Information Theory and Channel Coding

Prof. Rodrigo C. de Lamare
CETUC, PUC-Rio, Brazil
delamare@cetuc.puc-rio.br

XI. Polar codes

A. Introduction

B. Encoding and code structure

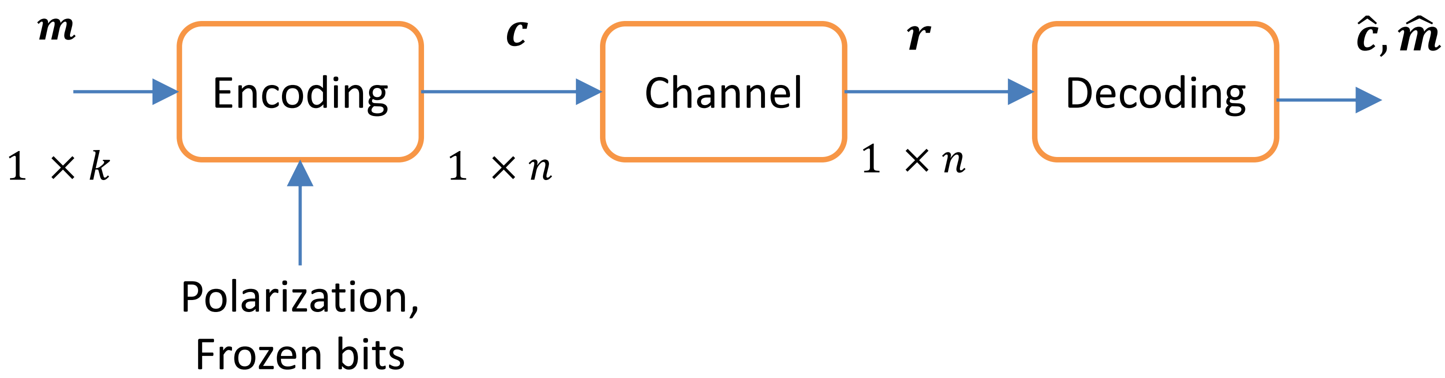
C. Decoding

PUC RIO

- Polar codes were invented by Erdal Arikan in 2009.
- Polar codes are the first provably capacity-achieving efficient coding scheme.
- The encoding of polar codes has low complexity $(n \log_2 n)$.
- The decoding of polar codes can be carried out via successive cancellation decoding with cost $(n \log_2 n)$, list decoding or belief propagation.
- They have been adopted for control channels in extended Mobile Broadband (eMBB) of 5G systems.

A. Introduction

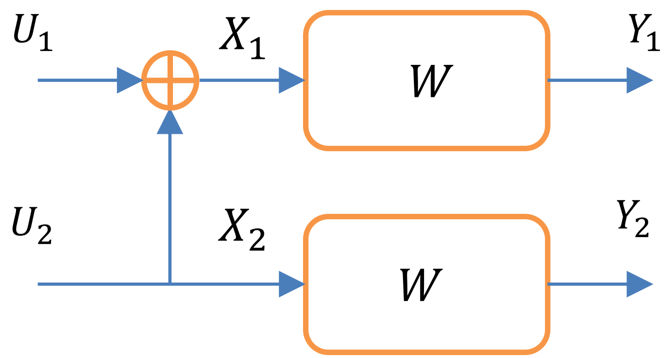
- Let us consider the block diagram of a polar coding system.



- A unique feature of polar codes is the use of polarization of bit channels, which results in noisy useless channels and error-free channels.

- Bits that are transmitted over useless channels are frozen.

- A building block of polar codes is the basic channel transformation illustrated by



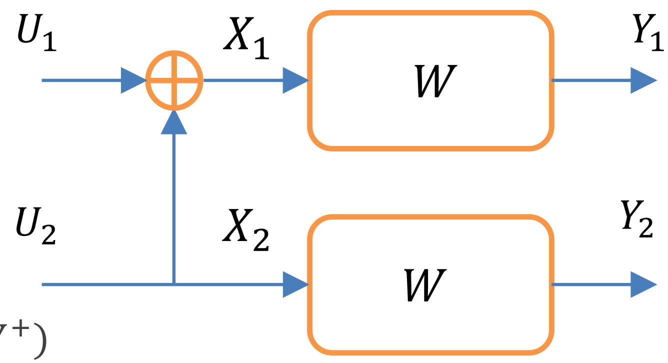
where W is a discrete memoryless channel (DMC).

- For an error-free/ perfect channel:
  - Y determines X
  - No need to encode data.

- For a useless channel:
  - Y is independent of X
  - No need to encode.

- Concatenation of DMCs:
  - Channels that are perfect or useless but nothing in between-> polarization
  - In a polar coding system a given DMC W is used n times to transmit a codeword.

