Forum of Mathematics, Pi (2023), **Vol. 11**:e17 1–30
doi:10.1017/fmp.2023.14

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

New lower bounds for matrix multiplication and $\det_3$

Austin Conner[1], Alicia Harper[2] and J.M. Landsberg[3]

[1]Department of Mathematics, Harvard University, 1 Oxford St, Cambridge, MA, 01238; E-mail: aconner@math.harvard.edu.
[2]Department of Mathematics, Texas A&M University, College Station, TX 77843-3368; E-mail: adharper@math.tamu.edu.
[3]Department of Mathematics, Texas A&M University, College Station, TX 77843-3368; E-mail: jml@math.tamu.edu.

Received: 12 March 2022; **Revised:** 30 November 2022; **Accepted:** 27 March 2023

2020 Mathematics subject classification: *Primary* – 68Q17, 15A69; *Secondary* – 14L35

Abstract

Let $M_{\langle \mathbf{u},\mathbf{v},\mathbf{w}\rangle} \in \mathbb{C}^{\mathbf{uv}}\otimes\mathbb{C}^{\mathbf{vw}}\otimes\mathbb{C}^{\mathbf{wu}}$ denote the matrix multiplication tensor (and write $M_{\langle \mathbf{n}\rangle} = M_{\langle \mathbf{n},\mathbf{n},\mathbf{n}\rangle}$), and let $\det_3 \in (\mathbb{C}^9)^{\otimes 3}$ denote the determinant polynomial considered as a tensor. For a tensor T, let $\underline{\mathbf{R}}(T)$ denote its border rank. We (i) give the first hand-checkable algebraic proof that $\underline{\mathbf{R}}(M_{\langle 2\rangle}) = 7$, (ii) prove $\underline{\mathbf{R}}(M_{\langle 223\rangle}) = 10$ and $\underline{\mathbf{R}}(M_{\langle 233\rangle}) = 14$, where previously the only nontrivial matrix multiplication tensor whose border rank had been determined was $M_{\langle 2\rangle}$, (iii) prove $\underline{\mathbf{R}}(M_{\langle 3\rangle}) \geq 17$, (iv) prove $\underline{\mathbf{R}}(\det_3) = 17$, improving the previous lower bound of 12, (v) prove $\underline{\mathbf{R}}(M_{\langle 2\mathbf{nn}\rangle}) \geq \mathbf{n}^2 + 1.32\mathbf{n}$ for all $\mathbf{n} \geq 25$, where previously only $\underline{\mathbf{R}}(M_{\langle 2\mathbf{nn}\rangle}) \geq \mathbf{n}^2 + 1$ was known, as well as lower bounds for $4 \leq \mathbf{n} \leq 25$, and (vi) prove $\underline{\mathbf{R}}(M_{\langle 3\mathbf{nn}\rangle}) \geq \mathbf{n}^2 + 1.6\mathbf{n}$ for all $\mathbf{n} \geq 18$, where previously only $\underline{\mathbf{R}}(M_{\langle 3\mathbf{nn}\rangle}) \geq \mathbf{n}^2 + 2$ was known. The last two results are significant for two reasons: (i) they are essentially the first nontrivial lower bounds for tensors in an "unbalanced" ambient space and (ii) they demonstrate that the methods we use (border apolarity) may be applied to sequences of tensors.

The methods used to obtain the results are new and "nonnatural" in the sense of Razborov and Rudich, in that the results are obtained via an algorithm that cannot be effectively applied to generic tensors. We utilize a new technique, called *border apolarity* developed by Buczyńska and Buczyński in the general context of toric varieties. We apply this technique to develop an algorithm that, given a tensor T and an integer r, in a finite number of steps, either outputs that there is no border rank r decomposition for T or produces a list of all normalized ideals which could potentially result from a border rank decomposition. The algorithm is effectively implementable when T has a large symmetry group, in which case it outputs potential decompositions in a natural normal form. The algorithm is based on algebraic geometry and representation theory.

Contents

1 Introduction **2**
1.1 Methods/History . . . 2
1.2 Results . . . 4
1.3 What comes next? . . . 5
1.3.1 Breaking the lower bound barriers . . . 5
1.3.2 Upper bounds, especially for tensors relevant for Strassen's laser method . . . 6
1.3.3 Geometrization of the (111) test for matrix multiplication . . . 6
1.4 Overview . . . 6
2 Preliminaries **6**
2.1 Definitions/Notation . . . 6
2.2 Border rank decompositions as curves in Grassmannians . . . 7

2.3 Multigraded ideal associated to a border rank decomposition 8
2.4 Lie's theorem and consequences 9
2.5 Borel fixed subspaces 9
2.5.1 Weights and weight diagrams 9
2.5.2 Parameterizing Borel fixed subspaces 10
2.5.3 $\mathfrak{gl}_m$ and $\mathfrak{sl}_m$ weights 10
3 The ideal enumeration algorithm **13**
4 Matrix multiplication **15**
5 Explanation of the proofs of Theorems 1.1 and 1.2 **16**
6 Representation theory relevant for matrix multiplication **18**
6.1 Refinement of the (210) test for matrix multiplication 18
6.2 $\underline{\mathbf{R}}(M_{\langle 2\rangle}) > 6$ revisited 20
7 Proofs of Theorems 1.5 and 1.6 **20**
7.1 Overview 20
7.2 Preliminaries 20
7.3 The local argument 22
7.4 The globalization 25
7.5 Proof of Lemma 7.7 26

1. Introduction

Over 50 years ago Strassen [40] discovered that the usual row-column method for multiplying $\mathbf{n}\times\mathbf{n}$ matrices, which uses $O(\mathbf{n}^3)$ arithmetic operations, is not optimal by exhibiting an explicit algorithm to multiply matrices using $O(\mathbf{n}^{2.81})$ arithmetic operations. Ever since then, substantial efforts have been made to determine just how efficiently matrices may be multiplied. See any of [12, 8, 31] for an overview. Matrix multiplication of $\mathbf{n}\times\boldsymbol{\ell}$ matrices with $\boldsymbol{\ell}\times\mathbf{m}$ matrices is a bilinear map, that is, a tensor $M_{\langle\boldsymbol{\ell},\mathbf{m},\mathbf{n}\rangle}\in\mathbb{C}^{\boldsymbol{\ell}\mathbf{m}}\otimes\mathbb{C}^{\mathbf{m}\mathbf{n}}\otimes\mathbb{C}^{\mathbf{n}\boldsymbol{\ell}}$, and since 1980 [6], the primary complexity measure of the matrix multiplication tensor has been its *border rank*, which is defined as follows.

A nonzero tensor $T\in\mathbb{C}^{\mathbf{a}}\otimes\mathbb{C}^{\mathbf{b}}\otimes\mathbb{C}^{\mathbf{c}}=:A\otimes B\otimes C$ has *rank one* if $T=a\otimes b\otimes c$ for some $a\in A$, $b\in B$, $c\in C$ and the *rank* of T, denoted $\mathbf{R}(T)$, is the smallest r such that T may be written as a sum of r rank one tensors. The *border rank* of T, denoted $\underline{\mathbf{R}}(T)$, is the smallest r such that T may be written as a limit of a sum of r rank one tensors. In geometric language, the border rank is smallest r such that $[T]\in\sigma_r(\mathrm{Seg}(\mathbb{P}A\times\mathbb{P}B\times\mathbb{P}C))$. Here, $\sigma_r(\mathrm{Seg}(\mathbb{P}A\times\mathbb{P}B\times\mathbb{P}C))$ denotes the r-th secant variety of the Segre variety of rank one tensors. For the relations between rank, border rank and other measures of complexity, see [12, Ch. 14-15].

Despite the vast literature on matrix multiplication, previous to this paper, the precise border rank of $M_{\langle\boldsymbol{\ell},\mathbf{m},\mathbf{n}\rangle}$ was known in exactly one nontrivial case, namely $M_{\langle 2\rangle}=M_{\langle 222\rangle}$ [29]. We determine the border rank in two new cases, $M_{\langle 223\rangle}$ and $M_{\langle 233\rangle}$. We prove new border rank lower bounds for $M_{\langle 3\rangle}$ and two infinite sequences of new cases, $M_{\langle 2\mathbf{nn}\rangle}$ and $M_{\langle 3\mathbf{nn}\rangle}$. Previous to this paper, there were no nontrivial lower bounds for these sequences. In fact, there were no nontrivial border rank lower bounds for *any tensor* in $\mathbb{C}^{\mathbf{a}}\otimes\mathbb{C}^{\mathbf{a}}\otimes\mathbb{C}^{\mathbf{b}}$, where $\mathbf{b}>2\mathbf{a}$ other than Lickteig's near trivial bound [37] $\underline{\mathbf{R}}(M_{\langle\mathbf{m},\mathbf{n},\mathbf{n}\rangle})\geq\mathbf{n}^2+1$ when $\mathbf{m}<\mathbf{n}$, (where the bound of $\mathbf{n}^2$ is trivial). We also determine the border rank of the 3×3 determinant considered as a tensor, which is important for proving *upper bounds* on the exponent of matrix multiplication as discussed below. See §1.2 below for precise statements.

1.1. Methods/History

This paper deals exclusively with lower bounds ("complexity theory's Waterloo" according to [5, Chap. 14]). For a history of upper bounds, see, for example, [8, 31].

Let $\sigma_r(\mathrm{Seg}(\mathbb{P}A\times\mathbb{P}B\times\mathbb{P}C))$ denote the set of tensors of border rank at most r, which is called the *r-th secant variety of the Segre variety*. Previously, border rank lower bounds for tensors were

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

primarily obtained by finding a polynomial vanishing on $\sigma_r(\mathrm{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C))$ and then showing the polynomial is nonzero when evaluated on the tensor in question. These polynomials were found by reducing multilinear algebra to linear algebra [41], and also exploiting the large symmetry group of $\sigma_r(\mathrm{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C))$ to help find the polynomials [35, 36]. Such methods are subject to *barriers* [18, 21]; see [32, §2.2] for an overview. A technique allowing one to go slightly beyond the barriers was introduced in [34]. The novelty there was, in addition to exploiting the symmetry group of $\sigma_r(\mathrm{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C))$, to also exploit the symmetry group of the tensor one wanted to prove lower bounds on. This *border substitution method* of [34] relied on first using the symmetry of the tensor to study its degenerations via the *Normal Form Lemma* 2.3, and then to use polynomials on the degeneration of the tensor.

The classical *apolarity method* studies the decompositions of a homogeneous polynomial of degree d into a sum of d-th powers of linear forms, (these are called *Waring rank decompositions*); see,for example, [27]. It was generalized to study ranks of points with respect to toric varieties [22, 23]. To prove rank lower bounds with it, one takes the ideal of linear differential operators annihilating a given polynomial P and proves it does not contain an ideal annihilating r distinct points. In [11], Buczyńska and Buczyński extend this classical method to the border rank setting. They also extend the Normal Form Lemma to the entire ideal associated to the border rank decomposition of the tensor, their *Fixed Ideal Theorem* (Theorem 2.4). (In the language introduced below, the Normal Form Lemma is the (111) case of the Fixed Ideal Theorem.) In the present work, we describe an algorithm to enumerate a set of parameterized families of ideals which together exhaust those which could satisfy the conclusion of the Fixed Ideal Theorem, and we show this enumeration fails to produce any candidates in important cases of interest.

The ideals subject to enumeration are homogeneous in three sets of variables, so we have a $\mathbb{Z}^3$-graded ring of polynomials, that is, $I = \bigoplus_{i,j,k} I_{ijk}$, and we may study a putative ideal I in each multidegree. Given r, the ideal enumeration algorithm builds a candidate ideal family step by step, starting in low (multi) degree and building upwards. At each building step, there are tests that restrict a so-far built family to a subfamily, and after these tests empty families are removed. If at any point there are no remaining candidates, one concludes there is no border rank r decomposition. For tensors with large symmetry groups, the dimensions of candidate ideal families one needs to consider during this enumeration are typically small. All the results of this paper require examining only the first few multigraded components of candidate ideal families.

The restrictions to subfamilies result from upper bounding the ranks of certain linear maps. The linear maps are multiplication maps. On one hand, in order for a candidate space of polynomials to be an ideal, it must be closed under multiplication. On the other hand, our hypothesis that the ideal arises via a border rank r decomposition upper-bounds its dimension in each nontrivial multidegree; in fact one may assume it has codimension r in each multidegree.

The fact that the elimination conditions are rank conditions implies that the lower bound barriers [18, 21] still hold for the technique as presented in this paper. In §1.3, we explain how we plan to augment these tests to go beyond the barriers in future work and how our techniques might be used to overcome *upper bound barriers* for the laser method as well.

We use *representation theory* at several levels. For tensors with symmetry, the Fixed Ideal Theorem significantly restricts the candidate I_{ijk}'s one must consider, namely to those that are fixed by a *Borel subgroup* of the *symmetry group of the tensor*. Without this additional condition, even low degree ideal enumeration would likely be impossible to carry out except for very small examples.

We also make standard use of representation theory to put the matrices whose ranks we need to lower-bound in block diagonal format via Schur's lemma. For example, to prove $\underline{\mathbf{R}}(M_{\langle 2\rangle}) > 6$, the border apolarity method produces three size 40×40 matrices whose ranks need to be lower bounded. Decomposing the matrices to maps between isotypic components reduces the calculation to computing the ranks of several matrices of size 4×8 with entries $0, \pm 1$, making the proof easily hand-checkable.

Our results for $M_{\langle 3\rangle}$ and $\det_3$ are obtained by a computer implementation of the ideal enumeration algorithm.

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

For $M_{\langle \mathbf{2n} \rangle}$ and $M_{\langle \mathbf{3n} \rangle}$, we must handle all $\mathbf{n}$ uniformly, and a computer calculation is no longer possible. To do this, we consider potential I_{110} candidates as a certain sum of 'local' contributions, which we analyze separately (Lemmas 7.2 and 7.4). Given this analysis, it is possible to give a purely combinatorial necessary condition for the suitability of a potential I_{110} candidate, and the analysis of all potential candidates then takes the form of a combinatorial optimization problem over filled Young diagrams (Lemma 7.7). This technique reduces the problem to checking three cases of local contribution for $M_{\langle \mathbf{2n} \rangle}$ and eight cases for $M_{\langle \mathbf{3n} \rangle}$. This method for proving lower bounds is completely different from previous techniques.

To enable a casual reader to see the various techniques we employ, we return to the proof that $\underline{\mathbf{R}}(M_{\langle 2 \rangle}) > 6$ multiple times: first using the general algorithm naïvely in §4, then working dually to reduce the calculation (Remark 4.1), then using representation theory to block diagonalize the calculation in §6.2 and finally we observe that the result is an immediate consequence of our localization principle and Lemma 7.2 (Remark 7.3).

1.2. Results

Theorem 1.1. $\underline{\mathbf{R}}(M_{\langle 3 \rangle}) \geq 17$.

The previous lower bounds were 14 [41] in 1983, 15 [36] in 2015 and 16 [34] in 2018.

Let $\det_3 \in \mathbb{C}^9 \otimes \mathbb{C}^9 \otimes \mathbb{C}^9$ denote the 3×3 determinant polynomial considered as a tensor. That is, as a bilinear map, it inputs two 3×3 matrices and returns a third such that if the input is (M, M), the output is the cofactor matrix of M.

Strassen's laser method [39] upper bounds the exponent of matrix multiplication using "simple" tensors. In [2, 3, 1, 13], barriers to proving further upper bounds with the method were found for many tensors. In [15], we showed that the (unique up to scale) skew-symmetric tensor in $\mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^3$, which we denote $T_{skewcw,2}$, is *not* subject to these upper bound barriers and could potentially be used to prove the exponent of matrix multiplication is two via its Kronecker powers. Explicitly, if one were to prove that $\lim_{k \to \infty} \underline{\mathbf{R}}(T_{skewcw,2}^{\boxtimes k})^{\frac{1}{k}}$ equals 3, that would imply the exponent is two. One has $\underline{\mathbf{R}}(T_{skewcw,2}) = 5$ and $T_{skewcw,2}^{\boxtimes 2} = \det_3$; see [15]. Thus, the following result is important for matrix multiplication complexity *upper bounds*:

Theorem 1.2. $\underline{\mathbf{R}}(\det_3) = 17$.

The upper bound was proved in [15]. In [9], a lower bound of 15 for the Waring rank of $\det_3$ was proven. The previous border rank lower bound was 12 as discussed in [19], which follows from the Koszul flattening equations [36]. Note that had the result here turned out differently, for example, were the border rank 16 or lower, $T_{skewcw,2}$ would have immediately been promoted to the most promising tensor for proving the exponent is two; see the discussion in [15].

Remark 1.3. The computation of the trilinear map associated to $\det_3$, which inputs three matrices and outputs a number, is different than the computation of the associated polynomial, which inputs a single matrix and outputs a number. The polynomial may be computed using 12 multiplications in the naïve algorithm and using 10 with the algorithm in [17].

Previous to this paper, $M_{\langle 2 \rangle}$ was the only nontrivial matrix multiplication tensor whose border rank had been determined, despite 50 years of work on the subject. We add two more cases to this list.

Theorem 1.4. $\underline{\mathbf{R}}(M_{\langle 223 \rangle}) = 10$.

The upper bound dates back to Bini et al. in 1980 [7]. Koszul flattenings [36] give $\underline{\mathbf{R}}(M_{\langle 22\mathbf{n} \rangle}) \geq 3\mathbf{n}$. Smirnov [38] showed that $\underline{\mathbf{R}}(M_{\langle 22\mathbf{n} \rangle}) \leq 3\mathbf{n} + 1$ for $\mathbf{n} \leq 7$, and we expect equality to hold for all $\mathbf{n}$.

Theorem 1.5.

1. $\underline{\mathbf{R}}(M_{\langle 233 \rangle}) = 14$.

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

2. *We have the following border rank lower bounds:*

$\mathbf{n}$	$\underline{\mathbf{R}}(M_{\langle 2\mathbf{n}\mathbf{n}\rangle}) \geq$	$\mathbf{n}$	$\underline{\mathbf{R}}(M_{\langle 2\mathbf{n}\mathbf{n}\rangle}) \geq$	$\mathbf{n}$	$\underline{\mathbf{R}}(M_{\langle 2\mathbf{n}\mathbf{n}\rangle}) \geq$
4	$22 = 4^2 + 6$	11	$136 = 11^2 + 15$	18	$348 = 18^2 + 24$
5	$32 = 5^2 + 7$	12	$161 = 12^2 + 17$	19	$387 = 19^2 + 26$
6	$44 = 6^2 + 8$	13	$187 = 13^2 + 18$	20	$427 = 20^2 + 27$
7	$58 = 7^2 + 9$	14	$215 = 14^2 + 19$	21	$470 = 21^2 + 29$
8	$75 = 8^2 + 11$	15	$246 = 15^2 + 21$	22	$514 = 22^2 + 30$
9	$93 = 9^2 + 12$	16	$278 = 16^2 + 22$	23	$561 = 23^2 + 32$
10	$114 = 10^2 + 14$	17	$312 = 17^2 + 23$	24	$609 = 24^2 + 33$.

