Problem Set 4

- Due Date: **21 Dec 2021**
- The points for each problem is indicated on the side. This problem set has **85** points but you may solve any **60** points worth of questions among these for a full score (the remaining **25** points are bonus). Any additional points obtained will still count towards your final aggregate.
- Turn in your problem sets electronically (PDF; either LaTeXed or scanned etc.) on Acadly.
- Collaboration is encouraged, but all writeups must be done individually and must include names of all collaborators.
- Referring to sources other than the text book and class notes is strongly discouraged. But if you do use an external source (eg., other text books, lecture notes, or any material available online), ACKNOWLEDGE all your sources in your writeup, with a brief remark on *why* you sought that source. This **will not** affect your grades. However, not acknowledging will be treated as a serious case of academic dishonesty.
- Be clear in your writing.

1. **[PRG for RL using extractors]** (10 + 5)

   In class we saw Nisan's PRG for RL using expanders, and you saw how to use pairwise independent hash families in problem set 3. In this problem, you will see how to use extractors.

   Suppose $\mathrm{Ext}\colon \{0,1\}^{m_1} \times \{0,1\}^d \to \{0,1\}^{m_2}$ is a strong (k, ε)-extractor with $k = m - c - \log \frac{1}{\varepsilon}$, define the following generator:

   $$G\colon \{0,1\}^{m_1} \times \{0,1\}^d \to \{0,1\}^{m_1+m_2}$$
   $$G(x,y) = x \circ E(x,y)$$

   (a) Show that the above G is a generator for 2-step communication-c algorithms with error 2ε, where the first step requires m_1 random bits and the second step requires m_2 random bits.

   (b) Can you say something (informally) about viewing Nisan's PRG that we discussed in class (and saw in the last problem set) in this light?

2. **[Impossibility of deterministic extractors]** (10 + 5)

   (a) Suppose X is a source on $\{0,1\}^n$ such that for all x, y we have

   $$\frac{\Pr[X = x]}{\Pr[X = y]} \leq \frac{1-\delta}{\delta}.$$

   Show that $X \in \text{Unpred-Bits}_\delta$.

   (b) For any $\delta > 0$, and any deterministic function $\mathrm{Ext} : \{0,1\}^n \to \{0,1\}$, construct an $\text{Unpred-Bits}_\delta$ source X such that

   $$\Pr_{x\sim X}[\mathrm{Ext}(x) = 1] \leq \delta \quad \text{or} \quad \Pr_{x\sim X}[\mathrm{Ext}(x) = 1] \geq 1 - \delta.$$

Base version: (2021-12-07 21:52:45 +0530) , ccd8d0b

[Hint: Try out X being uniform on $\mathrm{Ext}^{-1}(0)$ with probability δ and uniform on $\mathrm{Ext}^{-1}(1)$ with probability $1-\delta$, or vice-versa.]

3. **[Building the building-block-extractor]** (15)

We had outlined a construction of a building-block extractor in proof of the Guruswami-Umans-Vadhan theorem. In this question, you are to formalise that construction. Prove the following:

> Let $t > 1$ be an integer. For any $n \geq k$ and $\varepsilon > 0$, formally build an explicit (k, ε)-extractors $\text{BB-Ext}^{(t)}\colon [N] \times [D] \to [M]$ with $m \geq k/2$ and $d \leq \frac{k}{t} + O(\log \frac{n}{\varepsilon})$.

4. **[Extractors from almost pairwise independent hash functions]** (10)

In problem set 1, you came across ε-almost pairwise independent families. In this problem we will consider a minor relaxation and say a family $\mathcal{H} = \{h\colon [N] \to [M]\}$ is *ε-almost universal* if for any $x_1 \neq x_2 \in [N]$, we have

$$\Pr_{h \in \mathcal{H}}[H(x_1) \neq H(x_2)] \leq \frac{1+\varepsilon}{M}.$$

Note that $\frac{\varepsilon}{M^2}$-almost pairwise independent family is ε-almost universal. In problem set 1, you would have seen that there are such families $\mathcal{H}$ such that specifying an element of $\mathcal{H}$ only requires $O(m + \log n + \log \frac{1}{\varepsilon})$ as opposed to $O(m+n)$ for standard pairwise independent families.

