NEW LOWER BOUNDS FOR MATRIX MULTIPLICATION AND $\det_3$

AUSTIN CONNER, ALICIA HARPER, AND J. M. LANDSBERG

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 methods used to obtain the results are new, and "non-natural" in the sense of Razborov and Rudich, in that the results are obtained via an algorithm that cannot be 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 tensors with symmetry to obtain an algorithm that, given a tensor T with a large symmetry group 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 potential border rank r decompositions in a natural normal form. The algorithm is based on algebraic geometry and representation theory.

1. Introduction

Over fifty years ago Strassen [30] 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 [10, 7, 20] for an overview. Matrix multiplication of $\mathbf{n}\times\mathbf{l}$ matrices with $\mathbf{l}\times\mathbf{m}$ matrices is a bilinear map, i.e., a tensor $M_{\langle \mathbf{l},\mathbf{m},\mathbf{n}\rangle} \in \mathbb{C}^{\mathbf{lm}}\otimes\mathbb{C}^{\mathbf{mn}}\otimes\mathbb{C}^{\mathbf{nl}}$, and since 1980 [5], the primary complexity measure of the matrix multiplication tensor has been its *border rank*, which is defined as follows:

A 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.

Despite the vast literature on matrix multiplication, previous to this paper, the precise border rank of $M_{\langle \mathbf{l},\mathbf{m},\mathbf{n}\rangle}$ was known in exactly one nontrivial case, namely $M_{\langle 2\rangle} = M_{\langle 222\rangle}$ [19]. We

2010 *Mathematics Subject Classification.* 68Q15, 15A69, 14L35.

Key words and phrases. Tensor rank, Matrix multiplication complexity, determinant complexity, border rank.

Landsberg supported by NSF grant AF-1814254 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 supplied by the Simons Foundation and by the Mathematisches Forschungsinstitut Oberwolfach.

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. See §1.2 below for precise statements.

1.1. **Methods/History.** This paper deals exclusively with lower bounds ("Complexity theory's Waterloo" according to [4, Chap. 14]). For a history of upper bounds see e.g., [7, 20].

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 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 multi-linear algebra to linear algebra [29], 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 [23, 24]. Such methods are subject to *barriers* [14, 16] (see [21, §2.2] for an overview). A technique allowing one to go slightly beyond the barriers was introduced in [22]. 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 [22] relied on first using the symmetry of the tensor to study its degenerations (via the *Normal form lemma*), and then to use polynomials on the degeneration of the tensor.

The classical *apolarity method* was introduced for studying the decomposition of a homogeneous polynomial of degree d into a sum of d-th powers of linear forms (Waring rank), see, e.g., [18]. It was generalized to study ranks of points with respect to toric varieties. 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 [9], Buczyńska and Buczyński introduce new language that enables them to extend this classical method to the border rank setting. They then extend the normal form lemma to the entire ideal associated to the border rank decomposition of the tensor (their *Fixed ideal theorem*). (In the language introduced below, the Normal form lemma is the (111) case of the Fixed ideal theorem.) Our contribution to their theory is to convert their Fixed ideal theorem into an effective algorithm in the situation of tensors with large symmetry groups and to successfully apply it to important tensors. This contribution was obtained while [9] was being developed and in regular discussions with Buczyńska and Buczyński.

Given r, the algorithm builds a candidate ideal step by step, starting in low (multi)-degree and building upwards. At each building step, there is a test that, if the so-far built ideal fails to pass, it is eliminated from consideration. If at any point there are no candidates, one concludes there is no border rank r decomposition. All the results of this paper just use the first steps of this algorithm. For tensors with symmetry, the Fixed ideal theorem drastically reduces the candidates one needs to consider, see §2.4 restriction (iv).

The eliminations are obtained when the ranks of certain linear maps are too large. 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 multi-degree (in fact one may assume it has codimension r in each multi-degree).

The fact that the elimination conditions are rank conditions implies that the barriers [14, 16] 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.

We use *representation theory* at several levels: The border apolarity method applied to tensors involves the study of an ideal of polynomials in three sets of variables, so we have a $\mathbb{Z}^3$-graded ring of polynomials. This enables us to study a putative ideal I in each multi-degree. For tensors with "large" symmetry groups, for each $(i,j,k) \in \mathbb{Z}^3$ the Fixed ideal theorem reduces the possible candidate I_{ijk}'s to a short list. Given such data, one then must compute the ranks of the above-mentioned multiplication maps for each candidate. One can do this by computer. This is how we obtain our results for $M_{\langle 3\rangle}$ and $\det_3$, although, in both cases, since the matrices are large, numerous, and with parameters appearing, several innovations were required to perform the computations.

For $M_{\langle 2\mathbf{nn}\rangle}$ and $M_{\langle 3\mathbf{nn}\rangle}$ a computer calculation is not possible for all $\mathbf{n}$. Here we show there are no Borel-fixed (110)-spaces that could possibly be extended to ideals by splitting the problem into a local and a global problem: We show that the total contribution to a test can be upper-bounded by adding local contributions (Lemma 7.6). This enabled us to set-up an optimization problem to bound all possible sums of local contributions, which we then solved (Lemma 7.9) by showing a modification of it is convex. We emphasize that this method for proving lower bounds is completely different from previous techniques.

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 24×40 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.

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 7.1 (Remark 7.2).

1.2. **Results.**

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

The previous lower bounds were 14 [29] in 1983, 15 [24] in 2015, and 16 [22] 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. In [12] we explain how one could potentially use $\det_3$ to prove the exponent of matrix multiplication is two via upper bounds on the border rank of its Kronecker powers. Here we determine the base case:

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

The upper bound was proved in [12]. In [8] 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 [15], which follows from the Koszul flattening equations [24].

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

The upper bound dates back to Bini et. al. in 1980 [6]. Koszul flattenings [24] give $\underline{\mathbf{R}}(M_{\langle 22\mathbf{n}\rangle}) \geq 3\mathbf{n}$. Smirnov [28] 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 1.4.

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

(2) *We have the following border rank lower bounds:*

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

Previously only the near trivial result that $\underline{\mathbf{R}}(M_{\langle 2\mathbf{nn}\rangle}) \geq \mathbf{n}^2 + 1$ was known by [25, Rem. p175], see §8.

The upper bound in (1) is due to Smirnov [28], 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 $\underline{\mathbf{R}}(M_{\langle 2\mathbf{nn}\rangle}) \leq \frac{7}{4}\mathbf{n}^2$ by writing $M_{\langle 2\mathbf{nn}\rangle} = M_{\langle 222\rangle} \boxtimes M_{\langle 1\frac{\mathbf{n}}{2}\frac{\mathbf{n}}{2}\rangle}$, where $\boxtimes$ denotes Kronecker product of tensors, see, e.g., [12].

Theorem 1.5. *For all* $\mathbf{n} \geq 18$, $\underline{\mathbf{R}}(M_{\langle 3\mathbf{nn}\rangle}) \geq \mathbf{n}^2 + \sqrt{\frac{8}{3}}\mathbf{n} > \mathbf{n}^2 + 1.6\mathbf{n}$.

Previously the only bound was the the near trivial result that when $\mathbf{n} \geq 4$, $\underline{\mathbf{R}}(M_{\langle 3\mathbf{nn}\rangle}) \geq \mathbf{n}^2 + 2$ by [25, Rem. p175], see §8.

Using [25, Rem. p175], one obtains

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

Remark 1.7. Koszul flattenings [24] fail to give border rank lower bounds for tensors in $A\otimes B\otimes C$ when the dimension of one of A, B, C is much larger than that of the other two, such as $M_{\langle 2\mathbf{nn}\rangle} \in \mathbb{C}^{2\mathbf{n}}\otimes\mathbb{C}^{2\mathbf{n}}\otimes\mathbb{C}^{\mathbf{n}^2}$ and $M_{\langle 3\mathbf{nn}\rangle} \in \mathbb{C}^{3\mathbf{n}}\otimes\mathbb{C}^{3\mathbf{n}}\otimes\mathbb{C}^{\mathbf{n}^2}$. Theorems 1.4 and 1.5 show that the border apolarity method does not share this defect.

