1-SATURATING SETS, CAPS AND DOUBLING-CRITICAL SETS IN BINARY SPACES

DAVID J. GRYNKIEWICZ AND VSEVOLOD F. LEV

Abstract. We show that, for a positive integer r, every minimal 1-saturating set in $\mathrm{PG}(r-1,2)$ of size at least $\frac{11}{36}\,2^r+3$ either is a complete cap or can be obtained from a complete cap S by fixing some $s\in S$ and replacing every point $s'\in S\setminus\{s\}$ by the third point on the line through s and s'. Since, conversely, every set obtained in this way is a minimal 1-saturating set, and the structure of large sum-free sets in an elementary abelian 2-group is known, this provides a complete description of large minimal 1-saturating sets.

An algebraic restatement is as follows. Suppose that G is an elementary abelian 2-group and a subset $A\subseteq G\setminus\{0\}$ satisfies $A\cup 2A=G$ and is minimal subject to this condition. If $|A|\ge\frac{11}{36}\,|G|+3$, then either A is a maximal sum-free set, or there are a maximal sum-free set $S\subseteq G$ and an element $s\in S$ such that $A=\{s\}\cup\big(s+(S\setminus\{s\})\big)$.

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

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. Given an integer $\rho\ge 1$, one says that A is *ρ-saturating* if every point of $\mathrm{PG}(r-1,q)$ is contained in a subspace generated by $\rho+1$ points from 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 property of being ρ-saturating is inherited by supersets and that of being a cap is inherited by subsets, of particular interest are minimal ρ-saturating sets and complete caps.

In this paper, we are concerned with the case $\rho=1$ and the space $\mathrm{PG}(r-1,2)$ whose points are, essentially, non-zero elements of the elementary abelian 2-group of rank r, and whose lines are triples of points adding up to 0. A large random set in $\mathrm{PG}(r-1,2)$ is 1-saturating with very high probability, but the probability that it is a *minimal* 1-saturating set is extremely low; thus, one can expect that large minimal 1-saturating

2000 *Mathematics Subject Classification.* 51E20, 11B75, 11P70.

Key words and phrases. Additive combinatorics, blocking set, cap, saturating set, sum-free set, sumset.

The first author was supported by the Austrian Science Fund FWF (Project Number M1014-N13).

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 seems to be considerably more subtle, which is quite natural bearing in mind that complete caps in $\mathrm{PG}(r-1,2)$ can be characterized as those 1-saturating sets possessing the extra property of having no internal lines (as will be explained shortly).

With the exception of the next section where our result is discussed from the projective geometric viewpoint, we mostly use the language of abelian groups. Accordingly, denoting by $\mathbb{F}_2^r$ the elementary abelian 2-group of rank $r \ge 1$ and writing

$$2A := \{a_1 + a_2 \colon a_1, a_2 \in A\}$$

for a subset $A \subseteq \mathbb{F}_2^r$, we interpret 1-saturating sets in $\mathrm{PG}(r-1,2)$ as those subsets $A \subseteq \mathbb{F}_2^r \setminus \{0\}$ satisfying $A \cup 2A = \mathbb{F}_2^r$. Similarly, caps in $\mathrm{PG}(r-1,2)$ are interpreted as sets $A \subseteq \mathbb{F}_2^r \setminus \{0\}$ 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, they satisfy $A \cup 2A = \mathbb{F}_2^r$. Consequently, maximal sum-free sets are minimal 1-saturating sets without internal lines. 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 a minimal 1-saturating set;

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

An extension of construction (i) has just been mentioned: any maximal sum-free set is a minimal 1-saturating set. Construction (ii) can be extended by observing that if s is an element of a maximal sum-free set $S \subseteq \mathbb{F}_2^r$, then $A := \{s\} \cup \big((S+s) \setminus \{0\}\big)$ is a minimal 1-saturating set: 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 $B \cup 2B \ne \mathbb{F}_2^r$.

Indeed, a common description can be given to these two extensions: namely, if $S \subseteq \mathbb{F}_2^r$ is a maximal sum-free set and $s \in S \cup \{0\}$, then $A := \big(s + (S \cup \{0\})\big) \setminus \{0\}$ is a minimal 1-saturating set. 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 $r \geq 1$ be an integer. A set $A \subseteq \mathbb{F}_2^r \setminus \{0\}$ with $|A| > \frac{11}{36}\,2^r + 3$ is a minimal 1-saturating set 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 $A = \big(s + (S \cup \{0\})\big) \setminus \{0\}$.*

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. We record the following easy corollary of their result.

Fact 2 ([DT89])**.** *Let $r \geq 1$ be an integer. Every maximal sum-free set in $\mathbb{F}_2^r$ of size larger than $9 \cdot 2^{r-5}$ either is the non-zero coset of an index-2 subgroup, or has 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 five-element set with the elements adding up to 0 such that $\mathbb{F}_2^r = \langle B \rangle \oplus H$.*

In the statement of Fact 2 and below in the paper, for a set B of group elements, we write $\langle B \rangle$ to 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 $|A| \geq 2^{r-2} + 3$, provided that $r \geq 6$. (The group $\mathbb{F}_2^5$ contains an 11-element minimal 1-saturating set, but no 11-element maximal sum-free sets; see [DMP06].) If true, this is best possible.

Example 3. 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 a minimal 1-saturating set. Now, 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, whence $2S = H$ in view of $|H| \geq 4$, and therefore S 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 subgroups with $\mathbb{F}_2^r = F \oplus H$ and $|F|, |H| \geq 4$, then $(F \cup H) \setminus \{0\}$ is a minimal 1-saturating set which cannot be represented as in Theorem 1.

2. The Projective Geometry Viewpoint

We remark that Theorem 1 can be reformulated in purely geometrical terms, as in the abstract.

Theorem 1a. *For an integer $r \ge 1$, every minimal 1-saturating set in $\mathrm{PG}(r-1,2)$ of size at least $\frac{11}{36}\,2^r+3$ either is a complete cap, or can be obtained from a complete cap S by fixing some $s \in S$ and replacing every point $s' \in S \setminus \{s\}$ by the third point on the line through s and s'.*

Another reformulation, kindly pointed out by Simeon Ball, involves blocking sets. Recall that a set of points in a projective geometry is called a *blocking set* if it has a non-empty intersection with every line; consequently, a set in $\mathrm{PG}(r-1,2)$ is a (minimal) blocking set if and only if its complement is a (complete) cap. It is easy to derive that Theorem 1 is equivalent to the following assertion.

