Improved Bounds for the Sunflower Lemma[1]

Aditya Lonkar Aditya Subramanian

Indian Institute of Science, Bangalore

E0 224 Project Presentation
December 25, 2020

[1] Alweiss et al. 2020.

Table of Contents

Sunflowers

Definition 1 (Sunflower)

A collection of sets $S_1, \ldots, S_r$ is an r-sunflower if

$$S_i \cap S_j = S_1 \cap \cdots \cap S_r, \qquad \forall i \neq j.$$

We call $K = S_1 \cap \cdots \cap S_r$ the kernel and $S_1 \setminus K, \ldots, S_r \setminus K$ the petals of the sunflower.

Sunflowers

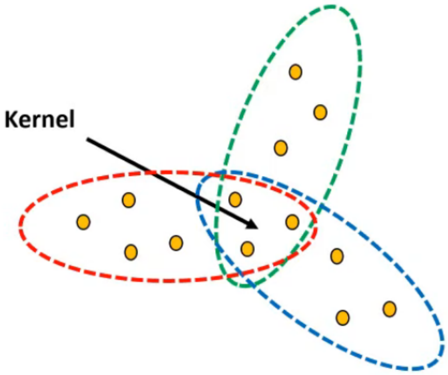

Figure: A 3-sunflower[2]

[2] Lovett 2020.

Sunflower Lemma

Erdős and Rado proved that large enough set systems must contain a sunflower.

Lemma 1 (Sunflower lemma[3])

Let $r \geq 3$ and $\mathcal{F}$ be a w-set system of size $|\mathcal{F}| \geq w! \cdot (r-1)^w$. Then $\mathcal{F}$ contains an r-sunflower.

They conjectured in the same paper that the bound in Lemma 1 can be drastically improved.

Conjecture 1 (Sunflower conjecture)

Let $r \geq 3$. There exists $c = c(r)$ such that any w-set system $\mathcal{F}$ of size $|\mathcal{F}| \geq c^w$ contains an r-sunflower.

[3] Erdős and Rado 1960.

Progress over the Years

The bound in Lemma 1 is of the form $w^{w(1+o(1))}$ where the $o(1)$ depends on r. Despite nearly 60 years of research, the best known bounds were still of the form $w^{w(1+o(1))}$, even for $r = 3$.

- Kostochka[4] proved that any w-set system of size $|\mathcal{F}| \geq cw! \cdot (\log\log\log w / \log\log w)^w$ must contain a 3-sunflower for some absolute constant c.
- Fukuyama[5] claimed an improved bound of $w^{(3/4+o(1))w}$ for $r = 3$, but this proof has not been verified.

[4] Kostochka 1997.

[5] Fukuyama 2018.

Main Result

In 2019 Ryan Alweiss, Shachar Lovett, Kewen Wu and Jiapeng Zhang vastly improved this bound. They proved that any w-set system of size $(\log w)^{w(1+o(1))}$ must contain a sunflower. More precisely:

Theorem 2 (Main theorem[6])

Let $r \geq 3$. For some constant C, any w-set system $\mathcal{F}$ of size $|\mathcal{F}| \geq (Cr^3 \log w \log\log w)^w$ contains an r-sunflower.

[6] Alweiss et al. 2020.

Follow up Works

The bound they obtained was of the form $(r \cdot \log w)^w (\log\log q)^{O(w)}$

- this was improved to $(cr \cdot \log w)^w$ by Frankston, Kahn, Narayanan, Park[7].
- a simplified proof using information theory was given by Anup Rao[8].
- An alternate proof using information theory was given by Terry Tao[9].

[7] Frankston et al. 2019.
[8] Rao 2020.
[9] Tao 2020.

Table of Contents

Notation

In this section we shall introduce some structures and their generalizations that are used in the proof.

- Link of a set system.
- Spreadness of a set system (& Weighted spread system).
- satisfying set systems.
- sunflowers $\rightarrow$ robust sunflowers.

