On symmetric intersecting families

David Ellis, Gil Kalai, and Bhargav Narayanan

Abstract. We make some progress on a question of Babai from the 1970s, namely: for $n, k \in \mathbb{N}$ with $k \leq n/2$, what is the largest possible cardinality $s(n,k)$ of an intersecting family of k-element subsets of $\{1, 2, \ldots, n\}$ admitting a transitive group of automorphisms? We give upper and lower bounds for $s(n,k)$, and show in particular that $s(n,k) = o(\binom{n-1}{k-1})$ as $n \to \infty$ if and only if $k = n/2 - \omega(n)(n/\log n)$ for some function $\omega(\cdot)$ that increases without bound, thereby determining the threshold at which 'symmetric' intersecting families are negligibly small compared to the maximum-sized intersecting families. We also exhibit connections to some basic questions in group theory and additive number theory, and pose a number of problems.

1. Introduction

A family of sets is said to be *intersecting* if any two sets in the family have nonempty intersection, and *uniform* if all the sets in the family have the same size. In this paper, we study uniform intersecting families. The most well-known result about such families is the Erdős–Ko–Rado theorem [11].

Theorem 1.1. *Let $n, k \in \mathbb{N}$ with $k \leq n/2$. If $\mathcal{A}$ is an intersecting family of k-element subsets of $\{1, 2, \ldots, n\}$, then $|\mathcal{A}| \leq \binom{n-1}{k-1}$. Furthermore, if $k < n/2$, then equality holds if and only if $\mathcal{A}$ is a star, meaning that $\mathcal{A}$ consists of all the k-element subsets of $\{1, 2, \ldots, n\}$ that contain some fixed element $i \in \{1, 2, \ldots, n\}$.*

Over the last fifty years, many variants of this theorem have been obtained. A common feature of many of these variants is that the extremal families are highly asymmetric; this is the case, for example, in the Erdős–Ko–Rado theorem itself, in the Hilton–Milner theorem [18], and in Frankl's generalisation [13] of these results. It is therefore natural to ask what happens to the maximum possible size of a uniform intersecting family when one imposes some kind of symmetry requirement on the family.

In the 1970s, Babai [1] posed the problem of determining the maximum possible size of a uniform intersecting family with transitive automorphism group; this is a very natural symmetry requirement to impose. In this paper, we make some progress on Babai's problem.

Date: 21st July 2018.

2010 *Mathematics Subject Classification.* Primary 05D05; Secondary 05E18.

Let us first give our definitions in full, and fix some notation. For a positive integer $n \in \mathbb{N}$, we denote the set $\{1, 2, \ldots, n\}$ by $[n]$. We write S_n for the symmetric group on $[n]$ and $\mathcal{P}_n$ for the power-set of $[n]$. For a permutation $\sigma \in S_n$ and a set $x \subset [n]$, we write $\sigma(x)$ for the image of x under σ, and if $\mathcal{A} \subset \mathcal{P}_n$, we write $\sigma(\mathcal{A}) = \{\sigma(x) : x \in \mathcal{A}\}$. We define the *automorphism group* of a family $\mathcal{A} \subset \mathcal{P}_n$ by

$$\mathrm{Aut}(\mathcal{A}) = \{\sigma \in S_n : \sigma(\mathcal{A}) = \mathcal{A}\}.$$

We say that $\mathcal{A} \subset \mathcal{P}_n$ is *symmetric* if $\mathrm{Aut}(\mathcal{A})$ is a transitive subgroup of S_n, i.e., if for all $i, j \in [n]$, there exists a permutation $\sigma \in \mathrm{Aut}(\mathcal{A})$ such that $\sigma(i) = j$.

For a pair of integers $n, k \in \mathbb{N}$ with $k \leq n$, let $[n]^{(k)}$ denote the family of all k-element subsets of $[n]$, and let

$$s(n, k) = \max\{|\mathcal{A}| : \mathcal{A} \subset [n]^{(k)} \text{ such that } \mathcal{A} \text{ is symmetric and intersecting}\}.$$

Of course, if $k > n/2$, then $[n]^{(k)}$ itself is a symmetric intersecting family, so $s(n, k) = \binom{n}{k}$; in studying $s(n, k)$, we may therefore restrict our attention to the case where $k \leq n/2$.

With these definitions in place, we may state the aforementioned question of Babai more precisely.

Problem 1.2. *For $n, k \in \mathbb{N}$ with $k \leq n/2$, determine $s(n, k)$.*

We remark that, in the non-uniform setting (where one studies families of sets not all of the same size, i.e., subsets of $\mathcal{P}_n$), several authors have obtained results on the maximum size of symmetric families that are intersecting (or satisfy some stronger intersection requirement); see for example the results of Frankl [12], Cameron, Frankl and Kantor [6], and the more recent results of the first and third authors [10]. Relatively little seems to be known in the uniform setting, however.

As a first step towards Problem 1.2, we focus on determining when a symmetric uniform intersecting family must be *significantly smaller* than the extremal families (of the same uniformity) in the Erdős–Ko–Rado theorem. A more precise formulation of this question is as follows.

Question 1.3. *For which $k = k(n) \leq n/2$ is $s(n, k) = o(\binom{n-1}{k-1})$ as $n \to \infty$?*

Utilising a well-known sharp threshold result of Friedgut and the second author [15], we prove the following.

Theorem 1.4. *There exists a universal constant $c > 0$ such that*

$$s(n, k) \leq 2 \exp\left(-\frac{c(n - 2k) \log n}{n}\right) \binom{n}{k}$$

for any $n, k \in \mathbb{N}$ with $k \leq n/2$.

We also give a construction showing that Theorem 1.4 is sharp up to the value of c in the regime where k/n is bounded away from zero. This construction, in conjunction with Theorem 1.4, provides a complete answer to Question 1.3.

Theorem 1.5. *If $k = k(n) \leq n/2$, then as $n \to \infty$, $s(n,k) = o(\binom{n-1}{k-1})$ if and only if*

$$k = \frac{n}{2} - \omega(n)\left(\frac{n}{\log n}\right)$$

for some function $\omega(\cdot)$ that increases without bound.

