A composition lemma for the Fourier Entropy-Influence Conjecture

Ryan O’Donnell (CMU) and Li-Yang Tan (Columbia)

ICALP
12 July 2013, Riga

Fourier analysis of Boolean functions

- Every Boolean function $f\colon\{-1,1\}^n \to \{-1,1\}$ can be expressed as a polynomial $p : \mathrm{R}^n \to \mathrm{R}$

$$f(x) = \sum_{S\subseteq[n]} \hat{f}(S) \prod_{i\in S} x_i$$

$$\mathrm{MAJ}(x_1, x_2, x_3) = \frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 - \frac{1}{2}x_1x_2x_3$$

- Introduced into TCS by Kahn-Kalai-Linial 1988.
- Indispensable tool in TCS: hardness of approximation, learning theory, circuit complexity, communication complexity, ….
- Analytic methods to study inherently combinatorial object.

Parseval's identity

For every $f : \{-1,1\}^n \to \{-1,1\}$,

$$\sum_{S \subseteq [n]} \hat{f}(S)^2 = 1\,.$$

$$\mathrm{MAJ}(x_1, x_2, x_3) = \frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 - \frac{1}{2}x_1x_2x_3$$

Every f induces distribution D_f over all $S \subseteq 2^{[n]}$ where $\Pr_{D_f}[S] = \hat{f}(S)^2$.

"Fourier weight of f on S"

This talk: understanding two fundamental properties of D_f

1. Is D_f concentrated on few sets, or spread out among many?

$\approx$ Is f close to a sparse polynomial?

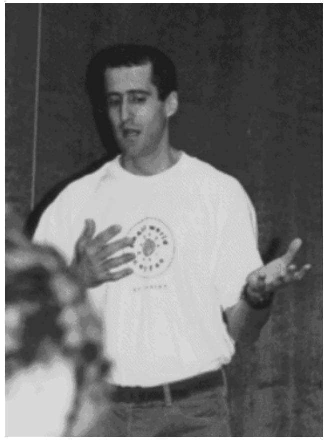
2. Does D_f place most of its weight on high or low degree or large sets S?

$\approx$ Is f close to a low-degree polynomial?

$$\text{Entropy}[f] \leq C \cdot \text{Influence}[f]$$

for some universal constant C?

"Fourier Entropy-Influence Conjecture"
Ehud Friedgut and Gil Kalai, 1996



outline

- The FEI conjecture.
- Motivation, applications, previous work.
- This work:
    1. Composition lemma for FEI
    2. FEI true for read-once formulas
- Thoughts about FEI.

the conjecture

Let $f : \{-1,1\}^n \to \{-1,1\}$, distribution D_f over $2^{[n]}$ where $\Pr_{D_f}[S] = \hat{f}(S)^2$.

$$\text{Entropy}[f] = \mathrm{H}\left[D_f\right] = \sum_{S\subseteq[n]} \hat{f}(S)^2 \cdot \log_2\left(\frac{1}{\hat{f}(S)^2}\right)$$

$$\text{Influence}[f] = \mathrm{E}_{S\sim D_f}[|S|] = \sum_{S\subseteq[n]} \hat{f}(S)^2 \cdot |S| \quad = \text{Avg-degree}[f]$$

$$\text{Entropy[MAJ]} = 4 \cdot \left(\frac{1}{2}\right)^2 \log_2(2^2) = 2$$

$$\mathrm{MAJ}(x_1, x_2, x_3) = \frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 - \frac{1}{2}x_1x_2x_3$$

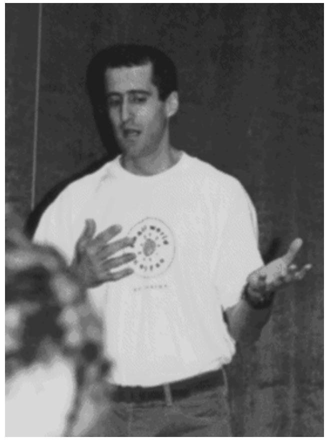
$$\text{Influence[MAJ]} = \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 \cdot 3$$

the conjecture

Let $f: \{-1,1\}^n \to \{-1,1\}$, distribution D_f over $2^{[n]}$ where $\Pr_{D_f}[S] = \hat{f}(S)^2$.

