Nearest common root of polynomials, approximate Greatest Common Divisor and the structured singular value

G. Halikias*, G. Galanis*, N. Karcanias* and E. Milonidis*

Abstract

In this paper the following problem is considered: Given two co-prime polynomials, find the smallest perturbation in the magnitude of their coefficients, such that the perturbed polynomials have a common root. It is shown that the problem is equivalent to the calculation of the structured singular value of a matrix arising in robust control and a numerical solution to the problem is developed. A simple numerical example illustrates the effectiveness of the method for two polynomials of low degree. Finally, problems involving the calculation of the approximate greatest common divisor (GCD) of univariate polynomials are considered, by proposing a generalization of the definition of the structured singular value involving additional rank constraints.

Keywords: Approximate common root of polynomials, approximate GCD, Sylvester resultant matrix, structured singular value, distance to singularity, structured approximations.

1. Introduction

The computation of the greatest common divisor (GCD) of two polynomials $a(s)$ and $b(s)$ is a non-generic problem, in the sense that a generic pair of polynomials is co-prime, i.e. has greatest common divisor equal to one. Thus, the notion of *approximate* common factors and approximate GCD's has to be introduced for the purpose of effective numerical calculations [1], [2], [3], [7], [10], [28], [30].

In the context of Systems and Control applications, the main motivation for developing non-generic techniques for calculating the nearest common root and the approximate GCD of polynomials arises from the study of almost zeros [18]. The numerical computation of the GCD of two (or more) polynomials in a robust way has been considered by many researchers using a variety of methodologies, such as ERES [27], extended ERES [27], matrix pencil methods [21], optimization techniques based on total least squares or alternating projection algorithms [9], Pade approximations [30] and "rank-revealing" matrix factorizations [10]. The common characteristic of all these techniques is that they transform exact procedures to their numerical versions. However, most of the above methods suffer from two fundamental limitations: (i) infinitesimally-small perturbations are typically assumed, and (ii) the structure of the Sylvester resultant matrix on which the solution is based is ignored. A numerical example given later in this work shows that these assumptions may lead to gross under-estimation of the distance to the nearest common root and the corresponding approximate GCD polynomial.

In this paper, we formalize the notion of "approximate co-primeness" of two polynomials, by considering the minimum-norm perturbation in the polynomial's coefficient vectors such that the

[1]Control Engineering Research Centre, School of Engineering and Mathematical Sciences, City University, Northampton Square, London EC1V 0HB, UK

perturbed polynomials have a common root. The calculation of the minimal perturbation is shown to correspond to the distance of a structured matrix from singularity, or, equivalently to the calculation of the structured singular value (μ) of a matrix under repeated scalar perturbations. An alternative notion of approximation based on the Frobenius norm and the "spherical structured singular value" of a matrix [8], [22], [31], [33], [34], is also briefly discussed. In contrast to alternative methods, our approach gives a precise meaning to the notion of approximate common roots of polynomials and the related notion of the "strength of approximation" of the approximate GCD of polynomials.

Problems involving the structured singular value of a matrix have been extensively studied in the context of robust-stability and robust-performance design of control systems, and efficient methods have been developed for its numerical solution. Although the problem is computationally NP hard [5], [16], tight bounds have been reported in the robust control literature using exclusively convex programming techniques. In a recent work [17] tests have been developed for certifying the absence of a duality gap between the structured singular value and its convex upper bound, along with a systematic procedure for reducing the duality gap when this is non-zero.

The approximate GCD problem has numerous applications in Numerical Mathematics and Engineering and various techniques and algorithms have been proposed for its solution. Recent work in this area includes references [4], [23], [24] and [25]. The general approach followed in these papers formulates the approximate GCD problem as a structured least-squares approximation which can be solved numerically (e.g. via a Newton-like iterative procedure). The structure of the approximation problem is captured by the "displacement structure" of the Sylvester or Bezout matrix corresponding to the polynomials whose (approximate) GCD is sought. There is no guarantee that this procedure results in the globally optimum solution and the overall algorithm relies on a number of heuristics and "almost" rank-revealing matrix factorizations [4]. The resulting algorithms presented have typically low complexity and, from the several numerical experiments performed so far, they appear to be robust and numerically stable [4], although "small" perturbations in the polynomials' coefficients are only considered. Quadratic (least-squares type) norms can only be used to quantify perturbation size in the polynomials' coefficients.

The method proposed in this work is motivated by structured perturbation problems in the area of robust control. Uncertainty in the polynomials' coefficients is captured via block-diagonal matrix perturbations consisting of (repeated) uncertain scalar parameters. The approach is not limited to"small" perturbations in the polynomials' coefficients and general vector norms can be used in principle to represent uncertainty (in the paper quadratic and l_∞ norms are considered). The resulting optimization is non-convex and computationally demanding for large problems, although tight convex bounds (and techniques for improving them) have been reported in the literature [16], [17], [14], [20], [22], [29], [33], [34], [35]. Algorithms combining the technique presented in this work with aspects of [24] and [25] also seem possible and will be reported in a future publication.

The layout of the paper is as follows: Section 2 surveys the theory of structured approximations arising in robust control theory, by reviewing the definitions and basic results related to the structured singular value (μ). To avoid unnecessary complications, the main aspects of the theory are presented in a simplified form, relevant to the problem addressed here, i.e. only repeated scalar uncertainties are considered in the corresponding uncertainty structure used in the definition of the structured

singular value. The main results of the paper are included in sections 3 and 4. Section 3 formulates the polynomial nearest common-root problem in an equivalent framework involving the distance of a matrix to singularity, subject to structured perturbations in its elements; this is shown to be equivalent to the calculation of a structured singular value (μ) of a matrix constructed from the coefficients of the two polynomials. Extensions to the case of more than two polynomials is possible via the theory of generalized resultants [3], [15], [19], [27]. Section 4 generalizes the "nearest common-root" problem to the computation of the approximate GCD, using two methods: The first, involves the sequential extraction of all roots of the approximate GCD, combined with polynomial division at each iteration, which terminates when an approximate numerical tolerance specification is satisfied. The second approach, which is theoretically more satisfying, proposes the extraction of the approximate GCD in a single step, by solving an optimization problem which corresponds to the minimization of a structured singular value of a matrix under additional rank constraints. This is reducible to a standard structured distance-to-singularity problem using the displacement structure [4], [24], [25] of the Sylvester resultant matrix associated with the problem. The main conclusions of the paper, along with suggestions for further work, appear in section 5.

The notation used in the paper is standard and is summarized here for convenience. If n is a positive integer, then $\underline{n} = \{1, 2, \ldots, n\}$. $\mathcal{R}$ and $\mathcal{C}$ denote the fields of real and complex numbers, respectively. $\mathcal{R}^n$ and $\mathcal{C}^n$ are the n-dimensional vector spaces over $\mathcal{R}$ and $\mathcal{C}$, respectively, while $\mathcal{R}^{m\times n}$ and $\mathcal{C}^{m\times n}$ the spaces of $m \times n$ matrices over $\mathcal{R}$ and $\mathcal{C}$, respectively. If A is a $m \times n$ matrix (either real or complex) $\mathcal{R}(A)$ denotes the range (column span) of A and $\mathcal{N}(A)$ denotes the (right) null-space (kernel) of A. It is well known that $\mathcal{R}(A)$ is a subspace of $\mathcal{R}^m$ ($\mathcal{C}^m$) and that $\mathcal{N}(A)$ is a subspace of $\mathcal{R}^n$ ($\mathcal{C}^n$). Further, $\text{rank}(A) := \dim(\mathcal{R}(A))$, $\text{null}(A) := \dim(\mathcal{N}(A))$ and from the rank-nullity theorem $\text{rank}(A) + \text{null}(A) = n$. If A is a square matrix, $\lambda(A)$ denotes its spectrum (set of eigenvalues) and $\rho(A) = \max\{|z| : z \in \lambda(A)\}$ is the spectral radius of A. For $A \in \mathcal{C}^{m\times n}$, A' denotes the transpose of A and A^* the complex conjugate transpose of A. A matrix $A \in \mathcal{C}^{n\times n}$ is Hermitian if $A = A^*$. A Hermitian matrix A is said to be positive definite ($A > 0$) if $x^*Ax > 0$ for all $x \neq 0$. In this case A has a spectral decomposition $A = U\Lambda U^*$ with U unitary and $\Lambda = \text{diag}(\Lambda) > 0$. Writing Λ in full as $\Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, we define $\Lambda^{1/2} := \text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n})$ and $A^{1/2} = U\Lambda^{1/2}U^*$, so that $A = A^{1/2}A^{1/2}$ with $A^{1/2}$ Hermitian and positive-definite. $A^{1/2}$ is referred to as the Hermitian square-root of A; we also denote $A^{-1/2} = (A^{1/2})^{-1}$. For $A \in \mathcal{C}^{m\times n}$, the singular values of A are denoted as $\sigma_i(A)$, $i = 1, 2, \ldots, \min(m, n)$, ordered in non-increasing order of magnitude, i.e., $\sigma_1(A) \geq \sigma_2(A) \geq \ldots \geq \sigma_{\min(m,n)}(A) \geq 0$. The spectral norm of a matrix (induced Euclidian norm) is its largest singular value and is also denoted as $\|A\| := \sigma_1(A)$. We also make use of the Frobenius norm of a matrix, denoted as $\|A\|_F$. If $A \in \mathcal{C}^{n\times n}$ and $\mathbf{U}$ is a nonempty set of $n \times n$ matrices, then $A\mathbf{U} := \{AU : U \in \mathbf{U}\}$. $\mathbf{1}_{n\times n}$ denotes the $n\times n$ matrix with all n^2 elements equal to 1. For two matrices A and B, $A \circ B$ denotes the Hadamard (element-wise) product and $A \otimes B$ the Kronecker product of the two matrices. If $A \in \mathcal{C}^{m\times n}$, $\text{vec}(A)$ denotes the mn-dimensional column vector obtained by "stacking" the columns of A one after the other. Let $a(s) = a_n s^n + a_{n-1}s^{n-1} + \ldots + a_0$ be a polynomial in s with either complex or real coefficients a_i. The degree of $a(s)$ is $\partial a(s) = n$, provided $a_n \neq 0$. By convention, if $a(s) \equiv 0$, $\partial a(s) = -\infty$.

