Multicolor Sunflowers

Dhruv Mubayi* Lujia Wang†

October 19, 2017

Abstract

A sunflower is a collection of distinct sets such that the intersection of any two of them is the same as the common intersection C of all of them, and $|C|$ is smaller than each of the sets. A longstanding conjecture due to Erdős and Szemerédi (solved recently in [9, 7], see also [23]) was that the maximum size of a family of subsets of $[n]$ that contains no sunflower of fixed size $k > 2$ is exponentially smaller than 2^n as $n \to \infty$. We consider the problems of determining the maximum sum and product of k families of subsets of $[n]$ that contain no sunflower of size k with one set from each family. For the sum, we prove that the maximum is

$$(k-1)2^n + 1 + \sum_{s=0}^{k-2} \binom{n}{s}$$

for all $n \geq k \geq 3$, and for the $k = 3$ case of the product, we prove that the maximum is

$$\left(\frac{1}{8} + o(1)\right) 2^{3n}.$$

We conjecture that for all fixed $k \geq 3$, the maximum product is $(1/8 + o(1))2^{kn}$.

1 Introduction

Throughout the paper, we write $[n] = \{1, \ldots, n\}$, $2^{[n]} = \{S : S \subset [n]\}$ and $\binom{[n]}{s} = \{S : S \subset [n], |S| = s\}$. A family $\mathcal{A} \subset 2^{[n]}$ is s-uniform if further $\mathcal{A} \subset \binom{[n]}{s}$. A *sunflower* (or *strong* Δ*-system*) with k petals is a collection of k sets $\mathcal{S} = \{S_1, \ldots, S_k\}$ such that $S_i \cap S_j = C$ for all $i \neq j$, and $S_i \setminus C \neq \emptyset$ for all $i \in [k]$. The common intersection C is called the *core* of the sunflower and the sets $S_i \setminus C$ are called the *petals*. In 1960, Erdős and Rado [11] proved a fundamental result regarding the existence of sunflowers in a large family of sets of uniform

*Department of Mathematics, Statistics, and Computer Science, University of Illinois, Chicago, IL 60607, Email:mubayi@uic.edu. Research partially supported by NSF grants DMS-0969092 and DMS-1300138

†Department of Mathematics, Statistics, and Computer Science, University of Illinois, Chicago, IL 60607, Email:lwang203@uic.edu.

size, which is now referred to as the *sunflower lemma*. It states that if $\mathcal{A}$ is a family of sets of size s with $|\mathcal{A}| > s!(k-1)^s$, then $\mathcal{A}$ contains a sunflower with k petals. Later in 1978, Erdős and Szemerédi [12] gave the following upper bound when the underlying set has n elements.

Theorem 1 (Erdős, Szemerédi [12]). *There exists a constant c such that if $\mathcal{A} \subset 2^{[n]}$ with $|\mathcal{A}| > 2^{n-c\sqrt{n}}$ then $\mathcal{A}$ contains a sunflower with 3 petals.*

In the same paper, they conjectured that for n sufficiently large, the maximum number of sets in a family $\mathcal{A} \subset 2^{[n]}$ with no sunflowers with three petals is at most $(2-\epsilon)^n$ for some absolute constant $\epsilon > 0$. This conjecture, often referred to as the *weak sunflower lemma*, is closely related to the algorithmic problem of matrix multiplication [1] and remained open for nearly forty years. Recently, this was settled via the polynomial method by Ellenberg–Gijswijt [9] and Croot–Lev–Pach [7] (see also Naslund–Sawin [23]).

A natural way to generalize problems in extremal set theory is to consider versions for multiple families or so-called multicolor or cross-intersecting problems. Beginning with the famous Erdős–Ko–Rado theorem [10], which states that an intersecting family of k-element subsets of $[n]$ has size at most $\binom{n-1}{k-1}$, provided $n \geq 2k$, several generalizations were proved for multiple families that are cross-intersecting. In particular, Hilton [16] showed in 1977 that if t families $\mathcal{A}_1, \ldots, \mathcal{A}_t \subset \binom{[n]}{k}$ are cross intersecting (meaning that $A_i \cap A_j \neq \emptyset$ for all $(A_i, A_j) \in \mathcal{A}_i \times \mathcal{A}_j$) and if $n/k \leq t$, then $\sum_{i=1}^{t} |\mathcal{A}_i| \leq t\binom{n-1}{k-1}$. On the other hand, results of Pyber [24] in 1986, that were later slightly refined by Matsumoto and Tokushige [20] and Bey [2], showed that if two families $\mathcal{A} \subset \binom{[n]}{k}$, $\mathcal{B} \subset \binom{[n]}{l}$ are cross-intersecting and $n \geq \max\{2k, 2l\}$, then $|\mathcal{A}||\mathcal{B}| \leq \binom{n-1}{k-1}\binom{n-1}{l-1}$. These are the first results about bounds on sums and products of the size of cross-intersecting families. More general problems were considered recently, for example for cross t-intersecting families (i.e. pair of sets from distinct families have intersection of size at least t) and r-cross intersecting families (any r sets have a nonempty intersection where each set is picked from a distinct family) and labeled crossing intersecting families, see [4, 14, 15]. A more systematic study of multicolored extremal problems (with respect to the sum of the sizes of the families) was initiated by Keevash, Saks, Sudakov, and Verstraëte [17], and continued in [3, 18]. Cross-intersecting versions of Erdős' problem on weak Δ-systems (for the product of the size of two families) were proved by Frankl and Rödl [13] and by the first author and Rödl [22].

In this note, we consider multicolor versions of sunflower theorems. Quite surprisingly, these basic questions appear not to have been studied in the literature.

Definition 2. Let $A_i \in \mathcal{A}_i \subset 2^{[n]}$ for $i = 1, \ldots, k$. Then $(A_i)_{i=1}^k$ is a *sunflower* with k petals if there exists $C \subset [n]$ such that

- $A_i \cap A_j = C$ for all $i \neq j$, and
- $A_i \setminus C \neq \emptyset$, for all $i \in [k]$.