While Question 1.3 is the most basic question in the regime where the uniformity k is large compared to the size n of the ground set, the most basic question in the regime where k is small compared to n concerns the *existence* of symmetric intersecting families. Note that if $s(n,k) > 0$, then $s(n,l) > 0$ for all $l > k$; indeed, if $\mathcal{A} \subset [n]^{(k)}$ is nonempty, symmetric and intersecting, then so is $\{y \in [n]^{(l)} : x \subset y \text{ for some } x \in \mathcal{A}\}$. With this in mind, for each $n \in \mathbb{N}$, we define

$$g(n) = \min\{k \in \mathbb{N} : s(n,k) > 0\}.$$

It turns out that problem of determining the function $g(\cdot)$ is intimately connected to some longstanding open problems in group theory and additive number theory. It is not hard to show, as we shall see, that $g(n) \geq \sqrt{n}$ for all $n \in \mathbb{N}$. It is then natural to ask whether this bound is asymptotically tight, and this prompts us to raise the following question.

Question 1.6. *Is it true that $g(n) = (1 + o(1))\sqrt{n}$ for all $n \in \mathbb{N}$?*

While we are able to show that the estimate in Question 1.6 holds along various (arithmetically special) sequences of positive integers, we are unfortunately unable to settle this question entirely.

The remainder of this paper is organised as follows. We give the proof of Theorem 1.4 in Section 2. In Section 3, we describe a combinatorial approach to constructing large symmetric intersecting families in the regime where k is comparable to n, and deduce Theorem 1.5 as a consequence. In Section 4, we turn to the regime where k is comparable to $\sqrt{n}$, and describe various algebraic constructions of symmetric intersecting families in this regime. We finally conclude in Section 5 with a discussion of some open problems and related work.

2. Upper bounds

We first describe briefly the notions and tools we will need for the proof of Theorem 1.4. In what follows, all logarithms are to the base e.

We begin with the following simple observation which may be found in [6], for example; we include a proof for completeness.

Proposition 2.1. *For all $n, k \in \mathbb{N}$ with $1 < k \leq \sqrt{n}$, we have $s(n,k) = 0$.*

Proof. The proposition follows from a simple averaging argument. Indeed, let $k \leq \sqrt{n}$, suppose for a contradiction that $\mathcal{A} \subset [n]^{(k)}$ is a nonempty, symmetric intersecting family, and let $x \in \mathcal{A}$. If we choose $\sigma \in \mathrm{Aut}(\mathcal{A})$ uniformly at random, then since $\mathrm{Aut}(\mathcal{A})$ is transitive, we have

$$\mathbb{E}[|x \cap \sigma(x)|] = \frac{k^2}{n} \leq 1,$$

where the first equality above depends on the fact that

$$|\sigma \in \mathrm{Aut}(\mathcal{A}) : \sigma(i) = j| = |\sigma \in \mathrm{Aut}(\mathcal{A}) : \sigma(i) = k|$$

for all $i, j, k \in [n]$. Since $|x \cap \mathrm{Id}(x)| = k > 1$, there must exist a permutation $\sigma \in \mathrm{Aut}(\mathcal{A})$ such that $x \cap \sigma(x) = \varnothing$, contradicting the fact that $\mathcal{A}$ is intersecting. $\square$

For $0 \leq p \leq 1$, we write μ_p for the *p-biased measure* on $\mathcal{P}_n$, defined by

$$\mu_p(\{x\}) = p^{|x|}(1-p)^{n-|x|}$$

for each $x \subset [n]$, and by

$$\mu_p(\mathcal{F}) = \sum_{x \in \mathcal{F}} \mu_p(\{x\})$$

for each $\mathcal{F} \subset \mathcal{P}_n$. We say that a family $\mathcal{F} \subset \mathcal{P}_n$ is *increasing* if it is closed under taking supersets, i.e., if $x \in \mathcal{F}$ and $x \subset y$, then $y \in \mathcal{F}$. It is easy to see that if $\mathcal{F} \subset \mathcal{P}_n$ is increasing, then $p \mapsto \mu_p(\mathcal{F})$ is a monotone non-decreasing function on $[0,1]$. For a family $\mathcal{F} \subset \mathcal{P}_n$, we write $\mathcal{F}^{\uparrow}$ for the smallest increasing family containing $\mathcal{F}$; in other words, $\mathcal{F}^{\uparrow} = \{y \subset [n] : x \subset y \text{ for some } x \in \mathcal{F}\}$.

We need the following fact, which allows one to bound from above the size of a family $\mathcal{F} \subset [n]^{(k)}$ in terms of $\mu_p(\mathcal{F}^{\uparrow})$, where $p \approx k/n$; this was proved in a slightly different form by Friedgut [14]. We provide a proof for completeness.

Lemma 2.2. *Let $n, k \in \mathbb{N}$ and suppose that $0 < p, \phi < 1$ satisfy*

$$p \geq \frac{k}{n} + \frac{\sqrt{2n \log(1/\phi)}}{n}.$$

Then for any family $\mathcal{F} \subset [n]^{(k)}$, we have

$$\mu_p(\mathcal{F}^{\uparrow}) > (1-\phi)\frac{|\mathcal{F}|}{\binom{n}{k}}.$$

Proof. Let $\delta = |\mathcal{F}|/\binom{n}{k}$ and let X be a binomial random variable with distribution $\mathrm{Bin}(n,p)$. First, for each $l \geq k$, the local LYM inequality implies that

$$\frac{|\mathcal{F}^{\uparrow} \cap [n]^{(l)}|}{\binom{n}{l}} \geq \frac{|\mathcal{F}|}{\binom{n}{k}} = \delta.$$

Now, for any $\eta > 0$, it follows from a standard Chernoff bound that

$$\mathbb{P}(X < (1-\eta)np) < \exp(-\eta^2 np/2).$$

Hence,

$$\begin{aligned}\mu_p(\mathcal{F}^{\uparrow}) &\geq \sum_{l=k}^{n} p^l (1-p)^{n-l} \binom{n}{l} \delta \\ &= \mathbb{P}(X \geq k)\delta \\ &> (1-\phi)\delta,\end{aligned}$$

where the last inequality above follows from a standard calculation. $\square$

We will also require the following sharp threshold result due to Friedgut and the second author [15].

