A Column Space Based Approach to Solve Systems of Multivariate Polynomial Equations ⋆

Christof Vermeersch * **Bart De Moor**, *Fellow, IEEE & SIAM* *

* *Center for Dynamical Systems, Signal Processing, and Data Analytics (STADIUS), Dept. of Electrical Engineering (ESAT), KU Leuven, Kasteelpark Arenberg 10, 3001 Leuven, Belgium (e-mail: {christof.vermeersch,bart.demoor}@esat.kuleuven.be)*

Abstract: We propose a novel approach to solve systems of multivariate polynomial equations, using the column space of the Macaulay matrix that is constructed from the coefficients of these polynomials. A multidimensional realization problem in the null space of the Macaulay matrix results in an eigenvalue problem, the eigenvalues and eigenvectors of which yield the common roots of the system. Since this null space based algorithm uses well-established numerical linear algebra tools, like the singular value and eigenvalue decomposition, it finds the solutions within machine precision. In this paper, on the other hand, we determine a complementary approach to solve systems of multivariate polynomial equations, which considers the column space of the Macaulay matrix instead of its null space. This approach works directly on the data in the Macaulay matrix, which is sparse and structured. We provide a numerical example to illustrate our new approach and to compare it with the existing null space based root-finding algorithm.

Keywords: Macaulay matrix, multivariate polynomials, numerical algorithms, realization theory, matrix algebra.

1. INTRODUCTION

Determining the roots of a multivariate polynomial, or the common roots of a system of multivariate polynomials, is one of the oldest problems in mathematics (Pan, 1997; Cox et al., 2004). Multivariate polynomial system solving arises in a myriad of applications in science and engineering, e.g., computational biology, machine learning, systems and control, and computer vision. It has a long and rich history that can be traced back to the Ancient Near East. For example, the Babylonians and Egyptians (about 2000 B.C.) already solved linear and quadratic equations by methods similar to those we use today (Pan, 1997).

Within the area of mathematics, algebraic geometry studies multivariate polynomial equations and algebraic varieties, i.e., the geometrical objects defined by the zero sets of these polynomials (Cox et al., 2004). The roots of algebraic geometry go back to Descartes' introduction of coordinates to describe points in the Euclidean space and his idea of describing curves and surfaces by algebraic equations. Except for the work done on resultants (e.g., Sylvester (1853) and Macaulay (1902, 1916)), the historical focus of algebraic geometry was initially more on abstract algebra than on multivariate polynomial system solving. However, in the 1960s, the computational aspects of algebraic geometry re-entered the scene with the work of Buchberger (1965). Buchberger's algorithm computes a so-called Gröbner basis, which has been one of the main tools to solve systems of multivariate polynomial equations ever since. The methods of Faugère (1999, 2002) and their extensions are currently the most efficient methods to compute a Gröbner basis. Gröbner bases have dominated computer algebra, but remain computationally very expensive and are symbolic in nature, which means that their extensions to floating-point arithmetic are known to be rather cumbersome (Kondratyev, 2003; Stetter, 2004).

On the other hand, iterative nonlinear root-finding algorithms, e.g., Newton and quasi-Newton methods, do not suffer from these floating-point issues, but are heuristic: they do not guarantee to find the exact solutions and strongly depend on their initial guesses. Nocedal and Wright (2006) give an extensive summary about these nonlinear root-finding algorithms.

Homotopy continuation methods (see for example Li (1997) and Verschelde (1996)) employ a mixture of tech-

⋆ This work was supported in part by the KU Leuven: Research Fund (projects C16/15/059, C3/19/053, C24/18/022, C3/20/117), Industrial Research Fund (fellowships 13-0260, IOF/16/004), and several Leuven Research and Development bilateral industrial projects, in part by Flemish Government agencies: FWO (EOS project G0F6718N (SeLMA), SBO project S005319, infrastructure project I013218N, TBM project T001919N, and PhD grants (SB/1SA1319N, SB/1S93918, SB/1S1319N)), EWI (Flanders AI Research Program), and VLAIO (City of Things (COT.2018.018), Baekeland PhD mandate (HBC.2019.2204), and innovation mandate (HBC.2019.2209)), and in part by the European Commission (EU Research Council under the European Union's Horizon 2020 research and innovation programme (ERC Adv. Grant under grant 885682). Other funding: Foundation "Kom op tegen Kanker", CM (Christelijke Mutualiteit). The work of Christof Vermeersch was supported by the FWO Strategic Basic Research fellowship under grant SB/1SA1319N. *(Corresponding author: Christof Vermeersch.)*

niques from algebraic geometry and nonlinear root-finding, and they can be seen as a hybrid technique to solve systems of multivariate polynomial equations. Although issues with ill-conditioning still exist, homotopy continuation methods are inherently parallel, i.e., each isolated solution can be computed independently, and are currently among the most competitive algorithms to solve systems of multivariate polynomial equations.

Despite their manifest common historical ground, the intimate link between polynomial equations and linear algebra has been neglected in most of the algebraic geometry literature since the end of the 19th century until well into the 20th century (Dreesen, 2013). During the 1980s, the work of Lazard and Stetter (and coworkers) revived the interest in matrix-based methods for solving systems of multivariate polynomial equations. Auzinger and Stetter (1988) rigorously established the link between polynomial system solving and eigenvalue decompositions. This link has been further explored by, among others, Corless et al. (1995), Emiris and Mourrain (1999), Mourrain and Pan (2000), Hanzon and Jibetean (2003), and Faugère (1999, 2002). Stetter (2004) observed that, at that time, the only way to obtain a basis for the quotient space using commonly available software was via Gröbner basis methods. Dreesen, Batselier, and De Moor (Dreesen, 2013; Dreesen et al., 2012, 2018) have overcome this hurdle and developed a pure linear algebra approach to solve systems of multivariate polynomial equations, using only the null space of the Macaulay matrix and techniques from systems theory and linear algebra. This numerical linear algebra approach yields results that are exact within machine precision, as it relies on well-established floating-point algorithms to compute the singular value or eigenvalue decomposition.

