TECHNICAL UNIVERSITY OF CRETE
ELECTRICAL AND COMPUTER ENGINEERING DEPARTMENT

Pascal-matrix polar coding for the wiretap erasure channel

by

Titos Agapakis

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DIPLOMA OF

ELECTRICAL AND COMPUTER ENGINEERING

THESIS COMMITTEE

Professor George N. Karystinos, *Thesis Supervisor*
Professor Aggelos Bletsas
Professor Athanasios Liavas

July 2024

Abstract

Wyner introduced a special case of the wiretap channel in 1975 which consists of two communications channels. The first channel is between the transmitter and the receiver whereas the second channel is between the transmitter and the eavesdropper. This work focuses on the utilization of coding schemes that reach the maximum transmission rate between the transmitter and the receiver while preserving the secrecy of the transmitted data from the eavesdropper. To achieve this we utilize polar codes, which were introduced by Arikan in 2009 and achieve Shannon's capacity with low encoding and decoding complexity. They were presented initially for binary-input discrete memoryless channels and later on for arbitrary-input channels. In this thesis, we begin by implementing the original polar codes for the binary symmetric, binary erasure, and q-ary erasure channels. Then, for the q-ary erasure channel, we present a different approach to construct the encoder and the decoder based on the Pascal-matrix. Finally, we apply both polar coding schemes on the wiretap channel and compare their performance.

Acknowledgements

I would like to thank my supervisor, Professor G. Karystinos for his guidance and patience during my thesis.
I would like to thank my thesis committee members, Professor A. Bletsas and Professor A. Liavas for their participation.
I would also like to thank my friends for their support and encouragement throughout my studies.
Finally, I would like to thank my family for their endless love and support and dedicate this thesis to them.

Table of Contents

1 Introduction **7**

2 Polarization for Binary-input Channels **8**
2.1 Symmetric Capacity . . . 8
2.2 Bhattacharyya Parameter . . . 8
2.3 Binary Symmetric Channel (BSC) . . . 8
2.4 Binary Erasure Channel (BEC) . . . 8
2.5 Basic Polarization . . . 9
2.5.1 Channel Combining . . . 9
2.5.2 Channel Splitting . . . 11
2.6 Encoding . . . 16
2.7 Decoding . . . 17
2.8 Performance on BEC . . . 19

3 Polarization for Prime-input Erasure Channels **20**
3.1 Symmetric Capacity . . . 20
3.2 q-ary input Erasure Channel . . . 20
3.3 Arikan's approach . . . 21
3.3.1 Performance on TEC . . . 22
3.4 Pascal-matrix approach . . . 23
3.4.1 Pascal-matrix based construction . . . 24
3.4.2 Encoding . . . 26
3.4.3 Decoding . . . 27
3.4.4 Performance on q-ary EC . . . 27

4 Polar Codes for Secrecy **28**
4.1 The Wiretap Communication Model . . . 28
4.2 The Secrecy Capacity of the Wiretap Channel . . . 29
4.3 Encoding and Decoding . . . 29
4.4 Performance Results . . . 31

Appendix **38**

References **47**

List of Figures

1 BSC(p). 9
2 BEC(ϵ). 9
3 The channel W_2. 9
4 The channel W_4 and its relation to W_2 and W. 10
5 Split channel $W_N^{(i)}$. 11
6 Split channels of the vector channel W_2. 12
7 Symmetric Capacity of BSC and BEC before and after the basic polarization step. 13
8 Bhattacharyya parameter of BSC and BEC before and after the basic polarization step. 13
9 Channel polarization for a BEC(0.5) with $N = 2^{12}$. 14
10 Channel polarization for a BEC(0.3) with $N = 2^{12}$. 15
11 Symmetric capacity vs Normalized channel index for different block lengths for the BEC. 15
12 Recursive construction of W_N from two copies of $W_{N/2}$. 16
13 The channel transformation process, $N = 8$. 18
14 BER for different block lengths over BEC. 19
15 q-ary input Erasure Channel. 20
16 Symmetric capacity for TEC before and after the basic polarization step. . 21
17 SER for different block lengths over TEC of the original scheme. 22
18 The Sierpinski triangle. 23
19 $G_{128}/[\text{Pascal matrix}]_{128}$ mod 2. 23
20 The 5×5 Pascal matrix. 23
21 Symmetric capacity of the proposed Pascal-matrix code for different block lengths over a TEC(0.5) and a 5-input EC, with $\epsilon = 0.5$. 24
22 Symmetric capacity comparison between the conventional polar code and the proposed polar code base on the Pascal-matrix for two different input sizes. 25
23 Recursive construction of W_N from q copies of $W_{N/q}$. 26
24 Symbol error rate as a function of erasure probability ϵ for transmission over a TEC. 27
25 Symbol error rate as a function of erasure probability ϵ for transmission over a 5-input erasure channel. 28
26 The wiretap channel. 29
27 Secrecy structure. 30
28 BER vs Normalized channel index for different block lengths over binary erasure wiretap channel. 31
29 Rate of the coding scheme as a function of the tolerance of the BER $\delta \in [0, 0.2]$. 31
30 SER vs Normalized channel index for different block lengths over Pascal-matrix ternary erasure wiretap channel. 32
31 Rate of the coding scheme as a function of the tolerance of the SER $\delta \in [0, 0.2]$. 32
32 SER of ternary erasure channel using the original scheme vs using the proposed pascal-matrix scheme. 33
33 Rate of the coding scheme as a function of the tolerance of the SER $\delta \in [0, 0.2]$. 33
34 The symmetric channel formed between the original message u and the estimated message $\hat{u}$. 34

35 Transition probability matrix for the Ternary Symmetric Channel (when the underlying ternary erasure channel has $\epsilon = 0.1$, I = 0.9, R = 0.85). 34
36 Transition probability matrix for the Ternary Symmetric Channel (when the underlying ternary erasure channel has $\epsilon = 0.4$, I = 0.6, R = 0.55). 34
37 Transition probability matrix for the Ternary Symmetric Channel (when the underlying ternary erasure channel has $\epsilon = 0.8$, I = 0.2, R = 0.15). . . 34
38 SER of 5-input erasure channel using the original scheme vs using the proposed pascal-matrix scheme. 35
39 Rate of the coding scheme as a function of the tolerance of the SER $\delta \in [0, 0.2]$. 35
40 Transition probability matrix for the 5-input Symmetric Channel (when the underlying 5-input erasure channel has $\epsilon = 0.1$, I = 0.9, R = 0.85). 36
41 Transition probability matrix for the 5-input Symmetric Channel (when the underlying 5-input erasure channel has $\epsilon = 0.4$, I = 0.6, R = 0.55). 36
42 Transition probability matrix for the 5-input Symmetric Channel (when the underlying 5-input erasure channel has $\epsilon = 0.8$, I = 0.2, R = 0.15). 36
43 The q-ary input symmetric channel with its transition probability matrix. . 37
44 Synthesized channel W_3. 41

1 Introduction

In 1948, Shannon demonstrated that for any communication channel there exists a coding scheme that allows information to be transmitted reliably over a noisy channel at a rate up to the capacity of the channel, known as Shannon capacity. The search for codes that achieve Shannon's capacity proved quite challenging. In 2009, Arikan introduced polar codes, which are the first provably capacity-achieving codes for binary-input discrete memoryless channels with low encoding and decoding complexity. Later, he generalized the proposed coding scheme to arbitrary-input discrete memoryless channels. One year after his landmark paper, Shannon introduced the idea of information theoretic security, showing that secure communication is possible if we utilize a shared secret key between the transmitter and the receiver.

In 1975, Wyner proposed an alternative approach to secure communication schemes by introducing the wiretap channel model. In this model, a transmitter is sending a message to the legitimate receiver. However, an eavesdropper receives the transmitted message as well. Wyner concluded that, even if the wiretapper is aware of the encoding scheme that the transmitter utilizes, secrecy is preserved by the presence of greater noise in the eavesdropper channel.

In this thesis, we begin by presenting the basic polarization techniques for binary symmetric channels. We extend the original scheme to q-ary input discrete memoryless erasure channels. We present an alternative approach to construct polar codes, called Pascal-matrix polar coding [11], that was introduced in 2016 and compare the performance of the two schemes. In the end, we apply all the above in Wyner's wiretap channel model and make performance comparisons between them.

2 Polarization for Binary-input Channels

2.1 Symmetric Capacity

Given a binary-discrete memoryless channel (B-DMC) $W : \mathcal{X} \to \mathcal{Y}$, with input alphabet $\mathcal{X} = \{0, 1\}$, we define the symmetric capacity as

$$I(W) \triangleq \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} \frac{1}{2} W(y|x) \log_2 \frac{W(y|x)}{\frac{1}{2}W(y|0) + \frac{1}{2}W(y|1)} \quad . \tag{1}$$

Symmetric capacity is the mutual information between the input and the output of the channel when the input is uniformly distributed ($P(\mathcal{X} = 0) = P(\mathcal{X} = 1) = \frac{1}{2}$). $I(W)$ is a measure of *rate* in a channel. It is well-known that reliable communication is possible over a symmetric B-DMC at any rates up to $I(W)$.

2.2 Bhattacharyya Parameter

The Bhattacharyya parameter between the input and output of a B-DMC W is defined as

$$Z(W) \triangleq \sum_{y \in \mathcal{Y}} \sqrt{W(y|0)W(y|1)}. \tag{2}$$

Its an upper bound on the probability of maximum-likelihood (ML) decision error when W is used only once to transmit a 0 or 1. Therefore, it will be used as a measure of *reliability* of a channel. The channel is reliable when $Z(W)$ is small.

2.3 Binary Symmetric Channel (BSC)

The binary symmetric channel, shown in Fig. 1, takes a binary input $\mathcal{X} = \{0, 1\}$ and produces the output $\mathcal{Y} = \{0, 1\}$. The input is transmitted either correctly with probability $1 - p$ or it is inverted with probability p. The symmetric capacity of the BSC is

$$I(W) = 1 + p \log_2(p) + (1 - p) \log_2(1 - p). \tag{3}$$

The proof is available in [2]. The Bhattacharyya parameter of the BSC is

$$Z(W) = 2\sqrt{p(1 - p)}. \tag{4}$$

2.4 Binary Erasure Channel (BEC)

The binary erasure channel, shown in Fig. 2, takes a binary input $\mathcal{X} = \{0, 1\}$ and produces the output $\mathcal{Y} = \{0, 1, ?\}$. The input is transmitted either correctly with probability $1 - \epsilon$ or it is erased with probability ϵ. The symmetric capacity of the BEC is

$$I(W) = 1 - \epsilon. \tag{5}$$

The proof is available in [2]. The Bhattacharyya parameter of the BEC is

$$Z(W) = \epsilon. \tag{6}$$

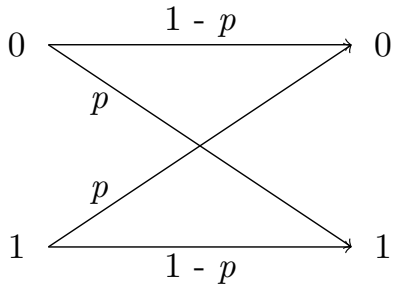


Figure 1: BSC(p).

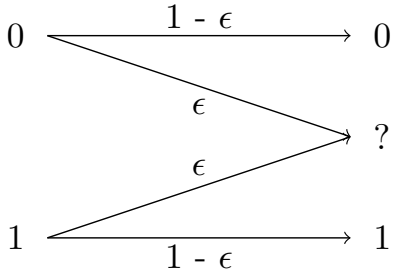


Figure 2: BEC(ϵ).

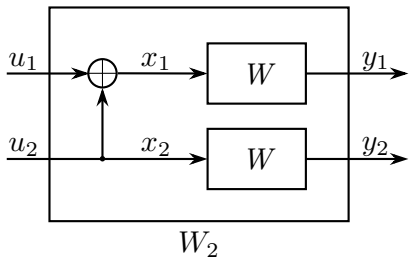


Figure 3: The channel W_2.

2.5 Basic Polarization

Channel polarization is an operation by which, out of N independent copies of a given discrete memory channel W, one manufactures a second set of N channels $\{W_N^{(i)} : 1 \leq i \leq N\}$ that show a polarization effect in the sense that, as N becomes large, the symmetric capacity terms $\{I(W_N^{(i)})\}$ tend towards 0 or 1 for all but a vanishing fraction of indices i. This operation consists of two phases, channel combining and channel splitting [1].

2.5.1 Channel Combining

This phase combines copies of a given B-DMC W in a recursive way to produce a vector channel $W_N : \mathcal{X}^N \to \mathcal{Y}^N$, where $N = 2^n, n \geq 0$. The first level of the recursion combines two independent copies of W, we set x_1 to be the result of the XOR between u_1 and u_2 and x_2 to be equal to u_2, to obtain the channel $W_2 : \mathcal{X}^2 \to \mathcal{Y}^2$, shown in Fig. 3. This operation is nothing but a linear transformation over GF(2).
The channel W_2 is defined as

$$\begin{aligned} W_2(y_1, y_2|u_1, u_2) &= W_2(y_1, y_2|x_1, x_2) \\ &= W(y_1|x_1)W(y_2|x_2) = W(y_1|u_1 \oplus u_2)W(y_2|u_2) \end{aligned} \tag{7}$$

for all $y_1, y_2 \in \mathcal{Y}^2$, $u_1, u_2 \in \mathcal{X}^2$ with u_1, u_2 independent.

The next level of the recursion is to combine two independent copies of W_2 to create the channel $W_4 : \mathcal{X}^4 \to \mathcal{Y}^4$, shown in Fig. 4. The permutation operation, illustrated as R_4, separates odd-indexed input from even-indexed input, (s_1, s_2, s_3, s_4) to $v_1^4 = (s_1, s_3, s_2, s_4)$. The mapping $u_1^4 \to x_1^4$ from the input of channel W_4 to the input of the four starting channels W, W^4, can be written also as $x_1^4 = u_1^4 G_4$. Matrix G_N is called generator matrix and is of size N and is defined in [1] as

$$G_N = B_N F^{\otimes n} \tag{8}$$

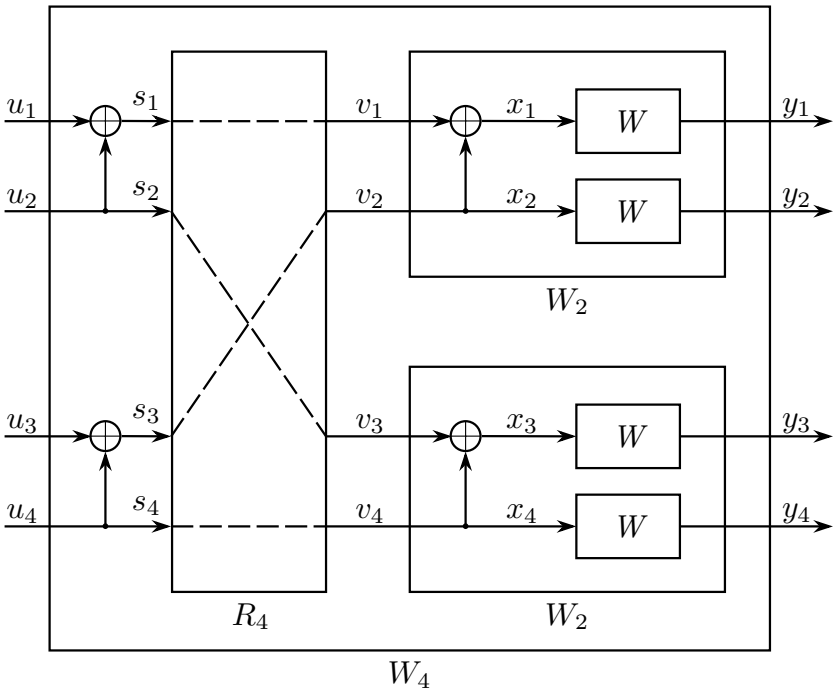


Figure 4: The channel W_4 and its relation to W_2 and W.

where $F^{\otimes n}$ is the n^{th} Kronecker product of the kernel matrix $F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, B_N is called bit-reversal and is defined as $B_N = R_N(I_2 \otimes B_{N/2})$, and R_N is constructed by rearranging the rows of the $N \times N$ identity matrix I_N according to the bit reversal order.

To compute G_4, we start by computing the Kronecker product $F^{\otimes 2}$ as

$$F^{\otimes 2} = F \otimes F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 \cdot \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} & 0 \cdot \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \\ 1 \cdot \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} & 1 \cdot \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$R_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad I_2 \otimes B_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The bit-reversal permutation matrix is

$$B_4 = R_4(I_2 \otimes B_2) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The generator matrix G_4 is

$$G_4 = B_4 F^{\otimes 2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$x_1^4 = u_1^4 G_4 = \begin{bmatrix} u_1 & u_2 & u_3 & u_4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} u_1 \oplus u_2 \oplus u_3 \oplus u_4 & u_3 \oplus u_4 & u_2 \oplus u_4 & u_4 \end{bmatrix}$$

We observe that the mapping $u_1^N \to v_1^N$ is linear over GF(2). Therefore, it follows by induction that the overall mapping $u_1^N \to x_1^N$, from the input of the synthesized channel W_N to the input underlying of raw channel W^N, is also linear thus $x_1^N = u_1^N G_N$. The relation between the channels W_N and W^N is given by

$$W_N(y_1^N|u_1^N) = W^N(y_1^N|u_1^N G_N) = W^N(y_1^N|x_1^N) \tag{9}$$

for all $y_1^N \in \mathcal{Y}^N, u_1^N \in \mathcal{X}^N$.