$$\text{Entropy}[f] = \mathrm{H}[D_f] = \sum_{S \subseteq [n]} \hat{f}(S)^2 \cdot \log_2\left(\frac{1}{\hat{f}(S)^2}\right)$$

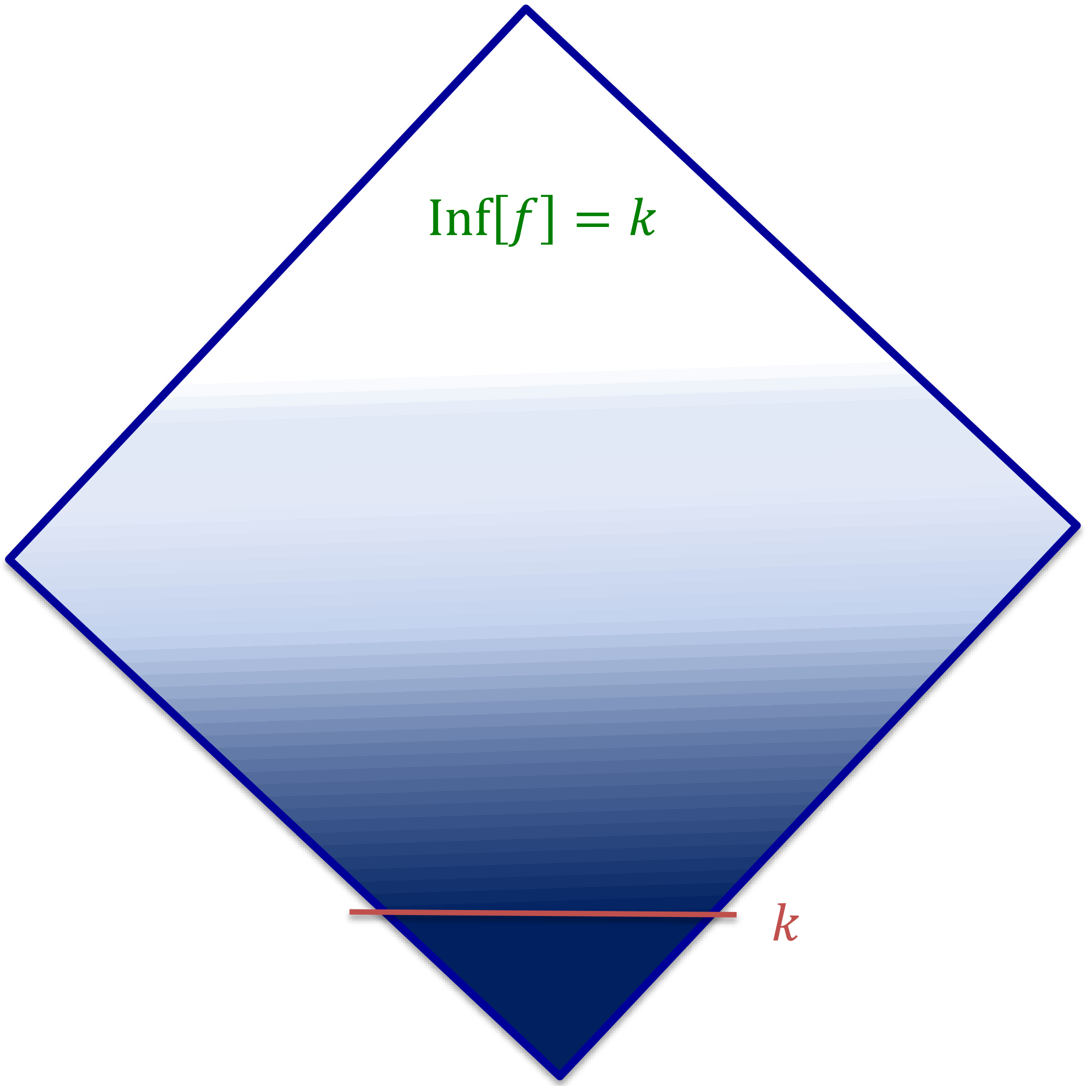
$$\text{Influence}[f] = \mathrm{E}_{S \sim D_f}[|S|] = \sum_{S \subseteq [n]} \hat{f}(S)^2 \cdot |S|$$

