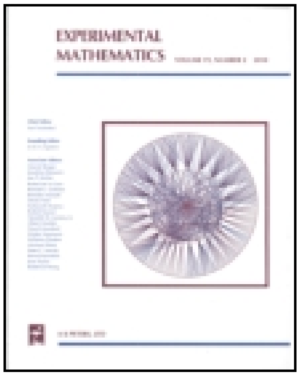

Experimental Mathematics

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/uexm20

Equations for Lower Bounds on Border Rank

Jonathan D. Hauenstein [a] , Christian Ikenmeyer [b] & J. M. Landsberg [b]

[a] Department of Mathematics , North Carolina State University , Box 8205, Raleigh , NC , 27695-8205 , USA

[b] Department of Mathematics , Texas A&M University , Mailstop 3368, College Station , TX , 77843-3368 , USA

Published online: 09 Dec 2013.

To cite this article: Jonathan D. Hauenstein , Christian Ikenmeyer & J. M. Landsberg (2013) Equations for Lower Bounds on Border Rank, Experimental Mathematics, 22:4, 372-383, DOI: 10.1080/10586458.2013.825892

To link to this article: http://dx.doi.org/10.1080/10586458.2013.825892

Experimental Mathematics, 22(4):372–383, 2013

ISSN: 1058-6458 print / 1944-950X online
DOI: 10.1080/10586458.2013.825892

Taylor & Francis
Taylor & Francis Group

Equations for Lower Bounds on Border Rank

Jonathan D. Hauenstein, Christian Ikenmeyer, and J. M. Landsberg

CONTENTS

We present new methods for determining polynomials in the ideal of the variety of bilinear maps of border rank at most r. We apply these methods to several cases including the case $r = 6$ in the space of bilinear maps $\mathbb{C}^4 \times \mathbb{C}^4 \to \mathbb{C}^4$. This space of bilinear maps includes the matrix multiplication operator M_2 for 2×2 matrices. We show that these newly obtained polynomials do not vanish on the matrix multiplication operator M_2, which gives a new proof that the border rank of the multiplication of 2×2 matrices is seven. Other examples are considered along with an explanation of how to implement the methods.

1. INTRODUCTION

Lower bounds in complexity theory are considered difficult to obtain. We describe a new method for obtaining lower bounds on the *border rank* that is based on a new way to find polynomials that vanish on bilinear maps $T : \mathbb{C}^{\mathbf{a}} \times \mathbb{C}^{\mathbf{b}} \to \mathbb{C}^{\mathbf{c}}$ of low border rank.

1.1. Rank and Border Rank

Let $\mathbb{C}^{\mathbf{a}*} := \{f : \mathbb{C}^{\mathbf{a}} \to \mathbb{C} \mid f \text{ is linear}\}$ denote the dual vector space to $\mathbb{C}^{\mathbf{a}}$. That is, if an element of $\mathbb{C}^{\mathbf{a}}$ is represented by a column vector of height $\mathbf{a}$, then $\mathbb{C}^{\mathbf{a}*}$ corresponds to row vectors, and the evaluation is just row–column matrix multiplication. A bilinear map $T : \mathbb{C}^{\mathbf{a}} \times \mathbb{C}^{\mathbf{b}} \to \mathbb{C}^{\mathbf{c}}$ has *rank one* if there exist $\alpha \in \mathbb{C}^{\mathbf{a}*}$, $\beta \in \mathbb{C}^{\mathbf{b}*}$, and $c \in \mathbb{C}^{\mathbf{c}}$ such that $T(a, b) = \alpha(a)\beta(b)c$. The rank-one bilinear maps are in some sense the simplest bilinear maps, and T is said to have *rank* r if r is the minimum number of rank-one bilinear maps that sum to T. This r is sometimes called the *tensor rank* of T.

If one views multiplication by constants as a "free" operation, then the rank differs by at most a factor of two from the minimal number of multiplications of variables that is needed to compute T; see [Bürgisser et al. 97, Chapter 14] for more information. Since the set of all bilinear maps $\mathbb{C}^{\mathbf{a}} \times \mathbb{C}^{\mathbf{b}} \to \mathbb{C}^{\mathbf{c}}$ is a vector space of dimension $\mathbf{abc}$, it is natural to talk about polynomials on the space of bilinear maps $\mathbb{C}^{\mathbf{a}} \times \mathbb{C}^{\mathbf{b}} \to \mathbb{C}^{\mathbf{c}}$.

2000 AMS Subject Classification: 68Q17
Keywords: border rank, matrix multiplication

Unfortunately, one cannot test directly for the tensor rank by the vanishing of polynomials, since the common zero locus of the set of all polynomials vanishing on the set of bilinear maps of rank at most r is, typically, larger than the set of bilinear maps of rank at most r. This may be described precisely using the language of algebraic geometry: for the purposes of this article, we define an *algebraic variety* (or simply a *variety*) to be the common zero locus of a collection of polynomials that is *irreducible*, in the sense that it cannot be written as a union of two zero loci.

A (proper) *Zariski closed subset* of a variety X is the common zero locus of a collection of polynomials restricted to X, and a *Zariski open subset* is the complement of a Zariski closed set. The *border rank* of a tensor T is defined to be the smallest r such that all polynomials vanishing on the set of bilinear maps of rank at most r also vanish at T, and one writes $\underline{\mathbf{R}}(T) = r$. In this case, T is arbitrarily close, in any reasonable measure, to a bilinear map of rank r (including the possibility that the rank of T is r). We let $\sigma_{r;\mathbf{a},\mathbf{b},\mathbf{c}} \subset \mathbb{C}^{\mathbf{a}} \otimes \mathbb{C}^{\mathbf{b}} \otimes \mathbb{C}^{\mathbf{c}}$ denote the set of bilinear maps of border rank at most r. It is the algebraic variety formed from the zero set of all the polynomials having the set of bilinear maps of rank at most r in their zero set.

When $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are understood from the context, we simply write σ_r. The set of bilinear maps of rank r is a Zariski open subset of the algebraic variety $\sigma_{r;\mathbf{a},\mathbf{b},\mathbf{c}}$. The set is open because the set of bilinear maps of border rank less than r is a closed subset of $\mathbb{C}^{\mathbf{a}} \otimes \mathbb{C}^{\mathbf{b}} \otimes \mathbb{C}^{\mathbf{c}}$, and the subset of $\sigma_{r;\mathbf{a},\mathbf{b},\mathbf{c}}$ of bilinear maps of rank greater than r is closed in $\sigma_{r;\mathbf{a},\mathbf{b},\mathbf{c}}$.

1.2. Results

We introduce a new technique based on numerical algebraic geometry and interpolation that finds, with high probability, where equations that vanish on the variety of bilinear maps of border rank at most r can be found. Once one knows where to look, we can use methods that were introduced in [Landsberg and Manivel 04] and refined in [Bates and Oeding 11, Bürgisser et al. 13] to find the actual equations and rigorously prove that they vanish on $\sigma_{r;\mathbf{a},\mathbf{b},\mathbf{c}}$. With these equations, one can then show that the border rank of a given tensor T is greater than r if T does not satisfy these equations. Of special interest in this paper will be the border rank of the matrix multiplication tensor

$$M_2 := \sum_{i,j,k=1}^{2} e_{i,j} \otimes e_{j,k} \otimes e_{k,i} \in \mathbb{C}^4 \otimes \mathbb{C}^4 \otimes \mathbb{C}^4,$$

where $(e_{i,j})$ is the standard basis of $\mathbb{C}^{2\times 2} = \mathbb{C}^4$.

It has been known since Strassen's fundamental breakthrough [Strassen 69] that $\underline{\mathbf{R}}(M_2) \leq 7$. Our main result is a new proof of the lower bound

$$\underline{\mathbf{R}}(M_2) \geq 7, \qquad (1\text{–}1)$$

which was originally proven in 2005 by Landsberg; see [Landsberg 06], where very different methods were used (see Section 5).

