18.218 Topics in Combinatorics Spring 2021 – Lecture 6

Dor Minzer

In this lecture, we will present more advanced applications of the hypercontractive inequality.

1 The FKN theorem

Recall that in the homework assignment, you have seen that degree 1 Boolean functions must be a dictatorship or an anti-dictatorship. The following theorem is a stability-version of that statement, showing that Boolean functions that are close to being degree 1 are close to dictatorships or anti-dictatorships.

Theorem 1.1. *Suppose a function $f\colon \{-1,1\}^n \to \{-1,1\}$ is ε-close to a degree 1 function in ℓ_2^2, i.e. $\|f - f^{=1}\|_2^2 \leqslant \varepsilon$. Then, there exists $b_i \in \{-1,1\}$ and $i \in [n]$ such that $\|f - b_i x_i\|_2 = O(\varepsilon)$.*

Proof. Let $\ell(x) = f^{=1}(x) = \sum\limits_{i=1}^{n} a_i x_i$. Expanding, we see that

$$\ell(x)^2 = \sum_{i=1}^{n} a_i^2 + 2\sum_{i<j} a_i a_j x_i x_j.$$

Therefore, $\mathsf{var}(\ell^2) = 4\sum\limits_{i<j} a_i^2 a_j^2 = 2\left(\left(\sum\limits_{i=1}^{n} a_i^2\right)^2 - \sum\limits_{i=1}^{n} a_i^4\right)$. We will argue that the variance of ℓ^2 is small, and $\sum\limits_{i=1}^{n} a_i^2 \approx 1$, from which the proof will quickly be concluded.

Bounding $\sum\limits_{i=1}^{n} a_i^2$. Note that $\sum\limits_{i=1}^{n} a_i^2 = \mathbb{E}\left[\ell^2\right] = \mathbb{E}\left[f^2\right] - \mathbb{E}\left[(f-\ell)^2\right] \geqslant 1 - \varepsilon$.

Bounding $\mathsf{var}(\ell^2)$. Letting $h(x) = \ell^2(x) - \mathbb{E}\left[\ell^2\right] = \sum\limits_{i \neq j} a_i a_j x_i x_j$, we have $\mathsf{var}(\ell^2) = \|h\|_2^2$, and our goal is to bound the 2-norm of h. By the 1-norm trick, as h is a degree 2 function $\|h\|_2 \leqslant 9\|h\|_1$, and computing

$$\|h\|_1 \leqslant \|\ell^2 - f^2\|_1 + \left|\mathbb{E}\left[\ell^2\right] - 1\right| \leqslant \|\ell - f\|_2 \|\ell + f\|_2 + \varepsilon \leqslant \sqrt{\varepsilon}2 + \varepsilon \leqslant 3\sqrt{\varepsilon}.$$

Thus, $\|h\|_2 \leqslant 27\sqrt{\varepsilon}$.

Finishing the proof. Staring now at $\mathsf{var}(\ell^2) = \left(\sum\limits_{i=1}^{n} a_i^2\right)^2 - \sum\limits_{i=1}^{n} a_i^4$, we get that

$$(27\sqrt{\varepsilon})^2 \geqslant (1-\varepsilon)^2 - \sum_{i=1}^{n} a_i^4,$$

so $\sum_{i=1}^{n} a_i^4 \geqslant 1 - O(\varepsilon)$. Thus,

$$\max_i a_i^2 \sum_{i=1}^{n} a_i^2 \geqslant 1 - O(\varepsilon),$$

and as $\sum_{i=1}^{n} a_i^2 \leqslant 1$, we get $\max_i a_i^2 \geqslant 1 - O(\varepsilon)$. This shows that there is $i^\star$ such that $|a_{i^\star}| \geqslant 1 - O(\varepsilon)$. Assume without loss of generality that $a_{i^\star} \geqslant 1 - O(\varepsilon)$; we thus get

$$1 - O(\varepsilon) \leqslant a_{i_\star} = \widehat{f}(\{i_\star\}) = 2\Pr_x[f(x) = x_{i_\star}] - 1,$$

and so $\Pr_x[f(x) = x_{i_\star}] \geqslant 1 - O(\varepsilon)$. □

Remark 1.2. *An interesting question which is not fully understood asks for extensions of this theorem to degree d functions. Namely, what can one say about a degree d function that is close to Boolean? The question however is more delicate, as the precise notion of closeness depends on d, and we will not elaborate on this further for now.*

2 The Fourier spectrum of small-sets

Suppose $S \subseteq \{-1,1\}^n$ is a small set, i.e. $|S| = \delta 2^n$ for a small δ. What can we say about the Fourier spectrum of 1_S?

Claim 2.1. $\mathsf{deg}(1_S) \geqslant \Omega(\log(1/\delta))$.

