Lecture 5: Block codes

Course instructor: Alexey Frolov

al.frolov@skoltech.ru

Teaching Assistant: Stanislav Kruglik

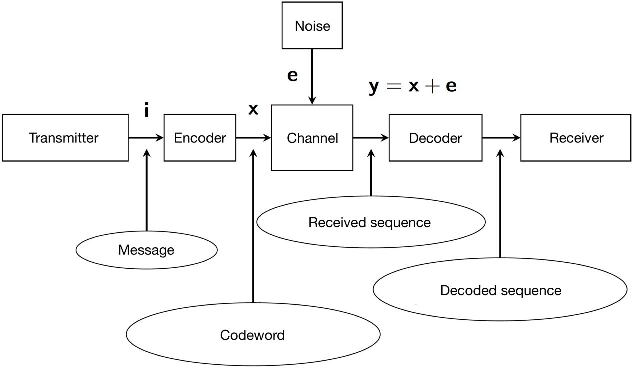
stanislav.kruglik@skolkovotech.ru

February 9, 2017

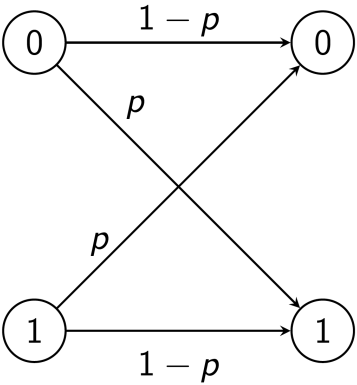
Outline

Outline

Noisy transmission



Binary symmetric channel



$$101101 \quad \rightsquigarrow \quad 10\underline{0}101$$

Why do we need encoder and decoder?

Example

Let $p = 10^{-3}$.

The probability of correct reception of n bits is equal to

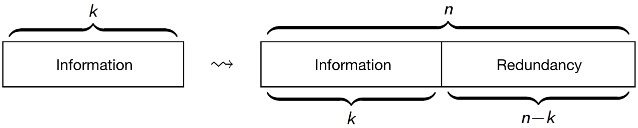
$$P_0(n) = (1 - p)^n = 0.999^n.$$

Note, that

- P_0 decreases exponentially;
- $P_0(10^3) < 0.37$;
- $P_0(10^5) < 5 \cdot 10^{-5}$;

Block and convolutional coding

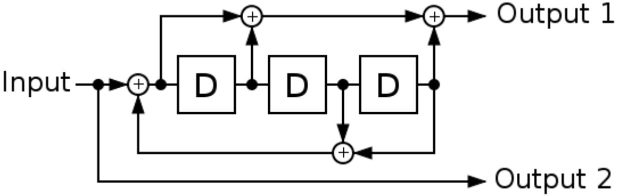
Main idea: add redundancy and use it to deal with errors.



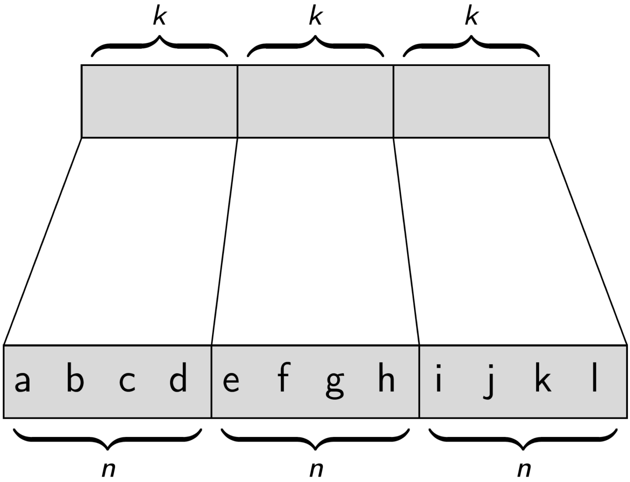
Coding methods:

1. **Block codes.** Information is split in blocks of k bits. Each block is encoded independently. As a result we obtain blocks of length n.
2. **Convolutional codes.** The output of a convolutional encoder (potentially) depends on all the previous input bits.

