Coding Theory 10 Nov, 2016

Problem Set 2

- Due Date: **15 Dec 2016**
- Turn in your problem sets electronically (LaTeX, pdf or text file) by email. If you submit handwritten solutions, start each problem on a fresh page.
- Collaboration is encouraged, but all writeups must be done individually and must include names of all collaborators.
- Refering 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 (including collaborators) in your writeup. This will not affect your grades. However, not acknowledging will be treated as a serious case of academic dishonesty.
- The points for each problem are indicated on the side.
- Be clear in your writing.
- Problems 1 is from Guruswami's course while problems 5, 6 are due to Ramprasad Saptharishi.

---

1. (15) **Parallel decoding of expander codes**

   Consider the binary expander code based on a (d, D)-bounded unbalanced bipartite $(d(1-\varepsilon), \delta)$-expander (L, R, E) as defined in lecture (i.e., the code whose parity check matrix is the bipartite adjacency matrix of the expander) for some $\varepsilon \in (0, 1/20)$. Let $|L| = n$. In lecture, we gave a sequential decoder that decoded as long as the fraction of errors is at most $\delta(1-2\varepsilon)$. In this problem, we will analyze the following parallel iterative decoder.

   > For $c \log n$ rounds (for a constant c chosen large enough), do the following in parallel for each variable node: If the variable is in at least $2d/3$ unsatisfied checks, flip its value.

   Prove that the above algorithm corrects any pattern of $\delta(1-3\varepsilon)n$ errors.

2. (3+4+3+5) **Exponential lower bounds for 2-query linear LDCs**

   In this problem, we will prove an exponential lower bound for 2-query linear locally decodable codes.

   Recall that a code $\mathcal{C} : \{0,1\}^k \to \{0,1\}^n$ is said to be (q, δ, ε)-locally decodable if there exists a (probabilistic) decoder D such that on oracle access to any $\mathbf{y} \in \{0,1\}^n$ that satisfies $\Delta(\mathbf{y}, \mathcal{C}(\mathbf{x})) \leq \delta n$, we have

   - $\forall i \in [k], \Pr\left[D^{\mathbf{y}}(i) = \mathbf{x}_i\right] \geq \frac{1}{2} + \varepsilon$.
   - D makes at most q probes into $\mathbf{y}$ on any input i and internal random coins.

For fixed $c \in \mathbb{R}$, $\varepsilon \in (0,1)$ and integer 2, we say that $\mathcal{C} : \{0,1\}^k \to \{0,1\}^n$ is a (q, c, ε)-smooth code if there exists a probabilistic oracle machine A such that:

- In every invocation, A makes at most q queries non-adaptively.
- For every $\mathbf{x} \in \{0,1\}^k$ and and for every $i \in [k]$, we have

$$\Pr[A^{\mathcal{C}(\mathbf{x})}(i) = \mathbf{x}_i] \geq \frac{1}{2} + \varepsilon.$$

- For every $i \in [k]$ and $j \in [n]$, the probability that on input i the oracle machine A queries index j is at most $\frac{c}{m}$.

(a) Show that if $\mathcal{C} : \{0,1\}^k \to \{0,1\}^n$ is a (q, δ, ε)-locally decodable code, then $\mathcal{C}$ is also a $(q, q/\delta, \varepsilon)$-smooth code.

Let $\mathcal{C} : \{0,1\}^k \to \{0,1\}^n$ be a linear code. Since $\mathcal{C}$ is linear, we might wlog. assume that there exist $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \{0,1\}^k$, such that for all $\mathbf{x} \in \{0,1\}^k$ and $j \in [n]$, we have $\mathcal{C}(\mathbf{x})_j = \langle \mathbf{a}_j, \mathbf{x} \rangle$. For simplicity, let us assume that all the $\mathbf{a}_i$'s are distinct. Suppose $\mathcal{C}$ is a $(2, \delta, \varepsilon)$-locally decodable for some $\delta, \varepsilon \in (0,1)$. Let us further make a simplifying assumption that the local D (corresponding to $\mathcal{C}$) makes exactly 2 probes every time and uses both the probes. It follows from 2a that $\mathcal{C}$ is $(2, 2/\delta, \varepsilon)$-smooth.

