

# Exact Bounds for Some Hypergraph Saturation Problems

Guy Moshkovitz\*

Asaf Shapira<sup>†</sup>

## Abstract

Let  $W_n(p, q)$  denote the minimum number of edges in an  $n \times n$  bipartite graph  $G$  on vertex sets  $X, Y$  that satisfies the following condition; one can add the edges between  $X$  and  $Y$  that do not belong to  $G$  one after the other so that whenever a new edge is added, a new copy of  $K_{p, q}$  is created. The problem of bounding  $W_n(p, q)$ , and its natural hypergraph generalization, was introduced by Balogh, Bollobás, Morris and Riordan. Their main result, specialized to graphs, used algebraic methods to determine  $W_n(1, q)$ .

Our main results in this paper give exact bounds for  $W_n(p, q)$ , its hypergraph analogue, as well as for a new variant of Bollobás's Two Families Theorem. In particular, we completely determine  $W_n(p, q)$ , showing that if  $1 \leq p \leq q \leq n$  then

$$W_n(p, q) = n^2 - (n - p + 1)^2 + (q - p)^2 .$$

Our proof applies a reduction to a multi-partite version of the Two Families Theorem obtained by Alon. While the reduction is combinatorial, the main idea behind it is algebraic.

## 1 Introduction

One of the most well-known results in Extremal Combinatorics is Bollobás's Two Families Theorem [3], which states<sup>1</sup> that if  $A_1, \dots, A_h$  and  $B_1, \dots, B_h$  are two families of sets satisfying  $A_i \cap B_j = \emptyset$  if and only if  $i = j$ , and if  $|A_i| \leq a$  and  $|B_i| \leq b$ , then  $h \leq \binom{a+b}{b}$ . While this theorem has many applications, Bollobás's motivation for proving it was an extremal graph/hypergraph *saturation* problem, which we further discuss in Subsection 1.1.

Our most general result in this paper is a variant of the Two Families Theorem for multi-partite sets. Our main motivation for proving this result was a problem considered by Balogh,

---

2010 *Mathematics Subject Classification*: 05D10, 05A17, 05C65

\*School of Mathematics, Tel-Aviv University, Tel-Aviv, Israel 69978. Email: [guymosko@tau.ac.il](mailto:guymosko@tau.ac.il). Supported in part by ISF grant 224/11.

<sup>†</sup>School of Mathematics, Tel-Aviv University, Tel-Aviv, Israel 69978, and Schools of Mathematics and Computer Science, Georgia Institute of Technology, Atlanta, GA 30332. Email: [asafico@tau.ac.il](mailto:asafico@tau.ac.il). Supported in part by NSF Grant DMS-0901355, ISF Grant 224/11 and a Marie-Curie CIG Grant 303320.

<sup>1</sup>This formulation, which is equivalent to the result of Bollobás, was actually conjectured by Ehrenfeucht and Mycielski, and confirmed by Katona [12], Jaeger and Payan [10], and Tarján [14].

Bollobás, Morris and Riordan [2] which is a variant of the saturation problem originally studied by Bollobás [3]. We elaborate on this aspect of the paper in Subsection 1.2. As we show later in the paper, our new Two Families Theorem can be used to fully resolve the problem considered in [2].

To state our new version of the Two Families Theorem we need the following definition, where here and throughout the paper, we use  $[n]$  to denote the set  $\{1, \dots, n\}$ .

**Definition 1.1.** *Let  $a_1, \dots, a_d$  and  $b_1, \dots, b_d$  be nonnegative integers with  $a := \max_i a_i$ , and take  $U_1, \dots, U_d$  to be disjoint sets, where  $|U_i| = a + b_i$ , and (with a slight abuse of notation) we think of each set  $U_i$  as  $[a + b_i]$ . We define  $Q(a_1, \dots, a_d, b_1, \dots, b_d)$  as the number of sets  $S \subseteq U_1 \cup \dots \cup U_d$  for which there is a permutation  $\pi : [d] \rightarrow [d]$  so that  $S \cap U_i$  is a subset of  $[a_{\pi(i)} + b_i]$  of size  $b_i$ .*

Our new Two Families Theorem is as follows.

**Theorem 1.** *Let  $X_1, \dots, X_d$  be  $d$  disjoint sets and let  $a_1, \dots, a_d, b_1, \dots, b_d$  be nonnegative integers. Suppose  $A_1, \dots, A_h$  and  $B_1, \dots, B_h$  are two families of subsets of  $X_1 \cup \dots \cup X_d$  that satisfy:*

1.  $A_i \cap B_i = \emptyset$  for every  $1 \leq i \leq h$ .
2.  $A_i \cap B_j \neq \emptyset$  for every  $1 \leq i < j \leq h$ .
3. For every  $1 \leq i \leq h$  and  $1 \leq j \leq d$  we have  $|B_i \cap X_j| \leq b_j$ .
4. For every  $1 \leq i \leq h$  there is a permutation  $\pi : [d] \rightarrow [d]$  so that for every  $1 \leq j \leq d$  we have  $|A_i \cap X_j| \leq a_{\pi(j)}$ .

Then

$$h \leq Q(a_1, \dots, a_d, b_1, \dots, b_d).$$

Moreover, this bound is best possible for any choice of  $a_1, \dots, a_d$  and  $b_1, \dots, b_d$ .

We refer the reader to the end of Section 3 for an explicit formula for  $Q(a_1, \dots, a_d, b_1, \dots, b_d)$ . Variants and special cases of Theorem 1 were proved by several authors. Most notably, Alon [1], using techniques from exterior algebra, proved a Two Families Theorem that differs from Theorem 1 in that the permutation  $\pi$  is restricted to be the identity permutation (i.e., the fourth condition is replaced by the requirement that for every  $1 \leq i \leq h$ ,  $1 \leq j \leq d$  we have  $|A_i \cap X_j| \leq a_j$ ). Alon's theorem then states that  $h \leq \prod_{i=1}^d \binom{a_i + b_i}{b_i}$ . Alon's theorem was preceded by a proof of the special case  $d = 1$ , which is often called the *skew Two Families*<sup>2</sup> theorem. This theorem was proved by Lovász [13], Frankl [9] and Kalai [11] in some of the classical applications of the linear algebra method in Combinatorics. Interestingly, finding a combinatorial proof for it is still open.

---

<sup>2</sup>Strictly speaking, the conditions in the skew version assume an ordering of the sets, so the two families are really two sequences; however, we keep the term "Two Families" for the sake of consistency.