Say that $(\mathcal{A}_i)_{i=1}^k$ is *sunflower-free* if it contains no sunflower with k petals.

For any k families that are sunflower-free, the problem of upper bounding the size of any single family is uninteresting, since there is no restriction on a particular family. So we are interested in the sum and product of the sizes of these families.

Given integers n and k, let

$$\mathcal{F}(n,k) = \{(\mathcal{A}_i)_{i=1}^k : \mathcal{A}_i \subset 2^{[n]} \text{ for } i \in [k] \text{ and } (\mathcal{A}_i)_{i=1}^k \text{ is sunflower-free}\},$$

$$S(n,k) = \max_{(\mathcal{A}_i)_{i=1}^k \in \mathcal{F}(n,k)} \sum_{i=1}^k |\mathcal{A}_i| \qquad \text{and} \qquad P(n,k) = \max_{(\mathcal{A}_i)_{i=1}^k \in \mathcal{F}(n,k)} \prod_{i=1}^k |\mathcal{A}_i|.$$

Our two main results are sharp or nearly sharp estimates on $S(n,k)$ and $P(n,3)$. By Theorem 1 (or [7, 9, 23]) we obtain that

$$S(n,3) \leq 2 \cdot 2^n + 2^{n-c\sqrt{n}}.$$

Indeed, if $|\mathcal{A}| + |\mathcal{B}| + |\mathcal{C}|$ is larger than the RHS above then $|\mathcal{A} \cap \mathcal{B} \cap \mathcal{C}| > 2^{n-c\sqrt{n}}$ by the pigeonhole principle and we find a sunflower in the intersection which contains a sunflower. Our first result removes the last term to obtain an exact result.

Theorem 3. *For* $n \geq k \geq 3$

$$S(n,k) = (k-1)2^n + 1 + \sum_{s=0}^{k-2} \binom{n}{s}.$$

The problem of determining $P(n,k)$ seems more difficult than that of determining $S(n,k)$. Our bounds for general k are quite far apart, but in the case $k = 3$ we can refine our argument to obtain an asymptotically tight bound.

Theorem 4.

$$P(n,3) = \left(\frac{1}{8} + o(1)\right) 2^{3n}.$$

We conjecture that a similar result holds for all $k \geq 3$.

Conjecture 1. *For each fixed* $k \geq 3$,

$$P(n,k) = \left(\frac{1}{8} + o(1)\right) 2^{kn}.$$

In the next two sections we give the proofs of Theorems 3 and 4.

2 Sums

In order to prove Theorem 3, we first deal with s-uniform families and prove a stronger result. Given a sunflower $\mathcal{S} = (A_i)_{i=1}^k$, define its *core size* to be $c(\mathcal{S}) = |C|$, where $C = A_i \cap A_j, i \neq j$.

Lemma 5. *Given integers $s \geq 1$ and c with $0 \leq c \leq s-1$, let n be an integer such that $n \geq c + k(s-c)$. For $i = 1, \ldots, k$, let $\mathcal{A}_i \subset \binom{[n]}{s}$ such that $(\mathcal{A}_i)_{i=1}^k$ contains no sunflower with k petals and core size c. Then*

$$\sum_{i=1}^{k} |\mathcal{A}_i| \leq (k-1)\binom{n}{s}.$$

Furthermore, this bound is tight.

Proof. Randomly take an ordered partition of $[n]$ into $k+2$ parts $X_1, X_2, \ldots, X_{k+2}$ such that $|X_1| = n - (c + k(s-c)), |X_2| = c$, and $|X_i| = s - c$ for $i = 3, \ldots, k+2$, with uniform probability for each partition. For each partition, construct the bipartite graph

$$G = (\{\mathcal{A}_i : i = 1, \ldots, k\} \cup \{X_2 \cup X_j : j \in [3, k+2]\}, E)$$

where a pair $\{\mathcal{A}_i, X_2 \cup X_j\} \in E$ if and only if $X_2 \cup X_j \in \mathcal{A}_i$. If there exists a perfect matching in G, then we will get a sunflower with k petals and core size c, since X_2 will be the core. This shows that G has matching number at most $k-1$. Then König's theorem implies that the random variable $|E(G)|$ satisfies

$$|E(G)| \leq (k-1)k. \tag{1}$$

Another way to count the edges of G is through the following expression:

$$|E(G)| = \sum_{i=1}^{k} \sum_{j=3}^{k+2} \chi_{\{X_2 \cup X_j \in \mathcal{A}_i\}},$$

where χ_S is the characteristic function of the event S. Taking expectations and using (1) we obtain

$$\mathbb{E}\left(\sum_{i=1}^{k} \sum_{j=3}^{k+2} \chi_{\{X_2 \cup X_j \in \mathcal{A}_i\}}\right) \leq (k-1)k. \tag{2}$$

By linearity of expectation,

$$\mathbb{E}\left(\sum_{i=1}^{k} \sum_{j=3}^{k+2} \chi_{\{X_2 \cup X_j \in \mathcal{A}_i\}}\right) = \sum_{i=1}^{k} \sum_{j=3}^{k+2} \mathbb{P}\left(X_2 \cup X_j \in \mathcal{A}_i\right) = \sum_{i=1}^{k} \sum_{j=3}^{k+2} \sum_{A \in \mathcal{A}_i} \mathbb{P}\left(A = X_2 \cup X_j\right).$$

Since the partition of $[n]$ is taken uniformly, for any j with $3 \leq j \leq k+2$, the set $X_2 \cup X_j$ covers all possible s-subsets of $[n]$ with equal probability. Hence for any $A \in \mathcal{A}_i$, we have

$$\mathbb{P}(A = X_2 \cup X_j) = \frac{1}{\binom{n}{s}}.$$

So we have

$$\mathbb{E}\left(\sum_{i=1}^{k} \sum_{j=3}^{k+2} \chi_{\{X_2 \cup X_j \in \mathcal{A}_i\}}\right) = \sum_{i=1}^{k} \sum_{j=3}^{k+2} \sum_{A \in \mathcal{A}_i} \frac{1}{\binom{n}{s}} = \sum_{i=1}^{k} |\mathcal{A}_i| \frac{k}{\binom{n}{s}}.$$

Hence by (2),

$$\sum_{i=1}^{k} |\mathcal{A}_i| \leq (k-1)\binom{n}{s}.$$

The bound shown above is tight, since we can take $\mathcal{A}_1 = \mathcal{A}_2 = \ldots = \mathcal{A}_{k-1} = \binom{[n]}{s}$, and $\mathcal{A}_k = \emptyset$. $\square$

Now we use this lemma to prove Theorem 3.

