

# General Mixed State Quantum Data Compression with and without Entanglement Assistance

Zahra Baghali Khanian<sup>1,2</sup> and Andreas Winter<sup>1,3</sup>

<sup>1</sup>*Física Teòrica: Informació i Fenòmens Quàntics, Departament de Física,  
Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain*

<sup>2</sup>*ICFO—Institut de Ciències Fotòniques,  
Barcelona Institute of Science and Technology, 08860 Castelldefels, Spain\**

<sup>3</sup>*ICREA—Institut Catalana de Recerca i Estudis Avançats,  
Pg. Lluís Companys, 23, 08010 Barcelona, Spain†*

(Dated: 16 December 2019)

We consider the most general (finite-dimensional) quantum mechanical information source, which is given by a quantum system  $A$  that is correlated with a reference system  $R$ . The task is to compress  $A$  in such a way as to reproduce the joint source state  $\rho^{AR}$  at the decoder with asymptotically high fidelity. This includes Schumacher’s original quantum source coding problem of a pure state ensemble and that of a single pure entangled state, as well as general mixed state ensembles. Here, we determine the optimal compression rate (in qubits per source system) in terms of the Koashi-Imoto decomposition of the source into a classical, a quantum, and a redundant part. The same decomposition yields the optimal rate in the presence of unlimited entanglement between compressor and decoder, and indeed the full region of feasible qubit-ebit rate pairs.

## I. WHAT IS A QUANTUM SOURCE?

A quantum source is a quantum system together correlations with a *reference system*. A criterion of how well a source is reproduced in a communication task is to measure how well the correlations are preserved with the reference system. Without correlation, the information does not make sense because a known quantum state without correlations can be reproduced at the destination without any communication.

To elaborate more on these notions, consider the source that Schumacher defined in his 1995 paper [1, 2] as an ensemble of pure states  $\{p(x), |\psi_x\rangle^A\}$ , where the source generates the state  $|\psi_x\rangle$  with probability  $p(x)$ . The figure of merit for the encoding-decoding process is to keep the decoded quantum states *on average* very close to the original states with respect to the fidelity, where the average is taken over the probability distribution  $p(x)$ . By basic algebra one can show that this is equivalent to preserving the classical-quantum state  $\rho^{AX} = \sum_x p(x) |\psi_x\rangle\langle\psi_x|^A \otimes |x\rangle\langle x|^X$ , where system  $A$  is the quantum system to be compressed. Another source model that Schumacher considered was the purification of the source ensemble, that is the state  $|\psi\rangle^{AR} = \sum_x \sqrt{p(x)} |\psi_x\rangle^A |x\rangle^R$ , where the figure of merit for the encoding-decoding process was to preserve the pure state correlations with the reference system  $R$  by maintaining a high fidelity between the decoded state and  $\psi$ . He showed that both definitions lead to the same compression rate, namely, the von Neumann entropy of the source  $S(A)_\rho = S(\rho^A)$ , where  $\rho^A = \text{Tr}_R \rho^{AR}$ . Incidentally, the full proof of optimality in the first model, without any additional restrictions on the encoder, had to wait until [3] (see also [5]); the strong converse, i.e. the optimality of the entropy rate even for constant error bounded away from 1, was eventually given in [4].

Another example of a quantum source is the mixed state source considered by Horodecki [5] and Barnum *et al.* [6], and finally solved by Koashi and Imoto [7], where the source is defined as an ensemble of mixed states  $\{p(x), \rho_x^A\}$ . Preserving these mixed quantum states, on average, in the process of encoding-decoding, the task is equivalent to preserving the state  $\rho^{AX} = \sum_x p(x) \rho_x^A \otimes |x\rangle\langle x|^X$ , that is the quantum system  $A$  together with its correlation with the classical reference system  $X$ .

The reference system is not usually considered in the description of classical information theory tasks, but arguably it is conceptually necessary in quantum information. This is because it allows us to present the figure of merit quantifying the decoding error as operationally accessible, for example via the probability of passing a test in the form of a measurement on the combined  $AR$ -system. This point is made eloquently in the early work of Schumacher on quantum information transmission [8, 9].

---

\* [zbkhanian@gmail.com](mailto:zbkhanian@gmail.com)

† [andreas.winter@uab.cat](mailto:andreas.winter@uab.cat)

In this work, we consider the most general finite-dimensional source in the realm of quantum mechanics, namely a quantum system  $A$  that is correlated with a reference system  $R$  in an arbitrary way, described by the overall state  $\rho^{AR}$ . In particular, the reference does not necessarily purify the source, nor is it assumed to be classical. The ensemble source and the pure source defined by Schumacher are special cases of this model, where the reference is a classical system in the former and a purifying system in the latter. So is the source considered by Koashi and Imoto in [7], where the reference system is classical, too.

Understanding the compression of the source  $\rho^{AR}$  has paramount importance in the field of quantum information theory and unifies all the models that have been considered in the literature. Schumacher's pure source model in a sense is the most stringent model because it requires preserving the correlations with a purifying reference system which implies that the correlations with any other reference system is preserved which follows from the fact that the fidelity is non-decreasing under quantum channels. However, the converse is not necessarily true: if in a compression task the parties are required to preserve the correlations with a given reference system which does not purify the source state, they might be able to compress more efficiently compared to the scenario where the reference system purifies the source. This is exactly what we show in this paper: we characterise the gap precisely depending on the reference system.

We find the optimal trade-off between the quantum and entanglement rates of the compression which are in terms of a decomposition of the state  $\rho^{AR}$  introduced in [10]. This decomposition is a generalization of the decomposition introduced by Koashi and Imoto for a set of quantum states in [11], so when the reference system is classical, the quantum rate reduces to the rate derived by Koashi and Imoto. We show the optimality of the rates with a new converse proof which is based on the decoupling of the environment systems of the encoding and decoding operations from the decoded systems and gives us an insight into how general mixed states are processed in an encoding-decoding task. Our results also cover the entanglement assisted compression task considered in [12] when the side information system is trivial, as well as the entanglement assisted version of the Koashi-Imoto compression.

The structure of the paper is as follows. In the remainder of this section, we introduce the notation that we use throughout the paper. In Sec. II, we rigorously define the task of the asymptotic compression of the source  $\rho^{AR}$ , where as for the communication purposes, we let the encoder and decoder share initial entanglement, and the encoder sends the compressed information to the decoder through a noiseless quantum channels. In Sec. III, we first introduce the Koashi-Imoto decomposition of the state  $\rho^{AR}$ , and then in Theorem 2 we state the main result of this paper, that is the optimal rate region for the compression of the source in terms of the trade-off between the entanglement and quantum rates, then we prove the achievability of the rates in the same section, but we leave the converse proofs for the subsequent sections which need more involved machinery. In Sec. IV, we define two functions which emerge in the converse proofs, and in Lemma 6 we state some important properties of these functions which then we use to prove the tight asymptotic converse bounds of Theorem 2. We prove Lemma 6 in Sec. V. Finally, in Sec. VI we discuss our results and some related open problems.

**Notation.** Quantum systems are associated with (finite dimensional) Hilbert spaces  $A$ ,  $R$ , etc., whose dimensions are denoted by  $|A|$ ,  $|R|$ , respectively. Since it is clear from the context, we slightly abuse the notation and let  $Q$  denote both a quantum system and a quantum rate. The von Neumann entropy is defined as  $S(\rho) = -\text{Tr}\rho \log \rho$  (throughout this paper,  $\log$  denotes by default the binary logarithm). The conditional entropy and the conditional mutual information,  $S(A|B)_\rho$  and  $I(A : B|C)_\rho$ , respectively, are defined in the same way as their classical counterparts:

$$\begin{aligned} S(A|B)_\rho &= S(AB)_\rho - S(B)_\rho, \text{ and} \\ I(A : B|C)_\rho &= S(A|C)_\rho - S(A|BC)_\rho \\ &= S(AC)_\rho + S(BC)_\rho - S(ABC)_\rho - S(C)_\rho. \end{aligned}$$

The fidelity between two states  $\rho$  and  $\xi$  is defined as  $F(\rho, \xi) = \|\sqrt{\rho}\sqrt{\xi}\|_1 = \text{Tr}\sqrt{\rho^{\frac{1}{2}}\xi\rho^{\frac{1}{2}}}$ , with the trace norm  $\|X\|_1 = \text{Tr}|X| = \text{Tr}\sqrt{X^\dagger X}$ . It relates to the trace distance in the following well-known way [13]:

$$1 - F(\rho, \xi) \leq \frac{1}{2} \|\rho - \xi\|_1 \leq \sqrt{1 - F(\rho, \xi)^2}. \quad (1)$$

