Small doubling, atomic structure and ℓ-divisible set families

Lior Gishboliner* Benny Sudakov* István Tomon*

Abstract

Let $\mathcal{F} \subset 2^{[n]}$ be a set family such that the intersection of any two members of $\mathcal{F}$ has size divisible by ℓ. The famous Eventown theorem states that if $\ell = 2$ then $|\mathcal{F}| \leq 2^{\lfloor n/2 \rfloor}$, and this bound can be achieved by, e.g., an 'atomic' construction, i.e. splitting the ground set into disjoint pairs and taking their arbitrary unions. Similarly, splitting the ground set into disjoint sets of size ℓ gives a family with pairwise intersections divisible by ℓ and size $2^{\lfloor n/\ell \rfloor}$. Yet, as was shown by Frankl and Odlyzko, these families are far from maximal. For infinitely many ℓ, they constructed families $\mathcal{F}$ as above of size $2^{\Omega(n \log \ell/\ell)}$. On the other hand, if the intersection of *any number* of sets in $\mathcal{F} \subset 2^{[n]}$ has size divisible by ℓ, then it is easy to show that $|\mathcal{F}| \leq 2^{\lfloor n/\ell \rfloor}$. In 1983 Frankl and Odlyzko conjectured that $|\mathcal{F}| \leq 2^{(1+o(1))n/\ell}$ holds already if one only requires that for some $k = k(\ell)$ any k distinct members of $\mathcal{F}$ have an intersection of size divisible by ℓ. We completely resolve this old conjecture in a strong form, showing that $|\mathcal{F}| \leq 2^{\lfloor n/\ell \rfloor} + O(1)$ if k is chosen appropriately, and the $O(1)$ error term is not needed if (and only if) $\ell \,|\, n$, and n is sufficiently large. Moreover the only extremal configurations have 'atomic' structure as above. Our main tool, which might be of independent interest, is a structure theorem for set systems with small 'doubling'.

1 Introduction

An *eventown* is a family $\mathcal{F} \subset 2^{[n]}$ such that $|A \cap B|$ is even for any $A, B \in \mathcal{F}$. The famous Eventown theorem of Berlekamp [2], also proved independently by Graver [13], states that if $\mathcal{F} \subset 2^{[n]}$ is an eventown, then $|\mathcal{F}| \leq 2^{\lfloor n/2 \rfloor}$. This bound is also the best possible, and a simple construction showing this can be obtained as follows. Say that a family $\mathcal{F} \subset 2^{[n]}$ is *atomic*, if there exist disjoint sets $A_1, \dots, A_d \subset [n]$ such that $\mathcal{F}$ is the family of all sets F satisfying that either $A_i \subset F$ or $A_i \cap F = \emptyset$ for every $i \in [d]$, and F contains no element not covered by the sets A_i. The sets $A_1, \dots, A_d$ are called the *atoms* of $\mathcal{F}$. Also, let $S(n, \ell)$ be the atomic family for which $d = \lfloor n/\ell \rfloor$ and all $A_i, i \in [d]$ have size exactly ℓ. Note that $|S(n, \ell)| = 2^{\lfloor n/\ell \rfloor}$, and the size of the intersection of any number of sets in $S(n, \ell)$ is divisible by ℓ. Therefore, the family $S(n, 2)$ is an eventown of size $2^{\lfloor n/2 \rfloor}$. Moreover, any eventown family can be completed to a maximal one of size $2^{\lfloor n/2 \rfloor}$, see e.g. the book of Babai and Frankl [1], which is also a general reference on intersection problems.

In general, one might be tempted to conjecture that the maximal families $\mathcal{F} \subset 2^{[n]}$, whose all pairwise intersections are divisible by ℓ, have size close to $2^{(1+o(1))n/\ell}$. However, this turns out to be far from the truth. Frankl and Odlyzko [8] proved that if there exists a Hadamard matrix of order 4ℓ, then there exists such a family of size $2^{\Omega(n \log \ell/\ell)}$, and this bound is also the best possible up to the constant factor. On the other hand, it follows from a result of Deza, Erdős and Frankl [6], proved also by Frankl and Tokushige [9], that if we consider uniform families, that is, $\mathcal{F} \subset [n]^{(r)}$, then $|\mathcal{F}| \leq \binom{\lfloor n/\ell \rfloor}{r/\ell}$ if n is sufficiently large given r and $\ell \mid r$. This bound is also the best possible as witnessed by the family $\mathcal{F} = [n]^{(r)} \cap S(n, \ell)$. Let us emphasize that the condition that n must be large compared to r is necessary, otherwise this would contradict the aforementioned construction of Frankl and Odlyzko.

Despite all the above, if we require that the intersection of *any number* of sets in $\mathcal{F} \subset 2^{[n]}$ must have size divisible by ℓ, then it is not difficult to show that $|\mathcal{F}| \leq 2^{\lfloor n/\ell \rfloor}$ for any n and ℓ. Moreover, in this case, $\mathcal{F}$ is contained in some isomorphic copy of $S(n, \ell)$ (we say that two families in $2^{[n]}$ are isomorphic if they are equal up to a permutation of $[n]$). In 1983, Frankl and Odlyzko [8] asked whether a similar conclusion holds if we only require that the intersection of any k distinct sets in $\mathcal{F}$ has size divisible by ℓ, where k is some constant only depending on ℓ. More precisely, they conjectured that for some k, we must have $|\mathcal{F}| \leq 2^{(1+o(1))n/\ell}$ for such a family $\mathcal{F}$. Until recently, it was not even known if the bound $2^{O(n \log \ell/\ell)}$ can be improved for any constant k. Indeed, while there are many tools to handle pairwise intersections as they correspond to the scalar product of characteristic vectors, k-wise intersections are usually harder to analyse,

*ETH Zurich, *e-mail*: **{benjamin.sudakov,lior.gishboliner,istvan.tomon}@math.ethz.ch**

see, e.g., [12, 15, 18, 19, 21] for related results. Also, it was shown in [18] that if the conjecture is true, k must depend on ℓ. In particular, if ℓ is a power of 2, there exist families $\mathcal{F} \subset 2^{[n]}$ such that the intersection of any k sets in $\mathcal{F}$ has size divisible by ℓ, and $|\mathcal{F}| \geq 2^{c_k n \log \ell/\ell}$, where $c_k > 0$ is a constant only depending on k. In this paper we resolve the conjecture of Frankl and Odlyzko in the following strong form.

Theorem 1. *Let ℓ be a positive integer, then there exists $k = k(\ell)$ such that for every positive integer n the following holds. Let $\mathcal{F} \subset 2^{[n]}$ such that the intersection of any k distinct elements of $\mathcal{F}$ is divisible by ℓ. Then $|\mathcal{F}| \leq 2^{\lfloor n/\ell \rfloor} + c$, where $c = c(\ell, k)$ is a constant, and $c = 0$ if $\ell \mid n$ and n is sufficiently large.*

Note that the error term c is needed if $\ell \nmid n$. Indeed, in this case $S(n, \ell)$ is not extremal, one can add a constant number of sets contained in the nonempty set not covered by members of $S(n, \ell)$ while retaining the property that the intersection of every k distinct sets has size divisible by ℓ.

1.1 Stability

As we mentioned above, maximum size eventowns are not unique. In particular, any eventown can be completed to an eventown of size $2^{\lfloor n/2 \rfloor}$. However, as it was proved in [18], if we require that the intersection of any three sets is even sized, then $S(n, 2)$ is the unique family achieving the maximum. Moreover, in this case we have stability, that is, if $\mathcal{F} \subset 2^{[n]}$ has this property and $|\mathcal{F}| \geq (1-\epsilon)2^{\lfloor n/2 \rfloor}$ for some small ϵ, then $\mathcal{F}$ is a subfamily of some isomorphic copy of $S(n, 2)$.

In Theorem 1, we also have stability. More precisely, if $\mathcal{F} \subset 2^{[n]}$ such that the intersection of any k distinct elements of $\mathcal{F}$ is divisible by ℓ, and $|\mathcal{F}| > \frac{1}{2} \cdot 2^{\lfloor n/\ell \rfloor}$, then one can remove a constant number of sets from $\mathcal{F}$ to make it a subfamily of some isomorphic copy of $S(n, \ell)$. However, somewhat surprisingly, a much more robust form of stability also holds. We show that under the substantially weaker condition $|\mathcal{F}| > 2^{\alpha n}$, the family $\mathcal{F}$ already highly resembles a subfamily of $S(n, \ell)$ if k is chosen appropriately with respect to α and ℓ. The following result showing this seems to be new even for $\ell = 2$.

