國立臺灣大學理學院數學系

碩士論文

Department of Mathematics
College of Science
National Taiwan University

Master Thesis

片秩法及其應用

Slice Rank Methods and Their Applications

張哲睿

Che-Jui Chang

指導教授：沈俊嚴 博士

Advisor: Chun-Yen Shen, Ph.D.

中華民國 111 年 06 月

June, 2022

doi:10.6342/NTU202200962

國立臺灣大學碩士學位論文
口試委員會審定書

Slice Rank Methods and Their Applications

本論文係張哲睿君(R09221001)在國立臺灣大學數學系完成之碩士學位論文，於民國111年05月31日承下列考試委員審查通過及口試及格，特此證明

口試委員：

沈俊嚴 (簽名)
(指導教授)

系主任、所長 ______________ (簽名)
(是否須簽章依各院系所規定)

序言

這是我人生中的第一篇論文，同時也是我很引以為傲的全新成果，但我碩士這兩年並非完全順遂，我曾因為懷疑自己能力而有想放棄學術轉去業界的想法，很幸運的我找到自己的方法走過來了，現在的我會認為，求學路上學到最有價值的不只是學術本身，更珍貴的是我學會了如何平衡學業與生活，這才是能讓我繼續在這條路上走更久的重要關鍵。

感謝沈俊嚴老師給我許多非常有用的建議以及想法，同時我也想感謝我的家人不斷地給我許多支持。

2022.06.12 張哲睿

doi:10.6342/NTU202200962

摘要

片秩法 (Slice Rank) 是 Croot, Lev 以及 Pach 於 2016 年提出的一個新的組合數學工具，許多極值組合學當中的問題都透過了這個方法有了新的進展。在這篇論文中，我們會介紹片秩法及其應用，同時我們也會介紹劃分秩法 (Partition Rank) 及其應用。最後我們會利用片秩法以及隨機圖的定理證明直角移除定理。

關鍵詞：片秩法、劃分秩法、極值組合、直角、移除引理。

doi:10.6342/NTU202200962

Abstract

Slice rank methods are new combinatorial tools introduced by Croot, Lev, and Pach [6] in 2016. Many problems in extremal combinatorics are improved by applying the slice rank methods. In this thesis, we'll introduce the slice rank methods and their applications. Moreover, the partition rank and one of its applications are also introduced. Finally, we use slice rank methods and a random graph theorem to prove the right angles removal lemma.

Keywords: *slice rank, partition rank, extremal combinatorics, right angle, removal lemma*

doi:10.6342/NTU202200962

Contents

doi:10.6342/NTU202200962

1 Introduction

1.1 Background

The **slice rank** methods are innovative ideas that improve many results in extremal combinatorics. They are first appeared in the work of Croot, Lev, and Pach [6] in 2016. They use this methods to obtain an exponential upper bound for progression-free subsets in $\mathbb{Z}_4^n$. Later that year, Ellenberg and Gijswijt [8] also used this methods to obtain an exponential upper bound for three-term-progression-free subsets, but this time the sets lie in $\mathbb{F}_q^n$, which is the case for the cap set problem in affine geometry if $q = 3$. Many new applications of this methods are found in these years. Such as finding solutions of certain linear system and finding upper bounds for right-angle-free sets,... etc. It is no doubt that this is one of the powerful methods in extremal combinatorics.

In this thesis, we first introduce the **slice rank of functions** and prove some of its basic properties, as well as two of its applications. See Section 2 for more details. In Section 3 we introduce the **slice rank of tensors**, which is in fact equivalent to that of functions but the basis independent property helps us to prove more advanced properties of the slice rank. In Section 4, we define the **partition rank** of functions, which is similar to the slice rank but it gives better upper bounds than the slice rank gives. An application of it is also provided.

Finally in section 5, we'll use the slice rank method to prove the **right angle removal lemma.** Roughly speaking, it says that if there are not too many right angles in a given set, then one can remove reasonable elements from it so that the remaining set is right-angle-free. See Theorem 5.2.1 for more detail.

1.2 General notations

Throughout this thesis, the symbol $[L]$ denotes the set $\{1, 2, \ldots, L\}$ for any positive integer L. For any set X and any positive integer k, we denote X^k the k-fold Cartesian product of X. For any prime p and any prime power q, we denote $\mathbb{F}_p$ and $\mathbb{F}_q$ to be the finite field with order p and q, respectively. In particular, $\mathbb{F}_q^n$ denotes the n-fold Cartesian product of $\mathbb{F}_q$, which can be considered

doi:10.6342/NTU202200962

as an n dimensional vector space of the field $\mathbb{F}_q$.

2 Slice Rank of Functions and Its Applications

The notion of the slice rank can be defined on functions or tensors. We'll introduce the slice rank of a function first.

2.1 Slice rank of functions

Definition 2.1.1. *Given finite sets $X_1, \dots, X_k$ and a field $\mathbb{F}$. Consider a function*

$$F : X_1 \times \cdots \times X_k \to \mathbb{F}.$$

We say F is of ***slice rank one*** *if it is non-zero and can be written as*

$$F(x_1, \dots, x_k) = f(x_i)g(x_{\hat{i}})$$

for some $i \in [k]$ and functions $f : X_i \to \mathbb{F}$, $g : X_1 \times \cdots X_{i-1} \times X_{i+1} \times \cdots \times X_k \to \mathbb{F}$ and $x_{\hat{i}}$ denotes the $(k-1)$-variables $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_k)$.

For general $F : X_1 \times \cdots \times X_k \to \mathbb{F}$, its ***slice rank****, denoted as* slice-rank(F), *is the least non-negative integer r so that F can be written as a sum of r slice rank one functions.*

Thus, for any non-negative integer r, slice-rank$(F) \leq r$ if and only if

$$F(x_1, \dots, x_k) = \sum_{i=1}^{k} \sum_{\alpha \in S_i} f_{i,\alpha}(x_i) g_{i,\alpha}(x_{\hat{i}})$$

for some $f_{i,\alpha} : X_i \to \mathbb{F}$ and $g_{i,\alpha} : X_1 \times \cdots X_{i-1} \times X_{i+1} \times \cdots \times X_k \to \mathbb{F}$ and the index sets S_i satisfying $|S_1| + \cdots + |S_k| \leq r$.

The reason why it is called the "slice rank" is because the function $F : X_1 \times \cdots \times X_k \to \mathbb{F}$ can be considered as a hypermatrix whose size is $|X_1| \times \cdots \times |X_k|$, and the non-zero function $f(x_i)g(x_{\hat{i}})$ is

doi:10.6342/NTU202200962

of slice rank 1 because its "slice" along the $i-$th dimension is a multiple of $g(x_{\hat{i}})$ as a hypermatrix.
Remark. Under the case of dimension $k = 2$, the slice rank of $F : X_1 \times X_2 \to \mathbb{F}$ equals to the rank of the corresponding $|X_1| \times |X_2|$ matrix. Indeed, from a basic property in linear algebra, the rank of a matrix equals to the dimension of its column space. Using the basis of the column space, we obtain that slice-rank(F) equals to the rank of the matrix form of F.

Similar to the rank of a matrix, the slice rank of a function is bounded above by its sizes of each dimension.

Lemma 2.1.2. *Given a function $F : X_1 \times \cdots \times X_k \to \mathbb{F}$, we have*

$$\text{slice-rank}(F) \leq \min_{1 \leq i \leq k} |X_i|.$$

Proof. For any $i \in [k]$, we can write

$$F(x_1, \ldots, x_k) = \sum_{a \in X_i} \mathbb{1}_{a=x_i} F(x_1, \ldots, x_{i-1}, a, x_{i+1}, \ldots, x_k).$$

Since each $\mathbb{1}_{a=x_i} F(x_1, \ldots, x_{i-1}, a, x_{i+1}, \ldots, x_k)$ is of slice rank 1, we have slice-rank$(F) \leq |X_i|$ for all $i \in [k]$.

□

Recall that in linear algebra, it is well known that the rank of a diagonal matrix equals to the number of non-zero indexes in its diagonal. The following theorem is an analogous result in the slice rank of a function. It turns out that this is one of the most important theorems in slice rank methods.

Theorem 2.1.3. *[31, Lemma 1] Let X be a finite non-empty subset and $k \geq 2$ be integers and let $\mathbb{F}$ be a field. Suppose that $F : X^k \to \mathbb{F}$ is diagonal. That is, suppose F can be written as*

$$F(x_1, \ldots, x_k) = \sum_{a \in A} c_a \mathbb{1}_{a=x_1=x_2=\cdots=x_k}$$

for some $A \subseteq X$ where $c_a \neq 0$ for each $a \in A$. Then

$$\text{slice-rank}(F) = |A|.$$

doi:10.6342/NTU202200962

Proof. Note that we can restrict F on A^k and its slice rank remains the same. By Lemma 2.1.2, we have slice-rank$(F) \leq L$. So it suffices to show slice-rank$(F) \geq L$.

Without loss of generality, set $X = [L]$. We'll use induction on k. The case $k = 2$ follows from the previous Remark. So we may assume $k \geq 3$. By definition of slice rank, we can write

$$F(x_1, \ldots, x_k) = \sum_{i=1}^{k} \sum_{\alpha \in S_i} f_{i,\alpha}(x_i) g_{i,\alpha}(x_{\hat{i}}), \tag{1}$$

for some $f_{i,\alpha} : [L] \to \mathbb{F}$ and $g_{i,\alpha} : [L]^{k-1} \to \mathbb{F}$ and the index sets satisfying $|S_1| + \cdots + |S_k| =$ slice-rank(F). We consider the orthogonal complement H of the vector subspace spanned by the functions $f_{k,\alpha}$, $\alpha \in S_k$ over $\mathbb{F}$. That is,

$$H = \{h : [L] \to \mathbb{F} \mid \sum_{x=1}^{L} f_{k,\alpha}(x) h(x) = 0 \ \forall \alpha \in S_k\}.$$

Then it has dimension at least $L - |S_k|$. So this space must contain an element $h_0 : [L] \to \mathbb{F}$ with $h_0(x) \neq 0$ for at least $L - |S_k|$ values of $x \in [L]$. Multiply by $h_0(x_k)$ and take the sum over $x_k \in [L]$ to (1), we have

$$\sum_{x_k=1}^{L} h_0(x_k) F(x_1, \ldots, x_k) = \sum_{i=1}^{k-1} \sum_{\alpha \in S_i} f_{i,\alpha}(x_i) \tilde{g}_{i,\alpha}(x_{\hat{i,k}}), \tag{2}$$

where $x_{\hat{i,k}}$ denotes $(k-2)$ variables $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{k-1})$ and

$$\tilde{g}_{i,\alpha}(x_{\hat{i,k}}) = \sum_{x_k=1}^{L} h_0(x_k) g_{i,\alpha}(x_{\hat{i}})$$

for each i and α. Since the left hand side of the equation (2) is diagonal with at least $L - |S_k|$ non-zero diagonal entries, its slice rank is at least $L - |S_k|$ by induction hypothesis. On the other hand, the right hand side is the sum of $|S_1| + \cdots + |S_{k-1}|$ slice rank one functions. Hence its slice rank is at most $|S_1| + \cdots + |S_{k-1}|$ by definition. Therefore, we obtain an inequality $L - |S_k| \leq |S_1| + \cdots + |S_{k-1}|$, or equivalently,

$$L \leq |S_1| + \cdots + |S_k| = \text{slice-rank}(F),$$

as desired. □

doi:10.6342/NTU202200962

Using this theorem, we can prove some results in combinatorics. In the next subsection, we'll introduce its application in proving the existence of non-trivial solutions of a given homogeneous linear system.

2.2 Application 1: Cap set problem

In affine geometry, a **cap set** is a subset of $\mathbb{F}_3^n$ which contains no lines, or equivalently no non-trivial arithmetic progressions of length three. We denote $r_3(n)$ to be the largest size of cap sets in $\mathbb{F}_3^n$. The cap set problem is the following:

Question 2.2.1. How large can $r_3(n)$ be?

The trivial bound is $r_3(n) \leq 3^n$. Back in 1987, Frankl Graham, and Rodl [11] showed that $r_3(n) = o(3^n)$. Years later, Meshulam [21] used the Fourier method to improve the bound to $r_3(n) = O(\frac{3^n}{n})$. The proof is a direct use of the ideas of Roth, who is famous for its theorem about the largest progression-free subsets in $\{1, \ldots, n\}$. This bound is then improved to $O(\frac{3^n}{n^{1+c}})$ for some constant $c > 0$ by Bateman and Katz [3]. It turns out that via the slice rank method, we can have an exponential bound $r_3(n) = o(2.756^n)$. This is first proved by Ellenberg and Gijswijt [8].

The proof is short but elegant. It also presents the standard process of the slice rank methods. The key lemma is the following:

Lemma 2.2.2. *[8] Given any non-empty set $A \subseteq \mathbb{F}_3^n$. Define a function $F : A^3 \to \mathbb{F}_3$ by*

$$F(x, y, z) = \begin{cases} 1 & \text{if } x + y + z = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

Then

$$\text{slice-rank}(F) \leq 3 \left(t^{-2/3} + t^{1/3} + t^{4/3}\right)^n,$$

where $t = \frac{-1+\sqrt{33}}{8}$. In particular, we have slice-rank$(F) = o(2.756^n)$.

doi:10.6342/NTU202200962

Proof. Fermat's little theorem states that for all $t \in \mathbb{F}_p$,

$$t^{p-1} = \begin{cases} 1 & \text{if } t \neq 0, \\ 0 & \text{if } t = 0. \end{cases}$$

Using this theorem for $p = 3$, we can rewrite F as

$$F(x, y, z) = \prod_{i=1}^{n} [1 - (x_i + y_i + z_i)^2],$$

where x_i, y_i, z_i denotes the i-th index of the vector x, y, z, respectively. By expanding the product, we can see that F is a linear combinations of

$$x_1^{d_1} \cdots x_n^{d_n} y_1^{e_1} \cdots y_n^{e_n} z_1^{f_1} \cdots z_n^{f_n}$$

for integers $d_i, e_i, f_i \in \{0, 1, 2\}$ satisfying the inequality

$$\sum_{i=1}^{n} (d_i + e_i + f_i) \leq 2n. \tag{4}$$

Thus the slice rank of F is at most the sum of the slice rank of those terms. From the inequality (4), we know that at least one of these three summations $\sum_{i=1}^{n} d_i, \sum_{i=1}^{n} e_i$ and $\sum_{i=1}^{n} f_i$ is at most $2n/3$. Hence F can be written as

$$\begin{aligned} F(x, y, z) = & \sum_{\sum d_i \leq 2n/3} \left[x_1^{d_1} \cdots x_n^{d_n} \right] \alpha_{d_1, \ldots, d_n}(y, z) \\ & + \sum_{\sum e_i \leq 2n/3} [y_1^{e_1} \cdots y_n^{e_n}] \, \beta_{e_1, \ldots, e_n}(x, z) \\ & + \sum_{\sum f_i \leq 2n/3} \left[z_1^{f_1} \cdots z_n^{f_n} \right] \gamma_{f_1, \ldots, f_n}(x, y). \end{aligned}$$