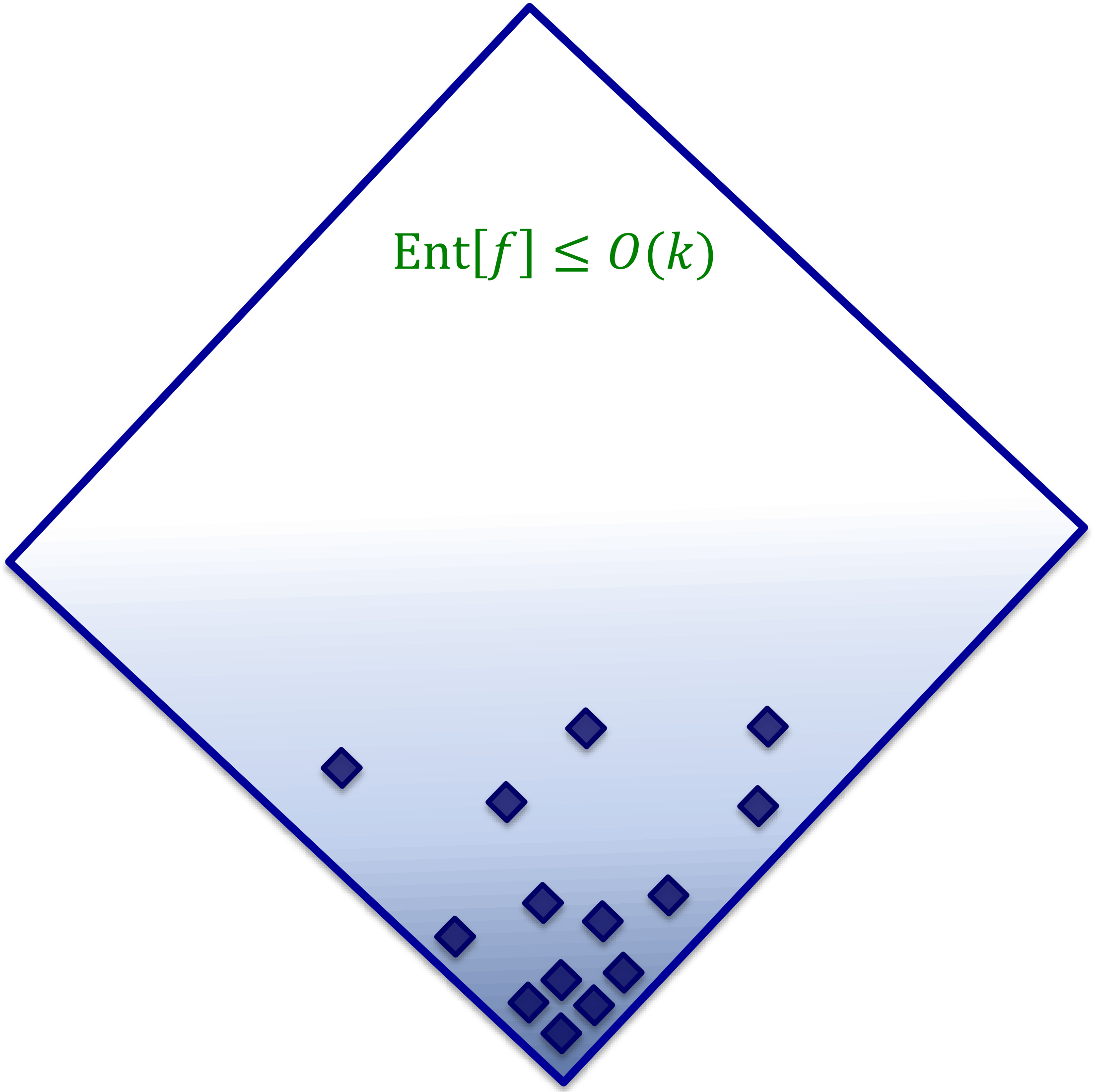


Fourier Entropy-Influence Conjecture:

There exists a universal constant $C > 0$ s.t.
$\text{Entropy}[f] \le C \cdot \text{Influence}[f]$.