Theorem 2.3. *There exists a universal constant $c_0 > 0$ such that the following holds for all $n \in \mathbb{N}$. Let $0 \leq p < 1$, $0 < \varepsilon < 1$ and let $\mathcal{F} \subset \mathcal{P}_n$ be a symmetric increasing family. If $\mu_p(\mathcal{F}) > \varepsilon$, then $\mu_q(\mathcal{F}) > 1 - \varepsilon$, where*

$$q = \min\left\{1, p + c_0 \frac{\log(1/\varepsilon)}{\log n}\right\}. \qquad \square$$

The idea of the proof of Theorem 1.4 is as follows. Let $\mathcal{A} \subset [n]^{(k)}$ be a symmetric intersecting family. We first use Lemma 2.2 to bound $|\mathcal{A}|/\binom{n}{k}$ from above in terms of $\mu_p(\mathcal{A}^{\uparrow})$, where $p \approx k/n$; we then use Theorem 2.3, together with the simple fact that $\mu_{1/2}(\mathcal{A}^{\uparrow}) \leq 1/2$, to bound $\mu_p(\mathcal{A}^{\uparrow})$, and hence $|\mathcal{A}|$, from above. Let us also remark that this strategy of 'approximating' the uniform measure on $[n]^{(k)}$ with the p-biased measure μ_p, where $p \approx k/n$, has proven useful on a number of different occasions in the study of uniform intersecting families; see [14, 8, 9], for example.

Proof of Theorem 1.4. Let $n, k \in \mathbb{N}$ with $k \leq n/2$, let $\mathcal{A} \subset [n]^{(k)}$ be a symmetric intersecting family, and set $\delta = |\mathcal{A}|/\binom{n}{k}$.

In the light of Proposition 2.1, we may suppose without loss of generality that $k > \sqrt{n}$. By the Erdős–Ko–Rado theorem, we have $\delta \leq k/n \leq 1/2$, so we may also assume (by choosing $c > 0$ to be sufficiently small, say $c < (\log 2)^2$) that $k \leq n/2 - 10n/\log n$.

First, applying Lemma 2.2 with $p = k/n + \sqrt{(2\log 2)n}/n$ and $\phi = 1/2$, we see that

$$\mu_p(\mathcal{A}^{\uparrow}) > \frac{\delta}{2}.$$

Next, since $\mathcal{A}$ is symmetric, so is $\mathcal{A}^{\uparrow}$. We may therefore apply Theorem 2.3 with $\varepsilon = \delta/2$ to deduce that $\mu_q(\mathcal{A}^{\uparrow}) > 1/2$, where

$$q = \min\left\{1, p + c_0\left(\frac{\log(2/\delta)}{\log n}\right)\right\}.$$

Since $\mathcal{A}^{\uparrow}$ is increasing, the function $r \mapsto \mu_r(\mathcal{A}^{\uparrow})$ is monotone non-decreasing on $[0,1]$. Also, since $\mathcal{A}$ is intersecting, so is $\mathcal{A}^{\uparrow}$, and therefore $\mu_{1/2}(\mathcal{A}^{\uparrow}) \le 1/2$. Now, as $\mu_{1/2}(\mathcal{A}^{\uparrow}) \le 1/2$ and $\mu_q(\mathcal{A}^{\uparrow}) > 1/2$, the monotonicity of $r \mapsto \mu_r(\mathcal{A}^{\uparrow})$ implies that

$$p + c_0\left(\frac{\log(2/\delta)}{\log n}\right) > \frac{1}{2}.$$

Rearranging this inequality, we see that

$$\delta < 2\exp\left(-\frac{(1-2p)\log n}{2c_0}\right) \le 2\exp\left(-\frac{c(n-2k)\log n}{n}\right),$$

where $c > 0$ is a universal constant; this proves the theorem. □

3. Lower bounds for large k

In this section, we give a construction showing that Theorem 1.4 is sharp up to the value of the constant c in the exponent for many choices of $k = k(n)$.

Given $n, k \in \mathbb{N}$ with $k \le n$, we identify $[n]$ with $\mathbb{Z}_n$, we identify a subset $S \subset \mathbb{Z}_n$ with its characteristic vector $\chi_S \in \{0,1\}^{\mathbb{Z}_n}$, and we define $\mathcal{F}(n,k)$ to be the family of all k-element subsets $S \subset \mathbb{Z}_n$ such that the longest run of consecutive ones in χ_S is longer than the longest run of consecutive zeros in χ_S. Slightly less formally, we take $\mathcal{F}(n,k)$ to consist of all the cyclic strings of n zeros and ones which contain exactly k ones and in which the longest run of consecutive ones is longer than the longest run of consecutive zeros.

It is clear that $\mathcal{F}(n,k)$ is symmetric. It is also easy to check that $\mathcal{F}(n,k)$ is intersecting. Indeed, given $S, T \in \mathcal{F}(n,k)$, suppose without loss of generality that the longest run of consecutive ones in S is at least as long as that in T. Choose a run of consecutive ones in S of the maximum length; these cannot be all zeros in T because otherwise T would have a longer run of consecutive ones than S. Therefore, $S \cap T \ne \varnothing$. An analysis of $\mathcal{F}(n,k)$ yields the following.

Lemma 3.1. *There exists a universal constant $\hat{C} > 0$ such that if $k = k(n) \in \mathbb{N}$ is such that $\sqrt{n}\log n \le k \le n/2$ for all $n \in \mathbb{N}$, then*

$$|\mathcal{F}(n,k)| \ge \exp\left(-(1+\hat{C}/\log n)\left(\frac{\log n - \log k}{\log n - \log(n-k)}\right)\log n + \log n\right)\binom{n}{k}. \qquad \square$$

A comparison of Theorem 1.4 and Lemma 3.1 shows that Theorem 1.4 is sharp up to the constant factor in the exponent when k/n is bounded away from zero and $1/2 - k/n = \Omega(1/\log n)$.

