Approximate Zero Polynomials of Polynomial Matrices and Linear Systems

Nicos Karcanias and George Halikias

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

Abstract

This paper introduces the notions of approximate and optimal approximate zero polynomial of a polynomial matrix by deploying recent results on the approximate GCD of a set of polynomials [1] and the exterior algebra [4] representation of polynomial matrices. The results provide a new definition for the “approximate”, or “almost” zeros of polynomial matrices and provide the means for computing the distance from non-coprimeness of a polynomial matrix. The computational framework is expressed as a distance problem in a projective space. The general framework defined for polynomial matrices provides a new characterization of approximate zeros and decoupling zeros [2], [4] of linear systems and a process leading to computation of their optimal versions. The use of restriction pencils provides the means for defining the distance of state feedback (output injection) orbits from uncontrollable (unobservable) families of systems, as well as the invariant versions of the “approximate decoupling polynomials”. The overall framework that is introduced provides the means for introducing measures for the distance of a system from different families of uncontrollable, or unobservable systems, which may be feedback dependent, or feedback invariant as well as the notion of “approximate decoupling polynomials”.

Keywords: approximate GCD, exterior algebra, Grassmann varieties, linear systems, determinantal assignment

1. Introduction

The notion of almost zeros and almost decoupling zeros for a linear system has been introduced in [4] and their properties have been linked to mobility of poles under compensation. The basis of that definition has been the representation of the
Plücker embedding by using the Grassmann polynomial vectors [3]. This process has introduced new system invariant and led to the definition of “almost zeros” of a set of polynomials as the minima of a function associated with the polynomial vector [2]. In this paper we develop the concept further by introducing the notion of “approximate zero polynomials” using the exterior algebra framework introduced in [4], and then by deploying the results on the approximate GCD defined in [1]. The notion of “approximate zero polynomials” (AZP) of a polynomial matrix and “optimal” AZP are defined in terms of an optimization expressing the computation of the distance of a point in a projective space from the intersection of two varieties. The first is the Grassmann variety [3], [18] and the second is the given degree GCD variety of the projective space. The results on polynomial matrices are then used to define the “approximate input, output decoupling zero polynomials” and “approximate zero polynomial” of a linear system. The framework that is introduced allows the characterisation of strength of the given order approximate zero polynomial, as well as permits the characterisation of the optimal approximate solutions of a given order as the solution of an optimization problem.

Defining the distance of a system described by the pair \((A, B)\) (pair \((A, C)\)) from the family of uncontrollable (unobservable) systems has been a subject under consideration for some time [20], [21], [22]. It is worth pointing that although the controllability (observability) properties are invariant under state feedback (output injection), their “strength” (measured with different criteria) is not. This raises the question of whether invariant measures can be defined. This problem is also addressed here from a much more general perspective. In fact, we introduce a new framework for evaluating such distances that allows the computation of the specific system \((A, B)\) (\((A, C)\)), as well as the state feedback (output injection) orbits \((A + BL, B)\) (\((A + KC, C)\)) from the uncontrollable (unobservable) systems. The latter is a new dimension to the problem and it is complemented by the definition of the corresponding decoupling zero polynomials. It is worth noting that the suggested approach allows the computation of decoupling polynomials of varying degrees.

We are using the exterior algebra framework by deploying the Plücker embedding [3] to associate polynomial vectors to polynomial matrices; thus we have a framework that allows a proper definition of the notion of “approximate matrix divisor” of polynomial models, as well as the notion of the distance of a polynomial matrix from families of non-coprime matrices. It is shown that the characterisation and computation of an “approximate matrix divisor” is equivalent to a distance problem of a general set of polynomials from the intersection of two
varieties, a GCD (defined by the degree of the desirable GCD) and the dynamic Grassmann variety that is defined by the Forney order [11] of the polynomial matrix. The notion of approximate matrix divisor introduced here refers to a family of square matrices all having the same polynomial as determinant.

The results introduce a computational framework that potentially can provide the means for defining “approximate zero polynomials” for linear systems and introduce new measures of distance of systems from uncontrollability, un-observability using the “strength” associated with a given approximate solution. The characterisation of distance from uncontrollability, un-observability uses the algebraic matrix pencil characterisation [8], [9] which is based on the properties of Grassmann vectors and associated Plücker matrices of the corresponding pencils [12]. This distance may be parameterised by the degree of the corresponding decoupling zero polynomial. Using the algebraic feedback free criteria introduced by the restriction pencils [17], [18] a new notion of distance that is invariant under feedback is introduced, which expresses distance from state feedback orbit (uncontrollability case), output injection orbit (un-observability case). The use of Grassmann vectors, that is polynomial vectors in a projective space, implies that the general results on the “strength” of approximation, defined in [1] for polynomial vectors, yield lower bounds for the corresponding approximate polynomials of polynomial matrices; for two boundary cases (systems with one input (output), or systems with \( n - 1 \) inputs (outputs), where \( n \) is the number of states), the results coincide.

2. Definitions and preliminary results

Consider the linear system \( S(A,B,C,D) \):

\[
S(A,B,C,D) : \quad \dot{x} = Ax + Bu, \quad y = Cx + Du
\] (1)

where \( A \in \mathbb{R}^{n \times n} \), \( B \in \mathbb{R}^{n \times p} \), \( C \in \mathbb{R}^{m \times n} \) and \( D \in \mathbb{R}^{m \times p} \). It is assumed that \( (A,B) \) is controllable and \( (A,C) \) is observable. Alternatively, \( S(A,B,C,D) \) is defined by the transfer function matrix represented in terms of left, right coprime matrix fraction descriptions (LCMFD, RCMFD), as

\[
G(s) = D_l(s)^{-1}N_l(s) = N_r(s)D_r(s)^{-1}
\] (2)

where \( N_l(s), N_r(s) \in \mathbb{R}^{m \times p}[s] \), \( D_l(s) \in \mathbb{R}^{m \times m}[s] \) and \( D_r(s) \in \mathbb{R}^{p \times p}[s] \). The system will be called square if \( m = p \) and non-square if \( m \neq p \). We shall denote by \( N \) a
left annihilator of $B$, i.e. $N \in \mathbb{R}^{(n-p) \times n}$, $NB = 0$ and by $M$ a right annihilator of $C$, i.e. $M \in \mathbb{R}^{n \times (n-m)}$, $CM = 0$, where both $N, M$ have full rank.

