ISSN 0032-9460, Problems of Information Transmission, 2016, Vol. 52, No. 3, pp. 214–238. © Pleiades Publishing, Inc., 2016.
Original Russian Text © S. Karumanchi, S. Mancini, A. Winter, D. Yang, 2016, published in Problemy Peredachi Informatsii, 2016, Vol. 52, No. 3, pp. 17–44.

INFORMATION THEORY

Classical Capacities of Quantum Channels with Environment Assistance

S. Karumanchi[a,1], **S. Mancini**[b], **A. Winter**[c,2], **and D. Yang**[d,3]

[a] *Faculty of Informatics, Masaryk University, Brno, Czech Republic*
e-mail: siddharth.karumanchi@unicam.it
[b] *School of Science and Technology, University of Camerino, Camerino, Italy*
INFN–Sezione Perugia, Perugia, Italy
e-mail: stefano.mancini@unicam.it
[c] *ICREA and Física Teòrica: Informació i Fenòmens Quàntics,*
Universitat Autònoma de Barcelona, Barcelona, Spain
e-mail: andreas.winter@uab.cat
[d] *Física Teòrica: Informació i Fenòmens Quàntics,*
Universitat Autònoma de Barcelona, Barcelona, Spain
Laboratory for Quantum Information, China Jiliang University, Hangzhou, Zhejiang, China
e-mail: dyang@cjlu.edu.cn

Received September 23, 2015; in final form, February 3, 2016

Abstract—A quantum channel physically is a unitary interaction between an information carrying system and an environment, which is initialized in a pure state before the interaction. Conventionally, this state, as also the parameters of the interaction, is assumed to be fixed and known to the sender and receiver. Here, following the model introduced by us earlier [1], we consider a benevolent third party, i.e., a helper, controlling the environment state, and show how the helper's presence changes the communication game. In particular, we define and study the classical capacity of a unitary interaction with helper, in two variants: one where the helper can only prepare separable states across many channel uses, and one without this restriction. Furthermore, two even more powerful scenarios of pre-shared entanglement between helper and receiver, and of classical communication between sender and helper (making them conferencing encoders) are considered.

DOI: 10.1134/S0032946016030029

1. INTRODUCTION

The noise in quantum communication is modeled by a quantum channel, which is a completely positive and trace-preserving (CPTP) map on the set of states (density operators) of a system devoted to carry information. Note that this view includes classical channels as a special case (cf. [2]). Every quantum channel can be viewed as a unitary interaction between the information carrying system and an environment, where the latter is customarily considered not under control. Therefore, the initial environment state together with the unitary defines the channel when the final environment is traced out. In this standard picture, the initial environment state is simply fixed,

[1] The work was done when the author was with the School of Science and Technology, University of Camerino, Camerino, Italy, and INFN–Sezione Perugia, Perugia, Italy.

[2] Supported by the European Commission, STREP "RAQUEL," the European Research Council, advanced grant "IRQUAT," the Spanish MINECO, project no. FIS2013-40627-P, with the support of FEDER funds, as well as by the Generalitat de Catalunya CIRIT, project no. 2014-SGR-966.

[3] Supported by the European Research Council, advanced grant "IRQUAT," and the NSFC, grant no. 11375165.

and the environment output is completely lost. However, the possibility of an *active helper*, one that reads information from the channel environment and communicates it to the channel receiver, is an interesting one that has been considered before with some success [3–7]. In the present paper, instead, we shall be concerned with a benevolent party (a helper, henceforth called Helen) setting the initial environment state in order to assist sender and receiver of the channel to communicate. We considered transmission of quantum information in this model in our earlier work [1]. Here we look at classical communication: as in [1], we have a model of *passive environment assistance*, where Helen simply sets an initial state of the environment as a part of the code, once and for all; likewise, we motivated to consider *passive environment assistance with entanglement* between Helen and the receiver Bob of the channel output. Because classical information, unlike quantum information, can be freely shared, here we can then contrast these passive models with one where Helen's state can also depend on the message to be sent; we call it *conferencing encoders*, allowing local operations and classical communication (LOCC) between the sender Alice and Helen.

The structure of the paper is as follows. Section 2 introduces the notation and provides details of the proposed models. Section 3 contains coding theorems for passive environment-assisted capacities. There, after making general observations, we also provide examples of superadditivity of capacities. Then, in Section 4 we go on to study the entanglement-environment-assisted capacities. In Section 5 we make general observations about the conferencing encoders model, and go on to show that for a unitary operator the classical capacity with conferencing encoders is nonzero. Finally, the Appendix give details of the parametrization for two-qubit unitaries.

2. NOTATION AND MODELS

Let A, E, B, F, etc. be finite-dimensional Hilbert spaces, $\mathcal{L}(X)$ denote the space of linear operators on the Hilbert space X, and $|X|$ denote the dimension of the Hilbert space. We denote the identity operator in $\mathcal{L}(X)$ by $\mathbb{1}^X$, and the ideal map $\mathrm{id}\colon \mathcal{L}(X) \to \mathcal{L}(X)$ is denoted by id^X. For any linear operator $\Lambda\colon A \to B$ we will use the *trace norm* defined as

$$\|\Lambda\|_1 = \operatorname{Tr}\sqrt{\Lambda^\dagger\Lambda} = \operatorname{Tr}|\Lambda|, \tag{1}$$

and the *operator norm* defined as

$$\|\Lambda\|_{\mathrm{op}} = \sup\left\{\|\Lambda a\| :\ a \in A,\ \|a\| = 1\right\}. \tag{2}$$

For any superoperator $\mathcal{N}\colon \mathcal{L}(A) \to \mathcal{L}(B)$ the *induced trace norm* is defined as

$$\|\mathcal{N}\|_{1\to1} = \max\left\{\|\mathcal{N}(\rho)\|_1,\ \rho \in \mathcal{L}(A),\ \|\rho\|_1 = 1\right\}. \tag{3}$$

Furthermore, for any superoperator $\mathcal{N}\colon \mathcal{L}(A) \to \mathcal{L}(B)$ we will make use of the *diamond norm*, defined as

$$\|\mathcal{N}\|_\diamond = \|\mathrm{id}^A \otimes \mathcal{N}\|_{1\to1} = \max\left\{\|\mathrm{id}^A \otimes \mathcal{N}(\rho)\|_1,\ \rho \in \mathcal{L}(A\otimes A),\ \|\rho\|_1 = 1\right\}. \tag{4}$$

Note that the maximum in this definition is attained on a rank-one operator $\rho = |\psi\rangle\langle\varphi|$, with unit vectors $|\psi\rangle$ and $|\varphi\rangle$. If $\mathcal{N}$ is Hermitian-preserving, then the maximum is indeed attained on a pure state $\rho = |\psi\rangle\langle\psi|$. For any superoperator $\mathcal{N}\colon \mathcal{L}(A) \to \mathcal{L}(B)$ the diamond norm and the induced trace norm are related as follows:

$$\|\mathcal{N}\|_{1\to1} \le \|\mathcal{N}\|_\diamond \le \|\mathcal{N}\|_{1\to1}\min(|A|,|B|). \tag{5}$$

For a density operator α^A the *von Neumann entropy* is defined as

$$S(A)_\alpha = S(\alpha) := -\operatorname{Tr}\alpha\log\alpha. \tag{6}$$

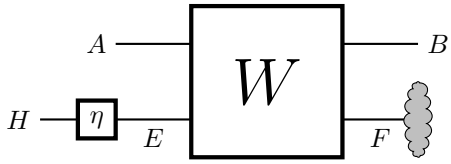

Fig. 1. Diagrammatic view of the three parties, Alice (A), Helen (H) and Bob (B), involved in the communication with a third party (H) controlling the environment input system whose aim is to enhance the communication between Alice and Bob. Since Helen has no further role to play after setting the initial environment state, its assistance is of passive nature. The inaccessible output-environment system is labeled by F.

For two density operators α and β, the *quantum relative entropy* of α with respect to β is defined as

$$D(\alpha \,\|\, \beta) := \operatorname{Tr} \alpha(\log \alpha - \log \beta). \tag{7}$$

Furthermore, for any density operator ρ^{AB} on a bipartite system, the *quantum mutual information* is defined as

$$I(A\!:\!B)_\rho := S(A)_\rho + S(B)_\rho - S(AB)_\rho. \tag{8}$$

We have three presumably inequivalent models of classical communication, depending on the role of the helper. In the first model under consideration, we assume that Helen sets the initial state of the environment to enhance the classical communication from Alice to Bob as is depicted in Fig. 1. Since Helen has no role in the protocol after setting the initial environment state, this model is thus referred to as *passive environment-assisted model*.

We assume that there are no quantum correlations between Alice's and Helen's inputs. Consider a unitary or, more generally, an isometry $W\colon A \otimes E \to B \otimes F$ which defines a channel (CPTP map) $\mathcal{N}\colon \mathcal{L}(A \otimes E) \to \mathcal{L}(B)$, whose action on the input state ρ on $A \otimes E$ is

$$\mathcal{N}^{AE\to B}(\rho) = \operatorname{Tr}_F W \rho W^\dagger. \tag{9}$$

Then an effective channel $\mathcal{N}_\eta\colon \mathcal{L}(A) \to \mathcal{L}(B)$ is established between Alice and Bob once the initial state η on E is set:

$$\mathcal{N}_\eta^{A\to B}(\rho) := \mathcal{N}^{AE\to B}(\rho \otimes \eta). \tag{10}$$

For a given CPTP map $\mathcal{N}\colon \mathcal{L}(A) \to \mathcal{L}(B)$, we can consider the Stinespring isometry $V\colon A \to B \otimes F$. This is a special case of the above model where the initial environment E is *one-dimensional*, i.e., Helen has no choice of an initial environment state. The classical capacity for the above case is given by the Holevo–Schumacher–Westmoreland theorem [8, 9] (cf. [2]),

$$C(\mathcal{N}) = \sup_n \max_{\{p_x, \rho_x^{A^n}\}} \frac{1}{n} I(X\!:\!B^n)_\sigma, \tag{11}$$

where the quantum mutual information is evaluated with respect to the state

$$\sigma = \sum_x p_x |x\rangle\langle x| \otimes \mathcal{N}^{\otimes n}(\rho_x^{A^n}). \tag{12}$$

The maximization is over the ensembles $\{p_x, \rho_x^{A^n}\}$ where the states $\rho_x^{A^n}$ are input across A^n. Here $\{|x\rangle\}$ are the orthonormal basis of the classical reference system X. It is known that the supremum over n (the "regularization") is necessary [10], except for some special channels [11].

When the encoding by the sender is restricted to separable states $\rho_x^{A^n}$, i.e., convex combinations of tensor products $\rho_x^{A^n} = \rho_{1x}^{A_1} \otimes \ldots \otimes \rho_{nx}^{A_n}$, the classical communication capacity admits a single-letter characterization, given by the so-called *Holevo information* of quantum channel,

$$\chi(\mathcal{N}) = \max_{\{p_x, \rho_x^A\}} I(X:B)_\sigma, \tag{13}$$

where the quantum mutual information is evaluated with respect to the state $\sigma := \sum_x p_x |x\rangle\langle x| \otimes \mathcal{N}(\rho_x^A)$.