2. Problem definition and the structured singular value

The main problem considered in this paper is the following:

Problem 2.1*: Let $a_0(s)$ and $b_0(s)$ be two monic and co-prime polynomials with complex coefficients, such that $\partial a_0(s) = m$ and $\partial b_0(s) = n$. What is the complex perturbation of minimum magnitude in the coefficients of $a_0(s)$ and $b_0(s)$ so that the perturbed polynomials have a common root?* □

Formally write:

$$a_0(s) = s^m + \alpha_{m-1}s^{m-1} + \alpha_{m-2}s^{m-2} + \ldots + \alpha_0 \tag{1}$$

and

$$b_0(s) = s^n + \beta_{n-1}s^{n-1} + \beta_{n-2}s^{n-2} + \ldots + \beta_0 \tag{2}$$

and assume that the coefficients $\{\alpha_i,\ i = 0, 1, \ldots, m-1\}$ and $\{\beta_i,\ i = 0, 1, \ldots, n-1\}$ are subjected to complex perturbations $\delta_0, \delta_1, \ldots, \delta_{m-1}$ and $\epsilon_0, \epsilon_1, \ldots, \epsilon_{n-1}$, respectively, i.e. the perturbed polynomials are:

$$a(s; \delta_0, \ldots, \delta_{m-1}) = s^m + (\alpha_{m-1} + \delta_{m-1})s^{m-1} + (\alpha_{m-2} + \delta_{m-2})s^{m-2} + \ldots + (\alpha_0 + \delta_0) \tag{3}$$

and

$$b(s; \epsilon_0, \ldots, \epsilon_{n-1}) = s^n + (\beta_{n-1} + \epsilon_{n-1})s^{n-1} + (\beta_{n-2} + \epsilon_{n-1})s^{n-1} + \ldots + (\beta_0 + \epsilon_0) \tag{4}$$

respectively. Let also:

$$\gamma = \max\{\ |\delta_0|,\ |\delta_1|, \ldots, |\delta_{m-1}|,\ |\epsilon_0|,\ |\epsilon_1|, \ldots, |\epsilon_{n-1}|\ \} \tag{5}$$

Then Problem 2.1 is equivalent to: $\min \gamma$ such that the two polynomials $a(s; \delta_0, \ldots, \delta_{m-1})$ and $b(s; \epsilon_0, \ldots, \epsilon_{n-1})$ have a common root.

It is shown in section 3 that Problem 2.1 is equivalent to the calculation of the structured singular value (μ) of an appropriate matrix. This is a problem arising in robust control which has been analyzed extensively over recent years [14], [17], [29], [33], [34]. A generalization of Problem 2.1 involving the computation of the numerical GCD of two polynomials is also introduced at a later section and is shown to correspond to the calculation of a generalized structured singular value subject to additional rank constraints (see section 4).

The remaining part of this section defines the structured singular value (μ) for a special class of uncertainty structures, relevant to the polynomial approximate common-root problem and reviews the relevant theory, first developed for the solution of stabilization problems in the area of robust control [14], [29], [33], [34]. We start by introducing the definition of the complex structured singular value for repeated scalar uncertainties. This specializes the more general definition given in [29], which also takes into account matrix (block) uncertainties. These are ignored in our exposition, as they are not relevant for the problems addressed in this paper.

The structured singular value of a matrix M is defined as follows:

Problem 2.2*: (structured singular value). Let $M \in \mathcal{C}^{n\times n}$ and define the "structured" set:*

$$\mathbf{\Delta} = \{\mathrm{diag}(\delta_1 I_{r_1},\ \delta_2 I_{r_2}, \ldots, \delta_s I_{r_s})\ :\ \delta_i \in \mathcal{C},\ i = 1, 2, \ldots, s\} \tag{6}$$

where the r_i are positive integers such that $\sum_{i=1}^{s} r_i = n$. The structured singular value of M (relative to $\mathbf{\Delta}$) is defined as:

$$\mu_{\mathbf{\Delta}}(M) = \frac{1}{\min\{\|\Delta\| \ : \ \Delta \in \mathbf{\Delta}, \ \det(I_n - M\Delta) = 0\}} \tag{7}$$

unless no $\Delta \in \mathbf{\Delta}$ makes $I_n - M\Delta$ singular, in which case $\mu_{\mathbf{\Delta}}(M) = 0$. Here $\|\cdot\|$ denotes the spectral norm of a matrix (largest singular value). The problem is to calculate $\mu_{\mathbf{\Delta}}(M)$ and, provided that $\mu_{\mathbf{\Delta}}(M) \neq 0$, to find a $\Delta \in \mathbf{\Delta}$ of minimal norm such that $\det(I_n - M\Delta) = 0$. □

Remark 2.1***:*** A more general definition of the structured singular value is given in [29] and includes both scalar and matrix uncertainty blocks in a block-diagonal structure. These are not considered here as matrix perturbations are not relevant for GCD problems. The definition is further generalized in [14] to include non-square uncertainty blocks and general induced norms (e.g., p-norms where $1 \leq p \leq \infty$). Although this more general formulation could have been followed here, general p-norms are difficult to compute [12] and so we have restricted the definition and all subsequent results to spectral (induced Euclidian) norms only. The use of the Frobenius norm and the real uncertainty case are briefly discussed at a later section of the paper. □

Remark 2.2***:*** Equivalently to the definition in Problem 2.2, the structured singular value of $M \in \mathcal{C}^{n\times n}$ can also be defined as:

$$\mu_{\mathbf{\Delta}}(M) = \left[\min \|\Delta\| : \Delta \in \mathbf{\Delta}, \det \begin{pmatrix} I_n & \Delta \\ M & I_n \end{pmatrix} = 0\right]^{-1}$$

Note also that it is possible to define the structured singular value $\mu_{\mathbf{\Delta}}(M)$ with a more general uncertainty structure $\mathbf{\Delta}$ compared to the one used in the definition of Problem 2.2. In this case, the minimum in the definition of $\mu_{\mathbf{\Delta}}(M)$ must be replaced by an infimum. This is because it is not possible, in general, to guarantee that the constraint set $\{\Delta \in \mathbf{\Delta} : \det(I_n - M\Delta) = 0\}$ is closed and non-empty, so that it contains at least one $\Delta \in \mathbf{\Delta}$ of minimum norm. This is in contrast to the problems considered in this work for which $\mathbf{\Delta}$ is a subspace of $\mathcal{C}^{n\times n}$ (and hence closed), and $\Delta_0 \in \mathbf{\Delta}$ of sufficiently large norm always exists such that $\det(I_n - M\Delta_0) = 0$, so that $\mu_{\mathbf{\Delta}}(M) > 0$. □

Remark 2.3***:*** Note that in the definition of the structured singular value $\mu_{\mathbf{\Delta}}(M)$ both M and $\mathbf{\Delta}$ are assumed to be complex-valued. It can be shown [29] that if M is actually real, $\mu_{\mathbf{\Delta}}(M)$ is not affected if Δ is further restricted to be real (in addition to the constraint $\Delta \in \mathbf{\Delta}$). For a discussion of the "real structured singular-value" problem ($M \in \mathcal{C}^{n\times n}$ and $\Delta \in \mathbf{\Delta} \cap \mathcal{R}^{n\times n}$) see [14]. □

Upper and lower bounds of the structured singular value can be established via the following Lemma:

Lemma 2.1 ***[14], [29]:*** *Let $\mathbf{\Delta}$ denote the subspace of $\mathcal{C}^{n\times n}$ defined in equation (6). Then for every M in $\mathcal{C}^{n\times n}$,*

$$\rho(M) \leq \mu_{\mathbf{\Delta}}(M) \leq \|M\| \tag{8}$$

where $\rho(\cdot)$ denotes the spectral radius and $\|\cdot\|$ the spectral norm of a matrix.

