Porteous Meets Walsh: Counting Critical Points for $\mathcal{H}_2$-Norm Model Reduction Using Algebraic Geometry Behind Multiparameter Eigenvalue Problems

Sibren Lagauw[1], Leonie Kayser[2], and Bart De Moor[1,3]

[1]STADIUS-ESAT, KU Leuven, Leuven.
[2]MPI for Mathematics in the Sciences, Leipzig.
[3]Fellow IEEE, IFAC, SIAM.

Abstract

Several polynomial systems arising in systems and control have been shown to exhibit a partially linear structure, which allows them to be reformulated as a multiparameter eigenvalue problem (MEP). In this work, we show how such MEP formulations can be leveraged to obtain refined information about the solution set of the original polynomial system. We demonstrate, in the setting of single-input single-output $\mathcal{H}_2$-norm model order reduction, how the Thom–Porteous formula can be used to derive an exact count of the complex solutions of the critical equations derived from Walsh's theorem. We prove, under suitable genericity assumptions, that their number of complex solutions is finite and that all solutions occur without multiplicity. Moreover, we derive an explicit formula for the total number of complex solutions as a function of the model orders. Throughout, we provide a self-contained and application-oriented exposition of the algebro-geometric tools required for our analysis, with the aim of making them accessible to the broader engineering community.

1 Introduction

Systems of polynomial equations arise frequently in engineering applications, yet solving them comprehensively—numerically or symbolically—is computationally challenging. For such systems, one typically has access to unconditional but only coarse upper bounds on the number of affine solutions, such as those provided by Bézout's theorem. Nevertheless, having a precise a priori count of the solutions is valuable both for the design and use of numerical methods and for validating computations by confirming that the full solution set has been recovered.

Several systems of polynomial equations arising in applications from systems and control have been shown to exhibit a partially linear structure, such that they can be reformulated as a multiparameter eigenvalue problem (MEP); see, for example, [BDM12; VD19; DeM20; AVM21; LVM26]. An MEP generalizes the standard eigenvalue problem by allowing multiple eigenvalues and by dropping the restriction of linearity in the parameters: a k-parameter eigenvalue problem has k-tuples of eigenvalues, which we refer to as *eigentuples* [CM21]. In these cases, one can alternatively solve the problem by computing the eigentuples using specialized MEP solvers (e.g., [VD23; HKP24; VM25]) instead of solving the system of equations directly with general-purpose polynomial system solvers (e.g., [BT18]).

In this work, we show how these MEP-based reformulations can also be leveraged to obtain finer theoretical information about the solution set of the original system of polynomial equations. In particular, we demonstrate how the algebraic geometry underlying MEPs can, in some cases, be exploited to obtain an exact count of the complex solutions via the Thom–Porteous formula.

We illustrate this approach in the context of the single-input single-output (SISO) $\mathcal{H}_2$-norm model order reduction problem. Using the corresponding MEP formulation proposed in [LAM23], we derive a formula that describes the number of complex solutions of the critical equations of the model reduction problem, i.e., the number of critical points. The critical equations follow from what is commonly referred to as Walsh's theorem, which motivates the title of this paper.

Our contributions are threefold:

(1) We present a new proof of finiteness of the number of (complex) critical points of the SISO $\mathcal{H}_2$-norm model order reduction problem. We show that, generically, the solutions of the critical equations occur without multiplicity (Corollary 3.3), a property favorable for numerical computation. Moreover, assuming the higher-order model has only simple poles, we derive an explicit formula for the number of complex critical points, i.e., for the total number of complex solutions to the critical equations, as a function of the full and reduced orders (Theorem 3.6).

(2) Our analysis is also applicable to related problems such as discrete-time models and higher order poles (Section 4). We illustrate all theoretical results by means of several (numerical) examples (Section 5).

(3) We provide a self-contained and application-oriented exposition of the key algebro-geometric tools, making them accessible to a broader engineering audience (Section 3 and Appendix A). We show how the fiber dimension theorem and the Thom–Porteous formula can be applied to MEPs and their associated polynomial systems. These techniques are not specific to the model reduction problem under consideration, but are equally applicable to other engineering problems where such MEP formulations arise.

Outline The paper is organized as follows. In Section 2, we give a formal statement of the model reduction problem and derive the critical equations of the optimization problem. The heart of the paper is Section 3, where we lay out a recipe for establishing finiteness and deriving a count of the solution set of MEPs. We implement this recipe for the model reduction problem and derive the corresponding solution count. Several further considerations are given in Section 4: the discrete-time case and the special case where the given higher-order model has multiple poles. We present numerical examples in Section 5 and conclude the paper in Section 6. The complementary Appendix A gives an elementary introduction to the essential algebraic geometry concepts relevant to our work.

Notation Throughout, scalars are denoted by lowercase letters (e.g., $s \in \mathbb{C}$), vectors by boldface lowercase letters (e.g., $\boldsymbol{a} \in \mathbb{C}^k$), and matrices by boldface uppercase letters (e.g., $\boldsymbol{M} \in \mathbb{C}^{k\times k}$). Let $\mathbb{C}[s]_{\leq k}$ denote the $(k+1)$-dimensional vector space of complex polynomials in s of degree at most k. For a polynomial $f(s) = \sum_{i=0}^{k} f_i s^i \in \mathbb{C}[s]_{\leq k}$, we write its coefficient vector in boldface $\boldsymbol{f} = (f_0, \ldots, f_k)^\mathsf{T} \in \mathbb{C}^{k+1}$, and conversely, any vector $\boldsymbol{f} \in \mathbb{C}^{k+1}$ determines a polynomial $f(s) \in \mathbb{C}[s]_{\leq k}$. Let $\boldsymbol{T}_\ell(f) \in \mathbb{C}^{\ell\times(\ell+k)}$ denote the banded Toeplitz matrix generated by ℓ forward-shifts of the coefficient vector $\boldsymbol{f} \in \mathbb{C}^{k+1}$:

$$\boldsymbol{T}_\ell(f) \coloneqq \begin{bmatrix} f_0 & \cdots & f_{k-1} & f_k & 0 & \cdots & 0 \\ 0 & f_0 & \cdots & f_{k-1} & f_k & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & f_0 & \cdots & f_{k-1} & f_k \end{bmatrix}.$$

Toeplitz matrices constructed in this way allow us to represent polynomial multiplication in terms of coefficient vectors. For instance, if $h \in \mathbb{C}[s]$, then the product $h(s)f(s)$ has coefficient vector $\boldsymbol{T}_{k+1}(h)^\mathsf{T}\boldsymbol{f}$. Algebraic varieties are denoted by uppercase letters, with calligraphic font occasionally used for emphasis (e.g., $\mathcal{W}$). Unless stated otherwise, all open and closed subsets mentioned in this paper are taken with respect to the Zariski topology.

2 $\mathcal{H}_2$-norm model order reduction and its critical equations

Throughout, we work with transfer functions of real, continuous-time, strictly causal, stable, single-input single-output (SISO) linear time-invariant (LTI) systems of finite order. For a fixed integer $N \geq 1$, let $\mathcal{M}_N$ denote the set of all such transfer functions of McMillan degree exactly N. Any model $H \in \mathcal{M}_N$ can be parametrized as

$$H(s) = \frac{d(s)}{c(s)} = \frac{d_{N-1}s^{N-1} + \cdots + d_1 s + d_0}{c_N s^N + c_{N-1}s^{N-1} + \cdots + c_1 s + c_0},$$

where $c(s)$ and $d(s)$ are real polynomials with $c_N \neq 0$ and $\gcd(c, d) = 1$. (Hurwitz-)Stability here means that all poles of H, the roots of $c(s)$, lie in the open left half-plane. Strict causality means $\deg c > \deg d$, so that the impulse response vanishes for negative time and is integrable for $t \geq 0$. Consequently, elements of $\mathcal{M}_N$ are real rational functions in the Hardy space $\mathcal{H}_2$ of functions analytic in the open right half-plane and square-integrable along the imaginary axis.

Given a target reduced order $n < N$, we seek an approximant $\widehat{H}$ of McMillan degree at most n, i.e., $\widehat{H}(s) \in \mathcal{M}_{\leq n}$. We parameterize $\widehat{H}$ accordingly:

$$\widehat{H}(s) = \frac{b(s)}{a(s)} = \frac{b_{n-1}s^{n-1} + \cdots + b_1 s + b_0}{a_n s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0},$$

while allowing for the possibility that $\widehat{H}$ has McMillan degree strictly smaller than n, for instance due to pole–zero cancellations between $a(s)$ and $b(s)$. The goal of the model order reduction problem under consideration is to minimize the $\mathcal{H}_2$-norm of the approximation error $\widetilde{H}(s) = H(s) - \widehat{H}(s)$. Formally, the problem corresponds to

$$\underset{\widehat{H}(s) \in \mathcal{M}_{\leq n}}{\text{minimize}} \ \|\widetilde{H}\|_{\mathcal{H}_2}^2, \qquad \|\widetilde{H}\|_{\mathcal{H}_2}^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} |H(\mathrm{j}\omega) - \widehat{H}(\mathrm{j}\omega)|^2 \mathrm{d}\omega = \int_0^{\infty} |h(t) - \widehat{h}(t)|^2 \mathrm{d}t, \tag{1}$$

where $h(t)$ and $\widehat{h}(t)$ denote the impulse responses of H and $\widehat{H}$, respectively, obtained via the inverse Laplace transform.

The model reduction problem (1) is non-convex and may therefore admit multiple local optima. Its first-order optimality conditions can be derived in two complementary ways: analytically, by partial differentiation of the $\mathcal{H}_2$-norm error, and geometrically, by exploiting orthogonality in the Hardy space $\mathcal{H}_2$ (see e.g., [GAB08; ML67]). By doing so, one obtains a set of Hermite interpolation constraints at the reflected poles of the reduced model $\widehat{H}$. This foundational result, originally due to Aigrain and Williams [AW49], was established for reduced models with distinct, real poles and later extended by Meier and Luenberger [ML67] to cover complex and higher-order poles. Model reduction for linear systems is closely related with the mathematical discipline of approximation on the complex domain. From this viewpoint it is unsurprising that similar optimality conditions were already known in the field of classical analysis in the 1920s, e.g., in work by Walsh [Wal60].

Theorem 2.1 (Meier and Luenberger [ML67])**.** *Consider, for a given $H \in \mathcal{M}_N$, the model reduction problem as formulated in (1). A necessary condition for a reduced model $\widehat{H} \in \mathcal{M}_{\leq n}$ to be a local minimizer is that, for each pole $p_i \in \mathbb{C}$ of $\widehat{H}$, the following interpolation conditions hold:*

$$H(-p_i)^{(j)} = \widehat{H}(-p_i)^{(j)}, \quad j = 0, \ldots, d_i, \tag{2}$$

where d_i is the pole order of p_i and $(-)^{(j)}$ denotes differentiation: $F(a)^{(j)} = \frac{\mathrm{d}^j F(s)}{\mathrm{d}s^j}|_{s=a}$.

We illustrate conceptually how the interpolatory conditions follow naturally from (1), assuming, for the ease of notation, that $\widehat{H}$ has simple poles only. Express $\widehat{H}$ in a partial fraction expansion

$$\widehat{H}(s) = \frac{r_1}{s - p_1} + \frac{r_2}{s - p_2} + \cdots + \frac{r_n}{s - p_n},$$

and consider the Wirtinger derivatives[1] of $J^2 \coloneqq \|\widetilde{H}\|_{\mathcal{H}_2}^2$ with respect to the conjugates $\overline{r_i}$,

$$\begin{aligned}\frac{\partial J^2}{\partial \overline{r_i}} &= \frac{1}{\pi \mathrm{j}} \int_{-\mathrm{j}\infty}^{\mathrm{j}\infty} \left(H(s) - \widehat{H}(s)\right) \frac{\partial}{\partial \overline{r_i}} \overline{\left(H(s) - \widehat{H}(s)\right)} \,\mathrm{d}s \\ &= \frac{-1}{\pi \mathrm{j}} \int_{-\mathrm{j}\infty}^{\mathrm{j}\infty} \left(H(s) - \widehat{H}(s)\right) \left(\frac{1}{\overline{s} - \overline{p_i}}\right) \mathrm{d}s,\end{aligned} \tag{3}$$

and $\overline{p_i}$,

$$\frac{\partial J^2}{\partial \overline{p_i}} = \frac{-1}{\pi \mathrm{j}} \int_{-\mathrm{j}\infty}^{\mathrm{j}\infty} \left(H(s) - \widehat{H}(s)\right) \left(\frac{\overline{r_i}}{(\overline{s} - \overline{p_i})^2}\right) \mathrm{d}s. \tag{4}$$

Because the objective function J^2 is real-valued, we know that the Wirtinger derivatives with respect to r_i and p_i are equal to the complex conjugates of (3) and (4), respectively:

$$\overline{\left(\frac{\partial J^2}{\partial r_i}\right)} = \frac{\partial J^2}{\partial \overline{r_i}} \quad \text{and} \quad \overline{\left(\frac{\partial J^2}{\partial p_i}\right)} = \frac{\partial J^2}{\partial \overline{p_i}}.$$

It therefore suffices to equate the expressions in (3)–(4) to zero for all $i = 1, \ldots, n$, to compose the first-order necessary conditions for optimality. Using properties of the Laplace transform, these $2n$ equations can be shown to be equivalent to the interpolatory conditions described in Theorem 2.1. See, e.g., [ML67], for more details.