Theorem 1b. *For an integer $r \ge 1$, every minimal 1-saturating set A in $\mathrm{PG}(r-1,2)$ of size at least $\frac{11}{36}\,2^r+3$ either is the complement of a minimal blocking set, or can be obtained from a minimal blocking set B by fixing a point $s \notin B$ and letting A consist of s along with all points $b \in B$ for which the line through s and b is tangent to B (i.e., passes through precisely one point of B).*

The equivalence between Theorems 1 and **??**b is a simple exercise, left to the reader.

Yet another consequence of Theorem 1 concerns the spectrum of possible sizes of minimal 1-saturating sets. As indicated in Section 1, the largest size of a minimal 1-saturating set in $\mathrm{PG}(r-1,2)$ is 2^{r-1}. The *second largest* size can be determined as an immediate corollary of Theorem 1 and Fact 2.

Corollary 4. *If $r \ge 9$ is an integer, then the second largest size of a minimal 1-saturating set in $\mathrm{PG}(r-1,2)$ is $5 \cdot 2^{r-4}$, and the third largest size is smaller than $\frac{11}{36}\,2^r+3$.*

It is observed in [DMP06] that, with a single exception for $r=5$, the spectrum of sizes of all known large minimal 1-saturating sets in $\mathrm{PG}(r-1,2)$ is contained in the spectrum of sizes of sum-free sets in $\mathrm{PG}(r-1,2)$. Theorem 1 and its above-mentioned conjectured strengthening provide, of course, an explanation to this phenomenon.

Finally, we note that Theorem 1 allows one to find all classes of projectively equivalent minimal 1-saturating sets in $\mathrm{PG}(r-1,2)$. For, it is not difficult to derive from Fact 2 that if $r \ge 6$ is an integer, S_1 and S_2 are (potentially identical) complete caps in $\mathrm{PG}(r-1,2)$ with $|S_1| = |S_2| > 9 \cdot 2^{r-5}$, and, for $i \in \{1,2\}$, the sets S_i' are obtained from S_i as described in Theorem **??**a, then S_1 and S_2 are projectively equivalent, as are S_1' and S_2', while S_1 is not equivalent to S_2'—regardless of the specific choice of the elements fixed in S_1 and S_2 to get S_1' and S_2'. (For the non-equivalence, one only needs to note that S_2' is not a cap for $r \ge 6$.) This leads to the following corollary.

Corollary 5. *For a positive integer $r \ge 9$, there are four projectively non-isomorphic minimal 1-saturating sets in $\mathrm{PG}(r-1,2)$ of size larger than $\frac{11}{36}\,2^r+3$: two are complete*

caps of sizes 2^{r-1} *and* $5 \cdot 2^{r-4}$*, and two more are obtained from them as in Theorem* **??***a.*

3. Doubling-Critical 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 doubling-critical 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 doubling-critical sets and unique representation graphs. Finally, we state a structure theorem for doubling-critical sets (Theorem 7 below) and show that it implies Theorem 1.

The remainder of the paper is structured as follows. Important auxiliary results are gathered in Section 4. In Section 5, 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 ($r \le 5$), it serves as a simplified model of our method, exhibiting many of the core ideas. Sections 6–8 are devoted to the proof of Theorem 7: in Section 6, the problem is reduced to the case where the unique representation graph has at least two isolated edges, Sections 7 and 8 present a treatment of this case.

Let $r \ge 1$ be an integer. We say that a set $A \subseteq \mathbb{F}_2^r$ is *doubling-critical* if, for every proper subset $B \subset A$, we have $2B \ne 2A$; 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 $A \subseteq \mathbb{F}_2^r \setminus \{0\}$ is a minimal 1-saturating set 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 further.

Lemma 6. *Let* $r \ge 1$ *be an integer. If* $A \subseteq \mathbb{F}_2^r \setminus \{0\}$ *is a minimal* 1*-saturating set, then either* A *or* $A \cup \{0\}$ *is doubling-critical.*

Remark. It is easy to derive from Theorem 1 and the observation following the proof below that if $A \subseteq \mathbb{F}_2^r \setminus \{0\}$ is a large minimal 1-saturating set, then, indeed, $A \cup \{0\}$ is doubling-critical.

Proof of Lemma 6. Suppose that $A \subseteq \mathbb{F}_2^r$ is a minimal 1-saturating set. If $A \cup \{0\}$ is not doubling-critical, 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 the minimality of A, we actually have $a_0 = 0$, whence $2A = \mathbb{F}_2^r$. Now if also A is not doubling-critical, 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 the minimality of A. □

Lemma 6 allows us to concentrate on studying large doubling-critical sets instead of large 1-saturating sets; indeed, we will hardly ever refer to 1-saturating sets from now on, except for the deduction of Theorem 1 from Theorem 7 at the end of this section.

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 doubling-critical. To verify this, we can assume $g = 0$ (as the property of being doubling-critical 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 $s_0 \in 2(S \cup \{0\})$. The heart of our paper is the following theorem, showing that, in fact, any large doubling-critical set has the structure just described.