## 1.1 Background on saturation problems

While our main motivation in this paper is a certain saturation problem in bipartite graphs (and more generally,  $d$ -uniform  $d$ -partite hypergraphs), we begin by mentioning some classical results on saturation problems in non-bipartite graphs. A graph  $G$  is *strongly saturated* with respect to a graph  $H$  (or strongly  $H$ -saturated) if  $G$  does not contain a copy of  $H$ , yet adding any new edge to  $G$  creates a copy of  $H$ . The problem of strong saturation asks for the *minimum* number of edges in an  $n$ -vertex graph that is strongly  $H$ -saturated, for different graphs  $H$  of interest (notice that the “dual” problem, of finding the *maximum* number of edges in an  $n$ -vertex  $H$ -saturated graph, is of course the classical Turán problem). Let  $S_n(p)$  be the minimum number of edges in an  $n$ -vertex graph that is strongly  $K_p$ -saturated, where  $2 \leq p \leq n$ . The problem of determining  $S_n(p)$  was considered already in the 1940’s by Zykov [17], and later by Erdős, Hajnal and Moon [8] who showed that  $S_n(p) = \binom{n}{2} - \binom{n-p+2}{2}$ . The upper bound on  $S_n(p)$  is easy, as removing the edges of a  $K_{n-p+2}$  from  $K_n$  clearly gives a strongly  $K_p$ -saturated graph. Bollobás’s Two Families Theorem, mentioned at the beginning of this paper, gives a tight lower bound for  $S_n(p)$  and for its natural hypergraph generalization (see the end of Section 3 for a similar reduction).

The following notion of saturation was originally introduced by Bollobás [5]. A graph  $G$  is *weakly saturated* with respect to a graph  $H$  (or weakly  $H$ -saturated) if all the non-edges of  $G$  can be added one at a time, in some order, so that each new edge creates a new copy of  $H$ . We refer to the corresponding ordering of the non-edges of  $G$  as a *saturation process* of  $G$  with respect to  $H$ . For example, it is not hard to see that the weakly  $K_3$ -saturated graphs with the minimum number of edges are precisely the trees; notice that already for  $K_3$  the extremal examples are not unique, suggesting that the general problem might be quite challenging.

Let  $W_n(p)$  be the minimum number of edges in an  $n$ -vertex graph that is weakly  $K_p$ -saturated, where  $2 \leq p \leq n$ . Notice that for any  $H$ , a strongly  $H$ -saturated graph is in particular weakly  $H$ -saturated, so  $W_n(p) \leq S_n(p)$ . It follows from the skew version of the Two Families Theorem that in fact  $W_n(p) = S_n(p)$ . That being said, the extremal graphs are not the same; there are weakly  $K_p$ -saturated graphs with  $\binom{n}{2} - \binom{n-p+2}{2}$  edges which are not strongly  $K_p$ -saturated.

## 1.2 Weak saturation in multi-partite hypergraphs

In this paper we focus on saturation problems in the setting of bipartite graphs, and more generally,  $d$ -uniform  $d$ -partite hypergraphs. This variant of the problem was first introduced in 1964 by Erdős, Hajnal and Moon [8]. Unlike the definition of saturation in the previous subsection, here (and henceforth) the only edges that are considered are those containing one vertex from each vertex class. Let  $H$  be a  $d$ -uniform  $d$ -partite hypergraph with vertex classes  $V_1, \dots, V_d$ . We say that  $H$  is weakly  $K_{p_1, \dots, p_d}^d$ -saturated<sup>3</sup> if all edges—containing one vertex from each  $V_i$ —that do not belong to  $H$  can be added to  $H$  one after the other so that whenever a new edge is added, a new copy of

---

<sup>3</sup> $K_{p_1, \dots, p_d}^d$  denotes the complete  $d$ -uniform  $d$ -partite hypergraph with vertex classes of sizes  $p_1, \dots, p_d$ .

$K_{p_1, \dots, p_d}^d$  is created.<sup>4</sup>

Our main motivation in this paper is the question of determining the following function. For integers  $1 \leq p_1, \dots, p_d \leq n$ , let  $W_n(p_1, \dots, p_d)$  be the smallest number of edges in a  $d$ -uniform  $d$ -partite hypergraph, with  $n$  vertices in each vertex class, that is weakly  $K_{p_1, \dots, p_d}^d$ -saturated. While this notion of weak saturation in  $d$ -uniform  $d$ -partite hypergraphs was introduced only recently by Balogh et al. [2], a similar notion of weak saturation, in which the copies of  $K_{p_1, \dots, p_d}^d$  are required to have  $p_i$  vertices in the  $i^{\text{th}}$  vertex class, was considered long before; we refer to this notion as *directed* weak saturation. For integers  $1 \leq p_1, \dots, p_d \leq n$  denote  $\vec{W}_n(p_1, \dots, p_d)$  the directed analogue of  $W_n(p_1, \dots, p_d)$ . Alon [1] determined  $\vec{W}_n(p_1, \dots, p_d)$  exactly, showing that  $\vec{W}_n(p_1, \dots, p_d) = n^d - \prod_{i=1}^d (n - p_i + 1)$ . Note that, by definition,  $W_n(p, \dots, p) = \vec{W}_n(p, \dots, p)$ , and so we can deduce from Alon's result a partial answer to the question considered in this paper, namely,

$$W_n(p, \dots, p) = n^d - (n - p + 1)^d. \quad (1)$$

A partial answer for a different setting of parameters was given by Balogh et al. [2]. Their main result, proved using linear algebraic techniques, determined  $W_n(p_1, \dots, p_d)$  when the  $p_i$  only take the two values 1 and  $q$ , for some positive integer  $q$ .

In this paper we determine  $W_n(p_1, \dots, p_d)$  for all values of  $p_1, \dots, p_d$  and  $n$ . To state our result we need the following definition.

**Definition 1.2.** For integers  $1 \leq p_1 \leq \dots \leq p_d \leq n$ , let  $q_n(p_1, \dots, p_d)$  be the number of  $d$ -tuples  $x \in [n]^d$  such that  $x_{(i)} \geq p_i$  for every  $1 \leq i \leq d$ , where  $x_{(i)}$  is the  $i^{\text{th}}$  smallest element in the sorted  $d$ -tuple of  $x$  (i.e., which includes repetitions).<sup>5</sup>

Our main theorem is as follows.

**Theorem 2.** For all integers  $1 \leq p_1 \leq \dots \leq p_d \leq n$  we have

$$W_n(p_1, \dots, p_d) = n^d - q_n(p_1, \dots, p_d).$$

It is of course interesting to find explicit formulas for  $q_n(p_1, \dots, p_d)$ , and thus for  $W_n(p_1, \dots, p_d)$ ; we do so in Section 2. By combining Theorem 2 and the explicit formulas, we obtain the interesting corollary that if  $p_d = o(n)$  then  $W_n(p_1, \dots, p_d)$  is asymptotically determined only by  $p_1$ , namely,

$$W_n(p_1, \dots, p_d) = (d(p_1 - 1) + o(1))n^{d-1}. \quad (2)$$

This should be compared with the fact that the directed analogue  $\vec{W}_n(p_1, \dots, p_d)$  is asymptotically determined by all  $p_1, \dots, p_d$ ; specifically, if  $p_1, \dots, p_d$  are all of order  $o(n)$  then  $\vec{W}_n(p_1, \dots, p_d) = (p_1 + \dots + p_d - d + o(1))n^{d-1}$ .

---

<sup>4</sup>Saturation in the setting of  $d$ -partite hypergraphs is referred to in some papers as  $d$ -saturation, or bi-saturation if  $d = 2$ . Since we henceforth only consider saturation in this setting, we prefer to keep using the term "saturation".

<sup>5</sup>For example, if  $x = (5, 2, 5, 1)$  then the sorted 4-tuple of  $x$  is  $(1, 2, 5, 5)$ .

