NEW RESULTS IN THE COMPLEXITY OF MATRIX MULTIPLICATION

A Dissertation

by

AUSTIN DANIEL CONNER

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Joseph M. Landsberg
Committee Members,	Laura F. Matusevich
	Eric C. Rowell
	Jianer Chen
Head of Department,	Sarah Witherspoon

May 2020

Major Subject: Mathematics

ABSTRACT

Determining the complexity of matrix multiplication has been a central problem in complexity theory ever since Strassen showed, in 1969, that one can multiply matrices in $O(\mathbf{n}^{2.81})$ arithmetic operations, strictly better than with the usual algorithm. Bini reduced the problem of the complexity of matrix multiplication to one in multilinear algebra, that of determining the border rank of the matrix multiplication tensor. In this thesis, I prove new border rank bounds, both upper and lower, on certain matrix multiplication tensors as well as on the little Coppersmith-Winograd tensor and its recently introduced skew variant, auxiliary tensors relevant to the study via Strassen's laser method. Upper bounds are obtained through explicit rank and border rank decompositions. The lower bounds are are obtained principally through representation theory, both of finite and Lie groups. In particular, I present new results for matrix multiplication coming from a recent development in lower bounds due to Buczyńska and Buczyński, the idea of border apolarity.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professor Joseph M. Landsberg, advisor, Professor Laura F. Matusevich, and Professor Eric C. Rowell of the Department of Mathematics, and Professor Jianer Chen of the Department of Computer Science.

The results of chapter 2 were obtained in collaboration with Joseph M. Landsberg, Emanuele Ventura, and Fulvio Gesmundo. The results of chapter 3 were obtained in collaboration with Joseph M. Landsberg and Alicia Harper. The results of chapter 4 is the independent work of the author.

Funding Sources

This material is based upon work supported by the National Science Foundation under Grant No. CCF-1814254.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

TABLE OF CONTENTS

LIST OF FIGURES

1. INTRODUCTION AND BACKGROUND

1.1 The complexity of matrix multiplication

Linear algebra is central to applications of mathematics, and matrix multiplication is the essential operation of linear algebra. The standard algorithm to multiply two $\mathbf{n}\times\mathbf{n}$ matrices uses $\mathbf{n}^3$ multiplications. In 1969, while attempting to show that the standard algorithm was optimal, V. Strassen [1] discovered an explicit algorithm to multiply 2×2 matrices using seven multiplications rather than eight. This algorithm may also be used to multiply $\mathbf{n}\times\mathbf{n}$ matrices using $O(\mathbf{n}^{2.81})$ arithmetic operations rather than the usual $O(\mathbf{n}^3)$ (see §1.2).

The exponent ω of matrix multiplication is defined as

$$\omega := \inf\{\tau \mid \mathbf{n}\times\mathbf{n} \text{ matrices may be multiplied using } O(\mathbf{n}^\tau) \text{ arithmetic operations}\}.$$

Trivially, $\omega \geq 2$, as any matrix multiplication must at least look at all $\mathbf{n}^2$ entries of the matrices to be multiplied, and it is conjectured that, in fact, $\omega = 2$. There was steady progress in the research for upper bounds from 1968 to 1988: after Strassen's famous $\omega < 2.81$, Bini et. al. [2], using border rank (see §1.4), showed $\omega < 2.78$, then a major breakthrough by Schönhage [3] (the asymptotic sum inequality, §1.6) was used to show $\omega < 2.55$, and then Strassen's *laser method* was introduced and used by Strassen to show $\omega < 2.48$, and refined by Coppersmith and Winograd to show $\omega < 2.3755$ [4] (§1.7). Then there was no progress until 2011 when a series of improvements by Stothers, Williams, and Le Gall [5, 6, 7] lowered the upper bound to the current state of the art $\omega < 2.373$.

1.2 Strassen's algorithm

Consider the product of a pair of $2\mathbf{n} \times 2\mathbf{n}$ matrices, where we have blocked them into $\mathbf{n} \times \mathbf{n}$ quadrants,

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}.$$

Computing the C_{ij} blockwise via the usual algorithm for 2×2 matrix multiplication takes eight block multiplications:

$$\begin{aligned} C_{11} &= A_{11}B_{11} + A_{12}B_{21} \\ C_{12} &= A_{11}B_{12} + A_{12}B_{22} \\ C_{21} &= A_{21}B_{11} + A_{22}B_{21} \\ C_{21} &= A_{21}B_{12} + A_{22}B_{22}, \end{aligned}$$

If, however, we instead compute

$$\begin{aligned} I &= (A_{11} + A_{22})(B_{11} + B_{22}) \\ II &= A_{11}(B_{12} + B_{22}) \\ III &= (A_{21} - A_{22})B_{11} \\ IV &= (A_{12} + A_{22})(B_{21} - B_{22}) \\ V &= A_{22}(B_{11} + B_{21}) \\ VI &= (-A_{11} + A_{12})B_{22} \\ VII &= (A_{11} + A_{21})(-B_{11} + B_{12}), \end{aligned} \tag{1.1}$$

we may form

$$\begin{aligned} C_{11} &= I + IV - V + VI \\ C_{12} &= II + VI \\ C_{21} &= III + V \\ C_{22} &= I - II + III + VII. \end{aligned} \tag{1.2}$$

That is, we may compute the product of a pair of matrices using only seven block multiplications. This is enough to obtain an algorithm for square matrix multiplication which uses asymptotically fewer arithmetic operations than the standard algorithm.

Let A and B be $\mathbf{n} \times \mathbf{n}$ matrices. First, if $\mathbf{n} = 1$, return the product. Otherwise, if $\mathbf{n}$ is odd, reduce to the even case by padding each of A and B with a row and column of zeroes. Divide A and B into quadrants as above, and compute the products I–VII by recursively calling this algorithm. Finally, form the C_{ij} as above and return the corresponding C.

We analyze the above algorithm. Let $T(\mathbf{n})$ denote the number of arithmetic operations required to carry out the above algorithm on $\mathbf{n}\times\mathbf{n}$ matrices. The formation of the appropriate linear combinations of blocks to compute the arguments to the recursive calls, and then to compute the C_{ij} from the results takes $O(\mathbf{n}^2)$ arithmetic operations. Then, we have $T(1) = 1$, $T(\mathbf{n}) = 7T(\lceil \frac{\mathbf{n}}{2} \rceil) + O(\mathbf{n}^2)$. Applying the master theorem of [8], we obtain $T(\mathbf{n}) = O(\mathbf{n}^{\log_2 7})$, where $\log_2 7 \cong 2.81$.

1.3 Reduction to tensor rank

The existence of Strassen's algorithm depends only on the existence of the equations 1.1 and 1.2. These equations have a special structure such that a corresponding algorithm for fast matrix multiplication can be derived. We will describe this special structure in geometric terms as a property of matrix multiplication considered as a tensor. The notion we define will be independent of coordinates, so our presentation will be in the coordinate free language of composition of linear maps, rather than multiplication of matrices.

In what follows, we will make extensive use of finite dimensional vector spaces over $\mathbb{C}$. We denote such vector spaces with capital letters and by convention use the corresponding bold lower case letter to denote dimension. Define the tensor product $U \otimes V$ as the $\mathbb{C}$-vector space with basis $\{(u,v) \mid u \in U, v \in V\}$ modulo the relations

$$
\begin{aligned}
(u+u',v) &= (u,v) + (u',v) \\
(u,v+v') &= (u,v) + (u,v') \\
(au,v) &= a(u,v) = (u,av),
\end{aligned}
$$

where $u, u' \in U$, $v, v' \in V$, and $a \in \mathbb{C}$. We write $u \otimes v$ for the equivalence class of (u,v) in $U \otimes V$. Then, for instance, $\dim U \otimes V = \mathbf{uv}$. Any bilinear map $\phi : U \times V \to W$ satisfies $\phi(u,v) = \widehat{\phi}(u \otimes v)$ for some unique linear map $\widehat{\phi} : U \otimes V \to W$, and $U \otimes V$ is characterized by this property.

Denote by $\mathrm{Hom}(U,V)$ the vector space of linear maps $U \to V$, and write U^* for $\mathrm{Hom}(U,\mathbb{C})$. When U and V are finite dimensional, the natural inclusion $V \otimes U^* \to \mathrm{Hom}(U,V)$, $v \otimes f \mapsto [u \mapsto vf(u)]$ is an isomorphism. Write $\mathrm{Id}_U \in \mathrm{Hom}(U,U) = U \otimes U^*$ as the identity operator under this identification. Similarly, the natural inclusion $V \to (V^*)^*$, $v \mapsto [f \mapsto f(v)]$ is an isomorphism. Finally the map $(\mathrm{Hom}(U,V) \otimes \mathrm{Hom}(V,U))^*$, $A \otimes B \mapsto \mathrm{trace}(AB)$ is a perfect pairing, yielding a natural isomorphism $\mathrm{Hom}(U,V)^* \to \mathrm{Hom}(V,U)$. We will freely identify spaces under these isomorphisms.

Linear maps $W \to V$ and $V \to U$ may be composed, and the composition operator is bilinear. In other words, there is a tensor $M_{\langle \mathbf{u},\mathbf{v},\mathbf{w} \rangle} \in \mathrm{Hom}(\mathrm{Hom}(V,U) \otimes \mathrm{Hom}(W,V), \mathrm{Hom}(W,U))$. This tensor is called the *matrix multiplication tensor*. There are additional ways to conceive of this tensor, modulo the identifications above. For instance, under the natural isomorphism to the space $(\mathrm{Hom}(V,U) \otimes \mathrm{Hom}(W,V) \otimes \mathrm{Hom}(U,W))^*$, $M_{\langle \mathbf{u},\mathbf{v},\mathbf{w} \rangle}$ has the form $A \otimes B \otimes C \mapsto \mathrm{trace}(ABC)$. Under the natural isomorphism to the space $(U \otimes U^*) \otimes (V \otimes V^*) \otimes (W \otimes W^*)$,

$M_{\langle \mathbf{u},\mathbf{v},\mathbf{w}\rangle} = \mathrm{Id}_U \otimes \mathrm{Id}_V \otimes \mathrm{Id}_W$. We will most frequently consider $M_{\langle \mathbf{u},\mathbf{v},\mathbf{w}\rangle}$ to lie in the space $(U^* \otimes V) \otimes (V^* \otimes W) \otimes (W^* \otimes U)$. The grouping of the terms is significant in what follows. We write $M_{\langle \mathbf{n}\rangle} = M_{\langle \mathbf{n},\mathbf{n},\mathbf{n}\rangle}$.

We define a notion of complexity of tensors, called *rank*, which, for the matrix multiplication tensor, corresponds up to a factor of two to the minimal number of multiplications needed to compute matrix multiplication via an arithmetic circuit [9, Equation 14.8]. More precisely, for a tensor $T \in A \otimes B \otimes C$, its *rank* $\mathbf{R}(T)$ is the smallest r such that $T = \sum_{j=1}^{r} a_j \otimes b_j \otimes c_j$ for some $a_j \in A$, $b_j \in B$, $c_j \in B$. We will call such an expression of T as a sum of rank one tensors a *rank decomposition*, even in cases where the number of terms is not minimal. For instance, $\mathbf{R}(M_{\langle 2,2,2\rangle}) \le 7$, as Strassen's algorithm (§1.2) may equivalently be written

$$
\begin{aligned}
M_{\langle 2\rangle} = {} & (u^1 v_1 + u^2 v_2) \otimes (v^1 w_1 + v^2 w_2) \otimes (w^1 u_1 + w^2 u_2) \\
& + u^1 v_1 \otimes (v^1 w_2 + v^2 w_2) \otimes (w^2 u_1 - w^2 u_2) \\
& + (u^2 v_1 - u^2 v_2) \otimes v^1 w_1 \otimes (w^1 u_2 + w^2 u_2) \\
& + (u^1 v_2 + u^2 v_2) \otimes (v^2 w_1 - v^2 w_2) \otimes w^1 u_1 \\
& + u^2 v_2 \otimes (v^2 w_1 + v^1 w_1) \otimes (w^1 u_2 - w^1 u_1) \\
& + (u^1 v_2 - u^1 v_1) \otimes v^2 w_2 \otimes (w^2 u_1 + w^1 u_1) \\
& + (u^2 v_1 + u^1 v_1) \otimes (v^1 w_2 - v^1 w_1) \otimes w^2 u_2.
\end{aligned}
\tag{1.3}
$$

Here we have written u_i, v_i, w_i for bases of U, V and W, $i \in \{1,2\}$, and u^i, v^i and w^i for the corresponding dual bases of U^*, V^* and W^*, and we have suppressed the tensor product sign for elements of $U^* \otimes V$, $V^* \otimes W$, and $W^* \otimes U$.

Given a rank r decomposition of $M_{\langle \mathbf{n}_0\rangle}$, there is a corresponding algorithm to multiply square $\mathbf{n} \times \mathbf{n}$ matrices. Namely, if $\mathbf{n} < \mathbf{n}_0$, return the matrix product via the usual algorithm. Otherwise pad with zero rows and columns to reduce to the case that $\mathbf{n}$ is divisible by $\mathbf{n}_0$, so that each matrix can be blocked into $\mathbf{n}_0^2$ blocks of dimensions $\lceil \frac{\mathbf{n}}{\mathbf{n}_0} \rceil \times \lceil \frac{\mathbf{n}}{\mathbf{n}_0} \rceil$. For each summand of the rank decomposition, form the appropriate corresponding linear combination of the blocks and recursively call the fast multiplication algorithm. Finally, form the blocks of the answer as appropriate linear combinations of the products as prescribed by the rank decomposition. As with Strassen's algorithm, the formation of all required linear combinations of blocks requires $O(\mathbf{n}^2)$ arithmetic operations, so if $T(\mathbf{n})$ is the number of arithmetic operations to multiply a pair of matrices via this algorithm, then $T(\mathbf{n})$ satisfies $T(\mathbf{n}) = rT(\lceil \frac{\mathbf{n}}{\mathbf{n}_0} \rceil) + O(\mathbf{n}^2)$, and by the master theorem, it follows $T(\mathbf{n}) = O(\mathbf{n}^{\log_{\mathbf{n}_0} r})$. Hence, if $\mathbf{R}(M_{\langle \mathbf{n} \rangle}) \leq r$, then $\omega \leq \log_{\mathbf{n}} r$. Strassen showed the converse holds in the following sense.

Theorem 1.3.1 (Strassen [1]). $\omega = \liminf_{\mathbf{n}\to\infty} \log_{\mathbf{n}} \mathbf{R}(M_{\langle \mathbf{n} \rangle})$

Hence, the study of the exponent of matrix multiplication can be entirely reduced to the study of the rank of the matrix multiplication tensor. In fact, the problem can be further reduced to a more geometric question. Unlike the matrix rank, the set of tensors in $A \otimes B \otimes C$ with tensor rank at most r does not form a closed set, either in the Euclidean or Zariski sense. There is thus an additional notion in the context of tensors, that of *border rank*.

1.4 Tensor border rank

For a tensor $T \in A \otimes B \otimes C$, its *border rank* $\underline{\mathbf{R}}(T)$ is the smallest r such that T lies in the closure of the set of rank at most r tensors, either in the Euclidean or Zariski sense. That is, the set of tensors of border rank at most r is precisely the cone over $\sigma_r(\mathrm{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C))$, the r-th secant variety of the Segre variety of rank one tensors. In fact, $\underline{\mathbf{R}}(T) \leq r$ if and only if there is an expression $T = \lim_{t\to 0} \sum_{i=1}^r a_i(t) \otimes b_i(t) \otimes c_i(t)$. We call such an expression for T a *border rank decomposition*. The set of tensors of border rank at most r is by definition a closed set. Thus, from the standpoint of geometry, border rank is a more natural measure of complexity. Moreover, from the standpoint of the complexity of matrix multiplication, there is no loss in generality, in view of the following.

Theorem 1.4.1 (Bini [10]). $\omega = \liminf_{\mathbf{n}\to\infty} \log_{\mathbf{n}} \underline{\mathbf{R}}(M_{\langle \mathbf{n} \rangle})$

One has $\underline{\mathbf{R}}(T) \leq \mathbf{R}(T)$ and the inequality can be strict: let $T = a_1 \otimes b_1 \otimes c_2 + a_1 \otimes b_2 \otimes c_1 + a_2 \otimes b_1 \otimes c_1$, then $\mathbf{R}(T) = 3$ and $\underline{\mathbf{R}}(T) = 2$ as

$$T = \lim_{t \to 0} \frac{1}{t}[(a_1 + ta_2) \otimes (b_1 + tb_2) \otimes (c_1 + tc_2) - a_1 \otimes b_1 \otimes c_1]. \tag{1.4}$$

In [11], it was shown that $\underline{\mathbf{R}}(M_{\langle 2 \rangle}) = 7$. Until the results of chapter 3, this was the only nontrivial matrix multiplication tensor whose border rank was known exactly. In chapter 3, we prove $\underline{\mathbf{R}}(M_{\langle 2,2,3 \rangle}) = 10$ and $\underline{\mathbf{R}}(M_{\langle 2,3,3 \rangle}) = 14$, expanding the list of tensors whose border rank is known from one to three, as well as giving other significant new bounds.

1.5 Waring (border) rank and symmetrized matrix multiplication

The symmetric group on k elements acts on the space $A^{\otimes k}$ by permuting factors, and we call tensors invariant under this action *symmetric*. The space of symmetric tensors is written $S^k(A) \subset A^{\otimes k}$. For symmetric tensors T, there is a more specialized notion of rank and border rank, namely, the *symmetric* or *Waring rank* $\mathbf{R}_s(T)$ is the smallest r so that T can be written as the sum of r symmetric rank one tensors, and we call such an expression a *Waring decomposition.* The *Waring border rank* $\underline{\mathbf{R}}_s(T)$ is defined analogously.

Define the symmetrized matrix multiplication tensor $sM_{\langle \mathbf{v} \rangle} \in S^3(V^* \otimes V)$ as the result of symmetrizing $M_{\langle \mathbf{v} \rangle} \in (V^* \otimes V)^{\otimes 3}$ under the action of the symmetric group on three elements. For the complexity of matrix multiplication, there is no loss in generality to study Waring (border) rank of $sM_{\langle \mathbf{n} \rangle}$, in view of the following.

Theorem 1.5.1 ([12])**.** $\omega = \liminf_{\mathbf{n}\to\infty} \log_{\mathbf{n}} \mathbf{R}_s(sM_{\langle \mathbf{n} \rangle}) = \liminf_{\mathbf{n}\to\infty} \log_{\mathbf{n}} \underline{\mathbf{R}}_s(sM_{\langle \mathbf{n} \rangle})$

In chapter 4, I show $\mathbf{R}_s(sM_{\langle 3 \rangle}) \leq 18$ and $\mathbf{R}_s(sM_{\langle 4 \rangle}) \leq 40$ by providing explicit Waring decompositions.

1.6 Asymptotic sum inequality

Schönhage's advance comes from his discovery that it can be more efficient to perform two matrix multiplications together than one at a time. For tensors $T \in A \otimes B \otimes C$ and $T' \in A' \otimes B' \otimes C'$, define a new tensor $T \oplus T' \in (A \oplus A') \otimes (B \oplus B') \otimes (C \otimes C')$ whose computation is equivalent to computing T and T'. He gave explicit examples of matrix multiplication tensors where $\underline{\mathbf{R}}(T \oplus T') \ll \underline{\mathbf{R}}(T) + \underline{\mathbf{R}}(T')$. To explain how he exploited this we need some more definitions:

Given $T \in A \otimes B \otimes C$ and $T' \in A' \otimes B' \otimes C'$, the *Kronecker product* of T and T' is the tensor $T \boxtimes T' := T \otimes T' \in (A \otimes A') \otimes (B \otimes B') \otimes (C \otimes C')$, regarded as 3-way tensor. Given $T \in A \otimes B \otimes C$, the *Kronecker powers* of T are $T^{\boxtimes N} \in A^{\otimes N} \otimes B^{\otimes N} \otimes C^{\otimes N}$, defined iteratively. We have $\mathbf{R}(T \boxtimes T') \leq \mathbf{R}(T)\,\mathbf{R}(T')$, and similarly for border rank. The matrix multiplication tensor has the following important self-reproducing property, corresponding to the fact that matrices may be multiplied block-wise: $M_{\langle \mathbf{l},\mathbf{m},\mathbf{n}\rangle} \boxtimes M_{\langle \mathbf{l}',\mathbf{m}',\mathbf{n}'\rangle} = M_{\langle \mathbf{l}\mathbf{l}',\mathbf{m}\mathbf{m}',\mathbf{n}\mathbf{n}'\rangle}$.

