7-15-2013

Polar codes for degradable quantum channels

Mark M. Wilde
Université McGill

Saikat Guha
BBN Technologies

Recommended Citation
Wilde, M., & Guha, S. (2013). Polar codes for degradable quantum channels. *IEEE Transactions on Information Theory, 59* (7), 4718-4729. https://doi.org/10.1109/TIT.2013.2250575

Polar codes for degradable quantum channels

Mark M. Wilde and Saikat Guha

***Abstract*—Channel polarization is a phenomenon in which a particular recursive encoding induces a set of synthesized channels from many instances of a memoryless channel, such that a fraction of the synthesized channels becomes near perfect for data transmission and the other fraction becomes near useless for this task. Mahdavifar and Vardy have recently exploited this phenomenon to construct codes that achieve the symmetric private capacity for private data transmission over a degraded wiretap channel. In the current paper, we build on their work and demonstrate how to construct quantum wiretap polar codes that achieve the symmetric private capacity of a degraded quantum wiretap channel with a classical eavesdropper. Due to the Schumacher-Westmoreland correspondence between quantum privacy and quantum coherence, we can construct quantum polar codes by operating these quantum wiretap polar codes in superposition, much like Devetak's technique for demonstrating the achievability of the coherent information rate for quantum data transmission. Our scheme achieves the symmetric coherent information rate for quantum channels that are degradable with a classical environment. This condition on the environment may seem restrictive, but we show that many quantum channels satisfy this criterion, including amplitude damping channels, photon-detected jump channels, dephasing channels, erasure channels, and cloning channels. Our quantum polar coding scheme has the desirable properties of being channel-adapted and symmetric capacity-achieving along with having an efficient encoder, but we have not demonstrated that the decoding is efficient. Also, the scheme may require entanglement assistance, but we show that the rate of entanglement consumption vanishes in the limit of large blocklength if the channel is degradable with classical environment.**

I. Introduction

In a seminal paper on quantum error correction, Shor set out the "goal of [defining] the quantum analog of the Shannon capacity [41] for a quantum channel, and [finding] encoding schemes which approach this capacity" [42]. At the time, it was not really clear how to define the quantum capacity of a quantum channel, but Shor's quantum error correction code [42] gave some clues for constructing more general encoding schemes. Subsequently, several authors contributed increasingly sophisticated quantum error correction codes [10], [46], [18] and others established a good definition of and upper bounds on the quantum capacity of a channel [37], [38], [4], [5], culminating in some high-performing quantum error-correction codes [30], [29], [33], [25] and random-coding based schemes for achieving the coherent information rate [38] of a quantum channel [28], [43], [11]. For some channels known as degradable quantum channels [12], in which the channel to the environment is noisier than the channel to the intended receiver, the random-coding based schemes from Refs. [28], [43], [11] achieve their quantum capacity, due to the particular structure of these channels.

In spite of the astounding progress in both quantum error correction [13] and quantum Shannon theory [48], none of the high performance codes constructed to date are provably capacity achieving, and none of the aforementioned schemes that achieve the capacity are explicit (the proofs instead exploit randomness to establish the existence of a code). Among the schemes that achieve the quantum capacity, perhaps Devetak's [11] provides the most clear recipe to a quantum code designer interested in constructing a capacity-achieving quantum code. His proof takes a cue from a certain security proof of quantum key distribution [44] and the Schumacher-Westmoreland correspondence between quantum privacy and quantum coherence [39], by first establishing the existence of codes that achieve the private capacity of a quantum wiretap channel and then demonstrating how to operate such a code in superposition so that it achieves the quantum capacity. It is also clear that the structure of his codes bears some similarities with Calderbank-Shor-Steane codes [10], [46].

Along with the above developments, there have been impressive breakthroughs in classical coding theory and information theory [14], one of which is Arikan's recent work on polar codes [2]. Arikan's polar codes exploit a phenomenon known as channel polarization, in which a particular recursive encoding induces a set of synthesized channels from the original memoryless noisy channels. The synthesized channels are such that a fraction of them are perfect for data transmission, while the other fraction are completely useless, and the fraction that are perfect is equal to the symmetric capacity of the original channel. The codes are channel adapted, in the sense that Arikan's "polar coding rule" establishes through which of the synthesized channels the sender should transmit data, and this polar coding rule depends on the particular channel being used. The codes are near explicit and have the desirable property that both the encoding and decoding are efficient (the complexity of each is $O\left(N \log N\right)$ where N is the blocklength of the code).

Arikan's work might make us wonder whether it would be possible to construct polar codes for transmitting quantum data over general quantum channels, and the development in the classical world most relevant for this task is due to Madhavifar and Vardy [31]. There, they established that a modification of Arikan's original polar coding scheme can achieve the symmetric private capacity of a degraded classical wiretap channel. (In order to make this statement, the sender and receiver actually require access to a small amount of secret key, but the rate of secret key needed vanishes when the code's blocklength becomes large.) Thus, with the Madhavifar-Vardy scheme for polar coding over classical degraded wiretap channels [31], the Devetak scheme for operating a quantum wiretap

Mark M. Wilde is with the School of Computer Science, McGill University, Montreal, Quebec H3A 2A7, Canada. Saikat Guha is with the Quantum Information Processing Group, Raytheon BBN Technologies, Cambridge, Massachusetts, USA 02138. (E-mail: mark.wilde@mcgill.ca; sguha@bbn.com)

code in superposition [11], and our recent work on polar codes for transmitting classical data over quantum channels [49], it should be evident that one could put these pieces together in order to construct polar codes for transmitting quantum data over degradable quantum channels.

In this paper, we pursue this direction by constructing polar codes that achieve a symmetric capacity for transmitting quantum data over particular degradable quantum channels. These channels should satisfy the property that encoding classical data in some orthonormal basis at their input leads to commuting states for the environment (essentially, the environment becomes classical), and we clarify later why this is important in our construction. Many degradable channels fall into this class, including amplitude damping channels [16], photon-detected jump channels [1], erasure channels [19], dephasing channels [11], and cloning channels [6] (channels induced by universal cloning machines [17], [9]). These noisy channels occur naturally in physical processes, with amplitude damping modeling photon loss or spontaneous emission, the photon-detected jump channel modeling the spontaneous decay of atoms with a detected photon emission [1], the erasure channel being a different model for photon loss [15], the dephasing channel modeling random phase noise in superconducting systems [7], and the cloning channel modeling stimulated emission from an atom [32], [45], [27]. Our codes are symmetric capacity achieving for all of the above channels, and this follows from analyzing a quantum polar coding rule for these channels.

We summarize briefly how the construction works. First, we consider a quantum wiretap channel with one classical input and two quantum outputs, one for the legitimate receiver (Bob) and the other for the wiretapper (Eve). We demonstrate that these channels polarize in four different ways, based on whether the channels are good or bad for the receiver or the wiretapper. In order to have strong security, we must guarantee that the bad channels for the wiretapper are in fact "really bad" in a precise sense, and it is for this reason that we consider channels with classical environment. (Interestingly, we know of quantum channels where it is not clear to us how to ensure that they become "really bad," and we prove some results in this direction in Appendix A.) The resulting coding scheme is to send the information bits through the channels which are good for Bob and bad for Eve, "frozen" bits through the channels that are bad for both, half of a secret key through channels which are bad for Bob but good for Eve, and randomized bits through the channels good for both. By an analysis similar to that of Madhavifar and Vardy [31], we can demonstrate that this scheme achieves the symmetric private capacity of a degraded quantum wiretap channel with classical environment, while the rate of secret key required vanishes in the limit of large blocklength.

The main idea for constructing quantum polar codes for degradable quantum channels with classical environment is just to operate the quantum wiretap code in superposition and exploit Arikan's encoding with CNOT gates with respect to some orthonormal basis. This amounts to sending information qubits through the channels good for the receiver Bob and bad for the *environment* Eve, frozen ancilla qubits through the channels that are bad for both, half of shared entanglement through the channels that are bad for Bob but good for Eve, and superposed ancilla qubits in the state $|+\rangle \equiv (|0\rangle + |1\rangle)/\sqrt{2}$ through the channels that are good for both. The resulting quantum polar codes are entanglement-assisted [8], but we can prove that the entanglement consumption rate required vanishes in the limit of large blocklength (in this regard, the codes here are similar to Hsieh *et al.*'s recent ones [24]). Operating the quantum successive cancellation decoder from Ref. [49] in a coherent fashion, followed by controlled "decoupling unitaries" allows us to exploit the properties of the quantum wiretap polar code in order to prove that the above quantum polar code performs well (we note that this decoder is similar to Devetak's [11]). The resulting quantum polar codes achieve the symmetric quantum capacity of degradable channels with classical environment with an encoding circuit that has complexity $O(N \log N)$. The decoding unfortunately remains inefficient, but further efforts may lead to an efficient realization of a decoder.

Recently, Renes *et al*. have independently constructed quantum polar codes that have both an efficient encoding and decoding, though they achieve the coherent information rate only for Pauli channels [35]. We should clarify the ways in which their scheme is different from ours. First, they restrict their construction to Pauli channels because they are considering the effective classical channels induced in the amplitude (Pauli-Z) and phase (Pauli-X) bases (not all channels, including some of the ones mentioned above, induce classical channels in complementary bases). As a result, they can directly import Arikan's ideas because they are dealing with classical channels in complementary bases. An additional bonus is that they obtain an efficient decoder as well as an efficient encoder, essentially because their decoder is Arikan's successive cancellation decoder implemented as an efficient unitary operation. Their codes require the assistance of shared entanglement, but there are some channels for which they can prove that it is not required.

Our scheme is different from theirs in several ways. First, we have a quantum polar coding rule that is adapted to a given quantum channel. In particular, the rule used for determining the good or bad channels is based on a quantum parameter (fidelity). Also, we demonstrate polarization in terms of two quantum parameters, the fidelity and the Holevo information, by building on our earlier results in Ref. [49]. It is for this reason that our scheme is symmetric capacity-achieving for a wide variety of quantum channels. Also, the first part of our decoder is a coherent version of the quantum successive cancellation decoder from Ref. [49], rather than one that is based directly on the classical decoder. Since we have not proven that the quantum successive cancellation decoder from Ref. [49] has an efficient implementation, the coherent version of it in this work is certainly not efficient. In spite of the inefficiency of the quantum successive cancellation decoder, this decoder is needed in order to achieve the symmetric quantum capacity of the channels considered here. Finally, all the channels we consider here only require a vanishing rate of entanglement assistance, due to an argument similar to Proposition 22 of Madhavifar and Vardy [31] and the fact that

they are degradable with a classical environment.

We structure this paper as follows. Section II reviews our work from Ref. [49] on polar codes for classical-quantum channels. We present in Section III our scheme for private communication over quantum wiretap channels that are degraded with a classical environment. Finally, we demonstrate how to construct quantum polar codes from quantum wiretap polar codes in Section IV. The last section concludes with a summary and some open questions.