In this paper, we consider the column space of the Macaulay matrix instead. We avoid the singular value decomposition to determine a numerical basis of the null space and work on the data in the column space directly. Many properties have a complementary interpretation in both subspaces of the Macaulay matrix. Our main contribution is a novel, complementary algorithm that finds the common roots of the system of multivariate polynomials, starting from the information in the column space of the Macaulay matrix.

This paper proceeds as follows: Section 2 rigorously defines the Macaulay matrix and its (right) null space. We show how to solve systems of multivariate polynomial equations using the null space of the Macaulay matrix in Section 3 and translate this approach to the column space in Section 4. Section 5 contains a numerical example to illustrate our new approach and to compare it with the existing null space based root-finding algorithm. We conclude this paper and point at ideas for future research in Section 6.

2. MACAULAY MATRIX AND ITS NULL SPACE

2.1 Systems of multivariate polynomial equations

A system of multivariate polynomial equations $\mathcal{S}$ defines a set of solutions, which are the common roots of the n different n-variate polynomials (with real coefficients) $p_i(x_1, \ldots, x_n)$. We denote this system $\mathcal{S}$ as

$$\mathcal{S} = \begin{cases} p_1(x_1, \ldots, x_n) = 0 \\ \quad\vdots \\ p_n(x_1, \ldots, x_n) = 0 \end{cases}$$

and refer to its solution set as $\mathcal{B}_\mathcal{S}$. The total degree d_i of every polynomial p_i corresponds to the highest degree among all monomials of that polynomial. We can rewrite a polynomial $p_i(x_1, \ldots, x_n)$ as its coefficient vector p multiplied by a vector v that contains all the monomials. For example, the univariate polynomial $p(x) = x^2 + 2x - 3$ can be represented by the vector $p = \begin{bmatrix}-3 & 2 & 1\end{bmatrix}^\mathrm{T}$. The polynomial $p(x)$ then corresponds to $p(x) = p^\mathrm{T} v$, with the vector of monomials $v = \begin{bmatrix}1 & x & x^2\end{bmatrix}^\mathrm{T}$. In order to have an unambiguous notation, this representation requires a consensus about the ordering of the multivariate monomials. Although we use the degree negative lexicographic ordering in this paper (Dreesen et al., 2018), the remainder of this paper remains valid for any (graded) multivariate monomial ordering.

In the one-dimensional case, the fundamental theorem of algebra states that a univariate polynomial $p(x)$ with complex coefficients of degree d has exactly d roots in the closed field of the complex numbers. The theorem of Bézout extends this primordial theorem in the multidimensional situation, where, due to algebraic interactions among the coefficients of the polynomials, also solutions at infinity can emerge (Cox et al., 2004). We assume in this paper that the system $\mathcal{S}$ has an isolated zero-dimensional solution set $\mathcal{B}_\mathcal{S}$. Then, the total number of solutions in the projective space $\#\mathcal{B}_\mathcal{S}$, counted with multiplicities, is given by the Bézout number m_b, which includes both the m_a affine solutions and the m_∞ solutions at infinity, or

$$m_b = m_a + m_\infty = \prod_{i=1}^{n} d_i,$$

with d_i the total degree of every polynomial p_i.

Example 1. As an example, we consider the following bivariate system:

$$\mathcal{S}_1 = \begin{cases} x_1 - 3x_2^2 = 0 \\ 2x_1 - 6x_2 = 0 \end{cases}. \tag{1}$$

It consists of two bivariate polynomials of total degree $d_1 = 2$ and $d_2 = 1$. Hence, the Bézout number equals $m_b = 2 \cdot 1 = 2$, which agrees with the fact that the system has two common roots $(0, 0)$ and $(3, 1)$.

Example 2. A slightly different bivariate system,

$$\mathcal{S}_2 = \begin{cases} x_1 - 3x_2^2 = 0 \\ 2x_1x_2 - 6x_2 = 0 \end{cases},$$

with polynomials of total degree $d_1 = 2$ and $d_2 = 2$, has four solutions, in accordance with the Bézout number $m_b = 2 \cdot 2 = 4$. Three solutions, namely $(0, 0)$, $(3, 1)$ and $(3, -1)$, are affine, while one solution lives at infinity.

2.2 Macaulay matrix and its null space

In order to solve a system of multivariate polynomial equations, we incorporate its polynomials in the Macaulay matrix (Macaulay, 1902, 1916).

Definition 1. (Macaulay matrix). The Macaulay matrix $M(d) \in \mathbb{R}^{p \times q}$ of degree d contains as its rows the coefficient vectors of the polynomials p_i and their shifts $\{x^{\alpha_i}\} p_i$:

$$M(d) = \begin{bmatrix} \{x^{\alpha_1}\} p_1 \\ \vdots \\ \{x^{\alpha_n}\} p_n \end{bmatrix}, \tag{2}$$

where every polynomial $p_i(x_1, \ldots, x_n)$, for $i = 1, \ldots, n$, is multiplied (or shifted) by all monomials $\{x^{\alpha_i}\}$ of total degree $\alpha_i \leq d - d_i$.