The family of frequency assignment problems has a common formulation that allows a unifying treatment in terms of the Abstract Determinantal Assignment Problem. Thus:

(i) Pole assignment by state feedback: Pole assignment by state feedback $L \in \mathbb{R}^{p \times n}$ reduces to the determinantal problem:

$$p_L(s) = \det \{sI - A - BL\} = \det \{B(s)\tilde{L}\}$$

where $B(s) = [sI - A, -B]$ is defined as the system controllability pencil and $\tilde{L} = [I_n, L']^t$. The zeros of $B(s)$ are the input decoupling zeros of the system [7], [23].

(ii) Design observers: The design of an $n$-state observer by an output injection $T \in \mathbb{R}^{n \times m}$ reduces to the determinantal problem:

$$p_T(s) = \det \{sI - A - TC\} = \det \{\tilde{T}C(s)\}$$

where $C(s) = [sI - A', -C']^t$ is the observability pencil and $\tilde{T} = [I_n, T]$ represents output injection. The zeros of $C(s)$ define the output decoupling zeros [7].

(iii) Zero assignment by squaring down: Given a system with $m > p$ and $c \in \mathbb{R}^p$ the vector of the variables which are to be controlled, then $c = Hy$ where $H \in \mathbb{R}^{p \times m}$ is a squaring down post-compensator, and $G'(s) = HG(s)$ is the squared down transfer function matrix [5]-[7]. A right MFD for $G'(s)$ is defined $G'(s) = H N_r(s) D_r(s)^{-1}$, $G(s) = N_r(s) D_r(s)^{-1}$. Finding $H$ such that $G'(s)$ has assigned zeros is defined as the zero assignment by squaring down problem [5], [7] and the zero polynomial of $S(A, B, HC, HD)$ is

$$z_K(s) = \det \{HN_r(s)\}$$

2.1. The Abstract Determinantal Assignment Problem (DAP):

DAP is a multilinear nature problem of a determinantal character. If $M(s) \in \mathbb{R}^{p \times r}[s]$, $r \leq p$ such that rank $\{M(s)\} = r$ and let $\mathcal{H}$ be a family of full rank $r \times p$ constant matrices having a certain structure then DAP is reduced to solve equation (6) with respect to $H \in \mathcal{H}$

$$f_M(s, H) = \det \{HM(s)\} = f(s)$$

where $f(s)$ is a real polynomial of some degree $d$. 
Note that if $M(s) = \overline{M}(s)Z(s)$ where $\overline{M}(s) \in \mathbb{R}^{p \times r}[s]$ and $Z(s) \in \mathbb{R}^{r \times r}[s]$ is a Greatest Right Divisor of $M(s)$ (the zeros of $Z(s)$ define the zeros of $M(s)$ [9]), then clearly

$$f_M(s, H) = \det\{HM(s)\} = \det\{H\overline{M}(s)\} \det\{Z(s)\} \quad (7)$$

**Remark 1:** The zeros of $M(s)$ defined by the zeros of $Z(s)$, are fixed zeros of all polynomial combinants $f(s)$. This implies that input (output) decoupling zeros are fixed zeros under state feedback (output injection) and nonsquare zeros are fixed zeros under all squaring down compensators. For the case of polynomial matrices, the zeros are expressed as zeros of matrix divisors [9], or as the roots of the GCD of a polynomial multi-vector [3], [4]. It is the latter formulation that allows the development of a framework for defining “almost zeros” in a way that also permits the quantification of the strength of approximation. The results in [1] are deployed here to provide a new characterisation of “approximate” zeros of different type for Linear Systems.

**Notation** [10]: Let $Q_{k,n}$ be the set of lexicographically ordered, strictly increasing sequences of $k$ integers from $1, 2, \ldots, n$. By $\chi$ we denote a vector and by $c$ a scalar. If $\{x_{i_1}, \ldots, x_{i_k}\}$ is a set of vectors of a vector space $\mathcal{V}$, $\omega = (i_1, \ldots, i_k) \in Q_{k,n}$, then $x_{i_1} \wedge \ldots \wedge x_{i_k} = x_{\omega} \wedge$ denotes the exterior product and by $\wedge^r \mathcal{V}$ we denote the $r$-th exterior power of $\mathcal{V}$. If $H \in F^{m \times n}$ and $r \leq \min\{m, n\}$, then by $C_r(H)$ we denote the $r$-th compound matrix of $H$ [10].

If $h^i_j, m^i_j(s), i \in I$, denote the rows of $H$, columns of $M(s)$ respectively, then

$$C_r(H) = h_1^r \wedge \ldots \wedge h_r^r = h^r \wedge \in \mathbb{R}^\sigma \quad (8)$$

and

$$C_r(M(s)) = m_1^r(s) \wedge \ldots \wedge m_r(s) = m(s) \wedge \in \mathbb{R}^\sigma[s], \quad \sigma = \binom{p}{r} \quad (9)$$

and by Binet-Cauchy theorem [10] we have that [4]:

$$f_M(s, H) = C_r(H)C_r(M(s)) = \langle h^r, m(s) \wedge \rangle = \sum_{\omega \in Q_{r,p}} h_{\omega} m_{\omega}(s)$$

$\omega = (i_1, \ldots, i_r) \in Q_{r,p}$, and $h_{\omega}, m_{\omega}(s)$ are the coordinates of $h^r \wedge, m(s) \wedge$, respectively. Note that $h_{\omega}$ is the $r \times r$ minor of $H$ which corresponds to the $\omega$ set of columns of $H$ and thus $h_{\omega}$ is a multilinear alternating function [3] of the entries $h_{ij}$ of $H$. 

