03683170: Expanders, Pseudorandomness and Derandomization ??/??/16

Lecture 6 – The Chattopadhyay-Zuckerman two-source extractor

Amnon Ta-Shma and Dean Doron

1 The construction

Our goal is to construct a two-source extractor. Let us start with a failed attempt to construct a 1-source extractor (which is bound to fail because we know there are no deterministic extractor for k_1-sources, even for k_1 as as large as $n_1 - 1$).

Let X_1 be a k_1-source over $\{0,1\}^{n_1}$. Take any strong, one output bit extractor $E : \{0,1\}^{n_1} \times \{0,1\}^{d_1} \to \{0,1\}$. E is a seeded extractor and requires randomness which we don't have. However, we may still compute a $D_1 \times 1$ table T, with $D_1 = 2^{d_1}$ rows, where in the i-th row we write the bit $E(x_1, i)$. I.e., we unfold E and write down all the choices E would have made had it had the randomness it requires. We know that almost all the rows in the table are close to uniform, namely, all rows except maybe a $\sqrt{\varepsilon}$ fraction are $\sqrt{\varepsilon}$-close to uniform.

Now take a resilient function $f : \{0,1\}^{D_1} \to \{0,1\}$ and apply it on the table T. We can think of the good rows as good players outputting random bits, and the bad rows as bad players outputting bits that are adversarially correlated with the good bits. Still, the number of bad players is small and the resilient function is deterministic and knows how to handle malicious players, so we may hope to indeed ensure the output bit is close to uniform.

To summarize, we wold like to output $f(E(x_1, 1), \ldots, E(x_1, D_1))$ and use the strong extractor and the resilient function properties to argue that the output bit is close to uniform.

The approach of course fails, and for two reasons:

1. It is true that each good row in the table is marginally distributed close to uniform. However, the good rows in the table are potentially correlated, e.g., the same bit may repeat in all the good rows. In fact, clearly, the entropy in the output cannot exceed that of the input, and since $D_1 \gg n_1 \geq k_1$, we must have many correlations among the good rows.
2. The number of bad players is quite large. Specifically, if, say, ε is a constant, then there is a constant fraction of bad players, whereas by KKL a resilient function can tolerate at most $\Theta\left(\frac{D_1}{\log D_1}\right)$ bad players.

Let us look at the first problem. Suppose we could guarantee that the good players are t-wise independent for some small (but super-constant) t. Further suppose that f can be computed by an AC^0 circuit of depth d. Then, Braverman's theorem tells us that f is fooled by t-wise independence (for t large enough), i.e., f cannot distinguish between the t-wise independent distribution and the actual distribution (except for a small loss).This means that for the analysis of how f behaves on the input, we can assume (for a small penalty) that all the good rows are independent. Thus, we expect our first problem to be solved! Of course we need to show such an extractor exists, but for the moment let us assume it does (and indeed it does and we will name it a t-non-malleable extractor) and go on with the construction.

The above argument is almost correct except for an annoying bug: The distribution we feed to f is (almost) t-wise independent on the good rows and not truly uniform. The solution here is simple and elegant (as the whole construction!). We use a function f that is also *monotone*. The strategy for the bad players is then immediate and independent of the good rows: Vote 1 if you wish to bias the result to 1, vote 0 if you wish to bias it to 0. Thus, the indicator function of whether the bad players can bias the result is itself an AC^0 circuit and we can again use Braverman's result.

We keep requiring more from our resilient function f, so we should now wonder whether such a resilient function exists at all, namely, a monotone resilient function that is computed by a small-depth circuit. The Tribes function, for instance, is monotone, in AC^0 and optimally resilient against one bad player, but it has very poor resiliency against a coalition of players. Remarkably, there exists an f that can sustain all the above requirements, and we use this fact as a black box without a proof. Thus, remarkably, using such a resilient function solves the first problem.

So far we are still working with just a single source. Indeed, we still face the second problem. The construction is going to use the second source to improve the fraction of bad players among all players, making the whole approach viable. The way this is done is by using the second source to *sample* the rows of the table T, getting a smaller number of rows with about the same fraction $\delta = \sqrt{\varepsilon}$ of bad players. To see that this improves things, assume we started with δD_1 bad rows out of the D_1 rows for some constant δ. Assume we sample a set of size $D_2 = \text{poly}(\frac{1}{\delta})$ (specifically, say $D_2 = (\frac{1}{\delta})^c$ for some constant c) with about 2δ fraction of bad players. The fraction of the bad players before the sampling is a constant which is too bad for us, while the fraction after the sampling is about $\frac{2D_2^{c-1}}{D_2^c} = D_2^{1-\alpha}$ for some $\alpha > 0$, which is tolerable by good resilient functions!

