

AN ODE BASED METHOD FOR COMPUTING THE DISTANCE OF COPRIME POLYNOMIALS TO COMMON DIVISIBILITY

NICOLA GUGLIELMI* AND IVAN MARKOVSKY ‡

Abstract. The problem of computing the distance of two real coprime polynomials to the set of polynomials with a nontrivial greatest common divisor (GCD) appears in computer algebra, signal processing, and control theory. It has been studied in the literature under the names approximate common divisor, ε -GCD, and distance to uncontrollability. Existing solution methods use different types of local optimization methods and require a user defined initial approximation. In this paper, we propose a new method that allows us to include constraints on the coefficients of the polynomials. Moreover, the method proposed in the paper is more robust to the initial approximation than Newton-type optimization methods available in the literature. Our approach consists of two steps: 1) reformulate the problem as the problem of determining the structured distance to singularity of an associated Sylvester matrix, and 2) integrate a system of ordinary differential equations, which describes the gradient associated to the functional to be minimized.

Key words. ε -GCD, Sylvester matrix, structured pseudospectrum, structured low-rank approximation, ODEs on matrix manifolds, structured distance to singularity.

1. Introduction. We indicate by \mathcal{P}_k the set of polynomials of degree at most k . Consider a pair of polynomials $p \in \mathcal{P}_n$ and $q \in \mathcal{P}_m$ ($m \leq n$) which are assumed to be coprime. An interesting problem discussed in the literature (see e.g. [3]) is that of determining the closest pair $\hat{p} \in \mathcal{P}_n, \hat{q} \in \mathcal{P}_m$ which admit a nontrivial greatest common divisor (GCD). In this paper, we assume that the polynomials p, q, \hat{p}, \hat{q} have real coefficients; however an extension to complex polynomials is straightforward.

Let $p \in \mathcal{P}_n$ and $q \in \mathcal{P}_m$

$$\begin{aligned} p(z) &= a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 \\ q(z) &= b_m z^m + b_{m-1} z^{m-1} + \cdots + b_1 z + b_0 \end{aligned} \tag{1.1}$$

and similarly \hat{p} and \hat{q} of the same degrees, respectively. Denote by

$$\begin{aligned} a &= (a_n \quad a_{n-1} \quad \cdots \quad a_1 \quad a_0)^\top \\ b &= (b_m \quad b_{m-1} \quad \cdots \quad b_1 \quad b_0)^\top \end{aligned}$$

the vectors of the coefficients of the polynomials p and q , respectively, and similarly \hat{a} and \hat{b} be the vectors of the coefficients of the polynomials \hat{p} and \hat{q} , respectively. Then, we define the distance measure

$$\text{dist}((p, q), (\hat{p}, \hat{q})) = \sqrt{\sum_{i=0}^n |a_i - \hat{a}_i|^2 + \sum_{j=0}^m |b_j - \hat{b}_j|^2} \tag{1.2}$$

that is the spectral norm of the vector $\begin{pmatrix} a - \hat{a} \\ b - \hat{b} \end{pmatrix}$.

The considered problem is defined as follows.

*Gran Sasso Science Institute, via Crispi 7, L'Aquila and Dipartimento di Ingegneria Scienze Informatiche e Matematica, Università degli Studi di L'Aquila, Via Vetoio - Loc. Coppito, and Gran Sasso Science Institute (GSSI), I-67010 L'Aquila, Italy. Email: guglielm@univaq.it

‡ELEC, Vrije Universiteit Brussel (VUB), Building K, Pleinlaan 2 1050 Brussels, Belgium. Email: imarkovs@vub.ac.be

PROBLEM 1. Given a pair of coprime polynomials $p \in \mathcal{P}_n, q \in \mathcal{P}_m$ find (for the given distance):

$$d_{CD}(p, q) = \inf_{\substack{\hat{p} \in \mathcal{P}_n, \hat{q} \in \mathcal{P}_m \\ \hat{p} \text{ and } \hat{q} \text{ not coprime}}} \text{dist}((p, q), (\hat{p}, \hat{q})) \quad (\text{coprimeness radius}). \quad (1.3)$$

The distance to common divisibility, defined by equation (1.3), is related to the problem of computing an approximate common divisor [26, 9] and a *closest ε -GCD* [24]. The approximate common divisor problem aims to compute a pair of polynomials (\hat{p}, \hat{q}) that has a common factor of a specified degree and minimizes $\text{dist}((p, q), (\hat{p}, \hat{q}))$. The closest ε -GCD problem aims to compute a pair of polynomials (\hat{p}, \hat{q}) , such that $\text{dist}((p, q), (\hat{p}, \hat{q})) \leq \varepsilon$ and the degree of the GCD of (\hat{p}, \hat{q}) is maximized. The approximate common divisor problem and the closest ε -GCD problems are equivalent optimization problems [46]. In the distance to common divisibility problem, we compute a pair of polynomials (\hat{p}, \hat{q}) with a common factor of degree 1 or 2 that minimize $\text{dist}((p, q), (\hat{p}, \hat{q}))$. Therefore, we are dealing with a special case of the approximate common divisor problem.

In many applications, coprimeness of a set of polynomials is an important property. However, the coefficients of the polynomials are often affected by uncertainties. A more reliable issue is that of deciding whether two polynomials remain coprime even after perturbations of coefficients, bounded in norm by some ε . This issue is considered for example in image processing, robotics and control theory [23, 38], where the input data are only known to a certain accuracy, or where input parameters are affected by noise. In this sense Problem 1.3 aims to compute the value ε which guarantees that if the perturbations in the polynomial coefficients are smaller than ε the perturbed polynomials remain coprime.

By representing the polynomials \hat{p} and \hat{q} as $\hat{p}(z) = (z - \lambda)r(z)$ and $b(z) = (z - \lambda)s(z)$, where λ is the common zero and eliminating the r and s polynomials by analytically minimizing $\text{dist}((p, q), (\hat{p}, \hat{q}))$ over them, we obtain (by using the variable projections principle) [13] an optimization problem equivalent to (1.3), i.e. $\min_{\lambda} f(\lambda)$, with

$$f(\lambda) = \frac{|p(\lambda)|^2}{1 + |\lambda|^2 + \dots + |\lambda^n|^2} + \frac{|q(\lambda)|^2}{1 + |\lambda|^2 + \dots + |\lambda^m|^2}. \quad (1.4)$$

As pointed out by an anonymous reviewer, formula (1.4) appears in [3] but has a longer history.

In this paper we follow a different approach, based on the Sylvester matrix of two polynomials, which is a fundamental tool in determining their greatest common divisor. In particular coprimeness is equivalent to the non-singularity of the associated Sylvester matrix. As a consequence several estimates of the distance to common divisibility of two polynomials are based on the magnitude of the inverse of the Sylvester matrix (see e.g. [3]).

The aim of this article is that of approximating this distance by computing suitable upper bounds with a local optimization procedure. Such computation is not straightforward since it has to do with an optimization problem in the sub-variety of polynomials having a GCD. The method we propose returns (in general) a local minimum for the considered distance but we cannot certify that it is a global minimum. As far as we know there are no methods able to guarantee in general global optimality of a computed solution (see e.g. [24]). Thus convergence to local optima

is an unavoidable property. However, trying different initial approximations, would increase the robustness of the method, which—in the cases of small dimension we are able to check—computes the exact distance.

An introductory example from [3]. Let us consider the following example:

$$p(z) = z^n, \quad q(z) = \left(\frac{1-z}{2}\right)^n \quad \text{with } n = 8.$$

Clearly the zeros of p and q are relatively far. However, as observed in [3], this does not imply that the distance to common divisibility is large. In fact, using the methods proposed in this article we are able to compute a sharp upper bound for $d_{CD}(p, q)$. We compute

$$d_{CD}(p, q) \leq 1.9798 \dots 10^{-4}.$$

This is obtained by determining a pair of polynomials \hat{p} and \hat{q} having the distance $1.9798 \dots 10^{-4}$ from p and q , and having a common real zero $z = 0.32495 \dots$. This is consistent with the results in [3], where a different distance (the maximum of the 1-norm of the two vectors $a - \hat{a}$ and $b - \hat{b}$ replaces the 2-norm in formula (1.2)) is determined exactly for this example, which turns out to be of the same order of magnitude. They determine in fact a distance $3^{-8} = 1.5241 \dots 10^{-4}$ and a common zero $z' = 1/3$. As observed by the authors of [3], classical estimates based on the norm of the inverse of the Sylvester matrix associated to the polynomials p and q , fail to give sharp bounds for this example. The spirit of this paper is similar to [3], where the authors propose (in a slightly different setting) sharper bounds for the distance with respect to the classical bounds proposed in the literature, based on a clever use of the structure of the Sylvester matrix. Here the computation of optimal (or suboptimal) bounds is also pursued by working on the structured set of Sylvester matrices and developing a numerical approach relying on the integration of a system of ODEs which identifies with a gradient system for a functional which is minimized in correspondence of a closest pair of polynomials to p and q , having at least a common zero.

Overview of the contribution and organization of the paper. The problem of computing a so-called ε -GCD of a pair of polynomials $p \in \mathcal{P}_n$ and $q \in \mathcal{P}_m$, that is a nearby pair of polynomials $\hat{p} \in \mathcal{P}_n, \hat{q} \in \mathcal{P}_m$ having a non-trivial GCD, has been studied in the literature, where several criteria have been used in order to specify the *nearness* property (see e.g. [3, 10, 11, 37, 47, 4, 5, 40] and the references therein). In many cases, for given polynomials p and q , and a tolerance ε , the methods aim to find the degree of an ε -GCD, a set of perturbations $\delta p, \delta q$ (such that $\hat{p} = p + \delta p, \hat{q} = q + \delta q$) and an ε -GCD (w.r.t. the perturbations) without addressing the minimization in Problem 1.3 directly. Here instead we look for the pair with minimal distance (in the 2-norm) and allow further constraints on the coefficients of the polynomials (see Section 6).

Problem 1.3 is a nonconvex optimization problem and can be approached by global optimization, local optimization [7, 29, 35], and convex relaxation methods [12]. The methods based on global optimization, such as the branch and bound method [2], are too expensive for most real-life problems.

In this paper, we consider the local optimization approach. A nonstandard feature of the problem is that the optimization is over a Grassman manifold [1, 44]. Our main contribution is a new method based on integration of a gradient system of ordinary

differential equations. The system of ordinary differential equations describes the gradient dynamics associated to an appropriate cost functional. It is given by the modulus of the smallest eigenvalue of the Sylvester matrix. (In principle, we may replace it by the smallest singular value and obtain a similar system of ODEs.) The method is globally convergent to a locally optimal solution. Our simulation results indicate that it is more robust to the initial approximation than the Newton-type methods.

In addition, we incorporate constraint that some of the polynomials' coefficients p and q are known exactly, e.g., p monic, q monic. Such constraints can not be treated by the existing alternative methods in the literature. Another extension, discussed in Section 6), is the case of complex polynomials.

Alternative methods for solving Problem 1.3 based on local optimization are developed in the structured low-rank approximation setting [30, 32]. In particular, the method of [33] which makes use of the kernel representation of the rank constraint and the variable projections as well as the method of [22] using the image representation of the rank constraint. Also homotopy methods can be used to compute a locally optimal solution of Problem 1.3. The methods of [33, 22] can not impose arbitrary constraints on the coefficients of the approximating polynomials \hat{p} and \hat{q} which is a limitation of these approaches with respect to the one proposed here. In all numerical examples, shown in Section 7, where the polynomials \hat{p} and \hat{q} are unconstrained or where only \hat{p} or \hat{q} is constrained to be monic, the proposed method finds a solution with the same value of the cost function as the one found by the methods of [33, 22].

The paper is organized as follows. In Section 2 we set the problem and the notation and give the mathematical framework. In Section 3 we introduce the structured ε -pseudospectrum of a Sylvester matrix, which plays a fundamental role in our methodology. In Section 4 we obtain a gradient system for the smallest eigenvalue(s) of a Sylvester matrix (under additive perturbations) which allows us to approximate the distance of the structured ε -pseudospectrum to the origin. In Section 5 we propose an iterative two-level convergent method to approximate the distance to singularity of a Sylvester matrix, which is equivalent to find a nearby pair of polynomials having a non-trivial GCD. In Section 6 we consider some extensions of the considered methodology to problems with additional constraints. Finally in Section 7 we present some numerical tests.