Convolutional coding



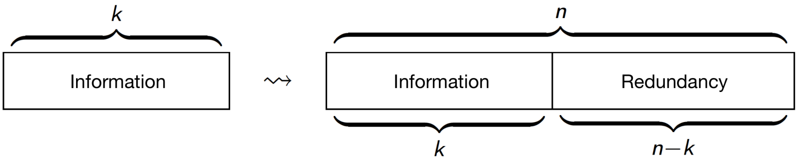
Block coding



Code rate

- $\{1, 2, \dots, M\}$ – message set;
- $Q = \{0, \dots, q-1\}$;
- $\mathbf{x} = \Psi(i) \in Q^n$ – codeword;
- $\mathcal{C} = \{\mathbf{x} = \Psi(i), i = 1, \dots, M\}$ – code;
- codebook – a table with all codewords listed;
- $\mathbf{y} \sim P(y^n|x^n)$ – received sequence;
- $\hat{i} = \Psi^{-1}(\mathbf{y})$ – decoding rule.
- $R = \frac{\log_q M}{n} = \frac{k}{n}$.

Systematic encoding



k information symbols, $n - k$ check symbols.

How to decode?

$$\mathbf{y} = 10101$$

$$\Psi : \begin{cases} 00 \longrightarrow 00001 \\ 01 \longrightarrow 01010 \\ 10 \longrightarrow 10111 \\ 11 \longrightarrow 11100 \end{cases}$$

$\mathbf{x}$	$P(\mathbf{y}\vert\mathbf{x})$
00001	$p^2(1-p)^3$
01010	p^5
10111	$p(1-p)^4$
11100	$p^2(1-p)^3$

$$p^5 < p^2(1-p)^3 < p(1-p)^4$$

$$\Rightarrow \mathbf{x} = 10111,\ \mathbf{i} = 10$$

Maximum likelihood decoding

$$\overbrace{\boxed{\quad\quad}}^{n}\overbrace{\boxed{\quad\quad}}^{n}\overbrace{\boxed{\quad\mathbf{y}\quad}}^{n}\overbrace{\boxed{\quad\quad}}^{n}\ \cdots \qquad «\Psi^{-1}(\mathbf{y}) = ?»$$

ML decoding:

1. $\mathbf{x} = \arg\max_{\mathbf{x}\in\mathcal{C}} P(\mathbf{y}|\mathbf{x})$;
2. $i = \Psi^{-1}(\mathbf{x})$.

Lemma

Let $\mathcal{C} = \{x_i\}$, $p < 0.5$ and $P(\mathbf{y}|\mathbf{x}) = \max_i P(\mathbf{y}|\mathbf{x}_i)$, then

$$d(\mathbf{y},\mathbf{x}) = \min_i d(\mathbf{y},\mathbf{x}_i),$$

where $d(\mathbf{y},\mathbf{x})$ denotes the number of elements in which $\mathbf{y}$ and $\mathbf{x}$ differ.

Hamming distance

Definition

Let $\alpha, \beta \in Q^n$.

$$d(\alpha, \beta) = |\{i : \alpha(i) \neq \beta(i)\}|.$$

Example

$$\begin{aligned} \alpha &= 01101 \\ \beta &= 00111 \end{aligned}$$

$$d(\alpha, \beta) = 2.$$

Weight and number

Definition

- $||\alpha|| = d(\alpha, \mathbf{0})$ – weight of α;
- $|\alpha| = \sum_{i=1}^{n} \alpha_i q^{n-i}$ – number (lexicographic order) of α;

