ACCURATE SOLUTIONS OF POLYNOMIAL EIGENVALUE PROBLEMS

YILING YOU, JOSE ISRAEL RODRIGUEZ, AND LEK-HENG LIM

Abstract. Quadratic eigenvalue problems (QEP) and more generally polynomial eigenvalue problems (PEP) are among the most common types of nonlinear eigenvalue problems. Both problems, especially the QEP, have extensive applications. A typical approach to solve QEP and PEP is to use a linearization method to reformulate the problem as a higher dimensional linear eigenvalue problem. In this article, we use homotopy continuation to solve these nonlinear eigenvalue problems without passing to higher dimensions. Our main contribution is to show that homotopy method produces substantially more accurate results, and finds all eigenvalues with a certificate of correctness via Smale's α-theory. To explain the superior accuracy, we show that the nonlinear eigenvalue problem we solve is better conditioned than its reformulated linear eigenvalue problem, and our homotopy continuation algorithm is more stable than QZ algorithm — theoretical findings that are borne out by our numerical experiments. Our studies provide yet another illustration of the dictum in numerical analysis that, for reasons of conditioning and stability, it is sometimes better to solve a nonlinear problem directly even when it could be transformed into a linear problem with the same solution mathematically.

1. Introduction

The study of polynomial eigenvalue problems (PEPs) is an important topic in numerical linear algebra. Such problems arise in partial differential equations and in various scientific and engineering applications (more on these below). A special case of PEP that has been extensively and thoroughly studied is the *quadratic eigenvalue problem* (QEP), often formulated as below, where we follow the notations in [27].

Problem 1.1 (QEP)**.** Let $M, C, K \in \mathbb{C}^{n\times n}$. The QEP corresponding to these matrices is to determine all solutions (x, λ) to the following equation

$$Q(\lambda)x = 0 \quad \text{where} \quad Q(\lambda) := \lambda^2 M + \lambda C + K. \tag{1}$$

The dimension of the QEP is n. We call a solution (x, λ) an eigenpair, λ an eigenvalue, and x an eigenvector.

More generally, we may consider *polynomial eigenvalue problems* (PEP), of which the QEP is the special case when $m = 2$.

Problem 1.2 (PEP)**.** Let $A_0, A_1, \dots, A_m \in \mathbb{C}^{n\times n}$. The PEP corresponding to these matrices is to determine all solutions (x, λ) to the equation

$$P(\lambda)x = 0 \quad \text{where} \quad P(\lambda) := \lambda^m A_m + \lambda^{m-1} A_{m-1} + \cdots + A_0. \tag{2}$$

Here m is the degree of the PEP and again the solutions are called eigenpairs. A general PEP has mn eigenpairs (we use 'general' in the sense of algebraic geometry; those unfamiliar with this usage may think of it as 'random').

Well-known examples where such problems arise include the acoustic wave problem [7, 10], which gives a QEP, and the planar waveguide problem [7, 25], which gives a PEP of higher degree; they

1991 *Mathematics Subject Classification.* 65H20, 65H17, 65H10, 65G20, 35P30 .

Key words and phrases. Homotopy continuation method, polynomial eigenvalue problem, quadratic eigenvalue problem, Shub–Smale's α-theory .

are discussed in section 9.1 and section 9.2 respectively. The encyclopedic survey [27] contains many more examples of QEPs.

There are several methods for solving the QEP, including iterative methods such as Arnoldi method [4], Jacobi–Davidson method [23, 24], and linearization method. However the existing methods invariably suffer from some form of inadequacies: Either (i) they do not apply to PEP of arbitrary degree, or (ii) they require matrices with special structures, or (iii) they only find the largest or smallest eigenvalues. Essentially the only existing method that potentially avoids these inadequacies is the linearization method, as it is based on a reduction to a usual (linear) eigenvalue problem. With this consideration, the linearization method forms the basis for comparison with our proposed method, which does not suffer from any of the aforementioned inadequacies. We test our method with the software Bertini [5] and compare numerical results obtained with those obtained using the linearization method described in section 2.

Our main contribution is to propose the use of homotopy method to directly solve a nonlinear problem and find *all* eigenpairs of a PEP or QEP, suitably adapted to take advantage of the special structures of these problems. Since this article is written primarily for numerical analysts, we would like to highlight that the study of *homotopy method for solving a system of multivariate polynomial equations* (as opposed to systems involving non-algebraic or transcendental functions) has undergone enormous progress within the past decade — both theoretically, with the import of powerful results from complex algebraic geometry, and practically, with the development of new softwares implementing greatly improved algorithms. Some of the recent milestones include the resolution of Smale's 17th Problem [9] and finding all tens of millions of solutions of kinematic problems in biologically-inspired linkage design [20]. Homotopy method has been used for symmetric EVPs and GEPs in the traditional, non-algebraic manner; for instance, [19, 29] rely on Raleigh quotient iterations rather than Newton's method as the corrector method and their results are not certified.

Fortunately for us, PEP and QEP fall in this realm of algebraic problems — they are special systems of multivariate polynomial equations. In fact, we will see that when formulated as such a system of polynomial equations, the PEP is better conditioned than its alternative formulation as a generalized eigenvalue problem obtained using linearization. Our numerical experiments also show that as the dimension n crosses a threshold of around 20, homotopy method begins to significantly outperform linearization method in terms of the normwise backward errors. Moreover, our homotopy method approach is not limited to a specific class of matrices but applies robustly to a wide range of matrices — dense sparse, badly scaled, etc. In addition, we will see that our homotopy method approach makes it perfect for parallelization. Perhaps most importantly, a unique feature of our approach is that our outputs are certifiable using Smale's α-theory.

2. Linearization method

A popular approach to solve QEP and PEP is the *linearization method*. The goal of which is to transform a PEP into a generalized eigenvalue problem (GEP) [27] involving an equivalent linear λ-matrix $A - \lambda B$. A $2n \times 2n$ linear λ-matrix is a *linearization* of $Q(\lambda)$ [12, 17] if

$$\begin{bmatrix} Q(\lambda) & 0 \\ 0 & I_n \end{bmatrix} = E(\lambda)(A - \lambda B)F(\lambda) \tag{3}$$

where $E(\lambda)$ and $F(\lambda)$ are $2n \times 2n$ λ-matrices with constant nonzero determinants. The eigenvalues of the quadratic λ-matrix $Q(\lambda)$ and the linear λ-matrix $A - \lambda B$ coincide.

Linearizations of PEP are not unique, but two linearizations commonly used in practice are the *first* and *second companion forms*:

$$\text{L1:} \quad \begin{bmatrix} 0 & N \\ -K & -C \end{bmatrix} - \lambda \begin{bmatrix} N & 0 \\ 0 & M \end{bmatrix}, \qquad \text{L2:} \quad \begin{bmatrix} -K & 0 \\ 0 & N \end{bmatrix} - \lambda \begin{bmatrix} C & M \\ N & 0 \end{bmatrix}, \tag{4}$$

where N can be any nonsingular $n \times n$ matrix. The choice between the two companion forms eq. (4) usually depends on the nonsingularity of M and K [1].

More generally, a PEP eq. (2) can also be transformed into a GEP of dimension mn. The most common linearization is called the *companion linearization* where A and B are:

$$A = \begin{bmatrix} A_0 & & & & \\ & I & & & \\ & & \ddots & & \\ & & & I & \\ & & & & I \end{bmatrix}, \qquad B = \begin{bmatrix} -A_1 & -A_2 & \cdots & \cdots & -A_m \\ I & 0 & \cdots & \cdots & 0 \\ 0 & I & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & I & 0 \end{bmatrix}. \tag{5}$$

We will call such a GEP corresponding to a PEP its *companion GEP*. One reason for the popularity of the companion linearization is that eigenvectors of PEP can be directly recovered from eigenvectors of this linearization (see [16] for details).

3. Stability

To assess the stability and quality of the numerical methods in this article, we use the backward error for PEP as defined and discussed in [26]. Since we do not run into zero or infinite eigenvalues in all the numerical experiments and application problems in this article, our backward errors are always well-defined. For an approximate eigenpair $(\widetilde{x}, \widetilde{\lambda})$ of $Q(\lambda)$, the *normwise backward error* is defined as

$$\eta(\widetilde{x}, \widetilde{\lambda}) := \min \big\{ \epsilon : \big(Q(\widetilde{\lambda}) + \Delta Q(\widetilde{\lambda})\big)\widetilde{x} = 0,\ \|\Delta M\| \le \epsilon\alpha_2,\ \|\Delta C\| \le \epsilon\alpha_1,\ \|\Delta K\| \le \epsilon\alpha_0 \big\},$$

where $\Delta Q(\lambda)$ denotes the perturbation

$$\Delta Q(\lambda) = \lambda^2 \Delta M + \lambda \Delta C + \Delta K,$$

and α_k's are nonnegative parameters that allow freedom in how perturbations are measure, e.g., in an absolute sense ($\alpha_k \equiv 1$) or a relative sense ($\alpha_2 = \|M\|$, $\alpha_1 = \|C\|$, $\alpha_0 = \|K\|$). In [26], the backward error is shown to be equal to the following scaled residual,

$$\eta(\widetilde{x}, \widetilde{\lambda}) = \frac{\big\|Q(\widetilde{\lambda})\widetilde{x}\big\|}{\big(|\widetilde{\lambda}|^2\alpha_2 + |\widetilde{\lambda}|\alpha_1 + \alpha_0\big)\|\widetilde{x}\|}, \tag{6}$$

which has the advantage of being readily computable.

We adopt usual convention and say that a numerical method is *numerically stable* if all of the computed eigenpairs have backwards errors that are of the same order as the unit roundoff. It is known that a numerically stable reduction of GEP may be obtained by computing the *generalized Schur decomposition*:

$$W^*AZ = S, \qquad W^*BZ = T, \tag{7}$$

where W and Z are unitary and S and T are upper triangular. Then $\lambda_i = s_{ii}/t_{ii}, i = 1, \ldots, 2n$, with the convention that $s_{ii}/t_{ii} = \infty$ whenever $t_{ii} = 0$.

The QZ algorithm [21, 13] is numerically stable for computing the decomposition eq. (7) and solving the GEP, but it is not stable for the solution of the QEP. To be more precise, one may solve a QEP via the linearization followed by QZ algorithm applied to resulting GEP. An approximate eigenvector $\widetilde{x}$ to the QEP can be recovered from either the first n components or the last n components of the approximate eigenvector of the GEP computed by the QZ algorithm, a $2n$-vector $\widetilde{\xi}^{\mathsf{T}} = (\widetilde{x}_1^{\mathsf{T}}, \widetilde{x}_2^{\mathsf{T}})$ that yields the smallest backward error eq. (6). Nevertheless, this method is in general unstable [26]. Even though its backward stability is not guaranteed, it is still the method-of-choice when all eigenpairs are desired and if the coefficient matrices have no special structure and are not too large.

The discussions above regarding QEP extends to PEP. For an approximate eigenpair $(\widetilde{x}, \widetilde{\lambda})$ of $P(\lambda)$, the normwise backward error is defined in [26] as

$$\eta(\widetilde{x}, \widetilde{\lambda}) := \min\{\epsilon : \big(P(\widetilde{\lambda}) + \Delta P(\widetilde{\lambda})\big)\widetilde{x} = 0,\ \|\Delta A_k\| \le \epsilon \|E_k\|,\ k = 0, \dots, m\}$$

where $\Delta P(\lambda)$ denotes the perturbation

$$\Delta P(\lambda) = \lambda^m \Delta A_m + \lambda^{m-1} \Delta A_{m-1} + \cdots + \Delta A_0$$

and the matrices $E_k, k = 0, \dots, m$ are arbitrary and represent tolerances against which the perturbations ΔA_k to A_k will be measured. As in the case of QEP, the normwise backward error of PEP may be computed [26] via the following expression

$$\eta(\widetilde{x}, \widetilde{\lambda}) = \frac{\big\|P(\widetilde{\lambda})\widetilde{x}\big\|}{\big(\|E_0\| + |\widetilde{\lambda}|\|E_1\| + |\widetilde{\lambda}|^2\|E_2\| \cdots + |\widetilde{\lambda}|^m\|E_m\|\big)\|\widetilde{x}\|}. \tag{8}$$

4. Conditioning

We will see examples in section 8 where a PEP as formulated in eq. (2) is far better conditioned than its companion GEP in eq. (5) obtained using linearization. Thus for accurate solutions, it is better to solve the nonlinear PEP problem directly rather than to first transform it into a mathematically equivalent GEP. As an analogy, this is much like solving a least squares problem directly versus solving its normal equations.

