Porteous Meets Walsh: Counting Critical Points for $\mathcal{H}_2$-Norm Model Order 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]Mathematical Institute, University of Bern, Bern.
[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 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 set 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; still, 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 computational methods. It provides a concrete proxy for the expected computational cost and can be used to validate computations by confirming that the full solution set has been recovered.

A *multiparameter eigenvalue problem* (MEP) generalizes the standard eigenvalue problem $(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{v} = \boldsymbol{0}$ by allowing multiple spectral parameters and by dropping the restriction of linearity in the parameters [Atk72; Vol88; DD21]. For rectangular MEPs [VD23; HKP24], one looks for all $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_k) \in \mathbb{C}^k$ and nonzero $\boldsymbol{v} \in \mathbb{C}^m$ such that

$$\boldsymbol{M}(\boldsymbol{\lambda})\boldsymbol{v} = \boldsymbol{0}, \tag{1}$$

where $\boldsymbol{M}(\boldsymbol{\lambda})$ is a tall, $\ell \times m$ matrix whose entries are polynomials in $\boldsymbol{\lambda}$. The solutions $\boldsymbol{\lambda}$ are called *eigentuples* [DD21] and $\boldsymbol{v}$ the corresponding eigenvector. For $k = 1$ and $\boldsymbol{M}$ linear in λ, we recover the (generalized) eigenvalue problem.

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 rectangular MEP, see for example [BDD12; VD19; DeM20; AVD21; LVD26]. In these cases, instead of solving the system directly using general-purpose polynomial system solvers

(e.g., [BT18]), one can alternatively obtain the solutions by computing the eigentuples of the corresponding MEP using specialized solvers (e.g., [VD23; HKP24; VD26]).

In this paper, we argue that the reformulation of the system of polynomial equations as an MEP can also yield more theoretical structural insights into the problem at hand. In particular, two central results in algebraic geometry, when applied in the context of these MEP reformulations, can be used to obtain finer information about the solution set of the polynomial system. When a polynomial system is expected to have finitely many solutions, the *fiber dimension theorem* is a useful tool to verify this. Consequently, where finiteness of the solutions can be established, the *Thom–Porteous formula*, applied to the determinantal variety associated with the MEP reformulation, can often be used to derive an exact count of the affine complex solutions.

To ground our discussion, we illustrate the use of these tools in the context of a problem from systems and control: the $\mathcal{H}_2$-norm model reduction problem for single-input single-output (SISO) linear time-invariant (LTI) models, in which a given model is approximated by one of lower order. It corresponds to a non-convex optimization problem and may therefore admit multiple local optima, which can be characterized as (a subset of the) real-valued common roots of a system of polynomial equations [DVS93; Reg95; HMC07; LAD23]. These equations, commonly referred to as the *critical equations*, follow from *Walsh's Theorem* [Wal60], which, together with the Thom–Porteous formula, motivates the title of this paper.

We study the complex solution set of the critical equations as a whole and refer to its elements as the *critical points*. When the given higher-order model is sufficiently generic, the set of critical points is known to be finite [Bar86, Theorem 3]; in fact, simple poles suffice for this to hold [HPB12, Theorem 3.2]. This observation has motivated several approaches that use algebraic and numerical linear algebra methods to enumerate all such critical points; see, e.g., [HMC07; PBH12; Als19; AVD21; LAD23]. When complete, such enumerations allow one to certify global optimality. However, they are computationally feasible only for relatively small problems and are therefore best regarded as theoretical tools, mostly used for benchmarking purposes. Practical solution approaches, such as the iterative rational Krylov algorithm (IRKA), scale to large problems by iteratively enforcing the first-order optimality conditions, but converge at best to a single local optimum, without any guarantee of global optimality; see, e.g., [GAB08] and references therein. Despite the extensive literature on $\mathcal{H}_2$-norm model reduction, an exact count of the critical points for arbitrary problem sizes is, to the best of our knowledge, not available. For the co-order one case, i.e., when the order is reduced by one, [HMC07] provides an upper bound.

Recent work showed that the critical equations allow for an MEP reformulation [LAD23]. By applying the algebro-geometric tools outlined above, we arrive at the following result:

Theorem (Theorem 3.10)**.** *For generic higher-order models of McMillan degree N, the number of critical points of the model order reduction problem to degree n, $1 \leq n < N$, is*

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

The main contributions of our present work are as follows:

(1) We give a new proof that, for higher-order models with simple poles, the set of critical points is finite (Corollary 3.9). Moreover, we show for generic higher-order models that the solutions of the critical equations occur without multiplicity, a property favorable for numerical computation (Corollary 3.5).

(2) Assuming the higher-order model has only simple poles, we derive an explicit formula for the number of critical points, that is, the total number of complex solutions to the critical equations, as a function of the full and reduced orders (Theorem 3.10). For higher-order models with one or more multiple poles, we illustrate how the aforementioned count can be

used to derive an upper bound on the number of critical points (Section 4.1). Although we consider the continuous-time case throughout, our analysis easily extends to a discrete-time setting (Section 4.2).

(3) We provide a self-contained, application-oriented exposition of the key algebro-geometric tools, making them accessible to a broader engineering audience (Section 3 and Appendix A). These techniques are not specific to the model reduction problem, but are equally applicable to other engineering problems where such MEP formulations arise.

We would like to note that our results connect to a broader line of work in algebraic geometry. Counting of critical points is related to the *Euclidean distance (ED) degree* of an algebraic variety, defined as the number of complex critical points of the squared Euclidean distance from a generic data point to the smooth part of the variety [Dra+16]. In discrete time, the squared $\mathcal{H}_2$-norm of a SISO LTI system is the sum of its squared impulse response coefficients. The discrete-time model reduction problem can be seen as an infinite-dimensional instance of structured low-rank approximation of *Hankel matrices*, whose ED degrees are studied in [OSS14; SP26] and [KL26+]. The latter employs similar algebro-geometric tools as used here, and can be seen as a companion to the present paper.

Outline The remainder of 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 special case where the given higher-order model has multiple poles and the discrete-time setting. We illustrate all theoretical results by means of several 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 $f = \sum_{i=0}^{k} f_i s^i \in \mathbb{C}[s]_{\leq k}$, 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 $f \in \mathbb{C}[s]_{\leq k}$. 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 hf has coefficient vector $\boldsymbol{T}_{k+1}(h)^\mathsf{T}\boldsymbol{f}$. For $f \in \mathbb{C}[s]_{\leq k}$, write $f^\mathrm{h} \coloneqq t^k f(s/t) \in \mathbb{C}[s,t]_k$ for its homogenization to degree k. 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

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 (but see

Section 4 for the discrete-time setting). 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

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

where $c \in \mathbb{R}[s]_{\leq N}$ and $d \in \mathbb{R}[s]_{\leq N-1}$ 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, 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 square-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 $1 \leq n < N$, we seek an approximant $\widehat{H} \in \mathcal{M}_n$, parametrized as

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

where $a \in \mathbb{R}[s]_{\leq n}$ and $b \in \mathbb{R}[s]_{\leq n-1}$ are real polynomials with $a_n \neq 0$, $\gcd(a, b) = 1$, and a Hurwitz stable. The goal of the model order reduction problem under consideration is to minimize the $\mathcal{H}_2$-norm of the *approximation error* $H - \widehat{H}$. Formally, the problem corresponds to

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

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

We restrict the search to reduced models of McMillan degree exactly n, rather than at most n. This is because the set $\mathcal{M}_n$ is open and dense in $\mathcal{M}_{\leq n} := \mathcal{M}_0 \cup \mathcal{M}_1 \cup \cdots \cup \mathcal{M}_n$, where $\mathcal{M}_0 := \{0\}$, so that every model of lower McMillan degree is a limit of models of degree n. Moreover, every local minimizer of (2), when it were to be considered over $\mathcal{M}_{\leq n}$, is guaranteed to have McMillan degree exactly n [Reg95, Theorem 5.3].[1] The set $\mathcal{M}_n$ is a smooth manifold of dimension $2n$: indeed, normalize a to be monic so that $a_n = 1$ and the remaining coefficients of a and b range over an open subset of $\mathbb{R}^{2n}$, obtained by removing those points where a and b share one or more common factors. We call $\widehat{H} \in \mathcal{M}_n$ a *stationary point* of (2) if the gradient of J^2 with respect to these $2n$ coefficients vanishes.

The first-order optimality conditions of (2) can be derived in two complementary ways: analytically, by partial differentiation of the $\mathcal{H}_2$-norm error [AW49; ML67], and geometrically, by exploiting orthogonality in the Hardy space $\mathcal{H}_2$ (see, e.g., [Reg95; GAB08]). 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.

Theorem 2.1 (Meier & Luenberger [ML67])**.** *Consider, for a given $H \in \mathcal{M}_N$, the model reduction problem as formulated in (2). A necessary condition for a reduced model $\widehat{H} \in \mathcal{M}_n$ to be a local minimizer is that, for each pole $\hat{p}_i \in \mathbb{C}$ of $\widehat{H}$, the following interpolation conditions hold:*

$$H^{(k)}(-\hat{p}_i) = \widehat{H}^{(k)}(-\hat{p}_i), \quad k = 0, \ldots, m_i, \tag{3}$$

where m_i is the pole order of $\hat{p}_i$ and $F^{(k)}(z)$ denotes the kth derivative of F, evaluated at z.

[1] The result is stated in [Reg95] for discrete time and is proven by showing that when a reduced model of lower McMillan degree is perturbed by adding a first-order term, then there always exists a perturbation for which the approximation error of the perturbed model decreases. A completely similar argument can be made in the present, continuous-time setting.

Model reduction for linear systems is closely related to the mathematical discipline of function approximation. From this viewpoint it is unsurprising that similar optimality conditions were already known in the field of classical analysis in the 1930s, e.g., in work by Walsh [Wal60]. Adapted to our context of model reduction, the result reads as follows.

Theorem 2.2 (Walsh's theorem)**.** *Consider a higher-order model $H \in \mathcal{M}_N$. An approximant $\widehat{H} \in \mathcal{M}_n$ is a stationary point of the model reduction problem (2) if and only if, as functions,*

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

with $\tilde{a} := \sum_{i=0}^{n} a_i(-s)^i \in \mathbb{R}[s]_{\leq n}$, which has the mirror images $-\hat{p}_i$ of the poles $\hat{p}_i$ as its roots with the same multiplicities, and R a real rational function in $\mathcal{H}_2$.

Theorem 2.2 is the continuous-time counterpart of [Reg95, Theorem 3.14]. It can be derived using a projection argument in the Hardy space $\mathcal{H}_2$ [HMC07, Section II], [Ble10, Section 8.3]. Similar results have been obtained using state-space realizations [VGA10, Theorem 4.5].

Proof sketch. Since $\mathcal{M}_n$ is a smooth submanifold of $\mathcal{H}_2$, the model $\widehat{H}$ is a stationary point if and only if the approximation error $H - \widehat{H}$ is orthogonal to the tangent space of $\mathcal{M}_n$ at $\widehat{H}$.

Normalize (a, b) so that a is monic. Differentiating $\widehat{H} = b/a$ along a perturbation $(a + \epsilon\alpha, b + \epsilon\beta)$, with $\alpha, \beta \in \mathbb{R}[s]_{\leq n-1}$, gives the tangent vector

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon} \frac{b + \epsilon\beta}{a + \epsilon\alpha}\bigg|_{\epsilon=0} = \frac{a\beta - b\alpha}{a^2}.$$

Since $\gcd(a, b) = 1$, the linear map $(\alpha, \beta) \mapsto a\beta - b\alpha$ is injective, and hence, by a dimension count, onto $\mathbb{R}[s]_{\leq 2n-1}$. The tangent space of $\mathcal{M}_n$ at $\widehat{H}$ therefore consists of the rational functions f/a^2 with $f \in \mathbb{R}[s]_{\leq 2n-1}$, regardless of the multiplicities of the poles of $\widehat{H}$. Over $\mathbb{C}$, it is spanned by the functions $1/(s - \hat{p}_i)^k$, $k = 1, \ldots, 2m_i$, where m_i is the order of the pole $\hat{p}_i$.

