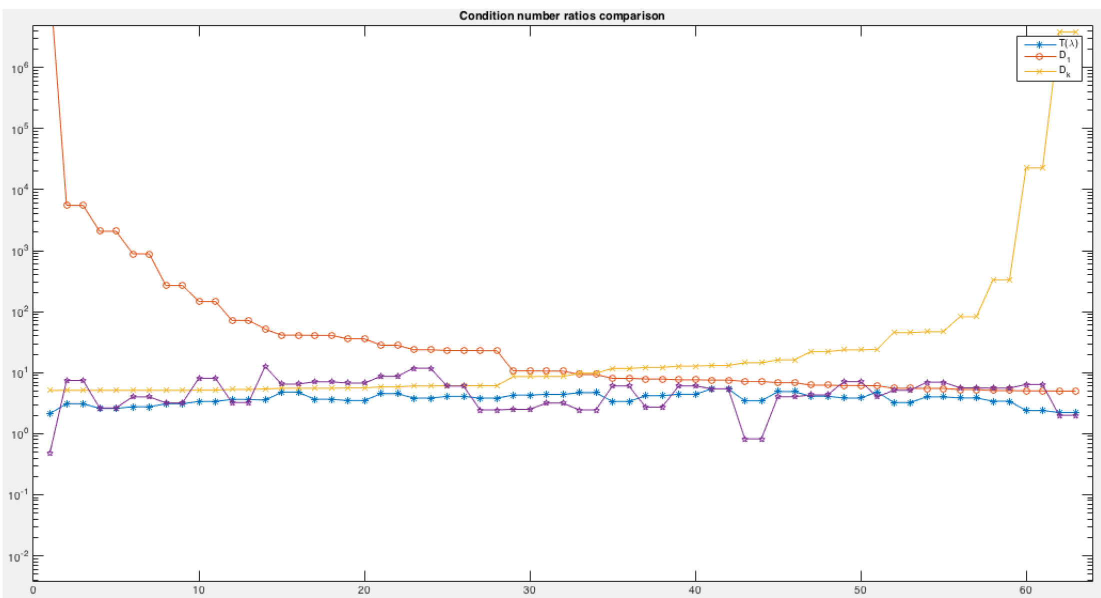

FIGURE 5. Condition number comparison after scaling a matrix polynomial P with $k = 21$ and $n = 3$.

One might expect $T(\lambda)$ to numerically behave poorly for large k, since the bounds for the condition number and backward error ratios depend on k. However optimal behavior is still observed for matrix polynomials of large degree k. Figure 5 and Figure 6 show the scaled results for a random matrix polynomial of degree $k = 21$ and with coefficients of

TABLE 1. Experiment results for random polynomials: $k = 3, n = 2$ and $k = 3, n = 20$.

