

ASYMPTOTIC ENUMERATION OF NON-UNIFORM LINEAR HYPERGRAPHS

MAHDIEH HASHEMINEZHAD

Department of Computer Science
Yazd University, Yazd, Iran
Combinatorial and Geometric Algorithms Lab
e-mail: hasheminezhadl@yazd.ac.ir

AND

BRENDAN D. MCKAY

Research School of Computer Science
Australian National University
Canberra, ACT 2601, Australia
e-mail: brendan.mckay@anu.edu.au

Abstract

A linear hypergraph, also known as a partial Steiner system, is a collection of subsets of a set such that no two of the subsets have more than one element in common. Most studies of linear hypergraphs consider only the uniform case, in which all the subsets have the same size. In this paper we provide, for the first time, asymptotically precise estimates of the number of linear hypergraphs in the non-uniform case, as a function of the number of subsets of each size.

Keywords: Steiner system, linear hypergraph, asymptotic enumeration, switching method.

2010 Mathematics Subject Classification: 05A16, 05C65.

1. INTRODUCTION

A (simple) *hypergraph* is a pair $H = (V, E)$, where V is a finite nonempty set of *vertices* and E is a set of non-empty subsets of V called *edges*. Hypergraphs, which are also known as *set systems* and *block designs*, are fundamental to the study of complex discrete systems. They arise naturally in many application areas such as ecology, chemistry, statistics and computer science.

If all the edges consist of two elements, a hypergraph is better known as a *graph*. In contrast to the vast literature on graphs, the literature on more general hypergraphs is poor in comparison, especially for enumerative problems.

A *Steiner system* (for $t = 2$) is a hypergraph with the property that every pair of vertices lies in exactly one edge. If also every edge has size r , it is a *Steiner r -tuple system*. Wilson proved in 1975 that such systems exist for fixed r and sufficiently large V provided the obvious divisibility condition is met [5].

If we replace “exactly one edge” by “at most one edge”, we have a *partial Steiner system*, which is also known as a *linear hypergraph*. In other words, a hypergraph is linear if no pair of edges have more than one vertex in common. Edges of size 1 can be added arbitrarily to a linear hypergraph without destroying the linearity, so we will disallow them without further comment.

Our aim is to study the following problem.

Problem 1. Let $n \geq 1$ and $K \geq 2$ be integers. For $2 \leq i \leq K$, let $m_i = m_i(n)$ be a nonnegative integer. How many linear hypergraphs have n vertices, exactly m_i edges of size i for $2 \leq i \leq K$, and no edges of size greater than K ?

To the best of our knowledge, there is no previous work on this problem except in the r -uniform case (i.e., when all the edges have the same size r). The logarithm of the total number of r -uniform linear hypergraphs with n vertices was determined asymptotically by Grable and Phelps [2]. A recent preprint of McKay and Tian, completed while this paper was in process, asymptotically counts r -uniform hypergraphs with $o(n^{3/2})$ edges [4]. Blinovskiy and Greenhill determined the asymptotic number of r -uniform linear hypergraphs with m edges and maximum degree k as a function of the vertex degrees, provided $r^3 k^4 (k+r) = o(m)$ [1]. In this article we are not concerned with the vertex degrees but only with the number of edges of each size.

Since there are only $\binom{n}{2}$ pairs of vertices, a necessary condition for the existence of a linear hypergraph satisfying our description is that $\sum_{i=2}^K m_i \binom{i}{2} \leq \binom{n}{2}$, which implies that $m_i = O(n^2)$ for each i . The total number of hypergraphs, not required to be linear, is

$$N = N(n, m_2, \dots, m_K) = \prod_{i=2}^K \binom{\binom{n}{i}}{m_i}.$$

We will prove the following result.