The function $(\tilde{a}/a)^2$ is analytic in the open right half-plane and has modulus one on the imaginary axis, and by the Beurling–Lax theorem [Lax59], see also [Par04, Theorem 3.1.5], the orthogonal complement in $\mathcal{H}_2$ of functions f/a^2 is the subspace $(\tilde{a}/a)^2\mathcal{H}_2$. Hence the approximation error $H - \widehat{H}$ is orthogonal to the tangent space if and only if $H - \widehat{H} = (\tilde{a}/a)^2 R$ for some $R \in \mathcal{H}_2$. Clearly, R must be a real rational function as well. □

Theorem 2.2 implies that the numerator of $H - \widehat{H} = (ad - bc)/(ac)$ satisfies

$$ad - bc = \tilde{a}^2 Rc/a. \tag{4}$$

The zeros of $\tilde{a}^2$ lie in the open right half-plane, whereas the poles of Rc/a, which are poles of R or zeros of a, lie in the open left half-plane by the stability assumption, so that none of the zeros of $\tilde{a}^2$ can be canceled. Hence, $\widehat{H} \in \mathcal{M}_n$ is a stationary point if and only if $\tilde{a}^2$ divides $ad - bc$, or, equivalently, if and only if the interpolation conditions (3) hold for $k = 0, \ldots, 2m_i - 1$. See Remark 2.3 at the end of this section on the subtle differences between Theorems 2.1 and 2.2.

The left-hand side of (4) is polynomial of degree $N + n - 1$, implying that $\widehat{H} \in \mathcal{M}_n$ is a stationary point if and only if there exists a polynomial $g \in \mathbb{R}[s]_{\leq N-n-1}$ such that

$$ad - bc = \tilde{a}^2 g. \tag{5}$$

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 $ad - bc - \tilde{a}^2 g$. 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 equivalently given by the following matrix expression:

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

We refer to this system of $N+n$ polynomial equations in the $N+n+1$ unknown coefficients a, b, and g as the *critical equations*. Since the system is invariant under scaling $(a, b, g) \mapsto (\lambda a, \lambda b, \lambda^{-1} g)$ for $\lambda \in \mathbb{C} \setminus \{0\}$, one can expect finitely many solutions up to normalizing (e.g., by $a_n = 1$). We call the complex solutions with $a \neq 0$, considered up to scaling, the *critical points*. By Theorem 2.2, real critical points with $\deg a = n$, $\gcd(a, b) = 1$, and a Hurwitz stable are exactly the stationary points of (2); and in particular, they include all local minimizers.

Note that the coefficients of b and g appear only linearly in (6). It is precisely this partially linear structure that is exploited in the corresponding MEP formulation (cf. Section 3.3).

Remark 2.3 (Multiplicity structure). For approximants with simple poles, Walsh's theorem enforces exactly the interpolatory conditions of Theorem 2.1. However, for a pole $\hat{p}_i$ of multiplicity $m_i \geq 2$, it imposes $m_i - 1$ additional conditions. The difference in both formulations stems from the parametrization considered. Meier and Luenberger [ML67] expand $\widehat{H}$ in partial fractions with a prescribed multiplicity structure: each distinct pole of multiplicity m_i is described by a single pole location and m_i residues. This parametrization only covers the models with that particular multiplicity structure, and therefore covers only a subset of $\mathcal{M}_n$ of dimension $\sum_i (m_i + 1)$, where the sum runs over the distinct poles, which is smaller than $2n$ as soon as one of the poles is multiple. Theorem 2.1 describes the stationarity conditions on this submanifold, i.e., the first-order optimality conditions for the model reduction problem constrained to approximants with the prescribed multiplicity structure. The $m_i - 1$ additional conditions in Walsh's theorem correspond to perturbations that split the multiple pole, and thereby leave this submanifold. Consequently, every stationary point of (2) satisfies the conditions of Theorem 2.1, but a reduced model with multiple poles may satisfy them without being a stationary point of (2).

3 Finiteness and degree of the critical points

In our analysis of the critical points, we 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. Our approach relies 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 admits a simpler description after fixing a different subset of variables (see, e.g., Lemma 3.12 and the proof of Theorem 3.4). This then allows us to derive key geometric properties of the incidence variety, such as its dimension, irreducibility, and reducedness (Theorem 3.4).

- 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 3.5). 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 3.7), so that the determinantal variety is in fact isomorphic to the generic fibers of the incidence variety over the input data (Corollary 3.8). Moreover, by Lemma A.8, reducedness of these fibers implies reducedness of the determinantal variety. Finally, we apply the Thom–Porteous formula (Theorem A.7) to compute the degree of the determinantal variety, obtaining the desired number of solutions (Theorem 3.10).

Remark 3.2. We have deliberately formulated the above strategy in a general way since it applies beyond the present setting of model reduction, 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 3.4) and of the bad locus (cf. Lemma 3.7). See, for instance, our related work [KL26+] 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 to characterize 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 two classical, yet powerful, algebro-geometric tools that 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 I.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 [CR25, Section 12.5].

Theorem (Fiber dimension; Theorem A.1)**.** *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 (even smooth!) fibers at $\boldsymbol{q} \in U$ and equality holds in the first statement.*

The second tool, the *Thom–Porteous formula* (see, e.g., [EH16, Theorem 12.4]), also referred to as the Giambelli–Thom–Porteous formula, started with [Gia07], and culminated in its modern form with [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 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 basic invariants of the corresponding determinantal variety, 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.

Example 3.3 (Polynomial eigenvalue problems). For square one-parameter MEPs $\boldsymbol{M}(\lambda)\boldsymbol{v} = \boldsymbol{0}$, better known as polynomial eigenvalue problems, the determinantal variety is exactly the set of roots of the determinant $\det(\boldsymbol{M}(\lambda))$. As the roots of a univariate polynomial, the eigenvalues are a finite set, whose count with multiplicity is determined by the degree of the determinant.

The Thom–Porteous formula gives a way to count solutions beyond polynomial eigenvalue problems. 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.7)**.** *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 each component $Z \subseteq D_{e-1}(\boldsymbol{M})$ of the determinantal variety of rank $e-1$ has the expected dimension $\dim Z = n - (f-e+1) \geq 0$. Then*

$$\deg D_{e-1}(\boldsymbol{M}) = T^{f-e+1}\text{-coefficient of } \frac{1}{(1-d_1T)\cdots(1-d_eT)} \geq 0.$$

3.2 Finiteness of the solution set

Analysis of the defining equations of $\mathcal{W}$ leads to the following structural description.

Theorem 3.4 (Dimension count)**.** *The Walsh variety $\mathcal{W}$ is an irreducible ideal-theoretic complete intersection in $A^\circ BCDG$ of dimension $2N+2$.*

In simple terms, this means that the $N+n$ defining equations cut down the $(3N+n+2)$-dimensional ambient space $A^\circ BCDG$ to the expected dimension $2N+2$, and do so while intersecting transversally almost everywhere. In order to avoid interrupting the main line of argument, we defer the proof to Section 3.5.

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 3.5 (Finiteness and reducedness of the critical points)**.** *Consider the projection*

$$\tau : \mathcal{W} \to CD,\ (a,b,c,d,g) \to (c,d). \tag{7}$$

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

$$\deg a = n, \qquad a \text{ has simple roots}, \qquad \gcd(a,b) = 1. \tag{8}$$

In particular, for generic $H \in \mathcal{M}_N$, the set of complex solutions (a,b,g) to the critical equations (6) is finite up to normalization of a (e.g., fixing $a_n = 1$) and 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 $\tau^{-1}(c,d) = \emptyset$.

Alternatively, 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 negations of the three additional properties claimed in the corollary are enforced by polynomial equations in the coefficients of (a, b, g): equating the coefficient a_n, the discriminant $\Delta(a)$, or Sylvester's resultant $\operatorname{res}_s(a, b)$, respectively, to zero. Since $\mathcal{W}$ is irreducible and none of these conditions is trivial[2] 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 $\tau_{\mathcal{V}} \coloneqq \tau|_{\mathcal{V}} \colon \mathcal{V} \to CD \colon (a, b, c, d, g) \mapsto (c, d)$. Since the solutions to the defining equations of $\mathcal{V}$ are invariant w.r.t. the transformation $(a, b, c, d, g) \to (\lambda a, \lambda b, c, d, \lambda^{-1} g)$ for any $\lambda \in \mathbb{C} \setminus \{0\}$, all non-empty fibers $\tau_{\mathcal{V}}^{-1}(c, d)$ have dimension at least one. Applying the fiber dimension theorem to every component of $\mathcal{V}$, one sees that $\dim \overline{\tau(\mathcal{V})} \leq 2N < \dim CD$, hence $CD \setminus \tau(\mathcal{V})$ contains a dense open set. We can conclude that for a generic choice of (c, d), the fiber $\tau^{-1}(c, d)$ does not intersect $\mathcal{V}$ and thus a, b, g satisfy the conditions in (8).

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 zero, which completes the proof. □

Remark 3.6. Note that this corollary does not claim the general fiber to be non-empty; though this will follow from Theorem 3.10 below. In fact, it is known that there always exists a stable global minimizer of full McMillan degree exactly n [Bar86, Theorem 1 & Proposition 1]. Baratchart works in the discrete-time setting, which is isometrically equivalent to the continuous-time setting (cf. Section 4.2 for details).

3.3 Multiparameter eigenvalue problem

The partially linear structure of the critical equations in (6) can be exploited to reformulate the critical equations as a multiparameter eigenvalue problem [LAD23]. Let

$$\boldsymbol{M}(a, c, d) \coloneqq \begin{bmatrix} \boldsymbol{T}_N(a)^{\mathsf{T}} \boldsymbol{d} & \boldsymbol{T}_n(c)^{\mathsf{T}} & \boldsymbol{T}_N(\tilde{a})^{\mathsf{T}} \boldsymbol{T}_{N-n}(\tilde{a})^{\mathsf{T}} \end{bmatrix} \in \mathbb{C}[\boldsymbol{a}, \boldsymbol{c}, \boldsymbol{d}]^{(N+n)\times(N+1)},$$

so that the critical equations correspond to the following matrix-vector product:

$$\boldsymbol{T}_N(a)^{\mathsf{T}} \boldsymbol{d} - \boldsymbol{T}_n(c)^{\mathsf{T}} \boldsymbol{b} - \boldsymbol{T}_N(\tilde{a})^{\mathsf{T}} \boldsymbol{T}_{N-n}(\tilde{a})^{\mathsf{T}} \boldsymbol{g} = \boldsymbol{0} \iff \boldsymbol{M}(a, c, d) \begin{pmatrix} -1 \\ \boldsymbol{b} \\ \boldsymbol{g} \end{pmatrix} = \boldsymbol{0}. \tag{9}$$

It follows 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 $\boldsymbol{M}(a, c, d)$ fails to have full column rank at such points.

Fix input data $(c, d) \in CD$ and consider the determinantal variety determined by the maximal minors of $\boldsymbol{M}(a, c, d)$,

$$\mathcal{D} \coloneqq D_N(\boldsymbol{M}(a, c, d)) = \{\, [a] \mid \operatorname{rank} \boldsymbol{M}(a, c, d) \leq N \,\} \subseteq \mathbb{P}(A). \tag{10}$$

Projectivizing is natural since all entries of $\boldsymbol{M}(a, c, d)$ are homogeneous in a and this removes the scaling freedom in a. The points of $\mathcal{D}$ are the eigentuples of the MEP and can be computed using multiparameter eigenvalue solvers (e.g., [VD23; HKP24; VD26]).