The proof outline is as follows. We start by proving, with the aid of a computer, that no nonconstant polynomial of degree less than 19 vanishes on $\sigma_{6;4,4,4}$. In the course of this computation, we compute the necessary data to perform the membership test of [Hauenstein and Sommese 13], which numerically shows (i.e., shows with extremely high probability) that (1–1) holds. Additionally, the same data give strong evidence that there is a 64-dimensional space of degree-19 equations that vanish on $\sigma_{6;4,4,4}$. The only 64-dimensional representation of GL_4 in $\mathbb{C}[\mathbb{C}^4 \otimes \mathbb{C}^4 \otimes \mathbb{C}^4]_{19}$ is of type $((5,5,5,4),(5,5,5,4),(5,5,5,4))$. By a randomized procedure, we then construct a basis for the 31-dimensional highest-weight vector space of weight

$$((5,5,5,4),(5,5,5,4),(5,5,5,4)) \text{ in } \mathbb{C}[\mathbb{C}^4 \otimes \mathbb{C}^4 \otimes \mathbb{C}^4]_{19}.$$

We show using numerical methods that the restriction of this 31-dimensional vector space to functions defined on $\sigma_{6;4,4,4}$ has a 1-dimensional kernel.

Since the highest-weight space of weight

$$((5,5,5,5),(5,5,5,5),(5,5,5,5)) \text{ in } \mathbb{C}[\mathbb{C}^4 \otimes \mathbb{C}^4 \otimes \mathbb{C}^4]_{20}$$

is only 4-dimensional, we focus on this space to develop a rigorous proof of (1–1). In particular, the restriction to functions defined on $\sigma_{6;4,4,4}$ also has a 1-dimensional kernel. This corresponds to a degree-20 polynomial that vanishes on $\sigma_{6;4,4,4}$ and does not vanish at M_2, thereby completing the proof.

We remark that while there is a large subspace of $\mathbb{C}[\mathbb{C}^4 \otimes \mathbb{C}^4 \otimes \mathbb{C}^4]_{20}$ that vanishes on $\sigma_{6;4,4,4}$, the polynomial we work with is distinguished in that it is the only one that is unchanged (up to scale) by changes of bases in each of the $\mathbb{C}^4$'s.

To make this approach computationally feasible, each polynomial is represented by a pair of permutations; see Section 4.3. These permutations provide all the information needed to evaluate the corresponding polynomial, but unfortunately, this means that we are unable to obtain additional information, such as the number of terms.

The technique we present can be applied to other *implicitization problems*. That is, we want to consider a variety

$$X := \overline{g(Y)}, \qquad (1\text{–}2)$$

where Y is (possibly a Zariski open subset of) a variety, and g is a system of rational functions defined on Y. In the bilinear case, the tensors of rank at most r form a dense subset of the algebraic set of tensors of border rank at most r, where each tensor of rank at most r can be written as a sum of r tensors of rank at most one. In this case, Y is simply the product of r copies of the variety of tensors of rank at most one, and g is the linear map corresponding to taking the sum. Another specific application arising in physics is the analysis of vacuum moduli spaces in (supersymmetric) field theories [Hauenstein et al. 12], which arise as the closure of the image under a polynomial map of an algebraic set.

Besides our main result, we have proved, using numerical methods, the following:

1. The degree of $\sigma_{6;4,4,4}$ is 15 456.
2. The degree of the codimension-three variety $\sigma_{15;4,8,9}$ is at least 83 000, and no nonconstant polynomial of degree at most 45 vanishes on $\sigma_{15;4,8,9}$.
3. The degree of the hypersurface $\sigma_{18;7,7,7}$ is at least 187 000.
4. The degree of the codimension-six variety $\sigma_{6;3,4,6}$ is 206 472, and no nonconstant polynomial of degree at most 14 vanishes on $\sigma_{6;3,4,6}$.
5. The degree of the codimension-three variety $\sigma_{7;4,4,5}$ is 44 000, and no nonconstant polynomial of degree at most 56 vanishes on $\sigma_{7;4,4,5}$.
6. The degree of the hypersurface $\sigma_{8;3,5,7}$ is 105.

The varieties $\sigma_{15;4,8,9}$ and $\sigma_{18;7,7,7}$ have applications to 3×3 matrix multiplication. Information about polynomials in their ideals could help to determine $\underline{\mathbf{R}}(M_3)$ more precisely, with the current known bounds being $15 \leq \underline{\mathbf{R}}(M_3) \leq 21$; see [Schönhage 81, Landsberg and Ottaviani 11]. The other varieties are presented, since they are at the margin of what is currently feasible. Results regarding the ideal being empty are potentially useful for finding further equations, since they provide a starting point for such an endeavor.

1.3. Other Methods for Finding Equations

Very little is known about the equations of σ_r in general. One can reduce to the case of $\mathbf{a} = \mathbf{b} = \mathbf{c} = r$ via a process called *inheritance*. Additionally, there is a systematic way to determine the equations in any given degree using *multiprolongation*. For a discussion on inheritance and multiprolongation, see [Landsberg 12, Section 3.7]. Even though multiprolongation is systematic, it is very difficult to utilize except in very small cases. Most known equations have been found by reducing multilinear algebra to linear algebra. See [Landsberg 12, Landsberg and Ottaviani 11] for the most recent equations, which go up to $\sigma_{2m-2;m,m,m}$. Some information about the ideal of $\sigma_{r;r,r,r}$ can be found using representation theory (via the algebraic Peter–Weyl theorem), since this case is an orbit closure; see [Bürgisser and Ikenmeyer 11] for an exposition. By inheritance, one can deduce the $\sigma_{r;m,n,p}$ case for any m, n, p from the $\sigma_{r;r,r,r}$ case.

1.4. Polynomials on Vector Spaces

We write $I(\sigma_r)$ for the set of all polynomials vanishing on σ_r, which forms an *ideal*. Since σ_r is invariant under rescaling, we may restrict our attention to homogeneous polynomials, since in this case, a polynomial will be in the ideal if and only if all of its homogeneous components are in the ideal.

Let V be a vector space. A subset $X \subset V$ is called an *algebraic set* if it is the common zero locus of a collection of polynomials on V. Recall that we say that an irreducible algebraic set is a *variety*. If $X \subset V$ is a variety that is invariant under rescaling, let $S^d V^*$ be the space of homogeneous polynomials of degree d on V and let $I_d(X) \subset S^d V^*$ be the component of the ideal of X in degree d.

Roughly speaking (see Section 2 for more details), our technique for studying the equations that vanish on a variety X of positive dimension is to apply techniques of numerical algebraic geometry to finite subsets of X that lie in a common linear space. That is, we aim to study finite subsets of algebraic sets of the form $Y = X \cap \mathcal{L} \subset \mathcal{L}$, where $\mathcal{L}$ is a general linear space of codimension at most $\dim X$. If the codimension of $\mathcal{L}$ is $\dim X$, then Y consists of $\deg X$ points. If the codimension of $\mathcal{L}$ is strictly less than $\dim X$, then Y is also a variety with the same degree as X. Moreover, if one considers $X \subset V$ and $Y \subset \mathcal{L}$, and defines d_X and d_Y to be the minimal degrees of the nonzero polynomials in $I(X)$ and $I(Y)$, respectively, then $d_X \geq d_Y$.

In particular, $\dim I_d(Y) \geq \dim I_d(X)$ for every $d \leq d_X$ with similar bounds for all $d \geq 0$ that can be developed from the corresponding Hilbert functions.

Once we have inferred information about polynomials in $I(Y)$, we use representation theory to identify which modules can appear. Finally, sample vectors from these modules are used to test whether the entire module is in the ideal $I(X)$.

2. DECIDING WHERE TO GO HUNTING

The basic idea of our algorithm is to combine the ability of numerical algebraic geometry to compute points on certain subsets of a variety with interpolation to obtain information about this subset from the computed points. We first describe needed concepts from numerical algebraic geometry and then give a brief discussion regarding interpolation.

At a basic level, the algorithms of numerical algebraic geometry (see [Sommese and Wampler 05] for general background information) perform numerical computations on varieties each of which is represented by a data structure called a *witness set*.

