KU LEUVEN

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Citation/Reference	Authors: Hans van Rooij, Bart De Moor Title: Numerical Elimination Theory via the Macaulay Matrix Journal: SIAM Journal on Matrix Analysis and Applications Pages: 20
Archived version	■ Submitted version ☐ Accepted (Author manuscript: the content is identical to the content of the published paper, but without the final typesetting by the publisher) ☐ Published ☐ Published (Open Access)
Acknowledgements	In this manuscript, the following funding is acknowledged: ■ EU H22020 ☐ EU Horizon Europe ■ FWO ☐ KULeuven Internal Funds ■ VLAIO ■ AI ☐ Other
Published version	N.A.
Copyright statement	©2026 van Rooij, H. De Moor, B. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
Author contact	hvanrooi@esat.kuleuven.be



NUMERICAL ELIMINATION THEORY VIA THE MACAULAY MATRIX*

HANS VAN ROOIJ† AND BART DE MOOR‡

Abstract. Two numerical approaches are presented to eliminate variables from a system of multivariate polynomial equations using the Macaulay matrix. Both approaches reformulate elimination as the identification of a structured subspace within the row space of the Macaulay matrix. The first approach constructs eliminant systems from left singular vectors of a submatrix of the Macaulay matrix, while the second approach uses principal angles and directions between subspaces, interpreting elimination in terms of the cosine–sine decomposition. The developed methods allow the elimination of an arbitrary subset of variables. For polynomial systems where the solution set is zero-dimensional, elimination results in a univariate eliminant polynomial or a reduced system in the remaining variables. Alternatively, when the solution set is positive-dimensional, the same methods produce implicit polynomial relations. Several numerical examples illustrate the application potential of our numerical elimination methodology, such as computing the critical value polynomial in multivariate polynomial optimization and parametric optimization problems with positive-dimensional solution sets.

Key words. numerical elimination theory, Macaulay matrix, singular value decomposition, principal angles and directions, cosine-sine decomposition, multivariate polynomial optimization

MSC codes. 15A03, 15B05, 15A18, 15A23

1. Introduction.

This paper presents a numerical elimination method for polynomial equations using the Macaulay matrix. Two approaches based on the singular value decomposition are discussed for eliminating variables from polynomial equations to reveal relationships among the remaining variables. The first uses left singular vectors of a submatrix of the Macaulay matrix constructed from the given polynomials. The second extracts information from principal vectors and principal angles between subspaces. The principal vectors correspond to zero principal angles between the row space of the Macaulay matrix and an elimination subspace. We show that the latter method is inherently connected to the cosine–sine decomposition, offering an interesting geometric interpretation of numerical elimination theory. Additionally, principal angles provide a quantitative algebraic perspective on numerical elimination theory. The concept of elimination can be understood through a simple example.

Example 1.1. Consider the following system of bivariate polynomials:

$$\begin{cases} f_1(x_1, x_2) = x_1^4 + x_2^4 - 1.4\,(x_1^2 + x_2^2) + 0.5 = 0, \\ f_2(x_1, x_2) = (x_1^2 + x_2^2)^2 - 2\,(x_1^2 - x_2^2) = 0. \end{cases} \tag{1.1}$$

Figure 1.1 shows the roots of the multivariate polynomials defined by (1.1). Their intersection points correspond to the common zeros of both polynomials. For this example, elimination aims to describe these intersection points in terms of x_1 by

*Submitted to the editors DATE.

Funding: This work was supported in part by the KU Leuven Research Fund (IDN/24/013, iBOF/23/064), the Industrial Research Fund, and several LRD bilateral industrial projects; in part by FWO (TBM T001725N; PhD grant 11K5623N); in part by EWI via the Flanders AI Research Program; in part by VLAIO (HBC.2021.0076); in part by FOD VVVL (AI-943); and in part by the ERC under the EU Horizon 2020 programme (Advanced Grant 885682).

†hans.vanrooij@kuleuven.be, Dept. of Electrical Engineering (ESAT–STADIUS), KU Leuven.

‡bart.demoor@kuleuven.be, Dept. of Electrical Engineering (ESAT–STADIUS), KU Leuven. Fellow IEEE, SIAM, and IFAC.

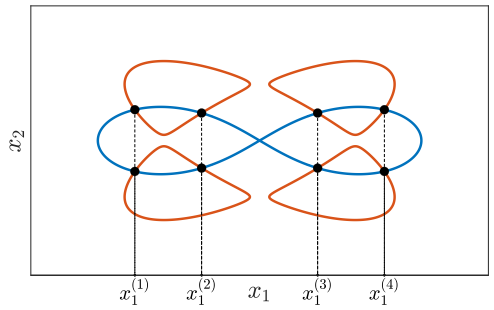

Fig. 1.1: The curves $f_1(x_1, x_2) = 0$ (■) and $f_2(x_1, x_2) = 0$ (■) with their intersection points (•). Elimination can be seen as projecting this intersection set onto the coordinate axis corresponding to the variable kept, here x_1.

removing x_2 from the system. Geometrically, this corresponds to projecting the intersection set onto the x_1-axis as seen on Figure 1.1. Algebraically, elimination boils down to combining $f_1(x_1, x_2)$ and $f_2(x_1, x_2)$, where the coefficients are multivariate polynomials, to obtain a polynomial in fewer variables, as shown in (1.2).

(1.2)
$$\begin{aligned} f_{\text{elim}}(x_1) &= \left(0.415 + 0.762x_1^2 + 0.080x_2^2 + 0.170x_1^4 - 0.057x_2^4\right) f_1(x_1, x_2) + \\ &\quad \left(0.270 + 0.046x_1^2 - 0.273x_2^2 + 0.057x_1^4 - 0.114x_1^2x_2^2 + 0.057x_2^4\right) f_2(x_1, x_2) \\ &= 0.255x_1^8 + 0.510x_1^6 - 0.433x_1^4 - 0.831x_1^2 + 0.232 = 0. \end{aligned}$$

The idea of eliminating variables from a system of polynomial equations traces back to the 18th century where Bézout's theorem provided early methods for understanding the structure of algebraic systems by counting the number of solutions of a multivariate polynomial system. In the 19th century, J. Sylvester introduced the resultant, a determinant-based method which helps detect when polynomials share common roots. Gröbner bases were introduced in the 20th-century. This method can be seen as a generalization of Gaussian elimination and it simplifies polynomial systems and helps elimination theory by systematically removing variables. An elaborate historical overview is provided in [1]. The numerical formulation of elimination theory is relatively recent. Batselier et al. [2] develop this viewpoint using the Macaulay matrix as is done in this paper as well. Zeng [3], on the other hand, derives elimination matrices from the condition that the eliminant has zero partial derivative with respect to the eliminated variables.

The remainder of this paper is organized as follows. Section 2 introduces the Macaulay matrix and some of its properties. Section 3 develops a numerical framework for elimination theory and presents two SVD-based approaches: one based on left singular vectors of an adapted Macaulay matrix, and another based on principal angles and vectors between subspaces, leading to a geometric interpretation via the cosine–sine decomposition. We also show that extracting information from the sine of the principal angles is more reliable than using their cosine counterparts. Section 4 illustrates numerical elimination theory through the introduction of the critical value polynomial (CVP). The roots of this univariate polynomial align with the critical values of a multivariate polynomial optimization problem. Additionally, Section 4 shows that numerical elimination theory can be used to find positive-dimensional solution sets, as demonstrated through parametric polynomial optimization problems.

Finally, Section 5 concludes the paper and outlines directions for future research.

2. Preliminary Material. Both our approaches to eliminate variables from a system of polynomial equations use the Macaulay matrix. The following example shows how this type of matrix naturally arises in the elimination problem and thus motivates its use. After the example, the Macaulay matrix will be defined more generally, and some of its properties that are useful for our numerical elimination approaches will be highlighted.

Example 1.1 *(continued).* Rewriting (1.2) as an inner product between a vector of scalar coefficients and a vector collecting the functions $f_1(x_1, x_2)$, $f_2(x_1, x_2)$ together with their monomial multiples, we obtain

$$f_{\text{elim}}(x_1) = \begin{bmatrix} 0.415 & 0.762 & \cdots & -0.114 & 0.057 \end{bmatrix} \begin{bmatrix} f_1(x_1, x_2) \\ x_1^2 f_1(x_1, x_2) \\ \vdots \\ x_1^2 x_2^2 f_2(x_1, x_2) \\ x_2^4 f_2(x_1, x_2) \end{bmatrix} = 0.$$

In general, neither the coefficients nor the monomial multiples (or shifts) of $f_1(x_1, x_2)$ and $f_2(x_1, x_2)$ used to eliminate x_1 are known in advance. Thus, the elimination problem in (1.1) amounts to finding a linear combination of monomial multiples of $f_1(x_1, x_2)$ and $f_2(x_1, x_2)$ that removes x_1:

$$f_{\text{elim}}(x_1) = \underbrace{\begin{bmatrix} c_1 & c_2 & \cdots & c_i & c_{i+1} & \cdots \end{bmatrix}}_{\boldsymbol{c}^T} \begin{bmatrix} f_1(x_1, x_2) \\ x_1 f_1(x_1, x_2) \\ \vdots \\ f_2(x_1, x_2) \\ x_1 f_2(x_1, x_2) \\ \vdots \end{bmatrix} = 0.$$

This formulation motivates the use of the Macaulay matrix $\boldsymbol{M}_d$ of degree d, which encodes the coefficients of all monomial multiples of the initial polynomials $f_1(x_1, x_2)$ and $f_2(x_1, x_2)$ into a quasi-Toeplitz structured matrix. This is visualized as

$$\begin{bmatrix} f_1(x_1, x_2) \\ x_1\, f_1(x_1, x_2) \\ x_2\, f_1(x_1, x_2) \\ x_1^2\, f_1(x_1, x_2) \\ \vdots \\ f_2(x_1, x_2) \\ x_1\, f_2(x_1, x_2) \\ x_2\, f_2(x_1, x_2) \\ x_1^2\, f_2(x_1, x_2) \\ \vdots \end{bmatrix} = \underbrace{\begin{array}{c}\textbf{Macaulay Matrix} \\ \includegraphics[height=3.5cm]{macaulay_matrix.pdf}\end{array}}_{\boldsymbol{M}_d} \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1 x_2 \\ x_2^2 \\ x_1^3 \\ \vdots \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \end{bmatrix}.$$

More general, the Macaulay matrix [4, 5, 6] is a useful tool to study systems of multivariate polynomial equations and their solutions. In this paper, the monomials are ordered in degree negative lexicographic order [6].