2.5.2 Channel Splitting

Having synthesized the vector channel W_N out of raw channels W_N, we split W_N channel back into a set of N bit-channels $W_N^{(i)}$ for $1 \leq i \leq N$. These channels are defined by the transition probabilities

$$W_N^{(i)}(y_1^N, u_1^{i-1}|u_i) \triangleq \sum_{u_{i+1}^N \in \mathcal{X}^{N-i}} \frac{1}{2^{N-1}} W_N(y_1^N|u_1^N) \tag{10}$$

where (y_1^N, u_1^{i-1}) denotes the output of $W_N^{(i)}$ and u_i its input.
For the proof, consult the Appendix.

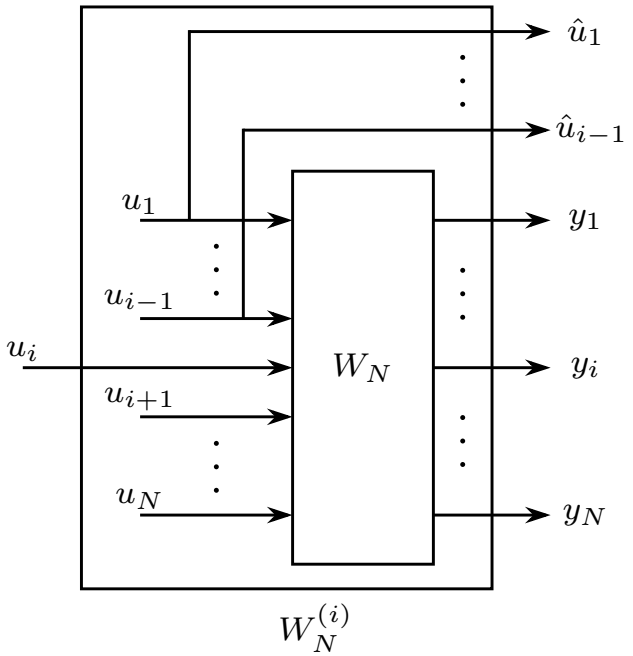


Figure 5: Split channel $W_N^{(i)}$.

Figure 6: Split channels of the vector channel W_2.

If we apply the above to the channel W_2 of Fig. 3, we get the split channels $W_2^{(1)}$ and $W_2^{(2)}$, as shown in Fig. 6. Note that after being estimated, u_1 is used as an output in the split channel $W_2^{(2)}$ in order to estimate u_2. The transition probabilities for each channel are the following

$$W_2^{(1)}(y_1, y_2|u_1) = \frac{1}{2}\sum_{u_2} W(y_1|u_1 \oplus u_2)W(y_2|u_2), \tag{11}$$

$$W_2^{(2)}(y_1, y_2, u_1|u_2) = \frac{1}{2}W(y_1|u_1 \oplus u_2)W(y_2|u_2). \tag{12}$$

For the proof of (11) and (12), consult the Appendix.

It is essential to highlight that given the synthetic channel W_2: $\{W, W\} \mapsto \left\{(W_2^{(1)}, W_2^{(2)})\right\} \iff \{W, W\} \mapsto \{W', W''\}$ the following relations hold

$$I(W') + I(W'') = 2I(W), \tag{13}$$

$$I(W') \leq I(W) \leq I(W''). \tag{14}$$

The equality in (13) indicates that channel polarization is a capacity preserving operation. If W is either a perfect channel($I(W) = 1$) or a completely noisy one($I(W) = 0$) then the symmetric capacity remains unchanged under a single-step transform, $I(W') = I(W'') = I(W)$.
Regarding the reliability parameter, the following relations hold

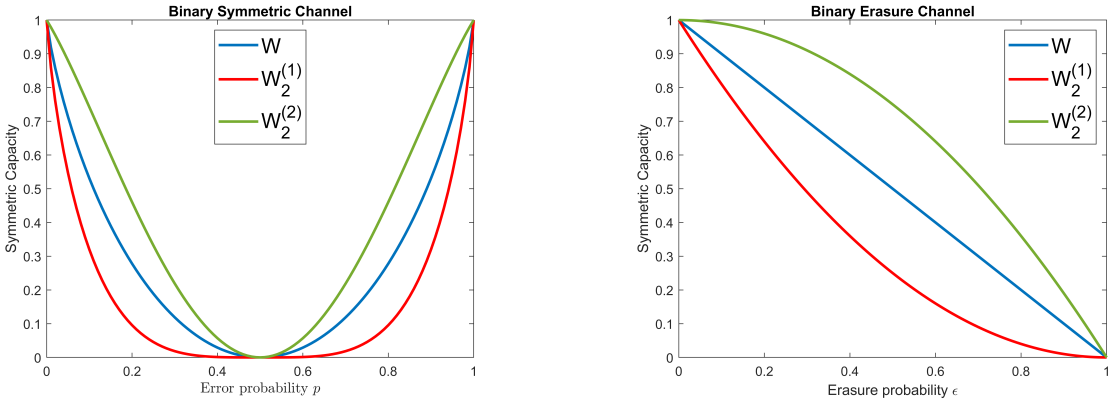
$$Z(W'') = Z(W)^2, \tag{15}$$

$$Z(W') \leq 2Z(W) - Z(W)^2, \tag{16}$$

$$Z(W') \geq Z(W) \geq Z(W''). \tag{17}$$

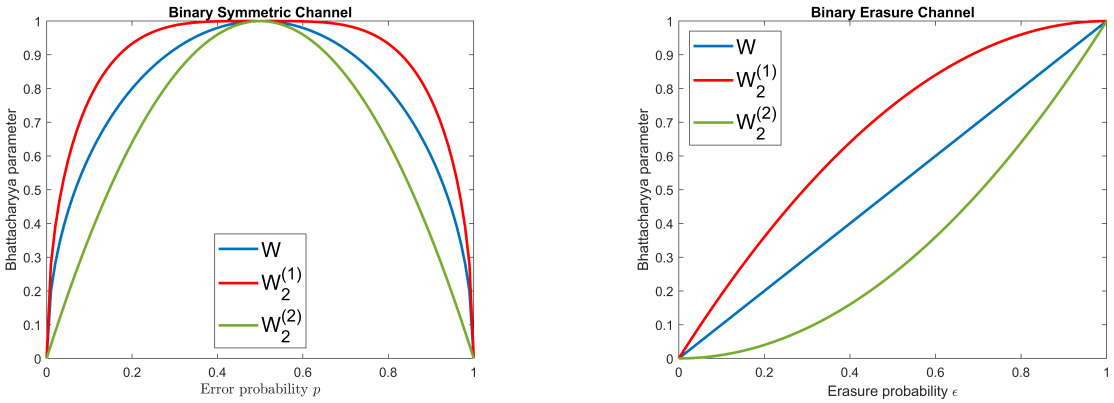
In the case that W is a BEC, equality holds for (16).

In order to illustrate the behaviour of $I(W)$ and $Z(W)$ of the split channels after the polarization step for the synthetic channel W_2: $\{W, W\} \mapsto \left\{(W_2^{(1)}, W_2^{(2)})\right\} \iff \{W, W\} \mapsto \{(y_1^2; u_1), (y_1^2, u_1; u_2)\}$ we need to work separately for the BSC and the BEC.
For the BSC, we have to calculate the transition probabilities $p(y_1, y_2|u_1)$, $p(y_1, y_2, u_1|u_2)$, one-by-one and then use them in (18) to (21). For the BEC, we work with (15) and (16) since (5) and (6) hold.



(a) BSC (b) BEC

Figure 7: Symmetric Capacity of BSC and BEC before and after the basic polarization step.



(a) BSC (b) BEC

Figure 8: Bhattacharyya parameter of BSC and BEC before and after the basic polarization step.

$$I(u_1; y_1, y_2) = \sum_{y_1, y_2, u_1} \frac{1}{2} p(y_1, y_2|u_1) \log_2 \frac{p(y_1, y_2|u_1)}{\frac{1}{2}p(y_1, y_2|u_1 = 0) + \frac{1}{2}p(y_1, y_2|u_1 = 1)}. \tag{18}$$

$$I(u_2; y_1, y_2, u_1) = \sum_{y_1, y_2, u_1, u_2} \frac{1}{2} p(y_1, y_2, u_1|u_2) \log_2 \frac{p(y_1, y_2, u_1|u_2)}{\frac{1}{2}p(y_1, y_2, u_1|u_2 = 0) + \frac{1}{2}p(y_1, y_2, u_1|u_2 = 1)}. \tag{19}$$

$$Z(W') = \sum_{y_1, y_2} \sqrt{p(y_1, y_2|u_1 = 0)p(y_1, y_2|u_1 = 1)}. \tag{20}$$

$$Z(W'') = \sum_{y_1, y_2, u_1} \sqrt{p(y_1, y_2, u_1|u_2 = 0)p(y_1, y_2, u_1|u_2 = 1)}. \tag{21}$$

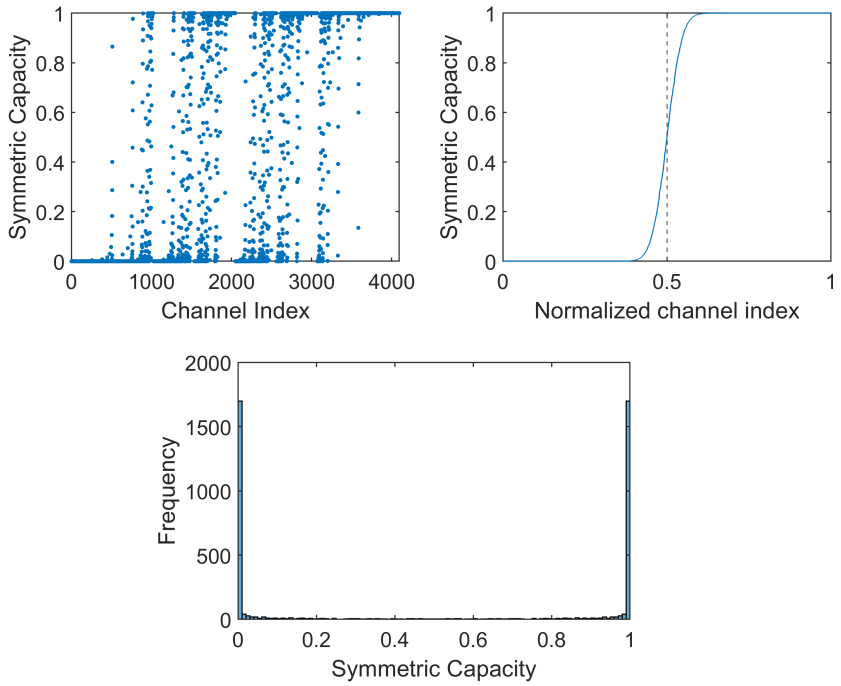


Figure 9: Channel polarization for a BEC(0.5) with $N = 2^{12}$.

The two bit channels created after the polarization step are defined as upgraded and degraded channels with respect to the original one in terms of symmetric capacity, as shown in Fig. 7. For both the BSC and the BEC, the upgraded channel $W_2^{(2)}$ has a higher capacity than the original channel while the degraded has a lower capacity, verifying (14).

In Fig. 8, we observe that Bhattacharyya parameter of the upgraded channel is lower for both the BSC and the BEC, making it a more reliable channel than the degraded channel which has a higher Bhattacharyya parameter, verifying (17).

By using (15) and (16), we can recursively calculate the symmetric capacities of the manufactured channels for the case that W is a BEC, using the following formulas and (5) [1].

$$I(W_N^{(2i-1)}) = I(W_{N/2}^{(i)})^2, \tag{22}$$

$$I(W_N^{(2i)}) = 2I(W_{N/2}^{(i)}) - I(W_{N/2}^{(i)})^2. \tag{23}$$

In Fig. 9, we analyze a BEC(0.5) with symmetric capacity $I(W) = 0.5$ bits/channel. Almost 50% of the channels are perfect($I(W)$ close to 1) and almost 50% are useless ($I(W)$ close to 0). Moreover, it is clear that small indexed channels are extremely unreliable, whereas the higher indexed channels tend to be more reliable. In Fig. 10, we analyze a BEC(0.3) with symmetric capacity $I(W) = 0.7$ bits/channel. Decreasing the erasure probability ϵ has a noteworthy impact on our results. Almost 70% of the channels are perfect and 30% are useless. The effect described above is called channel polarization.

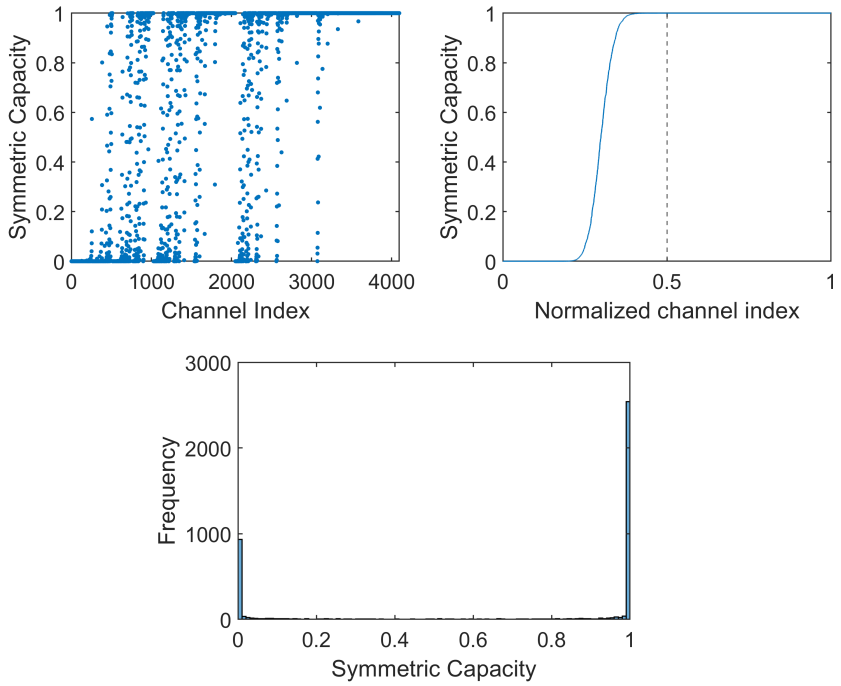


Figure 10: Channel polarization for a BEC(0.3) with $N = 2^{12}$.

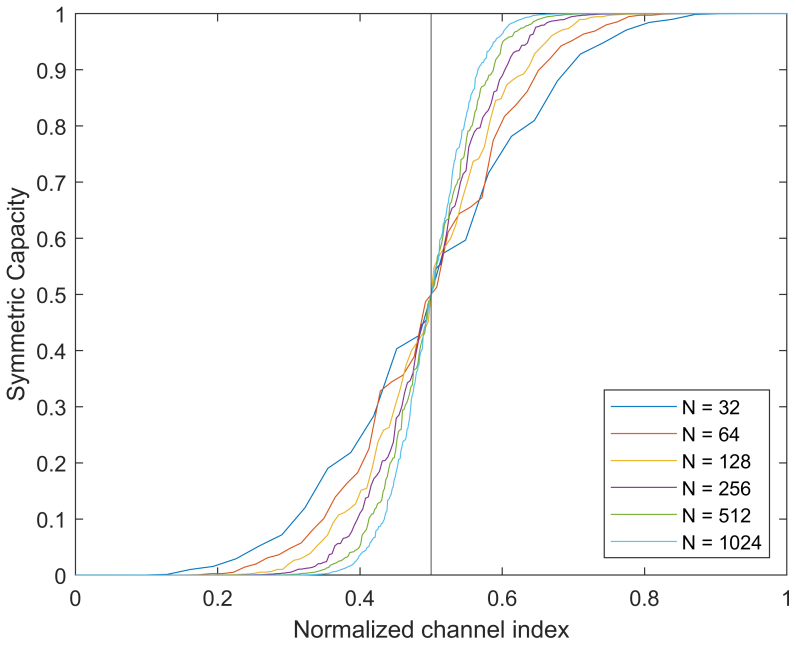


Figure 11: Symmetric capacity vs Normalized channel index for different block lengths for the BEC.

In Fig. 11, we observe that as the block length N increases the channels polarize better in the sense that more channels become either useless($I(W)$ close to 0) or perfect ($I(W)$ close to 1).

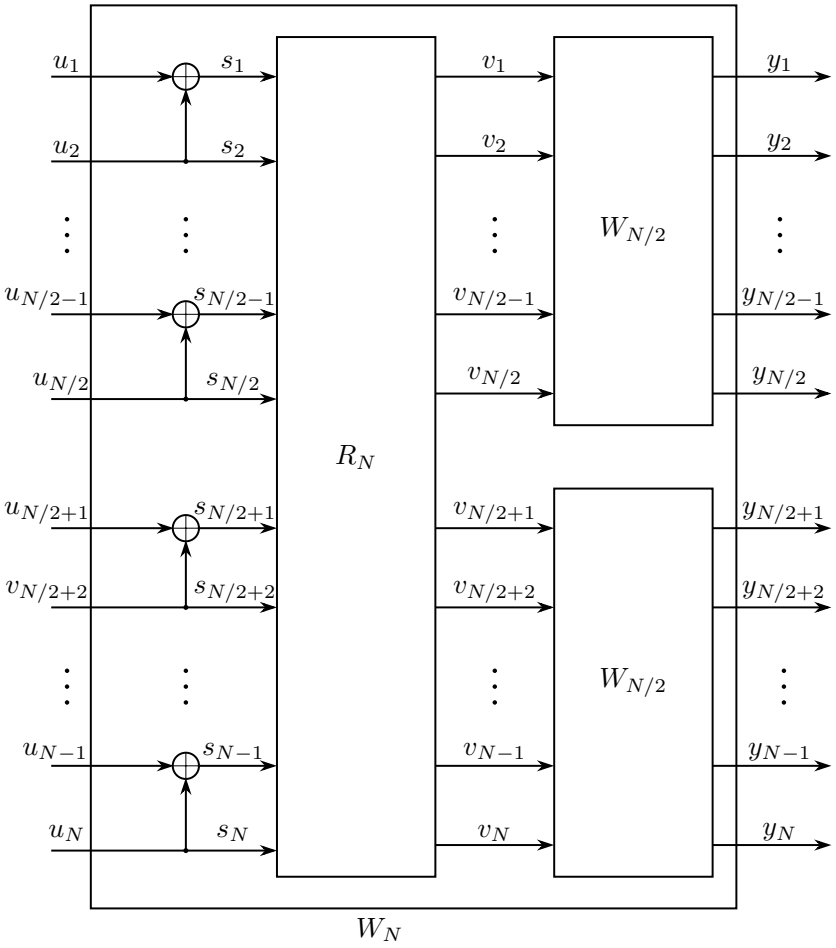


Figure 12: Recursive construction of W_N from two copies of $W_{N/2}$.

