Generalized remote state preparation: Trading cbits, qubits, and ebits in quantum communication

Anura Abeyesinghe* and Patrick Hayden†
Institute for Quantum Information, Physics Department, Caltech, 103-33, Pasadena, California 91125, USA
(Received 29 August 2003; published 24 December 2003)

We consider the problem of communicating quantum states by simultaneously making use of a noiseless classical channel, a noiseless quantum channel, and shared entanglement. We specifically study the version of the problem in which the sender is given knowledge of the state to be communicated. In this setting, a trade-off arises between the three resources, some portions of which have been investigated previously in the contexts of the quantum-classical trade-off in data compression, remote state preparation, and superdense coding of quantum states, each of which amounts to allowing just two out of these three resources. We present a formula for the triple resource trade-off that reduces its calculation to evaluating the data compression trade-off formula. In the process, we also construct protocols achieving all the optimal points. These turn out to be achievable by trade-off coding and suitable time sharing between optimal protocols for cases involving two resources out of the three mentioned above.

DOI: 10.1103/PhysRevA.68.062319 PACS number(s): 03.67.Hk, 03.65.Ta

I. INTRODUCTION

Quantum information theory can be described as the effort to identify and quantify the basic resources required to communicate or, more generally, process information in a quantum-mechanical setting. The dual goals of identifying new protocols and demonstrating their optimality have, respectively, helped to expose the surprising range of information processing tasks facilitated by quantum mechanics and highlighted the subtle ways in which physics dictates limitations on the transmission and processing of information.

Part of the appeal of the information theoretic paradigm is that it emphasizes the notions of interconvertibility and simulation. Identifying basic resources and evaluating their interconvertibility provides a general strategy for systematically charting the capabilities of quantum-mechanical systems. Some early successes of this approach include Schumacher's quantum noiseless coding theorem [1,2], which demonstrated that a single number quantifies the compressibility of memoryless sources of quantum states, and the theory of pure state bipartite entanglement, where a single number, likewise, determines the asymptotic interconvertibility of entanglement [3]. More recently, we have seen how to evaluate the interconvertibility of quantum memories [4] and even seen that the rate at which one noisy quantum channel can simulate any other (in the presence of entanglement and with certain restrictions on the input) is controlled again by a single number, the channel's entanglement-assisted capacity [5].

From the point of view of communication theory, these results identify three basic and inequivalent resources: noiseless classical channels, noiseless quantum channels and maximally entangled states. Other inequivalent resources exist, of course. One such, classically correlated bits, will prove useless for the problem we investigate. Noisy versions of the basic list of three resources identified above potentially adds many others but we do not study them here. Those caveats aside, the three basic resources serve as formalized versions of abstract ''classicality,'' ''quantumness,'' and ''nonlocality,'' quantifiable in units of classical bits (cbits), quantum bits (qubits), and maximally entangled qubits (ebits). While the three basic resources are inequivalent, relationships exist between them. Because cbits can be encoded in qubits and ebits can be established by sending qubits, the noiseless quantum channel is (in this narrow sense) the strongest of the three. Because it is impossible to establish entanglement using classical communication or to communicate using only entanglement, ebits and cbits are simply incomparable; neither is truly stronger than the other.

In the present work, we quantify the relationship between the three resources for a basic task in quantum information theory: communicating quantum states from a sender to a receiver (and, more generally, sharing entangled states between them). There are at least two variations on the task, depending on whether or not the sender has knowledge of the states she is required to communicate. If she is only given a copy of the quantum state and not a description, we describe the source as hidden and the encoding as oblivious (or blind). At the other extreme, if she is told which state she is required to transmit, we describe the source as visible and the encoding as nonoblivious. (Sometimes in the quantum information literature the adjective ''visible'' is also applied, somewhat nonsensically, to the encoding.) While the distinction makes no difference in classical information theory, quantum-mechanical restrictions on the sender's ability to measure without causing a disturbance lead to very different results for the two tasks in the quantum case. (Compare, for example, the results of Refs. [6–8].) Our emphasis here is on the visible scenario since there is generically only a trivial trade-off for the blind encoder case: using teleportation, two cbits and one ebit can be used to simulate a noiseless one-qubit channel but no other interesting trade-offs are possible.

In the visible scenario, the relationship between the three resources becomes much more varied. When no quantum channel is permitted, we recover the problem known as remote state preparation [9,10], while forbidding use of the

*Electronic address: anura@caltech.edu
†Electronic address: patrick@cs.caltech.edu

classical channel leads to superdense coding of quantum states [11,12]. Likewise, if entanglement is not permitted, we recover the trade-off between classical and quantum communication solved in Ref. [8]. The present paper completely solves the problem of trading all three resources against each other, finding that optimal protocols for any combination of resources can be constructed by appropriate combinations of the protocols representing the extremes identified above. Such a clean resolution in terms of previously discovered building blocks is encouraging: it confirms yet again the simplifying power of the resource-based approach, this time yielding a manageable taxonomy of optimal protocols for the triple trade-off problem.

The rest of the paper is structured as follows. Section II defines the problem rigorously and describes previous results for the cases when one of the three resources is not used, along with some minor extensions. Section III studies the relationship between the trade-off between qubits and cbits in quantum data compression (QCT) and the trade-off between ebits and cbits in remote state preparation (RSP). In Sec. IV these connections and the results described in Sec. II are used to obtain optimal protocols and optimal resource trade-offs for communicating quantum states when all three resources are used simultaneously: the full "triple trade-off."

