The OSSS Inequality and the FKN Theorem (lecture notes)

Course: Analysis of Boolean Functions, Autumn 2025, University of Chicago
Instructor: William Hoza (`williamhoza@uchicago.edu`)

1 The O'Donnell-Saks-Schramm-Servedio inequality

Previously, we proved the KKL theorem, which says that for every Boolean function $f\colon \{\pm1\}^n \to \{\pm1\}$, there is a variable i such that $\mathrm{Inf}_i[f] \ge \Omega(\mathrm{Var}[f] \cdot \frac{\log n}{n})$. The KKL theorem is tight, as demonstrated by the Tribes function, but we can improve the KKL theorem if we make extra assumptions about f. We will prove that if f can be computed by a size-s decision tree, then there is a variable $i \in [n]$ such that $\mathrm{Inf}_i[f] \ge \mathrm{Var}[f]/\log s$. Equivalently, our goal is to prove that $\mathrm{Var}[f] \le (\log s) \cdot \max_i \mathrm{Inf}_i[f]$. For the sake of induction, we will actually prove an upper bound on the *covariance* between f and g, where f is a size-s decision tree and g has small influences.

Definition 1.1 (Covariance)**.** Let $f, g\colon \{\pm1\}^n \to \mathbb{R}$. We define

$$\mathrm{Cov}[f,g] = \mathbb{E}[fg] - \mathbb{E}[f] \cdot \mathbb{E}[g].$$

The proof relies on the expectation operator. Recall that for a function $f\colon \{\pm1\}^n \to \mathbb{R}$ and $i \in \mathbb{N}$, we define $E_i f = \sum_{S \not\ni i} \widehat{f}(S) \cdot \chi_S$. The following lemma provides a more intuitive interpretation of the expectation operator.

Lemma 1.2. *If $f\colon \{\pm1\}^n \to \mathbb{R}$ and $i \in \mathbb{N}$, then*

$$(E_i f)(x) = \mathop{\mathbb{E}}_{b \in \{\pm1\}}[f(x^{(i\mapsto b)}].$$

Proof.

$$\begin{aligned}(E_i f)(x) &= f(x) - x_i \cdot (D_i f)(x) \\ &= f(x) - x_i \cdot \frac{f(x^{(i\mapsto 1)}) - f(x^{(i\mapsto -1)})}{2} \\ &= \mathop{\mathbb{E}}_{b \in \{\pm1\}}[f(x^{(i\mapsto b)})].\end{aligned}$$

□

Lemma 1.3. *Let $f, g\colon \{\pm1\}^n \to \mathbb{R}$ and $i \in [n]$. For each $b \in \{\pm1\}$, let f_b, g_b denote the restrictions of f in which we fix $x_i = b$. Then*

$$\mathrm{Cov}[f,g] = \mathop{\mathbb{E}}_{b,b'}[\mathrm{Cov}[f_b, g_{b'}]] + \langle f, x_i \cdot D_i g\rangle. \tag{1}$$

Proof. Without loss of generality, assume $\mathbb{E}[f] = \mathbb{E}[g] = 0$. (None of the terms in Eq. (1) are affected if we shift f or g by an additive constant.) Then

$$\begin{aligned}\mathrm{Cov}[f,g] = \langle f, g\rangle &= \langle f, x_i \cdot D_i g\rangle + \langle f, E_i g\rangle \\ &= \langle f, x_i \cdot D_i g\rangle + \langle x_i \cdot D_i f, E_i g\rangle + \langle E_i f, E_i g\rangle \\ &= \langle f, x_i \cdot D_i g\rangle + \langle E_i f, E_i g\rangle \\ &= \langle f, x_i \cdot D_i g\rangle + \mathop{\mathbb{E}}_{b,b'}[\langle f_b, g_{b'}\rangle] \qquad \text{by Lemma 1.2.}\end{aligned}$$