Theorem 7. *Let $r \ge 1$ be an integer and suppose that $A \subseteq \mathbb{F}_2^r$ is doubling-critical. 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 doubling-critical 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 8. *Let $r \ge 1$ be an integer, let $g \in \mathbb{F}_2^r$, and suppose that $A \subseteq \mathbb{F}_2^r$ satisfies $|A| \ge 2$. 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 $A \ne g + (S \cup \{0\})$; 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, $g + (g + s) \ne (g + s_1) + (g + s_2)$ whenever $s_1, s_2 \in S$. This is equivalent to S being sum-free. □

By Lemma 8, to prove Theorem 7, it suffices to show that if $A \subseteq \mathbb{F}_2^r$ is a large doubling-critical 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 9. *Let $r \geq 1$ be an integer and suppose that $A \subseteq \mathbb{F}_2^r$. If $|A| \geq 2^{r-2}+3$, then $\Gamma(A)$ is triangle-free. Moreover, if $|A| > 2^{r-2}+3$, then, indeed, $D(A)$ is sum-free.*

Remark. Observe that if $a_1, a_2, a_3 \in A$ induce a triangle in $\Gamma(A)$, then $D(A)$ is not sum-free in view of $(a_1 + a_2) + (a_2 + a_3) = a_1 + a_3$; thus, "$D(A)$ is sum-free" is a stronger conclusion than "$\Gamma(A)$ is triangle-free". We also notice that the bound $2^{r-2}+3$ is sharp. To see this, suppose that e_1, e_2, H, and A are as in Example 3, and set $A_0 := \langle e_1, e_2\rangle \cup H$. Then $|A| = 2^{r-2}+2$ and the vertices e_1, e_2, and e_1+e_2 of $\Gamma(A)$ induce a triangle, whereas $|A_0| = 2^{r-2}+3$ and $D(A_0)$ is not sum-free: for if h_1 and h_2 are distinct non-zero elements of H, then e_1+h_1, e_2+h_2 and $e_1+e_2+h_1+h_2$ belong to $D(A_0)$.

Proof of Proposition 9. 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 (namely b_1, b_2 and a), while every other coset of H contains at most two elements of A—both conclusions in view of d_1, $d_2 \in D(A)$. Thus, the assumption $|A| \geq 2^{r-2}+3$ implies that there is a coset containing exactly two elements of A. These two elements cannot differ by d_1 or d_2 (again, since d_1, $d_2 \in D(A)$); 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 $d_1+d_2 \notin D(A)$. The first assertion follows since if $\Gamma(A)$ were containing a triangle with two legs corresponding to d_1 and d_2, then the third leg would correspond to d_1+d_2.

Assuming now that the edges of $\Gamma(A)$ corresponding to 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 $d_2 = a_2+b_2$. (Note that a_1, a_2, b_1 and b_2 are all distinct.) Then 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, while every other coset of H contains at most two elements of A. If $|A| > 2^{r-2}+3$, then 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 distinct representations of d_1+d_2, leading, as above, to the conclusion $d_1+d_2 \notin D(A)$ and proving the second assertion. □

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

Proposition 10. *Let $r \geq 1$ 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*

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

We present two different proofs.

First proof of Proposition 10. 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 9. 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, with the fact that the last two are disjoint following from $a_1 + a_2 \in D(A)$ and Proposition 9, the fact that the first two are disjoint following from $a_1 + a_2 \in D(A)$, and the rest following from the definition of A'. 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 10. Since $a_1 + a_2 \in D(A)$ and the set $D(A)$ is sum-free by Proposition 9, 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 7. To this end, we first derive from Proposition 9 an interesting property of sum-free sets. Thinking projectively, if A is a large cap in $\mathrm{PG}(r-1, 2)$ and the point $p \notin A$ lies on the line determined by a pair of points in A, then in fact there are *many* pairs of points in A determining a line through p. (We remark that, for a *generic* subset of $\mathrm{PG}(r-1, 2)$, not assumed to be a cap, the same conclusion requires a much stronger assumption; cf. Lemma 12.)

In the definitions of a doubling-critical set and the set $D(A)$, given above in this section, we consider *unordered* representations of elements of $\mathbb{F}_2^r$, that is, representations which differ by the order of summands are considered identical. This convention is also extended onto the following corollary.

Corollary 11. *Let $r, \kappa \ge 2$ be integers and suppose that $S \subseteq \mathbb{F}_2^r$ is a sum-free set with $|S| > 2^{r-2} + \kappa$. Then every element of the sumset $2S$ has at least κ representations (distinct under permutation of summands) as a sum of two elements of S.*

Proof. Assuming that an element $c \in 2S$ has fewer than κ representations as a sum of two elements from S, we find a subset $S_0 \subseteq S$ with $|S_0| \geq |S| - (\kappa - 2)$ such that c has exactly one representation as a sum of two elements of S_0.

Let $A := S_0 \cup \{0\}$. Since $|S| > 2^{r-2} + \kappa$, we have $|A| > 2^{r-2} + 3$, so in view of $S_0 \subseteq D(A)$ and Proposition 9, 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, and therefore at least two representations as a sum of two elements from S_0 (since $(2S_0) \cap S_0 = \varnothing$), contradicting the choice of S_0. □

Deduction of Theorem 1 from Theorem 7. As we have already observed, if $S \subseteq \mathbb{F}_2^r$ is a maximal sum-free set and $s \in S \cup \{0\}$, then $\big(s + (S \cup \{0\})\big) \setminus \{0\}$ is a minimal 1-saturating set. Suppose now that $r \geq 1$ is an integer and $A \subseteq \mathbb{F}_2^r \setminus \{0\}$ is a minimal 1-saturating set with $|A| > \frac{11}{36}\, 2^r + 3$. By Lemma 6, either $A \cup \{0\}$ or A is doubling-critical. We show that, in the former case, A is of the form required, while the latter case cannot occur.

If $A \cup \{0\}$ is doubling-critical, then by Theorem 7 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 a *maximal* sum-free set (as remarked in Section 1), proving the assertion in this case.

Suppose now that A is doubling-critical, so that by Theorem 7 there exist a sum-free set $S \subseteq \mathbb{F}_2^r$ and an element $g \in \mathbb{F}_2^r$ with $A = g + (S \cup \{0\})$. In view of the previous paragraph, we may assume that $A \cup \{0\} = g + (S \cup \{0, g\})$ is *not* doubling-critical, whence $S \cup \{g\}$ is not sum-free (see the comment just above Theorem 7); 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 $g \neq 0$ and thus $s_1 \neq s_2$, 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

$$|S| = |A| - 1 > 2^{r-2} + 2,$$

it follows from Corollary 11 (applied with $\kappa = 2$) that $2S_1 = 2S$. Consequently,

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

On the other hand, as $g = s_1 + s_2$ with $s_1 \neq s_2$, we have $s_1, s_2 \in S_1 \cup (g + S_1)$, implying $S_1 \cup (g + S_1) = S \cup (g + S)$; therefore, $A_1 \cup 2A_1 = A \cup 2A$, contradicting the minimality of A. □

4. 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.