It remains to show that the finiteness and reducedness of the generic fibers of $\mathcal{W}$ (Theorem 3.4 and Corollary 3.5) carry over to the determinantal variety $\mathcal{D}$. Define

$$\phi \colon A^\circ BG \to \mathbb{P}(A), \qquad (a, b, g) \mapsto [a]. \tag{11}$$

[2] One can easily construct a point $(a, b, c, d, g) \in \mathcal{W}$ that satisfies all of the conditions in (8). For example, the point $a = (s + 1)(s + 2) \cdots (s + n)$ and $(b, c, d, g) = (1, 0, 0, 0)$.

It is immediate from (9) that $\phi(\mathcal{W}_{c,d}) \subseteq \mathcal{D}$. However, $\mathcal{D}$ may contain additional components, corresponding to the closed subset

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

which we refer to as the *bad locus*. This locus consists of points for which there exists a non-trivial vector in the null space of $\boldsymbol{M}(a,c,d)$ whose first entry vanishes. 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 get

$$\mathcal{D} = \phi(\mathcal{W}_{c,d}) \cup \mathcal{B}. \tag{12}$$

The following result shows that the bad locus $\mathcal{B}$ is in fact empty whenever c has N distinct roots. Under this assumption, $\mathcal{D}$ captures precisely the image of the critical points.

Lemma 3.7 (Bad locus)**.** *Consider $(a,c) \in A^\circ C^\circ$ and let $\tilde{a}^{\mathrm{h}}$ and c^{h} denote the homogenization of the polynomials $\tilde{a}$ and c, respectively. Then,*

$$\begin{aligned} \operatorname{nullity} \boldsymbol{B}(a,c) &= \max\Big\{0,\ \deg\gcd(c,\tilde{a}^2) + \min\big\{N - \deg c - n,\ n - 2\cdot\deg a\big\}\Big\} \\ &= \max\Big\{0,\ \deg\gcd(c^{\mathrm{h}},(\tilde{a}^{\mathrm{h}})^2) - n\Big\}. \end{aligned} \tag{13}$$

In particular, if $c \in C$ has N distinct roots, then $\operatorname{rank} \boldsymbol{B}(a,c) = N$ for all $a \in A^\circ$. In that case, $\mathcal{B} = \emptyset$ and $\operatorname{rank} \boldsymbol{M}(a,c,d) \in \{N, N+1\}$.

The proof uses similar ideas to Bézout's lemma on polynomial greatest common divisors.

Proof. Identifying $\boldsymbol{u} \in \mathbb{C}^n$ and $\boldsymbol{v} \in \mathbb{C}^{N-n}$ with u^{h} and v^{h}, homogeneous polynomials of degree $n-1$ and $N-n-1$, respectively, we have

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

Let $f := \gcd(c^{\mathrm{h}},(\tilde{a}^{\mathrm{h}})^2)$ and write $c^{\mathrm{h}} = fc'$ and $(\tilde{a}^{\mathrm{h}})^2 = fa'$, so that $\gcd(c',a') = 1$. Dividing by f gives $uc' + va' = 0$. Since a' divides $c'u$ and is coprime to c', it divides u. Hence $u = a'w$ and $v = -c'w$ for a $w \in \mathbb{C}[s,t]$ of degree $(n-1) - (2n - \deg f) = \deg f - n - 1$. Conversely, every such w yields a null vector of $\boldsymbol{B}(a,c)$ via the linear map $w \mapsto (a'w, -c'w)$. This map is hence an isomorphism from $\mathbb{C}[s,t]_{\deg f - n - 1}$ onto the null space of $\boldsymbol{B}(a,c)$. Since $\mathbb{C}[s,t]_k$ has dimension $k+1$ for $k \geq 0$ and $\mathbb{C}[s,t]_k = \{0\}$ for $k < 0$, the nullity equals $\max\{0, \deg f - n\}$.

If c has N distinct roots $p_1, \ldots, p_N$, then $c^{\mathrm{h}} = c_N \prod_{i=1}^N (s - p_i t)$. Since f divides c^{h}, it is a product of distinct factors $s - p_i t$, each of which divides $(\tilde{a}^{\mathrm{h}})^2$. Each factor being prime, it also divides $\tilde{a}^{\mathrm{h}}$; it follows that f divides $\tilde{a}^{\mathrm{h}}$. In particular $\deg f \leq n$, and the nullity in (13) vanishes. □

By Lemma 3.7, the bad locus is empty whenever c has N distinct roots, so that every point of $\mathcal{D}$ corresponds to a critical point.[3] The following corollary makes this precise: under the assumption, the critical equations and the determinantal variety describe the same solution set, preserving its reducedness, so that counting points of $\mathcal{D}$ is equivalent to counting the critical points.

[3] In model reduction practice, where (c,d) correspond to 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, so that the estimated c almost always has simple roots.

Corollary 3.8 (Equivalence of the MEP)**.** *Consider input data $(c,d) \in CD$ with c having N distinct roots. Then, 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 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 3.7, we have for any $a \in \widehat{\mathcal{D}}$ that $\operatorname{rank} \boldsymbol{B}(a,c) = N$, and hence the null space of $\boldsymbol{M}(a,c,d)$ is one-dimensional containing vectors of the form $(-\kappa, \kappa\boldsymbol{b}, \kappa\boldsymbol{g})^\mathsf{T}$, with $\kappa \in \mathbb{C} \setminus \{0\}$. Upon normalization ($\kappa = 1$), this yields a unique pair (b,g) for which $(a,b,g) \in \mathcal{W}_{c,d}$. Hence the inverse map $a \mapsto (a,b,g)$ is a morphism, given explicitly by Cramer's rule. This establishes the isomorphism.

Corollary 3.5 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.8, this reducedness carries over to the determinantal description. □

In fact, this argument yields a proof of finiteness of the critical locus for *arbitrary* higher-order models of McMillan degree exactly N and distinct poles. The following result sharpens Corollary 3.5, and yields a proof independent of [HPB12, Theorem 3.2]; note that it does not rely on Theorem 3.4 and hence is rather direct.

Corollary 3.9. *If c has N distinct roots and $\gcd(c,d) = 1$, then $\mathcal{D}$ is a finite set.*

Proof. Since c has N distinct roots, Corollary 3.8 provides the inverse morphism $\widehat{\mathcal{D}} \to \mathcal{W}_{c,d}$, $a \mapsto (a,b,g)$, defined on all of $\widehat{\mathcal{D}}$. For $0 \le i \le n$ and $0 \le j \le N-n-1$, consider the composite map

$$\varphi_{ij} \colon \widehat{\mathcal{D}} \to \mathcal{W}_{c,d} \to \mathbb{C}, \qquad a \mapsto (a,b,g) \mapsto a_i \cdot g_j.$$

Scaling a by $\lambda \in \mathbb{C}\setminus\{0\}$ scales g by λ^{-1}, that is, $\varphi_{ij}(\lambda a) = \varphi_{ij}(a)$, so φ_{ij} is a well-defined regular function on $\mathcal{D}$. Since $\mathcal{D} \subseteq \mathbb{P}(A)$ is a projective variety, Theorem A.4 indicates that φ_{ij} must be *constant* on each irreducible component of $\mathcal{D}$.

If $[a] \in \mathcal{D}$, then g is not identically zero, since by (5) we would otherwise have $ad = bc$, so $c \mid a$ by $\gcd(c,d) = 1$, contradicting $\deg c = N > n \ge \deg a$ (and $a \neq 0$). In particular, for any $[a] \in \mathcal{D}$ there exist i,j such that $\varphi_{ij}(a) \neq 0$. In a neighborhood of $[a]$ in $\mathcal{D}$, the coordinate ratio

$$[a] \mapsto \frac{a_k}{a_i} = \frac{a_k \cdot g_j}{a_i \cdot g_j} = \frac{\varphi_{kj}(a)}{\varphi_{ij}(a)}, \qquad k = 0, \dots, n$$

is thus constant, so $[a]$ is isolated in $\mathcal{D}$. This shows that $\mathcal{D} \subseteq \mathbb{P}(A)$ is discrete, hence finite. □

3.4 Degree of the solution set

To obtain an explicit count of the critical points, it remains to apply the Thom–Porteous formula (Theorem A.7) to the determinantal variety $\mathcal{D}$.

Theorem 3.10 (Solution count)**.** *Consider input data $(c,d) \in CD$ such that c has N distinct roots and $\gcd(c,d) = 1$, and let m denote the number of critical points, that is, the number of complex solutions to the critical equations (6) up to scaling. We have*

$$m = \#\mathcal{D} \le \deg \mathcal{D} = \sum_{j=0}^{n} \binom{N-n-1+j}{j} 2^j \tag{14}$$

and equality holds for generic (c,d).

Proof. The expected dimension of $\mathcal{D}$ is $n - (N+1-N)(N+n-N) = 0$. By Corollary 3.9, $\mathcal{D}$ is indeed finite, so that each component has the expected dimension zero, and we can apply the Thom–Porteous formula for the corank one case (cf. Theorem A.7). 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

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

By Corollary 3.8, $\mathcal{D}$ is in bijection with the critical points, so that $m = \#\mathcal{D} \leq \deg \mathcal{D}$. The same corollary shows that for generic (c, d), the maximal minors define $\mathcal{D}$ in a reduced way, so that $\#\mathcal{D} = \deg \mathcal{D}$. □

Several values of the formula from Theorem 3.10 are depicted in Table 1. If we consider the co-order one case $n = N - 1$, then the formula (14) reduces to $\sum_{j=0}^{n} 2^j = 2^{n+1} - 1$, which corresponds exactly to the upper bound obtained in [HMC07].

Table 1: Evaluation of the formula from Theorem 3.10 for various problem sizes (N, n). Under the assumption that c, d are generic, 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
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

To put these numbers in perspective, we compare (14) with the alternative bound provided by Bézout's theorem. Applying that result to the critical equations (6), in the chart $a_n = 1$, gives an upper bound of $2 \cdot 3^{N+n-2}$ complex solutions, since one equation is linear, one is quadratic, and the remaining $N + n - 2$ are cubic. For $(N, n) = (6, 2)$, the Bézout and Thom–Porteous numbers are 1458 and 49, respectively; for $(N, n) = (10, 5)$, they are 3 188 646 and 5503.

Remark 3.11. In contrast to the number of complex critical points, the number of real-valued and Hurwitz 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 3.4. We first point out a useful structure in the defining equations of the Walsh variety $\mathcal{W}$. Consider the projection

$$\pi\colon \mathcal{W} \to A^{\circ}BG,\ (a,b,c,d,g) \to (a,b,g). \tag{15}$$

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:

$$\begin{bmatrix} \boldsymbol{T}_N(a)^{\mathsf{T}} & -\boldsymbol{T}_{N+1}(b)^{\mathsf{T}} \end{bmatrix} \begin{pmatrix} \boldsymbol{d} \\ \boldsymbol{c} \end{pmatrix} = \boldsymbol{T}_N(\tilde{a})^{\mathsf{T}} \boldsymbol{T}_{N-n}(\tilde{a})^{\mathsf{T}} \boldsymbol{g}. \tag{16}$$

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

$$\pi^{-1}(a,b,g) = \{(a,b,g)\} \times \{\ (c,d) \mid (16) \text{ is satisfied}\ \}. \tag{17}$$

Lemma 3.12. *For given $(a,b,g) \in A^{\circ}BG$, the fiber $\pi^{-1}(a,b,g)$ is non-empty if and only if $\gcd(a,b)$ divides $\tilde{a}^2 g$. In that case, we have*

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

where we use the convention $\deg 0 = -\infty$ and, as usual, $\gcd(a,0) = a$.

The notation and construction in the proof is similar to Lemma 3.7.

Proof. Let $f \coloneqq \gcd(a,b)$ and $k \coloneqq \deg f$, and write $a' \coloneqq a/f$ and $b' \coloneqq b/f$ so that $\gcd(a',b') = 1$. It is immediate from

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

that $f \mid \tilde{a}^2 g$ is necessary for the fiber (17) to be non-empty.

