Polar Codes for Quantum Reading

Francisco Revson F. Pereira and Stefano Mancini

***Abstract*— Quantum readout provides a general framework for formulating statistical discrimination of quantum channels. Several paths have been taken for such this 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 out: how to construct low complexity encoding schemes that are interesting for channel discrimination and, more importantly, how to develop capacity-achieving protocols. This paper aims at presenting a solution to these questions using polar codes. Firstly, we characterize the information rate and reliability parameter of the channels under polar encoding. We also show that the error probability of the scheme proposed decays exponentially with the square root of the code length. Secondly, an analysis of the optimal quantum states to be used as probes is given.**

***Index Terms*— Quantum reading, polar codes, capacity-achieving protocols.**

I. Introduction

QUANTUM hypothesis testing aims to identify strategies 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. [15], [18]). The latter, under the name of quantum channel discrimination, has been recently addressed [1], [7], [9], [11], [13], [14], [31], [32], [38]. In its basic formulation, one has to identify a quantum channel selected from a set according to a probability measure. This should be done by using a suitable input state and output measurement. As such, it is a double-optimization problem and hence results in a daunting task. The performance is usually quantified in terms of the minimum error probability, and recently bounds on it were found for general strategies [21], [26]. 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 [28], where the secret information is encoded in a Gaussian ensemble of phase-space displacements. It also appears in the quantum illumination of targets [22], [37], 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 [25]. There, the data storage corresponds to a process of channel encoding, where information is recorded into a cell, a quantum-mechanical version of the classical memory cell, by storing a quantum channel picked from a given ensemble [25]. 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 surpass any classical strategy based on mixtures of coherent states, hence the name of "quantum" reading.

Efficient paths to quantum reading can also be envisaged by the use of coding techniques [8], [27]. 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. To decrease the error probability in discriminating the channels in the ensemble, we can use encoding and decoding schemes in the following manner. The classical information to be stored in cells is processed by the classical encoder to shield it from noise and measurement errors. The ratio between information bits and code length is denoted by the code rate in the paper. The output of the classical encoder, a codeword of the classical code used in the process, corresponds to the actual string of information to be stored in the cells. Suppose now that a measurement has been implemented in the cells and one has a string, with possible errors. Then, a classical decoder is used to retrieve the initially noiseless classical information. There are several classical codes that one could use for quantum reading, depending on the goal in mind. Denote by information rate the information-theoretical quantity describing a coding rate under which there exists a code with length going to infinity but attaining low error probability. Then, we are interested in information-theoretically provable codes that could be used to attain high information rates and error probability decreasing exponentially with respect to some power of the code length. So, 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 theoretically prove the existence of such codes. Fortunately, by developing the idea of information

Manuscript received December 13, 2020; revised January 31, 2022; accepted February 17, 2022. Date of publication March 16, 2022; date of current version June 15, 2022. This work was supported by the European Union Horizon 2020 Research and Innovation Programme, Quantum readout Techniques and Technologies (QUARTET) under Grant 862644. An earlier version of this paper was presented in part at the 2021 IEEE International Symposium on Information Theory [DOI: 10.1109/ISIT45174.2021.9517807]. *(Corresponding author: Francisco Revson F. Pereira.)*

The authors are with the School of Science and Technology, University of Camerino, 62032 Camerino, Italy, and also with the Istituto Nazionale di Fisica Nucleare Sezione di Perugia, 06123 Perugia, Italy (e-mail: revson.ee@gmail.com; stefano.mancini@unicam.it).

Communicated by P. K. Sarvepalli, Associate Editor for Quantum.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TIT.2022.3159488.

Digital Object Identifier 10.1109/TIT.2022.3159488

combining, Arikan was able to show that for binary memoryless symmetric (BMS) channels, such codes exist [2]. 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" channels[1]. 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 mentioning that the polar code construction highly depends on the channel in consideration.

The analysis of using polar codes for quantum memory cells in this paper is twofold. In the first part, for fixed probe states, polar coding is applied to the labels of the quantum memory cell to decrease the error probability in distinguishing them. In this direction, we show that the polarization phenomenon can be characterized in quantum memory cells 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-theoretical 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 information rate and reliability parameter of the synthesized channels polarize, where the reliability parameter quantifies the distinguishability of the channel outputs. This is later used in the asymptotic analysis. A connection between information rate and reliability parameter is also presented, showing that the information rate is inversely proportional to the reliability parameter. Next, our first major result is given. It characterizes the asymptotic distribution of the synthesized channel with the code length. 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 construct encoding and decoding schemes. 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 square root of the code length. The complexity of the encoding scheme is $O(N \log N)$, where N is the code length. However, we have not been able to compute the decoding scheme complexity. Construction of low or reduced decoding complexity schemes is the subject of current ongoing research of (classical and quantum) polar codes.

It is important to notice that the analysis using polar codes is elaborated with fixed probe states. This allows us to use the model of classical-quantum channels in the quantum reading framework. Indeed, quantum memory cells can be seen as a set of classical-quantum channels for fixed probe states. However, the choice of the probe state can drastically change the conclusion one can get. For example, the polarization phenomenon may not happen for a poor choice of probe states. Thus, to avoid any misunderstanding, we do not merely consider classical-quantum channels, but rather our formulation explicitly considers quantum memory cells and shows the probe state dependence.

The second part of this paper 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. Moreover, even though we have considered just the first level polar encoding, it seems satisfactory supposing this claim can be extended to any N-level polar encoding.

A. Relation With Previous Works

Several papers extend or apply the idea of polarization in a diversity of channels and areas [3], [10], [12], [17], [20], [23], [29], [33], [34], [39], [40]. In particular, polar codes have also been constructed for classical-quantum channels [23], [39], and quantum channels [17], [29], [40]. These papers aim at achieving the capacity of classical-quantum or quantum channels using polar coding. For classical-quantum channels, the ones relevant for this paper, Ref. [39] extends the fundamental ideas of Arikan's landmark paper. The main technical contributions of them are a characterization of information rate and reliability parameter for the paradigm in consideration, and the generalization of Arikan's successive cancellation decoder to the quantum case. In Ref. [23], the authors consider arbitrary[2] classical-quantum channels and arbitrary classical-quantum multiple access channels by using arbitrary Abelian group operations on the input alphabets. The distribution of the input alphabet is considered uniformly distributed.

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 by the energy constraint imposed over the channels. Since these two conditions are plausible hypotheses to be considered in classical digital memories, which are the main goal of the 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 designed to, primarily, asymmetric classical-quantum channels, differently from Refs.[23], [39]. In the classical paradigm, this has been done by Sutter *et al.* [35], and Honda and Yamamoto [20]. Some of the ideas used in Ref. [20] 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. This is particularly true for the measurement process implemented in the decoding part. Adequate definitions for information rate and reliability parameter are also needed. Thus, we are not only extending

[1]The reason for such names is due to the capacity of these new channels being close to their maximum value and close to zero, respectively.

[2]Not necessarily binary.

the results of Ref. [20], as Ref. [39] does in comparison to Arikan's work, but also applying them in a different paradigm.

The use and analysis of classical codes in the quantum reading paradigm have not been addressed until now. There was a persistent aim at showing that entanglement-assisted probing outperforms classical strategies, culminating with recent work on barcodes data decoding [5]. On the other hand, Refs. [8], [27] introduced the capacity of the quantum reading paradigm by just resorting to random coding arguments. Thus, our paper has taken the novelty path of proposing an explicit encoding and decoding scheme that exploits redundancy in the quantum reading framework. Both schemes have been analyzed in terms of coding and probing strategies, with an error probability decaying asymptotically with the square root of the code length.

B. Structure of the Work

This paper is organized as follows. In Section II we present the notation used through this paper, some definitions that are relevant to characterize quantum memory cells, and previous results on polar codes for classical-quantum channels. Next, the main results are shown in Section III. 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 IV. Lastly, we draw our conclusions and some final remarks in Section V.

II. Preliminaries

This section is devoted to introducing the main concepts of quantum reading and polar codes. We begin with the notation used in the paper. Afterward, a brief overview of binary polar codes and their desirable attributes is presented. They can be used to achieve the capacity of discrete memoryless channels (DMC); there are efficient encoding and decoding schemes; and it is possible to reach error probabilities that decay exponentially in the square root of the code length. In the following, the quantum reading paradigm 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 information rate and reliability parameter.

A. 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 $\mathcal{X} = \mathbb{Z}_2$ 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 $j \geq i$) is the sequence of random variables $X_i, X_{i+1}, \ldots, X_j$. X^N simply stands for the sequence $X_1, X_2, \ldots, X_N$. A classical memoryless channel is written as $W\colon \mathcal{X} \to \mathcal{Y}$ 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., $W^N\colon \mathcal{X}^N \to \mathcal{Y}^N$ with transition matrix $W^N(y^N|x^N) = \prod_{i=1}^N W(y_i|x_i)$, 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 $A^N := A_1 A_2 \cdots A_N$ 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. The von Neumann entropy $H(\rho)$ of the density operator ρ is defined as $H(\rho) := -\text{Tr}\{\rho\log\rho\}$. The quantum mutual information of a bipartite density operator $\rho_{AB} \in \mathcal{D}(\mathcal{H}_{AB})$ is defined as $I(A;B) := H(\text{Tr}_A\{\rho\}) + H(\text{Tr}_B\{\rho\}) - H(\rho)$. The fidelity between two operators $\rho, \sigma \in \mathcal{D}(\mathcal{H}_A)$ is given by $F(\rho,\sigma) := ||\sqrt{\rho}\sqrt{\sigma}||_1$, where $||E||_1$ is the Schatten 1-norm of an operator $E \in \mathcal{L}(\mathcal{H}_A)$. Lastly, a quantum channel $\mathcal{W}$ is a linear completely positive trace-preserving map from $\mathcal{L}(\mathcal{H}_A)$ to $\mathcal{L}(\mathcal{H}_B)$.

B. Quantum Memory Cell

We now formulate the general description of a 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

$$\mathcal{W}^x\colon \mathcal{D}(\mathcal{H}_{B'}) \to \mathcal{D}(\mathcal{H}_B) \quad (1)$$

$$\rho \mapsto \mathcal{W}^x(\rho), \quad (2)$$

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 $\mathcal{W}^x_{B'\to B}$ to highlight the input and output systems of $\mathcal{W}^x$. An important hypothesis is given 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 by the energy constraint imposing this distribution. Since these two reasonings 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.

The quantum reading framework can be explained by a two-part process. The first process consists of encoding the information bits. Suppose Alice has a string of information bits

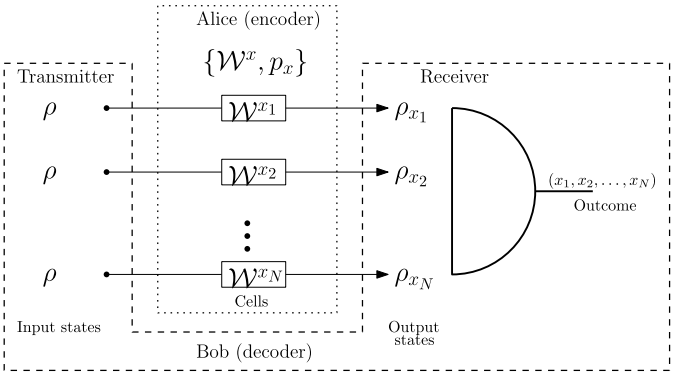


Fig. 1. Quantum reading scheme.

that Bob is planning to read in a future moment using quantum tools. Additionally, we are considering she has access to a quantum channel array that composes a memory. Then Alice stores the information bits by choosing an appropriate quantum channel according to the information bits. In the end, she will have an array of quantum channels labeled by the information bits. Finished this process, we go to the decoding part. Bob decodes the information stored in the memory by probing each of the quantum channels and making a measurement on the output state. Therefore, Bob's task is two-fold. He has to choose an optimal state to probe a quantum channel, which is unknown beforehand, and also choose a decoding strategy that minimizes the error probability in determining the information bit stored by Alice. We have schematized the process in Fig. 1.