Problem:	k=3 n=2		k=3 n=20	
	Unscaled	Scaled	Unscaled	Scaled
$\|\delta_{min}\|$	0.24	0.24	0.114	0.114
$\|\delta_{max}\|$	4.437	4.437	4.872	4.872
$\|A_3\|_2$	10.51	1	46.60	0.99002
$\|A_2\|_2$	7.77	0.74	47.07	1
$\|A_1\|_2$	10.44	0.99	46.39	0.986
$\|A_0\|_2$	10.06	0.96	47.05	0.9995
ρ_1	115.4	1.04	2238.17	1.01
ρ_2	0.095	0.96	0.0212	0.990
ρ_1'	10.98	1.04	47.55	1.01
T bounds for back. err.	[0.333, 345.5]	[0.333, 32.88]	[0.333, 1496.03]	[0.333, 31.78]
T bounds for cond. num.	[0.032, 18695.5]	[0.32, 169.3]	[0.007, 362583.1]	[0.33, 163.63]
$\min\{\eta_P/\eta_{D_1}\}$	0.53	0.53	0.53	0.53
$\max\{\eta_P/\eta_{D_1}\}$	21.98	21.98	75.69	75.69
$\min\{\eta_P/\eta_{D_k}\}$	0.59	0.59	0.48	0.48
$\max\{\eta_P/\eta_{D_k}\}$	9.78	9.78	22.57	22.57
$\min\{\eta_P/\eta_T\}$	0.597	0.65	2.02	0.65
$\max\{\eta_P/\eta_T\}$	5.04	1.56	32.39	1.27
$\min\{\eta_P/\eta_{C_1}\}$	1.43	0.603	1.01	0.5
$\max\{\eta_P/\eta_{C_1}\}$	2.12	1.00	8.07	1.08
$\min\{\kappa_{D_1}/\kappa_P\}$	1.57	1.57	1.42	1.42
$\max\{\kappa_{D_1}/\kappa_P\}$	30.04	30.04	125.45	125.45
$\min\{\kappa_{D_k}/\kappa_P\}$	1.36	1.36	1.35	1.35
$\max\{\kappa_{D_k}/\kappa_P\}$	29.16	29.16	37.1	37.1
$\min\{\kappa_T/\kappa_P\}$	1.52	1.55	7.88	1.63
$\max\{\kappa_T/\kappa_P\}$	32.31	2.12	429.3	1.94
$\min\{\kappa_{C_1}/\kappa_P\}$	1.34	0.44	3.2	0.152
$\max\{\kappa_{C_1}/\kappa_P\}$	17.94	2.61	93.96	4.11

size 3×3.

In the previous examples, experiments were run on randomly generated matrix polynomials whose matrix coefficients had norms close together. Experiments were also run on random polynomials with one coefficient of very large norm. In order to run these experiments, matrix polynomials were generated as before, except that a coefficient was modified by replacing one entry, which was originally between $[-10, 10]$, with 10^6. Figures 7-10 show results for a randomly generated matrix polynomial with modified A_0.

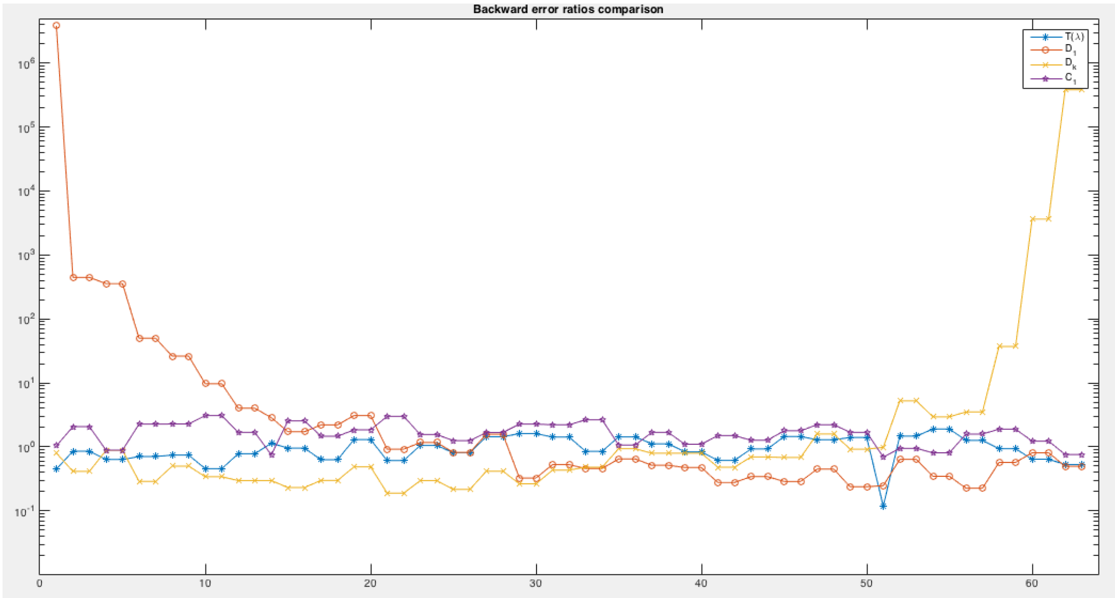

FIGURE 6. Backward error comparison after scaling the same P from Figure 5.

