# 1-SATURATING SETS, CAPS, AND ROUND SETS IN BINARY SPACES

DAVID J. GRYNKIEWICZ AND VSEVOLOD F. LEV

Abstract. We show that for positive integer $r$, every minimal 1-saturating set in $\mathrm{PG}(r-1,2)$ of size at least $\frac{11}{36}\,2^r+3$ can be obtained by adding the point at infinity to a complete cap, translating the resulting set by its element, and removing the infinite point from the translate. Stated algebraically: if $G$ is an elementary abelian 2-group, and a set $A \subseteq G \setminus \{0\}$ with $|A| \ge \frac{11}{36}\,|G|+3$ satisfies $A \cup 2A = G$ and is minimal subject to this condition, then there are a maximal sum-free set $S \subseteq G$ and an element $s \in S \cup \{0\}$ so that $A = \big(s+(S \cup \{0\})\big) \setminus \{0\}$.

Our approach is based on characterizing those sets $A$ in elementary abelian 2-groups such that for every proper subset $B$ of $A$, the sumset $2B$ is a proper subset of the sumset $2A$.

## 1. Saturating sets and caps. The main result.

Let $r \ge 1$ be an integer, $q$ a prime power, and $A \subseteq \mathrm{PG}(r-1,q)$ a set in the $(r-1)$-dimensional projective space over the $q$-element field $\mathbb{F}_q$. Given an integer $\rho \ge 1$, one says that $A$ is *$\rho$-saturating* if every point of $\mathrm{PG}(r-1,q)$ belongs to the subspace, generated by $\rho+1$ points of $A$. Furthermore, $A$ is said to be a *cap* if no three points of $A$ are collinear; a cap is *complete* if it is not properly contained in another cap. Since the properties of being $\rho$-saturating and that of being a cap are monotonic, of particular interest are minimal $\rho$-saturating sets and complete caps.

In this paper we are concerned exclusively with the case $q=2$ and $\rho=1$. A large random set in $\mathrm{PG}(r-1,2)$ is 1-saturating with very high probability, but the probability that it is *minimal* 1-saturating is extremely low; thus, one can expect that large minimal 1-saturating sets are well-structured and can be explicitly described. A similar heuristic applies to large complete caps and indeed, a classical result of Davydov and Tombak [DT89] establishes the structure of complete caps of size, larger than $2^{r-2}+1$. Classifying large 1-saturating sets in $\mathrm{PG}(r-1,2)$ seems to be considerably subtler, which is only natural bearing in mind that complete caps can be characterized as those 1-saturating sets, possessing the extra property of having no internal lines (as will be explained shortly).

Below we use the language of abelian groups, rather than projective geometries. Accordingly, denoting by $\mathbb{F}_2^r$ the elementary abelian 2-group of rank $r \ge 1$ and writing

$$2A := \{a_1+a_2 \colon a_1,a_2 \in A\}; \quad A \subseteq \mathbb{F}_2^r,$$

we interpret 1-saturating sets in $\mathrm{PG}(r-1,2)$ as those subsets $A \subseteq \mathbb{F}_2^r \setminus \{0\}$ with the property that $A \cup 2A = \mathbb{F}_2^r$. Similarly, caps in $\mathrm{PG}(r-1,2)$ are understood as sets $A \subseteq \mathbb{F}_2^r$ with $A \cap 2A = \varnothing$; such sets are customarily referred to as *sum-free*. Complete caps are thus identified with maximal (by inclusion) sum-free sets.

It is well known and easy to see that a sum-free set $A \subseteq \mathbb{F}_2^r$ is maximal if and only if the sets $A$ and $2A$ partition $\mathbb{F}_2^r$; that is, in addition to being disjoint, satisfy $A \cup 2A = \mathbb{F}_2^r$. Consequently, any maximal sum-free set is minimal 1-saturating. Beyond this simple observation, the only general result which seems to be known about minimal 1-saturating sets in $\mathbb{F}_2^r$ is established in [DMP03]; it asserts that the largest possible size of such a set is $2^{r-1}$, examples being furnished by the following two constructions:

(i) if $H < \mathbb{F}_2^r$ is an index-2 subgroup and $g \in \mathbb{F}_2^r \setminus H$, then $g + H$ is minimal 1-saturating;

(ii) with $H$ and $g$ as in (i), the set $\{g\} \cup (H \setminus \{0\})$ is minimal 1-saturating.

An extension of construction (i) has just been mentioned: any maximal sum-free set is minimal 1-saturating. Construction (ii) can be extended by observing that if $S$ is maximal sum-free and $s \in S$, then $A := \{s\} \cup \big((S+s) \setminus \{0\}\big)$ is minimal 1-saturating: for in this case

$$A \cup 2A = 2(A \cup \{0\}) = 2\big(s + (S \cup \{0\})\big) = 2(S \cup \{0\}) = S \cup 2S = \mathbb{F}_2^r,$$

and this computation also shows that for any proper subset $B \subset A$ we have $2B \ne \mathbb{F}_2^r$.

Indeed, a common description can be given to these two extensions: namely, if $S \subseteq \mathbb{F}_2^r$ is maximal sum-free and $s \in S \cup \{0\}$, then $A := \big(s + (S \cup \{0\})\big) \setminus \{0\}$ is minimal 1-saturating. In this paper we classify completely minimal 1-saturating sets in $\mathbb{F}_2^r$ of size at least $\frac{11}{36}\,2^r + 3$, showing that they all are of this form.

**Theorem 1.** *Let $r \ge 1$ be an integer. A set $A \subseteq \mathbb{F}_2^r \setminus \{0\}$ with $|A| > \frac{11}{36}\,2^r + 3$ is minimal 1-saturating if and only if there are a maximal sum-free set $S \subseteq \mathbb{F}_2^r$ and an element $s \in S \cup \{0\}$ such that $A = \big(s + (S \cup \{0\})\big) \setminus \{0\}$.*

We notice that Theorem 1 provides a comprehensive characterization of large minimal 1-saturating sets, as the structure of large maximal sum-free sets is known due to the result of Davydov and Tombak, mentioned at the beginning of this section. In particular, every maximal sum-free set in $\mathbb{F}_2^r$ of size larger than $9 \cdot 2^{r-5}$ is known either to be the non-zero coset of an index-2 subgroup, or to have the form $B + H$, where $H < \mathbb{F}_2^r$ is a subgroup of index 16, and $B \subset \mathbb{F}_2^r$ is a set of five elements, adding up to 0 and such that $\mathbb{F}_2^r = \langle B \rangle \oplus H$. (Here and below in the paper, for a set $B$ of group elements, by $\langle B \rangle$ we denote the subgroup, generated by $B$. Furthermore, given yet another subset $C$ of the same group, we write $B+C := \{b+c\colon b \in B,\, c \in C\}$. The set $B+C$ is commonly referred to as the *sumset* of $B$ and $C$. Notice, that $B+B = 2B$.)

We conjecture that the density assumption of Theorem 1 can actually be relaxed to $|A| > 2^{r-2} + 2$. If true, this is best possible.

**Example 1.** Given an integer $r \geq 4$, fix elements $e_1, e_2 \in \mathbb{F}_2^r$ and an index-4 subgroup $H < \mathbb{F}_2^r$ with $\mathbb{F}_2^r = \langle e_1, e_2 \rangle \oplus H$, and let $A := (\langle e_1, e_2 \rangle \cup H) \setminus \{0\}$. Straightforward verification shows that $A$ is minimal 1-saturating. Next, if $A = \big(s+(S\cup\{0\})\big)\setminus\{0\}$ for a subset $S \subseteq \mathbb{F}_2^r\setminus\{0\}$ and an element $s \in S\cup\{0\}$, then $S\cup\{0\} = s+(\langle e_1, e_2\rangle\cup H)$. Since this set contains 0, we have $s \in \langle e_1, e_2\rangle \cup H$. If $s \in H$, then $S$ contains all non-zero elements of $H$, and so it is not sum-free. If $s = e_1$, then $S = \{e_1, e_2, e_1+e_2\}\cup(e_1+H)$ is evidently not sum-free, and similarly it is not sum-free if $s = e_2$ or $s = e_1 + e_2$. Thus, $A$ cannot be represented as in Theorem 1.

More generally, if $F$ and $H$ are non-trivial subgroups of $\mathbb{F}_2^r$ with $\mathbb{F}_2^r = F \oplus H$, then the set $(F \cup H) \setminus \{0\}$ is minimal 1-saturating, but cannot be represented as in Theorem 1.

## 2. Round sets and the Unique Representation Graph.

In a paradoxical way, for a minimal 1-saturating set minimality seems to be more important than saturation. This idea is captured in the notion of a round set, introduced in the present section. We also bring into consideration unique representation graphs, which are of fundamental importance for our argument, and establish some basic properties of round sets and unique representation graphs. Finally, we state a structure theorem for round set (Theorem 2 below), and show that it implies Theorem 1.

