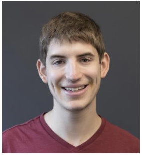
Improved bounds for the sunflower lemma

Kewen Wu
Peking University

Joint work with

Ryan Alweiss
Princeton



Shachar Lovett
UCSD

Jiapeng Zhang
Harvard

Definitions

Definition (w-set system and r-sunflower)

A w-set system is a family of sets of size at most w.
An r-sunflower is r sets $S_1, \ldots, S_r$ where

- **Kernel**: $Y = S_1 \cap \cdots \cap S_r$;
- **Petals**: $S_1 \setminus Y, \ldots, S_r \setminus Y$ are pairwise disjoint.

Example

$\{\{1,2\},\{1,3,4,6\},\{1,5\},\{2,3\}\}$ is a 4-set system of size 4.
It has a 3-sunflower $\{\{1,2\},\{1,3,4,6\},\{1,5\}\}$ with kernel $\{1\}$ and petals $\{2\},\{3,4,6\},\{5\}$.

Main result

Theorem (Erdős-Rado sunflower)

Any w-set system of size s has an r-sunflower.

Let's focus on $r = 3$.

- Erdős and Rado 1960: $s = w! \cdot 2^w \approx w^w$.
- Kostochka 2000: $s \approx (w \log\log\log w / \log\log w)^w$.
- Fukuyama 2018: $s \approx w^{0.75w}$.
- Now: $s \approx (\log w)^w$ and this is tight for our approach.

Actual bound and further refinement

Theorem (Improved sunflower lemma)

For some constant C, any w-set system of size s has an r-sunflower, where

$$s = \left(Cr^2 \cdot \left(\log w \log\log w + (\log r)^2\right)\right)^w .$$

Recently, Anup Rao improved it to

$$s = \left(Cr(\log w + \log r)\right)^w .$$

Applications – Theoretical computer science

- Circuit lower bounds
- Data structure lower bounds
- Matrix multiplication
- Pseudorandomness
- Cryptography
- Property testing
- Fixed parameter complexity
- ...

Applications – Combinatorics

- Erdős-Szemerédi sunflower lemma
- Intersecting set systems
- Packing Kneser graphs
- Alon-Jaeger-Tarsi nowhere-zero conjecture
- Thersholds in random graphs
- ...

Section 3

Proof overview

Make it robust

Assume $\mathcal{F} = \{S_1, \ldots, S_m\}$ is a w-set system. Define a width-w DNF (disjunctive normal form) $f_{\mathcal{F}}$ as $f_{\mathcal{F}} = \bigvee_{i=1}^{m} \bigwedge_{j \in S_i} x_j$.

Example

If $\mathcal{F} = \{\{1,2\}, \{1,3,4,6\}, \{1,5\}, \{2,3\}\}$, then
$f_{\mathcal{F}} = (x_1 \wedge x_2) \vee (x_1 \wedge x_3 \wedge x_4 \wedge x_6) \vee (x_1 \wedge x_5) \vee (x_2 \wedge x_3)$.

Definition (Satisfying system)

$\mathcal{F}$ is satisfying if $\Pr[f_{\mathcal{F}}(x) = 0] < 1/3$ with $\Pr[x_i = 1] = 1/3$, i.e., $\Pr[\forall i \in [m], S_i \not\subset S] < 1/3$ with $\Pr[x_i \in S] = 1/3$.

Satisfyingness implies sunflower

Assume $\mathcal{F}$ is a set system on ground set $\{x_1, \ldots, x_n\}$.

Lemma

If $\mathcal{F}$ is satisfying, then it has 3 pairwise disjoint sets.

Proof.

Color $x_1, \ldots, x_n$ to red, green, blue uniformly and independenty.
By definition, $\mathcal{F}$ contains a purely red (green/blue) set w.p $> 2/3$.
By union bound, $\mathcal{F}$ contains one purely red set, one purely green set, and one purely blue set w.p > 0. $\square$

In particular, 3 pairwise disjoint sets is a 3-sunflower.

Structure vs pseudorandomness

Assume $\mathcal{F} = \{S_1, \ldots, S_m\}, m > \kappa^w$ is a w-set system. Define link $\mathcal{F}_Y = \{S_i \backslash Y \mid Y \subset S_i\}$, which is a $(w - |Y|)$-set system.

Example

If $\mathcal{F} = \{\{1,2\}, \{1,3,4\}, \{1,5\}, \{2,3\}\}$, then $\mathcal{F}_{\{2\}} = \{\{1\}, \{3\}\}$.

If there exists Y such that $|\mathcal{F}_Y| \geq m/\kappa^{|Y|} > \kappa^{w-|Y|}$, then we can apply induction and find 3-sunflower in $\mathcal{F}_Y$.

So induction starts at such $\mathcal{F}$, that $|\mathcal{F}_Y| < m/\kappa^{|Y|}$ holds for any Y.

Lemma

Let $\kappa \geq (\log w)^{O(1)}$. If $|\mathcal{F}_Y| < m/\kappa^{|Y|}$ holds for any Y, then $\mathcal{F}$ is satisfying, which means there are 3 pairwise disjoint sets in $\mathcal{F}$.