Therefore the slice rank of F is at most

$$3 |\{(d_1, \ldots, d_n) \in \{0, 1, 2\}^n : \sum_{i=1}^{n} d_i \leq 2n/3\}|,$$

doi:10.6342/NTU202200962

which equals to

$$3^{1+n}\mathbb{P}[\sum_{i=1}^{n} d_i \leq \frac{2n}{3}],$$

where $\mathbb{P}$ denotes the probability measure of the uniformly random choices of $(d_1, \ldots, d_n) \in \{0, 1, 2\}^n$. For any $0 < t \leq 1$, we have

$$\begin{aligned}
\text{slice-rank}(F) &\leq 3^{1+n}\mathbb{P}[\sum_{i=1}^{n} d_i \leq 2n/3] \\
&\leq 3^{1+n}\mathbb{P}[t^{d_1+\cdots+d_n} \geq t^{2n/3}] \\
&\leq 3^{1+n}\frac{1}{t^{2n/3}}\mathbb{E}[t^{d_1+\cdots+d_n}] \\
&= 3^{1+n}\frac{1}{t^{2n/3}}\mathbb{E}[t^{d_1}]\cdots\mathbb{E}[t^{d_n}] \\
&= 3^{1+n}\frac{1}{t^{2n/3}}\left(\frac{1+t+t^2}{3}\right)^n \\
&= 3\left(t^{-2/3}+t^{1/3}+t^{4/3}\right)^n.
\end{aligned}$$

Using basic calculus, we obtain that $t = \frac{-1+\sqrt{33}}{8}$ gives us the best bound. This proves the lemma. □

Theorem 2.2.3. *Let $r_3(n)$ be the largest size of cap sets in $\mathbb{F}_3^n$. Then*

$$r_3(n) = o(2.756^n).$$

Proof. Given a cap set $A \subseteq \mathbb{F}_3^n$. We define F as (3). By lemma 2.2.2, we have slice-rank$(F) = o(2.756^n)$. Observe that F is diagonal since A contains no non-trivial solution for $x + y + z = 0$. Hence we can apply theorem 2.1.3 and obtain that $|A| \leq$ slice-rank(F). Thus $|A| = o(2.756^n)$. □

2.3 Application 2: Solutions of particular linear systems

Recall that the cap set problem can be considered as finding the size of largest solution-free subset $A \subseteq \mathbb{F}_3^n$ corresponding to a single equation $x+y+z = 0$. Now we'll consider a more general problem. Rather than one equation, we consider a system of equations.

doi:10.6342/NTU202200962

Given integers $m, k \geq 1$ and a prime p, we consider the following system:

$$\begin{cases} a_{1,1}x_1 + \cdots + a_{1,k}x_k & = 0, \\ \qquad\vdots \\ a_{m,1}x_1 + \cdots + a_{m,k}x_k & = 0, \end{cases}$$

where the coefficients $a_{i,j}$ are in $\mathbb{F}_p$ and the variables x_i are in $\mathbb{F}_p^n, n \in \mathbb{N}$.

A natural question is to ask how large does $A \subseteq \mathbb{F}_p^n$ need so that there exists a non-trivial solution $(x_1, \ldots, x_k) \in A^k$ to the system? We first observe that if the coefficients satisfying $a_{1,1}+\cdots+a_{1,k} \neq 0$, then we can take

$$A = \{ \begin{pmatrix} 1 \\ x_{1,2} \\ \vdots \\ x_{1,n} \end{pmatrix} \in \mathbb{F}_p^n : x_{1,2}, \ldots, x_{1,n} \in \mathbb{F}_p \}.$$

In this case, any element $(x_1, \ldots, x_k) \in A^k$ is not a solution of the system. Also, $|A| = p^{n-1} = \frac{1}{p}|\mathbb{F}_p^n|$, which is the same order as the size of the space $\mathbb{F}_p^n$ as $n \to \infty$.

Thus we may **assume** $a_{i,1} + \cdots + a_{i,k} = 0$ **for all** $i = 1, \ldots, m$. Note that in this case, $(x, \ldots, x)$ is a solution of the system for any $x \in A$. We say such solutions are **trivial**.

For later reference, we formulate our question again:

Question 2.3.1. Given integers $m, k \geq 1$ and a prime p, we consider the following system:

$$\begin{cases} a_{1,1}x_1 + \cdots + a_{1,k}x_k = 0, \\ \qquad\vdots \\ a_{m,1}x_1 + \cdots + a_{m,k}x_k = 0, \end{cases} \tag{5}$$

where $x_i \in \mathbb{F}_p^n, n \in \mathbb{N}$, and the coefficients $a_{i,j} \in \mathbb{F}_p$ satisfying $a_{i,1}+\cdots+a_{i,k} = 0$ for all $i = 1, \ldots, m$. Is it possible to find a constant $1 \leq C_{p,m,k}$ and $1 \leq \Gamma_{p,m,k} < p$ such that if $A \subseteq \mathbb{F}_p^n$ with $|A| >$

doi:10.6342/NTU202200962

$C_{p,m,k}(\Gamma_{p,m,k})^n$, then there exist a non-trivial solution $(x_1 \ldots, x_k) \in A^k$ to the system (5)?

This is in fact one of the fundamental questions in additive combinatorics. Currently, we only know that the answer is positive under some certain conditions. See [22, 23, 28]. The natural number version of this problem with sets $A \subseteq [N]$ is also studied in recent years. [5, 18, 19, 25, 26].

Example 2.3.2. *Here's a example of such system:*

$$\begin{cases} x_1 - 2x_2 + \ x_3 & = 0, \\ \quad x_2 - 2x_3 + \ x_4 & = 0, \\ \quad\quad x_3 - 2x_4 + x_5 & = 0, \\ \quad\quad\quad \ddots & \\ \quad\quad\quad\quad x_m - 2x_{m+1} + x_{m+2} & = 0. \end{cases}$$

We can see the set of all non-trivial solutions $(x_1, \ldots, x_{m+2})$ are precisely the set of all $(m+2)-$terms arithmetic progressions in $\mathbb{F}_p^n$. Thus one of the special case of Question 2.3.1 is to find an exponential bound for the size of k-term-progression-free subsets in $\mathbb{F}_p^n$. Many partial results are studied in [3, 7, 8, 9, 14, 15, 16, 20, 21], but this is still an open problem in general.

It turns out that under the case $k \geq 2m+1$, the slice rank methods can help us answer the question. Before showing the result, we first define the constant $\Gamma_{p,m,k}$, which strictly smaller than p if $k \geq 2m+1$.

Definition 2.3.3. *For a prime p and two positive integers m and k, define the constant*

$$\Gamma_{p,m,k} := \min_{0<t\leq 1} \frac{1+t+\cdots+t^{p-1}}{t^{m(p-1)/k}}.$$

Remark. The function $f(t) = \dfrac{1+t+\cdots+t^{p-1}}{t^{m(p-1)/k}}$ for $0 < t \leq 1$ indeed attains its minimum. To see this, we observe that $\lim_{t\to 0^+} f(t) = +\infty$. So there is some small enough $\epsilon_{p,m,k} > 0$ such that

$$\inf_{0<t\leq 1} f(t) = \inf_{\epsilon_{p,m,k}\leq t\leq 1} f(t).$$

doi:10.6342/NTU202200962

Since $f(t)$ is continuous and the interval $[\epsilon_{p,m,k}, 1]$ is compact, $f(t)$ indeed attains its minimum on $[\epsilon_{p,m,k}, 1]$, hence on $(0, 1]$.

Proposition 2.3.4. *If $k \geq 2m+1$, then $1 \leq \Gamma_{p,m,k} < p$.*

Proof. The first inequality follows from the observation

$$\Gamma_{p,m,k} \geq \min_{0<t\leq 1} \frac{1}{t^{m(p-1)/k}} \geq 1$$

since $m(p-1)/k > 0$. For the second inequality, we let $f(t) = \dfrac{1+t+\cdots+t^{p-1}}{t^{m(p-1)/k}}$ for $t > 0$. Then by direct computation, we have $f(1) = p$ and

$$f'(1) = [1+2+\cdots+(p-1)] - p\,(m(p-1)/k) = p(p-1)(\frac{1}{2} - \frac{m}{k}),$$

which is positive since $k \geq 2m+1$. So there is some $t < 1$ near 1 with $f(t) < f(1) = p$. Hence $\Gamma_{p,m,k} = \min_{0<t\leq 1} f(t) < p$.

□

Lemma 2.3.5. *[6, 31] Suppose $k \geq 2m+1$. We consider the system (5). Given an integer L and vectors $x_i^{(l)} \in \mathbb{F}_p^n$ for $i \in [k]$ and $l \in [L]$. We define the function $f : [L]^k \to \mathbb{F}$ as*

$$F(l_1, \ldots, l_k) = \begin{cases} 1, & \text{if } (x_1^{(l_1)}, \ldots, x_k^{(l_k)}) \text{is a solution to the system (5)}, \\ 0, & \text{otherwise.} \end{cases}$$

Then the slice rank of F has an upper bound:

$$\text{slice-rank}(F) \leq k\,(\Gamma_{p,m,k})^n\,.$$

Proof. Denote $x_i(s)$ the $s-$th index of the vector $x_i \in \mathbb{F}_p^n$, then we can rewrite F as a polynomial

doi:10.6342/NTU202200962

of the variables $x_i(s)$:

$$F(l_1,\dots,l_k) = \begin{cases} 1, & \text{if } (x_1^{(l_1)},\dots,x_k^{(l_k)}) \text{is a solution to the system (5),} \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

$$= \begin{cases} 1, & \text{if } a_{i,1}x_1^{(l_1)}(s)+\cdots+a_{i,k}x_k^{(l_k)}(s)=0 \text{for all } i\in[m], s\in[n], \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

$$= \prod_{i=1}^{m}\prod_{s=1}^{n}\left[1-\left(a_{i,1}x_1^{(l_1)}(s)+\cdots+a_{i,k}x_k^{(l_k)}(s)\right)^{p-1}\right]. \tag{8}$$

The last equality holds by the Fermat's little theorem. The equation (8) is a polynomial of total degree $mn(p-1)$ in the kn variables $x_i(s)$. Since $t^p = t$ for all $t \in \mathbb{F}_p$, we can further represent $F(x_1,\dots,x_k)$ such that each individual variable appears with degree at most $p-1$. Hence every monomial of (8) can be written as a constant multiple of

$$\prod_{i=1}^{k} x_i^{(l_i)}(1)^{d_1^{(i)}}\cdots x_i^{(l_i)}(n)^{d_n^{(i)}},$$

where integers $0 \le d_1^{(i)},\dots,d_n^{(i)} \le p-1$ satisfying the inequality

$$\sum_{i=1}^{k}\left(d_1^{(i)}+\cdots+d_n^{(i)}\right) \le mn(p-1).$$

Hence for each such monomial, there is some $i \in [k]$ such that

$$d_1^{(i)}+\cdots+d_n^{(i)} \le \frac{mn(p-1)}{k}. \tag{9}$$

Using this property, we can sort all the monomials of F and rewrite it as

$$F(l_1,\dots,l_k) = \sum_{i=1}^{k}\sum_{d_1,\dots,d_n}\left[x_i^{(l_i)}(1)^{d_1}\cdots x_i^{(l_i)}(n)^{d_n}\right] g_{i,d_1,\dots,d_k}(x_1^{(l_1)},\dots,x_{i-1}^{(l_{i-1})},x_{i+1}^{(l_{i+1})},\dots,x_k^{(l_k)}),$$

where the summation of the d_i runs over all the $(d_1,\dots,d_n) \in \{0,\dots,p-1\}^n$ satisfying the inequality

doi:10.6342/NTU202200962

(9). This shows that the slice rank of F is at most

$$k|\{(d_1,\ldots,d_n)\in\{0,\ldots,p-1\}^n|\sum_{i=1}^{k} d_i \le \frac{mn(p-1)}{k}\}.$$

We note that this value equals to

$$kp^n\mathbb{P}[\sum_{i=1}^{n} d_i \le \frac{mn(p-1)}{k}],$$

where $\mathbb{P}$ is the probability corresponding to the uniformly random choices of $d_i \in \{0,\ldots,p-1\}$. Fix any $0<t\le 1$, we have

$$\begin{aligned}
\text{slice-rank}(F) &\le kp^n\mathbb{P}[\sum_{i=1}^{n} d_i \le \frac{mn(p-1)}{k}] \\
&\le kp^n\mathbb{P}[t^{d_1+\cdots+d_n} \ge t^{mn(p-1)/k}] \\
&\le kp^n\frac{1}{t^{mn(p-1)/k}}\mathbb{E}[t^{d_1+\cdots+d_n}] \\
&= kp^n\frac{1}{t^{mn(p-1)/k}}\mathbb{E}[t^{d_1}]\cdots\mathbb{E}[t^{d_n}] \\
&= kp^n\frac{1}{t^{mn(p-1)/k}}\left(\frac{1+t+\cdots+t^{p-1}}{p}\right)^n \\
&= k\left(\frac{1+t+\cdots+t^{p-1}}{t^{m(p-1)/k}}\right)^n.
\end{aligned}$$

Since it holds for all $0<t\le 1$, we have

$$\text{slice-rank}(F) \le k\left(\Gamma_{p,m,k}\right)^n.$$

□

The following problem gives a positive answer to the question 2.3.1 under the case $k \ge 2m+1$.

Theorem 2.3.6. *[31] Suppose $k \ge 2m+1$, consider the constant $\Gamma_{p,m,k} \in [1,p)$ we defined in Definition 2.3.3. Then for any subset $A \subseteq \mathbb{F}_p^n$ of size $|A| > \left(\Gamma_{p,m,k}\right)^n$, the system (5) has a non-*

doi:10.6342/NTU202200962

trivial solution $(x_1, \ldots, x_k) \in A^k$.

Proof. Given $A \subseteq \mathbb{F}_p^n$, we defined $F : A^k \to \mathbb{F}_p$ as

$$F(x_1, \ldots, x_k) = \begin{cases} 1, & \text{if } (x_1, \ldots, x_k) \text{is a solution to the system (5),} \\ 0, & \text{otherwise.} \end{cases}$$

By Lemma 2.3.5, we have slice-rank$(F) \leq k\left(\Gamma_{p,m,k}\right)^n$. If the all solutions $(x_1, \ldots, x_k) \in A^k$ to (5) are trivial, then F is diagonal. Hence slice-rank$(F) = |A|$ by Theorem 2.1.3. This shows that $|A| \leq k\left(\Gamma_{p,m,k}\right)^n$. To remove the coefficient k, we consider the system 5 except the variables are now in F_p^{nt}, where $t \in \mathbb{N}$, and we replace A with $A^t \subseteq \mathbb{F}_p^{nt}$. Then the above argument still holds, and this time we obtain the inequality $|A|^t \leq k\left(\Gamma_{p,m,k}\right)^{nt}$ for all $t \in \mathbb{N}$, i.e.

$$|A| \leq k^{1/t}\left(\Gamma_{p,m,k}\right)^n$$

for all $t \in \mathbb{N}$. Taking $t \to \infty$, and we obtain that $|A| \leq \left(\Gamma_{p,m,k}\right)^n$. □

2.4 Application 3: Largest right-angle-free subsets of $\mathbb{F}_p^n$

Another application of slice rank methods is also an extremal problems in combinatorics. Let q be an odd prime power. We'll use the slice rank to obtain a polynomial bound on the size of the largest right-angles-free subsets in vector space $\mathbb{F}_q^n$. We denote $\langle \cdot, \cdot \rangle$ as the dot product in $\mathbb{F}_q^n$.

