A $2\mathbf{n}^2 - \log_2(\mathbf{n}) - 1$ LOWER BOUND FOR THE BORDER RANK OF MATRIX MULTIPLICATION

J.M. LANDSBERG AND MATEUSZ MICHAŁEK

Abstract. Let $M_{\langle \mathbf{n}\rangle} \in \mathbb{C}^{\mathbf{n}^2} \otimes \mathbb{C}^{\mathbf{n}^2} \otimes \mathbb{C}^{\mathbf{n}^2}$ denote the matrix multiplication tensor for $\mathbf{n}\times\mathbf{n}$ matrices. We use the border substitution method [2, 3, 6] combined with Koszul flattenings [8] to prove the border rank lower bound $\underline{\mathbf{R}}(M_{\langle \mathbf{n},\mathbf{n},\mathbf{n}\rangle}) \geq 2\mathbf{n}^2 - \lceil \log_2(\mathbf{n}) \rceil - 1$.

1. Introduction

Let A, B, C, U, V, W be vector spaces of dimensions $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{u}, \mathbf{v}, \mathbf{w}$. The matrix multiplication tensor $M_{\langle \mathbf{u},\mathbf{v},\mathbf{w}\rangle} \in (U^* \otimes V) \otimes (V^* \otimes W) \otimes (W^* \otimes U)$ is given in coordinates by

$$M_{\langle \mathbf{u},\mathbf{v},\mathbf{w}\rangle} = \sum_{i=1}^{\mathbf{u}} \sum_{j=1}^{\mathbf{v}} \sum_{k=1}^{\mathbf{w}} x^i_j \otimes y^j_k \otimes z^k_i .$$

Ever since Strassens' discovery [11] that the standard algorithm for multiplying matrices is not optimal, the matrix multiplication tensor has been a central object of study. We write $M_{\langle \mathbf{n}\rangle} = M_{\langle \mathbf{n},\mathbf{n},\mathbf{n}\rangle}$.

Let $T \in A \otimes B \otimes C$ be a tensor. The *rank* of T is the smallest r such that T may be written as a sum of r rank one tensors (tensors of the form $a \otimes b \otimes c$ for $a \in A, b \in B, c \in C$). The *border rank* of T is the smallest r such that T may be written as a limit of rank r tensors. We write $\underline{\mathbf{R}}(T) = r$. Border rank is a basic measure of the complexity of a tensor. For example, the exponent of matrix multiplication, the smallest ω such that $\mathbf{n}\times\mathbf{n}$ matrix multiplication can be computed with $O(\mathbf{n}^\omega)$ arithmetic operations, satisfies $\omega = \underline{\lim}_{\mathbf{n}\to\infty} \log_{\mathbf{n}}(\underline{\mathbf{R}}(M_{\langle \mathbf{n}\rangle}))$. All modern upper and lower bounds for the complexity of matrix multiplication rely implicitly or explicitly on border rank. Strassen showed $\underline{\mathbf{R}}(M_{\langle \mathbf{n}\rangle}) \geq \frac{3\mathbf{n}^2}{2}$ [10] and Lickteig improved this to $\underline{\mathbf{R}}(M_{\langle \mathbf{n}\rangle}) \geq \frac{3\mathbf{n}^2}{2} + \frac{\mathbf{n}}{2} - 1$ [9]. After that, progress stalled for nearly thirty years (other than showing $\underline{\mathbf{R}}(M_{\langle 2\rangle}) = 7$ [5]), until in 2012 the first author and Ottaviani showed $\underline{\mathbf{R}}(M_{\langle \mathbf{n}\rangle}) \geq 2\mathbf{n}^2 - \mathbf{n}$ [8]. In 2016 we improved this to $\underline{\mathbf{R}}(M_{\langle \mathbf{n}\rangle}) \geq 2\mathbf{n}^2 - \mathbf{n} + 1$ [7]. More important than the result in [7] was the method of proof - a border rank version of the *substitution method* [2, 3, 6] that we now review.

A tensor $T \in A \otimes B \otimes C$ is *A-concise* if it is not contained in any $\tilde{A} \otimes B \otimes C$ where $\tilde{A} \subsetneq A$. Let $G(k, V)$ denote the Grassmannian of k-planes through the origin in V.

Proposition 1.1. [3, 6] *Let $T \in A \otimes B \otimes C$ be A-concise. Fix $\mathbf{a}' \leq \mathbf{a}$. Then*

$$\underline{\mathbf{R}}(T) \geq \min_{A' \in G(\mathbf{a}', A^*)} \underline{\mathbf{R}}(T|_{A' \otimes B^* \otimes C^*}) + (\mathbf{a} - \mathbf{a}').$$

This method at first glance appears very hard to implement, as one would have to test every $\mathbf{a}'$-plane. As explained in [6], it is useful for tensors with symmetry because one can use the

Key words and phrases. matrix multiplication complexity, border rank, tensor, commuting matrices, Strassen's equations, Koszul flattening, substitution method, MSC 68Q17, 14L30, 15A69.