The conditions in Theorem 2.1 can be expressed compactly using the polynomial $\tilde{a}(s) \coloneqq a(-s) \in \mathbb{C}[s]_{\leq n}$, which has the mirror images of the poles p_i as its roots. Indeed, any model $\widehat{H}$ that satisfies the interpolatory conditions (2) must be such that the numerator of the residual $H(s) - \widehat{H}(s) = \frac{d(s)}{c(s)} - \frac{b(s)}{a(s)}$ is divisible by $\tilde{a}(s)^2$.

Theorem 2.2 (Walsh's theorem [Reg95])**.** *Consider a higher-order model $H \in \mathcal{M}_N$. If $\widehat{H} \in \mathcal{M}_{\leq n}$ is a local minimizer of the model reduction problem (1), then, for all $s \in \mathbb{C}$,*

$$H(s) - \widehat{H}(s) = \frac{d(s)}{c(s)} - \frac{b(s)}{a(s)} = \left(\frac{\tilde{a}(s)}{a(s)}\right)^2 G(s),$$

with $G(s)$ the Laplace-transform of some real-valued, stable and causal signal.

Clearing denominators in the above expression shows that, if $\widehat{H}$ is a locally minimizing solution, then there must exist a real polynomial $g(s) \in \mathbb{C}[s]_{\leq N-n-1}$ such that

$$a(s)d(s) - b(s)c(s) = \tilde{a}(s)^2 g(s). \tag{5}$$

Since (5) must hold identically as a polynomial identity in $s \in \mathbb{C}$, it is equivalent to the vanishing of each coefficient in the expansion of the polynomial $a(s)d(s) - b(s)c(s) - \tilde{a}(s)^2 g(s)$. Enforcing these coefficients to vanish gives a system of $N + n$ polynomial equations in the unknown coefficients of the polynomials a, b, and g. It is given by the following matrix expression:

$$\boldsymbol{T}_N(a)^\mathsf{T} \boldsymbol{d} - \boldsymbol{T}_n(c)^\mathsf{T} \boldsymbol{b} - \boldsymbol{T}_N(\tilde{a})^\mathsf{T} \boldsymbol{T}_{N-n}(\tilde{a})^\mathsf{T} \boldsymbol{g} = \boldsymbol{0}. \tag{6}$$

This system of polynomial equations characterizing the locally optimal solutions is referred to as the *critical equations*. Its complex common roots are the so-called *critical points*. Among these, only real critical points that correspond to stable reduced models represent genuine stationary points (local minima, saddles, or local maxima) of the original optimization problem. Note that the coefficients of b and g appear only linearly in (6). It is precisely this partially linear structure that will be exploited in the corresponding MEP formulation (cf. Section 3.3).

[1] The Wirtinger derivative of $f(s, \overline{s})$ wrt. s (respectively $\overline{s}$) is calculated using the standard rules of differentiation, while considering the variable $\overline{s}$ (resp. s) as a constant, i.e., s and $\overline{s}$ are treated as two independent variables for the purpose of differentiation. This is also called $\mathbb{CR}$-*calculus* [Kre09].

Under the assumption that all poles of H are simple, it is known that the critical points are a finite set of isolated points; see [Bar86, Theorem 3] and [HPB12, Theorem 3.2]. However, despite the extensive literature on $\mathcal{H}_2$-norm model reduction (see, e.g., [Ant05; GAB08] and references therein), an exact count of the critical points for arbitrary problem sizes was, to the best of our knowledge, not available. Beyond its intrinsic theoretical interest, such a count of the critical points is informative for computation. In particular, it provides a concrete proxy for the cost of globally optimal approaches that certify global optimality by enumerating all critical points (see e.g., [AVM21; Als19; LAM23; PBH12]), and it offers a benchmark for verifying that such enumerations are complete. Note that these globally optimal methods are computationally feasible only for relatively small problems and are therefore best regarded as theoretical tools, mostly used for benchmarking purposes.

3 Finiteness and degree of the critical points

In our analysis of the critical points, we will deal with several products of vector spaces. To ease notation, we use the following shorthands:

$$A \coloneqq \mathbb{C}[s]_{\leq n}, \quad B \coloneqq \mathbb{C}[s]_{\leq n-1}, \quad C \coloneqq \mathbb{C}[s]_{\leq N}, \quad D \coloneqq \mathbb{C}[s]_{\leq N-1}, \quad G \coloneqq \mathbb{C}[s]_{\leq N-n-1}.$$

Additionally, $A^\circ \coloneqq A \setminus 0$, and we use, for example, $A^\circ BG$ to denote the product $(A \setminus 0) \times B \times G$. Consider the incidence variety consisting of tuples (a, b, c, d, g) satisfying the critical equations.

Definition 3.1. The *Walsh variety* is $\mathcal{W} \coloneqq \{\, (a, b, c, d, g) \mid ad - bc = \tilde{a}^2 g \,\} \subseteq A^\circ BCDG$.

In this section, we use the Walsh variety to derive an exact count of the critical points of the model reduction problem under consideration. Our approach relies heavily on tools from algebraic geometry; to keep the exposition self-contained and accessible to a broader engineering audience, the required background is presented in Appendix A. The overall strategy can be summarized as follows:

We first analyze the incidence variety $\mathcal{W}$, whose variables naturally split into two groups: the *input variables* (c, d), specifying a problem instance, and the *unknown variables* (a, b, g), representing candidate solutions. While the fibers of the variety over fixed input data are difficult to analyze directly, the variety often admits a simpler description after fixing a different subset of variables (see e.g., Lemma 3.7 and the proof of Theorem 3.2). This then allows to derive key geometric properties of the incidence variety, such as its dimension, irreducibility, and reducedness (Theorem 3.2).

Once we show that the dimension of the incidence variety coincides with the dimension of the input space, the fiber dimension theorem (Theorem A.1) implies that for generic input data the corresponding fiber consists of (at most) finitely many isolated points (Corollary 3.3). Moreover, if the incidence variety was shown to be defined in a reduced way, then the fiber dimension theorem implies that these generic fibers are reduced as well.

To obtain an explicit solution count, we then exploit the partially linear structure in the system of equations by passing to its MEP formulation. This yields a determinantal description in which the variables that appear only linearly (b, g) are eliminated, at the cost of introducing potential extraneous solutions. These spurious points form what we call the *bad locus*. We show that for generic input data the bad locus is empty (Lemma 3.4), so that the determinantal variety is in fact isomorphic to the generic fibers of the incidence variety over the input data (Corollary 3.5). Moreover, by Lemma A.7, reducedness of these fibers implies reducedness of the determinantal variety. Finally, we apply the Thom–Porteous formula (Theorem A.6) to compute the degree of the determinantal variety, obtaining the desired number of solutions (Theorem 3.6).

Remark. We have deliberately formulated this strategy in a general way since it applies beyond the present setting, and can be used for other systems of equations and MEPs arising in related applications. The only problem-specific ingredients are the analysis of the incidence

variety (cf. Theorem 3.2) and of the bad locus (cf. Lemma 3.4). See, for instance, our related work [KLM26] on Euclidean distance degrees of secant varieties of the rational normal curve. In contrast to the present setting, reducedness, irreducibility, and the dimension of the incidence variety are relatively easily established in that context, whereas the corresponding bad locus is substantially more intricate than the one encountered here.

3.1 Fiber dimension theorem and Thom–Porteous

Before implementing the above strategy on the incidence variety $\mathcal{W}$, we recall from Appendix A the two classical, yet powerful, algebro-geometric tools that will play a central role in our analysis: the fiber dimension theorem and a special case of the Thom–Porteous formula.

The *fiber dimension theorem* (see e.g., [Sha13, Theorem 1.25]) describes (or bounds) the dimension of algebraic varieties. It arose in the context of the broader development of "dimension theory" in commutative algebra and algebraic geometry, which sought to formalize intuitive notions of dimension in the context of geometric objects. The theorem lies at the basis of many dimension-counting arguments and is an everyday tool for algebraic geometers [RM09].

Theorem (Fiber dimension; Theorem A.1)**.** *Let $\phi\colon X \to Y$ be a morphism of irreducible varieties.*

(i) *For every $\boldsymbol{q} \in \phi(X)$, one has $\dim \phi^{-1}(\boldsymbol{q}) \geq \dim X - \dim \overline{\phi(X)}$.*

(ii) *There is a dense open subset $U \subseteq \overline{\phi(X)}$ such that ϕ has reduced fibers at $\boldsymbol{q} \in U$ and equality holds in the first statement.*

The *Thom–Porteous formula* (see e.g., [EH16, Theorem 12.4]), also referred to as the Giambelli-Thom-Porteous formula, originally appeared in [Por71]. It gives an expression for the fundamental class of a degeneracy locus of a morphism of vector bundles in terms of Chern classes—we will clarify each part of this statement. For our purposes, the vector bundle morphism can be viewed concretely as a parameter-dependent matrix (cf. an MEP formulation): for each set of parameter values, the matrix defines a linear map between vector spaces of fixed dimension. In this setting, the degeneracy locus of the morphism corresponds to the determinantal variety associated with this matrix, that is, the set of parameter values for which the matrix drops rank (cf. the so-called eigentuples of the MEP). Such determinantal varieties are often highly singular, and their geometric structure can be difficult to study directly from their defining equations. Instead, the Thom–Porteous formula uses only structural information about the parameter-dependent matrix to compute the characteristic class of the corresponding determinantal variety, thereby providing a quantitative measure of "how large" and "how complicated" the variety is. In the special case where the determinantal variety is known to be zero-dimensional (i.e., it consists of finitely many points), the formula provides a way to count those points. Note that we only state the corank-one case; the general Thom–Porteous formula applies to arbitrary rank conditions, and for more general matrices (e.g., $e > f$ and/or mixed degrees, assuming their minors are still homogeneous).

Theorem (Thom–Porteous formula (corank one); Theorem A.6)**.** *Let $0 < e \leq f$ be integers, and let $\boldsymbol{M}(\boldsymbol{x}) \in \mathbb{C}[\boldsymbol{x}]^{f\times e}$ be a matrix of homogeneous polynomials in $\boldsymbol{x} \in \mathbb{P}^n$, whose columns have constant degree $d_1, \ldots, d_e$. Assume that the determinantal variety of rank $e-1$, $D_{e-1}(\boldsymbol{M})$, has the expected dimension*

$$\dim D_{e-1}(\boldsymbol{M}) = n - (f - e + 1).$$

Then

$$\deg D_{e-1}(\boldsymbol{M}) = T^{f-e+1}\text{-coefficient of } \frac{1}{(1-d_1T)\cdots(1-d_eT)}.$$

3.2 Finiteness of the solution set

Analysis of the defining equations of $\mathcal{W}$ leads to the following characterization, the proof of which is deferred to Section 3.5 in order to avoid interrupting the main line of argument.

Theorem 3.2 (Dimension count)**.** *The Walsh variety $\mathcal{W}$ is an irreducible ideal-theoretic complete intersection in $A^\circ BCDG$ of dimension $2N+2$.*

In practice, we treat (c,d) as the input data and (a,b,g) as the unknowns. For a specific problem instance $(c_0,d_0)\in CD$, the critical points are thus given by the intersection

$$\mathcal{W}_{c_0,d_0} \coloneqq \mathcal{W} \cap \{c=c_0,\, d=d_0\}.$$

Its dimension follows from the fiber dimension theorem.

Corollary 3.3 (Finiteness of the critical points)**.** *Consider the projection*

$$\tau\colon \mathcal{W} \to CD,\ (a,b,c,d,g)\to(c,d).$$

For almost all $(c,d)\in CD$, the fiber $\tau^{-1}(c,d)=\mathcal{W}_{c,d}$ is reduced, of dimension at most 1. If it is non-empty, it consists of tuples (a,b,c,d,g) satisfying

$$\deg a = n, \quad a(s) \text{ has simple roots}, \quad \gcd(a,b)=1. \tag{7}$$

In particular, for generic $H\in\mathcal{M}_N$, the set of complex solutions (a,b,g) to the critical equations (5) is finite up to normalization of a (e.g., fixing $a_n=1$) and is non-degenerate (i.e., consists of simple solutions without multiplicities).