Example 2.1 (Degree negative lexicographic order in three variables). The variables x_1, x_2 and x_3 of maximal degree two in degree negative lexicographic order are ordered as

$$\underbrace{1}_{\deg=0} < \underbrace{x_1 < x_2 < x_3}_{\deg=1} < \underbrace{x_1^2 < x_1x_2 < x_1x_3 < x_2^2 < x_2x_3 < x_3^2}_{\deg=2}$$

Extending the Macaulay matrix is accomplished by applying the forward shift recursion (FSR) to the initial equations, which means multiplying (shifting) the original polynomial system by all monomials up to a certain degree (this degree will be specified in Definition 2.3).

Example 2.2 (Constructing the Macaulay matrix). We build the Macaulay matrix for the system

$$\begin{cases} f_1(\boldsymbol{x}) = x_1^2 - x_2 = 0, \\ f_2(\boldsymbol{x}) = x_1 + x_2 = 0. \end{cases}$$

The degrees of the polynomials are 2 and 1, so the Macaulay matrix of degree $d = 2$, denoted $\boldsymbol{M}_2$, is constructed from all monomials of total degree up to 2. Writing f_1 and f_2 in this monomial basis yields the initial rows

$$\boldsymbol{M}_{2,\text{init}} = \begin{array}{c} \\ f_1(\boldsymbol{x}) \\ f_2(\boldsymbol{x}) \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 \end{array} \\ \begin{bmatrix} 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix} \end{array}.$$

To complete $\boldsymbol{M}_2$, we include the shifted equations $x_1 f_2(\boldsymbol{x})$ and $x_2 f_2(\boldsymbol{x})$, which contribute the rows corresponding to the monomials of total degree 2. This leads to

$$\boldsymbol{M}_2 = \begin{array}{c} \\ f_1(\boldsymbol{x}) \\ f_2(\boldsymbol{x}) \\ x_1 f_2(\boldsymbol{x}) \\ x_2 f_2(\boldsymbol{x}) \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 \end{array} \\ \begin{bmatrix} 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \end{array}.$$

Definition 2.3 (Macaulay matrix). *The Macaulay matrix $\boldsymbol{M}_d$ of degree d is built from the coefficients of a system of multivariate polynomials denoted as $\mathcal{S}$*

$$\mathcal{S} = \begin{cases} f_1(\boldsymbol{x}) &= 0, \\ f_2(\boldsymbol{x}) &= 0, \\ &\vdots \\ f_s(\boldsymbol{x}) &= 0, \end{cases} \tag{2.1}$$

of degree $d_1, \ldots, d_s$, respectively with $f_i(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}]$ for $i = 1, \ldots, s$ and $\boldsymbol{x} \in \mathbb{R}^n$ in a structured manner, where the columns are indexed by all monomials up to degree d in degree negative lexicographic order, and the rows are indexed by all forward shift recursions $\{\boldsymbol{x}^{\boldsymbol{\alpha}_i}\} f_i(\boldsymbol{x})$ where $\boldsymbol{\alpha}_i = (\alpha_{i,1}, \ldots, \alpha_{i,n})$ is an n-tuple and $\boldsymbol{x}^{\boldsymbol{\alpha}_i} = \prod_{j=1}^{n} x_j^{\alpha_{i,j}}$

with $\deg(\boldsymbol{x}^{\boldsymbol{\alpha}_i}) \leqslant d - d_i$ *for* $i = 1, \ldots, s$, *where* $\deg(\bullet)$ *denotes the total degree of* $\bullet$.

$$\boldsymbol{M}_d = \begin{bmatrix} \{\boldsymbol{x}^{\boldsymbol{\alpha}_1}\}\boldsymbol{f_1} \\ \vdots \\ \{\boldsymbol{x}^{\boldsymbol{\alpha}_s}\}\boldsymbol{f_s} \end{bmatrix}, \tag{2.2}$$

where $\{\boldsymbol{x}^{\boldsymbol{\alpha}_i}\}\boldsymbol{f_i}$ *represents the coefficients of* $\{\boldsymbol{x}^{\boldsymbol{\alpha}_i}\}f_i(\boldsymbol{x})$.

This way, the system of multivariate polynomial equations together with their respective monomial shifts can be represented as

$$\boldsymbol{M}_d \boldsymbol{\phi}_d = 0,$$

where $\boldsymbol{\phi}_d$ represents a multivariate Vandermonde vector:

$$\boldsymbol{\phi}_d = \begin{bmatrix} 1 & x_1 & x_2 & \cdots & x_n & x_1^2 & x_1 x_2 & \cdots & x_n^d \end{bmatrix}^T. \tag{2.3}$$

The size of the vector $\boldsymbol{x}$ as well as the degree of the Macaulay matrix, the number of equations in the system of multivariate polynomial equations and its corresponding total degrees influence the size of $\boldsymbol{M}_d$:

THEOREM 2.4 (Dimensions of the Macaulay matrix $\boldsymbol{M}_d$ [7]). *The number of rows* p_d *and the number of columns* q_d *of* $\boldsymbol{M}_d$ *can be represented using binomial coefficients:*

$$p_d = \sum_{i=1}^{s} \binom{d - d_i + n}{n} = \sum_{i=1}^{s} \frac{(d - d_i + n)!}{n!(d - d_i)!} \quad \textit{and} \quad q_d = \binom{d+n}{n} = \frac{(d+n)!}{n!d!}.$$

Furthermore, p_d *and* q_d *can also be expressed as a polynomial in* d, *given the number of variables* n *[8]:*

$$p_d = \frac{s}{n!} d^n + \mathcal{O}(d^{n-1}) \quad \textit{and} \quad q_d = \frac{1}{n!} d^n + \mathcal{O}(d^{n-1}).$$

The number of solutions to (2.1) can be described by its Bézout number, which provides a connection between the algebraic properties of the system of multivariate polynomials (their degrees) and the geometric properties (the number of intersection points) in case the solution set is zero-dimensional. This includes the m_a affine solutions along with the m_∞ solutions at infinity.

THEOREM 2.5 (Theorem of Bézout). *The Bézout number, denoted as* m_b, *represents the number of isolated complex solutions (counting multiplicities) that* (2.1) *has in the complex projective space for* $s = n$ *(square system) [9]. It can be calculated as follows:*

$$m_b = m_a + m_\infty = \prod_{i=1}^{n} \deg(f_i(\boldsymbol{x})),$$

Throughout this text, different subspaces of the Macaulay matrix $\boldsymbol{M}_d$ will be of interest. The fundamental theorem of linear algebra reveals that each real matrix has four fundamental subspaces, being the column space, row space, left null space and right null space [10]. From (2.2), it is clear that each row of the Macaulay matrix represents a multivariate polynomial $\boldsymbol{x}^{\boldsymbol{\alpha}_i} f_i(\boldsymbol{x})$. This results in the inclusion of the following multivariate polynomials in the row space of $\boldsymbol{M}_d$ [8]:

$$\mathcal{M}_d = \operatorname{row}(\boldsymbol{M}_d) = \left\{ \sum_{i=1}^{s} h_i(\boldsymbol{x}) f_i(\boldsymbol{x}) : h_i(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}] \wedge \deg(h_i(\boldsymbol{x})) \leqslant d - d_i \right\}.$$

3. Numerical Elimination Theory. This section presents the machinery underlying our numerical elimination procedure based on the Macaulay matrix. Before introducing the two SVD-based approaches, we first review the basic idea of numerical elimination via the Macaulay matrix and state a condition under which such elimination is possible in subsection 3.1.

On the one hand, subsection 3.2 describes a method for numerically eliminating variables from a system of polynomial equations using the left singular vectors of a submatrix of the Macaulay matrix. On the other hand, subsection 3.3 presents an alternative approach that uses the right singular vectors of the full Macaulay matrix to identify a positive-dimensional intersection between two subspaces. This is accomplished by computing the principal angles and directions between two subspaces. Using this latter method, we find an interesting interpretation of numerical elimination theory through the cosine–sine decomposition.