Landsberg supported by NSF grant DMS-1405348. Michalek was supported by Iuventus Plus grant 0301/IP3/2015/73 of the Polish Ministry of Science.

symmetry group to reduce testing to a finite number of $\mathbf{a}'$-planes, which is what we do in this article to prove the following theorem:

Theorem 1.2. *Let $0 < m < \mathbf{n}$. Then for all $\mathbf{w}$,*

$$\underline{\mathbf{R}}(M_{\langle \mathbf{n},\mathbf{n},\mathbf{w}\rangle}) \geq 2\mathbf{n}\mathbf{w} - \mathbf{w} + m - \lfloor \frac{\mathbf{w}\binom{\mathbf{n}-1+m}{m-1}}{\binom{2\mathbf{n}-2}{\mathbf{n}-1}} \rfloor.$$

In particular, taking $\mathbf{w} = \mathbf{n}$ and $m = \mathbf{n} - \lceil \log_2(\mathbf{n}) \rceil - 1$,

$$\underline{\mathbf{R}}(M_{\langle \mathbf{n}\rangle}) \geq 2\mathbf{n}^2 - \lceil \log_2(\mathbf{n}) \rceil - 1.$$

As can be seen in the proof, one can get a slightly better lower bound. Here are a few cases with optimal m and the improvement over the previous bound:

$\mathbf{n}$	$\underline{\mathbf{R}}(M_{\langle \mathbf{n}\rangle}) \geq$	improvement over $2\mathbf{n}^2 - \mathbf{n} + 1$
4	29	0
5	47	1
6	69	2
7	95	3
8	122	3
9	158	4
10	196	6
100	19,992	92
1000	1,999,989	989
10,000	199,999,985	9985.

One might expect that the substitution and border substitution methods could potentially be used to prove rank and border rank lower bounds up to $3\mathbf{m} - 3$ for tensors in $\mathbb{C}^{\mathbf{m}} \otimes \mathbb{C}^{\mathbf{m}} \otimes \mathbb{C}^{\mathbf{m}}$. We show this is not quite possible for border rank. We define a variety $X(\mathbf{a}', \mathbf{b}', \mathbf{c}') \subset \mathbb{P}(A \otimes B \otimes C)$ that corresponds to tensors where the border substitution method fails to provide lower bounds beyond $\mathbf{a} + \mathbf{b} + \mathbf{c} - \mathbf{a}' - \mathbf{b}' - \mathbf{c}'$. More precisely, $X(\mathbf{a}', \mathbf{b}', \mathbf{c}')$ is the variety of $(\mathbf{a}', \mathbf{b}', \mathbf{c}')$-*compressible tensors*, those for which there exists $A' \subset A^*$, $B' \subset B^*$, $C' \subset C^*$, respectively of dimensions $\mathbf{a}', \mathbf{b}', \mathbf{c}'$, such that T, considered as a linear form on $A^* \otimes B^* \otimes C^*$, satisfies $T\mid_{A' \otimes B' \otimes C'} = 0$. We show:

Proposition 1.3. *The set $X(\mathbf{a}', \mathbf{b}', \mathbf{c}') \subseteq \mathbb{P}(A \otimes B \otimes C)$ is Zariski closed. If*

$$\mathbf{a}\mathbf{a}' + \mathbf{b}\mathbf{b}' + \mathbf{c}\mathbf{c}' < (\mathbf{a}')^2 + (\mathbf{b}')^2 + (\mathbf{c}')^2 + \mathbf{a}'\mathbf{b}'\mathbf{c}' \tag{1}$$

then $X(\mathbf{a}', \mathbf{b}', \mathbf{c}') \subsetneq \mathbb{P}(A \otimes B \otimes C)$. In particular, in the range where (1) *holds, the substitution methods may be used to prove nontrivial lower bounds for border rank.*

The proof and examples show that beyond this bound one expects $X(\mathbf{a}', \mathbf{b}', \mathbf{c}') = \mathbb{P}(A \otimes B \otimes C)$, so that the method cannot be used.

If $\underline{\mathbf{R}}(T) \leq \mathbf{a} + \mathbf{b} + \mathbf{c} - (\mathbf{a}' + \mathbf{b}' + \mathbf{c}')$ then there exists $A' \subset A^*, B' \subset B^*, C' \subset C^*$ such that $T|_{A' \otimes B' \otimes C'} = 0$, because if $T = \sum_{j=1}^{\mathbf{a}+\mathbf{b}+\mathbf{c}-(\mathbf{a}'+\mathbf{b}'+\mathbf{c}')} a_j \otimes b_j \otimes c_j$, one can choose A' to annhilate $a_1, \cdots, a_{\mathbf{a}-\mathbf{a}'}$, B' to annihilate $b_{\mathbf{a}-\mathbf{a}'+1}, \cdots, b_{\mathbf{a}-\mathbf{a}'+\mathbf{b}-\mathbf{b}'}$ and C' to annihilate $c_{\mathbf{a}-\mathbf{a}'+\mathbf{b}-\mathbf{b}'+1}, \cdots, c_{\mathbf{a}+\mathbf{b}+\mathbf{c}-(\mathbf{a}'+\mathbf{b}'+\mathbf{c}')}$, and for a border rank decomposition one takes the limits of such planes from the sequence of rank decompositions converging to it. Let $\sigma_r(Seg(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)) \subset \mathbb{P}(A \otimes B \otimes C)$ denote the variety of tensors of border rank at most r, called the r-th secant variety of the Segre variety. The above remark may be restated as:

Proposition 1.4.