### 1.3 Proof overview

Let us finally remark on the proofs of the theorems stated above. Interestingly, our proof of Theorem 1 proceeds by an indirect argument that reduces Theorem 1 to Alon’s Two Families Theorem (see Theorem 3 for the exact statement). Actually, our proof proceeds by reducing an instance with  $a_1, \dots, a_d$  to one where all  $a_i$  are replaced by  $\max_i a_i$ . This might seem counter-intuitive since enlarging the  $a_i$  increases the upper bound in Theorem 3. The catch is that when we will come to apply the bound for Theorem 3, the fact that we have increased the  $a_i$  is going to allow us to add some “dummy” pairs of sets  $A'_i, B'_i$  so that the new instance will still satisfy the requirements of Theorem 3. Somehow the trade-off between increasing  $a_1, \dots, a_d$  and adding the dummy sets results in a tight bound. We prove Theorem 2 using a reduction along similar lines.

The alert reader has probably noticed that the “trick” we use here, namely, adding extra sets to the families before applying the upper bound, is somewhat reminiscent of the trick used by Blokhuis to improve the bound on the size of 2-distance sets in Euclidean space [6]. And indeed, our original proof was a direct one, applying the algebraic proof of Alon’s theorem by Blokhuis [7], via resultants of polynomial, together with the trick from [6] of adding extra polynomials in order to improve the upper bound. As it happens, we later realized that it is in fact possible to reduce the problem to Alon’s theorem—no algebraic machinery necessary! (at least not explicitly)

**Organization:** The rest of the paper is organized as follows. We prove Theorem 2 in Section 2. Section 2 also contains some explicit formulas for  $W_n(p_1, \dots, p_d)$ . In Section 3 we prove our new Two Families Theorem, Theorem 1. We also give an explicit formula for  $Q(a_1, \dots, a_d, b_1, \dots, b_d)$ , as well as briefly explain how can one obtain an alternative proof of Theorem 2 using Theorem 1. Section 4 contains some concluding remarks and open problems.

## 2 Undirected Weak Saturation of Hypergraphs

We begin by proving the upper bound in Theorem 2. As a warm-up, let us briefly describe the construction proving the upper bound for the special case of graphs. Specifically, we show that  $W_n(p, q) \leq n^2 - (n - p + 1)^2 + (q - p)^2$ . Consider the  $n \times n$  bipartite graph with vertex classes  $\{x_1, \dots, x_n\}$  and  $\{y_1, \dots, y_n\}$  which is the union of three complete graphs: two  $K_{p-1, n}$ ’s, one with edges  $\{x_i y_j : i < p\}$  and the other with edges  $\{x_i y_j : j < p\}$ , and a  $K_{q-p, q-p}$  with edges  $\{x_i y_j : p \leq i, j < q\}$ . To see that this graph is weakly  $K_{p, q}$ -saturated, simply add the edges  $x_i y_j$  with  $i, j \geq q$  only after the rest of the missing edges are added. The reader may easily verify that by adding the missing edges in this order, a new copy of  $K_{p, q}$  is indeed created upon each addition.

We now generalize the construction above to the case of hypergraphs. Let  $G_0 = G_0(p_1, \dots, p_d)$  be the  $d$ -uniform  $d$ -partite hypergraph whose non-edges are enumerated by  $q_n(p_1, \dots, p_d)$  (recall Definition 1.2). More formally, let the vertex classes  $V_1, \dots, V_d$  of  $G_0$  each contain  $n$  vertices, and let us label the vertices in each set by  $1, 2, \dots, n$  (abusing notation slightly). Let us henceforth

identify edges with  $d$ -tuples in  $[n]^d$ .<sup>6</sup> Then an edge  $(x_1, \dots, x_d) \in V_1 \times \dots \times V_d$  **does not** belong to  $G_0$  if and only if for every  $1 \leq i \leq d$ , the  $i^{\text{th}}$  smallest element (i.e., when  $x_1, \dots, x_d$  are sorted with repetitions) is at least  $p_i$ . We now show that  $G_0$  is weakly  $K_{p_1, \dots, p_d}^d$ -saturated, proving the upper bound in Theorem 2. Henceforth we use  $\|H\|$  for the number of edges in a hypergraph  $H$ .

**Lemma 2.1.**  $W_n(p_1, \dots, p_d) \leq \|G_0\|$ .

*Proof.* Call  $x_1 + \dots + x_d$  the *weight* of the edge  $e = (x_1, \dots, x_d)$ , and denote  $G_w$  the  $d$ -uniform  $d$ -partite hypergraph obtained from  $G_0$  by adding every edge of weight at most  $w$ . We next prove that adding any new edge of weight  $w$  to  $G_{w-1}$  creates a new copy of  $K_{p_1, \dots, p_d}^d$ . From this it clearly follows by induction on  $w$  that  $G_0$  is weakly  $K_{p_1, \dots, p_d}^d$ -saturated, as required.

Let  $e = (x_1, \dots, x_d)$  be an edge of weight  $w$  and suppose  $e$  is not in  $G_{w-1}$ . We next construct for each  $1 \leq i \leq d$  a set  $S_i$  of vertices from the  $i^{\text{th}}$  vertex class. Fix  $1 \leq i \leq d$ , and suppose that  $x_i$  is the  $j^{\text{th}}$  smallest among  $x_1, \dots, x_d$  (i.e., when ordered with repetitions). We let  $S_i$  be the set of vertices, from the  $i^{\text{th}}$  vertex class, labeled by  $1, 2, \dots, p_j - 1$  (recall  $p_1 \leq \dots \leq p_d$ ). Since  $e$  is not in  $G_w$  and hence not in  $G_0$ , it follows from the definition of  $G_0$  that  $x_i \geq p_j$ . Therefore,  $S_i \cup \{x_i\}$  has  $p_j$  (distinct) elements. Note that every edge spanned by  $\bigcup_{i=1}^d (S_i \cup \{x_i\})$ , except for  $e$ , is of weight smaller than that of  $e$ , and so is contained in  $G_{w-1}$ . This means that adding  $e$  to  $G_0$  creates a new copy of  $K_{p_1, \dots, p_d}^d$  spanned by the vertices  $\bigcup_{i=1}^d (S_i \cup \{x_i\})$ , thus completing the proof.  $\square$

We next turn to the proof of the lower bound in Theorem 2. Let us start with a quick argument showing that in the graph case we have  $W_n(p, q) \geq n^2 - (n - p + 1)^2 + (q - p)^2$ . Given a weakly  $K_{p, q}$ -saturated bipartite graph, add  $q - p$  new vertices to each vertex class, connecting each new vertex to all the original vertices in the other class. A moment's thought reveals that this new graph is weakly  $K_{q, q}$ -saturated. This means that the number of edges in the new graph is at least  $W_{n+q-p}(q, q)$ , and applying (1) we get the desired lower bound.

We will now show how one can use the hypergraph  $G_0$  we constructed earlier to prove that every weakly  $K_{p_1, \dots, p_d}^d$ -saturated hypergraph must have as many edges as  $G_0$ . First, we will need to use the property of  $G_0$  that its complement,<sup>7</sup> denoted  $\overline{G_0}$ , contains every possible ‘‘orientation’’ of  $K_{n-p_1+1, \dots, n-p_d+1}^d$ .