We start with an easy consequence of the pigeonhole principle; see, for instance, [N01, Lemma 2.1] or [GH01, Lemma 5.29].

Lemma 12. *Let B and C be non-empty subsets of a finite abelian group G. If $|B| + |C| \ge |G| + \kappa$ with an integer $\kappa \ge 1$, then every element of G has at least κ representations as a sum of an element from B and an element from C.*

We remark that, in Lemma 12 and in the vast majority of situations below, we consider representations of elements of $\mathbb{F}_2^r$ as sums of elements from two *potentially distinct* sets; therefore (in contrast with Section 3), representations are considered ordered.

Given a subgroup H of an abelian group G, by φ_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 13 (Kneser, [K53, K55]; see also [M76, N01, GH01])**.** *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 14. *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 chosen so 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 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} in Corollary 14 is sharp.

Proof of Corollary 14. We proceed by induction on r. The case $r = 1$ is immediate, and so we assume $r \ge 2$. 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 13, 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-holes" in B and C, showing that $B+H$ and $C+H$ are disjoint (since B and C are themselves disjoint). By the same reasoning, these two sets are also disjoint with $B+C$ (as $B+C$ is disjoint with both B and C, and $\pi(B+C)=H$, so there are no "H-holes" in $B+C$). 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

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

□

For an 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 15 (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 ρ in the statement of Theorem 15 and the part of the conclusion involving these quantities.)

Corollary 16. *If G is a finite abelian group and $B, C \subseteq G$ satisfy $\min\{|B|, |C|\} \ge 2$, then*

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

Proof. If $|B \overset{1}{+} C| + |B \overset{2}{+} C| \ge 2|B| + 2|C| - 4$, then $|B \overset{2}{+} C| \ge 2|B| + 2|C| - 4 - |G|$ follows trivially. Otherwise, we apply Theorem 15 to find $B' \subseteq B$ and $C' \subseteq C$ satisfying $|B \setminus B'| + |C \setminus C'| \le 1$ and $B' + C' = B' \overset{2}{+} C' = B \overset{2}{+} C$. Now

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

□

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 number of edges in a matching of the graph.

Lemma 17. *Let (V, E) be a triangle-free graph without isolated vertices, such that the matching number of (V, E) does not exceed 2. If $|V| \ge 6$, 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 two 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.

As a result, 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 $v_1 \ne v_2$ 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 18. *Let $\delta \ge 1$ be an integer and suppose that (V, E) is a graph such that $\deg(v_1) + \deg(v_2) \ge \delta$ holds for every edge $(v_1, v_2) \in E$. If (V, E) has no isolated vertices, then $|E| \ge (1 - \delta^{-1})|V|$.*

Remark. Equality is attained if (V, E) is a disjoint union of stars with δ vertices each.

Proof of Lemma 18. For $\delta \le 2$, the assertion is immediate. Assume therefore that $\delta \ge 3$ and, for each $j \in [1, \delta - 2]$, let $V_j := \{v \in V\colon \deg(v) = j\}$; also, let $V_+ := \{v \in V\colon \deg(v) \ge \delta - 1\}$, 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_+ in view of the hypothesis $\deg(v_1) + \deg(v_2) \geq \delta$. Taking the sum of inequality (2) with weight $2\delta^{-1}$, inequality (3) with weight $1-2\delta^{-1}$, and equalities (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 18 with $\delta = 2$ to the unique representation graph of a doubling-critical set, we get the following corollary.

Corollary 19. *If $r \geq 1$ is an integer and $A \subseteq \mathbb{F}_2^r$ is a doubling-critical set, then $|D(A)| \geq \frac{1}{2}|A|$.*

Lemma 20. *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 $|T| = t$ edges, then (V, E) has $|V| - 2t$ vertices not incident with the edges of T. By the maximality of T, no two of these vertices are adjacent, and thus each of them is incident to an edge from $E \setminus T$, since (V, E) contains no isolated vertices. Consequently,

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

□

Since the matching number of a graph does not exceed the number of edges in the graph, the following corollary strengthens Corollary 19.

Corollary 21. *If $r \geq 1$ is an integer and $A \subseteq \mathbb{F}_2^r$ is a doubling-critical set, then $|A| \leq |D(A)| + t$, where t is the matching number of $\Gamma(A)$.*

5. A "Light Version" of Theorem 1.

In this section, we combine the tools, developed so far, to prove the following, slightly weaker version of Theorem 1.

Theorem 1′. *Let $r \ge 1$ 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 7 at the end of Section 3, the reader will see that, in an identical way, Theorem **??′** can be obtained from the following "weak version" of Theorem 7 (the only difference being that for Theorem 1′ one needs the estimate $|A| \le 2^{r-1}$, mentioned in Section 1, to exclude the cases $r \le 3$ where $|A| > \frac{1}{3}\,2^r + 2$ does not imply $|A| > 2^{r-2} + 3$).

Theorem 7′. *Let $r \ge 1$ be an integer and suppose that $A \subseteq \mathbb{F}_2^r$ is doubling-critical. 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 **??′**.

Proof of Theorem **??′**. Suppose that $|A| > \frac{1}{3}\,2^r + 2$. As mentioned in Section 1, the size of a doubling-critical set in $\mathbb{F}_2^r$ does not exceed 2^{r-1}. Consequently, the hypotheses imply $r \ge 4$ and, furthermore, $|A| > 2^{r-2} + 3$; this is implicitly used below to invoke Propositions 9 and 10.