The ensemble that achieves the maximum in (13), say $\{p_x^*, \varphi_x^*\}$, is called the *optimal ensemble*. Let $\varphi_{\text{avg}}^* = \sum_x p_x^* \varphi_x^*$ be the average of the optimal ensemble. From [12] we know the existence of such an ensemble that achieves the Holevo information and has the following relative entropic formulation:

$$\chi(\mathcal{N}) = \sum_x p_x^* D\big(\mathcal{N}(\varphi_x^*) \,\|\, \mathcal{N}(\varphi_{\text{avg}}^*)\big). \tag{14}$$

For any ρ^A, we have the inequality

$$\chi(\mathcal{N}) \geq D\big(\mathcal{N}(\rho^A) \,\|\, \mathcal{N}(\varphi_{\text{avg}}^*)\big), \tag{15}$$

with the equality holding for any member of the optimal ensemble. Note also that for a given channel $\mathcal{N}$, though we can have many optimal ensembles that achieve the Holevo information, the average of the optimal ensemble is unique [13]. An important class of channels which admit single-letter characterization of classical capacity, i.e., $C(\mathcal{N}) = \chi(\mathcal{N})$, are the entanglement-breaking channels [14]. Actually, for any two entanglement-breaking channels, $\mathcal{N}_1$ and $\mathcal{N}_2$, we have the following additivity property:

$$\chi(\mathcal{N}_1 \otimes \mathcal{N}_2) = \chi(\mathcal{N}_1) + \chi(\mathcal{N}_2). \tag{16}$$

A variant of the passive environment-assisted model where Helen has pre-shared entanglement with Bob, thus referred to as *entanglement-environment-assisted model*, can also be considered. In such a case we can extend the notation of $\mathcal{N}_\eta = \mathcal{N}(\,\cdot \otimes \eta)$ and let, for a state $\varkappa$ on EK,

$$\mathcal{N}_\varkappa^{A\to BK}(\rho) := \big(\mathcal{N}^{AE\to B} \otimes \mathrm{id}^K\big)\big(\rho^A \otimes \varkappa^{EK}\big). \tag{17}$$

A new model, called *conferencing helper*, is introduced, where we contemplate the possibility of local operations and classical communication (LOCC) between Alice and Helen, thus allowing Helen to play an active role in the encoding process (in contrast to the previous models discussed above); see Fig. 2. Of course, the possibility that Alice and Helen share entanglement has to be excluded; otherwise, we recover the situation of a quantum channel determined by Alice and Helen input system and Bob output obtained by tracing away a part of the system. For a given unitary or, more generally, isometry $W\colon A \otimes E \to B \otimes F$, which defines the channel $\mathcal{N}\colon \mathcal{L}(A \otimes E) \to \mathcal{L}(B)$, whose action on a input state is

$$\mathcal{N}^{AE\to B}(\rho_i \otimes \eta_i) = \mathrm{Tr}_F\, W(\rho_i \otimes \eta_i) W^\dagger, \tag{18}$$

the state η_i can be adjusted according to the classical message with the proviso that the global input state of the systems A and E is separable.

Furthermore, we would mention that throughout the paper log stands for the logarithm to the base 2 and ln is the natural logarithm. The binary entropy is denoted by

$$H_2(p) = -p\log(p) - (1-p)\log(1-p). \tag{19}$$

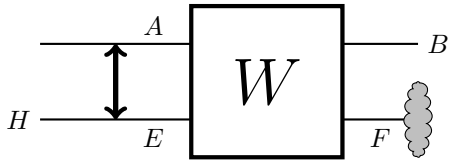

Fig. 2. Diagrammatic view of the three parties, Alice (A), Helen (H) and Bob (B), involved in the communication, with the party (H) controlling the environment input system and the sender (A) that can freely communicate classically. The inaccessible output-environment system is labeled by F.

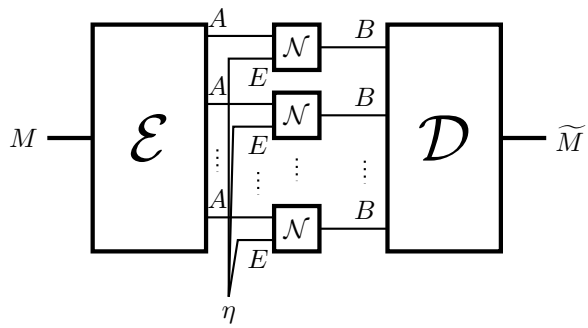

Fig. 3. Schematic of a general protocol to transmit classical information with passive assistance from the environment; $\mathcal{E}$ and $\mathcal{D}$ are the encoding and decoding maps, respectively. The initial state of the environment is η.

The cyclic shift operator $X(x)$ and the phase operator $Z(z)$ acting on the computational basis $\{|j\rangle\}_{0,1,2,\ldots,d-1}$ of a d-dimensional Hilbert space are defined in the following way:

$$\begin{aligned} X(x)|j\rangle &= |(x+j) \bmod d\rangle, \quad x = 0,1,2,\ldots,d-1, \\ Z(z)|j\rangle &= \omega^{zj}|j\rangle, \quad z = 0,1,2,\ldots,d-1. \end{aligned} \tag{20}$$

Here the complex number $\omega = \exp\left(\frac{2\pi i}{d}\right)$ is a primitive dth root of unity and i denotes the imaginary unit. Given the operators in (20), for each pair (x,z) we can identify the *discrete Weyl operator* $W(x,z) \in \mathrm{U}(d)$ defined as

$$W(x,z) := X(x)Z(z). \tag{21}$$

3. PASSIVE ENVIRONMENT-ASSISTED CAPACITIES

In this section we define the *passive environment-assisted* model rigorously and provide different notions of assisted codes, depending on the capabilities of Helen (whether she can input arbitrary entangled states across different instances of isometry or whether she is restricted to separable states). Furthermore, a capacity when Helen is restricted to separable states and Alice's encoding is restricted to product states across n instances of the channel is also considered.

3.1. Model for Transmission of Classical Information

By referring to Fig. 3, let Alice select some classical message m from the set of messages $\{1,2,\ldots,|M|\}$ to communicate to Bob. Let M denote the random variable corresponding to Alice's choice of message, and M corresponds to the associated Hilbert space with the orthonormal basis $\{|m\rangle\}$. An encoding CPTP map $\mathcal{E}\colon M \to \mathcal{L}(A^n)$ can be realized by preparing states $\{\alpha_m\}$ to be input across A^n of n instances of the channel. A decoding CPTP map $\mathcal{D}\colon \mathcal{L}(B^n) \to \widetilde{M}$ can be

realized by a positive operator-valued measure (POVM) $\{\Lambda_m\}$. Here $\widetilde{M}$ is the Hilbert space associated to the random variable $\widetilde{\mathrm{M}}$ for Bob's estimate of the message sent by Alice. The probability of error for a particular message m is

$$P_e(m) = 1 - \mathrm{Tr}\big[\Lambda_m \mathcal{N}^{\otimes n}(\alpha_m^{A^n} \otimes \eta^{E^n})\big]. \tag{22}$$

Definition 1. A *passive environment-assisted classical code* of block length n is a family of triples $\{\alpha_m^{A^n}, \eta^{E^n}, \Lambda_m\}$ with the error probability $\overline{P}_e := \frac{1}{|M|}\sum_m P_e(m)$ and the rate $\frac{1}{n}\log|M|$. A rate R is achievable if there is a sequence of codes over their block length n with $\overline{P}_e$ converging to 0 and rate converging to R. The passive environment-assisted classical capacity of W, denoted by $C_H(W)$ or, equivalently, $C_H(\mathcal{N})$, is the maximum achievable rate. If the helper is restricted to fully separable states η^{E^n}, i.e., convex combinations of tensor products $\eta^{E^n} = \eta_1^{E_1} \otimes \ldots \otimes \eta_n^{E_n}$, the largest achievable rate is denoted by $C_{H\otimes}(W) = C_{H\otimes}(\mathcal{N})$.

Since the error probability is linear in the environment state η, without loss of generality η may be assumed to be pure, for both unrestricted and separable helper. We shall assume this from now on, without necessarily specifying it each time.

Theorem 1. *For an isometry $W\colon AE \to BF$, the passive environment-assisted classical capacity is given by*

$$C_H(W) = \sup_n \max_{\eta^{(n)}} \frac{1}{n} C(\mathcal{N}_{\eta^{(n)}}^{\otimes n}) = \sup_n \max_{\{p(x),\alpha_x^{A^n}\},\eta^{E^n}} \frac{1}{n} I(X:B^n)_\sigma, \tag{23}$$

where the mutual information is evaluated with respect to the state

$$\sigma = \sum_x p(x)|x\rangle\langle x| \otimes \mathcal{N}_{\eta^{E^n}}^{\otimes n}(\alpha_x^{A^n})$$

and the maximization is over the ensemble $\{p(x), \alpha_x^{A^n}\}$ and pure environment input states $\eta^{(n)}$ on E^n.

Similarly, the capacity with separable helper is given by the formula

$$\begin{aligned} C_{H\otimes}(W) &= \sup_n \max_{\eta^{(n)}=\eta_1\otimes\ldots\otimes\eta_n} \frac{1}{n} C(\mathcal{N}_{\eta_1} \otimes \ldots \otimes \mathcal{N}_{\eta_n}) \\ &= \sup_n \max_{\{p(x),\alpha_x^{A^n}\},\eta^{E^n}} \frac{1}{n} I(X:B^n)_\sigma, \end{aligned} \tag{24}$$

where the maximum is only over (pure) product states, i.e., $\eta^{(n)} = \eta_1 \otimes \ldots \otimes \eta_n$.

As a consequence, $C_H(W) = \lim_{n\to\infty} \frac{1}{n} C_{H\otimes}(W^{\otimes n})$.

Proof. The direct part (the "$\geq$" inequality), follows immediately from the HSW theorem [8,9] applied to the channel $(\mathcal{N}^{\otimes n})_{\eta^{(n)}}$; to be precise, to asymptotically many copies of this block channel, so that the i.i.d. arguments hold true (cf. [2]).

For the converse part (the "$\leq$" inequality), consider a code of block length n with error probability $\overline{P}_e$. The state after an encoding operation and action of the channel is given by

$$\Phi^{MB^n} = \frac{1}{|M|}\sum_m p(m)|m\rangle\langle m| \otimes \mathcal{N}^{\otimes n}(\alpha_m^{A^n} \otimes \eta^{E^n}),$$

and the state after a decoding operation is given by

$$\omega^{M\widetilde{M}} = \mathbb{1}^M \otimes \mathcal{D}(\Phi^{MB^n}).$$

Then we have

$$\begin{aligned} nR &= H(M)_\omega \\ &= I(M:\widetilde{M})_\omega + H(M|\widetilde{M})_\omega \\ &\le I(M:\widetilde{M})_\omega + H(\overline{P}_e) + nR\overline{P}_e \\ &\le I(M:B^n)_\Phi + n\varepsilon. \end{aligned}$$

The first inequality follows from the application of Fano's inequality, and the second one follows from the data-processing inequality, where $\varepsilon = \dfrac{1}{n} + R\overline{P}_e$. Setting $M = X$, we have

$$R \le \frac{1}{n} I(X:B^n) + \varepsilon.$$

As $n \to \infty$ and $\overline{P}_e \to 0$, the upper bound on the rate follows, depending on C_H or $C_{H\otimes}$, without or with restrictions on $\eta^{(n)}$. $\triangle$

Remark 1. The channel whose inputs are Alice and Helen and outputs Bob can be viewed as a quantum version of a multiple access channels (MAC) with two senders and one receiver, which was studied in [15, 16]. In such a model both Alice and Helen try to communicate their individual independent messages to Bob. The rates R_A and R_H at which Alice and Helen can respectively communicate with Bob give the capacity region (R_A, R_H). If this capacity region is known, then the passive environment-assisted capacity is given by $\max\{R : (R, 0) \in \text{capacity region}\}$. Whenever a single-letter characterization for a MAC is available, this might be helpful in the evaluation of environment-assisted capacities, but in general, when the regularization is required, this view may not help.

For a separable helper, and when in addition Alice's encoding is restricted to product input states, i.e.,

$$\alpha_m = \alpha_{1m} \otimes \alpha_{2m} \ldots \otimes \alpha_{nm},$$

across A^n, we have the following:

$$\max_{\{p(x),\alpha_x^{A^n}\},\, \eta_1\otimes\ldots\otimes\eta_n} I(X:B^n)_\sigma = \sum_i \max_{\{p(x_i),\alpha_{x_i}^A\},\, \eta_i} I(X_i:B)_{\sigma_i}.$$

Then, from (24), we have the product-state capacity with separable helper given by

$$\chi_{H\otimes}(W) = \max_{\{p(x),\rho_x\},\, \eta^E} I(X:B)_\sigma, \tag{25}$$

where the mutual information is evaluated with respect to the state $\sigma := \sum_x p(x)|x\rangle\langle x| \otimes \mathcal{N}_\eta(\rho_x)$ and the maximization is over the ensemble $\{p(x), \rho_x\}$ and the state η^E.