We note that the non-uniform case of this construction, i.e., the family of all cyclic strings of n zeros and ones in which the longest run of consecutive ones is longer than the longest run of consecutive zeros, shows that the Kahn–Kalai–Linial theorem [19] cannot be improved by more than a constant factor for intersecting families; see [20] for more details.

In the case where k/n is close to $1/2$, Lemma 3.1 implies the following.

Lemma 3.2. *For each $C > 0$, there exists $c > 0$ such that for any $n, k \in \mathbb{N}$ with $1/2(1 - C/\log n) \leq k/n \leq 1/2$, we have*

$$|\mathcal{F}(n,k)| \geq c\binom{n-1}{k-1}. \qquad \square$$

To prove Lemma 3.1, we need the following.

Lemma 3.3. *Let $k < n$. The number of cyclic strings of n zeros and ones with exactly k ones and a run of consecutive zeros of length at least l is at most $\frac{1}{4}\binom{n}{k}$, provided*

$$l \geq \frac{\log n + 2\log 2}{\log n - \log(n-k)}.$$

Proof. The number of such strings is at most $n\binom{n-l}{k}$, since (possibly overcounting) there are n choices for the position of the run of l consecutive zeros, and then $\binom{n-l}{k}$ choices for the positions of the ones. We have

$$\frac{n\binom{n-l}{k}}{\binom{n}{k}} = \frac{n(n-k)(n-k-1)\dots(n-k-l+1)}{n(n-1)\dots(n-l+1)} \leq n\left(\frac{n-k}{n}\right)^l \leq \frac{1}{4}$$

provided $l \geq (\log n + 2\log 2)/(\log n - \log(n-k))$, as required. $\square$

We now observe the following.

Claim 3.4. *If $1 \leq l \leq k \leq n/2$, then the number of cyclic strings of length n with k ones and a run of consecutive ones of length at least l is at most the number of cyclic strings of length n with k ones and a run of consecutive zeros of length at least l.*

Proof of Claim 3.4. For any family $\mathcal{A} \subset [n]^{(k)}$, we write

$$\partial^+\mathcal{A} = \{x \in [n]^{(k+1)} : \ x \supset y \text{ for some } y \in \mathcal{A}\}$$

for the *upper shadow* of $\mathcal{A}$, and

$$\partial^{+(t)}(\mathcal{A}) = \{x \in [n]^{(k+t)} : \ x \supset y \text{ for some } y \in \mathcal{A}\} = \partial^+(\partial^{+(t-1)}\mathcal{A})$$

for each $t \in \mathbb{N}$ with $t \leq n-k$. The local LYM inequality (see [4], for example) states that for any integers $1 \leq k < n$ and any family $\mathcal{A} \subset [n]^{(k)}$, we have

$$\frac{|\partial^+\mathcal{A}|}{\binom{n}{k+1}} \geq \frac{|\mathcal{A}|}{\binom{n}{k}}.$$

If $k \le n/2$, then iterating this inequality $n-2k$ times, and using the fact that $\binom{n}{k} = \binom{n}{n-k}$, yields

$$\frac{|\partial^{+(n-2k)}\mathcal{A}|}{\binom{n}{k}} = \frac{|\partial^{+(n-2k)}\mathcal{A}|}{\binom{n}{n-k}} \ge \frac{|\mathcal{A}|}{\binom{n}{k}}. \tag{1}$$

Now let $1 \le l \le k \le n/2$, let $\mathcal{A}$ be the family of cyclic strings of length n with k ones and a run of consecutive ones of length at least l, and let $\mathcal{B}$ be the family of cyclic strings of length n with $n-k$ ones and a run of consecutive ones of length at least l. Clearly, we have $\mathcal{B} = \partial^{+(n-2k)}\mathcal{A}$, and therefore by (1), we have $|\mathcal{A}| \le |\mathcal{B}|$. But, by flipping zeros and ones, it is clear that $|\mathcal{B}|$ is precisely the number of cyclic strings of length n with k ones and a run of consecutive zeros of length at least l, proving the claim. $\square$

The following is immediate from Lemma 3.3 and Claim 3.4.

Corollary 3.5. *Let $k \le n/2$. The number of cyclic strings of length n with k ones and no run of l consecutive zeros or ones is at least $\frac{1}{2}\binom{n}{k}$, provided*

$$l \ge \frac{\log n + 2\log 2}{\log n - \log(n-k)}. \qquad \square$$

We are now ready to prove Lemma 3.1.

Proof of Lemma 3.1. Choose $l_0 \in \mathbb{N}$ such that

$$l_0 - 1 \ge \frac{\log(n-l_0-2) + 2\log 2}{\log(n-l_0-2) - \log(n-k-2)}. \tag{2}$$

Observe that $\mathcal{F}(n,k)$ contains all cyclic strings of length n with k ones, precisely one run of l_0 consecutive ones, all other runs of consecutive ones having length at most l_0-2, and no run of l_0 consecutive zeros. We claim that if $l_0 < n/2$, then the number of such strings is at least

$$\frac{n}{2}\binom{n-l_0-2}{k-l_0}. \tag{3}$$

Indeed, there are n choices for the position of the run of l_0 consecutive ones, and there must be a zero on each side of this run of ones. Now, there are at least $\frac{1}{2}\binom{n-l_0-2}{k-l_0}$ choices for the remainder of the cyclic string (by Corollary 3.5), since if we take a cyclic string of length $n-l_0-2$ which contains no run of l_0-1 consecutive ones or zeros, and then insert (at some point) a run of l_0 consecutive ones with a zero on either side into this string, then the resulting string has the desired property provided $l_0 < n/2$.

