Lecture 6: Local list decoding of polynomial codes

Topics in Pseudorandomness and Complexity Theory (Spring 2017)
Rutgers University
Swastik Kopparty
Scribe: Justin Semonsen

1 Recap

Previous classes we discussed an algorithm that, given a circuit C such that $\Pr_{x \in \mathbb{F}_q^n}[C(x) = g(x)] > 0.9$ for g a polynomial in $\mathbb{F}_q$ of degree $< d$ in each of m variables, we can produce a circuit $FIX(C)$ such that $\forall x \in \mathbb{F}_q^n, C(x) = g(x)$ and $size(FIX(C)) \leq \text{poly}(m) size(C)$.

In addition, we discussed the singe variable polynomial decoder algorithm, which works as follows: Given $S \subset \mathbb{F}_q^2$ with $|S| = n$, output all polynomials $\{Q(t)\}$ of degree $< d$ such that $\forall\, Q, |\{t : (t, Q(t)) \in S\}| \geq \epsilon n$. Note that this requires that $d = O(\epsilon^2 n)$ and ensures $|\{Q(t)\}| \leq \frac{2}{\epsilon}$.

In the upcoming proofs, we also use the low degree polynomial extension method discussed in a previous class. For simplicity, we define this extension of a function f to be $LDPE(f)$.

2 Correcting circuits that compute low degree polynomials in multiple variables

Theorem 1. *Suppose $g \in \mathbb{F}_q[x_1, \ldots x_m]$ is a polynomial of degree $< d$ in each variable, and C is a circuit of size s such that $\Pr_{x \in \mathbb{F}_q^n}[C(x) = g(x)] > \epsilon$. Then, using randomness, we can efficiently produce circuits $C_1, \ldots C_L$ of size $\text{poly}(q, m, \epsilon^{-1})$ (with $L = \text{poly}(q, m, \epsilon^{-1})$) such that with high probability $\exists i \in [L]$ such that $\forall\, x \in \mathbb{F}_q^m, C_i(x) = g(x)$.*

Proof. We define the function *pre-*$C_{y,a}(x)$ as follows:

1. Let l be the line $\{x + t(y - x) : t \in \mathbb{F}_q\}$, which is the line through x and y.
2. Using C, compute the set $S = \{(t, C(x + t(y - x))) : t \in \mathbb{F}_q \setminus \{0\}\}$.
3. Using the polynomial decoder algorithm, find $\{Q(t)\}$ of degree $< d$ with $|\{t : (t, Q(t)) \in S\}| \geq \epsilon q$.
4. If $\exists!\, Q^*(t) \in \{Q(t)\}$ such that $Q^*(1) = a$, output $Q^*(0)$. Otherwise, error.

Claim 2.

$$\Pr_{x,y \in \mathbb{F}_q^m}[\textit{pre-}C_{y,a}(x) = g(x)] \geq .99$$

The proof for this claim will follow in the next section.

Lemma 3.

$$\Pr_{y\in\mathbb{F}_q^m}\left[\Pr_{x\in\mathbb{F}_q^m}[\textit{pre-}C_{y,a}(x)=g(x)]\geq .9\right]\geq .99$$

Proof. Assume $\Pr_{y\in\mathbb{F}_q^m}\left[\Pr_{x\in\mathbb{F}_q^m}[\textit{pre-}C_{y,a}(x)=g(x)]\geq .9\right]<.9$.

By linearity of expectation, $\mathbb{E}_{y\in\mathbb{F}_q^m}\left[\mathbb{E}_{x\in\mathbb{F}_q^m}[1_{\{\textit{pre-}C_{y,a}(x)=g(x)\}}]\right]<.9+.9*.1=.99$.

However, from Claim 2, $\mathbb{E}_{x,y\in\mathbb{F}_q^m}[1_{\{\textit{pre-}C_{y,a}(x)=g(x)\}}]\geq .99$, proving the lemma by contradiction. □

This means that with high probability, $\Pr_{x\in\mathbb{F}_q^m}[\textit{pre-}C_{y,a}(x)=g(x)]\geq .9$, meaning that (with high probability) $\forall\, x\in\mathbb{F}_q^m, FIX(\textit{pre-}C_{y,a})(x)=g(x)$. Let $C_{y,a}$ be defined to be $FIX(\textit{pre-}C_{y,a})$.