$$\sigma_{\mathbf{a}+\mathbf{b}+\mathbf{c}-(\mathbf{a}'+\mathbf{b}'+\mathbf{c}')}Seg(\mathbb{P}A\times\mathbb{P}B\times\mathbb{P}C)\subset X(\mathbf{a}',\mathbf{b}',\mathbf{c}').$$

We expect the inequality in Proposition 1.3 to be sharp or nearly so. For tensors in $\mathbb{C}^{\mathbf{m}}\otimes\mathbb{C}^{\mathbf{m}}\otimes\mathbb{C}^{\mathbf{m}}$ the limit of this method alone would be a border rank lower bound of $3(\mathbf{m}-\sqrt{3\mathbf{m}+\frac{9}{4}}+\frac{3}{2})$. However, it is unlikely the method alone could attain such a bound due to technical difficulties in proving an explicit tensor does not belong to $X(\mathbf{a}',\mathbf{b}',\mathbf{c}')$.

The state of the art for matrix multiplication is such that on one hand, for upper bounds on the exponent there does not appear to be a viable path proposed for proving the exponent is less than 2.3, but on the other hand, none of the existing techniques provide a viable path for proving a border rank lower bound of $2\mathbf{n}^2$ for matrix multiplication.

1.1. **Acknowledgements.** We thank Jason Starr and Math Overflow for help with Example 4.5. Michalek is a member of the AGATES group and is supported by the Foundation for Polish Science (FNP).

2. Preliminaries

Let $A=U^*\otimes V$, $B=V^*\otimes W$, $C=U\otimes W^*$. For $v\in V$, we write $\hat{v}\subset V$ for the line it determines and $[v]\in\mathbb{P}V$ for the corresponding point in projective space.

Definition 2.1. For a tensor $T\in V_1\otimes\ldots\otimes V_n$, and $U\subset V_1$, let $T/U\in(V_1/U)\otimes V_2\otimes\ldots\otimes V_n$ denote $T|_{U^\perp\otimes V_2^*\otimes\cdots\otimes V_n^*}$, where we consider T as a linear form on $V_1^*\otimes\cdots\otimes V_n^*$. Define

$$B_k(T):=\{[v]\in\mathbb{P}V_1\mid \underline{\mathbf{R}}(T/\hat{v})\le k\}.$$

Lemma 2.2. *Let $T\in V_1\otimes\ldots\otimes V_n$ be a tensor, let $G_T\subset\mathrm{GL}(V_1)\times\cdots\times\mathrm{GL}(V_n)$ denote its stabilizer and let $G_1\subset GL(V_1)$ denote the projection of G_T to $GL(V_1)$. The set $B_k(T)$ is:*

- i *Zariski closed,*
- ii *a G_1-variety.*

Proof. Proof of (i): Let $\mathcal{L}$ be the total space of the quotient bundle over $\mathbb{P}V_1$ tensored with $V_2\otimes\cdots\otimes V_n$, i.e., the fiber over $[v]$ is $(V_1/\hat{v})\otimes V_2\otimes\ldots\otimes V_n$. We have a natural section $s:\mathbb{P}V_1\to\mathcal{L}$ defined by $s([v]):=T/\hat{v}$. Let $X\subset\mathcal{L}$ denote the sub-bundle whose fiber over $[v]\in\mathbb{P}V_1$ is the locus of tensors of border rank at most k in $(V_1/\hat{v})\otimes V_2\otimes\ldots\otimes V_n$. The set $B_k(T)$ is the projection to $\mathbb{P}V_1$ of the intersection of the image of the section s and X.

Proof of (ii). We need to show that for all $g_1\in G_1$ and $[v]\in B_k(T)$, that $g_1[v]\in B_k(T)$. Let $g=(g_1,\ldots,g_n)\in G_T$. Then $\underline{\mathbf{R}}(T/\hat{v})=\underline{\mathbf{R}}(gT/g_1\hat{v})=\underline{\mathbf{R}}(T/g_1\hat{v})$. □

To prove Theorem 1.2, we will use the Koszul flattening of [8]: for $T\in A\otimes B\otimes C$, define

$$T_A^{\wedge p}:B^*\otimes\Lambda^pA\to\Lambda^{p+1}A\otimes C \tag{2}$$

by first taking $T_B\otimes\mathrm{Id}_{\Lambda^p A}:B^*\otimes\Lambda^pA\to\Lambda^pA\otimes A\otimes C$, and then projecting to $\Lambda^{p+1}A\otimes C$. If $\{a_i\},\{b_j\},\{c_k\}$ are bases of A,B,C and $T=\sum_{i,j,k}t^{ijk}a_i\otimes b_j\otimes c_k$, then

$$T_A^{\wedge p}(\beta\otimes f_1\wedge\cdots\wedge f_p)=\sum_{i,j,k}t^{ijk}\beta(b_j)a_i\wedge f_1\wedge\cdots\wedge f_p\otimes c_k. \tag{3}$$

We have [8]:

$$\underline{\mathbf{R}}(T) \geq \frac{\operatorname{rank}(T_A^{\wedge p})}{\binom{\mathbf{a}-1}{p}}. \tag{4}$$

In practice the map $T_A^{\wedge p}$ is used after specializing T to a subspace of A of dimension $2p+1$ to get a potential $\frac{2p+1}{p+1}\mathbf{b}$ border rank lower bound.

