Combining the theorems of Turán and de Bruijn-Erdős

Sayok Chakravarty* Dhruv Mubayi†

October 22, 2025

Abstract

Fix an integer $s \geq 2$. Let $\mathcal{P}$ be a set of n points and let $\mathcal{L}$ be a set of lines in a linear space such that no line in $\mathcal{L}$ contains more than $(n-1)/(s-1)$ points of $\mathcal{P}$. Suppose that for every s-set S in $\mathcal{P}$, there is a pair of points in S that lies in a line from $\mathcal{L}$. We prove that $|\mathcal{L}| \geq (n-1)/(s-1)+s-1$ for n large, and this is sharp when $n-1$ is a multiple of $s-1$. This generalizes the de Bruijn-Erdős theorem which is the case $s=2$. Our result is proved in the more general setting of linear hypergraphs.

1 Introduction

A finite linear space over a set X is a family $\mathcal{L}$ of its subsets, called lines, such that every line contains at least two points, and any two points are on exactly one line. Two common examples of finite linear spaces are the near pencil and finite projective planes. A near pencil is a family $\mathcal{L}$ over a set X of size n comprising a line of size $n-1$ and $n-1$ lines of size two. A projective plane of order q is a family $\mathcal{L}$ over a set X of size q^2+q+1 with the properties that every line has $q+1$ points and every two points lie on a unique line. Note that these properties imply

$$|\mathcal{L}| = \binom{q^2+q+1}{2} \Big/ \binom{q+1}{2} = q^2+q+1.$$

The following fundamental theorem was proved by de Bruijn and Erdős [5].

Theorem 1 (de Bruijn-Erdős [5]). *If $\mathcal{L}$ is a finite linear space over a set X with $X \notin \mathcal{L}$, then $|\mathcal{L}| \geq |X|$, and equality holds if and only if either $\mathcal{L}$ is a near pencil or $|X| = q^2+q+1$ and $\mathcal{L}$ is a projective plane of order q.*

*Department of Mathematics, Statistics and Computer Science, University of Illinois, Chicago, IL 60607. Email: schakr31@uic.edu. Research partially supported by NSF Award DMS-2153576.

†Department of Mathematics, Statistics and Computer Science, University of Illinois, Chicago, IL 60607. Email: mubayi@uic.edu. Research partially supported by NSF Awards DMS-1952767 and DMS-2153576 and a Simons Fellowship.

This is often viewed as a statement in incidence geometry, in which case it states that the number of lines determined by n points in a projective plane is at least n. The result also has the following graph theoretic formulation: the minimum number of proper complete subgraphs of the complete graph K_n that are needed to partition its edge set is n (see [1] for an extension of this formulation to hypergraphs). Various other extensions have been studied. For instance, [4] considered the problem of determining the minimum number of lines determined by n points in general metric spaces and [6] defined a notion of de Bruijn-Erdős sets in measure spaces and bounded the Hausdorff dimension and Hausdorff measure of such sets. Theorem 1 is also a basic result in extremal set theory and design theory that has many extensions and generalizations. The most notable of these are due to Fisher [7], Bose [3], and Ray-Chaudhuri-Wilson [8].

Here we consider another natural generalization of Theorem 1. We relax the condition that every pair of points lies in a line as follows. An s-set is a set of size s.

Definition 1. *A collection of subsets (lines) $\mathcal{L}$ of a set X is an s-cover if every two lines in $\mathcal{L}$ have at most one point in common and for every s-set $S \subset X$, some pair of points from S lies in a line in $\mathcal{L}$.*

Note that when $s = 2$, this is the definition of a linear space (excluding trivial requirements). We are interested in the minimum size of an s-cover. When $\mathcal{L}$ comprises sets of size exactly two, this minimum is determined by Turán's theorem.

Theorem 2 (Turán [9])**.** *Let $n \geq s \geq 2$. Let G be a graph with vertex set X of size n and edge set $\mathcal{L}$ such that $\mathcal{L}$ is an s-cover of X. Then $|\mathcal{L}| \geq n^2/(2s-2) - n/2$. If $(s-1)|n$ and equality holds, then G is a collection of $s-1$ vertex disjoint cliques, each of size $n/(s-1)$.*

Theorem 2 shows that we can view Definition 1 through the lens of graph theory by considering the graph $G = (X, E)$ where E is the set of pairs not contained in any line in $\mathcal{L}$. Then G is K_s-free when $\mathcal{L}$ is an s-cover. Hence, the s-cover condition can be thought of as a Turán-type property.

As s becomes larger, the requirement for a family to be an s-cover becomes weaker, and hence the number of lines needed for an s-cover decreases. So a natural question is to ask for the size of a smallest s-cover. In order to make this problem nontrivial, we need to impose an upper bound on the size of subsets in $\mathcal{L}$. For example, if we allow sets of size $|X|$, then just one set suffices to cover every pair. Morever, if sets in $\mathcal{L}$ are allowed to be of size greater than $(n-1)/(s-1)$, then we can take a collection of $s-1$ pairwise disjoint sets that cover all the points. This is an s-cover, as any s points will contain two points in one of the sets and will be covered. As $n \to \infty$ this has constant size. Hence the natural condition to obtain a nontrivial result as $n \to \infty$ is that all sets in $\mathcal{L}$ have size at most $t = (n-1)/(s-1)$.

Under this condition, a straightforward construction reminiscient of the construction for Theorem 2 is the following. Assume that $t = (n-1)/(s-1)$ is an integer. The underlying set is a $t \times (s-1)$ grid with an additional new point z, and the line set comprises all

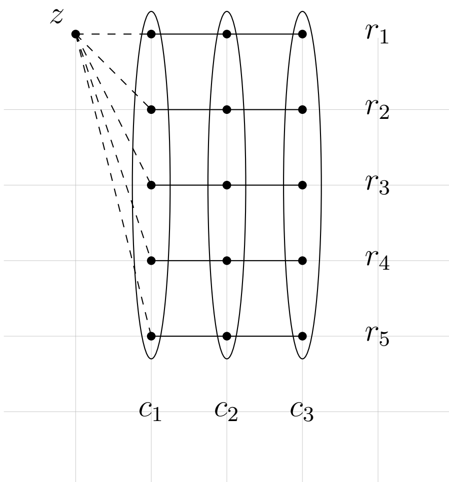

Figure 1: The construction of $\mathcal{L}$ when $t = 5$ and $s = 4$.

columns as well as all rows where we append z to each row (see Figure 1). Formally, $X = ([t] \times [s-1]) \cup \{z\}$, and

$$\mathcal{L} = \{c_i : 1 \le i \le s-1\} \cup \{r_j \cup \{z\} : 1 \le j \le t\},$$

where the ith column is $c_i := [t] \times \{i\}$ and the jth row is $r_j := \{j\} \times [s-1]$. This yields an s-cover with $|\mathcal{L}| = s - 1 + t$.

In this paper, we show that the above construction is tight. Our main result is the following theorem.

Theorem 3. *Fix $s \ge 2$. Let $\mathcal{L}$ be an s-cover of a set of size n. Suppose that each set in $\mathcal{L}$ has size at most $(n-1)/(s-1)$. Then $|\mathcal{L}| \ge (n-1)/(s-1)+s-1$ for n large. If $(s-1) \mid (n-1)$, then there exists an s-cover $\mathcal{L}$ of an n-set such that $|\mathcal{L}| = (n-1)/(s-1)+s-1$. If $(s-1) \nmid (n-1)$, then there exists an s-cover $\mathcal{L}$ of an n-set such that $|\mathcal{L}| \le (1+o(1))(n-1)/(s-1)$ as $n \to \infty$.*

The general framework of Theorem 3 specializes to give appealing geometric statements as given in the abstract or the more special form below.

Corollary 1. *Fix an integer $s \ge 2$. Let $\mathcal{P}$ be a set of n points and let $\mathcal{L}$ be a set of m lines in the plane such that no line in $\mathcal{L}$ contains more than $(n-1)/(s-1)$ points of $\mathcal{P}$. Suppose that for every s-set S of points from $\mathcal{P}$, there is a pair of points in S lies in some line from $\mathcal{L}$. Then $m \ge (n-1)/(s-1) + s - 1$ for n large and if $n-1$ is a multiple of $s-1$, this is tight.*

We remark that equality can hold above as some of the hypergraphs we construct to prove tightness for Theorem 3 can be realized as lines in the plane.

Our proof requires n to be large in terms of s and it remains an open problem to prove the result for small n.

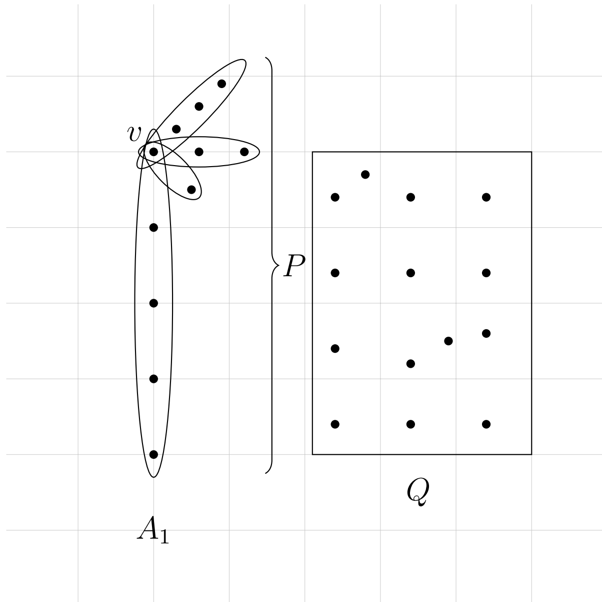

Figure 2: Setup for Lemma 1.

2 Proof of Theorem 3

We will prove Theorem 3 by induction on s. However, in order to facilitate the induction argument, we need to prove a slightly stronger statement for $s \geq 3$ as shown below.

Theorem 4. *Fix $s \geq 3$. Let $\mathcal{L}$ be an s-cover of a set of size n. Suppose that each set in $\mathcal{L}$ has size at most $(n-1)/(s-1)$. Then $|\mathcal{L}| \geq (n-1)/(s-1) + s - 1$ for n large. If each set has size at most $(n-1)/(s-1) - 1$, then $|\mathcal{L}| > (n-1)/(s-1) + s - 1$ for n large.*

2.1 Notation and Lemmas

Let $\mathcal{L} = \{A_1, A_2, \ldots, A_m\}$ be an s-cover on $X := [n]$. Assume $|A_i| = a_i$ and $(n-1)/(s-1) \geq |A_1| \geq |A_2| \geq \cdots \geq |A_m|$. For $x \in X$, the degree of x, written $d(x)$, is the number of A_i that contain x. The neighborhood of x, written $N(x)$, is the collection of A_i's that contain x. Let $d = \min_{w \in A_1} d(w)$ and suppose that $d(v) = d$ where $v \in A_1$. Let $\{A_{i_1}, \ldots, A_{i_d}\}$ be the neighborhood of v where $i_1 = 1$ and let $Q := [n] \setminus \bigcup_{j=1}^{d} A_{i_j}$ and $P := \bigcup_{j=1}^{d} A_{i_j}$. Set $p := |P|$ (see Figure 2). Throughout the proof, we say a subset $J \subset [n]$ is covered if there exists some A_i containing J.