Set $\delta := |A| + |D(A)| - 2^{r-1}$. By Corollary 19, we have

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

thus, Proposition 10 and Lemma 18 give $\delta|D(A)| \ge (\delta - 1)|A|$. Substituting the value of δ and rearranging the terms, we rewrite this estimate as

$$f(|D(A)|) \le |A|(2^{r-1} + 1 - |A|), \tag{4}$$

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 19, and

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

(which follows by a straightforward computation using the assumption on the size of A), we derive from (4) that

$$|D(A)| \ge |A| - 1. \tag{5}$$

In view of Lemma 8, it suffices to show that $\Gamma(A)$ has a spanning star; that is (since $\Gamma(A)$ 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 10,

incident with each of the two edges of T are $\delta-2$ edges of $\Gamma(A)$. Moreover, since $\Gamma(A)$ is triangle-free by Proposition 9, there are at most two edges of $\Gamma(A)$ incident with both edges of T. Recalling that the total number of edges of $\Gamma(A)$ is $|D(A)|$, we get

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

Using (5) and the assumption $|A| > \frac{1}{3}\,2^r+2$, we derive

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

a contradiction. □

6. Securing Two Isolated Edges.

In this section, we prove Theorem 7 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 7 and 8. We split the argument into two lemmas.

Lemma 22. *Let $r \geq 1$ be an integer and suppose that $A \subseteq \mathbb{F}_2^r$ is doubling-critical. If $\Gamma(A)$ has at most one isolated edge and $|A| > \frac{3}{5}\,2^{r-1}+\frac{13}{5}$, then the matching number of $\Gamma(A)$ is at most 2.*

The proof is a minor modification of that of Theorem **??′**.

Proof of Lemma 22. Since $|A| \leq 2^{r-1}+1$ by Lemma 12 and the definition of a doubling-critical set, we have $r \geq 4$, and thus $\frac{3}{5}\,2^{r-1}+\frac{13}{5} > 2^{r-2}+3$; this will allow us to apply Propositions 9 and 10.

If $\Gamma(A)$ does not have isolated edges then, applying Lemma 18 to the graph $\Gamma(A)$, we get $|D(A)| \geq \frac{2}{3}\,|A|$; if $\Gamma(A)$ has one isolated edge, then, applying Lemma 18 to the graph $\Gamma(A)$ with this edge removed, we get $|D(A)| \geq 1+\frac{2}{3}\,(|A|-2) = \frac{2}{3}\,|A|-\frac{1}{3}$. In any case, letting $\delta := |A|+|D(A)|-2^{r-1}$ and assuming $|A| > \frac{3}{5}\,2^{r-1}+\frac{13}{5}$, we have

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

Consequently, using Proposition 10 and Lemma 18, we obtain $\delta|D(A)| \geq (\delta-1)|A|$. Substituting the value of δ, 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, that

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

and that

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

where all three estimates follow from $|A| > \frac{3}{5}\,2^{r-1}+\frac{13}{5}$. Thus, in view of $|D(A)| \ge \frac{2}{3}\,|A|-\frac{1}{3}$, we conclude that, indeed,

$$|D(A)| \ge |A|-2. \tag{6}$$

Suppose now by contradiction that $\Gamma(A)$ possesses a three-edge matching T. Using Proposition 10 to count the edges of $\Gamma(A)$ incident to those in T, and also taking into account the three edges of T, we get

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

for any edge incident to two different edges from T joins two vertices from T while, since $\Gamma(A)$ is triangle-free (by Proposition 9) and T is a matching in $\Gamma(A)$, the graph induced by the six vertices of T has at most six edges not in T. Rearranging the terms and applying (6) gives

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

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

Lemma 23. *Let $r \ge 1$ be an integer. If $A \subseteq \mathbb{F}_2^r$ is a doubling-critical 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 doubling-critical, $|A| > 2^{r-2}+3$, and the matching number of $\Gamma(A)$ is equal to 2. From $2^{r-1}+1 \ge |A| > 2^{r-2}+3$ we derive $2^{r-2} > 2$, and then $|A| > 6$. Hence, by Proposition 9 and Lemma 17, 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 (up to permutation of summands) as a sum of two elements of A.

By Proposition 9, $\Gamma(A)$ 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 we consider the sets $B_i := \{0, a_1 + a_2\} + D_i$. Since $a_1 + a_2 \in D(A)$ and $D_1, D_2 \subseteq D(A)$, and since $D_1 \cap D_2 = \varnothing$, it follows in view of Proposition 9 that $B_1 \cap B_2 = \varnothing$ and

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

We claim now that the sumset $B_1 + B_2 = \{0, a_1 + a_2\} + D_1 + D_2$ is disjoint from the union $B_1 \cup B_2 = \{0, a_1 + a_2\} + (D_1 \cup D_2)$; which, since $\{0, a_1 + a_2\}$ is a subgroup, is equivalent to $\{0, a_1+a_2\}+D_1+D_2$ being disjoint with $D_1 \cup D_2$. To see this, assume that $\{0, a_1+a_2\}+D_1+D_2$ is not disjoint with, say, D_1. As $(D_1+D_2)\cap D_1 = \varnothing$ by Proposition 9, this assumption yields $(a_1 + a_2 + D_1 + D_2) \cap D_1 \neq \varnothing$; that is, $a_1 + a_2 + d_1 + d_2 = d_1'$ for 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 14 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. □

7. Using Two Isolated Edges: The Coset Structure.

As follows from Lemmas 8, 22 and 23, and since $\frac{11}{36} > \frac{3}{10}$, to complete the proof of Theorem 7, 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 doubling-critical set such that $\Gamma(A)$ has two (or more) isolated edges, and we show that $|A| < \frac{11}{36}\, 2^r + 3$.

