**03683170: Expanders, Pseudorandomness and Derandomization** 15/06/16

# Take-home exam

*Amnon Ta-Shma and Dean Doron*

**General instructions:**

1. The deadline for the exam is 17/07/16.
2. Submit your (typed) solution by mail to amnon@tau.ac.il and deandoron@mail.tau.ac.il.
3. Using electronic sources is allowed, but the work must be done alone. Please sign the attached statement that you indeed did it alone.
4. Many questions are based on published papers. We do not give the due credit, because we want to encourage you to try it yourself. If you want to get a hint, send us an email.
5. Try to solve all questions. The final grade will be normalized, and everyone who has done decent work will get a decent grade, with those who solved more getting a higher grade. So don't be discouraged if you don't solve all exercises. Sometimes, it is possible to solve parts of a question assuming the correctness of the previous parts.
6. If you find a mistake (or a typo), or you have a suggestion that may benefit others, please let us know as soon as possible.
7. You have a month, use it!
8. Enjoy!

# 1 Improving Braverman's result

## 1.1 $(\alpha, \beta)$ majority – Amplifying gaps

**Definition 1** (The $(\alpha, \beta)$-MAJ promise problem)**.** *The input to the promise problem is $x \in \{0,1\}^n$. The YES instances are all $x \in \{0,1\}^n$ such that $\sum x_i > \beta n$. The NO instances are all $x \in \{0,1\}^n$ such that $\sum x_i \leq \alpha n$.*

Although $(\frac{1}{2}, \frac{1}{2})$-MAJ is hard for $\mathsf{AC}^0$ (we proved it in the exercise), prove that:

**Question 1.1.** *For all constants $0 \leq \alpha < \beta < 1$ and large enough $n$, there exist constant $c, d$ such that $(\alpha, \beta)$-MAJ has an $\mathsf{AC}$ circuit with size $n^c$ and depth $d$.*

What is the constant $d$ you have found?

Hint: Find a probabilistic algorithm with error probability smaller than the number of inputs. Use alternate steps of powering and negation.

## 1.2 Amplifying the success of a distribution

We repeat a definition we gave in class.

**Definition 2.** *Let $f : \{0,1\}^n \to \{0,1\}$ and $P$ a distribution over real polynomials. We say:*

- *$P$ $\varepsilon$-approximates $f$ (worst-case and exact) if for every $x \in \{0,1\}^n$, $\Pr_{p \in P}[p(x) = f(x)] \geq 1 - \varepsilon$,*
- *We say $\| P \|_\infty \leq L$ if for every $p \in P$, $\max_{a \in \{0,1\}^n} |p(a)| \leq L$,*
- *We say $P$ has degree $D$ if every $p \in P$ has degree at most $D$.*

Prove:

**Question 1.2.** *Suppose $P$ $(\frac{1}{2} - \delta)$-approximates $f : \{0,1\}^n \to \{0,1\}$, has degree $D$ and norm $L \geq 2$. Then, there exists a distribution $P'$ that $\varepsilon$-approximates $f$, has degree at most $O(\frac{D}{\delta^2} \log \frac{1}{\varepsilon})$ and norm at most $L^{O(\frac{1}{\delta^2} \log \frac{1}{\varepsilon})}$.*

## 1.3 Amplifying the success of the error detector

In the lecture we proved:

**Theorem 3.** *Fix $\varepsilon > 0$. Any $\mathsf{AC}$ circuit $C$ of size $s$ and depth $d$ has a distribution $P$ over real polynomials such that:*

- *The degree of $P$ is at most $(\log \frac{s}{\varepsilon})^{O(d)}$.*
- *The norm of $P$ is at most $2^{(\log \frac{s}{\varepsilon})^{O(d)}}$.*
- *For every $p \in P$ there exists an "error-detecting" circuit $E_p$ such that:*

- *$E_p$ has size* $\text{poly}(s \log \frac{1}{\varepsilon})$ *and depth* $d + O(1)$,
- *(Small error) For every* $a \in \{0,1\}^n$, $\Pr_p[E_p(a) = 1] \leq \varepsilon$, *and,*
- *(One-sided error) For every* $a \in \{0,1\}^n$, *whenever* $E_p(a) = 0$, $p(a) = C(a)$.

