Exercise List 2

Sivakanth Gopi

Novermber 15, 2019

You are not required to turn in exercises. They are here if you want to practice your understanding of the concepts we discussed in class.

1. In this exercise, you will prove the Schwartz-Zippel lemma which says that a low-degree polynomial cannot have too many roots.

   (a) Let $g(x) \in \mathbb{F}_q[x_1, x_2, \ldots, x_n]$ be a non-zero polynomial where the individual degree of each variable is at most $q-1$. Show that there exists a point $a \in \mathbb{F}_q^n$ s.t. $g(a) \neq 0$. (Hint: Use induction on the number of variables.)

   (b) Let $f(x) \in \mathbb{F}_q[x_1, x_2, \ldots, x_n]$ be a non-zero polynomial of total degree $d < q$. Suppose $f(x) = f_d(x) + f_{d-1}(x) + \cdots + f_0(x)$ where $f_i(x)$ is homogeneous[1] degree i component of f.

   i. Show that there exists some $a^* \in \mathbb{F}_q^n$ s.t. $f_d(a^*) \neq 0$.

   ii. For $z \in \mathbb{F}_q^n$, let $\ell = \{z + \lambda a^* : \lambda \in \mathbb{F}_q\}$ be the line through z in direction a^*. Show that f can have at most d roots on the line ℓ.

   iii. Show that there are exactly q^{n-1} lines in direction a^* and they partition the space $\mathbb{F}_q^n$.

   iv. Combine the above observations to show that

$$\Pr_{z \in \mathbb{F}_q^n}[f(z) = 0] \leq \frac{d}{q}.$$

[1] A homogenous polyomial of degree k only has terms of degree exactly k.

2. In the class we have seen how to construct matching vector families (MVFs) from low-degree polynomial representations of OR mod m. In this exercise, you will show how to construct MVFs from sparse representations of OR mod m over $\{-1, 1\}$ basis. Suppose $p(x) \in \mathbb{Z}_m[x_1, x_2, \ldots, x_n]$ is a polynomial of sparsity[2] s s.t.

   (a) $p(x) = 0 \mod m$ if $x = (1, 1, \ldots, 1)$ and

   (b) $p(x) \neq 0 \mod m$ if $x \in \{-1, 1\}^n \setminus \{(1, 1, \ldots, 1)\}$.

   Show that there exists a MVF over $\mathbb{Z}_m^s$ of size 2^n.

[2] Sparsity is the number of monomials with non-zero coefficients.

3. In this exercise you will prove the edge isoperimetric inequality for the hypercube. Let $G = (\{0, 1\}^n, E)$ be the hypercube graph where $(x, y) \in E$ iff x, y differ in exactly one coordinate. Let $S \subset \{0, 1\}^n$ and let $E(S, S)$ denote the number of edges with both end points in S.

(a) Let $S_0 = \{x \in S : \ x_1 = 0\}$ and $S_1 = \{x \in S : \ x_1 = 1\}$. Show that
$$E(S,S) = E(S_0,S_0) + E(S_1,S_1) + E(S_0,S_1).$$

(b) Show that $E(S_0,S_1) \leq \min(|S_0|,|S_1|)$.

(c) Use induction on the dimension n, to prove that
$$E(S,S) \leq \frac{1}{2}|S|\log_2|S|.$$

(d) Show that the above inequality is tight for subcubes.

4. In this exercise, you will show how to construct 2-query LDCs from q-query LDCs. Let $C : \{-1,1\}^k \to \{-1,1\}^n$ be a 4-query LDC. Let $M_1, M_2, \ldots, M_k$ be q-matchings of size at least $\Omega(n)$ s.t. for every $i \in [k]$ and for every edge $(j_1, j_2, j_3, j_4) \in M_i$,
$$x_i = C(x)_{j_1} C(x)_{j_2} C(x)_{j_3} C(x)_{j_4}.$$
Define $C' : \{-1,1\}^k \to \{-1,1\}^N$ where $N = n^t$, $C'(x) = C(x)^{\otimes t}$ and $t = \sqrt{n}$.

(a) Fix $i \in [k]$. Pick $t = \sqrt{n}$ elements at random from $[n]$ (with repetition). Show that with constant probability you will pick at least two vertices of an edge of M_i.