II. Review of polar codes for classical-quantum channels

We begin by providing a brief review of polar codes constructed for classical-quantum channels [49]. There, we considered channels with binary classical inputs and quantum outputs of the form:

$$W : x \to \rho_x,$$

where W denotes the channel, $x \in \{0,1\}$, and ρ_x is a density operator. The relevant parameters that determine channel performance are the fidelity $F(W) \equiv \left\|\sqrt{\rho_0}\sqrt{\rho_1}\right\|_1^2$ and the symmetric Holevo information $I(W) \equiv H(\rho) - 1/2(H(\rho_0) + H(\rho_1))$ with $\rho = 1/2(\rho_0 + \rho_1)$ and the von Neumann entropy $H(\sigma) \equiv -\text{Tr}\{\sigma \log_2 \sigma\}$. Channels with $F(W) \leq \epsilon$ are nearly noiseless and those with $F(W) \geq 1 - \epsilon$ are near to being completely useless. The fidelity generalizes the Bhattacharya distance Z [2] in the sense that $\sqrt{F(W)} = Z(W)$ if the two states ρ_0 and ρ_1 commute (i.e., if the channel is classical).

In coding classical information for the above channel, we consider $N = 2^n$ copies of W, such that the resulting channel is of the form

$$x^N \equiv x_1 \cdots x_N \to \rho_{x^N} \equiv \rho_{x_1} \otimes \cdots \otimes \rho_{x_N},$$

where x^N is the length N input and ρ_{x^N} is the output state. We can extend Arikan's idea of channel combining to this classical-quantum channel, by considering the channels induced by a transformation on an input bit (row) vector u^N:

$$u^N \to \rho_{u^N G_N},$$

where $G_N = B_N F^{\otimes n}$, with B_N being a permutation matrix that reverses the order of the bits and

$$F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

This classical encoding is equivalent to a network of classical CNOT gates and permutation operations that can be implemented with complexity $O(N \log N)$ (see Figures 1 and 2 of Ref. [49] or Figures 1, 2, and 3 of Ref. [2]). We can also define the split channels from the above combined channels as

$$W_N^{(i)} : u_i \to \rho_{(i),u_i}^{U_1^{i-1} B^N}, \tag{1}$$

where

$$\rho_{(i),u_i}^{U_1^{i-1} B^N} \equiv \sum_{u_1^{i-1}} \frac{1}{2^{i-1}} \left|u_1^{i-1}\right\rangle \left\langle u_1^{i-1}\right|^{U_1^{i-1}} \otimes \overline{\rho}_{u_1^i}^{B^N}, \tag{2}$$

$$\overline{\rho}_{u_1^i}^{B^N} \equiv \sum_{u_{i+1}^N} \frac{1}{2^{N-i}} \rho_{u^N G_N}^{B^N}. \tag{3}$$

The interpretation of this channel is that it is the one "seen" by the bit u_i if all of the previous bits u_1^{i-1} are available and if we consider all the future bits u_{i+1}^N as randomized. This motivates the development of a quantum successive cancellation decoder [49] that attempts to distinguish $u_i = 0$ from $u_i = 1$ by adaptively exploiting the results of previous measurements and Helstrom-Holevo measurements [22], [23] for each bit decision.

Arikan's polar coding rule is to divide the channels into "good" ones and "bad" ones. Let $[N] \equiv \{1, \ldots, N\}$ and β be a real such that $0 < \beta < 1/2$. The polar coding rule divides the channels as follows:

$$\mathcal{G}_N(W, \beta) \equiv \left\{ i \in [N] : \sqrt{F(W_N^{(i)})} < 2^{-N^\beta} \right\}, \tag{4}$$

$$\mathcal{B}_N(W, \beta) \equiv \left\{ i \in [N] : \sqrt{F(W_N^{(i)})} \geq 2^{-N^\beta} \right\}, \tag{5}$$

so that the channels in $\mathcal{G}_N(W, \beta)$ are the good ones and those in $\mathcal{B}_N(W, \beta)$ are the bad ones. Observe that the quantum polar coding rule involves the quantum channel parameter F, rather than a classical one such as the Bhattacharya distance.

The following theorem is helpful in determining what fraction of the channels become good or bad [3]:

Theorem 1 (Convergence Rate): Let $\{X_n : n \geq 0\}$ be a random process with $0 \leq X_n \leq 1$ and satisfying

$$X_{n+1} \leq q X_n \quad \text{w.p.} \quad 1/2, \tag{6}$$

$$X_{n+1} = X_n^2 \quad \text{w.p.} \quad 1/2, \tag{7}$$

where q is some positive constant. Let $X_\infty = \lim_{n\to\infty} X_n$ exist almost surely with $\Pr\{X_\infty = 0\} = P_\infty$. Then for any $\beta < 1/2$,

$$\lim_{n\to\infty} \Pr\left\{ X_n < 2^{-2^{n\beta}} \right\} = P_\infty,$$

and for any $\beta > 1/2$,

$$\lim_{n\to\infty} \Pr\left\{ X_n < 2^{-2^{n\beta}} \right\} = 0.$$

One can then consider the channel combining and splitting mentioned above as a random birth process in which a channel W_{n+1} is constructed from two copies of a previous one W_n according to the rules in Section 4 of Ref. [49]. One can then consider the process $\{F_n : n \geq 0\} \equiv \left\{ \sqrt{F(W_n)} : n \geq 0 \right\}$ and prove that it is a bounded super-martingale by exploiting the relationships given in Proposition 10 of Ref. [49]. From the convergence properties of martingales, one can then conclude that F_∞ converges almost surely to a value in $\{0, 1\}$, and the probability that it equals zero is equal to the symmetric Holevo information $I(W)$. Furthermore, since the process F_n satisfies the relations in (6-7), the following proposition on the convergence rate of polarization holds:

Theorem 2: Given a binary input classical-quantum channel W and any $\beta < 1/2$,

$$\lim_{n\to\infty} \Pr\left\{ F_n < 2^{-2^{n\beta}} \right\} = I(W).$$

One of the important advances in Ref. [49] was to establish that a quantum successive cancellation decoder performs well for polar coding over classical-quantum channels. In this

case, the decoder is some positive operator-valued measure (POVM) $\{\Lambda_{u_{\mathcal{A}}}\}$ that attempts to decode the information bits $u_{\mathcal{A}}$ reliably. In particular, we showed the following bound on the performance of such a decoder (by exploiting Sen's "non-commutative union bound" [40]):

$$\Pr\{\widehat{U}_{\mathcal{A}} \neq U_{\mathcal{A}}\} \leq 2\sqrt{\sum_{i\in\mathcal{A}} \frac{1}{2}\sqrt{F(W_N^{(i)})}},$$

under the assumption that the sender chooses the information bits $U_{\mathcal{A}}$ according to a uniform distribution. Thus, by choosing the channels over which the sender transmits the information bits to be in $\mathcal{G}_N(W,\beta)$ and those over which she transmits agreed upon frozen bits to be in $\mathcal{B}_N(W,\beta)$, we obtain the following bound on the probability of decoding error:

$$\Pr\{\widehat{U}_{\mathcal{A}} \neq U_{\mathcal{A}}\} = o(2^{-\frac{1}{2}N^{\beta}}).$$

This completes the specification of a polar code for classical-quantum channels.

We end this section by stating a lemma that will prove useful for us:

Lemma 3: Let W and W^* both be binary-input classical-quantum channels, such that W^* is a degraded version of W, in the sense that

$$W^*(x) = \mathcal{D}(W(x)),$$

where x is the classical input to the channels and $\mathcal{D}$ is some degrading quantum channel from W to W^*. Let $W_N^{(1)}$, $\ldots$, $W_N^{(N)}$ and $W_N^{*(1)}$, $\ldots$, $W_N^{*(N)}$ denote the corresponding synthesized channels from channel combining and splitting. Then $W_N^{*(i)}$ is degraded with respect to $W_N^{(i)}$ for all $i \in [N]$ and furthermore, we have that $I(W_N^{(i)}) \geq I(W_N^{*(i)})$ and $F(W_N^{(i)}) \leq F(W_N^{*(i)})$.

Proof: This lemma follows straightforwardly from the definition in (1), and the fact that quantum mutual information and fidelity are monotone under quantum processing with the degrading map $\mathcal{D}$ [48]. ∎

We can then observe from the above lemma and (4) that if W^* is degraded with respect to W, the good channels for W^* are a subset of those that are good for W: $\mathcal{G}_N(W^*,\beta) \subseteq \mathcal{G}_N(W,\beta)$. Similarly, the following relationship holds as well: $\mathcal{B}_N(W,\beta) \subseteq \mathcal{B}_N(W^*,\beta)$.

III. Quantum wiretap polar codes

We now discuss how to construct polar codes that achieve the symmetric private information rate for a quantum wiretap channel. The results in this section build upon those of Mahdavifar and Vardy in Ref. [31].

The model for a binary-input quantum wiretap channel is as follows:

$$x \to \rho_x^{BE},$$

where $x \in \{0,1\}$ and ρ_x^{BE} is a density operator on a tensor product Hilbert space BE. The legitimate receiver Bob has access to the system B and the eavesdropper Eve has access to the system E. Thus, Bob's density operator is

$$\rho_x^B = \mathrm{Tr}_E\left\{\rho_x^{BE}\right\},$$

and Eve's density operator is

$$\rho_x^E = \mathrm{Tr}_B\left\{\rho_x^{BE}\right\}.$$

The quantum wiretap channel is degraded if there exists some quantum channel $\mathcal{D}$ such that the following condition holds for all x:

$$\rho_x^E = \mathcal{D}\left(\rho_x^B\right).$$

Let W denote the channel to Bob:

$$W : x \to \rho_x^B, \tag{8}$$

and let W^* denote the channel to Eve:

$$W^* : x \to \rho_x^E. \tag{9}$$

In order to make a statement about the strong security of a quantum wiretap polar code, we need to ensure that the channels over which the sender is transmitting information bits to Bob should be "really bad" for Eve. That is, it is not sufficient for the channels to satisfy (5), but they should be divided as to whether they are poor for Eve according to the following stronger criterion:

$$\mathcal{P}(W^*,\beta) \equiv \left\{i \in [N] : \sqrt{F(W_N^{*(i)})} > 1 - 2^{-N^{\beta}}\right\}.$$

Dividing the channels for Eve in this way makes it nearly impossible for her to determine whether the sender transmits a zero or one through these channels in the limit where N becomes large. In what follows, we say that the channels in $\mathcal{P}(W^*,\beta)$ are "bad" for Eve while those in $\mathcal{P}^c(W^*,\beta) \equiv [N] \setminus \mathcal{P}(W^*,\beta)$ are "good" for Eve.

