Shape of the asymptotic maximum sum-free sets in integer lattice grids

Hong Liu* Guanghui Wang† Laurence Wilkes‡ Donglei Yang§

19th August 2022

Abstract

We determine the shape of all sum-free sets in $\{1, 2, \ldots, n\}^2$ of size close to the maximum $\frac{3}{5}n^2$, solving a problem of Elsholtz and Rackham. We show that all such asymptotic maximum sum-free sets lie completely in the stripe $\frac{4}{5}n - o(n) \leq x + y \leq \frac{8}{5}n + o(n)$. We also determine for any positive integer p the maximum size of a subset $A \subseteq \{1, 2, \ldots, n\}^2$ which forbids the triple (x, y, z) satisfying $px + py = z$.

1 Introduction

A cornerstone result of Schur [Sch16] states that for sufficiently large integer n and a fixed integer r, any r-coloring of $[n] := \{1, 2, \ldots, n\}$ yields a monochromatic triple x, y, z such that $x + y = z$. For an integer $n \in \mathbb{N}$ a subset $A \subseteq [n]$ is *sum-free* if it has no solution for the equation $x + y = z$, i.e. for all $x, y \in A$ we have $x + y \notin A$. The topic of sum-free sets of integers is well-studied in combinatorial number theory and has a long history.

It is clear that the sets

$$S_1 = \left\{1, 3, 5, \ldots, 2\left\lfloor \frac{n-1}{2} \right\rfloor + 1\right\} \quad \text{and} \quad S_2 = \left\{\left\lceil \frac{n+1}{2} \right\rceil, \left\lceil \frac{n+1}{2} \right\rceil + 1, \ldots, n\right\}$$

are sum-free and of size $\left\lceil \frac{n}{2} \right\rceil$. If n is even, $S_3 = S_2 - 1$ is another one of the same size. Let us denote the density of a maximum sum-free subset of $[n]$ by $\mu([n]) := \max\{\frac{|S|}{n} \mid S \subseteq [n], S \textit{ is sum-free}\}$. If $S \subseteq [n]$ is a sum-free set and $a \in S$ is the largest element, then at most one of x or $a - x$ can be in S for each $x \leq a$. Therefore $|S| \leq \left\lceil \frac{a}{2} \right\rceil \leq \left\lceil \frac{n}{2} \right\rceil$. Together with the above examples, we see that

$$\mu([n]) = \begin{cases} \frac{1}{2} & \text{if } n \text{ even,} \\ \frac{1}{2} + \frac{1}{2n} & \text{if } n \text{ odd.} \end{cases}$$

*Extremal Combinatorics and Probability Group (ECOPRO), Institute for Basic Science (IBS), Daejeon, South Korea, Email: `hongliu@ibs.re.kr`. H.L. was supported by the Institute for Basic Science (IBS-R029-C4).

†School of Mathematics, Shandong University, China. Email: `ghwang@sdu.edu.cn`. G.W. was supported by National Key R&D Program of China (2020YFA0712400), Natural Science Foundation of China (11871311, 11631014) and seed fund program for international research cooperation of Shandong University.

‡Department of Computer Science, KU Leuven, Belgium. Email: `laurence.wilkes@kuleuven.be`.

§Data Science Institute, Shandong University, China, Email: `dlyang@sdu.edu.cn`. D.Y. was supported by the China Postdoctoral Science Foundation (2021T140413), Natural Science Foundation of China (12101365) and Natural Science Foundation of Shandong Province (ZR2021QA029).

1.1 Structure for large sum-free sets

Given the extremal result, great efforts has been made to better understand the general structure of large sum-free sets in $[n]$. The first result on this topic was due to Freiman [Fre92] who showed that if the size of a sum-free set in $[n]$ is large enough, then it will either consist of all odd numbers as in S_1 above or it will be close to the second half of the interval as S_2. We remark that more structural results are known for large sum-free sets in the 1-dimensional integer lattice (see [DFST99] and a recent progress [Tra18]). Such structural results are not only interesting on their own; they have been utilized e.g. in recent work on enumerating maximal sum-free sets (see [BLST18]).

The problem of sum-free sets has been generalized to higher dimensional lattice $\mathbb{Z}^d$, $d \geq 2$. Similarly, we define $\mu([n]^d) := \max\{\frac{|S|}{n^d} \mid S \subseteq [n]^d \textit{ is sum-free}\}$. In particular, for $d = 2$, the problem of finding the largest sum-free subset of $[n]^2 = \{1, 2, \ldots, n\}^2$ was firstly presented by Cameron as an unsolved problem in [Cam05].

Conjecture 1.1. [Cam05] *There exists a constant $c \in \mathbb{R}$ such that $\mu([n]^2) = c + O(1/n)$.*

Cameron later [Cam02] suggested that Conjecture 1.1 is true with $c = 0.6$ and gave a lower bound construction:

$$S_0 = \{(x, y) \in [n]^2 \mid u \leq x + y \leq 2u - 1\},$$

which has maximum density 0.6 when $u = \lfloor \frac{4n+7}{5} \rfloor$. Recently, Elsholtz and Rackham settled Conjecture 1.1 in [ER17], proving that indeed

$$\mu([n]^2) = 0.6 + O(1/n).$$

In the same paper, Elsholtz and Rackham [ER17] raised the problem of classifying the sum-free sets whose size are close to the extremal value.

In this paper, we resolve this problem by showing that any sum-free subset $S \subseteq [n]^2$ of size at least $(\frac{3}{5} - o(1))n^2$ will have all its points in the region $\{(x, y) \in [n]^2 \mid \frac{4n}{5} - o(n) \leq x + y < \frac{8n}{5} + o(n)\}$.

Theorem 1.2. *For all $\gamma > 0$ there exists $\delta > 0$ and $n_0 \in \mathbb{N}$ such that the following holds for all $n > n_0$. If $S \subseteq [n]^2$ is sum-free with $|S| > (\frac{3}{5} - \delta)n^2$, then*

$$S \subseteq \{(x, y) \in [n]^2 \mid \frac{4n}{5} - \gamma n \leq x + y < \frac{8n}{5} + \gamma n\}.$$

This gives a satisfying answer to the 2-dimension sum-free problem. The situation is, however, unclear for higher dimension. In particular, even the maximum density of a sum-free set in the 3-dimension grid $[n]^3$ is unknown.

1.2 (p, q)-sum-free sets

Given positive integers d, n and rational numbers p, q, a set $S \subseteq [n]^d$ is called (p, q)-*sum-free* if it has no solution for the equation $px + qy = z$. As a generalization of sum-free sets (i.e. (1,1)-sum-free sets), the notion of (p, q)-sum-free sets encapsulates many fundamental topics in combinatorial number theory. In particular, for $d = 1$, a $(\frac{1}{2}, \frac{1}{2})$-sum-free set is precisely a set without 3-term arithmetic progression, which has received considerable attention in recent decades. Therefore, it is a natural question to determine the size of the largest (p, q)-sum-free sets in $[n]^d$. Here one can similarly define

$$\mu_{[p,q]}([n]^d) := \max\left\{\frac{|S|}{n^d} \mid S \subseteq [n]^d \textit{ is } (p, q)\textit{-sum-free}\right\}.$$

By Roth's theorem [Rot53], $\mu_{[1/2,1/2]}([n]) = o(1)$. See [Blo16] for the best known upper bound for the size of a $(1/2, 1/2)$-sum-free set. In [Ruz93, Ruz95], instead of the form $x+y = z$, Ruzsa instigated the study of more general linear equations $a_1x_1 + \cdots + a_kx_k = b$. In particular, for more general cases when p, q are positive integers and $p \geq 2$, Hancock and Treglown [HT17] completely determined the value $\mu_{[p,q]}([n])$. For higher dimensional lattices, Choi, Kim and Park [CKP20] initiated the investigation of the form $x_1 + x_2 + \cdots + x_k = b$, where b is a prescribed point in $[n]^2$.

For 2-dimension (p, q)-sum-free problem, we make the first attempt to determine $\mu_{[p,p]}([n]^2)$ for any integer p.

Theorem 1.3. *Let $p \in \mathbb{N}$ and $S \subseteq [n]^2$ be a (p, p)-sum-free set. Then*

$$|S| \leq \Big(1 - \frac{2}{4p^2 + 1}\Big)n^2 + O(n).$$

We observe that the upper bound in Theorem 1.3 is optimal up to the error term $O(n)$, given by the following construction. For any positive integers p, q and positive real a, define $S = \{(x, y) \in [n]^2 \mid a < x + y < (p + q)a\}$. One can easily check that S is (p, q)-sum-free with size

$$|S| = \left(1 - \frac{2}{(p+q)^2 + 1}\right) n^2 + O(n),$$

when $a = \frac{2(p+q)}{(p+q)^2+1}n$, corresponding to the stripe

$$S = \left\{(x, y) \in [n]^2 \mid \frac{2(p+q)}{(p+q)^2+1}n < x + y < \frac{2(p+q)^2}{(p+q)^2+1}n\right\}.$$

We conjecture that for all integers p and q, the above construction provides the maximum (p, q)-sum-free set.