**Claim 2.2.** *For every permutation  $\pi : [d] \rightarrow [d]$ , the hypergraph  $\overline{G_0}$  contains a copy of the hypergraph  $K_{n-p_1+1, \dots, n-p_d+1}^d$  having  $n - p_{\pi(i)} + 1$  vertices in the  $i^{\text{th}}$  vertex class.*

*Proof.* We start with a simple observation, claiming that if two tuples of real numbers  $x = (x_1, \dots, x_d)$  and  $y = (y_1, \dots, y_d)$  satisfy  $x_i \geq y_i$  for every  $1 \leq i \leq d$ , then they satisfy  $x_{(i)} \geq y_{(i)}$  for every  $1 \leq i \leq d$  as well (where, as usual,  $x_{(i)}$  is the  $i^{\text{th}}$  smallest element in the sorted tuple of  $x$ , and similarly for  $y$ ). To see this, let  $\sigma : [d] \rightarrow [d]$  be a permutation sorting  $y$ , that is,  $y_{\sigma(1)} \leq \dots \leq y_{\sigma(d)}$ .

<sup>6</sup>I.e., the  $d$ -tuple  $(x_1, \dots, x_d) \in [n]^d$  is identified with the edge containing from each  $V_i$  the vertex labeled  $x_i$ .

<sup>7</sup>By complement we mean relative to  $K_{n, \dots, n}$ , that is, the hypergraph that contains an edge  $(x_1, \dots, x_d) \in V_1 \times \dots \times V_d$  if and only if  $G_0$  does not.

Now note that for every  $1 \leq i \leq d$  and  $i \leq j \leq d$  we have  $x_{\sigma(j)} \geq y_{\sigma(j)} \geq y_{\sigma(i)} = y_{(i)}$ . This means that  $x$  has at least  $d - i + 1$  elements that are at least as large as  $y_{(i)}$ , which means that we must have  $x_{(i)} \geq y_{(i)}$ .

Now, suppose without loss of generality that  $p_1 \leq \dots \leq p_d$ . Let  $\pi : [d] \rightarrow [d]$  be an arbitrary permutation and let  $S_i$  be the subset of vertices of the  $i^{\text{th}}$  vertex class containing those vertices labeled by  $p_{\pi(i)}, p_{\pi(i)} + 1, \dots, n$ . Then for every edge  $e = (x_1, \dots, x_d)$  spanned by the vertices in  $\bigcup_{i=1}^d S_i$  it holds that  $x_i \geq p_{\pi(i)}$ . It now follows from our observation above that  $x_{(i)} \geq p_i$ . By the definitions of  $q_n(p_1, \dots, p_d)$  and  $G_0$  we conclude that  $e \in \overline{G_0}$ . Hence,  $\bigcup_{i=1}^d S_i$  spans a copy of  $K_{n-p_1+1, \dots, n-p_d+1}^d$  in  $\overline{G_0}$  having  $n - p_{\pi(i)} + 1$  vertices in the  $i^{\text{th}}$  vertex class, as desired.  $\square$

**Lemma 2.3.**  $W_n(p_1, \dots, p_d) \geq \|G_0\|$ .

*Proof.* Let  $H$  be a  $d$ -uniform  $d$ -partite hypergraph that is weakly  $K_{p_1, \dots, p_d}^d$ -saturated, where its vertex classes  $(V_1, \dots, V_d)$  are each of cardinality  $n$ . We construct a hypergraph  $H'$  with  $2n$  vertices in each vertex class by combining it with  $\overline{G_0}$  as follows. Let  $U_1, \dots, U_d$  be  $d$  sets of new vertices (i.e., disjoint from  $\bigcup_{i=1}^d V_i$  and from each other) with  $|U_i| = n$ . We let  $H'$  be the  $d$ -uniform  $d$ -partite hypergraph with vertex classes  $(V_1 \cup U_1, \dots, V_d \cup U_d)$  whose edges are defined as follows. The edges of  $H'$  that are spanned by the vertices in  $V := \bigcup_{i=1}^d V_i$  are precisely those of  $H$ ; the edges of  $H'$  that are spanned by the vertices in  $U := \bigcup_{i=1}^d U_i$  are precisely those of  $\overline{G_0}$ ; finally, all other possible edges (i.e., those containing at least one vertex from  $V$  and at least one vertex from  $U$ ) appear in  $H'$  as well. Notice that by counting the non-edges of  $H'$  we get

$$(2n)^d - \|H'\| = (n^d - \|H\|) + \|G_0\| . \quad (3)$$

We claim that  $H'$  is weakly  $K_{n+1, \dots, n+1}^d$ -saturated. Observe that (1) and (3) would then give

$$(2n)^d - (n^d - \|H\|) - \|G_0\| = \|H'\| \geq W_{2n}(n+1, \dots, n+1) = (2n)^d - n^d ,$$

implying that  $\|H\| \geq \|G_0\|$ , thus completing the proof.

To show that  $H'$  is weakly  $K_{n+1, \dots, n+1}^d$ -saturated, we claim that one obtains a saturation process of  $H'$  with respect to  $K_{n+1, \dots, n+1}^d$  by first adding the non-edges of  $H$  in the same order they appear in some saturation process of  $H$  (with respect to  $K_{p_1, \dots, p_d}^d$ ), and then adding, in an arbitrary order, all edges of  $G_0$ . To see that this indeed defines a saturation process of  $H'$  with respect to  $K_{n+1, \dots, n+1}^d$ , let  $e$  be a non-edge of  $H$  added at some point. Then adding  $e$  to  $H'$  (after all the edges that precede  $e$  in the saturation process are added) creates a new copy of  $K_{p_1, \dots, p_d}^d$  in  $H'$ , which we denote  $C$ . Let  $\pi : [d] \rightarrow [d]$  be a permutation such that  $C$  contains  $p_{\pi(i)}$  vertices in the  $i^{\text{th}}$  vertex class for every  $1 \leq i \leq d$ . By Claim 2.2,  $\overline{G_0}$  contains a copy  $C'$  of  $K_{n-p_1+1, \dots, n-p_d+1}^d$  having  $n - p_{\pi(i)} + 1$  vertices in the  $i^{\text{th}}$  vertex class. It follows that when adding  $e$  we in fact create a new copy of  $K_{n+1, \dots, n+1}^d$  in  $H'$ , namely, the copy spanned by the union of the vertex sets of  $C$  and  $C'$ . To complete the proof of our claim we observe that, after all the edges over  $V$  are added to  $H'$ , each edge  $(x_1, \dots, x_d)$  of  $G_0$  is the only missing edge in the copy of  $K_{n+1, \dots, n+1}^d$  spanned by  $\bigcup_{i=1}^d (V_i \cup \{x_i\})$  (recall that  $|V_i| = n$  for every  $i$ ). This completes the proof of the statement.  $\square$



*Proof of Theorem 2.* Lemmas 2.1 and 2.3 give  $W_n(p_1, \dots, p_d) = \|G_0\| = n^d - q_n(p_1, \dots, p_d)$ .  $\square$

## 2.1 Explicit formulas for $W_n(p_1, \dots, p_d)$

We begin by computing  $q_n(p_1, \dots, p_d)$  in some easy special cases. When the  $p_i$  take only one value we clearly have  $q_n(v, \dots, v) = (n - v + 1)^d$ . When the  $p_i$  take two values  $v_1 \leq v_2$ , where  $v_1$  occurs  $r$  times, it is easy to see that  $x = (x_1, \dots, x_d) \in [n]^d$  is enumerated by  $q_n$  if and only if it holds that  $x_i \geq v_1$  for all  $i$  and the number of  $x_i$  smaller than  $v_2$  is at most  $r$ ; thus,  $q_n(v_1, \dots, v_1, v_2, \dots, v_2) = \sum_{i=0}^r \binom{d}{i} (v_2 - v_1)^i (n - v_2 + 1)^{d-i}$ .