3. Proof of Theorem 1.2

We first observe that the "In particular" assertion follows from the main assertion because, taking $m = \mathbf{n} - c$, we want c such that

$$\frac{\mathbf{n}\binom{2\mathbf{n}-1-c}{\mathbf{n}}}{\binom{2\mathbf{n}-2}{\mathbf{n}-1}} < 1.$$

This ratio is

$$\frac{(\mathbf{n}-1)\cdots(\mathbf{n}-c)}{(2\mathbf{n}-2)(2\mathbf{n}-3)\cdots(2\mathbf{n}-c)} = \frac{\mathbf{n}-c}{2^{c-1}}\frac{\mathbf{n}-1}{\mathbf{n}-\frac{2}{2}}\frac{\mathbf{n}-2}{\mathbf{n}-\frac{3}{2}}\frac{\mathbf{n}-3}{\mathbf{n}-\frac{4}{2}}\cdots\frac{\mathbf{n}-c+1}{\mathbf{n}-\frac{c}{2}}.$$

so if $c - 1 \geq \log_2(\mathbf{n})$ it is less than one.

For the rest of the proof, we first introduce notation: for a Young diagram λ, we picture it Russian style, as we think of it as representing entries in the south-west corner of an $\mathbf{n} \times \mathbf{n}$ matrix. More precisely for $(i,j) \in \lambda$ we number the boxes of λ by pairs (row,column) however we number the rows starting from $\mathbf{n}$, i.e. $i = \mathbf{n}$ is the first row. For example

$$\begin{array}{|c|c|}\hline x & y \\ \hline z & \multicolumn{1}{c}{} \\ \cline{1-1} w & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}$$

is labeled $x = (\mathbf{n}, 1), y = (\mathbf{n}, 2), z = (\mathbf{n}-1, 1), w = (\mathbf{n}-2, 1)$. Let $U_\lambda := \operatorname{span}\{u^i \otimes v_j \mid (i,j) \in \lambda\}$ and write $M^\lambda_{\langle \mathbf{n},\mathbf{n},\mathbf{w}\rangle} := M_{\langle \mathbf{n},\mathbf{n},\mathbf{w}\rangle}/U_\lambda$.

The proof consists of two parts. In the first, we prove by induction on k that for any $k < \mathbf{n}$ there exists a Young diagram λ with k boxes such that $\underline{\mathbf{R}}(M^\lambda_{\langle \mathbf{n},\mathbf{n},\mathbf{w}\rangle}) \leq \underline{\mathbf{R}}(M_{\langle \mathbf{n},\mathbf{n},\mathbf{w}\rangle}) - k$.

In the second part we estimate $\underline{\mathbf{R}}(M^\lambda_{\langle \mathbf{n},\mathbf{n},\mathbf{w}\rangle})$ for any λ by reducing to the case when λ has just one row (or column).

Part 1) First step: $k = 1$. By Proposition 1.1 there exists $a \in B_{\underline{\mathbf{R}}(M_{\langle \mathbf{n},\mathbf{n},\mathbf{w}\rangle})-1}(M_{\langle \mathbf{n},\mathbf{n},\mathbf{w}\rangle})$ such that the reduced tensor drops border rank. The group $\mathrm{GL}(U) \times \mathrm{GL}(V) \times \mathrm{GL}(W)$ stabilizes $M_{\langle \mathbf{n},\mathbf{n},\mathbf{w}\rangle}$. By Lemma 2.2 with $G_1 = \mathrm{GL}(U) \times \mathrm{GL}(V)$, we may act on a and pass to the limit. For example, we may first reduce the rank of a to 1, e.g., if a has a nonzero entry in the first row, by multiplying it on the left by the diagonal matrix with entries $(1, \epsilon, \cdots, \epsilon)$ and then letting ϵ go to zero, and then make it equal $u^{\mathbf{n}} \otimes v_1$ with an element of G_1.

Second step: We assume that $\underline{\mathbf{R}}(M^{\lambda'}_{\langle \mathbf{n},\mathbf{n},\mathbf{w}\rangle}) \leq \underline{\mathbf{R}}(M_{\langle \mathbf{n},\mathbf{n},\mathbf{w}\rangle}) - k + 1$, where λ' has $k-1$ parts. Again by Proposition 1.1 there exists $a \in B_{\underline{\mathbf{R}}(M_{\langle \mathbf{n},\mathbf{n},\mathbf{w}\rangle})-k}(M^{\lambda'}_{\langle \mathbf{n},\mathbf{n},\mathbf{w}\rangle})$ such that when we reduce by it the border rank drops. We no longer have the full action of $\mathrm{GL}(U) \times \mathrm{GL}(V)$. However, the product of Borel groups that stabilize the flags induced by λ' stabilizes $M^{\lambda'}_{\langle \mathbf{n},\mathbf{n},\mathbf{w}\rangle}$. By the torus action and Lemma 2.2 we may assume that a has just one nonzero entry outside of λ. Further, using the Borel action we can move the entry south-west to obtain the desired Young diagram λ.