Example 3. We resume Example 1 and build the Macaulay matrix $M(2)$ for the system $\mathcal{S}_1$ in Equation (1). The first polynomial has total degree $d_1 = 2$. Therefore, we do not need to multiply it. The second polynomial, on the other hand, has total degree $d_2 = 1$, which means that we have to multiply it by all monomials of total degree $\alpha_2 \leq 1$, which are x_1 and x_2. Hence, the Macaulay matrix $M(2)$ equals

$$M(2) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & -3 \\ 0 & 2 & -6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & -6 & 0 \\ 0 & 0 & 0 & 0 & 2 & -6 \end{bmatrix}. \tag{3}$$

Using the Macaulay matrix in Equation (2), we can now rewrite the system of multivariate polynomial equations as

$$\underbrace{\begin{bmatrix} \{x^{\alpha_1}\} p_1 \\ \vdots \\ \{x^{\alpha_n}\} p_n \end{bmatrix}}_{M(d)} \underbrace{\begin{bmatrix} 1 \\ \hline x_1 \\ \vdots \\ x_n \\ \hline x_1^2 \\ \vdots \end{bmatrix}}_{v(d)} = 0.$$

The vector $v(d)$ is a vector in the (right) null space of $M(d)$ and has a special multivariate Vandermonde structure. If we consider, for didactic purposes, only systems with simple, affine solutions (see Subsection 3.2 for systems with multiple solutions and solutions at infinity), then there exists such a vector for every solution of the system. Together, they span the entire null space. This leads naturally to the definition of the multivariate Vandermonde basis $V(d) \in \mathbb{C}^{q \times m_a}$ of degree d,

$$V(d) = \begin{bmatrix} 1 & \cdots & 1 \\ \hline x_1|_{(1)} & \cdots & x_1|_{(m_a)} \\ \vdots & & \vdots \\ x_n|_{(1)} & \cdots & x_n|_{(m_a)} \\ \hline x_1^2|_{(1)} & \cdots & x_1^2|_{(m_a)} \\ \vdots & & \vdots \end{bmatrix},$$

which has one column $v(d)|_{(j)}$ for every solution of the system.

Example 4. Since we know the common roots of the system $\mathcal{S}_1$ in Example 1, we can build the multivariate Vandermonde basis $V(2)$ of the null space directly, i.e.,

$$V(2) = \begin{bmatrix} 1 & 1 \\ \hline 0 & 3 \\ 0 & 1 \\ \hline 0 & 9 \\ 0 & 3 \\ 0 & 1 \end{bmatrix}.$$

One can easily check that the columns of this basis annihilate the Macaulay matrix in Equation (3).

3. NULL SPACE BASED ROOT-FINDING

We now exploit the special structure of the null space of the Macaulay matrix in order to find the solutions of the system of multivariate polynomial equations. For didactic purposes, we first assume that all solutions are simple and affine, which allows us to show that a multidimensional realization problem in the null space yields the exact solutions. Next, we show how to deal with multiplicities and how the solutions at infinity can be deflated. Finally, we summarize the null space based root-finding algorithm.

3.1 Multidimensional realization theory

We consider a system of multivariate polynomial equations that only has m_a simple, affine solutions (hence, we have an affine isolated zero-dimensional solution set $\mathcal{B}_\mathcal{S}$), e.g., the system $\mathcal{S}_1$ in Equation (1). As we iteratively increase the degree d of the Macaulay matrix $M(d)$, we notice that the nullity (the dimension of the null space) grows, until it stabilizes at the Bézout number m_b ($= m_a$, in this case). As mentioned in the previous section, the null space of the Macaulay matrix has a special multi-shift-invariant structure, which means that if we select a row from a basis of the null space and multiply (or shift) it by one of the variables, we find again a row from that basis. Note that this structure is a property of the null space as a vector space and not of the specific basis (Dreesen, 2013).

Example 5. To clarify, one could consider for example a vector of the multivariate Vandermonde basis $V(2)$ of degree $d = 2$, i.e.,

$$v(2) = \begin{bmatrix} 1 \\ \hline x_1 \\ x_2 \\ \hline x_1^2 \\ x_1 x_2 \\ x_2^2 \end{bmatrix},$$

and multiply the first three elements by x_1. The elements obtained after the multiplication are again three elements of that vector $v(2)$:

$$\begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix} \xrightarrow{x_1} \begin{bmatrix} x_1 \\ x_1^2 \\ x_1 x_2 \end{bmatrix}.$$

We can also write this multiplication, by means of two row selection matrices S_1 and S_{x_1}, as $S_1 v(2) x_1 = S_{x_1} v(2)$, with

$$S_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad S_{x_1} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

This property does not restrict itself to one variable. Any shift polynomial $g(x_1, \ldots, x_n)$ in the given variables results in a valid multiplication. If we consider this multi-shift-invariance for every column of the multivariate Vandermonde basis V (we no longer specify the degree d explicitly, but we assume it to be large enough), we obtain a generalized eigenvalue problem

$$S_1 V D_g = S_g V, \tag{4}$$

where the diagonal matrix D_g contains the evaluations of the shift polynomial $g(x_1, \ldots, x_n)$ in the different solutions. In order for this eigenvalue problem to not be degenerate, the matrix $S_1 V$ has to be square non-singular. This means that the row selection matrix S_1 must select at least m_a linearly independent rows from the multivariate Vandermonde basis V. Actually, from algebraic geometry, it follows that these linearly independent correspond to the standard monomials, and hence, to the solutions of the system (Cox et al., 2004; Dreesen, 2013). The matrix S_g, on the other hand, selects the rows obtained after the multiplication with the shift polynomial $g(x_1, \ldots, x_n)$, e.g., if we multiply in the previous example the first three monomials by the shift polynomial $g(x_1, x_2) = 2x_1 + 3x_2$, then the selection matrix S_g equals

$$S_g = \begin{bmatrix} 0 & 2 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 & 2 & 3 \end{bmatrix}.$$

In practice, we do not know the multivariate Vandermonde basis V of the null space in advance, since it is constructed from the unknown solutions. Therefore, as the multi-shift-invariance is a property of the null space, we work with a numerical basis Z, obtained for example via the singular value decomposition. There exists a relation between these two bases, namely $V = ZT$, with $T \in \mathbb{C}^{m_a \times m_a}$ a non-singular transformation matrix, which reduces Equation (4) to a solvable generalized eigenvalue problem

$$(S_1 Z)\, T D_g = (S_g Z)\, T, \tag{5}$$

where T contains the eigenvectors and D_g the eigenvalues of the matrix pencil $(S_1 Z, S_g Z)$, and yields alternatively the standard eigenvalue problem

$$T D_g T^{-1} = (S_1 Z)^{\dagger} (S_g Z)\,. \tag{6}$$

We can then use the matrix of eigenvectors T to retrieve the multivariate Vandermonde matrix V, via $V = ZT$, and to find the solutions of the system.