How are we going to do the sampling? We already know the answer. Sampling is almost equivalent to extracting. So, we take an extractor F with δ error. We need to sample from $\{0,1\}^{d_1}$, so this is the range of the extractor F. In the range there are good elements (that correspond to good rows of T, or equivalently, good seeds of E). The fraction of bad elements is δ. The domain of F is (slightly) larger to ensure almost all elements in the domain have about the right fraction δ of good neighbors. So take

$$F : \{0,1\}^{n_2} \times [D_2] \to \{0,1\}^{d_1}$$

that is a (k_2, δ) strong extractor (for $k_2 > d_1$ as small as we can support). Use the element x_2 from the second source to define the sample set

$$\{F(x_2, 1), \ldots, F(x_2, D_2)\}$$

and apply the resilient function f on it, i.e., output

$$f(E(x_1, F(x_2, 1)), \ldots, E(x_1, F(x_2, D_2))).$$

We still need to formally define a t-non-malleable extractor.

Definition 1. *Fix a function $E : \{0,1\}^n \times [D] \to \{0,1\}^m$ and a source X on $\{0,1\}^n$. The table $T = T(E, X)$ is a $D \times m$ table, where the i-th row of T (for $i \in [D]$) contains the distribution $E(X, i)$.*

Definition 2. *A function $E : \{0,1\}^n \times [D] \to \{0,1\}^m$ is a (k, t, ε)-non-malleable strong extractor if for every (n, k)-source X, the $D \times m$ table $T(E, X)$ has at least $(1 - \varepsilon)D$ rows that are $(t, t \cdot \varepsilon)$-independent.*

Notice that when $t = 1$ we recover (a variant) of the familiar strong extractors. The two-source extractor of Chattopadhyay and Zuckerman then does the following:

Input : $x_1 \in \{0,1\}^{n_1}$ sampled from a k_1-source, $x_2 \in \{0,1\}^{n_2}$ sampled from a k_2-source.

The objects we use : We need a t-non-malleable strong extractor with one output bit E, a strong extractor F and a resilient function f. Specifically,

- $E : \{0,1\}^{n_1} \times \{0,1\}^{d_1} \to \{0,1\}$ that is a (k_1, t, ε) non-malleable, strong extractor, where $t \geq \log^c n$ for some constant c to be determined.
- $F : \{0,1\}^{n_2} \times \{0,1\}^{d_2} \to \{0,1\}^{d_1}$ that is a (k_2, δ) extractor, $d_2 = O(\log \frac{n_2}{\delta})$ and δ to be determined.
- $f : \{0,1\}^{D_2} \to \{0,1\}$ for $D_2 = 2^{d_2}$, a (t,γ)-independent (q,δ)-resilient, monotone function that is computed by a depth 4 AC^0 circuit.

The construction : Output

$$EXT(x_1, x_2) \;=\; f(E(x_1, F(x_2, 1)), \ldots, E(x_1, F(x_2, D_2))).$$

2 The correctness proof (assuming the components)

2.1 Facts about f

Recall that a distribution $\mathcal{D}$ is called (t,γ) independent distribution if the restriction of $\mathcal{D}$ to every t coordinates is γ-close to U_t. We extend the notions of non-oblivious bit-fixing sources and resilient functions.

Definition 3. *A source X over $\{0,1\}^n$ is called a (q,t,γ) non-oblivious bit-fixing source if there exists a subset $Q \subseteq [n]$ of size at most q such that the joint distribution of the bits in $Q \setminus [n]$ is (t,γ) independent. The bits in Q are allowed to arbitrarily depends on the bits in $Q \setminus [n]$.*

Definition 4. *For a distribution $\mathcal{D}$ and $Q \subseteq [n]$ we let $I_{Q,\mathcal{D}}(f)$ denote the probability that f is underdetermined when the variables outside Q are sampled from $\mathcal{D}$. We define $I_{q,t,\gamma}(f)$ to the maximum of $I_{Q,\mathcal{D}}(f)$ over all $Q \subseteq [n]$ of size q and all $\mathcal{D}$ that is a (t,γ) independent distribution. We say that f is (t,γ)-independent (q,ε)-resilient if $I_{q,t,\gamma}(f) \leq \varepsilon$.*

The guarantee on our f is given by the following theorem:

Theorem 5. *For any $\delta > 0$ and large enough integer n there exists an explicit boolean function $f : \{0,1\}^n \to \{0,1\}$ satisfying:*

- *f is a depth 4 circuit of size $n^{O(1)}$.*
- *$|\mathbb{E}_{x \sim U_n}[f(x)] - 1/2| \leq \frac{1}{n^{\Omega(1)}}$.*
- *For any $q > 0$, $I_q(f) \leq \frac{q}{n^{1-\delta}}$.*

We will soon prove that that the above result also holds for (t,γ) independent distributions.