Let f be a polynomial system. The common zero locus of f is an algebraic set that can be decomposed uniquely into finitely many varieties, none of which is contained in the union of the others. If X is one of these varieties, called an *irreducible component* of the zero locus of f, then a witness set for X is the triple $\{f, L, W\}$, where the zero set of L defines a general linear subspace of codimension equal to the dimension of X, and W is the intersection of X with this linear subspace defined by L. Given one point in W, arbitrarily many points on X can be computed in a process called *sampling*. In numerical terms, computing a point "on" a variety means that we have a numerical approximation of the point along with an algorithm that can be used to approximate the point to arbitrary accuracy.

This witness set description is not useful for the problems at hand, since for each of the varieties X under consideration, we do not assume that we have access to a polynomial system f or even any nonzero polynomials that vanish on X. However, we do assume that we have a description of X in the form (1–2). In fact, by adding variables and clearing denominators, we can assume that $X := \overline{\pi(Z)}$, where π is a projection map and Z is an irreducible component of the zero locus for some polynomial system F. This is demonstrated in the following simple example.

Example 2.1. The set $X := \{(x, y) \in \mathbb{C}^2 \mid x^2 + y^2 = 1\}$ is equal to $\overline{g(Y)}$, where

$$g(t) = \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right) \quad \text{and} \quad Y := \mathbb{C} \setminus \{\pm i\}.$$

We also have $X = \overline{\pi(Z)}$, where $\pi(x, y, t) = (x, y)$ and Z is the zero locus (which is irreducible) of

$$F(x, y, t) = \begin{bmatrix} (1+t^2)x - (1-t^2) \\ (1+t^2)y - 2t \end{bmatrix}.$$

With this setup, we utilize a *pseudowitness set* [Hauenstein and Sommese 10, Hauenstein and Sommese 13] for $X = \overline{\pi(Z)}$, which is the quadruple $\{F, \pi, L, W\}$, where L defines a linear subspace of codimension equal to the dimension of Z, and W is the intersection of Z with this linear subspace defined by L. In this case, the linear polynomials L are constructed so that there are exactly $\dim X$ general linear polynomials in the image space of π, i.e., the linear space intersects X in $\deg X$ points, while the remaining linear polynomials are general. In particular, $\pi(W)$ consists of exactly $\deg X$ distinct points. As with traditional witness sets, one can sample and perform membership tests on X [Hauenstein and Sommese 13].

The key here is that once a single sufficiently general point is known on X, other points on X can be computed as well. In fact, these other points can be forced to live in a fixed general linear subspace of codimension at most $\dim X$, thereby simplifying the future computations, since one can work intrinsically on this linear subspace. If the intersection of the linear subspace and X is of positive dimension, then it is also a variety, and arbitrarily many points can be sampled from this variety. If the intersection is zero-dimensional, it consists of exactly $\deg X$ points, from which, after one point is computed, random monodromy loops [Sommese et al. 01] could be used to try to compute the other points. The trace test [Sommese et al. 02] provides a stopping criterion for deciding when exactly $\deg X$ points have been computed.

Clearly, any polynomial that vanishes on X must also vanish on a finite subset of X. Although we will not delve too deeply into the theory here, one can recover the invariants of X from a general linear subspace section of X when X is an *arithmetically Cohen–Macaulay* scheme (see [Migliore 98, Chapter 1]). Nonetheless, since our current focus is on developing a list of potential places where one might focus further representation-theoretic computations, we can consider all varieties and not just the arithmetically Cohen–Macaulay ones. Of course, this is at the expense of bounds rather than equality, as demonstrated in the following example.

Example 2.2. Consider the following varieties in $\mathbb{P}^3$:

$$X_1 := \left\{\left(s^3, s^2t, st^2, t^3\right) \mid (s,t) \in \mathbb{P}^1\right\}$$

and

$$X_2 := \left\{\left(s^4, s^3t, st^3, t^4\right) \mid (s,t) \in \mathbb{P}^1\right\}.$$

It is easy to verify that

- $\dim X_1 = 1$, $\deg X_1 = 3$, and $I(X_1)$ is generated by three quadratics;

- $\dim X_2 = 1$, $\deg X_2 = 4$, and $I(X_2)$ is generated by a quadratic and three cubics.

Let $Y_i = X_i \cap \mathcal{H}$ be the set of $\deg X_i$ points, where $\mathcal{H}$ is the hyperplane defined by the vanishing of $\ell(x) = x_0 + 2x_1 + 3x_2 + 5x_3$. If we consider $Y_i \subset \mathcal{H}$, then

- $I(Y_1)$ is generated by three quadratics;
- $I(Y_2)$ is generated by two quadratics.

To summarize, X_1 is the twisted cubic curve in $\mathbb{P}^3$ that is arithmetically Cohen–Macaulay, so that, for example, the dimension of $I_d(X_1)$ can be determined from $I_d(Y_1)$. However, X_2 is not arithmetically Cohen–Macaulay, which in this case, can be observed because $2 = \dim I_2(Y_2) > \dim I_2(X_2) = 1$. Even though one should expect only $d_{X_2} \geq d_{Y_2}$, we have equality in this case, namely $d_{X_2} = d_{Y_2} = 2$.

Once we have decided on our first finite set to consider, the next task is *polynomial interpolation*, that is, to compute polynomials that vanish on this finite set. Given a basis for the finite-dimensional space of polynomials under consideration, polynomial interpolation reduces to computing null vectors of a (potentially very large) matrix. From a numerical standpoint, as the degrees of the polynomials under consideration increase, preconditioning becomes essential if one is to perform reliable computations. For our computations, we use the approach of [Griffin et al. 14].

Each computation provides some restrictions as to which polynomials can be in $I(X)$. Nevertheless, we also consider what happens when we add new points to our finite set. For a particular degree, there are two possible choices: either the originally computed polynomials will vanish at the new points, or the dimension of the set of polynomials that vanish at all the points will decrease. In the former case, we can then move on to searching for higher-degree polynomials not generated by these polynomials. In the latter case, we continue adding new points. If no polynomials of a particular degree, say d, vanish on some finite set, then we know that $\dim I_d(X) = 0$ and $d_X > d$. Thus, we try again by considering polynomials of degree $d + 1$.

Variations of this approach involve sampling points from the intersection of X with linear spaces of increasing dimension to see how the dimension of the vanishing polynomials changes as fewer restrictions are placed on the sample points. The key, in the end, is to control the growth of the dimension of the space of polynomials under consideration, since this can quickly become unwieldy. In particular, this method is practical for varieties X of low codimension, since we can work implicitly on linear spaces of low dimension.

When the codimension is one, X is a hypersurface, so that the degree of X is equal to the degree of the polynomial defining X. In this case, one can simply compute a pseudowitness set to compute its degree rather then use this interpolation-based approach. For example, such an approach was used in [Blekherman et al. 12] for computing the degree of implicitly defined hypersurfaces that arise as the algebraic boundaries of Hilbert's sums-of-squares cones of degree 38 475 and 83 200.

Example 2.3. In Example 2.2, we considered finite sets obtained by intersecting the curves with a particular hyperplane. We now use this information to limit our focus when we add other points to our finite set. In four variables, there is a ten-dimensional space of homogeneous polynomials of degree 2, but with our previously computed information, this has already been reduced to seven- and six-dimensional spaces for X_1 and X_2, respectively, more specifically, the four-dimensional space arising from the linear polynomial $\ell(x)$ along with the three- and two-dimensional spaces, respectively, from $I_2(Y_1)$ and $I_2(Y_2)$, namely

$$\begin{aligned} &I_2(X_1) \\ &\quad \subset \operatorname{span}\{x_0\ell(x),\ x_1\ell(x),\ x_1x_2 + 2x_1x_3 + 3x_2x_3 + 5x_3^2, \\ &\qquad x_2\ell(x),\ x_3\ell(x),\ x_2^2 - x_1x_3, \\ &\qquad x_1^2 - x_1x_3 - x_2x_3 - 10x_3^2\} \end{aligned}$$

and

$$\begin{aligned} &I_2(X_2) \\ &\quad \subset \operatorname{span}\{x_0\ell(x),\ x_1\ell(x),\ x_1x_2 + 2x_1x_3 + 3x_2x_3 + 5x_3^2, \\ &\qquad x_2\ell(x),\ x_3\ell(x), x_1^2 - x_2^2 + 11x_1x_3 + 2x_2x_3 + 20x_3^2\}. \end{aligned}$$

By selecting additional random points, one indeed obtains $\dim I_2(X_1) = 3$ and $\dim I_2(X_2) = 1$.

