

Polar Codes for Quantum Reading

Francisco Revson F. Pereira^{*1,2} and Stefano Mancini^{1,2}

¹School of Science and Technology, University of Camerino, I-62032 Camerino, Italy

²INFN, Sezione di Perugia, I-06123 Perugia, Italy

Abstract

Quantum reading provides a general framework where to formulate the statistical discrimination of quantum channels. Several paths have been taken for such a problem. However, there is much to be done in the avenue of optimizing channel discrimination using classical codes. At least two open questions can be pointed to: how to construct low complexity encoding schemes that are interesting for channel discrimination and, more importantly, how to develop capacity-achieving protocols. The aim of this paper is to present a solution to these questions using polar codes. Firstly, we characterize the rate and reliability of the channels under polar encoding. We also show that the error probability of the scheme proposed decays exponentially with respect to the code length. Lastly, an analysis of the optimal quantum states to be used as probes is given.

Keywords: Quantum Reading, Polar Codes, Capacity-Achieving Protocols.

1 Introduction

Quantum hypothesis testing aims to identify strategies in order to statistically discriminate quantum states or processes. The former is called quantum state discrimination and has been largely analyzed in the literature (starting from Refs. [14, 16]). The latter, under the name of quantum channel discrimination, has been recently addressed [8, 2, 11, 30, 29, 36, 10, 13, 12]. In its basic formulation, one has to identify a quantum channel selected from a set accordingly to a probability measure. This should be done by using suitable input state and output measurement. As such, it is a double-optimization problem and hence results in a daunting task. Performance are usually quantified in terms of minimum error probability and recently bounds on it were found for general strategies [24, 19]. Although theoretically subtle, quantum channel discrimination is interesting for practical applications. For instance, it is at the basis of the decoding procedure of two-way quantum cryptography [26] where the secret information is encoded in a Gaussian ensemble of phase-space displacements. It also appears in the quantum illumination of targets [20, 35], where the sensing of a remote low-reflective object in a bright thermal environment corresponds to the binary discrimination between a lossy channel (presence of target) and a depolarizing channel (absence of target). Following this line, quantum channel discrimination can be reformulated in the framework of quantum reading [23]. There, the data storage corresponds to a process of channel encoding, where information is recorded into a cell by storing a quantum channel picked from a given ensemble. Then readout corresponds to the process of channel decoding, which is equivalent to discriminating between the various channels in the ensemble. In such a setting, using a quantum resource, such as entanglement, was shown to overcome any classical strategy based on mixtures of coherent states, hence the name of “quantum” reading.

Efficient paths to quantum reading can be also envisaged by the use of coding techniques [25, 9]. There are at least two possible approaches: classical coding on the quantum memory cell labels and quantum coding on the probe states used in the readout. This paper takes the former one. Similar to channel coding, where redundancy is added to protect the information to be transmitted, we use classical codes to encode the information to be recorded in the memory cell. The goal is to decrease the error probability in discriminating the channels in the ensemble. There are several classical codes that one could use for quantum reading, depending on the goal in mind. In this work, we are interested in information-theoretically provable codes that could be used to attain high

*revson.ee@gmail.com

information rates and low error probability. Therefore, suitable codes can be derived from the family of polar codes.

Creating capacity-achieving codes has been a challenge since the development of Classical Information Theory. Even more difficult was to show there exist theoretically-proven such codes. Fortunately, developing the idea of information combining, Arikan was able to show that for binary memoryless symmetric (BMS) channels such codes exist [3]. They are called polar codes. The major achievement of Arikan's paper is to show a clever way to combine information, which leads to new synthetic channels manifesting a polarization phenomenon. They are commonly divided into two groups named "good" and "bad" channel¹. Furthermore, the fractions of good and bad channels are related to the capacity of the original channel in consideration. Encoding and decoding schemes take advantage of these properties to attain low complexity. It is worth mention that polar code construction highly depends on the channel into consideration.

Several papers extend or apply the idea of polarization in a diversity of channels and areas [32, 4, 31, 18]. In particular, polar codes have also been constructed to classical-quantum channels [37], and quantum channels [38, 27]. However, this is not the case when one considers quantum reading, which is the goal of this paper. The distribution of the channels to be discriminated can be symmetric or asymmetric. Asymmetric quantum reading can be seen as a generalization of the former quantum reading and this formulation can be justified by the measurement strategy implemented in the decoding process or energy constrain imposed over the channels. Since these two conditions are plausible hypothesis to be considered in classical digital memories, which are the main goal of channel model used in this paper, we will formulate our results for asymmetric quantum reading and, where it makes necessary, an adaptation for symmetric quantum reading is given. Therefore, our polar code will be design to, primarily, asymmetric quantum channels. In the classical paradigm, this has been done by Sutter, et al. [33], and Honda and Yamamoto [18]. Some of the ideas used in Ref. [18] are applied in this paper. However, as we will see, the transition from the classical to quantum paradigm is subtle and major changes are needed to be done to derive our construction. This is particularly true for the measurement process implemented in the decoding part. An adequately definition for rate and reliability is also needed.

The analysis of quantum memory cell in this paper is twofold. Initially, polar coding is applied on the labels of the quantum memory cell in order to decrease the error probability in distinguishing them. In this direction, we show that the polarization phenomenon can be characterized in quantum memory cell by studying its composing parts; namely, channel combining and channel splitting. The former constitutes a systematic approach to combine source bits so the polarization phenomenon emerges. An encoding matrix, sometimes also called combining function, describing the process is given in this paper. The latter is an information-theoretically analysis of the synthesized channels created in the channel combining part. Initially, the first level encoding process combining two quantum channels is analyzed. We show that the rate and reliability of the synthesized channels polarize. This is later used in the asymptotic analysis. A connection between rate and reliability is also presented, showing that rate is inversely proportional to reliability. Next, our first major result is given. It characterizes the asymptotic distribution of the synthesized channel with respect to the length code. This result is based on proving the existence, shown in this paper, of a symmetric quantum channel that can be used to obtain the statistics of the (asymmetric) quantum channel of interest. Lastly, in this first part, we draw the encoding and decoding scheme. The strategy used for encoding the frozen bits does not follow the scheme used in previous works on polar codes for classical-quantum channels. Additionally, examining the error probability obtained after the decoding scheme, we see that it decays exponentially with respect to the length code. The second part addresses the optimization problem of probe states. As a first-order approximation of classical digital memories, we consider amplitude damping channels as our channel model. The fundamental result of this part is showing that the optimal probe states are pure states. Even though we have taken into consideration just the first level polar encoding, it seems satisfactory supposing this can be extended to any N -level polar encoding.

¹The reason for such names is due to the capacity of these new channels been close to its maximum value and close to zero, respectively.

This paper is organized as follows. In Section 2 we present previous results on polar codes for classical-quantum channels, and some important definition used to characterize quantum memory cells. This section constitutes the fundamentals that our work is based on. Next, the main results are shown in Section 3. Encoding and decoding schemes are described and analyzed in detail, showing that the use of polar coding is also interesting for quantum memory cells. The following subject treated in this work is the optimization of states used to probe the channels. This is drawn in Section 4. Lastly, we draw our conclusions and some final remarks in Section 5.

1.1 Notation

We denote classical random variables as X, Y, U , whose realizations are elements of the finite sets $\mathcal{X}, \mathcal{Y}, \mathcal{U}$, respectively. The probability distributions are respectively represented by $p_X(x)$, $p_Y(y)$, and $p_U(u)$ for the random variables X, Y, U . In particular, X is always assumed to be discrete and, in some parts of the text, to be a Bernoulli random variable with values in $\{0, 1\}$ and $P_X(0) = p$ by using the notation $X \sim \text{Ber}(p)$. For such random variables, its binary Shannon entropy is defined as $H(X) = h(p) := -p \log(p) - (1-p) \log(1-p)$. The use of subscript and superscript on a letter indicates a sequence starting with the element denoted by the subscript and ending with the element denoted by the superscript, e.g. X_i^j (with $j > i$) is the sequence of random variables X_i, X_{i+1}, \dots, X_j . X^N simply stands for the sequence X_1, X_2, \dots, X_N . A classical memoryless channel is written as $W: \mathcal{X} \rightarrow \mathcal{Y}$ from the input alphabet \mathcal{X} to the output alphabet \mathcal{Y} . We write W^N to denote N uses of the channel W ; i.e., $W^N: \mathcal{X}^N \rightarrow \mathcal{Y}^N$ with $W^N(y^N|x^N) = \prod_{i=1}^N W(y_i|x_i)$ since W is memoryless. Quantum systems A, B , and C correspond to Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B$, and \mathcal{H}_C . The notation $A^N := A_1 A_2 \dots A_N$ denotes a joint system consisting of N subsystems, each of which is isomorphic to \mathcal{H}_A . Let $\mathcal{L}(\mathcal{H}_A)$ denote the algebra of bounded linear operators acting on a Hilbert space \mathcal{H}_A . The subset $\mathcal{L}_+(\mathcal{H}_A)$ of $\mathcal{L}(\mathcal{H}_A)$ denotes the set of all positive semi-definite operators. A special and important class of operators in $\mathcal{L}_+(\mathcal{H}_A)$ is the one containing density operators $\mathcal{D}(\mathcal{H}_A)$. A density operator $\rho_A \in \mathcal{D}(\mathcal{H}_A)$ is a positive semi-definite operator with unit trace, $\text{Tr}\{\rho_A\} = 1$, and represents the state of a quantum system A . Lastly, a quantum channel \mathcal{W} is a linear completely positive trace-preserving map from $\mathcal{L}(\mathcal{H}_A)$ to $\mathcal{L}(\mathcal{H}_B)$.

2 Preliminaries

This section is devoted to introducing the main concepts of polar codes and quantum reading. We begin with a brief overview of binary polar codes and their desirable attributes. They can be used to achieve the capacity of discrete memoryless channels (DMC); there are encoding and decoding efficient schemes; and it is possible to reach error probabilities that decay exponentially in the square root of the blocklength. In the following, quantum reading problematic is introduced. The general concept is given, followed by the channel model adopted for analyzing probe states. In the end, we present the quantities assumed as rate and reliability.

2.1 Classical-Quantum Polar Codes

Let $W: \mathcal{X} \rightarrow \mathcal{D}(\mathcal{H}_B)$ be a binary-input memoryless classical-quantum channel (cq channel) with input and output alphabets given by \mathcal{X} and $\mathcal{D}(\mathcal{H}_B)$, respectively. Suppose the random variable X is $\text{Ber}(\frac{1}{2})$. Then the transition probabilities can be derived from the joint input-output state

$$\rho^{XB} = \frac{1}{2} |0\rangle\langle 0|^X \otimes \rho_0^B + \frac{1}{2} |1\rangle\langle 1|^X \otimes \rho_1^B, \quad (1)$$

where $\rho_x^B = W(x)$, for $x = 0, 1$. Notice that we are using superscripts to emphasize the system to which each part belongs. Whenever this is clear from context, they will be omitted. One important characterization can be given to this cq channel: since the possible outcomes are uniformly distributed over all possible labels $x \in \mathcal{X}$, we say that the cq channel W is symmetric. Similarly, a cq channel is called asymmetric when this does not happen. Though in this subsection we are going

to deal with a symmetric channel, the rest of the text adopts asymmetric channel in order to provide a more general formulation. As a last comment to be made, we adopt as input alphabet $\mathcal{X} = \mathbb{Z}_2$, and arbitrary finite-dimensional output alphabet $\mathcal{D}(\mathcal{H}_B)$ and transition probability. This choice of input alphabet allows us to operate with their elements; in particular, we can use XOR operations, or sum mod 2 operations, with the elements of the input alphabet. An appropriate combination of these operations leads to channel polarization.

Channel polarization consists mainly of two parts. The first one is named channel combining, which describes a method of combining inputs of N cq channels. The second is channel splitting. This part is an information-theoretical analysis of new inputs and outputs that the channel combining produces. These new inputs and outputs generate synthesized channels. With a careful examination the synthesized channels, it is possible to show that, for an arbitrarily large number of them, they fit into two sets called good and bad channels. The statistical behavior of them gives the desirable attributes of polar codes [3, 5, 18, 37, 38, 27]: they achieve the capacity when used for transmitting information over a cq channel; they can be encoded efficiently (with a complexity that is essentially linear in the blocklength); the error probability of the decoder decays exponentially in the square root of the blocklength. A descriptive explanation of channel polarization, synthesized channels, and some attributes of polar codes are given below.

Suppose there are N copies of a cq channel W , which we denote by W^N , and N realizations u^N of a random variable U representing the source. The general formulation of polar codes consists of applying a composition function to the input u^N , traditionally represented by a matrix G_N , and using the output of the composition function, defined as x^N , as the actual inputs for the channels W^N . This general procedure is called channel combining, and a scheme is shown in Fig. 1. Observe that this is the same scheme used in any linear channel coding; the type of the composition function is what determines the coding scheme to generate a polar code. For a particular example of coding scheme and composition function, with $N = 2$, see Fig. 2.

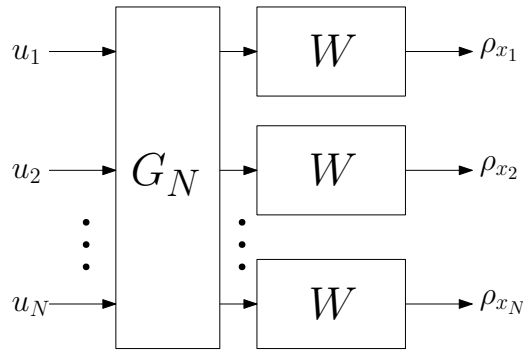


Figure 1: General polar encoding scheme for a classical-quantum channel. The matrix G_N represents the composition function applied on the input u^N and resulting in the vector $x^N = u^N G_N$. The density operator $\rho_{x^N} = \rho_{x_1} \otimes \cdots \otimes \rho_{x_N}$ is the output $W^N(x^N)$.