Introduce the structured ball of unit radius $\mathbf{B}_{\mathbf{\Delta}} = \{\Delta \in \mathbf{\Delta} : \|\Delta\| \leq 1\}$ and the set

$$\mathbf{Q} = \{Q \in \mathcal{C}^{n\times n} : QQ^* = I_n\} \cap \mathbf{\Delta} \tag{9}$$

i.e., the group of all unitary matrices with the structure inherited from the set $\mathbf{\Delta}$ defined in equation (6). Define also the complementary-structure set:

$$\mathbf{D} = \{\text{diag}[D_1, D_2, \ldots, D_s] : D_i \in \mathcal{C}^{r_i \times r_i}, D_i = D_i^* > 0, i = 1, 2, \ldots, s\} \tag{10}$$

Note that for every $\Delta \in \mathbf{\Delta}$, $D \in \mathbf{D}$ and $Q \in \mathbf{Q}$ we have: (i) $Q^* \in \mathbf{Q}$, (ii) $Q\Delta \in \mathbf{\Delta}$, (iii) $\Delta Q \in \mathbf{\Delta}$, (iv) $\|Q\Delta\| = \|\Delta Q\| = \|\Delta\|$ and (v) $D^{1/2}\Delta = \Delta D^{1/2}$. We also have the following alternative characterization of $\mu_{\mathbf{\Delta}}(M)$:

Lemma 2.2 ***[14]:*** *For all $M \in \mathcal{C}^{n\times n}$ and $\mathbf{\Delta}$ defined in equation (6),*

$$\mu_{\mathbf{\Delta}}(M) = \max_{\Delta \in \mathbf{B}_{\mathbf{\Delta}}} \rho(\Delta M) = \max_{\Delta \in \mathbf{\Delta},\, \Delta \neq 0} \frac{\rho(\Delta M)}{\|\Delta\|} \tag{11}$$

If $\mu_{\mathbf{\Delta}}(M) > 0$ we have:

$$\mu_{\mathbf{\Delta}}(M)^{-1} = \min_{\Delta \in \mathbf{\Delta},\, \rho(\Delta M) \geq 1} \|\Delta\| = \min_{\Delta \in \mathbf{\Delta},\, \rho(\Delta M) = 1} \|\Delta\| \tag{12}$$

The characterization of $\mu_{\mathbf{\Delta}}(M)$ given in equation (11) as $\mu_{\mathbf{\Delta}}(M) = \max_{\Delta \in \mathbf{B}_{\mathbf{\Delta}}} \rho(\Delta M)$ can be expressed as a maximum over a much smaller set (compared to $\mathbf{B}_{\mathbf{\Delta}}$).

Theorem 2.1 ***[29]:*** *For any $M \in \mathcal{C}^{n\times n}$ and $\mathbf{\Delta}$ defined in equation (6), we have*

$$\max_{Q \in \mathbf{Q}} \rho(MQ) = \max_{\Delta \in \mathbf{B}\mathbf{\Delta}} \rho(M\Delta) = \mu_{\mathbf{\Delta}}(M)$$

where $\mathbf{Q}$ is defined in equation (9).

The results presented in this section are summarized by the following Theorem:

Theorem 2.2***:*** *For all $Q \in \mathbf{Q}$, $D \in \mathbf{D}$:*

$$\max_{Q \in \mathbf{Q}} \rho(MQ) = \max_{\Delta \in \mathbf{B}_{\mathbf{\Delta}}} \rho(M\Delta) = \mu_{\mathbf{\Delta}}(M) \leq \inf_{D \in \mathbf{D}} \|D^{1/2} M D^{-1/2}\| := \overline{\mu}_{\mathbf{\Delta}}(M)$$

Remark 2.4***:*** The equality $\mu_{\mathbf{\Delta}}(M) = \max\{\rho(QM) : Q \in \mathbf{Q}\}$ suggests a possible optimization algorithm for calculating the structured singular value. Unfortunately, the function $\rho(QM)$ typically has many local maxima which are not global [29] and hence its maximization can only produce a lower bound of $\mu_{\mathbf{\Delta}}(M)$. The right-hand side inequality is in fact more interesting as a means of estimating $\mu_{\mathbf{\Delta}}(M)$. The following result shows that the function $D \to \|D^{1/2} M D^{-1/2}\|$ has convex level sets:

Proposition 2.1 ***[29]:*** *Let $M \in \mathcal{C}^{n\times n}$ be given, along with a scaling set $\mathbf{D}$ and $\beta > 0$. Then the sub-level set*

$$\mathcal{S}_\beta(D) = \{D \in \mathbf{D} : \|D^{1/2} M D^{-1/2}\| < \beta\}$$

is convex.

Remark 2.5***:*** Proposition 2.1 above suggests that calculating the infimizing scaling matrix D can be performed via an iterative procedure ("D-iteration") or any convex algorithm that can accommodate

Linear Matrix Inequality (LMI) constraints. Extensive computational experience suggests that the bound is often adequate for applications in robust control [29]. This is important since the computation of μ is an NP-hard problem [5]. If $\overline{\mu}_{\mathbf{\Delta}}(M)$ denotes the convex upper bound of $\mu_{\mathbf{\Delta}}(M)$ and $M \in \mathcal{C}^{n\times n}$, it has been shown that the ratio $g_n := \overline{\mu}_{\mathbf{\Delta}}(M)/\mu_{\mathbf{\Delta}}(M)$ grows no faster than a linear function of n. In [17] a systematic method is proposed for reducing the duality gap $\overline{\mu}_{\mathbf{\Delta}}(M) - \mu_{\mathbf{\Delta}}(M)$, by solving (at least approximately) a structured singular value problem of reduced rank compared to the original problem.

3. Minimum l_∞ distance to common root of polynomials

The main theory of structured approximation problems and the structured singular value has been outlined in section 2. In this section we establish the equivalence between Problem 2.1 and Problem 2.2. First we need the following result which relates the existence of common factors of two polynomials to the rank of their corresponding Sylvester resultant matrix:

Theorem 3.1***:*** *Consider the monic polynomials $a(s)$ and $b(s)$ with $\partial a(s) = m$ and $\partial b(s) = n$ and let $R_{m,n}(a,b)$ be their Sylvester resultant matrix,*

$$R_{m,n}(a,b) = \begin{pmatrix} 1 & \alpha_{m-1} & \alpha_{m-2} & \cdots & \alpha_0 & 0 & \cdots & 0 \\ 0 & 1 & \alpha_{m-1} & \cdots & \alpha_1 & \alpha_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \alpha_{m-1} & \cdots & \alpha_1 & \alpha_0 \\ 1 & \beta_{n-1} & \beta_{n-2} & \cdots & \beta_0 & 0 & \cdots & 0 \\ 0 & 1 & \beta_{n-1} & \cdots & \beta_1 & \beta_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \beta_{n-1} & \cdots & \beta_1 & \beta_0 \end{pmatrix} \tag{13}$$

Then, if $\phi(s)$ denotes the GCD of $a(s)$ and $b(s)$ the following properties hold true:

1. $\mathrm{rank}(R_{m,n}(a,b)) = n + m - \partial\phi(s)$.
2. *The polynomials $a(s)$ and $b(s)$ are co-prime if and only if* $\mathrm{rank}(R_{m,n}(a,b)) = n + m$.
3. *The GCD $\phi(s)$ is invariant under elementary row operations on $R_{m,n}(a,b)$. Furthermore, if we reduce $R_{m,n}(a,b)$ to its row echelon form, the last non-vanishing row defines the coefficients of $\phi(s)$.*

Proof: See [1, 27, 15]. □

Using Theorem 3.1, the equivalence of Problem 2.1 and Problem 2.2 can now be established:

Theorem 3.2***:*** *Let $a_0(s)$ and $b_0(s)$ be two co-prime polynomials defined in equations (1) and (2), respectively. Then, Problem* 2.1 *is equivalent to Problem* 2.2 *by defining:*

1. *The structured set $\mathbf{\Delta}$ as:*

$$\mathbf{\Delta} = \{\mathrm{diag}(\delta_{m-1}I_n,\ \delta_{m-2}I_n, \ldots, \delta_0 I_n, \epsilon_{n-1}I_m,\ \epsilon_{m-2}I_m, \ldots, \epsilon_0 I_m) : \delta_i \in \mathcal{C}, \epsilon_i \in \mathcal{C}\} \tag{14}$$

i.e., $s = m + n$, $r_i = n$ for $1 \le i \le m$ and $r_i = m$ for $m + 1 \le i \le m + n$.

