Sharp bound on the number of maximal sum-free subsets of integers

József Balogh * Hong Liu † Maryam Sharifzadeh ‡ Andrew Treglown §

February 26, 2015

Abstract

Cameron and Erdős [6] asked whether the number of *maximal* sum-free sets in $\{1, \dots, n\}$ is much smaller than the number of sum-free sets. In the same paper they gave a lower bound of $2^{\lfloor n/4 \rfloor}$ for the number of maximal sum-free sets. Here, we prove the following: For each $1 \leq i \leq 4$, there is a constant C_i such that, given any $n \equiv i \mod 4$, $\{1, \dots, n\}$ contains $(C_i + o(1))2^{n/4}$ maximal sum-free sets. Our proof makes use of container and removal lemmas of Green [10, 11], a structural result of Deshouillers, Freiman, Sós and Temkin [7] and a recent bound on the number of subsets of integers with small sumset by Green and Morris [12]. We also discuss related results and open problems on the number of maximal sum-free subsets of abelian groups.

1 Introduction

A triple x, y, z is a *Schur triple* if $x + y = z$ (note x, y and z may not necessarily be distinct). A set S is *sum-free* if S does not contain a Schur triple. Let $[n] := \{1, \dots, n\}$. We say that $S \subseteq [n]$ is a *maximal sum-free subset of* $[n]$ if it is sum-free and it is not properly contained in another sum-free subset of $[n]$. Let $f(n)$ denote the number of sum-free subsets of $[n]$ and $f_{\max}(n)$ denote the number of maximal sum-free subsets of $[n]$. The study of sum-free sets of integers has a rich history. Clearly, any set of odd integers and any subset of $\{\lfloor n/2 \rfloor + 1, \dots, n\}$ is a sum-free set, hence $f(n) \geq 2^{n/2}$. Cameron and Erdős [5] conjectured that $f(n) = O(2^{n/2})$. This conjecture was proven independently by Green [10] and Sapozhenko [17]. In fact, they showed that there are constants C_1 and C_2 such that $f(n) = (C_i + o(1))2^{n/2}$ for all $n \equiv i \mod 2$.

In a second paper, Cameron and Erdős [6] showed that $f_{\max}(n) \geq 2^{\lfloor n/4 \rfloor}$. Noting that all the sum-free subsets of $[n]$ described above lie in just two maximal sum-free sets, they asked whether $f_{\max}(n) = o(f(n))$ or even $f_{\max}(n) \leq f(n)/2^{\varepsilon n}$ for some constant $\varepsilon > 0$. Łuczak and Schoen [15] answered this question in the affirmative, showing that $f_{\max}(n) \leq 2^{n/2 - 2^{-28} n}$ for sufficiently large n. Later, Wolfovitz [19] proved that $f_{\max}(n) \leq 2^{3n/8 + o(n)}$. More recently, the authors [2] proved

*Department of Mathematical Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA jobal@math.uiuc.edu. Research is partially supported by Simons Fellowship, NSA Grant H98230-15-1-0002, NSF CAREER Grant DMS-0745185, Arnold O. Beckman Research Award (UIUC Campus Research Board 15006) and Marie Curie FP7-PEOPLE-2012-IIF 327763.

†Department of Mathematical Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA hliu36@illinois.edu.

‡Department of Mathematical Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA sharifz2@illinois.edu.

§School of Mathematics, University of Birmingham, United Kingdom a.c.treglown@bham.ac.uk.

that the lower bound is essentially tight, proving that $f_{\max}(n) = 2^{(1/4+o(1))n}$. In this paper we give the following exact solution to the problem.

Theorem 1.1. *For each $1 \leq i \leq 4$, there is a constant C_i such that, given any $n \equiv i \mod 4$, $[n]$ contains $(C_i + o(1))2^{n/4}$ maximal sum-free sets.*

The proof of Theorem 1.1 is given in Section 4, with the main work arising in Section 4.1. The proof draws on a number of ideas from [2]. In particular, as in [2] we make use of 'container' and 'removal' lemmas of Green [10, 11] as well as a result of Deshouillers, Freiman, Sós and Temkin [7] on the structure of sum-free sets. In order to avoid over-counting the number of maximal sum-free subsets of $[n]$, our present proof also develops a number of new ideas, thereby making the argument substantially more involved. We use a bound on the number of subsets of integers with small sumset by Green and Morris [12] as well as several new bounds on the number of maximal independent sets in various graphs. Further, the proof provides information about the typical structure of the maximal sum-free subsets of $[n]$. Indeed, we show that almost all of the maximal sum-free subsets of $[n]$ look like one of two particular extremal constructions (see Section 2.3 for more details).

In Section 2 we give an overview of the proof and highlight the new ideas that we develop. We state some useful results in Section 3 and prove Theorem 1.1 in Section 4. In Section 5 we give some results and open problems on the number of maximal sum-free subsets of abelian groups.

2 Background and an overview of the proof of Theorem 1.1

2.1 Independence and container theorems

An exciting recent development in combinatorics and related areas has been the emergence of 'independence' as a unifying concept. To be more precise, let V be a set and $\mathcal{E}$ a collection of subsets of V. We say that a subset I of V is an *independent set* if I does not contain any element of $\mathcal{E}$ as a subset. For example, if $V := [n]$ and $\mathcal{E}$ is the collection of all Schur triples in $[n]$ then an independent set I is simply a sum-free set. It is often helpful to think of $(V, \mathcal{E})$ as a hypergraph with vertex set V and edge set $\mathcal{E}$; thus an independent set I corresponds to an independent set in the hypergraph.

So-called 'container results' have emerged as a powerful tool for attacking many problems that concern counting independent sets. Roughly speaking, container results state that the independent sets of a given hypergraph H lie only in a 'small' number of subsets of the vertex set of H (referred to as *containers*), where each of these containers is an 'almost independent set'. Balogh, Morris and Samotij [3] and independently Saxton and Thomason [18], proved general container theorems for hypergraphs whose edge distribution satisfies certain boundedness conditions.

In the proof of Theorem 1.1 we will apply the following container theorem of Green [10].

Lemma 2.1 (Proposition 6 in [10]). *There exists a family $\mathcal{F}$ of subsets of $[n]$ with the following properties.*

(i) Every member of $\mathcal{F}$ has at most $o(n^2)$ Schur triples.
(ii) If $S \subseteq [n]$ is sum-free, then S is contained in some member of $\mathcal{F}$.
(iii) $|\mathcal{F}| = 2^{o(n)}$.
(iv) Every member of $\mathcal{F}$ has size at most $(1/2 + o(1))n$.

We refer to the sets in $\mathcal{F}$ as *containers*.

In [2] we used Lemma 2.1 to prove that $f_{\max}(n) = 2^{(1+o(1))n/4}$. Indeed, we showed that every $F \in \mathcal{F}$ contains at most $2^{(1+o(1))n/4}$ maximal sum-free subsets of $[n]$ which by (ii) and (iii) yields the desired result. To obtain an exact bound on $f_{\max}(n)$ it is not sufficient to give a tight general bound on the number of maximal sum-free subsets of $[n]$ that lie in a container $F \in \mathcal{F}$. Indeed, such an $F \in \mathcal{F}$ could contain $O(2^{n/4})$ maximal sum-free subsets of $[n]$, and thus together with (iii) this still gives an error term in the exponent. In general, since containers may overlap, applications of container results may lead to 'over-counting'.

We therefore need to count the number of maximal sum-free subsets of $[n]$ in a more refined way. To explain our method, we first need to describe the constructions which imply that $f_{\max}(n) \geq 2^{\lfloor n/4 \rfloor}$.

2.2 Lower bound constructions

The following construction of Cameron and Erdős [6] implies that $f_{\max}(n) \geq 2^{\lfloor n/4 \rfloor}$. Let $n \in \mathbb{N}$ and let $m = n$ or $m = n-1$, whichever is even. Let S consist of m together with precisely one number from each pair $\{x, m-x\}$ for odd $x < m/2$. Then S is sum-free. Moreover, although S may not be maximal, no further odd numbers less than m can be added, so distinct S lie in distinct maximal sum-free subsets of $[n]$.

The following construction from [2] also yields the same lower bound on $f_{\max}(n)$. Suppose that $4|n$ and set $I_1 := \{n/2+1, \dots, 3n/4\}$ and $I_2 := \{3n/4+1, \dots, n\}$. First choose the element $n/4$ and a set $S' \subseteq I_2$. Then for every $x \in I_2 \setminus S'$, choose $x - n/4 \in I_1$. The resulting set S is sum-free but may not be maximal. However, no further element in I_2 can be added, thus distinct S lie in distinct maximal sum-free sets in $[n]$. There are $2^{|I_2|} = 2^{n/4}$ ways to choose S.

2.3 Counting maximal sum-free sets

The following result provides structural information about the containers $F \in \mathcal{F}$. Lemma 2.2 is implicitly stated in [2] and was essentially proven in [10]. It is an immediate consequence of a result of Deshouillers, Freiman, Sós and Temkin [7] on the structure of sum-free sets and a removal lemma of Green [11]. Here O denotes the set of odd numbers in $[n]$.

Lemma 2.2. *If $F \subseteq [n]$ has $o(n^2)$ Schur triples then either*
(a) $|F| \leq 0.47n$;
or one of the following holds for some $-o(1) \leq \gamma = \gamma(n) \leq 0.03$:
(b) $|F| = \left(\frac{1}{2} - \gamma\right) n$ and $F = A \cup B$ where $|A| = o(n)$ and $B \subseteq [(1/2-\gamma)n, n]$ is sum-free;
(c) $|F| = \left(\frac{1}{2} - \gamma\right) n$ and $F = A \cup B$ where $|A| = o(n)$ and $B \subseteq O$.

The crucial idea in the proof of Theorem 1.1 is that we show 'most' of the maximal sum-free subsets of $[n]$ 'look like' the examples given in Section 2.2: We first show that containers of type (a) house only a small (at most $2^{0.249n}$) number of maximal sum-free subsets of $[n]$ (see Lemma 4.3). For type (b) containers we split the argument into two parts. More precisely, we count the number of maximal sum-free subsets S of $[n]$ with the property that (i) the smallest element of S is $n/4 \pm o(n)$ and (ii) the second smallest element of S is at least $n/2 - o(n)$. (For this we use a direct argument rather than counting such sets within the containers.) We then show that the number of maximal sum-free subsets of $[n]$ that lie in type (b) containers but that fail to satisfy one of (i) and (ii) is small ($o(2^{n/4})$). We use a similar idea for type (c) containers. Indeed, we show directly that the number of maximal sum-free subsets of $[n]$ that contain *at most* one even number is $O(2^{n/4})$. We then show that the number of maximal sum-free subsets of $[n]$ that lie in type (c) containers and which contain two or more even numbers is small ($o(2^{n/4})$).

In each of our cases, we give an upper bound on the number of maximal sum-free sets in a container by counting the number of maximal independent sets in various auxiliary graphs. (Similar techniques were used in [19, 2], and in the graph setting in [4].) In Section 3.3 we collect together a number of results that are useful for this.