Finally, this allows us to define the algorithm for the theorem as follows:

1. Pick $y\in\mathbb{F}_q^m$ uniformly at random.
2. Output $\{C_{y,a}:a\in\mathbb{F}_q\}$.

The list output by this algorithm then clearly is of the appropriate size, and contains a circuit that evaluates $g(x)$ for every $x\in\mathbb{F}_q^m$. □

3 Proof of Claim 2

Proof. Let $A=\{x\in\mathbb{F}_q^m:C(x)=g(x)\}$. Note that $|A|=q^m\Pr_{x\in\mathbb{F}_q^m}[C(x)=g(x)]\geq\epsilon q^m$.

Pick $x,y\in\mathbb{F}_q^m$ uniformly (and distinctly) at random. Let l be the line $\{x+t(y-x):t\in\mathbb{F}_q\}$, which is the line through x and y.

Note that for fixed $t_1\neq t_2$, $\Pr[x(1-t_1)+t_1y=a\wedge x(1-t_2)+t_2y=b]=\Pr[x=\frac{1}{t_2-t_1}(t_2a-t_1b)\wedge y=\frac{1}{t_2-t_1}(-(1-t_2)a+(1-t_1)b)]$. Since all of t_1,t_2,a, and b are fixed, and x and y are chosen independently, $\Pr[x(1-t_1)+t_1y=a\wedge x(1-t_2)+t_2y=b]=\Pr[x=\frac{1}{t_2-t_1}(t_2a-t_1b)]\Pr[y=\frac{1}{t_2-t_1}(-(1-t_2)a+(1-t_1)b)]$, and thus the points on the line are pairwise independent.

Since the points x and y are chosen uniformly at random, $x+t(y-x)$ is distributed uniformly at random and thus $Pr(x+t(y-x)\in A)\geq\epsilon$. This means that if we define the events $E_t=\{x+t(y-x)\in A\}$, these events are pairwise independent with probabilities $\geq\epsilon$.

Therefore, using that the variance of a Bernoulli random variable with probability p is $p(1-p)\leq p$, by Markov:

$$\begin{aligned}
\Pr\left[\sum_{t\in\mathbb{F}_q} 1_{E_t} < \frac{\epsilon q}{2}\right] &< \Pr\left[\left|\sum_{t\in\mathbb{F}_q} 1_{E_t} - \epsilon q\right| < \frac{\epsilon q}{2}\right] \\
&= \Pr\left[\left(\sum_{t\in\mathbb{F}_q} 1_{E_t} - \epsilon q\right)^2 < \frac{\epsilon^2 q^2}{4}\right] \\
&\leq \frac{\mathbb{E}\left[\left(\sum_{t\in\mathbb{F}_q} 1_{E_t} - \epsilon q\right)^2\right]}{\frac{\epsilon^2 q^2}{4}} \\
&\leq \frac{4\epsilon q}{\epsilon^2 q^2} = O\left(\frac{1}{\epsilon q}\right)
\end{aligned}$$

Therefore $\Pr_{x,y\in\mathbb{F}_q^m}\left[|l\cap A| < \frac{\epsilon q}{2}\right] < O\left(\frac{1}{\epsilon q}\right)$. This means that if $\epsilon q = \omega(1)$, with high probability we have that the polynomial decoder contains a polynomial $\overline{Q}(t) = g(x + t(y - x))$.

When this happens, and $a = g(y)$, then $\overline{Q}(1) = a$, then $\overline{Q}$ is a polynomial in the list with $Q(1) = a$.

Lemma 4. *Let $a = g(y)$ and $\{\overline{Q} = Q_0, Q_1, \ldots Q_r\}$ be the polynomials of degree $\leq d$ returned from the polynomial decoder as above. Then $\Pr[\exists\, i : \overline{Q}(1) = Q_i(1)] = o(1)$.*

Proof. Consider choosing l uniformly at random from lines in $\mathbb{F}_q^m$. Note that then choosing x and y uniformly at random (and distinctly) on the line is exactly the same as choosing x and y uniformly at random.