We use the following conventions throughout the paper. If $\mathcal{E}_{AB}=\{\varphi_i^{AB},p_i\}$ is an ensemble of bipartite states then we write $\mathcal{E}_A$ for the ensemble $\{\varphi_i^A,p_i\}$ of reduced states on system A. Sometimes we omit subscripts (or superscripts) labeling subsystems, in which case the largest subsystem on which the ensemble (or state) has been defined should be assumed: $\mathcal{E}=\mathcal{E}_{AB}$ and $\varphi_i=\varphi_i^{AB}$. We identify states with their density operators and if $|\varphi\rangle$ is a pure state, we use the notation $\varphi=|\varphi\rangle\langle\varphi|$ for its density operator. The function $S(\rho)$ is the von Neumann entropy $S(\rho)=-\mathrm{Tr}\rho\log\rho$ and $S(\mathcal{E})$ the von Neumann entropy of the average state of the ensemble $\mathcal{E}$. Functions like $S(A|B)_\rho$ and $S(A:B|C)_\rho$ are defined in the same way as their classical counterparts:

$$S(A:B|C)_\rho=S(\rho^{AC})+S(\rho^{BC})-S(\rho^{ABC})-S(\rho^C), \quad (1)$$

for example. $\chi(\mathcal{E})$ is the Holevo χ quantity of $\mathcal{E}$ [13]. Given a bipartite ensemble $\mathcal{E}_{AB}=\{\varphi_i^{AB},p_i\}$, we also make use the abbreviations $S=S(\mathcal{E}_B)$, $\overline{S}=\Sigma_i p_i\varphi_i^B$, $\chi=\chi(\mathcal{E}_B)$ and $H=H(p_i)$. Throughout, log and exp are taken base 2.

II. DEFINITION OF THE PROBLEM AND PREVIOUS RESULTS

We now give a more formal definition of the task to be completed by the sender and receiver, henceforth, respectively, Alice and Bob. The reader can also refer to Fig. 1, which illustrates the definition. We consider an ensemble of bipartite quantum states $\mathcal{E}=\{|\varphi_i\rangle^{AB},p_i\}$ on a finite-dimensional Hilbert space $\mathcal{H}_{AB}=\mathcal{H}_A\otimes\mathcal{H}_B$ and the product ensembles $\mathcal{E}^{\otimes n}=\{|\varphi_{i^n}\rangle^{AB},p_{i^n}\}$ on $\mathcal{H}_{AB}^{\otimes n}$, where

$$i^n=i_1i_2\cdots i_n,$$

$$p_{i^n}=p_{i_1}p_{i_2}\cdots p_{i_n},$$

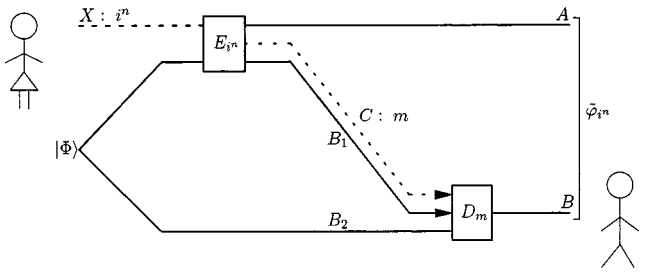

FIG. 1. In the above quantum circuit diagram for generalized remote state preparation, time goes from left to right, solid lines represent quantum registers, and dashed lines represent classical registers. The registers connected in the left represent a maximally entangled state of $\log d_E$ ebits initially shared between Alice and Bob. The $\log d_Q$-qubit quantum register B_1 is sent from Alice to Bob, as is the $\log d_C$ cbit classical message m. Alice's encoding operation is denoted by E_{i^n} and Bob's decoding operation, which is conditioned on m, by D_m.

and

$$|\varphi_{i^n}\rangle=|\varphi_{i_1}\rangle\otimes|\varphi_{i_2}\rangle\otimes\cdots\otimes|\varphi_{i_n}\rangle.$$

At the end of the protocol, Alice and Bob are to reproduce the states of the bipartite ensemble with high fidelity. (Regardless of whether pure states are *prepared* in Bob's system, or entangled states are *shared* between Alice and Bob, we will always refer to the task simply as *communicating* from Alice to Bob.) We imagine that there is a noiseless classical channel from Alice to Bob capable of sending one of d_C messages, a noiseless quantum channel capable of sending a d_Q-dimensional quantum system and a maximally entangled state $|\Phi\rangle=d_E^{-1/2}\Sigma_{i=1}^{d_E}|i\rangle|i\rangle$ of Schmidt rank d_E. A source provides Alice with i^n, drawn with probability p_{i^n}, at which point Alice applies a quantum operation E_{i^n} to her half of $|\Phi\rangle$ that without loss of generality has output of the form

$$\sum_{j=1}^{d_C}\rho_{i^n,j}^{AB_1B_2}\otimes q(j|i^n)|j\rangle\langle j|^C, \quad (2)$$

where B_1 is a d_Q-dimensional quantum system, B_2 is the quantum system supporting Bob's half of $|\Phi\rangle$, the states $\{|j\rangle\}$ are orthonormal (i.e., classical) and $q(\cdot|i^n)$ is a probability distribution. Alice then sends register B_1 to Bob over her noiseless quantum channel and C to Bob over the noiseless classical channel. The protocol is completed by Bob performing a quantum operation D_j on registers B_1 and B_2. Write $\tilde{\varphi}_{i^n}$ for the joint Alice-Bob output state averaged over different values of j. We say that the protocol has fidelity $1-\epsilon$ if

$$\sum_{i^n}p_{i^n}\langle\varphi_{i^n}|\tilde{\varphi}_{i^n}|\varphi_{i^n}\rangle\geq 1-\epsilon. \quad (3)$$

Likewise, (R,Q,E) is an achievable rate triple for the ensemble $\mathcal{E}$ if for all $\delta,\epsilon>0$ there exists N such that for all $n>N$ there is a protocol for $\mathcal{E}^{\otimes n}$ with fidelity $1-\epsilon$ and

$$\frac{1}{n}\log d_C \leqslant R+\delta, \quad \frac{1}{n}\log d_Q \leqslant Q+\delta, \quad \frac{1}{n}\log d_E \leqslant E+\delta. \tag{4}$$

Our goal will be to identify these achievable triples. In particular, we will find a formula for the function

$$E^*(R,Q) = \inf\{E:(R,Q,E) \text{ is achievable}\}. \tag{5}$$

We refer to rate triples of the form $(R,Q,E^*(R,Q))$ as optimal rate triples and the protocols that achieve them as optimal protocols. We will indicate that a rate triple (R,Q,E) is optimal by writing it as $(R,Q,E)^*$. Throughout the paper, unless otherwise stated, all entropic quantities will be taken with respect to 4-partite states ω of the following form:

$$\omega = \sum_i p_i |i\rangle\langle i|^X \otimes \varphi_i^{AB} \otimes \sum_{j=1}^{m+1} p(j|i)|j\rangle\langle j|^C, \tag{6}$$

where m is the number of states in $\mathcal{E}_{AB}$ (if that number is finite), and $p(\cdot|\cdot)$ is a classical noisy channel. Note that for all such states

$$S(X:B|C) = S(B|C) - \overline{S} \quad \text{where } \overline{S} = \sum_i p_i S(\varphi_i^B), \tag{7}$$

a fact that will be useful later. Before moving on to the general problem, we consider the special cases given by setting one of the three rates to zero.