Given $T, T' \in A \otimes B \otimes C$, we say that T *degenerates* to T' if $T' \in \overline{\mathrm{GL}(A) \times \mathrm{GL}(B) \times \mathrm{GL}(C) \cdot T}$, the closure of the orbit of T under the natural action of $\mathrm{GL}(A) \times \mathrm{GL}(B) \times \mathrm{GL}(C)$ on $A \otimes B \otimes C$. Here, $\mathrm{GL}(A)$ denotes the general linear group of invertible linear maps $A \to A$. We extend this notion to tensors $T \in A \otimes B \otimes C$ and $T' \in A' \otimes B' \otimes C'$ by picking inclusions A and A' into a common space A'', and likewise for B and C. Border rank is upper semi-continuous under degeneration: if T' is a degeneration of T, then $\underline{\mathbf{R}}(T') \leq \underline{\mathbf{R}}(T)$.

Schönhage observed that if one takes a high Kronecker power of $M_{\langle \mathbf{l},\mathbf{m},\mathbf{n}\rangle} \oplus M_{\langle \mathbf{l}',\mathbf{m}',\mathbf{n}'\rangle}$, that because of the reproducing property, it will be a sum of matrix multiplication tensors, some of them quite large. One can then perform a degeneration to obtain a single very large matrix multiplication tensor and exploit the strict sub-additivity to get an upper bound on this large matrix multiplication tensor. This reasoning results in the celebrated

Theorem 1.6.1 (Shönhage's asymptotic sum inequality [3]). *For all* $\mathbf{l}_i$, $\mathbf{m}_i$, $\mathbf{n}_i$ *with* $1 \le i \le s$,

$$\sum_{i=1}^{s} (\mathbf{m}_i \mathbf{n}_i \mathbf{l}_i)^{\frac{\omega}{3}} \le \underline{\mathbf{R}} \Big(\bigoplus_{i=1}^{s} M_{\langle \mathbf{m}_i, \mathbf{n}_i, \mathbf{l}_i \rangle} \Big)$$

1.7 Strassen's laser method

After Schönhage, Strassen realized that the starting tensor need not be a sum of matrix multiplication tensors, as long as some high power of it degenerates to a large matrix multiplication tensor. This gave rise to his *laser method*, where the starting tensor "resembles" the sum of disjoint matrix multiplication tensors. All upper bounds since 1984 are obtained via Strassen's laser method. The best starting tensor for Strassen's method (so far) was discovered by Coppersmith and Winograd, the *big Coppersmith-Winograd tensor*,

$$\begin{aligned} T_{CW,q} &:= \sum_{j=1}^{q} a_0 \otimes b_j \otimes c_j + a_j \otimes b_0 \otimes c_j + a_j \otimes b_j \otimes c_0 \\ &\quad + a_0 \otimes b_0 \otimes c_{q+1} + a_0 \otimes b_{q+1} \otimes c_0 + a_{q+1} \otimes b_0 \otimes c_0 \in (\mathbb{C}^{q+2})^{\otimes 3}. \end{aligned}$$

It was used to obtain the current world record $\omega < 2.373$ and all bounds below $\omega < 2.41$ [4].

In 2014 [13] gave an explanation for the limited progress since 1988, followed by further explanations in [14, 15, 16, 17]: there are limitations to the laser method applied to the big Coppersmith-Winograd tensor and other auxiliary tensors. These limitations are referred to as *barriers*. We are interested in two kinds of barriers: to proving the exponent is two, and barriers to proving the exponent is less than 2.3. In [13], it was shown that the big Coppersmith-Winograd tensor is subject to this second kind of barrier, that the laser method applied to it is not sufficient to prove $\omega < 2.3$.

The second best tensor for the laser method so far has been the *little Coppersmith-Winograd* *tensor*,

$$T_{cw,q} := \sum_{j=1}^{q} a_0 \otimes b_j \otimes c_j + a_j \otimes b_0 \otimes c_j + a_j \otimes b_j \otimes c_0 \in (\mathbb{C}^{q+1})^{\otimes 3}. \tag{1.5}$$

The laser method implies the following.

Theorem 1.7.1. *[4] For all k and q,*

$$\omega \le \log_q\left(\frac{4}{27}(\underline{\mathbf{R}}(T_{cw,q}^{\boxtimes k}))^{\frac{3}{k}}\right). \tag{1.6}$$

It is a fact that $\underline{\mathbf{R}}(T_{cw,q}) = q+2$ [4]. Applying Theorem 1.7.1 with $k=1$ and $q=8$ yields $\omega < 2.404$. In chapter 2, we address [18, Problem 9.8], which was motivated by Theorem 1.7.1: Is $\underline{\mathbf{R}}(T_{cw,q}^{\boxtimes 2}) < (q+2)^2$? We give an almost complete answer:

Theorem 1.7.2. *For all $q > 2$, $\underline{\mathbf{R}}(T_{cw,q}^{\boxtimes 2}) = (q+2)^2$, and $15 \le \underline{\mathbf{R}}(T_{cw,2}^{\boxtimes 2}) \le 16$.*

We also examine the Kronecker cube:

Theorem 1.7.3. *For all $q > 4$, $\underline{\mathbf{R}}(T_{cw,q}^{\boxtimes 3}) = (q+2)^3$.*

Proofs are given in §2.2.

Proposition 2.2.1 below, combined with the proofs of Theorems 1.7.3 and 1.7.2, implies

Corollary 1.7.4. *For all $q > 4$ and all N,*

$$\underline{\mathbf{R}}(T_{cw,q}^{\boxtimes N}) \ge (q+1)^{N-3}(q+2)^3,$$

and $\underline{\mathbf{R}}(T_{cw,4}^{\boxtimes N}) \ge 36 \times 5^{N-2}$.

Previously, in [19] it had been shown that $\underline{\mathbf{R}}(T_{cw,q}^{\boxtimes N}) \ge (q+1)^N + 2^N - 1$ for all q, N, whereas the bound in Corollary 1.7.4 is $(q+1)^N + 3(q+1)^{N-1} + 3(q+1)^{N-2} + (q+1)^{N-3}$.

Previous to this work one might have hoped to prove $\omega < 2.3$ simply by using the Kronecker square of, e.g., $T_{cw,7}$. Now, the smallest possible calculation to give a new upper bound on ω from a tensor that has been used in the laser method would be, e.g., to prove the fourth Kronecker power of a small Coppersmith-Winograd tensor achieves the lower bound of Corollary 1.7.4 (which we do not expect to happen). Of course, one could work directly with the matrix multiplication tensor, in which case the cheapest possible upper bound would come from proving the border rank of the 6×6 matrix multiplication tensor equaled its known lower bound of 69 from [20].

1.8 New tensor for the laser method not subject to barriers

In light of the bad news for upper bounds of §1.7, we have the following.

Problem 1.8.1 (AFL Challenge [13]). Find new tensors for Strassen's laser method not subject to known barriers.

One promising tensor is the following skew cousin of the little Coppersmith-Winograd tensor, first introduced in [21] and defined when q is even.

$$T_{skewcw,q} := \sum_{j=1}^{q} a_0 \otimes b_j \otimes c_j + a_j \otimes b_0 \otimes c_j + \sum_{\xi=1}^{\frac{q}{2}} (a_\xi \otimes b_{\xi+\frac{q}{2}} - a_{\xi+\frac{q}{2}} \otimes b_\xi) \otimes c_0 \in (\mathbb{C}^{q+1})^{\otimes 3}. \tag{1.7}$$

In the language of [9], $T_{skewcw,q}$ has the same "block structure" as $T_{cw,q}$, which immediately implies Theorem 1.7.1 also holds for $T_{skewcw,q}$:

Proposition 1.8.2. *For all k,*

$$\omega \leq \log_q\left(\tfrac{4}{27}(\underline{\mathbf{R}}(T_{skewcw,q}^{\boxtimes k}))^{\frac{3}{k}}\right). \tag{1.8}$$

In particular, the known barriers do not apply to $T_{skewcw,2}$ for proving $\omega = 2$ and to any $T_{skewcw,q}$ for $q \leq 10$ for proving $\omega < 2.3$. Unfortunately, we have

Proposition 1.8.3. $\underline{\mathbf{R}}(T_{skewcw,q}) \geq q+3$.

Proposition 1.8.3 is proved in §2.2.

Thus $\underline{\mathbf{R}}(T_{skewcw,q}) > \underline{\mathbf{R}}(T_{cw,q})$ for all q, and in particular $\underline{\mathbf{R}}(T_{skewcw,2}) = 5$.

However, unlike $T_{cw,2}$, substantial strict sub-multiplicativity holds for the Kronecker square of $T_{skewcw,2}$:

Theorem 1.8.4. $\underline{\mathbf{R}}(T_{skewcw,2}^{\boxtimes 2}) = 17$.

In fact, the tensor $T_{skewcw,2}^{\boxtimes 2}$ is much more familiar than it may at first seem. We say that two tensors are *isomorphic* if they are the same up to a change of bases in A, B and C. Let $\det_3 \in (\mathbb{C}^9)^{\otimes 3}$ and $\mathrm{perm}_3 \in (\mathbb{C}^9)^{\otimes 3}$ be the 3×3 determinant and permanent polynomials considered as symmetric tensors.

Proposition 1.8.5. *We have the following isomorphisms of tensors:*

$$T_{cw,2}^{\boxtimes 2} \cong \mathrm{perm}_3$$

$$T_{skewcw,2}^{\boxtimes 2} \cong \det_3.$$

Proposition 1.8.5 is proved in §2.1.1. Hence, Theorem 1.8.4 is a consequence of the following.

Theorem 1.8.6. $\underline{\mathbf{R}}_s(\det_3) = \underline{\mathbf{R}}(\det_3) = 17$.

Proof. $\underline{\mathbf{R}}_s(\det_3) \leq 17$ is proved in §2.3.2. $\underline{\mathbf{R}}(\det_3) \geq 17$ is Theorem 3.1.2 and proved in §3.5. The theorem then follows as $\underline{\mathbf{R}}(\det_3) \leq \underline{\mathbf{R}}_s(\det_3)$. □

We also prove the following in §2.3.1 (note that the first inequality is trivial).

Theorem 1.8.7. $\mathbf{R}(\det_3) \leq \mathbf{R}_s(\det_3) \leq 18$.

Write $\mathfrak{S}_k$ for the symmetric group on k letters. The wedge product $\alpha_1 \wedge \cdots \wedge a_k$ is defined as $\sum_{\sigma \in \mathfrak{S}_k} \mathrm{sgn}(\sigma)\, a_{\sigma(1)} \cdots a_{\sigma(k)} \in A^{\otimes k}$, where $a_i \in A$. The span of all such wedge products is denoted by $\Lambda^k A \subset A^{\otimes k}$. The special linear group $\mathrm{SL}(A) \subset \mathrm{GL}(A)$ is set of linear automorphisms with determinant one or, equivalently, those that preserve $a_1 \wedge \cdots \wedge a_{\mathbf{a}}$ under the induced action on $A^{\otimes \mathbf{a}}$, where $a_1, \ldots, a_{\mathbf{a}}$ is a basis of A. Write $\mathrm{GL}_n = \mathrm{GL}(\mathbb{C}^n)$ and $\mathrm{SL}_n = \mathrm{SL}(\mathbb{C}^n)$. For $X \subset A$, $X^\perp := \{\alpha \in A^* \mid \alpha(x) = 0, x \in X\}$ is its annihilator.

2.1 Symmetry groups of tensors and polynomials

The group $\mathrm{GL}(A) \times \mathrm{GL}(B) \times \mathrm{GL}(C)$ acts naturally on $A \otimes B \otimes C$. The map $\Phi : \mathrm{GL}(A) \times \mathrm{GL}(B) \times \mathrm{GL}(C) \to \mathrm{GL}(A \otimes B \otimes C)$ has a two dimensional kernel $\ker \Phi = \{(\lambda \mathrm{Id}_A, \mu \mathrm{Id}_B, \nu \mathrm{Id}_C) : \lambda\mu\nu = 1\} \simeq (\mathbb{C}^*)^2$.

In particular, the group $(\mathrm{GL}(A) \times \mathrm{GL}(B) \times \mathrm{GL}(C))/(\mathbb{C}^*)^{\times 2}$ is naturally identified with a subgroup of $\mathrm{GL}(A \otimes B \otimes C)$. Given $T \in A \otimes B \otimes C$, the *symmetry group* of a tensor T is the stabilizer of T in $(\mathrm{GL}(A) \times \mathrm{GL}(B) \times \mathrm{GL}(C))/(\mathbb{C}^*)^{\times 2}$, that is

$$G_T := \{g \in (\mathrm{GL}(A) \times \mathrm{GL}(B) \times \mathrm{GL}(C))/(\mathbb{C}^*)^{\times 2} \mid \ g \cdot T = T\}. \tag{2.1}$$

2.1.1 Proof of Proposition 1.8.5

Let

$$\det{}_3 = \sum_{\sigma,\tau \in \mathfrak{S}_3} \mathrm{sgn}(\tau) a_{\sigma(1)\tau(1)} \otimes b_{\sigma(2)\tau(2)} \otimes c_{\sigma(3)\tau(3)},$$

$$\mathrm{perm}_3 = \sum_{\sigma,\tau \in \mathfrak{S}_3} a_{\sigma(1)\tau(1)} \otimes b_{\sigma(2)\tau(2)} \otimes c_{\sigma(3)\tau(3)}$$

be the 3×3 determinant and permanent polynomials regarded as tensors in $\mathbb{C}^9 \otimes \mathbb{C}^9 \otimes \mathbb{C}^9$.

Proof of Lemma 1.8.5. After the change of basis $\tilde{b}_0 := -b_0$ and $\tilde{c}_1 := c_2$, $\tilde{c}_2 := -c_1$, we obtain

$$\begin{aligned}T_{skewcw,2} = a_0 \otimes b_1 \otimes \tilde{c}_2 - a_0 \otimes b_2 \otimes \tilde{c}_1 + a_2 \otimes \tilde{b}_0 \otimes c_1 \\ - a_1 \otimes \tilde{b}_0 \otimes \tilde{c}_2 + a_1 \otimes b_2 \otimes c_0 - a_2 \otimes b_1 \otimes c_0.\end{aligned}$$

This shows that, after identifying the three spaces, $T_{skewcw,2} = a_0 \wedge a_1 \wedge a_2$ is the unique (up to scale) skew-symmetric tensor in $\mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^3$. In particular, $T_{skewcw,2}$ is invariant under the action of SL_3 on $\mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^3$.

Consequently, the stabilizer of $T_{skewcw,2}^{\boxtimes 2}$ in $\mathrm{GL}(\mathbb{C}^9)$ contains (and in fact equals) $\mathrm{SL}_3^{\times 2} \rtimes \mathbb{Z}_2$. This is the stabilizer of the determinant polynomial $\det_3$. Since the determinant is characterized by its stabilizer, we conclude.

The tensor $T_{cw,2}$ is symmetric and, after identifying the three spaces, it coincides with $a_0(a_1^2 + a_2^2) \in S^3\mathbb{C}^3$. After the change of basis $\tilde{a}_1 := a_1 + a_2$, $\tilde{a}_2 := a_1 - a_2$, we obtain $T_{cw,2} = a_0\tilde{a}_1\tilde{a}_2 \in S^3\mathbb{C}^3$ is the square-free monomial of degree 3. The stabilizer of $T_{cw,2}$ under the action of GL_3 on $S^3\mathbb{C}^3$ is $\mathbb{T}_3^{\mathrm{SL}} \rtimes \mathfrak{S}_3$, where $\mathbb{T}_3^{\mathrm{SL}}$ denotes the torus of diagonal matrices with determinant one, and $\mathfrak{S}_3$ acts permuting the three basis elements.

Consequently, the stabilizer of $T_{cw,2}^{\boxtimes 2}$ in $\mathrm{GL}(\mathbb{C}^9)$ contains (and in fact equals) $(\mathbb{T}_3^{\mathrm{SL}} \rtimes \mathfrak{S}_3)^{\times 2} \rtimes \mathbb{Z}_2$. This is the stabilizer of the permanent polynomial perm_3. Since the permanent is characterized by its stabilizer, we conclude. ◻

Remark 2.1.1. For the reader's convenience, here are short proofs that $\det_m, \mathrm{perm}_m$ are characterized by their stabilizers: To see $\det_m$ is characterized by its stabilizer, note that $\mathrm{SL}_m \times \mathrm{SL}_m = \mathrm{SL}(E) \times \mathrm{SL}(F)$ acting on $S^m(E \otimes F)$ decomposes it to

$$\bigoplus_{|\pi|=m} S_\pi E \otimes S_\pi F$$

which is multiplicity free, with the only trivial module $S_{1^m}E \otimes S_{1^m}F = \Lambda^m E \otimes \Lambda^m F$. Here, we have written S_π for the Schur functor corresponding to the partition π. To see that perm_m is characterized by its stabilizer, take the above decomposition and consider the $\mathbb{T}^{\mathrm{SL}(E)} \times \mathbb{T}^{\mathrm{SL}(F)}$-invariants, these are the weight zero spaces $(S_\pi E)_0 \otimes (S_\pi F)_0$. By [22], one has the decomposition of the weight zero spaces as $\mathfrak{S}_m^E \times \mathfrak{S}_m^F$-modules to $(S_\pi E)_0 \otimes (S_\pi F)_0 = [\pi]_E \otimes [\pi]_F$. The only such that is trivial is the case $\pi = (d)$.

Remark 2.1.2. Even Kronecker powers of $T_{skewcw,2}$ are invariant under $\mathrm{SL}_3^{\times 2k}$, and coincide, up to a change of basis, with the *Pascal determinants* (see, e.g., [23, §8.3]), $T_{skewcw,2}^{\boxtimes 2k} = \mathrm{PasDet}_{k,3}$, the unique, up to scale, tensor spanning $(\Lambda^3\mathbb{C}^3)^{\otimes 2k} \subset S^3((\mathbb{C}^3)^{\otimes 2k})$.

Remark 2.1.3. One can regard the 3×3 determinant and permanent as trilinear maps $\mathbb{C}^3 \times \mathbb{C}^3 \times \mathbb{C}^3 \to \mathbb{C}$, where the three copies of $\mathbb{C}^3$ are the first, second and third column of a 3×3 matrix. From this point of view, the trilinear map given by the determinant is $T_{skewcw,2}$ as a tensor and the one given by the permanent is $T_{cw,2}$ as a tensor. This perspective, combined with the notion of product rank, immediately provides the upper bounds $\underline{\mathbf{R}}(\mathrm{perm}_3) \leq 16$ (which is also a consequence of Lemma 1.8.5) and $\underline{\mathbf{R}}(\det_3) \leq 20$, see [24, 25].

Remark 2.1.4. A similar change of basis as the one performed in the second part of proof of Lemma 1.8.5 shows that, up to a change of basis, $T_{skewcw,q} \in \Lambda^3\mathbb{C}^{q+1}$. In particular, its even Kronecker powers are symmetric tensors.

2.2 Koszul flattenings and lower bounds for Kronecker powers

In this section we review Koszul flattenings, prove a result on propagation of Koszul flattening lower bounds under Kronecker products, and prove Theorems 1.7.2 and 1.7.3.