3. *For* $0 < \epsilon < \frac{1}{4}$, *and* $\mathbf{n} > \frac{6}{\epsilon}\frac{3\sqrt{6}+6-\epsilon}{6\sqrt{6}-\epsilon}$, $\underline{\mathbf{R}}(M_{\langle 2\mathbf{n}\mathbf{n}\rangle}) \geq \mathbf{n}^2 + (3\sqrt{6} - 6 - \epsilon)\mathbf{n}$. *In particular,* $\underline{\mathbf{R}}(M_{\langle 2\mathbf{n}\mathbf{n}\rangle}) \geq \mathbf{n}^2 + 1.32\mathbf{n} + 1$ *when* $\mathbf{n} \geq 25$.

Previously, only the near trivial result that $\underline{\mathbf{R}}(M_{\langle 2\mathbf{n}\mathbf{n}\rangle}) \geq \mathbf{n}^2 + 1$ was known by [37, Rem. p175].

The upper bound in (1) is due to Smirnov [38], where he also proved $\underline{\mathbf{R}}(M_{\langle 244\rangle}) \leq 24$, and $\underline{\mathbf{R}}(M_{\langle 255\rangle}) \leq 38$. When $\mathbf{n}$ is even, one has the upper bound $\underline{\mathbf{R}}(M_{\langle 2\mathbf{n}\mathbf{n}\rangle}) \leq \frac{7}{4}\mathbf{n}^2$ by writing $M_{\langle 2\mathbf{n}\mathbf{n}\rangle} = M_{\langle 222\rangle} \boxtimes M_{\langle 1\frac{\mathbf{n}}{2}\frac{\mathbf{n}}{2}\rangle}$, where $\boxtimes$ denotes Kronecker product of tensors; see, for example, [15].

Theorem 1.6. *For all* $\mathbf{n} \geq 18$, $\underline{\mathbf{R}}(M_{\langle 3\mathbf{n}\mathbf{n}\rangle}) \geq \mathbf{n}^2 + \sqrt{\frac{8}{3}}\mathbf{n} > \mathbf{n}^2 + 1.6\mathbf{n}$.

Previously, the only bound was the near trivial result that when $\mathbf{n} \geq 4$, $\underline{\mathbf{R}}(M_{\langle 3\mathbf{n}\mathbf{n}\rangle}) \geq \mathbf{n}^2 + 2$ by [37, Rem. p175].

Using [37, Rem. p175], one obtains:

Corollary 1.7. *For all* $\mathbf{n} \geq 18$ *and* $\mathbf{m} \geq 3$, $\underline{\mathbf{R}}(M_{\langle \mathbf{m}\mathbf{n}\mathbf{n}\rangle}) \geq \mathbf{n}^2 + \sqrt{\frac{8}{3}}\mathbf{n} + \mathbf{m} - 3$.

Theorems 1.5 and 1.6 are the first nontrivial border rank lower bounds for *any* tensor in $\mathbb{C}^{\mathbf{a}}\otimes\mathbb{C}^{\mathbf{b}}\otimes\mathbb{C}^{\mathbf{c}}$ when $\mathbf{c} > 2\max\{\mathbf{a}, \mathbf{b}\}$ other than the above mentioned near trivial result of Lickteig, vastly expanding the classes of tensors for which lower bound techniques exist.

1.3. What comes next?

1.3.1. Breaking the lower bound barriers

The geometric interpretation of the border rank lower bound barriers of [18, 21] is that all equations obtained by taking minors, called *rank methods*, are actually equations for a larger variety than $\sigma_r(\mathrm{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C))$, called the *r-th cactus variety* [11]. This cactus variety agrees with the secant variety for $r < 13$, but it quickly fills the ambient space of tensors in $\mathbb{C}^{\mathbf{m}}\otimes\mathbb{C}^{\mathbf{m}}\otimes\mathbb{C}^{\mathbf{m}}$ at latest when $r = 6\mathbf{m} - 4$. Thus one cannot prove $\underline{\mathbf{R}}(T) > 6\mathbf{m} - 4$ for any tensor T via rank/determinantal methods, in particular, with border apolarity alone.

In brief, the r-th secant variety consists of points on limits of spans of zero-dimensional smooth schemes of length r. The r-th cactus variety consists of points on limits of spans of zero-dimensional schemes of length r.

The border apolarity algorithm produces ideals, and thus to break the barrier, one needs to distinguish ideals that are limits of smooth schemes from ideals that are limits of nonsmoothable schemes, and ideals that are not limits of any sequence of saturated ideals. In principle, this can be done using *deformation theory* (see, e.g., [25]). This is exciting, as *it is the first proposed path to overcoming the lower bound barriers.*

Remark 1.8. After this paper was posted on arXiv, we went on to find an ideal passing all border apolarity tests for $M_{\langle 3\rangle}$ with $r = 17$. We are currently working to effectively implement deformation theory to

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

determine if such an example comes from an actual border rank 17 decomposition. The obstruction to doing this is the effective implementation of the theory. The naïve implementation, even on a large computer cluster, is not feasible, and we are working to develop effective computational techniques.

1.3.2. Upper bounds, especially for tensors relevant for Strassen's laser method

There is intense interest in tensors not subject to the upper bound barriers for Strassen's laser method described in [4, 1, 3, 13]. All tensors used in or proposed for the laser method have positive-dimensional symmetry groups, so the border apolarity method potentially may be applied. For example, the small Coppersmith–Winograd tensor $T_{cw,q} := \sum_{j=1}^{q} a_0 \otimes b_j \otimes c_j + a_j \otimes b_0 \otimes c_j + a_j \otimes b_j \otimes c_0$ has a very large symmetry group, namely the orthogonal group $O(q)$ [14], which has dimension $\binom{q}{2}$. Since these tensors, and their Kronecker squares tend to have border rank below the cactus barrier, we expect to be able to effectively apply the method as is to determine the border rank at least for small Kronecker powers. After this paper was posted on arXiv, border apolarity was utilized to determine the border rank of $T_{cw,2}^{\boxtimes 2}$ in [16] and the answer ended up being the known upper bound. We are developing techniques to write down usual border rank decompositions guided by the ideals produced by border apolarity to potentially determine new upper bounds for higher Kronecker powers of $T_{cw,2}$ and $\det_3$ (or to show that the known bounds are sharp). In other words, we are working to use border apolarity to inject some "science" into the "art" of finding upper bounds.

1.3.3. Geometrization of the (111) test for matrix multiplication

Our results for $M_{\langle 2\mathbf{nn}\rangle}, M_{\langle 3\mathbf{nn}\rangle}$ for general $\mathbf{n}$ only use the (210) and (120) tests as defined in §3, and we expect to be able to prove stronger results for general $\mathbf{n}$ in these cases once we develop a proper geometric understanding of the (111) test like we have for the (210) test.

1.4. Overview

In §2, we review terminology regarding border rank decompositions of tensors, Borel subgroups and Borel fixed subspaces. We then describe a curve of multigraded ideals one may associate to a border rank decomposition. We also review Borel fixed subspaces and list them in the cases relevant for this paper. In §3, we describe the border apolarity algorithm and accompanying tests. In §4, we review the matrix multiplication tensor. In §5, we describe the computation to prove Theorems 1.1 and 1.2, which are computer assisted calculations, the code for which is available at github.com/adconner/chlbapolar. In §6, we discuss representation theory relevant for applying the border apolarity algorithm to matrix multiplication. In §7, we prove our localization and optimization algorithm and use it to prove Theorems 1.4, 1.5 and 1.6.

2. Preliminaries

2.1. Definitions/Notation

Throughout, A, B, C, U, V, W will denote complex vector spaces, respectively, of dimensions $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{u}, \mathbf{v}, \mathbf{w}$. The dual space to A is denoted A^*. The space of symmetric degree d tensors is denoted $S^d A$, which may also be viewed as the space of degree d homogeneous polynomials on A^*. Set $\mathrm{Sym}(A) := \bigoplus_d S^d A$. The identity map is denoted $\mathrm{Id}_A \in A \otimes A^*$. For $X \subset A$, $X^\perp := \{\alpha \in A^* \mid \alpha(x) = 0, \forall x \in X\}$ is its annihilator, and $\langle X \rangle \subset A$ denotes the linear span of X. Projective space is $\mathbb{P}A = (A\backslash\{0\})/\mathbb{C}^*$, and if $x \in A\backslash\{0\}$, we let $[x] \in \mathbb{P}A$ denote the associated point in projective space (the line through x). The general linear group of invertible linear maps $A \to A$ is denoted $\mathrm{GL}(A)$ and the special linear group of determinant one linear maps is denoted $\mathrm{SL}(A)$. The permutation group on r elements is denoted $\mathfrak{S}_r$.

For at tensor $T \in A \otimes B \otimes C$, define its *symmetry group*

$$G_T := \{(g_A, g_B, g_C) \in \mathrm{GL}(A) \times \mathrm{GL}(B) \times \mathrm{GL}(C)/(\mathbb{C}^*)^{\times 2} \mid (g_A, g_B, g_C) \cdot T = T\}. \tag{1}$$

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

One quotients by $(\mathbb{C}^*)^{\times 2} := \{(\lambda \mathrm{Id}_A, \mu \mathrm{Id}_B, \nu \mathrm{Id}_C) \mid \lambda\mu\nu = 1\}$ because $(\nu \mathrm{Id}_A, \mu \mathrm{Id}_B, \frac{1}{\nu\mu} \mathrm{Id}_C)$ is the kernel of the map $\mathrm{GL}(A) \times \mathrm{GL}(B) \times \mathrm{GL}(C) \to \mathrm{GL}(A\otimes B\otimes C)$. Lie algebras of Lie groups are denoted with corresponding symbols in old German script, for example, $\mathfrak{g}_T$ is the Lie algebra corresponding to G_T.

The Grassmannian of r planes through the origin is denoted $G(r, A)$, which we will view in its Plücker embedding $G(r, A) \subset \mathbb{P}\Lambda^r A$. That is, given $E \in G(r, A)$, that is, a linear subspace $E \subset A$ of dimension r, if $e_1, \ldots, e_r$ is a basis of E, we represent E as a point in $\mathbb{P}(\Lambda^r A)$ by $[e_1 \wedge \cdots \wedge e_r]$. Here, the wedge product is defined by $e_1 \wedge \cdots \wedge e_r := \sum_{\sigma\in\mathfrak{S}_r} \mathrm{sgn}(\sigma) e_{\sigma(1)}\otimes \cdots \otimes e_{\sigma(r)}$.

For a set $Z \subset \mathbb{P}A$, $\overline{Z} \subset \mathbb{P}A$ denotes its Zariski closure, $\hat{Z} \subset A$ denotes the cone over Z union the origin, $I(Z) = I(\hat{Z}) \subset \mathrm{Sym}(A^*)$ denotes the ideal of Z, that is, $I(Z) = \{P \in \mathrm{Sym}(A^*) \mid P(z) = 0 \forall z \in \hat{Z}\}$, and $\mathbb{C}[\hat{Z}] = \mathrm{Sym}(A^*)/I(Z)$, denotes the homogeneous coordinate ring of $\hat{Z}$. Both $I(Z)$ and $\mathbb{C}[\hat{Z}]$ are $\mathbb{Z}$-graded by degree.

We will be dealing with ideals on products of three projective spaces, that is, we will be dealing with polynomials that are homogeneous in three sets of variables, so our ideals with be $\mathbb{Z}^3$-graded. More precisely, we will study ideals $I \subset \mathrm{Sym}(A^*)\otimes \mathrm{Sym}(B^*)\otimes \mathrm{Sym}(C^*)$, and I_{ijk} denotes the component in $S^iA^*\otimes S^jB^*\otimes S^kC^*$.

Given $T \in A\otimes B\otimes C$, we may consider it as a linear map $T_C : C^* \to A\otimes B$, and we let $T(C^*) \subset A\otimes B$ denote its image and similarly for permuted statements. A tensor T is *concise* if the maps T_A, T_B, T_C are injective, that is, if it requires all basis vectors in each of A, B, C to write down in any basis.

We remark that the tensor T may be recovered up to isomorphism from any of the spaces $T(A^*), T(B^*), T(C^*)$; see, for example, [33].

Elements $P \in S^iA^*\otimes S^jB^*\otimes S^kC^*$ may be viewed as differential operators on elements $X \in S^sA\otimes S^tB\otimes S^uC$. Write $X \lrcorner P \in S^{s-i}A\otimes S^{t-j}B\otimes S^{u-k}C$ for the contraction operation. The annihilator of X, denoted $\mathrm{Ann}(X)$, is defined to be the ideal of all $P \in \mathrm{Sym}(A^*)\otimes \mathrm{Sym}(B^*)\otimes \mathrm{Sym}(C^*)$ such that $X \lrcorner P = 0$. In the case that $X = T \in A\otimes B\otimes C$, its annihilator consists of all elements in degree (i, j, k) with one of i, j, k greater than one and the annihilators in low degrees are just the usual linear annihilators defined above. Explicitly, the annihilators in low degree are $T(C^*)^\perp \subset A^*\otimes B^*$, $T(B^*)^\perp \subset A^*\otimes C^*$ and $T(A^*)^\perp \subset B^*\otimes C^*$ and $T^\perp \subset A^*\otimes B^*\otimes C^*$.

2.2. Border rank decompositions as curves in Grassmannians

A border rank r decomposition of a tensor T is normally viewed as a curve $T(t) = \sum_{j=1}^r T_j(t)$, where each $T_j(t)$ is rank one for all $t \neq 0$, and $\lim_{t\to 0} T(t) = T$. It will be useful to change perspective, viewing a border rank r decomposition of a tensor $T \in A\otimes B\otimes C$ as a curve $E_t \subset G(r, A\otimes B\otimes C)$ satisfying

(i) for all $t \neq 0$, E_t is spanned by r rank one tensors, and
(ii) $T \in E_0$.

Example 2.1. The border rank decomposition

$$a_1\otimes b_1\otimes c_2 + a_1\otimes b_2\otimes c_1 + a_2\otimes b_1\otimes c_1 = \lim_{t\to 0} \frac{1}{t}[(a_1 + ta_2)\otimes(b_1 + tb_2)\otimes(c_1 + tc_2) - a_1\otimes b_1\otimes c_1]$$

may be rephrased as the curve

$$\begin{aligned}
E_t &= [(a_1\otimes b_1\otimes c_1) \wedge (a_1 + ta_2)\otimes(b_1 + tb_2)\otimes(c_1 + tc_2)]\\
&= [(a_1\otimes b_1\otimes c_1) \wedge (a_1\otimes b_1\otimes c_1 + t(a_1\otimes b_1\otimes c_2 + a_1\otimes b_2\otimes c_1 + a_2\otimes b_1\otimes c_1)\\
&\qquad + t^2(a_1\otimes b_2\otimes c_2 + a_2\otimes b_1\otimes c_2 + a_2\otimes b_2\otimes c_1) + t^3 a_2\otimes b_2\otimes c_2)]\\
&= [(a_1\otimes b_1\otimes c_1) \wedge (a_1\otimes b_1\otimes c_2 + a_1\otimes b_2\otimes c_1 + a_2\otimes b_1\otimes c_1\\
&\qquad + t(a_1\otimes b_2\otimes c_2 + a_2\otimes b_1\otimes c_2 + a_2\otimes b_2\otimes c_1) + t^2 a_2\otimes b_2\otimes c_2)]\\
&\subset G(2, A\otimes B\otimes C).
\end{aligned}$$

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

Here,

$$E_0 = [(a_1\otimes b_1\otimes c_1)\wedge(a_1\otimes b_1\otimes c_2 + a_1\otimes b_2\otimes c_1 + a_2\otimes b_1\otimes c_1)].$$

2.3. Multigraded ideal associated to a border rank decomposition

Given a border rank r decomposition $T = \lim_{t\to 0}\sum_{j=1}^{r} T_j(t)$, we have additional information. Let

$$I_t \subset \mathrm{Sym}(A^*)\otimes \mathrm{Sym}(B^*)\otimes \mathrm{Sym}(C^*)$$

denote the $\mathbb{Z}^3$-graded ideal of the set of r distinct points $[T_1(t)]\cup\cdots\cup[T_r(t)]$, where $I_{ijk,t}\subset S^iA^*\otimes S^jB^*\otimes S^kC^*$. Sometimes, it is more convenient to work with $I^{\perp}_{ijk,t}$ which contains equivalent information.

Example 2.2. Consider the ideal of $[a_1\otimes b_1\otimes c_1]$. In degree (ijk), we have $I_{ijk} = \langle \alpha^M\otimes\beta^N\otimes\gamma^P\rangle$, where $\alpha^M = \alpha^{m_1}\cdots\alpha^{m_i}$ etc., and M,N,P ranges over those triples where at least one of the indices appearing is not equal to 1. Thus, $I^{\perp}_{ijk} = \langle a_1^i\otimes b_1^j\otimes c_1^k\rangle$.

When we take the ideal of the union of two points, the ideal is the intersection of the two ideals, and if the points are in general position, for example, $[a_1\otimes b_1\otimes c_1]\cup[a_2\otimes b_2\otimes c_2]$, in the notation above one of the indices appearing in M,N,P must not be 1 and one must not be 2, so $I^{\perp}_{ijk} = \langle a_1^i\otimes b_1^j\otimes c_1^k, a_2^i\otimes b_2^j\otimes c_2^k\rangle$.

Thus, in Example 2.1 above, $I^{\perp}_{ijk,t} = \langle a_1^i\otimes b_1^j\otimes c_1^k, (a_1+ta_2)^i\otimes(b_1+tb_2)^j\otimes(c_1+tc_2)^k\rangle$, where the role of a_2 in Example 2.2 is played by (a_1+ta_2) and similarly for b_2,c_2. As $t\to 0$, $I^{\perp}_{ijk,t}$ limits to $I^{\perp}_{ijk} = \langle a_1^i\otimes b_1^j\otimes c_1^k, ia_1^{i-1}a_2\otimes b_1^j\otimes c_1^k + ja_1^i\otimes b_1^{j-1}b_2\otimes c_1^k + ka_1^i\otimes b_1^j\otimes c_1^{k-1}c_2\rangle$.

If the r points are in general position, then $\mathrm{codim}(I_{ijk,t}) = r$ as long as $r\le \dim S^iA^*\otimes S^jB^*\otimes S^kC^*$; see, for example, [11, Lemma 3.9].

Let $\mathbf{a}\le\mathbf{b}\le\mathbf{c}$. If $r\le\binom{\mathbf{a}+1}{2}$, then for all (ijk) with $i+j+k>1$, one has $r\le\dim S^iA^*\otimes S^jB^*\otimes S^kC^*$.

In all the examples in this paper $r\le\binom{\mathbf{a}+1}{2}$. For example, for $M_{\langle 2\mathbf{nn}\rangle}$, $\binom{\mathbf{a}+1}{2} = 2\mathbf{n}^2+\mathbf{n}$ and we prove border rank lower bounds like $\mathbf{n}^2+1.32\mathbf{n}$.

Thus, in this paper we may and will assume $\mathrm{codim}(I_{ijk}) = r$ for all (ijk) with $i+j+k>1$.

Thus, in addition to $E_0 = I^{\perp}_{111,0}$ defined in §2.2, we obtain a limiting ideal I, where we define $I_{ijk} := \lim_{t\to 0} I_{ijk,t}$ and the limit is taken in the Grassmannian of codimension r subspaces in $S^iA^*\otimes S^jB^*\otimes S^kC^*$,

$$G(\dim(S^iA^*\otimes S^jB^*\otimes S^kC^*) - r, S^iA^*\otimes S^jB^*\otimes S^kC^*).$$

We remark that there are subtleties here: The limiting ideal may not be saturated. See [11] for a discussion.