It is easily checked that if $\sqrt{n}\log n \le k \le n/2$, then we may choose $l_0 \in \mathbb{N}$ satisfying (2) such that

$$l_0 = (1 + O(1/\log n))\frac{\log n}{\log n - \log(n-k)}. \tag{4}$$

Indeed, choose

$$l_0 = (1+\varepsilon)\frac{\log n}{\log n - \log(n-k)},$$

where $\varepsilon = \Theta(1/\log n)$ is to be chosen later. Then, using the fact that $\log n - \log(n-k) = \Omega(k/n)$, we have $l_0 = O((n\log n)/k) = O(\sqrt{n})$, and therefore

$$\begin{aligned}
\frac{\log(n-l_0-2)-\log(n-k-2)}{\log n - \log(n-k)} &= 1 - \frac{\log(1+\frac{l_0+2}{n-l_0-2}) - \log(1+\frac{2}{n-k-2})}{\log n - \log(n-k)} \\
&= 1 - O(\tfrac{l_0}{k}) \\
&= 1 - O((n\log n)/k^2) \\
&= 1 - O(1/\log n).
\end{aligned} \tag{5}$$

Taking $\varepsilon = L/\log n$ for some absolute constant L, we have

$$l_0 - 1 \geq \left(1 + \frac{L}{2\log n}\right)\frac{\log n}{\log n - \log(n-k)}. \tag{6}$$

Finally, we clearly have

$$\frac{\log n}{\log(n-l_0-2)+2\log 2} = 1 - O(1/\log n). \tag{7}$$

Putting (5), (6) and (7) together, we obtain

$$\begin{aligned}
l_0 - 1 &\geq \left(1+\frac{L}{2\log n}\right)\frac{\log n}{\log n - \log(n-k)} \\
&\geq \left(1+\frac{L}{2\log n}\right)(1-O(1/\log n))\frac{\log(n-l_0-2)+2\log 2}{\log(n-l_0-2)-\log(n-k-2)},
\end{aligned}$$

yielding (2) provided L is sufficiently large.

Provided n is at least an absolute constant (which we may assume), we have $l_0 < n/2$, and therefore, using (3), we have

$$\begin{aligned}
|\mathcal{F}(n,k)| &\geq \frac{n}{2}\binom{n-l_0-2}{k-l_0} \geq \frac{n}{2}\binom{n-l_0-2}{k-l_0-2} \\
&\geq \frac{n}{2}\left(\frac{k-l_0-2}{n-l_0-2}\right)^{l_0+2}\binom{n}{k} \\
&\geq \Omega(1)\cdot n \cdot (\tfrac{k}{n})^{l_0+2}\binom{n}{k} \\
&\geq n\cdot\exp\left(-(1+O(1/\log n))\left(\frac{\log n - \log k}{\log n - \log(n-k)}\right)\log n\right)\binom{n}{k},
\end{aligned}$$

proving the lemma. $\square$

We now deduce Lemma 3.2 from Lemma 3.1.

Proof of Lemma 3.2. Defining $\eta = 1/2 - k/n$, we have $\eta \le C/\log n$, and

$$\frac{\log n - \log k}{\log n - \log(n-k)} = \frac{\log 2 - \log(1-2\eta)}{\log 2 - \log(1+2\eta)} = 1 + O(\eta).$$

Hence, it follows from Lemma 3.1 that

$$|\mathcal{F}(n,k)| \ge n \cdot \exp\Big(-(1+\hat{C}/\log n)(1+O(C/\log n))\log n\Big)\binom{n}{k} \ge c\binom{n-1}{k-1}$$

provided c is sufficiently small depending on $\hat{C}$ and C, as required. □

With the above estimates in hand, Theorem 1.5 is now easily established.

Proof of Theorem 1.5. The result is immediate from Theorem 1.4 and Lemma 3.2. □

4. Lower bounds for small k

In the previous two sections, we focussed on estimating the largest possible cardinality $s(n,k)$ of a symmetric intersecting subfamily of $[n]^{(k)}$. We now turn our attention to estimating the smallest possible uniformity $k = g(n)$ for which there exists a nonempty, symmetric intersecting subfamily of $[n]^{(k)}$. To this end, we will investigate the set

$$\mathcal{S} = \{(n,k) \in \mathbb{N}^2 : s(n,k) > 0\}.$$

Along the way, we describe some constructions of symmetric intersecting families that are larger than $\mathcal{F}(n,k)$ for certain values of n and k.

We have already seen (in Proposition 2.1) that $g(n) > \sqrt{n}$ for all $n \ge 2$. Let us now consider upper bounds on $g(n)$.

It is easy to check that $\mathcal{F}(n,k) \ne \varnothing$ provided $n \le \lfloor k^2/4 \rfloor + k$. (Consider the cyclic string $1^\ell(0^{\ell-1}1)^t 0^{\ell-1}$, where $\ell = \lfloor k/2 \rfloor + 1$, $t = \lceil k/2 \rceil - 1$, and $n = \lfloor k^2/4 \rfloor + k$; here, as usual, if S is a string, S^N denotes S concatenated with itself N times.) This observation implies that

$$g(n) \le 2\sqrt{n} \tag{8}$$

for all $n \in \mathbb{N}$. To improve (8), we note a strong connection between the problem of determining $g(n)$ and the problem of finding a so-called *difference cover* in an Abelian group. If G is a finite Abelian group and $S \subset G$, we say that S is a *difference cover for* G if $S - S = G$, i.e., if $\{i - j : i,j \in S\} = G$; we then define

$$h(G) = \min\{|S| : S \text{ is a difference cover for } G\}.$$

Note that if $S \subset G$, then S is a difference cover for G if and only if the family of all the translates of S is an intersecting family of subsets of G. This observation yields the following.