Ball and sphere

Definition

Let us consider a metric space (Q^n, d), then a ball and sphere are defined as follows

$$B_r(\alpha) = \{\beta \in Q^n : d(\alpha, \beta) \leq r\}$$

and

$$S_r(\alpha) = \{\beta \in Q^n : d(\alpha, \beta) = r\}$$

Ball and sphere

$$|S_r(\alpha)| = \binom{n}{r}(q-1)^r$$

and

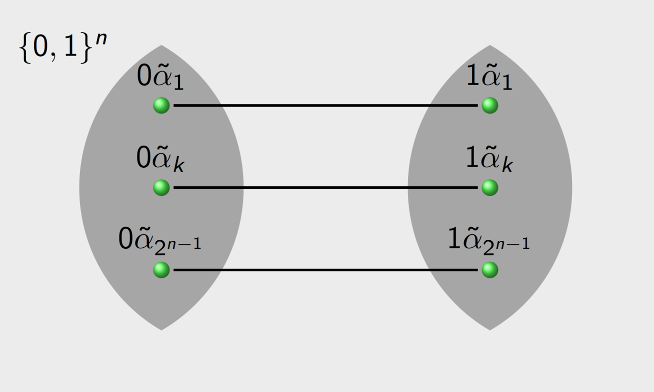
$$|B_r(\alpha)| = \sum_{i=0}^{r} |S_r(\alpha)| = \sum_{i=0}^{r} \binom{n}{i}(q-1)^i$$

$q = 2$

- $\{0,1\}^n$ – Boolean cube;
- $\{0,1\}^n_k = \{\alpha \in \{0,1\}^n : ||\alpha|| = k\}$ – Boolean cube layer;
- The set of points of $\{0,1\}^n$ with fixed $n-k$ coordinates is called k-dimensional facet.

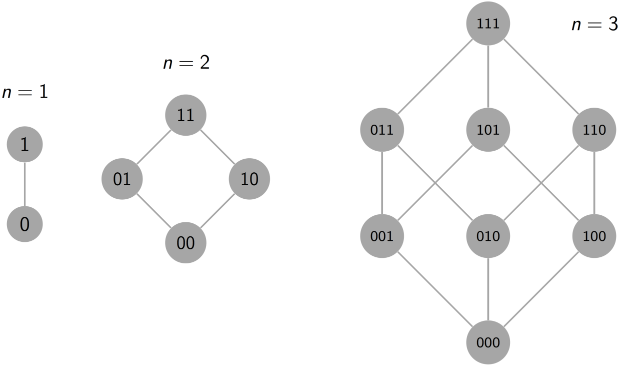
$$*0*10 = \left\{ \begin{array}{c} 00010 \\ 00110 \\ 10010 \\ 10110 \end{array} \right\}$$

$\{0,1\}^{n-1} \to \{0,1\}^n$

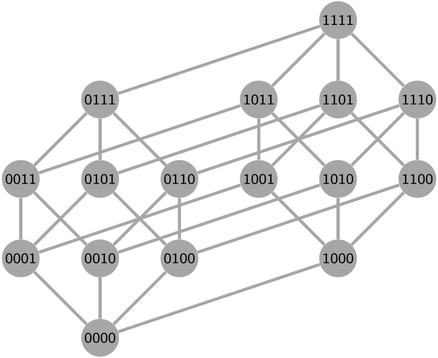


$\{0,1\}^n$
$0\tilde{\alpha}_1$
$1\tilde{\alpha}_1$
$0\tilde{\alpha}_k$
$1\tilde{\alpha}_k$
$0\tilde{\alpha}_{2^{n-1}}$
$1\tilde{\alpha}_{2^{n-1}}$