3.1. Introduction To Numerical Elimination. We will call the result of the numerical elimination procedure the eliminant system. This eliminant system can be a single polynomial, in which case we will refer to it as the eliminant polynomial. The goal now is to identify a vector space within the row space $\mathcal{M}_d$ of the Macaulay matrix $\boldsymbol{M}_d$ such that every polynomial in that vector space has a predetermined zero pattern. This vector space will only exhibit non-zero entries corresponding to the monomials selected by this zero pattern.

Example 3.1 (Predetermined zero pattern for $n = 2$). In case $\boldsymbol{x} = (x_1, x_2)^T \in \mathbb{R}^2$ and we want to eliminate x_2, the predetermined zero pattern to be found in $\mathcal{M}_d$ is

$$\boldsymbol{\xi}^T = \begin{bmatrix} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 & x_1^3 & x_1^2x_2 & x_1x_2^2 & x_2^3 & \cdots \\ \bullet & \bullet & 0 & \bullet & 0 & 0 & \bullet & 0 & 0 & 0 & \cdots \end{bmatrix}, \tag{3.1}$$

where $\bullet$ stands for any arbitrary real number.

Example 3.2 (Predetermined zero pattern for $n = 3$). In case $\boldsymbol{x} = (x_1, x_2, x_3)^T \in \mathbb{R}^3$ and we want to eliminate x_3, the predetermined zero pattern to be found in $\mathcal{M}_d$ is

$$\boldsymbol{\xi}^T = \begin{bmatrix} 1 & x_1 & x_2 & x_3 & x_1^2 & x_1x_2 & x_1x_3 & x_2^2 & x_2x_3 & x_3^2 & \cdots \\ \bullet & \bullet & \bullet & 0 & \bullet & \bullet & 0 & \bullet & 0 & 0 & \cdots \end{bmatrix},$$

where $\bullet$ stands for any arbitrary real number.

Denote the index set $\mathcal{I}$ as the set containing the indices corresponding to the monomials that are kept after the elimination procedure in the multivariate Vandermonde vector represented in (2.3). Now, define $\boldsymbol{E}_d \in \mathbb{R}^{k \times q_d}$ as a submatrix of the identity matrix only including the rows which are in $\mathcal{I}$:

$$\boldsymbol{E}_d = \boldsymbol{I}_{q_d}(\mathcal{I}, :) = \begin{bmatrix} \boldsymbol{I}_k & \boldsymbol{0} \end{bmatrix} \boldsymbol{P},$$

where $\boldsymbol{P}$ represents a column-wise permutation matrix and k is the number of monomials in $\boldsymbol{\xi}$ not forced to equal zero.

Example 3.1 *(continued).* The zero pattern vector $\boldsymbol{\xi}$ in (3.1) can be parametrized

as

$$\boldsymbol{\xi}^T = \begin{bmatrix} \bullet & \bullet & \bullet & \cdots \end{bmatrix} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}}_{\boldsymbol{E}_d}.$$

This way, the columns of $\boldsymbol{E}_d^T$ span an orthonormal basis encoding the predetermined zero pattern, called the elimination vector space, which we will denote as $\mathcal{E}_d$. As a consequence, finding the eliminant system means looking for a non-trivial intersection of $\mathcal{M}_d$ and $\mathcal{E}_d$. For this to happen, the dimension t of the intersection between these two subspaces should be positive:

$$t = \dim(\mathcal{M}_d \cap \mathcal{E}_d) > 0. \tag{3.2}$$

The numerical elimination procedure via the Macaulay matrix can be tackled by a variety of methods. This paper will distinguish two approaches. One of them uses the left singular vectors of an adapted Macaulay matrix, while the other method provides a geometric interpretation of numerical elimination theory by relying on principal angles and directions.

3.2. Numerical Elimination Theory via Left Singular Vectors. The condition in (3.2), together with Graßmann's dimension theorem [11], leads to a simple rank test for the existence of an eliminant system. Recall that for any two subspaces $\mathcal{U}, \mathcal{V} \subseteq \mathbb{R}^n$,

$$\dim(\mathcal{U} \cap \mathcal{V}) = \dim(\mathcal{U}) + \dim(\mathcal{V}) - \dim(\mathcal{U} + \mathcal{V}).$$

Applying this relation to the subspaces $\mathcal{M}_d$ and $\mathcal{E}_d$ gives

$$\dim(\mathcal{M}_d \cap \mathcal{E}_d) = \operatorname{rank}(\boldsymbol{M}_d) + \operatorname{rank}(\boldsymbol{E}_d) - \operatorname{rank}\begin{bmatrix} \boldsymbol{E}_d \\ \boldsymbol{M}_d \end{bmatrix}. \tag{3.3}$$

The requirement for this dimension to be positive guarantees a non-trivial intersection, and hence the existence of an eliminant system. Computing the dimension of the intersection is done by isolating the columns of $\boldsymbol{M}_d$ that correspond to the monomials not selected by the predetermined zero pattern. The resulting matrix is

$$\bar{\boldsymbol{M}}_d = \boldsymbol{M}_d(:, \mathcal{I}'),$$

where $\mathcal{I}'$ is the complement of $\mathcal{I}$.

Example 1.1 *(continued).* Figure 3.1 depicts the visual difference between the Macaulay matrix $\boldsymbol{M}_d$ and its filtered version $\bar{\boldsymbol{M}}_d \in \mathbb{R}^{p_d \times (q_d - k)}$.

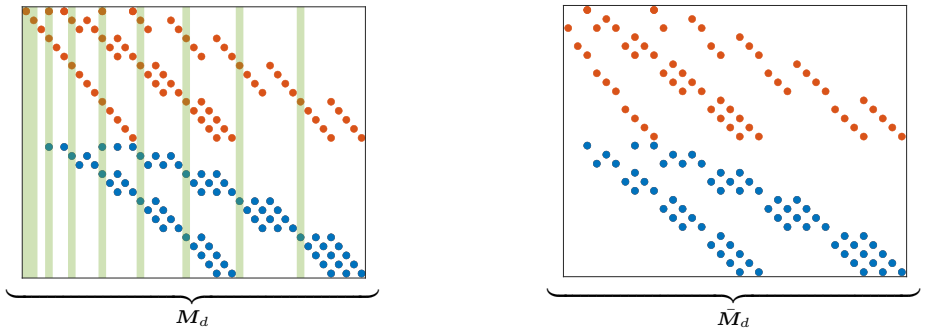

Fig. 3.1: Comparison between the Macaulay matrix $\boldsymbol{M}_d$ and its filtered version $\bar{\boldsymbol{M}}_d$, illustrating the preserved structure after removing the columns in $\mathcal{I}$, indicated by the green bands ■.

With this notation, (3.3) is rewritten as

$$\operatorname{rank}\begin{bmatrix}\boldsymbol{E}_d\\ \boldsymbol{M}_d\end{bmatrix} = \operatorname{rank}(\boldsymbol{E}_d) + \operatorname{rank}(\bar{\boldsymbol{M}}_d),$$

resulting in

$$\begin{aligned}\dim(\mathcal{M}_d \cap \mathcal{E}_d) &= \operatorname{rank}(\boldsymbol{M}_d) + \operatorname{rank}(\boldsymbol{E}_d) - (\operatorname{rank}(\boldsymbol{E}_d) + \operatorname{rank}(\bar{\boldsymbol{M}}_d))\\ &= \operatorname{rank}(\boldsymbol{M}_d) - \operatorname{rank}(\bar{\boldsymbol{M}}_d).\end{aligned} \tag{3.4}$$

Thus, an eliminant system exists when $\operatorname{rank}(\boldsymbol{M}_d) > \operatorname{rank}(\bar{\boldsymbol{M}}_d)$. Using the system in (1.1), we observe a rank drop for $d = 8$, and the corresponding matrices are those that were visualized in Figure 3.1.

To derive the eliminant system, we first construct a matrix $\Upsilon = \begin{bmatrix}\upsilon_1 & \dots & \upsilon_{l_d}\end{bmatrix} \in \mathbb{R}^{p_d \times l_d}$ whose columns form a basis for the left null space of $\bar{\boldsymbol{M}}_d$, where l_d denotes the dimension of this null space:

$$\Upsilon^T \bar{\boldsymbol{M}}_d = 0. \tag{3.5}$$

Example 1.1 *(continued).* Applied to system (1.1), this gives

$$\underbrace{\begin{bmatrix}0 & 0 & 0 & 0.445 & \dots\\ 0.415 & 0 & 0 & 0.762 & \dots\end{bmatrix}}_{\Upsilon^T}
\begin{bmatrix}\cdot & \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & \cdot\end{bmatrix}
\underbrace{\qquad\qquad}_{\bar{M}_d}
= \begin{bmatrix}0 & 0 & 0 & 0 & \cdots\\ 0 & 0 & 0 & 0 & \cdots\end{bmatrix}.$$

Finding such a matrix Υ using (3.5) ensures the correct zero-pattern to emerge. This can then be used to obtain a corresponding matrix $\Xi = \begin{bmatrix}\xi_1 & \dots & \xi_{l_d}\end{bmatrix} \in \mathbb{R}^{q_d \times l_d}$ satisfying the predetermined zero-pattern:

$$\Xi^T = \Upsilon^T \boldsymbol{M}_d.$$

Example 1.1 *(continued).* For the system (1.1), we obtain

$$\underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & \dots \\ 0.232 & 0 & 0 & -0.831 & \cdots \end{bmatrix}}_{\Xi^T} = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0.445 & \dots \\ 0.415 & 0 & 0 & 0.762 & \cdots \end{bmatrix}}_{\Upsilon^T} \underbrace{\quad}_{M_d}.$$

In the current setting, $\boldsymbol{\xi}$ can contain columns equal to the zero vector as can be seen in the example above. To avoid this, numerical elimination proceeds with the columns of Υ that do not lie in the left null space of $\boldsymbol{M}_d$. So finally, the eliminant system is constructed as

$$\mathcal{S}_{elim} = \Xi_{\setminus 0}^T \phi_d = 0,$$

where $\Xi_{\setminus 0}$ is the matrix $\boldsymbol{\xi}$ without the zero columns and $\mathcal{S}_{elim}$ represents the eliminant system containing the t multivariate polynomial equations only in the remaining variables.