To characterize the sensitivity of solutions to these problems, we will describe their condition numbers in this section. As previously mentioned, in all of our numerical experiments and applications, we will not encounter zero or infinite eigenvalues. So in principle we just need a notion of condition number [26] that is based on the nonhomogeneous matrix polynomial $P(\lambda)$ in eq. (2). However, since the MATLAB `polyeig` function that we use for comparison (see section 7.1) implements a more general notion of condition number [11, 16] based on the homogenized version (permitting zero and infinite eigenvalues), we will briefly review the homogenized eigenvalue problem and define its condition number accordingly.

We rewrite the polynomial matrix $P(\lambda)$ in eq. (2) in homogeneous form

$$P(\lambda_0, \lambda_1) = \sum_{i=0}^{m} \lambda_0^i \lambda_1^{m-i} A_i \tag{9}$$

and consider eigenvalues as pairs $(\lambda_0, \lambda_1) \neq (0, 0)$ that are solutions of the equation $\det P(\lambda_0, \lambda_1) = 0$. Let $T_{(\lambda_0,\lambda_1)}\mathbb{P}_1$ denote the tangent space at (λ_0, λ_1) to $\mathbb{P}_1$, the projective space of lines through the origin in $\mathbb{C}^2$. A condition operator is defined in [11] as $K(\lambda_0, \lambda_1) : (\mathbb{C}^{n\times n})^{m+1} \to T_{(\lambda_0,\lambda_1)}\mathbb{P}_1$ of the eigenvalue (λ_0, λ_1) as the differential of the map from the $(m+1)$-tuple $(A_0, \dots, A_m)$ to (λ_0, λ_1) in projective space. If we write a representative of an eigenvalue (λ_0, λ_1) as a row vector $[\lambda_0, \lambda_1] \in \mathbb{C}^{1\times 2}$, the condition number $\kappa_P(\lambda_0, \lambda_1)$ can be defined as a norm of the condition operator [16],

$$\kappa_P(\lambda_0, \lambda_1) := \max_{\|\Delta A\| \le 1} \frac{\|K(\lambda_0, \lambda_1)\Delta A\|_2}{\|[\lambda_0, \lambda_1]\|_2},$$

for any arbitrary norm on ΔA. We will choose the norm on $(\mathbb{C}^{n\times n})^{m+1}$ to be the ω-weighted Frobenius norm

$$\|A\| := \|(A_0, \dots, A_m)\| = \big\|[\omega_0^{-1}A_0, \dots, \omega_m^{-1}A_m]\big\|_F,$$

with weights $\omega_i > 0$, $i = 1, \dots, m$. If we define the operators $\partial_{\lambda_0} := \partial/\partial\lambda_0$ and $\partial_{\lambda_1} := \partial/\partial\lambda_1$, then the normwise condition number $\kappa_P(\lambda_0, \lambda_1)$ of a simple eigenvalue (λ_0, λ_1) is given by

$$\kappa_P(\lambda_0, \lambda_1) = \left(\sum_{i=0}^{m} |\lambda_0|^{2i}|\lambda_1|^{2(m-i)}\omega_i^2\right)^{1/2} \frac{\|y\|_2\|x\|_2}{|y^*(\overline{\lambda}_1\partial_{\lambda_0}P - \overline{\lambda}_0\partial_{\lambda_1}P)|_{(\lambda_0,\lambda_1)}x|},$$

where x, y are the corresponding right and left eigenvector respectively [16].

Remark 4.1. Homogenization allows one to better handle eigenvalues at infinity. In general, the characteristic polynomial is the determinant of the matrix polynomial and takes the form $\det(A_m)\lambda_1^{mn} + \cdots + \det(A_0)\lambda_0^{mn}$. Therefore, $P(\lambda_0, \lambda_1)$ has mn finite eigenvalues when the matrix A_m is nonsingular. However, when $\det(A_m) = 0$, the characteristic polynomial has degree $r < mn$ and there are r finite eigenvalues and $mn - r$ infinite eigenvalues. Those infinite eigenvalues correspond to $\lambda_0 = 0$. None of the numerical experiments we consider in this article has eigenvalues at infinity.

5. Certification

A major advantage of our homotopy method approach for solving the PEP is that we may use Smale's α-theory (also known as Shub–Smale's α-theory, see [8, Chapter 8]) to certify that the Newton iterations will converge quadratically to an eigenpair. In numerical analysis lingo, this means we can control the *forward error*, not just the backward error.

To apply Smale's α-theory, we view the PEP as a collection of n polynomial equations eq. (2) and an affine linear constraint $L(x) = 0$, which yields a polynomial map $f = (f_0, \ldots, f_n) : \mathbb{C}^{n+1} \to \mathbb{C}^{n+1}$. The affine linear polynomial $L(x)$ is chosen randomly as described in the numerical experiments section 7.

Let $\mathcal{V}(f) := \{\zeta \in \mathbb{C}^{n+1} : f(\zeta) = 0\}$ and let $Df(z)$ be the Jacobian matrix of the system f at $z = (x, \lambda)$. Consider the map $N_f : \mathbb{C}^{n+1} \to \mathbb{C}^{n+1}$ defined by

$$N_f(z) := \begin{cases} z - Df(z)^{-1} f(z) & \text{if } Df(z) \text{ is invertible,} \\ z & \text{otherwise.} \end{cases}$$

We say the point $N_f(z)$ is the *Newton iteration of f starting at z*. The kth Newton iteration of f starting at x is denoted by $N_f^k(z)$. Now we define precisely what we mean by an approximate solution to f.

Definition 5.1. [8, p. 155] With the notation above, a point $z \in \mathbb{C}^{n+1}$ is an *approximate solution to f with associated solution $\zeta \in \mathcal{V}(f)$*, if for every $k \in \mathbb{N}$,

$$\left\| N_f^k(z) - \zeta \right\| \leq \left(\frac{1}{2}\right)^{2^k - 1} \|z - \zeta\|, \tag{10}$$

where the norm is the 2-norm $\|z\| = \left(|z_1|^2 + \cdots + |z_{n+1}|^2\right)^{1/2}$.

Smale's α-theory gives a condition for when a given point z is an approximate solution to $f = 0$ using the following constants when $Df(z)$ is invertible:

$$\begin{aligned} \alpha(f, z) &:= \beta(f, z)\gamma(f, z), \\ \beta(f, z) &:= \|z - N_f(z)\| = \left\| Df(z)^{-1} f(z) \right\|, \\ \gamma(f, z) &:= \sup_{k \geq 2} \left\| \frac{Df(z)^{-1} D^k f(z)}{k!} \right\|^{1/(k-1)}. \end{aligned}$$

The following theorem from [15] is a version of Theorem 2 in [8, p. 160] and it provides a certificate that a point z is an approximate solution to $f = 0$.

Theorem 5.2. *If $f : \mathbb{C}^{n+1} \to \mathbb{C}^{n+1}$ is a polynomial system and $z \in \mathbb{C}^{n+1}$, with*

$$\alpha(f, z) < \frac{13 - 3\sqrt{17}}{4} \approx 0.157671,$$

then z is an approximate solution to $f = 0$.

The quantity $\gamma(f,z)$ is difficult to compute in general. In [15], this quantity is bounded in terms of an alternative that is more readily computable. Define

$$\mu(f,z) := \max\left\{1, \|f\| \cdot \|Df(z)^{-1}\Delta_{(d)}(z)\|\right\}, \tag{11}$$

where $\Delta_{(d)}(z)$ is an $(n+1)\times(n+1)$ diagonal matrix with ith diagonal entry $d_i^{1/2}(1+\|z\|^2)^{(d_i-1)/2}$ and $d_i := \deg(f_i)$. For a polynomial $g = \sum_{|\nu|\le d} a_\nu z^\nu$, we define the norm $\|g\|$ according to [15],

$$\|g\|^2 := \sum_{|\nu|\le d} |a_\nu|^2 \frac{\nu!(d-|\nu|)!}{d!}. \tag{12}$$

This can extended to polynomial system f via

$$\|f\|^2 := \sum_{i=0}^{n} \|f_i\|^2 \tag{13}$$

where each $\|f_i\|^2$ is as defined in eq. (12).

With the notations above, [15] gives the following bound:

$$\gamma(f,z) \le \frac{\mu(f,z) d_{\max}^{3/2}}{2(1+\|z\|^2)^{1/2}}, \tag{14}$$

where $d_{\max}$ is the maximal degree of a polynomial in the system.

When f comes from a PEP, theorem 5.3 follows from eq. (11)–eq. (13) and a straightforward calculation of $Df(z)^{-1}\Delta_{(d)}(z)$. This yields the value of $\mu(f,z)$ and thus a bound for $\gamma(f,z)$ for the PEP. We will rely on our bound in the acoustic wave problem fig. 9 to certify the eigenvalues of the PEP therein.

Corollary 5.3. *Let $f=0$ denote the PEP in theorem 1.2, and let L denote the affine constraint on the x. Then,*

$$\mu(f,x,\lambda) = \left[\|L\|^2 + \sum_{k=0}^{m} \frac{k!(m-k)!}{(m+1)!}\|A_k\|_F^2\right]^{1/2} \left\| \begin{bmatrix} \frac{P(\lambda)}{\sqrt{m+1}(1+\|[x^\intercal,\lambda]\|^2)^{m/2}} & P'(\lambda)x \\ \frac{(\nabla_x L)^\intercal}{\sqrt{m+1}(1+\|[x^\intercal,\lambda]\|^2)^{m/2}} & 0 \end{bmatrix}^{-1} \right\|. \tag{15}$$

6. Homotopy method for the polynomial eigenvalue problem

We now describe our approach of using *homotopy continuation method*, also called *homotopy method*, to solve PEPs. Our goal is to find all eigenpairs.

Homotopy method deforms solutions of a *start system* $S(z)=0$ that is easy to solve to solutions of a *target system* $T(z)=0$ that is of interest. More precisely, a *straight-line homotopy* with *path parameter* t is defined as

$$H(z,t) := (1-t)S(z) + tT(z), \quad t\in[0,1]. \tag{16}$$

When $t=0$ or $t=1$, the system $H(z,t)=0$ gives the start system $H(z,0)=S(z)=0$ or the target system $H(z,1)=T(z)=0$ respectively.

A start system for the homotopy (16) is said to be *chosen correctly* if the following properties [18] hold:

(i) the solution set of the start system $S(z)=0$ are known or easy to obtain;
(ii) the solution set of $H(z,t)=0$ for $0\le t<1$ consists of a finite number of smooth paths, each parametrized by t in $[0,1)$;
(iii) for each isolated solution of the target systems $T(z)=0$, there is some path originating at $t=0$, that is, a solution of the start systems $S(z)=0$.

Example 6.1. For an example of homotopy familiar to numerical linear algebraists, consider $P(\lambda) = I - \lambda A$, i.e., $m = 1$, $A_0 = I$, and $A_1 = -A$. Let D be the diagonal matrix whose diagonal entries are the diagonal entries of A. The proof of strengthened Gershgorin Circle Theorem [28] exactly illustrates a straight-line homotopy path $H(t) = (1-t)D + tA$, $t \in [0,1]$. Note however that such a D would be a poor choice for us as the start system for the homotopy is not guaranteed to be chosen correctly: The solution set $H(t) = 0$ for $0 \leq t < 1$ need not consist of a finite number of smooth paths.

In this article, we consider a target system to solve the PEP in (2) given by

$$T(z) = \begin{bmatrix} P(\lambda)x \\ L(x) \end{bmatrix} \tag{17}$$

where $z = (x, \lambda)$ and $L(x)$ is a general affine linear polynomial, chosen randomly so that we have a polynomial system $T : \mathbb{C}^{n+1} \to \mathbb{C}^{n+1}$ as defined in section 5. The requirement that $L(x) = 0$ also fixes the scaling indeterminacy in the polynomial eigenvector x. In more geometric language, x is a point in projective space and the random choice of $L(x)$ restricts this space to a general affine chart so that the eigenvectors are not at infinity. If one instead had chosen $L(x) = x_1 - 1$, then eigenvectors with a first coordinate equal to zero would not be solutions to the system.

