Coding Theory : Midsem Exam

---

1. All the statements proven in the class can be assumed without proofs.
2. To solve a sub-problem of a particular problem in the question paper, you can assume all its previous sub-problems without proof.
3. Statements mentioned in the appendix section of the question paper can be assumed without proof.
4. Other than that, anything you use needs to be proven.
5. Total marks: **50**
6. Duration: **Two hours**

---

1. A set of $S \subseteq \mathbb{F}_2^k$ vectors is called *ϵ-biased sample space* if the following property holds: Pick a vector $X = (x_1, x_2, \ldots, x_k)$ uniformly at random from S. Then, X has bias at most ϵ, that is, for every nonempty subset $I \subseteq [k]$,

$$\left| \Pr\left(\sum_{i \in I} x_i = 0 \right) - \Pr\left(\sum_{i \in I} x_i = 1 \right) \right| \leq \epsilon,$$

where the sum is over $\mathbb{F}_2$. Observe that $S = \mathbb{F}_2^k$ is an ϵ-biased sample space with $\epsilon = 0$. In this problem, we will look at some connections of ϵ-biased sample space to linear codes over $\mathbb{F}_2$.

   (a) **(5 marks)** Let C be an $[n,k]_2$ code such that all non-zero codewords have Hamming weight in the range $\left[\left(\frac{1-\epsilon}{2}\right)n, \left(\frac{1+\epsilon}{2}\right)n\right]$. Let $G \in \mathbb{F}_2^{k \times n}$ be a generator matrix of C. Then, show that the set of columns of G forms an ϵ-biased sample space of size n.

   (b) **(10 marks)** Let C be an $[n,k]_2$ code such that all nonzero codewords have Hamming weight in the range $\left[\left(\frac{1-\gamma}{2}\right)n, \left(\frac{1+\gamma}{2}\right)n\right]$ where $\gamma \in (0,1)$. Then, show that for every odd positive integer m, there exists an $[n^m, k]_2$ code C' such that all nonzero codewords have Hamming weight in the range

$$\left[\left(\frac{1-\gamma^m}{2}\right)n^m, \left(\frac{1+\gamma^m}{2}\right)n^m\right].$$

   (c) **(8 marks)** Let C be an $[n,k]_2$ code such that all nonzero codewords have Hamming weight in the range $\left[\left(\frac{1-\gamma}{2}\right)n, \left(\frac{1+\gamma}{2}\right)n\right]$ where $\gamma \in (\epsilon, 1)$. Then, show that there exists an ϵ-biased sample space of size

$$n^{O\left(\frac{\log 1/\epsilon}{\log 1/\gamma}\right)}.$$

2. **(7 marks)** For any $[n, k, n-k+1]_q$ code, show that its dual is an $[n, n-k, k+1]_q$ code.

3. Let $q \geq 2$ be an integer. As we have seen in the class, the *Gilbert-Varshamov bound* (GV bound) says that for every $\delta \in [0, 1 - \frac{1}{q}]$, there exists a q-ary code with the rate $R \geq 1 - H_q(\delta)$ and relative distance δ, where $H_q(\cdot)$ denotes the q-array entropy function defined in the class. In the class, we also saw a greedy construction-based proof for GV bound. Here, see a graph-theoretic proof for GV bound. Let $d = \delta n$, and Σ be an alphabet of size q. Let $G_{n,d,q} = (V, E)$ be a graph whose vertex set is Σ^n. Given vertices $\mathbf{u} \neq \mathbf{v} \in \Sigma^n$, we have the edge $\{\mathbf{u}, \mathbf{v}\} \in E$ if and only if $\Delta(\mathbf{u}, \mathbf{v}) < d$. A subset $I \subseteq V$ of vertices is called an *independent set* of $G_{n,d,q}$, if for every $\mathbf{u} \neq \mathbf{v} \in I$, $\{\mathbf{u}, \mathbf{v}\} \notin E$. Then, solve the following sub-problems.

   (a) (**5 marks**) Show that any independent set C of $G_{n,d,q}$ is a q-ary code of distance at least d.

   (b) (**10 marks**) The *degree* of vertex in a graph $G = (V, E)$ is the number of edges incident on that vertex. Let D be the maximum degree of any vertex in $G = (V, E)$. Then argue that G has an independent set of size at least $\frac{|V|}{D+1}$.

   (c) (**5 marks**) Using the parts (a) and (b), prove the GV bound.

Appendix

1. **Reducing bias:** Let $S = (s_1, s_2, \ldots, s_n) \in \{0,1\}^n$ be a binary string with pn many 1's for some $p \in (0, 1)$. For some positive integer m, let $S' \in \{0,1\}^{n^m}$ be a binary string defined as follows: Let the coordinates of the string S' be denoted by the elements of $[n]^m$. Then,

$$S' = \left(\bigoplus_{j=1}^{m} s_{i_j} \right)_{(i_1, i_2, \ldots, i_m) \in [n]^m}$$

   Then, the number of 1's in the string S' is

$$\Delta_p = \frac{1}{2} \cdot (1 - (1 - 2p)^m) \cdot n^m.$$

   Furthermore, if m is odd, then Δ_p is a non-decreasing function of p.

Coding Theory : Endsem Exam

M Tech (CS), 2025-26

Exam Date: 26/11/2025 Time: 2:30 PM-5:30 PM Total Marks: 60

---

1. All the statements proven in the class can be assumed without proofs.
2. To solve a sub-problem of a particular problem in the question paper, you can assume all its previous sub-problems without proof.
3. Other than that, anything you use needs to be proven.
4. During examination, *only* handwritten notes are allowed, printouts/electronic notes are *not* allowed. Calculators are also *not* allowed during the examination.