Example 1.1 *(continued).* Applied to example (1.1):

$$\mathcal{S}_{elim} = \xi_2^T \phi_8 = 0.255x_1^8 + 0.510x_1^6 - 0.433x_1^4 - 0.831x_1^2 + 0.232 = 0. \tag{3.6}$$

The procedure outlined in this section is summarized in Algorithm 3.1, where the dimension of the intersection given by (3.4) is used as a criterion for determining the eliminant system $\mathcal{S}_{elim}$. Since the degree of the Macaulay matrix at which this non-trivial intersection first appears is not known a priori, the initial degree d is chosen as the maximum degree among the given polynomials $f_1(\boldsymbol{x}), \dots, f_s(\boldsymbol{x})$. The degree is then increased iteratively until a non-trivial intersection is obtained.

3.3. Numerical Elimination Theory via Right Singular Vectors. The previous approach relied on the left singular vectors of an adapted Macaulay matrix. In this subsection, we develop an alternative formulation based on the right singular vectors of the full Macaulay matrix. This perspective views numerical elimination as the problem of detecting a non-trivial intersection between two subspaces: the row space of the Macaulay matrix $\mathcal{M}_d$ and the elimination subspace $\mathcal{E}_d$. The existence of such an intersection can be characterized in terms of the principal angles and principal directions between these subspaces, providing a geometric interpretation of the elimination process.

We first recall the basic notions of principal angles and directions in subsection 3.3.1 and describe how they are computed via the singular value decomposition. Subsequently, in subsection 3.3.2, we show how this framework naturally leads to the cosine–sine decomposition, which offers additional insight into the structure of numerical elimination.

3.3.1. Principal Angles and Directions. The idea is simple: if two subspaces intersect non-trivially, then some of their principal angles equal zero. Conversely, being orthogonal to a subspace results in a principal angle of $\pi/2$. Both scenarios can be exploited to detect eliminant systems numerically. Below, we briefly recall the

Algorithm 3.1 Numerical Elimination theory via Left Singular Vectors

Input: polynomials $f_1(\boldsymbol{x}), \ldots, f_s(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}]$, set of variables you want to eliminate
Output: $\mathcal{S}_{elim}$

$d \leftarrow \max(\deg(f_1(\boldsymbol{x})), \ldots, \deg(f_s(\boldsymbol{x}))$
$\mathcal{S}_{elim} \leftarrow \varnothing$
while $\mathcal{S}_{elim} = \varnothing$ **do**
  $\boldsymbol{M}_d \leftarrow$ Macaulay matrix of degree d
  $\boldsymbol{E}_d^T \leftarrow$ basis for $\mathcal{E}_d$ from $\mathcal{I}$
  if $t = \dim(\mathcal{M}_d \cap \mathcal{E}_d) = \text{rank}(\boldsymbol{M}_d) - \text{rank}(\bar{\boldsymbol{M}}_d) > 0$ **then**
    $\bar{\boldsymbol{M}}_d \leftarrow \boldsymbol{M}_d(:, \mathcal{I}')$
    $\Upsilon \leftarrow$ numerical basis for the left null space of $\bar{\boldsymbol{M}}_d$
    $\Xi^T \leftarrow \Upsilon^T \boldsymbol{M}_d$
    $\mathcal{S}_{elim} \leftarrow \Xi_{\setminus 0}^T \phi_d = 0$
  else
    $d \leftarrow d + 1$
  end if
end while

definition of principal angles and directions before applying them to the elimination problem.

DEFINITION 3.3 (Principal Angles). *The principal angles $\theta_i \in [0, \pi/2]$ for $i = 1, \ldots, \min(d_1, d_2)$ between the subspaces $\mathcal{S}_1$ and $\mathcal{S}_2$ of dimension d_1 and d_2 can be found recursively as [12]*

$$\cos\theta_i = \max_{\boldsymbol{u}\in\mathcal{S}_1, \boldsymbol{v}\in\mathcal{S}_2} \boldsymbol{u}^T\boldsymbol{v} = \boldsymbol{u}_i^T\boldsymbol{v}_i, \qquad ||\boldsymbol{u}||_2 = 1, ||\boldsymbol{v}||_2 = 1$$
$$\text{subject to} \quad \boldsymbol{u}^T\boldsymbol{u}_j = 0, \text{ for } j = 1, \ldots, i-1,$$
$$\boldsymbol{v}^T\boldsymbol{v}_j = 0, \text{ for } j = 1, \ldots, i-1.$$

THEOREM 3.4 (Principal angles and directions via the SVD). *Let $\boldsymbol{Q}_a \in \mathbb{R}^{k\times p}$ and $\boldsymbol{Q}_b \in \mathbb{R}^{k\times q}$ have orthonormal columns. Define*

$$\boldsymbol{A} = \boldsymbol{Q}_a^T\boldsymbol{Q}_b \in \mathbb{R}^{p\times q},$$

and let the singular value decomposition be

$$\boldsymbol{A} = \boldsymbol{Y}\,\Sigma\,\boldsymbol{Z}^T, \qquad \Sigma = \text{diag}(\sigma_1, \ldots, \sigma_r), \quad r = \min\{p, q\},$$

with $\sigma_1 \geqslant \cdots \geqslant \sigma_r \geqslant 0$ and $\boldsymbol{Y}^T\boldsymbol{Y} = \boldsymbol{Z}^T\boldsymbol{Z} = I_r$. Then the principal angles $\theta_i \in [0, \pi/2]$ are given by

$$\cos\theta_i = \sigma_i, \qquad i = 1, \ldots, r,$$

and the corresponding principal directions are

$$\boldsymbol{u}_i = \boldsymbol{Q}_a\boldsymbol{y}_i, \qquad \boldsymbol{v}_i = \boldsymbol{Q}_b\boldsymbol{z}_i,$$

where $\boldsymbol{y}_i$ and $\boldsymbol{z}_i$ are the i-th columns of $\boldsymbol{Y}$ and $\boldsymbol{Z}$.

It is worth noting that in practice, the computation of principal angles and directions between two subspaces involves three singular value decompositions in total: two to find the orthonormal bases $\boldsymbol{Q}_a$ and $\boldsymbol{Q}_b$ of the subspaces $\mathcal{S}_1$ and $\mathcal{S}_2$, and one additional SVD of $\boldsymbol{A} = \boldsymbol{Q}_a^T \boldsymbol{Q}_b$ to extract the principal angles and directions themselves.

3.3.2. Numerical Elimination via the Cosine-Sine Decomposition. To identify eliminant systems, one can inspect the intersection between the row space of the Macaulay matrix $\mathcal{M}_d$ and the elimination subspace $\mathcal{E}_d$. According to Theorem 3.4, a non-trivial intersection is present when at least one principal angle between these subspaces equals zero, and thus when its cosine equals 1. Consider the SVD of $\boldsymbol{M}_d$:

$$\boldsymbol{M}_d = \begin{bmatrix} \boldsymbol{U}_1 & \boldsymbol{U}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_1^T \\ \boldsymbol{V}_2^T \end{bmatrix} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T.$$

The right singular vectors are collected in the orthogonal matrix $\boldsymbol{V} \in \mathbb{R}^{q_d \times q_d}$. If r_M denotes the rank of the Macaulay matrix, then the columns of $\boldsymbol{V}_1 \in \mathbb{R}^{q_d \times r_M}$ form a basis for $\mathcal{M}_d$, while the columns of $\boldsymbol{V}_2 \in \mathbb{R}^{q_d \times (q_d - r_M)}$ span its null space. We assume $k < r_M$ and $q_d - k \geqslant r_M$. Furthermore, recall that the columns of $\boldsymbol{E}_d^T$ span an orthonormal basis for the elimination subspace $\mathcal{E}_d$. The SVD of

$$\boldsymbol{E}_d \boldsymbol{V}_1 = \begin{bmatrix} \boldsymbol{I}_k & \boldsymbol{0} \end{bmatrix} \boldsymbol{P}\boldsymbol{V}_1 = \begin{bmatrix} \boldsymbol{I}_k & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_1(\mathcal{I}, :) \\ \boldsymbol{V}_1(\mathcal{I}', :) \end{bmatrix} = \boldsymbol{V}_1(\mathcal{I}, :) = \boldsymbol{R}_{11}\boldsymbol{S}_{11}\boldsymbol{W}_{11}^T$$

has as singular values the cosines of the principal angles θ_i, $i = 1, \ldots, k$, between $\mathcal{M}_d$ and $\mathcal{E}_d$, where

$$\boldsymbol{S}_{11} = \begin{bmatrix} \boldsymbol{C} & 0 \end{bmatrix}, \qquad \boldsymbol{C} = \operatorname{diag}\big(\cos\theta_1, \ldots, \cos\theta_k\big). \tag{3.7}$$

Integrating angular information between $\mathcal{M}_d$ and $\mathcal{E}_d$ in the numerical elimination framework is very powerful as it offers a quantitative way of looking at the intersection requirement to eliminate. In contrast to symbolic elimination methods from algebraic geometry, we can track the alignment between $\mathcal{M}_d$ and $\mathcal{E}_d$ as a function of the degree d of the Macaulay matrix by looking at the principal angles.