Shifting A, if necessary, we assume that $0 \in A$ and that 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,

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

Our argument involves 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, from $\{a_1, a_2 + a_3\} \subseteq (2A_0) \cap D(A)$, it follows that

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

for $g \notin L$, and the fact that $(0, a_1)$ and (a_2, a_3) are isolated edges gives

$$(A_g + A_0) \cap D_g = \varnothing, \tag{8}$$

under the same assumption. Furthermore, in view of Proposition 9 and since a_1, $a_2 + a_3 \in D(A)$, we have

$$(2D_g) \cap \{a_1, a_2 + a_3\} = \varnothing, \tag{9}$$

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

An immediate corollary of (7) and Lemma 12 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* L-cosets of type i (so that L is not counted in n_4); hence,

$$n_0 + n_1 + n_2 + n_3 + n_4 = 2^{r-3} - 1 \tag{10}$$

and

$$n_1 + 2n_2 + 3n_3 + 4n_4 = |A| - 4. \tag{11}$$

We now introduce a manner of pictorially representing the distribution of subsets of $\mathbb{F}_2^r$ in L-cosets that will help elucidate the otherwise tedious arguments needed for this section, and which may be helpful to keep in mind for the next section as well. Specifically, given a set $X \subseteq \mathbb{F}_2^r$ and an element $g \in \mathbb{F}_2^r$, we represent the elements of X in the coset $g + L$ by a diagram like

$$X \cap (g + L): \qquad \begin{array}{rcl} g & \longrightarrow & \left(\begin{array}{c} \left(\begin{array}{c} \bullet \\ \bullet \end{array} \right) \\ \left(\begin{array}{c} \circ \\ \circ \end{array} \right) \end{array} \right) \\ g + a_1 + a_2 + a_3 & \longrightarrow & \\ g + a_1 & \longrightarrow & \\ g + a_2 + a_3 & \longrightarrow & \\ g + a_2 & \longrightarrow & \left(\begin{array}{c} \left(\begin{array}{c} \circ \\ \bullet \end{array} \right) \\ \left(\begin{array}{c} \circ \\ \circ \end{array} \right) \end{array} \right) \\ g + a_1 + a_3 & \longrightarrow & \\ g + a_1 + a_2 & \longrightarrow & \\ g + a_3 & \longrightarrow & \end{array}$$

where each filled dot represents an element (as labeled) contained in X, and each open dot represents an element not in X. Note that this representation depends, though only up to translation, on the choice of the element g within the L-coset.

We remark that two blocks of points enclosed by parentheses of the same level are cosets of the same subgroup; say, the four two-point blocks correspond to the four H-cosets contained in $g + L$.

As an example, the distribution of the elements of A in L can be depicted as

$$A \cap L : \quad \begin{array}{cc} 0 & \longrightarrow \\ a_1 + a_2 + a_3 & \longrightarrow \\ a_1 & \longrightarrow \\ a_2 + a_3 & \longrightarrow \\ a_2 & \longrightarrow \\ a_1 + a_3 & \longrightarrow \\ a_1 + a_2 & \longrightarrow \\ a_3 & \longrightarrow \end{array} \begin{array}{c} \left.\begin{array}{c} \left.\begin{array}{c} \left.\begin{array}{c} \bullet \\ \circ \end{array}\right\} H \\ \begin{array}{c} \bullet \\ \circ \end{array} \end{array}\right\} M \\ \begin{array}{c} \circ \\ \bullet \end{array} \end{array}\right\} L \end{array}$$

where we have used braces to label the subgroups H, L and $M := \langle a_1, a_2 + a_3 \rangle$. Furthermore, K^- and K^+ are located in L as follows:

$$K^- : \begin{array}{c} \begin{array}{cc} \bullet & \bullet \\ \bullet & \bullet \end{array} \\ \begin{array}{cc} \circ & \circ \\ \circ & \circ \end{array} \\ \begin{array}{cc} \circ & \circ \\ \circ & \circ \end{array} \\ \begin{array}{cc} \bullet & \bullet \\ \bullet & \bullet \end{array} \end{array} \qquad K^+ : \begin{array}{c} \begin{array}{cc} \bullet & \bullet \\ \bullet & \bullet \end{array} \\ \begin{array}{cc} \circ & \circ \\ \circ & \circ \end{array} \\ \begin{array}{cc} \bullet & \bullet \\ \bullet & \bullet \end{array} \\ \begin{array}{cc} \circ & \circ \\ \circ & \circ \end{array} \end{array}$$

With the above diagrams in mind, we see that (7) is just the statement that any two elements of A from the same M-coset, excepting the two M-cosets contained in L, are actually from the same H-coset. Thus $|(g + M) \cap A| \leq 2$ for each $g \in \mathbb{F}_2^r$, and consequently, given any $g \in \mathbb{F}_2^r \setminus L$, we can find $x, y \in g + L$ (one element for each of the two M-cosets contained in $g + L$) such that $A_g \subseteq (x + H) \cup (y + H)$. Since $(x + H) \cup (y + H)$ is either a K^+ or K^--coset for any choice of x and y in the same L-coset, we conclude that A_g is contained either in a single K^+-coset, or in a single K^--coset. Using the homomorphism notation from Section 4, we record this observation as follows.

Claim 24. *For every $g \in \mathbb{F}_2^r \setminus L$, we have $\min\{|\varphi_{K^-}(A_g)|, |\varphi_{K^+}(A_g)|\} \leq 1$.*

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 24, 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 25. *For every $g \in \mathbb{F}_2^r$, we have*

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

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 an H-coset as $|A_g| = 2$ and A_g is contained in the intersection of a K^--coset and a K^+-coset. If $g \notin L$ and $g+L$ is of type 3 or 4, then A_g has two elements in the same M-coset, hence A_g contains an H-coset by the above observation that two elements of A_g, falling into the same M-coset, are actually in the same H-coset. As a result, if $g \notin L$ and $g+L$ is of type $2^0, 3$, or 4, then A_g contains an H-coset, and without loss of generality we assume $g+H \subseteq A_g$. However, $g+H+A_0 = g+L$ (as is readily apparent from the diagram for A_0), whence $A_0 + A_g = g+L$ and thus (8) implies $|D_g| = 0$.

By (9) 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 $\{a_2, a_3, a_1+a_2, a_1+a_3\}$. (If, for instance, we had $a_2 \in D(A)$, then a_2 would be adjacent to 0; if we had $a_1+a_2 \in D(A)$, then a_2 would be adjacent to a_1, etc.) Consequently, if $g \in L$, then $D_g = \{a_1, a_2+a_3\}$ and thus $|D_g| = 2$.