Link

Definition 1

Given a set system $\mathcal{F}$ on X and a set $T \subset X$, the link of $\mathcal{F}$ at T is

$$\mathcal{F}_T = \{S \setminus T : S \in \mathcal{F}, T \subset S\}.$$

Example:

$$\mathcal{F} = \{\{1,2\},\{2,3,4\},\{1,3,5\},\{2,3,6\},\{3,5,7\}\}$$

$$\mathcal{F}_{\{1\}} = \{\{2\},\{3,5\}\}$$

$$\mathcal{F}_{\{2,3\}} = \{\{4\},\{6\}\}$$

Link

Definition 1

Given a set system $\mathcal{F}$ on X and a set $T \subset X$, the link of $\mathcal{F}$ at T is

$$\mathcal{F}_T = \{S \setminus T : S \in \mathcal{F}, T \subset S\}.$$

Observation: If $\mathcal{F}_T$ contains a r-sunflower, then $\mathcal{F}$ contains a r-sunflower.

$$S_1 \setminus T, S_2 \setminus T, S_3 \setminus T \text{ sunflower} \implies S_1, S_2, S_3 \text{ sunflower}$$

Spread Set Systems

Definition 2 (Spreadness,[10])

We say that a w-set system is κ-spread if $|\mathcal{F}| \geq \kappa^w$ and $|\mathcal{F}_T| \leq \kappa^{-|T|}|\mathcal{F}|$ for all non-empty T, where $|\mathcal{F}_T|$ is the size of the link at T.

Note that, if $\mathcal{F}$ is a w-set system of size $|\mathcal{F}| \geq \kappa^w$ on a ground set X. Then either $\mathcal{F}$ is κ-spread, or there is a link $\mathcal{F}_T$ of size $|\mathcal{F}_T| \geq \kappa^{w-|T|}$. Going forward we shall refer to the former as the "pseudorandom case" and the latter as the "structured case".

[10] Lovett, Solomon, and Zhang 2019.

Spread Weighted Set Systems

Definition 3 (Weighted set system)

Let $\mathcal{F}$ be a set system, and let $\sigma : \mathcal{F} \mapsto \mathbb{Q}_{\geq 0}$ be a weight function that assigns nonnegative rational weights to sets in $\mathcal{F}$ which are not all 0. We call the pair $(\mathcal{F}, \sigma)$ a weighted set system.

For a subset $\mathcal{F}' \subset \mathcal{F}$ we write $\sigma(\mathcal{F}') = \sum_{S \in \mathcal{F}'} \sigma(S)$ for the sum of the weights of the sets in $\mathcal{F}'$.

Definition 4 (Spread weighted set system)

Let $\mathsf{s} = (s_0; s_1, \ldots, s_w)$ be a weight profile. A weighted set system $(\mathcal{F}, \sigma)$ is s-spread if

1. $\sigma(\mathcal{F}) \geq s_0$;
2. $\sigma(\mathcal{F}_T) \leq s_{|T|}$ *for any link* $\mathcal{F}_T$ *with non-empty* T.

In particular, $\mathcal{F}$ is a w-set system.

Satisfying Set Systems

Now, we define a more "robust" version of the property of having disjoint sets. Given a finite set X, we denote by $\mathcal{U}(X, p)$ the distribution over subsets $R \subset X$, where each element $x \in X$ is included in R independently with probability p (i.e. p-biased inputs).

Definition 5 (Satisfying set system)

Let $0 < \alpha, \beta < 1$. A set system $\mathcal{F}$ on X is (α, β)-satisfying if

$$\Pr_{R \sim \mathcal{U}(X,\alpha)} [\exists S \in \mathcal{F}, S \subset R] > 1 - \beta.$$

Satisfying Set Systems

The explanation for the name "satisfying" is that if the set system is interpreted as a disjunctive normal form (DNF) formula, then this condition is that the formula has more than a $1 - \beta$ probability of being satisfied on α-biased inputs.

Definition 5 (Satisfying set system)

Let $0 < \alpha, \beta < 1$. A set system $\mathcal{F}$ on X is (α, β)-satisfying if

$$\Pr_{R \sim \mathcal{U}(X,\alpha)}[\exists S \in \mathcal{F}, S \subset R] > 1 - \beta.$$

Definition 6 (Satisfying weight profile)

Let $0 < \alpha, \beta < 1$. A weight profile s *is (α, β)-satisfying if any* s*-spread set system is (α, β)-satisfying.*

Satisfying Set Systems

We mentioned that this was a more “robust” version of the property of having disjoint sets.