Construct recovery graphs $\{G_i = ([n], E_i)\}_{i=1}^k$ based on the smooth decoder A for $\mathcal{C}$ as follows: the vertices of all the k graphs G_i's are $[n]$. Two vertices $j, j' \in [n]$ are connected in G_i if

$$\Pr[A^{\mathcal{C}(\mathbf{x})}(i) = \mathbf{x}_i | A \text{ queries } \mathcal{C}(\mathbf{x}) \text{ at indices } j \text{ and } j'] > \frac{1}{2}.$$

(b) If G is $(2, c, \varepsilon)$-smooth, show that for each $i \in [k]$, the graph G_i has a matching $M_i \subseteq E_i$ of size at least $\varepsilon n/c$.

(c) Argue that for each $i \in [k]$, if $(j, j') \in E_i$ then $\mathbf{e}_i \in \text{span}\{\mathbf{a}_j, \mathbf{a}_{j'}\}$. It then follows from our assumption ("the local D makes exactly 2 probes every time and uses both the probes") that $\mathbf{a}_j + \mathbf{a}_{j'} = \mathbf{e}_i$.

[For extra credit, do not make this simplifying assumption and modify the following part suitably to still yield an exponential lower bound.]

We can thus identify the vertices $[n]$ with the set $A = \{\mathbf{a}_j | j \in [m]\}$, a subset of the vertices of the hypercube $\{0,1\}^k$ and the edges (j, j') with the corresponding edges in the hypercube. Consider the graph $G = ([n], E_1 \cup \cdots \cup E_k)$. From the above identification, we get that G is a subgraph of the hypercube. Furthermore, from 2c, we get that the k edge-sets E_i are all distinct. Hence, from 2b, we have $|E(A, A)| \geq \sum_{i=1}^k |E_i| \geq k \cdot (\varepsilon n/c) = \varepsilon \delta k n/2$. Here, $E(A, A)$ refers to the edges in G both of whose endpoints in is A.

(d) Since G is a subgraph of the hypercube, use the upper bound on $E(A, A)$ to conclude that $n \geq 2^{\varepsilon \delta k}$.

This proves an exponential lower bound on the size of any 2-query linear LDC (provided all the codeword bits are distinct, ie. $\mathbf{a}_j's$ are distinct). For extra credit, see if you can remove this assumption of distinctness.

3. (4+7+8+1) **3-AP-free sets in $\mathbb{F}_3^n$ via the polynomial method**

Let $A \subseteq \mathbb{F}_3^n$. We say that A is 3-AP-free if there does not exist $x \neq y \in \mathbb{F}_3^n$ such that $x, (x+y)/2, y \in A$ (i.e., A does not contain any non-trivial arithmetic progression of length 3). In this problem, we will use the polynomial method to show that any 3-AP-free set is of size at most c^n for some fixed $c \in (2,3)$.

(a) (1+1+2) Let $0 \leq d \leq 2n$. Let $V_d(n)$ denote the set of all functions from $\mathbb{F}_3^n$ to $\mathbb{F}_3$ expressible as degree d polynomials. In other words, if $f \in V_d$, then f can be expressed as a polynomial of the form

$$f(x_1, \ldots, x_n) = \sum_{\mathbf{a}=(a_1,\ldots,a_n) \in \{0,1,2\}^n : \sum a_i \leq d} c_{\mathbf{a}} \prod_{i=1}^{n} x_i^{a_i}.$$

Let $m_d(n) = \dim(V_d(n))$. Prove the following facts about m_d.

i. $m_{2n}(n) = 3^n$.

ii. For all $0 \leq d \leq 2n$, $m_{2n-d}(n) = 3^n - m_d(n)$.

iii. There exists $c \in (2,3)$ such that $m_{2n/3}(n) \leq c^n$.

(b) (2+4+1) Let $A \subseteq \mathbb{F}_3^n$ be 3-AP-free.

i. Show that if $m_d > 3^n - |A|$, then there exists a non-zero $f \in V_d$ such that $f(\mathbf{x}) = 0$ for all $\mathbf{x} \notin A$.

ii. Strengthen the above to show that if $m_d > 3^n - |A|$, then there exists an $f \in V_d$ such that $f(\mathbf{x}) = 0$ for all $\mathbf{x} \notin A$ and f is non-zero on at least $(m_d + |A| - 3^n)$ points in A.

iii. Let $f : \mathbb{F}_3^n \to \mathbb{F}$ such that $f(\mathbf{x}) = 0$ for all $\mathbf{x} \notin A$. Define the matrix $M_f \in \mathbb{F}_3^{A \times A}$ as follows: $M_f(x,y) := f((x+y)/2)$ for all $x, y \in A$. Show that the rank of M_f is exactly $|\{\mathbf{x} \in A | f(\mathbf{x}) \neq 0\}|$.