Proof of Theorem 3. Recall that $n \geq k \geq 3$ and we are to show that

$$S(n,k) = (k-1)2^n + 1 + \sum_{s=n-k+2}^{n} \binom{n}{s}.$$

We first show the lower bound by the following example: Let $\mathcal{A}_i = 2^{[n]}$ for $i = 1 \ldots, k-1$ and $\mathcal{A}_k = \{\emptyset\} \cup \{S \subset [n] : |S| \geq n-k+2\}$. To see that $(\mathcal{A}_i)_{i=1}^k$ is sunflower-free, notice that any sunflower uses a set from $\mathcal{A}_k$. The empty set does not lie in any sunflower. So if a set of size at least $n-k+2$ appears in a sunflower S with k petals, it requires at least $k-1$ other points, but then the total number of points in S is at least $n+1$, a contradiction.

To see the upper bound, given families $(\mathcal{A}_i)_{i=1}^k \in \mathcal{F}(n,k)$, we define $\mathcal{A}_{i,s} = \mathcal{A}_i \cap \binom{[n]}{s}$ for each $i \in [k]$ and integer $s \in [0,n]$. This gives a partition of each family $\mathcal{A}_i$ into $n+1$ subfamilies. Since $(\mathcal{A}_i)_{i=1}^k$ is sunflower-free, so is $(\mathcal{A}_{i,s})_{i=1}^k$ for all $s \in [0,n]$. Now, for each $s = 1, 2, \ldots, n-k+1$, let $c = s-1$. Then $0 \leq c \leq s-1$, and $c + k(s-c) = s-1+k \leq n$. Therefore, by Lemma 5, $\sum_{i=1}^k |\mathcal{A}_{i,s}| \leq (k-1)\binom{n}{s}$ for all $s \in [n-k+1]$. For $s > n-k+1$, the trivial bound for this sum is $k\binom{n}{s}$. So we get,

$$\begin{aligned}
\sum_{i=1}^{k} |\mathcal{A}_i| = \sum_{i=1}^{k}\sum_{s=0}^{n} |\mathcal{A}_{i,s}| &= \sum_{s=0}^{n}\sum_{i=1}^{k} |\mathcal{A}_{i,s}| \\
&= \sum_{i=1}^{k} |\mathcal{A}_{i,0}| + \sum_{s=1}^{n-k+1}\sum_{i=1}^{k} |\mathcal{A}_{i,s}| + \sum_{s=n-k+2}^{n}\sum_{i=1}^{k} |\mathcal{A}_{i,s}| \\
&\leq k\binom{n}{0} + \sum_{s=1}^{n-k+1} (k-1)\binom{n}{s} + \sum_{s=n-k+2}^{n} k\binom{n}{s} \\
&\leq \sum_{s=0}^{n} (k-1)\binom{n}{s} + \binom{n}{0} + \sum_{s=n-k+2}^{n} \binom{n}{s} \\
&= (k-1)2^n + 1 + \sum_{s=n-k+2}^{n} \binom{n}{s} \\
&= (k-1)2^n + 1 + \sum_{s=0}^{k-2} \binom{n}{s}.
\end{aligned}$$

This completes the proof. $\square$

3 Products

From the bound on the sum of the families that do not contain a sunflower, we deduce the following bound on the product by using the AM-GM inequality.

Corollary 6. *Fix $k \geq 3$. As $n \to \infty$,*

$$\left(\frac{1}{8} + o(1)\right) 2^{kn} \leq P(n,k) \leq \left(\left(\frac{k-1}{k}\right)^k + o(1)\right) 2^{kn}.$$

Proof. The upper bound follows from Theorem 3 and the AM-GM inequality,

$$\prod_{i=1}^{k} |\mathcal{A}_i| \leq \left(\frac{\sum_{i=1}^{k} |\mathcal{A}_i|}{k}\right)^k \leq \left((1+o(1))\frac{(k-1)2^n}{k}\right)^k = (1+o(1))\left(\frac{k-1}{k}\right)^k 2^{kn}.$$

For the lower bound, we take

$$\mathcal{A}_1 = \mathcal{A}_2 = \{S \subset [n] : 1 \in S\} \cup \{[2,n]\},$$

$$\mathcal{A}_3 = \{S \subset [n] : 1 \notin S\} \cup \{S \subset [n] : |S| \geq n-1\},$$

and $\mathcal{A}_4 = \mathcal{A}_5 = \ldots = \mathcal{A}_k = 2^{[n]}$. A sunflower with k petals must use three sets from $\mathcal{A}_1, \mathcal{A}_2$, and $\mathcal{A}_3$, call them A_1, A_2, A_3 respectively. These three sets form a sunflower with three petals. If any of these sets is of size at least $n-1$, then it will be impossible to form a 3-petal sunflower with the other two sets. So by their definitions, we have $1 \in A_1 \cap A_2$, but $1 \notin A_3$, which implies $A_1 \cap A_2 \neq A_1 \cap A_3$, a contradiction. So $(\mathcal{A}_i)_{i=1}^k$ is sunflower-free. The sizes of these families are $|\mathcal{A}_1| = |\mathcal{A}_2| = 2^{n-1}+1, |\mathcal{A}_3| = 2^{n-1}+n$ and $|\mathcal{A}_i| = 2^n$ for $i \geq 4$. Thus,

$$\prod_{i=1}^{k} |\mathcal{A}_i| = \left(\frac{1}{8} + o(1)\right) 2^{kn}$$

as required. ∎

For any positive integer k we have $(\frac{k-1}{k})^k < 1/e$, so Corollary 6 implies the upper bound $(1/e + o(1))2^{kn}$ for all $k \geq 3$. For $k = 3$, we will improve the factor in the upper bound from $(2/3)^3 = 0.29629\cdots$ to our conjectured value of 0.125.