Remark 1: Before introducing the channel model used for numerical analysis in this paper, an important point-of-view over quantum memory cells needs to be given. A common approach to quantum memory cells 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, a quantum memory cell can be seen as a set of classical-quantum channels. Thus, a capacity-achieving approach using polar codes is possible when dealing with quantum memory cell elements with 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 the bosonic attenuator channel, which is the standard model for digital memory cells. 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:

$$\mathcal{W}^x(\rho) = A_0(x)\rho A_0(x)^\dagger + A_1(x)\rho A_1(x)^\dagger, \tag{3}$$

with

$$A_0(x) = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma(x)} \end{pmatrix}, \quad A_1(x) = \begin{pmatrix} 0 & \sqrt{\gamma(x)} \\ 0 & 0 \end{pmatrix}, \tag{4}$$

where $\gamma(x)$ is the decay probability with respect to x. Observe that the same notation for the 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 is possible to attain optimal rate with low reliability parameter. This is obtained 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, let X be a random variable with probability law p_X. We can write the joint density state describing the systems X and B as

$$\rho^{XB} = p_X(0)\,(0)\,0^X \otimes \mathcal{W}^0(\rho) + p_X(1)\,(1)\,1^X \otimes \mathcal{W}^1(\rho). \tag{5}$$

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

Definition 2: Let $X \sim \text{Ber}(p)$, $\mathcal{W}^x_{B'\to B}(\rho)$ be a quantum memory cell, where $x \in \mathcal{X}$, and $I(X;B)_\rho$ be the quantum mutual information of the state ρ^{XB}. The *information 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

$$\begin{aligned} I(\mathcal{W})_\rho &= H\Big(p\mathcal{W}^0(\rho) + (1-p)\mathcal{W}^1(\rho)\Big) \\ &- pH(\mathcal{W}^0(\rho)) - (1-p)H(\mathcal{W}^0(\rho)), \end{aligned} \tag{6}$$

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

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

Definition 3: Let $X \sim \text{Ber}(p)$ and $\mathcal{W}^x_{B'\to B}(\rho)$ be a quantum memory cell, where $x \in \mathcal{X}$. The *reliability parameter* of the quantum memory cell $\mathcal{W}$ is defined as

$$\begin{aligned} Z(\mathcal{W})_\rho &:= 2\sqrt{p(1-p)}F(\mathcal{W}^0(\rho), \mathcal{W}^1(\rho)) \\ &= 2\sqrt{p(1-p)}||\sqrt{\mathcal{W}^0(\rho)}\sqrt{\mathcal{W}^1(\rho)}||_1. \end{aligned} \tag{7}$$

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

C. Classical-Quantum Polar Codes

Let $W\colon \mathcal{X} \to \mathcal{D}(\mathcal{H}_B)$ 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 conditional input-output probabilities can be derived from the joint input-output state

$$\rho^{XB} = \frac{1}{2}\,(0)\,0^X \otimes \rho_0^B + \frac{1}{2}\,(1)\,1^X \otimes \rho_1^B, \tag{8}$$

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

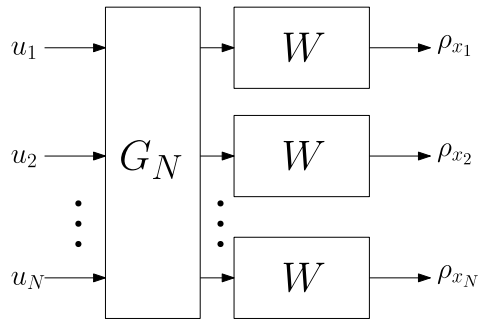


Fig. 2. General polar encoding scheme for a classical-quantum channel. The matrix G_N represents the composition function applied on the input u^N and resulting in the vector $x^N = u^N G_N$. The density operator $\rho_{x^N} = \rho_{x_1} \otimes \cdots \otimes \rho_{x_N}$ is the output $W^N(x^N)$.

omitted. One important characterization can be given to this cq channel: consider a binary random variable R on the sample space $\{\rho_0, \rho_1\}$. The probabilities of R are obtained from ρ^{XB}. Furthermore, from the density operator ρ^{XB}, we can also see that $P_{R|X}(0|0) = P_{R|X}(1|1)$ and $P_{R|X}(0|1) = P_{R|X}(1|0)$. From the above relation, 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 channels to provide a more general formulation. As the 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 of 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 [2], [4], [20], [29], [39], [40]: 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 code length); the error probability of the decoder decays exponentially in the square root of the code length. A descriptive explanation of channel polarization, synthesized channels, and some attributes of polar codes is 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. 2. 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. 3.

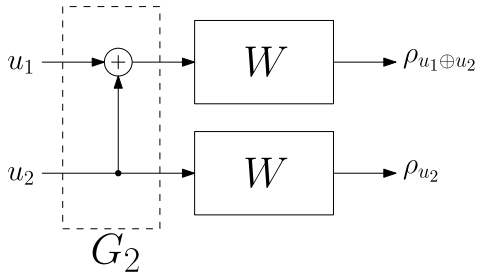


Fig. 3. Polar encoding scheme for $N = 2$. The choice of the composition function characterizes the encoding scheme to be polar.

Now, we can introduce channel splitting and the polarization phenomenon that emerge from it. For $i \in \{1, \ldots, N\}$, 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

$$W_N^{(i)}(u_i) = \sum_{u_1^{i-1}} \frac{1}{2^{i-1}} \left(u_1^{i-1}\right) u_1^{i-1 U_1^{i-1}} \otimes \overline{\rho}_{u_1^i}^{B^N}, \quad (9)$$

$$\overline{\rho}_{u_1^i}^{B^N} = \sum_{u_{i+1}^N} \frac{1}{2^{N-i}} \rho_{u^N G_N}^{B^N}, \quad (10)$$

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 [2] or its extension to classical-quantum channels in Ref. [39].

There are two important quantities used to quantify the polarization phenomenon and error probability decay in polar coding. These quantities are called information rate and reliability parameter. Their definition depends on the channel under consideration. For a classical-quantum channel W, information rate is defined as the mutual information $I(X;B)$ and we denote it as $I(W)$. Reliability parameter is adopted as the fidelity between the possible channel outputs, i.e., reliability parameter 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 [39]. Some interesting results are obtained studying the information rate and reliability parameter 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 the 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, one can demonstrate the polarization phenomenon and construct a capacity-achieving coding strategy for polar codes. See the

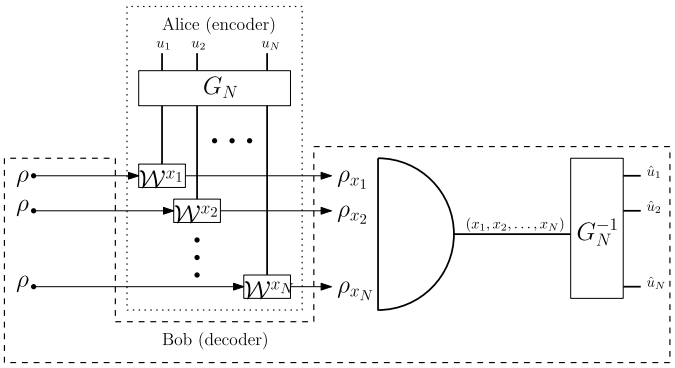


Fig. 4. Encoding and decoding protocols of polar coding applied to quantum reading paradigm.

following proposition for a formal formulation of the exposed ideas.

Proposition 4 ([39, 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 $i \in \{1,2,\ldots,N\}$ 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

$$P_e(N,K,\mathcal{A}) \leq 2 \sum_{i \in \mathcal{A}} \sqrt{F(W^{(i)}(0), W^{(i)}(1))}. \quad (11)$$

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

$$P_e(N,K) = o(2^{-N^\beta}), \quad (12)$$

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

We make use of Proposition 4 in the following section to show that polar codes can be applied in the task of quantum memory cell discrimination.

III. Polar Coding Scheme

This section considers 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.

Before presenting the results produced in this section, we need to explain the protocol in consideration. See the schematic of the protocol in Fig. 4. Polar codes are applied in the quantum reading framework by making the encoding and decoding processes corresponding to Alice and Bob's parts described in the previous section, respectively. First of all, Alice encodes the information bits by using a polar encoding scheme. The output of the encoding scheme, a codeword, is stored in the memory, and the first part ends. Suppose Bob probed each memory cell with a particular probing state, and he has access to the output states. Now, we proceed to the polar decoding process. A measurement strategy is constructed with low error probability. In particular, we will show a measurement strategy with error probability decaying exponentially in the square root of the code length. As a product of the measurements, an array of bits is obtained. We use this array of bits as noisy codeword to be fed in the polar decoder. After the polar decoder, the information bits are recovered with a low error probability. Our analysis using polar code in the quantum reading framework is given in the current section and the dependence of the probing state in Section IV.

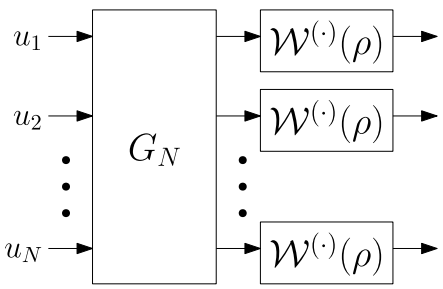


Fig. 5. Classical polar coding scheme.

A. Channel Polarization

1) Channel Combining: As mentioned before, for a fixed quantum probe state and $\mathcal{X} = \mathbb{Z}_2$, the elements $\mathcal{W}^x$, for $x \in \mathcal{X}$, 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. 5. 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 $u \in \mathcal{X}$. 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. 6.

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

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

We depict this scheme in Fig. 7.

Following the same procedure, we can derive the n-level channels. As described in [2], 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

$$G_N = R_N F^{\otimes n}, \quad (16)$$

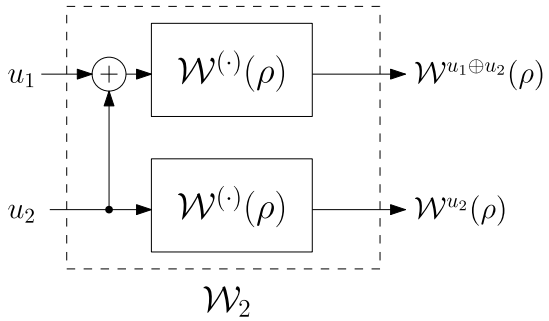


Fig. 6. Fundamental polar encoding block.