Notation.

- Π_s — set of polynomials of degree at most s
- $\mu(S)$ — inner spectral radius of the matrix S
- $\Lambda_\varepsilon^S(S)$ — structured ε -pseudospectrum
- $\mu_\varepsilon(S)$ — inner ε -pseudospectral radius of S
- $\Lambda(S)$ — spectrum of S
- $\text{Syl}(a, b)$ — Sylvester matrix (2.1)
- \mathcal{S} — set of Sylvester structured matrices
- $\|\cdot\|_F$ — Frobenius norm
- I — identity matrix
- $\mathbf{1} = (1 \ 1 \ \dots \ 1)^T$
- for real matrices A, B , $\langle A, B \rangle = \text{trace}(A^T B)$ — Frobenius inner product

2. Preliminaries. Consider the polynomials (1.1) with real coefficients $\{a_i\}$ and $\{b_i\}$ and with $m \leq n$, where we may set $a_n = 1$ (p monic). Also, by setting $b_{m+1} = \dots = b_n = 0$, we can always consider the case $m = n$. In order to check

whether the polynomials p and q are coprime it is natural to introduce the associated Sylvester matrix of dimension $2n \times 2n$,

$$S = \text{Syl}(a, b) := \begin{pmatrix} a_n & \dots & a_m & \dots & \dots & a_1 & a_0 & 0 & \dots & 0 \\ 0 & a_n & \dots & a_m & \dots & \dots & a_1 & a_0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_n & \dots & \dots & \dots & \dots & a_1 & a_0 \\ 0 & \dots & b_m & \dots & \dots & b_1 & b_0 & 0 & \dots & 0 \\ 0 & \dots & 0 & b_m & \dots & \dots & b_1 & b_0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 0 & b_m & \dots & \dots & b_1 & b_0 \end{pmatrix}. \quad (2.1)$$

Then, we have the following well-known result due to S. Sylvester [39].

THEOREM 2.1. *A pair of polynomials p, q is coprime if and only if the matrix $\text{Syl}(a, b)$ given by (2.1) is nonsingular.*

Let

$$\begin{aligned} \delta p(z) &= \delta a_s z^s + \delta a_{s-1} z^{s-1} + \dots + \delta a_1 z + \delta a_0 \\ \delta q(z) &= \delta b_m z^m + \delta b_{m-1} z^{m-1} + \dots + \delta b_1 z + \delta b_0 \end{aligned} \quad (2.2)$$

where $s = n - 1$ if $p + \delta p$ is constrained to be monic and $s = n$ otherwise, and the vectors of their coefficients $\delta a = (\delta a_0 \ \delta a_1 \ \dots \ \delta a_s)^\top$ and $\delta b = (\delta b_0 \ \delta b_1 \ \dots \ \delta b_m)^\top$. With this notation problem (1.3) can be restated as

$$\begin{aligned} d_{CD}(p, q) &= \sup\{\varepsilon : (p + \delta p, q + \delta q) \text{ is coprime for all } \delta p \in \Pi_s, \delta q \in \Pi_m, \\ &\text{such that } \text{dist}((p, q), (p + \delta p, q + \delta q)) < \varepsilon\}. \end{aligned} \quad (2.3)$$

We are mainly considering two different distances to common divisibility: one with p monic and one without this constraint. Another interesting case is when only a few coefficients of the polynomials are subject to perturbations; for an extension to this case see Section 6.

REMARK 2.1. Generically, the smallest perturbations which make a pair of polynomials to have a nontrivial GCD, creates either one real common root or a pair of complex conjugate common roots. It is well-known that in the first case the co-rank of the associated Sylvester matrix is one, while in the second case it is two. \diamond

3. General framework. The problem we consider here can be cast into a more widespread class of problems, which we are briefly introducing.

3.1. Structured ε -pseudospectrum. This paper treats a special instance of a more general problem, that is to determine the distance to singularity (with respect to Frobenius norm) of a matrix having an affine structure. To be precise, for a given real matrix A , we look for a singular matrix $B = A + X$, with X belonging to a real linear manifold of matrices \mathcal{M} , such that $\|X\|$ is minimal. Let us define the *inner spectral radius* of a matrix B as

$$\mu(B) = \min\{|\lambda| : \lambda \in \Lambda(B)\},$$

where $\Lambda(B)$ is the spectrum of B ; if $\mu(B) = 0$ then B is singular. With this notation the problem we aim to deal with is the following:

$$\min_{X \in \mathcal{M} : \mu(A+X)=0} \|X\|_F. \quad (3.1)$$

Let us write $X \in \mathcal{M}$ as $X = \varepsilon E$ with $E \in \mathcal{M}$ such that $\|E\|_F = 1$. The basic idea to solve this problem is that of introducing the structured ε -pseudospectrum of A ,

$$\Lambda_\varepsilon^{\mathcal{M}}(A) = \{ \lambda \in \Lambda(A + \varepsilon E) : E \in \mathcal{M}, \|E\|_F \leq 1 \} \quad (3.2)$$

and looking for the minimal value of ε such that $0 \in \Lambda_\varepsilon^{\mathcal{M}}(A)$. This requires—for a given ε —to minimize $|\lambda|$ for $\lambda \in \Lambda_\varepsilon^{\mathcal{M}}(A)$ (which means minimizing $\mu(A + \varepsilon E)$ for $E \in \mathcal{M}$, $\|E\|_F \leq 1$) and then iterating on ε . Usually it is possible to prove that it is equivalent to replace the inequality constraint $\|E\|_F \leq 1$ by the equality constraint $\|E\|_F = 1$. If such a property holds, the minimization step can be pursued by computing the gradient of the functional $F_\varepsilon(E) = \mu(A + \varepsilon E)$ for $E \in \mathcal{M}$, $\|E\|_F = 1$ and applying a descent method. Indicating by G the free gradient of $F_\varepsilon(E)$ in the space of complex matrices, which is a rank-1 matrix obtained by left and right eigenvectors of $A + \varepsilon E$, the orthogonal projection of the gradient onto \mathcal{M} , with the additional constraint $\|E\|_F = 1$ is simply given by

$$P_{\mathcal{M}}(G) - \langle E, P_{\mathcal{M}}(G) \rangle E$$

where for a pair of matrices A and B , $\langle A, B \rangle$ denotes the Frobenius inner product and $P_{\mathcal{M}}(G)$ denotes the orthogonal projection of the matrix G onto \mathcal{M} . Understanding the dependence of $\Lambda_\varepsilon^{\mathcal{M}}(A)$ on ε usually allows to obtain fast methods to compute the minimal ε such that $F_\varepsilon(E) = 0$. Naturally this approach provides usually a local extremizer which is not guaranteed to be global.

3.2. Distance to common divisibility. The basic observation used to calculate the distance to common divisibility $d_{CD}(p, q)$ is that (2.3) is equivalent to the following problem

$$d_{CD}(p, q) = \frac{1}{\sqrt{n}} \inf \left\{ \varepsilon : \mu(\text{Syl}(a, b) + \varepsilon E) > 0 \text{ for all } E \in \mathcal{S}^{n,m}, \|E\|_F \leq 1 \right\} \quad (3.3)$$

where (with $m \leq n$)

$$\mathcal{S}^{n,m} := \{ \text{Syl}(u, v) : u \in \mathbb{R}^{n+1}, v \in \mathbb{R}^{m+1} \} \subset \mathbb{R}^{2n \times 2n} \quad (3.4)$$

is the set of real Sylvester matrices (see (2.1)). For brevity in the sequel we simply denote $\mathcal{S}^{n,m}$ by \mathcal{S} .

This is a classical matrix nearness problem of the form considered in Section 3.1 (see e.g. [20]). Denote by

$$\Lambda_\varepsilon^{\mathcal{S}}(S) = \{ \lambda \in \Lambda(S + \varepsilon E) : E \in \mathcal{S}, \|E\|_F \leq 1 \}.$$

the structured ε -pseudospectrum (see [41]). Note that \mathcal{S} is a smooth linear manifold which implies that

$$S + \varepsilon E \in \mathcal{S}, \quad \text{if } E \in \mathcal{S}.$$

Example. Consider the two polynomials of degree 3,

$$\begin{aligned} p(z) &= z^3 + 2z^2 + 2z + 2 \\ q(z) &= 2z^3 + z - 2 \end{aligned}$$

where p is constrained to be monic. Here $a = (1 \ 1 \ 2 \ 2)^T$ and $b = (2 \ 0 \ 1 \ -2)^T$; the corresponding Sylvester matrix is given by

$$\text{Syl}(a, b) = \begin{pmatrix} 1 & 2 & 2 & 2 & 0 & 0 \\ 0 & 1 & 2 & 2 & 2 & 0 \\ 0 & 0 & 1 & 2 & 2 & 2 \\ 2 & 0 & 1 & -2 & 0 & 0 \\ 0 & 2 & 0 & 1 & -2 & 0 \\ 0 & 0 & 2 & 0 & 1 & -2 \end{pmatrix} \quad (3.5)$$

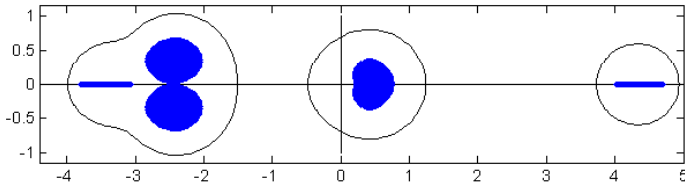


FIG. 3.1. The approximated structured ε -pseudospectrum for $\varepsilon = \frac{1}{2}$ for Example (3.5) is filled with blue; the boundary of the unstructured ε -pseudospectrum is drawn in black.

The set $\Lambda_\varepsilon^S(S)$ for $\varepsilon = \frac{1}{2}$ is approximated by dense sampling on the set of admissible perturbations and is plotted in blue in Figure 3.1. The black curve represents the boundary of the corresponding unstructured ε -pseudospectrum, which means that arbitrary complex perturbations of norm bounded by $\frac{1}{2}$ are considered.

Next, we define $\mu_\varepsilon(S)$, the *inner ε -pseudospectral radius* of S , which is the minimum of the modulus of the elements of the structured ε -pseudospectrum,

$$\mu_\varepsilon(S) = \min\{|\lambda| : \lambda \in \Lambda_\varepsilon^S(S)\}. \quad (3.6)$$

Note that the case $\varepsilon = 0$, reduces $\mu_\varepsilon(S)$ to the inner spectral radius $\mu(S)$.

With this notation we characterize the distance to common divisibility as

$$d_{CD}(p, q) = \frac{1}{\sqrt{n}} \min\{\varepsilon : \mu_\varepsilon(S) = 0\}, \quad \text{where } S = \text{Syl}(a, b).$$

If S is associated to a coprime pair, we have that

$$\mu_\varepsilon(S) > 0 \quad \iff \quad d_{CD}(p, q) > \varepsilon.$$

3.3. A 2-level methodology. To find the distance to common divisibility we have to solve the equation (w.r.t. ε), $\mu_\varepsilon(S) = 0$ and find the minimum root,

$$\varepsilon_{\text{opt}} = \min\{\varepsilon : \mu_\varepsilon(S) = 0\}.$$

This is a global optimization problem which we approach locally, by introducing a methodology which allows to compute suitable upper bounds for ε_{opt} . The following definition will be used often, where, for $\delta > 0$, $\mathcal{B}_\delta(\lambda) = \{z \in \mathbb{C} : |z - \lambda| \leq \delta\}$.

DEFINITION 3.1. A matrix E such that $\|E\|_F \leq 1$ and $S + \varepsilon E$ has a smallest eigenvalue z that locally minimizes the modulus of $\Lambda_\varepsilon^S(S)$, i.e. $\exists \delta > 0$ such that

$$\lambda = \arg \min\left\{|z| : z \in \Lambda_\varepsilon^S(S) \cap \mathcal{B}_\delta(\lambda)\right\}, \quad (3.7)$$

is called a local extremizer. Similarly λ is called a local minimum point of $\Lambda_\varepsilon^S(S)$.

We propose a two-level algorithm.