There is an obvious choice of start system — we choose random diagonal matrices D_i's to replace the coefficient matrices A_i's in $P(\lambda)$:

$$S(z) = \begin{bmatrix} (\lambda^m D_m + \lambda^{m-1} D_{m-1} + \cdots + D_0)x \\ L(x) \end{bmatrix} \tag{18}$$

One observes that $S(x, \lambda) = 0$ is a polynomial system with linear products of (x, λ). Specifically, let $d_{j,i}$ denote the ith diagonal entry of D_j and let x_i be the ith entry of x. Then we may factor the univariate polynomial

$$\big(\lambda^m d_{m,i} + \lambda^{m-1} d_{m-1,i} + \cdots + d_{0,i}\big)x_i = d_{m,i}(\lambda - r_{m,i})(\lambda - r_{m-1,i}) \cdots (\lambda - r_{1,i})x_i,$$

where $r_{j,i} \in \mathbb{C}$, $j = 1, \ldots, m$, are the roots of the respective monic polynomial (obtained via, say, the Schur decomposition of its companion matrix). The solutions to (18) are then simply

$$\lambda = r_{j,i}, \qquad x_k = 0 \text{ for all } k \neq i, \qquad L(0, \ldots, 0, x_i, 0, \ldots, 0) = 0,$$

for every $j = 1, 2, \ldots, m$ and $i = 1, 2, \ldots, n$.

From the known eigenpairs at $t = t_0$, solutions at $t = t_0 + \Delta t$ can be obtained by iterative methods [2, 3]. This process is called path tracking. The steps in iteration are repeated until t reaches 1 or t is sufficiently close to 1 by some criteria. The output is regarded as an approximation of a solution to the PEP.

The path tracking becomes numerically unstable as we get close to the space of ill-posed systems, also called the *discriminant locus*. In other words, this occurs when the Jacobian of $H(z, t_0)$ with respect to z has large condition number. Using randomization and adaptive-precision we are able to avoid these ill-conditioned problems in the middle of path tracking. This is because the discriminant locus has real codimension two in the space of coefficients, whereas the homotopy is over a real one-dimensional path.

Example 6.2. A QEP may be expressed as a univariate polynomial root finding problem by taking the determinant of $Q(\lambda)$. Say, $n = 2$, then we have

$$f(\lambda, M, C, K) := \det Q(\lambda) = \det\left(\lambda^2 \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} + \lambda \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} + \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix}\right).$$

Solving this univariate problem is numerically unstable — in fact this is a special case of the resultant method, which is unstable in general [22]. The discriminant locus is the set of $(M, C, K) \in \mathbb{C}^{2\times 2\times 3}$ is such that there exists a λ where the following is satisfied: $f(\lambda, M, C, K) = \partial f/\partial\lambda = 0$.

If we were to express f as a sum of monomials then it is the determinant of a Sylvester matrix. Recall that the entries of a Sylvester matrix of two polynomials are coefficients of the polynomials, which in our case are f and $\partial f/\partial\lambda$, i.e.,

$$\begin{bmatrix} s_4 & s_3 & s_2 & s_1 & s_0 & & \\ & s_4 & s_3 & s_2 & s_1 & s_0 & \\ & & s_4 & s_3 & s_2 & s_1 & s_0 \\ 4s_4 & 3s_3 & 2s_2 & s_1 & & & \\ & 4s_4 & 3s_3 & 2s_2 & s_1 & & \\ & & 4s_4 & 3s_3 & 2s_2 & s_1 & \\ & & & 4s_4 & 3s_3 & 2s_2 & s_1 \end{bmatrix},$$

where s_i's are polynomials in the entries of (M, C, K). The determinant of this matrix is zero when the two polynomials have a common root, i.e.,

$$f(\lambda, M, C, K) = s_4 \cdot \lambda^4 + s_3 \cdot \lambda^3 + s_2 \cdot \lambda^2 + s_1 \cdot \lambda + s_0.$$

So the vanishing of this determinant defines where the problem is ill-conditioned. By taking a randomized start system, the homotopy method will avoid this space.

7. Numerical experiments

In this section, we provide numerical experiments to compare the speed and accuracy of (i) solving the PEP as formulated in eq. (2) with homotopy continuation (henceforth abbreviated as *homotopy method*), and (ii) solving the companion GEP in eq. (5) with QZ algorithm (henceforth abbreviated as *linearization method*). We will also compare the conditioning of the PEP in eq. (2) and its companion GEP in eq. (5). All experiments were performed on a computer running Windows 8 with an Intel Core i7 processor. For the linearization method, we use the implementation in the MATLAB `polyeig` function; for the homotopy method, we implemented it with BERTINI 1.5.1. We tested both methods in serial on a common platform: MATLAB R2016a with the interface BERTINILAB 1.5 [6]. We also tested the homotopy method in parallel using Intel MPI Library 5.1 compiled with intel/16.0 on the Midway1 compute cluster[1] in the University of Chicago Research Computing Center.

7.1. The `polyeig` and `quadeig` functions.

In our numerical experiments, we rely on the `polyeig` function in MATLAB to solve PEPs with linearization method and compute condition numbers. This routine adopts the companion linearization, uses QZ factorization to compute generalized Schur decomposition, recovers the right eigenvectors that has the minimal residual, and gives the condition number for each eigenvalue. While different linearization forms can have widely varying eigenvalue condition numbers [16], `polyeig` does not allow for other choices.

A more recent linearization-based algorithm, `quadeig`, was proposed in [14] for QEPs. While `polyeig` is a built-in MATLAB function, `quadeig` is a third-party program implemented in MATLAB. It incorporates extra preprocessing steps that scale the problem's parameters and choose the linearization with favorable conditioning and backward stability properties. However, we remain in favor of using `polyeig` as the basis of comparison against our homotopy method for solving QEPs as (i) the scaling is redundant since the 2-norms of our random coefficient matrices are all approximately one; (ii) for a uniform comparison, we prefer a method that works for all m but `quadeig` does not extend to higher order PEP since the scaling process cannot be generalized. There will be further discussions of these issues in section 8.

[1] `https://rcc.uchicago.edu/resources/high-performance-computing`

7.2. **Speed comparisons for QEP.** We first test and compare timings for our homotopy method and the linearization method for general quadratic matrix polynomials. The matrix polynomial is generated at random[2] with coefficient matrices M, C, K in eq. (2) having independent entries following the standard complex Gaussian distribution $\mathcal{N}_{\mathbb{C}}(0,1)$, using the `randn` function in MATLAB. The coefficients of $L(x)$ are also chosen randomly in the same way. We performed three sets of numerical experiments:

(i) table 1 gives the elapsed timings for computing eigenpairs with the linearization method for dimensions $n = 2, \ldots, 100$.
(ii) table 2 gives the elapsed timings for computing eigenpairs with the homotopy method for dimensions $n = 2, \ldots, 80$.
(iii) table 3 gives the elapsed timings for computing the eigenpairs with homotopy method in parallel on 20 cores for dimensions $n = 20, 30 \ldots, 100$.

For item i and item ii the timings include both the setup and the solution process; for item iii, the timings only include the solution process running BERTINI in parallel. For each method and each dimension, we run our experiments ten times and record the best, average, median, and worst performance. The conclusions drawn from these experiments are discussed in section 7.5.

TABLE 1. SPEED OF QEP — LINEARIZATION. Elapsed timings (in seconds) for the linearization method with dimensions $n = 2, 3, \ldots, 100$ (not all displayed).

n	# ROOTS	BEST	AVERAGE	MEDIAN	WORST
2	4	2.4550E-4	0.0015	8.0279E-4	0.0082
3	6	3.1406E-4	6.7019E-4	6.0923E-4	0.0013
4	8	3.6414E-4	6.7064E-4	5.9055E-4	0.0011
5	10	4.2818E-4	6.7003E-4	5.2938E-4	0.0013
6	12	4.6144E-4	0.0010	7.3403E-4	0.0034
7	14	5.8542E-4	9.3642E-4	9.1938E-4	0.0015
8	16	6.7327E-4	9.5284E-4	8.5719E-4	0.0013
9	18	7.3280E-4	0.0027	0.0012	0.0158
10	20	9.2533E-4	0.0015	0.0013	0.0028
11	22	9.2246E-4	0.0014	0.0012	0.0027
12	24	0.0011	0.0017	0.0015	0.0027
13	26	0.0012	0.0017	0.0017	0.0025
14	28	0.0016	0.0020	0.0019	0.0025
15	30	0.0015	0.0023	0.0021	0.0035
16	32	0.0020	0.0030	0.0029	0.0050
17	34	0.0025	0.0037	0.0033	0.0058
18	36	0.0029	0.0037	0.0036	0.0056
19	38	0.0032	0.0048	0.0040	0.0091
20	40	0.0031	0.0038	0.0038	0.0048
30	60	0.0089	0.0109	0.0103	0.0164
40	80	0.0114	0.0157	0.0152	0.0220
50	100	0.0225	0.0253	0.0252	0.0295
60	120	0.0328	0.0407	0.0400	0.0511
70	140	0.0418	0.0527	0.0538	0.0624
80	160	0.0642	0.0762	0.0768	0.0850
90	180	0.0867	0.0983	0.0957	0.1191
100	200	0.1177	0.1285	0.1229	0.1504

[2]Results on matrix polynomials arising from actual applications will be presented in section 9.

TABLE 2. SPEED OF QEP — HOMOTOPY IN SERIAL. Elapsed timings (in seconds) for the homotopy method with dimensions $n = 2, 3, \ldots, 80$ (not all displayed).

n	# ROOTS	BEST	AVERAGE	MEDIAN	WORST
2	4	0.3692	0.4185	0.3950	0.6064
3	6	0.4717	0.5509	0.5225	0.7905
4	8	0.5005	1.1083	0.5553	5.9034
5	10	0.6316	0.7809	0.6735	1.0826
6	12	0.7767	1.0083	0.9946	1.2599
7	14	0.8593	1.0333	0.9418	1.4684
8	16	1.0043	1.4300	1.1941	2.6119
9	18	1.2054	1.7269	1.3505	2.8268
10	20	1.4373	1.6836	1.5428	2.9608
11	22	1.6972	2.2203	2.0245	3.3329
12	24	1.9914	2.7784	2.3702	4.6917
13	26	2.1979	3.3690	2.6251	7.5556
14	28	2.6687	3.8757	3.7469	5.3345
15	30	3.1057	5.2769	5.1792	9.3289
16	32	3.4994	7.2866	6.6201	16.3712
17	34	3.8719	6.1690	5.9380	9.3421
18	36	4.9051	7.8015	8.2610	13.4794
19	38	5.4613	10.4376	11.2833	13.4682
20	40	6.4268	14.0531	14.0381	23.9232
30	60	15.2266	32.1909	32.1480	63.8345
40	80	74.4945	103.9896	93.9979	161.6461
50	100	133.0395	244.3541	245.9194	394.0840
60	120	309.0921	532.1712	485.0595	943.0120
70	140	705.9720	1200.4053	1101.9838	1796.4483
80	160	1207.1300	1848.9342	1745.6093	2974.4295

TABLE 3. SPEED OF QEP — HOMOTOPY IN PARALLEL. Elapsed timings (in seconds) for the homotopy method ran in parallel on 20 cores and with dimensions $n = 20, 30, \ldots, 100$.

n	# ROOTS	BEST	AVERAGE	MEDIAN	WORST
20	40	2.5010	2.8091	2.7410	3.5160
30	60	2.5890	5.6856	4.6965	9.4390
40	80	8.3950	12.7100	13.2490	16.3700
50	100	16.1620	22.2175	21.8220	26.9460
60	120	18.2790	24.4677	23.4870	29.1530
70	140	36.5210	53.7870	54.7050	64.7610
80	160	63.6320	83.7379	77.9380	105.2450
90	180	101.4390	142.3974	149.6170	177.7420
100	200	164.3740	196.5620	199.4585	239.5180

7.3. **Accuracy comparisons for QEP.** Next we test and compare the absolute and relative backward errors of the computed eigenpairs corresponding to the smallest and largest eigenvalues for our homotopy method and the linearization method for randomly generated matrix polynomials. For each of the methods, we tabulated the absolute and relative backward errors of the computed eigenpairs corresponding to the smallest and largest eigenvalues with dimension $n = 2, \ldots, 100$. All tests are averaged over ten runs. They are compared side-by-side in table 4 (absolute error) and table 5 (relative error), and graphically in fig. 1 (absolute error) and fig. 2 (relative error). The conclusions drawn from these experiments are presented in section 7.5.