However, once the line is chosen, there are at most $O\left(\frac{1}{\epsilon}\right)$ polynomials Q of degree $< d$ such that $|\{t : (t, Q(t)) \in S\}| \geq \epsilon n$. Note that if $\overline{Q}(t) = Q_i(t)$ for some t, then $\overline{Q}(t) - Q_i(t) = 0$, and since $\deg(\overline{Q}), \deg(Q_i) \leq d$, there are at most d such t for which $\overline{Q}(t) = Q_i(t)$. This is true for every i, meaning that there are $\leq O\left(\frac{d}{\epsilon}\right)$ points on l for which $\overline{Q}(t) = Q_i(t)$ for some i. Since y is chosen uniformly at random on $l, \Pr[\exists\, i : \overline{Q}(1) = Q_i(1)] \leq O\left(\frac{d}{\epsilon q}\right)$. Since $d = O(\epsilon^2 q)$ for the polynomial decoder, $O\left(\frac{d}{\epsilon q}\right) = o(1)$. □

Note that in the above proof, we do not actually need $o(1)$, just a small constant to get high probability ($\geq .99$).

This means that if $a = g(y)$ with high probability $\overline{Q}(t) = g(x + t(y - x))$ is returned from the polynomial decoder and is the unique polynomial returned with $Q(1) = a$, and thus is returned with high probability.

Since $g(y) \in \mathbb{F}_q$, this means that one of the $C_{y,a}$, has $a = g(y)$, and thus $\overline{Q}(t)$ is in the list with high probability (notably $\geq .99$).

□

4 "Extreme" Hardness Amplification

Let $f : \{0,1\}^n \to \{0,1\}$ such that $\forall C$ circuits of size $\leq s, \exists x$ with $C(x) \neq f(x)$.

Using this, we define $\widetilde{f} : \{0,1\}^N \to \{0,1\}$e:

1. Define $g : \mathbb{F}_q^m \to \mathbb{F}_q$ by $g = LDPE(f)$.
2. Let $\widetilde{f} : \mathbb{F}_q^m \times 2^t \to \{0,1\}$ be defined by $\widetilde{f}(x,i) = Had(g(x))_i$ where Had is the Hadamard encoding. Note that this means that $\{0,1\}^N$ is given by identifying $\mathbb{F}_q^m$ with a product of booleans, and taking the binary representation of 2^t. This is done in the same way as in a previous class.

Theorem 5. *$\forall \widetilde{C}$ circuits of size $\leq \frac{s}{poly(n,\epsilon^{-1})}$, $\Pr_{y\in\{0,1\}^n}[\widetilde{f}(y) = \widetilde{C}(y)] < \frac{1}{2} + \epsilon$.*

Proof. Assume for contradiction $\Pr_{x\in\mathbb{F}_q^m, i\in[2^t]}[\widetilde{f}(x,i) = \widetilde{C}(x,i)] \geq \frac{1}{2} + \epsilon$.

Given the circuit $\widetilde{C}$, let $C_0 : \mathbb{F}_q^m \to \mathbb{F}$ be defined by setting $C_0(x)$ as follows:

1. Let $v \in \mathbb{F}_q^m$ be defined by $v_i = \widetilde{C}(x,i)$ by querying $\widetilde{C}$.
2. Let $S = \{r \in \mathbb{F}_q : \Pr_{i\in[2^t]}[Had(r)_i = v_i] \geq \frac{1}{2} + \frac{\epsilon}{2}\}$.
3. Output $r \in S$ chosen uniformly at random.

By Goldreich-Levin $|S| \leq poly(\epsilon^{-1})$. Since r is outputted randomly, and $g \in S$, trivially $C_0(x) = g(x)$ with probability $\geq \epsilon^{O(1)}$ for any $x \in \mathbb{F}_q^m$ with $\Pr_{i\in[2^t]}[\widetilde{C}(x,i) = Had(g(x))_i] \geq \frac{1}{2} + \frac{\epsilon}{2}$.

Claim 6. $\Pr_{x\in\mathbb{F}_q^m}\left[\Pr_{i\in[2^t]}[\widetilde{f}(x,i) = \widetilde{C}(x,i)] \geq \frac{1}{2} + \frac{\epsilon}{2}\right] \geq \frac{\epsilon}{2}$.