1.3. What comes next?

1.3.1. *Breaking the barriers.* The geometric interpretation of the barriers is that all equations obtained by taking minors are actually equations for a larger variety than $\sigma_r(\text{Seg}(\mathbb{P}A\times\mathbb{P}B\times\mathbb{P}C)$, called the *r-th cactus variety.* This cactus variety agrees with the secant variety for small r, but it quickly fills the ambient space of tensors in $\mathbb{C}^m\otimes\mathbb{C}^m\otimes\mathbb{C}^m$ at latest when $r = 6m - 4$. Thus after the filling dimension, one cannot prove lower bounds via minors of matrices. 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 algorithm produces ideals, and thus to break the barrier, one needs to distinguish limits of ideals of smooth schemes from limits of ideals of non-smoothable schemes. In general this is hopeless, but for matrix multiplication, the ideals are of a very special type of scheme: one supported at a single point and fixed by a Borel subgroup. Here there are techniques that we hope to apply effectively. This is work in progress with Buczyńska and Buczyński.

The matrix multiplication tensor may well have cactus border rank below its actual border rank and the general barrier, in which case the barrier for proving lower bounds for it using determinantal methods would occur below that of other tensors.

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 [3, 2, 1, 11]. These tensors have positive dimensional symmetry groups, so the border apolarity method potentially may be applied. For example, the small Coppersmith-Winograd tensor $T_{cw,q}$ has a very large symmetry group, namely the orthogonal group $O(q)$ [13], 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. The method will produce candidate ideals, and we are developing techniques to write down usual border rank decompositions guided by these ideals. This is joint work in progress with A. Huang, Buczyńska and Buczyński.

1.3.3. *Geometrization of the* (111) *test for matrix multiplication.* Our results for $M_{\langle \mathbf{2nn}\rangle}, M_{\langle \mathbf{3nn}\rangle}$ 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 multi-graded ideals one may associate to a border rank decomposition. We also review Borel fixed (highest weight) subspaces. 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 calculations, the code for which is available at `https://www.math.tamu.edu/~jml/bapolaritycode.html`. 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 1.3, 1.4 and 1.5. In §8, for the convenience of the reader, we give a proof of Lickteig's result that $\underline{\mathbf{R}}(M_{\langle \mathbf{l,m,n}\rangle}) \geq \underline{\mathbf{R}}(M_{\langle \mathbf{l-1,m,n}\rangle}) + 1$ for all $\mathbf{l}, \mathbf{m}, \mathbf{n}$ which is used in the proof of Corollary 1.6.

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 Oliviera and anonymous referees for significant help with the exposition.

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 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 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 $\mathfrak{S}_r$.

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 $E\in G(r,A)$, i.e., a linear subspace $E\subset A$ of dimension r, if $e_1,\cdots,e_r$ 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, $\widehat{Z}\subset A$ denotes the cone over Z union the origin, $I(Z)=I(\widehat{Z})\subset\mathrm{Sym}(A^*)$ denotes the ideal of Z, and $\mathbb{C}[\widehat{Z}]=\mathrm{Sym}(A^*)/I(Z)$, denotes the homogeneous coordinate ring of $\widehat{Z}$. Both $I(Z)$, $\mathbb{C}[\widehat{Z}]$ 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, i.e., if it requires all basis vectors in each of A,B,C to write down in any basis.

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

(1) for all $t\neq 0$, E_t is spanned by r rank one tensors, and

(2) $T\in E_0$.

For example 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

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

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. **Multi-graded 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

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

denote the $\mathbb{Z}^3$-graded ideal of the set of r points $[T_1(t)]\sqcup\cdots\sqcup[T_r(t)]$, where $I_{ijk,t}\subset S^iA^*\otimes S^jB^*\otimes S^kC^*$. If the r points are in general position, then $\operatorname{codim}(I_{ijk,t})=r$ as long as $r\le\dim S^iA^*\otimes S^jB^*\otimes S^kC^*$. In our situation r will be sufficiently small so that this will hold if at least two of i,j,k are nonzero, see e.g., [17, 16, 31]. For all (ijk) with $i+j+k>1$, we may choose the curves such that $\operatorname{codim}(I_{ijk})=r$ by [9, Thm. 1.2].

Thus, in addition to $E_0=I_{111,0}^\perp$ 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 $G(\dim(S^iA^*\otimes S^jB^*\otimes S^kC^*)-r, S^iA^*\otimes S^jB^*\otimes S^kC^*)$. We remark that there are subtleties here: the limiting ideal may not be saturated. See [9] for a discussion.

Thus we may assume a multi-graded 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. This condition 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 $i+j+k>1$, $\operatorname{codim} I_{ijk}=r$.

(iii) I is an ideal, so the multiplication maps

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

have image contained in I_{ijk}.

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 way to prove this, but for tensors with a nontrivial 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 solvable group, let W be an H-module, and let $[w]\in\mathbb{P}W$. Then the orbit closure $\overline{H\cdot[w]}$ contains an H-fixed point.

Assume G_T is reductive (or contains a nontrivial reductive subgroup). Let $\mathbb{B}_T\subset G_T$ be a maximal solvable subgroup, called a *Borel* subgroup. By Lie's theorem and the Normal Form

Lemma of [22], in order to prove $\underline{\mathbf{R}}(T) > r$, it is sufficient to disprove the existence of a border rank decomposition where E_0 is a $\mathbb{B}_T$-fixed point of $\mathbb{P}\Lambda^r(A\otimes B\otimes C)$.

By the same reasoning, as observed in [9], we may assume I_{ijk} is $\mathbb{B}_T$-fixed for all i,j,k. When G_T is large, this can reduce the problem significantly.

Thus we may assume a multi-graded 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, Borel fixed spaces are easy to list.

2.5. **Borel fixed subspaces.** We review standard facts about Borel fixed subspaces. In this paper only general and special linear groups and products of such appear. A Borel subgroup of GL_m is the group of invertible matrices that are zero below the diagonal, and in products of general linear groups, the product of Borel subgroups is a Borel subgroup. Let $\mathbb{C}^m$ have basis $e_1,\cdots,e_m$, with dual basis $e^1,\cdots,e^m$. Assign e_j weight $(0,\cdots,0,1,0,\cdots,0)$, where the 1 is in the j-th slot, and e^j weight $(0,\cdots,0,-1,0,\cdots,0)$. For vectors in $(\mathbb{C}^m)^{\otimes d}$, $wt(e_1^{\otimes a_1}\otimes\cdots\otimes e_m^{\otimes a_m}) = (a_1,\cdots,a_m)$ and the weight is unchanged under permutations of the $d = a_1 + \cdots + a_m$ factors. Partially order the weights so that $(i_1,\cdots,i_m) \geq (j_1,\cdots,j_m)$ if $\sum_{\alpha=1}^s i_\alpha \geq \sum_{\alpha=1}^s j_\alpha$ for all s. The action of the Borel on a monomial μ sends it to a sum of monomials whose weights are higher than that of μ in the partial order plus a monomial that is a scalar multiple of μ. Each irreducible GL_m module appearing in the tensor algebra of $\mathbb{C}^m$ has a unique highest weight which is given by a partition $\pi = (p_1,\cdots,p_m)$ and the module is denoted $S_\pi\mathbb{C}^m$. Write $d = |\pi| = \sum p_i$. See any of, e.g., [20, §8.7], [27, §9.1], or [26, I.A] for details. Let $\mathbb{T} \subset \mathrm{GL}_m$ denote the maximal torus of diagonal matrices. A vector w (or line $[w]$) is a *weight vector (line)* if the line $[w]$ is fixed by the action of $\mathbb{T}$. For reductive groups G, we let $\mathbb{B}$ denote a choice of Borel subgroup.

We will use SL_m weights, which we write as $c_1\omega_1+\cdots+c_{m-1}\omega_{m-1}$, where the ω_j are the fundamental weights. Here $wt(e_1) = \omega_1$, $wt(e_m) = -\omega_{m-1}$, for $2 \leq s \leq m-1$, $wt(e_s) = \omega_s - \omega_{s-1}$ and for all j, $wt(e^j) = -wt(e_j)$. See the above references for explanations.

After fixing a (weight) basis of $\mathbb{C}^m$, an irreducible G-submodule M of $(\mathbb{C}^m)^{\otimes d}$ has a basis of weight vectors, which is unique up to scale if M is multiplicity free, i.e., there is at most one weight line of any given weight. In this case the $\mathbb{B}$-fixed subspaces of dimension k, considered as elements of the Grassmannian $G(k,M)$, are just wedge products of choices of k-element subsets of the weight vectors of M such that no other element of $G(k,M)$, considered as a line in $\Lambda^k M$, has higher weight in the partial order. In the case a weight occurs with multiplicity in M, one has to introduce parameters in describing the subspaces. In the case of direct sums of irreducible modules $M_1 \oplus M_2$, a subspace is $\mathbb{B}$-fixed if it is spanned by weight vectors and, setting all the M_2-vectors in a basis of the subspace zero, what remains is a $\mathbb{B}$-fixed subspace of M_1 and similarly with the roles of M_1, M_2 reversed.