5
**DAP Linear sub-problem:** Set \( m(s) \land p(s) \in \mathbb{R}^\sigma [s], f(s) \in \mathbb{R}[s] \). Determine the existence of \( k \in \mathbb{R}^\sigma, k \neq 0 \), such that

\[
f_M(s, H) = k^T p(s) = \sum k_i p_i(s) = f(s), \quad i \in \sigma
\]  \hspace{1cm} (10)

**DAP Multilinear sub-problem:** Assume that \( \mathcal{X} \) is the family of solution vectors \( k \) of equation (5). Determine if there exists \( H' = [h_1, \ldots, h_r], H' \in \mathbb{R}^{p \times r} \), such that

\[
h_1 \land \ldots \land h_r = h \land k, \quad k \in \mathcal{X}
\]  \hspace{1cm} (11)

**Lemma 1** [3]: Let \( k \in \mathbb{R}^\sigma, \sigma = (p) \) and let \( k_{\omega}, \omega = (i_1, \ldots, i_r) \in Q_{r,p} \) be the Plücker coordinates of a point in \( P_{r-1}(\mathbb{R}) \). Necessary and sufficient condition for the existence of \( H \in \mathbb{R}^{r \times p}, H = [h_1, \ldots, h_r]^T \), such that

\[
h_1 \land \ldots \land h_r = k = [\ldots, k_{\omega}, \ldots]^T
\]  \hspace{1cm} (12)

is that the coordinates \( k_{\omega} \) satisfy the quadratics

\[
\sum_{k=1}^{r+1} (-1)^{\nu - 1} k_{i_1, \ldots, i_{\nu-1}, j_1, \ldots, j_{\nu-1}, j_{\nu+1}} = 0
\]  \hspace{1cm} (13)

where \( 1 \leq i_1 < i_2 < \ldots < i_{\nu-1} \leq n, 1 \leq j_1 < j_2 < \ldots < j_{\nu+1} \leq n \). \hfill \square

The quadratics defined by equation (13) are known as the **Quadratic Plücker Relations (QPR)** [3] and they define the Grassmann variety \( \Omega(r, p) \) of \( P_{r-1}(\mathbb{R}) \).

### 3. Grassman invariants of linear systems

Consider \( T(s) \in \mathbb{R}^{p \times r}[s] \), \( T(s) = [t_1(s), \ldots, t_r(s)] \), \( p \geq r \), rank \( \{ T(s) \} = r \), \( \mathcal{X}_r = \text{Range}_{\mathbb{R}[s]}(T(s)) \). If \( T(s) = M(s)D(s)^{-1} \) is a RCMFD of \( T(s) \), then \( M(s) \) is a polynomial basis for \( \mathcal{X}_r \). If \( Q(s) \) is a greatest right divisor of \( M(s) \) then \( T(s) = M(s)Q(s)D(s)^{-1} \), where \( M(s) \) is a least degree polynomial basis of \( \mathcal{X}_r \) [9]. A Grassmann Representative (GR) for \( \mathcal{X}_r \) is defined by [4]

\[
 t(s) \land = t_1(s) \land \ldots \land t_r(s) = \overline{m_1}(s) \land \ldots \land \overline{m_r}(s) \cdot z_t(s)/p_t(s)
\]  \hspace{1cm} (14)

where \( z_t(s) = \det \{ Q(s) \}, p_t(s) = \det \{ D(s) \} \) are the zero, pole polynomials of \( T(s) \) and \( \overline{m}(s) = \overline{m_1}(s) \land \ldots \land \overline{m_r}(s) \in \mathbb{R}^\sigma [s], \sigma = (p) \), is also a GR of \( \mathcal{X}_r \). Since \( \overline{M}(s) \) is a least degree polynomial basis for \( \mathcal{X}_r \), the polynomials of \( \overline{m}(s) \land \) are coprime and \( \overline{m}(s) \land \) will be referred to as a **reduced polynomial** GR (\( \mathbb{R} - \mathbb{R}[s] \)- GR) of \( \mathcal{X}_r \).
If $\delta = \deg \{\bar{m}(s)\wedge\}$, then $\delta$ is the Forney dynamical order [11] of $\mathcal{X}_t$. $\bar{m}(s)\wedge$ may be expressed as

$$\bar{m}(s)\wedge = p(s) = p_0 + p_1 s + \ldots + p_\delta s^\delta = P_\delta \cdot e_\delta(s)$$  \hspace{1cm} (15)

where $P_\delta \in \mathbb{R}^{m \times (\delta + 1)}$ is a basis matrix for $\bar{m}(s)\wedge$ and $e_\delta(s) = [1, s, \ldots, s^\delta]'$. All $\mathbb{R}[s]$-GRs of $\mathcal{X}_t$ differ only by a nonzero scalar factor $a \in \mathbb{R}$ and if $\|P_\delta\| = 1$, we define the canonical $\mathbb{R}[s]$-GR $g(\mathcal{X}_t)$ and the basis matrix $P_\delta$ is the Plücker matrix of $\mathcal{X}_t$ [4].

**Theorem 1:** $g(\mathcal{X}_t)$, or the associated Plücker matrix $P_\delta$, is a complete (basis free) invariant of $\mathcal{X}_t$. \Box

In fact, the Grassmann representative and associated Plücker matrices provide alternative complete invariants for rational vector spaces to those introduced by the Forney theory of minimal bases [11].