Let us consider the general case. Suppose  $p_1, \dots, p_d$  take  $m + 1$  distinct values  $v_1 < \dots < v_{m+1}$ , where  $v_i$  occurs  $r_i$  times,  $1 \leq i \leq m + 1$  (so  $r_{m+1} = d - r_1 - \dots - r_m$ ). For  $x = (x_1, \dots, x_d) \in [n]^d$  set  $i_j = |\{i : v_j \leq x_i < v_{j+1}\}|$ . Then it is not hard to see that  $x$  is enumerated by  $q_n$  if and only if  $x_i \geq v_1$  for all  $i$  and moreover  $i_1 \leq r_1, i_1 + i_2 \leq r_1 + r_2, \dots, i_1 + \dots + i_m \leq r_1 + \dots + r_m$ . This gives the following explicit formula;

$$q_n(p_1, \dots, p_d) = \sum_{i_1, \dots, i_m} \binom{d}{i_1, \dots, i_m} \prod_{j=1}^m (v_{j+1} - v_j)^{i_j} \cdot (n - v_{m+1} + 1)^{d - \sum_{k=1}^m i_k}$$

where the sum is over all  $i_1, \dots, i_m$  satisfying, for every  $1 \leq j \leq m$ , the inequality  $i_1 + \dots + i_j \leq r_1 + \dots + r_j$ .<sup>8</sup> We note that when all the  $p_i$  are distinct, that is, when  $r_1 = \dots = r_d = 1$ , the number of summands in the above formula is the  $d^{\text{th}}$  Catalan number. So in a sense,  $q_n(p_1, \dots, p_d)$  may be thought of as a “weighted” Catalan number.

An alternative description for  $W_n(p_1, \dots, p_d)$  can be obtained as follows. By Definition 1.2 and Theorem 2 we have that  $W_n(p_1, \dots, p_d)$  equals the number of  $d$ -tuples  $x \in [n]^d$  such that  $x_{(i)} < p_i$  holds for at least one  $1 \leq i \leq d$ . Consider now the set of  $d$ -tuples

$$L_i(t) = \left\{ (x_1, \dots, x_d) \in [n]^d : |\{j : x_j < t\}| = i \right\}.$$

A moment's thought reveals that<sup>9</sup>

$$W_n(p_1, \dots, p_d) = \left| \bigcup_{i=1}^d L_i(p_i) \right|. \quad (4)$$

We therefore obtain the inclusion-exclusion formula

$$W_n(p_1, \dots, p_d) = \sum_{\emptyset \neq I \subseteq [d]} (-1)^{|I|+1} \left| \bigcap_{i \in I} L_i(p_i) \right|. \quad (5)$$

It is easy to see that

$$|L_i(p_i)| = \binom{d}{i} (p_i - 1)^i (n - p_i + 1)^{d-i}, \quad (6)$$

<sup>8</sup>The notation  $\binom{n}{k_1, \dots, k_m}$  stands for the multinomial coefficient, that is,  $n! / (k_1! \dots k_m! \ell!)$  where  $\ell = n - \sum_{j=1}^m k_j$ .

<sup>9</sup>Indeed, if  $i$  is the largest such that  $x_{(i)} < p_i$  then clearly  $x \in L_i(p_i)$ ; conversely, if  $x \in L_i(p_i)$  then  $x_{(i)} < p_i$ .



and that for arbitrary  $I = \{i_1 < i_2 < \dots < i_t\}$  we have

$$\left| \bigcap_{i \in I} L_i(p_i) \right| = \binom{d}{i_1, \dots, i_t} (p_{i_1} - 1)^{i_1} \cdot \prod_{j=2}^t (p_{i_j} - p_{i_{j-1}})^{i_j - i_{j-1}} \cdot (n - p_{i_t} + 1)^{d - \sum_{k=1}^t i_k}. \quad (7)$$

Plugging (7) into (5) we get another explicit formula for  $W_n(p_1, \dots, p_d)$ .

Note that (4) gives the crude bound  $|L_1(p_1)| \leq W_n(p_1, \dots, p_d) \leq \sum_{i=1}^d |L_i(p_i)|$ . Combining it with (6) we get that

$$d(p_1 - 1)(n - p_1 + 1)^{d-1} \leq W_n(p_1, \dots, p_d) \leq \sum_{i=1}^d \binom{d}{i} (p_i - 1)^i (n - p_i + 1)^{d-i},$$

which implies the asymptotic formula (2) mentioned in the Introduction.

### 3 Undirected Two Families Theorem

In this section we prove Theorem 1, give an explicit formula for  $Q(a_1, \dots, a_d, b_1, \dots, b_d)$ , as well as give an alternative proof of Theorem 2. Our proof of Theorem 1 will follow by a reduction to Alon's Two Families Theorem which we repeat here.

**Theorem 3 (Alon [1]).** *Let  $X_1, \dots, X_d$  be  $d$  disjoint sets and let  $a_1, \dots, a_d, b_1, \dots, b_d$  be nonnegative<sup>10</sup> integers. Suppose  $A_1, \dots, A_h$  and  $B_1, \dots, B_h$  are two families of subsets of  $X_1 \cup \dots \cup X_d$  satisfying:*

1.  $A_i \cap B_i = \emptyset$  for every  $1 \leq i \leq h$ .
2.  $A_i \cap B_j \neq \emptyset$  for every  $1 \leq i < j \leq h$ .
3. For every  $1 \leq i \leq h$  and  $1 \leq j \leq d$  we have  $|A_i \cap X_j| \leq a_j$  and  $|B_i \cap X_j| \leq b_j$ .

Then  $h \leq \prod_{j=1}^d \binom{a_j + b_j}{b_j}$ .

Recall that the conditions of Theorem 1 differ from those in Alon's theorem only in that  $|A_i \cap X_j| \leq a_j$  need not hold, and instead it is only required that for every  $1 \leq i \leq h$  there is a permutation  $\pi : [d] \rightarrow [d]$  so that  $|A_i \cap X_j| \leq a_{\pi(j)}$  for  $1 \leq j \leq d$ .

#### 3.1 Proof of Theorem 1

We start with the proof of the upper bound.

---

<sup>10</sup>The original statement in [1] considers positive integers, but it is easy to see that it implies the statement with nonnegative integers.