Proof. Assume for the sake of contradiction that $\Pr_{x\in\mathbb{F}_q^m}\left[\Pr_{i\in[2^t]}[\widetilde{f}(x,i) = \widetilde{C}(x,i)] \geq \frac{1}{2} + \frac{\epsilon}{2}\right] < \frac{\epsilon}{2}$. This means by the law of total probability, $\Pr_{x\in\mathbb{F}_q^m, i\in[2^t]}[\widetilde{f}(x,i) = \widetilde{C}(x,i)] < \frac{\epsilon}{2}(1) + 1*(\frac{1}{2} + \frac{\epsilon}{2}) = \frac{1}{2} + \epsilon$. □

Now if we let R be the randomness in C_0, then we can let $C_1(x,R)$ be the deterministic function that simulates C_0 with the randomness given by R. Using the definition of C_0, we have $\Pr_{x\in\mathbb{F}_q^m}\left[\Pr_R[C_1(x,R) = g(x)] \geq \epsilon^{O(1)}\right] \geq \frac{\epsilon}{2}$. By a trivial multiplicative bound, this gives that $\Pr_{x\in\mathbb{F}_q^m, R}[C_1(x,R) = g(x)] \geq \epsilon^{O(1)}$.

By the pigeonhole principle, $\exists R'$ such that $\Pr_{x\in\mathbb{F}_q^m}[C_1(x,R') = g(x)] \geq \epsilon^{O(1)}$. By the previous theorem (using $C(x) = C_1(x,R')$ and an appropriate ϵ), we can generate a list of circuits within which $\exists C_2$ circuit such that $\forall x \in \mathbb{F}_q^m, C_2(x) = g(x)$. Since $size(C_2) \leq \text{poly}(n, \epsilon^{-1}) size(\widetilde{C})$, the circuit C_2 is a circuit of size $\leq s$ computing g, meaning f is not worst case hard. This contradiction proves the theorem. □

Let $\epsilon = 2^{-\delta n}$ and $s = 2^{\eta n}$ such that $\frac{s}{\text{poly}(n,\epsilon^{-1})} >> 2^{\eta' n}$ for some η'. If $f \in E$ (time complexity $2^{O(n)}$) such that $\forall C$ circuits of size $\leq 2^{\eta n}, \exists\, x \in \{0,1\}^n$ with $C(x) \neq f(x)$, then by the above theorem, $\exists\, \widetilde{f} \in E$ such that $\forall\, \widetilde{C}$ circuits of size $\leq 2^{\eta' n}, \Pr_{y \in \{0,1\}^N}[\widetilde{C}(y) \neq \widetilde{f}(y)] < \frac{1}{2} + 2^{-\delta n}$. This means that for exponentially computable functions, very (average-case) hard functions exist, assuming there are worst case hard functions.

5 Derandomization

We want to define a pseudorandom generator, but first we have to consider the context in which we care about it. A pseudorandom for cryptography would look different than the one we will create, which will be for randomized algorithms. With that in mind, we define a pseudorandom generator G to be a probability distribution over outputs $r_1, \ldots r_t \in \{0,1\}^n$ such that for any (sufficiently small) randomized algorithm C, $\Pr_{r \in \{0,1\}^n}[C(r) = 1] \approx \frac{1}{t}|\{i : C(r_i) = 1\}|$.

To formalize this, we note that we can replace any randomized algorithm of small size with a circuit of small size. In addition, to formalize the approximation, we use another parameter. This means that an ϵ-PRG against size s is a distribution over $r_1, \ldots r_t \in \{0,1\}^n$ such that $\left|\Pr_{r \in \{0,1\}^n}[C(r) = 1] - \frac{1}{t}|\{i : C(r_i) = 1\}\right| < \epsilon$.

Now we have an example of an pseudorandom generator: Let $h : \{0,1\}^{n-1} \to a\{0,1\}$ be such that $\Pr_{x \in \{0,1\}^{n-1}}[C(x) = h(x)] \leq \frac{1}{2} + \epsilon$ for every circuit C of size s. Output $(x, h(x))$ with x chosen uniformly at random from $\{0,1\}^{n-1}$.

Claim 7. *The above algorithm is an ϵ-PRG against size $\frac{s}{O(1)}$.*

Proof. Assume for the sake of contradiction that that it is not. This means that $\exists\, \widetilde{C}$ circuit of size $\frac{s}{O(1)}$ such that $\left|\Pr_{r \in \{0,1\}^n}[\widetilde{C}(r) = 1] - \frac{1}{2^{n-1}}|\{i : \widetilde{C}(r_i) = 1\}\right| > \epsilon$. By reversing the output of $\widetilde{C}$ if needed, we may assume that $\Pr_{x \in \{0,1\}^{n-1}}[\widetilde{C}(x, h(x)) = 1] - \Pr_{r \in \{0,1\}^n}[\widetilde{C}(r) = 1] > \epsilon$.