Conjecture 1.4. *Let p, q and n be positive integers and $S \subseteq [n]^2$ be a (p, q)-sum-free set. Then*

$$|S| \leq \left(1 - \frac{2}{(p+q)^2 + 1}\right) n^2 + O(n).$$

Organization. The rest of the paper will be organized as follows. Section 2 includes some notation and tools needed. Section 3 is devoted to the proof of Theorem 1.2. The proof of Theorem 1.3 is given in Section 4.

2 Preliminaries

Given a convex polygon P in $\mathbb{R}^2_{[0,n]}$, denote by $\Lambda(P)$ the number of lattice points contained within P, and by $\|P\|$ the area of P with respect to the Lebesgue measure. The *translate* of P by a vector $a \in \mathbb{R}^2_{[0,n]}$ is denoted as $P + a := \{a + (x, y) \mid (x, y) \in P\}$. Write $a - P := \{a - (x, y) \mid (x, y) \in P\}$. Throughout the proof, we always use the following result which is a corollary of Lemma 3.1 in [ER17].

Lemma 2.1. *If P is a convex polygon in $\mathbb{R}^2_{[0,n]}$ with finitely many sides, then $\Lambda(P) = \|P\| + O(n)$.*

This lemma implies that any convex polygon P, described above, satisfies that $\Lambda(P) = \|P\| + O(n)$, which allows us to focus on the area $\|P\|$ instead of $\Lambda(P)$.

For two points $p_1, p_2 \in \mathbb{R}^2_{[0,n]}$, denote by $m(p_1, p_2)$ the gradient and by $c(p_2, p_2)$ the y-intercept of the line in $\mathbb{R}^2$ passing through p_1 and p_2.

Definition 2.2 (Upper boundary)**.** Given a set $A \subseteq \mathbb{R}^2_{[0,n]}$, the *upper boundary* of A is a set of points in A, denoted by ∂A, such that for each $p_1 \in \partial A$ there exists a point $p_2 \in A \setminus \{p_1\}$ with the following properties:

- $m(p_1, p_2) < 0$;
- Let $T = \{(x, y) \subseteq \mathbb{R}^2_{[0,n]} \mid y > m(p_1, p_2)x + c(p_1, p_2)\}$. Then $|A \cap T| = 0$.

Any two such points p_1, p_2 are said to be *adjoint*, and the line passing through two points that are adjoint is called an *upper boundary line.* The second condition above states that there is no point of A strictly above any upper boundary line.

The following lemma shows that if the upper boundary of a set A is empty, then A has a 'top right corner'.

Lemma 2.3 (Lemma 5.1 in [ER17])**.** *Suppose $A \subseteq \mathbb{R}^2_{[0,n]}$ such that $\partial A = \emptyset$. Then there is a point $(a, b) \in A$ such that $a \geq x$ and $b \geq y$ for all $(x, y) \in A$.*

We also need the concept of pairing sets, which will be frequently used throughout the proof.

Definition 2.4. Given a point $(a_1, a_2) \in \mathbb{R}^2_{[0,n]}$ and a set $P \subseteq \mathbb{R}^2_{[0,n]}$, we call P a *pairing set for* (a_1, a_2) if for any $x \in P$, we have $(a_1, a_2) - x \in P$.

The following lemma guarantees that for any point in a sum-free set S, every pairing set for that point cannot intersect too much with S.

Lemma 2.5 (Lemma 3.4 in [ER17])**.** *Let S be a sum-free set in $[n]^2$. Then for any $a \in S$ and a pairing set P for a, we have $|P \cap S| \leq \frac{1}{2}\Lambda(P)$.*

The following lemma bounds the intersection of a set and its translate with a sum-free set.

Lemma 2.6. *Given two sets $S, T \subseteq [n]^2$, if S is sum-free, then for any $a \in S$, it holds that*

$$|S \cap (T \cup (a \pm T))| \leq |T|.$$

Proof. For each element $t \in T$ there is a corresponding element $a \pm t \in a \pm T$. Since $a \in S$, one can observe from sum-freeness that at most one of t and $a \pm t$ belongs to S. $\square$

3 Proof of Theorem 1.2

We carry out the proof in a few steps. First, using Lagrange multiplier, we show that any almost maximum-size sum-free set S in $[n]^2$ has an upper boundary line that is close to the line $y + x = \frac{8n}{5}$, see Lemma 3.3. Then we show that there is a point (x^*, y^*) in S close to $(\frac{4n}{5}, \frac{4n}{5})$, see Lemma 3.4. Finally, using this point (x^*, y^*), we show in Section 3.3 that S has no point below the line $y + x = \frac{4n}{5} - o(n)$, which, together with the upper boundary line close to $y + x = \frac{8n}{5}$, implies that S must be close to the extremal stripe $\frac{4n}{5} \leq x + y \leq \frac{8n}{5}$.

Throughout the proofs, when we write $\beta \ll \gamma$, we always mean that β, γ are constants in $(0, 1)$, and there exists $\beta_0 = \beta_0(\gamma)$ such that the subsequent arguments hold for all $0 < \beta \leq \beta_0$. Hierarchies of other lengths are defined analogously.

Definition 3.1. A sum-free set $S \subseteq [n]^2$ with $\partial S \neq \emptyset$ is of *Type* 1 if there exists a point $p_1 = (x_1, y_1) \in \partial S$ with $x_1 \leq y_1$ and $x_1 y_1 \geq xy$ for all $(x, y) \in \partial S$, and a point $p_2 = (x_2, y_2)$ adjoint to p_1 satisfying the following conditions, where we simply write $m = m(p_1, p_2)$ and $c = c(p_1, p_2)$.

(1) $x_2 > x_1$, $y_2 < y_1$ and $m < -\frac{y_1}{x_1} \le -1$;

(2) $c > n$ and $-c \le nm$.

In addition, S is of *Type* 2 if there exist two adjoint points $p_1 = (x_1, y_1)$ and $p_2 = (x_2, y_2)$ in ∂S satisfying the following conditions.

(1) $x_2 > x_1$, $y_2 < y_1$ and $-\frac{y_1}{x_1} \le m \le -\frac{y_2}{x_2}$;

(2) $y_2 \le \frac{c}{2} \le y_1$;

(3) $c > n$ and $-c < nm$.

For either type of the sum-free sets, we call the upper boundary lines passing through p_1 and p_2 *typical*. Let

$$A = \{(x, y) \in \mathbb{R}^2_{[0,n]} \mid y > mx + c\}$$

with m and c given as above. Then A is a triangle in both cases.

For the Type 1 set S, we claim that the upper boundary line $y = mx+c$ satisfies $x_1 > \frac{n}{2}$. In fact, since $m < -\frac{y_1}{x_1}$ and $y_1 = mx_1 + c$, we have that $x_1 > \frac{c}{-2m} > \frac{n}{2}$ because $-c < nm$.

If S is of Type 2, then it is straightforward to check that the following two sets are nonempty (see Figure 1).

$$T_1 = \left\{(x, y) \in \mathbb{R}^2_{[0,n]} \mid x \ge x_1, y - mx \le \frac{c}{2}\right\},$$

$$T_2 = \left\{(x, y) \in \mathbb{R}^2_{[0,n]} \mid y \ge y_2, y - mx \le \frac{c}{2}\right\}.$$

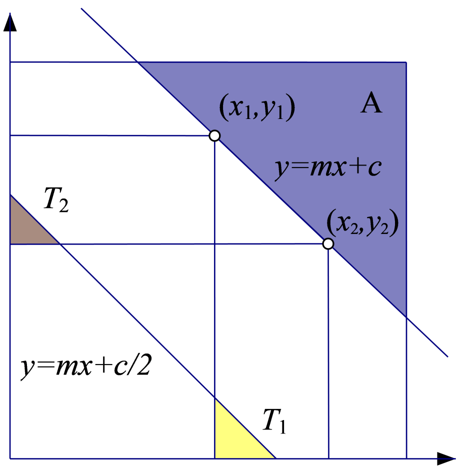

Figure 1: S is of Type 2

The two types we defined above correspond to the only two cases in [ER17] that attain the maximum density $\frac{3}{5}$. We will use the following bounds for these two types of sum-free sets.

Lemma 3.2 ([ER17])**.** *Given a sum-free set $S \subseteq [n]^2$, if $|S| > 0.56n^2$, then either*

(1) *S is of Type 1 and $|S| \le (n+1)^2 - \frac{1}{2}x_1y_1 + \frac{(c+nm-n)^2}{2m}$, or*

(2) *S is of Type 2 and $|S| \le (n+1)^2 + \frac{c^2}{8m} + \frac{1}{2m}(n - nm - c)^2$.*

3.1 Fixing an upper boundary line

Given constants ε and C, we call a line L *ε-close* to the line $x+y=C$ if the portion of L intersecting $\mathbb{R}^2_{[0,n]}$ lies entirely within the set $\{(x,y) \in \mathbb{R}^2_{[0,n]} \mid |x+y-C| \leq \varepsilon n\}$. Similarly, we call two points $p_1 = (x_1, y_1)$ and $p_2 = (x_2, y_2)$ *ε-close* to each other if $|x_1 - x_2| \leq \varepsilon n$ and $|y_1 - y_2| \leq \varepsilon n$.