Definition 2.4.1 (Right angle)**.** *A* ***right angle*** *in* $\mathbb{F}_q^n$ *is a triple* $(x, y, z) \in \mathbb{F}_q^n \times \mathbb{F}_q^n \times \mathbb{F}_q^n$ *of distinct elements satisfying*

$$\langle x - z, y - z \rangle = 0.$$

We say a set A *contains a right angle if* (x, y, z) *is a right angle for some* $x, y, z \in A$.

In 2015, Bennett [4] proved that any subset $A \subseteq \mathbb{F}_q^n$ of size

$$|A| \geq 4q^{\frac{n+2}{3}}$$

doi:10.6342/NTU202200962

contains a right angle. One year later, Ge and Shangguan [12] used the slice rank methods to improve this result for large n. They showed that if $A \subseteq \mathbb{F}_q^n$ satisfying

$$|A| > \binom{n+q}{q-1} + 3, \tag{10}$$

then it contains a right angle. Another year later, Naslund [24] also used the slice rank methods and slightly improved the bound:

Theorem 2.4.2. *[24] Let q be an odd prime power. If $A \subseteq \mathbb{F}_q^n$ satisfies*

$$|A| > \binom{n+q}{q-1} + 2 - \binom{n+q-2}{q-3},$$

then A contains a right angle.

We'll follow Naslund's proof.

Proof. Consider the function $F : \mathbb{F}_q^n \times \mathbb{F}_q^n \times \mathbb{F}_q^n \to \mathbb{F}_q^n$ defined by

$$F(x,y,z) = (1 - \mathbb{1}_{x=y} - \mathbb{1}_{x=z} - \mathbb{1}_{y=z})(1 - \langle x-z, y-z\rangle^{q-1}).$$

Then F is an indicator for distinct right angles since

$$F(x,y,z) = \begin{cases} -2 & \text{if } x=y=z, \\ 1 & \text{if } (x,y,z) \text{is a right angle}, \\ 0 & \text{otherwise.} \end{cases}$$

If $A \subseteq \mathbb{F}_q^n$ has no right angle, then $F|_{A\times A\times A}$ is a diagonal function since q is odd. Thus by Theorem 2.1.3,

$$|A| = \text{slice-rank}(F|_{A\times A\times A}) \leq \text{slice-rank}(F).$$

On the other hand, we will obtain an upper bound on slice-rank(F) by rewriting $F(x,y,z)$ into a

doi:10.6342/NTU202200962

linear combination of slice rank 1 functions. Note that

$$\begin{cases} \mathbb{1}_{x=z} \cdot (1 - \langle x-z, y-z \rangle^{q-1}) = \mathbb{1}_{x=z}, \\ \mathbb{1}_{y=z} \cdot (1 - \langle x-z, y-z \rangle^{q-1}) = \mathbb{1}_{y=z}, \end{cases}$$

thus both of them are of slice rank 1.

For the remaining part, we note that

$$(1 - \langle x-z, y-z \rangle^{q-1})$$

can be written as a linear combination of terms of the form

$$x_1^{d_1} \cdots x_n^{d_n} y_1^{e_1} \cdots y_n^{e_n} z_0^{f_0} z_1^{f_1} \cdots z_n^{f_n},$$

where x_i, y_i, z_i denotes the ith coordinate of x, y, z, for $i = 1, \ldots, n$ respectively, z_0 denotes $(z_1^2 + \cdots + z_n^2)$, and each d_i, e_j, f_k are non-negative integers satisfying that

$$\begin{cases} d_1 + d_2 + \cdots + d_n & \leq q-1, \\ e_1 + e_2 + \cdots + e_n & \leq q-1, \\ f_0 + f_1 + \cdots + f_n & \leq q-1. \end{cases}$$

Thus $(1 - \mathbb{1}_{x=y})(1 - \langle x-z, y-z \rangle^{q-1})$ can be written as a linear combination of terms of the form $g(x,y)h(z)$, where h is a polynomial in the space

$$\mathrm{Poly}_{q-1}^2(\mathbb{F}_q^n) := \mathrm{span}_{\mathbb{F}_q}\{(z_1^2 + \cdots z_n^2)^{f_0} z_1^{f_1} \cdots z_n^{f_n} \mid f_k \in \mathbb{Z}_{\geq 0}, \sum_{k=0}^{n} f_i \leq q-1\}.$$

Thus

$$\text{slice-rank}(F) \leq 2 + \dim \mathrm{Poly}_{q-1}^2(\mathbb{F}_q^n).$$

We use the following lemma:

doi:10.6342/NTU202200962

Lemma 2.4.3. *[2] Let $\mathbb{F}$ be a field and d, n be positive integers, then*

$$\dim \mathrm{Poly}_d^2(\mathbb{F}^n) = \binom{n+d}{d} + \binom{n+d-1}{d-1}.$$

Then we have

$$\begin{aligned} |A| \leq \text{slice-rank}(F) &\leq 2 + \binom{n+q-1}{q-1} + \binom{n+q-2}{q-2} \\ &= \binom{n+q}{q-1} + 2 - \binom{n+q-2}{q-3} \end{aligned}$$

by the binomial identity $\binom{n+q-1}{q-1} + \binom{n+q-2}{q-2} + \binom{n+q-2}{q-3} = \binom{n+q}{q-1}$. □

Remark. In later section, we'll show the analogous result for "k-right corner". See Theorem 4.3.2 for more detail.

3 Slice Rank of Tensors

Recall that in the previous section, we define the slice rank of functions, which can also be considered as hypermatrices. But both of them are basis *dependent*. It turns out that the slice rank of tensors would be basis *independent*. Using such definition, some properties of slice rank becomes much easier to prove. Before introducing the slice rank of tensors, we introduce the term "tensor product" first.

3.1 Tensor product of vector spaces

We first consider the tensor product between two vector spaces.

Definition 3.1.1 (Tensor product of 2 vector spaces)**.** *Given two vector spaces V, W over the same field $\mathbb{F}$, the tensor product $V \bigotimes W$ is a vector space over $\mathbb{F}$ generated by the elements $v \otimes w$, where $v \in V$ and $w \in W$, subject to the constrain that the following operation is bilinear:*

$$(v, w) \mapsto v \otimes w.$$

doi:10.6342/NTU202200962

Thus every elements in the tensor product $V \bigotimes W$ are of the form

$$\sum_{i=1}^{N} v_i \otimes w_i$$

where $v_i \in V, w_i \in W$ and $N \in \mathbb{N} \cup \{0\}$. The bilinear constrains allow us to do the following operations:

- $(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$,
- $v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2$,
- $(cv) \otimes w = v \otimes (cw) = c(v \otimes w)$,

for $v, v_1, v_2 \in V, w, w_1, w_2 \in W$ and $c \in \mathbb{F}$.

Note that in the definition of the tensor product, it is no need to specify the basis of the vector space V and W. But if we specify the basis, the elements in the tensor product would have a unique expression corresponding to the given basis. Say the vector spaces V, W have basis $\{v_1 \ldots, v_n\}, \{w_1, \ldots, w_m\}$, respectively, then the tensor product $V \bigotimes W$ has basis

$$\{v_i \otimes w_j : 1 \leq i \leq n \text{and } 1 \leq j \leq m\}.$$

So the elements in the tensor product $V \bigotimes W$ can be uniquely written as

$$\sum_{i=1}^{n} \sum_{j=1}^{m} c_{i,j} v_i \otimes w_j,$$

and this can be considered as a function $F : [n] \times [m] \to \mathbb{F}$ with $F(i, j) = c_{i,j}$ or an $n \times m$ matrix whose (i, j)-index is $c_{i,j}$. Notice that if we choose different basis for V and W, the corresponding matrix would be different, but those matrices are mutually similar.

Now we consider the tensor product of k vector spaces. We'll define the slice rank of tensors in these space.

Definition 3.1.2 (Tensor product of k vector spaces)**.** *Given vector spaces $V_1, \ldots, V_k$ over the same field $\mathbb{F}$, the tensor product $V_1 \bigotimes \cdots \bigotimes V_k = \bigotimes_{i=1}^{k} V_i$ is a vector space over $\mathbb{F}$ generated by all*

doi:10.6342/NTU202200962

the elements $v_1 \otimes \cdots \otimes v_k$, where $v_i \in V_i$, subject to the constrain that the following operation is multilinear:

$$(v_1, \ldots, v_k) \mapsto v_1 \otimes \cdots \otimes v_k.$$

Thus every elements in $\bigotimes_{i=1}^k V_i$ are of the form

$$\sum_{j=1}^{N} v_{1,j} \otimes \cdots \otimes v_{k,j},$$

where $v_{i,j} \in V_i$ and $N \in \mathbb{N} \cup \{0\}$. Similar to the $k = 2$ case, we can do the following operations:

- $v_1 \otimes \cdots \otimes (v_i + v_i') \otimes \cdots \otimes v_k = (v_1 \otimes \cdots \otimes v_i \otimes \cdots \otimes v_k) + (v_1 \otimes \cdots \otimes v_i' \otimes \cdots \otimes v_k)$,
- $v_1 \otimes \cdots \otimes (cv_i) \otimes \cdots \otimes v_k = c(v_1 \otimes \cdots \otimes v_k)$,

for $v_i, v_i' \in V_i$ and $c \in \mathbb{F}$.

We also don't need to specify the basis of the vector space V_i here. But if we specify the basis, the elements in the tensor product would have a unique expression. Say the vector spaces V_i have basis $\{v_{i,1} \ldots, v_{i,d_i}\}$ for each i, then the tensor product $\bigotimes_{i=1}^k V_i$ has basis

$$\{v_{1,t_1} \otimes \cdots \otimes v_{k,t_k} : 1 \leq t_i \leq d_i \text{for each } i\}.$$

So the elements in the tensor product $\bigotimes_{i=1}^k V_i$ can be uniquely written as

$$\sum_{t_1=1}^{d_1} \cdots \sum_{t_k=1}^{d_k} c_{t_1,\ldots,t_k} v_{1,t_1} \otimes \cdots \otimes v_{k,t_k},$$

and this can be considered as a function $F : [d_1] \times \cdots \times [d_k] \to \mathbb{F}$ or an $d_1 \times \cdots \times d_k$ hypermatrix $(c_{t_1,\ldots,t_k})_{t_1,\ldots,t_k}$.

In the tensor product $\bigotimes_{i=1}^k V_i$, we can define an operation: for each $1 \leq j \leq k$, we have the smaller tensor product $\bigotimes_{1 \leq i \leq k, i \neq j} V_i$ and the jth tensor product

$$\otimes_j : V_j \times \bigotimes_{1 \leq i \leq k, i \neq j} V_i \to \bigotimes_{i=1}^{k} V_i,$$

doi:10.6342/NTU202200962

which defined as the unique $\mathbb{F}$-bilinear operation satisfying that

$$v_j \otimes_j (v_1 \otimes \cdots \otimes v_{j-1} \otimes v_{j+1} \otimes \cdots \otimes v_k) = v_1 \otimes \cdots \otimes v_k$$

for all $v_i \in V_i$.

3.2 Slice rank of tensors

Now we can define the slice rank of tensors.

Definition 3.2.1 (Slice rank of tensors). *Given vector spaces $V_1, \ldots, V_k$ over the same field $\mathbb{F}$. We say a non-zero element v in the tensor product $\bigotimes_{i=1}^k V_i$ is of slice rank one if it is of the form*

$$v = v_j \otimes_j v_{\hat{j}}$$

for some $j \in [k]$, $v_j \in V_j$ and $v_{\hat{j}} \in \bigotimes_{1 \le i \le k, i \ne j} V_i$.

The slice rank of any element $v \in \bigotimes_{i=1}^k V_i$ is the least non-negative integer $r = \text{slice-rank}(v)$ such that v can be written as a sum of r slice rank one elements.

Remark. Compare to the definition of slice rank of functions (Definition 2.1.1). They are actually compatible. More specifically, given a function $f : X_1 \times \cdots \times X_k \to \mathbb{F}$, where $X_i = [d_i]$ are finite sets and $\mathbb{F}$ is a field. We can consider the tensor product

$$\bigotimes_{i=1}^k \mathbb{F}^{|X_i|},$$

and let the element

$$v = \sum_{1 \le x_i \le d_i, 1 \le i \le k} f(x_1, \ldots, x_k) e_{x_1}^{(1)} \otimes \cdots \otimes e_{x_k}^{(k)},$$

where $\{e_1^{(i)}, \ldots, e_{|X_i|}^{(i)}\}$ is any fixed basis of the vector space $\mathbb{F}^{|X_i|}$. Then the slice rank of the function f equals to the slice rank of the tensor v. Conversely, by specifying any basis of V_i, we can construct a corresponding function for any tensor element and their slice ranks are the same.

This allows us to switch the objects whenever we want and it makes the proof of properties of slice

doi:10.6342/NTU202200962

rank much easier to achieve. They have their own advantages: Functions on finite sets are easier to visualize and compute, while tensors are elements in a vector space, so we can use tools in linear algebra.

The following proposition states some characterization of the slice rank of tensors.

Proposition 3.2.2. *[29] Given vector spaces $V_1, \ldots, V_k$ over the same field $\mathbb{F}$. For any $v \in \bigotimes_{i=1}^k V_i$ and any non-negative integer r, the following statements are equivalent:*

(i) slice-rank$(v) \leq r$.

(ii) The element v can be written as

$$v = \sum_{j=1}^{k} \sum_{\alpha \in S_j} v_{\alpha,j} \otimes_j v_{\alpha,\hat{j}},$$

for some $v_{\alpha,j} \in V_j$, $v_{\alpha,\hat{j}} \in \bigotimes_{1 \leq i \leq k, i \neq j} V_i$ with $|S_1| + \cdots + |S_k| \leq r$.

(iii) We have

$$v \in \sum_{j=1}^{k} U_j \otimes_j \left(\bigotimes_{1 \leq i \leq k, i \neq j} V_i \right),$$

where U_j is a vector subspace of V_j for each j with $\dim U_1 + \cdots + \dim U_k \leq r$ and we view the tensor product $U_j \otimes_j \left(\bigotimes_{1 \leq i \leq k, i \neq j} V_i \right)$ as a subspace of $\bigotimes_{i=1}^k V_i$ in the obvious fashion.