(b) Use the above fact to construct 2-matchings $M'_1, M'_2, \ldots, M'_k$ on $[N] = [n]^t$ of size $\Omega(N)$ s.t. for every $i \in [k]$ and every edge $(a,b) \in M'_i$,
$$x_i = C'(x)_a C'(x)_b.$$

(c) By applying the 2-query exponential lower bound for C', conclude that
$$n \gtrsim (k/\log k)^2.$$

5. In pseudorandomness, we need to generate a sequence of n bits $X_1, X_2, \ldots, X_n$ which are k-wise independent (and uniform) using as little truly random bits as possible. We want a map (called a pseudorandom generator) $G : \Sigma^r \to \Sigma^n$ s.t. if $(X_1, X_2, \ldots, X_n) = G(U)$ for a uniformly random $U \in \Sigma^r$, then $X_1, X_2, \ldots, X_n$ are k-wise independent and uniform. The goal is to construct an explicit map G where r (called the seed length) is as small as possible. In this exercise, you will show how to do this using error-correcting codes.

(a) Suppose $G : \mathbb{F}^r \to \mathbb{F}^n$ is defined as $G(u) = (\langle v_i, u\rangle)_{i\in[n]}$ for some $v_i \in \mathbb{F}^r$ s.t. every k vectors among $v_1, v_2, \ldots, v_n$ are indepedent.. Show that $(X_1, X_2, \ldots, X_n) = G(U)$ are k-wise independent and uniform if U is chosen uniformly at random from $\mathbb{F}^r$.

(b) Let $C \subset \mathbb{F}^n$ be a linear error correcting code of codimension r (i.e., $\dim(C) = n - r$) and distance at least $k+1$. Let $H_{r\times n}$ be the parity check matrix of C, i.e., $C = \{x \in \mathbb{F}^n : Hx = 0\}$. Let $v_1, v_2, \ldots, v_n$ be the columns of H. Show that every k vectors among $v_1, v_2, \ldots, v_n$ are linearly independent.

(c) Suppose $q \geq n$. Show that there exists a psuedorandom generator $G : \mathbb{F}_q^k \to \mathbb{F}_q^n$ which generates k-wise independent and uniform symbols. (Hint: Reed-Solomon codes)

(d) Suppose there is a code over $\mathbb{F}_2$ with codimension $\lceil \frac{k}{2} \rceil \log n + O(k)$ and distance $\geq k+1$ (BCH codes achieve this). Show that this implies that one can generate n k-wise independent and uniform bits starting from $\lceil \frac{k}{2} \rceil \log n + O(k)$ truly random bits! Thus for small k, we have an exponential improvement!

Let $n = 2^m$. BCH code in $\mathbb{F}_2^n$ of distance D is obtained by taking all codewords with $\mathbb{F}_2$-coordinates from the Reed-Solomon code in $\mathbb{F}_n^n$ of distance D. Clearly, it will have distance at least D. It is non-trivial to show that it will have codimension at most $\lceil \frac{D-1}{2} \rceil \log n$. For constant D, BCH codes nearly achieve the Hamming bound we proved in the beginning of the course. So they are nearly optimal binary codes for constant distance.

6. In this exercise, you will construct ε-biased sets from codes. A subset $S \subset \mathbb{F}_2^k$ is called an ε-biased set if for every $z \in \mathbb{F}_2^k \setminus \{0\}$,

$$\left| \mathbb{E}_{x \in S}[(-1)^{\langle z,x \rangle}] \right| \leq \varepsilon.$$

Note that $S = \mathbb{F}_2^k$ is 0-biased. Our goal is to construct an ε-biased set of small size. Suppose $C : \mathbb{F}_2^k \to \mathbb{F}_2^n$ be an linear code s.t. every codeword has Hamming weight between $(\frac{1-\varepsilon}{2})n$ and $(\frac{1+\varepsilon}{2})n$.[3] Let $G_{n\times k}$ be the generator matrix of C i.e. $C = \{Gx : x \in \mathbb{F}_2^k\}$. Let $u_1, u_2, \ldots, u_n$ be the rows of G.

(a) Show that $S = \{u_1, u_2, \ldots, u_n\} \subset \mathbb{F}_2^k$ is an ε-biased set.

(b) Show that there exist ε-biased sets of size $O(k/\varepsilon^2)$.

[3] This implies that minimum distance is at least $(\frac{1}{2} - \varepsilon)n$, but this is a little stronger.