At the inner level, for a fixed ε we compute a (local) minimum point of $\Lambda_\varepsilon^S(S)$, which we denote by $\lambda(\varepsilon) \in \partial\Lambda_\varepsilon^S(S)$. If $\lambda(\varepsilon)$ is a global minimum point then $|\lambda(\varepsilon)| = \mu_\varepsilon(S)$, otherwise $|\lambda(\varepsilon)| > \mu_\varepsilon(S)$. The (inner) algorithm we propose finds local optima of problem (3.7) by determining the stationary point of a system of ODEs. In general there is no assurance that these are global minima, although this seems to be the case in all our experiments of small dimension (where we were able to perform a statistical investigation on a very large number of samples).

For the outer iteration, we indicate by $\lambda(\varepsilon)$ a continuous branch of local minima of $\Lambda_\varepsilon^S(S)$ (see (3.7)); the aim is to compute

$$\varepsilon^* = \min\{\varepsilon : \lambda(\varepsilon) = 0\}.$$

In order to compute ε^* we vary ε by an interpolation based iteration which exploits the knowledge of the exact derivative of $\lambda(\varepsilon)$ with respect to ε and exhibits fast convergence (a similar methodology has been exploited for different structures, see [28], and also for computing the H_∞ norm of a linear dynamical system [14] and the distance to instability of real matrices, see [17]).

For ε in a left neighborhood of ε^* we expect generically the occurrence of one of the following two situations:

- (i) There is a unique real local minimum $\lambda(\varepsilon)$. This means that there exists a matrix $E(\varepsilon) \in \mathcal{S}$ of unit norm such that $\lambda(\varepsilon)$ is a real simple eigenvalue of $S + \varepsilon E(\varepsilon)$, which implies that $S + \varepsilon^* E(\varepsilon^*)$ has co-rank equal to 1 and the two perturbed polynomials associated to $S + \varepsilon^* E(\varepsilon^*)$ have a real common root.
- (ii) There is a unique pair of complex conjugate local minima $\lambda(\varepsilon)$ and $\overline{\lambda(\varepsilon)}$. This means that there exists a matrix $E(\varepsilon) \in \mathcal{S}$ of unit norm such that $\lambda(\varepsilon), \overline{\lambda(\varepsilon)}$ is a pair of complex conjugate eigenvalues of $S + \varepsilon E(\varepsilon)$, which implies that $S + \varepsilon^* E(\varepsilon^*)$ has co-rank equal to 2 and the two perturbed polynomials associated to $S + \varepsilon^* E(\varepsilon^*)$ have two complex conjugate common roots.

This means that—contrarily to the case of unstructured perturbations—we expect that as $\varepsilon \rightarrow \varepsilon^*$ we have to expect a *non-defective coalescence* of two complex conjugate eigenvalues in zero.

4. Approximation of local minima of the structured ε -pseudospectrum.

We address here the minimization problem

$$\lambda = \arg \min\{|z| : z \in \Lambda_\varepsilon^S(S)\}.$$

We will use the convention that when an eigenvalue of minimum modulus λ is not real, which means it appears pairwise with $\overline{\lambda}$, then we select the eigenvalue with positive imaginary part.

The idea is that to make use of a continuous (in time) minimization method. Let us consider a smooth matrix valued function $S + \varepsilon E(t)$ where $E(t) \in \mathcal{S}$ and $\|E(t)\|_F \leq 1$ for all t .

Our goal is to find an optimal direction $\dot{E}(t) = Z$ such that the smallest eigenvalue λ of $S + \varepsilon E(t)$, is characterized (locally) by the maximal possible decrease.

We follow an approach which extends to structured pseudospectra some ideas developed in [15, 16, 18] and use of the following standard perturbation result for eigenvalues see, e.g., [27, Section II.1.1]. Here and in the following, we denote $\dot{\cdot} = d/dt$.

LEMMA 4.1. *Consider the differentiable matrix valued function $C(t)$ for t in a neighborhood of 0. Let $\lambda(t)$ be an eigenvalue of $C(t)$ converging to a simple eigenvalue*

λ_0 of $C_0 = C(0)$ as $t \rightarrow 0$. Let y_0 and x_0 be left and right eigenvectors, respectively, of C_0 corresponding to λ_0 , that is, $(C_0 - \lambda_0 I)x_0 = 0$ and $y_0^*(C_0 - \lambda_0 I) = 0$. Then, $y_0^*x_0 \neq 0$ and $\lambda(t)$ is differentiable near $t = 0$ with

$$\dot{\lambda}(0) = \frac{y_0^* \dot{C}(0)x_0}{y_0^*x_0}.$$

Assume that $E(t)$ is a smooth function and observe that for a simple eigenvalue $\lambda(t) = r(t)e^{i\theta(t)}$ of the matrix-valued function $S + \varepsilon E(t)$ ($r(t)$ denotes the modulus and $\theta(t)$ the phase), with associated left and right eigenvectors $y(t)$ and $x(t)$ respectively, we have (omitting the dependence on t)

$$\begin{aligned} \frac{d}{dt}|\lambda| &= \frac{1}{2|\lambda|} \frac{d}{dt}|\lambda|^2 = \frac{1}{|\lambda|} \operatorname{Re}(\bar{\lambda}\dot{\lambda}) = \frac{1}{|\lambda|} \operatorname{Re}\left(\bar{\lambda}\varepsilon \frac{y^* \dot{E}x}{y^*x}\right) \\ &= \frac{\varepsilon}{|\lambda|} \operatorname{Re}\left(\frac{(\lambda y)^* \dot{E}x}{y^*x}\right) = \varepsilon \operatorname{Re}\left(\frac{y^* \dot{E}x}{e^{i\theta} y^*x}\right). \end{aligned}$$

In the sequel of the paper we shall always impose the following normalization to the eigenvectors y and x ,

$$\|y\|_2 = \|x\|_2 = 1, \quad y^*x = |y^*x|e^{-i\theta} \quad (4.1)$$

which makes the denominator of (4.2) real and positive (note that $|y^*x| \neq 0$ is a consequence of the assumption that λ is simple). Hence we have

$$\frac{d}{dt}|\lambda| = \frac{\varepsilon}{|y^*x|} \operatorname{Re}\left(y^* \dot{E}x\right) \quad (4.2)$$

so that the optimal variation $Z = \dot{E}$ is obtained by minimizing the function $\operatorname{Re}(y^* \dot{E}x)$. Note that for $\dot{E} \in \mathcal{S}$ we have

$$\operatorname{Re}\left(y^* \dot{E}x\right) = \operatorname{Re}\langle yx^*, \dot{E} \rangle = \langle P_{\mathcal{S}}(yx^*), \dot{E} \rangle,$$

where we denote by $P_{\mathcal{S}}(B)$ the orthogonal projection of a matrix $B \in \mathbb{C}^{2n, 2n}$ onto \mathcal{S} . The following result gives an explicit formula for $P_{\mathcal{S}}$ (for Toeplitz matrices similar results are discussed in [36] and [8]).

LEMMA 4.2. *Let $\mathcal{S} \subset \mathbb{R}^{2n \times 2n}$ be the manifold of real Sylvester matrices of dimension $2n$ and $B \in \mathbb{C}^{2n \times 2n}$. The orthogonal projection (with respect to the Frobenius inner product $\langle \cdot, \cdot \rangle$) $P_{\mathcal{S}}(B)$ of B onto \mathcal{S} is given by*

$$B_{\mathcal{S}} = P_{\mathcal{S}}(B) = \operatorname{Syl}(\alpha, \beta) \quad (4.3)$$

where

$$\begin{aligned} \alpha_{n-k} &= \frac{1}{n} \sum_{l=1}^n \operatorname{Re}(B_{l, l+k}), \quad k = k_0, \dots, n \\ \beta_{m-k} &= \frac{1}{n} \sum_{l=1}^n \operatorname{Re}(B_{n+l, n-m+l+k}), \quad k = 0, \dots, m \end{aligned}$$

with $k_0 = 1$ if p is constrained to be monic, and $k_0 = 0$ otherwise (and $\alpha_n = 0$).

Proof. Being \mathcal{S} a manifold of real matrices, note the obvious property $P_{\mathcal{S}}(B) = P_{\mathcal{S}}(\operatorname{Re}(B))$. We have to find $B_{\mathcal{S}} \in \mathcal{S}$ such that

$$B_{\mathcal{S}} = \arg \min_{S \in \mathcal{S}} \|B - S\|_F = \arg \min_{S \in \mathcal{S}} \|\operatorname{Re}(B) - S\|_F.$$

The result follows directly from the property that for a real vector $x \in \mathbb{R}^d$,

$$\nu_* = \arg \min_{\nu \in \mathbb{R}} \|x - \nu \mathbf{1}\|_F = \frac{1}{d} \sum_{i=1}^d \operatorname{Re}(x_i),$$

being $\mathbf{1} = (1 \ 1 \ \dots \ 1)^T$. \square

The following result assures the important property $P_{\mathcal{S}}(yx^*) \neq 0$ (where x and y are the eigenvectors of $S + \varepsilon E$ associated to the smallest eigenvalue λ), which is considered later in Lemmas 4.4 and 4.5. In order to distinguish the case where p is unconstrained from the case where p is constrained to be monic, we introduce the set \mathcal{S}^* which is the submanifold of Sylvester matrices $\{\operatorname{Syl}(u, v)\}$ (2.1) given by $u_n = 0$.

LEMMA 4.3. *Let $S \in \mathcal{S}$ and either $E \in \mathcal{S}$ or $E \in \mathcal{S}^*$ of unit Frobenius norm, and $\varepsilon > 0$. If $\lambda \neq 0$ is a simple eigenvalue of $S + \varepsilon E$, with left and right eigenvectors y and x scaled according to (4.1), then*

$$P_{\mathcal{S}}(yx^*) \neq 0. \quad (4.4)$$

Proof. We analyze first the case where p is not constrained to be monic.

Let y and x be the left and right eigenvectors of $S + \varepsilon E$ associated to $\lambda = re^{i\theta}$, with $r > 0$. Assume — by contradiction — that $P_{\mathcal{S}}(yx^*) = 0$; this would imply

$$0 = \langle P_{\mathcal{S}}(yx^*), S + \varepsilon E \rangle = \langle yx^*, S + \varepsilon E \rangle = \langle \operatorname{Re}(yx^*), S + \varepsilon E \rangle. \quad (4.5)$$

Observing that

$$\begin{aligned} \langle \operatorname{Re}(yx^*), S + \varepsilon E \rangle &= \operatorname{Re} \langle yx^*, S + \varepsilon E \rangle \\ &= \operatorname{Re}(y^*(S + \varepsilon E)x) = \operatorname{Re}(re^{i\theta}y^*x), \end{aligned}$$

and exploiting the normalization (4.1), we obtain

$$\langle \operatorname{Re}(yx^*), S + \varepsilon E \rangle = r|y^*x| > 0 \quad (4.6)$$

where positivity follows by the simplicity assumption for λ . This would contradict (4.5) and consequently (4.4) holds true.

Second we consider the case where p is constrained to be monic. If we assume that $P_{\mathcal{S}^*}(yx^*) = 0$, where the projection $P_{\mathcal{S}^*}$ — which is given by (4.3) by imposing $\alpha_n = 0$ — is relevant to the monic case and is used here to distinguish it from $P_{\mathcal{S}}$, we get

$$P_{\mathcal{S}}(yx^*) = \begin{pmatrix} \beta I & 0 \\ 0 & 0 \end{pmatrix} \quad (4.7)$$

where $P_{\mathcal{S}}$ is the usual projection on the manifold \mathcal{S} to which belongs S (which now contains the submanifold \mathcal{S}^* to which belongs E) and

$$\beta = \frac{1}{n} \sum_{i=1}^n \operatorname{Re}(y_i \bar{x}_i).$$

Now, consider the matrix $C = S + \varepsilon E - I$, where I is the identity matrix, and define the matrix

$$\tilde{S}_\gamma := I + \gamma C$$

which preserves the structure of S and also the eigenvectors x and y associated to the shifted eigenvalue λ .

First — by (4.7) — we obtain (recall that $a_n = 1$ in (2.1))

$$\langle \operatorname{Re}(yx^*), \tilde{S}_\gamma \rangle = n\beta = \operatorname{Re} \left(\sum_{i=1}^n y_i \bar{x}_i \right)$$

which has modulus smaller than 1.