Thus, we may assume a multigraded ideal I coming from a border rank r decomposition of a concise tensor T satisfies the following conditions:

(i) I is contained in the annihilator of T, which by definition says $I_{110}\subset T(C^*)^{\perp}$, $I_{101}\subset T(B^*)^{\perp}$, $I_{011}\subset T(A^*)^{\perp}$ and $I_{111}\subset T^{\perp}\subset A^*\otimes B^*\otimes C^*$.
(ii) For all (ijk) with $i+j+k>1$, $\mathrm{codim} I_{ijk} = r$.
(iii) I is an ideal, so the multiplication maps

$$I_{i-1,j,k}\otimes A^*\oplus I_{i,j-1,k}\otimes B^*\oplus I_{i,j,k-1}\otimes C^*\to S^iA^*\otimes S^jB^*\otimes S^kC^* \tag{2}$$

have image contained in I_{ijk}.

Here, equation (2) is the sum of three maps, the first of which is the restriction of the symmetrization map $S^{i-1}A^*\otimes A^*\otimes S^jB^*\otimes S^kC^*\to S^iA^*\otimes S^jB^*\otimes S^kC^*$ to $I_{i-1,j,k}\otimes A^*$ and similarly for the others. When $i-1=0$, the first map is just the inclusion $I_{0jk}\otimes A^*\subset A^*\otimes S^jB^*\otimes S^kC^*$.

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

One may prove border rank lower bounds for T by showing that for a given r, no such I exists. For arbitrary tensors, we do not see any effective way to prove this, but for tensors with a large symmetry group, we have a vast simplification of the problem as described in the next subsection.

2.4. Lie's theorem and consequences

Lie's theorem may be stated as: Let H be a connected solvable group and W an H-module, then a closed H-fixed set $X \subset \mathbb{P}W$ contains an H-fixed point. Applying this fact to appropriately chosen sets X yield the Normal Form Lemma and its generalization, the Fixed Ideal Theorem.

Theorem 2.3 (Normal form lemma, tensor case [34]). *Let $T \in A\otimes B\otimes C$, and let $H \subset G_T$ be a connected solvable group. If $\underline{\mathbf{R}}(T) \leq r$, then there exists $E_0 \in G(r, A\otimes B\otimes C)$ corresponding to a border rank r decomposition of T as in §2.2 that is H-fixed, that is, $b \cdot E_0 = E_0$ for all $b \in H$.*

By the Normal Form Lemma, in order to prove $\underline{\mathbf{R}}(T) > r$, it is sufficient to rule out the existence of a border rank r decomposition E_t where E_0 is a H-fixed point of $G(r, A\otimes B\otimes C)$.

Theorem 2.4 (Fixed Ideal Theorem, tensor case [11]). *Let $T \in A\otimes B\otimes C$, and let $H \subset G_T$ be a connected solvable group. If $\underline{\mathbf{R}}(T) \leq r$, then there exists an ideal $I \in \mathrm{Sym}(A^*)\otimes \mathrm{Sym}(B^*)\otimes \mathrm{Sym}(C^*)$ as in §2.3 corresponding to a border rank r decomposition of a tensor T that is H-fixed, that is, $b \cdot I_{ijk} = I_{ijk}$ for all $b \in H$ and all (i, j, k).*

The conclusions of the theorems above are stronger for larger groups of symmetries H, so in this paper we will always fix a Borel subgroup $\mathbb{B}_T \subset G_T$, that is, a maximal connected solvable subgroup of G_T. Thus, we may assume a multigraded ideal I coming from a border rank r decomposition of T satisfies the additional condition:

(iv) Each I_{ijk} is $\mathbb{B}_T$-fixed.

As we explain in the next subsection, for the instances in considered in this paper, Borel fixed spaces are easy to list.

2.5. Borel fixed subspaces

All of the $\mathbb{B}_T$-modules for which we would like to study $\mathbb{B}_T$-fixed subspaces are also G_T-modules, where G_T is reductive. This fact simplifies the description of $\mathbb{B}_T$-fixed subspaces, so we will assume this in what follows.

It will be convenient for us to linearize the problem by considering Lie algebras instead of Lie groups. Let $\mathfrak{g}_T$ be the Lie algebra of G_T, and let $\mathfrak{b}_T \subset \mathfrak{g}_T$ be the Lie algebra of $\mathbb{B}_T \subset G_T$. A subspace $S \subset M$ is $\mathbb{B}_T$ fixed if and only if it is $\mathfrak{b}_T$ fixed.

2.5.1. Weights and weight diagrams

For more details on the material in this section, see any of [26, 28, 20, 10].

If one has a diagonalizable matrix, one may choose a basis of eigenvectors each of which has an associated eigenvalue. If one has a space $\mathfrak{t} \subset \mathfrak{gl}_m$ of simultaneously diagonalizable matrices, we may choose a basis of simultaneous eigenvectors, say $e_1, \ldots, e_m$. Instead of considering the eigenvalues of each individual matrix, it is convenient to think of all the eigenvalues simultaneously as elements of $\mathfrak{t}^*$, and these generalized eigenvalues are called *weights*. Write the weights as $\mu_1, \ldots, \mu_{\mathbf{v}} \in \mathfrak{t}^*$. Then, given $X \in \mathfrak{t}$, one has $Xe_j = \mu_j(X)e_j$, where the number $\mu_j(X)$ is X's usual eigenvalue for the eigenvector e_j. In this context, the eigenvectors are called *weight vectors*.

Since $\mathfrak{g}_T$ is reductive, there exists a unique up to conjugation *maximal torus* $\mathfrak{t} \subset \mathfrak{g}_T$, and the choice of $\mathfrak{b}_T$ fixes a unique $\mathfrak{t} \subset \mathfrak{b}_T$. The maximal torus is an abelian subalgebra such that the adjoint action on $\mathfrak{g}_T$ is simultaneously diagonalizable and the weight zero space under this action is exactly $\mathfrak{t}$. That is, $\mathfrak{g}_T = \mathfrak{t} \oplus \bigoplus_{\alpha\neq 0} \mathfrak{g}_\alpha$, where $\mathfrak{g}_\alpha$ is the weight space under the adjoint action of $\mathfrak{t}$ corresponding to weight α.

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

The nonzero weights under the adjoint action of $\mathfrak{t}$ are called the *roots* of $\mathfrak{g}_T$, and the corresponding $\mathfrak{g}_\alpha$ *root spaces*. For a root α, one has $\dim \mathfrak{g}_\alpha = 1$. We have $\mathfrak{b}_T = \mathfrak{t} \oplus \bigoplus_{\alpha \in P} \mathfrak{g}_\alpha$, for some subset P of roots which are called *positive*. The positive roots define a partial order on the set of all roots, where $\alpha < \beta$ when $\beta - \alpha \in P$. In this language, $\mathfrak{b}_T = \mathfrak{t} \oplus \bigoplus_{\alpha > 0} \mathfrak{g}_\alpha$. Call $\mathfrak{n} := \bigoplus_{\alpha > 0} \mathfrak{g}_\alpha$ the set of *raising operators*, which is a nilpotent Lie-subalgebra of $\mathfrak{g}_T$. Inside P are the *simple roots*, those which cannot be written as sums of two elements of P.

Any $\mathfrak{g}_T$-module M is also simultaneously diagonalizable under $\mathfrak{t}$, say $M = \bigoplus_\lambda M_\lambda$, and $\mathfrak{g}_\alpha . M_\lambda \subset M_{\lambda+\alpha}$. A weight vector in M is a *highest weight vector* if it is annihilated by the action of $\mathfrak{n}$. We can summarize the action of $\mathfrak{b}_T$ on M with a *weight diagram*, a graph with vertices corresponding to M_μ and edges corresponding to the action of the $\mathfrak{g}_\alpha$, where α is a simple root. Edges remain unlabelled as their weight is implicitly determined by their source and target. Since each $\mathfrak{g}_\alpha$ is one dimensional, edges may be interpreted as a single linear map up to scale from the source to the target. The partial order on roots extends naturally to a partial order on weights: $\lambda \geq \mu$ when $\lambda = \mu + \sum_{\alpha > 0} k_\alpha \alpha$, where $k_\alpha \geq 0$ (or equivalently, where the sum ranges over simple roots). We draw weight diagrams so that when $\lambda \geq \mu$, then M_λ is placed higher than M_μ.

Suppose $S \subset M$ is a $\mathfrak{b}_T$ fixed subspace. S is $\mathfrak{t}$-fixed, so it is spanned by weight vectors, that is, $S = \bigoplus_\lambda S_\lambda$, $S_\lambda = S \cap M_\lambda$. Furthermore, S is closed under raising operators, which means that $\mathfrak{g}_\alpha . S_\lambda \subset S_{\lambda+\alpha}$ for each positive (or each simple) root α. Thus, $\mathfrak{b}_T$ fixed subspaces of M are precisely those $S = \bigoplus_\lambda S_\lambda$ where S_λ maps inside S_μ under every arrow $M_\lambda \to M_\mu$ in the weight diagram of M.

2.5.2. Parameterizing Borel fixed subspaces

We may parameterize the $\mathfrak{b}_T$ fixed subspaces $S \subset M$ of dimension d as follows: Fix an assignment of dimensions $d_\lambda, 0 \leq d_\lambda \leq \dim M_\lambda, \sum_\lambda d_\lambda = d$. Choices of S_λ with $\dim S_\lambda = d_\lambda$ are parameterized by the product of Grassmannians $X = \prod_\lambda G(d_\lambda, M_\lambda)$. Given a raising operator x corresponding to an arrow $M_\lambda \to M_\mu$ in the weight diagram, the condition that $x.S_\lambda \subset S_\mu$ is an explicit polynomial condition on X. Cutting X by all such polynomials gives a description of the set of $\mathfrak{b}_T$ fixed subspaces with $\dim S_\lambda = d_\lambda$ (which can be empty). All Borel fixed subspaces are obtained as d_λ ranges over all such assignments. In small examples, a complete list of $\mathfrak{b}_T$ fixed subspaces may frequently be read off of the weight diagram.

2.5.3. $\mathfrak{gl}_m$ and $\mathfrak{sl}_m$ weights

All of the groups appearing as G_T in this paper are GL_m and SL_m and products of such. In this case, a Borel subgroup in some choice of basis is just the group of invertible upper triangular matrices (in the case of SL_m, with determinant one) or the product of such.

For $\mathbb{B}$ the invertible upper triangular matrices, $\mathfrak{b}$ is just all upper triangular matrices. Here, $\mathfrak{b} = \mathfrak{t} \oplus \mathfrak{n}$, where $\mathfrak{t}$ is the diagonal matrices and $\mathfrak{n}$ is the set of upper triangular matrices with zero on the diagonal.

Let $\epsilon_1, \ldots, \epsilon_m \in \mathfrak{t}^*$ be the basis dual to the basis $e_{11}, \ldots, e_{mm}$ of $\mathfrak{t} \subset \mathfrak{gl}_m$, and and write $\epsilon_j = (0, \ldots, 0, 1, 0, \ldots, 0)$, where the 1 is in the j-th slot. Let $\mathbb{C}^m$ have standard basis $e_1, \ldots, e_m$, with dual basis $e^1, \ldots, e^m$. Then $e_1, \ldots, e_m$ are weight vectors of $\mathfrak{t}$, and e_j has weight ϵ_j.

If $g \in G$ acts on V by $v \mapsto gv$, then the induced action on V^* is $\alpha \mapsto \alpha \circ g^{-1}$ so that $g \cdot (\alpha(v)) = (\alpha \circ g^{-1})(g \cdot v) = \alpha(v)$. When we differentiate this action, the induced Lie algebra action is $X.\alpha = -\alpha \circ X$. Thus, considering the action of $\mathfrak{t}$ on $(\mathbb{C}^m)^*$, $e^1, \ldots, e^j$ are the set of weight vectors and $wt(e^j) = -\epsilon_j = (0, \ldots, 0, -1, 0, \ldots, 0)$.

For vectors in $(\mathbb{C}^m)^{\otimes d}$, $wt(e_1^{\otimes a_1} \otimes \cdots \otimes e_m^{\otimes a_m}) = a_1\epsilon_1 + \cdots + a_m\epsilon_m$ and the weight is unchanged under permutations of the $d = a_1 + \cdots + a_m$ factors. The partial order on weights of §2.5.1 may be written $a_1\epsilon_1 + \cdots + a_m\epsilon_m \geq b_1\epsilon_1 + \cdots + b_m\epsilon_m$ if $\sum_{i=1}^s a_i \geq \sum_{i=1}^s b_i$ for all s.

The Lie algebra $\mathfrak{sl}_m$ corresponding to SL_m consists of the trace free $m \times m$ matrices. Here, $\mathfrak{t}$ is the set of diagonal matrices with trace zero, so the set of weights is defined only modulo $\epsilon_1 + \cdots + \epsilon_m$. We will write $\mathfrak{sl}_m$ weights as $c_1\omega_1 + \cdots + c_{m-1}\omega_{m-1}$, where the $\omega_j := \epsilon_1 + \cdots + \epsilon_j$ are called the *fundamental weights*. Thus, in terms of $\mathfrak{sl}_m$ weights, $wt(e_1) = \omega_1$, for $2 \leq s \leq m-1$, $wt(e_s) = \omega_s - \omega_{s-1}$, $wt(e_m) = -\omega_{m-1}$, and for all j, $wt(e^j) = -wt(e_j)$.

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

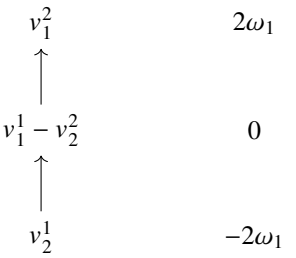

Figure 1. *Weight diagram for adjoint representation of* $\mathfrak{sl}_2$

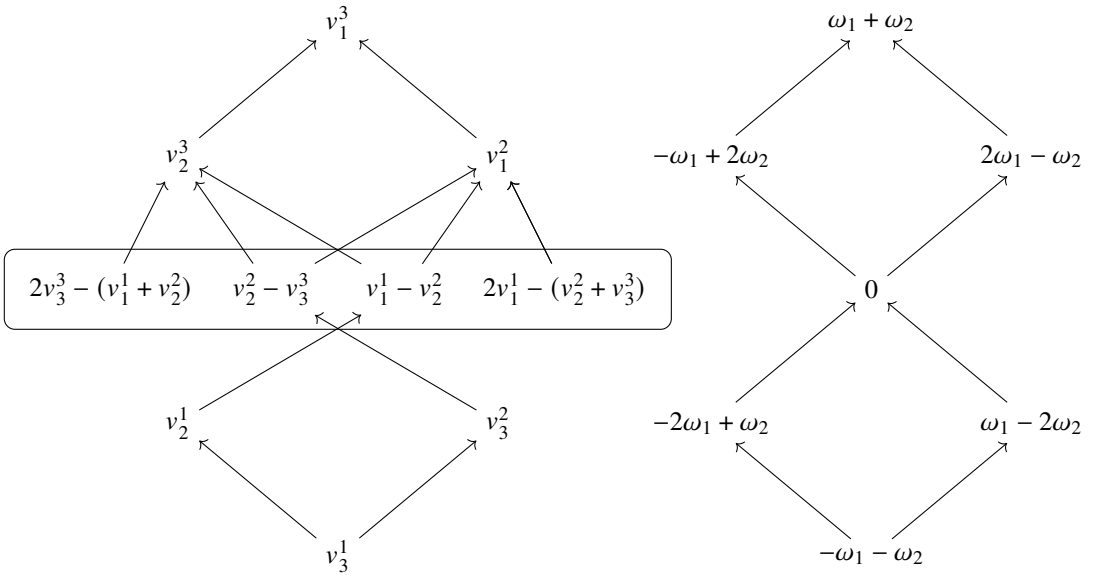

Figure 2. *Weight diagram for adjoint representation of* $\mathfrak{sl}_3$

In terms of $\mathfrak{sl}_m$ weights, the partial order is thus $a_1\omega_1 + \cdots + a_{m-1}\omega_{m-1} \geq b_1\omega_1 + \cdots + b_{m-1}\omega_{m-1}$ when $a_i \geq b_i$ for all i. For every $\mathfrak{sl}_m$ weight $\lambda \geq 0$, there is a unique irreducible module denoted V_λ containing a highest weight vector of weight λ. See, for example, [26, Chap. 6] or [28, Chap. 5, §2] for details.

Example 2.5 ($\mathfrak{sl}_2$ as an $\mathfrak{sl}_2$-module). This example will be used in the proofs of Theorems 1.4 and 1.5. Figure 1 gives the weight diagram for $\mathfrak{sl}_2 = \mathfrak{sl}(V)$ as a $\mathfrak{sl}_2$-module under the adjoint action, that is, for $X, Y \in \mathfrak{sl}_2$, $X.Y = XY - YX$. Here, v_1, v_2 is a basis of V with dual basis v^1, v^2 and $v_j^i := v_j \otimes v^i$.

The only $\mathbb{B}$-fixed subspaces are 0, $\langle v_1^2 \rangle$, $\langle v_1^2, v_1^1 - v_2^2 \rangle$ and $\langle v_1^2, v_1^1 - v_2^2, v_1^2 \rangle$.

Example 2.6 ($\mathfrak{sl}_3$ as an $\mathfrak{sl}_3$-module). This example will be used in the proofs of Theorems 1.1 and 1.6. Figure 2 gives the weight diagram for $\mathfrak{sl}_3$ as an $\mathfrak{sl}_3$-module under the adjoint action. As above $v_j^i = v_j \otimes v^i$. The oval is around the two-dimensional weight zero subspace, which has four distinguished one-dimensional subspaces: the images of the two raising operators in and the kernels of the two raising operators out. Additional arrows indicating these relationships have been added to the weight diagram.

The $\mathbb{B}$-fixed subspaces of dimension three are $\langle v_1^3, v_2^3, v_1^2 \rangle$, $\langle v_1^3, v_2^3, 2v_3^3 - (v_1^1 + v_2^2) \rangle$ and $\langle v_1^3, v_1^2, 2v_1^1 - (v_2^2 + v_3^3) \rangle$.

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

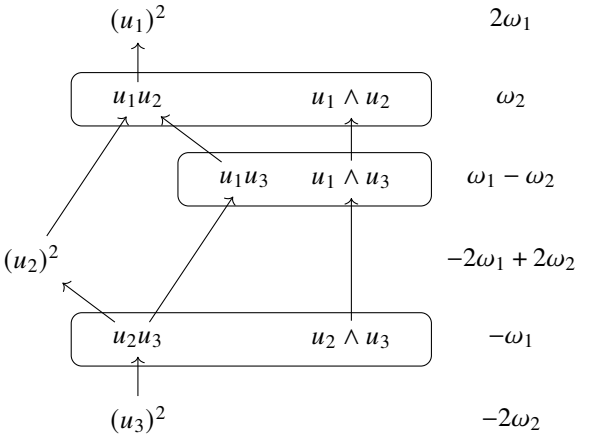

Figure 3. *Weight diagram for $U \otimes U$ when $U = \mathbb{C}^3$. There are six distinct weights appearing, indicated on the right. On the far left are the weight vectors in S^2U, and in the middle are the weight vectors in $\Lambda^2 U$*

The $\mathbb{B}$-fixed subspaces of dimension four are a family parametrized by $[s,t] \in \mathbb{P}^1$: $\langle v_1^3, v_2^3, v_1^2, s(v_2^2 - v_3^3) + t(v_1^1 - v_2^2)\rangle$. There are no others: We cannot include the entire weight zero space, as then we must also include all the positive weight vectors for a total dimension of five, exceeding our limit. If we include a negative weight vector, we must include its image in the weight zero space, which again raises to all positive weight vectors, exceeding our limit.