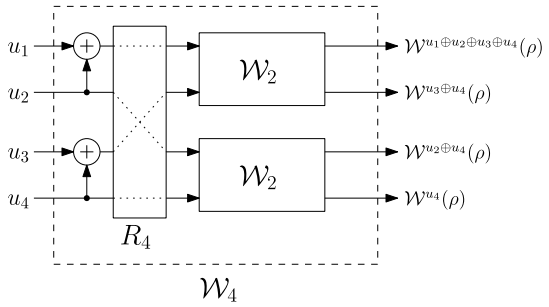


Fig. 7. Second level of polar coding scheme.

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

$$F = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \tag{17}$$

and R_N is the permutation matrix known as *bit-reversal* [2]. In particular, the second level combining function is

$$G_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}. \tag{18}$$

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

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 $(\mathcal{W}, \mathcal{W}) \to (\mathcal{W}^-, \mathcal{W}^+)$, 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

$$\mathcal{W}^-\colon u_1 \in \mathcal{X} \mapsto \mathcal{W}^{-,u_1}(\rho) \in \mathcal{D}(\mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2}), \tag{19a}$$

$$\mathcal{W}^+\colon u_2 \in \mathcal{X} \mapsto \mathcal{W}^{+,u_2}(\rho) \in \mathcal{D}(\mathcal{H}_{U_1} \otimes \mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2}), \tag{19b}$$

with

$$\mathcal{W}^{-,u_1}(\rho) = \sum_{u_2 \in \mathcal{X}} p_U(u_2) \mathcal{W}^{u_1 \oplus u_2}(\rho) \otimes \mathcal{W}^{u_2}(\rho) \tag{19c}$$

and

$$\mathcal{W}^{+,u_2}(\rho) = \sum_{u_1 \in \mathcal{X}} p_U(u_1)\,(u_1)\,u_1 \otimes \mathcal{W}^{u_1 \oplus u_2}(\rho) \otimes \mathcal{W}^{u_2}(\rho). \tag{19d}$$

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 $(\mathcal{W}, \mathcal{W}) \to (\mathcal{W}^-, \mathcal{W}^+)$ for some channels satisfying Eq. (19). Then the following rule holds for information rates:

$$I(\mathcal{W}^-)_\rho + I(\mathcal{W}^+)_\rho \leq 2I(\mathcal{W})_\rho \tag{20}$$

$$I(\mathcal{W}^+)_\rho \geq I(\mathcal{W})_\rho. \tag{21}$$

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

$$I(\mathcal{W}^-)_\rho + I(\mathcal{W}^+)_\rho = I(U_1; B_1B_2) + I(U_2; B_1B_2U_1) \tag{22a}$$

$$\stackrel{(i)}{=} I(U_1; B_1B_2) + I(U_2; B_1B_2|U_1) \tag{22b}$$

$$\stackrel{(ii)}{=} I(X_1X_2; B_1B_2) \tag{22c}$$

$$\stackrel{(iii)}{=} I(X_1; B_1) + I(X_2; B_2) \tag{22d}$$

$$= I(X_1; B_1) + I(\mathcal{W}). \tag{22e}$$

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 bijective 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 Ber$(p^2 + (1-p)^2)$ and X_2 is Ber(p). Then, it is possible to bound $I(X_2; B_2) - I(X_1; B_1)$ by the following inequality, (23) shown at the bottom of the page.

We have used the bounds

$$I(X_2; B_2) \geq -\log \mathrm{Tr}\{(p_{X_2}(0)\sqrt{\mathcal{W}^0(\rho)} + p_{X_2}(1)\sqrt{\mathcal{W}^1(\rho)})^2\}, \tag{24}$$

from [19, Prop. 1], (25a) and (25b) shown at the bottom of the next page, where (26), shown at the bottom of the next page, from [30, Thm. 3]. Calling the RHS of Eq. (23) by $f(p)$ and analyzing its first and second derivatives, 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 $I(\mathcal{W}^+)_\rho = I(U_2; B_1B_2U_1) \geq I(U_2; B_2) = I(X_2; B_2) = I(\mathcal{W})$. □

$$\begin{aligned} I(X_2; B_2) - I(X_1; B_1) \geq\ & -\log\{p^2 + (1-p)^2\} - \log\{p(1-p)\} - \log \mathrm{Tr}\{\sqrt{\mathcal{W}^0(\rho)}\sqrt{\mathcal{W}^1(\rho)}\} \\ & -2\sqrt{(1 - F(\mathcal{W}^0(\rho), \mathcal{W}^1(\rho))^2) 2p(1-p)(p^2 + (1-p)^2)} \end{aligned} \tag{23}$$

Remark 6: Proposition 5 shows that the polarization phenomenon can also be derived in quantum reading paradigm. However, it is worth mention the importance of the probe state used during the process. As explicitly shown in Section IV, 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. [29] relates the reliability parameter of the synthesized channels to the reliability parameter of the original channels.

Proposition 7 ([29, Prop. 4]): Consider the transformation $(\mathcal{W}, \mathcal{W}) \to (\mathcal{W}^-, \mathcal{W}^+)$ for some channels satisfying Eq. (19). Then the following rule holds for the reliability parameter:

$$Z(\mathcal{W}^+)_\rho = Z(\mathcal{W})_\rho^2 \tag{27}$$

$$Z(\mathcal{W}^-)_\rho \leq 2Z(\mathcal{W})_\rho - Z(\mathcal{W})_\rho^2. \tag{28}$$

Now, we can extend the splitting analysis to an N-fold combining of $\mathcal{W}$. Wilde and Guha have shown how to extend the previous characterization of synthesized channels [39]. 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 $i \in \{1, 2, \ldots, N\}$. The i-th synthesized channel is given by the map

$$\mathcal{W}_N^{(i)} : u_i \mapsto \mathcal{W}^{(i),u_i}(\rho^{\otimes N}), \tag{29}$$

where (30) and (31), shown at the bottom of the page.

From Eq. (30), we see that the description of $\mathcal{W}_N^{(i)}$ supposes the knowledge of previous inputs 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. [2], [39]. 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 information rate, reliability parameter, and error probability in the asymptotic scenario. See the next subsection.

3) Information Rate, Reliability Parameter, and Error Probability: The first result present in this section describes the connection between information rate and reliability parameter. It shows, as expected, that the information rate $I(\mathcal{W})_\rho \to 0$ (or $I(\mathcal{W})_\rho \to h(p)$) when the reliability parameter $Z(\mathcal{W})_\rho \to 2\sqrt{p(1-p)}$ (or $Z(\mathcal{W})_\rho \to 0$).

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

$$I(\mathcal{W})_\rho \geq h(p) - \log(1 + Z(\mathcal{W})_\rho), \tag{32}$$

$$I(\mathcal{W})_\rho \leq \sqrt{4p(1-p) - Z(\mathcal{W})_\rho^2}. \tag{33}$$

Proof: The first inequality follows from [16, Thm. C.1]. For the second one, we need to use the following inequality derived from Theorem 3 in Ref. [30]:

$$I(\mathcal{W})_\rho \leq H(\sigma), \tag{34}$$

where

$$\sigma = \begin{pmatrix} p & \frac{Z(\mathcal{W})_\rho}{2} \\ \frac{Z(\mathcal{W})_\rho}{2} & 1-p \end{pmatrix}. \tag{35}$$

Thus

$$I(\mathcal{W})_\rho \leq h\Big(\frac{1}{2}(1 - \sqrt{1 - 4p(1-p) + Z(\mathcal{W})_\rho^2})\Big) \tag{36}$$

$$\leq \sqrt{4p(1-p) - Z(\mathcal{W})_\rho^2}. \tag{37}$$

□

Computing the reliability parameter for asymmetric channels can be a demanding task. A method to solve this problem is to show a relation between the asymmetric channel of interest and a particular symmetric channel. Indeed, we are going to use this approach to show that there is a bigger symmetric channel that is equal to the asymmetric channel of interest by tracing over a particular subsystem. Since these two channels are connected, we can use some properties of symmetric channels in order to derive conclusions over the

$$I(X_1; B_1) \leq H(\sigma) = h\Big(\frac{1}{2}(1 - \sqrt{1 - 4(1 - F(\mathcal{W}^0(\rho), \mathcal{W}^1(\rho))^2)2p(1-p)(p^2 + (1-p)^2)})\Big) \tag{25a}$$

$$\leq 2\sqrt{(1 - F(\mathcal{W}^0(\rho), \mathcal{W}^1(\rho))^2)2p(1-p)(p^2 + (1-p)^2)} \tag{25b}$$

$$\sigma = \begin{pmatrix} p_{X_2}(0) & \sqrt{p_{X_2}(0)p_{X_2}(1)F(\mathcal{W}^0(\rho), \mathcal{W}^1(\rho))} \\ \sqrt{p_{X_2}(0)p_{X_2}(1)F(\mathcal{W}^0(\rho), \mathcal{W}^1(\rho))} & p_{X_2}(1) \end{pmatrix} \tag{26}$$

$$\mathcal{W}^{(i),u_i}(\rho^{\otimes N}) = \sum_{u_1^{i-1} \in \mathcal{X}^{i-1}} p_{U^{i-1}}(u_1^{i-1}) \left(u_1^{i-1}\right) u_1^{i-1} \otimes \overline{\mathcal{W}}^{u_1^i}(\rho^{\otimes N}), \tag{30}$$

$$\overline{\mathcal{W}}^{u_1^i}(\rho^{\otimes N}) = \sum_{u_{i+1}^N \in \mathcal{X}^{N-i}} p_{U_{i+1}^N}(u_{i+1}^N) \mathcal{W}^{u^N G_N}(\rho^{\otimes N}) \tag{31}$$

asymmetric channel. A similar approach has been followed for classical polar codes in Ref. [20]. Let $\tilde{U}_i \sim \text{Ber}(\frac{1}{2})$, $U_j \sim \text{Ber}(p)$ for all $i,j = 1,\ldots,N$. Assume that the output system of the classical-quantum channel $\tilde{\mathcal{W}}$ is given by $\mathcal{D}(\mathcal{H}_{\tilde{X}} \otimes \mathcal{H}_B)$, where $\sigma \in \mathcal{D}(\mathcal{H}_{\tilde{X}} \otimes \mathcal{H}_B)$ can be written as

$$\sigma = \sum_{\tilde{x} \in \mathcal{X}} (\tilde{x})\, \tilde{x}^{\tilde{X}} \otimes \rho_{\tilde{x}}^{B}. \tag{38}$$

Notice that $\tilde{X}$ represents the classical input system in the same way as the system X associated with the quantum channel $\mathcal{W}$. Therefore, the associated quantum systems are the same. We keep different notations 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 $u \in \mathcal{U}$. Then $\tilde{\mathcal{W}}$ is the map given by

$$\begin{aligned} &\tilde{\mathcal{W}}\colon \tilde{U} \to \mathcal{D}(\mathcal{H}_{\tilde{X}} \otimes \mathcal{H}_B) \\ &\quad \tilde{u} \mapsto \tilde{\mathcal{W}}^{\tilde{u}}(\rho) := \sum_{u \in \mathcal{U}} p_U(u)\, (\tilde{u} \oplus u)\, \tilde{u} \oplus u^{\tilde{X}} \otimes \mathcal{W}^{u}(\rho)^{B}. \end{aligned} \tag{39}$$

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 $\mathcal{H}_{\tilde{X}}$ part. The next step is to describe the synthesized channels generated in polar coding over $\tilde{\mathcal{W}}$. Following the characterization given in Eq. (29), we have that (40), shown at the bottom of the page.

