From DNF compression to sunflower theorems via regularity

Jiapeng Zhang

UC San Diego → Harvard University

July 20, 2019

Joint works with Shachar Lovett and Noam Solomon
Overview

1. Sunflower families
2. DNF sparsification
3. Regular set systems
4. Open problems
The Sunflower structure

Definition (Set systems)

Let $X$ be a set. A set system $\mathcal{F} = \{S_1, \ldots, S_m\}$ on $X$ is a collection of subsets of $X$. We call $\mathcal{F}$ a $w$-set system if each set of $\mathcal{F}$ has size at most $w$. 

Definition (Sunflower, Erdős and Rado)

Given $r$ sets $S_1, \ldots, S_r \subseteq X$ where $r \geq 3$. Denote as $B = S_1 \cap \cdots \cap S_r$.

We say it is an $r$-sunflower if for any $i, j \in [r]$, $S_i \cap S_j = B$.

We call $B$ the kernel of this sunflower. In this talk, we focus on $r = 3$. 

Jiapeng Zhang UCSD
The Sunflower structure

**Definition (Set systems)**

Let $X$ be a set. A set system $\mathcal{F} = \{S_1, \ldots, S_m\}$ on $X$ is a collection of subsets of $X$. We call $\mathcal{F}$ a $w$-set system if each set of $\mathcal{F}$ has size at most $w$.

**Definition (Sunflower, Erdős and Rado)**

Given $r$ sets $S_1, \ldots, S_r \subseteq X$ where $r \geq 3$. Denote as

$$B = S_1 \cap \cdots \cap S_r$$

We say it is a $r$-sunflower if for any $i, j \in [r]$, $S_i \cap S_j = B$. 
The Sunflower structure

Definition (Set systems)

Let \( X \) be a set. A set system \( \mathcal{F} = \{S_1, \ldots, S_m\} \) on \( X \) is a collection of subsets of \( X \). We call \( \mathcal{F} \) a \( w \)-set system if each set of \( \mathcal{F} \) has size at most \( w \).

Definition (Sunflower, Erdős and Rado)

Given \( r \) sets \( S_1, \ldots, S_r \subseteq X \) where \( r \geq 3 \). Denote as

\[
B = S_1 \cap \cdots \cap S_r
\]

We say it is a \( r \)-sunflower if for any \( i, j \in [r] \), \( S_i \cap S_j = B \). We call \( B \) the kernel of this sunflower. In this talk, we focus on \( r = 3 \).
Sunflowers: an example

An example

Let $\mathcal{F} = \{\{a, b, x\}, \{a, b, y\}, \{a, b, z\}\}$. Then $\mathcal{F}$ is a 3-sunflower. Its kernel is $\{a, b\}$. 

\[ \begin{align*} 
\text{\{x\}} & \quad \text{\{a, b\}} \\
\text{\{y\}} & \quad \text{\{a, b\}} \\
\text{\{z\}} & \quad \text{\{a, b\}} 
\end{align*} \]
The sunflower lemma (conjecture)

**Lemma (Erdős and Rado)**

Let $\mathcal{F}$ be a $w$-set system such that $|\mathcal{F}| \geq w! \cdot 2^w$, then it contains a sunflower of size 3.
The sunflower lemma (conjecture)

Lemma (Erdős and Rado)

Let $\mathcal{F}$ be a $w$-set system such that $|\mathcal{F}| \geq w! \cdot 2^w$, then it contains a sunflower of size 3. Notice that $w! \approx w^w$. 
The sunflower lemma (conjecture)

Lemma (Erdős and Rado)

Let $\mathcal{F}$ be a $w$-set system such that $|\mathcal{F}| \geq w! \cdot 2^w$, then it contains a sunflower of size 3. Notice that $w! \approx w^w$.

Conjecture (Erdős and Rado)

There is a constant $C > 0$ such that for any $w$-set system $\mathcal{F}$ with $|\mathcal{F}| \geq C^w$, it contains a sunflower of size 3.
DNF sparsification
Disjunctive Normal Forms

Definition

A DNF (Disjunctive Normal Form) is disjunction of conjunctive terms. The size of a DNF is the number of terms, and the width of a DNF is the maximal number of literals in a term.