2. $M = -ZR_{m,n}^{-1}(a_0, b_0)\Theta$ *where:*

$$\Theta = \left(\begin{array}{cccc|cccc} I_n & I_n & \cdots & I_n & O_{n,m} & O_{n,m} & \cdots & O_{n,m} \\ \hline O_{m,n} & O_{m,n} & \cdots & O_{m,n} & I_m & I_m & \cdots & I_m \end{array}\right) \in \mathcal{R}^{n+m,2nm} \tag{15}$$

$$Z' = \left(\begin{array}{cccc|cccc} (Z_{nm}^0)' & (Z_{nm}^1)' & \cdots & (Z_{nm}^{m-1})' & (Z_{mn}^0)' & (Z_{mn}^1)' & \cdots & (Z_{mn}^{n-1})' \end{array}\right) \in \mathcal{R}^{n+m,2nm} \tag{16}$$

in which:

$$Z_{nm}^k = \begin{pmatrix} O_{n,k+1} & I_n & O_{n,m-k-1} \end{pmatrix} \quad \textit{for } k = 0, 1, \ldots, m-1 \tag{17}$$

and $R_{m,n}(a_0, b_0)$ denotes the (non-singular) Sylvester's resultant matrix of polynomials $a_0(s)$ and $b_0(s)$ defined in Theorem 3.1.

3. *With the above definitions:*

$$\gamma^* = \frac{1}{\mu_{\mathbf{\Delta}}(M)} \tag{18}$$

where γ^* denotes the minimum-magnitude complex perturbation in the coefficients of the polynomials $a_0(s)$ and $b_0(s)$ such that the perturbed polynomials $a(s)$ and $b(s)$ defined in equations (3) and (4), respectively, have a common root.

Proof: Since $a_0(s)$ and $b_0(s)$ are assumed co-prime, their Sylvester resultant matrix $R_{m,n}(a_0, b_0)$ is nonsingular. The Sylvester resultant matrix $R_{m,n}(a, b)$ of the perturbed polynomials $a(s; \delta_0, \ldots, \delta_{m-1})$ and $b(s; \epsilon_0, \ldots, \epsilon_{n-1})$ can be decomposed as $R_{m,n}(a, b) = R_{m,n}(a_0, b_0) + E$ where E denotes the "perturbation matrix":

$$E = \begin{pmatrix} 0 & \delta_{m-1} & \delta_{m-2} & \cdots & \delta_0 & 0 & \cdots & 0 \\ 0 & 0 & \delta_{m-1} & \cdots & \delta_1 & \delta_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \delta_{m-1} & \cdots & \delta_1 & \delta_0 \\ 0 & \epsilon_{n-1} & \epsilon_{n-2} & \cdots & \epsilon_0 & 0 & \cdots & 0 \\ 0 & 0 & \epsilon_{n-1} & \cdots & \epsilon_1 & \epsilon_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \epsilon_{n-1} & \cdots & \epsilon_1 & \epsilon_0 \end{pmatrix} \tag{19}$$

Matrix E can now be factored as $E = \Theta\Delta Z$ where Θ and Z are defined in (15) and (16) respectively, and

$$\Delta = \text{diag}(\delta_{m-1}I_n,\ \delta_{m-2}I_n, \ldots, \delta_0 I_n, \epsilon_{n-1}I_m,\ \epsilon_{m-2}I_m, \ldots, \epsilon_0 I_m) \tag{20}$$

Clearly $\Delta \in \mathbf{\Delta}$, i.e. it has a block-diagonal structure with $s = m+n$, $r_i = n$ for $1 \le i \le m$ and $r_i = m$ for $m+1 \le i \le m+n$. Note also that:

$$\gamma = \max\{\ |\delta_0|,\ |\delta_1|, \ldots, |\delta_{m-1}|,\ |\epsilon_0|,\ |\epsilon_1|, \ldots, |\epsilon_{n-1}|\ \} = \|\Delta\| \tag{21}$$

Since the resultant Sylvester matrix $R_{m,n}(a, b)$ loses rank if and only if the polynomials $a(s; \delta_0, \ldots, \delta_{m-1})$ and $b(s; \epsilon_0, \ldots, \epsilon_{m-1})$ have a common root, Problem 2.1 is equivalent to:

$$\min \|\Delta\| \ \text{ such that } \ \det(R_{m,n}(a_0, b_0) + \Theta\Delta Z) = 0 \ \text{ and } \ \Delta \in \mathbf{\Delta} \tag{22}$$

Using the matrix identity

$$\det(I + BC) = \det(I + CB) \tag{23}$$

which holds for any two matrices B and C of compatible dimensions, and the fact that $R_{m,n}(a_0, b_0)$ is non-singular, we have that:

$$\det(R_{m,n}(a_0, b_0) + \Theta\Delta Z) = 0 \;\; \Leftrightarrow \;\; \det(I + ZR_{m,n}^{-1}(a_0, b_0)\Theta\Delta) = 0 \;\; \Leftrightarrow \;\; \det(I - M\Delta) = 0 \tag{24}$$

Hence Problem 2.1 becomes:

$$\min\{\|\Delta\| \; : \; \det(I - M\Delta) = 0, \;\; \Delta \in \mathbf{\Delta} \;\; \} \tag{25}$$

which is equivalent to Problem 2.2, the minimum being $\mu_{\mathbf{\Delta}}^{-1}(M)$. □

For the problem considered in this section there is always a $\Delta \in \mathbf{\Delta}$ of sufficiently large norm such that $\det(I_n - M\Delta) = 0$. This is formally established in the following result.

Lemma 3.1*: Consider Problem* 2.2 *with* $\mathbf{\Delta}$ *and* M *as defined in the statement of Theorem* 3.2*. Then* $\mu_{\mathbf{\Delta}}(M) > 0$ *and hence,*

$$\mu_{\mathbf{\Delta}}(M)^{-1} = \min_{\Delta \in \mathbf{\Delta}, \; \rho(\Delta M) \geq 1} \|\Delta\| = \min_{\Delta \in \mathbf{\Delta}, \; \rho(\Delta M) = 1} \|\Delta\| \tag{26}$$

Proof: Since the polynomials $a_0(s)$ and $b_0(s)$ are assumed co-prime and since there is always a perturbation of finite modulus such that the two perturbed polynomials $a(s)$ and $b(s)$ have a common root, γ^* defined in Theorem 3.2 is a finite positive number. Note that an immediate bound for γ^* is $\gamma^* \leq \max(|\alpha_0|, |\beta_0|)$, corresponding to a perturbation in the constant coefficients of the two polynomials for which the two perturbed polynomials have a common root at the origin. Note also that since $a_0(s)$ and $b_0(s)$ are assumed co-prime, at least one of $|\alpha_0|$ or $|\beta_0|$ is positive. From the equivalence between Problems 2.1 and 2.2 established in Theorem 3.2 we thus have

$$\mu_{\mathbf{\Delta}}(M) \geq \frac{1}{\max(|\alpha_0|, |\beta_0|)} > 0$$

Since $\mu_{\mathbf{\Delta}}(M) \neq 0$, Lemma 2.2 implies that (26) holds. □

Remark 3.1*:* The inverse of the structured singular value $\mu_{\mathbf{\Delta}}^{-1}(M)$ is an exact measure of the "strength of approximation" for the problem of finding the nearest common root of two co-prime polynomials which are subjected to perturbations in their coefficients. This is equivalent to the problem of calculating the numerical GCD of degree one of the two polynomials. Previous work in this area has proposed the inverse of the smallest singular value of the Sylvester resultant matrix $R_{m,n}(a_0, b_0)$ as a measure of the strength of approximation [10], [19], [21]. However, as is well known, the smallest singular value of a matrix measures only the distance to singularity of the matrix when this is subjected to unstructured perturbations, i.e.,

$$\sigma_{m+n}(R_{m,n}(a, b)) = \min\{\|\Delta\| : \Delta \in \mathcal{C}^{(m+n)\times(m+n)}, \det(R_{m,n}(a, b) + \Delta) = 0\}$$

Hence $\sigma_{m+n}^{-1}(R_{m,n}(a, b))$ as a measure of the "strength of approximation" ignores the uncertainty structure entering the resultant matrix due to perturbations in the coefficients of the two polynomials

(see equation 19). As shown in a numerical example included later in this section, the difference between these two measures can be significant. The exact notion of the "strength of approximation" used in this work can be easily generalized, if we assume that only a subset of the polynomial coefficients are subjected to perturbations or if we take into account "correlations" between the coefficients' perturbations. □

Remark 3.2*:* The proof of Theorem 3.2 is based on the observation that the perturbation matrix E can be factored in the form $E = \Theta\Delta Z$ where Θ and Z are constant matrices and Δ is a linear subspace of diagonal matrices. In fact this is always possible when the uncertainty enters the data in affine form [14]: Consider the general affine structure:

$$\delta = (\delta_1, \ldots, \delta_N) \rightarrow A(\delta_1, \ldots, \delta_N) = A_0 + \sum_{i=1}^{N} \delta_i A_i,\ \delta \in \mathcal{C}^N$$

where the matrices $A_i \in \mathcal{C}^{N\times N}$ are given. Let $A_i = \Theta_i Z_i$ be arbitrary factorizations with $\Theta_i \in \mathcal{C}^{n\times l_i}$ and $Z_i \in \mathcal{C}^{l_i\times n}$. Then setting

$$\Theta = (\Theta_1\ \Theta_2\ \ldots\ \Theta_N), \quad Z' = (Z_1'\ Z_2' \ldots\ Z_N')' \quad \text{and} \quad \Delta(\delta) = \mathrm{diag}(\delta_1 I_{l_1}, \delta_2 I_{l_2}, \ldots, \delta_N I_{l_N})$$

we obtain

$$A(\delta_1, \ldots, \delta_N) = A + \Theta\Delta(\delta)Z$$

in which $\Delta(\delta)$ is diagonal. In fact, as can be seen from the above derivation, the assumption that A is a square matrix is not required. Hence the method can be extended to generalized resultants of an arbitrary number of polynomials for which coefficient perturbations also enter in affine form. □

Remark 3.3*:* The formulation of Theorem 3.2 assumes that the data of the problem, i.e. the coefficients of the two nominal polynomials and their perturbations, are complex. The result remains identical if all data are assumed to be real (see Remark 2.3). The problem is more complicated if the coefficients of the two nominal polynomials are complex and their perturbations are restricted to be real. This case corresponds to the "real structured singular value" which is fully analyzed in [14].