The remainder of the paper is structured as follows. Important auxiliary results are gathered in Section 3. In Section 4 we prove a "light version" of Theorem 1, with the assumption on the size of $A$ strengthened to $|A| > \frac{1}{3}\,2^r + 2$; besides supplying a proof of Theorem 1 for small dimensions ($r = 4$ and $r = 5$), it serves a simplified model of our method, keeping many of the major ideas while avoiding most of the technicalities. Sections 5–7 are devoted to the proof of Theorem 2: in Section 5 the problem is reduced to the case where the unique representation graph is known to have at least two isolated edges, Sections 6 and 7 present a treatment of this case.

Let $r \geq 1$ be an integer. We say that a set $A \subseteq \mathbb{F}_2^r$ is *round* if for every proper subset $B \subset A$, we have $2B \neq 2A$; that is, for every $a \in A$ there exists $a' \in A$ such that $a + a'$ has a unique (up to the order of summands) representation as a sum of two elements of $A$.

It is immediate from the definition that a set $A \subseteq \mathbb{F}_2^r$ is minimal 1-saturating if and only if it satisfies $2(A \cup \{0\}) = \mathbb{F}_2^r$ and is minimal subject to this condition. The following simple lemma takes this observation a little bit further.

**Lemma 1.** *Let $r \ge 1$ be an integer. If $A \subseteq \mathbb{F}_2^r$ is minimal 1-saturating, then either $A$, or $A \cup \{0\}$ is round.*

*Remark.* As it is easy to derive from Theorem 1, if $A \subseteq \mathbb{F}_2^r$ is a large minimal 1-saturating set, then, indeed, $A \cup \{0\}$ is round.

*Proof of Lemma 1.* Suppose that $A \subseteq \mathbb{F}_2^r$ is minimal 1-saturating. If $A \cup \{0\}$ is not round, then there exists $a_0 \in A \cup \{0\}$ such that $2\big((A \cup \{0\}) \setminus \{a_0\}\big) = 2(A \cup \{0\}) = \mathbb{F}_2^r$. Since $a_0 \in A$ would contradict minimality of $A$, we actually have $a_0 = 0$, whence $2A = \mathbb{F}_2^r$. Now if also $A$ is not round, then there exists $a \in A$ with $2(A \setminus \{a\}) = 2A = \mathbb{F}_2^r$. This yields $2\big((A \setminus \{a\}) \cup \{0\}\big) = \mathbb{F}_2^r$ which, again, contradicts minimality of $A$. $\square$

Lemma 1 allows us to concentrate on studying large round sets instead of large 1-saturating sets; indeed, we will hardly recur to 1-saturating sets from now on, except for the deduction of Theorem 1 from Theorem 2 below.

We observe that if $S \subseteq \mathbb{F}_2^r$ is sum-free, then $0 \notin S$ and for each $g \in \mathbb{F}_2^r$, the set $g + (S \cup \{0\})$ is round. To verify this we can assume $g = 0$ (as roundness is translation invariant) and notice that, fixing arbitrarily $s_0 \in S$ and letting $S_0 := S \setminus \{s_0\}$, we have $s_0 \notin 2S$ and $s_0 \notin 2(S_0 \cup \{0\})$, whereas $s_0 \in 2(S \cup \{0\})$. In the heart of our paper is the following theorem, showing that, in fact, any large round set has the structure just described.

**Theorem 2.** *Let $r \ge 1$ be an integer and suppose that $A \subseteq \mathbb{F}_2^r$ is round. If $|A| > \frac{11}{36}\, 2^r + 3$, then there is a sum-free set $S \subseteq \mathbb{F}_2^r$ and an element $g \in \mathbb{F}_2^r$ such that $A = g + (S \cup \{0\})$.*

We now turn to the notion of a unique representation graph. Given an integer $r \ge 1$ and a set $A \subseteq \mathbb{F}_2^r$, we define $D(A)$ to be the set of all those elements of $\mathbb{F}_2^r$ with a unique, up to the order of summands, representation as a sum of two elements of $A$. By $\Gamma(A)$ we denote the graph on the vertex set $A$ in which two vertices $a_1, a_2 \in A$ are adjacent whenever $a_1 + a_2 \in D(A)$; if $|A| > 1$, then $\Gamma(A)$ is a simple, loopless graph (as all graphs below are tacitly assumed to be). We call $\Gamma(A)$ the *unique representation graph* of $A$. Notice that the number of edges of $\Gamma(A)$ is $|D(A)|$, and that for any $g \in \mathbb{F}_2^r$ we have $D(A+g) = D(A)$, while $\Gamma(g+A)$ is obtained from $\Gamma(A)$ by re-labeling the vertices.

Evidently, a set $A \subseteq \mathbb{F}_2^r$ with $|A| \ge 2$ is round if and only if $\Gamma(A)$ has no isolated vertices. Another indication of the importance of unique representation graphs is given by the following lemma.

**Lemma 2.** *Let $r \ge 1$ be an integer and suppose that $A \subseteq \mathbb{F}_2^r$ satisfies $|A| \ge 2$ and $g \in \mathbb{F}_2^r$. For $\Gamma(A)$ to have a spanning star with the center at $g$ it is necessary and sufficient that $A = g + (S \cup \{0\})$, where $S \subseteq \mathbb{F}_2^r$ is sum-free.*

*Proof.* If $g \notin A$, then $g$ is not a vertex of $\Gamma(A)$ and $A \neq g+(S\cup\{0\})$; thus, the assertion is immediate in this case. If $g \in A$, set $S := (A+g)\setminus\{0\}$, so that $A = g+(S\cup\{0\})$. The graph $\Gamma(A)$ has a spanning star with the center at $g$ if and only if for every $s \in S$ we have $g+(g+s) \in D(A)$; that is, $g+(g+s) \neq (g+s_1)+(g+s_2)$ with $s_1, s_2 \in S$. This is equivalent to $S$ being sum-free. □

By Lemma 2, to prove Theorem 2 it suffices to show that if $A \subseteq \mathbb{F}_2^r$ is a large round set, then $\Gamma(A)$ contains a spanning star. The following basic result shows that for the unique representation graph of a large set, *containing* a spanning star is equivalent to *being* a star.

**Proposition 1.** *Let $r \geq 1$ be an integer and suppose that $A \subseteq \mathbb{F}_2^r$. If $|A| > 2^{r-2}+3$, then $D(A)$ is sum-free and consequently, $\Gamma(A)$ is triangle-free.*

*Proof.* Fix two distinct elements $d_1, d_2 \in D(A)$ and consider the subgroup $H := \langle d_1, d_2\rangle$, generated by $d_1$ and $d_2$.

Suppose, to begin with, that the edges of $\Gamma(A)$, corresponding to $d_1$ and $d_2$, are incident; that is, there are $a, b_1, b_2 \in A$ such that $d_1 = a+b_1$ and $d_2 = a+b_2$. It is easy to see that the coset $a+H$ contains exactly three elements of $A$, while every other coset of $H$ contains at most two elements of $A$. Thus, the assumption $|A| > 2^{r-2}+3$ implies that there is a coset, containing exactly two elements of $A$. These two elements cannot differ by $d_1$ or $d_2$; therefore they differ by $d_1+d_2$, yielding a representation of $d_1+d_2$ as a sum of two elements of $A$. Another representation is $d_1+d_2 = b_1+b_2$, and the existence of two representations shows that $d_1+d_2 \notin D(A)$.

Assuming now that the edges of $\Gamma(A)$, corresponding $d_1$ and $d_2$, are *not* incident, find $a_1, a_2, b_1, b_2 \in A$ such that $d_1 = a_1+b_1$ and $d_2 = a_2+b_2$. There are two cosets of $H$, intersecting the set $\{a_1, a_2, b_1, b_2\}$; each of these cosets contains exactly two elements of $A$, and every other coset of $H$ contains at most two elements of $A$. The assumption $|A| > 2^{r-2}+3$ implies that there are at least two cosets, disjoint with $\{a_1, a_2, b_1, b_2\}$ and containing two elements of $A$. This yields two representations of $d_1+d_2$ leading, as above, to the conclusion $d_1+d_2 \notin D(A)$.

The second assertion of the proposition is now almost immediate: if $a_1, a_2, a_3 \in A$ were inducing a triangle in $\Gamma(A)$, then $d_1 := a_1+a_2$, $d_2 := a_2+a_3$, and $d_3 := a_3+a_1$ would be elements of $D(A)$ with $d_1+d_2 = d_3$. □

*Remark.* Examining carefully the proof of Proposition 1, one readily discovers that already the weaker assumption $|A| > 2^{r-2}+2$ suffices to ensure that $\Gamma(A)$ is triangle-free. This bound is sharp: with $e_1, e_2, H$, and $A$ as in Example 1, the vertices $e_1, e_2$, and $e_1+e_2$ of $\Gamma(A)$ induce a triangle.

Given a set $A \subseteq \mathbb{F}_2^r$, for each $a \in A$ by $\deg(a)$ we denote the degree of the vertex $a$ in $\Gamma(A)$. Yet another fundamental property of the unique representation graph is established by

**Proposition 2.** *Let $r \ge 1$ be an integer and suppose that $A \subseteq \mathbb{F}_2^r$ satisfies $|A| > 2^{r-2}+3$. If $(a_1, a_2)$ is an edge in $\Gamma(A)$, then*

$$\deg(a_1) + \deg(a_2) \ge |A| + |D(A)| - 2^{r-1}.$$

We present two different proofs.