Definition: Basic Channel Transformation (BCT)



$W \to (W^-, W^+)$

$W: \underbrace{\{0,1\}}_{\mathcal{X}} \to \mathcal{Y}$

$W^-: \{0,1\} \to \mathcal{Y}^2: \mathcal{U}_1 \to (Y_1, Y_2)$ (output)

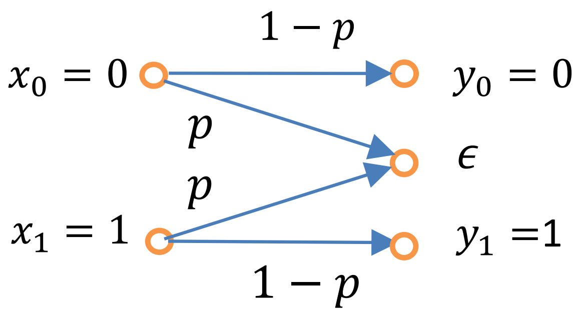
$W^+: \{0,1\} \to \mathcal{Y}^2 \times \{0,1\}: \mathcal{U}_2 \to (Y_1, Y_2, U_1)$ (output)

Inputs: $X_1 = U_1 \oplus U_2$

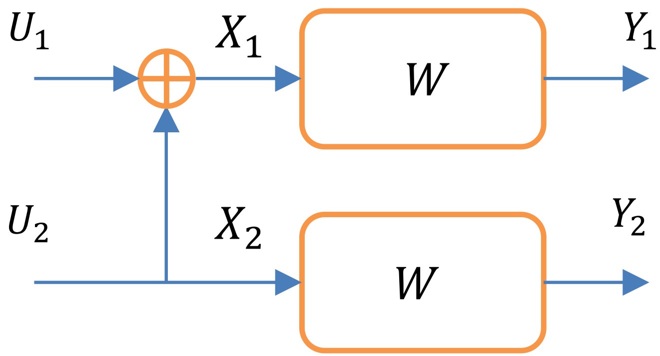
$X_2 = U_2$

Example 1

Consider that W is a binary erasure channel (BEC) with erasure probability ϵ

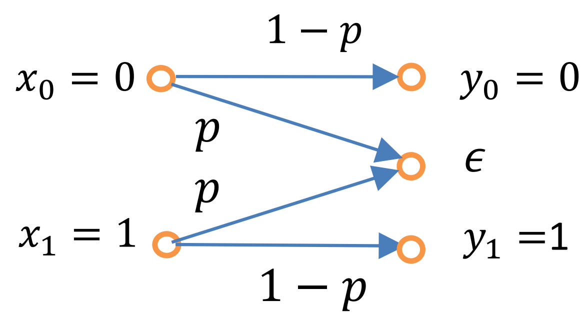


and the BCT

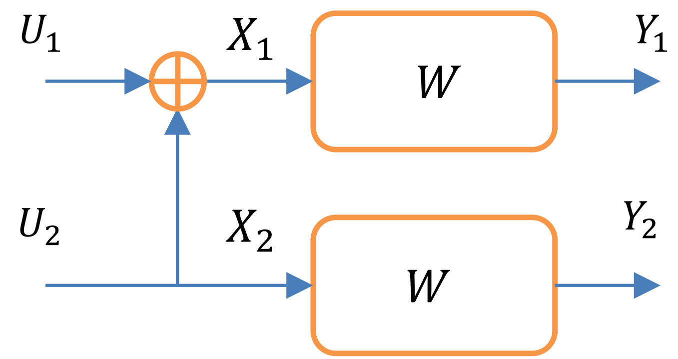


What are W^-, W^+ ? Consider $U_1, U_2 \in \{0,1\}$

Solution:



W^-: input $\mathcal{U}_1$, output (Y_1, Y_2)

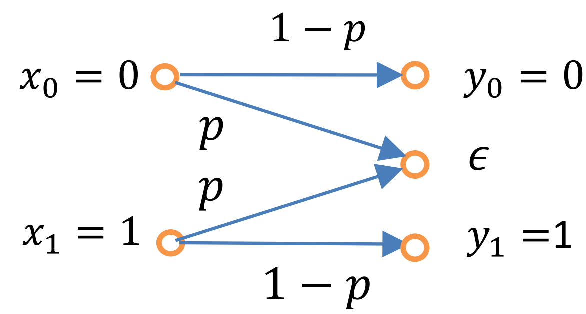


i) No erasure: $(Y_1, Y_2) = (\mathcal{U}_1 \oplus \mathcal{U}_2, \mathcal{U}_2)$

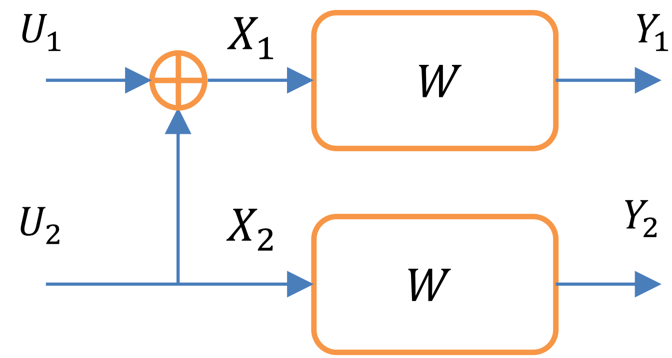
ii) Erasure on the 1st BEC: $(Y_1, Y_2) = (?\ , \mathcal{U}_2\ )$

iii) Erasure on the 2nd BEC: $(Y_1, Y_2) = (\mathcal{U}_1 \oplus \mathcal{U}_2, ?\ )$

iv) Two erasures: $(Y_1, Y_2) = (?\ , ?\ )$



W^+: input $\mathcal{U}_2$, output $(Y_1, Y_2, \mathcal{U}_1)$



i) No erasure: $(Y_1, Y_2, \mathcal{U}_1) = (\mathcal{U}_1 \oplus \mathcal{U}_2, \mathcal{U}_2, \mathcal{U}_1)$