In discussing weights, it is convenient to work with Lie algebras. Let $\mathfrak{b}$ denote the Lie algebra of $\mathbb{B}$ and let $\mathfrak{u} \subset \mathfrak{b}$ be the space of upper triangular matrices with zero on the diagonal. We refer to elements of $\mathfrak{u}$ as *raising operators*. A vector (or line) is a *highest weight vector (line)* if it is a weight vector (line) annihilated by the action of $\mathfrak{u}$. A subspace of M of dimension k is $\mathbb{B}$-fixed if and only if, considered as a line in $\Lambda^k M$, it is a highest weight line.

Example 2.1. When U,V,W each have dimension 2, Figure 1 gives the $\mathrm{SL}(U)\times\mathrm{SL}(V)\times\mathrm{SL}(W)$-weight diagram for $U^*\otimes\mathfrak{sl}(V)\otimes W$. Here, in each factor $\mathfrak{u}$ is spanned by the matrix

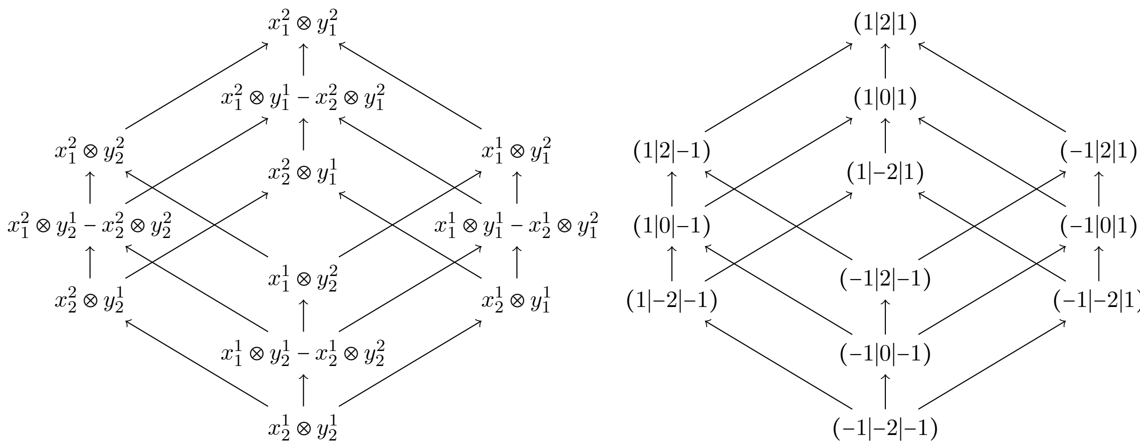

FIGURE 1. Weight diagram for $U^*{\otimes}\mathfrak{sl}(V){\otimes}W$ when $U = V = W = \mathbb{C}^2$. Left are the weight vectors and right the weights: since $\mathfrak{sl}_2$ weights are just $j\omega_1$, we have just written $(i|j|k)$ for the $\mathfrak{sl}(U)\oplus\mathfrak{sl}(V)\oplus\mathfrak{sl}(W)$ weight. Raisings in U^* correspond to NW (north-west) arrows, those in W to NE arrows and those in $\mathfrak{sl}(V)$ to upward arrows.

$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ (raising goes from bottom to top). There is a unique $\mathbb{B}$-fixed (highest weight) line, spanned by $x_1^2{\otimes}y_1^2$, (here $x_j^i = u^i{\otimes}v_j$, $y_j^i = v^i{\otimes}w_j$, and $z_j^i = w^i{\otimes}u_j$) three highest weight 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 $\langle x_1^2{\otimes}y_1^2, x_1^2{\otimes}y_1^1 - x_2^2{\otimes}y_1^2\rangle$, four highest weight 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 $\langle x_1^2{\otimes}y_1^2, x_1^2{\otimes}y_1^1 - x_2^2{\otimes}y_1^2, x_2^2{\otimes}y_1^1\rangle$, etc..

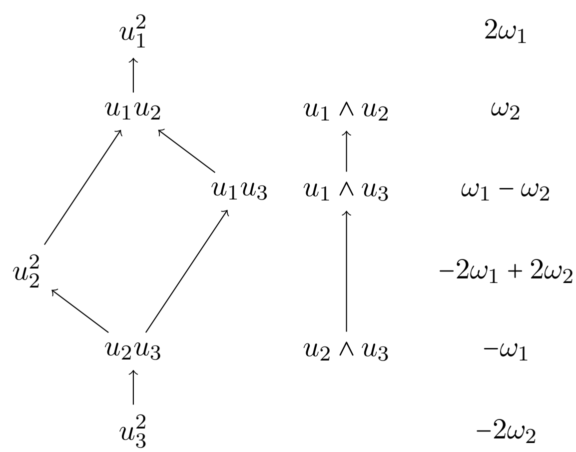

FIGURE 2. Weight diagram for $U{\otimes}U$ when $U = \mathbb{C}^3$. There are 6 distinct weights appearing, indicated on the right. On the far left are the weight vectors in S^2U and in the middle are the weight vectors in Λ^2U.

Example 2.2. Let $\dim U = 3$. Figure 2 gives the weight diagram for $U{\otimes}U = S^2U \oplus \Lambda^2U$. 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 $\langle u_1 \wedge u_2, u_1 \wedge u_3\rangle$ etc..

Example 2.3. Figure 3 gives the weight diagram for $\mathfrak{sl}_3$. Here $v^i_j = v_j \otimes v^i$. The oval is around the two-dimensional weight zero subspace, which has four distinguished vectors: two with only two weight vectors above them in the partial order, and two with only two weight vectors below them in the partial order. Equivalently, the distinguished vectors up to scale are images and kernels of the two raising operators.

The $\mathbb{B}$-fixed subspaces of dimension 3 are $X = \langle v^3_1, v^3_2, 2v^3_3 - (v^1_1 + v^2_2)\rangle$, $X = \langle v^3_1, v^2_1, 2v^1_1 - (v^2_2 + v^3_3)\rangle$ and $X = \langle v^3_1, v^2_1, v^3_2\rangle$.

The $\mathbb{B}$-fixed subspaces of dimension 4 are a family parametrized by $[s,t] \in \mathbb{P}^1$: $X = \langle v^3_1, v^3_2, s(2v^3_3 - (v^1_1 + v^2_2)) + t(2v^1_1 - (v^2_2 + v^3_3))\rangle$.

The $\mathbb{B}$-fixed subspaces of dimension 5 are, the weight ≥ 0 space, $X = \langle v^3_1, v^2_1, v^3_2, v^2_2 - v^3_3, v^2_3\rangle$, and $X = \langle v^3_1, v^2_1, v^3_2, v^1_1 - v^2_2, v^1_2\rangle$.

The other $\mathbb{B}$-fixed subspaces are clear from the picture.

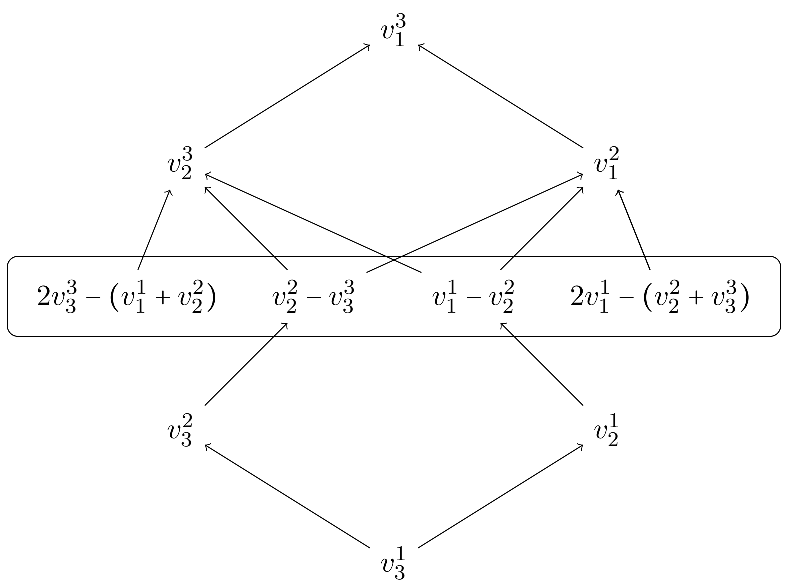

FIGURE 3. Weight diagram for $\mathfrak{sl}_3$.