Say that a family $\mathcal{F} \subset 2^{[n]}$ is *k-closed* $\pmod{\ell}$ if the intersection of any k (not necessarily distinct) sets in $\mathcal{F}$ has size divisible by ℓ. Later, we show that if $\mathcal{F}$ has the property that any k *distinct* sets in $\mathcal{F}$ have an intersection of size divisible by ℓ, then $\mathcal{F}$ can be made k-closed by removing $O(n)$ elements. The next statement will be more convenient to state for k-closed families, however, this only makes a small difference by the previous claim. If $X \subset [n]$ then $\mathcal{F}|_X = \{F \cap X : F \in \mathcal{F}\}$ denotes the *projection* of $\mathcal{F}$ onto X.

Theorem 2. *Let $\epsilon > 0$, and let ℓ be a positive integer, then there exist $k = k(\ell, \epsilon)$ and $c = c(\ell, \epsilon)$ such that the following holds. Let $\mathcal{F} \subset 2^{[n]}$ be k-closed $\pmod{\ell}$. Then there exist $X \subset [n]$ such that $|\mathcal{F}|_X| \geq 2^{-\epsilon n}|\mathcal{F}|$, and $\mathcal{F}|_X$ is a subfamily of an isomorphic copy of $S(|X|, \ell)$.*

The following construction shows that this theorem is also optimal in a certain sense. Suppose that $\ell = p$ is a prime and $2p^{k+1}$ divides n. Partition $\{1, \dots, n/2\}$ into sets $A_1, \dots, A_q$ of size p, and partition $\{n/2+1, \dots, n\}$ into sets $B_1, \dots, B_r$ of size p^{k+1}. Let $\mathcal{F}$ be the family of sets F of the following form.

- For $i \in [q]$, either $A_i \subset F$ or $A_i \cap F = \emptyset$.
- For $j \in [r]$, identify B_j with the vector space $\mathbb{F}_p^{k+1}$. Then $F \cap B_j$ is a k-dimensional subspace of $\mathbb{F}_p^{k+1}$.

Clearly, we have $|\mathcal{F}| = 2^q p^{(k+1)r} = 2^{n/2p} p^{n(k+1)/2p^{k+1}}$. Also, if $F_1, \dots, F_k \in \mathcal{F}$, then $|F_1 \cap \dots \cap F_k|$ is divisible by p. Indeed, $|F_1 \cap \dots \cap F_k \cap A_i| \in \{0, p\}$, and $F_1 \cap \dots \cap F_k \cap B_j$ is a subspace of $\mathbb{F}_p^{k+1}$ of dimension at least 1, so its size is also divisible by p. Finally, if $X \subset [n]$ is such that $\mathcal{F}|_X$ is a subfamily of some isomorphic copy of $S(|X|, p)$, then $X \subset \{1, \dots, n/2\}$ and $|\mathcal{F}|_X| \leq 2^{n/2p} = 2^{-n(\log_2 p)(k+1)/2p^{k+1}}|\mathcal{F}|$.

1.2 Tools

The main technical tool we develop to prove Theorem 1 is a structure theorem for set systems with small 'doubling'. This result is similar in spirit to the famous Freiman-Ruzsa type theorems in additive combinatorics. These theorems give a structure of a subset A of a group such that the sum-set $A + A$ has size not much larger than A (i.e. A has small doubling). The theorem of Freiman and Ruzsa [11, 17] states that a subset of the integers with small doubling must be contained in a so-called generalized arithmetic progression of bounded rank (see also [20]). This classical result was subsequently generalized to abelian groups by Green and Ruzsa [14] and to all groups by Breuillard, Green and Tao [3] (see also the survey [4]). These results show that sets with small doubling must be close in structure to one of few natural examples.

Given a set-family $\mathcal{F}$, our measure for the size of $\mathcal{F}$ will be the dimension of the subspace $\langle \mathcal{F} \rangle$ spanned by the characteristic vectors of the sets in $\mathcal{F}$ over some field $\mathbb{F}$. Let $\mathcal{F} \cdot \mathcal{F} = \{A \cap B : A, B \in \mathcal{F}\}$, note that, by definition, $\mathcal{F} \subset \mathcal{F} \cdot \mathcal{F}$. What can we say about $\mathcal{F}$ if the dimension of $\langle \mathcal{F} \cdot \mathcal{F} \rangle$ is not much larger than that of $\langle \mathcal{F} \rangle$? Observe that if S is an atomic set-family, then $S \cdot S = S$. Our structure theorem shows that this is essentially the only possible example: any set-family $\mathcal{F}$ with small 'doubling' must be close to being atomic. To make this more precise, we need the following definition. Given $i, j \in [n]$, say that i and j are *twins* for $\mathcal{F}$ if every $F \in \mathcal{F}$ either contains both i, j or none of them, and there is at least one $F \in \mathcal{F}$ such that $i, j \in F$. Note that being twins is an equivalence relation (on the set of $i \in [n]$ which are contained in at least one $F \in \mathcal{F}$). A set of coordinates $T \subset [n]$ is called a *set of twins* if any pair of elements in T are twins, or $|T| = 1$. Also, say that T is a maximal set of twins if it is a complete equivalence class of the twins relation. We can now state our structure theorem.

Theorem 3. *Let $\mathcal{F} \subset \{0,1\}^n$, let $\mathbb{F}$ be a field, and suppose that $\dim\langle \mathcal{F} \rangle = d$ and $\dim\langle \mathcal{F} \cdot \mathcal{F} \rangle = d + h$. Then $[n]$ can be partitioned into $d+1$ sets $A_1, \ldots, A_d, B$ such that A_i is a maximal set of twins for $\mathcal{F}$ for $i \in [d]$, and $\dim\langle \mathcal{F}|_B \rangle \leq 2h$.*

2 Small 'doubling' and twins

In this section we establish Theorem 3. We will actually prove a more general statement about arbitrary vector spaces. Let us introduce some notation.

As usual, $\mathbb{Z}_\ell$ denotes the ring of integers modulo ℓ, and if p is a prime, we write $\mathbb{F}_p$ instead of $\mathbb{Z}_p$ to emphasize that it is also a field. Fix any commutative ring $\mathcal{R}$ with a unity (in our case, $\mathcal{R}$ will be either a field or $\mathbb{Z}_\ell$ for some positive integer ℓ). For a vector $v \in \mathcal{R}^n$, we use $v(i)$ to denote the ith coordinate of v. The *support* of v is $\{i \in [n] : v(i) \neq 0\}$. For $\mathcal{F} \subset \mathcal{R}^n$, we use $\langle \mathcal{F} \rangle$ to denote the *span* (i.e. the set of all linear combinations) of the elements of $\mathcal{F}$. If $\mathcal{R} = \mathbb{Z}_\ell$, we might write $\langle \mathcal{F} \rangle_\ell$ instead of $\langle \mathcal{F} \rangle$ if $\mathcal{R}$ is not clear from the context. If $A \subset [n]$ and $F \in \mathcal{R}^n$, then $F|_A \in \mathcal{R}^A$ denotes the restriction of F to the coordinates in A, and if $\mathcal{F} \subset \mathcal{R}^n$, then $\mathcal{F}|_A = \{F|_A : F \in \mathcal{F}\}$.

Given vectors $v, w \in \mathcal{R}^n$, let $v \cdot w$ be the vector in $\mathcal{R}^n$ defined as $(v \cdot w)(i) = v(i)w(i)$ for $i \in [n]$. Note that if v and w are characteristic vectors of sets A and B, then $v \cdot w$ is the characteristic vector of $A \cap B$. For $V, W \subset \mathcal{R}^n$, let $V \cdot W = \{v \cdot w : v \in V, w \in W\}$. Given $V \subset \mathcal{R}^n$ and $i, j \in [n]$, say that i and j are *twins* for V if $v(i) = v(j)$ for all $v \in V$ and $v(i) \neq 0$ for at least one $v \in V$. Observe that if V is a subspace generated by some family $\mathcal{F} \subset \{0,1\}^n$, then this definition of twins exactly coincides with the one given in the previous section.