2.6 Encoding

The code construction method, based on matrix G, we presented earlier is of $\mathcal{O}(N^2)$ complexity. Consequently, it is not suitable for large block lengths N that we are interested in. For this reason, a recursive approach of the encoding procedure is needed, illustrated in Fig. 12. The Vector channel W_N is formed by two independent copies of $W_{N/2}$, the operator R_N is a permutation, known as reverse shuffle operation that splits odd-indexed from even-indexed channels. The odd-index group enters the first copy of $W_{N/2}$, whereas the even-index group enters the second copy $W_{N/2}$. We repeat this process recursively until vector channel consists of two copies of the original channel $W \triangleq W_1$. We transmit with rate $R = \lfloor K/N \rfloor$, where K is the number of information bits. These bits will be placed to the information vector $u_A \in \mathcal{X}^K$ and will be send for transmission through the K-most reliable split channels. The unreliable channels will have as input, the vector $u_{A^c} \in \mathcal{X}^{N-K}$ filled with frozen(known to the decoder) bits.

If we assume that the complexity of a scalar mod-2 addition is 1 unit, the complexity of the reverse shuffle operation R_N is N units of time, the encoding complexity with the help of Master Theorem is

$$\begin{aligned} &\mathcal{T}(N) = \frac{N}{2} + \mathcal{O}(N) + 2\mathcal{T}(\frac{N}{2}) \Rightarrow \\ &\mathcal{T}(N) = \mathcal{O}(N \log_2 N). \end{aligned} \tag{24}$$

2.7 Decoding

Arikan proposed an algorithm for the decoding process called successive cancellation decoder [1]. The decoder sequentially calculates an estimation $\hat{u}_i$ of the transmitted bit u_i, based on the frozen and the previously estimated bits. If u_i is a frozen bit, $i \in A^c$, the decoder will set $\hat{u}_i$ to zero. If u_i is an information bit, $i \in A$, the decoder will decide on i^{th} bit ($1 \leq i \leq N$) after estimating all the previous bits u_1^{i-1} based on either the likelihood ratios (LRs) or the recursive formulas (28) and (29).
The likelihood ratio (LR) is defined as

$$L_N^{(i)}(y_1^N, \hat{u}_1^{i-1}) \triangleq \frac{W_N^{(i)}(y_1^N, \hat{u}_1^{i-1}|0)}{W_N^{(i)}(y_1^N, \hat{u}_1^{i-1}|1)}. \tag{25}$$

To estimate $L_N^{(i)}(y_1^N, \hat{u}_1^{i-1})$, we use the recursive formulas

$$L_N^{(2i-1)}(y_1^N, \hat{u}_1^{2i-2}) = \frac{L_{N/2}^{(i)}(y_1^{N/2}, \hat{u}_{1,o}^{2i-2} \oplus \hat{u}_{1,e}^{2i-2}) \cdot L_{N/2}^{(i)}(y_{N/2+1}^N, \hat{u}_{1,e}^{2i-2}) + 1}{L_{N/2}^{(i)}(y_{N/2+1}^N, \hat{u}_{1,e}^{2i-2}) + L_{N/2}^{(i)}(y_1^{N/2}, \hat{u}_{1,o}^{2i-2} \oplus \hat{u}_{1,e}^{2i-2})}, \tag{26}$$

$$L_N^{(2i)}(y_1^N, \hat{u}_1^{2i-1}) = [L_{N/2}^{(i)}(y_1^{N/2}, \hat{u}_{1,o}^{2i-2} \oplus \hat{u}_{1,e}^{2i-2})]^{1-2\hat{u}_{2i-1}} \cdot L_{N/2}^{(i)}(y_{N/2+1}^N, \hat{u}_{1,e}^{2i-2}). \tag{27}$$

For the proof of (26) and (27), consult the Appendix.

The SC Decoder generates its decision for $\hat{u}_i$ as follows

$$\hat{u}_i = \begin{cases} 0, & \text{if } u_i \text{ is a frozen bit} \\ 0, & \text{if } L_N^{(i)}(y_1^N, \hat{u}_1^{i-1}) \geq 1 \\ 1, & \text{otherwise.} \end{cases}$$

If we inspect (26) and (27), we notice that each LR value in the pair ($L_N^{(2i-1)}(y_1^N, \hat{u}_1^{2i-2})$, $L_N^{(2i)}(y_1^N, \hat{u}_1^{2i-1})$), is assembled from the same pair of LRs ($L_{N/2}^{(i)}(y_1^{N/2}, \hat{u}_{1,o}^{2i-2} \oplus \hat{u}_{1,e}^{2i-2})$, $L_{N/2}^{(i)}(y_{N/2+1}^N, \hat{u}_{1,e}^{2i-2})$). For this reason, we decide to use one shared matrix to store the LRs that we compute each time. This way, if a LR that is requested already exists in the matrix, we do not need to compute it again. The matrix is of size $N \times (\log_2 N + 1)$, implying that the decoding complexity is $\mathcal{O}(N \log_2 N)$.

The recursive formulas for transition probabilities are defined as

$$W_{2N}^{(2i-1)}(y_1^{2N}, u_1^{2i-2}|u_{2i-1}) = \sum_{u_{2i}} \frac{1}{2} W_N^{(i)}(y_1^N, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}|u_{2i-1} \oplus u_{2i}) W_N^{(i)}(y_1^N, u_{1,e}^{2i-2}|u_{2i}), \tag{28}$$

$$W_{2N}^{(2i)}(y_1^{2N}, u_1^{2i-1}|u_{2i}) = \frac{1}{2} W_N^{(i)}(y_1^N, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}|u_{2i-1} \oplus u_{2i}) W_N^{(i)}(y_1^N, u_{1,e}^{2i-2}|u_{2i}). \tag{29}$$

For the proof of (28) and (29), consult the Appendix.

The SC Decoder generates its decision for $\hat{u}_i$ as

$$\hat{u}_i = \begin{cases} 0, & \text{if } u_i \text{ is a frozen bit} \\ \text{argmax}_{x \in \{0,1\}} W_N^{(i)}(y_1^N, u_1^{i-1}|x), & \text{otherwise.} \end{cases}$$

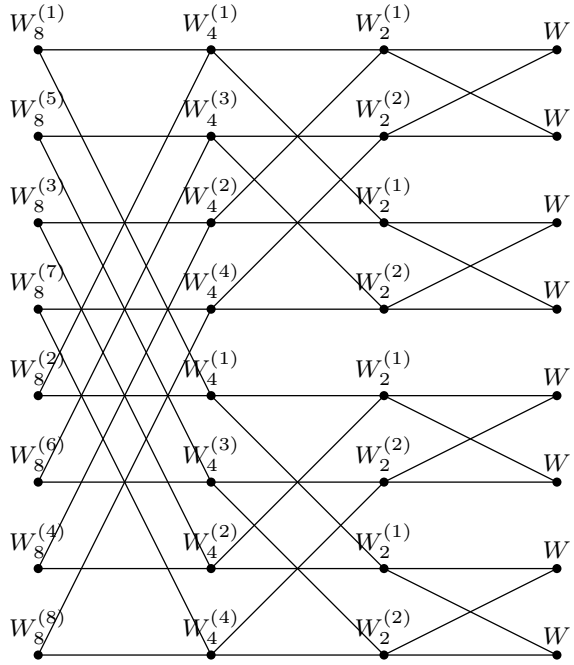


Figure 13: The channel transformation process, $N = 8$.

In Fig. 13, we observe that every transition probability after the first stage of the recursion is used over one time. For instance, the value of $W_4^{(1)}$ is needed for the calculation of both $W_8^{(1)}$ and $W_8^{(2)}$. If we exploit this effect, we can save in terms of decoding complexity. Therefore, we create two matrices: one for the probabilities $W_N^{(i)}(y_1^N, u_1^{i-1}|0)$ and one for the $W_N^{(i)}(y_1^N, u_1^{i-1}|1)$, each of size $N \times (\log_2 N + 1)$. This results in a decoding complexity of $\mathcal{O}(N \log_2 N)$.

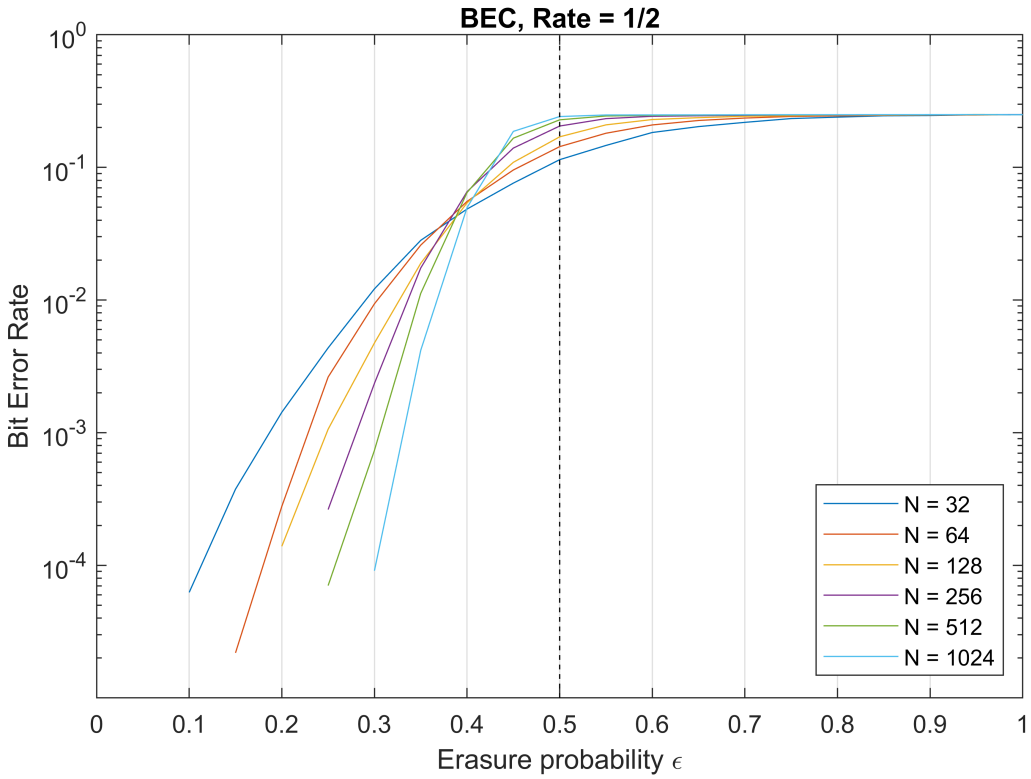


Figure 14: BER for different block lengths over BEC.

2.8 Performance on BEC

In Fig. 14, we illustrate the behavior of a binary erasure channel with code rate $\mathrm{R} = \frac{1}{2}$. As we grow larger the block N, the bit error rate improves causing polar codes to approach Shannon capacity as N increases towards infinity.

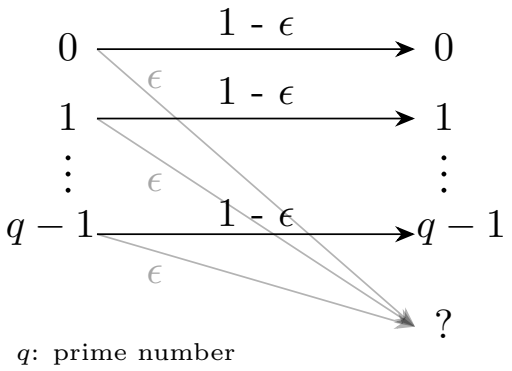


Figure 15: q-ary input Erasure Channel.

3 Polarization for Prime-input Erasure Channels

The polar scheme considered in Section 2 can be generalized to arbitrary q-input discrete memoryless channels, where q is a prime number [3]. In this work, we examine the prime input erasure channels.

3.1 Symmetric Capacity

Given a q-ary input discrete memoryless channel(DMC) $W : \mathcal{X} \to \mathcal{Y}$, with input alphabet $\mathcal{X} = \{0, 1, ..., q-1\}$ we define the symmetric capacity as

$$I(W) \triangleq \sum_{y\in\mathcal{Y}} \sum_{x\in\mathcal{X}} \frac{1}{q} W(y|x) \log_q \frac{W(y|x)}{\sum_{x'\in\mathcal{X}} \frac{1}{q} W(y|x')}. \tag{30}$$

We use q as the base of the logarithm so that $0 \leq I(W) \leq 1$.

3.2 q-ary input Erasure Channel

The q-ary input EC, shown in Fig. 15, takes a q-ary input $\mathcal{X} = \{0, 1, ..., q-1\}$ and produces the output $\mathcal{Y} = \{0, 1, ..., q-1, ?\}$. The input is transmitted either correctly with probability $1 - \epsilon$ or it is erased with probability ϵ. The symmetric capacity of the q-ary input EC is

$$I(W) = 1 - \epsilon. \tag{31}$$

For the proof of (31), consult the Appendix.
We notice that the symmetric capacity is the same for the BEC and the q-ary input EC. Moreover, the erasure probability ϵ is independent of the cardinality of the channel for the conventional transformation shown in Fig. 3. Hence, we conclude that the recursive formulas (22) and (23) can also be applied to the q-ary EC.

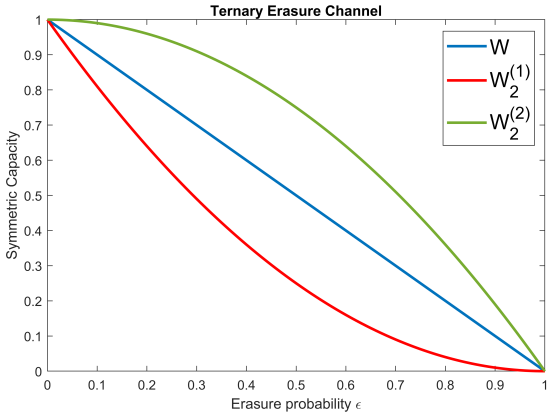


Figure 16: Symmetric capacity for TEC before and after the basic polarization step.

3.3 Arikan's approach

Polarization in q-ary input EC, as shown for a TEC in Fig. 16, occurs in exactly the same way as we mentioned earlier for BEC. We only need to modify the operations in the encoder and the decoder to be over GF(q). The recursive formulas for transition probabilities are defined as

$$W_{2N}^{(2i-1)}(y_1^{2N}, u_1^{2i-2}|u_{2i-1}) = \sum_{u_{2i}} \frac{1}{q} W_N^{(i)}(y_1^N, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}|u_{2i-1} \oplus u_{2i}) W_N^{(i)}(y_1^N, u_{1,e}^{2i-2}|u_{2i}), \tag{32}$$

$$W_{2N}^{(2i)}(y_1^{2N}, u_1^{2i-1}|u_{2i}) = \frac{1}{q} W_N^{(i)}(y_1^N, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}|u_{2i-1} \oplus u_{2i}) W_N^{(i)}(y_1^N, u_{1,e}^{2i-2}|u_{2i}). \tag{33}$$

The SC Decoder generates its decision for $\hat{u}_i$ as follows

$$\hat{u}_i = \begin{cases} 0, & \text{if } u_i \text{ is a frozen symbol} \\ \text{argmax}_{x \in \{0,1,\ldots,q-1\}} W_N^{(i)}(y_1^N, u_1^{i-1}|x), & \text{otherwise.} \end{cases}$$

Likelihood ratio is based on the binary alphabet and does not generalize for larger alphabets.

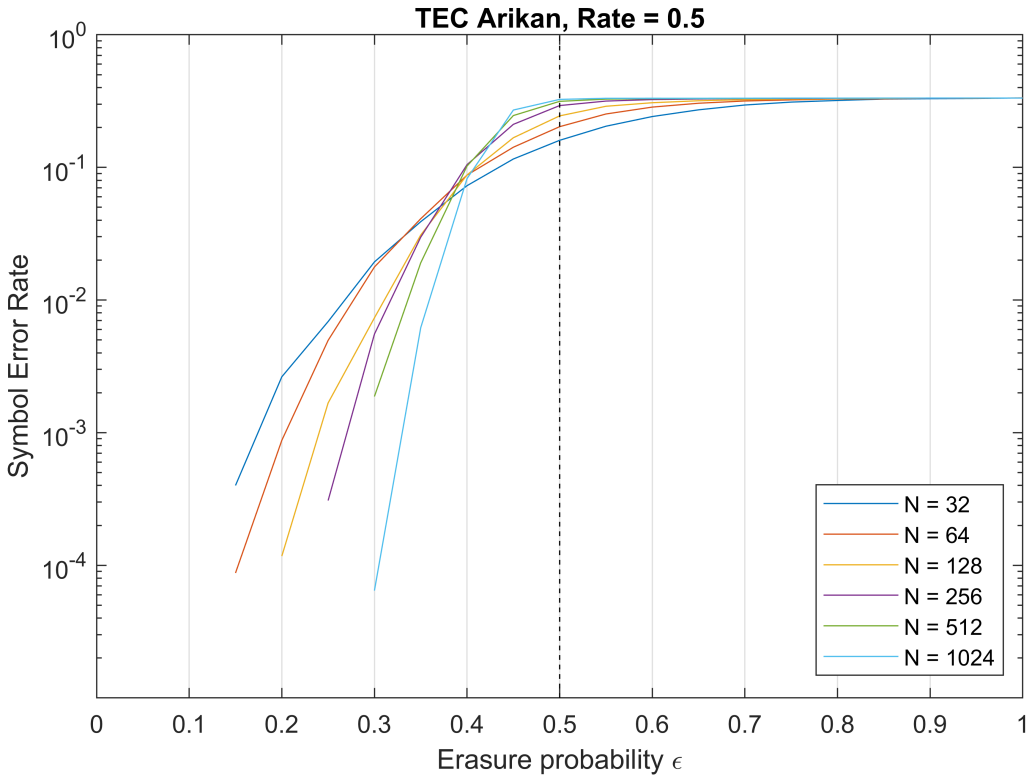


Figure 17: SER for different block lengths over TEC of the original scheme.