Lemma 2 ([11])

If $\mathcal{F}$ is a $(1/r, 1/r)$-satisfying set system and $\emptyset \notin \mathcal{F}$, then $\mathcal{F}$ contains r pairwise disjoint sets.

Let $\mathcal{F}$ be a set system on X. Consider a random coloring of elements of X with r colors, so that each $x \in X$ has an equal probability of being colored with any of the r colors, independently of the other elements. For $1 \leq i \leq r$, let Y_i denote the subset of X colored with color i and let $\mathcal{E}_i$ denote the event that $\mathcal{F}$ contains an i-monochromatic set, namely,

$$\mathcal{E}_i = [\exists S \in \mathcal{F}, S \subset Y_i] .$$

Satisfying Set Systems (continued)

Note that $Y_i \sim \mathcal{U}(X, 1/r)$, and since we assume $\mathcal{F}$ is $(1/r, 1/r)$-satisfying, we have

$$\Pr[\mathcal{E}_i] > 1 - 1/r.$$

By the union bound, with positive probability all of $\mathcal{E}_1, \ldots, \mathcal{E}_r$ hold. In this case, $\mathcal{F}$ contains a set which is i-monochromatic for each $i = 1, \ldots, r$. Such sets must be pairwise disjoint.

[11] Lovett, Solomon, and Zhang 2019.

Robust Sunflower

Definition 7 (Robust sunflower)

Let $0 < \alpha, \beta < 1$, $\mathcal{F}$ be a set system, and let $K = \bigcap_{S \in \mathcal{F}} S$ be the common intersection of all sets in $\mathcal{F}$. $\mathcal{F}$ is an (α, β)-robust sunflower if (i) $K \notin \mathcal{F}$, and (ii) $\mathcal{F}_K$ is (α, β)-satisfying. We call K the kernel.

Robust sunflowers are a generalization of satisfying set systems. In particular, an (α, β)-satisfying w-uniform system $\mathcal{F}$ will be an (α, β)-robust sunflower as long as $\mathcal{F}$ contains at least two distinct sets, because if $R \supset S$ for some $S \in F$, then clearly $R \supset S \setminus K$.

Lemma 3 ([12])

Any $(1/r, 1/r)$-robust sunflower contains an r-sunflower.

[12] Lovett, Solomon, and Zhang 2019.

Table of Contents

Main Theorem

The proof of 2 follows from the following stronger theorem, by setting $\alpha = \beta = 1/r$ and applying Lemma 3.

Theorem 3 (Main theorem, robust sunflowers)

Let $0 < \alpha, \beta < 1$. For some constant C, any w-uniform set system $\mathcal{F}$ of size $|\mathcal{F}| \geq \left(C\alpha^{-2} \cdot \left(\log w \log\log w + \left(\log \frac{1}{\beta} \right)^2 \right) \right)^w$ *contains an (α, β)-robust sunflower.*

Table of Contents

Structured Case

In the structured case, the set family has links of "large size". Formally let $\mathcal{F}$ be a w-set system of size $|\mathcal{F}| \geq \kappa^w$ on a ground set X. We shall show that it has a r-sunflower by induction on w.

1. **Base Case:** when $w = 1$, the set system has singleton sets. For a sufficiently large κ, these form a r-sunflower with an empty kernel.
2. **Hypothesis:** for $i < w$, let every i-set system of size $\geq \kappa^i$ have an r-sunflower.
3. **Induction:** consider a w-set system $\mathcal{F}$. In the structured case, there is a link $\mathcal{F}_T$ of size $|\mathcal{F}_T| \geq \kappa^{w-|T|}$ (if not, we are in the pseudorandom case). Now, the $(w - |T|)$-set system, $\mathcal{F}_T$ has an r-sunflower by the induction hypothesis. But we know that if $\mathcal{F}_T$ contains an r-sunflower, then $\mathcal{F}$ contains an r-sunflower. Hence we are done. $\square$

Table of Contents

Encoding Argument

Let $\mathrm{s} = (s_0; s_1, \ldots, s_w)$ be a weight profile, and $w' \leq w$. In the pseudorandom case we consider w-set systems which are s-spread. Let $W \subset X$. Given a set $S \in \mathcal{F}$, the pair (W, S) is said to be *good* if there exists a set $S' \in \mathcal{F}$ (possibly with $S' = S$) such that

1. $S' \setminus W \subset S \setminus W$;
2. $|S' \setminus W| \leq w'$.

