Sunflowers in Combinatorics

- Let $\mathcal{F}$ be a $k$-uniform family of subsets of $X$, i.e., $|S| = k$ and $S \subseteq X$ for all $S \in \mathcal{F}$.
- $\mathcal{F}$ is a sunflower with $p$ petals if $|\mathcal{F}| = p$ and there exists $Y \subseteq X$ with $Y = S_1 \cap S_2$ for all distinct $S_1, S_2 \in \mathcal{F}$.
- $Y$ is the core and $S_1 \setminus Y$ are the petals.

Note that $p$ disjoint sets forms a sunflower with $p$ petals and empty core.

Research Question: What is the smallest $r = r(p, k)$ such that every $k$-uniform family with at least $r^p$ sets must contain a sunflower with $p$ petals?

Erdős–Rado (1960): Basic Result

(a) $r = pk$ is sufficient to guarantee a sunflower:

- every family with more than $(pk)^2 > k(p-1)^k$ sets contains a sunflower

(b) $r > p - 1$ is necessary to guarantee a sunflower:

- there is a family of $(p - 1)^k$ sets without sunflower.

Erdős conjectured $r = r(p)$ is sufficient (no $k$ dependency).

Until 2019, best known upper bound on $r$ was still $k^{1-o(1)}$ with respect to $k$.

“The sunflower problem has fascinated me greatly – I really do not see why this question is so difficult.”

–Paul Erdős (1981)

Recent Developing Developments


$r = \Theta((\log p)^{1+o(1)})$ is sufficient to guarantee a sunflower

New papers built off their ideas:

- Sep 2019: Rao used Shannon’s coding theorem for a cleaner proof and slightly better bound
- Oct 2019: Frankston–Kahn–Narayanan–Park improved a key lemma of ALWZ, enabling them to prove a long-standing conjecture of Talagrand regarding thresholds functions
- Jan 2020: Rao improved $r = O(\max(\log p, \log k))$ by incorporating ideas from FKNP

Our Results

- Proved that $r = p\max\{\log p, \log k\}$ is sufficient to guarantee a sunflower, without using Shannon’s noiseless coding theorem
- Showed that best known bound cannot be improved without change of proof strategy: proved existence of a $p\max\{\log p, \log k\}$-spread family where main lemma fails
- Clarified barriers of current approaches: explicitly formulated inequalities that $r$ must satisfy for the proof technique to work

Strategy: Proving that a Sunflower Exists

- Key Definition: $\mathcal{F}$ is $r$-spread if $|\mathcal{F}| \geq r^p$ and for every nonempty $S \subseteq X$, the number of sets in $\mathcal{F}$ which contain $S$ is at most $r^{p-|S|}$.

Inductive Reduction to $r$-spread family

If every $r$-spread family contains $p$ disjoint sets, then $r^p$ sets is sufficient to guarantee a sunflower.

Proof: Induction on $k$.

The Probabilistic Method: find $p$ disjoint sets

- Consider a random partition of $X$ into $X_1, X_2, \ldots, X_p$ with equal size.
- Use probabilistic method to prove that there is a partition where there exists a set $S_i \in \mathcal{F}$ such that $S_i \subseteq X_i$, for all $i$.
- Then $S_1, \ldots, S_p$ are disjoint sets in $\mathcal{F}$.
- Main Lemma gives conditions for this step to work.

- We only need to work with one petal rather than $p$ petals.

Main Lemma

$\mathbb{P}(\text{There does not exist } S_i \in \mathcal{F} \text{ such that } S_i \subseteq X_i) < \frac{1}{p}$

Proof: Partition $X$ to $V_1, V_2$ with equal size, so $|V_1| = |V_2| = |X|/(2p)$.

- Key Definition: Given $S \in \mathcal{F}$ and $W \subseteq X$, $(S, W)$ is $m$-good if there exists $S' \in \mathcal{F}$ such that $S' \subseteq W \cup S$ and $|S' \setminus W| \leq m$.

Iteration: $\mathbb{P}(\text{Less than half of sets in } \mathcal{F} \text{ are } m\text{-good with respect to } V_1) \leq \frac{1}{p}$

- Partition $V_i$ to $W_1, W_2, \ldots, W_i$ with equal size
- Iteratively replace each good $(S, \cup_{i=1}^{k} W_i)$ pair with the guaranteed $S'$
- Bound the number of bad pairs by a key counting lemma & Markov’s inequality
- Moving from $S$ to $S'$ reduces the set sizes at each step as $\cup_{i=1}^{k} W_i$ expands

Final Step: $\mathbb{P}(V_1 \cup V_2 \text{ does not contain a set in } \mathcal{F}) < \frac{1}{p}$

- Construct an $m$-uniform, $\mathcal{F}'$ from sets in $\mathcal{F}$ which are $m$-good with respect to $V_1$
- Apply Janson’s Inequality with $V_1$ and $\mathcal{F}'$ to bound $\mathbb{P}(\exists S \in \mathcal{F}' \text{ s.t. } S \subseteq V_2)$

Applications

Sunflowers have many uses in computer science:

- Fast algorithms for matrix multiplication
- Cryptography
- Pseudorandomness
- Lower bounds on circuitry
- Data structure efficiency
- Random approximations

References

- Erdős (1981). On the combinatorial problem which I would most like to see solved. Combinatorica.

We thank Lutz Warnke (Georgia Tech) for mentorship and NSF Award #1851843 for funding.