A. $Q=0$: Remote state preparation (RSP)

This problem was studied extensively in Ref. [14]. It is impossible to achieve an entanglement rate of less than $\Sigma_i p_i \varphi_i^B$, essentially because that is the amount of entanglement shared between Alice and Bob at the end of any successful protocol. The optimal cbit rate when the entanglement is minimal is just $H(p_i)$, meaning that the simple protocol consisting of Alice communicating i^n to Bob and then the pair performing entanglement dilution is optimal. At the other extreme, the cbit rate is minimized (at least for irreducible sources) by a protocol achieving the rate $(\chi(\mathcal{E}_B),0,S(\mathcal{E}_B))$. In general, we introduce the function

$$E^*(R) = \inf\{E : (R,0,E) \text{ is achievable}\}. \tag{8}$$

This choice, a slight abuse of notation given our earlier definition of a function E^* with two arguments, is chosen for consistency with the remote state preparation paper. Note that $E^*(R) = E^*(R,0)$. We have the following theorem from Ref. [14].

Theorem II.1. For the ensemble $\mathcal{E}=\{|\varphi_i\rangle^{AB},p_i\}$ of pure bipartite states and $R \geqslant 0$,

$$E^*(R) = \min\{S(B|C) : S(X:BC) \leqslant R\}, \tag{9}$$

where the entropic quantities are with respect to the state ω, minimization is over all 4-partite states ω of the form of Eq. (6) with classical channels $p(j|i)$, and m the number of states in $\mathcal{E}$. E^* is convex, continuous, and strictly decreasing in the interval in which it takes positive values.

We will also use the simple fact that the inequality in Eq. (9) can be replaced by equality.

B. $E=0$: Quantum-classical trade-off (QCT)

The case where the ensemble $\mathcal{E}$ consists only of product states $|\varphi_i\rangle^{AB} = |0\rangle^A|\varphi_i\rangle^B$ was the focus of Ref. [8]. At the extreme when $R=0$, only quantum communication is permitted so the problem of finding achievable rates is answered by the quantum noiseless coding theorem: $(0,S(\mathcal{E}_B),0)$ is an *optimal point*, in the sense that none of the three rates can be reduced. Likewise, the optimal point when $Q=0$ is given by $(H(p_i),0,0)$, meaning that Alice has no better strategy than to communicate the label i^n to Bob. More generally, when the ensemble is allowed to contain entangled states, the techniques of Refs. [8], [14] are easily adapted to yield a formula for

$$Q^*(R) = \inf\{Q:(R,Q,0) \text{ is achievable}\}. \tag{10}$$

In particular, we have the following analog of Theorem II.1.

Theorem II.2. For the ensemble $\mathcal{E}=\{|\varphi_i\rangle^{AB},p_i\}$ of pure bipartite states and $R \geqslant 0$,

$$Q^*(R) = \min\{S(B|C):S(X:C) \leqslant R\}, \tag{11}$$

where the entropic quantities are with respect to the state ω, minimization is over all 4-partite states ω of the form of Eq. (6) with classical channels $p(j|i)$, and m the number of states in $\mathcal{E}$. Q^* is convex, continuous, and strictly decreasing in the interval in which it takes positive values. There exists a critical value of R, hereafter referred to as H_c such that $R+Q^*(R)=S(B)$ for $R \leqslant H_c$ and $R+Q^*(R) > S(B)$ otherwise.

As before, the inequality in Eq. (11) can be replaced by an equality.

C. $R=0$: Superdense coding of quantum states (SDC)

Reference [12] showed that it is possible to communicate arbitrary d^2-dimensional quantum states using $\log d + o(\log d)$ qubits, $\log d + o(\log d)$ ebits and shared random bits. For exploring the trade-off of quantum resources, we need a variation on this result that applies to ensembles of entangled states: using his coherent classical communication technique, Harrow has shown that

$$(0,\tfrac{1}{2}\chi(\mathcal{E}_B), S(\mathcal{E}_B) - \tfrac{1}{2}\chi(\mathcal{E}_B)) \tag{12}$$

is an achievable rate triple [15]. Using his construction, we can easily find the $R=0$ trade-off curve.

Theorem II.3. For the ensemble $\mathcal{E}=\{|\varphi_i\rangle^{AB},p_i\}$ of pure bipartite states and $Q \geqslant 0$,

$$E^*(0,Q) = \begin{cases} S(\mathcal{E}_B) - Q & \text{if } Q \geqslant \chi(\mathcal{E}_B)/2 \\ +\infty & \text{otherwise.} \end{cases} \tag{13}$$

Proof. Since $(0,S,0)$ and $(0,\chi/2,S-\chi/2)$ (S and χ are defined in the Introduction) are both achievable rate triples, any

convex combination of the two is an achievable rate triple corresponding to a time-shared protocol. Thus, if $0 \leqslant \lambda \leqslant 1$,

$$(0, \lambda S + (1-\lambda)\chi/2, (1-\lambda)(S - \chi/2)) \tag{14}$$

is achievable. Suppose these points are not optimal. Then there exists $\epsilon > 0$ such that

$$(0, \lambda S + (1-\lambda)\chi/2, (1-\lambda)(S - \chi/2) - \epsilon) \tag{15}$$

is optimal. By using quantum communication to establish entanglement, however, protocols achieving this rate can be converted into protocols with the rate triple

$$(0, \lambda S + (1-\lambda)\chi/2 + (1-\lambda)(S - \chi/2) - \epsilon, 0) = (0, S - \epsilon, 0), \tag{16}$$

contradicting the optimality of Schumacher compression. We conclude that $E^*(0,Q) = S - Q$ when this conversion is possible, that is, when $Q \geqslant \chi/2$. This condition is required by causality. (For a detailed proof, see Sec. IV C.) ■

The simple argument used in the proof of Theorem II.3 is characteristic of what will follow. Our evaluation of $E^*(R,Q)$ will be accomplished via operational reductions to the three extremal cases we have now completed, just as Theorem II.3 was demonstrated using a reduction from the unknown $E^*(0,Q)$ curve to the known Schumacher compression point.