Theorem 2. *If $K \geq 2$ is fixed, and $m_i = o(n^{4/3})$ as $n \rightarrow \infty$ for $2 \leq i \leq K$, then the number of linear hypergraphs with n vertices, exactly m_i edges of size i for $2 \leq i \leq K$, and no edges of size greater than K is*

$$N \exp \left(- \sum_{3 \leq a \leq K} \frac{a^2(a-1)^2}{2n^2} \binom{m_a}{2} - \sum_{2 \leq a < b \leq K} \frac{a(a-1)b(b-1)}{2n^2} m_a m_b + O(m^2/n^3 + m^3/n^4) \right),$$

where $m = \sum_{i=2}^K m_i$.

2. OUTLINE AND DEFINITIONS

Considering a hypergraph $H = (V, E)$, let G be the graph whose vertex set is E and two vertices of G are adjacent if they intersect in at least two vertices of H . Let $t \geq 2$ be an integer. A set S of t edges of H is called a t -cluster of H if S is the vertex set of a component of G . A 2-cluster is *strong* if its two edges have more than two vertices in common; otherwise it is *weak*. For integers a and b with $2 \leq a \leq b \leq K$, a weak 2-cluster with edges of sizes a and b is called an *ab-pair*. (Note that 22-pairs cannot happen for simple hypergraphs.) A set B of three edges of H is called a *3-bunch* if the induced subgraph of G with vertex set B is connected.

Define $J = \{\{a, b\} \mid 2 \leq a \leq b \leq K, b \geq 3\}$, which is a set of $\frac{1}{2}(K-2)(K+1)$ pairs that we always use in lexicographic order denoted by \preceq . For notational convenience, we will often write $\{a, b\}$ as ab or ba and $\{a, a\}$ as aa . If they exist, the element preceding ab in lexicographic order is denoted by $p(ab)$ and the element following ab in lexicographic order is denoted by $n(ab)$. The notations $p(23)$ and $n(KK)$ are left undefined, but a list like $(x_{23}, \dots, x_{p(23)})$ can appear as a special case and represents a list $()$ with no entries.

Define

$$M = \lceil 55K^4 m^2 / n^2 + \log n \rceil$$

and denote by X the product space $\{0, 1, \dots, M\}^{\frac{1}{2}(K-2)(K+1)}$ corresponding to vectors $\mathbf{x} = (x_{23}, \dots, x_{KK})$, where the subscripts are all the elements of J in lexicographic order.

We approach Problem 1 by generating a random hypergraph H with the required number of edges of each size and seeking the probability that it is linear. We achieve this goal in two steps.

1. With probability $1 - O(m^2/n^3 + m^3/n^4)$, H has no t -clusters for $t > 2$, no strong 2-clusters, and at most M ab -pairs for each $ab \in J$.
2. For $\mathbf{x} = (x_{23}, \dots, x_{KK}) \in X$, let $\mathcal{L}(\mathbf{x})$ be the set of hypergraphs on n vertices with no t -clusters for $t > 2$, no strong 2-clusters, m_i edges of size i for $2 \leq i \leq K$,

and x_{ab} ab -pairs for $ab \in J$. Define $L(\mathbf{x}) = |\mathcal{L}(\mathbf{x})|$, so in particular the number of linear hypergraphs with the specified edge sizes is $L(0, \dots, 0)$. Let

$$T = T(n, m_2, \dots, m_k) = \sum_{\mathbf{x} \in X} L(\mathbf{x})$$

be the quantity we estimated in Step 1. In this step we estimate the value of $T/L(0, \dots, 0)$ by the switching method. Figure 1 shows an example of an operation that removes one 35-pair.

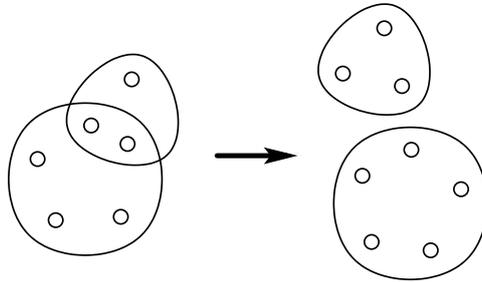


Figure 1. Switching operation to remove a 35-pair.