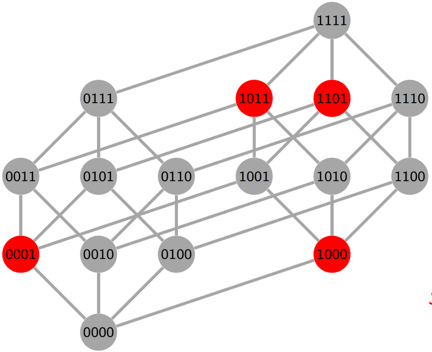
Small dimensions



$\{0,1\}^4$


1111
0111
1011
1101
1110
0011
0101
0110
1001
1010
1100
0001
0010
0100
1000
0000

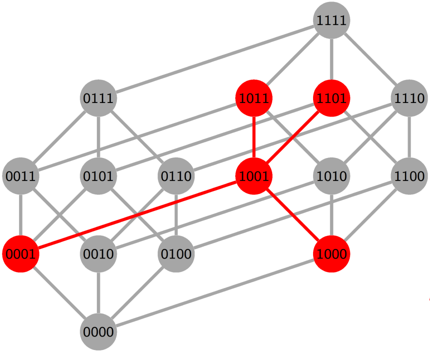
$\{0,1\}^4$



Sphere

$S_1(1001)$

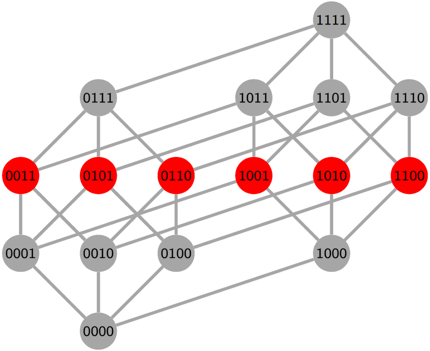
$\{0,1\}^4$



Ball

$B_1(1001)$

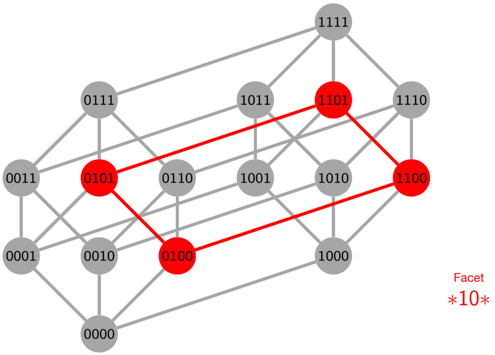
$\{0,1\}^4$



Layer

$\{0,1\}^4_2$

$\{0,1\}^4$


1111
0111
1011
1101
1110
0011
0101
0110
1001
1010
1100
0001
0010
0100
1000
0000
Facet
$*10*$

Code

Definition

- Code $\mathcal{C} \subseteq Q^n$;
- Minimum code distance

$$d(\mathcal{C}) = \min_{a,b \in \mathcal{C}; a \neq b} d(a, b).$$

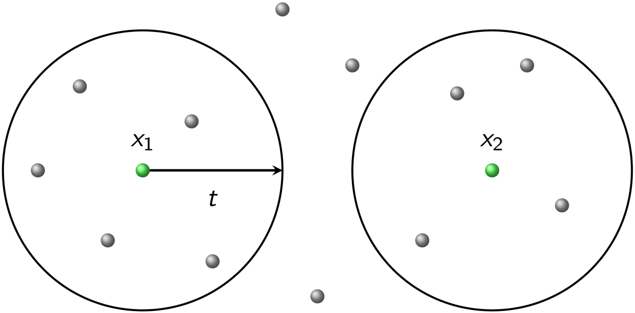
Detection and correction of errors

Theorem

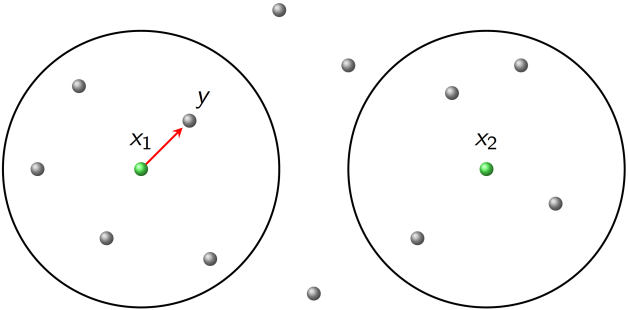
Assume the code $\mathcal{C}$ can correct t errors, then

