Special Topics in Complexity Theory, Fall 2017. Instructor: Emanuele Viola

1 Xuangui Huang's presentation, Scribe: Chin Ho Lee

1.1 Robustifying polynomials

In this lecture, we show how to make a polynomial robust to noise by proving the following theorem by Sherstov [She13].

Theorem 1. Let $p\colon \{-1,1\}^n \to [-1,1]$ be a degree-d polynomial. There exists an explicit degree-$O(d)$ polynomial $\tilde{p}\colon \mathbb{R}^n \to \mathbb{R}$ such that for every $x \in X^n$, where $X = [-4/3,-2/3] \cup [2/3,4/3]$,

$$|p(\operatorname{sgn}(x_1), \operatorname{sgn}(x_2), \dots, \operatorname{sgn}(x_n)) - \tilde{p}(x)| \le 2^{-\Omega(d)}.$$

We will prove Theorem 1 in 3 steps: where (1) p is a monomial, (2) p is a homogeneous polynomial of degree d, i.e., every monomial of p has degree exactly d, and (3) p is a general polynomial. We first prove (1), then prove (3) assuming (2), and defer the proof of (2) to the end.

1.2 Monomial

Let us now consider the case when $p(x) := \prod_{i=1}^d x_j$ is the parity function. We will use the following Taylor's expansion of the function $(1+t)^\alpha$.

Claim 2. For every $t \in (-1,1)$ and $\alpha \in \mathbb{R}$, we have $(1+t)^\alpha = \sum_{i=0}^\infty \binom{\alpha}{i} t^i$, where $\binom{\alpha}{i} := \frac{\alpha(\alpha-1)\cdots(\alpha-i+1)}{1\cdot 2\cdots i}$ is the extension of the binomial coefficients to the real numbers.

Using Claim 2, we obtain the follow Taylor's expansion for $\operatorname{sgn}(t)$.

Claim 3. For $0 < |t| < \sqrt{2}$, $\operatorname{sgn}(t) = t \sum_{i=0}^\infty \binom{-1/2}{i} (t^2-1)^i$.

Proof.

$$\operatorname{sgn}(t) = \frac{t}{\sqrt{t^2}} = \frac{t}{\sqrt{(1+(t^2-1))}} = t\sum_{i=0}^\infty \binom{-1/2}{i}(t^2-1)^i.$$

□

We can now derive the Taylor approximation of $\prod_{j=1}^{d} \operatorname{sgn}(x_j)$:

$$\begin{aligned}\prod_{j=1}^{d} \operatorname{sgn}(x_j) &= \prod_{j=1}^{d} \left(x_j \sum_{i=0}^{\infty} \binom{-1/2}{i} (x_j^2 - 1)^i \right) \\ &= \left(\prod_{j=1}^{d} x_j \right) \sum_{0 \le i_1, \ldots, i_d \le \infty} \prod_{j=1}^{d} \binom{-1/2}{i_j} (x_j^2 - 1)^{i_j}.\end{aligned}$$

We now define $\tilde{p}$. Let $d' = Cd$ for a sufficiently large constant C. We define $\tilde{p} \colon \mathbb{R}^n \to \mathbb{R}$ to be the truncation of the above infinite series up to the indices that satisfy $i_1 + \cdots + i_d \le d'$, that is,

$$\tilde{p}(x_1, \ldots, x_d) := \left(\prod_{j=1}^{d} x_j \right) \sum_{i_1 + \cdots + i_d \le d'} \prod_{j=1}^{d} \binom{-1/2}{i_j} (x_j^2 - 1)^{i_j}.$$

Clearly, $\tilde{p}$ has degree $d + 2d' = O(d)$. It remains to analyze the approximation error. First we need a simple bound for $\binom{-1/2}{i_j}$.

Claim 4. For every $k \ge 1$, $\binom{-1/2}{k} = (-4)^{-k} \binom{2k}{k} \le 1/2$.