Example 3.1*:* Let $a_0(s) = s^3 + \alpha_2 s^2 + \alpha_1 s + \alpha_0$ $(m = 3)$ and $b_0(s) = s^2 + \beta_1 s + \beta_0$ $(n = 2)$. The Sylvester resultant matrix of the perturbed polynomials $a(s)$ and $b(s)$ is:

$$R_{3,2}(a,b) = \begin{pmatrix} 1 & \alpha_2+\delta_2 & \alpha_1+\delta_1 & \alpha_0+\delta_0 & 0 \\ 0 & 1 & \alpha_2+\delta_2 & \alpha_1+\delta_1 & \alpha_0+\delta_0 \\ 1 & \beta_1+\epsilon_1 & \beta_0+\epsilon_0 & 0 & 0 \\ 0 & 1 & \beta_1+\epsilon_1 & \beta_0+\epsilon_0 & 0 \\ 0 & 0 & 1 & \beta_1+\epsilon_1 & \beta_0+\epsilon_0 \end{pmatrix} \tag{27}$$

which can be written as:

$$R_{3,2}(a,b) = R_{3,2}(a_0,b_0) + E = \begin{pmatrix} 1 & \alpha_2 & \alpha_1 & \alpha_0 & 0 \\ 0 & 1 & \alpha_2 & \alpha_1 & \alpha_0 \\ 1 & \beta_1 & \beta_0 & 0 & 0 \\ 0 & 1 & \beta_1 & \beta_0 & 0 \\ 0 & 0 & 1 & \beta_1 & \beta_0 \end{pmatrix} + \begin{pmatrix} 0 & \delta_2 & \delta_1 & \delta_0 & 0 \\ 0 & 0 & \delta_2 & \delta_1 & \delta_0 \\ 0 & \epsilon_1 & \epsilon_0 & 0 & 0 \\ 0 & 0 & \epsilon_1 & \epsilon_0 & 0 \\ 0 & 0 & 0 & \epsilon_1 & \epsilon_0 \end{pmatrix} \tag{28}$$

The "perturbation" matrix E can be factored as:

$$E = \left(\begin{array}{cc|cc|cc|ccc|ccc} 1&0&1&0&1&0&0&0&0&0&0&0\\ 0&1&0&1&0&1&0&0&0&0&0&0\\ \hline 0&0&0&0&0&0&1&0&0&1&0&0\\ 0&0&0&0&0&0&0&1&0&0&1&0\\ 0&0&0&0&0&0&0&0&1&0&0&1 \end{array}\right) \left(\begin{array}{ccccc} \delta_2 I_2 & 0 & 0 & 0 & 0\\ 0 & \delta_1 I_2 & 0 & 0 & 0\\ 0 & 0 & \delta_0 I_2 & 0 & 0\\ 0 & 0 & 0 & \epsilon_1 I_3 & 0\\ 0 & 0 & 0 & 0 & \epsilon_0 I_3 \end{array}\right) \left(\begin{array}{ccccc} 0&1&0&0&0\\ 0&0&1&0&0\\ \hline 0&0&1&0&0\\ 0&0&0&1&0\\ \hline 0&0&0&1&0\\ 0&0&0&0&1\\ \hline 0&1&0&0&0\\ 0&0&1&0&0\\ 0&0&0&1&0\\ \hline 0&0&1&0&0\\ 0&0&0&1&0\\ 0&0&0&0&1 \end{array}\right)$$

which is of the required form $E = \Theta\Delta Z$ with $\Delta \in \mathbf{\Delta}$. The minimum coefficient perturbation is the inverse of the structured singular value of the matrix:

$$M = -\left(\begin{array}{ccccc} 0&1&0&0&0\\ 0&0&1&0&0\\ \hline 0&0&1&0&0\\ 0&0&0&1&0\\ \hline 0&0&0&1&0\\ 0&0&0&0&1\\ \hline 0&1&0&0&0\\ 0&0&1&0&0\\ 0&0&0&1&0\\ \hline 0&0&1&0&0\\ 0&0&0&1&0\\ 0&0&0&0&1 \end{array}\right) \left(\begin{array}{ccccc} 1&\alpha_2&\alpha_1&\alpha_0&0\\ 0&1&\alpha_2&\alpha_1&\alpha_0\\ 1&\beta_1&\beta_0&0&0\\ 0&1&\beta_1&\beta_0&0\\ 0&0&1&\beta_1&\beta_0 \end{array}\right)^{-1} \left(\begin{array}{ccc|ccc|ccc|ccc} 1&0&1&0&1&0&0&0&0&0&0&0\\ 0&1&0&1&0&1&0&0&0&0&0&0\\ \hline 0&0&0&0&0&0&1&0&0&1&0&0\\ 0&0&0&0&0&0&0&1&0&0&1&0\\ 0&0&0&0&0&0&0&0&1&0&0&1 \end{array}\right)$$

and can be computed using standard numerical techniques [33],[34].

Example 3.2***:*** Here the effectiveness of the method is tested with a numerical example. Consider the polynomials

$$a(s) = s^3 - 6s^2 + 11s - 6 \quad \text{and} \quad b(s) = s^2 - 6s + 8 \tag{29}$$

with roots $r_a = \{3, 2, 1\}$ and $r_b = \{4, 2\}$ respectively. Since there is a common root ($s = 2$) the resultant Sylvester matrix $R_{3,2}(a, b)$ is singular. The polynomials were perturbed to:

$$a_0(s) = s^3 - 6.05s^2 + 11.1s - 5.95 \quad \text{and} \quad b_0(s) = s^2 - 6.06s + 8.1 \tag{30}$$

which have roots $r_{a_0} = \{3.0512, 2.0454, 0.9534\}$ and $r_{b_0} = \{4.0301,\ 2.0099\}$ respectively. The singular values of the corresponding Sylvester resultant matrix,

$$R_{3,2}(a_0, b_0) = \left(\begin{array}{ccccc} 1&-6.05&11.1&-5.95&0\\ 0&1&-6.05&11.1&-5.95\\ 1&-6.04&8.1&0&0\\ 0&1&-6.04&8.1&0\\ 0&0&1&-6.04&8.1 \end{array}\right)$$

were obtained as $\lambda(R_{3,2}(a_0, b_0)) = \{22.7997, 12.3247, 5.4710, 0.2264, 0.0007\}$ indicating a numerical rank of 4 and, hence, an approximate GCD of degree one (Theorem 3.1), as expected.

Next the results of Theorem 3.2 were applied to $a_0(s)$ and $b_0(s)$. Note that since the maximum perturbation of the coefficients of $a_0(s)$ and $b_0(s)$ from those of $a(s)$ and $b(s)$ (which have a common root) is 0.1, we expect that $\gamma^* \leq 0.1$.

Two functions from Matlab's μ-optimization toolbox (*mu.m* and *unwrap.m*) were used to calculate the structured singular value of matrix M and the corresponding minimum-norm singularizing perturbation Δ_0. The lower and upper bounds of the structured singular value were obtained as:

$$222.991497161 \leq \mu_{\mathbf{\Delta}}(M) = \frac{1}{\gamma^*} \leq 222.991497162$$

corresponding to $\gamma^* = 0.0044844759227$. Note that although the smallest singular value of $R_{3,2}(a_0, b_0)$ indicates numerical singularity, it is almost six times smaller than γ^* which is the exact distance-to-singularity measure since it takes into account the uncertainty structure of the problem.

It was checked that the singularizing perturbation Δ_0 corresponding to $\mu_{\mathbf{\Delta}}(M)$ had the right structure (i.e. $\Delta_0 \in \mathbf{\Delta}$) and in fact $\delta_0 = \delta_1 = \delta_2 = -\gamma^*$ and $\epsilon_0 = \epsilon_1 = \gamma^*$. Polynomials with a common factor "nearest" to $a_0(s)$ and $b_0(s)$ were obtained with the help of Δ_0 as:

$$\begin{aligned}
\hat{a}(s) &= s^3 - 6.05448447592s^2 + 11.0955155241s - 5.95448447592, \text{ and} \\
\hat{b}(s) &= s^2 - 6.03551552408s + 8.10448447592
\end{aligned}$$

The roots of $\hat{a}(s)$ and $\hat{b}(s)$ were calculated as:

$$r(\hat{a}) = \{3.07886773569, \ \ 2.01656975051, \ \ 0.959046989722\}$$

and

$$r(\hat{b}) = \{4.01894577356, \ \ 2.01656975051\}$$

respectively corresponding to an "optimal" approximate GCD, $\phi(s) = s - 2.01656975051$.