TABLE 4. ACCURACY OF QEP — LINEARIZATION VS HOMOTOPY. Absolute backward errors of computed smallest and largest eigenpairs with dimensions $n = 2, \ldots, 100$.

	LINEARIZATION AVG. BK. ERR.		HOMOTOPY AVG. BK. ERR.	
n	SMALLEST	LARGEST	SMALLEST	LARGEST
2	3.39024E-16	2.01468E-16	2.81624E-16	2.13291E-16
5	1.03302E-15	8.68237E-16	4.98414E-16	3.21280E-16
10	2.34081E-15	2.43521E-15	1.03488E-15	7.96172E-16
20	5.21350E-15	4.81234E-15	1.20237E-15	1.25302E-15
30	8.24479E-15	6.06398E-15	1.31006E-15	3.27204E-15
40	1.18830E-14	8.06968E-15	1.83843E-15	1.56400E-15
50	1.37506E-14	9.77673E-15	2.66411E-15	3.04532E-15
60	1.69521E-14	9.92973E-15	2.37774E-15	3.37814E-15
70	1.99600E-14	1.12571E-14	2.66186E-15	3.414932E15
80	2.27208E-14	1.61510E-14	2.91539E-15	4.21745E-15
90	2.41366E-14	1.63954E-14	3.23766E-15	3.10200E-15
100	2.85801E-14	1.75552E-14	3.51036E-15	3.62939E-15

FIGURE 1. Graphs for table 4. Blue for linearization and red for homotopy.

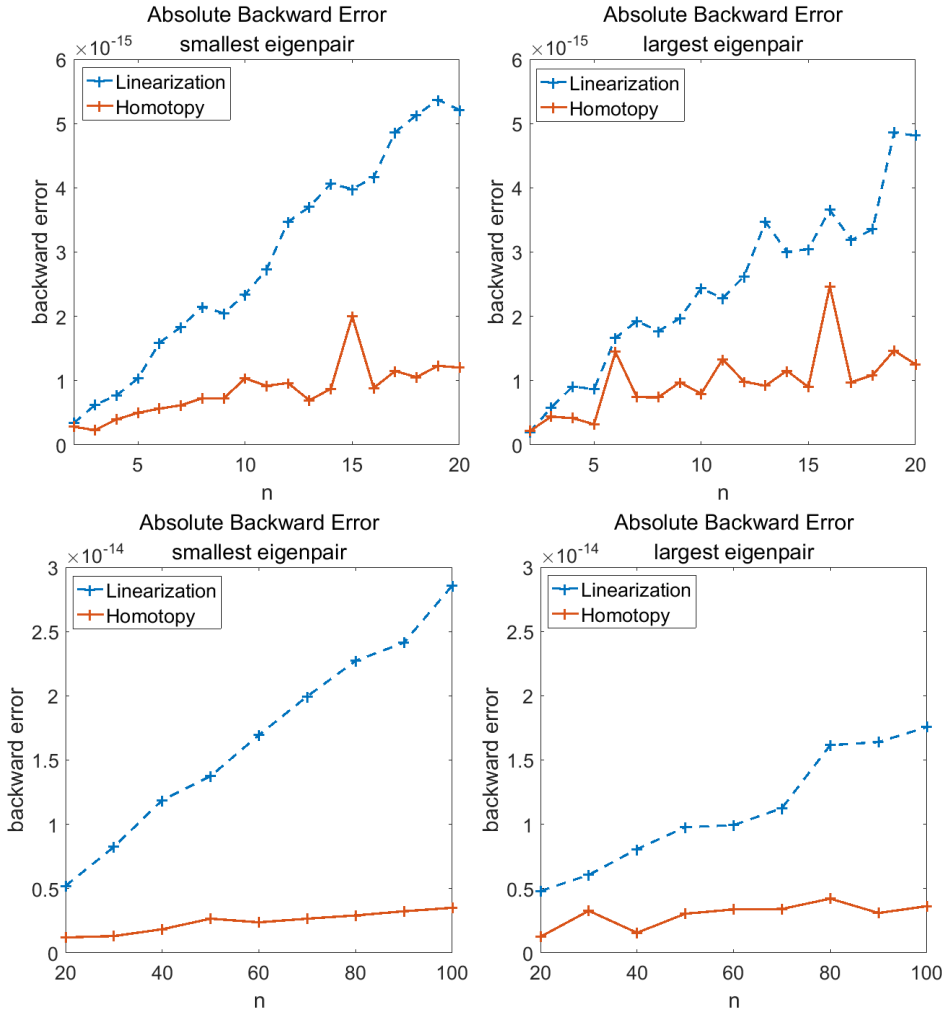
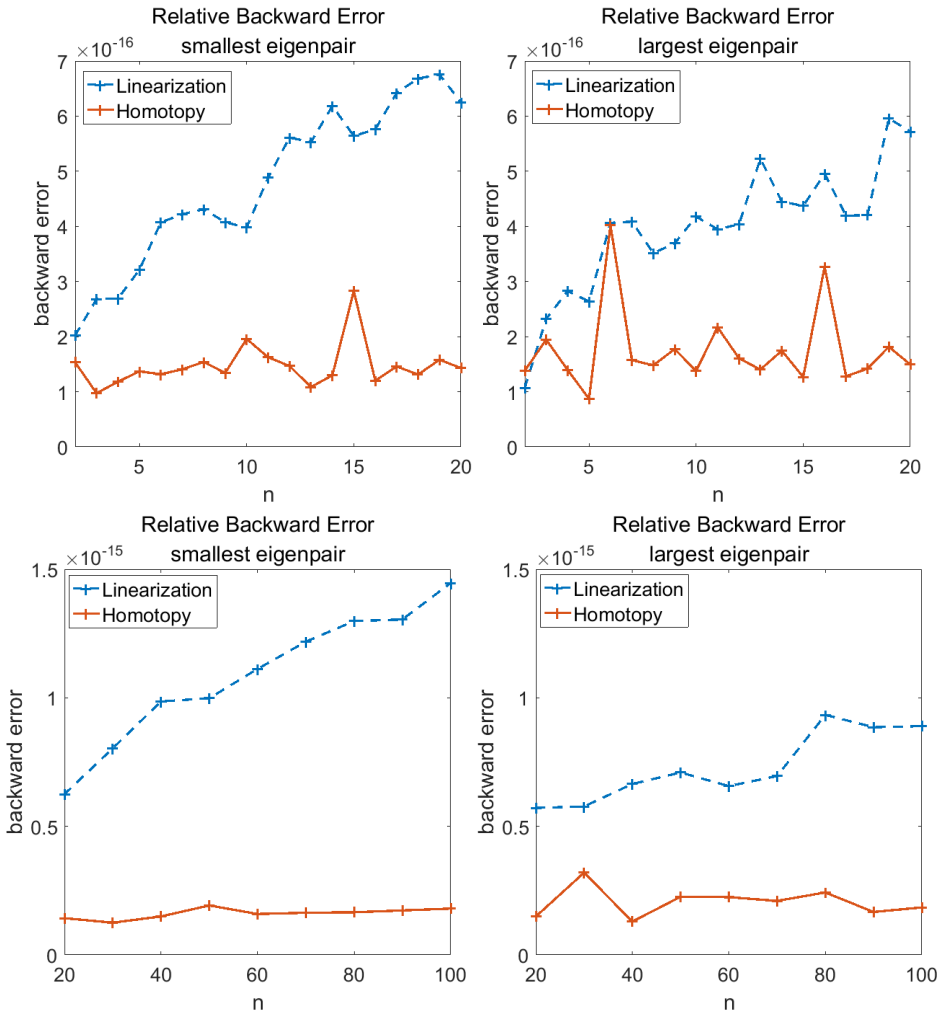


TABLE 5. ACCURACY OF QEP — LINEARIZATION VS HOMOTOPY. Relative backward errors of computed smallest and largest eigenpairs with dimensions $n = 2, \ldots, 100$.

	LINEARIZATION AVG. BK. ERR.		HOMOTOPY AVG. BK. ERR.	
n	SMALLEST	LARGEST	SMALLEST	LARGEST
2	2.02195E-16	1.07268E-16	1.53449E-16	1.38260E-16
5	3.21357E-16	2.63883E-16	1.36922E-16	8.70572E-17
10	3.97923E-16	4.17780E-16	1.95933E-16	1.37749E-16
15	5.64010E-16	4.37375E-16	2.82767E-16	1.26614E-16
20	6.25007E-16	5.71766E-16	1.42798E-16	1.49327E-16
30	8.03800E-16	5.77483E-16	1.25286E-16	3.20068E-16
40	9.85766E-16	6.64878E-16	1.49583E-16	1.29559E-16
50	9.98415E-16	7.09502E-16	1.92299E-16	2.26032E-15
60	1.11238E-15	6.56447E-16	1.59075E-16	2.24618E-16
70	1.21839E-15	6.95665E-16	1.63888E-16	2.10081E-16
80	1.29842E-15	8.85652E-16	1.65893E-16	2.42794E-16
90	1.30478E-15	8.85652E-16	1.72905E-16	1.66882E-16
100	1.44644E-15	8.90062E-16	1.80208E-16	1.84720E-16

FIGURE 2. Graphs for table 5. Blue for linearization and red for homotopy.



7.4. **Speed and accuracy comparisons for PEP.** We repeat the timing and accuracy tests in section 7.2 and section 7.3 for PEP where $m = 4$. The timing results are presented in table 6 and table 7. The accuracy results are presented side-by-side in tabulated form in table 8 (absolute error) and table 9 (relative error), and graphically in fig. 3 (absolute error) and fig. 4 (relative error). As in the case of QEP, for speed comparisons, we run our experiments ten times and record the best, average, median, and worst performance for each method and each dimension; for accuracy comparisons, the tests are averaged over ten runs. The conclusions drawn from these experiments are discussed in section 7.5.

TABLE 6. SPEED OF PEP — LINEARIZATION. Elapsed timings (in seconds) for the linearization method with dimension $n = 2, 3, \ldots, 100$ (not all displayed).

n	# ROOTS	BEST	AVERAGE	MEDIAN	WORST
2	8	3.6209E-4	7.3132E-4	6.2914E-4	0.0016
4	16	6.4617E-4	9.8930E-4	8.6807E-4	0.0017
6	24	0.0010	0.0017	0.0015	0.0028
8	32	0.0017	0.0022	0.0022	0.0031
10	40	0.0022	0.0030	0.0031	0.0042
12	48	0.0030	0.0040	0.0037	0.0055
14	56	0.0045	0.0054	0.0050	0.0074
16	64	0.0068	0.0080	0.0080	0.0100
18	72	0.0091	0.0108	0.0099	0.0141
20	80	0.0106	0.0121	0.0116	0.0149
30	120	0.0262	0.0337	0.0349	0.0373
40	160	0.0656	0.0751	0.0695	0.0929
50	200	0.0920	0.1225	0.1095	0.1724
60	240	0.1606	0.1906	0.1807	0.2513
70	280	0.2401	0.2719	0.2675	0.3262
80	320	0.3757	0.4189	0.4140	0.4986
90	360	0.4947	0.5507	0.5296	0.6968
100	400	0.7903	0.8285	0.8210	0.9458

TABLE 7. SPEED OF PEP — HOMOTOPY IN SERIAL AND IN PARALLEL. Elapsed timings (in seconds) for the homotopy method with dimension $n = 2, 3, \ldots, 100$; in serial for $n = 2, 3, \ldots, 20$ (not all displayed); in parallel on 20 cores for $n = 30, 40, \ldots, 100$.

n	# ROOTS	BEST	AVERAGE	MEDIAN	WORST
2	8	0.4449	0.5051	0.4725	0.6503
4	16	0.6237	0.8137	0.6946	1.2260
6	24	1.0655	2.4848	2.0590	4.9239
8	32	1.6349	3.5440	2.7851	7.6895
10	40	2.7746	8.6236	9.2868	16.6303
12	48	7.3469	16.3581	16.8977	25.2626
14	56	10.0197	26.8512	18.3648	57.4097
16	64	22.8207	45.5760	46.3289	66.0404
18	72	19.6662	61.4427	56.9810	94.0309
20	80	47.8803	119.9619	98.6018	245.0955
30	120	14.0910	30.9721	30.2710	56.8700
40	160	35.1230	55.5960	51.7820	77.9600
50	200	75.2800	111.3593	111.8870	148.0980
60	240	165.5460	214.3197	205.0230	290.4680
70	280	290.4960	397.6789	421.8780	474.0510
80	320	479.8200	616.7326	574.7630	811.5990
90	360	684.9450	1037.8848	1114.4720	1258.1440
100	400	1046.2700	1452.7988	1423.2730	1947.1600

TABLE 8. ACCURACY OF PEP — LINEARIZATION VS HOMOTOPY. Absolute backward errors of computed smallest and largest eigenpairs with dimension $n = 2, \ldots, 100$.

