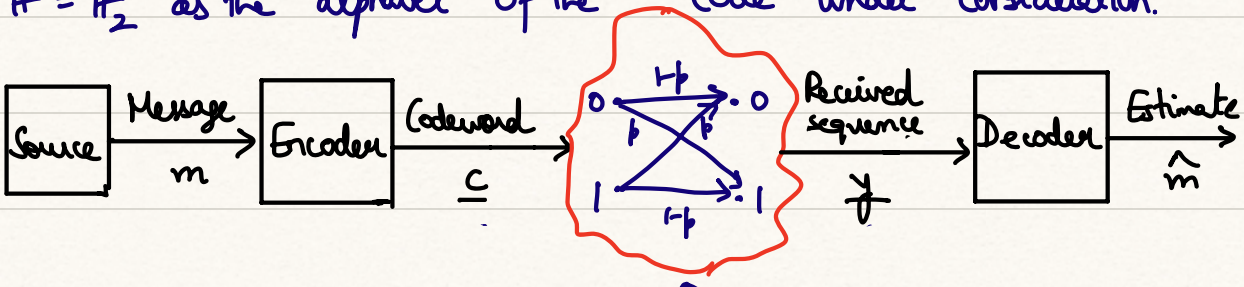
Lecture : Decoding linear codes

In this lecture (and elsewhere in classical coding theory), we shall consider as motivation the problem of decoding a codeword of a linear code over a binary symmetric channel (BSC) with some cross-over probability $0 < p < \frac{1}{2}$.

Fix $\mathbb{F} = \mathbb{F}_2$ as the alphabet of the code under consideration.



The "binary symmetric channel" (BSC).

Recall the structure of the ML (or MAP) decoder over the BSC(p), given $\underline{y}$:

$$ML(\underline{y}) = \underset{\underline{c} \in \mathcal{L}}{\operatorname{argmax}} \; P(\underline{y} | \underline{c}) = \underset{\underline{c} \in \mathcal{L}}{\operatorname{argmin}} \; d_H(\underline{y}, \underline{c}).$$

This task of finding the codeword that is "closest" to $\underline{y}$, is also called "minimum distance decoding" (MDD).

In this lecture, we shall explore what MDD means, in the context of $\mathcal{L}$ being a linear code.

Note first that

$$\text{MDD}(\underline{y}) \equiv \text{Find an "error vector" } \underline{e} \text{ of smallest Hamming weight such that } \underline{y} - \underline{e} \in \mathcal{C}. \text{ Decode to } \hat{\underline{c}} = \underline{y} - \underline{e}.$$

Now, given $\underline{y}$, it is of interest to characterize the set

$$\mathcal{E}(\underline{y}) \triangleq \{\underline{e} \in \mathbb{F}_2^n : \underline{y} - \underline{e} \in \mathcal{C}\}.$$

This is the set of "candidate" error vectors. Clearly, MDD can correct those sequences $\underline{y}$ caused by error vectors such that there exists a unique error vector $\underline{e}$ of smallest weight in $\mathcal{E}(\underline{y})$.

Furthermore, note that

$$\mathcal{E}(\underline{y}) = \{\underline{e} : \underline{e} \in \underline{y} + \mathcal{C}\} \quad [\text{linearity of } \mathcal{C}]$$

and hence since $\mathcal{C}$ is a subgroup of $\mathbb{F}_2^n$ (verify this!), it follows that $\mathcal{E}(\underline{y})$ is a coset of $\mathcal{C}$ in $\mathbb{F}_2^n$.

> MDD hence reduces to finding a coset (corresp. to $\underline{y}$) of $\mathcal{C}$ in $\mathbb{F}_2^n$ and obtaining an error vector $\underline{e}$ of min. weight in this coset.

Standard array and complexity of MDD

The task above requires us to find a specific coset, $\underline{y} + \mathcal{C}$, among all cosets

of $\mathcal{L}$ in $\mathbb{F}_2^n$. Let $\mathcal{L}$ be an $[n,k]_2$ code.

From Lecture 3, recall that the cosets of $\mathcal{L}$ in $\mathbb{F}_2^n$ partition $\mathbb{F}_2^n$; furthermore, each coset has size exactly $|\mathcal{L}| = 2^k$. There are hence exactly $2^n/2^k = 2^{n-k}$ cosets of $\mathcal{L}$ in $\mathbb{F}_2^n$.

Example: Consider the (systematic) linear code $\mathcal{L}$ generated by

$$G = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

Clearly, we have $\dim(\mathcal{L}) = 2 = k$. The cosets of $\mathcal{L}$ in $\mathbb{F}_2^n$ can be written as follows:

$\mathcal{L}\ (=\underline{0}+\mathcal{L})$	00000	01001	10101	11100
$00001 + \mathcal{L}$	00001	01000	10100	11101 ← Two candidate coset leaders!
$00010 + \mathcal{L}$	00010	01011	10111	11110
$00011 + \mathcal{L}$	00011	01010	10110	11111
$00100 + \mathcal{L}$	00100	01101	10001	11000
$00101 + \mathcal{L}$	00101	01100	10000	11001
$00110 + \mathcal{L}$	00110	01111	10011	11010
$00111 + \mathcal{L}$	00111	01110	10010	11011

Standard array table for $\mathcal{L}$