If $M(s) \in \mathbb{R}^{p \times r}[s]$, $p \geq r$, rank$\{M(s)\} = r$, is a polynomial basis of $\mathcal{X}_t$, then $M(s) = \bar{M}(s)Q(s)$, where $\bar{M}(s)$ is a least degree polynomial basis and $Q(s)$ is a greatest right divisor of the rows of $M(s)$ and thus

$$\bar{m}(s)\wedge = \bar{m}(s)\wedge \cdot \det\{Q(s)\} = P_\delta \cdot e_\delta(s)z_m(s)$$  \hspace{1cm} (16)

For the control problems discussed before the vector $m(s)\wedge$ has important properties which stem from the properties of the corresponding system. A number of Plücker type matrices are:

(a) **Controllability Plücker Matrices**: For the pair $(A, B)$, $b(s)\wedge$ denotes the exterior product of the rows of $B(s) = [sI - A, -B]$ and $P(A, B)$ is the basis matrix of $b(s)\wedge$, then $P(A, B)$ is the controllability Plücker matrix and its rank characterises the controllability properties. If $N$ is an $(n - p) \times n$ left annihilator of $B \ (NB = 0)$ and $N$ basis matrix for left null space of $B$, the pencil $R_c(s) = sN - NA \in \mathbb{R}^{(n-p) \times n}[s]$ provides the “feedback free” description of the original forced system and it is referred to as restricted input state pencil [19]; this pencil is invariant under the state feedback orbit and its elementary divisors define the set of input-decoupling zeros [8] of the system. Similarly, if $M$ is an $n \times (n - m)$ basis matrix for the right null space of (case of a strictly proper system) the pencil $R_o(s) = sM - AM$ provides the “output injection free” description of the original forced system and it is referred to as restricted state output pencil [19] and its zeros (roots of elementary divisors) define the set of output decoupling zeros of the system [8]. The pencils $R_c(s)$ and $R_o(s)$ provide alternative transformation free characterisations of controllability, respectively observability.
and these allow the introduction of new measures for relative controllability respectively observability. If we denote by $\mathbf{r}(s)^\mathbf{\wedge}$ the exterior product of the rows of $R(s)$ and $P(N, NA)$ is the basis matrix of $\mathbf{r}(s)^\mathbf{\wedge}$, then $P(N, NA)$ will be referred to as the restricted controllability Plücker matrix of the system.

**Theorem 2** [12]: $S(A, B)$ is controllable, iff $P(A, B)$ or equivalently $P(N, NA)$ has full rank.

(b) Observability Plücker Matrix: For the pair $(A, C)$, $\mathbf{c}(s)^\mathbf{\wedge}$ denotes the exterior product of the columns of $C(s) = [sI - A^\mathbf{T}, -C]^\mathbf{T}$ and $P(A, C)$ is the basis matrix of $\mathbf{c}(s)^\mathbf{\wedge}$. $P(A, C)$ is the observability Plücker matrix and its rank characterises system observability. For the linear system an equivalent “output injection feedback-free” characterisation of observability [19] is provided by the output-restricted pencil $Q(s) = sM - AM \in \mathbb{R}^{n \times (n-m)}$; this pencil is invariant of the output-injection orbit and its elementary divisors define the set of output-decoupling zeros of the system. If we denote by $\mathbf{q}(s)^\mathbf{\wedge}$ the exterior product of the rows of $Q(s)$ and $P(M, AM)$ is the basis matrix of $\mathbf{q}(s)^\mathbf{\wedge}$, then $P(M, AM)$ will be referred to as the restricted observability Plücker matrix of the system.

**Theorem 3** [12]: $S(A, C)$ is observable, iff $P(A, C)$ or equivalently $P(M, AM)$ has full rank.

The above results provide alternative new tests for the controllability and observability properties, which are now based on Plücker matrices which are directly related the corresponding problems of frequency assignment. Although the controllability (observability) property is invariant under state feedback (output injection), the relative degrees of such properties may change. The use of "feedback free" descriptions provided by the restriction pencils allows the extension of distance on system orbits under feedback transformations.

**Remark 2**: As far as the exact properties of controllability (observability) the pencils $B(s), R(s), C(s), Q(s)$ provide equivalent characterisations. The invariance of $R(s), Q(s)$ under feedback has significant differences when it comes to characterising the “relative degree” of controllability, observability, respectively. The relative rank properties of the matrices $P(A, B), P(N, NA)$ and $P(A, C), P(M, AM)$ as defined by the singular values characterise respectively different system properties. In fact, rank properties of:

(i) $P(A, B), P(A, C)$ provide an indication for relative controllability and observability respectively.

(ii) $P(N, NA), P(M, AM)$ provide an indication for relative controllability and observability of the state feedback, output injection orbit respectively.  

8
It is worth noting that the vectors $\mathbf{b}^\wedge$, $\mathbf{r}^\wedge$, $\mathbf{c}^\wedge$, and $\mathbf{q}^\wedge$ are decomposable multivectors [3] and thus the corresponding matrix coefficients should satisfy special conditions (based on the Quadratic Plücker Relations (QPRs) [3]) and thus they are sub-families of the corresponding general sets of matrices. This leads to:

**Theorem 4:** The smallest singular values of the Plücker matrices may be used to provide lower bounds for the distance from the family uncontrollable (unobservable) systems. In particular, the smallest singular values of:

(i) $P(A, B) (P(A, C))$ provide a lower bound for distance of the system $S(A, B)$ ($S(A, C)$) from the family of uncontrollable (unobservable) systems.

(ii) $P(N, NA) (P(M, AM))$ provide a lower bound for distance of the state feedback (output injection) system orbit $S(A + BL, B)$ ($S(A + KC, C)$) from the family of uncontrollable (unobservable) systems.

**Proof:** (i) It is well known [24] that the minimal singular value defines the distance of a matrix $T$ from the set of singular matrices of the same dimension. However, this result assumes that the elements of $T$ are not constrained. The matrices $P(A, B)$, $P(A, C)$ are the coefficient matrices of Grassmann representatives [3] (that is they are obtained as the coefficient matrices of polynomial vectors and thus they are decomposable matrices [3]). As such the elements of such matrices are not free and thus their minimal singular value is a lower bound. (ii) The restricted input state pencil has the same value $R_c(s) = sN - NA \in \mathbb{R}^{(n-p) \times n}[s]$ when defined on the open loop or the closed-loop under state feedback $L$, since $R_c(s) = sN - NA = sN - N(A + BL)$, since $NB = 0$. Clearly, $R_c(s) = sN - NA$ characterises the state feedback orbit and not an individual system. The singular value bound then clearly follows from the same arguments on the structured nature of matrices given in part (i). The proof for observability follows along similar lines.