We first prove a lemma giving various bounds on m depending on d, the a_i's, and $|Q|$. We will assume below that Theorem 4 holds for all $s' \leq s - 1$ by induction on s and that n is sufficiently large in terms of s to apply induction and any further inequalities that require this.

The base case $s = 2$ of the induction follows from Theorem 1, so we assume $s > 2$. We note that the stronger statement of Theorem 4 holds only for $s \geq 3$, and we will take care of the specific case $s = 3$ during the proof.

Let $C_1, \delta, \delta_1, \delta_2, \delta_3, \delta_4 > 0$ be constants that follow the hierarchy

$$\frac{1}{C_1^{1/10}} \ll \delta_2 \ll \delta_1 \ll \delta \ll \delta_4 \ll \delta_3 \ll \frac{1}{s^2}$$

where $\xi \ll \eta$ simply means that ξ is a sufficiently small function of η that is needed to satisfy some inequality in the proof. In particular, set

$$\delta_3 = \delta^{1/4} \qquad \text{and} \qquad \delta_4 = \delta^{1/2}.$$

Let $n_0(s)$ be sufficiently large in terms of δ and δ_i above such that all inequalities used in the ensuing proof are valid. In addition, let $n_0(s) \geq C_1 s^8$. We will repeatedly use the fact that if $\beta > 1/10$ and $0 < \alpha < 8\beta$,

$$\frac{s^\alpha}{n^\beta} \leq \frac{s^\alpha}{(C_1 s^8)^\beta} \leq \frac{s^{\alpha-8\beta}}{C_1^\beta} \leq \frac{1}{C_1^\beta} \ll \delta_2. \tag{1}$$

Assume that $n \geq n_0(s)$. We will show that $|\mathcal{L}| \geq (n-1)/(s-1) + s - 1$. Moreover, we will show that $|\mathcal{L}| > (n-1)/(s-1) + s - 1$ when $a_1 \leq (n-1)/(s-1) - 1$.

Lemma 1. *The following bounds hold for $n \geq n_0(s)$.*

1. $m \geq a_1(d-1) + 1$
2. *If $|Q| > n(s-2)/(s-1)$, then*

$$m \geq \frac{|Q|-1}{s-2} + s - 2 + d.$$

3. *For distinct $i_1, i_2, ..., i_s$, let $A'_{i_k} \subset A_{i_k}$ be pairwise disjoint subsets, $a'_{i_k} = |A'_{i_k}|$. Then,*

$$m \geq \frac{a'_{i_1} a'_{i_2} a'_{i_3} ... a'_{i_s}}{e_{s-2}(a'_{i_1}, a'_{i_2}, a'_{i_3}, \ldots, a'_{i_s})}$$

   where $e_k(x_1, \ldots, x_n) = \sum_{1 \leq i_1 < \cdots < i_k \leq n} x_{i_1} \cdots x_{i_k}$ is the k-th elementary symmetric polynomial.

4. $\sum_{i=1}^m \binom{a_i}{2} \geq \frac{n^2}{2(s-1)} - \frac{n}{2}$.
5. *Suppose $a_1 \geq (1-\delta) \cdot \sqrt{n}$. Write $n - 1 + (s-1)(s-2) = (s-1)a_1 q + r$ for integers q, r with $0 \leq r < (s-1)a_1$. If $d \notin \{q, q+1\}$, then $m > \frac{n-1}{s-1} + s - 1$.*

Proof. We prove each statement of the lemma.

1. The number of sets containing a point in A_1 is at least $a_1(d-1)+1$ as d is the minimum degree of points in A_1.

2. Observe that

$$\frac{|Q|-1}{s-2} > \frac{\frac{n(s-2)}{s-1}-1}{s-2} = \frac{n}{s-1} - \frac{1}{s-2}.$$

This implies $(|Q|-1)(s-1) \geq n(s-2)-(s-1)+1 = (n-1)(s-2)$. Rearranging, we get $(n-1)/(s-1) \leq (|Q|-1)/(s-2)$. This shows the size of all sets not in $N(v)$ is less than $(|Q|-1)/(s-2)$. Also,

$$|Q| > \frac{s-2}{s-1}n \geq \frac{s-2}{s-1}n_0(s) \geq n_0(s-1).$$

For any $s-1$ distinct points $x_1,\dots,x_{s-1} \in Q$, there must be a set A_t containing a pair from $x_1,\dots,x_{s-1}$ as the s-set $\{x_1,\dots,x_{s-1},v\}$ must be covered and $x_1,\dots,x_{s-1} \notin P$. Therefore the collection of sets $A_i \setminus P$ that have at least one point in Q is an $(s-1)$-cover of Q. By the induction hypothesis, the number of these sets is at least $(|Q|-1)/(s-2)+s-2$. Hence $m \geq (|Q|-1)/(s-2)+s-2+d$ as there are d sets containing v in addition to these A_i.

3. Consider the collection of s-sets $B = \{\{x_1,\dots,x_s\} : x_j \in A'_{i_j}, j = 1,\dots,s\}$. A particular A_t covers at most $e_{s-2}(a'_{i_1},\dots,a'_{i_s})$ such s-sets as it has at most one point in each of $A_{i_1}, \dots, A_{i_s}$. The number of s-sets in B is $a'_{i_1}\cdots a'_{i_s}$. It follows that

$$m \geq \frac{a'_{i_1}\cdots a'_{i_s}}{e_{s-2}(a'_{i_1},\dots,a'_{i_s})}.$$

4. Let $G = ([n], E)$ be the graph where E is the set of pairs not contained in any A_i. Then G is K_s-free. Since every pair in $[n]$ is either in some A_i or in E, we have

$$\binom{n}{2} = \sum_{i=1}^{m}\binom{a_i}{2} + |E|$$

Since $|E| \leq n^2/2 \cdot (1-1/(s-1))$ by Theorem 2,

$$\sum_{i=1}^{m}\binom{a_i}{2} \geq \binom{n}{2} - \frac{n^2}{2}\left(1-\frac{1}{s-1}\right) = \frac{n^2}{2(s-1)} - \frac{n}{2}.$$

5. We have $n-1+(s-1)(s-2) = (s-1)a_1q + r$ for integers q, r with $0 \leq r < (s-1)a_1$. If $d \geq q+2$, then part 1 implies

$$\begin{aligned} m &\geq a_1(q+1)+1 \\ &= \frac{n-1+(s-1)(s-2)-r}{s-1} + a_1 + 1 \\ &= \frac{n-1}{s-1} + s - 1 + a_1 - \frac{r}{s-1} \\ &> \frac{n-1}{s-1} + s - 1. \end{aligned}$$

If $d \le q-1$, then

$$\begin{aligned}
|Q| &\geq n - d(a_1 - 1) - 1 \\
&\geq n - (q-1)(a_1 - 1) - 1 \\
&= n - qa_1 + q + a_1 - 1 - 1 \\
&= n - \frac{n - 1 + (s-1)(s-2) - r}{s-1} + q + a_1 - 2 \\
&> \frac{s-2}{s-1} n - (s-2) + q + a_1 - 2 \\
&= \frac{s-2}{s-1} n + q + a_1 - s.
\end{aligned}$$

Note that

$$a_1 - s \geq (1-\delta)\sqrt{n_0(s)} - s > s.$$

It follows that $q + a_1 - s \geq s$, so $|Q| > n(s-2)/(s-1) + s$. By part 2

$$m \geq \frac{|Q| - 1}{s-2} + s - 2 \geq \frac{\frac{s-2}{s-1} n + s}{s-2} + s - 2 > \frac{n-1}{s-1} + s - 1.$$

□

2.2 Sketch of Proof of Lower Bound of Theorem 4

In this section, we give a sketch of the proof of the lower bound for $s = 3$. This will serve to illustrate the main ideas of the proof without dealing with cumbersome calculations. These will appear later in Section 2.3. Define $n_0 := n_0(3)$ and let $n \geq n_0$. Let $\varepsilon > 0$ be a small constant. We consider different ranges for a_1.

Case 1: $a_1 < (1-\delta)\sqrt{n}$.

Recall the graph from Lemma 1 part 4: $G = ([n], E)$ where E is the set of pairs that are not contained in any A_i. Observe that $\mathcal{L}$ being a 3-cover implies that G is triangle free, so $|E| \leq n^2/4$ by Theorem 2. Since every pair is either in E or contained in some A_i, we have

$$\binom{n}{2} = \sum_{i=1}^{m} \binom{a_i}{2} + |E| \leq m \binom{a_1}{2} + |E|.$$

Hence, the bound $|E| \leq n^2/4$ yields a lower bound for m. By a short calculation, this implies that $m \geq 1/(1-\delta)^2 \cdot (n/2 - 1)$. Since $\delta > 0$, for n large enough we have $m > (n-1)/2 + 2$.

Case 2: $(1-\delta)\sqrt{n} < a_1 < 10\sqrt{3n}$.

By Lemma 1 part 5, we can assume d is either q or $q+1$. Recall that $n - 1 + (s-1)(s-2) = (s-1)a_1 q + r$ for integers q, r with $0 \leq r < (s-1)a_1$. In particular,

$$q = \left\lfloor \frac{n-1}{2a_1} \right\rfloor = \Omega\left(\sqrt{n}\right).$$

If $|Q| > n/2$, then by Lemma 1 part 2

$$m \geq |Q| - 1 + 1 + d > \frac{n}{2} + d.$$

Since $d \geq q = \Omega(\sqrt{n})$, this implies $m > (n-1)/2 + 2$ if $|Q| > n/2$. Note that we did not need the precise constant $10\sqrt{3}$ for this argument as long as n is sufficiently large. On the other hand, if $|Q| \leq n/2$, then $p := |P| = n - |Q| \geq n/2$. Let G_P be the subgraph of G from Lemma 1 part 4 induced by P. Let P' be the edge set of the complement of G_P. Then G_P is triangle free, so it has at most $p^2/4$ edges by Theorem 2. It follows that $|P'| \geq p^2/4 - p/2$. We will now bound the number of elements in $\mathcal{L}$ it takes to cover the elements of P'.

The number of pairs covered by $A_{i_1}, \ldots, A_{i_d}$ is

$$\sum_{j=1}^{d} \binom{a_{i_j}}{2} \leq d\binom{a_1}{2} \leq \frac{a_1^2 d}{2}.$$

Since any $A_i \in \mathcal{L} \setminus N(v)$ has at most d points in P,

$$m \geq \frac{\frac{p^2}{4} - \frac{p}{2} - \frac{a_1^2 d}{2}}{\binom{d}{2}}$$

by bounding the number of $A_i \in \mathcal{L} \setminus N(v)$ that are needed to cover the pairs in P'. A simple calculation yields

$$m \geq (1-\delta_2)\frac{p^2}{(s-1)d^2} \geq \frac{1-\delta_2}{(1+\delta_2)^2}\frac{a_1^2}{2}. \tag{2}$$

If $a_1 \geq (1+\delta_1)\sqrt{n}$, then by (2), we get

$$m \geq \frac{(1-\delta_2)(1+\delta_1)^2}{(1+\delta_2)^2}\frac{n}{2}.$$

Our choice of constants yields $(1-\delta_2)(1+\delta_1)^2/(1+\delta_2)^2 > 1$; consequently $m > (n-1)/2+2$. Hence we may assume that $a_1 < (1+\delta_1)\sqrt{n}$.