The $\mathbb{B}$-fixed subspaces of dimension five are $\langle v_1^3, v_2^3, v_1^2, v_2^2 - v_3^3, v_3^2\rangle$, $\langle v_1^3, v_2^3, v_1^2, v_1^1 - v_2^2, v_2^1\rangle$ and the span of the weight ≥ 0 space $\langle v_1^3, v_2^3, v_1^2, v_2^2 - v_3^3, v_1^1 - v_2^2\rangle$. This is easy to see as were v_2^1, v_3^2 both present we would need the full weight zero space making the dimension six, and v_3^1 can be included only if the whole Lie algebra is included.

Example 2.7 (Bilinear forms on U^*). This example will be used in the proof of Theorem 1.2. Let $\dim U = 3$ with basis u_1, u_2, u_3. Figure 3 gives the weight diagram for $U \otimes U = S^2U \oplus \Lambda^2 U$ as an $\mathfrak{sl}(U)$-module. The action of $X \in \mathfrak{sl}(U)$ on a matrix $Z \in U \otimes U$ is $Z \mapsto XZ + ZX^{\mathbf{t}}$. There are two $\mathbb{B}$-fixed lines $\langle (u_1)^2\rangle$ and $\langle u_1 \wedge u_2\rangle$, there is a 1-(projective) parameter $[s,t] \in \mathbb{P}^1$ space of $\mathbb{B}$-fixed 2-planes, $\langle (u_1)^2, su_1u_2 + tu_1 \wedge u_2\rangle$ plus an isolated one $\langle u_1 \wedge u_2, u_1 \wedge u_3\rangle$.

Example 2.8 (Tensor products of modules for different groups). Suppose M and N are modules for groups G and H, respectively. Then $M \otimes N$ is a $G \times H$ module with weight spaces $M_\lambda \otimes N_\mu$, as λ and μ range over all pairs of weights of M and N. For each arrow $M_\lambda \to M_\nu$ in the weight diagram of M corresponding to the raising operator x, there is an edge $M_\lambda \otimes N_\mu \to M_\nu \otimes N_\mu$ corresponding to the raising operator $x \oplus 0 \in \mathfrak{g} \oplus \mathfrak{h}$. Similarly, for each arrow $N_\mu \to N_\nu$ in the weight diagram of N corresponding to the raising operator y, there is an edge $M_\lambda \otimes N_\mu \to M_\lambda \otimes N_\nu$ corresponding to the raising operator $0 \oplus y \in \mathfrak{g} \oplus \mathfrak{h}$. Example 2.9 is a special case of this applied twice to obtain the diagram of a triple tensor product $U^* \otimes \mathfrak{sl}(V) \otimes W$ with $\mathbf{u} = \mathbf{v} = \mathbf{w} = 2$. This example with $\mathbf{u} = \mathbf{w} = \mathbf{n}$ and $\mathbf{v} = 2$ (resp. $\mathbf{v} = 3$) is used in the proof of Theorem 1.5 (resp. 1.6).

Example 2.9 ($U^* \otimes \mathfrak{sl}(V) \otimes W$ as an $\mathfrak{sl}(U) \times \mathfrak{sl}(V) \times \mathfrak{sl}(W)$-module). This example is crucial for the case of $M_{\langle \mathbf{u},\mathbf{v},\mathbf{w}\rangle}$ as then $A \otimes B = (U^* \otimes V) \otimes (V^* \otimes W) = U^* \otimes \mathfrak{sl}(V) \otimes W \oplus M_{\langle \mathbf{u},\mathbf{v},\mathbf{w}\rangle}(C^*)$. When U, V, W each have dimension two, Figure 4 gives the $\mathfrak{sl}(U) \times \mathfrak{sl}(V) \times \mathfrak{sl}(W)$-weight diagram for $U^* \otimes \mathfrak{sl}(V) \otimes W$. Set $x_j^i = u^i \otimes v_j$ and $y_j^i = v^i \otimes w_j$. There is a unique $\mathbb{B}$-fixed line, $\langle x_1^2 \otimes y_1^2\rangle$, three $\mathbb{B}$-fixed 2-planes, $\langle x_1^2 \otimes y_1^2, x_1^1 \otimes y_1^2\rangle$, $\langle x_1^2 \otimes y_1^2, x_1^2 \otimes y_2^2\rangle$, and $\langle x_1^2 \otimes y_1^2, x_1^2 \otimes y_1^1 - x_2^2 \otimes y_1^2\rangle$, and four $\mathbb{B}$-fixed 3-planes, $\langle x_1^2 \otimes y_1^2, x_1^1 \otimes y_1^2, x_1^2 \otimes y_1^1 - x_2^2 \otimes y_1^2\rangle$, $\langle x_1^2 \otimes y_1^2, x_1^2 \otimes y_1^1 - x_2^2 \otimes y_1^2, x_1^2 \otimes y_2^2\rangle$, $\langle x_1^2 \otimes y_1^2, x_1^1 \otimes y_1^2, x_1^2 \otimes y_2^2\rangle$, and $\langle x_1^2 \otimes y_1^2, x_1^2 \otimes y_1^1 - x_2^2 \otimes y_1^2, x_2^2 \otimes y_1^1\rangle$.

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

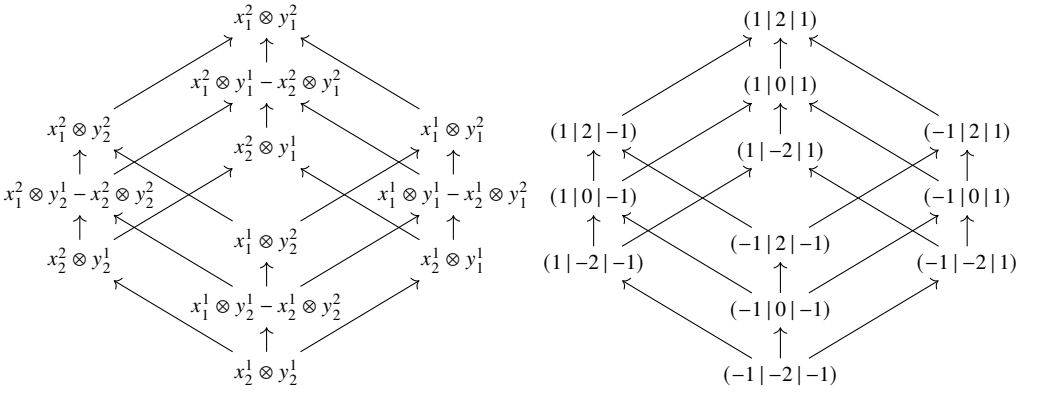

Figure 4. *Weight diagram for $U^*\otimes\mathfrak{sl}(V)\otimes W$ when $U = V = W = \mathbb{C}^2$. Left are the weight vectors and right the weights: Since $\mathfrak{sl}_2$ weights are just $j\omega_1$, we have just written $(i \mid j \mid k)$ for the $\mathfrak{sl}(U)\oplus\mathfrak{sl}(V)\oplus\mathfrak{sl}(W)$ weight. Raisings in U^* correspond to NW (northwest) arrows, those in W to NE (northeast) arrows and those in $\mathfrak{sl}(V)$ to upward arrows*

3. The ideal enumeration algorithm

Input: An integer r, a concise tensor $T \in A\otimes B\otimes C$, and a (possibly trivial) Borel subgroup $\mathbb{B}_T \subset G_T$.

Output: A list of parameterized families of ideals which together exhaust those satisfying conditions (i)-(iv) in §2.3 and §2.4.

Remark 3.1. This algorithm may find that there are no such ideals, in which case $\underline{\mathbf{R}}(T) > r$. If the output is a nonempty set of Borel-fixed ideals, without any further work one cannot conclude anything. As mentioned above, techniques exist that in principle will determine if an ideal deforms to an ideal of r distinct points (in which case the border rank is at most r) or does not (if one proves that all ideals on the list fail to deform to an ideal of r distinct points, then one concludes the border rank is greater than r), but these techniques are not implementable in the examples of interest such as $M_{\langle 3\rangle}$ at this writing. However, since there is no theoretical obstruction to the computation, we have a potential path forward for further lower bounds, and even in principle superlinear lower bounds. To our knowledge, no other path to further lower bounds has been proposed.

In what follows, we take for granted that a suitable description of the variety of $\mathbb{B}_T$-fixed subspaces of given dimension of any G_T-module M may always be computed. When G_T is reductive, a convenient such description is described in §2.5.2.

In fact, such a description is always available in general. For instance, we may represent subspaces in Plücker coordinates and observe that a subspace $S \subset M$ of dimension $\mathbf{s}$ is $\mathbb{B}_T$-fixed if and only if $[\Lambda^{\mathbf{s}}S] \subset \mathbb{P}(\Lambda^{\mathbf{s}}M)$ is $\mathbb{B}_T$-fixed. Equivalently, $[\Lambda^{\mathbf{s}}S]$ is fixed under the Lie algebra of $\mathbb{B}_T$. The condition of being fixed under one element of the Lie algebra is a quadratic condition in the Plücker coordinates of S and being fixed under the whole Lie algebra is the same as being fixed under a basis.

Here is the algorithm to build an ideal I in each multidegree. We initially have $I_{100} = I_{010} = I_{001} = 0$ (by conciseness), so the first spaces to build are in total degree two.

(i) For each $\mathbb{B}_T$-fixed family of subspaces F_{110} of codimension $r - \mathbf{c}$ in $T(C^*)^{\perp} \subset A^*\otimes B^*$ (and codimension r in $A^*\otimes B^*$), restrict the family to the closed set on which the following symmetrization maps have images of codimension at least r.

$$F_{110}\otimes A^* \to S^2A^*\otimes B^*, \text{ and} \tag{3}$$

$$F_{110}\otimes B^* \to A^*\otimes S^2B^*. \tag{4}$$

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

After this restriction, we have a (possibly empty) candidate family of components I_{110}. Call these maps the (210) and (120) maps and the rank conditions the (210) and (120) tests.

(ii) Perform the analogous tests for spaces $F_{101} \subset T(B^*)^\perp$ and $F_{011} \subset T(A^*)^\perp$ to obtain candidate families I_{101}, I_{011}.

(iii) For each triple $F_{110}, F_{101}, F_{011}$ of families passing the above tests, restrict the product of these families to the closed set on which the following addition map has image of codimension at least r.

$$F_{110}\otimes C^* \oplus F_{101}\otimes B^* \oplus F_{011}\otimes A^* \to A^*\otimes B^*\otimes C^*. \tag{5}$$

After this restriction, we have a (possibly empty) candidate family of compatible triples. Call this map the (111) map and the rank condition the (111) test.

(iv) In the language of [24, §3], let D be a set of multidegrees which is very supportive for the Hilbert function corresponding to our codimension r condition. Such a set D may be effectively constructed by following the proof of [24, Proposition 3.2]. By [24, Theorem 3.6], an ideal generated in multidegrees D satisfying the codimension condition in D satisfies it in every multidegree. For simplicity, assume further D is closed under taking smaller multidegrees in the partial order. Fix an ordered list $(\alpha_s)_s$ of the remaining undetermined multidegrees in D which respects the partial order in $\mathbb{Z}^3$.

For each t, write $(ijk) = \alpha_t$, and do the following to determine the families of candidate sets $\{F_{\alpha_s}\}_{s\le t}$. For all pairs of (i) a family of candidate sets $\{F_{\alpha_s}\}_{s\le t-1}$ and (ii) an $\mathbb{B}_T$-fixed family of subspaces $F_{i,j,k} \subset S^iA^*\otimes S^jB^*\otimes S^kC^*$ of codimension r, restrict the product of these families to the closed set on which the symmetrization and addition map

$$F_{i-1,j,k}\otimes A^* \oplus F_{i,j-1,k}\otimes B^* \oplus F_{i,j,k-1}\otimes C^* \to S^iA^*\otimes S^jB^*\otimes S^kC^* \tag{6}$$

has image contained in $F_{i,j,k}$. The output of the algorithm consists of the family of candidate sets $\{F_{\alpha_s}\}_{\alpha_s\in D}$. The conditions on D ensure that this output is correct and exhaustive.

Remark 3.2. All the results of this paper with the exception of Theorems 1.1 and 1.2 require only step (i) of this algorithm. Theorems 1.1 and 1.2 require steps (i), (ii) and (iii) only, and are carried out via computer implementation.

Remark 3.3. Only step (iv) is needed for the algorithm to be complete and correct. Applying the tests of steps (i)–(iii) is an attempt to rule out bad candidates early and avoid costly redundant work. This heuristic in practice greatly simplifies the initial steps of the search (e.g., the previous remark).

Proposition 3.4. *The algorithm terminates in a finite number of steps.*

Proof. All of the steps of the above algorithm which manipulate infinite families of candidates may be accomplished in finite time using the standard technology of computational algebraic geometry, for example, Gröbner bases. As only the finitely many components with multidegree in D must be determined, and each has only finitely many parameterized families of $\mathbb{B}_T$-fixed subspaces, complete enumeration requires only finitely many steps. □

Sometimes, it is more convenient to perform the tests dually.

Proposition 3.5. *The codimension of the image of the* (210)*-map is the dimension of the kernel of the skew-symmetrization map*

$$F_{110}^\perp\otimes A \to \Lambda^2 A\otimes B. \tag{7}$$

The kernel of the transpose of the (ijk)*-map (6) is*

$$(F_{i,j,k-1}^\perp\otimes C) \cap (F_{i,j-1,k}^\perp\otimes B) \cap (F_{i-1,j,k}^\perp\otimes A). \tag{8}$$

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

Remark 3.6. The expression (8) should be interpreted in view of the canonical embedding $S^iA \subset S^{i-1}A\otimes A$ and its analogues for B and C, with the intersection taking place in

$$(S^{i-1}A\otimes A)\otimes(S^{j-1}B\otimes B)\otimes(S^{k-1}C\otimes C).$$

That this intersection lies in $S^iA\otimes S^jB\otimes S^kC$ is part of the assertion.

Proof. The transpose of the (210) map (3) is

$$\begin{aligned} S^2A\otimes B \to F_{110}^*\otimes A &= [(A\otimes B)/F_{110}^\perp]\otimes A \\ &= A\otimes A\otimes B/(F_{110}^\perp\otimes A) \\ &= (\Lambda^2A\otimes B\oplus S^2A\otimes B)/(F_{110}^\perp\otimes A). \end{aligned}$$

Since the source maps to $S^2A\otimes B$, the kernel equals $(S^2A\otimes B)\cap(F_{110}^\perp\otimes A)$, which in turn is the kernel of equation (7).

We now prove the assertion regarding equation (8). Let $X\in S^iA\otimes S^jB\otimes S^kC$. Write $\mathrm{Proj}_{i,j,k-1}(X) = X + F_{i,j,k-1}^\perp\otimes C$, $\mathrm{Proj}_{i,j-1,k}(X) = X + F_{i,j-1,k}^\perp\otimes B$, and $\mathrm{Proj}_{i-1,j,k}(X) = X + F_{i-1,j,k}^\perp\otimes A$. The transpose of equation (8) is the map

$$\begin{aligned} S^iA\otimes S^jB\otimes S^kC &\to F_{i,j,k-1}^*\otimes C\oplus F_{i,j-1,k}^*\otimes B\oplus F_{i-1,j,k}^*\otimes A \\ X &\mapsto \mathrm{Proj}_{i,j,k-1}(X)\oplus\mathrm{Proj}_{i,j-1,k}(X)\oplus\mathrm{Proj}_{i-1,j,k}(X) \end{aligned}$$

so X is in the kernel if and only if all three projections are zero. The kernels of the three projections, respectively, are $F_{i,j,k-1}^\perp\otimes C$, $F_{i,j-1,k}^\perp\otimes B$, and $F_{i-1,j,k}^\perp\otimes A$, each intersected with $S^iA\otimes S^jB\otimes S^kC$. Take intersections term by term in the tensor product to get $(F_{i,j,k-1}^\perp\otimes C)\cap(F_{i,j-1,k}^\perp\otimes B)\cap(F_{i-1,j,k}^\perp\otimes A)\subset S^iA\otimes S^jB\otimes S^kC$, and we conclude. □

4. Matrix multiplication

Let $A = U^*\otimes V$, $B = V^*\otimes W$, $C = W^*\otimes U$. The matrix multiplication tensor $M_{\langle\mathbf{u},\mathbf{v},\mathbf{w}\rangle}\in A\otimes B\otimes C$ is the re-ordering of $\mathrm{Id}_U\otimes\mathrm{Id}_V\otimes\mathrm{Id}_W\in(U^*\otimes U)\otimes(V^*\otimes V)\otimes(W^*\otimes W)$. Thus, $G_{M_{\langle\mathbf{u},\mathbf{v},\mathbf{w}\rangle}}\supseteq \mathrm{PGL}(U)\times\mathrm{PGL}(V)\times\mathrm{PGL}(W) =: G$, where here $\mathrm{PGL}(V) = \mathrm{GL}(V)/\mathbb{C}^*$. As a G-module $A^*\otimes B^* = U\otimes\mathfrak{sl}(V)\otimes W^*\oplus U\otimes\mathrm{Id}_V\otimes W^*$. We have $M_{\langle\mathbf{u},\mathbf{v},\mathbf{w}\rangle}(C^*) = U^*\otimes\mathrm{Id}_V\otimes W$. We fix bases and let $\mathbb{B}$ denote the induced Borel subgroup of G of triples of upper-triangular, $\mathbf{u}\times\mathbf{u}$, $\mathbf{v}\times\mathbf{v}$, and $\mathbf{w}\times\mathbf{w}$ matrices.

For dimension reasons, it will be easier to describe $E_{ijk} := F_{ijk}^\perp\subset S^iA\otimes S^jB\otimes S^kC$ than F_{ijk}. Note that E_{ijk} is $\mathbb{B}$-fixed if and only if F_{ijk} is. Any $\mathbb{B}$-fixed candidate E_{110} is an enlargement of $U^*\otimes\mathrm{Id}_V\otimes W$ obtained from choosing a $\mathbb{B}$-fixed $(r-\mathbf{wu})$-plane inside $U^*\otimes\mathfrak{sl}(V)\otimes W$. This is because $F_{110}\subseteq T(C^*)^\perp$ says that $E_{110} := F_{110}^\perp\supseteq T(C^*) = U^*\otimes\mathrm{Id}_V\otimes W$. Write $E_{110} = (U^*\otimes\mathrm{Id}_V\otimes W)\oplus E'_{110}$, where $E'_{110}\subset U^*\otimes\mathfrak{sl}(V)\otimes W$ and $\dim E'_{110} = r-\mathbf{wu}$.