It is again important for us to know what fraction of the channels $W_N^{*(i)}$ become bad for Eve in order to establish that the quantum wiretap polar codes are symmetric capacity-achieving—i.e., it would be good to have another theorem similar to Theorem 2 for this case. In order to have such a theorem, we would require a birth process that obeys the properties in (6-7). Fortunately, in the case that ρ_0^E and ρ_1^E commute, we have the following proposition:

Proposition 4: Suppose that the states ρ_0^E and ρ_1^E for the binary-input classical-quantum channel W^* commute. Then for any $\beta < 1/2$,

$$\lim_{n\to\infty} \Pr\left\{F_n^* > 1 - 2^{-2^{n\beta}}\right\} = 1 - I(W^*),$$

where F_n^* is the process $\{F_n^* : n \geq 0\} \equiv \{\sqrt{F(W_n^*)} : n \geq 0\}$.

Proof: The proof proceeds along similar lines as Theorem 3.15 in Ref. [26]. Since the states ρ_0^E and ρ_1^E commute, they are effectively classical, and the fidelities in F_n^* reduce to the classical Bhattacharya parameters. It is then possible to show that this process satisfies

$$F_{n+1}^* \geq F_n^*\sqrt{2 - (F_n^*)^2} \quad \text{w.p.} \quad 1/2, \tag{10}$$

$$F_{n+1}^* = (F_n^*)^2 \quad \text{w.p.} \quad 1/2. \tag{11}$$

The first relation follows from Lemma 3.16 in Ref. [26], and the second follows from Lemma 2.16 in Ref. [26]. We can

then rewrite the above conditions as follows:

$$1-\left(F_{n+1}^{*}\right)^{2} \leq \left(1-\left(F_{n}^{*}\right)^{2}\right)^{2} \quad \text{w.p.} \quad 1/2, \tag{12}$$

$$1-\left(F_{n+1}^{*}\right)^{2} = 1-\left(F_{n}^{*}\right)^{4} \leq 2\left(1-\left(F_{n}^{*}\right)^{2}\right) \quad \text{w.p.} \quad 1/2. \tag{13}$$

Defining X_n by $X_n \equiv 1-\left(F_n^*\right)^2$, it is now clear the process X_n satisfies the conditions in (6-7). Since we know that F_n^* converges almost surely to a random variable F_∞^* taking values in $\{0,1\}$ with $\Pr\{F_\infty^*=1\} = 1-I(W^*)$, it follows that X_n converges almost surely to X_∞ with $\Pr\{X_\infty=0\} = 1-I(W^*)$. The process X_n then satisfies all the requirements needed to apply Theorem 1, so that

$$\lim_{n\to\infty} \Pr\left\{X_n < 2^{-2^{n\beta}}\right\} = 1-I(W^*),$$

which in turn, from the relation $X_n = 1-\left(F_n^*\right)^2 \geq 1-F_n^*$, implies that

$$\lim_{n\to\infty} \Pr\left\{1-F_n^* < 2^{-2^{n\beta}}\right\} = 1-I(W^*),$$

giving the statement of the proposition. ∎

It is worthwhile to discuss why we specialized the above proposition to the case where the states ρ_0^E and ρ_1^E are commuting. First, as we demonstrate in Appendix B, there are many examples of natural quantum channels for which this condition holds, including amplitude damping channels, photon-detected jump channels, dephasing channels, erasure channels, and cloning channels. Thus, the quantum wiretap polar coding scheme in this section and the quantum polar coding scheme in the next section works well for these channels. On the other hand, there exist quantum wiretap channels for which the critical inequality in (10) does not hold. For example, Appendix A demonstrates a violation of the inequality whenever the states ρ_0^E and ρ_1^E are pure and such that $\mathrm{Tr}\{\rho_0^E\rho_1^E\} \notin \{0,1\}$. So the scheme given in this section does not necessarily achieve the symmetric private capacity for such channels because it is not clear how to guarantee that the fraction of bad channels for Eve is equal to $1-I(W^*)$.

We can now establish our scheme for a quantum wiretap polar code. We divide the set $[N]$ into four different subsets:

$$\begin{aligned}
\mathcal{A} &\equiv \mathcal{P}(W^*,\beta)\cap\mathcal{G}_N(W,\beta),\\
\mathcal{B} &\equiv \mathcal{P}(W^*,\beta)\cap\mathcal{B}_N(W,\beta),\\
\mathcal{X} &\equiv \mathcal{P}^c(W^*,\beta)\cap\mathcal{B}_N(W,\beta),\\
\mathcal{Y} &\equiv \mathcal{P}^c(W^*,\beta)\cap\mathcal{G}_N(W,\beta).
\end{aligned}$$

Observe that $\mathcal{A}$, $\mathcal{B}$, $\mathcal{X}$, and $\mathcal{Y}$ form a partition of $[N]$ because they are all pairwise disjoint and $\mathcal{A}\cup\mathcal{B}\cup\mathcal{X}\cup\mathcal{Y}=[N]$. Thus, the set $\mathcal{A}$ consists of channels that are good for Bob and bad for Eve, $\mathcal{B}$ has the channels that are bad for both, $\mathcal{X}$ has the channels that are good for Eve and bad for Bob, and $\mathcal{Y}$ has the channels that are good for Eve and good for Bob. The Mahdavifar-Vardy coding scheme is then straightforward:

1) Send the information bits through the channels in $\mathcal{A}$.
2) Send the frozen bit vector $u_{\mathcal{B}}$ through the channels in $\mathcal{B}$.
3) Send randomized bits through the channels in $\mathcal{Y}$.
4) We suppose that Alice and Bob have access to a secret key before communication begins. Alice inputs her half of the secret key into the channels in $\mathcal{X}$.[1] Mahdavifar and Vardy demonstrated that the fraction $|\mathcal{X}|/N$ tends to zero in the limit $N\to\infty$ [31], and a slight modification of their argument demonstrates that the codes constructed here have the same property. This implies that the rate of secret key needed to ensure reliability and strong security for this coding scheme vanishes and is thus negligible in the asymptotic limit (we require the strong security criterion for when we produce quantum polar codes from quantum wiretap polar codes).

The following theorem guarantees that the rate of the quantum wiretap polar code is equal to the symmetric private information:

Theorem 5: For the quantum wiretap polar coding scheme given above, a degraded quantum wiretap channel with W and W^* as defined in (8-9), with W^* having a classical output, and for sufficiently large N, its rate $R_N = |\mathcal{A}|/N$ converges to the symmetric private information:

$$\lim_{N\to\infty} R_N = I(W) - I(W^*).$$

Proof: We just need to determine the size of the set $\mathcal{A}$. From basic set theory, we know that

$$\begin{aligned}
\frac{|\mathcal{A}|}{N} &= \frac{1}{N}\left|\mathcal{P}(W^*,\beta)\cap\mathcal{G}_N(W,\beta)\right|\\
&= \frac{|\mathcal{P}(W^*,\beta)|}{N} + \frac{|\mathcal{G}_N(W,\beta)|}{N}\\
&\quad - \frac{1}{N}\left|\mathcal{P}(W^*,\beta)\cup\mathcal{G}_N(W,\beta)\right|.
\end{aligned}$$

Consider that

$$\begin{aligned}
\mathcal{P}(W^*,\beta)\cup\mathcal{G}_N(W,\beta) &= [N]\setminus\left(\mathcal{P}(W^*,\beta)\cup\mathcal{G}_N(W,\beta)\right)^c\\
&= [N]\setminus\left(\mathcal{P}^c(W^*,\beta)\cap\mathcal{B}_N(W,\beta)\right)\\
&= [N]\setminus\mathcal{X}
\end{aligned}$$

So it follows that

$$\begin{aligned}
\frac{|\mathcal{A}|}{N} &= \frac{|\mathcal{P}(W^*,\beta)|}{N} + \frac{|\mathcal{G}_N(W,\beta)|}{N} - \frac{1}{N}\left|[N]\setminus\mathcal{X}\right|\\
&= \frac{|\mathcal{P}(W^*,\beta)|}{N} + \frac{|\mathcal{G}_N(W,\beta)|}{N} - 1 + \frac{|\mathcal{X}|}{N}.
\end{aligned}$$

In the limit as N becomes large, we know from Proposition 4 that

$$\lim_{N\to\infty}\frac{|\mathcal{P}(W^*,\beta)|}{N} = 1-I(W^*),$$

and from Theorem 2 that

$$\lim_{N\to\infty}\frac{|\mathcal{G}_N(W,\beta)|}{N} = I(W).$$

[1]This is a slight variation of the Mahdavifar-Vardy coding scheme that ensures both reliability and strong security. Mahdavifar and Vardy were inconclusive about the reliability of their coding scheme because they were unable to make any statements about the reliability of the channels in $\mathcal{X}$. This minor variation with the addition of a secret key ensures security and reliability because a secret key is a hybrid of a frozen bit and a randomized bit. It is similar to a frozen bit in that its value is available to Bob and thus he does not need to decode the bit channels with secret key input. It is similar to a randomized bit from the assumption that its value is uniform and unknown to Eve.

Finally, we later show that $\lim_{N\to\infty} |\mathcal{X}|/N = 0$. The statement of the theorem then follows. ∎

The following theorem demonstrates that the quantum wiretap polar coding scheme has strong security:

Theorem 6: For the quantum wiretap polar coding scheme given above, a degraded quantum wiretap channel with W and W^* as defined in (8-9), with W^* having a classical output, and for sufficiently large N, it satisfies the following strong security criterion:

$$I\left(U_{\mathcal{A}}; E^n\right) = o\left(2^{-\frac{1}{2}N^\beta}\right).$$

Proof: Consider that

$$\begin{aligned} I\left(U_{\mathcal{A}}; E^n\right) &= \sum_{i\in\mathcal{A}} I\left(U_i; E^n | U_{\mathcal{A}_i^-}\right) \\ &= \sum_{i\in\mathcal{A}} I\left(U_i; E^n U_{\mathcal{A}_i^-}\right) \\ &\leq \sum_{i\in\mathcal{A}} I\left(U_i; E^n U_1^{i-1}\right) \\ &= \sum_{i\in\mathcal{A}} I(W_N^{*(i)}) \end{aligned}$$

The first equality is from the chain rule for quantum mutual information and by defining $\mathcal{A}_i^-$ to be the indices in $\mathcal{A}$ preceding i. The second equality follows from the assumption that the bits in $U_{\mathcal{A}_i^-}$ are chosen uniformly at random. The firts inequality is from quantum data processing. The third equality is from the definition of the synthesized channels $W_N^{*(i)}$. Continuing, we have

$$\begin{aligned} &\leq \sum_{i\in\mathcal{A}} \sqrt{1 - F(W_N^{*(i)})} \\ &\leq \sum_{i\in\mathcal{A}} \sqrt{1 - \left(1 - 2^{-N^\beta}\right)^2} \\ &= o\left(2^{-\frac{1}{2}N^\beta}\right). \end{aligned}$$

The first inequality is from Proposition 1 in Ref. [49]. The final inequality follows from the definition of the set $\mathcal{A}$. ∎

We also know that the code has good reliability, in the sense that there exists a POVM $\left\{\Lambda_{u_{\mathcal{A}}, u_{\mathcal{Y}}}^{(u_{\mathcal{X}})}\right\}$ such that

$$\begin{aligned} \Pr\{\widehat{U}_{\mathcal{A}\cup\mathcal{Y}} \neq U_{\mathcal{A}\cup\mathcal{Y}}\} &\leq 2\sqrt{\sum_{i\in\mathcal{A}\cup\mathcal{Y}} \frac{1}{2}\sqrt{F(W_N^{(i)})}} \\ &= o\left(2^{-\frac{1}{2}N^\beta}\right). \end{aligned}$$

This POVM is the quantum successive cancellation decoder established in Ref. [49]. The quantum successive cancellation decoder operates exactly as before, but it needs to decode both the information bits in $\mathcal{A}$ and the randomized bits in $\mathcal{Y}$. It also exploits the frozen bits in $\mathcal{B}$ and the secret key bits in $\mathcal{X}$ to help with decoding.