Proof. We distinguish two cases. When τ is not dominant, there is a dense open subset of points $(c,d)\in CD$ lying outside the image of τ, so that for generic choice of (c,d) we have $\tau^{-1}(c,d)=\emptyset$.

Conversely, if τ is dominant, i.e., $\overline{\tau(\mathcal{W})}=CD$, most fibers are not empty. In this case, the fiber dimension theorem (cf. Theorem A.1) implies that generic fibers are reduced of dimension

$$\dim \tau^{-1}(c,d) = \dim \mathcal{W} - \dim CD = (2N+2)-(2N+1)=1.$$

The (negation of the) three additional properties claimed in the corollary are given by polynomial equations in the coefficients of (a,b,g), such as the discriminant $\Delta(a)$ and the resultant[2] $\operatorname{res}_s(a,b)$. Since $\mathcal{W}$ is irreducible and none of these conditions is trivial[3] on $\mathcal{W}$, they form a smaller-dimensional subset; of dimension at most $2N+1$. Let

$$\mathcal{V} \coloneqq \mathcal{W} \cap (\{a_n=0\}\cup\{\Delta(a)=0\}\cup\{\operatorname{res}_s(a,b)=0\})$$

and consider $\pi\colon \mathcal{V}\to CD\colon (a,b,c,d,g)\mapsto(c,d)$. Since the solutions to the defining equations of $\mathcal{V}$ are invariant wrt. the transformation $(a,b,c,d,g)\to(\lambda a,\lambda b,c,d,\frac{1}{\lambda^2}g)$ for any $\lambda\in\mathbb{C}\setminus\{0\}$, all non-empty fibers $\pi^{-1}(c,d)$ have dimension at least one. This then shows by the fiber dimension theorem that the closure of the image $\pi(\mathcal{V})$ has dimension strictly smaller than $\dim\mathcal{V}$, hence $CD\setminus\pi(\mathcal{V})$ contains a dense open set. We can conclude that for a generic choice of (c,d), the fiber $\tau^{-1}(c,d)$ will not intersect $\mathcal{V}$ and thus satisfy the conditions in (7).

Finally, since a generic fiber $\tau^{-1}(c,d)$ is one-dimensional and the condition $a_n\neq 0$ holds, we can normalize by setting $a_n=1$. This reduces the dimension of the generic fiber to 0, which completes the proof. □

Remark. Note that this corollary does not claim the general fiber to be non-empty; though this will follow from Theorem 3.6 below.

[2] The resultant $\operatorname{res}_s(a,b)$ is the determinant of the Sylvester matrix of a and b.

[3] One can easily construct a point $(a,b,c,d,g)\in\mathcal{W}$ that satisfies none of the conditions in (7). For example, the point $a=(s+1)(s+2)\cdots(s+n)$ and $(b,c,d,g)=(1,0,0,0)$.

3.3 Multiparameter eigenvalue problem

The partially linear structure of the critical equations in (6) can be exploited to obtain a multiparameter eigenvalue formulation. Let

$$\boldsymbol{M}(a,c,d) \coloneqq \begin{bmatrix} \boldsymbol{T}_N(a)^\mathsf{T}\boldsymbol{d} & \boldsymbol{T}_n(c)^\mathsf{T} & \boldsymbol{T}_N(\tilde{a})^\mathsf{T}\boldsymbol{T}_{N-n}(\tilde{a})^\mathsf{T} \end{bmatrix},$$

so that the critical equations can be written as the following matrix-vector product:

$$\boldsymbol{T}_N(a)^\mathsf{T}\boldsymbol{d} - \boldsymbol{T}_n(c)^\mathsf{T}\boldsymbol{b} - \boldsymbol{T}_N(\tilde{a})^\mathsf{T}\boldsymbol{T}_{N-n}(\tilde{a})^\mathsf{T}\boldsymbol{g} = \boldsymbol{0} \iff \boldsymbol{M}(a,c,d)\begin{pmatrix} -1 \\ \boldsymbol{b} \\ \boldsymbol{g} \end{pmatrix} = \boldsymbol{0}. \tag{8}$$

This implies that for each $(a,b,c,d,g) \in \mathcal{W}$, the vector $\boldsymbol{T}_N(a)^\mathsf{T}\boldsymbol{d}$ lies in the column space of

$$\boldsymbol{B}(a,c) \coloneqq \begin{bmatrix} \boldsymbol{T}_n(c)^\mathsf{T} & \boldsymbol{T}_N(\tilde{a})^\mathsf{T}\boldsymbol{T}_{N-n}(\tilde{a})^\mathsf{T} \end{bmatrix},$$

and hence that the matrix $\boldsymbol{M}(a,c,d)$ fails to have full column rank at such points.

The values $a \in A^\circ$ for which $\boldsymbol{M}(a,c,d)$ drops rank—equivalently, for which its null space contains a non-trivial vector—are the so-called *eigentuples* of the MEP. For fixed $(c,d) \in CD$, we therefore consider the determinantal variety determined by the maximal minors of $\boldsymbol{M}(a,c,d)$,

$$\mathcal{D} \coloneqq D_N(\boldsymbol{M}(a,c,d)) = \{\, [a] \mid \operatorname{rank}\boldsymbol{M}(a,c,d) \le N \,\} \subseteq \mathbb{P}(A). \tag{9}$$

The projectivization is natural since all entries of the matrix $\boldsymbol{M}(a,c,d)$ are homogeneous in a and this removes the degree of freedom from scaling a.

Consider the projection

$$\phi\colon A^\circ BG \to \mathbb{P}(A), \qquad (a,b,g) \mapsto [a]. \tag{10}$$

It is immediate from (8) that the image $\phi(\mathcal{W}_{c,d})$ is contained in the determinantal variety $\mathcal{D}$. However, $\mathcal{D}$ may contain additional components besides these images of the critical points. These extraneous solutions are given by the closed subset

$$\mathcal{B} \coloneqq \{\, [a] \mid \operatorname{rank}\boldsymbol{B}(a,c) < N \,\} \subseteq \mathbb{P}(A),$$

to which we will refer as the *bad locus*. The bad locus $\mathcal{B}$ consists of those points for which there exists a non-trivial vector in the null space of $\boldsymbol{M}(a,c,d)$ whose first entry vanishes. In this case, the corresponding solutions (a,b,c,d,g) do not satisfy the critical equations (6), but instead solve the alternative system

$$\boldsymbol{T}_n(c)^\mathsf{T}\boldsymbol{b} - \boldsymbol{T}_N(\tilde{a})^\mathsf{T}\boldsymbol{T}_{N-n}(\tilde{a})^\mathsf{T}\boldsymbol{g} = \boldsymbol{0}.$$

Consequently, we obtain the set-theoretic decomposition: $\mathcal{D} = \phi(\mathcal{W}_{c,d}) \cup \mathcal{B}$.

We now proceed by studying the determinantal variety $\mathcal{D}$, since this formulation eliminates the variables (b,g). The following result shows that this reduction is valid: for generic $(c,d) \in CD$, the bad locus $\mathcal{B}$ is empty, and hence $\mathcal{D}$ captures precisely the image of the critical points.

Lemma 3.4 (Bad locus is generically empty)**.** *For given $(a,c) \in A^\circ C$,*

$$\operatorname{nullity}(\boldsymbol{B}(a,c)) = \max\Big\{0,\ \deg\gcd(c,\tilde{a}^2) + \min\{N - \deg c - n,\ n - 2\cdot\deg a\}\Big\}. \tag{11}$$

It follows naturally that if $c \in C$ has N distinct roots, then

$$\operatorname{rank}\boldsymbol{B}(a,c) = N, \quad \forall\, a \in A^\circ,$$

and consequently, $\operatorname{rank}\boldsymbol{M}(a,c,d) \in \{N, N+1\}$.

Proof. We use Bézout's relation to inspect the nature of rank $\boldsymbol{B}(a,c)$. For arbitrary $\boldsymbol{u} \in \mathbb{C}^n$, $\boldsymbol{v} \in \mathbb{C}^{N-n}$,

$$\boldsymbol{B}(a,c)\begin{pmatrix}\boldsymbol{u}\\ \boldsymbol{v}\end{pmatrix} = \boldsymbol{0} \iff cu + \tilde{a}^2 v = 0.$$

Let $f \coloneqq \gcd(c, \tilde{a}^2)$. Pruning f gives $uc' + va' = 0$, where $c = fc'$ and $\tilde{a}^2 = fa'$. For this to be satisfied for nonzero v and u, we need $v = -c'w$ and $u = a'w$ for some polynomial $w \neq 0$. Taking into account the maximum degrees of u and v gives

$$\deg w \leq \min\{\underbrace{(N-n-1)}_{\deg v} - \underbrace{(\deg c - \deg f)}_{\deg c'}, \underbrace{(n-1)}_{\deg u} - \underbrace{(2\cdot\deg a - \deg f)}_{\deg a'}\}.$$

Since $w \neq 0$ is parametrized using $\deg w{+}1$ coefficients, and each choice of w with the appropriate degree constitutes an element in the null space of $\boldsymbol{B}(a,c)$, we conclude that

$$\operatorname{nullity}(\boldsymbol{B}(a,c)) = \max\Big\{0,\ \deg f + \min\{N - \deg c - n,\ n - 2\cdot\deg a\}\Big\}.$$

When $c \in C$ is such that $\deg c = N$, the minimum is upper bounded by $-n$. If in addition, c has distinct roots only, $\deg f \leq n$ by construction, so that $\boldsymbol{B}(a,c)$ has full column rank for all $a \in A^\circ$. In this case $\mathcal{B} = \emptyset$, which completes the proof. □

Corollary 3.5 (Equivalence of the MEP)**.** *Consider the determinantal variety $\mathcal{D}$ and projection ϕ from (9) and (10), respectively. For given $(c,d) \in CD$ with c having N distinct roots, the projection $\mathcal{W}_{c,d} \to A^\circ$ is an isomorphism onto the affine determinantal variety $\widehat{\mathcal{D}} = \{\, a \in A^\circ \mid [a] \in \mathcal{D} \,\}$. Consequently, if (c,d) are generic, then $\mathcal{D}$ is a finite set (finite up to scaling of a) of reduced points; that is, the maximal minors of $\boldsymbol{M}(a,c,d)$ define the elements in a reduced way.*

Proof. Fix $(c,d) \in CD$ with c having N distinct roots. For any $(a,b,g) \in \mathcal{W}_{c,d}$, the defining equations of $\mathcal{W}_{c,d}$ imply that $(-1,b,g)$ is a nonzero null vector of $\boldsymbol{M}(a,c,d)$, so $\boldsymbol{M}(a,c,d)$ is rank-deficient. Conversely, by the assumption on c and Lemma 3.4, we have for any $a \in \widehat{\mathcal{D}}$ that rank $\boldsymbol{B}(a,c) = N$, and hence the null space of $\boldsymbol{M}(a,c,d)$ is one-dimensional containing vectors of the form $(-\lambda, \lambda\boldsymbol{b}, \lambda\boldsymbol{g})^\mathsf{T}$, with $\lambda \in \mathbb{C}\setminus\{0\}$. Upon normalization ($\lambda = 1$), this yields a unique pair (b,g) for which $(a,b,g) \in \mathcal{W}_{c,d}$. This establishes the isomorphism.

Corollary 3.3 further shows that, for generic (c,d), the fiber $\mathcal{W}_{c,d}$ is finite up to scaling of a. Moreover, the same corollary asserts that the critical equations define the generic fibers in a reduced way. By Lemma A.7, this reducedness carries over to the determinantal description. □

Note that in model reduction practice, the conditions from Lemma 3.4 are not restrictive. As (c,d) are treated as the input data, the requirement $c_N \neq 0$ is trivially satisfied. Furthermore, the higher-order model H often originates from a system identification experiment based on noisy measurements. In these cases, the estimated c almost always has simple roots.

Remark. The MEP formulation highlights that the nonlinearity of the model reduction problem is contained in the search for the minimizing values of a; for each such a, the corresponding optimal values for (b,g) can be found uniquely by solving a linear system of equations. This is discussed in more detail in Example 5.1 and is used there to construct the plot in Figure 1.

3.4 Degree of the solution set

To obtain an explicit count of the solutions, it remains to apply the Thom–Porteous formula (Theorem A.6) to the determinantal variety $\mathcal{D}$.

Table 1: Evaluation of the formula from Theorem 3.6 for various problem sizes (N, n). Under the assumption that c, d are generic, so that c has N distinct roots, each entry gives the number of complex solutions to the critical equations of the model reduction problem.