Second, exploiting $\langle \operatorname{Re}(yx^*), \tilde{S}_\gamma \rangle = \operatorname{Re} \left(y^* \tilde{S}_\gamma x \right)$, we get

$$\langle \operatorname{Re}(yx^*), \tilde{S}_\gamma \rangle = \tilde{r}_\gamma \quad (4.8)$$

where $|\tilde{r}_\gamma|$ can be chosen arbitrarily large if $|\gamma|$ is chosen large enough.

This leads to a contradiction. As a consequence we have that $P_{S^*}(yx^*) \neq 0$. \square

4.1. Minimizing on the sphere. The problem we are considering is

$$\min \{ |\lambda| : \lambda \in \Lambda(S + \varepsilon E), \quad E \in \mathcal{S}, \quad \|E\|_F \leq 1 \}. \quad (4.9)$$

Let us show that extremizers are located on the sphere, that is we may replace (4.9) by

$$\min \{ |\lambda| : \lambda \in \Lambda(S + \varepsilon E), \quad E \in \mathcal{S}, \quad \|E\|_F = 1 \}. \quad (4.10)$$

In order to do this let us state the following result.

LEMMA 4.4. *Assume that E is a (local) extremizer for Problem (4.9) where $\lambda \in \Lambda_\varepsilon^S(S)$, $\lambda \neq 0$ has (locally) minimum modulus and is simple. Then $\|E\|_F = 1$.*

Proof. Assume by contradiction $\|E\|_F < 1$. Let y and x be the left and right eigenvectors of $S + \varepsilon E$ associated to λ . By Lemma 4.3 we have that $P_S(yx^*) \neq 0$. Note that

$$\frac{d}{dt} |\lambda| = \frac{\varepsilon}{|y^* x|} \langle P_S(yx^*), \dot{E} \rangle$$

implies that $-P_S(yx^*)$ is a descent direction for $|\lambda|$; as a consequence consider the matrix $E_\delta = E - \delta P_S(yx^*)$ and denote by λ_δ its eigenvalue of smallest modulus. For a sufficiently small $\delta > 0$, the matrix $E_\delta = E - \delta P_S(yx^*)$ would be such that $|\lambda_\delta| < |\lambda|$ and $\|E_\delta\|_F < 1$, contradicting optimality of E . \square

As a consequence we are justified in looking extremizers on the set

$$\mathcal{S}_1 = \mathcal{S} \cap \{E \in \mathbb{R}^{2n \times 2n} : \|E\|_F = 1\}. \quad (4.11)$$

4.2. Steepest descent direction. Let $\lambda = re^{i\theta} \neq 0$ be the eigenvalue of minimum modulus of $S + \varepsilon E$. Then the optimal steepest descent direction for $|\lambda|$ (see (4.2) and (4.1)), with $Z = \dot{E} \in \mathcal{S}$, is given by:

$$\begin{aligned} Z_* &= \arg \min_{Z \in \mathcal{S}_1} \operatorname{Re}(y^* Z x) \\ &\text{subject to } \langle E, Z \rangle = 0, \end{aligned} \quad (4.12)$$

where the constraint $\langle E, Z \rangle = 0$ guarantees norm conservation of E and the normalization $Z \in \mathcal{S}_1$, is considered for convenience to obtain a unique solution (since Z represents indeed a direction).

The solution to (4.12) is given in the following lemma.

LEMMA 4.5. *Let $E \in \mathcal{S}_1$ be a $2n \times 2n$ real matrix of unit Frobenius norm, and $y, x \in \mathbb{C}^{2n}$ be non-zero complex vectors. Assume that $P_{\mathcal{S}}(yx^*) \neq 0$ and E is not proportional to $P_{\mathcal{S}}(yx^*)$. Then the solution of the optimization problem (4.12) is given by*

$$\nu Z_* = -P_{\mathcal{S}}(yx^*) + \langle E, P_{\mathcal{S}}(yx^*) \rangle E \quad (4.13)$$

where ν is the Frobenius norm of the matrix on the right hand side.

Proof. We have for the function to minimize,

$$\operatorname{Re}(y^* Z x) = \operatorname{Re}\langle Z, yx^* \rangle,$$

that is a linear function with respect to Z . Since \mathcal{S} and $E_{\perp} = \{Z : \langle E, Z \rangle = 0\}$ are linear subspaces, it is direct to see, by the fact that the inner product with a given vector is minimized over a subspace by orthogonally projecting the vector onto that subspace, that the solution to (4.12) is given by a matrix proportional to the orthogonal projection of the rank-1 matrix yx^* onto the linear subspace $\mathcal{S} \cap E_{\perp}$, which we denote by $P_{\mathcal{S} \cap E_{\perp}}(yx^*)$, scaled to have unit norm (unless such projection is zero). Let $P_{\mathcal{S}}(\cdot)$ and $P_{E_{\perp}}(\cdot)$ denote the orthogonal projections onto \mathcal{S} and E_{\perp} respectively. By the well-known Von Neumann iterative formula (see e.g. [42]), we have that if $P_{\mathcal{S}}$ and $P_{E_{\perp}}$ commute then

$$P_{\mathcal{S} \cap E_{\perp}}(\cdot) = P_{\mathcal{S}}(P_{E_{\perp}}(\cdot)) = P_{E_{\perp}}(P_{\mathcal{S}}(\cdot)).$$

Let $B \in \mathbb{C}^{2n \times 2n}$; then

$$P_{E_{\perp}}(P_{\mathcal{S}}(B)) = P_{\mathcal{S}}(B) - \langle E, P_{\mathcal{S}}(B) \rangle E.$$

On the other hand (since $E \in \mathcal{S}$)

$$P_{\mathcal{S}}(P_{E_{\perp}}(\cdot)) = P_{\mathcal{S}}(B - \langle E, B \rangle E) = P_{\mathcal{S}}(B) - \langle E, P_{\mathcal{S}}(B) \rangle E,$$

which proves the commutativity. Hence $P_{\mathcal{S} \cap E_{\perp}}(B) = P_{\mathcal{S}}(B) - \langle E, P_{\mathcal{S}}(B) \rangle E$. Since—by assumption— $P_{\mathcal{S}}(yx^*) - \langle E, P_{\mathcal{S}}(yx^*) \rangle E$ does not vanish we obtain the solution to (4.12)

$$\nu Z_* = -P_{\mathcal{S}}(yx^*) - \langle E, P_{\mathcal{S}}(yx^*) \rangle E \neq 0,$$

where ν is the reciprocal of the Frobenius norm of the right hand side. \square

Optimality conditions. The application of Karush-Kuhn-Tucker local optimality conditions for the optimization problem (4.10),

$$\min_{E \in \mathcal{S}, \|E\|_F=1} |\lambda|$$

in the case where the minimum is not zero, read as follows:

$$P_{\mathcal{S}}(yx^*) = sE_*, \quad s < 0 \quad (4.14)$$

$$\|E_*\|_F = 1 \quad (4.15)$$

where E_* denotes a extremizer. Naturally $s > 0$ would characterize a local maximum.

4.3. The gradient system associated to the minimization problem.

Lemma 4.5 and formula (4.2) suggest to consider the following differential equation on the manifold \mathcal{S}_1 (see (4.11)),

$$\dot{E} = -P_S(yx^*) + \langle E, P_S(yx^*) \rangle E \quad (4.16)$$

where $y(t), x(t)$ are left and right eigenvectors of unit norm respectively to a simple eigenvalue $\lambda(t)$ of $S + \varepsilon E(t)$, and with $y^*x = |y^*x|e^{-i\theta}$, where ε is fixed.

REMARK 4.1. Indeed Lemma 4.5 gives us a steepest descent direction and (4.16) scales this direction in a way that the the projected gradient has not necessarily unit norm, but has the order of magnitude of $P_S(yx^*)$ so that to prevent a big amplification of this vector-field when this is small (close to convergence). \diamond

We are in the position to prove the monotonic decrease of $|\lambda(t)|$ along every solution of (4.16).

THEOREM 4.6. *Let $E(t)$ of unit Frobenius norm satisfy the differential equation (4.16). If $\lambda(t)$ is a simple eigenvalue of $S + \varepsilon E(t)$, then*

$$\frac{d}{dt}|\lambda(t)| \leq 0. \quad (4.17)$$

Proof. Note that

$$\operatorname{Re}(y^* P_S(yx^*) x) = \operatorname{Re}\langle yx^*, P_S(yx^*) \rangle = \langle P_S(yx^*), P_S(yx^*) \rangle = \|P_S(yx^*)\|_F^2,$$

and (since $E \in \mathcal{S}$)

$$\operatorname{Re}(y^* E x) = \langle E, P_S(yx^*) \rangle.$$

By the Cauchy–Schwarz inequality,

$$|\langle E, P_S(yx^*) \rangle| \leq \|E\|_F \|P_S(yx^*)\|_F = \|P_S(yx^*)\|_F.$$

Finally, by (4.16),

$$\operatorname{Re}(y^* \dot{E} x) = \left(-\|P_S(yx^*)\|_F^2 + \langle E, P_S(yx^*) \rangle^2 \right) \leq 0, \quad (4.18)$$

implying (4.17) by Lemma 4.1. \square

Since we are interested to minimize $|\lambda|$ we address our attention to the stationary points of (4.16).

4.4. Stationary points. Since stationary points of (4.16) are potential extremizers for $\mu_\varepsilon(S)$, we give the following result for their characterization.

THEOREM 4.7. *Assume $\lambda \neq 0$. The following are equivalent on solutions of (4.16):*

- (1) $\frac{d}{dt}|\lambda| = 0$;
- (2) $\dot{E} = 0$;
- (3) E is a real multiple of $P_S(yx^*)$.

Proof. The proof follows directly by combining (4.2) and Lemma 4.5. \square

The following result characterizes the local extremizers.

THEOREM 4.8. *Let $E_* \in \mathcal{S}$ with $\|E_*\|_F = 1$. Let $\lambda_* = re^{i\theta} \neq 0$ be a simple eigenvalue of $S + \varepsilon E_*$ with minimum modulus, with left and right eigenvectors y and x , respectively, both of unit norm and with the normalization $y^*x = |y^*x|e^{-i\theta}$. Then the following two statements are equivalent:*

- (i) Every differentiable path $(E(t), \lambda(t))$ (for small $t \geq 0$) such that $\|E(t)\|_F \leq 1$ and $\lambda(t)$ is an eigenvalue of $S + \varepsilon E(t)$, with $E(0) = E_*$ and $\lambda(0) = \lambda_*$, has

$$\left. \frac{d}{dt} |\lambda(t)| \right|_{t=0} \geq 0.$$

- (ii) E_* is a negative multiple of $P_{\mathcal{S}}(yx^*)$.

Proof. Assume that (i) does not hold true. Then there is some path $E(t)$ through E_* such that $\left. \frac{d}{dt} |\lambda(t)| \right|_{t=0} < 0$; thus Lemma 4.5 together with Lemma 4.1 shows that also the solution path of (4.16) passing through E_* is such a path. Consequently E_* is not a stationary point of (4.16), and Theorem 4.7 then yields that E_* is not a real multiple of $P_{\mathcal{S}}(yx^*)$. This implies that also (ii) does not hold true.

Vice versa, if E_* is not a real multiple of $P_{\mathcal{S}}(yx^*)$, then E_* is not a stationary point of (4.16), and Theorems 4.6 and 4.7 yield that $\left. \frac{d}{dt} |\lambda(t)| \right|_{t=0} < 0$ along the solution path of (4.16). Moreover, using a similar argument to [16, Theorem 2.2], if

$$E_* = \gamma P_{\mathcal{S}}(yx^*), \quad \text{with } \gamma > 0,$$

then along the path $E(t) = (1-t)E_*$, $t \geq 0$, we have that

$$\operatorname{Re}(y^* \dot{E}(0)x) = -\gamma \|P_{\mathcal{S}}(yx^*)\|_F^2 < 0$$

and thus, by exploiting Lemma 4.1, $\left. \frac{d}{dt} |\lambda(t)| \right|_{t=0} < 0$, which contradicts (i). \square

As a consequence, if in Theorem 4.7 $\lambda \neq 0$ is locally minimal (in modulus),

$$E = E_* = -P_{\mathcal{S}}(yx^*) / \|P_{\mathcal{S}}(yx^*)\|_F$$

that is the projection onto \mathcal{S} of a real matrix of either rank 1 (if λ , x and y are real) or rank 2 (if λ is non real, and consequently also y and x).