Proof. By definition,

$$\begin{aligned}\binom{-1/2}{k} &= \frac{(-1/2) \cdot (-3/2) \cdots (-1/2 - k + 1)}{k!} \\ &= 2^{-k} \cdot \frac{1 \cdot 3 \cdots (2k-1)}{k!} \\ &= 2^{-k} \cdot \frac{1}{2^k k!} \cdot \frac{(2k)!}{k!} \\ &= (-4)^{-k} \binom{2k}{k}.\end{aligned}$$

The inequality follows from $\binom{2k}{k} \le 2^{2k}/2$. $\square$

Note that the approximation error $\delta(x) := \prod_{j=1}^{d} \operatorname{sgn}(x_j) - \tilde{p}(x_1, \ldots, x_d)$ is simply the remaining sum in the infinite series after the truncation, that is

$$\delta(x) = \left(\prod_{j=1}^{d} x_j \right) \sum_{i_1 + \cdots + i_d > d'} \prod_{j=1}^{d} \binom{-1/2}{i_j} (x_j^2 - 1)^{i_j}. \tag{1}$$

The R.H.S. is at most

$$\Big|\prod_{j=1}^{d} x_j\Big| \cdot \Big| \sum_{i_1+\cdots+i_d>d'} \prod_{j=1}^{d} \binom{-1/2}{i_j} (x_j^2-1)^{i_j}\Big| \le (4/3)^d \sum_{i_1+\cdots+i_d>d'} \prod_{j=1}^{d} \binom{-1/2}{i_j} |x_j^2-1|^{i_j}$$
$$\le (4/3)^d \cdot (1/2)^d \sum_{i_1+\cdots+i_d>d'} \prod_{j=1}^{d} |x_j^2-1|^{i_j}$$
$$\le \sum_{i_1+\cdots+i_d>d'} (7/9)^{i_1+\cdots+i_d},$$

The first inequality is because $|x_j| \le 4/3$ for $x \in X$. The second inequality is because of Claim 4, and the last inequality is because $|x_j^2 - 1| \le 7/9$ for $x \in X$.

Now, for every k, there are $\binom{k+d-1}{d}$ choices of $i_1, \dots, i_d$ for which $i_1 + \cdots + i_d = k$. Hence, the summation is equal to

$$\sum_{k=d'+1}^{\infty} \sum_{i_1+\cdots+i_d=k} (7/9)^k = \sum_{k=d'+1}^{\infty} \binom{k+d-1}{k} (7/9)^k$$
$$\le \sum_{k=d'+1}^{\infty} (2k)^d (7/9)^k$$
$$= 2^{-\Omega(d')}.$$

This finishes the proof for the case when p is a monomial.

1.3 General case assuming homogeneous case

We now prove Theorem 1 assuming the same conclusion holds for case (2), when p is a homogeneous polynomial.

First we can rewrite p as $p = \sum_{i=0}^{d} p_i$, where p_i is the degree-i homogeneous polynomial of p. Note that while p is bounded by 1, p_i may not be. So, we instead apply Theorem 1 to $p_i/\|p_i\|_\infty$, where $\|p_i\|_\infty := \max_{x\in\{-1,1\}} |p_i(x)|$, and obtain $\tilde{p}_i$ such that

$$\max_{x\in X^n} |\tilde{p}_i(x) - p_i(\mathrm{sgn}(x_1), \mathrm{sgn}(x_2), \dots, \mathrm{sgn}(x_n))| \le \|p_i\|_\infty \cdot 2^{-\Omega(d)}.$$

If we assume $\sum_{i=0}^{d}\|p_i\|_\infty \leq 2^{O(d)}$ and define $\tilde{p} := \sum_{i=0}^{d}\tilde{p}_i$, then we have

$$\begin{aligned}
|p(\mathrm{sgn}(x_1),\ldots,\mathrm{sgn}(x_n)) - \tilde{p}(x)| &\leq \sum_{i=0}^{d} |p_i(\mathrm{sgn}(x_1),\ldots,\mathrm{sgn}(x_n)) - \tilde{p}(x)| \\
&\leq \sum_{i=0}^{d}\|p_i\|_\infty \cdot 2^{-\Omega(d)} \\
&\leq (d+1)\cdot 4^d \cdot 2^{-\Omega(d)} \\
&\leq 2^{-\Omega(d)}.
\end{aligned}$$

We now prove that $\sum_{i=0}^{d}\|p_i\|_\infty \leq 2^{O(d)}$ whenever p has output $[-1,1]$. We first prove the result for *univariate* polynomials and then reduce the above problem to it. The univariate version in fact follows by a theorem due to Vladimir Markov which gives a tight upper bound [**?**]:

Theorem 5. If $p\colon [-1,1] \to [-1,1]$ is a univariate degree-d polynomial, then the sum of its $d+1$ coefficients in absolute values is bounded by $O((1+\sqrt{2})^d/\sqrt{d})$.