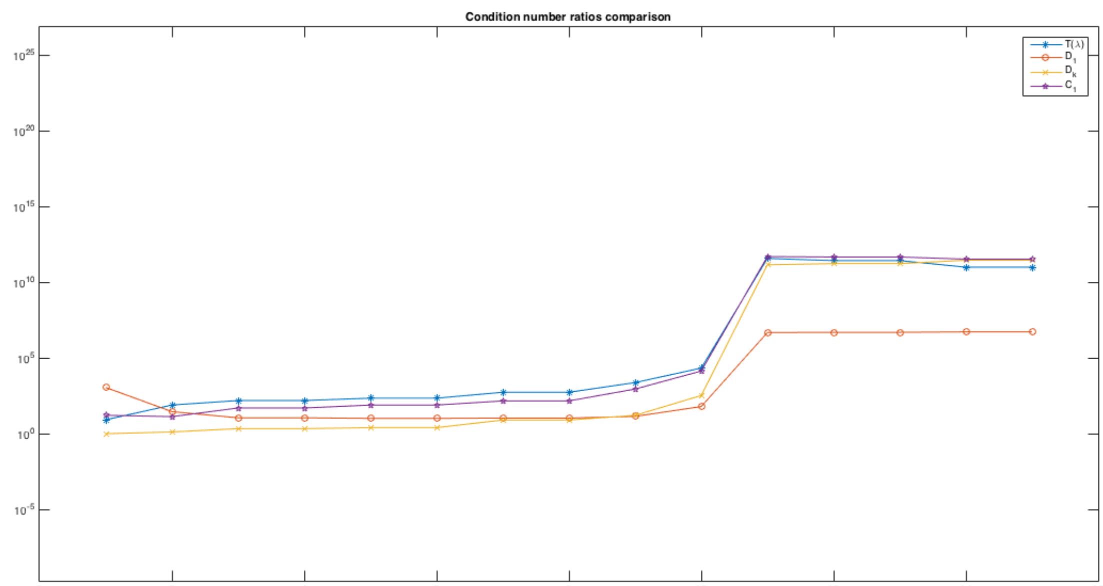

FIGURE 7. Condition number comparison for P with $k = 5$ and $n = 3$ with modified A_0

Based on these figures, there is reason to believe that T has ratios close to 1 only when the norms of the coefficients of P are close together. However, this is an open topic that must be investigated theoretically in the future. However, we see that in this particular case, $T(\lambda)$ remains comparable with C_1.