We will carry out Step 1 in Section 3. In Section 4, we will analyse the switchings that are key to our result. Then in Section 5 we will carry out Step 2.

3. STEP ONE

Lemma 3. *Let H be a random hypergraph H with exactly m_i edges of size i for $2 \leq i \leq K$ and no larger edges. Then, with probability $1 - O(m^2/n^3 + m^3/n^4)$, H has no t -clusters for $t \geq 3$, no strong 2-clusters, and at most M ab -pairs for each $ab \in J$. That is,*

$$T = (1 - O(m^2/n^3 + m^3/n^4)) N.$$

Proof. For $2 \leq i \leq K$, let $0 \leq b_i \leq m_i$. Consider a set S of edges consisting of b_i edges of size i for each i . Then the probability that H contains S is

$$(1) \quad \prod_{i=2}^K \frac{\binom{\binom{n}{i} - b_i}{m_i - b_i}}{\binom{\binom{n}{i}}{m_i}} \leq \prod_{i=2}^K \left(\frac{m_i}{\binom{n}{i}} \right)^{b_i} \leq \prod_{i=2}^K \left(\frac{i! m_i}{(n-i)^i} \right)^{b_i}.$$

Since every t -cluster for $t \geq 3$ contains a 3-bunch, we can prove the first part of the lemma by showing that there are no 3-bunches with high probability. Consider a 3-bunch B with edges of size a, b, c , where $2 \leq a, b, c \leq K$. Since it

occupies at most $s = a + b + c - 4$ vertices and K is bounded, the number of possible locations for such 3-bunches is $O(n^s)$. Moreover, the probability that H contains B is $O(m^3/n^{s+4})$, by (1). Therefore, the expected number of 3-bunches with edge sizes a, b, c is $O(m^3/n^4)$. Since there are only boundedly many choices of a, b, c , and $m^3/n^4 = o(1)$ by assumption, the probability that there are no 3-bunches (and therefore no t -clusters for $t \geq 3$) is $1 - O(m^3/n^4)$.

Similarly, a strong 2-cluster with edges of size a and b occupies at most $a + b - 3$ vertices. So the number of all possible locations of a strong 2-cluster with edges of sizes a and b is $O(n^{a+b-3})$ and the probability of each being in H is $O(m^2/n^{a+b})$. Since there is a constant number of choices for the values of a and b , the expected number of strong 2-clusters is $O(m^2/n^3) = o(1)$. Therefore, with probability $1 - O(m^2/n^3)$, H has no strong 2-clusters.

Let a and b be two integers such that $2 \leq a < b \leq K$, and let S be a set of M ab -pairs. Recall that by definition ab -pairs cannot have edges in common. If $\min\{m_a, m_b\} < M$, then S cannot exist; so assume this is not the case. The number of possible locations of S is at most $\binom{N_{a,b}}{M} \leq N_{a,b}^M/M! \leq (eN_{a,b}/M)^M$, where

$$N_{a,b} = \binom{n}{n-a-b+2, a-2, b-2, 2} \leq \frac{n^{a+b-2}}{2(a-2)!(b-2)!}.$$

Multiplying by the probability that S is in H , as given by (1), we find that the expected number of sets of M ab -pairs in H is at most

$$\left(\frac{2a^2b^2m_a m_b}{n^2M}\right)^M \leq \left(\frac{2}{55}\right)^{\log n} \leq n^{-3}.$$

Therefore, the probability that there are more than M ab -pairs is $O(n^{-3})$. The same argument holds for $a = b > 2$, and so simultaneously for $ab \in J$. Applying the union bound completes the proof. ■

4. ANALYSIS OF SWITCHINGS

For $\mathbf{x} \in X$ and $2 \leq a \leq K$, define

$$x_a(\mathbf{x}) = x_{aa} + \sum_{b:ab \in J} x_{ab},$$

using $x_{22} = 0$. Note that $x_a(\mathbf{x})$ is the total number of edges of size a that appear in the 2-clusters of $H \in \mathcal{L}(\mathbf{x})$, if such a hypergraph exists. Indeed, an obvious necessary condition for $L(\mathbf{x}) > 0$ is that

$$(2) \quad x_a(\mathbf{x}) \leq m_a \quad (2 \leq a \leq K).$$

In Lemma 4 below we will show that (2) is also sufficient.