Example 1.1 *(continued).* Table 3.1 depicts how increasing the degree d of the Macaulay matrix for (1.1) when eliminating x_2 leads to a tighter geometric alignment between $\mathcal{M}_d$ and $\mathcal{E}_d$, as shown by the principal angles that decrease monotonically.

Table 3.1: Principal angle θ_1 between $\mathcal{M}_d$ and $\mathcal{E}_d$ for increasing Macaulay matrix degree d for (1.1) when eliminating x_2, obtained from the diagonal elements of $\boldsymbol{S}_{11}$. Smaller angles correspond to stronger subspace alignment and at $d = 8$ we notice a non-trivial intersection due to $\theta_1 = 0^\circ$.

d	4	5	6	7	8
θ_1	32.716°	32.716°	0.985°	0.985°	0°

The t principal vectors associated with the zero principal angles are obtained from Theorem 3.4 as

$$\Xi = \begin{bmatrix} \xi_1 & \ldots & \xi_t \end{bmatrix} = \boldsymbol{E}_d^T \boldsymbol{R}_{11}(:, 1\!:\!t).$$

while the resulting eliminant system is

$$\mathcal{S}_{elim} = \Xi^T \phi_d = 0.$$

Example 1.1 *(continued).* Applied to (1.1) this gives

$$\Xi^T = \xi_1^T = \begin{bmatrix} 0.207 & 0 & 0 & -0.741 & \dots \end{bmatrix}$$

and the eliminant system is

$$\mathcal{S}_{elim} = \xi_1^T \phi_8 = 0.227x_1^8 + 0.455x_1^6 - 0.386x_1^4 - 0.741x_1^2 + 0.207 = 0. \tag{3.8}$$

Notice that the same monomials appear as in (3.6), but the coefficients differ slightly. This difference arises because the coefficients derived from the principal vectors are normalized, whereas the previous approach does not enforce such normalization.

However, as noted by [12], determining θ_i from $\cos\theta_i$ becomes ill-conditioned for small values of θ_i. Instead of studying the principal angles between $\mathcal{M}_d$ and $\mathcal{E}_d$, one may consider the subspace orthogonal to $\mathcal{M}_d$, namely $\mathcal{Z}_d = \text{null}(\boldsymbol{M}_d) = \text{col}(\boldsymbol{V}_2)$. In this setting, an eliminant system appears when a principal angle between $\mathcal{Z}_d$ and $\mathcal{E}_d$ equals $\pi/2$. This information can be obtained from

$$\boldsymbol{E}_d \boldsymbol{V}_2 = \begin{bmatrix} \boldsymbol{I}_k & \boldsymbol{0} \end{bmatrix} \boldsymbol{P} \boldsymbol{V}_2 = \begin{bmatrix} \boldsymbol{I}_k & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_2(\mathcal{I}, :) \\ \boldsymbol{V}_2(\mathcal{I}', :) \end{bmatrix} = \boldsymbol{V}_2(\mathcal{I}, :) = \boldsymbol{R}_{12} \boldsymbol{S}_{12} \boldsymbol{W}_{12}^T.$$

The singular values of $\boldsymbol{S}_{12}$ equal the cosines of the principal angles ψ_i between $\mathcal{Z}_d$ and $\mathcal{E}_d$, with $i = 1, \dots, k$, where

$$\boldsymbol{S}_{12} = \begin{bmatrix} \boldsymbol{S} & 0 \end{bmatrix}, \qquad \boldsymbol{S} = \text{diag}\big(\cos\psi_1, \dots, \cos\psi_k\big). \tag{3.9}$$

This decomposition can be interpreted in two equivalent ways: the diagonal entries of $\boldsymbol{S}$ correspond either to the cosines of the principal angles between $\mathcal{Z}_d$ and $\mathcal{E}_d$, or to the sines of the principal angles θ_i between $\mathcal{M}_d$ and $\mathcal{E}_d$. In particular, if t singular values in $\boldsymbol{S}_{12}$ equal zero, then $\mathcal{Z}_d$ and $\mathcal{E}_d$ have t orthogonal principal vector pairs, and hence $\mathcal{M}_d$ and $\mathcal{E}_d$ intersect in a subspace of dimension t.

Example 1.1 *(continued).* Table 3.2 shows the evolution of the angles between $\mathcal{Z}_d$ and $\mathcal{E}_d$ for (1.1) as the degree d increases. An orthogonal component between these subspaces appears at $d = 8$.

Table 3.2: Principal angle θ_{end} between $\mathcal{Z}_d$ and $\mathcal{E}_d$ for increasing Macaulay matrix degree d for (1.1) when eliminating x_2, obtained from the diagonal elements of $\boldsymbol{S}_{12}$. Bigger angles correspond to stronger orthogonal subspace alignment and at $d = 8$ we notice an orthogonal component between the two subspaces due to $\theta_r = 90°$.

d	4	5	6	7	8
θ_r	57.284°	57.284°	89.015°	89.015°	90°

The t principal vectors associated with the principal angles equal to $\pi/2$ are obtained from Theorem 3.4 as

$$\Xi = \begin{bmatrix} \xi_1 & \dots & \xi_t \end{bmatrix} = \boldsymbol{E}_d^T \boldsymbol{R}_{12}(:, k - t + 1 : k).$$

while the resulting eliminant system is again

$$\mathcal{S}_{elim} = \Xi^T \phi_d = 0.$$

Example 1.1 *(continued).* Applied to (1.1) this gives

$$\Xi^T = \xi_1^T = \begin{bmatrix} 0.207 & 0 & 0 & -0.741 & \dots \end{bmatrix}$$

and the eliminant system is

$$\mathcal{S}_{elim} = \xi_1^T \phi_8 = 0.227x_1^8 + 0.455x_1^6 - 0.386x_1^4 - 0.741x_1^2 + 0.207 = 0,$$

where you notice that up to three digits after the decimal, we get the same eliminant polynomial as in (3.8).

Interestingly, the cosines and the sines of the principal angles can also be retrieved as the singular values of the following SVD's respectively:

$$\boldsymbol{E}_d^{'}\boldsymbol{V}_1 = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{I}_{q_d-k} \end{bmatrix} \boldsymbol{P}\boldsymbol{V}_1 = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{I}_{q_d-k} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_1(\mathcal{I},:) \\ \boldsymbol{V}_1(\mathcal{I}^{'},:) \end{bmatrix} = \boldsymbol{V}_1(\mathcal{I}^{'},:) = \boldsymbol{R}_{21}\boldsymbol{S}_{21}\boldsymbol{W}_{21}^T,$$

$$\boldsymbol{E}_d^{'}\boldsymbol{V}_2 = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{I}_{q_d-k} \end{bmatrix} \boldsymbol{P}\boldsymbol{V}_2 = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{I}_{q_d-k} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_2(\mathcal{I},:) \\ \boldsymbol{V}_2(\mathcal{I}^{'},:) \end{bmatrix} = \boldsymbol{V}_2(\mathcal{I}^{'},:) = \boldsymbol{R}_{22}\boldsymbol{S}_{22}\boldsymbol{W}_{22}^T,$$

where

$$\boldsymbol{S}_{21} = \begin{bmatrix} \boldsymbol{I}_{r_M-k} & 0 \\ 0 & \boldsymbol{S} \\ 0 & 0 \end{bmatrix}, \qquad \boldsymbol{S}_{22} = \begin{bmatrix} \boldsymbol{I}_{q_d-r_M-k} & 0 \\ 0 & \boldsymbol{C} \\ 0 & 0 \end{bmatrix}$$

and $\boldsymbol{C}$, $\boldsymbol{S}$ the same as in (3.7) and (3.9). This naturally leads to the cosine-sine decomposition (CSD) of the 2×2 orthogonal block matrix $\boldsymbol{PV}$, again assuming $k < r_M$ and $q_d - k \geqslant r_M$:

$$\begin{array}{c} \\ k \\ q_d-k \end{array} \begin{array}{c} \begin{array}{cc} r_M & q_d-r_M \end{array} \\ \begin{bmatrix} \boldsymbol{V}_1(\mathcal{I},:) & \boldsymbol{V}_2(\mathcal{I},:) \\ \boldsymbol{V}_1(\mathcal{I}^{'},:) & \boldsymbol{V}_2(\mathcal{I}^{'},:) \end{bmatrix} \end{array} = \begin{bmatrix} \boldsymbol{P}_1 & 0 \\ 0 & \boldsymbol{P}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{D}_{11} & \boldsymbol{D}_{12} \\ \boldsymbol{D}_{21} & \boldsymbol{D}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{Q}_1^T & 0 \\ 0 & \boldsymbol{Q}_2^T \end{bmatrix}, \tag{3.10}$$

where

$$\begin{bmatrix} \boldsymbol{D}_{11} & \boldsymbol{D}_{12} \\ \boldsymbol{D}_{21} & \boldsymbol{D}_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} \boldsymbol{C} & 0 & \hat{\boldsymbol{S}} & 0 \\ \hline 0 & \boldsymbol{I}_{r_M-k} & 0 & 0 \\ -\hat{\boldsymbol{S}} & 0 & \boldsymbol{C} & 0 \\ 0 & 0 & 0 & \boldsymbol{I}_{q_d-k-r_M} \end{array}\right].$$