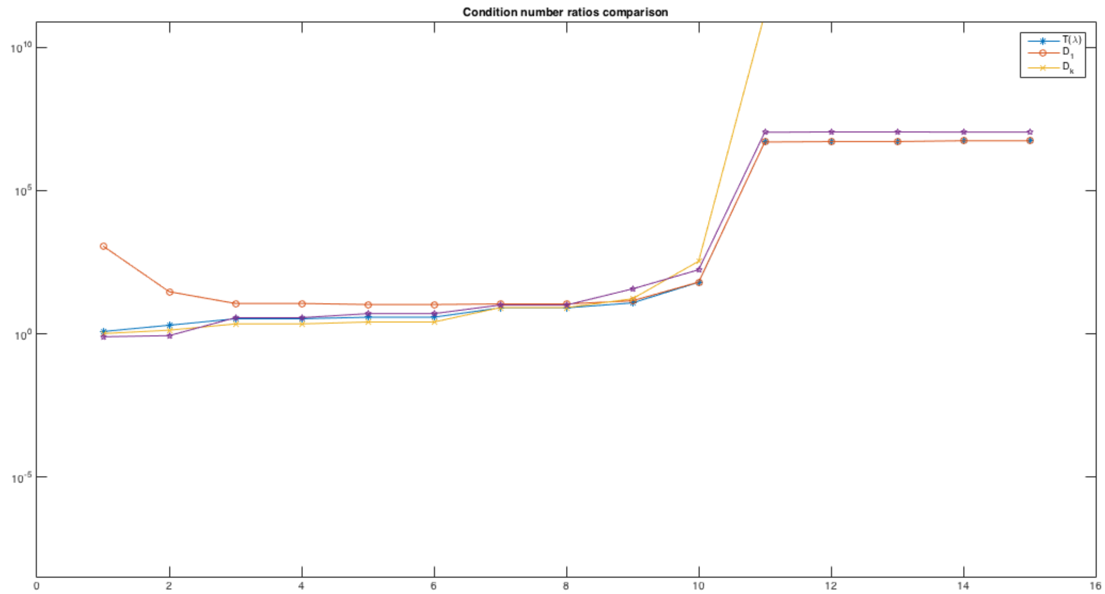

FIGURE 8. Condition number comparison after scaling the same P.

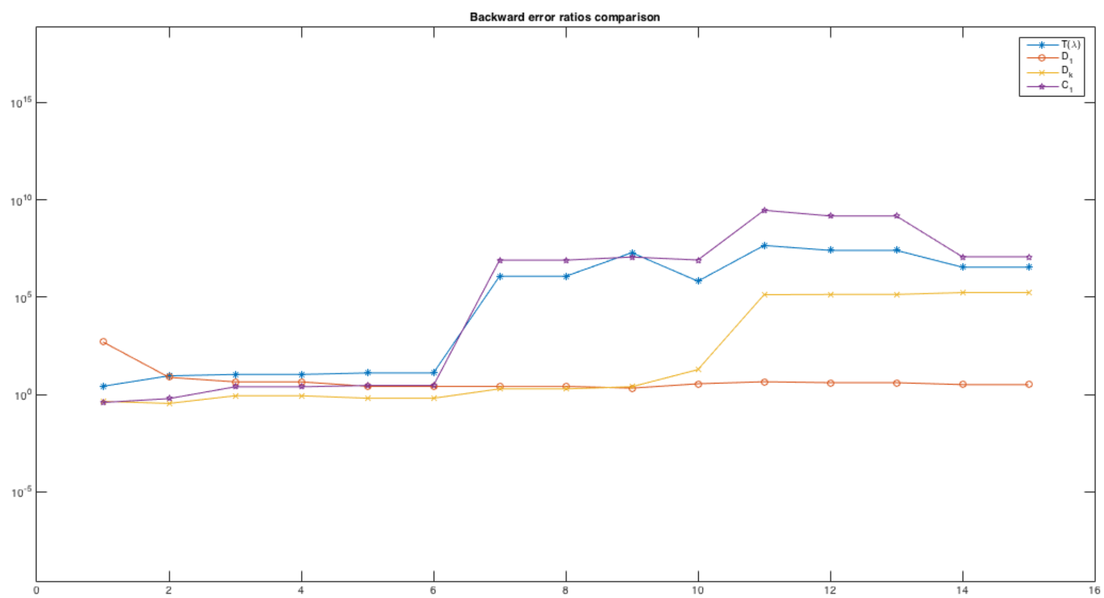

FIGURE 9. Backward error comparison before scaling.

It was mentioned that good theoretical bounds for the condition number and backward error of T could not be found for the case in which k is even. Though poor bounds do not necessarily imply poor results, numerical experiments for k even demonstrate poor behavior for condition number ratios and even less desirable behavior for backward error ratios. Figure 11 shows the graphic for condition number ratios created after scaling a matrix polynomial with $k = 4$ and $n = 15$. Most of the condition number ratios for T are

