HW 2 - CORRECTION

1 Homework 2

1. Let $A_q(n,d)$ be the largest k such that a code over alphabet $\{1,\dots,q\}$ of block length n, dimension k and minimum distance d exists (recall that this corresponds to the notation $(n,k,d)_q$). Determine $A_2(3,d)$ for all integers $d \geq 1$.

A: *We know that $\forall\ [n,k,d]_q-$ code, we have:*

$$k \leq n - \log_q \left(\sum_{i=1}^{\lfloor \frac{d-1}{2} \rfloor} \binom{n}{i} (q-1)^i \right)$$

- *Since $n=3$, for $d>3$, we have $A_2(3,d)=0$ (cannot have two words with 3 bits but having Hamming distance $d>3$).*
- *For $d=1$, we have $k \leq 3$, and we can achieve the equality by taking $C=\{0,1\}^3$, so we can encode all words with Hamming distance 1, and $A_2(3,1)=3$.*
- *For $d=2$, we have $k \leq 3$, but $k \neq 3$ because we cannot encode all 3-bits codewords with Hamming distance 2. But we can achieve $k=2$ by taking $C=\{000,011,101,110\}$. So, $A_2(3,2)$.*
- *For $d=3$, then $k \leq 1$, and it is achievable by taking $C=\{000,111\}$, so $A_2(3,3)=1$.*

2. By constructing the columns of a parity check matrix in a greedy fashion, show that there exists a binary linear code $[n,k,d]_2$ provided that

$$2^{n-k} > 1 + \binom{n-1}{1} + \cdots + \binom{n-1}{d-2}. \tag{1}$$

This is a small improvement compared to the general Gilbert-Varshamov bound. In particular, it is tight for the $[7,4,3]_2$ Hamming code.

A: *Consider $\mathbb{F}_2^{n-k}$ as the set of column vector of length $(n-k)$ over $\mathbb{F}_2$. Construct parity check matrix H as follows.*

1. *Begin with $H=h_1$, where h_1 is any nonzero vector in $\mathbb{F}_2^{n-k}$.*
2. *$\forall i \geq 2$, choose h_i as the vector in $\mathbb{F}_2^{n-k} \setminus H$ such that h_i cannot be written as a linear combination of $(d-2)$ or fewer of the vectors in H (recall that $H=\{h_1,\dots,h_{i-1}\}$).*
3. *Set $H \leftarrow H \cup \{h_i\}$.*
4. *Repeat step (2) until n column vectors are constructed (i.e. $|H|=n$).*

Now, we show that the matrix H composed by the column vectors $\{h_1,\dots,h_n\}$ is the PCM of an $[n,k,d]_2$-linear code.

In the end of the procedure, we have matrix H of size $(n-k)\times n$, and every subset of $(d-1)$ vectors of $\{h_1,\dots,h_n\}$ are linearly independent. Moreover, H is a full-rank matrix, i.e. $dim(H)=n-k$.

So we can construct an $[n,k,d]_2$-linear code by taking the generator matrix $G = kernel(H)$ which is of size $k \times n$, and $dim(G) = k$, and defining $C = x \cdot G$, with x is taken over $\mathbb{F}_2^k$.

Since any subset of $(d-1)$ column vectors of H are linearly independent, then we know that the minimum distance of C is d. So, C is an $[n,k,d]_2$-linear code.

Now we show that H can be constructed if:

$$2^{n-k} > 1 + \binom{n-1}{1} + \cdots + \binom{n-1}{d-2} \tag{2}$$

Assume that by running the algorithm we have found vectors $\{h_1, \ldots, h_j\}$ with $1 \leq j \leq n-1$. The number of different linear combinations of $(d-2)$ of fewer of the set $\{h_1, \ldots, h_j\}$ is:

$$\sum_{i=0}^{d-2} \binom{j}{i} \leq \sum_{i=0}^{d-2} \binom{n-1}{i} = \binom{n-1}{1} + \cdots + \binom{n-1}{d-2}$$

So if the inequality 2 holds, we know that there is a vector $h_{j+1} \in \mathbb{F}_2^{n-k}$ which is not a linear combination of $(d-2)$ or fewer vectors of $\{h_1, \ldots, h_j\}$ (i.e. h_{j+1} is independent of $\{h_{i_1}, \ldots, h_{i_k}\}\ ;\ k \leq d-2$).

Thus, by induction on j, we can conclude that we can obtain $\{h_1, \ldots, h_n\}$.

For the particular case of $[7,4,3]_2$-Hamming code, we have $2^{7-4} > 1 + \binom{7-1}{1}$ (so, we can use the algorithm to get its PCM).

3. A well-studied family of codes is called cyclic codes. Their defining property is that if $(c_0, \ldots, c_{n-1}) \in C$ then $(c_{n-1}, c_0, \ldots, c_{n-2}) \in C$. Show that if β is a generator of $\mathbb{F}_q^*$ and $\alpha_i = \beta^{i-1}$ with $n = q-1$, then the $[n,k]_q$ Reed-Solomon code is cyclic.

