Channel Polarization

By: Himanshu Sharma **Roll No.: 1610110149**

Channel polarization is a process in which N channels are generated from N independent copies of a B-DMC W. The newly generated channel is $\{W_N^{(i)} : 1 \leq i \leq N\}$. As $N \to \infty$, the symmetric capacity $I(W_N^{(i)})$ either becomes 0 or it becomes 1 for all vanishing fractions of i. Note that the symmetric capacity of a B-DMC W is given by

$$I(W) = \sum_{x \in X} \sum_{y \in Y} \frac{1}{2} W(y|x) \log_2 \frac{W(y|x)}{0.5W(y|0) + 0.5W(y|1)} \tag{1}$$

1 Channel Combining

Channel combining is a phase operation used in channel polarization wherein **copies** of B-DMC W are combined in a recursive manner to produce a vector channel $W_N : X^N \to Y^N$. The value of N is 2^n where $n \geq 0$. That means, the very first copy of the channel $W_1 = W$, because $n = 0$ for $N = 1$. Similarly, $W_2 : X^2 \to Y^2$.

The transition probability is given by $W_2(y_1, y_2|x_1, x_2)$, note the W_2, its not W. Asking for the value $W_2(y_1, y_2|x_1, x_2)$ is same as asking $W(y_1|x_1)W(y_2|x_2)$, becuase sending u_1 and u_2 in channel W_2 is same as sending x_1 and x_2 in two copies of channel W separately. That means,

$$W_2(y_1, y_2|x_1, x_2) = W(y_1|x_1)W(y_2|x_2) = W(y_1|u_1 \oplus u_2)W(y_2|u_2) \text{ because } x_1 = u_1 \oplus u_2.$$