Example

The function \( f = (x_1 \land x_2) \lor (x_2 \land x_3 \land x_4) \lor (x_1 \land x_4) \lor (x_2 \land x_5) \) is a DNF of size 4 and width 3.
Set systems and monotone DNFs

A DNF is *monotone* if it contains no negated variables. Monotone DNFs are in one-to-one correspondence with set systems. Formally,

\[
\mathcal{F} = \{S_1, \ldots, S_m\} \iff f_\mathcal{F}(x) := \bigvee_{j \in [m]} \bigwedge_{x_i \in S_j} x_i.
\]
Set systems and monotone DNFs

A DNF is *monotone* if it contains no negated variables. Monotone DNFs are in one-to-one correspondence with set systems. Formally,

\[ \mathcal{F} = \{ S_1, \ldots, S_m \} \iff f_\mathcal{F}(x) := \bigvee_{j \in [m]} \bigwedge_{x_i \in S_j} x_i. \]

\{x_1, x_2\}, \{x_2, x_3, x_4\}, \{x_1, x_4\}, \{x_2, x_5\}

\iff (x_1 \land x_2) \lor (x_2 \land x_3 \land x_4) \lor (x_1 \land x_4) \lor (x_2 \land x_5) \]
DNF compression from (approximate) sunflowers

Theorem (Gopalan-Meka-Reingold)

Let $f$ be a width-$w$ DNF. Then for every $\varepsilon > 0$ there exist two width-$w$ DNFs, $f_{\text{lower}}$ and $f_{\text{upper}}$ such that

(i) $f_{\text{lower}}(x) \leq f(x) \leq f_{\text{upper}}(x)$ for all $x$.

(ii) $\Pr[f_{\text{lower}}(x) \neq f_{\text{upper}}(x)] \leq \varepsilon$ for a uniform random $x$.

(iii) $f_{\text{lower}}$ and $f_{\text{upper}}$ have size $(w \log(1/\varepsilon))^{O(w)}$. 

Remark: The term $w$ comes Rossman's approximate sunflower bound.
Theorem (Gopalan-Meka-Reingold)

Let $f$ be a width-$w$ DNF. Then for every $\varepsilon > 0$ there exist two width-$w$ DNFs, $f_{\text{lower}}$ and $f_{\text{upper}}$ such that

(i) $f_{\text{lower}}(x) \leq f(x) \leq f_{\text{upper}}(x)$ for all $x$.

(ii) $\Pr[f_{\text{lower}}(x) \neq f_{\text{upper}}(x)] \leq \varepsilon$ for a uniform random $x$.

(iii) $f_{\text{lower}}$ and $f_{\text{upper}}$ have size $(w \log(1/\varepsilon))^{O(w)}$.

Remark:

The term $w^w$ comes Rossman’s approximate sunflower bound.
Theorem (Lovett-Zhang)

Let $f$ be a width-$w$ DNF. Then for every $\varepsilon > 0$ there exists a width-$w$ DNFs $f_{\text{lower}}$ such that

(i) $f_{\text{lower}}(x) \leq f(x)$ for all $x$.

(ii) $\Pr[f_{\text{lower}}(x) \neq f(x)] \leq \varepsilon$ for a uniform random $x$.

(iii) $f_{\text{lower}}$ has size $(1/\varepsilon)^{O(w)}$. 

Conjecture

Could we find a $f_{\text{upper}}$ with the same bound as $f_{\text{lower}}$. 

Jiapeng Zhang UCSD
DNF compression beyond sunflowers

**Theorem (Lovett-Zhang)**

Let $f$ be a width-$w$ DNF. Then for every $\varepsilon > 0$ there exists a width-$w$ DNFs $f_{\text{lower}}$ such that

1. $f_{\text{lower}}(x) \leq f(x)$ for all $x$.
2. $\Pr[f_{\text{lower}}(x) \neq f(x)] \leq \varepsilon$ for a uniform random $x$.
3. $f_{\text{lower}}$ has size $(1/\varepsilon)^{O(w)}$.

**Conjecture**

Could we find a $f_{\text{upper}}$ with the same bound as $f_{\text{lower}}$. 
Regular set systems
Regular set systems

Definition

Let $\mathcal{F}$ be a set system on $X$, and $\kappa > 0$. We say that $\mathcal{F}$ is $\kappa$-regular if for every $A \subseteq X$,

$$\frac{|\{ S \in \mathcal{F} : A \subseteq S \}|}{|\mathcal{F}|} \leq \kappa^{-|A|}$$

That is, each variable is in $1/\kappa$ fraction sets, each pair in $1/\kappa^2$ fraction sets, and so on.
Regular set systems

Definition

Let $\mathcal{F}$ be a set system on $X$, and $\kappa > 0$. We say that $\mathcal{F}$ is $\kappa$-regular if for every $A \subseteq X$,

$$\frac{|\{S \in \mathcal{F} : A \subseteq S\}|}{|\mathcal{F}|} \leq \kappa^{-|A|}$$

That is, each variable is in $1/\kappa$ fraction sets, each pair in $1/\kappa^2$ fraction sets, and so on.