(iv) There exist subspaces W_i of the dual space V_i^* for each $i \in [k]$ respectively such that

$$\dim W_1 + \cdots + \dim W_k \geq \dim V_1 + \cdots + \dim V_k - r$$

and v is orthogonal to $\bigotimes_{i=1}^k W_i$ in the sense that

$$\langle w_1 \otimes \cdots \otimes w_k, v \rangle = 0 \qquad \forall w_i \in W_i,$$

where the dual pair $\langle \cdot, \cdot \rangle : \bigotimes_{i=1}^k V_i^* \times \bigotimes_{i=1}^k V_i \to \mathbb{F}$ is the obvious pairing.

Proof. The equivalence between (i) and (ii) follows from the definition of the slice rank of tensors.

doi:10.6342/NTU202200962

We suppose (ii) is true. For each $j \in [k]$, we let U_j be the subspace of V_j generated by the elements $v_{\alpha,j}, \alpha \in S_j$. Then

$$\sum_{j=1}^{k} \dim U_j \leq \sum_{j=1}^{k} |S_j| \leq r.$$

Also, by our construction of U_j, we obtain that

$$\sum_{\alpha \in S_j} v_{\alpha,j} \otimes_j v_{\alpha,\hat{j}} \in U_j \otimes_j \left(\bigotimes_{1 \leq i \leq k, i \neq j} V_i \right).$$

Hence $v \in \sum_{j=1}^{k} U_j \otimes_j \left(\bigotimes_{1 \leq i \leq k, i \neq j} V_i \right)$. So (ii) implies (iii).

On the other hand, we suppose (iii) is true. Then we take a basis of U_j. Say $\{v_{1,j}, \ldots, v_{\dim U_j, j}\}$ is a basis of U_j, we set $S_j = \{1, \ldots, \dim U_j\}$. Then we have

$$\sum_{j=1}^{k} |S_j| = \sum_{j=1}^{k} \dim U_j \leq r.$$

and every elements in $U_j \otimes_j \left(\bigotimes_{1 \leq i \leq k, i \neq j} V_i \right)$ can be written as

$$\sum_{\alpha=1}^{\dim U_j} v_{\alpha,j} \otimes_j v_{\alpha,\hat{j}},$$

for some $v_{\alpha,j} \in V_j$, and $v_{\alpha,\hat{j}} \in \bigotimes_{1 \leq i \leq k, i \neq j} V_i$. Thus (iii) implies (ii).

Now we suppose (iii) is true. For each $i \in [k]$, we let W_i be the annihilator of U_i. That is, the subspace consists of all elements in V_i^* that are orthogonal to U_i, which can be written as

$$W_i = \{w_i \in V_i^* \mid \langle w_i, u_i \rangle = 0 \ \forall u_i \in U_i\},$$

where $\langle \cdot, \cdot \rangle$ denotes the dual pair. Then $\dim W_i = \dim V_i^* - \dim U_i = \dim V_i - \dim U_i$. Hence

$$\sum_{i=1}^{k} \dim W_i = \sum_{i=1}^{k} (\dim V_i - \dim U_i) \geq \sum_{i=1}^{k} \dim V_i - r.$$

By definition, W_i is orthogonal to U_i for each i w.r.t to the dual pair, so $\bigotimes_{i=1}^{k} W_j$ is orthogonal to

doi:10.6342/NTU202200962

$U_j \otimes_j \left(\bigotimes_{1\le i\le k, i\ne j} V_i\right)$ for each j. Since v is generated by the elements in $U_j \otimes_j \left(\bigotimes_{1\le i\le k, i\ne j} V_i\right)$ with $j \in [k]$, $\bigotimes_{i=1}^k W_j$ is orthogonal to v. So (iii) implies (iv).

Finally, we suppose (iv) is true. In this case, we let U_i be the annihilator of W_i. Then $\dim U_i = \dim V_i^* - \dim W_i = \dim V_i - \dim W_i$. Hence

$$\sum_{i=1}^k \dim U_i = \sum_{i=1}^k (\dim V_i - \dim W_i) \le r.$$

Recall a result in linear algebra [17, Page 27, Exercise 8(c)], which states that if V_1 and V_2 are subspaces of a finite dimensional vector space V, and V_1°, V_2° are the annihilator of V_1, V_2 respectively. Then

$$(V_1 \cap V_2)^\circ = V_1^\circ + V_2^\circ. \tag{11}$$

Since $U_j = W_j^\circ$ for each j, we have

$$U_j \otimes_j \left(\bigotimes_{1\le i\le k, i\ne j} V_i\right) = \left(W_j \otimes_j \left(\bigotimes_{1\le i\le k, i\ne j} V_i\right)\right)^\circ.$$

Thus by (11),

$$v \in \{w_1 \otimes \cdots \otimes w_k \mid w_i \in W_i\}^\circ = \left(\bigcap_{j=1}^k W_j \otimes_j \left(\bigotimes_{1\le i\le k, i\ne j} V_i\right)\right)^\circ = \sum_{j=1}^k U_j \otimes_j \left(\bigotimes_{1\le i\le k, i\ne j} V_i\right),$$

as desired. □

Using the above characterizations of the slice rank of tensors, we can generalize Theorem 2.1.3. First, we recall the definition of an antichain.

Definition 3.2.3 (antichain)**.** *Given a partial order $\preceq$ on a set S. A subset $A \subseteq S$ is called an **antichain** if any two distinct elements in A are not comparable, that is, for all distinct $x, y \in A$, we have both $x \not\preceq y$ and $y \not\preceq x$.*

Theorem 3.2.4. *[29] Let $L, k \in \mathbb{N}$ and let $\mathbb{F}$ be a field. Fix k total ordering $\preceq^1, \ldots, \preceq^k$ on $[L]$ and we consider the corresponding product partial ordering $\preceq$ on $[L]^k$, i.e. $(a_1, \ldots, a_k) \preceq (b_1, \ldots, b_k)$ iff*

doi:10.6342/NTU202200962

$a_i \preceq^i b_i$ for each $i = 1, \ldots, k$. Suppose that $F : [L]^k \to \mathbb{F}$ is a function with its support

$$S = \{(\ell_1, \ldots, \ell_k) \in [L]^k : F(\ell_1, \ldots, \ell_k) \neq 0\}$$

is an antichain w.r.t to $\preceq$, Then

$$\text{slice-rank}(F) = \min_{S = S_1 \cup \cdots \cup S_k} \left(|\pi_1(S_1)| + \cdots + |\pi_k(S_k)| \right),$$

where $\pi_i : [L]^k \to [L]$ is the ith projection. That is, $\pi_i(\ell_1, \ldots, \ell_k) = \ell_i$ for each i.

Proof. Consider the vector space $\mathbb{F}^L$ over $\mathbb{F}$ with standard basis $e_1, \ldots, e_L$, and let

$$\begin{aligned} v &= \sum_{(\ell_1,\ldots,\ell_k) \in [L]^k} F(\ell_1, \ldots, \ell_k) e_{\ell_1} \otimes \cdots \otimes e_{\ell_k} \\ &= \sum_{(\ell_1,\ldots,\ell_k) \in S} F(\ell_1, \ldots, \ell_k) e_{\ell_1} \otimes \cdots \otimes e_{\ell_k}. \end{aligned} \tag{12}$$

Then v is an element in the tensor product $\bigotimes_{i=1}^k \mathbb{F}^L$ and slice-rank(v) = slice-rank(F). Given any partition $S = S_1 \cup \cdots \cup S_k$, we have

$$v = \sum_{j=1}^{k} \sum_{\ell_j \in \pi_j(S_j)} e_{\ell_j} \otimes_j v_{\hat{j}, \ell_j}$$

for some $v_{\hat{j},\ell_j} \in \bigotimes_{1 \le i \le k, i \neq j} \mathbb{F}^L$. Thus we have an upper bound

$$\text{slice-rank}(v) \le |\pi_1(S_1)| + \cdots + |\pi_k(S_k)|.$$

Note that this inequality also holds if S is not an antichain. Now it suffices to show

$$\text{slice-rank}(v) \ge |\pi_1(S_1)| + \cdots + |\pi_k(S_k)|$$

for some partition $S = S_1 \cup \cdots \cup S_k$. Without loss of generality, assume the total orderings $\preceq^i$ are all defined as $1 \preceq^i \cdots \preceq^i L$ for each i. Let r = slice-rank(v). By the statement (iv) of Proposition 3.2.2,

doi:10.6342/NTU202200962

there exist subspaces W_i of the dual space $(\mathbb{F}^L)^*$ for each $i = 1, \ldots, k$ so that v is orthogonal to $\bigotimes_{i=1}^k W_i$ and

$$\dim W_1 + \cdots + \dim W_k \geq \sum_{i=1}^{k} \dim \mathbb{F}^L - r = kL - r.$$

Let $d_i := \dim W_i$. By Gaussian elimination, we can find basis $\{w_{i,1}, \ldots, w_{i,d_i}\}$ of W so that they're in the row echelon form with respect to the standard dual basis $e_1^*, \ldots, e_L^*$ of $(\mathbb{F}^L)^*$. Hence there exists indexes

$$1 \leq t_{i,1} < \cdots < t_{i,d_i} \leq L$$

so that $w_{i,j}$ is a linear combination of $e_{t_{i,j}}^*, e_{t_{i,j}+1}^*, \ldots, e_L^*$ and the coefficient of $e_{t_{i,j}}^*$ is 1. We claim that $S \cap \prod_{i=1}^k \{t_{i,1}, \ldots, t_{i,d_i}\} = \emptyset$. Suppose not, then there exist $1 \leq r_i \leq d_i$ such that $(t_{1,r_1}, \ldots, t_{1,r_k}) \in S$. Since v is orthogonal to $\bigotimes_{i=1}^k W_i$, we have

$$\langle w_{1,r_1} \otimes \cdots \otimes w_{k,r_k}, v \rangle = 0.$$

On the other hand, the element $w_{1,r_1} \otimes \cdots \otimes w_{k,r_k}$ is $e_{t_{1,r_1}}^* \otimes \ldots \otimes e_{t_{k,r_k}}^*$ plus a linear combination of those $e_{t'_1}^* \otimes \ldots \otimes e_{t'_k}^*$ with

$$(t'_1, \ldots, t'_k) > (t_{1,r_1}, \ldots, t_{k,r_k}).$$

Since S is an antichain and $(t_{1,r_1}, \ldots, t_{1,r_k}) \in S$, we obtain that $(t'_1, \ldots, t'_k) \notin S$. Hence $F(t'_1, \ldots, t'_k) = 0$. This shows that the coefficient of $e_{t'_1} \otimes \ldots \otimes e_{t'_k}$ in (12) is 0 for all $(t'_1, \ldots, t'_k) > (t_{1,r_1}, \ldots, t_{k,r_k})$. Therefore

$$0 = \langle w_{1,r_1} \otimes \cdots \otimes w_{k,r_k}, v \rangle = \langle e_{t_{1,r_1}}^* \otimes \ldots \otimes e_{t_{k,r_k}}^*, v \rangle = F(t_{1,r_1}, \ldots, t_{k,r_k}),$$

which implies $(t_{1,r_1}, \ldots, t_{k,r_k}) \notin S$, a contradiction. Therefore $S \cap \prod_{i=1}^k \{t_{i,1}, \ldots, t_{i,d_i}\} = \emptyset$ is true. Since $S \cap \prod_{i=1}^k \{t_{i,1}, \ldots, t_{i,d_i}\} = \emptyset$, we may let

$$S_i := \{(\ell_1, \ldots, \ell_k) \in S : \ell_i \notin \{t_{i,1}, \ldots, t_{i,d_i}\}\}$$

doi:10.6342/NTU202200962

for each i. Then $S = S_1 \cup \cdots \cup S_k$ and $|\pi_i(S_i)| \leq L - d_i$ for each i, which implies that

$$|\pi_1(S_1)| + \cdots + |\pi_k(S_k)| \leq \sum_{i=1}^{k}(L - d_i) \leq kL - (kL - r) = r,$$

as desired. □

Remark. The above theorem is stronger than Theorem 2.1.3. Indeed, suppose $k \geq 2$ and we define the total orderings $\preceq^1$ and $\preceq^2$ as

$$1 \preceq^1 \cdots \preceq^1 L \qquad \text{and} \qquad L \preceq^2 \cdots \preceq^2 1,$$

and the remaining orderings $\preceq^i$ are defined arbitrary. If $F : [L]^k \to \mathbb{F}$ is diagonal, then its support

$$S \subseteq \{(\ell, \ldots, \ell) \in [L]^k : \ell \in [L]\}$$

is an antichain w.r.t the corresponding product partial order, and for any partition $S = S_1 \cup \cdots \cup S_k$,

$$|S| = |\pi_1(S_1)| + \cdots + |\pi_k(S_k)|,$$

which implies that slice-rank$(F) = |S|$.

This theorem can also help us understand the structure of the solutions of linear system we discussed in Section 2.3.

Corollary 3.2.5. *[28, Corollary 3.7] Suppose we are given a linear system of equations with coefficients in $\mathbb{F}_p$ and constant terms in $\mathbb{F}_p^n$, consisting of $m \geq 1$ equations in $k \geq 2m+1$ variables. Let $(x_1^{(l)}, \ldots, x_k^{(l)}) \in (\mathbb{F}_p^n)^k$ for $l = 1 \ldots, L$ be solutions in $\mathbb{F}_p^n$ to this system of equations. Suppose that there exists a disjoint partition $[k] = J_1 \cup \cdots \cup J_t$ with $|J_h| \geq 2$ for each h such that the following condition holds: For any choice of $l_1, \ldots, l_k \in [L]$ such that $(x_1^{(l_1)}, \ldots, x_k^{(l_k)})$ is a solution to the given system of equations, we have $|\{l_j \mid j \in J_h\}| = 1$ for all $h = 1, \ldots, t$. Then we must have $L \leq k \cdot (\Gamma_{p,m,k})^n$.*

doi:10.6342/NTU202200962

Proof. Define the function $f : [L]^k \to \mathbb{F}_p$ as

$$f(l_1, \ldots, l_k) = \begin{cases} 1, & \text{if } (x_1^{(l_1)}, \ldots, x_k^{(l_k)}) \text{is a solution of the given system,} \\ 0, & \text{otherwise.} \end{cases}$$

By Lemma 2.3.5, we have slice-rank$(F) \le k \cdot (\Gamma_{p,m,k})^n$. So it suffices to show that $L =$ slice-rank(F). We'll define some total orderings on $[L]$ such that the support of F is an antichain with respect to the corresponding product ordering, which allows us to use the Theorem 3.2.4.