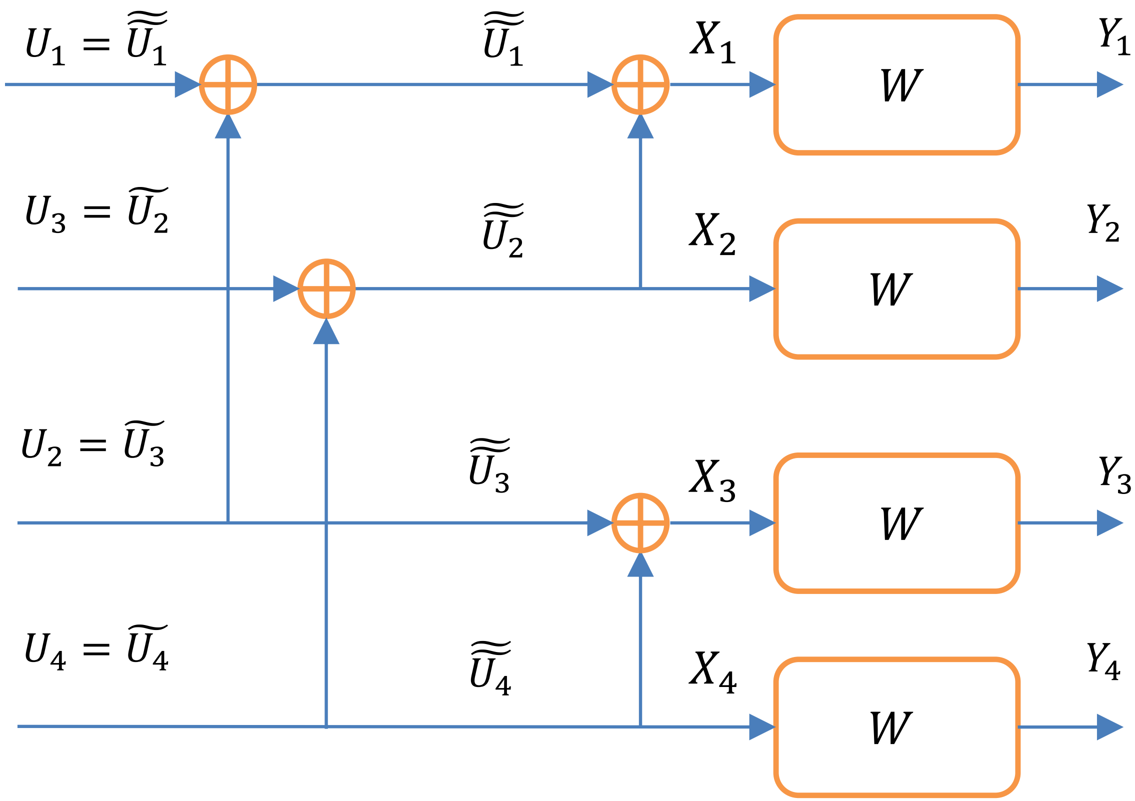
ii) Erasure on the 1st BEC: $(Y_1, Y_2, \mathcal{U}_1) = (?\ , \mathcal{U}_2\ , \mathcal{U}_1)$

iii) Erasure on the 2nd BEC: $(Y_1, Y_2, \mathcal{U}_1) = (\mathcal{U}_1 \oplus \mathcal{U}_2, ?\ , \mathcal{U}_1)$

iv) Two erasures: $(Y_1, Y_2, \mathcal{U}_1) = (?\ , ?\ , \mathcal{U}_1)$

Polarization

- The polarization effect can be obtained by recursively applying the BCT.