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 $\tilde{\mathcal{W}}_N^{(i)}\colon \tilde{U}_i \to \mathcal{D}(\mathcal{H}_{\tilde{U}_1^{i-1}} \otimes \mathcal{H}_{B^N} \otimes \mathcal{H}_{\tilde{X}^N})$, where $\tilde{U}_i \sim \text{Ber}(\frac{1}{2})$ for all $i = 1,\ldots,N$. Then, $\mathcal{W}_N^{(i),u_i}(\rho^N) = \text{Tr}_{\tilde{X}^N}\{\tilde{\mathcal{W}}_N^{(i),u_i}(\rho^N)\,(0)\,0^{X^N}\}$.

Proof: Let $\tilde{\rho}^{(i)}$ be the joint input-output density operator of $\tilde{\mathcal{W}}_N^{(i),\tilde{u}_i}$. Thus, (41a)–(41c) shown at the bottom of the page.

Defining $\tilde{x}^N = (\tilde{u}^N \oplus u^N)G_N$, it is possible to rearrange the sums as (42), shown at the bottom of the page.

Now, we can see that $\text{Tr}_{\tilde{X}^N}\{\tilde{\rho}^{(i)}\,(0)\,0^{X^N}\} = \sum_{u_1^i} p_{U_1^i}(u_1^i)\,(u_1^i)\,u_1^i \otimes \Big(\sum_{u_{i+1}^N} p_{U_{i+1}^N}(u_{i+1}^N)\mathcal{W}^{u^N G_N}(\rho^N)\Big)$, which is the joint input-output state of $\mathcal{W}_N^{(i),u_i}(\rho^N)$. □

Proposition 10: Let $\mathcal{W}_N^{(i)}$ and $\tilde{\mathcal{W}}_N^{(i)}$ be the synthesized quantum channels described in Eq. (29) and Proposition 9, respectively. Then $Z(\tilde{\mathcal{W}}_N^{(i)})_\rho = Z(\mathcal{W}_N^{(i)})_\rho$.

Proof: From the definition of reliability parameter for symmetric classical-quantum channels in Ref. [39] and Eq (42), we have that, (43) and (44) shown at the bottom of the page, where, (45) and (46) shown at the bottom of the next page.

$$\tilde{\mathcal{W}}_N^{(i),\tilde{u}_i}(\rho^N) = \sum_{\tilde{u}_1^{i-1}} \frac{1}{2^{i-1}}\,(\tilde{u}_1^{i-1})\,\tilde{u}_1^{i-1} \otimes \Big(\sum_{u^N} p_{U^N}(u^N)\mathcal{W}^{u^N G_N}(\rho^N) \otimes \Big(\sum_{\tilde{u}_{i+1}^N} \frac{1}{2^{N-i}}\,((\tilde{u}^N \oplus u^N)G_N)\,(\tilde{u}^N \oplus u^N)G_N\Big)\Big). \tag{40}$$

$$\begin{aligned} \tilde{\rho}^{(i)} &= \frac{1}{2}\sum_{\tilde{u}_i \in \mathcal{X}} (\tilde{u}_i)\,\tilde{u}_i \otimes \tilde{\mathcal{W}}_N^{(i),\tilde{u}_i}(\rho^N) && \text{(41a)} \\ &= \sum_{\tilde{u}^N} \frac{1}{2^N}\,(\tilde{u}_1^i)\,\tilde{u}_1^i \otimes \Big(\sum_{u^N} p_{U^N}(u^N)\mathcal{W}^{u^N G_N}(\rho^N) \otimes \big((\tilde{u}^N \oplus u^N)G_N\big)\,(\tilde{u}^N \oplus u^N)G_N\Big) && \text{(41b)} \\ &= \sum_{u^N} p_{U^N}(u^N)\mathcal{W}^{u^N G_N}(\rho^N) \otimes \Big(\sum_{\tilde{u}^N} \frac{1}{2^N}\,(\tilde{u}_1^i)\,\tilde{u}_1^i \otimes \big((\tilde{u}^N \oplus u^N)G_N\big)\,(\tilde{u}^N \oplus u^N)G_N\Big) && \text{(41c)} \end{aligned}$$

$$\begin{aligned} \tilde{\rho}^{(i)} &= \sum_{u^N} p_{U^N}(u^N)\mathcal{W}^{u^N G_N}(\rho^N) \otimes \Big(\sum_{\substack{\tilde{x}^N \\ [u^N \oplus \tilde{x}^N G_N]_1^i = \tilde{u}_1^i}} \frac{1}{2^N}\,([u^N \oplus \tilde{x}^N G_N]_1^i)\,[u^N \oplus \tilde{x}^N G_N]_1^i \otimes (\tilde{x}^N)\,\tilde{x}^N\Big) \\ &= \sum_{\tilde{x}^N} \frac{1}{2^N}\,(\tilde{x}^N)\,\tilde{x}^N \otimes \Big(\sum_{\tilde{u}_1^i} p_{U_1^i}(\tilde{u}_1^i \oplus [\tilde{x}^N G_N]_1^i)\,(\tilde{u}_1^i)\,\tilde{u}_1^i \otimes \mathcal{W}^{[(\tilde{u}_1^i, 0_{i+1}^N)G_N]_1^i \oplus \tilde{x}_1^i}(\rho^i) \\ &\otimes \Big(\sum_{u_{i+1}^N} p_{U_{i+1}^N}(u_{i+1}^N)\mathcal{W}^{[(0_1^i, u_{i+1}^N)G_N]_{i+1}^N}(\rho^{N-i})\Big)\Big). \end{aligned} \tag{42}$$

$$\begin{aligned} Z(\tilde{\mathcal{W}}_N^{(i)})_\rho &= F(\tilde{\mathcal{W}}_N^{(i),0}(\rho^N), \tilde{\mathcal{W}}_N^{(i),1}(\rho^N)) && \text{(43)} \\ &= \sum_{\tilde{x}^N}\sum_{\tilde{u}_1^{i-1}} \frac{p_{U_1^{i-1}}(\tilde{u}_1^{i-1} \oplus [\tilde{x}^N G_N]_1^{i-1})\sqrt{p_U(0 \oplus [\tilde{x}^N G_N]_i)p_U(1 \oplus [\tilde{x}^N G_N]_i)}}{2^{N-1}} F(w_0^{(i)}, w_1^{(i)}) && \text{(44)} \end{aligned}$$

Let $x^N = \tilde{x}^N G_N$ and $u_1^{i-1} = \tilde{u}_1^{i-1} \oplus [\tilde{x}^N G_N]_1^{i-1} = \tilde{u}_1^{i-1} \oplus x_1^{i-1}$. Since the map $(\tilde{x}_1^N, \tilde{u}_1^{i-1}) \mapsto (x_1^N, u_1^{i-1})$ is a bijection, we see that (47), shown at the bottom of the page, where $\overline{\mathcal{W}}^{u_1^i}(\rho^{\otimes N})$ is defined in Eq. (30). □

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. (29) and $Z(X|Y)$ be the reliability parameter 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 (48) and (49), shown at the bottom of the page.

Proof: Firstly, notice that

$$\begin{aligned} I(\tilde{\mathcal{W}}) &= I(\tilde{X}; \tilde{X} \oplus X, B) \\ &= H(\tilde{X} \oplus X, B) - H(\tilde{X} \oplus X, B|\tilde{X}) \\ &= 1 + H(B) - H(X,B) = 1 - H(X|B). \end{aligned} \tag{50}$$

Applying Eq. (50) to Proposition 4, we can deduce that

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{2^n} |\{i \in \mathcal{X}^n \colon Z(\mathcal{W}_N^{(i)})_\rho \le 2^{-2^{n\beta}}\}| \\ = 1 - H(X|B)_\rho, \end{aligned} \tag{51a}$$

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{2^n} |\{i \in \mathcal{X}^n \colon Z(\mathcal{W}_N^{(i)})_\rho \ge 1 - 2^{-2^{n\beta}}\}| \\ = H(X|B)_\rho. \end{aligned} \tag{51b}$$

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. (51) is derived

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{2^n} |\{i \in \mathcal{X}^n \colon Z(U_i|U_1^{i-1}) \le 2^{-2^{n\beta}}\}| \\ = 1 - H(X), \end{aligned} \tag{52a}$$

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{2^n} |\{i \in \mathcal{X}^n \colon Z(U_i|U_1^{i-1}) \ge 1 - 2^{-2^{n\beta}}\}| \\ = H(X). \end{aligned} \tag{52b}$$

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

$$A = \{i \colon Z(\mathcal{W}_N^{(i)})_\rho \le 2^{-2^{n\beta}}\}, \tag{53a}$$

$$B = \{i \colon Z(\mathcal{W}_N^{(i)})_\rho \ge 1 - 2^{-2^{n\beta}}\}, \tag{53b}$$

$$C = \{i \colon Z(U_i|U_1^{i-1}) \le 2^{-2^{n\beta}}\}, \tag{53c}$$

$$D = \{i \colon Z(U_i|U_1^{i-1}) \ge 1 - 2^{-2^{n\beta}}\}. \tag{53d}$$

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

$$\lim_{n\to\infty} \frac{|A| + |B|}{2^n} = \lim_{n\to\infty} \frac{|C| + |D|}{2^n} = 1. \tag{54}$$

Therefore, the claim is derived from

$$\lim_{n\to\infty} \frac{|B \cup C|}{2^n} = \lim_{n\to\infty} \frac{|B| + |C|}{2^n} = 1 - I(X;B)_\rho. \tag{55}$$

and

$$\lim_{n\to\infty} \frac{|A \cap D|}{2^n} = 1 - \lim_{n\to\infty} \frac{|B \cup C|}{2^n} = I(X;B)_\rho. \tag{56}$$

□

As can be seen in the theorem statement and elaborated in the proof, we had to impose an additional constraint on the reliability parameter of U_i given the previous U_1^{i-1}

$$w_0^{(i)} = \mathcal{W}^{[(\tilde{u}_1^{i-1}, 0_i^N)G_N]_1^i \oplus \tilde{x}_1^i}(\rho^i) \sum_{u_{i+1}^N} p_{U_{i+1}^N}(u_{i+1}^N) \mathcal{W}^{[(0_1^i, u_{i+1}^N)G_N]_{i+1}^N}(\rho^{N-i}), \tag{45}$$

$$w_1^{(i)} = \mathcal{W}^{[(\tilde{u}_1^{i-1}, 1, 0_{i+1}^N)G_N]_1^i \oplus \tilde{x}_1^i}(\rho^i) \sum_{u_{i+1}^N} p_{U_{i+1}^N}(u_{i+1}^N) \mathcal{W}^{[(0_1^i, u_{i+1}^N)G_N]_{i+1}^N}(\rho^{N-i}) \tag{46}$$