Later we will also have occasion to make use of the following analog of the QCT and RSP constructions. Given a state ω of the form of Eq. (6), the trade-off coding technique from Ref. [8] then gives protocols achieving all the rate triples of the form

$$(S(X:C), \tfrac{1}{2} S(X:B|C), S(B|C) - \tfrac{1}{2} S(X:B|C)). \tag{17}$$

Briefly, once a channel $p(j|i)$ is chosen, Alice and Bob can share (typical) $j^n = j_1 \cdots j_n$ at a cost of $nS(X:C) + o(n)$ bits of communication plus shared random bits using the Reverse Shannon Theorem [16]. Harrow's protocol is then used on the induced ''conditional'' ensembles

$$\{|\varphi_{i^n}\rangle^{AB}, q(i^n|j^n) = q(i_1|j_1) \cdots q(i_n|j_n)\},$$

where

$$q(i|j) = \left(\sum_{i'} p_{i'} p(j|i') \right)^{-1} p(j|i) p_i . \tag{18}$$

The shared random bits are then seen to be unnecessary because we only require high fidelity on average (so that some particular value of the shared random bits can be used). Evaluation of the rates for the approach gives exactly Eq. (17).

Given any $(R, Q^*(R), 0)$ there is a state ω of the form Eq. (6) for which $(S(X:C), S(B|C), 0) = (R, Q^*(R), 0)$. For this state, we therefore find a new achievable rate triple:

$$(S(X:C), \tfrac{1}{2} S(X:B|C), S(B|C) - \tfrac{1}{2} S(X:B|C))$$

$$= (R, \tfrac{1}{2}(Q^*(R) - \overline{S}), \tfrac{1}{2}(Q^*(R) + \overline{S})), \tag{19}$$

where we have used Eq. (7) to arrive at the expression on the right-hand side.

III. RELATING OPTIMAL QCT AND OPTIMAL RSP

Any protocol for quantum-classical compression can be converted into a RSP protocol by using a RSP to send the compressed qubits. One might hope that if the original QCT point was optimal that the resulting RSP point would also be optimal. For classical rates above H_c this is indeed the case but otherwise it need not be. Consider, for example, the ensemble consisting of the orthonormal states $|0\rangle$ and $|1\rangle$, each occurring with probability 1/2. In this case, $Q^*(0) = 1$ but the corresponding RSP protocol would wastefully consume 1 cbit and 1 ebit per signal when 1 cbit and no entanglement are sufficient.

As an aside, while there is a natural way to convert optimal QCT protocols into optimal RSP protocols (for unentangled ensembles and $R \geqslant H_c$), there is no known way to do the opposite. An appendix to Ref. [14], however, demonstrates the existence of just such an operational reduction but only under the assumption that the mixed state compression conjecture is true. (See Refs. [17–19] for more details on the conjecture.)

The following two lemmas formally express the relationship between optimal QCT and optimal RSP.

Lemma III.1. When $R \geqslant H_c$, $E^*[R + Q^*(R) - \overline{S}] = Q^*(R)$. Otherwise, $E^*[R + Q^*(R) - \overline{S}] = Q^*(H_c)$.

Proof. We begin by showing that $E^*[R + Q^*(R) - \overline{S}] \leqslant Q^*(R)$. We know that $(S(X:BC), 0, S(B|C))$ is an achievable rate triple for any ω of the form of Eq. (6). In particular, it is achievable when $(S(X:C), S(B|C), 0) = (R, Q^*(R), 0)$, in which case

$$(S(X:BC), 0, S(B|C)) = (S(X:C) + S(B|C) - \overline{S}, 0, S(B|C)) \tag{20}$$

$$= (R + Q^*(R) - \overline{S}, 0, Q^*(R)). \tag{21}$$

This proves the claim. Note that this inequality is true regardless of whether R is greater or less than H_c.

We now prove the opposite inequality: $E^*[R + Q^*(R) - \overline{S}] \geqslant Q^*(R)$ when $R \geqslant H_c$. Substituting our expressions for $E^*(R)$ and $Q^*(R)$ shows that what we need to prove is that

$$\min\{S(B|C) : S(X:C) + S(B|C) = R + Q^*(R)\} \tag{22}$$

$$\geqslant \min\{S(B|C) : S(X:C) = R\}. \tag{23}$$

Let ω be the state that minimizes the first expression for fixed R. If $S(X:C)_\omega \leqslant R$ then we are done so we may suppose not: $S(X:C)_\omega = R + \Delta$ for some $\Delta > 0$. By convexity and the definition of H_c, for any $R \geqslant H_c$,

$$\frac{Q^*(R + \Delta) - Q^*(R)}{\Delta} > -1. \tag{24}$$

Rearranging this inequality yields

$$(R+\Delta)+Q^*(R+\Delta)>R+Q^*(R). \tag{25}$$

Using the hypothesis $S(X:C)_\omega = R+\Delta$ and the fact that the right-hand side of the above inequality is $S(X:C)_\omega + S(B|C)_\omega$, we find that $S(B|C)_\omega < Q^*(R+\Delta)$. But, again by hypothesis, $S(X:C)_\omega = R+\Delta$ so we have a contradiction of the definition of $Q^*(R+\Delta)$. We conclude that $S(X:C)_\omega \leqslant R$.

Finally, $R+Q^*(R)-\overline{S}=\chi$ when $R<H_c$ so $E^*(R) = E^*(\chi)$ is constant. Using the first half of the lemma, we then find $E^*(\chi)=E^*[H_c+Q^*(H_c)-\overline{S}]=Q^*(H_c)$. ■

Lemma III.2. $Q^*[R-E^*(R)+\overline{S}]=E^*(R)$ when $R \geqslant \chi$. Otherwise $E^*(R)=+\infty$.