Part 2) We use (2) and recall from [8] that for the matrix multiplication operator, the Koszul flattening factors as $M_{\langle \mathbf{n},\mathbf{n},\mathbf{w}\rangle} = M_{\langle \mathbf{n},\mathbf{n},1\rangle}\otimes \mathrm{Id}_W$, and the rank of $(M_{\langle \mathbf{n},\mathbf{n},\mathbf{w}\rangle})^{\wedge k}_A$ is the rank of $(M_{\langle \mathbf{n},\mathbf{n},1\rangle})^{\wedge k}_A$ times $\mathbf{w}$, and this continues to hold when we restrict the evaluation to subspaces $A' \subset A^*$. We apply the Koszul flattening to $M_{\langle \mathbf{n},\mathbf{n},1\rangle} \in (U^*\otimes V)\otimes V^*\otimes U$, where $\mathbf{u} = \mathbf{v} = \mathbf{n}$. We need to show that for all λ of size m,

$$\underline{\mathbf{R}}(M^{\lambda}_{\langle \mathbf{n},\mathbf{n},1\rangle}) \geq 2\mathbf{n} - 1 - \frac{\binom{\mathbf{n}-1+m}{m-1}}{\binom{2\mathbf{n}-2}{\mathbf{n}-1}}.$$

We will accomplish this by projecting to a suitable subspace $\tilde{A}$ of dimension $2\mathbf{n} - 1$ via the projection map $p_{\tilde{A}} : A \to \tilde{A}$, such that

$$\mathrm{rank}([p_{\tilde{A}}(M^{\lambda}_{\langle \mathbf{n},\mathbf{n},1\rangle}))]^{\wedge \mathbf{n}-1}_{\tilde{A}} \geq \binom{2\mathbf{n}-1}{\mathbf{n}-1}\mathbf{n} - \binom{\mathbf{n}-1+m}{m-1},$$

and then apply (4). By our choice of basis we may consider $M^{\lambda}_{\langle \mathbf{n},\mathbf{n},1\rangle} \in (A/U_\lambda)\otimes B\otimes C$ in $A\otimes B\otimes C$, with specific coordinates equal to 0. We need to show

$$\dim\ker([p_{\tilde{A}}(M_{\langle \mathbf{n},\mathbf{n},1\rangle})^{\lambda}]^{\wedge \mathbf{n}-1}_{\tilde{A}}) \leq \binom{\mathbf{n}-1+m}{m-1}.$$

Consider the map $\phi : A \to \mathbb{C}^{2\mathbf{n}-1}$ given by $u^i\otimes v_j \mapsto e_{i+j-1}$. The rank of the reduced Young flattening $\Lambda^{n-1}\mathbb{C}^{2\mathbf{n}-1}\otimes V \to \Lambda^n\mathbb{C}^{2\mathbf{n}-1}\otimes U$ could *a priori* go down. However, for $M_{\langle \mathbf{n},\mathbf{n},1\rangle}$, as was shown in [8, 7], the new map is surjective. We recall the argument from [7], as a similar argument will finish the proof.

Write $e_S = e_{s_1} \wedge \cdots \wedge e_{s_{\mathbf{n}-1}}$, where $S \subset [2\mathbf{n} - 1]$ has cardinality $\mathbf{n} - 1$. For $1 \leq \eta \leq \mathbf{n}$ the reduced Koszul flattening is given by:

$$e_S\otimes v_\eta \mapsto \sum_{j=1}^{\mathbf{n}} \phi(u^j\otimes v_\eta) \wedge e_S\otimes u_j = \sum_{j=1}^{\mathbf{n}} e_{j+\eta-1} \wedge e_S\otimes u_j.$$

We index a basis of the source by pairs (S, k), with $k \in [\mathbf{n}]$, and the target by (P, l) where $P \subset [2\mathbf{n}-1]$ has cardinality $\mathbf{n}$ and $l \in [\mathbf{n}]$. Define an order on the target basis vectors as follows: For (P_1, l_1) and (P_2, l_2), set $l = \min\{l_1, l_2\}$, and declare $(P_1, l_1) < (P_2, l_2)$ if and only if

(1) In lexicographic order, the set of l minimal elements of P_1 is strictly after the set of l minimal elements of P_2 (i.e. the smallest element of P_2 is smaller than the smallest of P_1 or they are equal and the second smallest of P_2 is smaller or equal etc. up to l-th), or
(2) the l minimal elements in P_1 and P_2 are the same, and $l_1 < l_2$.
(3) the l minimal elements in P_1 and P_2 are the same, and $l_1 = l_2$, and the set of $\mathbf{n} - l$ tail elements of P_1 are after the set of $\mathbf{n} - l$ tail elements of P_2.