FEI: “If f is close to low-degree, then f is close to sparse”





motivation and applications

- Friedgut-Kalai: threshold phenomena in random graphs.
  - e.g. probability that $G \sim G(n, p)$ is connected.
  - sharp threshold: probability jumps from 0.1 to 0.9 by increasing p a little.

- Implies Mansour's conjecture on Fourier spectrum of DNFs.

  $\Longrightarrow$ efficient algorithm for agnostically learning DNFs.

  $\Longrightarrow$ improved PRGs for DNFs.

- Implies the Kahn-Kalai-Linial theorem.
  - Fundamental inequality in analysis of Boolean functions.
  - Inapproximability, metric embeddings, social choice, …

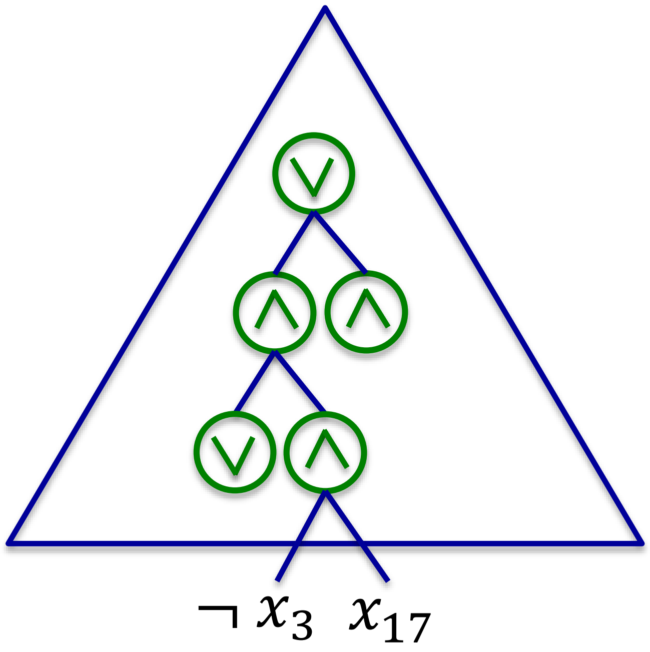
previous work

- Folklore: $\text{Entropy}[f] \le O(\log n) \cdot \text{Influence}[f]$.
- No published progress for 15 years!
- Klivans-Lee-Wan 2010: random $\text{poly}(n)$-term DNFs.
- O’Donnell-Wright-Zhou 2011:
  - symmetric functions
  - read-once decision trees






read-once de Morgan formulas



- Formula over the basis $\{\mathrm{AND}_2, \mathrm{OR}_2, \mathrm{NOT}\}$.
- Binary tree with
  - Internal nodes = $\{\mathrm{AND}_2, \mathrm{OR}_2\}$
  - Leaves = $\{x_i\ , \neg\, x_i\ :\ i \in [n]\}$
- Every variable appears exactly once.

- $\mathrm{AND}_2, \mathrm{OR}_2$: extremely simple Fourier spectra, FEI trivially holds.
- Yet prior to this work, FEI open for read-once formulas.

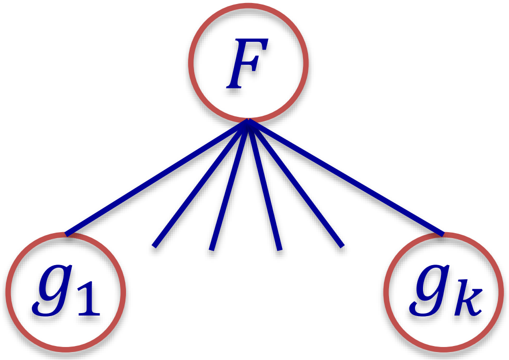
Starting point of this research:

Is FEI preserved under disjoint composition?