Proof. Let $H_c \leqslant R_1$ and consider $R=R_1+Q^*(R_1)-\overline{S}$. R is a strictly increasing function of R_1 by the definition of H_c, taking all values $\chi \leqslant R$. Substituting into Lemma III.1 gives

$$Q^*[R-E^*(R)+\overline{S}]=Q^*[R_1+Q^*(R_1)-\overline{S}-Q^*(R_1)+\overline{S}] \tag{26}$$

$$=Q^*(R_1) \tag{27}$$

$$=E^*[R_1+Q^*(R_1)-\overline{S}] \tag{28}$$

$$=E^*(R). \tag{29}$$

Also, $R<\chi$ is not achievable (by causality, see Sec. IV C), yielding the second half of the lemma. ■

IV. THE TRIPLE TRADE-OFF

The following theorem is the main result of the paper: a prescription for calculating the minimal amount of entanglement required given any cbit and qubit rate.

Theorem IV.1. We have

$$E^*(R,Q) = \begin{cases} 0 & \text{if } Q^*(R)<Q \\ Q^*(R)-Q & \text{if } \frac{1}{2}[Q^*(R)-\overline{S}] \leqslant Q \leqslant Q^*(R) \\ E^*(R+2Q)-Q & \text{if } \frac{1}{2}(\chi-R) \leqslant Q < \frac{1}{2}[Q^*(R)-\overline{S}] \\ +\infty & \text{if } Q<\frac{1}{2}(\chi-R) \end{cases}.$$

We discuss each of the four ranges for Q separately, referring to them, in order, as the *QCT region*, the *low-entanglement region*, the *high-entanglement region*, and the *forbidden region*. The names of the first and last regions should be self-explanatory. (QCT is optimal by definition in the QCT region and no amount of entanglement is sufficient in the forbidden region.) In the low-entanglement region we will find that optimal protocols can be found by time-sharing between QCT and SDC (the first of which does not use entanglement) while the optimal protocols for the high-entanglement region are found by time-sharing between RSP and SDC, *both* of which rely on entanglement.

While H_c does not appear explicitly in our formula, it once again delineates the boundary between two qualitatively different regimes: for $R<H_c$ we have that $\frac{1}{2}[Q^*(R) - \overline{S}] = \frac{1}{2}(\chi-R)$ so there is no high-entanglement region in this case. The region defined by $R<H_c$ and $Q \geqslant \frac{1}{2}(\chi-R)$ is entirely contained in the low-entanglement region.

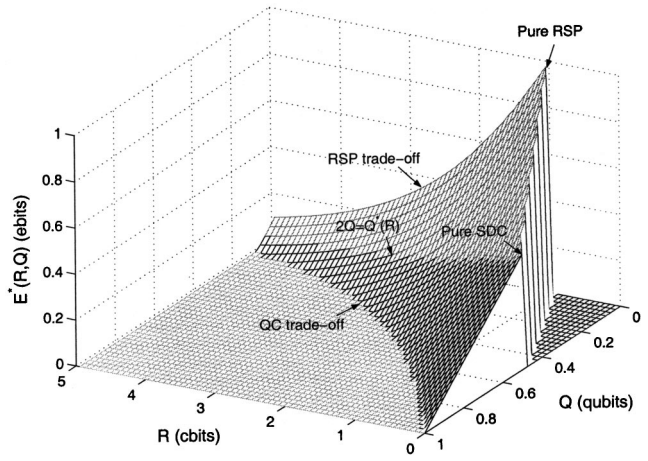

FIG. 2. Trade-off surface for the uniform qubit ensemble. The region on the left for which $E^*(R,Q)=0$ is the QCT region, whose boundary with the low-entanglement region is given by the curve $(R,Q^*(R),0)$. The transition to the high-entanglement region then occurs when $2Q=Q^*(R)$; note that the surface is not smooth at the transition. Finally the points corresponding to pure RSP, (1, 0, 1), and pure SDC, (0, 1/2, 1/2), define the boundary of the forbidden region. The surface matches the convex hull of the QCT, SDC, and RSP curves.

Before giving a proof of Theorem IV.1, we consider the standard example: $\mathcal{E}_{AB}$ being the uniform (unitarily invariant) ensemble over qubit states on B. Devetak and Berger gave an explicit parametrization [20] of the function identified as $Q^*(R)$ for this ensemble in Ref. [8] and the corresponding RSP curve appeared in Ref. [14]. We present the full trade-off surface $E^*(R,Q)$ in Fig. 2. (In the case of an infinite ensemble, Theorems II.1 and II.2 need to be slightly modified: the min should be replaced by an inf as explained in theorem 10.1 of Ref. [8]. The only significant modification required to the argument of this paper is in the second half of Lemma III.1, where a sequence of ω_n needs to be considered instead of a fixed minimizing ω.)

We also summarize for convenience in Table I all the rate triples and conversions between them that we will use in the proof. We use the notation $(R,Q,E) \rightarrow (R',Q',E')$ to indicate that if the rate triple (R,Q,E) is achievable then so is the rate triple (R',Q',E'), i.e., (R,Q,E) can be *converted* into (R',Q',E'). Similarly, if we write $(R,Q,E)^* \rightarrow (R',Q',E')$ then the conversion is possible conditional on (R,Q,E) being optimal.

A. The low-entanglement region: $\frac{1}{2}[Q^*(R)-\overline{S}] \leqslant Q \leqslant Q^*(R)$

Define $\lambda = 2[Q^*(R)-Q]/[Q^*(R)+\overline{S}]$. By the definition of the low-entanglement region, $0 \leqslant \lambda \leqslant 1$. Both $(R,Q^*(R),0)$ and $(R,\frac{1}{2}[Q^*(R)-\overline{S}],1/2[Q^*(R)+\overline{S}])$ are achievable so the convex combination

TABLE I. Achievable rate triples and conversions.