We show that it is also sufficient. Let $\boldsymbol{S}(a,b) \coloneqq \begin{bmatrix} \boldsymbol{T}_N(a)^{\mathsf{T}} & -\boldsymbol{T}_{N+1}(b)^{\mathsf{T}} \end{bmatrix} \in \mathbb{C}^{(N+n)\times(2N+1)}$. An 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 gives $a'u = b'v$, and since $\gcd(a',b') = 1$, this holds if and only if there exists a polynomial r such that $u = b'r$ and $v = a'r$. Taking the degrees of u and v into account gives

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

so that $r \mapsto (b'r, a'r)$ is an isomorphism from $\mathbb{C}[s]_{\le\rho}$ onto the null space of $\boldsymbol{S}(a,b)$. Hence we obtain the first expression in the lemma:

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

Since $\gcd(a^{\mathrm{h}}, b^{\mathrm{h}}) = t^{\min\{n-\deg a, n-1-\deg b\}} f^{\mathrm{h}}$ (where f^{h} is homogenized to degree $\deg f$), we have

$$\deg\gcd(a^{\mathrm{h}}, b^{\mathrm{h}}) = k + n - 1 - \max\{\deg b, \deg a - 1\},$$

which gives the second expression for the dimension in the lemma. Taking into account the size of $\boldsymbol{S}(a,b)$, we obtain[4]

$$\operatorname{rank} \boldsymbol{S}(a,b) = N + \max\{\deg b, \deg a - 1\} - k + 1 = N + n - \deg\gcd(a^{\mathrm{h}}, b^{\mathrm{h}}). \tag{18}$$

[4]Note that (18) generalizes the basic fact that the Sylvester matrix of a and b has nullity equal to $\deg\gcd(a^{\mathrm{h}}, b^{\mathrm{h}})$.

The image of $\boldsymbol{S}(a,b)$ is the subspace of coefficient vectors corresponding to polynomials $pa - qb$ for suitable $p \in \mathbb{C}[s]_{\leq N-1}$ and $q \in \mathbb{C}[s]_{\leq N}$. Each such polynomial is divisible by f and has degree at most $N + \max\{\deg b, \deg a - 1\}$, so that

$$\operatorname{image} \boldsymbol{S}(a,b) \subseteq \left\{ f \cdot h \;\middle|\; h \in \mathbb{C}[s]_{\leq N+\max\{\deg b,\, \deg a-1\}-k} \right\}.$$

By (18), both sides have dimension $N+\max\{\deg b,\ \deg a-1\}-k+1$, so that this inclusion is in fact an equality. Moreover, since $\deg a \leq n$, we have $\deg(\tilde{a}^2 g) \leq 2\deg a + N - n - 1 \leq N + \deg a - 1$. Hence, if f divides $\tilde{a}^2 g$, then $\tilde{a}^2 g$ lies in the image of $\boldsymbol{S}(a,b)$, which shows that the divisibility constraint is also sufficient. If it is satisfied, the affine space of solutions $\pi^{-1}(a,b,g)$ has dimension nullity $\boldsymbol{S}(a,b)$. □

Now that we have established Lemma 3.12, we can present the proof of Theorem 3.4. It proceeds in the following steps. First, we stratify $\mathcal{W}$ by the degree of the greatest common divisor of a and b, this allows us to determine the dimension of $\mathcal{W}$. Then, we show that $\mathcal{W}$ is a set-theoretic complete intersection and irreducible. Lastly, reducedness follows from a bundle argument, so $\mathcal{W}$ is an ideal-theoretic complete intersection.

Proof of Theorem 3.4. For $0 \leq k \leq n$, consider the following subsets of $A^\circ BG$:

$$\mathcal{S}_{\geq k} \coloneqq \left\{ (a,b,g) \;\middle|\; \deg \gcd(a^{\mathrm{h}}, b^{\mathrm{h}}) \geq k \right\}, \qquad \mathcal{S}_k \coloneqq \mathcal{S}_{\geq k} \setminus \mathcal{S}_{\geq k+1},$$

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

The set $\mathcal{S}_{\geq k}$ is defined by the vanishing of the $(2n-k)$-minors of the Sylvester matrix constructed from the polynomials a^{h} and b^{h}. On $\mathcal{S}_k$, we have $\operatorname{rank} \boldsymbol{S}(a,b) = N + n - k$ by (18), and, by the proof of Lemma 3.12, the divisibility constraint holds if and only if the coefficient vector of $\tilde{a}^2 g$ lies in the column space of $\boldsymbol{S}(a,b)$. Within $\mathcal{S}_k$, the set $\mathcal{T}_k$ is thus defined by the vanishing of the $(N+n-k+1)$-minors of the matrix $\boldsymbol{S}(a,b)$ concatenated with the coefficient vector of $\tilde{a}^2 g$.

Using the convention $\mathbb{C}[s]_{\leq -1} \coloneqq \{0\}$, the variety $\mathcal{S}_{\geq k}$ is the image of the polynomial map

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

The image of this map is exactly $\mathcal{S}_{\geq k}$, which is therefore irreducible.

Given $(a,b,g) \in \mathcal{S}_{\geq k}$, 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, so that $\mathcal{S}_{\geq k}$ is an irreducible variety of dimension $N + n + 1 - k$ (Theorem A.1). Since $\mathcal{S}_{\geq k+1}$ has smaller dimension, $\mathcal{S}_k$ is a non-empty open subset of $\mathcal{S}_{\geq k}$, and thus irreducible of the same dimension.

If $k = 0$, then the divisibility constraint is trivial and $\mathcal{T}_0 = \mathcal{S}_0$, a non-empty open subset of $A^\circ BG$. For $k \geq 1$, $\mathcal{T}_k$ is a proper closed subset of $\mathcal{S}_k$, since there are tuples $(a,b,g) \in \mathcal{S}_k$ that do *not* satisfy the divisibility constraint, for example,

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

for $1 \leq k \leq n-1$, and similarly for $k = n$. Therefore $\dim \mathcal{T}_0 = N+n+1$ and $\dim \mathcal{T}_k \leq N+n-k$ for $k \geq 1$. Consider again the map π from (15); we stratify $\mathcal{W}$ according to the degree of $\gcd(a^{\mathrm{h}}, b^{\mathrm{h}})$:

$$\mathcal{W} = \bigcup_{k=0}^{n} \mathcal{W}_k, \qquad \mathcal{W}_k \coloneqq \left\{ (a,b,c,d,g) \in \mathcal{W} \;\middle|\; \deg \gcd(a^{\mathrm{h}}, b^{\mathrm{h}}) = k \right\} = \pi^{-1}(\mathcal{T}_k)$$

where the last equality holds by Lemma 3.12. By said lemma, the fibers of π over $\mathcal{T}_k$ are the solution sets of the linear system (16) of constant rank $N + n - k$. Hence, $\pi \colon \mathcal{W}_k \to \mathcal{T}_k$ is an affine bundle of rank $N + k - (n-1)$, and

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

where equality holds exactly for $k = 0$. This establishes $\dim \mathcal{W} = \max_k \dim \mathcal{W}_k = 2N + 2$. Moreover, $\mathcal{W}_0$, as an affine bundle over the irreducible base $\mathcal{T}_0$, is irreducible.

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, $\mathcal{W} = \bigcup_{k=0}^{n} \overline{\mathcal{W}_k}$ is a union of subvarieties, so none of the $\mathcal{W}_k$, $k \geq 1$, are (unions of) components of $\mathcal{W}$, as they have strictly 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) along $\mathcal{W}_0$ 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 Higher-order models with multiple poles

If $\deg c = N$ but c has multiple roots, the bad locus $\mathcal{B}$ may be non-empty 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 $a = (s + p_1)(s + p_2)(s - p_3)$, for arbitrary $p_3 \in \mathbb{C} \cup \{\infty\}$. △

When the bad locus is non-empty, it is natural to ask whether the images of the critical points can still be distinguished from it. The following corollary shows that for reasonable input data—higher-order models of McMillan degree exactly N—the bad locus never overlaps with the images of the critical points.

Corollary 4.2 (Critical points avoid the bad locus)**.** *Let ϕ be the projection from (11) and consider input data $(c, d) \in CD$ with $\deg c = N$ and $\gcd(c, d) = 1$. Then $\phi(\mathcal{W}_{c,d}) \cap \mathcal{B} = \emptyset$, so that every point of $\mathcal{D}$ is either the image of a critical point or an element of the bad locus, but not both. Moreover, for every $[a] \in \phi(\mathcal{W}_{c,d})$, the null space of $\boldsymbol{M}(a, c, d)$ is spanned by a single vector $(-1, b, g)$, so that the corresponding critical point (a, b, g) is unique up to scaling.*

Proof. Let $(a, b, g) \in \mathcal{W}_{c,d}$ and $f := \gcd(c, \tilde{a}^2)$. Since $ad = bc + \tilde{a}^2 g$, the polynomial f divides ad. By the assumption, $\gcd(c, d) = 1$, so that f is coprime to d and hence f divides a. Therefore $\deg f \leq \deg a \leq n$. Since $\deg c = N$, t does not divide c^{h}, so that $\gcd(c^{\mathrm{h}}, (\tilde{a}^{\mathrm{h}})^2)$ also has degree $\deg f \leq n$. This implies, by Lemma 3.7, that $\boldsymbol{B}(a, c)$ has full column rank N, i.e., $[a] \notin \mathcal{B}$. Consequently, the null space of $\boldsymbol{M}(a, c, d) = \begin{bmatrix} \boldsymbol{T}_N(a)^{\mathsf{T}} \boldsymbol{d} & \boldsymbol{B}(a, c) \end{bmatrix}$ is at most one-dimensional, and any nonzero null vector has a nonzero first entry. It therefore contains $(-1, b, g)$ and is spanned by it. □

If $\mathcal{B}$ is non-empty but finite, part of the generic count for $\deg \mathcal{D}$ in Theorem 3.10 is accounted for by points in the bad locus; subtracting their contribution decreases the count for the critical points accordingly. In this case, the combination of Corollary 4.2 and Theorem 3.10 provides an upper bound on the number of critical points.

For $[a] \in \mathcal{D}$, let $\operatorname{mult}_a I_{N+1}(\boldsymbol{M}(a, c, d))$ denote its multiplicity with respect to the ideal generated by the maximal minors of $\boldsymbol{M}(a, c, d)$, and let $\deg \mathcal{B}$ denote the degree of the ideal $I_N(\boldsymbol{B})$ generated by the maximal minors of $\boldsymbol{B}(a, c)$.

Corollary 4.3 (Adjusted count for a non-empty bad locus)**.** *Consider input data $(c,d) \in CD$ with $\deg c = N$ and $\gcd(c,d) = 1$, and assume that $\mathcal{D}$ is finite. Then the number m of critical points satisfies*

$$m = \#\mathcal{D} - \#\mathcal{B} \leq \deg \mathcal{D} - \sum_{[a]\in\mathcal{B}} \operatorname{mult}_a I_{N+1}(\boldsymbol{M}(a,c,d)) \leq \deg \mathcal{D} - \deg \mathcal{B}, \tag{19}$$

where $\deg \mathcal{D}$ is still given by (14). Equality holds in the first inequality if and only if every point of $\mathcal{D} \setminus \mathcal{B}$ has multiplicity one.

Proof. By the decomposition $\mathcal{D} = \phi(\mathcal{W}_{c,d}) \cup \mathcal{B}$ and Corollary 4.2, the critical points are in bijection with the points of $\mathcal{D} \setminus \mathcal{B}$, so that $m = \#\mathcal{D} - \#\mathcal{B}$. Since $\mathcal{D}$ is finite, the Thom–Porteous formula applies and $\deg \mathcal{D}$ equals (14). Each point of $\mathcal{D}$ contributes its multiplicity, which is at least one, to $\deg \mathcal{D}$. This gives the first inequality, with equality if and only if every point of $\mathcal{D}\setminus\mathcal{B}$ has multiplicity one. The second inequality holds because $I_{N+1}(\boldsymbol{M}(a,c,d)) \subseteq I_N(\boldsymbol{B}(a,c))$, so that the multiplicity of every $[a] \in \mathcal{B}$ with respect to the maximal minors of $\boldsymbol{M}(a,c,d)$ is bounded below by the multiplicity of the point with respect to the maximal minors of $\boldsymbol{B}(a,c)$. □

In particular, for every given model $H \in \mathcal{M}_N$ for which $\mathcal{D}$ is finite, the count (14) remains an upper bound on the number of critical points, also when H has multiple poles. One can deduce from Lemma 3.7 that $\mathcal{B}$ is finite if c has at most one multiple root, which may be of multiplicity at most 3. We analyze the example of one double root below.