3 Notation and preliminaries

3.1 Notation

For a set $F \subseteq [n]$, denote by $\mathrm{MSF}(F)$ the set of all maximal sum-free subsets of $[n]$ that are contained in F and let $f_{\max}(F) := |\mathrm{MSF}(F)|$. Also, denote by $\min(F)$ and $\max(F)$ the minimum and the maximum element of F respectively. Let $\min_2(F)$ denote the second smallest element of F. Denote by E the set of all even and by O the set of all odd numbers in $[n]$. Given sets A, B, we let $A + B := \{a + b : a \in A, b \in B\}$. We say a real valued function $f(n)$ is exponentially smaller than another real valued function $g(n)$ if there exists a constant $\varepsilon > 0$ such that $f(n) \leq g(n)/2^{\varepsilon n}$ for n sufficiently large. We use log to denote the logarithm function of base 2.

Throughout, all graphs considered are simple unless stated otherwise. We say that G is a *graph possibly with loops* if G can be obtained from a simple graph by adding at most one loop at each vertex. We write $e(G)$ for the number of edges in G. Given a vertex x in G, we write $\deg_G(x)$ for the *degree* of x in G. Note that a loop at x contributes two to the degree of x. We write $\delta(G)$ for the *minimum degree* and $\Delta(G)$ for the *maximum degree* of G. Denote by $G[T]$ the induced subgraph of G on the vertex set T and $G \setminus T$ the induced subgraph of G on the vertex set $V(G) \setminus T$. Given $x \in V(G)$, we write $N_G(x)$ for the *neighourhood of x in G*. Given $S \subseteq V(G)$, we write $N_G(S)$ for the set of vertices $y \in V(G)$ such that $xy \in E(G)$ for some $x \in S$.

We write C_m for the cycle, and P_m for the path on m vertices. Given graphs G and H we write $G \square H$ for the *cartesian product graph*. So $G \square H$ has vertex set $V(G) \times V(H)$ and (x, y) and (x', y') are adjacent in $G \square H$ if (i) $x = x'$ and y and y' are adjacent in H or (ii) $y = y'$ and x and x' are adjacent in G.

Throughout the paper we omit floors and ceilings where the argument is unaffected. We write $0 < \alpha \ll \beta \ll \gamma$ to mean that we can choose the constants α, β, γ from right to left. More precisely, there are increasing functions f and g such that, given γ, whenever we choose some $\beta \leq f(\gamma)$ and $\alpha \leq g(\beta)$, all calculations needed in our proof are valid. Hierarchies of other lengths are defined in the obvious way.

3.2 The number of sets with small sumset

We need the following lemma of Green and Morris [12], which bounds the number of sets with small sumset.

Lemma 3.1. *Fix $\delta > 0$ and $R > 0$. Then the following hold for all integers $s \geq s_0(\delta, R)$. For any $D \in \mathbb{N}$ there are at most*

$$2^{\delta s} \binom{\frac{1}{2}Rs}{s} D^{\lfloor R+\delta \rfloor}$$

sets $S \subseteq [D]$ with $|S| = s$ and $|S + S| \leq R|S|$.

3.3 Maximal independent sets in graphs

In this section we collect together results on the number of maximal independent sets in a graph. Let $\mathrm{MIS}(G)$ denote the number of maximal independent sets in a graph G.

Moon and Moser [16] showed that for any simple graph G, $\mathrm{MIS}(G) \leq 3^{|G|/3}$. When a graph is triangle-free, this bound can be improved significantly: A result of Hujter and Tuza [14] states that for any triangle-free graph G,

$$\mathrm{MIS}(G) \leq 2^{|G|/2}. \tag{1}$$

The next result implies that the bound given in (1) can be further lowered if G is additionally not too sparse.

Lemma 3.2. *Let $n, D \in \mathbb{N}$ and $k \in \mathbb{R}$. Suppose that G is a triangle-free graph on n vertices with $\Delta(G) \leq D$ and $e(G) \geq n/2 + k$. Then*

$$\mathrm{MIS}(G) \leq 2^{n/2 - k/(100D^2)}.$$

The following result for 'almost triangle-free' graphs follows from Lemma 3.2.

Corollary 3.3. *Let $n, D \in \mathbb{N}$ and $k \in \mathbb{R}$. Suppose that G is a graph and T is a set such that $G' := G \setminus T$ is triangle-free. Suppose that $\Delta(G) \leq D$, $|G'| = n$ and $e(G') \geq n/2 + k$. Then*

$$\mathrm{MIS}(G) \leq 2^{n/2 - k/(100D^2) + 101|T|/100}.$$

We defer the proofs of Lemma 3.2 and Corollary 3.3 to the appendix.

The following result gives an improvement on the Moon–Moser bound for graphs that are not too sparse, almost regular and of large minimum degree. (The result is proven as equation (3) in [2].)

Lemma 3.4 ([2]). *Let $k \geq 1$ and let G be a graph on n vertices possibly with loops. Suppose that $\Delta(G) \leq k\delta(G)$ and set $b := \sqrt{\delta(G)}$. Then*

$$\mathrm{MIS}(G) \leq \sum_{0 \leq i \leq n/b} \binom{n}{i} 3^{\left(\frac{k}{k+1}\right)\frac{n}{3} + \frac{2n}{3b}}.$$

Fact 3.5. *Suppose that G' is a (simple) graph. If G is a graph obtained from G' by adding loops at some vertices $x \in V(G')$ then*

$$\mathrm{MIS}(G) \leq \mathrm{MIS}(G').$$

The following lemma from [1] gives an improvement on (1) when G additionally contains many vertex disjoint P_3s. Its proof is similar to that of Lemma 3.2.

Lemma 3.6 ([1]). *Let G be an n-vertex triangle-free graph, possibly with loops. If G contains k vertex-disjoint P_3s, then*

$$\mathrm{MIS}(G) \leq 2^{\frac{n}{2} - \frac{k}{25}}.$$

4 Proof of Theorem 1.1

Let $1 \leq i \leq 4$ and $0 < \eta < 1$. To prove Theorem 1.1, we must show that there is a constant C_i (dependent only on i) such that if n is sufficiently large and $n \equiv i \mod 4$ then

$$(C_i - \eta)2^{n/4} \leq f_{\max}(n) \leq (C_i + \eta)2^{n/4}. \tag{2}$$

Given $\eta > 0$ and sufficiently large n with $n \equiv i \mod 4$, define constants $\alpha, \delta, \varepsilon > 0$ so that

$$0 < 1/n \ll \alpha \ll \delta \ll \varepsilon \ll \eta < 1. \tag{3}$$

Let $\mathcal{F}$ be the family of containers obtained from Lemma 2.1. Since n is sufficiently large, Lemma 2.2 implies that $|\mathcal{F}| \leq 2^{\alpha n}$ and for every $F \in \mathcal{F}$ either

(a) $|F| \leq 0.47n$;

or one of the following holds for some $-\alpha \leq \gamma = \gamma(n) \leq 0.03$:

(b) $|F| = \left(\frac{1}{2} - \gamma\right) n$ and $F = A \cup B$ where $|A| \leq \alpha n$ and $B \subseteq [(1/2-\gamma)n, n]$ is sum-free;

(c) $|F| = \left(\frac{1}{2} - \gamma\right) n$ and $F = A \cup B$ where $|A| \leq \alpha n$ and $B \subseteq O$.

Throughout the rest of the paper we refer to such containers as type (a), type (b) and type (c), respectively.

For any subsets $B, S \subseteq [n]$, let $L_S[B]$ be the *link graph of S on B* defined as follows. The vertex set of $L_S[B]$ is B. The edge set of $L_S[B]$ consists of the following two types of edges:
(i) Two vertices x and y are adjacent if there exists an element $z \in S$ such that $\{x, y, z\}$ forms a Schur triple;
(ii) There is a loop at a vertex x if $\{x, x, z\}$ forms a Schur triple for some $z \in S$ or if $\{x, z, z'\}$ forms a Schur triple for some $z, z' \in S$.

The following simple lemma from [2] will be applied in many cases throughout the proof.

Lemma 4.1 ([2])**.** *Suppose that B and S are both sum-free subsets of $[n]$. If $I \subseteq B$ is such that $S \cup I$ is a maximal sum-free subset of $[n]$, then I is a maximal independent set in $G := L_S[B]$.*

The next lemma will allow us to apply (1) to certain link graphs.

Lemma 4.2. *Suppose that $B, S \subseteq [n]$ such that S is sum-free and* $\max(S) < \min(B)$. *Then $G := L_S[B]$ is triangle-free.*

Proof. Suppose to the contrary that $z > y > x > \max(S)$ form a triangle in G. Then there exists $a, b, c \in S$ such that $z - y = a, y - x = b$ and $z - x = c$, which implies $a + b = c$ with $a, b, c \in S$. This is a contradiction to S being sum-free. $\square$

In the proof we will use the simple fact that if $S \subseteq T \subseteq [n]$ then

$$f_{\max}(S) \leq f_{\max}(T). \tag{4}$$

The following lemma is a slightly stronger form of Lemma 3.2 from [2], which deals with containers of 'small' size. The proof is exactly the same as in [2].

Lemma 4.3. *If $F \in \mathcal{F}$ has size at most $0.47n$, then $f_{\max}(F) \leq 2^{0.249n}$.*

Thus, to show that (2) holds it suffices to show that there is a constant C_i such that in total, type (b) and (c) containers house $(C_i \pm \eta/2)2^{n/4}$ maximal sum-free subsets of $[n]$. In Section 4.1 we deal with containers of type (b) and in Section 4.2 we deal with containers of type (c).

4.1 Type (b) containers

The following lemma allows us to restrict our attention to type (b) containers that have at most εn elements from $[n/2]$.

Lemma 4.4. *Let $F \in \mathcal{F}$ be a container of type (b) so that $|F \cap [n/2]| \geq \varepsilon n$. Then $f_{\max}(F) \leq 2^{(1/4-\delta)n}$.*

Proof. Define $c \geq \varepsilon$ so that $|F \cap [n/2]| = cn$. Since F is of type (b), $F = A \cup B$ where $|A| \leq \alpha n$ and B is sum-free where $\min(B) \geq 0.47n$. Therefore $cn \leq (0.03 + \alpha)n$.

As $|F \cap [n/2]| = cn$, $|B \cap [0.47n, n/2]| \geq (c-\alpha)n$ and so trivially $|(B+B) \cap [0.94n, n]| \geq (2c-4\alpha)n$. Therefore, since B is sum-free, F is missing at least $(2c-4\alpha)n$ numbers from $[0.94n, n]$. Partition $F = F_1 \cup F_2$ where $F_1 := F \cap [n/2]$ and $F_2 := F \setminus F_1$. Note that $|F_2| \leq (1/2 - 2c + 4\alpha)n$.

The following observation is a key idea for the proof of this lemma. Every maximal sum-free subset of $[n]$ in F can be built in the following two steps. First, fix an arbitrary sum-free set $S \subseteq F_1$. Next, extend S in F_2 to a maximal one. Since $|F_1| = cn$, there are at most 2^{cn} ways to pick S. By Lemma 4.1, the number of choices for the second step is at most the number of maximal independent sets I in $L_S[F_2]$.