We now prove the theorem above with the upper bound replaced by the crude bound of $2^{O(d)}$, which is sufficient for our purpose.

Claim 6. If $p\colon [-1,1] \to [-1,1]$ is a univariate degree-d polynomial, then the sum of its coefficients in absolute values is at most $2^{O(d)}$.

Proof. Let $t_0, t_1, \ldots, t_d$ be the $d+1$ points that are evenly spaced in the interval $[-1,1]$, so $t_i := -1 + 2i/t$. By interpolation, we can write p as

$$p(t) = \sum_{i=0}^{d} p(t_i)\frac{\prod_{j\neq i}(t-t_j)}{\prod_{j\neq i}(t_i - t_j)}.$$

We first bound below $\prod_{j\neq i}^{d}(t_i - t_j)$. Since every distinct pair t_i and t_j differ by $2/d$, This product is smallest when t_i is closest to 0, and so is at least $(2/d)^d(\frac{d}{2})!^2$ when d is even and is at least $(2/d)^d(\frac{d+1}{2})(\frac{d-1}{2})!^2$ when d is odd. By Stirling's formula, in both cases we have

$$\prod_{j\neq i}^{d}(t_i - t_j) \geq (2/d)^d(d/2e)^d \geq e^{-d}.$$

Hence the sum of the coefficients in absolute values is at most

$$e^d \sum_{i=0}^{d} \prod_{j \neq i} (1 + |t_j|) \leq (d+1)(2e)^d \leq 2^{O(d)}.$$

$\square$

We now bound above $\sum_{i=0}^{d} \|p_i\|_\infty$ by a reduction to Claim 6.

Claim 7. $\sum_{i=0}^{d} \|p_i\|_\infty \leq 2^{O(d)}$.

Proof. Fix any $x \in \{-1,1\}^n$. Define the univariate polynomial $q_x \colon [-1,1] \to [-1,1]$ by $q_x(t) := \sum_{i=0}^{d} p_i(x) \cdot t^i$. We will show that $|q_x(t)| \leq 1$ for every $x \in \{-1,1\}^n$. Then the rest simply follows from Claim 6.

Let $Z = (Z_1, \ldots, Z_n) \in \{-1,1\}^n$ be independent random variables with $\mathbb{E}[Z_i] = t$. Write p in its Fourier expansion $p(x) = \sum_{|S| \leq d} \hat{p}(S) \prod_{i \in S} x_i$. We have

$$\begin{aligned}
\mathbb{E}_Z[p(x_1 Z_1, \ldots, x_n Z_n)] &= \mathbb{E}_Z\Big[\sum_{|S| \leq d} \hat{p}(S) \prod_{i \in S} x_i Z_i\Big] \\
&= \sum_{|S| \leq d} \hat{p}(S) \prod_{i \in S} x_i \cdot \prod_{i \in S} \mathbb{E}_Z[Z_i] \\
&= \sum_{|S| \leq d} \hat{p}(S) \prod_{i \in S} x_i \cdot t^{|S|} \\
&= \sum_{i=0}^{d} p_i(x) t^i \\
&= q_x(t).
\end{aligned}$$

This shows $|q_x(t)| \leq 1$ as the L.H.S. is at most $\max_{y \in \{-1,1\}^n} |p(y)| \leq 1$. $\square$

1.4 Homogeneous polynomial

Let $p \colon \{-1,1\}^n \to [-1,1]$ be a homogeneous polynomial of degree d. We can write p as $p(x) = \sum_{|S|=d} \hat{p}(S) \chi_S(x)$, where $\chi_S(x) := \prod_{j \in S} x_j$. In this way we can regard p as a function from $\mathbb{R}^n$ to $\mathbb{R}$. We will apply the robustification in the monomial case to each χ_S. More specifically, we define $\tilde{p}$ to be $\tilde{p}(x) :=$