Proposition 4.4. *Let $(c,d) \in CD$, where c has exactly one double root and $N-2$ simple roots. Then $\#\mathcal{B} = \binom{N-2}{n-1}$. If, moreover, $\gcd(c,d) = 1$ and $\mathcal{D}$ is finite, then the number m of critical points satisfies*

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

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 3.7 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 $\mathcal{B}$ consists of exactly $\binom{N-2}{n-1}$ points, disregarding potential multiplicity. The result is then an application of Corollary 4.3. □

4.2 Discrete-time models

All results presented in this paper have a direct counterpart in the discrete-time setting. They are obtained via identical derivations but adapted notation, or by direct transformation of our results via the isomorphism on spaces $\mathcal{H}_2(\mathbb{C}^+)$ and $\mathcal{H}_2(\mathbb{D})$, the Hardy spaces on the half-plane and the unit disk, respectively [Par04, Theorem 1.2.5]. A treatment of this discrete-time case can be found in [Kay26, Chapter 5].

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 $1 \leq n < N$, the $\mathcal{H}_2$-norm model reduction problem searches for a solution

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

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

In this setting, the interpolatory conditions in Theorems 2.1 and 2.2 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 $-\hat{p}_i$ are replaced by $1/\hat{p}_i$. Accordingly, our notation $\tilde{a} = \sum_{i=0}^{n} a_i(-s)^i$ used in (5) must be adapted to the discrete-time setting and becomes $\tilde{a}(z) = z^n a(1/z) = \sum_{i=0}^{n} a_i z^{n-i}$.

Theorem 4.5 (Discrete-time count)**.** *With adapted notation, all results in Sections 3 and 4.1, including Corollaries 3.5, 3.9, 4.2 and 4.3, lemma 3.7, and theorem 3.10, hold verbatim for the discrete-time model order reduction problem.*

5 Examples: the good, the bad and the degenerate

We present some (numerical) examples to illustrate our results. The code is available in a public repository [KL26]. Throughout, numerical values are rounded for displaying purposes.

We start with Examples 5.1 and 5.2, illustrating the behavior for generic input data; the remaining examples consider special behavior. In Examples 5.3 and 5.4 the input data is such that the bad locus is non-empty, whose presence turns the exact count into an upper bound (Section 4.1). Example 5.5 considers special input data for which a critical point appears with multiplicity, contrary to the generic case characterized in Corollary 3.5. Examples 5.6 and 5.8 investigate the subtle differences between the different optimality conditions from Theorems 2.1 and 2.2 for approximants with multiple poles. Finally, Example 5.6 and Remark 5.7 briefly touch upon the phenomenon of *degenerate* stationary points, which are reduced models at which the gradient of J^2 with respect to the coefficients of (a, b) vanishes, but $\gcd(a, b) \neq 1$ so that $\widehat{H} \notin \mathcal{M}_n$.

Example 5.1 (Continuous time). We revisit the example of [AVD21, Section 5] and [LAD23, Example 1], which seeks the optimal first-order approximation ($n = 1$) of the third-order model ($N = 3$)

$$H(s) = \frac{d}{c} = \frac{s^2 + 9s - 10}{s^3 + 12s^2 + 49s + 78} = \frac{(s+10)(s-1)}{(s+6)(s^2+6s+13)} \in \mathcal{M}_3,$$

whose poles are $s \in \{-6, -3 \pm 2\mathrm{j}\}$. For this setup, the matrix $\boldsymbol{M}(a, c, d)$ corresponds to

$$\boldsymbol{M}(a,c,d) = \begin{bmatrix} -10a_0 & 78 & a_0^2 & 0 \\ 9a_0 - 10a_1 & 49 & -2a_0a_1 & a_0^2 \\ a_0 + 9a_1 & 12 & a_1^2 & -2a_0a_1 \\ a_1 & 1 & 0 & a_1^2 \end{bmatrix}.$$

The determinantal variety $\mathcal{D}$, defined by the determinant of $\boldsymbol{M}(a, c, d)$, consists of five reduced points $[a] = (a_0 : a_1) \in \mathbb{P}^1$:

$$\mathcal{D} = \{\, (9.6796 : 1), (0.2671 : 1), (-16.6189 : 1), (-4.1639 + 0.9027\mathrm{j} : 1), (-4.1639 - 0.9027\mathrm{j} : 1) \,\}.$$

Since c has N distinct roots, the bad locus $\mathcal{B}$ is empty (cf. Lemma 3.7). The number of obtained critical points is in agreement with the count from Theorem 3.10. The two real points in $\mathcal{D}$ with $a_0/a_1 > 0$ correspond to models

$$\widehat{H}_1 = \frac{1.279}{s + 9.679}, \qquad \widehat{H}_2 = \frac{-0.043}{s + 0.267}$$

in $\mathcal{M}_1$, both of which are local minimizers of the $\mathcal{H}_2$-error; $\widehat{H}_1$ being the global minimizer. △

Example 5.2 (Discrete-time). We consider the example given in [LAD23, Example 2], in which the discrete-time model

$$H(z) = \frac{0.0448z^5 + 0.2368z^4 + 0.0013z^3 + 0.0211z^2 + 0.2250z + 0.0219}{z^6 - 1.2024z^5 + 2.3675z^4 - 2.0039z^3 + 2.2337z^2 - 1.0420z + 0.8513}, \tag{20}$$

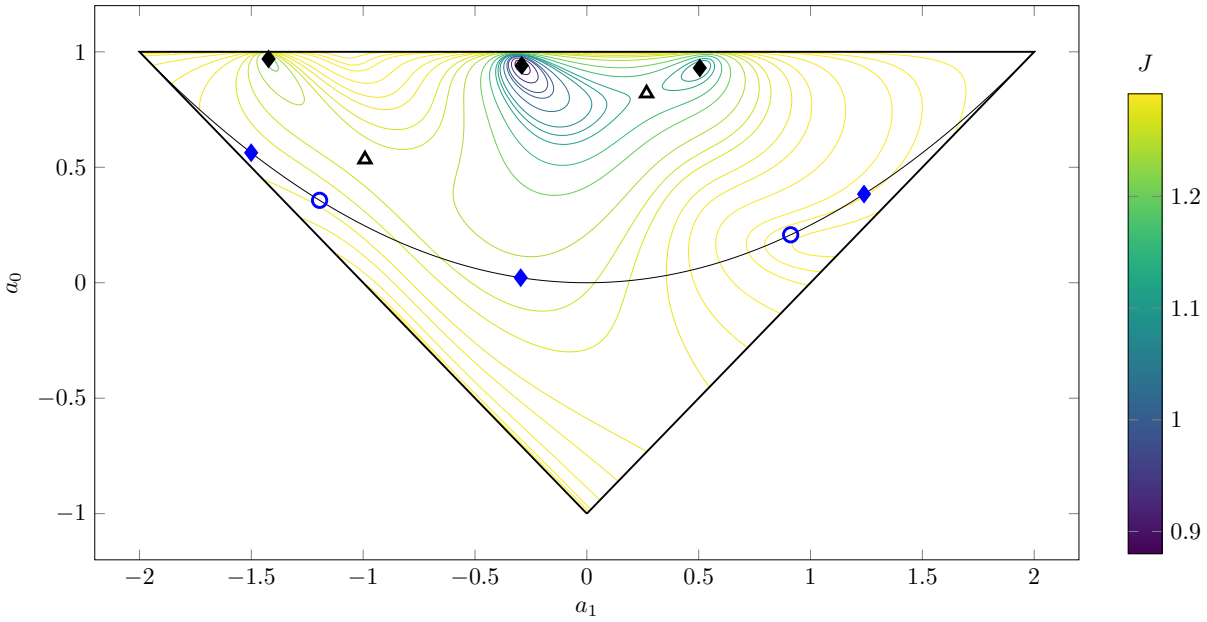


Figure 1: Contour plot of the objective function value $\|H - \widehat{H}\|_{\mathcal{H}_2}$ for varying reduced-order models $\widehat{H}$. The problem setup is described in Examples 5.2 and 5.6. Each point $(a_0 : a_1 : 1) \in \mathbb{P}^2$ uses the optimal numerator b, obtained from solving (21). The local minimizers (⧫) and saddle points (△) in the displayed region are highlighted. The black line depicts those points for which the discriminant $\Delta(a)$ vanishes. When the model reduction problem is restricted to this discriminant, a different set of local minimizers (⧫) and maximizers (◯) is obtained.

is approximated by a second-order model $\widehat{H}(z) = \frac{b_1 z + b_0}{a_2 z^2 + a_1 z + a_0}$. The authors report that the critical equations admit 49 isolated solutions, matching the prediction from Theorem 3.10. For this problem instance, 11 critical points are real-valued, 5 of which correspond to stable models.

For a fixed stable a there is a unique b that minimizes the error J^2. The derivation of the critical equations in Section 2, when carried out for b only, shows that this optimal b satisfies $ad - bc = \tilde{a}q$ for some $q \in \mathbb{C}[z]_{\leq N-1}$, with $\tilde{a}(z) = z^n a(1/z)$ as in Section 4.2. The optimal b can thus be found from the unique solution of the linear system

$$\begin{bmatrix} \boldsymbol{T}_n(c)^\mathsf{T} & \boldsymbol{T}_N(\tilde{a})^\mathsf{T} \end{bmatrix} \begin{bmatrix} \boldsymbol{b} \\ \boldsymbol{q} \end{bmatrix} = \boldsymbol{T}_N(a)^\mathsf{T} \boldsymbol{d}. \tag{21}$$

Its coefficient matrix, the Sylvester matrix of c and $\tilde{a}$, is invertible for all points of interest because c and $\tilde{a}$ have no common roots when H and $\widehat{H}$ are stable.

This observation allows us to draw a contour plot of the objective function value J for various reduced models $\widehat{H}$, shown in Figure 1. We work in the affine chart $a_2 = 1$ and the plot over $(a_1, a_0) \in \mathbb{R}^2$ is restricted to the triangular region corresponding to stable $\widehat{H}$; in the discrete-time setting, stability means that all roots of a lie strictly inside the complex unit disk. △

We now turn our attention to approximation problems where the bad locus $\mathcal{B}$ is not empty and hence the solution count has to be adapted. Such examples are mainly of theoretical interest and illustrate the assumptions and bounds obtained in the theorems of Sections 3 and 4.

Example 5.3 (Non-empty bad locus). Consider the third-order model ($N = 3$)

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

which has two distinct poles, one of which has multiplicity two. We compute its optimal first-order approximant ($n = 1$). The determinantal variety $\mathcal{D}$ contains the following approximate

values of $[a] = (a_0 : a_1) \in \mathbb{P}^1$:

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

The generic count from Table 1 lists 5 solutions for $(N, n) = (3, 1)$. In our example, however, c fails the genericity condition because it has a double root. Since $\mathcal{D}$ is finite, we may apply Proposition 4.4 to adjust the generic count: we must subtract at least $\binom{N-2}{n-1} = \binom{1}{0} = 1$ to account for the non-empty bad locus $\mathcal{B}$. Indeed, symbolic computation shows that $\mathcal{B}$ consists of the single point $(-2 : 1)$. The adjusted count yields 4 critical points.