For any η^E and for all ρ^A we have

$$D\big(\mathcal{N}_\eta{}^{A\to B}(\rho^A) \,\|\, \omega_\eta^B\big) \le \chi_{H\otimes}(W), \tag{26}$$

where $\omega_\eta^B := \mathcal{N}_\eta^{A\to B}(\rho^A_{\mathrm{avg},\eta})$ and $\rho^A_{\mathrm{avg},\eta}$ is the average of the optimal ensemble that achieves the Holevo information for $\mathcal{N}_\eta$.

We can further contemplate the scenario where the roles of A and E are exchanged, i.e., Alice tries to set an initial state in A, thereby establishing a channel between E and B. Then, the quantities of interest are $\chi^A_{H\otimes}$ and $\chi^H_{A\otimes}$, the product-state capacities with a separable helper when

the sender is Alice and Helen, respectively. They are given by the formulas[4]

$$\begin{aligned}\chi^A_{H\otimes}(W) &= \max_{\{p(x),\alpha^A_x\},\eta^E} I(X:B)_\sigma,\\ \chi^H_{A\otimes}(W) &= \max_{\{p(x),\eta^E_x\},\alpha^A} I(X:B)_\mu,\end{aligned} \tag{27}$$

where the mutual information for the former case is evaluated with respect to the state $\sigma := \sum_x p(x)|x\rangle\langle x| \otimes \mathcal{N}_\eta(\alpha_x)$ and the maximization is over the ensemble $\{p(x), \alpha_x\}$ and the state η^E. In the latter case, when Helen is the sender, the mutual information is evaluated with respect to the state $\mu := \sum_x p(x)|x\rangle\langle x| \otimes \mathcal{M}_\alpha(\eta_x)$ and the maximization is over the ensemble $\{p(x), \eta_x\}$ and the state α^A. Here, the effective channels $\mathcal{N}_\eta\colon \mathcal{L}(A) \to \mathcal{L}(B)$ and $\mathcal{M}_\alpha\colon \mathcal{L}(E) \to \mathcal{L}(B)$ are, respectively,

$$\mathcal{N}_\eta(\rho) = \mathrm{Tr}_F\big(W(\rho\otimes\eta)W^\dagger\big), \qquad \mathcal{M}_\alpha(\nu) = \mathrm{Tr}_F\big(W(\alpha\otimes\nu)W^\dagger\big).$$

From (25) we see that

$$\begin{aligned}\chi^A_{H\otimes}(W) &\le \log|B| - S_{\min}(W),\\ \chi^H_{A\otimes}(W) &\le \log|B| - S_{\min}(W),\end{aligned} \tag{28}$$

where

$$S_{\min}(W) := \min_{\alpha^A,\eta^E} S\big(\mathrm{Tr}_F\big(W(\alpha^A\otimes\eta^E)W^\dagger\big)\big) \tag{29}$$

is the minimum output entropy of the given unitary W.

Lemma 1 [17] (continuity of the Stinespring dilation). *For any two quantum channels $\mathcal{N}_1, \mathcal{N}_2\colon \mathcal{L}(A) \to \mathcal{L}(B)$ with Stinespring dilations $V_1, V_2\colon A \to B\otimes F$, we have*

$$\inf_U \big\|(\mathbb{1}^B\otimes U^F)V_1 - V_2\big\|^2_{\mathrm{op}} \le \|\mathcal{N}_1 - \mathcal{N}_2\|_\diamond \le 2\inf_U \big\|(\mathbb{1}^B\otimes U^F)V_1 - V_2\big\|_{\mathrm{op}}, \tag{30}$$

where the infimum is over the unitaries $U\colon F \to F$.

Theorem 2. *For any unitary $W\colon A\otimes E \to B\otimes F$, with $|A| = |E| = |B| = |F| = d$, we have*

$$\chi^A_{H\otimes} + \chi^H_{A\otimes} \ge \frac{1}{2^{13}d^2\ln 2}\left(\frac{\sqrt{2+2(\log d)^2} - \sqrt{2}}{\log d}\right)^8. \tag{31}$$

This is a kind of an uncertainty relation for $\chi^A_{H\otimes}$ and $\chi^H_{A\otimes}$, saying that not both of them can be arbitrarily small.

Proof. Let $\mathrm{SWAP}\colon A\otimes E \to B\otimes F$ be the *swap operator*, defined by $\mathrm{SWAP}(|\psi\rangle^A\otimes|\varphi\rangle^E) := |\varphi\rangle^B\otimes|\psi\rangle^F$. Let us define a quantum channel which has SWAP as its dilation as follows:

$$\mathcal{M}^{A\to B}_\sigma(\rho^A) := \mathrm{Tr}_F\big(\mathrm{SWAP}(\rho\otimes\sigma)\,\mathrm{SWAP}^\dagger\big) = \sigma^B\,\mathrm{Tr}(\rho^A).$$

Assume that $\chi^A_{H\otimes} = \varepsilon$. Then, from (26), for all η^E on E and ρ^A on A,

$$D\big({\mathcal{N}_\eta}^{A\to B}(\rho^A)\,\|\,\mathcal{M}^{A\to B}_{\omega_\eta}(\rho^A)\big) \le \varepsilon,$$

where $\omega^B_\eta := {\mathcal{N}_\eta}^{A\to B}(\rho^A_{\mathrm{avg},\eta})$ and $\rho^A_{\mathrm{avg},\eta}$ is the average of the optimal ensemble that achieves the Holevo information for $\mathcal{N}_\eta$. From the quantum Pinsker inequality [18], for all ρ^A on A

$$\big\|{\mathcal{N}_\eta}^{A\to B}(\rho^A) - \mathcal{M}^{A\to B}_{\omega_\eta}(\rho^A)\big\|_1 \le \sqrt{2\varepsilon\ln 2}.$$

[4] When we omit the superscripts, it is implicitly understood that Alice is the sender and Helen sets an initial state in E.

From (3), we have

$$\|\mathcal{N}_\eta{}^{A\to B} - \mathcal{M}_{\omega_\eta}^{A\to B}\|_{1\to 1} \le \sqrt{2\varepsilon \ln 2}.$$

Using the relation between the induced trace norm and the diamond norm as expressed by (5) gives

$$\|\mathcal{N}_\eta{}^{A\to B} - \mathcal{M}_{\omega_\eta}^{A\to B}\|_\diamond \le \sqrt{2\varepsilon \ln 2}\cdot d.$$

From the left-hand side of the continuity bound in Lemma 1, we have

$$\|(\mathbb{1}^B \otimes U^F)W - \mathrm{SWAP}\|_{\mathrm{op}} \le (2\varepsilon \ln 2)^{\frac{1}{4}}\sqrt{d},$$

where the minimum is achieved by U^F. The channels $\mathcal{P}_\alpha^{E\to B}(\eta^E) := \mathrm{Tr}_F(W(\alpha\otimes\eta)W^\dagger)$, and $\mathrm{id}^{E\to B}$ have the dilations W and SWAP, respectively. From the right-hand side of the continuity bound, we get

$$\left\|\mathcal{P}_\alpha^{E\to B} - \mathrm{id}^{E\to B}\right\|_\diamond \le 2(2\varepsilon \ln 2)^{\frac{1}{4}}\sqrt{d} =: \Delta, \quad \forall \alpha^A.$$

Thus, from the continuity of χ [19, 20] (see (51))

$$|\chi(\mathcal{P}_\alpha) - \chi(\mathrm{id})| \le 2\Delta \log d + (2+\Delta)H_2\Big(\frac{\Delta}{2+\Delta}\Big), \tag{32}$$

which gives us

$$\chi(\mathcal{P}_\alpha) \ge \max\Big\{(1-2\Delta)\log d - (2+\Delta)H_2\Big(\frac{\Delta}{2+\Delta}\Big), 0\Big\}.$$

Using

$$H_2(\Delta) \le 2\sqrt{\Delta(1-\Delta)},$$

and $\chi_{A\otimes}^H \ge \chi(\mathcal{P}_\alpha)$ (see (25)), we have

$$\chi_{A\otimes}^H \ge \max\{\log d - 2\Delta \log d - \sqrt{8\Delta}, 0\}.$$

Then, rewriting the above inequality in terms of ε, we obtain

$$\chi_{A\otimes}^H \ge \max\{f(\varepsilon), 0\}, \tag{33}$$

where

$$f(\varepsilon) := \log d - \big(2^9 d^2 \varepsilon \ln 2\big)^{\frac{1}{4}} \log d - \big(2^{17} d^2 \varepsilon \ln 2\big)^{\frac{1}{8}}.$$

The function $f(\varepsilon)$ is nonnegative for $\varepsilon \in [0, \varepsilon_0]$ with

$$\varepsilon_0 := \frac{1}{2^{13} d^2 \ln 2}\Big(\frac{\sqrt{2+2(\log d)^2} - \sqrt{2}}{\log d}\Big)^8.$$

Putting everything together for $\varepsilon \in [0, \varepsilon_0]$, we arrive at

$$\chi_{H\otimes}^A + \chi_{A\otimes}^H \ge \min_{\varepsilon>0}\{\varepsilon + f(\varepsilon)\}.$$

Since the function $\varepsilon + f(\varepsilon)$ is monotonically decreasing in the interval $\varepsilon \in (0, \varepsilon_0]$, the minimum is attained at ε_0, and thus we obtain

$$\chi_{H\otimes}^A + \chi_{A\otimes}^H \ge \frac{1}{2^{13} d^2 \ln 2}\Big(\frac{\sqrt{2+2(\log d)^2} - \sqrt{2}}{\log d}\Big)^8, \tag{34}$$

which is nonzero. $\triangle$

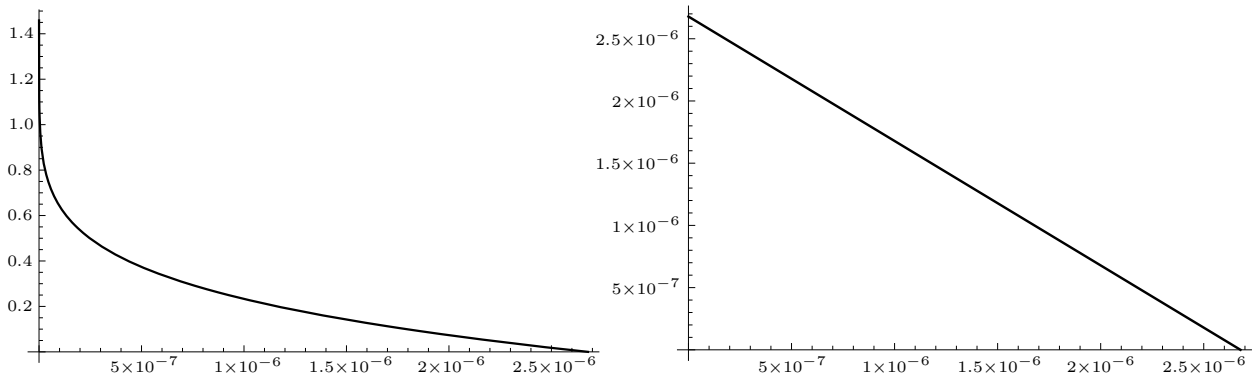

Fig. 4. Plots of $\chi^H_{A\otimes}$ versus $\chi^A_{H\otimes}$ for dimension $d = 3$. The left plot corresponds to (33), and the right one to (34).

Remark 2. We present the uncertainty relation in the form of (34), motivated by the well-known entropic uncertainty relation [21]. Actually we get a tighter lower bound on $\chi^H_{A\otimes}$ as a function of $\chi^A_{H\otimes} := \varepsilon$ from (33), as can be seen in Fig. 4.