Respectively fix bases $\{a_i\}$, $\{b_j\}$, $\{c_k\}$ of the vector spaces A, B, C. Given $T = \sum_{ijk} T^{ijk} a_i \otimes$

$b_j \otimes c_k \in A \otimes B \otimes C$, define the linear map

$$T_A^{\wedge p} : \Lambda^p A \otimes B^* \to \Lambda^{p+1} A \otimes C$$
$$X \otimes \beta \mapsto \sum_{ijk} T^{ijk} \beta(b_j)(a_i \wedge X) \otimes c_k.$$

Then [26, Proposition 4.1.1] states

$$\underline{\mathbf{R}}(T) \geq \frac{\operatorname{rank}(T_A^{\wedge p})}{\binom{\dim(A)-1}{p}}. \tag{2.2}$$

This type of lower bound has a long history: in general, one takes the space $A \otimes B \otimes C$ and linearly embeds it into a large space of matrices. Then if a rank one tensor maps to a rank q matrix, a rank r tensor maps to a rank at most rq matrix, so the size $rq+1$ minors give equations testing for border rank r. In this case the size of the matrices is $\binom{\mathbf{a}}{p}\mathbf{b} \times \binom{\mathbf{a}}{p+1}\mathbf{c}$ and a rank one tensor maps to a matrix of rank $\binom{\mathbf{a}-1}{p}$.

In practice, one takes a subspace $A'^* \subset A^*$ of dimension $2p+1$ and restricts T (considered as a trilinear form) to $A'^* \times B^* \times C^*$ to get an optimal bound, so the denominator $\binom{\dim(A)-1}{p}$ is replaced by $\binom{2p}{p}$ in (2.2). Write $\phi : A \to A/(A'^*)^\perp =: A'$ for the projection onto the quotient: the corresponding Koszul flattening map gives a lower bound for $\underline{\mathbf{R}}(\phi(T))$, which, by linearity, is a lower bound for $\underline{\mathbf{R}}(T)$. The case $p=1$ is equivalent to Strassen's equations [27]. There are numerous expositions of Koszul flattenings and their generalizations, see, e.g., [23, §7.3], [28, §7.2], [29], [30, §2.4], or [31].

Proof of Proposition 1.8.3. Write $q = 2u$. Fix a space $A' = \langle e_0, e_1, e_2\rangle$. Define $\phi : A \to A'$ by

$$\phi(a_0) = e_0,$$
$$\phi(a_i) = e_1 \quad \text{for } i = 1, \dots, u,$$
$$\phi(a_s) = e_2 \quad \text{for } s = u+1, \dots, q.$$

As an element of $\Lambda^3 A$, we have $T_{skewcw,q} = a_0 \wedge \sum_{i=1}^{u} a_i \wedge a_{u+i}$.

We prove that for $T = T_{skewcw,q}$, one obtains $\mathrm{rank}(T_{A'}^{\wedge 1}) = 2(q+2)+1$. This provides the lower bound $\underline{\mathbf{R}}(T) \geq \left\lceil \frac{2(q+2)+1}{2} \right\rceil = q+3$.

We record the images via $T_{A'}^{\wedge 1}$ of a basis of $A' \otimes B^*$. Fix the range of $i = 1, \dots, u$:

$$T_{A'}^{\wedge 1}(e_0 \otimes \beta_0) = (e_0 \wedge e_1) \otimes \textstyle\sum_{i=1}^{u} c_{u+i} - (e_0 \wedge e_2) \otimes \sum_{i=1}^{u} c_i,$$
$$T_{A'}^{\wedge 1}(e_0 \otimes \beta_i) = (e_0 \wedge e_2) \otimes c_0,$$
$$T_{A'}^{\wedge 1}(e_0 \otimes \beta_{u+i}) = (e_0 \wedge e_1) \otimes c_0,$$
$$T_{A'}^{\wedge 1}(e_1 \otimes \beta_0) = (e_1 \wedge e_2) \otimes \textstyle\sum_{i=1}^{u} c_{u+i},$$
$$T_{A'}^{\wedge 1}(e_1 \otimes \beta_i) = (e_0 \wedge e_1) \otimes c_{u+i} + e_1 \wedge e_2 \otimes c_0,$$
$$T_{A'}^{\wedge 1}(e_1 \otimes \beta_{u+i}) = e_0 \wedge e_1 \otimes c_i,$$
$$T_{A'}^{\wedge 1}(e_2 \otimes \beta_0) = (e_1 \wedge e_2) \otimes \textstyle\sum_{i=1}^{u} c_i,$$
$$T_{A'}^{\wedge 1}(e_2 \otimes \beta_i) = e_0 \wedge e_2 \otimes c_{u+i},$$
$$T_{A'}^{\wedge 1}(e_2 \otimes \beta_{u+i}) = (e_0 \wedge e_2) \otimes c_i - e_1 \wedge e_2 \otimes c_0.$$

Notice that the image of $\sum_{i=1}^{u}(e_1 \otimes \beta_i) - \sum_{i=1}^{u}(e_2 \otimes \beta_{u+i}) - e_0 \otimes \beta_0$ is (up to scale) $e_1 \wedge e_2 \otimes c_0$.

This shows that the image of $T_{A'}^{\wedge 1}$ contains

$$\Lambda^2 A' \otimes c_0 + e_1 \wedge e_2 \otimes \langle \sum_{i=1}^u c_i, \sum_{i=1}^u c_{u+i} \rangle + \langle e_0 \wedge e_1, e_0 \wedge e_2 \rangle \otimes \langle c_1, \dots, c_q \rangle.$$

These summands are in disjoint subspaces, so we conclude

$$\operatorname{rank}(T_{A'}^{\wedge 1}) \geq 3 + 2 + 2q = 2q + 5.$$

□

2.2.1 Propagation of lower bounds under Kronecker products

A tensor $T \in A \otimes B \otimes C$, with $\dim B = \dim C$ is 1_A-*generic* if $T(A^*) \subset B \otimes C$ contains a full rank element. Here is a partial multiplicativity result for Koszul flattening lower bounds under Kronecker products:

Proposition 2.2.1. *Let $T_1 \in A_1 \otimes B_1 \otimes C_1$ with $\dim B_1 = \dim C_1$ be a tensor with a Koszul flattening lower bound for border rank $\underline{\mathbf{R}}(T) \geq r$ given by ${T_1}_{A_1}^{\wedge p}$ (possibly after a restriction ϕ). Let $T_2 \in A_2 \otimes B_2 \otimes C_2$, with $\dim B_2 = \dim C_2 = \mathbf{b}_2$ be 1_{A_2}-generic. Then*

$$\underline{\mathbf{R}}(T_1 \boxtimes T_2) \geq \left\lceil \frac{\operatorname{rank}({T_1}_{A_1}^{\wedge p}) \cdot \mathbf{b}_2}{\binom{2p}{p}} \right\rceil. \tag{2.3}$$

In particular, if $\frac{\operatorname{rank}({T_1}_{A_1}^{\wedge p})}{\binom{2p}{p}} \in \mathbb{Z}$, then $\underline{\mathbf{R}}(T_1 \boxtimes T_2) \geq r\mathbf{b}_2$.

Proof. After applying a restriction ϕ as described above, we may assume $\dim A_1 = 2p + 1$ so that the lower bound for T_1 is

$$\underline{\mathbf{R}}(T_1) \geq \left\lceil \frac{\operatorname{rank}({T_1}_{A_1}^{\wedge p})}{\binom{2p}{p}} \right\rceil.$$

Let $\alpha \in A_2^*$ be such that $T(\alpha) \in B_2 \otimes C_2$ has full rank $\mathbf{b}_2$, which exists by 1_{A_2}-genericity.

Define $\psi : A_1 \otimes A_2 \to A_1$ by $\psi = \mathrm{Id}_{A_1} \otimes \alpha$ and set $\Psi := \psi \otimes \mathrm{Id}_{B_1 \otimes C_1 \otimes B_2 \otimes C_2}$. Then $(\Psi(T_1 \boxtimes T_2)^{\wedge p}_{A_1})$ provides the desired lower bound.

Indeed, the linear map $(\Psi(T_1 \boxtimes T_2)^{\wedge p}_{A_1})$ coincides with $T_1{}^{\wedge p}_{A_1} \boxtimes T_1(\alpha)$. Since matrix rank is multiplicative under Kronecker product, we conclude. □

2.2.2 A short detour on computing ranks of equivariant maps

We briefly explain how to exploit Schur's Lemma (see, e.g., [32, §1.2]) to compute the rank of an equivariant linear map. This is a standard technique, used extensively e.g., in [33, 34] and will reduce the proof of Theorems 1.7.2 and 1.7.3 to the computation of the ranks of specific linear maps in small dimension.

Let G be a reductive group. In the proof of Theorems 1.7.2 and 1.7.3, G will be the product of symmetric groups. Let Λ_G be the set of irreducible representations of G. For $\lambda \in \Lambda_G$, let W_λ denote the corresponding irreducible module.

Suppose U, V are two representations of G. Write $U = \bigoplus_{\lambda \in \Lambda_G} W_\lambda^{\oplus m_\lambda}$, $V = \bigoplus_{\lambda \in \Lambda_G} W_\lambda^{\oplus \ell_\lambda}$, where m_λ is the multiplicity of W_λ in U and ℓ_λ is the multiplicity of W_λ in V. The direct summand corresponding to λ is called the *isotypic component* of type λ.

Let $f : U \to V$ be a G-equivariant map. By Schur's Lemma [32, §1.2], f decomposes as $f = \oplus f_\lambda$, where $f_\lambda : W_\lambda^{\oplus m_\lambda} \to W_\lambda^{\oplus \ell_\lambda}$. Consider multiplicity spaces M_λ, L_λ with $\dim M_\lambda = m_\lambda$ and $\dim L_\lambda = \ell_\lambda$ so that $W_\lambda^{\oplus m_\lambda} \simeq M_\lambda \otimes W_\lambda$ as a G-module, where G acts trivially on M_λ and similarly $W_\lambda^{\oplus \ell_\lambda} \simeq L_\lambda \otimes W_\lambda$.

By Schur's Lemma, the map $f_\lambda : M_\lambda \otimes W_\lambda \to L_\lambda \otimes W_\lambda$ decomposes as $f_\lambda = \phi_\lambda \otimes \mathrm{Id}_{[\lambda]}$, where $\phi_\lambda : M_\lambda \to L_\lambda$. Thus $\mathrm{rank}(f)$ is uniquely determined by $\mathrm{rank}(\phi_\lambda)$ for $\lambda \in \Lambda_G$.

The ranks $\mathrm{rank}(\phi_\lambda)$ can be computed via restrictions of f. For every λ, fix a vector $w_\lambda \in W_\lambda$, so that $M_\lambda \otimes \langle w_\lambda \rangle$ is a subspace of U. Here and in what follows, for a subset $X \subset V$, $\langle X \rangle$

denotes the span of X. Then the rank of the restriction of f to $M_\lambda \otimes \langle w_\lambda \rangle$ coincides with the rank of ϕ_λ.

We conclude

$$\operatorname{rank}(f) = \textstyle\sum_\lambda \operatorname{rank}(\phi_\lambda) \cdot \dim W_\lambda.$$

The proof of Theorem 1.7.2 and proof of Theorem 1.7.3 will follow the algorithm described above, exploiting the symmetries of $T_{cw,q}$. Consider the action of the symmetry group $\mathfrak{S}_q$ on $A \otimes B \otimes C$ defined by permuting the basis elements with indices $\{1, \dots, q\}$. More precisely, a permutation $\sigma \in \mathfrak{S}_q$ induces the linear map defined by $\sigma(a_i) = a_{\sigma(i)}$ for $i = 1, \dots, q$ and $\sigma(a_0) = a_0$. The group $\mathfrak{S}_q$ acts on B, C similarly, and the simultaneous action on the three factors defines an $\mathfrak{S}_q$-action on $A \otimes B \otimes C$. The tensor $T_{cw,q}$ is invariant under this action.

2.2.3 Proof of Theorem 1.7.2

When $q = 3$, the result is true by a direct calculation using the $p = 2$ Koszul flattening with a sufficiently generic $\mathbb{C}^5 \subset A^*$, which is left to the reader. In what follows we treat the case $q > 3$.

Write $a_{ij} = a_i \otimes a_j \in A^{\otimes 2}$ and similarly for $B^{\otimes 2}$ and $C^{\otimes 2}$. Let $A' = \langle e_0, e_1, e_2 \rangle$ and define the linear map $\phi_2 : A^{\otimes 2} \to A'$ via

$$\begin{aligned}
&\phi_2(a_{00}) = \phi_2(a_{01}) = \phi_2(a_{10}) = e_0 + e_1,\\
&\phi_2(a_{11}) = e_0,\\
&\phi_2(a_{03}) = \phi_2(a_{30}) = e_1 + e_2\\
&\phi_2(a_{22}) = \phi_2(a_{31}) = e_2\\
&\phi_2(a_{02}) = \phi_2(a_{20}) = e_1\\
&\phi_2(a_{0i}) = \phi_2(a_{i0}) = e_1 \quad \text{for } i = 4, \dots, q\\
&\phi_2(a_{ij}) = 0 \quad \text{for all other pairs } (i, j).
\end{aligned}$$

For $q \geq 4$, we apply the $p = 1$ Koszul flattening map to the restriction of $T_{cw,q}^{\boxtimes 2}$ determined by ϕ_2.

The tensor $T_{cw,q}$ is invariant under the action of $\mathfrak{S}_q$ acting on the indices $\{1, \dots, q\}$ of the basis elements of $\mathbb{C}^{q+1}$. Therefore $T_{cw,q}^{\boxtimes 2}$ is invariant under the action of $\mathfrak{S}_q \times \mathfrak{S}_q$ on $A^{\otimes 2} \otimes B^{\otimes 2} \otimes C^{\otimes 2}$. Let $\Gamma := \mathfrak{S}_{q-3} \times \mathfrak{S}_{q-3}$ where $\mathfrak{S}_{q-3}$ is the permutation group on $\{4, \dots, q\}$; $T_{cw,q}^{\boxtimes 2}$ is invariant under the action of Γ.

The projection ϕ_2 is invariant under the action of Γ, so $(\phi_2(T_{cw,q}^{\boxtimes 2})_{A'}^{\wedge 1}$ is Γ-equivariant, because in general Koszul flattenings are equivariant under the product of the three general linear groups, which is $\mathrm{GL}(A') \times \mathrm{GL}(B^{\otimes 2}) \times \mathrm{GL}(C^{\otimes 2})$ in our case.

We now apply the method described in §2.2.2 to compute $\mathrm{rank}((T_q)_{A'}^{\wedge 1})$.

Let $[\mathrm{triv}]$ denote the trivial $\mathfrak{S}_{q-3}$-representation and let V denote the standard representation, that is the Specht module associated to the partition $(q-4, 1)$ of $q-3$. We have $\dim[\mathrm{triv}] = 1$ and $\dim V = q - 4$. When $q = 4$ only the trivial representation appears.

The spaces B, C are isomorphic as $\mathfrak{S}_{q-3}$-modules and they decompose as $B = C = [\mathrm{triv}]^{\oplus 5} \oplus V$. After fixing a 5-dimensional multiplicity space $\mathbb{C}^5$ for the trivial isotypic component, we write $B^* = C = \mathbb{C}^5 \otimes [\mathrm{triv}] \oplus V$. To distinguish the two $\mathfrak{S}_{q-3}$-actions, we write $B^{*\otimes 2} = ([\mathrm{triv}]_L^{\oplus 5} \oplus V_L) \otimes ([\mathrm{triv}]_R^{\oplus 5} \oplus V_R)$ and similarly for $C^{\otimes 2}$

Thus,

$$
\begin{aligned}
B^{*\otimes 2} = C^{\otimes 2} = & \mathbb{C}^{5^{\otimes 2}} \otimes ([\mathrm{triv}]_L \otimes [\mathrm{triv}]_R) \oplus \\
& \mathbb{C}^5 \otimes ([\mathrm{triv}]_L \otimes V_R) \quad \oplus \\
& \mathbb{C}^5 \otimes (V_L \otimes [\mathrm{triv}]_R) \quad \oplus \\
& (V_L \otimes V_R).
\end{aligned}
$$

Write $W_1, \ldots, W_4$ for the four irreducible representations in the decomposition above and let $M_1, \ldots, M_4$ be the four corresponding multiplicity spaces.

Recall from [35] that a basis of V is given by standard Young tableaux of shape $(q-4, 1)$ (with entries in $4, \ldots, q$ for consistency with the action of $\mathfrak{S}_{q-3}$); let w_{std} be the vector corresponding to the standard tableau having $4, 6, \ldots, q$ in the first row and 5 in the second row. We refer to [35, §7] for the straightening laws of the tableaux. Let w_{triv} be a generator of the trivial representation $[\mathrm{triv}]$. Writing $\mathbb{C}^{q+1} = \langle e_0, \ldots, e_q \rangle$, we explicitly have $w_{std} = e_5 - e_4$ and the multiplicity space 5-dimensional multiplicity space of the trivial representation is $\langle e_0, \ldots, e_3, \sum_4^q e_j \rangle$.

For each of the four isotypic components in the decomposition above, we fix a vector $w_i \in W_i$ and explicitly realize the subspaces $M_i \otimes \langle w_i \rangle$ of $B^{*\otimes 2}$ as follows:

W_i	w_i	$\dim M_i$	$M_i \otimes \langle w_i \rangle$
$[\mathrm{triv}]_L \otimes [\mathrm{triv}]_R$	$w_{\mathrm{triv}} \otimes w_{\mathrm{triv}}$	25	$\langle \beta_{ij} : i,j=0,\ldots,3 \rangle \oplus$ $\langle \sum_{j=4}^q \beta_{ij} : i=0,\ldots,3 \rangle \oplus$ $\langle \sum_{i=4}^q \beta_{ij} : j=0,\ldots,3 \rangle \oplus$ $\langle \sum_{i,j=4}^q \beta_{ij} \rangle$
$[\mathrm{triv}]_L \otimes V_R$	$w_{\mathrm{triv}} \otimes w_{std}$	5	$\langle \beta_{i5} - \beta_{i4} : i=0,\ldots,3 \rangle \oplus$ $\langle \sum_{i=4}^q (\beta_{i5} - \beta_{i4}) \rangle$
$V_L \otimes [\mathrm{triv}]_R$	$w_{std} \otimes w_{\mathrm{triv}}$	5	$\langle \beta_{5j} - \beta_{4j} : j=0,\ldots,3 \rangle \oplus$ $\langle \sum_{j=4}^q (\beta_{5j} - \beta_{4j}) \rangle$
$V_L \otimes V_R$	$w_{std} \otimes w_{std}$	1	$\langle \beta_{55} - \beta_{45} - \beta_{54} + \beta_{44} \rangle$.

The subspaces in $C^{\otimes 2}$ are realized similarly.

Since $(T_{cw,q}^{\boxtimes 2})^{\wedge 1}_{A'}$ is Γ-equivariant, by Schur's Lemma, it has the isotypic decomposition $(T_{cw,q}^{\boxtimes 2})^{\wedge 1}_{A'} =$

$f_1 \oplus f_2 \oplus f_3 \oplus f_4$, where

$$f_i : A' \otimes (M_i \otimes W_i) \to \Lambda^2 A' \otimes (M_i \otimes W_i). \tag{2.4}$$

As explained in §2.2.2, it suffices to compute the ranks of the four restrictions $\Phi_i : A' \otimes M_i \otimes \langle w_i \rangle \to \Lambda^2 A' \otimes M_i \otimes \langle w_i \rangle$ to the multiplicities spaces.

The four matrices representing $\Phi_1, \ldots, \Phi_4$ are computed by a routine which exploits their structure. The script to compute the matrices and their ranks is available at `https://www.math.tamu.edu/~jml/CGLVkronsupp.html`, Appendix D. The method to compute the matrices is explained in Section 2.4.

The script provides an expression for the entries of the matrices Φ_i which are univariate polynomials in q up to a global univariate polynomial factor. The expressions are valid for $q \geq 5$. The rank of the Koszul flattening in the cases $q = 3$ and $q = 4$ is computed directly.