Lemma 3.3. *Given $\varepsilon > 0$, there exist $\delta > 0$ and $n_0 \in \mathbb{N}$ such that the following holds for all $n > n_0$. If $S \subseteq [n]^2$ is sum-free and $|S| > (\frac{3}{5} - \delta)n^2$, then there is a typical upper boundary line for S which is ε-close to $x + y = \frac{8n}{5}$.*

Proof. Given $\varepsilon > 0$, let $\delta = \frac{\varepsilon^2}{100}$ and n be sufficiently large with respect to ε. Let $S \subseteq [n]^2$ be a sum-free set with $|S| > (\frac{3}{5} - \frac{\varepsilon^2}{100})n^2$.

Suppose for contradiction that any upper boundary line $y = mx + c$ for S is not ε-close to $x + y = \frac{8n}{5}$. That is, either the y- or the x-intercept is far from where it should be:

$$\text{either} \quad |c - 8n/5| > \varepsilon n \quad \text{or} \quad |c/m + 8n/5| > \varepsilon n.$$

In both cases we shall obtain a contradiction by showing that $|S| \leq (3/5 - \varepsilon^2/100)n^2$.

Considering the typical upper boundary line $y = mx + c$ passing through p_1 and p_2 given in Definition 3.1, we will finish the case when the y-intercept is too far, that is, $|c - 8n/5| > \varepsilon n$, whose proof will be divided into two cases depending on the type of S. The case when the x-intercept is too far (that is, $|c/m + 8n/5| > \varepsilon n$) is similar and we omit the details.

Suppose first that S is of Type 1, then by Lemma 3.2(1), we have

$$|S| \leq (n+1)^2 - \frac{1}{2}\left(x_1(mx_1 + c) - \frac{(c + mn - n)^2}{m}\right) =: f(x, m, c).$$

To simplify the presentation, we introduce a new variable η with $\eta \in (-\infty, -\varepsilon) \cup (\varepsilon, +\infty)$ and define

$$f_\eta := \max\{f(x, m, c) \mid c - 8n/5 = \eta n\}.$$

Let $L := f(x, m, c) - \lambda g$, where $g = c - 8n/5 - \eta n$. By solving $\frac{\partial L}{\partial x} = 0$, $\frac{\partial L}{\partial m} = 0$, $\frac{\partial L}{\partial c} = 0$ and $\frac{\partial L}{\partial \lambda} = 0$, we obtain $m = -\sqrt{1 + 2\eta + \frac{5\eta^2}{4}}$ and $x = \frac{\frac{4}{5} + \frac{\eta}{2}}{\sqrt{1 + 2\eta + \frac{5\eta^2}{4}}} n$, and thus the maximum value is

$$f_\eta = \left(8/5 + \eta - \sqrt{1 + 2\eta + 5\eta^2/4}\right) n^2 + O(n).$$

As η takes values over $(-\infty, -\varepsilon) \cup (\varepsilon, +\infty)$, we get

$$f_\eta \leq \left(8/5 + \varepsilon - \sqrt{1 + 2\varepsilon + 5\varepsilon^2/4}\right) n^2 + O(n) \leq (3/5 - \varepsilon^2/100)n^2.$$

For the second case when S is of Type 2, by Lemma 3.2(2), we have:

$$|S| \leq (n+1)^2 + \frac{c^2}{8m} + \frac{(n - nm - c)^2}{2m}.$$

Using Lagrange multiplier again, we arrive at the same bound $\left(8/5 + \varepsilon - \sqrt{1 + 2\varepsilon + 5\varepsilon^2/4}\right) n^2 + O(n) \leq (3/5 - \varepsilon^2/100)n^2$ as desired. □

3.2 Top right corner

Lemma 3.4. *For any $\beta > 0$, there exist $\delta > 0$ and $n_0 \in \mathbb{N}$ such that for all $n > n_0$, if $S \subseteq [n]^2$ is sum-free with $|S| > (\frac{3}{5} - \delta)n^2$, then there is a point $(x^*, y^*) \in S$ which is β-close to the point $(\frac{4n}{5}, \frac{4n}{5})$.*

Proof. We first handle Type 1 sum-free sets. Given $\beta > 0$, we have constants $\delta = \delta_{3.3} > 0$ and $n_0 \in \mathbb{N}$ returned from Lemma 3.3 with $\varepsilon = \beta/6$. Let $S \subseteq [n]^2$ be a sum-free set of Type 1 with $|S| > (\frac{3}{5} - \delta)n^2$. Then Lemma 3.3 gives a typical upper boundary line $y = mx + c$ that is ε-close to $x + y = \frac{8n}{5}$ and let $p_1 = (x_1, y_1)$, $p_2 = (x_2, y_2)$ be the two points involved. Therefore, $|c - \frac{8n}{5}| < \varepsilon n$, $|x_1 + y_1 - \frac{8n}{5}| < \varepsilon n$. Consequently, by triangle inequality we have

$$|m + 1| = \frac{|x_1 + y_1 - c|}{x_1} < \frac{2\varepsilon n}{x_1} < 4\varepsilon,$$

where the last inequality follows since $x_1 > n/2$. Recall that $m \leq -\frac{y_1}{x_1} \leq -1$. Then we have that $|m + \frac{y_1}{x_1}| < 4\varepsilon$.

Using these facts we can write $m = -\frac{y_1}{x_1} - \gamma_1$ and $c = (\frac{8}{5} + \gamma_2)n$ for constants $0 \leq \gamma_1 < 4\varepsilon$ and $|\gamma_2| < \varepsilon$. Using the equation $y_1 = mx_1 + c$, we obtain that $y_1 = \frac{4}{5}n + \frac{\gamma_2 n - \gamma_1 x_1}{2}$. As $x_1 \leq n$, by triangle inequality, we have

$$\left| y_1 - \frac{4n}{5} \right| < \frac{5\varepsilon n}{2} < \beta n.$$

Moreover, since $-\frac{y_1}{x_1} \geq m > -1 - 4\varepsilon$ and $x_1 \leq y_1$, we can easily obtain that $|x_1 - \frac{4n}{5}| < 6\varepsilon n = \beta n$. So (x_1, y_1) is β-close to the point $(\frac{4n}{5}, \frac{4n}{5})$ as desired.

Let us turn to Type 2 sum-free sets. Now, given $\beta > 0$, choose positive constants ε, δ with $\delta \ll \varepsilon \ll \beta$. Let S be a sum-free set of Type 2 with $|S| > (\frac{3}{5} - \delta)n^2$. Then applying Lemma 3.3 with $\sqrt{2}\varepsilon$ playing the role of ε gives a typical upper boundary line $y = mx + c$ passing through $p_1 = (x_1, y_1)$ and $p_2 = (x_2, y_2)$ (see Figure 2), which is $\sqrt{2}\varepsilon$-close to $x + y = \frac{8n}{5}$. This implies that the line $y = mx + \frac{c}{2}$ is $\frac{\varepsilon}{\sqrt{2}}$-close to $x + y = \frac{4n}{5}$. We may assume for contradiction that S has no points in the region

$$T_\beta = \left\{ (x, y) \in \mathbb{R}^2_{[0,n]} \mid x, y \geq \frac{4n}{5} - \beta n\,,\ \ y + x \leq \frac{8n}{5} \right\}.$$

Redefine the regions as follows:

$$A = \left\{ (x, y) \in \mathbb{R}^2_{[0,n]} \mid y + x \geq \frac{8n}{5} + \sqrt{2}\varepsilon n \right\},$$

$$T_1 = \left\{ (x, y) \in \mathbb{R}^2_{[0,n]} \mid y + x \leq \frac{4n}{5} - \frac{\varepsilon n}{\sqrt{2}},\ x \geq x_1 \right\},$$

$$T_2 = \left\{ (x, y) \in \mathbb{R}^2_{[0,n]} \mid y + x \leq \frac{4n}{5} - \frac{\varepsilon n}{\sqrt{2}},\ y \geq y_2 \right\}.$$

Note that

$$T_1 + T_2 = \left\{ (x, y) \in \mathbb{R}^2_{[0,n]} \mid y + x \leq \frac{8n}{5} - \sqrt{2}\varepsilon n,\ x \geq x_1,\ y \geq y_2 \right\}.$$

We now proceed by considering the areas which may be excluded from S. Firstly, we show that S has two points in T_1 that are far apart.

Claim 3.5. *There are two points in $T_1 \cap S$ which are at least βn far apart.*

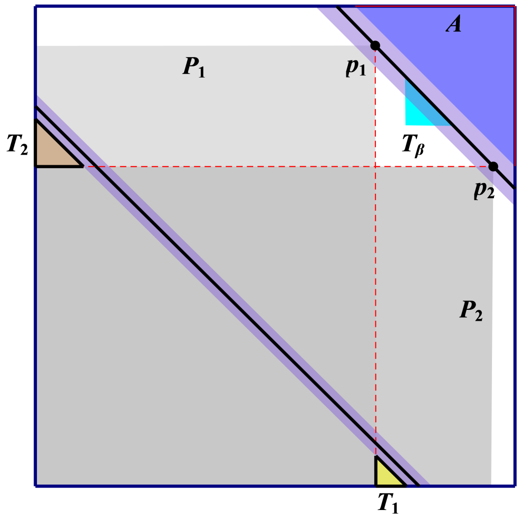

Figure 2: S is of Type 2: the two purple stripes are $\{(x,y) \mid |x+y-\frac{8n}{5}| \le \sqrt{2}\varepsilon n\}$ (on the top right) and $\{(x,y) \mid |x+y-\frac{4n}{5}| \le \frac{\varepsilon n}{\sqrt{2}}\}$.