$$\begin{aligned} Z(\tilde{\mathcal{W}}_N^{(i)})_\rho &= \sum_{x^N} \sum_{u_1^{i-1}} \frac{p_{U_1^{i-1}}(u_1^{i-1}) \sqrt{p_U(0)p_U(1)}}{2^{N-1}} F(\mathcal{W}^{(u_1^{i-1},0)}(\rho^i) \sum_{u_{i+1}^N} p_{U_{i+1}^N}(u_{i+1}^N) \mathcal{W}^{[(0_1^i, u_{i+1}^N)G_N]_{i+1}^N}(\rho^{N-i}), \\ &\quad \mathcal{W}^{(u_1^{i-1},1)}(\rho^i) \sum_{u_{i+1}^N} p_{U_{i+1}^N}(u_{i+1}^N) \mathcal{W}^{[(0_1^i, u_{i+1}^N)G_N]_{i+1}^N}(\rho^{N-i})) \\ &= 2\sqrt{p_U(0)p_U(1)} \sum_{u_1^{i-1}} p_{U_1^{i-1}}(u_1^{i-1}) F(\overline{\mathcal{W}}^{(u_1^{i-1},0)}(\rho^{\otimes N}), \overline{\mathcal{W}}^{(u_1^{i-1},1)}(\rho^{\otimes N})) \\ &= 2\sqrt{p_U(0)p_U(1)} F(\mathcal{W}^{(i),0}(\rho^{\otimes N}), \mathcal{W}^{(i),1}(\rho^{\otimes N})) = Z(\mathcal{W}_N^{(i)})_\rho \end{aligned} \tag{47}$$

$$\lim_{n\to\infty} \frac{1}{2^n} |\{i \in \mathcal{X}^n \colon Z(\mathcal{W}_N^{(i)})_\rho \le 2^{-2^{n\beta}} \text{ and } Z(U_i|U_1^{i-1}) \ge 1 - 2^{-2^{n\beta}}\}| = I(X;B)_\rho, \tag{48}$$

$$\lim_{n\to\infty} \frac{1}{2^n} |\{i \in \mathcal{X}^n \colon Z(\mathcal{W}_N^{(i)})_\rho \ge 1 - 2^{-2^{n\beta}} \text{ and } Z(U_i|U_1^{i-1}) \le 2^{-2^{n\beta}}\}| = 1 - I(X;B)_\rho \tag{49}$$

to derive our result. This is because we are dealing with asymmetric quantum reading. For symmetric quantum reading, the constraint is not needed.

Remark 12: Subsections III-A.1 and III-A.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, 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. [39].

B. Polar Coding and Decoding Schemes

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 latter 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, in our case, the latter strategy is more suitable. See Ref. [20] for more explanations.

The construction of a codeword is done as follows. The source generates an 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 $x^N = u^N G_N$. Clearly, 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 set $\mathcal{A}$. A common approach is to select the coordinates with smallest reliability parameters. 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 constraint to characterize the asymptotic rates in polar coding. If any of these two constraints are not satisfied, we decrease the cardinality of $\mathcal{A}$ down to when they are satisfied.

The functions λ_i used in this paper are 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

$$p_{\Lambda_i}[\lambda_i(u_1^{i-1})] = p_{U_i|U_1^{i-1}}(1|u_1^{i-1}), \tag{57}$$

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 [20], 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

$$p_{\tilde{\Lambda}_i}[u] = \frac{1}{2}, \tag{58}$$

for $u = 0, 1$. Secondly, the proposed strategy for defining the set $\mathcal{A}$ is applicable for symmetric channels discrimination. We only need to drop the constraint on $Z(U_i|U_1^{i-1})$. Approximation techniques in symmetric polar codes [36] can compose the decoding scheme 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

$$\hat{u}_i = \lambda_i^{-1}(\hat{u}_1^{i-1}). \tag{59}$$

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 here. First of all, this makes our decoding strategy and analysis completely different from any decoding method used for classical polar codes. Secondly, because of the constructive approach adopted in the paper, 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 cells and having desirable error probabilities in Proposition 13.

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

$$\rho_x^{UB} = \sum_{u\in\mathcal{U}} p_U(u)\,(u)\,u^U \otimes \rho_{x,u}^B, \tag{60}$$

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, we can construct a POVM $\{\Lambda_x^{UB}\}_{x\in\mathcal{X}}$ satisfying

$$1 - \sum_{x\in\mathcal{X}} p_X(x)\mathrm{Tr}\{\Lambda_x\rho_x\} < \frac{1}{2}Z(\mathcal{W}). \tag{61}$$

Proof: To derive our claim, we need to invoke a result from Barnum and Knill [6]. 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

$$\begin{aligned} P_{succ} &= \sum_{u\in\mathcal{U}} p_U(u)\mathrm{Tr}\{\rho_u^{\otimes N}\Lambda_u^*\} \\ &\geq 1 - \sum_{u\neq v} \sqrt{p_U(u)p_U(v)}F(\rho_u^{\otimes N}, \rho_v^{\otimes N}). \end{aligned} \tag{62}$$

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

$$\begin{aligned} P_{err} &= 1-\sum_{u\in\mathcal{U}} p_U(u)\text{Tr}\{\rho_u\Lambda_u\} \\ &\leq \sum_{u\neq v}\sqrt{p_U(u)p_U(v)}F(\rho_u,\rho_v). \end{aligned} \tag{63}$$

Now, we return to our proof. Assume that $\mathcal{W}^u\colon x\to\rho_{x,u}$ is a cq channel, where $x\in\mathcal{X}$ is a realization of a discrete random variable X. The result in Eq. (63) says that there exist POVMs $\{\Lambda_{x,u}\}_{x\in\mathcal{X}}$ satisfying

$$1-\sum_{x\in\mathcal{X}} p_X(x)\text{Tr}\{\rho_{x,u}\Lambda_{x,u}\} < \frac{1}{2}Z(\mathcal{W}^u). \tag{64}$$

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

$$\Lambda_x := \sum_{u\in\mathcal{U}} \Lambda_{x,u}\otimes(u)\,u, \tag{65}$$

we have that (66)–(68), shown at the bottom of the page.

The equality in (i) follows from the definitions of ρ_x and Λ_x in Eq. (60) and (65), respectively. The inequality in Eq. (64) is used to obtain (ii). The item (iii) is just the computation of the reliability parameter for $\mathcal{W}^u$. Lastly, we use in (iv) the fact that $F(\sum_x p_X(x)\,(x)\,x\otimes\sigma_x, \sum_x q_X(x)\,(x)\,x\otimes\tau_x) = \sum_x\sqrt{p_X(x)q_X(x)}F(\sigma_x,\tau_x)$. □

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 ([24, 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 $i\in\{1,\ldots,L\}$, and let $c>0$. Then the following quantum union bound holds

$$\begin{aligned} \text{Tr}\{\rho\} \quad & -\text{Tr}\{\Pi_{\Lambda_L}\cdots\Pi_{\Lambda_1}(\rho\otimes(\overline{0})\,\overline{0}_{P^L})\Pi_{\Lambda_1}\cdots\Pi_{\Lambda_L}\} \\ \leq \quad & (1+c)\text{Tr}\{(I-\Lambda_L)\rho\} \\ + \quad & (2+c+c^{-1})\sum_{i=2}^{L-1}\text{Tr}\{(I-\Lambda_i)\rho\} \\ + \quad & (2+c^{-1})\text{Tr}\{(I-\Lambda_1)\rho\}, \end{aligned} \tag{69}$$

where $|\overline{0}_{P_L}\rangle \equiv |0_{P_1}\rangle\otimes\cdots\otimes|0_{P_L}\rangle$ is an auxiliary state of L qubit probe systems and Π_{Λ_i} is a projector satisfying $\text{Tr}\{\Pi_{\Lambda_i}(\rho\otimes(\overline{0})\,\overline{0}_{P^L})\} = \text{Tr}\{\Lambda_i\rho\}$. In particular, for any set of positive semi-definite operators, there exists a set of projectors satisfying Eq. (69).

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 in the square root of the code length. This result motivates the use of polar codes in practical protocols devoted to discriminate quantum memory cells.

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

$$P_e(N,K,\mathcal{A}) \leq \gamma\sum_{i\in\mathcal{A}} Z_\rho(\mathcal{W}^{(i)}). \tag{70}$$

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

$$P_e(N,K) = o(2^{-N^\beta}), \tag{71}$$

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

Proof: First of all, suppose we have access to an auxiliary system $|\overline{0}_{P_N}\rangle = |0_{P_1}\rangle\otimes\cdots\otimes|0_{P_N}\rangle$. 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 (72)–(74), shown on the next page, where c is the constant in Lemma 14, and the 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 (75), shown on the next page.

Continuing, (76) and (77) shown on the next page, where the first equality follows from Eq. (31). For the second equality, we have used the fact that [39] $\sum_x p_X(x)\text{Tr}\{A_x\rho_x\} = \text{Tr}\Big\{\Big(\sum_x(x)\,x\otimes A_x\Big)\Big(\sum_{x'}p_X(x)\,(x')\,x'\otimes\rho_{x'}\Big)\Big\}$. Now, from the definition of synthesized channels present in Eq. (30) the next equality is produced (78), shown on the next page.

Now, the number of terms in the previous sum can be reduced by means of a simple observation. Since the frozen

$$1 \quad -\sum_{x\in\mathcal{X}} p_X(x)\text{Tr}\{\Lambda_x\rho_x\} \overset{(i)}{=} 1-\sum_{x\in\mathcal{X}} p_X(x)\sum_{u\in\mathcal{U}} p_U(u)\text{Tr}\{\Lambda_{x,u}\rho_{x,u}\} = \sum_{u\in\mathcal{U}} p_U(u)\sum_{x\in\mathcal{X}} p_X(x)\Big(1-\text{Tr}\{\Lambda_{x,u}\rho_{x,u}\}\Big) \tag{66}$$

$$\overset{(ii)}{<} \sum_{u\in\mathcal{U}} p_U(u)\frac{1}{2}Z(\mathcal{W}^u) \overset{(iii)}{=} \sum_{\substack{x,x'\in\mathcal{X}\\ x\neq x'}}\sqrt{p_X(x)p_X(x')}F\Big(\sum_{u\in\mathcal{U}} p_U(u)\,(u)\,u\otimes\rho_{x,u}, \sum_{u\in\mathcal{U}} p_U(u)\,(u)\,u\otimes\rho_{x',u}\Big) \tag{67}$$

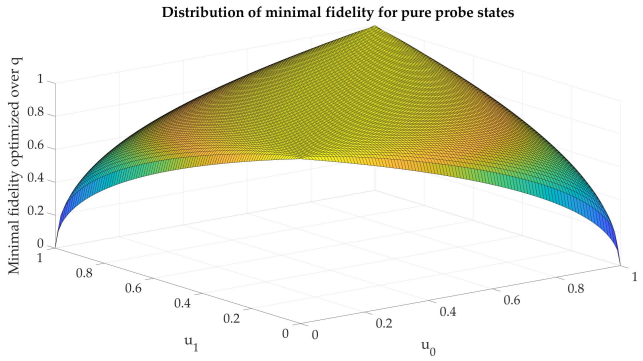
$$\overset{(iv)}{=} \sum_{\substack{x,x'\in\mathcal{X}\\ x\neq x'}}\sqrt{p_X(x)p_X(x')}F\Big(\rho_x,\rho_{x'}\Big) = \frac{1}{2}Z(\mathcal{W}) \tag{68}$$

$$
\begin{aligned}
P_e(N,K,\mathcal{A}) &= 1-\sum_{u^N} p_{U^N}(u^N)\mathrm{Tr}\{\Pi_{\Lambda_{u^N}}\mathcal{W}^{u^N}(\rho^N)\otimes\left(\overline{0}_{P_N}\right)\overline{0}_{P_N}\} && (72)\\
&= 1-\sum_{u^N} p_{U^N}(u^N)\mathrm{Tr}\{\Pi^{B^N}_{\Lambda_{u_1^{N-1}u_N}}\cdots\Pi^{B^N}_{\Lambda_{u_1}}\mathcal{W}^{u^N}(\rho^N)\otimes\left(\overline{0}_{P_N}\right)\overline{0}_{P_N}\Pi^{B^N}_{\Lambda_{u_1}}\cdots\Pi^{B^N}_{\Lambda_{u_1^{N-1}u_N}}\} && (73)\\
&\le \sum_{u^N} p_{U^N}(u^N)\Big((1+c)\mathrm{Tr}\Big\{(I-\Lambda^{B^N}_{u_1^{N-1}u_N})\mathcal{W}^{u^N}(\rho^N)\Big\}\\
&+ (2+c+c^{-1})\sum_{i=2}^{N-1}\mathrm{Tr}\Big\{(I-\Lambda^{B^N}_{u_1^{i-1}u_i})\mathcal{W}^{u^N}(\rho^N)\Big\}\\
&+ (2+c^{-1})\mathrm{Tr}\Big\{(I-\Lambda^{B^N}_{u_1})\mathcal{W}^{u^N}(\rho^N)\Big\}\Big) && (74)
\end{aligned}
$$