Claim 4.5. *There are at most $2^{(1/4-\varepsilon/20)n}$ maximal sum-free subsets M of $[n]$ in F such that $|M \cap F_1| \leq cn/4$.*

Proof. Choose an arbitrary sum-free set $S \subseteq F_1$ such that $|S| \leq cn/4$ (there are at most $cn\binom{cn}{cn/4}/4$ choices for S). By Lemma 4.2, $L := L_S[F_2]$ is triangle-free. So $\mathrm{MIS}(L) \leq 2^{|F_2|/2} \leq 2^{(1/4-c+2\alpha)n}$ by (1). Thus, the number of maximal sum-free subsets of $[n]$ in F with at most $cn/4$ elements from F_1 is at most

$$\frac{cn}{4}\binom{cn}{\frac{cn}{4}} \cdot 2^{(1/4-c+2\alpha)n} \leq 2^{(1/4-c/10+2\alpha)n} \leq 2^{(1/4-\varepsilon/20)n},$$

where the last inequality follows since $\alpha \ll \varepsilon \leq c$. □

Let $S \subseteq F_1$ be sum-free such that $|S| > cn/4$. Claim 4.5 together with our earlier observation implies that to prove the lemma it suffices to show that $\mathrm{MIS}(L_S[F_2]) \leq 2^{(1/4-c-2\delta)n}$.

By Lemma 4.2, $L_S[F_2]$ is triangle-free. We may assume that F is missing at most $(2c+4\delta)n$ numbers from $[0.94n, n]$. Indeed, otherwise by (1), $\mathrm{MIS}(L_S[F_2]) \leq 2^{(1/4-c-2\delta)n}$, as required.

Claim 4.6. *We may assume that $(2c - 4\alpha)n \leq |[n/2+1, n] \setminus F| \leq (2c+9\delta)n$.*

Proof. Since we already know that $(2c-4\alpha)n \leq |[0.94n, n] \setminus F| \leq (2c+4\delta)n$, to prove the claim we only need to prove that F is missing at most $5\delta n$ elements from $[0.5n, 0.94n]$. Suppose to the contrary that F is missing at least $5\delta n$ numbers from $[0.5n, 0.94n]$. Then $|F_2| \leq (1/2 - 2c + 4\alpha - 5\delta)n \leq (1/2 - 2c - 4\delta)n$ and so by (1), $\mathrm{MIS}(L_S[F_2]) \leq 2^{(1/4-c-2\delta)n}$. □

Claim 4.7. *Set $m := \min(S)$. Suppose that $m < (1/4 - 2c)n$ or $m > (1/4+\varepsilon)n$. Then $\mathrm{MIS}(L_S[F_2]) \leq 2^{(1/4-c-2\delta)n}$.*

Proof. Suppose that $m > (1/4+\varepsilon)n$. Then in $L := L_S[F_2]$ a vertex $x \in [(3/4-\varepsilon)n, (3/4+\varepsilon)n] =: N$ is either isolated or adjacent only to itself. Thus $\mathrm{MIS}(L) = \mathrm{MIS}(L')$ where $L' := L \setminus N$. Recall that $(2c-4\alpha)n \leq |[0.94n, n] \setminus F|$. Hence, (1) implies that, $\mathrm{MIS}(L) \leq 2^{(1/4-c+2\alpha-\varepsilon)n} \leq 2^{(1/4-c-2\delta)n}$.

Now suppose that $m < (1/4 - 2c)n$. Then $L := L_S[F_2]$ contains at least $100\delta n$ vertex-disjoint copies of P_3. Indeed, consider the set of all P_3s with vertex set $\{n/2+i, n/2+m+i, n/2+2m+i\}$ for all $1 \leq i \leq n/2 - 2m$. Since $m \leq (1/4 - 2c)n$, we have at least $n/2 - 2m \geq 4cn$ such P_3s. By Claim 4.6, at most $(2c + 9\delta)n$ elements from $[n/2 + 1, n]$ are not in F. Hence, L contains at least $(2c - 9\delta)n \geq 700\delta n$ of these copies of P_3. Note that these copies of P_3 may not be vertex-disjoint, but given one of these copies P of P_3, there are at most 6 copies of P_3 of this type that intersect P in L. So L contains a collection of $100\delta n$ vertex-disjoint copies of P_3. Using Lemma 3.6, we have $\mathrm{MIS}(L) \leq 2^{(1/4-c+2\alpha)n-4\delta n} \leq 2^{(1/4-c-2\delta)n}$. □

By Claim 4.7 we may now assume that $(1/4 - 2c)n \leq m \leq (1/4 + \varepsilon)n$.

Claim 4.8. *Set* $b := \min_2(S)$*. If* $b \leq (1/2 - 4c)n$ *then* $\mathrm{MIS}(L_S[F_2]) \leq 2^{(1/4-c-2\delta)n}$.

Proof. We claim that $L := L_S[F_2]$ contains at least $100\delta n$ vertex-disjoint copies of P_3. Consider the set of all P_3s with vertex set $\{n/2 + i, n/2 + b + i, n/2 + b - m + i\}$ for all $1 \leq i \leq n/2 - b$. Since $b \leq n/2 - 4cn$, we have at least $n/2 - b \geq 4cn$ such P_3s. Note that F might be missing up to $(2c+9\delta)n$ elements from $[n/2+1, n]$. Hence, L contains at least $(2c-9\delta)n \geq 700\delta n$ of these copies of P_3. Note that these copies of P_3 may not be vertex-disjoint, but given one of these copies P of P_3, there are at most 6 copies of P_3 of this type that intersect P in L. So L contains a collection of $100\delta n$ vertex-disjoint copies of P_3. Hence, Lemma 3.6 implies that $\mathrm{MIS}(L_S[F_2]) \leq 2^{(1/4-c-2\delta)n}$. □

So now we may assume that $|S| > cn/4$, $(1/4 - 2c)n \leq m \leq (1/4 + \varepsilon)n$ and $b \geq (1/2 - 4c)n$. Thus, at least $cn/4$ elements from $[(3/4 - 6c)n, (3/4 + \varepsilon)n]$ lie in $S + m$. Every element of $S + m$ is either missing from F_2 or has a loop in $L_S[F_2]$. Recall that F_2 is missing $(2c - 4\alpha)n$ elements from $[0.94n, n]$. Thus, altogether at least $2cn - 4\alpha n + cn/4 \geq 2cn + 4\delta n$ elements from $[n/2 + 1, n]$ are either missing from F_2 or have a loop in $L_S[F_2]$. Hence, we have,

$$\mathrm{MIS}(L_S[F_2]) \leq 2^{(1/4-c-2\delta)n}.$$

□

Lemma 4.9. *Let* $F \in \mathcal{F}$ *be a container of type (b) so that* $|F \cap [n/2]| \leq \varepsilon n$*. Let* $f^*_{\max}(F)$ *denote the number of maximal sum-free subsets* M *of* $[n]$ *in* F *that satisfy at least one of the following properties:*

(i) $\min(M) > (1/4 + 2\varepsilon)n$ *or* $\min(M) < (1/4 - 175\varepsilon)n$*;*

(ii) $\min_2(M) \leq (1/2 - 350\varepsilon)n$*.*

Then $f^*_{\max}(F) \leq 2^{(1/4-\varepsilon)n}$*.*

Proof. Since F is of type (b), $F = A \cup B$ for some A, B where $|A| \leq \alpha n$ and B is sum-free where $\min(B) \geq 0.47n$. Partition $F = F_1 \cup F_2$ where $F_1 := F \cap [n/2]$ and $F_2 := F \setminus F_1$. So $|F_1| \leq \varepsilon n$ by the hypothesis of the lemma. By (4) we may assume that $F_2 = [n/2 + 1, n]$.

Every maximal sum-free subset of $[n]$ in F that satisfies (i) or (ii) can be built in the following two steps. First, fix a sum-free set $S \subseteq F_1$. Next, extend S in F_2 to a maximal one. To give an upper bound on the sets M satisfying (i) we choose $S \subseteq F_1$ where $m := \min(S)$ is such that $m > (1/4 + 2\varepsilon)n$ or $m < (1/4 - 175\varepsilon)n$ (there are at most $2^{|F_1|} \leq 2^{\varepsilon n}$ choices for S). Then by arguing similarly to Claim 4.7 we have that $\mathrm{MIS}(L_S[F_2]) \leq 2^{(1/4-2\varepsilon)n}$.

To give an upper bound on the sets M satisfying (ii) we choose $S \subseteq F_1$ where $b := \min_2(S)$ is such that $b \leq n/2 - 350\varepsilon n$ (there are at most $2^{|F_1|} \leq 2^{\varepsilon n}$ choices for S). Then by arguing similarly to Claim 4.8 we have that $\mathrm{MIS}(L_S[F_2]) \leq 2^{(1/4-2\varepsilon)n}$.

Altogether, this implies that $f^*_{\max}(F) \leq 2^{(1/4-\varepsilon)n}$ as desired. $\square$

Throughout this subsection, given a maximal sum-free set M we write $m := \min(M)$ and $b := \min_2(M)$ and define $S := (M \cap [n/2]) \setminus \{m\}$. Lemmas 4.4 and 4.9 imply that, to count the number of maximal sum-free subsets of $[n]$ lying in type (b) containers, it now suffices to count the number of maximal sum-free sets M with the following structure:

(α) $m \in [(1/4 - 175\varepsilon)n, (1/4 + 175\varepsilon)n]$.

(β) $b \geq (1/2 - 350\varepsilon)n$.

In particular, the next lemma shows that almost all of the maximal sum-free subsets of $[n]$ that satisfy (α) and (β) lie in type (b) containers only.

Lemma 4.10. *There are at most $\varepsilon 2^{n/4}$ maximal sum-free subsets of $[n]$ that satisfy (α) and (β) and that lie in type (a) or (c) containers.*

Proof. By Lemma 4.3, at most $2^{0.249n} \leq \varepsilon 2^{n/4}/2$ such maximal sum-free subsets of $[n]$ lie in type (a) containers.

Suppose that M is a maximal sum-free subset of $[n]$ that satisfies (α) and (β) and lies in a type (c) container F. Thus, $F = A \cup B$ where $|A| \leq \alpha n$ and $B \subseteq O$. Define $F' := B \cap [n/2 - 350\varepsilon n, n]$. So, $|F'| \leq (1/4 + 175\varepsilon)n$. By Lemma 4.1, $M = I \cup S$ where $\min(S) = m$ for some $m \in [(1/4 - 175\varepsilon)n, (1/4 + 175\varepsilon)n]$, $(S \setminus \{m\}) \subseteq A$ and I is a maximal independent set in $G := L_S[F']$. By the Moon–Moser bound,

$$\mathrm{MIS}(G) \leq 3^{(1/12+60\varepsilon)n} \leq 2^{(1/4-\varepsilon)n}.$$

In total, there are at most $2^{\alpha n}$ choices for F, at most $350\varepsilon n$ choices for m and at most $2^{\alpha n}$ choices for $S \setminus \{m\}$. Thus, there are at most

$$2^{\alpha n} \times 350\varepsilon n \times 2^{\alpha n} \times 2^{n/4-\varepsilon n} \leq \varepsilon 2^{n/4}/2$$

maximal sum-free subsets of $[n]$ that satisfy (α) and (β) and that lie in type (c) containers, as desired. $\square$

For the rest of this subsection, we focus on counting the maximal sum-free sets that satisfy (α) and (β). Fix m, b such that $m \in [(1/4 - 175\varepsilon)n, (1/4 + 175\varepsilon)n]$ and $b \geq (1/2 - 350\varepsilon)n$. Define $t := |m - n/4|$ and $D := n/2 - b$, so $t, D \leq 350\varepsilon n$. (Notice that if $b > n/2$, then D is negative.) Let $S \subseteq [b, n/2]$ such that $b \in S$, $S \cup \{m\}$ is sum-free and set $s := |S| \leq D$. In the case when $b > n/2$, we define $S := \emptyset$.