We determine a lower bound on $\mathrm{rank}(\Phi_i)$ by computing a matrix $P_i \cdot \Phi_i \cdot Q_i$, where P_i is a rectangular matrix whose entries are rational functions of q (well defined for $q \geq 5$) and Q_i is a rectangular matrix whose entries are constant. The resulting matrix $P_i \cdot \Phi_i \cdot Q_i$ is a square matrix, upper triangular with ± 1 on the diagonal, so that the size of $P_i \Phi_i Q_i$ gives a lower bound on $\mathrm{rank}(\Phi_i)$.

We summarize the results of the script in the following table.

W_i	$\dim W_i$	$\dim M_i$	$\mathrm{rank}(\Phi_i)$	contribution to total rank
$[\mathrm{triv}]_L \otimes [\mathrm{triv}]_R$	1	25	72	72
$[\mathrm{triv}]_L \otimes V_R$	$q-4$	5	12	$12(q-4)$
$V_L \otimes [\mathrm{triv}]_R$	$q-4$	5	12	$12(q-4)$
$V_L \otimes V_R$	1	$(q-4)^2$	2	$2(q-4)^2$

Adding the total contributions, we obtain

$$\mathrm{rank}(T_{A'}^{\wedge 1}) = 2\cdot(q-4)^2 + 12\cdot(q-4) + 12\cdot(q-4) + 72\cdot 1 = 2(q+2)^2.$$

This conclude the proof of Theorem 1.7.2.

2.2.4 Proof of Theorem 1.7.3

We will give a lower bound on $\underline{\mathbf{R}}(T_{cw,q}^{\boxtimes 3})$ by computing its Koszul flattening for $p=2$. Write $a_{ijk} = a_i \otimes a_j \otimes a_k \in A^{\otimes 3}$ and similarly for $B^{\otimes 3}$ and $C^{\otimes 3}$. Let $\{\alpha_{ijk}\} \subset A^{*\otimes 3}$ be the dual basis to $\{a_{ijk}\} \subset A^{\otimes 3}$. Let $A' = \langle e_0, \dots, e_4\rangle$ be a 5-dimensional space and let $\{e^0, \dots, e^4\}$ be the dual basis of $\{e_0, \dots, e_4\}$ and define $\phi_3 : A^{\otimes 3} \to A'$ to be the linear map whose transpose $\phi_3^T : A'^* \to A^{*\otimes 3}$ is given by

$$\phi_3^T(e^0) = \alpha_{000}$$
$$\phi_3^T(e^1) = \textstyle\sum_{i=1}^q (\alpha_{i00} + \alpha_{0i0} + \alpha_{00i})$$
$$\phi_3^T(e^2) = \alpha_{001} + \alpha_{010} + \alpha_{012} + \alpha_{102} + \alpha_{110} + \alpha_{121} + \alpha_{200} + \alpha_{211}$$
$$\phi_3^T(e^3) = \alpha_{022} + \alpha_{030} + \alpha_{031} + \alpha_{100} + \alpha_{103} - \alpha_{120} + \alpha_{210} + \alpha_{212} + \alpha_{300}$$
$$\phi_3^T(e^4) = \alpha_{002} + \alpha_{004} + \alpha_{011} + \alpha_{014} + \alpha_{020} + \alpha_{023} + \alpha_{032} + \alpha_{040} + \alpha_{100} + \alpha_{122} + \alpha_{220} + \alpha_{303}.$$

Let $T_q = \phi_3(T_{cw,q}^{\boxtimes 3}) \in A' \otimes B^{\otimes 3} \otimes C^{\otimes 3}$ and consider the Koszul flattening

$$(T_q)_{A'}^{\wedge 2} : \Lambda^2 A' \otimes B^{*\otimes 3} \to \Lambda^3 A' \otimes C^{\otimes 3}.$$

We will show $\mathrm{rank}((T_q)_{A'}^{\wedge 2}) = 6(q+2)^3$, which implies $\underline{\mathbf{R}}(T_{cw,q}^{\boxtimes 3}) \geq (q+2)^3$.

We employ the same method as in Section 2.2.3 in the case of $T_{cw,q}^{\boxtimes 2}$. The Koszul flattening is equivariant for the action of $\Gamma = \mathfrak{S}_{q-4}^{\times 3}$ where $\mathfrak{S}_{q-4}$ acts on $\{5, \dots, q\}$. In particular $\mathbb{C}^{q+1}$

splits under the action of $\mathfrak{S}_{q-4}$ into a 6-dimensional subspace of invariants $\mathbb{C}^6 \otimes [\mathrm{triv}] = \langle e_0, \dots, e_4, e_5 + \dots + e_q \rangle$ and a copy of the standard representation $V = \langle e_i - e_5 : i = 6, \dots, q \rangle$, with $\dim V = q - 5$.

Hence, the spaces $B^{\otimes 3}$ and $C^{\otimes 3}$ split into the direct sum of 8 isotypic components for the action of Γ as follows (we use indices $1, 2, 3$ to denote the trivial or the standard representation on the first, second or third factor):

$$\begin{aligned}
B^{*\otimes 3} \simeq C^{\otimes 3} = &(\mathbb{C}^6)^{\otimes 3} \otimes ([\mathrm{triv}]_1 \otimes [\mathrm{triv}]_2 \otimes [\mathrm{triv}]_3) \\
&\oplus (\mathbb{C}^6)^{\otimes 2} \otimes \Big[([\mathrm{triv}]_1 \otimes [\mathrm{triv}]_2 \otimes V_3) \\
&\qquad\qquad \oplus ([\mathrm{triv}]_1 \otimes V_2 \otimes [\mathrm{triv}]_3) \\
&\qquad\qquad \oplus (V_1 \otimes [\mathrm{triv}]_2 \otimes [\mathrm{triv}]_3)\Big] \\
&\oplus (\mathbb{C}^6) \otimes \Big[([\mathrm{triv}]_1 \otimes V_2 \otimes V_3) \\
&\qquad\qquad \oplus (V_1 \otimes V_2 \otimes [\mathrm{triv}]_3) \\
&\qquad\qquad \oplus (V_1 \otimes [\mathrm{triv}]_2 \otimes V_3)\Big] \\
&\oplus V_1 \otimes V_2 \otimes V_3
\end{aligned}$$

Similarly to the square case, for each of the eight isotypic components, we consider $w_i \in W_i$ where W_i is the corresponding irreducible and we compute the rank of the restriction $\Psi_i : \Lambda^2 A' \otimes M_i \otimes \langle w_i \rangle \to \Lambda^3 A' \otimes M_i \otimes \langle w_i \rangle$ of the Koszul flattening.

The matrices representing the maps Ψ_i are computed exploiting the structure of the tensors involved, following the method described in Section 2.4. The expression computed by the script is valid for $q \geq 6$. The case $q = 5$ is computed explicitly. Their ranks are computed by reducing Ψ_i to a triangular matrix as in the previous case.

The ranks of the restrictions are recorded in the following table:

W_i	$\dim W_i$	$\dim M_i$	$\mathrm{rank}(\Psi_i)$	total contribution
$[\mathrm{triv}]_1 \otimes [\mathrm{triv}]_2 \otimes [\mathrm{triv}]_3$	1	$6^3 = 216$	2058	2058
$[\mathrm{triv}]_1 \otimes [\mathrm{triv}]_2 \otimes V_3$ (and permutations)	$(q-5)$ (three times)	$6^2 = 36$ (three times)	294 (three times)	$3 \cdot 294(q-5)$
$[\mathrm{triv}]_1 \otimes V_2 \otimes V_3$ (and permutations)	$(q-5)^2$ (three times)	6 (three times)	42 (three times)	$3 \cdot 42(q-5)^2$
$V_1 \otimes V_2 \otimes V_3$	$(q-5)^3$	1	6	$6(q-5)^3$

Adding all the contributions together, we obtain

$$\mathrm{rank}(T_{A'}^{\wedge 2}) = 6(q-5)^3 + 3 \cdot 42(q-5)^2 + 3 \cdot 294(q-5) + 2058 \cdot 1 = 6 \cdot (q+2)^3.$$

This concludes the proof of Theorem 1.7.3.

2.3 Upper bounds for Waring rank and border rank of $\det_3$

2.3.1 Proof of upper bound in Theorem 1.8.7

We give the rank 18 decomposition for $\det_3$ explicitly, as a collection of 18 linear forms on $\mathbb{C}^9 = \mathbb{C}^3 \otimes \mathbb{C}^3$ whose cubes add up to $\det_3$. The linear forms are given in coordinates recorded in the matrices below: the 3×3 matrix (ζ_{ij}) represents the linear forms $\sum_{ij} \zeta_{ij} x_{ij}$. This presentation highlights some of the symmetries of the decomposition.

Let $\vartheta = \exp(2\pi i/6)$ and let $\overline{\vartheta}$ be its inverse. The tensor $\det_3 = T_{skewcw,2}^{\boxtimes 2} = \det(x_{ij}) \in S^3(\mathbb{C}^3 \otimes \mathbb{C}^3)$ satisfies

$$\det_3 = \sum_1^{18} L_i^3$$

where $L_1, \ldots, L_{18}$ are the 18 linear forms given by the following coordinates:

$$
\begin{array}{lll}
L_1 = \begin{pmatrix} -\vartheta & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & \overline{\vartheta} \end{pmatrix} & L_2 = \begin{pmatrix} -\overline{\vartheta} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & \vartheta \end{pmatrix} & L_3 = \begin{pmatrix} -\overline{\vartheta} & 0 & 0 \\ 0 & \frac{1}{3}\overline{\vartheta} & 0 \\ 0 & 0 & \overline{\vartheta} \end{pmatrix} \\
L_4 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -\overline{\vartheta} \\ 0 & -\frac{1}{3}\vartheta & 0 \end{pmatrix} & L_5 = \begin{pmatrix} \overline{\vartheta} & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -\frac{1}{3}\vartheta & 0 \end{pmatrix} & L_6 = \begin{pmatrix} \vartheta & 0 & 0 \\ 0 & 0 & -\vartheta \\ 0 & -\frac{1}{3}\vartheta & 0 \end{pmatrix} \\
L_7 = \begin{pmatrix} 0 & \frac{1}{3}\overline{\vartheta} & 0 \\ -\vartheta & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} & L_8 = \begin{pmatrix} 0 & \frac{1}{3}\overline{\vartheta} & 0 \\ -\overline{\vartheta} & 0 & 0 \\ 0 & 0 & -\overline{\vartheta} \end{pmatrix} & L_9 = \begin{pmatrix} 0 & \frac{1}{3}\vartheta & 0 \\ -\overline{\vartheta} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
L_{10} = \begin{pmatrix} 0 & -\frac{1}{3}\vartheta & 0 \\ 0 & 0 & \overline{\vartheta} \\ -1 & 0 & 0 \end{pmatrix} & L_{11} = \begin{pmatrix} 0 & -\frac{1}{3}\overline{\vartheta} & 0 \\ 0 & 0 & \vartheta \\ -1 & 0 & 0 \end{pmatrix} & L_{12} = \begin{pmatrix} 0 & \frac{1}{3} & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{pmatrix} \\
L_{13} = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \end{pmatrix} & L_{14} = \begin{pmatrix} 0 & 0 & 1 \\ \overline{\vartheta} & 0 & 0 \\ 0 & \frac{1}{3}\vartheta & 0 \end{pmatrix} & L_{15} = \begin{pmatrix} 0 & 0 & 1 \\ \vartheta & 0 & 0 \\ 0 & \frac{1}{3}\overline{\vartheta} & 0 \end{pmatrix} \\
L_{16} = \begin{pmatrix} 0 & 0 & \overline{\vartheta} \\ 0 & -\frac{1}{3}\vartheta & 0 \\ 1 & 0 & 0 \end{pmatrix} & L_{17} = \begin{pmatrix} 0 & 0 & \overline{\vartheta} \\ 0 & -\frac{1}{3}\overline{\vartheta} & 0 \\ -\overline{\vartheta} & 0 & 0 \end{pmatrix} & L_{18} = \begin{pmatrix} 0 & 0 & \vartheta \\ 0 & -\frac{1}{3}\overline{\vartheta} & 0 \\ 1 & 0 & 0 \end{pmatrix}
\end{array}.
$$

The equality can be verified by hand. A Macaulay2 file performing the calculation is available at `https://www.math.tamu.edu/~jml/CGLVkronsupp.html`, Appendix B.

□

2.3.2 Proof of Theorem 1.8.6

As in the case of Theorem 1.8.7, we prove Theorem 1.8.6 by explicitly giving 17 linear forms, depending on a parameter t, whose cubes provide a border rank 17 expression for $\det_3$. The algebraic numbers involved are more complicated than in the previous case.

The result was achieved by numerical methods, which allowed us to sparsify the decomposition and ultimately determine the value of the coefficients. The linear forms in the decomposition are described below.

Consider

$$L_1(t) = \begin{pmatrix} z_1 & 0 & 0 \\ 0 & z_2 t & 0 \\ -1 & 0 & 0 \end{pmatrix} \qquad L_2(t) = \begin{pmatrix} z_3 & 0 & 0 \\ z_4 & 0 & z_5 t \\ z_6 & 0 & 0 \end{pmatrix} \qquad L_3(t) = \begin{pmatrix} -z_{36} & z_7 t & 0 \\ -z_{38} & 0 & -z_{39} t \\ 0 & 0 & t \end{pmatrix}$$

$$L_4(t) = \begin{pmatrix} 0 & 0 & t \\ -z_{34} & 0 & 0 \\ 0 & z_8 t & -z_{35} t \end{pmatrix} \qquad L_5(t) = \begin{pmatrix} 0 & -z_{19} t & -z_{20} t \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \qquad L_6(t) = \begin{pmatrix} -z_{22} & z_9 t & 0 \\ -z_{23} & 0 & -z_{24} t \\ -z_{25} & 0 & 0 \end{pmatrix}$$

$$L_7(t) = \begin{pmatrix} z_{10} & z_{11} t & 0 \\ z_{12} & 0 & z_{13} t \\ z_{14} & 0 & 0 \end{pmatrix} \qquad L_8(t) = \begin{pmatrix} z_{15} & -t & 0 \\ z_{16} & 0 & z_{17} t \\ z_{18} & 0 & 0 \end{pmatrix} \qquad L_9(t) = \begin{pmatrix} 0 & z_{19} t & z_{20} t \\ 0 & z_{21} t & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$L_{10}(t) = \begin{pmatrix} -z_{41} & 0 & 0 \\ 0 & 0 & 0 \\ -z_{44} & 0 & 0 \end{pmatrix} \qquad L_{11}(t) = \begin{pmatrix} z_{22} & 0 & 0 \\ z_{23} & 0 & z_{24} t \\ z_{25} & 0 & 0 \end{pmatrix} \qquad L_{12}(t) = \begin{pmatrix} -z_{31} & z_{26} t & 0 \\ 0 & z_{27} t & 0 \\ 0 & 0 & t \end{pmatrix}$$

$$L_{13}(t) = \begin{pmatrix} z_{28} & z_{29} t & 0 \\ z_{30} & 0 & -t \\ 0 & t & 0 \end{pmatrix} \qquad L_{14}(t) = \begin{pmatrix} z_{31} & z_{32} t & 0 \\ 0 & 0 & 0 \\ 0 & z_{33} t & -t \end{pmatrix} \qquad L_{15}(t) = \begin{pmatrix} 0 & 0 & -t \\ z_{34} & 0 & 0 \\ 0 & 0 & z_{35} t \end{pmatrix}$$

$$L_{16}(t) = \begin{pmatrix} z_{36} & z_{37} t & 0 \\ z_{38} & 0 & z_{39} t \\ 0 & z_{40} t & -t \end{pmatrix} \qquad L_{17}(t) = \begin{pmatrix} z_{41} & z_{42} t & 0 \\ 0 & z_{43} t & 0 \\ z_{44} & 0 & 0 \end{pmatrix}$$

The coefficients $z_1, \ldots, z_{44}$ are algebraic numbers described as follows. Let y_* be a real root of the polynomial

$$x^{27} - 2x^{26} + 17x^{25} - 29x^{24} + 81x^{23} + 52x^{22} - 726x^{21} + 3451x^{20} - 10901x^{19} + 25738x^{18} -$$
$$50663x^{17} + 72133x^{16} - 72973x^{15} + 10444x^{14} + 138860x^{13} - 308611x^{12} + 427344x^{11}$$
$$- 267416x^{10} - 196096x^{9} + 762736x^{8} - 1236736x^{7} + 1092352x^{6} - 537600x^{5} - 42240x^{4} +$$
$$684032x^{3} - 1136640x^{2} + 1146880x - 520192.$$

For $i = 1, \dots, 44$, we consider algebraic numbers y_j in the field extension $\mathbb{Q}[y_*]$, described as a polynomial of degree (at most) 26 in y_* with rational coefficients. Notice that all the y_j's are real. The expressions of the $y_1, \dots, y_{44}$ in terms of y_* are provided in the file `yy_exps.m2` at `https://www.math.tamu.edu/~jml/CGLVkronsupp.html`, Appendix C. Let z_j be the unique real cubic root of y_j.

We are going to prove that, with this choice of coefficients z_j,

$$t^2 \det_3 + O(t^3) = \sum_{i=1}^{17} L_i(t)^3. \tag{2.5}$$

The condition $t^2 \det_3 + O(t^3) = \sum_{i=1}^{17} L_i(t)^3$ is equivalent to the fact that the degree 0 and the degree 1 components of $\sum_{i=1}^{17} L_i(t)^3$ vanish and that the degree 2 component equals $\det_3$. Given the sparse structure of the $L_i(t)$, this reduces to a system of 54 cubic equations in the 44 unknowns $z_1, \dots, z_{44}$. Our goal is to show that the algebraic numbers described above are a solution of this system.

We show that the z_i's satisfy each equation as follows. After evaluating the equations at the z_i's, there are two possible cases

1. all monomials appearing in the equation are elements of $\mathbb{Q}[y_*]$; we say that this is an

equation of type 1; there are 14 such equations;

2. at least one monomial appearing in the equation is not an element of $\mathbb{Q}[y_*]$; we say that this is an equation of type 2; there are 40 such equations.

For equations of type 1, we provide expressions of each monomial in terms of y_*. To verify that each expression is indeed equal to the corresponding monomial, it suffices to compare the cube of the given expression and the expression obtained by evaluating the monomial at the y_j's. Finally, the equation can be verified in $\mathbb{Q}[y_*]$. This is performed by the file `checkingType1eqns.m2`.

For equations of type 2, let u be one of the monomials which do not belong to $\mathbb{Q}[y_*]$. We claim that it is possible to choose the monomial in such a way that $\mathbb{Q}[u^3] = \mathbb{Q}[y_*]$. For each equation, we choose one of the monomials and we verify the claim as follows. The element u^3 has an expression in terms of y_* which equals the chosen monomial evaluated at the y_i's. Let M_u be the 27×27 matrix with rational entries such that

$$(1, u^3, \cdots, u^{3\cdot 26}) = \left(1, y_*, \ldots, y_*^{26}\right) \cdot M_u;$$

M_u can be computed directly by considering the expressions of the powers of u^3 in terms of y_*. Then $\mathbb{Q}[u^3] = \mathbb{Q}[y_*]$ if and only if M_u is full rank.

In particular y_* has an expression in terms of u^3, which can be computed inverting the matrix M_u. A consequence of this is that $\mathbb{Q}[u] = \mathbb{Q}[y_*, u]$.

At this point, we observe that $\mathbb{Q}[u]$ contains the other monomials occurring in the equation as well. To see this, we proceed as in the case of equations of type 1. For each monomial occurring in the equation, we provide an expression in terms of u (in fact, to speed up the calculation, we provide an expression in terms of u and y_*, which is equivalent to an expression in u because $\mathbb{Q}[u^3] = \mathbb{Q}[y_*]$ and y_* has a unique expression in terms of u^3); we compare the cube of this expression (appropriately reduced modulo the minimal polynomial of y_* and the relation between u^3 and y_*) with the expression obtained by evaluating the monomial at the y_i's (expressed in terms of y_*). This shows that all monomials occurring in the expression belong to $\mathbb{Q}[u]$, and verifies that the given expressions are indeed equal to the corresponding monomials. Finally, the equation is verified in $\mathbb{Q}[u]$ as in the case of type 1. This is performed by the file `checkingType2eqns.m2`. $\square$

2.3.3 Discussion of how the decomposition was obtained

Many steps were accomplished by finding solutions of polynomial equations by nonlinear optimization. In each case, this was accomplished using a variant of Newton's method applied to the mapping of variable values to corresponding polynomial values. The result of this procedure in each case is limited precision machine floating point numbers.