2.2 Facts about F

F will be an optimal strong seeded-extractor – the GUV extractor [2].

Theorem 6. *For any $\varepsilon > 0$, any constant $\alpha > 0$ and all integers $n, k > 0$ there exists an explicit strong (k, ε) seeded-extractor $Ext : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with $d = O(\log \frac{n}{\varepsilon})$ and $m = (1-\alpha)k$. Further, for all $x \in \{0,1\}^n$, $Ext(x, s_1) \neq Ext(x, s_2)$ whenever $s_1 \neq s_2$.*

We will also use the fact that every extractor is a good sampler. Formally:

Claim 7. *Let $Ext : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ be a (k, ε) seeded extractor. Identify $\{0,1\}^d$ with $1, \ldots, D$ where $D = 2^d$. Define $Samp(x) = \{Ext(x,1), \ldots, Ext(x,D)\}$. Let X be an $(n, 2k)$-source. Then, for any $R \subseteq \{0,1\}^m$,*

$$\Pr_{x \sim X}[||Samp(x) \cap R| - \mu_R D| > \varepsilon D] \;<\; 2^{-k},$$

where $\mu_R = |R|/2^m$.

2.3 Facts about E

We will use the following non-malleable extractor by Chattopadhyay, Goyal and Li [1].

Theorem 8. *There exists a constant $c > 0$ such that for all $n, t > 0$ there exists an explicit (t, k, ε) non-malleable extractor $nmExt : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}$, where $k \geq ct \log^2 \frac{n}{\varepsilon}$ and $d = O(t^2 \log^2 \frac{n}{\varepsilon})$.*

2.4 Reducing to a bit-fixing extractor (resolving issue I)

Denote

$$S(x_1, x_2) = E(x_1, F(x_2, 1)), \ldots, E(x_1, F(x_2, D_2)),$$

so $EXT(x_1, x_2) = f(S(x_1, x_2))$. Continuing with our previous notations, we prove:

Lemma 9. *There exists a constant $0 < \alpha < 1$ such that with probability at least $1 - n_1^{-\omega(1)}$ over $x_2 \sim X_2$, $S(X_1, x_2)$ is a (q, t, γ) non-oblivious bit-fixing source, where $q = D_2^{1-\alpha}$ and $\gamma = \frac{1}{D_2^{t+1}}$.*

Proof. Let $Samp(x_2) = \{F(x_2, 1), \ldots, F(x_2, D_2)\}$ and let M be the set of good seeds of E (respectively, indices of good rows of $T(E, x_1)$). Note that all elements in $Samp(x_2)$ are distinct. By the properties of E, $|M| \geq (1-\varepsilon)D_1$. By the extractor's sampling property (Claim 7),

$$\Pr_{x_2 \sim X_2}[|Samp(x_2) \cap M| - (1-\varepsilon)D_2 < -\delta D_2] < 2^{-k_2/2}.$$

Thus, with probability at least $1 - 2^{-k_2/2}$, $Samp(x_2)$ contains at least $(1 - \varepsilon - \delta)D_2$ good seeds.

Fix such x_2 and let $Z_{x_2} = S(X_1, x_2)$. The bits corresponding to good rows are $(t, \varepsilon \cdot t)$-independent and there are at most $(\varepsilon + \delta)D_2$ "malicious" bits, so Z_{x_2} is a $(q = (\varepsilon + \delta)D_2, t, t \cdot \varepsilon)$ non-oblivious bit-fixing source.

We choose ε, δ and the exctractors' constants such that $(\varepsilon + \delta)D_2 \leq D_2^{1-\alpha}$ for some constant α and $t \cdot \varepsilon \leq \frac{1}{D_2^{t+1}}$. Specifically, we need to take $\delta = \text{poly}(n_2^{-1})$ and $\varepsilon = \text{poly}(n_1^{-t})$, observing that $D_2 = \text{poly}(n_2)$, E works for min-entropy $k_1 = \text{polylog}(n_1)$ and F works for min-entropy $k_2 = O(D_1) = \text{polylog}(n_1)$ as well. □

2.5 A bit-fixing extractor for (t, γ) independence (resolving issue II)

We prove:

Lemma 10. *Let $C : \{0,1\}^n \to \{0,1\}$ be a monotone AC^0 circuit of depth d and size s such that $|\mathbb{E}_{x\sim U_n}[C(x)] - \frac{1}{2}| \leq \zeta_1$. Suppose $q > 0$ is such that $I_q(C) \leq \zeta_2$ and let $t = \mathrm{poly}((\log \frac{s}{\zeta_3})^{d^2})$ that is guaranteed by Braverman's result for fooling depth $d+5$ circuits of size $3s+6$ with ζ_3 error. Then,*

- *For any distribution $\mathcal{D}$ that is (t,γ) independent, $|\mathbb{E}_{x\sim\mathcal{D}}[C(x)] - \frac{1}{2}| \leq \zeta_1 + \zeta_3 + \gamma n^t$.*
- *$I_{q,t}(C) \leq \zeta_2 + \zeta_3$ and $I_{q,t,\gamma}(C) \leq \zeta_2 + \zeta_3 + \gamma n^t$.*

Proof. Item (1) follows directly from Braverman's result and the fact that $\mathcal{D}$ is $n^t\gamma$-close to a t-wise independent distribution – a fact we proved in the exercise.