Now let $ab \in J$ be such that $x_{ab} > 0$ and let $\mathbf{x}' = (x_{23}, \dots, x_{ab} - 1, \dots, x_{KK})$. We define a “forward” switching that converts a hypergraph $H \in \mathcal{L}(\mathbf{x})$ into a hypergraph in $\mathcal{L}(\mathbf{x}')$. Choose an ab -pair $\{e, e'\}$. Remove e and e' from H , then insert two edges of the same sizes (a and b) into the new hypergraph in such a way that no t -clusters for any $t \geq 2$ are created. The case of $a = 3, b = 5$ is shown in Figure 1, where it should be noted that the replacement edges might have one vertex in common but not more.

Consider the case $a \neq b$ first. We can choose $\{e, e'\}$ in x_{ab} ways. The number of places to insert the new edge of size a is at most $\binom{n}{a}$. From this we must subtract the number of positions that overlap some other edge in two or more vertices, which is at most $m \binom{K}{2} \binom{n}{a-2} = O(m/n^2) \binom{n}{a}$. Similarly for the edge of size b .

In the case $a = b$, the two new edges can be inserted in either order to achieve the same hypergraph. Otherwise, the estimates are the same. Consequently, the number of forward switchings is

$$(3) \quad W_F = \begin{cases} \frac{1}{2} x_{ab} \binom{n}{a}^2 (1 + O(m/n^2)), & \text{if } a = b, \\ x_{ab} \binom{n}{a} \binom{n}{b} (1 + O(m/n^2)), & \text{if } a \neq b. \end{cases}$$

Now consider a hypergraph $H' \in \mathcal{L}(\mathbf{x}')$. We will count the number of reverse switchings that convert H' into a member of $\mathcal{L}(\mathbf{x})$. Let $x_a = x_a(\mathbf{x}')$ and $x_b = x_b(\mathbf{x}')$. A reverse switching consists of removing two edges, one of size a and one of size b , and then inserting a new ab -pair.

Consider $a \neq b$ first. We can choose edges of size a and b , neither part of a 2-cluster, in $(m_a - x_a)(m_b - x_b)$ ways. Then we need to choose a location for the new ab -pair. Two sets of size a and b that overlap in 2 vertices can be chosen in $\binom{n}{a} \binom{n-a}{b-2} \binom{a}{2}$ ways. This pair of sets might be unsuitable due to one of them overlapping an edge in 2 or more vertices. We can bound the number of such bad choices by noting that the number of choices of $a + b - 2$ vertices which overlap an edge by 2 or more vertices is at most $m \binom{K}{2} \binom{n}{a+b-4} = O(m/n^2) \binom{n}{a} \binom{n-a}{b-2}$.

In the case $a = b$, we can choose two edges of size a , not part of a 2-cluster, in $\binom{m_a - x_a}{2}$ ways. When we form an aa -pair within a set of $2a - 2$ vertices, the number of ways is $\frac{(2a-2)!}{4(a-2)!^2}$. Otherwise, the argument is the same.

In summary, the number of reverse switchings is

$$(4) \quad W_R = \begin{cases} \binom{m_a - x_a}{2} \binom{n}{2a-2} \frac{(2a-2)!}{4(a-2)!^2} (1 + O(m/n^2)), & \text{if } a = b, \\ (m_a - x_a)(m_b - x_b) \binom{n}{a} \binom{n-a}{b-2} \binom{a}{2} (1 + O(m/n^2)), & \text{if } a \neq b. \end{cases}$$

Lemma 4. *Assume the conditions of Theorem 2. For $\mathbf{x} \in X$, $L(\mathbf{x}) > 0$ if and only if condition (2) is satisfied. In particular, $L(0, \dots, 0) > 0$.*