Proof of claim. If this is not true, then there are less than $\pi(\frac{\beta n}{2})^2 \le \beta^2 n^2$ points of S in T_1, given by the number of points in a square around a circle of diameter βn in T_1. Since $\|T_\beta\| = \frac{1}{2}(2\beta)^2 n^2 = 2\beta^2 n^2$, we then use the pairing set P_1 for (x_1, y_1) and thus

$$\begin{aligned}|S| &\le n^2 - \Lambda(A) - \frac{1}{2}\Lambda(P_1) - \Lambda(T_1) + \beta^2 n^2 - \Lambda(T_\beta)\\ &= n^2 - \|A\| - \frac{1}{2}\|P_1\| - \|T_1\| + \beta^2 n^2 - \|T_\beta\| + O(n)\\ &= n^2 - \frac{1}{2}\left(\frac{2}{5} - \sqrt{2}\varepsilon\right)^2 n^2 - \frac{1}{2}x_1 y_1 - \frac{1}{2}\left(\frac{4}{5} - \frac{\varepsilon}{\sqrt{2}} - \frac{x_1}{n}\right)^2 n^2 + \beta^2 n^2 - 2\beta^2 n^2 + O(n).\end{aligned}$$

It is easy to see this is maximized when y_1 is minimal and $x_1 + y_1 = \frac{8n}{5} - \sqrt{2}\varepsilon n$. Then

$$|S| \le \left(\frac{3}{5} - \beta^2 + 10\varepsilon\right) n^2.$$

Therefore, we reach a contradiction by the fact that $\delta \ll \varepsilon \ll \beta$. ■

By Claim 3.5, we let s and t be two points in T_1 with distance greater than βn, and let

$$T_2^s := s + T_2 \quad \text{and} \quad T_2^t := t + T_2.$$

Claim 3.6. *$\Lambda(T_\beta \setminus T_2^s), \Lambda(T_\beta \setminus T_2^t) < \frac{\beta^2}{L} n^2$, where $L = \frac{4}{5\sqrt{3}-6}$.*

Proof of claim. Suppose to the contrary that either $\Lambda(T_\beta \setminus T_2^s) \ge \frac{\beta^2}{L} n^2$ or $\Lambda(T_\beta \setminus T_2^t) \ge \frac{\beta^2}{L} n^2$, and by symmetry we may assume the first inequality holds. Considering the pairing set P_2 for (x_2, y_2) and T_2 paired with T_2^s, we can obtain from Lemmas 2.5 and 2.6 that

$$\begin{aligned}|S| &\le n^2 - \Lambda(A) - \frac{1}{2}\Lambda(P_2) - \Lambda(T_2) - \Lambda(T_\beta \setminus T_2^s)\\ &\le \left(\frac{3}{5} - \frac{\beta^2}{L} + O(\varepsilon)\right) n^2,\end{aligned}$$

which once again gives a contradiction as $\delta \ll \beta$. ■

In the rest of the proof, we shall find a partition $T_2 = T_{2,1} \cup T_{2,2}$ into two regions such that their corresponding translates $T_{2,1}^s = s + T_{2,1}$ and $T_{2,2}^t = t + T_{2,2}$ are distantly separated in $T_1 + T_2$, which provides a significant portion of points in $T_\beta \setminus (T_{2,1}^s \cup T_{2,2}^t)$ that are to be excluded from S.

Write $s = (x_s, y_s)$ and $t = (x_t, y_t)$. By Claim 3.6, we can find that the two points $s+(0, y_2)$ and $t+(0, y_2)$ belong to the region $\{(x, y) \in [n]^2 \mid x, y \le \frac{4n}{5}\}$. We may assume $x_s+y_s \ge x_t+y_t$ and let $d := x_s + y_s - (x_t + y_t)$. It is easy to see in Figure 3 that d is the difference between the corresponding y-intercepts of the red diagonal and the blue diagonal. By the symmetry of all the shapes involved, we can further assume that $x_s \ge x_t$.

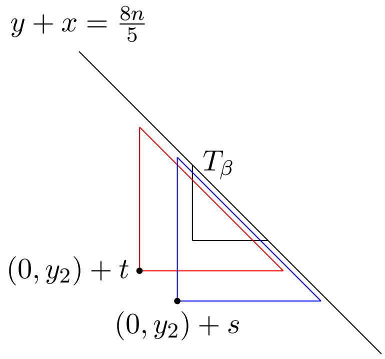

Figure 3: The red triangle represents T_2^t, the blue one represents T_2^s and the black one represents T_β.

Claim 3.7. $d \le \left(2 - 2\sqrt{\frac{2L-1}{2L}}\right)\beta n = \frac{\sqrt{3}-1}{2}\beta n.$

Proof of claim. It is easy to see the region $T_\beta \cap T_2^t$ is a triangle similar to T_β. Note that the area of $T_\beta \setminus T_2^t$ is at least

$$\frac{1}{2}(2\beta n)^2 - \frac{1}{2}(2\beta n - d)^2 = \left(2\beta n - \frac{d}{2}\right)d.$$

By Claim 3.6, we have that $\left(2\beta n - \frac{d}{2}\right) d \le \frac{\beta^2 n^2}{L}$, which yields the bound on d as desired. ∎

Define points

$$X_1 = (x_s, y_s + y_2), \quad X_2 = (x_t, y_t + y_2) \quad \text{and} \quad X_3 = (x_s, y_t + y_2).$$

Let P_1X_1 and P_2X_2 be line segments which are parallel to PX_3 (see Figure 4). Construct a line passing through $(0, y_2)$ which is also parallel to the line segments PX_3, where $P = (\frac{4n}{5}, \frac{4n}{5})$. Such a line separates T_2 into two parts, and we denote by $T_{2,2}$ the part above the line and $T_{2,1}$ for the rest.

Claim 3.8. *There exists a triangle $T_{\beta,1} \subseteq T_\beta$ similar to T_β such that $T_{\beta,1}$ does not intersect with either of the regions $T_{2,1} + s$ or $T_{2,2} + t$ and $\Lambda(T_{\beta,1}) \ge \frac{\beta^2 n^2}{8}$.*

Proof of claim. Let $h := y_t - y_s$. Since s and t are of distance at least βn far apart, that is, $(x_s - x_t)^2 + (y_s - y_t)^2 = (h + d)^2 + h^2 \ge \beta^2 n^2$, together with Claim 3.7, we obtain that either $h \ge \frac{\sqrt{2\beta^2 - d^2}}{2} - \frac{d}{2} \ge \frac{\beta n}{2}$ or $h \le -\frac{\sqrt{2\beta^2 - d^2}}{2} - \frac{d}{2} < -d$, where the latter contradicts with the assumption that $x_s - x_t = h + d \ge 0$. Thus, $h \ge \frac{\beta n}{2}$ and the segment P_1P_2 has length at least $\frac{\sqrt{2}}{2}\beta n$. Let $T_{\beta,1}$ be the rectangle triangle P_1P_2Q with diagonal line P_1P_2. Then $T_{\beta,1}$ has area at least $\frac{\beta^2 n^2}{8}$ and does not intersect either of the regions $T_{2,1} + s$ or $T_{2,2} + t$. ∎

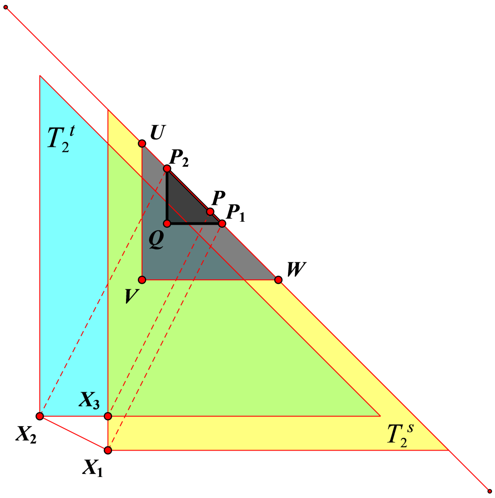

Figure 4: The triangle UVW represents T_β, in which $P = (\frac{4n}{5}, \frac{4n}{5})$ is the median point for the line segment UW.