*First proof of Proposition 2.* Let $A'$ denote the set of those elements of $A$, neighboring neither $a_1$, nor $a_2$ in $\Gamma(A)$; thus, $|A'| = |A| - \deg(a_1) - \deg(a_2)$ by Proposition 1. Then the sets

$$a_1 + A',\ a_2 + A',\ D(A),\ \text{and}\ a_1 + a_2 + D(A)$$

are easily seen to be pairwise disjoint, the fact that the last two are disjoint following from Proposition 1. Hence

$$2^r \ge 2|A'| + 2|D(A)| = 2\big(|A| + |D(A)| - \deg(a_1) - \deg(a_2)\big).$$

□

*Second proof of Proposition 2.* Since the set $D(A)$ is sum-free by Proposition 1, it contains at most one element from each coset of the two-element subgroup $\langle a_1 + a_2 \rangle$. On the other hand, $D(A)$ has exactly $\deg(a_1) + \deg(a_2) - 2$ elements in common with the set $\{a_1, a_2\} + A \setminus \{a_1, a_2\}$, the size of which is $2(|A| - 2|)$, and which is a union of cosets of $\langle a_1 + a_2 \rangle$. It follows that

$$|D(A)| \le (\deg(a_1) + \deg(a_2) - 2) + \frac{1}{2}\left(2^r - 2(|A| - 2)\right) = \deg(a_1) + \deg(a_2) + 2^{r-1} - |A|.$$

□

We conclude this section deducing Theorem 1 from Theorem 2. To this end, we first derive from Proposition 1 an interesting property of sum-free sets. Thinking projectively, if $A$ is a large cap in $\mathrm{PG}(r-1, 2)$ and the point $p$ lies on the line, determined by a pair of points in $A \setminus \{p\}$, then in fact there are *many* pairs of points in $A \setminus \{p\}$ determining a line through $p$.

**Corollary 1.** *Let $r \ge 1$ be an integer and suppose that $S \subseteq \mathbb{F}_2^r$ is sum-free. If $|S| > 2^{r-2} + \kappa$ with an integer $\kappa$, then every element of the sumset $2S$ has at least $\kappa$ representations as a sum of two elements of $S$ (such that none of these representations can be obtained from another one by permuting the summands).*

*Proof.* For $\kappa \le 1$ there is nothing to prove. Assuming that $\kappa \ge 2$ and that an element $c \in 2S$ has with fewer, than $\kappa$ representations as a sum of two elements from $S$, we find a subset $S_0 \subseteq S$ with $|S_0| \ge |S| - (\kappa - 2)$ such that $c$ has exactly one representation as a sum of two elements of $S_0$.

Let $A := S_0 \cup \{0\}$. If $|S| > 2^{r-2} + \kappa$, then $|A| > 2^{r-2} + 3$ and in view of $S_0 \subseteq D(A)$, applying Proposition 1 we get $(2S_0) \cap D(A) = \varnothing$. Thus, every element of $2S_0$ has at least two representations as a sum of two elements from $A$, therefore at least two representations as a sum of two elements from $S_0$, contradicting the choice of $S_0$. □

*Deduction of Theorem 1 from Theorem 2.* As we have already observed, if $S \subseteq \mathbb{F}_2^r$ is maximal sum-free and $s \in S \cup \{0\}$, then the set $\big(s + (S \cup \{0\})\big) \setminus \{0\}$ is minimal 1-saturating. Suppose now that $r \ge 1$ is an integer and $A \subseteq \mathbb{F}_2^r$ is a minimal 1-saturating set with $|A| > \frac{11}{36}\, 2^r + 3$. By Lemma 1, either $A \cup \{0\}$, or $A$ is round, and we show that in the former case $A$ is of the form required, while the latter case cannot occur.

If $A \cup \{0\}$ is round, then by Theorem 2 there exist a sum-free set $S \subseteq \mathbb{F}_2^r$ and an element $g \in \mathbb{F}_2^r$ such that $A \cup \{0\} = g + (S \cup \{0\})$. From $0 \in g + (S \cup \{0\})$ it follows that $g \in S \cup \{0\}$, and $\mathbb{F}_2^r = 2(A \cup \{0\}) = 2(S \cup \{0\}) = S \cup 2S$ implies that $S$ is *maximal* sum-free, proving the assertion in this case.

Suppose now that $A$ is round, so that by Theorem 2 there exist a sum-free set $S \subseteq \mathbb{F}_2^r$ and an element $g \in \mathbb{F}_2^r$ with $A = g + (S \cup \{0\})$. Without loss of generality we assume that $A \cup \{0\} = g + (S \cup \{0, g\})$ is *not* round, whence $S \cup \{g\}$ is not sum-free; that is, $g \in 2S$ and we write $g = s_1 + s_2$ with $s_1, s_2 \in S$. Notice, that $0 \notin A$ yields $s_1 \ne s_2$ and that $2A = S \cup 2S$ and

$$A \cup 2A = S \cup (g + S) \cup 2S.$$

Let $S_1 := S \setminus \{s_1\}$ and $A_1 := g + (S_1 \cup \{0\})$. Since $|S| = |A| - 1 > 2^{r-2} + 2$, by Corollary 1 (applied with $\kappa = 2$) we have $2S_1 = 2S$. It follows that

$$A_1 \cup 2A_1 = S_1 \cup (g + S_1) \cup 2S.$$

On the other hand, $s_1, s_2 \in S_1 \cup (g + S_1)$ implies $S_1 \cup (g + S_1) = S \cup (g + S)$ and thus $A_1 \cup 2A_1 = A \cup 2A$, contradicting minimality of $A$. □

## 3. Notation and auxiliary results.

In this section we deviate slightly from the flow of the proof to introduce some important notation and results, preparing the ground for the rest of the argument.

Given a subgroup $H$ of an abelian group $G$, by $\varphi_H$ we denote the canonical homomorphism from $G$ onto the quotient group $G/H$.

For a subset $B$ of an abelian group $G$, the (maximal) period of $B$ will be denoted by $\pi(B)$; recall that this is the subgroup of $G$ defined by

$$\pi(B) := \{g \in G\colon B + g = B\},$$

and that $B$ is called *periodic* if $\pi(B) \ne \{0\}$ and *aperiodic* otherwise. Thus, $B$ is a union of $\pi(B)$-cosets, and $\pi(B)$ lies above every subgroup $H \le G$ such that $B$ is a union of $H$-cosets. Observe also that $\pi(B) = G$ if and only if either $B = \varnothing$ or $B = G$, and that $\varphi_{\pi(B)}(B)$ is an aperiodic subset of the group $G/\pi(B)$.

**Theorem 3** (Kneser, [Kn53, Kn55]; see also [Mn76])**.** *Let $B$ and $C$ be finite, non-empty subsets of an abelian group $G$. If*

$$|B + C| \le |B| + |C| - 1,$$

*then, letting $H := \pi(B + C)$, we have*

$$|B + C| = |B + H| + |C + H| - |H|.$$

**Corollary 2.** *Let $r \ge 1$ be an integer and suppose that the sets $B, C \subseteq \mathbb{F}_2^r$ are disjoint and non-empty. If $|B| + |C| > 2^{r-1}$, then $B \cup C$ is not disjoint with $B + C$.*

*Remark.* If the elements $e_1, e_2 \in \mathbb{F}_2^r$ and the subgroup $H < \mathbb{F}_2^r$ of index 4 are so chosen that $\mathbb{F}_2^r = \langle e_1, e_2\rangle \oplus H$, then the sets $B := e_1 + H$ and $C := e_2 + H$ are disjoint, and so are their union $B \cup C = \{e_1, e_2\} + H$ and their sumset $B + C = e_1 + e_2 + H$; at the same time, $|B| + |C| = 2^{r-1}$. This shows that the bound $2^{r-1}$ of Corollary 2 is sharp.

*Proof of Corollary 2.* We use induction by $r$. The case $r = 1$ is immediate, and we assume that $r \ge 2$. Assuming, furthermore, that $B \cup C$ and $B + C$ are disjoint, whereas $|B| + |C| > 2^{r-1}$, we derive

$$|B + C| \le 2^r - |B| - |C| < |B| + |C| - 1.$$

Set $H := \pi(B + C)$. By Theorem 3, the subgroup $H$ is non-trivial and

$$|(B + H) \setminus B| + |(C + H) \setminus C| = |B + C| - |B| - |C| + |H| < |H| - 1.$$

The left-hand side can be interpreted as the total number of "$H$-gaps" in $B$ and $C$, showing that $B + H$ and $C + H$ are disjoint. Evidently, these two sets are also disjoint with $B + C$. Consequently, $\varphi_H(B)$ and $\varphi_H(C)$ are disjoint, non-empty subsets of the group $\mathbb{F}_2^r/H$, and $\varphi_H(B) \cup \varphi_H(C)$ is disjoint with $\varphi_H(B) + \varphi_H(C) = \varphi_H(B + C)$. This contradicts the induction hypothesis in view of

$$\begin{aligned}|\varphi_H(B)| + |\varphi_H(C)| &= (|B + H| + |C + H|)/|H| \\ &\ge (|B| + |C|)/|H| > 2^{r-1}/|H| = \frac{1}{2}\,|\mathbb{F}_2^r/H|.\end{aligned}$$

□

For integer $k$ and subsets $B$ and $C$ of an additively written group, let $B \overset{k}{+} C$ denote the set of all those group elements with at least $k$ representations as $b+c$ with $b \in B$ and $c \in C$; thus, for instance, $B \overset{1}{+} C = B + C$. We need a corollary of the following theorem, which is (a refinement of) a particular case of the main result of [G].