3.2 Multiple solutions and solutions at infinity

Multiple solutions When all solutions are simple, we find one column in the multivariate Vandermonde basis V of the null space for every solution of the system and every column contributes to the nullity of the Macaulay matrix. However, if multiple solutions prevail, the null space of the Macaulay matrix no longer contains only the multivariate Vandermonde solution vectors $v|_{(j)}$, but also linear combinations of the partial derivatives of these solution vectors, i.e., a confluent Vandermonde matrix (Dreesen, 2013). Möller and Stetter (1995) and Dayton et al. (2011) elaborate in more detail on the consequences of multiple solutions. Except for the well-known loss of accuracy in computing multiple eigenvalues, multiplicity poses no problem for the above-described null space based root-finding approach (Dreesen et al., 2012).

Solutions at infinity Systems of multivariate polynomial equations can have solutions at infinity. As in the affine case and for systems with an isolated zero-dimensional solution set $\mathcal{B}_{\mathcal{S}}$, when the degree of the Macaulay matrix increases, the nullity grows with it, until it stabilizes at the Bézout number m_b ($d = d^*$). The Bézout number now includes both the affine solutions and the solutions at infinity. In that null space, we find not only linearly independent rows that correspond to affine solutions, but also linearly independent rows that correspond to solutions at infinity. When the degree d of the Macaulay matrix $M(d)$ further increases ($d > d^*$), some of the linearly independent rows stabilize at their position (they correspond to the affine solutions), while other linearly independent rows keep on moving to higher degree blocks (they correspond to the solutions at infinity) [1]. A gap without any solutions eventually emerges and helps to separate the affine solutions from the solutions at infinity. Fig. 1 visualizes this behavior. We actually know that, when the nullity of the Macaulay matrix stabilizes, its null space can be modeled as the column space of an observability matrix of a multidimensional descriptor system, where the dimension corresponds the number of variables n of the system (Dreesen, 2013). The column space of such an observability matrix is the union of two subspaces: one that is forward multi-shift-invariant and corresponds to the affine solutions (with the causal part of the observability matrix), and one that is backward multi-shift-invariant and corresponds to the solutions at infinity (with the acausal part of the observability matrix).

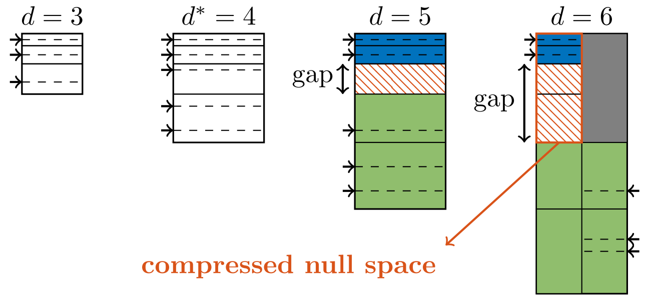

Fig. 1. The null space of the Macaulay matrix $M(d)$ grows as its degree d increases. At a certain degree d^*, the nullity stabilizes at the Bézout number m_b. From that degree on, some linearly independent rows (that correspond to the affine solutions) stabilize, while the other linearly independent rows (that correspond to the solutions at infinity) move to higher degree blocks. A gap separates these linearly independent rows. The influence of the solutions at infinity can be removed via a column compression. The affine root-finding procedure can then be applied straightforwardly on the compressed null space.

Theorem 6. (Column compression). The numerical basis $Z = \begin{bmatrix} Z_1^{\mathrm{T}} & Z_2^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ of the null space is a $q \times m_b$ matrix, which can be partitioned into a $k \times m_b$ matrix Z_1 (that contains the part with the affine solutions and the gap) and a $(q-k) \times m_b$ matrix Z_2 (that contains the part with the solutions at infinity), with rank $(Z_1) = m_a < m_b$. Furthermore, let the singular value decomposition of $Z_1 = U \Sigma Q^{\mathrm{T}}$. Then, $W = ZQ$ is called the column compression of Z and can be partitioned as

$$W = \begin{bmatrix} W_{11} & 0 \\ W_{21} & W_{22} \end{bmatrix},$$

where W_{11} is the $k \times m_a$ matrix that corresponds to the compressed numerical basis of the null space.

This column compression deflates the solutions at infinity and allows us to straightforwardly use the above-described affine null space based root-finding approach as if no

[1] A degree block contains all rows (or columns) that correspond to monomials with the same degree (e.g., x_1^2, $x_1 x_2$, and x_2^2).

solutions at infinity were present (we simply replace Z in Equation (5) by W_{11}), provided that the gap can accommodate for the shift polynomial $g(x_1,\ldots,x_n)$ (a shift polynomial of total degree d_g requires a gap of at least d_g degree blocks).

3.3 Null space based root-finding algorithm

Algorithm 1. (Null space based root-finding).