## II. THE COMPRESSION TASK

We will consider the information theoretic limit of many copies of the source  $\rho^{AR}$ , i.e.  $\rho^{A^n R^n} = (\rho^{AR})^{\otimes n}$ . We assume that the encoder, Alice, and the decoder, Bob, have initially a maximally entangled state  $\Phi_K^{A_0 B_0}$  on registers  $A_0$  and  $B_0$  (both of dimension  $K$ ). The encoder, Alice, performs the encoding compression operation  $\mathcal{C} : A^n A_0 \rightarrow M$  on the system  $A^n$  and her part  $A_0$  of the entanglement, which is a quantum channel, i.e. a completely positive and trace preserving (CPTP) map. Notice that as functions CPTP maps act on the operators (density matrices) over the respective input and output Hilbert spaces, but as there is no risk of confusion, we will simply write the Hilbert spaces when denoting a CPTP map. Alice's encoding operation produces the state  $\sigma^{M B_0 R^n}$  with  $M$  and  $B_0$  as the compressed system of Alice and Bob's part of the entanglement, respectively. The dimension of the compressed system is without loss of generality not larger than the dimension of the original source, i.e.  $|M| \leq |A|^n$ . We call  $\frac{1}{n} \log K$  and  $\frac{1}{n} \log |M|$  the *entanglement rate* and *quantum rate* of the compression protocol, respectively. The system  $M$  is then sent to Bob via a noiseless quantum channel, who performs a decoding operation  $\mathcal{D} : M B_0 \rightarrow \hat{A}^n$  on the system  $M$  and his part of the entanglement  $B_0$ . We say the encoding-decoding scheme has *fidelity*  $1 - \epsilon$ , or *error*  $\epsilon$ , if

$$F(\rho^{A^n R^n}, \xi^{\hat{A}^n R^n}) \geq 1 - \epsilon, \quad (2)$$

where  $\xi^{\hat{A}^n R^n} = ((\mathcal{D} \circ \mathcal{C}) \otimes \text{id}_{R^n}) \rho^{A^n R^n}$ . Moreover, we say that  $(E, Q)$  is an (asymptotically) achievable rate pair if for all  $n$  there exist codes such that the fidelity converges to 1, and the entanglement and quantum rates converge to  $E$  and  $Q$ , respectively. The rate region is the set of all achievable rate pairs, as a subset of  $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$ .

According to Stinespring's theorem [14], a CPTP map  $\mathcal{T} : A \rightarrow \hat{A}$  can be dilated to an isometry  $U : A \hookrightarrow \hat{A}E$  with  $E$  as an environment system, called an isometric extension of a CPTP map, such that  $\mathcal{T}(\rho^A) = \text{Tr}_E(U\rho^A U^\dagger)$ . Therefore, the encoding and decoding operations are in general be viewed as isometries  $U_{\mathcal{E}} : A^n A_0 \hookrightarrow MW$  and  $U_{\mathcal{D}} : MB_0 \hookrightarrow \hat{A}^n V$ , respectively, with the systems  $W$  and  $V$  as the environment systems of Alice and Bob, respectively.

We say a source  $\omega^{BR}$  is equivalent to a source  $\rho^{AR}$  if there are CPTP maps  $\mathcal{T} : A \rightarrow B$  and  $\mathcal{R} : B \rightarrow A$  in both directions taking one to the other:

$$\omega^{BR} = (\mathcal{T} \otimes \text{id}_R) \rho^{AR} \quad \text{and} \quad \rho^{AR} = (\mathcal{R} \otimes \text{id}_R) \omega^{BR}. \quad (3)$$

The rate regions of equivalent sources are the same, because any achievable rate pair for one source is achievable for the other source as well. This follows from the fact that for any code  $(\mathcal{C}, \mathcal{D})$  of block length  $n$  and error  $\epsilon$  for  $\rho^{AR}$ , concatenating the encoding and decoding operations with  $\mathcal{T}$  and  $\mathcal{R}$ , i.e. letting  $\mathcal{C}' = \mathcal{C} \circ \mathcal{R}^{\otimes n}$  and  $\mathcal{D}' = \mathcal{T}^{\otimes n} \circ \mathcal{D}$ , we get a code of the same error  $\epsilon$  for  $\omega^{BR}$ . Analogously we can turn a code for  $\omega^{BR}$  into one for  $\rho^{AR}$ .

## III. THE QUBIT-EBIT RATE REGION

The idea behind the compression of the source  $\rho^{AR}$  is based on a decomposition of this state introduced in [10], which is a generalization of the decomposition introduced by Koashi and Imoto in [11]. Namely, for any set of quantum states  $\{\rho_x\}$ , there is a unique decomposition of the Hilbert space describing the structure of CPTP maps which preserve the set  $\{\rho_x^A\}$ . This idea was generalized in [10] for a general mixed state  $\rho^{AR}$  describing the structure of CPTP maps acting on system  $A$  which preserve the overall state  $\rho^{AR}$ . This was achieved by showing that any such map preserves the set of all possible states on system  $A$  which can be obtained by measuring system  $R$ , and conversely any map preserving the set of all possible states on system  $A$  obtained by measuring system  $R$ , preserves the state  $\rho^{AR}$ , thus reducing the general case to the case of classical-quantum states

$$\rho^{AY} = \sum_y q(y) \rho_y^A \otimes |y\rangle\langle y|^Y = \sum_y \text{Tr}_R \rho^{AR} (\mathbb{1}_A \otimes M_y^R) \otimes |y\rangle\langle y|^Y,$$

which is the ensemble case considered by Koashi and Imoto. As a matter of fact, looking at the algorithm presented in [11] to compute the decomposition, it is enough to consider an informationally complete POVM  $(M_y)$  on  $R$ , with no more than  $|R|^2$  many outcomes. The properties of this decomposition are stated in the following theorem.

**Theorem 1** ([10, 11]) *Associated to the state  $\rho^{AR}$ , there are Hilbert spaces  $C$ ,  $N$  and  $Q$  and an isometry  $U_{KI}: A \hookrightarrow CNQ$  such that:*

1. *The state  $\rho^{AR}$  is transformed by  $U_{KI}$  as*

$$(U_{KI} \otimes \mathbb{1}_R) \rho^{AR} (U_{KI}^\dagger \otimes \mathbb{1}_R) = \sum_j p_j |j\rangle\langle j|^C \otimes \omega_j^N \otimes \rho_j^{QR} =: \omega^{CNQR}, \quad (4)$$

where the set of vectors  $\{|j\rangle^C\}$  form an orthonormal basis for Hilbert space  $C$ , and  $p_j$  is a probability distribution over  $j$ . The states  $\omega_j^N$  and  $\rho_j^{QR}$  act on the Hilbert spaces  $N$  and  $Q \otimes R$ , respectively.

2. *For any CPTP map  $\Lambda$  acting on system  $A$  which leaves the state  $\rho^{AR}$  invariant, that is  $(\Lambda \otimes \text{id}_R) \rho^{AR} = \rho^{AR}$ , every associated isometric extension  $U: A \hookrightarrow AE$  of  $\Lambda$  with the environment system  $E$  is of the following form*

$$U = (U_{KI} \otimes \mathbb{1}_E)^\dagger \left( \sum_j |j\rangle\langle j|^C \otimes U_j^N \otimes \mathbb{1}_j^Q \right) U_{KI}, \quad (5)$$

where the isometries  $U_j: N \hookrightarrow NE$  satisfy  $\text{Tr}_E[U_j \omega_j U_j^\dagger] = \omega_j$  for all  $j$ . The isometry  $U_{KI}$  is unique (up to trivial change of basis of the Hilbert spaces  $C$ ,  $N$  and  $Q$ ). Henceforth, we call the isometry  $U_{KI}$  and the state  $\omega^{CNQR} = \sum_j p_j |j\rangle\langle j|^C \otimes \omega_j^N \otimes \rho_j^{QR}$  the Koashi-Imoto (KI) isometry and KI-decomposition of the state  $\rho^{AR}$ , respectively.

3. *In the particular case of a tripartite system  $CNQ$  and a state  $\omega^{CNQR}$  already in Koashi-Imoto form (4), property 2 says the following: For any CPTP map  $\Lambda$  acting on systems  $CNQ$  with  $(\Lambda \otimes \text{id}_R) \omega^{CNQR} = \omega^{CNQR}$ , every associated isometric extension  $U: CNQ \hookrightarrow CNQE$  of  $\Lambda$  with the environment system  $E$  is of the form*

$$U = \sum_j |j\rangle\langle j|^C \otimes U_j^N \otimes \mathbb{1}_j^Q, \quad (6)$$

where the isometries  $U_j: N \hookrightarrow NE$  satisfy  $\text{Tr}_E[U_j \omega_j U_j^\dagger] = \omega_j$  for all  $j$ .