$$d(\mathcal{C}) \geq 2t + 1.$$

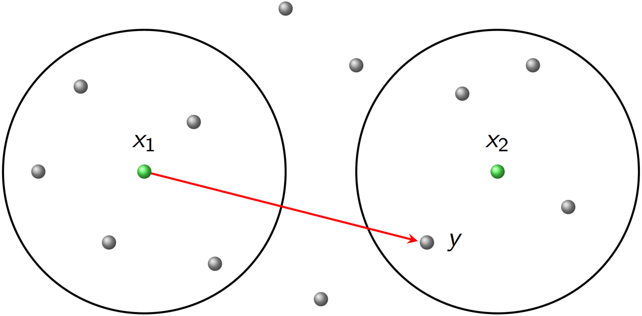
Geometric interpretation



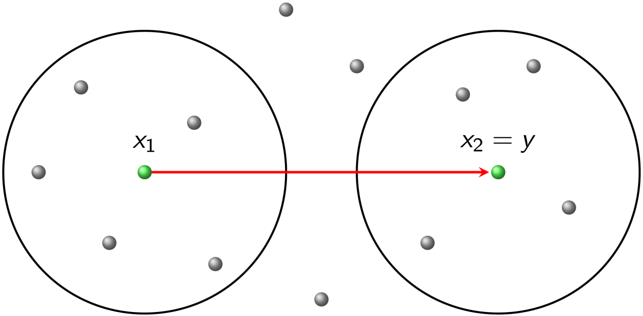
Error corrected



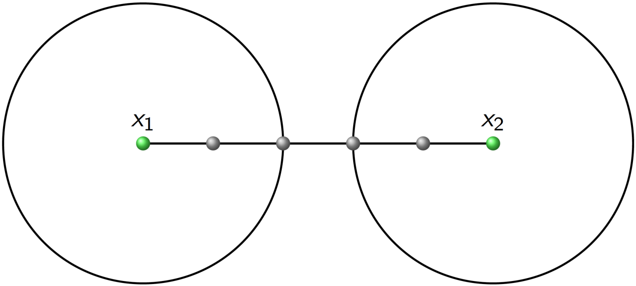
Error detected



Error undetected

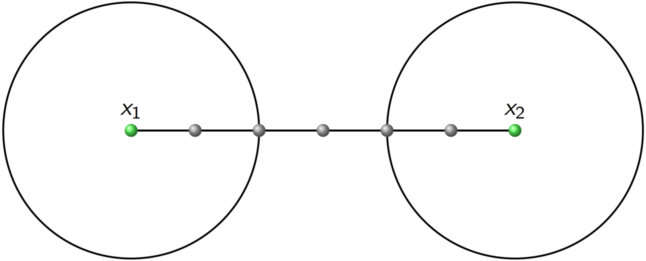


Odd distance



$$d = 5 \;\Rightarrow\; t = 2,\; s = 4,\; s' = 2$$

even distance



$$d = 6 \;\Rightarrow\; t = 2,\; s = 5,\; s' = 3$$

Error correction and error detection

Theorem

Assume $d(\mathcal{C}) = d$, *then*

$$t = \left\lfloor \frac{d-1}{2} \right\rfloor$$

and

$$s = d - 1.$$

Outline

Definition of $A_q(n, d)$

Definition

$$A_q(n, d) = \max_{\mathcal{C} \subseteq Q^n, d(\mathcal{C}) = d} |\mathcal{C}|.$$

Note, that size and rate maximization are equal tasks.

In what follows we omit the index q in case of $q = 2$.

Hamming bound

Let $\alpha \in Q^n$. Let us introduce a notation

$$V_t = V_q(t) = |B_t(\alpha)| = \sum_{i=0}^{n} \binom{n}{i} (q-1)^i.$$

Theorem (Hamming bound)

$$A_q(n,d) \le \frac{q^n}{V_t}.$$

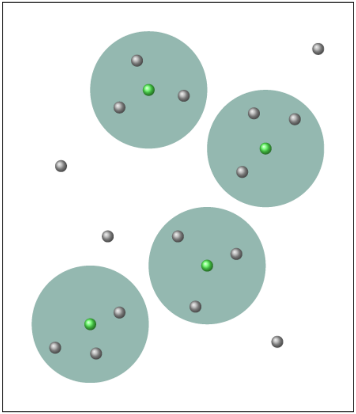
Definition

The code is called a *perfect* code if $|\mathcal{C}| = \frac{q^n}{V_t}$.

Example

A code $\mathcal{C} = \{000, 111\} \subset \{0,1\}^3$ is a perfect code.

Proof of Hamming bound



Balls of radius t do not intersect!