First proof that $\underline{\mathbf{R}}(M_{\langle 2\rangle}) = 7$. Here, $\mathbf{u} = \mathbf{v} = \mathbf{w} = 2$. We show $\underline{\mathbf{R}}(M_{\langle 2\rangle}) > 6$ by checking that no $\mathbb{B}$-fixed 10-dimensional F_{110} (i.e., six-dimensional E_{110} or two-dimensional E'_{110}) passes both the (210) and (120) tests. The weight diagram for $U^*\otimes\mathfrak{sl}(V)\otimes W$ appears in Figure 4.

By Figure 4, there are three $\mathbb{B}$-fixed 2-planes E'_{110} in $U^*\otimes\mathfrak{sl}(V)\otimes W$. For each, we compute the ranks of the corresponding maps $m_1 : F_{110}\otimes A^*\to S^2A^*\otimes B^*$ and $m_2 : F_{110}\otimes B^*\to A^*\otimes S^2B^*$, which are given by 40×40 matrices:

E'_{110}	m_1 rank	m_2 rank
$\langle x_1^2\otimes y_1^2, x_1^2\otimes y_2^2\rangle$	36	34
$\langle x_1^2\otimes y_1^2, x_1^2\otimes y_1^1 - x_2^2\otimes y_1^2\rangle$	35	35
$\langle x_1^2\otimes y_1^2, x_1^1\otimes y_1^2\rangle$	34	36

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

We see that for each candidate E'_{110}, at least one of the maps has rank strictly greater than $34 = 40 - 6$, and we conclude.

For readers unhappy with computing the rank of a sparse 40×40 matrix whose entries are all $0, \pm 1$, Remark 4.1 below reduces to 24×24 matrices, and in §6.2, using more representation theory, we reduce to 4×8 matrices whose entries are all $0, \pm 1$. Finally, we give a calculation free proof in Remark 7.3.

Remark 4.1. We may also proceed according to Proposition 3.5 and instead compute the ranks of the maps $E_{110}\otimes A \to \Lambda^2 A\otimes B$ and $E_{110}\otimes B \to A\otimes\Lambda^2 B$. The images of the basis vectors of $E_{110}\otimes A$ in the case $E'_{110} = \langle x_1^2\otimes y_1^2, x_1^1\otimes y_1^2\rangle$ are

$$
\begin{aligned}
&x_1^1 \wedge x_1^2\otimes y_1^2, x_2^1 \wedge x_1^2\otimes y_1^2, x_2^2 \wedge x_1^2\otimes y_1^2,\\
&x_2^1 \wedge x_1^1\otimes y_1^2, x_2^2 \wedge x_1^1\otimes y_1^2,\\
&x_1^1 \wedge (x_1^1\otimes y_1^1 + x_2^1\otimes y_1^2), x_2^1 \wedge (x_1^1\otimes y_1^1 + x_2^1\otimes y_1^2), x_1^2 \wedge (x_1^1\otimes y_1^1 + x_2^1\otimes y_1^2), x_2^2 \wedge (x_1^1\otimes y_1^1 + x_2^1\otimes y_1^2),\\
&x_1^1 \wedge (x_1^2\otimes y_1^1 + x_1^2\otimes y_1^2), x_2^1 \wedge (x_1^2\otimes y_1^1 + x_1^2\otimes y_1^2), x_1^2 \wedge (x_1^2\otimes y_1^1 + x_2^2\otimes y_1^2), x_2^2 \wedge (x_1^2\otimes y_1^1 + x_2^2\otimes y_1^2),\\
&x_1^1 \wedge (x_1^2\otimes y_2^1 + x_2^2\otimes y_2^2), x_2^1 \wedge (x_1^2\otimes y_2^1 + x_2^2\otimes y_2^2), x_1^2 \wedge (x_1^2\otimes y_2^1 + x_2^2\otimes y_2^2), x_2^2 \wedge (x_1^2\otimes y_2^1 + x_2^2\otimes y_2^2)\\
&x_1^1 \wedge (x_1^2\otimes y_2^1 + x_2^2\otimes y_2^2), x_2^1 \wedge (x_1^2\otimes y_2^1 + x_2^2\otimes y_2^2), x_1^2 \wedge (x_1^2\otimes y_2^1 + x_2^2\otimes y_2^2), x_2^2 \wedge (x_1^2\otimes y_2^1 + x_2^2\otimes y_2^2)
\end{aligned}
$$

and if we remove one of the two $x_1^2 \wedge (x_1^2\otimes y_1^1 + x_2^2\otimes y_1^2)$'s we obtain a set of 20 independent vectors.

After choosing isomorphisms $U \cong V \cong W$, the square matrix multiplication tensor $M_{\langle \mathbf{n}\rangle}$ has $\mathbb{Z}_3$-symmetry via cyclic permutation of factors. If the isomorphisms $U \cong V \cong W$ are chosen (uniquely) to identify $\mathbb{B}$-fixed subspaces with $\mathbb{B}$-fixed subspaces, cyclic permutation gives a correspondence between the candidate F_{110}, F_{101} and F_{011} sets. This fact is used in §5 to simplify the calculation, as there it is necessary to carry out the ideal enumeration algorithm up to the (111) test.

Similarly, when $\mathbf{u} = \mathbf{w}$, given a choice of isomorphism $U \cong W$ there is a corresponding transpose symmetry $A\otimes B\otimes C \leftrightarrow B^*\otimes A^*\otimes C^*$ of $M_{\langle \mathbf{u},\mathbf{v},\mathbf{u}\rangle}$. If the (unique) isomorphism $U \cong V$ identifying $\mathbb{B}$-fixed subspaces with $\mathbb{B}$-fixed subspaces is chosen, the corresponding transpose symmetry gives an isomorphism between the list of candidate F_{101}'s the list of candidate F_{011}'s. Furthermore, such a transposition gives an involution of the set of $\mathbb{B}_T$-fixed F_{110}'s so that the application of the (210) test to F_{110} is equivalent to the application of the (120) test to its transpose. This symmetry may be observed in the three pair of equal numbers in the table above and will play a critical role in §7.

5. Explanation of the proofs of Theorems 1.1 and 1.2

The proofs of the theorems are achieved by a computer implementation of the ideal enumeration algorithm up to the (111) test to rule out any candidate ideals when $r = 16$ for each of $M_{\langle 3\rangle}$ and $\det_3$ (see §3). Each of $\det_3$ and $M_{\langle 3\rangle}$ has a reductive symmetry group, so candidate F_{110} families can be enumerated from the weight diagram of $T(C^*)^\perp$ as described in §2.5.2. In order to carry out these steps as described tractably, two additional ideas are needed.

The first is in the combinatorial part of the enumeration of the (110)-components. In §2.5.2, the $\mathbb{B}_T$-fixed subspaces of a G_T-module are first parameterized by an integer function on weights d_λ and then by a subproduct Y_{d_λ} of Grassmannians. In our case, we wish to enumerate $\mathbb{B}_T$-fixed $65 = 81 - 16$ dimensional subspaces of the 72-dimensional space $M = T(C^*)^\perp$. When $T = M_{\langle 3\rangle}$, there are 54 weight spaces of dimension one and nine weight spaces of dimension two, and for $T = \det_3$, there are nine weight spaces of dimension one, 18 weight spaces of dimension two, and nine weight spaces of dimension three. In either case, it is intractable to enumerate on a computer all assignments d_λ summing to 65 and consistent with these data.

Fortunately, there are additional linear inequalities one can derive from the weight diagram between the values d_λ which are necessary for $Y_{d_\lambda} \neq \emptyset$. For example, if in the weight diagram $x : M_\lambda \to M_\mu$ corresponds to a linear inclusion since any $\mathbb{B}$-fixed subspace S satisfies $x.S_\lambda \subset S_\mu$ we have

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

$d_\lambda = \dim S_\lambda \leq \dim S_\mu = d_\mu$. This reasoning can be generalized to any x, not necessarily an inclusion, by applying the rank nullity theorem. An arrow $x : M_\lambda \to M_\mu$ in the weight diagram restricts to an arrow $S_\lambda \to S_\mu$, and the rank nullity theorem implies $d_\lambda + \dim \ker x \leq d_\mu$. More generally, consider the map $\bigoplus_i x_i : M_\lambda \to \bigoplus_i M_{\mu_i}$, where μ_i ranges over any set of weights with arrows out of M_λ. For any $\mathbb{B}$-fixed S, this map restricts to $S_\lambda \to \bigoplus_i S_{\mu_i}$, and the rank nullity theorem implies $d_\lambda + \dim \ker(\bigoplus_i x_i) \leq \sum_i d_{\lambda_i}$. Dually, we can consider the sum of transpose maps $\bigoplus_i x_i^{\mathbf{t}} : M_\mu^* \to \bigoplus_i M_{\lambda_i}^*$, where λ_i ranges over any set of weights with arrows into M_μ. For any $\mathbb{B}$-fixed S this map restricts to $S_\mu^\perp \to \bigoplus_i S_{\lambda_i}^\perp$, and we obtain $\dim M_\mu - d_\mu + \dim \ker(\bigoplus x_i^{\mathbf{t}}) \leq \sum_i \dim M_{\lambda_i} - d_{\lambda_i}$.

The assignments d_λ can thus be restricted to lie in a particular explicit and computable rational polytope P determined by the weight diagram, integer points of which are sufficiently small in number to completely enumerate. One can efficiently enumerate the integer points of such a polytope by recursively fixing one coordinate at a time, stopping early when the corresponding cut of P is empty (checked by solving the corresponding linear program).

The second idea needed is in how to efficiently apply the (210) and (120) tests to parameterized families F_{110}. Concretely, this corresponds to finding the variety on which a 405×585 matrix has rank at most 389. Doing this by explicitly enumerating minors is intractable due to the combinatorially huge number. Since we only care about the variety set theoretically cut out by minors, we may arrange the computation in a manner more analogously with how one would find the rank of a constant matrix: using row reduction.

Given an $m \times n$ matrix M with entries in some polynomial ring, we wish to find the equations describing the set where M has rank at most r. First, generalize to matrix coefficients in some quotient of some ring of fractions of the polynomial ring, say R. If there is any matrix coefficient which is a unit in R, row reduce using it and pass to the problem of finding equations of an $m - 1 \times n - 1$ matrix having rank at most $r - 1$. Otherwise, heuristically pick a matrix coefficient f, for example, the most common nonzero entry, and recursively continue the computation in two cases which geometrically correspond to the terms in the decomposition of the target variety X as $(X \cap V(f)) \cup (X \setminus V(f))$. The case analyzing $X \cap V(f)$ algebraically corresponds to recursively continuing the computation with R replaced by $R/(f)$, and the case analyzing $X \setminus V(f)$ algebraically corresponds to recursively continuing the computation with R replaced by R_f. In both cases, progress is made, since in the first at least one entry is zeroed, and in the second at least one entry is made into a unit. Given the resulting ideals J_1 and J_2 from these cases, report our result as $J_1 \cap J_2$.

Carrying out the algorithm, one finds that the $\mathbb{B}_T$-fixed subspaces of dimension 65 in $T(C^*)^\perp$ sometimes occur in positive-dimensional families. The following table records the number of irreducible families of each dimension, those which pass the (210) test only, and those which pass both the (210) and (120) tests.

T	Dimension	$\mathbb{B}_T$-fixed	(210) test	(210) and (120) tests
$M_{\langle 3\rangle}$	0	132	53	8
	1	13	6	0
$\det_3$	0	342	54	4
	1	187	18	0
	2	44	0	0
	3	6	0	0

Remark 5.1. In the case of $M_{\langle 3\rangle}$, all families either entirely passed or entirely failed each of the (210) and (120) tests. In the case of $\det_3$, some families split into a number of smaller-dimensional families upon application of the tests. Two of the four final candidates for $\det_3$ started as isolated $\mathbb{B}_T$-fixed subspaces, and two are from one-dimensional families of $\mathbb{B}_T$-fixed subspaces.

To avoid redundant work we make use of the observation that both $M_{\langle 3\rangle}$ and $\det_3$ are invariant under cyclic permutation of the factors, so once we have the F_{110} candidates we automatically obtain the F_{101}

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

and F_{011} candidates. For each triple of candidates, in these cases with no remaining parameters, one checks that the (111) test is not passed, proving the theorems.

The module structure for matrix multiplication was discussed in §4. We now describe the relevant module structure for the determinant: Write $U, V = \mathbb{C}^m$ and $A_1 = \cdots = A_m = U \otimes V$. The determinant $\det_m$, considered as a tensor, spans the line $\Lambda^m U \otimes \Lambda^m V \subset A_1 \otimes \cdots \otimes A_m$. Explicitly, letting A_α have basis x^α_{ij},

$$\det_m = \sum_{\sigma,\tau\in\mathfrak{S}_m} \operatorname{sgn}(\sigma\tau) x^1_{\sigma(1)\tau(1)} \otimes \cdots \otimes x^m_{\sigma(m)\tau(m)}.$$

We will be concerned with the case $m = 3$, and we write $A_1 \otimes A_2 \otimes A_3 = A \otimes B \otimes C$. As a tensor, $\det_3$ is invariant under $(\mathrm{SL}(U) \times \mathrm{SL}(V)) \rtimes \mathbb{Z}_2$ as well as $\mathfrak{S}_3$. As an $\mathrm{SL}(U) \times \mathrm{SL}(V)$-module, $A \otimes B$ is $U^{\otimes 2} \otimes V^{\otimes 2} = S^2U \otimes S^2V \oplus S^2U \otimes \Lambda^2V \oplus \Lambda^2U \otimes S^2V \oplus \Lambda^2U \otimes \Lambda^2V$, and $\det_3(C^*) = \Lambda^2U \otimes \Lambda^2V$. As $\mathrm{SL}(U) \times \mathrm{SL}(V)$-modules, $\det_3(C^*)^\perp$ is the dual of the complement to $\det_3(C^*)$ in $A \otimes B$, and the weight diagram of $A \otimes B$ is the tensor product of the diagram in Example 2.7 with the same diagram for $V \otimes V$. Each of the three modules in the complement to $\det_3(C^*)$ in $A \otimes B$ are multiplicity free, but there are weight multiplicities up to three, for example, $u_1u_2 \otimes v_1v_2$, $u_1u_2 \otimes v_1 \wedge v_2$, and $u_1 \wedge u_2 \otimes v_1v_2$ each have weight $(\omega_2^U | \omega_2^V)$. Consequently, there are more and larger-dimensional $\mathbb{B}_T$-fixed subspaces, as observed in the table above.

For the code and further discussion of the implementation details, see the supplemental materials at github.com/adconner/chlbapolar.

6. Representation theory relevant for matrix multiplication

Theorems 1.4 and 1.5(1),(2) may also be proved using computer calculations, but we present hand-checkable proofs to both illustrate the power of the method and lay groundwork for future results. This section establishes the representation theory needed for those proofs.

6.1. *Refinement of the* (210) *test for matrix multiplication*

Recall $A = U^* \otimes V$, $B = V^* \otimes W$, $C = W^* \otimes U$ and $M_{\langle \mathbf{u},\mathbf{v},\mathbf{w}\rangle} = \mathrm{Id}_U \otimes \mathrm{Id}_V \otimes \mathrm{Id}_W$ and the notation ω_j for the fundamental $\mathfrak{sl}$-weights from §2.5. Let V_μ denote the irreducible $\mathfrak{sl}(V)$-module with highest weight μ. We have the following decompositions as $\mathrm{SL}(U) \times \mathrm{SL}(V)$-modules: (note $V_{\omega_2+\omega_{\mathbf{v}-1}}$ does not appear when $\mathbf{v} = 2$, and when $\mathbf{v} = 3$, $V_{\omega_2+\omega_{\mathbf{v}-1}} = V_{2\omega_2}$):

$$\begin{aligned}
\Lambda^2(U^* \otimes V) \otimes V^* &= (S^2U^* \otimes V_{\omega_1}) \oplus (\Lambda^2U^* \otimes V_{\omega_1}) \\
&\quad \oplus (S^2U^* \otimes V_{\omega_2+\omega_{\mathbf{v}-1}}) \oplus (\Lambda^2U^* \otimes V_{2\omega_1+\omega_{\mathbf{v}-1}}), && (9)\\
S^2(U^* \otimes V) \otimes V^* &= (S^2U^* \otimes V_{2\omega_1+\omega_{\mathbf{v}-1}}) \oplus (\Lambda^2U^* \otimes V_{\omega_2+\omega_{\mathbf{v}-1}}) \\
&\quad \oplus (S^2U^* \otimes V_{\omega_1}) \oplus (\Lambda^2U^* \otimes V_{\omega_1}), && (10)\\
A \otimes M_{\langle \mathbf{u},\mathbf{v},\mathbf{w}\rangle}(C^*) &= (U^* \otimes V) \otimes (U^* \otimes \mathrm{Id}_V \otimes W) \\
&= (S^2U^* \otimes V_{\omega_1} \otimes W) \oplus (\Lambda^2U^* \otimes V_{\omega_1} \otimes W), && (11)
\end{aligned}$$

$$\begin{aligned}
V \otimes \mathfrak{sl}(V) &= V_{\omega_1} \oplus V_{2\omega_1+\omega_{\mathbf{v}-1}} \oplus V_{\omega_2+\omega_{\mathbf{v}-1}}, && (12)\\
(U^* \otimes V) \otimes (U^* \otimes \mathfrak{sl}(V)) &= (S^2U^* \otimes V_{2\omega_1+\omega_{\mathbf{v}-1}}) \oplus (\Lambda^2U^* \otimes V_{2\omega_1+\omega_{\mathbf{v}-1}}) \\
&\quad \oplus (S^2U^* \otimes V_{\omega_1}) \oplus (\Lambda^2U^* \otimes V_{\omega_1}) \\
&\quad \oplus (S^2U^* \otimes V_{\omega_2+\omega_{\mathbf{v}-1}}) \oplus (\Lambda^2U^* \otimes V_{\omega_2+\omega_{\mathbf{v}-1}}). && (13)
\end{aligned}$$

These formulas follow from the following basic formulas: for any vector spaces U, V, one has the $\mathrm{GL}(U) \times \mathrm{GL}(V)$ decompositions $S^2(U \otimes V) = S^2U \otimes S^2V \oplus \Lambda^2U \otimes \Lambda^2V$; see, for example, [30, §2.7.1]

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

and the decomposition $\Lambda^2(U\otimes V) = S^2U\otimes\Lambda^2V \oplus S^2U\otimes\Lambda^2V$ is derived similarly, and one has the $\mathrm{GL}(U)$ decomposition $U\otimes U = S^2U \oplus \Lambda^2U$. Finally, the Pieri formula (see, e.g., [20, §6.1, eqns 6.8,6.9]) gives $S^2V\otimes V^* = V_{2\omega_1}\otimes V_{\omega_{\mathbf{v}-1}} = V_{2\omega_1+\omega_{\mathbf{v}-1}} \oplus V_{\omega_1}$ and $\Lambda^2V\otimes V^* = V_{\omega_2}\otimes V_{\omega_{\mathbf{v}-1}} = V_{\omega_2+\omega_{\mathbf{v}-1}} \oplus V_{\omega_1}$.