According to the discussion at the end of Sec. II, the sources  $\rho^{AR}$  and  $\omega^{CNQR}$  are equivalent because there are the isometry  $U_{KI}$  and the reversal CPTP map  $\mathcal{R}: CNQ \rightarrow A$ , which reverses the action of the KI isometry, such that:

$$\begin{aligned} \omega^{CNQR} &= (U_{KI} \otimes \mathbb{1}_R) \rho^{AR} (U_{KI}^\dagger \otimes \mathbb{1}_R) \\ \rho^{AR} &= (\mathcal{R} \otimes \text{id}_R) \omega^{CNQR} = (U_{KI}^\dagger \otimes \mathbb{1}_R) \omega^{CNQR} (U_{KI} \otimes \mathbb{1}_R) + \text{Tr}[(\mathbb{1}_{CNQ} - \Pi_{CNQ}) \omega^{CNQR}] \sigma, \end{aligned} \quad (7)$$

where  $\Pi_{CNQ} = U_{KI} U_{KI}^\dagger$  is the projection onto the subspace  $U_{KI} A \subset C \otimes N \otimes Q$ , and  $\sigma$  is an arbitrary state acting on  $A \otimes R$ . Henceforth we assume that the source is  $\omega^{CNQR}$ , which is convenient because our main result is expressed in terms of the systems  $C$  and  $Q$ . Notice that the source  $\omega^{CNQR}$  is in turn equivalent to  $\omega^{CQR}$ , a fact we will exploit in the proof.

Moreover, since the information in  $C$  is classical, we can reduce the compression rate even more if the sender and receiver share entanglement, by using dense coding of  $j$ . In the following theorem we show the optimal qubit-ebit rate tradeoff for the compression of the source  $\rho^{AR}$ .

**Theorem 2** *For the compression of the source  $\rho^{AR}$ , all asymptotically achievable entanglement and quantum rate pairs  $(E, Q)$  satisfy*

$$\begin{aligned} Q &\geq S(CQ)_\omega - \frac{1}{2} S(C)_\omega, \\ Q + E &\geq S(CQ)_\omega, \end{aligned}$$

where the entropies are with respect the KI decomposition of the state  $\rho^{AR}$ , i.e. the state  $\omega^{CNQR}$ . Conversely, all the rate pairs satisfying the above inequalities are asymptotically achievable.

**Remark 3** This theorem implies that the optimal asymptotic quantum rates for the compression of the source  $\rho^{AR}$  with and without entanglement assistance are  $S(CQ)_\omega - \frac{1}{2}S(C)_\omega$  and  $S(CQ)_\omega$  qubits, respectively, and  $\frac{1}{2}S(C)_\omega$  ebits of entanglement are sufficient and necessary in the entanglement assisted case.

**Remark 4** If in the compression task the parties were required to preserve the correlations with a purifying reference system, then due to Schumacher compression the optimal qubit rate would be  $S(A)_\rho = S(CNQ)_\omega$ . However, Theorem 2 shows that the parties can compress more if they are only required to preserve the correlations with a mixed state reference. This gap can be strictly positive if the redundant system  $N$  is mixed given the classical information  $j$  in system  $C$ , that is  $S(CNQ)_\omega - S(CQ)_\omega = S(N|CQ)_\omega > 0$ .

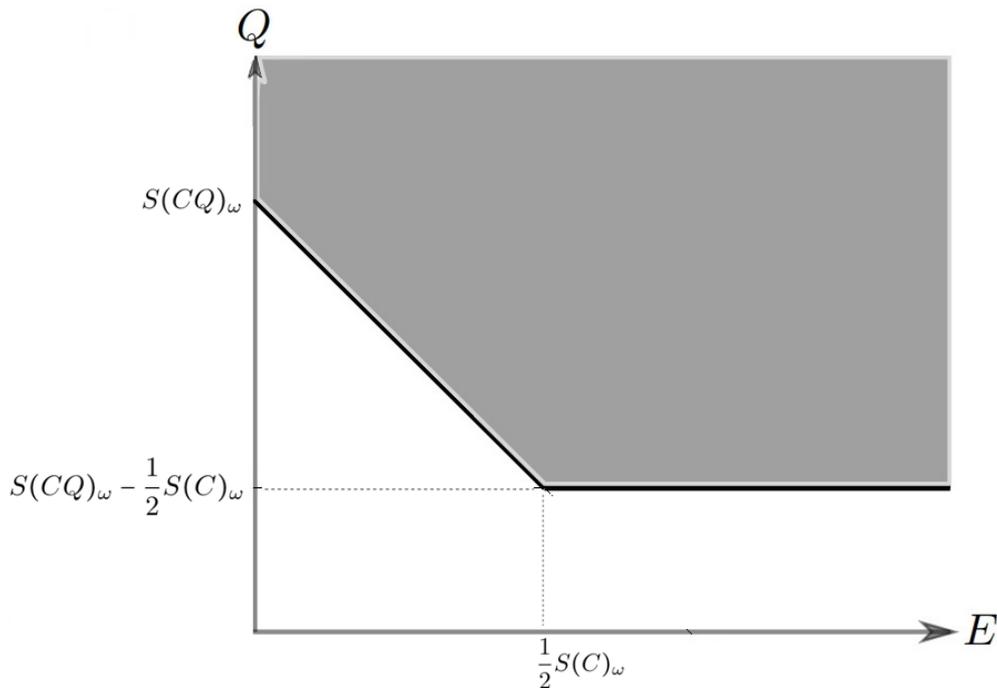


Figure 1. The achievable rate region of the entanglement and quantum rates.

*Proof.* We start with the achievability of these rates. The converse proofs need more tools, so we will leave them to the subsequent sections. Looking at Fig. 1, it will be enough to prove the achievability of the corresponding corner points  $(E, Q) = (0, S(CQ)_\omega)$  and  $(E, Q) = (\frac{1}{2}S(C)_\omega, S(CQ)_\omega - \frac{1}{2}S(C)_\omega)$  for the unassisted and entanglement assisted cases, respectively. This is because by definition (and the time-sharing principle) the rate region is convex and upper-right closed. Indeed, all the points on the line  $Q + E = S(CQ)_\omega$  for  $Q \geq S(CQ)_\omega - \frac{1}{2}S(C)_\omega$  are achievable because one ebit can be distributed by sending a qubit. All other rate pairs are achievable by resource wasting. The rate region is depicted in Fig. 1.

As we discussed, we can assume that the source is  $(\omega^{CNQR})^{\otimes n} = \omega^{C^n N^n Q^n R^n}$ . To achieve the point  $(0, S(CQ)_\omega)$ , Alice traces out the redundant part  $N^n$  of the source, to get the state  $\omega^{C^n Q^n R^n}$  and applies Schumacher compression to send the systems  $C^n Q^n$  to Bob. Since the Schumacher compression preserves the purification of the systems  $C^n Q^n$ , it preserves the state  $\omega^{C^n Q^n R^n}$  as well. To be more specific, let  $\Lambda_S$  denote the composition of the encoding and decoding operations for the Schumacher compression of the state  $|\omega\rangle^{C^n Q^n R^n R'^n}$  where the system  $R'^n$  is a purifying reference system which of course the parties do not have access to. The Schumacher compression preserves the following fidelity on the left member of the equation, therefore it preserves the fidelity on the right member:

$$1 - \epsilon \leq F\left(\omega^{C^n Q^n R^n R'^n}, (\Lambda_S \otimes \text{id}_{R^n R'^n})\omega^{C^n Q^n R^n R'^n}\right) \leq F\left(\omega^{C^n Q^n R^n}, (\Lambda_S \otimes \text{id}_{R^n})\omega^{C^n Q^n R^n}\right),$$

where the inequality is due to monotonicity of the fidelity under partial trace. The rate achieved by this scheme is  $S(CQ)_\omega$ . After applying this scheme, Bob has access to the systems  $\hat{C}^n \hat{Q}^n$ , which is correlated

with the reference system  $R^n$ :

$$\zeta^{\hat{C}^n \hat{Q}^n R^n} = (\Lambda_S \otimes \text{id}_{R^n}) \omega^{C^n Q^n R^n}.$$

Then, to reconstruct the system  $N^n$ , Bob applies the CPTP map  $\mathcal{N} : CQ \rightarrow CNQ$  to each copy, which acts as follows:

$$\mathcal{N}(\rho^{CQ}) = \sum_j (|j\rangle\langle j|^C \otimes \mathbb{1}_Q) \rho^{CQ} (|j\rangle\langle j|^C \otimes \mathbb{1}_Q) \otimes \omega_j^N.$$

This map satisfies the fidelity criterion of Eq. (8) because of monotonicity of the fidelity under CPTP maps:

$$\begin{aligned} 1 - \epsilon &\leq F\left(\omega^{C^n Q^n R^n}, \zeta^{\hat{C}^n \hat{Q}^n R^n}\right) \\ &\leq F\left((\mathcal{N}^{\otimes n} \otimes \text{id}_{R^n}) \omega^{C^n Q^n R^n}, (\mathcal{N}^{\otimes n} \otimes \text{id}_{R^n}) \zeta^{\hat{C}^n \hat{Q}^n R^n}\right) \\ &= F\left(\omega^{C^n N^n Q^n R^n}, \tau^{\hat{C}^n \hat{N}^n \hat{Q}^n R^n}\right). \end{aligned} \quad (8)$$

To achieve the point  $(\frac{1}{2}S(C)_\omega, S(CQ)_\omega - \frac{1}{2}S(C)_\omega)$ , Alice applies dense coding to send the classical system  $C^n$  to Bob which requires  $\frac{n}{2}S(C)_\omega$  ebits of initial entanglement and  $\frac{n}{2}S(C)_\omega$  qubits [15]. When both Alice and Bob have access to system  $C^n$ , Alice can send the quantum system  $Q^n$  to Bob by applying Schumacher compression, which requires sending  $nS(Q|C)$  qubits to Bob. Therefore, the overall qubit rate is  $\frac{1}{2}S(C)_\omega + S(Q|C) = S(CQ)_\omega - \frac{1}{2}S(C)_\omega$ . ■

#### IV. CONVERSE

In this section, we will provide the converse bounds for the qubit rate  $Q$  and the sum rate  $Q + E$  of Theorem 2. We obtain these bounds based on the structure of the CPTP maps which preserve the source state  $\omega^{CNQR}$ . Namely, according to Theorem 1 the CPTP maps acting on systems  $CNQ$ , which preserve the state  $\omega^{CNQR}$ , act only on the redundant system  $N$ . This implies that the environment systems of such CPTP maps are decoupled from systems  $QR$  given the classical information  $j$  in the classical system  $C$ . This gives us an insight into the structure of the encoding-decoding maps, which preserve the overall state *asymptotically* intact.