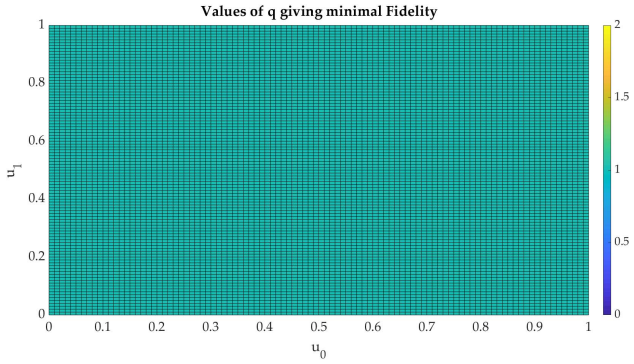
$$
\begin{aligned}
\text{RHS of Eq. (74)} &= (1+c)\sum_{u^N} p_{U^N}(u^N)\mathrm{Tr}\Big\{(I-\Lambda^{B^N}_{u_1^{N-1}u_N})\mathcal{W}^{u^N}(\rho^N)\Big\}\\
&+ (2+c+c^{-1})\sum_{i=2}^{N-1} p_U(u_i)\sum_{u_1^{i-1}} p_{U_1^{i-1}}(u_1^{i-1})\mathrm{Tr}\Big\{(I-\Lambda^{B^N}_{u_1^{i-1}u_i})\sum_{u_1^{i-1}} p_{U_{i+1}^N}(u_{i+1}^N)\mathcal{W}^{u^N}(\rho^N)\Big\}\\
&+ (2+c^{-1})\sum_{u_1} p_U(u_1)\mathrm{Tr}\Big\{(I-\Lambda^{B^N}_{u_1})\sum_{u_2^{i-1}} p_{U_2^N}(u_2^N)\mathcal{W}^{u^N}(\rho^N)\Big\} && (75)
\end{aligned}
$$

$$
\begin{aligned}
\text{RHS of Eq. (75)} &= (1+c)\sum_{u^N} p_{U^N}(u^N)\mathrm{Tr}\Big\{(I-\Lambda^{B^N}_{u_1^{N-1}u_N})\overline{\mathcal{W}}^{u_1^N}(\rho^N)\Big\}\\
&+ (2+c+c^{-1})\sum_{i=2}^{N-1} p_U(u_i)\sum_{u_1^{i-1}} p_{U_1^{i-1}}(u_1^{i-1})\mathrm{Tr}\Big\{(I-\Lambda^{B^N}_{u_1^{i-1}u_i})\overline{\mathcal{W}}^{u_1^i}(\rho^N)\Big\}\\
&+ (2+c^{-1})\sum_{u_1} p_U(u_1)\mathrm{Tr}\Big\{(I-\Lambda^{B^N}_{u_1})\overline{\mathcal{W}}^{u_1}(\rho^N)\Big\} && (76)\\
&= (1+c)\sum_{u^N} p_{U^N}(u^N)\mathrm{Tr}\Big\{\Big(\sum_{u_1^{N-1}}\left(u_1^{N-1}\right)u_1^{N-1}\otimes(I-\Lambda^{B^N}_{u_1^{N-1}u_N})\Big)\\
&\Big(\sum_{u_1^{N-1}} p_{U_1^{N-1}}(u_1^{N-1})\left(u_1^{N-1}\right)u_1^{N-1}\otimes\overline{\mathcal{W}}^{u_1^N}(\rho^N)\Big)\Big\}\\
&+ (2+c+c^{-1})\sum_{i=2}^{N-1} p_U(u_i)\mathrm{Tr}\Big\{\Big(\sum_{u_1^{i-1}}\left(u_1^{i-1}\right)u_1^{i-1}\otimes(I-\Lambda^{B^N}_{u_1^{i-1}u_i})\Big)\\
&\Big(\sum_{u_1^{i-1}} p_{U_1^{i-1}}(u_1^{i-1})\left(u_1^{i-1}\right)u_1^{i-1}\otimes\overline{\mathcal{W}}^{u_1^i}(\rho^N)\Big)\Big\}\\
&+ (2+c^{-1})\sum_{u_1} p_U(u_1)\mathrm{Tr}\Big\{(I-\Lambda^{B^N}_{u_1})\overline{\mathcal{W}}^{u_1}(\rho^N)\Big\} && (77)
\end{aligned}
$$

$$
\begin{aligned}
\text{RHS of Eq. (77)} &= (1+c)\sum_{u^N} p_{U^N}(u^N)\mathrm{Tr}\Big\{\Big(\sum_{u_1^{N-1}}\left(u_1^{N-1}\right)u_1^{N-1}\otimes(I-\Lambda^{B^N}_{u_1^{N-1}u_N})\Big)\mathcal{W}^{(N),u_N}(\rho^N)\Big\}\\
&+ (2+c+c^{-1})\sum_{i=2}^{N-1} p_U(u_i)\mathrm{Tr}\Big\{\Big(\sum_{u_1^{i-1}}\left(u_1^{i-1}\right)u_1^{i-1}\otimes(I-\Lambda^{B^N}_{u_1^{i-1}u_i})\Big)\mathcal{W}^{(i),u_i}(\rho^N)\Big\}\\
&+ (2+c^{-1})\sum_{u_1} p_U(u_1)\mathrm{Tr}\Big\{(I-\Lambda^{B^N}_{u_1})\mathcal{W}^{(1),u_1}(\rho^N)\Big\} && (78)
\end{aligned}
$$



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



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

Fig. 8. Analysis of fidelity for the pure probe state.

bits are obtained by using a classical function, there is no need to make a measurement on the frozen bits position; i.e., $\Lambda^{B^N}_{u_1^{i-1}u_i} = I$ for any $i \in \mathcal{A}^c$. This leads to

$$\begin{aligned} \text{RHS of Eq. (78)} &< \frac{(1+c+c^{-1})}{2} \sum_{i\in\mathcal{A}} Z_\rho(\mathcal{W}^{(i)}) \\ &= o(2^{-N^\beta}), \end{aligned} \tag{79}$$

where the definition of reliability parameter for synthesized channels have been applied. The fact that this sum equals to $o(2^{-N^\beta})$ is a consequence of Proposition 4. □

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 a pretty good measurement for the symmetric channels under consideration.

IV. Probe State Analysis

We are going to analyze how reliability parameter and information 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 for the quantity under consideration.

First of all, we adopt

$$|\psi\rangle = \sqrt{1-q}\,|0\rangle + e^{-i\phi}\sqrt{q}\,|1\rangle \tag{80}$$

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

$$|\Psi^{xz}\rangle = (Z^z X^x \otimes I)\,|\Phi\rangle\,, \tag{81}$$

where $z, x \in \mathcal{X}$, 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.

Remark 16: Before presenting some of the computations and analysis obtained from them, we need to explain the notation used here. As described in Section II-B, The evolution of an AD channel depends on the decay probability associated with the parameter u of the channel. Since we are dealing with two possible values of u, $u_0, u_1 = \{0, 1\}$, there are two values for the decay probability to analyze. These values of the decay probability can be any value $\gamma \in [0, 1]$. So, to not overcrowd our notation, computations, and figures, we have opted to use u_0, u_1 in this section instead of $\gamma(u_0), \gamma(u_1)$. Since we do not process the values of u_0 and u_1, there is no confusion in what follows.

Now, consider pure probe states. Substituting Eq. (80) into Eq. (3), it is possible to see that (82) and (83), shown at the bottom of the page.

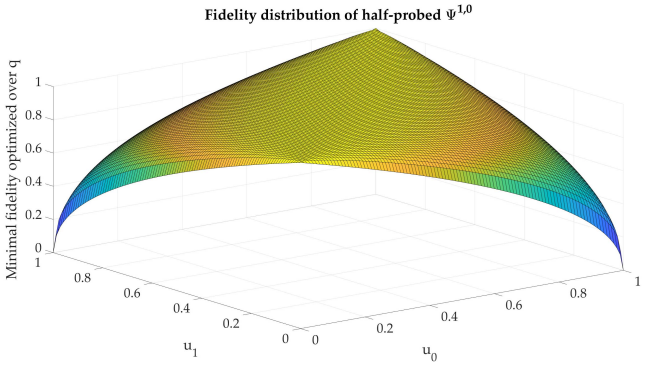
Now, the minimal value of the reliability parameter with respect to q can be computed. First of all, notice that the probability law p_X is arbitrary. Therefore, the conclusions derived below need to be independent of such a choice. So, we have opted to study a normalized version of the reliability parameter given by

$$\begin{aligned} \overline{Z}(\mathcal{W})_{(\psi)\psi} &= \min_q \frac{Z(\mathcal{W})_{(\psi)\psi}}{2\sqrt{q(1-q)}} \\ &= \min_q F(\mathcal{W}^{u_0}((\psi)\,\psi), \mathcal{W}^{u_1}((\psi)\,\psi)). \end{aligned} \tag{84}$$

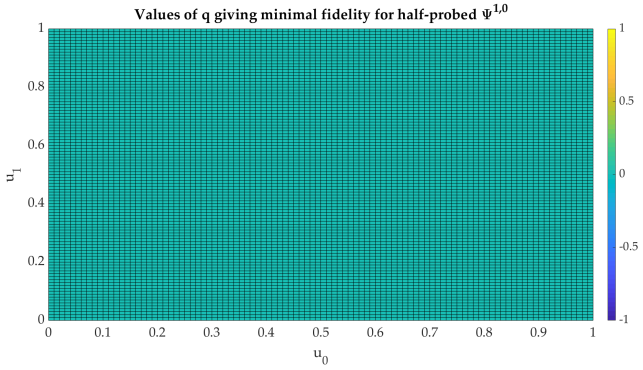
So, the procedure goes as follows. For every pair (u_0, u_1), we compute the value of q that minimizes the normalized

$$\mathcal{W}^{u_0}((\psi)\,\psi) = \begin{pmatrix} 1-(1-u_0)q & \sqrt{(1-u_0)q(1-q)} \\ \sqrt{(1-u_0)q(1-q)} & (1-u_0)q \end{pmatrix} \tag{82}$$

$$\mathcal{W}^{u_1}((\psi)\,\psi) = \begin{pmatrix} 1-(1-u_1)q & \sqrt{(1-u_1)q(1-q)} \\ \sqrt{(1-u_1)q(1-q)} & (1-u_1)q \end{pmatrix} \tag{83}$$

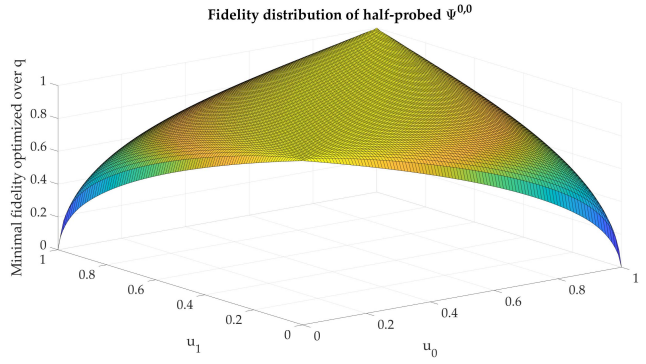


(a) Minimal value of fidelity w.r.t. q.