Lemma 4.1. *For all $n \in \mathbb{N}$, we have $g(n) \le h(\mathbb{Z}_n)$, with equality holding in the case where n is prime.*

Proof. Let $h = h(\mathbb{Z}_n)$ and write $\mathbb{Z}_n^{(h)}$ for the family of h-element subsets of $\mathbb{Z}_n$. By definition, there exists $S \in \mathbb{Z}_n^{(h)}$ such that $S - S = \mathbb{Z}_n$. Let $\mathcal{A} = \{S + j : j \in \mathbb{Z}_n\} \subset \mathbb{Z}_n^{(h)}$ denote the family of all the translates of S. Then $\mathcal{A}$ is clearly symmetric and intersecting. Hence, $g(n) \leq h$, proving the first part of the claim.

Now suppose that n is prime, and let $g(n) = k$. Let $\mathcal{A} \subset [n]^{(k)}$ be a nonempty, symmetric intersecting family. Since $\mathrm{Aut}(\mathcal{A}) \leq S_n$ is transitive, the orbit-stabilizer theorem implies that n divides $|\mathrm{Aut}(\mathcal{A})|$, and therefore by Cauchy's theorem, $\mathrm{Aut}(\mathcal{A})$ has a cyclic subgroup H of order n. Let $\sigma \in S_n$ be a generator of H; then σ is an n-cycle, and by relabelling the ground set $[n]$ if necessary, we may assume that $\sigma = (1\ 2\ \ldots\ n)$ (in the standard cycle notation). Fix $x \in \mathcal{A}$ and note that $\mathcal{B} = \{x, \sigma(x), \ldots, \sigma^{n-1}(x)\}$ is also a nonempty, symmetric intersecting family as $H \leq \mathrm{Aut}(\mathcal{B})$. Clearly, $\mathcal{B}$ consists of all the cyclic translates, modulo n, of x. If we regard x as a subset of $\mathbb{Z}_n$, then since $\mathcal{B}$ is intersecting, we have $x - x = \mathbb{Z}_n$, i.e., x is a difference cover for $\mathbb{Z}_n$. Hence, $h(\mathbb{Z}_n) \leq k$ and it follows that $h(\mathbb{Z}_n) = g(n)$ when n is prime, as required. □

We now describe how existing constructions of difference covers lead to an improvement of (8). We say that $S \subset \mathbb{Z}$ is a *difference cover for n* if $[n] \subset S - S$. For each $n \in \mathbb{N}$, let $\pi_n \colon \mathbb{Z} \to \mathbb{Z}_n$ denote the natural projection modulo n defined by $\pi_n(i) = i \pmod{n}$ for all $i \in \mathbb{Z}$. Note that if $S \subset \mathbb{Z}$ is a difference cover for $\lfloor n/2 \rfloor$, then $\pi_n(S)$ is a difference cover for $\mathbb{Z}_n$. Building on work of Rédei and Rényi [23] and of Leech [21], Golay [17] proved that for any $n \in \mathbb{N}$, there exists a difference cover for n of size at most $\sqrt{cn}$, where $c < 2.6572$ is an absolute constant. It follows that for any $n \in \mathbb{N}$, we have

$$g(n) \leq h(\mathbb{Z}_n) \leq 1.1527\sqrt{n}.$$

Unfortunately, one cannot hope to answer Question 1.6 in the affirmative purely by projecting difference covers for $\lfloor n/2 \rfloor$ into $\mathbb{Z}_n$ and using the fact that $g(n) \leq h(\mathbb{Z}_n)$; this is a consequence of a result of Rédei and Rényi [23] which asserts that if $S \subset \mathbb{Z}$ is a difference cover for n, then

$$|S| \geq \sqrt{\left(2 + \frac{4}{3\pi}\right)n}.$$

In view of Lemma 4.1, we are led to the following question, which being a natural question in its own right, has also occurred independently to others; see [2], for instance.

Question 4.2. *Is it true that $h(\mathbb{Z}_n) = (1 + o(1))\sqrt{n}$ for all $n \in \mathbb{N}$?*

By Lemma 4.1, an affirmative answer to this question would imply an affirmative answer to Question 1.6. We remark that Question 4.2 is a 'covering' problem whose 'packing' counterpart has received a lot of attention. If G is an Abelian group and $S \subset G$, we say that S is a *Sidon set* in G if for any non-identity element $g \in G$, there exists at most one ordered pair $(s_1, s_2) \in S^2$ such that $g = s_1 - s_2$. For $n \in \mathbb{N}$, let

$$\lambda(n) = \max\{|S| : S \subset \mathbb{Z}_n \text{ such that } S \text{ is a Sidon set}\}.$$

The determination of $\lambda(n)$ is a well-known open problem; see [7], for example. In particular, the following remains open.

Question 4.3. *Is it true that $\lambda(n) = (1 - o(1))\sqrt{n}$ for all $n \in \mathbb{N}$?*

The constructions of Singer [25] and Bose [5] yield affirmative answers to Question 4.3 when n is of the form $q^2 + q + 1$ or $q^2 - 1$ respectively, where q is a prime power, and a construction due to Ruzsa [24] does so when n is of the form $p^2 - p$, where p is prime; as observed by Banakh and Gavrylkiv [2], these constructions of Singer, Bose and Ruzsa yield efficient difference covers as well, so we also have affirmative answers to Questions 4.2 and 1.6 for all n of the aforementioned form.

Returning to the question of determining $g(n)$, we have shown that

$$\lfloor \sqrt{n} \rfloor + 1 \leq g(n) \leq 1.1527\sqrt{n} \tag{9}$$

for all $n \geq 2$. It turns out that the precise value of $g(n)$ has a nontrivial dependence on the arithmetic properties of n; indeed, the lower bound in (9) is sharp for some positive integers, but strict for others. We record these facts, as well as some other properties of $g(\cdot)$, below. Since these observations don't seem to be enough resolve Question 1.6 completely, we chose not to include detailed proofs of the claims below.