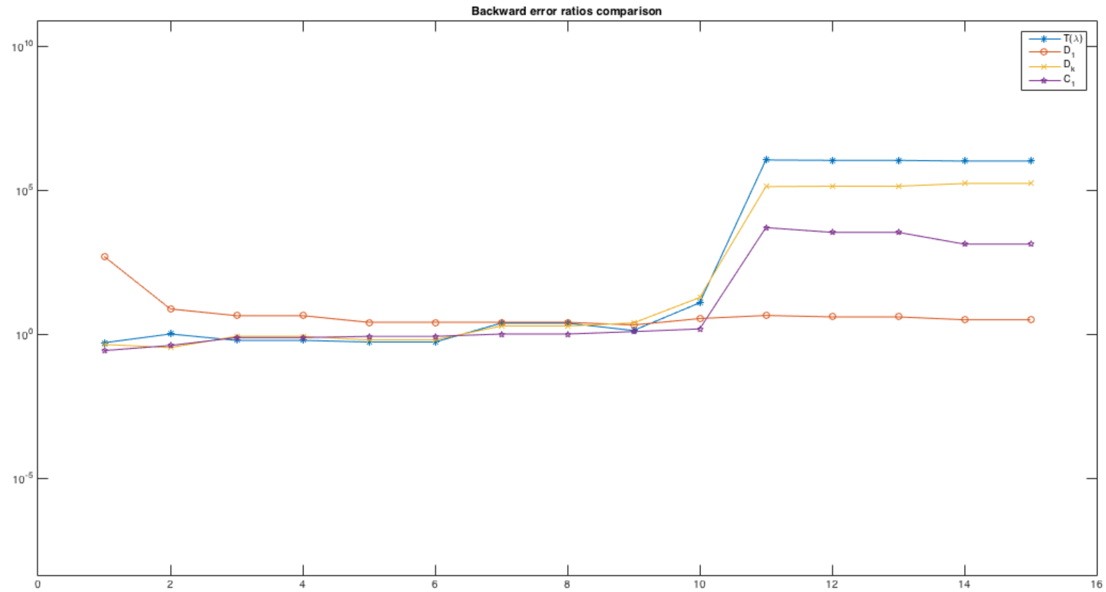

FIGURE 10. Backward error comparison after scaling.

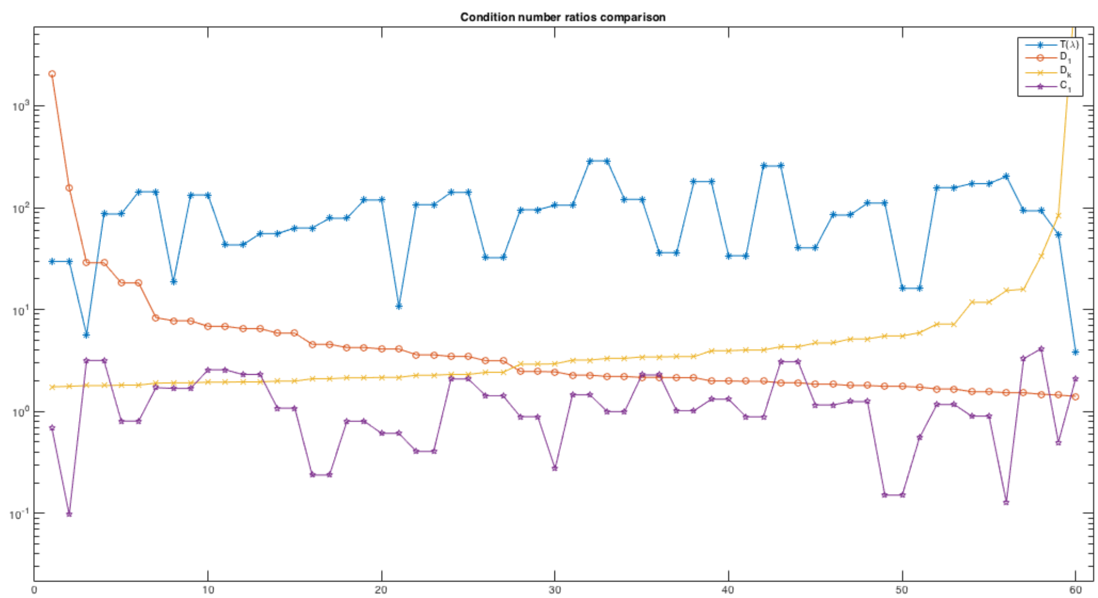

FIGURE 11. Condition number comparison after scaling random P with $k = 4$ and $n = 15$.