	LINEARIZATION AVG. BK. ERR.		HOMOTOPY AVG. BK. ERR.	
n	SMALLEST	LARGEST	SMALLEST	LARGEST
2	3.97673E-16	4.06234E-16	1.93084E-16	4.72180E-16
5	1.30486E-15	1.01101E-15	5.14466E-16	5.52493E-16
10	3.40010E-15	2.43585E-15	6.66926E-16	9.29498E-16
15	4.06465E-15	3.66940E-15	1.13208E-15	9.14666E-16
20	5.58414E-15	5.16130E-15	1.81496E-15	1.81759E-15
30	8.03892E-15	6.24574E-15	1.38205E-15	1.39749E-15
40	1.13754E-14	8.64433E-15	1.70770E-15	1.81052E-15
50	1.38394E-14	1.07375E-14	2.23095E-15	2.08415E-15
60	1.76932E-14	1.24398E-14	2.45192E-15	3.05775E-15
70	1.90902E-14	1.42487E-14	2.61274E-15	2.70850E-15
80	2.24766E-14	1.57303E-14	2.80187E-15	3.28150E-15
90	2.56978E-14	1.77254E-14	3.14128E-15	3.88126E-15
100	2.64759E-14	1.94732E-14	3.01357E-15	2.92889E-15

FIGURE 3. Graphs for table 8. Blue for linearization and red for homotopy.

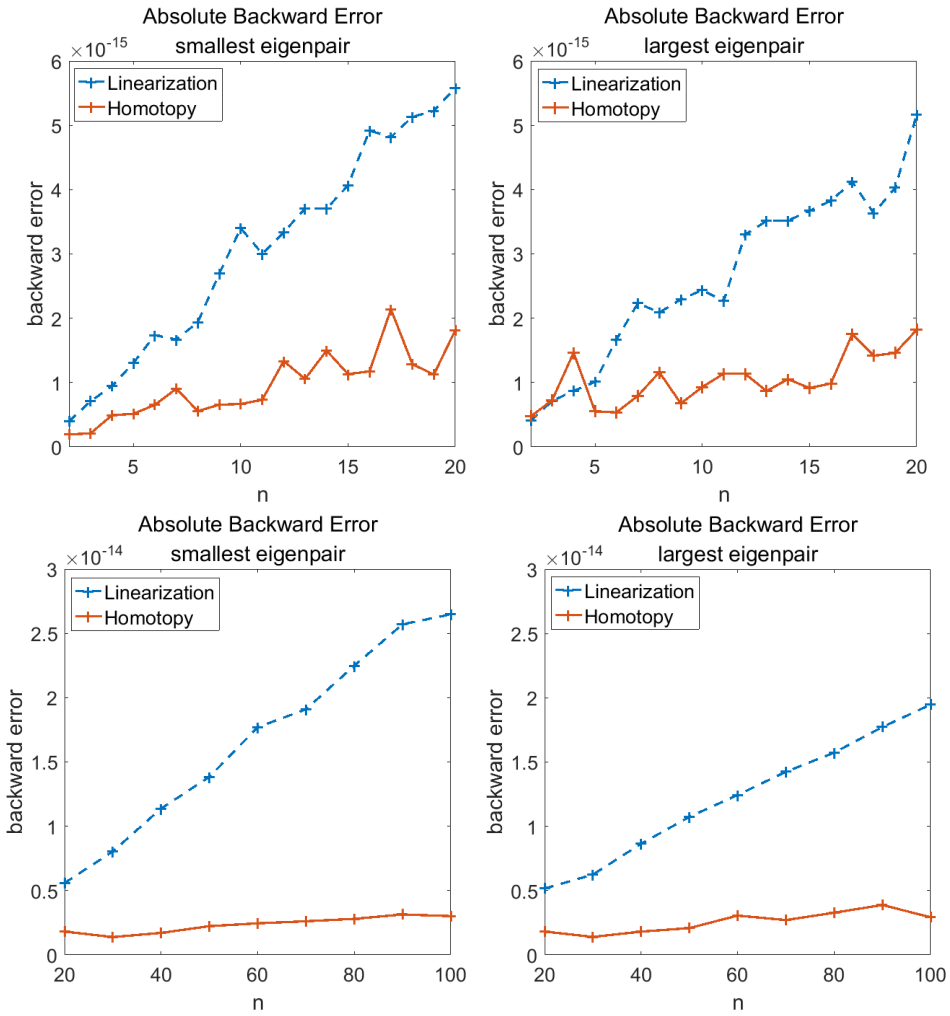
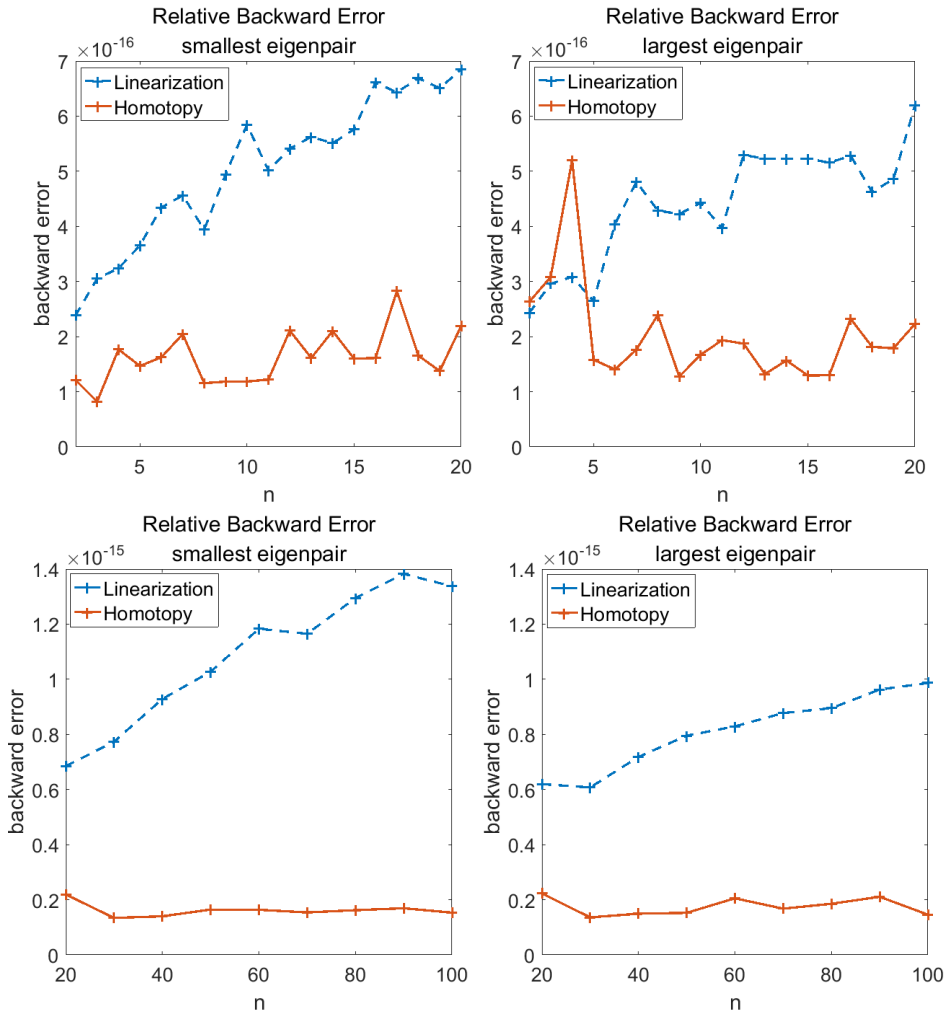


TABLE 9. ACCURACY OF PEP — LINEARIZATION VS HOMOTOPY. Relative backward errors of computed smallest and largest eigenpairs with dimension $n = 2, \ldots, 100$.

	LINEARIZATION AVG. BK. ERR.		HOMOTOPY AVG. BK. ERR.	
n	SMALLEST	LARGEST	SMALLEST	LARGEST
2	2.38800E-16	2.42932E-16	1.20966E-16	2.63005E-16
5	3.65363E-16	2.64597E-16	1.46731E-16	1.58114E-16
10	5.84645E-16	4.42737E-16	1.18572E-16	1.67121E-16
15	5.76099E-16	5.22943E-16	1.60022E-16	1.29557E-16
20	6.84554E-16	6.19752E-16	2.19221E-16	2.22975E-16
30	7.73547E-16	6.07382E-16	1.33975E-16	1.36049E-16
40	9.28358E-16	7.18160E-16	1.39950E-16	1.49331E-16
50	1.02753E-15	7.95390E-16	1.64387E-16	1.52488E-16
60	1.18312E-15	8.28990E-16	1.63573E-16	2.05105E-16
70	1.16505E-15	8.77305E-16	1.53855E-16	1.67640E-16
80	1.29478E-15	8.95204E-16	1.62206E-16	1.85393E-16
90	1.38225E-15	9.62668E-16	1.69063E-16	2.10661E-16
100	1.33725E-15	9.86613E-16	1.53193E-16	1.45843E-16

FIGURE 4. Graphs for table 9. Blue for linearization and red for homotopy.



7.5. **Conclusions.** From the results of the numerical experiments in preceding subsections, we may draw several conclusions regarding the speed and accuracy of our proposed homotopy method versus linearization method, the existing method-of-choice.

Speed: From the elapsed times, the linearization method is consistently much faster than the homotopy method, even when the latter is run in parallel.

Accuracy: From the normwise backward errors, homotopy method is more accurate than the linearization method across coefficient matrices of all dimensions.

Dimension: The gap in accuracy increases significantly with the dimension of the coefficient matrices, favoring homotopy method for higher dimensional problems, particularly when all eigenpairs are needed.

Stability: Our results reflect what is known about linearization method (see section 3), namely, it is numerically unstable for both QEP and PEP.

Initialization: Our results reflect what is known about homotopy method (see section 6), namely, its dependence on the choice of start system — one that leads to a path passing near a singularity takes longer to converge.

Parallelism: The timings of homotopy method can be improved substantially with parallel computing. On the other hand, the linearization method exhibits no obvious parallelism.

It is also evident from these results that both the linearization and homotopy methods take considerably longer to solve a PEP with $m = 4$ than to solve a QEP. This can be attributed to the fact that a PEP with $m = 4$ has $5/3$ times more parameters than a QEP: $\dim(\mathbb{C}^{n\times n\times 5}) = 5n^2$ versus $\dim(\mathbb{C}^{n\times n\times 3}) = 3n^2$, and that a PEP with $m = 4$ has twice as many eigenpairs as a QEP: $4n$ versus $2n$ eigenpairs.

It is interesting to observe that accuracy does not exhibit a similar deterioration with increasing m — for a fixed n, the averaged backward error for a PEP with $m = 4$ can be smaller than that for a QEP.

8. Conditioning and accuracy

It has been shown [16] that if the 2-norms of the coefficient matrices in a PEP are all approximately 1, then the companion linearization and original PEP have similar condition numbers. In particular, define

$$\rho = \frac{\max_i \|A_i\|_2}{\min(\|A_0\|_2, \|A_m\|_2)} \geq 1. \tag{19}$$

When ρ is of order 1, there exists a linearization for a particular eigenvalue that is about as well conditioned as the original PEP itself for that eigenvalue, to within a small constant factor. However this is no longer true when the A_i's vary widely in norm: The companion GEP is potentially far more ill-conditioned than the PEP.