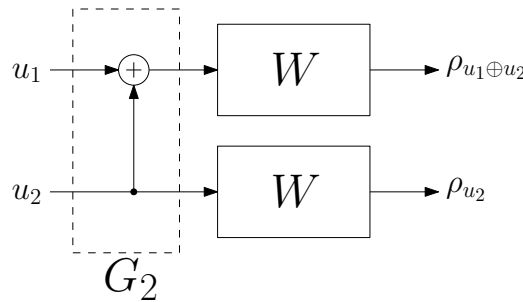


Figure 2: Polar encoding scheme for $N = 2$. The choice of the composition function characterizes the encoding scheme to be polar.

Now, we can introduce channel splitting and the polarization phenomenon that emerges from it. For $i \in \{1, \dots, N\}$, we define the i -th synthesized channel $W_N^{(i)}$ with input alphabet \mathcal{U} and output $\mathcal{D}(\mathcal{H}_{U_1^{i-1}} \mathcal{H}_{B^N})$ as

$$W_N^{(i)}(u_i) = \sum_{u_1^{i-1}} \frac{1}{2^{i-1}} |u_1^{i-1}\rangle\langle u_1^{i-1}|^{U_1^{i-1}} \otimes \bar{\rho}_{u_1^{i-1}}^{B^N}, \quad (2)$$

$$\bar{\rho}_{u_1^{i-1}}^{B^N} = \sum_{u_{i+1}^N} \frac{1}{2^{N-i}} \rho_{u_{i+1}^N}^{B^N} G_N, \quad (3)$$

where N is a power of two and G_N is the composition function. This formulation comes from analyzing a successive decoder acting on the output channel with the help of a genie. For more details, have a look at the original formulation by Arikan [3] or its extension to classical-quantum channels in Ref. [37].

There are two important quantities used to quantify the polarization phenomenon and error probability decay in polar coding. These quantities are called rate and reliability. Their definition depends on the channel under consideration. For a classical-quantum channel W , rate is defined as the mutual information $I(X; B)$ and we denote it as $I(W)$. Reliability is adopted as the fidelity between the possible channel outputs, i.e., reliability is given by $F(\rho_0, \rho_1) = \|\sqrt{\rho_0} \sqrt{\rho_1}\|_1$, where $\|A\|_1$ is the Schatten 1-norm of an operator $A \in \mathcal{L}(\mathcal{H}^B)$. It is possible to show that these quantities are inversely proportional to each other in the sense that when one has a value close to its maximum, the other has a value close to zero [37]. Some interesting results are obtained studying the rate and reliability over channels produced in the channel splitting part. In particular, for N sufficiently large, it is shown that the channels $W_N^{(i)}$ are divided in two sets: one set with $I(W_N^{(i)})$ close to unit, these channels are called ‘‘good’’; and a set with $I(W_N^{(i)})$ close to zero, called ‘‘bad’’ channels. We denote by \mathcal{A} the set with indexes labeling good channels. Analyzing how the fraction of good and bad channels grow when N goes to infinity, it is shown the polarization phenomenon and a capacity-achieving coding strategy for polar codes. See the following proposition for a formal formulation of the exposed ideas.

Proposition 1. [37, Thm. 2, Prop. 4] *Let W be a classical-quantum channel. Then the following is true:*

1. *The channels $\{W_N^{(i)}\}$ polarize in the sense that, for any $\delta \in (0, 1)$, as N goes to infinity through powers of two, the fraction of indexes $i \in \{1, 2, \dots, N\}$ for which $I(W_N^{(i)}) \in (1 - \delta, 1]$ goes to $I(W)$ and the fraction for which $I(W_N^{(i)}) \in [0, \delta)$ goes to $1 - I(W)$;*
2. *For any choice of parameters (N, K, \mathcal{A}) for a classical-quantum polar code, the probability of error is bounded above by*

$$P_e(N, K, \mathcal{A}) \leq 2 \sum_{i \in \mathcal{A}} \sqrt{F(W^{(i),0}, W^{(i),1})}. \quad (4)$$

In particular, for any fixed $R = K/N < I(W)$ and $\beta < 1/2$, block error probability for polar coding under sequential decoding satisfies

$$P_e(N, K) = o(2^{-N^\beta}), \quad (5)$$

where $o(\cdot)$ is the little-O notation from complexity theory.

We make use of Proposition 1 in the following section to show that polar codes can be applied in the task of quantum memory cell discrimination.

2.2 Quantum Memory Cell

We now formulate the general description of quantum memory cell used in this paper. A *quantum memory cell* is defined as the set $\{\mathcal{W}^x\}_{x \in \mathcal{X}}$ of quantum channels. For a fixed x , we have

$$\mathcal{W}^x : \mathcal{D}(\mathcal{H}_{B'}) \rightarrow \mathcal{D}(\mathcal{H}_B) \quad (6)$$

$$\rho \mapsto \mathcal{W}^x(\rho), \quad (7)$$

where $\mathcal{D}(\mathcal{H}_{B'})$, $\mathcal{D}(\mathcal{H}_B)$ are the sets of input and output density states of the quantum channel \mathcal{W}^x . We call $x \in \mathcal{X}$ the *quantum memory cell index*. Sometimes we denote \mathcal{W}^x as $\mathcal{W}_{B' \rightarrow B}^x$ to highlight the input and output systems of \mathcal{W}^x . An important hypothesis is given in here. We are supposing that the distribution of the random variable X describing the label of the quantum channels is non-uniform. This can be justified by the measurement strategy implemented in the decoding process or energy constrain imposing this distribution. Since these two reasoning are plausible in this paper, we are going to adopt, in most parts of this paper, that we are dealing with asymmetric quantum reading. The definition of asymmetric quantum reading goes similarly to the definition of asymmetric cq channels. A quantum memory cell $\{\mathcal{W}^x\}_{x \in \mathcal{X}}$ is called asymmetric if the possible outcomes are non-uniformly distributed over all possible labels $x \in \mathcal{X}$. Lastly, notice that minor changes over our results lead to the applicability of them to symmetric quantum reading.

Remark 2. *Before introducing the channel model used for numerical analysis in this paper, an important point-of-view over quantum memory cell need to be given. A common approach to quantum memory cell is to study how to discriminate their elements using different probe states but having the labels fixed. However, in this paper, we are also interested in improving quantum memory cell discrimination by working on the channel labels. This optimization is performed by fixing the probe states used. Along these lines, quantum memory cell can be seen as a set of classical-quantum channels. Thus, a capacity-achieving approach using polar codes is possible when one deals with quantum memory cell elements having indexes obeying some rule. The rule used in this paper is given by the composition function of polar codes.*

For this paper, we will adopt the *amplitude damping* (AD) channel as our channel model. This is because AD channel is the finite-dimensional first-order approximation of bosonic attenuator channel, which is the standard model for digital memory cell. They can be described in the following manner. Let $\rho \in \mathcal{D}(\mathcal{H})$ be a single-qubit density state, then the AD channel can be described by the following Kraus expression:

$$\mathcal{W}^x(\rho) = A_0 \rho A_0^\dagger + A_1 \rho A_1^\dagger, \quad (8)$$

where

$$A_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-x} \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & \sqrt{x} \\ 0 & 0 \end{pmatrix}. \quad (9)$$

Observe that the same notation for general quantum memory cell and AD channel has been used. However, through the paper will be clear when we are talking about one or the other.

The goal of the following sections is to show that using polar codes it is possible to attain optimal rate with low reliability. This is possible by using synthesized channels formulation and connecting its properties with the original channel under analysis. A method for approaching this is firstly introducing the joint input-output density state and characterizing the channel via this density state. Thus, lets X be a random variable with probability law p_X . We can write the joint density state describing the systems X and B as

$$\rho^{XB} = p_X(0) |0\rangle\langle 0|^X \otimes \mathcal{W}^0(\rho) + p_X(1) |1\rangle\langle 1|^X \otimes \mathcal{W}^1(\rho). \quad (10)$$

As outlined previously, there are two parameters at the center of polarization phenomenon: rate and reliability. Rate is defined in this work as the quantum mutual information between the source X and the output system B :

Definition 3. *Let $X \sim \text{Ber}(p)$ and $\mathcal{W}_{B' \rightarrow B}^x(\rho)$ be a quantum memory cell, where $x \in \mathcal{X}$. The rate of \mathcal{W} is defined as $I(\mathcal{W})_\rho := I(X; B)_\rho$. A direct computation of this quantum mutual information shows that*

$$I(\mathcal{W})_\rho = H\left(p\mathcal{W}^0(\rho) + (1-p)\mathcal{W}^1(\rho)\right) - pH(\mathcal{W}^0(\rho)) - (1-p)H(\mathcal{W}^1(\rho)), \quad (11)$$

where $H(\sigma)$ is the von Neumann entropy for a density operator $\sigma \in \mathcal{D}(\mathcal{H})$.

For reliability of the quantum memory cell \mathcal{W} , it is used the fidelity between the possible output states:

Definition 4. Let $X \sim \text{Ber}(p)$ and $\mathcal{W}_{B' \rightarrow B}^x(\rho)$ be a quantum memory cell, where $x \in \mathcal{X}$. The reliability of the quantum memory cell \mathcal{W} is defined as

$$Z(\mathcal{W})_\rho := 2\sqrt{p(1-p)}F(\mathcal{W}^0(\rho), \mathcal{W}^1(\rho)) = 2\sqrt{p(1-p)}\|\sqrt{\mathcal{W}^0(\rho)}\sqrt{\mathcal{W}^1(\rho)}\|_1, \quad (12)$$

where $F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1$ is the fidelity, and $\|A\|_1$ is the Schatten 1-norm of a operator $A \in \mathcal{L}(\mathcal{H})$.

The following section describes the polar coding scheme proposed in this paper and applies the rate and reliability defined above to quantify the goodness of the codes created.

3 Polar Coding Scheme

In this section, we consider the polarization phenomenon induced by a combining function acting on the quantum memory cells indexes. This approach is similar to the classical-quantum polarization explained in the previous section. However, significant refinements in the arguments and proofs are needed for the results to be valid in our situation.

3.1 Channel Polarization

3.1.1 Channel Combining

As mentioned before, for a fixed quantum probe state, the elements \mathcal{W}^x , for $x \in \mathbb{Z}_2$, of a quantum memory cell can be seen as a classical-quantum channel. Then, without loss of generality, we will treat them in this form through this section. For an illustrative description of our coding scheme in conjunction with our hypothesis, see Fig. 3. Let \mathcal{W} be a classical-quantum channel from which we derive an N -fold classical-quantum channel \mathcal{W}_N recursively, where $N = 2^n$ with $n \in \mathbb{N}_0$. The zeroth level of recursion gives solely the channel $\mathcal{W}_1(u) = \mathcal{W}^u(\rho)$, for all $u \in \mathbb{Z}_2$. The first level is a composition of two zeroth level channels; i.e., the classical-quantum channel \mathcal{W}_2 is given by

$$\mathcal{W}_2(u_1, u_2) = \mathcal{W}_1(u_1 \oplus u_2) \otimes \mathcal{W}_1(u_2) \quad (13)$$

$$= \mathcal{W}^{u_1 \oplus u_2}(\rho) \otimes \mathcal{W}^{u_2}(\rho). \quad (14)$$

This is shown in Fig. 4.

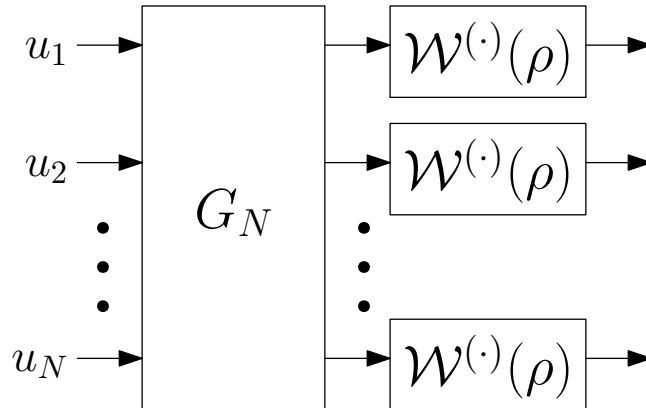


Figure 3: Classical polar coding scheme.

The second level follows from two copies of first level channels. The rule is

$$\mathcal{W}_4(u_1, u_2, u_3, u_4) = \mathcal{W}_2(u_1 \oplus u_2, u_3 \oplus u_4) \otimes \mathcal{W}_2(u_2, u_4) \quad (15)$$

$$= \mathcal{W}^{u_1 \oplus u_2 \oplus u_3 \oplus u_4}(\rho) \otimes \mathcal{W}^{u_3 \oplus u_4}(\rho) \otimes \mathcal{W}^{u_2 \oplus u_4}(\rho) \otimes \mathcal{W}^{u_4}(\rho). \quad (16)$$

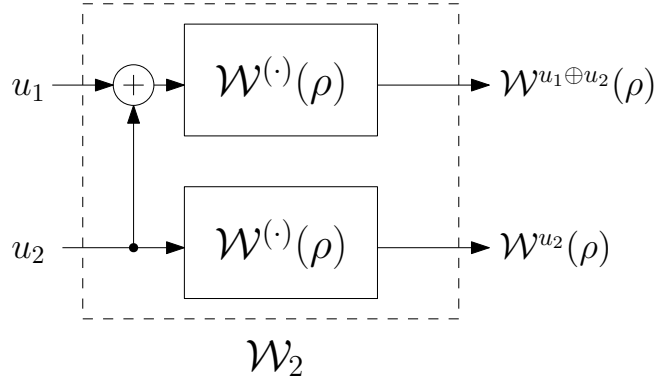


Figure 4: Fundamental polar encoding block

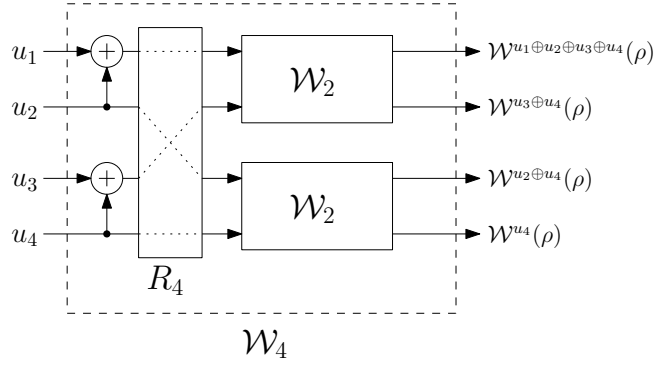


Figure 5: Second level of polar coding scheme.