Rate triple	Description
$(R, Q^*(R), 0)$	QCT
$(R, 0, E^*(R))$	RSP
$(R, \frac{1}{2}[Q^*(R)-S], \frac{1}{2}[Q^*(R)+S])$	SDC on QCT: Eq. (19)
$(R+Q^*(R)-\overline{S}, 0, Q^*(R))$ for $R \geqslant H_c$	QCT to RSP: Lemma III.1
$(R-E^*(R)+\overline{S}, E^*(R), 0)$	RSP to QCT: Lemma III.2
$(R,Q,E) \rightarrow (R+2Q, 0, E+Q)$	Teleportation (of qubits)
$(R,Q,E) \rightarrow (0, Q+\frac{1}{2}R+Q, \frac{1}{2}R+E)$	Superdense coding (of cbits)
(R_1,Q_1,E_1) and (R_2,Q_2,E_2) $\rightarrow \lambda(R_1,Q_1,E_1)+(1-\lambda)(R_2,Q_2,E_2)$	Time sharing
$(R,Q,E) \rightarrow (R, Q+E, 0)$	Sending entanglement using qubits
$(R,Q,E)^* \rightarrow (R-E+Q-\overline{S}, Q+E, 0)$ if $R \geqslant \overline{S}$ and $E > Q+\overline{S}$	Lemma IV.2

$$[R,Q,Q^*(R)-Q] = \lambda(R,Q^*(R),0) + (1-\lambda) \times (R, \tfrac{1}{2}[Q^*(R)-\overline{S}], \tfrac{1}{2}[Q^*(R)+\overline{S}]) \tag{30}$$

is achievable by time sharing.

The proof that these points are optimal is very simple. Suppose they are not. Then there would exist an ϵ such that $(R,Q,Q^*(R)-Q-\epsilon)$ were optimal. Now, using the conversion $(R,Q,E) \rightarrow (R,Q+E,0)$, it follows that $(R,Q^*(R)-\epsilon,0)$ is achievable, which is a contradiction of the definition of Q^*.

B. The high-entanglement region: $\frac{1}{2}(\chi - R) \leqslant Q < \frac{1}{2}[Q^*(R)-\overline{S}]$

This region seems to require a more elaborate analysis. We first define two new variables R_1 and R_2 which are functions of R and Q but easier to work with

$$R_1 = R + 2Q - E^*(R+2Q) + \overline{S}, \tag{31}$$

$$R_2 = R - R_1 + \overline{S} = E^*(R+2Q) - 2Q. \tag{32}$$

We collect for future use some simple facts about R_1 and R_2.

(1) $R_1 \geqslant H_c$. The function $R' - E^*(R') + \overline{S}$ is a monotonically increasing function of R'. By causality, therefore, the minimum of this function over achievable R' occurs when $R' = \chi$. From Lemma III.1, $E^*(\chi) = Q^*(H_c) = S - H_c$, so $R' - E^*(R') + \overline{S} \geqslant H_c$. Since $R + 2Q \geqslant \chi$ in the high-entanglement region, we conclude that $R_1 \geqslant H_c$.

(2) $Q = \frac{1}{2}[Q^*(R_1) - R_2]$. This follows by Lemma III.2: $Q^*(R_1) = E^*(R+2Q) = R_2 + 2Q$.

(3) $E^*(R+2Q) - Q = R_2 + Q = \frac{1}{2}[Q^*(R_1) + R_2]$. This follows by the definition of R_2 and the previous fact.

(4) $R_2 \leqslant Q^*(R_1)$. By fact (1), $R_2 = Q^*(R_1) - 2Q$.

(5) $Q^*(R_1) \geqslant \overline{S}$. $Q^*(R_1) - \overline{S} = S(B|C) - \overline{S} = S(X:B|C) \geqslant 0$ (for optimal ω).

(6) $R_2 \geqslant \overline{S}$ (for $Q \leqslant \frac{1}{2}[Q^*(R) - \overline{S}]$). This is equivalent to $E^*(R+2Q) \geqslant 2Q + \overline{S}$. Since $2Q \leqslant Q^*(R) - \overline{S}$ in this region, we have by the monotonicity of E^* and by Lemma III.1 that

$$E^*(R+2Q) \geqslant E^*[R + Q^*(R) - \overline{S}] \tag{33}$$

$$= Q^*(R) \tag{34}$$

$$\geqslant 2Q + \overline{S}. \tag{35}$$

Equipped with these observations we can now proceed to the proof of Theorem IV.1 in the high-entanglement region. That is, we will prove that $E^*(R,Q) = E^*(R+2Q) - Q$ when $\frac{1}{2}(\chi - R) \leqslant Q < \frac{1}{2}[Q^*(R) - \overline{S}]$. Note that

$$(R,Q,E^*(R+2Q)-Q) = (R_1 + R_2 - \overline{S}, \tfrac{1}{2}[Q^*(R_1) - R_2], \tfrac{1}{2}[Q^*(R_1) + R_2]) \tag{36}$$

in terms of the new variables, by the definition of R_1 and R_2 as well as facts 2 and 3.

1. Proof of achievability

$(R_1, \frac{1}{2}[Q^*(R_1) - \overline{S}], \frac{1}{2}[Q^*(R_1) + \overline{S}])$ is achievable by Eq. (19) and $(R_1 + Q^*(R_1) - \overline{S}, 0, Q^*(R_1))$ is achievable by Lemma III.1. By facts (4), (5), and (6), $\lambda = [Q^*(R_1) - R_2]/[Q^*(R_1) - \overline{S}]$ is between 0 and 1. Therefore, the convex combination

$$(R_1 + R_2 - \overline{S}, \tfrac{1}{2}[Q^*(R_1) - R_2], \tfrac{1}{2}[Q^*(R_1) + R_2]) \tag{37}$$

$$= \lambda(R_1, \tfrac{1}{2}[Q^*(R_1) - \overline{S}], \tfrac{1}{2}[Q^*(R_1) + \overline{S}]) + (1-\lambda)(R_1 + Q^*(R_1) - \overline{S}, 0, Q^*(R_1)) \tag{38}$$

is also achievable by time sharing.

2. Proof of optimality

We defer the proof of the following lemma, which is at the heart of our optimality proof, to the end of the section.

Lemma IV.2. If R_1, $Q \geqslant 0$ and $R_2 > \overline{S}$, then there is a conversion

$$(R_1+R_2,Q,R_2+Q)^* \rightarrow (R_1+\overline{S},R_2+2Q,0). \tag{39}$$

(Note that when $R_2=\overline{S}$, the conversion always exists, regardless of the optimality of the first rate triple.) Now suppose that points of the form of Eq. (36) are not optimal. Then there exists some $\epsilon>0$ such that

$$(R_1+R_2-\overline{S},\tfrac{1}{2}[Q^*(R_1)-R_2],\tfrac{1}{2}[Q^*(R_1)+R_2]-\epsilon) \tag{40}$$

is optimal. We handle the cases $R_2 > \overline{S}+\epsilon$ and $R_2 \leqslant \overline{S}+\epsilon$ separately.