3.3.1 Performance on TEC

In Fig. 17, we illustrate the behavior of a ternary erasure channel, $\mathcal{X} = \{0, 1, 2\}$ with code rate $\mathrm{R} = \frac{1}{2}$. We observe the improvement of the symbol error rate, as we increase the block length N.

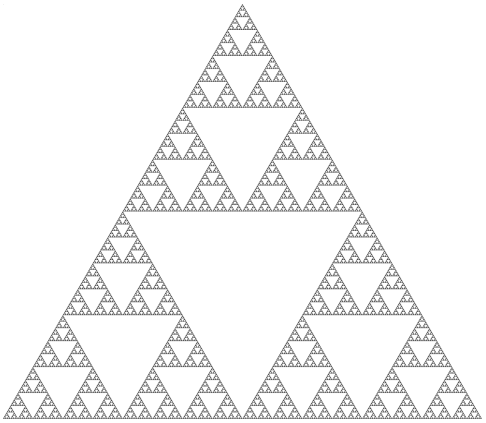


Figure 18: The Sierpinski triangle.

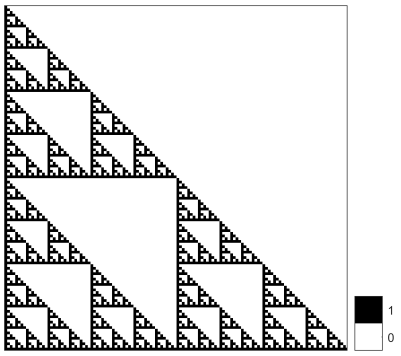


Figure 19: $G_{128}/[\text{Pascal matrix}]_{128} \bmod 2$.

$$\begin{bmatrix} 1 & 5 & 15 & 35 & 70 \\ 1 & 4 & 10 & 20 & 35 \\ 1 & 3 & 6 & 10 & 15 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Figure 20: The 5×5 Pascal matrix.

3.4 Pascal-matrix approach

In the previous sections, we presented the polarization effect using the original construction method of Polar codes. It is proved that for all prime q, an invertible $\mathbb{F}_q$ matrix is polarizing if and only if its not upper-triangular [4]. Consequently, the number of matrices that polarize is vast and we may find matrices that polarize better in terms of error probability. Nevertheless, we will focus in matrices that reproduces themselves since the recursive reproduction of generator matrix we presented earlier gives polar codes good performance under low complexity.

If we depict the generator matrix G for $N = 128$ as defined earlier, we observe that we get the pattern seen in Fig. 19. This never-ending pattern is a fractal (meaning that it reproduces itself on re-scaling), also known as the Sierpinski triangle illustrated in Fig. 18. This fractal can be produced by using the modulus operation on the Pascal matrix [11]. If we apply modulus 2 operation on the Pascal matrix, we obtain exactly the same pattern as of G_{128}, as shown in Fig. 19. The Pascal matrix is an infinite matrix containing the binomial coefficients as its elements and it is an alternative formulation of the Pascal's triangle. The expression for the generator matrix G_N is

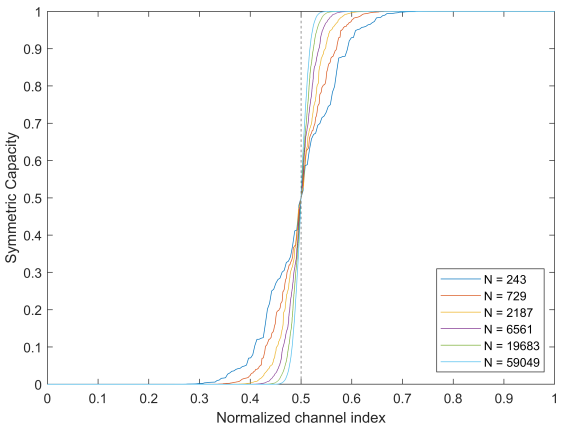
$$G_N = [\text{Pascal matrix}]_N \mod q \quad (\text{for prime } q). \tag{34}$$

The above method is not unique [5]. We can construct any fractal that originates from the Pascal matrix using the Kronecker power on the kernel matrix F which is defined as

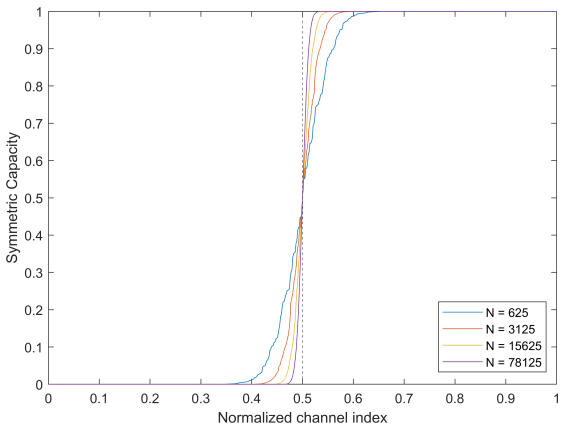
$$F = [\text{Pascal matrix}]_q \mod q \quad (\text{for prime } q). \tag{35}$$

By using (35), we can calculate recursively the generator matrix G_N as

$$G_N = F^{\otimes n} \text{ over } \mathrm{GF}(q). \tag{36}$$



(a) TEC(0.5) for block length N = $3^5 = 243$, 3^6, ..., $3^{10} = 59049$.



(b) 5-input EC(0.5) for block length N = $5^4 = 625$, 5^5, 5^6, $5^7 = 78125$.

Figure 21: Symmetric capacity of the proposed Pascal-matrix code for different block lengths over a TEC(0.5) and a 5-input EC, with $\epsilon = 0.5$.

3.4.1 Pascal-matrix based construction

In [11], the Pascal-matrix construction was presented for ternary input erasure channels. We extend this construction method for prime input erasure channels. We begin the construction of the code by choosing the size q of the channel's input alphabet and by applying (35), we obtain the kernel matrix F of our scheme. This time, we will need q recursive formulas in order to compute the symmetric capacities of the synthesized channels. The following algorithm describes the procedure for computing the capacities of the fabricated channels regardless of the channel's input size. To get an intuitive understanding of the recursive formulas, a detailed analysis for the case of a ternary erasure channel, $q = 3$, is given in the Appendix.

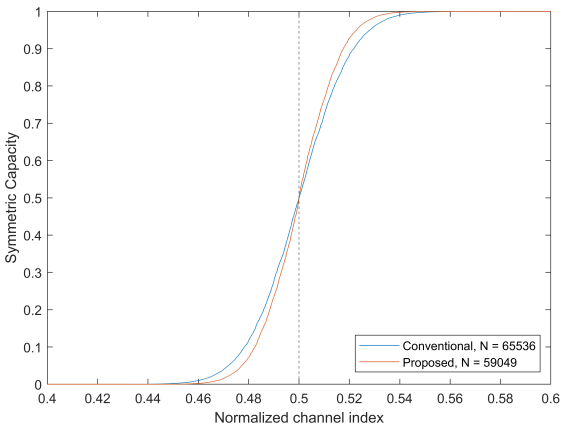
Algorithm 1 Calculation of $I(W)$ of the fabricated channels

Input: Size q of the channel's input alphabet , the result *pos* of the channel index mod q and the previous estimated capacity b.

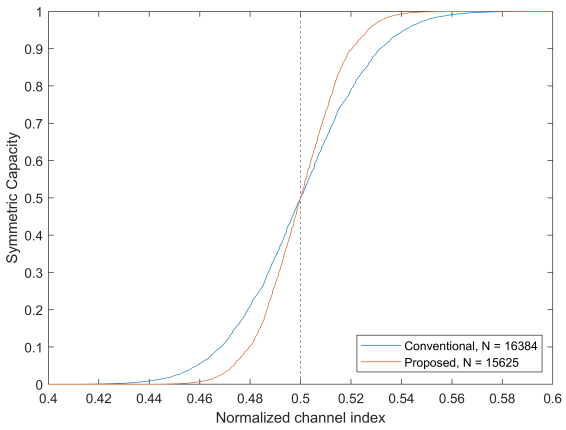
Output: Symmetric capacity of the ith channel.

1: $ep \leftarrow 0$
2: **for** $i \leftarrow$ pos + 1 **to** q + 1 **do**
3: $\quad coef \leftarrow$ binomial_coefficient$(q, i-1)$
4: $\quad ep \leftarrow ep + coef \cdot (1-b)^{i-1} \cdot (b)^{q-i+1}$
5: **end for**
6: $cap \leftarrow 1 - ep$
7: **return** *cap*

In Fig. 21, we observe that as we increase the block length N through powers of 3 and 5, respectively, the amount of mediocre channels in terms of symmetric capacity shrinks. In the Fig. 22, we compare the behavior of symmetric capacities between the original implementation of polar codes and the proposed implementation based on the Pascal-matrix, for two cases, a transmission over a TEC and a transmission over a 5-input EC. We observe the dominance of the proposed Pascal-matrix scheme compared to the original scheme, even when an advantage of longer block length N is given to the original scheme.



(a) Conventional polar code with block length $N = 2^{16} = 65536$ vs the proposed Pascal-matrix polar code with block length $N = 3^{10} = 59049$ over a TEC(0.5).



(b) Conventional polar code with block length $N = 2^{14} = 16384$ vs the proposed Pascal-matrix polar code with block length $N = 5^{6} = 15625$ over a 5-input EC(0.5).

Figure 22: Symmetric capacity comparison between the conventional polar code and the proposed polar code base on the Pascal-matrix for two different input sizes.

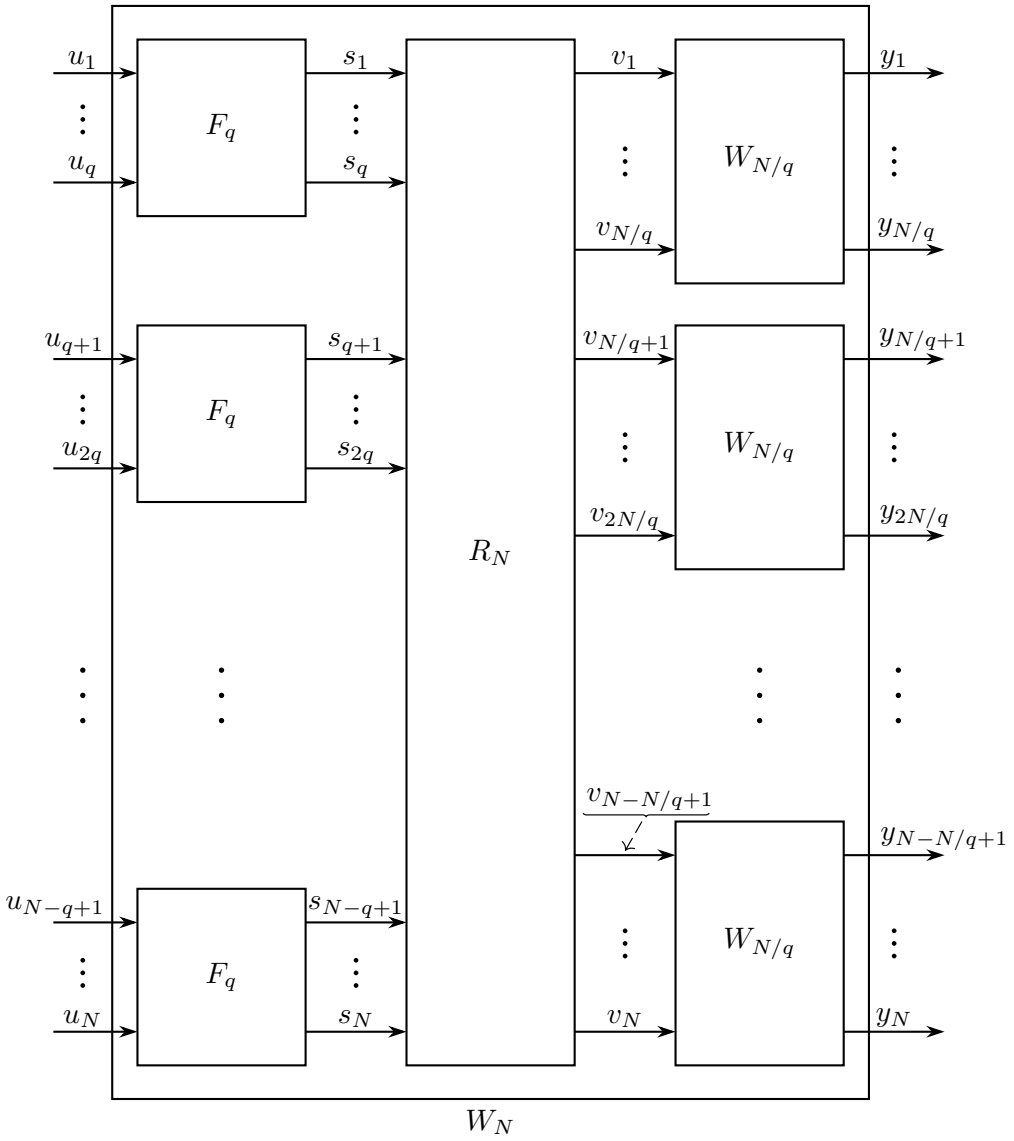
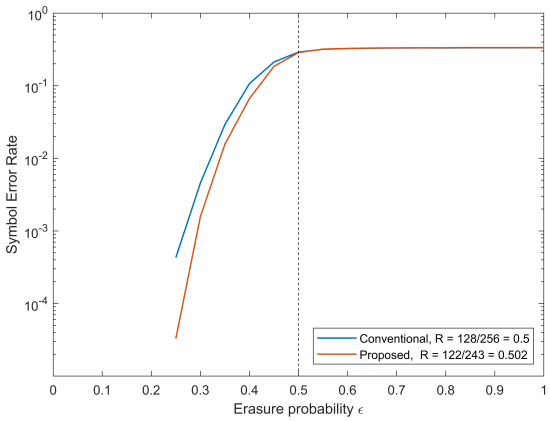


Figure 23: Recursive construction of W_N from q copies of $W_{N/q}$.

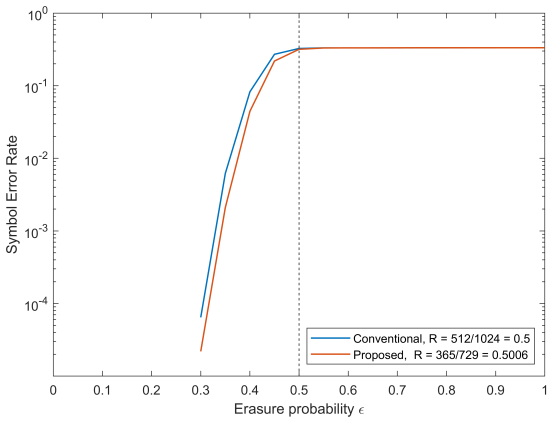
3.4.2 Encoding

As we illustrate in Fig. 23, we combine q copies of $W_{N/q}$ synthesized channels to construct W_N. The permutation R_N sorts its input s_1^N based on the signal's index mod-q result. The first $W_{N/q}$ channel corresponds to result 1 and we continue to assign the signals to each $W_{N/q}$ channel until we reach to result 0. If we assume that the complexity of a scalar mod-q addition is 1 unit, the complexity of the reverse shuffle operation R_N is N units of time, the encoding complexity with the help of Master Theorem is

$$\begin{aligned} &\mathcal{T}(N) = N + \mathcal{O}(N) + q\mathcal{T}(\frac{N}{q}) \Rightarrow \\ &\mathcal{T}(N) = \mathcal{O}(N \log_q N). \end{aligned} \tag{37}$$



(a) Symbol error rate of conventional polar code with block length $N = 2^8 = 256$ and rate $R = 0.5$ and the proposed Pascal-matrix polar code with block length $N = 3^5 = 243$ and rate $R = 0.502$.



(b) Symbol error rate of conventional polar code with block length $N = 2^{10} = 1024$ and rate $R = 0.5$ and the proposed Pascal-matrix polar code with block length $N = 3^6 = 729$ and rate $R = 0.5006$.

Figure 24: Symbol error rate as a function of erasure probability ϵ for transmission over a TEC.

3.4.3 Decoding

Since the kernel matrix F we are using has always $q \times q$ dimensions, we define q transition probabilities which we derive from (38). The transition probabilities for the case of a ternary erasure channel, $q = 3$, are given in the Appendix.

$$W_N^{(i)}(y_1^N, u_1^{i-1}|u_i) = \sum_{u_{i+1}^N} \frac{1}{q^{N-1}} W_N(y_1^N|u_1^N). \tag{38}$$

The SC Decoder generates its decision for $\hat{u}_i$ as follows

$$\hat{u}_i = \begin{cases} 0, & \text{if } u_i \text{ is a frozen symbol} \\ \text{argmax}_{x\in\{0,1,\ldots,q-1\}} W_N^{(i)}(y_1^N, u_1^{i-1}|x), & \text{otherwise.} \end{cases}$$

Since we use q matrices of size $N \times (\log_q N + 1)$, the decoding complexity is $\mathcal{O}(N \log_q N)$.