Proof. By Lemma 3, there is some $\hat{\mathbf{x}} \in X$ such that $L(\hat{\mathbf{x}}) > 0$. We can move from $\hat{\mathbf{x}}$ to \mathbf{x} by a sequence of forward and reverse switchings while staying within X . Since the values given in (3) or (4) at each step of this path are positive, we must have $L(\mathbf{x}) > 0$. ■

5. STEP TWO

In the second step, we calculate $T/L(0, 0, \dots, 0)$. Recall that the denominator is nonzero by Lemma 4. By the definition of T , we have

$$\frac{T}{L(0, \dots, 0)} = \sum_{\mathbf{x} \in X} \frac{L(\mathbf{x})}{L(0, \dots, 0)}.$$

Define

$$R_{ab}(x_{23}, \dots, x_{ab}) = \begin{cases} 0, & \text{if } L(x_{23}, \dots, x_{ab}, 0, \dots, 0) = 0, \\ \frac{L(x_{23}, \dots, x_{ab}, 0, \dots, 0)}{L(x_{23}, \dots, x_{ab}-1, 0, \dots, 0)}, & \text{otherwise.} \end{cases}$$

Note that Lemma 4 ensures that we cannot have a nonzero numerator and a zero denominator in the second case. Then

$$(5) \quad \frac{T}{L(0, \dots, 0)} = \sum_{\mathbf{x} \in X} \prod_{ab \in J} \prod_{j=1}^{x_{ab}} R_{ab}(x_{23}, \dots, x_{p(ab)}, j).$$

We will evaluate (5) using a recursion. For $ab \in J$, define

$$(6a) \quad S_{ab}(x_{23}, \dots, x_{p(ab)}) = \begin{cases} \sum_{x_{KK}=0}^M \prod_{j=1}^{x_{KK}} R_{KK}(x_{23}, \dots, x_{p(KK)}, j), & \text{if } a = b = K, \\ \sum_{x_{ab}=0}^M S_{n(ab)}(x_{23}, \dots, x_{ab}) \prod_{j=1}^{x_{ab}} R_{ab}(x_{23}, \dots, x_{p(ab)}, j), & \text{otherwise.} \end{cases}$$

Then $T/L(0, \dots, 0) = S_{23}()$.

Lemma 5. *Assume the conditions of Theorem 2 and define*

$$x_a = x_a(x_{23}, \dots, x_{p(ab)}, 0, \dots, 0) \quad \text{and} \quad x_b = x_b(x_{23}, \dots, x_{p(ab)}, 0, \dots, 0).$$

If $j \geq 1$ and $L(x_{23}, \dots, x_{p(ab)}, j - 1, 0, \dots, 0) > 0$, then

$$R_{ab}(x_{23}, \dots, x_{p(ab)}, j) = \begin{cases} \binom{m_a - x_a - 2j + 2}{2} \frac{(a - 1)^2 a^2}{2jn^2} (1 + O(1/n + m/n^2)), & \text{if } a = b, \\ (m_a - x_a - j + 1)(m_b - x_b - j + 1) \cdot \frac{(a - 1)a(b - 1)b}{2jn^2} (1 + O(1/n + m/n^2)), & \text{if } a \neq b. \end{cases}$$

Proof. Note that $x_a(x_{23}, \dots, x_{p(ab)}, j - 1, 0, \dots, 0)$ is equal to $x_a + 2j - 2$ if $a = b$ and $x_a + j - 1$ if $a \neq b$, and similarly for x_b . The lemma now follows on dividing (4) by (3). ■

We will need the following summation lemma from [3, Corollary 4.5].