If no such S' exists, we say that (W, S) is *bad*. Note that if W contains a set in $\mathcal{F}$ (i.e., $S' \subset W$ for some $S' \in \mathcal{F}$) then all pairs (W, S) are good.

Lemma 4

Let $(\mathcal{F}, \sigma)$ be an $\mathrm{s} = (s_0; s_1, \ldots, s_w)$-spread weighted w-set system on X. Let $W \subset X$ be a uniform subset of size $|W| = p|X|$ and $\mathcal{B}(W) = \{S \in \mathcal{F} : (W, S) \text{ is bad}\}$. Then $\mathbb{E}_W[\sigma(\mathcal{B}(W))] \leq (4/p)^w s_{w'}$.

Encoding Argument

Proof. First, we simplify the setting a bit. We may assume by scaling σ and s by the same factor that $\sigma(S) = N_S$ is an integer for each $S \in \mathcal{F}$. Let $N = \sum N_S \geq s_0$. By scaling back down by N, we can then identify $(\mathcal{F}, \sigma)$ with the uniform distribution on the multi-set system $\mathcal{F}' = \{S_1, \ldots, S_N\}$, where each set $S \in \mathcal{F}$ is repeated N_S times in $\mathcal{F}'$. Thus

$$\sigma(\mathcal{B}(W)) = |\{i : S_i \in \mathcal{F}' \text{ and } (W, S_i) \text{ is bad}\}|.$$

Assume that (W, S_i) is bad in $\mathcal{F}'$. In particular, this means that W does not contain any set in $\mathcal{F}$. We describe (W, S_i) with a small amount of information. Let $|X| = n$ and $|W| = pn$. We encode (W, S_i) as follows:

Encoding Argument (continued)

1. The first piece of information is $W \cup S_i$, a subset of X of size between pn and $pn + w$. The number of options for this is at most:
$$\sum_{i=0}^{w} \binom{n}{pn+i} \le \sum_{i=0}^{w} \left(\frac{1-p}{p}\right)^i \binom{n}{pn}$$
$$\le \left(\frac{1-p}{p} + 1\right)^w \binom{n}{pn} = p^{-w} \binom{n}{pn}.$$
2. Given $W \cup S_i$, let j be minimal such that $S_j \subset W \cup S_i$, so that j depends only on $W \cup S_i$. Given S_j, there are fewer than 2^w possibilities for $A = S_i \cap S_j$. As such, we will let A be the second piece of information.

Encoding Argument (continued)

3. Because $S_j \subset W \cup S_i$ and $A = S_i \cap S_j$, we have $S_j \setminus W \subset A$. Since (W, S_i) is bad, $|A| \geq |S_j \setminus W| > w'$. The number of the sets in $\mathcal{F}'$ which contain A is $|\mathcal{F}'_A| \leq s_{w'}$. The third piece of information will be which one of these is S_i.
4. Finally, once we have specified S_i, we will specify $S_i \cap W$, which is one of 2^w possible subsets of S_i.

From these four pieces of information one can uniquely reconstruct (W, S_i). Thus the total number of bad pairs (W, S_i) is bounded by

$$p^{-w}\binom{n}{pn} \cdot 2^w \cdot s_{w'} \cdot 2^w = (4/p)^w s_{w'}\binom{n}{pn}.$$

Because the number of sets $W \subset X$ of size $|W| = p|X|$ is $\binom{n}{pn}$, the lemma follows by taking expectation over W.

A Reduction Step

The following lemma is about reducing an (α, β)-satisfying set system $\mathcal{F}$ to a smaller set system $\mathcal{F}'$ which is (α', β')-satisfying.

Lemma 5 (Reduction Lemma[13])

Let $\mathrm{s} = (s_0; s_1, \ldots, s_w)$ *be a weight profile,* $w' \leq w$*,* $\delta > 0$ *and define* $\mathrm{s}' = ((1-\delta)s_0; s_1, \ldots, s_{w'})$*. Assume* s' *is* (α', β')*-satisfying. Then for any* $p > 0$*,* s *is* (α, β)*-satisfying for*

$$\alpha = p + (1-p)\alpha', \quad \beta = \beta' + \frac{(4/p)^w s_{w'}}{\delta s_0}.$$

[13] Alweiss et al. 2020.