(N,n)	1	2	3	4	5	6	7	8	9	10	11
2	3										
3	5	7									
4	7	17	15								
5	9	31	49	31							
6	11	49	111	129	63						
7	13	71	209	351	321	127					
8	15	97	351	769	1023	769	255				
9	17	127	545	1471	2561	2815	1793	511			
10	19	161	799	2561	5503	7937	7423	4097	1023		
11	21	199	1121	4159	10625	18943	23297	18943	9217	2047	
12	23	241	1519	6401	18943	40193	61183	65537	47103	20481	4095

Theorem 3.6 (Solution count)**.** *Let m denote the number critical points, that is, the number of complex solutions to the critical equations (6). Assume that (c, d) are such that the set $\mathcal{D}$ is finite, then*

$$m \leq \deg \mathcal{D} = \sum_{j=0}^{n} \binom{N-n-1+j}{j} 2^j. \tag{12}$$

Moreover, if (c, d) are generic, such that $\mathcal{B} = \emptyset$, then equality holds.

Proof. The expected dimension of $\mathcal{D}$ corresponds to $n-(N+1-N)(N+n-N) = 0$. Thus, under the assumption that $\mathcal{D}$ is finite, we can apply the Thom–Porteous formula for the corank one case (cf. Theorem A.6). The coefficient of T^n in the series expansion of $(1-T)^{-1}(1-2T)^{-(N-n)}$ can be determined by expressing each factor as a power series. Specifically,

$$\frac{1}{1-T} = \sum_{j=0}^{\infty} T^j, \qquad \frac{1}{(1-2T)^{N-n}} = \sum_{j=0}^{\infty} \binom{N-n+j-1}{j} (2T)^j,$$

such that by the Cauchy product

$$\frac{1}{(1-T)(1-2T)^{N-n}} = \sum_{k=0}^{\infty} \left(\sum_{j=0}^{k} \binom{N-n+j-1}{j} 2^j \right) T^k.$$

In this series, the coefficient of T^n corresponds to

$$\sum_{j=0}^{n} \binom{N-n-1+j}{j} 2^j.$$

If (c, d) are generic, then by Corollary 3.5 the maximal minors define $\mathcal{D}$ in a reduced way, so that $\deg \mathcal{D} = \#\mathcal{D}$, and $\mathcal{D}$ is in bijection with the solutions of the critical equations. □

Several values of the formula from Theorem 3.6 are depicted in Table 1. Notice for example that, if we reduce the order of the given model by one ($n = N - 1$) the formula (12) reduces to $\sum_{j=0}^{n} 2^j = 2^{n+1} - 1$.

Remark. In contrast to the number of complex critical points, the number of real-valued and stable solutions—corresponding to the genuine stationary points of the optimization problem (e.g., local minima, saddle points)—depends on the specific input data c, d and cannot, in general, be determined solely from the problem size (N, n).

3.5 Irreducibility of the Walsh variety

We complete this section by proving Theorem 3.2. We first point out a useful structure in the defining equations of the Walsh variety $\mathcal{W}$. Consider the projection

$$\pi\colon \mathcal{W} \to ABG,\ (a,b,c,d,g) \to (a,b,g).$$

For fixed (a,b,g), the right-hand side of the defining equation of $\mathcal{W}$ is constant, so the unknowns (c,d) are constrained by an affine system of linear equations:

$$\begin{bmatrix} \boldsymbol{T}_N(a)^\mathsf{T} & -\boldsymbol{T}_{N+1}(b)^\mathsf{T} \end{bmatrix} \begin{pmatrix} \boldsymbol{d} \\ \boldsymbol{c} \end{pmatrix} = \boldsymbol{T}_N(\tilde{a})^\mathsf{T} \boldsymbol{T}_{N-n}(\tilde{a})^\mathsf{T} \boldsymbol{g}. \tag{13}$$

Consequently, each fiber of π is an affine linear subspace:

$$\pi^{-1}(a,b,g) = (a,b) \times \{\ (c,d) \mid (13) \text{ is satisfied } \} \times (g). \tag{14}$$

Lemma 3.7. *For given $(a,b,g) \in A^\circ BG$, the fiber $\pi^{-1}(a,b,g)$ is non-empty if and only if the greatest common divisor of (a,b) divides $\tilde{a}^2 g$, i.e., $\gcd(a,b) \mid \tilde{a}^2 g$. In that case, we have that*

$$\begin{aligned} \dim \pi^{-1}(a,b,g) &= N + \deg\gcd(a,b) - \max\{\deg b,\ \deg a - 1\} \\ &= N + \deg\gcd(a^\mathrm{h}, b^\mathrm{h}) - (n-1), \end{aligned}$$

where $a^\mathrm{h} \in \mathbb{C}[s,t]_{\le n}$ and $b^\mathrm{h} \in \mathbb{C}[s,t]_{\le n-1}$ denote the homogenization of the polynomials a and b, respectively.

Proof. Let $a' = a/\gcd(a,b)$ and $b' = b/\gcd(a,b)$ such that $\gcd(a',b') = 1$. It is immediate from

$$\gcd(a,b) \cdot (a'd - b'c) = \tilde{a}^2 g$$

that $\gcd(a,b) \mid \tilde{a}^2 g$ is necessary for the fiber (14) to be non-empty.

We show that it is also sufficient. Let $\boldsymbol{S}(a,b) \coloneqq \begin{bmatrix} \boldsymbol{T}_N(a)^\mathsf{T} & -\boldsymbol{T}_{N+1}(b)^\mathsf{T} \end{bmatrix} \in \mathbb{C}^{(N+n)\times(2N+1)}$. Any element $(\boldsymbol{u}^\mathsf{T}, \boldsymbol{v}^\mathsf{T})^\mathsf{T} \in \mathbb{C}^{2N+1}$ in the null space of $\boldsymbol{S}(a,b)$ corresponds to polynomials $u \in \mathbb{C}[s]_{\le N-1}$ and $v \in \mathbb{C}[s]_{\le N}$ satisfying $au + bv = 0$. Removal of the common factor of a and b gives $a'u + b'v = 0$, which implies $a'u = -b'v$.

Since $\gcd(a',b') = 1$, this can only be satisfied if there exists a polynomial $r \neq 0$ such that $u = b'r$ and $v = -a'r$. Let $k \coloneqq \deg\gcd(a,b)$, then taking the degrees of u and v into account gives

$$\deg r \le \min\{\ \underbrace{(N-1)}_{\deg u} - \underbrace{(\deg b - k)}_{\deg b'},\ \underbrace{(N)}_{\deg v} - \underbrace{(\deg a - k)}_{\deg a'}\ \}.$$

Because r is determined by $(\deg r)+1$ coefficients and each choice of r leads to a unique element in the null space of the matrix $\boldsymbol{S}(a,b)$, we have

$$\begin{aligned} \operatorname{nullity}(\boldsymbol{S}(a,b)) &= \min\{N - \deg b + k,\ N - (\deg a - 1) + k\} \\ &= N + k - \max\{\deg b,\ \deg a - 1\}. \end{aligned}$$

Taking into account the size of $\boldsymbol{S}(a,b)$ gives[4]

$$\operatorname{rank} \boldsymbol{S}(a,b) = N + \max\{\deg b,\ \deg a - 1\} + 1 - k. \tag{15}$$

The image of $\boldsymbol{S}(a,b)$ is the subspace of coefficient vectors corresponding to polynomials $\{\ pa + qb : \deg p \le N-1, \deg q \le N\ \}$. Each element of this set is divisible by $\gcd(a,b)$, so that

$$\operatorname{image} \boldsymbol{S}(a,b) \subseteq \left\{\ \gcd(a,b) \cdot l \ \mid\ l \in \mathbb{C}[s]_{\le N + \max\{\deg b,\ \deg a - 1\} - k}\ \right\}.$$

[4]Note that this rank formula generalizes the familiar fact that the rank of the classical Sylvester matrix of a and b of size $(2n-1)\times(2n-1)$ equals $2n - 1 - \deg\gcd(a,b)$.

We know from (15) that the dimension of image $\boldsymbol{S}(a,b)$ equals $N+\max\{\deg b,\ \deg a-1\}-k+1$, so that this inclusion is in fact equality. This shows that the divisibility constraint is also sufficient. If it is satisfied, the affine space of solutions $\pi^{-1}(a,b,g)$ has dimension equal to the dimension of the null space of $\boldsymbol{S}(a,b)$. □

Now that we have established Lemma 3.7, we can present the proof of Theorem 3.2.

Proof of Theorem 3.2. The proof will proceed in the following steps: First, we stratify $\mathcal{W}$ by the degree of the greatest common divisor of a and b, this will allow us to determine the dimension of $\mathcal{W}$. Then, we show that $\mathcal{W}$ is a set-theoretic complete intersection and irreducible. Lastly, reducedness will follow from a bundle argument, so $\mathcal{W}$ is an ideal-theoretic complete intersection.

For $0 \le k \le n-1$, consider the following two subsets of $A^\circ BG$:

$$\mathcal{S}_k \coloneqq \mathcal{S}_{\ge k} \setminus \mathcal{S}_{\ge k+1},$$

where $\mathcal{S}_{\ge k} \coloneqq \left\{ (a,b,g) \;\middle|\; \deg \gcd(a^{\mathrm{h}}, b^{\mathrm{h}}) \ge k \right\}$, and

$$\mathcal{T}_k \coloneqq \left\{ (a,b,g) \in \mathcal{S}_k \;\middle|\; \gcd(a,b) \text{ divides } \tilde{a}^2 g \right\}.$$

The defining equations of $\mathcal{S}_{\ge k}$ are given by the vanishing of $(k+1)$-minors of the (square) Sylvester matrix constructed from the polynomials a and b. Within $\mathcal{S}_k$, the divisibility constraint of $\mathcal{T}_k$ is given by the $(k+1)$-minors of the matrix $\boldsymbol{S}(a,b)$ concatenated with the coefficient vector of $\tilde{a}^2 g$.

The variety $\mathcal{S}_{\ge k}$ is the image of the polynomial map

$$\underbrace{(\mathbb{C}[s]_k \setminus 0)}_{f} \times \underbrace{(\mathbb{C}[s]_{n-k} \setminus 0)}_{u} \times \underbrace{\mathbb{C}[s]_{n-k-1}}_{v} \times G \to A^\circ BG, \qquad (f,u,v,g) \mapsto (fu, fv, g),$$

and thus irreducible. Given $(a,b,g) \in \mathcal{S}_{\ge k}$, there are only finitely many factorizations $a = fu$, $b = fv$ up to rescaling f by $\lambda \in \mathbb{C} \setminus \{0\}$, hence, the fibers of this map are 1-dimensional. From this we conclude that $\mathcal{S}_{\ge k}$ is an irreducible variety of dimension $N+n+1-k$.

If $k = 0$, then the divisibility constraint is trivial and $\mathcal{T}_0 = \mathcal{S}_0$. For $k \ge 1$, $\mathcal{T}_k \subseteq \mathcal{S}_k$ is a strict subvariety, since there are tuples $(a,b,g) \in \mathcal{S}_k$ that do *not* satisfy the divisibility constraint, for example,

$$a = (s-1)^n, \quad b = (s-1)^k (s-2)^{n-1-k}, \quad g = (s-3)^{N-n-1}.$$

Therefore $\dim \mathcal{T}_0 = N+n+1$ and $\dim \mathcal{T}_k \le N+n-k$ for $k \ge 1$.

Stratifying $\mathcal{W}$ by the degree of $\gcd(a,b)$ then gives us this (locally closed) decomposition

$$\mathcal{W} = \bigcup_{k=0}^{n-1} \mathcal{W}_k, \qquad \mathcal{W}_k = \{\, (a,b,c,d,g) \in \mathcal{W} \mid \deg \gcd(a,b) = k \,\}.$$

By Lemma 3.7, $\mathcal{W}_k$ is an affine bundle over $\mathcal{T}_k$ of rank at most $N+k-(n-1)$. In particular, each (non-empty) $\mathcal{W}_k$ is irreducible of dimension at most

$$\dim \mathcal{W}_k = \dim \mathcal{T}_k + \operatorname{rank}_{\mathcal{T}_k} \mathcal{W}_k \le (N+n-k+1) + (N+k-(n-1)) = 2N+2,$$

where equality holds only for $k = 0$, so that $\dim \mathcal{W} = \max_k \dim \mathcal{W}_k = 2N+2$.

From this we can conclude that $\mathcal{W}$ is a set-theoretic complete intersection:

$$\operatorname{codim}_{A^\circ BCDG} \mathcal{W} = \underbrace{(n+1)+n+(N+1)+N+(N-n)}_{\text{ambient dimension}} - \underbrace{2N+2}_{\dim \mathcal{W}} = \underbrace{N+n}_{\#\text{equations}}.$$

Hence, by Theorem A.2, every irreducible component of $\mathcal{W}$ has dimension exactly $2N+2$.