The hat notation is used to emphasize that the diagonal matrix $\hat{\boldsymbol{S}}$ contains the same sine values of the principal angles between $\mathcal{M}_d$ and $\mathcal{E}_d$ as $\boldsymbol{S}$, but now appearing in increasing order instead of decreasing.

$$\hat{\boldsymbol{S}} = \operatorname{diag}\big(\cos\psi_r, \dots, \cos\psi_1\big) = \operatorname{diag}\big(\sin\theta_r, \dots, \sin\theta_1\big).$$

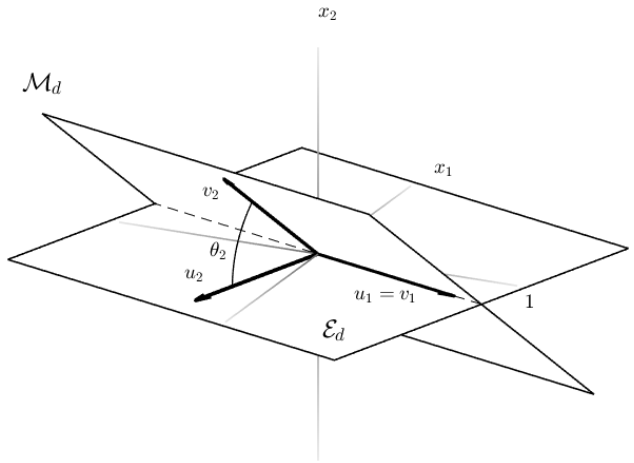

Fig. 3.2: Geometric visualization of the numerical elimination procedure for the case where the variable x_2 is eliminated. The subspace $\mathcal{M}_d$ represents the row space of the Macaulay matrix, while $\mathcal{E}_d$ denotes the subspace of monomials that do not involve x_2. The principal direction $u_1 = v_1$ indicates how $\mathcal{M}_d$ intersects the subspace $\mathcal{E}_d$, while the principal directions u_2 and v_2 correspond to the principal angle θ_2.

This results in the following equality:

$$\boldsymbol{C}^2 + \hat{\boldsymbol{S}}^2 = \boldsymbol{I}_k.$$

In [13], the authors mention the following about the CSD: "This powerful theoretical tool is worth trying whenever it might even remotely be applicable; it brings such clarity and simplicity.". Let us dissect this quote.

The simplicity arises because $\boldsymbol{PV}$ is orthogonal, which forces the singular values of the blocks to be geometrically linked. As a result, only four singular vector matrices are needed instead of eight.

The clarity comes from the geometric interpretation it provides in the numerical elimination setting. From (3.10) we see that numerical elimination corresponds to identifying the intersections of $\mathcal{M}_d$ with the coordinate axes labelled by monomials not to be eliminated and thus the subspace $\mathcal{E}_d$. Figure 3.2 illustrates this viewpoint for the case where the variable being eliminated is x_2.

The numerical elimination procedure based on right singular vectors is summarized in Algorithm 3.2. As in the left singular vector approach, the degree d of the Macaulay matrix is initialized as the maximum degree among the input polynomials and increased iteratively until an eliminant system $\mathcal{S}_{elim}$ is obtained. For each degree d, elimination is detected by inspecting the sine values of the principal angles between the null space of the Macaulay matrix and the elimination subspace, computed from the singular value decomposition of $\boldsymbol{E}_d\boldsymbol{V}_2$, which is numerically preferable for small principal angles compared to the cosine variant. For conciseness, only this principal angle based formulation is presented.

4. Numerical Examples. Numerical elimination theory has several applications in multivariate polynomial optimization. In this paper, we present two applica-

Algorithm 3.2 Numerical Elimination Theory via Right Singular Vectors

Input: polynomials $f_1(\boldsymbol{x}), \ldots, f_s(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}]$, set of variables you want to eliminate
Output: $\mathcal{S}_{elim}$

$d \leftarrow \max(\deg(f_1), \ldots, \deg(f_s))$
$\mathcal{S}_{elim} \leftarrow \varnothing$
while $\mathcal{S}_{elim} = \varnothing$ **do**
  $\boldsymbol{M}_d \leftarrow$ Macaulay matrix of degree d
  $\boldsymbol{E}_d^T \leftarrow$ basis for $\mathcal{E}_d$ from $\mathcal{I}$
  $\boldsymbol{M}_d = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T, \quad \boldsymbol{V}_2 \leftarrow \boldsymbol{V}(:, r_M + 1 : q_d)$
  $\boldsymbol{E}_d\boldsymbol{V}_2 = \boldsymbol{R}_{12}\boldsymbol{S}_{12}\boldsymbol{W}_{12}^T$ $\{\sigma_i(\boldsymbol{S}_{12}) = \sin\theta_i\}$
  $t \leftarrow \#\{i : 0 \leqslant \sin\theta_i \leqslant 0 + \tau\}$ {tolerance $\tau > 0$}
  if $t > 0$ **then**
    $\Xi \leftarrow \boldsymbol{E}_d^T\boldsymbol{R}_{12}(:, k - t + 1 : k)$
    $\mathcal{S}_{elim} \leftarrow \Xi^T\phi_d = 0$
  else
    $d \leftarrow d + 1$
  end if
end while

tions. The first concerns the computation of the critical value polynomial, obtained by eliminating the decision variables and Lagrange multipliers from the KKT system. The second application addresses parametric polynomial optimization problems, where elimination is used to derive implicit relations between the decision variables and the parameters in positive-dimensional settings.

4.1. The Critical Value Polynomial. In the context of multivariate polynomial optimization problems (MPOPs), one is often interested not only in the critical points that satisfy the Karush–Kuhn–Tucker (KKT) conditions, but also in the corresponding objective values attained at these points.

DEFINITION 4.1 (Critical points and critical values [14]). *Let $J(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}]$ denote the objective function of a polynomial optimization problem. All points $\boldsymbol{x}^* \in \mathbb{R}^n$ for which $\nabla_{\boldsymbol{x}} J(\boldsymbol{x}^*) = 0$ are defined as critical points, and $J(\boldsymbol{x}^*)$ represents the corresponding critical values.*

To encode these critical values algebraically, we augment the KKT system with an additional polynomial equation that links the objective function to a new scalar variable σ. This leads to the following system of polynomial equations:

$$\begin{cases} \nabla_{\boldsymbol{x}}\mathcal{L}(\boldsymbol{x}, \Lambda) &= 0 &= \nabla_{\boldsymbol{x}} J(\boldsymbol{x}) + \nabla_{\boldsymbol{x}}\mathbf{g}(\boldsymbol{x})^{\mathrm{T}}\Lambda, \\ \nabla_{\boldsymbol{\Lambda}}\mathcal{L}(\boldsymbol{x}, \Lambda) &= 0 &= \mathbf{g}(\boldsymbol{x}), \\ \sigma - J(\boldsymbol{x}) &= 0 & . \end{cases}$$

where $\mathbf{g}(\boldsymbol{x}) = [g_1(\boldsymbol{x}), \ldots, g_{n_g}(\boldsymbol{x})]^T$ are the equality constraints[1]. Here, the variable σ captures the objective value corresponding to each solution of the KKT equations. By eliminating the variables $\boldsymbol{x}$ and the multipliers Λ, one obtains a univariate polynomial

[1]Inequality constraints $\mathbf{h}(\boldsymbol{x}) = [h_1(\boldsymbol{x}), \ldots, h_{n_h}(\boldsymbol{x})]^T$, with $h_i(\boldsymbol{x}) \geqslant 0 \in \mathbb{R}[\boldsymbol{x}]$ can be incorporated as well by introducing slack variables $\mathbf{s} = [s_1, \ldots, s_{n_h}]^T$. This way, the inequality constraints can be transformed into equality constraints as $h_i(\boldsymbol{x}) - \mu_i^2 = 0$, for $i = 1, \ldots, n_h$.

in σ whose roots are precisely the critical values of the optimization problem. This polynomial is referred to as the critical value polynomial.

DEFINITION 4.2 (Critical Value Polynomial). *The critical value polynomial or CVP is a univariate polynomial in σ. Suppose the objective function has r critical values $\sigma_1, \ldots, \sigma_r$, then the CVP is defined as*

$$\sigma_{\text{CVP}} = \prod_{i=1}^{r} (\sigma - \sigma_i).$$

It is interesting to note that for unconstrained optimization problems, the CVP seems to result in a lowest degree eliminant polynomial; each critical value only appears once as a root of the CVP even if it is obtained at multiple critical points. The CVP of an unconstrained MPOP is a square-free univariate polynomial, meaning that no repeated roots are present. Consequently the following statement holds:

$$\text{GCD}(\sigma_{\text{CVP}}, \sigma'_{\text{CVP}}) = 1,$$

with $\text{GCD}(p(\boldsymbol{x}), q(\boldsymbol{x}))$ the greatest common divisor between $p(\boldsymbol{x})$ and $q(\boldsymbol{x})$ and σ'_{CVP} the derivative of σ_{CVP} with respect to σ.