Denote by $L := L(n, m, S)$ the link graph of $S \cup \{m\}$ on vertex set $[n/2+1, n]$. So L is triangle-free by Lemma 4.2. We will need the following two bounds on the number of maximal independent sets in L.

Lemma 4.11. *We have the following two bounds on* $\mathrm{MIS}(L)$.

(i) $\mathrm{MIS}(L) \leq 2^{n/4-D/25}$*;*

(ii) Let R be defined so that $|S+S| = Rs$. Then $\mathrm{MIS}(L) \leq 2^{n/4-(R+1)s/2}$.

Proof. If $D \leq 0$ then (i) follows from (1). So assume $D > 0$. Notice that there are D vertex-disjoint P_3s in L: $\{n/2+i, n+i-D, n+i-D-m\}$ for each $1 \leq i \leq D$. (These paths are vertex-disjoint since $D \leq 350\varepsilon n$ and $m \in [(1/4 - 175\varepsilon)n, (1/4 + 175\varepsilon)n]$.) The bound follows immediately from Lemma 3.6.

For (ii), notice that in L we have loops at all vertices in $S+S$ and $S+m$ (in total $(R+1)s$ vertices). $\mathrm{MIS}(L)=\mathrm{MIS}(L')$ where L' is the graph obtained from L by deleting all the vertices with loops. The bound then follows from (1). □

The following lemma bounds the number of maximal sum-free sets M satisfying (α) and (β) and with b sufficiently bounded away from $n/2$ from above.

Lemma 4.12. *There exists a constant $K=K(\varepsilon)$ such that the number of maximal sum-free sets M in $[n]$ that satisfy (α), (β) and $b\leq n/2-K$ is at most $\varepsilon 2^{n/4}$.*

Proof. Let K be such that $\delta \ll 1/K \ll \varepsilon$. Our first claim implies that there are not too many maximal sum-free subsets of $[n]$ with t or D 'large'.

Claim 4.13. *There are at most $\varepsilon 2^{n/4}/5$ maximal sum-free sets M which satisfy (α) and (β) and with*

(a) $b\leq n/2-K$;

(b) $t\geq 3D$ *or* $D\geq 10^9 s$.

Proof. Fix any m,b such that $m\in[(1/4-175\varepsilon)n,(1/4+175\varepsilon)n]$ and $n/2-350\varepsilon n\leq b\leq n/2-K$. Define t and D as before. Let $S\subseteq[b,n/2]$ such that $b\in S$, $S\cup\{m\}$ is sum-free and set $s:=|S|\leq D$. Define the link graph L as before.

Suppose that $t\geq 3D$. If $m=n/4-t$ then for each i with $D+1\leq i\leq 2t-D$ consider the subgraph H_i of L induced by $\{n/2+i,3n/4+i-t,n+i-2t\}$. Ignoring loops, H_i spans a P_3 component in L and so $\mathrm{MIS}(H_i)\leq 2$. Indeed, since $t,D\leq 350\varepsilon n$ and $\min(S)=b=n/2-D$, the vertex $3n/4+i-t$ has no neighbour in L generated by S. Also, since $n/2+i+b=n+i-D>n$ and $n+i-2t-b=n/2+i-2t+D\leq n/2$, neither $n/2+i$ nor $n+i-2t$ has a neighbour generated by S in L. Recall L and thus $L':=L\setminus\cup_{i=D+1}^{2t-D}H_i$ is triangle-free. Thus by (1) we have

$$\mathrm{MIS}(L)\leq \mathrm{MIS}(L')\cdot\prod_i \mathrm{MIS}(H_i)\leq 2^{[n/2-3(2t-2D)]/2}\cdot 2^{2t-2D}\leq 2^{n/4-(t-D)}\leq 2^{n/4-2t/3}.$$

Otherwise $m=n/4+t$ and then there are $2t$ isolated vertices $\{3n/4-t+1,\dots,3n/4+t\}$ in L. Then by (1), $\mathrm{MIS}(L)\leq 2^{n/4-t}$.

Given fixed t, there are 2 choices for m. There are at most $2^{t/3}$ choices for S so that $D\leq t/3$. Further, fixing S determines b and D. Altogether, this implies that the number of maximal sum-free subsets M of $[n]$ that satisfy (α), (β), (a) and $t\geq 3D$ is at most

$$2\cdot\sum_{t\geq 3D\geq 3K}2^{t/3}\cdot 2^{n/4-2t/3}\leq 2\cdot\sum_{t\geq 3K}2^{n/4-t/3}\leq\frac{\varepsilon}{10}\cdot 2^{n/4},\tag{5}$$

where the last inequality follows since $1/K\ll\varepsilon$ and n is sufficiently large.

Suppose now that $t\leq 3D$ and $D/s\geq 10^9$. For fixed $D\geq K$ there are $3D$ choices for t and so at most $6D\leq 2^{2\log D}$ choices for m. Given fixed D, there are $D=2^{\log D}$ choices for s. For fixed D,s there are $\binom{D}{s}\leq\left(\frac{eD}{s}\right)^s\leq 2^{s\log(eD/s)}$ choices for S. Note that when $D/s\geq 10^9$, $3\log D+s\log(eD/s)\leq D/50$. Together, with Lemma 4.11(i), this implies that the number of maximal sum-free subsets M of $[n]$ that satisfy (α), (β), (a) and with $t \leq 3D$ and $D/s \geq 10^9$ is at most

$$\sum_{D \geq K} 2^{2\log D} \cdot 2^{\log D} \cdot 2^{s \log(eD/s)} \cdot 2^{n/4 - D/25} \leq \sum_{D \geq K} 2^{n/4 - D/50} \leq \frac{\varepsilon}{10} \cdot 2^{n/4}. \tag{6}$$

□

By Claim 4.13, to complete the proof of the lemma it suffices to count the number of maximal sum-free subsets M of $[n]$ that satisfy (α), (β) and

(γ_1) $b \leq n/2 - K$;

(γ_2) $s \geq D/10^9 \geq K/10^9$;

(γ_3) $t < 3D$.

Fix any m, b such that $m \in [(1/4 - 175\varepsilon)n, (1/4 + 175\varepsilon)n]$ and $n/2 - 350\varepsilon n \leq b \leq n/2 - K$. Let $S \subseteq [b, n/2]$ such that $b \in S$, $S \cup \{m\}$ is sum-free and set $s := |S| \leq D$. Define the link graph L as before.

Choose s and D such that $s \geq D/10^9$. For each fixed s there are at most $10^9 s$ choices for D. For a fixed $s \geq D/10^9$, there are at most $6D \leq 10^{10}s \leq 2^{2\log s}$ choices for m so that $t < 3D$ and at most $\binom{10^9 s}{s}$ choices for S. So there are at most

$$10^9 s \cdot 2^{2\log s} \cdot \binom{10^9 s}{s} \leq 10^9 s \cdot 2^{2\log s} \cdot 2^{s \log(e \cdot 10^9)} \leq 2^{49s} \tag{7}$$

choices for the pair S, m given fixed s. Let R be defined so that $|S + S| = Rs$. We now distinguish two cases depending on the size of $S + S$.

The number of maximal sum-free subsets M in $[n]$ that satisfy (α), (β), (γ_1)–(γ_3) and $R \geq 100$ is at most

$$\sum_{s \geq K/10^9} 2^{49s} \cdot 2^{n/4 - 50s} \leq \sum_{s \geq K/10^9} 2^{n/4 - s} \leq \frac{\varepsilon}{10} \cdot 2^{n/4}. \tag{8}$$

(Here we have applied (7) and Lemma 4.11 (ii).)

Let $s_0(1/9, 100)$ be the constant returned from Lemma 3.1. Since we chose K sufficiently large, we have that $s \geq K/10^9 \geq s_0(1/9, 100)$.

Now suppose $R \leq 100$. Then by Lemma 3.1 the number of choices for S is at most

$$2^{s/9} \binom{\frac{1}{2}Rs}{s} D^{\lfloor R + 1/9 \rfloor} \leq 2^{s/9} \cdot 2^{Rs/2} \cdot 2^{4R \log s} \leq 2^{Rs/2 + 2s/9}. \tag{9}$$

Recall that for a fixed s, the number of choices for m is at most $2^{2\log s}$. Together with Lemma 4.11(ii) and (9), we have that the number of maximal sum-free subsets M in $[n]$ that satisfy (α), (β), (γ_1)–(γ_3) and $R \leq 100$ is at most

$$\begin{aligned} & \sum_{s \geq K/10^9} 2^{2\log s} \cdot 2^{Rs/2 + 2s/9} \cdot 2^{n/4 - (R+1)s/2} \leq \sum_{s \geq K/10^9} 2^{n/4 - s/2 + s/3} \\ \leq & \sum_{s \geq K/10^9} 2^{n/4 - s/6} \leq \frac{\varepsilon}{10} \cdot 2^{n/4}. \end{aligned} \tag{10}$$

Thus by Claim 4.13, (8) and (10), we have that the number of maximal sum-free sets that satisfy (α), (β) and $b \leq n/2 - K$ is at most $\varepsilon \cdot 2^{n/4}$. □

The following lemma bounds the number of maximal sum-free sets when t is large.