To proceed with the proof, we first define two functions that emerge in the converse bounds. Then, we state some important properties of these functions in Lemma 6 which we will use to compute the tight asymptotic converse bounds.

**Definition 5** For the KI decomposition  $\omega^{CNQR} = \sum_j p_j |j\rangle\langle j|^C \otimes \omega_j^N \otimes \rho_j^{QR}$  of the state  $\rho^{AR}$  and  $\epsilon \geq 0$ , define

$$\begin{aligned} J_\epsilon(\omega) &:= \max I(\hat{N}E : \hat{C}\hat{Q}|C')_\tau \text{ s.t. } U : CNQ \rightarrow \hat{C}\hat{N}\hat{Q}E \text{ is an isometry with } F(\omega^{CNQR}, \tau^{\hat{C}\hat{N}\hat{Q}R}) \geq 1 - \epsilon, \\ Z_\epsilon(\omega) &:= \max S(\hat{N}E|C')_\tau \text{ s.t. } U : CNQ \rightarrow \hat{C}\hat{N}\hat{Q}E \text{ is an isometry with } F(\omega^{CNQR}, \tau^{\hat{C}\hat{N}\hat{Q}R}) \geq 1 - \epsilon, \end{aligned}$$

where

$$\begin{aligned} \omega^{CNQRC'} &= \sum_j p_j |j\rangle\langle j|^C \otimes \omega_j^N \otimes \rho_j^{QR} \otimes |j\rangle\langle j|^{C'}, \\ \tau^{\hat{C}\hat{N}\hat{Q}ERC'} &= (U \otimes \mathbb{1}_{RC'}) \omega^{CNQRC'} (U^\dagger \otimes \mathbb{1}_{RC'}), \\ \tau^{\hat{C}\hat{N}\hat{Q}R} &= \text{Tr}_{EC'}[\tau^{\hat{C}\hat{N}\hat{Q}ERC'}]. \end{aligned}$$

In this definition, the dimension of the environment is w.l.o.g. bounded as  $|E| \leq (|C||N||Q|)^2$  because the input and output dimensions of the channel are fixed as  $|C||N||Q|$ ; hence, the optimisation is of a continuous function over a compact domain, so we have a maximum rather than a supremum.

**Lemma 6** The functions  $Z_\epsilon(\omega)$  and  $J_\epsilon(\omega)$  have the following properties:

1. They are non-decreasing functions of  $\epsilon$ .

2. They are concave in  $\epsilon$ .

3. They are continuous for  $\epsilon \geq 0$ .

4. For any two states  $\omega_1^{C_1 N_1 Q_1 R_1}$  and  $\omega_2^{C_2 N_2 Q_2 R_2}$  and for  $\epsilon \geq 0$ ,

$$\begin{aligned} J_\epsilon(\omega_1 \otimes \omega_2) &\leq J_\epsilon(\omega_1) + J_\epsilon(\omega_2), \\ Z_\epsilon(\omega_1 \otimes \omega_2) &\leq Z_\epsilon(\omega_1) + Z_\epsilon(\omega_2). \end{aligned}$$

5. At  $\epsilon = 0$ ,  $Z_0(\omega) = S(N|C)_\omega$  and  $J_0(\omega) = 0$ .

The proof of this lemma follows in the next section. Now we show how it is used to prove the converse (optimality) of Theorem 2. As a guide to reading the subsequent proof, we remark that in Eqs. (24) and (28), the environment systems  $VW$  of the encoding-decoding operations appear in the terms  $I(\hat{N}^n VW : \hat{C}^n \hat{Q}^n | C'^m)$  and  $S(\hat{N}^n VW | C'^m)$ , which are bounded by the functions  $J_\epsilon(\omega^{\otimes n})$  and  $Z_\epsilon(\omega^{\otimes n})$ , respectively. As stated in point 4 of Lemma 6, these functions are sub-additive, so basically we can single-letterize the terms appearing in the converse. Moreover, from point 3 of Lemma 6, we know that these functions are continuous for  $\epsilon \geq 0$ ; therefore, the limit points of these functions are equal to the values of these functions at  $\epsilon = 0$ . When the fidelity is equal to 1 ( $\epsilon = 0$ ), the structure of the CPTP maps preserving the state  $\omega^{C^N Q^R}$  in Theorem 1 implies that  $J_0(\omega) = 0$  and  $Z_0(\omega) = S(N|C)_\omega$ , as stated in point 5 of Lemma 6. Thereby, we conclude the converse bounds in Eqs. (27) and (31).

*Proof of Theorem 2 (converse).* We first get the following chain of inequalities considering the process of the decoding of the information:

$$nQ + S(B_0) \geq S(M) + S(B_0) \tag{9}$$

$$\geq S(MB_0) \tag{10}$$

$$= S(\hat{C}^n \hat{N}^n \hat{Q}^n V) \tag{11}$$

$$= S(\hat{C}^n \hat{Q}^n) + S(\hat{N}^n V | \hat{C}^n \hat{Q}^n) \tag{12}$$

$$\geq nS(CQ) + S(\hat{N}^n V | \hat{C}^n \hat{Q}^n) - n\delta(n, \epsilon) \tag{13}$$

$$\geq nS(CQ) + S(\hat{N}^n V | \hat{C}^n \hat{Q}^n C'^m) - n\delta(n, \epsilon) \tag{14}$$

$$= nS(CQ) + S(\hat{N}^n V | \hat{C}^n \hat{Q}^n C'^m) - S(\hat{N}^n V | C'^m) + S(\hat{N}^n V | C'^m) - n\delta(n, \epsilon)$$

$$= nS(CQ) - I(\hat{N}^n V : \hat{C}^n \hat{Q}^n | C'^m) + S(\hat{N}^n V | C'^m) - n\delta(n, \epsilon)$$

$$\geq nS(CQ) - I(\hat{N}^n VW : \hat{C}^n \hat{Q}^n | C'^m) + S(\hat{N}^n V | C'^m) - n\delta(n, \epsilon) \tag{15}$$

where Eq. (9) follows because the entropy of a system is bounded by the logarithm of the dimension of that system; Eq. (10) is due to sub-additivity of the entropy; Eq. (11) follows because the decoding isometry  $U_{\mathcal{D}} : MB_0 \rightarrow \hat{C}^n \hat{N}^n \hat{Q}^n V$  does not change the entropy; Eq. (12) is due to the chain rule; Eq. (13) follows from the decodability: the output state on systems  $\hat{C}^n \hat{Q}^n$  is  $2\sqrt{2\epsilon}$ -close to the original state  $C^n Q^n$  in trace norm; then the inequality follows by applying the Fannes-Audenaert inequality [16, 17], where  $\delta(n, \epsilon) = \sqrt{2\epsilon} \log(|C||Q|) + \frac{1}{n} h(\sqrt{2\epsilon})$ ; Eq. (14) is due to strong sub-additivity of the entropy, and system  $C'$  is a copy of classical system  $C$ ; Eq. (15) follows from data processing inequality where  $W$  is the environment system of the encoding isometry  $U_{\mathcal{E}} : C^n N^n Q^n A_0 \rightarrow MW$ .

