MTAT.05.082: Introduction to Coding Theory **University of Tartu**

Final exam

Instructors: Vitaly Skachek, Yauhen Yakimenka December 28th, 2016

Student name: _____

Student ID: _____

1. This exam contains 10 pages. Check that no pages are missing.
2. It is possible to collect up to 120 points. Try to collect as many points as possible.
3. Justify and prove all your answers.
4. All facts and results that were proved or stated in the class can be used in your solution without a proof. Such results need to be rigorously formulated.
5. Any printed and written material is allowed in the class. No electronic devices are allowed.
6. Exam duration is 3 hours.
7. Good luck!

Question 1	
Question 2	
Question 3	
Question 4	
Total	

Question 1 (20 points).

A code $\mathcal{C}$ is defined as the following set of vectors over $\mathbb{F}_5 = \{0, 1, 2, 3, 4\}$:

$$\mathcal{C} = \{\mathbf{c} \mid \boldsymbol{H}\mathbf{c}^\intercal = \mathbf{0}^\intercal\} \ ,$$

where

$$\boldsymbol{H} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 & 2 \\ 0 & 0 & 1 & 3 & 3 \\ 0 & 0 & 0 & 4 & 4 \end{pmatrix} .$$

(a) What is the length n, dimension k and minimum distance d of the code $\mathcal{C}$? Justify your answer.

(b) Find a generator matrix of the code $\mathcal{C}$.

Question 2 (30 points).

Definition: An (n, M, d) code $\mathcal{C}$ over $\mathbb{F} = \mathbb{F}_q$, $q = n$, is called a *permutation code* if each codeword in $\mathcal{C}$ has n different symbols from $\mathbb{F}$. (Please note that the code is generally non-linear.)

Example: for $n = 3$, $\mathbb{F} = \mathbb{F}_3 = \{0, 1, 2\}$, the following code is a permutation code of Hamming distance 3:

$$\{(012), (201), (120)\} \ .$$

Let $\mathcal{C}$ be a permutation code of length n over $\mathbb{F}$.

(a) Show that for any two codewords $\bar{\boldsymbol{x}}$ and $\bar{\boldsymbol{y}}$ in $\mathcal{C}$, $\bar{\boldsymbol{x}} \neq \bar{\boldsymbol{y}}$, the Hamming distance $\mathsf{d}_H(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}) \geq 2$.

(b) What is the maximal possible size of $\mathcal{C}$ for $d = 2$?

(c) Let $\bar{\boldsymbol{x}} \in \mathcal{C}$. Show that the maximal possible number of codewords in $\mathcal{C}$ at the Hamming distance at most $r = 2$ from $\bar{\boldsymbol{x}}$ is

$$1 + \binom{n}{2} \ .$$

(d) What is the maximal possible size of a permutation code $\mathcal{C}$ with $d = 5$? Derive an analogue of the sphere-packing (Hamming) bound for the permutation codes.

Question 3 (40 points).

Let $\mathbb{F} = \mathbb{F}_2$ be a finite field with two elements and $\mathcal{C}$ be an $[n, k, d]$ code over $\mathbb{F}$, defined using the following generator matrix

$$\boldsymbol{G} = \left(\begin{array}{c|c} 0\,0\,\cdots\,0 & 1\,1\,\cdots\,1 \\ \hline \boldsymbol{G}_1 & \boldsymbol{G}_2 \end{array}\right),$$

where the number of 1's in the first row equals d. Let $\mathcal{C}_1$ be the linear $[n_1 = n - d, k_1, d_1]$ code over $\mathbb{F}$, which is spanned by the rows of $\boldsymbol{G}_1$, and $\mathcal{C}_2$ be the linear $[n_2 = d, k_2, d_2]$ code over $\mathbb{F}$, which is spanned by the rows of $\boldsymbol{G}_2$.

(a) Take

$$\boldsymbol{G}_1 = \boldsymbol{G}_2 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

What are the corresponding values of n, k and d?

(b) (1) Show that if the rows of $\boldsymbol{G}_1$ are linearly dependent, then there exist two different non-zero codewords $\bar{\boldsymbol{b}}, \bar{\boldsymbol{c}} \in \mathcal{C}$, whose first n_1 coordinates are zeros (one of them is the first row of $\boldsymbol{G}$).

(2) By using the result in (b-1), show that if the rows of $\boldsymbol{G}_1$ are linearly dependent, then there exists a codeword $\bar{\boldsymbol{b}} \in \mathcal{C}$ with weight strictly less than d. Observe a contradiction, and conclude that the rows of $\boldsymbol{G}_1$ must be linearly independent.

(c) (1) Show that if $(\bar{\boldsymbol{c}}_1|\bar{\boldsymbol{c}}_2) \in \mathcal{C}$, where $\bar{\boldsymbol{c}}_1 \in \mathcal{C}_1$, then either $\bar{\boldsymbol{c}}_2 \in \mathcal{C}_2$ or $\bar{\boldsymbol{c}}_2 + \bar{\boldsymbol{e}} \in \mathcal{C}_2$, where $\bar{\boldsymbol{e}} = (1\,1\,\cdots\,1)$. Conclude that if $\bar{\boldsymbol{c}}_1$ has weight exactly d_1, then both $\bar{\boldsymbol{c}}_2$ and $\bar{\boldsymbol{c}}_2 + \bar{\boldsymbol{e}}$ have weight at least $d - d_1$.

(2) Show that $d_1 \geq d_2$.

Question 4 (30 points).

Let $\mathbb{F} = \mathbb{F}_7$ be a field of integer residues modulo 7. Suppose that $\mathcal{C}$ is a $[6, 2, 5]$ Reed-Solomon code over $\mathbb{F}$, with a parity-check matrix of the code given by

$$\boldsymbol{H} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ 1^2 & 2^2 & 3^2 & 4^2 & 5^2 & 6^2 \\ 1^3 & 2^3 & 3^3 & 4^3 & 5^3 & 6^3 \end{pmatrix}.$$

This means that the code locators are $\alpha_1 = 1$, $\alpha_2 = 2$, $\alpha_3 = 3$, $\alpha_4 = 4$, $\alpha_5 = 5$, $\alpha_6 = 6$, and the column multipliers are $v_1 = v_2 = v_3 = v_4 = v_5 = v_6 = 1$.

Assume that $\bar{\boldsymbol{c}} \in \mathcal{C}$ is transmitted, and $\bar{\boldsymbol{y}} = (3, 1, 5, 6, 3, 1) \in \mathbb{F}^6$ is received. In this question, you will decode $\bar{\boldsymbol{y}}$.

(a) Find the syndrome polynomial $S(x)$.

(b) Find the error-locator and the error-evaluator polynomials by either Peterson-Gorenstein-Zierler method or by Euclid's algorithm.

(c) What are the error locations and error values?

(d) What is $\bar{\boldsymbol{c}}$ if we assume that there were at most $\lfloor (d-1)/2 \rfloor$ errors?

(e) Compute $\boldsymbol{H} \cdot \bar{\boldsymbol{c}}^T$ and show that indeed $\bar{\boldsymbol{c}} \in \mathcal{C}$.