First, we attempted to solve the equations describing a Waring rank 17 decomposition of $\det_3$ with nonlinear optimization, namely, $\det_3 = \sum_{i=1}^{17}(w_i')^{\otimes 3}$, where $w_i' \in \mathbb{C}^{3\times 3}$. Instead of finding a solution to working precision, we obtained a sequence of local refinements to an approximate solution where the norm of the defect is slowly converging to to zero, and some of the parameter values are exploding to infinity. Numerically, these are Waring decompositions of polynomials very close to $\det_3$.

Next, this approximate solution needed to be upgraded to a solution to equation (2.5).

We found a choice of parameters in the neighborhood of a solution, and then applied local optimization to solve to working precision. We used the following method: Consider the linear mapping $M : \mathbb{C}^{17} \to S^3(\mathbb{C}^{3\times 3})$, $M(e_i) = (w_i')^{\otimes 3}$, and let $M = U\Sigma V^*$ be its singular value decomposition (with respect to the standard inner products for the natural coordinate systems). We observed that the singular values seemed to be naturally partitioned by order of magnitude. We estimated this magnitude factor as $t_0 \approx 10^{-3}$, and wrote Σ' as Σ where we multiplied each singular value by $(t/t_0)^k$, with k chosen to agree with this observed partitioning, so that the constants remaining were reasonably sized. Finally, we let $M' = U\Sigma' V^*$, which has entries in $\mathbb{C}[[t]]$. Thus M' is a representation of the map M with a parameter t.

Next, for each i, we optimized to find a best fit to the equation $(a_i + tb_i + t^2 c_i)^{\otimes 3} = M'(e_i)$, which is defined by polynomial equations in the entries of a_i, b_i and c_i. The a_i, b_i and c_i we constructed in this way proved to be a good initial guess to optimize equation (2.5), and we immediately saw quadratic convergence to a solution to machine precision. At this point, we greedily sparsified the solution by speculatively zero-ing values and re-optimizing, rolling back one step in case of failure. After sparsification, it turned out the c_i were not needed. The resulting matrices are those given in the proof.

To compute the minimal polynomials and other integer relationships between quantities, we used Lenstra-Lenstra-Lovász integer lattice basis reduction [36]. As an example, let $\zeta \in \mathbb{R}$ be approximately an algebraic number of degree k. Let N be a large number inversely proportional to the error of ζ. Consider the integer lattice with basis $\{e_i + \lfloor N\zeta^i \rfloor e_{k+1}\} \subset \mathbb{Z}^{k+2}$, for $0 \leq i \leq k$. Then elements of this lattice are of the form $v_0 e_0 + \cdots + v_k e_k + E e_{k+1}$, where $E \approx Np(\zeta)$, $p = v_0 + v_1 x + \cdots x_k x^k$. Polynomials p for which ζ is an approximate root are distinguished by the property of having relatively small Euclidean norm in this lattice. Computing a small norm vector in an integer lattice is accomplished by LLL reduction of a known basis.

For example, the fact that the number field of degree 27 obtained by adjoining any z_α^3 to $\mathbb{Q}$ contains all the rest was determined via LLL reduction, looking for expressions of z_α^3 as a polynomial in z_β^3 for some fixed β. These expressions of z_α^3 in a common number field can be checked to have the correct minimal polynomial, and thus agree with our initial description of the z_α. LLL reduction was also used to find the expressions of values as polynomials in the primitive root of the various number fields.

After refining the known value of the parameters to $10,000$ bits of precision using Newton's method, LLL reduction was successful in identifying the minimal polynomials. The degrees were simply guessed, and the results checked by evaluating the computed polynomials in the parameters to higher precision.

Remark 2.3.1. With the minimal polynomial information, it is possible to check that equation (2.5) is satisfied to any desired precision by the parameters.

2.4 A method to compute flattenings of structured tensors

In this section, we explain how to compute the matrices $\Phi_1, \dots, \Phi_4$ in Section 2.2.3 and the matrices $\Psi_1, \dots, \Psi_8$ in Section 2.2.4.

The matrices $\Phi_1, \dots, \Phi_4$ and $\Psi_1, \dots, \Psi_8$ arise via a series of tensor contractions of highly structured tensors. In this section, we introduce the notion of *box parametrized* sequence of tensors. Lemma 2.4.2 below shows that contraction of box parametrized tensors gives rise to box parametrized tensors; in addition, the expression of the tensors resulting from the contraction is particularly easy to control.

We will then show that the tensors in Section 2.2.3 and Section 2.2.4 which give rise to the matrices $\Phi_1, \dots, \Phi_4$ and $\Psi_1, \dots, \Psi_8$ are box parametrized. This allows us to track down the entries of the final matrices as functions of the dimension q.

The full calculation of the matrices is left to the scripts available in Appendix D at `https://www.math.tamu.edu/~jml/CGLVkronsupp.html`.

The point of view is partially inspired to the interpretation of tensors in communication models, where a tensor on k factors is regarded as a function from $\underbrace{\mathbb{N} \times \cdots \times \mathbb{N}}_{k} \to \mathbb{C}$ with finite support sending a k-tuple of integers to the corresponding coefficient of the tensor. Explicitly, for every $j = 1, \dots, k$ fix a basis $\{v_i^{(j)}\}$ on the j-th factor: given a finite support $\Sigma \subset \mathbb{N}^{\times k}$, the tensor $T = \sum_{(i_1,\dots,i_k)\in\Sigma} t_{i_1,\dots,i_k} v_{i_1}^{(1)} \otimes \cdots \otimes v_{i_k}^{(k)}$ corresponds to the function defined by $T(i_1,\dots,i_k) = t_{i_1,\dots,i_k}$. We do not explicitly write the dimensions of the factors.

Let $\mathcal{T} = \{T_q : q \in \mathbb{N}\}$ be a sequence of tensors of order k. We say that $\mathcal{T}$ is *basic box-parametrized* if, for every q

$$T_q = p(q) \sum_{(i_1,\dots,i_k)\in\Sigma_q} v_{i_1}^{(1)} \otimes \cdots \otimes v_{i_k}^{(k)}$$

where $p(q)$ is a univariate polynomial in q and the support Σ_q is defined by conditions $\eta_j q + \vartheta_j \le i_j \le H_j q + \Theta_j$ for $\eta_j, H_j \in \{0,1\}$ and $\vartheta_j, \Theta_j \in \mathbb{Z}_{\ge 0}$, and any number (not depending on q) of equalities $i_j = i_{j'}$ among indices. Without loss of generality, assume that the inequalities are sharp for every j, in the sense that for every i_j satisfying the j-th inequality, the basis element $v_{i_j}^{(j)}$ does appear in T_q. We often say that $\mathcal{T}$ is basic box-parametrized for $q \ge q_0$ for some q_0, in the sense that the sequence has the desired structure for $q \ge q_0$.

Example 2.4.1. The sequence $T_q = v_0^{(1)} \otimes \sum_{i=1}^q v_i^{(2)} \otimes v_i^{(3)}$ is basic box-parametrized for $q \ge 1$, with support Σ_q defined by the conditions

$$0 \le i_1 \le 0, \quad 1 \le i_2 \le q, \quad 1 \le i_2 \le q, \quad i_2 = i_3.$$

We define a contraction operation between the j_1-th and the j_2-th factor of $\mathcal{T}$, obtained by summing over the corresponding indices: in other words, the contraction is the image of T via the trace map $\sum u_i^{(j_1)} \otimes u_i^{(j_2)}$ applied to the j_1-th and j_2-th factors, where $\{u_i^{(j)}\}$ is the dual basis to the fixed basis $\{v_i^{(j)}\}$ on the j-th factor.

Lemma 2.4.2. *Let $\mathcal{T}$, $\mathcal{T}'$ be basic box-parametrized tensors for $q \ge q_0$ and $q \ge q_0'$ respectively. Then*

- *$\mathcal{T} \otimes \mathcal{T}'$ is basic box-parametrized for $q \ge \max\{q_0, q_0'\}$;*
- *the contraction of $\mathcal{T}$ on factors j_1 and j_2 is basic box-parametrized for $q \ge \max\{|\vartheta_{j_1} -$*

$\vartheta_{j_2}|, |\Theta_{j_1} - \Theta_{j_2}|, q_0\}$; moreover, if the univariate coefficient $p(q)$ of $\mathcal{T}$ is a polynomial of degree e, then the coefficient of the tensor resulting from the contraction has degree at most $e+1$.

Proof. The first statement is immediate.

For the second statement, without loss of generality assume $j_1 = 1$ and $j_2 = 2$. First observe that if $\mathcal{T}$ is basic box-parametrized, then summing over the first index, or equivalently applying the linear map $\sum_i u_i^{(1)}$, generates a basic box-parametrized tensor; the coefficient of this tensor has the same degree as the coefficient of $\mathcal{T}$ unless the first index i_1 is not related by equality to any other index, and $\eta_1 = 0$ and $H_1 = 1$; in the latter case, the degree of the coefficient is increased by one.

Now, contraction of $\mathcal{T}$ on factors 1 and 2 is equivalent to first imposing the equality $i_1 = i_2$ on the support Φ_q and then summing up on the first and second index. Imposing the equality $i_1 = i_2$ effects the inequalities of i_1 and i_2 as follows:

$$\max\{\eta_1 q + \vartheta_1, \eta_2 q + \vartheta_2\} \leq i_1 = i_2 \leq \min\{H_1 q + \Theta_1, H_2 q + \Theta_2\}.$$

Each of the two bounds can be replaced by one the two linear functions (uniformly in q) whenever $q \geq \{|\vartheta_1 - \vartheta_2|, |\Theta_1 - \Theta_2|\}$. This, together with the previous observation, concludes the proof. $\square$

Given two sequences of tensors $\mathcal{T}^{(1)}, \mathcal{T}^{(2)}$ of order k, we define their sum as $\mathcal{T}_1 + \mathcal{T}_2 = \{T_q^{(1)} + T_q^{(2)} : q \in \mathbb{N}\}$. We say that a sequence $\mathcal{T}$ is *box parametrized* (for $q \geq q_0$) if $\mathcal{T}$ is a finite sum of basic box-parametrized sequences of tensors(for $q \geq q_0$). Observe that a sequence of tensors with constant dimensions is box parametrized if and only if its coefficients are univariate polynomials in q.

We will show that the maps $\Phi_1, \dots, \Phi_4$ in the proof of Theorem 1.7.2 in Section 2.2.3 and the maps $\Psi_1, \dots, \Psi_8$ in the proof of Theorem 1.7.3 in Section 2.2.4 are box parametrized.

The scripts in Appendix D perform the contraction of box parametrized tensors according to Lemma 2.4.2, keeping track of the univariate polynomial coefficients and of the lower bound q_0 for which the expressions are valid. The final result is that the maps $\Phi_1, \dots, \Phi_4$ are box parametrized for $q \geq 5$ and the maps $\Psi_1, \dots, \Psi_8$ are box parametrized for $q \geq 6$.

In the following, we show that the tensors involved in the various contractions are box parametrized. Lemma 2.4.2 guarantees that the results of the contractions are box parametrized as well.

First, notice that $T_{cw,q}$ is box parametrized for $q \geq 1$, as it is the sum of three tensors as the ones described in Example 2.4.1. By Lemma 2.4.2, we deduce that $T_{cw,q}^{\otimes 2}$ (regarded as a tensor of order 6) and $T_{cw,q}^{\otimes 3}$ (regarded as a tensor of order 9) are box parametrized. In all three cases, the polynomials defining the coefficients have degree 0.

2.4.1 Restriction

We show that the two restriction maps $\phi_2 : A^{\otimes 2} \to \mathbb{C}^3$ and $\phi_3 : A^{\otimes 3} \to \mathbb{C}^5$ are box parametrized as tensors of order 3 and 4 respectively.

Write $\phi_2 = X_0 \otimes e_0 + X_1 \otimes e_1 + X_2 \otimes e_2$, where $\mathbb{C}^3 = \langle e_0, e_1, e_2 \rangle$ and $X_0, X_1, X_2 \in A^{\otimes 2 *}$. It suffices to show that X_0, X_1, X_2 are box parametrized, regarded as tensors of order two. Using a basis dual to the basis of $A^{\otimes 2}$, we have

$$
\begin{aligned}
X_0 &= \alpha_0 \otimes \alpha_1 + \alpha_1 \otimes \alpha_0 + \alpha_1 \otimes \alpha_1 \\
X_1 &= \alpha_0 \otimes \textstyle\sum_1^q \alpha_i + \sum_1^q \alpha_i \otimes \alpha_0 \\
X_2 &= \alpha_0 \otimes \alpha_2 + \alpha_2 \otimes \alpha_0 + \alpha_2 \otimes \alpha_1 + \alpha_3 \otimes \alpha_3.
\end{aligned}
$$

This shows that X_0, X_1, X_2 are box parametrized.

Similarly, write $\phi_3 = Y_0 \otimes e_0 + \cdots + Y_4 \otimes e_4$, where $\mathbb{C}^5 = \langle e_0, \dots, e_4 \rangle$ and $Y_0, \dots, Y_4 \in A^{\otimes 3*}$. Directly from the definition in Section 2.2.4, it is immediate that $Y_0, \dots, Y_4$ are box parametrized and therefore ϕ_3 is box parametrized as well.

Applying Lemma 2.4.2, we deduce that the two sequences $\phi_2(T_{cw,q}^{\otimes 2})$ and $\phi_3(T_{cw,q}^{\otimes 3})$ are box parametrized.

2.4.2 Koszul maps

The Koszul differentials on $\mathbb{C}^3$ and $\mathbb{C}^5$ used in the definition of the Koszul flattenings are the skew-symmetric projections $\mathbb{C}^3 \otimes \mathbb{C}^3 \to \Lambda^2\mathbb{C}^3$ and $\Lambda^2\mathbb{C}^5 \otimes \mathbb{C}^5 \to \Lambda^3\mathbb{C}^5$. They are both fixed size, therefore they are box parametrized.

By Lemma 2.4.2, we deduce that the resulting Koszul flattenings $(\phi_2(T_{cw,q}^{\boxtimes 2}))^{\wedge 1}$ and $(\phi_3(T_{cw,q}^{\boxtimes 3}))^{\wedge 2}$ are box parametrized, regarded as tensors of order 6 and 8 respectively.

2.4.3 Diagonalizing maps

Recall that the maps $\Phi_1, \dots, \Phi_4$ in the proof of Theorem 1.7.2 and the maps $\Psi_1, \dots, \Psi_8$ in the proof of Theorem 1.7.3 are the restrictions of $(\phi_2(T_{cw,q}^{\boxtimes 2}))^{\wedge 1}$ and $(\phi_3(T_{cw,q}^{\boxtimes 3}))^{\wedge 2}$ to the multiplicity spaces of the isotypic components for the action of $\mathfrak{S}_{q-3}$ and $\mathfrak{S}_{q-5}$.

We analyze the square case in detail. For the square case, let $\mathcal{M}$ be the matrix of change of basis on $\mathbb{C}^q$ from the basis $\{e_1, \dots, e_q\}$ to the basis $\{e_1, e_2, e_3, \sum_4^q e_i, e_5 - e_4, \dots, e_q - e_{q-1}\}$. Explicitly

$$\mathcal{M} = \begin{bmatrix} \mathrm{Id}_3 & & & & \\ & 1 & 1 & \cdots & 1 \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix}.$$

In particular, $\mathcal{M}$ diagonalizes the action of $\mathfrak{S}_{q-3}$ and therefore the change of basis defined by $\mathrm{Id}_{\mathbb{C}^3} \boxtimes \mathcal{M}^{\boxtimes 2}$ on $\mathbb{C}^3 \otimes B^{\otimes 2}$ brings the matrix representing $(\phi_2(T_{cw,q}^{\boxtimes 2}))^{\wedge 1}$ into a block diagonal matrix, whose diagonal blocks are matrices representing the maps $f_i : \mathbb{C}^3 \otimes (M_i \otimes W_i) \to \Lambda^2\mathbb{C}^3 \otimes (M_i \otimes W_i)$ from (2.4); denote the diagonal blocks by $f_1^{\mathcal{M}}, \dots, f_4^{\mathcal{M}}$.

Because of our choice of basis, the multiplicity subspaces $\mathbb{C}^3 \otimes \langle w_i \rangle \otimes M_i$ and $\Lambda^2\mathbb{C}^3 \otimes \langle w_i \rangle \otimes M_i$ described in Section 2.2.3 are spanned by basis vectors, so that the matrices representing $\Phi_1, \dots, \Phi_4$ are given by submatrices of $f_1^{\mathcal{M}}, \dots, f_4^{\mathcal{M}}$. More precisely, setting π_{inv}, π_{std} to be the matrices of the two coordinate projections of $\mathbb{C}^q$ onto $\langle e_1, \dots, e_4 \rangle$ and $\langle e_5 \rangle$, we have

$$\Phi_1 = (\mathrm{Id}_{\Lambda^2\mathbb{C}^3} \boxtimes \pi_{inv} \boxtimes \pi_{inv}) \circ f_1^{\mathcal{M}} \circ (\mathrm{Id}_{\mathbb{C}^3} \boxtimes \pi_{inv} \boxtimes \pi_{inv})^T,$$
$$\Phi_2 = (\mathrm{Id}_{\Lambda^2\mathbb{C}^3} \boxtimes \pi_{inv} \boxtimes \pi_{std}) \circ f_2^{\mathcal{M}} \circ (\mathrm{Id}_{\mathbb{C}^3} \boxtimes \pi_{inv} \boxtimes \pi_{std})^T,$$
$$\Phi_3 = (\mathrm{Id}_{\Lambda^2\mathbb{C}^3} \boxtimes \pi_{std} \boxtimes \pi_{inv}) \circ f_3^{\mathcal{M}} \circ (\mathrm{Id}_{\mathbb{C}^3} \boxtimes \pi_{std} \boxtimes \pi_{inv})^T,$$
$$\Phi_4 = (\mathrm{Id}_{\Lambda^2\mathbb{C}^3} \boxtimes \pi_{std} \boxtimes \pi_{std}) \circ f_4^{\mathcal{M}} \circ (\mathrm{Id}_{\mathbb{C}^3} \boxtimes \pi_{std} \boxtimes \pi_{std})^T.$$

Since the composition can be performed on the single factors, by Lemma 2.4.2 it suffices to show that the four matrices $\mathcal{M}^{-1} \circ \pi_{inv}^T$, $\mathcal{M}^{-1} \circ \pi_{std}^T$, $\pi_{inv} \circ \mathcal{M}$ and $\pi_{std} \circ \mathcal{M}$ are box parametrized.

From the structure of $\mathcal{M}$, it is clear that $\pi_{inv} \circ \mathcal{M}$ and $\pi_{std} \circ \mathcal{M}$ are box parametrized. The computation of $\mathcal{M}^{-1}$ is straightforward, and it is easy to see that $\mathcal{M}^{-1} \circ \pi_{inv}^T$, $\mathcal{M}^{-1} \circ \pi_{std}^T$ are box parametrized.

This shows that $\Phi_1, \dots, \Phi_4$ are box parametrized. The script available in Appendix D computes the box parametrized representation of $\Phi_1, \dots, \Phi_4$ starting from the box parametrized version of T_{cw}, the restriction map ϕ_2, the Koszul differential and the four matrices $\mathcal{M}^{-1} \circ \pi_{inv}^T$, $\mathcal{M}^{-1} \circ \pi_{std}^T$, $\pi_{inv} \circ \mathcal{M}$ and $\pi_{std} \circ \mathcal{M}$.