*Proof.* Put  $a = \max_i a_i$ . Let  $U_1, \dots, U_d$  be  $d$  mutually disjoint sets, where  $U := \bigcup_{j=1}^d U_j$  is also disjoint from  $X_1 \cup \dots \cup X_d$ , and where each  $U_j$  contains  $a + b_j$  elements labeled  $1, 2, \dots, a + b_j$ . Setting  $X'_i := X_i \cup U_i$ , the idea is to transform any two families over  $X_1 \cup \dots \cup X_d$  satisfying the conditions of Theorem 1 into two larger families over  $X'_1 \cup \dots \cup X'_d$  satisfying the conditions of Theorem 3.

For a permutation  $\pi : [d] \rightarrow [d]$  we denote  $C_\pi$  the subset of  $U$  containing from each  $U_j$  its last  $a - a_{\pi(j)}$  members, that is, those labeled by  $a_{\pi(j)} + b_j + 1, \dots, a + b_j$ . For each  $1 \leq i \leq h$  fix a permutation  $\pi = \pi_i : [d] \rightarrow [d]$  satisfying  $|A_i \cap X_j| \leq a_{\pi(j)}$  (one exists by the fourth condition in the statement in Theorem 1) and set  $A'_i = A_i \cup C_\pi$ . Note that  $|A'_i \cap X'_j| \leq a$  for every  $1 \leq j \leq d$ . It is thus clear that the two families  $A'_1, \dots, A'_h$  and  $B_1, \dots, B_h$  satisfy the first and second conditions in the statement of Theorem 3 with respect to  $X'_1, \dots, X'_d$ , simply because every  $C_\pi$  is disjoint from  $X_1 \cup \dots \cup X_d$ ; furthermore, they also satisfy the third condition with  $a_1 = \dots = a_d = a$  and the same  $b_1, \dots, b_d$ .

Now, suppose that we are able to add  $h'$  new sets to each family—with the sets in the first family (i.e., the  $A_i$ 's in the statement) containing at most  $a$  elements from each part  $X'_j$  and the sets in the second family (i.e., the  $B_i$ 's) containing at most  $b_j$  elements from  $X'_j$ —while still satisfying the first and second conditions of Theorem 3. Applying Theorem 3 would then yield the upper bound  $h + h' \leq \prod_{j=1}^d \binom{a+b_j}{b_j}$ . Therefore, to complete the proof it suffices to show that we may extend the two families by  $h'$  new sets where  $h' = \prod_{j=1}^d \binom{a+b_j}{b_j} - Q(a_1, \dots, a_d, b_1, \dots, b_d)$ .

Note that  $Q(a_1, \dots, a_d, b_1, \dots, b_d)$  equals the number of sets  $B \subseteq U$ , with  $|B \cap U_j| = b_j$ , for which there is a permutation  $\pi : [d] \rightarrow [d]$  so that  $B \cap C_\pi = \emptyset$ . Thus,  $h' = \prod_{j=1}^d \binom{a+b_j}{b_j} - Q(a_1, \dots, a_d, b_1, \dots, b_d)$  is the number of sets  $B \subseteq U$ , with  $|B \cap U_j| = b_j$ , such that for every permutation  $\pi$ ,  $B \cap C_\pi \neq \emptyset$ . Let  $B'_1, \dots, B'_{h'}$  denote the sets enumerated by  $h'$ , and consider the two families<sup>11</sup>

$$A'_1, \dots, A'_h, \overline{B'_1}, \dots, \overline{B'_{h'}}$$

and

$$B_1, \dots, B_h, B'_1, \dots, B'_{h'}.$$

Since any  $B'_i$  contains  $b_j$  elements from each  $U_j$ , we have that  $\overline{B'_i}$  contains  $a$  elements from each  $U_j$ . Therefore, every set in the first family contains at most  $a$  members from each part  $X'_j$ , and every set in the second family contains at most  $b_j$  members from each part  $X'_j$ , as desired. We claim that the above two families satisfy the conditions of Theorem 3, which would complete the proof by applying that theorem as discussed above. The first condition in the statement is clearly satisfied, as  $A'_i \cap B_i = \emptyset$  and  $\overline{B'_i} \cap B'_i = \emptyset$ . As for the second condition, recall that, as observed above,  $A'_i \cap B_j \neq \emptyset$  when  $i \neq j$ ; moreover, it is clear that for any  $i \neq j$  we have  $\overline{B'_i} \cap B'_j \neq \emptyset$ , as  $B'_j \not\subseteq B'_i$ . It remains to show that for every  $1 \leq i \leq h$  and every  $1 \leq j \leq h'$  we have  $A'_i \cap B'_j \neq \emptyset$ . Indeed, any  $A'_i$  contains some  $C_\pi$  and any  $B'_j$  intersects every  $C_\pi$ . This completes the proof.  $\square$

<sup>11</sup>We write  $\overline{S}$  for the complement of  $S$  in  $U$ , that is,  $U \setminus S$ .

We now show that the bound in Theorem 1 is best possible for any choice of  $a_1, \dots, a_d$  and  $b_1, \dots, b_d$ .

*Proof.* Given  $a_1, \dots, a_d$  and  $b_1, \dots, b_d$  let  $h = Q(a_1, \dots, a_d, b_1, \dots, b_d)$ . We need to construct two families of  $h$  sets satisfying the four conditions of Theorem 1. For a set  $B \subseteq U_1 \cup \dots \cup U_d$  (where  $U_1, \dots, U_d$  are as in the definition of  $Q$ ), let  $w(B)$  be the sum of the labels of its members, that is,  $w(B) = \sum_{j=1}^d \sum_{x \in B \cap U_j} x$ . Let  $B_1, \dots, B_h$  be the sets enumerated by  $Q(a_1, \dots, a_d, b_1, \dots, b_d)$ , ordered by decreasing weight (breaking ties arbitrarily). For each  $B_i$ , fix a permutation  $\pi : [d] \rightarrow [d]$  so that  $B_i \cap U_j$  is a subset of  $[a_{\pi(j)} + b_j]$  of size  $b_j$ , and let  $A_i = A_i(\pi)$  be the set satisfying for every  $1 \leq j \leq d$  that  $A_i \cap U_j = [a_{\pi(j)} + b_j] \setminus (B_i \cap U_j)$ . The proof would follow by showing that  $A_1, \dots, A_h$  and  $B_1, \dots, B_h$  satisfy the conditions of Theorem 1.