For the second item, let Q be any set of indices, and let $\bar{Q} = Q \setminus [n]$. We construct $E_Q : \{0,1\}^{n-q} \to \{0,1\}$ such that $E(y) = 1$ iff C is underdetermined when $\bar{Q}$ is set to y. How do we do that? Let C_0 and C_1 be the circuits obtained from C by fixing the bits in Q to 0 and 1, correspondingly. Define $E_Q = (C_0 \neq C_1)$. Verify to yourself that this indeed satisfies the requirement, as C is monotone. Also, E_Q is an AC^0 circuit of depth $d+5$ and size at most $3s+6$.

Having E_Q, it follows that

$$\mathop{\mathbb{E}}_{y\sim U_{n-q}}[E_Q(y)] \;=\; \Pr_{y\sim U_{n-q}}[E_Q(y)=1] \;\leq\; I_q(C) \;\leq\; \zeta_2.$$

Let $\mathcal{D}$ be any t-wise independent distribution over $n-q$ bits. Similarly, $\mathbb{E}_{y\sim\mathcal{D}}[E_Q(y)] \leq I_{q,t}(C)$. Thus, to prove that $I_{q,t}(C) \leq \zeta_2+\zeta_3$ it is enough to prove that $|\mathbb{E}_{y\sim U_{n-q}}[E_Q(y)] - \mathbb{E}_{y\sim\mathcal{D}}[E_Q(y)]| \leq \zeta_3$, which follows directly from Braverman's result. The fact that $I_{q,t,\gamma}(C) \leq \zeta_2+\zeta_3+\gamma n^t$ again follows from the exercise. □

In light of Lemma 10 and Theorem 5, we can infer:

Theorem 11. *There exists a constant c such that for any $\delta > 0$ and every large enough integer n there exists an explicit boolean function $f : \{0,1\}^n \to \{0,1\}$ satisfying: For every $q > 0$, $t \geq c(\log n)^{18}$ and $\gamma < \frac{1}{n^{t+1}}$,*

- *f is a depth 4 circuit of size $n^{O(1)}$.*
- *For any (t,γ)-wise independent distribution $\mathcal{D}$, $|\mathbb{E}_{x\sim\mathcal{D}}[f(x)] - 1/2| \leq \frac{1}{n^{\Omega(1)}}$.*
- *$I_{q,t,\gamma}(f) \leq \frac{q}{n^{1-\delta}}$.*

2.6 Putting it together

For the simplicity of presentation, let us assume $n_1 = n_2 = n$ and $k_1 = k_2 = k$. We are now ready to prove:

Theorem 12. *There exists a constant $c > 0$ such that $EXT : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ is a two-source extractor for min-entropy at least $\log^c n$ and error $n^{-\Omega(1)}$.*

Proof. Recall that $EXT(x_1, x_2) = f(S(x_1, x_2))$. Let X_1 and X_2 two independent (n, k) sources. By Lemma 9, with probability at least $1 - n^{-\omega(1)}$ over x_2, $S(X_1, x_2)$ is a (q, t, γ) non-oblivious bit-fixing source for $q = D_2^{1-\alpha}$, $t = \text{polylog}(n)$ and $\gamma \leq \frac{1}{D_2^{t+1}}$.Thus, by Theorem 11, for each such good x_2,

$$|f(S(X_1, x_2)) - U_1| \ \leq \ \frac{1}{n^{\Omega(1)}}.$$

Thus, we have that

$$|E(X_1, X_2) - U_1| \ = \ |f(S(X_1, X_2)) - U_1| \ \leq \ \frac{1}{n^{\omega(1)}} + \frac{1}{n^{\Omega(1)}} \ = \ \frac{1}{n^{\Omega(1)}},$$

as required. □

3 A closer look at the components

non-malleable - we want to do. separate section. We need to understand it first.

resilient.

References

[1] Eshan Chattopadhyay, Vipul Goyal, and Xin Li. Non-malleable extractors and codes, with their many tampered extensions. *arXiv preprint arXiv:1505.00107*, 2015.

[2] Venkatesan Guruswami, Christopher Umans, and Salil Vadhan. Unbalanced expanders and randomness extractors from parvaresh–vardy codes. *Journal of the ACM (JACM)*, 56(4):20, 2009.