This procedure develops ideas on the degrees d_j for which generators of the ideal appear.

The next section summarizes the numerical evidence for the results presented in Section 1.2. From these data, the next step is to determine conclusively the linear subspace of the space of polynomials of degrees d_j that are in the ideal. For this, one uses representation theory, as we describe in Section 4.

3. REVIEW OF NUMERICAL RESULTS

We summarize the six varieties presented in Section 1.2. In all these cases, the codimension of the variety is the expected codimension, namely $\operatorname{codim}\sigma_{r;\mathbf{a},\mathbf{b},\mathbf{c}} = \mathbf{abc} - r(\mathbf{a}+\mathbf{b}+\mathbf{c}-2)$. The points on the varieties were computed using Bertini [Bates et al. 06], with the linear algebra computations performed in Matlab.

3.1. The Variety $\sigma_{6;4,4,4}$

As discussed above, the codimension-four variety $\sigma_{6;4,4,4}$ provides information about $\underline{\mathbf{R}}(M_2)$, that is, showing that $M_2 \notin \sigma_{6;4,4,4} \subset \mathbb{P}^{63}$ shows that $\underline{\mathbf{R}}(M_2) = 7$. We first fix a random linear space $\mathcal{L} \subset \mathbb{P}^{63}$ of dimension 4 and consider the finite set $W := \sigma_{6;4,4,4} \cap \mathcal{L}$. The first objective is to compute points in W, with a goal of computing every point in W. To this end, we first computed one point in W as follows. We chose a random point $x^* \in \sigma_{6;4,4,4}$, which is trivial, since a dense subset is $\sigma_{6;4,4,4}$ is parameterizable. Let L be a system of 59 linear forms, so that $\mathcal{L}$ is the zero locus of L, and let $\mathcal{L}_{t,x^*}$ be the zero locus of $L(x) - t \cdot L(x^*)$. Since $x^* \in \sigma_{6;4,4,4} \cap \mathcal{L}_{1,x^*}$, a point in W is the endpoint of the path defined by $\sigma_{6;4,4,4} \cap \mathcal{L}_{t,x^*}$ at $t = 0$ starting from x^* at $t = 1$.

Even though the above process could be repeated for different x^* to compute points in W, we instead used monodromy loops [Sommese et al. 01] for generating more points in W. After 21 loops had been performed, the number of points in W that had been computed stabilized at 15 456. The trace test [Sommese et al. 02] shows that 15 456 is indeed the degree of $\sigma_{6;4,4,4}$, thereby showing that we had indeed computed W.

From W, we performed two computations. The first was the membership test of [Hauenstein and Sommese 13] for deciding whether $M_2 \in \sigma_{6;4,4,4}$, which requires tracking 15 456 homotopy paths that start at the points of W. In this case, each of these 15 456 paths converged to points in $\sigma_{6;4,4,4}$ distinct from M_2, providing a numerical proof that $M_2 \notin \sigma_{6;4,4,4}$.

The second was to compute the minimal degree of nonzero polynomials vanishing on $W \subset \mathcal{L}$. This sequence of polynomial interpolation problems showed that no nonconstant polynomials of degree ≤ 18 vanish on W, and hence on $\sigma_{6;4,4,4}$. The $15\,456 \times 8855$ matrix resulting from polynomial interpolation of homogeneous forms of degree 19 in five variables using the approach of [Griffin et al. 14] has a 64-dimensional null space. Thus, the minimal degree of nonzero polynomials vanishing on $W \subset \mathcal{L}$ is 19.

The next objective was to verify that the minimal degree of nonzero polynomials vanishing on the curve $C := \sigma_{6;4,4,4} \cap \mathcal{K} \subset \mathcal{K}$ for a fixed random linear space $\mathcal{K} \subset \mathbb{P}^{63}$ of dimension 5 is also 19. We used 50 000 points on C, and the $50\,000 \times 42\,504$ matrix resulting from polynomial interpolation of homogeneous forms of degree 19 in six variables using the approach of [Griffin et al. 14] also has a 64-dimensional null space. With this agreement, we proceeded to use representation theory, described in Section 4, to understand these polynomials and prove that M_2 is indeed not contained in $\sigma_{6;4,4,4}$.

3.2. The Variety $\sigma_{15;4,8,9}$

For the codimension-three variety $\sigma_{15;4,8,9}$, we followed a similar computation as above for computing 83 000 points in $W := \sigma_{15;4,8,9} \cap \mathcal{L}$, where $\mathcal{L} \subset \mathbb{P}^{287}$ is a random linear space of dimension 3. Using polynomial interpolation on these points, we were able to show that no nonconstant polynomial of degree ≤ 45 vanishes on $\sigma_{15;4,8,9}$.

3.3. The Variety $\sigma_{18;7,7,7}$

For the hypersurface $\sigma_{18;7,7,7}$, we computed 187 000 points in $W := \sigma_{18;7,7,7} \cap \mathcal{L}$, where $\mathcal{L} \subset \mathbb{P}^{342}$ is a random line. This shows that 187 000 is a lower bound on the degree of $\sigma_{18;7,7,7}$ and the degree of the polynomial that defines it.

3.4. The Variety $\sigma_{6;3,4,6}$

For the codimension-six variety $\sigma_{6;3,4,6}$, we followed a similar computation as above for computing $W := \sigma_{6;3,4,6} \cap \mathcal{L}$, where $\mathcal{L} \subset \mathbb{P}^{71}$ is a random linear space of dimension 6. In this case, the trace test shows that the set of 206 472 points computed by monodromy indeed equals W. Polynomial interpolation showed that no nonconstant polynomial of degree ≤ 14 vanishes on $\sigma_{6;3,4,6}$. We stopped at degree 14 due to memory limitations of the numerical linear algebra routines. However, even though we were unable to compute the minimal degree of nonconstant polynomials vanishing on $W \subset \mathcal{L}$, we note that W with [Hauenstein and Sommese 13] can be used to decide membership in $\sigma_{6;3,4,6}$.

3.5. The Variety $\sigma_{7;4,4,5}$

For the codimension-three variety $\sigma_{7;4,4,5}$, we followed a similar computation as above for computing $W := \sigma_{7;4,4,5} \cap \mathcal{L}$, where $\mathcal{L} \subset \mathbb{P}^{79}$ is a random linear space of dimension 3. In this case, the trace test shows that the set of 44 000 points computed by monodromy indeed equals W. Polynomial interpolation showed that no nonconstant polynomial of degree ≤ 56 vanishes on $\sigma_{7;4,4,5}$. We stopped at degree 56 due to additional conditioning problems arising from the numerical linear algebra routines. As with $\sigma_{6,3,4,6}$, W with [Hauenstein and Sommese 13] still can be used to decide membership in $\sigma_{7;4,4,5}$.

3.6. The Variety $\sigma_{8;3,5,7}$

For the hypersurface $\sigma_{8;3,5,7}$, the trace showed that the set of 105 points computed by monodromy is equal to $W := \sigma_{8;3,5,7} \cap \mathcal{L}$, where $\mathcal{L} \subset \mathbb{P}^{104}$ is a random line. In particular, this shows that there is a degree-105 polynomial vanishing on $\sigma_{8;3,5,7}$.