The bad locus point can also be read off from Lemma 3.7. By the lemma, $[a] \in \mathcal{B}$ if and only if $\deg \gcd(c^{\mathrm{h}}, (\tilde{a}^{\mathrm{h}})^2) > 1$. Since $\deg c = 3$, c has no roots at infinity, so $[a] \in \mathcal{B}$ requires $(s+2)^2 \mid \tilde{a}^2$, i.e., $\tilde{a}$ is proportional to $s+2$ and hence $a = a_1(s-2)$ for some $a_1 \in \mathbb{C}$. △

Example 5.4 (Multiplicity in the bad locus points). Reconsider the third-order ($N = 3$) model H from Example 5.3, but now compute its optimal second-order approximant ($n = 2$). In this case, 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 $\boldsymbol{M}(a, c, d)$. Hence, $\#\mathcal{D} = 6$ but $\deg \mathcal{D} = 7$, the latter of which coincides with the entry in Table 1 for $(N, n) = (3, 2)$.

By Lemma 3.7, $[a] \in \mathcal{B}$ if and only if $\deg \gcd(c^{\mathrm{h}}, (\tilde{a}^{\mathrm{h}})^2) \geq 3$. Because $\deg c^{\mathrm{h}} = 3$, this occurs only when c^{h} divides $(\tilde{a}^{\mathrm{h}})^2$, which requires $\tilde{a}$ to be proportional to $(s+2)(s+3)$, i.e., $a = a_2(s-2)(s-3) = a_2(s^2 - 5s + 6)$. Hence $\mathcal{B}$ consists of the single point $[a] = (6 : -5 : 1)$. It has multiplicity two with respect to the maximal minors of $\boldsymbol{B}(a, c)$ as well, so that $\#\mathcal{B} = 1$ and $\deg \mathcal{B} = 2$. We conclude that there are $\#\mathcal{D} - \#\mathcal{B} = 5$ critical points.

Proposition 4.4 adjusts the generic count of 7 by subtracting $\binom{N-2}{n-1} = \binom{1}{1} = 1$, yielding an upper bound of 6 critical points. The sharper bound from Corollary 4.3, which takes the multiplicity of the bad locus points into account, gives $\deg \mathcal{D} - \deg \mathcal{B} = 7 - 2 = 5$. △

Example 5.5 (Critical point with multiplicity). Consider the higher-order model ($N = 2$) defined by

$$H(s) = \frac{11s + 13}{(s+7)(s+\frac{1}{5})}$$

and consider reduction to order $n = 1$. Because c has distinct roots and $\gcd(c, d) = 1$, the bad locus is empty and the determinantal variety $\mathcal{D}$ is finite (Lemma 3.7 and Corollary 3.9, respectively). The matrix $\boldsymbol{M}(a, c, d) \in \mathbb{C}[\boldsymbol{a}]^{3\times 3}$ is square and its determinant is of degree 3. It has a double root at $a_0/a_1 = 1$ and a single root at $a_0/a_1 = \frac{91}{55}$. Hence, there are $m = \#\mathcal{D} = 2$ critical points, both real and stable. Because of the root with multiplicity, the inequality in (14) is strict: $2 = \#\mathcal{D} < \deg \mathcal{D} = 3$.

These input data (c, d) are thus not generic in the sense of the fiber dimension theorem (Theorem A.1): they lie outside the dense open subset on which the fibers of τ, from (7), are reduced. Hence the conclusion from Corollary 3.5 does not hold and the fiber is *not* reduced. △

Note that (c, d) were deliberately chosen to have two critical points coincide. Such input data lie on a hypersurface in CD, which is a set of measure zero, and an arbitrarily small perturbation of (c, d) splits the double critical point into two simple ones.

Example 5.6 (Prescribed multiplicity structure). Reconsider the setup from Example 5.2. The black curve in Figure 1 is the zero-set of the discriminant $\Delta(a) = a_1^2 - 4a_0a_2$, depicting all reduced models with a multiple pole. It is the submanifold described in Remark 2.3.

The local minimizers of J^2, when restricted to this curve, satisfy (the discrete-time counterpart of) the conditions of Theorem 2.1 (Meier–Luenberger), rather than those of Theorem 2.2 (Walsh). Indeed, when one parametrizes $\hat{H}$ with a pole of multiplicity $m = 2$, there remain only $m + 1 = 3$ interpolation conditions at $1/p$, instead of the $2m = 4$ conditions from Theorem 2.2. Hence, when parametrizing $a := a_2(z-p)^2$, the factor $\tilde{a}^2 = (1-pz)^4$ in (5) is to be replaced by $\tilde{a}' := (1-pz)^3$:

$$ad - bc = (1-pz)^3 h, \qquad h \in \mathbb{R}[z]_{\leq N-2}. \tag{22}$$

Upon normalization (e.g., $a_2 = 1$), these are 8 polynomial equations in the 8 unknowns $p, \boldsymbol{b}$ and $\boldsymbol{h}$. The system (22) is equivalent to

$$\boldsymbol{M}'(p) \begin{pmatrix} -1 \\ \boldsymbol{b} \\ \boldsymbol{h} \end{pmatrix} = \boldsymbol{0}, \qquad \boldsymbol{M}'(p) \coloneqq \begin{bmatrix} \boldsymbol{T}_N(a)^\mathsf{T} \boldsymbol{d} & \boldsymbol{T}_2(c)^\mathsf{T} & \boldsymbol{T}_{N-1}(\tilde{a}')^\mathsf{T} \end{bmatrix} \in \mathbb{R}[p]^{8\times 8}.$$

This matrix is square and depends on the single parameter p; it is a polynomial eigenvalue problem (cf. Example 3.3).

For the higher-order model in (20), $\det \boldsymbol{M}'(p)$ has $3N-1 = 17$ distinct roots, 7 of which are real, and 3 of which lie in the unit disk: $p \in \{-0.6196, 0.1478, 0.7503\}$.[5] These constitute the three local minimizers of the restricted problem, depicted in blue in Figure 1.

The restricted problem has two local maxima, $p \in \{-0.4556, 0.5974\}$, which do not appear as a solution to (22). At both points, the optimal numerator b is such that a and b share a common factor—the stationary point is *degenerate*—so that $\widehat{H} = b_1/(z-p)$ has McMillan degree one instead of two. These local maximizers are stationary points of the (unrestricted) first-order problem ($n = 1$). △

Remark 5.7 (Degenerate stationary points). Also in the unrestricted case, stationary points of a lower-order problem can reappear in the higher-order problem as degenerate stationary points, see e.g., [BOW92, Proposition 2.7] for a characterization in discrete time. Because such points lie outside of $\mathcal{M}_n$, Theorem 2.2 does not apply to them, and they need not appear as a solution of the critical equations (6). For model reduction purposes, this poses no problem as [Reg95, Theorem 5.3] guarantees that such degenerate points can never be local minimizers.

Example 5.8 (Optimal reduced model with a triple pole). We revisit the example of [VGA10, Section 6], following their Theorem 6.2, where

$$H(s) = \frac{\frac{1}{256}(s^3 - 9s^2 + 295s + 2449)}{s^4 + 13s^3 + \frac{513}{16}s^2 + \frac{239}{8}s + \frac{153}{16}}$$

has four distinct poles. For $n = 3$, the bad locus $\mathcal{B}$ is therefore empty, and the determinantal variety $\mathcal{D}$ consists of $15 = 2^4 - 1$ points of multiplicity one, in agreement with Theorem 3.10. Three of these points are real, and only one of them corresponds to a stable model: $a = (s+1)^3$ with $b = 1$. Since a global minimizer exists (cf. Remark 3.6), $\widehat{H} = 1/(s+1)^3$ is the global minimizer.

For this particular higher-order model H, the global minimizer thus lies on the curve of polynomials $a = (s-p)^3$ with a triple root, which is contained in the discriminant locus $\Delta(a) = 0$. Being a stationary point of the unrestricted problem, $\widehat{H}$ satisfies Theorem 2.2: the interpolation conditions (3) hold at $s = 1$ for $k = 0, \ldots, 5$. These include the conditions for $k = 0, \ldots, 3$ required by Theorem 2.1, so that the latter are automatically satisfied as well.

Hence, this stationary point would also have been obtained by enforcing the multiplicity structure from the start, as in Example 5.6, i.e., by replacing $\tilde{a}^2 = (s+p)^6$ in (5) with $(s+p)^4$. One may verify that in this case, $\det \boldsymbol{M}'(p)$ has 15 distinct roots, 5 of which are real and 3 of which are stable: $p \in \{-1, -0.084, -0.476\}$. Contrary to the point $p = -1$, which corresponds to the global minimizer of the unrestricted problem, the other two are stationary points of the restricted problem only. △

6 Conclusions and outlook

We have 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

[5] Similar to $\boldsymbol{M}(a, c, d)$, the matrix $\boldsymbol{M}'(p)$ may have a bad locus, consisting of the values of p for which its last $N+1$ columns, $[\boldsymbol{T}_2(c)^\mathsf{T}\ \boldsymbol{T}_{N-1}(\tilde{a}')^\mathsf{T}]$, drop rank. Its behavior can be studied similarly as in Lemma 3.7.

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—corresponding to extraneous solutions introduced by the MEP formulation—was favorable, as we were able to show that the bad locus is most often 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 [KL26+] we employed a similar strategy in the context of a system identification problem.

In Examples 5.3 and 5.4, we illustrated that for specific problem instances, in particular when the higher-order model has multiple poles, the bad locus $\mathcal{B}$ can be non-empty and contain points of higher multiplicity. Moreover, Example 5.6 touches on how the critical equations change when the poles of the reduced model follow a prescribed multiplicity structure. A detailed understanding of these special cases is an interesting direction for future work.

In the case of a higher-order model of full McMillan degree N and simple poles, we have reproven the main result of [HPB12] on finiteness of critical points (cf. Corollary 3.9). It remains an open question whether the critical points can ever form an infinite set, for *any* higher-order model (of degree $N \geq n$). We conjecture that the critical points *always* form a finite set, and have verified this in the case of a finite bad locus.

Example 5.6 also illustrated that stationary points of a lower-order problem can reappear in a (possibly restricted) higher-order problem as *degenerate* stationary points, that is, as models with pole-zero cancellations and thus McMillan degree strictly smaller than n. For the (unrestricted) problem in discrete time, such points were characterized in [BOW92, Proposition 2.7]. By formulating the model order reduction problem (2) over $\mathcal{M}_n$, we have excluded these degenerate points from the search space. However, even though they are never local minimizers [Reg95, Theorem 5.3], they can appear as saddle points or local maxima in the optimization landscape and may therefore still be of practical interest. A systematic understanding of how many such degenerate stationary points arise, and how they influence the optimization landscape, is another interesting research direction. In our related work [KL26+], lower-order solutions similarly appear as separate components of the critical equations.

Another natural extension is to investigate how the presented 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. We expect the weights to alter the structure of the critical equations, and thereby to complicate the analysis of the Walsh variety $\mathcal{W}$ and corresponding bad locus.

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: 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 Sibren Lagauw was supported by the FWO Fundamental Research Fellowship under Grant 11K5625N.

Declaration of AI use

GenAI-based tools were used solely in the proofreading process.

A Relevant background in Algebraic Geometry

This appendix 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. For readers interested in applications, [SW05] provides an excellent introduction to polynomial system solving in the sciences and engineering. The book [CR25] also provides a beginner-friendly introduction to algebraic geometry. Basic results in algebraic geometry are introduced in e.g., [Har92; Sha13; CLO15], while more advanced treatments can be found in [Har77; EH16].

A.1 Affine varieties

Let $R_n \coloneqq \mathbb{C}[x_1, \ldots, x_n]$ 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 $F = \{f_1, \ldots, f_s\} \subseteq R_n$ 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 $F \subseteq \mathbb{C}[x_1, \ldots, x_n]$; their complements are called *open sets*.

A *subvariety* or *closed subset* Y of X is an affine variety $Y \subseteq X$; an *open subset* is of the form $X \setminus Y$. 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 property holds for *generic* points $\boldsymbol{x} \in X$ if it holds for all points in a 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*; almost all of the definitions in this section apply also to such varieties.