As aforementioned, now we are ready to finish the proof. Applying Lemma 2.6 to $T_{2,1}, T_{2,2}$ and their translates $T_{2,1} + s$, $T_{2,2} + t$, we obtain that

$$\begin{aligned}
|S| &\leq n^2 - \Lambda(A) - \frac{1}{2}\Lambda(P_2) - \Lambda(T_2) - \Lambda(T_{\beta,1}) \\
&= n^2 - \|A\| - \frac{1}{2}\|P_2\| - \|T_2\| - \|T_{\beta,1}\| + O(n) \\
&\leq n^2 - \frac{1}{2}\left(\frac{2}{5} - \sqrt{2}\varepsilon\right)^2 n^2 - \frac{1}{2}x_2y_2 - \frac{1}{2}\left(\frac{4}{5} - \frac{\varepsilon}{\sqrt{2}} - \frac{x_2}{n}\right)^2 n^2 - \frac{\beta^2n^2}{8} + O(n).
\end{aligned}$$

The right-hand side above is maximized when y_2 is minimal and $x_2 + y_2 = \frac{8n}{5} - \sqrt{2}\varepsilon n$. Thus,

$$|S| \leq \left(\frac{3}{5} - \frac{\beta^2}{8} + O(\varepsilon)\right) n^2,$$

a final contradiction. $\square$

3.3 Putting things together

We are now ready to prove our main result, knowing that any almost maximum sum-free set contains an upper boundary line $o(1)$-close to $y + x = \frac{8n}{5}$ and a point $o(1)$-close to $(\frac{4n}{5}, \frac{4n}{5})$.

Proof of Theorem 1.2. Given $\gamma > 0$, choose $\delta \ll \varepsilon \ll \beta \ll \gamma$. Let $S \subseteq [n]^2$ be a sum-free set of size at least $(3/5 - \delta)n^2$. Then by Lemma 3.3, S has a typical upper boundary line $y = mx + c$ which is ε-close to $y + x = \frac{8n}{5}$. Now it suffices to show that S has no point below the line $x + y = \frac{4n}{5} - \gamma n$ (see the red line in Figure 5).

Note that Lemma 3.4 ensures the existence of a point (x_1, y_1) in S that is β-close to $(\frac{4n}{5}, \frac{4n}{5})$. Suppose to the contrary that $p_0 = (x_0, y_0) \in S$ is such a point below the line $x + y = \frac{4n}{5} - \gamma n$, and without loss of generality we may assume that $y_0 \geq x_0$.

Let

$$A =: \left\{(x, y) \in [n]^2 \mid y + x > \frac{8n}{5} + \varepsilon n\right\}.$$

Considering the pairing set $P := \{(x,y) \in [n]^2 \mid x \le x_1, y \le y_1\}$ for (x_1,y_1), there are at most

$$n^2 - \Lambda(A) - \frac{1}{2}\Lambda(P) \le \left(\frac{3}{5} + \left(\frac{2}{5} - \frac{\varepsilon}{2}\right)\varepsilon + \left(\frac{4}{5} - \frac{\beta}{2}\right)\beta\right) n^2 + O(n) \tag{1}$$

points which may be included in S; and all these points are below the line $x + y = \frac{8n}{5} + \varepsilon n$. Then, writing

$$D_1 := \left\{(x,y) \in [n]^2 \mid y > \frac{4n}{5} + \beta n,\ y + x < \frac{8n}{5} - \varepsilon n\right\}$$

and

$$D_2 := \left\{(x,y) \in [n]^2 \mid x > \frac{4n}{5} + \beta n,\ y + x < \frac{8n}{5} - \varepsilon n\right\},$$

it follows from the assumption $|S| \ge (3/5 - \delta)n^2$ and (1) that

$$\frac{1}{n^2}|(D_1 \cup D_2) \setminus S| \le \delta + \left(\frac{2}{5} - \frac{\varepsilon}{2}\right)\varepsilon + \left(\frac{4}{5} - \frac{\beta}{2}\right)\beta =: v(\delta,\varepsilon,\beta). \tag{2}$$

Note that we can choose $\delta, \varepsilon, \beta$ small enough such that $v(\delta,\varepsilon,\beta) = o(\gamma^2)$. In the remaining proof, we shall find in $D_1 \cup D_2$ (or its translate) a relatively large subset of lattice points which are to be excluded from S, yielding a contradiction.

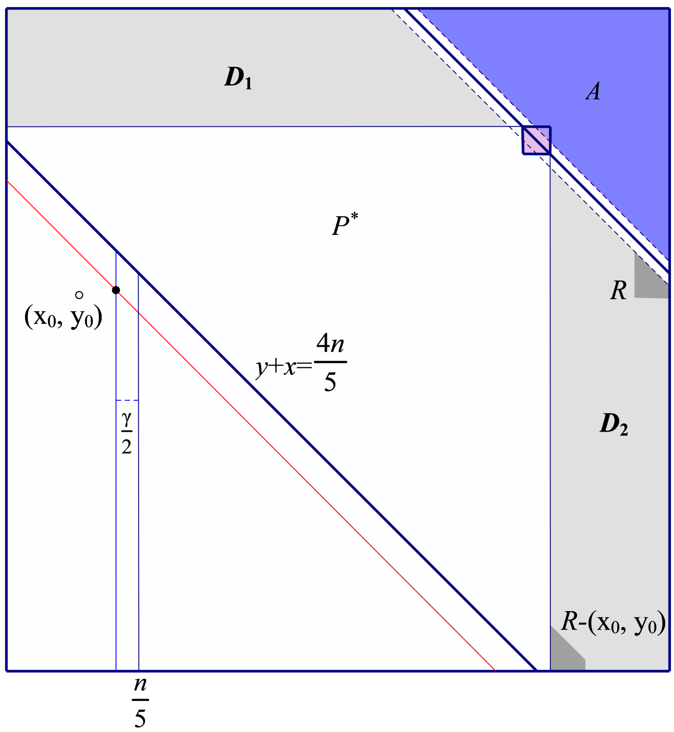

Figure 5: The case when $x_0 < \frac{n}{5} - \frac{\gamma}{2}n$: The two grey regions $R := (D_2 + p_0) \cap D_2$ and its translate $R - p_0$ form a pairing, which excludes from the sum-free set S the amount of points which fit in one of the regions.

First assume that p_0 is such that $x_0 < \frac{n}{5} - \frac{\gamma}{2}n$. Then the region $D_2 + p_0$ intersects D_2 on a set of lattice points, denoted by R. Since $R, R - p_0 \subseteq D_1 \cup D_2$, applying Lemma 2.6 with $a = p_0$ and $T = R - p_0$ gives that $|(R \cup (R - p_0)) \cap S| \le |R|$, and thus

$$|(D_1 \cup D_2) \setminus S| \ge |(R \cup (R - p_0)) \setminus S| \ge |R \cup (R - p_0)| - |R|.$$

It is easy to observe that $|R \cup (R - p_0)| - |R|$ is minimized when p_0 is close to the point $(\frac{n}{5} - \frac{\gamma n}{2}, \frac{3n}{5} - \frac{\gamma n}{2})$, yielding an area of size at least $\left(\frac{3}{8}\gamma^2 + \frac{\gamma - 2\beta}{4}(\beta - 2\varepsilon)\right) n^2 + O(n)$ (See Figure 5). Thus $|(D_1 \cup D_2) \setminus S| \geq \left(\frac{3}{8}\gamma^2 + \frac{\gamma - 2\beta}{4}(\beta - 2\varepsilon)\right) n^2 + O(n) > \upsilon(\delta, \varepsilon, \beta)n^2$, a contradiction to (2).

Now it remains to consider the case when p_0 satisfies $x_0 \geq \frac{n}{5} - \frac{\gamma n}{2}$. We consider the overlap of $(D_1 \cup D_2) - p_0$ with $(x_1, y_1) - ((D_1 \cup D_2) - p_0)$ and denote by $\mathcal{O}$ the set of lattice points in the overlap (see Figure 6). Let

$$D := ((D_1 \cup D_2) - p_0) \setminus (D_1 \cup D_2).$$

Then it is easy to verify that $\mathcal{O} \subseteq D$. Note that by Lemma 2.6 with $a = p_0$ and $T = D_1 \cup D_2$, one has that

$$|(D \cup D_1 \cup D_2) \cap S| \leq |D_1 \cup D_2|.$$

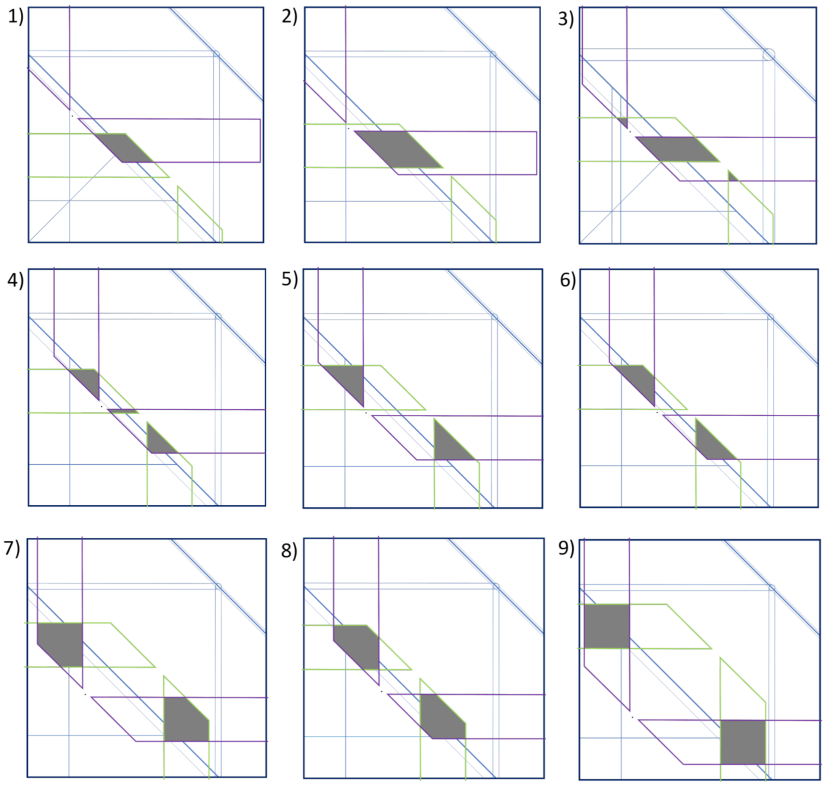

Figure 6: All possible shapes of $\mathcal{O}$: the green lines frame the region $(D_1 \cup D_2) - p_0$, whilst the purple lines frame $(x_1, y_1) - ((D_1 \cup D_2) - p_0)$. The trivial cases where the overlap is cut off by the x- and y-axes are not shown.