Moreover, the set of critical points can even be positive dimensional, as example Example 4.4 will illustrate, but the number of critical values will always be finite:

THEOREM 4.3 (Sard's Theorem [15, 16]). *Let $J(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}]$. The set of critical values from the smooth function $J(\boldsymbol{x})$ forms a set of Lebesgue measure zero. The ideal $< \nabla_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}, \Lambda), \nabla_{\Lambda} \mathcal{L}(\boldsymbol{x}, \Lambda), \sigma - J(\boldsymbol{x}) >$ contains the univariate CVP.*

Example 4.4 (Infinitely many critical points). Given the following unconstrained optimization problem

$$\underset{\boldsymbol{x}}{\text{minimize}}\ J(\boldsymbol{x}) = x_1^4 + x_2^4 - 4x_1^2 - 4x_2^2 + 2x_1^2 x_2^2 + 15. \tag{4.1}$$

The set of critical points of this objective function is positive dimensional. This can be seen by deriving the FONC for optimality of (4.1):

$$\nabla_{\boldsymbol{x}} J(\boldsymbol{x}) = \boldsymbol{0} \iff \begin{cases} \frac{\partial J(\boldsymbol{x})}{\partial x_1} = 0 = 4x_1^3 + 4x_1 x_2^2 - 8x_1 = 4x_1(x_1^2 + x_2^2 - 2), \\ \frac{\partial J(\boldsymbol{x})}{\partial x_2} = 0 = 4x_2^2 + 4x_1^2 x_2 - 8x_2 = 4x_2(x_1^2 + x_2^2 - 2), \end{cases} \tag{4.2}$$

where the common factor $(x_1^2 + x_2^2 - 2)$ causes the positive dimensionality. The CVP can be obtained if we change (4.2) to:

$$\begin{cases} \frac{\partial J(\boldsymbol{x})}{\partial x_1} = 0 = 4x_1^3 + 4x_1 x_2^2 - 8x_1 = 4x_1(x_1^2 + x_2^2 - 2), \\ \frac{\partial J(\boldsymbol{x})}{\partial x_2} = 0 = 4x_2^2 + 4x_1^2 x_2 - 8x_2 = 4x_2(x_1^2 + x_2^2 - 2), \\ \sigma - J(\boldsymbol{x}) = 0 = \sigma - (x_1^4 + x_2^4 - 4x_1^2 - 4x_2^2 + 2x_1^2 x_2^2 + 15). \end{cases} \tag{4.3}$$

The evolution of the principal angles between $\mathcal{Z}_d$ and $\mathcal{E}_d$ when eliminating x_1 and x_2 is shown in Table 4.1.

Table 4.1: Principal angle ψ_k between $\mathcal{Z}_d$ and $\mathcal{E}_d$ for increasing Macaulay matrix degree d for (4.3) when eliminating x_1 and x_2. At $d = 6$ we find an orthogonal complement between $\mathcal{Z}_d$ and $\mathcal{E}_d$ due to $\psi_k = 90°$.

d	4	5	6
ψ_k	82.978°	83.198°	90°

The principal vector in $\mathcal{E}_d$ that corresponds to this principal angle at $d = 6$ gives the CVP (after normalizing the coefficient of σ^2 to one):

$$\sigma_{CVP} = \sigma^2 - 26\sigma + 165 = 0.$$

The roots of this CVP are $\{11, 15\}$. The critical value 11 is obtained for all critical points on the circle $x_1^2 + x_2^2 - 2 = 0$, but only shows up as a single root of the CVP satisfying Theorem 4.3.

Example 4.5 (Minimization of a quartic on the sphere). The CVP can also be computed for multivariate polynomial optimization problems subject to constraints. Consider minimizing the trivariate quartic subject to a sphere.

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & J(x_1, x_2, y) = x_1^4 + x_2^4 + x_3^4 \\ \text{subject to} \quad & x_1^2 + x_2^2 + x_3^2 - 1 = 0. \end{aligned} \tag{4.4}$$

The Lagrange function is

$$\mathcal{L}(\boldsymbol{x}, \lambda_1) = x_1^4 + x_2^4 + x_3^4 + \lambda_1(x_1^2 + x_2^2 + x_3^2 - 1).$$

The CVP and the critical values of (4.4) can be found by eliminating $\boldsymbol{x}$ and λ_1 from

$$\begin{cases} \nabla_{x_1}\mathcal{L}(\boldsymbol{x}, \lambda_1) & = 0 = 4x_1^3 + 2\lambda_1 x_1, \\ \nabla_{x_2}\mathcal{L}(\boldsymbol{x}, \lambda_1) & = 0 = 4x_2^3 + 2\lambda_1 x_2, \\ \nabla_{x_3}\mathcal{L}(\boldsymbol{x}, \lambda_1) & = 0 = 4x_3^3 + 2\lambda_1 x_3, \\ \nabla_{\lambda_1}\mathcal{L}(\boldsymbol{x}, \lambda_1) & = 0 = x_1^2 + x_2^2 + x_3^2 - 1, \\ \sigma - J(\boldsymbol{x}) & = 0 = \sigma - (x_1^4 + x_2^4 + x_3^4). \end{cases} \tag{4.5}$$

The evolution of the principal angles between $\mathcal{M}_d$ and $\mathcal{E}_d$ is shown in Table 4.2.

Table 4.2: Principal angle θ_1 between $\mathcal{M}_d$ and $\mathcal{E}_d$ for increasing Macaulay matrix degree d for (4.5) when eliminating $\boldsymbol{x}$ and λ_1. At $d = 8$ we notice a non-trivial intersection due to $\theta_1 = 0^\circ$.

d	4	5	6	7	8
θ_1	22.068°	10.504°	5.192°	2.589°	0°

Finding the principal vector in $\mathcal{E}_d$ corresponding to the zero principal angle results in the CVP:

$$\sigma_{CVP} = 0.431\sigma^3 - 0.790\sigma^2 + 0.431\sigma - 0.072 = 0,$$

and hence the CVP roots (critical values) are $\{0.333, 0.500, 1.000\}$.

4.2. Parametric Polynomial Optimization. Many optimization problems involve parameters and can be posed as parametric polynomial optimization problems (see Appendix A). The goal is to determine the optimal critical points $\boldsymbol{x}$, to evaluate the corresponding optimal cost, and to recover the associated Lagrange multipliers $\boldsymbol{\lambda}$, all expressed parametrically with respect to $\boldsymbol{y}$. The difficulty in parametric optimization lies in the fact that the solution set is positive-dimensional. Consequently, the dimension of $\mathcal{Z}_d$ (the nullity of the corresponding Macaulay matrix) does not stabilize. As hinted in Example 4.4, numerical elimination methods are able to deal with positive-dimensionalities in their solution set.

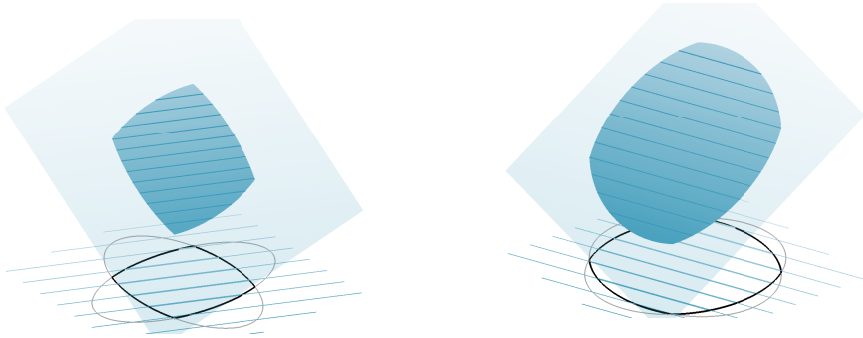

Fig. 4.1: Visualization of the parametric optimization problem for $y = 0.3$ (left) and $y = 0.7$ (right). The bilinear objective function forms a plane ■, while the gray boundaries of the inequality constraints define two elliptic cylinders. The dark blue region ■ indicates the portion of the objective function restricted to the feasible set.

Example 4.6 (Bilinear cost function constrained within two elliptic cylinders). Consider the parametric optimization problem of a bilinear cost function constrained by two elliptic cylinders ([17], p. 230):

$$\begin{aligned} &\underset{\boldsymbol{x}}{\text{minimize}} && J(\boldsymbol{x}, y) = yx_1 + (1-y)x_2, \\ &\text{subject to} && yx_1^2 + x_2^2 - y \leqslant 0 \;\; \rightarrow \;\; yx_1^2 + x_2^2 - y + \mu_1^2 = 0, \\ &&& x_1^2 + yx_2^2 - y \leqslant 0 \;\; \rightarrow \;\; x_1^2 + yx_2^2 - y + \mu_2^2 = 0. \end{aligned} \tag{4.6}$$

The optimization landscape depends on the parameter y and Figure 4.1 visualizes this dependency for parameter values $y = 0.3$ and $y = 0.7$.

Finding the parametric critical points means eliminating $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$ and either x_1 or x_2. Table 4.3 shows the evolution of the largest principal angle between $\mathcal{Z}_d$ and $\mathcal{E}_d$.