REMARK 4.2. The test $E = -P_{\mathcal{S}}(yx^*) / \|P_{\mathcal{S}}(yx^*)\|_F$ might be useful to recognize that a (local) extremizer has been computed, also in different algorithms with respect to the ones presented in this article. \diamond

4.5. The system of ODEs. We can write (4.16) in a compact form for the coefficients $\{\delta a_i\}$ and $\{\delta b_i\}$ of the polynomials δp and δq (see (2.2)), which form E ,

$$E = \operatorname{Syl}(\delta a, \delta b). \quad (4.19)$$

This is given by

$$\begin{aligned} \frac{d}{dt} \delta a_k &= (\alpha_k - \eta \delta a_k) & k = 0, \dots, n - k_0 \\ \frac{d}{dt} \delta b_k &= (\beta_k - \eta \delta b_k) & k = 0, \dots, m, \end{aligned} \quad (4.20)$$

where $k_0 = 1$ if p is monic (in which case $\delta a_n \equiv 0$) and $k_0 = 0$ otherwise, α_k and β_k are the elements of $P_{\mathcal{S}}(yx^*)$,

$$P_{\mathcal{S}}(yx^*) = \operatorname{Syl}(\alpha, \beta) = \begin{pmatrix} \alpha_n & \dots & \alpha_m & \dots & \dots & \alpha_1 & \alpha_0 & 0 & \dots & 0 \\ 0 & \alpha_n & \dots & \alpha_m & \dots & \dots & \alpha_1 & \alpha_0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \alpha_n & \dots & \dots & \dots & \dots & \alpha_1 & \alpha_0 \\ 0 & \dots & \beta_m & \dots & \dots & \beta_1 & \beta_0 & 0 & \dots & 0 \\ 0 & \dots & 0 & \beta_m & \dots & \dots & \beta_1 & \beta_0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \beta_m & \dots & \dots & \beta_1 & \beta_0 \end{pmatrix}$$

and $\eta = \langle E, P_S(yx^*) \rangle$. This means we have to solve a system of $(m+1) + (n+1-k_0)$ ordinary differential equations.

REMARK 4.3. The ODEs (4.16) and (4.20) follow the smallest eigenvalue (for a complex conjugate pair the one with positive imaginary part) so that they detect automatically whether the ε -GCD has either degree 1 or 2. \diamond

In the next section we propose a numerical method to approximate the solution of the ODEs (4.20) until either a stationary point is reached or the followed eigenvalue λ goes to zero. As suggested by an anonymous referee, whom we thank, a classical alternative to Algorithm 1 would be that of using projected gradient type methods, like Zoutendijk approach (see e.g. [49]). Exploring these approaches and comparing them to the methodology we propose would be certainly an interesting development.

Algorithm 1: Euler step applied to the ODEs (4.20) with step-size control

Data: $\alpha^{(\ell)}, \beta^{(\ell)}, \lambda_\ell, y_\ell, x_\ell$ and \tilde{h}_ℓ (step size predicted by the previous step), tol (stopping tolerance).

Result: $E_{\ell+1}, y_{\ell+1}, x_{\ell+1}, \lambda_{\ell+1}$ and $\tilde{h}_{\ell+1}$.

begin

1 Set $h = \tilde{h}_\ell$.

2 Compute $Z_\ell = P_S(y_\ell x_\ell^*) := S(\alpha^{(\ell)}, \beta^{(\ell)})$ and $\eta_\ell = \langle E_\ell, Z_\ell \rangle$.

3 Compute

$$\delta a_k^{(\ell+1)} = \delta a_k^{(\ell)} - h \left(\alpha_k^{(\ell)} - \eta_\ell \delta a_k^{(\ell)} \right), \quad k = 0, \dots, n - k_0$$

$$\delta b_k^{(\ell+1)} = \delta b_k^{(\ell)} - h \left(\beta_k^{(\ell)} - \eta_\ell \delta b_k^{(\ell)} \right), \quad k = 0, \dots, m.$$

4 Compute $\sigma_{\ell+1} = n \sqrt{\sum_{k=k_0}^n \left(\delta a_k^{(\ell+1)} \right)^2 + \sum_{k=0}^m \left(\delta b_k^{(\ell+1)} \right)^2}$.

5 Normalize as

$$\delta a_k^{(\ell+1)} = \delta a_k^{(\ell+1)} / \sigma_{\ell+1}, \quad \delta b_k^{(\ell+1)} = \delta b_k^{(\ell+1)} / \sigma_{\ell+1}.$$

6 Set $E_{\ell+1} = \text{Syl}(\delta a^{(\ell+1)}, \delta b^{(\ell+1)})$.

7 Compute the eigenvalue of minimum modulus $\hat{\lambda}$ of $S + \varepsilon E_{\ell+1}$, and the left and right eigenvectors \hat{y}, \hat{x} .

8 **if** $|\hat{\lambda}| \geq |\lambda_\ell|$ **then**

 | reject the step: reduce the step size as $h := h/\gamma$ and repeat from 3;

else

 | accept the step: set $h_{\ell+1} = h$, $\lambda_{\ell+1} = \hat{\lambda}$, $y_{\ell+1} = \hat{y}$ and $x_{\ell+1} = \hat{x}$.

9 **if** $|\lambda_{\ell+1} - \lambda_\ell| \leq \text{tol}$ **or** $|\lambda_\ell| \leq \text{tol}$ **then**

 | **return**

10 **if** $h_{\ell+1} = \tilde{h}_\ell$ **then**

 | increase the step-size as $\tilde{h}_{\ell+1} := \gamma \tilde{h}_\ell$;

else

 | set $\tilde{h}_{\ell+1} = \tilde{h}_\ell$.

11 Proceed to next step

4.6. Numerical integration. Given $E_\ell \approx E(t_\ell)$ of unit Frobenius norm, and given y_ℓ and x_ℓ left and right eigenvectors of $S + \varepsilon E_\ell$ associated with its eigenvalue λ_ℓ of minimum modulus (if λ_ℓ is not real we choose $\lambda_\ell = r_\ell e^{i\theta_\ell}$ with positive imaginary part), with $y_\ell^* x_\ell = |y_\ell^* x_\ell| e^{-i\theta_\ell}$,

$$\alpha^{(\ell)} = \{\alpha_k^{(\ell)}\}_{k=0}^{n-k_0}, \quad \beta^{(\ell)} = \{\beta_k^{(\ell)}\}_{k=0}^m$$

and

$$\delta a^{(\ell)} = \{\delta a_k^{(\ell)}\}_{k=0}^{n-k_0}, \quad \delta b^{(\ell)} = \{\delta b_k^{(\ell)}\}_{k=0}^m,$$

we determine all numerical approximations at time $t_{\ell+1} = t_\ell + h_\ell$ by applying a step of the Euler method with step-size h_ℓ to (4.20), which is fully described by Algorithm 1.

In order to control the step size we simply require that the monotonicity property of the exact flow, that is $|\lambda(t_{\ell+1})| < |\lambda(t_\ell)|$ is preserved by the numerical solution $|\lambda_{\ell+1}| < |\lambda_\ell|$. Since we are only interested in stationary points we can neglect the classical error control estimate on the solution, that is we do not estimate $\|E(t_{\ell+1}) - E_{\ell+1}\|$.

At line **9** of Algorithm 1 we have introduced a stopping criterion which activates either when the difference of two subsequent iterates goes under a given accuracy tol (this is an indicator of the fact that the sequence λ_ℓ has converged) or when λ_ℓ approaches zero to a certain tolerance tol . If $|\lambda_\ell| > \text{tol}$ it means that it has reached a local minimum for $|\lambda|, \lambda \in \Lambda_\varepsilon^S(S)$; if $|\lambda_\ell| \leq \text{tol}$ it indicates that the global minimum has been reached (to the given tolerance since the exact minimum value is zero). In both cases E_ℓ approximates the corresponding extremizer.

REMARK 4.4. The normalization step at lines **4** and **5** of Algorithm 1 ensures that $\|E_\ell\|_F = 1$ for all ℓ . Since the exact solution of the ODEs (4.16) (as well as (4.20)) preserves the norm of E , after the application of an Euler step the error on $\|E\|$ is $\mathcal{O}(h^2)$, which makes it very small if h is small (which certainly occurs close to convergence). Hence we do not expect a significant slowing down of the algorithm (and we do not observe this in our numerical experiments). As remarked by an anonymous referee, whom we thank, the choice of the stepsize might be improved by suitable line search algorithms. In the present paper we have selected the stepsize according to the preservation of the monotonicity of $|\lambda_\ell|$, in agreement with the monotonicity property of the exact flow. Investigating more efficient step size control strategies would be certainly an interesting topic for a future investigation. \diamond

5. An iterative method for approximating $d_{CD}(p, q)$. In this section, we discuss the outer algorithm and make use of the following notation: all quantities written as $g(\varepsilon)$, like $\lambda(\varepsilon), E(\varepsilon)$ and so on, are intended to be exact and associated to local minima/extremizers for the optimization problem (4.9).

In order to compute the distance to common divisibility we should consider equation $\mu_\varepsilon(S) = 0$ and minimize its solution. As a surrogate of this problem, which is of global optimization, we try to compute a value, say ε^* , such that the boundary of the corresponding structured ε -pseudospectrum, $\partial\Lambda_\varepsilon^S(S)$, crosses the origin.

This would provide an upper bound for the distance; repeating such a search over different regions of the ε -pseudospectrum would increase the probability of computing the exact distance, and hence would improve the robustness of the method.

REMARK 5.1. Let $\lambda(\varepsilon) \in \partial\Lambda_\varepsilon^S(S)$ a branch of points of locally minimum modulus, with $\lambda(\varepsilon) \neq 0$ for $\varepsilon < \varepsilon^*$ and $\lambda(\varepsilon^*) = 0$ (see e.g. Figure 7.1). Consider the function

$$\varepsilon \longrightarrow |\lambda(\varepsilon)|, \quad \varepsilon < \varepsilon^*.$$

Its continuity is a consequence of continuity of eigenvalues; however, in principle, such a function does not need to be differentiable, which would prevent our arguments from applying, since they are based on its differentiability. In fact one might switch from one local minimum to another one located in a different region of the structured pseudospectrum, which would imply generically a jump of its derivative.

Hence we shall further assume some smoothness, more specifically that $\lambda(\varepsilon)$ is a smooth and continuously differentiable curve of the complex plane. If $\lambda(\varepsilon)$ is a continuous curve of simple eigenvalues then its differentiability is implied by the simplicity. Under these assumptions we may think to approach ε^* from the left, by a repeated integration of the ODEs 4.20, and exploit the knowledge of the derivative of $\lambda(\varepsilon)$ w.r.t. ε . \diamond

To summarize, in order to proceed we indicate by

$$\lambda(\varepsilon) = \arg \min_{\lambda \in \Lambda_\varepsilon^S(S)} |\lambda|$$

a smooth branch of (local) minima parametrized by ε and computed by determining the stationary point of the system of ODEs (4.16) (or equivalently (4.20)) which we denote by $E(\varepsilon)$, and make the following generic assumption.

ASSUMPTION 5.1. *Let $\lambda(\varepsilon) \neq 0$ be a point of locally minimum modulus of $\Lambda_\varepsilon^S(S)$ (with ε fixed), that is an eigenvalue with minimum modulus of the matrix $S + \varepsilon E(\varepsilon)$ (where $E(\varepsilon)$ denotes the corresponding (local) extremizer). Then $\lambda(\varepsilon)$ is simple.*

Moreover we assume that $E(\varepsilon)$ and $\lambda(\varepsilon)$ are smooth with respect to ε (at least in a neighbourhood of ε^).* \diamond

Assumption 5.1 states that the eigenvalue $\lambda(\varepsilon)$ of minimum modulus of $S + \varepsilon E(\varepsilon)$ is a smooth function of ε in a left neighbourhood of ε^* .

5.1. A key variational formula. The following result provides us an explicit and easily computable expression for the derivative of $|\lambda(\varepsilon)|$ (and thus also $\mu_\varepsilon(S)$) w.r.t. ε .