3.4.4 Performance on q-ary EC

In the following figures, we examine the performance of the proposed Pascal-matrix polar codes versus the original polar codes for transmission over a ternary erasure channel. The block lengths N cannot be equal due to the distinct construction of each code. Nevertheless, we always give an advantage to the conventional scheme. In Fig. 24(a), we plot the Symbol Error Rate of both codes as a function of the channel's erasure probability ϵ. Despite the fact that we give shorter block length N and slightly higher rate in the proposed code, it performs better than the original one. In Fig. 24(b), we give considerably shorter block length to the proposed code and its superiority remains.

In Fig. 25, we examine the performance over a 5-input erasure channel. We verify that the superior performance of the proposed Pascal-matrix polar code remains over the original one despite the disadvantage in terms of block length N and rate R.

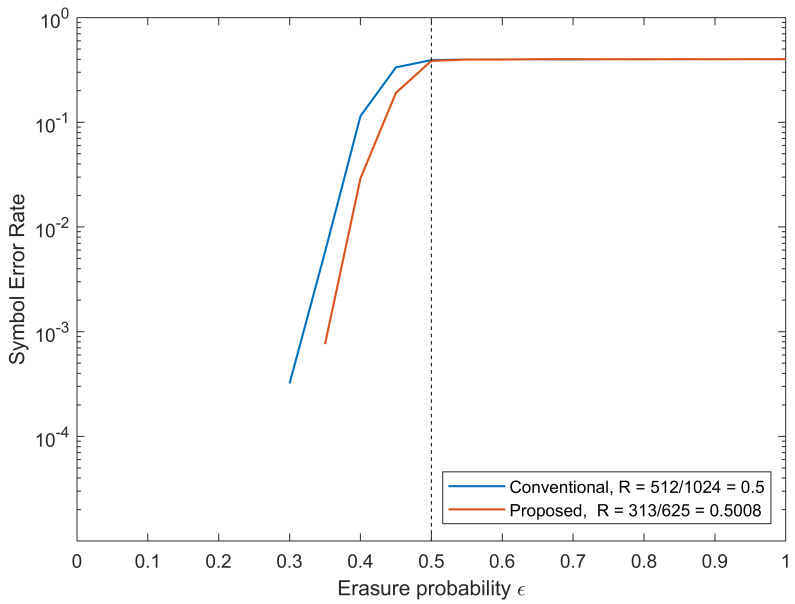


Figure 25: Symbol error rate as a function of erasure probability ϵ for transmission over a 5-input erasure channel.

4 Polar Codes for Secrecy

In this section, we will introduce Wyner's approach to achieving secure communication over a transmission in the presence of a wire-tapper. We will use the polar coding techniques on q-ary input erasure channels, both the original introduced by Arikan and the proposed one based on the Pascal-matrix, for the encoding and decoding procedure and present the performance results.

4.1 The Wiretap Communication Model

In the wiretap communication model, Alice wishes to send messages to Bob through a communication channel C_1, called the main channel, but her transmissions also reach an adversary Eve through another channel C_2, called the wiretap channel. This is illustrated in Fig. 26 , wherein u denotes the k-bit length message that Alice wishes to communicate to Bob. We think of u as a data sequence which consists of independent copies of the random variable $\boldsymbol{U} = \{0, 1, ..., q-1\}$, where $\mathbf{U}$ is equally distributed. The encoder maps u into a sequence x of n channel symbols. This sequence is transmitted across the main channel and the wiretap channel resulting in the corresponding channel outputs y and z, respectively. Finally the decoder maps y into an estimate $\hat{u}$ of the original message. The goal is to design a coding scheme-namely, an encoding algorithm and an decoding algorithm-that makes it possible to communicate both *reliably* and *securely*, as the message length k becomes large. Reliability is measured in terms of the *probability of error* in recovering the message. Specifically the objective is to satisfy the following:

$$\textbf{Reliability Condition: } \lim_{k\to\infty} \Pr\{\hat{u}^k \neq u^k\} = 0. \tag{39}$$

where the probability is over all the relevant coin tosses in the system: in the generation of u^k, in the encoder, and in the main channel. Security is usually measured in terms of the *normalized mutual information* between the message u^k and Eve's observations z.

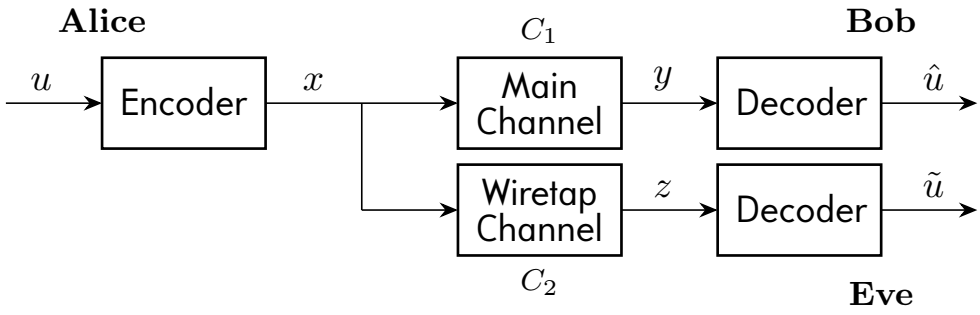


Figure 26: The wiretap channel.

Specifically, one is interested in encoding algorithms that satisfy the following:

$$\textbf{Security Condition: } \lim_{k\to\infty} \frac{I(u^k; z)}{k} = 0. \tag{40}$$

Note that $I(u^k; z)$ is equal to the difference between the *a priori* entropy $H(u^k)$ and the conditional entropy $H(u^k|z)$. Thus, intuitively, (40) means that observing z does not provide any information about u^k beyond what is available *a priori* as compared to the message length k.

Wyner proved that if both C_1 and C_2 are discrete memoryless channels and moreover, C_2 is degraded with respect to C_1 then the system is characterized by a single constant C_s, called the secrecy capacity, which has the following meaning. For all $\epsilon > 0$, there exist coding schemes of information rate $R \geq C_s - \epsilon$ that satisfy (39) and (40); conversely, it is not possible to satisfy both (39) and (40) at rates greater than C_s.

4.2 The Secrecy Capacity of the Wiretap Channel

A simple expression for secrecy capacity C_s was given by Leung-Yan-Cheong in [7]. If C_1 and C_2 are symmetric and C_2 is degraded with respect to C_1 then

$$C_s = C_1 - C_2. \tag{41}$$

The q-ary input erasure channel we discussed earlier is symmetric. Therefore, if the main channel is an erasure channel with erasure probability ϵ_1 and the wiretap channel is an erasure channel with erasure probability ϵ_2, with $\epsilon_2 \geq \epsilon_1$, then the secrecy capacity $C_s = \epsilon_2 - \epsilon_1$.

4.3 Encoding and Decoding

We define the main channel $C_1 = W^*(y|x)$ and the wiretap channel $C_2 = W(z|x)$, shown in Fig. 26. As we observe in Fig. 27, our secret message is transmitted over those bit-channels $W_N^{(i)}$ that are bad for the eavesdropper-Eve, while flooding the bit-channels that are good for Eve with random bits.

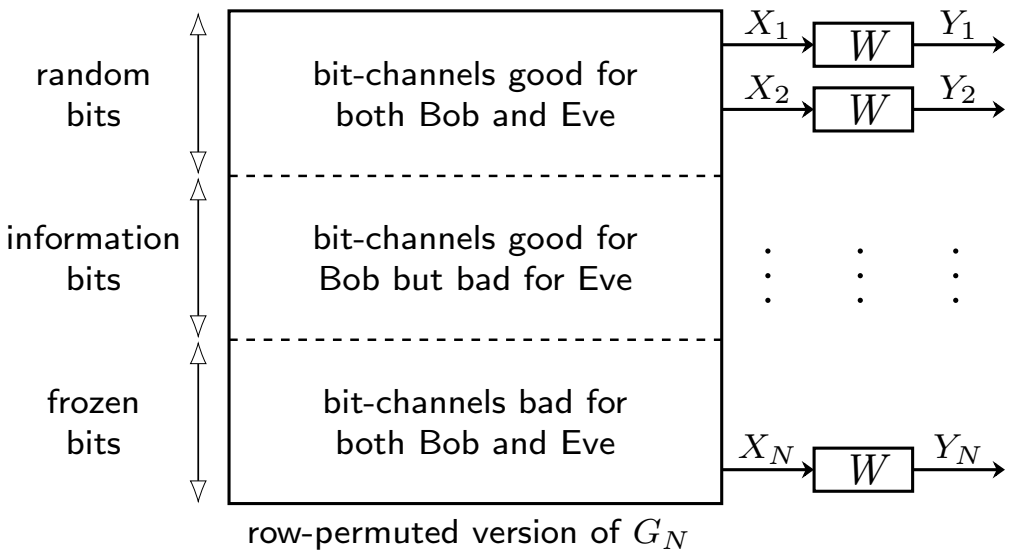


Figure 27: Secrecy structure.

We define three subsets of the block length N as follows:

$$\mathcal{R} \triangleq A(W) \tag{42}$$

$$\mathcal{S} \triangleq A(W^*) \backslash A(W) \tag{43}$$

$$\mathcal{B} \triangleq A^c(W^*) \tag{44}$$

where $A(W)$ is the information set of the good channel indexes for the wiretap channel, $A(W^*)$ is the information set of the good channel indexes for the main channel, A^c is the complement of A over N. Notice that the sets $\mathcal{R}$, $\mathcal{S}$, $\mathcal{B}$ are disjoint and $\mathcal{R} \cup \mathcal{S} \cup \mathcal{B} = [N]$. Let $|\mathcal{R}| = r$ and $|\mathcal{S}| = k$.
The code rate is $R = \frac{r+k}{N}$ and the secrecy rate is $R_s = \frac{k}{N}$.

The encoder is a function $\mathcal{E} : \{0, 1, ..., q-1\}^k \times \{0, 1, ..., q-1\}^r \to \{0, 1, ..., q-1\}^N$. It accepts as input a message $\boldsymbol{u} \in \{0, 1, ..., q-1\}^k$ and a vector $\mathbf{e} \in \{0, 1, ..., q-1\}^r$. It has been proved that $\mathbf{e}$ must be selected by Alice uniformly at random from $\{0, 1, ..., q-1\}^r$ in order to help keep Eve ignorant [8]. The encoder first constructs the vector $\boldsymbol{v} \in \{0, 1, ..., q-1\}^N$, by setting $\boldsymbol{v}_{\mathcal{R}} = \mathbf{e}$, $\boldsymbol{v}_{\mathcal{S}} = \boldsymbol{u}$, and $\boldsymbol{v}_{\mathcal{B}} = \mathbf{0}$. The encoder then outputs $\mathcal{E}(\mathbf{u}, \mathbf{e}) := \boldsymbol{v}\, G_n$ as we described in Section 2. The decoder invokes successive cancellation decoding to produce the vector $\hat{u}^k$ for the main channel and $\tilde{u}^k$ for the wiretap channel. Earlier we defined the secrecy capacity for the wiretap channel, as $C_s = C_1 - C_2$, (41). Let $R_s = k/n$ denote the code rate of our scheme, then following theorem holds

$$\lim_{N \to \infty} R_s = C(W^*) - C(W). \tag{45}$$

For the proof of (45), consult the Appendix.

4.4 Performance Results

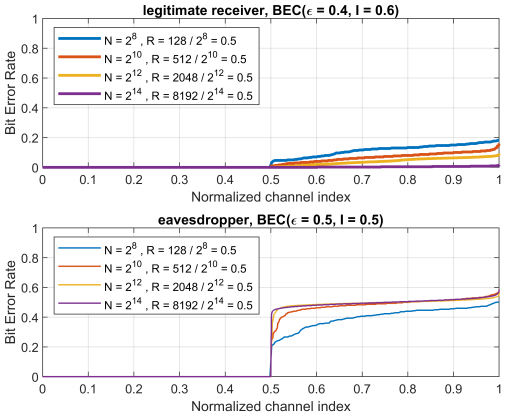


Figure 28: BER vs Normalized channel index for different block lengths over binary erasure wiretap channel.

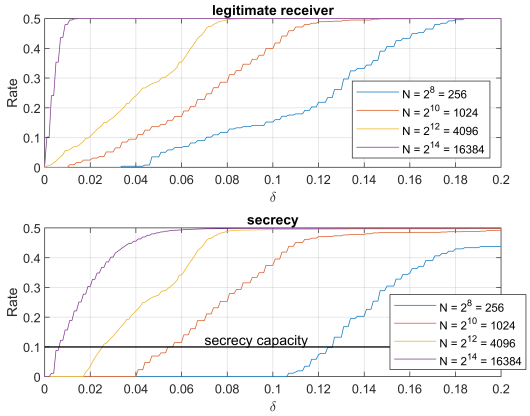


Figure 29: Rate of the coding scheme as a function of the tolerance of the BER $\delta \in [0, 0.2]$.

In Fig. 28, we illustrate the performance of the original polar codes for different block lengths N for the wiretap channel of Fig. 26. The main channel is a BEC with erasure probability $\epsilon = 0.4$, and the wiretap channel is a BEC with erasure probability $\epsilon = 0.5$. In order to examine the performance for different block lengths N, the horizontal axis is marked with all channel indexes, normalized to the closed interval $[0, 1]$. The vertical axis is marked with the bit error rate. The code rate is $R = \frac{1}{2}$, resulting in the half of the channels having BER $= 0$. We observe that on the legitimate receiver's side, we get better performance as we increase the block length N in terms of bit error rate. On the eavesdropper's side, as the block length N increases, the bit error rate approaches 0.5, making the channel completely unreliable and more difficult for the eavesdropper to decode our secret message.

In the first subplot of Fig. 29, we illustrate the rate in the legitimate receiver as a function of the tolerance of the bit error rate δ, $\delta \in [0, 0.2]$. The tolerance of the BER is the acceptable range within which the BER can vary. If $\delta = 0.2$, we take into consideration all the channels that have BER up to 0.2 in the legitimate receiver. We observe that as we increase δ, the rate increases since more channels are counted as good channels in the estimation. We observe that larger block lengths N achieve higher rates more quickly and with smaller tolerance values δ.

In the second subplot of Fig. 29, we illustrate the secrecy rate as a function of the tolerance of the bit error rate δ, $\delta \in [0, 0.2]$. To estimate the secrecy rate we take into consideration the channels that have BER smaller than the tolerance δ in the legitimate receiver's side and at the same time their BER falls within the range $0.5 - \delta <$ BER $< 0.5 + \delta$ in the eavesdropper side (for BER$= 0.5$ the security condition holds). We observe that larger block lengths N achieve higher rates more quickly and with smaller tolerance values δ. As a reference, we highlight the secrecy capacity $C_s = 0.6 - 0.5 = 0.1$.

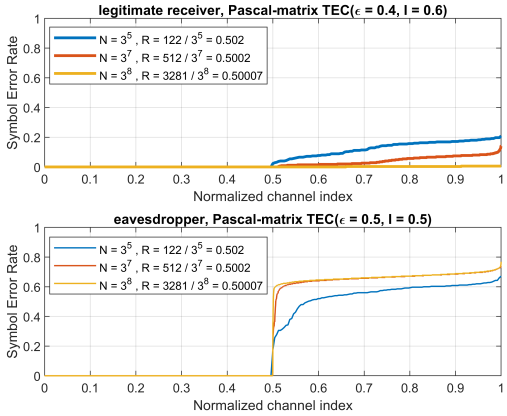


Figure 30: SER vs Normalized channel index for different block lengths over Pascal-matrix ternary erasure wiretap channel.

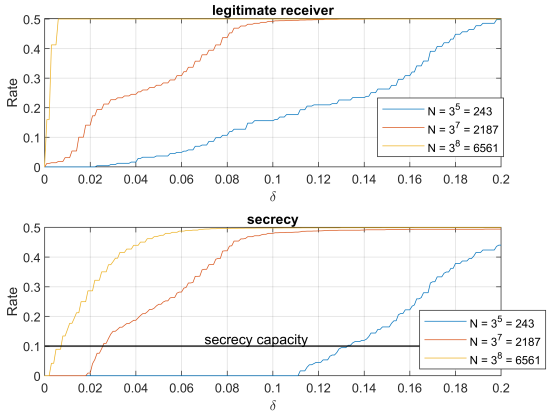


Figure 31: Rate of the coding scheme as a function of the tolerance of the SER $\delta \in [0, 0.2]$.

In Fig. 30, we illustrate the performance of the polar codes based on the Pascal-matrix for different block lengths N for the wiretap channel of Fig. 26. The main channel is a Pascal-matrix TEC with erasure probability $\epsilon = 0.4$, and the wiretap channel is a Pascal-matrix TEC with erasure probability $\epsilon = 0.5$. In order to examine the performance for different block lengths N, the horizontal axis is marked with all channel indexes, normalized to the closed interval $[0, 1]$. The vertical axis is marked with the symbol error rate. The code rate is $R = \frac{1}{2}$, resulting in the half of the channels having SER $= 0$. We observe that on the legitimate receiver's side, we get better performance as we increase the block length N in terms of symbol error rate. On the eavesdropper's side, as the block length N increases, the symbol error rate approaches 0.7, making the channel completely unreliable and more difficult for the eavesdropper to decode our secret message.

In the first subplot of Fig. 31, we illustrate the rate in the legitimate receiver as a function of the tolerance of the SER δ, $\delta \in [0, 0.2]$. The tolerance of the SER is the acceptable range within which the SER can vary. If $\delta = 0.2$, we take into consideration all the channels that have SER up to 0.2 in the legitimate receiver. We observe that as we increase δ, the rate increases since more channels are counted as good channels in the estimation. We observe that larger block lengths N achieve higher rates more quickly and with smaller tolerance values δ.