One can show that to cover the pairs in P', there must be at least $(1-o(1))n/2$ elements of $\mathcal{L} \setminus N(v)$ with at least $(1-o(1))\sqrt{n}/2$ points in P. If $|P'| \approx n^2/16$, then the total number of pairs covered by these sets is

$$\frac{n}{2}\binom{\sqrt{n}/2}{2} \approx |P'|.$$

Note that $A_{i_1}, \ldots, A_{i_d}$ cover at most $O(n^{3/2})$ pairs so they do not contribute significantly to this calculation. The precise statement we will need is the following. We omit the somewhat technical proof here.

Claim 1. *At least $(1-\delta_3)n/2$ sets in $\mathcal{L} \setminus N(v)$ have at least $(1/2-\delta_4)\sqrt{n}$ points in P.*

We now use Claim 1 to get a contradiction. In particular, we will show that not enough pairs in Q are covered. Suppose $A \in \mathcal{L} \setminus N(v)$. If A has at least $(1/2-\delta_4)\sqrt{n}$ points in P, then it has at most $(1+\delta_1-1/2+\delta_4)\sqrt{n}$ points in Q as $a_1<(1+\delta_1)\sqrt{n}$. Hence, by Claim 1 and the fact that there are at most $n/2$ sets in $\mathcal{L}\setminus N(v)$, the number of covered pairs in Q is at most

$$\begin{aligned}&\frac{(1-\delta_3)\,n}{2}\binom{(1/2+\delta_1+\delta_4)\sqrt{n}}{2}+\frac{\delta_3 n}{2}\binom{(1+\delta_1)\sqrt{n}}{2}\\ &\leq\frac{1}{2}\left(\frac{1}{2}+\delta_1+\delta_4\right)^2\frac{(1-\delta_3)n^2}{2}+\frac{(1+\delta_1)^2\delta_3 n^2}{4}. \qquad (3)\end{aligned}$$

On the other hand, $|Q|\geq n-d(a_1-1)-1$. Using a more precise bound for q, one can show that this is at least $(n-1)/2-a_1$. Since at least $|Q|^2/4-|Q|/2$ pairs in Q must be covered by Theorem 2 and

$$\frac{|Q|^2}{4}-\frac{|Q|}{2}\geq(1-o(1))\frac{1}{16}n^2,$$

we have a contradiction as it can be shown that the coefficient of n^2 in (3) is smaller than $1/16$.

Case 3: $10\sqrt{3n}\leq a_1\leq\left(\frac{1}{2}-\varepsilon\right)n$

By Lemma 1 part 5, we can assume $d\in\{q,q+1\}$. Suppose $d=1$. Then, $|Q|=n-a_1\geq(1/2+\varepsilon)n$, so by Lemma 1 part 2 for $n\geq n_0$

$$m>|Q|-1+1\geq\left(\frac{1}{2}+\varepsilon\right)n>\frac{n-1}{s-1}+s-1.$$

Hence, we can assume $d\geq 2$. If $|Q|>n/2$, then for $n\geq n_0$

$$m\geq|Q|-1+1+d\geq\left(\frac{1}{2}+\varepsilon\right)n+d>\frac{n-1}{2}+2. \qquad (4)$$

by Lemma 1 part 2. Hence, we can assume $p=n-|Q|\geq n/2$. We consider the case $d\geq 5$ and $d\leq 4$ separately. Let $d\geq 5$. We will apply Lemma 1 part 3 on three sets in P. It can be shown that there are at most 2 elements of $\mathcal{L}$ in $N(v)$ that have size less than $a_1/2$. As we are assuming $d\geq 5$, there are at least 3 sets with size more than $a_1/2-1\geq 4\sqrt{3n}$. Taking disjoint subsets of these 3 sets of size at least $4\sqrt{3n}$ and applying Lemma 1 part 3 we get

$$m\geq\frac{(4\sqrt{3n})^3}{\binom{3}{2}(4\sqrt{3n})}=\frac{48n}{3}=16n>\frac{n-1}{2}+2$$

for $n\geq n_0$.

We now consider the case in which $d\leq 4$. Since $n/2\leq p\leq a_1 d$, we have

$$a_1\geq\frac{n}{2d}\geq\frac{n}{8}.$$

Let $A_1=\{x_1,\ldots,x_{a_1}\}$. For $1\leq i\leq a_1$, let T_i be the number of covered pairs in the neighborhood of x_i excluding those containing x_i. Then,

$$T_i=\sum_{j:\,x_i\in A_j,j\neq 1}\binom{a_j-1}{2}.$$

Let $B = \{i : d(x_i) \leq 9\}$. Then we have

$$\sum_{i \in B} T_i < n^2$$

Note that we may assume $|B| = \Omega(n)$ as if not there will more that $(n-1)/2 + 2$ elements of $\mathcal{L}$ containing points in A_1. It follows that there is $\ell \in B$ so that $T_\ell = O(n)$. By Jensen's inequality,

$$\sum_{k:\, x_\ell \in A_k, k \neq 1} (a_k - 1) = O(\sqrt{n}).$$

Note that the above sum is the number of points in $N(x_i)$ excluding points in A_1. Let Q_1 be the set of points outside of the neighborhood of x_ℓ. Then every pair in $[n] \setminus Q_1$ is covered. Hence, we can apply Theorem 1 on Q_1. Since $|Q_1| \geq n - a_1 - O(\sqrt{n}) \geq (1/2 + \varepsilon)n - O(\sqrt{n})$, this shows that $m > (n-1)/2 + 2$.

Case 4: $(1/2 - \varepsilon)n < a_1 < \lfloor (n-1)/2 \rfloor$

We first note that if $d = 1$, then by Lemma 1 part 2 we have $m > (n-1)/2 + 2$. If $d \geq 3$, then $m > 2a_1 > (1 - 2\varepsilon)n > (n-1)/2 + 2$. It follows that we can assume $d = 2$. We can also assume that the number of points in A_1 with degree at least 3 is at most $(\varepsilon + 1/8)n$ as otherwise we get $m > (n-1)/2 + 1$ by counting the elements in $\mathcal{L}$ that intersect A_1. We also observe that if all of the pairs in $[n] \setminus A_1$ are covered, then $m \geq |Q| + 1 > (n-1)/2 + 2$ by Theorem 1. Consequently, we may assume that

1. $d = 2$
2. There are at most $(\varepsilon + 1/8)n$ points in A_1 with degree at least 3
3. There is a pair $\{x_1, x_2\}$ in $[n] \setminus A_1$ that is not covered

For any $x_0 \in A_1$ the 3-set $\{x_0, x_1, x_2\}$ is covered, so there is a set containing a pair x_0, x_i for some $i \in \{1, 2\}$. Set

$$B_{x_i} = \{x_0 \in A_1 : x_0, x_i \in A_j \text{ for some } j\}$$

for $1 \leq i \leq 2$. Without loss of generality, assume $|B_{x_1}| \geq |B_{x_2}|$. Then

$$|B_{x_1}| \geq \frac{a_1}{2} > \left(\frac{1}{4} - \frac{\varepsilon}{2}\right) n.$$

Let $B'_{x_1} \subset B_{x_1}$ be the points in B_{x_1} that have degree two. Since the number of points in A_1 with degree greater than two is at most $(\varepsilon + 1/8)n$, we have $|B'_{x_1}| \geq (1/8 - 3\varepsilon/2)n$.

Let $Q_2 := [n] \setminus (A_1 \cup \{x_1\})$. It can be shown that all pairs in Q_2 are covered. We leave this argument for the detailed proof, but the idea is that if some pair $\{u_1, u_2\}$ in Q_2 is uncovered, we can consider triples $\{u_0, u_1, u_2\}$ where $u_0 \in B'_{x_1}$ and get a contradiction using the fact that points in B'_{x_1} have degree 2. If $a_1 < (n-1)/2 - 1$, then we get $m > (n-1)/2 + 2$ by using Theorem 1 on Q_2. When $a_1 = (n-1)/2 - 1$, then this argument does not yield a strict inequality for m, so we need to use a more subtle argument involving the extremal constructions for Theorem 1.

Suppose $a_1 = (n-1)/2 - 1$, $|Q_2| = (n-1)/2 + 1$, and every pair in Q_2 is covered. Define

$$\mathcal{L}_2 := \{A \cap Q_2 : A \in \mathcal{L},\ |A \cap Q_2| \geq 2\}.$$

By Theorem 1, $|\mathcal{L}_2| \geq (n-1)/2 + 1$. If $|\mathcal{L}_2| > (n-1)/2 + 1$, then $m > (n-1)/2 + 2$ as $A_1 \cap Q_2 = \emptyset$. Hence we can assume $|\mathcal{L}_2| = (n-1)/2 + 1$. By Theorem 1, $\mathcal{L}_2$ is either a near pencil or projective plane. If $\mathcal{L}_2$ is a near pencil then, there is a set of size $(n-1)/2$ in $\mathcal{L}_2$. This is a contradiction as the largest set in $\mathcal{L}$ has size $a_1 = (n-1)/2 - 1$. Suppose now that $\mathcal{L}_2$ is a projective plane, so $|\mathcal{L}_2| = |Q_2| = (n-1)/2$. Recall that x_1 has degree at least $|B_{x_1}| \geq a_1/2 \geq 3$. If $N(x_1)$ contains A_i, A_j such that neither $A_i \cap Q_2$ nor $A_j \cap Q_2$ is in $\mathcal{L}_2$, then $m \geq |\mathcal{L}_2| + 3 > (n-1)/2 + 2$. So $N(x_1)$ contains A_i, A_j such that both $A_i \cap Q_2$ and $A_j \cap Q_2$ are in $\mathcal{L}_2$. But $\mathcal{L}_2$ is a projective plane hence $(A_i \cap Q_2) \cap (A_j \cap Q_2) \neq \emptyset$, which means that $|A_i \cap A_j| \geq 2$, a contradiction. Hence $m > (n-1)/2 + 2$ when $a_1 = (n-1)/2 - 1$.

Case 5: $a_1 = \lfloor (n-1)/2 \rfloor$.

We note that for $x > 1$, we have $\lfloor x \rfloor > x - 1$ so in this case $a_1 > (n-1)/2 - 1$ and we only need to prove that $m \geq (n-1)/2 + 2$. By the same argument as Case 4 we can assume that

1. $d = 2$
2. There are at most $s - 1$ points in A_1 with degree at least 3
3. There is a pair $\{x_1, x_2\}$ in A_1 that is not covered

Define $B_{x_1}, B_{x_2}, B'_{x_1}$ as in Case 4. Let $Q_2 = [n] \setminus A_1 \cup \{x_1\}$. It can be shown that all pairs in Q_2 are covered by the same argument as Case 4.