4. POLYNOMIALS ON THE SPACE OF BILINEAR MAPS

4.1. Tensors

In order to explain the polynomials, it will be useful to work more invariantly, so instead of $\mathbb{C}^{\mathbf{a}}$, $\mathbb{C}^{\mathbf{b}}$, etc., we write A, B, etc., for complex vector spaces of dimensions $\mathbf{a}, \mathbf{b}$ etc. It will also be useful to introduce the language of *tensors*. A bilinear map $A^* \times B^* \to C$ may also be viewed as a trilinear map $A^* \times B^* \times C^* \to \mathbb{C}$, as well as in numerous other ways. To avoid prejudicing ourselves, we simply write $T \in A \otimes B \otimes C$ for any of these manifestations and call T a *tensor*. Just as we may view a linear map as a matrix after fixing bases, such a T may be viewed as a three-dimensional matrix after fixing bases. Note that $A \otimes B \otimes C$, the set of all such tensors, is a vector space of dimension $\mathbf{abc}$. More generally, given vector spaces $A_1, \ldots, A_k$, one can define the space of tensors $A_1 \otimes \cdots \otimes A_k$. There is a natural map

$$\begin{aligned} A_1 \otimes \cdots \otimes A_k \times B_1 \otimes \cdots \otimes B_l \\ \to A_1 \otimes \cdots \otimes A_k \otimes B_1 \otimes \cdots \otimes B_l, \\ (f, g) \mapsto f \otimes g, \end{aligned}$$

where

$$f \otimes g(\alpha_1, \ldots, a_k, \beta_1, \ldots, \beta_l) := f(\alpha_1, \ldots, a_k) g(\beta_1, \ldots, \beta_l).$$

4.2. Remarks on the Theory

We briefly review the representation theory underlying the algorithm. For more details, see [Landsberg 12, Chapter 6]. Let $S^d(A \otimes B \otimes C)^*$ denote the vector space of homogeneous polynomials of degree d on $A \otimes B \otimes C$. The variety $\sigma_{r;\mathbf{a},\mathbf{b},\mathbf{c}}$ is mapped to itself under changes of bases in each of the vector spaces, and thus if we have one equation, we can obtain many more by changing bases. That is, let $\mathrm{GL}(A)$ denote the set of invertible linear maps $A \to A$, and similarly for B, C. The group $G := \mathrm{GL}(A) \times \mathrm{GL}(B) \times \mathrm{GL}(C)$ acts on $A \otimes B \otimes C$ by

$$(g_A, g_B, g_C) \cdot \Big(\sum_i a_i \otimes b_i \otimes c_i \Big) = \sum_i g_A a_i \otimes g_B b_i \otimes g_C c_i,$$

and $\mathrm{GL}(V)$ acts on $S^d V^*$ by $g \cdot P(x) = P(g^{-1} \cdot x)$. Letting $V = A \otimes B \otimes C$ and noting that $G \subset \mathrm{GL}(V)$, we have a G-action on $S^d(A \otimes B \otimes C)^*$. If $P \in I(\sigma_r)$, then $g \cdot P \in I(\sigma_r)$ for all $g \in G$. Since ideals are in particular vector spaces, the linear span of the orbit of P in $S^d(A \otimes B \otimes C)^*$ will be in $I(\sigma_r)$.

We will use the action of the group G to organize our calculations. A group G is said to *act* on a vector space V if there is a group homomorphism $\rho : G \to \mathrm{GL}(V)$. Then V is called a *G-module*. The G-module V is said to be *irreducible* if there is no nontrivial subspace of V invariant under the action of G. The irreducible polynomial $\mathrm{GL}(V)$-modules are indexed by partitions $\pi = (p_1, \ldots, p_{\mathbf{v}})$, where $p_1 \geq \cdots \geq p_{\mathbf{v}} \geq 0$. We write $|\pi| = p_1 + \cdots + p_{\mathbf{v}}$, and we say that π is a partition of $|\pi|$.

Let $S_\pi V$ denote the corresponding irreducible $\mathrm{GL}(V)$-module. It occurs in $V^{\otimes |\pi|}$ and no other tensor power, though not uniquely: there is a vector space's worth of realizations except in the cases $\pi = (d)$ and $\pi = (1, \ldots, 1)$. The irreducible $\mathrm{GL}(A) \times \mathrm{GL}(B) \times \mathrm{GL}(C)$-modules are all of the form $V_A \otimes V_B \otimes V_C$, where V_A is an irreducible $\mathrm{GL}(A)$-module, etc. For $G = \mathrm{GL}(A) \times \mathrm{GL}(B) \times \mathrm{GL}(C)$, every G-module decomposes into a direct sum of irreducible submodules. This decomposition is not unique in general, but the *isotypic* decomposition, in which all isomorphic modules are grouped together, is.

We are interested in the homogeneous polynomials of degree say d on $A \otimes B \otimes C$, denoted by $S^d(A \otimes B \otimes C)^*$. Via *polarization*, a polynomial may be considered a symmetric tensor, so

$$\begin{aligned} S^d(A \otimes B \otimes C)^* &\subset (A \otimes B \otimes C)^{*\otimes d} \\ &\simeq A^{*\otimes d} \otimes B^{*\otimes d} \otimes C^{*\otimes d}. \end{aligned}$$

Thus, the isomorphism types of irreducible G-modules in $S^d(A \otimes B \otimes C)^*$ are described by triples (π, μ, ν) of partitions of d whose number of parts satisfies $\ell(\pi) \leq \mathbf{a}$, etc.

Let $k_{\pi,\mu,\nu}$ denote the multiplicity of

$$S_\pi A^* \otimes S_\mu B^* \otimes S_\nu C^* \text{ in } S^d(A \otimes B \otimes C)^*,$$

that is, the dimension of the space of realizations of $S_\pi A^* \otimes S_\mu B^* \otimes S_\nu C^*$ in $S^d(A \otimes B \otimes C)^*$. The integers $k_{\pi,\mu,\nu}$ are called *Kronecker coefficients* and can be computed combinatorially. The programs Schur and Sage as well as several others will compute them for you in small cases. We used a program written by Harm Derksen, which is based on characters of the symmetric group.

There is a simple formula for $\dim S_\pi A^*$, namely

$$\dim S_{(p_1,\dots,p_{\mathbf{a}})}A^* = \prod_{1\le i<j\le \mathbf{a}} \frac{p_i - p_j + j - i}{j-i};$$

see, e.g., [Fulton and Harris 91, Theorem 6.3]. We will be interested in cases in which the dimension is small.

Let $\mathfrak{S}_d$ denote the group of permutations of d elements. If $\mathbf{a}=\mathbf{b}=\mathbf{c}$, then $\sigma_{r,\mathbf{a},\mathbf{a},\mathbf{a}}$ is also invariant under the $\mathfrak{S}_3$-action permuting the vector spaces. Thus whenever $S_{\pi_1}A^*\otimes S_{\pi_2}B^*\otimes S_{\pi_3}C^*$ is in the ideal of σ_r, the module $S_{\pi_{\sigma(1)}}A^*\otimes S_{\pi_{\sigma(2)}}B^*\otimes S_{\pi_{\sigma(3)}}C^*$ will be as well, for every $\sigma\in\mathfrak{S}_3$.

4.3. First Algorithm: To Obtain a Sample Collection of Polynomials

In Algorithm 1 we present an algorithm for computing a basis of highest-weight vectors for each isotypic component in $S^d(A\otimes B\otimes C)^*$. Once one has these, for each isotypic component, one can test whether there are modules in the ideal of $\sigma_{r;\mathbf{a},\mathbf{b},\mathbf{c}}$ (or any G-variety for that matter) by sampling random points on $\sigma_{r;\mathbf{a},\mathbf{b},\mathbf{c}}$ as described in the second algorithm.