Lemma 4.14. *There are at most $\varepsilon 2^{n/4}$ maximal sum-free sets in $[n]$ that satisfy (α) and (β) and with $|m-n/4|=t$ and $b=n/2-D$ such that $D\le K$ and $t\ge 50K$.*

Proof. Let us first assume that $m=n/4+t$. If $b\le n/2$ then let $S\subseteq[b,n/2]$ where $b\in S$. Otherwise let $S=\emptyset$. Then in the link graph $L:=L(n,m,S)$, every vertex in $\{3n/4-t+1,3n/4+t\}=:N$ is either isolated or adjacent only to itself. Since $D\le K$, the number of choices for S is at most 2^K. Let $L':=L\setminus N$, then by (1) the number of maximal sum-free sets in this case is at most

$$\sum_{t\ge 50K} 2^K\cdot \mathrm{MIS}(L')\le \sum_{t\ge 50K} 2^K\cdot 2^{n/4-t}\le \varepsilon 2^{n/4}/2.$$

Otherwise, suppose $m=n/4-t$. If $b\le n/2$ then let $S\subseteq[b,n/2]$ where $b\in S$. Otherwise let $S=\emptyset$. The link graph $L:=L(n,m,S)$ contains $2t$ vertex-disjoint P_3s on the vertex set $\{n/2+i,3n/4-t+i,n-2t+i\}$ where $1\le i\le 2t$. Then by Lemma 3.6, the number of maximal sum-free sets in this case is at most

$$\sum_{t\ge 50K} 2^K\cdot \mathrm{MIS}(L)\le \sum_{t\ge 50K} 2^K\cdot 2^{n/4-2t/25}\le \varepsilon 2^{n/4}/2.$$

□

By Lemmas 4.12 and 4.14, we now need only focus on maximal sum-free sets with

$$t,D\le 50K,\quad \textit{i.e.}\quad S\subseteq[n/2-50K,n/2]\quad\text{and}\quad m\in[n/4-50K,n/4+50K],\tag{11}$$

where here D may be negative and $S=\emptyset$. Given any m,S satisfying (11) so that $2m\notin S$, define $C(n,m,S):=\frac{|\mathrm{MIS}(L(n,m,S))|}{2^{n/4}}$. Notice that not every maximal independent set in $L(n,m,S)$ necessarily gives a maximal sum-free set in $[n]$. This happens exactly when a set I is a maximal independent set in both $L(n,m,S)$ and $L(n,m,S^*)$ for some sum-free $S^*\supset S$ such that $S^*\subseteq[n/2]\setminus\{m,2m\}$. Let $\mathcal{I}(n,m,S)$ be the set of all maximal independent sets in $L(n,m,S)$ that do not correspond to maximal sum-free sets in $[n]$. For each $I\in\mathcal{I}(n,m,S)$, define $S^*(I)$ to be the largest sum-free set such that $S\subseteq S^*(I)\subseteq[n/2]\setminus\{m,2m\}$ and I is also a maximal independent set in $L(n,m,S^*(I))$. Further partition $\mathcal{I}(n,m,S):=\mathcal{I}_1(n,m,S)\cup\mathcal{I}_2(n,m,S)$, in which $\mathcal{I}_1(n,m,S)$ consists of all those $I\in\mathcal{I}(n,m,S)$ with $S^*(I)\subseteq[n/2-50K,n/2]$. Let $\mathrm{MSF}(n,m,S)$ be the number of maximal sum-free sets M in $[n]$ that satisfy (α) and (β) with $\min(M)=m$ and $(M\cap[n/2])\setminus\{m\}=S$. For $i=1,2$, further define $C_i(n,m,S):=\frac{|\mathcal{I}_i(n,m,S)|}{2^{n/4}}$. Then clearly by the definition we have

$$\mathrm{MSF}(n,m,S)=[C(n,m,S)-C_1(n,m,S)-C_2(n,m,S)]2^{n/4}.$$

Notice that every set $I\in\mathcal{I}_2(n,m,S)$ is a maximal independent set in $L(n,m,S^*(I))$ with $\min(S^*(I))\le n/2-50K$, it then follows from Lemma 4.12 that $\sum_{m,S:\ t,D\le 50K}C_2(n,m,S)\le\varepsilon$.

Thus, the number of maximal sum-free sets M in $[n]$ that satisfy (α) and (β) is at least

$$\begin{aligned}\sum_{m,S:\ t,D\le 50K}\mathrm{MSF}(n,m,S) &= \sum_{m,S:\ t,D\le 50K}[C(n,m,S)-C_1(n,m,S)-C_2(n,m,S)]2^{n/4}\\ &\ge \sum_{m,S:\ t,D\le 50K}[C(n,m,S)-C_1(n,m,S)]2^{n/4}-\varepsilon 2^{n/4}.\end{aligned}$$

On the other hand, by Lemmas 4.12 and 4.14, the number of maximal sum-free sets M in $[n]$ that satisfy (α) and (β) is at most

$$\begin{aligned}\sum_{m,S}\mathrm{MSF}(n,m,S) &= \sum_{m,S\colon\ t,D\le 50K}\mathrm{MSF}(n,m,S)+\sum_{m,S\colon\ \max\{t,D\}>50K}\mathrm{MSF}(n,m,S)\\ &\le \sum_{m,S\colon\ t,D\le 50K}[C(n,m,S)-C_1(n,m,S)]2^{n/4}+2\varepsilon 2^{n/4}.\end{aligned}$$

By defining $C(n) := \sum_{m,S\colon\ t,D\le 50K}[C(n,m,S)-C_1(n,m,S)]$, together with Lemmas 4.4, 4.9 and 4.10, we have that the number of maximal sum-free sets of $[n]$ contained in type (b) containers is $(C(n)\pm 4\varepsilon)2^{n/4}$.

We now proceed to prove that for any $n'\equiv n \mod 4$, $C(n')=C(n)$. We need the following lemma, which roughly states that for any "fixed" choice of m and S, the link graphs on $[n/2+1,n]$ and $[n'/2+1,n']$ differ by a component consisting of an induced matching of size $(n'-n)/4$. To be formal, fix $t\in[-50K,50K]$, $S_0\subseteq[50K]$ and $\ell\in\mathbb{N}$. Define

$$n':=n+4\ell,\quad m:=n/4-t,\quad m':=n'/4-t,\quad S:=n/2-S_0,\quad S':=n'/2-S_0. \tag{12}$$

The proof of the following lemma for the case $m=n/4+t$ and $m'=n'/4+t$ is almost identical except only simpler, we omit it here.

Lemma 4.15. *Let n',m,m',S,S' be given as in* (12)*. Then $L(n',m',S')$ is isomorphic to the disjoint union of $L(n,m,S)$ and a matching of size ℓ.*

Proof. Let $I_1:=[n'/2+200K+1,3n'/4-200K+t]$ and $I_2:=[3n'/4+200K+1-t,n'-200K]$. Notice first that the induced subgraph of $L':=L(n',m',S')$ on $I_1\cup I_2$ is a matching: $\{n'/2+200K+1,3n'/4+200K+1-t\},\dots,\{3n'/4-200K+t,n'-200K\}$. Let $\mathcal{M}$ be the first ℓ matching edges in $L'[I_1\cup I_2]$, i.e. $\{n'/2+200K+1,3n'/4+200K+1-t\},\dots,\{n'/2+200K+\ell,3n'/4+200K+\ell-t\}$. Define $L'':=L'\setminus\mathcal{M}$. It is a straightforward but tedious task to see that L'' is isomorphic to $L:=L(n,m,S)$. We give here only the mapping $f:V(L)\to V(L'')$ that defines an isomorphism:

- $[n/2+1,n/2+200K] \quad\to\quad [n'/2+1,n'/2+200K]$;
- $[n/2+200K+1,3n/4+200K-t] \quad\to\quad [n'/2+200K+\ell+1,3n'/4+200K-t]$;
- $[3n/4+200K-t+1,n-200K] \quad\to\quad [3n'/4+200K+\ell-t+1,n'-200K]$;
- $[n-200K+1,n] \quad\to\quad [n'-200K+1,n']$.

$\square$

Fix n',m,m',S,S' satisfying (11) and (12). By the definition of $C(n)$, to show that $C(n)=C(n')$, it suffices to show that $C(n,m,S)=C(n',m',S')$ and $C_1(n,m,S)=C_1(n,m,S)$. Let $\mathcal{M}$ and f be the matching of size ℓ and the mapping from Lemma 4.15. As an immediate consequence of Lemma 4.15, we have

$$C(n',m',S')=\frac{|\mathrm{MIS}(L(n',m',S'))|}{2^{n'/4}}=\frac{|\mathrm{MIS}(L(n,m,S))|\cdot\mathrm{MIS}(\mathcal{M})}{2^{n/4}\cdot 2^{\ell}}=C(n,m,S).$$

As for $C_1(n,m,S)$, it suffices to show that every $I\in\mathcal{I}_1(n,m,S)$ corresponds to precisely 2^{ℓ} sets in $\mathcal{I}_1(n',m',S')$. Fix an arbitrary $I\in\mathcal{I}_1(n,m,S)$ and recall that $S\subseteq S^*(I)\subseteq[n/2-50K,n/2]$. Let

S^{**} be the "counterpart" (as in S' to S in (12)) of $S^*(I)$ in $[n']$, i.e. $S^{**} := n'/2 - (n/2 - S^*(I)) \subseteq [n'/2 - 50K, n'/2]$. By the definition of $\mathcal{M}$, edges generated by $S', S^{**} \subseteq [n'/2 - 50K, n'/2]$ on $[n'/2, n']$ are not incident to any vertex in $\mathcal{M}$. Hence by adding any maximal independent set of $\mathcal{M}$ to $f(I)$, we obtain $|\mathrm{MIS}(\mathcal{M})| = 2^\ell$ many maximal independent sets I' in $\mathcal{I}_1(n', m', S')$ with $S^*(I') = S^{**}$ as required. We have concluded the following main result of this subsection.

Lemma 4.16. *For each $1 \le i \le 4$, there is a constant D_i such that, if $n \equiv i \mod 4$ then the number of maximal sum-free subsets of $[n]$ in type (b) containers is $(D_i \pm 4\varepsilon)2^{n/4}$.*

4.2 Type (c) containers

The next result implies that the number of maximal sum-free subsets of $[n]$ that contain at least two even numbers and that lie in type (c) containers is 'small'.

Lemma 4.17. *Let $F \in \mathcal{F}$ be a container of type (c). Then F contains at most $2^{(1/4-\varepsilon/2)n}$ maximal sum-free subsets of $[n]$ that contain at least two even numbers.*

Proof. Let $F \in \mathcal{F}$ be as in the statement of the lemma. Let K be a sufficiently large constant so that

$$\sum_{0 \le i \le n/K} \binom{n}{i} 3^{\frac{5n}{36} + \frac{n}{3K}} \le 2^{0.249n}. \tag{13}$$

Since $1/n \ll \varepsilon \ll 1$, we have that $\varepsilon \ll 1/K^2$. By (4), we may assume that $F = O \cup C$ with $C \subseteq E$ and $|C| \le \alpha n$. Similarly as before, every maximal sum-free subset of $[n]$ in F can be built from choosing a sum-free set $S \subseteq C$ (at most $2^{|C|} \le 2^{\alpha n}$ choices) and extending S in O to a maximal one. Fix an arbitrary sum-free set S in C where $|S| \ge 2$ and let $G := L_S[O]$ be the link graph of S on vertex set O. Since O is sum-free and $\alpha \ll \varepsilon$, Lemma 4.1 implies that, to prove the lemma, it suffices to show that $\mathrm{MIS}(G) \le 2^{(1/4-\varepsilon)n}$. We will achieve this in two cases depending on the size of S.

Case 1: $|S| \ge 2K^2$.

In this case, we will show that G is 'not too sparse and almost regular'. Then we apply Lemma 3.4.

We first show that $\delta(G) \ge |S|/2$ and $\Delta(G) \le 2|S| + 2$, thus $\Delta(G) \le 5\delta(G)$. Let x be any vertex in O. If $s \in S$ such that $s < \max\{x, n-x\}$ then at least one of $x - s$ and $x + s$ is adjacent to x in G. If $s \in S$ such that $s \ge \max\{x, n-x\}$ then $s - x$ is adjacent to x in G. By considering all $s \in S$ this implies that $\deg_G(x) \ge |S|/2$ (we divide by 2 here as an edge xy may arise from two different elements of S). For the upper bound consider $x \in O$. If $xy \in E(G)$ then $y = x + s$, $x - s$ or $s - x$ for some $s \in S$ and only two of these terms are positive. Further, there may be a loop at x in G (contributing 2 to the degree of x in G). Thus, $\deg_G(x) \le 2|S| + 2$, as desired.