Lemma 6. Let $Z \geq 2$ be an integer and, for $1 \leq i \leq Z$, let real numbers $A(i), B(i)$ be given such that $A(i) \geq 0$ and $1 - (i - 1)B(i) \geq 0$. Define $A_1 = \min_{i=1}^Z A(i)$, $A_2 = \max_{i=1}^Z A(i)$, $C_1 = \min_{i=1}^Z A(i)B(i)$ and $C_2 = \max_{i=1}^Z A(i)B(i)$. Suppose that there exists \hat{c} with $0 < \hat{c} < \frac{1}{3}$ such that $\max\{A/Z, |C|\} \leq \hat{c}$ for all $A \in [A_1, A_2]$, $C \in [C_1, C_2]$. Define n_0, \dots, n_Z by $n_0 = 1$ and

$$n_i/n_{i-1} = \frac{1}{i}A(i)(1 - (i - 1)B(i))$$

for $1 \leq i \leq Z$, with the following interpretation: if $A(i) = 0$ or $1 - (i - 1)B(i) = 0$, then $n_j = 0$ for $i \leq j \leq Z$. Then

$$\Sigma_1 \leq \sum_{i=0}^Z n_i \leq \Sigma_2,$$

where

$$\Sigma_1 = \exp(A_1 - \frac{1}{2}A_1C_2) - (2e\hat{c})^Z$$

and

$$\Sigma_2 = \exp(A_2 - \frac{1}{2}A_2C_1 + \frac{1}{2}A_2C_1^2) + (2e\hat{c})^Z.$$

For $ab, rs \in J$, define

$$g_{rs}(x_{23}, \dots, x_{p(ab)}) = \begin{cases} 0, & \text{if } L(x_{23}, \dots, x_{p(ab)}, 0, \dots, 0) = 0, \\ \text{otherwise,} \\ \binom{m_r - x_r}{2} \frac{r^2(r - 1)^2}{2n^2}, & \text{if } r = s, \\ (m_r - x_r)(m_s - x_s) \frac{r(r - 1)s(s - 1)}{2n^2}, & \text{if } r \neq s, \end{cases}$$

where $x_r = x_r(x_{23}, \dots, x_{p(ab)}, 0, \dots, 0)$ and $x_s = x_s(x_{23}, \dots, x_{p(ab)}, 0, \dots, 0)$.

Lemma 7. *Assume the conditions of Theorem 2. Then, for $ab \in J$,*

$$(7) \quad S_{ab}(x_{23}, \dots, x_{p(ab)}) = (1 + O(m^2/n^3 + m^3/n^4)) S'_{ab}(x_{23}, \dots, x_{p(ab)}),$$

where

$$S'_{ab}(x_{23}, \dots, x_{p(ab)}) = \exp \left(\sum_{rs \in J: rs \succeq ab} g_{rs}(x_{23}, \dots, x_{p(ab)}) \right).$$

In particular,

$$\frac{T}{L(0, \dots, 0)} = (1 + O(m^2/n^3 + m^3/n^4)) \exp \left(\sum_{rs \in J} g_{rs}(\cdot) \right).$$

Proof. For convenience we will also define an artificial value

$$S'_{n(KK)}(x_{23}, \dots, x_{KK}) = 1.$$

We will prove the lemma by reverse induction on ab .

Suppose (7) holds for $S_{n(ab)}$. Applying recurrence (6b), and noting that we are dealing with summations of nonnegative terms, we have

$$(8) \quad \begin{aligned} S_{ab}(x_{23}, \dots, x_{p(ab)}) &= \sum_{x_{ab}=0}^M S_{n(ab)}(x_{23}, \dots, x_{ab}) \prod_{j=1}^{x_{ab}} R_{ab}(x_{23}, \dots, x_{p(ab)}, j) \\ &= (1 + O(m^2/n^3 + m^3/n^4)) \\ &\quad \cdot \sum_{x_{ab}=0}^M S'_{n(ab)}(x_{23}, \dots, x_{ab}) \prod_{j=1}^{x_{ab}} R_{ab}(x_{23}, \dots, x_{p(ab)}, j) \\ &= (1 + O(m^2/n^3 + m^3/n^4)) \\ &\quad \cdot S'_{n(ab)}(x_{23}, \dots, x_{p(ab)}, 0) \sum_{i=0}^M \prod_{j=1}^i Q_{ab}(x_{23}, \dots, x_{p(ab)}, j), \end{aligned}$$

where

$$Q_{ab}(x_{23}, \dots, x_{p(ab)}, j) = \frac{S'_{n(ab)}(x_{23}, \dots, x_{p(ab)}, j)}{S'_{n(ab)}(x_{23}, \dots, x_{p(ab)}, j-1)} R_{ab}(x_{23}, \dots, x_{p(ab)}, j).$$