- 1st copy of W: input X_1, output Y_1
- 2nd copy of W: input X_2, output Y_2

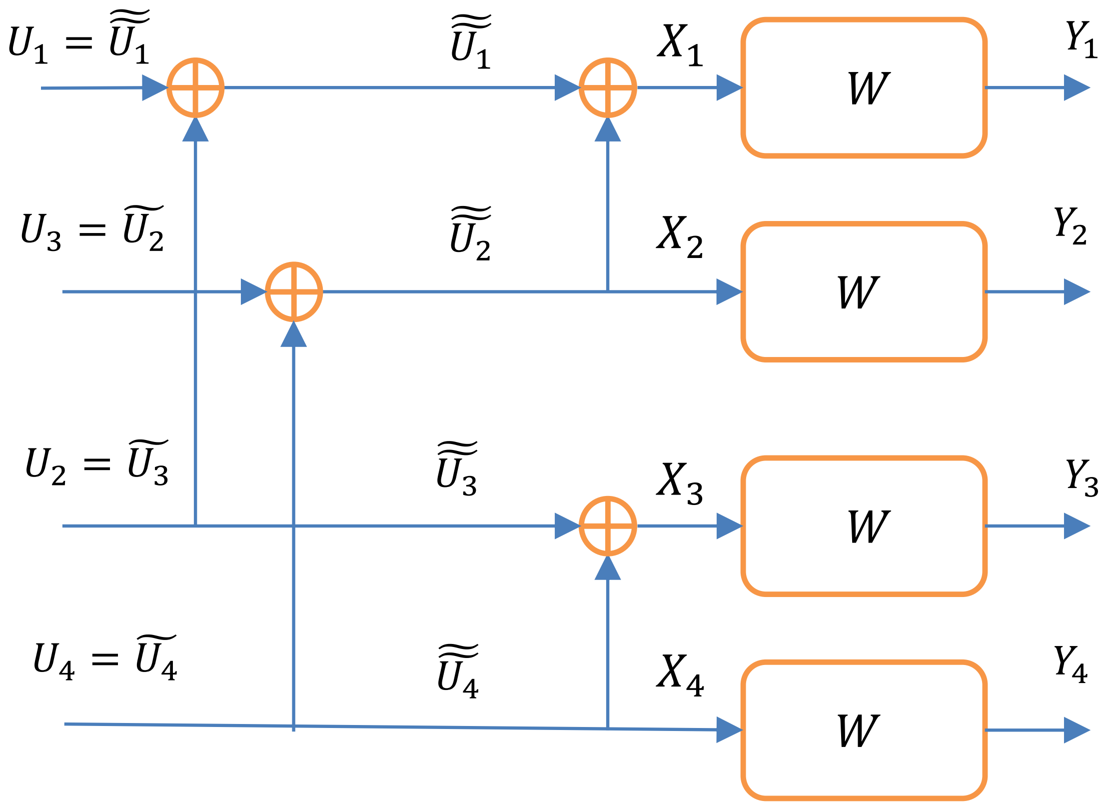
$$W^-: \tilde{\tilde{U}}_1 \to (Y_1, Y_2),\ \text{where } X_1 = \tilde{\tilde{U}}_1 \oplus \tilde{\tilde{U}}_2$$

$$W^+: \tilde{\tilde{U}}_2 \to \left(Y_1, Y_2, \tilde{\tilde{U}}_1\right),\ \text{where } X_2 = \tilde{\tilde{U}}_2$$

- 3rd copy of W: input X_3, output Y_3
- 4th copy of W: input X_4, output Y_4

$$W^-: \tilde{\tilde{U}}_1 \to (Y_1, Y_2),\ \text{where } X_1 = \tilde{\tilde{U}}_1 \oplus \tilde{\tilde{U}}_2$$

$$W^+: \tilde{\tilde{U}}_2 \to \left(Y_1, Y_2, \tilde{\tilde{U}}_1\right),\ \text{where } X_2 = \tilde{\tilde{U}}_2$$



- Application of BCT to W^{-}:

1st copy of W: input $\tilde{\tilde{U}}_1$, output (Y_1, Y_2)

2nd copy of W: input $\tilde{\tilde{U}}_3$, output (Y_3, Y_4)

$W^{--}: \tilde{\tilde{U}}_1 \rightarrow (Y_1, Y_2, Y_3, Y_4)$, where $\tilde{\tilde{U}}_1 = \tilde{U}_1 \oplus \tilde{U}_3$

$W^{-+}: \tilde{\tilde{U}}_2 \rightarrow (Y_1, Y_2, Y_3, Y_4, \tilde{U}_1)$, where $\tilde{\tilde{U}}_3 = \tilde{U}_3$

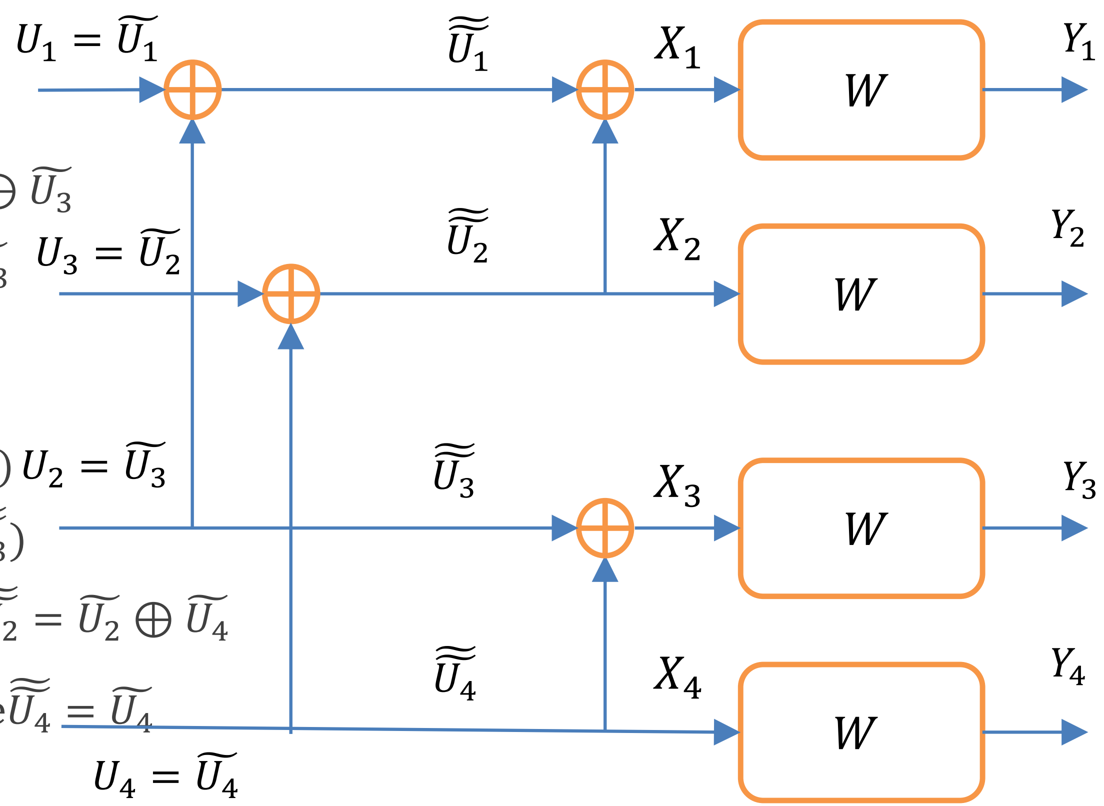
- Application of BCT to W^{+}:

1st copy of W: input $\tilde{\tilde{U}}_2$, output (Y_1, Y_2, U_1)

2nd copy of W: input $\tilde{\tilde{U}}_4$, output $(Y_3, Y_4, \tilde{\tilde{U}}_3)$

$W^{+-}: \tilde{\tilde{U}}_2 \rightarrow \left(Y_1, Y_2, Y_3, Y_4, \tilde{\tilde{U}}_1, \tilde{\tilde{U}}_3, \right)$, where $\tilde{\tilde{U}}_2 = \tilde{U}_2 \oplus \tilde{U}_4$

$W^{++}: \tilde{\tilde{U}}_2 \rightarrow \left(Y_1, Y_2, Y_3, Y_4, \tilde{\tilde{U}}_1, \tilde{\tilde{U}}_3, U_2^{\sim}\right)$, where $\tilde{\tilde{U}}_4 = \tilde{U}_4$



- By renaming $U_1 = \widetilde{U_1}, U_2 = \widetilde{U_2}, U_3 = \widetilde{U_3}, U_4 = \widetilde{U_4}$, we obtain

$W^{--}: U_1 \rightarrow (Y_1, Y_2, Y_3, Y_4)$

$W^{-+}: U_2 \rightarrow (Y_1, Y_2, Y_3, Y_4, U_1)$

$W^{+-}: U_3 \rightarrow (Y_1, Y_2, Y_3, Y_4, U_1, U_2)$

$W^{++}: U_4 \rightarrow (Y_1, Y_2, Y_3, Y_4, U_1, U_2, U_3)$

General procedure

- $W_n^{s_1 \dots s_l} : \mathcal{U}_{(s_1 \dots s_l)+1} \to (Y_1, \dots, Y_{2^l}, \mathcal{U}_1, \dots, \mathcal{U}_{(s_1 \dots s_l)}),$

where $s_1 \dots s_l \in \{0, -, 1, +\}$
$n = 2^l$
Inputs are $u_{(s_1 \dots s_l)}$
Binary string is $(s_1 \dots s_l)$

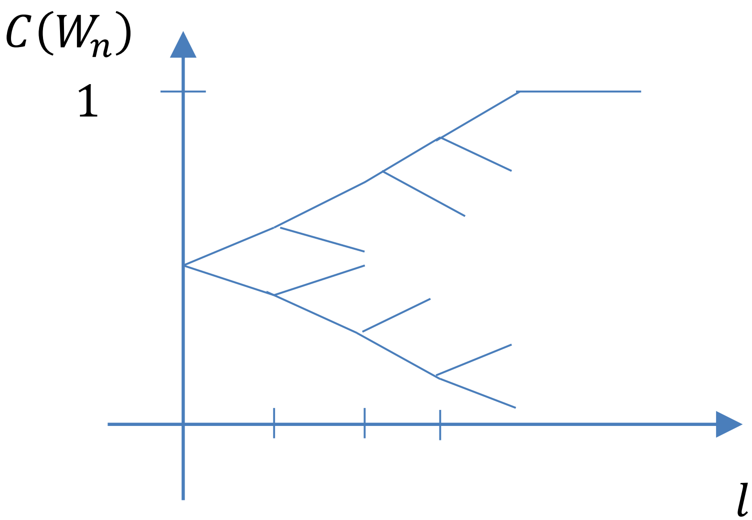
- The mutual information associated with this procedure is given by

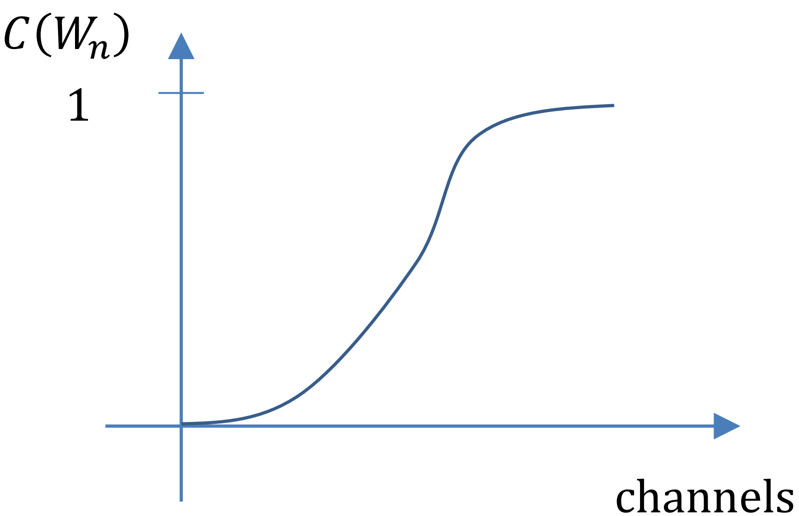
$$I\left(W_n^{s_1 \dots s_l}\right) = I(\mathcal{U}_{(s_1 \dots s_l)}, Y_1, \dots, Y_{2^l}, \mathcal{U}_1, \dots, \mathcal{U}_{(s_1 \dots s_l)})$$