Note that $\delta(G)^{1/2} \ge K$. Thus, applying Lemma 3.4 to G with $k = 5$ we obtain that

$$\mathrm{MIS}(G) \le \sum_{0 \le i \le n/K} \binom{n}{i} 3^{\frac{5n}{36} + \frac{n}{3K}} \overset{(13)}{\le} 2^{0.249n}.$$

Case 2: $2 \le |S| \le 2K^2$.

As in Case 1 we have that $\Delta(G) \le 2|S| + 2 \le 5K^2$. Additionally, we need to count triangles in G.

Claim 4.18. *G contains at most $24|S|^3$ triangles.*

The claim is shown in the proof of Lemma 3.4 in [2], so we omit the proof here. Let $T \subseteq V(G)$ such that $|T| \leq 24|S|^3$ and $G \setminus T$ is triangle-free.

Let G_1 denote the graph obtained from G by removing all loops. Given any $x \in O$ and $s \in S$, one of $x-s, s-x$ is adjacent to x in G. In particular, if $2x \neq s$, then one of $x-s, s-x$ is adjacent to x in G_1. Therefore each $s \in S$ gives arise to at least $(|O|-1)/2$ edges in G_1. Given distinct $s, s' \in S$, there is at most one pair $x, y \in O$ such that s, x, y and s', x, y are both Schur triples. Thus, since $|S| \geq 2$, this implies that $e(G_1) \geq |O| - 2$. Set $G' := G_1 \setminus T$. Note that $\Delta(G_1) \leq 5K^2$, $|G'| \leq |O|$ and $e(G') \geq |O| - 2 - |T|5K^2 \geq 3|O|/4$. Thus Corollary 3.3 implies that $\mathrm{MIS}(G_1) \leq 2^{(1/4-\varepsilon)n}$. Fact 3.5 therefore implies that $\mathrm{MIS}(G) \leq 2^{(1/4-\varepsilon)n}$, as desired. □

Note that the argument in Case 2 of Lemma 4.17 immediately implies the following result.

Lemma 4.19. *Given any distinct $x, x' \in E$,*

$$\mathrm{MIS}(L_{\{x,x'\}}[O]) \leq 2^{(1/4-\varepsilon)n}.$$

Given $n \in \mathbb{N}$, let $f'_{\max}(n)$ denote the number of maximal sum-free subsets of $[n]$ that contain precisely one even number. The next result implies that $f'_{\max}(n)$ is approximately equal to the number of maximal independent sets in the link graphs $L_x[O]$ where $x \in E$.

Lemma 4.20.

$$\sum_{x \in E} \mathrm{MIS}(L_x[O]) - 2 \cdot \sum_{x \neq x' \in E} \mathrm{MIS}(L_{\{x,x'\}}[O]) \leq f'_{\max}(n) \leq \sum_{x \in E} \mathrm{MIS}(L_x[O]). \tag{14}$$

In particular,

$$\sum_{x \in E} \mathrm{MIS}(L_x[O]) - 2^{(1/4-\varepsilon/2)n} \leq f'_{\max}(n) \leq \sum_{x \in E} \mathrm{MIS}(L_x[O]). \tag{15}$$

Proof. Given any maximal sum-free subset M of $[n]$ that contains precisely one even number x, $M \setminus \{x\}$ is a maximal independent set in $L_x[O]$. So the upper bound in (14) follows.

Claim 4.21. *Suppose $x \in E$ and S is a maximal independent set in $L_x[O]$. Let M denote the maximal sum-free subset of $[n]$ that contains $S \cup \{x\}$. Then $M \setminus S \subseteq E$.*

Proof. Suppose not. Then there exists $S' \subseteq M$ such that $S \subset S' \subseteq O$. But as M is sum-free, S' is an independent set in $L_x[O]$, a contradiction to the maximality of S. □

Suppose $y \in E$ and S is a maximal independent set in $L_y[O]$. If $S \cup \{y\}$ is not a maximal sum-free subset of $[n]$ then Claim 4.21 implies that there exists $y' \in E \setminus \{y\}$ such that $S \cup \{y, y'\}$ is sum-free. In particular, S is a maximal independent set in $L_{\{y,y'\}}[O]$. In total there are at most

$$2 \cdot \sum_{x \neq x' \in E} \mathrm{MIS}(L_{\{x,x'\}}[O])$$

such pairs S, y. Thus, the lower bound in (14) follows.

The lower bound in (15) follows since, by Lemma 4.19,

$$2 \cdot \sum_{x \neq x' \in E} \mathrm{MIS}(L_{\{x,x'\}}[O]) \leq 2n^2 \cdot 2^{(1/4-\varepsilon)n} \leq 2^{(1/4-\varepsilon/2)n},$$

where the last inequality follows since n is sufficiently large. □

The next result determines $\sum_{x\in E}\mathrm{MIS}(L_x[O])$ asymptotically and thus, together with Lemma 4.20 determines, asymptotically, $f'_{\max}(n)$.

Lemma 4.22. *Given $1 \le i \le 4$, there exists a constant D'_i such that, if $n \equiv i \mod 4$,*

$$(D'_i - \varepsilon)2^{n/4} \le \sum_{x\in E}\mathrm{MIS}(L_x[O]) \le (D'_i + \varepsilon)2^{n/4}.$$

Proof. Suppose that $n \equiv 0 \mod 4$. The proofs for the other cases are essentially identical, so we omit them. Let $2n/3 < m \le n$ be even. Consider $G := L_m[O]$. The edge set of G consists of precisely the following edges:

- An edge between i and $m-i$ for every odd $i < m/2$;
- A loop at $m/2$ if $m/2$ is odd;
- An edge between i and $m+i$ for all odd $i \le n-m < n/3$.

In particular, since $m > 2n/3$, if $i < m/2$ is odd then in G, $m-i$ is only adjacent to i. Altogether this implies that if $m/2$ is even then G is the disjoint union of:

- $(n-m)/2$ copies of P_3;
- A matching containing $(3m-2n)/4$ edges.

In this case $\mathrm{MIS}(G) = 2^{(n-m)/2} \times 2^{(3m-2n)/4} = 2^{m/4}$. If $m/2$ is odd then G is the disjoint union of:

- $(n-m)/2$ copies of P_3;
- A single loop;
- A matching containing $(3m-2n-2)/4$ edges.

In this case $\mathrm{MIS}(G) = 2^{(m-2)/4}$.

Thus,

$$\begin{aligned}\sum_{m\in E\,:\,m>2n/3}\mathrm{MIS}(L_m[O]) &\le \sum_{m=4\,:\,m\equiv 0 \bmod 4}^{n} 2^{m/4} + \sum_{m=2\,:\,m\equiv 2 \bmod 4}^{n} 2^{(m-2)/4}\\ &= \sum_{m=1}^{n/4} 2^m + \sum_{m=0}^{n/4-1} 2^m \le (3+\varepsilon/2)2^{n/4}. \end{aligned} \tag{16}$$

Further,

$$\sum_{m\in E\,:\,m>2n/3}\mathrm{MIS}(L_m[O]) \ge (3-\varepsilon/2)2^{n/4} - \sum_{m=1}^{2n/3} 2^{m/4} \ge (3-\varepsilon)2^{n/4}. \tag{17}$$

Consider $m \in E$ where $m \le 2n/3$ and set $G := L_m[O]$. It is easy to see that G is the disjoint union of paths that contain at least 3 vertices and in the case when $m/2$ is odd, an additional path of length at least 2 which contains a vertex (namely $m/2$) with a loop. Every such graph on $n/2$

vertices contains at least $n/10-1$ vertex-disjoint copies of P_3. Therefore, by Lemma 3.6 we have that

$$\sum_{m\in E\,:\,m\leq 2n/3} \mathrm{MIS}(L_m[O]) \leq n2^{n/4-n/250+1}. \tag{18}$$

Overall, we have that

$$(3-\varepsilon)2^{n/4} \overset{(17)}{\leq} \sum_{x\in E} \mathrm{MIS}(L_x[O]) \overset{(16),(18)}{\leq} (3+\varepsilon/2)2^{n/4} + n2^{n/4-n/250+1} \leq (3+\varepsilon)2^{n/4},$$

as desired. □

We showed that the constant D'_4 in Lemma 4.22 is equal to 3. By following the argument given in the proof, it is easy to see that $D'_1 = 3\cdot 2^{-1/4}$, $D'_2 = 2^{3/2}$ and $D'_3 = 2^{5/4}$.

The next lemma shows that almost all of the maximal sum-free subsets of $[n]$ that contain precisely one even number lie in type (c) containers only.

Lemma 4.23. *There are at most $\varepsilon 2^{n/4}$ maximal sum-free subsets of $[n]$ that contain precisely one even number and that lie in type (a) or (b) containers.*

Proof. By Lemma 4.3, at most $2^{0.249n} \leq \varepsilon 2^{n/4}/2$ such maximal sum-free subsets of $[n]$ lie in type (a) containers.

Suppose that M is a maximal sum-free subset of $[n]$ that lies in a type (b) container F and only contains one even number. Define $F' := F\cap O$. Since F is of type (b), $|F'| \leq (0.53n)/2+\alpha n \leq 0.27n$. By Lemma 4.1, $M = I\cup\{m\}$ where m is even and I is a maximal independent set in $G := L_m[F']$. By the Moon–Moser bound,

$$\mathrm{MIS}(G) \leq 3^{0.09n} \leq 2^{(1/4-\varepsilon)n}.$$

In total, there are at most $2^{\alpha n}$ choices for F and at most $n/2$ choices for m. Thus, there are at most

$$2^{\alpha n} \times \frac{n}{2} \times 2^{n/4-\varepsilon n} \leq \varepsilon 2^{n/4}/2$$

maximal sum-free subsets of $[n]$ that that lie in type (b) containers and only contain one even number, as desired. □

Notice that this completes the proof of Theorem 1.1. Indeed, for each $1\leq i\leq 4$, set $C_i := D_i + D'_i$. Lemmas 4.3, 4.16, 4.17, 4.20, 4.22 and 4.23 together imply that if $n \equiv i \mod 4$, then

$$(C_i-\eta)2^{n/4} \leq f_{\max}(n) \leq (C_i+\eta)2^{n/4},$$

as desired.

5 Maximal sum-free sets in abelian groups

Throughout this section, unless otherwise specified, G will be an abelian group of order n and we denote by $\mu(G)$ the size of the largest sum-free subset of G. Denote by $f(G)$ the number of sum-free subsets of G and by $f_{\max}(G)$ the number of maximal sum-free subsets of G. Given a set $F\subseteq G$, we write $f_{\max}(F)$ for the number of maximal sum-free subsets of G that lie in F.

The study of sum-free sets in abelian groups dates back to the 1960s. Although Diananda and Yap [8] determined $\mu(G)$ for a large class of abelian groups G, it was not until 2005 that Green and Ruzsa [13] determined $\mu(G)$ for all such G. In particular, for every finite abelian group G, $2n/7 \leq \mu(G) \leq n/2$. Further, Green and Ruzsa [13] determined $f(G)$ up to an error term in the exponent for all G, showing that $f(G) = 2^{(1+o(1))\mu(G)}$.