---

1. Let $\mathbb{F}_q$ be the finite field of size q. Let $E = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a subset of $\mathbb{F}_q$, that is, $n \leq q$. Let $k \in \mathbb{Z}_{\geq 1}$ such that $k \leq n$. Let $\mathrm{RS}(E, k, q)$ be the set of Reed-Solomon codes of block length n, dimension k, and alphabet $\mathbb{F}_q$, defined using E as the evaluation points. That is,

$$\mathrm{RS}(E, k, q) = \Big\{(f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_n)) \mid f(x) \in \mathbb{F}_q[x] \text{ such that } \deg(f) < k\Big\}.$$

Let $n, \ell, \delta \in \mathbb{Z}_{\geq 1}$ with $\gcd(n, q) = 1$, $n \mid (q^m - 1)$, and $\delta \leq n$. Let $\beta \in \mathbb{F}_{q^m}$ be an n-th primitive root of unity in $\mathbb{F}_{q^m}$, that is, $\beta^n = 1$ but for any integer $0 < m < n$, $\beta^m \neq 1$. Then, the BCH code over the alphabet $\mathbb{F}_q$ is defined as follows:

$$\mathrm{BCH}(n, \delta, q, \ell) = \left\{(c_0, c_1, \ldots, c_{n-1}) \in \mathbb{F}_q^n \,\middle|\, \forall\, i \in \{0, 1, 2, \ldots, \delta - 2\},\ \sum_{j=0}^{n-1} c_j \beta^{(\ell+i)j} = 0,\right\}.$$

   (a) **(20 marks)** Consider the code $\mathrm{RS}(E, k, q)^\perp$, that is, the dual of $\mathrm{RS}(E, k, q)$. Design an error-correction algorithm $\mathcal{A}$ for $\mathrm{RS}(E, k, q)^\perp$ that runs in $\mathrm{poly}(n)$ $\mathbb{F}_q$-operations and can correct less than $\frac{k+1}{2}$ many errors. More specifically, given a $\mathbf{y} = (y_1, y_2, \ldots, y_n) \in \mathbb{F}_q^n$ as input to $\mathcal{A}$ with the promise that there exists a codeword $\mathbf{c} \in \mathrm{RS}(E, k, q)^\perp$ such that $\Delta(\mathbf{y}, \mathbf{c}) < \frac{k+1}{2}$, it outputs $\mathbf{c}$ in $\mathrm{poly}(n)$ $\mathbb{F}_q$-operations.

   (b) **(10 marks)** Design an error-correction algorithm $\mathcal{B}$ for $\mathrm{BCH}(n, \delta, q, 1)$ that runs in $\mathrm{poly}(n)$ $\mathbb{F}_q$-operations and can correct less than $\frac{\delta}{2}$ many errors. More specifically, given a $\mathbf{y} = (y_1, y_2, \ldots, y_n) \in \mathbb{F}_q^n$ as input to $\mathcal{B}$ with the promise that there exists a codeword $\mathbf{c} \in \mathrm{BCH}(n, k, q)^\perp$ such that $\Delta(\mathbf{y}, \mathbf{c}) < \frac{\delta}{2}$, it outputs $\mathbf{c}$ in $\mathrm{poly}(n)$ $\mathbb{F}_q$-operations.

2. Let $G = (L, R, E)$ be a c-left-regular and d-right-regular bipartite graph with the left vertex set $L = [n]$, the right vertex set $R = [m]$, and the set of edges $E \subseteq L \times R$. For any $r \in R$ and $j \in [d]$, let $N_j(r)$ denote the j-th smallest neighbor of r, that is, if $\{i_1, i_2, i_3, \ldots, i_d\} \subset L$ be the set of neighbors of r with $i_1 < i_2 < \cdots < i_d$, then $N_j(r) = i_j$. Let $\Sigma = \{0, 1\}^d$. For all $\mathbf{u} = (u_1, u_2, \ldots, u_n) \in \{0, 1\}^n$, define $G(\mathbf{u}) \in \Sigma^m$ as follows: For all $r \in [m]$, $G(\mathbf{u})_r$, the r-th coordinate of $G(\mathbf{u})$, is

$$G(\mathbf{u})_r = \left(u_{N_1(r)}, u_{N_2(r)}, \ldots, u_{N_d(r)}\right).$$

Let $C \subseteq \{0, 1\}^n$ be a binary code of block length n. Let $G(C) \subseteq \Sigma^m$ be a code over the alphabet Σ, defined as follows:

$$G(C) = \Big\{G(\mathbf{c}) \,\Big|\, \mathbf{c} \in C\Big\}.$$

Now show the following.

   (a) **(10 marks)** $R(G(C)) = \frac{1}{c} \cdot R(C)$, where $R(C)$ and $R(G(C))$ are the rates of C and $G(C)$, respectively.

   (b) **(10 marks)** Suppose that G and C satisfies the following property: For some $\gamma, \epsilon \in (0, 1)$, let $\mathrm{dist}(C) \geq \gamma n$, and for every $S \subseteq L$ with $|S| \geq \gamma n$, let $|N(S)| \geq (1 - \epsilon)m$, where $N(S)$ denotes the set of neighbors of S. Then,

$$\mathrm{dist}(G(C)) \geq (1 - \epsilon)m.$$

3. **(10 marks)** Let q be a prime power. Let C_{in} be an $[n, k]_q$ code and C_{out} be an $[N, K]_Q$ with $Q = q^k$. Let $C = C_{\mathrm{in}} \circ C_{\mathrm{out}}$ be the concatenated code of C_{out} and C_{in}. As discussed in the class, C is an $[Nn, Kk]_q$ code. Let $k < n$. Then, show that

$$\mathrm{dist}\left(C^\perp\right) \leq k + 1,$$

where $C^\perp$ is the dual code of C.