THEOREM 5.1. *Assume the following:*

1. $\varepsilon \in (0, \varepsilon^*)$ such that $\lambda(\varepsilon) \neq 0$,
2. $\lambda(\varepsilon)$ be a smooth branch of points of (locally) minimum modulus of $\Lambda_\varepsilon^S(S)$,
3. Assumption 5.1 holds, i.e. $\lambda(\varepsilon)$ and $E(\varepsilon)$ are smooth w.r.t. ε , and let $y(\varepsilon)$ and $x(\varepsilon)$ be corresponding left and right eigenvectors of $S + \varepsilon E(\varepsilon)$ (where $E(\varepsilon)$ is a local extremizer), scaled according to (4.1), with $\|E(\varepsilon)\|_F = 1$ for all $\varepsilon \in (0, \varepsilon^*)$.

Then the following holds, with $s(\varepsilon) = -\|P_S(y(\varepsilon)x(\varepsilon)^)\|_F$,*

$$\frac{d|\lambda(\varepsilon)|}{d\varepsilon} = \frac{s(\varepsilon)}{|y(\varepsilon)^*x(\varepsilon)|} < 0, \quad \text{for all } \varepsilon.$$

Proof. By Theorem 4.8 we have

$$P_S(y(\varepsilon)x(\varepsilon)^*) = s(\varepsilon)E(\varepsilon), \quad \text{with } s(\varepsilon) = -\|P_S(y(\varepsilon)x(\varepsilon)^*)\|_F \quad \text{and} \quad \|E(\varepsilon)\| = 1,$$

the latter implying

$$0 = \langle E(\varepsilon), E'(\varepsilon) \rangle = \langle \text{Re}(y(\varepsilon)x(\varepsilon)^*), P_S(E'(\varepsilon)) \rangle = \text{Re}\langle y(\varepsilon)x(\varepsilon)^*, E'(\varepsilon) \rangle$$

where $' = d/d\varepsilon$. The proof follows by observing that

$$\begin{aligned} \frac{d}{d\varepsilon} |\lambda(\varepsilon)| &= \frac{1}{|y(\varepsilon)^* x(\varepsilon)|} \operatorname{Re} \langle y(\varepsilon) x(\varepsilon)^*, E(\varepsilon) + \varepsilon E'(\varepsilon) \rangle \\ &= \frac{1}{|y(\varepsilon)^* x(\varepsilon)|} \langle P_S(y(\varepsilon) x(\varepsilon)^*), E(\varepsilon) \rangle = \frac{s(\varepsilon)}{|y(\varepsilon)^* x(\varepsilon)|}. \end{aligned}$$

Since $P_S(y(\varepsilon) x(\varepsilon)^*)$ does not vanish by Lemma 4.3, the previous is strictly negative and the proof is complete. \square

Algorithm 2: Basic algorithm for computing ε^*

Data: $\text{tol} > 0$ and $\varepsilon_0, \varepsilon_1, \varepsilon_u$ (such that $|\lambda(\varepsilon_0)| > |\lambda(\varepsilon_1)| > \text{tol}$, and $|\lambda(\varepsilon_u)| < \text{tol}$).

Result: ε_f (approximation of ε^*).

begin

```

1  Set Reject = False and  $k = 1$ .
2  while  $|\varepsilon_k - \varepsilon_u| \geq \text{tol}$  do
3      if Reject = False. then
4          Store  $\varepsilon_k$  and  $\lambda(\varepsilon_k)$  into the memory.
5          Compute the polynomial  $d_k(\varepsilon)$  (see (5.2)).
6          Compute  $\tilde{\varepsilon}_{k+1}$  the real root of  $d_k(\varepsilon)$  closest to  $\varepsilon_k$ .
7          if  $\tilde{\varepsilon}_{k+1} > \varepsilon_u$  then
8              Set  $\tilde{\varepsilon}_{k+1} = (\varepsilon_u + \varepsilon_k)/2$ .
9          else
10             Set  $\tilde{\varepsilon}_{k+1} = (\varepsilon_u + \varepsilon_k)/2$ .
11             Compute  $\lambda(\tilde{\varepsilon}_{k+1})$  by integrating (4.20) (equivalently (4.16)) with
12             initial datum  $E(\varepsilon_k)$  (i.e., the previously computed extremizer).
13             if  $|\lambda(\tilde{\varepsilon}_{k+1})| < \text{tol}$  then
14                 Set Reject = True.
15                 Set  $\varepsilon_u = \tilde{\varepsilon}_{k+1}$ .
16             else
17                 Set Reject = False.
18                 Set  $\varepsilon_{k+1} = \tilde{\varepsilon}_{k+1}$ .
19                 Order the array  $\{\varepsilon_j\}_{j=0}^{k+1}$  in ascending order,  $\varepsilon_{j+1} > \varepsilon_j$ .
20                 Set  $k = k + 1$ .
21         Set  $\varepsilon_f = \varepsilon_k$ .
```

5.2. The outer iteration. As a consequence the function $\varepsilon \rightarrow |\lambda(\varepsilon)|$ is smooth for $\varepsilon < \varepsilon^*$ (where $|\lambda(\varepsilon)| > 0$); applying a Newton's iterate yields, for $\varepsilon_k < \varepsilon^*$:

$$\varepsilon_{k+1} = \varepsilon_k + \left(\frac{\|P_S(y(\varepsilon_k) x(\varepsilon_k)^*)\|_F}{|y(\varepsilon_k)^* x(\varepsilon_k)|} \right)^{-1} |\lambda(\varepsilon_k)| \quad (5.1)$$

where $\lambda(\varepsilon_k)$ is the eigenvalue of smallest modulus of $S + \varepsilon_k E(\varepsilon_k)$, $E(\varepsilon_k)$ being the extremizer computed by the inner method, which integrates numerically the ODE (4.16). Likely the value ε_{k+1} will be closer to ε^* than ε_k but might lie on the right of ε^* , where the function $|\lambda(\varepsilon)|$ is identically zero; hence it needs a correction to

provide a lower bound to ε^* . This would certainly occur when the function $\lambda(\varepsilon)$ is concave for $\varepsilon < \varepsilon^*$.

An alternative, which allows to obtain a sequence of lower bounds which is more rapidly convergent to ε^* , is that of interpolating pairs $(\varepsilon_{k-1}, |\lambda(\varepsilon_{k-1})|)$, $(\varepsilon_k, |\lambda(\varepsilon_k)|)$ for values $\varepsilon_{k-1}, \varepsilon_k < \varepsilon^*$, implying $|\lambda(\varepsilon_{k-1})|, |\lambda(\varepsilon_k)| > \text{tol}$, tol being a suitable tolerance. Setting $d_k(\varepsilon)$ the cubic Hermite polynomial, such that

$$\begin{aligned} d_k(\varepsilon_\ell) &= |\lambda(\varepsilon_\ell)|, & \ell &= k-1, k \\ d'_k(\varepsilon_\ell) &= -\frac{\|P_{\mathcal{S}}(y(\varepsilon_\ell)x(\varepsilon_\ell)^*)\|_F}{|y(\varepsilon_\ell)^*x(\varepsilon_\ell)|}, & \ell &= k-1, k \end{aligned} \quad (5.2)$$

we define $\widehat{\varepsilon}_{k+1}$ as the solution of $d_k(\varepsilon) = 0$.

Then, if $|\lambda(\widehat{\varepsilon}_{k+1})| > \text{tol}$ we set $\varepsilon_{k+1} = \widehat{\varepsilon}_{k+1}$, otherwise a bisection technique defines ε_{k+1} . This prevents from quadratic convergence when several bisection steps are taken by the method.

An algorithm for the approximation of ε^* . Algorithm 2 is devised to approximate ε^* .

It makes use of an upper bound ε_u such that $\lambda(\varepsilon_u) = 0$ and construct a sequence $\{\varepsilon_k\}$ in the region where $|\lambda(\varepsilon_k)|$ is strictly monotonically decreasing, by successively finding zeros of the polynomials $d_k(\varepsilon)$, $k = 1, 2, \dots$

A natural upper bound is $\varepsilon_u = \sqrt{n}\|a - b\|_F$ (where a and b are the vectors of coefficients of p and q); a natural lower bound $\varepsilon_0 = \sigma_{\min}(\text{Syl}(a, b))$, where $\sigma_{\min}(\cdot)$ indicates the smallest singular value, i.e. the unconstrained distance to singularity of the Sylvester matrix $\text{Syl}(a, b)$.

6. Possible extensions. We consider now a few natural extensions of the proposed methodology to deal with a wider class of problems, and related applications, the distance to uncontrollability of a controllable single-input single-output system.

6.1. An extension to constrained systems. Assume that only certain subsets of the coefficients $\{a_i\}$, $\{b_j\}$ of the polynomials p and q are allowed to be perturbed in order to find a close-by uncontrollable pair. Then the method has the same structure and only the projection changes. In fact, if $\{a_i\}$ does not vary for $i \notin \mathcal{I}$ and the same holds for $\{b_j\}$ for $j \notin \mathcal{J}$, where $\mathcal{I} \subseteq \{0, 1, \dots, n\}$ and $\mathcal{J} \subseteq \{0, 1, \dots, m\}$ are the sets of indices corresponding to the coefficients of the polynomials which are allowed to be perturbed, we have simply to consider in (4.16), for $B \in \mathbb{C}^{2n \times 2n}$, the new projection $P_{\mathcal{S}(\mathcal{I}, \mathcal{J})}(B)$ given by (4.3) with

$$\alpha_{n-k} = \begin{cases} \frac{1}{n} \sum_{l=1}^n \text{Re}(B_{l, l+k}) & k \in \mathcal{I} \\ 0 & k \notin \mathcal{I} \end{cases}$$

and

$$\beta_{m-k} = \begin{cases} \frac{1}{n} \sum_{l=1}^n \text{Re}(B_{n+l, n-m+l+k}) & k \in \mathcal{J} \\ 0 & k \notin \mathcal{J}. \end{cases}$$

Note that the proof that the $P_{\mathcal{S}(\mathcal{I}, \mathcal{J})}(yx^*) \neq 0$ is not obtained as a direct extension of Lemma 4.3.

The system of ODEs we have to solve is still (4.20) but now the number of ordinary differential equations is $|\mathcal{I}| + |\mathcal{J}|$.

An illustrative example. We consider the following example of two monic quadratic polynomials considered in [26],

$$p(z) = z^2 - 6z + 5, \quad q(z) = z^2 - 6.3z + 5.72$$

- (i) Applying Algorithm 2, without constraining the perturbed polynomials to be monic we obtain

$$\begin{aligned} \hat{p}(z) &= 0.985005935828721 z^2 - 6.002940644075092 z + 4.999423279273879 \\ \hat{q}(z) &= 1.014952404182629 z^2 - 6.297067526304693 z + 5.720575118346765, \end{aligned}$$

such that $d_{CD}(p, q) = \varepsilon^*/\sqrt{2} \approx 0.021594147$ and $z \approx 5.098904194$ is the common root. These results well agree with those in [26] and also in [40].

- (ii) Applying Algorithm 2, constraining only the first polynomial to be monic, we obtain

$$\begin{aligned} \hat{p}(z) &= z^2 - 6.005752814118045 z + 4.998851154980102 \\ \hat{q}(z) &= 1.028807031186749 z^2 - 6.294247195609338 z + 5.721148843077327, \end{aligned}$$

such that $d_{CD}(p, q) = \varepsilon^*/\sqrt{2} \approx 0.029977897$ and $z \approx 5.00747501054342$ is the common root.

- (iii) Applying Algorithm 2, constraining both polynomials to be monic (i.e. $\mathcal{I} = \{0, 1\}, \mathcal{J} = \{0, 1\}$), we obtain

$$\begin{aligned} \hat{p}(z) &= z^2 - 6.075037558842548 z + 4.985277938401874 \\ \hat{q}(z) &= z^2 - 6.222182433577167 z + 5.735267487695024, \end{aligned}$$

such that $d_{CD}(p, q) = \varepsilon^*/\sqrt{2} \approx 0.11016371$ and $z \approx 5.0969464661670$ is the common root. These results well agree with those in [40].

- (iv) Applying Algorithm 2, constraining the first polynomial to be unperturbed (i.e. $\mathcal{I} = \emptyset, \mathcal{J} = \{0, 1, 2\}$), we obtain

$$\hat{q}(z) = 1.029953916904784 z^2 - 6.294009215915526 z + 5.721198156957997,$$

such that $d_{CD}(p, q) = \varepsilon^*/\sqrt{2} \approx 0.030570610$ and $z = 5$ is the common root.

