Homework #11. Due Tue, May 4th by 11:59pm in filedrop

Reading for this homework:

Online Lectures 23-25 and Chapter 7 in Lindell's notes.

Problems:

1.

(a) Deduce from the binary Plotkin bound that for any $n, d \in \mathbb{N}$ with $n \leq 2d$ we have $A_2(n, d) \leq 4d$.

(b) Now assume that $n > 2d$. Prove that $B_2(n, d) \leq 2^{n-2d+2} \cdot d$, that is, for any binary linear code C of length n and distance d we have $|C| \leq 2^{n-2d+2} \cdot d$.

Hint: Given $m \leq n$, let us think of $\mathbb{F}_2^m$ as the subspace of $\mathbb{F}_2^n$ consisting of all vectors whose last $n - m$ coordinates are zero. Now consider the code $C' = C \cap \mathbb{F}_2^{2d}$. Show that the size of C' can be bounded above using the Plotkin bound. Then use the fact that $\dim(U \cap W) = \dim(U) + \dim(W) - \dim(U + W)$ for any vector subspaces U and W of a finite-dimensional vector space V, deduce the desired bound on $|C|$.

(c) Now let $\mathcal{C} = \{C_m\}_{m=1}^{\infty}$ be a sequence of binary linear codes. Assume that C_m is $[n_m, k_m, d_m]$-linear where $n_m \to \infty$ as $m \to \infty$. Also assume that the asymptotic relative distance $\delta(\mathcal{C}) = \liminf_{m\to\infty} \delta(C_m) = \liminf_{m\to\infty} \frac{d_m - 1}{n_m}$ satisfies $\delta(\mathcal{C}) \geq \frac{1}{2}$. Use (a) and (b) to prove that $R(\mathcal{C}) = 0$.

Recall that $R(\mathcal{C}) = \liminf_{m\to\infty} R(C_m) = \liminf_{m\to\infty} \frac{k_m}{n_m}$. If you are not comfortable with liminf, you may assume that the limit in the definition of $\delta(\mathcal{C})$ exists.

2. Prove Lemma 24.1 from online notes.

3. Prove that an $[n, k, d]$-linear code over $\mathbb{F}_q$ satisfying the Gilbert-Varshamov bound can be constructed using at most $q^{n-k} \cdot k(n - k)$ additions in $\mathbb{F}_q$.

Hint: For the proof of the Gilbert-Varshamov bound see 5.2.2 in the book or online Lecture 23 (up to minor variations, both are essentially the same as the proof we gave in class in Lecture 15). To prove the above bound estimate (1) the number of columns of PCM that need to be computed, (2) the number of prohibited vectors at each step and (3) the number of additions needed to calculate each prohibited vector.

4. Let $B = PCC_3$, the parity-check code of length 3, and let A be the zero-sum code of length 4 over $\mathbb{F}_4$. Verify that the concatenated code $B \circ A$ is defined and compute its length, dimension, distance AND a generator matrix (equivalently, find its basis).

Hint: If φ is the map from the definition of the concatenated code (in the notations from class) and S is a basis of A, then $\varphi(S)$ is linearly independent, but it will usually not be a basis for $\varphi(A) = B \circ A$. However, it is easy to describe a basis for $B \circ A$ in terms of S and multiplication in $\mathbb{F}_{q^k}$ (the field over which A is defined).

Warning: The notations in class differ from those in online Lecture 25 and Lindell's Section 7.2. The map φ from class is called φ_* in online notes and Lindell's notes. It is not hard to see that the map φ in the online notes is given by $\varphi(v) = R(v)G_B$ where R is the restriction of scalars and G_B is the generator matrix for B whose rows are elements of the chosen basis.

5. Let $A = B = Rep(1, 5)$, the simple binary repetition code of length 5 (thus $Rep(1, 5) = \{0^5, 1^5\}$).

(a) Prove that the concatenated code $B \circ A$ is defined and is equal to $Rep(1, 25)$.

The remainder of the problem investigates the following decoding rules for $B \circ A$:

- NND decoding (treating $B \circ A$ as a single code)
- N2SD decoding – naive 2-stage decoding as defined in Lecture 27 on April 28.
- EED_s decoding – the errors and erasures decoding with the threshold $e_{min} = s$ (the definition, which we briefly discussed in class, is recalled below). In this example we can consider EED_s for $s = 0, 1$ and 2.

Note that in this example NND can correct up to $\lfloor \frac{25-1}{2} \rfloor = 12$ errors, while N2SD can correct up to $(\lfloor \frac{5-1}{2} \rfloor + 1)^2 - 1 = 8$ errors (as proved in Lecture 27).

(b) Explain why in this example EED_2 coincides with N2SD. What property of $Rep(1, 5)$ does this reflect?

(c) Give an example where 9 errors occur during the transmission, N2SD works incorrectly while EED_1 works correctly.

(d) Give an example where 12 errors occur during the transmission, N2SD works correctly while EED_1 works incorrectly.

(e) Now give an example where 12 errors occur during the transmission, N2SD and EED_1 both work incorrectly, but EED_0 works correctly

(f) (bonus) Prove that if ≤ 12 erros occur during the transmission, then at least 1 of the following 3 rules – N2SD, EED_1 or EED_0 works correctly.

Definitions of N2SD and EED_s. We first briefly recall the definition of N2SD applied to the concatenated code $B \circ A$ where B is $[n, k]$-linear over $\mathbb{F}_q$ and A is $[N, m]$-linear over $\mathbb{F}_{q^k}$. Take the received word $w \in \mathbb{F}_q^{nN}$, write it as $w = w_1 \dots w_N$ with $w_i \in \mathbb{F}_q^n$, and for each i decode w_i using NND for B to get the word $\gamma = \gamma_1 \dots \gamma_N$ with $\gamma_i \in \mathbb{F}_q^n$. Then write each γ_i as $\gamma_i = R(v_i)G_B$ (where $R : \mathbb{F}_{q^k} \to \mathbb{F}_q^k$ is the restriction of scalars) and decode $v = v_1 \dots v_N \in \mathbb{F}_{q^k}^N$ using NND for A.

Now fix a non-negative integer s and define EED_s as follows. For each $1 \leq i \leq N$ define γ_i as in N2SD, compute $d(\gamma_i, w_i)$ and call γ_i *questionable* if $d(\gamma_i, w_i) > s$. Now let v' be the word obtained from $v = v_1 \dots v_N$ defined in N2SD by replacing v_i by the erasure symbol ? (we will use ? instead of the symbol used in class) whenever γ_i is questionable. Now decode v' using NND for A ignoring the positions with the erasure symbol. Formally this means that we remove the erasure symbols ? from v' and then decode using NND for A' where A' is the code obtained from A by puncturing the coordinates where the erasure symbols occurred.

Let us see an example how this works for A and B in this problem. Suppose the received word is $w = (1^40)(1^30^2)(1^30^2)(1^40)(10^4)$. N2SD will first decode it to $(1^5)(1^5)(1^5)(1^5)(0^5)$ (this is our γ). The corresponding v is then 11110 which will be decoded to $1^5 \in A$ using NND for A. If we use EED_1, v' will be 1??10 since we have to correct more than 1 error when B-decoding w_2 and w_3. Since this word has more 1's than 0's, it will still be decoded to 11111.

Now suppose $w = (1^40)(1^30^2)(1^30^2)(1^30^2)(10^4)$. If we apply N2SD, v and the decoded word will be the same as in the previous case. However, if we apply EDD_1, v' will be 1???0 – in this case NND for A will randomly choose between 0^5 and 1^5.