$\sum_{|S|=d} \hat{p}(S)\tilde{\chi}_S(x)$. Let $\delta(x_S)$ be the approximation error of $\tilde{\chi}_S$, i.e., the expression in Equation (1). Then $\forall x \in X^n$,

$$\begin{aligned}|p(\operatorname{sgn}(x_1), \operatorname{sgn}(x_2), \ldots, \operatorname{sgn}(x_n)) - \tilde{p}(x)| &= \Big| \sum_{|S|=d} \hat{p}(S)\big(\prod_{j\in S} \operatorname{sgn}(x_j) - \prod_{j\in S} x_j\big)\Big| \\ &= \Big| \sum_{|S|=d} \hat{p}(S)\delta(x_S)\Big|.\end{aligned}$$

Therefore to prove Theorem 1 in the homogeneous case we need to show $\max_{x\in X^n} \Big| \sum_{|S|=d} \hat{p}(S)\delta(x_S)\Big| \le 2^{-\Omega(d)}$.

We first show that one cannot get anything just by naïvely summing up all the error $\delta(x_S)$ for each S.

Claim 8. There exists a homogeneous degree-d polynomial $p\colon \{-1,1\}^n \to [-1,1]$ such that $\hat{p}(S) = \pm(2n\binom{n}{d})^{-1/2}$.

The error of $\tilde{p}$ for the polynomial p in the claim would be $\sum_{|S|=d} |\hat{p}(S)| \cdot 2^{-\Omega(d)} = \binom{n}{d}(2n\binom{n}{d})^{-1/2} \cdot 2^{-\Omega(d)} > 1$.

1.4.1 Error cancellation

We now do a more refined analysis on the error by proving the following theorem, showing that the errors in different terms in fact cancel out each other.

Theorem 9.(Warm-up) Let $p\colon \{-1,1\}^n \to [-1,1]$ be a homogeneous degree-d polynomial. Let $\delta\colon \{-1,1\}^d \to \mathbb{R}$ be a symmetric function. Then

$$\max_{x\in\{-1,1\}^n} \Big| \sum_{|S|=d} \hat{p}(S)\delta(x_S)\Big| \le \frac{d^d}{d!}\|\hat{\delta}\|_1,$$

where $\|\hat{\delta}\|_1 = \sum_S |\hat{\delta}(S)|$ is the sum of the magnitude of the coefficients in the Fourier expansion of $\delta(x) = \sum_S \hat{\delta}(S)\prod_{j\in S} x_j$.

For the specific δ given in Equation (1) we have $\|\hat{\delta}\|_1 \le 2^{-Cd}$. Hence the maximum error is $d^d/d! \cdot 2^{-Cd} \le 2^{-\Omega(d)}$ for a sufficiently large constant C.

But this is only a warm-up theorem: the maximum is taken over $\{-1,1\}^n$ instead of X^n. At the end we will briefly mention the changes required to prove Theorem 1 in the homogeneous case.

The crucial tool in proving Theorem 9 is the following operator.

Definition 10. For every $v \in \{0,1\}^d$, we define the operator $A_v \colon \mathbb{R}^{\{-1,1\}^n} \to \mathbb{R}^{\{-1,1\}^n}$ by

$$(A_v f)(x) = \mathbb{E}_{z \sim \{-1,1\}^d} \left[z_1 \cdots z_d \, f\Big(\frac{1}{d} \sum_{i=1}^{d} z_i x_1^{v_i}, \dots, \frac{1}{d} \sum_{i=1}^{d} z_i x_n^{v_i} \Big) \right].$$

Note that we can identify f with its multilinear extension on $[-1,1]^n$ using its Fourier expansion so the term "$f\Big(\frac{1}{d}\sum_{i=1}^d z_i x_1^{v_i}, \dots, \frac{1}{d}\sum_{i=1}^d z_i x_n^{v_i}\Big)$" makes sense. We will use the following properties of A_v.

Claim 11. The operator A_v is

(1) linear;
(2) for every f we have $\|A_v f\|_\infty \le \|f\|_\infty$, and
(3) for every subset $S \subseteq \{1,\dots,n\}$ of size d,

$$A_v \chi_S(x) = \frac{d!}{d^d} \cdot \mathbb{E}_{\tau\colon S \to \{1,\dots d\} \text{ bijective}} \Big[\prod_{j \in S} x_j^{v_{\tau(j)}} \Big].$$

Proof. (1) is clear.