For the QEP in eq. (1), the quantity

$$\rho = \frac{\max(\|M\|, \|C\|, \|K\|)}{\min(\|M\|, \|K\|)}$$

is of order 1 if $\|C\| \lesssim \max(\|M\|, \|K\|)$ and $\|M\| \approx \|K\|$. When these are not satisfied, a scaling of $Q(\lambda)$ will typically improve the conditioning of the linearization — provided that $Q(\lambda)$ is not too heavily damped, i.e., $\|C\| \lesssim \sqrt{\|M\|\|K\|}$. However, in general such a scaling is unavailable; for instance, it is still not known how one should scale a heavily damped QEP.

We compare how linearization and homotopy methods perform on damped QEPs:

(i) We generate random 20×20 coefficient matrices M, C, K having independent entries that follow the standard real Gaussian distribution $\mathcal{N}_{\mathbb{R}}(0, 1)$.

(ii) For $Q_k(\lambda) = \lambda^2 M + \lambda(2^k \cdot C) + K$, $k = 0, 1, \ldots, 5$, we determine the relative backward errors of all computed eigenpairs for both linearization and homotopy methods. The results are in fig. 5.

(iii) At the same time, for each $Q_k(\lambda)$, we compute the condition number of each eigenvalue in both the original QEP (used in homotopy method) and its companion GEP (used in linearization method). The results are in fig. 6.

In both fig. 5 and fig. 6, the horizontal axis is the index of eigenvalues in ascending order of magnitude, the vertical axis is on a log scale.

FIGURE 5. Relative backward errors (in Frobenius norm) of all computed eigenpairs. Blue dots: linearization method. Red crosses: homotopy method.

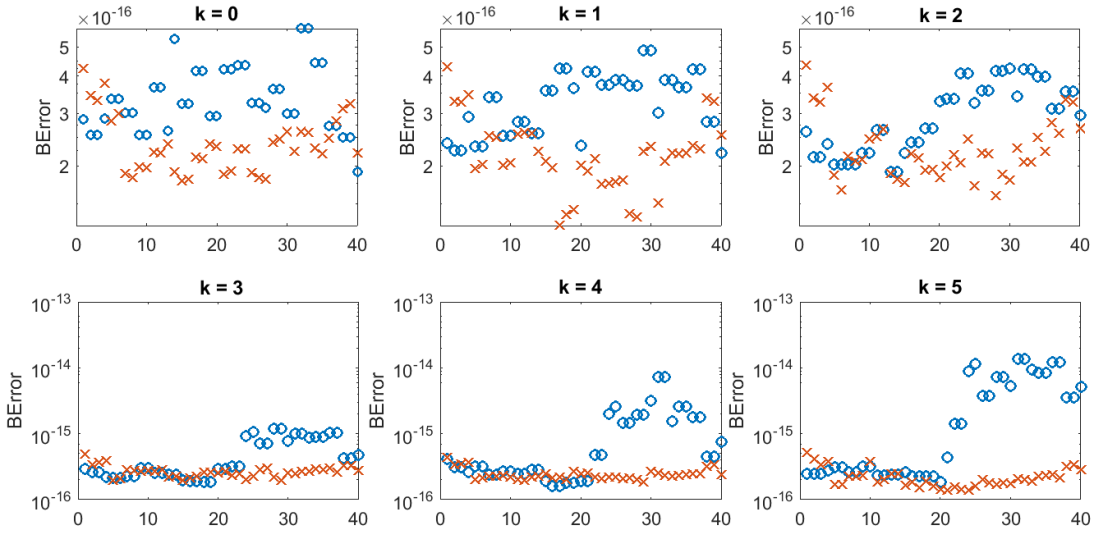
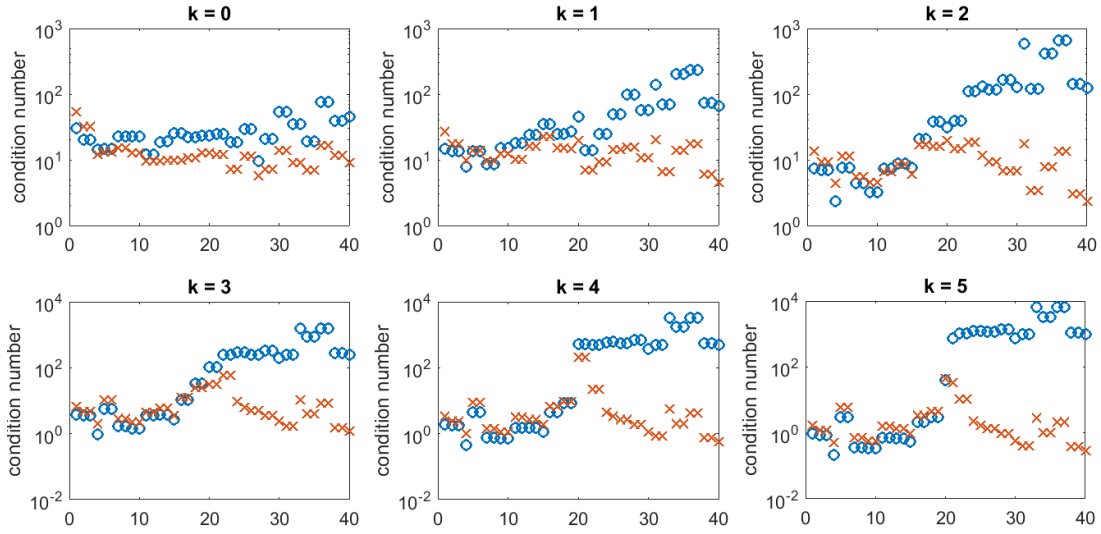


FIGURE 6. Condition numbers of all computed eigenpairs. Blue dots: companion GEPs. Red crosses: original QEPs.



We may deduce the following from these results:

(a) From fig. 5 we see that homotopy method is backward stable for all eigenvalues and all $k = 0, 1, \ldots, 5$; on the other hand, linearization method becomes significantly less stable for the larger eigenvalues as k increases past 3. Note that the larger the value of k, the heavier damped the QEP.
(b) From fig. 6 we see that the larger eigenvalues in the original QEP are far better conditioned than in its companion GEP, whereas the smaller eigenvalues are similarly conditioned in both problems.
(c) From fig. 6 we see that as the QEP becomes more heavily damped, the companion GEP becomes exceedingly worse-conditioned than the original QEP.

In summary, homotopy method is evidently the preferred method for accurate determination of large eigenvalues in heavily damped QEPs.

Although it is possible to find an alternative linearization of a QEP into a GEP that is better conditioned towards the larger eigenvalues [26], there is no known linearization with conditioning comparable to the original QEP across *all* eigenvalues (in fact, such an ideal linearization quite likely does not exist). In principle, one might use different linearizations to determine eigenvalues in different ranges, and then combine the results to obtain a full set of eigenvalues. However this is not only impractical but suffers from a fallacy — we do not know a priori which eigenvalues from which linearizations are more accurate.

9. Applications

In this section we provide numerical experiments on real data (as opposed to randomly generated data in section 7) arising from the two application problems mentioned in section 1. All experiments here are conducted in the same environment as described in section 7, except that the matrices defining the problems are generated with the `nlevp` function in the MATLAB toolbox NLEVP [7].

9.1. **Acoustic wave problem.** This application is taken from [10]. Consider an acoustic medium with constant density and space-varying sound speed $c(x)$ occupying the volume $\Omega \subseteq \mathbb{R}^d$. The homogeneous wave equation for the acoustic pressure $p(x,t) = \widehat{p}(x)e^{\widehat{\lambda}t}$ has a factored form that simplifies the wave equation to the following, where $\widehat{\lambda}$ is the eigenvalue:

$$-\Delta\widehat{p}(x) - \left(\widehat{\lambda}/c\right)^2 \cdot \widehat{p}(x) = 0. \tag{20}$$

For our purpose, it suffices to consider the one-dimensional acoustic wave problem, i.e., $d = 1$ and $\Omega = [0, 1]$, with Dirichlet boundary condition $\widehat{p} = 0$ and impedance boundary condition $\partial p/\partial n + i\widehat{\lambda}p/\zeta = 0$.

9.1.1. *Numerical results.* The quadratic matrix polynomial $Q(\lambda) = \lambda^2 M + \lambda C + K$ arises from a finite element discretization of eq. (20). In our numerical experiments we set impedance $\zeta = 1$. The three $n \times n$ matrices in this QEP are

$$M = -4\pi^2 \frac{1}{n}\left(I_n - \frac{1}{2}e_n e_n^{\mathsf{T}}\right), \quad C = 2\pi i \frac{1}{\zeta} e_n e_n^{\mathsf{T}}, \quad K = n\begin{bmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & 2 & -1 \\ & & -1 & 1 \end{bmatrix}.$$

As before, we compare the accuracy and timings of homotopy and linearization methods on this problem. We tabulate the absolute and relative backward errors of the computed eigenpairs corresponding to the smallest and largest eigenvalues for dimensions $n = 20, 30, \ldots, 100$ in table 10 and table 11, and plot them graphically in fig. 7. All accuracy tests are averaged over ten runs.

The elapsed timings are tabulated in table 12 and table 13. All speed tests are run ten times with the best, average, median and worst timings recorded.

TABLE 10. ACOUSTIC WAVE PROBLEM — ABSOLUTE BACKWARD ERRORS. Computed smallest and largest eigenpairs with dimensions $n = 20, 30, \ldots, 100$ via homotopy and linearization methods.

	LINEARIZATION AVG. BK. ERR.		HOMOTOPY AVG. BK. ERR.	
n	SMALLEST	LARGEST	SMALLEST	LARGEST
20	1.91540E-14	3.26558E-14	5.50499E-15	1.20102E-15
30	2.48304E-14	3.98265E-14	8.69382E-15	6.41943E-16
40	4.10844E-14	4.28048E-14	1.09178E-14	4.76802E-16
50	4.78653E-14	4.97587E-14	1.04868E-14	4.50403E-16
60	6.85333E-14	5.29234E-14	1.40925E-14	1.91169E-16
70	7.68982E-14	4.65017E-14	1.90018E-14	1.07398E-16
80	8.80830E-14	4.12431E-14	1.58457E-14	8.02790E-17
90	1.11451E-13	4.90212E-14	1.37846E-14	9.11717E-17
100	1.14562E-13	4.49546E-14	1.52145E-14	1.52150E-16

TABLE 11. ACOUSTIC WAVE PROBLEM — RELATIVE BACKWARD ERRORS. Computed smallest and largest eigenpairs with dimensions $n = 20, 30, \ldots, 100$ via homotopy and linearization methods.