Finally, we can prove that the rate of secret key bits required by the scheme vanishes in the limit as N becomes large:

Proposition 7: For the quantum wiretap polar coding scheme given above, a degraded quantum wiretap channel with W and W^* as defined in (8-9), with W^* having a classical output, the rate $|\mathcal{X}|/N$ vanishes as N becomes large:

$$\lim_{N\to\infty} \frac{|\mathcal{X}|}{N} = 0.$$

Proof: This result follows by an argument similar to that for Proposition 22 in Ref. [31], but we need to modify it slightly. We prove that the sets $\mathcal{X}$, $\mathcal{G}_N(W^*, \beta)$, and $\mathcal{P}_N(W^*, \beta)$ are pairwise disjoint (note that we define the set $\mathcal{G}_N(W^*, \beta)$ as in (4), but with respect to the channel W^*). It then follows that

$$\frac{|\mathcal{X}|}{N} + \frac{|\mathcal{G}_N(W^*, \beta)|}{N} + \frac{|\mathcal{P}_N(W^*, \beta)|}{N} \leq 1. \tag{14}$$

So we prove that these sets are disjoint. First, consider that $\mathcal{X}$ and $\mathcal{P}_N(W^*, \beta)$ are disjoint by definition because $\mathcal{X}$ is formed from an intersection with $\mathcal{P}_N^c(W^*, \beta)$. Next, observe that for sufficiently large N, $\mathcal{P}_N(W^*, \beta)$ and $\mathcal{G}_N(W^*, \beta)$ are disjoint by definition. Observe that $\mathcal{X}$ and $\mathcal{G}_N(W^*, \beta)$ are disjoint because

$$\mathcal{B}_N(W, \beta) \subseteq \mathcal{B}_N(W^*, \beta),$$

which follows from W^* being a degraded version of W^* and Lemma 3 (also, $\mathcal{B}_N(W^*, \beta)$ is the complement of $\mathcal{G}_N(W^*, \beta)$). By applying Proposition 4, we know that

$$\lim_{N\to\infty} \frac{|\mathcal{P}(W^*, \beta)|}{N} = 1 - I(W^*),$$

and from Theorem 2, we know that

$$\lim_{N\to\infty} \frac{|\mathcal{G}_N(W^*, \beta)|}{N} = I(W^*).$$

Thus, the statement of the proposition follows from (14) and the above asymptotic limits. ∎

IV. Quantum polar codes

From such a scheme for private classical communication over a quantum wiretap channel with classical environment, we can readily construct a quantum polar code achieving the coherent information of a degradable quantum channel by exploiting Devetak's ideas for quantum coding [11] and the recent quantum successive cancellation decoder from Ref. [49]. First, recall that a quantum channel W is specified by a completely-positive trace-preserving map (we consider quantum channels with qubit inputs in this work). Any such map has a dilation to a larger system in which the dynamics over a tensor product space are unitary, i.e., it holds that

$$W(\rho) = \mathrm{Tr}_E\left\{U_W^{A'\to BE} \rho (U_W^{A'\to BE})^\dagger\right\},$$

where $U_W^{A'\to BE}$ is the isometric extension of the channel W. The complementary channel W^* is the map obtained by tracing over Bob's system

$$W^*(\rho) = \mathrm{Tr}_B\left\{U_{W^*}^{A'\to BE} \rho (U_{W^*}^{A'\to BE})^\dagger\right\}.$$

Such a realization makes the quantum coding setting analogous to the quantum wiretap setting. We also define the symmetric coherent information of the channel as

$$I_c(W) = H(B) - H(AB),$$

where the entropies result from sending half of a maximally entangled Bell state through the input of the channel. It is straightforward to verify that

$$I_c(W) = I(W) - I(W^*).$$

Our strategy for achieving the coherent information is to operate the quantum wiretap polar code in superposition (just as Devetak does [11]).

We can similarly specify a quantum polar code by the parameter vector $(N, K, \mathcal{A}, \mathcal{B}, \mathcal{X}, \mathcal{Y}, u_{\mathcal{B}})$ where these parameters are all the same as in the quantum wiretap polar code. The encoder is a coherent version of Arikan's encoder where we replace classical CNOT gates with quantum CNOT gates that act as follows on a two-qubit state:

$$\sum_{x,y} \alpha_{x,y} |x\rangle |y\rangle \rightarrow \sum_{x,y} \alpha_{x,y} |x\rangle |x \oplus y\rangle . \tag{15}$$

It is important to choose the orthonormal basis for the CNOT to be the one such that the induced states for the environment commute. That is, consider the following induced classical-quantum channel for the environment Eve:

$$x \rightarrow W^*(|x\rangle\langle x|) \equiv \rho_x^E,$$

where the basis $\{|x\rangle\}$ is the same as in (15). The scheme in this section works at the claimed rates if ρ_0^E and ρ_1^E commute (so that we can exploit the result in Proposition 4). We prove in Appendix B that many important channels satisfy this criterion.

We can now state our quantum polar coding theorem:

Theorem 8 (Quantum Polar Coding): For any degradable qubit-input quantum channel W with classical environment, there exists a quantum polar coding scheme for entanglement generation that achieves the symmetric coherent information, in the sense that the fidelity between the input entanglement and the generated entanglement is equal to $1 - o\left(2^{-\frac{1}{4}N^\beta}\right)$ where N is the blocklength of the code and β is some real such that $0 < \beta < 1/2$. The scheme may require entanglement assistance, but the entanglement consumption rate vanishes in the limit of large blocklength.

Proof: We assume that our task is merely to generate entanglement between Alice and Bob.[2] Alice begins by preparing Bell states locally on her side of the channel. Also, we assume that Alice and Bob share a small number of ebits before communication begins. We have the following structure for our code:

1) Alice sends half of the locally prepared Bell states through the channels in $\mathcal{A}$.
2) Alice sends the frozen ancilla qubits $|u_{\mathcal{B}}\rangle$ through the channels in $\mathcal{B}$.
3) Alice sends $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ states through the channels in $\mathcal{Y}$ (these are bits frozen in the Hadamard basis).
4) Alice sends her shares of the ebits through the channels in $\mathcal{X}$.

[2]The task of entanglement generation is equivalent to the task of quantum communication if forward classical communication from sender to receiver is available. Furthermore, forward classical communication does not increase the capacity of a quantum channel [4], [48].

Thus, the state before it is input to the encoder is as follows:

$$\frac{1}{\sqrt{2^{|\mathcal{A}|}}} \sum_{u_{\mathcal{A}}} |u_{\mathcal{A}}\rangle |u_{\mathcal{A}}\rangle |u_{\mathcal{B}}\rangle \frac{1}{\sqrt{2^{|\mathcal{Y}|}}} \sum_{u_{\mathcal{Y}}} |u_{\mathcal{Y}}\rangle \frac{1}{\sqrt{2^{|\mathcal{X}|}}} \sum_{u_{\mathcal{X}}} |u_{\mathcal{X}}\rangle |u_{\mathcal{X}}\rangle ,$$

where Alice possesses both shares of $\frac{1}{\sqrt{2^{|\mathcal{A}|}}} \sum_{u_{\mathcal{A}}} |u_{\mathcal{A}}\rangle |u_{\mathcal{A}}\rangle$, Alice possesses $|u_{\mathcal{B}}\rangle$ and the superposed state $\frac{1}{\sqrt{2^{|\mathcal{Y}|}}} \sum_{u_{\mathcal{Y}}} |u_{\mathcal{Y}}\rangle$, and Alice and Bob share the entangled state $\frac{1}{\sqrt{2^{|\mathcal{X}|}}} \sum_{u_{\mathcal{X}}} |u_{\mathcal{X}}\rangle |u_{\mathcal{X}}\rangle$. We can also write the above state as

$$\frac{1}{\sqrt{2^{|\mathcal{A}|+|\mathcal{Y}|+|\mathcal{X}|}}} \sum_{u_{\mathcal{A}},u_{\mathcal{Y}},u_{\mathcal{X}}} |u_{\mathcal{A}}\rangle |u_{\mathcal{A}}\rangle |u_{\mathcal{B}}\rangle |u_{\mathcal{Y}}\rangle |u_{\mathcal{X}}\rangle |u_{\mathcal{X}}\rangle ,$$

and we furthermore require Alice to apply some gates that realize some relative phases $\gamma_{u_{\mathcal{Y}}}$ so that the above state becomes

$$\frac{1}{\sqrt{2^{|\mathcal{A}|+|\mathcal{Y}|+|\mathcal{X}|}}} \sum_{u_{\mathcal{A}},u_{\mathcal{Y}},u_{\mathcal{X}}} e^{i\gamma_{u_{\mathcal{Y}}}} |u_{\mathcal{A}}\rangle |u_{\mathcal{A}}\rangle |u_{\mathcal{B}}\rangle |u_{\mathcal{Y}}\rangle |u_{\mathcal{X}}\rangle |u_{\mathcal{X}}\rangle .$$