The structured singular value introduced above optimizes the largest magnitude perturbation in the polynomial coefficients. It is also possible to consider alternative perturbation norms, e.g. the weighted Euclidean norm. In this case the corresponding distance-to-singularity measure is given by the "spherical structured singular value" (spherical μ) introduced in [22], [31], [32]. A formal definition follows:

Problem 3.1*: (spherical structured singular value). Let $M \in \mathcal{C}^{n \times n}$ and define the "structured" set:*

$$\mathbf{\Delta} = \{\text{diag}(\delta_1 I_{r_1}, \ \delta_2 I_{r_2}, \ldots, \delta_s I_{r_s}) \ : \ \delta_i \in \mathcal{C}, \ i = 1, 2, \ldots, s\} \tag{31}$$

where the r_i are positive integers such that $\sum_{i=1}^{s} r_i = n$. The spherical structured singular value of M (relative to "structure" $\mathbf{\Delta}$) is defined as:

$$\mu^f_{\mathbf{\Delta}}(M) = \frac{1}{\min\{\|\Delta\|_F \ : \ \Delta \in \mathbf{\Delta}, \ \det(I_n - M\Delta) = 0\}} \tag{32}$$

unless no $\Delta \in \mathbf{\Delta}$ makes $I_n - M\Delta$ singular, in which case $\mu^f_{\mathbf{\Delta}}(M) = 0$. Here $\| \cdot \|_F$ denotes the Frobenius norm of a matrix (i.e. $\|M\|_F^2 = \sum_{i,j} |m_{ij}|^2$). □

The corresponding result in this case is as follows:

Theorem 3.3 *Let all variables be defined as in Theorem* 3.2*. Define also*

$$\gamma^f = \sqrt{n \sum_{i=0}^{m-1} |\delta_i|^2 + m \sum_{i=0}^{n-1} |\epsilon_i|^2} \tag{33}$$

where $m = \partial a_0(s)$ *and* $n = \partial b_0(s)$*. Then the minimum value of* γ^f*,* γ^f_{min}*, such that the perturbed polynomials* $a(s)$ *and* $b(s)$ *defined in equations (3) and (4), respectively, have a common root is given by:*

$$\gamma^f_{min} = \frac{1}{\mu^f_{\mathbf{\Delta}}(M)} \geq \frac{1}{\overline{\mu}^f_{\mathbf{\Delta}}} \tag{34}$$

where M *is defined in the statement of Theorem* 3.2 *part* 2*, and*

$$\overline{\mu}^f_{\mathbf{\Delta}} := inf_{D>0}\{\gamma : M^*(Q \circ D)M - \gamma^2 D < 0\} = \sqrt{\rho((M' \otimes M^*) diag(vec(Q)))} \tag{35}$$

where $Q = diag\{\mathbf{1}_{n\times n}, \ldots, \mathbf{1}_{n\times n}, \mathbf{1}_{m\times m}, \ldots, \mathbf{1}_{m\times m}\}$*, in which the* $\mathbf{1}_{n\times n}$ *blocks appear* m *times and the* $\mathbf{1}_{m\times m}$ *appear* n *times.*

Proof: The procedure for transforming the problem to a spherical structured singular value calculation is essentially the same with the procedure in the proof of Theorem 3.2, on noting that if,

$$\Delta = \text{diag}(\delta_{m-1} I_n,\ \delta_{m-2} I_n, \ldots, \delta_0 I_n, \epsilon_{n-1} I_m,\ \epsilon_{m-2} I_m, \ldots, \epsilon_0 I_m) \tag{36}$$

then

$$\|\Delta\|_F = \sqrt{n \sum_{i=0}^{m-1} |\delta_i|^2 + m \sum_{i=0}^{n-1} |\epsilon_i|^2} \tag{37}$$

The fact that γ^f_{min} is the inverse of $\mu^f_{\mathbf{\Delta}}(M)$ follows immediately from the definition given in Problem 3.1. Finally, the proof of the bound in equations (34) and (35) can be found in [31], [11]. □

4. Approximate GCD of polynomials

The technique developed in section 3 may be used to define a conceptual algorithm to calculate the numerical GCD of any two polynomials $a(s)$ and $b(s)$. This sequentially extracts approximate common factors $\phi_i(s)$ from the two polynomials, by calculating at each step a structured singular value $\mu_{\mathbf{\Delta}}(M)$ of an appropriate matrix M and the corresponding minimum-norm singularizing matrix perturbation Δ_0. After extracting each factor, the quotients $a_{i+1}(s) = a_i(s)/\phi_i(s)$ and $b_{i+1}(s) = b_i(s)/\phi_i(s)$ are calculated, ignoring possible (small) remainders of the division. The procedure is initialized by setting $a_0(s) = a(s)$, $b_0(s) = b(s)$, and iterates by constructing at each step of the algorithm the reduced-dimension Sylvester matrix corresponding to the polynomial pair $(a_{i+1}(s), b_{i+1}(s))$, followed by calculating a new $(\mu_{\mathbf{\Delta}}(M), \Delta_0)$ pair, which in turn leads to the extraction of a new approximate factor $\phi_{i+1}(s)$. The whole process is repeated until a tolerance condition is met (defined by the value of $\mu_{\mathbf{\Delta}}(M)$), at which stage the approximate GCD, $\phi(s)$, can be constructed by accumulating the extracted common factors $\phi_i(s)$. Special care is needed in the real case, to ensure that any complex roots in $\phi(s)$ appear in conjugate pairs.

The algorithm described above essentially consists of the repeated solution of a sub-problem, involving the estimation of the distance to the nearest common root of polynomials whose coefficients are subject

to perturbations. A clear disadvantage of the proposed method is that the optimality properties associated with each sub-problem (which involves the extraction of a degree-one numerical GDC) are lost when the algorithm is considered in total. Formally, we can restore optimality to the problem of extracting a numerical GCD of arbitrary degree by formulating the following optimization problem:

Problem 4.1*: Let $a_0(s)$ and $b_0(s)$ be two monic co-prime polynomials with complex coefficients with $\partial a_0(s) = m$ and $\partial b_0(s) = n$ defined as*

$$a_0(s) = s^m + \alpha_{m-1}s^{m-1} + \alpha_{m-2}s^{m-2} + \ldots + \alpha_0 \tag{38}$$

and

$$b_0(s) = s^n + \beta_{n-1}s^{n-1} + \beta_{n-2}s^{n-2} + \ldots + \beta_0 \tag{39}$$

and assume that the coefficients $\{\alpha_i\}$, $i = 0, 1, \ldots, m-1$ and $\{\beta_i\}$, $i = 0, 1, \ldots, n-1$, are subjected to complex perturbations $\{\delta_0, \delta_1, \ldots, \delta_{m-1}\}$ and $\{\epsilon_0, \epsilon_1, \ldots, \epsilon_{n-1}\}$, respectively, i.e. the perturbed polynomials are:

$$a(s; \delta_0, \ldots, \delta_{m-1}) = s^m + (\alpha_{m-1} + \delta_{m-1})s^{m-1} + (\alpha_{m-2} + \delta_{m-2})s^{m-2} + \ldots + (\alpha_0 + \delta_0) \tag{40}$$

and

$$b(s; \epsilon_0, \ldots, \epsilon_{n-1}) = s^n + (\beta_{n-1} + \epsilon_{n-1})s^{n-1} + (\beta_{n-2} + \epsilon_{n-2})s^{n-2} + \ldots + (\beta_0 + \epsilon_0) \tag{41}$$

respectively. Define also for each positive integer k, $0 < k \leq \min(m, n)$, the minimum infinity norm of the coefficients' perturbation vector:

$$\begin{aligned}\gamma(k) &= \min \left\| \begin{pmatrix} \delta_0 & \delta_1 & \ldots & \delta_{m-1} & \epsilon_0 & \epsilon_1 & \ldots & \epsilon_{n-1} \end{pmatrix} \right\|_\infty \\ &= \min \max\{|\delta_{i-1}| : i \in \underline{m}\} \cup \{|\epsilon_{i-1}| : i \in \underline{n}\}\end{aligned}$$

Then, the problem is to minimize $\gamma(k)$ for each k, $0 \leq k \leq \min(m, n) - 1$, such that the two perturbed polynomials $a(s)$ and $b(s)$ have at least $k + 1$ common roots. □

Compared to the iterative procedure described in the beginning of the section, Problem 4.1 is a more elegant and precise definition of the numerical GCD and replaces the iterative algorithmic approach by the solution of an optimization problem. To link this problem to our approach we need to refine the definition of structured singular values by introducing rank (or equivalently nullity) constraints:

Definition 4.1*: (Generalized structured singular value). Let $M \in \mathcal{C}^{n\times n}$ and define the "structured" set:*

$$\mathbf{\Delta} = \{\text{diag}(\delta_1 I_{r_1},\ \delta_2 I_{r_2}, \ldots, \delta_s I_{r_s}) \ :\ \delta_i \in \mathcal{C},\ i = 1, 2, \ldots, s\}$$

where the r_i are positive integers such that $\sum_{i=1}^{s} r_i = n$. Note that $\mathbf{\Delta}$ is a subspace of $\mathcal{C}^{n\times n}$. The generalized structured singular value of M relative to "structure" $\mathbf{\Delta}$ and for a non-negative integer k, $k \in \{0, 1, \ldots, \min(m, n) - 1\}$, is defined as:

$$\hat{\mu}_{\mathbf{\Delta},k}(M) = \frac{1}{\min\{\|\Delta\| \ :\ \Delta \in \mathbf{\Delta},\ null(I_n - M\Delta) > k\}} \tag{42}$$

unless there does not exist a $\Delta \in \mathbf{\Delta}$ such that $null(I_n - M\Delta) > k$, in which case $\hat{\mu}_{\mathbf{\Delta},k}(M) = 0$.

