# Introduction to Coding Theory (January 2007) : Final exercise

Ronen Shaltiel

January 22, 2007

## Instructions

**General:**

- You must hand in printed solutions. (I'm willing to accept calculations in handwriting if you don't know how to print them but the rest must be printed).
- Please write clearly and precisely!
- I will not accept solutions that are longer than 10 pages.

**Rules:**

- This is a test! You are not allowed to collaborate with other people and must do the work on your own. For some of the questions it may be the case that there are solutions on-line. Don't use them!
- Please include in your submission a **signed statement** saying that: I (*include name and ID*) hereby declare that I did not discuss this project with any other people and the solution that I am submitting is my own work.

**Grading:**

- There are 7 questions and you are supposed to answer only 3 of them. Each question $i$ comes with two numbers $(a_i, b_i)$. The final score will be given by summing the $a_i$'s of the questions you answer to compute $a$, summing the $b_i$'s of the questions you answer to compute $b$ and the final score is $min(a, 70) + b$.
- A 5 point bonus will be given to submissions which are clear and well written.
- You are allowed to answer one additional question over the 3 questions required (that is you can answer 4 questions), and I will base the scoring on the best 3 questions.

Good Luck!

## Definitions

This section gives additional definitions needed to solve the questions:
We use $[n]$ to denote the set $\{1, \dots, n\}$.

**Definition 1 (Dual code).** *Given a linear $[n, k, d]_q$-code $C$ we define its dual $C' = \left\{y \in F_q^n : \forall x \in C,\ < x, y >= 0\right\}$.*

**Definition 2 (1-way randomized communication protocols for comparison with 1-sided error).** *A 1-way randomized protocol for comparison is a protocol between two players: Alice that holds a string $x \in \{0,1\}^k$ and Bob that holds a string $y \in \{0,1\}^k$. The purpose of the protocol is that Alice and Bob determine whether $x = y$ while sending as few bits as possible. The protocol is defined given a function $m : \{0,1\}^k \times [n] \to [q]$ where $n$ is called the* randomness size *and $q$ is called the* message size*. The protocol works as follows: Alice chooses a random element $i$ in $[n]$, and computes $m(x, i)$ and sends $m(x, i)$ and $i$ it to Bob. Bob then decides whether $x = y$ based on $y$ and $m(x, i)$ and $i$. To completely define the protocol one needs to choose a specific rule for Bob.*

*The protocol has success probability $1 - \epsilon$ if Bob always decides correctly when $x = y$ and for every $x \neq y$, $\Pr[$Bob decides incorrectly$] < \epsilon$ where the probability os over the choice of $i$.*

**Definition 3 ($\ell$-wise independence).** *A distribution $X = (X_1, \ldots, X_n)$ over $n$ bit strings is $\ell$-wise independent if for any $i_1, \ldots, i_\ell \in [n]$ the distribution $X_{i_1}, \ldots, X_{i_\ell}$ is uniformly distributed.*

**Definition 4 (bit-fixing sources).** *A set $S \subseteq \{0,1\}^n$ is called a bit-fixing source with threshold $\ell$ if there exist $\ell$ indices $i_1, \ldots, i_\ell \in [n]$ such that for every string $z \in \{0,1\}^\ell$ there exists a string $y \in S$ such that $y|_{i_1,\ldots,i_\ell} = z$. (In words, restricting $y$ to the indices $i_1, \ldots, i_\ell$ gives $z$).*

*A function $f : \{0,1\}^n \to \{0,1\}^k$ is called an $\ell$-*disperser *if for every bit-fixing source $S$ with threshold $\ell$ the set $f(S) = \{f(y) : y \in S\}$ is $\{0,1\}^k$.*

# Questions

1. (20,0) (Properties of Parity Check matrix) Let $C$ be a linear $[n,k,d]_q$ code and let $H$ be its parity check matrix. Show that any $d-1$ rows of $H$ are independent.

   *(Hint:) I essentially did this in class.*

2. (25,3) (Dual of Hadamard) Consider the following version of Hadamard Code (which is a $[2^k-1, k, 2^{k-1}-1]_2$-code). Given a message $x_1, \ldots, x_k$ the encoding of $x$ has $2^k-1$ coordinates (one for each choice of $0 \neq y \in F_2^k$) and $E(x)_y = \Sigma_{1 \le i \le k} x_i \cdot y_i$ (where the computation is in $F_2$). Show that the dual code to this code is a Hamming code.

   *(Hint:) What is the generating and parity check matrices of the dual code?*

3. (25,5) (Codes give 1-way protocols)

   (a) (15,5) Show how to use a $(n,k,(1-\epsilon)n)_q$-code to give a 1-way randomized protocol for comparison that has success probability $1-\epsilon$ and has randomness size $n$ and message size $q$.

   (b) (10,0) Use RS codes to conclude that there are protocols which give $n = q = O(k/\epsilon)$. Note that such protocols allow passing only $O(\log k + \log(1/\epsilon))$ bits!

   *(Hint:) I sketched how to solve this question in class and you are only required to fill in the details.*

4. (28,8) (1-way protocols give codes) Show how to transform any 1-way randomized protocol for comparison into a $(n,k,(1-\epsilon)n)_q$-code.

   *(Hint:) It may be easier (though not necessary) to solve question 3 first.*

5. (28,10) An $(n,k,d)_q$-code is called an MDS code if $d = n-k+1$ (that is the code meets the singleton bound). Show that if a linear code $C$ is an MDS code then so does the dual code $C'$.

   *(Hint:) You may want to use question 1. See also the hint for 2.*

6. (28,10) (Codes give $\ell$-wise independence)

   (a) (3,0) Show that if $A$ is an $\ell \times \ell$ regular matrix with 0/1 entries and $Y \in F_2^\ell$ is a uniformly distributed vector then $A \cdot Y$ is also uniformly distributed.

   (b) (15,5) Let $H$ be the parity check matrix of an $[n,k,d]_2$-code. Recall that $H$ is a $n \times n-k$-matrix. Consider the following distribution: Choose at random $Y \in F_2^{n-k}$ and compute $X = H \cdot Y$. Show that $X$ is $\ell$-wise independent for $\ell = d-1$.

   *(Hint:) You probably want to use question 1.*

   (c) (10,5) Conclude that there are 2-wise independent distributions on $n = 2^\ell - 1$ bits that can be generated using only $\ell$ random bits.

7. (28,15) (Dispersers for bit-fixing sources) Show that there exist constants $c_1, c_2 > 0$ such that for any $n$ there is an $\ell$-disperser $f : \{0,1\}^n \to \{0,1\}^k$ with $\ell \ge (1-c_1)n$ and $k = c_2 n$ that is computable in polynomial time.

   *(Hint:) Consider a binary code that is asymptotically good and has $k = c_2 n$...*