

# Tensor Rank: Some Lower and Upper Bounds

by

Michael Andrew Forbes

S.B., Massachusetts Institute of Technology (2009)

Submitted to the Department of Electrical Engineering and Computer  
Science

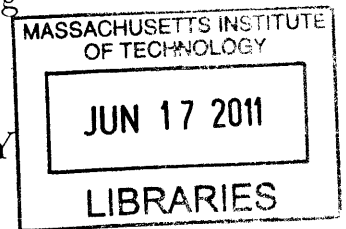
in partial fulfillment of the requirements for the degree of  
Master of Science in Computer Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2011

**ARCHIVES**



© Massachusetts Institute of Technology 2011. All rights reserved.

Author ... ..  
Department of Electrical Engineering and Computer Science  
May 20, 2011

Certified by.. ..  
Scott Aaronson  
Associate Professor  
Thesis Supervisor

Accepted by .. ..  
| U Leslie A. Kolodziejski  
Chair, Department Committee on Graduate Students



# Tensor Rank: Some Lower and Upper Bounds

by

Michael Andrew Forbes

Submitted to the Department of Electrical Engineering and Computer Science  
on May 20, 2011, in partial fulfillment of the  
requirements for the degree of  
Master of Science in Computer Science and Engineering

## Abstract

The results of Strassen [25] and Raz [19] show that good enough *tensor rank* lower bounds have implications for algebraic circuit/formula lower bounds.

We explore tensor rank lower and upper bounds, focusing on explicit tensors. For odd  $d$ , we construct field-independent explicit 0/1 tensors  $T : [n]^d \rightarrow \mathbb{F}$  with rank at least  $2n^{\lfloor d/2 \rfloor} + n - \Theta(d \lg n)$ . This improves the lower-order terms in known lower bounds for any odd  $d \geq 3$ .

We also explore a generalization of permutation matrices, which we denote permutation tensors. We show, by applying known counting lower bounds, that there exist order-3 permutation tensors with super-linear rank as well as order- $d$  permutation tensors with high rank. We also explore a natural class of permutation tensors, which we call group tensors. For any group  $G$ , we define the group tensor  $T_G^d : G^d \rightarrow \mathbb{F}$ , by  $T_G^d(g_1, \dots, g_d) = 1$  iff  $g_1 \cdots g_d = 1_G$ . We give two upper bounds for the rank of these tensors. The first uses representation theory and works over “large” fields  $\mathbb{F}$ , showing (among other things) that  $\text{rank}_{\mathbb{F}}(T_G^d) \leq |G|^{d/2}$ . In the case that  $d = 3$ , we are able to show that  $\text{rank}_{\mathbb{F}}(T_G^3) \leq O(|G|^{\omega/2}) \leq O(|G|^{1.19})$ , where  $\omega$  is the exponent of matrix multiplication. The next upper bound uses interpolation and only works for abelian  $G$ , showing that over any field  $\mathbb{F}$  that  $\text{rank}_{\mathbb{F}}(T_G^d) \leq O(|G|^{1+\lg d} \lg^{d-1} |G|)$ . In either case, this shows that many permutation tensors have far from maximal rank, which is very different from the matrix case and thus eliminates many natural candidates for high tensor rank.

We also explore monotone tensor rank. We give explicit 0/1 tensors  $T : [n]^d \rightarrow \mathbb{F}$  that have tensor rank at most  $dn$  but have monotone tensor rank exactly  $n^{d-1}$ . This is a nearly optimal separation.

Thesis Supervisor: Scott Aaronson

Title: Associate Professor



# Acknowledgments

This work was made possible through the support of many people. I would like to thank my advisor, Scott Aaronson, for much of the advice, as well as independence, that he gave me in exploring this thesis. My father, Kevin Forbes, has been a role model for me throughout my life, and undoubtedly much of my research career is due to his guidance. My mother, Lorraine Daly, has always been supportive of whatever I do, and has thus made many of my daily challenges easier. Sunny Lou, my significant other, has made life so much more beautiful through her sunshine.

Many thanks to the loving care of my longtime babysitter Asia Martinis, as well for the warmth of her (now-deceased) husband, Paul Martinis.

I would also like to thank the family and friends who have supported me in the past two (or more) years: Doug Adams, John Chai, Kendra Leigh, Randy Li, Joann Lin, as well as Mary Jo Braid-Forbes and Thomas Forbes.

On a more academic level, I would primarily like to thank Boris Aleexev and Jacob Tsimmerman for their collaboration in the paper [1, 2] upon which this thesis is based. I would also like to thank Swastik Kopparty for point to the standard construction presented in Proposition 3.2.3, Guy Moshkovitz for pointing out the relation of permutation tensors and quasigroups, Amir Shpilka for point out the improvement in Theorem 4.1.7 for  $d = 3$ , and Madhu Sudan for pointing us to the existence of Lemma 3.2.6. I would also like to thank Arnab Bhattacharyya, Andy Drucker, Zeev Dvir, Kevin Hughes, Neeraj Kayal, Satya Lokam, Jakob Nordstrom and Pavel Pudlák, as well as anonymous reviewers, for various constructive comments.

This work was supported by NSF grant 6919791, MIT CSAIL and a fellowship from the Siebel Foundation.



# Contents

<b>1</b>	<b>Introduction</b>	<b>9</b>
1.1	Motivation . . . . .	9
1.2	Prior Work . . . . .	11
1.3	Our Results . . . . .	13
<b>2</b>	<b>Structural Results</b>	<b>15</b>
2.1	Definitions and Notation . . . . .	15
2.2	Basic Facts about Tensors . . . . .	16
2.3	Layer Reduction . . . . .	22
<b>3</b>	<b>Tensor Rank Lower Bounds</b>	<b>27</b>
3.1	Combinatorially-defined Tensors . . . . .	27
3.2	Algebraically-defined Tensors . . . . .	32
3.3	Higher-Order Tensors . . . . .	37
<b>4</b>	<b>Tensor Rank Upper Bounds</b>	<b>41</b>
4.1	Permutation Tensors . . . . .	41
4.1.1	Permutation Tensors: Rank Lower Bounds . . . . .	42
4.1.2	Permutation Tensors: Rank Upper Bounds . . . . .	44
<b>5</b>	<b>Monotone Tensor Rank</b>	<b>55</b>





# Chapter 1

## Introduction

### 1.1 Motivation

Most real-world computing treats data as boolean, and thus made of bits. However, for some computational problems this viewpoint does not align with algorithm design. For example, the determinant is a polynomial, and computing it typically does not require knowledge of the underlying bit representation and rather treats the inputs as numbers in some field. In such settings, it is natural to consider the computation of the determinant as computing a polynomial over the underlying field, as opposed to computing a boolean function.

When computing polynomials, just as when computing boolean functions, there are many different models of computation to choose. The most general is the *algebraic circuit* model. Specifically, to compute a polynomial  $f$  over a field  $\mathbb{F}$  in variables  $x_1, \dots, x_n$ , one defines a directed acyclic graph, with exactly  $n$  source nodes (each labeled with a distinct variable), a single sink node (which is thought of as the output), and internal nodes labeled with either  $+$ , meaning addition, or  $\times$ , meaning multiplication. Further, each non-leaf is restricted to have at most two children nodes. Computation is defined in the natural way: each non-source node computes the (polynomial) function of its children according to its label, and source nodes compute the variable they are labeled with. One can also consider the *algebraic formula* model, which requires the underlying graph to be a tree. In both of these models, we define

the *size* of the circuit/formula to be the total number of nodes in the graph (other natural measures of size are polynomially related).

Neither the algebraic circuit nor the formula model are well understood, in the sense that while it can be shown that there exist polynomials which require large circuits for their computation, no explicit<sup>1</sup> examples of such polynomials are known. Indeed, finding such lower bounds for explicit functions is considered one of the most difficult problems in computational complexity theory. Several lower bounds are known, such as Strassen's [26] result (using the result of Baur-Strassen [4]) that the degree  $n$  polynomial  $\sum_{i=1}^n x_i^n$  requires  $\Omega(n \lg n)$  size circuits. However, no super-linear size lower bounds are known for constant-degree polynomials. In the case of formulas, Kalorkoti [16] proved a quadratic-size lower bound for an explicit function.

One approach for improvements on known lower bounds in both of these models is by proving lower bounds for *tensor rank*. A tensor is a generalization of a matrix, and an order- $d$  tensor is defined as a function  $T : [n]^d \rightarrow \mathbb{F}$ , where  $[n]$  denotes the set  $\{1, \dots, n\}$ . A tensor is *rank one* if it can be factorized as  $T(i_1, \dots, i_d) = \prod_{j=1}^d \vec{v}_j(i_j)$  for  $\vec{v}_j \in \mathbb{F}^n \setminus \{\vec{0}\}$ , where  $\vec{v}_j(i_j)$  is the  $i_j$ -th entry of the vector  $\vec{v}_j$ . The rank of a tensor is the minimum  $r$  such that  $T = \sum_{k=1}^r S_k$  for rank one tensors  $S_k$ . It can be seen that an order-2 tensor is a matrix, and the notions of rank coincide. It can be observed that the rank of an  $[n]^d$  tensor is always at most  $n^{d-1}$ , and a counting-type argument shows that over any field there exist tensors of rank at least  $n^{d-1}/d$ . A tensor is called *explicit* if  $T(i_1, \dots, i_d)$  can be computed by algebraic circuits of size at most polynomial in  $\text{poly}(d \lg n)$ , that is, at most polynomial in the size of the input  $(i_1, \dots, i_d)$ . All explicit tensors in this paper will also be uniformly explicit.

Interest in tensors arise from their natural correspondence with certain polynomials (a class known as set-multilinear polynomials). Consider the sets of variables  $\{X_{i,j}\}_{i \in [n], j \in [d]}$ . Given a tensor  $T : [n]^d \rightarrow \mathbb{F}$ , one can define the polynomial

$$f_T(\{X_{i,j}\}_{i \in [n], j \in [d]}) = \sum_{i_1, \dots, i_d \in [n]} T(i_1, \dots, i_d) \prod_{j=1}^d X_{i_j, j}$$

---

<sup>1</sup>A polynomial is said to be *explicit* if the coefficient of a monomial  $\vec{X}^{\vec{\alpha}}$  is computable by algebraic circuits of size at most  $\text{poly}(|\vec{\alpha}|)$ .

This connection was used in the following two results. First, Strassen [25] showed that

**Theorem** (Strassen [25], see also [31]). *For a tensor  $T : [n]^3 \rightarrow \mathbb{F}$ , the circuit size complexity of  $f_T$  is  $\Omega(\text{rank}(T))$ .*

Thus, any super-linear lower-bound for order-3 tensor rank gives a super-linear lower bound for general arithmetic circuits, even for the constant degree polynomials. More recently, Raz [19] proved

**Theorem** (Raz [19]). *For a family of tensors  $T_n : [n]^{d(n)} \rightarrow \mathbb{F}$  with  $\text{rank}(T_n) \geq n^{(1-o(1))d(n)}$  and  $\omega(1) \leq d(n) \leq O(\log n / \log \log n)$ , the formula-size complexity of  $f_{T_n}$  is super-polynomial.*

Thus, while Strassen's result cannot be used to prove super-quadratic circuit-size lower bounds (because of the upper bounds on order-3 tensor rank), Raz's result shows that tensor rank could be used to prove very strong lower-bounds. These results motivate a study of tensor rank as a model of computation, which this work pursues.

## 1.2 Prior Work

Strassen's connection between order-3 tensor rank and circuit complexity further established a close connection between tensor rank to what is known as *bilinear complexity*. As several important problems, such as matrix multiplication and polynomial multiplication, are bilinear, one can study their bilinear complexity, and thus their order-3 tensor rank. We interpret various prior results in the language of tensor rank. For the matrix multiplication (which corresponds to a tensor of size  $[n^2] \times [n^2] \times [n^2]$ ), Shpilka [23] showed that the tensor rank is at least  $3n^2 - O(n^{5/3})$  over  $\mathbb{F}_2$ , and Bläser [5] earlier showed that over any field the tensor rank is at least  $2.5n^2 - \Theta(n)$ . For polynomial multiplication (which corresponds to a tensor of size  $[2n-1] \times [n] \times [n]$ ), Kaminski [17] showed that the tensor rank over  $\mathbb{F}_q$  is known to be  $(3 + 1/\Theta(q^3))n - o(n)$  and earlier work by Brown and Dobkin [8] showed that over  $\mathbb{F}_2$  the tensor rank is at

least  $3.52n$ . Lower bounds for these problems seem difficult, in part because strong upper bounds exist for both matrix multiplication and polynomial multiplication.

This work attempts to prove tensor rank lower bounds for *any* explicit function, not just problems of prior interest such as matrix or polynomial multiplication. Previous work in this realm include that of Ja'Ja' [15] (see their Theorem 3.6), who used the Kronecker theory of pencils to show tensor rank lower bounds of  $1.5n$  for  $[n] \times [n] \times [2]$  tensors, for large fields. The work was later expanded by Sumi, Miyazaki, and Sakata [27] to smaller fields. However, in these works the rank is shown to be at most  $1.5n$ , so seemingly cannot be pushed further. Most notably, Bläser [6] exhibited an explicit algebra whose multiplication tensor is of rank at least  $3n - O(\sqrt{n})$ , which is the best tensor rank lower bound we are aware of prior to this work (in this work we improve the lower order term to  $-\Theta(\log n)$ ).

It is also worth noting that Håstad proved [12, 13] that determining if the tensor rank of  $T : [n]^3 \rightarrow \mathbb{F}$  is at most  $r$  is NP-hard, for  $\mathbb{F}$  finite or the rationals (the problem is also known to be within NP for finite  $\mathbb{F}$ , but not known for the rationals). Implicit in his work is a tensor rank lower bound (for explicit order-3 tensors) of  $4n/3$ . To the best of our knowledge, the hardness of approximating tensor rank is an open question. Part of its difficulty is that any gap-preserving reduction from NP to tensor rank would automatically yield lower bounds for explicit tensors, and few lower bound techniques for tensor rank are known.

It is also a folklore result (eg. see Raz [19], or the proof in Section 3.3) that one can reshape, or embed, a  $n^{\lfloor d/2 \rfloor} \times n^{\lfloor d/2 \rfloor}$  size matrix into a order- $d$  tensor, thus achieving a  $n^{\lfloor d/2 \rfloor}$  rank lower bound for  $[n]^d$  size tensors. Independent of this work, Weitz [32] and Bläser [7] use this idea to transform rank lower bounds for order-3 tensors into rank lower bounds for order- $(2d + 1)$  tensors for any  $d$ . In particular, they prove  $2n^d - O(n^{d-1})$  rank lower bounds for explicit order- $(2d + 1)$  tensors. We also use this idea to achieve  $2n^d + n - \Theta(d \lg n)$  rank lower bounds for explicit order- $(2d + 1)$  tensors.

## 1.3 Our Results

Our work<sup>2</sup> has several components, each studying different aspects of the tensor rank problem. We first give two new methods in proving tensor rank lower bounds. In Section 3.1, we detail the first construction, which proves the best known<sup>3</sup> tensor rank lower bound for a tensor of size  $[n] \times [n] \times [n]$  (over any field). In particular, using a generalization of Gaussian elimination we prove

**Theorem** (Corollary 3.1.10). *Let  $\mathbb{F}$  be an arbitrary field. There are explicit  $\{0, 1\}$ -tensors  $T_n : [n]^3 \rightarrow \mathbb{F}$  such that  $\text{rank}(T_n) = 3n - \Theta(\lg n)$ .*

Our analysis of this construction is in fact exact. In Section 3.2, we give a different order-3 tensor construction with a  $3n - \Theta(\lg n)$  rank lower bound over  $\mathbb{F}_2$  that has no matching upper bound, and leave as an open question what is the exact rank. In Section 3.3, we show how to extend the order-3 tensor rank lower bounds to yield a lower bound (for odd  $d$ ) of  $2n^{\lfloor d/2 \rfloor} + n - \Theta(d \lg n)$  for the tensor rank of an explicit 0/1 size  $[n]^d$  tensor, which improves by a factor of 2 on the folklore reshaping lower bound of  $n^{\lfloor d/2 \rfloor}$ , and improves in the lower-order terms of other works [32, 7] that use reshaping.

In Section 4.1, we explore the tensor rank of permutation tensors. For matrices, permutation matrices are all full-rank and have a tight connection with the determinant. Consequently, it is natural to conjecture that a generalization of permutation matrices, which we call permutation tensors, have high rank. In particular, using known counting lower bounds for Latin squares and  $k$ -ary quasigroups, we show that indeed there are order-3 permutation tensors with super-linear tensor rank and order- $d$  permutation tensors with high rank (over finite fields).

A natural class of permutation tensors are those constructed from groups. That is, for a finite group  $G$  we can define the *group tensor*  $T_G^d : G^d \rightarrow \mathbb{F}$ , which is a 0/1 tensor defined by  $T_G^d(g_1, \dots, g_d) = 1$  iff  $g_1 \cdots g_d = 1_G$ . It seems natural to conjecture

---

<sup>2</sup>Parts of this work have appeared online [1], and will appear in a conference [2].

<sup>3</sup>When comparing this result to those listed in the prior work, it is helpful to note the differences in size of the tensors, such as comparing  $[n^2]^3$  (for matrix multiplication) to our  $[n]^3$ . For example, over  $\mathbb{F}_2$ , we match Shpilka's  $3n^2 - O(n^{5/3})$  lower bound (and improve on low-order terms).

that these tensors might also have high-rank. However, using representation theory we can give a strong upper bound on the rank of any group tensor (over “large” fields such as  $\mathbb{C}$ ). To prove results over any field, we use interpolation methods and field-transfer results to bound the rank of any group tensor arising from an abelian group. In particular, we have the following theorem.