In [7] we showed that when one orders the basis as above, the reduced Koszul flattening for $M_{\langle \mathbf{n}\rangle}$ has an upper triangular structure. More explicitly, let $P = (p_1, \cdots, p_{\mathbf{n}})$ with $p_i < p_{i+1}$. Identifying basis vectors with their indices, the image of $(P\backslash\{p_l\}, 1 + p_l - l)$ is $\pm(P, l)$ plus smaller terms in the order. The crucial part is to control how the projection of $M^{\lambda}_{\langle \mathbf{n},\mathbf{n},\mathbf{w}\rangle}$ to the complement of $u^j\otimes v_{\mathbf{n}+1-i}$ effects the reduced Koszul flattening. We determine the number of additional zeros on the diagonal. Note that (P, l) will not appear as the leading term in the reduced map if and only if $l = j$ and $\mathbf{n} + 1 - i + j - 1 = p_l$. Hence, the number of additional zeros on the diagonal equals the number of $\mathbf{n}$ element subsets of $[2\mathbf{n} - 1]$ that have the j-th entry equal to $\mathbf{n} - i + j$, which is $\binom{\mathbf{n}-i+j-1}{j-1}\binom{\mathbf{n}+i-j-1}{i-1} := g(i, j)$. So it is enough to prove that $\sum_{(i,j)\in\lambda} g(i, j) \leq \binom{\mathbf{n}-1+m}{m-1}$. Note that $\sum_{i=1}^{m} g(i, 1) = \sum_{j=1}^{m} g(1, j) = \binom{\mathbf{n}-1+m}{m-1}$. Thus we have to prove that the Young diagram that maximizes $f_\lambda := \sum_{(i,j)\in\lambda} g(i,j)$ has one row or column. We prove it inductively on the size of λ, the case $|\lambda| = 1$ being trivial.

Suppose now that $\lambda = \lambda' + (i,j)$. By induction it is sufficient to show that:

$$g(1,ij) = \binom{\mathbf{n}-1+ij-1}{ij-1} \geq \binom{\mathbf{n}-j+i-1}{i-1}\binom{\mathbf{n}-i+j-1}{j-1} = g(i,j), \tag{5}$$

where $\mathbf{n} > ij$. Without loss of generality we may assume $2 \leq i \leq j$. For $j = 2,3$ the inequality is straightforward to check, so we assume $j \geq 4$. We prove the inequality (5) by induction on $\mathbf{n}$. For $\mathbf{n} = ij$ the inequality follows from the combinatorial interpretation of binomial coefficients and the fact that the middle one is the largest.

We have $\binom{\mathbf{n}+1-1+ij-1}{ij-1} = \binom{\mathbf{n}-1+ij-1}{ij-1}\frac{\mathbf{n}-1+ij}{\mathbf{n}}$, $\binom{\mathbf{n}+1-j+i-1}{i-1} = \binom{\mathbf{n}-j+i-1}{i-1}\frac{\mathbf{n}-j+i}{n-j+1}$ and $\binom{\mathbf{n}+1-i+j-1}{j-1} = \binom{\mathbf{n}-i+j-1}{j-1}\frac{\mathbf{n}-i+j}{\mathbf{n}-i+1}$. By induction it is enough to prove that:

$$\frac{\mathbf{n}-1+ij}{n} \geq \frac{\mathbf{n}-j+i}{\mathbf{n}-j+1}\frac{\mathbf{n}-i+j}{\mathbf{n}-i+1}. \tag{6}$$

This is equivalent to:

$$ij - 1 \geq \frac{\mathbf{n}(i-1)}{\mathbf{n}-j+1} + \frac{\mathbf{n}(j-1)}{\mathbf{n}-i+1} + \frac{\mathbf{n}(i-1)(j-1)}{(\mathbf{n}-j+1)(\mathbf{n}-i+1)}.$$

As the left hand side is independent from $\mathbf{n}$ and each fraction on the right hand side decreases with growing $\mathbf{n}$ (differentiate each term with respect to $\mathbf{n}$ and use that $\mathbf{n} > ij$ in the last case to see all derivatives are negative), we may set $\mathbf{n} = ij$ in inequality 6. Thus it is enough to prove:

$$2 - \frac{1}{ij} \geq (1 + \frac{i-1}{ij-j+1})(1 + \frac{j-1}{ij-i+1}).$$

Then the inequality is straightforward to check for $i = 2$, so we assume $i \geq 3$. Then:

$$(1 + \frac{i-1}{ij-j+1})(1 + \frac{j-1}{ij-i+1}) \leq (1 + \frac{j-1}{j^2-j+1})(1 + \frac{j-1}{3j-2}) \leq \frac{16}{13}\cdot\frac{4}{3} = \frac{64}{39}.$$

However,

$$\frac{64}{39} \leq 2 - \frac{1}{12} \leq 2 - \frac{1}{3j} \leq 2 - \frac{1}{ij},$$

which finishes the proof.

Remark 3.1. Note that we made two kinds of restrictions:

(1) projecting A to A/U_λ and
(2) projecting A/U_λ to $\tilde{A}$.

The first one corresponds to deleting rows (specified by λ) in the matrix representation of $M_{\langle \mathbf{n},\mathbf{n},1\rangle}$. The second one takes $2\mathbf{n}-1$ linear combinations of rows as explained below.

Since linear projections commute, one might try to first apply the second projection and then the first one. This is not feasible for two reasons. First, after applying the second projection we lose symmetry. Second, our method removes whole rows in the matrix representation of the tensor in the first projection (not just specific entries). Hence it is much better to first remove rows (when the matrix has mostly zeros) and then use the second projection, than to remove rows when the matrix is dense (after the second projection).