Define

$$\mathcal{L}_2 := \{A \cap Q_2 : A \in \mathcal{L},\ |A \cap Q_2| \geq 2\}.$$

Suppose first that n is odd, say $n = 2k + 1$ for some integer k. Then, $a_1 = k$ and $|Q_2| = k$. Then either $Q_2 \in \mathcal{L}$ or by Theorem 1 $|\mathcal{L}_2| \geq k$. Suppose $Q_2 \in \mathcal{L}$. Since A_1 has minimum degree 2, we have $m \geq k + 2 = (n-1)/2 + 2$ as required. Now suppose $Q_2 \notin \mathcal{L}$. By the Theorem 1, $|\mathcal{L}_2| \geq k$. If $|\mathcal{L}_2| \geq k + 1$, we have $m \geq k + 2$ as $A_1 \cap Q_2 = \emptyset$. If $|\mathcal{L}_2| = k$, then $\mathcal{L}_2$ is either a near pencil or projective plane. In either case, any pair of sets in Q_2 intersects in exactly one point. Suppose $m = k + 1$ for the sake of contradiction. Then since x_1 has degree at least $a_1/(s-1) = a_1/2$, there must be a collection of pairwise disjoint sets of size $a_1/2$ in Q_2 (i.e. $a_1/2$ pairwise disjoint sets of the form $A_i \cap Q_2$). This is a contradiction, so we have $m \geq k + 2$.

Now, suppose $n = 2k$. Then, $a_1 = k - 1$ and $|Q_2| = k$. Since $|Q_2| > a_1$, by Theorem 1 $|\mathcal{L}_2| \geq k$. This implies $m \geq k + 1$. By the same argument as above using the fact that x_1 has degree at least $a_1/2$, we have $m \neq k + 1$, so $m \geq k + 2$.

2.3 Proof of Lower Bound for Theorem 3

We prove Theorem 1 by considering different ranges for a_1. Let $n \geq n_0(s)$ and set $\varepsilon = 1/10s^2$.

Case 1: $a_1 < (1-\delta)\sqrt{n}$.

From Lemma 1 part 4, we have $\sum_{i=1}^m \binom{a_i}{2} \geq n^2/2(s-1) - n/2$. Since $\sum_{i=1}^m \binom{a_i}{2} \leq ma_1^2/2 < m(1-\delta)^2 n/2$, we have

$$\frac{(1-\delta)^2}{2} nm \geq \frac{n^2}{2(s-1)} - \frac{n}{2}.$$

This implies

$$m \geq \frac{\frac{n^2}{2(s-1)} - \frac{n}{2}}{\frac{(1-\delta)^2}{2} n} = \frac{1}{(1-\delta)^2}\left(\frac{n}{s-1} - 1\right).$$

For $n \geq n_0(s)$, this is greater than $(n-1)/(s-1) + s - 1$.

Case 2: $(1-\delta)\sqrt{n} \leq a_1 \leq 10\sqrt{sn}$.

By Lemma 1 part 5, we can assume d is either q or $q+1$. Suppose $|Q| > n\left(1 - 1/(s-1)\right)$. Then by Lemma 1 part 2

$$m \geq \frac{|Q|-1}{s-2} + s - 2 + d > \frac{n}{s-1} - \frac{1}{s-2} + s - 2 + d.$$

Since

$$\frac{n-1}{(s-1)a_1} < \frac{n-1}{(s-1)a_1} + \frac{s-2}{a_1} = q + \frac{r}{(s-1)a_1} < q+1,$$

we have

$$d \geq q \geq \frac{n-1}{a_1(s-1)} - 1 \geq \frac{n-1}{10\sqrt{s}(s-1)\sqrt{n}} - 1. \tag{5}$$

Hence, for $n \geq n_0(s)$

$$m > \frac{n}{s-1} - \frac{1}{s-2} + s - 2 + d > \frac{n-1}{s-1} + s - 1.$$

Therefore $|Q| \leq n(1 - 1/(s-1))$ and $p := |P| = n - |Q| \geq n/(s-1)$. By Theorem 2, at least $p^2/2(s-1) - p/2$ of the pairs in P must be covered. The number of pairs covered by $A_{i_1}, \dots, A_{i_d}$ is

$$\sum_{j=1}^d \binom{a_{i_j}}{2} \leq d\binom{a_1}{2} \leq \frac{a_1^2 d}{2}.$$

Since a set A_i in $\mathcal{L} \setminus N(v)$ has at most d points from P, it covers at most $\binom{d}{2}$ pairs. So

$$m \geq \frac{\frac{p^2}{2(s-1)} - \frac{p}{2} - \sum_{j=1}^d \binom{a_{i_j}}{2}}{\frac{d^2}{2}} \geq \frac{p^2}{(s-1)d^2}\left(1 - \frac{s-1}{p} - \frac{a_1^2 d(s-1)}{p^2}\right). \tag{6}$$

Note that

$$\frac{a_1^2 d(s-1)}{p^2} \leq \frac{100snd(s-1)}{n^2/(s-1)^2} = \frac{100ds(s-1)^3}{n}.$$

Observe that

$$d \le q+1 \le \frac{n-1}{a_1(s-1)} + \frac{s-2}{a_1} + 1 \tag{7}$$

and by (1)

$$\begin{aligned}\frac{n-1}{a_1(s-1)} + \frac{s-2}{a_1} + 1 &\le \frac{n-1}{(1-\delta)\sqrt{n}(s-1)} + \frac{s-2}{(1-\delta)\sqrt{n}} + 1 \\ &\le \frac{\sqrt{n}}{(1-\delta)(s-1)} + \delta_2 + 1.\end{aligned} \tag{8}$$

Consequently, by (7),

$$d(s-1) \le \frac{\sqrt{n}}{1-\delta} + (s-1)(1+\delta_2).$$

This and (1) imply

$$\begin{aligned}\frac{100ds(s-1)^3}{n} &\le \frac{100s^3}{n}d(s-1) \\ &\le \frac{100s^3}{n}\left(\frac{\sqrt{n}}{1-\delta} + (s-1)(1+\delta_2)\right) \\ &= \frac{100s^3}{\sqrt{n}(1-\delta)} + \frac{100s^3(s-1)(1+\delta_2)}{n} \\ &\le \frac{\delta_2}{2}.\end{aligned}$$

It follows that $a_1^2d(s-1)/p^2 \le \delta_2/2$. As $p \ge n/(s-1)$, we have $(s-1)/p \le (s-1)^2/n \le s^2/n \le \delta_2/2$ by (1) and hence by (6)

$$m \ge (1-\delta_2)\frac{p^2}{(s-1)d^2}. \tag{9}$$

We will now prove a lower bound for $p^2/(s-1)d^2$. By (7),

$$d(s-1) \le \frac{n-1}{a_1} + \frac{(s-2)(s-1)}{a_1} + s - 1 < \frac{n}{a_1} + \frac{s^2}{a_1} + s,$$

and this, $a_1 \le 10\sqrt{sn}$, and (1) yield

$$\begin{aligned}\frac{p^2}{(s-1)d^2} &\ge \frac{(n/(s-1))^2}{(s-1)d^2} \\ &= \frac{n^2}{s-1}\frac{1}{d^2(s-1)^2} \\ &\ge \frac{n^2}{s-1}\left(\frac{n}{a_1} + \frac{s^2}{a_1} + s\right)^{-2} \\ &= \frac{a_1^2}{s-1}\left(1 + \frac{s^2}{n} + \frac{a_1 s}{n}\right)^{-2} \\ &\ge \frac{a_1^2}{s-1}\left(1 + \frac{s^2}{n} + \frac{10s^{3/2}}{\sqrt{n}}\right)^{-2} \\ &\ge \frac{a_1^2}{s-1}(1+\delta_2)^{-2}.\end{aligned} \tag{10}$$

Combining (10) and (9), we get

$$m \geq \frac{1-\delta_2}{(1+\delta_2)^2} \frac{a_1^2}{s-1}. \tag{11}$$

If $a_1 \geq (1+\delta_1)\sqrt{n}$, then by (11),

$$m \geq (1+\delta_1)^2 \frac{1-\delta_2}{(1+\delta_2)^2} \frac{n}{s-1}.$$

Since $\delta_1 \gg \delta_2$, we have $(1+\delta_1)^2(1-\delta_2)/(1+\delta_2)^2 > 1$. It follows that $m > (n-1)/(s-1)+s-1$ for $n \geq n_0(s)$.

Hence we may assume $a_1 < (1+\delta_1)\sqrt{n}$. We can also assume that there are at most $n/(s-1)$ sets in $\mathcal{L} \setminus N(v)$ as otherwise $m \geq n/(s-1) + d > (n-1)/(s-1) + s - 1$ by (5). Recall that $\delta_3 = \delta^{1/4}$ and $\delta_4 = \delta^{1/2}$.

Claim 2. *At least $(1-\delta_3)n/(s-1)$ sets in $\mathcal{L} \setminus N(v)$ have at least $(1/(s-1) - \delta_4)\sqrt{n}$ points in P.*

Proof. Assume this is not true. Since every $A \in \mathcal{L} \setminus N(v)$ has at most one point in common with every set in $N(v)$, we conclude that $|A \cap P| \leq d$. Hence the number of covered pairs in P is at most

$$(1-\delta_3)\frac{n}{s-1}\frac{d^2}{2} + \frac{\delta_3 n}{s-1}\left(\frac{1}{s-1} - \delta_4\right)^2 \frac{n}{2} + \sum_{j=1}^{d} \binom{a_{i_j}}{2}. \tag{12}$$

Recall that $d \leq \sqrt{n}/((1-\delta)(s-1)) + 1 + \delta_2$ by (8). By this bound and (1) we have

$$\begin{aligned}
(1-\delta_3)\frac{n}{(s-1)}\frac{d^2}{2} &\leq (1-\delta_3)\frac{n}{2(s-1)}\left(\frac{\sqrt{n}}{(1-\delta)(s-1)} + \delta_2 + 1\right)^2 \\
&= (1-\delta_3)\frac{n}{2(s-1)}\frac{n}{(1-\delta)^2(s-1)^2}\left(1 + \frac{(\delta_2+1)(1-\delta)(s-1)}{\sqrt{n}}\right)^2 \\
&\leq \frac{1-\delta_3}{(1-\delta)^2}\frac{n^2}{2(s-1)^3}(1+\delta_2)^2.
\end{aligned} \tag{13}$$

We also have

$$\frac{\delta_3 n}{s-1}\left(\frac{1}{s-1} - \delta_4\right)^2 \frac{n}{2} = \frac{\delta_3 n^2}{2(s-1)^3}\left(1 - \delta_4(s-1)\right)^2. \tag{14}$$

Since $a_1 < (1+\delta_1)\sqrt{n}$ and $d \leq 2\sqrt{n}/((1-\delta)(s-1))$ by (7), we have

$$\begin{aligned}
\sum_{j=1}^{d} \binom{a_{i_j}}{2} &\leq \frac{a_1^2}{2} d \\
&\leq \frac{(1+\delta_1)^2 n}{2}\frac{2\sqrt{n}}{(1-\delta)(s-1)} \\
&\leq \frac{n^2}{2(s-1)^3}\frac{2(s-1)^2(1+\delta_1)^2}{(1-\delta)\sqrt{n}} \\
&\leq \delta_2 \frac{n^2}{2(s-1)^3}.
\end{aligned} \tag{15}$$