(c) (4+4) Let $g : \mathbb{F}_3^n \to \mathbb{F}_3$ be a function in $V_d(n)$. Consider the matrix M_g given by $M_g(x,y) := g(x+y)$. Prove the following facts about the rank of the matrix M_g.

i. $\text{rank}(M_g) \leq m_d(n)$.

ii. Strenghten the above to show that $\text{rank}(M_g) \leq 2 \cdot m_{d/2}(n)$.

Hint: Recall that if a $t \times t$-matrix M can be decomposed as $M = UV$ where U is a $t \times r$-matrix and V is a $r \times t$ matrix (or equivalently there exists $2t$ r-dimensional vectors $\mathbf{u}_1, \ldots, \mathbf{u}_t, \mathbf{v}_1, \ldots, \mathbf{v}_t$ such that $M(i,j) = u_i^T v_i$), then $\text{rank}(M) \leq r$.

(d) Conclude from the above parts that if A is 3-AP-free, then $|A| \leq m_{2n-d} + 2m_{d/2}$. Setting $d = 4n/3$ show that $|A| \leq 3c^n$ where c is as in Part 3(a)iii

4. (3+2+2+3) **List-decodability of the Hadamard code via Fourier analysis**

In this problem, we will prove the list-decodability of the Hadamard code via Fourier analysis (this was proved via the Goldreich-Levin theorem in class).

Let $\mathcal{F}$ denote the set of all functions from $\{0,1\}^k$ to $\mathbb{R}$. Note $\mathcal{F}$ is a 2^k-dimensional vector space over $\mathbb{R}$. Define an inner product on this space as follows:

$$\langle f, g \rangle := \mathop{\mathbb{E}}_{\mathbf{x}}[f(x)g(x)].$$

For any $\mathbf{y} \in \{0,1\}^k$, define $\chi_{\mathbf{y}} \in \mathcal{F}$ as follows: $\chi_{\mathbf{y}}(\mathbf{x}) := (-1)^{\sum_{i \in [k]} x_i y_i \pmod 2}$.

(a) Show that for all $\mathbf{y} \neq \mathbf{z}$, we have $\langle \chi_{\mathbf{y}}, \chi_{\mathbf{z}} \rangle = 0$. Conclude that the 2^k functions $\chi_{\mathbf{y}}$ form an orthonormal basis of functions for the vector space $\mathcal{F}$.

Hence, conclude that any function $f \in \mathcal{F}$ can be expressed uniquely as follows:

$$f(\mathbf{x}) = \sum_{\mathbf{y}} \widehat{f}(\mathbf{y}) \cdot \chi_{\mathbf{y}}(\mathbf{x}).$$

where $\widehat{f}(\mathbf{y}) = \langle f, \chi_{\mathbf{y}} \rangle$. These real numbers $\widehat{f}(\mathbf{y})$ are called the Fourier coefficients of f.

(b) (Parsevals' equation). Show that for $f \in \mathcal{F}$, we have $\|f\|_2^2 = \langle f, f \rangle = \sum_{\mathbf{y}} |\widehat{f}(\mathbf{y})|^2$. Hence, for any Boolean function $f : \{0,1\}^k \to \{1,-1\}$, we have $\sum_{\mathbf{y}} |\widehat{f}(\mathbf{y})|^2 = 1$.

It will be convenient to express the range of a Boolean function as $\{1,-1\}$ instead of $\{0,1\}$. We move from $\{0,1\}$ to $\{1,-1\}$ using the transformation $b \mapsto (-1)^b$. Observe that with this notation in place, the $\chi_{\mathbf{y}}$'s exactly correspond to all the linear functions (and thus all the Hadamard codewords).

(c) Let $f : \{0,1\}^k \to \{1,-1\}$ be any Boolean function and $\mathbf{y} \in \{0,1\}^k$ such that $\Pr_{\mathbf{x}}[f(\mathbf{x}) = \chi_{\mathbf{y}}(\mathbf{x})] \geq \frac{(1+\delta)}{2}$. Conclude that $\widehat{f}(\mathbf{y}) \geq \delta$.