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):

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

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,

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

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 an affine variety $X \subseteq \mathbb{C}^n$ 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 $F = \{f_1, \dots, f_s\} \subseteq I$ 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

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

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, if X is irreducible, is a connected complex manifold of the same (complex) 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 the defining polynomials are real and $\boldsymbol{p} \in X^{\mathrm{sm}}(\mathbb{R})$, then $X^{\mathrm{sm}}(\mathbb{R}) \subseteq \mathbb{R}^n$ is a smooth (real) submanifold in a neighborhood of $\boldsymbol{p}$.

For polynomial equations F we defined what it means to defined $X = V(F)$ in a reduced way. We now turn from this global viewpoint to a local one. Globally, we consider polynomial functions defined on all of $\mathbb{C}^n$, and reducedness is expressed by the defining ideal being radical. Locally, however, we focus on the behavior near a single point. Functions in this local setting are defined only on a small neighborhood 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 such local functions.

Let $X \subseteq \mathbb{C}^n$ be an affine variety and $\boldsymbol{p} \in X$. A regular function germ $\varphi_{\boldsymbol{p}}$ at $\boldsymbol{p}$ is a function $\varphi\colon U \to \mathbb{C}$ defined on some open set $\boldsymbol{p} \in U \subseteq X$ which, on U, can be represented as $\varphi = \frac{f}{g}$, $f, g \in R_n$ and $V(g) \cap U = \emptyset$. Two such function germs are considered equal if they agree on a common open neighborhood of $\boldsymbol{p}$. That is, germs are functions on an arbitrarily small neighborhood of $\boldsymbol{p}$, so it does not make sense to evaluate $\varphi_{\boldsymbol{p}}$ at any point of X other than $\boldsymbol{p}$, though local properties like order of vanishing, gradient and such are still well-defined. The

collection of regular function germs form the *local ring of X at $\boldsymbol{p}$* denoted by $\mathcal{O}_{X,\boldsymbol{p}}$. For example, regular function germs on $X = \mathbb{C}^n$ at $\mathbf{0}$ are rational functions without a pole in $\mathbf{0}$.

Every polynomial $f \in R_n$ defines function germs $f_{\boldsymbol{p}} \in \mathcal{O}_{\mathbb{C}^n,\boldsymbol{p}}$. The *germ* of an ideal $\langle F\rangle_{R_n}$ is the ideal $F\mathcal{O}_{\mathbb{C}^n,\boldsymbol{p}}$ in the ring $\mathcal{O}_{\mathbb{C}^n,\boldsymbol{p}}$ generated by $\{\, f_{\boldsymbol{p}} \mid f \in F \,\}$. Now F defines the variety $X = V(F) \subseteq \mathbb{C}^n$ *in a reduced way at* $\boldsymbol{p}$ if $F\mathcal{O}_{\mathbb{C}^n,\boldsymbol{p}} = I(X)\mathcal{O}_{\mathbb{C}^n,\boldsymbol{p}}$, equivalently if $F\mathcal{O}_{\mathbb{C}^n,\boldsymbol{p}}$ is a radical ideal. For example, if the Jacobian matrix of F at $\boldsymbol{p}$ is of maximal rank, then F defines $X = V(F)$ in a reduced way at $\boldsymbol{p}$ and furthermore $\boldsymbol{p}$ is a smooth point of X.

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

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

Polynomial coordinate functions are what sets a morphism apart from general maps; it ensures that the preimages of subvarieties are again subvarieties, and that the composition of ϕ with a regular function germ $\varphi_{\boldsymbol{q}}$, $\phi(\boldsymbol{p}) = \boldsymbol{q}$, yields another regular function germ $(\varphi\circ\phi)_{\boldsymbol{p}}$. 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 *isomorphic*.

The following two fundamental results of algebraic geometry can be seen as extensions of the rank-nullity theorem from linear algebra to non-linear equations.

Theorem A.1 (Fiber dimension theorem [Sha13, Theorem I.1.25], [Har77, Corollary III.10.7])**.** *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 (even smooth!) fibers at $\boldsymbol{q} \in U$ and equality holds in the first statement.*

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

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

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

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

With minor additional technicalities (defining morphisms of quasi-affine varieties), 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 exists a set $\{f_1, \dots, f_s\} \subseteq R_n$ such that $X = V(f_1, \dots, f_s) \cap U$. In this case, X can be defined by s equations, and no strictly smaller number of equations suffices to cut out X set-theoretically (by Theorem A.2). If, in addition, $f_1, \dots, f_s$ generate the defining ideal $I(X)$ of X in $\mathcal{O}_{\mathbb{C}^n,\boldsymbol{p}}$ (i.e., they define X by a radical ideal in $\mathcal{O}_{\mathbb{C}^n,\boldsymbol{p}}$) for all $\boldsymbol{p} \in U$, then X is an *ideal-theoretic complete intersection in U*.

Theorem A.3 (Unmixedness theorem [Eis95, Corollary 18.14]). *Let $X = V(f_1, \dots, f_s)$ 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

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

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 $(\lambda x_0, \lambda x_1, \dots, \lambda x_n) \in \mathbb{C}^{n+1}$ 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 nonzero vectors modulo scaling, in other words 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 $F = \{f_1, \dots, f_s\} \subseteq R_{n+1} = \mathbb{C}[x_0, \dots, x_n]$. Notions such as open/closed subsets are defined analogously. 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 $\dim \widehat{X} = \dim X + 1$ 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 affine part on $\mathbb{P}^n$. A *standard affine chart* of a projective variety X is an open subset of the form $X \setminus V(x_i)$. Affine charts can be naturally identified with affine varieties. For example, in the case $X = \mathbb{P}^n$, 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 $Y = V(f_1, \ldots, f_s)$, then it is usually *not* the case that $\overline{Y} = V(f_1^{\mathrm{h}}, \ldots, f_s^{\mathrm{h}})$, even when the f_i define Y in a reduced way; the latter projective variety may contain extra components in H_∞.

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.

Let $X \subseteq \mathbb{P}^n$ be a quasi-projective variety. A function $\varphi \colon X \to \mathbb{C}$ is *regular* if on every standard affine chart $X \setminus V(x_i)$ it defines a regular function germ at every point $\boldsymbol{p} \in X_i$ in the sense of Section A.2. Locally, though not necessarily globally, it can be written as a fraction of two homogeneous polynomials $\varphi = f(x)/g(x)$ of the same degree (which hence is well-defined for projective points).

Affine varieties carry many regular functions, for instance all polynomials. Projective varieties, by contrast, carry very few. The following result is an algebraic counterpart of the fact that a holomorphic function on a compact, connected complex manifold is constant (Liouville's theorem/the maximum modulus principle).

Theorem A.4 ([Sha13, Corollary I.5.1.1])**.** *Every regular function on an irreducible projective variety is constant.*

An important advantage of working in projective space is that some 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.5 ([Sha13, Theorem I.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 $d_1, \ldots, d_e$:

$$\boldsymbol{M}(\boldsymbol{x}) = (h_{i,j})_{1 \leq i \leq f,\, 1 \leq j \leq e}, \qquad h_{i,j} \in R_{n+1} \text{ is homogeneous of degree } d_j. \tag{23}$$

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.6 (Dimension of determinantal varieties [Ful98, Theorem 14.4(b)])**.** *Let $\boldsymbol{M}$ be the matrix polynomial from (23). For every irreducible component $Z \subseteq D_k(\boldsymbol{M})$ one has*

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

One can show, if the expected dimension $n-(e-k)(f-k)$ is non-negative and under suitable positivity assumptions on the degrees d_j, that $D_k(\boldsymbol{M})$ is always non-empty, in analogy to Theorem A.5. 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.6 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 (23)) the rank conditions (potentially) impose fewer independent constraints, and the codimension can only become smaller. Theorem A.6 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 (potentially empty), 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 $\deg X \coloneqq \#X = \deg I(X)$. In this case, $\deg\langle F\rangle$ corresponds to the number of points in X, counted with multiplicity, and we have

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

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

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

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 $X = V(f)$ with f square-free (e.g., irreducible), 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 in R_{n+1} 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. This construction also gives a general definition of the degree of a homogeneous ideal $I \subseteq R_{n+1}$. The degree of a non-empty projective variety is strictly positive, while the empty set has degree 0 (its Hilbert function is the constant 0).

A version of Bézout's theorem is that if $X = V(f_1, \dots, f_n) \subseteq \mathbb{P}^n$ is finite, then

$$\#X \le \deg\langle f_1, \dots, f_n\rangle = \deg(f_1) \cdots \deg(f_n).$$

This has a vast generalization, of which we consider a special case. Recall from Theorem A.6 that determinantal varieties have a general lower bound on their dimension. The following result tells us the degree of the resulting variety (or, to be precise, of the ideal generated by the minors) in case this bound is matched exactly. Here and everywhere else in the paper, by $\deg D_{k-1}(\boldsymbol{M})$ we mean the degree of the *ideal* generated by k-minors; if the minors define the determinantal variety in a reduced way, then this agrees with the degree of the variety.

Theorem A.7 (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 each component $Z \subseteq D_{e-1}(\boldsymbol{M})$ of the determinantal variety of rank $e-1$ has the expected dimension $\dim Z = n - (f-e+1) \ge 0$. Then*

$$\deg D_{e-1}(\boldsymbol{M}) = T^{f-e+1}\text{-coefficient of } \frac{1}{(1-d_1T)\cdots(1-d_eT)} \ge 0.$$

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 submaximal 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.8 ([Kay26, Proposition 0.4.15])**.** *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 $I(X)$ and 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.

In this lemma, "defined in a reduced way" is understood to be in the affine sense on each standard affine chart $\mathbb{C}^{n+e-1} \cong \mathbb{C}^n \times (\mathbb{P}^{e-1} \setminus V(x_i)) \subseteq \mathbb{C}^n \times \mathbb{P}^{e-1}$.

References

[Als19] Fawwaz Fayiz Alsubaie. "$\mathcal{H}_2$ Optimal Model Reduction for Linear Dynamic Systems and the Solution of Multiparameter Matrix Pencil Problems". PhD thesis. London, UK: Imperial College London, 2019. DOI: 10.25560/94222 (cit. on p. 2).

[Atk72] Frederick V. Atkinson. *Multiparameter Eigenvalue Problems. Volume I: Matrices and Compact Operators*. Vol. 82. Mathematics in Science and Engineering. New York: Academic Press, 1972 (cit. on p. 1).

[AVD21] Oscar Mauricio Agudelo, Christof Vermeersch, and Bart De Moor. "Globally Optimal $\mathcal{H}_2$-Norm Model Reduction: A Numerical Linear Algebra Approach". In: *IFAC-PapersOnLine* 54.9 (2021), pp. 564–571. DOI: 10.1016/j.ifacol.2021.06.117 (cit. on pp. 1, 2, 17).

[AW49] Pierre R. Aigrain and Everard M. Williams. "Synthesis of n-Reactance Networks for Desired Transient Response". In: *Journal of Applied Physics* 20.6 (1949), pp. 597–600. DOI: 10.1063/1.1698434 (cit. on p. 4).

[Bar86] Laurent Baratchart. "Existence and Generic Properties of L^2 Approximants for Linear Systems". In: *IMA Journal of Mathematical Control and Information* 3.2–3 (1986), pp. 89–101. DOI: 10.1093/imamci/3.2-3.89 (cit. on pp. 2, 9).

[BDD12] Kim Batselier, Philippe Dreesen, and Bart De Moor. "Prediction Error Method Identification is an Eigenvalue Problem". In: *IFAC Proceedings Volumes* 45.16 (2012), pp. 221–226. DOI: 10.3182/20120711-3-be-2027.00132 (cit. on p. 1).