**Theorem 4** (Grynkiewicz, [G, Theorem 1.2])**.** *Let $G$ be an abelian group and suppose that $B, C \subseteq G$ are finite and satisfy $\min\{|B|, |C|\} \geq 2$. Then either*

$$|B \overset{1}{+} C| + |B \overset{2}{+} C| \geq 2|B| + 2|C| - 4,$$

*or there exist subsets $B' \subseteq B$ and $C' \subseteq C$ with*

$$l := |B \setminus B'| + |C \setminus C'| \leq 1,$$

$$B' + C' = B' \overset{2}{+} C' = B \overset{2}{+} C,$$

*and*

$$\begin{aligned} |B \overset{1}{+} C| + |B \overset{2}{+} C| &\geq 2|B| + 2|C| - (2-l)(|H| - \rho) - 2l \\ &\geq 2|B| + 2|C| - 2|H|, \end{aligned}$$

*where $H = \pi(B \overset{2}{+} C)$ and $\rho = |(B' + H) \setminus B'| + |(C' + H) \setminus C'|$.*

(For our present purposes, the reader can completely ignore the definitions of $H$ and $\rho$ in the statement of the theorem and the part of the conclusion, involving these quantities.)

**Corollary 3.** *If $G$ is a finite abelian group and $B, C \subseteq G$ satisfy $\min\{|B|, |C|\} \geq 2$, then*

$$|B \overset{2}{+} C| \geq \min\{2|B| + 2|C| - 4 - |G|, |B| - 1\}.$$

*Proof.* If $|B \overset{1}{+} C| + |B \overset{2}{+} C| \geq 2|B| + 2|C| - 4$, then $|B \overset{2}{+} C| \geq 2|B| + 2|C| - 4 - |G|$ follows trivially. Otherwise we apply Theorem 4 to find $B' \subseteq B$ and $C' \subseteq C$, satisfying $|B \setminus B'| + |C \setminus C'| \leq 1$ and $B' + C' = B' \overset{2}{+} C' = B \overset{2}{+} C$. Now

$$|B \overset{2}{+} C| = |B' + C'| \geq |B'| \geq |B| - 1.$$

□

Finally, we prove several simple graph-theoretic lemmas and apply them to the unique representation graph.

Recall, that the matching number of a graph is the largest size of its matching.

**Lemma 3.** *Let $(V, E)$ be a triangle-free graph without isolated vertices, such that the matching number of $(V, E)$ does not exceed 2. If $|V| \geq 6$, then $(V, E)$ is either a star, or a union of two stars, possibly with an edge between their centers. More precisely,*

*there is a partition* $V = \{v_1, v_2\} \cup V_0 \cup V_1 \cup V_2$ *such that* $E$ *consists of all pairs* $(v_1, v)$ *with* $v \in V_0 \cup V_1$*, all pairs* $(v_2, v)$ *with* $v \in V_0 \cup V_2$*, and, possibly, the pair* $(v_1, v_2)$*.*

*Proof.* We notice that $(V, E)$ does not contain a pentagon: for otherwise one could construct a matching of size 3 using the edges of the pentagon and an edge, incident with a vertex outside the pentagon. Furthermore, $(V, E)$ does not contain cycles of length 6 or more. Consequently, $(V, E)$ contains no odd cycles; hence it is bipartite.

By König's theorem, $(V, E)$ has a vertex cover of size at most 2. Now if $\{v\}$ is a vertex cover, then $(V, E)$ is a star with the center at $v$, and if $\{v_1, v_2\}$ with $v_1 \neq v_2$ is a vertex cover, then the assertion follows by letting $V_0$ be the set of common neighbors of $v_1$ and $v_2$, and for $i \in \{1, 2\}$ defining $V_i$ to be the set of all neighbors of $v_i$ in $V \setminus (V_0 \cup \{v_1, v_2\})$. □

**Lemma 4.** *Let* $\delta \geq 1$ *be an integer, and suppose that* $(V, E)$ *is a graph such that for every edge* $(v_1, v_2) \in E$ *holds* $\deg(v_1) + \deg(v_2) \geq \delta$*. If* $(V, E)$ *has no isolated vertices, then* $|E| \geq (1 - \delta^{-1})|V|$*.*

*Remark.* Equality is attained if $(V, E)$ is a disjoint union of $\delta$-vertex stars.

*Proof of Lemma 4.* For $\delta \leq 2$ the assertion is immediate. Assume therefore that $\delta \geq 3$ and for each $j \in [1, \delta - 2]$ let $V_j := \{v \in V \colon \deg(v) = j\}$; also, let $V_+ := \{v \in V \colon \deg(v) \geq \delta - 1\}$, so that $V$ is the disjoint union of $V_1, \dots, V_{\delta-2}$, and $V_+$. Evidently, we have

$$\sum_{v \in V_j} \deg(v) = j|V_j|; \quad j \in [1, \delta - 2] \tag{1}$$

and

$$\sum_{v \in V_+} \deg(v) \geq (\delta - 1)|V_+|. \tag{2}$$

Also,

$$\sum_{v \in V_+} \deg(v) \geq |V_1|, \tag{3}$$

as every vertex from $V_1$ is adjacent to a vertex from $V_+$. Taking the sum of inequality (2) with weight $2\delta^{-1}$, inequality (3) with weight $1 - 2\delta^{-1}$, and equations (1) with weight 1 for each $j \in [1, \delta - 2]$, we get

$$2|E| = \sum_{v \in V} \deg(v) \geq (2 - 2\delta^{-1})|V_1| + \sum_{j=2}^{\delta-2} j|V_j| + (2 - 2\delta^{-1})|V_+| \geq (2 - 2\delta^{-1})|V|.$$

□

Applying Lemma 4 with $\delta = 2$ to the unique representation graph of a round set, we get

**Corollary 4.** *If $r \geq 1$ is an integer and $A \subseteq \mathbb{F}_2^r$ is a round set, then $|D(A)| \geq \frac{1}{2}|A|$.*

**Lemma 5.** *Let $t$ be the matching number of a graph $(V, E)$. If $(V, E)$ does not have isolated vertices, then*

$$|V| \leq |E| + t.$$

*Proof.* If $T$ is a matching with $t = |T|$, then $(V, E)$ has $|V| - 2t$ vertices, not incident with the edges of $T$. Each of these vertices is incident with an edge from $E \setminus T$, without two vertices sharing an edge. Consequently,

$$|E| = |T| + |E \setminus T| \geq t + (|V| - 2t) = |V| - t.$$

□

Since the matching number of a graph does not exceed the number of its edges, the following corollary strengthens Corollary 4.

**Corollary 5.** *If $r \geq 1$ is an integer and $A \subseteq \mathbb{F}_2^r$ is a round set, then $|A| \leq |D(A)| + t$, where $t$ is the matching number of $\Gamma(A)$.*

## 4. A "light version" of Theorem 1.

In this section we combine the tools, developed so far, to prove

**Theorem 1′.** *Let $r \geq 1$ be an integer. A set $A \subseteq \mathbb{F}_2^r \setminus \{0\}$ with $|A| > \frac{1}{3}2^r + 2$ is minimal 1-saturating if and only if there are a maximal sum-free set $S \subseteq \mathbb{F}_2^r$ and an element $s \in S \cup \{0\}$ such that $A = \big(s + (S \cup \{0\})\big) \setminus \{0\}$.*

Examining the deduction of Theorem 1 from Theorem 2 at the end of Section 2, the reader will see that in a completely identical way Theorem 1′ follows from

**Theorem 2′.** *Let $r \geq 1$ be an integer and suppose that $A \subseteq \mathbb{F}_2^r$ is round. If $|A| > \frac{1}{3}2^r + 2$, then there is a sum-free set $S \subseteq \mathbb{F}_2^r$ and an element $g \in \mathbb{F}_2^r$ such that $A = g + (S \cup \{0\})$.*

Thus, all we need is to prove Theorem 2′.

*Proof of Theorem 2′.* Suppose that $|A| > \frac{1}{3}2^r + 2$. As mentioned in Section 1, the size of a round set in $\mathbb{F}_2^r$ does not exceed $2^{r-1}$. Consequently, the assumptions imply $r \geq 4$; this is implicitly used below when we invoke Proposition 2.

