CMSC818K/ENEE729A. Final examination **Instructor: A. Barg**

Posted on 12/14/2020, 12pm. Date due 12/17/2020, 12pm.

Please submit your work as a **single PDF file** to ELMS/Canvas under the Assignments tab

- Papers submitted as multiple pictures of individual pages are difficult for grading and **will not be accepted**.
- Justification of solutions is required.
- Max score 60 points

Problem 1. (20pts) The proofs in this problem rely on random codes with independently chosen coordinates of the codewords.

(A) Call a binary code $\mathcal{C} \subset \{0,1\}^n$ $(2,2)$-separating if for any 4 codewords $a_1, a_2, b_1, b_2 \in \mathcal{C}$ there exists a coordinate i such that either $(a_{1,i}, a_{2,i}) = (0,0)$ and $(b_{1,i}, b_{2,i}) = (1,1)$ or $(a_{1,i}, a_{2,i}) = (1,1)$ and $(b_{1,i}, b_{2,i}) = (0,0)$.

(1) Bound above the average number of subsets $\{(a_1, a_2), (b_1, b_2)\}$ that violate the separating property.

(2) Define

$$R_{2,2} = \liminf_{n\to\infty} \max_{\substack{\mathcal{C}\subset\{0,1\}^n \\ \mathcal{C} \text{ is (2,2) separating}}} \frac{\log_2 |\mathcal{C}|}{n}$$

Prove that

$$R_{2,2} \geq 1 - \frac{1}{3}\log_2 7.$$

(Hint: the method is from HW2, problem 1; note that to break up a 4-tuple that violates the separating property it suffices to delete one vector from the code.)

(B) Let $Q, |Q| = q$ be a finite set. A perfect t-hash family of functions is defined as a code $\mathcal{C} \subset Q^n$ such that for every t-tuple of vectors $x_1, x_2, \ldots, x_t \in \mathcal{C}$ there is a coordinate $i \in [n]$ such that $|\{x_{1,i}, x_{2,i}, \ldots, x_{t,i}\}| = t$ (all the entries are distinct). Such coordinates are called *hash*. We say that $\mathcal{C}$ has *t-hash distance* d_t if every t-tuple has $\geq d_t$ hash coordinates. In this problem you will prove a lower bound on the rate of codes with a given t-hash distance.

(1) Prove that for a given t-tuple of codewords of $\mathcal{C}$ the probability that a given coordinate is hash equals $a_t := \prod_{i=1}^{t-1}(1 - i/q))$.

(2) Let P_b be the probability that a t-tuple of codewords contains $\leq d_t - 1$ hash coordinates. Use the Chernov-Hoeffding inequality to argue that

$$P_b \leq e^{-2n(\delta - a_t)^2}.$$

(3) Define

$$R_t(\delta) = \liminf_{n\to\infty} \max_{\substack{\mathcal{C}\subset\{0,1\}^n \\ \mathcal{C} \text{ has } t\text{-hash dist.} \geq \delta n}} \frac{\log_q |\mathcal{C}|}{n}$$

Prove that

$$R_t(\delta) \geq 2\frac{(\delta - a_t)^2}{(t-1)\ln q}.$$

Problem 2. (10pts) Let $G(V, E)$ be a simple undirected graph on n vertices (i.e., $|V| = n$). An edge $e \in E$ connects a pair of vertices u, v, and we denote this by $e = (u, v)$. A cycle in G is a sequence of edges $e_1 = (v_1, v_2), e_2 = (v_2, v_3), \ldots, e_t = (v_t, v_1)$, where all the vertices $v_2, \ldots, v_t$ are distinct.

(a) Consider characteristic vectors of cycles $c \in \{0,1\}^{|E|}$, where $c_i = 1$ if the edge e_i is contained in the cycle and 0 if not. Consider an $\mathbb{F}_2$-linear space spanned by these vectors, i.e., a binary code. Find the dimension and the distance of this code.

(b) The dual code of the code in part (a) can be also described in terms of characteristic vectors of some well-defined objects in G. Give a precise description of this code in graph-theoretic terms, and find its dimension and distance.

Problem 3. (10pts) (a) Given a binary linear code $\mathcal{C}$ of length n. Prove that the average weight of codewords $\sum i(A_i/|\mathcal{C}|)$ is $n/2$ if the dual distance $d^\perp > 1$. Find the average weight of codewords in $\mathcal{C}$ if $d^\perp = 1$. Find the second moment of the weight distribution of the code. (Hint: start with the MacWilliams identities in the form $\sum_{i=0}^n A_i y^i = \frac{1}{|\mathcal{C}^\perp|} \sum_{i=0}^n A_i^\perp (1+y)^{n-i}(1-y)^i$.)

(b) Given a linear code $\mathcal{C}$ of length n and dimension k over a field $\mathbb{F}_q$ with dual distance $d^\perp$. Let $I \subset [n], |I| = t \leq d^\perp - 1$ be a subset of coordinates. Write all the codewords in a $q^k \times n$ matrix and consider the submatrix formed by the columns with indices in I. Prove that every q-ary vector $x \in \mathbb{F}_q^t$ appears as a row of this submatrix the same number of times, i.e., q^{k-t} times.

Problem 4. (20pts)

Let $F = \mathbb{F}_q$ and let $\alpha \in \mathbb{F}_{q^2}$ be an element of multiplicative order $q+1$. Denote by $m_i(x), i = 0, 1, \ldots, q$ the minimal polynomial of α^i over F (this is a polynomial of the smallest degree with coefficients in F and such that $m_i(\alpha^i) = 0$).

(a) Show that the polynomial $x^{q+1} - 1$ can be decomposed into linear factors over $\mathbb{F}_{q^2}$ (i.e., that all the roots of $x^{q+1} - 1$ lie in $\mathbb{F}_{q^2}$).

(b) Show that $m_i(x) = (x - \alpha^i)(x - \alpha^{-i})$ for all $i \neq 0, (q+1)/2$.

(c) Let q be odd. Prove that for every odd $k, 1 \leq k \leq q$, the cyclic code of length $q+1$ over F with generator polynomial

$$\prod_{i=0}^{(q-k)/2} m_i(x)$$

has dimension k and is MDS.

(d) Let q be even. Prove that for every $k, 1 \leq k \leq q+1$ there exists a cyclic $[q+1, k]$ MDS code over F.