Note that V_{ω_1} is isomorphic to V and

$$\dim(V_{2\omega_1+\omega_{\mathbf{v}-1}}) = \tfrac{1}{2}\mathbf{v}^3 + \tfrac{1}{2}\mathbf{v}^2 - \mathbf{v}, \quad \dim(V_{\omega_2+\omega_{\mathbf{v}-1}}) = \tfrac{1}{2}\mathbf{v}^3 - \tfrac{1}{2}\mathbf{v}^2 - \mathbf{v}.$$

Proposition 6.1. *Write $E_{110} := M_{\langle \mathbf{u},\mathbf{v},\mathbf{w}\rangle}(C^*) \oplus E'_{110}$, where $E'_{110} \subset U^*\otimes\mathfrak{sl}(V)\otimes W$. The dimension of the kernel of the map (7) $E_{110}\otimes A \to \Lambda^2A\otimes B$ equals the dimension of the kernel of the skew symmetrization followed by projection map*

$$E'_{110}\otimes A \to S^2U^*\otimes V_{\omega_2+\omega_{\mathbf{v}-1}}\otimes W \oplus \Lambda^2U^*\otimes V_{2\omega_1+\omega_{\mathbf{v}-1}}\otimes W, \tag{14}$$

and the kernel of equation (14) is

$$(E'_{110}\otimes A) \cap [U^{*\otimes 2}\otimes V_{\omega_1}\otimes W \oplus S^2U^*\otimes V_{2\omega_1+\omega_{\mathbf{v}-1}}\otimes W \oplus \Lambda^2U^*\otimes V_{\omega_2+\omega_{\mathbf{v}-1}}\otimes W]. \tag{15}$$

Proof. Write M for the target of equation (14). We have the following commutative diagram, where horizontal arrows form exact sequences:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & (U^*\otimes \mathrm{Id}_V\otimes W)\otimes A & \longrightarrow & E_{110}\otimes A & \longrightarrow & E'_{110}\otimes A & \longrightarrow & 0\\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & (U^*)^{\otimes 2}\otimes V_{\omega_1}\otimes W & \longrightarrow & \Lambda^2(U^*\otimes V)\otimes V^*\otimes W & \longrightarrow & M & \longrightarrow & 0.
\end{array}$$

The bottom row reflects the decomposition (9) tensored with W. The middle vertical arrow is the skew symmetrization map (7), and since it is the restriction of a $\mathrm{GL}(U)\times\mathrm{GL}(V)\times\mathrm{GL}(W)$ equivariant map, by Schur's lemma, its submodule $(U^*\otimes \mathrm{Id}_V\otimes W)\otimes A = (U^*\otimes V)\otimes(U^*\otimes \mathrm{Id}_V\otimes W)$ must have image contained in $(U^*)^{\otimes 2}\otimes V_{\omega_1}\otimes W$. The induced right vertical arrow is the map (14).

We show the left vertical arrow is an isomorphism, from which the claim on the kernel dimension of equation (14) will follow by, for example, the snake lemma. We have the decomposition into irreducible modules

$$(U^*\otimes V)\otimes(U^*\otimes \mathrm{Id}_V\otimes W) = S^2U^*\otimes V\otimes \mathrm{Id}_V\otimes W \oplus \Lambda^2U^*\otimes V\otimes \mathrm{Id}_V\otimes W.$$

The vertical left arrow is an equivariant map, so by Schur's lemma, it is sufficient to see that a single vector in each of the modules on the right has nonzero image. We check the highest weight vectors:

$$\begin{aligned}
&(u^{\mathbf{u}}\otimes v_1)\otimes u^{\mathbf{u}}\otimes(\sum_j v_j\otimes v^j)\otimes w_1 \mapsto \sum_{\rho>1}(u^{\mathbf{u}}\otimes v_1)\wedge(u^{\mathbf{u}}\otimes v_\rho)\otimes v^\rho\otimes w_1, \text{ and}\\
&\left[(u^{\mathbf{u}}\otimes v_1)\otimes u^{\mathbf{u}-1} - (u^{\mathbf{u}-1}\otimes v_1)\otimes u^{\mathbf{u}}\right]\otimes(\sum_j v_j\otimes v^j)\otimes w_1 \mapsto\\
&\qquad\qquad \sum_j\left[(u^{\mathbf{u}}\otimes v_1)\wedge(u^{\mathbf{u}-1}\otimes v_j) - (u^{\mathbf{u}-1}\otimes v_1)\wedge(u^{\mathbf{u}}\otimes v_j)\right]\otimes v^j\otimes w_1
\end{aligned}$$

Now, equation (14) is a restriction of the surjective equivariant map $U^*\otimes\mathfrak{sl}(V)\otimes W\otimes A \to M$. Comparing modules in the irreducible decompositions of the source and target in view of equation (13), we obtain that equation (15) is the kernel of equation (14). □

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

6.2. $\underline{\mathbf{R}}(M_{\langle 2\rangle}) > 6$ *revisited*

In this case, the map (14) takes image in $\Lambda^2 U^* \otimes S^2 V \otimes V^* \otimes W$. We have the following images.

For the highest weight vector $x_1^2 \otimes y_1^2$ times the four basis vectors of A (with their $\mathfrak{sl}(V)$-weights in the second column), the image of equation (14) is spanned by

$$\begin{array}{ll} x_1^1 \wedge x_1^2 \otimes y_1^2 & 3\omega_1 \\ x_2^1 \wedge x_1^2 \otimes y_1^2 & \omega_1. \end{array}$$

(Note, e.g., $x_2^2 \otimes x_1^2 \otimes y_1^2$ maps to zero under the skew-symmetrization map as $u^2 \otimes u^2$ projects to zero in $\Lambda^2 U^*$.) For $A \otimes (x_1^2 \otimes y_1^1 - x_2^2 \otimes y_1^2)$ (the lowering of $x_1^2 \otimes y_1^2$ under $\mathfrak{sl}(V)$), the image is spanned by

$$\begin{array}{ll} x_1^1 \wedge (x_1^2 \otimes y_1^1 - x_2^2 \otimes y_1^2) & \omega_1 \\ x_2^1 \wedge (x_1^2 \otimes y_1^1 - x_2^2 \otimes y_1^2) & -\omega_1. \end{array}$$

Since W has nothing to do with the map, we don't need to compute the image of, for example, $A \otimes x_1^2 \otimes y_2^2$ to know its contribution to the kernel, as it must be the same dimension as that of $A \otimes x_1^2 \otimes y_1^2$, just with a different W-weight.

Were $\underline{\mathbf{R}}(M_{\langle 2\rangle}) = 6$, E'_{110} would have dimension two, spanned by the highest weight vector and one lowering of it, and in order to be a candidate, its image in $\Lambda^2 U^* \otimes S^3 V \otimes W$ would have to have dimension at most two. Taking $E'_{110} = \langle x_1^2 \otimes y_1^2, x_1^2 \otimes y_1^1 - x_2^2 \otimes y_1^2 \rangle$, the image of equation (14) has dimension three. Taking $E'_{110} = \langle x_1^2 \otimes y_1^2, x_1^2 \otimes y_2^2 \rangle$, the image of equation (14) has dimension four. Finally, taking $E'_{110} = \langle x_1^1 \otimes y_1^2, x_1^2 \otimes y_1^2 \rangle$, by transpose symmetry (see §4), the image of the (120)-version of equation (14) must have dimension four and we conclude.

7. Proofs of Theorems 1.5 and 1.6

7.1. Overview

To prove border rank lower bounds for a fixed tensor using border apolarity, one checks a list of candidates for components of a multigraded ideal. It is not immediate how to extend the technique to sequences of tensors in $\mathbf{n}$. Even in good situations such as in Theorems 1.5 and 1.6 where there are large Borel subgroups, candidate components can still occur in positive-dimensional families, and there is an exponential growth in $\mathbf{n}$ in the number of families to check. We overcome this problem by introducing several new ideas.

We restrict attention to only the (110)-graded ideal component and the application of the dual form of the (120) and (210) tests of §3. For each given candidate component, we forget everything about it except for the dimensions of certain internal weight spaces. We then analyze the kernels of Proposition 3.5 as sums of "local" contributions from each internal weight space. As we consider only dimension information, we determine upper bounds on the contributions. At this point, there are still many discrete cases of possible choices of these internal dimensions to consider. We use techniques from convex optimization to show that the relevant kernel contributions for any choice is no better than a constant more than that of a small fixed number of choices. We call this step the "globalization". These choices can then be completely analyzed as functions of $\mathbf{n}$.

7.2. Preliminaries

Recall that in these proofs $\mathbf{u} = \mathbf{w} = \mathbf{n}$ and $\mathbf{v}$ is 2 or 3. Let $E'_{110} \subset U^* \otimes \mathfrak{sl}(V) \otimes W$ be a $\mathbb{B}$-fixed subspace. In particular, E'_{110} is fixed under the torus of $\mathrm{GL}(U) \times \mathrm{GL}(W)$, so we may write $E'_{110} = \bigoplus_{s,t} u^{\mathbf{n}-s+1} \otimes X_{s,t} \otimes w_t$, where $X_{s,t} \subset \mathfrak{sl}(V)$. Since E'_{110} is fixed under the Borel subgroups of $\mathrm{GL}(U)$, $\mathrm{GL}(V)$ and $\mathrm{GL}(W)$, for each s and t we have

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

1. $X_{s,t} \subset \mathfrak{sl}(V)$ is $\mathbb{B}_V \subset \mathrm{GL}(V)$ fixed,
2. $X_{s,t} \supset X_{s+1,t}$ and
3. $X_{s,t} \supset X_{s,t+1}$.

Define the *outer structure* of E'_{110} to be the data

$$(s, t, \dim X_{s,t}), \quad 0 \le s, t \le \mathbf{n}.$$

Define the *inner structure at site* (s,t) to be $X_{s,t}$.

We may consider the outer structure of E'_{110} as an $\mathbf{n} \times \mathbf{n}$ grid, with each grid point (s,t) labelled by the dimension of the corresponding $X_{s,t}$. We will represent such filled grids by the corresponding Young diagrams on the nonzero labels so that the upper left box corresponds with the highest weight. Here, labels weakly decrease going to the right and down. It is reasonable to imagine such a filled Young diagram rotated 45° clockwise to put the highest weight at the top, as in the corresponding weight diagram (see Example 2.9, where $\mathbf{n} = \mathbf{v} = 2$).

In the case of $\mathfrak{sl}_2$, each $X_{s,t}$ is determined by its dimension, so an outer structure completely specifies a corresponding E'_{110}. In the case of $\mathfrak{sl}_3$, information about the particular inner structures is lost passing from E'_{110} to its outer structure.

Example 7.1. Here are three examples with $\mathbf{v} = 2$ and $\rho = 4$.

The diagram $\begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline 1 \\ \cline{1-1}\end{array}$ corresponds to

$$E'_{110} = \langle u^{\mathbf{n}} \otimes v^2 \otimes v_1 \otimes w_1, u^{\mathbf{n}} \otimes v^2 \otimes v_1 \otimes w_2, u^{\mathbf{n}} \otimes v^2 \otimes v_1 \otimes w_3, u^{\mathbf{n}-1} \otimes v^2 \otimes v_1 \otimes w_1 \rangle.$$

The diagram $\begin{array}{|c|c|}\hline 2 & 1 \\ \hline 1 \\ \cline{1-1}\end{array}$ corresponds to

$$E'_{110} = \langle u^{\mathbf{n}} \otimes v^2 \otimes v_1 \otimes w_1, u^{\mathbf{n}} \otimes (v^1 \otimes v_1 - v^2 \otimes v_2) \otimes w_1, u^{\mathbf{n}} \otimes v^2 \otimes v_1 \otimes w_2, u^{\mathbf{n}-1} \otimes v^2 \otimes v_1 \otimes w_1 \rangle.$$

The diagram $\begin{array}{|c|c|}\hline 3 & 1 \\ \hline\end{array}$ corresponds to

$$E'_{110} = \langle u^{\mathbf{n}} \otimes v^2 \otimes v_1 \otimes w_1, u^{\mathbf{n}} \otimes (v^1 \otimes v_1 - v^2 \otimes v_2) \otimes w_1, u^{\mathbf{n}} \otimes v^1 \otimes v_2 \otimes w_1, u^{\mathbf{n}} \otimes v^2 \otimes v_1 \otimes w_2 \rangle.$$

The transpose symmetry discussed in §4 maps $E'_{110} = \bigoplus_{s,t} u^{\mathbf{n}-s+1} \otimes X_{s,t} \otimes w_t$ to $\bigoplus_{s,t} u^{\mathbf{n}-t+1} \otimes X^{\mathbf{t}}_{s,t} \otimes w_s$, that is, the inner structure at site (s,t) becomes the transpose of the inner structure at site (t,s). In particular, transpose symmetry conjugates the diagram corresponding to the outer structure. In view of this symmetry, it is sufficient to study the (210) test only, for then everything we can say is also a statement about the (120) test under this transpose.

As mentioned above, we split the calculation of the kernel into a local and global computation. We bound the local contribution to the kernel at site (s,t) by a function of s, t and $\dim X_{s,t}$. Once this is done, the theorems are proved by solving the resulting combinatorial optimization problem over outer structures.

Recall the expression (15) and let K denote the term in brackets, that is,

$$\begin{aligned} K &= (U^*)^{\otimes 2} \otimes V_{\omega_1} \otimes W \oplus S^2 U^* \otimes V_{2\omega_1 + \omega_{\mathbf{v}-1}} \otimes W \oplus \Lambda^2 U^* \otimes V_{\omega_2 + \omega_{\mathbf{v}-1}} \otimes W \\ &\subset (U^*)^{\otimes 2} \otimes V \otimes \mathfrak{sl}(V) \otimes W. \end{aligned} \tag{16}$$

We may filter E'_{110} by $\mathbb{B}$-fixed subspaces such that each quotient corresponds to the inner structure contribution over some site (s,t). Call such a filtration *admissible*. Let $\Sigma_1 \subset \Sigma_2 \subset \cdots \subset \Sigma_f = E'_{110}$ be an admissible filtration, and put

$$K_g = (\Sigma_g \otimes A) \cap K. \tag{17}$$

Then the dimension of equation (15) may be written as the sum over g of $\dim (K_g / K_{g-1})$, and we may upper bound the dimension of equation (15) by upper bounding each $\dim (K_g / K_{g-1})$. We obtain bounds

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

on $\dim(K_g/K_{g-1})$ which depend only on s and $j = j_g := \dim(\Sigma_g/\Sigma_{g-1})$. For $\mathfrak{sl}_2$, this is Lemma 7.2, and for $\mathfrak{sl}_3$, this is Lemma 7.4. As discussed above, bounds on the kernel of the (120) map are obtained by symmetry; specifically, the bound is the same with s replaced by t.

Once these lemmas are established, the claims on fixed finite values of $\mathbf{n}$ may be immediately settled by enumerating the finitely many possible outer structures and checking that none gives a large enough kernel for both the (210) and (120) maps. The claims on infinite sequences of $\mathbf{n}$ require us to work more carefully, and we prove the required bounds on the solution to such problems parameterized by $\mathbf{n}$ in Lemma 7.7.

7.3. The local argument

Lemma 7.2. *Let* $\dim V = 2$, $\dim U = \mathbf{n}$. *Fix an admissible filtration such that* $\Sigma_g \subset E'_{110}$ *contains the* $\mathfrak{sl}(V)$*-subspace at site* (s,t) *and* Σ_{g-1} *does not. Write* j *for the dimension of the* $\mathfrak{sl}(V)$*-subspace at site* (s,t). *Then*

$$\dim(K_g/K_{g-1}) = a_j s + b_j,$$

where a_j, b_j *are the following functions of* j*:*

j	a_j	b_j
1	2	0
2	3	$\mathbf{n}$
3	4	$2\mathbf{n}$.

Lemma 7.2 is proved later this section.

Remark 7.3. Revisiting the proof that $\underline{\mathbf{R}}(M_{\langle 2\rangle}) > 6$ in this language, the possible outer structures of $\mathbb{B}$-fixed two planes are $\ydiagram{2}$, $\ydiagram{1,1}$, $\ydiagram{1,1,1}$, which, according to Lemma 7.2, have (210) map kernel dimensions $5 = 3(1)+2$, $6 = (1(2)+0)+(2(2)+0)$, and $4 = (1(2)+0)+(1(2)+0)$, respectively. The first and third are smaller than 6 and the choice of $\ydiagram{1,1}$ fails the (120) test by transpose symmetry. This gives our shortest proof that $\underline{\mathbf{R}}(M_{\langle 2\rangle}) > 6$.

Proof of Theorem 1.4. Here, we take $\mathbf{u} = 2$, $\mathbf{w} = 3$, $\mathbf{v} = 2$. We show that there is no E'_{110} of dimension $3 = 9-6$ passing the (210) and (120) tests. The possible outer structures are $\ydiagram{3}$, $\ydiagram{2,1}$, $\ydiagram{1,1,1,1}$ and $\ydiagram{2,1,1}$. Applying Lemma 7.2 with $\mathbf{n} = 2$, the corresponding (210) map kernel dimensions are eight, seven, six and nine, respectively, so only $\ydiagram{2,1}$ passes. However, $\ydiagram{2,1}$ has (120) kernel dimension eight and fails this test. □

Proof of Theorem 1.5(1),(2). For Theorem 1.5(1), $\mathbf{u} = \mathbf{w} = 3$, $\mathbf{v} = 2$. The outer structures corresponding to $13 - 9 = 4$ dimensional subpaces of $U^* \otimes \mathfrak{sl}(V) \otimes W$ are $\ydiagram{3,1,1}$, $\ydiagram{2,1,1}$, $\ydiagram{2,2,1}$, $\ydiagram{2,2}$, $\ydiagram{3,2}$, $\ydiagram{3,1}$, $\ydiagram{4,1}$. Of these, $\ydiagram{2,2,1}$, $\ydiagram{3,2}$, $\ydiagram{3,1}$, and $\ydiagram{4,1}$ pass the (210) test with kernel dimensions of size 14, 16, 15, and 14, respectively. However, none of these pass the (120) test as none appear in this list whose conjugate tableau also appear.

For Theorem 1.5(2), the result follows by similar complete enumeration of outer structures on a computer. □

Lemma 7.4. *Let* $\dim V = 3$, $\dim U = \mathbf{n}$. *Fix an admissible filtration such that* $\Sigma_g \subset E'_{110}$ *contains the* $\mathfrak{sl}(V)$*-subspace at site* (s,t) *and* Σ_{g-1} *does not. Write* j *for the dimension of the* $\mathfrak{sl}(V)$*-subspace at site* (s,t). *Then*

$$\dim(K_g/K_{g-1}) \le a_j s + b_j,$$

where a_j, b_j *are the following functions of* j*:*

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

j	a_j	b_j	j	a_j	b_j
1	3	-2	5	14	$\mathbf{n}$
2	6	0	6	17	$\mathbf{n}$
3	8	$\mathbf{n}$	7	21	$2\mathbf{n}-6$
4	11	$\mathbf{n}$	8	21	$3\mathbf{n}-6$.

In order to prove Lemmas 7.2 and 7.4, we first observe the following.