**Theorem** (Theorem 4.1.7, and Corollary 4.1.16). *Let  $G$  be a finite group. For “large” fields  $\mathbb{F}$  (such as algebraically closed fields with characteristic exceeding  $|G|$ ),  $\text{rank}_{\mathbb{F}}(T_G^d) \leq |G|^{d/2}$ . Further, for  $d = 3$  we have that  $\text{rank}_{\mathbb{F}}(T_G^3) \leq O(|G|^{\omega/2}) \leq O(|G|^{1.19})$ , where  $\omega$  is the exponent of matrix multiplication. Moreover, for any field  $\mathbb{F}$ , if  $G$  is abelian then  $\text{rank}_{\mathbb{F}}(T_G^d) \leq O(|G|^{1+\lg d} \lg^{d-1} |G|)$ .*

In each case, we show that group tensors have rank far from the maximal  $\Theta_d(|G|^{d-1})$ , and thus are not good candidates for high tensor rank for large  $d$  (which are needed for Raz’s application) as well as possibly not being good candidates for Strassen’s application (assuming the conjecture that matrix multiplication exponent  $\omega$  is in fact 2). We are unable to place non-trivial upper bounds on the rank of  $T_G^d$  for  $G$  non-abelian and small  $\mathbb{F}$ , but it seems natural to conjecture that strong upper-bounds exist given the above results. Moreover, these results indicate that searching for candidate high-rank tensors should be done among less structured tensors.

Finally, in Chapter 5 we explore *monotone* tensor rank. Monotone computation exploits the idea that if a polynomial only uses positive coefficients (over an ordered field such as  $\mathbb{Q}$ ), then one might try to compute this polynomial only using positive field elements. Previous researchers have tried, in various models, to show that such restricted computation is much more inefficient than unrestricted computation. Indeed, for general algebraic circuits Valiant [28] has shown that allowing negative field elements allows for an exponential improvement in the efficiency of computing certain polynomials. We continue in this line of work. In particular, we can show the following nearly optimal separation.

**Theorem** (Theorem 5.0.3). *Let  $\mathbb{F}$  be any ordered field. There is a explicit 0/1 tensor  $T : [n]^d \rightarrow \mathbb{F}$  such that  $\text{rank}_{\mathbb{F}}(T) \leq dn$ , but the monotone rank of  $T$  is  $n^{d-1}$ .*

# Chapter 2

## Structural Results

### 2.1 Definitions and Notation

We first define tensors, and give some basic facts about them. Throughout this work,  $[n]$  shall denote the set  $\{1, \dots, n\}$ ,  $\llbracket n \rrbracket$  shall denote the set  $\{0, \dots, n-1\}$ , and  $\lg n$  shall denote the logarithm of  $n$  base 2. Further, the notation  $\llbracket E \rrbracket$  (the Iverson bracket) will often be used as an indicator variable for the event  $E$ , and can be distinguished from  $\llbracket n \rrbracket$  by context.

**Definition 2.1.1.** A **tensor** over a field  $\mathbb{F}$  is a function  $T : \prod_{j=1}^d [n_j] \rightarrow \mathbb{F}$ . It is said to have order  $d$  and size  $(n_1, \dots, n_d)$ . If all of the  $n_j$  are equal to  $n$ , then  $T$  is said to have size  $[n]^d$ .  $T$  is said to belong to the **tensor product space**  $\otimes_{j=1}^d \mathbb{F}^{n_j}$ .

In later sections, the input space of a tensor will sometimes be a group or a set  $\llbracket n \rrbracket$  instead of the set  $[n]$ . Throughout this paper  $\mathbb{F}$  shall denote an arbitrary field, the variable  $n$  (or  $(n_1, \dots, n_d)$ ) shall be reserved for the tensor size, and  $d$  shall be reserved for the order.  $\mathbb{F}_q$  will denote the field on  $q$  elements. We can now define the notion of rank for tensors.

**Definition 2.1.2.** A tensor  $T : \prod_{j=1}^d [n_j] \rightarrow \mathbb{F}$  is **simple** if for  $j \in [d]$  there are vectors  $\vec{v}_j \in \mathbb{F}^{n_j}$  such that  $T = \otimes_{j=1}^d \vec{v}_j$ . That is, for all  $i_j \in [n_j]$ ,  $T(i_1, \dots, i_d) = \prod_{j=1}^d \vec{v}_j(i_j)$  where  $\vec{v}_j(i_j)$  denotes the  $i_j$ -th coordinate of  $\vec{v}_j$ .

**Definition 2.1.3.** The **rank** of a tensor  $T : \prod_{j=1}^d [n_j] \rightarrow \mathbb{F}$ , is defined as the minimum number of terms in a summation of simple tensors expressing  $T$ , that is,

$$\text{rank}_{\mathbb{F}}(T) = \min \left\{ r : T = \sum_{k=1}^r \otimes_{j=1}^d \vec{v}_{j,k}, \vec{v}_{j,k} \in \mathbb{F}^{n_j} \right\}$$

Notice that by definition, a non-zero tensor is simple iff it is of rank one.

The next definition shows how identically sized order- $(d-1)$  tensors can be combined into an order- $d$  tensor.

**Definition 2.1.4.** For  $T_1, \dots, T_{n_d} \in \otimes_{j=1}^{d-1} \mathbb{F}^{n_j}$  define  $T = [T_1 | \dots | T_{n_d}]$  by the equation  $T(i_1, \dots, i_{d-1}, i_d) = T_{i_d}(i_1, \dots, i_{d-1})$ . The  $T_{i_j}$  are said to be the **layers** of  $T$  (along the  $d$ -th axis). Layers along other axes are defined analogously.

Conversely, given  $T \in \otimes_{j=1}^d \mathbb{F}^{n_j}$ , define **the  $l$ -th layer of  $T$  (along the  $d$ -axis)**, sometimes denoted  $T_l \in \otimes_{j=1}^{d-1} \mathbb{F}^{n_j}$ , to be the tensor defined by  $T_l(i_1, \dots, i_{d-1}) = T(i_1, \dots, i_{d-1}, l)$ .

## 2.2 Basic Facts about Tensors

We now prove some standard facts about tensors.

**Lemma 2.2.1.** *The  $\otimes_{j=1}^d \mathbb{F}^{n_j}$  tensor product space is an  $\prod_{j=1}^d n_j$  dimensional  $\mathbb{F}$ -vector space, with standard basis  $\{\otimes_{j=1}^d \vec{e}_{i_j,j}\}_{i_j \in [n_j]}$  where  $\{e_{i_j,j}\}_{i_j \in [n_j]}$  is the standard basis for  $\mathbb{F}^{n_j}$ .*

*Thus, the space  $\otimes_{j=1}^d \mathbb{F}^{n_j}$  tensor product space is isomorphic (as a vector space) to  $\mathbb{F}^{\prod_{j=1}^d n_j}$ .*

*Proof.* Recall that the tensor product space  $\otimes_{j=1}^d \mathbb{F}^{n_j}$  is the set of functions from  $\prod_{j=1}^d [n_j]$  to  $\mathbb{F}$ . As a  $\mathbb{F}$ -valued function space, it is thus an  $\mathbb{F}$ -vector space. That it is  $\prod_{j=1}^d n_j$  dimensional follows from the fact that this is the cardinality of the domain.

To see that the basis is as claimed, note that the function  $\otimes_{j=1}^d \vec{e}_{i_j,j}$  is equal to the tensor  $T(i'_1, \dots, i'_d) = \prod_{j=1}^d \mathbb{1}[i'_j = i_j]$ . It is then not hard to see that these tensors are a basis for the tensor product space.  $\square$



**Lemma 2.2.2** (Multilinearity of Tensor Product). *Suppose  $j \in [d]$ ,  $\vec{v}_j \in \mathbb{F}^{n_j}$ , and  $a, b \in \mathbb{F}$ . In the tensor product space  $\otimes_{j=1}^d \mathbb{F}^{n_j}$ , for any  $j_0 \in [d]$  and  $\vec{w} \in \mathbb{F}^{n_{j_0}}$  the following identity holds:*

$$\vec{v}_1 \otimes \cdots \otimes (a\vec{v}_{j_0} + b\vec{w}) \otimes \cdots \otimes \vec{v}_d = a(\vec{v}_1 \otimes \cdots \otimes \vec{v}_{j_0} \otimes \cdots \otimes \vec{v}_d) + b(\vec{v}_1 \otimes \cdots \otimes \vec{w} \otimes \cdots \otimes \vec{v}_d)$$

*Proof.* This follows directly from Definition 2.1.2. □

We now use these properties to establish a class of rank-preserving maps on tensors.

**Lemma 2.2.3.** *For  $j \in [d]$ , consider linear maps  $A_j : \mathbb{F}^{n_j} \rightarrow \mathbb{F}^{n'_j}$ .*

1. *The  $A_j$  induce a function on simple tensors  $\otimes_{j=1}^d \vec{v}_j \mapsto \otimes_{j=1}^d A_j \vec{v}_j$  which uniquely extends to a linear map on the tensor product space which is denoted  $\otimes_{j=1}^d A_j : \otimes_{j=1}^d \mathbb{F}^{n_j} \rightarrow \otimes_{j=1}^d \mathbb{F}^{n'_j}$ .*
2. *If the  $A_j$  are invertible, then so is  $\otimes_{j=1}^d A_j$  and its inverse is given by  $\otimes_{j=1}^d A_j^{-1}$ .*
3. *For  $T : \prod_{j=1}^d [n_j]^d \rightarrow \mathbb{F}$ ,  $\text{rank}(T) \geq \text{rank}((\otimes_{j=1}^d A_j)(T))$ , with equality if the  $A_j$  are invertible.*

*Proof.* (1): By Lemma 2.2.1 the tensor product space  $\otimes_{j=1}^d \mathbb{F}^{n_j}$  has a basis consisting entirely of simple tensors. Thus by standard linear algebra, the map  $\otimes_{j=1}^d \vec{v}_j \mapsto \otimes_{j=1}^d A_j \vec{v}_j$  on this basis extends uniquely to a linear map  $\otimes_{j=1}^d A_j$  on the entire tensor product space.

It must also be shown that the map  $\otimes_{j=1}^d A_j$  induced from the basis elements is also compatible with the map  $\otimes_{j=1}^d \vec{v}_j \mapsto \otimes_{j=1}^d A_j \vec{v}_j$  defined on the rest of the simple tensors. This fact follows from the linearity of the  $A_j$  and the multilinearity of the tensor product, Lemma 2.2.2. That is, we first use that each  $\vec{v}_j$  can be expressed in terms of the basis elements  $\vec{v}_j = \sum_{i_j=1}^{n_j} c_{i_j, j} \vec{e}_{i_j, j}$  and then notice that by multilinearity

of the tensor product we have

$$\begin{aligned}
\bigotimes_{j=1}^d A_j \vec{v}_j &= \bigotimes_{j=1}^d A_j \left( \sum_{i_j=1}^{n_j} c_{i_j,j} \vec{e}_{i_j,j} \right) = \bigotimes_{j=1}^d \left( \sum_{i_j=1}^{n_j} c_{i_j,j} A_j \vec{e}_{i_j,j} \right) \\
&= \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} c_{i_1,1} \cdots c_{i_d,d} \bigotimes_{j=1}^d A_j \vec{e}_{i_j,j} \\
&= \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} c_{i_1,1} \cdots c_{i_d,d} \left( \bigotimes_{j=1}^d A_j \right) \left( \bigotimes_{j=1}^d \vec{e}_{i_j,j} \right)
\end{aligned}$$

We observe similarly that  $\bigotimes_{j=1}^d \vec{v}_j = \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} c_{i_1,1} \cdots c_{i_d,d} (\bigotimes_{j=1}^d \vec{e}_{i_j,j})$ . As the unique linear map induced above defines  $(\bigotimes_{j=1}^d A_j)(\bigotimes_{j=1}^d \vec{v}_j)$  as

$$\sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} c_{i_1,1} \cdots c_{i_d,d} \bigotimes_{j=1}^d A_j (\bigotimes_{j=1}^d \vec{e}_{i_j,j})$$

, this shows that  $\bigotimes_{j=1}^d A_j \vec{v}_j = (\bigotimes_{j=1}^d A_j)(\bigotimes_{j=1}^d \vec{v}_j)$ , and so the two maps agree on the simple tensors.

It should also be noted that this argument is independent of the basis chosen, as long as the basis is chosen among the simple tensors. This fact follows from the fact that the induced map on the entire space agrees with the map only defined on the simple tensors. Thus, the map  $\bigotimes_{j=1}^d A_j$  is well-defined.

(2): Denote the linear maps  $A := \bigotimes_{j=1}^d A_j$ , and  $A^{-1} := \bigotimes_{j=1}^d A_j^{-1}$ . Part 1 of this lemma shows that the maps  $A$  and  $A^{-1}$  compose, in either order, to be the identity on the simple tensors. As there is a basis among the simple tensors, by Lemma 2.2.1, this means that  $A \circ A^{-1}$  and  $A^{-1} \circ A$  are both identity maps. Thus  $A^{-1}$  is indeed the inverse map of  $A$ .

(3): Consider a minimal simple tensor decomposition of  $T$ , so that  $T = \sum_{l=1}^r \bigotimes_{j=1}^d \vec{v}_{j,l}$ . By part 1 of this lemma, we have a simple tensor decomposition  $(\bigotimes_{j=1}^d A_j)T = \sum_{l=1}^r \bigotimes_{j=1}^d A_j \vec{v}_{j,l}$ . This establishes the desired rank inequality. To establish equality when the  $A_j$  are invertible it is enough to run the inequality in the opposite direction using the linear map  $\bigotimes_{j=1}^d A_j^{-1}$  and using part 2 of this lemma.  $\square$

We now use these rank-preserving maps to establish facts about tensors and their layers.

**Lemma 2.2.4.** *Consider  $T = \otimes_{j=1}^d \vec{v}_j \in \otimes_{j=1}^d \mathbb{F}^{n_j}$ , where  $T$  is split into layers as  $T = [T_1 | \cdots | T_{n_d}]$ . Then  $T_l = (\vec{v}_1 \otimes \cdots \otimes \vec{v}_{d-1}) \cdot \vec{v}_d(l) \in \otimes_{j=1}^{d-1} \mathbb{F}^{n_j}$ .*

*Proof.* Definition 2.1.4 and Definition 2.1.2 show that  $T_l(i_1, \dots, i_{d-1}) := T(i_1, \dots, i_{d-1}, l) = \vec{v}_1(i_1) \cdots \vec{v}_{d-1}(i_{d-1}) \cdot \vec{v}_d(l)$ . We can then note that this is exactly the function  $\vec{v}_1 \otimes \cdots \otimes \vec{v}_{d-1}$ , multiplied by the scalar  $\vec{v}_d(l)$ , as desired.  $\square$

**Lemma 2.2.5.** *Consider the operation of taking the  $l$ -th layer (along the  $d$ -th axis). This is a linear map  $L_l : \otimes_{j=1}^d \mathbb{F}^{n_j} \rightarrow \otimes_{j=1}^{d-1} \mathbb{F}^{n_j}$ .*

*Proof.* Given the tensor  $T(\cdot, \dots, \cdot)$ . Taking the  $l$ -th layer yields  $T_l(\cdot, \dots, \cdot) := T(\cdot, \dots, \cdot, l)$ . Thus, the statements  $T = S + R \implies T_l = S_l + R_l$ , and  $c \in \mathbb{F}, T = cS \implies T_l = cS_l$  hold because they are simply a restriction of the above identity.  $\square$

We can now prove the main lemma of this appendix, on how applying linear maps interacts with the layers of a tensor.

**Lemma 2.2.6.** *Consider  $T \in \otimes_{j=1}^d \mathbb{F}^{n_j}$ . Expand  $T$  into layers, so  $T = [T_1 | \cdots | T_{n_d}]$ .*

*Let  $(a_{i,j})_{i,j} \in \mathbb{F}^{m \times n_d}$  be a matrix. Define  $A : \mathbb{F}^{n_d} \rightarrow \mathbb{F}^m$  to be the linear map the matrix  $(a_{i,j})_{i,j}$  induces via the standard basis. Then,*

$$(I \otimes \cdots \otimes I \otimes A)(T) = \left[ \sum_{i_1=1}^{n_d} a_{1,i_1} T_{i_1} \mid \cdots \mid \sum_{i_m=1}^{n_d} a_{m,i_m} T_{i_m} \right]$$

*Proof.* The proof is in two parts. The first part proves the claim for simple tensors, and the second part extends the claim, using the linearity shown in Lemma 2.2.5, to general case.