Set $\delta := |A| + |D(A)| - 2^{r-1}$. By Corollary 4, we have

$$\delta \geq \frac{3}{2}|A| - 2^{r-1} > 0;$$

thus, Proposition 2 and Lemma 4 give $\delta|D(A)| \ge (\delta-1)|A|$. Substituting the value of $\delta$ and rearranging the terms, we rewrite this estimate as

$$f(|D(A)|) \le |A|(2^{r-1}+1-|A|),$$

where $f$ is the real function, defined by $f(x) := x(2^{r-1}-x)$.

Since $f$ is concave, $|D(A)| \ge \frac{1}{2}\,|A|$ by Corollary 4, and

$$\min\{f(|A|/2),\, f(|A|-2)\} > |A|(2^{r-1}+1-|A|)$$

(as it follows by a straightforward verification using the assumption on the size of $A$), we have

$$|D(A)| \ge |A| - 1. \tag{4}$$

In view of Lemma 2, it suffices to show that $\Gamma(A)$ has a spanning star; that is (since $\Gamma(A)$ is triangle-free and has no isolated vertices), that the matching number of $\Gamma(A)$ is equal to 1. Suppose, for a contradiction, that $\Gamma(A)$ has a two-edge matching $T$. By Proposition 2, incident with each of the two edges of $T$ are $\delta-2$ edges of $\Gamma(A)$. Moreover, since $\Gamma(A)$ is triangle-free, there are at most two edges of $\Gamma(A)$, incident with both edges of $T$. Consequently, the total number of edges of $\Gamma(A)$ is

$$|D(A)| \ge 2(\delta-2) + |T| - 2 = 2|A| + 2|D(A)| - 2^r - 4.$$

Using (4) and the assumption $|A| > \frac{1}{3}\,2^r + 2$ we derive

$$2^r + 4 \ge 2|A| + |D(A)| \ge 3|A| - 1 > 2^r + 5,$$

a contradiction. □

## 5. Securing two isolated edges.

In this section we prove Theorem 2 under the extra assumption that $\Gamma(A)$ has at most one isolated edge; the case where $\Gamma(A)$ has two or more isolated edges is dealt with in Sections 6 and 7. We split the argument into two lemmas.

**Lemma 6.** *Let $r \ge 1$ be an integer and suppose that $A \subseteq \mathbb{F}_2^r$ is round. If $\Gamma(A)$ has at most one isolated edge and $|A| > 0.3 \cdot 2^r + 2.6$, then the matching number of $\Gamma(A)$ is at most 2.*

The proof is a minor modification of that of Theorem 2′.

*Proof of Lemma 6.* If $\Gamma(A)$ does not have isolated edges then, applying Lemma 4 to the graph $\Gamma(A)$, we get $|D(A)| \ge \frac{2}{3}\,|A|$; if $\Gamma(A)$ has one isolated edge, then $|D(A)| \ge 1 + \frac{2}{3}\,(|A|-2) = \frac{2}{3}\,|A| - \frac{1}{3}$. In any case, letting $\delta := |A| + |D(A)| - 2^{r-1}$ and assuming $|A| > 0.3 \cdot 2^r + 2.6$ we have

$$\delta \ge \frac{5}{3}\,|A| - 2^{r-1} - \frac{1}{3} > 0.$$

Consequently, applying Proposition 2 and Lemma 4 we obtain $\delta|D(A)| \geq (\delta-1)|A|$. Substituting the value of $\delta$, rearranging the terms, and letting $f(x) := (2^{r-1}-x)x$, we re-write this estimate as

$$f(|D(A)|) \leq (2^{r-1}+1-|A|)|A|.$$

We notice that $f(x)$ is concave,

$$\begin{aligned} f\left(\frac{2}{3}|A|-\frac{1}{3}\right) &= \frac{2}{3}|A|\left(2^{r-1}-\frac{2}{3}|A|+\frac{1}{3}\right)-\frac{1}{3}\left(2^{r-1}-\frac{2}{3}|A|+\frac{1}{3}\right) \\ &> \left(\frac{2}{3}2^{r-1}-\frac{4}{9}|A|+\frac{2}{9}\right)|A|-\frac{1}{3}|A| \\ &= \left(\frac{2}{3}2^{r-1}-\frac{4}{9}|A|-\frac{1}{9}\right)|A| \\ &> (2^{r-1}+1-|A|)|A|, \end{aligned}$$

and

$$\begin{aligned} f(|A|-3) &= (2^{r-1}-|A|+3)(|A|-3) \\ &= (2^{r-1}-|A|+1)|A|+5|A|-3\cdot 2^{r-1}-9 \\ &> (2^{r-1}-|A|+1). \end{aligned}$$

Thus, in view of $|D(A)| \geq \frac{2}{3}|A|-\frac{1}{3}$, we conclude that, indeed,

$$|D(A)| \geq |A|-2.$$

Suppose now that $\Gamma(A)$ possesses a three-edge matching $T$. Using Proposition 2 to count the edges of $\Gamma(A)$, incident to those in $T$, and taking into account also the three edges of $T$, we get

$$|D(A)| \geq 3(|A|+|D(A)|-2^{r-1}-2)+3-6;$$

for any edge, incident to two different edges from $T$, joins two vertices from $T$, while being triangle-free and possessing the perfect matching $T$, the graph induced by the six vertices of $T$ has at most six edges, not in $T$. Rearranging the terms gives

$$2^{r-1} \geq |A|+\frac{2}{3}|D(A)|-3 \geq \frac{5}{3}|A|-\frac{13}{3},$$

which contradicts the assumption on $|A|$. □

**Lemma 7.** *Let $r \geq 1$ be an integer. If $A \subseteq \mathbb{F}_2^r$ is a round set with $|A| > 2^{r-2}+3$, then the matching number of $\Gamma(A)$ is distinct from 2.*

*Proof.* Assume for a contradiction that $A \subseteq \mathbb{F}_2^r$ is round, $|A| > 2^{r-2}+3$, and the matching number of $\Gamma(A)$ is equal to 2. By Lemma 3, there exist distinct elements

$a_1, a_2 \in A$ and disjoint subsets $A_0, A_1, A_2 \subseteq A \setminus \{a_1, a_2\}$ such that $A = \{a_1, a_2\} \cup A_0 \cup A_1 \cup A_2$ and

$$D(A) \setminus \{a_1 + a_2\} = \big(a_1 + (A_1 \cup A_0)\big) \cup \big(a_2 + (A_2 \cup A_0)\big).$$

Indeed, $a_1 + a_2 \in D(A)$ holds: else for some $a', a'' \in A_0 \cup A_1 \cup A_2$ we would have $a_1 + a_2 = a' + a''$, contradicting the fact that either $a_1 + a'$, or $a_2 + a'$ is uniquely representable as a sum of two elements of $A$.

By Proposition 1, $\Gamma(A)$ is triangle-free, and consequently, $A_0 = \varnothing$: for $a_1$ and $a_2$ are joined by an edge in $\Gamma(A)$, and therefore have no common neighbors. Hence,

$$D(A) = \{a_1 + a_2\} \cup (a_1 + A_1) \cup (a_2 + A_2),$$

where the union is disjoint by the definition of $D(A)$. For $i \in \{1, 2\}$ we write $D_i := a_i + A_i$ and consider the sets $B_i := \{0, a_1 + a_2\} + D_i$. Since $a_1 + a_2 \in D(A)$ and $D_1, D_2 \subseteq D(A)$, by Proposition 1 we have $B_1 \cap B_2 = \varnothing$ and

$$|B_1| + |B_2| = 2|D_1| + 2|D_2| = 2(|A| - 2) > 2^{r-1}.$$

We claim now that the sumset $B_1 + B_2$ is disjoint with the union $B_1 \cup B_2$; equivalently, $\{0, a_1 + a_2\} + D_1 + D_2$ is disjoint with both $D_1$ and $D_2$. To see this assume, for instance, that $\{0, a_1 + a_2\} + D_1 + D_2$ is not disjoint with $D_1$. As $0 \notin 2D_1 + D_2$ by Proposition 1, this assumption yields $a_1 + a_2 \in 2D_1 + D_2$; that is, $a_1 + a_2 = d_1 + d_1' + d_2$ with some $d_1, d_1' \in D_1$ and $d_2 \in D_2$. Letting $\alpha_i := a_i + d_i$ ($i \in \{1, 2\}$), we re-write this equality as $\alpha_1 + \alpha_2 = d_1'$ and obtain a contradiction observing that $\alpha_1 \in A_1 \subseteq A \setminus \{a_1\}$ and $\alpha_2 \in A_2 \subseteq A \setminus \{a_1\}$, whereas $d_1' \in D_1$ shows that the only representation of $d_1'$ as a sum of two elements of $A$ involves $a_1$ as a summand.

Applying Corollary 2 to the sets $B_1$ and $B_2$, we conclude that one of them is empty. Consequently, either $A_1$, or $A_2$ is empty. Thus, $\Gamma(A)$ is a star, whence the matching number of $\Gamma(A)$ is 1, contrary to an assumption at the beginning of the proof. □

## 6. Using two isolated edges: the coset structure.

As it follows from Lemmas 2, 6, and 7, and since $\frac{11}{36} > 0.3$, to complete the proof of Theorem 2 it remains to consider the case where $\Gamma(A)$ has at least two isolated edges. Accordingly, we assume in this and the next section that $r \geq 1$ is an integer and that $A \subseteq \mathbb{F}_2^r$ is a round set such that $\Gamma(A)$ has two (or more) isolated edges, and show that $|A| < \frac{11}{36}\, 2^r + 3$.