For $\sigma\in\mathfrak{S}_d$, we write

$$\sigma(v_1\otimes\cdots\otimes v_d) := v_{\sigma(1)}\otimes\cdots\otimes v_{\sigma(d)}.$$

Once and for all fix bases $a^1,\dots,a^{\mathbf{a}}$ of A^*, and similarly for B,C. Let $\pi=(p_1,\dots,p_\ell,0,\dots,0)$ be a partition as above. Write $\pi=(p_1,\dots,p_\ell)$ and $\ell(\pi)=\ell$. Define $F_{A,\pi}\in A^{*\otimes d}$ by

$$F_{A,\pi} := (a^1)^{\otimes(p_1-p_2)}\otimes(a^1\wedge a^2)^{\otimes(p_2-p_3)}\otimes\cdots \otimes(a^1\wedge\cdots\wedge a^f)^{\otimes(p_f-p_{f-1})}.$$

Here

$$v_1\wedge\cdots\wedge v_k := \frac{1}{k!}\sum_{\sigma\in\mathfrak{S}_k}\operatorname{sgn}(\sigma)\sigma(v_1\otimes\cdots\otimes v_k).$$

Example 4.1. $d=2$, $(\pi,\mu,\nu)=((2),(1,1),(1,1))$. Here $k_{(2),(1,1),(1,1)}=1$, so we are looking for a single polynomial. We have $F_{A,(2)}=(a^1)^2$, $F_{B,(1,1)}=b^1\wedge b^2$, and $F_{C,(1,1)}=c^1\wedge c^2$. Try $\tau_1=\tau_2=\mathrm{Id}$. Then

$$\begin{aligned}
&F^{\mathrm{Id},\mathrm{Id}}_{(2),(1,1),(1,1)}\\
&\quad= (a^1\otimes a^1)\otimes(b^1\otimes b^2 - b^2\otimes b^1)\otimes(c^1\otimes c^2 - c^2\otimes c^1)\\
&\quad= (a^1\otimes b^1\otimes c^1)\otimes(a^1\otimes b^2\otimes c^2)\\
&\qquad - (a^1\otimes b^1\otimes c^2)\otimes(a^1\otimes b^2\otimes c^1)\\
&\qquad - (a^1\otimes b^2\otimes c^1)\otimes(a^1\otimes b^1\otimes c^2)\\
&\qquad + (a^1\otimes b^2\otimes c^2)\otimes(a^1\otimes b^1\otimes c^1).
\end{aligned}$$

Thus

$$P^{\mathrm{Id},\mathrm{Id}}_{(2),(1,1),(1,1)}(x^{ijk}a_i\otimes b_j\otimes c_k) = 2x^{111}x^{122} - 2x^{112}x^{121}.$$

Here and throughout, repeated indices are to be summed over. Note that if $T=x^{ijk}a_i\otimes b_j\otimes c_k$ has rank one, then $P^{\mathrm{Id},\mathrm{Id}}_{(2),(1,1),(1,1)}(T)=0$, but P will evaluate to be nonzero on a general rank-two tensor.

Example 4.2. $d=3$, $(\pi,\mu,\nu)=((2,1),(2,1),(2,1))$. Here $k_{(2,1),(2,1),(2,1)}=1$, so again we are looking for a single polynomial. We have $F_{A,(2,1)}=a^1\otimes(a^1\wedge a^2)$, and similarly for B,C. Try $\tau_1=\tau_2=\mathrm{Id}$. Then

$$\begin{aligned}
&F^{\mathrm{Id},\mathrm{Id}}_{(21),(21),(21)}\\
&\quad= (a^1\otimes a^1\otimes \alpha^2 - a^1\otimes a^2\otimes a^1)\\
&\qquad \otimes(b^1\otimes b^1\otimes b^2 - b^1\otimes b^2\otimes b^1)\\
&\qquad \otimes(c^1\otimes c^1\otimes c^2 - c^1\otimes c^2\otimes c^1)\\
&\quad= (a^1\otimes b^1\otimes c^1)\otimes(a^1\otimes b^1\otimes c^1)\otimes(a^2\otimes b^2\otimes c^2)\\
&\qquad - (a^1\otimes b^1\otimes c^1)\otimes(a^1\otimes b^1\otimes c^2)\otimes(a^2\otimes b^2\otimes c^1)\\
&\qquad - (a^1\otimes b^1\otimes c^1)\otimes(a^1\otimes b^2\otimes c^1)\otimes(a^2\otimes b^1\otimes c^2)\\
&\qquad + (a^1\otimes b^1\otimes c^1)\otimes(a^1\otimes b^2\otimes c^2)\otimes(a^2\otimes b^1\otimes c^1)\\
&\qquad - (a^1\otimes b^1\otimes c^1)\otimes(a^2\otimes b^1\otimes c^1)\otimes(a^1\otimes b^2\otimes c^2)\\
&\qquad + (a^1\otimes b^1\otimes c^1)\otimes(a^2\otimes b^1\otimes c^2)\otimes(a^1\otimes b^2\otimes c^1)\\
&\qquad + (a^1\otimes b^1\otimes c^1)\otimes(a^2\otimes b^2\otimes c^1)\otimes(a^1\otimes b^1\otimes c^2)\\
&\qquad - (a^1\otimes b^1\otimes c^1)\otimes(a^2\otimes b^2\otimes c^2)\otimes(a^1\otimes b^1\otimes c^1)
\end{aligned}$$

Thus $P^{\mathrm{Id},\mathrm{Id}}_{(21),(21),(21)}\equiv 0$, so we need to try different τ_1,τ_2. Take $\tau_1=\mathrm{Id}$ and $\tau_2=(12)$. Then

$$\begin{aligned}
&F^{\mathrm{Id},(1,2)}_{(21),(21),(21)}\\
&\quad= (a^1\otimes a^1\otimes a^2 - a^1\otimes a^2\otimes a^1)\otimes(b^1\otimes b^1\otimes b^2\\
&\qquad - b^1\otimes b^2\otimes b^1)\otimes(c^1\otimes c^1\otimes c^2 - c^2\otimes c^1\otimes c^1)\\
&\quad= (a^1\otimes b^1\otimes c^1)\otimes(a^1\otimes b^1\otimes c^1)\otimes(a^2\otimes b^2\otimes c^2)\\
&\qquad - (a^1\otimes b^1\otimes c^2)\otimes(a^1\otimes b^1\otimes c^1)\otimes(a^2\otimes b^2\otimes c^1)\\
&\qquad - (a^1\otimes b^1\otimes c^1)\otimes(a^1\otimes b^2\otimes c^1)\otimes(a^2\otimes b^1\otimes c^2)\\
&\qquad + (a^1\otimes b^1\otimes c^2)\otimes(a^1\otimes b^2\otimes c^1)\otimes(a^2\otimes b^1\otimes c^1)\\
&\qquad - (a^1\otimes b^1\otimes c^1)\otimes(a^2\otimes b^1\otimes c^1)\otimes(a^1\otimes b^2\otimes c^2)\\
&\qquad + (a^1\otimes b^1\otimes c^2)\otimes(a^2\otimes b^1\otimes c^1)\otimes(a^1\otimes b^2\otimes c^1)\\
&\qquad + (a^1\otimes b^1\otimes c^1)\otimes(a^2\otimes b^2\otimes c^1)\otimes(a^1\otimes b^1\otimes c^2)\\
&\qquad - (a^1\otimes b^1\otimes c^2)\otimes(a^2\otimes b^2\otimes c^1)\otimes(a^1\otimes b^1\otimes c^1).
\end{aligned}$$

Thus

$$\begin{aligned}
&P^{\mathrm{Id},(1,2)}_{(21),(21),(21)}\left(\sum_{i,j,k=1}^{2} x^{ijk}a_i\otimes b_j\otimes c_k\right)\\
&\quad= x^{111}x^{111}x^{222} + 2x^{112}x^{121}x^{211}\\
&\qquad - \left(x^{111}x^{121}x^{212} + x^{111}x^{211}x^{122} + x^{111}x^{112}x^{221}\right).
\end{aligned}$$

Algorithm 1 ConstructBasis.

Input: Degree d and partitions π, μ, ν of d.

Output: A basis P of the highest-weight space vector of the isotypic component of $S_\pi A^* \otimes S_\mu B^* \otimes S_\nu C^*$ in $S^d(A \otimes B \otimes C)^*$.