It follows immediately from the definition that $\hat{\mu}_{\mathbf{\Delta},0}(M) = \mu_{\mathbf{\Delta}}(M)$ and that $\hat{\mu}_{\mathbf{\Delta},k}(M) \geq \hat{\mu}_{\mathbf{\Delta},k+1}(M)$ for each integer $k \geq 0$. Further if for some integer k, $\hat{\mu}_{\mathbf{\Delta},k}(M) > 0$ and $\hat{\mu}_{\mathbf{\Delta},k+1}(M) = 0$, then any $\Delta \in \mathbf{\Delta}$ that minimizes the denominator in the RHS of equation (42) has $\text{null}(I_n - M\Delta) = k + 1$.

Theorem 4.1***:*** *Let $a_0(s)$ and $b_0(s)$ be two monic coprime polynomials of degrees $\partial a_0(s) = m$ and $\partial b_0(s) = n$ defined in equations (38) and (39) respectively. Let $a(s)$ and $b(s)$ be the two perturbed polynomials defined in equations (40) and (41), and set*

$$\gamma = \max\{\ |\delta_0|,\ |\delta_1|, \ldots, |\delta_{m-1}|,\ |\epsilon_0|,\ |\epsilon_1|, \ldots, |\epsilon_{n-1}|\ \} \tag{43}$$

where $\{\delta_{i-1}\}$, $i \in \underline{m}$ and $\{\epsilon_i\}$, $i \in \underline{n}$, denote the perturbations in the coefficients of $a_0(s)$ and $b_0(s)$, respectively. Further, let $\gamma^*(k)$ denote the the minimum value of γ such that $a(s)$ and $b(s)$ have a common factor $\phi(s)$ of degree $\partial\phi(s) > k$, $k = 0, 1, \ldots, \min(m, n) - 1$. Then,

$$\gamma^*(k) = \frac{1}{\hat{\mu}_{\mathbf{\Delta},k}(M)} \tag{44}$$

where $\hat{\mu}_{\mathbf{\Delta},k}(M)$ denotes the generalized structured singular value of $M = -ZR_{m,n}^{-1}(a_0, b_0)\Theta$ with respect to the structure $\mathbf{\Delta}$ and $R_{m,n}(a_0, b_0)$, Θ and Z are as defined in Theorem 3.2. Further, $\phi(s)$ may be constructed from any $\Delta \in \mathbf{\Delta}$ which minimizes $\|\Delta\|$ subject to the constraint null$(I_n - M\Delta) > k$.

Proof: This is almost identical to the proof of Theorem 3.2, on noting that the transformations in equation (24) do not change the nullity of the corresponding matrices. □

Theorem 4.1 suggests that the numerical GCD of two polynomials $a(s)$ and $b(s)$ can be obtained by calculating successively $\hat{\mu}_{\mathcal{D},k}(M)$ for each $k = 0, 1, \ldots, \min(m, n) - 1$, and checking if its value is sufficiently large, which indicates the existence of a numerical GCD with degree at least $k + 1$. The procedure terminates when either $k = \min(m, n) - 1$ is reached (in which case the numerical GCD is the polynomial of minimal degree among $a(s)$ and $b(s)$), or when the generalized structured singular value falls below a pre-specified tolerance level.

We can also state the following Lemma related to the approximate GCD of two polynomials:

Lemma 4.1***:*** *Consider two monic co-prime polynomials $a_0(s)$ and $b_0(s)$ as defined in Problem 4.1. In the notation of Theorem 4.1 we have $\hat{\mu}_{\mathbf{\Delta},k}(M) > 0$ for each $k = 0, 1, \ldots, \min(m, n) - 1$.*

Proof: Since the polynomials $a_0(s)$ and $b_0(s)$ are assumed co-prime, there is always a perturbation of finite non-zero modulus such that $a(s)$ and $b(s)$ have a common root and thus $\gamma^*(k)$ defined in Theorem 4.1 is a positive number. Note that an immediate upper bound for $\gamma^*(k)$ is:

$$\gamma_0^*(k) \leq \max\{|\alpha_0|, \ldots, |\alpha_k|, |\beta_0|, \ldots, |\beta_k|\} \tag{45}$$

corresponding to perturbations $\delta_{i-1} = -\alpha_{i-1}$ and $\epsilon_{i-1} = -\beta_{i-1}$, $i \in \underline{k+1}$, in the coefficients of $a_0(s)$ and $b_0(s)$, such that the two perturbed polynomials $a(s)$ and $b(s)$ have $k + 1$ common roots at the origin. Since clearly the RHS of equation (45) is finite (for otherwise the two polynomials $a(s)$ and $b(s)$ would not be co-prime) we have that

$$\hat{\mu}_{\mathbf{\Delta},k}(M) \geq \frac{1}{\max\{|\alpha_0|, \ldots, |\alpha_k|, |\beta_0|, \ldots, |\beta_k|\}} > 0$$

as required. □

An upper bound of $\hat{\mu}_{\mathbf{\Delta},k}(M)$ can be easily established (see Lemma 4.3) below in terms of the $(k+1)$-th singular value of M. First we need the following result.

Lemma 4.2*: Let $M \in \mathcal{C}^{n\times n}$. Then for every $k \in \underline{n}$,*

$$\begin{aligned}\sigma_k(M) &= \min\{\|\Delta\| : \Delta \in \mathcal{C}^{n\times n}, \mathrm{rank}(M-\Delta) < k\} \\ &= [\inf\{\|\Delta\| : \Delta \in \mathcal{C}^{n\times n}, \mathrm{null}(I_n - \Delta M) \geq k\}]^{-1}\end{aligned}$$

where, by convention, $\inf(\emptyset) = \infty$ and $\infty^{-1} = 0$.

Proof: See [14] Theorem 4.3.13. □

Lemma 4.3*: Let $M \in \mathcal{C}^{n\times n}$ and let $\hat{\mu}_{\mathbf{\Delta},k}(M)$ be as defined in equation (42). Then $\hat{\mu}_{\mathbf{\Delta},k}(M) \leq \sigma_{k+1}(M)$.*

Proof: Assuming that $\hat{\mu}_{\mathbf{\Delta},k}(M)$ is positive,

$$\begin{aligned}\hat{\mu}^{-1}_{\mathbf{\Delta},k}(M) &= \min\{\|\Delta\| : \Delta \in \mathbf{\Delta},\ \mathrm{null}(I_n - M\Delta) > k\} \\ &= \min\{\|\Delta\| : \Delta \in \mathbf{\Delta},\ \mathrm{rank}(I_n - M\Delta) \leq n-k\} \\ &\geq \min\{\|\Delta\| : \Delta \in \mathcal{C}^{n\times n},\ \mathrm{rank}(I_n - M\Delta) \leq n-k\} \\ &= \sigma^{-1}_{k+1}(M)\end{aligned}$$

as required. □

The bound of Lemma 4.3 can be sharpened as follows:

Lemma 4.4*: Let $M \in \mathcal{C}^{n\times n}$ and define the sets $\mathbf{\Delta}$ and $\mathbf{D}$ according to equations (6) and (10) respectively. Then*

$$\hat{\mu}_{\mathbf{\Delta},k}(M) \leq \inf_{D\in\mathbf{D}} \sigma_{k+1}(DMD^{-1})$$

Proof: The Lemma is a straightforward generalization of the inequality in Theorem 2.2. For all $D \in \mathbf{D}$, $\Delta \in \mathbf{\Delta}$,

$$\det(I_n - M\Delta) = \det(I_n - MD^{-1/2}\Delta D^{1/2}) = \det(I_n - D^{1/2}MD^{-1/2}\Delta)$$

since D commutes with Δ. Thus:

$$\begin{aligned}\hat{\mu}_{\mathbf{\Delta},k}(M) &= [\min\{\|\Delta\| : \Delta \in \mathbf{\Delta},\ \mathrm{null}(I_n - M\Delta) > k\}]^{-1} \\ &= \Big[\min\{\|\Delta\| : \Delta \in \mathbf{\Delta},\ \mathrm{null}(I_n - MD^{-1/2}\Delta D^{1/2}) > k\}\Big]^{-1} \\ &= \Big[\min\{\|\Delta\| : \Delta \in \mathbf{\Delta},\ \mathrm{null}(I_n - D^{1/2}MD^{-1/2}\Delta) > k\}\Big]^{-1} \\ &= \hat{\mu}_{\mathbf{\Delta},k}(D^{1/2}MD^{-1/2})\end{aligned}$$