(It is possible to realize these phases with only linear overhead in the encoding. Also, we specify how to choose these phases later.) Alice then applies a coherent version of Arikan's CNOT encoder [2], leading to the following encoded state:

$$\frac{1}{\sqrt{2^{|\mathcal{A}|}}} \sum_{u_{\mathcal{A}}} |u_{\mathcal{A}}\rangle |\phi_{u_{\mathcal{A}}}\rangle ,$$

where $\{|\phi_{u_{\mathcal{A}}}\rangle\}_{u_{\mathcal{A}}}$ are the *entanglement-assisted quantum codewords*, given by

$$\begin{aligned} |\phi_{u_{\mathcal{A}}}\rangle &= \frac{1}{\sqrt{2^{|\mathcal{Y}|+|\mathcal{X}|}}} \sum_{u_{\mathcal{Y}},u_{\mathcal{X}}} e^{i\gamma_{u_{\mathcal{Y}}}} |\psi_{u_{\mathcal{A}},u_{\mathcal{B}},u_{\mathcal{Y}},u_{\mathcal{X}}}\rangle |u_{\mathcal{X}}\rangle \\ &\equiv U \frac{1}{\sqrt{2^{|\mathcal{Y}|+|\mathcal{X}|}}} \sum_{u_{\mathcal{Y}},u_{\mathcal{X}}} e^{i\gamma_{u_{\mathcal{Y}}}} |u_{\mathcal{A}}\rangle |u_{\mathcal{B}}\rangle |u_{\mathcal{Y}}\rangle |u_{\mathcal{X}}\rangle |u_{\mathcal{X}}\rangle , \end{aligned}$$

with the coherent Arikan CNOT encoder U acting only on Alice's registers. Observe that the above state is encoded in a Calderbank-Shor-Steane code [10], [46] (modulo the relative phases) because the inputs are either information qubits, ancillas in a fixed state $|0\rangle$ or $|1\rangle$, ancillas in a state $|+\rangle$, or shares of ebits, and furthermore, the encoder consists of just SWAP gates and CNOT gates. Alice transmits the register containing the states $|\phi_{u_{\mathcal{A}}}\rangle$ over the quantum channel, leading to the following state

$$\frac{1}{\sqrt{2^{|\mathcal{A}|}}} \sum_{u_{\mathcal{A}}} |u_{\mathcal{A}}\rangle |\phi_{u_{\mathcal{A}}}\rangle^{B^N E^N} ,$$

where

$$|\phi_{u_{\mathcal{A}}}\rangle^{B^N E^N} \equiv U_W^{A^N \rightarrow B^N E^N} |\phi_{u_{\mathcal{A}}}\rangle^{A^N} .$$

Bob then applies the following coherent quantum successive cancellation decoder to his systems B^N and his half of the entanglement $B^{\mathcal{X}}$:

$$\sum_{u_{\mathcal{A}},u_{\mathcal{Y}},u_{\mathcal{X}}} \sqrt{\Lambda_{u_{\mathcal{A}},u_{\mathcal{Y}}}^{(u_{\mathcal{X}})}}^{B^N} \otimes |u_{\mathcal{X}}\rangle\langle u_{\mathcal{X}}| \otimes |u_{\mathcal{A}}, u_{\mathcal{Y}}\rangle^{\widehat{B}} . \tag{16}$$

The POVM elements $\left\{\Lambda_{u_{\mathcal{A}},u_{\mathcal{Y}}}^{(u_{\mathcal{X}})}\right\}_{u_{\mathcal{A}},u_{\mathcal{Y}},u_{\mathcal{X}}}$ are the same as those in the incoherent quantum successive cancellation decoder from Ref. [49]. Placing them in the above operation

allows this first part of the decoder to be an isometry and for the code to thus operate in superposition. We claim that the following state has fidelity $1 - o\left(2^{-\frac{1}{2}N^\beta}\right)$ with the state after Bob applies the above coherent quantum successive cancellation decoder:

$$\frac{1}{\sqrt{2^{|\mathcal{A}|+|\mathcal{Y}|+|\mathcal{X}|}}} \sum_{u_{\mathcal{A}},u_{\mathcal{Y}},u_{\mathcal{X}}} e^{i\delta_{u_{\mathcal{Y}}}} |u_{\mathcal{A}}\rangle |\psi_{u_{\mathcal{A}},u_{\mathcal{B}},u_{\mathcal{Y}},u_{\mathcal{X}}}\rangle^{B^N E^N} \otimes |u_{\mathcal{X}}\rangle |u_{\mathcal{A}}, u_{\mathcal{Y}}\rangle^{\widehat{B}}, \quad (17)$$

where we specify the phases $e^{i\delta_{u_{\mathcal{Y}}}}$ later. Indeed, consider the following states

$$|\chi_{u_{\mathcal{Y}}}\rangle \equiv \frac{1}{\sqrt{2^{|\mathcal{A}|+|\mathcal{X}|}}} \sum_{u_{\mathcal{A}},u_{\mathcal{X}},u'_{\mathcal{A}},u'_{\mathcal{Y}}} |u_{\mathcal{A}}\rangle \otimes \sqrt{\Lambda^{(u_{\mathcal{X}})}_{u'_{\mathcal{A}},u'_{\mathcal{Y}}}}^{B^N} |\psi_{u_{\mathcal{A}},u_{\mathcal{B}},u_{\mathcal{Y}},u_{\mathcal{X}}}\rangle^{B^N E^N} |u_{\mathcal{X}}\rangle \left|u'_{\mathcal{A}}, u'_{\mathcal{Y}}\right\rangle^{\widehat{B}},$$

$$|\varphi_{u_{\mathcal{Y}}}\rangle \equiv \frac{1}{\sqrt{2^{|\mathcal{A}|+|\mathcal{X}|}}} \sum_{u_{\mathcal{A}},u_{\mathcal{X}}} |u_{\mathcal{A}}\rangle |\psi_{u_{\mathcal{A}},u_{\mathcal{B}},u_{\mathcal{Y}},u_{\mathcal{X}}}\rangle^{B^N E^N} \otimes |u_{\mathcal{X}}\rangle |u_{\mathcal{A}}, u_{\mathcal{Y}}\rangle^{\widehat{B}}.$$

Their overlap is as follows:

$$\begin{aligned}
\langle\varphi_{u_{\mathcal{Y}}}|\chi_{u_{\mathcal{Y}}}\rangle &= \frac{1}{2^{|\mathcal{A}|+|\mathcal{X}|}} \sum_{u''_{\mathcal{A}},u''_{\mathcal{X}},u_{\mathcal{A}},u_{\mathcal{X}},u'_{\mathcal{A}},u'_{\mathcal{Y}}} \left(\langle u''_{\mathcal{A}}| \left\langle\psi_{u''_{\mathcal{A}},u_{\mathcal{B}},u_{\mathcal{Y}},u''_{\mathcal{X}}}\right|^{B^N E^N} \langle u''_{\mathcal{X}}| \langle u''_{\mathcal{A}}, u_{\mathcal{Y}}|^{\widehat{B}}\right) \times \\
&\left(\sqrt{\Lambda^{(u_{\mathcal{X}})}_{u'_{\mathcal{A}},u'_{\mathcal{Y}}}}^{B^N} |u_{\mathcal{A}}\rangle |\psi_{u_{\mathcal{A}},u_{\mathcal{B}},u_{\mathcal{Y}},u_{\mathcal{X}}}\rangle^{B^N E^N} |u_{\mathcal{X}}\rangle \left|u'_{\mathcal{A}}, u'_{\mathcal{Y}}\right\rangle^{\widehat{B}}\right) \\
&= \frac{1}{2^{|\mathcal{A}|+|\mathcal{X}|}} \sum_{u''_{\mathcal{A}},u''_{\mathcal{X}},u_{\mathcal{A}},u_{\mathcal{X}},u'_{\mathcal{A}},u'_{\mathcal{Y}}} \langle u''_{\mathcal{A}}|u_{\mathcal{A}}\rangle \times \\
&\left\langle\psi_{u''_{\mathcal{A}},u_{\mathcal{B}},u_{\mathcal{Y}},u''_{\mathcal{X}}}\right| \sqrt{\Lambda^{(u_{\mathcal{X}})}_{u'_{\mathcal{A}},u'_{\mathcal{Y}}}}^{B^N} |\psi_{u_{\mathcal{A}},u_{\mathcal{B}},u_{\mathcal{Y}},u_{\mathcal{X}}}\rangle^{B^N E^N} \times \\
&\langle u''_{\mathcal{X}}|u_{\mathcal{X}}\rangle \langle u''_{\mathcal{A}}|u'_{\mathcal{A}}\rangle \langle u_{\mathcal{Y}}|u'_{\mathcal{Y}}\rangle \\
&= \frac{1}{2^{|\mathcal{A}|+|\mathcal{X}|}} \times \\
&\sum_{u_{\mathcal{A}},u_{\mathcal{X}}} \langle\psi_{u_{\mathcal{A}},u_{\mathcal{B}},u_{\mathcal{Y}},u_{\mathcal{X}}}| \sqrt{\Lambda^{(u_{\mathcal{X}})}_{u_{\mathcal{A}},u_{\mathcal{Y}}}}^{B^N} |\psi_{u_{\mathcal{A}},u_{\mathcal{B}},u_{\mathcal{Y}},u_{\mathcal{X}}}\rangle^{B^N E^N} \\
&\geq \frac{1}{2^{|\mathcal{A}|+|\mathcal{X}|}} \sum_{u_{\mathcal{A}},u_{\mathcal{X}}} \operatorname{Tr}\left\{\Lambda^{(u_{\mathcal{X}})B^N}_{u_{\mathcal{A}},u_{\mathcal{Y}}} \psi^{B^N}_{u_{\mathcal{A}},u_{\mathcal{B}},u_{\mathcal{Y}},u_{\mathcal{X}}}\right\}.
\end{aligned}$$

We can then consider the average overlap:

$$\begin{aligned}
&\frac{1}{2^{|\mathcal{Y}|}} \sum_{u_{\mathcal{Y}}} \langle\varphi_{u_{\mathcal{Y}}}|\chi_{u_{\mathcal{Y}}}\rangle \\
&\geq \frac{1}{2^{|\mathcal{A}|+|\mathcal{X}|+|\mathcal{Y}|}} \sum_{u_{\mathcal{A}},u_{\mathcal{X}},u_{\mathcal{Y}}} \operatorname{Tr}\left\{\Lambda^{(u_{\mathcal{X}})B^N}_{u_{\mathcal{A}},u_{\mathcal{Y}}} \psi^{B^N}_{u_{\mathcal{A}},u_{\mathcal{B}},u_{\mathcal{Y}},u_{\mathcal{X}}}\right\} \\
&\geq 1 - o\left(2^{-\frac{1}{2}N^\beta}\right),
\end{aligned}$$

where the last inequality follows from the performance of the corresponding quantum wiretap polar code. Thus, there must exist some phases $\delta_{u_{\mathcal{Y}}}$ and $\gamma_{u_{\mathcal{Y}}}$ such that the state in (17) has a large overlap with the state resulting from the action of the coherent quantum successive cancellation decoder in (16) (this follows from Lemma 4 of Ref. [11]). The state resulting from the coherent quantum successive cancellation decoder is then close to the state in (17), which we repeat below:

$$\frac{1}{\sqrt{2^{|\mathcal{A}|+|\mathcal{Y}|+|\mathcal{X}|}}} \sum_{u_{\mathcal{A}},u_{\mathcal{Y}},u_{\mathcal{X}}} e^{i\delta_{u_{\mathcal{Y}}}} |u_{\mathcal{A}}\rangle |\psi_{u_{\mathcal{A}},u_{\mathcal{B}},u_{\mathcal{Y}},u_{\mathcal{X}}}\rangle^{B^N E^N} \otimes |u_{\mathcal{X}}\rangle |u_{\mathcal{A}}, u_{\mathcal{Y}}\rangle^{\widehat{B}}.$$