Use your favorite method to compute $k_{\pi,\mu,\nu}$.
Set $k = 0$.
while $k < k_{\pi,\mu,\nu}$ **do**
  repeat
    Choose permutations $\tau_1, \tau_2 \in \mathfrak{S}_d$.
    Define

$$F^{\tau_1,\tau_2}_{\pi,\mu,\nu} := F_{A,\pi} \otimes (\tau_1 \cdot F_{B,\mu}) \otimes (\tau_2 \cdot F_{C,\nu}) \in A^{*\otimes d} \otimes B^{*\otimes d} \otimes C^{*\otimes d},$$

    and rearrange the factors so it is expressed as an element of $(A \otimes B \otimes C)^{*\otimes d}$, and symmetrize to get

$$P^{\tau_1,\tau_2}_{\pi,\mu,\nu} := \sum_{\sigma\in\mathfrak{S}_d} \sigma \cdot F^{\tau_1,\tau_2}_{\pi,\mu,\nu} = \sum_{\sigma\in\mathfrak{S}_d} (\sigma \cdot F_{A,\pi}) \otimes (\sigma \cdot \tau_1 \cdot F_{B,\mu}) \otimes (\sigma \cdot \tau_2 \cdot F_{C,\nu}) \in A^{*\otimes d} \otimes B^{*\otimes d} \otimes C^{*\otimes d},$$

    where we recall that $\sigma \cdot (a^1 \otimes \cdots \otimes a^d) := a^{\sigma(1)} \otimes \cdots \otimes a^{\sigma(d)}$.
  until $P^{\tau_1,\tau_2}_{\pi,\mu,\nu}$ is linearly independent of $P_1, \ldots, P_{k-1}$.
  Increase $k = k + 1$.
  Set $P_k = P^{\tau_1,\tau_2}_{\pi,\mu,\nu}$.
end while

Note that if T has rank one, then

$$P^{\mathrm{Id},(1,2)}_{(21),(21),(21)}(T) = 0,$$

but P will evaluate to be nonzero on a general rank-two tensor.

4.3.1. Permutation Pairs to Avoid

We want to avoid the case that occurs in the first try of Example 4.1, i.e., that $P^{\tau_1,\tau_2}_{\pi,\mu,\nu} = 0$. Although a complete classification of the cases in which this happens is unknown, an easy necessary condition for $P^{\tau_1,\tau_2}_{\pi,\mu,\nu} \neq 0$ is the following [Ikenmeyer 12, Lemma 7.2.7]: When we write $1, 2, \ldots, d$ in a columnwise tableau starting with the longest column and we write $\tau_1(1), \tau_1(2), \ldots, \tau_1(d)$ in a second columnwise tableau, and we do the same for τ_2 in a third tableau, then it is necessary that there exist no pair of numbers that lies in the same column in all three tableaux. If this occurs, we call this situation a *zero pattern*. We can choose random permutations that avoid the zero pattern by just choosing random permutations and choosing again if the one chosen contains a zero pattern.

4.3.2. Implementation

The algorithm is not complicated to implement, but the following details are paramount for its running time.

What is crucial in our implementation is that we avoid writing down the P_i as polynomials. A polynomial P_i is stored only as its permutation pair $(\tau_1, \tau_2) \in \mathfrak{S}_d \times \mathfrak{S}_d$. To prove linear independence among polynomials P_i, the $P_i \in S^d(A \otimes B \otimes C)^*$ are contracted with random tensors $t_j = w_j^{\otimes d}$ with $w_j \in A \otimes B \otimes C$ having low rank, which is the same as evaluating the function P_i on w_j. If the resulting matrix $(\langle P_i, t_j \rangle)_{i,j}$ consisting of the contractions $\langle P_i, t_j \rangle$ has full rank, then the P_i are linearly independent.

If $a \in (A \otimes B \otimes C)^{\otimes d}$ is of rank 1, then the contraction $\langle P_i, a \rangle$ is a product of $\ell \times \ell$ determinants, which can be efficiently computed. Hence to compute a contraction $\langle P_i, t_j \rangle$, it would suffice to expand t_j into rank-1 tensors and sum over the products of determinants. But to make this method computationally feasible, we do not expand t_j completely, since $\langle P_i, t_j \rangle$ would consist of a huge quantity of zero summands. We use a standard divide and conquer method to expand t_j partially and prune the computation whenever at least one determinant is seen to be zero.

To avoid numerical errors, it suffices for proving linear independence to work over a finite field or ring. The same method can be used to evaluate at the matrix multiplication tensor M_2.

4.4. Second Algorithm: To Test on the Secant Variety

Once one has a basis of highest-weight vectors for an isotypic component, one needs to determine which linear

Algorithm 2 EvaluateBasis.

Input: The output Algorithm 1 for some (π,μ,ν), i.e., a collection $P_1,\dots,P_k = P_{k_{\pi,\mu,\nu}} \in S^d(A\otimes B\otimes C)^*$ and r, where we will test for polynomials in $I(\sigma_{r;\mathbf{a},\mathbf{b},\mathbf{c}})$.

Output: with probability as high as you like, the component of $I(\sigma_{r;\mathbf{a},\mathbf{b},\mathbf{c}})$ in $S_\pi A^*\otimes S_\mu B^*\otimes S_\nu C^*$. If the component is zero, then the answer is guaranteed correct, and more generally, the algorithm can overestimate the component only if the points on σ_r are not chosen randomly enough.

Set $P = c_1P_1+\cdots+c_kP_k$, where $c_1,\dots,c_k$ are variables.

Chose "random" vectors

$$v_j = \sum_{i=1}^{\mathbf{a}}\sum_{k=1}^{\mathbf{b}}\sum_{l=1}^{\mathbf{c}}(\alpha^i_{1,j}a_i)\otimes(\beta^k_{1,j}b_k)\otimes(\gamma^l_{1,j}c_l)+\cdots+(\alpha^i_{r,j}a_i)\otimes(\beta^k_{r,j}b_k)\otimes(\gamma^l_{r,j}c_l),$$

where the $\alpha^i_{\delta,j},\beta^k_{\delta,j},\gamma^l_{\delta,j}$ are "random" numbers.

Evaluate P at these k points.

if there exists a solution $c_1,\dots,c_k$ such that all the evaluations are zero **then**

  If there is an m-dimensional solution space, then with reasonable probability one has m copies of the module in the ideal.

else

  No module in this isotypic component is in $I(\sigma_{r;\mathbf{a},\mathbf{b},\mathbf{c}})$.

end if

combinations of basis vectors vanish on σ_r. The is standard linear algebra and is accomplished in Algorithm 2.

For the implementation, it is again crucial to store the P_i only as permutation pairs. Evaluation at points works as described in Section 4.3.2. Unlike linear independence, we need stronger methods to prove linear dependence. One can parameterize σ_6 and use the fact that the relations between all determinants that appear during the calculation are given by Young tableau relations; cf. [Fulton 97, p. 110]. No particular optimization was done during this step, which renders it the slowest part of our algorithm.

4.5. Our Run

Let $d=19$, $(\pi,\mu,\nu)=((5554),(5554),(5554))$. Here $k_{(5554),(5554),(5554)}=31$. We found 31 pairs τ_1,τ_2 that result in 31 linearly independent polynomials by choosing τ_1 and τ_2 randomly, but avoiding the zero pattern. As expected, the linear combination has support 31 and no evident structure other than that it "magically" vanishes on $\sigma_{6;4,4,4}$. A somewhat nicer description (smaller support) of a polynomial vanishing on $\sigma_{6;4,4,4}$ is obtained in the following $d=20$ case.

Let $d=20$, $(\pi,\mu,\nu)=((5555),(5555),(5555))$. Here $k_{(5555),(5555),(5555)}=4$. The following random choices of pairs τ_1,τ_2 give four linearly independent polynomials:

$$\begin{aligned}\tau_1 &= (\tau_1(1),\tau_1(2),\dots,\tau_1(20))\\ &= (10,15,5,9,13,4,17,14,7,20,19,11,2,12,8,3,16,18,6,1),\\ \tau_2 &= (10,11,6,2,8,9,4,20,15,16,13,18,14,19,7,5,17,3,12,1)\end{aligned}$$

$$\begin{aligned}\tau_1 &= (19,10,1,5,7,12,2,13,16,6,18,9,11,20,3,17,14,8,15,4),\\ \tau_2 &= (10,5,13,6,3,16,11,1,4,18,15,17,9,2,8,12,19,7,14,20)\end{aligned}$$

$$\begin{aligned}\tau_1 &= (16,20,9,13,8,1,4,19,11,17,7,2,14,3,6,5,12,15,18,10),\\ \tau_2 &= (1,20,11,19,5,16,17,2,18,13,7,12,14,10,8,15,6,9,3,4)\end{aligned}$$

$$\begin{aligned}\tau_1 &= (11,5,2,1,16,10,20,3,17,19,12,18,13,9,14,4,8,6,15,7),\\ \tau_2 &= (1,6,15,13,20,3,18,11,14,2,9,5,4,17,12,8,19,16,7,10).\end{aligned}$$

This gives rise to four polynomials $f_1,\dots,f_4$. If we restrict these functions to σ_6, the second algorithm yields the following linear combination, which vanishes on σ_6: $-266054\,f_1+421593\,f_2+755438\,f_3+374660\,f_4$. The coefficients look random, as is expected, since the permutation pairs were chosen at random. The computation took several hours on 16 processors, the symbolic proof of vanishing at σ_6 being by far the slowest part.