Proposition 7.5. *The included module $V_{\omega_1} \subset V\otimes\mathfrak{sI}(V)$ has weight basis*

$$\overline{v}_i := \sum_{j\neq i}[\mathbf{v}v_j\otimes(v_i\otimes v^j) - v_i\otimes(v_j\otimes v^j)] + (\mathbf{v}-1)v_i\otimes v_i\otimes v^i,\quad 1\leq i\leq \mathbf{v}.$$

Proof. The line $[\overline{v}_1]$ has weight $\omega_1 = \epsilon_1$ and is $\mathbb{B}$-stable, the lines $[\overline{v}_i]$ are lowerings of the line $[\overline{v}_1]$ and have weight ϵ_i. □

Proof of Lemmas 7.2 and 7.4. We begin in somewhat greater generality, not fixing $\mathbf{v} = \dim V$. We must bound $\dim K_g - \dim K_{g-1}$, where K_g is given by equation (17). Write $X \subset \mathfrak{sI}(V)$ for the inner structure at (s,t) so that $\Sigma_g = \Sigma_{g-1}\oplus u^{\mathbf{n}-s+1}\otimes X\otimes w_t$. Write

$$\begin{aligned}
V_0 &= \varnothing,\\
V_1 &= V_{\omega_1},\\
V_2 &= V_{\omega_1}\oplus V_{2\omega_1+\omega_{\mathbf{v}-1}},\\
V_3 &= V_{\omega_1}\oplus V_{2\omega_1+\omega_{\mathbf{v}-1}}\oplus V_{\omega_2+\omega_{\mathbf{v}-1}} = V\otimes\mathfrak{sI}(V).
\end{aligned}$$

Note that $V_2 = V_3$ when $\mathbf{v} = 2$. Then $\{V_f\}_f$ is a (partial) flag for $V\otimes\mathfrak{sI}(V)$, and

$$S_f := U^*\otimes U^{*(s-1)}\otimes V_3\otimes W + U^{*\otimes 2}\otimes V_f\otimes W + U^{*\otimes 2}\otimes V_3\otimes W_{(t-1)}$$

is a flag for $U^{*\otimes 2}\otimes V_3\otimes W$, where we have written $U^{*(s)} = \operatorname{span}\{u^{\mathbf{n}},\dots,u^{\mathbf{n}-s+1}\}$ and $W_{(t)} = \operatorname{span}\{w_1,\dots,w_t\}$. Hence, $S_f\cap K_g$ is a flag for K_g with $S_3\cap K_g = K_g$ and $S_0\cap K_g\subseteq K_{g-1}$. The fact that the inclusion $S_0\cap K_g\subseteq K_{g-1}$ may be strict is the only place in this argument we prove an inequality rather than equality. Use the isomorphism of quotient vector spaces

$$\frac{K_g\cap S_f}{K_g\cap S_{f-1}} = \frac{(K_g\cap S_f)+S_{f-1}}{S_{f-1}} \tag{18}$$

to obtain the successive quotients of $\{S_f\cap K_g\}_f$ as subspaces of

$$\frac{U^{*\otimes 2}\otimes V_3\otimes W}{S_{f-1}} = \frac{U^{*\otimes 2}}{U^*\otimes U^{*(s-1)}}\otimes\frac{V_3}{V_{f-1}}\otimes\frac{W}{W_{(t-1)}}. \tag{19}$$

Write K^f for the f-th summand of equation (16) so that $K\cap S_f = K^f + K\cap S_{f-1}$. Intersecting with $\Sigma_g\otimes A$ and adding S_{f-1}, we obtain

$$\begin{aligned}
K_g\cap S_f + S_{f-1} &= (K^f + S_{f-1})\cap(\Sigma_g\otimes A) + S_{f-1}\\
&= (K^f + S_{f-1})\cap(U^*\otimes u^{\mathbf{n}-s+1}\otimes V\otimes X\otimes w_t + S_{f-1}).
\end{aligned}$$

We may now pass in each side of the intersection to the right-hand side of equation (19), after which the intersection may be computed term by term. To compute the intersection in the $U^{*\otimes 2}/(U^*\otimes U^{*(s-1)})$ term, momentarily write $\overline{Z} = Z + U^*\otimes U^{*(s-1)}$ for $Z\in U^{*\otimes 2}$ and observe that

$$\overline{S^2U^*}\cap\overline{U^*\otimes u^{\mathbf{n}-s+1}} = \overline{U^{*s}\otimes u^{\mathbf{n}-s+1}},\ \text{and}\ \overline{\Lambda^2U^*}\cap\overline{U^*\otimes u^{\mathbf{n}-s+1}} = \overline{U^{*(s-1)}\otimes u^{\mathbf{n}-s+1}}.$$

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

Therefore, the right-hand side of equation (18) may be written, for $f = 1, 2$ and 3 respectively,

$$U^* \otimes (u^{\mathbf{n}-s+1} + U^{*(s-1)}) \otimes [(V \otimes X) \cap V_1] \otimes (w_t + W_{(t-1)})$$
$$U^{*s} \otimes (u^{\mathbf{n}-s+1} + U^{*(s-1)}) \otimes [(V \otimes X + V_1) \cap V_2] \otimes (w_t + W_{(t-1)})$$
$$U^{*(s-1)} \otimes (u^{\mathbf{n}-s+1} + U^{*(s-1)}) \otimes [V \otimes X + V_2] \otimes (w_t + W_{(t-1)}).$$

Write

$$Y = (V \otimes X) \cap V_1,$$
$$Y' = ((V \otimes X + V_1) \cap V_2)/V_1, \text{ and}$$
$$Y'' = (V \otimes X + V_2)/V_2,$$

and write their dimensions, respectively, as $\mathbf{y}, \mathbf{y}', \mathbf{y}''$. We obtain

$$\dim K_g = \dim (S_0 \cap K_g) + \mathbf{yn} + \mathbf{y}'s + \mathbf{y}''(s-1) \leq \dim K_{g-1} + \mathbf{yn} + \mathbf{y}'s + \mathbf{y}''(s-1),$$

the sum of the successive quotient dimensions of $\{S_f \cap K_g\}_f$.

Thus, when $j = \mathbf{v}^2 - 1$, that is, $X = \mathfrak{sl}(V)$, the desired result follows from $\mathbf{y} = \mathbf{v}$, $\mathbf{y}' = \dim V_{2\omega_1+\omega_{\mathbf{v}-1}}$ and $\mathbf{y}'' = \dim V_{\omega_2+\omega_{\mathbf{v}-1}}$.

In all cases, Y has a basis consisting of weight vectors and is closed under raising operators. Hence, by Proposition 7.5, $Y = \operatorname{span}\{\overline{v}_i \mid i \leq \mathbf{y}\}$.

Consider the case $j = \mathbf{v}^2 - 2$, that is X is the span of all weight vectors of $\mathfrak{sl}(V)$ except $v_{\mathbf{v}} \otimes v^1$. Then $\overline{v}_{\mathbf{v}}$ is not an element of Y because in the monomial basis, the monomial $v_1 \otimes (v_{\mathbf{v}} \otimes v^1)$ fails to have a nonzero coefficient in any element of Y. Hence, $\mathbf{y} \leq \mathbf{v} - 1$, and the trivial $\mathbf{y}' \leq \dim V_{2\omega_1+\omega_{\mathbf{v}-1}}$, and $\mathbf{y}'' \leq \dim V_{\omega_2+\omega_{\mathbf{v}-1}}$ give the asserted upper bounds.

By similar reasoning, when $\mathbf{v} = 3$, considering Example 2.6, we obtain the bounds $\mathbf{y} = 0$ when $j = 1, 2$ and $\mathbf{y} \leq 1$ when $j = 3, 4, 5, 6$. For all values of j except 1, the result then follows from

$$\begin{aligned} \dim K_g - \dim K_{g-1} &\leq (j\mathbf{v} - \mathbf{y})s + \mathbf{yn} - \mathbf{y}'' \\ &\leq (j\mathbf{v} - \mathbf{y})s + \mathbf{yn}, \end{aligned} \tag{20}$$

as $\mathbf{y} + \mathbf{y}' + \mathbf{y}'' = j\mathbf{v}$. The upper bound for $\mathbf{v} = 2$, $j = 1$, is settled similarly.

We must argue more for the $j = 1$ upper bound for $\mathbf{v} = 3$, namely that $\mathbf{y}'' \geq 2$. For this, consider

$$V \otimes \mathfrak{sl}(V) \oplus V_{\omega_1} = V \otimes V \otimes V^* = S^2V \otimes V^* \oplus \Lambda^2 V \otimes V^* \text{ and } \Lambda^2 V \otimes V^* = V_{\omega_2+\omega_{\mathbf{v}-1}} \oplus V_{\omega_1}.$$

Because we have $\mathbf{y} = 0$, the dimension $\mathbf{y}''$ of the projection of $V \otimes X$ onto $V_{\omega_2+\omega_{\mathbf{v}-1}}$ is the same as that onto $\Lambda^2 V \otimes V^*$. We have the images $v_2 \wedge v_1 \otimes v^3$ and $v_3 \wedge v_1 \otimes v^3$ of $v_2 \otimes v_1 \otimes v^3$ and $v_3 \otimes v_1 \otimes v^3$, respectively, whence $\mathbf{y}'' \geq 2$ as required.

To see the upper bounds in the $\mathbf{v} = 2$ cases are sharp, note that in this case $V_{\omega_2+\omega_{\mathbf{v}-1}} = \varnothing$, so $\mathbf{y}'' = 0$. The $j = 1$ case is thus automatic from equation (20), and for $j = 2$, we must show $\mathbf{y} \geq 1$. In this case, however, we have $\overline{v}_1 = 2v_2 \otimes (v_1 \otimes v^2) + v_1 \otimes (v_1 \otimes v^1 - v_2 \otimes v^2) \in V \otimes X$, as required. □

Remark 7.6. Although the bounds are essentially sharp when one assumes nothing about previous sites (σ, t) for $\sigma < s$, with knowledge of them one can get a much sharper estimate, although it is more complicated to implement the local/global principle. For example, if we are at a site (s, t) with $\mathbf{v} = 3$, $j = 1$ and for (σ, t) with $\sigma < t$ one also has $j = 1$, then the new contribution at site (s, t) is just s, not $3s - 2$.

In Lemma 7.7 below, the linear functions of s in Lemmas 7.2 and 7.4 appear as $a_{\mu_{s,t}}s + b_{\mu_{s,t}}$.

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

7.4. The globalization

Write μ for a Young diagram filled with nonnegative integer labels. The label in position (s,t) is denoted $\mu_{s,t}$, and sums over s,t are to be taken over the boxes of μ. As before, each μ will correspond to a possible outer structure.

We remark that the lemmas in this section and the next may be used for $M_{\langle \mathbf{mnn} \rangle}$ for any $\mathbf{n} \geq \mathbf{m}$.

The following lemma allows us to reduce from considering all possible outer structures and the corresponding bounds on the dimension of the kernels of the (210) and (120) tests to considering three (resp. eight) possible kernel dimensions in the case of $\mathbf{v} = 2$ (resp. $\mathbf{v} = 3$).

Lemma 7.7. *Fix $k \in \mathbb{N}$, $0 \leq a_1 \leq \cdots \leq a_k$, and $b_i \in \mathbb{R}$, $1 \leq i \leq k$. Let μ be a Young diagram filled with labels in the set $\{1, \ldots, k\}$, nonincreasing in rows and columns. Write $\rho = \sum_{s,t} \mu_{s,t}$. Then*

$$\min\left\{ \sum_{s,t} a_{\mu_{s,t}} s + b_{\mu_{s,t}}, \sum_{s,t} a_{\mu_{s,t}} t + b_{\mu_{s,t}} \right\} \leq \max_{1 \leq j \leq k} \left\{ \frac{a_j \rho^2}{8j^2} + (a_j + b_j)\frac{\rho}{j} \right\}. \tag{21}$$

Remark 7.8. The bound in the lemma is nearly tight. Taking μ to be a balanced hook filled with j makes the left-hand side equal $\frac{a_j}{8}(\frac{\rho^2}{j^2} - 1) + (a_j + b_j)\frac{\rho}{j}$. Hence, for any fixed ρ, a_i, b_i, the maximum of the left-hand side is within $\frac{1}{8}\max_j a_j$ of the right hand side.

Lemma 7.7 is proved in §7.5.

Proof of Theorem 1.5(3). Let $E'_{110} \subset U^* \otimes \mathfrak{sl}(V) \otimes W$ be a $\mathbb{B}$-fixed subspace, and let μ be the corresponding outer structure. We apply Lemma 7.7 with $k = 3$ and a_i and b_i from Lemma 7.2 to obtain an upper bound on the smaller of the kernel dimensions of the (120) and (210) maps. The resulting upper bound is $\max\{\frac{1}{4}\rho^2 + 2\rho, \frac{3}{32}\rho^2 + \frac{3+\mathbf{n}}{2}\rho, \frac{1}{18}\rho^2 + \frac{4+2\mathbf{n}}{3}\rho\}$.

Fix $\epsilon > 0$. We must show that if $\rho = (3\sqrt{6} - 6 - \epsilon)\mathbf{n}$, then each of $\frac{1}{4}\rho^2 + 2\rho$, $\frac{3}{32}\rho^2 + \frac{3+\mathbf{n}}{2}\rho$, and $\frac{1}{18}\rho^2 + \frac{4+2\mathbf{n}}{3}\rho$ is strictly smaller than $\mathbf{n}^2 + \rho$. Substituting and solving for $\mathbf{n}$, we obtain that this holds for the last expression when

$$\mathbf{n} > \frac{6}{\epsilon}\frac{3\sqrt{6} + 6 - \epsilon}{6\sqrt{6} - \epsilon},$$

and when $\epsilon < \frac{1}{4}$, this condition implies the other two inequalities. □

Proof of Theorem 1.6. Proceeding in the same way as in the proof of Theorem 1.5(3), we apply Lemma 7.7 with μ the outer structure corresponding to an arbitrary $\mathbb{B}$-fixed subspace $E'_{110} \subset U^* \otimes \mathfrak{sl}(V) \otimes W$, $k = 8$, and a_i and b_i corresponding to the inner structure contribution upper bounds obtained in Lemma 7.4. We obtain the smaller of the kernel dimensions of the (120) and (210) maps is at most the largest of the following,

j	Lemma 7.7	j	Lemma 7.7
1	$\frac{3}{8}\rho^2 + \rho$	5	$\frac{7}{100}\rho^2 + \frac{14+\mathbf{n}}{5}\rho$
2	$\frac{3}{16}\rho^2 + \frac{6}{2}\rho$	6	$\frac{17}{288}\rho^2 + \frac{17+\mathbf{n}}{6}\rho$
3	$\frac{1}{9}\rho^2 + \frac{8+\mathbf{n}}{3}\rho$	7	$\frac{3}{56}\rho^2 + \frac{15+2\mathbf{n}}{7}\rho$
4	$\frac{11}{128}\rho^2 + \frac{11+\mathbf{n}}{4}\rho$	8	$\frac{21}{512}\rho^2 + \frac{15+3\mathbf{n}}{8}\rho$.

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

Now, if one takes $\rho = \lfloor \sqrt{\frac{8}{3}}\mathbf{n} \rfloor$, the kernel upper bound for each j is strictly less than $\mathbf{n}^2 + \rho$. This follows for $j = 1$ because $\sqrt{\frac{8}{3}}\mathbf{n}$ is irrational. This follows for $2 \leq j \leq 8$ because $\mathbf{n} \geq 18$. Hence, at least one of the kernels of the (120) and (210) maps is too small, and $\underline{\mathbf{R}}(M_{\langle 3\mathbf{nn}\rangle}) > \mathbf{n}^2 + \rho$, as required. □

7.5. Proof of Lemma 7.7

We will reduce Lemma 7.7 to the following, which may be viewed as a continuous reformulation. Its proof depends on a delicate perturbation argument.

Lemma 7.9. *Fix $k \in \mathbb{N}$, $c_i \geq 0$, $d_i \in \mathbb{R}$, for $1 \leq i \leq k$. Write $C_j = \sum_{i=1}^{j} c_i$ and $D_j = \sum_{i=1}^{j} d_i$. For all choices of x_i, y_j satisfying the constraints $x_1 \geq \cdots \geq x_k \geq 0$, $y_1 \geq \cdots \geq y_k \geq 0$, and $\sum_i x_i + y_i = \rho$, the following inequality holds:*

$$\min\left\{\sum_{i\leq k} c_i x_i^2 + d_i(x_i + y_i), \sum_{i\leq k} c_i y_i^2 + d_i(x_i + y_i)\right\} \leq \max_{1\leq j\leq k}\left\{\frac{\rho^2}{4j^2}C_j + \frac{\rho}{j}D_j\right\}. \tag{22}$$

Remark 7.10. The inequality is tight. Choose j so that the maximum on the right-hand side is achieved. Then equality is achieved when $x_1 = \cdots = x_j = y_1 = \cdots = y_j = \frac{\rho}{2j}$ and $x_s, y_s = 0$ for $s > j$.

Proof. As both the left- and right-hand sides are continuous in the c_i, it suffices to prove the result under the assumption $c_i > 0$. The idea of the proof is the following: Any choice of x_i and y_i which has at least two degrees of freedom inside its defining polytope can be perturbed in such a way that the local linear approximations to the two polynomials on the left -hand side do not decrease; that is, two closed half planes in $\mathbb{R}^2$ containing $(0, 0)$ also intersect aside from $(0, 0)$. Each polynomial on the left strictly exceeds its linear approximation at any point, and thus one can strictly improve the left-hand side with a perturbation. The case of at most one degree of freedom is settled directly.

Write $x_{k+1} = y_{k+1} = 0$, and define $x_i' = x_i - x_{i+1}$ and $y_i' = y_i - y_{i+1}$ so that $x_i = \sum_{j=i}^{k} x_j'$ and $y_i = \sum_{j=i}^{k} y_j'$. Then $x_i', y_i' \geq 0$ and $\sum_{i=1}^{k} i(x_i' + y_i') = \rho$. Suppose at least three of the x_i', y_j' are nonzero, we will show the expression on the left-hand side of equation (22) is not maximal. Write three of the nonzero x_i', y_j' as $\overline{x}, \overline{y}, \overline{z}$. Replace them by $\overline{x} + \epsilon_1$, $\overline{y} + \epsilon_2$, $\overline{z} + \epsilon_3$, with the ϵ_i to be determined. This will preserve the equation $\sum_i x_i + y_i = \rho$ only if $\epsilon_1 + \epsilon_2 + \epsilon_3 = 0$, so we require this. Substitute these values into

$$E_L := \sum_{i\leq k} c_i x_i^2 + d_i(x_i + y_i) \text{ and } E_R := \sum_{i\leq k} c_i y_i^2 + d_i(x_i + y_i).$$

View E_L, E_R as polynomial expressions in the ϵ_j. Then

$$E_L = \sum_i c_i S_{L,i}^2 + L_L + d,\ E_R = \sum_i c_i S_{R,i}^2 + L_R + d,$$

where $S_{L,i}, S_{R,i}$ and L_L, L_R are linear forms in the ϵ_i, and $d \in \mathbb{R}$. Each $S_{L,i}, S_{R,i}$ is a sum of some subset of the ϵ_i, and the union of them span the 2-plane $\langle \epsilon_1, \epsilon_2, \epsilon_3\rangle / \langle \sum \epsilon_j = 0\rangle$. Consider the linear map $T = L_L \oplus L_R : \langle \epsilon_1, \epsilon_2, \epsilon_3\rangle / \langle \sum \epsilon_j = 0\rangle \to \mathbb{R}^2$. If T is nonsingular, then for any $\epsilon > 0$, there are constants $\overline{\epsilon}_j$, with $\sum \overline{\epsilon}_j = 0$ so that $T(\overline{\epsilon}_1, \overline{\epsilon}_2, \overline{\epsilon}_3) = (\epsilon, \epsilon)$, and it is possible to choose ϵ so that $\overline{x} + \overline{\epsilon}_1, \overline{y} + \overline{\epsilon}_2, \overline{z} + \overline{\epsilon}_3 \geq 0$. Then this new assignment strictly improves the old one. Otherwise, if T is singular, then there is an admissible $(\overline{\epsilon}_1, \overline{\epsilon}_2, \overline{\epsilon}_3) \neq 0$ in the kernel of T, where again we may assume the same nonnegativity condition. The corresponding assignment does not change L_L, L_R, but as the $S_{L,i}, S_{R,i}$ span the linear forms, at least one them is nonzero. Consequently, at least one of the modified E_L, E_R is strictly larger after the perturbation, and neither is smaller. If, say, only E_L is strictly larger, and $x_i' > 0$, we may substitute $x_i' - \epsilon$ and $y_i' + \epsilon$ for x_i' and y_i' for some $\epsilon > 0$ to make both E_L and E_R strictly larger.