Next, if $g \in A$ and $g \notin L$, then by (8) the set D_g is disjoint with $A_0 + A_g \supseteq \{g, g+a_1, g+a_2, g+a_3\}$. Also, (9) 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 (9) and Lemma 12. □

Claim 26. *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 $|\widetilde{D}_g| = |\varphi_H(\widetilde{D}_g)| = 4$;*
(ii) *if $g + L$ is of type 2^-, then $|\widetilde{D}_g| = |\varphi_H(\widetilde{D}_g)| = 2$ and $|\varphi_{K^-}(\widetilde{D}_g)| = 1$;*
(iii) *if $g + L$ is of type 2^+, then $|\widetilde{D}_g| = |\varphi_H(\widetilde{D}_g)| = 2$ and $|\varphi_{K^+}(\widetilde{D}_g)| = 1$.*

Proof. If $g+L$ is not of type 0 and $g \notin L$, then by (8) the set D_g is disjoint with the set $A_g + A_0 \subseteq g + L$, which contains 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 $\widetilde{D}_g \subseteq g + L$ with $D_g \subseteq \widetilde{D}_g$ and $|\widetilde{D}_g| = |\varphi_H(\widetilde{D}_g)|$, and thus 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 (8), 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 K^--coset, but not contained in an H-coset.

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

8. Using Two Isolated Edges: Completion of the Proof.

In this section, we complete the proof of Theorem 7. We keep the notation and assumptions of the previous section, and since $\frac{11}{36}\,2^r + 3 > \frac{1}{3}\,2^r + 2$ for $r \in [1, 5]$, we may and do assume, in view of Theorem **??′**, that $r \ge 6$. 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. To this end, we further refine our knowledge of the distribution of the elements of A and $D(A)$ in L-cosets. In particular, we derive a number of estimates involving the quantities n_i, n_i^+, and n_i^- and relate them to the matching number of the graph $\Gamma(A)$.

Claim 27. *We have* $\min\{n_4^-, n_4^+\} < n_0 + 3$.

Proof. Suppose by contradiction 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 Lemma 12 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 $\varphi_L(A^-)$.

Hence, 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 the 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 26 we conclude that $|D_g| \le 1$ if $g+L$ is of type $1, 2^-$, or 2^+, and that $|D_g| \le 2$ if $g+L$ is of type 0. Combining this observation with Claim 25 and using (10) and (11), we derive

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

Compared with Corollary 19, 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 28. *We have* $\max\{n_4^-, n_4^+\} < n_0 + 3$.

Proof. Switching the notation, if necessary, and in view of Claim 27, we assume by contradiction that

$$n_4^- < n_0 + 3 \le n_4^+. \tag{12}$$

Let A^+ be the union of all sets A_g such that $g+L$ is of type 4^+. As in the proof of Claim 27, 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 $D(A)$. Consequently, in view of Claim 26 (ii), we have $|D_g| \le 1$ whenever $g+L$ is of type 2^-. Thus, by Claim 25,

$$|D(A)| \le 4n_0 + 2n_1 + n_2^- + 2n_2^+ + 2. \tag{13}$$

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 $|(b+K^+) \cap A| = 1$ (else b could not be adjacent to an element from A^+); it follows that B is disjoint with A^+ and, since there are precisely $n_1 + 2n_2^- + n_3^-$ elements $b \in A$ such that $|(b+K^+) \cap A| = 1$, that

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

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

$$t' \le n_1 + 2n_2^- + n_3^-. \tag{14}$$

Let t be the matching number of $\Gamma(A)$ and let T be a matching in $\Gamma(A)$ with $|T| = t$ edges. As $\Gamma(A)$ has no isolated vertices, the number of edges between A^+ and B in $\Gamma(A)$ is at least $|A^+| = 4n_4^+$, and the definition of t' ensures that at most t' of these edges belong to T; thus,

$$t \le |D(A)| - 4n_4^+ + t'. \tag{15}$$

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 K^+-coset (as shown earlier), and since $a + b \in D_{a+b}$, this coset is $a + b + K^+ = b + A_a$. Now, fix a matching T' of Γ'. As any edge of T corresponds to an element from $D(A)$, we see from Claim 25 that corresponding to the edges of T 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 (as all edges in D_{a+b} are adjacent to the same vertex), we conclude that

$$t \le 4n_0 + 2n_1 + 2n_2^- + 2n_2^+ + 2 - t'. \tag{16}$$

We complete the proof of Claim 28 showing that an appropriate combination of the estimates (12)–(16) yields a contradiction to the assumption on the size of A, made at the beginning of this section. Specifically, substituting (11) into the estimate of Corollary 21, we get

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

Taking the sum of this inequality with weight 4, identity (10) with weight 44, the first inequality in (12) in the form $-n_0 + n_4^- \le 2$ with weight 12, and inequalities (13), (14), (15), and (16) with 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.) In view of (11), 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 29. *We have* $n_4 \ge n_0 + n_1 + n_2^0 + 4$.

Proof. Applying Claim 25 and using (10) and (11), we get

$$\begin{aligned}|D(A)| &\le 4n_0 + 2n_1 + 2n_2^- + 2n_2^+ + 2\\ &= 6(n_0+n_1+n_2+n_3+n_4) - 2(n_1+2n_2+3n_3+4n_4)\\ &\qquad + 2(n_4 - n_0 - n_1 - n_2^0 + 1)\\ &= 3\cdot 2^{r-2} - 2|A| + 2(n_4 - n_0 - n_1 - n_2^0 + 2). \end{aligned} \tag{17}$$

Comparing with Corollary 19, we obtain

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

Hence

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

and if $r \ge 7$, then the result follows from

$$|A| \ge \left\lceil \frac{11}{36}\,2^r + 3 \right\rceil > \frac{3}{5}\,2^{r-1} + 4.$$

If $r = 6$ and $|A| \ge 24$, then the estimate $|A| > \frac{3}{5}\,2^{r-1} + 4$ remains valid, proving the result in this case, too.