Then, using (2), we have

$$\begin{aligned} |\mathcal{O} \cap S| \leq |D \cap S| &= |(D \cup D_1 \cup D_2) \cap S| - |(D_1 \cup D_2) \cap S| \\ &\leq |(D_1 \cup D_2) \setminus S| \leq \upsilon(\delta, \varepsilon, \beta)n^2. \end{aligned}$$

Moreover, by definition we know that $(x_1, y_1) - \mathcal{O} \subseteq \mathcal{O}$, that is, $\mathcal{O}$ (and also $P \setminus \mathcal{O}$) is a pairing set for (x_1, y_1). It follows from Lemma 2.5 that

$$
\begin{aligned}
|S| &\le n^2 - \Lambda(A) - \frac{1}{2}\Lambda(P \setminus \mathcal{O}) - (|\mathcal{O}| - |\mathcal{O} \cap S|) \\
&\le n^2 - \|A\| - \frac{1}{2}\|P\| - \frac{1}{2}|\mathcal{O}| + \upsilon(\delta, \varepsilon, \beta)n^2 \\
&\le n^2 - \frac{1}{2}\Big(\frac{2n}{5} - \varepsilon n\Big)^2 - \frac{1}{2}\Big(\frac{4n}{5} - \beta n\Big)^2 - \frac{1}{2}|\mathcal{O}| + \upsilon(\delta, \varepsilon, \beta)n^2 \\
&= \frac{3}{5}n^2 + o(\gamma^2)n^2 - \frac{1}{2}|\mathcal{O}|.
\end{aligned}
$$

Therefore, it suffices to show that $|\mathcal{O}| = \Omega(\gamma^2)n^2$, and in the remaining proof we shall verify this by considering all possible shapes of $\mathcal{O}$.

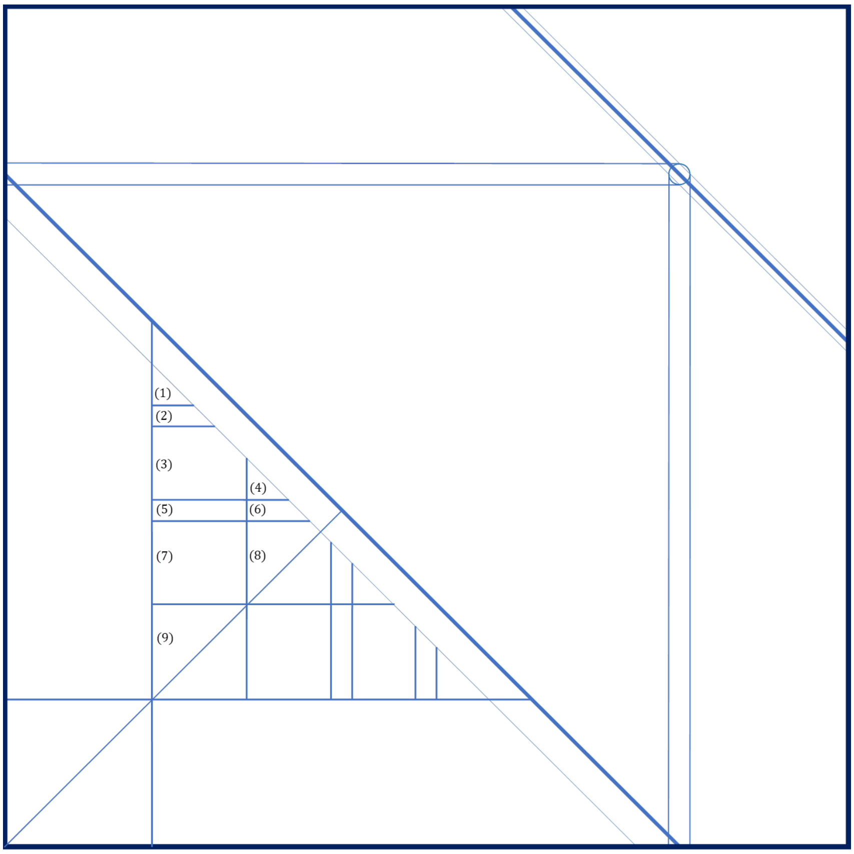

Figure 7: each numbered region will produce a unique shape of the overlap.

Since (x_1, y_1) is β-close to $(\frac{4n}{5}, \frac{4n}{5})$ and $\beta \ll \gamma$, we may further assume that $(x_1, y_1) = (\frac{4n}{5}, \frac{4n}{5})$ in order not to cluster the presentation. We list in Figure 6 all possible shapes of the overlap $\mathcal{O}$, which originate from the location of the point (x_0, y_0) (see Figure 7). In particular, the area of the overlap in each of these cases is given as follows:

(1) $4\left(\frac{3}{5}n - y_0\right)\left(\frac{4}{5}n - y_0 - x_0 - \varepsilon n\right)$, where $y_0 \ge \frac{n}{2} + \frac{\beta}{2}n$, $x_0 \ge \frac{1}{5}n - \frac{\gamma}{2}n$.

(2) $4\left(y_0 - \frac{2}{5}n - \beta n\right)\left(\frac{4}{5}n - y_0 - x_0 - \varepsilon n\right)$, where $y_0 \in [\frac{1}{2}n - \frac{\beta+\varepsilon}{2}n, \frac{1}{2}n + \frac{\beta}{2}n]$, $x_0 \ge \frac{1}{5}n - \frac{\gamma}{2}n$.

(3) $(n - 2y_0 - \beta n - \varepsilon n)^2 + 4\left(y_0 - \frac{2}{5}n - \beta n\right)\left(\frac{4}{5}n - y_0 - x_0 - \varepsilon n\right)$, where

$$
y_0 \in \left[\frac{2}{5}n + \beta n, \frac{1}{2}n - \frac{\beta + \varepsilon}{2}n\right], x_0 \in \left[\frac{1}{5}n - \frac{\gamma}{2}n, \frac{3}{10}n + \frac{\beta - \varepsilon}{2}n\right].
$$

(4) $4\left(\frac{4}{5}n-y_0-x_0-\varepsilon n\right)\left(x_0-\frac{1}{5}n-2\beta n\right)$, where $y_0 \geq \frac{2}{5}n+\beta n, x_0 \geq \frac{3}{10}n+\frac{\beta-\varepsilon}{2}n$.

(5) $(n-2y_0-\beta n-\varepsilon n)^2$, where

$$y_0 \in \left[\frac{2}{5}n-\frac{\varepsilon}{2}n, \frac{2}{5}n+\beta n\right], x_0 \in \left[\frac{1}{5}n-\frac{\gamma}{2}n, \frac{3}{10}n+\frac{\beta-\varepsilon}{2}n\right].$$

(6) $4\left(\frac{4}{5}n-y_0-x_0-\varepsilon n\right)\left(\frac{1}{5}n+x_0-y_0-\beta n\right)$, where

$$y_0 \in \left[\frac{2}{5}n-\frac{\varepsilon}{2}n, \frac{2}{5}n+\beta n\right], x_0 \geq \frac{3}{10}n+\frac{\beta-\varepsilon}{2}n.$$

(7) $2\left(\frac{1}{5}n-\beta n\right)^2-\left(2y_0-\frac{3}{5}n+\varepsilon n-\beta n\right)^2$, where

$$y_0 \in \left[\frac{3}{10}n+\frac{\beta-\varepsilon}{2}n, \frac{2}{5}n-\frac{\varepsilon}{2}n\right], x_0 \in \left[\frac{1}{5}n-\frac{\gamma}{2}n, \frac{3}{10}n+\frac{\beta-\varepsilon}{2}n\right].$$

(8) $2\left(\frac{1}{5}n-\beta n\right)^2-\left(2y_0-\frac{3}{5}n+\varepsilon n-\beta n\right)^2-\left(2x_0-\frac{3}{5}n+\varepsilon n-\beta n\right)^2$, where

$$\frac{3}{10}n+\frac{\beta-\varepsilon}{2}n \leq x_0 \leq y_0 \leq \frac{2}{5}n-\frac{\varepsilon}{2}n.$$

(9) $2\left(\frac{1}{5}n-\beta n\right)^2$, where $\frac{1}{5}n-\frac{\gamma}{2}n \leq x_0 \leq y_0 \leq \frac{3}{10}n+\frac{\beta-\varepsilon}{2}n$.