For a particular value of $u_{\mathcal{A}}$, the above state is

$$\frac{1}{\sqrt{2^{|\mathcal{Y}|+|\mathcal{X}|}}} \sum_{u_{\mathcal{Y}},u_{\mathcal{X}}} e^{i\delta_{u_{\mathcal{Y}}}} |u_{\mathcal{A}}\rangle |\psi_{u_{\mathcal{A}},u_{\mathcal{B}},u_{\mathcal{Y}},u_{\mathcal{X}}}\rangle^{B^N E^N} \otimes |u_{\mathcal{X}}\rangle |u_{\mathcal{A}}, u_{\mathcal{Y}}\rangle^{\widehat{B}}.$$

Tracing over Bob's systems gives the following state

$$\psi^{E^N}_{u_{\mathcal{A}}} \equiv \frac{1}{2^{|\mathcal{X}|+|\mathcal{Y}|}} \sum_{u_{\mathcal{X}},u_{\mathcal{Y}}} \psi^{E^N}_{u_{\mathcal{A}},u_{\mathcal{B}},u_{\mathcal{Y}},u_{\mathcal{X}}}.$$

We know from the quantum wiretap polar code that the following property holds

$$I\left(U_{\mathcal{A}}; E^n\right) = o\left(2^{-\frac{1}{2}N^\beta}\right).$$

Thus, from the fact that relative entropy upper bounds the trace distance (Theorem 11.9.2 of Ref. [48]), we know that

$$\begin{aligned}
2\ln 2\sqrt{I\left(U_{\mathcal{A}}; E^n\right)} &\geq \left\|\rho^{U_{\mathcal{A}}} \otimes \rho^{E^n} - \rho^{U_{\mathcal{A}}E^n}\right\|_1 \\
&= \frac{1}{2^{|\mathcal{A}|}} \sum_{u_{\mathcal{A}}} \left\|\psi^{E^N}_{u_{\mathcal{A}}} - \psi^{E^N}\right\|_1,
\end{aligned}$$

where

$$\psi^{E^N} \equiv \frac{1}{2^{|\mathcal{A}|}} \sum_{u_{\mathcal{A}}} \psi^{E^N}_{u_{\mathcal{A}}}.$$

Thus, it follows that

$$\frac{1}{2^{|\mathcal{A}|}} \sum_{u_{\mathcal{A}}} \left\|\psi^{E^N}_{u_{\mathcal{A}}} - \psi^{E^N}\right\|_1 = o\left(2^{-\frac{1}{4}N^\beta}\right),$$

and from the relationship between trace distance and fidelity [48], we have that

$$\frac{1}{2^{|\mathcal{A}|}} \sum_{u_{\mathcal{A}}} F(\psi^{E^N}_{u_{\mathcal{A}}}, \psi^{E^N}) = 1 - o\left(2^{-\frac{1}{4}N^\beta}\right).$$

So, for each $u_{\mathcal{A}}$, consider that Bob possesses the purification of the state $\psi^{E^N}_{u_{\mathcal{A}}}$ or ψ^{E^N} and by Uhlmann's theorem [47], [48], there exist isometries $U_{u_{\mathcal{A}}}$ such that

$$F(\psi^{E^N}_{u_{\mathcal{A}}}, \psi^{E^N}) = F(U_{u_{\mathcal{A}}} \psi^{B^N E^N}_{u_{\mathcal{A}}} U^\dagger_{u_{\mathcal{A}}}, \psi^{B^N E^N}),$$

for all $u_{\mathcal{A}}$, where we observe that $\psi^{B^N E^N}_{u_{\mathcal{A}}}$ is a purification of $\psi^{E^N}_{u_{\mathcal{A}}}$, defined as

$$|\psi_{u_{\mathcal{A}}}\rangle^{B^N E^N} = \sum_{u_{\mathcal{Y}},u_{\mathcal{X}}} \frac{1}{\sqrt{2^{|\mathcal{Y}|+|\mathcal{X}|}}} e^{i\delta_{u_{\mathcal{Y}}}} \left(|\psi_{u_{\mathcal{A}},u_{\mathcal{B}},u_{\mathcal{Y}},u_{\mathcal{X}}}\rangle^{B^N E^N} |u_{\mathcal{X}}\rangle |u_{\mathcal{Y}}\rangle^{\widehat{B}}\right).$$

Thus, Bob's final task is to apply the controlled unitary on his state

$$\sum_{u_{\mathcal{A}}} |u_{\mathcal{A}}\rangle \langle u_{\mathcal{A}}|^{\widehat{B}} \otimes U_{u_{\mathcal{A}}},$$

leading to the following state

$$\frac{1}{\sqrt{2^{|\mathcal{A}|}}} \sum_{u_{\mathcal{A}}} |u_{\mathcal{A}}\rangle |u_{\mathcal{A}}\rangle^{\widehat{B}} U_{u_{\mathcal{A}}} \sum_{u_{\mathcal{Y}}, u_{\mathcal{X}}} \frac{1}{\sqrt{2^{|\mathcal{Y}|+|\mathcal{X}|}}} e^{i\delta_{u_{\mathcal{Y}}}} \times \left(|\psi_{u_{\mathcal{A}}, u_{\mathcal{B}}, u_{\mathcal{Y}}, u_{\mathcal{X}}}\rangle^{B^N E^N} |u_{\mathcal{X}}\rangle |u_{\mathcal{Y}}\rangle^{\widehat{B}} \right).$$

The desired, ideal state is as follows:

$$\frac{1}{\sqrt{2^{|\mathcal{A}|}}} \sum_{u_{\mathcal{A}}} |u_{\mathcal{A}}\rangle |\psi\rangle^{B^N E^N} |u_{\mathcal{A}}\rangle^{\widehat{B}},$$

and, after a straightforward calculation, the overlap between the actual state and the desired state is equal to

$$\begin{aligned} &\frac{1}{2^{|\mathcal{A}|}} \sum_{u_{\mathcal{A}}} F(U_{u_{\mathcal{A}}} \psi_{u_{\mathcal{A}}}^{B^N E^N} U_{u_{\mathcal{A}}}^{\dagger}, \psi^{B^N E^N}) \\ &= \frac{1}{2^{|\mathcal{A}|}} \sum_{u_{\mathcal{A}}} F(\psi_{u_{\mathcal{A}}}^{E^N}, \psi^{E^N}) \\ &= 1 - o\left(2^{-\frac{1}{4} N^{\beta}}\right). \end{aligned}$$

Bob's final move is to discard his register in B^N, leading to a state close to the following entangled state shared between Alice and Bob:

$$\frac{1}{\sqrt{2^{|\mathcal{A}|}}} \sum_{u_{\mathcal{A}}} |u_{\mathcal{A}}\rangle |u_{\mathcal{A}}\rangle^{\widehat{B}}$$

Putting everything together, we have a scheme that generates entanglement with a fidelity equal to

$$1 - o\left(2^{-\frac{1}{4} N^{\beta}}\right),$$

where the error results from two parts in the protocol: the first error is from the reliability of the quantum wiretap polar code and the second is from the strong security of the quantum wiretap polar code. ∎

The above scheme achieves the symmetric coherent information rate and has an efficient encoder, but it is unclear to us right now how to improve the efficiency of the two-step decoder. If one were able to improve the efficiency of the quantum successive cancellation decoder, this would lead to an efficient implementation of the first part of the decoding. Improving the efficiency of the second part might be possible by taking into account the particular structure of these quantum polar codes.

As a final point, we should note that the symmetric coherent information is equal to the quantum capacity of the quantum erasure channel, the quantum dephasing channel, and the universal cloning machine channel. This is not the case for the amplitude damping channel and the photon-detected jump channel, but the ratio between the true quantum capacity and the symmetric coherent information is close to one for these channels (see Figure 1).

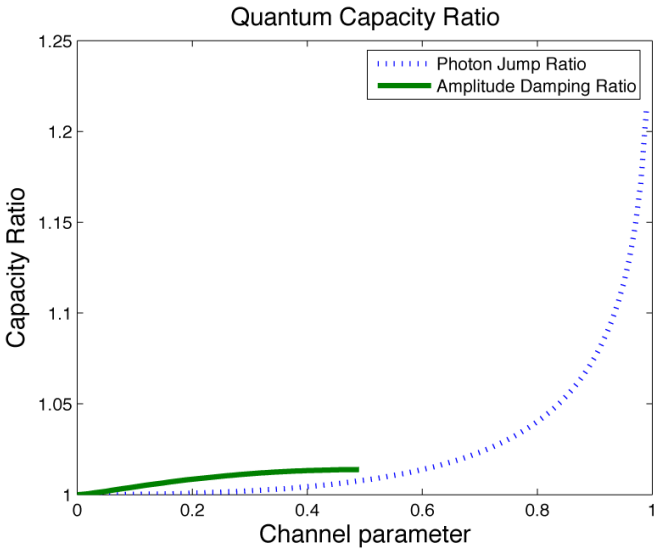

Fig. 1. The figure plots the ratio of the true quantum capacity to the symmetric coherent information as a function of the channel parameter for both the amplitude damping channel and the photon-detected jump channel (see Appendix B for definitions of these channels). Observe that the ratio is close to one for many values of the channel parameters.

V. Conclusion

We constructed polar codes for transmitting classical data privately over a quantum wiretap channel and for transmitting quantum data over a quantum channel. The rates achievable are respectively equal to the symmetric private information and the symmetric coherent information. In both settings, we require that the channel to the environment be effectively classical in order to guarantee that the rates are as claimed. The codes exploit the channel polarization phenomenon, first observed by Arikan in the classical setting [2] and later in Ref. [49] for classical-quantum channels.

Two of the most important problems left open in this paper are to determine an efficient quantum decoder and to find other channels outside of the class discussed here for which the symmetric coherent information is achievable. Progress on the first question is outlined in Ref. [50], though the main question of an efficient decoder still remains open. The recent work of Wilde and Renes resolves the second question [52], [51], [36], by adapting an earlier coding scheme of Renes and Boileau [34] to the polar coding setting.

One might question whether we should call the codes developed here and in Ref. [35] "quantum polar codes," if the criterion for a quantum polar code is that it be symmetric capacity-achieving, channel-adapted, and possessing efficient encoders and decoders. The codes constructed here are symmetric capacity-achieving and channel-adapted, but do not have efficient decoders. The codes from Ref. [35] have efficient encoders and decoders, but they are capacity-achieving and channel-adapted only for Pauli channels. Though, both constructions certainly take advantage of the channel polarization effect, which gives the code construction its name. Solving the open problems posed here would certainly lead to quantum polar codes that possess all desiderata.

ACKNOWLEDGEMENTS

MMW acknowledges financial support from the MDEIE (Québec) PSR-SIIRI international collaboration grant. SG was supported by the DARPA Information in a Photon (InPho) program under contract number HR0011-10-C-0159. The authors are grateful to S. Hamed Hassani for helpful discussions concerning Theorem 3 in Ref. [21] and Omar Fawzi and Patrick Hayden for useful discussions.