Note that we used (1) in the last step. Combining (13), (14), and (15), we deduce that the number of covered pairs in P is at most

$$\frac{n^2}{2(s-1)^3}\left(\frac{(1+\delta_2)^2(1-\delta_3)}{(1-\delta)^2}+\delta_3\left(1-\delta_4(s-1)\right)^2+\delta_2\right). \tag{16}$$

Observe that

$$(1+\delta_2)^2 = 1+2\delta_2+\delta_2^2 \le 1+2\delta_2+\delta_2 = 1+3\delta_2$$

and

$$\frac{1}{(1-\delta)^2} = \frac{1}{1-2\delta+\delta^2} \le 1+3\delta$$

since $(1-2\delta+\delta^2)(1+3\delta) = 1+\delta-5\delta^2+3\delta^3 \ge 1$ for δ small. As $\delta \gg \delta_2$, we obtain

$$\begin{aligned}\frac{(1+\delta_2)^2(1-\delta_3)}{(1-\delta)^2} &\le (1+3\delta_2)(1+3\delta)(1-\delta_3)\\ &\le (1+4\delta)(1-\delta_3)\\ &= 1-\delta_3(1+4\delta-4\delta/\delta_3).\end{aligned}$$

As $\delta_3\delta_4 = \delta^{1/4}\delta^{1/2} = \delta^{3/4} \gg \delta$, we have $4\delta+\delta_4(s-1)/2 > 4\delta/\delta_3$ and hence

$$1-\delta_3(1+4\delta-4\delta/\delta_3) \le 1-\delta_3(1-\delta_4(s-1)/2).$$

We also have $\delta_3\left(1-\delta_4(s-1)\right)^2 \le \delta_3\left(1-\delta_4(s-1)\right)$. It follows that (16) is at most

$$\begin{aligned}&\frac{n^2}{2(s-1)^3}\left(1-\delta_3\left(1-\frac{\delta_4(s-1)}{2}\right)+\delta_3\left(1-\delta_4(s-1)\right)+\delta_2\right)\\ &\le \frac{n^2}{2(s-1)^3}\left(1-\frac{\delta_3\delta_4(s-1)}{2}+\delta_2\right).\end{aligned}$$

Note that $1-\delta_3\delta_4(s-1)/2+\delta_2 < 1$ as $\delta_3\delta_4 \gg \delta \gg \delta_2$. Since $p^2/2(s-1) \ge n^2/2(s-1)^3$, this implies that for $n \ge n_0(s)$ the number of covered pairs in P is less that $p^2/2(s-1)-p/2$. This contradiction completes the proof of the claim. □

Suppose $A \in \mathcal{L}\setminus N(v)$. If A has at least $(1/(s-1)-\delta_4)\sqrt{n}$ points in P, then it has at most $(1+\delta_1-1/(s-1)+\delta_4)\sqrt{n}$ points in Q as $a_1 < (1+\delta_1)\sqrt{n}$. Hence, by Claim 2 the number of covered pairs in Q is at most

$$\begin{aligned}&\frac{(1-\delta_3)\,n}{s-1}\binom{(1+\delta_1-1/(s-1)+\delta_4)\sqrt{n}}{2}+\frac{\delta_3 n}{s-1}\binom{(1+\delta_1)\sqrt{n}}{2}\\ &\le \frac{1}{2}\left(1+\delta_1+\delta_4-\frac{1}{s-1}\right)^2\frac{(1-\delta_3)n^2}{s-1}+\frac{(1+\delta_1)^2\delta_3 n^2}{2(s-1)}.\end{aligned} \tag{17}$$

Note that by (7)

$$\begin{aligned}|Q| &\ge n-d(a_1-1)-1\\ &\ge n-da_1\\ &\ge n-\left(\frac{n-1}{s-1}+s-2+a_1\right)\\ &= \left(1-\frac{1}{s-1}\right)n+\frac{1}{s-1}-s+2-a_1.\end{aligned} \tag{18}$$

Since every $(s-1)$-set in Q is covered, at least $|Q|^2/2(s-2) - |Q|/2$ pairs in Q must be covered by Theorem 2. We now show that

$$\frac{1}{2}\left(1+\delta_1+\delta_4-\frac{1}{s-1}\right)^2\frac{(1-\delta_3)}{s-1}+\frac{(1+\delta_1)^2\delta_3}{2(s-1)}<\frac{1}{2(s-2)}\left(1-\frac{1}{s-1}\right)^2. \qquad (19)$$

To see this, first note that $(1+\delta_1)^2\delta_3/2(s-1) \le 1/100s^3$ as $\delta_3 \ll 1/s^2$. Next

$$\begin{aligned}
&\frac{1}{2}\left(1+\delta_1+\delta_4-\frac{1}{s-1}\right)^2\frac{(1-\delta_3)}{s-1}\\
&\le \frac{1}{2(s-1)}\left(\left(1-\frac{1}{s-1}\right)^2+3(\delta_1+\delta_4)\right)(1-\delta_3)\\
&\le \frac{1}{2(s-1)}\left(1-\frac{1}{s-1}\right)^2-\frac{\delta_3}{2(s-1)}\left(1-\frac{1}{s-1}\right)^2+3(\delta_1+\delta_4)\\
&\le \frac{1}{2(s-1)}\left(1-\frac{1}{s-1}\right)^2-\frac{\delta_3}{3s}\\
&= \frac{1}{2(s-2)}\left(1-\frac{1}{s-1}\right)^2-\frac{1}{2(s-1)(s-2)}\left(1-\frac{1}{s-1}\right)^2-\frac{\delta_3}{3s}
\end{aligned}$$

as $\delta_3 \gg \delta_4 \gg \delta_1$. Since $[1/2(s-2)(s-1)]\cdot(1-1/(s-1))^2 > 1/100s^3$, this proves (19). This means the quadratic coefficient of the lower bound of $|Q|^2/2(s-2)-|Q|/2$ from (18) is larger than the quadratic coefficient of the upper bound for the number of covered pairs in Q from (17). It follows that for $n \ge n_0(s)$, the number of covered pairs in Q is less than $|Q|^2/2(s-2)-|Q|/2$. Hence we have a contradiction, so it is not possible that $(1-\delta)\sqrt{n} \le a_1 < (1+\delta_2)\sqrt{n}$.

Case 3: $10\sqrt{sn} \le a_1 \le \left(\frac{1}{s-1}-\varepsilon\right)n$, where $\varepsilon = 1/10s^2$.

By Lemma 1 part 5, we can assume $d \in \{q, q+1\}$. Suppose $d = 1$. Then, $|Q| = n - a_1 \ge (1-1/(s-1)+\varepsilon)n$, so by Lemma 1 part 2 for $n \ge n_0(s)$

$$m > \frac{|Q|-1}{s-2}+s-2 \ge \frac{n}{s-1}+\frac{\varepsilon n}{s-2}+s-2-\frac{1}{s-2} > \frac{n-1}{s-1}+s-1.$$

Hence, we can assume $d \ge 2$. If $|Q| > n\,(1-1/(s-1))$, then for $n \ge n_0(s)$

$$m \ge \frac{|Q|-1}{s-2}+s-2+d \ge \frac{n}{s-1}+s-2-\frac{1}{s-2}+d > \frac{n-1}{s-1}+s-1. \qquad (20)$$

by Lemma 1 part 2. Hence, we can assume $p \ge n/(s-1)$. We consider the cases $d \ge s+2$ and $d \le s+1$ separately. Suppose $d \ge s+2$. Observe that

$$d \le q+1 \le \frac{n-1}{a_1(s-1)}+\frac{s-2}{a_1}+1$$

and

$$\frac{n}{s-1} \le p \le a_1 d \le \frac{n-1}{s-1}+s-2+a_1.$$

Suppose there are three sets in the neighborhood of v with size less than $a_1/2$. Then,

$$p \leq a_1(d-3) + \frac{3a_1}{2} \leq \frac{n-1}{s-1} + s - 2 - \frac{a_1}{2}.$$

Since $a_1 \geq 10\sqrt{sn}$, this upper bound for p is smaller than $n/(s-1)$, so we have a contradiction. Hence there are at most two sets in the neighborhood of v with size less that $a_1/2$. As we are assuming $d \geq s+2$, there are at least s sets with size more than $a_1/2 - 1 \geq 4\sqrt{sn}$. Taking disjoint subsets of these s sets of size $4\sqrt{sn}$ and applying Lemma 1 part 3 we get

$$m \geq \frac{(4\sqrt{sn})^s}{\binom{s}{2}(4\sqrt{sn})^{s-2}} = \frac{16sn}{\binom{s}{2}} = \frac{32n}{s-1} > \frac{n-1}{s-1} + s - 1$$

for $n \geq n_0(s)$.
We now consider the case in which $d \leq s+1$. Since $n/(s-1) \leq p \leq a_1 d$, we have

$$a_1 \geq \frac{n}{d(s-1)} \geq \frac{n}{s^2-1}.$$

Let $A_1 = \{x_1, \ldots, x_{a_1}\}$. If there are at least $2n/s^3$ points in A_1 with degree at least s^2+1, then $m > 2n/s > (n-1)/(s-1) + s - 1$, so we can assume that there are at most $2n/s^3$ points in A_1 with degree at least s^2+1. This implies that there are at least $n/(s^2-1) - 2n/s^3$ points in A_1 with degree at most s^2. For $1 \leq i \leq a_1$, let

$$T_i = \sum_{j:\, x_i \in A_j, j \neq 1} \binom{a_j - 1}{2}$$

and let $B = \{i : d(x_i) \leq s^2\}$. Then we have

$$\sum_{i \in B} T_i < n^2$$

as each pair is in at most one A_i. It follows that there is some $\ell \in B$ such that

$$T_\ell \leq \frac{n^2}{|B|} \leq \frac{n^2}{\frac{n}{s^2-1} - \frac{2n}{s^3}} = \frac{s^3(s^2-1)}{s^3 - 2s^2 + 2} n \leq 4s^2 n.$$

By Jensen's inequality and the inequality $\binom{x}{2} \geq x^2/16$,

$$T_\ell \geq (d(x_\ell) - 1)\binom{\frac{1}{d(x_\ell)-1}\sum_{k:\, x_\ell \in A_k, k \neq 1}(a_k - 1)}{2} \geq \frac{1}{16(d(x_\ell)-1)}\left(\sum_{k:\, x_\ell \in A_k, k \neq 1}(a_k - 1)\right)^2.$$

Comparing the lower bound and upper bound for T_ℓ yields

$$\sum_{k:\, x_\ell \in A_k, k \neq 1} (a_k - 1) \leq 4\sqrt{d(x_\ell) - 1} \cdot 2s\sqrt{n} \leq 8s^2\sqrt{n}.$$

Let Q_1 be the set of points outside of the neighborhood of x_ℓ. Then every $(s-1)$-set in $[n] \setminus Q_1$ is covered. Furthermore, since $a_1 \leq (1/(s-1) - \varepsilon)n$ and $\varepsilon > \delta_2$

$$\begin{aligned}|Q_1| \geq n - a_1 - 8s^2\sqrt{n} &\geq \left(1 - \frac{1}{s-1} + \varepsilon - \frac{8s^2}{\sqrt{n}}\right) n \\ &\geq \left(1 - \frac{1}{s-1} + \varepsilon - \delta_2\right) n \\ &> \left(1 - \frac{1}{s-1}\right) n,\end{aligned}$$

so by Lemma 1 part 2 with Q replaced with Q_1 we get $m > (n-1)/(s-1) + s - 1$ by the same computation as (20).