We depict this scheme in Fig. 5.

Following the same procedure, we can derive the n -level channels. As described in [3], the matrix G_N , which connects the source output u^N to the channel input x^N by $x^N = u^N G_N$, can be expressed as

$$G_N = R_N F^{\otimes n}, \quad (17)$$

with $F^{\otimes n}$ being the n -fold Kronecker product of the matrix

$$F = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad (18)$$

and R_N is a permutation matrix known as *bit-reversal* [3]. In particular, the second level combining function is

$$G_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}. \quad (19)$$

The next step is to characterize the synthesized channels produced by the action of G_N . This is the goal of the next subsection.

3.1.2 Channel Splitting

First of all, consider two realizations of the channel \mathcal{W} . A new input-output relation can be generated through the channel combining procedure described before. These channels are called synthesized since they are not real channels but new point-of-views obtained from the relations created. We

denote this transformation by $(\mathcal{W}, \mathcal{W}) \rightarrow (\mathcal{W}^-, \mathcal{W}^+)$, where \mathcal{W}^- and \mathcal{W}^+ are the synthesized channels. Fixing each input and examining the corresponding outputs, we come to the definition of the synthesized channels below

$$\mathcal{W}^- : u_1 \in \mathbb{Z}_2 \mapsto \mathcal{W}^{-,u_1}(\rho) \in \mathcal{D}(\mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2}), \quad (20a)$$

$$\mathcal{W}^+ : u_2 \in \mathbb{Z}_2 \mapsto \mathcal{W}^{+,u_2}(\rho) \in \mathcal{D}(\mathcal{H}_{U_1} \otimes \mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2}), \quad (20b)$$

with

$$\mathcal{W}^{-,u_1}(\rho) = \sum_{u_2 \in \mathbb{Z}_2} p_U(u_2) \mathcal{W}^{u_1 \oplus u_2}(\rho) \otimes \mathcal{W}^{u_2}(\rho) \quad (20c)$$

and

$$\mathcal{W}^{+,u_2}(\rho) = \sum_{u_1 \in \mathbb{Z}_2} p_U(u_1) |u_1\rangle\langle u_1| \otimes \mathcal{W}^{u_1 \oplus u_2}(\rho) \otimes \mathcal{W}^{u_2}(\rho). \quad (20d)$$

Before presenting channel splitting description for N copies of the channel \mathcal{W} , some properties of the 2-fold case need to be given.

Proposition 5. Consider the transformation $(\mathcal{W}, \mathcal{W}) \rightarrow (\mathcal{W}^-, \mathcal{W}^+)$ for some channels satisfying 20. Then the following rule holds for the rates:

$$I(\mathcal{W}^-)_\rho + I(\mathcal{W}^+)_\rho \leq 2I(\mathcal{W})_\rho \quad (21)$$

$$I(\mathcal{W}^+)_\rho \geq I(\mathcal{W})_\rho. \quad (22)$$

Proof. For the first statement, it is easy to see that

$$I(\mathcal{W}^-)_\rho + I(\mathcal{W}^+)_\rho = I(U_1; B_1 B_2) + I(U_2; B_1 B_2 U_1) \quad (23a)$$

$$\stackrel{(i)}{=} I(U_1; B_1 B_2) + I(U_2; B_1 B_2 | U_1) \quad (23b)$$

$$\stackrel{(ii)}{=} I(X_1 X_2; B_1 B_2) \quad (23c)$$

$$\stackrel{(iii)}{=} I(X_1; B_1) + I(X_2; B_2) = I(X_1; B_1) + I(\mathcal{W}). \quad (23d)$$

The equality in (i) follows from U_1 and U_2 being independent. (ii) is derived from the chain rule for mutual information and the existence of a one-to-one function from U_1, U_2 to X_1, X_2 . Lastly, the independence between X_1 and X_2 is applied in (iii). Now, notice that we are dealing with a (possibly non-uniform) Bernoulli random variable X_1 , which implies $I(X_1; B_1) \neq I(\mathcal{W})$ in general. From $X_1 = U_1 + U_2$ and $X_2 = U_2$, we have that X_1 is $\text{Ber}(p^2 + (1-p)^2)$ and X_2 is $\text{Ber}(p)$. Then, it is possible to bound $I(X_2; B_2) - I(X_1; B_1)$ by the following inequality

$$I(X_2; B_2) - I(X_1; B_1) \geq -\log\{p^2 + (1-p)^2\} - \log\{p(1-p)\} - \log \text{Tr}\{\sqrt{\mathcal{W}^0(\rho)}\sqrt{\mathcal{W}^1(\rho)}\} \\ - 2\sqrt{(1 - F(\mathcal{W}^0(\rho), \mathcal{W}^1(\rho)))^2 2p(1-p)(p^2 + (1-p)^2)}. \quad (24)$$

We have used the bounds

$$I(X_2; B_2) \geq -\log \text{Tr}\{(p_{X_2}(0)\sqrt{\mathcal{W}^0(\rho)} + p_{X_2}(1)\sqrt{\mathcal{W}^1(\rho)})^2\}, \quad (25)$$

from [17, Prop. 1], and

$$I(X_1; B_1) \leq H(\sigma) = h\left(\frac{1}{2}(1 - \sqrt{1 - 4(1 - F(\mathcal{W}^0(\rho), \mathcal{W}^1(\rho)))^2 2p(1-p)(p^2 + (1-p)^2)})\right) \quad (26a)$$

$$\leq 2\sqrt{(1 - F(\mathcal{W}^0(\rho), \mathcal{W}^1(\rho)))^2 2p(1-p)(p^2 + (1-p)^2)}, \quad (26b)$$

where

$$\sigma = \left(\frac{p_{X_2}(0)}{\sqrt{p_{X_2}(0)p_{X_2}(1)F(\mathcal{W}^0(\rho), \mathcal{W}^1(\rho))}} \frac{\sqrt{p_{X_2}(0)p_{X_2}(1)F(\mathcal{W}^0(\rho), \mathcal{W}^1(\rho))}}{p_{X_2}(0)} \right), \quad (27)$$

from [28, Thm. 3]. Calling the RHS of Eq. 24 by $f(p)$ and analyzing its first and second derivative, we can conclude that $f(p) \geq 0$ and, thus, $I(\mathcal{W}^-)_\rho + I(\mathcal{W}^+)_\rho \leq 2I(\mathcal{W})_\rho$.

The second statement is derived from

$$I(\mathcal{W}^+)_\rho = I(U_2; B_1 B_2 U_1) \geq I(U_2; B_2) = I(X_2; B_2) = I(\mathcal{W}). \quad (28)$$

□

Remark 6. Proposition 5 shows that the polarization phenomenon can also be derived in quantum reading problematic. However, it is worth mention the importance of the probe state used during the process. As explicitly shown in Section 4, there are probe states that can polarize “faster” than other, in the sense that $I(\mathcal{W}^+)_\rho - I(\mathcal{W}^-)_\rho > I(\mathcal{W}^+)_\sigma - I(\mathcal{W}^-)_\sigma$ for some probe states ρ and σ . Therefore, even though any probe state generates the polarization phenomenon and, thus, can be used to achieve a nonzero communication rate with arbitrarily low error probability, it may not be optimal. There may exist another probe state that produces the same results but demanding a polar code with lower length.

The following proposition shown in Ref. [27] relates the reliability of the synthesized channels to the reliability of the original one.

Proposition 7. [27, Prop. 4] Consider the transformation $(\mathcal{W}, \mathcal{W}) \rightarrow (\mathcal{W}^-, \mathcal{W}^+)$ for some channels satisfying Eq. 20. Then the following rule holds for the reliability:

$$Z(\mathcal{W}^+)_\rho = Z(\mathcal{W})_\rho^2 \quad (29)$$

$$Z(\mathcal{W}^-)_\rho \leq 2Z(\mathcal{W})_\rho - Z(\mathcal{W})_\rho^2. \quad (30)$$

Now, we can extend the splitting analysis to a N -fold combining of \mathcal{W} . Wilde and Guha have shown how to extend the previous characterization of synthesized channels [37]. We are going to follow a similar path but using a proper description for quantum memory cells. Let $N = 2^n$, where $n \in \mathbb{N}_0$. First of all, in this general case is preferable to label the synthesized channels by natural numbers instead of $\{+, -\}^n$. Let $i \in \{1, 2, \dots, N\}$. The i -th synthesized channel is given by the map

$$\mathcal{W}_N^{(i)}: u_i \mapsto \mathcal{W}^{(i), u_i}(\rho^{\otimes N}), \quad (31)$$

where

$$\mathcal{W}^{(i), u_i}(\rho^{\otimes N}) = \sum_{u_1^{i-1} \in \mathbb{Z}_2^{i-1}} p_{U^{i-1}}(u_1^{i-1}) |u_1^{i-1}\rangle\langle u_1^{i-1}| \otimes \overline{\mathcal{W}}^{u_1^i}(\rho^{\otimes N}), \quad (32)$$

$$\overline{\mathcal{W}}^{u_1^i}(\rho^{\otimes N}) = \sum_{u_{i+1}^N \in \mathbb{Z}_2^{N-i}} p_{U_{i+1}^N}(u_{i+1}^N) \mathcal{W}^{u_{i+1}^N}(\rho^{\otimes N}). \quad (33)$$

From Eq. 32 we see that the description of $\mathcal{W}_N^{(i)}$ supposes the knowledge of the previous input u_1^{i-1} . For a finite-length analysis, this hypothesis may hold with the use of a “genie-aided” successive cancellation decoder similar to Refs. [3, 37]. In the asymptotic analysis, this is not needed.

Having described channel combining and splitting, and given important definitions like the meaning of synthesized channels, our next step is to characterize the behavior of rate, reliability, and error probability in the asymptotic scenario. See the next subsection.

3.1.3 Rate, Reliability, and Error Probability

The first result present in this section describes the connection between rate and reliability. It shows, as expected, that the rate $I(\mathcal{W})_\rho \rightarrow 0$ (or $I(\mathcal{W})_\rho \rightarrow h(p)$) when the reliability $Z(\mathcal{W})_\rho \rightarrow 2\sqrt{p(1-p)}$ (or $Z(\mathcal{W})_\rho \rightarrow 0$).

Proposition 8. Let $X \sim \text{Ber}(p)$ and $\mathcal{W}^x(\rho)$ be a quantum memory cell, where $x \in \mathbb{Z}_2$. Then the following holds

$$I(\mathcal{W})_\rho \geq h(p) - \log(1 + Z(\mathcal{W})_\rho), \quad (34)$$

$$I(\mathcal{W})_\rho \leq \sqrt{4p(1-p) - Z(\mathcal{W})_\rho^2}. \quad (35)$$

Proof. The first inequality follows from [15, Thm. C.1]. For the second one, we need to use the following inequality derived from Theorem 3 in Ref. [28]:

$$I(\mathcal{W})_\rho \leq H(\sigma), \quad (36)$$

where

$$\sigma = \begin{pmatrix} p & \frac{Z(\mathcal{W})_\rho}{2} \\ \frac{Z(\mathcal{W})_\rho}{2} & 1-p \end{pmatrix}. \quad (37)$$

Thus

$$I(\mathcal{W})_\rho \leq h\left(\frac{1}{2}(1 - \sqrt{1 - 4p(1-p) + Z(\mathcal{W})_\rho^2})\right) \quad (38)$$

$$\leq \sqrt{4p(1-p) - Z(\mathcal{W})_\rho^2}. \quad (39)$$

□

Before presenting the next result, we need to introduce a new classical-quantum symmetric channel. Let $\tilde{U}_i \sim \text{Ber}(\frac{1}{2})$, $U_j \sim \text{Ber}(p)$ for all $i, j = 1, \dots, N$. Assume that the output system of the classical-quantum channel $\tilde{\mathcal{W}}$ is given by $\mathcal{D}(\mathcal{H}_{\tilde{Z}} \otimes \mathcal{H}_B)$, where $\sigma \in \mathcal{D}(\mathcal{H}_{\tilde{Z}} \otimes \mathcal{H}_B)$ can be written as

$$\sigma = \sum_{\tilde{z} \in \mathbb{Z}_2} |\tilde{z}\rangle\langle\tilde{z}|^{\tilde{Z}} \otimes \rho^B. \quad (40)$$

Notice that \tilde{Z} represents the classical input system in the same way as the system Z associated with the quantum channel \mathcal{W} . Therefore, the associated quantum systems are the same. We keep different notation here to emphasize that one is associated with a symmetric channel and the other with an asymmetric channel. Now, we can properly describe the classical-quantum channel $\tilde{\mathcal{W}}$. Fix $\rho \in \mathcal{D}(\mathcal{H}_A)$ and let $u \in \mathbb{Z}_2$. Then $\tilde{\mathcal{W}}$ is the map given by

$$\begin{aligned} \tilde{\mathcal{W}}: \tilde{U} &\rightarrow \mathcal{D}(\mathcal{H}_{\tilde{Z}} \otimes \mathcal{H}_B) \\ \tilde{u} &\mapsto \tilde{\mathcal{W}}^{\tilde{u}}(\rho) := \sum_{u \in \mathcal{U}} p_U(u) |\tilde{u} \oplus u\rangle\langle\tilde{u} \oplus u|^{\tilde{Z}} \otimes \mathcal{W}^u(\rho)^B. \end{aligned} \quad (41)$$

It is easy to see that $\tilde{\mathcal{W}}$ is a symmetric classical-quantum channel. Indeed, this follows from the uniform distribution of \tilde{U} and the construction of the $\mathcal{H}_{\tilde{Z}}$ part. The next step is to describe the synthesized channels generated in polar coding on $\tilde{\mathcal{W}}$. Following the characterization given in Eq. 31, we have that

$$\tilde{\mathcal{W}}_N^{(i), \tilde{u}_i}(\rho^N) = \sum_{\tilde{u}_1^{i-1}} \frac{1}{2^{i-1}} |\tilde{u}_1^{i-1}\rangle\langle\tilde{u}_1^{i-1}| \otimes \left(\sum_{u^N} p_{U^N}(u^N) \mathcal{W}^{u^N G_N}(\rho^N) \otimes \left(\sum_{\tilde{u}_{i+1}^N} \frac{1}{2^{N-i}} |(\tilde{u}^N \oplus u^N) G_N\rangle\langle(\tilde{u}^N \oplus u^N) G_N| \right) \right). \quad (42)$$

With these tools, we show in the next proposition that any asymmetric classical-quantum channel can be described via a symmetric one. Moreover, in the following, a relation between the reliabilities of these channels is provided.