In the second subplot of Fig. 31, we illustrate the secrecy rate as a function of the tolerance of the SER δ, $\delta \in [0, 0.2]$. To estimate the secrecy rate we take into consideration the channels that have SER smaller than the tolerance δ in the legitimate receiver's side and at the same time their SER falls within the range $0.67 - \delta <$ SER $< 0.67 + \delta$ in the eavesdropper side (for SER$= 0.67$ the security condition holds). We observe that larger block lengths N achieve higher rates more quickly and with smaller tolerance values δ. As a reference, we highlight the secrecy capacity $C_s = 0.6 - 0.5 = 0.1$.

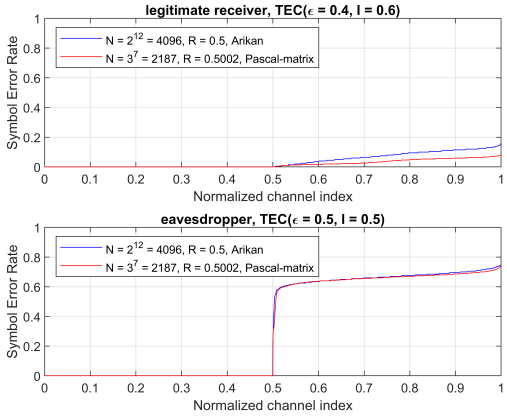


Figure 32: SER of ternary erasure channel using the original scheme vs using the proposed pascal-matrix scheme.

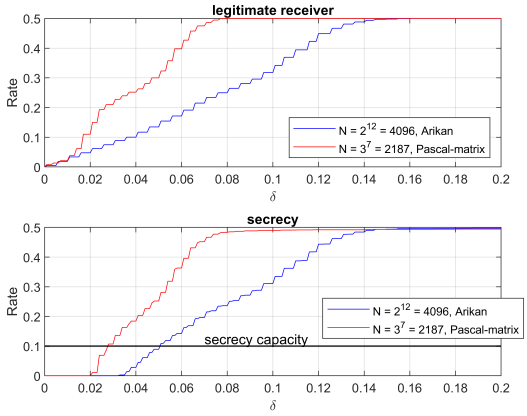


Figure 33: Rate of the coding scheme as a function of the tolerance of the SER $\delta \in [0, 0.2]$.

In Fig. 32, we illustrate the performance of the original polar codes versus the proposed implementation based on the Pascal-matrix for the wiretap channel of Fig. 26. The main channel is a TEC with erasure probability $\epsilon = 0.4$, and the wiretap channel is a TEC with erasure probability $\epsilon = 0.5$. Since the two implementations have different block lengths N, the horizontal axis is marked with all channel indexes, normalized to the closed interval $[0, 1]$. The vertical axis is marked with the symbol error rate. The code rate is $R = \frac{1}{2}$, resulting in the half of the channels having SER $= 0$. We observe that in the legitimate receiver's side, the Pascal-matrix implementation performs better than the original one in terms of symbol error rate. In the eavesdropper's side, the symbol error rate of both implementations has the same behaviour and SER is approximately around 0.7 for all channels. In order to satisfy the security condition (40), the mutual information between the transmitted secret message and the observation of the eavesdropper needs to be zero. We have to examine the channel formed between the secret message and the estimation of the eavesdropper so as to conclude to the mutual information formula. As we observe in the transition probabilities matrices of Fig. 35 to Fig. 37, between the transmitted secret message u and the estimation $\hat{u}$ of the eavesdropper it is formed a ternary channel which is symmetric (each row is a permutation of each other row and each column is a permutation of each other column), illustrated in Fig. 34. The symmetric capacity of a ternary symmetric channel with error probability p is

$$I(u; \tilde{u}) = 1 + p \log_3(\frac{p}{2}) + (1 - p) \log_3(1 - p). \tag{46}$$

For the proof of (46), consult the Appendix.
For $p = \frac{2}{3} = 0.67$, (46) equals 0, the proof is provided in the Appendix.
Since SER is around 0.67 for the most channels, our estimation is verified, meaning that the eavesdropper is being successfully kept ignorant about the secret message.

In the first subplot of Fig. 33, we illustrate the rate in the legitimate receiver as a function of the tolerance of the SER δ, $\delta \in [0, 0.2]$. The tolerance of the SER is the acceptable range within which the SER can vary. If $\delta = 0.2$, we take into consideration all the channels that have SER up to 0.2 in the legitimate receiver. We observe that as we increase δ, the rate increases since more channels are counted as good channels in the estimation. The Pascal-matrix implementation performs better than the original's scheme

since for smaller tolerance values the former achieves greater rate.

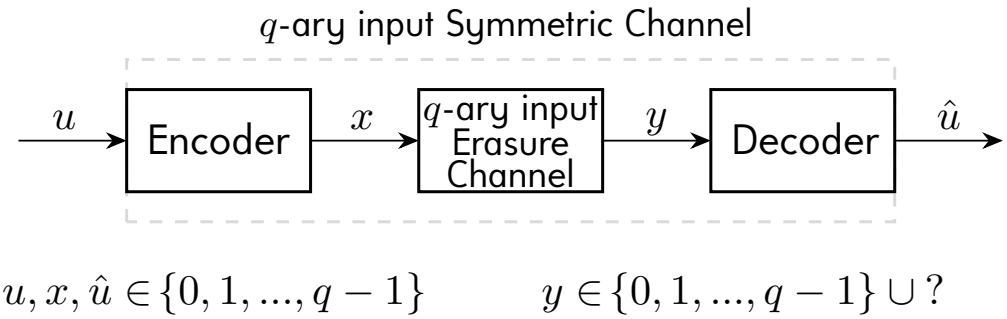


Figure 34: The symmetric channel formed between the original message u and the estimated message $\hat{u}$.

y \ x	**0**	**1**	**2**
0	0.9230	0.0388	0.0390
1	0.0385	0.9224	0.0382
2	0.0385	0.0388	0.9229

Figure 35: Transition probability matrix for the Ternary Symmetric Channel (when the underlying ternary erasure channel has $\epsilon = 0.1$, I = 0.9, R = 0.85).

y \ x	**0**	**1**	**2**
0	0.5499	0.2279	0.2245
1	0.2244	0.5465	0.2240
2	0.2257	0.2256	0.5515

Figure 36: Transition probability matrix for the Ternary Symmetric Channel (when the underlying ternary erasure channel has $\epsilon = 0.4$, I = 0.6, R = 0.55).

y \ x	**0**	**1**	**2**
0	0.8675	0.0659	0.0651
1	0.0660	0.8699	0.0648
2	0.0665	0.0642	0.8701

Figure 37: Transition probability matrix for the Ternary Symmetric Channel (when the underlying ternary erasure channel has $\epsilon = 0.8$, I = 0.2, R = 0.15).

In the second subplot of Fig. 33, we illustrate the secrecy rate as a function of the tolerance of the SER δ, $\delta \in [0, 0.2]$. To estimate the secrecy rate we take into consideration the channels that have SER smaller than the tolerance δ in the legitimate receiver's side and at the same time their SER falls within the range $0.67 - \delta < \text{SER} < 0.67 + \delta$ in the eavesdropper side (for SER = 0.67 the security condition holds). We observe again better performance for the Pascal-matrix implementation as we increase the tolerance δ. As a reference, we highlight the secrecy capacity $C_s = 0.6 - 0.5 = 0.1$.

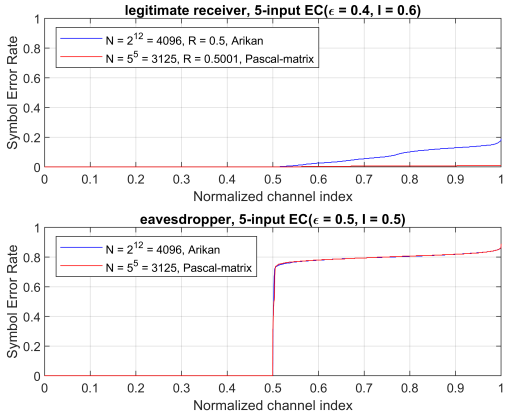


Figure 38: SER of 5-input erasure channel using the original scheme vs using the proposed pascal-matrix scheme.

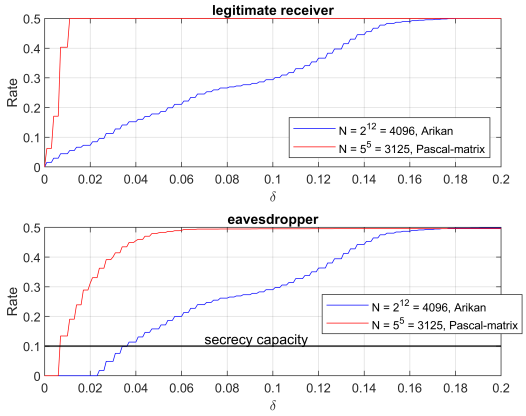


Figure 39: Rate of the coding scheme as a function of the tolerance of the SER $\delta \in [0, 0.2]$.

In Fig. 38, we illustrate the performance of the original polar codes versus the proposed implementation based on the Pascal-matrix for the wiretap channel of Fig. 26. The main channel is a 5-input EC with erasure probability $\epsilon = 0.4$, and the wiretap channel is a 5-input EC with erasure probability $\epsilon = 0.5$. Since the two implementations have different block lengths N, the horizontal axis is marked with all channel indexes, normalized to the closed interval $[0, 1]$. The vertical axis is marked with the symbol error rate. The code rate is $R = \frac{1}{2}$, resulting in the half of the channels having SER $= 0$. We observe that in the legitimate receiver's side, the Pascal-matrix implementation performs better than the original one in terms of symbol error rate. In the eavesdropper's side, the symbol error rate of both implementations has the same behaviour and SER is approximately around 0.8 for all channels.
In order to satisfy the security condition (40), the mutual information between the transmitted secret message and the observation of the eavesdropper needs to be zero. We have to examine the channel formed between the secret message and the estimation of the eavesdropper so as to conclude to the mutual information formula. As we observe in the transition probabilities matrices of Fig. 40 to Fig. 42, between the transmitted secret message u and the estimation $\hat{u}$ of the eavesdropper it is formed a 5-input channel which is symmetric (each row is a permutation of each other row and each column is a permutation of each other column), illustrated in Fig. 34. The symmetric capacity of a 5-input symmetric channel with error probability p is

$$I(W) = 1 + p\log_5(\frac{p}{4}) + (1-p)\log_5(1-p). \tag{47}$$

For the proof of (47), consult the Appendix.
For $p = \frac{4}{5} = 0.8$, (47) equals 0, the proof is provided in the Appendix.
Since SER is around 0.8 for the most channels, our estimation is verified, meaning that the eavesdropper is being successfully kept ignorant about the secret message.

In the first subplot of Fig. 39, we illustrate the rate in the legitimate receiver as a function of the tolerance of the SER δ, $\delta \in [0, 0.2]$. The tolerance of the SER is the acceptable range within which the SER can vary. If $\delta = 0.2$, we take into consideration all the channels that have SER up to 0.2 in the legitimate receiver. We observe that as we increase δ, the rate increases since more channels are counted as good channels in the

estimation. The Pascal-matrix implementation performs better than the original's scheme since for smaller tolerance values the former achieves greater rate.

y \ x	**0**	**1**	**2**	**3**	**4**
0	0.9920	0.0022	0.0020	0.0021	0.0015
1	0.0019	0.9923	0.0018	0.0020	0.0021
2	0.0018	0.0018	0.9923	0.0019	0.0017
3	0.0020	0.0019	0.0017	0.9918	0.0020
4	0.0022	0.0018	0.0021	0.0022	0.9926

Figure 40: Transition probability matrix for the 5-input Symmetric Channel (when the underlying 5-input erasure channel has $\epsilon = 0.1$, I = 0.9, R = 0.85).

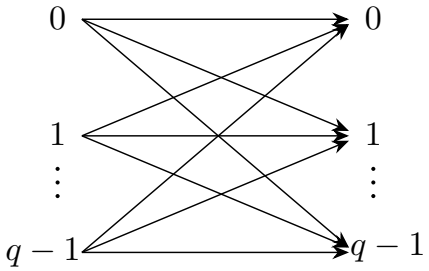
y \ x	**0**	**1**	**2**	**3**	**4**
0	0.7218	0.0715	0.0703	0.0714	0.0693
1	0.0698	0.7158	0.0689	0.0728	0.0719
2	0.0713	0.0685	0.7220	0.0713	0.0722
3	0.0685	0.0703	0.0701	0.7165	0.0706
4	0.0686	0.0739	0.0688	0.0687	0.7160

Figure 41: Transition probability matrix for the 5-input Symmetric Channel (when the underlying 5-input erasure channel has $\epsilon = 0.4$, I = 0.6, R = 0.55).

y \ x	**0**	**1**	**2**	**3**	**4**
0	0.9868	0.0034	0.0035	0.0031	0.0019
1	0.0033	0.9869	0.0038	0.0042	0.0024
2	0.0037	0.0042	0.9862	0.0033	0.0034
3	0.0029	0.0028	0.0021	0.9868	0.0028
4	0.0032	0.0027	0.0044	0.0026	0.9895

Figure 42: Transition probability matrix for the 5-input Symmetric Channel (when the underlying 5-input erasure channel has $\epsilon = 0.8$, I = 0.2, R = 0.15).

In the second subplot of Fig. 39, we illustrate the secrecy rate as a function of the tolerance δ, $\delta \in [0, 0.2]$. To estimate the secrecy rate we take into consideration the channels that have SER smaller than the tolerance δ in the legitimate receiver's side and at the same time their SER falls within the range $0.8-\delta < \text{SER} < 0.8+\delta$ in the eavesdropper side (for SER = 0.8 the security condition holds). We observe again better performance for the Pascal-matrix implementation as we increase the tolerance δ. As a reference, we highlight the secrecy capacity $C_s = 0.6 - 0.5 = 0.1$.



$\hat{u}$ \ u	**0**	**1**	$\cdots$	$\boldsymbol{q-1}$
0	$1-p$	$p/(q-1)$	$\cdots$	$p/(q-1)$
1	$p/(q-1)$	$1-p$	$\cdots$	$p/(q-1)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\boldsymbol{q-1}$	$p/(q-1)$	$p/(q-1)$	$\cdots$	$1-p$

q: prime number

Figure 43: The q-ary input symmetric channel with its transition probability matrix.

We expect similar performance for all q-ary input erasure channels, when q is a prime number. In Fig. 43, we illustrate the q-ary input symmetric channel, when q is a prime and its transition probability matrix. The symmetric capacity of the q-ary symmetric channel is

$$I(W) = 1 + p\log_q\left(\frac{p}{q-1}\right) + (1-p)\log_q(1-p). \tag{48}$$

and for $p = \frac{q-1}{q}$, the symmetric capacity equals 0.
For the proof, consult the Appendix.

Appendix

Proof of Eq. (10):

$$\begin{aligned} W_N^{(i)}(y_1^N, u_1^{i-1}|u_i) &= \frac{p(y_1^N, u_1^{i-1}, u_i)}{p(u_i)} = \frac{p(y_1^N, u_1^i)}{p(u_i)} = \frac{\sum_{u_{i+1}^N} p(y_1^N, u_1^i, u_{i+1}^N)}{p(u_i)} \\ &= \frac{\sum_{u_{i+1}^N} p(y_1^N, u_1^N)}{p(u_i)} = \frac{\sum_{u_{i+1}^N} p(y_1^N|u_1^N)p(u_1^N)}{p(u_i)} = \frac{\sum_{u_{i+1}^N} p(y_1^N|u_1^N)2^{-N}}{2^{-1}} \\ &= \sum_{u_{i+1}^N} \frac{1}{2^{N-1}} W_N(y_1^N|u_1^N). \end{aligned}$$

Proof of Eq. (11):

$$W_2^{(1)}(y_1, y_2|u_1) = \frac{1}{2}\sum_{u_2} W(y_1|u_1 \oplus u_2)W(y_2|u_2)$$

$$\left.\begin{aligned} p(y_1, y_2|u_1) &= \frac{p(y_1, y_2, u_1)}{p(u_1)} \\ p(x, y) &= \sum_z p(x, y, z) \end{aligned}\right\} \qquad p(y_1, y_2|u_1) = \frac{\sum_{u_2} p(y_1, y_2, u_1, u_2)}{p(u_1)}$$

$$\begin{aligned} p(y_1, y_2|u_1) &= \frac{1}{p(u_1)}\sum_{u_2} p(y_1, y_2|u_1, u_2)p(u_1, u_2) = \frac{1}{p(u_1)}\sum_{u_2} p(y_1, y_2|u_1, u_2)p(u_1)p(u_2) \\ &= p(u_2)\sum_{u_2} p(y_1, y_2|u_1, u_2) = \frac{1}{2}\sum_{u_2} p(y_1, y_2|x_1, x_2) = \frac{1}{2}\sum_{u_2} p(y_1|x_1)p(y_2|x_2) \\ &= \frac{1}{2}\sum_{u_2} p(y_1|u_1 \oplus u_2)p(y_2|u_2). \end{aligned}$$

Proof of Eq. (12):

$$W_2^{(2)}(y_1, y_2, u_1|u_2) = \frac{1}{2}W(y_1|u_1 \oplus u_2)W(y_2|u_2)$$

Applying the Chain rule: $p(A, B|C) = p(A|C)p(B|A, C)$

$$\begin{aligned} p(y_1, y_2, u_1|u_2) &= p(u_1|u_2)p(y_1, y_2|u_1, u_2) = p(u_1)p(y_1, y_2|u_1, u_2) \\ &= \frac{1}{2}p(y_1, y_2|x_1, x_2) = \frac{1}{2}p(y_1|x_1)p(y_2|x_2) \\ &= \frac{1}{2}p(y_1|u_1 \oplus u_2)p(y_2|u_2). \end{aligned}$$