Case 4: $(1/(s-1) - \varepsilon)\, n < a_1 < \lfloor (n-1)/(s-1) \rfloor$.

Suppose $d = 1$. Then $|Q| = n - a_1 > (s-2)n/(s-1)$ as $a_1 \leq (n-1)/(s-1) - 1$, so by Lemma 1 part 2 we have

$$m \geq \frac{n - a_1 - 1}{s-2} + s - 2 + 1 > \frac{n-1}{s-1} + s - 1.$$

We can assume $d = 2$ as if $d \geq 3$, then $m > 2a_1 > (2/(s-1) - 2\varepsilon)n > (n-1)/(s-1) + s - 1$. Furthermore, we can assume the number of points in A_1 with degree greater than two is at most $(\varepsilon + 1/s^3)n$ as if not the number of sets that intersect A_1 is at least

$$\left(\frac{1}{s-1} - \varepsilon\right) n + \left(\varepsilon + \frac{1}{s^3}\right) n = \left(\frac{1}{s-1} + \frac{1}{s^3}\right) n > \frac{n-1}{s-1} + s - 1.$$

Suppose all the $(s-1)$-sets in $[n] \setminus A_1$ are covered. Observe that

$$|[n] \setminus A_1| = n - a_1 > n - \frac{n-1}{s-1} > \frac{s-2}{s-1} n > n_0(s-1),$$

and $a_1 \leq (n - a_1 - 1)/(s-2)$ by the same inequalities used in the proof of Lemma 1 part 2 to show

$$|Q| > \frac{n(s-2)}{s-1} \Rightarrow \frac{|Q| - 1}{s-2} \geq \frac{n-1}{s-1}.$$

Define

$$\mathcal{L}' := \{A \cap ([n] \setminus A_1) : A \in \mathcal{L},\ |A \cap ([n] \setminus A_1)| \geq 2\}.$$

By induction on s

$$|\mathcal{L}'| \geq \frac{n - a_1 - 1}{s-2} + s - 2 > \frac{n-1}{s-1} + s - 2.$$

Since $A_1 \cap ([n] \setminus A_1) = \emptyset$, we have

$$m > \frac{n-1}{s-1} + s - 1.$$

Hence we can assume there is some $(s-1)$-set $x_1, x_2, \ldots, x_{s-1}$ in $[n] \setminus A_1$ that is not covered. For any $x_0 \in A_1$ the s-set $\{x_0, x_1, x_2, \ldots, x_{s-1}\}$ is covered, so there is a set containing a pair x_0, x_i for some $i \in [s-1]$. Set

$$B_{x_i} = \{x_0 \in A_1 : x_0, x_i \in A_j \text{ for some } j\}$$

for $1 \leq i \leq s-1$. Without loss of generality, assume $|B_{x_1}| \geq |B_{x_2}| \geq \ldots \geq |B_{x_{s-1}}|$. Then

$$|B_{x_1}| \geq \frac{a_1}{s-1} > \left(\frac{1}{(s-1)^2} - \frac{\varepsilon}{s-1}\right) n.$$

Let $B'_{x_1} \subset B_{x_1}$ be the points in B_{x_1} that have degree two. Since the number of points in A_1 with degree greater than two is at most $(\varepsilon + 1/s^3)n$, we have

$$|B'_{x_1}| \geq \left(\frac{1}{(s-1)^2} - \frac{1}{s^3} - \varepsilon\left(1 + \frac{1}{s-1}\right)\right) n. \tag{21}$$

Let $Q_2 = [n] \setminus (A_1 \cup \{x_1\})$ (see Figure 3). Suppose $\{u_1, \ldots, u_{s-1}\} \subset Q_2$ is uncovered and let $u_0 \in B'_{x_1}$. Then the s-set $\{u_0, u_1, u_2, \ldots, u_{s-1}\}$ must be covered, so there is a set A_i containing u_0, u_i for some $i \in [s-1]$. Since $|B'_{x_1}| > s-1$ by (21), there are $p_1, p_2 \in B'_{x_1}$ and $1 \leq j \leq s-1$ so that the pairs $p_1 u_j$ and $p_2 u_j$ are both covered. The sets containing these pairs are distinct as $p_1, p_2 \in A_1$. This is a contradiction as any set containing a point from B_{x_1} contains x_1, so it implies that the pair $u_j x_1$ is in two distinct sets. Hence, all $(s-1)$-sets in Q_2 are covered. Observe that since $a_1 \leq (n-1)/(s-1)$,

$$|Q_2| = n - a_1 - 1 \geq n - \frac{n-1}{s-1} - 1 = \frac{s-2}{s-1} n + \frac{1}{s-1} - 1 \geq n_0(s-1). \tag{22}$$

Consider the collection of sets $\{A_i \cap Q_2\}_{i=1}^m$. Since $a_1 \leq (n-1)/(s-1) - 1 \leq (n-2)/(s-1)$,

$$a_1 \leq \frac{n - a_1 - 2}{s-2} = \frac{|Q_2| - 1}{s-2},$$

so $|A_i \cap Q_2| \leq (|Q_2| - 1)/(s-2)$ for all i. Suppose $a_1 < (n-1)/(s-1) - 1$. Then, by induction on s, the number of sets in $\mathcal{L}$ containing at least 2 points in Q_2 is at least

$$\frac{|Q_2| - 1}{s-2} + s - 2 = \frac{n - a_1 - 2}{s-2} + s - 2 > \frac{n-1}{s-1} + s - 2, \tag{23}$$

Since $A_1 \cap Q_2 = \emptyset$, $m > (n-1)/(s-1) + s - 1$.

We now consider the case $a_1 = (n-1)/(s-1) - 1$. Note that $(s-1) \mid (n-1)$ in this case. By the inequality in (23), we have $m \geq (n-1)/(s-1) + s - 1$. As stipulated by the induction statement, we are required to show that this inequality is strict.

Claim 3. *If $a_1 \leq (n-1)/(s-1) - 1$, then $m > (n-1)/(s-1) + s - 1$.*

Proof. If $a_1 < (1/(s-1) - \varepsilon)n$, then the claim follows by the arguments in Cases 1-3. If $(1/(s-1) - \varepsilon)n \leq a_1 < (n-1)/(s-1) - 1$, then (23) shows $m > (n-1)/(s-1) + s - 1$, so we can assume $a_1 = (n-1)/(s-1) - 1$. Define

$$\mathcal{L}_2 := \{A \cap Q_2 : A \in \mathcal{L},\ |A \cap Q_2| \geq 2\}.$$

We first consider the case $s = 3$. Then $a_1 = (n-1)/2 - 1$, $|Q_2| = n - (n-1)/(s-2) + 1 - 1 = (n-1)/2 + 1$, and every pair in Q_2 is covered. By Theorem 1, $|\mathcal{L}_2| \geq (n-1)/2 + 1$. If

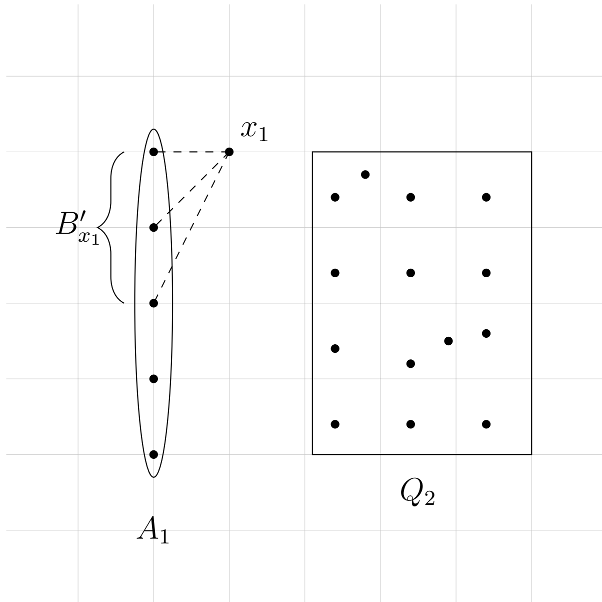

Figure 3: Q_2 and B'_{x_1}

$|\mathcal{L}_2| > (n-1)/2 + 1$, then $m > (n-1)/2 + 2$ as $A_1 \cap Q_2 = \emptyset$. Hence we can assume $|\mathcal{L}_2| = (n-1)/2 + 1$. By Theorem 1, $\mathcal{L}_2$ is either a near pencil or projective plane. If $\mathcal{L}_2$ is a near pencil then, there is a set of size $(n-1)/2$ in $\mathcal{L}_2$. This is a contradiction as the largest set in $\mathcal{L}$ has size $a_1 = (n-1)/2 - 1$. Suppose now that $\mathcal{L}_2$ is a projective plane, so $|\mathcal{L}_2| = |Q_2| = (n-1)/2$. Recall that x_1 has degree at least $|B_{x_1}| \geq a_1/(s-1) = a_1/2 \geq 3$. If $N(x_1)$ contains A_i, A_j such that neither $A_i \cap Q_2$ nor $A_j \cap Q_2$ is in $\mathcal{L}_2$, then $m \geq |\mathcal{L}_2| + 3 > (n-1)/2 + 2$. So $N(x_1)$ contains A_i, A_j such that both $A_i \cap Q_2$ and $A_j \cap Q_2$ are in $\mathcal{L}_2$. But $\mathcal{L}_2$ is a projective plane hence $(A_i \cap Q_2) \cap (A_j \cap Q_2) \neq \emptyset$, which means that $|A_i \cap A_j| \geq 2$, a contradiction.

Now, suppose $s \geq 4$. Suppose $a_1 = (n-1)/(s-1) - 1$. Then every $(s-1)$-set in Q_2 is covered. We also have

$$\frac{|Q_2| - 1}{s-2} = \frac{n - a_1 - 2}{s-2} = \frac{n-1}{s-1},$$

so $a_1 \leq (|Q_2| - 1)/(s-2) - 1$. This shows $|A| \leq (|Q_2| - 1)/(s-2) - 1$ for all $A \in \mathcal{L}_2$. Since $|Q_2| \geq n_0(s-1)$ by (22), the inductive hypothesis implies

$$|\mathcal{L}_2| > \frac{|Q_2| - 1}{s-2} + s - 2 = \frac{n-1}{s-1} + s - 2$$

As $A_1 \notin \mathcal{L}_2$, it follows that $m > (n-1)/(s-1) + s - 1$ and this concludes the proof of the claim and Case 4. □

Case 5: $a_1 = \lfloor (n-1)/(s-1) \rfloor$.

We note that for $x > 1$, we have $\lfloor x \rfloor > x - 1$ so in this case $a_1 > (n-1)/(s-1) - 1$ and we only need to prove that $m \geq (n-1)/(s-1) + s - 1$.