A Reduction Step(Proof[14])

Let $(\mathcal{F}, \sigma)$ be an $\mathrm{s} = (s_0; s_1, \ldots, s_w)$-spread weighted set system on X. Let $W \sim \mathcal{U}(X, p)$. Say that W is δ-bad if $\sigma(\mathcal{B}(W)) \geq \delta s_0$. By applying the Corollary and Markov's inequality, we obtain that

$$\Pr[W \text{ is } \delta\text{-bad}] \leq \frac{\mathbb{E}[\sigma(\mathcal{B}(W))]}{\delta s_0} \leq \frac{(4/p)^w s_{w'}}{\delta s_0}.$$

Fix W which is not δ-bad. If for some $S \in \mathcal{F}$, the pair (W, S) is good, then there exists $\pi(S) = S' \in \mathcal{F}$ (possibly with $S' = S$) such that (i) $S' \setminus W \subset S \setminus W$ and (ii) $|S' \setminus W| \leq w'$.
Define a new weighted set system $(\mathcal{F}', \sigma')$ on $X' = X \setminus W$ as follows:

$$\mathcal{F}' = \{\pi(S) \setminus W : S \in \mathcal{F} \setminus \mathcal{B}(W)\}, \quad \sigma'(S' \setminus W) = \sigma(\pi^{-1}(S')).$$

[14] Alweiss et al. 2020.

A Reduction Step(Proof[15])

The claim is that $\mathcal{F}'$ is $\mathrm{s}' = ((1-\delta)s_0; s_1, \ldots, s_{w'})$-spread. To see that, note that $\sigma'(\mathcal{F}') = \sigma(\mathcal{F} \setminus \mathcal{B}(W)) \geq (1-\delta)s_0$ and that for any set $T \subset X'$,

$$\sigma'(\mathcal{F}'_T) = \sum_{S' \supset T} \sigma'(S') = \sum_{S : \pi(S) \supset T} \sigma(S) \leq \sum_{S \supset T} \sigma(S) = \sigma(\mathcal{F}_T) \leq s_{|T|}.$$

Finally, all sets in $\mathcal{F}'$ have size at most w'. Thus, if we choose $W' \sim \mathcal{U}(X', \alpha')$ then we obtain that with probability more than $1 - \beta'$, there exists $S^* \in \mathcal{F}'$ such that $S^* \subset W'$. Recall that $S^* = S \setminus W$ for some $S \in \mathcal{F}$. Thus $S \subset W \cup W'$, which is distributed according to $\mathcal{U}(X, p + (1-p)\alpha')$. The value of β' accounts for the case where W is δ-bad.

[15] Alweiss et al. 2020.

Janson's Inequality[16]

- Janson's Inequality falls under "The Poisson Paradigm". The idea is when X is the sum of "mostly independent" indicator random variables with $\mathbb{E}[X] = \mu$, then X is close to a Poisson distribution with mean μ. Hence, $Pr[X = 0] \sim e^{-\mu}$.
- Basically, we have a set of bad events B_i and we would like to bound the probability that none of these bad events occur.
- We would make use of Janson's inequality to show that a randomly chosen set from a p-biased distribution W is such that for some $S_i \in \mathcal{F}$, $S_i \subset W$ with high probability.

[16] Alon and Spencer 2004.

Janson's Inequality[17](Continued)

To explain the setting for Janson's Inequality,
Ω is a finite universal set and R is a random set of Ω given by

$$Pr[r \in R] = p_r$$

and these events are mutually independent over $r \in \Omega$. A_i, $i \in I$ be subsets of Ω and I is a finite index set. B_i is the event that $A_i \subset R$(That is, each point $r \in R$ "flips a coin" to determine if it is in R and B_i is is the event that the coins for all $r \in A_i$ came up "heads"). Define $\Delta = \sum_{i \neq j} Pr[B_i \wedge B_j]$. By Janson's Inequality,

$$Pr[\wedge_{i \in I} \overline{B_i}] \leq e^{-\frac{\mu^2}{2\delta}}$$

when $\Delta \geq \mu$.

[17] Alon and Spencer 2004.