(d) Let $f : \{0,1\}^k \to \{1,-1\}$ be any Boolean function. Conclude that there at most $1/\delta^2$ linear functions which have agreement at least $\frac{(1+\delta)}{2}$ with f.

We have thus proved that for any Boolean function f, there are at most $1/\delta^2$ linear functions which are within $\frac{1-\delta}{2}$ fractional distance from f. We had proved this fact via the Goldreich-Levin list-decoding algorithm. The above Fourier-based argument can also be used to construct a list-decoding algorithm for Hadamard codes.

5. (5 + 5 + 5) **Rough polarization without martingales**

In the lectures on polar codes, we studied a sequence of random variables $Z_1, Z_2, \cdots$, with $Z_1 = \alpha$, that evolved as:

$$Z_{n+1} \quad = \quad \begin{cases} 2Z_n - Z_n^2 & \text{with probability } 1/2 \\ Z_n^2 & \text{with probability } 1/2 \end{cases}.$$

(a) Let $\Phi_n = \sqrt{Z_n(1-Z_n)}$. Show that

$$\mathbb{E}[\Phi_{n+1} \mid \Phi_n] \leq \left(\frac{\sqrt{3}}{2}\right) \cdot \Phi_n.$$

(b) Show that for any $\frac{3}{4} < \delta < 1$,

$$\Pr[Z_n(1-Z_n) > \delta^n] \quad \leq \quad \frac{1}{2} \cdot \left(\frac{3}{4\delta}\right)^{n/2}.$$

Hint: Jensen, Markov

(c) For some $\varepsilon > 0$, how high should n be (asymptotically, in terms of ε) to ensure that

$$\Pr[Z_n \in (\varepsilon, 1-\varepsilon)] < \frac{1}{1000}?$$

6. (1+4+5+4+3+3+5) **Sharp thresholds for transitive monotone functions**

For a $p \in [0,1]$, define the probability measure μ_p on length n strings $\mathbf{x} \in \{0,1\}^n$ as

$$\mu_p(\mathbf{x}) := p^{|\mathbf{x}|}(1-p)^{n-|\mathbf{x}|}$$

where $|\mathbf{x}|$ refers to the Hamming weight or the number of non-zero coordinates in $\mathbf{x}$. By abusing notation, define μ_p of a Boolean function $g : \{0,1\}^n \to \{0,1\}$ as

$$\mu_p(g) := \Pr_{\mathbf{x}\sim\mu_p}[g(\mathbf{x}) = 1] = \sum_{\mathbf{x}:g(\mathbf{x})=1} \mu_p(\mathbf{x}).$$

In this problem, we will study the behaviour of $\mu_p(g)$ for monotone functions as p increases from 0 to 1. You may want to check the behaviour for some simple monotone functions: dictators ($\mathrm{DICT}_i(x_1,\ldots,x_n) = x_i$), majority function ($\mathrm{maj}_n$), OR of a few variables (eg., $x_1 \vee x_2$), OR of several variables (eg., $\bigvee_{i=1}^n x_i$).

First for some notation. For any string $\mathbf{x}$, the sensitivity, $\mathrm{sens}_g(\mathbf{x})$, of g at $\mathbf{x}$ is defined as follows:

$$\mathrm{sens}_g(\mathbf{x}) := \left|\{i \in [n] \ : \ g(\mathbf{x}+e_i) \neq g(x)\}\right|,$$

the number of coordinates of $\mathbf{x}$ which when flipped changes the value of g. The averarge sensitivity, $\mathrm{as}_p(g)$ of g with respect to the distribution μ_p is defined as follows:

$$\mathrm{as}_p(g) := \mathop{\mathbb{E}}_{\mathbf{x}\sim\mu_p}\left[\mathrm{sens}_g(\mathbf{x})\right] = \sum_{\mathbf{x}} \mu_p(\mathbf{x}) \cdot \mathrm{sens}_g(\mathbf{x}) = \frac{1}{p} \sum_{\mathbf{x}:g(x)=1} \mu_p(\mathbf{x}) \cdot \left|\{i \in [n] \ : \ g(\mathbf{x}+e_i) = 0\}\right|.$$

(a) Let $k \in [n]$. We say that a Boolean function $g : \{0,1\}^n \to \{0,1\}$ is a k-junta if it depends only on at most k co-ordinates. In other words, there exists a subset $I \in [n]$ of size k such that for all $\mathbf{x}, \mathbf{x}' \in \{0,1\}^n$ such that $\mathbf{x}|_I = \mathbf{x}'|_I$, we have $g(\mathbf{x}) = g(\mathbf{x}')$. Show that if g is a k-junta, then for all p, $\mathrm{as}_p(g) \leq k$.