	LINEARIZATION AVG. BK. ERR.		HOMOTOPY AVG. BK. ERR.	
n	SMALLEST	LARGEST	SMALLEST	LARGEST
20	4.19838E-16	7.81869E-15	1.20664E-16	2.87557E-16
30	3.87748E-16	1.35042E-14	1.35761E-16	2.17667E-16
40	5.03756E-16	1.88157E-14	1.33868E-16	2.09588E-16
50	4.86198E-16	2.68895E-14	1.06521E-16	2.43396E-16
60	5.96657E-16	3.39445E-14	1.22691E-16	1.22614E-16
70	5.87499E-16	3.45264E-14	1.45173E-16	7.97402E-17
80	6.00842E-16	3.47937E-14	1.08089E-16	6.77252E-17
90	6.87825E-16	4.63156E-14	8.50724E-17	8.61397E-17
100	6.46406E-16	4.70234E-14	8.58465E-17	1.59152E-16

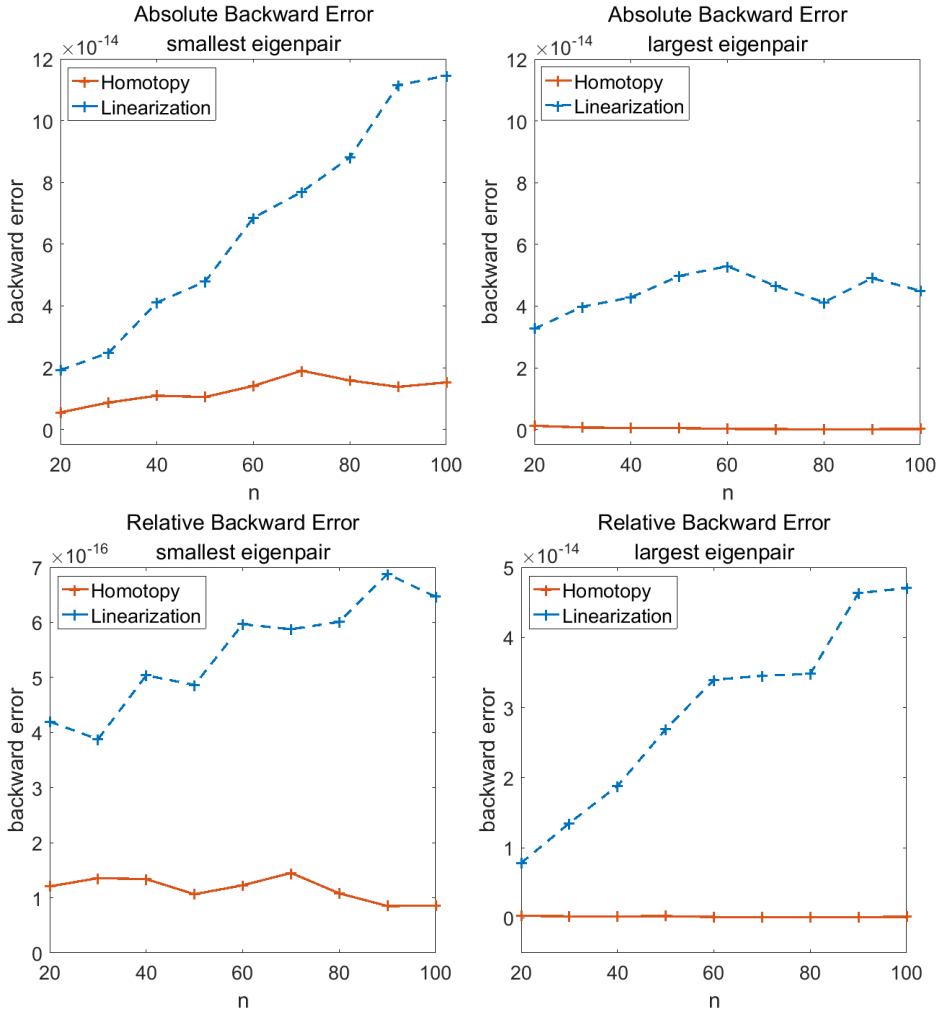
TABLE 12. ACOUSTIC WAVE PROBLEM — LINEARIZATION METHOD. Elapsed timings (in seconds) for the linearization method with dimensions $n = 20, 30, \ldots, 100$.

n	# ROOTS	BEST	AVERAGE	MEDIAN	WORST
20	40	0.0056	0.0139	0.0071	0.0754
30	60	0.0107	0.0175	0.0134	0.0592
40	80	0.0202	0.0310	0.0221	0.1127
50	100	0.0365	0.0499	0.0387	0.1464
60	120	0.0469	0.0673	0.0600	0.1380
70	140	0.0680	0.0826	0.0830	0.1040
80	160	0.0933	0.1192	0.1004	0.2491
90	180	0.1304	0.1549	0.1445	0.2135
100	200	0.1723	0.2015	0.1818	0.3137

TABLE 13. ACOUSTIC WAVE PROBLEM — HOMOTOPY METHOD. Elapsed timings (in seconds) for the homotopy method with dimensions $n = 20, 30, \ldots, 100$, in serial for $n = 20, 30, 40$, and in parallel on 20 cores for $n = 50, 60, \ldots, 100$.

n	# ROOTS	BEST	AVERAGE	MEDIAN	WORST
20	40	7.7472	19.5057	16.7535	34.7300
30	60	104.2553	166.2379	155.9914	257.1145
40	80	265.2321	557.2342	526.8778	1003.7452
50	100	30.0660	37.9267	35.5240	49.5630
60	120	74.2800	92.1235	94.7845	113.7100
70	140	149.3000	189.5564	182.7230	280.6460
80	160	252.4920	378.7414	370.6855	519.2240
90	180	530.7180	691.8141	736.6715	841.8600
100	200	559.9340	1079.0177	1154.3995	1464.4170

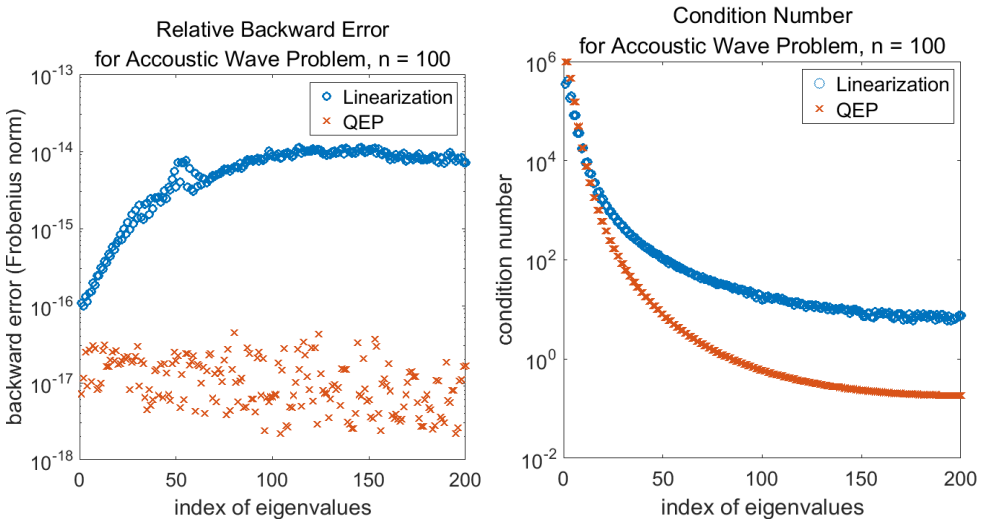
FIGURE 7. ACOUSTIC WAVE PROBLEM — ABSOLUTE AND RELATIVE BACKWARD ERRORS. Computed smallest and largest eigenpairs with dimensions $n = 20, 30, \ldots, 100$. Blue dashed lines: linearization method. Red solid lines: homotopy method.



9.1.2. *Conclusions.* From these results, we may draw essentially the same conclusions in section 7.5 that we made about our numerical experiments on randomly generated data. In particularly, homotopy method vastly outperforms linearization method in accuracy, as measured by normwise backward errors. The difference in this case is that we have very sparse, highly structured matrices M, C, K.

9.1.3. *Conditioning and accuracy.* We next examine the accuracy of the two methods for the case $n = 100$ more closely — computing all eigenvalues, instead of just the largest and smallest. We will also compute the condition number for each eigenvalue in both the original QEP formulation and in the companion GEP formulation. The results are plotted in fig. 8. In this case, $\|M\|_F = 3.9$, $\|C\|_F = 6.3$, $\|K\|_F = 2439.3$, which is not regarded as a heavily damped QEP.

FIGURE 8. ACOUSTIC WAVE PROBLEM WITH $n = 100$ — ACCURACY AND CONDITIONING. Horizontal axes: index of eigenvalues in ascending order of magnitude. Vertical axis: log scale. LEFT PLOT: relative backward errors of computed eigenpairs; blue dots for the linearization method, red crosses for the homotopy method. RIGHT PLOT: condition numbers of each eigenvalue; blue dots for the companion GEP, red crosses for the original QEP.

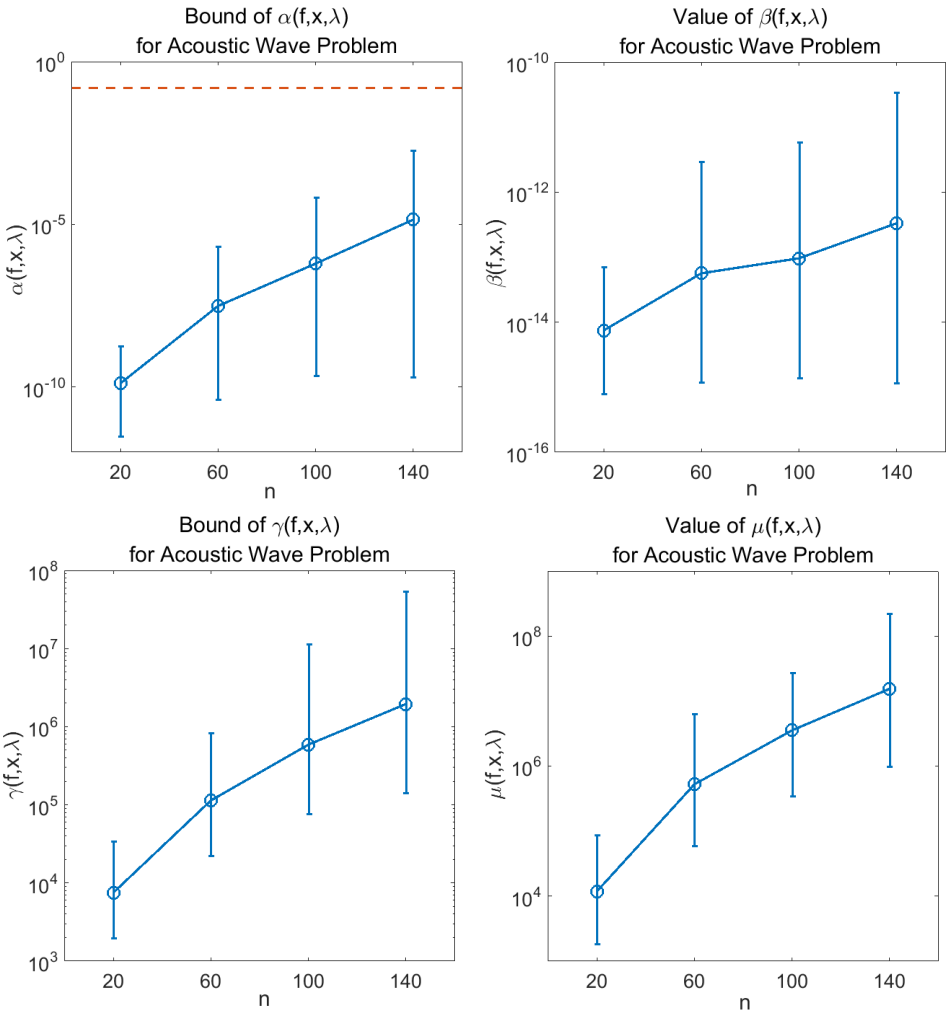


Again, we see that homotopy method is numerically stable across all eigenpairs whereas the linearization method is unstable for nearly 90% of eigenpairs. The original QEP is also better-conditioned than its companion GEP: the condition number of each eigenvalue of the former is consistently smaller than, or very close to, that of the same eigenvalue in the latter.

9.1.4. *Certification.* Last but not least, we will use Smale's α-theory, discussed in section 5, to deduce that for $n = 20, 60, 100, 140$, the Newton iterations in an application of homotopy method to the acoustic wave problem will converge quadratically to all eigenpairs.

Specifically, for $n = 20, 60, 100, 140$, we solve for all eigenpairs, compute $\mu(f, x, \lambda)$ using eq. (15), and obtain a bound for $\gamma(f, x, \lambda)$ according to theorem 5.3. The Newton residual $\beta(f, x, \lambda)$ is also computed, which together with $\gamma(f, x, \lambda)$ yields an upper bound for $\alpha(f, x, \lambda)$. We present our results in fig. 9.

FIGURE 9. ACOUSTIC WAVE PROBLEM — CERTIFICATION. The values of $\beta(f, x, \lambda)$, $\mu(f, x, \lambda)$, and upper bound values of $\alpha(f, x, \lambda)$, $\gamma(f, x, \lambda)$ for the acoustic wave problem with $n = 20, 60, 100, 140$. The circle in the middle, the whiskers above and below represent the mean, maximum, and minimum value of each of these quantities taken over all eigenpairs. The red horizontal dashed line in the plot for $\alpha(f, x, \lambda)$ represents the threshold in theorem 5.2. Note that the vertical axis is in log scale.



For every dimension that we test, the value of $\alpha(f, x, \lambda)$ for each eigenpair is much smaller than the threshold given in theorem 5.2. In other words, this certifies that all eigenpairs that we computed using the homotopy method are accurate solutions to the PEP, for $n = 20, 60, 100, 140$.