3. THE ALGORITHM

Input: An integer r and a concise tensor $T \in A \otimes B \otimes C$ whose symmetry group that contains a reductive group with Borel subgroup $\mathbb{B}_T$.

Output: Either a proof that $\underline{\mathbf{R}}(T) > r$ or a list of all Borel-fixed ideals that could potentially arise in a border rank r decomposition of T.

The following steps build an ideal I in each multi-degree. 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 weight subspace F_{110} of codimension $r-\mathbf{c}$ in $T(C^*)^\perp \subset A^*\otimes B^*$ (and codimension r in $A^*\otimes B^*$) compute the ranks of the multiplication maps

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

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

If both have images of codimension at least r, then F_{110} is a candidate 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 potential $I_{101} \subset T(B^*)^\perp$ and $I_{011} \subset T(A^*)^\perp$ to obtain spaces F_{101}, F_{011}.

(iii) For each triple $F_{110}, F_{101}, F_{011}$ passing the above tests, compute the rank of the map

$$F_{110}\otimes C^* \oplus F_{101}\otimes B^* \oplus F_{011}\otimes A^* \to A^*\otimes B^*\otimes C^*. \tag{4}$$

If the codimension of the image is at least r, then one has a candidate triple. Call this map the (111)-map and the rank condition the (111)-test. A space F_{111} is a candidate for I_{111} if it is of codimension r, contains the image of (4) and it is contained in $T^\perp$.

(iv) For each candidate triple $F_{110}, F_{101}, F_{011}$ obtained in the previous step, and for each $\mathbb{B}_T$-fixed subspace $F_{200} \subset S^2A^*$ of codimension r, compute the rank of the maps $F_{110}\otimes A^* \oplus F_{200}\otimes B^* \to S^2A^*\otimes B^*$ and $F_{101}\otimes A^* \oplus F_{200}\otimes C^* \to S^2A^*\otimes B^*$. If the codimension of these images is at least r, then one may add F_{200} to the candidate set.

Do the same for $\mathbb{B}_T$-fixed subspaces F_{020} and F_{002}, and collect all total degree two candidate sets.

(v) Given an up until this point candidate set $\{F_{uvw}\}$ including degrees $(i-1,j,k)$, $(i,j-1,k)$, and $(i,j,k-1)$, compute the rank of the map

$$F_{i-1,j,k}\otimes A^* \oplus F_{i,j-1,k}\otimes B^*\otimes F_{i,j,k-1}\otimes C^* \to S^iA^*\otimes S^jB^*\otimes S^kC^*. \tag{5}$$

If the codimension of the image of this map is less than r, the set is not a candidate. Say the codimension of the image is $\xi \geq r$. The image will be $\mathbb{B}_T$-fixed by Schur's Lemma, as (5) is a $\mathbb{B}_T$-module map. Each $(\xi-r)$-dimensional $\mathbb{B}_T$-fixed subspace of the image (i.e., codimension r $\mathbb{B}_T$-fixed subspace of $S^iA^*\otimes S^jB^*\otimes S^kC^*$ in the image) is a candidate F_{ijk}.

(vi) If at any point there are no such candidates, we conclude $\underline{\mathbf{R}}(T) > r$.

Despite appearances, the algorithm is finite: it must stabilize at latest in multi-degree (r,r,r), see [9]. That is, if all maps up to that point have the correct ranks, the higher degree maps also will and there will be no new generators of the ideal in higher multi-degrees. Thus the output is either a certificate that $\underline{\mathbf{R}}(T) > r$ or a collection of multi-graded ideals representing all possible candidates for a $\mathbb{B}_T$-fixed border rank decomposition. In current work with Buczyńska and Buczyński we are developing tests to determine if a given multi-graded ideal comes from a border rank decomposition.

The algorithm above in total degree three suffices to obtain the lower bounds proved in this article.

Sometimes it is more convenient to perform the tests dually:

Proposition 3.1. *The codimension of the image of the* (210)*-map is the dimension of the kernel of the skew-symmetrization map*

$$F_{110}^{\perp}\otimes A \to \Lambda^2 A\otimes B. \tag{6}$$

The codimension of the image of the (ijk)*-map is the dimension of*

$$(F_{ij,k-1}^{\perp}\otimes C)\cap(F_{i,j-1,k}^{\perp}\otimes B)\cap(F_{i-1,j,k}^{\perp}\otimes A). \tag{7}$$

Proof. The codimension of the image of the (210)-map is the dimension of the kernel of its transpose,

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

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 (6).

The codimension of the image of the (ijk)-map is the dimension of the kernel of its transpose. Let $X\in S^iA\otimes S^jB\otimes S^kC$. Write $\mathrm{Proj}_{ij,k-1}(X) = X \bmod F_{ij,k-1}^{\perp}\otimes C$, $\mathrm{Proj}_{ij-1,k}(X) = X \bmod F_{ij-1,k}^{\perp}\otimes B$, and $\mathrm{Proj}_{i-1,jk}(X) = X \bmod F_{i-1,jk}^{\perp}\otimes A$. The transpose is the map

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

so X is in the kernel if and only if all three projections are zero. The kernels of the three projections are respectively $(F_{ij,k-1}^{\perp}\otimes C)$, $(F_{i,j-1,k}^{\perp}\otimes B)$, and $(F_{i-1,jk}^{\perp}\otimes A)$, so 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 $\mathrm{Id}_U\otimes\mathrm{Id}_V\otimes\mathrm{Id}_W$. Thus $G_{M_{\langle \mathbf{u},\mathbf{v},\mathbf{w}\rangle}}\supseteq GL(U)\times GL(V)\times GL(W) =: G$. 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.

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 $E_{ijk}^{\perp}$ is.

Any 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 $U^*\otimes\mathfrak{sl}(V)\otimes W$. 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 $\dim E'_{110} = r-\mathbf{wu}$.

Since $M_{\langle \mathbf{n}\rangle}$ has $\mathbb{Z}_3$-symmetry (via cyclic permutation of factors), to determine the candidate I_{110}, I_{101} and I_{011} it will suffice to determine the candidate I_{110}'s. Similarly, since $M_{\langle \mathbf{n},\mathbf{l},\mathbf{n}\rangle}$ has $\mathbb{Z}_2$-symmetry, the list of candidate I_{110}'s is isomorphic to the list of candidate I_{011}'s.

First proof that $\underline{\mathbf{R}}(M_{\langle 2\rangle}) = 7$. Here $\mathbf{u}=\mathbf{v}=\mathbf{w}=2$. We disprove border rank at most six by showing no $\mathbb{B}$-fixed six dimensional F_{110} (i.e., 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 1.

By Figure 1, there are three $\mathbb{B}$-fixed 2-planes in $U^*\otimes\mathfrak{sl}(V)\otimes W$:

$$\langle(u^2\otimes v_1)\otimes(v^2\otimes w_1),(u^1\otimes v_1)\otimes(v^2\otimes w_1)\rangle,$$
$$\langle(u^2\otimes v_1)\otimes(v^2\otimes w_1),(u^2\otimes v_1)\otimes(v^2\otimes w_2)\rangle,$$
$$\text{and } \langle(u^2\otimes v_1)\otimes(v^2\otimes w_1),(u^2\otimes v_1)\otimes(v^1\otimes w_1)-(u^2\otimes v_2)\otimes(v^2\otimes w_1)\rangle.$$

For the first, the rank of the 24×40 matrix of the map $E_{110}^{\perp}\otimes A^*\to S^2A^*\otimes B^*$ is $20>24-6=18$. For the second, by symmetry, the rank of the (120)-map is also 20. For the third the rank of the (210)-map is 19 and we conclude. □

For readers unhappy with computing the rank of a sparse 40×24 matrix whose entries are all $0,\pm1$, the following remark reduces to 24×24 matrices, and in §6.2, using more representation theory, we reduce to 4×8 matrices whose entries are all $0,\pm1$. Finally we give a calculation free proof in Remark 7.2.