A: *Since β is the generator of $\mathbb{F}_q^*$, $\{1, 2, \ldots, q-1\} = \{1, \beta^1, \beta^2, \ldots, \beta^{q-2}\}$. Moreover, $\beta^{q-1} = \beta^0 = 1$, and in general $\beta^i = \beta^i + k(q-1);\ k \in \mathbb{Z}$.*

To prove that $C = [n,k]_q$ R-S is cyclic, we need to show that:

$$\forall (c_0, c_1, \ldots, c_{n-1}) \in C,\ then\ (c_{n-1}, c_0, \ldots, c_{n-2}) \in C$$

Indeed: $\forall (c_0, c_1, \ldots, c_{n-1}) \in C$, we can write it as:

$$\begin{aligned}(c_0, c_1, \ldots, c_{n-1}) &= (f_m(\alpha_1), \ldots, f_m(\alpha_n)) \\ &= (f_m(\beta^0), \ldots, f_m(\beta^{n-1}))\end{aligned}$$

where $f_m(x) = \sum_{j=0}^{k-1} m_j x^j$, $\forall x \in \{\beta^0, \ldots, \beta^{n-1}\}$ for some $m = (m_0, \ldots, m_{k-1}) \in \mathbb{F}_q^k$.

Then, showing $(c_{n-1}, c_0, \ldots, c_{n-2}) \in C$ is equivalent to showing that:

$$(c_{n-1}, c_0, \ldots, c_{n-2}) = (f_{m'}(\beta^0), f_{m'}(\beta^1), \ldots, f_{m'}(\beta^{n-1}))$$

for some $m' = (m'_0, \ldots, m'_{k-1}) \in \mathbb{F}_q^k$.

Consider $m' = (m'_0, \ldots, m'_{k-1})$ where $\forall j \in \{0, 1, \ldots, k-1\}$, $m'_j = m_j \cdot \beta^{-j}$. Clearly, $m' \in \mathbb{F}_q^k$. Then, $\forall i \in \{1, 2, \ldots, n\}$, we have:

$$f_{m'}(\beta^i) = \sum_{j=0}^{k-1} m'_j (\beta^i)^j = \sum_{j=0}^{k-1} m_j \cdot \beta^{-j} \cdot (\beta^i)^j = m_j (\beta^{i-1})^j = f_m(\beta^{i-1})$$

and $f_{m'}(\beta^0) = f_{m'}(\beta^{q-1}) = f_{m'}(\beta^n) = f_m(\beta^{n-1})$.

Therefore,

$$(c_{n-1}, c_0, \ldots, c_{n-2}) = (f_m(\beta^{n-1}), f_m(\beta^0), \ldots, f_m(\beta^{n-2}))$$
$$= (f_{m'}(\beta^0), f_{m'}(\beta^1), \ldots, f_{m'}(\beta^{n-1}))$$

So, $(c_{n-1}, c_0, \ldots, c_{n-2}) \in C$*, hence* C *is cyclic.*

4. The Hadamard code has a nice property that it can be locally decoded. Let $C_{Had,r} : \{0,1\}^r \to \{0,1\}^{2^r}$ be the encoding function of the Hadamard code. Suppose you are interested only in the i-th bit x_i of the message $x \in \{0,1\}^r$. The challenge is that you only have access to $y \in \{0,1\}^{2^r}$ such that $\Delta(C_{Had,r}(x), y) \leq \frac{2^r}{10}$ and you would like to look only at a few bits of y. Show that by querying only 2 well-chosen positions (the choice will involve some randomization) of y, you can determine x_i correctly with probability $4/5$ (the probability here is over the choice of the queries, in particular x, y and i are fixed).

   Hint: You might want to query y at the position labelled by $u \in \{0,1\}^r$ at random and the position $u + e_i$ where $e_i \in \{0,1\}^r$ is the binary representation of i

A: *We will query* y_u *and* y_{u+e_i}*, where* y_u *and* y_{u+e_i} *is the bit of* y *corresponds to the decimal value of* u *and* $u+e_i$ *respectively, with* u *is chosen randomly over* $\{0,1\}^r$ *and* $e_i = (0\ldots010\ldots0)$ *(with 1 at the* i*-th position).*

Note that every k*-th bit of* $C_{Had,r}(x)$ *corresponds to one of* $k \in \{0,1\}^r$ *and the message* x*, i.e. we can write:*

$$C_{Had,r}(x)_k = x \odot k$$

with $x \odot k = (\sum_{i=1}^{r} x_i \cdot k_i)$ *(mod 2).*

Now notice that:

$$(x \odot u) + (x \odot (u + e_i)) \equiv (x \odot u) + (x \odot u) + (x \odot e_i)$$
$$\equiv (x \odot e_i)(\textit{mod } 2)$$
$$\equiv x_i$$

So we can determine x_i *correctly if and only if we can determine both* $(x \odot u)$ *and* $(x \odot (u + e_i))$ *correctly.*

Note that u *is picked randomly (also uniformly) from the set* $\{0,1\}^r$*. Then, since we have:* $\Delta(C_{Had,r}(x), y) \leq \frac{2^r}{10}$*, we know that:*

$$\mathbb{P}(x \odot u \textit{ is wrong}) = \mathbb{P}(x \odot (u + e_i) \textit{ is wrong}) \leq \frac{1}{10}$$

Therefore:

$$\mathbb{P}(x_i \textit{ is correct}) = 1 - \mathbb{P}(x \odot u \textit{ is wrong or } x \odot (u + e_i) \textit{ is wrong})$$
$$\geq 1 - (\mathbb{P}(x \odot u \textit{ is wrong}) + \mathbb{P}(x \odot (u + e_i) \textit{ is wrong}))$$
$$\geq 1 - (\frac{1}{10} + \frac{1}{10})$$
$$= \frac{4}{5}$$

So, $\mathbb{P}(\textit{we can determine } x_i \textit{ correctly}) \geq \frac{4}{5}$.