The cube case is similar. Now, restriction space $\mathbb{C}^3$ is a $\mathbb{C}^5$, the top left block in the matrix $\mathcal{M}$ is a 5×5 identity block, the result of the conjugation by $\mathcal{M}$ is block diagonal with 8 blocks, corresponding to the eight isotypic components. The coordinate projections π_{inv} and

π_{std} are onto $\langle e_1, \ldots, e_6 \rangle$ and $\langle e_7 \rangle$. The script computes the box parametrized representation of the matrices $\Psi_1, \ldots, \Psi_8$.

3. BORDER APOLARITY OF TENSORS

3.1 History of border rank lower bounds

This chapter deals exclusively with border rank lower bounds. Initially border rank lower bounds for tensors were obtained by finding a polynomial vanishing on the set of tensors of border rank at most r, $\sigma_r(\mathrm{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C))$, 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 [37], 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 [38, 33]. Such methods are subject to *barriers* [31, 39] (see [40, §2.2] for an overview). A technique allowing one to go slightly beyond the barriers was introduced in [41]. 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 [41] 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). 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 [42], 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 [42] 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 chapter 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 §3.2.3 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).

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 $M_{\langle 2,\mathbf{n},\mathbf{n}\rangle}$ and $M_{\langle 3,\mathbf{n},\mathbf{n}\rangle}$ 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 computed by adding local contributions. This enabled us to set-up an optimization problem to bound all possible sums of local contributions, which we then solved (Lemma 3.7.9) 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 §3.4, then working dually to reduce the calculation (Remark 3.4.1), then using representation theory to block diagonalize the calculation in §3.6.2, and finally we observe that the result is an immediate consequence of our localization principle and Lemma 3.7.1 (Remark 3.7.2).

3.1.1 Results

Theorem 3.1.1. $\underline{\mathbf{R}}(M_{\langle 3\rangle}) \geq 17$.

The previous lower bounds were 14 [37] in 1983, 15 [33] in 2015, and 16 [41] in 2018.

Theorem 3.1.2. $\underline{\mathbf{R}}(\det_3) \geq 17$.

In [43] 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 [44], which follows from the Koszul flattening equations (§2.2).

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

Theorem 3.1.3. $\underline{\mathbf{R}}(M_{\langle 2,2,3\rangle}) = 10$.

The upper bound dates back to Bini et. al. in 1980 [2]. Koszul flattenings give $\underline{\mathbf{R}}(M_{\langle 2,2,\mathbf{n}\rangle}) \geq 3\mathbf{n}$ [33]. Smirnov [45] showed that $\underline{\mathbf{R}}(M_{\langle 2,2,\mathbf{n}\rangle}) \leq 3\mathbf{n}+1$ for $\mathbf{n} \leq 7$, and we expect equality to hold for all $\mathbf{n}$.

Theorem 3.1.4.

1. $\underline{\mathbf{R}}(M_{\langle 2,3,3\rangle}) = 14$.

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 [46, Rem.

p175], see §3.8.

The upper bound in (1) is due to Smirnov [45], where he also proved $\underline{\mathbf{R}}(M_{\langle 2,4,4\rangle}) \leq 24$, and $\underline{\mathbf{R}}(M_{\langle 2,5,5\rangle}) \leq 38$. 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 $M_{\langle 2,\mathbf{n},\mathbf{n}\rangle} = M_{\langle 2,2,2\rangle} \boxtimes M_{\langle 1,\frac{\mathbf{n}}{2},\frac{\mathbf{n}}{2}\rangle}$.

Theorem 3.1.5. *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 [46, Rem. p175], see §3.8.

Using [46, Rem. p175], one obtains

Corollary 3.1.6. *For all* $\mathbf{n} \geq 18$ *and* $\mathbf{m} \geq 3$, $\underline{\mathbf{R}}(M_{\langle \mathbf{m},\mathbf{n},\mathbf{n}\rangle}) \geq \mathbf{n}^2 + \sqrt{\frac{8}{3}}\mathbf{n} + \mathbf{m} - 3$.

Remark 3.1.7. Koszul flattenings 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 $M_{\langle 2,\mathbf{n},\mathbf{n}\rangle} \in \mathbb{C}^{2\mathbf{n}} \otimes \mathbb{C}^{2\mathbf{n}} \otimes \mathbb{C}^{\mathbf{n}^2}$ and $M_{\langle 3,\mathbf{n},\mathbf{n}\rangle} \in \mathbb{C}^{3\mathbf{n}} \otimes \mathbb{C}^{3\mathbf{n}} \otimes \mathbb{C}^{\mathbf{n}^2}$. Theorems 3.1.4 and 3.1.5 show that the border apolarity method does not share this defect.

3.2 Preliminaries

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). 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.

3.2.1 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

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

Here

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

3.2.2 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

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

denote the $\mathbb{Z}^3$-graded ideal of the set of r points $[T_1(t)] \sqcup \cdots \sqcup [T_r(t)]$, where $I_{ijk,t} \subset S^i A^* \otimes S^j B^* \otimes S^k C^*$. If the r points are in general position, then $\mathrm{codim}(I_{ijk,t}) = r$ as long as $r \leq \dim S^i A^* \otimes S^j B^* \otimes S^k C^*$ (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., [47, 39, 48]). For all (ijk) with $i + j + k > 1$, we may choose the curves such that $\mathrm{codim}(I_{ijk}) = r$ by [42, Thm. 1.2].

Thus, in addition to $E_0 = I_{111,0}^{\perp}$ defined in §3.2.1, 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 $G(\dim(S^i A^* \otimes S^j B^* \otimes S^k C^*) - r, S^i A^* \otimes S^j B^* \otimes S^k C^*)$. We remark that there are subtleties here: the limiting ideal may not be saturated. See [42] 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 $i + j + k > 1$, $\mathrm{codim}\, I_{ijk} = r$.

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

$$I_{i-1,j,k} \otimes A^* \oplus I_{i,j-1,k} \otimes B^* \oplus I_{i,j,k-1} \otimes C^* \to S^i A^* \otimes S^j B^* \otimes S^k C^* \tag{3.1}$$

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.

3.2.3 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 (see §2.1) 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 [41], 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 [42], 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.

3.2.4 Borel fixed subspaces

We review standard facts about Borel fixed subspaces. In this chapter 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 $e_1, \ldots, e_m$, with dual basis $e^1,\dots,e^m$. Assign e_j weight $(0,\dots,0,1,0,\dots,0)$, where the 1 is in the j-th slot, and e^j weight $(0,\dots,0,-1,0,\dots,0)$. For vectors in $(\mathbb{C}^m)^{\otimes d}$, $wt(e_1^{\otimes a_1}\otimes\cdots\otimes e_m^{\otimes a_m}) = (a_1,\dots,a_m)$ and the weight is unchanged under permutations of the $d = a_1+\cdots+a_m$ factors. Partially order the weights so that $(i_1,\dots,i_m)\geq(j_1,\dots,j_m)$ 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 $\pi = (p_1,\dots,p_m)$ and the module is denoted $S_\pi\mathbb{C}^m$. Write $d = |\pi| = \sum p_i$. See any of, e.g., [30, §8.7], [49, §9.1], or [50, 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 $\mathrm{wt}(e_1)=\omega_1$, $\mathrm{wt}(e_m)=-\omega_{m-1}$, for $2\leq s\leq m-1$, $\mathrm{wt}(e_s)=\omega_s-\omega_{s-1}$ and for all j, $\mathrm{wt}(e^j) = -\mathrm{wt}(e_j)$. 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.

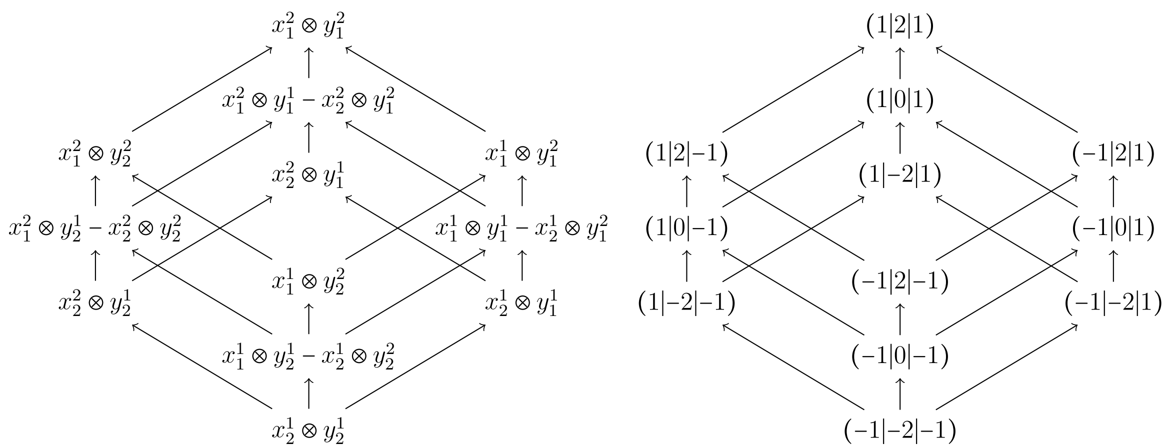

Figure 3.1: Weight diagram for $U^* \otimes \mathfrak{sl}(V) \otimes W$ when $U = V = W = \mathbb{C}^2$. Left are the weight vectors and right the weights: since $\mathfrak{sl}_2$ weights are just $j\omega_1$, we have just written $(i|j|k)$ for the $\mathfrak{sl}(U) \oplus \mathfrak{sl}(V) \oplus \mathfrak{sl}(W)$ 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.

Example 3.2.1. When U, V, W each have dimension 2, Figure 3.1 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 $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ (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..

Example 3.2.2. Let $\dim U = 3$. Figure 3.2 gives the weight diagram for $U \otimes U = S^2 U \oplus \Lambda^2 U$. 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, $\langle (u_1)^2, s u_1 u_2 + t u_1 \wedge u_2 \rangle$ plus an isolated one $\langle u_1 \wedge u_2, u_1 \wedge u_3 \rangle$ etc..

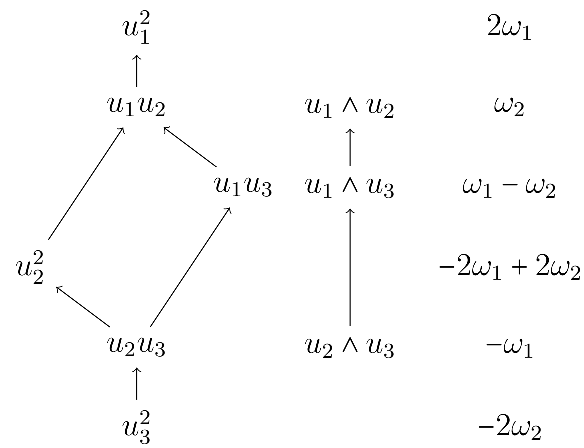

Figure 3.2: Weight diagram for $U \otimes U$ when $U = \mathbb{C}^3$. There are 6 distinct weights appearing, indicated on the right. On the far left are the weight vectors in S^2U and in the middle are the weight vectors in $\Lambda^2 U$.

Example 3.2.3. Figure 3.3 gives the weight diagram for $\mathfrak{sl}_3$. Here $v_j^i = v_j \otimes v^i$. 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 $X = \langle v_1^3, v_2^3, 2v_3^3 - (v_1^1 + v_2^2)\rangle$, $X = \langle v_1^3, v_1^2, 2v_1^1 - (v_2^2 + v_3^3)\rangle$ and $X = \langle v_1^3, v_1^2, v_2^3\rangle$.

The $\mathbb{B}$-fixed subspaces of dimension 4 are a family parametrized by $[s,t] \in \mathbb{P}^1$: $X = \langle v_1^3, v_2^3, s(2v_3^3 - (v_1^1 + v_2^2)) + t(2v_1^1 - (v_2^2 + v_3^3))\rangle$.

The $\mathbb{B}$-fixed subspaces of dimension 5 are, the weight ≥ 0 space, $X = \langle v_1^3, v_1^2, v_2^3, v_2^2 - v_2^3, v_3^2\rangle$, and $X = \langle v_1^3, v_1^2, v_2^3, v_1^1 - v_2^2, v_2^1\rangle$.

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

3.3 The algorithm

Input: An integer r and a concise tensor $T \in A \otimes B \otimes C$ whose symmetry group contains a

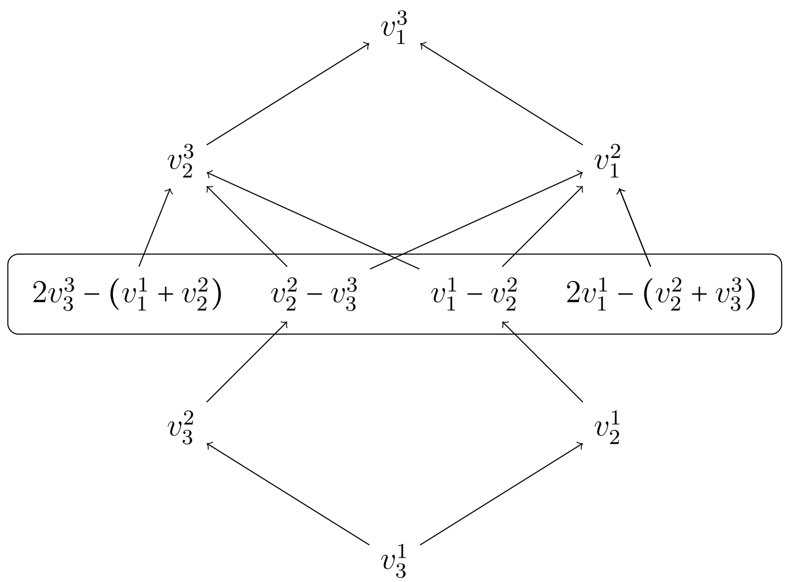

Figure 3.3: Weight diagram for $\mathfrak{sl}_3$.

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

$$F_{110} \otimes A^* \to S^2 A^* \otimes B^*, \text{ and} \tag{3.2}$$

$$F_{110} \otimes B^* \to A^* \otimes S^2 B^*. \tag{3.3}$$

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

$$F_{110} \otimes C^* \oplus F_{101} \otimes B^* \oplus F_{011} \otimes A^* \to A^* \otimes B^* \otimes C^*. \tag{3.4}$$

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 (3.4) 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

$$F_{i-1,j,k} \otimes A^* \oplus F_{i,j-1,k} \otimes B^* \otimes F_{i,j,k-1} \otimes C^* \to S^iA^* \otimes S^jB^* \otimes S^kC^*. \tag{3.5}$$

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 (3.5) 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 [42]. 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 3.3.1. *The codimension of the image of the* (210)*-map is the dimension of the kernel of the skew-symmetrization map*

$$F_{110}^{\perp} \otimes A \to \Lambda^2 A \otimes B. \tag{3.6}$$

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

$$(F_{ij,k-1}^{\perp} \otimes C) \cap (F_{i,j-1,k}^{\perp} \otimes B) \cap (F_{i-1,j,k}^{\perp} \otimes A). \tag{3.7}$$

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

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

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 (3.6).

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. $\square$

3.4 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 $G_{M_{\langle \mathbf{u},\mathbf{v},\mathbf{w}\rangle}} \supseteq \mathrm{GL}(U) \times \mathrm{GL}(V) \times \mathrm{GL}(W) =: G$. 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 $M_{\langle n \rangle}$ 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 3.1.

By Figure 3.1, there are three $\mathbb{B}$-fixed 2-planes in $U^* \otimes \mathfrak{sl}(V) \otimes W$:

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

For the first, the rank of the 24×40 matrix of the map $E^{\perp}_{110} \otimes A^* \to S^2 A^* \otimes B^*$ 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 the result follows. □

For readers unhappy with computing the rank of a sparse 40×24 matrix whose entries are all $0, \pm 1$, the following remark reduces to 24×24 matrices, and in §3.6.2, 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 3.7.2.

Remark 3.4.1. One can simplify the calculation of the rank of the map $E^{\perp}_{110} \otimes A^* \to S^2 A^* \otimes B^*$ by using the map (3.6). 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.

3.5 Explanation of the proofs of Theorems 3.1.1 and 3.1.2

The actual proofs to these theorems are in the code at the webpage `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 $E_{110}^{\perp} \otimes A^* \to S^2 A^* \otimes B^*$, and $E_{110}^{\perp} \otimes B^* \to A^* \otimes S^2 B^*$. 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 recall the relevant module structure for the determinant, first discussed in §2.1.1:

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 $\mathfrak{S}_3$. 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 $U^{\otimes 2}\otimes V^{\otimes 2}=S^2U\otimes S^2V\oplus S^2U\otimes\Lambda^2V\oplus\Lambda^2U\otimes S^2V\oplus\Lambda^2U\otimes\Lambda^2V$, and $\det_3(C^*)=\Lambda^2U\otimes\Lambda^2V$.

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 $u_1\wedge u_2\otimes v_1v_2$ each have weight $(\omega_2^U|\omega_2^V)$. We examine all 7-dimensional $\mathbb{B}$-fixed subspaces of $S^2U\otimes S^2V\oplus S^2U\otimes\Lambda^2V\oplus\Lambda^2U\otimes S^2V$. 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, the tests all reduce to determining the ideal in which a given matrix with polynomial entries drops rank to at most a given value r. If the matrix or r is small enough, we can simply take the ideal of $r+1\times r+1$ minors. Sometimes, however, this is computationally infeasible. In this case, 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_1J_2. Its zero set is the rank $<r$ locus and computing with it is tractable.

3.6 Representation theory relevant for matrix multiplication

Theorems 3.1.3 and 3.1.4(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.

3.6.1 Refinement of the (210) test for matrix multiplication

Recall $A = U^* \otimes V$, $B = V^* \otimes W$, $C = W^* \otimes U$. 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 $\mathbf{v} = 3$, $V_{\omega_2+\omega_{\mathbf{v}-1}} = V_{2\omega_2}$):

$$\Lambda^2(U^* \otimes V) \otimes V^* = (S^2U^* \otimes V_{\omega_1}) \oplus (\Lambda^2 U^* \otimes V_{\omega_1}) \oplus (S^2U^* \otimes V_{\omega_2+\omega_{\mathbf{v}-1}}) \oplus (\Lambda^2 U^* \otimes V_{2\omega_1+\omega_{\mathbf{v}-1}}), \tag{3.8}$$

$$S^2(U^* \otimes V) \otimes V^* = (S^2U^* \otimes V_{2\omega_1+\omega_{\mathbf{v}-1}}) \oplus (\Lambda^2 U^* \otimes V_{\omega_2+\omega_{\mathbf{v}-1}}) \oplus (S^2U^* \otimes V_{\omega_1}) \oplus (\Lambda^2 U^* \otimes V_{\omega_1}), \tag{3.9}$$

$$A \otimes M_{\langle \mathbf{u},\mathbf{v},\mathbf{w}\rangle}(C^*) = (U^* \otimes V) \otimes (U^* \otimes \mathrm{Id}_V \otimes W) = (S^2U^* \otimes V_{\omega_1} \otimes W) \oplus (\Lambda^2 U^* \otimes V_{\omega_1} \otimes W), \tag{3.10}$$

$$V \otimes \mathfrak{sl}(V) = V_{\omega_1} \oplus V_{2\omega_1+\omega_{\mathbf{v}-1}} \oplus V_{\omega_2+\omega_{\mathbf{v}-1}}, \tag{3.11}$$

$$\begin{aligned}(U^* \otimes V) \otimes (U^* \otimes \mathfrak{sl}(V)) = {} & (S^2U^* \otimes V_{2\omega_1+\omega_{\mathbf{v}-1}}) \oplus (\Lambda^2 U^* \otimes V_{2\omega_1+\omega_{\mathbf{v}-1}}) \oplus (S^2U^* \otimes V_{\omega_1}) \\ & \oplus (\Lambda^2 U^* \otimes V_{\omega_1}) \oplus (S^2U^* \otimes V_{\omega_2+\omega_{\mathbf{v}-1}}) \oplus (\Lambda^2 U^* \otimes V_{\omega_2+\omega_{\mathbf{v}-1}}).\end{aligned} \tag{3.12}$$

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 3.6.1. *Write* $E_{110} := M_{\langle \mathbf{u},\mathbf{v},\mathbf{w}\rangle}(C^*) \oplus E'_{110}$. *The dimension of the kernel of the map* (3.6) $E_{110} \otimes A \to \Lambda^2 A \otimes B$ *equals the dimension of the kernel of the map*

$$E'_{110} \otimes A \to S^2U^* \otimes V_{\omega_2+\omega_{\mathbf{v}-1}} \otimes W \oplus \Lambda^2U^* \otimes V_{2\omega_1+\omega_{\mathbf{v}-1}} \otimes W, \tag{3.13}$$

and the kernel of (3.13) *is*

$$(E'_{110} \otimes A) \cap [U^{*\otimes 2} \otimes V_{\omega_1} \otimes W \oplus S^2U^* \otimes V_{2\omega_1+\omega_{\mathbf{v}-1}} \otimes W \oplus \Lambda^2U^* \otimes V_{\omega_2+\omega_{\mathbf{v}-1}} \otimes W]. \tag{3.14}$$

The second assertion follows by applying Schur's lemma using (3.12) as the map (3.13) is the restriction of an equivariant map.