Remark 4.1. One can simplify the calculation of the rank of the map $E_{110}^{\perp}\otimes A^*\to S^2A^*\otimes B^*$ by using the map (6). In the case above, the resulting matrix is of size 24×24. 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

$x_1^1\wedge x_1^2\otimes y_1^2, x_2^1\wedge x_1^2\otimes y_1^2, x_2^2\wedge x_1^2\otimes y_1^2,$

$x_2^1\wedge x_1^1\otimes y_1^2, x_2^2\wedge x_1^1\otimes y_1^2,$

$x_1^1\wedge(x_1^1\otimes y_1^1+x_2^1\otimes y_1^2), x_2^1\wedge(x_1^1\otimes y_1^1+x_2^1\otimes y_1^2), x_1^2\wedge(x_1^1\otimes y_1^1+x_2^1\otimes y_1^2), x_2^2\wedge(x_1^1\otimes y_1^1+x_2^1\otimes y_1^2),$

$x_1^1\wedge(x_1^2\otimes y_1^1+x_1^2\otimes y_1^2), x_2^1\wedge(x_1^2\otimes y_1^1+x_1^2\otimes y_1^2), x_1^2\wedge(x_1^2\otimes y_1^1+x_2^2\otimes y_1^2), x_2^2\wedge(x_1^2\otimes y_1^1+x_2^2\otimes y_1^2),$

$x_1^1\wedge(x_1^2\otimes y_2^1+x_2^2\otimes y_2^2), x_2^1\wedge(x_1^2\otimes y_2^1+x_2^2\otimes y_2^2), x_1^2\wedge(x_1^2\otimes y_2^1+x_2^2\otimes y_2^2), x_2^2\wedge(x_1^2\otimes y_2^1+x_2^2\otimes y_2^2)$

$x_1^1\wedge(x_1^2\otimes y_2^1+x_2^2\otimes y_2^2), x_2^1\wedge(x_1^2\otimes y_2^1+x_2^2\otimes y_2^2), x_1^2\wedge(x_1^2\otimes y_2^1+x_2^2\otimes y_2^2), x_2^2\wedge(x_1^2\otimes y_2^1+x_2^2\otimes y_2^2)$

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.

5. Explanation of the proofs of Theorems 1.1 and 1.2

The actual proofs to these theorems are in the code at the web page `https://www.math.tamu.edu/~jml/bapolaritycode.html`. What follows are explanations of what is carried out.

In the case of $M_{\langle 3\rangle}$, the weight zero subspace of $\mathfrak{sl}_3$ has dimension two, so there are $\mathbb{B}$-fixed spaces of dimension $7=16-3\cdot 3$ in $U^*\otimes\mathfrak{sl}(V)\otimes W\subset A\otimes B$ that arise in positive dimensional families. (Here 16 is the border rank we wish to rule out and $3\cdot 3=\dim U^*\otimes \mathrm{Id}_V\otimes W$.) Fortunately the set of 7-planes that pass the (210) and (120) tests is finite. When there are no parameters present, these tests consist of computing the ranks of the 144×405 matrices of $E_{110}^{\perp}\otimes A^*\to S^2A^*\otimes B^*$, and $E_{110}^{\perp}\otimes B^*\to A^*\otimes S^2B^*$. When there are parameters, one determines the ideal in which the rank drops to the desired value. There are eight 7-planes that do pass the test, giving rise to 512 possible triples (or 176 triples taking symmetries into account). Among the candidate triples, none pass the (111)-test.

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 $\Lambda^mU\otimes\Lambda^mV\subset A_1\otimes\cdots\otimes A_m$. Explicitly, letting A_α have basis x_{ij}^{α},

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

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 $\mathfrak{S}_3$. In particular, to determine the candidate E_{110}'s it is sufficient to look in $A \otimes B$, which, as an $\mathrm{SL}(U) \times \mathrm{SL}(V)$-module is $U^{\otimes 2} \otimes V^{\otimes 2} = S^2U \otimes S^2V \oplus S^2U \otimes \Lambda^2 V \oplus \Lambda^2 U \otimes S^2 V \oplus \Lambda^2 U \otimes \Lambda^2 V$, and $\det_3(C^*) = \Lambda^2 U \otimes \Lambda^2 V$.

In the case of $\det_3$, 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, e.g., $u_1u_2 \otimes v_1v_2, u_1u_2 \otimes v_1 \wedge v_2$, and $u_1 \wedge u_2 \otimes v_1 v_2$ each have weight $(\omega_2^U | \omega_2^V)$. We examine all 7-dimensional $\mathbb{B}$-fixed subspaces of $S^2U \otimes S^2V \oplus S^2U \otimes \Lambda^2 V \oplus \Lambda^2 U \otimes S^2 V$. There are four candidates passing the (210) and (120) tests, but no triples passed the (111) test.

In both cases, for the E'_{110} with parameters, to perform the test we first perform row reduction by constant entries. This usually reduces the problem enough to take minors even with parameters. If it does not, we use the following algorithm, which effectively allows us to do row reduction: First, generalize to matrix entries in some quotient of some ring of fractions of the polynomial ring, say R. If there is a matrix entry which is a unit, pivot by it, reducing the problem. Otherwise, select a nonzero entry, say p. Recursively compute the target ideal in two cases: 1. Pass to $R/(p)$, the computation here is smaller because the entry is zeroed. 2. Pass to R_p, the computation here is smaller because now p is a unit, and one can pivot by it. Finally lift the ideals obtained by 1 and 2 back to R, say to J_1 and J_2, and take $J_1 J_2$. Its zero set is the rank $< r$ locus and computing with it is tractable.

6. Representation theory relevant for matrix multiplication

Theorems 1.3 and 1.4(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 $M_{\langle \mathbf{u},\mathbf{v},\mathbf{w}\rangle} = \mathrm{Id}_U \otimes \mathrm{Id}_V \otimes \mathrm{Id}_W$. 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 $\mathbf{v} = 3$, $V_{\omega_2 + \omega_{\mathbf{v}-1}} = V_{2\omega_2}$):

$$\Lambda^2(U^* \otimes V) \otimes V^* = (S^2 U^* \otimes V_{\omega_1}) \oplus (\Lambda^2 U^* \otimes V_{\omega_1}) \oplus (S^2 U^* \otimes V_{\omega_2 + \omega_{\mathbf{v}-1}}) \oplus (\Lambda^2 U^* \otimes V_{2\omega_1 + \omega_{\mathbf{v}-1}}), \tag{8}$$

$$S^2(U^* \otimes V) \otimes V^* = (S^2 U^* \otimes V_{2\omega_1 + \omega_{\mathbf{v}-1}}) \oplus (\Lambda^2 U^* \otimes V_{\omega_2 + \omega_{\mathbf{v}-1}}) \oplus (S^2 U^* \otimes V_{\omega_1}) \oplus (\Lambda^2 U^* \otimes V_{\omega_1}), \tag{9}$$

$$A \otimes M_{\langle \mathbf{u},\mathbf{v},\mathbf{w}\rangle}(C^*) = (U^* \otimes V) \otimes (U^* \otimes \mathrm{Id}_V \otimes W) = (S^2 U^* \otimes V_{\omega_1} \otimes W) \oplus (\Lambda^2 U^* \otimes V_{\omega_1} \otimes W), \tag{10}$$

$$V \otimes \mathfrak{sl}(V) = V_{\omega_1} \oplus V_{2\omega_1 + \omega_{\mathbf{v}-1}} \oplus V_{\omega_2 + \omega_{\mathbf{v}-1}}, \tag{11}$$

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

Here we have written V_{ω_1} for embedded submodules isomorphic to V. Note that

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

The map $(U^* \otimes V) \otimes (U^* \otimes \mathrm{Id}_V \otimes W) \to \Lambda^2(U^* \otimes V) \otimes (V^* \otimes W)$ is injective, which implies:

Proposition 6.1. *Write $E_{110} \coloneqq M_{\langle \mathbf{u},\mathbf{v},\mathbf{w}\rangle}(C^*) \oplus E'_{110}$. The dimension of the kernel of the map (6) $E_{110} \otimes A \to \Lambda^2 A \otimes B$ equals the dimension of the kernel of the map*

$$E'_{110} \otimes A \to S^2 U^* \otimes V_{\omega_2 + \omega_{\mathbf{v}-1}} \otimes W \oplus \Lambda^2 U^* \otimes V_{2\omega_1 + \omega_{\mathbf{v}-1}} \otimes W. \tag{13}$$

and the kernel of (13) *is*

$$(14) \qquad (E'_{110} \otimes A) \cap [U^{*\otimes 2} \otimes V_{\omega_1} \otimes W \oplus S^2 U^* \otimes V_{2\omega_1 + \omega_{\mathbf{v}-1}} \otimes W \oplus \Lambda^2 U^* \otimes V_{\omega_2 + \omega_{\mathbf{v}-1}} \otimes W].$$

The second assertion follows by applying Schur's lemma using (12) as the map (13) is the restriction of an equivariant map.