We first prove the claim for simple tensors. Let  $T = \otimes_{j=1}^d \vec{v}_j$  be a simple tensor. Let  $\{\vec{e}_{i,d}\}_{i \in [n_d]}$  be the standard basis for  $\mathbb{F}^{n_d}$  and  $\{\vec{e}_{i',d}\}_{i' \in [m]}$  be the standard basis for  $\mathbb{F}^m$ . Then by expanding out in terms of the basis elements and using multilinearity,

we have

$$\begin{aligned} T &= \vec{v}_1 \otimes \cdots \otimes \vec{v}_d \\ &= \sum_{i=1}^{n_d} \vec{v}_1 \otimes \cdots \otimes \vec{v}_{d-1} \otimes (\vec{v}_d(i) \vec{e}_{i,d}) \end{aligned}$$

Denote  $T' := (I \otimes \cdots \otimes I \otimes A)(T)$ . So then,

$$\begin{aligned} T' &= \sum_{i=1}^{n_d} \vec{v}_1 \otimes \cdots \otimes \vec{v}_{d-1} \otimes A(\vec{v}_d(i) \vec{e}_{i,d}) \\ &= \sum_{i=1}^{n_d} \vec{v}_1 \otimes \cdots \otimes \vec{v}_{d-1} \otimes (\vec{v}_d(i) \cdot A(\vec{e}_{i,d})) \\ &= \sum_{i=1}^{n_d} \vec{v}_1 \otimes \cdots \otimes \vec{v}_{d-1} \otimes \left( \vec{v}_d(i) \cdot \sum_{i'=1}^m a_{i',i} \vec{e}_{i',d} \right) \\ &= \sum_{i=1}^{n_d} \sum_{i'=1}^m \vec{v}_d(i) \cdot a_{i',i} \cdot (\vec{v}_1 \otimes \cdots \otimes \vec{v}_{d-1} \otimes \vec{e}_{i',d}) \end{aligned}$$

By Lemma 2.2.4 and Lemma 2.2.5, we have,

$$T'_l = \sum_{i=1}^{n_d} \sum_{i'=1}^{n'_d} \vec{v}_d(i) \cdot a_{i',i} \cdot (\vec{v}_1 \otimes \cdots \otimes \vec{v}_{d-1}) \cdot \vec{e}_{i',d}(l)$$

and using that  $\vec{e}_{i',d}(l) = \llbracket i' = l \rrbracket$ ,

$$\begin{aligned} &= \sum_{i=1}^{n_d} \vec{v}_d(i) \cdot a_{l,i} \cdot (\vec{v}_1 \otimes \cdots \otimes \vec{v}_{d-1}) \\ &= \sum_{i=1}^{n_d} a_{l,i} T_i \end{aligned}$$

which establishes the claim for simple tensors.

Now let  $T \in \otimes_{j=1}^d \mathbb{F}^{n_j}$  be an arbitrary tensor. Consider a simple tensor expansion  $T = \sum_{k=1}^r S_k$  for  $S_k = \otimes_{j=1}^d \vec{v}_{j,k}$ . Denote  $S_{k,l}$  to be the  $l$ -th layer of  $S_k$ . So then as the  $S_k$  are simple, we have that  $(I \otimes \cdots \otimes I \otimes A)(S_k) = \left[ \sum_{i_1=1}^{n_d} a_{1,i_1} S_{k,i_1} \mid \cdots \mid \sum_{i_m=1}^{n_d} a_{m,i_m} S_{k,i_m} \right]$

by the above analysis. So then,

$$\begin{aligned} (I \otimes \cdots \otimes I \otimes A)(T) &= (I \otimes \cdots \otimes I \otimes A) \left( \sum_{k=1}^r S_k \right) \\ &= \sum_{k=1}^r \left[ \sum_{i_1=1}^{n_d} a_{1,i_1} S_{k,i_1} \mid \cdots \mid \sum_{i_m=1}^{n_d} a_{m,i_m} S_{k,i_m} \right] \end{aligned}$$

by linearity of taking layers, Lemma 2.2.5, we get

$$\begin{aligned} &= \left[ \sum_{k=1}^r \sum_{i_1=1}^{n_d} a_{1,i_1} S_{k,i_1} \mid \cdots \mid \sum_{k=1}^r \sum_{i_m=1}^{n_d} a_{m,i_m} S_{k,i_m} \right] \\ &= \left[ \sum_{i_1=1}^{n_d} a_{1,i_1} \left( \sum_{k=1}^r S_{k,i_1} \right) \mid \cdots \mid \sum_{i_m=1}^{n_d} a_{m,i_m} \left( \sum_{k=1}^r S_{k,i_m} \right) \right] \\ &= \left[ \sum_{i_1=1}^{n_d} a_{1,i_1} T_{i_1} \mid \cdots \mid \sum_{i_m=1}^{n_d} a_{m,i_m} T_{i_m} \right] \end{aligned}$$

which is the desired result.  $\square$

We now apply this to get a symmetry lemma.

**Corollary 2.2.7.** *Consider  $T \in \otimes_{j=1}^d \mathbb{F}^{n_j}$ . Expand  $T$  into layers, so  $T = [T_1 \mid \cdots \mid T_{n_d}]$ . For any permutation  $\sigma : [n_d] \rightarrow [n_d]$ ,*

$$\text{rank}([T_{\sigma(1)} \mid \cdots \mid T_{\sigma(n_d)}]) = \text{rank}([T_1 \mid \cdots \mid T_{n_d}])$$

*Proof.* Let  $P$  be the linear transformation defined by the permutation that  $\sigma$  induces on the basis vectors of  $\mathbb{F}^{n_d}$ . Then  $P$  is invertible, and so by Lemma 2.2.3.2 the induced transformation  $I \otimes \cdots \otimes I \otimes P$  is also invertible and so  $\text{rank}((I \otimes \cdots \otimes I \otimes P)(T)) = \text{rank}(T)$  by Lemma 2.2.3.2. Then, by Lemma 2.2.5 we see that  $(I \otimes \cdots \otimes I \otimes P)(T) = [T_1 \mid \cdots \mid T_{n_d}]$ .  $\square$

We also need another symmetry lemma.

**Lemma 2.2.8.** *For  $T \in \otimes_{j=1}^d \mathbb{F}^{n_j}$  and a permutation  $\sigma : [d] \rightarrow [d]$ , define  $T' \in \otimes_{j=1}^d \mathbb{F}^{n_{\sigma(j)}}$  by  $T'(i_1, \dots, i_d) = T(i_{\sigma^{-1}(1)}, \dots, i_{\sigma^{-1}(d)})$ . Then,  $\text{rank}(T) = \text{rank}(T')$ .*

*Proof.* We show  $\text{rank}(T) \geq \text{rank}(T')$ , and the equality follows by symmetry as  $\sigma$  is invertible. Consider a simple tensor decomposition  $T = \sum_{k=1}^r \otimes_{j=1}^d \vec{v}_{j,k}$ . It is then easy to see that  $T' = \sum_{k=1}^r \otimes \vec{v}_{\sigma(j),k}$  by considering the equation pointwise:  $T'(i_1, \dots, i_d) = T(i_{\sigma^{-1}(1)}, \dots, i_{\sigma^{-1}(d)}) = \sum_{k=1}^r \prod_{j=1}^d \vec{v}_{j,k}(i_{\sigma^{-1}(j)}) = \sum_{k=1}^r \prod_{j=1}^d \vec{v}_{\sigma(j),k}(i_j)$ . The conclusion then follows by considering a minimal rank expansion.  $\square$

Finally, we need a corollary about how dropping layers from a tensor affects rank.

**Corollary 2.2.9.** *For layers  $S_1, \dots, S_{n_d}, S' \in \otimes_{j=1}^{d-1} \mathbb{F}^{n_j}$ , we have that*

$$\text{rank}([S_1 | \dots | S_{n_d}]) \leq \text{rank}([S_1 | \dots | S_{n_d} | S'])$$

*with equality if  $S'$  is the zero layer.*

*Proof.* ( $\leq$ ): The projection map  $P$  induces the map  $(I \otimes \dots \otimes I \otimes P)$  which takes  $[S_1 | \dots | S_{n_d} | S']$  to  $[S_1 | \dots | S_{n_d}]$  by Lemma 2.2.5 and so Lemma 2.2.3.3 implies that the rank has not increased.

( $\geq$ ): So now assume  $S'$  is the zero layer. Then again we apply Lemma's 2.2.5 and Lemma 2.2.3.3 but now extend the natural inclusion map  $\iota : \mathbb{F}^{n_d} \rightarrow \mathbb{F}^{n_d+1}$  to a linear map  $(I \otimes \dots \otimes I \otimes \iota)$  on the tensors which takes  $[S_1 | \dots | S_{n_d}]$  to  $[S_1 | \dots | S_{n_d} | 0]$ , again showing that the rank has not increased.  $\square$

## 2.3 Layer Reduction

This section details a generalization of row-reduction, which we call layer-reduction. We show that layer-reduction can alter a tensor in such a way to provably reduce its rank. By showing this process can be repeated many times, a rank lower bound can be established.

The following lemma is the main technical part of this section. Håstad implicitly used<sup>1</sup> a version of this lemma in his proof that tensor rank is NP-Complete [12, 13]

---

<sup>1</sup>Håstad's usage, and proof, is reflected by Lemmas 2, 3 and 4 (and the following discussion) of the conference version [12]. The journal version [13] ascribes the origin of these lemmas to Lemma 2 in the work of Hopcroft and Kerr [14]

However, Håstad's usage requires that  $S_{n_d}$  is a rank-one tensor. This special case does not seem to directly imply our lemma, which was independently proven. While the special case is sufficient to lower-bound the combinatorially-constructed tensors of Section 3.1, the full lemma is needed to lower-bound the rank of the algebraically-constructed tensors of Section 3.2.

**Lemma 2.3.1** (Layer Reduction). *For layers  $S_1, \dots, S_{n_d} \in \bigotimes_{j=1}^{d-1} \mathbb{F}^{n_j}$  with  $S_{n_d}$  non-zero, there exist constants  $c_1, \dots, c_{n_d-1} \in \mathbb{F}$  such that*

$$\text{rank}([S_1 | \dots | S_{n_d}]) \geq \text{rank}([S_1 + c_1 S_{n_d} | \dots | S_{n_d-1} + c_{n_d-1} S_{n_d}]) + 1$$

*Proof.* Denote  $T := [S_1 | \dots | S_{n_d}]$ . The proof is in two steps. The first step defines a linear transformation  $A$  on  $\mathbb{F}^{n_d}$  such that the linear transformation  $I \otimes \dots \otimes I \otimes A$  is a higher-dimensional analogue of a row-reduction step in Gaussian elimination. That is, for  $T'$  the image of  $T$ , it is seen that  $T' = [S_1 + c_1 S_{n_d} | \dots | S_{n_d-1} + c_{n_d-1} S_{n_d} | S_{n_d}]$  by Lemma 2.2.6. The  $c_i$  are chosen in such a way so that  $T'$  has a minimal simple tensor expansion where some simple tensor  $R$  is non-zero only on the  $S_{n_d}$ -layer. In the second step, the  $S_{n_d}$ -layer is dropped and the remaining tensor  $T'' = [S_1 + c_1 S_{n_d} | \dots | S_{n_d-1} + c_{n_d-1} S_{n_d}]$  no longer requires  $R$  in its simple tensor expansion and so  $\text{rank}(T) \geq \text{rank}(T'') + 1$ .

Consider a minimal simple tensor expansion  $T = \sum_{k=1}^r \bigotimes_{j=1}^d \vec{v}_{j,k}$ . Expanding the  $\vec{v}_{d,k}$  in terms of basis vectors yields

$$T = \sum_{k=1}^r (\vec{v}_{1,k} \otimes \dots \otimes \vec{v}_{d-1,k}) \otimes (\vec{v}_{d,k}(1) \cdot \vec{e}_{1,d} + \dots + \vec{v}_{d,k}(n_d) \cdot \vec{e}_{n_d,d})$$

and in particular Lemma 2.2.4 shows that  $S_{n_d} = \sum_{k=1}^r (\vec{v}_{1,k} \otimes \dots \otimes \vec{v}_{d-1,k}) \cdot \vec{v}_{d,k}(n_d)$ . As  $S_{n_d}$  is non-zero there must be some  $k_0$  such that  $\vec{v}_{d,k_0}(n_d) \neq 0$ . Define  $A : \mathbb{F}^{n_d} \rightarrow \mathbb{F}^{n_d}$  to be the linear transformation defined by its action on the standard basis

$$A(\vec{e}_{i,d}) = \begin{cases} \vec{e}_{n_d,d} - \frac{\vec{v}_{d,k_0}(1)}{\vec{v}_{d,k_0}(n_d)} \vec{e}_{1,d} - \dots - \frac{\vec{v}_{d,k_0}(n_d-1)}{\vec{v}_{d,k_0}(n_d)} \vec{e}_{n_d-1,d} & \text{if } i = n_d \\ \vec{e}_{i,d} & \text{else} \end{cases}$$

Letting  $I$  denote the identity transformation, consider the tensor  $T' := (I \otimes \cdots \otimes I \otimes A)(T) \in \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_d}$ . By Lemma 2.2.6, we observe that  $T' = [S_1 + c_1 S_{n_d} | \cdots | S_{n_d-1} + c_{n_d-1} S_{n_d} | S_{n_d}]$ , where  $c_i = -\frac{\vec{v}_{d,k_0}(j)}{\vec{v}_{d,k_0}(n_d)}$ .

By Lemma 2.2.3 we have the simple tensor expansion  $T' = \sum_{k=1}^r \vec{v}_{1,k} \otimes \cdots \otimes \vec{v}_{d-1,k} \otimes A\vec{v}_{d,k}$ . By construction,  $A(\vec{v}_{d,k_0}) = \vec{v}_{d,k_0}(n_d) \cdot \vec{e}_{n_d,d}$ . Using Lemma 2.2.4 we observe that the simple tensor  $\vec{v}_{1,k_0} \otimes \cdots \otimes \vec{v}_{d-1,k_0} \otimes A\vec{v}_{d,k_0}$  has non-zero entries only on the  $S_{n_d}$ -layer.

We now define the linear transformation  $A' : \mathbb{F}^{n_d} \rightarrow \mathbb{F}^{n_d-1}$  defined by

$$A'(\vec{e}_{i,d}) = \begin{cases} \vec{0} & \text{if } i = n_d \\ \vec{e}_{i,d} & \text{else} \end{cases}$$

This will correspond to dropping the  $S_{n_d}$ -layer. We can compose this with  $A$  to get  $A'' = A' \circ A$ , defined by

$$A''(\vec{e}_{i,d}) = \begin{cases} -\frac{\vec{v}_{d,k_0}(1)}{\vec{v}_{d,k_0}(n_d)} \vec{e}_{1,d} - \cdots - \frac{\vec{v}_{d,k_0}(n_d-1)}{\vec{v}_{d,k_0}(n_d)} \vec{e}_{n_d-1,d} & \text{if } i = n_d \\ \vec{e}_{i,d} & \text{else} \end{cases}$$

So now we take  $T'' = (I \otimes \cdots \otimes I \otimes A'')(T)$ . By Lemma 2.2.6 we see that  $T'' = [S_1 + c_1 S_{n_d} | \cdots | S_{n_d-1} + c_{n_d-1} S_{n_d}]$ . Further, we observe now that by construction  $A''(\vec{v}_{d,k_0}) = \vec{0}$ . This leads to the simple tensor expansion,

$$\begin{aligned} T'' &= \sum_{k=1}^r \vec{v}_{1,k} \otimes \cdots \otimes \vec{v}_{d-1,k} \otimes A''\vec{v}_{d,k} \\ &= \vec{v}_{1,k_0} \otimes \cdots \otimes \vec{v}_{d-1,k_0} \otimes A''\vec{v}_{d,k_0} + \sum_{k=1, k \neq k_0}^r \vec{v}_{1,k} \otimes \cdots \otimes \vec{v}_{d-1,k} \otimes A''\vec{v}_{d,k} \\ &= \vec{v}_{1,k_0} \otimes \cdots \otimes \vec{v}_{d-1,k_0} \otimes \vec{0} + \sum_{k=1, k \neq k_0}^r \vec{v}_{1,k} \otimes \cdots \otimes \vec{v}_{d-1,k} \otimes A''\vec{v}_{d,k} \\ &= \sum_{k=1, k \neq k_0}^r \vec{v}_{1,k} \otimes \cdots \otimes \vec{v}_{d-1,k} \otimes A''\vec{v}_{d,k} \end{aligned}$$

Therefore  $\text{rank}(T'') \leq r - 1 = \text{rank}(T) - 1$ , and thus  $\text{rank}(T) \geq \text{rank}(T'') - 1$ .  $\square$



The layer-reduction lemma will mostly be used via the following extension.

**Corollary 2.3.2** (Iterative Layer-Reduction). *For layers  $S_1, \dots, S_{n_d} \in \mathbb{F}^{n_1} \otimes \dots \otimes \mathbb{F}^{n_{d-1}}$  with  $S_1, \dots, S_m$  linearly independent (as vectors in the space  $\mathbb{F}^{n_1 \dots n_{d-1}}$ ), there exist constants  $c_{i,j} \in \mathbb{F}$ ,  $i \in \{1, \dots, m\}$ ,  $j \in \{m+1, \dots, n_d\}$ , such that*

$$\text{rank}([S_1 | \dots | S_{n_d}]) \geq \text{rank} \left( \left[ \begin{array}{c} S_{m+1} + \sum_{i=1}^m c_{i,m+1} S_i \\ \vdots \\ S_{n_d} + \sum_{i=1}^m c_{i,n_d} S_i \end{array} \right] \right) + m \quad (2.3.3)$$

*Proof.* The proof is by induction on  $m$ .

$m = 1$ : This is Lemma 2.3.1, up to reordering of the layers, with the observation that the singleton set  $\{S_1\}$  is linearly-independent iff  $S_1$  is non-zero. The reordering of layers is justified by Lemma 2.2.7.