Lemma 2 [22]. *Let U be a random gate in* $\mathrm{U}(d^2)$ *according to the Haar measure, then for any* $\delta > 0$,

$$\Pr\{|S_{\min}(U) - \mathbf{E}(S_{\min}(U))| \geq \delta\} \leq 2\exp\left(-\frac{d^2\delta^2}{64(\log d)^2}\right),$$

where $\mathbf{E}(S_{\min}(U))$ *is the expectation value of the minimum output entropy. Here* [22, Corollary 44],

$$\mathbf{E}(S_{\min}(U)) \geq \log d - \frac{1}{\ln 2} - 1.$$

Remark 3. When the U are chosen according to the Haar measure on $\mathrm{U}(d^2)$, from Lemma 2 and (28) we can give an upper bound on $\mathbf{E}(\chi_{H\otimes}(U))$, the expectation value of the product-state capacity with separable helper, which reads as follows:

$$\mathbf{E}(\chi^A_{H\otimes}(U)) \leq 1 + \frac{1}{\ln 2}, \qquad \mathbf{E}(\chi^H_{A\otimes}(U)) \leq 1 + \frac{1}{\ln 2}.$$

It follows that when $d \to \infty$, by the concentration of measure phenomenon [23], with overwhelming probability

$$\chi^A_{H\otimes}, \chi^H_{A\otimes} < 2.5.$$

Remark 4. The classical capacity of a quantum channel is zero if and only if the channel maps all inputs to a constant output, i.e., the output of the channel is independent of the input. This helps us to identify the unitaries which have $C_{H\otimes} = 0$. These unitaries must have effective channels with constant output for every choice of the initial environment state. At least in the case where $|A| = |F|$ and $|B| = |E|$, the unitary is the SWAP. Furthermore, for these unitaries, $C_H(\mathrm{SWAP}) = C_{H\otimes}(\mathrm{SWAP}) = 0$.

3.2. Controlled-Unitaries

As we have noted, the above-defined passive assisted capacities, like the standard classical capacity of a quantum channel [10] (cf. [2]), admit multi-letter characterizations, thus posing a hard optimization problem. It is therefore important to single out classes of unitaries, if any, for which we can reduce to the single-letter case. We focus on controlled-unitaries, which, apart from allowing for a simple characterization of capacities, provide examples for interesting phenomena like superadditivity.

Definition 2. We say that a unitary operator U is *universally entanglement-breaking* (respectively, *universally classical-quantum*) if for every $|\eta\rangle \in E$, the effective channel $\mathcal{N}_\eta\colon \mathcal{L}(A) \to \mathcal{L}(B)$ is entanglement-breaking (respectively, classical-quantum). The set of universally entanglement-breaking (respectively, universally classical-quantum) unitaries is denoted by $\mathfrak{E}$ (respectively, $\mathfrak{CQ}$).

For these unitaries, $C_{H\otimes}$ reduces to the single-letter case. Indeed, for any $W \in \mathfrak{E}$, we have

$$C_{H\otimes}(W) = \max_{\{p(x),\rho_x\},\, \eta^E} I(X:B)_\sigma,$$

where the mutual information is evaluated with respect to the state $\sigma := \sum_x p(x)|x\rangle\langle x| \otimes \mathcal{N}_\eta(\rho_x)$ and the maximization is over the ensemble $\{p(x), \rho_x\}$. This follows from additivity of the Holevo information for entanglement-breaking channels,

$$\max_{\{p(x^n),\rho_{x^n}^{A^n}\},\, \eta^{E^n}} I(X^n:B^n)_\sigma = \sum_i \max_{\{p(x),\rho_x^A\},\, \eta_i} I(X:B)_{\sigma_i}$$

Let us define a unitary operator $U_c\colon A \otimes E \to B \otimes F$ with $|A| = |B| = |E| = |F| = d$ as follows:

$$U_c := \sum_i |i\rangle^F \langle i|^A \otimes U_i^{E\to B}. \tag{35}$$

Here $\{|i\rangle\}$ denotes an orthonormal basis of A and $U_i \in \mathrm{U}(d)$. When the initial environment state is $|\eta\rangle$, the Kraus operators of the effective channel $\mathcal{N}_\eta\colon \mathcal{L}(A) \to \mathcal{L}(B)$ are given by $K_i = U_i|\eta\rangle\langle i|$. Thus, $\mathcal{N}_\eta$ is a classical-quantum channel for each choice of $|\eta\rangle$, and consequently $U_c \in \mathfrak{CQ}$. Hence,

$$C_{H\otimes}(U_c) = \max_{p_i,\eta} S\Big(\sum_i p_i U_i|\eta\rangle\langle\eta|U_i^\dagger\Big). \tag{36}$$

For $U_c \otimes V_c\colon A'E'AE \to B'F'BF$, the Kraus operators of the effective channel $\mathcal{N}_\eta\colon \mathcal{L}(A \otimes A') \to \mathcal{L}(B \otimes B')$, when the initial state of the environments $E'E$ is $|\eta\rangle$, are $K_{ij} = (U_i \otimes V_j)|\eta\rangle\langle ij|$, which is also a classical-quantum channel. Hence, $U_c \otimes V_c \in \mathfrak{CQ}$. We can also say that

$$C_{H\otimes}(U_c \otimes V_c) = \max_{p_{ij},\eta} S\Big(\sum_{i,j} p_{ij}(U_i \otimes V_j)|\eta\rangle\langle\eta|(U_i \otimes V_j)^\dagger\Big). \tag{37}$$

Remark 5. Universal properties of bipartite unitary operators have been studied in [24], although with a different motivation than in this manuscript. Since we are interested in evaluating environment-assisted capacities, we restrict the universal properties to pure environment states. We can extend the universal properties to a general density operators in the case of entanglement-breaking and classical-quantum channels because of the convexity of these sets of maps. In particular, they treat in full generality the question of bipartite unitaries which give constant channels for all input-environment states (see [24, Theorem 2.4 and Remark 2.5]; cf. Remark 4, in which we restricted to the case where $|A| = |F|$ and $|B| = |E|$).

Remark 6. We have identified unitaries with $C_H = 0$. It is much harder to characterize unitaries with quantum capacity $Q_H = 0$ [1]. SWAP was the only unitary which was identified to have $Q_H = 0$. It was also conjectured there that $\sqrt{\mathrm{SWAP}}$ has zero passive environment-assisted capacity. From the previous discussions, U_c has zero *passive environment-assisted* quantum capacity. When an arbitrary initial environment state $|\eta\rangle^{(n)}$ is input across E^n, the effective channel $\mathcal{N}_{\eta^{(n)}}\colon A^n \to B^n$ is classical-quantum; thus, the quantum capacity of the effective channel is $Q(\mathcal{N}_{\eta^{(n)}}) = 0$. As a consequence of coding theorems for transmission of quantum information with a passive separable helper ($Q_{H\otimes}$) and passive helper (Q_H), they are related by $Q_H(U_c) = \lim_{n\to\infty} \frac{1}{n} Q_{H\otimes}(U_c^{\otimes n})$. Thus, $Q_H(U_c) = 0$.

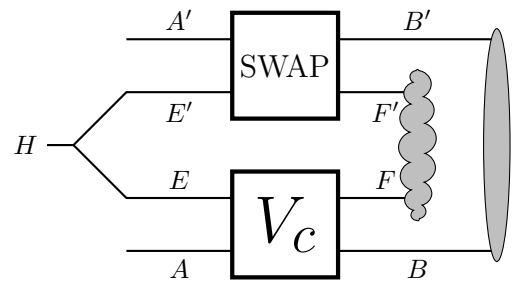

Fig. 5. Inputs controlled by Alice are A' and A. Helen controls E' and E, Bob's systems are labeled as B' and B. The inaccessible output-environment systems are labeled by F' and F. Helen inputs an entangled state in $E'E$.

Two-qubit unitaries. Here we evaluate the passive environment-assisted classical capacity with separable helper for universally classical-quantum two-qubit unitaries. A two-qubit controlled-unitary is of the form $U_{c(2)} := \sum_{i=0}^{1} |i\rangle^F\langle i|^A \otimes U_i^{E\to B}$, where $U_i \in \mathrm{SU}(2)$. Then, according to the parametrization reported in the Appendix, the $U_{c(2)}$ has a parametric representation $\left(\frac{\pi}{2}, \frac{\pi}{2}, u\right)$ where $0 \le u \le \frac{\pi}{2}$. From the previous discussion we know that $U_{c(2)} \in \mathfrak{CU}$ when $0 \le u \le \frac{\pi}{2}$. Let the initial state of the environment be $|\psi\rangle^E = c_0|0\rangle + c_1|1\rangle$ with $c_0, c_1 \in \mathbb{C}$ and $|c_0|^2 + |c_1|^2 = 1$. The input ensemble is $|0\rangle, |1\rangle$ with probability $1-q, q$, respectively; then the capacity is given by

$$C_{H\otimes}(U_{c(2)}) = \max_{|c_0|^2, q} H_2\Big(\frac{1}{2} + \sqrt{\frac{1}{4} - 4|c_0|^2|c_1|^2 q(1-q)\cos^2 u}\Big). \tag{38}$$

The maximum is attained at $|c_0|^2 = q = \frac{1}{2}$, and it is

$$C_{H\otimes}(U_{c(2)}) = H_2\Big(\frac{1+\sin u}{2}\Big). \tag{39}$$

3.3. Superadditivity

In this subsection, we find two unitaries such that when they are used in conjunction, and their initial environments are entangled, they transmit more classical information than the sum of the classical information transferred by them individually. This phenomenon is called *superadditivity*.

The following examples use the setting and notation of Fig. 5.

1. Superadditivity for $C_{H\otimes}$.

Let $V_c = \sum_{i=0}^{2} |i\rangle^F\langle i|^A \otimes V_i^{E\to B}$ act on 2-qutrit systems, and let V_i be given as

$$V_0 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad V_1 := \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad V_2 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We can use (36) to evaluate $C_{H\otimes}(V_c)$, namely

$$C_{H\otimes}(V_c) = \max_{p_i, \eta} S\Big(\sum_i p_i V_i|\eta\rangle\langle\eta|V_i^\dagger\Big),$$

where the maximization is over the initial state of the environment $|\eta\rangle^E = a|0\rangle + b|1\rangle + c|2\rangle$ with $|a|^2 + |b|^2 + |c|^2 = 1$ and $a, b, c \in \mathbb{C}$. Let $|\psi_i\rangle = V_i|\eta\rangle$. There are no $a, b, c \in \mathbb{C}$ such that the $|\psi_i\rangle$ are mutually orthogonal, and thus $C_{H\otimes}(V_c) < \log 3$.

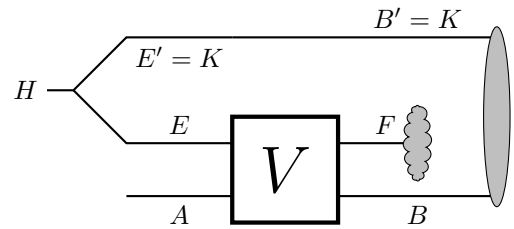

Fig. 6. In Fig. 5, when Helen inputs an entangled state across $E'E$ and an arbitrary state in A', the SWAP acts like a "dummy" channel but helps to establish entanglement between the receiver BB' and the environment E. This is equivalent to sharing an entangled state between Helen and Bob.

Consider the scenario in Fig. 5, where Helen inputs a state $|\Phi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ across EE'. In such a scenario the effective channel is $\mathcal{N}_\Phi\colon AA' \to BB'$, and when we input $\{|00\rangle, |01\rangle, |02\rangle\}$ in $A'A$, the outputs in $B'B$ result, respectively, in

$$(\mathbb{1} \otimes V_0)|\Phi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \qquad (\mathbb{1} \otimes V_1)|\Phi\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}}, \qquad (\mathbb{1} \otimes V_2)|\Phi\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}},$$

which are orthogonal, thus making the classical capacity of the effective channel equal to $\log 3$. Therefore,

$$C_{H\otimes}(\mathrm{SWAP} \otimes V_c) > C_{H\otimes}(\mathrm{SWAP}) + C_{H\otimes}(V_c),$$

since SWAP has zero passive environment-assisted capacities (see Remark 4).