For each $h \in [t]$, we label the sets $J_h = \{j_{h,1}, j_{h,2}, \ldots, j_{h,|J_h|}\}$. Since $|J_h| \ge 2$, we can define the total orderings $\preceq^{j_{h,1}}, \preceq^{j_{h,2}}$ on $[L]$ as

$$1 \preceq^{j_{h,1}} 2 \preceq^{j_{h,1}} \cdots \preceq^{j_{h,1}} L$$

and

$$L \preceq^{j_{h,2}} L-1 \preceq^{j_{h,2}} \cdots \preceq^{j_{h,2}} 1.$$

The remaining orderings $\preceq^{j_{h,3}}, \ldots, \preceq^{j_{h,|J_h|}}$ can be any arbitrary total orderings on $[L]$. Given two distinct elements $(l_1, \ldots, l_k), (l'_1, \ldots, l'_k)$ in the support of F. By definition of F, we obtain that $(x_1^{(l_1)}, \ldots, x_k^{(l_k)})$ and $(x_1^{(l'_1)}, \ldots, x_k^{(l'_k)})$ are both solutions of the given system. Thus by assumption, we have

$$|\{l_j \mid j \in J_h\}| = |\{l'_j \mid j \in J_h\}| = 1$$

for each h. Since $(l_1, \ldots, l_k)$ and $(l'_1, \ldots, l'_k)$ are distinct, there is some $i \in [k] = J_1 \cup \cdots \cup J_t$ such that $l_i \ne l'_i$. Say $i \in J_h$ for some $h \in [t]$. Using $|\{l_j \mid j \in J_h\}| = |\{l'_j \mid j \in J_h\}| = 1$, we have

$$l_{j_{h,1}} = l_{j_{h,2}} = l_i \ne l'_i = l'_{j_{h,1}} = l'_{j_{h,2}}.$$

There are only two possibilities:

$$\begin{cases} l_{j_{i,1}} \preceq^{j_{i,1}} l'_{j_{i,1}}, \\ l_{j_{i,2}} \not\preceq^{j_{i,2}} l'_{j_{i,2}}, \end{cases} \quad \text{or} \quad \begin{cases} l_{j_{i,1}} \not\preceq^{j_{i,1}} l'_{j_{i,1}}, \\ l_{j_{i,2}} \preceq^{j_{i,2}} l'_{j_{i,2}}. \end{cases}$$

doi:10.6342/NTU202200962

But either case shows that $(l_1, \dots, l_k)$ and $(l'_1, \dots, l'_k)$ are not comparable with respect to the product ordering we've constructed. Therefore the support of F is an antichain. By Theorem 3.2.4, we have

$$\text{slice-rank}(F) = \min_{S=S_1\cup\cdots\cup S_k} \left(|\pi_1(S_1)| + \cdots + |\pi_k(S_k)|\right).$$

By assumption and the definition of F, we know that

$$F(l, l, \dots, l) = 1$$

for $l \in [L]$. Thus $\pi_1(S_1) \cup \cdots \cup \pi_k(S_k) = [L]$ for any disjoint unions $S_1, \dots, S_k$ of S. This forces $|\pi_1(S_1)| + \cdots + |\pi_k(S_k)| \geq L$. Together with slice-rank$(F) \leq L$, we obtain that slice-rank$(F) = L$.

□

The above corollary plays an important role in the paper of Sauermann [28]. For reference, she used this result to prove the following theorem.

Theorem 3.2.6. *[28] For any fixed integer $m \geq 1$ and $k \geq 3m$ and a fixed prime p, there exists constants $C_{p,m,k} \geq 1$ and $1 \leq \Gamma^*_{p,m,k} < p$ such that the following holds: Consider the system* (5) *with every $m \times m$ minor of the $m \times k$ matrix $(a_{j,i})_{j,i}$ is invertible. Then for any $n \in \mathbb{N}$ and any subset $A \subseteq \mathbb{F}_p^n$ of size $|A| > C_{p,m,k} \cdot (\Gamma^*_{p,m,k})^n$, the system* (5) *has a solution $(x_1, \dots, x_k) \in A^k$ such that the vectors $x_1, \dots, x_k$ are all distinct.*

4 Partition Rank

In this section, we'll introduce another way to define the rank of functions (and thus tensors), which we called it the partition rank.

4.1 Definitions and basic properties

Given variables $x_1, \dots, x_n$ and a subset $S \subseteq [n]$. Write $S = \{s_1, \dots, s_k\}$ with $1 \leq s_1 < \cdots < s_k \leq n$. We use $\vec{x_S}$ to denote the subset of variables $x_{s_1}, \dots, x_{s_k}$. So for a function g of k variables, $g(\vec{x_S})$

doi:10.6342/NTU202200962

denotes $g(x_{s_1}, \dots, x_{s_k})$.

Definition 4.1.1 (Partition)**.** *A **partition** of $[n]$ is a collection P of non-empty, pairwise disjoint, subsets of $[n]$, satisfying*

$$\bigcup_{S \in P} S = [n].$$

*We say that P is the **trivial partition** if P only contains one set $[n]$.*

Definition 4.1.2 (Partition rank)**.** *Given finite sets $X_1, \dots, X_k$ and a field $\mathbb{F}$. Consider a function*

$$F : X_1 \times \cdots \times X_k \to \mathbb{F}.$$

*We say F is of **partition rank one** if it is non-zero and can be written as*

$$F(x_1, \dots, x_k) = \prod_{S \in P} f_S(\vec{x_S})$$

for some functions f_S and a non-trivial partition P of $[n]$.

*For general $F : X_1 \times \cdots \times X_k \to \mathbb{F}$, its **partition rank** partition-rank(F) is the least non-negative integer r so that F can be written as a sum of r partition rank one functions.*

Remark. We remark that a non-zero $F : X_1 \times \cdots \times X_k \to \mathbb{F}$ has partition rank 1 if and only if it can be written as

$$F(x_1, \dots, x_n) = f(\vec{x_S})g(\vec{x_T})$$

for some f, g and some disjoint $S, T \neq \emptyset$ with $S \cup T = [n]$. Furthermore, if either $|S| = 1$ or $|T| = 1$, then F is of slice rank 1. Hence when the dimension k is less than or equal to 3, the partition rank coincides with the partition rank. For general positive integer k, we have the following inequality

$$\text{partition-rank}(F) \leq \text{slice-rank}(F) \leq \min_i |A_i|.$$

Example 4.1.3. *In some case, the partition rank and the slice rank can differ a lot. For example, if X is a finite set and $F : X^4 \to \mathbb{F}$ is defined as*

$$F(x_1, x_2, x_3, x_4) = \mathbb{1}_{x_1 = x_2} \mathbb{1}_{x_3 = x_4},$$

doi:10.6342/NTU202200962

or equivalently,

$$F(x_1,x_2,x_3,x_4)=\begin{cases}1 & \text{if } x_1=x_2 \text{and } x_3=x_4,\\ 0 & \text{otherwise.}\end{cases}$$

Then partition-rank$(F)=1$, *while* slice-rank$(F)=|X|$ *since its support is an antichain and we can apply Theorem 3.2.4.*

A key observation is that the analogous result of Theorem 2.1.3 also holds for partition rank.

Theorem 4.1.4. *[24] Let X be a non-empty finite set, $k\geq 2$ be integers and let $\mathbb{F}$ be a field. Suppose that $F:X^k\to\mathbb{F}$ is diagonal, that is,*

$$F(x_1,\dots,x_k)=\sum_{a\in A}c_a\mathbb{1}_{a=x_1=\cdots=x_n} \tag{13}$$

for some $A\subseteq X$ where $c_a\neq 0$. Then

$$\text{partition-rank}(F)=|A|.$$

Proof. Let $r=$ partition-rank(F). By the inequality partition-rank$(F)\leq$ slice-rank(F) and Theorem 2.1.3, we have $r\leq$ slice-rank$(F)=|A|$. So our goal is to prove $r\geq|A|$. We'll use induction on k. When $k=2$, the slice rank is equivalent to the partition rank, so the result follows from Theorem 2.1.3. Assume $k\geq 3$. Since partition-rank$(F)\leq$ slice-rank$(F)\leq\min_i|A_i|$ we can write:

$$F(x_1,\dots,x_n)=\sum_{i=1}^{r}f_i(\vec{x}_{S_i})g_i(\vec{x}_{T_i}), \tag{14}$$

where S_i,T_i are non-empty sets with $S_i\cap T_i=\emptyset$ and $S_i\cup T_i=[n]$. By switching the labeling if needed, we may assume that $|S_i|\leq n/2$ for each i. We consider two cases:

Case 1: $|S_i|\geq 2$ for each i.

By taking the summation $\sum_{x_1\in X}$ to both sides of the equation (14), we have

$$\sum_{x_1\in X}\sum_{a\in A}c_a\mathbb{1}_{a=x_1\cdots=x_n}=\sum_{x_1\in X}F(x_1,\dots,x_n)=\sum_{i=1}^{r}\tilde{f}_i(\vec{x}_{S_i\setminus\{1\}})\tilde{g}_i(\vec{x}_{T_i\setminus\{1\}}),$$

doi:10.6342/NTU202200962

where

$$\tilde{f}_i(\vec{x}_{S_i\setminus\{1\}}) := \begin{cases} \sum_{x_1\in X} f_i(\vec{x}_{S_i}) & \text{if } 1\in S_i, \\ f_i(\vec{x}_{S_i}) & \text{if } 1\notin S_i, \end{cases}$$

and $\tilde{g}_i(\vec{x}_{T_i\setminus\{1\}})$ is defined similarly. Observe that

$$\sum_{x_1\in X}\sum_{a\in A} c_a \mathbb{1}_{a=x_1\cdots=x_n} = \sum_{a\in A} c_a \mathbb{1}_{a=x_2=\cdots=x_n},$$

So $\sum_{a\in A} c_a \mathbb{1}_{a=x_2=\cdots=x_n}$ has partition rank at most r. By induction hypothesis, its partition rank is precisely $|A|$. Thus $|A|\le r$.

Case 2: $|S_i|=1$ for some i.

Then we have $S_i=\{j\}$ for some $j\in[n]$. Let $U=\{u\in[r]: S_u=\{j\}\}$. Define the following vector space

$$V=\{h: X\to\mathbb{F}: \sum_{x_j\in X} f_u(x_j)h(x_j)=0 \ \ \forall u\in U\}.$$

This vector space has dimension at least $|X|-|U|$. Let v be an element in V whose support $\operatorname{supp}(v) := \{x\in X: v(x)\neq 0\}$ has the largest cardinality among elements in V. Then

$$|\operatorname{supp}(v)|\ge \dim V\ge |X|-|U|.$$

Multiplying both sides of (13) by $v(x_j)$ and sum over $x_j\in X$, we obtain

$$\sum_{x_j\in X} v(x_j)F(x_1,\dots,x_n) = \sum_{a\in A} v(a)c_a \mathbb{1}_{a=x_1=\cdots=x_{j-1}=x_{j+1}=\cdots=x_n}.$$

By induction hypothesis, the partition rank of this diagonal function is the number of $a\in A$ such that $v(a)c_a\neq 0$. Thus

$$\text{partition-rank}\left(\sum_{x_j\in X} v(x_j)F(x_1,\dots,x_n)\right) = |\operatorname{supp}(v)\cap A| \ge |\operatorname{supp}(v)|+|A|-|X| \ge |A|-|U|. \quad (15)$$

On the other hand, we can do the similar operation on (14). Since $\sum_{x_j\in X} f_u(x_j)v(x_j)=0$ for all

doi:10.6342/NTU202200962

$u \in U$, we have

$$\sum_{x_j \in X} v(x_j) F(x_1, \ldots, x_n) = \sum_{i=1, i \notin U}^{r} \sum_{x_j \in X} v(x_j) c_i f_i(\vec{x}_{S_i}) g_i(\vec{x}_{T_i}).$$

Since S_i and T_i are disjoint for each i, every $\sum_{x_j \in X} v(x_j) c_i f_i(\vec{x}_{S_i}) g_i(\vec{x}_{T_i})$ are of partition rank at most 1. Thus

$$\text{partition-rank}\left(\sum_{x_j \in X} v(x_j) F(x_1, \ldots, x_n)\right) \le r - |U|. \tag{16}$$

By (15) and (16), we have $|A| \le r$, as desired. $\square$

Remark. Using the inequality partition-rank$(F) \le$ slice-rank(F), we can see that the above Theorem is stronger than Theorem 2.1.3. In practice, if we construct a function $F : X \times \cdots \times X \to \mathbb{F}$ for some finite set X, and F is diagonal with all its diagonal indexes are non-zero when X satisfies some constraint. Then both slice rank methods and partition rank methods can give us an upper bound for $|X|$. However, every upper bound we obtained from slice rank can be obtained from the partition rank. Thus the partition rank methods are better when we are dealing with diagonal functions.

4.2 Distinctness indicator function

We introduce a special function, which we called the distinctness indicator function. In previous sections, we've introduced the functions $F : X^k \to \mathbb{F}$ whose slice ranks or partition ranks are related to $|X|$. The distinctness indicator functions however, are functions whose partition rank is bounded above by a function of dimension k, instead of $|X|$.

Definition 4.2.1 (Distinctness indicator function)**.** *Let X be a finite set, $k \ge 2$ be an integer and $\mathbb{F}$ be a field whose characteristic is at least k. We define the **distinctness indicator function** $H_k : X^k \to \mathbb{F}$ as*

$$H_k(x_1, \ldots, x_k) = \begin{cases} 1 & \text{if } x_1, \ldots, x_k \text{are all distinct,} \\ (-1)^k (k-1)! & \text{if } x_1 = \cdots = x_k, \\ 0 & \text{otherwise.} \end{cases}$$

doi:10.6342/NTU202200962

The functions H_k can also be defined from the permutations in symmetric groups. Denote S_k to be the symmetric group of degree k. For any permutation $\sigma \in S_k$, we define

$$f_\sigma : X^k \to \mathbb{F}$$

to be the function that is 1 if $(x_1, \dots, x_k)$ is a fixed point of σ, and 0 otherwise. We also let $\mathrm{Cyc}_k \subset S_k$ be the set of all $k-$cycles in S_k. Then $|\mathrm{Cyc}_k| = (k-1)!$ by direct computation. Then we have the following equality:

Proposition 4.2.2. *[24] For $k \geq 2$, we have*

$$H_k(x_1, \dots, x_k) = \sum_{\sigma \in S_k, \sigma \notin \mathrm{Cyc}_k} \mathrm{sgn}(\sigma) f_\sigma(x_1, \dots, x_k),$$

where $\mathrm{sgn}(\sigma)$ *is the sign of the permutation* σ.