The main part of our proof is Lemma 7 below, which proves a much better bound than $S(n,3) = (2+o(1))2^n$ for the sum of three sunflower-free families under the assumption that all of them contain a positive proportion of sets.

Lemma 7. *For all $\epsilon > 0$ there exists $n_0 = n_0(\epsilon) > 0$ such that the following holds for $n > n_0$. Let $\mathcal{A}_i \subset 2^{[n]}$ with $|\mathcal{A}_i| \geq \epsilon 2^n$ for $i \in [3]$, and suppose that $(\mathcal{A}_i)_{i=1}^3$ is sunflower-free. Then*

$$|\mathcal{A}_1| + |\mathcal{A}_2| + |\mathcal{A}_3| \leq \left(\frac{3}{2} + \epsilon\right) 2^n.$$

Lemma 7 immediately implies Theorem 4 by the AM-GM inequality as shown below.

Proof of Theorem 4. Let $\epsilon \in (0, 1/8)$, n_0 be obtained from Lemma 7 and $n > n_0$. Suppose there is an i, such that $|\mathcal{A}_i| < \epsilon 2^n$. Then

$$\prod_{i=1}^{3} |\mathcal{A}_i| < \epsilon 2^n \cdot 2^n \cdot 2^n < \frac{1}{8} \cdot 2^{3n}.$$

So we may assume that $|\mathcal{A}_i| \geq \epsilon 2^n$ for all i. Thus, by the AM-GM inequality and Lemma 7,

$$\prod_{i=1}^{3} |\mathcal{A}_i| \leq \left(\frac{|\mathcal{A}_1| + |\mathcal{A}_2| + |\mathcal{A}_3|}{3} \right)^3 \leq \left(\frac{1}{2} + \frac{\epsilon}{3} \right)^3 2^{3n} < \left(\frac{1}{8} + \epsilon \right) 2^{3n}$$

which is the bound sought. □

In the rest of this section we prove Lemma 7.

3.1 Proof of Lemma 7

We begin with the following lemma, which uses ideas similar to those used in the proof of Lemma 2.1 of [17].

Lemma 8. *Let $k \geq 3$, $\mathcal{A}_1, \dots \mathcal{A}_k$ be families of subsets of $[n]$ that are sunflower-free. For any real number $\epsilon > 0$, if $|\mathcal{A}_i| \geq \epsilon 2^n$ for all i, then there exists $\delta = \delta(\epsilon) > 0$ and k families $\mathcal{B}_1, \dots, \mathcal{B}_k$ such that the following holds.*

- $|\mathcal{B}_i| \geq \delta 2^n$ *for* $i = 1, \dots, k$,
- $\sum_{i=1}^{k} |\mathcal{A}_i| \leq \sum_{i=1}^{k} |\mathcal{B}_i| + \left(\frac{\epsilon}{2}\right) 2^n$,
- $(\mathcal{B}_i)_{i=1}^{k}$ *is sunflower-free,*
- $\mathcal{B}_1, \dots, \mathcal{B}_k$ *form a laminar system, that is, either* $\mathcal{B}_i \cap \mathcal{B}_j = \emptyset$, $\mathcal{B}_i \subset \mathcal{B}_j$, *or* $\mathcal{B}_j \subset \mathcal{B}_i$ *for all* $i \neq j$.

Proof. The families $\mathcal{A}_i$, $i = 1, \dots, k$ form a collection of subsets of $2^{[n]}$, hence they induce a partition of $2^{[n]}$ into at most 2^k parts. More precisely, the disjoint parts (some may be empty) in this partition are

$$\mathcal{X}_I = \bigcap_{i \in I} \mathcal{A}_i \bigcap_{i \in [k] \setminus I} \mathcal{A}_i^c, \text{ where } I \subset [k].$$

Take $\delta = \epsilon/(k 2^k)$. For each $I \subset [k]$, if $|\mathcal{X}_I| < \delta 2^n$, update the $\mathcal{A}_i$s by deleting $\mathcal{X}_I$ from all $\mathcal{A}_i$s that contain it, that is, all $\mathcal{A}_i$s with $i \in I$. At the end of this process, let the resulting families be $\mathcal{A}'_i, i = 1 \dots, k$. Now, all $\mathcal{X}_I$s that are nonempty have size at least $\delta 2^n$.

For each original $\mathcal{A}_i$, the families $\mathcal{A}_i \cap \mathcal{A}_j, j \in [k] \setminus \{i\}$ induce a partition on it into at most 2^{k-1} parts. So, after the above deletion steps the remaining $\mathcal{A}'_i$ has size at least

$$|\mathcal{A}'_i| \geq \epsilon 2^n - 2^{k-1}\delta 2^n = \epsilon 2^n - 2^{k-1}\frac{\epsilon}{k2^k}2^n = \left(1 - \frac{1}{2k}\right)\epsilon 2^n.$$

If $\mathcal{X}_I < \delta 2^n$, it is deleted from all $|I|$ of the $\mathcal{A}_i$s that contain it. Hence, the total number of deleted parts with repetition is at most

$$\sum_{i=1}^{n} i\binom{k}{i} = k2^{k-1}.$$

So, we have

$$\sum_{i=1}^{k} |\mathcal{A}_i| \leq \sum_{i=1}^{k} |\mathcal{A}'_i| + k2^{k-1}\delta 2^n = \sum_{i=1}^{k} |\mathcal{A}'_i| + \left(\frac{\epsilon}{2}\right)2^n.$$

Two families $\mathcal{A}$ and $\mathcal{B}$ are said to be *crossing* if all three of $\mathcal{A} \cap \mathcal{B}$, $\mathcal{A} \setminus \mathcal{B}$ and $\mathcal{B} \setminus \mathcal{A}$ are nonempty. For each pair of crossing families $\mathcal{A}'_i$ and $\mathcal{A}'_j$, replace $\mathcal{A}'_i$ and $\mathcal{A}'_j$ by $\mathcal{A}'_i \cap \mathcal{A}'_j$ and $\mathcal{A}'_i \cup \mathcal{A}'_j$. Call the resulting families $\mathcal{B}_1, \ldots, \mathcal{B}_k$.