2. Superadditivity for C_H.

Let us consider $V_c\colon AE \to BF$ with $|A| = |F| = d^2$, $|E| = |B| = d$, given by

$$V_c = \sum_{x,z} |xz\rangle^F \langle xz|^A \otimes (W(x,z))^{E\to B}.$$

Here $W(x,z)$ are the discrete Weyl operators. From (36) we have $C_{H\otimes}(V_c) = \log d$. This is also the capacity with an unrestricted Helen, since it saturates the dimension of B. Thus, $C_H(V_c) = \log d$. Now consider $\mathrm{SWAP} \otimes V_c$ where $\mathrm{SWAP}\colon A'E' \to B'F'$ with $|A'| = |B'| = |E'| = |F'| = d$. When Helen inputs a maximally entangled state $|\Phi\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |ii\rangle$ across $E'E$, Alice inputs $\{|0ij\rangle\}_{i,j=0}^{d-1}$ in $A'A$ (note that $|A| = d^2$), the outputs in $B'B$ are the set of states $\{\mathbb{1} \otimes X(i)Z(j)|\Phi\rangle\}$, which are d^2 orthonormal maximally entangled states in $B'B$. Thus, $C_H(\mathrm{SWAP} \otimes V_c) = 2\log d$ is achieved by the above inputs when they are chosen with equal probability of $\frac{1}{d^2}$. In conclusion,

$$C_H(\mathrm{SWAP} \otimes V_c)) > C_H(V_c) + C_H(\mathrm{SWAP}).$$

4. ENTANGLEMENT-ENVIRONMENT-ASSISTED CAPACITY

As we have noted in Section 4, SWAP, in spite of having no communication capabilities with passive environment assistance on its own, can indeed enhance the classical communication when used in conjunction with other specific unitaries. In other words, SWAP acts like a "dummy" channel but helps to establish entanglement between the receiver and the initial environment, as is shown in Fig. 6. This is equivalent to sharing an entangled state between Helen and Bob, which motivates us to rigorously define the following model of communication.

By referring to Fig. 7, an encoding CPTP map $\mathcal{E}\colon M \to \mathcal{L}(A^n)$ can be realized by preparing states $\{\alpha_m\}$ to be input across A^n of n instances of the channel. Let M denote the random

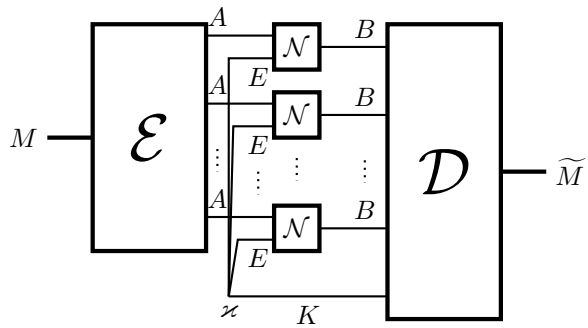

Fig. 7. The general form of a protocol to transmit classical information when the helper and receiver pre-share entanglement; $\mathcal{E}$ and $\mathcal{D}$ are the encoding and decoding maps respectively, and $\varkappa$ is the initial state of the environments and system K.

variable corresponding to Alice's choice of message and M be the associated Hilbert space with the orthonormal basis $\{|m\rangle\}$. A decoding CPTP map $\mathcal{D}\colon \mathcal{L}(B^n \otimes K) \to \widetilde{M}$ can be realized by a POVM $\{\Lambda_m\}$. Here $\widetilde{M}$ is Hilbert space associated to the random variable $\widetilde{\mathrm{M}}$ for Bob's estimate of the message sent by Alice. The probability of error for a particular message m is

$$P_e(m) = 1 - \mathrm{Tr}\big((\Lambda_m)(\mathcal{N}^{\otimes n} \otimes \mathrm{id}_H)(\alpha_m^{A^n} \otimes \varkappa^{E^n K})\big).$$

Definition 3. An *entanglement-environment-assisted classical code* of block length n is a family of triples $\{\alpha_m^{A^n}, \varkappa^{E^n K}, \Lambda_m\}$ with error probability $\overline{P}_e := \frac{1}{|M|}\sum_m P_e(m)$ and rate $\frac{1}{n}\log|M|$. A rate R is achievable if there is a sequence of codes over their block length n with $\overline{P}_e$ converging to 0 and rate converging to R. The entanglement-assisted environment classical capacity of W, denoted by $C_{EH}(W)$ or, equivalently, $C_{EH}(\mathcal{N})$, is the maximum achievable rate.

Theorem 3. *For an isometry $W\colon AE \to BF$, the entanglement-environment-assisted classical capacity is given by*

$$C_{EH}(W) = \sup_n \max_{\varkappa^{(n)}} \frac{1}{n} C(\mathcal{N}^{\otimes n}_{\varkappa^{(n)}}) = \sup_n \max_{\{p(x),\alpha_x^{A^n}\},\varkappa^{E^n K}} \frac{1}{n} I(X : B^n K)_\sigma, \tag{40}$$

where the mutual information is evaluated with respect to the state

$$\sigma = \sum_x p(x)|x\rangle\langle x| \otimes \mathcal{N}^{\otimes n}_{\varkappa^{E^n K}}(\alpha_x^{A^n})$$

and the maximization is over the ensemble $\{p(x), \alpha_x^{A^n}\}$ and over pure environment input states $\varkappa^{(n)}$ on $E^n K$.

Proof. The direct part (the "$\geq$" inequality) follows immediately from the HSW Theorem [8,9] (cf. [2]).

For the converse part (the "$\leq$" inequality), consider a code of block length n with error probability $\overline{P}_e$. The state after an encoding operation and action of the channel is given by

$$\Phi^{MB^n K} = \frac{1}{|M|}\sum_m p(m)|m\rangle\langle m| \otimes (\mathcal{N}^{\otimes n} \otimes \mathrm{id}^K)(\alpha_m^{A^n} \otimes \varkappa^{E^n K}),$$

and the state after a decoding operation is given by

$$\omega^{M\widetilde{M}} = \mathbb{1}^M \otimes \mathcal{D}(\Phi^{MB^n K}).$$

Then we have

$$\begin{aligned} nR &= H(M)_\omega \\ &= I(M:\widetilde{M})_\omega + H(M|\widetilde{M})_\omega \\ &\le I(M:\widetilde{M})_\omega + H(\overline{P}_e) + nR\overline{P}_e \\ &\le I(M:B^nK)_\Phi + n\varepsilon. \end{aligned}$$

The first inequality follows from the application of Fano's inequality, and the second one follows from the data-processing inequality, where $\varepsilon = \dfrac{1}{n} + R\overline{P}_e$. Setting $M = X$, we have

$$R \le \frac{1}{n} I(X:B^nK) + \varepsilon.$$

As $n \to \infty$ and $\varepsilon \to 0$, the upper bound on the rate follows. $\triangle$

The classical capacity assisted by entangled states of the form $\varkappa^{E^nK^n} = \varkappa^{E_1K_1} \otimes \ldots \otimes \varkappa^{E_nK_n}$ in Definition 3 is denoted by $C_{EH\otimes}(W)$, in analogy with $C_{H\otimes}(W)$. Thus, we can say for U_c (from (37)) that

$$C_{EH\otimes}(U_c) = C_{H\otimes}(\mathrm{SWAP} \otimes U_c).$$

As a consequence, $C_{EH\otimes}$ admits a single-letter characterization for U_c given by

$$C_{EH\otimes}(U_c) = \max_{p_i,|\eta\rangle} S\Big(\sum_i p_i(\mathbb{1} \otimes U_i)|\eta\rangle\langle\eta|(\mathbb{1} \otimes U_i)^\dagger\Big).$$

Lemma 3 [25]. *For two unitary operators* $\{U_1, U_2\} \in \mathrm{SU}(2)$, *with probability* $\{p_1, p_2\}$, *we have*

$$\max_{|\mu\rangle} S\left(\sum_i p_i U_i|\mu\rangle\langle\mu|U_i^\dagger\right) = \max_{|\gamma\rangle} S\left(\sum_i p_i(\mathbb{1} \otimes U_i)|\gamma\rangle\langle\gamma|(\mathbb{1} \otimes U_i)^\dagger\right),$$

where $|\mu\rangle$ *is a pure state in* A *and* $|\gamma\rangle$ *is a pure state in* AR.

We can use Lemma 3 to evaluate $C_{EH\otimes}$ for the universally classical-quantum two-qubit unitary interactions. This results in

$$C_{EH\otimes}(U_{c(2)}) = C_{H\otimes}(U_{c(2)}) = H_2\Big(\frac{1+\sin u}{2}\Big),$$

which shows that entanglement does not enhance the classical capacity in this case. But, clearly, from the examples of superadditivity presented in Section 3.3, we can see that pre-shared entanglement between Helen and Bob does indeed increase the classical communication capability.

5. CONFERENCING SENDER AND HELPER

In this section we define the capacity with conferencing encoders, that is, when Alice and Helen can freely communicate classical messages. A product-state capacity with conferencing encoders is also defined when Alice and Helen are respectively restricted to product-state encoding.

By referring to Fig. 8, an encoding CPTP map $\mathcal{E}\colon M \to \mathcal{L}(A^n) \otimes \mathcal{L}(E^n)$ can be thought of as two local encoding maps performed by Alice and Helen, respectively, and given by $\mathcal{E}_\mathcal{A}\colon M \to \mathcal{L}(A^n)$ and $\mathcal{E}_\mathcal{H}\colon M \to \mathcal{L}(E^n)$. These can be realized by preparing pure product states $\{|\alpha_m\rangle \otimes |\eta_m\rangle\}$ to be input across A^n and E^n of n instances of the channel. A decoding CPTP map $\mathcal{D}\colon \mathcal{L}(B^n) \to \widetilde{M}$ can be realized by a POVM $\{\Lambda_m\}$. The probability of error for a particular message m is

$$P_e(m) = 1 - \mathrm{Tr}\big(\Lambda_m \mathcal{N}^\otimes(\alpha_m^{A^n} \otimes \eta_m^{E^n})\big).$$

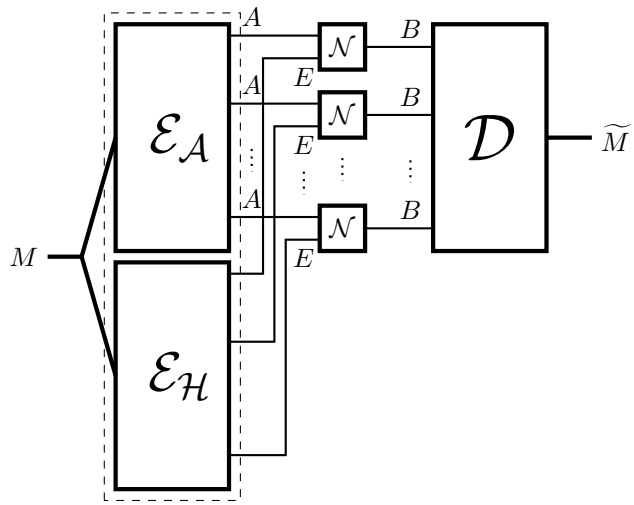

Fig. 8. Schematic of a general protocol to transmit classical information with conferencing encoders; $\mathcal{E}_{\mathcal{A}}$ and $\mathcal{E}_{\mathcal{H}}$ are the encoding maps of Alice and Helen, respectively. The decoding map is $\mathcal{D}$.

Definition 4. A *classical code for conferencing encoders* of block length n is a family of triples $(|\alpha_m\rangle^{A^n}, |\eta_m\rangle^{E^n}, \Lambda_m)$ with the error probability $\overline{P}_e := \frac{1}{|M|}\sum_m P_e(m)$ and rate $\frac{1}{n}\log|M|$. A rate R is achievable if there is a sequence of codes over their block length n with $\overline{P}_e$ converging to 0 and rate converging to R. The classical capacity with conferencing encoders of W, denoted by $C_{\mathrm{conf}}(W)$ or, equivalently, $C_{\mathrm{conf}}(\mathcal{N})$, is the maximum achievable rate. If the sender and helper are restricted to fully separable states $\alpha_m^{A^n}$ and $\eta_m^{E^n}$, i.e., convex combinations of tensor products $\eta_m^{E^n} = (\eta_{1m}^{E_1} \otimes \ldots \otimes \eta_{nm}^{E_n})$, and $\alpha_m^{A^n} = (\alpha_{1m}^{A_1} \otimes \ldots \otimes \alpha_{nm}^{A_n})$ for all m, the largest achievable rate is denoted by $C_{\mathrm{conf}\otimes}(W) = C_{\mathrm{conf}\otimes}(\mathcal{N})$ and is henceforth referred to as *classical capacity with product conferencing encoders.*

Theorem 4. *For an isometry $W\colon AE \to BF$, the classical capacity of the conferencing encoders model is given by*

$$C_{\mathrm{conf}}(W) = \sup_n \max_{\{p(x),\alpha_x^{A^n}\otimes\eta_x^{E^n}\}} \frac{1}{n} I(X:B^n)_\sigma, \tag{41}$$

where the mutual information is evaluated with respect to the state

$$\sigma = \sum_x p(x)|x\rangle\langle x| \otimes \mathcal{N}^{\otimes n}(\alpha_x^{A^n} \otimes \eta_x^{E^n})$$

and the maximization is over the ensemble $\{p(x), \alpha_x^{A^n} \otimes \eta_x^{E^n}\}$. The classical capacity with product conferencing encoders is given by

$$C_{\mathrm{conf}\otimes}(W) = \max_{\{p(x),\alpha_x^{A}\otimes\eta_x^{E}\}} I(X:B)_\sigma, \tag{42}$$

where the mutual information is evaluated with respect to the state

$$\sigma = \sum_x p(x)|x\rangle\langle x| \otimes \mathcal{N}(\alpha_x^{A} \otimes \eta_x^{E})$$

and the maximization is over the ensemble $\{p(x), \alpha_x^{A} \otimes \eta_x^{E}\}$.