$m > 1$ : By the induction hypothesis we have that

$$\text{rank}([S_1 | \dots | S_{n_d}]) \geq \text{rank} \left( \left[ \begin{array}{c} S_m + \sum_{i=1}^{m-1} c_{i,m} S_i \\ \vdots \\ S_{n_d} + \sum_{i=1}^{m-1} c_{i,n_d} S_i \end{array} \right] \right) + m - 1 \quad (2.3.4)$$

for the appropriate set of constants  $c_{i,j}$ . As the  $S_i$  are linearly independent,  $S_m + \sum_{i=1}^{m-1} c_{i,m} S_i$  is non-zero and so we can eliminate this layer from

$$\left[ \begin{array}{c} S_m + \sum_{i=1}^{m-1} c_{i,m} S_i \\ \vdots \\ S_{n_d} + \sum_{i=1}^{m-1} c_{i,n_d} S_i \end{array} \right]$$

by Lemma 2.3.1 and consequently have

$$\begin{aligned} & \text{rank} \left( \left[ \begin{array}{c} S_m + \sum_{i=1}^{m-1} c_{i,m} S_i \\ \vdots \\ S_{n_d} + \sum_{i=1}^{m-1} c_{i,n_d} S_i \end{array} \right] \right) \\ & \geq \text{rank} \left( \left[ \left( \left( S_{m+1} + \sum_{i=1}^{m-1} c_{i,m+1} S_i \right) + c_{m,m+1} \left( S_m + \sum_{i=1}^{m-1} c_{i,m} S_i \right) \right) \right. \right. \\ & \quad \left. \left. \dots \left( \left( S_{n_d} + \sum_{i=1}^{m-1} c_{i,n_d} S_i \right) + c_{m,n_d} \left( S_m + \sum_{i=1}^{m-1} c_{i,m} S_i \right) \right) \right] \right) + 1 \quad (2.3.5) \end{aligned}$$

where the  $c_{m,j}$  are new constants. Now define

$$c'_{i,j} = \begin{cases} c_{i,j} + c_{m,j}c_{i,m} & \text{if } i \neq m \\ c_{m,j} & \text{else} \end{cases} \quad (2.3.6)$$

Combining Equations (2.3.4), (2.3.5), and (2.3.6) yields the desired Equation (2.3.3).  $\square$

Notice that by Lemma 2.2.8 we can in fact use Lemma 2.3.1 and Corollary 2.3.2 along any axis, not just the  $d$ -th one.

*Remark 2.3.7.* Lemma 2.3.1 shows that the rank of  $T : \prod_{j=1}^d [n_j] \rightarrow \mathbb{F}$  is at least 1 more than the rank of some  $T' : \prod_{j=1}^d [n'_j] \rightarrow \mathbb{F}$ , where  $n'_j = n_j$  for all  $j \neq j_0$ , and  $n'_{j_0} = n_{j_0} - 1$ . In using this lemma, the quantity  $\sum_{j=1}^d n_j$  decreases by one. Therefore, we can never hope to apply this lemma more than  $\sum_{j=1}^d n_j$  many times, and thus using this lemma alone will never produce lower bounds larger than this quantity. Corollary 2.3.2 simply applies Lemma 2.3.1, so the same barriers apply.

# Chapter 3

## Tensor Rank Lower Bounds

### 3.1 Combinatorially-defined Tensors

In this section, we construct combinatorially-defined tensors and prove linear lower bounds for their rank. To do so, we use the follow fact about tensors, which is proved in Section 2.3. For matrices, this can be seen as a statement about Gaussian elimination.

**Corollary** (Iterative Layer Reduction, Corollary 2.3.2). *For layers  $S_1, \dots, S_{n_d} \in \mathbb{F}^{n_1} \otimes \dots \otimes \mathbb{F}^{n_{d-1}}$  with  $S_1, \dots, S_m$  linearly independent (as vectors in the space  $\mathbb{F}^{n_1 \cdots n_{d-1}}$ ), there exist constants  $c_{i,j} \in \mathbb{F}$ ,  $i \in \{1, \dots, m\}$ ,  $j \in \{m+1, \dots, n_d\}$ , such that*

$$\text{rank}([S_1 | \dots | S_{n_d}]) \geq \text{rank} \left( \left[ \begin{array}{c|c} S_{m+1} + \sum_{i=1}^m c_{i,m+1} S_i & \dots & S_{n_d} + \sum_{i=1}^m c_{i,n_d} S_i \end{array} \right] \right) + m \quad (3.1.1)$$

The idea of this section is to construct tensors such that we can apply Corollary 2.3.2 as many times as possible. As mentioned in Remark 2.3.7, for a  $[n]^d$  tensor, the lemma can be applied at most  $dn$  times, and thus the lower bounds can at best be  $dn$ . In general, the lemma may not be able to be applied this much because the elimination of layers zeroes out too much of the tensor. However, in this section we construct tensors (for  $d = 3$ ) such that we can almost apply the lemma  $dn$  times. The result is that we give explicit (order 3) 0/1-tensors with tensor rank exactly

$3n - \Theta(\lg n)$  over any field. To begin, we apply the above corollary twice, along two different axes, to get the following lemma.

**Lemma 3.1.2.** *Let  $A_1, \dots, A_k$  be  $n \times n$  sized  $\mathbb{F}$ -matrices. Let  $I_m$  denote the  $m \times m$  identity matrix, and  $0_m$  denote the  $m \times m$  zero matrix. Then,*

$$\text{rank} \left( \left[ \begin{array}{cc|cc} I_n & 0_n & 0_n & 0_n \\ 0_n & I_n & A_1 & 0_n \end{array} \middle| \dots \middle| \begin{array}{cc} 0_n & 0_n \\ A_k & 0_n \end{array} \right] \right) \geq \text{rank}([A_1 | \dots | A_k]) + 2n \quad (3.1.3)$$

and

$$\text{rank} \left( \left[ \begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ I_n & 0_n & 0 & 0_n & 0_n & 0 \\ 0_n & I_n & 0 & A_1 & 0_n & 0 \end{array} \middle| \dots \middle| \begin{array}{ccc} 0 & 0 & 0 \\ 0_n & 0_n & 0 \\ A_k & 0_n & 0 \end{array} \right] \right) \geq \text{rank}([A_1 | \dots | A_k]) + 2n \quad (3.1.4)$$

where the left-hand side of Equation 3.1.4 expresses the tensor rank of a  $[2n + 1] \times [2n + 1] \times [k + 1]$ -sized tensor.

*Proof.* Notice that the left hand sides of Equation 3.1.3 and Equation 3.1.4 are equal. This follows from applying Corollary 2.2.9 twice (using that this corollary extends to layers along any axes, not just the  $d$ -th, by applying Lemma 2.2.8), once on the layers slicing the page vertically, and once on the layers slicing the page horizontally. Thus, it is enough to show Equation 3.1.3.

We now apply Corollary 2.3.2. First, we use it on the layers slicing the page vertically and deriving that

$$\text{rank} \left( \left[ \begin{array}{cc|cc} I_n & 0_n & 0_n & 0_n \\ 0_n & I_n & A_1 & 0_n \end{array} \middle| \dots \middle| \begin{array}{cc} 0_n & 0_n \\ A_k & 0_n \end{array} \right] \right) \geq \text{rank} \left( \left[ \begin{array}{c|c} I_n & 0_n \\ C & A_1 \end{array} \middle| \dots \middle| \begin{array}{c} 0_n \\ A_k \end{array} \right] \right) + n \quad (3.1.5)$$

where  $C$  is an  $n \times n$  matrix of field elements defined by the constants  $c_{i,j}$  of Corollary 2.3.2. Notice that the layers being dropped in the use of this corollary must be linearly independent. However, as they are the layers of  $[I_n | 0_n | \dots | 0_n]$  which slice the page vertically, they have exactly one 1 in the first row<sup>1</sup>, and have 0 entries

---

<sup>1</sup>It is immaterial whether we call this a “row” or “column”, as no specific orientation of these

elsewhere. As their non-zero entries are in different positions, they are linearly independent. Similarly, we can apply the corollary again on the remaining layers that slice the page horizontally to see that

$$\text{rank} \left( \left[ \begin{array}{c|c|c} I_n & 0_n & \cdots & 0_n \\ \hline C & A_1 & \cdots & A_k \end{array} \right] \right) \geq \text{rank}([C + C'|A_1| \cdots |A_k]) + n \quad (3.1.6)$$

where  $C'$  is yet another  $n \times n$  matrix of field elements produced by Corollary 2.3.2. We now invoke Corollary 2.2.9 to observe that

$$\text{rank}([C + C'|A_1| \cdots |A_k]) \geq \text{rank}([A_1| \cdots |A_k]) \quad (3.1.7)$$

Combining Equations (3.1.5), (3.1.6), and (3.1.7) yields Equation (3.1.3) and thus the claim.  $\square$

Applying this lemma recursively yields the following construction.

**Definition 3.1.8.** Let  $H : \mathbb{N} \rightarrow \mathbb{N}$  denote the Hamming weight function. That is,  $H(n)$  is the number of 1's in the binary expansion of  $n$ .

**Theorem 3.1.9.** For  $i \in \{0, \dots, \lfloor \lg n \rfloor\}$ , let  $S_{n,i}$  be an  $n \times n$  matrix defined in the following recursive manner.

- $S_{1,0} = [1]$
- For  $2n > 1$ ,

$$S_{2n,i} = \begin{cases} \begin{bmatrix} 0_n & 0_n \\ S_{n,i} & 0_n \end{bmatrix} & \text{if } i < \lfloor \lg n \rfloor \\ \begin{bmatrix} I_n & 0_n \\ 0_n & I_n \end{bmatrix} & \text{if } i = \lfloor \lg n \rfloor \end{cases}$$

---

tensors was chosen.

- For  $2n + 1 > 1$ ,

$$S_{2n+1,i} = \begin{cases} \begin{bmatrix} 0 & 0 & 0 \\ 0_n & 0_n & 0 \\ S_{n,i} & 0_n & 0 \end{bmatrix} & \text{if } i < \lfloor \lg n \rfloor \\ \begin{bmatrix} 0 & 0 & 0 \\ I_n & 0_n & 0 \\ 0_n & I_n & 0 \end{bmatrix} & \text{if } i = \lfloor \lg n \rfloor \end{cases}$$

Then, denoting  $T_n = [S_{n,0} | \cdots | S_{n,\lfloor \lg n \rfloor}]$ ,

1.  $T_n$  has size  $[n] \times [n] \times [\lfloor \lg n \rfloor + 1]$ .
2.  $\text{rank}(T_n) = 2n - 2H(n) + 1$ .
3. On inputs  $n$  and  $(i, j, k) \in [n] \times [n] \times [\lfloor \lg n \rfloor + 1]$ ,  $T_n(i, j, k)$  can be computed in polynomial time. That is, in time  $O(\text{polylog}(n))$ .

*Proof.* (1): This is clear from construction.

(2): We first note that  $\lfloor \lg 2n \rfloor = \lfloor \lg n + 1 \rfloor = \lfloor \lg n \rfloor + 1$ . We first prove the upper bound, and then the lower bound.

To see that  $\text{rank}(T_n) \leq 2n - 2H(n) + 1$  we observe that  $T_n$  has exactly this many non-zero entries. Denote this quantity  $r_n$ . We proceed by induction on the recursive definition of the  $S_{n,i}$ . For  $n = 1$ , there is clearly exactly  $2 \cdot 1 - 2H(1) + 1 = 1$  non-zero entry. For  $2n > 1$ ,  $r_{2n} = r_n + 2n$  which by induction yields  $r_{2n} = (2n - 2H(n) + 1) + 2n$ . Observing that  $H(n) = H(2n)$ , we see that  $r_{2n} = 2(2n) - 2H(2n) + 1$ . For  $2n + 1 > 1$ ,  $r_{2n+1} = r_n + 2n$ , which by induction yields  $r_{2n+1} = (2n - 2H(n) + 1) + 2n$ . Noticing that  $H(2n + 1) = H(n) + 1$  we have that  $r_{2n+1} = 2n - 2(H(2n + 1) - 1) + 1 + 2n = 4n + 2 - 2H(2n + 1) + 1 = 2(2n + 1) - 2H(2n + 1) + 1$ . Thus, the induction hypothesis shows that  $r_n = 2n - 2H(n) + 1$  for all  $n$ , and thus upper-bounding the rank by this quantity.

For the rank lower bound, we use Lemma 3.1.2 and induction on the recursive definition of the  $S_{n,i}$ . Clearly,  $\text{rank}(T_1) \geq 1$ . Then for  $2n > 1$ ,  $\text{rank}(T_{2n}) \geq \text{rank}(T_n) + 2n$ , and for  $2n + 1 > 1$ ,  $\text{rank}(T_{2n+1}) \geq \text{rank}(T_n) + 2n$ . These are exactly the same recurrences from the proceeding paragraph, and so they have the same solution:  $\text{rank}(T_n) \geq 2n - 2H(n) + 1$ .

Combining these two bounds shows that  $\text{rank}(T_n) = 2n - 2H(n) + 1$ .

(3): This is clear from the equations defining the  $S_{n,i}$ . □

Another application of Corollary 2.3.2 (along the one axis it has not yet been applied) yields the following claim.

**Corollary 3.1.10.** *Define  $S_{n,i}$  as in Theorem 3.1.9. Let  $n \in \mathbb{N}$  and restrict to  $n \geq 2$ . Then, for  $i \in [n]$  define  $n \times n$  matrices  $S'_{n,i}$  by*

$$S'_{n,i} = \begin{cases} \begin{bmatrix} S_{n-1,i-1} & 0 \\ 0 & 0 \end{bmatrix} & \text{if } i \in [\lfloor \lg(n-1) \rfloor + 1] \\ \begin{bmatrix} 0_{n-1} & \vec{e}_{i - (\lfloor \lg(n-1) \rfloor + 1)} \\ 0 & 0 \end{bmatrix} & \text{else} \end{cases}$$

where  $\vec{e}_j \in \mathbb{F}^{n-1}$  is the indicator column vector where  $\vec{e}_j(k) = \llbracket j = k \rrbracket$ . (Notice that  $\lfloor \lg(n-1) \rfloor + 1 \leq n-1$  for all  $n \geq 2$ .) Then, denoting  $T'_n = [S'_{n,1} | \cdots | S'_{n,n}]$ ,

1.  $T'_n$  has size  $[n]^3$ .
2.  $\text{rank}(T'_n) = 3n - 2H(n-1) - \lfloor \lg(n-1) \rfloor - 2 \geq 3n - \Theta(\lg n)$ .
3. On inputs  $n$  and  $(i, j, k) \in [n] \times [n] \times [n]$ ,  $T'_n(i, j, k)$  can be computed in polynomial time, that is,  $O(\text{polylog}(n))$ .

*Proof.* (1): This is clear from construction.

(2): Observe that in the construction of  $T'_n$ , the matrices  $S'_{n,i}$  for  $i > \lfloor \lg(n-1) \rfloor + 1$  are linearly independent. Thus, applying Corollary 2.3.2, we see that

$$\text{rank}(T'_n) \geq \text{rank}([\tilde{S}'_{n,1} | \cdots | \tilde{S}'_{n, \lfloor \lg(n-1) \rfloor + 1}]) + n - (\lfloor \lg(n-1) \rfloor + 1)$$

where

$$\tilde{S}'_{n,i} = \begin{bmatrix} S_{n-1,i-1} & \vec{c}_i \\ 0 & 0 \end{bmatrix}$$

for some arbitrary vectors  $\vec{c}_i \in \mathbb{F}^{n-1}$ . It follows from Corollary 2.2.9 that we can drop the bottom row and last column of each of the  $\tilde{S}'_{n,i}$  without increasing the rank, so that

$$\text{rank}([\tilde{S}'_{n,1} | \cdots | \tilde{S}'_{n, \lfloor \lg(n-1) \rfloor + 1}]) \geq \text{rank}([S_{n-1,0} | \cdots | S_{n-1, \lfloor \lg(n-1) \rfloor}])$$

where the  $S_{n-1,i-1}$  are as defined in Theorem 3.1.9, and as such,

$$\text{rank}([S_{n-1,0} | \cdots | S_{n-1, \lfloor \lg(n-1) \rfloor}]) = 2(n-1) + 2H(n-1) + 1$$

. Combining these inequalities yields the rank lower bound for  $T'_n$ .

(3): This is clear from the equations defining  $T'_n$ , and using the explicitness of the  $S_{n-1,i-1}$  as seen from Theorem 3.1.9.  $\square$

Note that this analysis is exact. The next section contains similar lower bounds over  $\mathbb{F}_2$  using different methods, where no non-trivial upper bound is known.

## 3.2 Algebraically-defined Tensors

The results of this section will be field-specific, and so we no longer work over an arbitrary field.

**Lemma 3.2.1.** *Let  $\mathbb{F}_q$  be the field of  $q$  elements. Consider  $n \times n$  matrices  $M_1, \dots, M_k$  over  $\mathbb{F}_q$  such that all non-zero linear combinations have full-rank. Then the tensor  $T = [M_1 | \cdots | M_k]$  has tensor rank at least  $\frac{q^k - 1}{q^k - q^{k-1}} n$ .*

*Proof.* The proof is via the probabilistic method, using randomness to perform an analogue of gate elimination. For non-zero  $\vec{c} \in \mathbb{F}_q^k$ , the summation  $\vec{c} \cdot \vec{M}$  will nullify terms in a simple tensor expansion with some probability. This will in expectation reduce the rank. We then invoke the hypothesis that the result is full-rank, to conclude the bound on the original rank.