Theorem 4. *Let $\mathbb{F}$ be a field, $V < \mathbb{F}^n$, $d = \dim(V)$ and $\dim(\langle V \cup (V \cdot V) \rangle) = d + h$. Then $[n]$ can be partitioned into $d+1$ sets $A_1, \ldots, A_d, B$ such that A_i is a maximal set of twins for V for each $i \in [d]$, and $\dim(V|_B) \leq 2h$.*

Proof. For $i, j \in [n]$, say that i and j are *siblings* for V if there exists $\lambda \in \mathbb{F}, \lambda \neq 0$ such that $v(i) = \lambda v(j)$ for all $v \in V$, and $v(i) \neq 0$ for at least one $v \in V$. Note that if $\lambda = 1$ then i, j are twins.

Let $v_1, \ldots, v_d \in V$ be a basis of V, and let M be the $d \times n$ sized matrix, whose rows are $v_1, \ldots, v_d$. It is possible to choose the basis $v_1, \ldots, v_d$ such that after possibly rearranging the columns of M, the restriction of M to the first d columns is a diagonal matrix.

We can assume without loss of generality that M has no all-zero columns, i.e. that there is no coordinate $i \in [n]$ such that $v(i) = 0$ for all $v \in V$. Note that the indices $i \in [d]$ and $j \in [n] \setminus [d]$ are siblings for V if and only if $v_i(j) \neq 0$ and the j-th column of M contains exactly one nonzero entry. For $i \in [d]$, let S_i contain all the siblings of i, and let $B' = [n] \setminus (S_1 \cup \cdots \cup S_d)$. Note that for every $j \in B'$, the j-th column of M contains at least two nonzero entries.

Let $r = \dim(V|_{B'})$. Let $C \subset B'$ such that $|C| = r$ and $\dim(V|_C) = r$. Also, let $w_i = v_i|_C$ for $i \in [d]$. For $c \in C$, let $\mathbb{1}_c \in \mathbb{F}^C$ be the characteristic vector of the single element set $\{c\}$.

Claim 5. *For every $c \in C$ there exists $\lambda \in \mathbb{F}$, $c' \in C, c' \neq c$, and two vectors of coefficients $x, y \in \mathbb{F}^d$ such that x, y have disjoint support and $\mathbb{1}_c + \lambda \mathbb{1}_{c'} = (\sum_{i=1}^d x(i) w_i) \cdot (\sum_{i=1}^d y(i) w_i)$.*

Here, as always, $\cdot$ denotes the coordinate-wise product of vectors.

Proof. Let $K \subset [d]$ be the set of indices i such that $w_i(c) \neq 0$. Note that $|K| \geq 2$. Also, without loss of generality, suppose that $w_1, \ldots, w_r$ is a basis of $\mathbb{F}^C$. For every $u \in C$, we can write $\mathbb{1}_u = \sum_{i=1}^r \lambda_{u,i} w_i$ with suitable $\lambda_{u,1}, \ldots, \lambda_{u,r} \in \mathbb{F}$. Consider two cases.

Case 1. There exists $k \in K$ such that $k \notin [r]$ or $\lambda_{c,k} = 0$. Recall that, by definition of K, $w_k(c) \neq 0$. In this case, the choices $\lambda = 0$, $x = (\lambda_{c,1}, \ldots, \lambda_{c,r}, 0, \ldots, 0)$ and y defined as

$$y(i) = \begin{cases} \frac{1}{w_k(c)} & \text{if } i = k, \\ 0 & \text{otherwise} \end{cases}$$

suffice. Indeed, $\sum_{i=1}^d x(i) w_i = \mathbb{1}_c$, $\sum_{i=1}^d y(i) w_i = \frac{1}{w_k(c)} w_k$, $(\sum_{i=1}^d x(i) w_i) \cdot (\sum_{i=1}^d y(i) w_i) = \mathbb{1}_c$, and x and y have disjoint supports by our assumption on k.

Case 2. $K \subset [r]$ and $\lambda_{c,k} \neq 0$ for every $k \in K$. There exists some $c' \in C, c' \neq c$ such that not all $|K|$ coefficients $\lambda_{c',k}$ for $k \in K$ vanish. This is true as $(\mathbb{1}_u)_{u \in C}$ is also a basis of $\mathbb{F}^C$ and $|K| \geq 2$. Therefore, if we had $\lambda_{c',k} = 0$ for all $k \in K$ and $c' \in C \setminus \{c\}$, then the $r - |K| \leq r - 2$ vectors $w_j, j \in [r] \setminus K$ would span the $r - 1$ independent vectors $\mathbb{1}_{c'}, c' \in C \setminus \{c\}$. Choose such a c' and let $k \in K$ such that $\lambda_{c',k} \neq 0$. Let $\alpha = -\lambda_{c,k}/\lambda_{c',k}$, $x = (\lambda_{c,1} + \alpha \lambda_{c',1}, \ldots, \lambda_{c,r} + \alpha \lambda_{c',r}, 0, \ldots, 0)$, and define y as

$$y(i) = \begin{cases} \frac{1}{w_k(c)} & \text{if } i = k, \\ 0 & \text{otherwise} \end{cases}.$$

Then $\sum_{i=1}^d x(i) w_i = \mathbb{1}_c + \alpha \mathbb{1}_{c'}$, $\sum_{i=1}^d y(i) w_i = \frac{1}{w_k(c)} w_k$, $(\sum_{i=1}^d x(i) w_i) \cdot (\sum_{i=1}^d y(i) w_i) = \mathbb{1}_c + \frac{\alpha w_k(c')}{w_k(c)} \mathbb{1}_{c'}$, and x and y have disjoint supports because $x(k) = 0$ by our choice of α.

□

For $c \in C$, let $x_c, y_c \in \mathbb{F}^d$ be the vectors of coordinates x and y given by the previous claim. Also, let $z_c = (\sum_{i=1}^d x_c(i) v_i) \cdot (\sum_{i=1}^d y_c(i) v_i)$, then $z_c \in V \cdot V$. Note that $z_c|_{[n] \setminus B'} = \mathbf{0}$, because x_c and y_c have disjoint supports, and every column in $M|_{[n] \setminus B'}$ has only one non-zero entry. Also, $z_c|_C = \mathbb{1}_c + \lambda \mathbb{1}_{c'}$ for some $c' \in C, c \neq c'$ and $\lambda \in \mathbb{F}$. Let W be the vector-space generated by the vectors $(z_c)_{c \in C}$.

Claim 6. $\dim(W) \geq r/2$.

Proof. Suppose this does not hold, and without loss of generality, let $D \subset C$ such that $\{z_c\}_{c \in D}$ is a basis of W, and $|D| < r/2$. As $z_c|_C$ vanishes in all but at most two coordinates, there exists $u \in C$ such that z_c vanishes on u for every $c \in D$. But then z_u is not contained in W, contradiction. □

Let I be the set of indices $i \in [d]$ such that the entries in $v_i|_{S_i}$ are *not* all equal, that is, S_i is not a set of twins. For $i \in [d]$, let A_i be a maximal set of twins in S_i, then $S_i \neq A_i$ if and only if $i \in I$. If $i \in I$, and v_i is the constant s vector on A_i, then define $v'_i = s v_i - v_i \cdot v_i \in \langle V \cup (V \cdot V) \rangle$. Then v'_i vanishes on A_i and on S_j for every $j \in [d] \setminus \{i\}$, and v'_i does not vanish on $S_i \setminus A_i$. Therefore, the $d + |I|$ vectors $v_1, \ldots, v_d, \{v'_i\}_{i \in I}$ are linearly independent and do not vanish on $[n] \setminus B'$. Set $V' = \langle \{v_i : i \in [d]\} \cup \{v'_i : i \in I\} \rangle$, then $\dim(V') = d + |I|$. Also, we have $W \cap V' = \{\mathbf{0}\}$, as $z_c|_{[n] \setminus B'} = \mathbf{0}$ for every $c \in C$, but every vector in V' other than $\mathbf{0}$ does not completely vanish on $[n] \setminus B'$. But then as $V' + W < \langle V \cup (V \cdot V) \rangle$, we have

$$\dim(\langle V \cup (V \cdot V) \rangle) \geq \dim(V' + W) = \dim(V') + \dim(W) \geq d + |I| + r/2.$$

Therefore, $|I| + r/2 \leq h$. Let $B = B' \cup \bigcup_{i \in I} (S_i \setminus A_i)$, then $\dim(V|_B) \leq r + |I| \leq 2h$, so the sets $A_1, \ldots, A_d, B$ satisfy the desired properties. □

3 k-closed families are atomic

In this section, we prove a theorem which implies Theorem 1 after a small amount of work. Before stating it, we need some additional notation. Say that $\mathcal{F}$ is *non-reducible* if $\mathcal{F}$ does not vanish on any of the coordinates (namely, if there is no i such that $v(i) = 0$ for all $v \in \mathcal{F}$). Recall that $\mathcal{F} \cdot \mathcal{F}$ is the set of all $v \cdot w$, $v, w \in \mathcal{F}$, where $v \cdot w$ is the coordinate-wise product. We also put $v^k = v \cdot \ldots \cdot v$ and $\mathcal{F}^k = \mathcal{F} \cdot \ldots \cdot \mathcal{F}$, where the products contain k terms. Finally, let $||v|| = \sum_{i=1}^n v(i)$.