Moreover, considering the process of encoding the information,  $Q$  is bounded as follows:

$$\begin{aligned} nQ &\geq S(M) \\ &\geq S(M|WC'^m) \end{aligned} \quad (16)$$

$$= S(MWC'^m) - S(WC'^m) \quad (17)$$

$$= S(C^n N^n Q^n A_0 C'^m) - S(WC'^m) \quad (18)$$

$$= S(C^n N^n Q^n C'^m) + S(A_0) - S(WC'^m) \quad (19)$$

$$= S(C^n N^n Q^n C'^m) + S(A_0) - S(C'^m) - S(W|C'^m) \quad (20)$$

$$= S(C^n N^n Q^n) + S(A_0) - S(C'^m) - S(W|C'^m) \quad (21)$$

$$= nS(CQ) + nS(N|CQ) + S(A_0) - nS(C') - S(W|C'^m) \quad (22)$$

$$= nS(CQ) + nS(N|C) + S(A_0) - nS(C') - S(W|C'^m), \quad (23)$$

where Eq. (16) is due to sub-additivity of the entropy; Eq. (17) is due to the chain rule; Eq. (18) follows because the encoding isometry  $U_{\mathcal{E}} : C^n N^n Q^n A_0 \leftrightarrow MW$  does not change the entropy; Eq. (19) follows because the initial entanglement  $A_0$  is independent from the source; Eq. (20) is due to the chain rule; Eq. (21) follows because  $C'$  is a copy of the system  $C$ , so  $S(C'|CNQ) = 0$ ; Eq. (22) is due to the chain rule and the fact that the entropy is additive for product states; Eq. (23) follows because conditional on system  $C$  the system  $N$  is independent from system  $Q$ .

Now, we add Eqs. (15) and (23); the entanglement terms  $S(A_0)$  and  $S(B_0)$  cancel out, and by dividing by  $2n$  we obtain

$$\begin{aligned} Q &\geq S(CQ) - \frac{1}{2}S(C) + \frac{1}{2}S(N|C) - \frac{1}{2n}I(\hat{N}^n VW : \hat{C}^n \hat{Q}^n | C'^m) + \frac{1}{2n}S(\hat{N}^n V | C'^m) - \frac{1}{2n}S(W | C'^m) - \frac{1}{2}\delta(n, \epsilon) \\ &\geq S(CQ) - \frac{1}{2}S(C) + \frac{1}{2}S(N|C) - \frac{1}{2n}I(\hat{N}^n VW : \hat{C}^n \hat{Q}^n | C'^m) - \frac{1}{2n}S(\hat{N}^n VW | C'^m) - \frac{1}{2}\delta(n, \epsilon) \end{aligned} \quad (24)$$

$$\geq S(CQ) - \frac{1}{2}S(C) + \frac{1}{2}S(N|C) - \frac{1}{2n}J_{\epsilon}(\omega^{\otimes n}) - \frac{1}{2n}Z_{\epsilon}(\omega^{\otimes n}) - \frac{1}{2}\delta(n, \epsilon) \quad (25)$$

$$\geq S(CQ) - \frac{1}{2}S(C) + \frac{1}{2}S(N|C) - \frac{1}{2}J_{\epsilon}(\omega) - \frac{1}{2}Z_{\epsilon}(\omega) - \frac{1}{2}\delta(n, \epsilon), \quad (26)$$

where Eq. (24) follows from strong sub-additivity of the entropy,  $S(\hat{N}^n V | C'^m) + S(\hat{N}^n V | WC'^m) \geq 0$ ; Eq. (25) follows from Definition 5; Eq. (26) is due to point 4 of Lemma 6.

In the limit of  $\epsilon \rightarrow 0$  and  $n \rightarrow \infty$ , the qubit rate is thus bounded by

$$\begin{aligned} Q &\geq S(CQ) - \frac{1}{2}S(C) + \frac{1}{2}S(N|C) - \frac{1}{2}J_0(\omega) - \frac{1}{2}Z_0(\omega) \\ &= S(CQ) - \frac{1}{2}S(C), \end{aligned} \quad (27)$$

where the equality follows from point 5 of Lemma 6.

Moreover, from Eq. (15) we have:

$$\begin{aligned} nQ + S(B_0) &= nQ + nE \\ &\geq nS(CQ) - I(\hat{N}^n VW : \hat{C}^n \hat{Q}^n | C'^m) + S(\hat{N}^n V | C'^m) - n\delta(n, \epsilon) \\ &\geq nS(CQ) - I(\hat{N}^n VW : \hat{C}^n \hat{Q}^n | C'^m) - n\delta(n, \epsilon) \end{aligned} \quad (28)$$

$$\geq nS(CQ) - J_{\epsilon}(\omega^{\otimes n}) - n\delta(n, \epsilon) \quad (29)$$

$$\geq nS(CQ) - nJ_{\epsilon}(\omega) - n\delta(n, \epsilon), \quad (30)$$

where Eq. (28) follows because the entropy conditional on a classical system is positive,  $S(\hat{N}^n V | C'^m) \geq 0$ ; Eq. (29) follows from Definition 5; Eq. (30) is due to point 4 of Lemma 6.

In the limit of  $\epsilon \rightarrow 0$  and  $n \rightarrow \infty$ , we thus obtain the following bound on the rate sum:

$$Q + E \geq S(CQ) - J_0(\omega) = S(CQ), \quad (31)$$

where the equality follows from point 5 of Lemma 6.  $\blacksquare$

**Remark 7** Our lower bound on  $Q + E$  in Eq. (31) reproduces the result of Koashi and Imoto [7] for the case of a classical-quantum source  $\rho^{AX} = \sum_x p(x) \rho_x^A \otimes |x\rangle\langle x|^X$ . This is because a code with qubit-ebit rate pair  $(Q, E)$  gives rise to a compression code in the sense of Koashi and Imoto using a rate of qubits  $Q + E$  and no prior entanglement, simply by first distributing  $E$  ebits and then using the entanglement assisted code.

It is worth noting that conversely, Eq. (31) can be obtained from the Koashi-Imoto result, as follows. Any good code for  $\rho^{AR}$  is automatically a good code for the classical-quantum source of mixed states

$$\rho^{AY} = \sum_y q(y) \rho_y^A \otimes |y\rangle\langle y|^Y = \sum_y \text{Tr}_R \rho^{AR} (\mathbb{1}_A \otimes M_y^R) \otimes |y\rangle\langle y|^Y,$$

for any POVM  $(M_y)$  on  $R$ , simply by the monotonicity of the fidelity under CPTP maps. As discussed before, by choosing an informationally complete measurement, the KI-decomposition of the ensemble  $\{q(y), \rho_y^A\}$  is identical to that of  $\rho^{AR}$  in Theorem 1. Thus the unassisted qubit compression rate of  $\rho^{AY}$  and of  $\rho^{AR}$  are lower bounded by the same quantity, the right hand side of Eq. (31).

## V. PROOF OF LEMMA 6

1. The definitions of the functions  $J_\epsilon(\omega)$  and  $Z_\epsilon(\omega)$  directly imply that they are non-decreasing functions of  $\epsilon$ .
2. We first prove the concavity of  $Z_\epsilon(\omega)$ . Let  $U_1 : CNQ \mapsto \hat{C}\hat{N}\hat{Q}E$  and  $U_2 : CNQ \mapsto \hat{C}\hat{N}\hat{Q}E$  be the isometries attaining the maximum for  $\epsilon_1$  and  $\epsilon_2$ , respectively, which act as follows on the purification  $|\omega\rangle^{CNQRC'R'}$  of the previously introduced state  $\omega^{CNQRC'}$ :

$$|\tau_1\rangle^{\hat{C}\hat{N}\hat{Q}ERC'R'} = (U_1 \otimes \mathbb{1}_{RC'R'})|\omega\rangle^{CNQRC'R'} \quad \text{and} \quad |\tau_2\rangle^{\hat{C}\hat{N}\hat{Q}ERC'R'} = (U_2 \otimes \mathbb{1}_{RC'R'})|\omega\rangle^{CNQRC'R'},$$

where  $\text{Tr}_{R'}[|\omega\rangle\langle\omega|^{CNQRC'R'}] = \omega^{CNQRC'}$ . For  $0 \leq \lambda \leq 1$ , define the isometry  $U_0 : CNQ \mapsto \hat{C}\hat{N}\hat{Q}EFF'$  which acts as

$$U_0 := \sqrt{\lambda}U_1 \otimes |11\rangle^{FF'} + \sqrt{1-\lambda}U_2 \otimes |22\rangle^{FF'}, \quad (32)$$

where systems  $F$  and  $F'$  are qubits, and which leads to the state

$$(U_0 \otimes \mathbb{1}_{RC'R'})|\omega\rangle^{CNQRC'R'} = \sqrt{\lambda}|\tau_1\rangle^{\hat{C}\hat{N}\hat{Q}ERC'R'}|11\rangle^{FF'} + \sqrt{1-\lambda}|\tau_2\rangle^{\hat{C}\hat{N}\hat{Q}ERC'R'}|22\rangle^{FF'}. \quad (33)$$

Then,  $U_0$  defines its state  $\tau$ . for which the reduced state on the systems  $\hat{C}\hat{N}\hat{Q}RC'$  is

$$\tau^{\hat{C}\hat{N}\hat{Q}RC'} = \lambda\tau_1^{\hat{C}\hat{N}\hat{Q}RC'} + (1-\lambda)\tau_2^{\hat{C}\hat{N}\hat{Q}RC'}. \quad (34)$$