It is clear from the definition of  $A_i$  and  $B_i$  (viewed as subsets of  $U_1 \cup \dots \cup U_d$ ) that the third and fourth conditions in the statement are satisfied. As for the first condition, note that for every  $1 \leq i \leq h$  we have (by definition)  $A_i \cap B_{i'} = \emptyset$ . Hence, it remains to show that if  $A_i \cap B_{i'} = \emptyset$  and  $i \neq i'$  then  $i' < i$ . Fixing  $i \neq i'$  for which  $A_i \cap B_{i'} = \emptyset$ , we will show that  $w(B_{i'}) > w(B_i)$ , thus completing the proof. Observe that for every  $x \in B_{i'} \cap U_j$  either  $x \in B_i \cap U_j$  or else  $x > a_{\pi(j)} + b_j$  (i.e.,  $x \in U_j \setminus [a_{\pi(j)} + b_j]$ ), as otherwise  $x \in A_i$ . Since  $|B_{i'} \cap U_j| = |B_i \cap U_j| (= b_j)$ , it follows that  $\sum_{x \in B_{i'} \cap U_j} x \geq \sum_{x \in B_i \cap U_j} x$ , and moreover, this inequality is strict if  $B_{i'} \cap U_j \neq B_i \cap U_j$ . Since  $B_{i'} \neq B_i$  there must be at least one  $j$  for which  $B_{i'} \cap U_j \neq B_i \cap U_j$ , implying that  $w(B_{i'}) > w(B_i)$ , as desired.  $\square$

### 3.2 Explicit formula for $Q(a_1, \dots, a_d, b_1, \dots, b_d)$

For a permutation  $\pi : [d] \rightarrow [d]$ , let  $\mathcal{F}_\pi$  be the family of sets  $B \subseteq U_1 \cup \dots \cup U_d$  satisfying  $B \cap U_i \subseteq [a_{\pi(i)} + b_i]$  and  $|B \cap U_i| = b_i$ . Denoting  $S_d$  the set of permutations on  $[d]$ , we have the inclusion-exclusion formula

$$Q(a_1, \dots, a_d, b_1, \dots, b_d) = \left| \bigcup_{\pi \in S_d} \mathcal{F}_\pi \right| = \sum_{\emptyset \neq I \subseteq S_d} (-1)^{|I|+1} \left| \bigcap_{\pi \in I} \mathcal{F}_\pi \right|. \quad (8)$$

Notice that for every  $\pi \in S_d$  we have  $|\mathcal{F}_\pi| = \prod_{i=1}^d \binom{a_{\pi(i)} + b_i}{b_i}$ . More generally, for any  $\emptyset \neq I \subseteq S_d$ , putting  $a_i^I = \min_{\pi \in I} a_{\pi(i)}$  we clearly have

$$\left| \bigcap_{\pi \in I} \mathcal{F}_\pi \right| = \prod_{i=1}^d \binom{a_i^I + b_i}{b_i}. \quad (9)$$

Plugging (9) into (8) gives an explicit formula for  $Q(a_1, \dots, a_d, b_1, \dots, b_d)$ . As an example, we get for  $d = 2$  that if  $a_1 \leq a_2$  then

$$Q(a_1, a_2, b_1, b_2) = \binom{a_1 + b_1}{b_1} \binom{a_2 + b_2}{b_2} + \binom{a_2 + b_1}{b_1} \binom{a_1 + b_2}{b_2} - \binom{a_1 + b_1}{b_1} \binom{a_1 + b_2}{b_2}.$$

### 3.3 Alternative proof of Theorem 2

At the beginning of Section 1 we claimed that Theorem 1 is the most general result of this paper. Let us briefly explain how can one derive the lower bound part of Theorem 2 from Theorem 1. Let  $G$  be a  $d$ -uniform  $d$ -partite hypergraph that is  $K_{p_1, \dots, p_d}^d$ -saturated, and suppose  $e_1, \dots, e_h$  is a corresponding saturation process, that is, an ordering of the non-edges of  $G$  such that, after all edges  $e_{i'}$  with  $i' < i$  are added to  $G$ , adding  $e_i$  creates a new copy  $C_i$  of  $K_{p_1, \dots, p_d}^d$ . Note that for each edge  $e \in C_i$ , either  $e \in G$  or else  $e = e_{i'}$  for some  $i' \leq i$ . Put  $A_i = V(G) \setminus V(C_i)$  and observe that  $A_1, \dots, A_h$  and  $e_1, \dots, e_h$  satisfy the first and second conditions of Theorem 1; indeed, we have  $e_i \subseteq V(C_i)$  so  $A_i \cap e_i = \emptyset$ , while for  $i < j$  we have  $e_j \not\subseteq V(C_i)$  so  $A_i \cap e_j \neq \emptyset$ . Now, denote  $V_1, \dots, V_d$  the vertex classes of  $G$ , and suppose all are of size  $n$ . Since  $C_i$  is a copy of  $K_{p_1, \dots, p_d}^d$ , there is a permutation  $\pi : [d] \rightarrow [d]$  so that  $|C_i \cap V_j| = p_{\pi(j)}$ . It follows that for every  $1 \leq i \leq h$  there is a permutation  $\pi$  so that  $|A_i \cap V_j| = n - p_{\pi(j)}$ ; moreover, clearly  $|e_i \cap V_j| = 1$ . Therefore, by applying Theorem 1 we deduce that the number of edges of  $G$ , which is  $n^d - h$ , is at least  $n^d - Q(n - p_1, \dots, n - p_d, 1, \dots, 1)$ . This can be used to give an alternative proof of the lower bound part of Theorem 2 since it can be shown that  $Q(n - p_1, \dots, n - p_d, 1, \dots, 1) = q_n(p_1, \dots, p_d)$ .

## 4 Concluding Remarks and Open Problems

**Undirected strong saturation in bipartite graphs:** Recall that our main result (specialized to graphs) shows that in the setting of weak saturation, the undirected version  $W_n(p, q)$  requires much fewer edges than its directed analogue  $\vec{W}_n(p, q)$ . It is thus natural to ask what happens in the setting of strong saturation. As we now discuss, we conjecture that while the undirected version is easier than the directed one, it is only easier by an additive constant factor.

Let  $S_n(p, q)$  be the minimum number of edges in an  $n \times n$  bipartite graph such that any addition of a new edge between its two classes creates a copy of  $K_{p, q}$  (i.e., the graph is strongly  $K_{p, q}$ -saturated). Let  $\vec{S}_n(p, q)$  denote the *directed* analogue<sup>12</sup> of  $S_n(p, q)$ . Answering a conjecture of Erdős-Hajnal-Moon [8],  $\vec{S}_n(p, q)$  was completely determined by Wessel [16] and Bollobás [4] to be  $(p + q - 2)n - (p - 1)(q - 1)$ .<sup>13</sup> Perhaps surprisingly, there are constructions showing that  $S_n(p, q)$  is, in general, strictly smaller than this (of course, there is no distinction between the undirected and directed versions when  $p = q$ ). To see this, suppose  $p \leq q$  and let  $G_{p, q}^k$  be any  $n \times n$  bipartite graph having  $p - 1$  vertices in each class complete to the other class, some  $k$  additional vertices in each class spanning a  $K_{k, k}$ , and where the remaining vertices have degree  $q - 1$ . Note that  $G_{p, q}^k$  has the property that any new edge one adds to it has an endpoint of degree at least  $q$ . The  $q$  neighbors, together with the  $p$  complete vertices from the other class, then form a  $K_{p, q}$ , implying that  $G_{p, q}^k$  is strongly  $K_{p, q}$ -saturated. One can check that  $G_{p, q}^k$  has in fact  $\vec{S}_n(p, q) - k(q - p - k)$  edges. Optimizing using  $k = \lfloor (q - p)/2 \rfloor$  gives  $S_n(p, q) \leq \vec{S}_n(p, q) - \lfloor (q - p)^2/4 \rfloor$ , for every  $n$  large

<sup>12</sup>I.e., where the copies of  $K_{p, q}$  must have their  $p$  vertices in the first class and their  $q$  vertices in the second class.