Proposition 9. Let $\tilde{\mathcal{W}}_N^{(i)}: \tilde{U}_i \rightarrow \mathcal{D}(\mathcal{H}_{\tilde{U}_1^{i-1}} \otimes \mathcal{H}_{B^N} \otimes \mathcal{H}_{\tilde{Z}^N})$, where $\tilde{U}_i \sim \text{Ber}(\frac{1}{2})$ for all $i = 1, \dots, N$. Then,

$$\mathcal{W}_N^{(i), u_i}(\rho^N) = \text{Tr}_{\tilde{Z}^N} \{ \tilde{\mathcal{W}}_N^{(i), u_i}(\rho^N) |0\rangle\langle 0|^{Z^N} \}. \quad (43)$$

Proof. Let $\tilde{\rho}^{(i)}$ be the joint input-output density operator of $\tilde{\mathcal{W}}_N^{(i), \tilde{u}_i}$. Thus,

$$\tilde{\rho}^{(i)} = \frac{1}{2} \sum_{\tilde{u}_i \in \mathbb{Z}_2} |\tilde{u}_i\rangle\langle\tilde{u}_i| \otimes \tilde{\mathcal{W}}_N^{(i), \tilde{u}_i}(\rho^N) \quad (44a)$$

$$= \sum_{\tilde{u}_1^N} \frac{1}{2^N} |\tilde{u}_1^N\rangle\langle\tilde{u}_1^N| \otimes \left(\sum_{u^N} p_{U^N}(u^N) \mathcal{W}^{u^N G_N}(\rho^N) \otimes |(\tilde{u}^N \oplus u^N) G_N\rangle\langle(\tilde{u}^N \oplus u^N) G_N| \right) \quad (44b)$$

$$= \sum_{u^N} p_{U^N}(u^N) \mathcal{W}^{u^N G_N}(\rho^N) \otimes \left(\sum_{\tilde{u}_1^N} \frac{1}{2^N} |\tilde{u}_1^N\rangle\langle\tilde{u}_1^N| \otimes |(\tilde{u}^N \oplus u^N) G_N\rangle\langle(\tilde{u}^N \oplus u^N) G_N| \right). \quad (44c)$$

Defining $\tilde{z}^N = (\tilde{u}^N \oplus u^N)G_N$, it is possible to rearrange the sums as

$$\begin{aligned}
\tilde{\rho}^{(i)} &= \sum_{u^N} p_{U^N}(u^N) \mathcal{W}^{u^N G_N}(\rho^N) \otimes \left(\sum_{\substack{\tilde{z}^N \\ [u^N \oplus \tilde{z}^N G_N]_1^i = \tilde{u}_1^i}} \frac{1}{2^N} |[u^N \oplus \tilde{z}^N G_N]_1^i\rangle \langle [u^N \oplus \tilde{z}^N G_N]_1^i| \otimes |\tilde{z}^N\rangle \langle \tilde{z}^N| \right) \\
&= \sum_{\tilde{z}^N} \frac{1}{2^N} |\tilde{z}^N\rangle \langle \tilde{z}^N| \otimes \left(\sum_{\tilde{u}_1^i} p_{U_1^i}(\tilde{u}_1^i \oplus [\tilde{z}^N G_N]_1^i) |\tilde{u}_1^i\rangle \langle \tilde{u}_1^i| \otimes \mathcal{W}^{[(\tilde{u}_1^i, 0_{i+1}^N)G_N]_1^i \oplus \tilde{z}_1^i}(\rho^i) \right. \\
&\quad \left. \otimes \left(\sum_{u_{i+1}^N} p_{U_{i+1}^N}(u_{i+1}^N) \mathcal{W}^{[(0_1^i, u_{i+1}^N)G_N]_{i+1}^N}(\rho^{N-i}) \right) \right). \tag{45}
\end{aligned}$$

Now, we can see that

$$\text{Tr}_{\tilde{z}^N} \{ \tilde{\rho}^{(i)} |0\rangle \langle 0| \} = \sum_{u_1^i} p_{U_1^i}(u_1^i) |u_1^i\rangle \langle u_1^i| \otimes \left(\sum_{u_{i+1}^N} p_{U_{i+1}^N}(u_{i+1}^N) \mathcal{W}^{u^N G_N}(\rho^N) \right), \tag{46}$$

which is the joint input-output state of $\mathcal{W}_N^{(i), u_i}(\rho^N)$. \square

Proposition 10. Let $\mathcal{W}_N^{(i)}$ and $\tilde{\mathcal{W}}_N^{(i)}$ be the synthesized quantum channels described in Eq. 31 and Proposition 9, respectively. Then

$$Z(\tilde{\mathcal{W}}_N^{(i)})_\rho = Z(\mathcal{W}_N^{(i)})_\rho. \tag{47}$$

Proof. From the definition of reliability for symmetric classical-quantum channels in Ref. [37] and Eq 45, we have that

$$\begin{aligned}
Z(\tilde{\mathcal{W}}_N^{(i)})_\rho &= F(\tilde{\mathcal{W}}_N^{(i),0}(\rho^N), \tilde{\mathcal{W}}_N^{(i),1}(\rho^N)) \\
&= \sum_{\tilde{z}^N} \sum_{\tilde{u}_1^{i-1}} \frac{p_{U_1^{i-1}}(\tilde{u}_1^{i-1} \oplus [\tilde{z}^N G_N]_1^{i-1}) \sqrt{p_U(0 \oplus [\tilde{z}^N G_N]_i) p_U(1 \oplus [\tilde{z}^N G_N]_i)}}{2^{N-1}} F(w_0^{(i)}, w_1^{(i)}) \tag{49}
\end{aligned}$$

where

$$w_0^{(i)} = \mathcal{W}^{[(\tilde{u}_1^{i-1}, 0_{i+1}^N)G_N]_1^i \oplus \tilde{z}_1^i}(\rho^i) \sum_{u_{i+1}^N} p_{U_{i+1}^N}(u_{i+1}^N) \mathcal{W}^{[(0_1^i, u_{i+1}^N)G_N]_{i+1}^N}(\rho^{N-i}), \tag{50}$$

$$w_1^{(i)} = \mathcal{W}^{[(\tilde{u}_1^{i-1}, 1_{i+1}^N)G_N]_1^i \oplus \tilde{z}_1^i}(\rho^i) \sum_{u_{i+1}^N} p_{U_{i+1}^N}(u_{i+1}^N) \mathcal{W}^{[(0_1^i, u_{i+1}^N)G_N]_{i+1}^N}(\rho^{N-i}). \tag{51}$$

Let $z^N = \tilde{z}^N G_N$ and $u_1^{i-1} = \tilde{u}_1^{i-1} \oplus [\tilde{z}^N G_N]_1^{i-1} = \tilde{u}_1^{i-1} \oplus z_1^{i-1}$. Since the map $(\tilde{z}_1^N, \tilde{u}_1^{i-1}) \mapsto (z_1^N, u_1^{i-1})$ is a bijection, we see that

$$\begin{aligned}
Z(\tilde{\mathcal{W}}_N^{(i)})_\rho &= \sum_{z^N} \sum_{u_1^{i-1}} \frac{p_{U_1^{i-1}}(u_1^{i-1}) \sqrt{p_U(0) p_U(1)}}{2^{N-1}} F(\mathcal{W}^{(u_1^{i-1}, 0)}(\rho^i) \sum_{u_{i+1}^N} p_{U_{i+1}^N}(u_{i+1}^N) \mathcal{W}^{[(0_1^i, u_{i+1}^N)G_N]_{i+1}^N}(\rho^{N-i}), \\
&\quad \mathcal{W}^{(u_1^{i-1}, 1)}(\rho^i) \sum_{u_{i+1}^N} p_{U_{i+1}^N}(u_{i+1}^N) \mathcal{W}^{[(0_1^i, u_{i+1}^N)G_N]_{i+1}^N}(\rho^{N-i})) \\
&= 2\sqrt{p_U(0)p_U(1)} \sum_{u_1^{i-1}} p_{U_1^{i-1}}(u_1^{i-1}) F(\overline{\mathcal{W}}^{(u_1^{i-1}, 0)}(\rho^{\otimes N}), \overline{\mathcal{W}}^{(u_1^{i-1}, 1)}(\rho^{\otimes N})) \\
&= 2\sqrt{p_U(0)p_U(1)} F(\mathcal{W}^{(i),0}(\rho^{\otimes N}), \mathcal{W}^{(i),1}(\rho^{\otimes N})) \\
&= Z(\mathcal{W}_N^{(i)})_\rho, \tag{52}
\end{aligned}$$

where $\overline{\mathcal{W}}^{u_1^i}(\rho^{\otimes N})$ is defined in Eq. 32. \square

The next theorem shows that the asymptotically fraction of good channels is equals to the mutual information between the classical system X and the quantum system B .

Theorem 11. Let $\mathcal{W}_N^{(i)}$ be the synthesized quantum channels described in Eq. 31 and $Z(X|Y)$ be the reliability of two Bernoulli random variables X and Y ; i.e., $Z(X|Y) = 2 \sum_y \sqrt{p_{X,Y}(0,y)p_{X,Y}(1,y)}$. For every $\beta < 1/2$, we have

$$\lim_{n \rightarrow \infty} \frac{1}{2^n} |\{i \in \mathbb{Z}_2^n : Z(\mathcal{W}_N^{(i)})_\rho \leq 2^{-2^{n\beta}} \text{ and } Z(U_i|U_1^{i-1}) \geq 1 - 2^{-2^{n\beta}}\}| = I(X; B)_\rho, \quad (53)$$

$$\lim_{n \rightarrow \infty} \frac{1}{2^n} |\{i \in \mathbb{Z}_2^n : Z(\mathcal{W}_N^{(i)})_\rho \geq 1 - 2^{-2^{n\beta}} \text{ and } Z(U_i|U_1^{i-1}) \leq 2^{-2^{n\beta}}\}| = 1 - I(X; B)_\rho, \quad (54)$$

Proof. Firstly, notice that

$$\begin{aligned} I(\tilde{\mathcal{W}}) &= I(\tilde{X}; \tilde{X} \oplus X, B) \\ &= H(\tilde{X} \oplus X, B) - H(\tilde{X} \oplus X, B|\tilde{X}) \\ &= 1 + H(B) - H(X, B) = 1 - H(X|B). \end{aligned} \quad (55)$$

Applying Eq. 55 to Proposition 1, we can deduce that

$$\lim_{n \rightarrow \infty} \frac{1}{2^n} |\{i \in \mathbb{Z}_2^n : Z(\mathcal{W}_N^{(i)})_\rho \leq 2^{-2^{n\beta}}\}| = 1 - H(X|B)_\rho, \quad (56a)$$

$$\lim_{n \rightarrow \infty} \frac{1}{2^n} |\{i \in \mathbb{Z}_2^n : Z(\mathcal{W}_N^{(i)})_\rho \geq 1 - 2^{-2^{n\beta}}\}| = H(X|B)_\rho. \quad (56b)$$

Additionally, it is possible to derive $Z(\mathcal{W}_N^{(i)})_\rho = Z(U_i|U_1^{i-1})$ and $H(X|B) = H(X)$ if the output probe state of \mathcal{W} is independent of X . Thus, a similar result to Eq. 56 is derived

$$\lim_{n \rightarrow \infty} \frac{1}{2^n} |\{i \in \mathbb{Z}_2^n : Z(U_i|U_1^{i-1}) \leq 2^{-2^{n\beta}}\}| = 1 - H(X), \quad (57a)$$

$$\lim_{n \rightarrow \infty} \frac{1}{2^n} |\{i \in \mathbb{Z}_2^n : Z(U_i|U_1^{i-1}) \geq 1 - 2^{-2^{n\beta}}\}| = H(X). \quad (57b)$$

Now, define A, B, C , and D as the sets

$$A = \{i : Z(\mathcal{W}_N^{(i)})_\rho \leq 2^{-2^{n\beta}}\}, \quad (58a)$$

$$B = \{i : Z(\mathcal{W}_N^{(i)})_\rho \geq 1 - 2^{-2^{n\beta}}\}, \quad (58b)$$

$$C = \{i : Z(U_i|U_1^{i-1}) \leq 2^{-2^{n\beta}}\}, \quad (58c)$$

$$D = \{i : Z(U_i|U_1^{i-1}) \geq 1 - 2^{-2^{n\beta}}\}. \quad (58d)$$

From Proposition 8 and $H(U_i|U_1^{i-1}, B^N) \leq H(U_i|U_1^{i-1})$, it is possible to see that $B \cap C$ is empty for sufficiently large n . Furthermore,

$$\lim_{n \rightarrow \infty} \frac{|A| + |B|}{2^n} = \lim_{n \rightarrow \infty} \frac{|C| + |D|}{2^n} = 1. \quad (59)$$

Therefore, the claim is derived from

$$\lim_{n \rightarrow \infty} \frac{|B \cup C|}{2^n} = \lim_{n \rightarrow \infty} \frac{|B| + |C|}{2^n} = 1 - I(X; B)_\rho. \quad (60)$$

and

$$\lim_{n \rightarrow \infty} \frac{|A \cap D|}{2^n} = 1 - \lim_{n \rightarrow \infty} \frac{|B \cup C|}{2^n} = I(X; B)_\rho. \quad (61)$$

□

As can be seen in the theorem statement and elaborated in the proof, we had to impose an additional constraint on the reliability of U_i given the previous U_1^{i-1} in order to derive our result. This is because we are dealing with asymmetric quantum reading. For symmetric quantum reading, the constrain is not needed.