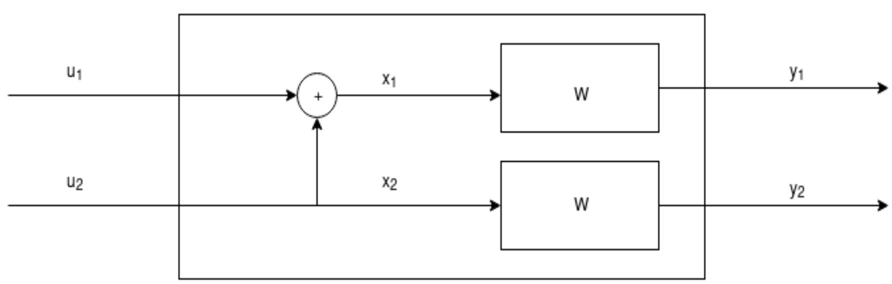

Figure 1: Channel W_2

The mapping between u^N and y^N is done by,

$$y_i^N = u_i^N G_N \ \forall i \in \{1, ..., N\} \tag{2}$$

Where

$$G_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

Similarly, for W_4 also, the kernel matrix is G_4 where,

$$G_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

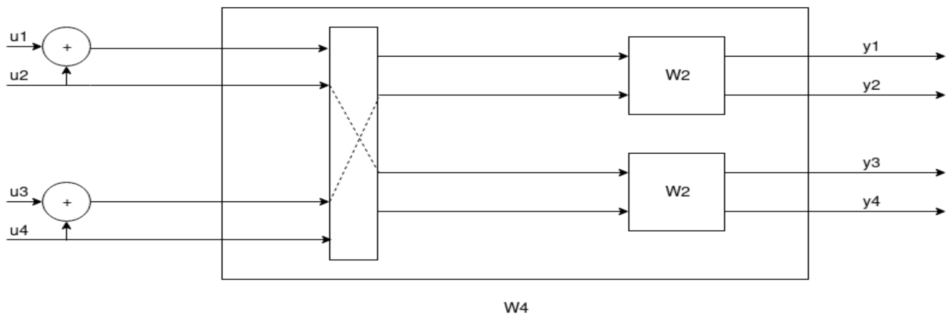

Figure 2: Channel W_4

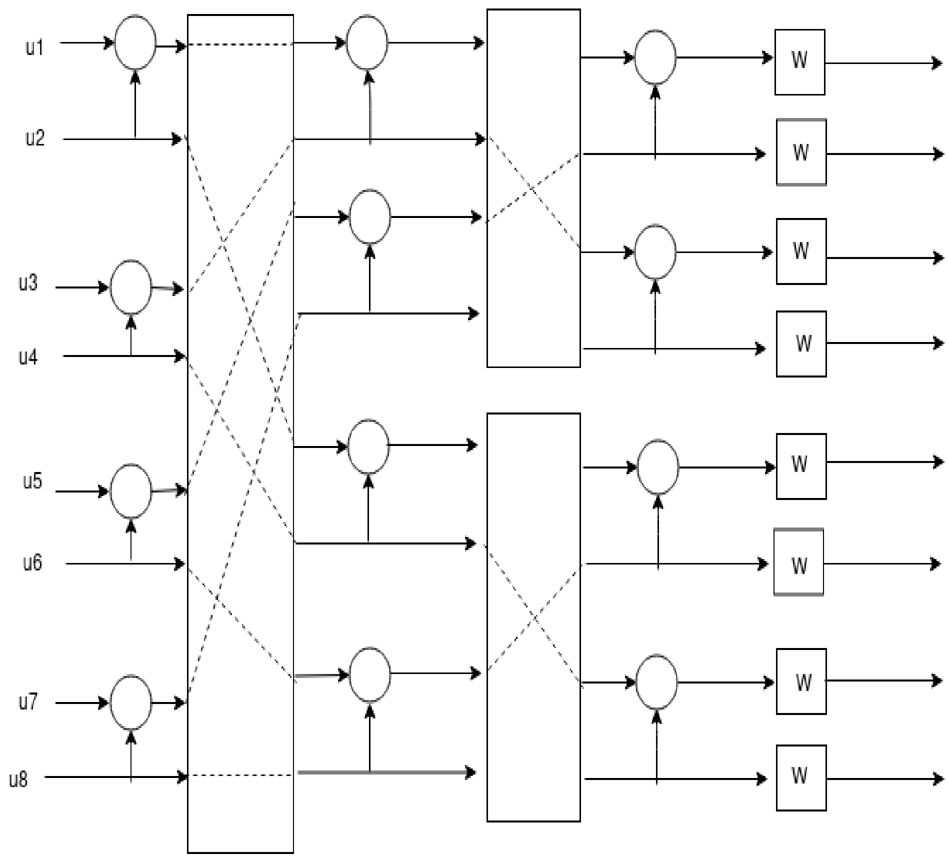

Figure 3: Channel W_8

The vertical rectangular box shown in the channel W_4 is called the **reverse shuffle operator** which takes odd indices at one side and even on the other. If one carefully goes through both W_4 and W_2, then the outputs $[y_1, y_2, y_3, y_4]$ are $[u_1 \oplus u_2 \oplus u_3 \oplus u_4, u_3 \oplus u_4, u_2 \oplus u_4, u_4]$. The same can be obtained from the kernel matrix G_4.

$$\begin{bmatrix} u_1 & u_2 & u_3 & u_4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} u_1 \oplus u_2 \oplus u_3 \oplus u_4 & u_3 \oplus u_4 & u_2 \oplus u_4 & u_4 \end{bmatrix}$$

2 Channel Polarization

Channel polarization is a process in which a B-DMC W gives rise to N channels such that $W_N^{(i)}$ $\forall i \in [1, N]$. In this process, the generated channels either go to zero information state or pure information state of $I(W) = 1/0$ as $N \to \infty$. The following diagram shows the polarization effect.

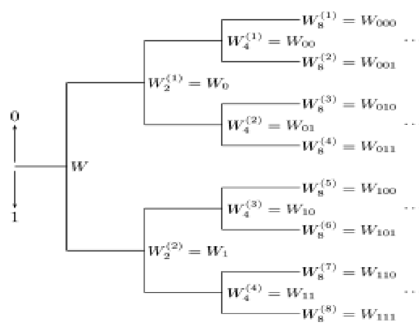

Figure 4: Channel Polarization Effect

Considering the channel W_2, we can see that it has two copies of the original channel W and therefore, it has a capacity of $2I(W)$ where $I(W) = 1 - f(p)$ is the Shannon's capacity for a BEC and p denotes the transition probability of the BEC.

2.1 Single Step Transform

Consider figure 1 again. This time, refer to the below shown diagram also, taken from the reference.