Consider a minimal simple tensor decomposition  $T = \sum_{i=1}^r \vec{u}_i \otimes \vec{v}_i \otimes \vec{w}_i$ . For  $\vec{c} \in \mathbb{F}_q^k$ , consider (notation-abused) dot-product  $\langle \vec{c}, \vec{M} \rangle$ , which can also be written as the matrix  $\sum_{i=1}^k c_i M_i$ . By Lemma 2.2.6 it can be seen that this is the image of  $T$  under the linear transformation  $I \otimes I \otimes A$ , where  $A$  is the linear transformation that sends the basis element  $\vec{e}_i$  to  $c_i \vec{e}_1$ . Consequently, we have that  $\langle \vec{c}, \vec{M} \rangle = (I \otimes I \otimes A)(T) = \sum_{i=1}^r \vec{u}_i \otimes \vec{v}_i \otimes (A\vec{w}_i)$ . Noticing that  $A\vec{w}_i = \langle \vec{c}, \vec{w}_i \rangle$ , and that we can then treat this as a matrix instead of one-layer tensor, we see that  $\langle \vec{c}, \vec{M} \rangle = \sum_{i=1}^r \langle \vec{c}, \vec{w}_i \rangle \vec{u}_i \otimes \vec{v}_i$ .

Minimality implies that  $\vec{u}_i \neq 0$  for all  $i$ . So for a fixed  $i$ , the set of  $\vec{c}$  such that  $\langle \vec{c}, \vec{w}_i \rangle = 0$  is a 1-dimensional subspace by the Rank-Nullity theorem. Using that the field size is  $q$ , this shows that

$$\Pr_{\vec{c} \in {}_u\mathbb{F}_q^k \setminus \{0\}} [\langle \vec{c}, \vec{w}_i \rangle \neq 0] = \frac{q^k - q^{k-1}}{q^k - 1}$$

Now define  $S_{\vec{c}} := \{i | \langle \vec{c}, \vec{w}_i \rangle \neq 0\}$ . By linearity of expectation,  $\mathbb{E}_{\vec{c} \in {}_u\mathbb{F}_q^k \setminus \{0\}} [|S_{\vec{c}}|] = \frac{q^k - q^{k-1}}{q^k - 1} r$ . Thus, there exists a non-zero  $\vec{c}_0$  such that  $|S_{\vec{c}_0}| \leq \frac{q^k - q^{k-1}}{q^k - 1} r$ . Therefore, we can write the matrix  $\langle \vec{c}, \vec{M} \rangle$  as  $\langle \vec{c}, \vec{M} \rangle = \sum_{i=1}^r \langle \vec{c}, \vec{w}_i \rangle \vec{u}_i \otimes \vec{v}_i = \sum_{i \in S} \langle \vec{c}, \vec{w}_i \rangle \vec{u}_i \otimes \vec{v}_i$ . The hypothesis on the  $M_i$  says that  $\langle \vec{c}, \vec{M} \rangle$  is of full-rank, and therefore we have that  $n \leq \text{rank}(\langle \vec{c}, \vec{M} \rangle) \leq |S| \leq \frac{q^k - q^{k-1}}{q^k - 1} r$ . As  $\text{rank}(T) = r$ , we have that  $\text{rank}(T) \geq \frac{q^k - 1}{q^k - q^{k-1}} n$ .  $\square$

**Corollary 3.2.2.** *Let  $\mathbb{F}_q$  be the field of  $q$  elements. Consider  $n \times n$  matrices  $M_1, \dots, M_n$  over  $\mathbb{F}_q$  such that all non-zero linear combinations have full-rank. Then the tensor  $T = [M_1 | \dots | M_n]$  has tensor rank at least  $\frac{2q-1}{q-1}n - \lceil \log_q n \rceil - \frac{q}{q-1} = \frac{2q-1}{q-1}n - \Theta(\log_q n)$ .*

*Proof.* Let  $k \leq n$  be a parameter, to be optimized over later.

Notice that the hypothesis show that the matrices  $M_i$  are linearly independent and so Corollary 2.3.2 shows that

$$\text{rank}([M_1 | \dots | M_n]) \geq \text{rank}([M_1 + M'_1 | \dots | M_k + M'_k]) + (n - k)$$

where the  $M'_i$  are linear combinations of the  $M_{k+1}, \dots, M_n$ . Thus, any non-zero linear combination of the  $(M_i + M'_i)$  is necessarily a non-zero linear combination of the  $M_i$ .

In particular, this shows that any non-zero linear combination of the  $(M_i + M'_i)$  has full-rank. Thus, by Lemma 3.2.1

$$\text{rank}([M_1 + M'_1 | \cdots | M_k + M'_k]) \geq \frac{q^k - 1}{q^k - q^{k-1}} n$$

and so

$$\text{rank}([M_1 | \cdots | M_n]) \geq n - k + \frac{q^k - 1}{q^k - q^{k-1}} n =: f(k)$$

One can observe that  $k = \log_q \left( \frac{n \ln q}{1 - 1/q} \right)$  maximizes  $f$ , but asymptotically it is sufficient to take  $k = \lceil \log_q n \rceil$ . Then

$$\begin{aligned} f(k) &= n - k + \frac{1 - q^{-k}}{1 - 1/q} n \\ &\geq n - \lceil \log_q n \rceil + \frac{1 - 1/n}{1 - 1/q} n \\ &\geq n - \lceil \log_q n \rceil + \frac{1}{1 - 1/q} n - \frac{1/n}{1 - 1/q} n \\ &\geq n - \lceil \log_q n \rceil + \frac{q}{q-1} n - \frac{q}{q-1} \\ &\geq \frac{2q-1}{q-1} n - \lceil \log_q n \rceil - \frac{q}{q-1} \end{aligned}$$

As  $f(k)$  lower-bounds the rank by the above, this establishes the claim.  $\square$

The above lemma and its corollary establish a property implying tensor rank lower bounds. We now turn to constructing tensors that have this property. Clearly we seek explicit tensors, and by this we mean that each entry of the tensor is efficiently computable.

We first observe that the property can be easily constructed given explicit field extensions of the base field  $\mathbb{F}$ .

**Proposition 3.2.3.** *Let  $\mathbb{F}$  be a field and  $f \in \mathbb{F}[x]$  be an irreducible polynomial of degree  $n$ . Then there exists  $n \times n$   $\mathbb{F}$ -matrices  $M_1, \dots, M_n$ , such that all non-zero  $\mathbb{F}$ -linear combinations of the  $M_i$  have full-rank. Furthermore, the entries of each matrix are computable in algebraic circuits of size  $O(\text{polylog}(n)\text{poly}(\|f\|_0))$ , where  $\|f\|_0$  is*

the number of non-zero coefficients of  $f$ .

*Proof.* Let  $f(x) = a_n x^n + \dots + a_1 x + a_0$ . Recall that  $\mathbb{K} = \mathbb{F}[x]/(f)$  is a field, and because  $\deg f = n$ ,  $\mathbb{K}$  is a  $n$ -dimensional  $\mathbb{F}$ -vector space, where we choose  $1, x, \dots, x^{n-1}$  as the basis. This gives an  $\mathbb{F}$ -algebra isomorphism  $\mu$  between  $\mathbb{K}$  and a sub-ring  $M$  of the  $m \times m$   $\mathbb{F}$ -matrices, where  $M$  is defined as the image of  $\mu$ . The map  $\mu$  is defined by associating  $\alpha \in \mathbb{K}$  with the matrix inducing the linear map  $\mu(\alpha) : \mathbb{F}^m \rightarrow \mathbb{F}^m$ , where  $\mu(\alpha)$  is the multiplication map of  $\alpha$ . That is, using that  $\mathbb{K} = \mathbb{F}^m$  we can see that the map  $\beta \mapsto \alpha\beta$  for  $\beta \in \mathbb{K}$  is an  $\mathbb{F}$ -linear map, and thus defines  $\mu(\beta)$  over  $\mathbb{F}^m$ .

That the map is injective follows from the fact that  $\mu(\alpha)$  must map  $1 \in \mathbb{K}$  to  $\alpha \in \mathbb{K}$ , so  $\alpha$  is recoverable from  $\mu(\alpha)$  (and surjectivity follows by definition of  $M$ ). To see the required homomorphism properties is also not difficult. As  $(\alpha + \gamma)\beta = \alpha\beta + \gamma\beta$  for any  $\alpha, \beta, \gamma \in \mathbb{K}$ , this shows that  $\mu(\alpha + \gamma) = \mu(\alpha) + \mu(\gamma)$  as linear maps, and thus as matrices. Similarly, as  $(\alpha\gamma)\beta = \alpha(\gamma\beta)$  for any  $\alpha, \beta, \gamma \in \mathbb{K}$  it must be that  $\mu(\alpha\gamma) = \mu(\alpha)\mu(\gamma)$ . That this map interacts linearly in  $\mathbb{F}$  implies that it is an  $\mathbb{F}$ -algebra homomorphism, as desired.

In particular, this means that  $\alpha \in \mathbb{K}$  is invertible iff the matrix  $\mu(\alpha) \in M \subseteq \mathbb{F}^{n \times n}$  is invertible. As  $\mathbb{K}$  is a field, the only non-invertible matrix in  $M$  is  $\mu(0)$ . The  $\mathbb{F}$ -algebra homomorphism means that for  $a_i \in \mathbb{F}$  and  $\alpha_i \in \mathbb{K}$ , the linear combination  $\sum a_i \mu(\alpha_i)$  equals  $\mu(\sum a_i \alpha_i)$  and so the matrix  $\sum a_i \mu(\alpha_i)$  is invertible iff  $\sum a_i \alpha_i \neq 0$ . Thus, as  $1, x, \dots, x^{n-1}$  are  $\mathbb{F}$ -linearly independent in  $\mathbb{K}$ , it follows that the matrix  $\mu(1), \mu(x), \dots, \mu(x^{n-1})$  have that all non-zero  $\mathbb{F}$ -linear combinations are invertible, as desired.

We now study how to compute  $\mu(x^i)$ . Observe that acting as a linear map on  $\mathbb{F}^n$ ,  $\mu(x^i)$  sends  $x^{i+j} \pmod{f}$ . To read off the  $x^k$  component can be done with a lookup table to the coefficients of  $f$ , and thus in  $O(\text{poly}(\|f\|_0))$  size circuits.  $\square$

To make the above construction explicit, we need to show that the irreducible polynomial  $f$  can be found efficiently. We now cite the following result of Shoup [22]. It says that we can find irreducible polynomials in finite fields in polynomial-time provided that the field size is fixed.

**Theorem 3.2.4** ([22], Theorem 4.1). *For any prime or prime power  $q$ , an irreducible polynomial of degree  $n$  in  $\mathbb{F}_q[x]$  can be found in time  $O(\text{poly}(nq))$ .*

Combining Corollary 3.2.2, Proposition 3.2.3, and Theorem 3.2.4, we arrive at the following result.

**Corollary 3.2.5.** *For any fixed prime or prime power  $q$ , over the field  $\mathbb{F}_q$  there is a family of tensors  $T_n$  of size  $[n]^3$  such that*

1.  $\text{rank}(T_n) \geq \frac{2q-1}{q-1}n - \Theta(\log_q n)$
2. *On inputs  $n$  and  $(i, j, k) \in [n]^3$ ,  $T_n(i, j, k)$  is computable in  $O(\text{poly}(nq))$*

Note that this is strictly worse than Corollary 3.1.10 in two respects. First, while this result asymptotically matches the lower bound of Corollary 3.1.10 over  $\mathbb{F}_2$ , the above result is only valid over finite fields, and as the field size grows, the lower bound approaches  $2n - o(n)$ . This seems inherent in the approach.

Further, the given construction is less explicit as computing even a single entry of the tensor might require examining all  $n$  of the coefficients in the irreducible polynomial  $f$ , preventing a  $O(\text{polylog}(n))$  runtime. One method of circumventing this problem is to use *sparse irreducible polynomials*. In particular, we use the following well-known construction.

**Lemma 3.2.6** ([29], Theorem 1.1.28). *Over  $\mathbb{F}_2[x]$ , the polynomial*

$$f(x) = x^{2 \cdot 3^l} + x^{3^l} + 1$$

*is irreducible for any  $l \geq 0$ .*

Observe that this allows for much faster arithmetic in the extension field, and this leads to the following result when applying the above results.

**Corollary 3.2.7.** *Over the field  $\mathbb{F}_2$ , there is a family of tensors  $T_n$  of size  $[n]^3$  defined for  $n = 2 \cdot 3^l$ , such that*

1.  $\text{rank}(T_n) \geq 3n - \Theta(\lg n)$

2. On inputs  $n$  and  $(i, j, k) \in [n]^3$ ,  $T_n(i, j, k)$  is computable in  $O(\text{polylog}(n))$

Thus, this algebraic construction is also explicit, at least for some values of  $n$ . Also, this corollary is not limited to  $\mathbb{F}_2$ . Other constructions [11] are known over some other fields. However, unlike the results of Section 3.1, it is not clear if better lower bounds exist for the tensors in this section. Indeed, we do not at present know non-trivial upper bounds for the tensors given here.

### 3.3 Higher-Order Tensors

In this section we investigate order- $d$  tensors, particularly when  $d$  is odd. It is a folklore result, eg. see [19], that we can always “reshape” a lower-order tensor into a higher-order tensor without decreasing rank. Reshaping an order- $d$  tensor into an order-2 tensor (a matrix) shows that there are explicit order- $d$  tensors with rank  $n^{\lfloor d/2 \rfloor}$ . Other works used reshaping of order- $(2d + 1)$  tensors into order-3 tensors to derive rank lower bounds of  $2n^d - O(n^{d-1})$  for explicit order- $(2d + 1)$  tensors. We also use this idea to derive rank lower bounds of  $2n^d + n - \Theta(d \lg n)$  for explicit order- $(2d + 1)$  tensors, improving the lower order terms in known lower bounds.

We first state our reshaping lemma, keeping in mind that we again now work over an arbitrary field.

**Lemma 3.3.1.** *Let  $T$  be an order-3 tensor of size  $[n^d] \times [n^d] \times [n]$ . Then define the order- $(2d + 1)$  tensor  $T'$  of size  $[n]^{2d+1}$  by*

$$T'(i_1, \dots, i_{2d+1}) = T \left( 1 + \sum_{j=0}^{d-1} (i_{j+1} - 1)n^j, 1 + \sum_{j=0}^{d-1} (i_{j+(d+1)} - 1)n^j, i_{2d+1} \right)$$

*Then  $T'$  has rank at least  $\text{rank}(T)$ .*

*Further, if  $T(\cdot, \cdot, \cdot)$  is computable in time  $f(n)$ , then  $T'(\cdot, \dots, \cdot)$  is computable in time  $O(\text{poly}(d)\text{polylog}(n) + f(n))$ .*

*Proof.* First observe that the map  $(i_1, \dots, i_d) \mapsto 1 + \sum_{j=0}^{d-1} (i_{j+1} - 1)n^j$  is a bijection from  $[n]^d$  to  $[n^d]$  as this is simply the base- $n$  expansion. That this map is computable

in time  $O(\text{poly}(d)\text{polylog}(n))$  establishes the claim about efficiency.

Now consider a rank  $r'$  decomposition of  $T'$

$$T' = \sum_{l=1}^{r'} \vec{v}_{l,1} \otimes \cdots \otimes \vec{v}_{l,2d+1}$$

We now define  $\vec{u}_{l,1} \in \mathbb{F}^{n^d}$ . Via the bijection from above, we can write

$$\vec{u}_{l,1} \left( 1 + \sum_{j=0}^{d-1} (i_{j+1} - 1)n^j \right) = \vec{v}_{l,1}(i_1) \cdots \vec{v}_{l,d}(i_d)$$

and similarly we define  $\vec{u}_{l,2} \in \mathbb{F}^{n^d}$  by

$$\vec{u}_{l,2} \left( 1 + \sum_{j=0}^{d-1} (i_{j+(d+1)} - 1)n^j \right) = \vec{v}_{l,d+1}(i_{d+1}) \cdots \vec{v}_{l,2d}(i_{2d})$$

, and we take  $\vec{u}_{l,3} = \vec{v}_{l,2d+1} \in \mathbb{F}^n$ . Thus we see that

$$T = \sum_{l=1}^{r'} \vec{u}_{l,1} \otimes \vec{u}_{l,2} \otimes \vec{u}_{l,3}$$

by examining the equation pointwise, and thus  $r' \geq \text{rank}(T)$ . The conclusion thus follows when taking  $r' = \text{rank}(T')$ .  $\square$

The above lemma shows that rank lower bounds for low-order tensors extend (weakly) to rank lower bounds of higher-order tensors. We now apply this lemma to the tensor rank lower bounds of Section 3.1. It is possible to do similarly with the results of Section 3.2, but a weaker conclusion would result as those lower bounds are weaker.

**Corollary 3.3.2.** *For every  $d \geq 1$ , there is a family of  $\{0,1\}$ -tensors  $T_n$  of size  $[n]^{2d+1}$  such that  $\text{rank}(T_n) = 2n^d + n - \Theta(d \lg n)$ . Further, given  $n$  and  $i_1, \dots, i_{2d+1}$ ,  $T_n(i_1, \dots, i_{2d+1})$  is computable in polynomial time (that is, in time  $O(\text{poly}(d)\text{polylog}(n))$ ).*

*Proof.* We first observe that the proof of Corollary 3.1.10 extends to give a family of tensors  $T_n$  with size  $[n] \times [n] \times [f(n)]$ , where  $\text{rank}(T_n) = 2n + f(n) - \Theta(\lg n)$ ,

where  $\omega(\lg n) \leq f(n) \leq n$ . Further, these tensors have their entries computed in  $O(\text{polylog}(n))$  time.

Thus, this leads to tensors of size  $[n^d] \times [n^d] \times [n]$  of rank  $2n^d + n - \Theta(d \lg n)$ . By Lemma 3.3.1 we can then see that these tensors can be reshaped into the desired tensors, establishing the claim on the rank as well as the explicitness.  $\square$





# Chapter 4

## Tensor Rank Upper Bounds

### 4.1 Permutation Tensors

One of the most natural families of full-rank matrices are permutation matrices. This section examines a natural generalization of permutation matrices to tensors, which we call permutation tensors. A counting argument (Proposition 4.1.4) shows that there exists order-3 permutation tensors of super-linear rank (over any fixed finite field) as well as order- $d$  permutation tensors with high rank, so it is natural to conjecture that permutation tensors may all have near-maximal rank, just as in the matrix setting. However, we show (Subsection 4.1.2) that this is false: we give tensor rank upper bounds proving that permutation tensors constructed from groups have rank far less than maximal.