(1) Construct the Macaulay matrix $M(d) \in \mathbb{R}^{p\times q}$ of large-enough degree $d > d^*$.
(2) Compute a numerical basis Z of the null space of $M(d)$.
(3) Determine the gap and the number of affine solutions m_a via rank tests.
(4) Use Theorem 6 to obtain the compressed numerical basis W_{11} of the null space (note that if $m_b = m_a$, then $W_{11} = Z$).
(5) For a user-defined shift polynomial $g(x_1,\ldots,x_n)$, solve the eigenvalue problem
$$S_1 W_{11} T D_g = S_g W_{11} T,$$
where the matrices S_1, S_g, T, and D_g are defined as in Equation (5).
(6) Retrieve the different components of the solutions from the multivariate Vandermonde basis $V = W_{11}T$.

4. COLUMN SPACE BASED ROOT-FINDING

In this section, we consider the column space of the Macaulay matrix instead of its null space. The complementarity between both subspaces enables a novel, complementary root-finding algorithm that works on the data in the column space of the Macaulay matrix directly.

4.1 Complementarity between both subspaces

The null space and column space of a matrix share an intrinsic complementarity (Dreesen, 2013):

Lemma 7. (Complementary subspaces). Let us consider an arbitrary matrix $M \in \mathbb{R}^{p\times q}$, with $\text{rank}(M) = r$, and a basis of its null space $Z \in \mathbb{R}^{q\times m_b}$, with $\text{rank}(Z) = q - r$, because of the rank-nullity theorem. We reorder the columns of the matrix as $[M_A \; M_B]$, where the $p \times r$ matrix M_B contains r linearly independent columns, and we partition the rows of $Z = \begin{bmatrix} Z_A^{\mathrm{T}} & Z_B^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}$ accordingly. This reordering and partitioning are generally not unique, but always exist. One can easily prove that
$$\text{rank}(M_B) = r \Leftrightarrow \text{rank}(Z_A) = q - r.$$

We can now apply Lemma 7 to the Macaulay matrix and its null space. The solutions of a system of multivariate polynomial equations give rise to the linearly independent rows of the basis of the null space. When we check the rank of this basis row-wise from top to bottom, every linearly independent row corresponds to one solution. If we gather these linearly independent rows in a matrix Z_A, which has full rank $q - r$, then we know, because of Lemma 7, that there exists a partitioning M_B of the columns of the Macaulay matrix, which has full rank r. Consequently, the remaining columns M_A of the Macaulay matrix linearly depend on the columns of M_B. They correspond to the linearly independent rows of the basis of the null space, and hence to the solutions of the system.

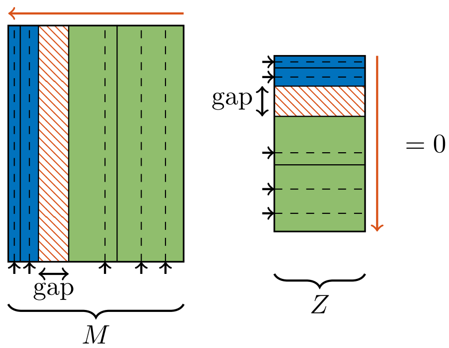

Fig. 2. The solutions of a system of multivariate polynomial equations give rise to linearly independent rows of the basis of the null space of the Macaulay matrix. If we check the rank of this basis row-wise from top to bottom, every linearly independent row corresponds to one solution. Because of the complementarity between the null space and column space of the Macaulay matrix, the linearly dependent columns of the Macaulay matrix, checked column-wise from right to left, correspond to the same solutions.

Corollary 8. The solutions of a system of multivariate polynomial equations give rise to both the linearly dependent columns of the Macaulay matrix (checked column-wise from right to left) and to linearly independent rows of the basis of its null space (checked row-wise from top to bottom). Fig. 2 visualizes this complementarity.

4.2 Column space based root-finding

If we consider a Macaulay matrix $M(d) \in \mathbb{R}^{p\times q}$, with large-enough degree $d > d^*$, such that the nullity has stabilized at the Bézout number m_b and a large-enough gap has emerged, then we know that
$$MW = M\begin{bmatrix} W_{11} & 0 \\ W_{21} & W_{22}\end{bmatrix} = 0,$$
where $W \in \mathbb{C}^{q\times m_b}$ is a special compressed multivariate Vandermonde basis of the null space, in which the matrix W_{11} contains the part with the affine solutions and the gap, while the matrices W_{21} and W_{22} correspond to the part with the solutions at infinity.

Next, we define two new matrices A and B. The matrix $A \in \mathbb{C}^{m_a\times m_a}$ contains all the rows of the basis W that correspond to the affine standard monomials, i.e., the monomials that lead to the affine solutions. If we multiply (or shift) these rows by the user-defined shift polynomial $g(x_1,\ldots,x_n)$, we obtain $(m_a - m_h)$ rows that are again present in the matrix A and m_h rows that are not. We gather these m_h non-present rows in the matrix $B \in \mathbb{C}^{m_h\times m_a}$ and rewrite this shift property as
$$AD_g = S_g\begin{bmatrix} A \\ B\end{bmatrix}, \tag{7}$$
with S_g an $m_a \times (m_a + m_h)$ matrix that selects the m_a combinations of rows obtained after the multiplication. The matrix D_g is a diagonal matrix that contains again the evaluations of the shift polynomial $g(x_1,\ldots,x_n)$. If we extract the matrix A from the right-hand side of Equation (7), an eigenvalue problem appears
$$AD_g = S_g\begin{bmatrix} I \\ BA^{-1}\end{bmatrix} A,$$

or

$$AD_gA^{-1} = S_g \begin{bmatrix} I \\ BA^{-1} \end{bmatrix}. \tag{8}$$

The matrix A is invertible because it contains exactly m_a linearly independent rows. Of course, since we do not know the matrices A and B in advance, we can not solve this eigenvalue problem right away. In the remainder of this subsection, we circumvent this problem via the complementarity between the null space and column space.