Notice first that at the end of the process (which terminates after at most $\binom{k}{2}$ steps, because it increases the number of inclusion related pairs at every step), the families $\mathcal{B}_1, \ldots, \mathcal{B}_k$ contain no crossing pairs, hence form a laminar system. Secondly, the sum of the sizes of the families remains the same, since $|X \cap Y| + |X \cup Y| = |X| + |Y|$ for all sets X, Y. Hence we get

$$\sum_{i=1}^{k} |\mathcal{A}_i| \leq \sum_{i=1}^{k} |\mathcal{A}'_i| + \left(\frac{\epsilon}{2}\right)2^n = \sum_{i=1}^{k} |\mathcal{B}_i| + \left(\frac{\epsilon}{2}\right)2^n.$$

Next, notice that all parts of the partition induced by $\mathcal{A}'_i, i = 1, \ldots, k$ have size at least $\delta 2^n$. Moreover, the steps of replacing two crossing families by their intersection and union only create new families that consists of the union of nonempty parts. This yields that $|\mathcal{B}_i| \geq \delta 2^n$ for all $i \in [k]$.

Finally, we claim that $(\mathcal{B}_i)_{i=1}^k$ is sunflower-free. The families $(\mathcal{A}'_i)_{i=1}^k$ are certainly sunflower-free because $\mathcal{A}'_i \subset \mathcal{A}_i$ for all i and $(\mathcal{A}_i)_{i=1}^k$ is sunflower-free. So we are left to show that the steps of removing crossing pairs do not introduce sunflowers.

Suppose we have families $(\mathcal{C}_i)_{i=1}^k$, w.l.o.g, the crossing pair $\mathcal{C}_1, \mathcal{C}_2$ are replaced by $\mathcal{C}_1 \cap \mathcal{C}_2$ and $\mathcal{C}_1 \cup \mathcal{C}_2$, and suppose that $C_i, i = 1, \ldots, k$ with $C_1 \in \mathcal{C}_1 \cap \mathcal{C}_2$, $C_2 \in \mathcal{C}_1 \cup \mathcal{C}_2$ and $C_i \in \mathcal{C}_i, i \geq 3$ is a sunflower in the resulting families. Then, w.l.o.g, C_2 is in $\mathcal{C}_2$. Thus we find that $C_i, i = 1, \ldots, k$ is also a sunflower in $(\mathcal{C}_i)_{i=1}^k$. This completes the proof. □

We will use the following lemma which follows from well-known properties of binomial coefficients (we omit the standard proofs).

Lemma 9. *For each $\delta > 0$, there exists a real number $\alpha = \alpha(\delta)$ and integer n_0 such that for $n > n_0$, every family $\mathcal{A}$ of subsets of $[n]$ with size $|\mathcal{A}| \geq \delta 2^n$ contains a set S with $|S| \in [n/2 - \alpha\sqrt{n}, n/2 + \alpha\sqrt{n}]$. Further, for each $\gamma \in (0, \delta)$, there exists a $\beta = \beta(\gamma)$, such that all but at most $\gamma 2^n$ elements in $\mathcal{A}$ have size in $[n/2 - \beta\sqrt{n}, n/2 + \beta\sqrt{n}]$.*

Now we have all the necessary ingredients to prove Lemma7.

Proof of Lemma 7. Let $\delta = \epsilon/(3 \cdot 2^3) = \epsilon/(24)$ as in the proof of Lemma 8. By Theorem 1, we have $|\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{A}_3| \leq 2^{n - c\sqrt{n}} < \delta 2^n$ for large enough n. Apply Lemma 8 to obtain families $\mathcal{B}_i, i = 1, 2, 3$ such that

- $|\mathcal{B}_i| \geq \delta 2^n$ for $i = 1, 2, 3$,
- $\sum_{i=1}^{3} |\mathcal{A}_i| \leq \sum_{i=1}^{3} |\mathcal{B}_i| + \left(\frac{\epsilon}{2}\right) 2^n$,
- $(\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3)$ is sunflower-free,
- $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$ form a laminar system.

Moreover, since $|\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{A}_3| < \delta 2^n$, the intersection of all three families is deleted from all three of them in the process of forming $\mathcal{B}_i$s which yields $\mathcal{B}_1 \cap \mathcal{B}_2 \cap \mathcal{B}_3 = \emptyset$. The rest of the proof is devoted to showing the claim below.

Claim 10.
$$|\mathcal{B}_1| + |\mathcal{B}_2| + |\mathcal{B}_3| \leq \left(\frac{3}{2} + \frac{\epsilon}{2}\right) 2^n.$$

Proof. The laminar system formed by the three families with an empty common intersection falls into the following three types. Let $\{\mathcal{A}, \mathcal{B}, \mathcal{C}\} = \{\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3\}$ and $a := |\mathcal{A}|, b := |\mathcal{B}|$, and $c := |\mathcal{C}|$.

Case I. $\mathcal{A}, \mathcal{B}, \mathcal{C}$ are mutually disjoint.

In this case, trivially we have $a + b + c \leq 2^n$ which is even better than what we need.

Case II. $\mathcal{A} \supset \mathcal{B}$ and $\mathcal{A} \cap \mathcal{C} = \emptyset$. Since $|\mathcal{C}| \geq \delta 2^n$, we may pick an $S \in \mathcal{C}$ with $|S| \in [n/2 - \alpha\sqrt{n}, n/2 + \alpha\sqrt{n}]$ by Lemma 9. Now for each subset $T \subset S$, consider the subfamily of $\mathcal{B}$ defined by
$$\mathcal{B}_T = \{B \in \mathcal{B} : B \cap S = T\}.$$
Clearly, these subfamilies form a partition of $\mathcal{B}$, i.e. $\mathcal{B} = \bigsqcup_{T \subset S} \mathcal{B}_T$. Now we define a new family derived from $\mathcal{B}'_T$
$$\mathcal{B}'_T = \{B \setminus T : B \in \mathcal{B}_T\}.$$
There is a naturally defined bijection between $\mathcal{B}_T$ and $\mathcal{B}'_T$, so $|\mathcal{B}_T| = |\mathcal{B}'_T|$.