We begin with the formal definition of permutation tensors.

**Definition 4.1.1.** Let  $\mathbb{F}$  be a field, and  $T$  be a tensor  $T : [n]^d \rightarrow \mathbb{F}$ .  $T$  is a **permutation tensor** if  $T$  assumes only 0/1 values, and  $T$  has exactly one 1 in each generalized row. (A **generalized row**, sometimes just “row”, is the set of  $n$  inputs to  $T$  resulting from fixing  $d - 1$  of the coordinates, and varying the remaining coordinates).

It is not hard to see that order-2 permutation tensors are permutation matrices: permutation matrices are those 0/1-matrices such that each row and column have exactly one 1.

### 4.1.1 Permutation Tensors: Rank Lower Bounds

We will show that for any  $d$ , there are high rank permutation tensors of order  $d$ . This will follow from counting permutation tensors and applying standard a counting argument. To count order- $d$  permutation tensors we exploit the fact that they are identical to  $(d - 1)$ -ary quasigroups, we we now define.

**Definition 4.1.2.** A function  $f : [n]^k \rightarrow [n]$  is an  $k$ -ary **quasigroup** if the unary function resulting from fixing any  $k - 1$  inputs to any values results in a bijection.

Observe that 2-ary quasigroups are Latin squares. We now show the equivalence to permutation tensors.

**Proposition 4.1.3.** *Let  $T : [n]^d \rightarrow \{0, 1\}$  be an order- $d$  tensor. Then  $T$  is a permutation tensor iff  $\{(i_1, \dots, i_d) : T(i_1, \dots, i_d) = 1\}$  is the graph of an  $(d - 1)$ -ary quasigroup  $f : [n]^{d-1} \rightarrow [n]$ .*

*Proof.*  $\implies$ : As fixing the first  $d - 1$  coordinates of  $T$  leaves exactly one choice for the remaining coordinate so that  $T$  evaluates to 1, this means that the graph of  $T$  defines a function  $f : [n]^{d-1} \rightarrow [n]$ . To see that  $f$  is a  $(d - 1)$ -ary quasigroup, note that if we fix any  $d - 2$  of the first  $d - 1$  coordinates,  $T$  ensures that the resulting tensor is a permutation matrix, and thus the graph of a bijection. This implies that each desired restriction of  $f$  is a bijection, so  $f$  is a  $(d - 1)$ -ary quasigroup.

$\impliedby$ : As  $T$  is defined as the graph of a function, we see that the generalized rows resulting from fixing the first  $d - 1$  coordinates will have exactly one 1 (as a function assumes exactly one value). If we restrict any other  $d - 1$  coordinates, then that  $f$  is a quasigroup implies that the remaining coordinate has exactly one choice such that  $f(i_1, \dots, i_{d-1}) = i_d$  because of  $f$ 's induced bijections. This says that the desired generalized row has exactly one 1. As this is for any generalized row,  $T$  is a permutation tensor.  $\square$

We now cite two lower bounds for the number of quasigroups. The first uses well-known lower bounds on the permanents of doubly-stochastic matrices, but only

applies to 2-ary quasigroups. The second applies to  $k$ -ary quasigroups but is not as strong for  $k = 2$  as the first bound.

**Theorem** ([30]). *The number of  $n \times n$  Latin squares (and thus 2-ary quasigroups) is at least  $(n!)^{2n}/n^{n^2}$ .*

**Theorem** ([18]). *The number of  $k$ -ary quasigroups on  $n$  symbols is at least  $2^{1.5 \lfloor n/3 \rfloor^k}$*

We now combine these results with a standard counting argument.

**Proposition 4.1.4.** *Let  $\mathbb{F}$  be a finite field.*

- *There exist order-3 permutation tensors of rank  $\Omega(n \log_{|\mathbb{F}|} n)$ .*
- *There exist order- $d$  permutation tensors of rank  $\Omega(\frac{1}{d \lg |\mathbb{F}|} (\frac{n}{3})^{d-2})$ .*

*Proof.* We apply a standard counting argument along with the lower bounds for the number of permutation tensors, as given by the above facts.

Over  $\mathbb{F}$ , there are at most  $|\mathbb{F}|^{dn}$  simple tensors of order  $d$ . Thus, there are at most  $|\mathbb{F}|^{dnr}$  tensors  $T : [n]^d \rightarrow \mathbb{F}$  of rank at most  $r$ . Consider the least  $r$  such that every permutation tensor has rank at most  $r$ . Then, we must have that  $|\mathbb{F}|^{dnr}$  bounds the number of permutation tensors. Solving for  $r$  yields the lower bound.

To reach the first claim, we use the standard lower bound of  $n! \geq (n/e)^n$  in application to the above lower bounds on the number of Latin squares, and this shows that the number of order-3 permutation tensors is at least  $(n/e^2)^{n^2}$ . Thus, we have that  $r \geq n/3 \cdot \log_{|\mathbb{F}|}(n/e^2) = \Omega(n \log_{|\mathbb{F}|} n)$ .

To reach the second claim, we see that the above facts imply that the number of order- $d$  permutation tensors is at least  $2^{1.5 \lfloor n/3 \rfloor^{d-1}}$ . This implies that  $r \geq 1/(dn) \cdot 1/\lg |\mathbb{F}| \cdot 1.5 \lfloor n/3 \rfloor^{d-1} = \Omega((n/3)^{d-2}/(d \lg |\mathbb{F}|))$ .  $\square$

Given this proposition, it may seem natural to restrict to searching within permutation tensors for high-rank tensors. However, the next section shows that a natural class of permutation tensors, called group tensors, do not have high-rank.

### 4.1.2 Permutation Tensors: Rank Upper Bounds

In this section we define a class of permutation tensors constructed from finite groups, and show that these tensors have rank far from maximal. We will give two rank upper-bound methods. The first method uses representation theory and accordingly only works where the group has a complete set of irreducible representations (which usually means “large” fields). The second method is based on polynomial interpolation, and while it gives worse upper bounds and only works for finite abelian groups, it gives results over any field. Neither of these methods applies to all finite non-abelian groups over small fields, and the rank of the corresponding tensors is unclear.

**Definition 4.1.5.** Let  $G$  be a finite group (written multiplicatively, with identity  $1_G$ ), and  $\mathbb{F}$  a field. Define the order- $d$  **group tensor**  $T_G^d : G^d \rightarrow \mathbb{F}$  by

$$T_G^d(g_1, \dots, g_d) = \llbracket g_1 \cdots g_d = 1_G \rrbracket$$

We first explore the representation-theory based upper bound. To do so, we first cite relevant facts from representation theory.

**Theorem 4.1.6** ([21]). *Let  $G$  be a finite group and  $\mathbb{F}$  a field. A **representation** of  $G$  is a homomorphism  $\rho : G \rightarrow \mathbb{F}^{d \times d}$ , where  $d$  is the **dimension** of the representation and is denoted  $\dim \rho$ . The **character of a representation**  $\rho$  is a map  $\chi_\rho : G \rightarrow \mathbb{F}$  defined by  $\text{tr} \circ \rho$ , that is, taking the trace of the resulting matrix of the representation.*

*If  $\text{char}(\mathbb{F})$  is coprime to  $|G|$ , and  $\mathbb{F}$  contains  $N$ -th roots of unity, for  $N$  equal to the least common multiple of all of the orders of elements of  $G$ , then there exists a **complete set of irreducible representations**. In particular, for  $c$  denoting the number of conjugacy classes of  $G$ , there is a set of representations  $\rho_1, \dots, \rho_c$  and associated characters such that (among other properties) we have*

1.  $\frac{1}{|G|} \sum_{i=1}^c (\dim \rho_i) \cdot \chi_i(g) = \llbracket g = e \rrbracket$
2.  $\sum_{i=1}^c (\dim \rho_i)^2 = |G|$
3.  $\dim \rho_i$  divides  $|G|$

In particular, for finite abelian groups,  $c = n$  and  $\dim \rho_i = 1$  for all  $i$ .

Notice that property (1) in the above theorem is an instance of the column orthonormality relations of character tables, which follow from the more commonly mentioned row orthonormality relations. We now use these facts to derive upper bounds on the rank of  $T_G^k$  when the conditions to the above theorem hold.

**Theorem 4.1.7.** *Let  $G$  be a finite group,  $d \geq 2$  and  $\mathbb{F}$  a field, such that  $\text{char}(\mathbb{F})$  is coprime to  $|G|$ , and  $\mathbb{F}$  contains  $N$ -th roots of unity, for  $N$  equal to the least common multiple of the orders of elements of  $G$ . Then given the irreducible representations  $\rho_1, \dots, \rho_c$  for  $G$  over  $\mathbb{F}$ , the order- $d$  group tensor has  $|G| \leq \text{rank}_{\mathbb{F}}(T_G^d) \leq \sum_{i=1}^c (\dim \rho_i)^d \leq |G|^{d/2}$ .*

Furthermore, for  $d = 3$ , we have that  $\text{rank}_{\mathbb{F}}(T_G^3) \leq \sum_{i=1}^c \text{MM}(\dim \rho_i) \leq O(|G|^{\omega/2}) = O(|G|^{1.19})$ , where  $\text{MM}(n)$  is the tensor rank of matrix-multiplication and  $\text{MM}(n) = O(n^{\omega})$ .

In particular, for finite abelian groups,  $\text{rank}_{\mathbb{F}}(T_G^k) = |G|$ .

*Proof.*  $\text{rank}_{\mathbb{F}}(T_G^d) \geq |G|$ : This follows from observing that for fixed  $g_3, \dots, g_d$ , the matrix  $T_G^d(\cdot, \cdot, g_3, \dots, g_d)$  is a permutation matrix, and thus its rank (of  $|G|$ ) lower bounds the rank of  $T_G^d$  (over any field). This can also be seen by induction on Corollary 2.2.9.

$G$  abelian  $\implies \text{rank}_{\mathbb{F}}(T_G^k) \leq |G|$ : Theorem 4.1.6 further implies  $\dim \rho_i = 1$  for all irreducible representations of finite abelian groups, which implies  $\text{rank}_{\mathbb{F}}(T_G^d) \leq |G|$  for abelian groups.

$\sum_{i=1}^c (\dim \rho_i)^d \leq |G|^{d/2}$ : Theorem 4.1.6(2) shows that  $\sum_{i=1}^c (\dim \rho_i)^2 = n$ . The claim will follow from proving that for  $d \in \mathbb{R}$ ,  $d \geq 1$  and  $n_i \in \mathbb{R}_{\geq 0}$ , we have  $\sum n_i^d \leq (\sum n_i)^d$ . To see the claim, we note that as we have  $d \geq 2$  in this case, we have that  $\sum ((\dim \rho_i)^2)^{d/2} \leq (\sum (\dim \rho_i)^2)^{d/2} = n^{d/2}$ .

We now show  $\sum n_i^d \leq (\sum n_i)^d$  under the above conditions. We first show that  $(n + m)^d + 0^d = (n + m)^d \geq n^d + m^d$ . Without loss of generality,  $n \geq m$ . Define  $f(x) = (n + x)^d$ . Expanding  $f$  around 0 using Taylor's theorem with exact error term [20] we have that  $(n + x)^d = n^d + f'(z)x$ , for some  $z \in [0, x]$ . Taking this at

$x = m$  we have  $(n+x)^d = n^d + d(n+z)^{d-1}m \geq n^d + d(m+z)^{d-1} \geq n^d + m^d$  as desired (using  $d \geq 1$ ,  $n \geq m$  and  $z \geq 0$ ). Thus, given non-negative  $n_i$  summing to  $n$ , one can iteratively zero out certain  $n_i$  while increasing the sum  $\sum n_i^d$ , until only  $n_1 = n$  and thus  $\sum n_i^d = n^d$ . Thus, this is a bound on the initial sum of  $\sum n_i^d$ .

$\text{rank}_{\mathbb{F}}(T_G^d) \leq \sum_{i=1}^c (\dim \rho_i)^d$ : The result will follow by constructing, for each  $i$ , the order- $d$  tensor

$$T_{\rho_i}^d(g_1, \dots, g_d) = \chi_i(g_1 \cdots g_d)$$

in rank  $(\dim \rho_i)^d$ . Theorem 4.1.6(1) shows that  $T_G^d = \frac{1}{|G|} \sum_{i=1}^c (\dim \rho_i) \cdot T_{\rho_i}^d$  and so distributing the  $\dim \rho_i/|G|$  term inside the simple tensors yields the result (where we crucially use the restriction on the field characteristic).

Thus, all that remains is to show that  $\text{rank}_{\mathbb{F}}(T_{\rho_i}^d) \leq (\dim \rho_i)^d$ . Using the group homomorphism properties of the representations and expanding the definition of the trace through the matrix multiplication we see

$$T_{\rho_i}^d(g_1, \dots, g_d) = \sum_{k_1=1}^{\dim \rho_i} \cdots \sum_{k_d=1}^{\dim \rho_i} (\rho_i(g_1))_{k_1, k_2} \cdots (\rho_i(g_{d-1}))_{k_{d-1}, k_d} (\rho_i(g_d))_{k_d, k_1}$$

and one can observe that for fixed  $k_1, \dots, k_d$ , the function

$$(\rho_i(g_1))_{k_1, k_2} \cdots (\rho_i(g_{d-1}))_{k_{d-1}, k_d} (\rho_i(g_d))_{k_d, k_1}$$

is a simple tensor so the above shows  $\text{rank}_{\mathbb{F}}(T_{\rho_i}^d) \leq (\dim \rho_i)^d$  as desired.

$\text{rank}_{\mathbb{F}}(T_G^3) \leq \sum_{i=1}^c \text{MM}(\dim \rho_i)$ : We now prove that  $\text{rank}_{\mathbb{F}}(T_G^3) \leq \sum_{i=1}^c \text{MM}(\dim \rho_i)$ , where  $\text{MM}(n)$  is the tensor-rank of matrix-multiplication of  $n \times n$  matrices (which is defined below). By the above, it is sufficient to show that  $\text{rank}_{\mathbb{F}}(T_{\rho_i}^3) \leq \text{MM}(\dim \rho_i)$ . To do this, we start with the observation that  $T_{\rho_i}^3 = \text{tr}(\rho_i(g_1)\rho_i(g_2)\rho_i(g_3))$ .

Now, we observe that for  $n \times n$  matrices  $X, Y, Z$ , the polynomial  $\text{tr}(XYZ^T) = \sum_{i,j,k=1}^{n,n,n} X_{i,j}Y_{j,k}Z_{i,k}$  exactly corresponds to the matrix multiplication tensor. That is, the tensor  $M : [n^2]^3 \rightarrow \{0, 1\}$  defined by  $M((i_1, i_2), (j_1, j_2), (k_1, k_2)) = \llbracket i_1 = k_1 \rrbracket \llbracket i_2 = j_1 \rrbracket \llbracket j_2 = k_2 \rrbracket$  is in natural correspondence with the bilinear form of matrix multiplication (see [9] for reference), and the correspondence implies that  $\text{rank}_{\mathbb{F}}(M) = \text{MM}(n)$ .

Thus, suppose we have a rank  $\text{MM}(n)$  decomposition, so  $M = \sum_{l=1}^{\text{MM}(n)} \vec{a}_l \otimes \vec{b}_l \otimes \vec{c}_l$ , for  $\vec{a}_l, \vec{b}_l, \vec{c}_l \in \mathbb{F}^{n^2}$ . Then we have that  $\text{tr}(XYZ^T) = \sum_{l=1}^{\text{MM}(n)} \langle \vec{a}_l, X \rangle \langle \vec{b}_l, Y \rangle \langle \vec{c}_l, Z^T \rangle$ , where we take the inner products as the natural inner products over the space  $\mathbb{F}^{n^2}$ .

Applying this to computing  $T_{\rho_i}^3$ , we see

$$T_{\rho_i}^3(g_1, g_2, g_3) = \sum_{l=1}^{\text{MM}(\dim(\rho_i))} \langle \vec{a}_l, \rho_i(g_1) \rangle \langle \vec{b}_l, \rho_i(g_2) \rangle \langle \vec{c}_l, \rho_i(g_3)^T \rangle$$

which is a rank  $\text{MM}(\dim(\rho_i))$  computation of  $T_{\rho_i}^3$  as each summand is a simple tensor. Thus,  $\text{rank}_{\mathbb{F}}(T_{\rho_i}^3) \leq \text{MM}(\dim(\rho_i))$ . By the previous work in this proof, the current claim follows.

$\sum_{i=1}^c \text{MM}(\dim \rho_i) \leq O(|G|^{\omega/2}) = O(|G|^{1.19})$ : We note that  $\text{MM}(n) \leq O(n^{2.376})$  follows from Coppersmith-Winograd's result [10], and that in general any exact algorithm for matrix-multiplication will result in the same (up to constant factors) complexity tensor rank for the matrix multiplication tensor, as per the connection between bilinear computation and tensor rank (again, see [9]). This proves the second inequality.