- **G1** Observe that if $d \geq 2$ and there exists a transitive projective plane of order d, then writing $n = d^2 + d + 1$, we have $s(n, k) > 0$ if and only if $k \geq d + 1$. Indeed, if $k \leq d$, then $s(n, k) = 0$ by Proposition 2.1, while if $k \geq d + 1$, then we start with a transitive projective plane $\mathbb{P}$ of order d and take the family of all k-element subsets of the points of $\mathbb{P}$ containing a line of $\mathbb{P}$ to see that $s(n, k) > 0$ in this case. In particular, for any odd prime power q, we have $s(q^2 + q + 1, k) > 0$ if and only if $k \geq q + 1$; it follows that the lower bound in (9) is sharp for any $n = q^2 + q + 1$ with q an odd prime power, and consequently, we also get an affirmative answer to Question 1.6 for all positive integers of this form.
- **G2** On the other hand, the lower bound in (9) is not tight for $n = 43$, for example. It was shown by Lovász [22] and Füredi [16] that if $d \geq 2$ is such that $n = d^2 + d + 1$ is prime, then $s(n, d + 1) > 0$ if and only if there exists a transitive projective plane of order d. Consequently, $s(43, 7) = 0$ since 43 is prime and there exists no projective plane of order 6, so the lower bound in (9) is not sharp in general. The Bateman–Horn conjecture [3] would imply that $d^2 + d + 1$ is prime for infinitely many positive integers d which are not themselves prime powers; taken together with the aforementioned observation of Lovász and Füredi along with the non-existence conjecture for projective planes whose order is not a prime

power, this would imply that $s(d^2+d+1, d+1) = 0$ for infinitely many $d \in \mathbb{N}$, and consequently that the lower bound in (9) is not sharp for infinitely many positive integers.

We can use other finite geometries in the place of projective planes to bound $g(\cdot)$; this allows us to answer Question 1.6 in the affirmative for various sequences of positive integers with suitable 'arithmetic structure'.

G3 For any prime power q, by taking the dual affine plane $\mathbb{DA}^2(\mathbb{F}_q)$ over the finite field $\mathbb{F}_q$ and considering the family of lines of $\mathbb{DA}^2(\mathbb{F}_q)$, then writing $n = q^2 + q$, we have $s(n, k) > 0$ if $k \geq q+1$; this yields an affirmative answer to Question 1.6 for any $n \in \mathbb{N}$ of this form.

These constructions based on projective planes and dual affine planes have natural analogues based upon higher-dimensional projective spaces and higher-dimensional dual affine spaces, enabling us to answer Question 1.6 affirmatively for some other infinite sequences of integers.

G4 Fix $r \in \mathbb{N}$, let q be an prime power and consider the $(2r)$-dimensional projective space $\mathbb{P}^{2r}(\mathbb{F}_q)$ over $\mathbb{F}_q$. Then the family of all r-dimensional projective subspaces of $\mathbb{P}^{2r}(\mathbb{F}_q)$ gives us an affirmative answer to Question 1.6 for all $n = (q^{2r+1} - 1)/(q-1)$ with q a prime power.

G5 Next, fix $r \in \mathbb{N}$, let q be a prime power, and consider the $(2r)$-dimensional dual affine space $\mathbb{DA}^{2r}(\mathbb{F}_q)$ over $\mathbb{F}_q$. Then the family of all r-flats of $\mathbb{DA}^{2r}(\mathbb{F}_q)$ gives us an affirmative answer to Question 1.6 for all $n = q(q^{2r}-1)/(q-1)$ with q a prime power.

For completeness, let us also record the following fact.

G6 The observation of Banakh and Gavrylkiv [2] mentioned earlier shows that $g(n) = (1+o(1))\sqrt{n}$ whenever $n = q^2 - 1$ for some prime power q, or $n = p^2 - p$ for some prime p. Consequently, we have an affirmative answer to Question 1.6 for any $n \in \mathbb{N}$ of the aforementioned forms.

Finally, we demonstrate using a tensor product construction that $\mathcal{S}$ is closed under taking pointwise products. For a set $x \subset [n]$, we define its *characteristic vector* $\chi_x \in \{0,1\}^n$ by $(\chi_x)_i = 1$ if $i \in x$ and $(\chi_x)_i = 0$ otherwise. Given two sets $x \subset [n]$ and $y \subset [m]$, we define their *tensor product* $x \otimes y$ to be the subset of $[nm]$ whose characteristic vector $\chi_{x \otimes y}$ is given by

$$(\chi_{x\otimes y})_{(i-1)m+j} = (\chi_x)_i (\chi_y)_j$$

for all $i \in [n]$ and $j \in [m]$. For two families $\mathcal{A} \subset \mathcal{P}_n$ and $\mathcal{B} \subset \mathcal{P}_m$, we define their tensor product by

$$\mathcal{A} \otimes \mathcal{B} = \{x \otimes y : x \in \mathcal{A}, y \in \mathcal{B}\};$$

note that $\mathcal{A} \otimes \mathcal{B} \subset \mathcal{P}_{nm}$ and that $|\mathcal{A} \otimes \mathcal{B}| = |\mathcal{A}||\mathcal{B}|$. Now observe that if $\mathcal{A} \subset [n]^{(k)}$ and $\mathcal{B} \subset [m]^{(l)}$, then $\mathcal{A} \otimes \mathcal{B} \subset [nm]^{(kl)}$, and furthermore, if $\mathcal{A}$ and $\mathcal{B}$ are symmetric and intersecting, then so is $\mathcal{A} \otimes \mathcal{B}$. It follows that

$$s(nm, kl) \geq s(n, k)s(m, l)$$

for all $k, l, m, n \in \mathbb{N}$, and in particular, if $(n, k), (m, l) \in \mathcal{S}$, then $(nm, kl) \in \mathcal{S}$.

G7 The above observation implies that $g(\cdot)$ is submultiplicative, i.e., we have

$$g(nm) \leq g(n)g(m)$$

for all $n, m \in \mathbb{N}$. This fact may be used to answer Question 1.6 affirmatively for some additional sequences of positive integers; for example, we conlcude that the answer to Question 1.6 is in the affirmative for all $n = (q_1^2 + q_1 + 1)(q_2^2 + q_2 + 1)$ with q_1 and q_2 both prime powers, and so on.