Proof. Let d be the degree of 1_S. Then

$$\delta = \|1_S\|_2^2 = \langle 1_S, 1_S\rangle \leqslant \|1_S\|_{4/3}\|1_S\|_4 \leqslant \|1_S\|_{4/3}\sqrt{3}^d\|1_S\|_2 = \sqrt{3}^d\delta^{5/4},$$

so $\sqrt{3}^d \geqslant \delta^{-1/4}$, hence $d \geqslant \frac{1}{4\log\sqrt{3}}\log(1/\delta)$. □

While this proof is very simple, there is an even simpler argument to prove this statement based on the Schwarz-Zippel argument (which says that a degree d function on the Boolean cube must be non-zero on at least 2^{-d} fraction of the points). However, one can adapt the above argument to say something much stronger: not only is the degree of 1_S must be $\Omega(\log(1/\delta))$, but in fact most of its Fourier mass lies on such levels.

For technical reasons, we will prove the following slightly more general statement.

Lemma 2.2. *Let $f\colon \{-1,1\}^n \to \{-1,0,1\}$ be a function such that $0 < \Pr_x[f(x) \neq 0] \leqslant \delta$. Then*

$$\sum_{|S| \leqslant \frac{1}{20}\log(1/\delta)} \widehat{f}(S)^2 \leqslant \delta^{24/20}.$$

Proof. Let $d = \frac{1}{20}\log(1/\delta)$. Introducing the notation $f^{\leqslant d}(x) = \sum_{|S|\leqslant d}\widehat{f}(S)\chi_S(x)$, our quantity of interest to bound is $\|f^{\leqslant d}\|_2^2 = \sum_{|S|\leqslant d}\widehat{f}(S)^2$. We do that as follows:

$$\|f^{\leqslant d}\|_2^2 = \langle f^{\leqslant d}, f^{\leqslant d}\rangle = \langle f^{\leqslant d}, f\rangle \leqslant \|f^{\leqslant d}\|_4\|f\|_{4/3} \leqslant \sqrt{3}^d\|f^{\leqslant d}\|_2\|f\|_{4/3} \leqslant \sqrt{3}^d\|f\|_2\|f\|_{4/3}.$$

By the premise $\|f\|_2 \leqslant \delta^{1/2}$ and $\|f\|_{4/3} \leqslant \delta^{3/4}$. We thus get $\|f^{\leqslant d}\|_2^2 \leqslant e^d\delta^{5/4} \leqslant \delta^{-1/20}\delta^{5/4} \leqslant \delta^{24/20}$. □

Note that as the overall Fourier mass of f is δ, and $\delta^{24/20} \ll \delta$ for small δ, the lemma says that a (signed) indicator of a small set has almost all of its mass on the high-degrees.

Remark 2.3. *With a bit more effort, one may even show a bound of the form $\delta^2 \log^d(1/\delta)$, and for some applications this quantitative difference is important; see homework assignment.*

3 The KKL theorem

What can we say about Boolean functions that have small average degree, i.e. $I[f] \leqslant K\mathsf{var}(f)$?

Theorem 3.1. *Let $f\colon \{-1,1\}^n \to \{-1,1\}$ be such that $I[f] \leqslant K\mathsf{var}(f)$. Then there exists $i \in [n]$ such that*

$$I_i[f] \geqslant e^{-O(K)}.$$

Proof overview. Before giving the proof, we will give the rough intuition. First, as $I[f] \leqslant K\mathsf{var}(f)$ and $I[f]$ is the average degree, a Markov-inequality type bound shows that all but little bit of the Fourier mass of f lies on degrees $O(K)$, and hence it makes sense to consider the low-degree part of f. We will decompose this low-degree part according to the contribution of different coordinates to it (via the derivatives), and then upper bound each one of these contributions separately using the tools we have developed so far. Since the total Fourier mass we have is roughly $\mathsf{var}(f)$, the contribution of at least one of these coordinates is meaningful, and that will be the influential coordinate we are looking for.

Proof. Suppose towards contradiction that $I_i[f] \leqslant e^{-C\cdot K} =: \delta$ for all $i \in [n]$, where C is an absolute constant to be determined later. Fix i, and consider the function $g = \partial_i f(x)$; note that g is $-1, 0, 1$ valued, and the probability it is non-zero is $I_i[f] \leqslant \delta$, so by Lemma 2.2 we have

$$\sum_{S:|S|\leqslant \frac{1}{20}\log(1/\delta)} \widehat{g}(S)^2 \leqslant I_i[f]^{24/20}.$$

Let us translate this now into information about the Fourier spectrum of f. If $i \in S$, $\widehat{g}(S) = \widehat{f}(S)$ and otherwise it is 0, so we get that

$$\sum_{\substack{S:|S|\leqslant \frac{1}{20}\log(1/\delta) \\ i\in S}} \widehat{f}(S)^2 \leqslant I_i[f]^{24/20}.$$