Proof. Note that the right hand side can be written as

$$\sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) f_\sigma(x_1, \dots, x_k) - \sum_{\sigma \in \mathrm{Cyc}_k} \mathrm{sgn}(\sigma) f_\sigma(x_1, \dots, x_k).$$

By definition of f_σ, we have

$$\sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) f_\sigma(x_1, \dots, x_k) = \sum_{\sigma \in S_k, \sigma \in \mathrm{Stab}(\vec{x})} \mathrm{sgn}(\sigma),$$

where $\mathrm{Stab}(\vec{x}) := \{\sigma \in S_k : (x_1, \dots, x_k)\text{is a fixed point of } \sigma\}$ is the stabilizer of $\vec{x}$. Since the stabilizer is a product of symmetric groups, the number of even cycle equals to the number of odd cycle if it

doi:10.6342/NTU202200962

is not the trivial group. Hence

$$\sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) f_\sigma(x_1, \ldots, x_k) = \begin{cases} 1 & \text{if } \operatorname{Stab}(\vec{x}) \text{is the trivial group,} \\ 0 & \text{otherwise.} \end{cases}$$

$$= \begin{cases} 1 & \text{if } x_1, \ldots, x_k \text{are distinct,} \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

Now we consider the remaining part

$$\sum_{\sigma \in \mathrm{Cyc}_k} \operatorname{sgn}(\sigma) f_\sigma(x_1, \ldots, x_k).$$

If $\sigma \in \mathrm{Cyc}_k$, then $(x_1, \ldots, x_k)$ is a fixed point of σ if and only if $x_1 = \cdots = x_k$. Also, for every k-cycle $\sigma \in \mathrm{Cyc}_k$, we have $\operatorname{sgn}(\sigma) = (-1)^{k-1}$. Thus

$$\sum_{\sigma \in \mathrm{Cyc}_k} \operatorname{sgn}(\sigma) f_\sigma(x_1, \ldots, x_k) = \begin{cases} \sum_{\sigma \in \mathrm{Cyc}_k} \operatorname{sgn}(\sigma) & \text{if } x_1 = \cdots = x_k, \\ 0 & \text{otherwise.} \end{cases}$$

$$= \begin{cases} (-1)^{k-1}(k-1)! & \text{if } x_1 = \cdots = x_k, \\ 0 & \text{otherwise.} \end{cases} \tag{18}$$

Hence by (17) and (18), we get our desired equality. □

Using this property, we can obtain an upper bound for the partition rank of H_k. We denote $\mathbb{1}$ to be the indicator function defined by

$$\mathbb{1}(\vec{x_S}) = \begin{cases} 1 & \text{if } \vec{x_S} = (x_{s_1}, \ldots, x_{s_{|S|}}) \text{with } x_{s_1} = \cdots = x_{s_{|S|}}, \\ 0 & \text{otherwise.} \end{cases}$$

For convenience, we set $\mathbb{1}(\vec{x_S}) = 1$ if $|S| = 1$.

doi:10.6342/NTU202200962

Proposition 4.2.3. *[24] For each $k \geq 2$, the function H_k can be written as*

$$H_k(x_1, \ldots, x_k) = \sum_{P \in \mathcal{P}_k} c_P \prod_{A \in P} \mathbb{1}(\vec{x}_A), \tag{19}$$

where $\mathcal{P}_k$ denotes the set of non-trivial partitions of $[k]$ and c_P are constants. Moreover, we have

$$\text{partition-rank}(H_k) \leq 2^{k-1} - 1.$$

Proof. For each $\sigma \in S_k$, the function f_σ can always be written as a product of indicator functions corresponding to the cycle notation of σ. For example, the function f_σ with $\sigma = (1)(23)(456) \in S_6$ can be written as

$$f_{(1)(23)(456)}(x_1, \ldots, x_6) = \mathbb{1}(x_1)\mathbb{1}(x_2, x_3)\mathbb{1}(x_4, x_5, x_6).$$

If $\sigma \notin \text{Cyc}_k$, then all the length of cycles in σ are strictly less than k. So each f_σ can be written as

$$f_\sigma(x_1, \ldots, x_k) = \prod_{A \in P} \mathbb{1}(\vec{x}_A)$$

for the corresponding non-trivial partition P of $[k]$. By Proposition 4.2.2, we have

$$H_k(x_1, \ldots, x_k) = \sum_{P \in \mathcal{P}_k} c_P \prod_{A \in P} \mathbb{1}(\vec{x}_A),$$

for some constants c_P.

Now we claim the second part. For each non-trivial partition $P \in \mathcal{P}_k$, there is a unique $A_1 \in P$ with $1 \in A_1 \subsetneq [k]$. Thus we obtain that

$$\prod_{A \in P} \mathbb{1}(\vec{x}_A) = \mathbb{1}(\vec{x}_{A_1}) \prod_{A \in P, A \neq A_1} \mathbb{1}(\vec{x}_A)$$

for such A_1. By sorting the terms of (19) for various A_1 satisfying $1 \in A_1 \subsetneq [k]$, we have

$$H_k(x_1, \ldots, x_k) = \sum_{\sigma \in S_k, \sigma \notin \text{Cyc}_k} \text{sgn}(\sigma) f_\sigma(x_1, \ldots, x_k) = \sum_{1 \in A_1 \subsetneq [k]} \mathbb{1}(\vec{x}_{A_1}) g_{A_1}(\vec{x}_{[k] \setminus A_1})$$

doi:10.6342/NTU202200962

for some functions g_{A_1}. Since the partition rank of each $\mathbb{1}(\vec{x}_{A_1})g_{A_1}(\vec{x}_{[k]\setminus A_1})$ is at most 1, we obtain that

$$\text{partition-rank}(H_k) \leq |\{A_1 : 1 \in A_1 \subsetneq [k]\}| = 2^{k-1} - 1.$$

□

4.3 Application 4: Largest k-right-corner-free subsets of $\mathbb{F}_p^n$

Recall that in Section 2.4, we use the slice rank methods to show that if q is an odd power and $A \subseteq \mathbb{F}_q^n$ is large enough, then A contains a right angle. In this section, we'll generalize the notion of right angle, which we call it the k-right corner, and prove the analogous result of Theorem 2.4.2.

Definition 4.3.1. *Given $k \geq 2$. An ordered pair of vectors $(x_1, \ldots, x_{k+1}) \in (\mathbb{F}_q^n)^{k+1}$ is called a **k-right corner** if they are distinct and the k-vectors $x_1 - x_{k+1}, \ldots, x_k - x_{k+1}$ are mutually orthogonal. In other words, $(x_1, \ldots, x_{k+1})$ is a **k-right corner** if and only if $x_1, \ldots, x_{k+1}$ are distinct and*

$$\langle x_i - x_{k+1}, x_j - x_{k+1}\rangle = 0 \quad \forall i, j \in [k], i \neq j.$$

We say a set A contains a k-right corner if $(x_1, \ldots, x_{k+1})$ is a k-right corner for some $x_1, \ldots, x_{k+1} \in A$.

Naslund used the partition rank methods to obtain a bound, polynomial in n, for the size of the largest subset of $\mathbb{F}_q^n$ that does not contain a k-right corner.

Theorem 4.3.2. *[24] Let $k \geq 3$ be given, and let $q = p^r$ be a prime power with $p > k$. If $A \subseteq \mathbb{F}_q^n$ satisfies*

$$|A| > \binom{n + (k-1)q}{(k-1)(q-1)},$$

then A contains a k-right corner.

doi:10.6342/NTU202200962

Proof. Define the function $R_{k+1} : \left(\mathbb{F}_q^n\right)^{k+1} \to \mathbb{F}_q$ by

$$R_{k+1}(x_1, \dots, x_k) = \begin{cases} 1 & x_1 - x_{k+1}, \dots, x_k - x_{k+1} \text{are mutually orthogonal,} \\ 0 & \text{otherwise.} \end{cases}$$

It can also be written as

$$R_{k+1}(x_1, \dots, x_k) = \prod_{1 \le i < j \le k} \left(1 - \langle x_i - x_{k+1}, x_j - x_{k+1} \rangle^{q-1}\right).$$

Note that $x_1, \dots, x_{k+1}$ are not required to be distinct. So this is not the indicator function of k-right corner yet. To handle this issue, we multiply it by the distinctness indicator function from Section 4.2, and define

$$J_k = H_{k+1} R_{k+1}.$$

Then the function J_k can be written as

$$J_k(x_1, \dots, x_{k+1}) = \begin{cases} 1 & \text{if } (x_1, \dots, x_{k+1}) \text{is a } k\text{-right corner,} \\ (-1)^k k! & \text{if } x_1 = \cdots = x_{k+1}, \\ 0 & \text{otherwise.} \end{cases}$$

If $A \subseteq \mathbb{F}_q^n$ doesn't contain any k-right corner, then $J_k|_{A^{k+1}}$ is a diagonal function. Note that the diagonal indexes $(-1)^k k!$ are non-zero in $\mathbb{F}_q$ since $p > k$. Thus by Theorem 4.1.4, we have

$$|A| \le \text{partition-rank}(J_k).$$

So it suffices to show that partition-rank$(J_k) \le \binom{n+(k-1)q}{(k-1)(q-1)}$. By Proposition 4.2.3, we could write J_k as a linear combination of the functions of the following form:

$$R_{k+1}(x_1, \dots, x_{k+1}) \prod_{A \in P} \mathbb{1}(\vec{x}_A), \tag{20}$$

where P is a non-trivial partition of $[k+1]$. The following lemma allows us to rewrite it into a lower

doi:10.6342/NTU202200962

degree polynomial.

Lemma 4.3.3. *[24] Let $P = \{A_1, \ldots, A_r\}$ be a non-trivial partition of $[k+1]$. Suppose without loss of generality that $k+1 \in A_r$, we let $a_r = k+1$. For each $i = 1, \ldots, r-1$, we let $a_i = \min A_i$ be the minimal element of A_i. Then*

$$R_{k+1}(x_1, \ldots, x_{k+1}) \prod_{A \in P} \mathbb{1}(\vec{x}_A) = R_r(x_{a_1}, \ldots, x_{a_r}) \prod_{i=1}^{r} \mathbb{1}(\vec{x}_{A_i}) \Pi_2^P, \tag{21}$$

where

$$\Pi_2^P = \prod_{|A_i| \geq 2, i \neq r} \left(1 - \langle x_{a_i} - x_{a_r}, x_{a_i} - x_{a_r} \rangle^{q-1}\right).$$

Proof. By the definition of R_{k+1}, (20) can be written as

$$\left(\prod_{i=1}^{r} \mathbb{1}(\vec{x}_{A_i})\right) \left(\prod_{1 \leq j < l \leq k} \left(1 - \langle x_j - x_{a_r}, x_l - x_{a_r} \rangle^{q-1}\right)\right). \tag{22}$$

Observe that for a set A, $a \in A$ and a function Q on $|A|$ variables, we have $\mathbb{1}(\vec{x}_A) Q(\vec{x}_A) = \mathbb{1}(\vec{x}_A) \tilde{Q}(a)$ for a single variable function $\tilde{Q}$ defined by $\tilde{Q}(a) = Q(a, \ldots, a)$. Using this property and the equality

$$(1 - \langle x, y \rangle^{q-1})^2 = (1 - \langle x, y \rangle^{q-1}),$$

we can rewrite (22) into

$$\left(\prod_{i=1}^{r} \mathbb{1}(\vec{x}_{A_i})\right) \left(\prod_{1 \leq j < l \leq r} \left(1 - \langle x_{a_j} - x_{a_r}, x_{a_l} - x_{a_r} \rangle^{q-1}\right)\right) \left(\prod_{i=1, |A_i| \geq 2}^{r-1} \left(1 - \langle x_{a_i} - x_{a_r}, x_{a_i} - x_{a_r} \rangle^{q-1}\right)\right).$$

The second product is just the function

$$R_r(x_{a_1}, \ldots, x_{a_r})$$

and the third product is the definition of Π_2^P. So the lemma holds. □

Let $\mathrm{Poly}_d^2(\mathbb{F}_q^n)$ be the polynomial space defined in the proof of Theorem 2.4.2, and we also define

doi:10.6342/NTU202200962

another polynomial space $\mathrm{Poly}_d(\mathbb{F}_q^n)$ as the space of $n-$variable polynomials over the field $\mathbb{F}_q$ of total degree at most d. Then

$$\dim \mathrm{Poly}_d(\mathbb{F}_q^n) = \binom{n+d}{d}. \tag{23}$$

Applying the equation (21), we have the following lemma

Lemma 4.3.4. *[24] Let $P = \{A_1, \ldots, A_r\}$ be a non-trivial partition of $[k+1]$. Suppose without loss of generality that $k+1 \in A_r$, we let $a_r = k+1$. For each $i = 1, \ldots, r-1$, we let $a_i = \min A_i$ be the minimal element of A_i. Then*

$$R_{k+1}(x_1, \ldots, x_{k+1}) \prod_{i=1}^{r} \mathbb{1}(\vec{x}_{A_i}) = \sum_j \left[\prod_{i=1}^{r} \mathbb{1}(\vec{x}_{A_i}) Q_{i,j}(x_{a_i})\right],$$

where for each $1 \le i \le r-1$,

$$Q_{i,j} \in \begin{cases} \mathrm{Poly}_d(\mathbb{F}_q^n) \text{with } d = (r-2)(q-1) & \text{if } |A_i| = 1, \\ \mathrm{Poly}_d^2(\mathbb{F}_q^n) \text{with } d = (r-1)(q-1) & \text{if } |A_i| \ge 2. \end{cases}$$

Proof. By expanding the product of (21), we may write

$$R_r(x_{a_1}, \ldots, x_{a_r}) \prod_{i=1}^{r} \mathbb{1}(\vec{x}_{A_i}) \Pi_2^P = \sum_j \left[\prod_{i=1}^{r} \mathbb{1}(\vec{x}_{A_i}) Q_{i,j}(x_{a_i})\right]$$

for some polynomials $Q_{i,j}$. Recall that

$$R_r(x_{a_1}, \ldots, x_{a_r}) = \prod_{1 \le j < l \le r-1} \left(1 - \langle x_{a_j} - x_{a_r}, x_{a_l} - x_{a_r}\rangle^{q-1}\right).$$

For any $1 \le i \le r-1$, there are exactly $r-2$ terms each of degree $q-1$ in the product definition of R_r that contain x_{a_i}, namely $\left(1 - \langle x_{a_j} - x_{a_r}, x_{a_l} - x_{a_r}\rangle^{q-1}\right)$ for $i = j$ or $i = l$. If $|A_i| = 1$, then x_{a_i} doesn't appear in Π_2^P, which implies that $\deg Q_{i,j} \le (r-2)(q-1)$ and hence

$$Q_{i,j} \in \mathrm{Poly}_d(\mathbb{F}_q^n) \text{with } d = (r-2)(q-1) \text{if } |A_i| = 1.$$

doi:10.6342/NTU202200962

If $|A_i| \geq 2$, then the product Π_2^P will contain an additional term that contains x_{a_i}, namely

$$\left(1 - \langle x_{a_i} - x_{a_r}, x_{a_i} - x_{a_r}\rangle^{q-1}\right).$$

If we expand the it in terms of the $(n+1)$ variables $x_{a_i,1}, \ldots, x_{a_i,n}$ and $x_{a_i,1}^2 + \cdots + x_{a_i,n}^2$, then it will have degree $q-1$ only. Thus

$$Q_{i,j} \in \mathrm{Poly}_d^2(\mathbb{F}_q^n)\text{with } d = (r-1)(q-1)\text{if } |A_i| \geq 2.$$

□

Using this lemma, we can obtain an upper bound for partition-rank(J_k).