Claim. $b \leq (1 + \frac{\epsilon}{2}) 2^{n-1}$.

Proof. We first show that $\mathcal{B}'_T \setminus \{\emptyset\}$ is an intersecting family if $T \subsetneq S$. Indeed, suppose there are disjoint nonempty sets $B_1, B_2 \in \mathcal{B}'_T$, then we find a sunflower consisting of $B_1 \cup T \in \mathcal{B} \subset \mathcal{A}, B_2 \cup T \in \mathcal{B}$ and $S \in \mathcal{C}$. So $|\mathcal{B}'_T| \le 2^{n-|S|-1} + 1$, which yields the following upper bound for $|\mathcal{B}|$:

$$\begin{aligned} b &= \sum_{T \subset S} |\mathcal{B}_T| = \sum_{T \subset S} |\mathcal{B}'_T| = \sum_{T \subsetneq S} |\mathcal{B}'_T| + |\mathcal{B}'_S| \\ &\le (2^{|S|} - 1)(2^{n-|S|-1} + 1) + 2^{n-|S|} = 2^{n-1} + 2^{|S|} - 2^{n-|S|-1} - 1 + 2^{n-|S|} \\ &\le 2^{n-1} + 2^{n/2+\alpha\sqrt{n}} - 2^{n-(n/2+\alpha\sqrt{n})-1} + 2^{n-(n/2-\alpha\sqrt{n})} \\ &= 2^{n-1} + 2^{n/2+\alpha\sqrt{n}} - 2^{n/2-\alpha\sqrt{n}-1} + 2^{n/2+\alpha\sqrt{n}} \\ &\le \left(1 + \frac{\epsilon}{2}\right) 2^{n-1}, \end{aligned}$$

where the last inequality holds for large enough n. □

Since $\mathcal{A} \cap \mathcal{C} = \emptyset$, the Claim implies that

$$a + b + c \le 2^n + \left(1 + \frac{\epsilon}{2}\right) 2^{n-1} = \left(\frac{3}{2} + \frac{\epsilon}{2}\right) 2^n.$$

Case III. $\mathcal{A} \supset (\mathcal{B} \cup \mathcal{C})$ **and** $\mathcal{B} \cap \mathcal{C} = \emptyset$**.** We first fix $\gamma = \min\{\delta, \epsilon/12\}$, find $\beta = \beta(\gamma)$ as in Lemma 9. Then all but at most $\gamma 2^n \le (\epsilon/12)2^n$ sets in each family are of size in $[n/2 - \beta\sqrt{n}, n/2 + \beta\sqrt{n}]$. Hence we have

$$a + b + c \le |\mathcal{A}_\beta| + |\mathcal{B}_\beta| + |\mathcal{C}_\beta| + \frac{\epsilon}{4} \cdot 2^n,$$

where $\mathcal{F}_\beta = \{F \in \mathcal{F} : n/2 - \beta\sqrt{n} \le |F| \le n/2 + \beta\sqrt{n}\}$. It remains to show that

$$|\mathcal{A}_\beta| + |\mathcal{B}_\beta| + |\mathcal{C}_\beta| \le \left(\frac{3}{2} + \frac{\epsilon}{4}\right) 2^n.$$

We may assume $\mathcal{A} = \mathcal{A}_\beta, \mathcal{B} = \mathcal{B}_\beta$ and $\mathcal{C} = \mathcal{C}_\beta$; our task is to prove $a+b+c \le (3/2 + \epsilon/4)\, 2^n$.

Consider a pair of sets $(B, C) \in \mathcal{B} \times \mathcal{C}$ which satisfies the following two conditions:

- $B \cup C \ne [n]$,
- $B \setminus C \ne \emptyset$ and $C \setminus B \ne \emptyset$.

Let $A = \overline{B \triangle C} = (B \cap C) \cup \overline{B \cup C}$. Then $A \notin \mathcal{A}$, otherwise A, B, C together form a sunflower. Hence the number of such As is at most $2^n - a$.

We claim that for each such A, there are at most $(1 + \epsilon/4)2^{n-1}$ pairs $(B, C) \in \mathcal{B} \times \mathcal{C}$ with the two properties above such that $A = \overline{B \triangle C}$. Indeed, for a given A, we first partition it into two ordered parts X_1, X_2 with $X_2 \ne \emptyset$ (here X_2 corresponds to $\overline{B \cup C}$). There are $2^{|A|} - 1$ ways to do so. Next we count the number of such pairs (B, C) such that

$B \cap C = X_1$ and $\overline{B \cup C} = X_2$. This number at most 1/2 of the number of ordered partitions of $[n] \setminus A$ into two nonempty parts. The ratio 1/2 comes from the fact that for each ordered bipartition $[n] \setminus A = X_3 \sqcup X_4$, if $(X_3 \cup X_1, X_4 \cup X_1) \in (\mathcal{B} \times \mathcal{C})$, then we cannot also have $(X_4 \cup X_1, X_3 \cup X_1) \in (\mathcal{B} \times \mathcal{C})$, because $\mathcal{B}$ and $\mathcal{C}$ are disjoint. So only half of the ordered bipartitions could actually become desired pairs. Consequently, the number of such pairs (B, C) is $(2^{n-|A|} - 2)/2 = 2^{n-|A|-1} - 1$. The total number (B, C) that give the same A is therefore at most

$$\begin{aligned}(2^{|A|} - 1)(2^{n-|A|-1} - 1) &= 2^{n-1} - 2^{|A|} - 2^{n-|A|-1} + 1 \\ &\le 2^{n-1} - 2^{n/2-\beta\sqrt{n}} - 2^{n-(n/2+\beta\sqrt{n})-1} + 1 \\ &\le \left(1 + \frac{\epsilon}{4}\right) 2^{n-1}.\end{aligned}$$

Here we use the assumption that $\mathcal{A} = \mathcal{A}_\beta$, which implies $|A| \in [n/2 - \beta\sqrt{n}, n/2 + \beta\sqrt{n}]$, and n is large enough. This yields

$$bc \le (2^n - a)\left(1 + \frac{\epsilon}{4}\right) 2^{n-1} + 3^{n+1}$$

where the error term 3^{n+1} arises from the number of pairs $(B, C) \in \mathcal{B} \times \mathcal{C}$ such that either $B \cup C = [n]$, $B \subset C$ or $C \subset B$.