For (2), we have for every $x \in \{-1,1\}^n$,

$$\begin{aligned} |(A_v f)(x)| &= \left| \mathbb{E}_{z \sim \{-1,1\}^d} \left[z_1 \cdots z_d \, f\Big(\frac{1}{d} \sum_{i=1}^{d} z_i x_1^{v_i}, \dots, \frac{1}{d} \sum_{i=1}^{d} z_i x_n^{v_i} \Big) \right] \right| \\ &\le \mathbb{E}_{z \sim \{-1,1\}^d} \left[\left| f\Big(\frac{1}{d} \sum_{i=1}^{d} z_i x_1^{v_i}, \dots, \frac{1}{d} \sum_{i=1}^{d} z_i x_n^{v_i} \Big) \right| \right] \\ &\le \max_{x \in [-1,1]^n} |f(x)|. \end{aligned}$$

It remains to show that $\max_{x \in [-1,1]^n} f(x) \le \max_{x \in \{-1,1\}^n} f(x)$. This follows from the following claim, which says for *multilinear* polynomials, the maximum value can always be attained in $\{-1,1\}^n$.

Claim 12. Let $p\colon [-1,1]^n \to [-1,1]$ be any multilinear polynomial. Then $\max_{x \in [-1,1]^n} |p(x)| = \max_{x \in \{-1,1\}} |p(x)|$.

Proof. It suffices to show that $\max_{x\in[-1,1]^n} |p(x)| \le \max_{x\in\{-1,1\}} |p(x)|$. Fix any $x = (x_1, \ldots, x_n) \in [-1,1]^n$. Let $X = (X_1, \ldots, X_n) \in \{-1,1\}^n$ be n independent random variables with $\mathbb{E}[X_i] = x_i$ for each $i \in \{1, 2, \ldots, n\}$. Since p is multilinear, we have that $\mathbb{E}[p(X)] = p(x)$. Hence there exists a fixing of $X \in \{-1,1\}^n$ such that $p(x) \le p(X)$. $\square$

For (3), without loss of generality assume $S = \{1, \ldots, d\}$. Then

$$\begin{aligned} A_v \chi_S(x) &= \mathbb{E}_{z\in\{-1,1\}^d} \left[z_1 \cdots z_d \prod_{j=1}^{d} \Big(\frac{1}{d} \sum_{i=1}^{d} z_i x_j^{v_i} \Big) \right] \\ &= \frac{1}{d^d} \cdot \mathbb{E}_{z\in\{-1,1\}^d} \left[z_1 \cdots z_d \sum_{1\le i_1,\ldots,i_d \le d} z_{i_1} \cdots z_{i_d} \cdot \prod_{j=1}^{d} x_j^{v_{i_j}} \right]. \end{aligned}$$

If some z_k does not appear in the product $z_{i_1} \cdots z_{i_d}$, then we can factor out $E[z_k]$ from the expression and so the whole summand is zero. Hence the summation only contains terms that are distinct, i.e., $z_{i_j} = z_{\tau(j)}$ for some permutation τ. So the expression becomes

$$\begin{aligned} &\frac{1}{d^d} \cdot \mathbb{E}_{z\in\{-1,1\}^d} \Big[z_1 \cdots z_d \sum_{\tau \text{ bijective}} z_{\tau(1)} \cdots z_{\tau(d)} \cdot \prod_{j=1}^{d} x_j^{v_{\tau(j)}} \Big] \\ &= \frac{1}{d^d} \sum_{\tau \text{ bijective}} \prod_{j=1}^{d} x_j^{v_{\tau(j)}} \\ &= \frac{d!}{d^d} \cdot \mathbb{E}_{\tau \text{ bijective}} \Big[\prod_{j=1}^{d} x_j^{v_{\tau(j)}} \Big], \end{aligned}$$

where the first equality is because each $z_i \in \{-1,1\}$ appears twice and $z_i^2 = 1$. $\square$

We now prove Theorem 9.