It is obvious that for the regions 5, 7 and 9, the area of the overlap has size $\Omega(\gamma^2)n^2$. The only regions which interest us are the ones bordering the line $y+x=\frac{4n}{5}-\gamma n$. Moreover, the regions in question are 1, 2, 3, 4, 6 and 8. Among them, the minimum overlap is achieved in region 1 by letting $(x_0,y_0)=(\frac{n}{5}-\frac{\gamma n}{2}, \frac{3n}{5}-\frac{\gamma n}{2})$, which yields a value of $|\mathcal{O}| \geq 2\gamma(\gamma-\varepsilon)n^2$ as desired. This completes the proof of Theorem 1.2. $\square$

4 Proof of Theorem 1.3

In this section we investigate the maximum size of a (p,p)-sum-free set S. To simplify the presentation, we write p-sum-free for (p,p)-sum-free. Our proof builds on the techniques developed in the work of Elsholtz and Rackham [ER17]. We need a variant notion of pairing set as follows.

Definition 4.1. For any $(a_1,a_2) \in \mathbb{R}^2_{[0,n]}$, $P \subseteq \mathbb{R}^2_{[0,n]}$ is a p-pairing set for (a_1,a_2) if, for any $(x_1,x_2) \in P$, we have $(\frac{a_1}{p}-x_1, \frac{a_2}{p}-x_2) \in P$.

Similar to Lemmas 2.5 and 2.6, the following lemma guarantees that for any point $a \in S$ and its p-pairing set P, at least half of the points in P are excluded from S. Similar statement also holds when we consider a set and its translate dilated by p. We omit the proof.

Lemma 4.2. *Let $S \subseteq [n]^2$ be a p-sum-free set.*

(1) *If P is a p-pairing set for some $a \in S$, then we have $|S \cap P| \leq \frac{1}{2}\Lambda(P)$.*

(2) *If $T \subseteq \mathbb{R}^2_{[0,n]}$ and $a \in S$, then $|S \cap (p(a+T) \cup T)| \leq \Lambda(T)$.*

Proof of Theorem 1.3. Let $S \subseteq [n]^2$ be a p-sum-free set. Our goal is to show that $|S| \leq \left(1 - \frac{2}{4p^2+1}\right) n^2 + O(n)$ for $p \geq 2$. We may neglect any boundary effects as they give error terms $O(n)$ for the size of S, which will be omitted so as to ease the presentation. We consider cases depending on the placement of upper boundary lines.

Case 1: $|\partial S| \leq 1$. As vertices in the upper boundary come in (adjoint) pairs, we see that in this case $\partial S = \emptyset$, and thus Lemma 2.3 ensures the existence a point $p_1 = (x_1, y_1) \in S$ such that $x_1 \geq x$ and $y_1 \geq y$ for all $(x, y) \in S$. Let $P := \{(x, y) \mid 0 \leq x \leq \frac{x_1}{p}, 0 \leq y \leq \frac{y_1}{p}\}$. Then P is a p-pairing set for p_1 and thus by Lemma 4.2, we have that

$$\begin{aligned}|S| &\leq (n+1)^2 - (n-x_1)n - (n-y_1)x_1 - \frac{1}{2}\Lambda(P)\\ &= \left(1 - \frac{1}{2p^2}\right) x_1 y_1 + O(n) \leq \left(1 - \frac{1}{2p^2}\right) n^2 + O(n) < \left(1 - \frac{2}{4p^2+1}\right) n^2.\end{aligned}$$

Case 2: $|\partial S| \geq 2$ and for every two points $p_1 = (x_1, y_1), p_2 = (x_2, y_2)$ that are adjoint in ∂S with $x_1 < x_2$ and $y_1 > y_2$, we have either $m(p_1, p_2) > -\frac{y_2}{x_2}$ or $m(p_1, p_2) < -\frac{y_1}{x_1}$.

In this case, we choose $p_1 = (x_1, y_1) \in \partial S$ such that $x_1 y_1 \geq xy$ holds for every $(x, y) \in \partial S$ and $P_1 := \{(x, y) \mid 0 \leq x \leq \frac{x_1}{p}, 0 \leq y \leq \frac{y_1}{p}\}$. By symmetry, we may further assume that $y_1 \geq x_1$. If there does not exist $p_2 = (x_2, y_2) \in \partial S$ adjoint to p_1 with $x_2 > x_1$ and $y_2 < y_1$, then by Lemma 4.2 and that $y_1 \geq x_1$, we have

$$|S| \leq n^2 - (n - x_1)n - \frac{1}{2}\Lambda(P_1) \leq nx_1 - \frac{x_1^2}{2p^2} \leq \left(1 - \frac{1}{2p^2}\right) n^2.$$

Thus, we may assume that there exists $p_2 = (x_2, y_2) \in \partial S$ adjoint to p_1 with $x_2 > x_1$ and $y_2 < y_1$. Let $L : y = mx + c$ be the line passing through p_1, p_2 and define

$$A = \{(x, y) \in \mathbb{R}^2_{[0,n]} \mid y > mx + c\}.$$

We claim that $m < -\frac{y_1}{x_1} \leq -1$. Indeed, by the assumption of Case 2, assume for contradiction that $m > -\frac{y_2}{x_2}$, then

$$x_2 y_2 = x_2(y_1 + m(x_2 - x_1)) \geq x_2 y_1 - y_2(x_2 - x_1) = x_1 y_1 + (y_1 - y_2)(x_2 - x_1) > x_1 y_1,$$

contrary to the choice of p_1.

We split into two subcases depending on the x- and y-intercept of L. Note first that, if $c \leq n$, then we have $-\frac{c}{m} \leq n$ because $m \leq -1$, and so $|S| < \frac{1}{2}n^2$ as $A \cap S = \emptyset$.

(I). If $c > n$ and $-\frac{c}{m} \leq n$, then

$$|S| \leq n^2 - \Lambda(A) - \frac{1}{2}\Lambda(P_1) = \frac{n}{m}\left(\frac{n}{2} - c\right) - \frac{1}{2p^2} x_1 y_1 = \frac{n}{m}\left(\frac{n}{2} - y_1\right) + x_1 n - \frac{1}{2p^2} x_1 y_1.$$

Now if $y_1 \leq \frac{n}{2}$, then as $m < -1$ and $x_1 \leq y_1 \leq \frac{n}{2}$, we observe that $|S| \leq x_1 n \leq \frac{1}{2}n^2$. We may then assume $y_1 > \frac{n}{2}$.

If $x_1 < \frac{n}{2}$, then by the assumption that $m < -\frac{y_1}{x_1}$, we have

$$|S| \leq \frac{nx_1}{y_1}\left(y_1 - \frac{n}{2}\right) + x_1 n - \frac{1}{2p^2} x_1 y_1 \leq \left(2n - \frac{n^2}{2y_1} - \frac{y_1}{2p^2}\right)\frac{n}{2} \leq \left(1 - \frac{1}{2p}\right) n^2.$$

where the last inequality follows from $\frac{n^2}{2y_1} + \frac{y_1}{2p^2} \geq 2\sqrt{\frac{n^2}{2y_1}\frac{y_1}{2p^2}} = \frac{n}{p}$.

Assume then $x_1 \geq \frac{n}{2}$. Note that as $-\frac{c}{m} \leq n$, the slope of L is smaller than the slope of the line passing through p_1 and $(n,0)$, and so $m \leq \frac{-y_1}{n-x_1}$. Thus, we have

$$\begin{aligned}|S| &\leq \frac{n(n-x_1)}{y_1}\left(y_1 - \frac{n}{2}\right) + x_1 n - \frac{1}{2p^2}x_1y_1 \leq n^2 - \left(\frac{n-x_1}{2n}n^2 + \frac{x_1y_1}{2p^2}\right)\\ &= \frac{n^2}{2} + \left(\frac{n}{2} - \frac{y_1}{2p^2}\right)x_1 \leq \frac{n^2}{2} + \left(\frac{n}{2} - \frac{y_1}{2p^2}\right)y_1 \leq \left(1 - \frac{1}{2p^2}\right)n^2,\end{aligned}$$

where the second last inequality follows since $x_1 \leq y_1$ and the last one follows from $p \geq 2$.