Define the circuit C by setting $C(x)$ as follows:

1. Compute $\widetilde{C}(x, 0)$ and $\widetilde{C}(x, 1)$.
2. If $\widetilde{C}(x, 0) = 1$ and $\widetilde{C}(x, 1) = 0$, then output 0.
3. If $\widetilde{C}(x, 0) = 0$ and $\widetilde{C}(x, 1) = 1$, then output 1.
4. Otherwise output a random bit.

Claim 8. *For the above circuit C, $\Pr_{x \in \{0,1\}^{n-1}}[C(x) = h(x)] > \frac{1}{2} + \epsilon$.*

Proof. Let $A_{ij} = \Pr_{x \in \{0,1\}^{n-1}}[\widetilde{C}(x, i) = 1 \wedge h(x) = j]$ for $i, j \in \{0,1\}$. Note that This means that $\Pr_{x \in \{0,1\}^{n-1}}[\widetilde{C}(x, h(x)) = 1] = A_{00} + A_{11}$ and $\Pr_{r \in \{0,1\}^n}[\widetilde{C}(x, i) = 1] = \frac{1}{2}\Pr_{x \in \{0,1\}^{n-1}}[\widetilde{C}(x, 0) = 1] + \frac{1}{2}\Pr_{x \in \{0,1\}^{n-1}}[\widetilde{C}(x, 1) = 1] = \frac{1}{2}(A_{00} + A_{01} + A_{10} + A_{11})$.

This means that $A_{00} + A_{11} - A_{01} - A_{10} > 2\epsilon$. Since $A_{01} - A_{10} = \Pr_{x \in \{0,1\}^{n-1}}[\widetilde{C}(x, 1 - h(x)) = 1]$, we have that $\Pr_{x \in \{0,1\}^{n-1}}[\widetilde{C}(x, h(x)) = 1] - \Pr_{x \in \{0,1\}^{n-1}}[\widetilde{C}(x, 1 - h(x)) = 1] > 2\epsilon$.

By subtracting the probability of intersection, $\Pr_{x \in \{0,1\}^{n-1}}[\widetilde{C}(x, h(x)) = 1 \wedge \widetilde{C}(x, 1 - h(x)) = 0] - \Pr_{x \in \{0,1\}^{n-1}}[\widetilde{C}(x, h(x)) = 0 \wedge \widetilde{C}(x, 1 - h(x)) = 1] > 2\epsilon$. By definition of $C(x)$, this means $\Pr_{x \in \{0,1\}^{n-1}}[C(x) = h(x) \wedge \widetilde{C}(x, h(x)) \neq \widetilde{C}(x, 1 - h(x))] - \Pr_{x \in \{0,1\}^{n-1}}[C(x) \neq h(x) \wedge \widetilde{C}(x, h(x)) \neq \widetilde{C}(x, 1 - h(x))] > 2\epsilon$.

Since $C(x)$ is random when $\widetilde{C}(x, h(x)) = \widetilde{C}(x, 1 - h(x))$, this means that $\Pr_{x \in \{0,1\}^{n-1}}[C(x) = h(x) | \widetilde{C}(x, h(x)) = \widetilde{C}(x, 1 - h(x))] = \frac{1}{2}$. Therefore we have that $\Pr_{x \in \{0,1\}^{n-1}}[C(x) = h(x)] - \Pr_{x \in \{0,1\}^{n-1}}[C(x) \neq h(x)] > .2\epsilon$, and thus $\Pr_{x \in \{0,1\}^{n-1}}[C(x) = h(x)] > \frac{1}{2} + \epsilon$.

□

Now note that C is a circuit of size $O(1) * size(\widetilde{C})$ such that $\Pr_{x \in \{0,1\}^{n-1}}[C(x) = h(x)] > \frac{1}{2} + \epsilon$. This means that if $size(\widetilde{C}) \leq \frac{s}{O(1)}$, then C is a circuit of size $\leq s$ and thus a counterexample to the assumption that $\Pr_{x \in \{0,1\}^{n-1}}[C(x) = h(x)] \leq \frac{1}{2} + \epsilon$ for all circuits C of size $\leq s$. □