Remark 12. Subsections 3.1.1 and 3.1.2, and Proposition 8 are easily applied to symmetric quantum channels discrimination, having it as a particular case when $p_X(0) = p_X(1) = \frac{1}{2}$. For the result in Theorem 11, there is no need to introduce the new type of channel and, furthermore, the asymptotic analysis of the rate of good and bad channels does not impose anything on $Z(U_i|U_1^{i-1})$. The symmetric quantum reading treatment of Theorem 11 goes similarly to Section IV of Ref. [37].

3.2 Polar Coding and Decoding Scheme

The encoding protocol of polar codes consists of setting up the labels for information bits and frozen bits. The former label is denoted by \mathcal{A} and the later by \mathcal{A}^c . Thus, we use bits $u_{\mathcal{A}} = \{u_i\}_{i \in \mathcal{A}}$ to transmit information. On the frozen bits, they can be fixed for the whole transmission or can depend on the previous u_1^{i-1} bits. For asymmetric channels, as our case, the latter strategy is more suitable. See Ref. [18] for more explanations.

The construction of a codeword is done as follows. The source generates a uniform sequence $u_1^{|\mathcal{A}|}$. Next, the encoder determines the value u_i , $i \in \mathcal{A}^c$, of the frozen bits in the ascending order by $u_i = \lambda_i(u_1^{i-1})$, where λ_i is a function from $\{0, 1\}^{i-1}$ to $\{0, 1\}$. Putting the sequence $u_1^{|\mathcal{A}|}$ in the information bits $u_{\mathcal{A}}$ and the frozen bits in the remaining coordinates, we have the vector u^N . Now, the codeword is given by $x^N = u^N G_N$. It is clear that the code rate is $R = |\mathcal{A}|/N$. It remains to describe how the set \mathcal{A} is determined and which functions λ_i do we use.

Determining the set of information bits is crucial for the performance of a polar code. Firstly, for a sufficiently large code length N , we choose the cardinality of \mathcal{A} so that the code rate $R < I(\mathcal{W})_{\rho}$. Next, we have to choose coordinates to constitute the set \mathcal{A} . A common approach is to select the coordinates with smallest reliabilities. Formally, this selection goes as follows. Let \mathcal{A} be the set that for any $j \in \mathcal{A}^c$ we have $Z(\mathcal{W}_N^{(j)}) \geq Z(\mathcal{W}_N^{(i)})$ for all $i \in \mathcal{A}$. Additionally, because the construction under consideration is for asymmetric channels, we also impose $Z(U_i|U_1^{i-1})$, for all $i \in \mathcal{A}$, to be large when compared with the elements in \mathcal{A}^c . Theorem 11 makes use of this constrain to characterize the asymptotic rates in polar coding. If any of these two constrains is not satisfied, we decrease the cardinality of \mathcal{A} down to when they are satisfied.

The functions λ_i used in this paper is such that optimize the probability of frozen bit output. Let $\Lambda_{\mathcal{A}^c} = \{\Lambda_i\}_{i \in \mathcal{A}^c}$ be random variables which are independent of each other and input and output systems, and satisfy

$$p_{\Lambda_i}[\Lambda_i(u_1^{i-1})] = p_{U_i|U_1^{i-1}}(1|u_1^{i-1}), \quad (62)$$

for all $u_1^{i-1} \in \{0, 1\}^{i-1}$. Then, λ_i is a realization of the random variable Λ_i . There are practical methods to generate in practice the functions λ_i using pseudorandom number generators [18], but there is no need to address it in this paper.

Considering symmetric channel discrimination, a similar encoding scheme can be proposed. First of all, the algorithm for choosing the frozen bits is exactly the one presented here without the dependence of the previous bits. Thus, the encoding map is a realization of the random variable $\tilde{\Lambda}_i$, where

$$p_{\tilde{\Lambda}_i}[\tilde{\Lambda}_i = i] = \frac{1}{2}, \quad (63)$$

for $i = 1, 2$. Secondly, the proposed strategy for defining the set \mathcal{A} is applicable for symmetric channels discrimination. We only need to drop the constrain on $Z(U_i|U_1^{i-1})$. Approximation techniques in symmetric polar codes [34] can compose the decoding scheme in order to derive a faster encoding scheme.

Now we describe the decoding process. Suppose the sequential decoder has obtained, up to this moment, the vector \hat{u}_1^{i-1} and plans to obtain \hat{u}_i . Then the decoding process divides into two

cases. For the coordinates in \mathcal{A}^c , we apply the inverse encoding function depending on the previous coordinates. Namely, we employ

$$\hat{u}_i = \lambda_i^{-1}(\hat{u}_1^{i-1}). \quad (64)$$

Notice that no measurement is implemented in this step and, for sufficiently large N , the error probability is arbitrarily low. See Theorem 11 for the proof. Next, we deal with the information bits. Quantum successive cancellation decoder is used in here. First of all, this makes our decoding strategy and analysis totally different from any decoding method used for classical polar codes. Secondly, because of the constructive approach adopted in here, with measurements created for our specific task, the polar decoding strategy differs from previous works on quantum polar codes.

For characterizing the error probability decay of the polar codes constructed, we firstly show the existence of “pretty good measurements” design to decode quantum memory cell and having desirable error probabilities in Proposition 13.

Proposition 13. *Let $\mathcal{W}: x \in \mathcal{X} \rightarrow \rho_x^{UB} \in \mathcal{D}(\mathcal{H}_{UB})$ be a cq channel such that*

$$\rho_x^{UB} = \sum_{u \in \mathcal{U}} p_U(u) |u\rangle\langle u|^U \otimes \rho_{x,u}^B, \quad (65)$$

where X is a discrete random variable, and $\{|u\rangle\}_{u \in \mathcal{U}}$ is an orthonormal basis of a finite-dimensional Hilbert space. Then, there exists a POVM $\{\Lambda_x^{UB}\}_{x \in \mathcal{X}}$ satisfying

$$1 - \sum_{x \in \mathcal{X}} p_X(x) \text{Tr}\{\Lambda_x \rho_x\} < \frac{1}{2} Z(\mathcal{W}). \quad (66)$$

Proof. To derive our claim, we need to invoke a result from Barnum and Knill [7]. Let $p_U(u)$ be the probability of a system to be found in the state ρ_u , for $u \in \mathcal{U}$, then there exists a POVM $\{\Lambda_u^*\}_{u \in \mathcal{U}}$ such that the average success probability is lower bounded as

$$P_{succ} = \sum_{u \in \mathcal{U}} p_U(u) \text{Tr}\{\rho_u^{\otimes M} \Lambda_u^*\} \geq 1 - \sum_{u \neq v} \sqrt{p_U(u)p_U(v)} F(\rho_u^M, \rho_v^M). \quad (67)$$

In particular, there exists a POVM $\{\Lambda_u\}_{u \in \mathcal{U}}$ such that the probability of error follows the following inequality

$$P_{err} = 1 - \sum_{u \in \mathcal{U}} p_U(u) \text{Tr}\{\rho_u \Lambda_u\} \leq \sum_{u \neq v} \sqrt{p_U(u)p_U(v)} F(\rho_u, \rho_v). \quad (68)$$

Now, we return to our proof. Assume that $\mathcal{W}^u: x \rightarrow \rho_{x,u}$ is a cq channel, where $x \in \mathcal{X}$ is a realization of a discrete random variable X . The result in Eq. 68 says that there exist POVMs $\{\Lambda_{x,u}\}_{x \in \mathcal{X}}$ satisfying

$$1 - \sum_{x \in \mathcal{X}} p_X(x) \text{Tr}\{\rho_{x,u} \Lambda_{x,u}\} < \frac{1}{2} Z(\mathcal{W}^u). \quad (69)$$

Defining, for every $u \in \mathcal{U}$, the POVM

$$\Lambda_x := \sum_{u \in \mathcal{U}} \Lambda_{x,u} \otimes |u\rangle\langle u|, \quad (70)$$

we have that

$$1 - \sum_{x \in \mathcal{X}} p_X(x) \text{Tr}\{\Lambda_x \rho_x\} \stackrel{(i)}{=} 1 - \sum_{x \in \mathcal{X}} p_X(x) \sum_{u \in \mathcal{U}} p_U(u) \text{Tr}\{\Lambda_{x,u} \rho_{x,u}\} \quad (71)$$

$$= \sum_{u \in \mathcal{U}} p_U(u) \sum_{x \in \mathcal{X}} p_X(x) \left(1 - \text{Tr}\{\Lambda_{x,u} \rho_{x,u}\}\right) \quad (72)$$

$$\stackrel{(ii)}{<} \sum_{u \in \mathcal{U}} p_U(u) \frac{1}{2} Z(\mathcal{W}^u) \quad (73)$$

$$\stackrel{(iii)}{=} \sum_{\substack{x, x' \in \mathcal{X} \\ x \neq x'}} \sqrt{p_X(x) p_X(x')} F \left(\sum_{u \in \mathcal{U}} p_U(u) |u\rangle\langle u| \otimes \rho_{x,u}, \sum_{u \in \mathcal{U}} p_U(u) |u\rangle\langle u| \otimes \rho_{x',u} \right) \quad (74)$$

$$\stackrel{(iv)}{=} \sum_{\substack{x, x' \in \mathcal{X} \\ x \neq x'}} \sqrt{p_X(x) p_X(x')} F(\rho_x, \rho_{x'}) \quad (75)$$

$$= \frac{1}{2} Z(\mathcal{W}). \quad (76)$$

The equality in (i) follows from the definitions of ρ_x and Λ_x in Eq. 65 and 70, respectively. The inequality in Eq. 69 is used to obtain (ii). The item (iii) is just the computation of the reliability for \mathcal{W}^u . Lastly, we use in (iv) the fact that $F(\sum_x p_X(x) |x\rangle\langle x| \otimes \sigma_x, \sum_x q_X(x) |x\rangle\langle x| \otimes \tau_x) = \sum_x \sqrt{p_X(x) q_X(x)} F(\sigma_x, \tau_x)$. \square

Analyzing a quantum successive cancellation decoder employing the measurements in Proposition 13 needs the use of a quantum union bound. Lemma 14 introduces the one used in this paper. As it is shown in the proof of Theorem 15, the inequality in Lemma 14 allows us to derive a bound for the error probability in accordance to what is expected using polar codes.

Lemma 14. [21, Lemma 4.1] *Let ρ be a positive semi-definite operator acting on a separable Hilbert space \mathcal{H}_B , let $\{\Lambda_i\}_{i=1}^L$ denote a set of positive semi-definite operators such that $0 \leq \Lambda_i \leq I$ for all $i \in \{1, \dots, L\}$, and let $c > 0$. Then the following quantum union bound holds*

$$\begin{aligned} \text{Tr}\{\rho\} - \text{Tr}\{\Pi_{\Lambda_L} \cdots \Pi_{\Lambda_1} (\rho \otimes |\bar{0}\rangle\langle\bar{0}|_{PL}) \Pi_{\Lambda_1} \cdots \Pi_{\Lambda_L}\} &\leq (1+c) \text{Tr}\{(I - \Lambda_L)\rho\} \\ &+ (2+c+c^{-1}) \sum_{i=2}^{L-1} \text{Tr}\{(I - \Lambda_i)\rho\} + (2+c^{-1}) \text{Tr}\{(I - \Lambda_1)\rho\}, \end{aligned} \quad (77)$$

where $|\bar{0}_{PL}\rangle \equiv |0_{P_1}\rangle \otimes \cdots \otimes |0_{P_L}\rangle$ is an auxiliary state of L qubit probe systems and Π_{Λ_i} is a projector defined as $\Pi_{\Lambda_i} := U_i^\dagger P_i U_i$ for some unitary U_i and projector P_i such that $\text{Tr}\{\Pi_{\Lambda_i} (\rho \otimes |\bar{0}\rangle\langle\bar{0}|_{PL})\} = \text{Tr}\{\Lambda_i \rho\}$. In particular, for any set of positive semi-definite operators, there exists a set of projectors satisfying Eq. 77.

An important attribute of polar codes is the error probability decay when the length of the code grows. We show in Theorem 15 that error probability decays exponentially fast in the code length. This result motivates the use of polar codes in practical protocols devoted to discriminate quantum memory cell.

Theorem 15. *Let $\{\mathcal{W}^x\}_{x \in \mathcal{X}}$ be a quantum memory cell and $\gamma \in \mathbb{R}$ a positive constant. Then there exists a polar code with parameters (N, K, \mathcal{A}) such that the error probability is bounded above by*

$$P_e(N, K, \mathcal{A}) \leq \gamma \sum_{i \in \mathcal{A}} Z_\rho(\mathcal{W}^{(i)}). \quad (78)$$

In particular, for any fixed $R = K/N < I(W)$ and $\beta < 1/2$, block error probability for polar coding under sequential decoding satisfies

$$P_e(N, K) = o(2^{-N^\beta}), \quad (79)$$

where $o(\cdot)$ is the little-O notation from complexity theory.