Given G, what can we say about $f_{\max}(G)$? Is it also the case that $f_{\max}(G)$ is exponentially smaller than $f(G)$? Wolfovitz [19] proved that $f_{\max}(G) \leq 2^{0.406n+o(n)}$ for every finite group G. For even order abelian groups G this answers the second question in the affirmative since $\mu(G) = n/2$ for such groups.

Our next result strengthens the result of Wolfovitz for abelian groups, and implies that indeed $f_{\max}(G)$ is exponentially smaller than $f(G)$ for all finite abelian groups G. Let G be fixed. By a container lemma [13, Proposition 2.1] and a removal lemma [11, Theorem 1.4] for abelian groups, there exists a collection of containers $\mathcal{F}$ such that:

(i) $|\mathcal{F}| = 2^{o(n)}$ and $F \subseteq G$ for all $F \in \mathcal{F}$;

(ii) Given any $F \in \mathcal{F}$, $F = B \cup C$ where B is sum-free with size $|B| \leq \mu(G)$ and $|C| = o(n)$;

(iii) Given any sum-free subset S of G, there is an $F \in \mathcal{F}$ such that $S \subseteq F$.

Given sets $S, T \subseteq G$, we can define the link graph $L_S[T]$ analogously to the integer case. In particular, it is easy to check that an analogue of Lemma 4.1 holds for such link graphs.

Let $F \in \mathcal{F}$ be fixed. Every maximal sum-free subset of G contained in F can be chosen by picking a sum-free set S in C (at most $2^{o(n)}$ choices by (ii)), and extending it in B (at most $\mathrm{MIS}(L_S[B]) \leq 3^{|B|/3} \leq 3^{\mu(G)/3}$ choices by Lemma 4.1 for abelian groups and the Moon-Moser theorem). Therefore, together this implies the following result.

Proposition 5.1. *Let G be an abelian group of order n. Then*

$$f_{\max}(G) \leq 3^{\mu(G)/3+o(n)}. \tag{19}$$

We do not know how far from tight the bound in Proposition 5.1 is. In particular, it would be interesting to establish whether the following bound holds.

Question 5.2. *Given an abelian group G of order n, is it true that $f_{\max}(G) \leq 2^{\mu(G)/2+o(n)}$?*

For the group $Z_2^k := Z_2 \otimes Z_2 \otimes \cdots \otimes Z_2$, the answer to the above question is affirmative and the upper bound is essentially tight.

Proposition 5.3. *The number of maximal sum-free subsets of Z_2^k is $2^{(1+o(1))\mu(Z_2^k)/2}$.*

Proof. Let $n := |Z_2^k|$. It is known that $\mu(Z_2^k) = n/2$. We first give a lower bound $f_{\max}(Z_2^k) \geq 2^{n/4}$. Write $Z_2^k = Z_2 \otimes Z_2 \otimes H$, where $H := Z_2^{k-2}$. Let $x := (0, 1, 0_H)$ and $U := \{1\} \otimes Z_2 \otimes H$. Notice that the link graph $L_x[U]$ is a perfect matching. Indeed, for any vertex $y = (1, a, h) \in U$, all of its possible neighbours in U are $x + y = (1, 1 + a, h)$, $x - y = (1, 1 - a, -h)$ and $y - x = (1, a - 1, h)$ and these elements of Z_2^k are identical. To build a collection of sum-free subsets, we first pick x and then pick exactly one of the endpoints of each edge in $L_x[U]$. Since $|U| = n/2$, we obtain $2^{n/4}$ sum-free subsets S in this way. These sets might not be maximal, but no further elements from U can be added into any of these sets. Hence distinct S lie in distinct maximal sum-free subsets. Therefore we have

$$f_{\max}(Z_2^k) \geq 2^{n/4}.$$

We now proceed with the proof of the upper bound. Let $\mathcal{F}$ be the family of $2^{o(n)}$ containers defined before Proposition 5.1. It suffices to show that $f_{\max}(F) \leq 2^{(1/4+o(1))n}$ for every container $F \in \mathcal{F}$. Fix a container $F \in \mathcal{F}$. We have $F = B \cup C$ with B sum-free, $|B| \leq \mu(Z_2^k) = n/2$ and $|C| = o(n)$. Every maximal sum-free subset of Z_2^k in F can be built by choosing a sum-free set S in C and extending S in B to a maximal one. The number of choices for S is at most $2^{|C|} = 2^{o(n)}$. For a fixed S, let $\Gamma := L_S[B]$ be the link graph of S on B. Then Lemma 4.1 (for abelian groups) implies that the number of extensions is at most $\mathrm{MIS}(\Gamma)$. Observe that Γ is triangle-free. Indeed, suppose to the contrary that there exists a triangle on vertices $a, b, c \in B \subseteq Z_2^k$. Since for any $x \in Z_2^k$, $x = -x$, we may assume that $a + b = s_1$, $b + c = s_2$ and $a + c = s_3$ for some $s_1, s_2, s_3 \in S$. Furthermore, s_1, s_2, s_3 are distinct elements in S since a, b, c are distinct in B. Then we have $s_1 + s_2 = a + 2b + c = a + c = s_3$, contradicting S being sum-free. Thus by (1), we have

$$\mathrm{MIS}(\Gamma) \leq 2^{|B|/2} \leq 2^{n/4}$$

and so

$$f_{\max}(F) \leq 2^{|C|} \cdot 2^{n/4} = 2^{(1/4+o(1))n},$$

as desired. $\square$

The following construction gives a lower bound $f_{\max}(Z_n) \geq 6^{(1/18-o(1))n}$. Let $n = 9k+i$ for some $0 \leq i \leq 8$ and $M := [3k+1, 6k]$. Set $\Gamma := L_{\{k,-2k\}}[M]$. Then $|M|/6 - o(n)$ components of Γ are copies of $K_3 \square K_2$ as there are at most a constant number of components of Γ that are not copies of $K_3 \square K_2$. Observe that $K_3 \square K_2$ contains 6 maximal independent sets. Thus, $\mathrm{MIS}(\Gamma) \geq 6^{(1/18-o(1))n}$, yielding the desired lower bound on $f_{\max}(Z_n)$. It is known that $\mu(Z_p) = (1/3 + o(1))p$, if p is *prime*, so together with (19), we obtain the following result.

Proposition 5.4. *If p is prime then*

$$1.1^{p-o(p)} \leq 6^{(1/18-o(1))p} \leq f_{\max}(Z_p) \leq 3^{(1/9+o(1))p} \leq 1.13^{p+o(p)}.$$

It would be interesting to close the gap in Proposition 5.4.

We end this section with two more constructions that would match the upper bound in Question 5.2 if it is true. For this, we need the following simple fact.

Fact 5.5. *Suppose G is an abelian group of odd order. Then given a fixed $x \in G$, there is a unique solution in G to the equation $2y = x$.*

Notice that Fact 5.5 is false for abelian groups of even order.

Proposition 5.6. *Suppose that $3|n$ where n is not divisible by a prime p with $p \equiv 2 \mod 3$. Then $f_{\max}(G) \geq 2^{(n-9)/6} = 2^{(\mu(G)-3)/2}$.*

Proof. First note that $\mu(G) = n/3$ for such groups (see [13]). Let $H \leq G$ be a subgroup of index 3. Then there are three cosets $0 + H, 1 + H, 2 + H$. Pick some $x \in 2 + H$. Then consider the link graph $\Gamma := L_x[1 + H]$ on $n/3$ vertices. There is a loop at $2x \in V(\Gamma)$. For every $y \in 1 + H$, $x + y \in 0 + H$, $y - x \in 2 + H$ and $x - y \in 1 + H$. So y has only one neighbour $x - y$ in $1 + H$ (unless $y = 2x$, which has a loop). By Fact 5.5, there is a unique $y \in 1 + H$ such that $x - y = y$. Overall this implies that Γ consists of the disjoint union of a matching M of size $(n-3)/6$, with a loop at at most one of the vertices in M, together with an additional vertex with a loop. Clearly $\mathrm{MIS}(\Gamma) \geq 2^{(n-9)/6}$ and so $f_{\max}(G) \geq 2^{(n-9)/6}$. $\square$

Proposition 5.7. *Suppose that n is only divisible by primes p such that $p \equiv 1 \mod 3$. Suppose further that the exponent of G (the largest order of any element of G) is 7. Then $f(G) \geq 2^{n/7-1} = 2^{\mu(G)/2-1}$.*

Proof. First note that $\mu(G) = 2n/7$ for such groups (see [13]). Let $H \leq G$ be a subgroup of index 7. Then pick some $x \in 1 + H$. Consider the link graph $\Gamma := L_x[(2+H) \cup (3+H)]$ on $2n/7$ vertices. There is a loop at $2x \in 2 + H$ in Γ. The remaining edges of Γ form a perfect matching between $2 + H$ and $3 + H$. Therefore $\text{MIS}(\Gamma) = 2^{n/7-1}$ and so $f_{\max}(G) \geq 2^{n/7-1}$. □

We conclude the section with two conjectures.

Conjecture 5.8. *For every abelian group G of order n,*

$$2^{n/7} \leq f_{\max}(G) \leq 2^{n/4+o(n)},$$

where the bounds, if true, are best possible.

We also suspect that there is an infinite class of finite abelian groups for which the upper bounds in Conjecture 5.8 and Question 5.2 are far from tight.

Conjecture 5.9. *There is a sequence of finite abelian groups $\{G_i\}$ of increasing order such that for all i,*

$$f_{\max}(G_i) \leq 2^{\mu(G_i)/2.01}.$$

Acknowledgements

The authors are grateful to the BRIDGE strategic alliance between the University of Birmingham and the University of Illinois at Urbana-Champaign. This research was conducted as part of the 'Building Bridges in Mathematics' BRIDGE Seed Fund project.

References

[1] J. Balogh, H. Liu, S. Petříčková and M. Sharifzadeh, The typical structure of maximal triangle-free graphs, submitted.

[2] J. Balogh, H. Liu, M. Sharifzadeh and A. Treglown, The number of maximal sum-free subsets of integers, *Proc. Amer. Math. Soc.*, to appear.

[3] J. Balogh, R. Morris and W. Samotij, Independent sets in hypergraphs, *J. Amer. Math. Soc.*, to appear.

[4] J. Balogh and S. Petříčková, The number of the maximal triangle-free graphs, *Bull. London Math. Soc.*, 46, (2014), 1003–1006.

[5] P. Cameron and P. Erdős, On the number of sets of integers with various properties, in Number Theory (R.A. Mollin, ed.), 61–79, Walter de Gruyter, Berlin, 1990.

[6] P. Cameron and P. Erdős, Notes on sum-free and related sets, *Combin. Probab. Comput.*, 8, (1999), 95–107.

[7] J. Deshouillers, G. Freiman, V. Sós and M. Temkin, On the structure of sum-free sets II, *Astérisque*, 258, (1999), 149–161.

[8] P.H. Diananda and H.P. Yap, Maximal sum-free sets of elements of finite groups, *Proc. Japan Acad.*, **45**, (1969), 1–5.