A Final Step

Repeated applications of the reduction step yield a set system which is almost as spread as the original set system, but whose sets are much smaller. Thus the spreadness guarantee becomes good compared to the size of the sets in the system.

Lemma 6 (Satisfiability Lemma[18])

Let $0 < \alpha, \beta < 1$, $w \geq 2$, and set $\kappa = \max\{4\log(1/\beta), 2\} \cdot w/\alpha$. Let $(\mathcal{F}, \sigma)$ be an $\mathrm{s} = (s_0; s_1, \ldots, s_w)$-spread weighted set system where $s_i < \kappa^{-i} s_0$. Then $\mathcal{F}$ is (α, β)-satisfying.

[18] Alweiss et al. 2020.

A Final Step(Continued)

Proof: The proof proceeds by converting $(\mathcal{F}, \sigma)$ to an equivalent uniformly weighted set system $(\mathcal{F}', \sigma')$. Then we consider $W \sim \mathcal{U}(X, \alpha)$ and define indicator random variables Z_i corresponding to sets $S_i \in \mathcal{F}$ such that $Z_i = 1$ when $S_i \subset W$. Then defining μ and Δ as the mean and the covariance parts resp.,

$$\Delta = \sum_{\ell=1}^{\ell=w} N^2 p_\ell \alpha^{2w-\ell}$$

where p_ℓ is the fraction of the N^2 pairs (i, j) such that $|S_i \cap S_j| = \ell$. The crucial part is because of the spread property, p_ℓ and subsequently Δ can be upper bounded. Finally we can apply Janson's inequality for $Pr[\forall i \ Z_i = 0]$.

Putting it all together

Now, all of these steps are clubbed together so as to finally obtain a bound on $\kappa(w, \alpha, \beta)$.

1. Apply Lemma 5 iteratively on the weighted set system $(\mathcal{F}, \sigma)$ until we reach a sufficiently small weighted set system $(\mathcal{F}', \sigma)$.
2. Finally, Apply Lemma 6 on this reduced set system $(\mathcal{F}', \sigma)$.

These two steps are applied in an optimal manner to get the best bound on $\kappa(w, \alpha, \beta)$.

Optimizing for $\kappa(w, \alpha, \beta)$

Let $w \geq 2$ be fixed throughout, and let $\kappa > 1$ be optimized later. We first introduce some notation. For $0 < \Delta < 1$ and $\ell \geq 1$, let $\mathsf{s}(\Delta, \ell) = (1 - \Delta; \kappa^{-1}, \ldots, \kappa^{-\ell})$ be a weight profile. Let $A(\Delta, \ell), B(\Delta, \ell)$ be some bounds such that any $\mathsf{s}(\Delta, \ell)$-spread set system is $(A(\Delta, \ell), B(\Delta, \ell))$-satisfying.
Lemma 5 applied to $w' \geq w''$ and p, δ shows that we may take

$$A(\Delta, w') \leq A(\Delta + \delta, w'') + p,$$

$$B(\Delta, w') \leq B(\Delta + \delta, w'') + \frac{(4/p)^{w'}}{\delta(1 - \Delta)\kappa^{w''}}.$$

We apply this iteratively for some widths $w_0, \ldots, w_r$. Set $w_0 = w$ and $w_{i+1} = \lceil (1 - \varepsilon) w_i \rceil$ for some small ε as long as $w_i > w^*$ for some w^*. In particular, we need $w^* \geq 1/\varepsilon$ to ensure $w_{i+1} < w_i$, and we will optimize ε, w^* later.

Optimizing for $\kappa(w, \alpha, \beta)$ (continued)

The number of steps is thus $r \leq (K \log w)/\varepsilon$ for some constant $K > 0$. Let $p_1, \ldots, p_r$ and $\delta_1, \ldots, \delta_r$ be the values we use for p, δ at each iteration. To simplify the notation, let $\Delta_i = \delta_1 + \cdots + \delta_i$ and $\Delta_0 = 0$. Furthermore, define

$$\gamma_i = \frac{(4/p_i)^{w_{i-1}}}{\kappa^{w_i}}.$$

Then for $i = 1, \ldots, r$, we may take

$$A(\Delta_{i-1}, w_{i-1}) \leq A(\Delta_i, w_i) + p_i,$$
$$B(\Delta_{i-1}, w_{i-1}) \leq B(\Delta_i, w_i) + \frac{\gamma_i}{\delta_i(1 - \Delta_{i-1})}.$$