6.2. The case of complex polynomials. If the polynomials p and q have complex coefficients, a similar gradient system to (4.16) can be derived for the smallest eigenvalue of the Sylvester matrix. The only difference here is that the projection P_S has to be replaced by a projection onto the set of complex valued Sylvester matrices (a result analogous to Lemma 4.2 holds where the coefficients $\{\alpha_{n-k}\}$ and $\{\beta_{m-k}\}$ are given by the same expressions as in Lemma 4.2 but for the real parts). In this case we have generically that the distance to singularity of the matrix $S = \text{Syl}(a, b)$ is attained by a matrix $E(\varepsilon^*)$ which is the projection of a rank-1 matrix yx^* and the corresponding matrix $S + \varepsilon^*E(\varepsilon^*)$ has co-rank equal to 1. The same replacement of P_S has to be considered in Theorem 5.1.

7. Illustrative examples. In this section, we illustrate the performance of the method proposed, although it has not yet been optimized in our implementation, on synthetic examples and compare it with alternative methods:

- the `uvGCD` function of the Numerical Algebraic Computing Toolbox (NAClab) for Matlab, developed by T.-Y. Li and Z. Zeng [47], and available from

<http://homepages.niu.edu/~naclab/>

- the methods of [46], based on the structured low-rank approximation (SLRA) toolbox, developed by K. Usevich and I. Markovsky [34], and available from <https://arxiv.org/abs/1304.6962v1>
<http://slra.github.io/>

The uvGCD method does not allow us to include constraint on the coefficients of p and q (e.g., p monic or q monic). Similarly, the SLRA approach has limitations imposed by its solution method [45, 13]. On the contrary, the ODE method allows to include arbitrary constraints on the coefficients. In the two considered unconstrained examples, all methods give the same results, but the ODE method is slower. An optimization of the implementation and the investigation of different gradient-based optimization techniques is a plan of the authors.

In all examples we made fine sampling of the parameter space (we randomly generated 10^6 matrices E with Sylvester structure and unit norm and computed the spectra of the corresponding matrices $S + \varepsilon E$) in order to accurately approximate the structured ε -pseudospectrum so that we can assert that what we compute is indeed the distance and not simply an upper-bound. Hence the figures illustrate the effective behavior of $|\lambda(\varepsilon)| = \mu_\varepsilon(S)$ as a function of ε and the first intersection to the horizontal axis provides the value ε^* which approximates the distance $d_{CD}(p, q)$.

In the case where $\lambda(\varepsilon)$ is a real eigenvalue for $\varepsilon \rightarrow \varepsilon^*$, we expect generically that $S + \varepsilon^* E$ has a simple zero eigenvalue and hence has rank $2n - 1$. This is illustrated by the following examples.

7.1. Example (common ε -GCD of degree 1). Consider the polynomials (1.1) of degree 5 with coefficients

$$\begin{array}{cccccc} a_5 = 1 & a_4 = 0 & a_3 = 1 & a_2 = 0 & a_1 = 2 & a_0 = 1 \\ b_5 = -2 & b_4 = 1 & b_3 = 1 & b_2 = -1 & b_1 = 0 & b_0 = 1. \end{array}$$

Note that the p polynomial is monic. This property will be preserved in the approximation \hat{p} .

The computed matrix $S + \varepsilon^* E(\varepsilon^*)$ has rank- $2n - 1$ due to a simple zero eigenvalue. The perturbed polynomials' $\hat{p} = p + \delta p$, $\hat{q} = q + \delta q$ coefficients are shown (with five digit accuracy) in Table 7.1.

TABLE 7.1
Coefficients of the perturbed polynomials $\hat{p} = p + \delta p$, $\hat{q} = q + \delta q$ in the example of 7.1.

$$\begin{array}{cccccc} \hat{a}_5 = 1 & \hat{a}_4 = 0.0144 & \hat{a}_3 = 0.9729 & \hat{a}_2 = 0.0510 & \hat{a}_1 = 1.9039 & \hat{a}_0 = 1.1811 \\ \hat{b}_5 = -1.9778 & \hat{b}_4 = 0.9583 & \hat{b}_3 = 1.0787 & \hat{b}_2 = -1.1483 & \hat{b}_1 = 0.2795 & \hat{b}_0 = 0.4732 \end{array}$$

The common zero of \hat{p} , \hat{q} is

$$z_1 = -0.530278660.$$

The value ε^* and the estimated distance to common divisibility are

$$\begin{aligned} \varepsilon^* &= 1.468981057767730 \\ d_{CD}(p, q) &= 0.656948300565638 \end{aligned}$$

The function $\varepsilon \mapsto |\lambda(\varepsilon)|$ is shown in Figure 7.1.

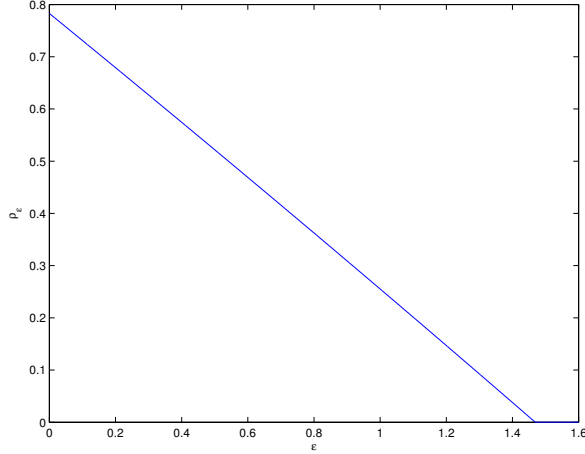


FIG. 7.1. The function $\varepsilon \rightarrow |\lambda(\varepsilon)|$ for Example 7.1

Using two different tolerances we computed two pairs of perturbed polynomials, the first pair with a common real zero $z'_1 = -0.579049166$ and with a distance 0.721916532883628 , the second with common complex conjugate roots $z''_{1,2} = 0.649831247 \pm 0.809734098i$ and with a distance 1.181337128476824 . Both pairs of polynomials, although being computed in an unconstrained setting, have a larger distance from (p, q) than the one computed by the present method. Indeed the non local optimality of the computed pairs can be checked by seeing that the corresponding Sylvester matrix $S + \varepsilon E$ is such that E is not negatively proportional to $P_S(yx^*)$ (see Remark 4.2). However, they both provide close (although not optimal) pairs with a common ε -GCD in a fast way, which is certainly a commendable quality of the method in [47].

The results obtained by the three methods—ODE method, SLRA method, and the uvGCD function—are summarized in Table 7.2.

TABLE 7.2

The numerical results for Example 7.1 show that all compared methods find the same (locally) optimal solution, however, the computation time differs by orders of magnitude. Most efficient is uvGCD, while least efficient is the ODE method. This is partly due to the unoptimized software implementation of the latter.

	SLRA	uvGCD	ODE
distance $d_{CD}(p, q)$	0.6569	0.65702	0.6569
computation time, sec.	0.32856	0.07814	2.1755

7.2. Example (common ε -GCD of degree 2). We consider the two polynomials of degree 3 with coefficients

$$\begin{array}{cccc} a_3 = 1 & a_2 = 2 & a_1 = 2 & a_0 = 2 \\ b_3 = 2 & b_2 = 0 & b_1 = 1 & b_0 = -2 \end{array}$$

First we consider the fully unconstrained case. The perturbed polynomials $\hat{p} = p + \delta p$, $\hat{q} = q + \delta q$ have coefficients

$$\begin{array}{cccc} \hat{a}_3 = 0.7389 & \hat{a}_2 = 2.1037 & \hat{a}_1 = 2.1263 & \hat{a}_0 = 1.8434 \\ \hat{b}_3 = 1.9539 & \hat{b}_2 = -0.05119 & \hat{b}_1 = 1.0663 & \hat{b}_0 = -2.0063. \end{array}$$

They are not anymore co-prime; in fact they both have the common zeros equal to

$$z_{1,2} = -0.4057918541 \pm 1.0300446514i.$$

The value ε^* and the estimated distance to common divisibility are

$$\varepsilon^* = 0.621061904239760, \quad d_{CD}(p, q) = 0.358570257596247.$$

We also applied the code *uvGCD* and computed in one run the pair of polynomials

$$\begin{aligned} \hat{a}'_3 &= 0.8949 & \hat{a}'_2 &= 2.1412 & \hat{a}'_1 &= 2.2491 & \hat{a}'_0 &= 1.7182 \\ \hat{b}'_3 &= 1.9385 & \hat{b}'_2 &= -0.03200 & \hat{b}'_1 &= 1.0339 & \hat{b}'_0 &= -1.9863. \end{aligned}$$

with common complex conjugate roots $z''_{1,2} = -0.4108991889 \pm 1.0263677806i$ and with an estimated distance 0.422550599355414, which is slightly larger than the one computed by the method presented in this paper. The results obtained by the three methods—ODE method, SLRA method, and the *uvGCD* function—are summarized in Table 7.3.

TABLE 7.3

The numerical results for Example 7.2 (unconstrained case) are consistent with the ones reported in Table 7.2.

	SLRA	uvGCD	ODE
distance $d_{CD}(p, q)$	0.35684	0.35684	0.35689
computation time, sec.	0.13727	0.05428	26.642

Next, we consider the case where p is constrained to be monic. The perturbed polynomials $\hat{p} = p + \delta p$, $\hat{q} = q + \delta q$ have coefficients

$$\begin{aligned} \hat{a}_3 &= 1 & \hat{a}_2 &= 2.1680 & \hat{a}_1 &= 2.2569 & \hat{a}_0 &= 1.6991 \\ \hat{b}_3 &= 1.9637 & \hat{b}_2 &= -0.1619 & \hat{b}_1 &= 1.1315 & \hat{b}_0 &= -1.9469. \end{aligned}$$

They are not anymore co-prime; in fact they both have the common zeros equal to

$$z_{1,2} = -0.373421293 \pm 1.0276668040i.$$

The value ε^* and the estimated distance to common divisibility are

$$\varepsilon^* = 0.835047606282059, \quad d_{CD}(p, q) = 0.482114960273099.$$

Figure 7.2 illustrates that the structured ε -pseudospectrum has the origin on its boundary, which implies that the computed value ε^* truly determines the distance to common divisibility $d_{CD}(p, q)$.

7.3. Example (constrained common ε -GCD). Consider again Example 7.1 but now assume that the only coefficients that can be perturbed are a_1, a_3, a_5 and b_1, b_3 and b_5 ; this corresponds to setting $\mathcal{I} = \{1, 3, 5\}$ and $\mathcal{J} = \{1, 3, 5\}$ in the projection $P_{S(\mathcal{I}, \mathcal{J})}$ considered in Section 6.1.

The computed matrix $S + \varepsilon^* E(\varepsilon^*)$ turns out to have rank- $2n - 1$ due to a simple zero eigenvalue. The coefficients of the perturbed polynomials $\hat{p} = p + \delta p$, $\hat{q} = q + \delta q$ are given in Table 7.4.