Proposition 4.3.5. *[24] Let $\mathcal{P}_{k+1}$ be the set of non-trivial partitions of $[k+1]$. Suppose that $B_1, \ldots, B_l$ is a sequence of non-empty subsets of $[k]$ such that for every $P \in \mathcal{P}_{k+1}$, there exists $A \in P$ with $k+1 \notin A$ and $A = B_i$ for some i. For each i, set*

$$r_i = \max\{|P| : P \in \mathcal{P}_{k+1}, B_i \in P, \text{and } B_j \notin P \text{for } j < i\}. \tag{24}$$

and let

$$\mathcal{V}_i = \begin{cases} \mathrm{Poly}_{(r_i-2)(q-1)}(\mathbb{F}_q^n) & \text{if } |B_i| = 1, \\ \mathrm{Poly}^2_{(r_i-1)(q-1)}(\mathbb{F}_q^n) & \text{if } |B_i| \geq 2. \end{cases}$$

Then we have

$$\text{partition-rank}(J_k) \leq \sum_{i=1}^{l} \dim \mathcal{V}_i. \tag{25}$$

Proof. For any $P \in \mathcal{P}_{k+1}$ there exists some index j such that $B_j \in P$ by assumption. Let i be the minimal index such that $B_i \in P$. Then $|P| \leq r_i$ by definition of r_i. Together with Lemma 4.3.4, we have

$$R_{k+1}(x_1, \ldots, x_{k+1}) \prod_{A \in P} \mathbb{1}(\vec{x}_A) = \sum_j \mathbb{1}(\vec{x}_{B_i}) Q_j(x_{b_i}) T_j(\vec{x}_{[k+1]\setminus B_i}),$$

where for each j, $Q_j(x_{b_i}) \in \mathcal{V}_i$. and T_j consists of the product $\prod_{A \in P, A \neq B_i} \mathbb{1}(\vec{x}_A)$. Thus J_k is a linear

doi:10.6342/NTU202200962

combination of the terms of the form

$$\mathbb{1}(\vec{x}_{B_i})Q(x_{b_i})T(\vec{x}_{[k+1]\backslash B_i}),$$

where $Q \in \mathcal{V}_i$. Note that $T(\vec{x}_{[k+1]\backslash B_i})$ involves only the variables x_i for $i \in [k+1]\backslash B_i$. Since B_i is a non-empty proper subset of $[k+1]$, each $\mathbb{1}(\vec{x}_{B_i})Q(x_{b_i})T(\vec{x}_{[k+1]\backslash B_i})$ is of partition rank at most 1. Together with $Q(x_{b_i}) \in \mathcal{V}_i$, this implies (25). □

Now we can complete the proof of Theorem 4.3.2. Let the sequence of subsets $B_1, \ldots, B_{2^k-1}$ be defined by listing all the non-empty subsets of $[k]$ in order by their cardinality with tie broken by lexicographical order of the elements of the set. That is,

$$B_1 = \{1\}, B_2 = \{2\}, \ldots, B_k = \{k\}, B_{k+1} = \{1,2\}, \ldots, B_{2^k-1} = \{1, \ldots, k\}.$$

By Proposition 4.3.5, (23) and Lemma 2.4.3, we have the upper bound

$$\begin{aligned}
&\text{partition-rank}(J_k)\\
\leq &\sum_{i=1}^{l} \dim \mathcal{V}_i\\
= &\sum_{i=1}^{k} \dim \text{Poly}_{(r_i-2)(q-1)}(\mathbb{F}_q^n) + \sum_{i=k+1}^{2^k-1} \dim \text{Poly}^2_{(r_i-1)(q-1)}(\mathbb{F}_q^n)\\
= &\sum_{i=1}^{k} \binom{n+(r_i-2)(q-1)}{(r_i-2)(q-1)} + \sum_{i=k+1}^{2^k-1} \left[\binom{n+(r_i-1)(q-1)}{(r_i-1)(q-1)} + \binom{n+(r_i-1)(q-1)-1}{(r_i-1)(q-1)-1}\right],
\end{aligned}$$

where r_i is defined in (24). To obtain the value of r_i, we shall find the largest partitions $P \in \mathcal{P}_{k+1}$ with the property in the definition of r_i. For $i = 1$, the partition $P = \{\{1\}, \{2\}, \ldots, \{k+1\}\}$ satisfies the criteria $B_1 \in P$ and this is the partition with the largest cardinality, so $r_1 = k+1$.

For $i = 2, \ldots, k$, since $B_1, \ldots, B_{i-1}$ can not contained in the partition, it can only has at most $k+1-(i-1)$ singletons $\{i\}, \{i+1\}, \ldots, \{k+1\}$. The first $i-1$ elements $1, \ldots, i-1$ can form at most $\lfloor \frac{i-1}{2} \rfloor$ pairs. Thus

$$r_i \leq (k+1) - (i-1) + \lfloor \frac{i-1}{2} \rfloor \leq k.$$

doi:10.6342/NTU202200962

Thus we can bound the first sum:

$$\sum_{i=1}^{k} \binom{n+(r_i-2)(q-1)}{(r_i-2)(q-1)} \le \binom{n+(k-1)(q-1)}{(k-1)(q-1)} + (k-1)\binom{n+(k-2)(q-1)}{(k-2)(q-1)}. \tag{26}$$

Consider $i \ge k+1$, say $|B_i| = j \ge 2$. Then the partition can not contain any sets of size less than j that do not contain $k+1$. Thus

$$r_i \le 1 + \lfloor \frac{k}{j} \rfloor.$$

Thus we can bound the second sum

$$\begin{aligned} &\sum_{i=k+1}^{2^k-1} \left[\binom{n+(r_i-1)(q-1)}{(r_i-1)(q-1)} + \binom{n+(r_i-1)(q-1)-1}{(r_i-1)(q-1)-1} \right] \\ \le &\sum_{j=2}^{k} \binom{k}{j} \left[\binom{n+\lfloor \frac{k}{j} \rfloor(q-1)}{\lfloor \frac{k}{j} \rfloor(q-1)} + \binom{n+\lfloor \frac{k}{j} \rfloor(q-1)-1}{\lfloor \frac{k}{j} \rfloor(q-1)-1} \right]. \end{aligned}$$

We use the following weak bound for convenience.

$$\binom{n+\lfloor \frac{k}{j} \rfloor(q-1)}{\lfloor \frac{k}{j} \rfloor(q-1)} + \binom{n+\lfloor \frac{k}{j} \rfloor(q-1)-1}{\lfloor \frac{k}{j} \rfloor(q-1)-1} \le \binom{n+1+\lfloor \frac{k}{j} \rfloor(q-1)}{\lfloor \frac{k}{j} \rfloor(q-1)},$$

This gives us an upper bound for the second sum

$$\sum_{i=k+1}^{2^k-1} \left[\binom{n+(r_i-1)(q-1)}{(r_i-1)(q-1)} + \binom{n+(r_i-1)(q-1)-1}{(r_i-1)(q-1)-1} \right] \le \sum_{j=2}^{k} \binom{k}{j} \binom{n+1+\lfloor \frac{k}{j} \rfloor(q-1)}{\lfloor \frac{k}{j} \rfloor(q-1)}. \tag{27}$$

To combine the upper bound of these two sums, we use the Vandermonde identity, which states that for any non-negative integers a, b, c, we have

$$\binom{a+b}{c} = \sum_{j=0}^{c} \binom{a}{j} \binom{b}{c-j}.$$

Let $a = k, b = n+(k-1)(q-1)-1$ and $c = (k-1)(q-1)$, we have

$$\binom{n+(k-1)q}{(k-1)(q-1)} = \sum_{j=0}^{k} \binom{k}{j} \binom{n+(k-1)(q-1)-1}{(k-1)(q-1)-j}, \tag{28}$$

doi:10.6342/NTU202200962

where the sum ends at k since $\binom{k}{j} = 0$ for $j > k$. The terms for $j = 0, 1$ can be written as

$$\begin{aligned}&\binom{n + (k-1)(q-1) - 1}{(k-1)(q-1)} + k\binom{n + (k-1)(q-1) - 1}{(k-1)(q-1) - 1}\\ =&\binom{n + (k-1)(q-1)}{(k-1)(q-1)} + (k-1)\binom{n + (k-1)(q-1) - 1}{(k-1)(q-1) - 1},\end{aligned}$$

which is an upper bound for the first sum by (26). For $k \geq 3$ and $j \geq 2$, we claim that

$$\lfloor \frac{k}{j} \rfloor (q-1) \leq (k-1)(q-1) - j. \tag{29}$$

For $k \geq 3$ and $2 \leq j \leq k-1$, we have $\lfloor \frac{k}{j} \rfloor \leq k - j$ and thus

$$\lfloor \frac{k}{j} \rfloor (q-1) \leq (k-j)(q-1) = (k-1)(q-1) - (j-1)(q-1) \leq (k-1)(q-1) - j.$$

For $k \geq 3$ and $j = k$, we also have

$$\lfloor \frac{k}{j} \rfloor (q-1) = q - 1 \leq (k-2)(q-1) \leq (k-1)(q-1) - j$$

since $q - 1 \geq k = j$ by assumption. Therefore for every $k \geq 3$ and $j \geq 2$, we can apply the inequality (29) and obtain that

$$\binom{n + 1 + \lfloor \frac{k}{j} \rfloor (q-1)}{\lfloor \frac{k}{j} \rfloor (q-1)} \leq \binom{n + 1 + (k-1)(q-1) - j}{(k-1)(q-1) - j}.$$

Together with (26), (27) and (28), we have

$$\text{partition-rank}(J_k) \leq \sum_i \dim \mathcal{V}_i \leq \binom{n + (k-1)q}{(k-1)(q-1)}.$$

This completes the proof. □

doi:10.6342/NTU202200962

5 Right Angle Removal Lemma

5.1 Background

In graph theory, the triangle removal lemma of Ruzsa and Szemerédi [27] states that any graph on n vertices with $o(n^3)$ triangles can be made triangle-free by removing $o(n^2)$ edges. This lemma can be proved by the Szemerédi regularity lemma [10, 30]. Green [13] proved the analogous theorem for abelian group. Let G be an abelian group with cardinality N, and let $A \subseteq G$. A triple $(x, y, z) \in A^3$ is called a **triangle** if $x + y + z = 0$. Then the arithmetic triangle removal lemma states that

Theorem 5.1.1. *[13] Suppose that $A \subseteq G$ is a set with $o(N^2)$ triangles. Then we may remove $o(N)$ elements from A to leave a set which is triangle-free.*

In fact, Green's result is more general:

Theorem 5.1.2. *[13] Let $k \geq 3$ be a fixed integer, and suppose that $A_1, \ldots, A_k$ are subsets of G such that there are $o(N^{k-1})$ solutions to the equation $a_1 + \cdots + a_k = 0$ with $a_i \in A_i$ for all i. Then we may remove $o(N)$ elements from each A_i so as to leave sets A'_i, such that there are no solutions to $a'_1, \ldots, a'_k = 0$ with $a'_i \in A'_i$ for each i.*

5.2 Main result and its proof

Throughout this section, the number q denotes a prime power and the logarithm function log is of base 2. We consider $\mathbb{F}_q^n$, which is an n-dimensional vector space over the finite field $\mathbb{F}_q$. Denote $N = |\mathbb{F}_q^n| = q^n$. Our main result is the following.

Theorem 5.2.1 (Right angle removal lemma)**.** *Suppose that X, Y, Z are subsets of $\mathbb{F}_q^n$ such that there are $o(N^3(\log N)^{-(4q-4)/3})$ solutions to the equation $\langle x - z, y - z\rangle = 0$, with $x \in X, y \in Y, z \in Z$. Then we may remove $o(N)$ elements from each X, Y, Z so as to leave sets X', Y', Z', respectively, such that there are no solutions to $\langle x' - z', y' - z'\rangle = 0$, with $x' \in X', y' \in Y', z' \in Z'$.*

Note that this is similar to Theorem 5.1.2 except that we replace the *linear* equation $a_1 + \cdots + a_k = 0$ with a *non-linear* equation $\langle x - z, y - z\rangle = 0$, where $\langle \cdot, \cdot \rangle$ denotes the dot product in the vector space $\mathbb{F}_q^n$. Recall that we say the triple (x, y, z) is a right angle if x, y, z are distinct and satisfy the equation.

doi:10.6342/NTU202200962

Thus we called this result the "right angle removal lemma."

To prove this result, we introduce some definitions and notations.

Definition 5.2.2. *We say a triple $(x, y, z) \in \mathbb{F}_q^n \times \mathbb{F}_q^n \times \mathbb{F}_q^n$ is a **weak right angle** if*

$$\langle x - z, y - z \rangle = 0.$$

Remark. Note that the definition of weak right angle is similar to right angle (Definition 2.4.1) except that x, y, z are not required to be distinct here.

For $X, Y, Z \subseteq \mathbb{F}_q^n$, we define

$$\eta(X, Y, Z) = |\{(x, y, z) \in X \times Y \times Z \mid (x, y, z) \text{is a weak right angle.}\}|$$

to be the number of all the weak right angles in $X \times Y \times Z$. The following theorem is a quantitative version of our main result Theorem 5.2.1. We'll prove this statement.

Theorem 5.2.3. *Given any $0 < \epsilon < 1$. Let $\delta = \epsilon^3$ and let n be a large integer so that*

$$(2q - 5)/3 < n \quad \text{and} \quad (n + (2q - 5)/3)^{(2q-2)/3} \leq \frac{11}{36} \epsilon q^n. \tag{30}$$

Denote $N = |\mathbb{F}_q^n| = q^n$. If $X, Y, Z \subseteq \mathbb{F}_q^n$ satisfying

$$\eta(X, Y, Z) \leq \delta C_q \frac{N^3}{(\log N)^{(4q-4)/3}},$$

where C_q is a positive constant depend only on q. Then there exist $X' \subseteq X, Y' \subseteq Y$, $Z' \subseteq Z$ with

$$|X \backslash X'| + |Y \backslash Y'| + |Z \backslash Z'| \leq \epsilon N$$

such that

$$\eta(X', Y', Z') = 0.$$

doi:10.6342/NTU202200962

Remark. The constant C_q can be chosen to be

$$C_q = \frac{1}{243}\left(\frac{\log q}{2}\right)^{(4q-4)/3}.$$

Note that $\eta(X, Y, Z)$ is greater than or equal to the number of all (not necessarily weak) right angles in $X \times Y \times Z$. But the theorem can be applied to right angles too. Indeed, the number of triples (x, y, z) with x, y, z are not all distinct is less than $3N^2$, which increases slower than $\Theta(\frac{N^3}{(\log N)^{(4q-4)/3}})$ as $n \to \infty$. Thus the theorem is still valid if replace "weak right angle" with simply "right angle."