6.2. $\underline{\mathbf{R}}(M_{\langle 2 \rangle}) > 6$ **revisited.** In this case the map (13) 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, where we suppress the ω_1 from the notation), the image of (13) is spanned by

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

(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 $x_1^2 \otimes y_1^1 - x_2^2 \otimes y_1^2$ (the lowering of $x_1^2 \otimes y_1^2$ under $\mathfrak{sl}(V)$), the image is spanned by

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

Since W has nothing to do with the map, we don't need to compute the image of, e.g., $A \otimes x_1^2 \otimes y_2^2$ 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 (13) 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 (13) has dimension four. Finally, taking $E'_{110} = \langle x_1^1 \otimes y_1^2, x_1^2 \otimes y_1^2 \rangle$, by symmetry (swapping the roles of U^* and W, which corresponds to taking transpose), the image of the (120)-version of (13) must have dimension four and we conclude.

7. Proofs of Theorems 1.4 and 1.5

Let $E'_{110} \subset U^* \otimes \mathfrak{sl}(V) \otimes W$ be a $\mathbb{B}$-fixed subspace. Define the *outer structure* of E'_{110} to be the set of $\mathfrak{sl}(U) \oplus \mathfrak{sl}(W)$ weights appearing in E'_{110}, counted with multiplicity. We identify the $\mathfrak{sl}(U)$ weights of U^* and the $\mathfrak{sl}(W)$ weights of W each with $\{1, \ldots, \mathbf{n}\}$, where 1 corresponds to the highest weight. In this way we consider the outer structure of E'_{110} as an $\mathbf{n} \times \mathbf{n}$ grid, with each grid point labelled by the dimension of the corresponding weight space. In what follows, we will represent such filled grids by the corresponding Young diagrams on the nonzero labels, where the upper left box corresponds with the highest weight. Here, labels weakly decrease going to the right and down. We speak of the *inner structure* of E'_{110} to be the particular $\mathfrak{sl}(V)$-weight spaces which occur at each weight $(s,t) \in \mathbf{n} \times \mathbf{n}$. The set of possible inner structures over a grid point (s,t) corresponds to the set of $\mathbb{B}$-fixed subspaces of $\mathfrak{sl}(V)$ that are contained in or equal to the chosen $\mathbb{B}$-fixed subspaces at sites $(s-1,t)$ and $(s,t-1)$.

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 $\Sigma_g \subset E'_{110}$ be an admissible filtration, and put

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

Then the dimension of (14) can be written as the sum over g of $\dim K_g/K_{g-1}$, and we may upper bound the dimension of (14) by upper bounding each $\dim K_g/K_{g-1}$. We obtain bounds on $\dim K_g/K_{g-1}$ which depend only on s and $j := \dim \Sigma_g/\Sigma_{g-1}$. For $\mathfrak{sl}_2$, this is Lemma 7.1, and for $\mathfrak{sl}_3$, this is Lemma 7.3. Bounds on the kernel of the (120) map are obtained by symmetry; specifically, the bound is the same as that on (14) with s replaced by t.

These lemmas reduce the problem to a combinatorial optimization problem over possible outer structures of fixed total dimension. In particular, 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 7.6.

7.1. **The local argument.**

Lemma 7.1. *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 the differences in the dimensions of the kernels of* (13) *with Σ_g and Σ_{g-1} in the place of E'_{110} equals the function $a_j s + b_j$ where*

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

Lemma 7.1 is proved later this section.

Remark 7.2. 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 $\ytableaushort{2}$, $\ytableaushort{1,1}$, $\ytableaushort{11}$, which, according to Lemma 7.1, have (210) map kernel dimensions 5, 4, and 4, respectively, all of which are smaller than 6. This gives our shortest proof that $\underline{\mathbf{R}}(M_{\langle 2\rangle}) > 6$.

Proof of Theorem 1.3. 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 $\ytableaushort{3}$, $\ytableaushort{21}$, $\ytableaushort{111}$, and $\ytableaushort{2,1}$. From Lemma 7.1, the corresponding (210) map kernel dimensions are 8, 7, 6, and 9, respectively, so only $\ytableaushort{2,1}$ passes. However, $\ytableaushort{2,1}$ has (120) kernel dimension 8, and fails this test. □

Proof of Theorem 1.4(1),(2). For Theorem 1.4(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 $\ytableaushort{111,1}$, $\ytableaushort{11,11}$, $\ytableaushort{11,1,1}$, $\ytableaushort{211}$, $\ytableaushort{22}$, $\ytableaushort{2,1,1}$, $\ytableaushort{21,1}$, $\ytableaushort{2,2}$, $\ytableaushort{31}$, $\ytableaushort{3,1}$. Of these, $\ytableaushort{11,1,1}$, $\ytableaushort{2,1,1}$, $\ytableaushort{2,2}$, and $\ytableaushort{3,1}$ pass the (210) test with kernel dimensions of size 14, 16, 15, and 14, respectively. However, none of these pass the (120) test (this can be seen as none appear in this list whose conjugate tableau also appear).

For Theorem 1.4(2), the result follows by similar complete enumeration of outer structures on a computer. □

Lemma 7.3. *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). *The differences in the dimensions of the kernels of* (13) *with* Σ_g *and* Σ_{g-1} *in the place of* E'_{110} *is bounded above by a function* $a_j s + b_j$ *where*

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 7.1 and 7.3, we first observe the following:

Proposition 7.4. *The included module* $V_{\omega_1} \subset V \otimes \mathfrak{sl}(V)$ *has weight basis* $\overline{v}_i = \sum_{j\neq i} [\mathbf{v} v_j \otimes (v_i \otimes v^j) - v_i \otimes (v_j \otimes v^j)] + (\mathbf{v}-1) v_i \otimes v_i \otimes v^i$, $1 \le i \le \mathbf{v}$.

Proof. The line $[\overline{v}_1]$ has weight ω_1 and is $\mathbb{B}$-stable, and the span of the $\overline{v}_j$ is fixed under the action of $\mathrm{SL}(V)$. □

Proof of Lemmas 7.1 and 7.3. We begin in somewhat greater generality, not fixing $\mathbf{v} = \dim V$. We must bound $\dim K_g - \dim K_{g-1}$, K_g given by (15). Write

$$K = U^{*\otimes 2} \otimes V_{\omega_1} \otimes W \oplus S^2 U^* \otimes V_{2\omega_1 + \omega_{\mathbf{v}-1}} \otimes W \oplus \Lambda^2 U^* \otimes V_{\omega_2 + \omega_{\mathbf{v}-1}} \otimes W \tag{16}$$

so that $K_g = \Sigma_g \otimes A \cap K$. Write $X \subset \mathfrak{sl}(V)$ 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 $V_0 = \varnothing$, $V_1 = V_{\omega_1}$, $V_2 = V_{\omega_1} \oplus V_{2\omega_1+\omega_{\mathbf{v}-1}}$, and $V_3 = V_{\omega_1} \oplus V_{2\omega_1+\omega_{\mathbf{v}-1}} \oplus V_{\omega_2+\omega_{\mathbf{v}-1}} = V \otimes \mathfrak{sl}(V)$. Note that $V_2 = V_3$ when $\mathbf{v} = 2$. Then $\{V_f\}_f$ is a flag of $V \otimes \mathfrak{sl}(V)$, 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 of $U^{*\otimes 2} \otimes V_3 \otimes W$, where we have written $U^{*s} = \mathrm{span}\{u^{\mathbf{n}}, \dots, u^{\mathbf{n}-s+1}\}$ and $W_{(t-1)} = \mathrm{span}\{w_1, \dots, w_t\}$. Hence, $S_f \cap K_g$ is a flag of K_g with $K_{g-1} = S_0 \cap K_g$. Use the isomorphism

$$\frac{K_g \cap S_f}{K_g \cap S_{f-1}} = \frac{K_g \cap S_f + S_{f-1}}{S_{f-1}} \tag{17}$$

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

$$\frac{U^{*\otimes 2} \otimes V_3 \otimes W}{S_{f-1}} = \frac{U^{*\otimes 2}}{U^* \otimes U^{*(s-1)}} \otimes \frac{V_3}{V_{f-1}} \otimes \frac{W}{W_{(t-1)}}. \tag{18}$$

Write K^f for the f-th summand of (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 $K_g \cap S_f + S_{f-1} = (K^f + S_{f-1}) \cap (\Sigma_g \otimes A) + S_{f-1} = (K^f + S_{f-1}) \cap (U^* \otimes u^{\mathbf{n}-s+1} \otimes V \otimes X \otimes w_t + S_{f-1})$. We may now pass in each side of the intersection to the right hand side of (18), 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 $\overline{S^2 U^*} \cap \overline{U^* \otimes u^{\mathbf{n}-s+1}} = \overline{U^{*s} \otimes u^{\mathbf{n}-s+1}}$ and $\overline{\Lambda^2 U^*} \cap \overline{U^* \otimes u^{\mathbf{n}-s+1}} = \overline{U^{*(s-1)} \otimes u^{\mathbf{n}-s+1}}$.