Improve the above theorem and prove that:

**Question 1.3.** *Fix* $\varepsilon > 0$. *Any* AC *circuit* $C$ *of size* $s$ *and depth* $d$ *has a distribution* $P$ *over real polynomials such that:*

- *The degree of* $P$ *is at most* $(\log s)^{O(d)} \log(\frac{1}{\varepsilon})$.
- *The norm of* $P$ *is at most* $(\frac{1}{\varepsilon})^{(\log s)^{O(d)}}$.
- *For every* $p \in P$ *there exists an "error-detecting" circuit* $E_p$ *such that:*
  - *$E_p$ has size* $\text{poly}(s \log \frac{1}{\varepsilon})$ *and depth* $d + O(1)$,
  - *(Small error) For every* $a \in \{0,1\}^n$, $\Pr_p[E_p(a) = 1] \leq \varepsilon$, *and,*
  - *(One-sided error) For every* $a \in \{0,1\}^n$, *whenever* $E_p(a) = 0$, $p(a) = C(a)$.

Using this result,

**Question 1.4.** *Re-prove Braverman's result that* $t = t(s, d, \varepsilon)$*-wise independence* $\varepsilon$*-fools* AC *circuits of size* $s$ *and depth* $d$, *but improve the dependence* $t$ *has on* $\varepsilon$.

## 2 Derandomizing the Ajtai-Linial function

Ajtai and Linial proved:

**Theorem 4.** *There exists an almost balanced boolean function $f$ on $n$ variables such that for every $\varepsilon > 0$, it holds that $I_q(f) = O(\varepsilon)$ for $q = \frac{\varepsilon n}{\log^2 n}$.*

The Ajtai-Linial construction is both non-explicit and non-monotone. Chattopadhyay and Zuckerman de-randomized the construction and made it *explicit* and also constructed a function $f$ that is *monotone*, while keeping it in $\mathsf{AC}^0$. Their construction works for $q = n^{1-\alpha}$ (for every fixed constant $\alpha > 0$). Here, we will present the construction, and do part of the analysis.

Specifically, we will show that if the good players are Bernoulli $p$ (i.e., each good player picks 1 with probability $p$ and the good players are independent) then for every coalition of size $q$, the probability the coalition can influence the result is small. We will not show:

- How to move to the uniform distribution (i.e., Bernoulli $p = 1/2$), and,
- How to show that the function is almost balanced (which requires new ideas).

We start with the construction. Let $[M]$ be some universe, for $M$ that will be determined later. Take a strong $(k =?, \varepsilon =?)$ extractor $E : [R] \times [B] \to [M]$ such that $n = B \cdot M$, for parameters that you will have to choose later.

**Question 2.1.** *Prove that for every $v \in [R]$, the sets $\{(j, E(v,j) \oplus w) \mid j \in [B]\}_{w \in [M]}$ form a partition of $[n]$.*

Fix $v \in [R]$. We will say the set

$$P_w^v = \left\{ x_{j,E(v,j)\oplus w} \mid j \in [B] \right\},$$

for $w \in [M]$, is the $w$-th tribe of $v$. We let $f_v : \{0,1\}^n \to \{0,1\}$ be the Tribes function for the above partition, i.e.,

$$f_v(x_1, \ldots, x_n) \;=\; \bigvee_{w \in [M]} \bigwedge_{j \in [B]} x_{j,E(v,j)\oplus w}.$$

We define $f : \{0,1\}^n \to \{0,1\}$ by

$$f(x_1, \ldots, x_n) \;\stackrel{\text{def}}{=}\; \bigwedge_{v \in [R]} f_v(x_1, \ldots, x_n)$$

As always, we say that the (good players) leave $f$ undetermined if the malicious players can set the function to either zero or one after the good players set their values. We say that $f$ is set to 1 (or 0) if any completion of the good players results in $f$ being 1 (or 0).