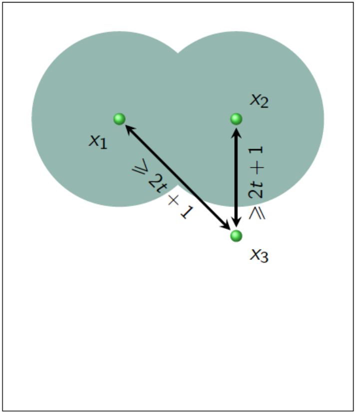
$$A_q(n,d)V_q(t) \le q^n.$$

Varshamov–Gilbert bound

Theorem (Varshamov–Gilbert bound)

$$A_q(n, d) \geq \frac{q^n}{V_{2t}}.$$

Proof



$$\mathbf{x}_3 \notin B_{2t}(\mathbf{x}_1) \cup B_{2t}(\mathbf{x}_2)$$

$\mathcal{C}_3 = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ corrects t errors.

Assume we constructed m codewords and can not add more

$$q^n = |\bigcup_{i=1}^{m} B_{2t}(\mathbf{x}_i)| \leq mV_{2t}.$$

Richard Wesley
Hamming

Edgar Nelson
Gilbert

Singleton bound

Theorem (Singleton bound)

$$A_q(n, d) \leq q^{n-d(\mathcal{C})+1}.$$

Proof.

Consider the codebook and delete $d - 1$ columns from it. All the words are different in the resulting table. □

Plotkin bound

Theorem (Plotkin bound)

$$d(\mathcal{C}) \le \frac{q-1}{q}\frac{M}{M-1}n.$$

Proof

$$S = \sum_{u,v \in \mathcal{C}} d(u,v).$$

Note, that

$$S \geq M(M-1)d$$

Consider the first column of the codebook. Let t_i be the number of times i appears in the first column.

$$\sum_{i=0}^{q-1} t_i(M - t_i) = \ell.$$

Finally,

$$\ell = M^2 - \sum_{i=0}^{q-1} t_i^2 \leq M^2 - q\left(\frac{M}{q}\right)^2 = M^2 \frac{q-1}{q}.$$

Asymptotic regime, $n \to \infty$

$\frac{d}{n} \to \delta$, $\frac{\log_q M}{n} = \frac{k}{n} \to R$

Definition

A code family $\{\mathcal{C}_n\}$ is said to be *asymptotically good* if there exist constants $R, \delta > 0$:

- $\frac{\log_q M_n}{n} = \frac{k_n}{n} \geq R > 0$;
- $\frac{d_n}{n} \geq \delta > 0$;

Asymptotic regime, $n \to \infty$, $q = 2$

Hamming bound

$$R \leq 1 - h(\delta/2).$$

Varshamov–Gilbert bound

$$R \geq 1 - h(\delta).$$

Singleton bound

$$R \leq 1 - \delta.$$

Plotkin bound

$$R \leq \frac{1}{2}(1 - \delta).$$

Proof hints

To derive asymptotic form of Hamming and Varshamov–Gilbert bounds use the following inequality

$$\sum_{i=0}^{W} \binom{n}{i} \le 2^{nh\left(\frac{W}{n}\right)} \quad \text{for} \quad W \le n/2.$$

Proof hints

To derive asymptotic form of Plotkin bound use the shortening method

Lemma

$$A_q(n, d) \leq qA_q(n-1, d).$$

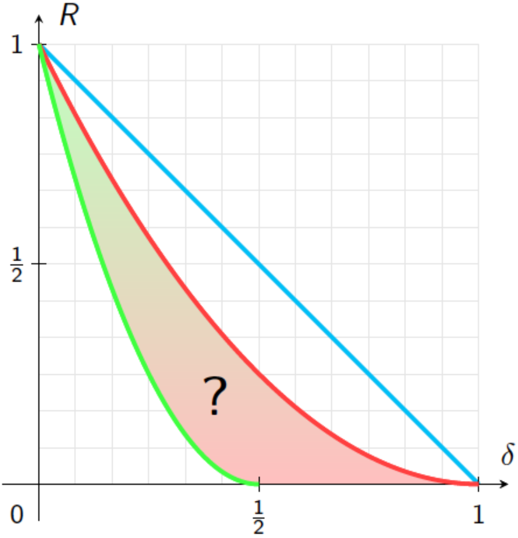
Proof.