We say that a set $\mathcal{F} \subset \mathbb{Z}_\ell^n$ is *k-closed* if $||v|| = 0$ for every $1 \leq i \leq k$ and $v \in \mathcal{F}^i$. Note that if $\mathcal{F} \subset \{0,1\}^n$, then this is the same as saying that the intersection of any k not necessarily distinct sets from $\mathcal{F}$ is divisible by ℓ. We will use the following simple, but important observation repeatedly.

Claim 7. *If $\mathcal{F} \subset \mathbb{Z}_\ell^n$ is k-closed, then $\langle \mathcal{F} \rangle$ is also k-closed. Also, if $\mathcal{F}, \mathcal{F}' \subset \mathbb{Z}_\ell^n$, then $\langle \mathcal{F} \cdot \mathcal{F}' \rangle = \langle \mathcal{F} \rangle \cdot \langle \mathcal{F}' \rangle$.*

For what follows, let us record some simple but important properties of twins.

Claim 8. *Let $\mathcal{F} \subset \{0,1\}^n$.*

1. *If $\mathcal{F}$ is non-reducible, then the maximal sets of twins for $\mathcal{F}$ form a partition of $[n]$.*
2. *For every $k \geq 1$, the family $\bigcup_{i=1}^k \mathcal{F}^i$ has the same sets of twins as $\mathcal{F}$.*

Recall that $S(n, \ell) \subset 2^{[n]}$ is the atomic set-family with $\lfloor n/\ell \rfloor$ atoms of size ℓ each. The main result of this section is the following variant of Theorem 1. We show that if $\mathcal{F} \subset 2^{[n]}$ is such that the intersection of any k not necessarily distinct elements of $\mathcal{F}$ has size divisible by ℓ, then $|\mathcal{F}| \leq 2^{\lfloor n/\ell \rfloor}$, given k is sufficiently large with respect to ℓ. We also show that if $\mathcal{F}$ is close to being extremal, then $\mathcal{F}$ must be a subfamily of (an isomorphic copy of) $S(n, \ell)$.

Theorem 9. *Let ℓ be a positive integer, then there exists k such that the following holds. Let $\mathcal{F} \subset \{0,1\}^n$ such that $\mathcal{F}$ is k-closed over $\mathbb{Z}_\ell$. Then $|\mathcal{F}| \leq 2^{\lfloor n/\ell \rfloor}$. Also, if $|\mathcal{F}| > 2^{\lfloor n/\ell \rfloor - 1}$, then $[n]$ can be partitioned into sets $A_1, \ldots, A_d, A'$ such that A_i is a maximal set of twins for $\mathcal{F}$ for $i \in [d]$, $|A_i| = \ell$, $|A'| \leq \ell - 1$, and $\mathcal{F}$ vanishes on A'.*

The Eventown theorem mentioned in the introduction can be easily extended to vector spaces, where we replace intersections with scalar products. We will make use of the following simple extension of this in which we consider certain bilinear forms instead of scalar product.

Lemma 10. *Let $\mathbb{F}$ be a field, let $b_1, \ldots, b_n \in \mathbb{F}$, let z be the number of zeros among $b_1, \ldots, b_n$, and let $b : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ be the bilinear form defined as $b(v, w) = \sum_{i=1}^n b_i v(i) w(i)$. Let $V < \mathbb{F}^n$ such that $b(v, w) = 0$ for every $v, w \in V$. Then $\dim(V) \leq \frac{1}{2}(n + z)$.*

Proof. Let M be the $n \times n$ diagonal matrix with diagonal entries $b_1, \ldots, b_n$, and let $W = \{Mv : v \in V\} < \mathbb{F}^n$. Then $\dim(\ker M) = z$, so $\dim(W) \geq \dim(V) - z$. By definition, V and W are orthogonal spaces (with respect to the standard inner product). Therefore, $\dim(V) + \dim(W) \leq n$, which implies $\dim(V) \leq \frac{1}{2}(n + z)$. $\square$

In what comes, we show that if $\ell = p^\alpha$ is a prime power, and $\mathcal{F} \subset \{0,1\}^n$ is k-closed over $\mathbb{Z}_\ell$ for some large constant k, then most sets of maximal twins for $\mathcal{F}$ must have size divisible by ℓ, provided that the dimension of $\langle \mathcal{F} \rangle_p$ is large. We start with the case when ℓ is a prime.

Lemma 11. *Let $V < \mathbb{F}_p^n$, let $A_1, \ldots, A_d$ be a partition of $[n]$ into twins for V, and suppose that V is 2-closed. If $\dim(V) = d - h$, then at least $d - 2h$ of the numbers $|A_1|, \ldots, |A_d|$ are divisible by p.*

Proof. For $i \in [d]$, let $b_i = |A_i|$ and let b be the bilinear form defined as in Lemma 10. Let $\phi : V \to \mathbb{F}_p^d$ be the linear map defined as $\phi(v)(i) = s$ if $v|_{A_i}$ is the constant s vector. Then ϕ is an injection, so $\dim(\phi(V)) = \dim(V) = d - h$. Also, for every $u, v \in V$, we have $||u \cdot v|| = b(\phi(u), \phi(v))$, so we have $b(x, y) = 0$ for every $x, y \in \phi(V)$. But then by Lemma 10, if z is the number of zeros among $b_1, \ldots, b_d$, then $\dim(\phi(V)) \leq \frac{1}{2}(d + z)$, which gives $z \geq d - 2h$. $\square$

Lemma 12. *Let p be a prime and $\alpha \in \mathbb{Z}^+$. Let $\mathcal{F} \subset \{0,1\}^n$ be $2(p + \alpha)$-closed over $\mathbb{Z}_{p^\alpha}$, let $\dim(\langle \mathcal{F} \rangle_p) = d$, and let $A_1, \ldots, A_d, B$ be a partition of $[n]$ such that A_i is a set of twins, and $\dim(\langle \mathcal{F}|_B \rangle_p) \leq h$. Then at least $d - 2\alpha h$ of the numbers $|A_1|, \ldots, |A_d|$ are divisible by p^α.*

Proof. Let $V = \langle \mathcal{F} \rangle_p$. We prove this by induction on α. The case $\alpha = 1$ follows from Lemma 11. Suppose that $\alpha > 1$. Then, by our induction hypothesis, at least $k = d - (2\alpha - 2)h$ of the sets $A_1, \dots, A_d$ have size divisible by $p^{\alpha-1}$, without loss of generality, let these sets be $A_1, \dots, A_k$. Also, let $B' = A_{k+1} \cup \dots \cup A_d \cup B$. Note that

$$\dim(V|_{B'}) \leq \dim(V|_B) + (d - k) \leq h + d - k.$$

Therefore, V contains a subspace W such that $\dim(W) \geq d - \dim(V|_{B'}) \geq k - h$ and W vanishes on B'.

Note that for every $w \in W$ there exists some $w' \in \langle \mathcal{F} \rangle_{p^\alpha}$ such that $w' \equiv w \pmod{p}$. Let β be the smallest number such that $\beta > \alpha$ and $\beta = 1 \pmod{p-1}$, then $\beta < \alpha + p$. As $\mathcal{F}$ is 2β-closed, for every $u, v \in W$ we have that $||(u')^\beta \cdot (v')^\beta||$ is divisible by p^α. However, note that $(w')^\beta \equiv w \pmod{p}$, and if $w(i) = 0$ (over $\mathbb{F}_p$) for some $i \in [n]$, then $p^\alpha \mid (w')^\beta(i)$.