Table 4.3: Principal angle ψ_k between $\mathcal{Z}_d$ and $\mathcal{E}_d$ for increasing Macaulay matrix degree d for (4.5) when eliminating x_2 (or x_1), $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$. At $d = 12$, $\mathcal{Z}_d$ and $\mathcal{E}_d$ meet orthogonal and thus an eliminant parametric polynomial relating x_1 (or x_2) and y can be found.

d	3	4	5	6	7
ψ_k	52.372°	58.085°	67.786°	71.752°	81.516°

d	8	9	10	11	12
ψ_k	83.759°	87.454°	88.566°	89.383°	90.000°

The parametric eliminant equations $p(x_1, y) = 0$ and $p(x_2, y) = 0$ are depicted in Figure 4.2. Once these polynomials are obtained, a specific parameter value y can be substituted to yield a univariate polynomial whose roots correspond to the critical points. To illustrate this procedure, a dashed line is included in Figure 4.2 for the parameter value $y = 0.8$. The projections of the intersection points of this line onto

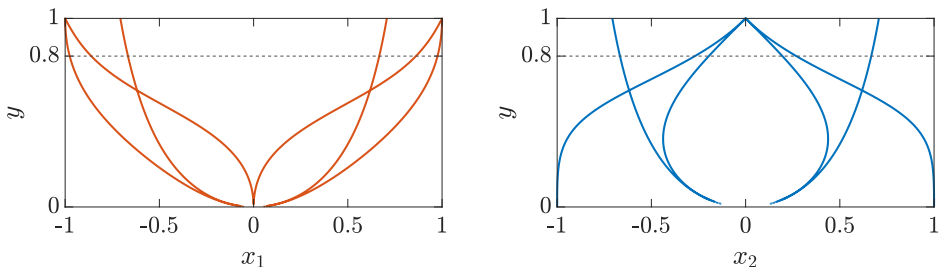

Fig. 4.2: Plots of the eliminant polynomials $p_{x_1}(x_1, y) = 0$ (■) and $p_{x_2}(x_2, y) = 0$ (■) obtained from the parametric optimization problem. Each curve shows the dependence of the critical points x_1 and x_2 on the parameter y, illustrating how these variables can be expressed as implicit functions of y. The dashed line marks a particular parameter value used for evaluation.

Table 4.4: For the parameter value $y = 0.8$, the table lists the solutions corresponding to the critical points x_1 and x_2, the associated Lagrange multipliers λ_1 and λ_2, and the critical values σ. The critical points x_1 and x_2 are highlighted in bold.

λ_1	λ_2	$\mathbf{x_1}$	$\mathbf{x_2}$	σ
$\mp$1.7500	$\pm$2.0000	$\mp$**0.6667**	$\pm$**0.6667**	$\mp$0.4000
$\pm$0.5123	$\pm$0.0000	$\mp$**0.9759**	$\mp$**0.1952**	$\mp$0.8198
$\pm$0.5123	$\pm$0.0000	$\mp$**0.9759**	$\mp$**0.1952**	$\mp$0.8198
$\pm$0.0000	$\pm$0.4644	$\mp$**0.8614**	$\mp$**0.2692**	$\mp$0.7430
$\pm$0.0000	$\pm$0.4644	$\mp$**0.8614**	$\mp$**0.2692**	$\mp$0.7430
$\mp$0.9167	$\pm$1.3333	$\mp$**0.6667**	$\mp$**0.6667**	$\mp$0.6667

the x_1- and x_2-axes represent the critical points corresponding to this parameter value. The numerical values of these solutions are listed in Table 4.4, obtained by solving the complete system in (4.6) with $y = 0.8$. In this case, the system becomes zero-dimensional, providing a direct verification of the parametric eliminant polynomials.

5. Conclusions and Future Work. This work introduced a numerical framework for elimination theory based on the Macaulay matrix. By reformulating variable elimination in systems of multivariate polynomials through tools from numerical linear algebra, the proposed approach differs from traditional symbolic methods.

Two formulations were developed. The first relies on the left singular vectors of an adapted Macaulay matrix, while the second provides a geometric perspective by expressing the elimination process in terms of principal angles between subspaces. In this context, the cosine–sine decomposition arises naturally and enables a quantitative algebraic geometric interpretation, as the evolution of these principal angles can be tracked as a function of the Macaulay matrix degree d.

Importantly, the framework extends to systems with positive-dimensional solution sets, as demonstrated through a parametric optimization example.

Future work will focus on improving computational efficiency by exploiting the fact that only the t relevant principal angles and vectors (corresponding to the dimension of the intersection) need to be computed, rather than the full set. In addition, the methodology presented in this paper can be extended to block Macaulay matrices.

Appendix A. Parametric Multivariate Polynomial Optimization. Consider the following parametric multivariate polynomial optimization problem:

$$\begin{aligned} &\underset{\boldsymbol{x}}{\text{minimize}}\ J(\boldsymbol{x},\boldsymbol{y}), \quad \boldsymbol{x}\in\mathbb{R}^n,\ \boldsymbol{y}\in\mathcal{Y}\subset\mathbb{R}^p, \\ &\text{subject to}\quad g_i(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\mu})=0, \quad i=1,\ldots,n_g,\ \boldsymbol{\mu}\in\mathbb{R}^{n_h}. \end{aligned}$$

Here, $J(\boldsymbol{x},\boldsymbol{y})\in\mathbb{R}[\boldsymbol{x},\boldsymbol{y}]$ denotes the polynomial cost function, and the equality constraints are collected in $\mathbf{g}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\mu})=[g_1,\ldots,g_{n_g}]^{\mathrm{T}}$. The variables $\boldsymbol{x}$ represent the decision variables, $\boldsymbol{\mu}$ are the slack variables, $\boldsymbol{y}$ are the parameters, and $\mathcal{Y}$ is the parameter set. Introducing Lagrange multipliers $\boldsymbol{\lambda}=[\lambda_1,\ldots,\lambda_{n_g}]^{\mathrm{T}}\in\mathbb{R}^{n_g}$, the Lagrangian $\mathcal{L}(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{y}):\mathbb{R}^{n+p+n_g+n_h}\to\mathbb{R}$ is given by

$$\mathcal{L}(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{y})=J(\boldsymbol{x},\boldsymbol{y})+\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{g}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\mu}).$$

The first-order optimality conditions form a structured polynomial system in $(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu})$, depending parametrically on $\boldsymbol{y}$:

$$\begin{cases} \nabla_{\boldsymbol{x}}\mathcal{L}(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{y})=0, \\ \nabla_{\boldsymbol{\lambda}}\mathcal{L}(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{y})=0, \\ \nabla_{\boldsymbol{\mu}}\mathcal{L}(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{y})=0. \end{cases}$$

REFERENCES

[1] D. Cox, *Applications of polynomial systems*. American Mathematical Soc., 2020, vol. 134.
[2] K. Batselier, P. Dreesen, and B. De Moor, "A geometrical approach to finding multivariate approximate lcms and gcds," *Linear Algebra and its Applications*, vol. 438, no. 9, pp. 3618–3628, 2013.
[3] Z. Zeng, "A numerical elimination method for polynomial computations," *Theoretical Computer Science*, vol. 409, no. 2, pp. 318–331, 2008.
[4] F. S. Macaulay, "Some Formulæ in Elimination," *Proc. of the London Mathematical Society*, vol. s1-35, no. 1, pp. 3–27, 1902.
[5] ——, *The Algebraic Theory of Modular Systems*. Cambridge University Press, 1994.
[6] P. Dreesen, "Back to the roots - polynomial system solving using linear algebra," Ph.D. dissertation, KU Leuven (Leuven, Belgium), 2013.
[7] C. Vermeersch, "The (block) Macaulay matrix: Solving systems of multivariate polynomial equations and multiparameter eigenvalue problems," Ph.D. dissertation, KU Leuven (Leuven, Belgium), 2023.
[8] K. Batselier, P. Dreesen, and B. De Moor, "On the null spaces of the Macaulay matrix," *Linear Algebra and its Applications*, vol. 460, pp. 259–289, 2014.
[9] E. Bézout, *General theory of algebraic equations*. Princeton University Press, Princeton, NJ, 2006, translated from the 1779 French original by Eric Feron.
[10] G. Strang, "The fundamental theorem of linear algebra," *The American Mathematical Monthly*, vol. 100, no. 9, pp. 848–855, 1993.
[11] H. Grassmann, *Die Ausdchnungslehre Von 1844, Oder Die Lineale Ausdehnungslehre: Ein Neuer Zweig Der Mathematik, Da*. BiblioBazaar, 2009.
[12] Åke Björck and G. H. Golub, "Numerical methods for computing angles between linear subspaces," *Mathematics of Computation*, vol. 27, no. 123, pp. 579–594, 1973.
[13] C. C. Paige and M. Wei, "History and generality of the CS decomposition," *Linear Algebra and its Applications*, vol. 208, pp. 303–326, 1994.
[14] J. W. Milnor, *Morse theory*. Princeton university press, 1963, no. 51.
[15] B. Hanzon and A. Mustata, "On the role of the critical value polynomial in algebraic optimization." 2014.
[16] J. M. Lee, "Sard's theorem," *Introduction to smooth manifolds*, pp. 125–149, 2012.
[17] J. B. Lasserre, *An introduction to polynomial and semi-algebraic optimization*. Cambridge University Press, 2015, vol. 52.