(*i.e.* Can we prove a *composition lemma* for FEI?)

a first attempt, dream version

For all $C > 0$, suppose $F: \{-1,1\}^k \to \{-1,1\}$ satisfies $\mathrm{Ent}[F] \leq C \cdot \mathrm{Inf}[F]$,
and $g_1, \ldots, g_k: \{-1,1\}^\ell \to \{-1,1\}$ satisfy $\mathrm{Ent}[g_i] \leq C \cdot \mathrm{Inf}[g_i]$.
Then $f = F(g_1(x^1), \ldots, g_k(x^k))$ satisfies $\mathrm{Ent}[f] \leq C \cdot \mathrm{Inf}[f]$.



- Would immediately imply FEI for read-once formulas
- Unfortunately, easily seen to be false ☹

(e.g. $F = \mathrm{OR}_2, g_1, g_2 = \mathrm{AND}_2$)

Solution: reformulate FEI carefully so that it "bootstraps itself" this work

1. Strengthen FEI to “FEI+”

$$\mathrm{Ent}[F] \le C \cdot (\mathrm{Inf}[f] - \mathrm{Var}[f])$$

2. Generalize FEI+ to “μ-biased FEI+”

$$\boxed{\begin{array}{c}\mathrm{Ent}^{\mu}[F] \le C \cdot (\mathrm{Inf}^{\mu}[f] - \mathrm{Var}^{\mu}[f]) \\ \text{for all product distributions } \mu \text{ over } \{-1,1\}^n\end{array}}$$

3. Prove composition lemma for μ-biased FEI+

composition theorem for FEI

Let $\mu = (\mu_1, \ldots, \mu_n)$ be a product distribution over $\{-1,1\}^n$.

The μ-biased Fourier transform of $f: \{-1,1\}^n \to \{-1,1\}$ is:

$$f(x) = \sum_{S \subseteq [n]} \tilde{f}(S) \prod_{i \in S} \frac{x_i - \mu_i}{1 - \mu_i^2}$$

- Simply the "usual" Fourier transform when $\mu = (0, \ldots, 0)$.
- $\mathrm{Ent}^{\mu}[f]$ and $\mathrm{Inf}^{\mu}[f]$ defined analogously w.r.t. $\tilde{f}(S)^2$'s.

Main Theorem [O'Donnell-T.].

Suppose $g_1, \ldots, g_k$ satisfy $\mathrm{Ent}[g_i] \le C \cdot (\mathrm{Inf}[g_i] - \mathrm{Var}[g_i])$,

and F satisfies $\mathrm{Ent}^{\mu}[F] \le C \cdot (\mathrm{Inf}^{\mu}[F] - \mathrm{Var}^{\mu}[F])$ where $\mu = (\mathrm{E}[g_1], \ldots, \mathrm{E}[g_k])$

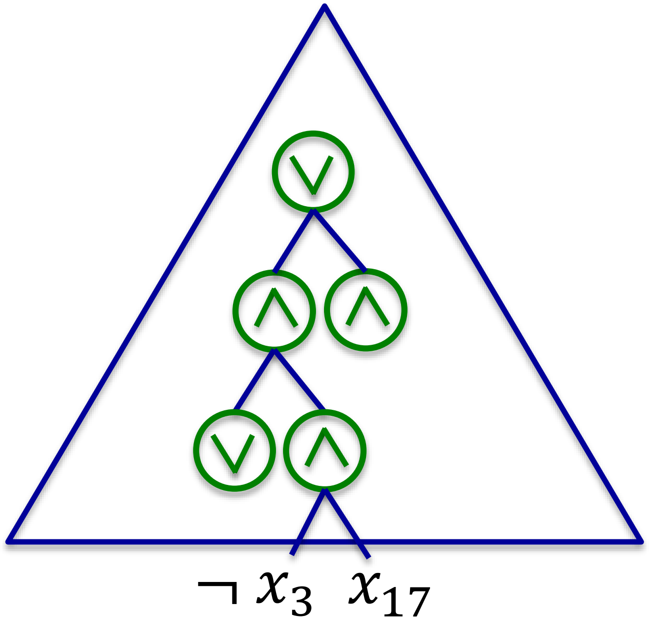
Then $f = F(g_1(x^1), \ldots, g_k(x^k))$ satisfies $\mathrm{Ent}[f] \le C \cdot (\mathrm{Inf}[f] - \mathrm{Var}[f])$.