For $i \in [k]$, let A_i' be a set of size $|A_i|/p^{\alpha-1}$, and let $A' = \bigcup_{i=1}^k A_i'$. Define the linear map $\phi : W \to \mathbb{F}_p^{A'}$ as follows. If $w \in W$, $i \in [k]$, and $w|_{A_i}$ is the constant s vector, then $\phi(w)|_{A_i'}$ is the constant s vector. Then ϕ is an injection, so $\dim(W) = \dim(\phi(W))$. Also, for every $u, v \in W$, we have

$$||(u')^\beta \cdot (v')^\beta|| \equiv p^{\alpha-1}||\phi(u) \cdot \phi(v)|| \pmod{p^\alpha}.$$

Therefore, we must have $||x \cdot y|| = 0$ for every $x, y \in \phi(W)$. Let z be the number of sets among $A_1', \dots, A_k'$, whose size is divisible by p. We can apply Lemma 11 to conclude that

$$z \geq k - 2(k - \dim(\phi(W))) \geq k - 2h \geq d - 2\alpha h.$$

As z is also the number of sets among $A_1, \dots, A_k$ whose size is divisible by p^α, this finishes the proof. $\square$

Lemma 13. *Let p be a prime, and $\alpha, t \in \mathbb{Z}^+$. Let $\mathcal{F} \subset \{0,1\}^n$ such that $\mathcal{F}$ is non-reducible and $2^{t+1}(p+\alpha)$-closed over $\mathbb{Z}_{p^\alpha}$. Let $A_1, \dots, A_d$ be the unique partition of $[n]$ into maximal sets of twins, and let*

$$B = \bigcup_{\substack{i \in [d] \\ |A_i| \not\equiv 0 \pmod{p^\alpha}}} A_i.$$

Then $\dim(\langle \mathcal{F}|_B \rangle_p) \leq \frac{6n\alpha}{t}$.

Proof. Let $\mathcal{F}_0 = \mathcal{F}$, and for $i = 1, 2, \dots, t$, let $\mathcal{F}_i = \mathcal{F}_{i-1} \cdot \mathcal{F}_{i-1}$. Note that $\mathcal{F}_{i-1} \subset \mathcal{F}_i$ and $\mathcal{F}_i$ is $2^{t+1-i}(p+\alpha)$-closed over $\mathbb{Z}_{p^\alpha}$, and $A_1, \dots, A_d$ is also the unique partition of $[n]$ into maximal sets of twins for $\mathcal{F}_i$. As $\dim(\langle \mathcal{F}_r \rangle_p)$ is monotone increasing, there exists $0 \leq r < t$ such that

$$\dim(\langle \mathcal{F}_{r+1} \rangle_p) \leq \dim(\langle \mathcal{F}_r \rangle_p) + \frac{n}{t}.$$

Let $d' = \dim(\langle \mathcal{F}_r \rangle_p)$. Applying Theorem 3, we get that $[n]$ can be partitioned into $d'+1$ sets $A_1, \dots, A_{d'}, C$ such that A_i is a maximal set of twins for $\mathcal{F}_r$, and $\dim(\langle \mathcal{F}_r|_C \rangle_p) \leq \frac{2n}{t}$. But then as $\mathcal{F}_r$ is $2(p+\alpha)$-closed, we can apply Lemma 12 to conclude that at least $q = d' - \frac{4n\alpha}{t}$ of the numbers $|A_1|, \dots, |A_{d'}|$ are divisible by p^α. Without loss of generality, let $A_1, \dots, A_q$ be the sets of twins whose sizes are divisible by p^α. Let $D = C \cup A_{q+1} \cup \dots \cup A_{d'}$. Then $B \subset D$, and noting that $\dim(\langle \mathcal{F}_r|_{D \setminus C} \rangle_p) \leq d' - q$, and $\mathcal{F} \subset \mathcal{F}_r$, we get the chain of inequalities

$$\dim(\langle \mathcal{F}|_B \rangle_p) \leq \dim(\langle \mathcal{F}_r|_D \rangle_p) \leq (d' - q) + \dim(\langle \mathcal{F}_r|_C \rangle_p) \leq \frac{4n\alpha + 2n}{t} \leq \frac{6n\alpha}{t}.$$

This finishes the proof. $\square$

The final ingredient we need for the proof of Theorem 9 is the following well known result, see e.g. the work of Odlyzko [16].

Lemma 14. *Let p be a prime and $V < \mathbb{F}_p^n$. Then $|V \cap \{0,1\}^n| \leq 2^{\dim(V)}$.*

Proof of Theorem 9. Write $\ell = p_1^{\alpha_1} \dots p_s^{\alpha_s}$, where $p_1, \dots, p_s$ are distinct primes. We show that $k = 2^{t+1} \max_{r \in [s]}(p_r + \alpha_r)$ suffices, where $t = 12\ell \sum_{r=1}^s \alpha_r$. More precisely, we show the following two statements. Let $\mathcal{F} \subset \{0,1\}^n$ such that $\mathcal{F}$ is k-closed over $\mathbb{Z}_\ell$.

(1) Then $|\mathcal{F}| \leq 2^{\lfloor n/\ell \rfloor}$.

(2) If $|\mathcal{F}| > 2^{\lfloor n/\ell \rfloor - 1}$, then $[n]$ can be partitioned into sets $A_1, \dots, A_{\lfloor n/\ell \rfloor}, A'$ such that A_i is a maximal set of twins for $i \in [\lfloor n/\ell \rfloor]$, $|A_i| = \ell$, $|A'| \leq \ell - 1$, and $\mathcal{F}$ vanishes on A'.

We proceed by induction on n. If $n \leq 6s\ell$, the statements are easy to show. Indeed, let $A_1, \dots, A_d, A'$ be a partition of $[n]$ such that A_i is a maximal set of twins for $\mathcal{F}$ for $i \in [d]$, and $\mathcal{F}$ vanishes on A'. Then $|\mathcal{F}| \leq 2^d$. As $k \geq n$, the characteristic vector of A_i is contained in $(\langle \mathcal{F} \rangle_\ell)^k$. Indeed, take $v \in \mathcal{F}$ such that $v|_{A_i}$ is the all 1 vector $\mathbf{1}$, and let J be the set of $j \in [d] \setminus \{i\}$ such that $v|_{A_j}$ is $\mathbf{1}$. For each $j \in J$, since A_i, A_j are maximal sets of twins, there is $u_j \in \mathcal{F}$ such that either $u_j|_{A_i} = \mathbf{1}$ and $u_j|_{A_j} = \mathbf{0}$, or the other way around. Let J_1 be the set of j of the first type, and J_2 the set of j of the second type. Then the product of u_j over $j \in J_1$ and $(v - u_j)$ over $j \in J_2$ is the characteristic vector of A_i, as required. Now, since $\mathcal{F}$ is k-closed, ℓ divides $|A_i|$ for $i \in [d]$. But then $d \leq \lfloor n/\ell \rfloor$, and we are done with (1). Also, if $|\mathcal{F}| > 2^{\lfloor n/\ell \rfloor - 1}$, we must have $d = \lfloor n/\ell \rfloor$, which is only possible if all the sets $A_1, \dots, A_d$ have size ℓ. Therefore, (2) also holds.

Let $n > 6s\ell$. First, suppose that there exists $A \subset [n]$ such that ℓ divides $|A|$ and A is a set of twins for $\mathcal{F}$. Then the family $\mathcal{F}' = \mathcal{F}|_{[n]\setminus A}$ is also k-closed over $\mathbb{Z}_\ell$ and $|\mathcal{F}'| \geq \frac{1}{2}|\mathcal{F}|$. By our induction hypothesis, we have $|\mathcal{F}'| \leq 2^{\lfloor (n-\ell)/\ell \rfloor}$, so we get $|\mathcal{F}| \leq 2^{\lfloor n/\ell \rfloor}$, and (1) indeed holds. If $|\mathcal{F}| > 2^{\lfloor n/\ell \rfloor - 1}$, then $|\mathcal{F}'| > 2^{\lfloor (n-\ell)/\ell \rfloor - 1}$, so by our induction hypothesis there exists a partition of $[n] \setminus A$ into sets $A_1, \dots, A_{\lfloor (n-\ell)/\ell \rfloor}, A'$ satisfying (2) with respect to $\mathcal{F}'$. Setting $A_{\lfloor n/\ell \rfloor} = A$, the sets $A_1, \dots, A_{\lfloor n/\ell \rfloor}, A'$ satisfy (2) with respect to $\mathcal{F}$.

Therefore, in order to finish the proof, it is enough to show that if $|\mathcal{F}| > 2^{\lfloor n/\ell \rfloor - 1}$, then $\mathcal{F}$ has a set of twins of size divisible by ℓ. Next, we show that if $I \subset [n]$ is large, then the dimension of $\langle \mathcal{F}|_I \rangle_p$ cannot be too small for any prime p.