Therefore, the right hand side of (17) 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 $Y = (V \otimes X) \cap V_1$, $Y' = ((V \otimes X + V_1) \cap V_2)/V_1$, and $Y'' = (V \otimes X + V_2)/V_2$. We obtain $\dim K_g = \dim K_{g-1} + \mathbf{yn} + \mathbf{y}'s + \mathbf{y}''(s-1)$, the sum of the successive quotient dimensions of $\{T_f \cap K_g\}_f$.

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 7.4, $Y = \mathrm{span}\{\overline{v}_i \mid i \leq \mathbf{y}\}$.

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 2.3, 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

$$\dim K_g - \dim K_{g-1} = (j\mathbf{v} - \mathbf{y})s + \mathbf{yn} - \mathbf{y}'' \leq (j\mathbf{v} - \mathbf{y})s + \mathbf{yn}, \tag{19}$$

as $\mathbf{y} + \mathbf{y}' + \mathbf{y}'' = j\mathbf{v}$. The only remaining 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 $V \otimes \mathfrak{sl}(V) \oplus V_{\omega_1} = V \otimes V \otimes V^* = S^2 V \otimes V^* \oplus \Lambda^2 V \otimes V^*$ 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 (19), 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 7.5. 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 7.6 below the linear functions of s in the lemmas above appear as $a_{\mu_{s,t}} s + b_{\mu_{s,t}}$.

7.2. **The globalization.** Write μ for a Young diagram filled with non-negative 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, we take such μ to correspond to outer structures.

Lemma 7.6. *Fix $k \in \mathbb{N}$, $0 \le a_1 \le \cdots \le a_k$, and $b_i \in \mathbb{R}$, $1 \le i \le k$. Let μ be a Young diagram filled with labels in the set $\{1, \dots, k\}$, non-increasing in rows and columns. Write $\rho = \sum_{s,t} \mu_{s,t}$. Then*

$$\min\Big\{\sum_{s,t} a_{\mu_{s,t}} s + b_{\mu_{s,t}}, \sum_{s,t} a_{\mu_{s,t}} t + b_{\mu_{s,t}}\Big\} \le \max_{1\le j\le k}\left\{\frac{a_j\rho^2}{8j^2} + (a_j + b_j)\frac{\rho}{j}\right\}. \tag{20}$$

Remark 7.7. 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 7.6 is proved in §7.3.

Proof of Theorem 1.4(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 7.6 with $k = 3$ and a_i and b_i from Lemma 7.1 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 1.5. Proceeding in the same way as in the proof of Theorem 1.4(3), we apply Lemma 7.6 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 7.3. We obtain the smaller of the kernel dimensions of the (120) and (210) maps is at most the largest of the following,

j	Lemma 7.6	j	Lemma 7.6
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	$\frac{21}{512}\rho^2 + \frac{15+3\mathbf{n}}{8}\rho.$

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 fact for $j = 1$ follows as $\sqrt{\frac{8}{3}}\mathbf{n}$ is irrational. This fact for $2 \le j \le 8$ follows from the restriction on $\mathbf{n}$. Hence, at least one of the kernels of the (120) and (210) maps is too small, and $\underline{\mathbf{R}}(M_{\langle \mathbf{3nn}\rangle}) > \mathbf{n}^2 + \rho$, as required. □

7.3. **Proof of Lemma 7.6.** For a partition $\lambda = (\lambda_1, \cdots, \lambda_q)$, write $\ell(\lambda) = q$ and $n(\lambda) = \sum_i (i-1)\lambda_i$. Let λ' denote the conjugate partition. We remark that the results in this section may be used for $M_{\langle\mathbf{mnn}\rangle}$ for any $\mathbf{n} \ge \mathbf{m}$.

To establish Lemma 7.6 we need two additional lemmas:

Lemma 7.8. *Let λ be a partition not of the form $(n,2)$. 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 $n(\lambda) = \binom{\lambda_1'}{2} = \frac{1}{2}(\lambda_1' - \frac{1}{2})^2 - \frac{1}{8} = \frac{1}{8}(|\lambda| + \lambda_1' - \lambda_1)^2 - \frac{1}{8}$, 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.

$$\begin{aligned} n(\lambda) &= n(\mu) + \binom{\lambda_k'}{2} \\ &\le \tfrac{1}{8}(|\mu| + \mu_1' - \mu_1)^2 - \tfrac{1}{8} + \binom{\lambda_k'}{2} \\ &= \tfrac{1}{8}(|\lambda| - \lambda_k' + \lambda_1' - (\lambda_1 - 1))^2 - \tfrac{1}{8} + \tfrac{1}{2}\lambda_k'(\lambda_k' - 1) \\ &= \tfrac{1}{8}(|\lambda| + \lambda_1' - \lambda_1)^2 - \tfrac{1}{8} - \tfrac{1}{4}(|\lambda| + \lambda_1' - \lambda_1 - \tfrac{5}{2}\lambda_k' + \tfrac{1}{2})(\lambda_k' - 1) \end{aligned}$$

We must show the right hand term is non-positive. If $\lambda_k' = 1$, this is immediate; otherwise, we show the first factor is nonnegative. We have $|\lambda| - \lambda_1 \ge k\lambda_k' - k$, so $|\lambda| + \lambda_1' - \lambda_1 - \frac{5}{2}\lambda_k' + \frac{1}{2} \ge (\lambda_1' - \lambda_k') + \frac{2k-3}{2}(\lambda_k' - 1) - 1$. If $k = 2$, then by assumption $\lambda_1' \ge 3$, and considering separately the cases $\lambda_2' = 2$ and $\lambda_2' \ge 3$ yields that the first factor is nonnegative. Otherwise $k \ge 3$, and because $\lambda_k' \ge 2$, the first factor is nonnegative. This completes the proof. □

Lemma 7.9. *Fix $k \in \mathbb{N}$, $c_i \ge 0$, $d_i \in \mathbb{R}$, for $1 \le i \le k$. 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 $x_1 \ge \cdots \ge x_k \ge 0$, $y_1 \ge \cdots \ge y_k \ge 0$, and $\sum_i x_i + y_i = \rho$, the following inequality holds:*

$$\min\Big\{\sum_{i\le k} c_i x_i^2 + d_i(x_i + y_i), \sum_{i\le k} c_i y_i^2 + d_i(x_i + y_i)\Big\} \le \max_{1\le j\le k}\left\{\frac{\rho^2}{4j^2}C_j + \frac{\rho}{j}D_j\right\}. \tag{21}$$

Remark 7.10. The maximum is achieved when $x_1 = \cdots = x_j = y_1 = \cdots = y_j = \frac{\rho}{2j}$ and $x_s, y_s = 0$ for $s > j$, for some 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 $x_i', y_i' \ge 0$ 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 (21) 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 summation to ρ only if $\epsilon_1 + \epsilon_2 + \epsilon_3 = 0$, so we require this. Substitute these values into $E_L := \sum_{i\le k} c_i x_i^2 + d_i(x_i + y_i)$ and $E_R := \sum_{i\le k} c_i y_i^2 + d_i(x_i + y_i)$. View E_L, E_R as two 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 space $\langle \epsilon_1, \epsilon_2, \epsilon_3\rangle / \langle \sum \epsilon_j = 0\rangle$. 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 the same non-negativity 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.