9.2. **Planar waveguide problem.** This example is taken from [25]. The 129×129 quartic matrix polynomial $P(\lambda) = \lambda^4 A_4 + \lambda^3 A_3 + \lambda^2 A_2 + \lambda A_1 + A_0$ arises from a finite element solution of the equation for the modes of a planar waveguide using piecewise linear basis φ_i, $i = 0, \ldots, 128$. The coefficient matrices are defined by:

$$A_1 = \frac{\delta^2}{4}\operatorname{diag}(-1, 0, 0, \ldots, 0, 0, 1), \quad A_3 = \operatorname{diag}(1, 0, 0, \ldots, 0, 0, 1),$$

$$A_0(i, j) = \frac{\delta^4}{16}\langle \varphi_i, \varphi_j \rangle, \quad A_2(i, j) = \langle \varphi_i', \varphi_j' \rangle - \langle q\varphi_i, \varphi_j \rangle, \quad A_4(i, j) = \langle \varphi_i, \varphi_j \rangle.$$

The parameter δ describes the difference in refractive index between the cover and substrate of the waveguide and q is a function from the variational formulation.

The dimension of this PEP is fixed at $n = 129$. As before, we compare the accuracy and timings of homotopy and linearization methods on this problem. We present the absolute and relative backward errors for all computed eigenpairs in fig. 10 and tabulate the best, average, median, worst performance in table 14. The elapsed timings are given in table 15. All speed and accuracy tests are run ten times with the best, average, median and worst results recorded.

TABLE 14. PLANAR WAVEGUIDE PROBLEM — ABSOLUTE AND RELATIVE BACKWARD ERRORS. Best, average, median, worst absolute (top table) and relative (bottom table) backward errors of all eigenpairs computed via homotopy and linearization methods.

$n = 129$	LINEARIZATION ABS. BK. ERR.	HOMOTOPY ABS. BK. ERR.
BEST	4.77790E-15	6.06981E-19
MEAN	5.38568E-13	2.05656E-15
MEDIAN	2.83418E-13	6.90006E-16
WORST	1.76856E-12	1.166201E-14
$n = 129$	LINEARIZATION REL. BK. ERR.	HOMOTOPY REL BK. ERR.
BEST	1.14289E-15	1.72613E-17
MEAN	2.53466E-12	9.48695E-17
MEDIAN	5.37243E-14	7.64027E-17
WORST	1.87416E-11	4.24275E-16

FIGURE 10. PLANAR WAVEGUIDE PROBLEM — SORTED ABSOLUTE AND RELATIVE BACKWARD ERRORS. Sorted absolute and relative backward errors of all 516 computed eigenpairs. Blue dashed lines represent linearization method; red solid lines represent homotopy method. Tests are averaged over ten runs and vertical axis is in log scale.

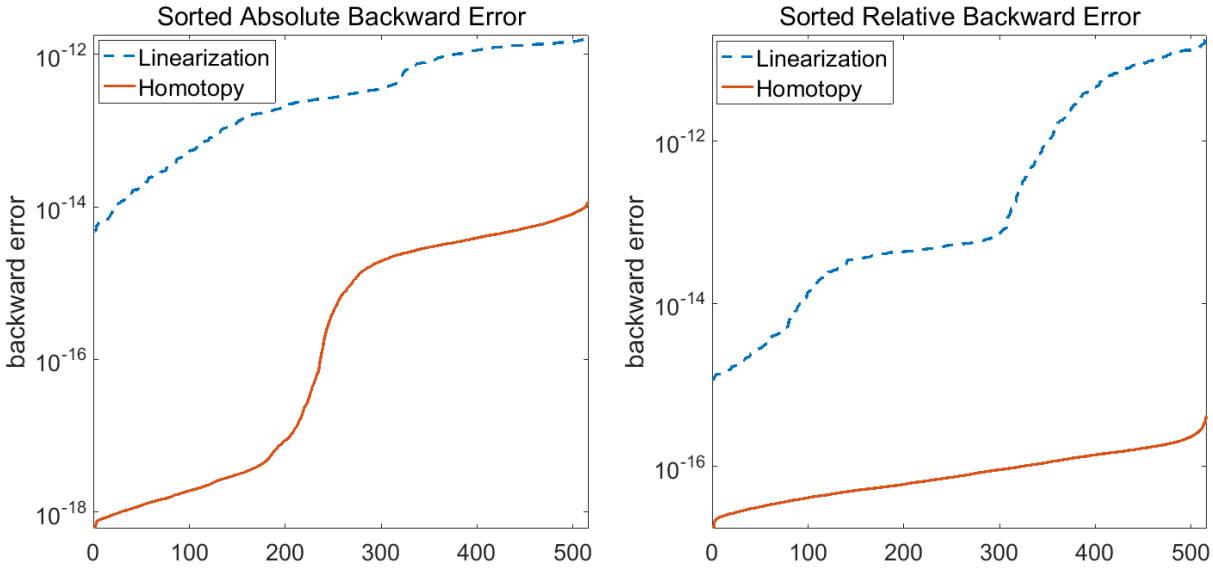


TABLE 15. PLANAR WAVEGUIDE PROBLEM — SPEED. Elapsed timings (in seconds) for homotopy and linearization methods. The homotopy method is run in parallel on 80 cores.

$n = 129$	LINEARIZATION TIMINGS	HOMOTOPY TIMINGS
BEST	2.6670	2673.0350
MEAN	2.7321	3013.9718
MEDIAN	2.7213	2973.6600
WORST	2.8414	3460.5290

The results obtained for this quartic PEP arising from the planar waveguide problem is consistent with what we have observed for the acoustic wave QEP in section 9.1 as well as the randomly generated PEPs and QEPs in section 7 — while homotopy method requires much longer running times than the linearization method, its results are also vastly superior in terms of accuracy.

In summary, if our main goal is to obtain accurate solutions to polynomial eigenvalue problems, particularly when all eigenpairs are needed, then expending additional resources (more cores and longer computing time) to employ the homotopy method is not only worthwhile but perhaps inevitable — we know of no other alternative that would achieve the same level of accuracy.

ACKNOWLEDGMENT

The work in this article is generously supported by DARPA D15AP00109 and NSF IIS 1546413. LHL is supported by a DARPA Director's Fellowship. JIR is supported by a University of Chicago Provost Postdoctoral Scholarship.

REFERENCES

[1] D Afolabi. Linearization of the quadratic eigenvalue problem. *Computers & structures*, 26(6):1039–1040, 1987.
[2] Eugene L. Allgower and Kurt Georg. *Numerical continuation methods*, volume 13 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1990. An introduction.
[3] Eugene L. Allgower and Kurt Georg. Continuation and path following. In *Acta numerica, 1993*, Acta Numer., pages 1–64. Cambridge Univ. Press, Cambridge, 1993.
[4] W. E. Arnoldi. The principle of minimized iteration in the solution of the matrix eigenvalue problem. *Quart. Appl. Math.*, 9:17–29, 1951.
[5] Daniel J. Bates, Jonathan D. Hauenstein, Andrew J. Sommese, and Charles W. Wampler. *Numerically solving polynomial systems with Bertini*, volume 25 of *Software, Environments, and Tools*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2013.
[6] Daniel J. Bates, Andrew J. Newell, and Matthew Niemerg. BertiniLab: a MATLAB interface for solving systems of polynomial equations. *Numer. Algorithms*, 71(1):229–244, 2016.
[7] Timo Betcke, Nicholas J. Higham, Volker Mehrmann, Christian Schröder, and Françoise Tisseur. NLEVP: a collection of nonlinear eigenvalue problems. *ACM Trans. Math. Software*, 39(2):Art. 7, 28, 2013.
[8] Lenore Blum, Felipe Cucker, Michael Shub, and Steve Smale. *Complexity and real computation*. Springer-Verlag, New York, 1998. With a foreword by Richard M. Karp.
[9] James Case and Michael Shub. The final touch? Solving smale's 17th challenge problem. *SIAM News*, 49(5):9, 2016.
[10] F. Chaitin-Chatelin and M. B. van Gijzen. Analysis of parameterized quadratic eigenvalue problems in computational acoustics with homotopic deviation theory. *Numer. Linear Algebra Appl.*, 13(6):487–512, 2006.
[11] Jean-Pierre Dedieu and Françoise Tisseur. Perturbation theory for homogeneous polynomial eigenvalue problems. *Linear Algebra Appl.*, 358:71–94, 2003. Special issue on accurate solution of eigenvalue problems (Hagen, 2000).
[12] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix polynomials*, volume 58 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2009. Reprint of the 1982 original [MR0662418].
[13] Gene H. Golub and Charles F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, fourth edition, 2013.
[14] Sven Hammarling, Christopher J. Munro, and Françoise Tisseur. An algorithm for the complete solution of quadratic eigenvalue problems. *ACM Trans. Math. Software*, 39(3):Art. 18, 19, 2013.

[15] Jonathan D. Hauenstein and Frank Sottile. Algorithm 921: alphaCertified: certifying solutions to polynomial systems. *ACM Trans. Math. Software*, 38(4):Art. ID 28, 20, 2012.
[16] Nicholas J. Higham, D. Steven Mackey, and Françoise Tisseur. The conditioning of linearizations of matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 28(4):1005–1028, 2006.
[17] Peter Lancaster and Miron Tismenetsky. *The theory of matrices*. Computer Science and Applied Mathematics. Academic Press, Inc., Orlando, FL, second edition, 1985.
[18] T. Y. Li. Numerical solution of multivariate polynomial systems by homotopy continuation methods. In *Acta numerica, 1997*, volume 6 of *Acta Numer.*, pages 399–436. Cambridge Univ. Press, Cambridge, 1997.
[19] S. H. Lui and G. H. Golub. Homotopy method for the numerical solution of the eigenvalue problem of self-adjoint partial differential operators. *Numer. Algorithms*, 10(3-4):363–378, 1995.
[20] Mark M. Plecnik and J McCarthy. Controlling the movement of a TRR spatial chain with coupled six-bar function generators for biomimetic motion. *Journal of Mechanisms and Robotics*, 8(5):051005–051005–10, 05 2016.
[21] C. B. Moler and G. W. Stewart. An algorithm for generalized matrix eigenvalue problems. *SIAM J. Numer. Anal.*, 10:241–256, 1973. Collection of articles dedicated to the memory of George E. Forsythe.
[22] Vanni Noferini and Alex Townsend. Numerical instability of resultant methods for multidimensional rootfinding. *SIAM J. Numer. Anal.*, 54(2):719–743, 2016.
[23] Gerard L. G. Sleijpen, Albert G. L. Booten, Diederik R. Fokkema, and Henk A. Van der Vorst. Jacobi-Davidson type methods for generalized eigenproblems and polynomial eigenproblems. *BIT*, 36(3):595–633, 1996. International Linear Algebra Year (Toulouse, 1995).
[24] Gerard L. G. Sleijpen and Henk A. Van der Vorst. A Jacobi-Davidson iteration method for linear eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 17(2):401–425, 1996.
[25] David Stowell and Johannes Tausch. Guided and leaky modes of planar waveguides: Computation via high order finite elements and iterative methods. *PIERS Online*, 6(7):669–673, 2010.
[26] Françoise Tisseur. Backward error and condition of polynomial eigenvalue problems. *Linear Algebra Appl.*, 309(1-3):339–361, 2000.
[27] Françoise Tisseur and Karl Meerbergen. The quadratic eigenvalue problem. *SIAM Rev.*, 43(2):235–286, 2001.
[28] Richard S. Varga. *Geršgorin and his circles*, volume 36 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2004.
[29] T. Zhang, K. H. Law, and G. H. Golub. On the homotopy method for perturbed symmetric generalized eigenvalue problems. *SIAM J. Sci. Comput.*, 19(5):1625–1645, 1998.

Department of Statistics, University of Chicago, Chicago, IL
E-mail address: ylyou@uchicago.edu

Computational and Applied Mathematics Initiative, Department of Statistics, University of Chicago, Chicago, IL
E-mail address, corresponding author: joisro@uchicago.edu

Computational and Applied Mathematics Initiative, Department of Statistics, University of Chicago, Chicago, IL
E-mail address: lekheng@galton.uchicago.edu