Proof. The direct part, the "$\geq$" inequality, of the coding theorem follows from the HSW Theorem [8, 9] (cf. [2]). For the converse part, the "$\leq$" inequality, consider a code of block length n

with error probability $\overline{P}_e$. The state after an encoding operation and action of the channel is given by

$$\Phi^{MB^n} = \frac{1}{|M|}\sum_m p(m)|m\rangle^M\langle m| \otimes \mathcal{N}^{\otimes n}(\alpha_m^{A^n} \otimes \eta_m^{E^n}),$$

and the state after a decoding operation is given by

$$\omega^{M\widetilde{M}} = \mathbb{1}^M \otimes \mathcal{D}(\Phi^{MB^n}).$$

We then have

$$\begin{aligned} nR &= H(M)_\omega \\ &= I(M:\widetilde{M})_\omega + H(M|\widetilde{M})_\omega \\ &\le I(M:\widetilde{M})_\omega + H(\overline{P}_e) + nR\overline{P}_e \\ &\le I(M:B^n)_\Phi + n\varepsilon. \end{aligned}$$

The first inequality follows from the application of Fano's inequality, and the second one follows from the data-processing inequality, where $\varepsilon = \frac{1}{n} + R\overline{P}_e$. Setting $M = X$, we have

$$R \le \frac{1}{n} I(X:B^n) + \varepsilon.$$

As $n \to \infty$ and $\overline{P}_e \to 0$, the upper bound on the rate follows for C_{conf}. For $C_{\mathrm{conf}\otimes}$ we have an additional step, namely the additivity of mutual information. $\triangle$

From Theorem 4 it is also clear that $C_{\mathrm{conf}}(W) = \lim_{n\to\infty} \frac{1}{n} C_{\mathrm{conf}\otimes}(W^{\otimes n})$.

Remark 7. In classical information theory, conferencing encoders for MAC were introduced in [26], where coding theorems were provided. Here each sender can gain partial knowledge of the other sender(s) message through conferencing, i.e., noiseless exchange of messages, eventually constrained to occur at a given rate. This is an example of "cooperation," which is receiving an increasing attention in classical communication systems (see, e.g., [27]), while it is still very rarely considered in the quantum domain. An exception is provided by [28], where the results of [26] have been extended to classical-quantum MAC (both inputs being classical and the output quantum). In the conferencing encoders model, unlike [28], we assume free classical communication between Alice and Helen, with both of them aiming to send the same message. We must remark here that the use of this resource, i.e., free classical communication between Alice and Helen, does not trivialize the task, since the global input state is still restricted to the set of separable states.

5.1. Role of Entanglement in Conferencing Models (Superadditivity)

Entanglement played a peculiar role in the passive environment-assisted capacities and entanglement-environment-assisted capacities. We shall see in this section that this is also true for the case of conferencing encoders. We consider the following example to highlight the role of entanglement with conferencing encoders.

Let us assume that $|A| = |B| = |E| = |F| = d$. From (42) we see that

$$C_{\mathrm{conf}\otimes}(U) \le \log d - S_{\min}(U), \tag{43}$$

where $S_{\min}(U)$ is the minimum output entropy of U as defined in (29).

Lemma 4 [29]. *For a given unitary $U\colon A\otimes E \to B\otimes F$, Shor's augmented unitary $U^{\mathrm{aug}}\colon A_0 \otimes E \to B\otimes F_0$ with $A_0 = L\otimes A$ and $F_0 = F\otimes L$, where $|L| = d^2$, is depicted by the quantum circuit in Fig. 9. Then for any environment state η, the effective channel for U^{aug} is given by*

$$\mathcal{N}_\eta^{\mathrm{aug}}(\sigma^L \otimes \rho^A) := \sum_{x,z} W(x,z)\mathcal{N}_\eta(\rho^A)W(x,z)^\dagger \langle xz|\sigma|xz\rangle.$$

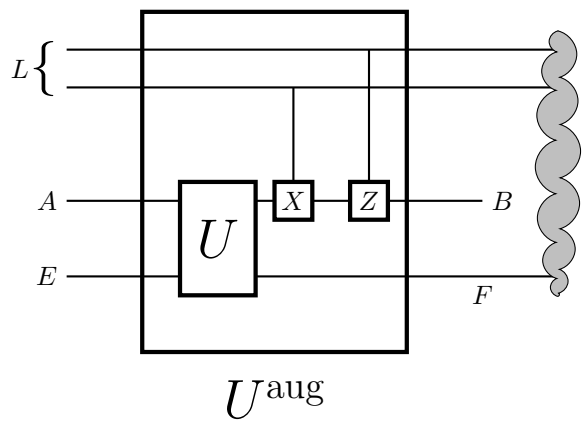

Fig. 9. Shor's augmented unitary U^{aug} of a given unitary U is depicted by the above quantum circuit. Here X is the cyclic shift operator, and Z is the phase operator as defined in (20).

Here $W(x,z)$ are the discrete Weyl operators. Then, for the augmented unitary,

$$C_{\text{conf}\otimes}(U^{\text{aug}}) = \log d - S_{\min}(U). \tag{44}$$

For *any* given unitaries $W_1\colon A\otimes E \to B\otimes F$ and $W_2\colon A'\otimes E' \to B'\otimes F'$, we can ask whether the product conferencing encoders capacity is additive, i.e., whether the following equality holds true:

$$C_{\text{conf}\otimes}(W_1\otimes W_2) \stackrel{?}{=} C_{\text{conf}\otimes}(W_1) + C_{\text{conf}\otimes}(W_2).$$

It is trivial to see that

$$C_{\text{conf}\otimes}(W_1\otimes W_2) \geq C_{\text{conf}\otimes}(W_1) + C_{\text{conf}\otimes}(W_2).$$

Now with the following example we show that the above inequality is strict in general.

Example. From Lemma 2 we can guarantee the existence of a unitary $V\colon A\otimes E \to B\otimes F$ (here all the parties are of equal dimension d) with the following lower bound on the minimum entropy:

$$S_{\min}(V) \geq \log d - \frac{1}{\ln 2} - 1.$$

Now consider $V^*\colon A'\otimes E' \to B'\otimes F'$, where the primed systems are isomorphic to the unprimed ones. Here V^* is the conjugate of V. It is useful to note that

$$S_{\min}(V) = S_{\min}(V^*).$$

The unitaries of interest are the Shor augmented unitaries V^{aug} and $V^{*\text{aug}}$. From (44) we have

$$C_{\text{conf}\otimes}(V^{\text{aug}}) \leq \frac{1}{\ln 2} + 1, \qquad C_{\text{conf}\otimes}(V^{*\text{aug}}) \leq \frac{1}{\ln 2} + 1.$$

Let us evaluate the product conferencing encoders capacity of $V^{\text{aug}}\otimes V^{*\text{aug}}$. Since $V^{\text{aug}}\otimes V^{*\text{aug}}$ is isomorphic to $(V\otimes V^*)^{\text{aug}}$, from (44) we have

$$C_{\text{conf}\otimes}(V^{\text{aug}}\otimes V^{*\text{aug}}) = C_{\text{conf}\otimes}((V\otimes V^*)^{\text{aug}}) = 2\log d - S_{\min}(V\otimes V^*). \tag{45}$$

When Alice inputs a maximally entangled state across AA', denoted by $|\Phi\rangle^{AA'}$, and Helen inputs a maximally entangled state across EE', denoted by $|\Phi\rangle^{EE'}$, the following holds true:

$$|\Phi\rangle^{AA'}\otimes|\Phi\rangle^{EE'} = |\Phi\rangle^{(AE)(A'E')}.$$

Also,

$$V \otimes V^*\big(|\Phi\rangle^{(AE)(A'E')}\big) = |\Phi\rangle^{(AE)(A'E')} = |\Phi\rangle^{AA'} \otimes |\Phi\rangle^{EE'}.$$

Thus, $S_{\min}(V \otimes V^*) = 0$, and from (45) we have

$$C_{\mathrm{conf}\otimes}(V^{\mathrm{aug}} \otimes V^{*\mathrm{aug}}) = 2\log d,$$

exhibiting the role of entanglement in enhancing conferencing communication. We would like to emphasize that in this example entanglement enables us to send the entire bandwidth, without which we can only send paltry amount of information.

5.2. The Classical Capacity with Conferencing Encoders is Always Nonzero

From Remark 4, we have seen that SWAP has $C_H = 0$. Now, when conferencing is allowed, i.e., Alice and Helen are on the same footing as the sender of information, we can send classical information at the maximum rate. This motivates us to study whether some positive amount of classical information can always be transmitted with conferencing encoders.

From the definition of $C_{\mathrm{conf}\otimes}$ and the previously defined quantities $\chi^A_{H\otimes}$, $\chi^H_{A\otimes}$ (see Section 3.1), we can see that

$$C_{\mathrm{conf}\otimes} \geq \max\{\chi^A_{H\otimes}, \chi^H_{A\otimes}\}.$$

Thus, for a unitary $U\colon A \otimes E \to B \otimes F$ with $|A| = |B| = |E| = |F| = d$, we can invoke the uncertainty relation of Theorem 2 to give a lower bound on the $C_{\mathrm{conf}\otimes}$ which reads as

$$C_{\mathrm{conf}\otimes} \geq \frac{\chi^A_{H\otimes} + \chi^H_{A\otimes}}{2} \geq \frac{1}{2^{14} d^2 \ln 2}\left(\frac{\sqrt{2 + 2(\log d)^2} - \sqrt{2}}{\log d}\right)^8.$$

Now we derive a lower bound when the dimensions of A, B, E, F are not equal.