On the other hand, $\mathcal{W} = \bigcup_{k=0}^{n-1} \overline{\mathcal{W}_k}$ is a union of irreducible subvarieties, so none of the $\mathcal{W}_k$, $k \geq 1$, are components, as they have smaller dimension. In this way we conclude that $\mathcal{W} = \overline{\mathcal{W}_0}$ is irreducible.

Finally, we have established that $\mathcal{W}_0$ is a dense open subset of $\mathcal{W}$, and it is a bundle over the smooth variety $\mathcal{T}_0$, an open subset of $\mathbb{C}^{N+n+1}$. As such, $\mathcal{W}$ is reduced (even smooth) in $\mathcal{W}_0$ and by the unmixedness Theorem A.3, the variety $\mathcal{W}$ is reduced everywhere; meaning that the critical equations define $\mathcal{W}$ in a reduced way. □

4 Further considerations

We present extensions and refinements of the results from Section 3.

4.1 Discrete-time models

All results presented in this paper extend without difficulty to the setting of discrete-time models. Only the notation requires an adjustment; all derivations, formulations, and results remain valid.

In the discrete-time case, the transfer functions under consideration are strictly proper rational functions of the complex variable z, analytic outside the closed unit disk. Given such an Nth order discrete-time model $H \in \mathcal{M}_{N,\mathrm{d}}$ and a desired reduced order $n < N$, the $\mathcal{H}_2$-norm model reduction problem searches for a solution

$$\widehat{H}(z) \in \operatorname*{argmin}_{\widehat{H} \in \mathcal{M}_{\leq n,\mathrm{d}}} \; J_{\mathrm{d}}^2,$$

where the subscript d indicates the discrete-time domain. For example,

$$J_{\mathrm{d}}^2 = \|H(z) - \widehat{H}(z)\|_{\mathcal{H}_2}^2 = \frac{1}{2\pi} \int_0^{2\pi} \left| H(e^{\mathrm{j}\omega}) - \widehat{H}(e^{\mathrm{j}\omega}) \right|^2 \mathrm{d}\omega = \sum_{k=1}^{\infty} \left(h_k - \widehat{h}_k \right)^2.$$

The samples $\{h_k\}_{k=1,\ldots,\infty}$ and $\{\widehat{h}_k\}_{k=1,\ldots,\infty}$ are the impulse responses of H and $\widehat{H}$, respectively.

In this setting, the interpolatory conditions in Theorem 2.1 change: instead of reflecting the poles across the imaginary axis, one now inverts them with respect to the unit circle, so that the interpolation points $-p_i$ are replaced by $1/p_i$. Accordingly, our notation $\tilde{a}(s) = a(-s)$ used in (5) must be adapted to the discrete-time setting and becomes $\tilde{a}(z) = z^n a(z^{-1})$.

A complete treatment of the discrete-time case can be found in [Kay26, Chapter 5].

4.2 Full-order models with multiple poles

If $\deg c = N$ but c has multiple roots, the bad locus $\mathcal{B}$ may be finite or even positive-dimensional. This is illustrated in the following example.

Example 4.1 (Multiple poles). Apply the affine chart $a_1 = 1$ and consider $(N, n) = (2, 1)$ with $c = (s-p)^2$ for some $p \in \mathbb{C}$. In this case, $\boldsymbol{B}(a, c)$ drops rank if and only if $a = (s+p)$. Alternatively, let $(N, n) = (4, 3)$ and $c = (s-p_1)^2(s-p_2)^2$ for some $p_1, p_2 \in \mathbb{C}$. Here, $\mathcal{B}$ consists of polynomials $a = (s+p_1)(s+p_2)$ and $a = (s+p_1)(s+p_2)(s-p_3)$, for arbitrary $p_3 \in \mathbb{C}$. △

If $\mathcal{B}$ is non-empty but finite, the Thom–Porteous formula provides an upper bound on the number of complex critical points. Let $\operatorname{mult}_a I_{N+1}(\boldsymbol{M}(a, c, d))$ denote the multiplicity of a with respect to the ideal generated by the maximal minors of $\boldsymbol{M}(a, c, d)$. The number of critical points m satisfies

$$m = \deg \mathcal{D} - \sum_{[a] \in \mathcal{B}} \operatorname{mult}_a I_{N+1}(\boldsymbol{M}(a, c, d)) \leq \deg \mathcal{D} - \deg \mathcal{B}.$$

Indeed, part of the generic count $\deg \mathcal{D}$ (cf. Theorem 3.6) is accounted for by points in the bad locus; subtracting their contribution decreases the count for the critical points accordingly.

The contribution of each $[a] \in \mathcal{B}$ on $\deg \mathcal{D}$ is governed by its multiplicity with respect to the maximal minors of $\boldsymbol{M}(a, c, d)$, which is in turn bounded below by the multiplicity of the point with respect to the maximal minors of $\boldsymbol{B}(a, c)$.

Special cases in which c leads to a finite but nonempty $\mathcal{B}$ can be analyzed individually. We present one such example below.

Proposition 4.2. *Let m denote the number of complex critical points. If (c, d) is such that $\mathcal{D}$ is finite and $\deg c = N$ but c has exactly one double root, then $\#\mathcal{B} = \binom{N-2}{n-1}$, so that*

$$m \leq \sum_{j=0}^{n} \binom{N-n-1+j}{j} 2^j - \binom{N-2}{n-1}.$$

Proof. Write $c = (s-p_1)^2 \prod_{i=2}^{N-1}(s-p_i)$ with $p_1, \ldots, p_{N-1}$ distinct. The set $\mathcal{D}$ can be split into critical points and bad locus points. From Lemma 3.4 we deduce that $\operatorname{rank} \boldsymbol{B}(a, c) < N$ happens if and only if $\tilde{a} = (s-p_1)\prod_{i\in I}(s-p_i)$, where $I \subseteq \{2, \ldots, N-1\}$ is a selection of $n-1$ roots of c. Thus $\mathcal{B}$ corresponds of exactly $\binom{N-2}{n-1}$ points, potentially with multiplicity. Taking into account the first part of Theorem 3.6 gives the required result. □

5 Numerical examples

We present some (numerical) examples to validate our results. The code is available in a public repository [KL26].

Example 5.1 (Discrete-time). We consider the example given in [LAM23, Example 2], in which the discrete-time model

$$H(z) = \frac{0.0448z^5 + 0.2368z^4 + 0.0013z^3 + 0.0211z^2 + 0.2250z + 0.0219}{z^6 - 1.2024z^5 + 2.3675z^4 - 2.0039z^3 + 2.2337z^2 - 1.0420z + 0.8513},$$

with $\|H(z)\|_{\mathcal{H}_2} = 1.2919$, is approximated by a second-order model $\widehat{H}(z) = \frac{b_0+b_1 z}{a_0+a_1 z+a_2 z^2}$. The authors report that the critical equations admit 49 isolated solutions, exactly matching the value predicted by (12). For this particular problem instance, 11 of the critical points are real-valued, 5 of which corresponded to stable reduced models.

A contour plot of the objective function value J for various reduced models $\widehat{H}$ is shown in Figure 1. We work in the affine chart $a_2 = 1$, and for each choice of the other parameters (a_0, a_1) the corresponding optimal values for (b_0, b_1) are obtained by solving for the unique vector $(\boldsymbol{b}^\mathsf{T}, \boldsymbol{g}^\mathsf{T})^\mathsf{T}$ satisfying the linear system equations:

$$\begin{bmatrix} \boldsymbol{T}_n(c)^\mathsf{T} & \boldsymbol{T}_N(\tilde{a})^\mathsf{T} \end{bmatrix} \begin{pmatrix} \boldsymbol{b} \\ \boldsymbol{g} \end{pmatrix} = \boldsymbol{T}_N(a)^\mathsf{T} \boldsymbol{d}.$$

The contour plot is restricted to the triangular region corresponding to stable models $\widehat{H}$; in the discrete-time setting, stability means that all roots of $a(z)$ lie strictly inside the complex unit disk. The black curve depicts the zero-set of the discriminant $\Delta(a) = a_1^2 - 4a_0a_2$, highlighting the reduced models $\widehat{H}$ that have a double pole. △

Example 5.2 (Higher-order model with multiple poles). Consider the third-order model

$$H = \frac{d(s)}{c(s)} = \frac{s^2 - 9s + 20}{s^3 + 7s^2 + 16s + 12} = \frac{(s-4)(s-5)}{(s+2)^2(s+3)},$$

which has two distinct poles, one of which has multiplicity two. We compute its optimal first-order approximant ($n = 1$). For this setup, the matrix $\boldsymbol{M}(a, c, d)$ corresponds to

$$\boldsymbol{M}(a, c, d) = \begin{bmatrix} 20a_0 & 12 & a_0^2 & 0 \\ -9a_0 + 20a_1 & 16 & -2a_0a_1 & a_0^2 \\ a_0 - 9a_1 & 7 & a_1^2 & -2a_0a_1 \\ a_1 & 1 & 0 & a_1^2 \end{bmatrix}.$$

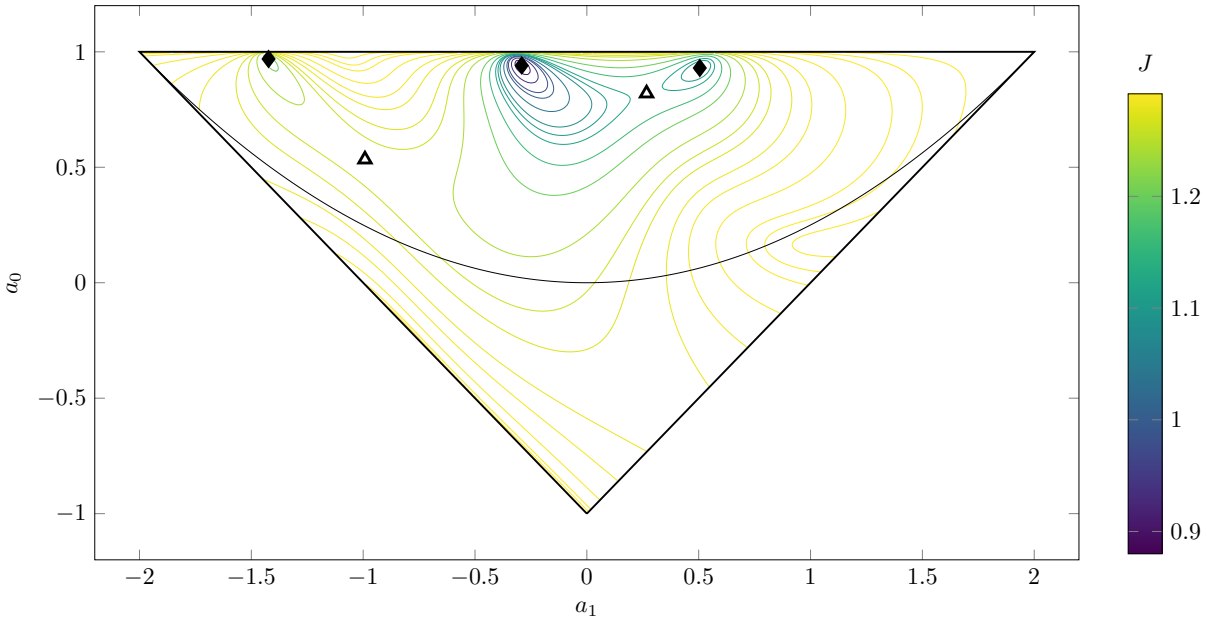

Figure 1: Contour plot of the objective function value J for varying reduced order models $\widehat{H}$. The problem setup corresponds to Example 5.1. The local minimizers (⧫) and saddle points (△) in the displayed region are highlighted.

The determinantal variety $\mathcal{D}$ contains the following approximate values (rounded for displaying purposes) of $(a_0 : a_1)$:

$$\mathcal{D} = \{\, (-2.6300 : 1),\ (-2.0000 : 1),\ (0.3021 : 1),\ (4.4588 : 1),\ (33.8690 : 1) \,\}.$$

Table 1 lists 5 solutions for $(N, n) = (3, 1)$. In our example, however, c fails the genericity condition because it contains a double root. Since $\mathcal{D}$ is finite, we may apply Proposition 4.2 to adjust the generic count: we must subtract $\binom{N-2}{n-1} = \binom{1}{0} = 1$ to account for the non-empty bad locus. Indeed, (symbolic) computation shows that the bad locus $\mathcal{B}$ consists of the single point $(-2 : 1)$. The adjusted count yields 4 critical points.