Proof. First of all, suppose we have access to an auxiliary system $|\bar{0}_{P_N}\rangle = |0_{P_1}\rangle \otimes \cdots \otimes |0_{P_N}\rangle$. Lemma 14 guarantees the existence of projective measurements obtained by extending the pretty good measurements of Proposition 13. We denote the projective and pretty good measurements by the sets $\{\Pi_{\Lambda_i}\}_{i=1}^N$ and $\{\Lambda_i\}_{i=1}^N$, respectively. Then, using these projective measurements on a quantum successive cancellation strategy for the information bits, a bound for the error probability is obtained as follows

$$\begin{aligned}
P_e(N, K, \mathcal{A}) &= 1 - \sum_{u^N} p_{U^N}(u^N) \text{Tr}\{\Pi_{\Lambda_{u^N}} \mathcal{W}^{u^N}(\rho^N) \otimes |\bar{0}_{P_N}\rangle\langle\bar{0}_{P_N}|\} \\
&= 1 - \sum_{u^N} p_{U^N}(u^N) \text{Tr}\{\Pi_{\Lambda_{u_1^{N-1}u_N}}^{B^N} \cdots \Pi_{\Lambda_{u_1}}^{B^N} \mathcal{W}^{u^N}(\rho^N) \otimes |\bar{0}_{P_N}\rangle\langle\bar{0}_{P_N}|\} \\
&\leq \sum_{u^N} p_{U^N}(u^N) \left((1+c) \text{Tr}\left\{ (I - \Lambda_{u_1^{N-1}u_N}^{B^N}) \mathcal{W}^{u^N}(\rho^N) \right\} \right. \\
&\quad + (2+c+c^{-1}) \sum_{i=2}^{N-1} \text{Tr}\left\{ (I - \Lambda_{u_1^{i-1}u_i}^{B^N}) \mathcal{W}^{u^N}(\rho^N) \right\} \\
&\quad \left. + (2+c^{-1}) \text{Tr}\left\{ (I - \Lambda_{u_1}^{B^N}) \mathcal{W}^{u^N}(\rho^N) \right\} \right), \tag{80}
\end{aligned}$$

where the second inequality is obtained from the definition of a quantum successive cancellation decoder and the inequality from Lemma 14. Distributing and rearranging the sums, the following equality holds

$$\begin{aligned}
\text{RHS of Eq. 82} &= (1+c) \sum_{u^N} p_{U^N}(u^N) \text{Tr}\left\{ (I - \Lambda_{u_1^{N-1}u_N}^{B^N}) \mathcal{W}^{u^N}(\rho^N) \right\} \\
&\quad + (2+c+c^{-1}) \sum_{i=2}^{N-1} p_U(u_i) \sum_{u_1^{i-1}} p_{U_1^{i-1}}(u_1^{i-1}) \text{Tr}\left\{ (I - \Lambda_{u_1^{i-1}u_i}^{B^N}) \sum_{u_1^{i-1}} p_{U_{i+1}^N}(u_{i+1}^N) \mathcal{W}^{u^N}(\rho^N) \right\} \\
&\quad + (2+c^{-1}) \sum_{u_1} p_U(u_1) \text{Tr}\left\{ (I - \Lambda_{u_1}^{B^N}) \sum_{u_2^{i-1}} p_{U_2^N}(u_2^N) \mathcal{W}^{u^N}(\rho^N) \right\}. \tag{83}
\end{aligned}$$

Continuing,

$$\begin{aligned}
\text{RHS of Eq. 83} &= (1+c) \sum_{u^N} p_{U^N}(u^N) \text{Tr}\left\{ (I - \Lambda_{u_1^{N-1}u_N}^{B^N}) \bar{\mathcal{W}}^{u_1^N}(\rho^N) \right\} \\
&\quad + (2+c+c^{-1}) \sum_{i=2}^{N-1} p_U(u_i) \sum_{u_1^{i-1}} p_{U_1^{i-1}}(u_1^{i-1}) \text{Tr}\left\{ (I - \Lambda_{u_1^{i-1}u_i}^{B^N}) \bar{\mathcal{W}}^{u_1^i}(\rho^N) \right\} \\
&\quad + (2+c^{-1}) \sum_{u_1} p_U(u_1) \text{Tr}\left\{ (I - \Lambda_{u_1}^{B^N}) \bar{\mathcal{W}}^{u_1}(\rho^N) \right\} \tag{84}
\end{aligned}$$

$$\begin{aligned}
&= (1+c) \sum_{u^N} p_{U^N}(u^N) \text{Tr}\left\{ \left(\sum_{u_1^{N-1}} |u_1^{N-1}\rangle\langle u_1^{N-1}| \otimes (I - \Lambda_{u_1^{N-1}u_N}^{B^N}) \right) \right. \\
&\quad \left. \left(\sum_{u_1^{N-1}} p_{U_1^{N-1}}(u_1^{N-1}) |u_1^{N-1}\rangle\langle u_1^{N-1}| \otimes \bar{\mathcal{W}}^{u_1^N}(\rho^N) \right) \right\} \\
&\quad + (2+c+c^{-1}) \sum_{i=2}^{N-1} p_U(u_i) \text{Tr}\left\{ \left(\sum_{u_1^{i-1}} |u_1^{i-1}\rangle\langle u_1^{i-1}| \otimes (I - \Lambda_{u_1^{i-1}u_i}^{B^N}) \right) \right. \\
&\quad \left. \left(\sum_{u_1^{i-1}} p_{U_1^{i-1}}(u_1^{i-1}) |u_1^{i-1}\rangle\langle u_1^{i-1}| \otimes \bar{\mathcal{W}}^{u_1^i}(\rho^N) \right) \right\} \\
&\quad + (2+c^{-1}) \sum_{u_1} p_U(u_1) \text{Tr}\left\{ (I - \Lambda_{u_1}^{B^N}) \bar{\mathcal{W}}^{u_1}(\rho^N) \right\}, \tag{85}
\end{aligned}$$

where the first equality follows from Eq. 33. For the second equality, we have used the fact that [37]

$$\sum_x p_X(x) \text{Tr}\{A_x \rho_x\} = \text{Tr}\left\{\left(\sum_x |x\rangle\langle x| \otimes A_x\right)\left(\sum_{x'} p_X(x) |x'\rangle\langle x'| \otimes \rho_{x'}\right)\right\}. \quad (86)$$

Now, from the definition of synthesized channels present in Eq. 32 the next equality is produced

$$\begin{aligned} \text{RHS of Eq. 85} &= (1+c) \sum_{u^N} p_{U^N}(u^N) \text{Tr}\left\{\left(\sum_{u_1^{N-1}} |u_1^{N-1}\rangle\langle u_1^{N-1}| \otimes (I - \Lambda_{u_1^{N-1}u_N}^{B^N})\right) \mathcal{W}^{(N),u_N}(\rho^N)\right\} \\ &+ (2+c+c^{-1}) \sum_{i=2}^{N-1} p_U(u_i) \text{Tr}\left\{\left(\sum_{u_1^{i-1}} |u_1^{i-1}\rangle\langle u_1^{i-1}| \otimes (I - \Lambda_{u_1^{i-1}u_i}^{B^N})\right) \mathcal{W}^{(i),u_i}(\rho^N)\right\} \\ &+ (2+c^{-1}) \sum_{u_1} p_U(u_1) \text{Tr}\left\{(I - \Lambda_{u_1}^{B^N}) \mathcal{W}^{(1),u_1}(\rho^N)\right\}. \end{aligned} \quad (87)$$

Now, the number of terms in the previous sum can be reduced by means of a simple observation. Since the frozen bits are obtained by using a classical function, there is no need to make a measurement on the frozen bits position; i.e., $\Lambda_{u_1^{i-1}u_i}^{B^N} = I$ for any $i \in \mathcal{A}^c$. This leads to

$$\text{RHS of Eq. 87} < \frac{(1+c+c^{-1})}{2} \sum_{i \in \mathcal{A}} Z_\rho(\mathcal{W}^{(i)}) = o(2^{-N^\beta}), \quad (88)$$

where the definition of reliability for synthesized channels have been applied. The fact that this sum equals to $o(2^{-N^\beta})$ is a consequence of Proposition 1. \square

When one considers this decoding scheme for symmetric quantum reading, the same result as in Theorem 15 is obtained. Thus, the strategy used is the same, and there is only the need to properly design the pretty good measurement for the symmetric channels under consideration.

4 Probe State Analysis

We are going to analyze how reliability and rate depend on the probe state used. The comparison will be between pure and entangled probe states. In both cases, a parameter of the probe states is maximized to give the optimal value of the quantity under consideration.

First of all, we adopt

$$|\psi\rangle = \sqrt{1-q}|0\rangle + e^{-i\phi}\sqrt{q}|1\rangle \quad (89)$$

as the pure probe state, where $q \in [0, 1]$ and $\phi \in [0, 2\pi)$. Because of the symmetric action of the AD channel with respect to the z -axis in the Bloch sphere representation, we can assume, without loss of generality, that $\phi = 0$. Secondly, the entangled state used as probe is given by

$$|\Psi^{xz}\rangle = (Z^z X^x \otimes I) |\Phi\rangle, \quad (90)$$

where $z, x \in \mathbb{Z}_2$, and $|\Phi\rangle = \sqrt{1-q}|00\rangle + \sqrt{q}|11\rangle$, with $q \in [0, 1]$. Notice that for $q = 1/2$, we have all four Bell states. Using these pure and entangled states, we can compare their goodness for quantum reading assisted by polar codes. This comparison is made by optimizing the quantity under consideration with respect to q and q .

Lets start considering pure probe states. Substituting Eq. 89 into Eq. 8, it is possible to see that

$$\mathcal{W}^{u_0}(|\psi\rangle\langle\psi|) = \begin{pmatrix} 1 - (1-u_0)q & \sqrt{(1-u_0)q(q-1)} \\ \sqrt{(1-u_0)q(q-1)} & (1-u_0)q \end{pmatrix}, \quad (91)$$

$$\mathcal{W}^{u_1}(|\psi\rangle\langle\psi|) = \begin{pmatrix} 1 - (1-u_1)q & \sqrt{(1-u_1)q(q-1)} \\ \sqrt{(1-u_1)q(q-1)} & (1-u_1)q \end{pmatrix}. \quad (92)$$

Now, the minimal value of reliability with respect to q can be computed. We are going to study a normalized version of reliability given by

$$\bar{Z}(\mathcal{W})_{|\psi\rangle\langle\psi|} = \max_q \frac{Z(\mathcal{W})_{|\psi\rangle\langle\psi|}}{2\sqrt{p(p-1)}} = \max_q F(\mathcal{W}^{u_0}(|\psi\rangle\langle\psi|), \mathcal{W}^{u_1}(|\psi\rangle\langle\psi|)). \quad (93)$$

For the current probe state, it is shown in Fig. 6 the minimal values of $\bar{Z}(\mathcal{W})_{|\psi\rangle\langle\psi|}$ and the corresponding q . An important fact can be seen in Fig. 6, the optimal pure probe state is given by the excited state $|1\rangle$.

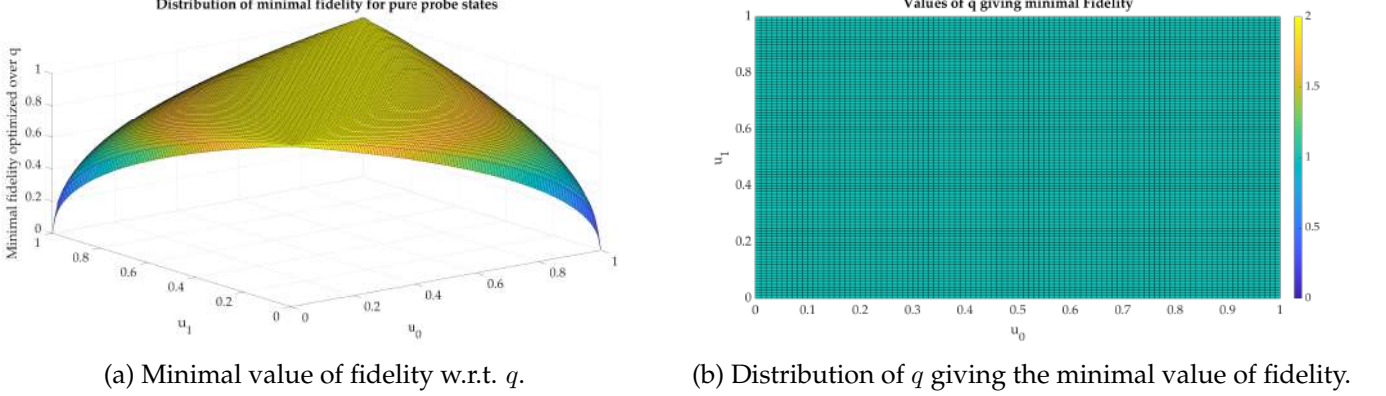


Figure 6: Analysis of fidelity for the pure probe state.