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 $\sum_{i\leq k} c_i x_i^2 + d_i(x_i + y_i) = \sum_{i\leq k} c_i y_i^2 + d_i(x_i + y_i)$, from which it follows that $\alpha^2 C_s = \sum_{i\leq k} c_i x_i^2 = \sum_{i\leq k} c_i y_i^2 = \beta^2 C_t$ 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

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

By the arithmetic mean-harmonic mean inequality, we have

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

so that

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

with the last inequality from the fact that $\frac{s\alpha}{\rho} + \frac{t\beta}{\rho} = 1$. Multiplying both sides by $\frac{\rho}{s\sqrt{C_t} + t\sqrt{C_s}}$, we conclude the optimal value is achieved at one of the claimed values. □

Proof of Lemma 7.6. 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

$$\begin{aligned} \textstyle\sum_{s,t} a_{\mu_{s,t}} s + b_{\mu_{s,t}} &= \textstyle\sum_{s,t} \sum_{i=1}^{\mu_{s,t}} (a_i - a_{i-1}) s + b_i - b_{i-1} \\ &= \textstyle\sum_{i=1}^{k} \sum_{s,t\in\lambda^i} (a_i - a_{i-1}) s + b_i - b_{i-1} \\ &= \textstyle\sum_{i=1}^{k} (a_i - a_{i-1}) n(\lambda^i) + (a_i - a_{i-1} + b_i - b_{i-1}) |\lambda^i| \\ &\leq \textstyle\sum_{i=1}^{k} \left[\tfrac{1}{2}(a_i - a_{i-1})\right] \left(\tfrac{1}{2}(|\lambda^i| + (\lambda^i)'_1 - \lambda^i_1)\right)^2 + \left[a_i - a_{i-1} + b_i - b_{i-1}\right] |\lambda^i| \qquad (22) \end{aligned}$$

where we have used Lemma 7.8 to obtain the last inequality. Set

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

Then (22) becomes

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

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

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

as required. □

8. Proof that $\underline{\mathbf{R}}(M_{\langle \mathbf{l},\mathbf{m},\mathbf{n}\rangle}) \ge \underline{\mathbf{R}}(M_{\langle \mathbf{l}-1,\mathbf{m},\mathbf{n}\rangle}) + 1$.

Here is a simple proof of the statement, which was originally shown in [25]. By the border substitution method [22], for any tensor $T \in A \otimes B \otimes C$

$$\underline{\mathbf{R}}(T) \ge \min_{A' \subset A^*} \underline{\mathbf{R}}(T|_{A' \otimes B^* \otimes C^*}) + 1,$$

where $A' \subset A^*$ is a hyperplane. Moreover, if T has symmetry group G_T, and G_T has a unique closed orbit in $\mathbb{P}A^*$, then we may restrict A' to be a point of that closed orbit by the Normal Form Lemma of [22]. In the case of matrix multiplication, $G_{M_{\langle \mathbf{l},\mathbf{m},\mathbf{n}\rangle}} \supset \mathrm{SL}(U) \times \mathrm{SL}(V)$ can degenerate any point in $\mathbb{P}A = \mathbb{P}(U^* \otimes V)$ to the annihilator of $x^1_{\mathbf{1}}$, so it amounts to taking $T|_{A' \otimes B^* \otimes C^*}$ to be the reduced matrix multiplication tensor with $x^1_{\mathbf{1}} = 0$. But now we may (using $\mathrm{GL}(A) \times \mathrm{GL}(B) \times \mathrm{GL}(C)$) degenerate this tensor further to set all $x^i_{\mathbf{1}}$ and $y^{\mathbf{1}}_j$ to zero to obtain the result.

References

1. J. Alman and V. Vassilevska Williams, *Limits on All Known (and Some Unknown) Approaches to Matrix Multiplication*, 2018 IEEE 59th Annual Symposium on Foundations of Computer Science (2018).
2. Josh Alman and Virginia 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, USA, 2018, pp. 25:1–25:15.
3. Andris Ambainis, Yuval Filmus, and François 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, pp. 585–593. MR 3388238
4. Sanjeev Arora and Boaz Barak, *Computational complexity*, Cambridge University Press, Cambridge, 2009, A modern approach. MR 2500087 (2010i:68001)
5. D. Bini, *Relations between exact and approximate bilinear algorithms. Applications*, Calcolo **17** (1980), no. 1, 87–97. MR 605920 (83f:68043b)
6. Dario Bini, Grazia Lotti, and Francesco Romani, *Approximate solutions for the bilinear form computational problem*, SIAM J. Comput. **9** (1980), no. 4, 692–697. MR MR592760 (82a:68065)
7. Markus Bläser, *Fast matrix multiplication*, Graduate Surveys, no. 5, Theory of Computing Library, 2013.
8. Mats Boij and Zach Teitler, *A bound for the Waring rank of the determinant via syzygies*, Linear Algebra and its Applications **587** (2020), 195 – 214.

9. Weronika Buczyńska and Jarosław Buczyński, *Apolarity, border rank and multigraded Hilbert scheme*, arXiv:1910.01944.
10. Peter Bürgisser, Michael Clausen, and M. Amin 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
11. Matthias Christandl, Péter Vrana, and Jeroen Zuiddam, *Barriers for fast matrix multiplication from irreversibility*, CoRR **abs/1812.06952** (2018).
12. Austin Conner, Fulvio Gesmundo, J.M. Landsberg, and Emanuele Ventura, *Kronecker powers of tensors and Strassen's laser method*, arXiv:1909.04785.
13. Austin Conner, Fulvio Gesmundo, Joseph M. Landsberg, and Emanuele Ventura, *Tensors with maximal symmetries*, arXiv e-prints (2019), arXiv:1909.09518.
14. Klim Efremenko, Ankit Garg, Rafael Oliveira, and Avi 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, pp. Art. No. 1, 19. MR 3761737
15. Cameron Farnsworth, *Koszul-Young flattenings and symmetric border rank of the determinant*, J. Algebra **447** (2016), 664–676. MR 3427655
16. Maciej Gałazka, *Vector bundles give equations of cactus varieties*, Linear Algebra Appl. **521** (2017), 254–262. MR 3611482
17. Matteo Gallet, Kristian Ranestad, and Nelly Villamizar, *Varieties of apolar subschemes of toric surfaces*, Ark. Mat. **56** (2018), no. 1, 73–99. MR 3800460
18. Anthony Iarrobino and Vassil Kanev, *Power sums, Gorenstein algebras, and determinantal loci*, Lecture Notes in Mathematics, vol. 1721, Springer-Verlag, Berlin, 1999, Appendix C by Iarrobino and Steven L. Kleiman. MR MR1735271 (2001d:14056)
19. J. M. Landsberg, *The border rank of the multiplication of* 2×2 *matrices is seven*, J. Amer. Math. Soc. **19** (2006), no. 2, 447–459. MR 2188132 (2006j:68034)
20. ______, *Geometry and complexity theory*, Cambridge Studies in Advanced Mathematics, vol. 169, Cambridge University Press, Cambridge, 2017. MR 3729273
21. J. M. Landsberg, *Tensors: Asymptotic Geometry and Developments 20162018*, CBMS Regional Conference Series in Mathematics, vol. 132, AMS, 2019.
22. J. M. Landsberg and Mateusz Michałek, *On the geometry of border rank decompositions for matrix multiplication and other tensors with symmetry*, SIAM J. Appl. Algebra Geom. **1** (2017), no. 1, 2–19. MR 3633766
23. J. M. Landsberg and Giorgio Ottaviani, *Equations for secant varieties of Veronese and other varieties*, Ann. Mat. Pura Appl. (4) **192** (2013), no. 4, 569–606. MR 3081636
24. Joseph M. Landsberg and Giorgio Ottaviani, *New lower bounds for the border rank of matrix multiplication*, Theory Comput. **11** (2015), 285–298. MR 3376667
25. Thomas Lickteig, *A note on border rank*, Inform. Process. Lett. **18** (1984), no. 3, 173–178. MR 86c:68040
26. I. G. Macdonald, *Symmetric functions and Hall polynomials*, second ed., Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 1995, With contributions by A. Zelevinsky, Oxford Science Publications. MR 1354144 (96h:05207)
27. Claudio Procesi, *Lie groups*, Universitext, Springer, New York, 2007, An approach through invariants and representations. MR MR2265844 (2007j:22016)
28. A. V. Smirnov, *The bilinear complexity and practical algorithms for matrix multiplication*, Comput. Math. Math. Phys. **53** (2013), no. 12, 1781–1795. MR 3146566
29. V. Strassen, *Rank and optimal computation of generic tensors*, Linear Algebra Appl. **52/53** (1983), 645–685. MR 85b:15039
30. Volker Strassen, *Gaussian elimination is not optimal*, Numer. Math. **13** (1969), 354–356. MR 40 #2223
31. Zach Teitler, *Geometric lower bounds for generalized ranks*, arXiv e-prints (2014), arXiv:1406.5145.

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

E-mail address, A. Conner: `connerad@math.tamu.edu`

E-mail address, A. Harper: `adharper@math.tamu.edu`

E-mail address, J.M. Landsberg: `jml@math.tamu.edu`