Optimizing for $\kappa(w, \alpha, \beta)$ (continued)

Set $p_i = p = \frac{\alpha}{2r}$ and $\delta_i = \sqrt{\gamma_i}$, where $r \leq (K \log w)/\varepsilon$ is the number of steps. We will select the parameters so that $\Delta_i \leq 1/2$ for all i. Thus we may take $A(1/2, w^*)$ and $B(1/2, w^*)$ such that

$$\begin{aligned} A(0, w) &\leq A(\Delta_r, w_r) + \alpha/2 \leq A(1/2, w^*) + \alpha/2, \\ B(0, w) &\leq B(\Delta_r, w_r) + 2\Delta_r \leq B(1/2, w^*) + 2\Delta_r. \end{aligned}$$

Plugging in the values for δ_i, we compute the sum

$$\begin{aligned} \Delta_r &= \sum_{i=1}^{r} \delta_i \leq \sum_{i=1}^{r} \sqrt{\frac{(4/p)^{w_{i-1}}}{\kappa^{(1-\varepsilon)w_{i-1}}}} = \sum_{1 \leq i \leq r} \left(\frac{4/p}{\kappa^{1-\varepsilon}} \right)^{\frac{w_{i-1}}{2}} \leq \sum_{k \geq w^*} \left(\frac{8K \log w}{\varepsilon \alpha \kappa^{1-\varepsilon}} \right)^{\frac{k}{2}} \\ &= \left(\frac{8K \log w}{\varepsilon \alpha \kappa^{1-\varepsilon}} \right)^{w^*/2} \sum_{k \geq 0} \left(\frac{8K \log w}{\varepsilon \alpha \kappa^{1-\varepsilon}} \right)^{k/2} \leq 2^{1-w^*}, \end{aligned}$$

Optimizing for $\kappa(w, \alpha, \beta)$ (continued)

assuming $\kappa^{1-\varepsilon} \geq (32K \log w)/\varepsilon\alpha$. Apply Lemma 6 to bound $A(1/2, w^*) \leq \alpha/2$ and $B(1/2, w^*) \leq \beta/2$. We use the simple observation that $(1/2; \kappa^{-1}, \ldots, \kappa^{-w^*})$-spread set systems are also $(1; (\kappa/2)^{-1}, \ldots, (\kappa/2)^{-w^*})$-spread, in which case we can apply Lemma 6

$$\kappa/2 \geq (2 + 4\log(2/\beta)) \cdot 2w^*/\alpha.$$

We still have the freedom to choose $\varepsilon > 0$ and $w^* \geq 1/\varepsilon$. To obtain $A(0, w) \leq \alpha, B(0, w) \leq \beta$, we also need $\Delta_r \leq \beta/4 < 1/2$. Thus all the constraints are:

1. $w^* \geq 1/\varepsilon$;
2. $\kappa^{1-\varepsilon} \geq (32K \log w)/\varepsilon\alpha$ where $K > 0$ is a constant;
3. $\kappa \geq 4 \cdot (2 + 4\log(2/\beta)) \cdot w^*/\alpha$;
4. $2^{1-w^*} \leq \beta/4$.

Optimizing for $\kappa(w, \alpha, \beta)$ (continued)

Set $\varepsilon = 1/\log\log w$ and $w^* = c \cdot \max\{\log(1/\beta), \log\log w\}$ for some large enough constant $c \geq 1$. This ensures that the first and last condition hold.

Thus we obtain that the result holds whenever

$$\kappa = \Omega\Bigg(\max\Bigg\{ \left(\frac{1}{\alpha}\right)^{1+2/\log\log w} \log w \log\log w, \frac{1}{\alpha}\left(1 + \log\frac{1}{\beta}\right)^2, \frac{1}{\alpha}\left(1+\log\frac{1}{\beta}\right)\log\log w \Bigg\}\Bigg)$$

In particular, it suffices to set

$\kappa = O\left(\alpha^{-2} \cdot \left(\log w \log\log w + \left(\log\frac{1}{\beta}\right)^2\right)\right)$.