discussion

Main Theorem [O'Donnell-T.].
Suppose $g_1, \dots, g_k$ satisfy $\mathrm{Ent}[g_i] \leq C \cdot (\mathrm{Inf}[g_i] - \mathrm{Var}[g_i])$,
and F satisfies $\mathrm{Ent}^{\mu}[F] \leq C \cdot (\mathrm{Inf}^{\mu}[F] - \mathrm{Var}^{\mu}[F])$ where $\mu = (\mathrm{E}[g_1], \dots, \mathrm{E}[g_k])$
Then $f = F(g_1(x^1), \dots, g_k(x^k))$ satisfies $\mathrm{Ent}[f] \leq C \cdot (\mathrm{Inf}[f] - \mathrm{Var}[f])$.

- μ-biased generalization natural on hindsight:
  - $F(g_1(x^1), \dots, g_k(x^k))$ with *uniform* $x^1, \dots, x^k$ induces *biased* product distribution $\mu = (\mathrm{E}[g_1], \dots, \mathrm{E}[g_k])$ on F.
- Main conceptual contribution: reformulation and strengthening of FEI
  - Composition lemma: evidence that this is the "correct" formulation?
- Difficulty was in finding right statement to prove.
  - Proof = careful Fourier-analytic computations.

read-once formulas revisited



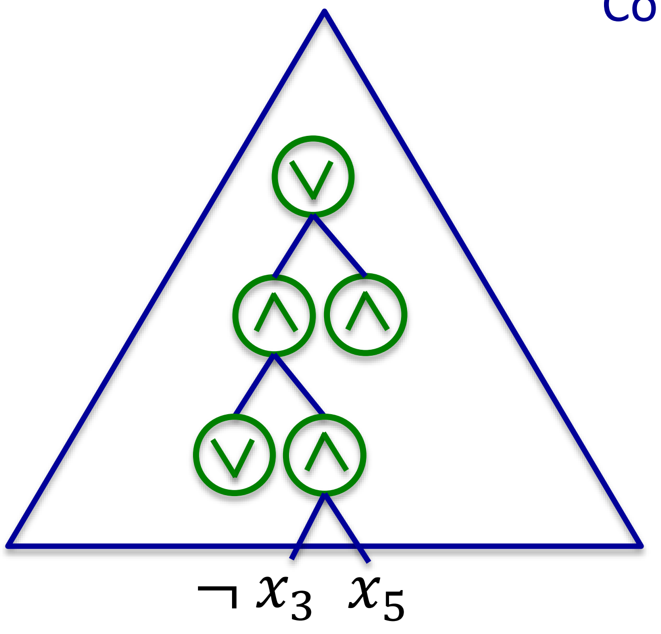
- Dream version of composition $\Longrightarrow$ FEI by structural induction.
- Need to prove $\mathrm{Ent}^\mu[\mathrm{AND}_2] \leq C \cdot (\mathrm{Inf}^\mu[\mathrm{AND}_2] - \mathrm{Var}^\mu[\mathrm{AND}_2])$ for some constant C *independent* of μ. Likewise for OR_2.
- Surprisingly tricky, even for 2-variable functions!
- But after much Fourier analysis…

Theorem [O'Donnell-T.].
Every $F: \{-1,1\}^k \to \{-1,1\}$ satisfies $\mathrm{Ent}^\mu[F] \leq 2^k \cdot (\mathrm{Inf}^\mu[F] - \mathrm{Var}[F])$ for *all* product distributions $\mu = (\mu_1, \ldots, \mu_k)$.

read-once formulas revisited

Theorem [O'Donnell-T.].
Every $F: \{-1,1\}^k \to \{-1,1\}$ satisfies $\mathrm{Ent}^\mu[F] \le 2^k \cdot (\mathrm{Inf}^\mu[F] - \mathrm{Var}[F])$
for *all* product distributions $\mu = (\mu_1, \ldots, \mu_k)$.

Combining with composition lemma, we get:



Corollary [O'Donnell-T.].
FEI holds for the class of read-once de Morgan formulas.

Corollary [O'Donnell-T.].
FEI holds for *any* class read-once formulas over *arbitrary* basis.


thank you!