The matrices A and B contain rows of the basis W of the null space, in particular of the matrix W_{11}. If we define the matrix $C \in \mathbb{C}^{(k-m_a-m_h)\times m_a}$ as the matrix that contains the remaining rows of W_{11}, then we can reorder the basis W as

$$PW = \begin{bmatrix} A & 0 \\ B & 0 \\ C & 0 \\ W_{21} & W_{22} \end{bmatrix}.$$

The matrix P is a $q \times q$ row-permutation matrix. We can rearrange the columns of the Macaulay matrix in accordance to the reordered basis of the null space and obtain

$$\underbrace{[M_1\ M_2\ M_3\ M_4]}_{N} \begin{bmatrix} A & 0 \\ B & 0 \\ C & 0 \\ W_{21} & W_{22} \end{bmatrix} = 0, \tag{9}$$

where every matrix M_i corresponds to a subset of the columns of the Macaulay matrix. We could even replace the Macaulay matrix M by the upper triangular matrix R of its QR-decomposition and reorder this upper triangular matrix R instead as

$$\underbrace{[R_1\ R_2\ R_3\ R_4]}_{N} \begin{bmatrix} A & 0 \\ B & 0 \\ C & 0 \\ W_{21} & W_{22} \end{bmatrix} = 0. \tag{10}$$

This initial QR-decomposition serves as a pre-processing step and reduces the number of rows of the resulting matrix N. We call the matrix $N = MP^{\mathrm{T}}$ or $N = RP^{\mathrm{T}}$ in both situations the reordered matrix.

We now apply the so-called backward QR-decomposition [2] on this reordered matrix N, which yields

$$\begin{bmatrix} R_{14} & R_{13} & R_{12} & R_{11} \\ R_{24} & R_{23} & R_{22} & 0 \\ R_{34} & R_{33} & 0 & 0 \\ R_{44} & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} A & 0 \\ B & 0 \\ C & 0 \\ W_{21} & W_{22} \end{bmatrix} = 0.$$

We notice that $R_{33}B = -R_{34}A$, which means that

$$BA^{-1} = -R_{33}^{-1}R_{34},$$

because R_{33} is always of full rank (since B is not of full rank and Lemma 7). This relation helps to remove the dependency on the null space in Equation (8) and yields a solvable standard eigenvalue problem

$$AD_gA^{-1} = S_g \begin{bmatrix} I \\ -R_{33}^{-1}R_{34} \end{bmatrix}. \tag{11}$$

This eigenvalue problem yields the solutions of the system of multivariate polynomial equations via the eigenvalues in D_g and the eigenvectors in A. Notice that this complementary column space approach does not require a column compression to remove the influence of the solutions at infinity, because the backward QR-decomposition already separates them implicitly.

Remark 9. Contrary to the null space based root-finding approach, the user-defined shift polynomial $g(x_1, \ldots, x_n)$ has an influence on the reconstruction of the solutions. If not all components of the solution vector are present in the matrix A, we must select the shift polynomial such that we can reconstruct the entire solution vector from the eigenvalues and eigenvectors (and sometimes even use multiple shift polynomials).

4.3 Complementary column compression

In the null space based root-finding algorithm, we use a column compression of the numerical basis of the null space to deflate the solutions at infinity. Because of the structure of the reversed QR-decomposition, the influence of the solutions at infinity disappear implicitly when working in the column space. However, there exists a complementary column compression in the column space that compresses the Macaulay matrix and reduces the computational complexity of the column space based approach.

Theorem 10. (Complementary column compression). The Macaulay matrix $M = [M_1\ M_2]$ of appropriate degree d is a $p\times q$ matrix, which can be partitioned into a $p\times(q-l)$ matrix M_1 (that contains the columns that correspond to the affine solutions and the gap) and a $p \times l$ matrix M_2 (that contains the columns that corresponds to the solutions at infinity), with $\operatorname{rank}(M_2) = l - m_\infty$. Furthermore, let the QR-decomposition of $M_2 = QR = [Q_1\ Q_2]\, R$. The matrix $Q_2 \in \mathbb{C}^{p\times(p-l+m_\infty)}$ is an orthogonal basis of the left null space of M_2. Then, $N = Q_2^{\mathrm{T}}M$ is called the complementary column compression of M and can be partitioned as

$$N = [N_1\ 0],$$

where N_1 is the $(p - l + m_\infty) \times (q - l)$ matrix that corresponds to the compressed Macaulay matrix.

Proof. If we partition the Macaulay matrix $M = [M_1\ M_2]$ and premultiply by the matrix Q_2^{T}, we obtain $N = Q_2^{\mathrm{T}}M = \left[Q_2^{\mathrm{T}}M_1\ Q_2^{\mathrm{T}}M_2\right]$. Since the matrix Q_2 is an orthogonal basis of the left null space of M_2, the matrix $Q_2^{\mathrm{T}}M_2 = 0$ and the theorem is proven. □

Note that this matrix Q_2 does not have to be calculated explicitly. When we look for the gap, we determine, via the singular value decomposition or QR-decomposition, at a certain point this orthonormal basis.

4.4 Column space based root-finding algorithm

Algorithm 2. (Column space based root-finding).

(1) Construct the Macaulay matrix $M(d) \in \mathbb{R}^{p\times q}$ of large-enough degree $d > d^*$.
(2) Replace the Macaulay matrix M by the upper triangular matrix R of its QR-decomposition (optional).
(3) Determine the linear dependent columns from right to left and reorder the Macaulay matrix M or its upper triangular matrix R as in Equations (9)-(10).
(4) Use Theorem 10 to obtain the compressed reordered matrix N_1 (optional).

[2] The backward QR-decomposition corresponds to the ordinary, forward QR-decomposition of a matrix, but starts with the last column and ends with the first column. It yields a backward upper triangular matrix R, analogue to the forward QR-decomposition, but with all the columns mirrored.