Theorem 5. *Given a unitary $U\colon A \otimes E \to B \otimes F$ with $|A||E| = |B||F|$, we have*

$$C_{\mathrm{conf}\otimes} \geq \frac{3}{8\ln 2}\left(\frac{1}{|A||E|}\right)^4.$$

Proof. Let $C_{\mathrm{conf}\otimes}(U) = \delta$. Then, from the quantum Pinsker inequality [18], we have

$$\big\|\mathcal{N}(\alpha^A \otimes \eta^E) - \Omega^B\big\|_1 \leq \sqrt{2\delta \ln 2}, \quad \forall \alpha^A \otimes \eta^E, \tag{46}$$

where $\Omega^B := \sum p_i \mathcal{N}(\alpha^A_i \otimes \eta^E_i)$, the output of average of the ensemble $\{p_i, \alpha^A_i \otimes \eta^E_i\}$ which achieves the product conferencing capacity. Let us now consider the set of density operators $\{\sigma^Y_{m.n}\}$ defined as follows:

$$\sigma^Y_{m,n} := \begin{cases} |m\rangle\langle m|, & \text{when } m = n, \\ \frac{1}{2}(|m\rangle + |n\rangle)(\langle m| + \langle n|), & \text{when } m < n, \\ \frac{1}{2}(|m\rangle + i|n\rangle)(\langle m| - i\langle n|), & \text{when } m > n. \end{cases} \tag{47}$$

Here $\{|m\rangle\}$ denotes the computational basis of the Hilbert space Y. The set $\{\sigma_{m,n}\}$ spans $\mathcal{L}(Y)$. Also, $\{\sigma^A_{m,n}\} \otimes \{\sigma^E_{o,p}\}$ spans $\mathcal{L}(A \otimes E)$. Thus, for an arbitrary state on AE, we can write

$$\rho^{AE} = \sum_{m,n=1}^{|A|} \sum_{o,p=1}^{|E|} \lambda_{m,n,o,p} \sigma^A_{m.n} \otimes \sigma^E_{o,p}, \qquad \sum_{m,n=1}^{|A|} \sum_{o,p=1}^{|E|} \lambda_{m,n,o,p} = 1,$$

where

$$|\lambda_{m,n,o,p}|^2 \leq \frac{1}{1 - \max\big|\mathrm{Tr}(\sigma^A_{m,n}\sigma^A_{m',n'})\,\mathrm{Tr}(\sigma^E_{o,p}\sigma^E_{o',p'})\big|^2} = \frac{4}{3}. \tag{48}$$

The maximization is over all the indices with at least one of the primed indices not equal to the unprimed indices. From (47) we can see that the maximum is indeed reached for the case where exactly one primed index is different from the unprimed ones. Now

$$\begin{aligned}\|\mathcal{N}(\rho^{AE}) - \Omega^B\|_1 &= \left\| \sum_{m,n=1}^{|A|} \sum_{o,p=1}^{|E|} \lambda_{m,n,o,p} \mathcal{N}(\sigma^A_{m.n} \otimes \sigma^E_{o,p}) - \sum_{m,n=1}^{|A|} \sum_{o,p=1}^{|E|} \lambda_{m,n,o,p} \Omega^B \right\|_1 \\ &\le \sum_{m,n=1}^{|A|} \sum_{o,p=1}^{|E|} |\lambda_{m,n,o,p}| \|\mathcal{N}(\sigma^A_{m.n} \otimes \sigma^E_{o,p}) - \Omega^B\|_1,\end{aligned}$$

which is due to the application of the triangle inequality. From (46) and (48) we have

$$\|\mathcal{N}(\rho^{AE}) - \Omega^B\|_1 \le \sqrt{\frac{8}{3}\delta \ln 2}\,(|A||E|)^2.$$

Let us further choose two states $\rho_i^{AE} := U^\dagger(\omega_i^B \otimes \varkappa_i^F)U$ with the property that $\|\omega_1 - \omega_2\|_1 = 2$; i.e., they are perfectly distinguishable states. Hence we have

$$\left\|\mathcal{N}(\rho_i^{AE}) - \Omega^B\right\|_1 \le \sqrt{\frac{8}{3}\delta \ln 2}\,(|A||E|)^2, \quad i = 1, 2,$$

which implies

$$\|\mathcal{N}(\rho_1^{AE}) - \mathcal{N}(\rho_2^{AE})\|_1 \le 2\sqrt{\frac{8}{3}\delta \ln 2}\,(|A||E|)^2.$$

Hence, we must have $\sqrt{\frac{8}{3}\delta \ln 2}\,(|A||E|)^2 \ge 1$, since otherwise we have a contradiction. This leads to $\delta \ge \frac{3}{8\ln 2}\Big(\frac{1}{|A||E|}\Big)^4$. $\triangle$

Remark 8. When the U are chosen according to the Haar measure on $\mathrm{U}(d^2)$, from Lemma 2 and (43) we can give an upper bound on $\mathbf{E}(C_{\mathrm{conf}\otimes}(U))$, the expectation value of the classical capacity with product conferencing encoders, which reads as

$$\mathbf{E}(C_{\mathrm{conf}\otimes}(U)) \le 1 + \frac{1}{\ln 2}.$$

It follows that when $d \to \infty$, by the concentration of measure phenomenon [23], with overwhelming probability

$$C_{\mathrm{conf}\otimes}(U) < 2.5.$$

For two-qubit unitaries a much tighter lower bound can be found, which actually coincides with the upper bound, thus giving the classical capacity with conferencing encoders.

Theorem 6. *In the qubit case, i.e.,* $|A| = |E| = |B| = |F| = 2$, *for any unitary* $U\colon A \otimes E \to B \otimes F$ *we have*

$$C_{\mathrm{conf}\otimes}(U) = C_{\mathrm{conf}}(U) = 1. \tag{49}$$

Proof. Let U be a two-qubit unitary. For its adjoint $U^\dagger$ we have

$$U^\dagger(|\varphi\rangle^B \otimes |0\rangle^F) = |\Phi_0\rangle^{AE}, \qquad U^\dagger(|\varphi\rangle^B \otimes |1\rangle^F) = |\Phi_1\rangle^{AE}, \tag{50}$$

where $|\Phi_i\rangle^{AE}$ are generically entangled across AE.

Now note that the subspace spanned by $|\Phi_0\rangle^{AE}$ and $|\Phi_1\rangle^{AE}$ contains at least one product state [30]. Let, say,

$$c_0|\Phi_0\rangle^{AE} + c_1|\Phi_1\rangle^{AE}$$

be a product state in AE with $c_0, c_1 \in \mathbb{C}$. Thus, from (50)

$$U^\dagger(|\varphi\rangle^B \otimes (c_0|0\rangle^F + c_1|1\rangle^F))$$

is a product state in AE. For each choice of $|\varphi\rangle^B$ we can find

$$|\psi\rangle^F := c_0|0\rangle^F + c_1|1\rangle^F$$

such that $U^\dagger(|\varphi\rangle^B \otimes |\psi\rangle^F)$ is a product state in AE. Let $|\psi_0\rangle^F$ and $|\psi_1\rangle^F$ be such states for the choices $|0\rangle^B$ and $|1\rangle^B$, respectively, of $|\varphi\rangle^B$. Hence, for a given U, we can find two input states which are product across AE,

$$U^\dagger(|0\rangle^B \otimes |\psi_0\rangle^F), \qquad U^\dagger(|1\rangle^B \otimes |\psi_1\rangle^F),$$

such that we have two orthogonal output signals in system B, thus achieving the capacity of 1 bit. $\triangle$

Remark 9. The capacities C_H, $C_{H\otimes}$, C_{EH}, C_{conf}, and C_{conf} are continuous in the channel with respect to the diamond norm. Concretely, if $\|\mathcal{N} - \mathcal{M}\|_\diamond \le \varepsilon$, then we have

$$\begin{aligned}
|C_{H\otimes}(\mathcal{N}) - C_{H\otimes}(\mathcal{M})| &\le 2\varepsilon \log|B| + (2+\varepsilon)H_2\Big(\frac{\varepsilon}{2+\varepsilon}\Big),\\
|C_{H}(\mathcal{N}) - C_{H}(\mathcal{M})| &\le 2\varepsilon \log|B| + (2+\varepsilon)H_2\Big(\frac{\varepsilon}{2+\varepsilon}\Big),\\
|C_{EH}(\mathcal{N}) - C_{EH}(\mathcal{M})| &\le 2\varepsilon \log|B| + (2+\varepsilon)H_2\Big(\frac{\varepsilon}{2+\varepsilon}\Big),\\
|C_{\mathrm{conf}}(\mathcal{N}) - C_{\mathrm{conf}}(\mathcal{M})| &\le 2\varepsilon \log|B| + (2+\varepsilon)H_2\Big(\frac{\varepsilon}{2+\varepsilon}\Big),\\
|C_{\mathrm{conf}\otimes}(\mathcal{N}) - C_{\mathrm{conf}\otimes}(\mathcal{M})| &\le 2\varepsilon \log|B| + (2+\varepsilon)H_2\Big(\frac{\varepsilon}{2+\varepsilon}\Big).
\end{aligned} \tag{51}$$

Since each of these capacities is expressed in terms of the quantum mutual information of a classical-quantum state, and the optimization is over extra parameters due to the initial environment state, the above results can be obtained following the same arguments as in [19]; cf. [1]. One distinction is the use of the improved Alicki–Fannes continuity bound for conditional entropy [20] compared to the original form of Alicki–Fannes [31] as used in [19].

6. CONCLUSIONS

We have laid foundations of classical communication with environment assistance at the input. In such a model, a benevolent helper is able to select an initial environment state of the channel, modeled as unitary interaction. Capacities admit multi-letter formulas, both for the unrestricted and separable helper, which are hard to compute. These capacities are continuous like the unassisted ones, which are a special case of our model. Further, we have identified a class of unitaries which admit a single-letter formula for the transmission of classical capacity with separable helper. Also, we have shown superadditivity for both $C_{H\otimes}$ and C_H. Due to the unique role that SWAP plays in the examples of superadditivity, we considered entanglement-environment-assisted capacities, where there is a pre-shared entanglement between the helper and receiver.

The U_c (as defined in Section 3.2) constitute an interesting class of unitaries which are universally classical-quantum ($\in \mathfrak{CQ}$). In fact the $C_{H\otimes}$ and $C_{EH\otimes}$ admit single-letter characterizations. The capacity can be related to the problem of distinguishability of unitaries, when the Holevo quantity is a measure of distinguishability. When we consider the distinguishability as mentioned in [32] (i.e., with an ancillary system), this is equal to $C_{EH\otimes}(U_c)$. Thus, the additivity of these quantities can be related to the additivity of distinguishability for unitary operations.

We have introduced a conferencing encoders model where the sender and the helper are equipped with LOCC. Like the previous environment-assisted models, they admit a regularized formulas and are continuous. For a given unitary we can always transmit nonzero amount of classical information using a conferencing helper model. It would be interesting to find unitaries (if they exist) such that C_H^A and C_A^H are small but C_{conf} is large. At least in the case of unitaries where all the parties have equal dimensions, we can rule out such a possibility. This is due to the fact that a small C_H^A implies a large C_A^H due to an uncertainty type relation, thus making C_{conf} large. Furthermore, we have evaluated the classical capacity for conferencing encoders for two-qubit unitaries, which turns out to be 1 bit. The computation of unrestricted helper capacities $C_H, C_{EH}, C_{\text{conf}}$ is a major open problem.

Finally, it is worth noting that if Helen exploits entanglement across channel uses, we get memory effects on communication; hence, the present study can shed further light on the subject of memory quantum channels [33].

The first author thanks the Universitat Autònoma de Barcelona for kind hospitality.

APPENDIX

PARAMETRIZATION OF TWO-QUBIT UNITARIES

A general two-qubit unitary interaction can be described by 15 real parameters. For the analysis of classical capacities under consideration we follow the arguments used in [34] to reduce the number of parameters to 3 by the action of local unitaries with some further observations in [1]. According to the definition of capacities, the local unitaries on A, B, E, and F do not affect the environment-assisted classical capacity, since they could be incorporated into the encoding and decoding maps, respectively, or can be reflected in a different choice of environment state.

Lemma 5 [34]. *Any two-qubit unitary interaction V^{AE} is equivalent, up to local unitaries before and after the V^{AE}, to one of the forms*

$$U^{AE} = \sum_k e^{-i\lambda_k} |\Phi_k\rangle\langle\Phi_k| = \exp -\frac{i}{2}(\alpha_x \sigma_x \otimes \sigma_x + \alpha_y \sigma_y \otimes \sigma_y + \alpha_z \sigma_z \otimes \sigma_z) =: U(\alpha_x, \alpha_y, \alpha_z),$$

where σ_x, σ_y, and σ_z are the Pauli operators and $\frac{\pi}{2} \geq \alpha_x \geq \alpha_y \geq |\alpha_z| \geq 0$. Furthermore, the λ_k are

$$\lambda_1 := \frac{\alpha_x - \alpha_y + \alpha_z}{2}, \qquad \lambda_2 := \frac{-\alpha_x + \alpha_y + \alpha_z}{2},$$
$$\lambda_3 := \frac{-\alpha_x - \alpha_y - \alpha_z}{2}, \qquad \lambda_4 := \frac{\alpha_x + \alpha_y - \alpha_z}{2},$$

and the $|\Phi_k\rangle$ are the so-called "magic basis" vectors [35]

$$|\Phi_1\rangle := \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \qquad |\Phi_2\rangle := \frac{-i(|00\rangle - |11\rangle)}{\sqrt{2}},$$
$$|\Phi_3\rangle := \frac{|01\rangle - |10\rangle}{\sqrt{2}}, \qquad |\Phi_4\rangle := \frac{-i(|01\rangle + |10\rangle)}{\sqrt{2}}. \tag{52}$$

This is of course the familiar Bell basis, but with peculiar phases.