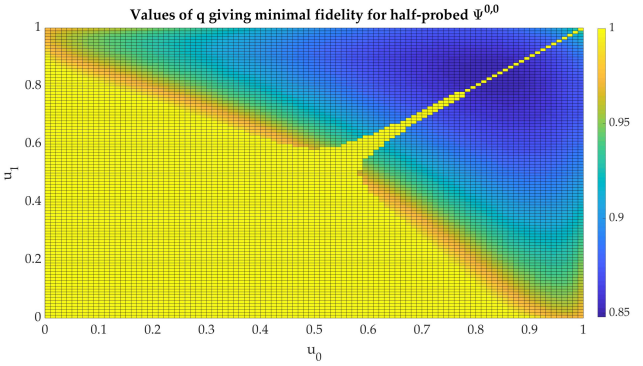


(b) Distribution of q giving the minimal value of fidelity.

Fig. 9. 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.

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

reliability parameter (or fidelity). Now, considering the pure probe state, it is shown in Fig. 8 the minimal values $\overline{Z}(\mathcal{W})_{(\psi)\psi}$. The value of q giving minimum fidelity is $q = 1$, which turns the optimal pure probe state to be the excited state $|1\rangle$.

Considering mixed probe states, the same analysis can be done. However, there is a difference on how to probe the channel. Since we have an entangled bipartite state, it can be used to probe the channel once or twice. The channel output for probing once is given by (85), shown at the bottom of the next page, where $i = 0, 1$ and δ_{ab} is the Kronecker delta, and for probing twice we have (86), shown at the bottom of the next page.

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. (85) and the second to Eq. (86). Our first description is over the half-probed. See Figs. 9 and 10 below. Notice that we have shown just the instance $z = 0$, since the result for $z = 1$ is the same.

Two conclusions can be obtained from Figs. 9 and 10. Firstly, the shape of the fidelity of half-probed state $|\Psi^{10}\rangle$ is the same as for the pure probe state. This is clear from the fact that the value of q giving minimum fidelity is $q = 0$, which turns the probe state to be $|1\rangle$[3]. As the second point, we see in

[3]For $q = 0$, we have that $|\Psi^{10}\rangle = X \otimes I\,|\Phi\rangle = |10\rangle$.

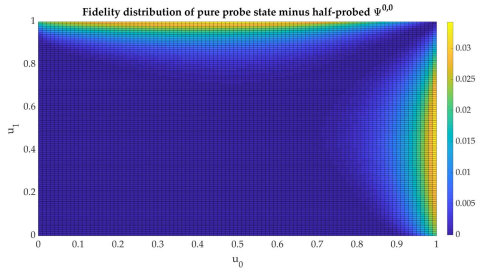
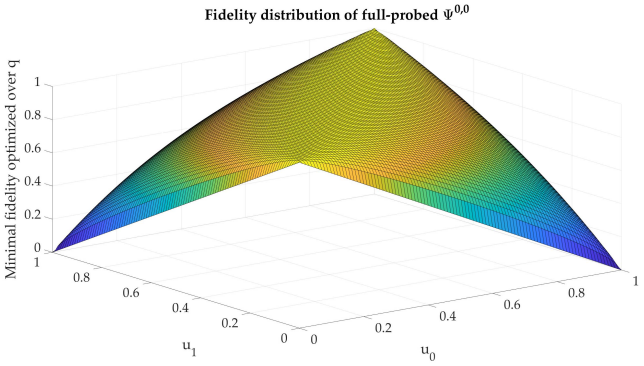


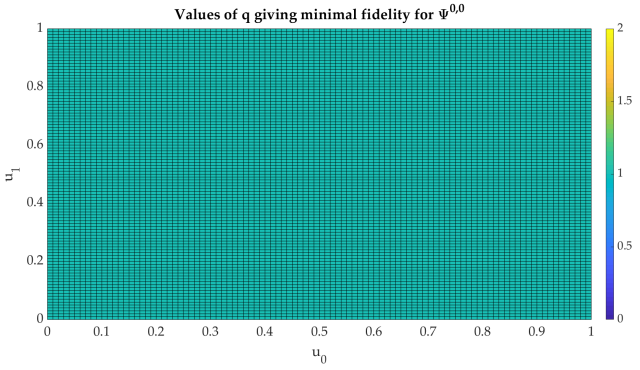
Fig. 11. Fidelity of pure state $|1\rangle$ minus fidelity of half-probed state $|\Psi^{0,0}\rangle$.

Figs. 9 and 10 that half-probing with the state $|\Psi^{00}\rangle$ can give some improvement in the channel discrimination task. So, even though the shape of these two figures looks similar, they are not the same. The improvement obtained for probing with the state $|\Psi^{00}\rangle$ is obtained when $q \neq 1$. However, the difference is not so significant, making the use of $|\Psi^{00}\rangle$ justifiable only if the entanglement cost is negligible.

We have shown the distribution of q, for every pair (u_0, u_1), which minimizes the fidelity for three cases, a pure probe state and two half-probed states. The state $|1\rangle$ is the optimal probe candidate for the pure probed state over every pair (u_0, u_1). The same result is produced for half-probed state $|\Psi^{10}\rangle$. This means that the optimization strategy gives the value $q = 0$ for every pair (u_0, u_1), which implies the same probing strategy as for the pure probed state, and, therefore,

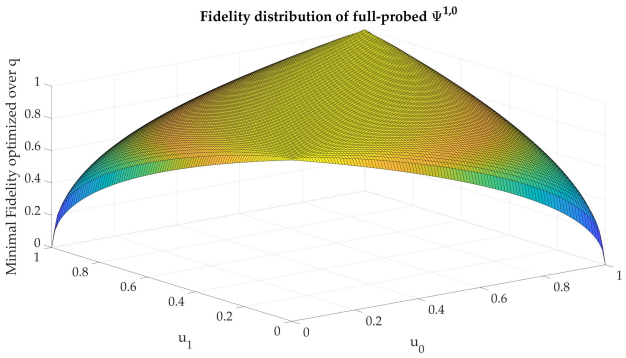


(a) Minimal value of fidelity w.r.t. q.

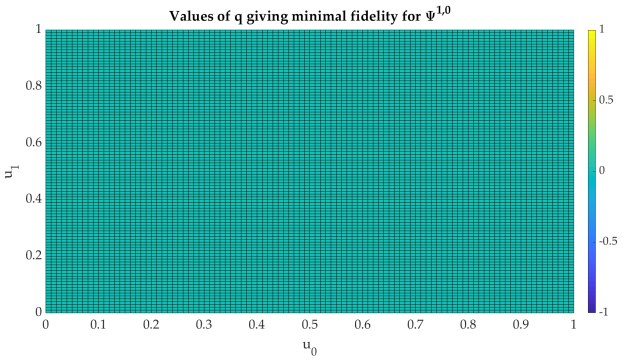


(b) Distribution of q giving the minimal value of fidelity.

Fig. 12. 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.

Fig. 13. Analysis of fidelity for the full-probed state $|\Psi^{1,0}\rangle$.

same fidelity. One can see this in Figs. 8 and 9. Probing with half-probe state $|\Psi^{00}\rangle$ we obtain a different result. As one can see in Fig. 10b, there are pairs of (u_0, u_1) where the optimal value of q is different of 1. This region is the one we obtain improvements when compared with the pure probed state and half-probed state $|\Psi^{10}\rangle$. Such improvements are not so significant, as Fig. 11 shows. Notice that the caption of Fig. 11 states "pure probe state", but the same result holds by replacing the pure probe state with the half-probed state $|\Psi^{10}\rangle$, since the optimal fidelities of these two probing strategies are the same.

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. 12 shows that $q = 1$ gives the minimal value of fidelity in this case. This leads to the fidelity having values equal to the square

$$
\begin{aligned}
\mathcal{W}^{u_i} \otimes \mathrm{id}_{A_2}((\Psi^{xz})\,\Psi^{xz}) \;=\; & \Big[\Big((1-q)\delta_{0x}\,(0)\,0_{A_2} + q\delta_{1x}\,(1)\,1_{A_2}\Big) \\
& + u_i\Big((1-q)\delta_{1x}\,(0)\,0_{A_2} + q\delta_{0x}\,(1)\,1_{A_2}\Big)\Big]\,(0)\,0_{A_1} \\
& + (\delta_{0z} - \delta_{1z})\sqrt{(1-u_i)q(1-q)}\Big[\delta_{0x}\,(0)\,1_{A_2} + \delta_{1x}\,(1)\,0_{A_2}\Big]\,(0)\,1_{A_1} \\
& + (\delta_{0z} - \delta_{1z})\sqrt{(1-u_i)q(1-q)}\Big[\delta_{1x}\,(0)\,1_{A_2} + \delta_{0x}\,(1)\,0_{A_2}\Big]\,(1)\,0_{A_1} \\
& + (1-u_i)\Big[(1-q)\delta_{1x}\,(0)\,0_{A_2} + q\delta_{0x}\,(1)\,1_{A_2}\Big]\,(1)\,1_{A_1} \qquad (85) \\
\mathcal{W}^{u_0} \otimes \mathcal{W}^{u_1}((\Psi^{xz})\,\Psi^{xz}) \;=\; & [(1-q)\delta_{0x} + u_0(1-q)\delta_{1x}]\,(0)\,0_{A_1} \otimes (0)\,0_{A_2} \\
& + q(u_0\delta_{0x} + \delta_{1x})\,(0)\,0_{A_1} \otimes [u_1\,(0)\,0_{A_2} + (1-u_1)\,(1)\,1_{A_2}] \\
& + (-1)^z\sqrt{(1-u_0)(1-u_1)q(1-q)}[\delta_{0x}\,(0)\,1_{A_1} + \delta_{1x}\,(1)\,0_{A_1}] \otimes (0)\,1_{A_2} \\
& + (-1)^z\sqrt{(1-u_0)(1-u_1)q(1-q)}[\delta_{0x}\,(1)\,0_{A_1} + \delta_{1x}\,(0)\,1_{A_1}] \otimes (1)\,0_{A_2} \\
& + (1-u_0)q\delta_{0x}\,(1)\,1_{A_1} \otimes [u_1\,(0)\,0_{A_2} + (1-u_1)\,(1)\,1_{A_2}] \\
& + (1-u_0)(1-q)\delta_{1x}\,(1)\,1_{A_1} \otimes (0)\,0_{A_2} \qquad (86)
\end{aligned}
$$

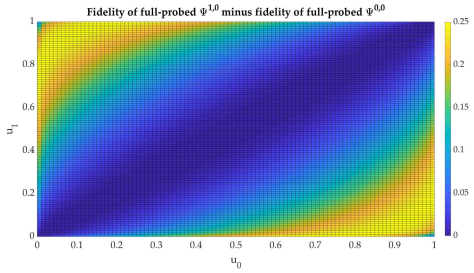


Fig. 14. Fidelity of full-probed state $|\Psi^{1,0}\rangle$ minus fidelity of full-probed state $|\Psi^{0,0}\rangle$.

of the pure probe state. Thus, optimizing the full-probed strategy using $|\Psi^{00}\rangle$ leads to the same result as for pure probe state. Lastly, we consider the full-probed strategy using $|\Psi^{10}\rangle$. We have $q = 0$ giving the optimal value of fidelity. A consequence of this is probing the quantum channel twice using states $|0\rangle$ and $|1\rangle$, respectively, in each round. Since for pure probes, using the state $|1\rangle$ is optimal, then the round when we probe with $|0\rangle$ contributes by increasing 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. 14.