Thus, using the result of Lemma 4.3

$$\hat{\mu}_{\mathbf{\Delta},k}(M) = \hat{\mu}_{\mathbf{\Delta},k}(D^{1/2}MD^{-1/2}) \leq \sigma_{k+1}(D^{1/2}MD^{-1/2})$$

for every $\Delta \in \mathbf{\Delta}$, and hence

$$\hat{\mu}_{\mathbf{\Delta},k}(M) \leq \inf_{D \in \mathbf{D}} \sigma_{k+1}(DMD^{-1})$$

since $D \in \mathbf{D}$ if and only if $D^{1/2} \in \mathbf{D}$. □

The effective numerical calculation of $\hat{\mu}_{\mathbf{\Delta},k}$ (or at least of tight upper and lower bounds) is a challenging non-convex problem and will be investigated in future research. Exploiting the displacement structure of the Sylvester matrix seems to be particularly relevant for this purpose [4], [24]. A further generalization of our method involves the calculation of the approximate GCD of an arbitrary number of polynomials (rather than just two). This problem can also be translated to our framework using some recent results on generalized resultants [15] and will also be elaborated in future work.

5. Conclusions

In this paper we have proposed a new method for calculating numerically the approximate GCD of two polynomials. It was shown that, for two co-prime polynomials, the problem of calculating the smallest l_∞-norm (weighted l_2-norm) perturbation in the polynomials' coefficient vector so that the perturbed polynomials have a common root, is equivalent to the calculation of the structured singular value (spherical structured singular value) of an appropriate matrix. This is a fundamental problem in the area of robust control and various techniques have been successfully developed for its solution. The effectiveness of one such method for calculating the GCD of low-degree polynomials has been demonstrated via a numerical example. We have further shown that calculating the minimum l_∞-norm perturbation in the coefficient vector of two co-prime polynomials so that the perturbed polynomials have a GCD of degree at least k, reduces to a structured singular value-type calculation with additional rank constraints. This is a non-standard problem and developing effective algorithms for its solution will be the topic of future research work, along with extensions of the method to multiple polynomials.

References

[1] S. Barnett (1983), Polynomials and Linear Control Systems, *Marcel Dekker, Inc.,* New York.

[2] S. Barnett (1970), Greatest common divisor of two polynomials, *Linear Algebra Appl.*, 3:7-9.

[3] S. Barnett (1971), Greatest common divisor of several polynomials, *Proc. Camb. Phil. Soc.*, 70:263-268.

[4] D.A. Bini and P. Boito (2007), Structured matrix-based methods for polynomial ϵ-gcd: analysis and comparisons, *Brown, C. W., Ed. ISSAC 2007 Proc. Internat. Symp. Symbolic Algebraic Comput.*, New York, N. Y., ACM Press.

[5] R. P. Braatz, P.M Young, J.C. Doyle and M. Morari (1994), Computational complexity of μ calculation, *IEEE Trans. Automat. Control*, AC-39:1000-1002.

[6] B. Bernhardsson, A. Rantzer and L. Qiu (1998), Real Perturbation Values and Real Quadratic Forms in a Complex Vector Space, *Linear Algebra and its Applications*, 270:131–154.

[7] D. Bini and V.Y. Pan (1994), Polynomial and matrix computations, *Fundamental Algorithms*, Vol. 1, Birkhauser.

[8] J. Chen, M.K.H. Fan and C.N. Nett (1994), Structured singular values and stability analysis of uncertain polynomials, Part 1: the generalized μ, *Systems and Control Letters*, 23:53–65.

[9] M.T. Chu, R.E. Funderlic and R.J. Plemmons (2003) Structured low rank approximation, *Linear Algebra and its Applications*, 366:157-172.

[10] R.M. Corless, S. M. Watt, and L. Zhi (2004), QR factoring to compute the GCD of univariate approximate polynomials, *IEEE Transactions on Signal Processing*, 52:3394-3402.

[11] J.C. Doyle (1998), Robust Control Theory and Applications, *Report AFRL-SR-BL-TR-98*, California Institute of Technology, USA.

[12] N.J. Higham (1992), Estimating the matrix p-norm *Int. J. Control*, 55(4):809–844.

[13] G.Halikias, S. Fatouros and N. Karcanias (2003), Approximate greatest common divisor of polynomials and the structured singular value, *Proc. European Contr. Conf., ECC03*, Cambridge, UK.

[14] D. Hinrichsen and A.J. Pritchard (2000), Mathematical Systems Theory I: Modelling, State-Space Analysis, Stability and Robustness, *Springer Verlag*, New York, USA.

[15] S. Fatouros and N. Karcanias (2002), The GCD of many polynomials and the factorisation of generalised resultants, *Research Report,* Control Engineering Research Centre, City University.

[16] M. Fu and S. Dasgupta (2000), Computational Complexity of Real Structured Singular Value in Setting, *IEEE Tran. Auto. Control*, AC-45(11):2173–2176

[17] I.M. Jaimoukha, G.D. Halikias, U. Malik and S.K Gungah (2006), On the gap between the complex structured singular value and its convex upper bound, *SIAM Journal on Control and Optimization*, 45(4):1251–1278.

[18] N. Karcanias, C. Giannakopoulos and M. Hubbard (1984), Almost zeros of a set of polynomials, *Int. J. Control*, 40:673–698.

[19] N. Karcanias, S. Fatouros, M Mitrouli and G.D. Halikias (2006), Approximate greatest common divisor of many polynomials, generalised resultants and strength of approximation, *Computers and Mathematics with Applications*, 51(12):1817–1830.

[20] M. Karow, E. Kokiopoulou and D. Kressner (2010) On the computation of structured singular values and pseudospectra, *Systems and Control Letters*, 59:122–129.

[21] N. Karcanias and M. Mitrouli (1994), A matrix pencil-based numerical method for the computation of the GCD of polynomials, *IEEE Tran. Auto. Contr.*, AC-39:977–981.

[22] S. Khatri and P. A. Parrilo (1998), Spherical μ, *Proc. American Contr. Conf*, Philadelphia, Pennsylvania, USA, 2314–2318.

[23] E. Kaltofen, Z. Yang and L. Zhi (2007), Structured low rank approximation of a Sylvester matrix, *In Dongming Wang and Lihong Zhi, Trends in Mathematics*, 69–83, Birkhauser Verlag, Basel, Switzerland.

[24] B. Li, Z. Liu and L. Zhi (2007), A Fast Algorithm for Solving the Sylvester Structured TotalLeast Squares Problem, *Signal Processing*, 87:2313–2319

[25] B. Li, Z. Yang and L. Zhi (2005), Fast Low Rank Approximation of a Sylvester Matrix by Structure Total Least Norm *Journal of Japan Society for Symbolic and Algebraic Computation*, 11(3,4):165–174.

[26] A. Megretski A (1993), On the gap between structures singular values and their upper bounds, *Proc. 32nd IEEE Conf. Decison and Control*, 3461–3462, San Antonio, Texas, USA.

[27] M. Mitrouli and N. Karcanias (1993), Computation of the GCD of polynomials using Gaussian transformations and shifting, *Int. J. Control*, 58:211–228.

[28] T. Noda and T. Sasaki (1991), Approximate GCD and its application to ill-conditioned algebraic equations, *J. Comput. Appl. Math.*, 38:335-351.

[29] A. Packard and J. Doyle (1992), The complex structured singular value, *Automatica*, 29(1):71–109.

[30] V. Y. Pan (2001), Numerical computation of a polynomial GCD and extensions, *Information and computation*, 167:71-85.

[31] P. A. Parrilo and S. Khatri (2000), On Cone-Invariant Linear Matrix Inequalities, *IEEE Tran. Auto. Contr.*, AC-45(8):1558–1563.

[32] P.A. Parrilo and S. Khatri (1998), Closed form solutions for a class of LMIs, *Proceedings of the American Control Conference*, 87-91, Philadelphia, Pennsylvania, USA.

[33] P.M. Young and J.C Doyle (1996), Properties of the mixed μ problem and its bounds, *IEEE Tran. Auto. Contr.*, 41(1):155–159.

[34] K. Zhou and J.C Doyle (1998), Essentials of Robust Control, *Prentice Hall, Inc.,* Upper Saddle River, NJ, USA.

[35] X. Zhu, S. H. Khatri and P. A. Parrilo (1999), μ with Linear Cuts: Upper Bound Computation, *Proceedings of the American Control Conference*, 2370–2374, San Diego, California, USA.