Shifting $A$, if necessary, we assume that $0 \in A$ and $a_1, a_2$, and $a_3$ are elements of $A$, distinct from 0 and each other, such that $(0, a_1)$ and $(a_2, a_3)$ are isolated edges of $\Gamma(A)$. We consider the subgroups $L = \langle a_1, a_2, a_3 \rangle$, $K^- = \langle a_3, a_1 + a_2 \rangle$, $K^+ = \langle a_2, a_1 + a_3 \rangle$, and $H = \langle a_1 + a_2 + a_3 \rangle$; thus,

$$H = K^- \cap K^+,\ |H| = 2,\ |K^-| = |K^+| = 4,\ \text{and } |L| = 8.$$

Our argument is based on a careful study of the distribution of the elements of $A$ and $D(A)$ in the cosets of $L$. The goal of the present section is to establish some basic facts about this distribution.

For $g \in \mathbb{F}_2^r$ we write $A_g := A \cap (g + L)$ and $D_g := D(A) \cap (g + L)$. Evidently, we have $\{0, a_1, a_2, a_3\} \subseteq A_0$, and it is easy to see that, indeed, $A_0 = \{0, a_1, a_2, a_3\}$. Next, if $g \notin L$, then from $\{a_1, a_2 + a_3\} \subseteq (2A_0) \cap D(A)$ it follows that

$$(2A_g) \cap \{a_1, a_2 + a_3\} = \varnothing, \tag{5}$$

and the fact that $(0, a_1)$ and $(a_2, a_3)$ are isolated edges gives

$$(A_g + D_g) \cap \{0, a_1, a_2, a_3\} = \varnothing. \tag{6}$$

Furthermore, using Proposition 1 we obtain

$$(2D_g) \cap \{a_1, a_2 + a_3\} = \varnothing, \tag{7}$$

for each $g \in \mathbb{F}_2^r$.

An immediate corollary of (5) is that $|A_g| \le 4$ holds for every element $g \in \mathbb{F}_2^r$. With this in mind, for $g \in \mathbb{F}_2^r$ and $i \in [0, 4]$ we say that the coset $g + L$ is of type $i$ if $|A_g| = i$, and we denote by $n_i$ the number of *non-zero* cosets of type $i$ (so that $L$ is not counted in $n_4$); hence,

$$n_0 + n_1 + n_2 + n_3 + n_4 = 2^{r-3} - 1 \tag{8}$$

and

$$n_1 + 2n_2 + 3n_3 + 4n_4 = |A| - 4. \tag{9}$$

Recall, that for a subgroup $F$ of an abelian group $G$, by $\varphi_F$ we denote the canonical homomorphism from $G$ onto the quotient group $G/F$. Thus, if $B$ is a subset of $G$, then $|\varphi_F(B)|$ is the number of cosets of $F$, intersecting $B$ non-trivially.

For the proof of the following three claims the reader is strongly encouraged to use a pictorial "parallelepiped" representation of the cosets of $L$.

**Claim 1.** *For every $g \in \mathbb{F}_2^r \setminus L$ we have* $\min\{|\varphi_{K^-}(A_g)|, |\varphi_{K^+}(A_g)|\} \le 1$.

*Proof.* Without loss of generality, we assume $g \in A_g$, whence $g + a_1 \notin A_g$ and $g + a_2 + a_3 \notin A_g$ by (5). Thus, if $A_g$ has non-empty intersection with both cosets of $K^-$, contained in $g+L$, then $A_g$ contains at least one of the elements $g+a_2$ and $g+a_1+a_3$. Similarly, if $A_g$ has non-empty intersection with both cosets of $K^+$, contained in $g+L$, then $A_g$ contains at least one of the elements $g + a_3$ and $g + a_1 + a_2$. However, the two conditions just mentioned cannot hold simultaneously, as the sum of an element from $\{g + a_2, g + a_1 + a_3\}$ and an element from $\{g + a_3, g + a_1 + a_2\}$ is either $a_1$, or $a_2 + a_3$; cf. (5). □

Refining our classification of cosets of $L$, for $i \in [2,4]$ and $g \in \mathbb{F}_2^r$ we say that the coset $g+L$ is of type $i^0$ if it is of type $i$ and, in addition,

$$|\varphi_{K^-}(A_g)| = |\varphi_{K^+}(A_g)| = 1;$$

that $g+L$ is of type $i^-$ if it is of type $i$ and, in addition,

$$|\varphi_{K^+}(A_g)| > |\varphi_{K^-}(A_g)| = 1;$$

and finally, that $g+L$ is of type $i^+$ if it is of type $i$ and

$$|\varphi_{K^-}(A_g)| > |\varphi_{K^+}(A_g)| = 1.$$

Let $n_i^0, n_i^-$, and $n_i^+$ denote the number of non-zero cosets of the corresponding types. From this definition, Claim 1, and the observation that if $|\varphi_{K^-}(A_g)| = |\varphi_{K^+}(A_g)| = 1$, then $|\varphi_H(A_g)| = 1$ and thus $|A_g| \le 2$, it follows that

$$n_2 = n_2^0 + n_2^- + n_2^+,\ n_3^0 = n_4^0 = 0,\ n_3 = n_3^- + n_3^+,\ \text{and } n_4 = n_4^- + n_4^+.$$

**Claim 2.** *For every $g \in \mathbb{F}_2^r$ we have*

$$|D_g| = \begin{cases} 0, & \text{if } g+L \text{ is of type } 2^0,\ 3, \text{ or } 4, \text{ and } g \notin L; \\ 2, & \text{if } g \in L; \end{cases}$$

*furthermore,*

$$|D_g| \le \begin{cases} 2, & \text{if } g+L \text{ is of type } 1, 2^-, \text{ or } 2^+; \\ 4 & \text{if } g+L \text{ is of type } 0. \end{cases}$$

*Proof.* If $g+L$ is of type $2^0$, then $A_g$ is a coset of $H$, and without loss of generality we assume that $A_g = g+H$. By (6), the assumption $g \in A$ yields $\{g, g+a_1, g+a_2, g+a_3\} \cap D_g = \varnothing$. Similarly, $g+a_1+a_2+a_3 \in A$ yields $\{g+a_1+a_2+a_3, g+a_2+a_3, g+a_1+a_3, g+a_1+a_2\} \cap D_g = \varnothing$. Thus, $D_g = \varnothing$.

Suppose now that $|A_g| \ge 3$, $g \notin L$ and show that in this case $D_g = \varnothing$, too. If $A$ contains a coset of $H$, we argue exactly as above. Assume therefore that $A$ does not contain a coset of $H$, and assume also, for definiteness, that $g \in A$. By (5), we have $\{g+a_1, g+a_2+a_3\} \cap A_g = \varnothing$ and $a_1 \notin 2A_g$. It is easy to derive that either $A_g = \{g, g+a_2, g+a_3\}$, or $A_g = \{g, g+a_1+a_2, g+a_1+a_3\}$ holds; however, each of these options is incompatible with $a_2+a_3 \notin 2A_g$, following from (5).

By (7) and since $\{a_1, a_2+a_3\} \subseteq D(A)$, the set $D(A)$ is disjoint with $\{0, a_1+a_2+a_3\}$, and the assumption that the edge $(a_2, a_3)$ is isolated shows that $D(A)$ is also disjoint with $\{a_2, a_3, a_1+a_2, a_1+a_3\}$; consequently, if $g \in L$, then $D_g = \{a_1, a_2+a_3\}$ and $|D_g| = 2$.

Next, if $g \in A$ and $g \notin L$, then by (6) we have $\{g, g+a_1, g+a_2, g+a_2, g+a_3\} \cap D_g = \varnothing$. Also, (7) shows that $D_g$ can possibly contain at most one of $g+a_2+a_3$ and

$g+a_1+a_2+a_3$, and similarly $D_g$ can possibly contain at most one of $g+a_1+a_3$ and $g+a_1+a_2$. It follows that $|D_g|\le 2$ whenever $g\notin L$ and $g+L$ is not of type 0.

Finally, the fact that $|D_g|\le 4$ for each $g\in\mathbb{F}_2^r$ is a direct consequence of (7) and the box principle. □

**Claim 3.** *For every $g\in\mathbb{F}_2^r$ such that $g+L$ is of type $1, 2^-$, or $2^+$ there exists a subset $\widetilde{D}_g\subseteq g+L$ with $D_g\subseteq\widetilde{D}_g$ and $|\widetilde{D}_g|=|\varphi_H(\widetilde{D}_g)|$; moreover,*

(i) *if $g+L$ is of type 1, then $|\widetilde{D}_g|=4$;*
(ii) *if $g+L$ is of type $2^-$, then $|\widetilde{D}_g|=2$ and $|\varphi_{K^-}(\widetilde{D}_g)|=1$;*
(iii) *if $g+L$ is of type $2^+$, then $|\widetilde{D}_g|=2$ and $|\varphi_{K^+}(\widetilde{D}_g)|=1$.*

*Proof.* If $g+L$ is not of type 0 and $g\notin L$ then, by (6), the set $D_g$ is disjoint with the set $A_g+A_0\subseteq g+L$, containing a translate of $A_0$. However, $A_0$ intersects non-trivially each of the four cosets of $H$, contained in $L$. Thus, the complement of $D_g$ in $g+L$ contains an element in each coset of $H$, contained in $g+L$. This shows the existence of $\widetilde{D}\subseteq g+L$ with $D_g\subseteq\widetilde{D}_g$ and $|\widetilde{D}_g|=|\varphi_H(\widetilde{D}_g)|$, and proves (i).