Hence, the parameter space given by

$$\mathfrak{T}_{\text{total}} = \left\{ (\alpha_x, \alpha_y, \alpha_z) : \frac{\pi}{2} \geq \alpha_x \geq \alpha_y \geq |\alpha_z| \geq 0 \right\} \tag{53}$$

describes all two-qubit unitaries up to a local basis choice. This forms a tetrahedron with vertices $(0,0,0)$, $(\frac{\pi}{2},0,0)$, $(\frac{\pi}{2},\frac{\pi}{2},-\frac{\pi}{2})$, and $(\frac{\pi}{2},\frac{\pi}{2},\frac{\pi}{2})$.

Since we are interested in evaluating the capacities of unitaries, we use

$$U\Big(\alpha_x, \alpha_y, \frac{\pi}{2} + \alpha_z\Big) = -i(\sigma_z \otimes \mathbb{1})U^*\Big(\alpha_x, \alpha_y, \frac{\pi}{2} - \alpha_z\Big)(\mathbb{1} \otimes \sigma_z), \tag{54}$$

where U^* is the complex conjugate of U. Note that the latter has the same environment-assisted classical capacities; indeed, any code for U is transformed into a code for U^* by taking complex conjugates. The reduced parameter space given by

$$\mathfrak{T} = \left\{(\alpha_x, \alpha_y, \alpha_z) : \frac{\pi}{2} \geq \alpha_x \geq \alpha_y \geq \alpha_z \geq 0\right\} \tag{55}$$

describes all two-qubit unitaries up to a local basis choice and complex conjugation (we should note that in general $U \otimes V$ and $U \otimes V^*$ have different environment-assisted capacities, and in such cases we should consider $\mathfrak{T}_{\text{total}}$, say, for example, to provide a complete characterization of superadditivity). This forms a tetrahedron with vertices $(0,0,0)$, $(\frac{\pi}{2},0,0)$, $(\frac{\pi}{2},\frac{\pi}{2},0)$, and $(\frac{\pi}{2},\frac{\pi}{2},\frac{\pi}{2})$.

Familiar two-qubit gates can easily be identified within this parameter space: for instance, $(0,0,0)$ represents the identity $\mathbb{1}$, $(\frac{\pi}{2},0,0)$ the CNOT, $(\frac{\pi}{2},\frac{\pi}{2},0)$ the DCNOT (double controlled-not), and $(\frac{\pi}{2},\frac{\pi}{2},\frac{\pi}{2})$ the SWAP gate, respectively.

Consider the unitaries $U_c'(d)$ with $\alpha_x = 0$, $\alpha_y = 0$, $\alpha_z = d$, where $0 \leq d \leq \frac{\pi}{2}$. These unitaries lie outside the tetrahedron $\mathfrak{T}$. When expressed in matrix form, these unitaries are diagonal in the magic basis (same order as in (52)) with the diagonal elements $\{e^{-i\frac{d}{2}}, e^{-i\frac{d}{2}}, e^{i\frac{d}{2}}, e^{i\frac{d}{2}}\}$. To see their parametric representation in the tetrahedron $\mathfrak{T}$, we follow the argument used in [36, Appendix A] (cf. [1, Example 12]). Observe that the spectrum of ${U_c'}^T U_c'$ is $(e^{-2i\lambda_1}, e^{-2i\lambda_2}, e^{-2i\lambda_3}, e^{-2i\lambda_4})$, where the transpose operator is with respect to the magic basis. The spectrum of ${U_c'}^T U_c'$ is thus $(e^{id}, e^{id}, e^{-id}, e^{-id})$. Using the order property, $\frac{\pi}{2} \geq \lambda_4 \geq \lambda_1 \geq \lambda_2 \geq \lambda_3 \geq -\frac{3\pi}{4}$ (condition (55) written in terms of λ_k), and by solving the linear equations in α_x, α_y, and α_z, we get the parametric points as $(d,0,0)$, which correspond to the edge joining the identity $\mathbb{1}$ and CNOT; i.e., these unitaries are controlled-unitaries of the form $\sum_{i=0}^{1} |i\rangle^B\langle i|^A \otimes U_i^{E\to F}$, where $U_i \in \mathrm{SU}(2)$. Here the parameter d is given by $d = t$ when $t \leq \frac{\pi}{2}$ and $d = \pi - t$ when $t \geq \frac{\pi}{2}$, where $2\cos t = \operatorname{Tr} U_0^\dagger U_1$.

Now, when we apply SWAP to $U_c'(d)$, i.e., the unitary of interest, SWAP $\cdot U_c'(d) = U\Big(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2} + d\Big)$ is outside the parameter tetrahedron $\mathfrak{T}$. From (54), we get $U\Big(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2} + d\Big) = U^*\Big(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2} - d\Big)$, up to local unitaries, which lie in the parameter space. In essence, up to local unitaries and complex conjugation $U_{c(2)} := \sum_{i=0}^{1} |i\rangle^F\langle i|^A \otimes U_i^{E\to B}$ has parameters $(\frac{\pi}{2}, \frac{\pi}{2}, u)$, where $u = \frac{\pi}{2} - d$.

REFERENCES

1. Karumanchi, S., Mancini, S., Winter, A., and Yang, D., Quantum Channel Capacities with Passive Environment Assistance, *IEEE Trans. Inform. Theory*, 2016, vol. 62, no. 4, pp. 1733–1747.

2. Wilde, M.M., *Quantum Information Theory*, Cambridge: Cambridge Univ. Press, 2013.

3. Gregoratti, M. and Werner, R.F., Quantum Lost and Found, *J. Mod. Opt.*, 2003, vol. 50, no. 6–7, pp. 915–933.

4. Gregoratti, M. and Werner, R.F., On Quantum Error-Correction by Classical Feedback in Discrete Time, *J. Math. Phys.*, 2004, vol. 45, no. 7, pp. 2600–2612.

5. Hayden, P. and King, C., Correcting Quantum Channels by Measuring the Environment, *Quantum Inf. Comput.*, 2005, vol. 5, no. 2, pp. 156–160.

6. Smolin, J.A., Verstraete, F., and Winter, A., Entanglement of Assistance and Multipartite State Distillation, *Phys. Rev. A*, 2005, vol. 72, no. 5, p. 052317.

7. Winter, A., On Environment-Assisted Capacities of Quantum Channels, *Markov Process. Related Fields*, 2007, vol. 13, no. 2, pp. 297–314.

8. Holevo, A.S., The Capacity of the Quantum Channel with General Signal States, *IEEE Trans. Inform. Theory*, 1998, vol. 44, no. 1, pp. 269–273.

9. Schumacher, B. and Westmoreland, M.D., Sending Classical Information via Noisy Quantum Channels, *Phys. Rev. A*, 1997, vol. 56, no. 1, pp. 131–138.

10. Hastings, M.B., Superadditivity of Communication Capacity Using Entangled Inputs, *Nat. Phys.*, 2009, vol. 5, no. 4, pp. 255–257.

11. Amosov, G.G. and Mancini, S., The Decreasing Property of Relative Entropy and the Strong Superadditivity of Quantum Channels, *Quantum Inf. Comput.*, 2009, vol. 9, no. 7, pp. 594–609.

12. Schumacher, B. and Westmoreland, M.D., Optimal Signal Ensembles, *Phys. Rev. A*, 2001, vol. 63, no. 2, p. 022308.

13. Cortese, J.A., Quantum Information Theory: Classical Communication over Quantum Channels, *PhD Thesis*, Pasadena, CA: California Inst. of Technology, 2003. Available at `http://thesis.library.caltech.edu/649/1/thesis.pdf`.

14. Horodecki, M., Shor, P.W., and Ruskai, M.B., Entanglement Breaking Channels, *Rev. Math. Phys.*, 2003, vol. 15, no. 6, pp. 629–641.

15. Winter, A., The Capacity of the Quantum Multiple-Access Channel, *IEEE Trans. Inform. Theory*, 2001, vol. 47, no. 7, pp. 3059–3065.

16. Yard, J., Hayden, P., and Devetak, I., Capacity Theorems for Quantum Multiple-Access Channels: Classical-Quantum and Quantum-Quantum Capacity Regions, *IEEE Trans. Inform. Theory*, 2008, vol. 54, no. 7, pp. 3091–3113.

17. Kretschmann, D., Schlingemann, D., and Werner, R.F., The Information-Disturbance Tradeoff and the Continuity of Stinespring's Representation, *IEEE Trans. Inform. Theory*, 2008, vol. 54, no. 4, pp. 1708–1717.

18. Ohya, M. and Petz, D., *Quantum Entropy and Its Use*, Berlin: Springer-Verlag, 1993, 2nd ed.

19. Leung, D. and Smith, G., Continuity of Quantum Channel Capacities, *Comm. Math. Phys.*, 2009, vol. 292, no. 1, pp. 201–215.

20. Winter, A., Tight Uniform Continuity Bounds for Quantum Entropies: Conditional Entropy, Relative Entropy Distance and Energy Constraints, `arXiv:1507.07775v6 [quant-ph]`, 2015.

21. Maassen, H. and Uffink, J.B.M., Generalized Entropic Uncertainty Relations, *Phys. Rev. Lett.*, 1988, vol. 60, no. 12, pp. 1103–1106.

22. Chen, J., Ji, Z., Kribs, D.W., and Zeng, B., Minimum Entangling Power is Close to Its Maximum, `arXiv:1210.1296 [quant-ph]`, 2012.

23. Hayden, P., Leung, D.W., and Winter, A., Aspects of Generic Entanglement, *Comm. Math. Phys.*, 2006, vol. 265, no. 1, pp. 95–117.

24. Deschapms, J., Nechita, I., and Pellegrini, C., On Some Classes of Bipartite Unitary Operators, `arXiv:1509.06543 [quant-ph]`, 2015.

25. D'Ariano, G.M., Lo Presti, P., and Paris, M.G.A., Using Entanglement Improves the Precision of Quantum Measurements, *Phys. Rev. Lett.*, 2001, vol. 87, no. 27, p. 270404.

26. Willems, F.M.J., The Discrete Memoryless Multiple Access Channel with Partially Cooperating Encoders, *IEEE Trans. Inform. Theory*, 1983, vol. 29, no. 3, pp. 441–445.

27. Kramer, G., Marić, I., and Yates, R.D., Cooperative Communications, *Found. Trends Network.*, 2006, vol. 1, no. 3–4, pp. 271–425.

28. Boche, H. and Nötzel, J., The Classical-Quantum Multiple Access Channel with Conferencing Encoders and with Common Messages, *Quantum Inf. Process.*, 2014, vol. 13, no. 12, pp. 2595–2617.

29. Shor, P.W., Equivalence of Additivity Questions in Quantum Information Theory, *Comm. Math. Phys.*, 2004, vol. 246, no. 3, pp. 453–472.

30. Sanpera, A., Tarrach, R., and Vidal, G., Local Description of Quantum Inseparability, *Phys. Rev. A (3)*, 1998, vol. 58, no. 2, pp. 826–830.

31. Alicki, R. and Fannes, M., Continuity of Quantum Conditional Information, *J. Phys. A*, 2004, vol. 37, no. 5, pp. L55–L57.

32. Yang, D., Distinguishability, Classical Information of Quantum Operations, `arXiv:quant-ph/0504073`, 2005.

33. Caruso, F., Giovannetti, V., Lupo, C., and Mancini, S., Quantum Channels and Memory Effects, *Rev. Mod. Phys.*, 2014, vol. 86, no. 4, pp. 1203–1259.

34. Kraus, B. and Cirac, J.I., Optimal Creation of Entanglement Using a Two-Qubit Gate, *Phys. Rev. A*, 2001, vol. 63, no. 6, p. 062309.

35. Hill, S. and Wootters, W.K., Entanglement of a Pair of Quantum Bits, *Phys. Rev. Lett.*, 1997, vol. 78, no. 26, pp. 5022–5025.

36. Hammerer, K., Vidal, G., and Cirac, J.I., Characterization of Nonlocal Gates, *Phys. Rev. A*, 2002, vol. 66, no. 6, p. 062321.