Theorem 5.2.4. *Given any $0 < \epsilon < 1$. Let $\delta = \epsilon^3$ and let n be an integer satisfying* (30). *Denote $N = |\mathbb{F}_q^n| = q^n$. Let m be an integer with $m > \epsilon N$. Let $X = \{x^{(i)}\}_{i=1}^m, Y = \{y^{(i)}\}_{i=1}^m, Z = \{z^{(i)}\}_{i=1}^m$ be three subsets of $\mathbb{F}_q^n$ with m elements. If for each $i \in \{1, \ldots, m\}$, we have*

$$(x^{(i)}, y^{(i)}, z^{(i)})\text{is a weak right angle,}$$

then

$$\eta(X, Y, Z) > \delta C_q \frac{N^3}{(\log N)^{(4q-4)/3}}.$$

We can prove Theorem 5.2.3 if Theorem 5.2.4 holds.

Proof of Theorem 5.2.3 assuming Theorem 5.2.4. Suppose we can not remove ϵN elements from $X, Y, Z \subseteq \mathbb{F}_q^n$ so that there are no weak right angle triples (x, y, z) remain, then there are more than ϵN weak right angles $(x^{(i)}, y^{(i)}, z^{(i)}) \in X \times Y \times Z$ with $|\{x^{(i)}\}_{i=1}^m| = |\{y^{(i)}\}_{i=1}^m| = |\{z^{(i)}\}_{i=1}^m| = m$. By Theorem 5.2.4,

$$\eta(X, Y, Z) \geq \eta(\{x^{(i)}\}_{i=1}^m, \{y^{(i)}\}_{i=1}^m, \{z^{(i)}\}_{i=1}^m) > \delta C_q \frac{N^3}{(\log N)^{(4q-4)/3}}.$$

This proves Theorem 5.2.3. □

Hence it suffices to prove Theorem 5.2.4. The following lemma uses the slice rank methods.

Lemma 5.2.5. *Given a collection of triples $\{(x^{(i)}, y^{(i)}, z^{(i)})\}_{i=1}^m$ in $\mathbb{F}_q^n$ such that for $i, j, k \in \{1, \ldots, m\}$,*

doi:10.6342/NTU202200962

we have

$$(x^{(i)}, y^{(j)}, z^{(k)}) \text{is a weak right angle} \quad \text{if and only if} \quad i = j = k.$$

Then the size of the collection satisfies the bound

$$m \leq 3\binom{n + (2q-5)/3}{(2q-2)/3} \leq 3(n + (2q-5)/3)^{(2q-2)/3}.$$

Proof. Let X, Y and Z to be the sets $\{x^{(1)}, \dots, x^{(m)}\}, \{y^{(1)}, \dots, y^{(m)}\}$ and $\{z^{(1)}, \dots, z^{(m)}\}$ respectively. Consider the function $F : \left(\mathbb{F}_q^n\right)^3 \to \mathbb{F}_q^n$ defined by

$$F(x, y, z) = \begin{cases} 1 & \text{if } (x, y, z) \text{is a weak right angle,} \\ 0 & \text{otherwise.} \end{cases}$$

Then $F|_{X \times Y \times Z}$ is diagonal by assumption. Thus $m \leq \text{slice-rank}(F)$. On the other hand, F can be written as

$$F(x, y, z) = 1 - \langle x - z, y - z \rangle^{q-1},$$

whose expansion is a linear combination the following monomials

$$x_1^{d_1} \cdots x_n^{d_n} y_1^{e_1} \cdots y_n^{e_n} z_1^{f_1} \cdots z_n^{f_n},$$

where x_i, y_i, z_i denotes the ith coordinate of x, y, z, for $i = 1, \dots, n$ respectively, and each d_i, e_j, f_k are non-negative integers satisfying

$$\sum_{i=1}^{n} d_i + \sum_{i=1}^{n} e_i + \sum_{i=1}^{n} f_i \leq 2(q-1).$$

doi:10.6342/NTU202200962

Thus F can be written as

$$F(x,y,z) = \sum_{\sum d_i \le (2q-2)/3} \left[x_1^{d_1} \cdots x_n^{d_n}\right] \alpha_{d_1,\dots,d_n}(y,z) + \sum_{\sum e_i \le (2q-2)/3} \left[y_1^{e_1} \cdots y_n^{e_n}\right] \beta_{e_1,\dots,e_n}(x,z) + \sum_{\sum f_i \le (2q-2)/3} \left[z_1^{f_1} \cdots z_n^{f_n}\right] \gamma_{f_1,\dots,f_n}(x,y).$$

So

$$\begin{aligned} \text{slice-rank}(F) &\le 3|\{(d_1,\dots,d_n) : \sum_{i=1}^{n} d_i \le (2q-2)/3\}| \\ &\le 3\binom{n + \lfloor(2q-2)/3\rfloor - 1}{\lfloor(2q-2)/3\rfloor} \\ &\le 3(n+(2q-5)/3)^{(2q-2)/3}, \end{aligned}$$

as desired. □

Now we use a particular property in graph theory. A **three-uniform hypergraph** is a hypergraph whose edges contain exactly 3 distinct vertices. For a hypergraph H, we denote $\alpha(H)$ to be the size of the largest independent set (i.e., a set of vertices containing no edges). We will use the following lemma. The proof uses probability method.

Lemma 5.2.6. *[1] Every three-uniform hypergraphs $H = (V,E)$ with $|E| \ge |V|/3$ satisfy the inequality*

$$\alpha(H) \ge \frac{2|V|^{3/2}}{3\sqrt{3|E|}}.$$

Proof. Let $H' = (V', E')$ be a random subhypergraph of H induced by $V' \subseteq V$, where every vertices occurred in V has independent probability p to be in V'. The probability $p \in [0,1]$ will be determined later. Since H' is induced by the vertices set V', we have

$$\mathbb{E}[|V'|] = |V|p \quad \text{and} \quad \mathbb{E}[|E'|] = |E|p^3.$$

doi:10.6342/NTU202200962

By deleting one vertex from each edge, the remaining vertices will be an independent set. Thus

$$\alpha(H) \geq \mathbb{E}[\alpha(H')] \geq \mathbb{E}[|V'|] - \mathbb{E}[|E'|] = |V|p - |E|p^3$$

for all $0 < p < 1$. By differentiate it with respect to p, we obtain that $p = \sqrt{\frac{|V|}{3|E|}} \in [0,1]$ gives us an extreme value. Thus

$$\alpha(H) \geq \mathbb{E}[\alpha(H')] \geq |V|\sqrt{\frac{|V|}{3|E|}} - |E|\left(\sqrt{\frac{|V|}{3|E|}}\right)^3 = \frac{2|V|^{3/2}}{3\sqrt{3|E|}}.$$

□

Now we have all the tools to prove Theorem 5.2.4.

Proof of Theorem 5.2.4. We define a three-uniform hypergraph $H = (V, E)$. Let the vertex set be $V = \{\pm 1, \dots, \pm m\}$. In particular, $|V| = 2m$. The edge set is defined as follow:

$$E = \{\{i, j, k\} \mid (x_{|i|}, y_{|j|}, z_{|k|}) \text{is a weak right angle and } i, j, k \text{are distinct.}\}$$

We note that only i, j, k needs to be distinct, so $i = -j$ is allowed. Since every weak right angles in $X \times Y \times Z$ contribute at most $2^3 = 8$ pairs to the edge set, we have the inequality

$$|E| \leq 8\eta(X, Y, Z). \tag{31}$$

For convenience, let

$$\gamma(n, q) := 3(n + (2q-5)/3)^{(2q-2)/3}.$$

which is the upper bound appeared in Lemma 5.2.5. Hence second condition of (30) shows that

$$\gamma(n, q) \leq \frac{11}{12}\epsilon N < \frac{11}{12}m \tag{32}$$

Consider the set $\{1, 2, \dots, \gamma(n, q) + 1\} \subseteq V$. By Lemma 5.2.5, it contains at least one edge. Without loss of generality, say it is $\{1, 2, 3\}$. We then consider the set $\{2, 3, \dots, \gamma(n, q) + 2\}$. Continue this

doi:10.6342/NTU202200962

process, we obtain $m-\gamma(n,q)$ edges in $\{1,\dots,m\}$. Note that if $\{i,j,k\}\in E$ with $i,j,k>0$, then $\{e_1 i, e_2 j, e_3 k\}\in E$ for all $e_1,e_2,e_3\in\{\pm 1\}$. Thus, we have

$$|E| \geq 8(m-\gamma(n,q)) > \frac{2m}{3} = \frac{|V|}{3},$$

where the second inequality follows from the condition (32). This allows us to use Lemma 5.2.6 and obtain the following inequality

$$\frac{2|V|^{3/2}}{3\sqrt{3|E|}} \leq \alpha(H). \tag{33}$$

For any independent set $I\subseteq V$, If $\{i,j,k\}\in I$ such that $(x_{|i|}, y_{|j|}, z_{|k|})$ is a weak right angle, then $|i|=|j|=|k|$. Thus by Lemma 5.2.5 again, we have

$$\frac{1}{2}\alpha(H) \leq \gamma(n,q). \tag{34}$$

By (33) and (34), we have

$$\frac{|V|^3}{27\gamma(n,q)^2} \leq |E|.$$

Using $|V| = 2m = 2\epsilon N$ and the inequality (31) of $|E|$, we obtain

$$\frac{\epsilon^3 N^3}{27\gamma(n,q)^2} \leq \eta(X,Y,Z).$$

Use the inequality (30), we have $n+(2q-5)/3 < 2n$ and thus

$$\eta(X,Y,Z) \geq \frac{\epsilon^3}{243}\frac{N^3}{(n+(2q-5)/3)^{(4q-4)/3}} > \frac{\epsilon^3}{243}\frac{N^3}{(2n)^{(4q-4)/3}} = \frac{\delta}{243}\left(\frac{\log q}{2}\right)^{(4q-4)/3}\frac{N^3}{(\log N)^{(4q-4)/3}},$$

as desired. □

Acknowledgement

I would like to thank my supervisor professor Chun-Yen Shen for many helpful suggestions and comments. I would also like to thank my family for their continuous support.

doi:10.6342/NTU202200962

References

[1] N. Alon and J. H. Spencer, *The Probabilistic Method*, Wiley Publishing, 4th ed., 2016.

[2] E. Bannai and E. Bannai, *An upper bound for the cardinality of an s-distance subset in real Euclidean space*, Combinatorica, 1 (1981), pp. 99–102.

[3] M. Bateman and N. H. Katz, *New bounds on cap sets*, J. Amer. Math. Soc., 25 (2012), pp. 585–613.

[4] M. Bennett, *Occurrence of right angles in vector spaces over finite fields*, European J. Combin., 70 (2018), pp. 155–163.

[5] B. Bukh, *Non-trivial solutions to a linear equation in integers*, Acta Arith., 131 (2008), pp. 51–55.

[6] E. Croot, V. F. Lev, and P. P. Pach, *Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small*, Ann. of Math. (2), 185 (2017), pp. 331–337.

[7] Y. Edel, *Extensions of generalized product caps*, Des. Codes Cryptogr., 31 (2004), pp. 5–14.

[8] J. S. Ellenberg and D. Gijswijt, *On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression*, Ann. of Math. (2), 185 (2017), pp. 339–343.

[9] C. Elsholtz and P. P. Pach, *Caps and progression-free sets in $\mathbb{Z}_m^n$*, Des. Codes Cryptogr., 88 (2020), pp. 2133–2170.

[10] J. Fox, *A new proof of the graph removal lemma*, Ann. of Math. (2), 174 (2011), pp. 561–579.

[11] P. Frankl, R. L. Graham, and V. Rödl, *Quantitative theorems for regular systems of equations*, J. Combin. Theory Ser. A, 47 (1988), pp. 246–261.

[12] G. Ge and C. Shangguan, *Maximum subsets of $\mathbb{F}_q^n$ containing no right angles*, J. Algebraic Combin., 52 (2020), pp. 455–460.

[13] B. Green, *A Szemerédi-type regularity lemma in abelian groups, with applications*, Geom. Funct. Anal., 15 (2005), pp. 340–376.

doi:10.6342/NTU202200962

[14] B. Green and T. Tao, *An inverse theorem for the Gowers $U^3(G)$ norm*, Proc. Edinb. Math. Soc. (2), 51 (2008), pp. 73–153.

[15] ——, *New bounds for Szemerédi's theorem. I. Progressions of length 4 in finite field geometries*, Proc. Lond. Math. Soc. (3), 98 (2009), pp. 365–392.

[16] ——, *New bounds for Szemerédi's theorem. II. A new bound for $r_4(N)$*, in Analytic number theory, Cambridge Univ. Press, Cambridge, 2009, pp. 180–204.

[17] P. Halmos, *Finite-Dimensional Vector Spaces*, Springer, New York, NY, 2 ed., 1987.

[18] R. Hancock and A. Treglown, *On solution-free sets of integers*, European J. Combin., 66 (2017), pp. 110–128.

[19] ——, *On solution-free sets of integers II*, Acta Arith., 180 (2017), pp. 15–33.

[20] Y. Lin and J. Wolf, *On subsets of $\mathbb{F}_q^n$ containing no k-term progressions*, European J. Combin., 31 (2010), pp. 1398–1403.

[21] R. Meshulam, *On subsets of finite abelian groups with no 3-term arithmetic progressions*, J. Combin. Theory Ser. A, 71 (1995), pp. 168–172.

[22] M. Mimura and N. Tokushige, *Avoiding a shape, and the slice rank method for a system of equations*. preprint on webpage at `https://arxiv.org/abs/1909.10509`, 2019.

[23] ——, *Solving linear equations in a vector space over a finite field*, Discrete Math., 344 (2021), pp. Paper No. 112603, 11.

[24] E. Naslund, *The partition rank of a tensor and k-right corners in $\mathbb{F}_q^n$*, J. Combin. Theory Ser. A, 174 (2020), pp. 105190, 25.

[25] I. Z. Ruzsa, *Solving a linear equation in a set of integers. I*, Acta Arith., 65 (1993), pp. 259–282.

[26] ——, *Solving a linear equation in a set of integers. II*, Acta Arith., 72 (1995), pp. 385–397.

[27] I. Z. Ruzsa and E. Szemerédi, *Triple systems with no six points carrying three triangles*, Combinatorica, 18 (1976), pp. 939–945.

doi:10.6342/NTU202200962

[28] L. Sauermann, *Finding solutions with distinct variables to systems of linear equations over* $\mathbb{F}_p$. preprint on webpage at `https://arxiv.org/abs/2105.06863`, 2021.

[29] W. Sawin and T. Tao, *Notes on the "slice rank" of tensors*, blog post, (2016). `https://terrytao.wordpress.com/2016/08/24/notes-on-the-slice-rank-of-tensors/`.

[30] E. Szemerédi, *Regular partitions of graphs*, in Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), vol. 260 of Colloq. Internat. CNRS, CNRS, Paris, 1978, pp. 399–401.

[31] T. Tao, *A symmetric formulation of the croot-lev-pach-ellenberg-gijswijt capset bound*, blog post, (2016). `https://terrytao.wordpress.com/2016/05/18/a-symmetric-formulation-of-the-croot-lev-pach-ellenberg-gijswijt-capset-bound/`.

doi:10.6342/NTU202200962