close to 10^2, and are not comparable to D_1 or D_k. In this case. C_1 performs much better than T with the furthest ratios close to 10^{-1}. Table 2 shows the results for the backward error for T.

So far results for randomly generated matrix polynomials have been shown, but we also ran experiments on well-known applied problems found in literature. The two problems

TABLE 2. Backward error results for a random matrix polynomial with $k = 4$ and $n = 15$.

Problem:	k=4 n=15	
	Unscaled	Scaled
$\lvert\delta_{min}\rvert$	0.095	0.095
$\lvert\delta_{max}\rvert$	20.76	20.76
$\min\{\eta_P/\eta_{D_1}\}$	0.477	0.477
$\max\{\eta_P/\eta_{D_1}\}$	849.35	849.35
$\min\{\eta_P/\eta_{D_k}\}$	0.448	0.448
$\max\{\eta_P/\eta_{D_k}\}$	9348.03	9348.03
$\min\{\eta_P/\eta_T\}$	7.236e+14	5.331e+14
$\max\{\eta_P/\eta_T\}$	4.744e+16	4.195e+15
$\min\{\eta_P/\eta_{C_1}\}$	1.13	0.47
$\max\{\eta_P/\eta_{C_1}\}$	10.597	1.08

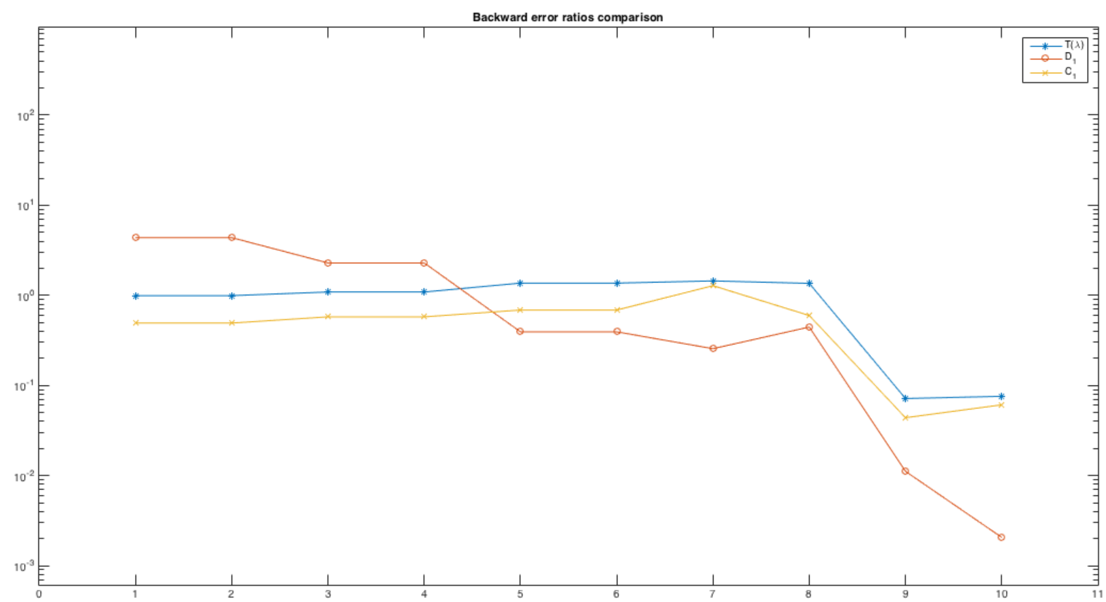

FIGURE 12. Backward error comparison after scaling Relative Pose 5pt Problem.

we investigated were the "Relative Pose 5pt Problem", which uses a matrix polynomial of degree $k = 3$ and with $n = 10$, and the "Plasma-drift Problem", which uses a matrix polynomial of degree $k = 3$ and with $n = 128$. Both of these problems appear in [6]. Figure 12 and Figure 13 show the scaled backward error results for the "Relative Pose 5pt Problem" and the "Plasma-drift Problem" respectively, and Table 3 shows more data related to these Figures.