Claim 15. *Let p be a prime and $I \subset [n]$ such that $|I| \geq \ell$. Then*

$$|I| \leq \ell \dim(\langle \mathcal{F}|_I \rangle_p) + 3\ell.$$

Proof. Let $V = \langle \mathcal{F}|_I \rangle_p$ and $d = \dim(V)$. Then $|V \cap \{0,1\}^I| \leq 2^d$ by Lemma 14. This means that there exists some $v \in \{0,1\}^I$ and $\mathcal{F}' \subset \mathcal{F}$ such that $w|_I = v$ for every $w \in \mathcal{F}'$, and $|\mathcal{F}'| \geq |\mathcal{F}|/2^d$. Let $0 \leq m \leq \ell - 1$ such that $||v|| \equiv m \pmod{\ell}$, and in every $w \in \mathcal{F}'$, replace the coordinates in I with m coordinates of 1 entries. This gives a family $\mathcal{F}'' \subset \{0,1\}^{n-|I|+m}$ such that $\mathcal{F}''$ is k-closed over $\mathbb{Z}_\ell$ and $|\mathcal{F}''| = |\mathcal{F}'| \geq |\mathcal{F}|/2^d$. Therefore, by our induction hypothesis, we have

$$2^{\lfloor n/\ell \rfloor - d - 1} < \frac{|\mathcal{F}|}{2^d} \leq |\mathcal{F}''| \leq 2^{\lfloor (n-|I|+m)/\ell \rfloor} < 2^{\lfloor n/\ell \rfloor + 2 - |I|/\ell}.$$

Comparing the left- and right-hand-side gives the desired inequality $|I| \leq \ell d + 3\ell$. □

We can assume that $\mathcal{F}$ is non-reducible, because otherwise we are immediately done by applying our induction hypothesis. Let $A_1, \dots, A_d$ be the unique partition of $[n]$ such that A_i is a maximal set of twins for $\mathcal{F}$. Let $r \in [s]$, and apply Lemma 13 to $\mathcal{F}$ with respect to the prime power $p_r^{\alpha_r}$. Let

$$B_r = \bigcup_{\substack{i \in [d] \\ |A_i| \not\equiv 0 \pmod{p_r^{\alpha_r}}}} A_i.$$

As $\mathcal{F}$ is $2^{t+1}(p_r + \alpha_r)$-closed, we get that $\dim(\langle \mathcal{F}|_{B_r} \rangle_{p_r}) \leq \frac{6n\alpha_r}{t}$. But then by Claim 15, we also have

$$|B_r| \leq \frac{6n\alpha_r \ell}{t} + 3\ell.$$

Let $B = \bigcup_{r=1}^s B_r$, then

$$|B| \leq \sum_{r=1}^s |B_r| \leq 3s\ell + \frac{6n\ell}{t} \sum_{r=1}^s \alpha_r < n,$$

where the last inequality holds by the choice of t and noting that $n > 6sl$. Observe that B is the union of those maximal sets of twins A_i where $|A_i|$ is not divisible by ℓ. Therefore, as $|B| < n$ and $A_1, \dots, A_d$ form a partition of $[n]$, there must exists $j \in [d]$ such that ℓ divides $|A_j|$, finishing the proof. □

Let us remark that in case ℓ is a prime power, we can prove something slightly stronger following the same proof. This might be of independent interest.

Theorem 16. *Let p be a prime, $\ell = p^\alpha$, then there exists k such that the following holds. Let $\mathcal{F} \subset \{0,1\}^n$ such that $\mathcal{F}$ is k-closed over $\mathbb{Z}_\ell$. Then $\dim(\langle \mathcal{F} \rangle_p) \leq \lfloor n/\ell \rfloor$.*

Proof. We show that $k = 2^{t+1}(p+\alpha)$ suffices, where $t = 6\alpha + 1$. We will proceed by induction on n. In case $n \leq \ell$, the statement is trivial, so assume that $n > \ell$. Assume that $\mathcal{F}$ is maximal k-closed over $\mathbb{Z}_\ell$. Let $V = \langle \mathcal{F} \rangle_p$.

Suppose that there exists a set $A \subset [n]$ of twins of size ℓ for $\mathcal{F}$. If $v \in \{0,1\}^n$ is the characteristic vector of A, then $v \in \mathcal{F}$, otherwise $\{v\} \cup \mathcal{F}$ contradicts the maximality of $\mathcal{F}$. Let $\mathcal{F}' = \mathcal{F}|_{[n]\setminus A}$ and $V' = \langle \mathcal{F}' \rangle_p$, then $\dim(V') = \dim(V) - 1$ (as v is an element of V) and $\mathcal{F}'$ is also k-closed. Therefore, by our induction hypothesis, $\dim(V') \leq \lfloor (n-\ell)/\ell \rfloor$, so $\dim(V) \leq \lfloor n/\ell \rfloor$.

In the rest of the proof, we show that if $\dim(V) \geq \lfloor n/\ell \rfloor$, then there exists a set $A \subset [n]$ of twins of size ℓ for $\mathcal{F}$. We can assume that $\mathcal{F}$ is non-reducible, otherwise apply our induction hypothesis. Let $A_1, \dots, A_d$ be the unique partition of $[n]$ such that A_i is a maximal set of twins for $\mathcal{F}$. Apply Lemma 13 to $\mathcal{F}$ with respect to the prime power p^α. Let

$$B = \bigcup_{\substack{i \in [d] \\ |A_i| \not\equiv 0 \pmod{p^\alpha}}} A_i.$$

As $\mathcal{F}$ is $2^{t+1}(p+\alpha)$-closed, we get that

$$\dim(V|_B) \leq \frac{6n\alpha}{t} < \dim(V).$$

Therefore, $B \neq [n]$, and at least one of $|A_1|, \dots, |A_d|$ is divisible by p^α. Hence, $\mathcal{F}$ has a set of twins of size $\ell = p^\alpha$. This finishes the proof. □

Finally, we note that if ℓ is a prime, a similar proof shows that Theorem 16 holds for every $\mathcal{F} \subset \mathbb{F}_\ell^n$ (that is, the elements of $\mathcal{F}$ need not be 0-1 vectors).

Theorem 17. *Let p be a prime, then there exists k such that the following holds. Let $\mathcal{F} \subset \mathbb{F}_p^n$ such that $\mathcal{F}$ is k-closed. Then $\dim(\langle \mathcal{F} \rangle_p) \leq \lfloor n/p \rfloor$. In particular, $|\mathcal{F}| \leq p^{\lfloor n/p \rfloor}$, and this bound is the best possible.*

4 Proof of the main result

In this section, we show how to deduce Theorem 1 from Theorem 9. Let us start with the following variant of the well known Oddtown theorem [2], see also [1] for related results.

Lemma 18. *Let ℓ, m, n be positive integers, and let $A_1, \dots, A_m, B_1, \dots, B_m \subset [n]$ such that $\ell \nmid |A_i \cap B_i|$ for $i \in [m]$, but ℓ divides $|A_i \cap B_j|$ for $i \neq j$. Then $m \leq sn$, where s is the number of distinct prime divisors of ℓ.*

Proof. Write $\ell = p_1^{\alpha_1} \dots p_s^{\alpha_s}$, where $p_1, \dots, p_s$ are distinct primes. Let v_i and w_i be the characteristic vectors of A_i and B_i over $\mathbb{Q}$, respectively. Let $t = \lceil m/s \rceil$, then there exists $r \in [s]$ such that for at least t of the indices $i \in [m]$, we have that $|A_i \cap B_i|$ is not divisible by $p_r^{\alpha_r}$. Without loss of generality let these t indices be $1, \dots, t$. We show that $v_1, \dots, v_t$ are linearly independent (over $\mathbb{Q}$), which then implies $t \leq n$ and $m \leq sn$.