Table 1. An overview of the size, rank, and nullity of the Macaulay matrix $M(d)$, for increasing degree d, that comprises system $\mathcal{S}_3$.

degree	**size**	**rank**	**nullity**
3	5×56	5	51
4	30×126	30	96
5	105×252	105	147
6	280×462	270	192
7	630×792	570	222
8	1260×1287	1050	237
9	2310×2002	1760	242
10	3960×3003	2760	243
11	6435×4368	4125	243

(5) Compute the (Q-less) backward QR-decomposition of the reordered matrix N (or the compressed N_1).

(6) For a user-defined shift polynomial $g(x_1, \ldots, x_n)$, solve the eigenvalue problem

$$AD_g A^{-1} = S_g \begin{bmatrix} I \\ -R_{33}^{-1} R_{34} \end{bmatrix},$$

where the matrices S_g, R_{33}, R_{34}, and D_g are defined as in Equation (11).

(7) Retrieve the different components of the solutions from the eigenvalues in the matrix D_g and the eigenvectors in the matrix A.

5. NUMERICAL EXAMPLE

In this section, we consider a realistic system of multivariate polynomial equations to illustrate our new column space based approach and to compare it with the existing null space based root-finding algorithm.

Example 11. (Verschelde, 1999) The following system of 5-variate polynomial equations (with maximum total degree $d_{\max} = 3$) models a neural network by an adaptive Lotka–Volterra system:

$$\mathcal{S}_3 = \begin{cases} x_1x_2^2 + x_1x_3^2 + x_1x_4^2 + x_1x_5^2 - 1.1x_1 + 1 = 0 \\ x_2x_1^2 + x_2x_3^2 + x_2x_4^2 + x_2x_5^2 - 1.1x_2 + 1 = 0 \\ x_3x_1^2 + x_3x_2^2 + x_3x_4^2 + x_3x_5^2 - 1.1x_3 + 1 = 0 \\ x_4x_1^2 + x_4x_2^2 + x_4x_3^2 + x_4x_5^2 - 1.1x_4 + 1 = 0 \\ x_5x_1^2 + x_5x_2^2 + x_5x_3^2 + x_5x_4^2 - 1.1x_5 + 1 = 0 \end{cases}.$$

This system has an isolated zero-dimensional solution set with 233 affine solutions and 10 solutions at infinity.

First, we iteratively build Macaulay matrices $M(d)$ of increasing degree d that comprise the system $\mathcal{S}_3$. Table 1 contains the size, rank, and nullity of these Macaulay matrices. When the Macaulay matrix has degree $d^* = 10$, the nullity is equal to the Bézout number $m_b = 3^5 = 243$, which corresponds to the totale number of solutions. This nullity remains the same if the degree further increases.

Next, we determine the gap in the null space or in the column space. At degree $d^* = 10$, the Macaulay matrix does not contain a gap yet, but a gap emerges for degrees $d \geq 11$. Table 2 summarizes, when we check the rows of the numerical basis Z of the null space degree block-wise from top to bottom, the number of linearly independent rows and shows that the basis of the null space contains a gap at the ninth degree block. The $m_a = 233$ linearly independent rows before the gap correspond to the affine solutions, while the $m_\infty = 10$ linearly independent rows after the gap correspond to the solutions at infinity. Analogously, because of the complementarity between both subspaces, we can also identify the gap while checking the columns of the Macaulay matrix directly. Table 3 summarizes, when we check the columns of the Macaulay matrix degree blockwise from right to left, the number of linearly dependent columns. We observe again that the ninth degree block corresponds to the gap. Notice that, when we consider the column space of the Macaulay matrix, we can use the QR-decomposition as a straightforward preprocessing step, as was mentioned in Section 4. If we replace in this example the 6435×4368 Macaulay matrix $M(11)$ by the 4368×4368 upper triangular matrix R of its QR-decomposition, we already reduce the number of rows with almost 33%. This reduction immediately proves to be useful, both in computational complexity and memory usage, e.g., when we want to determine the gap.

Table 2. A summary of the linearly independent rows of the basis of the null space of the Macaulay matrix $M(11)$ that comprises system $\mathcal{S}_3$.

degree block(s)	**rows**	**lin. indep. rows**	**increase**
0	1	1	1
$0-1$	$1-6$	6	5
$0-2$	$1-21$	21	15
$0-3$	$1-56$	51	30
$0-4$	$1-126$	96	45
$0-5$	$1-252$	147	51
$0-6$	$1-462$	192	45
$0-7$	$1-792$	222	30
$0-8$	$1-1287$	233	11
$0-9$	$1-2002$	233	0
$0-10$	$1-3003$	238	5
$0-11$	$1-4368$	243	5

Table 3. A summary of the linearly dependent columns of the Macaulay matrix $M(11)$ that comprises system $\mathcal{S}_3$.

degree block(s)	**columns**	**lin. dep. cols.**	**increase**
11	$3004-4368$	5	5
$10-11$	$2003-4368$	10	5
$9-11$	$1288-4368$	10	0
$8-11$	$793-4368$	21	11
$7-11$	$463-4368$	51	30
$6-11$	$253-4368$	96	45
$5-11$	$127-4368$	147	51
$4-11$	$57-4368$	192	45
$3-11$	$22-4368$	222	30
$2-11$	$7-4368$	237	15
$1-11$	$2-4368$	242	5
$0-11$	$1-4368$	243	1

Finally, after determining the gap, we remove the influence of the solutions at infinity and solve the eigenvalue problems that yield the solutions of the system of multivariate polynomial equations. In the null space, we perform a column compression on the top part of the numerical basis Z of the null space, which contains all degree blocks up to the ninth degree block (i.e., the gap), and obtain the compressed numerical basis W_{11}. In the column space, on the other hand, we reorder the columns of the upper triangular matrix R as shown in Equation (10). A backward QR-decomposition yields the matrices R_{33} and R_{34} and removes the influence of the solutions at infinity implicitly. The eigenvalues and eigenvectors of the eigenvalue problems in Equation (6) and Equation (11) yield the 233 affine solutions of the system, for the null space based approach and the column space based approach, respectively.