Since A_k is nonsingular for the matrix polynomial in the "Relative Pose 5pt Problem," D_k is not a linearization for P. For this reason, Figure 13 and Table 3 do not show data

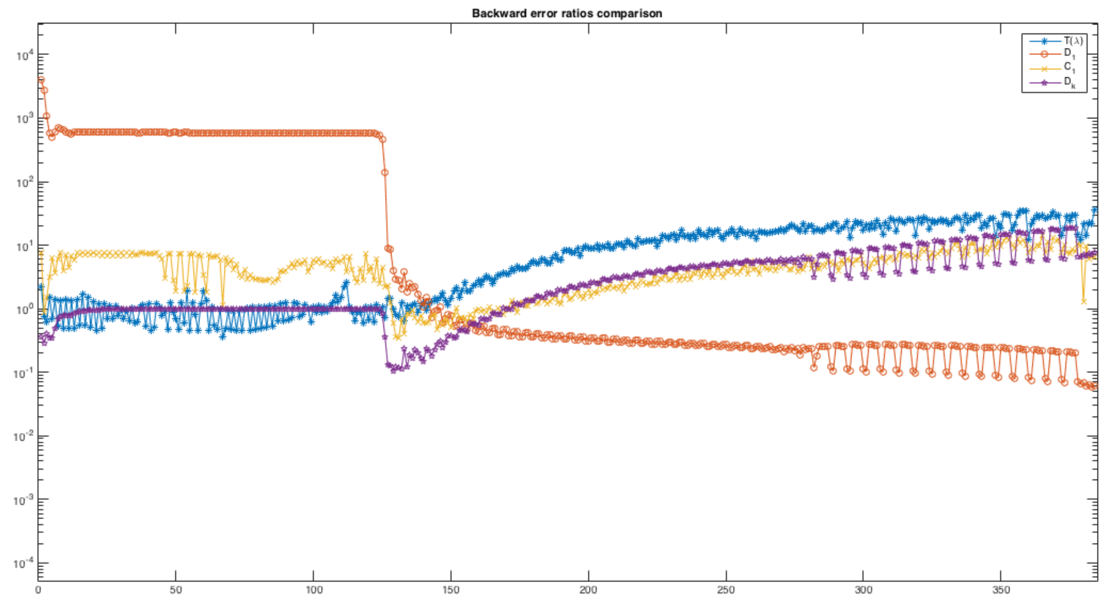

FIGURE 13. Backward error comparison after scaling Plasma Drift Problem.

TABLE 3. Backward error results for two problems in NLEVP: A collection of nonlinear eigenvalue problems [6].