3.6.2 $\underline{\mathbf{R}}(M_{\langle 2\rangle}) > 6$ revisited

In this case the map (3.13) takes image in $\Lambda^2U^* \otimes S^2V \otimes V^* \otimes W$. 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 (3.13) 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 Λ^2U^*.) 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

$$x_1^1 \wedge (x_1^2 \otimes y_1^1 - x_2^2 \otimes y_1^2) \quad 1$$
$$x_2^1 \wedge (x_1^2 \otimes y_1^1 - x_2^2 \otimes y_1^2) \quad -1$$

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 and be 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 (3.13) 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 (3.13) 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 (3.13) must have dimension four, and the result follows.

3.7 Proofs of Theorems 3.1.4 and 3.1.5

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

$$K_g = (\Sigma_g \otimes A) \cap [U^{*\otimes 2} \otimes V_{\omega_1} \otimes W \oplus S^2U^* \otimes V_{2\omega_1+\omega_{\mathbf{v}-1}} \otimes W \oplus \Lambda^2 U^* \otimes V_{\omega_2+\omega_{\mathbf{v}-1}} \otimes W]. \quad (3.15)$$

Then the dimension of (3.14) can be written as the sum over g of $\dim K_g/K_{g-1}$, and we may upper bound the dimension of (3.14) 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 3.7.1, and for $\mathfrak{sl}_3$, this is Lemma 3.7.3. Bounds on the kernel of the (120) map are obtained by symmetry; specifically, the bound is the same as that on (3.14) 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 3.7.6.

3.7.1 The local argument

Lemma 3.7.1. *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* (3.13) *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 3.7.1 is proved later this section.

Remark 3.7.2. 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 3.7.1, 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 3.1.3. 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 3.7.1, 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 3.1.4(1),(2). For Theorem 3.1.4(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 3.1.4(2), the result follows by similar complete enumeration of outer structures on a computer. □

Lemma 3.7.3. *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* (3.13) *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 3.7.1 and 3.7.3, we first observe the following:

Proposition 3.7.4. *The included module $V_{\omega_1} \subset V \otimes \mathfrak{sl}(V)$ has weight basis $\overline{v}_i = \sum_{j \neq i} [\mathbf{v} v_j \otimes (v_i \otimes v^j) - v_i \otimes (v_j \otimes v^j)] + (\mathbf{v}-1) v_i \otimes v_i \otimes v^i$, $1 \leq i \leq \mathbf{v}$.*

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 3.7.1 and 3.7.3. 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 (3.15). Write

$$K = U^{*\otimes 2} \otimes V_{\omega_1} \otimes W \oplus S^2 U^* \otimes V_{2\omega_1 + \omega_{\mathbf{v}-1}} \otimes W \oplus \Lambda^2 U^* \otimes V_{\omega_2 + \omega_{\mathbf{v}-1}} \otimes W \tag{3.16}$$

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 $U^{*s} = \mathrm{span}\{u^{\mathbf{n}}, \ldots, u^{\mathbf{n}-s+1}\}$ and $W_{(t-1)} =$

$\mathrm{span}\{w_1, \ldots, w_t\}$. Hence, $S_f \cap K_g$ is a flag of K_g with $K_{g-1} = S_0 \cap K_g$. Use the isomorphism

$$\frac{K_g \cap S_f}{K_g \cap S_{f-1}} = \frac{K_g \cap S_f + S_{f-1}}{S_{f-1}} \tag{3.17}$$

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

$$\frac{U^{*\otimes 2} \otimes V_3 \otimes W}{S_{f-1}} = \frac{U^{*\otimes 2}}{U^* \otimes U^{*(s-1)}} \otimes \frac{V_3}{V_{f-1}} \otimes \frac{W}{W_{(t-1)}}. \tag{3.18}$$

Write K^f for the f-th summand of (3.16), 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 (3.18), 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 (3.17) 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 3.7.4, $Y = \operatorname{span}\{\overline{v}_i \mid i \le \mathbf{y}\}$.

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 $\mathbf{y} \le \mathbf{v} - 1$, and the trivial $\mathbf{y}' \le \dim V_{2\omega_1 + \omega_{\mathbf{v}-1}}$, and $\mathbf{y}'' \le \dim V_{\omega_2 + \omega_{\mathbf{v}-1}}$ give the asserted upper bounds.

By similar reasoning when $\mathbf{v} = 3$, considering Example 3.2.3, we obtain the bounds $\mathbf{y} = 0$ when $j = 1, 2$ and $\mathbf{y} \le 1$ when $j = 3, 4, 5, 6$. For all values of j except 1, the result then follows from

$$\dim K_g - \dim K_{g-1} = (j\mathbf{v} - \mathbf{y})s + \mathbf{y}\mathbf{n} - \mathbf{y}'' \le (j\mathbf{v} - \mathbf{y})s + \mathbf{y}\mathbf{n}, \tag{3.19}$$

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 $\mathbf{y}'' \ge 2$. 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 $\mathbf{y}'' \ge 2$ 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 (3.19), and for $j = 2$, we must show $\mathbf{y} \ge 1$. 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. $\square$

Remark 3.7.5. 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 3.7.6 below the linear functions of s in the lemmas above appear as $a_{\mu_{s,t}}s + b_{\mu_{s,t}}$.

3.7.2 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 3.7.6. *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*

$$\min\Big\{\sum_{s,t} a_{\mu_{s,t}}s + b_{\mu_{s,t}}, \sum_{s,t} a_{\mu_{s,t}}t + b_{\mu_{s,t}}\Big\} \le \max_{1\le j\le k}\left\{\frac{a_j\rho^2}{8j^2} + (a_j+b_j)\frac{\rho}{j}\right\}. \tag{3.20}$$

Remark 3.7.7. 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 3.7.6 is proved in §3.7.3.

Proof of Theorem 3.1.4(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 3.7.6 with $k = 3$ and a_i and b_i from Lemma 3.7.1 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 3.1.5. Proceeding in the same way as in the proof of Theorem 3.1.4(3), we apply Lemma 3.7.6 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 3.7.3. We obtain the smaller of the kernel dimensions of the (120) and (210) maps is at most the largest of the following,

j	Lemma 3.7.6	j	Lemma 3.7.6
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 $2 \leq j \leq 8$ 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. □

3.7.3 Proof of Lemma 3.7.6

For a partition $\lambda = (\lambda_1, \ldots, \lambda_q)$, 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 $\mathbf{n} \geq \mathbf{m}$.

To establish Lemma 3.7.6 we need two additional lemmas:

Lemma 3.7.8. *Let λ be a partition not of the form $(n,2)$. Then $n(\lambda) \leq \frac{1}{8}(|\lambda| + \lambda'_1 - \lambda_1)^2 - \frac{1}{8}$. In particular, for all λ, $n(\lambda) \leq \frac{1}{8}(|\lambda| + \lambda'_1 - \lambda_1)^2$.*

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.

$$\begin{aligned}
n(\lambda) &= n(\mu) + \tbinom{\lambda'_k}{2} \\
&\le \tfrac{1}{8}(|\mu| + \mu'_1 - \mu_1)^2 - \tfrac{1}{8} + \tbinom{\lambda'_k}{2} \\
&= \tfrac{1}{8}(|\lambda| - \lambda'_k + \lambda'_1 - (\lambda_1 - 1))^2 - \tfrac{1}{8} + \tfrac{1}{2}\lambda'_k(\lambda'_k - 1) \\
&= \tfrac{1}{8}(|\lambda| + \lambda'_1 - \lambda_1)^2 - \tfrac{1}{8} - \tfrac{1}{4}(|\lambda| + \lambda'_1 - \lambda_1 - \tfrac{5}{2}\lambda'_k + \tfrac{1}{2})(\lambda'_k - 1)
\end{aligned}$$

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 3.7.9. *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:*

$$\min\Big\{\sum_{i\le k} c_i x_i^2 + d_i(x_i + y_i), \sum_{i\le k} c_i y_i^2 + d_i(x_i+y_i)\Big\} \le \max_{1\le j\le k}\left\{\frac{\rho^2}{4j^2}C_j + \frac{\rho}{j}D_j\right\}. \tag{3.21}$$

Remark 3.7.10. 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 $x'_i, y'_i \geq 0$ 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 (3.21) 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 $E_L \coloneqq \sum_{i \leq k} c_i x_i^2 + d_i(x_i + y_i)$ and $E_R \coloneqq \sum_{i \leq k} c_i y_i^2 + d_i(x_i + y_i)$. 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 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 $\sum_{i\le k} c_i x_i^2 + d_i(x_i+y_i) = \sum_{i\le k} c_i y_i^2 + d_i(x_i+y_i)$, 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 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 3.7.6. For each $1 \le i \le k$, let λ^i be the partition corresponding to the boxes of μ labeled $\geq i$. Write $a_0 = b_0 = 0$. Then

$$\begin{aligned}
\textstyle\sum_{s,t} a_{\mu_{s,t}} s + b_{\mu_{s,t}} &= \textstyle\sum_{s,t} \sum_{i=1}^{\mu_{s,t}} (a_i - a_{i-1})s + b_i - b_{i-1} \\
&= \textstyle\sum_{i=1}^{k} \sum_{s,t\in\lambda^i} (a_i - a_{i-1})s + b_i - b_{i-1} \\
&= \textstyle\sum_{i=1}^{k} (a_i - a_{i-1}) n(\lambda^i) + (a_i - a_{i-1} + b_i - b_{i-1})|\lambda^i| \\
&\leq \textstyle\sum_{i=1}^{k} \left[\frac{1}{2}(a_i - a_{i-1})\right] \left(\frac{1}{2}(|\lambda^i| + (\lambda^i)'_1 - \lambda^i_1)\right)^2 + \left[a_i - a_{i-1} + b_i - b_{i-1}\right]|\lambda^i| \qquad (3.22)
\end{aligned}$$

where we have used Lemma 3.7.8 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 (3.22) becomes

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

Similarly, $\sum_{s,t} a_{\mu_{s,t}} t + b_{\mu_{s,t}} \leq \sum_{i=1}^{k} c_i y_i^2 + d_i(x_i + y_i)$. Now, $\sum_i x_i + y_i = \sum_i |\lambda^i| = \rho$ and the x_i and y_i are each nonnegative and non-increasing. Hence, by Lemma 3.7.9,

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

as required. □

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

Here is a simple proof of the statement, which was originally shown in [46]. By the border substitution method [41], for any tensor $T \in A \otimes B \otimes C$

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

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 [41]. 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.

4.1 A rank 18 Waring decomposition of $sM_{\langle 3\rangle}$*

Let V have dimension three. Pick a basis of V to fix an identification of 3×3 matrices with $V^*\otimes V$, and consider the 18 matrices $m_1,\ldots,m_{18}$ below, where $\zeta = e^{2\pi i/3}$ and $a = -2^{1/3}$.

$$m_1 = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad m_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -\zeta \\ 0 & -\zeta^2 & 1 \end{pmatrix} \quad m_3 = \begin{pmatrix} 1 & 0 & -\zeta \\ 0 & 0 & 0 \\ -\zeta^2 & 0 & 1 \end{pmatrix}$$

$$m_4 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -\zeta^2 \\ 0 & -\zeta & 1 \end{pmatrix} \quad m_5 = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} \quad m_6 = \begin{pmatrix} 1 & -\zeta & 0 \\ -\zeta^2 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$m_7 = \begin{pmatrix} 1 & 0 & -\zeta^2 \\ 0 & 0 & 0 \\ -\zeta & 0 & 1 \end{pmatrix} \quad m_8 = \begin{pmatrix} 1 & -\zeta^2 & 0 \\ -\zeta & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad m_9 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

$$m_{10} = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix} \quad m_{11} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & \zeta \\ \zeta^2 & 0 & 0 \end{pmatrix} \quad m_{12} = \begin{pmatrix} 0 & 0 & 1 \\ \zeta^2 & 0 & 0 \\ 0 & \zeta & 0 \end{pmatrix}$$

$$m_{13} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & \zeta^2 \\ \zeta & 0 & 0 \end{pmatrix} \quad m_{14} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad m_{15} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \zeta & 0 \\ 0 & 0 & \zeta^2 \end{pmatrix}$$

$$m_{16} = \begin{pmatrix} 0 & 0 & 1 \\ \zeta & 0 & 0 \\ 0 & \zeta^2 & 0 \end{pmatrix} \quad m_{17} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \zeta^2 & 0 \\ 0 & 0 & \zeta \end{pmatrix} \quad m_{18} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

* **Acknowledgement:** This section is derived in part from an article published in Experimental Mathematics April 5, 2019 ©Taylor & Francis 2019, available online: 10.1080/10586458.2018.1547231

Theorem 4.1.1. $sM_{\langle 3\rangle} = \frac{1}{6}\sum_{i=1}^{18} m_i^{(3)}$. *That is, the m_i form a rank 18 Waring decomposition of $6sM_{\langle 3\rangle}$.*

The group $\Gamma = \mathrm{GL}(V^* \otimes V)$ naturally acts on $S^3(V^* \otimes V)$, and the stabilizer of $sM_{\langle n\rangle}$ is $\Gamma_{sM_{\langle n\rangle}} = \mathrm{PGL}(V) \rtimes \mathbb{Z}_2$ [51]. Here the action of $\mathrm{PGL}(V)$ is induced by its natural action on $V^* \otimes V$, and, after choosing a basis and its dual, $\mathbb{Z}_2$ acts as matrix transposition. Such a choice of matrix transposition is not canonical, but any choice generates the same group modulo $\mathrm{PGL}(V)$.

Notice that any m_i could be replaced by ζm_i as these matrices define the same rank 1 tensor.

4.1.1 Symmetries of the decomposition

To study symmetry, we wish to consider the m_i modulo this identification. Therefore, write $T_i = m_i^{(3)}$, the rank one symmetric tensors corresponding to the m_i, and define the *symmetry of the decomposition* as the subgroup Γ_S of $\Gamma_{sM_{\langle n\rangle}}$ which leaves the set $S = \{T_1, \ldots, T_{18}\}$ invariant under the natural induced action on subsets of $S^3(V^* \otimes V)$. A symmetry of the decomposition preserves the set $\{m_1, \ldots, m_{18}\}$ up to powers of ζ.

Theorem 4.1.2. *The symmetry group $\Gamma_S \cong (\mathbb{Z}_3^2 \rtimes \mathrm{SL}(2, \mathbb{F}_3)) \rtimes \mathbb{Z}_2$, which has order 432.*

The expression in parentheses is the $\mathrm{PGL}(V)$ action, and the $\mathbb{Z}_2$ is generated by matrix transposition with respect to the basis of the decomposition. To describe the $\mathrm{PGL}(V)$ part of the action, we label each 3×3 block of matrices with elements of the vector space $\mathbb{F}_3^2$ as follows:

$$\begin{array}{ccc} (0,0) & (0,1) & (0,2) \\ (1,0) & (1,1) & (1,2) \\ (2,0) & (2,1) & (2,2). \end{array}$$

Then $\mathbb{Z}_3^2 \rtimes \mathrm{SL}(2,\mathbb{F}_3)$ acts on the first 3×3 block as affine transformations of $\mathbb{F}_3^2$ according to this labelling: $\mathbb{Z}_3^2$ acts by translation and $\mathrm{SL}(2,\mathbb{F}_3)$ acts by linear transformation. On the second 3×3 block, $\mathbb{Z}_3^2$ acts trivially and $\mathrm{SL}(2,\mathbb{F}_3)$ again acts as linear transformations. One can alternatively view the action of $\mathbb{Z}_3^2 \rtimes \mathrm{SL}(2,\mathbb{F}_3)$ on the second 3×3 block as equivalent to the action on its normal subgroup $\mathbb{Z}_3^2$ by conjugation.

The decomposition is also closed under complex conjugation, which acts by transposing each 3×3 block. A Galois-type symmetry like this is not in general in Γ and represents another kind of symmetry of decompositions of tensors defined over $\mathbb{Q}$. There are no other nontrivial Galois symmetries for this decomposition, for any such symmetry must be an automorphism of $\mathbb{Q}[\zeta]$ fixing $\mathbb{Q}$. Including complex conjugation as a symmetry of the decomposition yields a group of order 864.

Proof of Theorem 4.1.2. We first describe the representation $\rho : \mathbb{Z}_3^2 \rtimes \mathrm{SL}(2,\mathbb{F}_3) \to \mathrm{PGL}(V)$ explicitly by giving the images of a generating set. These elements of $\mathrm{PGL}(V)$ can then be observed to act as claimed on the 3×3 blocks. Let e_r and e_d denote the generators of $\mathbb{Z}_3^2$ corresponding to translation right and down, respectively, and denote elements of $\mathrm{SL}(2,\mathbb{F}_3)$

by their matrices with respect to the standard basis of $\mathbb{F}_3^2$. Then

$$\rho(e_r) = \begin{pmatrix} 0 & 0 & 1 \\ \zeta^2 & 0 & 0 \\ 0 & \zeta & 0 \end{pmatrix} \quad \rho(e_d) = \begin{pmatrix} 0 & 0 & 1 \\ \zeta & 0 & 0 \\ 0 & \zeta^2 & 0 \end{pmatrix}$$

$$\rho\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -\zeta+1 & \zeta^2-1 & 2\zeta+1 \\ \zeta^2-1 & -\zeta+1 & 2\zeta+1 \\ -\zeta+1 & -\zeta+1 & -\zeta+1 \end{pmatrix}$$

$$\rho\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} -\zeta^2+1 & \zeta-1 & -\zeta^2+1 \\ \zeta-1 & -\zeta^2+1 & -\zeta^2+1 \\ \zeta-1 & \zeta-1 & -2\zeta-1 \end{pmatrix}$$

$$\rho\begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \zeta^2 \end{pmatrix}.$$

It remains to show there are no symmetries of the decomposition other than those claimed. Name the entries of a 3×3 block by the numbers $1, \ldots, 9$, like on a telephone. Since all symmetries of $\Gamma_{sM_{\langle 3 \rangle}}$ preserve matrix rank, we first observe that any symmetry of the decomposition must preserve in particular the first 3×3 block. This, combined with the fact that there is evidently a matrix transposition in Γ_S, shows it is sufficient to check the set of $\mathrm{PGL}(V)$ symmetries of the first 3×3 block is as claimed. Call this group G. We wish to show $G = \mathbb{Z}_3^2 \rtimes \mathrm{SL}(2, \mathbb{F}_3)$. The first block consists of only rank 1 matrices, so they uniquely determine column vectors up to multiplication by scalars. Let H denote the symmetry group of the corresponding projective configuration of points in $\mathbb{P}^2$. The vectors corresponding to matrices $(1, 3, 4, 6)$ are in general linear position, so each element of H determines at most one element of $\mathrm{PGL}(V)$ which induces it. Hence, the natural homomorphism $G \to H$ is injective, so it suffices to show $H \leq \mathbb{Z}_3^2 \rtimes \mathrm{SL}(2, \mathbb{F}_3)$.