Suppose this is not the case, then there exist $c_1, \dots, c_t \in \mathbb{Z}$, not all zero, such that $\sum_{i=1}^t c_i v_i = 0$. We can assume that at least one of $c_1, \dots, c_t$ is not divisible by p_r, otherwise we can replace c_i with $c_i' = c_i/p_r$ for every $i \in [t]$. Let $k \in [t]$ be an index such that $p_r \nmid c_k$. Consider the equality

$$0 = \left\langle \sum_{i=1}^t c_i v_i, w_k \right\rangle = \sum_{i=1}^t c_i |A_i \cap B_k|.$$

We have $p_r^{\alpha_r} \mid c_i |A_i \cap B_k|$ if $i \neq k$, and $p_r^{\alpha_r} \nmid c_k |A_k \cap B_k|$, so $p_r^{\alpha_r} \nmid \langle \sum_{i=1}^t c_i v_i, w_k \rangle$, contradiction. □

Say that $\mathcal{F} \subset 2^{[n]}$ is *weakly k-closed over $\mathbb{Z}_\ell$* if the intersection of any k distinct elements of $\mathcal{F}$ is divisible by ℓ. Also, say that $\mathcal{F} \subset 2^{[n]}$ is *k-closed over $\mathbb{Z}_\ell$* if the family formed by the characteristic vectors of the elements of $\mathcal{F}$ is k-closed over $\mathbb{Z}_\ell$. So $\mathcal{F}$ is k-closed over $\mathbb{Z}_\ell$ if and only if the intersection of any k not necessarily distinct elements of $\mathcal{F}$ is divisible by ℓ. We need the following useful observation, which for $\ell = 2$ appears in [18].

Lemma 19. *Let ℓ, k be positive integers, and let s be the number of distinct prime divisors of ℓ. Let $\mathcal{F} \subset 2^{[n]}$ such that $\mathcal{F}$ is weakly k-closed over $\mathbb{Z}_\ell$. Then there exists $\mathcal{F}' \subset \mathcal{F}$ such that $|\mathcal{F}'| \geq |\mathcal{F}| - sk^2n$, and $\mathcal{F}$ is k-closed over $\mathbb{Z}_\ell$.*

Proof. Repeat the following removal operation. Suppose that $\mathcal{F}$ is not k-closed, and let t be maximal such that some t distinct elements of $\mathcal{F}$ have an intersection not divisible by ℓ. So $t < k$ because $\mathcal{F}$ is weakly k-closed. Let $\mathcal{H}$ be the t-uniform hypergraph on $\mathcal{F}$ in which $\{C_1, \dots, C_t\}$ is an edge if $|C_1 \cap \dots \cap C_t|$ is not divisible by ℓ. We claim that $\mathcal{H}$ contains no matching of size more than sn. Indeed, suppose otherwise, let $\{C_{i,1}, \dots, C_{i,t}\}$, $i \in [m]$, be the edges of a matching of size $m > sn$. For $i \in [m]$, let $A_i = C_{i,1}$ and $B_i = C_{i,1} \cap \dots \cap C_{i,t}$. Then $\ell \nmid |A_i \cap B_i|$, but ℓ divides $|A_i \cap B_j|$ for every $i \neq j$ by the maximality of t. Therefore, by Lemma 18 we get $m \leq sn$, contradiction.

Consider a maximal matching of $\mathcal{H}$, and remove every element of $\mathcal{F}$ that appears in this matching. Then $\mathcal{F}$ no longer contains t distinct sets, whose intersection is not divisible by ℓ. Repeating this procedure at most $k-1$ times, we get a family $\mathcal{F}' \subset \mathcal{F}$ such that $\mathcal{F}'$ is k-closed over $\mathbb{Z}_\ell$, and $|\mathcal{F}'| \geq |\mathcal{F}| - sk^2n$. □

Lemma 19 combined with Theorem 9 immediately implies that if $\mathcal{F} \subset 2^{[n]}$ is weakly k-closed over $\mathbb{Z}_\ell$, then $|\mathcal{F}| \leq 2^{\lfloor n/\ell \rfloor} + sk^2n$. In order to improve the term sk^2n to a constant, we use the second part of Theorem 9.

Proof of Theorem 1. Let $d = \lfloor n/\ell \rfloor$. Let $\mathcal{F} \subset \{0,1\}^n$ such that $\mathcal{F}$ is weakly k-closed over $\mathbb{Z}_\ell$. Then by Lemma 19, there exists $\mathcal{F}' \subset \mathcal{F}$ such that $|\mathcal{F}'| \geq |\mathcal{F}| - sk^2n$, and $\mathcal{F}$ is k-closed over $\mathbb{Z}_\ell$, where s is the number of distinct prime divisors of ℓ. If $|\mathcal{F}'| \leq 2^{d-1}$, then

$$|\mathcal{F}| \leq 2^{d-1} + sk^2n < 2^d,$$

where the last inequality holds if n is sufficiently large.

Suppose that $|\mathcal{F}'| > 2^{d-1}$ and $|\mathcal{F}| \geq 2^d$, otherwise we are done. Then by Theorem 9, $[n]$ can be partitioned into sets $A_1, \dots, A_d, A'$ such that A_i is a maximal set of twins for $\mathcal{F}'$ for $i \in [d]$, $|A_i| = \ell$, $|A'| \leq \ell - 1$, and $\mathcal{F}'$ vanishes on A'. Let $S \subset \{0,1\}^n$ be the atomic family containing all possible 2^d sets C such that $C \cap A_i \in \{\emptyset, A_i\}$ for every $i \in [d]$. Then $\mathcal{F}' \subset S$ and $|S \setminus \mathcal{F}'| \leq sk^2n$. Also, if n is sufficiently large, for every $i \in [d]$ we can find $k-1$ distinct sets $B_{i,1}, \dots, B_{i,k-1} \in \mathcal{F}'$ such that $A_i = \bigcap_{j=1}^{k-1} B_{i,j}$. Indeed, $\mathcal{F}'$ contains a set of the form $A_i \cup A_a \cup A_b$ for some a, b, as the number of such sets in S is at least $\binom{d-1}{2} > sk^2n$, let this set be $B_{i,1}$. Also $\mathcal{F}'$ contains $k-2$ sets that contain A_i but do not contain A_a and A_b as the number of such sets in S is $2^{d-3} > sk^2n + k$. Let these $k-2$ sets be $B_{i,2}, \dots, B_{i,k-1}$. Then $A_i = \bigcap_{j=1}^{k-1} B_{i,j}$.

Let $F \in \mathcal{F} \setminus S$. For every $i \in [d]$, we have $A_i \subset F$ or $A_i \cap F = \emptyset$, as the size of $A_i \cap F = B_{i,1} \cap \dots \cap B_{i,k-1} \cap F$ must be divisible by ℓ. Now, as $F \notin S$, we must have $F \cap A' \neq \emptyset$. But for any $H \subset A'$, where $H \neq \emptyset$, there are at most $k-1$ elements $F \in \mathcal{F} \setminus S$ such that $F \cap A' = H$, because otherwise we would have k distinct $F_1, \dots, F_k \in \mathcal{F}$ with $|F_1 \cap \dots \cap F_k| \equiv |H| \not\equiv 0 \pmod{\ell}$, a contradiction. So we see that $|\mathcal{F} \setminus S| \leq k2^{|A'|} \leq k2^{\ell-1}$, and if $\ell \mid n$ then $\mathcal{F} \subset S$. This finishes the proof. □

5 Stability

In this section, we prove Theorem 2. The proof follows easily from our earlier results.

Proof of Theorem 2. Write $\ell = p_1^{\alpha_1} \dots p_s^{\alpha_s}$, where $p_1, \dots, p_s$ are distinct primes. Let $t = \lceil 6\epsilon^{-1} \sum_{r \in [s]} \alpha_i \rceil$, then we show that $k = 2^{t+1} \max_{r \in [s]}(p_r + \alpha_r)$ suffices.