**Question 2.2.** *Fix $v \in [R]$. Say a tribe of $v$ is* good *if all the variables in it are good, and* bad *otherwise. Assume the good players have chosen their values some way. Prove that if $f_v$ is undetermined then:*

1. *In every good tribe of $v$ there is a player that voted* $0$*, and,*

2. *There exists a bad tribe of $v$ in which all the good players voted* $1$*.*

Let $GT(v)$ denote the number of good tribes of $v$ and $BT(v)$ the number of bad tribes of $v$. Prove:

**Question 2.3.**

1. *For every $v$, the probability (over the votes of the good players) that $f_v$ is not set to* $1$ *is exactly* $(1-p^B)^{GT(v)}$*.*

2. *There exists a set $BV \subseteq [R]$ of cardinality at most $2^k M$ such that for every $v \notin BV$, the probability (over the votes of the good players) that $f_v$ is undetermined is at most $BT(v) \cdot p^{(1-2\varepsilon)B}$.*

3. *The probability that $f$ is undetermined is at most $2^k M(1-p^B)^{M-q} + q \cdot p^{(1-2\varepsilon)B}$.*

Set the parameters $R, B, M, k, \varepsilon$ and $p$ in order to prove:

**Question 2.4.** *For any $\delta > 0$, and every large enough integer $n$, there exists a polynomial time computable monotone, balanced boolean function $f : \{0,1\}^n \to \{0,1\}$ satisfying:*

- *$f$ is a depth 4 circuit of size $n^{O(1)}$.*
- *For any $q > 0$, $I_q(f) \leq \frac{q}{n^{1-\delta}}$ where the good players are Bernoulli $p$.*

*You may assume you have the best non-explicit extractor.*

# 3 Leader election

**Question 3.1.** *We describe the baton passing game. The game starts with player number* 1 *holding the baton. Every player is supposed to choose uniformly at random a player that has not received the baton yet and pass it to him. The elected leader is the player who receives the baton last.*

*The adversarial model: The adversary may corrupt $q$ players of his choice. He also chooses the first player. Bad players play arbitrarily, may collude, and may base their actions on previous history.*

- *Find the optimal strategy for the adversary, and prove it is optimal.*
- *Prove the probability the leader is bad is a function of only:*
  - *The number of good and bad players,*
  - *Whether the player holding the baton at the first stage is bad or not.*
- *Prove that for any adversary, the probability (over the good players' coin tosses) that the leader is bad is $O(\frac{q \log q}{s})$, where $q$ is the number of bad players and $s$ the number of good players.*

**Question 3.2.** *Solve question 6 in the questions pool.*

## 4 A two-source extractor from $\varepsilon$-biased sample spaces

We recall the definition we gave in class:

**Definition 5.** *Let $X$ be a distribution over $\{0,1\}^n$. We say $X$ is* $(k,\varepsilon)$-biased*, if it is at most $\varepsilon$-biased with respect to all non-empty, linear tests of size at most $k$.*

Recall, that we have proved:

**Theorem 6.** *There exists an explicit distribution that is $(k,\varepsilon)$-biased over $\{0,1\}^n$ and has support size at most $\left(\frac{k\log n}{\varepsilon}\right)^2$.*

We now give a two-source extractor construction. We use the first source to sample a row from a $(k,\varepsilon)$ biased distribution, and the second source to sample bits from it. Specifically, fix $n_1, n_2, k_1, k_2, m, k$ and $\varepsilon$ such that $n_1$ is the number of bits required to construct a $(km,\varepsilon)$-biased sample space with $mN_2$ variables ($N_2 = 2^{n_2}$). We also assume $km$ is even. Define $E : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ by:

- Use $x_1 \in \{0,1\}^{n_1}$ to sample $Z(x) = Z = (Z_{1,1}, \ldots, Z_{i,j}, \ldots, Z_{m,N_2})$ from the distribution.
- Sample $x_2 \in \{0,1\}^{n_2}$, and, output,

$$E(x,y) = Z_{1,y}(x) \circ \ldots \circ Z_{m,y}(x).$$

We want to prove $E$ is a good two-source extractor (when one source is dense). Before we start prove:

**Question 4.1.** *$E$ is a $((n_1,k_1),(n_2,k_2) \to_\gamma m)$ two-source extractor, iff it is a two source extractors for all flat sources $A$ and $B$ over $K_1 = 2^{k_1}$, $K_2 = 2^{k_2}$ elements, respectively.*

From now on we assume $X$ is a flat distribution over a set $A$ of size $K_1 = 2^{k_1}$, and $Y$ is a flat distribution over a set $B$ of size $K_2 = 2^{k_2}$. We will start with the $m = 1$ case. Prove:

**Question 4.2.** *$E(X,Y)$ is $\gamma$-biased for $\gamma = 2^{\frac{n_1-k_1}{k}} \cdot \left(\varepsilon^{1/k} + k \cdot 2^{-k_2/2}\right)$.*

If you wish, you may follow the following proof framework. For $x \in \{0,1\}^{n_1}, y \in \{0,1\}^{n_2}$ define

$$e(x,y) = (-1)^{E(x,y)}.$$

Our goal is to bound $|\mathbb{E}_{a\in A, b\in B}\, e(a,b)|^k$, and to use the $k$-wise independence for that (in a way similar to what we did when proving concentration bounds for $k$-wise independence). Use Jensen's inequality, and the fact that $k$ is even, to prove $(\mathbb{E}_{a\in A}\sum_{b\in B} e(a,b))^k \le \mathbb{E}_{a\in A}(\sum_{b\in B} e(a,b))^k$. Use the $k$-wise independence to prove that for any $r \le k$, and any different $b_1,\ldots,b_r \in \{0,1\}^{n_2}$, $\mathbb{E}_{x\in\{0,1\}^{n_1}} \prod_{j=1}^r e(x,b_j) \le \varepsilon$.

We now want to augment this in several ways: first, we would like to output many bits, and second we want to prove the construction is strong.

**Question 4.3.** *Let $E$ be as above with $m$ output bits. Prove that the output of $E$ is $2^{m/2}\gamma$ close to uniform.*

finally, prove it is strong (with weaker parameters), and choose parameters. Prove:

**Question 4.4.** *Prove that there exist constants $c_1, c_2 > 0$ such that for any $n$ there exists an explicit $((n, k_1 = \frac{3}{4}n), (n, k_2 = c_1 \log n) \to_\gamma m)$ two-source extractor $E : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ for $m = c_2 k_2$ that is strong in the first source, and where $\gamma = 2^{-c_2 k_2}$.*

So far we have extracted $\Omega(k_2)$ bits (and we think of $k_2$ as being small). Now, we would like to compose it with a good seeded extractor to output $\Omega(k_1 + k_2)$ bits. Suppose we have:

- $2EXT : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^d$ be a $((n_1, b_1), (n_2, b_2) \to_{\varepsilon_1} d)$ two-source extractor, that is strong in the first source.
- $EXT : \{0,1\}^{n_1} \times \{0,1\}^d \to \{0,1\}^m$ be a $(b_1, \varepsilon_2)$ extractor.

Define $E : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ to be the following composition of $2EXT$ and $EXT$:

$$E(x, y) = EXT(x, 2EXT(x, y)).$$

**Question 4.5.** *Prove that $E$ is a $((n_1, b_1), (n_2, b_2) \to_{\varepsilon_1 + \varepsilon_2} m)$ two-source extractor.*

## 5 The above extractor is non-malleable

We now claim that the two-source extractor given in the previous section is, in fact, non-malleable (when reducing the output length appropriately). The idea is that if an adversary can associate a seed to another in a way that the outputs are correlated, then the parity of the two-outputs is biased. To prevent that, in the extractor of the previous section, instead of by picking $x$ from a $(k = k_1 m, \varepsilon)$-wise independent distribution, we will chose it from a $(2k, \varepsilon)$-wise independent distribution, and this should make us immune against such adversaries.