Meanwhile,

$$\begin{aligned}
\mathop{\mathbb{E}}_{b,b'}[\mathrm{Cov}[f_b, g_{b'}]] &= \mathop{\mathbb{E}}_{b,b'}\left[\mathop{\mathbb{E}}_{x}[f_b(x)\cdot g_{b'}(x)] - \mathop{\mathbb{E}}_{x}[f_b(x)]\cdot \mathop{\mathbb{E}}_{x'}[g_{b'}(x')]\right] \\
&= \mathop{\mathbb{E}}_{b,b'}[\langle f_b, g_{b'}\rangle] - \mathop{\mathbb{E}}_{b,x}[f_b(x)]\cdot \mathop{\mathbb{E}}_{b',x'}[g_{b'}(x')] \\
&= \mathop{\mathbb{E}}_{b,b'}[\langle f_b, g_{b'}\rangle] - \mathbb{E}[f]\cdot\mathbb{E}[g] \\
&= \mathop{\mathbb{E}}_{b,b'}[\langle f_b, g_{b'}\rangle]. \qquad \square
\end{aligned}$$

Theorem 1.4 (OSSS Inequality)**.** *Let $f, g\colon \{\pm1\}^n \to \{\pm1\}$, let T be a decision tree computing f, and let $\delta_i(T)$ be the probability that T queries x_i when we plug in a uniform random x. Then*

$$\mathrm{Cov}[f, g] \le \sum_{i=1}^{n} \delta_i(T)\cdot \mathrm{Inf}_i[g].$$

Proof. We will prove it by induction on the depth of T. If T has depth zero, the theorem is trivial, so assume T has depth $D > 0$. Let i_* be the variable queried by the root. For each $b \in \{\pm1\}$, let f_b, g_b be the restrictions of f, g given by fixing $x_{i_*} = b$, and let T_b be the depth-$(D-1)$ subtree of T computing f_b. Then

$$\begin{aligned}
\mathrm{Cov}[f, g] &= \mathop{\mathbb{E}}_{b,b'}[\mathrm{Cov}[f_b, g_{b'}]] + \langle f, x_{i_*}\cdot D_{i_*}g\rangle \\
&\le \mathop{\mathbb{E}}_{b,b'}\left[\sum_{i\ne i_*} \delta_i(T_b)\cdot \mathrm{Inf}_i[g_{b'}]\right] + \mathop{\mathbb{E}}_{x}[|D_{i_*}g|] \\
&= \sum_{i\ne i_*} \mathop{\mathbb{E}}_{b}[\delta_i(T_b)]\cdot \mathop{\mathbb{E}}_{b'}[\mathrm{Inf}_i[g_{b'}]] + \mathrm{Inf}_{i_*}[g] \\
&= \sum_{i\ne i_*} \delta_i(T)\cdot \mathrm{Inf}_i[g] + \mathrm{Inf}_{i_*}[g] \\
&= \sum_{i=1}^{n} \delta_i(T)\cdot \mathrm{Inf}_i[g]. \qquad \square
\end{aligned}$$

Corollary 1.5. *Let $f\colon \{\pm1\}^n \to \{\pm1\}$ be a size-s decision tree. Then there is some $i \in [n]$ such that $\mathrm{Inf}_i[f] \ge \mathrm{Var}[f]/\log s$.*

Proof. Let T be the decision tree. By the OSSS inequality, we have

$$\begin{aligned}
\mathrm{Var}[f] = \mathrm{Cov}[f, f] \le \sum_{i=1}^{n} \delta_i(T)\cdot \mathrm{Inf}_i[f] &\le \left(\max_i \mathrm{Inf}_i[f]\right)\cdot \sum_{i=1}^{n}\delta_i(T) \\
&= \left(\max_i \mathrm{Inf}_i[f]\right)\cdot \mathop{\mathbb{E}}_{x}[\mathrm{cost}_T(x)] \\
&\le \left(\max_i \mathrm{Inf}_i[f]\right)\cdot \log s. \qquad \square
\end{aligned}$$

2 The Friedgut-Kalai-Naor theorem

Let $f\colon \{\pm1\}^n \to \{\pm1\}$. Recall that Condorcet's paradox is the situation $f(x^{ab}) = f(x^{bc}) = f(x^{ca})$ for some $x \in S_3^n$. Recall that Arrow's theorem say that if Condorcet's paradox *never* happens, then f or $-f$ is a dictator function. In this section, as another application of hypercontractivity, we will prove a robust version of Arrow's theorem, saying that if Condorcet's paradox *rarely* happens, then f or $-f$ is *close* to a dictator function.