As an overall picture, we can conclude the following. The best optimized probing strategy uses the half-probed state $|\Psi^{00}\rangle$ with the distribution of q, for every pair (u_0, u_1), given by Fig. 10b. However, since the improvement is not so significant, see Fig. 11, and manipulating entangled state is typically costly, a sub-optimal strategy uses the pure state $|1\rangle$ in each round.

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 information rate and reliability parameter are connected in a way that when the information rate is close to its maximum, then the reliability parameter 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 the unit and the other is close to zero, the mutual information attains its maximum possible value. Furthermore, probing the quantum memory cell using the half-probed $|\Psi^{00}\rangle$ strategy is again the optimal procedure for quantum reading.

V. 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 of information rate and reliability parameter of a quantum channel. The polarization phenomenon produced by channel combining and analyzed in the channel 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. We have shown that when the number of channels is arbitrarily large, the set of synthesized channels can be divided into 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 literature, we have shown that our decoding scheme has error probability that decays exponentially fast with the square root of the code length. In the end, optimizations over probe states are investigated, leading to the conclusion that, in general, half-probed states are the best choice.

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, can be applied here. Secondly, we can consider the set of quantum memory cells to be composed of generalized amplitude damping channels. This class of channel can be seen as a second-order approximation of classical digital memory, where the model considers the environment temperature. 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 is still a research topic even for classical Gaussian channels. The Gaussian bosonic channel is the ultimate goal as a quantum channel model for classical digital memories. Because this class of channels is defined over continuous variables, attacking it needs to be two-fold. Firstly, the polar code scheme has to take into account the channel to prove the polarization phenomenon. Secondly, an optimization over the probe states is necessary under the energy constraint. 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 context above. Uniqueness also needs to be verified.

Acknowledgment

The authors thank the reviewers for their help in improving this paper.

References

[1] A. Acín, "Statistical distinguishability between unitary operations," *Phys. Rev. Lett.*, vol. 87, no. 17, Oct. 2001.
[2] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, Jul. 2009.
[3] E. Arıkan, "Source polarization," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 899–903.
[4] E. Arikan and E. Telatar, "On the rate of channel polarization," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2009, pp. 1493–1495.
[5] L. Banchi, Q. Zhuang, and S. Pirandola, "Quantum-enhanced barcode decoding and pattern recognition," *Phys. Rev. A, Gen. Phys.*, vol. 14, no. 6, Dec. 2020, Art. no. 064026.
[6] H. Barnum and E. Knill, "Reversing quantum dynamics with near-optimal quantum and classical fidelity," *J. Math. Phys.*, vol. 43, no. 5, p. 2097, 2002.
[7] A. M. Childs, J. Preskill, and J. Renes, "Quantum information and precision measurement," *J. Mod. Opt.*, vol. 47, nos. 2–3, pp. 155–176, Feb. 2000.

[8] S. Das and M. M. Wilde, "Quantum rebound capacity," *Phys. Rev. A, Gen. Phys.*, vol. 100, no. 3, Sep. 2019, Art. no. 030302.
[9] R. Duan, Y. Feng, and M. Ying, "Perfect distinguishability of quantum operations," *Phys. Rev. Lett.*, vol. 103, no. 21, Nov. 2009, Art. no. 210501.
[10] F. Dupuis, A. Goswami, M. Mhalla, and V. Savin, "Polarization of quantum channels using Clifford-based channel combining," *IEEE Trans. Inf. Theory*, vol. 67, no. 5, pp. 2857–2877, May 2021.
[11] A. Gilchrist, N. K. Langford, and M. A. Nielsen, "Distance measures to compare real and ideal quantum processes," *Phys. Rev. A, Gen. Phys.*, vol. 71, no. 6, Jun. 2005, Art. no. 062310.
[12] A. Goswami, M. Mhalla, and V. Savin, "Multilevel polarization for quantum channels," *Quantum Inf. Comput.*, vol. 21, nos. 7–8, pp. 577–606, Jun. 2021.
[13] A. W. Harrow, A. Hassidim, D. W. Leung, and J. Watrous, "Adaptive versus nonadaptive strategies for quantum channel discrimination," *Phys. Rev. A, Gen. Phys.*, vol. 81, no. 3, Mar. 2010, Art. no. 032339.
[14] M. Hayashi, "Discrimination of two channels by adaptive methods and its application to quantum system," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3807–3820, Aug. 2009.
[15] C. W. Helstrom, "Quantum detection and estimation theory," *J. Statist. Phys.*, vol. 1, no. 2, pp. 231–252, 1969.
[16] C. Hirche, "From asymptotic hypothesis testing to entropy inequalities," Ph.D. dissertation, Dept. de Física, Univ. Autónoma de Barcelona, Bellaterra, Spain, 2018.
[17] C. Hirche, C. Morgan, and M. M. Wilde, "Polar codes in network quantum information theory," *IEEE Trans. Inf. Theory*, vol. 62, no. 2, pp. 915–924, Feb. 2016.
[18] A. S. Holevo, "Statistical decision theory for quantum systems," *J. Multivariate Anal.*, vol. 3, pp. 337–394, Dec. 1973.
[19] A. S. Holevo, "Reliability function of general classical-quantum channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 2256–2261, Sep. 2000.
[20] J. Honda and H. Yamamoto, "Polar coding without alphabet extension for asymmetric models," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 7829–7838, Dec. 2013.
[21] V. Katariya and M. M. Wilde, "Evaluating the advantage of adaptive strategies for quantum channel distinguishability," *Phys. Rev. A, Gen. Phys.*, vol. 104, no. 5, p. 052406, Nov. 2021.
[22] S. Lloyd, "Enhanced sensitivity of photodetection via quantum illumination," *Science*, vol. 321, no. 5895, pp. 1463–1465, Sep. 2008.
[23] R. Nasser and J. M. Renes, "Polar codes for arbitrary classical-quantum channels and arbitrary cq-macs," *IEEE Trans. Inf. Theory*, vol. 64, no. 11, pp. 7424–7442, Nov. 2018.
[24] S. Khabbazi Oskouei, S. Mancini, and M. M. Wilde, "Union bound for quantum information processing," *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 475, no. 2221, Jan. 2019, Art. no. 20180612.
[25] S. Pirandola, "Quantum reading of a classical digital memory," *Phys. Rev. Lett.*, vol. 106, no. 9, Mar. 2011, Art. no. 090504.
[26] S. Pirandola, R. Laurenza, C. Lupo, and J. L. Pereira, "Fundamental limits to quantum channel discrimination," *npj Quantum Inf.*, vol. 5, no. 1, pp. 1–8, Jun. 2019.
[27] S. Pirandola, C. Lupo, V. Giovannetti, S. Mancini, and S. L. Braunstein, "Quantum reading capacity," *New J. Phys.*, vol. 13, no. 11, Nov. 2011, Art. no. 113012.
[28] S. Pirandola, S. Mancini, S. Lloyd, and S. L. Braunstein, "Continuous-variable quantum cryptography using two-way quantum communication," *Nature Phys.*, vol. 4, pp. 726–730, Jul. 2008.
[29] J. M. Renes, D. Sutter, F. Dupuis, and R. Renner, "Efficient quantum polar codes requiring, no preshared entanglement," *IEEE Trans. Inf. Theory*, vol. 61, no. 11, pp. 6395–6414, Nov. 2015.
[30] W. Roga, M. Fannes, and K. Äyczkowski, "Universal bounds for the Holevo quantity, coherent information, and the Jensen-Shannon divergence," *Phys. Rev. Lett.*, vol. 105, no. 4, Jul. 2010, Art. no. 040505.
[31] M. F. Sacchi, "Entanglement can enhance the distinguishability of entanglement-breaking channels," *Phys. Rev. A, Gen. Phys.*, vol. 72, no. 1, Jul. 2005, Art. no. 014305.
[32] M. F. Sacchi, "Optimal discrimination of quantum operations," *Phys. Rev. A, Gen. Phys.*, vol. 71, no. 6, Jun. 2005, Art. no. 062340.
[33] A. G. Sahebi and S. S. Pradhan, "Multilevel polarization of polar codes over arbitrary discrete memoryless channels," in *Proc. 49th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Sep. 2011, pp. 1718–1725.
[34] E. Sasoglu, E. Telatar, and E. Arikan, "Polarization for arbitrary discrete memoryless channels," in *Proc. IEEE Inf. Theory Workshop*, Oct. 2009, pp. 144–148.
[35] D. Sutter, J. M. Renes, F. Dupuis, and R. Renner, "Achieving the capacity of any DMC using only polar codes," in *Proc. IEEE Inf. Theory Workshop*, Sep. 2012, pp. 114–118.
[36] I. Tal and A. Vardy, "How to construct polar codes," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6562–6582, Oct. 2013.
[37] S.-H. Tan *et al.*, "Quantum illumination with Gaussian states," *Phys. Rev. Lett.*, vol. 101, no. 25, Dec. 2008, Art. no. 253601.
[38] G. Wang and M. Ying, "Unambiguous discrimination among quantum operations," *Phys. Rev. A, Gen. Phys.*, vol. 73, no. 4, Apr. 2006, Art. no. 042301.
[39] M. M. Wilde and S. Guha, "Polar codes for classical-quantum channels," *IEEE Trans. Inf. Theory*, vol. 59, no. 2, pp. 1175–1187, Feb. 2013.
[40] M. M. Wilde and S. Guha, "Polar codes for degradable quantum channels," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4718–4729, Jul. 2013.

Francisco Revson F. Pereira received the M.Sc. and Ph.D. degrees in electrical engineering from the Federal University of Campina Grande, Brazil, in 2016 and 2019, respectively. From 2018 to 2019, he was a Visiting Ph.D. Student with the Technical University of Eindhoven, The Netherlands. Currently, he is a Postdoctoral Fellow with the University of Camerino, Italy. His research interests include coding theory, information theory, and quantum information processing.

Stefano Mancini received the Ph.D. degree in physics from the University of Perugia, Italy, in 1998. Then, he spent three years as a Postdoctoral Researcher at the University of Milan, Italy. Subsequently, with temporary lecturer positions, he contributed to establish the first Italian academic courses on quantum information and computation. From 2004 to 2010, he was a Researcher of theoretical physics and mathematical methods of physics with the University of Camerino, Italy, where he has been a Professor since 2010. His main research interests include quantum and geometric aspects of information theory.