Show that if $\mathcal{H} = \{h\colon [N] \to [M]\}$ is ε^2-almost universal, then the map $\mathrm{Ext}(x, h) := (h, h(x))$ is a (k, ε)-extractor for $k = m + 2\log \frac{1}{\varepsilon} + O(1)$.

Use this to deduce that, for all $n \geq k$ and $\varepsilon > 0$, there is a strong (k, ε)-extractor $\mathrm{Ext} : [N] \times [D] \to [M]$ with $d = O(k + \log \frac{n}{\varepsilon})$ with $m = k - O(\log \frac{1}{\varepsilon})$.

5. **[The Hadamard Matrix and a two-source extractor]** (3 + 5 + 7)

The construction in this problem is due to Chor and Goldreich.

For any power of two, $N = 2^n$, define the $N \times N$ *Hadamard Matrix* H_n inductively as follows:

$$H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$H_{i+1} = \begin{bmatrix} H_i & H_i \\ H_i & -H_i \end{bmatrix} \quad \text{for all } i \geq 1.$$

Another interpretation of this matrix is that if we index the rows and columns as vectors in $\mathbb{F}_2^n$, then the $(\mathbf{u}, \mathbf{v})$-th entry is $(-1)^{\langle \mathbf{u}, \mathbf{v} \rangle}$.

(a) Show that the rows of H_n are pairwise orthogonal and have ℓ_2-norm $\sqrt{N}$ each.

(b) Suppose $S, T \subseteq [N]$, show that

$$\sum_{i \in S, j \in T} (H_n)_{i,j} \leq \sqrt{|S| \cdot |T| \cdot N}.$$

[Hint: Write the LHS as an inner-product of two vectors, one of which is just a sum of $|S|$ rows of H_n. Then, use the fact that the rows of H_n are orthogonal to bound the norm of the sum of „few" rows of H_n.]

(c) Consider the map Ext: $[N] \times [N] \to \{0,1\}$ given by

$$\mathrm{Ext}(\mathbf{u}, \mathbf{v}) = \langle \mathbf{u}, \mathbf{v} \rangle \bmod 2$$

Show that is a $(k, k; \varepsilon)$-two-source extractor for $k = \frac{n}{2} + O\left(\log \frac{1}{\varepsilon}\right)$ (that is, if the two independent sources have min-entropy at least k each, then the output is ε-close to uniform.)

6. **[A construction of 3AP-free sets]** (3 + 3 + 9)

The construction in this problem is due to Behrend.

Let $S_d(k) = \left\{x \in \mathbb{R}^d \ : \ \|x\|_2 = k\right\}$, the d-dimension sphere of radius k. Note that this is 3AP-free (i.e., there are no distinct points $x, y, z \in S_d(k)$ such that $x + z = 2y$ since the line joining x and z intersects the sphere at exactly the two end-points and hence its midpoint cannot lie on the sphere).

(a) For any positive integer r, consider $U_r = [r]^d$, the set of integer points all of whose coordinates are from $\{1, 2, \ldots, r\}$. Show that for any r, there is some k such that $S_d(k)$ contains at least r^{d-2}/d points of U_r.

[Hint: Pigeon-hole principle]

(b) Consider the following map of converting a vector in U_r into an integer:

$$\Phi : (u_1, \ldots, u_d) \mapsto u_1 + (2r)u_2 + (2r)^2 u_3 + \cdots + (2r)^{d-1} u_d$$

Show that this is an injective map (that is, no two vectors of U_r are mapped to the same integer) to integers less than $(2r)^d$. Furthermore, show that if $u, v, w \in U_r$ with $u + w = 2v$, then we also have $\Phi(u) + \Phi(w) = 2\Phi(v)$ as integers.

[Hint: Look at numbers in base $2r$.]

(c) Show that there infinitely many N with subsets $A_N \subseteq [N]$ that are 3AP-free with

$$\delta_N = \frac{|A_N|}{N} = \exp(-c\sqrt{\log N})$$

for some constant absolute $c > 0$.

[Hint: Set $d = \sqrt{\log N}$ and $r = 2^{\sqrt{\log N}}$ and use the previous parts.]