We want to find a minimum $\kappa$ so that any $\kappa$-regular set system contains a three sunflower.
Definition

Let $\mathcal{F}$ be a set system on $X$, and $\kappa > 0$. We say that $\mathcal{F}$ is $\kappa$-regular if for every $A \subseteq X$,

$$\frac{\left|\{S \in \mathcal{F} : A \subseteq S\}\right|}{|\mathcal{F}|} \leq \kappa^{-|A|}$$

That is, each variable is in $1/\kappa$ fraction sets, each pair in $1/\kappa^2$ fraction sets, and so on.

We want to find a minimum $\kappa$ so that any $\kappa$-regular set system contains a three sunflower. If $\mathcal{F}$ is not $\kappa$-regular, then we can apply induction on $\mathcal{F}_A := \{S \setminus A : (S \in \mathcal{F}) \land (A \subseteq S)\}$.
From DNF sparsification to sunflowers

**Conjecture: Upper bound DNF sparsification**

Let $f$ be a width-$w$ monotone DNF. Then for every $\varepsilon > 0$ there exists a monotone width-$w$ DNFs $f_{upper}$ such that

1. $f(x) \leq f_{upper}(x)$ for all $x$.
2. $\Pr[f_{upper}(x) \neq f(x)] \leq \varepsilon$ for a uniform random $x$.
3. $f_{upper}$ has size $(1/\varepsilon)^{O(w)}$.

**Theorem (This work)**

Assume the above conjecture is true. There is a constant $c > 1$ such that, for any $w$-set system $\mathcal{F}$ with $|\mathcal{F}| \geq (\log w)^{c\cdot w}$, it contains a 3-sunflower.
Theorem

Assume the upper bound DNF sparsification conjecture holds. Then exists a constant $c > 0$, such that for any $w$-set system $F$, if it is $(\log w)^c$-regular, then

$$\Pr_{W}[\exists S \in F, S \subseteq W] \geq 0.99$$

equivalently, $\Pr_x[f_F(x) = 1] \geq 0.99$
Theorem

Assume the upper bound DNF sparsification conjecture holds. Then exists a constant $c > 0$, such that for any $w$-set system $\mathcal{F}$, if it is $(\log w)^c$-regular, then

$$\Pr_{W}[\exists S \in \mathcal{F}, S \subseteq W] \geq 0.99$$

equivalently, $\Pr_{x}[f_{\mathcal{F}}(x) = 1] \geq 0.99$

Corollary

Assume the above conjecture holds. Then for any $(\log w)^c$-regular set system $\mathcal{F}$, it contains two disjoint sets.
From two disjoint sets to \( r \) disjoint sets.

**Lemma**

Assume the upper bound DNF sparsification conjecture holds. Then exists a constant \( c > 0 \), such that for any \( w \)-set system \( \mathcal{F} \), if it is \((\log w)^{c \cdot r}\)-regular, then \( \mathcal{F} \) contains \( r \) disjoint sets.
From two disjoint sets to $r$ disjoint sets.

**Lemma**

Assume the upper bound DNF sparsification conjecture holds. Then exists a constant $c > 0$, such that for any $w$-set system $\mathcal{F}$, if it is $(\log w)^{c \cdot r}$-regular, then $\mathcal{F}$ contains $r$ disjoint sets.

**Theorem**

Assume the upper bound DNF sparsification conjecture holds. For any $w$-set system $\mathcal{F}$ with $|\mathcal{F}| \geq (\log w)^{c \cdot r \cdot w}$, it contains a $r$-sunflower.
Open problems
Open problems

Question

Could we confirm the upper bound DNF sparsification conjecture?
Open problems

Question
Could we confirm the upper bound DNF sparsification conjecture?

Highlight conjecture
There is a constant $c$. For any $(\log w)^c$-regular $w$-set system $\mathcal{F}$,

$$\Pr_{x}[f_\mathcal{F}(x) = 1] \geq 0.99$$
Open problems

Question
Could we confirm the upper bound DNF sparsification conjecture?

Highlight conjecture
There is a constant $c$. For any $(\log w)^c$-regular $w$-set system $\mathcal{F}$,

$$\Pr_{x}[f_{\mathcal{F}}(x) = 1] \geq 0.99$$

I guess for each variable $x_i$, the $l_{f_{\mathcal{F}}}(x_i) \ll 1/w$. 
Thanks!