Now suppose that $g+L$ is of type $2^-$. We assume, without loss of generality, that $g\in A$ and consequently, that either $A_g=\{g,g+a_3\}$, or $A_g=\{g,g+a_1+a_2\}$ holds. By (6), the set $D_g$ is contained in the complement of $A_g+A_0$ in $g+L$, which in the former case is $\{g+a_1+a_2, g+a_1+a_2+a_3\}$, and in the latter case $\{g+a_1+a_3, g+a_2+a_3\}$. To prove (ii) it remains to observe that each of these sets is contained in a coset of $K^-$, but not contained in a coset of $H$.

The proof of (iii) goes along similar lines. □

## 7. Using two isolated edges: completion of the proof.

In this section we complete the proof of Theorem 2. We keep using the notation and assumptions of the previous section, and since $\frac{11}{36}\,2^r+3>\frac{1}{3}\,2^r+2$ for $r\in[1,5]$, in view of Theorem 2′ we assume $r\ge 6$. To argue by contradiction, we also assume that $|A|>\frac{11}{36}\,2^r+3$. Our goal is to show that these assumptions are inconsistent.

**Claim 4.** *We have* $\min\{n_4^-,n_4^+\}<n_0+3$.

*Proof.* Suppose that $n_4^-\ge n_0+3$ and $n_4^+\ge n_0+3$, and let $A^-$ denote the union of all sets $A_g$ such that $g+L$ is of type $4^-$. Since $|\varphi_L(A)|=2^{r-3}-n_0$ and $|\varphi_L(A^-)|=n_4^-\ge n_0+3$, by the box principle every element of $\mathbb{F}_2^r/L$ is representable in at least three ways as a sum of an element from $\varphi_L(A)$ and an element from $\varphi_L(A^-)$. Observing that $A^-$ is a union of $K^-$-cosets and that each $L$-coset is a union of two $K^-$-cosets, we conclude that every $L$-coset contains a $K^-$-coset, disjoint from $D(A)$. Similarly, every $L$-coset contains a $K^+$-coset, disjoint with $D(A)$. As a union of a $K^-$-coset and a $K^+$-coset, contained in the same $L$-coset, covers all this $L$-coset, with the exception

of an $H$-coset, applying Claim 3 we conclude that $|D_g| \le 1$ if $g+L$ is of type 1 or 2, and $|D_g| \le 2$ if $g+L$ is of type 0. Hence, using (8) and (9), we obtain

$$\begin{aligned} |D(A)| &\le 2n_0 + n_1 + n_2 + 2 \\ &\le 2(n_0+n_1+n_2+n_3+n_4) - \frac{1}{2}\,(n_1+2n_2+3n_3+4n_4) + 2 \\ &= 2^{r-2} - \frac{1}{2}\,|A| + 2. \end{aligned}$$

Compared with Corollary 4 this yields $|A| \le 2^{r-2}+2$, a contradiction. □

Being the only place where the factor $\frac{11}{36}$ emerges, the following claim can be considered the bottleneck of our method.

**Claim 5.** *We have* $\max\{n_4^-, n_4^+\} < n_0 + 3$.

*Proof.* Switching the notation, if necessary, and in view of Claim 4, we assume that

$$n_4^- < n_0 + 3 \le n_4^+. \tag{10}$$

Let $A^+$ be the union of all sets $A_g$ such that $g+L$ of type $4^+$. As in the proof of Claim 4, every element of $\mathbb{F}_2^r/L$ is representable in at least three ways as a sum of an element from $\varphi_L(A)$ and an element from $\varphi_L(A^+)$, and $A^+$ is a union of $K^+$-cosets; hence every $L$-coset contains a $K^+$-coset, disjoint from $D(A)$. In view of Claim 3 (ii), we have $|D_g| \le 1$ whenever $g+L$ is of type $2^-$. Thus, by Claim 2,

$$|D(A)| \le 4n_0 + 2n_1 + n_2^- + 2n_2^+ + 2. \tag{11}$$

Let $B$ denote the set of all those elements of $A$, adjacent in $\Gamma(A)$ to an element from $A^+$. As $A^+$ is a union of $K^+$-cosets, for any $b \in B$ we have $|(b+K^+) \cap A| = 1$; it follows that $B$ is disjoint with $A^+$ and

$$|B| \le n_1 + 2n_2^- + n_3^-.$$

Consider the subgraph $\Gamma'$ of $\Gamma(A)$, induced by the elements of $A^+ \cup B$. Since $B$ is a vertex cover of $\Gamma'$, the matching number $t'$ of $\Gamma'$ does not exceed $|B|$; hence,

$$t' \le n_1 + 2n_2^- + n_3^-. \tag{12}$$

Let $t$ be the matching number of $\Gamma(A)$ and let $T$ be a matching in $\Gamma(A)$ with $|T| = t$. As the number of edges between $A^+$ and $B$ in $\Gamma(A)$ is at least $|A^+| = 4n_4^+$, at most $t'$ of these edges being in $T$, we have

$$t \le |D(A)| - 4n_4^+ + t'. \tag{13}$$

To obtain another relation between $t$ and $t'$ we notice that if $b \in B$ is adjacent in $\Gamma(A)$ to $a \in A^+$, then in fact every element of $D_{a+b}$ corresponds to an edge, incident with $b$: for, all elements of $D_{a+b}$ are contained in a coset of $K^+$, and since $a+b \in D_{a+b}$, this coset is $a+b+K^+ = b+A_a$. Now, fix a matching $T'$ of $\Gamma'$. As any edge of $T$

corresponds to an element from $D(A)$, by Claim 2 corresponding to the edges of $T$ are are at most four elements from every $L$-coset of type 0, two elements from $L$, and at most two elements from every $L$-coset of type $1, 2^-$, or $2^+$. Taking into account that for each edge $(a, b)$ of $T'$, there is actually at most one element in the coset $a+b+L$, corresponding to an edge of $T$, we conclude that

$$t \le 4n_0 + 2n_1 + 2n_2^- + 2n_2^+ + 2 - t'. \tag{14}$$

We complete the proof of Claim 5 showing that an appropriate combination of estimates (10)–(14) yields a contradiction with the assumption on the size of $A$, made at the beginning of this section. Specifically, substituting (9) into the estimate of Corollary 5 we get

$$n_1 + 2n_2 + 3n_3 + 4n_4 \le |D(A)| + t - 4.$$

Taking the sum of this inequality with the weight 4, identity (8) with the weight 44, the first inequality in (10) in the form $-n_0 + n_4^- \le 2$ with the weight 12, and inequalities (11), (12), (13), and (14) with the weights 7, 2, 3, and 1, respectively, we obtain

$$30n_1 + 52n_2^0 + 39n_2^- + 36n_2^+ + 54n_3^- + 56n_3^+ + 72n_4^- + 72n_4^+ \le 44 \cdot 2^{r-3} - 20.$$

(The weights were found by solving the corresponding linear program to yield the best possible bound.) Evidently, the left-hand side is at least as large as

$$18(n_1 + 2n_2 + 3n_3 + 4n_4) = 18(|A| - 4);$$

thus

$$|A| \le \frac{22}{9} \cdot 2^{r-3} - \frac{10}{9} + 4 < \frac{11}{36} \cdot 2^r + 3,$$

a contradiction. □

**Claim 6.** *We have* $n_4 \ge n_0 + n_1 + n_2^0 + 4$.

*Proof.* Applying Claim 2 and using (8) and (9) we get

$$\begin{aligned} |D(A)| &\le 4n_0 + 2n_1 + 2n_2^- + 2n_2^+ + 2 \\ &= 6(n_0 + n_1 + n_2 + n_3 + n_4) - 2(n_1 + 2n_2 + 3n_3 + 4n_4) \\ &\qquad + 2(n_4 - n_0 - n_1 - n_2^0 + 1) \\ &= 3 \cdot 2^{r-2} - 2|A| + 2(n_4 - n_0 - n_1 - n_2^0 + 2). \end{aligned}$$

Comparing with Corollary 4, we get

$$\frac{1}{2}\,|A| \le 3 \cdot 2^{r-2} - 2|A| + 2(n_4 - n_0 - n_1 - n_2^0 + 2).$$