Note that (8) holds for $ab = KK$, by (6a) and our adopted meaning of $S'_{n(KK)}$, which gets the induction started.

We can also calculate that, if $R_{ab}(x_{23}, \dots, x_{p(ab)}, j) > 0$,

$$\frac{S'_{n(ab)}(x_{23}, \dots, x_{p(ab)}, j)}{S'_{n(ab)}(x_{23}, \dots, x_{p(ab)}, j-1)} = 1 + O(m/n^2).$$

Therefore, $Q_{ab}(x_{23}, \dots, x_{p(ab)}, j) = (1 + O(m/n^2))R_{ab}(x_{23}, \dots, x_{p(ab)}, j)$.

For simplicity, define $Q(j) = Q_{ab}(x_{23}, \dots, x_{p(ab)}, j)$, $x_a = x_a(x_{23}, \dots, x_{p(ab)}, 0, \dots, 0)$ and $x_b = x_b(x_{23}, \dots, x_{p(ab)}, 0, \dots, 0)$. By Lemma 5, there are functions $\delta_{ab}(j) = \delta_{ab}(x_{23}, \dots, x_{p(ab)}, j)$ which are $O(1/n + m/n^2)$ uniformly over all \mathbf{x}, ab, j , such that, if $Q(j) > 0$, then

$$Q(j) = \begin{cases} \binom{m_a - x_a - 2j + 2}{2} \frac{(a-1)^2 a^2}{2jn^2} (1 + \delta_{aa}(j)), & \text{if } a = b, \\ (m_a - x_a - j + 1)(m_b - x_b - j + 1) \frac{(a-1)a(b-1)b}{2jn^2} (1 + \delta_{ab}(j)), & \text{if } a \neq b. \end{cases}$$

Note that $m_a - x_a - 2j + 2 \geq 2$ in the first case, or $m_a - x_a - j + 1 \geq 1$ and $m_b - x_b - j + 1 \geq 1$ in the second case, or else $Q(j) = 0$ by Lemma 4.

If $Q(1) = 0$, the sum in (8) equals 1. If not, we will use Lemma 6 to complete the induction proof by evaluating the sum. The value n_i in the lemma corresponds to $\prod_{j=1}^i Q(j)$. Then, for $1 \leq i \leq M$ define

$$A(i) = Q(1) (1 + \delta_{ab}(i)) (1 + \delta_{ab}(1))^{-1}$$

$$B(i) = \begin{cases} \frac{2(2m_a - 2x_a - 2i + 1)}{(m_a - x_a)(m_a - x_a - 1)}, & \text{if } a = b \text{ and } Q(i) > 0, \\ \frac{m_a + m_b - x_a - x_b - i + 1}{(m_a - x_a)(m_b - x_b)}, & \text{if } a \neq b \text{ and } Q(i) > 0, \\ (i-1)^{-1}, & \text{otherwise.} \end{cases}$$

It can now be verified that $Q(i) = \frac{1}{i} A(i) (1 - (i-1)B(i))$ for $1 \leq i \leq M$. In the case that $Q(i) > 0$, we have

$$A(i)B(i) = \begin{cases} \frac{a^2(a-1)^2(2m_a - 2x_a - 2i + 1)}{2n^2} (1 + \delta_{aa}(i)), & \text{if } a = b, \\ \frac{a(a-1)b(b-1)(m_a + m_b - x_a - x_b - i + 1)}{2n^2} (1 + \delta_{ab}(i)), & \text{if } a \neq b. \end{cases}$$

So, for $Q(i) > 0$ we have $A(i)B(i) = O(m/n^2)$.