APPENDIX A
FIDELITY OF PURE STATE CQ CHANNELS REMAINS INVARIANT UNDER CHANNEL COMBINING

In the theorem below, we demonstrate that the fidelity for the "worse" channel W^- [49] is equal to the fidelity of the original channel W if the original channel is a classical-quantum state with pure state outputs. The implication is that the inequality in (10) fails for this case.

Theorem 9: Suppose we have two pure states $|\psi_0\rangle$ and $|\psi_1\rangle$ such that the classical-quantum channel W outputs $|\psi_0\rangle$ if zero is input and $|\psi_1\rangle$ if one is input. The fidelity between these two pure states is as follows:

$$F\left(|\psi_0\rangle, |\psi_1\rangle\right) = |\langle\psi_0|\psi_1\rangle|^2 .$$

The states arising from channel combining in the worse direction W^- [49] are as follows:

$$\rho_0^- = \frac{1}{2}\left(|\psi_0\rangle\langle\psi_0| \otimes |\psi_0\rangle\langle\psi_0| + |\psi_1\rangle\langle\psi_1| \otimes |\psi_1\rangle\langle\psi_1|\right),$$
$$\rho_1^- = \frac{1}{2}\left(|\psi_0\rangle\langle\psi_0| \otimes |\psi_1\rangle\langle\psi_1| + |\psi_1\rangle\langle\psi_1| \otimes |\psi_0\rangle\langle\psi_0|\right).$$

Then the following relation holds

$$F\left(W\right) = F\left(W^-\right).$$

Proof: Recall that the Uhlmann fidelity $F\left(\rho_0^-, \rho_1^-\right)$ is equal to the maximum squared overlap between purifications of ρ_0^- and ρ_1^-, where the maximum is over all purifications. Since both of the above states are rank two, it suffices to consider a two-dimensional purifying system for both. Furthermore, we can fix one purification while varying the other one. So, one purification of ρ_0^- is as follows:

$$|\phi_{\rho_0^-}\rangle = \frac{1}{\sqrt{2}}\left(|\psi_0\rangle|\psi_0\rangle|0\rangle + |\psi_1\rangle|\psi_1\rangle|1\rangle\right),$$

and a varying purification of ρ_1^- is as follows:

$$|\phi_{\rho_1^-}\rangle = \frac{1}{\sqrt{2}}\left(|\psi_0\rangle|\psi_1\rangle|\varphi\rangle + |\psi_1\rangle|\psi_0\rangle|\varphi^\perp\rangle\right),$$

where we can set

$$|\varphi\rangle = \alpha|0\rangle + \beta|1\rangle,$$
$$|\varphi^\perp\rangle = \beta^*|0\rangle - \alpha^*|1\rangle.$$

Our goal is to maximize the overlap $\left|\langle\phi_{\rho_0^-}|\phi_{\rho_1^-}\rangle\right|^2$ over all legitimate choices of α and β. The overlap $\left|\langle\phi_{\rho_0^-}|\phi_{\rho_1^-}\rangle\right|^2$ is as follows:

$$\begin{aligned}
&\left|\langle\phi_{\rho_0^-}|\phi_{\rho_1^-}\rangle\right|^2 \\
&= \frac{1}{4}\left|\begin{array}{c}\langle\psi_0|\psi_0\rangle\langle\psi_0|\psi_1\rangle\langle 0|\varphi\rangle + \langle\psi_1|\psi_0\rangle\langle\psi_1|\psi_1\rangle\langle 1|\varphi\rangle + \\ \langle\psi_0|\psi_1\rangle\langle\psi_0|\psi_0\rangle\langle 0|\varphi^\perp\rangle + \langle\psi_1|\psi_1\rangle\langle\psi_1|\psi_0\rangle\langle 1|\varphi^\perp\rangle\end{array}\right|^2 \\
&= \frac{1}{4}\left|\begin{array}{c}\langle\psi_0|\psi_1\rangle\alpha + \langle\psi_1|\psi_0\rangle\beta + \\ \langle\psi_0|\psi_1\rangle\beta^* - \langle\psi_1|\psi_0\rangle\alpha^*\end{array}\right|^2 \\
&= \frac{1}{4}\left|\langle\psi_0|\psi_1\rangle\left(\alpha+\beta^*\right) + \langle\psi_1|\psi_0\rangle\left(\beta-\alpha^*\right)\right|^2 \\
&= \frac{1}{4}\left|\langle\psi_0|\psi_1\rangle\left(\alpha+\beta^*\right) + \langle\psi_1|\psi_0\rangle\left(\beta-\alpha^*\right)\right|^2
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{4}\left|\langle\psi_0|\psi_1\rangle\right|^2\left|\alpha+\beta^*\right|^2 + \frac{1}{4}\left|\langle\psi_0|\psi_1\rangle\right|^2\left|\beta-\alpha^*\right|^2 \\
&\quad + \frac{1}{4}2\,\mathrm{Re}\left\{\langle\psi_1|\psi_0\rangle^2\left(\alpha^*+\beta\right)\left(\beta-\alpha^*\right)\right\} \\
&= \frac{1}{4}\left|\langle\psi_0|\psi_1\rangle\right|^2\left(\left|\alpha+\beta^*\right|^2 + \left|\beta-\alpha^*\right|^2\right) \\
&\quad + \frac{1}{4}2\,\mathrm{Re}\left\{\langle\psi_1|\psi_0\rangle^2\left(\beta^2 - \left(\alpha^*\right)^2\right)\right\} \\
&= \frac{1}{4}\left|\langle\psi_0|\psi_1\rangle\right|^2\left(|\alpha|^2 + |\beta|^2 + \right. \\
&\quad \left. 2\,\mathrm{Re}\left\{\alpha\beta^*\right\} + |\alpha|^2 + |\beta|^2 - 2\,\mathrm{Re}\left\{\alpha\beta^*\right\}\right) \\
&\quad + \frac{1}{4}2\,\mathrm{Re}\left\{\langle\psi_1|\psi_0\rangle^2\left(\beta^2 - \left(\alpha^*\right)^2\right)\right\} \\
&= \frac{1}{2}\left|\langle\psi_0|\psi_1\rangle\right|^2 + \frac{1}{2}\,\mathrm{Re}\left\{\langle\psi_1|\psi_0\rangle^2\left(\beta^2 - \left(\alpha^*\right)^2\right)\right\}
\end{aligned}$$

We set $\langle\psi_1|\psi_0\rangle = r_\psi e^{i\phi}$ (with $r_\psi = |\langle\psi_1|\psi_0\rangle|$) so that we have

$$\begin{aligned}
&\frac{1}{2}\,\mathrm{Re}\left\{\langle\psi_1|\psi_0\rangle^2\left(\beta^2 - \left(\alpha^*\right)^2\right)\right\} \\
&= \frac{1}{2}\,\mathrm{Re}\left\{r_\psi^2 e^{i2\phi}\left(r_1^2 e^{i2\theta_1} - r_2^2 e^{i2\theta_2}\right)\right\} \\
&= \frac{1}{2}r_\psi^2\,\mathrm{Re}\left\{r_1^2 e^{i2(\theta_1+\phi)} - r_2^2 e^{i2(\theta_2+\phi)}\right\} \\
&= \frac{1}{2}\left|\langle\psi_1|\psi_0\rangle\right|^2\left(r_1^2\cos\left(2\left(\theta_1+\phi\right)\right) - r_2^2\cos\left(2\left(\theta_2+\phi\right)\right)\right) \\
&\leq \frac{1}{2}\left|\langle\psi_1|\psi_0\rangle\right|^2\left(r_1^2 + r_2^2\right) \\
&= \frac{1}{2}\left|\langle\psi_1|\psi_0\rangle\right|^2
\end{aligned}$$

By choosing an appropriate α and β such that $\cos\left(2\left(\theta_1+\phi\right)\right) = 1$ and $\cos\left(2\left(\theta_2+\phi\right)\right) = -1$, this demonstrates that the result holds for any pure states $|\psi_0\rangle$ and $|\psi_1\rangle$. ∎

APPENDIX B
EXAMPLES OF CHANNELS WITH CLASSICAL ENVIRONMENT

We prove in this section that several important degradable channels have a classical environment, in the sense that the classical-quantum channel induced by inputting classical orthonormal states at the input leads to commuting output states. That is, we would like to prove that $\left[\rho_0^E, \rho_1^E\right] = 0$ for several important channels, where

$$\rho_x^E = W^*\left(|x\rangle\langle x|\right),$$

W^* is the complementary channel, and $\{|x\rangle\}$ is some orthonormal basis.

We begin with the amplitude damping channel. The complement of an amplitude damping channel with damping parameter γ has the following action on a qubit input [48]:

$$\begin{bmatrix} 1-p & \eta^* \\ \eta & p \end{bmatrix} \to \begin{bmatrix} 1-\gamma p & \sqrt{\gamma}\eta^* \\ \sqrt{\gamma}\eta & \gamma p \end{bmatrix},$$

where $0 \leq p, \gamma \leq 1$, η is a complex number such that the input matrix is positive, and the matrix representations are with respect to the computational basis $\{|0\rangle, |1\rangle\}$. (The complement is effectively an amplitude damping channel with damping parameter $1-\gamma$.) The result follows by observing that

$$|0\rangle\langle 0| \to |0\rangle\langle 0|,$$
$$|1\rangle\langle 1| \to (1-\gamma)|0\rangle\langle 0| + \gamma|1\rangle\langle 1|.$$

Consider the photon-detected jump channel from Refs. [1], [20]. The authors of Ref. [20] demonstrated that the complement of this channel is as follows:

$$\begin{bmatrix} 1-p & \eta^* \\ \eta & p \end{bmatrix} \to (1-\gamma p)|0\rangle\langle 0|^E + \gamma p|1\rangle\langle 1|^E.$$

So it is again clear that the computational basis suffices to make the environment outputs commute.

The complement of an erasure channel with erasure parameter ϵ is just an erasure channel with erasure parameter $1-\epsilon$ [48]:

$$\rho \to \epsilon\rho + (1-\epsilon)|e\rangle\langle e|,$$

where $|e\rangle$ is some erasure symbol orthogonal to the space of ρ. Thus, any basis suffices to demonstrate that the states for the environment commute.

A qubit dephasing channel with parameter p has the following form:

$$\rho \to (1-p)\rho + p\ \sigma_i\rho\sigma_i,$$

where σ_i is one of the Pauli operators. The complement of this channel has the following form for a pure state input $|\psi\rangle$ [48]:

$$(1-p)|0\rangle\langle 0| + \sqrt{p(1-p)}\langle\psi|\sigma_i|\psi\rangle(|0\rangle\langle 1| + |1\rangle\langle 0|) + p|1\rangle\langle 1|.$$

Thus, choosing the input basis to be the one for which σ_i acts as a "bit-flipping" operator leads to commuting outputs for the environment. For example, if $\sigma_i = X$, then the basis is $\{|0\rangle, |1\rangle\}$, while if $\sigma_i = Z$, the basis is $\{|+\rangle, |-\rangle\}$.