Therefore, the fidelity for the state  $\tau$  is bounded as follows:

$$\begin{aligned} F(\omega^{CNQR}, \tau^{\hat{C}\hat{N}\hat{Q}R}) &= F(\omega^{CNQR}, \lambda\tau_1^{\hat{C}\hat{N}\hat{Q}R} + (1-\lambda)\tau_2^{\hat{C}\hat{N}\hat{Q}R}) \\ &= F(\lambda\omega^{CNQR} + (1-\lambda)\omega^{CNQR}, \lambda\tau_1^{\hat{C}\hat{N}\hat{Q}R} + (1-\lambda)\tau_2^{\hat{C}\hat{N}\hat{Q}R}) \\ &\geq \lambda F(\omega^{CNQR}, \tau_1^{\hat{C}\hat{N}\hat{Q}R}) + (1-\lambda)F(\omega^{CNQR}, \tau_2^{\hat{C}\hat{N}\hat{Q}R}) \\ &\geq 1 - (\lambda\epsilon_1 + (1-\lambda)\epsilon_2). \end{aligned} \quad (35)$$

The first inequality is due to simultaneous concavity of the fidelity in both arguments; the last line follows by the definition of the isometries  $U_1$  and  $U_2$ . Thus, the isometry  $U_0$  yields a fidelity of at least  $1 - (\lambda\epsilon_1 + (1-\lambda)\epsilon_2) =: 1 - \epsilon$ . Now let  $E' = EFF'$  denote the environment of the isometry  $U_0$  defined above. According to Definition 5, we obtain

$$\begin{aligned} Z_\epsilon(\omega) &\geq S(\hat{N}E'|C')_\tau \\ &= S(\hat{N}EFF'|C')_\tau \\ &= S(F|C')_\tau + S(\hat{N}E|FC')_\tau + S(F'|\hat{N}EFC')_\tau \end{aligned} \quad (36)$$

$$\geq S(\hat{N}E|FC')_\tau \quad (37)$$

$$= \lambda S(\hat{N}E|C')_{\tau_1} + (1-\lambda)S(\hat{N}E|C')_{\tau_2} \quad (38)$$

$$= \lambda Z_{\epsilon_1}(\omega) + (1-\lambda)Z_{\epsilon_2}(\omega), \quad (39)$$

where the state  $\tau$  in the entropies is given in Eq. (34); Eq. (36) is due to the chain rule; Eq. (37) follow because for the state on systems  $\hat{N}EFFC'$  we have  $S(F'|C') + S(F'|\hat{N}EFC') \geq 0$  which follows from strong sub-additivity of the entropy; Eq. (38) follows by expanding the conditional entropy on the classical system  $F$ ; Eq. (39) follows from the definitions of the isometries  $U_1$  and  $U_2$ .

Moreover, let  $U_1 : CNQ \leftrightarrow \hat{C}\hat{N}\hat{Q}E$  and  $U_2 : CNQ \leftrightarrow \hat{C}\hat{N}\hat{Q}E$  be the isometries attaining the maximum for  $\epsilon_1$  and  $\epsilon_2$  in the definition of  $J_\epsilon(\omega)$ , respectively. Again, define the isometry  $U_0$  as in Eq. (32), which leads to the bound on the fidelity as in Eq. (35), letting  $E' = EFF'$  be the environment of the isometry  $U_0$ . According to Definition 5, we obtain

$$\begin{aligned} J_\epsilon(\omega) &\geq I(\hat{N}EFF' : \hat{C}\hat{Q}|C')_\tau \\ &\geq I(\hat{N}EF : \hat{C}\hat{Q}|C')_\tau \end{aligned} \quad (40)$$

$$= I(F : \hat{C}\hat{Q}|C')_\tau + I(\hat{N}E : \hat{C}\hat{Q}|FC')_\tau \quad (41)$$

$$\geq I(\hat{N}E : \hat{C}\hat{Q}|FC')_\tau \quad (42)$$

$$= \lambda I(\hat{N}E : \hat{C}\hat{Q}|C')_{\tau_1} + (1 - \lambda) I(\hat{N}E : \hat{C}\hat{Q}|C')_{\tau_2} \quad (43)$$

$$= \lambda J_{\epsilon_1}(\omega) + (1 - \lambda) J_{\epsilon_2}(\omega), \quad (44)$$

where Eq. (40) follows from data processing; Eq. (41) is due to the chain rule for mutual information; Eq. (42) follows from strong sub-additivity of the entropy,  $I(F : \hat{C}\hat{Q}|C')_\tau \geq 0$ ; Eq. (43) is obtained by expanding the conditional mutual information on the classical system  $F$ ; finally, Eq. (44) follows from the definitions of the isometries  $U_1$  and  $U_2$ .

3. The functions are non-decreasing and concave for  $\epsilon \geq 0$ , so they are continuous for  $\epsilon > 0$ . The concavity implies furthermore that  $J_\epsilon$  and  $Z_\epsilon$  are lower semi-continuous at  $\epsilon = 0$ . On the other hand, since the fidelity, the conditional entropy and the conditional mutual information are all continuous functions of CPTP maps, and the domain of both optimizations is a compact set, we conclude that  $J_\epsilon(\omega)$  and  $Z_\epsilon$  are also upper semi-continuous at  $\epsilon = 0$ , so they are continuous at  $\epsilon = 0$  [18, Thms. 10.1 and 10.2].
4. We first prove  $Z_\epsilon(\omega_1 \otimes \omega_2) \leq Z_\epsilon(\omega_1) + Z_\epsilon(\omega_2)$ . In the definition of  $Z_\epsilon(\omega_1 \otimes \omega_2)$ , let the isometry  $U_0 : C_1N_1Q_1C_2N_2Q_2 \leftrightarrow \hat{C}_1\hat{N}_1\hat{Q}_1\hat{C}_2\hat{N}_2\hat{Q}_2E$  be the one attaining the maximum, which acts on the following purified source states with purifying systems  $R'_1$  and  $R'_2$ :

$$|\tau\rangle^{\hat{C}_1\hat{N}_1\hat{Q}_1\hat{C}_2\hat{N}_2\hat{Q}_2ER_1C'_1R'_1R_2C'_2R'_2} = (U_0 \otimes \mathbb{1}_{R_1C'_1R'_1R_2C'_2R'_2})|\omega_1\rangle^{C_1N_1Q_1R_1C'_1R'_1} \otimes |\omega_2\rangle^{C_2N_2Q_2R_2C'_2R'_2}. \quad (45)$$

By definition, the fidelity is bounded by

$$F(\omega_1^{C_1N_1Q_1R_1} \otimes \omega_2^{C_2N_2Q_2R_2}, \tau^{\hat{C}_1\hat{N}_1\hat{Q}_1\hat{C}_2\hat{N}_2\hat{Q}_2R_1R_2}) \geq 1 - \epsilon.$$

Now, we can define an isometry  $U_1 : C_1N_1Q_1 \leftrightarrow \hat{C}_1\hat{N}_1\hat{Q}_1E_1$  acting only on systems  $C_1N_1Q_1$ , by letting  $U_1 = (U_0 \otimes \mathbb{1}_{R_2C'_2R'_2})(\mathbb{1}_{C_1N_1Q_1} \otimes |\omega_2\rangle^{C_2N_2Q_2R_2C'_2R'_2})$  and with the environment  $E_1 := \hat{C}_2\hat{N}_2\hat{Q}_2ER_2C'_2R'_2$ . It has the property that  $|\tau\rangle^{\hat{C}_1\hat{N}_1\hat{Q}_1R_1C'_1R'_1E} = (U_1 \otimes \mathbb{1}_{R_1C'_1R'_1})|\omega_1\rangle^{C_1N_1Q_1R_1C'_1R'_1}$  has the same reduced state on  $\hat{C}_1\hat{N}_1\hat{Q}_1R_1$  as  $\tau$  from Eq. (45). This isometry preserves the fidelity for  $\omega_1$ , which follows from monotonicity of the fidelity under partial trace:

$$\begin{aligned} F(\omega_1^{C_1N_1Q_1R_1}, \tau_1^{\hat{C}_1\hat{N}_1\hat{Q}_1R_1}) &= F(\omega_1^{C_1N_1Q_1R_1}, \tau^{\hat{C}_1\hat{N}_1\hat{Q}_1R_1}) \\ &\geq F(\omega_1^{C_1N_1Q_1R_1} \otimes \omega_2^{C_2N_2Q_2R_2}, \tau^{\hat{C}_1\hat{N}_1\hat{Q}_1\hat{C}_2\hat{N}_2\hat{Q}_2R_1R_2}) \\ &\geq 1 - \epsilon. \end{aligned}$$