**Corollary 1:** For the case where: (i) $p = 1$ or $p = n - 1$ for $(A, B)$, or (ii) $m = 1$ or $m = n - 1$ for $(A, C)$, the lower bounds defined by the minimal singular bounds become exact.

**Proof:** It is clear that for the case $p = 1$, $m = 1$ the associated matrices $P(A, B)$, $P(A, C)$, or respectively $P(N, NA)$, $P(M, AM)$, are unstructured and thus the minimal singular value is not any more a bound and defines the distance. For the cases $p = n - 1$, $m = n - 1$ it is known [3] that the corresponding multi-vectors are always decomposable, as well as the corresponding matrices. The respective matrices are thus unstructured and the corresponding minimal singular values define the distance.
The use of the restriction pencils allows for the first time the definition and estimation of the distance of orbits of systems from uncontrollability, unobservability respectively.

**Remark 3:** Improving the above distance results requires using the properties of the structured nature of the related matrices. Such problems may be reduced to problems of computation of structural singular values [25], [26].

(c) **Column Plücker Matrices:** For the transfer function \( G(s), m \geq p, n(s)\wedge \) is the exterior product of the columns of the numerator \( N_r(s), \) of a RCMFD and \( P(N) \) is the basis matrix of \( n(s)\wedge. \) Note that \( d = \delta, \) the Forney order of \( \mathcal{H}_r, \) if \( G(s) \) has no finite zeros and \( d = \delta + k, \) where \( k \) is the number of finite zeros of \( G(s), \) otherwise. If \( N_r(s) \) is least degree, then \( P_c(N) \) is the column space Plücker matrix.

**Theorem 5** [13]: For a generic system with \( m > p, \) for which \( p(m - p) > \delta + 1, \) where \( \delta \) is the Forney order, \( P_c(N) \) has full rank. □

The rank properties of \( P_c(N) \) matrix characterize the properties of “almost zeros” [4] and “approximate zero polynomial” of a linear system based on the notion of “approximate GCD” [1] which is discussed next.

**4. Approximate GCD of Polynomial Sets**

Consider a set \( P = \{a(s), b_i(s) \in \mathbb{R}[s], i = 1, 2, \ldots, h\} \) of polynomials which has \( h + 1 \) elements and with the two largest values of degrees \((n, p), \) which is also denoted as \( P_{h+1, n}. \) The greatest common divisor (GCD) of \( P \) will be denoted by \( \varphi(s). \) For any \( P_{h+1, n} \) we define a vector representative \( p_{h+1, n}(s) \) and a basis matrix \( P_{h+1}. \) The classical approaches for the study of coprimeness and determination of the GCD makes use of the Sylvester Resultant, \( S_P \) and the GCD properties are summarised below [14], [15]:

**Theorem 6:** For a set of polynomials \( P_{h+1, n} \) with a resultant \( S_P \) the following properties hold true:

1. Necessary and sufficient condition for a set of polynomials to be coprime is that:

   \[
   \text{rank}(S_P) = n + p \tag{17}
   \]

2. Let \( \varphi(s) \) be the GCD of \( P. \) Then:

   \[
   \text{rank}(S_P) = n + p - \deg \varphi(s) \tag{18}
   \]

3. If we reduce \( S_P, \) by using elementary row operations, to its row echelon form, the last non-vanishing row defines the coefficients of the GCD. □
The results in [15] establish a matrix based representation of the GCD, which is equivalent to the standard algebraic factorisation of the GCD of polynomials. This new GCD representation provides the means to define the notion of the “approximate GCD” subsequently in a formal way, and thus allows the definition of the optimal solution.

**Theorem 7:** Let \( P = \{a(s), b_1(s), \ldots, b_h(s)\} \), deg \( a(s) = n \), deg \( b_i(s) \leq p \leq n \), \( i = 1, \ldots, h \) be a polynomial set, \( S_P \) the respective Sylvester matrix, \( \varphi(s) = \lambda_k s^k + \cdots + \lambda_1 s + \lambda_0 \) be the GCD of the set and let \( k \) be its degree. Then there exists transformation matrix \( \Phi_\varphi \in \mathbb{R}^{(n+p) \times (n+p)} \) such that:

\[
\tilde{S}_P^{(k)} = S_P \Phi_\varphi = \begin{bmatrix} 0_k \mid \tilde{S}_P \end{bmatrix}
\]  

or

\[
S_P = \tilde{S}_P^{(k)} \Phi_\varphi^{-1} = \begin{bmatrix} 0_k \mid \tilde{S}_P \end{bmatrix} \Phi_\varphi
\]

where \( \Phi_\varphi = \hat{\Phi}_\varphi^{-1}, \hat{\Phi}_\varphi \) being the Toeplitz form of \( \varphi(s) \) [15] and

\[
\tilde{S}_P^{(k)} = \begin{bmatrix}
0 & S_0^{(k)} \\
0 & S_1^{(k)} \\
\vdots & \vdots \\
0 & S_h^{(k)}
\end{bmatrix} = [0 \ 0 \ 0 \ \cdots \ 0 \ \tilde{S}_P]
\]

where \( S_i^{(k)} \) are Toeplitz blocks corresponding to the coprime polynomials obtained from the original set after the division by the GCD.