6. CONCLUSION AND FUTURE WORK

In this paper, we revised the Macaulay matrix approach that uses its null space to solve systems of multivariate polynomial equations. We pointed at the complementarity of the null space and column space of this Macaulay matrix and proposed a novel, complementary algorithm that considers the columns space of the Macaulay matrix instead. Contrary to null space based root-finding, this column space based approach does not require an explicit calculation of a numerical basis of the null space, i.e., an expensive singular value decomposition, but directly works on the data in the Macaulay matrix. In that context, we also proposed the complementary column space compression, which compresses the Macaulay matrix and removes the influence of the solutions at infinity. We provided a realistic numerical example to illustrate our new approach and compared it with the existing null space based root-finding algorithm.

This complementary column space based root-finding algorithm has created several research opportunities. First of all, a fast and sparse implementation of the (Q-less) QR-decomposition is much faster than the traditional singular value decomposition. Therefore, one of our current research efforts is to improve current implementations and to exploit both the structure and the sparsity of the Macaulay matrix. Furthermore, the complementarity of both subspaces may yield even more interesting properties in the column space. Together with more efficient implementations, this will give us the machinery to tackle much larger problems in the future.

REFERENCES

Auzinger, W. and Stetter, H.J. (1988). An elimination algorithm for the computation of all zeros of a system of multivariate polynomial equations. In *Proc. of the International Conference on Numerical Mathematics*, 11–30. Singapore.

Buchberger, B. (1965). *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem Nulldimensionalen Polynomideal*. Ph.D. thesis, Universitat Innsbruck, Innsbruck, Austria.

Corless, R.M., Gianni, P.M., Trager, B.M., and Watt, S.M. (1995). The singular value decomposition for polynomial systems. In *Proc. of the 1995 International Symposium on Symbolic and Algebraic Computation (ISSAC)*, 195–207. Montreal, Canada.

Cox, D.A., Little, J.B., and O'Shea, D. (2004). *Using Algebraic Geometry*, volume 185 of *Graduate Texts in Mathematics*. Springer, New York, NY, USA, 2nd edition.

Dayton, B.H., Li, T.Y., and Zeng, Z. (2011). Multiple zeros of nonlinear systems. *Mathematics of Computation*, 80(276), 2143–2168.

Dreesen, P. (2013). *Back to the Roots: Polynomial System Solving Using Linear Algebra*. Ph.D. thesis, KU Leuven, Leuven, Belgium.

Dreesen, P., Batselier, K., and De Moor, B. (2012). Back to the roots: Polynomial system solving, linear algebra, systems theory. In *Proc. of the 16th IFAC Symposium on System Identification*, 1203–1208. Brussels, Belgium.

Dreesen, P., Batselier, K., and De Moor, B. (2018). Multidimensional realisation theory and polynomial system solving. *International Journal of Control*, 91(12), 2692–2704.

Emiris, I.Z. and Mourrain, B. (1999). Matrices in elimination theory. *Journal of Symbolic Computation*, 28(1–2), 3–44.

Faugère, J.C. (1999). A new efficient algorithm for computing Gröbner bases (f4). *Journal of Pure and Applied Algebra*, 139(1–3), 61–88.

Faugère, J.C. (2002). A new efficient algorithm for computing Gröbner bases without reduction to zero (f5). In *Proc. of the 2002 International Symposium on Symbolic and Algebraic Computation (ISSAC)*, 75–83. Lille, France.

Hanzon, B. and Jibetean, D. (2003). Global minimization of a multivariate polynomial using matrix methods. *Journal of Global Optimization*, 27(1), 1–23.

Kondratyev, A. (2003). *Numerical Computation of Groebner Bases*. Ph.D. thesis, Johannes Kepler University, Linz, Austria.

Li, T.Y. (1997). Numerical solution of multivariate polynomial systems by homotopy continuation methods. *Acta Numerica*, 6, 399–436.

Macaulay, F.S. (1902). Some formulae in elimination. *Proc. of the London Mathematical Society*, 1(1), 3–27.

Macaulay, F.S. (1916). *The Algebraic Theory of Modular Systems*, volume 19 of *Cambridge Tracts in Mathematics and Mathematical Physics*. Cambridge University Press, London, UK.

Möller, M.H. and Stetter, H.J. (1995). Multivariate polynomial equations with multiple zeros solved by matrix eigenproblems. *Numerische Mathematik*, 70(3), 311–329.

Mourrain, B. and Pan, V.Y. (2000). Multivariate polynomials, duality, and structured matrices. *Journal of Complexity*, 16(1), 110–180.

Nocedal, J. and Wright, S.J. (2006). *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering. Springer, New York, NY, USA, 2nd edition.

Pan, V.Y. (1997). Solving a polynomial equation: Some history and recent progress. *SIAM Review*, 39(2), 187–220.

Stetter, H.J. (2004). *Numerical Polynomial Algebra*. SIAM, Philadelphia, PA, USA.

Sylvester, J.J. (1853). On a theory of the syzygetic relations of two rational integral functions, comprising an application to the theory of Sturm's functions, and that of the greatest algebraical common measure. *Philosophical Transactions of the Royal Society of London*, 143, 407–548.

Verschelde, J. (1996). *Homotopy Continuation Methods for Solving Polynomial Systems*. Ph.D. thesis, KU Leuven, Leuven, Belgium.

Verschelde, J. (1999). The test database of polynomial systems. `http://homepages.math.uic.edu/~jan/PHCpack/node10.html`. (Visited on April 1, 2021).