<sup>13</sup>In fact, this is a special case of Alon's Theorem 3; see [1].

enough such that  $G_{p,q}^k$  is well defined. We conjecture that this upper bound is best possible for large enough  $n$ , that is, the “gain” over the directed version is an additive constant.

**Conjecture 1.** *For every  $p, q$  there is an integer  $n_0$  such that for every  $n \geq n_0$  we have*

$$S_n(p, q) = \vec{S}_n(p, q) - \left\lfloor \frac{(q-p)^2}{4} \right\rfloor. \quad (10)$$

The only case for which we can confirm the above conjecture is when  $p = 1$  and  $q$  is arbitrary. Consider any  $n \times n$  bipartite graph that is strongly  $K_{1,q}$ -saturated, and observe that the set of vertices of degree strictly smaller than  $q - 1$  must span a clique. Consider the class containing the fewest such low-degree vertices, and let  $k$  be their number there. Summing the degrees of all vertices in that class, we get that the number of edges in the graph is at least

$$k^2 + (n - k)(q - 1) = (q - 1)n - k(q - 1 - k) \geq (q - 1)n - \left\lfloor \frac{(q - 1)^2}{4} \right\rfloor,$$

which agrees with (10).

We note that for general  $2 \leq p < q$  we do not have any (non trivial) lower bound for  $S_n(p, q)$ . It will thus be interesting to prove even a weaker version of the above conjecture, by establishing that  $S_n(p, q) \geq \vec{S}_n(p, q) - C$  for some constant  $C = C(p, q)$  that depends on  $p$  and  $q$  (and is independent of  $n$ ).

**Variants of Theorem 1:** It would be interesting to know what is the best possible bound one gets in Theorem 1 if, for example, one replaces the first and second conditions with a non-skew one, that is,  $A_i \cap B_j = \emptyset$  if and only if  $i = j$ , as in Bollobás’s Two Families Theorem. Such a variant would have implications for strong saturation; indeed, a special case of this was described in the previous item. Note that the bound in Alon’s Two Families Theorem cannot be improved if one replaces the skew condition with a non-skew one (as the natural extremal construction satisfies the non-skew condition as well). Interestingly, it can be shown that the bound in Theorem 1 is generally not best possible if one requires the non-skew condition instead (see the construction in the previous item). Nevertheless, it seems reasonable to conjecture that the correct bound in the non-skew case with, e.g.,  $b_i = b$  constant should not in general be much larger than the bound  $\prod_{i=1}^d \binom{a_i+b}{b}$  implied by Alon’s theorem.

It would also be interesting to consider variants of the third and fourth conditions of Theorem 1 that are symmetric with respect to  $a_i$  and  $b_i$ . For example, one might instead require that there be a permutation  $\pi : [d] \rightarrow [d]$  so that both  $|A_i \cap X_j| \leq a_{\pi(j)}$  and  $|B_i \cap X_j| \leq b_{\pi(j)}$ . As another example, we may instead require the existence of two permutations  $\pi, \sigma : [d] \rightarrow [d]$  such that  $|A_i \cap X_j| \leq a_{\pi(j)}$  and  $|B_i \cap X_j| \leq b_{\sigma(j)}$ . Notice that these two variants would have implications for weak saturation of classes of hypergraphs more general than  $d$ -uniform  $d$ -partite hypergraphs (this can be easily seen by following the reduction at the end of Section 3).

**$H$ -free saturation:** Notice that when defining whether a hypergraph  $G$  is strongly  $H$ -saturated, one may or may not require that  $G$  be  $H$ -free. Indeed, some authors make this requirement (e.g., [15]) and some do not (e.g., [1]). It would be interesting to know in this regard whether there is some  $H$  for which requiring  $H$ -freeness changes the corresponding strong saturation number.

## References

- [1] N. Alon, An extremal problem for sets with applications to graph theory, *J. Combin. Theory, Ser. A*, **40** (1985), 82–89. [1](#), [1.2](#), [3](#), [10](#), [13](#), [4](#)
- [2] J. Balogh, B. Bollobás, R. Morris and O. Riordan, Linear algebra and bootstrap percolation, *J. Combin. Theory, Ser. A*, **119** (2012), 1328–1335. [1](#), [1.2](#), [1.2](#)
- [3] B. Bollobás, On generalized graphs, *Acta Math. Acad. Sci. Hungar.*, **16** (1965), 447–452. [1](#)
- [4] B. Bollobás, On a conjecture of Erdős, Hajnal, and Moon, *Amer. Math. Monthly*, **74** (1967), 178–179. [4](#)
- [5] B. Bollobás, Weakly  $k$ -saturated graphs, In *Beiträge zur Graphentheorie* (Kolloquium, Manebach, 1967), 25–31. Teubner, Leipzig, 1968. [1.1](#)
- [6] A. Blokhuis, A new upper bound for the cardinality of 2-distance sets in Euclidean space, in *Convexity and Graph Theory*, **87** (1984), 65–67. [1.3](#)
- [7] A. Blokhuis, Solution of an extremal problem for sets using resultants of polynomials, in *Combinatorica*, **10** (1990), 393–396. [1.3](#)
- [8] P. Erdős, A. Hajnal and J. W. Moon, A problem in graph theory, *Amer. Math. Monthly* (1964), 1107–1110. [1.1](#), [1.2](#), [4](#)
- [9] P. Frankl, An extremal problem for two families of sets, *Europ. J. Math.*, **3** (1982), 125–127. [1](#)
- [10] F. Jaeger and CH. Payan, Nombre maximal d’arêtes d’un hypergraphe  $t$ -critique de rang  $h$ , *C. R. Acad. Sci. Paris*, **273** (1971), 221–223. [1](#)
- [11] G. Kalai, Weakly saturated graphs are rigid, *Convexity and Graph Theory*, **87** (1984), 189–190. [1](#)
- [12] G. Katona, Solution of a problem of A. Ehrenfeucht and J. Mycielski, *J. Combin. Theory, Ser. A*, **17** (1974), 265–266. [1](#)
- [13] L. Lovász, Flats in matroids and geometric graphs, in “Proc. 6<sup>th</sup> British Combin. Conf.” (P. J. Cameron, Ed.), pp. 45–86, Academic Press, 1977. [1](#)

- [14] T.G. Tarján, Complexity of lattice-configurations, *Studia Sci. Math. Hungar.*, **10** (1975), 203–211. [1](#)
- [15] Z. Tuza, Helly-type hypergraphs and Sperner families, *European J. Combin.*, **5** (1984), 185–187. [4](#)
- [16] W. Wessel, Über eine klasse paarer graphen, I: Beweis einer Vermutung von Erdős, Hajnal, und Moon, *Wiss. Z. Tech. Hochsch. Ilmenau* 12 (1966), 253–256. [4](#)
- [17] A. Zykov, On some properties of linear complexes (in Russian), *Mat. Sbornik N.S.*, **24** (1949), 163–188. [1.1](#)