CT-216

LAB GROUP 1

Project Group-2

Assigned by Professor
Yash Vasavada

POLAR CODES

POLAR CODES ACHIEVE SHANNON'S CAPACITY:

HOW??
First let’s think intuitively….

SHANNON’S CHANNEL CAPACITY THEOREM: “It gives us the maximum rate at which information can be transmitted over a communication channel of a specified bandwidth in the presence of noise.”

POLARIZATION THEOREM: “The channel polarization theorem states that, as the code-length N goes to infinity, a polarized bit-channel becomes either a noiseless channel or a pure noise channel. The information bits are transmitted over the noiseless bit-channels and the pure noise bit-channels are set to zero (frozen bits).”

We need to show that the bit error rate in the polar encoding scheme is almost equal to the BER of Shannon's capacity. This implies that no information is passed via noisy channels, and noise is the factor that introduces error. If there is almost zero noise when N approaches infinity, the probability of error also approaches 0. Hence, we can say that Polar codes achieve Shannon’s capacity!

We need to show that the bit error rate (BER) in the polar encoding scheme is almost equal to the BER of Shannon's capacity. This implies that no information is passed via noisy channels, and noise is the factor that introduces error. If there is almost zero noise when N approaches infinity, the probability of error also approaches 0. Hence, we can say that Polar codes achieve Shannon's capacity!

But, This is a blind guess. Let's look into its actual process of polarization through which encoding is done to achieve Shannon's channel capacity.

MATHEMATICAL SOLUTION:

Before getting into the mathematics let's look at the notations used

$P(X,Y)$ is $P_{x,y}$

$$H(X|Y) = \sum_{X,Y} P(x,y) log \frac{P(x,y)}{P(x|y)}$$

$$I(X|Y) = \sum_{X,Y} P(x,y) \log \frac{P(x,y)}{P(x)P(y)}$$

(X^N, Y^N) is $P(x^N, y^N)$

$$H(X^N|Y^N) = \sum_{i=1}^{N} H(X_i|Y^N \cdot X^{i-1})$$

$$I(X^N;Y^N) = \sum_{i=1}^{N} I(X_i;Y^N|X^{i-1})$$

$X^{i-1} = (X_1, X_2 \ldots\ldots X_{i-1})$

Given $U_1, U_2 \ldots\ldots U_N$ iid unit{0,1}

$X_1, X_2 \ldots.. X_N$ iid unit{0,1}

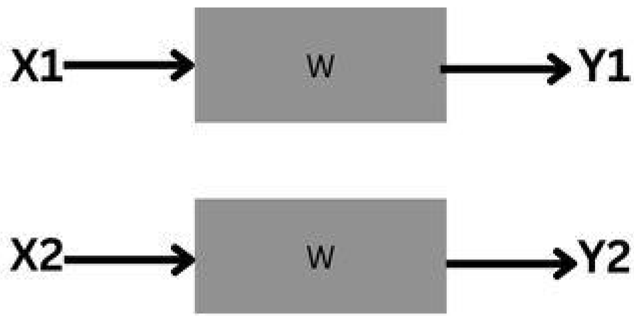
Proof:

$$H(U^N) = H(X^N) = N \cdot H(X)$$

Since it follows 1-1 transformation

$$H(U^N) = \sum_{i=1}^{N} H(U^i | U^{i-1})$$

$$\Rightarrow \sum_{i=1}^{N} H(U^i | U^{i-1}) = N \cdot H(X)$$

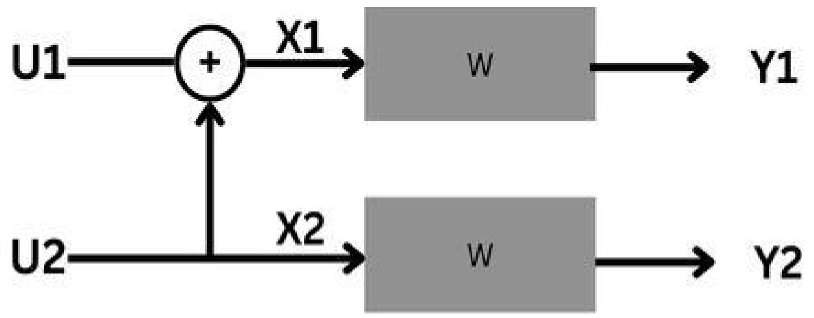


Let,

$$X_1 = U_1 \oplus U_2$$

$$X_2 = U_2$$

Here, (U_1, U_2) are uniform on F_2^2



As $(X_1,\ X_2)$ is also uniform on $\mathcal{F}_z^2$, we have:

$$
\begin{aligned}
2I(W) &= I(X_1,\ X_2;\ Y_1,\ Y_2) \qquad\qquad I(W) = I(XY)\ i.e.\ mutual\ info \\
&= I(U_1,\ U_2;\ Y_1,\ Y_2) \\
&= I(U_1;\ Y_1 Y_2) + I(U_2;\ Y_1 Y_2/\mathrm{U}_1) \\
&= I(U_1;\ Y_1 Y_2) + I(U_2;\ Y_1 Y_2 \mathrm{U}_1) \\
&= I(W^-) + I(W^+)
\end{aligned}
$$