If $(2^n - a)(\epsilon/4)\, 2^{n-1} < 3^{n+1}$, then $bc < (10/\epsilon)3^{n+1}$ and this contradicts $b, c \ge \delta 2^n$. Therefore

$$bc \le (2^n - a)\left(1 + \frac{\epsilon}{4}\right) 2^{n-1} + 3^{n+1} \le (2^n - a)\left(1 + \frac{\epsilon}{2}\right) 2^{n-1}.$$

Consequently, we have

$$a \le 2^n - \frac{bc}{(1 + \epsilon/2)2^{n-1}}.$$

By the same argument used for the proof of the Claim in Case II, we can show that $b \le (1 + \epsilon/2)2^{n-1}$ and $c \le (1 + \epsilon/2)2^{n-1}$. Now we obtain

$$a + b + c \le 2^n - \frac{bc}{(1 + \epsilon/2)2^{n-1}} + b + c = f(b, c) \le \left(\frac{3}{2} + \frac{\epsilon}{4}\right) 2^n$$

where the last inequality follows by maximizing the function $f(b, c)$ subject to the constraints $b, c \in I = [\delta 2^n, (1 + \epsilon/2)2^{n-1}]$. Indeed, setting $\partial_b f = \partial_c f = 0$ we conclude that the extreme points occur at the boundary of $I \times I$. In fact, the maximum is achieved at $(b, c) = ((1+\epsilon/2)2^n, (1+\epsilon/2)2^n)$, and $f((1+\epsilon/2)2^n, (1+\epsilon/2)2^n) = (3/2+\epsilon/4)2^n$ as claimed above. □

4 Concluding remarks

• The definition of sunflower can be generalized as follows: Let $0 \le t \le k$ and $A_i \in \mathcal{A}_i \subset 2^{[n]}$ for $i \in [k]$. Then $(A_i)_{i=1}^k$ is a *t-sunflower* if

- $A_i \cap A_j = C$ for all $i \neq j$, and
- $A_i \setminus C \neq \emptyset$ holds for at least t indices $i \in [k]$.

Note that a $(t+1)$-sunflower is a t-sunflower but the converse need not hold. Let

$$\mathcal{F}(n,k,t) = \{(\mathcal{A}_i)_{i=1}^k : \mathcal{A}_i \subset 2^{[n]} \text{ for } i \in [k] \text{ and } (\mathcal{A}_i)_{i=1}^k \text{ is } t\text{-sunflower-free}\},$$

$$S(n,k,t) = \max_{(\mathcal{A}_i)_{i=1}^k \in \mathcal{F}(n,k,t)} \sum_{i=1}^k |\mathcal{A}_i| \qquad \text{and} \qquad P(n,k,t) = \max_{(\mathcal{A}_i)_{i=1}^k \in \mathcal{F}(n,k,t)} \prod_{i=1}^k |\mathcal{A}_i|.$$

Using the ideas in this paper, one can show that for each $0 \leq t < k$,

$$S(n,k,t) = (k-1)2^k + \sum_{s=0}^{t-2} \binom{n}{s}.$$

The details can be found in [21].

• By the monotonicity of the function $P(n,k,t)$ in t, Theorem 4 implies for each fixed $0 \leq t \leq 3$,

$$P(n,3,t) = \left(\frac{1}{8} + o(1)\right) 2^{3n}.$$

The case $t=0$ is particularly interesting. Let $P^*(n,k) = P(n,k,0)$, $p^*(n,k) = P^*(n,k)/2^{kn}$ and $p(n,k) = P(n,k)/2^{kn}$. As pointed out by a referee, it is easy to show that $p^*(n,k)$ is monotone increasing as a function of n for each fixed $k \geq 3$, while $p(n,k)$ is not. Indeed, given a collection of optimal families $(\mathcal{A}_i)_{i=1}^k$ for $P^*(n,k)$, we can construct k families of subsets of $[n+1]$ that are 0-sunflower-free with the product of their sizes at least $2^k P^*(n,k)$ as follows. We "double" each $\mathcal{A}_i$ in the following way to get new families:

$$\mathcal{B}_i = \mathcal{A}_i \cup \{A \cup \{n+1\} : A \in \mathcal{A}_i\}, \quad i \in [k].$$

Clearly, $\prod_{i=1}^k |\mathcal{B}_i| = \prod_{i=1}^k 2|\mathcal{A}_i| = 2^k P^*(n,k)$ and it is an easy exercise to show that $(\mathcal{B}_i)_{i=1}^k$ contains no 0-sunflower. Since $p^*(n,k) \leq 1$, we conclude that $p^*(k) := \lim_{n\to\infty} p^*(n,k)$ exists. Clearly $p^*(3) = 1/8$, and in general $1/8 \leq p^*(k) \leq (1-1/k)^k < 1/e$. Further, for a fixed $k \geq 4$, if one can show that there exists a single value n_0 such that $p^*(n_0,k) > 1/8$, then by the monotonicity of $p^*(n,k)$ and $P^*(n,k) \leq P(n,k)$, Conjecture 1 would be disproved.