- By measuring the capacity of these concatenated channels, we obtain

$$C(W_n) = I\left(W_n^{s_1 \dots s_l}\right),$$

which is illustrated by





Example 2

Consider the BCT as the following Kronecker matrix

$$\boldsymbol{T} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

Illustrate how this matrix can be used for polarization with $l = 1,2$

Solution:

We can write for $l = 1$ the inputs X_1 and X_2 as follows:

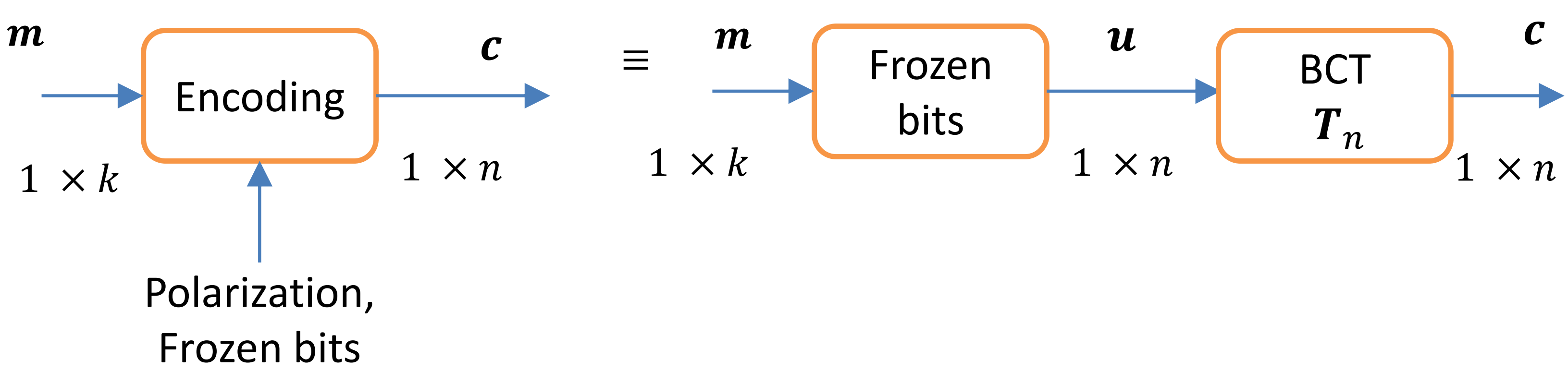
$$\boldsymbol{x} = [X_1\ X_2] = [\mathcal{U}_1\ \mathcal{U}_2]\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \boldsymbol{u}\,\boldsymbol{T}$$

For $l = 2$ the inputs X_1, X_2, X_3 and X_4 as follows:

$$\boldsymbol{x} = [X_1\ X_2\ X_3\ X_4] = [\mathcal{U}_1\ \mathcal{U}_2 \mathcal{U}_3\ \mathcal{U}_4]\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} = \widetilde{\boldsymbol{u}}\,(\boldsymbol{T} \otimes \boldsymbol{I}) = \widetilde{\boldsymbol{u}}\begin{bmatrix} \boldsymbol{T} & \boldsymbol{T} \\ \boldsymbol{T} & \boldsymbol{0} \end{bmatrix}$$

B. Encoding

- The encoding stage of polar codes relies on the following steps:
  - To repeatedly apply the BCT to n DMCs.
  - To select among the polarized channels those that are best.
  - The message bits $\boldsymbol{m}$ are transmited over good channels
  - The useless channels transmit the frozen bits $\boldsymbol{u}_f$.
- The encoding can be illustrated by



- For a codeword with length $n = 2^l$, we define the $(n, k, \mathcal{F}, \boldsymbol{u}_f)$ polar coding scheme $\mathcal{C}_{T_n}$ as follows.

- The length n codeword $\boldsymbol{c}$ of the polar code is given by

$$\boldsymbol{c} = \boldsymbol{u}\boldsymbol{T}_n,$$

where $\boldsymbol{T}_n = \boldsymbol{P}_n \boldsymbol{T}^{\odot l}$ is the l-fold BCT transformation, $\boldsymbol{P}_n$ is a permutation matrix and $\boldsymbol{u} \in \mathbb{R}^n$ is the input vector structured as

$$\boldsymbol{u} = [\boldsymbol{m} \mid \boldsymbol{u}_f],$$

where $\boldsymbol{u}_f \in \mathbb{R}^{n-k}$ contains the $n - k$ frozen bits.