In the case that $Q(i) = 0$, consider $a = b$ for example. We only define $B(i)$ when $Q(1) > 0$. If $Q(1) > 0$ but $Q(i) = 0$, then we must have $L(x_{23}, \dots, x_{p(aa)}, 0) > 0$ but $L(x_{23}, \dots, x_{p(aa)}, i) = 0$. By Lemma 4 this can only happen if $x_a + 2i > m_a$, which again implies that $A(i)B(i) = O(m/n^2)$.

We can now apply Lemma 6 using

$$\begin{aligned} A_1, A_2 &= Q(1)(1 + O(1/n + m/n^2)) \\ &= g_{ab}(x_{23}, \dots, x_{p(ab)}) + O(m^2/n^3 + m^3/n^4), \\ C_1, C_2 &= O(m/n^2), \\ \hat{c} &= \frac{1}{110}. \end{aligned}$$

Since $M \geq \log n$ by definition, we have $(2e\hat{c})^M \leq (e/55)^{\log n} < n^{-3}$. In all cases, we conclude that

$$\sum_{i=0}^M \prod_{j=1}^i Q(j) = \exp(g_{ab}(x_{23}, \dots, x_{p(ab)}) + O(m^2/n^3 + m^3/n^4)).$$

Now insert this value into (8) using

$$S'_{n(ab)}(x_{23}, \dots, x_{p(ab)}, 0) \exp(g_{ab}(x_{23}, \dots, x_{p(ab)})) = S'_{ab}(x_{23}, \dots, x_{p(ab)})$$

and the fact that $g_{rs}(x_{23}, \dots, x_{p(ab)}, 0) = g_{rs}(x_{23}, \dots, x_{p(ab)})$ for $ab, rs \in J$. This completes the induction step and the proof. ■

Proof of Theorem 2. In Lemma 3, we estimated T , and in Lemma 7, we estimated $T/L(0, \dots, 0)$. Dividing the first quantity by the second proves the theorem. ■

6. CONCLUDING REMARKS

We have determined the asymptotic number of partial Steiner systems with a given number of edges of each size, provided there are not too many edges. If the number of edges is increased further, it will be necessary to deal with clusters of more than two edges. This is plausible but challenging. Another open problem is to explore the properties of a random member of this class of hypergraphs, for example its subgraph counts and connectivity.

We thank the referees for their very useful suggestions.

REFERENCES

- [1] V. Blinovsky and C. Greenhill, *Asymptotic enumeration of sparse uniform linear hypergraphs with given degrees*, Electron. J. Combin. **23** (2016) #P3.17. <https://doi.org/10.37236/5512>
- [2] D.A. Grable and K.T. Phelps, *Random methods in design theory: A survey*, J. Combin. Des. **4** (1996) 255–273. [https://doi.org/10.1002/\(SICI\)1520-6610\(1996\)4:4<255::AID-JCD4i3.0.CO;2-E](https://doi.org/10.1002/(SICI)1520-6610(1996)4:4<255::AID-JCD4i3.0.CO;2-E)

- [3] C. Greenhill, B.D. McKay and X. Wang, *Asymptotic enumeration of sparse 0 – 1 matrices with irregular row and column sums*, J. Combin. Theory Ser. A **113** (2006) 291–324.
<https://doi.org/10.1016/j.jcta.2005.03.005>
- [4] B.D. McKay and Fang Tian, *Asymptotic enumeration of linear hypergraphs with given number of vertices and edges* (2019).
arXiv:1908.06333
- [5] R.M. Wilson, *An existence theory for pairwise balanced designs, III: Proof of the existence conjectures*, J. Combin. Theory Ser. A **18** (1975) 71–79.
[https://doi.org/10.1016/0097-3165\(75\)90067-9](https://doi.org/10.1016/0097-3165(75)90067-9)

Received 12 January 2019
Revised 21 September 2019
Accepted 21 September 2019