5. REVIEW OF THE ORIGINAL PROOF THAT $M_2 \notin \sigma_{6;4,4,4}$

The essence of the proof that the border rank of M_2 is not 6 in [Landsberg 06] is as follows: There is a now standard argument due to Baur for proving lower bounds for rank by splitting a putative computation into two parts using the algebra structure on the space of matrices. The argument in [Landsberg 06] was to apply the same type of argument to each component of the variety consisting of subvarieties whose ranks are greater than the border rank. The article [Landsberg 06] contained a gap in the proof that was filled in [Landsberg 04] but not published in the *Journal of the American Mathematical Society* because the editor was concerned that the erratum was almost as long as the original article, and the author did not see a way to shorten it.

The gap in [Landsberg 06] was caused by overlooking the possibility of certain types of components, where the limiting 6-planes are not formed by points coming together but by some other unusual configuration of points. Not all such components of $\sigma_{6;4,4,4}$ are known explicitly, but the correction used only qualitative aspects of how the limiting 6-plane arose.

There were three basic cases: that every set of five of the limit points is linearly independent; that there is a subset of five that is not, but every subset of four is; and that there is a subset of four that is not, but every subset of three is. In each of these cases, one is forced to have a limit taking place among rank-one tensors in a much smaller space, which is what made the analysis tractable. The computations performed above provide an explicit polynomial vanishing on $\sigma_{6;4,4,4}$ that does not vanish at M_2, providing a significantly shorter proof of this fact.

ACKNOWLEDGMENTS

We thank Peter Bürgisser for important discussions and suggestions, and the anonymous reviewer for many helpful comments.

Hauenstein's research was supported in part by AFOSR grant FA8650-13-1-7317 and NSF grant DMS-1262428. Ikenmeyer's research was supported in part by DFG grant BU 1371/3-2. Landsberg's research was supported in part by NSF grant DMS-1006353.

REFERENCES

[Bates and Oeding 11] D. Bates and L. Oeding. "Toward a Salmon Conjecture." *Exp. Math.* 20:3 (2011), 358–370.

[Bates et al. 06] Daniel J. Bates, Jonathan D. Hauenstein, Andrew J. Sommese, and Charles W. Wampler. "Bertini: Software for Numerical Algebraic Geometry." 2006.

[Blekherman et al. 12] G. Blekherman, J. Hauenstein, J. C. Ottem, K. Ranestad, and B. Sturmfels. "Algebraic Boundaries of Hilbert's SOS Cones." *Compositio Mathematica* 148 (2012), 1717–1735.

[Bürgisser and Ikenmeyer 11] Peter Bürgisser and Christian Ikenmeyer. "Geometric Complexity Theory and Tensor Rank." In *Proceedings 43rd Annual ACM Symposium on Theory of Computing 2011*, pp. 509–518. 2011.

[Bürgisser et al. 97] Peter Bürgisser, Michael Clausen, and M. Amin Shokrollahi. *Algebraic Complexity Theory*, Grundlehren der Mathematischen Wissenschaften 315. Springer-Verlag, 1997.

[Bürgisser et al. 13] Peter Bürgisser, Michael Clausen, and M. Amin Shokrollahi. "Explicit Lower Bounds via Geometric Complexity Theory." In *Proceedings 45th Annual ACM Symposium on Theory of Computing 2013*, pp. 141–150. 2013.

[Fulton 97] William Fulton. *Young Tableaux*, London Mathematical Society Student Texts 35. Cambridge University Press, 1997.

[Fulton and Harris 91] William Fulton and Joe Harris. *Representation Theory*, Graduate Texts in Mathematics 129. Springer-Verlag, 1991.

[Griffin et al. 14] Z. A. Griffin, J. D. Hauenstein, C. Peterson, and A. J. Sommese. "Numerical Computation of the Hilbert Function of a Zero-Scheme." Available online (www.math.ncsu.edu/~jdhauens/preprints). To appear in *Springer Proc. Math. Stat.*, 2014.

[Hauenstein and Sommese 10] J. D. Hauenstein and A. J. Sommese. "Witness Sets of Projections." *Appl. Math. Comput.* 217 (2010), 3349–3354.

[Hauenstein and Sommese 13] J. D. Hauenstein and A. J. Sommese. "Membership Tests for Images of Algebraic Sets by Linear Projections." *Appl. Math. Comput.* (2013), 6809–6818.

[Hauenstein et al. 12] J. D. Hauenstein, Y.-H. He, and D. Mehta. "Numerical Analyses on Moduli Space of Vacua." arXiv:1210.6038, 2012.

[Ikenmeyer 12] Christian Ikenmeyer. "Geometric Complexity Theory, Tensor Rank, and Littlewood–Richardson Coefficients." PhD thesis, Institute of Mathematics, University of Paderborn, 2012.

[Landsberg 04] J. M. Landsberg. "The Border Rank of the Multiplication of 2×2 Matrices Is Seven." arXiv:math/0407224, 2004.

[Landsberg 06] J. M. Landsberg. "The Border Rank of the Multiplication of 2×2 Matrices Is Seven." *J. Amer. Math. Soc.* 19 (2006), 447–459.

[Landsberg 12] J. M. Landsberg. "Explicit Tensors of Border Rank at Least $2n-1$." arXiv:1209.1664, 2012.

[Landsberg 12] J. M. Landsberg. *Tensors: Geometry and Applications*, Graduate Studies in Mathematics 128. American Mathematical Society, 2012.

[Landsberg and Manivel 04] J. M. Landsberg and Laurent Manivel. "On the Ideals of Secant Varieties of Segre Varieties." *Found. Comput. Math.* 4 (2004), 397–422.

[Landsberg and Ottaviani 11] J. M. Landsberg and Giorgio Ottaviani. *New Lower Bounds for the Border Rank of Matrix Multiplication.* arXiv:1112.6007, 2011.

[Migliore 98] J. C. Migliore. *Introduction to Liaison Theory and Deficiency Modules*, Progress in Mathematics 165. Birkhäuser, 1998.

[Schönhage 81] A. Schönhage. "Partial and Total Matrix Multiplication." *SIAM J. Comput.* 10 (1981), 434–455.

[Sommese and Wampler 05] A. J. Sommese and C. W. Wampler, II. *The Numerical Solution of Systems of Polynomials.* World Scientific, 2005.

[Sommese et al. 01] A. J. Sommese, J. Verschelde, and C. W. Wampler. "Using Monodromy to Decompose Solution Sets of Polynomial Systems into Irreducible Components." In *Applications of Algebraic Geometry to Coding Theory, Physics and Computation (Eilat, 2001)*, NATO Sci. Ser. II Math. Phys. Chem. 36, pp. 297–315. Kluwer Acad. Publ., 2001.

[Sommese et al. 02] A. J. Sommese, J. Verschelde, and C. W. Wampler. "Symmetric Functions Applied to Decomposing Solution Sets of Polynomial Systems." *SIAM J. Numer. Anal.* 40 (2002), 2026–2046.

[Strassen 69] Volker Strassen. *Gaussian Elimination Is Not Optimal. Numer. Math.* 13 (1969), 354–356.

Jonathan D. Hauenstein, Department of Mathematics, North Carolina State University, Box 8205, Raleigh, NC 27695-8205, USA (hauenstein@ncsu.edu)

Christian Ikenmeyer, Department of Mathematics, Texas A&M University, Mailstop 3368, College Station, TX 77843-3368, USA (ciken@math.tamu.edu)

J. M. Landsberg, Department of Mathematics, Texas A&M University, Mailstop 3368, College Station, TX 77843-3368, USA (jml@math.tamu.edu)