Summing this over i, we get that

$$\sum_{S:0<|S|\leqslant \frac{1}{20}\log(1/\delta)} |S|\,\widehat{f}(S)^2 \leqslant \sum_{i=1}^{n} I_i[f]^{24/20} \leqslant \delta^{1/5} I[f] \leqslant \delta^{1/5} K\mathsf{var}(f) \leqslant e^{-CK} K\mathsf{var}(f).$$

Therefore,

$$\sum_{S:|S|\leqslant \frac{C}{20}K} \widehat{f}(S)^2 \leqslant \delta^{1/5} K\mathsf{var}(f),$$

and on the other hand

$$\sum_{S:|S|>\frac{C}{20}K}\widehat{f}(S)^2 \leqslant \frac{\sum_{S:|S|>\frac{C}{20}K}|S|\,\widehat{f}(S)^2}{\frac{C}{20}K} \leqslant \frac{I[f]}{\frac{C}{20}K} \leqslant \frac{20}{C}\mathsf{var}(f).$$

Combining the two inequalities, we get that

$$\sum_{S:0<|S|}\widehat{f}(S)^2 \leqslant \frac{20}{C}\mathsf{var}(f) + e^{-CK}K\mathsf{var}(f) < \mathsf{var}(f),$$

where the last inequality holds for appropriate C ($C = 40$ will do), and contradiction. □

As an immediate corollary, we get a more standard formulation of the KKL theorem.

Corollary 3.2. *For any $f\colon \{-1,1\}^n \to \{-1,1\}$, there is $i \in [n]$ such that $I_i[f] \geqslant \Omega\left(\frac{\log n}{n}\mathsf{var}(f)\right)$.*

Proof. Let $C > 0$ be the implicit constant from Theorem 3.1, i.e. absolute C such that $\max_i I_i[f] \geqslant e^{-C\frac{I[f]}{\mathsf{var}(f)}}$. If $I[f] \leqslant \frac{1}{2C}\mathsf{var}(f)\log n$, we get from Theorem 3.1 that

$$\max_i I_i[f] \geqslant e^{-\log n/2} = \frac{1}{\sqrt{n}} \geqslant \frac{\log n}{n}\mathsf{var}(f).$$

Otherwise, $I[f] \geqslant \frac{1}{2C}\mathsf{var}(f)\log n$, and so

$$\max_i I_i[f] \geqslant \frac{I[f]}{n} \geqslant \frac{1}{2C}\frac{\log n}{n}\mathsf{var}(f).$$ □

4 Tightness of the KKL theorem

The following example, called the "Tribes" function, shows that the KKL theorem as well as Friedgut's theorem are tight.

Claim 4.1. *There exists $f\colon \{0,1\}^n \to \{0,1\}$ with $\mathsf{var}(f) \geqslant \Omega(1)$ and $I_i[f] = O(\log n/n)$ for all $i \in [n]$.*

Proof. Take $k, \ell \in \mathbb{N}$ such that $\ell k \leqslant n$, and take $I_1, \ldots, I_k \subseteq [n]$ disjoint each of size ℓ. Define the function $f(x) = \bigvee_{j=1}^{k}\bigwedge_{i\in I_j} x_i$. Note that

$$\mathbb{E}[f] = \Pr_x\left[f(x) = 1\right] = 1 - \Pr_x\left[f(x) = 0\right] = 1 - \prod_{j=1}^{k}\Pr\left[\bigwedge_{i\in I_j} x_i = 0\right] = 1 - (1-2^{-\ell})^k,$$

so if we take $k = 2^\ell$ we will have that $\mathbb{E}[f]$ is bounded away from 0 and 1, and hence $\mathsf{var}(f) \geqslant \Omega(1)$.

Indeed, we choose $k = 2^\ell$, and then the constraint on ℓ, k turns into $\ell 2^\ell \leqslant n$, and it is enough to choose $\ell = \lfloor \log n - \log\log n\rfloor$.

We finish the proof by computing the influences of f. Fix i, and assume without loss of generality $i \in I_1$. Note that i is influential on x if and only if $\bigwedge_{q\in I_j} x_q = 0$ for all $j > 1$, and $\bigwedge_{q\in I_1\setminus\{i\}} x_q = 1$, so

$$\Pr_x\left[f(x) \neq f(x\oplus e_i)\right] = \Pr_x\left[\bigwedge_{q\in I_1\setminus\{i\}} x_q = 1\right]\prod_{j=2}^{k}\Pr_x\left[\bigwedge_{q\in I_j} x_q = 0\right] = 2^{\ell-1}(1-2^{-\ell})^{k-1} = \Theta(\log n/n).$$

□

18.218 Topics in Combinatorics: Analysis of Boolean Functions
Spring 2021