Proof of Eq. (26):

$$L_N^{(2i-1)}(y_1^N, \hat{u}_1^{2i-2}) \overset{(25)}{=} \frac{W_N^{(2i-1)}(y_1^N, \hat{u}_1^{2i-2}|u_{2i-1}=0)}{W_N^{(2i-1)}(y_1^N, \hat{u}_1^{2i-2}|u_{2i-1}=1)}$$
$$\overset{(28)}{=} \frac{\sum_{u_{2i}} \frac{1}{2} W_{N/2}^{(i)}(y_1^{N/2}, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}|u_{2i-1}=0\oplus u_{2i}) W_{N/2}^{(i)}(y_{N/2+1}^N, u_{1,e}^{2i-2}|u_{2i})}{\sum_{u_{2i}} \frac{1}{2} W_{N/2}^{(i)}(y_1^{N/2}, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}|u_{2i-1}=1\oplus u_{2i}) W_{N/2}^{(i)}(y_{N/2+1}^N, u_{1,e}^{2i-2}|u_{2i})}$$

$$= \frac{W_{N/2}^{(i)}(y_1^{N/2}, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}|0) W_{N/2}^{(i)}(y_{N/2+1}^N, u_{1,e}^{2i-2}|0) + W_{N/2}^{(i)}(y_1^{N/2}, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}|1) W_{N/2}^{(i)}(y_{N/2+1}^N, u_{1,e}^{2i-2}|1)}{W_{N/2}^{(i)}(y_1^{N/2}, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}|1) W_{N/2}^{(i)}(y_{N/2+1}^N, u_{1,e}^{2i-2}|0) + W_{N/2}^{(i)}(y_1^{N/2}, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}|0) W_{N/2}^{(i)}(y_{N/2+1}^N, u_{1,e}^{2i-2}|1)}$$

$$= \frac{\frac{W_{N/2}^{(i)}(a|0) W_{N/2}^{(i)}(b|0)}{W_{N/2}^{(i)}(a|1) W_{N/2}^{(i)}(b|1)} + \frac{W_{N/2}^{(i)}(a|1) W_{N/2}^{(i)}(b|1)}{W_{N/2}^{(i)}(a|1) W_{N/2}^{(i)}(b|1)}}{\frac{W_{N/2}^{(i)}(a|1) W_{N/2}^{(i)}(b|0)}{W_{N/2}^{(i)}(a|1) W_{N/2}^{(i)}(b|1)} + \frac{W_{N/2}^{(i)}(a|0) W_{N/2}^{(i)}(b|1)}{W_{N/2}^{(i)}(a|1) W_{N/2}^{(i)}(b|1)}} = \frac{\frac{W_{N/2}^{(i)}(a|0)}{W_{N/2}^{(i)}(a|1)} \cdot \frac{W_{N/2}^{(i)}(b|0)}{W_{N/2}^{(i)}(b|1)} + 1}{\cancel{\frac{W_{N/2}^{(i)}(a|1)}{W_{N/2}^{(i)}(a|1)}} \cdot \frac{W_{N/2}^{(i)}(b|0)}{W_{N/2}^{(i)}(b|1)} + \frac{W_{N/2}^{(i)}(a|0)}{W_{N/2}^{(i)}(a|1)} \cdot \cancel{\frac{W_{N/2}^{(i)}(b|1)}{W_{N/2}^{(i)}(b|1)}}}$$

$$\overset{(25)}{=} \frac{L_{N/2}^{(i)}(y_1^{N/2}, \hat{u}_{1,o}^{2i-2} \oplus \hat{u}_{1,e}^{2i-2}) \cdot L_{N/2}^{(i)}(y_{N/2+1}^N, \hat{u}_{1,e}^{2i-2}) + 1}{L_{N/2}^{(i)}(y_{N/2+1}^N, \hat{u}_{1,e}^{2i-2}) + L_{N/2}^{(i)}(y_1^{N/2}, \hat{u}_{1,o}^{2i-2} \oplus \hat{u}_{1,e}^{2i-2})}.$$

Proof of Eq. (27):

$$L_N^{(2i)}(y_1^N, \hat{u}_1^{2i-1}) \overset{(25)}{=} \frac{W_N^{(2i)}(y_1^N, \hat{u}_1^{2i-1}|u_{2i}=0)}{W_N^{(2i)}(y_1^N, \hat{u}_1^{2i-1}|u_{2i}=1)}$$
$$\overset{(29)}{=} \frac{\frac{1}{2} W_{N/2}^{(i)}(y_1^{N/2}, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}|u_{2i-1}\oplus u_{2i}=0) W_{N/2}^{(i)}(y_{N/2+1}^N, u_{1,e}^{2i-2}|u_{2i}=0)}{\frac{1}{2} W_{N/2}^{(i)}(y_1^{N/2}, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}|u_{2i-1}\oplus u_{2i}=1) W_{N/2}^{(i)}(y_{N/2+1}^N, u_{1,e}^{2i-2}|u_{2i}=1)},$$

If $u_{2i-1}=0$, then we write

$$L_N^{(2i)}(y_1^N, \hat{u}_1^{2i-1}) = \frac{W_{N/2}^{(i)}(y_1^{N/2}, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}|0)}{W_{N/2}^{(i)}(y_1^{N/2}, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}|1)} \cdot \frac{W_{N/2}^{(i)}(y_{N/2+1}^N, u_{1,e}^{2i-2}|0)}{W_{N/2}^{(i)}(y_{N/2+1}^N, u_{1,e}^{2i-2}|1)}$$
$$\overset{(25)}{=} L_{N/2}^{(i)}(y_1^{N/2}, \hat{u}_{1,o}^{2i-2} \oplus \hat{u}_{1,e}^{2i-2}) \cdot L_{N/2}^{(i)}(y_{N/2+1}^N, \hat{u}_{1,e}^{2i-2}). \tag{49}$$

If $u_{2i-1}=1$, then we write

$$L_N^{(2i)}(y_1^N, \hat{u}_1^{2i-1}) = \frac{W_{N/2}^{(i)}(y_1^{N/2}, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}|1)}{W_{N/2}^{(i)}(y_1^{N/2}, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}|0)} \cdot \frac{W_{N/2}^{(i)}(y_{N/2+1}^N, u_{1,e}^{2i-2}|0)}{W_{N/2}^{(i)}(y_{N/2+1}^N, u_{1,e}^{2i-2}|1)}$$
$$\overset{(25)}{=} L_{N/2}^{(i)}(y_1^{N/2}, \hat{u}_{1,o}^{2i-2} \oplus \hat{u}_{1,e}^{2i-2})^{-1} \cdot L_{N/2}^{(i)}(y_{N/2+1}^N, \hat{u}_{1,e}^{2i-2}). \tag{50}$$

Combining (49) *and* (50)*, we obtain (27):*
$L_N^{(2i)}(y_1^N, \hat{u}_1^{2i-1}) = [L_{N/2}^{(i)}(y_1^{N/2}, \hat{u}_{1,o}^{2i-2} \oplus \hat{u}_{1,e}^{2i-2})]^{1-2\hat{u}_{2i-1}} \cdot L_{N/2}^{(i)}(y_{N/2+1}^N, \hat{u}_{1,e}^{2i-2})$.

Proof of Eq. (28):

$$\begin{aligned}
W_{2N}^{(2i-1)}(y_1^{2N}, u_1^{2i-2}|u_{2i-1}) &\overset{(10)}{=} \sum_{u_{2i}^{2N}} \frac{1}{2^{2N-1}} W_{2N}(y_1^{2N}|u_1^{2N}) \\
&= \sum_{u_{2i,e}^{2N},\, u_{2i,o}^{2N}} \frac{1}{2^{2N-1}} W_N(y_1^N|u_{1,o}^{2N} \oplus u_{1,e}^{2N}) W_N(y_{N+1}^{2N}|u_{1,e}^{2N}) \\
&= \sum_{u_{2i}} \frac{1}{2} \sum_{u_{2i+1,e}^{2N}} \frac{1}{2^{N-1}} W_N(y_{N+1}^{2N}|u_{1,e}^{2N}) \sum_{u_{2i+1,o}^{2N}} \frac{1}{2^{N-1}} W_N(y_1^N|u_{1,o}^{2N} \oplus u_{1,e}^{2N}) \\
&\overset{(10)}{=} \sum_{u_{2i}} \frac{1}{2} W_N^{(i)}(y_1^N, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}|u_{2i-1} \oplus u_{2i}) \sum_{u_{2i+1,e}^{2N}} \frac{1}{2^{N-1}} W_N(y_{N+1}^{2N}|u_{1,e}^{2N}) \\
&= \sum_{u_{2i}} \frac{1}{2} W_N^{(i)}(y_1^N, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}|u_{2i-1} \oplus u_{2i}) W_N^{(i)}(y_1^N, u_{1,e}^{2i-2}|u_{2i}).
\end{aligned}$$

Proof of Eq. (29):

$$\begin{aligned}
W_{2N}^{(2i)}(y_1^{2N}, u_1^{2i-1}|u_{2i}) &\overset{(10)}{=} \sum_{u_{2i+1}^{2N}} \frac{1}{2^{2N-1}} W_{2N}(y_1^{2N}|u_1^{2N}) \\
&\overset{(\star)}{=} \sum_{u_{2i+1,e}^{2N},\, u_{2i+1,o}^{2N}} \frac{1}{2^{2N-1}} W_N(y_1^N|u_{1,o}^{2N} \oplus u_{1,e}^{2N}) W_N(y_{N+1}^{2N}|u_{1,e}^{2N}) \\
&= \frac{1}{2} \sum_{u_{2i+1,e}^{2N}} \frac{1}{2^{N-1}} W_N(y_{N+1}^{2N}|u_{1,e}^{2N}) \sum_{u_{2i+1,o}^{2N}} \frac{1}{2^{N-1}} W_N(y_1^N|u_{1,o}^{2N} \oplus u_{1,e}^{2N}) \\
&\overset{(10)}{=} \frac{1}{2} W_N^{(i)}(y_1^N, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}|u_{2i-1} \oplus u_{2i}) \sum_{u_{2i+1,e}^{2N}} \frac{1}{2^{N-1}} W_N(y_{N+1}^{2N}|u_{1,e}^{2N}) \\
&= \frac{1}{2} W_N^{(i)}(y_1^N, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}|u_{2i-1} \oplus u_{2i}) W_N^{(i)}(y_1^N, u_{1,e}^{2i-2}|u_{2i}).
\end{aligned}$$

$$\begin{aligned}
(\star)\ p(y_1^{2N}|u_1^{2N}) &= p(y_1^{2N}|v_1^{2N}) = p(y_1^N, y_{N+1}^{2N}|v_{1,o}^{2N}, v_{1,e}^{2N}) = p(y_1^N|v_{1,o}^{2N}) \cdot p(y_{N+1}^{2N}, v_{1,e}^{2N}) \\
&= p(y_1^N|u_{1,o}^{2N} \oplus u_{1,e}^{2N}) \cdot p(y_{N+1}^{2N}|u_{1,e}^{2N}).
\end{aligned}$$

Proof of Eq. (31):
We first calculate the denominator of (30). If $y \in \{0, 1, \cdots, q-1\}$, then

$$\begin{aligned}
\sum_{y=0} \sum_{x' \in \mathcal{X}} \frac{1}{q} W(0|x') &= \frac{1}{q} W(0|0) + \frac{1}{q} W(0|1) + \cdots + \frac{1}{q} W(0|q-1) \\
&= \frac{1}{q} \cdot (1-\epsilon) + \frac{1}{q} \cdot (1-\epsilon) + \cdots \frac{1}{q} \cdot (1-\epsilon) = q \cdot \frac{1}{q} \cdot (1-\epsilon) = 1-\epsilon.
\end{aligned}$$

If $y =?$, then

$$\begin{aligned}
\sum_{y=?} \sum_{x' \in \mathcal{X}} \frac{1}{q} W(?|x') &= \frac{1}{q} W(?|0) + \frac{1}{q} W(?|1) + \cdots + \frac{1}{q} W(?|q-1) \\
&= \frac{1}{q} \cdot \epsilon + \frac{1}{q} \cdot \epsilon + \cdots \frac{1}{q} \cdot \epsilon = q \cdot \frac{1}{q} \cdot \epsilon = \epsilon.
\end{aligned}$$

$$
\begin{aligned}
I(W) &\overset{(30)}{=} \sum_{x\in\mathcal{X}} \Bigg[\frac{1}{q}W(0|x)\log_q\frac{q\cdot W(0|x)}{1-\epsilon} + \frac{1}{q}W(1|x)\log_q\frac{q\cdot W(1|x)}{1-\epsilon} + \\
&\qquad \cdots + \frac{1}{q}W(q-1|x)\log_q\frac{q\cdot W(q-1|x)}{1-\epsilon} + \frac{1}{q}W(?|x)\log_q\frac{W(?|x)}{\epsilon}\Bigg] \\
&= \frac{1}{q}(1-\epsilon)\log_q\frac{q\cdot(1-\epsilon)}{1-\epsilon} + \frac{1}{q}(1-\epsilon)\log_q\frac{q\cdot(1-\epsilon)}{1-\epsilon} + \cdots + \frac{1}{q}(1-\epsilon)\log_q\frac{q\cdot(1-\epsilon)}{1-\epsilon} \\
&= q\cdot\frac{1}{q}\cdot(1-\epsilon) = 1-\epsilon.
\end{aligned}
$$

For q = 3 and input alphabet X = {0, 1, 2} and from (35), the kernel matrix is defined as $\mathrm{F} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix}$.

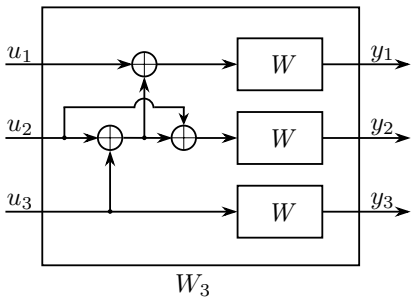


Figure 44: Synthesized channel W_3.

First, we will prove that single step channel transformation preserves the symmetric capacity in q-ary input erasure channels as well.

Since X_0, X_1, X_2 are independent, the mutual information between the joint channel input X_0, X_1, X_2 and the joint channel output Y_0, Y_1, Y_2 is

$$
\begin{aligned}
I(X_0, X_1, X_2; Y_0, Y_1, Y_2) &\overset{(\star)}{=} H(Y_0, Y_1, Y_2) - H(Y_0, Y_1, Y_2|X_0, X_1, X_2) \\
&= H(Y_0) + H(Y_1|Y_0) + H(Y_2|Y_1, Y_0) \\
&\quad - H(Y_0|X_0, X_1, X_2) - H(Y_1|Y_0, X_0, X_1, X_2) \\
&\quad - H(Y_2|Y_0, Y_1, X_0, X_1, X_2) \\
&\overset{(\star\star)}{=} H(Y_0) + H(Y_1) + H(Y_2) - H(Y_0|X_0) - H(Y_1|X_1) - H(Y_2|X_2) \\
&\overset{(\star)}{=} I(X_0; Y_0) + I(X_1; Y_1) + I(X_2; Y_2) \\
&= 3I(X_0; Y_0) \\
&= 3I(W). \qquad (51)
\end{aligned}
$$

($\star$) Mutual information is symmetric, therefore $I(X;Y) = I(Y;X)$ and $I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$.
($\star\star$) Since, by the definition of the discrete memoryless channel, Y_i depends only on X_i and is conditionally independent of anything else.

$\mathrm{F} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix}$ is invertible.

Thus, X_0, X_1, and X_2 are in invertible correspondence with U_0, U_1, and U_2, hence

$$I(X_0, X_1, X_2; Y_0, Y_1, Y_2) = I(U_0, U_1, U_2; Y_0, Y_1, Y_2). \tag{52}$$

By applying the chain rule for mutual information,

$$\begin{aligned} I(U_0^2; Y_0^2) &= I(U_0; Y_0^2) + I(U_1^2; Y_0^2 | U_0) \overset{U_0, U_1^2 i.i.d.}{=} \\ &= I(U_0; Y_0^2) + I(U_1^2; Y_0^2, U_0) = \\ &= I(U_0; Y_0^2) + I(U_1; Y_0^2, U_0) + I(U_2; Y_0^2, U_0 | U_1) \overset{U_1, U_2 i.i.d.}{=} \\ &= I(U_0; Y_0^2) + I(U_1; Y_0^2, U_0) + I(U_2; Y_0^2, U_0^1) = \\ &= I(W') + I(W'') + I(W'''). \end{aligned} \tag{53}$$

Then, from (51),(52), and (53), we have

$$I(W) = \frac{I(W') + I(W'') + I(W''')}{3}. \tag{54}$$

To find the recursive formulas of the symmetric capacities of the fabricated channels we start by calculating the erasure probabilities.

The table below shows different output values, where • indicates that y_i has been received as transmitted and ? indicates that a y_i is an erasure. The symbol ✔ indicates that the received data are sufficient to decode u_1^3 and ✘ indicates that the received data are insufficient to decode u_1^3.