In view of the assumption $r \ge 6$ made at the beginning of this section, and since for $r = 6$ we have $\left\lceil \frac{11}{36}\,2^r + 3 \right\rceil = 23$, we are left with the case where $r = 6$ and $|A| = 23$, which we proceed to consider. By Corollary 19, we have $|D(A)| \ge 12$, whence Proposition 10 gives

$$\deg(a_1) + \deg(a_2) \ge |A| + |D(A)| - 2^{r-1} \ge 3$$

for every edge (a_1, a_2) of $\Gamma(A)$. Now Lemma 18, applied with $\delta = 3$, yields $|D(A)| \ge \frac{2}{3}\,|A|$, and hence in fact $|D(A)| \ge 16$. Substituting into (17), we get

$$16 \le 3\cdot 16 - 2\cdot 23 + 2(n_4 - n_0 - n_1 - n_2^0 + 2),$$

leading to

$$n_4 - n_0 - n_1 - n_2^0 + 2 \ge 7$$

and implying the result. □

Claim 30. *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 $H = \{0, \sigma\}$. Observe that $\sigma \notin D(A)$, in view of Proposition 9 and since $a_1, a_2 + a_3 \in D(A)$. If (a, b) is an edge in $\Gamma(A)$, then either $a+\sigma \notin A$ or $b+\sigma \notin A$: otherwise $a+b \in D(A)$ would be represented as $(a+\sigma)+(b+\sigma)$ with both summands in A and distinct from a and b. This shows that every edge of $\Gamma(A)$ is incident with an element of the set $B := \{a \in A\colon a+\sigma \notin A\}$. However, by the definition of the quantities n_i, n_i^0, n_i^+, and n_i^-, 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 $\Gamma(A)$, no two distinct edges can be incident to the same element of B, and that any maximal matching contains the isolated edges $(0, a_1)$ and (a_2, a_3), both incident to two elements of B. □

By Claims 28 and 29, it remains to consider the case where

$$n_4^- \le n_0 + 2,\ n_4^+ \le n_0 + 2 \tag{18}$$

and

$$n_4 \ge n_0 + n_1 + 4, \tag{19}$$

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

$$\min\{n_4^-, n_4^+\} \ge 2. \tag{20}$$

As above, we define A^- to be the union of all sets A_g such that $g + L$ is of type 4^-, and A^+ to be 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,

$$|\varphi_L(A^-)| = n_4^-,\ |\varphi_L(A^+)| = n_4^+, \text{ and } |\varphi_L(B)| = n_2 + n_3 + n_4 + 1. \tag{21}$$

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 $\varphi_L(B)$.

If $A_g \subseteq A^-$, for some $g \in \mathbb{F}_2^r$, and if $b_1, b_2 \in B$ are distinct and belong to the same L-coset, then the K^--cosets $b_1 + A_g$ and $b_2 + A_g$ either coincide or cover the whole L-coset $g + \{b_1, b_2\} + L$. It follows that if $g \in C^- + L$, then $g + L$ contains a K^--coset disjoint from $D(A)$. Likewise, if $g \in C^+ + L$, then $g + L$ contains a K^+-coset disjoint from $D(A)$. Hence, for $g \in (C^- + L) \cap (C^+ + L)$, the set D_g is contained in an H-coset, and thus, by Claim 26,

$$|D_g| \le \begin{cases} 1 & \text{if } g + L \text{ is of type 1 or 2,} \\ 2 & \text{if } g + L \text{ is of type 0.} \end{cases} \tag{22}$$

By the pigeonhole principle, we have

$$|\varphi_L(C^-) \cap \varphi_L(C^+)| \ge |\varphi_L(C^-)| + |\varphi_L(C^+)| - 2^{r-3}, \tag{23}$$

while, by Corollary 16 and (21),

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

and

$$|\varphi_L(C^+)| \ge \min\{2n_2 + 2n_3 + 2n_4 + 2n_4^+ - 2^{r-3} - 2, n_2 + n_3 + n_4\}. \tag{24}$$

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

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

and

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

both hold true, then taking their sum we obtain

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

which, in view of (10), can be re-written as

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

this, however, is inconsistent with (19).

By symmetry, we can assume that

$$|\varphi_L(C^-)| \ge n_2 + n_3 + n_4, \tag{25}$$

and we consider two cases, according to the value in the right-hand side of (24).

If $|\varphi_L(C^+)| \ge n_2 + n_3 + n_4$, then from (23) and (25) we derive

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

Consequently, 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 25 and estimate (22), we see that

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

Combining this estimate with Corollary 21 and Claim 30, we get

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

and furthermore, substituting the value of $|A|$ from (11),

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

Taking the sum of this estimate, inequalities (18), and identity (10) multiplied by 6, we obtain

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

By (11), the expression in the left-hand side is at least $3(|A| - 4)$; consequently,

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

(in view of $r \ge 6$), a contradiction.

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

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

and hence

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

Taking the sum of the last inequality, the first of the inequalities (18), and identity (10) multiplied by 5, we obtain

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

In view of (20) and (11), this yields

$$3(|A| - 4) \leq 7 \cdot 2^{r-3} - 2,$$

leading to a contradiction as above. This completes the proof of Theorem 7.

Acknowledgement

We are grateful to Alexander Davydov for attracting our attention to the problem of studying 1-saturating sets and for several useful remarks, including mentioning to us the exceptional minimal 1-saturating set of size 11 in $\mathbb{F}_2^5$. We also thank the referees for the helpful suggestions.

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.

[GH01] A. Geroldinger and F. Halter-Koch, *Non-unique factorizations: Algebraic, combinatorial and analytic theory.* Pure and Applied Mathematics (Boca Raton), 278. Chapman & Hall/CRC, Boca Raton, FL, 2006.

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

[K53] M. Kneser, Abschätzung der asymptotischen Dichte von Summenmengen, *Math. Z.* **58** (1953), 459–484.

[K55] M. Kneser, Ein Satz über abelsche Gruppen mit Anwendungen auf die Geometrie der Zahlen, *Math. Z.* **61** (1955), 429–434.

[M76] H.B. Mann, Addition theorems: the addition theorems of group theory and number theory. Robert E. Krieger Publishing Co., Huntington, N.Y., 1976.

[N01] M.B. Nathanson, *Additive Number Theory: Inverse Problems and the Geometry of Sumsets*, Graduate Texts in Mathematics **165**, Springer-Verlag, New York, 1996.

E-mail address: diambri@hotmail.com

Institut für Mathematik und Wissenschaftliches Rechnen, Karl-Franzens-Universität Graz, Heinrichstrasse 36, 8010 Graz, Austria

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

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