Considering mixed probe states, the same analysis can be done. There is a difference now on how to probe the channel. Since we have a bipartite entangled state, it can be used to probe the channel once or twice. The channel output in both of these situations is given by

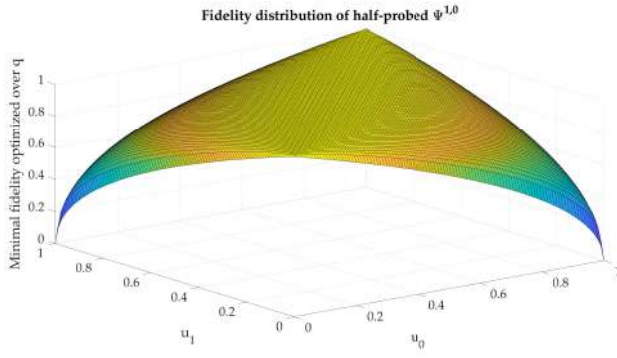
$$\begin{aligned} \mathcal{W}^{u_i} \otimes \text{id}_{A_2}(|\Psi^{xz}\rangle\langle\Psi^{xz}|) &= \left[\left((1-q)\delta_{0x} |0\rangle\langle 0|_{A_2} + q\delta_{1x} |1\rangle\langle 1|_{A_2} \right) \right. \\ &+ \left. u_i \left((1-q)\delta_{1x} |0\rangle\langle 0|_{A_2} + q\delta_{0x} |1\rangle\langle 1|_{A_2} \right) \right] |0\rangle\langle 0|_{A_1} \\ &+ (\delta_{0z} - \delta_{1z})\sqrt{(1-u_i)q(1-q)} \left[\delta_{0x} |0\rangle\langle 1|_{A_2} + \delta_{1x} |1\rangle\langle 0|_{A_2} \right] |0\rangle\langle 1|_{A_1} \\ &+ (\delta_{0z} - \delta_{1z})\sqrt{(1-u_i)q(1-q)} \left[\delta_{1x} |0\rangle\langle 1|_{A_2} + \delta_{0x} |1\rangle\langle 0|_{A_2} \right] |1\rangle\langle 0|_{A_1} \\ &+ (1-u_i) \left[(1-q)\delta_{1x} |0\rangle\langle 0|_{A_2} + q\delta_{0x} |1\rangle\langle 1|_{A_2} \right] |1\rangle\langle 1|_{A_1}, \end{aligned} \quad (94)$$

where $i = 0, 1$ and δ_{ab} is the Kronecker delta, and

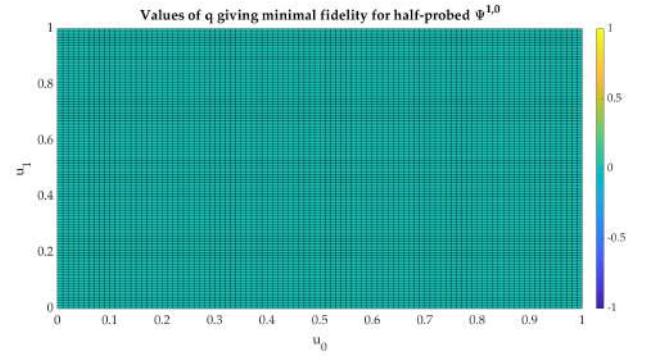
$$\begin{aligned} \mathcal{W}^{u_0} \otimes \mathcal{W}^{u_1}(|\Psi^{xz}\rangle\langle\Psi^{xz}|) &= [(1-q)\delta_{0x} + u_0(1-q)\delta_{1x}] |0\rangle\langle 0|_{A_1} \otimes |0\rangle\langle 0|_{A_2} \\ &+ q(u_0\delta_{0x} + \delta_{1x}) |0\rangle\langle 0|_{A_1} \otimes [u_1 |0\rangle\langle 0|_{A_2} + (1-u_1) |1\rangle\langle 1|_{A_2}] \\ &+ (-1)^z \sqrt{(1-u_0)(1-u_1)q(1-q)} [\delta_{0x} |0\rangle\langle 1|_{A_1} + \delta_{1x} |1\rangle\langle 0|_{A_1}] \otimes |0\rangle\langle 1|_{A_2} \\ &+ (-1)^z \sqrt{(1-u_0)(1-u_1)q(1-q)} [\delta_{0x} |1\rangle\langle 0|_{A_1} + \delta_{1x} |0\rangle\langle 1|_{A_1}] \otimes |1\rangle\langle 0|_{A_2} \\ &+ (1-u_0)q\delta_{0x} |1\rangle\langle 1|_{A_1} \otimes [u_1 |0\rangle\langle 0|_{A_2} + (1-u_1) |1\rangle\langle 1|_{A_2}] \\ &+ (1-u_0)(1-q)\delta_{1x} |1\rangle\langle 1|_{A_1} \otimes |0\rangle\langle 0|_{A_2}. \end{aligned} \quad (95)$$

For each value of x, z , we minimize the fidelity with respect to q considering these two cases, which we call half-probed and full-probed cases. The first one corresponds to Eq. 94 and the second to Eq. 95. Our first description is over the half-probed case. See Figs. 7 and 8 below. Notice that we have shown just the instance $z = 0$, since it gives the same results as for $z = 1$.

Two conclusions can be obtained from Figs. 7 and 8. Firstly, the shape of the fidelity of half-probed state $|\Psi^{10}\rangle$ is exactly the same as for the pure probe state case. This is clear from the fact that the value of q giving maximal fidelity is $q = 0$, which turns the probe state to be $|1\rangle$. As the second point, we see in Figs. 7 and 8 that half-probing with the state $|\Psi^{00}\rangle$ can give some improvement in

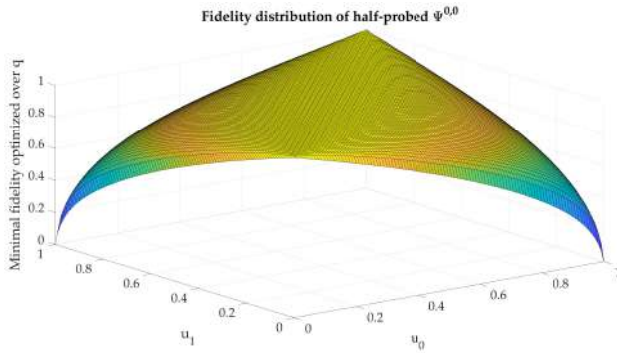


(a) Minimal value of fidelity w.r.t. q .

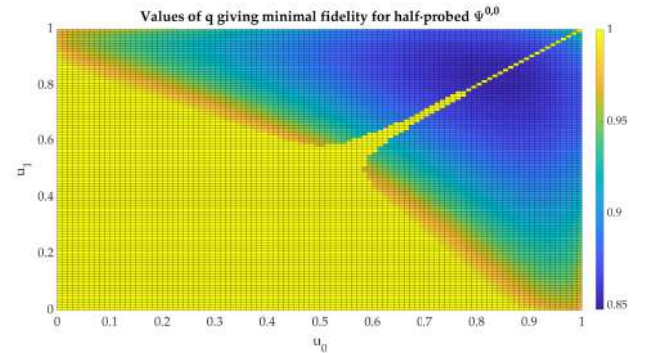


(b) Distribution of q giving the minimal value of fidelity.

Figure 7: Analysis of fidelity for the half-probed state $|\Psi^{1,0}\rangle$



(a) Minimal value of fidelity w.r.t. q .



(b) Distribution of q giving the minimal value of fidelity.

Figure 8: Analysis of fidelity for the half-probed state $|\Psi^{0,0}\rangle$.

the channel discrimination task. This improvement is obtained where $q \neq 1$. However, the difference is not so significant, making the use of $|\Psi^{00}\rangle$ justified only if entanglement cost is negligible.

The subsequent analysis is over a full-probed state. In opposition to what has been shown in the half-probed case, there is no improvement in using full-probed $|\Psi^{00}\rangle$. Fig. 10 shows that $q = 1$ gives the minimal value of fidelity in this case. This leads to the fidelity having values equal to the square of the pure probe state. Thus, optimizing full-probed strategy using $|\Psi^{00}\rangle$ leads to the same strategy for pure probe state. Lastly, we consider full-probed strategy using $|\Psi^{10}\rangle$. We have $q = 0$ giving the optimal value of fidelity. A consequence of this is that now we probe the quantum channel twice using the state $|0\rangle$ and $|1\rangle$ in each round. Since for pure probe, we are using the state $|1\rangle$, then the round probing with $|0\rangle$ contribute degrading the fidelity. Therefore, the optimal fidelity obtained using full-probed $|\Psi^{10}\rangle$ is much higher than using full-probed $|\Psi^{00}\rangle$, as can be seen in Fig. 12.

The previous results can be extended to the mutual information between the classical system X and the quantum channel output system B . In Proposition 8 we have seen that rate and reliability are connected in a way that when the rate is close to its maximum, then the reliability is close to its minimum, and vice versa. Thus, we can conclude the following. For channels with values of u_0 and u_1 close, the mutual information will be close to zero. On the other hand, when one of the u_i , $i = 0, 1$, is close to unit and the other is close to zero, the mutual information attains its maximum value possible. Furthermore, probing the quantum memory cell using the half-probed $|\Psi^{00}\rangle$ strategy is again the optimal procedure for quantum reading.

5 Final Remarks

We have demonstrated a new polar coding scheme for quantum memory cell discrimination. To achieve this goal, we had to introduce new definitions on the rate and reliability of a quantum channel. The polarization phenomenon produced by channel combining and analyzed in the channel

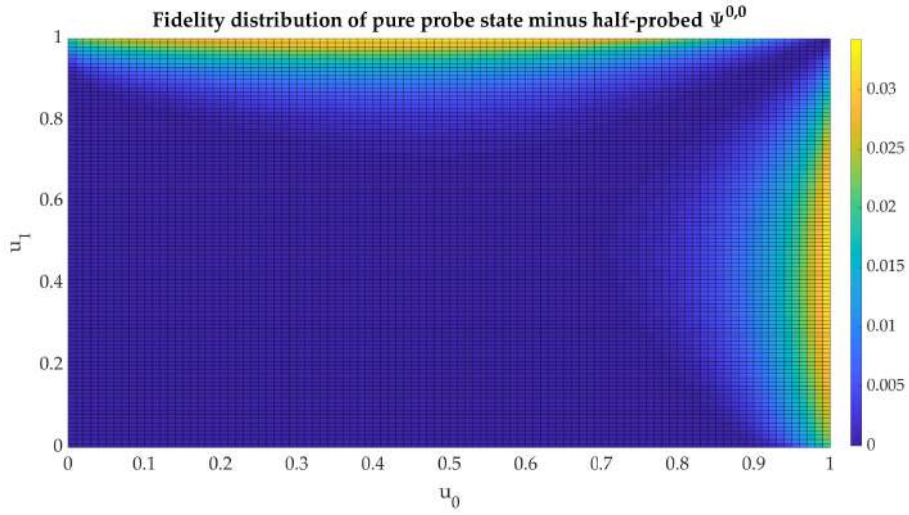
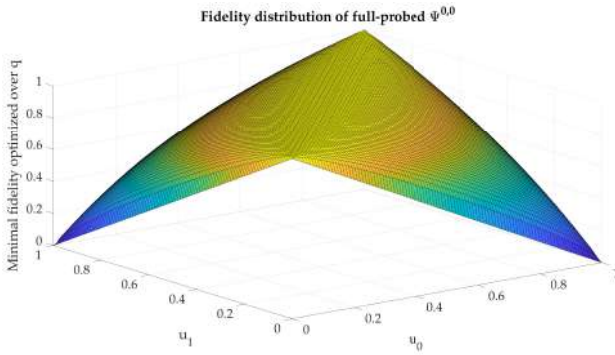
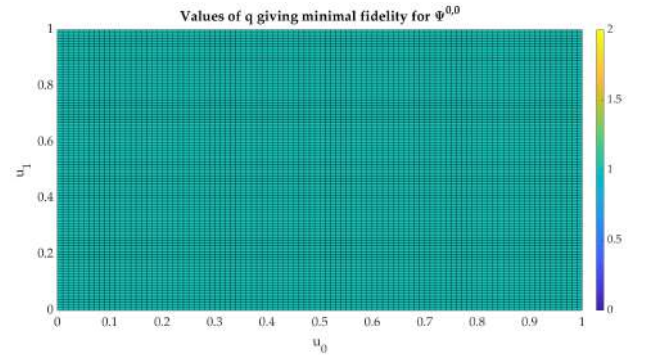


Figure 9: Fidelity of pure state minus fidelity of half-probed state $|\Psi^{0,0}\rangle$.

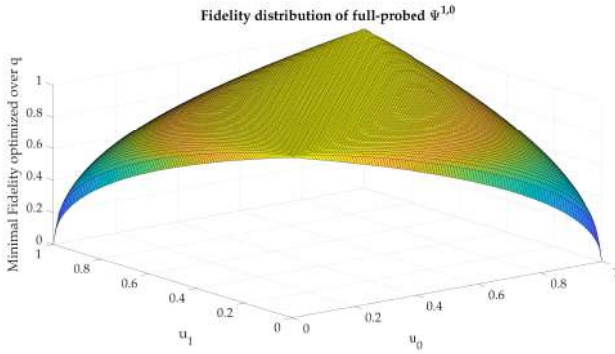


(a) Minimal value of fidelity w.r.t. q .

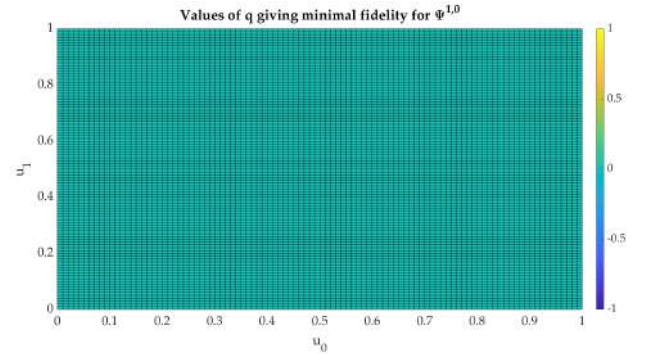


(b) Distribution of q giving the minimal value of fidelity.

Figure 10: Analysis of fidelity for the full-probed state $|\Psi^{0,0}\rangle$.



(a) Minimal value of fidelity w.r.t. q .



(b) Distribution of q giving the minimal value of fidelity.