First we show $\mathbb{Z}_3^2 \rtimes \mathrm{GL}(2, \mathbb{F}_3)$ is the symmetry group of the combinatorial affine plane consisting of 9 points and 12 lines determined by the points and collinearity relations of the configuration (Figure 4.1). Clearly $\mathbb{Z}_3^2$ are symmetries of this configuration. To see that $\mathrm{GL}(2, \mathbb{F}_3)$ are also symmetries, notice that we may identify points of the configuration with the group $\mathbb{Z}^2/3\mathbb{Z}^2$, and any line through points in the lattice $\mathbb{Z}^2$ projects down to one of our 12 lines when modding out by $3\mathbb{Z}^2$. Then since $\mathrm{GL}(2, \mathbb{Z})$ preserves the lines of $\mathbb{Z}^2$, it must be that $\mathrm{GL}(2, \mathbb{F}_3)$ preserves the lines of our configuration $\mathbb{Z}^2/3\mathbb{Z}^2$, as desired. Observe that any symmetry is determined by the image of 3 noncollinear points. For instance, fixing the image of 1,2, and 5 determines by collinearity the image of 3,8, and 9, which in turn determines the image the remaining 3 points. Then, since $\mathbb{Z}_3^2 \rtimes \mathrm{GL}(2, \mathbb{F}_3)$ are all indeed symmetries of the configuration, the full symmetry group has order at most $9 \cdot 8 \cdot 7 = 504$ and contains a subgroup of size $9 \cdot 48 = 432$. The only possibility is then equality with $\mathbb{Z}_3^2 \rtimes \mathrm{GL}(2, \mathbb{F}_3)$, as claimed.

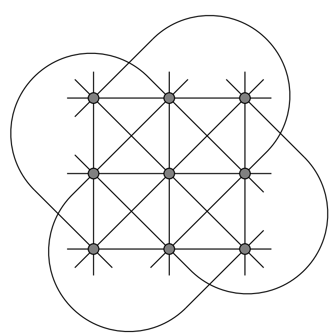

Figure 4.1: The configuration determined by column vectors of the rank one block. This is classically known as the Hesse configuration [52].

Now we show that the elements of $\mathbb{Z}_3^2 \rtimes \mathrm{GL}(2, \mathbb{F}_3)$ where the second factor has determinant -1 do not induce symmetries of the projective configuration of points. Because $\mathbb{Z}_3^2 \rtimes \mathrm{SL}(2, \mathbb{F}_3)$ does induce such symmetries, it suffices to show the failure for only one element. A convenient choice is the map $\mathbb{F}_3^2 \to \mathbb{F}_3^2$ which interchanges coordinates. The unique matrix taking the general frame $(1,2,7,8)$ to $(1,4,3,6)$ is

$$\begin{pmatrix} 0 & -\zeta^2 & -1 \\ -\zeta^2 & 0 & -\zeta \\ 0 & 0 & \zeta^2 \end{pmatrix},$$

and one readily checks this matrix does not send, e.g. 3 to any of the other points. Hence $\mathbb{Z}_3^2 \rtimes \mathrm{SL}(2,\mathbb{F}_3) \leq G \leq H \leq \mathbb{Z}_3^2 \rtimes \mathrm{SL}(2,\mathbb{F}_3)$, and the full symmetry group is $(\mathbb{Z}_3^2 \rtimes \mathrm{SL}(2,\mathbb{F}_3)) \rtimes \mathbb{Z}_2$, as claimed. □

The rank one block of the decomposition consists of orthogonal projections onto one dimensional subspaces times a factor of two. In this sense, each such matrix is determined by its column space. We have already seen that these 9 points of $\mathbb{P}^2$ form a certain projective configuration (Figure 4.1). It is a classical fact that any set of 9 points in this configuration are the inflections points of a plane cubic. Indeed, our configuration is precisely the inflection points of $x^3 + y^3 + z^3 = 0$. Equivalently, it is determined as the zeros of $x^3 + y^3 + z^3 = 0$ and its Hessian $xyz = 0$.

The Waring decomposition presented here was derived from a numerical decomposition given in [12]. I would like to thank Grey Ballard for his work transforming that numerical decomposition into a sparse numerical one.

4.2 A rank 40 Waring decomposition of $sM_{\langle 4 \rangle}$

Let V have dimension four, and fix a basis, with respect to which all matrices in $\mathrm{PGL}(V)$ and $V^* \otimes V$ will be presented. Let $\Gamma \leq \mathrm{PGL}(V)$, be the group generated by the following matrices,

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \begin{pmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ i & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \end{pmatrix} \quad \begin{pmatrix} \frac{1}{2}i+\frac{1}{2} & i & 0 & \frac{1}{2}i-\frac{1}{2} \\ -\frac{1}{2}i-\frac{1}{2} & 0 & 0 & -\frac{1}{2}i-\frac{1}{2} \\ \frac{1}{2}i-\frac{1}{2} & 0 & 0 & -\frac{1}{2}i+\frac{1}{2} \\ \frac{1}{2}i+\frac{1}{2} & 0 & i & -\frac{1}{2}i+\frac{1}{2} \end{pmatrix}.$$

Consider the following matrices m_1, m_2, and m_3 of $V^* \otimes V$ (here $i^2 = -1$).

$$m_1 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2}i \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2}i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad m_2 = \sqrt[3]{2} \begin{pmatrix} \frac{1}{4}i+\frac{1}{4} & -\frac{1}{4}i-\frac{1}{4} & -\frac{1}{4}i-\frac{1}{4} & 0 \\ \frac{1}{4}i-\frac{1}{4} & -\frac{1}{4}i+\frac{1}{4} & 0 & \frac{1}{4}i+\frac{1}{4} \\ 0 & -\frac{1}{4}i+\frac{1}{4} & 0 & 0 \\ \frac{1}{4}i-\frac{1}{4} & 0 & 0 & 0 \end{pmatrix},$$

$$m_3 = \sqrt[3]{2} \begin{pmatrix} 0 & 0 & 0 & -\frac{1}{4}i-\frac{1}{4} \\ 0 & 0 & \frac{1}{4}i+\frac{1}{4} & 0 \\ \frac{1}{4}i+\frac{1}{4} & -\frac{1}{2} & \frac{1}{4}i-\frac{1}{4} & -\frac{1}{4}i+\frac{1}{4} \\ -\frac{1}{2}i & \frac{1}{4}i+\frac{1}{4} & \frac{1}{4}i+\frac{1}{4} & -\frac{1}{4}i-\frac{1}{4} \end{pmatrix}$$

Theorem 4.2.1. $\Gamma \cong \mathfrak{S}_4$, *and the orbits of the* m_i *under* Γ *acting by conjugation form a rank 40 Waring decomposition of* $sM_{\langle 4 \rangle}$, *with orbit sizes 8, 8, and 24, respectively.*

We omit the proof, but the claims can be verified entirely computationally.

4.2.1 Symmetries of the decomposition

The decomposition of §4.2 there is an orbit consisting of rank one matrices, similar to the decomposition of §4.1. If we consider the eight column vectors associated to this rank one orbit as points of $\mathbb{P}^3$ and compute which subsets of them lie on a common plane, we find that there are exactly six nontrivial such planes which contain four points. Moreover, one can identify the eight points of $\mathbb{P}^3$ with the vertices of a cube in such a way that the coplanarity relations precisely correspond to the faces of the cube.

Now, any symmetry of the decomposition must preserve this configuration, so must in particular must be contained inside the symmetries of the cube, $\mathfrak{S}_4 \times \mathbb{Z}_2$. One can then argue similarly to the proof of Theorem 4.1.2 to see that in fact there are no other symmetries of the decomposition beside Γ (that is, one checks that a nonproper symmetry of the cube does not actually induce a symmetry of the decomposition).

What is not as obvious here is that there is a transpose symmetry. However, the action following element of $\mathrm{PGL}(V)$ after transposition preserves the decomposition,

$$\begin{pmatrix} 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & \frac{1}{2}i \\ \frac{1}{2} & 0 & 0 & \frac{1}{2}i \\ 0 & \frac{1}{2}i & \frac{1}{2}i & 0 \end{pmatrix}.$$

More specifically, this symmetry fixes the first orbit pointwise, and transposes pairs of points in the other two orbits.

We have proved

Theorem 4.2.2. *The symmetry group $\Gamma_S \cong \mathfrak{S}_4 \rtimes \mathbb{Z}_2$, which has order 48.*

5. CONCLUSIONS

In this chapter I discuss possible improvements and future work regarding the contents of this thesis.

Theorem 1.7.2 and 1.7.3 were proved with Koszul flattenings, but this technique could not prove results for the smallest values of q. It remains to decide the border rank of $T_{cw,2}^{\boxtimes 2}$ and $T_{cw,q}^{\boxtimes 3}$ for $q = 2, 3, 4$. Assuming that multiplicativity also holds for these values, this result requires proving lower bounds, and a natural idea is to apply the lower bound technique of chapter 3. Carrying out this process is future work.

The technique of §2.3.3 remains to be fully understood and automated; I anticipate that it may be used to obtain border rank decompositions for many other small tensors.

The border apolarity technique of chapter 3 is still new, and it is current work to obtain further lower bounds for the ordinary matrix multiplication tensor. In §3.7, the technique was used in its most basic form to obtain results on an infinite sequence of tensors, but only the (210) and (120) tests were applied. One direction of generalization is therefore to apply the higher tests of border apolarity, in particular the (111) test, to infinite families of tensors. Another direction for future work is implied by the algorithm of §3.3, which can fail to prove a border rank lower bound by producing a list of candidate ideals. If this happens, the next step is to either rule out these candidate ideals via some other method or to use them to construct a border rank decomposition.

The results of chapter 4 are suggestive that it may be profitable to search for decompositions with large symmetry groups for symmetrized matrix multiplication. For instance, if a symmetry group and orbit structure may be guessed in advance, the search space for a fixed small decomposition may be reduced to the point of tractability. If one can find small decomposition examples whose corresponding symmetry groups and representations occur in natural infinite families, one may be able to generalize to an infinite sequence of decompositions. If this program were successful, it would imply an upper bound to ω by Theorem 1.5.1.

REFERENCES

[1] V. Strassen, “Gaussian elimination is not optimal,” *Numer. Math.*, vol. 13, pp. 354–356, 1969.

[2] D. Bini, G. Lotti, and F. Romani, “Approximate solutions for the bilinear form computational problem,” *SIAM J. Comput.*, vol. 9, no. 4, pp. 692–697, 1980.

[3] A. Schönhage, “Partial and total matrix multiplication,” *SIAM J. Comput.*, vol. 10, no. 3, pp. 434–455, 1981.

[4] D. Coppersmith and S. Winograd, “Matrix multiplication via arithmetic progressions,” *J. Symbolic Comput.*, vol. 9, no. 3, pp. 251–280, 1990.

[5] A. Stothers, *On the Complexity of Matrix Multiplication.* PhD thesis, University of Edinburgh, 2010.

[6] V. Williams, “Multiplying matrices in $o(n^{2.373})$ time.” `https://people.csail.mit.edu/virgi/matrixmult-f.pdf`, 2014.

[7] F. Le Gall, “Powers of tensors and fast matrix multiplication,” in *Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation*, ISSAC ’14, (New York, NY, USA), pp. 296–303, ACM, 2014.

[8] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms.* MIT Press and McGraw–Hill, second ed., 2001.

[9] P. Bürgisser, M. Clausen, and M. A. Shokrollahi, *Algebraic complexity theory*, vol. 315 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Math-*

ematical Sciences]. Berlin: Springer-Verlag, 1997. With the collaboration of Thomas Lickteig.

[10] D. Bini, M. Capovani, F. Romani, and G. Lotti, "$O(n^{2.7799})$ complexity for $n \times n$ approximate matrix multiplication," *Inform. Process. Lett.*, vol. 8, no. 5, pp. 234–235, 1979.

[11] J. M. Landsberg, "The border rank of the multiplication of 2×2 matrices is seven," *J. Amer. Math. Soc.*, vol. 19, no. 2, pp. 447–459, 2006.

[12] L. Chiantini, J. D. Hauenstein, C. Ikenmeyer, J. M. Landsberg, and G. Ottaviani, "Polynomials and the exponent of matrix multiplication," *Bull. Lond. Math. Soc.*, vol. 50, no. 3, pp. 369–389, 2018.

[13] A. Ambainis, Y. Filmus, and F. Le Gall, "Fast matrix multiplication: limitations of the Coppersmith-Winograd method (extended abstract)," in *STOC'15—Proceedings of the 2015 ACM Symposium on Theory of Computing*, pp. 585–593, ACM, New York, 2015.

[14] J. Alman and V. V. Williams, "Further limitations of the known approaches for matrix multiplication," in *9th Innovations in Theoretical Computer Science Conference, ITCS 2018, January 11-14, 2018, Cambridge, MA, USA*, pp. 25:1–25:15, 2018.

[15] J. Alman and V. V. Williams, "Limits on all known (and some unknown) approaches to matrix multiplication," in *2018 IEEE 59th Ann. Symp. Found. Comp. Sc. (FOCS)*, pp. 580–591, 2018.

[16] M. Christandl, P. Vrana, and J. Zuiddam, "Barriers for fast matrix multiplication from irreversibility," *CoRR*, vol. abs/1812.06952, 2018.

[17] J. Alman, "Limits on the universal method for matrix multiplication," in *34th Comp. Compl. Conf. (CCC 2019)*, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019.

[18] M. Bläser, *Fast Matrix Multiplication.* No. 5 in Graduate Surveys, Theory of Computing Library, 2013.

[19] M. Bläser and V. Lysikov, "On degeneration of tensors and algebras," in *41st International Symposium on Mathematical Foundations of Computer Science*, vol. 58 of *LIPIcs. Leibniz Int. Proc. Inform.*, pp. Art. No. 19, 11, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2016.

[20] J. M. Landsberg and M. Michałek, "A $2n^2 - \log_2(n) - 1$ lower bound for the border rank of matrix multiplication," *Int. Math. Res. Not. IMRN*, no. 15, pp. 4722–4733, 2018.

[21] A. Conner, F. Gesmundo, J. M. Landsberg, and E. Ventura, "Tensors with maximal symmetries," *arXiv e-prints*, p. arXiv:1909.09518, Sep 2019.

[22] D. A. Gay, "Characters of the Weyl group of $SU(n)$ on zero weight spaces and centralizers of permutation representations," *Rocky Mountain J. Math.*, vol. 6, no. 3, pp. 449–455, 1976.

[23] J. M. Landsberg, *Tensors: geometry and applications*, vol. 128 of *Graduate Studies in Mathematics.* Providence, RI: American Mathematical Society, 2012.

[24] H. Derksen, "On the nuclear norm and the singular value decomposition of tensors," *Found. Comp. Math.*, vol. 16, no. 3, pp. 779–811, 2016.

[25] N. Ilten and Z. Teitler, "Product ranks of the 3×3 determinant and permanent," *Canad. Math, Bull.*, vol. 59, no. 2, pp. 311–319, 2016.

[26] J. M. Landsberg and G. Ottaviani, "Equations for secant varieties of Veronese and other varieties," *Ann. Mat. Pura Appl. (4)*, vol. 192, no. 4, pp. 569–606, 2013.

[27] V. Strassen, "Rank and optimal computation of generic tensors," *Lin. Alg. Appl.*, vol. 52/53, pp. 645–685, 1983.

[28] E. Ballico, A. Bernardi, M. Christandl, and F. Gesmundo, "On the partially symmetric rank of tensor products of W-states and other symmetric tensors," *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.*, vol. 30, pp. 93–124, 2019.

[29] H. Derksen and V. Makam, "On non-commutative rank and tensor rank," *Linear Multilinear Algebra*, vol. 66, no. 6, pp. 1069–1084, 2018.

[30] J. M. Landsberg, *Geometry and complexity theory*, vol. 169 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2017.

[31] K. Efremenko, A. Garg, R. Oliveira, and A. Wigderson, "Barriers for rank methods in arithmetic complexity," in *9th Innovations in Theoretical Computer Science*, vol. 94 of *LIPIcs. Leibniz Int. Proc. Inform.*, pp. Art. No. 1, 19, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2018.

[32] W. Fulton and J. Harris, *Representation theory*, vol. 129 of *Graduate Texts in Mathematics*. New York: Springer-Verlag, 1991. A first course, Readings in Mathematics.

[33] J. M. Landsberg and G. Ottaviani, "New lower bounds for the border rank of matrix multiplication," *Theory Comput.*, vol. 11, pp. 285–298, 2015.

[34] F. Gesmundo, C. Ikenmeyer, and G. Panova, "Geometric complexity theory and matrix powering," *Diff. Geom. Appl.*, vol. 55, pp. 106–127, 2017.

[35] W. Fulton, *Young tableaux. With applications to representation theory and geometry*, vol. 35 of *London Mathematical Society Student Texts*. Cambridge: Cambridge University Press, 1997.

[36] A. K. Lenstra, H. W. L. Jr, and L. Lovász, "Factoring Polynomials with Rational Coefficients," *Math. Ann*, vol. 261, no. 4, p. 515534, 1982.

[37] V. Strassen, "Rank and optimal computation of generic tensors," *Linear Algebra Appl.*, vol. 52/53, pp. 645–685, 1983.

[38] J. M. Landsberg and G. Ottaviani, "Equations for secant varieties of Veronese and other varieties," *Ann. Mat. Pura Appl. (4)*, vol. 192, no. 4, pp. 569–606, 2013.

[39] M. Gał azka, "Vector bundles give equations of cactus varieties," *Linear Algebra Appl.*, vol. 521, pp. 254–262, 2017.

[40] J. M. Landsberg, *Tensors: Asymptotic Geometry and Developments 20162018*, vol. 132 of *CBMS Regional Conference Series in Mathematics.* AMS, 2019.

[41] 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.*, vol. 1, no. 1, pp. 2–19, 2017.

[42] W. Buczyńska and J. Buczyński, "Apolarity, border rank and multigraded Hilbert scheme," *arXiv:1910.01944*, 2019.

[43] M. Boij and Z. Teitler, "A bound for the Waring rank of the determinant via syzygies," *Linear Algebra and its Applications*, vol. 587, pp. 195 – 214, 2020.

[44] C. Farnsworth, "Koszul-Young flattenings and symmetric border rank of the determinant," *J. Algebra*, vol. 447, pp. 664–676, 2016.

[45] A. V. Smirnov, "The bilinear complexity and practical algorithms for matrix multiplication," *Comput. Math. Math. Phys.*, vol. 53, no. 12, pp. 1781–1795, 2013.

[46] T. Lickteig, "A note on border rank," *Inform. Process. Lett.*, vol. 18, no. 3, pp. 173–178, 1984.

[47] M. Gallet, K. Ranestad, and N. Villamizar, "Varieties of apolar subschemes of toric surfaces," *Ark. Mat.*, vol. 56, no. 1, pp. 73–99, 2018.

[48] Z. Teitler, "Geometric lower bounds for generalized ranks," *arXiv e-prints*, p. arXiv:1406.5145, Jun 2014.

[49] C. Procesi, *Lie groups.* Universitext, New York: Springer, 2007. An approach through invariants and representations.

[50] I. G. Macdonald, *Symmetric functions and Hall polynomials.* Oxford Mathematical Monographs, New York: The Clarendon Press Oxford University Press, second ed., 1995. With contributions by A. Zelevinsky, Oxford Science Publications.

[51] G. P. Fulvio Gesmundo, Christian Ikenmeyer, "Geometric complexity theory and matrix powering," *Differential Geometry and its Applications*, vol. 55, pp. 106–127, 2017.

[52] O. Hesse, "über die elimination der variabeln aus drei algebraischen gleichungen vom zweiten grade mit zwei variabeln," *Journal für die Reine und Angewandte Mathematik*, vol. 28, pp. 68–96, 1844.