To make this work, in the sum $\mathbb{E}_{x \in \{0,1\}^{n_1}} \prod_{j=1}^{r} e(x, b_j) e(x, \phi(b_j))$, where $\phi(b_j)$ is the seed the adversary associates with the seed $b_j$, it seems we need, somehow, that the set of all $b_i$ is different than any $\phi(b_j)$, which is impossible (why?). For that they use a "representing sample", for which this holds.

We cite the following graph theoretic lemma:

**Lemma 7.** *Let $G = (V, E)$ be a* directed *graph with no self-loops (but possibly with parallel edges) where every vertex has regular out-degree $t$. Let $w : V \to \mathbb{R}$ be a weight function. Then, there exists a subset $A \subseteq V$ such that*

- *The induced subgraph of $G$ on $A$ is a-cyclic.*
- $|A| \geq \frac{|V|}{t+1}$, *and,*
- *Let $\overline{w(S)}$ denote the average weight of vertices in $S \subseteq V$, i.e., $\overline{w(S)} = \frac{1}{|S|} \sum_{s \in S} w(s)$. Then, $\overline{w(A)} \geq \frac{\overline{w(V)}}{t+1}$.*

You do not need to prove the lemma (but, of course, you may prove it if you wish).

With that prove:

**Question 5.1.** *There exists a constants $c > 0$ such that for any $n$ there exists an explicit $(k = \frac{3}{4}n, \varepsilon)$ 1-non-malleable extractor $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ for $d = O(\log n)$, $m = \Omega(d)$ and $\varepsilon = 2^{-m}$.*

You do not need to write the whole proof from the beginning, just mention the places where there is a change. Also specify what is the graph $G$ and how you define the weight of vertices in the graph.

Doing it for $t$-wise, would give an extra bonus.

## 6 A condenser from Trevisan's reconstruction scheme

In the lecture notes you will find a section on reconstructive extractors (though, we did not go over it in class). We use the definitions given in the lecture notes.

Let $(E, A, R)$ be a $(p, q)$ reconstructive extractor, where:

- $E : \{0,1\}^n \times \{0,1\}^{r_E} \to \{0,1\}^m$,
- $A : \{0,1\}^n \times \{0,1\}^{r_A} \to \{0,1\}^a$, and,
- $R : \{0,1\}^a \times \{0,1\}^{r_A} \times \{0,1\}^{r_R} \to \{0,1\}^n$.

Let $X$ be an $(n, k)$-source and $Y$ be the uniform distribution over $\{0,1\}^{r_A}$.

**Question 6.1.** *Prove that the distribution $A(X, Y) \circ Y$ is $(1-q)$-close to a distribution with min-entropy at least $k + r_A - \log \frac{1}{q} - 1$.*

We now want to use the above fact to prove that Trevisan's advice function is a *loseless condenser.*

**Definition 8.** *A function $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a $k_1 \to_\varepsilon k_2$ condenser if for every $(n, k_1)$-source $X$, $C(X, U_d)$ is $\varepsilon$-close to a distribution with $k_2$ min-entropy. We say $C$ is lossless if $k_1 = k_2$.*

In Lecture 5, our next-bit predictor had a relatively small success probability, and consequently the success of the reconstruction scheme ($q$) was polynomially-small. Consider the following Lemma:

**Lemma 9.** *If a flat distribution $Y$ over $\{0,1\}^m$ has min-entropy at most $\varepsilon m$, then there is a next-bit predictor $T$ for $Y$ with success probability $1 - \varepsilon$.*

Use Trevisan's scheme and the previous question to obtain a small-error lossless condenser. Specifically, prove the following:

**Question 6.2.** *Assume there exists a weak $(\ell = \log \bar{n} = \log n + O(1), \rho)$ design $Z_1, \ldots, Z_m \subseteq [t]$ with $m \geq \frac{k+t}{\varepsilon}$. Then, there exists a function $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^a$ that is a $k \to_{O(\varepsilon)} k$ condenser, with $a = \rho m$ and $d \leq t$.*