4. Compression of tensors: limits of the border substitution method

Consider the product of Grassmannians $\mathbf{G} := G(\mathbf{a}', A^*) \times G(\mathbf{b}', B^*) \times G(\mathbf{c}', C^*)$ with three projections π_i. Let $\mathcal{E} = \mathcal{E}(\mathbf{a}', \mathbf{b}', \mathbf{c}') := \otimes_{i=1}^3 \pi_i^*(\mathcal{S}_i)$ be the vector bundle that is the tensor product of the pullbacks of universal subspace bundles $\mathcal{S}_i$. Let $\mathbf{P} \to \mathbf{G}$ denote the projective bundle with fiber over (A', B', C') equal to $Seg(\mathbb{P}A' \times \mathbb{P}B' \times \mathbb{P}C')$, so $\mathbf{P} \subset \mathbb{P}\mathcal{E}$.

Definition 4.1. A tensor $T \in A \otimes B \otimes C$ is $(\mathbf{a}', \mathbf{b}', \mathbf{c}')$-*compressible* if there are subspaces $A' \subset A^*, B' \subset B^*, C' \subset C^*$ of respective dimensions $\mathbf{a}', \mathbf{b}', \mathbf{c}'$ such that $T|_{A' \otimes B' \otimes C'} = 0$, i.e., there exists $(A', B', C') \in \mathbf{G}$, such that $A' \otimes B' \otimes C' \subset T^\perp$, where $T^\perp \subset (A \otimes B \otimes C)^*$ is the hyperplane annihilating T. If T is not $(\mathbf{a}', \mathbf{b}', \mathbf{c}')$-compressible, we say it is $(\mathbf{a}', \mathbf{b}', \mathbf{c}')$-*compression generic (cg)*. Let $X(\mathbf{a}', \mathbf{b}', \mathbf{c}')$ denote the set of all tensors that are $(\mathbf{a}', \mathbf{b}', \mathbf{c}')$-compressible.

Proof of Proposition 1.3. Let

$$Y := \{(y, [T]) \in \mathbf{G} \times \mathbb{P}(A \otimes B \otimes C) \mid \mathcal{E}_y \subset T^\perp\}.$$

Each fiber of the projection $Y \to \mathbf{G}$ is a projective space of dimension $\mathbf{abc} - \mathbf{a}'\mathbf{b}'\mathbf{c}' - 1$, so

$$\dim Y := (\mathbf{abc} - \mathbf{a}'\mathbf{b}'\mathbf{c}' - 1) + (\mathbf{a} - \mathbf{a}')\mathbf{a}' + (\mathbf{b} - \mathbf{b}')\mathbf{b}' + (\mathbf{c} - \mathbf{c}')\mathbf{c}'.$$

On the other hand $X(\mathbf{a}', \mathbf{b}', \mathbf{c}')$ is the generically one to one projection of Y to $\mathbb{P}(A \otimes B \otimes C)$, which proves both claims since Y is Zariski closed. □

Recall the inequality

$$\mathbf{a}\mathbf{a}' + \mathbf{b}\mathbf{b}' + \mathbf{c}\mathbf{c}' < (\mathbf{a}')^2 + (\mathbf{b}')^2 + (\mathbf{c}')^2 + \mathbf{a}'\mathbf{b}'\mathbf{c}' \tag{7}$$

from Proposition 1.3.

Corollary 4.2.

i *If (7) holds then a generic tensor is* $(\mathbf{a}', \mathbf{b}', \mathbf{c}')$*-cg.*

ii *If (7) does not hold then* $\operatorname{rank} \mathcal{E}^* \leq \dim G(\mathbf{a}', A^*) \times G(\mathbf{b}', B^*) \times G(\mathbf{c}', C^*)$. *If the top Chern class of* $\mathcal{E}^*$ *is nonzero, then no tensor is* $(\mathbf{a}', \mathbf{b}', \mathbf{c}')$*-cg.*

Proof. The first assertion is a restatement of Proposition 1.3.

For the second, notice that T induces a section $\tilde{T}$ of the vector bundle $\mathcal{E}^* \to \mathbf{G}$. The zero locus of $\tilde{T}$ is $\{(A', B', C') \in \mathbf{G} \mid A' \otimes B' \otimes C' \subset T^\perp\}$. In particular, $\tilde{T}$ is non-vanishing if and only if T is $(\mathbf{a}', \mathbf{b}', \mathbf{c}')$-cg. If the top Chern class is nonzero, there cannot exist a non-vanishing section. □

Example 4.3. Let $\mathbf{a} = \mathbf{b} = \mathbf{c}$ and $\mathbf{a}' = \mathbf{b}' = \mathbf{c}'$. Then the border substitution method can be applied as long as

$$\mathbf{a}' \geq \lceil \sqrt{3\mathbf{a} + \frac{9}{4}} - \frac{3}{2} \rceil.$$

Thus by this method alone, one potentially gets border rank equations in $\mathbb{C}^{\mathbf{a}} \otimes \mathbb{C}^{\mathbf{a}} \otimes \mathbb{C}^{\mathbf{a}}$ up to

$$3(\mathbf{a} - \lceil (\sqrt{3\mathbf{a} + \frac{9}{4}} - \frac{3}{2}) \rceil).$$

For example, if $\mathbf{a} = 9$, we may take $\mathbf{a}' = 4$ and get equations up to σ_{15}.

Example 4.4. Let $\mathbf{a} = \mathbf{b} = \mathbf{c} = 3$. As pointed out by J. Kileel, the variety $X(2,2,3)$ equals the trifocal variety. By the results of C. Aholt and L. Oeding [1] the ideal of this variety is defined by 10 cubics, 81 quintics and 1980 sextics.