The bad locus point follows naturally from Lemma 3.4. For our setup, $\deg c = 3$, so that when $\deg a = 0$ the expression (11) for nullity of $\boldsymbol{B}(a, c)$ evaluates to zero. Alternatively, if $\deg a = 1$, a bad locus occurs precisely when $\deg \gcd(c, \tilde{a}^2) = 2$. This happens only when $\tilde{a}(s) = (s + 2)$, which implies $a(s) = (s - 2)$. △

Example 5.3 (Multiplicity of the bad locus points). Consider again the model from the previous example, but now compute the optimal second-order approximant ($n = 2$). The determinantal variety $\mathcal{D}$ consists of 6 isolated points, one of which—the point $(a_0 : a_1 : a_2) = (6 : -5 : 1)$—occurs with multiplicity two with respect to the ideal of maximal minors of $\boldsymbol{M}(a, c, d)$, so that $\#\mathcal{D} = 6$ and $\deg \mathcal{D} = 7$. This last value coincides with the entry in Table 1 for $(N, n) = (3, 2)$.

We consider Lemma 3.4: For our setup, $\deg c = 3$, so that for both $\deg a = 0$ and $\deg a = 1$, the expression (11) evaluates to zero; Alternatively, if $\deg a = 2$, a bad locus occurs precisely when $\deg \gcd(c, \tilde{a}^2) > 2$. This is the case only when $\tilde{a}(s) = (s + 2)(s + 3)$, so that $a(s) = (s - 2)(s - 3) = 6 - 5s + s^2$. As expected, the point in $\deg \mathcal{D}$ occurring with multiplicity belongs to the bad locus. It has multiplicity two with respect to the maximal minors of $\boldsymbol{B}(a, c)$, so that $\#\mathcal{B} = 1$ and $\deg \mathcal{B} = 2$. We conclude that there are $\#\mathcal{D} - \#\mathcal{B} = 5$ critical points.

Proposition 4.2 tells us to adjust the generic count of 7 by subtracting $\binom{N-2}{n-1} = \binom{1}{1} = 1$, giving an upper bound of 6 critical points. Due to the multiplicity of the bad locus point, the inequality in Proposition 4.2 is not tight. △

6 Conclusions and outlook

We illustrated how tools from algebraic geometry can facilitate the analysis of systems of polynomial equations arising in engineering applications. In the specific context of the $\mathcal{H}_2$-norm model order reduction problem, we applied the fiber dimension theorem and the Thom–Porteous formula to obtain an exact a priori count of the complex solutions to the critical equations. This result is substantially sharper than the coarse bounds available from, for instance, Bézout's theorem. Our formula shows that the number of critical points grows rapidly with the problem size, reflecting the computational complexity observed in engineering practice.

The most demanding part of the analysis was proving that the incidence variety $\mathcal{W}$ defined by the critical equations is irreducible, reduced and of the same dimension as the space of input variables. By contrast, the analysis of the bad locus $\mathcal{B}$—corresponding to extraneous solutions introduced by the MEP formulation $\mathcal{D}$—was favorable, as we showed that it is usually empty.

The methodology presented here is not restricted to the model reduction problem under consideration, but can be applied more broadly to other polynomial systems and multiparameter eigenvalue problems arising in related applications. For instance, in our related work [KLM26] we employed a similar strategy in the context of a system identification problem.

In Example 5.2 we exhibit that in specific problem instances, the bad locus can indeed introduce additional *multiple* solutions to the MEP formulation $\mathcal{D}$. A detailed understanding of this multiplicity structure is an interesting direction for future work. Another natural extension is to investigate how the analysis generalizes to frequency weighted $\mathcal{H}_2$-norm model order reduction (see e.g., [HYT01]), where a nontrivial weight function is incorporated in the approximation criterion.

Acknowledgements

We would like to thank Simon Telen, and Fulvio Gesmundo for helpful discussions and feedback on this manuscript. Parts of this article are based on the authors' PhD theses: [Kay26, Chapter 5] and [Lag26, Chapter 5].

This work was supported in part by the KU Leuven Research Fund (projects iBOF/24/064 and IDN/24/013), the Industrial Research Fund and several LRD bilateral industrial projects; in part by the Flemish Government agencies: FWO (TBM T001725N), EWI through the Flanders AI Research Program and VLAIO (CSBO - HBC.2021.0076); and in part by the European Commission (ERC Adv. Grant) under Grant 885682. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or ERC. Neither the European Union nor the granting authority can be held responsible for them. The work of S. Lagauw was supported by the FWO Fundamental Research Fellowship under Grant 11K5625N.

A Relevant background in Algebraic Geometry

This section reviews the algebro-geometric concepts that lie at the basis of our analysis. The exposition is intentionally limited in scope and presents only the essential ideas for our approach in a accessible form. For readers interested in applications, [SW05] provides an excellent introduction to polynomial system solving in the sciences and engineering. A complete reference for basic results in algebraic geometry are the introductory texts [Har95; Sha13; CLO15], while a more advanced treatment can be found in [EH16].

A.1 Affine varieties

Let $R_n := \mathbb{C}[x_1, \ldots, x_n]$ denote the ring of (multivariate) polynomials in the variables $x_1, \ldots, x_n$. An *affine variety* $X \subseteq \mathbb{C}^n$ is the set of all points in $\mathbb{C}^n$ at which a given collection of polynomials

$F = \{f_1, \dots, f_s\} \subseteq R_n$ vanishes simultaneously

$$X = V(F) \coloneqq \{\, \boldsymbol{x} \in \mathbb{C}^n \mid f(\boldsymbol{x}) = 0 \text{ for all } f \in F \,\}, \qquad F \subseteq R_n.$$

Closed subsets of $\mathbb{C}^n$ are precisely the sets of the form $V(F)$ for some $F \subseteq \mathbb{C}[x_1, \dots, x_n]$; their complements are called *open sets*.

A *subvariety* or *closed subset* Y of X is an algebraic subset $Y \subseteq X$, meaning that $X \setminus Y$ is open. The closure $\overline{Y}$ of a subset $Y \subseteq X$ is the smallest closed subset $\overline{Y} \subseteq X$ containing Y. The set Y is *dense* in X if $\overline{Y} = X$. A point $\boldsymbol{x} \in X$ is called *generic* for a given property if it lies in some dense open subset of X, so that "almost all" points of X share that property. The variety X is *irreducible* if $X \neq \emptyset$ and whenever $X = Y_1 \cup Y_2$ with subvarieties $Y_1, Y_2 \subseteq \mathbb{C}^n$, then $X = Y_1$ or $X = Y_2$. Equivalently, every two non-empty open subsets of X intersect.

These closed sets, as defined above, endow $\mathbb{C}^n$ with the so-called *Zariski topology* (see e.g., [Sha13] for a formal treatment). This topology is much coarser than the usual (Euclidean) topology on $\mathbb{C}^n$; for instance, any non-empty Zariski-open subset of an irreducible variety X, e.g., $X = \mathbb{C}^n$, is dense. An open subset of an affine variety is a *quasi-affine variety*.

Every variety X can be uniquely written as a finite union of its *irreducible components*, in such a way that no component is contained in the union of the others (an *irredundant* decomposition):

$$X = X_1 \cup \ldots \cup X_r, \qquad X_i \text{ irreducible, closed subset } X_i \not\subseteq X_j \text{ for all } i, j.$$

The *dimension* of an irreducible variety X is the length of the longest strictly increasing chain of irreducible subvarieties

$$\dim X = \max\{\, d \in \mathbb{N} \mid \exists Y_0 \subsetneq Y_1 \subsetneq \cdots \subsetneq Y_d = X,\ Y_i \text{ irreducible subvarieties} \,\}.$$

By convention $\dim \emptyset \coloneqq -1$. Let X be an arbitrary variety, and let $X = X_1 \cup \cdots \cup X_r$ be its decomposition into irreducible components. Then,

$$\dim X \coloneqq \max_{1 \le i \le r} \dim X_i.$$

It turns out that any chain of irreducible varieties can be extended (by suitably inserting more irreducible subvarieties) to a maximal one, and if X is irreducible, then all such maximal chains have the same length. In particular, any closed proper subvariety of an irreducible variety has strictly smaller dimension than the variety itself.

A variety X is *equidimensional* if every irreducible component $X_i \subseteq X$ has the same dimension. The *codimension* of an irreducible subvariety Y in a variety X is $\operatorname{codim}_X(Y) = \dim X' - \dim Y$, where $X' \subseteq X$ is an irreducible component of largest dimension containing Y. The dimension at a point $\boldsymbol{p} \in X$ is $\operatorname{codim}_X\{\boldsymbol{p}\}$; if X is irreducible then this is just the dimension of X.

A.2 Reducedness and smoothness

The *defining ideal* of a variety X is the set of all polynomials vanishing on X,

$$I(X) = \{\, f \in R_n \mid f(\boldsymbol{x}) = 0 \text{ for all } \boldsymbol{x} \in X \,\}.$$

For a set of polynomials $F \subseteq R_n$, denote by $\langle F \rangle$ the ideal generated by F. Conversely, for an ideal $I \subseteq R_n$, a subset $F = \{f_1, \dots, f_s\} \subset I$ is called a *generating set* of I if $I = \langle F \rangle$.

For any ideal $J \subseteq R_n$, its *radical* $\sqrt{J}$ is defined by

$$\sqrt{J} \coloneqq \left\{\, f \in R_n \mid \exists\, k \in \mathbb{N} : f^k \in J \,\right\}.$$

A set of polynomials $F \subseteq R_n$ is said to define a variety X *in a reduced way* if $V(F) = X$ and the ideal $\langle F \rangle$ is radical, i.e., $\sqrt{\langle F \rangle} = \langle F \rangle$. By *Hilbert's Nullstellensatz* we have

$$I(V(F)) = \sqrt{\langle F \rangle},$$

so if $V(F) = X$, then F defines X in a reduced way if and only if $\langle F \rangle = I(X)$, that is, if F is a *generating set* of the defining ideal $I(X)$.

If $F = \{f_1, \dots, f_m\}$ is a generating set of the defining ideal $I(X)$, then the *smooth locus* of X is

$$X^{\mathrm{sm}} = \{\, \boldsymbol{p} \in X \mid \operatorname{rank} \boldsymbol{J}_F(\boldsymbol{p}) \geq n - \operatorname{codim}_X\{\boldsymbol{p}\} \,\},$$

where $\boldsymbol{J}_F(\boldsymbol{p})$ is the Jacobian matrix of F at $\boldsymbol{p}$:

$$\boldsymbol{J}_F(\boldsymbol{p}) = \left[\frac{\partial f_i}{\partial x_j}(\boldsymbol{p})\right]_{i,j}.$$

The smooth locus is an open dense set $X^{\mathrm{sm}} \subseteq X$, and is isomorphic to a complex manifold of the same dimension as X. A point $\boldsymbol{p} \in X$ is called *singular* if it does not lie in X^{sm}. We denote $X^{\mathrm{sm}}(\mathbb{R}) \coloneqq X^{\mathrm{sm}} \cap \mathbb{R}^n$, that is, the set of smooth points of X with real coordinates. If $\boldsymbol{p} \in X^{\mathrm{sm}}(\mathbb{R})$, then $X^{\mathrm{sm}}(\mathbb{R}) \subseteq \mathbb{R}^n$ is a smooth (real) submanifold in a neighborhood of $\boldsymbol{p}$.

For polynomial equations we understand what it means to be defined in a reduced way. We now turn from this global viewpoint to a local one. Globally, we consider functions defined on the entire variety, and reducedness is expressed by the defining ideal being radical. Locally, however, we focus on the behaviour near a single point. Functions in this local setting are defined only on a small neighbourhood and are described algebraically by localizations of rings, which allow denominators that do not vanish at the point. To formulate the local analogue of reducedness, we therefore begin by clarifying what we mean by local polynomial functions.

Let $U \subset \mathbb{C}^n$ be an open subset and consider $F = \{f_1, \dots, f_s\}$. We denote by $\mathcal{O}(U)$ the ring of *regular functions* on U. These are functions $f : U \to \mathbb{C}$ which can locally be written as a quotient g/h of complex polynomials with h nonvanishing on U. In particular, when $U = \mathbb{C}^n$ itself, every regular function is a polynomial, because a denominator that never vanishes on all of $\mathbb{C}^n$ must be a nonzero constant.

For each $f \in F$ we write $f|_U \in \mathcal{O}(U)$ for its *restriction* to U, obtained by evaluating f only on points of U. We say that F defines X *in a reduced way on* U if the ideal in $\mathcal{O}(U)$ generated by the restrictions $f|_U$ coincides with the defining ideal of $X \cap U$ in $\mathcal{O}(U)$, i.e., every function in $\mathcal{O}(U)$ that vanishes on $X \cap U$ can be written (locally) as an $\mathcal{O}(U)$-linear combination of the $f|_U$.