Without loss of generality, $\mathcal{F}$ is non-reducible. Let $A_1, \dots, A_d$ be the unique partition of $[n]$ into maximal sets of twins for $\mathcal{F}$. For $r \in [s]$, let

$$B_r = \bigcup_{\substack{i \in [d] \\ |A_i| \not\equiv 0 \pmod{p_r^{\alpha_r}}}} A_i.$$

Then by Lemma 13, we have $\dim(\langle \mathcal{F}|_{B_r} \rangle_{p_r}) \leq \frac{6n\alpha_r}{t}$. This implies, using Lemma 14, that

$$|\mathcal{F}|_{B_r}| \leq 2^{6n\alpha_r/t}.$$

Let $B = \bigcup_{r \in [s]} B_r$. Then B is the union of those sets A_i, whose size is not divisible by ℓ. Also, we have

$$|\mathcal{F}|_B| \leq \prod_{r \in [s]} |\mathcal{F}|_{B_r}| \leq 2^{6n(\alpha_1 + \dots + \alpha_s)/t} \leq 2^{\epsilon n}.$$

Set $X = [n] \setminus B$, then $|\mathcal{F}|_X| \geq 2^{-\epsilon n} |\mathcal{F}|$. But X is the union of sets of twins of size divisible by ℓ, so we can partition X into sets of twins of size exactly ℓ. Note that $\mathcal{F}|_X$ is a subfamily of some isomorphic copy of $S(|X|, \ell)$, finishing the proof. $\square$

6 Concluding remarks

For many problems in extremal set theory, it is natural to consider their *multipartite* (i.e. *cross*) variant. For example, Frankl and Kupavskii [7] proved a tight bound on $|\mathcal{F}_1| + \dots + |\mathcal{F}_k|$ for k set-families $\mathcal{F}_1, \dots, \mathcal{F}_k$ with no choice of disjoint $F_1, \dots, F_k$, $F_i \in \mathcal{F}_i$. In a similar vein, Bucić, Letzter, Sudakov and Tran [5] proved a tight result for the multipartite version of the Erdős–Kleitman conjecture. We refer the reader to the book [10] for additional examples.

One might wonder what can be said about the cross-version of Theorem 1. That is, let $\mathcal{F}_1, \dots, \mathcal{F}_k \subset 2^{[n]}$ such that $|F_1 \cap \dots \cap F_k|$ is divisible by ℓ for every $F_1 \in \mathcal{F}_1, \dots, F_k \in \mathcal{F}_k$. What is the maximum of $|\mathcal{F}_1| \dots |\mathcal{F}_k|$? It follows from Theorem 9 that if $\mathcal{F}_1 = \dots = \mathcal{F}_k$, then this maximum is $2^{k\lfloor n/\ell \rfloor}$, given k is sufficiently large with respect to ℓ. However, if the families $\mathcal{F}_1, \dots, \mathcal{F}_k$ are not necessarily equal, the answer is very different. Indeed, let $A_1, \dots, A_k$ be an arbitrary partition of $[n]$, and let $\mathcal{F}_i = 2^{[n] \setminus A_i}$ for $i \in [k]$. Then $F_1 \cap \dots \cap F_k = \emptyset$ for every $F_i \in \mathcal{F}_i$, $i \in [k]$, and $|\mathcal{F}_1| \dots |\mathcal{F}_k| = 2^{(k-1)n}$. We can also show that $2^{(k-1)n}$ is the maximum, which somewhat surprisingly does not depend on ℓ.

Theorem 20. *Let $\ell, k \geq 2$, and let $\mathcal{F}_1, \dots, \mathcal{F}_k \subset 2^{[n]}$ such that ℓ divides $|F_1 \cap \dots \cap F_k|$ for every $F_1 \in \mathcal{F}_1, \dots, F_k \in \mathcal{F}_k$. Then $|\mathcal{F}_1| \dots |\mathcal{F}_k| \leq 2^{(k-1)n}$.*

Proof. Let p be any prime divisor of ℓ, and let $V_i = \langle \mathcal{F}_i \rangle_p$ for $i \in [k]$ (where we identify the members of $\mathcal{F}_i$ with their characteristic vectors). Note that for every $v \in V_1 \cdot \ldots \cdot V_k$, we have $||v|| = 0$. Our goal is to show the inequality

$$\sum_{i=1}^{k} \dim(V_i) \leq (k-1)n,$$

which then immediately implies the desired bound $|\mathcal{F}_1| \dots |\mathcal{F}_k| \leq 2^{(k-1)n}$ noting that $|\mathcal{F}_i| \leq 2^{\dim(V_i)}$.

Let $V = V_1 \cap \dots \cap V_{k-1}$, then $\dim(V) \geq \sum_{i=1}^{k-1} \dim(V_i) - (k-2)n$. Also, V^{k-1} and V_k are orthogonal spaces, so $\dim(V^{k-1}) + \dim(V_k) \leq n$. Therefore, in order to finish the proof, it is enough to show that $\dim(V^{k-1}) \geq \dim(V)$, as then

$$\sum_{i=1}^{k} \dim(V_i) \leq \dim(V) + (k-2)n + \dim(V_k) \leq \dim(V^{k-1}) + \dim(V_k) + (k-2)n \leq (k-1)n.$$

Let $d = \dim(V)$, then there exist $1 \leq i_1 < \dots < i_d \leq n$ and a basis $v_1, \dots, v_d \in V$ such that $v_r(i_r) = 1$ and $v_r(i_j) = 0$ for $r \in [d]$, $j \in [d] \setminus \{r\}$. But then $v_r^{k-1} \in V^{k-1}$, and $v_r^{k-1}(i_j) = v_r(i_j)$ for $j \in [d]$. Therefore, the vectors $v_1^{k-1}, \dots, v_d^{k-1}$ are linearly independent in V^{k-1}, hence $\dim(V^{k-1}) \geq d$. $\square$

Acknowledgements. All authors were supported by the SNSF grant 200021_196965. István Tomon also acknowledges the support of Russian Government in the framework of MegaGrant no 075-15-2019-1926, and the support of MIPT Moscow.

References

[1] L. Babai and P. Frankl. "Linear Algebra Methods in Combinatorics. (With Applications to Geometry and Computer Science)." Preliminary Version, Department of Computer Science, The University of Chicago, 1992. (Version 2.1, 2020)

[2] E. R. Berlekamp. "On subsets with intersections of even cardinality." Canad. Math. Bull. 12 (1969): 471–474.

[3] E. Breuillard, B. Green and T. Tao. "The structure of approximate groups." Publications mathématiques de l'IHÉS, 116(1) (2012): 115–221.

[4] E. Breuillard, B. Green and T. Tao. "Small doubling in groups." In Erdős Centennial (pp. 129-151). Springer, Berlin, Heidelberg, 2013.

[5] M. Bucić, S. Letzter, B. Sudakov and T. Tran. "Minimum saturated families of sets." Bulletin of the London Mathematical Society, 50(4) (2018): 725–732.

[6] M. Deza, P. Erdős and P.Frankl. "Intersection properties of systems of finite sets." Proc. Lond. Math. Soc. 36 (1978): 369–384.

[7] P. Frankl and A. Kupavskii. "Two problems on matchings in set families — In the footsteps of Erdős and Kleitman." Journal of Combinatorial Theory, Series B, 138 (2019): 286–313.

[8] P. Frankl and A. M. Odlyzko. "On subsets with cardinalities of intersections divisible by a fixed integer." European Journal of Combinatorics 4(3) (1983): 215–220.

[9] P. Frankl and N. Tokushige. "Uniform eventown problems." European Journal of Combinatorics 51 (2016): 280–286.

[10] P. Frankl and N. Tokushige. "Extremal problems for finite sets." (Vol. 86). American Mathematical Soc., 2018.

[11] G. A. Freiman. "Foundations of a structural theory of set addition." American Mathematical Society, Providence, R. I., 1973. Translated from the Russian, Translations of Mathematical Monographs, Vol 37.

[12] Z. Füredi and B. Sudakov. "Extremal set-systems with restricted k-wise intersections." J. Combinatorial Theory Ser. A 105 (2004): 143–159.

[13] J. E. Graver. "Boolean designs and self-dual matroids." Linear Algebra Appl. 10 (1975): 111–128.

[14] B. Green and I. Z. Ruzsa. "Freiman's theorem in an arbitrary abelian group." J. Lond. Math. Soc. (2), 75(1) (2007): 163–175.

[15] V. Grolmusz and B. Sudakov. "On k-wise set-intersections and k-wise hamming-distances." J. Combinatorial Theory Ser. A 99 (2002): 180–190.

[16] A. M. Odlyzko. "On the ranks of some $(0, 1)$-matrices with constant row sums." J. Austral. Math. Soc. 31 (1981): 193–201.

[17] I. Z. Ruzsa. "Sums of finite sets." In Number theory (New York, 1991–1995), pages 281–293. Springer, New York, 1996.

[18] B. Sudakov and P. Vieira. "Two remarks on even and oddtown problems." SIAM J. of Discrete Math. 32 (2018): 280–295.

[19] T. Szabó and V. H. Vu. "Exact k-wise intersection theorems". Graphs Combin., 21 (2005): 247–261.

[20] T. Tao and V. H. Vu. "Additive combinatorics" (Vol. 105). Cambridge University Press, 2006.

[21] V. H. Vu. "Extremal set systems with weakly restricted intersections." Combinatorica 19 (1999): 567–587.