- The code rate is given by

$$R = \frac{k}{n}$$

- The encoding complexity is given by $O(n\, log_2 n)$, which involves saving arithmetic operations with the l-fold BCT.

- It is common to employ other channel codes such as cyclic redundancy check (CRC) codes to further enhance the performance of polar codes.

Example 3

Consider the BCT as the following Kronecker matrix

$$\boldsymbol{T} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

Show how this matrix can be used to produce a codeword with length $n = 4$ using bit reversal ordering via the matrix

$$\boldsymbol{P}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Solution:

We can write for $l = 1$ the inputs X_1 and X_2 as follows:

$$\boldsymbol{c} = [X_1\ X_2] = [\mathcal{U}_1\ \mathcal{U}_2] \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \boldsymbol{u}\,\boldsymbol{T}$$

For $l = 2$ the inputs X_1, X_2, X_3 and X_4 constitute the codeword as follows:

$$\boldsymbol{c} = [X_1\ X_2\ X_3\ X_4] = [\mathcal{U}_1\ \mathcal{U}_2 \mathcal{U}_3\ \mathcal{U}_4] \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} = \tilde{\boldsymbol{u}}\,(\boldsymbol{T} \otimes \boldsymbol{I}) = \tilde{\boldsymbol{u}} \begin{bmatrix} \boldsymbol{T} & \boldsymbol{T} \\ \boldsymbol{T} & \boldsymbol{0} \end{bmatrix}$$

Using the bit-reversal matrix, we can write

$$\widetilde{\boldsymbol{u}} = \boldsymbol{u}\boldsymbol{P}_4$$

and

$$\begin{aligned} \boldsymbol{c} = \widetilde{\boldsymbol{u}}\,(\boldsymbol{T} \otimes \boldsymbol{I}) &= \boldsymbol{u}\boldsymbol{P}_4 \begin{bmatrix} \boldsymbol{T} & \boldsymbol{T} \\ \boldsymbol{T} & \boldsymbol{0} \end{bmatrix} \\ &= \boldsymbol{u}\boldsymbol{P}_4\boldsymbol{T}^{\odot 2} = \boldsymbol{u}\boldsymbol{T}_4 \\ &= [\mathcal{U}_1\ \mathcal{U}_2\mathcal{U}_3\ \mathcal{U}_4] \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

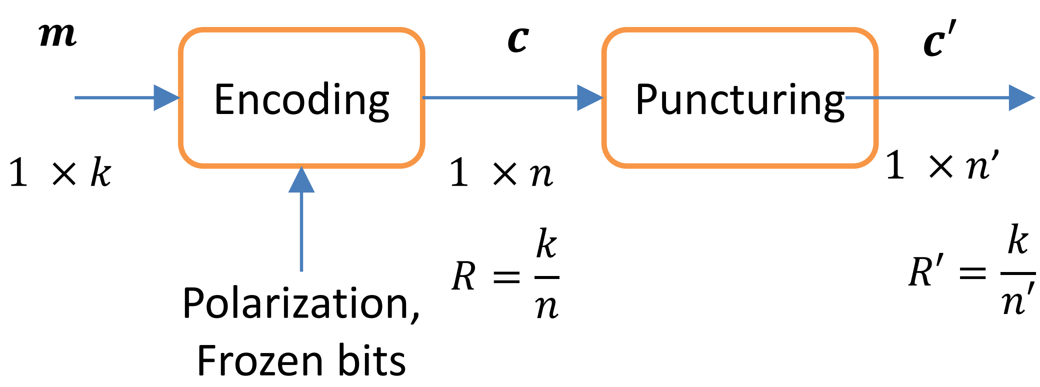
Therefore, in general we have

$$\boldsymbol{T}_n = \boldsymbol{P}_n\boldsymbol{T}^{\odot \boldsymbol{l}}$$

Puncturing

- Puncturing is important to adjust the codeword and/or the rate to the requirements of standards and applications.

- Due to the nature of polar codes they have codewords that are powers of 2, i.e., $n = 2^l$, which often requires puncturing of the frozen bits.



Example 4

Encode the message $\boldsymbol{m} = [1\ 0\ 0\ 1\ 0] = [\mathcal{U}_3\ \mathcal{U}_5 \mathcal{U}_6\ \mathcal{U}_7\ \mathcal{U}_8]$ using the frozen bits $\boldsymbol{u_f} = [\ 0\ 1\ 0]$ designated by the set $\mathcal{F} = \{1,2,4\}$ and the matrix

$$\boldsymbol{T}_8 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Consider also puncturing to produce a rate $\mathrm{R} = \frac{5}{7}$ polar code

Solution:

The codeword is given by

$$\boldsymbol{c} = \tilde{\boldsymbol{u}}\,(\boldsymbol{T} \otimes \boldsymbol{I}) = \boldsymbol{u}\boldsymbol{P}_8\boldsymbol{T}^{\odot 3} = \boldsymbol{u}\boldsymbol{T}_8$$