The problem which is addressed next is the formal definition of the notion of the “approximate GCD” [1] and the evaluation of its strength. We shall denote by \( \Pi(n, p; h+1) \) the set of all polynomial sets \( P_{h+1,n} \) with the \( (n,p) \) the maximal two degrees and \( h+1 \) elements. If \( P_{h+1,n} \in \Pi(n, p; h+1) \) we can define an \( (n,p) \)-ordered perturbed set

\[
P_{h+1,n}' = P_{h+1,n} - Q_{h+1,n} \in \Pi(n, p; h+1)
\]

\[
= \{ p_i'(s) = p_i(s) - q_i(s) : \deg q_i(s) \leq \deg p_i(s) \}
\]

The above process is described by Figure 1 at the end of the paper.

**Lemma 2** [1]: For a set \( P_{h+1,n} \in \Pi(n, p; h+1) \) and an \( \omega(s) \in \mathbb{R}[s] \) with \( \deg \omega(s) \leq p \), there always exists a family of \((n,p)\)-ordered perturbations \( Q_{h+1,n} \) and for every element of this family \( P_{h+1,n}' = P_{h+1,n} - Q_{h+1,n} \) has a GCD divisible by \( \omega(s) \).
**Definition 2:** Let \( P_{h+1,n} \in \Pi(n, p; h + 1) \) and \( \omega(s) \in \mathbb{R}[s] \) be a given polynomial with \( \deg \omega(s) = r \leq p \). Let \( \Sigma_\omega = \{ Q_{h+1,n} \} \) be the set of all \((n, p)\)-order perturbations

\[
P'_{h+1,n} = P_{h+1,n} - Q_{h+1,n} \in \Pi(n, p; h + 1)
\]

(24)

with the property that \( \omega(s) \) is a common factor of the elements of \( P'_{h+1,n} \). If \( Q^*_{h+1,n} \) is the minimal norm element of the set \( \Sigma_\omega \), then \( \omega(s) \) is referred as an \( r \)-order almost common factor of \( P_{h+1,n} \), and the norm of \( Q^*_{h+1,n} \), denoted by \( \| Q^* \| \), as the strength of \( \omega(s) \). If \( \omega(s) \) is the GCD of

\[
P^*_n = P_{h+1,n} - Q^*_{h+1,n}
\]

(25)

then \( \omega(s) \) will be called an \( r \)-order almost GCD of \( P_{h+1,n} \) with strength \( \| Q^* \| \). A polynomial \( \hat{\omega}(s) \) of degree \( r \) for which the strength \( \| Q^* \| \) is a global minimum will be called the \( r \)-order optimal almost GCD (OA-GCD) of \( P_{h+1,n} \).

The above definition suggests that any polynomial \( \omega(s) \) may be considered as an “approximate GCD”, as long as \( \deg \omega(s) \leq p \). Important issues in the definition of approximate (optimal approximate) GCD are the parameterisation of the \( \Sigma_\omega \) set, the definition of an appropriate metric for \( Q_{h+1,n} \) and the solution of the optimization problem to define \( Q^*_{h+1,n} \). The set of all resultants corresponding to \( \Pi(n, p; h + 1) \) set, will be denoted by \( \Psi(n, p; h + 1) \).

**Remark 4:** If \( P_{h+1,n}, Q_{h+1,n}, P'_{h+1,n} \in \Pi(n, p; h + 1) \) are sets of polynomials and \( S_P, S_Q, S_P' \) denote their generalised resultants, then these resultants are elements of \( \Psi(n, p; h + 1) \) and \( S_P' = S_P - S_Q \).

**Theorem 8:** Let \( P_{h+1,n} \in \Pi(n, p; h + 1) \) be a set, \( S_P \in \Psi(n, p; h + 1) \) be the corresponding generalised resultant and let \( \nu(s) \in \mathbb{R}[s] \), \( \deg \nu(s) = r \leq p \), \( \nu(0) \neq 0 \). Any perturbation set \( Q_{h+1,n} \in \Pi(n, p; h + 1) \), i.e. \( P'_{h+1,n} = P_{h+1,n} - Q_{h+1,n} \), which has \( \nu(s) \) as common divisor, has a generalized resultant \( S_Q \in \Psi(n, p; h + 1) \) that is expressed as

\[
S_Q = S_P - S_P^{(r)} \Phi_v = [0_r \mid S_P^\nu] \Phi_v
\]

(26)

where \( \Phi_v \) is the Toeplitz representation of \( \nu(s) \) and \( S_P^\nu \in \mathbb{R}^{(p+hn) \times (n+p-r)} \) the \((n, p)\)-expanded resultant of a \( P^*_{h+1,n} = P_{h+1,n} - Q_{h+1,n} \), has the Toepplitz representation of \( S_P^\nu \) is such that \( S_P^\nu \) has full rank, then \( \nu(s) \) is a GCD of the generalized

**Remark 5:** The result provides a parameterisation of all perturbations \( Q_{h+1,n} \in \Pi(n, p; h + 1) \) which yield sets \( P'_{h+1,n} \) having a GCD with degree at least \( r \) and divided by the given polynomial \( \nu(s) \). The free parameters are the coefficients of
the $P_{h+1,n-r}^* \in \Pi(n-r,p-r,h+1)$ set of polynomials. For a set of parameters, $\nu(s)$ is a divisor of $P_{h+1,n}';$ for generic sets, $\nu(s)$ is a GCD of $P_{h+1,n}^*.$ 

A useful metric for evaluation of strength of “approximate GCD” has to relate to the coefficients of the polynomials and the Frobenius norm seems to be an appropriate choice.

**Corollary 2**: Let $P_{h+1,n}^* \in \Pi(n,p;h+1)$ and $\nu(s) \in \mathbb{R}[s]$, $\deg \nu(s) = r \leq p$. The polynomial $\nu(s), \nu(0) \neq 0$ is an $r$-order almost common divisor of $P_{h+1,n}^*$ and its strength is defined as a solution of the following minimization problem:

$$f(P,P^*) = \min_{\forall P^*} \| \tilde{S}_{P^*} \Phi_{\nu} \|_F$$

where $P^* \in \Pi(n,p;h+1)$. Furthermore $\nu(s)$ is an $r$-order almost GCD of $P_{h+1,n}$ if the minimal corresponds to a coprime set $P^*$ or to full rank $S_{P^*}$. □

The optimization problem defining the strength of any order approximate GCD is now used to investigate the “best” amongst all approximate GCDs of a degree $r$. We consider polynomials $\nu(s), \nu(0) \neq 0$. Then $\hat{\Phi}_{\nu}$ is nonsingular and has fixed norm $\| \hat{\Phi}_{\nu} \|_F$ for all $\nu(s)$.