Finally, the complement of the cloning channel is available in Ref. [20]. Due to the covariance of the cloning channel and its complement, inputting any orthonormal basis leads to the following states on the output:

$$\psi_0^E = \sum_{i=0}^{N-1} \frac{i+1}{\Delta_N} |i\rangle\langle i|^E, \tag{18}$$

$$\psi_1^E = \sum_{i=0}^{N-1} \frac{i+1}{\Delta_N} |N-1-i\rangle\langle N-1-i|^E, \tag{19}$$

where N is the number of clones, $\Delta_N = N(N+1)/2$, and the states $\{|i\rangle^E\}$ form an orthonormal basis. Thus, the environmental states commute for these channels.

References

[1] G. Alber, Th. Beth, Ch. Charnes, A. Delgado, M. Grassl, and M. Mussinger. Stabilizing distinguishable qubits against spontaneous decay by detected-jump correcting quantum codes. *Physical Review Letters*, 86:4402–4405, May 2001. arXiv:quant-ph/0103042.

[2] Erdal Arikan. Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels. *IEEE Transactions on Information Theory*, 55(7):3051–3073, July 2009. arXiv:0807.3917.

[3] Erdal Arikan and Emre Telatar. On the rate of channel polarization. In *Proceedings of the 2009 International Symposium on Information Theory*, pages 1493–1495, Seoul, Korea, June 2009. arXiv:0807.3806.

[4] Howard Barnum, Emanuel Knill, and Michael A. Nielsen. On quantum fidelities and channel capacities. *IEEE Transactions on Information Theory*, 46:1317–1329, 2000.

[5] Howard Barnum, M. A. Nielsen, and Benjamin Schumacher. Information transmission through a noisy quantum channel. *Physical Review A*, 57(6):4153–4175, June 1998.

[6] Kamil Brádler. An infinite sequence of additive channels: The classical capacity of cloning channels. *IEEE Transactions on Information Theory*, 57(8):5497–5503, August 2011. arXiv:0903.1638.

[7] Frederico Brito, David P. DiVincenzo, Roger H. Koch, and Matthias Steffen. Efficient one- and two-qubit pulsed gates for an oscillator-stabilized Josephson qubit. *New Journal of Physics*, 10(3):033027 (33pp), 2008.

[8] Todd A. Brun, Igor Devetak, and Min-Hsiu Hsieh. Correcting quantum errors with entanglement. *Science*, 314(5798):436–439, October 2006.

[9] Vladimir Buzek and Mark Hillery. Universal optimal cloning of arbitrary quantum states: From qubits to quantum registers. *Physical Review Letters*, 81:5003–5006, November 1998.

[10] A. Robert Calderbank and Peter W. Shor. Good quantum error-correcting codes exist. *Physical Review A*, 54(2):1098–1105, August 1996.

[11] Igor Devetak. The private classical capacity and quantum capacity of a quantum channel. *IEEE Transactions on Information Theory*, 51(1):44, January 2005. arXiv:quant-ph/0304127.

[12] Igor Devetak and Peter W. Shor. The capacity of a quantum channel for simultaneous transmission of classical and quantum information. *Communications in Mathematical Physics*, 256:287–303, 2005.

[13] Simon J. Devitt, Kae Nemoto, and William J. Munro. Quantum error correction for beginners. May 2009. arXiv:0905.2794.

[14] G. David Forney and Daniel J. Costello. Channel coding: The road to channel capacity. *Proceedings of the IEEE*, 95(6):1150–1177, June 2007.

[15] Robert M. Gingrich, Pieter Kok, Hwang Lee, Farrokh Vatan, and Jonathan P. Dowling. All linear optical quantum memory based on quantum error correction. *Physical Review Letters*, 91:217901, November 2003.

[16] Vittorio Giovannetti and Rosario Fazio. Information-capacity description of spin-chain correlations. *Physical Review A*, 71(3):032314, March 2005.

[17] N. Gisin and S. Massar. Optimal quantum cloning machines. *Physical Review Letters*, 79:2153–2156, September 1997. arXiv:quant-ph/9705046.

[18] Daniel Gottesman. *Stabilizer Codes and Quantum Error Correction*. PhD thesis, California Institute of Technology, 1997. arXiv:quant-ph/9705052.

[19] Markus Grassl, Thomas Beth, and Thomas Pellizzari. Codes for the quantum erasure channel. *Physical Review A*, 56(1):33–38, July 1997.

[20] Markus Grassl, Zhengfeng Ji, Zhaohui Wei, and Bei Zeng. Quantum-capacity-approaching codes for the detected-jump channel. *Physical Review A*, 82:062324, December 2010.

[21] S. Hamed Hassani and Ruediger Urbanke. On the scaling of polar codes: I. the behavior of polarized channels. *Proceedings of the 2010 IEEE International Symposium on Information Theory*, pages 874–878, June 2010. Austin, Texas, USA. arXiv:1001.2766.

[22] Carl W. Helstrom. Quantum detection and estimation theory. *Journal of Statistical Physics*, 1:231–252, 1969.

[23] Alexander S. Holevo. An analog of the theory of statistical decisions in noncommutative theory of probability. *Trudy Moscov Mat. Obsc.*, 26:133–149, 1972. English translation: Trans. Moscow Math Soc. 26, 133–149 (1972).

[24] Min-Hsiu Hsieh, Wen-Tai Yen, and Li-Yi Hsu. High performance entanglement-assisted quantum LDPC codes need little entanglement. *IEEE Transactions on Information Theory*, 57(3):1761–1769, 2011. arXiv:0906.5532.

[25] Kenta Kasai, Manabu Hagiwara, Hideki Imai, and Kohichi Sakaniwa. Quantum error correction beyond the bounded distance decoding limit. *IEEE Transactions on Information Theory*, 58(2):1223–1230, February 2012. arXiv:1007.1778.

[26] Satish Babu Korada. *Polar Codes for Channel and Source Coding*. PhD thesis, École Polytechnique Fédérale de Lausanne, July 2009.

[27] Antía Lamas-Linares, Christoph Simon, John C. Howell, and Dik Bouwmeester. Experimental quantum cloning of single photons. *Science*, 296:712–714, 2002.

[28] Seth Lloyd. Capacity of the noisy quantum channel. *Physical Review A*, 55(3):1613–1622, March 1997.

[29] Zhicheng Luo. Quantum error correcting codes based on privacy amplification. August 2008. arXiv:0808.1392.

[30] David J.C. MacKay, Graeme Mitchison, and Paul L. McFadden. Sparse graph codes for quantum error-correction. *IEEE Transactions on Information Theory*, 50(10):2315, October 2004.

[31] Hessam Mahdavifar and Alexander Vardy. Achieving the secrecy capacity of wiretap channels using polar codes. *IEEE Transactions on Information Theory*, 57(10):6428–6443, October 2011. arXiv:1001.0210.

[32] P. W. Milonni and M. L. Hardies. Photons cannot always be replicated. *Physics Letters A*, 92(7):321–322, November 1982.

[33] David Poulin, Jean-Pierre Tillich, and Harold Ollivier. Quantum serial turbo-codes. *IEEE Transactions on Information Theory*, 55(6):2776–2798, June 2009.

[34] Joseph M. Renes and Jean-Christian Boileau. Physical underpinnings of privacy. *Physical Review A*, 78:032335, September 2008.

[35] Joseph M. Renes, Frédéric Dupuis, and Renato Renner. Efficient quantum polar coding. *Physical Review Letters*, 109(5):050504, August 2012. arXiv:1109.3195.

[36] Joseph M. Renes and Mark M. Wilde. Polar codes for private and quantum communication over arbitrary channels. December 2012. arXiv:1212.2537.

[37] Benjamin Schumacher. Sending entanglement through noisy quantum channels. *Physical Review A*, 54(4):2614–2628, October 1996.

[38] Benjamin Schumacher and Michael A. Nielsen. Quantum data processing and error correction. *Physical Review A*, 54(4):2629–2635, October 1996.

[39] Benjamin Schumacher and Michael D. Westmoreland. Quantum privacy and quantum coherence. *Physical Review Letters*, 80(25):5695–5697, June 1998.

[40] Pranab Sen. Achieving the Han-Kobayashi inner bound for the quantum interference channel by sequential decoding. September 2011. arXiv:1109.0802.

[41] Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 1948.

[42] Peter W. Shor. Scheme for reducing decoherence in quantum computer memory. *Physical Review A*, 52:R2493–R2496, October 1995.

[43] Peter W. Shor. The quantum channel capacity and coherent information. In *Lecture Notes, MSRI Workshop on Quantum Computation*, 2002.

[44] Peter W. Shor and John Preskill. Simple proof of security of the BB84 quantum key distribution protocol. *Physical Review Letters*, 85(2):441–444, July 2000.

[45] Christoph Simon, Gregor Weihs, and Anton Zeilinger. Optimal quantum cloning via stimulated emission. *Physical Review Letters*, 84(13):2993–2996, March 2000.

[46] Andrew M. Steane. Error correcting codes in quantum theory. *Physical Review Letters*, 77(5):793–797, July 1996.

[47] Armin Uhlmann. The "transition probability" in the state space of a *-algebra. *Reports on Mathematical Physics*, 9(2):273–279, 1976.

[48] Mark M. Wilde. *From Classical to Quantum Shannon Theory*. June 2011. arXiv:1106.1445.

[49] Mark M. Wilde and Saikat Guha. Polar codes for classical-quantum channels. *IEEE Transactions on Information Theory*, 59(2):1175–1187, February 2013. arXiv:1109.2591.

[50] Mark M. Wilde, Olivier Landon-Cardinal, and Patrick Hayden. Towards efficient decoding of classical-quantum polar codes. February 2013. arXiv:1302.0398.

[51] Mark M. Wilde and Joseph M. Renes. Polar codes for private classical communication. In *Proceedings of the 2012 International Symposium on Information Theory and its Applications*, pages 745–749, Honolulu, Hawaii, USA, October 2012. arXiv:1203.5794.

[52] Mark M. Wilde and Joseph M. Renes. Quantum polar codes for arbitrary channels. In *Proceedings of the 2012 International Symposium on Information Theory*, pages 334–338, Boston, Massachusetts, USA, July 2012. arXiv:1201.2906.

Mark M. Wilde (M'99) was born in Metairie, Louisiana, USA. He received the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, California, in 2008. He is a Postdoctoral Fellow at the School of Computer Science, McGill University and will start in August 2013 as an Assistant Professor in the Department of Physics and Astronomy and the Center for Computation and Technology at Louisiana State University. His current research interests are in quantum Shannon theory, quantum optical communication, quantum computational complexity theory, and quantum error correction.

Saikat Guha was born in Patna, India, on July 3, 1980. He received the Ph.D. degree in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology (MIT), Cambridge, MA in 2008. He is currently a Senior Scientist with Raytheon BBN Technologies, Cambridge, MA, USA. His current research interest surrounds the application of quantum information and estimation theory to fundamental limits of optical communication and imaging. He is also interested in classical and quantum error correction, network information and communication theory, and quantum algorithms.