To prove that  $\sum_{i=1}^c \text{MM}(\dim \rho_i) \leq O(|G|^{\omega/2})$ , we use the above proof that  $\sum n_i^d \leq (\sum n_i)^d$  for any  $d \geq 1$  and  $n_i \geq 0$ , while noting that  $\omega \geq 2$ . In particular, there is some constant  $C$  such that  $\text{MM}(\dim \rho_i) \leq C(\dim \rho_i)^\omega$  by assumption. Thus we have that  $\sum_{i=1}^c \text{MM}(\dim \rho_i) \leq \sum_{i=1}^c C((\dim \rho_i)^2)^{\omega/2} \leq C(\sum_{i=1}^c (\dim \rho_i)^2)^{\omega/2} = O(|G|^{\omega/2})$ .  $\square$

We remark that it is not clear whether fast matrix multiplication algorithms can improve the above results for  $d > 3$ .

Applying tensor rank lower bounds to Raz's [19] result requires order- $d$  tensors of rank  $n^{(1-o(1))d}$ , and Theorem 4.1.7 shows that no group tensor can achieve this rank over "large" fields. In particular, for the purposes of tensor rank lower bounds, the lower bounds of Corollary 3.3.2 are asymptotically (in  $d$ ) as good as the rank achievable by any group tensor (over "large" fields).

Further, Theorem 4.1.7 also possibly rules out group tensors as candidates for high rank order-3 tensors (and thus cannot be used for Strassen's result). That is, it

is conjectured that  $\text{MM}(n) \leq \tilde{O}(n^2)$ , which would imply that  $\text{rank}_{\mathbb{F}}(T_G^3) \leq \tilde{O}(|G|)$  for any group  $G$  over a “large” field  $\mathbb{F}$ .

Theorem 4.1.7 only works over “large” fields in general. In particular, it does not (in general) give insight into the rank of group tensors over fixed finite fields, or even over the rationals. To take an example, the cyclic group  $\mathbb{Z}_n$  requires  $n$ -th roots of unity for its irreducible representations. While Lemma 4.1.15 does show a relation between  $\text{rank}_{\mathbb{Q}}(T_{\mathbb{Z}_n}^d)$  and  $\text{rank}_{\mathbb{Q}[x]/\langle x^n - 1 \rangle}(T_{\mathbb{Z}_n}^d)$  (where  $\mathbb{Q}[x]/\langle x^n - 1 \rangle$  is the field of rationals adjoined with a  $n$ -th primitive root of unity, so Theorem 4.1.7 applies) this relationship implies nothing beyond trivial rank upper bounds. Thus, to achieve rank upper bounds for group tensors over small fields we take a different approach, one using polynomial interpolation. Our result only applies to finite abelian groups, but is able to show that have “low” rank in this regime.

**Proposition 4.1.8.** *Let  $\mathbb{F}$  be a field with at least  $d(n - 1) + 1$  elements. Let  $T : \llbracket n \rrbracket^d \rightarrow \mathbb{F}$  be a tensor such that*

$$T(i_1, \dots, i_d) = \sum_{m=0}^{d(n-1)} c_m \llbracket i_1 + i_2 + \dots + i_d = m \rrbracket$$

for constants  $c_m \in \mathbb{F}$ . Then,  $\text{rank}(T) \leq d(n - 1) + 1$ .

*Proof.* The proof follows the result of Ben-Or (as reported in Shpilka-Wigderson [24]) on computing the symmetric polynomials efficiently over large-sized fields. To compute a desired polynomial  $f(\vec{x})$ , one can introduce a new variable  $\alpha$  and an auxiliary polynomial  $P(\alpha, \vec{x})$  such that

- $P$  is efficiently computable, of degree at most  $d'$  in  $\alpha$
- For some  $m$ ,  $f(\vec{x}) = C_{\alpha^m}(P(\alpha, \vec{x}))$ . That is,  $f$  equals the coefficient of  $\alpha^m$  in  $P$ .

To compute  $f$  on input  $\vec{x}$ , we can then evaluate  $P$  on  $(\alpha_1, \vec{x}), \dots, (\alpha_{d'+1}, \vec{x})$  and then use interpolation to recover  $C_{\alpha^m} = f(x)$ .

To apply Ben-Or’s interpolation idea to tensors, we first note the connection



between tensors and polynomials. Consider the space of polynomials

$$\mathcal{P} := \mathbb{F}[\{X_i^{(1)}\}_{i=0}^{n-1}, \dots, \{X_i^{(d)}\}_{i=0}^{n-1}]$$

, that is, polynomials on the variables  $X_i^{(j)}$  that are set-multilinear with respect to the sets  $\{X_i^{(j)}\}_{i=0}^{n-1}$ . One can call such a polynomial *simple* if it can be written as

$$\prod_{j=1}^d \left( a_{j,0} X_0^{(j)} + \dots + a_{j,n-1} X_{n-1}^{(j)} \right)$$

One can then define the  $\text{rank}(p)$ , for  $p \in \mathcal{P}$ , as the least number of simple polynomials needed to sum to  $p$ . One can observe that  $\mathcal{P}$  is a  $[n]^d$  tensor product space, and the notions of rank coincide. In this language, we seek to upper-bound the rank of the polynomial

$$T(\{X_i^{(j)}\}_{i,j}) = \sum_{m=0}^{d(n-1)} c_m \sum_{i_1 + \dots + i_d = m} X_{i_1}^{(1)} \dots X_{i_d}^{(d)}$$

To implement Ben-Or's method, define the auxiliary polynomial  $P$  by

$$P(\alpha, \{X_j^{(i)}\}_{i,j}) := \prod_{i=1}^d \left( X_0^{(i)} + \alpha X_1^{(i)} + \alpha^2 X_2^{(i)} + \dots + \alpha^{n-1} X_{n-1}^{(i)} \right)$$

For fixed  $\alpha$ , this polynomial is simple. When  $\alpha$  is considered a variable, this polynomial has degree  $d(n-1)$  in  $\alpha$ . Further, the coefficient of  $\alpha^m$  is

$$C_{\alpha^m}(P(\alpha, \{X_i^{(j)}\}_{i,j})) = \sum_{i_1 + \dots + i_d = m} X_{i_1}^{(1)} \dots X_{i_d}^{(d)}$$

which corresponds exactly to tensors of the desired form. We can now interpret the auxiliary polynomials as polynomials in  $\mathbb{F}(\{X_i^{(j)}\}_{i,j})[\alpha]$ , the polynomial ring in the variable  $\alpha$  over the field of rational functions in the  $X_i^{(j)}$ . As  $|\mathbb{F}| > d(n-1)$ , we can consider the evaluations  $\{P(\alpha_l, \{X_i^{(j)}\}_{i,j})\}_{l=0}^{d(n-1)} \in \mathbb{F}[\{X_i^{(j)}\}_{i,j}]$  for distinct  $\alpha_l \in \mathbb{F}$ . Polynomial interpolation means that the coefficients  $C_{\alpha^m}(P)$  are recoverable from linear combinations of the evaluations of  $P$ . As the  $\alpha_l \in \mathbb{F}$ , the (linear) evaluation map from the coefficients of  $P$  to the evaluations is defined by a  $\mathbb{F}$ -matrix. Therefore, the

inverse of this map is also defined by an  $\mathbb{F}$ -matrix. Specifically, there are coefficients  $a_{m,l} \in \mathbb{F}$  such that

$$C_{\alpha^m}(P(\alpha, \{X_i^{(j)}\}_{i,j})) = \sum_{l=0}^{d(n-1)} a_{m,l} P(\alpha_l, \{X_i^{(j)}\}_{i,j})$$

Therefore,

$$\begin{aligned} T(\{X_i^{(j)}\}_{i,j}) &= \sum_{m=0}^{d(n-1)} c_m \sum_{l=0}^{d(n-1)} a_{m,l} P(\alpha_l, \{X_i^{(j)}\}_{i,j}) \\ &= \sum_{l=0}^{d(n-1)} \left( \sum_{m=0}^{d(n-1)} c_m a_{m,l} \right) P(\alpha_l, \{X_i^{(j)}\}_{i,j}) \end{aligned}$$

Thus,  $T$  is in the span of  $d(n-1) + 1$  simple polynomials. By moving the coefficients on the simple polynomials inside the product, this shows that  $T$  is expressible as the sum of  $d(n-1) + 1$  simple polynomials. Using the above connection with tensors, this shows that the rank is at most  $d(n-1) + 1$ .  $\square$

We now prove that Proposition 4.1.8 is essentially tight. The proof uses Corollary 2.3.2.

**Proposition 4.1.9.** *Let  $\mathbb{F}$  be a field. Let  $T$  be a tensor  $T : \llbracket n \rrbracket^d \rightarrow \mathbb{F}$  such that*

$$T(i_1, \dots, i_d) = \begin{cases} 1 & \text{if } i_1 + \dots + i_d = n \\ 0 & \text{if } i_1 + \dots + i_d > n \\ \text{unconstrained} & \text{else} \end{cases}$$

then  $\text{rank}_{\mathbb{F}}(T) \geq (d-1)(n-1) + 1$ .

*Proof.* The proof is by induction on  $d$ , using Corollary 2.3.2 to achieve a lower bound.

$d = 1$ : As  $T \neq 0$ , its rank must be at least 1, so the result follows.

$d \geq 2$ : Decompose  $T$  into layers along the  $d$ -th axis, so that  $T = [T_0] \cdots [T_{n-1}]$ . Observe that the hypothesis on  $T$  implies that for any linear combination  $S = \sum_{k=0}^{n-1} c_k T_k$  it must be that  $S \neq 0$ . For if not, one may consider the smallest  $i$  such that  $c_i \neq 0$ .

Then  $S(n-i, 0, \dots, 0) = c_i$  by the hypothesis on  $T$  and the construction of  $S$ , which is a contradiction as  $c_i \neq 0$ .

Thus, Corollary 2.3.2 implies that  $\text{rank}_{\mathbb{F}}(T) \geq \text{rank}_{\mathbb{F}}(T_0 + \sum_{i=1}^{n-1} a_i T_i) + (n-1)$  for some  $a_i \in \mathbb{F}$ . However, observing that  $T' = T_0 + \sum_{i=1}^{n-1} a_i T_i$  is an order- $(d-1)$  tensor fitting the hypothesis of the induction we see that  $\text{rank}_{\mathbb{F}}(T') \geq (d-2)(n-1) + 1$ . Combining the above equations finishes the induction.  $\square$

The above proposition shows that Proposition 4.1.8 is nearly tight, because together they show that defining  $T : \llbracket n \rrbracket^d \rightarrow \mathbb{F}$  by  $T(i_1, \dots, i_d) = \llbracket i_1 + \dots + i_d = n \rrbracket$ , we see that  $(d-1)(n-1) + 1 \leq \text{rank}_{\mathbb{F}}(T) \leq d(n-1) + 1$ .

We now turn to using Proposition 4.1.8 to upper bound the rank group tensors formed from cyclic groups.

**Corollary 4.1.10.** *Let  $\mathbb{F}$  be a field with at least  $d(n-1) + 1$  elements. Then,  $\text{rank}(T_{\mathbb{Z}_n}^d) \leq d(n-1) + 1$ .*

*Proof.* Notice that  $T_{\mathbb{Z}_n}^d$  is a tensor fitting the hypothesis of Proposition 4.1.8, where  $c_m = \llbracket m \equiv 0 \pmod{n} \rrbracket$ .  $\square$

Proposition 4.1.8 was done by interpolating a univariate polynomial. By interpolating multivariate polynomials one may obtain an upper bound for the rank over group tensors arising from the direct product of cyclic groups. However, the same result is derivable in a more modular fashion, which we now present. We start with the folklore fact that tensoring two tensors multiplies their rank bounds.

**Lemma 4.1.11.** *Let  $\mathbb{F}$  be a field. Let  $T : [n]^d \rightarrow \mathbb{F}$  and  $S : [m]^d \rightarrow \mathbb{F}$  be two tensors. Define  $(T \otimes S) : ([n] \times [m])^d \rightarrow \mathbb{F}$  by*

$$(T \otimes S)((i_1, i'_1), \dots, (i_d, i'_d)) = T(i_1, \dots, i_d) \cdot S(i'_1, \dots, i'_d)$$

*Then  $\text{rank}_{\mathbb{F}}(T \otimes S) \leq \text{rank}_{\mathbb{F}}(T) \text{rank}_{\mathbb{F}}(S)$ .*

*Proof.* Suppose  $T = \sum_{l=1}^r \otimes_{j=1}^d \vec{a}_{j,l}$  and  $S = \sum_{l'=1}^{r'} \otimes_{j'=1}^d \vec{a}'_{j',l'}$ . Then  $T(i_1, \dots, i_d) = \sum_l \prod_j \vec{a}_{j,l}(i_j)$  and  $S(i'_1, \dots, i'_d) = \sum_{l'} \prod_{j'} \vec{a}'_{j',l'}(i'_{j'})$ . Thus,  $(T \otimes S)((i_1, i'_1), \dots, (i_d, i'_d))$

equals  $\sum_{l,l'} \prod_{j,j'} \vec{a}_{j,l}(i_j) \vec{a}'_{j',l'}(i'_{j'}) = \sum_{l,l'} \prod_{(j,j')} (\vec{a}_{j,l} \otimes \vec{a}'_{j',l'})_{i_j, i'_{j'}}$ . Thus, as for fixed  $l, l'$  the tensor  $\prod_{(j,j')} (\vec{a}_{j,l} \otimes \vec{a}'_{j',l'})_{i_j, i'_{j'}}$  is simple (as a  $([n] \times [m])^d \rightarrow \mathbb{F}$  tensor), this shows the claim.  $\square$

We now apply this to the direct product construction of groups.

**Corollary 4.1.12.** *Consider integers  $n_1, \dots, n_m \in \mathbb{Z}_{\geq 2}$  and consider the finite abelian group  $G = \mathbb{Z}_{n_1} \times \dots \times \mathbb{Z}_{n_m}$ . Let  $\mathbb{F}$  be a field with at least  $\max_i(d(n_i - 1) + 1)$  elements. Then,  $\text{rank}(T_G^d) \leq \prod_i (d(n_i - 1) + 1)$ .*

*Proof.* First observe the relevant definitions imply that  $T_{\mathbb{Z}_n}^d \otimes T_{\mathbb{Z}_m}^d = T_{\mathbb{Z}_n \times \mathbb{Z}_m}^d$ . Thus, the claim follows directly from Corollary 4.1.10 and Lemma 4.1.11.  $\square$

We now recall the Structure Theorem of Finite Abelian Groups.

**Theorem 4.1.13** (Structure Theorem of Finite Abelian Groups (see, e.g. [3])). *Let  $G$  be a finite abelian group. Then there are (not necessarily distinct) prime powers  $n_1, \dots, n_m \in \mathbb{Z}_{\geq 2}$  such that  $G = \mathbb{Z}_{n_1} \times \dots \times \mathbb{Z}_{n_m}$ .*

This theorem shows that Corollary 4.1.12 extends to general groups. One can get better bounds if more information is known about the group, or if results such as Theorem 4.1.7 apply, but the next result shows that even without such information group tensors from finite abelian groups have “low” rank.

Using the Structure Theorem of Abelian Groups the following can now be shown.

**Corollary 4.1.14.** *Let  $G$  be a finite abelian group, and  $\mathbb{F}$  be a field with at least  $|G|$  elements. Then  $\text{rank}_{\mathbb{F}}(T_G^d) \leq |G|^{1+\lg d}$ .*

*Proof.* Observe that using the Structure Theorem of Finite Abelian groups, we can apply Corollary 4.1.12 to  $G = \prod_{i=1}^m \mathbb{Z}_{n_i}$ , and using that  $d(n_i - 1) + 1 \leq dn$  (as  $d \geq 2$ ) shows that  $\text{rank}_{\mathbb{F}}(T_G^d) \leq d^m \prod n_i$ . As  $m \leq \lg |G|$  and  $|G| = \prod n_i$ , the result follows.  $\square$

While all of the results based on Proposition 4.1.8 do not require the field to have large roots of unity, they still require the field to have large size. Thus, they seemingly

do not answer the question of the rank of group tensors over small fields. However, as the next lemma shows, one can transfer results over large-sized fields to small-sized fields with a minor overhead.

**Lemma 4.1.15.** *Let  $\mathbb{K}$  a field that extends  $\mathbb{F}$ . Then for any tensor  $T : [n]^d \rightarrow \mathbb{F}$ ,  $\text{rank}_{\mathbb{F}}(T) \leq (\dim_{\mathbb{F}} \mathbb{K})^{d-1} \cdot \text{rank}_{\mathbb{K}}(T)$ , where  $\dim_{\mathbb{F}} \mathbb{K}$  is the dimension of  $\mathbb{K}$  as an  $\mathbb{F}$ -vector space.*

*Proof.* Define  $m := \dim_{\mathbb{F}} \mathbb{K}$ . Thus, we can identify  $\mathbb{K} = \mathbb{F}^m$  as vector spaces, where we choose that  $1 \in \mathbb{K}$  is the first element in the  $\mathbb{F}$ -basis for  $\mathbb{K}$ . This gives an  $\mathbb{F}$ -algebra isomorphism  $\mu$  between  $\mathbb{K}$  and a sub-ring  $M$  of the  $m \times m$   $\mathbb{F}$ -matrices, where  $M$  is defined as the image of  $\mu$ . The map  $\mu$  is defined by associating  $x \in \mathbb{K}$  with the matrix inducing the linear map  $\mu(x) : \mathbb{F}^m \rightarrow \mathbb{F}^m$ , where  $\mu(x)$  is the multiplication map of  $x$ . That is, using that  $\mathbb{K} = \mathbb{F}^m$  we can see that the map  $y \mapsto xy$  for  $y \in \mathbb{K}$  is an  $\mathbb{F}$ -linear map, and thus defines  $\mu(x)$  over  $\mathbb{F}^m$ .