A.3 Dimension of fibers and intersections

A *morphism* of affine varieties $X \subseteq \mathbb{C}^n$ and $Y \subseteq \mathbb{C}^m$ is a coordinate-wise polynomial map

$$\phi = (\phi_1, \dots, \phi_m) \colon \mathbb{C}^n \to \mathbb{C}^m, \qquad \phi(X) \subseteq Y.$$

Having polynomial coordinate functions is what sets a morphism apart from general maps; it ensures that ϕ_i is defined on all of $\mathbb{C}^n$ and not merely on X, giving morphisms a global character. The morphism is *dominant* if $\overline{\phi(X)} = Y$. The *fiber* $\phi^{-1}(\boldsymbol{q})$ of $\boldsymbol{q} \in Y$ is the subset $\{\, \boldsymbol{x} \in X \mid \phi(\boldsymbol{x}) = \boldsymbol{q} \,\}$. Equations of $\phi^{-1}(\boldsymbol{q})$ are given by

$$I(X) \cup \{\phi_1 - q_1, \dots, \phi_m - q_m\} \subseteq \mathbb{C}[x_1, \dots, x_n],$$

and ϕ has *reduced fibers at* $\boldsymbol{q}$ if this set of equations defines $\phi^{-1}(\boldsymbol{q})$ in a reduced way. An *isomorphism* is a morphism $\phi \colon X \to Y$ that is bijective and whose inverse $\phi^{-1} \colon Y \to X$ is also a morphism; in this case X and Y are said to be algebraically identical.

Theorem A.1 (Fiber dimension theorem [Sha13, Theorem 1.25])**.** *Let $\phi \colon X \to Y$ be a morphism of varieties and X irreducible.*

(i) *For every $\boldsymbol{q} \in \phi(X)$, one has* $\dim \phi^{-1}(\boldsymbol{q}) \geq \dim X - \dim \overline{\phi(X)}$.

(ii) *There is a dense open subset* $U \subseteq \overline{\phi(X)}$ *such that* ϕ *has reduced fibers at* $\boldsymbol{q} \in U$ *and equality holds in the first statement.*

Theorem A.2 (Dimension of intersection [Sha13, Theorem 1.24])**.**

(i) *If* $X, Y \subseteq \mathbb{C}^n$ *are varieties, then every irreducible component* Z *of the intersection* $X \cap Y$ *satisfies*

$$\dim Z \geq \dim X + \dim Y - n.$$

(ii) *If* $f_1, \ldots, f_s \in R_n$*, then* $V(f_1, \ldots, f_s)$ *is empty or each of its irreducible components has dimension* $\geq n - s$*.*

With minor additional technicalities, the theorem remains valid after replacing $\mathbb{C}^n$ by any non-empty open subset $U \subseteq \mathbb{C}^n$.

Consider a non-empty open subset $U \subseteq \mathbb{C}^n$ and a variety $X \subseteq U$ with $\operatorname{codim}_U X = s \leq n$. We say that X is a *set-theoretic complete intersection in* U if there exist a set $\{f_1, \ldots, f_s\} \subseteq R_n$ such that $X = V(f_1, \ldots, f_s) \cap U$. In this case, X (up to taking radicals) can be defined by s equations, and no strictly smaller number of equations suffices to cut out X set-theoretically. If, in addition, $f_1, \ldots, f_s$ generate the defining ideal $I(X)$ of X in $\mathcal{O}(U)$ (i.e., they define X by a radical ideal in U), then X is an *ideal-theoretic complete intersection in* U.

Theorem A.3 (Unmixedness theorem [Eis95, Corollary 18.14])**.** *Let* $X = V(f_1, \ldots, f_s)$ *be a set-theoretic complete intersection. If* X *is defined in a reduced way somewhere on every irreducible component of* X*, then it is reduced everywhere, so actually an ideal-theoretic complete intersection.*

We now introduce affine bundles as a special and useful class of subvarieties. Take a (quasi) affine variety $B \subseteq \mathbb{C}^m$. Fix an $(\ell \times n)$-matrix $\boldsymbol{M}(\boldsymbol{y})$ whose entries are polynomials in the coordinates $\boldsymbol{y} \in B$. Assume that $\operatorname{rank} \boldsymbol{M}(\boldsymbol{y}) = k$ at every point $\boldsymbol{y} \in B$. Choose a polynomial column vector $\boldsymbol{b}(\boldsymbol{y}) \in \mathbb{C}[\boldsymbol{y}]^\ell$ so that for all $\boldsymbol{y} \in B$, the vector $\boldsymbol{b}(\boldsymbol{y})$ lies in the image of $\boldsymbol{M}(\boldsymbol{y})$, ensuring that the linear system $\boldsymbol{M}(\boldsymbol{y})\boldsymbol{x} = \boldsymbol{b}(\boldsymbol{y})$ admits solutions for every $\boldsymbol{y} \in B$. Under these assumptions, let

$$X = \{\, (\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{C}^n \times B \;\mid\; \boldsymbol{M}(\boldsymbol{y})\boldsymbol{x} = \boldsymbol{b}(\boldsymbol{y}) \,\}$$

and consider the projection $\pi : X \to B, (\boldsymbol{x}, \boldsymbol{y}) \to \boldsymbol{y}$. For a chosen point $\boldsymbol{y} \in B$, the fiber

$$\pi^{-1}(\boldsymbol{y}) = \{\, (\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{C}^n \times \mathbb{C}^m \;\mid\; \boldsymbol{M}(\boldsymbol{y})\boldsymbol{x} = \boldsymbol{b}(\boldsymbol{y}) \,\},$$

is an affine linear subspace of $\mathbb{C}^n$ of dimension $r = n - k$, hence isomorphic to $\mathbb{C}^r$. We say that $\pi : X \to B$ is an *affine bundle of rank* r. In that context, X is called the *total space* of the bundle, B is the *base space*, and π is the *bundle projection*.

Establishing that a given morphism $\pi : X \to B$ is a rank r affine bundle is advantageous. Because all fibers have dimension r, we have $\dim X = \dim B + r$. Moreover, the local product structure of the bundle—namely, that X is locally isomorphic to $B \times \mathbb{C}^r$—ensures that properties such as irreducibility, equidimensionality, reducedness, and smoothness pass from the base B to the total space X. For instance, if B is irreducible, then so is X.

A.4 Projective space and determinantal varieties

Projective space $\mathbb{P}^n$ is defined as $(\mathbb{C}^{n+1} \setminus \{0\})$ modulo the equivalence relation that identifies two nonzero vectors $\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{C}^{n+1}$ whenever they are scalar multiples of each other. A projective point $(x_0 : x_1 : \cdots : x_n) \in \mathbb{P}^n$ represents the equivalence class $[\boldsymbol{x}]$ of all nonzero vectors $(\lambda x_0, \lambda x_1, \ldots, \lambda x_n) \in \mathbb{C}^{n+1}$ with $\lambda \in \mathbb{C} \setminus \{0\}$. The notation $\mathbb{P}(V)$ denotes the *projectivization* of a vector space V, i.e., the set of one-dimensional linear subspaces of V. In particular, $\mathbb{P}(\mathbb{C}^{n+1}) = \mathbb{P}^n$.

A *projective variety* X is the zero-set $V(F) \subseteq \mathbb{P}^n$ of a set of homogeneous polynomials $F = \{f_1, \dots, f_s\} \subseteq R_{n+1}$. Its defining ideal $I(X)$ is defined as the defining ideal of the *affine cone* $\widehat{X} = \{\, \boldsymbol{x} \in \mathbb{C}^{n+1} \mid [\boldsymbol{x}] \in X \,\} \cup \{0\}$, and it is generated by all homogeneous polynomials in R_{n+1} vanishing on X. If $X \subseteq \mathbb{P}^n$ is a projective variety and $\widehat{X} \subseteq \mathbb{C}^{n+1}$ denotes its affine cone, then by construction $\dim \widehat{X} = \dim X + 1$ and hence the codimension is preserved:

$$\operatorname{codim}_{\mathbb{P}^n}(X) = n - \dim X = (n+1) - (\dim X + 1) = \operatorname{codim}_{\mathbb{C}^{n+1}}(\widehat{X}).$$

Projective space $\mathbb{P}^n$ can be decomposed as the disjoint union of the hyperplane at infinity $H_\infty \coloneqq V(x_0)$ and its complement

$$U_0 = \{x_0 \neq 0\} = \{\, (1 : x_1 : \cdots : x_n) \mid \boldsymbol{x} \in \mathbb{C}^n \,\},$$

which is called the standard projective chart on $\mathbb{P}^n$. An *affine chart* of a projective variety X is an open subset of X that is identified with an affine variety. In this case, normalizing $x_0 = 1$ gives the isomorphism $U_0 \cong \mathbb{C}^n$.

In this way, we can associate to a projective variety $X \subseteq \mathbb{P}^n$ not contained in H_∞ the affine variety

$$X' = X \cap U_0 = \{\, \boldsymbol{x} \in \mathbb{C}^n \mid (1 : x_1 : \cdots : x_n) \in X \,\}, \qquad I(X') = I(X)|_{x_0=1}.$$

Conversely, if $Y \subseteq \mathbb{C}^n$, then one can consider the closure $\overline{Y} \subseteq \mathbb{P}^n$. Its defining ideal is given by

$$I(\overline{Y}) = \langle f^{\mathrm{h}} = \sum_{i=0}^{d} x_0^{d-i} f^{(i)} \mid f = f^{(0)} + \cdots + f^{(d)} \in I(Y),\ \deg f = d \rangle,$$

where in this case, $f^{(i)}$ denotes the component of f of total degree i. Note, however, that if $Y = V(f_1, \dots, f_s)$, then it is usually *not* the case that $\overline{Y} = V(f_1^{\mathrm{h}}, \dots, f_s^{\mathrm{h}})$, even when the f_i define Y in a reduced way.

A *quasi-projective variety* is an open subset of a projective variety. Quasi-projective varieties include all previously introduced types of varieties. Indeed, if $W \subseteq \mathbb{C}^n$ is an affine variety, and $\overline{W} \subseteq \mathbb{P}^n$ denotes its projective closure, then $W = \overline{W} \cap \mathbb{C}^n$, so an affine variety W is an open subset of the projective variety $\overline{W}$ and hence is quasi-projective. It is the complement of the hyperplane at infinity.

An important advantage of working in projective space is that many equations admit solutions that do not exist in affine space. For example, parallel lines in the affine plane intersect at H_∞. Formally, this principle is captured by the following theorem.

Theorem A.4 ([Sha13, Theorem 1.24])**.** *If $X, Y \subseteq \mathbb{P}^n$ are projective varieties with $\dim X + \dim Y \geq n$, then $X \cap Y \neq \emptyset$ (and $\dim X \cap Y \geq \dim X + \dim Y - n$).*

Determinantal varieties are varieties defined by the vanishing of the minors of a matrix. In this work, we focus on determinantal varieties associated with matrices $\boldsymbol{M}(\boldsymbol{x}) \in R_{n+1}^{f \times e}$ $(e \leq f)$ of homogeneous polynomials in $\boldsymbol{x} \in \mathbb{P}^n$, whose columns have constant degree $d_1, \dots, d_e$:

$$\boldsymbol{M}(\boldsymbol{x}) = (h_{i,j})_{1 \leq i \leq f,\ 1 \leq j \leq e}, \quad h_{i,j} \in R_{n+1} \text{ is homogeneous of degree } d_j. \tag{16}$$

The determinantal variety of rank k associated to $\boldsymbol{M}$ is the projective variety

$$D_k(\boldsymbol{M}) = \{\, \boldsymbol{x} \in \mathbb{P}^n \mid \operatorname{rank} \boldsymbol{M}(\boldsymbol{x}) \leq k \,\}.$$

It is a closed set for which the equations are given by the $k+1$-minors of $\boldsymbol{M}$ (which are homogeneous by assumption on the degrees of the entries of $\boldsymbol{M}$).

Theorem A.5 (Dimension of determinantal varieties)**.** *Let $\boldsymbol{M}$ be the matrix polynomial from (16). For every irreducible component $Z \subseteq D_k(\boldsymbol{M})$ one has*

$$\dim Z \geq n - (e-k)(f-k).$$

Note that for $e = 1$ and $k = 0$, one obtains a projective analogue of Theorem A.2-(ii).

A helpful way to interpret Theorem A.5 is to compare its conclusion with the *expected* geometry of determinantal varieties. In the space of all $(f \times e)$-matrices, the locus of matrices of rank at most k is defined by the $(k+1)$-minors and has codimension $(f-k)(e-k)$, hence dimension $k(f+e-k)$. One can see this dimension directly by parametrizing a rank-k matrix via a k-dimensional row-space, a k-dimensional column space, and an invertible $(k \times k)$-core, which together give $k(f+e-k)$ degrees of freedom. When the entries of the matrix vary independently, this "expected" dimension is achieved exactly. In contrast, when the entries arise from prescribed homogeneous polynomials (e.g., the matrix $\boldsymbol{M}(\boldsymbol{x})$ from (16)) the rank conditions (potentially) impose fewer independent constraints, and the codimension can only become smaller. Theorem A.5 formalizes this: every component of $D_k(\boldsymbol{M})$ has codimension no greater than in the unrestricted case.