Randomness preserves pseudorandomness

Let $\mathcal{F} = \{S_1, \dots, S_m\}$ be a w-(multi-)set system.
Assume $|\mathcal{F}_Y| < m/\kappa^{|Y|}$ holds for any Y. $\Leftarrow$ $\mathcal{F}$ is pseudorandom
Take $\approx 1/\sqrt{\kappa}$-fraction of the ground set as W,
and construct a $w/2$-(multi-)set system $\mathcal{F}'$ from each S_i:

- Good: If there exists $|S_j \backslash W| \le w/2$ and $S_j \backslash W \subset S_i \backslash W$, then put $S_j \backslash W$ into $\mathcal{F}'$; (j may equal i)
  To satisfy $\{\{1\}, \{1,2,3\}\}$, it suffices to satisfy $\{\{1\}, \{1\}\}$.
- Bad: otherwise, we do nothing for S_i.

Example

If $\mathcal{F} = \{\{1,2\}, \{1,3\}, \{2,3,4\}, \{4,5,6,7\}\}$ and $w = 4, W = \{1\}$, then $\mathcal{F}' = \{\{1,2\}, \{1,3\}, \{2,3,4\}, \{4,5,6,7\}\}$.

One reduction step

Then $|\mathcal{F}'| \approx |\mathcal{F}|$ and $|\mathcal{F}'_Y| \leq |\mathcal{F}_Y|, \forall Y$. $\Leftarrow$ $\mathcal{F}'$ is also pseudorandom
Prove by encoding bad $(W, i) \to (W' = W \cup S_i, \mathsf{aux}_1, k, \mathsf{aux}_2)$,
where $S_j \backslash W \subset S_i \backslash W$ and S_i ranks $k < |\mathcal{F}|/\kappa^{w/2}$ in $\mathcal{F}_{S_j \cap S_i}$.

Example

$\mathcal{F}' = \{\{1,2\}, \{2,3,4\}, \{1,4,5,6\}, \{4,5,6,7\}\}, W = \{1\}, i = 4.$
Encode/decode bad pair (W, i):

- $W' = W \cup S_i = \{1,4,5,6,7\}$ — we find $j = 3$ with $S_j \subset W'$
- $\mathsf{aux}_1 = *\$\$\$$ with at least $w/2$ \$s — we know $S_j \cap S_i = \{4,5,6\}$
- $k = 2$ — S_i ranks 2 in $\mathcal{F}_{\{4,5,6\}}$, we recover $i = 4$
- $\mathsf{aux}_2 = \$\$\$\$$ — we recover $W = W' \backslash \{4,5,6,7\}$

Reductions

Let $\mathcal{F} = \{S_1, \ldots, S_m\}$ be a w-(multi-)set system on $\{x_1, \ldots, x_n\}$.
Assume $|\mathcal{F}_Y| < m/\kappa^{|Y|}$ holds for any Y, and $\kappa = (\log w)^{O(1)}$.
It suffices to prove

- $\mathcal{F}$ is satisfying
  $\iff$ w.h.p S contains some set of $\mathcal{F}$, and $\Pr[x_i \in S] = 1/3$.

Split S to several steps,

- $\Pr[x_i \in S] = 1/3$
  $\approx$ take $1/3$-fraction of the ground set as S
  $\approx$ view S as $W_1, W_2, \ldots, W_{\log w}$, each of $\approx 1/\sqrt{\kappa}$-fraction

Then we iteratively apply reductions,

$$\mathcal{F} \xrightarrow{W_1} \mathcal{F}' \xrightarrow{W_2} \mathcal{F}'' \xrightarrow{W_3} \cdots \xrightarrow{W_{\log w}} \mathcal{F}^{\text{last}}.$$

Final step

Recall $S = W_1 \cup \cdots \cup W_{\log w}$ and

$$\mathcal{F} \xrightarrow{W_1} \mathcal{F}' \xrightarrow{W_2} \mathcal{F}'' \xrightarrow{W_3} \cdots \xrightarrow{W_{\log w}} \mathcal{F}^{\text{last}}.$$

- either we stop at W_i when some set is contained in $\bigcup_{j<i} W_j$,
  $\Rightarrow$ S contains some set of $\mathcal{F}$
- or, $\mathcal{F}^{\text{last}}$ is a width-0 (multi-)set system of size $\approx m > \kappa^w$, and $\left|\mathcal{F}^{\text{last}}_Y\right| \lessapprox \left|\mathcal{F}^{\text{last}}\right| / \kappa^{|Y|}$ still holds for any Y.
  $\Rightarrow$ Impossible

Thus, (informally) we proved such $\mathcal{F}$ is satisfying, which means $\mathcal{F}$ has 3-sunflower (3 pairwise disjoint sets).

Section 4

Open problems

Problem 1 – Erdős-Rado sunflower

Problem (Erdős-Rado sunflower conjecture)

Any w-set system of size $O_r(1)^w$ has r-sunflower.

- Our robust sunflower cannot overcome $(\log w)^{(1-o(1))w}$.
  We need new ideas.
- Lift the sunflower size?
  $r = 3 \implies r = 4$
- Is $(\log w)^{(1-o(1))w}$ actually tight? Counterexamples?

Problem 2 – Erdős-Szemerédi sunflower

Assume $\mathcal{F} = \{S_1, \ldots, S_m\}$ and $S_i \subset \{1, 2. \ldots, n\}$.

Problem (Erdős-Szemerédi sunflower conjecture)

There exists function $\varepsilon = \varepsilon(r) > 0$, such that, if $m > 2^{n(1-\varepsilon)}$, then $\mathcal{F}$ has r-sunflower.

- Now:
  - general r: $\varepsilon = O_r(1/\log n)$ from ER sunflower.
  - $r = 3$: Naslund proved it using polynomial method.
- ER sunflower conjecture $\implies$ ES sunflower conjecture.

Section 5

Thanks