Suppose all the $(s-1)$-sets in $[n]\setminus A_1$ are covered. Observe that $n-a_1 \geq n_0(s-1)$ by (22) and $a_1 \leq (n-a_1-1)/(s-2)$ as $a_1 \leq (n-1)/(s-1)$. Define

$$\mathcal{L}' := \{A \cap ([n]\setminus A_1) : A \in \mathcal{L},\ |A \cap ([n]\setminus A_1)| \geq 2\}.$$

By induction on s,

$$|\mathcal{L}'| \geq \frac{n-a_1-1}{s-2} + s - 2 \geq \frac{n-1}{s-1} + s - 2.$$

Since $A_1 \cap ([n]\setminus A_1) = \emptyset$, we have $m \geq (n-1)/(s-1) + s - 1$.

Now, we consider the case in which there is an $(s-1)$-set $\{x_1, \ldots, x_{s-1}\}$ in $[n]\setminus A_1$ that is uncovered. Note that we can assume $d := \min_{w \in A_1} d(w) = 2$ in this case as if $d = 1$, then all the $(s-1)$-sets in $[n]\setminus A_1$ are covered and if $d \geq 3$, then $m \geq 2a_1 > (n-1)/(s-1) + s - 1$. Furthermore, we may assume that the number of points with degree at least 3 in A_1 is at most $s-1$, as if there are at least s points with degree at least 3, we have $m \geq 1 + a_1 + s \geq (n-1)/(s-1) + s - 1$. For any $x_0 \in A_1$ the s-set $\{x_0, x_1, x_2, \ldots, x_{s-1}\}$ is covered, so $\{x_0, x_i\}$ is covered for some $i \in [s-1]$. For $1 \leq i \leq s-1$, set

$$B_{x_i} = \{y \in A_1 : \{y, x_i\} \subset A_j \text{ for some } j\}.$$

Without loss of generality, assume $|B_{x_1}| \geq |B_{x_2}| \geq \cdots \geq |B_{x_{s-1}}|$. Then

$$|B_{x_1}| \geq \frac{a_1}{s-1}.$$

Let $B'_{x_1} \subset B_{x_1}$ be the set of points in B_{x_1} that have degree two. Since the number of points in A_1 with degree greater than two is at most $s-1$, we have

$$|B'_{x_1}| \geq \frac{a_1}{s-1} - s + 1. \tag{24}$$

Let $Q_2 = [n] \setminus (A_1 \cup \{x_1\})$. Suppose $\{u_1, \ldots, u_{s-1}\} \subset Q_2$ is uncovered and let $u_0 \in B'_{x_1}$. Then the s-set $\{u_0, u_1, u_2, \ldots, u_{s-1}\}$ must be covered, so there a set A_{i_j} containing u_0 and u_j for some $j \in [s-1]$. Since $|B'_{x_1}| > s-1$ by (24), there are $p_1, p_2 \in B'_{x_1}$ and $1 \leq j \leq s-1$ so that the pairs $p_1 u_j$ and $p_2 u_j$ are both covered. The sets containing these pairs are distinct as $p_1, p_2 \in A_1$. This is a contradiction as any set containing a point from B_{x_1} contains x_1, so it implies that the pair u_j, x_1 is in two distinct sets.

$$\text{Hence all } (s-1)\text{-sets in } Q_2 \text{ are covered.} \tag{25}$$

Define

$$\mathcal{L}_2 := \{A \cap Q_2 : A \in \mathcal{L},\ |A \cap Q_2| \geq 2\}.$$

We first consider the case $s = 3$. Suppose first that n is odd, say $n = 2k+1$ for some integer k. Then, $a_1 = k$ and $|Q_2| = k$. Then either $Q_2 \in \mathcal{L}$ or by Theorem 1 $|\mathcal{L}_2| \geq k$. Suppose $Q_2 \in \mathcal{L}$. Since A_1 has minimum degree 2, we have $m \geq k+2 = (n-1)/2 + 2$ as required. Now suppose $Q_2 \notin \mathcal{L}$. By Theorem 1, $|\mathcal{L}_2| \geq k$. If $|\mathcal{L}_2| \geq k+1$, we have $m \geq k+2$ as $A_1 \cap Q_2 = \emptyset$. If $|\mathcal{L}_2| = k$, then $\mathcal{L}_2$ is either a near pencil or projective plane. In either case, any pair of sets in Q_2 intersects in exactly one point. Suppose $m = k+1$ for the sake of contradiction. Then since x_1 has degree at least $a_1/(s-1) = a_1/2$, there must be a collection of pairwise disjoint sets of size $a_1/2$ in Q_2 (i.e. $a_1/2$ pairwise disjoint sets of the form $A_i \cap Q_2$). This is a contradiction, so we have $m \geq k+2$.

Now, suppose $n = 2k$. Then, $a_1 = k-1$ and $|Q_2| = k$. Since $|Q_2| > a_1$, by Theorem 1 $|\mathcal{L}_2| \geq k$. This implies $m \geq k+1$. By the same argument as above using the fact that x_1 has degree at least $a_1/2$, we have $m \neq k+1$, so $m \geq k+2$.

We now consider $s \geq 4$. Write $n-1 = (s-1)\ell + r$ for integers ℓ, r with $0 \leq r < s-1$ so that $a_1 = \ell$. We will show that $m \geq \ell + s \geq (n-1)/(s-1) + s - 1$.

Case 5.1: $r \geq 2$.

Note that $|Q_2| = n - a_1 - 1 = (s-2)\ell + r$ and

$$\frac{|Q_2|-1}{s-2} = \ell + \frac{r-1}{s-2} \geq \ell.$$

For $i > 1$, $|A_i \cap Q_2| \leq a_1 = \ell \leq (|Q_2|-1)/(s-2)$. Since $|Q_2| \geq n_0(s-1)$, by induction on s,

$$|\mathcal{L}_2| \geq \frac{|Q_2|-1}{s-2} + s - 2 = \ell + \frac{r-1}{s-2} + s - 2.$$

Since $A_1 \cap Q_2 = \emptyset$,

$$m \geq \ell + \frac{r-1}{s-2} + s - 1.$$

This shows that $m \geq \ell + s$ as $r \geq 2$.

Case 5.2: $r \in \{0, 1\}$.

Case 5.2.1: There is no $B_1 \in \mathcal{L}$ of size ℓ such that $B_1 \subset Q_2$.

Suppose $r = 1$. Note that $|Q_2| = (s-2)\ell + 1$ and

$$\frac{|Q_2|-1}{s-2} = \ell.$$

We also have $|A_i \cap Q_2| \leq \ell - 1$ for all i. Since $|Q_2| \geq n_0(s-1)$, by induction we have the strict inequality

$$|\mathcal{L}_2| > \frac{|Q_2|-1}{s-2} + s - 2 = \ell + s - 2.$$

Since $A_1 \cap Q_2 = \emptyset$, this implies $m \geq \ell + s$.

Suppose $r = 0$. Note that $\ell = (n-1)/(s-1)$, $|Q_2| = (s-2)\ell$ and

$$\left\lfloor \frac{|Q_2|-1}{s-2} \right\rfloor = \ell - 1.$$

We also have $|A_i \cap Q_2| \leq \ell - 1$ for all i and $|Q_2| \geq n_0(s-1)$. By induction on s,

$$|\mathcal{L}_2| \geq \frac{|Q_2|-1}{s-2} + s - 2 = \ell - \frac{1}{s-2} + s - 2.$$

Since $A_1 \cap Q_2 = \emptyset$, this implies

$$m \geq \ell - \frac{1}{s-2} + s - 1.$$

Since $s \geq 4$, this shows $m \geq \ell + s - 1 = (n-1)/(s-1) + s - 1$.

Case 5.2.2: There exists $B_1 \in \mathcal{L}$ of size ℓ such that $B_1 \subset Q_2$.

We use an iterative argument for this case. Recall (25) that all $(s-1)$-sets in Q_2 are covered. Let $2 \leq j \leq s-2$. Suppose all $(s-j+1)$-sets in Q_j are covered and $B_{j-1} \subset Q_j$ is an ℓ-set in $\mathcal{L}$. Let $Q_{j+1} = Q_j \setminus B_{j-1}$ so that $A_1, B_1, \ldots, B_{j-1}$ are pairwise disjoint (see Figure 4). We can assume the minimum degree among points in B_{j-1} is 2 as we did with A_1. Similarly we can also assume that the number of points in B_{j-1} with degree at least 3 is at most $s-1$. We can also assume the maximum degree of points in B_{j-1} is at most s, as otherwise $m \geq 1 + \ell + (s+1-2) = \ell + s$. Recall that B'_{x_1} is the set of $v \in A_1$ such that $\{v, x_1\}$ is covered and $d(v) = 2$. We first construct a matching (i.e. collection of pairwise disjoint pairs) $M_{j-1} \subset B'_{x_1} \times B_{j-1}$ of covered pairs (see Figure 4). For every $v \in B'_{x_1}$, let $X_v \in \mathcal{L}$ be the set containing x_1 and v. We will assume at most $s-1$ of the sets $\{X_v\}_{v \in B'_{x_1}}$ are disjoint from B_{j-1} as otherwise we have $m \geq \ell + s$ since points in B_{j-1} have minimum degree 2. This means that there are at least $|B'_{x_1}| - s + 1$ covered pairs in $B'_{x_1} \times B_{j-1}$. Since the maximum degree of points in B_{j-1} is at most s, there is a matching $M \subset B'_{x_1} \times B_{j-1}$ of size at least $(|B'_{x_1}| - s + 1)/s$. Let M_{j-1} be the matching formed by deleting pairs (g, h) with $d(h) \geq 3$ from M. Note that (24) implies

$$|M_{j-1}| \geq \frac{|B'_{x_1}| - s + 1}{s} - s + 1 > s - j.$$

We now claim that the $(s-j)$-sets in Q_{j+1} are covered. Suppose $\{h_1, \ldots, h_{s-j}\}$ is an uncovered set in Q_{j+1}. Since every $(s-j+1)$-set in Q_j is covered, for every $h_0 \in B_{j-1}$, the $(s-1)$-set $h_0, h_1, \ldots, h_{s-j}$ is covered. It follows that the pair h_0, h_i is covered for some $i \in [s-j]$. Since $|M_{j-1}| > s-j$, there is h_k and points x, y in the restriction of M_{j-1} to B_{j-1} so that the pairs x, h_k and y, h_k are both covered. This is a contradiction as $d(x) = d(y) = 2$, so sets in $\mathcal{L}$ that contain x, h_k and y, h_k both contain x_1. Hence the pair h_k, x_1 is in two sets in $\mathcal{L}$. Consequently, all $(s-j)$-sets in Q_{j+1} are covered. Define

$$\mathcal{L}_{j+1} := \{A \cap Q_{j+1} : A \in \mathcal{L},\, |A \cap Q_{j+1}| \geq 2\}.$$

Let us first suppose that $r = 1$. Assume $j < s-2$. If $|A_i \cap Q_{j+1}| \leq \ell - 1$ for all i, then $|A_i \cap Q_{j+1}| < \ell = (|Q_{j+1}| - 1)/(s-j-1)$. Since

$$|Q_{j+1}| = (s-j-1)\ell + 1 = (s-j-1)\frac{n-2}{s-1} \geq n_0(s-j),$$

and by induction,

$$|\mathcal{L}_{j+1}| > \frac{|Q_{j+1}| - 1}{s-j-1} + s - j - 1 = \ell + s - j - 1.$$