Assume first that $R_2 > \overline{S}+\epsilon$, then define $R_1' = R_1 - \overline{S} + \epsilon$ and $R_2' = R_2 - \epsilon$. Rewriting the triple (40) in terms of R_1' and R_2', we have that

$$(R_1'+R_2',\tfrac{1}{2}[Q^*(R_1)-R_2],R_2'+\tfrac{1}{2}[Q^*(R_1)-R_2]) \tag{41}$$

is optimal. Since $R_2' > \overline{S}$, we can use Lemma IV.2 to obtain that $(R_1+\epsilon,Q^*(R_1)-\epsilon,0)$ is achievable. This implies that $Q^*(R_1+\epsilon) \leqslant Q^*(R_1)-\epsilon$, which is a contradiction since, by fact (1), $R_1 \geqslant H_c$.

If instead $R_2 \leqslant \overline{S}+\epsilon$, we apply the conversion $(R,Q,E) \rightarrow (R,Q+E,0)$ obtained by using quantum communication to establish entanglement,

$$(R_1+R_2-\overline{S},\tfrac{1}{2}[Q^*(R_1)-R_2],\tfrac{1}{2}[Q^*(R_1)+R_2]-\epsilon)^* \rightarrow (R_1+R_2-\overline{S},Q^*(R_1)-\epsilon,0). \tag{42}$$

This implies that $Q^*(R_1+R_2-\overline{S}) \leqslant Q^*(R_1)-\epsilon$. We also have $Q^*(R_1+\epsilon) \leqslant Q^*(R_1+R_2-\overline{S})$ by assumption and the monotonicity of Q^*. As before, we find that $Q^*(R_1+\epsilon) \leqslant Q^*(R_1)-\epsilon$, which is a contradiction.

Proof (of Lemma IV.2). Performing teleportation yields the conversion

$$(R_1+R_2,Q,R_2+Q) \rightarrow (R_1+R_2+2Q,0,R_2+2Q). \tag{43}$$

(Note that teleportation is appropriate here instead of RSP because the encoding map corresponding to the first triple will generally produce complicated entangled states between Alice and Bob, conditioned on the classical bits being communicated. Teleportation will preserve this entanglement.) It will suffice to prove that the resulting triple is optimal because an application of Lemma III.2 would then show that $(R_1+\overline{S},R_2+2Q,0)$ is achievable.

Suppose then that $(R_1+R_2+2Q,0,R_2+2Q)$ is not optimal so that there exists some $\epsilon>0$ such that $(R_1+R_2+2Q,0,R_2+2Q-\epsilon)$ is optimal. By Lemma III.2 and then Eq. (19), there is a sequence of conversions

$$(R_1+R_2+2Q,0,R_2+2Q-\epsilon)^* \tag{44}$$

$$\rightarrow (R_1+\epsilon+\overline{S},R_2+2Q-\epsilon,0)^* \tag{45}$$

$$\rightarrow (R_1+\epsilon+\overline{S},\tfrac{1}{2}(R_2+2Q-\epsilon-\overline{S}),\tfrac{1}{2}(R_2+2Q-\epsilon+\overline{S})). \tag{46}$$

We handle the cases $R_2 \geqslant \overline{S}+\epsilon$ and $R_2 < \overline{S}+\epsilon$ separately.

Assume first that $R_2 \geqslant \overline{S}+\epsilon$. Then if we define $\lambda = (R_2 - \overline{S} - \epsilon)/(R_2+2Q-\epsilon-\overline{S})$, we have $0 \leqslant \lambda \leqslant 1$ so the convex combination

$$(R_1+R_2,Q,R_2+Q-\epsilon) \tag{47}$$

$$= \lambda(R_1+R_2+2Q,0,R_2+2Q-\epsilon) \tag{48}$$

$$+ (1-\lambda)(R_1+\epsilon+\overline{S},\tfrac{1}{2}(R_2+2Q - \epsilon-\overline{S}),\tfrac{1}{2}(R_2+2Q-\epsilon+\overline{S})) \tag{49}$$

is achievable, contradicting the optimality of (R_1+R_2,Q,R_2+Q).

Now suppose that $R_2 < \overline{S}+\epsilon$ and consider $\alpha = \epsilon + \overline{S} - R_2$, which is by definition positive. Rewriting the triple (45) in terms of α, applying the SDC conversion of Eq. (19) and then regular superdense coding of the cbits gives

$$(R_1+R_2+\alpha,2Q-\alpha+\overline{S},0)^* \tag{50}$$

$$\rightarrow (R_1+R_2+\alpha,Q-\alpha/2,Q-\alpha/2+\overline{S}) \tag{51}$$

$$\rightarrow (0,Q+\tfrac{1}{2}(R_1+R_2),Q+\tfrac{1}{2}(R_1+R_2)+\overline{S}). \tag{52}$$

Choosing $\lambda = \alpha/(R_1+R_2+\alpha)$, we can time share to achieve

$$(R_1+R_2,Q,Q+\overline{S}) \tag{53}$$

$$= \lambda(0,Q+\tfrac{1}{2}(R_1+R_2),Q+\tfrac{1}{2}(R_1+R_2)+\overline{S}) \tag{54}$$

$$+ (1-\lambda)(R_1+R_2+\alpha,Q-\alpha/2,Q-\alpha/2+\overline{S}), \tag{55}$$

contradicting again the optimality of (R_1+R_2,Q,R_2+Q) since $\overline{S} < R_2$ by the hypotheses of the lemma. ■

C. The forbidden region: $Q < \frac{1}{2}(\chi - R)$

In keeping with the operational spirit of the other arguments in this paper, we argue that achievability in this region would lead to a violation of causality. A classical channel of dimension d_C and a quantum channel of dimension d_Q can be used to transmit at most $\log d_C + 2\log d_Q$ bits of classical information by the optimality of superdense coding [11,13]. Success in the ensemble communication task, however, results in Bob holding a high-fidelity copy of $\mathcal{E}_B$. By using coding, Alice could then about communicate $\chi(\mathcal{E}_B)$ classical bits to Bob per usage of the protocol [21,22], a violation of causality (for sufficiently high fidelity and small δ in the notation of Sec. II) if $\chi(\mathcal{E}_B) > R+2Q$.