By the same argument, there is an isometry  $U_2 : C_2N_2Q_2 \leftrightarrow \hat{C}_1\hat{N}_1\hat{Q}_1\hat{C}_2\hat{N}_2\hat{Q}_2ER_1C'_1R'_1$  with output system  $\hat{C}_2\hat{N}_2\hat{Q}_2$  and environment  $E_2 := \hat{C}_1\hat{N}_1\hat{Q}_1ER_1C'_1R'_1$ , such that

$$\begin{aligned} F(\omega_2^{C_2N_2Q_2R_2}, \tau_2^{\hat{C}_2\hat{N}_2\hat{Q}_2R_2}) &= F(\omega_2^{C_2N_2Q_2R_2}, \tau^{\hat{C}_2\hat{N}_2\hat{Q}_2R_2}) \\ &\geq F(\omega_1^{C_1N_1Q_1R_1} \otimes \omega_2^{C_2N_2Q_2R_2}, \tau^{\hat{C}_1\hat{N}_1\hat{Q}_1\hat{C}_2\hat{N}_2\hat{Q}_2R_1R_2}) \\ &\geq 1 - \epsilon. \end{aligned}$$

Therefore, we obtain:

$$\begin{aligned} Z_\epsilon(\omega_1) + Z_\epsilon(\omega_2) - Z_\epsilon(\omega_1 \otimes \omega_2) \\ \geq S(\hat{N}_1 E_1 | C'_1)_\tau + S(\hat{N}_2 E_2 | C'_2)_\tau - S(\hat{N}_1 \hat{N}_2 E | C'_1 C'_2)_\tau \end{aligned} \quad (46)$$

$$= S(\hat{N}_1 E_1 C'_1)_\tau + S(\hat{N}_2 E_2 C'_2)_\tau - S(\hat{N}_1 \hat{N}_2 E C'_1 C'_2)_\tau - S(C'_1) - S(C'_2) + S(C'_1 C'_2) \quad (47)$$

$$= S(\hat{N}_1 E_1 C'_1)_\tau + S(\hat{N}_2 E_2 C'_2)_\tau - S(\hat{N}_1 \hat{N}_2 E C'_1 C'_2)_\tau \quad (48)$$

$$= S(\hat{C}_1 \hat{Q}_1 R_1 R'_1) + S(\hat{C}_2 \hat{Q}_2 R_2 R'_2) - S(\hat{C}_1 \hat{Q}_1 \hat{C}_2 \hat{Q}_2 R_1 R'_1 R_2 R'_2) \quad (49)$$

$$\begin{aligned} &= I(\hat{C}_1 \hat{Q}_1 R_1 R'_1 : \hat{C}_2 \hat{Q}_2 R_2 R'_2) \\ &\geq 0, \end{aligned} \quad (50)$$

where Eq. (46) is due to Definition 5; Eq. (47) is due to the chain rule; Eq. (48) because the systems  $C'_1$  and  $C'_2$  are independent from each other; Eq. (49) follows because the overall state on systems  $\hat{C}_1 \hat{N}_1 \hat{Q}_1 \hat{C}_2 \hat{N}_2 \hat{Q}_2 E R_1 C'_1 R'_1 R_2 C'_2 R'_2$  is pure; Eq. (50) is due to sub-additivity of the entropy.

To prove  $J_\epsilon(\omega_1 \otimes \omega_2) \leq J_\epsilon(\omega_1) + J_\epsilon(\omega_2)$ , let the isometry  $U_0 : C_1 N_1 Q_1 C_2 N_2 Q_2 \mapsto \hat{C}_1 \hat{N}_1 \hat{Q}_1 \hat{C}_2 \hat{N}_2 \hat{Q}_2 E$  be the one attaining the maximum in definition of  $J_\epsilon(\omega_1 \otimes \omega_2)$ , which acts on the following purified source states with purifying systems  $R'_1$  and  $R'_2$ , as in Eq. (45). By definition, the fidelity is bounded as

$$F(\omega_1^{C_1 N_1 Q_1 R_1} \otimes \omega_2^{C_2 N_2 Q_2 R_2}, \tau^{\hat{C}_1 \hat{N}_1 \hat{Q}_1 \hat{C}_2 \hat{N}_2 \hat{Q}_2 R_1 R_2}) \geq 1 - \epsilon.$$

Now define  $U_1 : C_1 N_1 Q_1 \mapsto \hat{C}_1 \hat{N}_1 \hat{Q}_1 \hat{C}_2 \hat{N}_2 \hat{Q}_2 E R_2 C'_2 R'_2$  and  $U_2 : C_2 N_2 Q_2 \mapsto \hat{C}_1 \hat{N}_1 \hat{Q}_1 \hat{C}_2 \hat{N}_2 \hat{Q}_2 E R_1 C'_1 R'_1$  as in the above discussion, with the environments  $E_1 := \hat{C}_2 \hat{N}_2 \hat{Q}_2 E R_2 C'_2 R'_2$  and  $E_2 := \hat{C}_1 \hat{N}_1 \hat{Q}_1 E R_1 C'_1 R'_1$ , respectively. Recall that the fidelity for the states  $\omega_1$  and  $\omega_2$  is at least  $1 - \epsilon$ , because of the monotonicity of the fidelity under partial trace. Thus we obtain

$$\begin{aligned} J_\epsilon(\omega_1) + J_\epsilon(\omega_2) - J_\epsilon(\omega_1 \otimes \omega_2) \\ \geq I(\hat{N}_1 E_1 : \hat{C}_1 \hat{Q}_1 | C'_1)_\tau + I(\hat{N}_2 E_2 : \hat{C}_2 \hat{Q}_2 | C'_2)_\tau - I(\hat{N}_1 \hat{N}_2 E : \hat{C}_1 \hat{Q}_1 \hat{C}_2 \hat{Q}_2 | C'_1 C'_2)_\tau \end{aligned} \quad (51)$$

$$\begin{aligned} &= S(\hat{N}_1 E_1 C'_1) + S(\hat{C}_1 \hat{Q}_1 C'_1) - S(\hat{C}_1 \hat{N}_1 \hat{Q}_1 E_1 C'_1) - S(C'_1) \\ &\quad + S(\hat{N}_2 E_2 C'_2) + S(\hat{C}_2 \hat{Q}_2 C'_2) - S(\hat{C}_2 \hat{N}_2 \hat{Q}_2 E_2 C'_2) - S(C'_2) \\ &\quad - S(\hat{N}_1 \hat{N}_2 E C'_1 C'_2) - S(\hat{C}_1 \hat{Q}_1 \hat{C}_2 \hat{Q}_2 C'_1 C'_2) + S(\hat{C}_1 \hat{N}_1 \hat{Q}_1 \hat{C}_2 \hat{N}_2 \hat{Q}_2 E C'_1 C'_2) + S(C'_1 C'_2) \end{aligned} \quad (52)$$

$$\begin{aligned} &= S(\hat{C}_1 \hat{Q}_1 R_1 R'_1) + S(\hat{C}_1 \hat{Q}_1 C'_1) - S(R_1 R'_1) - S(C'_1) \\ &\quad + S(\hat{C}_2 \hat{Q}_2 R_2 R'_2) + S(\hat{C}_2 \hat{Q}_2 C'_2) - S(R_2 R'_2) - S(C'_2) \\ &\quad - S(\hat{C}_1 \hat{Q}_1 \hat{C}_2 \hat{Q}_2 R_1 R'_1 R_2 R'_2) - S(\hat{C}_1 \hat{Q}_1 \hat{C}_2 \hat{Q}_2 C'_1 C'_2) + S(R_1 R'_1 R_2 R'_2) + S(C'_1 C'_2) \end{aligned} \quad (53)$$

$$\begin{aligned} &= I(\hat{C}_1 \hat{Q}_1 R_1 R'_1 : \hat{C}_2 \hat{Q}_2 R_2 R'_2) - I(R_1 R'_1 : R_2 R'_2) + I(\hat{C}_1 \hat{Q}_1 C'_1 : \hat{C}_2 \hat{Q}_2 C'_2) - I(C'_1 : C'_2) \\ &\geq I(R_1 R'_1 : R_2 R'_2) - I(R_1 R'_1 : R_2 R'_2) + I(C'_1 : C'_2) - I(C'_1 : C'_2) \\ &= 0, \end{aligned} \quad (54)$$

where Eq. (51) is due to Definition 5; In Eq. (52) we expand the mutual informations in terms of entropies; Eq. (53) follows because the overall state on systems  $\hat{C}_1 \hat{N}_1 \hat{Q}_1 \hat{C}_2 \hat{N}_2 \hat{Q}_2 E R_1 C'_1 R'_1 R_2 C'_2 R'_2$  is pure; Eq. (54) is due to data processing.