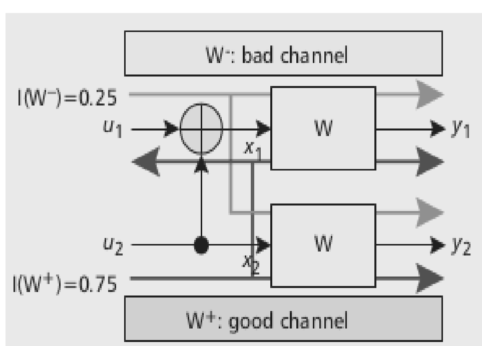

Figure 5: Polarized Channel [1]

By applying the chain rule of the mutual information, this channel W_2 can be decomposed into two BEC with capacities $I(W^-)$ and $I(W^+)$ where $I(W^-) + I(W^+) = 2I(W)$. Also, note the following identities,

$$I(W^-) = I(W)^2 \tag{3}$$

and

$$I(W^+) = 2I(W) - I(W)^2 \tag{4}$$

In the figure shown above, $I(W) = 0.5$. This proves that the bad channel W^- has a smaller capacity than the given BEC W, whereas the good channel W^+ has a larger capacity, that is, $I(W^-) \leq I(W) \leq I(W^+)$.

Let us take over with the same case of $I(W) = 0.5$, for BEC W. When this channel is polarized once, then $I(W^-) = 0.25$ and $I(W^+) = 0.75$. Consider '+' to be equivalent of 0 and '-' to be equivalent of 1, then the next will be of order 00, 01, 10, 11, i.e., ++, +-, -+ and –, with following values,

$$I(W^{++}) = 2I(W^+) - I(W^+)^2 = 0.9375$$
$$I(W^{+-}) = I(W^+)^2 = 0.5625$$
$$I(W^{-+}) = 2I(W^-) - I(W^-)^2 = 0.4375$$
$$I(W^{--}) = I(W^-)^2 = 0.0625$$

As we can see, the more we polarize, the more good and bad channels are generated with their respective channel capacities.

2.2 The Matthew Effect

The Matthew Effect is the summary of channel polarization. It says that as the length of the codeword goes to infinity, the capacity of the most good channel tends to one. The below plot shows the Matthew effect.

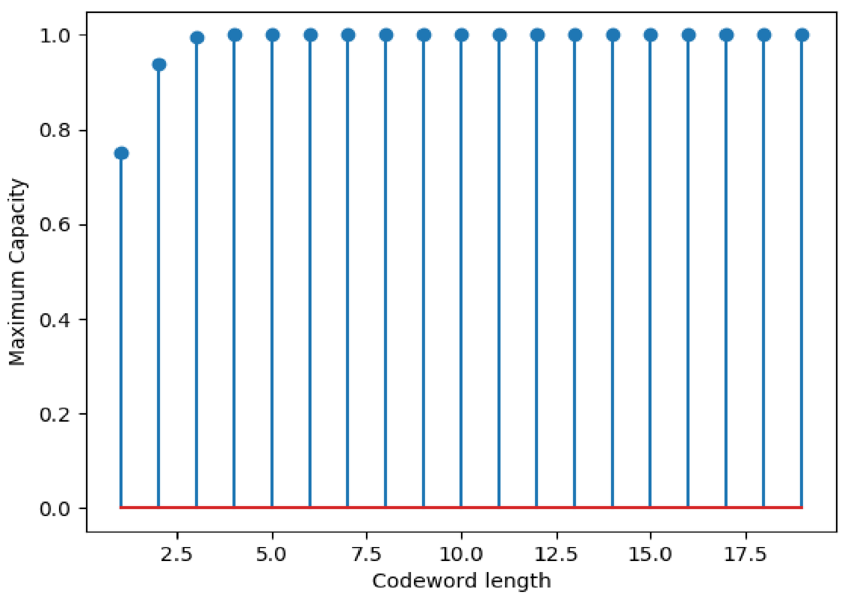

Figure 6: The Matthew Effect

Clearly, as the length of the codeword reaches 3, the capacity of the *good* channel tends to 1. That is, the more codeword length you have, the better will be the capacity of the good channel. Hence, we find that the capacities of most of the polarized channels tend to either 1 (good channels with little noise) or 0 (bad channels with full noise). Equivalently, the error probabilities of the noiseless channels or noisy channels go to 0 or 1.