Proof of Theorem 9. First we apply Claim 11 (3) with $v = 1^k 0^{d-k}$. We have

$$\frac{d^d}{d!} \cdot A_{1^k 0^{d-k}} \chi_S(x) = \mathbb{E}_{\tau \text{ bijective}} \Big[\prod_{j\in S} x_j^{v_{\tau(j)}} \Big] = \frac{1}{\binom{d}{k}} \sum_{T\subseteq S : |T| = k} \chi_T(x).$$

Because δ is symmetric, the coefficients $\hat{\delta}(T)$ are equal for subsets T of the same size. So,

$$\sum_{k=0}^{d} \hat{\delta}(\{1,\ldots,k\}) \sum_{T\subseteq S:|T|=k} \chi_T(x) = \sum_{k=0}^{d} \hat{\delta}(\{1,\ldots,k\})\binom{d}{k} \cdot \frac{d^d}{d!} A_{1^k 0^{d-k}} \chi_S(x).$$

Hence we can express the error term as

$$\begin{aligned}
\sum_{|S|=d} \hat{p}(S)\delta(x_S) &= \sum_{|S|=d} \hat{p}(S) \sum_{k=0}^{d} \binom{d}{k} \hat{\delta}(\{1,\ldots,k\}) \sum_{T\subseteq S,|S|=k} \chi_T(x) \\
&= \sum_{|S|=d} \hat{p}(S) \sum_{k=0}^{d} \binom{d}{k} \hat{\delta}(\{1,\ldots,k\}) \cdot \frac{d^d}{d!} \cdot A_{1^k 0^{d-k}} \chi_S(x) \\
&= \frac{d^d}{d!} \sum_{k=0}^{d} \binom{d}{k} \hat{\delta}(\{1,\ldots,k\}) \cdot A_{1^k 0^{d-k}} \Big(\sum_{|S|=d} \hat{p}(S)\chi_S(x) \Big) \\
&= \frac{d^d}{d!} \sum_{k=0}^{d} \binom{d}{k} \hat{\delta}(\{1,\ldots,k\}) \cdot A_{1^k 0^{d-k}} p(x).
\end{aligned}$$

where the last equality is because $A_{1^k 0^{d-k}}$ is linear. Since $\|A_v p\|_\infty \leq \|p\|_\infty \leq 1$, we have

$$\Big| \sum_{|S|=d} \hat{p}(S)\delta(x_S) \Big| \leq \frac{d^d}{d!} \|\hat{\delta}\|_1.$$

$\square$

To generalize the proof to real-valued inputs X'^n, where $X' = [-1.1, -0.9] \cup [0.9, 1.1]$. In the definition of the operator A_v, we replace $v \in \{0,1\}^d$ with $v \in \mathbb{N}^d$, and the j-th argument of the input for f becomes

$$\frac{1}{d} \sum_{i=1}^{d} z_i x_j (x_j^2 - 1)^{v_i} \cdot 4^{v_i}.$$

This term is bounded by 1 in absolute value for $x \in X'^n$, hence Property (2) in Claim 11 still holds. Finally, Property (3) in Claim 11 becomes

$$A_v \chi_S(x) = \frac{d!}{d^d} \mathbb{E}_{\tau:\, S \to \{1,\ldots,d\} \text{ bijective}} \Big[\prod_{j\in S} x_j (x_j^2 - 1)^{v_{\tau(j)}} \Big] \cdot 4^{v_1 + \cdots + v_d}.$$

Similarly, for the specific δ in Equation (1) we can prove

$$\begin{aligned}\sum_{|S|=d} \hat{p}(S)\delta(x_S) &= \sum_{|S|=d} \hat{p}(S) \sum_{v_1+\cdots+v_d>d'} \binom{-1/2}{v_1} \cdots \binom{-1/2}{v_d} 4^{-(v_1+\cdots+v_d)} \frac{d^d}{d!} A_v \chi_S(x) \\ &= \sum_{v_1+\cdots+v_d>d'} \binom{-1/2}{v_1} \cdots \binom{-1/2}{v_d} 4^{-(v_1+\cdots+v_d)} \frac{d^d}{d!} A_v p(x),\end{aligned}$$

which can be bounded by $2^{-\Omega(d)}$ given $d' = C \cdot d$ for sufficiently large C.

References

[She13] Alexander A. Sherstov. Making polynomials robust to noise. *Theory of Computing*, 2013.