**Optimisation Problem**: This can be expressed as

$$f_1(P,P^*) \triangleq \| \hat{\Phi}_{\nu} \|_F \cdot f(P,P^*)$$

$$= \min_{\forall P^*} \{ \| \tilde{S}_P - [0, \tilde{S}_{P^*}^2] \hat{\Phi}_{\nu} \|_F \cdot \| \Phi_{\nu} \|_F \}$$

$$= \min_{\forall P^*} \| \tilde{S}_{P^*} \hat{\Phi}_{\nu} - [0, \tilde{S}_{P^*}^2] \|_F$$

where $P, \hat{\Phi}_{\nu}$ have the structure as previously defined and $\nu(s)$ has degree $r$.

We may summarise the main results as follows [1]:

**Theorem 9**: Consider the set of polynomials $P \in \Pi(n,p;h+1)$ and $S_P$ be its Sylvester matrix. Then the following hold true:

1. For a certain approximate GCD $\nu(s)$ of degree $k$, the perturbed set $\tilde{P}$ corresponding to minimal perturbation applied on $P$, such that $\nu(s)$ becomes an exact GCD, defined by:

$$S_{\tilde{P}} = \tilde{S}_P \hat{\Phi}_{\nu} = [0_k | \tilde{S}_{P}^2] \hat{\Phi}_{\nu}$$

2. The strength of an arbitrary $\nu(s)$ of degree $k$ is then given by:

$$f(P,P^*) = \min_{\forall P^*} \| S_{P^*} \hat{\Phi}_{\nu} \|_F$$
3. The optimal approximate GCD of degree $k$ is a $\varphi(s)$ defined by solving:

$$f(P, P^*) = \min_{\forall P: \deg \varphi(s) = k} \left\{ \| \tilde{S}_P^\varphi \Phi \|_F \right\}$$  \hspace{1cm} (33)

The computation of the optimal approximate GCD is equivalent to a non-convex optimization problem and can be defined as a structural singular value problem [26]. Numerical techniques linked to the computation of “approximate GCD” [1], [16] have been developed in [27]-[30].

5. Grassmann Invariants, approximate zero polynomials and distance problems

The characterisation of the “approximate GCD” and its “optimal” version provides the means to define the respective approximate zero polynomials for different classes of linear systems properties, which cover the cases:

(a) **Approximate zero polynomial** based on the coprime polynomial multi-vector $n(s)\wedge$.

(b) **Approximate input decoupling zero polynomial** based on the coprime polynomial multi-vector $b(s)\wedge$.

(c) **Approximate invariant input decoupling zero polynomial** based on the coprime polynomial multi-vector $r(s)\wedge$.

(d) **Approximate output decoupling zero polynomial** based on the coprime polynomial multi-vector $c(s)\wedge$.

(e) **Approximate invariant output decoupling zero polynomial** based on the coprime polynomial multi-vector $q(s)\wedge$.

Note that such polynomial multi-vectors have to satisfy the corresponding set of QPRs. It is that property that makes the computation of the approximate polynomials a more difficult problem. We shall develop the results for the case of a general polynomial matrix. We note first that an obvious corollary of Theorem 6 is:

**Corollary 3:** Let $\Pi(n, p; h + 1)$ be the set of all polynomial sets $P_{h+1,n}$ with $h + 1$ elements and with the two higher degrees $(n, p)$, $n \geq p$ and let $S_P$ be the Sylvester
resultant of the general set $P_{h+1,n}$. The variety of $P^{N-1}$ which characterise all sets $P_{h+1,n}$ having a GCD with degree $d$, $0 < d \leq p$ is defined by the set of equations

$$C_{n+p-d+1}(S_p) = 0$$

(34)

The above defines a variety $\Delta_d(n, p; h + 1)$ described by the polynomial equations in the coefficients of the vector $p_{h+1,n}$, or the point $P_{h+1,n}$ of $P^{N-1}$, and will be called the $d$-GCD variety of $P^{N-1}$. This characterises all sets in $\Pi(n, p; h + 1)$ with a GCD of degree $d$. The definition of the “optimal GCD” is thus a problem of finding the distance of a given set $P_{h+1,n}$ from the variety $\Delta_d(n, p; h + 1)$. For any $P_{h+1,n} \in \Pi(n, p; h + 1)$ this distance is defined by

$$d(P, \Delta) = \min_{\forall \phi^* \in \Phi} \|S_p - [0_k | \hat{S}p^*] \hat{\phi}\|_F$$

(35)

$\phi(s) \in \mathbb{R}[s]$, $P^* \in \Pi(n - k, p - k; h + 1)$, $\deg \phi(s) = k$, the $k$-distance of $P_{h+1,n}$ from the $k$-GCD variety $\Delta_k(n, p; h + 1)$ and $\phi(s)$ emerges as a solution to an optimisation problem and it is the $k$-optimal approximate GCD and the value $d(P, \Delta)$ is its $k$-strength.

For polynomial matrices we can extend the scalar definition of the approximate GCD as follows:

**Definition 3:** Consider the coprime polynomial matrix $T(s) \in \mathbb{R}^{q \times r}[s]$ and let $\Delta T(s) \in \mathbb{R}^{q \times r}[s]$ be an arbitrary matrix such that

$$T(s) + \Delta T(s) = \hat{T}(s) = \hat{T}(s)R(s)$$

(36)

where $R(s) \in \mathbb{R}^{r \times r}[s]$. Then $R(s)$ will be called an approximate matrix divisor of $T(s)$.