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

Thus, the left-hand side is maximized at an assignment where at most two of x_i' and y_i' are nonzero. It is clear that at least one of each of x_i' and y_i' must be nonzero, so there is exactly one of each, say $x_s' = \alpha$ and $y_t' = \beta$. It is clear at the maximum that $\sum_{i\le k} c_i x_i^2 + d_i(x_i + y_i) = \sum_{i\le k} c_i y_i^2 + d_i(x_i + y_i)$, from which it follows that $\alpha^2 C_s = \sum_{i\le k} c_i x_i^2 = \sum_{i\le k} c_i y_i^2 = \beta^2 C_t$ and $\alpha\sqrt{C_s} = \beta\sqrt{C_t}$. We also have $s\alpha + t\beta = \rho$. Notice that

$$\alpha = \frac{\rho\sqrt{C_t}}{s\sqrt{C_t} + t\sqrt{C_s}}, \quad \beta = \frac{\rho\sqrt{C_s}}{s\sqrt{C_t} + t\sqrt{C_s}}$$

satisfy the equations so that the optimal value obtained is

$$\sum_{i\le k} c_i x_i^2 + d_i(x_i + y_i) = \alpha^2 C_s + \alpha D_s + \beta D_t = \frac{\rho}{s\sqrt{C_t} + t\sqrt{C_s}}\left(\frac{\rho C_s C_t}{s\sqrt{C_t} + t\sqrt{C_s}} + \sqrt{C_t}D_s + \sqrt{C_s}D_t\right).$$

By the arithmetic mean-harmonic mean inequality, we have

$$\frac{\rho C_s C_t}{s\sqrt{C_t} + t\sqrt{C_s}} = \frac{\rho}{\frac{s}{C_s\sqrt{C_t}} + \frac{t}{C_t\sqrt{C_s}}} \le \frac{\rho}{4}\left[\frac{C_s\sqrt{C_t}}{s} + \frac{C_t\sqrt{C_s}}{t}\right]$$

so that

$$\begin{aligned}
\frac{\rho C_s C_t}{s\sqrt{C_t} + t\sqrt{C_s}} + \sqrt{C_t}D_s + \sqrt{C_s}D_t &\le \frac{\rho}{4}\left[\frac{C_s\sqrt{C_t}}{s} + \frac{C_t\sqrt{C_s}}{t}\right] + \sqrt{C_t}D_s + \sqrt{C_s}D_t \\
&= \frac{s\sqrt{C_t} + t\sqrt{C_s}}{\rho}\left[\frac{s\alpha}{\rho}\left(\frac{\rho^2}{4s^2}C_s + \frac{\rho}{s}D_s\right) + \frac{t\beta}{\rho}\left(\frac{\rho^2}{4t^2}C_t + \frac{\rho}{t}D_t\right)\right] \\
&\le \frac{s\sqrt{C_t} + t\sqrt{C_s}}{\rho}\max\left\{\frac{\rho^2}{4s^2}C_s + \frac{\rho}{s}D_s, \frac{\rho^2}{4t^2}C_t + \frac{\rho}{t}D_t\right\},
\end{aligned}$$

with the last inequality holding because $\frac{s\alpha}{\rho} + \frac{t\beta}{\rho} = 1$. Multiplying both sides by $\frac{\rho}{s\sqrt{C_t}+t\sqrt{C_s}}$, we conclude. □

We prove one final lemma on partitions that will enable the reduction of Lemma 7.7 to an instance of Lemma 7.9.

For a partition $\lambda = (\lambda_1, \dots, \lambda_q)$, write $\ell(\lambda) = q$ and define

$$n(\lambda) := \sum_i (i-1)\lambda_i.$$

Let λ' denote the conjugate partition.

Lemma 7.11. *Let λ be a partition not of the form $(|\lambda| - 2, 2)$. Then $n(\lambda) \le \frac{1}{8}(|\lambda| + \lambda_1' - \lambda_1)^2 - \frac{1}{8}$. In particular, for all λ, $n(\lambda) \le \frac{1}{8}(|\lambda| + \lambda_1' - \lambda_1)^2$.*

Proof. We prove the result by induction on $\lambda_1 = \ell(\lambda')$. When $\ell(\lambda') = 1$, we have

$$n(\lambda) = \binom{\lambda_1'}{2} = \tfrac{1}{2}(\lambda_1' - \tfrac{1}{2})^2 - \tfrac{1}{8} = \tfrac{1}{8}(|\lambda| + \lambda_1' - \lambda_1)^2 - \tfrac{1}{8},$$

as required. Now, assume $k = \ell(\lambda') > 1$. Write μ for the partition where $\ell(\mu') = k - 1$ and $\mu_i' = \lambda_i'$, $i \le k - 1$. If $\lambda = (3, 3)$, we are done by direct calculation; hence, otherwise we may assume the result holds for μ by the induction hypothesis.

$$\begin{aligned}
n(\lambda) &= n(\mu) + \binom{\lambda_k'}{2} \\
&\le \tfrac{1}{8}(|\mu| + \mu_1' - \mu_1)^2 - \tfrac{1}{8} + \binom{\lambda_k'}{2}
\end{aligned}$$

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

$$= \tfrac{1}{8}(|\lambda| - \lambda'_k + \lambda'_1 - (\lambda_1 - 1))^2 - \tfrac{1}{8} + \tfrac{1}{2}\lambda'_k(\lambda'_k - 1)$$
$$= \tfrac{1}{8}(|\lambda| + \lambda'_1 - \lambda_1)^2 - \tfrac{1}{8} - \tfrac{1}{4}(|\lambda| + \lambda'_1 - \lambda_1 - \tfrac{5}{2}\lambda'_k + \tfrac{1}{2})(\lambda'_k - 1)$$

We must show $\frac{1}{4}(|\lambda| + \lambda'_1 - \lambda_1 - \frac{5}{2}\lambda'_k + \frac{1}{2})(\lambda'_k - 1) \geq 0$. If $\lambda'_k = 1$, this is immediate; otherwise, we show the first factor is nonnegative. We have $|\lambda| - \lambda_1 \geq k\lambda'_k - k$, so

$$|\lambda| + \lambda'_1 - \lambda_1 - \tfrac{5}{2}\lambda'_k + \tfrac{1}{2} \geq (\lambda'_1 - \lambda'_k) + \tfrac{2k-3}{2}(\lambda'_k - 1) - 1.$$

If $k = 2$, then by assumption $\lambda'_1 \geq 3$, and considering separately the cases $\lambda'_2 = 2$ and $\lambda'_2 \geq 3$ yields the result. Otherwise $k \geq 3$, and because $\lambda'_k \geq 2$, we again conclude. □

Proof of Lemma 7.7. For each $1 \leq i \leq k$, let λ^i be the partition corresponding to the boxes of μ labeled $\geq i$. Write $a_0 = b_0 = 0$. Then

$$\begin{aligned}
\textstyle\sum_{s,t} a_{\mu_{s,t}} s + b_{\mu_{s,t}} &= \textstyle\sum_{s,t} \sum_{i=1}^{\mu_{s,t}} (a_i - a_{i-1})s + b_i - b_{i-1} \\
&= \textstyle\sum_{i=1}^{k} \sum_{s,t \in \lambda^i} (a_i - a_{i-1})s + b_i - b_{i-1} \\
&= \textstyle\sum_{i=1}^{k} (a_i - a_{i-1})n(\lambda^i) + (a_i - a_{i-1} + b_i - b_{i-1})|\lambda^i| \\
&\leq \textstyle\sum_{i=1}^{k} \left[\frac{1}{2}(a_i - a_{i-1})\right] \left(\frac{1}{2}(|\lambda^i| + (\lambda^i)'_1 - \lambda^i_1)\right)^2 + [a_i - a_{i-1} + b_i - b_{i-1}]\,|\lambda^i| \quad (23)
\end{aligned}$$

where we have used Lemma 7.11 to obtain the last inequality. Set

$$\begin{aligned}
c_i &= \tfrac{1}{2}(a_i - a_{i-1}) \\
d_i &= a_i - a_{i-1} + b_i - b_{i-1} \\
x_i &= \tfrac{1}{2}(|\lambda^i| + (\lambda^i)'_1 - \lambda^i_1) \\
y_i &= \tfrac{1}{2}(|\lambda^i| - (\lambda^i)'_1 + \lambda^i_1).
\end{aligned}$$

Then equation (23) becomes

$$\sum_{i=1}^{k} c_i x_i^2 + d_i(x_i + y_i).$$

Similarly, $\sum_{s,t} a_{\mu_{s,t}} t + b_{\mu_{s,t}} \leq \sum_{i=1}^{k} c_i y_i^2 + d_i(x_i + y_i)$. Now, $\sum_i x_i + y_i = \sum_i |\lambda^i| = \rho$ and the x_i and y_i are each nonnegative and nonincreasing. Hence, by Lemma 7.9,

$$\min\left\{\sum_{s,t} a_{\mu_{s,t}} s + b_{\mu_{s,t}}, \sum_{s,t} a_{\mu_{s,t}} t + b_{\mu_{s,t}}\right\} = \max_{1 \leq j \leq k}\left\{\frac{a_j \rho^2}{8j^2} + (a_j + b_j)\frac{\rho}{j}\right\},$$

as required. □

Acknowledgements. This project began when the third author visited Institute of Mathematics of Polish Academy of Sciences (IMPAN), which he thanks for their hospitality, excellent environment for research and support. We are deeply indebted to Buczyńska and Buczyński for sharing their preprint with us and discussions. We thank Mateusz Michałek for insight in settling Lemma 7.9 and Rafael Oliveira and anonymous referees for significant help with the exposition.

Funding statement. Conner supported by NSF grant 2002149. Landsberg supported by NSF grants AF-1814254 and AF-2203618 and the grant 346300 for IMPAN from the Simons Foundation and the matching 2015–2019 Polish MNiSW fund as well as a Simons Visiting Professor grant.

Competing interest. The authors have no competing interest to declare.

References

[1] J. Alman, *Limits on the Universal Method for Matrix Multiplication*, 34th Computational Complexity Conference, LIPIcs. Leibniz Int. Proc. Inform., vol. 137 (Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2019), Art. No. 12, 24. MR 3984617

[2] J. Alman and V. Vassilevska Williams, *Further Limitations of the Known Approaches for Matrix Multiplication*, 9th Innovations in Theoretical Computer Science Conference, ITCS 2018, January 11–14, 2018 (Cambridge, MA, 2018), **25**:1–25: 15.

[3] J. Alman and V. Vassilevska Williams, *Limits on All Known (and Some Unknown)* Approaches to Matrix Multiplication, 59th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2018 (IEEE Computer Soc., Los Alamitos, CA, 2018), 580–591. MR 3899623

[4] A. Ambainis, Y. Filmus and F. Le Gall, *Fast Matrix Multiplication: Limitations of the Coppersmith–Winograd Method (Extended Abstract)*, STOC'15—Proceedings of the 2015 ACM Symposium on Theory of Computing (ACM, New York, 2015), 585–593. MR 3388238

[5] S. Arora and B. Barak, *Computational Complexity. A Modern Approach.* (Cambridge University Press, Cambridge, 2009). MR 2500087 (2010i,68001)

[6] D. Bini, 'Relations between exact and approximate bilinear algorithms', *Applications*, *Calcolo* **17**(1) (1980), 87–97. MR 605920 (83f,68043b)

[7] D. Bini, G. Lotti and F. Romani, 'Approximate solutions for the bilinear form computational problem', *SIAM J. Comput.* **9**(4) (1980), 692–697. MR MR592760 (82a:68065)

[8] M. Bläser, Fast Matrix Multiplication, Graduate Surveys, no. 5, Theory of Computing Library, 2013.

[9] M. Boij and Z. Teitler, 'A bound for the Waring rank of the determinant via syzygies', *Linear Algebra and its Applications* **587** (2020), 195 – 214.

[10] N. Bourbaki, *Lie Groups and Lie Algebras,* Chapters 4–6, Elements of Mathematics (Berlin) (Springer-Verlag, Berlin, 2002). Translated from the 1968 French original by Andrew Pressley. MR MR1890629 (2003a:17001)

[11] W. Buczyńska and J. Buczyński, 'Apolarity, border rank, and multigraded Hilbert scheme', *Duke Math. J.* **170**(16) (2021), 3659–3702. MR 4332674

[12] P. Bürgisser, M. Clausen and M. A. Shokrollahi, *Algebraic Complexity Theory*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 315 (Springer-Verlag, Berlin, 1997). With the collaboration of Thomas Lickteig. MR 99c:68002

[13] M. Christandl, P. Vrana and J. Zuiddam, *Barriers for Fast Matrix Multiplication from Irreversibility*, 34th Computational Complexity Conference, LIPIcs. Leibniz Int. Proc. Inform., vol. 137 (Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2019), Art. No. 26, 17. MR 3984631

[14] A. Conner, F. Gesmundo, J. M. Landsberg and E. Ventura, 'Tensors with maximal symmetries', Preprint, 2019, arXiv:1909.09518.

[15] A. Conner, F. Gesmundo, J. M. Landsberg and E. Ventura, 'Rank and border rank of Kronecker powers of tensors and Strassen's laser method', *Comput. Complexity* **31**(1) (2022), Paper No. 1, 40. MR 4356227

[16] A. Conner, H. Huang and J. M. Landsberg, 'Bad and good news for Strassen's laser method: Border rank of $Perm_3$ and strict submultiplicativity', *Foundations of Computational Mathematics* (2022).

[17] H. Derksen, 'On the nuclear norm and the singular value decomposition of tensors', *Found. Comput. Math.* **16**(3) (2016), 779–811. MR 3494510

[18] K. Efremenko, A. Garg, R. Oliveira and A. Wigderson, *Barriers for Rank Methods in Arithmetic Complexity*, 9th Innovations in Theoretical Computer Science, LIPIcs. Leibniz Int. Proc. Inform., vol. 94 (Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2018), Art. No. 1, 19. MR 3761737

[19] C. Farnsworth, 'Koszul–Young flattenings and symmetric border rank of the determinant', *J. Algebra* **447** (2016), 664–676. MR 3427655

[20] W. Fulton and J. Harris, *Representation Theory*, Graduate Texts in Mathematics, vol. 129 (Springer-Verlag, New York, 1991), A first course, Readings in Mathematics. MR 1153249 (93a,20069)

[21] M. Gałązka, 'Vector bundles give equations of cactus varieties', *Linear Algebra Appl.* **521** (2017), 254–262. MR 3611482

[22] M. Gałązka, 'Multigraded apolarity', *Math. Nachr.* **296**(1) (2023), 286–313. MR 4553599

[23] M. Gallet, K. Ranestad and N. Villamizar, 'Varieties of apolar subschemes of toric surfaces', *Ark. Mat.* **56**(1) (2018), 73–99. MR 3800460

[24] M. Haiman and B. Sturmfels, 'Multigraded Hilbert schemes', *J. Algebraic Geom.* **13**(4) (2004), 725–769. MR 2073194

[25] R. Hartshorne, *Deformation Theory,* Graduate Texts in Mathematics, vol. 257 (Springer, New York, 2010). MR 2583634

[26] J. E. Humphreys, *Introduction to Lie Algebras and Representation Theory*, Graduate Texts in Mathematics, vol. 9 (Springer-Verlag, New York, 1978), Second printing, revised. MR MR499562 (81b:17007)

[27] A. Iarrobino and V. Kanev, *Power Sums, Gorenstein Algebras, and Determinantal Loci*, Lecture Notes in Mathematics, vol. 1721 (Springer-Verlag, Berlin, 1999), Appendix C by Iarrobino and S. L. Kleiman. MR MR1735271 (2001d:14056)

[28] A. W. Knapp, *Lie Groups Beyond an Introduction*, second edn., Progress in Mathematics, vol. 140 (Birkhäuser Boston, Inc., Boston, MA, 2002). MR MR1920389 (2003c:22001)

[29] J. M. Landsberg, 'The border rank of the multiplication of 2×2 matrices is seven', *J. Amer. Math. Soc.* **19**(2) (2006), 447–459. MR 2188132 (2006j:68034)

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press

[30] J. M. Landsberg, *Tensors: Geometry and Applications*, Graduate Studies in Mathematics, vol. 128 (American Mathematical Society, Providence, RI, 2012). MR 2865915
[31] J. M. Landsberg, *Geometry and complexity theory,* Cambridge Studies in Advanced Mathematics, vol. 169 (Cambridge University Press, Cambridge, 2017). MR 3729273
[32] J. M. Landsberg, *Tensors: Asymptotic Geometry and Developments 2016–2018*, CBMS Regional Conference Series in Mathematics, vol. **132** (AMS, 2019).
[33] J. M. Landsberg and M. Michałek, 'Abelian tensors', *J. Math. Pures Appl. (9)* **108**(3) (2017), 333–371. MR 3682743
[34] J. M. Landsberg and M. Michałek, 'On the geometry of border rank decompositions for matrix multiplication and other tensors with symmetry', *SIAM J. Appl. Algebra Geom.* **1**(1) (2017), 2–19. MR 3633766
[35] J. M. Landsberg and G. Ottaviani, 'Equations for secant varieties of Veronese and other varieties', *Ann. Mat. Pura Appl. (***4**) **192(4**) (2013), 569–606. MR 3081636
[36] J. M. Landsberg and G. Ottaviani, 'New lower bounds for the border rank of matrix multiplication', *Theory Comput.* **11** (2015), 285–298. MR 3376667
[37] T. Lickteig, 'A note on border rank', *Inform. Process. Lett*. **18**(3) (1984), 173–178. MR 86c:68040
[38] A. V. Smirnov, 'The bilinear complexity and practical algorithms for matrix multiplication', *Comput. Math. Math. Phys.* **53**(12) (2013), 1781–1795. MR 3146566
[39] V. Strassen, 'Relative bilinear complexity and matrix multiplication', *J. Reine Angew. Math.* **375/**376 (1987), 406–443. MR MR882307 (88h:11026)
[40] V. Strassen, 'Gaussian elimination is not optimal', *Numer. Math.* **13** (1969), 354–356. MR 248973
[41] V. Strassen, 'Rank and optimal computation of generic tensors', *Linear Algebra Appl.* **52/53** (1983), 645–685. MR 85b:15039

https://doi.org/10.1017/fmp.2023.14 Published online by Cambridge University Press