For $W_3^{(1)}$,

$$(y_1, y_2, y_3) = \begin{cases} u_1 \oplus u_2 \oplus u_3 & 2u_2 \oplus u_3 & u_3 & w.p. & decodable \\ \bullet & \bullet & \bullet & (1-\epsilon)^3 & \checkmark \\ ? & \bullet & \bullet & \epsilon(1-\epsilon)^2 & \times \\ \bullet & ? & \bullet & (1-\epsilon)\epsilon(1-\epsilon) & \times \\ \bullet & \bullet & ? & (1-\epsilon)^2\epsilon & \times \\ ? & ? & \bullet & \epsilon^2(1-\epsilon) & \times \\ ? & \bullet & ? & \epsilon(1-\epsilon)\epsilon & \times \\ \bullet & ? & ? & (1-\epsilon)\epsilon^2 & \times \\ ? & ? & ? & \epsilon^3 & \times \end{cases}$$

$$\begin{aligned} P_e(W_3^{(1)}) &= 1 - (1-\epsilon)^3 = 1 - (1 - 3\epsilon + \epsilon^2 - \epsilon^3) \\ &= 3\epsilon - 3\epsilon^2 + \epsilon^3. \end{aligned} \tag{55}$$

For $W_3^{(2)}$,

$$(y_1, y_2, y_3, u_1) = \begin{cases} u_1 \oplus u_2 \oplus u_3 & 2u_2 \oplus u_3 & u_3 & u_1 & w.p. & decodable \\ \bullet & \bullet & \bullet & \bullet & (1-\epsilon)^3 & \checkmark \\ ? & \bullet & \bullet & \bullet & \epsilon(1-\epsilon)^2 & \checkmark \\ \bullet & ? & \bullet & \bullet & (1-\epsilon)\epsilon(1-\epsilon) & \checkmark \\ \bullet & \bullet & ? & \bullet & (1-\epsilon)^2\epsilon & \checkmark \\ ? & ? & \bullet & \bullet & \epsilon^2(1-\epsilon) & \times \\ ? & \bullet & ? & \bullet & \epsilon(1-\epsilon)\epsilon & \times \\ \bullet & ? & ? & \bullet & (1-\epsilon)\epsilon^2 & \times \\ ? & ? & ? & \bullet & \epsilon^3 & \times \end{cases}$$

$$\begin{aligned} P_e(W_3^{(2)}) &= \epsilon^3 + 3\epsilon^2(1-\epsilon) = \epsilon^3 + 3\epsilon^2 - 3\epsilon^3 \\ &= 3\epsilon^2 - 2\epsilon^3. \end{aligned} \tag{56}$$

For $W_3^{(3)}$,

$$(y_1, y_2, y_3, u_1, u_2) = \left\{ \begin{array}{ccccccc} u_1 \oplus u_2 \oplus u_3 & 2u_2 \oplus u_3 & u_3 & u_1 & u_2 & w.p. & decodable \\ \bullet & \bullet & \bullet & \bullet & \bullet & (1-\epsilon)^3 & \checkmark \\ ? & \bullet & \bullet & \bullet & \bullet & \epsilon(1-\epsilon)^2 & \checkmark \\ \bullet & ? & \bullet & \bullet & \bullet & (1-\epsilon)\epsilon(1-\epsilon) & \checkmark \\ \bullet & \bullet & ? & \bullet & \bullet & (1-\epsilon)^2\epsilon & \checkmark \\ ? & ? & \bullet & \bullet & \bullet & \epsilon^2(1-\epsilon) & \checkmark \\ ? & \bullet & ? & \bullet & \bullet & \epsilon(1-\epsilon)\epsilon & \checkmark \\ \bullet & ? & ? & \bullet & \bullet & (1-\epsilon)\epsilon^2 & \checkmark \\ ? & ? & ? & \bullet & \bullet & \epsilon^3 & \text{✘} \end{array} \right.$$

$$P_e(W_3^{(3)}) = \epsilon^3. \tag{57}$$

Given $P_e(W) = \epsilon$, $I(W) = 1 - \epsilon$, and the erasure probabilities of the synthesized channels that we calculated above, we have

$$\begin{aligned} P_e(W') &= P_e(W)^3 - 3P_e(W)^2 + 3P_e(W) \\ 1 - I(W') &= (1 - I(W))^3 - 3(1 - I(W))^2 + 3(1 - I(W) \\ 1 - I(W') &= 1 - 3I(W) + 3I(W)^2 - I(W)^3 - 3 + 6I(W) - 3I(W)^2 + 3 - 3I(W) \\ I(W') &= I(W)^3 \\ i.e., \quad I(W_N^{(3i-2)}) &= I(W_{N/3}^{(i)})^3, \end{aligned} \tag{58}$$

$$\begin{aligned} P_e(W'') &= 3P_e(W)^2 - 2P_e(W)^3 \\ 1 - I(W'') &= 3(1 - I(W))^2 - 2(1 - I(W))^3 \\ 1 - I(W'') &= 3 - 6I(W) + 3I(W)^2 - 2 + 6I(W) - 6I(W)^2 + 2I(W)^3 \\ I(W'') &= 3I(W)^2 - 2I(W)^3 \\ i.e., \quad I(W_N^{(3i-1)}) &= 3I(W_{N/3}^{(i)})^2 - 2I(W_{N/3}^{(i)})^3, \end{aligned} \tag{59}$$

$$\begin{aligned} P_e(W''') &= P_e(W)^3 \\ 1 - I(W''') &= (1 - I(W))^3 \\ 1 - I(W''') &= 1 - 3I(W) + 3I(W)^2 - I(W)^3 \\ I(W''') &= 3I(W) - 3I(W)^2 + I(W)^3 \\ i.e., \quad I(W_N^{(3i)}) &= 3I(W_{N/3}^{(i)}) - 3I(W_{N/3}^{(i)})^2 + I(W_{N/3}^{(i)})^3. \end{aligned} \tag{60}$$

For ternary input, (38) becomes

$$W_N^{(i)}(y_1^N, u_1^{(i-1)} | u_i) = \sum_{u_{i+1}^N} \frac{1}{3^{N-1}} W_N(y_1^N | u_1^N). \tag{61}$$

The transition probabilities for the successive cancellation decoder are calculated below [11]. Symbols m_0 stand for *index* mod 3 = 1, symbols m_1 stand for *index* mod 3 = 2, and symbols m_2 stand for *index* mod 3 = 0. The operations are over GF(3).

$$
\begin{aligned}
W_{3N}^{(3i)}(y_1^{3N}, u_1^{3i-1}|u_{3i}) &\overset{(61)}{=} \sum_{u_{3i+1}^{3N}} \frac{1}{3^{3N-1}} W_{3N}(y_1^{3N}|u_1^{3N}) = \\
&= \sum_{u_{3i+1}^{3N}} \frac{1}{3^{3N-1}} W_N(y_1^N|u_{1,m_0}^{3N} \oplus u_{1,m_1}^{3N} \oplus u_{1,m_2}^{3N}) \\
&\qquad \cdot W_N(y_{N+1}^{2N}|2u_{1,m_1}^{3N} \oplus u_{1,m_2}^{3N}) W_N(y_{2N+1}^{3N}|u_{1,m_2}^{3N}) \\
&= \frac{1}{9} \sum_{u_{3i+1,m_0}^{3N}} \frac{1}{3^{N-1}} W_N(y_1^N|u_{1,m_0}^{3N} \oplus u_{1,m_1}^{3N} \oplus u_{1,m_2}^{3N}) \\
&\quad \cdot \sum_{u_{3i+1,m_1}^{3N}} \frac{1}{3^{N-1}} W_N(y_{N+1}^{2N}|2u_{1,m_1}^{3N} \oplus u_{1,m_2}^{3N}) \\
&\quad \cdot \sum_{u_{3i+1,m_2}^{3N}} \frac{1}{3^{N-1}} W_N(y_{2N+1}^{3N}|u_{1,m_2}^{3N}) \\
&= \frac{1}{9} W_N^{(i)}(y_1^N, u_{1,m_0}^{3i-3} \oplus u_{1,m_1}^{3i-3} \oplus u_{1,m_2}^{3i-3}|u_{3i-2} \oplus u_{3i-1} \oplus u_{3i}) \\
&\quad \cdot W_N^{(i)}(y_{N+1}^{2N}, 2u_{1,m_1}^{3i-3} \oplus u_{1,m_2}^{3i-3}|2u_{3i-1} \oplus u_{3i}) W_N^{(i)}(y_{2N+1}^{3N}, u_{1,m_2}^{3i-3}|u_{3i}),
\end{aligned}
$$

$$
\begin{aligned}
W_{3N}^{(3i-1)}(y_1^{3N}, u_1^{3i-2}|u_{3i-1}) &\overset{(61)}{=} \sum_{u_{3i}^{3N}} \frac{1}{3^{3N-1}} W_{3N}(y_1^{3N}|u_1^{3N}) = \\
&= \sum_{\substack{u_{3i,m_0}^{3N}, u_{3i,m_1}^{3N}, \\ u_{3i,m_2}^{3N}}} \frac{1}{3^{3N-1}} W_N(y_1^N|u_{1,m_0}^{3N} \oplus u_{1,m_1}^{3N} \oplus u_{1,m_2}^{3N}) \\
&\qquad \cdot W_N(y_{N+1}^{2N}|2u_{1,m_1}^{3N} \oplus u_{1,m_2}^{3N}) W_N(y_{2N+1}^{3N}|u_{1,m_2}^{3N}) \\
&= \sum_{u_{3i}} \frac{1}{9} \sum_{u_{3i+1,m_0}^{3N}} \frac{1}{3^{N-1}} W_N(y_1^N|u_{1,m_0}^{3N} \oplus u_{1,m_1}^{3N} \oplus u_{1,m_2}^{3N}) \\
&\quad \cdot \sum_{u_{3i+1,m_1}^{3N}} \frac{1}{3^{N-1}} W_N(y_{N+1}^{2N}|2u_{1,m_1}^{3N} \oplus u_{1,m_2}^{3N}) \\
&\quad \cdot \sum_{u_{3i+1,m_2}^{3N}} \frac{1}{3^{N-1}} W_N(y_{2N+1}^{3N}|u_{1,m_2}^{3N}) \\
&= \sum_{u_{3i}} \frac{1}{9} W_N^{(i)}(y_1^N, u_{1,m_0}^{3i-3} \oplus u_{1,m_1}^{3i-3} \oplus u_{1,m_2}^{3i-3}|u_{3i-2} \oplus u_{3i-1} \oplus u_{3i}) \\
&\qquad \cdot W_N^{(i)}(y_{N+1}^{2N}, 2u_{1,m_1}^{3i-3} \oplus u_{1,m_2}^{3i-3}|2u_{3i-1} \oplus u_{3i}) W_N^{(i)}(y_{2N+1}^{3N}, u_{1,m_2}^{3i-3}|u_{3i}),
\end{aligned}
$$

$$
\begin{aligned}
W_{3N}^{(3i-2)}(y_1^{3N}, u_1^{3i-3}|u_{3i-2}) &\overset{(61)}{=} \sum_{u_{3i-1}^{3N}} \frac{1}{3^{3N-1}} W_{3N}(y_1^{3N}|u_1^{3N}) = \\
&= \sum_{u_{3i-1}^{3N}} \frac{1}{3^{3N-1}} W_N(y_1^N|u_{1,m_0}^{3N} \oplus u_{1,m_1}^{3N} \oplus u_{1,m_2}^{3N}) \\
&\qquad \cdot W_N(y_{N+1}^{2N}|2u_{1,m_1}^{3N} \oplus u_{1,m_2}^{3N}) W_N(y_{2N+1}^{3N}|u_{1,m_2}^{3N}) \\
&= \sum_{u_{3i}, u_{3i-1}} \frac{1}{9} \sum_{u_{3i+1,m_0}^{3N}} \frac{1}{3^{N-1}} W_N(y_1^N|u_{1,m_0}^{3N} \oplus u_{1,m_1}^{3N} \oplus u_{1,m_2}^{3N}) \\
&\qquad \cdot \sum_{u_{3i+1,m_1}^{3N}} \frac{1}{3^{N-1}} W_N(y_{N+1}^{2N}|2u_{1,m_1}^{3N} \oplus u_{1,m_2}^{3N}) \\
&\qquad \cdot \sum_{u_{3i+1,m_2}^{3N}} \frac{1}{3^{N-1}} W_N(y_{2N+1}^{3N}|u_{1,m_2}^{3N}) \\
&= \sum_{u_{3i}, u_{3i-1}} \frac{1}{9} W_N^{(i)}(y_1^N, u_{1,m_0}^{3i-3} \oplus u_{1,m_1}^{3i-3} \oplus u_{1,m_2}^{3i-3}|u_{3i-2} \oplus u_{3i-1} \oplus u_{3i}) \\
&\qquad \cdot W_N^{(i)}(y_{N+1}^{2N}, 2u_{1,m_1}^{3i-3} \oplus u_{1,m_2}^{3i-3}|2u_{3i-1} \oplus u_{3i}) W_N^{(i)}(y_{2N+1}^{3N}, u_{1,m_2}^{3i-3}|u_{3i}).
\end{aligned}
$$

Proof of (45):
Observe that

$$R_s = \frac{|\mathcal{S}|}{N} = \frac{|A(W^*)|}{N} - \frac{|A(W)|}{N} = C(W^*) - C(W)$$

by using (43), the fact that $A(W) \subseteq A(W^*)$, and $\lim_{N\to\infty} \frac{A(W)}{N} = C(W)$ [8].

Proof of Eq. (46):
In the case that W is TSC(p), we have

$$
\begin{aligned}
I(W) &\overset{(30)}{=} \sum_{x\in\mathcal{X}} \left[\frac{1}{3}W(0|x)\log_3(3W(0|x)) + \frac{1}{3}W(1|x)\log_3(3W(1|x)) + \frac{1}{3}W(2|x)\log_3(3W(2|x))\right] \\
&= (1-p)\log_3(3(1-p)) + p\log_3(\frac{3p}{2}) \\
&= (1-p)(\log_3 3 + \log_3(1-p)) + p(\log_3 3 + \log_3 \frac{p}{2}) \\
&= 1 + p\log_3(\frac{p}{2}) + (1-p)\log_3(1-p),
\end{aligned}
$$

$$
\begin{aligned}
I(W) = 0 \Leftrightarrow \\
1 + p\log_3(\frac{p}{2}) + (1-p)\log_3(1-p) &= 0 \\
\log_3 3 + \log_3\left(\frac{p}{2}\right)^p + \log_3(1-p)^{(1-p)} &= \log_3 1 \\
\log_3(3\cdot\left(\frac{p}{2}\right)^p \cdot (1-p)^{(1-p)}) &= \log_3 1 \\
3\cdot\left(\frac{p}{2}\right)^p \cdot (1-p)^{(1-p)} &= 1.
\end{aligned}
$$

For $p = \frac{2}{3}$, we get $3\cdot\frac{1}{3}^{2/3}\cdot\frac{1}{3}^{1/3} = 3\cdot\frac{1}{\sqrt[3]{3^2}}\cdot\frac{1}{\sqrt[3]{3}} = 3\cdot\frac{1}{\sqrt[3]{3^3}} = \frac{3}{3} = 1.$

In the case that W is a q-input symmetric channel with crossover probability p, we have

$$\begin{aligned}
I(W) &\overset{(30)}{=} \sum_{x\in\mathcal{X}} \left[\frac{1}{q}W(0|x)\log_q(qW(0|x)) + \frac{1}{q}W(1|x)\log_q(qW(1|x)) + ... + \frac{1}{q}W(q-1|x)\log_q(qW(q-1|x))\right] \\
&= (1-p)\log_q(q(1-p)) + p\log_q\left(q\cdot\frac{p}{q-1}\right) \\
&= (1-p)(\log_q q + \log_q(1-p)) + p\left(\log_q q + \log_q\frac{p}{q-1}\right) \\
&= 1 + p\log_q\left(\frac{p}{q-1}\right) + (1-p)\log_q(1-p),
\end{aligned}$$

$$\begin{aligned}
&I(W) = 0 \Leftrightarrow \\
&1 + p\log_q\left(\frac{p}{q-1}\right) + (1-p)\log_q(1-p) = 0 \\
&\log_q q + \log_q\left(\frac{p}{q-1}\right)^p + \log_q(1-p)^{(1-p)} = \log_q 1 \\
&\log_q\left(q\cdot\left(\frac{p}{q-1}\right)^p\cdot(1-p)^{(1-p)}\right) = \log_q 1 \\
&q\cdot\left(\frac{p}{q-1}\right)^p\cdot(1-p)^{(1-p)} = 1.
\end{aligned}$$

For $p = \dfrac{q-1}{q}$, we get $q\cdot\left(\dfrac{\frac{q-1}{q}}{q-1}\right)^{(q-1)/q}\cdot\left(1-\dfrac{q-1}{q}\right)^{1-(q-1)/q} = q\cdot\left(\dfrac{1}{q}\right)^{(q-1)/q}\cdot\left(\dfrac{1}{q}\right)^{1-(q-1)/q} = q\cdot\left(\dfrac{1}{q}\right)^1 = 1.$

References

[1] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3051–3073, July 2009.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory, 2nd Edition*, ISBN: 978-920-524-434-7.

[3] E. Şaşoğlu, E. Telatar and E. Arikan, "Polarization for arbitrary discrete memoryless channels," *IEEE Information Theory Workshop*, pp. 144–148, Oct. 2009.

[4] E. Şaşoğlu Ph.D. Thesis pp. 47–50, Switzerland, 2011.

[5] D. Doan, T. Sykes, and J. Smith, The Modular Pascal's Triangle, *http://orion.math.iastate.edu/reu/oldREU/modupasc.htm*

[6] A.D. Wyner, "The Wire-tap Channel," *Bell Syst. Tech J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[7] S. Leung-Yan-Cheong, "On a special class of wiretap channels," *IEEE Transactions on Information Theory*, vol. 23, no. 5, pp. 625–627, Sept. 1977.

[8] H. Mahdavifar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6428–6443, Oct. 2011.

[9] A. Torfi, S. Soleymani, S.M. Iranmanesh, H. Kazemi, R.A. Shirvani and V. Vakili, "Polar coding for achieving the capacity of marginal channels in nonbinary-input setting," *51st Annual Conference on Information Sciences and Systems (CISS)*, 2017.

[10] Tood K. Moon, *Error correction coding: mathematical methods and algorithms, 2nd Edition.*

[11] Ioannis-Themistoklis Papoutsidakis, "Pascal-matrix Polar Coding for Prime-input Channels," *Technical University of Crete*, July 2016.