• Our approach for $S(n,k)$ is simply to average over a suitable family of partitions. It can be applied to a variety of other extremal problems, for example, it yields some results about cross intersecting families proved by Borg [5]. It also applies to the situation when the number of colors is more than the size of the forbidden configuration. In particular, the proof of Lemma 5 yields the following more general statement.

Lemma 11. *Given integers $s \geq 1$, $1 \leq t \leq k$ and $0 \leq c \leq s-1$, let n be an integer such that $n \geq c + t(s-c)$. For $i = 1, \dots, k$, let $\mathcal{A}_i \subset \binom{[n]}{s}$ such that $\{\mathcal{A}_i\}_{i=1}^k$ contains no sunflower with t petals and core size c. Then,*

$$\sum_{i=1}^k |\mathcal{A}_i| \leq \begin{cases} \frac{(t-1)k}{m}\binom{n}{s}, & \text{if } c + t(s-c) \leq n \leq c + k(s-c) \\ (t-1)\binom{n}{s}, & \text{if } n \geq c + k(s-c), \end{cases}$$

where $m = \lfloor (n-c)/(s-c) \rfloor$*.*

Note that both upper bounds can be sharp. For the first bound, when $c = 0$, $m = t < k$ and $n = ms$, let each $\mathcal{A}_i$ consist of all s-sets omitting the element 1. A sunflower with $t = m$ petals and core size $c = 0$ is a perfect matching of $[n]$. Since every perfect matching has a set containing 1, there is no sunflower. Clearly $\sum_i |\mathcal{A}_i| = k\binom{n-1}{s} = ((t-1)k/m)\binom{n}{s}$. For the second bound, we can just take $t-1$ copies of $\binom{[n]}{s}$ to achieve equality.

• Another general approach that applies to the sum of the sizes of families was initiated by Keevash–Saks–Sudakov–Verstraëte [17]. We used the idea behind this approach in Lemma 7. Both methods can be used to solve certain problems. For example, as pointed out to us by Benny Sudakov, the approach in [17] can be used to prove the $k = 3$ case of Theorem 3 (and perhaps other cases too).

Acknowledgment. We thank a referee for very helpful comments.

References

[1] N. Alon, A. Shpilka, C. Umans, *On Sunflowers and Matrix Multiplication*, Computational Complexity, 22 (2013), 219–243.

[2] C. Bey, *On Cross-Intersecting Families of Sets*, Graphs and Combinatorics, 21 (2005), 161–168.

[3] B. Bollobás, P. Keevash, B. Sudakov, *Multicolored extremal problems*, Journal of Combinatorial Theory, Series A 107 (2004), 295–312.

[4] P. Borg, *Intersecting and Cross-Intersecting Families of Labeled Sets*, The Electronic Journal of Combinatorics, 15, #N9, 2008.

[5] P. Borg, *The maximum sum and the maximum product of sizes of cross-intersecting families*, European Journal of Combinatorics 35 (2014), 117–130.

[6] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press (2004), 243.

[7] E. Croot, V. Lev, and P. Pach. *Progression-free sets in* $\mathbb{Z}_4^n$ *are exponentially small* (2016). URL:https://arxiv.org/pdf/1605.01506.pdf.

[8] M. Deza, *Une Propriete Extremale des Plans Projectifs Finis Dans une Classe de Codes Equidestants*, Discrete Mathematics, 6 (1973), 343–352.

[9] J. Ellenberg, D. Gijswijt, *Large subsets of* $\mathbb{F}_q^n$ *with no three-term arithmetic progression*, (2016). URL:https://arxiv.org/pdf/1605.09223.pdf.

[10] P. Erdős, C. Ko, R. Rado, *Intersection Theorems for Systems of Finite Sets*, Quarterly Journal of Mathematics: Oxford Journals, 12 (1) (1961), 313–320.

[11] P. Erdős, R. Rado, *Intersection Theorems for Systems of Finite Sets*, Journal of London Mathematical Society, Second Series 35 (1) (1960), 85–90.

[12] P. Erdős, E. Szemerédi, *Combinatorial Properties of Systems of Sets*, Journal of Combinatorial Theory, Series A 24 (3) (1978), 308–313.

[13] P. Frankl, V. Rödl, *Forbidden Intersections*, Transactions of the American Mathematical Society, 300 (1) (1987), 259–286.

[14] P. Frankl, N. Tokushige, *On r-Cross Intersecting Families of Sets*, Combinatorics, Probability and Computing 20 (2011), 749–752.

[15] P. Frankl, S.J. Lee, M. Siggers, N. Tokushige, *An Erdős-Ko-Rado Theorem for Cross t-Intersecting Families*, Journal of Combinatorial Theory, Series A 128 (2014), 207-249.

[16] A.J.W. Hilton, *An Intersection Theorem for a Collection of Families of Subsets of a Finite Set*, Journal of London Mathematical Society, 2 (1977), 369–384.

[17] P. Keevash, M. Saks, B. Sudakov, J. Verstraëte, *Multicolor Turán problems*, Advances in Applied Mathematics 33 (2004), 238–262.

[18] P. Keevash, B. Sudakov, *Set systems with restricted cross-intersections and the minimum rank of inclusion matrices*, Siam. J. Discrete Math. Vol. 18, No. 4 (2005), 713–727.

[19] J.H. van Lint, *A Theorem on Equidistant Codes*, Discrete Mathematics, 6 (1973), 353–358.

[20] M. Matsumoto, N. Tokushige, *The Exact Bound in the Erdős-Rado-Ko Theorem for Cross-Intersecting Families*, Journal of Combinatorial Theory, Series A 52 (1989), 90–97.

[21] D. Mubayi, L. Wang, Multicolor Sunflowers, arXiv.

[22] D. Mubayi, V. Rödl, *Specified Intersections*, Transactions of the American Mathematical Society 366 (2014), No. 1, 491–504.

[23] E. Naslund, W.F. Sawin, *Upper bounds for sunflower-free sets*, (2016). URL: https://arxiv.org/pdf/1606.09575.pdf.

[24] L. Pyber, *A New Generalization of the Erdős-Ko-Rado Theorem*, Journal of Combinatorial Theory, Series A 43 (1986), 85–90.