(II). If $c > n$ and $-\frac{c}{m} > n$, then A is a triangle and thus

$$\begin{aligned}|S| &\leq n^2 - \Lambda(A) - \frac{1}{2}\Lambda(P_1)\\ &= n^2 + \frac{(n-y_1)^2}{2m} + \frac{m(n-x_1)^2}{2} - (n-x_1)(n-y_1) - \frac{x_1y_1}{2p^2}.\end{aligned}$$

The right-hand side above is increasing when $m \leq -\frac{n-y_1}{n-x_1}$. Since $\frac{n-y_1}{n-x_1} \leq \frac{y_1}{x_1} \leq -m$, it follows that

$$\begin{aligned}|S| &\leq n^2 - \frac{(n-y_1)^2}{\frac{2y_1}{x_1}} - \frac{y_1(n-x_1)^2}{2x_1} - (n-x_1)(n-y_1) - \frac{x_1y_1}{2p^2}\\ &\leq 2(x_1+y_1)n - n^2 - \left(2 + \frac{1}{2p^2}\right)x_1y_1,\end{aligned}$$

where the right-hand side of the last inequality is maximized when $x_1 = y_1 = \frac{4p^2}{4p^2+1}n$, and thus $|S| \leq \left(1 - \frac{2}{4p^2+1}\right)n^2$.

Case 3: There exist $p_1 = (x_1,y_1), p_2 = (x_2,y_2)$ adjoint in ∂S such that $x_1 < x_2$, $y_1 > y_2$ and $-\frac{y_1}{x_1} \leq m(p_1,p_2) \leq -\frac{y_2}{x_2}$.

For each p_i with $i \in [2]$, define $P_i := \{(x,y) \mid 0 \leq x \leq \frac{x_i}{p}, 0 \leq y \leq \frac{y_i}{p}\}$ and set $A = \{(x,y) \in \mathbb{R}^2_{[0,n]} \mid y > mx + c\}$ (see Figure 8). Since $m \leq -\frac{y_2}{x_2}$ and $y_2 = mx_2 + c$, we have that $y_2 \leq \frac{c}{2}$. Similarly, by the condition $m \geq -\frac{y_1}{x_1}$, we have that $y_1 \geq \frac{c}{2}$.

Define

$$T_1 = \{(x,y) \in \mathbb{R}^2_{[0,n]} \mid x \geq \frac{x_1}{p}, y \leq mx + \frac{c}{2p}\},$$

and

$$T_2 = \{(x,y) \in \mathbb{R}^2_{[0,n]} \mid y \geq \frac{y_2}{p}, y \leq mx + \frac{c}{2p}\}.$$

We claim that $T_1, T_2 \neq \emptyset$. These amount to proving $-\frac{c}{2mp} \geq \frac{x_1}{p}$ and $\frac{c}{2p} \geq \frac{y_2}{p}$, which in turn follows from the fact that $y_2 \leq \frac{c}{2} \leq y_1$.

If $T_1 \cap S = \emptyset$, then a short calculation shows

$$|S| \leq n^2 - \Lambda(T_1) - \Lambda(A) - \frac{1}{2}\Lambda(P_1) \leq n^2 + \frac{c^2}{8p^2m} - \|A\|.$$

If $T_1 \cap S \neq \emptyset$, then take a point $a \in T_1 \cap S$, then one can check that $p(a+T_2) \cap T_2 = \emptyset$. By Lemma 4.2, we have $|S \cap (p(a+T_2) \cup T_2)| \leq \Lambda(T_2)$. By the definition of T_2, any point $(x,y) \in p(a+T_2)$ satisfies that $y \leq mx + c$ and $x \geq x_1, y \geq y_2$. We again arrive to

$$|S| \leq n^2 - \Lambda(T_2) - \Lambda(A) - \frac{1}{2}\Lambda(P_2) \leq n^2 + \frac{c^2}{8p^2m} - \|A\|.$$

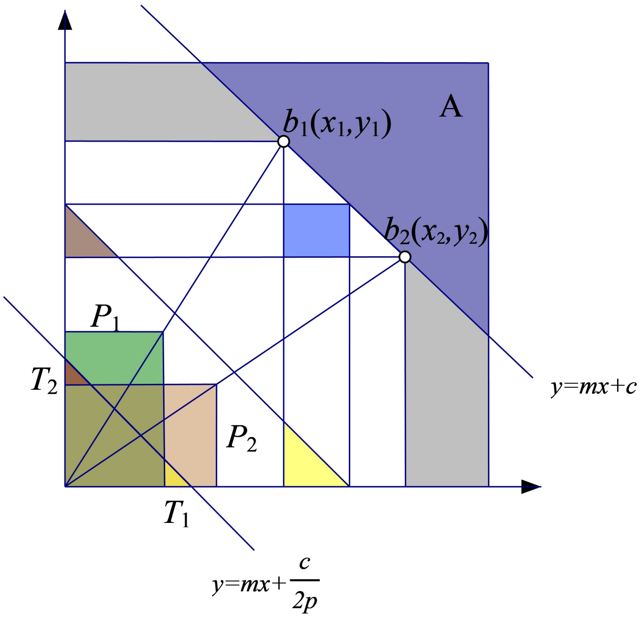

Figure 8: $T_1, T_2 \neq \emptyset$.

Suppose now that $c \leq n$. If $-\frac{c}{m} \leq n$, then $|S| \leq \frac{1}{2}n^2$ by excluding A alone. So $-\frac{c}{m} > n$. Then $\|A\| = \frac{n(2n-mn-2c)}{2}$ and we get, using $c \leq n$ and $x+y \geq 2\sqrt{xy}$ for $x, y > 0$,

$$|S| \leq n^2 + \frac{c^2}{8p^2m} - \|A\| = \frac{c^2}{8p^2m} + \frac{n^2m}{2} + cn \leq cn - \frac{cn}{2p} \leq \left(1 - \frac{1}{2p}\right)n^2.$$

We may then assume $c > n$. The case $-\frac{c}{m} \leq n$ can be handled as the above $c \leq n$ and $-\frac{c}{m} \geq n$ case. Thus, we can assume $-\frac{c}{m} > n$. Then A is a triangle with $\|A\| = -\frac{(n-mn-c)^2}{2m}$ and

$$|S| \leq n^2 + \frac{c^2}{8p^2m} + \frac{(n-mn-c)^2}{2m} = \left(\frac{1}{8p^2} + \frac{1}{2}\right)\frac{c^2}{m} + \frac{n(m-1)}{m}c + \frac{n^2m}{2} + \frac{n^2}{2m}.$$

The quadratic function of c above is maximized when $c = -\frac{(m-1)n}{1+\frac{1}{4p^2}}$. Thus

$$|S| \leq n^2\left[\frac{4p^2}{4p^2+1} + \left(\frac{1}{2} - \frac{2p^2}{4p^2+1}\right)\left(m + \frac{1}{m}\right)\right] \leq \left(1 - \frac{2}{4p^2+1}\right)n^2,$$

where the maximum is achieved when we choose $m = -1$ and thus $c = \frac{8p^2}{4p^2+1}n$.

This completes the proof. □

References

[Blo16] T. F. Bloom, *A quantitative improvement for Roth's theorem on arithmetic progressions*, J. Lond. Math. Soc. (2) **93** (2016), no. 3, 643–663. MR 3509957

[BLST18] József Balogh, Hong Liu, Maryam Sharifzadeh, and Andrew Treglown, *Sharp bound on the number of maximal sum-free subsets of integers*, J. Eur. Math. Soc. (JEMS) **20** (2018), no. 8, 1885–1911. MR 3854894

[Cam02] Peter J. Cameron, *Sum-free sets of a square*, manuscript (2002).

[Cam05] ______, *Research problems from the 19th British Combinatorial Conference*, Discrete Math. **293** (2005), no. 1-3, 313–320. MR 2136071

[CKP20] Ilkyoo Choi, Ringi Kim, and Boram Park, *Maximum k-sum n-free sets of the 2-dimensional integer lattice*, Electron. J. Combin. **27** (2020), no. 4, Paper No. 4.2, 12. MR 4245177

[DFST99] Jean-Marc Deshouillers, Gregory A. Freiman, Vera Sós, and Mikhail Temkin, *On the structure of sum-free sets. II*, no. 258, 1999, Structure theory of set addition, pp. xii, 149–161. MR 1701193

[ER17] Christian Elsholtz and Laurence Rackham, *Maximal sum-free sets of integer lattice grids*, J. Lond. Math. Soc. (2) **95** (2017), no. 2, 353–372. MR 3656272

[Fre92] Gregory A. Freiman, *On the structure and the number of sum-free sets*, no. 209, 1992, Journées Arithmétiques, 1991 (Geneva), pp. 13, 195–201. MR 1211012

[HT17] Robert Hancock and Andrew Treglown, *On solution-free sets of integers*, European J. Combin. **66** (2017), 110–128. MR 3692141

[Rot53] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 104–109. MR 51853

[Ruz93] Imre Z. Ruzsa, *Solving a linear equation in a set of integers. I*, Acta Arith. **65** (1993), no. 3, 259–282. MR 1254961

[Ruz95] ______, *Solving a linear equation in a set of integers. II*, Acta Arith. **72** (1995), no. 4, 385–397. MR 1348205

[Sch16] I. Schur, *Uber die kongruenz $x^m + y^m \equiv z^m \pmod p$*, Jahresber. Deutsch. Math.-Verein. **25** (1916), 114–117.

[Skr93] Maxim Skriganov, *On integer points in polygons*, Ann. Inst. Fourier (Grenoble) **43** (1993), no. 2, 313–323. MR 1220271

[Tra18] Tuan Tran, *On the structure of large sum-free sets of integers*, Israel J. Math. **228** (2018), no. 1, 249–292. MR 3874843