Consider the codebook and split it into q parts in dependence on the first symbol, i.e.

$$\mathcal{C} = \begin{bmatrix} 0 & \mathcal{C}'_0 \\ 1 & \mathcal{C}'_1 \\ \dots & \\ q-1 & \mathcal{C}'_{q-1} \end{bmatrix}$$

At least one of the codes $\mathcal{C}'_i$ contains $|\mathcal{C}|/q$ codewords. At the same time $d(\mathcal{C}'_i) \geq d(\mathcal{C})$ for all i. □

Asymptotic regime, $n \to \infty$



Outline

Linear code

Definition

A subgroup of Abelian group $\mathbb{F}_q^n$ is called a group code.

Definition

A subspace $\mathcal{C}$ of a vector space $\mathbb{F}_q^n$ is called a linear (n, k) code, where $k = \dim \mathcal{C}$.

Definition

G is a generator matrix for the code $\mathcal{C}$ if the rows of G form a basis in $\mathcal{C}$.

Minimum distance

Lemma

Let $\mathcal{C}$ be a linear code, then

$$d(\mathcal{C}) = \min_{a \in \mathcal{C}, a \neq 0} ||a||.$$

Dual code and parity check matrix

Definition (Dual code)

$$C^{\perp} = \{v \in \mathbb{F}_q^n : v \perp \mathcal{C}\}.$$

Definition (Parity check matrix)

H is a basis of $C^{\perp}$.

Varshamov–Gilbert bound

Theorem

If

$$\sum_{j=0}^{d-2} \binom{n-1}{j} (q-1)^j < q^{n-k},$$

then the exists a linear $(n, k, \geq d)_q$ code.

Proof

Proof.

Construct a parity check matrix column by column. Assume we have already added $\hat{n}$ columns and want to add one more column. As any $d-1$ columns are linerly independent (because of the distance) we can not add this number of columns (linear combinations of $d-2$ columns)

$$\sum_{j=0}^{d-2} \binom{\hat{n}}{j} (q-1)^j,$$

If this number is less, than the total number of columns (q^{n-k}) we can continue the procedure. $\square$

Practically all linear codes are good!

We randomly generate each bit of the parity-check matrix $\mathbf{H}$ of size $(n-k) \times n$ according to a Bernoulli distribution $\mathrm{Bern}(1/2)$. Let us consider a fixed word $\mathbf{x}$ of length n and weight $W > 0$. The probability of this word to be a codeword (or a probability of the syndrome to be equal to zero) can be calculated as follows

$$\Pr(\mathbf{H}\mathbf{x} = \mathbf{0}) = 2^{-(n-k)}.$$

Indeed, let us find any non-zero bit (say, bit i) in $\mathbf{x}$. Choose all the elements (except the column i) in $\mathbf{H}$ arbitrarily. The probability to choose the i-th column such, that the syndrome is equal to zero is $2^{-(n-k)}$.

Practically all linear codes are good!

Now consider the following event E: the code includes at least one codeword with weight $W \leq \delta n$. We have

$$\Pr(E) \leq \sum_{i=1}^{\delta n} \binom{n}{i} 2^{-(n-k)} \quad \text{(union bound)}.$$

Finally,

$$\sum_{i=1}^{\delta n} \binom{n}{i} 2^{-(n-k)} \leq 2^{-n(1-R-h(\delta))}$$

and we see, that the probability (or fraction) of bad codes decrease exponentially with n for any $R < 1 - h(\delta)$.

Exercise, 2 points

Prove

Lemma (Bassalygo's lemma)

Let $L \subset \mathbb{F}_q^n$, $A_q^{(L)}(n,d)$ is the maximal number of words with distance d in L, then

$$\frac{A_q(n,d)}{q^n} \leq \frac{A_q^{(L)}(n,d)}{|L|}$$

Thank you for your attention!