This shows that juntas have low average sensitivity. Friedgut proved a converse of the above fact in the sense that if g has small average sensitivity, then g is close to some function f which is a junta.

Lemma (Friedgut)**.** *Let $p \in (0,1)$. There exists a constant C_p such that the following is true. Let $g : \{0,1\}^n \to \{0,1\}$ be any Boolean function with and $\delta \in (0,1)$ any approximation parameter. Then there exists another function $f : \{0,1\}^n \to \{0,1\}$ such that (a) f is a $(C_p)^{\mathrm{as}_p(g)/\delta}$-junta and (b) $\Pr_{\mathbf{x}\sim\mu_p}[f(\mathbf{x}) \neq g(\mathbf{x})] \leq \delta$.*

(b) Show that if g is a monotone Boolean function, then $\mu_p(g)$ is an increasing function with p.

The following problem expresses the rate of growth of $\mu_p(g)$ with p in terms of the average sensitivity $\mathrm{as}_p(g)$.

(c) (The Margulis-Russo Lemma) Let $g : \{0,1\}^n \to \{0,1\}$ be any monotone Boolean function. Show that the average sensitivity measures the derivative of the measure μ_p of the function g, i.e.,

$$\frac{d}{dp}\mu_p(g) = \mathrm{as}_p(g).$$

Hint: This can be proved by expanding either side and doing a tedious calculation. For a simpler proof, recall how we proved the *Area Theorem* in class and follow similar ideas.

(d) Let $g : \{0,1\}^n \to \{0,1\}$ be any monotone Boolean function and $\varepsilon \in (0,1)$. Let $p \in (0, 1-\varepsilon)$. Show that there exists a $q \in [p, p+\varepsilon]$ such that $\mathrm{as}_q(g) \leq 1/\varepsilon$.

(e) Use all the above parts and Friedgut's Lemma to conclude the following. Let $g : \{0,1\}^n \to \{0,1\}$ be any monotone Boolean function, $\varepsilon, \delta \in (0,1)$ and $p \in (0, 1-\varepsilon)$. Show that there exists a $q \in [p, p+\varepsilon]$ and a function $f : \{0,1\}^n \to \{0,1\}$ such that (a) f is a $(C_q)^{1/\varepsilon\delta}$-junta and (b) $\Pr_{\mathbf{x}\sim\mu_q}[f(\mathbf{x}) \neq g(\mathbf{x})] \leq \delta$.

In other words, every monotone function is well-approximated by a junta.

(f) Explain why the following does not contradict the previous part (6e). Let n be odd and $\mathrm{maj}_n : \{0,1\}^n \to \{0,1\}$ be the majority function. Let $p = 1/2$. Clearly, maj_n is a monotone function but is certainly not a junta or close to any junta (as it depends on all the variables).

(g) Assume the following fact (a corollary of a result of [Bourgain-Kahn-Kalai-Katznelson-Linial])

Lemma. *There is a constant $c > 0$ such that for every transitive monotone function $g : \{0,1\}^n \to \{0,1\}$ we have*

$$\mathrm{as}_p(g) \quad \geq \quad c \cdot \min(\mu_p(g), \mu_p(\overline{g})) \cdot \log n.$$

Using this lemma, prove the following sharp threshold property of transitive monotone Boolean functions.

Theorem. *There exists a constant $c' > 0$ such that for any transitive monotone function $g : \{0,1\}^n \to \{0,1\}$, if $\mu_p(g) > \varepsilon$ for some $\varepsilon > 0$ then $\mu_q(g) > 1 - \varepsilon$ for some q satisfying*

$$q \quad \leq \quad p + c' \frac{\log(1/2\varepsilon)}{\log n}.$$

Hint: First increase p to q so that $\mu_q(g) = \frac{1}{2}$. Looking at $\int_p^q d(\log \mu_p(g))$ should help.