A simple entropic argument is also possible. Consider the state

$$\rho=\sum_{i^n,j} p_{i^n}|i^n\rangle\langle i^n|^X\otimes\rho_{i^n,j}^{AB_1B_2}\otimes q(j|i^n)|j\rangle\langle j|^C, \quad (56)$$

which represents the output of Alice's encoding operation for a given (unspecified) protocol of the form of Fig. 1. We can estimate

$$\frac{1}{n}\chi(\{\tilde{\varphi}_{i^n}^B,p_{i^n}\})\leqslant S(X:B_1B_2C) \quad \text{(by monotonicity of } \chi) \quad (57)$$

$$=S(X:B_2)+S(X:C|B_2)+S(X:B_1|B_2C) \quad (58)$$

$$\leqslant \log d_C+2\log d_Q, \quad (59)$$

using Lemma IV.3 (see below) twice and the fact that $S(X:B_2)=0$ since B_2 is maximally mixed for all i^n. On the other hand, applying the Fannes inequality [23] and the fidelity condition implies that

$$\frac{1}{n}\chi(\{\tilde{\varphi}_{i^n}^B,p_{i^n}\})\xrightarrow{\epsilon\to 0}\chi, \quad (60)$$

giving the constraint $\chi\leqslant R+2Q$.

Lemma IV.3. Let ρ be a tripartite density operator of the form

$$\rho=\sum_i p_i|i\rangle\langle i|^X\otimes\rho_i^{AB}, \quad (61)$$

where the states $\{|i\rangle^X\}$ are orthonormal and the p_i are probabilities. Then

$$S(X:A|B)\leqslant\min(\log\dim X,2\log\dim A). \quad (62)$$

Proof. We can expand $S(X:A|B)=S(X|B)-S(X|AB)$. By subadditivity of the von Neumann entropy, the first term is less than or equal to $S(X)$, which is in turn no more than $\log\dim X$. Moreover, because ρ is separable across the X/AB cut, $S(X|AB)\geqslant 0$. (This follows immediately from concavity of the entropy [24,25].)

To prove the second inequality, we expand the definition of $S(X:A|B)$ differently,

$$S(X:A|B)=S(A|B)_{\rho^{AB}}+\sum_i p_i S(A|B)_{\rho_i^{AB}}. \quad (63)$$

Using subadditivity of the von Neumann entropy again, $S(A|B)\leqslant S(A)$ for any density operator. $S(A)$, in turn, is always less than or equal to $\log\dim A$. ■

V. DISCUSSION

The problem we posed here, communication using noiseless classical and quantum channels in addition to maximally entangled states, is the natural setting in which to unify many pre-existing results on quantum-classical compression, remote state preparation and quantum state superdense coding. While our goal was to provide a unified synthesis of these disparate results, our conclusion was ultimately that the general problem can be understood in terms of those basic building blocks—the surface of optimal rate triples for the triple resource problem can be assembled by time sharing appropriately between protocols designed for the special cases. Such a neat resolution confirms the simplifying power of the resource-based approach and justifies viewing trade-off coding, remote state preparation and quantum state superdense coding as fundamental primitives instead of special cases of a more general problem.

ACKNOWLEDGMENTS

We thank Debbie Leung for many helpful conversations. The authors acknowledge the support of the U.S. National Science Foundation under Grant No. EIA-0086038. P.H. was also supported by the Sherman Fairchild Foundation.

[1] B. Schumacher, Phys. Rev. A **51**, 2738 (1995).
[2] M. Ohya and D. Petz, *Quantum Entropy and Its Use*, Texts and Monographs in Physics (Springer-Verlag, Berlin, 1993).
[3] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Phys. Rev. A **53**, 2046 (1996).
[4] G. Kuperberg, e-print quant-ph/0203105.
[5] C. H. Bennett, I. Devetak, A. Harrow, P. W. Shor, and Winter A (unpublished).
[6] H. Barnum, P. Hayden, R. Jozsa, and A. Winter, Proc. R. Soc. London, Ser. A **457**, 2019 (2001).
[7] M. Koashi and N. Imoto, Phys. Rev. A **66**, 022318 (2002).
[8] P. Hayden, R. Jozsa, and A. Winter, J. Math. Phys. **43**, 4404 (2002).
[9] H.-K. Lo, Phys. Rev. A **62**, 012313 (2000).
[10] C. H. Bennett, D. P. DiVincenzo, P. W. Shor, J. A. Smolin, B. M. Terhal, and W. K. Wootters, Phys. Rev. Lett. **87**, 077902 (2001).
[11] C. H. Bennett and S. Wiesner, Phys. Rev. Lett. **69**, 2881 (1992).
[12] A. Harrow, P. Hayden, and D. Leung, e-print quant-ph/0307221.
[13] A. S. Holevo, Probl. Peredachi Inf. **9**, 3 (1973).
[14] C. H. Bennett, P. Hayden, D. Leung, P. W. Shor, and A. Winter, e-print quant-ph/0307100 (2003).
[15] A. Harrow, e-print quant-ph/0307091.
[16] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, IEEE Trans. Inf. Theory **48**, 2637 (2002).
[17] M. Horodecki, Phys. Rev. A **61**, 052309 (2000).
[18] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. W. Schumacher, J. Phys. A **34**, 6767 (2001).

[19] A. Winter, e-print quant-ph/0208131.
[20] I. Devetak and T. Berger, Phys. Rev. Lett. **87**, 197901 (2001).
[21] B. Schumacher and M. D. Westmoreland, Phys. Rev. A **56**, 131–138 (1997).
[22] A. S. Holevo, IEEE Trans. Inf. Theory **44**, 269–273 (1998).
[23] M. Fannes, Commun. Math. Phys. **31**, 291 (1973).
[24] N. J. Cerf and C. Adami, Phys. Rev. A **60**, 893 (1999).
[25] P. Horodecki, R. Horodecki, and M. Horodecki, Acta Phys. Slov. **48**, 141 (1998).