5. According to Theorem 1 [10, 11], any isometry  $U : CNQ \rightarrow \hat{C} \hat{N} \hat{Q} E$  acting on the state  $\omega^{CNQRC'}$  which preserves the reduced state on systems  $CNQR C'$  ( $C'$  here is considered as a part of the reference system), acts as the following:

$$(U \otimes \mathbb{1}_{RC'}) \omega^{CNQRC'} (U^\dagger \otimes \mathbb{1}_{RC'}) = \sum_j p_j |j\rangle\langle j|^C \otimes U_j \omega_j^N U_j^\dagger \otimes \rho_j^{QR} \otimes |j\rangle\langle j|^{C'},$$

where the isometry  $U_j : N \rightarrow \hat{N} E$  satisfies  $\text{Tr}_E[U_j \omega_j^N U_j^\dagger] = \omega_j$ . Therefore, in Definition 5 for  $\epsilon = 0$ , the final state is

$$\tau^{\hat{C} \hat{N} \hat{Q} E R C'} = \sum_j p_j |j\rangle\langle j|^C \otimes U_j \omega_j^N U_j^\dagger \otimes \rho_j^{QR} \otimes |j\rangle\langle j|^{C'}.$$

Thus we can directly evaluate

$$Z_0(\omega) = S(\hat{N}E|C')_\tau = S(N|C)_\omega \text{ and } J_0(\omega) = I(\hat{N}E : \hat{C}\hat{Q}|C')_\tau = 0,$$

concluding the proof.  $\blacksquare$

## VI. DISCUSSION

We have introduced a common framework for all single-source quantum compression problems, i.e. settings without side information at the encoder or the decoder, by defining the compression task as the reproduction of a given bipartite state between the system to be compressed and a reference. That state, which defines the task, can be completely general, and special instances recover Schumacher's quantum source compression (in both variants of a pure state ensemble and of a pure entangled state) [1] and compression of a mixed state ensemble source in the blind variant [5, 7].

Our general result gives the optimal quantum compression rate in terms of qubits per source, both in the settings without and with entanglement, and indeed the entire qubit-ebit rate region, reproducing the aforementioned special cases, along with other previously considered problems [12]. Despite the technical difficulties in obtaining it, the end result has a simple and intuitive interpretation. Namely, the given source  $\rho^{AR}$  is equivalent to a source in standard Koashi-Imoto form,

$$\omega^{CQR} = \sum_j p_j |j\rangle\langle j|^C \otimes \rho_j^{QR},$$

so that  $j$  has to be compressed as classical information, at rate  $S(C)$ , and  $Q$  as quantum information, at rate  $S(Q|C)$ ; in the presence of entanglement, the former rate is halved while the latter is maintained. Indeed, what our Theorem 2 shows is that the original source has the same qubit-ebit rate region as the clean classical-quantum mixed source

$$\Omega^{CQRR'C'} = \sum_j p_j |j\rangle\langle j|^C \otimes |\psi_j\rangle\langle\psi_j|^{QRR'} \otimes |j\rangle\langle j|^{C'},$$

where  $|\psi_j\rangle^{QRR'}$  purifies  $\rho_j^{QR}$ , and  $RR'C'$  is considered the reference. In  $\Omega$ ,  $C$  is indeed a manifestly classical source, since it is duplicated in the reference system, and conditional on  $C$ ,  $Q$  is a genuinely quantum source since it is purely entangled with the reference system. As  $\text{Tr}_{R'C'} \Omega^{CQRR'C'} = \omega^{CQR}$ , any code and any achievable rates for  $\Omega$  are good for  $\omega$ , and that is how the achievability of the rate region in Theorem 2 can be described. The opposite, that a code good for  $\omega$  should be good for  $\Omega$ , is far from obvious. Indeed, if that were true, it would not only yield a quick and simple proof of our converse bounds, but would imply that the rate region of Theorem 2 satisfies a strong converse! However, as we do not know this reduction to the source  $\Omega$ , our converse proceeds via a more complicated, indirect route, and yields only a weak converse. Whether the strong converse holds, and what the detailed relation between the sources  $\omega^{CQR}$  and  $\Omega^{CQRR'C'}$  is, remain open questions.

As we were finishing the write-up of the present paper, we became aware of related work by Anshu *et al.* [19], who consider a source consisting of a commuting mixed state ensemble, with the aim of showing a large separation between the Holevo information of the ensemble and the actual (blind) compression rate of the ensemble, even at non-zero error. Unlike our work, which follows source coding convention by considering block error, they define the error as "error (infidelity) per letter", which is a weaker requirement, and prove a rate lower bound in their [19, Thm. 2]. It is worth noting that first, our lower bounds in Theorem 2 only require the error per letter criterion, and that indeed Eqs. (26) and (30) give rate lower bounds for asymptotically large  $n$  and non-zero error  $\epsilon$ , which in addition in the limit  $\epsilon \rightarrow 0$  become tight:

$$\begin{aligned} Q &\geq S(CQ) - \frac{1}{2}S(C) - \frac{1}{2}J_\epsilon(\omega) - \frac{1}{2}(Z_\epsilon(\omega) - S(N|C)) - \frac{1}{2}\delta(\epsilon), \\ Q + E &\geq S(CQ) - J_\epsilon(\omega) - \delta(\epsilon), \end{aligned}$$

where  $J_\epsilon$  and  $Z_\epsilon$  are as in Definition 5, and  $\delta(\epsilon) = \sqrt{2\epsilon} \log(|C||Q|)$ .

**Acknowledgments.** The authors were supported by the Spanish MINECO (projects FIS2016-80681-P, FISICATEAMO FIS2016-79508-P and SEVERO OCHOA No. SEV-2015-0522, FPI) with the support of FEDER funds, the Generalitat de Catalunya (projects 2017-SGR-1127, 2017-SGR-1341 and CERCA/Program), ERC AdG OSYRIS, EU FETPRO QUIC, and the National Science Centre, Poland-Symfonia grant no. 2016/20/W/ST4/00314.

- 
- [1] B. Schumacher, “Quantum coding,” *Phys. Rev. A*, vol. 51, no. 4, pp. 2738-2747, Apr 1995.
  - [2] R. Jozsa and B. Schumacher, “A new proof of the quantum noiseless coding theorem,” *J. Mod. Optics*, vol. 41, no. 12, pp. 2343-2349, 1994.
  - [3] H. Barnum, C. A. Fuchs, R. Jozsa, and B. Schumacher, “General fidelity limit for quantum channels,” *Phys. Rev. A*, vol. 54, no. 6, pp. 4707-4711, Dec 1996.
  - [4] A. Winter, *Coding Theorems of Quantum Information Theory*, Ph.D. dissertation, Universität Bielefeld, Department of Mathematics, Germany, July 1999, arXiv:quant-ph/9907077.
  - [5] M. Horodecki, “Limits for compression of quantum information carried by ensembles of mixed states,” *Phys. Rev. A*, vol. 57, no. 5, pp. 3364-3369, May 1998.
  - [6] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa and B. Schumacher, “On quantum coding for ensembles of mixed states,” *J. Phys. A: Math. Gen.*, vol. 34, no. 35, pp. 6767-6785, 2001.
  - [7] M. Koashi and N. Imoto, “Compressibility of Quantum Mixed-State Signals,” *Phys. Rev. Lett.*, vol. 87, no. 1, 017902, July 2001.
  - [8] B. Schumacher, “Sending entanglement through noisy quantum channels,” *Phys. Rev. A*, vol. 54, no. 4, pp. 2614-2628, 1996.
  - [9] H. Barnum, M. A. Nielsen and B. W. Schumacher, “Information transmission through a noisy quantum channel,” *Phys. Rev. A*, vol. 57, no. 6, pp. 4153-4175, June 1998.
  - [10] P. Hayden, R. Jozsa, D. Petz and A. Winter, “Structure of States Which Satisfy Strong Subadditivity of Quantum Entropy with Equality,” *Commun. Math. Phys.*, vol. 246, no. 2, pp. 359-374, Apr 2004.
  - [11] M. Koashi and N. Imoto, “Operations that do not disturb partially known quantum states,” *Phys. Rev. A*, vol. 66, no. 2, 022318, Aug 2002.
  - [12] Z. B. Khanian and A. Winter, “Entanglement-Assisted Quantum Data Compression,” in: *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Paris, France, pp. 1147-1151, July 2019.
  - [13] C. A. Fuchs and J. van de Graaf, “Cryptographic distinguishability measures for quantum-mechanical states,” *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1216-1227, May 1999.
  - [14] W. F. Stinespring, “Positive Functions on  $C^*$ -Algebras,” *Proc. Amer. Math. Society*, vol. 6, no. 2, pp. 211-216, 1955.
  - [15] C. H. Bennett and S. J. Wiesner, “Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states,” *Phys. Rev. Lett.*, vol. 69, no. 20, pp. 2881-2884, Nov 1992.
  - [16] M. Fannes, “A continuity property of the entropy density for spin lattice systems,” *Commun. Math. Phys.*, vol. 31, no. 4, pp. 291-294, Dec 1973.
  - [17] K. M. R. Audenaert, “A sharp continuity estimate for the von Neumann entropy,” *J. Phys. A: Math. Theor.*, vol. 40, no. 28, pp. 8127-8136, 2007.
  - [18] R. T. Rockafeller, *Convex Analysis*, Princeton University Press, 1970.
  - [19] A. Anshu, D. Leung and D. Touchette, “Incompressibility of classical distributions,” arXiv[quant-ph]:1911.09126, Nov 2019.