5. Conclusion

A number of interesting open problems remain. Theorem 1.4 and Lemma 3.1 together determine the order of magnitude of $\log(\binom{n}{k}/s(n, k))$ when k/n is bounded away from zero by a positive constant. The gap between our upper and lower bounds for $s(n, k)$ is somewhat worse for smaller k, and it would be of interest to improve Theorem 1.4 in the regime where $k = o(n)$.

Determining $s(n, k)$ precisely for all $k \leq n/2$ would appear to be a challenging problem. We conjecture that for any $\delta > 0$, if n is sufficiently large depending on δ and $(1 + \delta)\sqrt{n}\log n \leq k \leq n/2$, then

$$s(n, k) = |\mathcal{F}(n, k)|.$$

Note that if n is sufficiently large depending on δ and $(1 + \delta)\sqrt{n}\log n \leq k \leq n/2$, then the family $\mathcal{F}(n, k)$ yields a larger symmetric intersecting family than any of the algebraic constructions in Section 4.

Determining the asymptotic behaviour of $g(n)$ is another problem that merits further investigation. We have established various estimates in Section 4, but even the fundamental question of deciding whether $g(n)/\sqrt{n}$ converges in the limit as $n \to \infty$ still remains open.

Acknowledgements

The second author wishes to acknowledge support from ERC Advanced Grant 320924, NSF grant DMS-1300120 and BSF grant 2014290, and the third author was partially supported by NSF Grant DMS-1800521. We would also like to thank two anonymous referees for their careful reading of the paper.

References

1. L. Babai, Personal communication. 1
2. T. Banakh and V. Gavrylkiv, *Difference bases in cyclic groups*, J. Algebra Appl. **18** (2019), No. 5, 11pp. 11, 12, 13
3. P. T. Bateman and R. A. Horn, *A heuristic asymptotic formula concerning the distribution of prime numbers*, Math. Comp. **16** (1962), 363–367. 12
4. B. Bollobás, *Combinatorics: Set systems, hypergraphs, families of vectors and combinatorial probability*, Cambridge University Press, Cambridge, 1986. 7
5. R. C. Bose, *An affine analogue of Singer's theorem*, J. Indian Math. Soc. **6** (1942), 1–15. 12
6. P. J. Cameron, P. Frankl, and W. M. Kantor, *Intersecting families of finite sets and fixed-point-free* 2*-elements*, European J. Combin. **10** (1989), 149–160. 2, 3
7. J. Cilleruelo, I. Z. Ruzsa, and C. Vinuesa, *Generalized Sidon sets*, Adv. Math. **225** (2010), 2786–2807. 12
8. I. Dinur and E. Friedgut, *Intersecting families are essentially contained in juntas*, Combin. Probab. Comput **18** (2009), 107–122. 5
9. D. Ellis, N. Keller, and N. Lifshitz, *Stability versions of Erdős–Ko–Rado type theorems, via isoperimetry*, Preprint, arXiv:1604.02160. 5
10. D. Ellis and B. Narayanan, *On symmetric* 3*-wise intersecting families*, Proc. Amer. Math. Soc. **145** (2017), 2843–2847. 2
11. P. Erdős, C. Ko, and R. Rado, *Intersection theorems for systems of finite sets*, Quart. J. Math. Oxford **12** (1961), 313–320. 1
12. P. Frankl, *Regularity conditions and intersecting hypergraphs*, Proc. Amer. Math. Soc. **82** (1981), 309–311. 2
13. ______, *Erdős-Ko-Rado theorem with conditions on the maximal degree*, J. Combin. Theory Ser. A **46** (1987), 252–263. 1
14. E. Friedgut, *On the measure of intersecting families, uniqueness and stability*, Combinatorica **28** (2008), 503–528. 4, 5
15. E. Friedgut and G. Kalai, *Every monotone graph property has a sharp threshold*, Proc. Amer. Math. Soc. **124** (1996), 2993–3002. 2, 5
16. Z. Füredi, *Maximum degree and fractional matchings in uniform hypergraphs*, Combinatorica **1** (1981), 155–162. 12
17. M. J. E. Golay, *Notes on the representation of* $1, 2, \ldots, n$ *by differences*, J. London Math. Soc. **4** (1972), 729–734. 11
18. A. J. W. Hilton and E. C. Milner, *Some intersection theorems for systems of finite sets*, Quart. J. Math. Oxford, Series 2 **18** (1967), 369–384. 1
19. J. Kahn, G. Kalai, and N. Linial, *The influence of variables on Boolean functions*, Proceedings of the 29th Annual Symposium on Foundations of Computer Science, IEEE Computer Society, Washington, DC, 1988, pp. 68–80. 6

20. G. Kalai, *Answer to "Kahn–Kalai–Linial for intersecting upsets"*, MathOverflow, `https://mathoverflow.net/questions/105086/`. 7
21. J. Leech, *On the representation of* $1, 2, \ldots, n$ *by differences*, J. London Math. Soc. **43** (1938), 377–385. 11
22. L. Lovász, *On minimax theorems of combinatorics*, Mat. Lapok **26** (1975), 209–264. 12
23. L. Rédei and A. Rényi, *On the representation of the numbers* $1, 2, \ldots, n$ *by differences*, Mat. Sb. **24** (1949), 385–389. 11
24. I. Z. Ruzsa, *Solving an equation in a set of integers, I*, Acta Arith. **65** (1993), 259–282. 12
25. J. Singer, *A theorem in finite projective geometry and some applications to number theory*, Trans. Amer. Math. Soc. **43** (1938), 377–385. 12

School of Mathematics, University of Bristol, Fry Building, Woodland Road, Bristol BS8 1UG, UK

Email address: `david.ellis@bristol.ac.uk`

Einstein Institute of Mathematics, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Jerusalem 91904, Israel

Email address: `kalai@math.huji.ac.il`

Department of Mathematics, Rutgers University, Piscataway NJ 08854, USA

Email address: `narayanan@math.rutgers.edu`