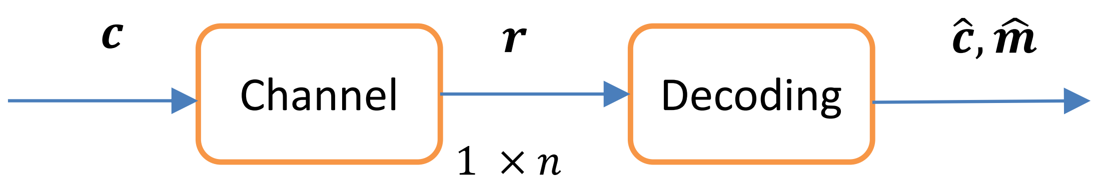
$$= [0\ 1\ 1\ 0\ 0\ 0\ 1\ 0]\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$= [1\ 0\ 1\ 1\ 0\ 1\ 1\ 1]$$

In order to puncture this code to obtain a rate $\mathrm{R} = \frac{5}{7}$ polar code, we need to discard 1 frozen bit, which would result in

$$\boldsymbol{c} = [0\ 1\ 1\ 0\ 1\ 1\ 1]$$

C. Decoding



- The most used decoding strategy is based on successive cancellation and is given by the following decisions:

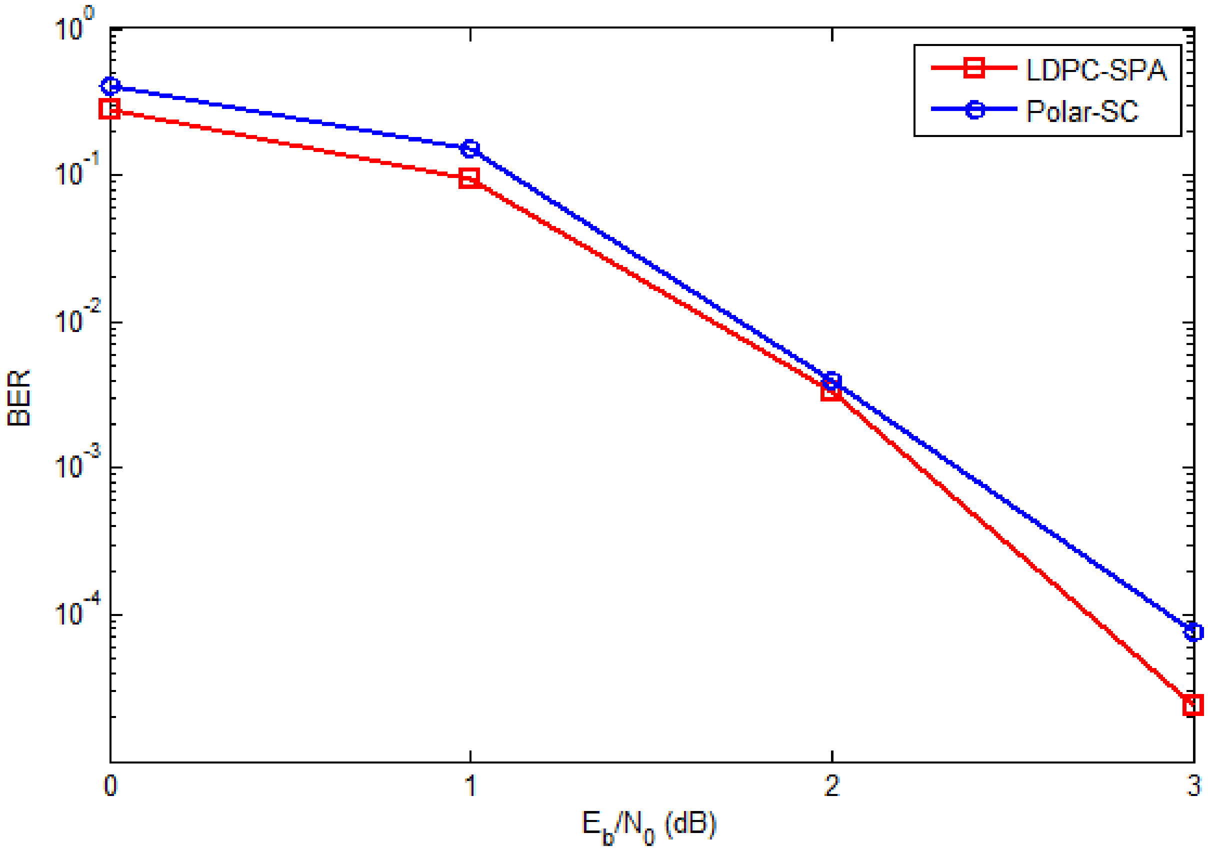
$$\hat{u}_k = \begin{cases} u_{k,} & if\ k \in \mathcal{F} \\ \psi_k(\boldsymbol{r}, \hat{u}_{k-1}) & if\ k \notin \mathcal{F}' \end{cases}$$

where the decoding functions is given by

$$\psi_k(\boldsymbol{r}, \hat{u}_{k-1}) = \begin{cases} 1, & if \log\left(\frac{P(\boldsymbol{r}, \hat{u}_{k-1} | m = 1)}{P(\boldsymbol{r}, \hat{u}_{k-1} | m = -1)}\right) \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Example 5

Compare the performance of an LDPC code and a polar code with puncturing, $n = 1920$ and rate $R = \frac{1}{4}$.



10^0
10^-1
10^-2
10^-3
10^-4
BER
0
1
2
3
Eb/N0 (dB)
LDPC-SPA
Polar-SC