The key is the Friedgut-Kalai-Naor (FKN) theorem. Recall that in the proof of Arrow's theorem, we showed that if the probability of the Condorcet paradox is at most ε, then $W^1[f] \geq 1 - O(\varepsilon)$. The FKN theorem says that this condition implies that f or $-f$ is $O(\varepsilon)$-close to a dictator.

Theorem 2.1 (Friedgut-Kalai-Naor)**.** *Let $f\colon \{\pm 1\}^n \to \{\pm 1\}$. Suppose $W^1[f] \geq 1 - \varepsilon$. Then there is some $i \in [n]$ and some $b \in \{\pm 1\}$ such that f is $O(\varepsilon)$-close to $b\chi_i$.*

Proof. Our goal is to show that there is some i such that $|\widehat{f}(i)| \approx 1$. We have

$$\begin{aligned}\max_i \widehat{f}(i)^2 \geq \left(\max_i \widehat{f}(i)^2\right) \cdot \sum_{i=1}^n \widehat{f}(i)^2 \geq \sum_{i=1}^n \widehat{f}(i)^4 &= \left(\sum_{i=1}^n \widehat{f}(i)^2\right)^2 - 2 \sum_{1 \leq i < j \leq n} \widehat{f}(i)^2 \cdot \widehat{f}(j)^2 \\ &\geq (1-\varepsilon)^2 - 2 \sum_{1 \leq i < j \leq n} \widehat{f}(i)^2 \cdot \widehat{f}(j)^2.\end{aligned}$$

So we would like to show that $\sum_{1 \leq i < j \leq n} \widehat{f}(i)^2 \cdot \widehat{f}(j)^2$ is small. Define

$$h(x) = 2 \sum_{1 \leq i < j \leq n} \widehat{f}(i) \cdot \widehat{f}(j) \cdot x_i x_j,$$

so our goal is to bound $\|h\|_2^2$. Recall that we used hypercontractivity to prove $\|h\|_2 \leq 2^{O(\deg(h))} \cdot \|h\|_1$. In our case, $\deg(h) = 2$, so $\|h\|_2 = O(\|h\|_1)$, and our new goal is to bound $\|h\|_1$.

Define $\ell(x) = \sum_{i=1}^n \widehat{f}(i) \cdot x_i$. Then $h = \ell^2 - \mathbb{E}[\ell^2]$, so

$$\begin{aligned}\|h\|_1 = \|\ell^2 - \mathbb{E}[\ell^2]\|_1 &\leq \|\ell^2 - f^2\|_1 + \|f^2 - \mathbb{E}[\ell^2]\|_1 \\ &= \|(\ell - f) \cdot (\ell + f)\|_1 + |1 - W^1[f]| \\ &\leq \|\ell - f\|_2 \cdot \|\ell + f\|_2 + \varepsilon \\ &\leq \sqrt{\varepsilon} \cdot 2 + \varepsilon \\ &\leq O(\sqrt{\varepsilon}).\end{aligned}$$

Putting everything together, we get $\max_i \widehat{f}(i)^2 \geq (1-\varepsilon)^2 - O(\varepsilon) = 1 - O(\varepsilon)$, hence $\max_i |\widehat{f}(i)| \geq 1 - O(\varepsilon)$. $\square$

Corollary 2.2 (Robust Arrow's theorem)**.** *Let $f\colon \{\pm 1\}^n \to \{\pm 1\}$. Suppose*

$$\Pr_{x \in S_3^n}[f(x^{ab}) = f(x^{bc}) = f(x^{ca})] \leq \varepsilon.$$

Then f is $O(\varepsilon)$-close to either χ_i or $-\chi_i$ for some $i \in [n]$.