Their common zero is

$$z = -0.5899110938.$$

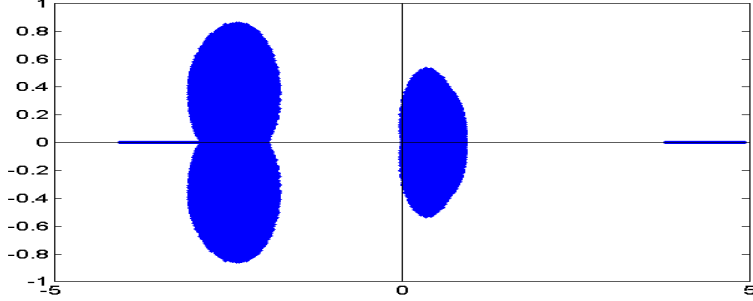


FIG. 7.2. Approximated structured ε -pseudospectrum for $\varepsilon = \varepsilon^*$ for Example 7.2. The origin lies on the boundary of $\Lambda_\varepsilon^S(S)$

TABLE 7.4

Coefficients of the perturbed polynomials $\hat{p} = p + \delta p$, $\hat{q} = q + \delta q$ in the example of 7.3.

$$\begin{array}{cccccc} \hat{a}_5 = 0.9175 & \hat{a}_4 = 0 & \hat{a}_3 = 0.7629 & \hat{a}_2 = 0 & \hat{a}_1 = 1.3186 & \hat{a}_0 = 1 \\ \hat{b}_5 = -1.8715 & \hat{b}_4 = 1 & \hat{b}_3 = 1.3691 & \hat{b}_2 = -1 & \hat{b}_1 = 1.0607 & \hat{b}_0 = 1 \end{array}$$

The value ε^* and the estimated (constrained) distance to common divisibility are

$$\begin{aligned} \varepsilon^* &= 3.004405111510952 \\ d_{CD}(p, q) &= 1.343610812257265. \end{aligned}$$

7.4. Example of small highest degree coefficients. In this subsection, we illustrate the behavior of the method on problems with “very small” highest degree coefficient a_n and $m < n$. In the extreme case of $a_n = 0$, p and q have a common zero at infinity and are therefore not co-prime. In this case the method proposed correctly yields a zero distance to common divisibility. Consider, next, the polynomials (1.1) of degrees $n = 3$ and $m = 2$ with coefficients

$$\begin{array}{cccc} a_3 \in [10^{-16}, 1] & a_2 = 2 & a_1 = 5 & a_0 = 3 \\ & b_2 = 1 & b_1 = 3.1 & b_0 = 2.2 \end{array}$$

Figure 7.3 shows the distance to common divisibility as $a_3 \rightarrow 0$. The method correctly determines that the distance to common divisibility goes to zero.

8. Discussion and outlook. We considered the distance to common divisibility of a pair of polynomials and discussed a new local optimization method for computing $d_{CD}(p, q)$ based on integration of a system of ordinary differential equations, which describes the gradient associated to the cost functional. The overall methodology consists of a two-level iteration, an inner level where we determine extremizers for the associated functional over the set of perturbations of a given norm ϵ by integrating a system of differential equations up to a stationary point, and an outer level where we optimize with respect to epsilon. The method allows specification of exactly known coefficients of the polynomials p and q . The presented numerical examples show the robustness of the method to the initial approximation. The method was applied to the problem of computing the distance of a single-input single-output linear time-invariant system to uncontrollability.

Future work will focus of generalization of the method to multiple polynomials, computing multiple common zeros (approximate common divisor of specified degree),

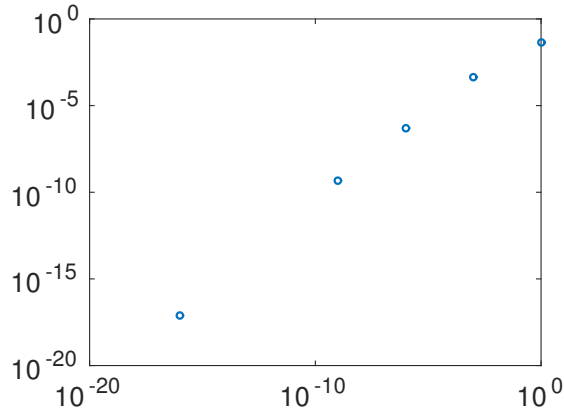


FIG. 7.3. The distance to common divisibility for Example 7.4.

multivariable polynomials, and application of the method in systems and control theory for computing the distance of a given system to the set of uncontrollable systems.

Acknowledgments. The authors thank the anonymous Referees for their valuable comments and suggestions. The first author thanks the Italian M.I.U.R. and the INdAM GNCS for financial support and also the Center of Excellence DEWS. The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement number 258581 “Structured low-rank approximation: Theory, algorithms, and applications”; Fund for Scientific Research (FWO-Vlaanderen), FWO project G028015N “Decoupling multivariate polynomials in nonlinear system identification”; and the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme initiated by the Belgian Science Policy Office.

REFERENCES

- [1] P.-A. ABSIL, R. MAHONY, AND R. SEPULCHRE, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ, 2008.
- [2] V. Balakrishnan, S. Boyd, and S. Balemi. Branch and bound algorithm for computing the minimum stability degree of parameter-dependent linear systems, *Int. J. of Robust and Nonlinear Control*, 1(4):295–317, 1991.
- [3] B. Beckermann and G. Labahn. When are two numerical polynomials relatively prime?, *J. Symbolic Comput.*, 26(6):677–689, 1998.
- [4] D. A. Bini and P. Boito. Structured matrix-based methods for polynomial ϵ -gcd: analysis and comparisons, ISSAC 2007, ACM, New York 9:16, 2007.
- [5] D. A. Bini and P. Boito. A fast algorithm for approximate polynomial GCD based on structured matrix computations, Numerical methods for structured matrices and applications, *Oper. Theory Adv. Appl.*, Birkhäuser Verlag, Basel, 199:155–173, 2010.
- [6] J. Burke, M. L. Overton, and A. Lewis. Pseudospectral components and distance to controllability, *SIAM J. Matrix Anal. Appl.*, 26:350–361, 2004.
- [7] D. MARQUARDT, *An algorithm for least-squares estimation of nonlinear parameters*, SIAM J. Appl. Math., 11 (1963), pp. 431–441.

- [8] P. Buttà, N. Guglielmi, and S. Noschese. Computing the structured pseudospectrum of a Toeplitz matrix and its extreme points, *SIAM J. Matrix Anal. Appl.*, 33:300–319, 2012.
- [9] P. Chin, R. Corless, and G. Corliss, “Optimization strategies for the floating-point GCD,” in *Proc. Int. Symp. Symbolic Algebraic Comp.*, O. Gloor, Ed., 1998, pp. 228–235.
- [10] R. M. Corless, P. M. Gianni, B. M. Trager and S. M. Watt. The singular value decomposition for polynomial systems, *Proc. 1995 Internat. Symp. Symbolic Algebraic Comput. ISSAC95*, ACM Press, New York, NY 96–103, 1995.
- [11] I.Z. Emiris, A. Galligo and H. Lombardi. Certified approximate univariate GCDs, *J. Pure Appl. Algebra*, 117/118:229–251, 1997.
- [12] M. FAZEL, *Matrix rank minimization with applications*, PhD thesis, Elec. Eng. Dept., Stanford University, 2002.
- [13] G. GOLUB AND V. PEREYRA, *Separable nonlinear least squares: the variable projection method and its applications*, Institute of Physics, Inverse Problems, 19 (2003), pp. 1–26.
- [14] N. Guglielmi, M. Gürbüzbalaban and M. Overton. Fast approximation of the H_∞ norm via optimization over spectral value sets, *SIAM J. Matrix Anal. Appl.*, 34:709–737, 2013.
- [15] N. Guglielmi and C. Lubich. Differential equations for roaming pseudospectra: paths to extremal points and boundary tracking, *SIAM J. Numer. Anal.*, 49:1194–1209, 2011.
- [16] N. Guglielmi and C. Lubich. Low-rank dynamics for computing extremal points of real pseudospectra, *SIAM J. Matrix Anal. Appl.*, 34:40–66, 2013.
- [17] N. Guglielmi and M. Manetta. Approximating real stability radii, *IMA J. Numer. Anal.* 35(3):1402–1425, 2015.
- [18] N. Guglielmi and M. Overton. Fast algorithms for the approximation of the pseudospectral abscissa and pseudospectral radius of a matrix, *SIAM J. Matrix Anal. Appl.*, 32:1166–1192, 2011.
- [19] N. Guglielmi, M. Overton, and G. Stewart. An efficient algorithm for computing the generalized null space decomposition, *SIAM J. Matrix Anal. Appl.*, 36(1):38–54, 2015.
- [20] N. Higham. Matrix nearness problems and applications, in *Applications of Matrix Theory*, M. Gover and S. Barnett, Eds. Oxford University Press, 1–27, 1989.
- [21] G. Hu and E. Davison. Real controllability/stabilizability radius of LTI systems, *IEEE Trans. Automat. Control*, 49:254–258, 2004.
- [22] M. Ishteva, K. Usevich, and I. Markovsky. Factorization approach to structured low-rank approximation with applications, *SIAM J. Matrix Anal. Appl.*, 35(3):118–1204, 2014.
- [23] T. Kailath. *Linear Systems*. Prentice Hall, 1981.
- [24] E. Kaltofen, R. M. Corless and D. J. Jeffrey. Challenges of symbolic computation: my favorite open problems, *J. Symbolic Comput.* 29(6):891–919, 2000.
- [25] E. Kaltofen, Z. Yang, L. Zhi. Approximate greatest common divisors of several polynomials with linearly constrained coefficients and singular polynomials, *Proc. International Symposium on Symbolic and Algebraic Computations*, 2006.
- [26] N. K. Karmarkar and Y. N. Lakshman. On Approximate GCDs of univariate polynomials, *J. Symbolic Comp.* 26:653–666, 1998.
- [27] T. Kato. *Perturbation theory for linear operators*. Springer-Verlag, 1995.
- [28] M. Manetta. Computation of pseudospectral matrix distances and stability measures, *PhD thesis*, University of L’Aquila, 2014.
- [29] C. LAWSON AND R. HANSON, *Solving Least Squares Problems*, Classics in Applied Mathematics, Society for Industrial and Applied Mathematics, 1987.
- [30] I. Markovsky. Structured low-rank approximation and its applications, *Automatica*, 44(4):891–909, 2008.
- [31] I. Markovsky. A software package for system identification in the behavioral setting, *Control Engineering Practice*, 21:1422–1436, 2013.
- [32] I. Markovsky, *Low Rank Approximation: Algorithms, Implementation, Applications*. Springer, 2012.
- [33] I. Markovsky and K. Usevich. Structured low-rank approximation with missing data, *SIAM J. Matrix Anal. Appl.*, 34(2):814–830, 2013.
- [34] I. MARKOVSKY AND K. USEVICH, *Software for weighted structured low-rank approximation*, *J. Comput. Appl. Math.*, 256 (2014), pp. 278–292.
- [35] J. Nocedal and S. Wright. *Numerical optimization*. Springer-Verlag, 1999.
- [36] S. Noschese and L. Pasquini. Eigenvalue patterned condition numbers: Toeplitz and Hankel cases, *J. Comput. Appl. Math.*, 206:615–624, 2007.
- [37] D. Rupprecht. An algorithm for computing certified approximate GCD of n univariate polynomials, *J. Pure Appl. Algebra*, 139(1-3):255–284, 1999.
- [38] T. W. Sederberg and G. Chang. Best linear common divisors for approximate degree reduction, *Computer-Aided Design*, 25(3): 163-168, 1993.

- [39] J. J. Sylvester. On a Theory of the Syzygetic relations of two rational integral functions, comprising an application to the theory of Sturm's functions, and that of the greatest algebraical common measure, *Philosophical Trans.* 143:407-548, 1853.
- [40] A. Terui. GPGCD: an iterative method for calculating approximate GCD of univariate polynomials, *Theoret. Comput. Sci.*, 479:127–149, 2013.
- [41] L. N. Trefethen and M. Embree. *Spectra and pseudospectra: The behavior of nonnormal matrices and operators.* Princeton University Press, 1999.
- [42] J. von Neumann. *Functional Operators. II. The Geometry of Orthogonal Spaces.* Annals of Mathematics Studies, no. 22. Princeton University Press, 1950.
- [43] A. I. G. Vardulakis and P. N. R. Stoye. Generalized resultant theorem, *J. Inst. Math. Appl.*, 22(3):331–335, 1978.
- [44] K. USEVICH AND I. MARKOVSKY, *Optimization on a Grassmann manifold with application to system identification*, *Automatica*, 50 (2014), pp. 1656–1662.
- [45] K. USEVICH AND I. MARKOVSKY, *Variable projection for affinely structured low-rank approximation in weighted 2-norms*, *J. Comput. Appl. Math.*, 272 (2014), pp. 430–448.
- [46] ———, “Variable projection methods for approximate (greatest) common divisor computations,” *Theoretical Computer Science*, 2016.
- [47] Z. Zeng. The approximate GCD of inexact polynomials. Part I: A univariate algorithm, *Manuscript*, 2004.
- [48] Z. Zeng. The numerical greatest common divisor of univariate polynomials, in Randomization, Relaxation, and Complexity in Polynomial Equation Solving (L. Gurvits, P. Pbay, J. M. Rojas, D. Thompson Eds.) *Contemporary Mathematics*, AMC, 556:187–217, 2011.
- [49] G. Zoutendijk. *Mathematical programming methods.* North-Holland Publishing Co., 1976.