1 Friedgut and KKL

Lemma 1.1. *For any* $f : \{0,1\}^n \to \{-1,1\}$, $0 \le \rho \le 1$,

$$\mathrm{Inf}_i^\rho(f) = \sum_{S \ni i} \rho^{|S|} \hat{f}(S)^2 \le (\mathrm{Inf}_i(f))^{2/(1+\rho)}$$

Remarks:

- Small influence gets much smaller after noise
- Not a Fourier statement
- Means that small influences cannot come from low levels

Proof. Define $f_i : \{0,1\}^n \to \{-1,0,1\}$ by $f_i = \frac{1}{2}(f(x) - f(x \oplus e_i))$. Then $f_i = \sum_{S \ni i} \hat{f}(S)\chi_S$, so

$$\mathrm{Inf}_i^\rho(f) = \|T_{\sqrt{\rho}}(f_i)\|_2^2 \overset{HC}{\le} \|f_i\|_{\rho+1}^2 \overset{\text{range of } f}{=} \|f_i\|_2^{4/(\rho+1)} = (\mathrm{Inf}_i(f))^{2/(\rho+1)}.$$

□

Corollary 1.2. *For all* i, d, ρ,

$$\sum_{\substack{S \ni i \\ |S| \le d}} \hat{f}(S)^2 \le \rho^{-d} (\mathrm{Inf}_i(f))^{2/(1+\rho)}.$$

Lemma 1.3. *For* $f : \{0,1\}^n \to \{-1,1\}$, $\varepsilon \in (0,1)$, *let* $d = 2\mathbb{I}(f)/\varepsilon$ *and*

$$J = \{j \in [n] \ : \ \mathrm{Inf}_j(f) \ge 64^{-d}\}.$$

Then f *is* ε*-concentrated on*

$$\mathcal{S} = \{S \subseteq J \ : \ |S| \le d\}.$$

Corollary 1.4 (Friedgut's theorem [Fri98])**.** *For any* $f : \{0,1\}^n \to \{-1,1\}$, *and* $\varepsilon \in (0,1)$, f *is* ε*-close to a* $2^{O(\mathbb{I}(f)/\varepsilon)}$ *junta.*

Proof. Define $g = \mathrm{sign}(\sum_{\substack{|S| \le d \\ S \subseteq J}} \hat{f}(S)\chi_S)$. Since f is ε-concentrated on $\mathcal{S}$, g is ε-close to f (as we saw in a previous class), and clearly g is a $|J|$-junta, and $|J| = 2^{O(\mathbb{I}(f)/\varepsilon)}$. □

Proof of lemma.

$$\sum_{S \notin \mathcal{S}} \hat{f}(S)^2 = \underbrace{\sum_{|S| > d} \hat{f}(S)^2}_{\le \varepsilon/2 \text{ since } d = 2\mathbb{I}(f)/\varepsilon} + \sum_{|S| \le d, S \not\subseteq J} \hat{f}(S)^2.$$

Now,

$$\begin{aligned}\sum_{|S|\le d, S\not\subseteq J} \hat{f}(S)^2 &\le \sum_{i\notin J}\sum_{S\ni i,|S|\le d} \hat{f}(S)^2 \\ &\overset{\rho=1/2}{\le} \sum_{i\notin J} 2^d(\mathrm{Inf}_i(f))^{4/3} \\ &\overset{*}{\le} 2^d \sum_{i\notin J} 4^{-d}\,\mathrm{Inf}_i(f) \\ &\le 2^{-d}\mathbb{I}(f) = 2^{-2\mathbb{I}(f)/\varepsilon}\mathbb{I}(f) \\ &\overset{2^{-x}\le 1/x}{\le} \varepsilon/2.\end{aligned}$$

(*): at this point we could just write $(64^{-d})^{4/3}$, but not good enough... □

Corollary 1.5 (Kahn, Kalai, and Linial [KKL88]). *For any balanced $f : \{0,1\}^n \to \{-1,1\}$, there exists i s.t.* $\mathrm{Inf}_i(f) \ge \log n/(48n)$.

Proof. Assume by contradiction that $\forall i, \mathrm{Inf}_i(f) < \log n/(48n)$. Then $\mathbb{I}(f) < \frac{1}{48}\log n$. So if we take $\varepsilon = 1/2$ we get that f is $1/2$-concentrated on subsets of

$$J = \{j \mid \mathrm{Inf}_j(f) \ge 64^{-d} = 64^{-4\mathbb{I}(f)} > 64^{-\log n/12} = 1/\sqrt{n}\} = \emptyset,$$

in contradiction to $\sum_{S\neq\emptyset} \hat{f}(S)^2 = 1$. □

2 Friedgut-Kalai-Naor 2002

Recall that we showed that

$$\Pr[\text{NAE accepts } f] = \frac{3}{4} - \frac{3}{4}\sum_S (-\frac{1}{3})^{|S|}\hat{f}(S)^2 \le \frac{7}{9} + \frac{2}{9}W_1(f).$$

So if the test accepts with probability $1-\varepsilon$ then $W_1(f) \ge 1 - \frac{9}{2}\varepsilon$. We already know that if $W_1(f) = 1$ (i.e., when all its Fourier mass is in the first level) then f is a dictator or antidictator, but let's show it again with a different proof: (Notice that the claim is false for non-Boolean functions, i.e., $f : \{0,1\}^n \to \mathbb{R}$.)