In each particular case when there are a finite number of $A'\otimes B'\otimes C'$ annhilating a generic T, we may explicitly compute how many different $A'\otimes B'\otimes C'$ a generic hyperplane may contain as follows: The Chern polynomial of the dual of the universal bundle is $\sum_{j=0}^{k} p_{1^j} t^j$, where p_{1^j} is the class corresponding to the Young diagram 1^j. These classes multiply by the Littlewood-Richardson rule (in our cases this is the iterated Pieri rule).

Example 4.5. Let $\mathbf{a} = \mathbf{b} = \mathbf{c} = 5$ and $\mathbf{a}' = 2, \mathbf{b}' = 1, \mathbf{c}' = 5$. The bundle $\mathcal{E}^*$ has rank ten: it is the tensor product of a rank 2 bundle (for $\mathbf{a}'$), rank 1 bundle (for $\mathbf{b}'$) and the trivial rank 5 bundle (for $\mathbf{c}'$). This example already appeared in [4]. Here $\mathbf{G} = G(2,5) \times \mathbb{P}^5$ as the last Grassmannian degenerates to a point. The second Chern class of the tensor product of pull-backs equals:

$$c_2(\pi_1^*(\mathcal{S}_1)\otimes\pi_2^*(\mathcal{S}_2)) = (\ydiagram{1,1}, 1) + (\ydiagram{1}, \ydiagram{1}) + (1, \ydiagram{1})^2,$$

where respective Young diagrams represent Schubert classes on $G(2,5)$ and $\mathbb{P}^5$. E.g. $(1, \ydiagram{1})$ is $G(2,5)$ times a hyperplane in $\mathbb{P}^5$. To compute the top Chern class of $\mathcal{E}^*$ we need to compute the 5-th power of the above expression. It will be proportional to the class of a point $(\ydiagram{3,3}, \ydiagram{5})$ and we just have to compute the coefficient.

We obtain the following contributions:

- $5(\ydiagram{1,1}, 1)(\ydiagram{1}, \ydiagram{1})^4 = 5 \cdot 2 = 10$. Indeed, on the second coordinate corresponding to $\mathbb{P}^5$ we just have to fill, one by one starting from left, the diagram $\ydiagram{4}$. On $G(2,5)$ we must start by filling the two left most entries, by the contribution of $(\ydiagram{1,1}, 1)$ obtaining: $\young(xoo,xoo)$. The remaining square (filled with o before) has to be filled with four unit squares. There are two ways to do this: $\young(12,34)$ and $\young(13,24)$.
- $5\binom{4}{2}(\ydiagram{1,1}, 1)^2(\ydiagram{1}, \ydiagram{1})^2(1, \ydiagram{1})^2 = 30$, because there is a unique filling here,
- $\binom{5}{2}$ corresponding to $(\ydiagram{1,1}, 1)^3(1, \ydiagram{1})^4$.

This gives the grand total of 50. Hence, in this case the map $Y \to \mathbf{P}(A\otimes B\otimes C)$ is surjective, finite with generic fiber of degree 50.

References

1. Chris Aholt and Luke Oeding, *The ideal of the trifocal variety*, Math. Comp. **83** (2014), no. 289, 2553–2574. MR 3223346
2. Boris Alexeev, Michael A. Forbes, and Jacob Tsimerman, *Tensor rank: some lower and upper bounds*, 26th Annual IEEE Conference on Computational Complexity, IEEE Computer Soc., Los Alamitos, CA, 2011, pp. 283–291. MR 3025382
3. M. Bläser and V. Lysikov, *On degeneration of tensors and algebras*, ArXiv e-prints (2016).
4. Shinobu Hosono and Hiromichi Takagi, *Mirror symmetry and projective geometry of Reye congruences I*, J. Algebraic Geom. **23** (2014), no. 2, 279–312. MR 3166392
5. J. M. Landsberg, *The border rank of the multiplication of* 2×2 *matrices is seven*, J. Amer. Math. Soc. **19** (2006), no. 2, 447–459. MR 2188132 (2006j:68034)
6. J. M. Landsberg and M. Michalek, *Abelian Tensors*, ArXiv e-prints, to appear in JMPA (2015).

7. ______, *On the geometry of border rank algorithms for matrix multiplication and other tensors with symmetry*, ArXiv e-prints, to appear in SIAGA (2016).
8. Joseph M. Landsberg and Giorgio Ottaviani, *New lower bounds for the border rank of matrix multiplication*, Theory Comput. **11** (2015), 285–298. MR 3376667
9. Thomas Lickteig, *A note on border rank*, Inform. Process. Lett. **18** (1984), no. 3, 173–178. MR 86c:68040
10. V. Strassen, *Rank and optimal computation of generic tensors*, Linear Algebra Appl. **52/53** (1983), 645–685. MR 85b:15039
11. Volker Strassen, *Gaussian elimination is not optimal*, Numer. Math. **13** (1969), 354–356. MR 40 #2223

Department of Mathematics, Texas A&M University, Mailstop 3368, College Station, TX 77843-3368, USA
E-mail address: `jml@math.tamu.edu`

Polish Academy of Sciences, ul. Śniadeckich 8, 00-956 Warsaw, Poland
E-mail address: `wajcha2@poczta.onet.pl`