Therefore, $I(W^-) \leq I(W) \leq I(W^+)$

<u>EXPLANATION USING THE EXAMPLE OF BEC CHANNEL:</u>

For N=2

Suppose W is a BEC(p) i.e.,

$$
Y = \begin{cases} \mathrm{X} & \text{with probability } (1-p), \\ \mathrm{e} & \text{with probability } p \end{cases}
$$

Now, let it be that

W^- has input U_1,output

$$
(Y_1, Y_2) = \begin{cases} (U_1 \oplus U_2, U_2) & (1-p)^2 \\ (e, U_2) & p(1-p) \\ (U_1 \oplus U_2, e) & (1-p)p \\ (e, e) & p^2 \end{cases}
$$

$\mathbf{W}^-$ is BEC (p^-): $p^- = 2p - p^2$

After estimating the value of U1 we move to U2.
U1 is known parameter hence a non random value, so we do not include it in the probability calculation of U2

W^+ has input U_2, output

$$(Y_1, Y_2, U_1) = \begin{cases} (\mathrm{U}_1 \oplus U_2, U_2, U_1) & : (1-p)^2 \\ (\mathrm{e},\ \mathrm{U}_2, U_1) & : p(1-p) \\ (\mathrm{U}_1 \oplus U_2, e, U_1) & : (1-p)p \\ (\mathrm{e},\ \mathrm{e},\ \mathrm{U}_1) & : p^2 \end{cases}$$

So,

We can determine U_2

$\mathbf{W}^+$ is BEC (p^+): $p^+ = p^2$

$\Rightarrow \mathbf{W}^+$ is better than $\mathbf{W}$

$\Rightarrow \mathbf{W}^-$ is worse than $\mathbf{W}$

Source coding:

Coordinates of extreme capacities

$$H = \{i : H(U^i|U^{i-1}) \approx 1\}$$
$$(H) \approx N \cdot H(X)$$

Capacity:

Input|Output symmetry

$C(W) = I(X;Y)$ with X is unit {0,1}

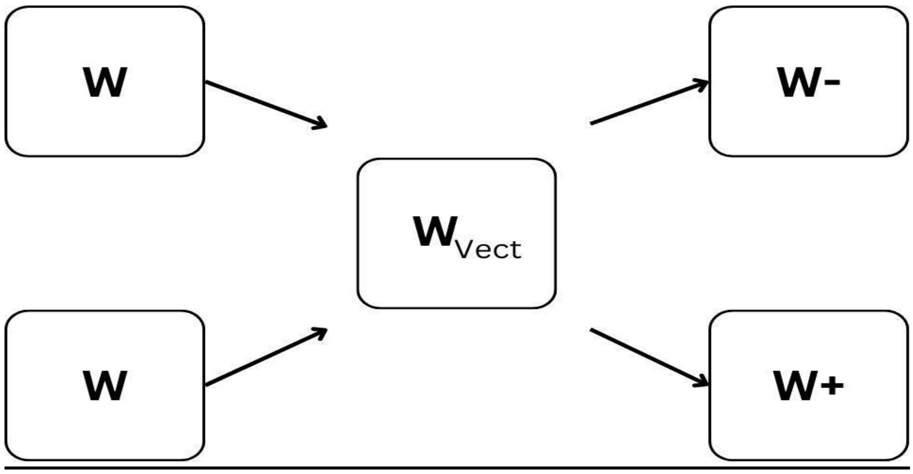
Channel:

For Perfect Channel: C(W)=1

For Useless Channel:C(W)=0

Combining and Splitting:

Polarization is achieved with the help of splitting and combining:



Begin with N copies of W

Use 1-1 mapping

$G_N : \{0,1\}^N \to \{0,1\}^N$ to create vector channels

$W_{vec} : X^N \to Y^N$

- Combining

  Combining operation is lossless

  Take $U_1,U_2.....U_N$ iid unit{0,1}

  then $X_1,X_2.....X_N$ iid unit{0,1}

$$C(W_{vec}) = I(0^N; Y^N)$$
$$= I(X^N; Y^N)$$
$$= NC(W)$$

- <u>Splitting:</u>

$$C(W_{vec}) = I(U^N; Y^N)$$
$$= \sum_{i=1}^{N} I(U_i; Y^N, U^{i-1})$$
$$= \sum_{i=1}^{N} I(W_i)$$

$\Rightarrow$ We got N polarized channels from N identical channels.

As N→∞,

A fraction of new bit channels becomes nearly perfect

$$I(W_i) \approx 1$$

and a fraction becomes nearly useless

$$I(W_i) \approx 0$$

We can extend this explanation for any given N=2^n n=1,2,3…

For N→∞,

Where we get channels that are highly informative or least informative, which reduces the probability of error that approximates to nearly zero.

Hence, Polar codes achieve Shannon's Channel Capacity.