Problem:	Relative Pose 5pt		Plasma-drift	
	Unscaled	Scaled	Unscaled	Scaled
$\lvert\delta_{min}\rvert$	0.573	0.573	0.028	0.028
$\lvert\delta_{max}\rvert$	29.71	29.71	11.745	11.745
$\lvert\lvert A_3\rvert\rvert_2$	0.708	0.336	12.698	0.0103
$\lvert\lvert A_2\rvert\rvert_2$	1.29	0.614	5.304	0.0043
$\lvert\lvert A_1\rvert\rvert_2$	1.88	0.891	1233.03	1.0
$\lvert\lvert A_0\rvert\rvert_2$	2.11	1	123.23	0.0999
ρ_2	0.336	0.336	0.0008	0.0103
ρ_1'	6.263	2.974	119730.83	97.103
Bounds for back. err.	[0.333, 197.07]	[0.333,9 3.59]	[0.333, 3.77e+6]	[0.333, 3055.3]
$\min\{\eta_P/\eta_{D_1}\}$	0.002	0.002	0.0577	0.0577
$\max\{\eta_P/\eta_{D_1}\}$	4.372	4.372	3980.4	3980.4
$\min\{\eta_P/\eta_{D_k}\}$	N/A	N/A	0.104	0.104
$\max\{\eta_P/\eta_{D_k}\}$	N/A	N/A	18.749	18.749
$\min\{\eta_P/\eta_T\}$	0.033	0.072	0.857	0.359
$\max\{\eta_P/\eta_T\}$	1.18	1.45	190738.36	35.28
$\min\{\eta_P/\eta_{C_1}\}$	0.112	0.044	13.79	0.341
$\max\{\eta_P/\eta_{C_1}\}$	1.19	1.28	1883.16	14.24

for D_k under "Relative Pose 5pt". Interestingly, $T(\lambda)$ outperforms D_1 for $|\delta| \leq 1$ and serves as a valid substitute for D_k.

In the "Plasma-drift Problem", A_k and A_0 are nonsingular, and so we analyze the behavior of D_1 and D_k(since they are linearizations). In this problem, there is more evidence to believe that dispersed coefficient norms in P produce poor behavior in T and C_1. In Table 3, we see that the norms of the coefficients of P are not close together. In Figure 13, we notice that T and C_1 behave poorly and seem to mimic the behavior of D_k. Again, this is still an open topic for investigation and should be verified theoretically.

6. Conclusion

In this paper, we examined the conditioning and backward error of the Mackey pencil and compared this pencil's properties with well-known pencils studied intensivel in literature such as D_1, D_k, and C_1. Through our proofs of the backward error bounds for the Mackey pencil and through our numerical experiments, we have presented strong evidence that the Mackey pencil produces optimal condition number and backward error ratios when the degree of P is odd and when the norms of the coefficients of P are close together. The experiments discussed in this report also imply that the optimality of T does not depend on the modulus of the eigenvalue when norms of the coefficients of P are close. Based on poor theoretical bounds for T for the case when k is even, it was conjectured that when k is even, T numerically behaves poorly and numerical experiments support this conjecture. Proving these observations theoretically is open and is a topic of interest. Other open questions include: the relationship between similar coefficient norms and the numerical behavior of T and C_1, calculating the backward error bounds for left eigenvectors of P and for a linearization, and applying different scalings in order to improve the numerical behavior of T under more general conditions. We would like to thank Maribel Bueno for advising us this summer and supporting us throughout the REU program.

References

[1] N.Higham, R-C.Li, F.Tisseur. *Backward Error of Polynomial Eigenproblems Solved by Linearization,*, SIAM J. Matrix Anal. Appl, 29,(2007) pp. 1218-1240.

[2] M.Bueno, F.Dopico, S.Furtago. *Conditioning of Block-Symmetric Fiedler-like Linearizations of Matrix Polynomials*, Pre-print.(2015) pp. 1-51.

[3] F.Tisseur. *Backward Error and Condition of Polynomial Eigenvalue Problems*, Linear Algebra Appl., 309 (2000), pp.339-361.

[4] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. *Vector Spaces of Linearizations for Matrix Polynomials*, SIAM J. Matrix Anal. Appl., 28(4), (2006) pp. 971-1004.

[5] N.J. Higham, D.S. Mackey, F.Tisseur. *The Conditioning of Linearizations of Matrix Polynomials,* SIAM J. Matrix Anal. Appl., 28(4), (2006) pp. 1005-1028.

[6] Betcke, Timo and Higham, Nicholas J and Mehrmann, Volker and Schröder, Christian and Tisseur, Françoise *NLEVP: A collection of nonlinear eigenvalue problems*, ACM Transactions on Mathematical Software (TOMS), 39(2), (2013).