That the map is injective follows from the fact that  $\mu(x)$  must map  $1 \in \mathbb{K}$  to  $x \in \mathbb{K}$ , so  $x$  is recoverable from  $\mu(x)$  (and surjectivity follows by definition of  $M$ ). To see the required homomorphism properties is also not difficult. As  $(x+z)y = xy + zy$  for any  $x, y, z \in \mathbb{K}$ , this shows that  $\mu(x+z) = \mu(x) + \mu(z)$  as linear maps, and thus as matrices. Similarly, as  $(xz)y = x(zy)$  for any  $x, y, z \in \mathbb{K}$  it must be that  $\mu(xz) = \mu(x)\mu(z)$ . That this map interacts linearly in  $\mathbb{F}$  implies that it is an  $\mathbb{F}$ -algebra homomorphism, as desired.

Now consider a tensor  $T : [n]^d \rightarrow \mathbb{F}$  with simple tensor decomposition  $T = \sum_{l=1}^{\text{rank}_{\mathbb{K}}(T)} \otimes_{j=1}^d \vec{a}_{j,l}$  over  $\mathbb{K}$ . First observe that if we define the map  $\pi : \mathbb{K} \rightarrow \mathbb{F}$  defined by

$$\pi(x) = \begin{cases} x & \text{if } x \in \mathbb{F} \\ 0 & \text{else} \end{cases}$$

then  $T = \sum_{l=1}^{\text{rank}_{\mathbb{K}}(T)} \pi(\otimes_{j=1}^d \vec{a}_{j,l})$ . Thus for each  $l$ ,  $\pi(\otimes_{j=1}^d \vec{a}_{j,l})$  is a tensor  $T_l : [n]^d \rightarrow \mathbb{F}$ .

We now show that  $\text{rank}_{\mathbb{F}}(T_l) \leq m^{d-1}$ . First observe that for  $x \in \mathbb{F}$ ,  $\mu(x)$  is a diagonal matrix. In particular, because we chose  $1 \in \mathbb{K}$  to be the first element in the  $\mathbb{F}$ -basis for  $\mathbb{K}$ , for  $x \in \mathbb{F}$ ,  $\pi(x)$  is equal to the  $(1,1)$ -th entry in  $\mu(x)$ . Thus, it

follows that  $T_l(i_1, \dots, i_d) = (\mu(\vec{a}_{1,l}(i_1)) \cdots \mu(\vec{a}_{d,l}(i_d)))_{1,1}$ . By expanding out the matrix multiplication we can see that  $T_l$  is expressible as

$$T_l(i_1, \dots, i_d) = \sum_{k_1=1}^m \cdots \sum_{k_{d-1}=1}^m \mu(\vec{a}_{1,l}(i_1))_{1,k_1} \cdot \mu(\vec{a}_{2,l}(i_2))_{k_1,k_2} \cdots \mu(\vec{a}_{d,l}(i_d))_{k_{d-1},1}$$

and just as in Theorem 4.1.7 we see that for fixed  $k_j$  the summands are simple  $\mathbb{F}$ -tensors, and thus  $\text{rank}_{\mathbb{F}}(T_l) \leq m^{d-1}$ .

Using the observation that  $T = \sum_{l=1}^{\text{rank}_{\mathbb{K}}(T)} T_l$  and the above bound for the  $\mathbb{F}$ -rank of  $T_l$ , we then see that  $\text{rank}_{\mathbb{F}}(T) \leq (\dim_{\mathbb{F}} \mathbb{K})^{d-1} \text{rank}_{\mathbb{K}}(T)$ , as desired.  $\square$

With this field-transfer result, we can now state rank upper bounds for group tensors (for finite abelian groups) for any field.

**Corollary 4.1.16.** *Let  $\mathbb{F}$  be any field, and  $G$  be a finite abelian group. Then  $\text{rank}_{\mathbb{F}}(T_G^d) \leq |G|^{1+\lg d} \lceil \lg |G| \rceil^{d-1}$ .*

*In particular, if  $G$  is cyclic, then  $\text{rank}_{\mathbb{F}}(T_G^d) \leq d|G| \lceil \lg |G| \rceil^{d-1}$ .*

*Proof.* First observe that, if needed, we can extend  $\mathbb{F}$  to a field  $\mathbb{K}$ , such that  $|\mathbb{K}| \geq |G|$  and (as  $|\mathbb{F}| \geq 2$ ) such that  $\dim_{\mathbb{F}} \mathbb{K} \leq \lceil \lg |G| \rceil$ . Thus, applying Lemma 4.1.15 to Corollaries 4.1.10 and 4.1.14 yields the result.  $\square$

This last result shows that for any finite abelian group, any field and any large  $d$ , the rank of the corresponding group tensor is far from possible  $\Omega(n^{d-1})$ . These results do not settle the rank of group tensors for non-abelian groups over small fields, and leaves the open question whether the methods of Theorem 4.1.7 or Proposition 4.1.8 (or other methods) can resolve this case.

# Chapter 5

## Monotone Tensor Rank

We now explore a restricted notion of tensor rank, that of monotone tensor rank. In algebraic models of computation, monotone computation requires that the underlying field is ordered, which we now define.

**Definition 5.0.1.** Let  $\mathbb{F}$  be a field.  $\mathbb{F}$  is **ordered** if there is a linear order  $<$  such that

- For all  $x, y, z \in \mathbb{F}$ ,  $x < y \implies x + z < y + z$ .
- For all  $x, y \in \mathbb{F}$  and  $z \in \mathbb{F}_{>0}$ ,  $x < y \implies xz < yz$ .

where  $\mathbb{F}_{>0} = \{x \mid x \in \mathbb{F}, x > 0\}$ .

Recall that every ordered field has characteristic zero, and thus is infinite.

Over ordered fields, computation of polynomials that only use positive coefficients can be done using only positive field constants, but many works (such as [28]) have shown that the circuit model of computation, the restriction to positive field constants in computation leads drastically worse efficiency as compared to unrestricted computation. In this section, we show that in the tensor rank model of computation, monotone computation is also much less efficient than unrestricted computation. We first define the notion of monotone tensor rank.

**Definition 5.0.2.** Let  $\mathbb{F}$  be an ordered field. Consider a tensor  $T : \prod_{i=1}^d [n_i] \rightarrow \mathbb{F}_{\geq 0}$ .

Define the **monotone tensor rank** of  $T$ , denoted  $\text{m-rank}(T)$ , to be

$$\text{m-rank}(T) = \min \left\{ r : T = \sum_{l=1}^r \vec{v}_{l,1} \otimes \cdots \otimes \vec{v}_{l,d}, \vec{v}_{l,i} \in (\mathbb{F}_{\geq 0})^{n_i} \right\}$$

We now show an essentially maximal separation between monotone tensor rank and unrestricted tensor rank, for the explicit group tensor  $T_{\mathbb{Z}_n}^d$ .

**Theorem 5.0.3.** *Let  $\mathbb{F}$  be an ordered field. Consider the group tensor  $T_{\mathbb{Z}_n}^d$ . Then*

1.  $\text{rank}_{\mathbb{F}}(T_{\mathbb{Z}_n}^d) \leq d(n-1) + 1$

2.  $\text{m-rank}_{\mathbb{F}}(T_{\mathbb{Z}_n}^d) = n^{d-1}$

*Proof.*  $\text{m-rank}_{\mathbb{F}}(T_{\mathbb{Z}_n}^d) \geq n^{d-1}$ : We remark that the following lower bound will only rely on the fact that  $T_{\mathbb{Z}_n}^d$  is a permutation tensor, and no other properties.

In monotone computation, there is no cancellation of terms. Thus, in a monotone simple tensor decomposition  $T = \sum_{l=1}^r T_l$ , one can see that the partial sums  $T_{\leq m} = \sum_{l=1}^m T_l$  successively cover more and more of the non-zero entries of  $T$ . We will show that in any monotone decomposition of  $T_{\mathbb{Z}_n}^d$ , at most one non-zero entry can be covered by any  $T_l$ , which implies that the monotone rank is at least the number of non-zero entries, which is  $n^{d-1}$ .

We now prove that in any monotone simple tensor decomposition  $T_{\mathbb{Z}_n}^d = \sum_{l=1}^r \otimes_{j=1}^d \vec{a}_{j,l}$ , each simple tensor  $T_l := \otimes_{j=1}^d \vec{a}_{j,l}$  can cover at most one non-zero entry of  $T_{\mathbb{Z}_n}^d$ . Suppose not, for contradiction. Then there is a simple tensor  $T_l$  that covers at least two non-zero entries  $(i_1, \dots, i_d)$  and  $(i'_1, \dots, i'_d)$  of  $T_{\mathbb{Z}_n}^d$ . However, these tuples must differ in at least one index, we we assume without loss of generality to be index 1, so that  $i_1 \neq i'_1$ . Consequently, we must have that  $\vec{a}_{1,l}(i_1), \vec{a}_{1,l}(i'_1) > 0$  (as all field constants are positive in monotone computation). As  $\vec{a}_{j,l}(i_j) > 0$  for  $j > 1$  (as  $T_l(i_1, i_2, \dots, i_d) = 1$ ), it must be that  $T_l(i'_1, i_2, \dots, i_d) = \vec{a}_{1,l}(i'_1) \prod_{j>1}^d \vec{a}_{j,l}(i_j) > 0$ . However, this is a contradiction. For now this positive number at  $T_l(i'_1, i_2, \dots, i_d)$  cannot be canceled out by other simple tensors in a monotone computation and we must have  $T_{\mathbb{Z}_n}^d(i'_1, i_2, \dots, i_d) = 0$  by the fact that this is a permutation tensor. Thus, it must be that each simple tensor



in this monotone computation can only cover a single non-zero entry of  $T_{\mathbb{Z}_n}^d$ , which implies the lower bound by the above argument.  $\square$



# Bibliography

- [1] Boris Alexeev, Michael Forbes, and Jacob Tsimmerman. Tensor rank: Some lower and upper bounds. *Electronic Colloquium on Computational Complexity (ECCC)*, 18:10, 2011. Available from: <http://eccc.hpi-web.de/report/2011/010>.
- [2] Boris Alexeev, Michael Forbes, and Jacob Tsimmerman. Tensor rank: Some lower and upper bounds. In *IEEE Conference on Computational Complexity*, 2011.
- [3] M. Artin. *Algebra*. Prentice Hall Inc., Englewood Cliffs, NJ, 1991.
- [4] Walter Baur and Volker Strassen. The complexity of partial derivatives. *Theoret. Comput. Sci.*, 22(3):317–330, 1983. Available from: [http://dx.doi.org/10.1016/0304-3975\(83\)90110-X](http://dx.doi.org/10.1016/0304-3975(83)90110-X), doi:10.1016/0304-3975(83)90110-X.
- [5] Markus Bläser. A  $\frac{5}{2}n^2$ -lower bound for the rank of  $n \times n$ -matrix multiplication over arbitrary fields. In *40th Annual Symposium on Foundations of Computer Science (New York, 1999)*, pages 45–50. IEEE Computer Soc., Los Alamitos, CA, 1999. Available from: <http://dx.doi.org/10.1109/SFFCS.1999.814576>, doi:10.1109/SFFCS.1999.814576.
- [6] Markus Bläser. Improvements of the Alder-Strassen bound: algebras with nonzero radical. In *Automata, languages and programming*, volume 2076 of *Lecture Notes in Comput. Sci.*, pages 79–91. Springer, Berlin, 2001. Available from: [http://dx.doi.org/10.1007/3-540-48224-5\\_7](http://dx.doi.org/10.1007/3-540-48224-5_7), doi:10.1007/3-540-48224-5\_7.
- [7] Markus Bläser. A note on tensor rank. Personal Communication, 2011.
- [8] Mark R. Brown and David P. Dobkin. An improved lower bound on polynomial multiplication. *IEEE Trans. Comput.*, 29(5):337–340, 1980. Available from: <http://dx.doi.org/10.1109/TC.1980.1675583>, doi:10.1109/TC.1980.1675583.
- [9] Peter Bürgisser, Michael Clausen, and M. Amin Shokrollahi. *Algebraic complexity theory*, volume 315 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1997. With the collaboration of Thomas Lickteig.

- [10] Don Coppersmith and Shmuel Winograd. Matrix multiplication via arithmetic progressions. *J. Symbolic Comput.*, 9(3):251–280, 1990. doi:10.1016/S0747-7171(08)80013-2.
- [11] Shuhong Gao and Daniel Panario. Tests and constructions of irreducible polynomials over finite fields. In *Foundations of computational mathematics (Rio de Janeiro, 1997)*, pages 346–361. Springer, Berlin, 1997.
- [12] J. Håstad. Tensor rank is NP-complete. In *ICALP '89: Proceedings of the 16th International Colloquium on Automata, Languages and Programming*, volume 372 of *Lecture Notes in Comput. Sci.*, pages 451–460. Springer, Berlin, 1989. Available from: <http://dx.doi.org/10.1007/BFb0035776>, <http://dx.doi.org/10.1007/BFb0035776> doi:10.1007/BFb0035776.
- [13] J. Håstad. Tensor rank is NP-complete. *J. Algorithms*, 11(4):644–654, 1990. doi:10.1016/0196-6774(90)90014-6.
- [14] J. E. Hopcroft and L. R. Kerr. On minimizing the number of multiplications necessary for matrix multiplication. *SIAM J. Appl. Math.*, 20:30–36, 1971. Available from: <http://www.jstor.org/stable/2099880>.
- [15] Joseph Ja'Ja'. Optimal evaluation of pairs of bilinear forms. *SIAM J. Comput.*, 8(3):443–462, 1979. Available from: <http://dx.doi.org/10.1137/0208037>, <http://dx.doi.org/10.1137/0208037> doi:10.1137/0208037.
- [16] K. A. Kalorkoti. A lower bound for the formula size of rational functions. *SIAM J. Comput.*, 14(3):678–687, 1985. Available from: <http://dx.doi.org/10.1137/0214050>, <http://dx.doi.org/10.1137/0214050> doi:10.1137/0214050.
- [17] Michael Kaminski. A lower bound on the complexity of polynomial multiplication over finite fields. *SIAM J. Comput.*, 34(4):960–992 (electronic), 2005. Available from: <http://dx.doi.org/10.1137/S0097539704442118>, doi:10.1137/S0097539704442118.
- [18] Denis S. Krotov, Valdimir N. Potapov, and Polina V. Sokolova. On reconstructing reducible  $n$ -ary quasigroups and switching subquasigroups. *Quasigroups Related Systems*, 16(1):55–67, 2008. arXiv:<http://arxiv.org/abs/math/0608269>.
- [19] R. Raz. Tensor-rank and lower bounds for arithmetic formulas. In *Proceedings of the 42nd ACM symposium on Theory of computing*, STOC '10, pages 659–666, New York, NY, USA, 2010. ACM. Available from: <http://doi.acm.org/10.1145/1806689.1806780>, doi:<http://doi.acm.org/10.1145/1806689.1806780>.
- [20] W. Rudin. *Principles of mathematical analysis*. McGraw-Hill Book Co., New York, third edition, 1976. International Series in Pure and Applied Mathematics.

- [21] Jean-Pierre Serre. *Linear representations of finite groups*. Springer-Verlag, New York, 1977. Translated from the second French edition by Leonard L. Scott, Graduate Texts in Mathematics, Vol. 42.
- [22] V. Shoup. New algorithms for finding irreducible polynomials over finite fields. *Math. Comp.*, 54(189):435–447, 1990. Available from: <http://dx.doi.org/10.2307/2008704>, <http://dx.doi.org/10.2307/2008704> doi:10.2307/2008704.
- [23] Amir Shpilka. Lower bounds for matrix product. *SIAM J. Comput.*, 32(5):1185–1200 (electronic), 2003. Available from: <http://dx.doi.org/10.1137/S0097539702405954>, doi:10.1137/S0097539702405954.
- [24] Amir Shpilka and Avi Wigderson. Depth-3 arithmetic formulae over fields of characteristic zero. *Journal of Computational Complexity*, 10:1–27, 1 2001.
- [25] Volker Strassen. Vermeidung von Divisionen. *J. Reine Angew. Math.*, 264:184–202, 1973.
- [26] Volker Strassen. Die Berechnungskomplexität der symbolischen Differentiation von Interpolationspolynomen. *Theor. Comput. Sci.*, 1(1):21–25, 1975.
- [27] Toshio Sumi, Mitsuhiro Miyazaki, and Toshio Sakata. Rank of 3-tensors with 2 slices and Kronecker canonical forms. *Linear Algebra Appl.*, 431(10):1858–1868, 2009. Available from: <http://dx.doi.org/10.1016/j.laa.2009.06.023>, doi:10.1016/j.laa.2009.06.023.
- [28] L. G. Valiant. Negation can be exponentially powerful. *Theoret. Comput. Sci.*, 12(3):303–314, 1980. Available from: [http://dx.doi.org/10.1016/0304-3975\(80\)90060-2](http://dx.doi.org/10.1016/0304-3975(80)90060-2), doi:10.1016/0304-3975(80)90060-2.
- [29] J. H. van Lint. *Introduction to coding theory*, volume 86 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, third edition, 1999.
- [30] J. H. van Lint and R. M. Wilson. *A course in combinatorics*. Cambridge University Press, Cambridge, second edition, 2001.
- [31] Joachim von zur Gathen. Algebraic complexity theory. In *Annual review of computer science, Vol. 3*, pages 317–347. Annual Reviews, Palo Alto, CA, 1988.
- [32] Benjamin Weitz. An improvement on ranks of explicit tensors. *CoRR*, abs/1102.0580:1–8, 2011. Available from: <http://arxiv.org/abs/1102.0580>.