[Ble10] Ivo W. M. Bleylevens. "Algebraic Polynomial System Solving and Applications". PhD thesis. Maastricht, The Netherlands: Maastricht University, 2010. DOI: 10.26481/dis.20101209ib (cit. on p. 5).

[BOW92] Laurent Baratchart, Martine Olivi, and Franck Wielonsky. "On a rational approximation problem in the real Hardy space H_2". In: *Theoretical Computer Science* 94.2 (1992), pp. 175–197. DOI: 10.1016/0304-3975(92)90033-c (cit. on pp. 20, 21).

[BT18] Paul Breiding and Sascha Timme. "HomotopyContinuation.jl: A Package for Homotopy Continuation in Julia". In: *Mathematical Software – ICMS 2018*. Ed. by James H. Davenport, Manuel Kauers, George Labahn, and Josef Urban. Vol. 10931. Lecture Notes in Computer Science. Cham: Springer, 2018, pp. 458–465. DOI: 10.1007/978-3-319-96418-8_54 (cit. on p. 2).

[CLO15] David A. Cox, John Little, and Donal O'Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. 4th ed. Undergraduate Texts in Mathematics. Cham: Springer, 2015. ISBN: 9783319167213. DOI: 10.1007/978-3-319-16721-3 (cit. on p. 22).

[CR25] Emily Clader and Dustin Ross. *Beginning in Algebraic Geometry*. Undergraduate Texts in Mathematics. Cham: Springer, 2025. DOI: 10.1007/978-3-031-88819-9 (cit. on pp. 7, 22).

[DD21] Katrien De Cock and Bart De Moor. "Multiparameter Eigenvalue Problems and Shift-invariance". In: *IFAC-PapersOnLine* 54.9 (2021), pp. 159–165. DOI: 10.1016/j.ifacol.2021.06.071 (cit. on p. 1).

[DeM20] Bart 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. DOI: 10.4310/cis.2020.v20.n2.a4 (cit. on p. 1).

[Dra+16] Jan Draisma, Emil Horobeţ, Giorgio Ottaviani, Bernd Sturmfels, and Rekha R. Thomas. "The Euclidean Distance Degree of an Algebraic Variety". In: *Foundations of Computational Mathematics* 16.1 (2016), pp. 99–149. DOI: 10.1007/s10208-014-9240-x (cit. on p. 3).

[DVS93] Bart De Moor, Peter Van Overschee, and Geert Schelfhout. "H_2 Model Reduction for SISO Systems". In: *IFAC Proceedings Volumes* 26.2 (1993), pp. 895–898. DOI: 10.1016/s1474-6670(17)49262-7 (cit. on p. 2).

[EH16] David Eisenbud and Joe Harris. *3264 and All That: A Second Course in Algebraic Geometry*. Cambridge: Cambridge University Press, 2016. ISBN: 9781107602724. DOI: 10.1017/cbo9781139062046 (cit. on pp. 7, 22, 27).

[Eis95] David Eisenbud. *Commutative Algebra: with a View Toward Algebraic Geometry*. Vol. 150. Graduate Texts in Mathematics. New York: Springer, 1995. ISBN: 9781461253501. DOI: 10.1007/978-1-4612-5350-1 (cit. on p. 25).

[Ful98] William Fulton. *Intersection Theory*. 2nd ed. Vol. 2. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. Berlin: Springer, 1998. ISBN: 9781461217008. DOI: 10.1007/978-1-4612-1700-8 (cit. on p. 26).

[GAB08] Serkan Gugercin, Athanasios C. Antoulas, and Christopher 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. DOI: 10.1137/060666123 (cit. on pp. 2, 4).

[Gia07] Giovanni Z. Giambelli. "Sul luogo dei punti di contatto delle ipersuperficie di due dati sistemi lineari". In: *Rendiconti del Circolo Matematico di Palermo* 23.1 (1907), pp. 302–306. DOI: 10.1007/bf03013525 (cit. on p. 7).

[Har77] Robin Hartshorne. *Algebraic Geometry*. Vol. 52. Graduate Texts in Mathematics. New York: Springer, 1977. ISBN: 9781475738490. DOI: 10.1007/978-1-4757-3849-0 (cit. on pp. 22, 24).

[Har92] Joe Harris. *Algebraic Geometry: A First Course*. Vol. 133. Graduate Texts in Mathematics. New York: Springer, 1992. ISBN: 9781475721898. DOI: 10.1007/978-1-4757-2189-8 (cit. on p. 22).

[HKP24] Michiel E. Hochstenbach, Tomaž Košir, and Bor 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), e2540. DOI: 10.1002/nla.2540 (cit. on pp. 1, 2, 9).

[HMC07] Bernard Hanzon, Jan M. Maciejowski, and Chun Tung Chou. *Optimal $\mathcal{H}_2$ order-one reduction by solving eigenproblems for polynomial equations*. 2007. arXiv: 0706.1862 [math.OC] (cit. on pp. 2, 5, 12).

[HPB12] Bernard Hanzon, Ralf Peeters, and Ivo Bleylevens. "On a Finiteness Result for the Number of Critical Points of the $\mathcal{H}_2$ Approximation Criterion". In: *IFAC Proceedings Volumes* 45.16 (2012), pp. 728–733. DOI: 10.3182/20120711-3-be-2027.00156 (cit. on pp. 2, 11, 21).

[HYT01] Xue-Xiang Huang, Wei-Yong Yan, and Kok Lay Teo. "A new approach to frequency weighted L_2 optimal model reduction". In: *International Journal of Control* 74.12 (2001), pp. 1239–1246. DOI: 10.1080/00207170110063499 (cit. on p. 21).

[Kay26] Leonie Kayser. "Computation and Complexity in Algebraic Geometry". PhD thesis. Leipzig, Germany: Universität Leipzig, 2026. URL: https://nbn-resolving.org/urn:nbn:de:bsz:15-qucosa2-1067809 (cit. on pp. 16, 21, 28).

[KL26] Leonie Kayser and Sibren Lagauw. *Code repository accompanying the paper "Porteous Meets Walsh: Counting Critical Points for $\mathcal{H}_2$-Norm Model Order Reduction Using Algebraic Geometry Behind Multiparameter Eigenvalue Problems"*. 2026. URL: https://github.com/leokayser/h2mor (cit. on p. 17).

[KL26+] Leonie Kayser and Sibren Lagauw. "Generic and specific Euclidean distance degrees of secant varieties to the rational normal curve". Manuscript in preparation. 2026 (cit. on pp. 3, 7, 21).

[LAD23] Sibren Lagauw, Oscar Mauricio Agudelo, and Bart De Moor. "Globally Optimal SISO $\mathcal{H}_2$-Norm Model Reduction Using Walsh's Theorem". In: *IEEE Control Systems Letters* 7 (2023), pp. 1670–1675. DOI: 10.1109/lcsys.2023.3276299 (cit. on pp. 2, 9, 17).

[Lag26] Sibren Lagauw. "Global Solution to Problems in System Identification and Model Order Reduction". PhD thesis. Leuven, Belgium: KU Leuven, 2026 (cit. on p. 21).

[Lax59] Peter D. Lax. "Translation invariant spaces". In: *Acta Mathematica* 101.3–4 (1959), pp. 163–178. DOI: 10.1007/bf02559553 (cit. on p. 5).

[LVD26] Sibren Lagauw, Lukas Vanpoucke, and Bart De Moor. "Exact solution to the least squares realization problem as a multiparameter eigenvalue problem". In: *Automatica* 185 (2026), p. 112792. DOI: 10.1016/j.automatica.2025.112792 (cit. on p. 1).

[ML67] Lewis Meier III and David G. Luenberger. "Approximation of linear constant systems". In: *IEEE Transactions on Automatic Control* 12.5 (1967), pp. 585–588. DOI: 10.1109/tac.1967.1098680 (cit. on pp. 4, 6).

[OSS14] Giorgio Ottaviani, Pierre-Jean Spaenlehauer, and Bernd Sturmfels. "Exact Solutions in Structured Low-Rank Approximation". In: *SIAM Journal on Matrix Analysis and Applications* 35.4 (2014), pp. 1521–1542. DOI: 10.1137/13094520x (cit. on p. 3).

[Par04] Jonathan R. Partington. *Linear Operators and Linear Systems: An Analytical Approach to Control Theory*. Vol. 60. London Mathematical Society Student Texts. Cambridge: Cambridge University Press, 2004. DOI: 10.1017/cbo9780511616693 (cit. on pp. 5, 16).

[PBH12] Ralf Peeters, Ivo Bleylevens, and Bernard Hanzon. "On algebraic and linear algebraic aspects of co-order three $\mathcal{H}_2$-model order reduction". In: *IFAC Proceedings Volumes* 45.16 (2012), pp. 704–709. DOI: 10.3182/20120711-3-be-2027.00379 (cit. on p. 2).

[Por71] Ian R. Porteous. "Simple singularities of maps". In: *Proceedings of Liverpool Singularities—Symposium I*. Ed. by Charles T. C. Wall. Vol. 192. Lecture Notes in Mathematics. Berlin: Springer, 1971, pp. 286–307. DOI: 10.1007/bfb0066829 (cit. on p. 7).

[Reg95] Phillip A. Regalia. *Adaptive IIR Filtering in Signal Processing and Control*. Vol. 90. Electrical Engineering and Electronics. New York: Marcel Dekker, 1995. DOI: 10.1201/9781315136653 (cit. on pp. 2, 4, 5, 20, 21).

[Sha13] Igor R. Shafarevich. *Basic Algebraic Geometry 1: Varieties in Projective Space*. 3rd ed. Berlin: Springer, 2013. ISBN: 9783642379567. DOI: 10.1007/978-3-642-37956-7 (cit. on pp. 7, 22, 24, 26).

[SP26] Belal Sefat Panah. "The computation of the Euclidean distance degree for the middle Catalacticant for the binary forms". In: *Communications in Algebra* 54.2 (2026), pp. 610–618. DOI: 10.1080/00927872.2025.2532175 (cit. on p. 3).

[SW05] Andrew J. Sommese and Charles W. Wampler II. *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science*. River Edge, NJ: World Scientific, 2005. ISBN: 9789812567727. DOI: 10.1142/5763 (cit. on p. 22).

[VD19] Christof Vermeersch and Bart De Moor. "Globally Optimal Least-Squares ARMA Model Identification is an Eigenvalue Problem". In: *IEEE Control Systems Letters* 3.4 (2019), pp. 1062–1067. DOI: 10.1109/lcsys.2019.2920765 (cit. on p. 1).

[VD23] Christof Vermeersch and Bart De Moor. "Two Double Recursive Block Macaulay Matrix Algorithms to Solve Multiparameter Eigenvalue Problems". In: *IEEE Control Systems Letters* 7 (2023), pp. 319–324. DOI: 10.1109/lcsys.2022.3182632 (cit. on pp. 1, 2, 9).

[VD26] Christof Vermeersch and Bart De Moor. *Solving Multivariate Polynomial Systems and Rectangular Multiparameter Eigenvalue Problems with MacaulayLab*. 2026. arXiv: 2605.20884 [cs.MS] (cit. on pp. 2, 9).

[VGA10] Paul Van Dooren, Kyle A. Gallivan, and Pierre-Antoine Absil. "$\mathcal{H}_2$-Optimal Model Reduction with Higher-Order Poles". In: *SIAM Journal on Matrix Analysis and Applications* 31.5 (2010), pp. 2738–2753. DOI: 10.1137/080731591 (cit. on pp. 5, 20).

[Vol88] Hans Volkmer. *Multiparameter Eigenvalue Problems and Expansion Theorems*. Vol. 1356. Lecture Notes in Mathematics. Berlin: Springer, 1988. ISBN: 9783540460152. DOI: 10.1007/bfb0089295 (cit. on p. 1).

[Wal60] Joseph L. Walsh. *Interpolation and Approximation by Rational Functions in the Complex Domain*. 3rd ed. Vol. 20. American Mathematical Society Colloquium Publications. Providence, RI: American Mathematical Society, 1960 (cit. on pp. 2, 5).