A.5 Degree and a special case of Thom–Porteous

If $X = V(F) \subseteq \mathbb{C}^n$ is a finite set, then the quotient space $R_n/\langle F\rangle$ has finite dimension.

The degree of the ideal, $\deg\langle F\rangle$, is the dimension of this quotient space and we define the *degree* of X as

$$\deg X \coloneqq \deg\langle F\rangle.$$

In this case, $\deg X$ corresponds to the number of points in X, counted with multiplicity, and therefore

$$\deg X = \#X \quad \iff \quad F \text{ defines } X \text{ in a reduced way.}$$

When X is infinite (its dimension is positive) the definition above is no longer meaningful. In this case, we pick $\dim X$ generic affine-linear polynomials $\ell_1, \dots, \ell_{\dim X}$ and define

$$\deg X := \deg\langle I(X), \ell_1, \dots, \ell_{\dim X}\rangle.$$

The degree is independent of the particular choice of the ℓ_i, and this construction agrees with the familiar notion of the degree of a hypersurface: if $X = V(f)$, then $\deg X = \deg f$.

If no irreducible component of $X \subseteq \mathbb{P}^n$ is contained in the hyperplane at infinity H_∞, then its degree is the degree of $X \cap \mathbb{C}^n$. The degree (and dimension) of a projective variety can alternatively be characterized via the leading coefficient of the *Hilbert function*:

$$h_X(t) \coloneqq \dim_{\mathbb{C}} \frac{S_t}{I(X)_t} = \frac{\deg X}{(\dim X)!} t^{\dim X} + O(t^{\dim X - 1}),$$

where S_t denotes the homogeneous polynomials of degree t and $I(X)_t$ the subspace of S_t containing all elements vanishing on X. The term $O(t^{\dim X-1})$ is in the standard big-O notation, meaning that the remainder is bounded by a constant multiple of $t^{\dim X-1}$ for all sufficiently large t.

A version of Bézout's theorem is that if $X = V(f_1, \dots, f_n)$ is finite, then

$$\#X \le \deg X = \deg(f_1) \cdots \deg(f_n).$$

This has a vast generalization (cf. [EH16, Theorem 12.4]), of which we consider a special case.

Theorem A.6 (Thom–Porteous formula (corank one) [EH16, Theorem 12.4])**.** *Let $0 < e \le f$ be integers, and let $\boldsymbol{M}(\boldsymbol{x}) \in \mathbb{C}[\boldsymbol{x}]^{f\times e}$ be a matrix of homogeneous polynomials in $\boldsymbol{x} \in \mathbb{P}^n$, whose columns have constant degree $d_1, \dots, d_e$. Assume that the determinantal variety of rank $e-1$, $D_{e-1}(\boldsymbol{M})$, has the expected dimension*

$$\dim D_{e-1}(\boldsymbol{M}) = n - (f - e + 1).$$

Then

$$\deg D_{e-1}(\boldsymbol{M}) = T^{f-e+1}\text{-coefficient of } \frac{1}{(1-d_1T)\cdots(1-d_eT)}.$$

If the dimension of $D_{e-1}(\boldsymbol{M})$ is zero, that is, if the determinantal variety consists of finitely many elements, then this gives a way to count the points $\boldsymbol{x} \in \mathbb{P}^n$ for which $\boldsymbol{M}(\boldsymbol{x})$ is of sub-maximal rank. This count, however, is only correct if we take the multiplicities into account, and is an overcount if the maximal minors of $\boldsymbol{M}$ do not define $D_{e-1}(\boldsymbol{M})$ in a reduced way. A sufficient criterion for the maximal minors to define $D_{e-1}(\boldsymbol{M})$ in a reduced way is the following.

Lemma A.7 ([KLM26, Theorem ??]). *Let $X \subseteq \mathbb{C}^n$ be an affine variety and consider a matrix polynomial $\boldsymbol{M}(\boldsymbol{x}) \in R_n^{f\times e}$, with $e \leq f$. Define a (quasi-projective) variety $\mathcal{K}$ by the f equations $\boldsymbol{M}(\boldsymbol{x})\boldsymbol{v} = \boldsymbol{0}$, that is,*

$$\mathcal{K} = \left\{ (\boldsymbol{x}, [\boldsymbol{v}]) \in X \times \mathbb{P}^{e-1} \;\middle|\; \boldsymbol{M}(\boldsymbol{x}) \cdot \boldsymbol{v} = \boldsymbol{0} \right\}.$$

If these f equations define $\mathcal{K}$ in a reduced way and if $\forall\, \boldsymbol{x} \in X : \operatorname{rank} \boldsymbol{M}(\boldsymbol{x}) \in \{e, e-1\}$, then the e-minors of $\boldsymbol{M}$ define $D_{e-1}(\boldsymbol{M})$ in a reduced way as well.

Specify citation once the other paper has converged.

References

[Als19] F. F. Alsubaie. "$\mathcal{H}_2$ Optimal Model Reduction for Linear Dynamical Systems and the Solution of Multiparameter Matrix Pencil Problems". PhD thesis. Imperial College London, 2019 (cit. on p. 5).

[Ant05] A. C. Antoulas. *Approximation of Large-Scale Dynamical Systems*. SIAM, Philadelphia, PA, 2005 (cit. on p. 5).

[AVM21] O. M. Agudelo, C. Vermeersch, and B. De Moor. "Globally Optimal H_2-Norm Model Reduction: A Numerical Linear Algebra Approach". In: *IFAC PapersOnLine (Special Issue: Proceedings of MTNS 2020)* 54.9 (2021), pp. 564–571 (cit. on pp. 1, 5).

[AW49] P. R. Aigrain and E. M. Williams. "Synthesis of n-Reactance Networks for Desired Transient Response". In: *Journal of Applied Physics* 20.6 (1949), pp. 597–600 (cit. on p. 3).

[Bar86] L. Baratchart. "Existence and Generic Properties of L_2 Approximants for Linear Systems". In: *IMA Journal of Mathematical Control & Information* 3.2-3 (1986), pp. 89–101 (cit. on p. 5).

[BDM12] K. Batselier, P. Dreesen, and B. De Moor. "Prediction error method identification is an eigenvalue problem." In: *IFAC Proceedings Volumes* 45.16 (2012), pp. 221–226 (cit. on p. 1).

[BT18] P. Breiding and S. Timme. "HomotopyContinuation.jl: A Package for Homotopy Continuation in Julia". In: *International Congress on Mathematical Software*. 2018, pp. 458–465 (cit. on p. 1).

[CLO15] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms*. Undergraduate Texts in Mathematics. Springer Cham, 2015 (cit. on p. 16).

[CM21] K. De Cock and B. De Moor. "Multiparameter Eigenvalue Problems and Shift-Invariance". In: *IFAC-PapersOnLine* 54.9 (2021), pp. 159–165 (cit. on p. 1).

[DeM20] B. De Moor. "Least squares optimal realisation of autonomous LTI systems is an eigenvalue problem". In: *Communications in Information and Systems* 20.2 (2020), pp. 163–207 (cit. on p. 1).

[EH16] D. Eisenbud and J. Harris. *3264 and All That: A Second Course in Algebraic Geometry*. Cambridge University Press, 2016 (cit. on pp. 6, 16, 21).

[Eis95] D. Eisenbud. *Commutative Algebra*. Springer New York, 1995 (cit. on p. 19).

[GAB08] S. Gugercin, A. C. Antoulas, and C. Beattie. "$\mathcal{H}_2$ Model Reduction for Large-Scale Linear Dynamical Systems". In: *SIAM Journal on Matrix Analysis and Applications* 30.2 (2008), pp. 609–638 (cit. on pp. 3, 5).

[Har95] J. Harris. *Algebraic Geometry: A first course*. Graduate Texts in Mathematics. Springer, New York, 1995 (cit. on p. 16).

[HKP24] M. E. Hochstenbach, T. Košir, and B. Plestenjak. "Numerical methods for rectangular multiparameter eigenvalue problems, with applications to finding optimal ARMA and LTI models". In: *Numerical Linear Algebra with Applications* 31.2 (2024) (cit. on p. 1).

[HPB12] B. Hanzon, R. Peeters, and I. Bleylevens. "On a Finiteness Result for the Number of Critical Points of the H2 Approximation Criterion". In: *IFAC Proceedings Volumes* 45.16 (2012), pp. 728–733 (cit. on p. 5).

[HYT01] X.-X. Huang, W.-Y. Yan, and K. L. Teo. "A new approach to frequency weighted L_2 optimal model reduction". In: *International Journal of Control* 74.12 (2001), pp. 1239–1246 (cit. on p. 16).

[Kay26] E. L. Kayser. "Computation and Complexity in Algebraic Geometry". PhD thesis. Universität Leipzig, 2026 (cit. on pp. 13, 16).

[KL26] E. L. Kayser and S. Lagauw. *Code repository accompanying the paper "Porteous Meets Walsh: Counting Critical Points for H_2-Norm Model Reduction Using Algebraic Geometry Behind Multiparameter Eigenvalue Problems"*. 2026. URL: https://github.com/leokayser/h2mor (cit. on p. 14).

[KLM26] E. L. Kayser, S. Lagauw, and B. De Moor. *Generic and specific Euclidean distance degrees of secant varieties to the rational normal curve*. Tech. rep. 2026 (cit. on pp. 6, 16, 22).

[Kre09] K. Kreutz-Delgado. *The Complex Gradient Operator and the $\mathbb{CR}$-Calculus*. 2009. URL: https://arxiv.org/abs/0906.4835 (cit. on p. 4).

[Lag26] S. Lagauw. "Global Solution to Problems in System Identification and Model Order Reduction". PhD thesis. KU Leuven, 2026 (cit. on p. 16).

[LAM23] S. Lagauw, O. M. Agudelo, and B. De Moor. "Globally Optimal SISO H_2-Norm Model Reduction Using Walsh's Theorem". In: *IEEE Control Systems Letters* 7 (2023), pp. 1670–1675 (cit. on pp. 2, 5, 14).

[LVM26] S. Lagauw, L. Vanpoucke, and B. De Moor. "Exact Solution to the Least Squares Realization Problem as a Multiparameter Eigenvalue Problem". In: *Automatica* 185 (2026), p. 112792 (cit. on p. 1).

[ML67] L. Meier and D. G. Luenberger. "Approximation of Linear Constant Systems". In: *IEEE Transactions on Automatic Control* 12.5 (Oct. 1967), pp. 585–588 (cit. on pp. 3, 4).

[PBH12] R. L. M Peeters, I. W. M. Bleylevens, and B. Hanzon. "On Algebraic and Linear Algebraic Aspects of Co-Order Three H2 Model Order Reduction". In: *IFAC Proceedings Volumes* 45.16 (2012), pp. 704–709 (cit. on p. 5).

[Por71] J. R. Porteous. "Simple singularities of maps". In: *Singularities—Symposium I*. Ed. by C. T. C. Wall. Vol. 192. Lecture Notes in Mathematics. Springer, 1971, pp. 286–307 (cit. on p. 6).

[Reg95] P. A. Regalia. *Adaptive IIR Filtering in Signal Processing and Control*. Marcel Dekker Inc., New York, NY, 1995 (cit. on p. 4).

[RM09] J. F. Rojas and R. Mendoza. "A note on the fiber dimension theorem". In: *Proyecciones* 28.1 (2009), pp. 57–73 (cit. on p. 6).

[Sha13] I. R. Shafarevich. *Basic Algebraic Geometry 1: Varieties in Projective Space*. 3rd ed. Springer Berlin Heidelberg, 2013 (cit. on pp. 6, 16–20).

[SW05] A. J. Sommese and C. W. Wampler. *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science*. World Scientific, 2005 (cit. on p. 16).

[VD19] C. Vermeersch and B. De Moor. "Globally Optimal Least-Squares ARMA Model Identification Is an Eigenvalue Problem". In: *IEEE Control Systems Letters* 3.4 (2019), pp. 1062–1067 (cit. on p. 1).

[VD23] C. Vermeersch and B. De Moor. "Two Double Recursive Block Macaulay Matrix Algorithms to Solve Multiparameter Eigenvalue Problems". In: *IEEE Control Systems Letters* 7 (2023), pp. 319–324 (cit. on p. 1).

[VM25] C. Vermeersch and B. De Moor. *Solving Multivariate Polynomial Systems and Rectangular Multiparameter Eigenvalue Problems with MacaulayLab*. Tech. rep. Leuven, Belgium: KU Leuven, 2025 (cit. on p. 1).

[Wal60] J. L. Walsh. *Interpolation and Approximation by Rational Functions in the Complex Domain*. 3th. American Mathematical Society, 1960 (cit. on p. 3).