Figure 11: Analysis of fidelity for the full-probed state $|\Psi^{1,0}\rangle$

splitting part has been shown for these two quantities. This was established due to an inequality connecting both of them. In the channel splitting part, we have also introduced the synthesized channels created by the polar coding. It had been shown that when the number of channels is arbitrarily large, the set of synthesized channels can be divided in two groups, good and bad channels. Additionally, the fraction of such channels is related to the mutual information of the original quantum channel in consideration. This result has motivated us to construct an optimal encoding scheme for quantum reading. A decoding scheme was introduced and analyzed, as well. Using an existence proof of pretty good measurements given in this paper and a previous quantum union bound from the liter-

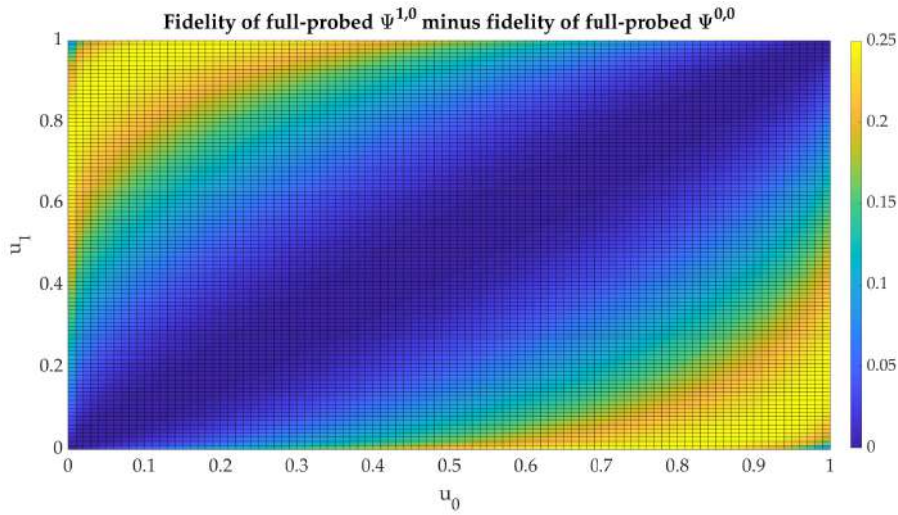


Figure 12: Fidelity of full-probed state $|\Psi^{1,0}\rangle$ minus fidelity of full-probed state $|\Psi^{0,0}\rangle$.

ature, we have shown that our decoding scheme has error probability that decays exponentially fast with respect to the code length. At the end, optimizations over probe states are investigated, leading to the conclusion of half-probed states as the best choice, in general, of probe states to be used.

This paper has given some future investigation topics. A question that can be stated is how to apply polar coding to more general classical and quantum systems. First of all, we can extend the binary discrimination to a d -ary discrimination task. The mathematical equivalence of this is considering a classical system with a larger alphabet. Some classical polar codes have been proposed in the literature and, with the proper adjustments, applied here. Secondly, we can consider the set of quantum memory cells to be composed of generalized amplitude damping channel. This class of channel can be seen as a second-order approximation of classical digital memory, where the model takes the environment temperature into consideration. In this open question, the task would be to find the optimal probe state for channel discrimination. Third, and more important, how can one construct efficient polar codes for Gaussian bosonic channel discrimination. This still a research topic even for classical Gaussian channel. Gaussian bosonic channel is the ultimate goal as quantum channel model for classical digital memories. Because this class of channels is defined over continuous variables, attacking it needs to be two-fold. The polar code scheme has to take into account the channel to prove the polarization phenomenon. Additionally, an optimization over the probe states is necessary under energy constrain. These two approaches are not independent, which makes the task even more difficult. Lastly, we do not know if there exists a provable optimal probe state in the contexts mentioned. Uniqueness also needs to be verified.

6 Acknowledgments

The authors acknowledge the financial support of the Future and Emerging Technologies (FET) programme, within the Horizon-2020 Programme of the European Commission, under the FET-Open grant agreement QUARTET, number 862644.

References

- [1] Emmanuel Abbe. "Randomness and dependencies extraction via polarization". In: *2011 Information Theory and Applications Workshop*. IEEE, 2011. DOI: [10.1109/ita.2011.5743556](https://doi.org/10.1109/ita.2011.5743556).
- [2] A. Acín. "Statistical Distinguishability between Unitary Operations". In: *Physical Review Letters* 87.17 (2001). DOI: [10.1103/physrevlett.87.177901](https://doi.org/10.1103/physrevlett.87.177901).

- [3] Erdal Arıkan. “Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels”. In: *IEEE Transactions on Information Theory* 55.7 (2009), pp. 3051–3073. DOI: [10.1109/tit.2009.2021379](https://doi.org/10.1109/tit.2009.2021379).
- [4] Erdal Arıkan. “Source polarization”. In: *2010 IEEE International Symposium on Information Theory*. IEEE, 2010. DOI: [10.1109/isit.2010.5513567](https://doi.org/10.1109/isit.2010.5513567).
- [5] Erdal Arıkan and Emre Telatar. “On the rate of channel polarization”. In: *2009 IEEE International Symposium on Information Theory*. IEEE, 2009. DOI: [10.1109/isit.2009.5205856](https://doi.org/10.1109/isit.2009.5205856).
- [6] Leonardo Banchi, Quntao Zhuang, and Stefano Pirandola. “Quantum-enhanced barcode decoding and pattern recognition”. In: (Oct. 7, 2020). arXiv: [2010.03594v1](https://arxiv.org/abs/2010.03594v1) [quant-ph].
- [7] H. Barnum and E. Knill. “Reversing quantum dynamics with near-optimal quantum and classical fidelity”. In: *Journal of Mathematical Physics* 43.5 (2002), p. 2097. DOI: [10.1063/1.1459754](https://doi.org/10.1063/1.1459754).
- [8] Andrew M. Childs, John Preskill, and Joseph Renes. “Quantum information and precision measurement”. In: *Journal of Modern Optics* 47.2-3 (2000), pp. 155–176. DOI: [10.1080/09500340008244034](https://doi.org/10.1080/09500340008244034).
- [9] Siddhartha Das and Mark M. Wilde. “Quantum rebound capacity”. In: *Physical Review A* 100.3 (2019). DOI: [10.1103/physreva.100.030302](https://doi.org/10.1103/physreva.100.030302).
- [10] Runyao Duan, Yuan Feng, and Mingsheng Ying. “Perfect Distinguishability of Quantum Operations”. In: *Physical Review Letters* 103.21 (2009). DOI: [10.1103/physrevlett.103.210501](https://doi.org/10.1103/physrevlett.103.210501).
- [11] Alexei Gilchrist, Nathan K. Langford, and Michael A. Nielsen. “Distance measures to compare real and ideal quantum processes”. In: *Physical Review A* 71.6 (2005). DOI: [10.1103/physreva.71.062310](https://doi.org/10.1103/physreva.71.062310).
- [12] Aram W. Harrow, Avinandan Hassidim, Debbie W. Leung, and John Watrous. “Adaptive versus nonadaptive strategies for quantum channel discrimination”. In: *Physical Review A* 81.3 (2010). DOI: [10.1103/physreva.81.032339](https://doi.org/10.1103/physreva.81.032339).
- [13] Masahito Hayashi. “Discrimination of Two Channels by Adaptive Methods and Its Application to Quantum System”. In: *IEEE Transactions on Information Theory* 55.8 (2009), pp. 3807–3820. DOI: [10.1109/tit.2009.2023726](https://doi.org/10.1109/tit.2009.2023726).
- [14] Carl W. Helstrom. “Quantum detection and estimation theory”. In: *Journal of Statistical Physics* 1.2 (1969), pp. 231–252. DOI: [10.1007/bf01007479](https://doi.org/10.1007/bf01007479).
- [15] Christoph Hirche. “From asymptotic hypothesis testing to entropy inequalities”. PhD thesis. Universitat Aut3noma de Barcelona, 2018.
- [16] A. S. Holevo. “Statistical decision theory for quantum systems”. In: *Journal of Multivariate Analysis* 3.4 (1973), pp. 337–394. DOI: [10.1016/0047-259x\(73\)90028-6](https://doi.org/10.1016/0047-259x(73)90028-6).
- [17] A.S. Holevo. “Reliability function of general classical-quantum channel”. In: *IEEE Transactions on Information Theory* 46.6 (2000), pp. 2256–2261. DOI: [10.1109/18.868501](https://doi.org/10.1109/18.868501).
- [18] Junya Honda and Hirosuke Yamamoto. “Polar Coding Without Alphabet Extension for Asymmetric Models”. In: *IEEE Transactions on Information Theory* 59.12 (2013), pp. 7829–7838. DOI: [10.1109/tit.2013.2282305](https://doi.org/10.1109/tit.2013.2282305).
- [19] Vishal Katariya and Mark M. Wilde. “Evaluating the Advantage of Adaptive Strategies for Quantum Channel Distinguishability”. In: (Jan. 15, 2020). arXiv: [2001.05376v1](https://arxiv.org/abs/2001.05376v1) [quant-ph].
- [20] Seth Lloyd. “Enhanced Sensitivity of Photodetection via Quantum Illumination”. In: *Science* 321.5895 (2008), pp. 1463–1465. DOI: [10.1126/science.1160627](https://doi.org/10.1126/science.1160627).
- [21] Samad Khabbazi Oskouei, Stefano Mancini, and Mark M. Wilde. “Union bound for quantum information processing”. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 475.2221 (2019), p. 20180612. DOI: [10.1098/rspa.2018.0612](https://doi.org/10.1098/rspa.2018.0612).
- [22] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd. “Advances in photonic quantum sensing”. In: *Nature Photonics* 12.12 (2018), pp. 724–733. DOI: [10.1038/s41566-018-0301-6](https://doi.org/10.1038/s41566-018-0301-6).

- [23] Stefano Pirandola. “Quantum Reading of a Classical Digital Memory”. In: *Physical Review Letters* 106.9 (2011). DOI: [10.1103/physrevlett.106.090504](https://doi.org/10.1103/physrevlett.106.090504).
- [24] Stefano Pirandola, Riccardo Laurenza, Cosmo Lupo, and Jason L. Pereira. “Fundamental limits to quantum channel discrimination”. In: *npj Quantum Information* 5.1 (2019). DOI: [10.1038/s41534-019-0162-y](https://doi.org/10.1038/s41534-019-0162-y).
- [25] Stefano Pirandola, Cosmo Lupo, Vittorio Giovannetti, Stefano Mancini, and Samuel L Braunstein. “Quantum reading capacity”. In: *New Journal of Physics* 13.11 (2011), p. 113012. DOI: [10.1088/1367-2630/13/11/113012](https://doi.org/10.1088/1367-2630/13/11/113012).
- [26] Stefano Pirandola, Stefano Mancini, Seth Lloyd, and Samuel L. Braunstein. “Continuous-variable quantum cryptography using two-way quantum communication”. In: *Nature Physics* 4.9 (2008), pp. 726–730. DOI: [10.1038/nphys1018](https://doi.org/10.1038/nphys1018).
- [27] Joseph M. Renes, David Sutter, Frederic Dupuis, and Renato Renner. “Efficient Quantum Polar Codes Requiring No Preshared Entanglement”. In: *IEEE Transactions on Information Theory* 61.11 (2015), pp. 6395–6414. DOI: [10.1109/tit.2015.2468084](https://doi.org/10.1109/tit.2015.2468084).
- [28] Wojciech Roga, Mark Fannes, and Karol Życzkowski. “Universal Bounds for the Holevo Quantity, Coherent Information, and the Jensen-Shannon Divergence”. In: *Physical Review Letters* 105.4 (2010). DOI: [10.1103/physrevlett.105.040505](https://doi.org/10.1103/physrevlett.105.040505).
- [29] Massimiliano F. Sacchi. “Entanglement can enhance the distinguishability of entanglement-breaking channels”. In: *Physical Review A* 72.1 (2005). DOI: [10.1103/physreva.72.014305](https://doi.org/10.1103/physreva.72.014305).
- [30] Massimiliano F. Sacchi. “Optimal discrimination of quantum operations”. In: *Physical Review A* 71.6 (2005). DOI: [10.1103/physreva.71.062340](https://doi.org/10.1103/physreva.71.062340).
- [31] Aria G. Sahebi and S. Sandeep Pradhan. “Multilevel polarization of polar codes over arbitrary discrete memoryless channels”. In: *2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2011. DOI: [10.1109/allerton.2011.6120375](https://doi.org/10.1109/allerton.2011.6120375).
- [32] Eren Sasoglu, Emre Telatar, and Erdal Arıkan. “Polarization for arbitrary discrete memoryless channels”. In: *2009 IEEE Information Theory Workshop*. IEEE, 2009. DOI: [10.1109/itw.2009.5351487](https://doi.org/10.1109/itw.2009.5351487).
- [33] David Sutter, Joseph M. Renes, Frederic Dupuis, and Renato Renner. “Achieving the capacity of any DMC using only polar codes”. In: *2012 IEEE Information Theory Workshop*. IEEE, 2012. DOI: [10.1109/itw.2012.6404638](https://doi.org/10.1109/itw.2012.6404638).
- [34] Ido Tal and Alexander Vardy. “How to Construct Polar Codes”. In: *IEEE Transactions on Information Theory* 59.10 (2013), pp. 6562–6582. DOI: [10.1109/tit.2013.2272694](https://doi.org/10.1109/tit.2013.2272694).
- [35] Si-Hui Tan, Baris I. Erkmen, Vittorio Giovannetti, Saikat Guha, Seth Lloyd, Lorenzo Maccone, Stefano Pirandola, and Jeffrey H. Shapiro. “Quantum Illumination with Gaussian States”. In: *Physical Review Letters* 101.25 (2008). DOI: [10.1103/physrevlett.101.253601](https://doi.org/10.1103/physrevlett.101.253601).
- [36] Guoming Wang and Mingsheng Ying. “Unambiguous discrimination among quantum operations”. In: *Physical Review A* 73.4 (2006). DOI: [10.1103/physreva.73.042301](https://doi.org/10.1103/physreva.73.042301).
- [37] Mark M. Wilde and Saikat Guha. “Polar Codes for Classical-Quantum Channels”. In: *IEEE Transactions on Information Theory* 59.2 (2013), pp. 1175–1187. DOI: [10.1109/tit.2012.2218792](https://doi.org/10.1109/tit.2012.2218792).
- [38] Mark M. Wilde and Saikat Guha. “Polar Codes for Degradable Quantum Channels”. In: *IEEE Transactions on Information Theory* 59.7 (2013), pp. 4718–4729. DOI: [10.1109/tit.2013.2250575](https://doi.org/10.1109/tit.2013.2250575).