Hence

$$n_4 - n_0 - n_1 - n_2^0 \ge \frac{5}{4}|A| - 3 \cdot 2^{r-3} - 2,$$

and the result follows from $|A| > 0.3 \cdot 2^r + 4$. $\square$

**Claim 7.** *The matching number of* $\Gamma(A)$ *does not exceed* $n_1 + 2n_2^- + 2n_2^+ + n_3 + 2$.

*Proof.* Write $\sigma := a_1 + a_2 + a_3$, so that $H = \{0, \sigma\}$. Clearly, if $(a, b)$ is an edge in $\Gamma(A)$, then either $a + \sigma \notin A$, or $b + \sigma \notin A$. This shows that every edge of $\Gamma(A)$ is incident with an element of the set $B := \{a \in A \colon a + \sigma \notin A\}$. By Claim 1, the total number of elements of $B$ is

$$n_1 + 2n_2^- + 2n_2^+ + n_3 + 4.$$

It remains to notice that in a matching of $\Gamma(A)$, there can be at most one edge, incident with a given element of $B$, and that any maximal matching constants the edges $(0, a_1)$ and $(a_2, a_3)$, both incident to two elements of $B$. $\square$

By Claims 5 and 6, it remains to consider the case where

$$n_4^- \le n_0 + 2, \ n_4^+ \le n_0 + 2 \tag{15}$$

and

$$n_4 \ge n_0 + n_1 + 4, \tag{16}$$

which from now on we assume to hold. Notice, that these assumptions imply

$$\min\{n_4^-, n_4^+\} \ge 2.$$

We define $A^-$ to be the union of all sets $A_g$ such that $g + L$ is of type $4^-$, and $A_g^+$ — the union of those $A_g$ with $g + L$ of type $4^+$. Next, let $B$ be the union of all $A_g$ with $|A_g| \ge 2$; thus,

$$|\varphi_L(A^-)| = n_4^-, \ |\varphi_L(A^+)| = n_4^+, \text{ and } |\varphi_L(B)| = n_2 + n_3 + n_4 + 1.$$

Furthermore, let $C^-$ denote the set of all those $g \in A^- + B$ with the property that $\varphi_L(g)$ has at least two representations as an element from $\varphi_L(A^-)$ and an element from $\varphi_L(B)$; similarly, denote by $C^+$ the set of those $g \in A^+ + B$ such that $\varphi_L(g)$ has at least two representations as an element from $\varphi_L(A^+)$ and an element from $\varphi_L(B)$.

Notice, that if $g \in C^- + L$, then $g + L$ contains a coset of $K^-$, disjoint from $D(A)$, and if $g \in C^+ + L$, then $g + L$ contains a coset of $K^+$, disjoint from $D(A)$. Hence, for $g \in (C^- + L) \cap (C^+ + L)$, the set $D_g$ is contained in a coset of $H$; thus by Claim 3 if $g + L$ is of type 1 or 2, then $|D_g| \le 1$, and if $g + L$ is of type 0, then $|D_g| \le 2$.

By the box principle, we have

$$|\varphi_L(C^-) \cap \varphi_L(C^+)| \ge |\varphi_L(C^-)| + |\varphi_L(C^+)| - 2^{r-3}, \tag{17}$$

whereas by Corollary 3,

$$|\varphi_L(C^-)| \ge \min\{2n_2 + 2n_3 + 2n_4 + 2n_4^- - 2^{r-3} - 2, n_2 + n_3 + n_4\}$$

and

$$|\varphi_L(C^+)| \geq \min\{2n_2 + 2n_3 + 2n_4 + 2n_4^+ - 2^{r-3} - 2, n_2 + n_3 + n_4\}. \tag{18}$$

We notice that at least one of these minima is attained on its second term: for if

$$2n_2 + 2n_3 + 2n_4 + 2n_4^- - 2^{r-3} - 2 \leq n_2 + n_3 + n_4$$

and

$$2n_2 + 2n_3 + 2n_4 + 2n_4^+ - 2^{r-3} - 2 \leq n_2 + n_3 + n_4$$

both hold true, then taking the sum we obtain

$$2n_2 + 2n_3 + 4n_4 \leq 2^{r-2} + 4$$

which, in view of (8), can be re-written as

$$n_4 \leq n_0 + n_1 + 3;$$

this, however, is inconsistent with (16).

By symmetry, we can assume that

$$|\varphi_L(C^-)| \geq n_2 + n_3 + n_4, \tag{19}$$

and we consider two cases, according to the value in the right-hand side of (18).

If $|\varphi_L(C^+)| \geq n_2 + n_3 + n_4$, then by (17) and (19),

$$|\varphi_L(C^-) \cap \varphi_L(C^+)| \geq 2n_2 + 2n_3 + 2n_4 - 2^{r-3}.$$

It follows that there are at least

$$(2n_2 + 2n_3 + 2n_4 - 2^{r-3}) - (n_3 + n_4 + 1) = 2n_2 + n_3 + n_4 - 2^{r-3} - 1$$

$L$-cosets of type $0, 1$, or $2$, contained in $(C^- + L) \cap (C^+ + L)$. Hence, by Claim 2 and an observation above,

$$\begin{aligned} |D(A)| &\leq 4n_0 + 2n_1 + 2n_2 + 2 - (2n_2 + n_3 + n_4 - 2^{r-3} - 1) \\ &= 4n_0 + 2n_1 - n_3 - n_4 + 2^{r-3} + 3. \end{aligned}$$

Combining this estimate with Corollary 5 and Claim 7, we get

$$|A| \leq 4n_0 + 3n_1 + 2n_2^- + 2n_2^+ - n_4 + 2^{r-3} + 5$$

and furthermore, substituting the value of $|A|$ from (9),

$$-4n_0 - 2n_1 + 2n_2^0 + 3n_3 + 5n_4 \leq 2^{r-3} + 1.$$

Taking the sum of this inequality, inequalities (15), and identity (8), multiplied by 6, we obtain

$$4n_1 + 2n_2^0 + 6n_2 + 9n_3 + 12n_4 \leq 7 \cdot 2^{r-3} - 1.$$

By (9), the expression in the left-hand side is at least $3(|A|-4)$; consequently,

$$|A| \le \frac{7}{24}\,2^r - \frac{1}{3} + 4 < \frac{11}{36}\,2^r + 3,$$

a contradiction.

Finally, suppose that $|\varphi_L(C^+)| \ge 2n_2 + 2n_3 + 2n_4 + 2n_4^+ - 2^{r-3} - 2$. Arguing as in the previous case, we get

$$\begin{aligned}
|\varphi_L(C^-) \cap \varphi_L(C^+)| &\ge 3n_2 + 3n_3 + 3n_4 + 2n_4^+ - 2^{r-2} - 2,\\
|D(A)| &\le 4n_0 + 2n_1 + 2n_2 + 2\\
&\qquad - \big((3n_2 + 3n_3 + 3n_4 + 2n_4^+ - 2^{r-2} - 2) - (n_3 + n_4 + 1)\big)\\
&= 4n_0 + 2n_1 - n_2 - 2n_3 - 2n_4 - 2n_4^+ + 2^{r-2} + 5,\\
|A| &\le 4n_0 + 3n_1 - n_2^0 + n_2^- + n_2^+ - n_3 - 2n_4 - 2n_4^+ + 2^{r-2} + 7,
\end{aligned}$$

and hence

$$-4n_0 - 2n_1 + 3n_2^0 + n_2^- + n_2^+ + 4n_3 + 6n_4 + 2n_4^+ \le 2^{r-2} + 3.$$

Taking the sum of the last inequality, the first of the inequalities (15), and identity (8), multiplied by 5, we obtain

$$3n_1 + 2n_2^0 + 6n_2 + 9n_3 + 12n_4 + n_4^+ \le 7 \cdot 2^{r-3} - 2,$$

which leads to a contradiction as above. This completes the proof of Theorem 2.

## References

[DMP03] A.A. Davydov, S. Marcugini, and F. Pambianco, On saturating sets in projective spaces, *J. Combin. Theory, Ser. A* **103** (1) (2003), 1–15.

[DMP06] A.A. Davydov, S. Marcugini, and F. Pambianco, Minimal 1-saturating sets and complete caps in binary projective spaces, *J. Combin. Theory, Ser. A* **113** (2006), 647–663.

[DT89] A.A. Davydov and L.M. Tombak, Quasi-perfect linear binary codes with distance 4 and complete caps in projective geometry, *Problemy Peredachi Informatzii* **25** (4) (1989), 11–23.

[G] D.J. Grynkiewicz, On extending Pollard's theorem for t-representable sums, *submitted*.

[Kn53] M. Kneser, Abschätzung der asymptotischen Dichte von Summenmengen, *Math. Z.* **58** (1953), 459–484.

[Kn55] M. Kneser, Ein Satz über abelsche Gruppen mit Anwendungen auf die Geometrie der Zahlen, *Math. Z.* **61** (1955), 429–434.

[Mn76] H.B. Mann, Addition theorems: the addition theorems of group theory and number theory. Robert E. Krieger Publishing Co., Huntington, N.Y., 1976.

*E-mail address*: diambri@hotmail.com

(To fill in)

*E-mail address*: seva@math.haifa.ac.il

Department of Mathematics, The University of Haifa at Oranim, Tivon 36006, Israel