[9] Z. Füredi, The number of maximal independent sets in connected graphs, *J. Graph Theory*, **11**, (1987), 463–470.

[10] B. Green, The Cameron-Erdős conjecture, *Bull. London Math. Soc.*, 36, (2004), 769–778.

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

[12] B. Green, R. Morris, Counting sets with small sumset and applications, *Combinatorica*, to appear.

[13] B. Green and I. Ruzsa, Sum-free sets in abelian groups, *Israel J. Math.*, 147, (2005), 157–189.

[14] M. Hujter and Z. Tuza, The number of maximal independent sets in triangle-free graphs, *SIAM J. Discrete Math.*, 6, (1993), 284–288.

[15] T. Łuczak and T. Schoen, On the number of maximal sum-free sets, *Proc. Amer. Math. Soc.*, 129, (2001), 2205–2207.

[16] J.W. Moon and L. Moser, On cliques in graphs, *Israel J. Math.*, 3, (1965), 23–28.

[17] A.A. Sapozhenko, The Cameron-Erdős conjecture, (Russian) *Dokl. Akad. Nauk.*, 393, (2003), 749–752.

[18] D. Saxton and A. Thomason, Hypergraph containers, submitted.

[19] G. Wolfovitz, Bounds on the number of maximal sum-free sets, *European J. Combin.*, 30, (2009), 1718–1723.

A Appendix

Here we give the proofs of Lemma 3.2 and Corollary 3.3. The following simple facts will be used in the proof of Lemma 3.2.

Fact A.1. *Suppose that G is a graph. For any maximal independent set I in G that contains x, $I \setminus \{x\}$ is a maximal independent set in $G \setminus (N_G(x) \cup \{x\})$.*

Given $x \in V(G)$, let $\mathrm{MIS}_G(x)$ denote the number of maximal independent sets in G that contain x.

Fact A.2. *Suppose that G is a graph. Given any $x \in V(G)$,*

$$\mathrm{MIS}(G) \leq \mathrm{MIS}_G(x) + \sum_{v \in N_G(x)} \mathrm{MIS}_G(v).$$

Notice that Fact A.2 is not true in general if G is a graph with loops.

Lemma A.3 (Füredi [9])**.** *For $m \geq 6$,* $\mathrm{MIS}(C_m) = \mathrm{MIS}(C_{m-2}) + \mathrm{MIS}(C_{m-3})$.

Lemma A.3 implies the following simple result.

Lemma A.4. *For all $m \geq 4$,* $\mathrm{MIS}(C_m) < 2^{0.49m}$.

Proof. It is easy to check that the lemma holds for $m = 4, 5, 6$. For $m \geq 7$, by induction, Lemma A.3 implies that

$$\mathrm{MIS}(C_m) = \mathrm{MIS}(C_{m-2}) + \mathrm{MIS}(C_{m-3}) < 2^{0.49m}(2^{-0.98} + 2^{-1.47}) < 2^{0.49m}.$$

$\square$

Corollary A.5. *If G is the vertex-disjoint union of cycles of length at least 4 then* $\mathrm{MIS}(G) < 2^{0.49|G|}$.

We now combine the previous results to prove Lemma 3.2.

Proof of Lemma 3.2. We proceed by induction on n. The case when $n \leq 4$ is an easy calculation. We split the argument into several cases.

Case 1: There is a vertex $x \in V(G)$ of degree 0.

By induction $G' := G \setminus \{x\}$ is such that $\mathrm{MIS}(G') \leq 2^{(n-1)/2-k/(100D^2)}$ and clearly $\mathrm{MIS}(G) = \mathrm{MIS}(G')$.

Case 2: There is a vertex $x \in V(G)$ of degree 1.

First suppose that x is adjacent to a vertex y of degree 1. Then consider $G' := G \setminus \{x, y\}$. Note that $\mathrm{MIS}(G) = 2 \cdot \mathrm{MIS}(G')$. Further, $|G'| = n - 2$, $e(G') \geq (n-2)/2 + k$ and $\Delta(G') \leq D$. Thus, by induction we have that

$$\mathrm{MIS}(G) = 2 \cdot \mathrm{MIS}(G') \leq 2 \times 2^{(n-2)/2-k/(100D^2)} = 2^{n/2-k/(100D^2)},$$

as desired.

Otherwise x is adjacent to a vertex y of degree $d \geq 2$. Consider $G' := G \setminus \{x, y\}$. So $|G'| = n-2$, $e(G') \geq (n-2)/2 + k - d + 1$ and $\Delta(G') \leq D$. Therefore by induction and Fact A.1,

$$\mathrm{MIS}_G(x) \leq \mathrm{MIS}(G') \leq 2^{(n-2)/2-(k-d+1)/(100D^2)} \leq 2^{n/2-k/(100D^2)}(2^{-1+d/(100D^2)}). \tag{20}$$

Consider $G'' := G \setminus (N_G(y) \cup \{y\})$. So $|G''| = n - d - 1$, $e(G'') \geq n/2 + k - (d-1)D - 1 \geq (n-d-1)/2 + (k-(d-1)D)$ and $\Delta(G'') \leq D$. Thus, by induction and Fact A.1,

$$\begin{aligned} \mathrm{MIS}_G(y) \leq \mathrm{MIS}(G'') &\leq 2^{(n-d-1)/2-(k-(d-1)D)/(100D^2)} \\ &= 2^{n/2-k/(100D^2)}(2^{-(d+1)/2+(d-1)/100D}). \end{aligned} \tag{21}$$

Now as $2 \leq d \leq D$ we have that

$$2^{-1+d/(100D^2)} + 2^{-(d+1)/2+(d-1)/100D} \leq 2^{-1+1/100} + 2^{-3/2+1/100} < 1.$$

So (20) and (21) together with Fact A.2 imply that

$$\mathrm{MIS}(G) \leq \mathrm{MIS}_G(x) + \mathrm{MIS}_G(y) < 2^{n/2-k/(100D^2)},$$

as desired.

Case 3: $\delta(G) \geq 4$.
Let $v \in V(G)$ be the vertex of smallest degree in G and write $\deg_G(v) = i - 1 \geq 4$. Given any $w \in N_G(v) \cup \{v\}$ let $G' := G \setminus (N_G(w) \cup \{w\})$. So $|G'| = n - \deg_G(w) - 1$, $e(G') \geq n/2 + (k - \deg_G(w)D) \geq |G'|/2 + (k - \deg_G(w)D)$ and $\Delta(G') \leq D$. Hence by induction and Fact A.1

$$\mathrm{MIS}_G(w) \leq \mathrm{MIS}(G') \leq 2^{(n-\deg_G(w)-1)/2-(k-\deg_G(w)D)/100D^2)} \leq 2^{(n-i)/2-(k-iD)/(100D^2)}.$$

Thus by Fact A.2 we have that

$$\mathrm{MIS}(G) \leq i \times 2^{(n-i)/2-(k-iD)/(100D^2)} \leq (i2^{-i/2+i/100})2^{n/2-k/(100D^2)} < 2^{n/2-k/(100D^2)},$$

as desired. (Here we used that for $i \geq 5$, $i2^{-i/2+i/100} < 1$.)
Case 4: $\delta(G) = 2$ and there exist $v, w \in V(G)$ such that $\deg_G(v) = 2$, $\deg_G(w) \geq 3$ and $vw \in E(G)$.
By arguing as before (using induction and Facts A.1 and A.2) we have that

$$\begin{aligned}\mathrm{MIS}(G) &\leq \mathrm{MIS}_G(v) + \sum_{u \in N_G(v)} \mathrm{MIS}_G(u) \leq 2 \times 2^{(n-3)/2-(k-2D)/(100D^2)} + 2^{(n-4)/2-(k-3D)/(100D^2)} \\ &< 2^{n/2-k/(100D^2)},\end{aligned}$$

as desired. (Here we have used that $2 \cdot 2^{-3/2+1/50} + 2^{-2+3/100} < 1$.)

Cases 1–4 imply that we may now assume that G consists precisely of 2-regular components and components of minimum degree at least 3.
Case 5: There exist $v, w \in V(G)$ such that $\deg_G(v) = 3$, $\deg_G(w) \geq 4$ and $vw \in E(G)$.
By arguing similarly to before (using induction and Facts A.1 and A.2) we have that

$$\begin{aligned}\mathrm{MIS}(G) &\leq \mathrm{MIS}_G(v) + \sum_{u \in N_G(v)} \mathrm{MIS}_G(u) \leq 3 \times 2^{(n-4)/2-(k-3D)/(100D^2)} + 2^{(n-5)/2-(k-4D)/(100D^2)} \\ &< 2^{n/2-k/(100D^2)},\end{aligned}$$

as desired. (Here we have used that $3 \cdot 2^{-2+3/100} + 2^{-5/2+1/25} < 1$.)

We may now assume that G consists only of 2- and 3-regular components and components of minimum degree at least 4. However, if there is a component of minimum degree at least 4 then by arguing precisely as in Case 3, we obtain that $\mathrm{MIS}(G) \leq 2^{n/2-k/(100D^2)}$. So we may now assume G consists of 2- and 3-regular components only.
Case 6: G contains a 3-regular component.
Here we use the fact that $\mathrm{MIS}(G) \leq \mathrm{MIS}(G \setminus \{v\}) + \mathrm{MIS}(G \setminus (N_G(v) \cup \{v\}))$ for any $v \in V(G)$. Indeed, by induction we have

$$\mathrm{MIS}(G) \leq 2^{(n-1)/2-(k-5/2)/(100D^2)} + 2^{(n-4)/2-(k-7)/(100D^2)} < 2^{n/2-k/(100D^2)},$$

as desired. (Here we have used that $2^{-1/2+1/40} + 2^{-2+7/100} < 1$.)
Case 7: G is 2-regular.
Since G is triangle-free, Corollary A.5 implies that $\mathrm{MIS}(G) \leq 2^{0.49n} \leq 2^{n/2-k/(100D^2)}$, as desired. □

Finally, we show that Corollary 3.3 follows from Lemma 3.2.
Proof of Corollary 3.3. Every maximal independent set in G can be obtained in the following two steps:

(1) Choose an independent set $S \subseteq T$.

(2) Extend S in $V(G) \setminus T = V(G')$, i.e. choose a set $R \subseteq V(G')$ such that $R \cup S$ is a maximal independent set in G.

Note that although every maximal independent set in G can be obtained in this way, it is not necessarily the case that given an arbitrary independent set $S \subseteq T$, there exists a set $R \subseteq V(G')$ such that $R \cup S$ is a maximal independent set in G. Notice that if $R \cup S$ is maximal, R is also a maximal independent set in $G'' := G \setminus (T \cup N_G(S))$. The number of choices for S in (1) is at most $2^{|T|}$. Note that G'' is triangle-free, $\Delta(G'') \leq D$ and $e(G'') \geq e(G') - |T|D^2 \geq |G''|/2 + (k - |T|D^2)$. Thus, Lemma 3.2 implies that the number of extensions in (2) is at most $2^{n/2-(k-|T|D^2)/(100D^2)}$. Therefore, we have $\mathrm{MIS}(G) \leq 2^{|T|} \cdot 2^{n/2-(k-|T|D^2)/(100D^2)}$, as desired. $\square$