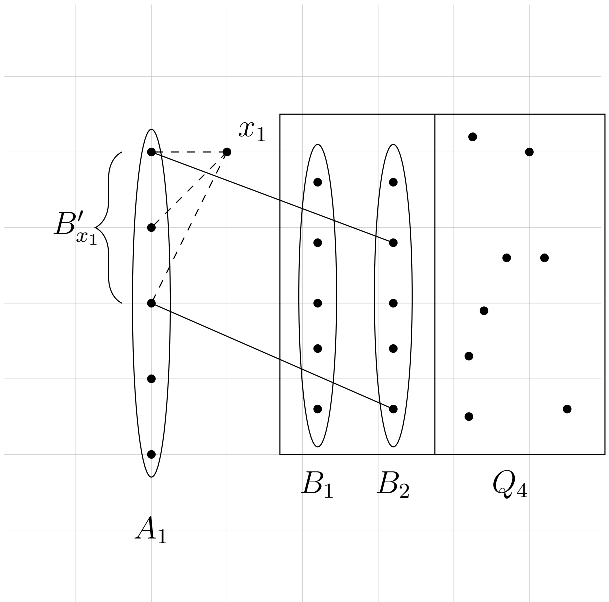

Figure 4: Iterative Procedure with edges in matching shown by solid lines

Since $A_1, B_1, \ldots, B_{j-1}$ are disjoint from Q_{j+1}, we get $m \geq \ell + s$. Otherwise, there is a set $B_j \subset Q_{j+1}$ with size ℓ in $\mathcal{L}$ and we continue the procedure. Now, suppose the procedure terminates at $j = s-2$. Then, we have ℓ-sets $A_1, B_1, \ldots, B_{s-3}$ and all the pairs in $Q_{s-1} = [n] \backslash (A_1 \cup \{x_1\} \cup B_1 \cup \ldots \cup B_{s-3})$ are covered. Since $|Q_{s-1}| = \ell+1$, by Theorem 1, $|\mathcal{L}_{s-1}| \geq \ell+1$. Since $A_1, B_1, \ldots, B_{s-3}$ are disjoint from Q_{s-1}, we have $m \geq \ell + s - 1$. We now show that $m \geq \ell + s$. Suppose $m = \ell + s - 1$ for the sake of contradiction. Then, $\mathcal{L}_{s-1}$ is either a near pencil or projective plane. Since x_1 has degree at least $\ell/(s-1)$ and $m = \ell + s - 1$, there must be a collection of at least $\ell/(s-1)$ sets in $\mathcal{L}_2$. Since $B_1, \ldots, B_{s-3}$ have size ℓ, they cannot contain x_1. It follows that there must be a collection of at least $\ell/(s-1)$ pairwise disjoint sets in $\mathcal{L}_{s-1}$. This is a contradiction as any 2 sets in $\mathcal{L}_{s-1}$ intersect in exactly one point. This shows $m \geq \ell + s$.

Let us now suppose that $r = 0$. Assume $j < s-2$. If $|A_i \cap Q_{j+1}| \leq \ell - 1$ for all i, then $|A_i \cap Q_{j+1}| \leq \lfloor (|Q_{j+1} - 1)/(s-j-1) \rfloor$ due to

$$\frac{|Q_{j+1}| - 1}{s-j-1} = \ell - \frac{1}{s-j-1}.$$

Since

$$|Q_{j+1}| = (s-j-1)\ell = (s-j-1) \cdot \frac{n-1}{s-1} \geq n_0(s-j),$$

we may apply induction to obtain

$$|\mathcal{L}_{j+1}| \geq \frac{|Q_{j+1}| - 1}{s-j-1} + s - j - 1 = \ell - \frac{1}{s-j-1} + s - j - 1.$$

Since $A_1, B_1, \ldots, B_{j-1}$ are disjoint from Q_{j+1}, we have $m \geq \ell + s - 1 = (n-1)/(s-1) + s - 1$ as $s - j - 1 > 1$. Otherwise, there is a set $B_j \subset Q_{j+1}$ with size ℓ in $\mathcal{L}$ and we continue the procedure.

Now, suppose the procedure goes up to $j = s-2$. Then we have ℓ-sets $A_1, B_1, \ldots, B_{s-3}$ and all the pairs in $Q_{s-1} = [n] \setminus (A_1 \cup \{x_1\} \cup B_1 \cup \ldots \cup B_{s-3})$ are covered. Since $|Q_{s-1}| = \ell$, either $Q_{s-1} \in \mathcal{L}$ or by Theorem 1, $|\mathcal{L}_{s-1}| \geq \ell$. If $Q_{s-1} \in \mathcal{L}$, then $m \geq \ell + s - 1$ as we have $s-1$ sets of size ℓ and an additional ℓ sets from the fact that A_1 has minimum degree 2. If $Q_{s-1} \notin \mathcal{L}$, then $|\mathcal{L}_{s-1}| \geq \ell$, so we have $m \geq \ell + s - 2$. Suppose $m = \ell + s - 2$. Since x_1 has degree at least $\ell/(s-1)$ and $m = \ell + s - 2$, there must be a collection of at least $\ell/(s-1)$ pairwise disjoint sets in $\mathcal{L}_2$. Since $B_1, \ldots, B_{s-3}$ have size ℓ, they cannot contain x_1. It follows that there must be a collection of at least $\ell/(s-1)$ pairwise disjoint sets in $\mathcal{L}_{s-1}$. This is a contradiction as any two sets in $\mathcal{L}_{s-1}$ intersect in exactly one point. This shows $m \geq \ell + s - 1 = (n-1)/(s-1) + s - 1$. This concludes the proof of the lower bound in Theorem 3. □

3 Construction for Theorem 3.

Suppose $(s-1) \,|\, (n-1)$. We construct a family of s covers $\mathcal{L}_n(s)$ on $[n]$ with size $(n-1)/(s-1) + s - 1$. Recall that a near pencil on $[n]$ comprises n sets $A, B_1, \ldots, B_{n-1}$ where $A = [n-1]$ and $B_i = \{i, n\}$. By Theorem 1, $\mathcal{L}_n(2)$ consists of the near pencil or a finite projective plane. For $s \geq 3$, and $t = (n-1)/(s-1)$, let $\mathcal{L}_n(s)$ consists of families obtained by taking the disjoint union of some $\mathcal{L} \in \mathcal{L}_{n-t}(s-1)$ with a t-set T, and possibly enlarging each set in $\mathcal{L}$ by a point in T while ensuring that no two sets in our family have more than one point in common. It clear that members of $\mathcal{L}_n(s)$ are s-covers as any s-set contains either 2 points in T or $s-1$ points in the $\mathcal{L} \in \mathcal{L}_{n-t}(s-1)$. Moreover,

$$|\mathcal{L}| + 1 = \frac{n-t-1}{s-2} + (s-2) + 1 = \frac{n-1}{s-1} + s - 1.$$

This shows Theorem 3 is tight if $(s-1) \,|\, (n-1)$.

We now show Theorem 3 is tight asymptotically. Suppose $(s-1) \nmid (n-1)$ and n is sufficiently large in terms of s. We construct an s-cover of size $n/(s-1) \cdot (1 + o(1))$ as $n \to \infty$. In [2], Baker, Harman, and Pintz showed there is a prime number in the interval $[x, x + x^{0.525}]$ for x sufficiently large. Setting $x = \sqrt{n/(s-1)}$ implies that there is a prime q such that

$$\sqrt{\frac{n}{s-1}} \leq q \leq \sqrt{\frac{n}{s-1}} + \left(\frac{n}{s-1}\right)^{0.2625}. \tag{26}$$

Let $A_1, A_2, \ldots, A_{s-3}$ be pairwise disjoint sets of size $\lfloor (n-1)/(s-1) \rfloor$. Let $x = (s-3)\lfloor (n-1)/(s-1) \rfloor + (q^2 + q + 1)$ and A_{s-2} be a set of size $n - x$ with no points from the previous A_i's. Let $A_{s-1}, \ldots, A_m$ be a projective plane on the remaining $q^2 + q + 1$ points. Let $\mathcal{L} = \{A_1, \ldots, A_m\}$. Note that any s-set has either 2 points in some A_i where $1 \leq i \leq s-2$ or 2 points in the projective plane formed by $A_{s-1}, \ldots, A_m$, so all s-sets are covered. We now show that $|A_{s-2}| = n - x \leq \lfloor (n-1)/(s-1) \rfloor$. This is equivalent to

$$n - (s-2)\left\lfloor \frac{n-1}{s-1} \right\rfloor \leq q^2 + q + 1.$$

This holds since

$$n - (s-2)\left\lfloor \frac{n-1}{s-1} \right\rfloor \le \frac{n-1}{s-1} + s - 1 < \frac{n}{s-1} + \sqrt{\frac{n}{s-1}} + 1 \le q^2 + q + 1.$$

For $s - 1 \le t \le m$,

$$|A_t| = q + 1 \le \sqrt{\frac{n}{s-1}} + \left(\frac{n}{s-1}\right)^{0.2625} + 1 < \frac{n-1}{s-1}.$$

This shows $\mathcal{L}$ is an s-cover of size

$$q^2 + q + 1 + s - 2 = \frac{n}{s-1}\left(1 + o(1)\right)$$

by (26). □

References

[1] Noga Alon, Keith E. Mellinger, Dhruv Mubayi, and Jacques Verstraëte. The de Bruijn-Erdős theorem for hypergraphs. *Des. Codes Cryptogr.*, 65(3):233–245, 2012.

[2] R. C. Baker, G. Harman, and J. Pintz. The difference between consecutive primes. II. *Proc. London Math. Soc. (3)*, 83(3):532–562, 2001.

[3] R. C. Bose. A note on fisher's inequality for balanced incomplete block designs. *Ann. Math. Statistics*, 20:619–620, 1949.

[4] Ehsan Chiniforooshan and Vašek Chvátal. A de bruijn-Erdős theorem and metric spaces. *Discrete Math. Theor. Comput. Sci.*, 13(1):67–74, 2011.

[5] N. G. de Bruijn and P. Erdős. On a combinatorial problem. *Nederl. Akad. Wetensch., Proc.*, 51:1277–1279 = Indagationes Math. 10, 421–423, 1948.

[6] M. Doležal, T. Mitsis, and Ch. Pelekis. The de bruijn–Erdős theorem from a hausdorff measure point of view. *Acta Math. Hungar.*, 159(2):400–413, 2019.

[7] R. A. Fisher. An examination of the different possible solutions of a problem in incomplete blocks. *Ann. Eugenics*, 10:52–75, 1940.

[8] Dijen K. Ray-Chaudhuri and Richard M. Wilson. On t-designs. *Osaka Math. J.*, 12(3):737–744, 1975.

[9] Paul Turán. Eine Extremalaufgabe aus der Graphentheorie. *Mat. Fiz. Lapok*, 48:436–452, 1941.