Summary

- $|\mathcal{F}| \sim \kappa^w$ where $\kappa \sim O(\log w)$
- **Structured case:** $\mathcal{F}$ has large link and hence there exists set T such that $|\mathcal{F}_T| \geq |\mathcal{F}|/\kappa^{|T|}$.
- Apply induction to show that $\mathcal{F}_T$ contains a r-sunflower.
- **Pseudorandom case:** All links are "small" and hence, $\mathcal{F}$ is κ-spread.
- Lemma: $\mathcal{F}$ is $(1/r, 1/r)$ satisfying.
- Corollary: $\mathcal{F}$ has a r-sunflower.

Summary(continued)

In particular, restating the results:

Lemma 7

If $\mathcal{F}$ is a $(1/r, 1/r)$-satisfying set system and $\emptyset \notin \mathcal{F}$, then $\mathcal{F}$ contains r pairwise disjoint sets.

Lemma 8

Any $(1/r, 1/r)$-robust sunflower contains an r-sunflower.

Lemma 9

Let $0 < \alpha, \beta < 1$ and $w \geq 2$. Let $\kappa = \kappa(w) > 1$ be a nonstrictly increasing function of w such that the weight profile $(1; \kappa^{-1}, \ldots, \kappa^{-w})$ is (α, β)-satisfying. Then any w-uniform set system $\mathcal{F}$ of size $|\mathcal{F}| > \kappa^w$ must contain an (α, β)-robust sunflower.

Summary(continued)

Theorem 4

Let $0 < \alpha, \beta < 1$. For some constant C, any w-uniform set system $\mathcal{F}$ of size $|\mathcal{F}| \geq \left(C\alpha^{-2} \cdot \left(\log w \log \log w + \left(\log \frac{1}{\beta} \right)^2 \right) \right)^w$ contains an (α, β)-robust sunflower.

Theorem 5

$\kappa(w, \alpha, \beta) = O\left(\alpha^{-2} \cdot \left(\log w \log \log w + \left(\log \frac{1}{\beta} \right)^2 \right) \right)$. *Here $\kappa(w, \alpha, \beta)$ is the least κ such that $(1; \kappa^{-1}, \ldots, \kappa^{-w})$ is (α, β)-satisfying*

References

Alon, Noga and Joel H Spencer (2004). *The probabilistic method*. John Wiley & Sons.

Alweiss, Ryan et al. (2020). "Improved Bounds for the Sunflower Lemma". In: *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2020. Chicago, IL, USA: Association for Computing Machinery, pp. 624–630. ISBN: 9781450369794. DOI: 10.1145/3357713.3384234. URL: https://doi.org/10.1145/3357713.3384234.

Erdős, Paul and Richard Rado (1960). "Intersection theorems for systems of sets". In: *Journal of the London Mathematical Society* 35.1, pp. 85–90.

Frankston, Keith et al. (2019). "Thresholds Versus Fractional Expectation-Thresholds". In: *arXiv preprint arXiv:1910.13433*.

References (continued)

Fukuyama, Junichiro (2018). "Improved Bound on Sets Including No Sunflower with Three Petals". In: *arXiv preprint arXiv:1809.10318*.

Kostochka, AV (1997). "A bound of the cardinality of families not containing Δ-systems". In: *The Mathematics of Paul Erdős II*. Springer, pp. 229–235.

Lovett, Shachar (2020). *Simon's Institute Talk*. URL: https://www.youtube.com/watch?v=cnH_-rwQzls.

Lovett, Shachar, Noam Solomon, and Jiapeng Zhang (2019). "From DNF Compression to Sunflower Theorems via Regularity". In: *34th Computational Complexity Conference, CCC 2019, July 18-20, 2019, New Brunswick, NJ, USA.* 5:1–5:14. DOI: 10.4230/LIPIcs.CCC.2019.5. URL: https://doi.org/10.4230/LIPIcs.CCC.2019.5.

References (continued)

Rao, Anup (2020). "Coding for Sunflowers". In: *Discrete Analysis*. URL: http://arxiv.org/abs/1909.04774.

Tao, Terence (2020). *The sunflower lemma via Shannon entropy*. URL: https://terrytao.wordpress.com/2020/07/20/the-sunflower-lemma-via-shannon-entropy/.

Thank You!