Claim 2.1. *If $f : \{0,1\}^n \to \{-1,1\}$ is such that $W_1(f) = 1$ then f is a dictator or antidictator.*

Proof. Since f is Boolean, f^2 is the constant 1 function, i.e., $\chi_\emptyset$. Therefore

$$\left(\sum_{i=1}^n \hat{f}(\{i\})^2\right)\chi_\emptyset + 2\sum_{i<j}\hat{f}(\{i\})\hat{f}(\{j\})\chi_{\{i,j\}} = \chi_\emptyset.$$

Hence for all $i < j$, we have $\hat{f}(\{i\})\hat{f}(\{j\}) = 0$ so at most one of the $\hat{f}(\{i\})$s is nonzeo. □

What can we say when $W_1(f) \geq 1 - \varepsilon$?

Theorem 2.2 ([FKN02]). *If $f : \{0,1\}^n \to \{-1,1\}$ has $\sum_{|S|>1} \hat{f}(S)^2 < \varepsilon$ then f is $O(\varepsilon)$ close to a 1-junta.*

This shows that NAE is a valid test for "dictator or antidictator" (think why). It also implies an approximate Arrow theorem: the only election functions having $1 - \varepsilon$ probability of a reasonable outcome are those close to a dictator or an antidictator.

Proof. First we notice that without loss of generality we can assume that $\hat{f}(\emptyset) = 0$ (a trick due to Guy Kindler). Indeed we can define $g : \{0,1\}^{n+1} \to \{-1,1\}$ by

$$g(x,y) = \begin{cases} f(x) & \text{if } y = 0 \\ -f(x + (1,\ldots,1)) & \text{if } y = 1. \end{cases}$$

This transformation sends χ_S for $S \subseteq [n]$ of odd size to itself and sends χ_S for $S \subseteq [n]$ of even size to $\chi_{S\cup\{n+1\}}$. In particular, we get $\hat{g}(\emptyset) = 0$ and $\sum_{|S|>1} \hat{g}(S)^2 < \varepsilon$. Moreover, if g is close to a 1-junta, then so is f. (Think: why don't we just take the odd part of f?)

Write $f = \ell + h$ with

$$\ell = \sum_{i=1}^{n} \hat{f}(\{i\})\chi_{\{i\}} \qquad \text{and} \qquad h = \sum_{|S|>1} \hat{f}(S)\chi_S.$$

Then,

$$1 = f^2 = \ell^2 + 2\ell h + h^2 = \ell^2 + h(2f - h).$$

Since $\mathbb{E}[h(x)^2] < \varepsilon$, $\Pr[|h(x)| \geq 10\sqrt{\varepsilon}] \leq 1/100$. Hence,

$$\Pr[|h(x) \cdot (2f(x) - h(x))| > 21\sqrt{\varepsilon}] < 1/100.$$

Moreover,

$$\ell^2 = \underbrace{\left(\sum_{i=1}^{n} \hat{f}(\{i\})^2\right)}_{\in(1-\varepsilon,1)} \chi_\emptyset + \underbrace{2\sum_{i<j} \hat{f}(\{i\})\hat{f}(\{j\})\chi_{\{i,j\}}}_{q}.$$

Therefore, $\Pr[|q(x)| > 11\sqrt{\varepsilon}] < 1/100$, or equivalently, $\Pr[q(x)^2 > 121\varepsilon] < 1/100$. By the hypercontractive inequality, $\mathbb{E}[q(x)^4] \leq 81(\mathbb{E}[q(x)^2])^2$. Using the following claim from the homework (with $X \leftarrow q^2$, $K \leftarrow 121\varepsilon$, $L \leftarrow \mathbb{E}[q^2]$, and $\delta \leq 1/100$), we get that $\mathbb{E}[q(x)^2] < 1000\varepsilon$ (since otherwise $\mathbb{E}[q^4] \geq 100(\mathbb{E}[q^2] - 121\varepsilon)^2 > 81(\mathbb{E}[q^2])^2$, in contradiction).

Claim 2.3. *If X is a random variable with $\Pr[X > K] = \delta$ and $\mathbb{E}[X] \geq L > K$ then $\mathbb{E}[X^2] \geq (L-K)^2/\delta$.*

Therefore,

$$\begin{aligned}
1000\varepsilon > \mathbb{E}[q(x)^2] &= \sum_{i<j} \hat{f}(\{i\})^2 \hat{f}(\{j\})^2 \\
&= \frac{1}{2}\left(\left(\sum_i \hat{f}(\{i\})^2\right)^2 - \sum_i \hat{f}(\{i\})^4\right) \\
&\geq \frac{1}{2}\left((1-\varepsilon)^2 - \sum_i \hat{f}(\{i\})^4\right),
\end{aligned}$$

which implies that $\sum_i \hat{f}(\{i\})^4 > 1 - 2002\varepsilon$. But since $\sum_i \hat{f}(\{i\})^4 \leq (\sum \hat{f}(\{i\})^2) \max \hat{f}(\{i\})^2$, we get that there exists an i such that $|\hat{f}(\{i\})| \geq 1 - 1002\varepsilon$. □

References

[FKN02] E. Friedgut, G. Kalai, and A. Naor. Boolean functions whose Fourier transform is concentrated on the first two levels. *Advances in Applied Mathematics*, 29:427–437, 2002.

[Fri98] E. Friedgut. Boolean functions with low average sensitivity depend on few coordinates. *Combinatorica*, 18(1):474–483, 1998.

[KKL88] J. Kahn, G. Kalai, and N. Linial. The influence of variables on boolean functions. In *Proceedings of the 29th Annual Symposium on Foundations of Computer Science*, pages 68–80. 1988.