The above definition may be interpreted using exterior products as an extension of the problem defined for polynomial vector sets. The difference between general sets of vectors and those generated from polynomial matrices by taking exterior products is that the latter must satisfy the decomposability conditions [3] and in turn they define another variety of the Grassmann type.

Consider now the set of polynomial vectors $\Pi(n, p; h + 1)$ and let $\Pi^\Delta(n, p; h + 1)$ be its subset of the decomposable polynomial vectors $p(s) \in \mathbb{R}^{\sigma}[s]$, which correspond to the $q \times r$ polynomial matrices with degree $n$. The set $\Pi^\Delta(n, p; h + 1)$ is defined as the Grassmann variety $G(q, r; \mathbb{R}[s])$ of the projective space $P^{\sigma-1}(\mathbb{R}[s])$. The way we can extend the scalar results is based on:
(i) Parameterise the perturbations that move a general set \( P_{\sigma,n} \), to a set \( P'_{\sigma,n} = P_{\sigma,n} + Q_{\sigma,n} \in \Delta_k(n,p;\sigma) \) where initially \( Q_{\sigma,n} \) and \( P'_{\sigma,n} \) are free.

(ii) For the scalar results to be transferred back to the polynomial matrices the sets \( P_{\sigma,n} \) have to be decomposable multi-vectors which are denoted by \( \Pi^\wedge(n,p;\sigma) \). The latter set will be referred to as the \( n \)-order subset of the Grassmann variety \( G(q,r;\mathbb{R}[s]) \) and the sets \( P'_{\sigma,n} \) must be such that

\[
P'_{\sigma,n} \in \Pi(n,p;\sigma) \bigcap \Delta_k(n,p;\sigma) = \Delta^n_k \Pi(n,p;\sigma)
\]

where \( \Delta^n_k \Pi(n,p;\sigma) \) is the decomposable subset of \( \Delta_k(n,p;\sigma) \). Parameterising all sets \( P'_{\sigma,n} \) provides the means for posing a distance problem as before. This is clearly a constrained distance problem since now we have to consider the intersection variety defined by the corresponding set of QPRs and the equations of the GCD variety. Some preliminary results on this problem are stated below:

**Lemma 3:** The following properties hold true:

1. \( \Pi^\wedge(n,p,h+1) \) is proper subset \( \Pi(n,p,h+1) \) if \( r \neq 1 \) and \( q \neq r - 1 \).
2. \( \Pi^\wedge(n,p,h+1) = \Pi(n,p,h+1) \) if either \( r = 1 \) or \( q = r - 1 \).
3. The set \( \Delta^n_k \Pi(n,p;\sigma) \) is always nonempty. \( \square \)

The result is a direct implication of the decomposability conditions for multi-vectors \([3]\).

**Theorem 10:** Let \( P_{\sigma,n} \in \Pi^\wedge(n,p;\sigma) \) and denote by \( d(P,\Delta_k) \), \( d(P,\Delta_k^n) \) the distance from \( \Delta_k(n,p;\sigma) \) and \( \Delta_k^n \Pi(n,p;\sigma) \) respectively. The following hold true:

1. If \( q = r - 1 \) or \( r = 1 \), then the solutions of the two optimisation problems are identical and \( d(P,\Delta_k) = d(P,\Delta_k^n) \).
2. If \( q \neq r - 1 \) and \( r \neq 1 \), then \( d(P,\Delta_k) \leq d(P,\Delta_k^n) \). \( \square \)

**Remark 6:** The definition of the almost zero polynomials in the matrix case is clearly a distance problem. For polynomial matrices this distance problem is defined on the set \( P_{h+1,n} \) of \( \Pi(n,p,h+1) \) from the intersection of the varieties \( \Delta_d(n,p,h+1) \) and \( G(q,r;\mathbb{R}[s]) \).

The above suggests that the Grassmann distance problem has to be considered only when \( q \neq r - 1 \) and \( r \neq 1 \). The Grassmann distance problem requires the study of some additional topics linked to algebraic geometry and exterior algebra such as: (i) Parameterisation of all decomposable sets \( P \) with a fixed order \( n \); (ii) Characterisation of the set \( \Delta^n_k \Pi(n,p;\sigma) \) and its properties. The above issues are central for the solution of the GDP and are topics for further research. For the special case \( r = 1, q = r - 1 \) the distance \( d(P,\Delta_k) \) is reduced to that of the polynomial vector case since we guarantee decomposability and this leads to the definition of almost zeros and almost decoupling zeros.
6. Conclusions

The paper uses the recently introduced notion of “approximate GCD” of a set of polynomials [1] and the characterization of controllability and observability properties in terms of exterior products and associated Plücker matrices of controllability and observability pencils [12] to define distance from the set of uncontrollable, unobservable systems, as well as the corresponding approximate decoupling polynomials; furthermore, the use of the restriction pencils $R(s)$ and $Q(s)$ allows the definition of the distance of the state feedback, output injection orbits from uncontrollable, unobservable families respectively. The main distinctive feature of the approach, is the definition of distance of the orbits of systems from the uncontrollable, unobservable sets, as well as the definition of the approximate decoupling polynomials. The paper also extends the notion of approximate GCD of a set of polynomials to the case of approximate matrix divisors. It has been shown that this problem is equivalent to a distance problem from the intersection of two varieties and it is much harder than the polynomial vectors case. Our approach is based on the optimal approximate GCD and when this is applied to linear systems introduces new system invariants with significance in defining system properties under parameter variations on the corresponding model. The natural way for introducing such notions has been the notion of the polynomial Grassmann representative [4], which introduces new sets of polynomials. The na-
ture of the new distance problem stems from the fact that the polynomial vectors are exterior products of the columns of a polynomial matrix and thus the distance has to be considered from a sub-variety of the general $k$-GCD variety that is the intersection with the Grassmann variety. The optimization problem is non-convex and developing methodology for computing this distance is a problem of current research.

7. References


