Improving Zero-Error Classical Communication with Entanglement

Toby S. Cubitt,[1] Debbie Leung,[2] William Matthews,[2,*] and Andreas Winter[1,3]

[1]Department of Mathematics, University of Bristol, Bristol BS8 1TW, United Kingdom
[2]Institute for Quantum Computing, University of Waterloo, Waterloo N2L 3G1, Ontario, Canada
[3]Centre for Quantum Technologies, National University of Singapore, 2 Science Drive 3, Singapore 117542

(Received 21 February 2010; published 8 June 2010)

Given one or more uses of a classical channel, only a certain number of messages can be transmitted with zero probability of error. The study of this number and its asymptotic behavior constitutes the field of classical zero-error information theory. We show that, given a single use of certain classical channels, entangled states of a system shared by the sender and receiver can be used to increase the number of (classical) messages which can be sent without error. In particular, we show how to construct such a channel based on any proof of the Kochen-Specker theorem. We investigate the connection to pseudotelepathy games. The use of generalized nonsignaling correlations to assist in this task is also considered. In this case, an elegant theory results and, remarkably, it is sometimes possible to transmit information with zero error using a channel with no unassisted zero-error capacity.

DOI: 10.1103/PhysRevLett.104.230503 PACS numbers: 03.67.Hk, 03.67.Ac, 89.70.Kn

It is well known that if two parties share an entangled quantum state, they may be able to achieve tasks which would be otherwise impossible. For instance, without communicating they can violate Bell inequalities [1], and with classical communication they can teleport the state of a quantum system [2]. Here we show that quantum effects can sometimes give an advantage in the context of zero-error coding [3,4]: A classical channel $\mathcal{N}$ connects a sender (Alice) to a receiver (Bob). It has a finite number of inputs and outputs and its behavior is fully described by the conditional probability distribution over outputs given the input; i.e., it is discrete and memoryless. Given one use of $\mathcal{N}$, the maximum number of different messages Alice can send to Bob if there is to be no chance of an error is known as the one-shot zero-error capacity of $\mathcal{N}$.

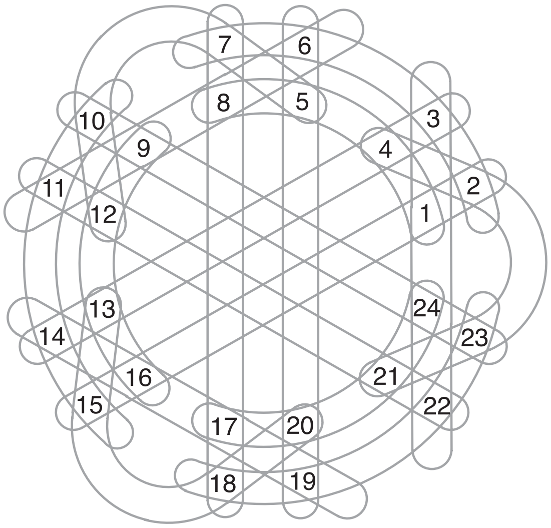
The main contribution of this Letter is to show that, for certain classical channels, entanglement between Alice and Bob can be used to increase the one-shot zero-error capacity for classical messages. This is in contrast to interesting recent work considering zero-error coding for classical and quantum data over quantum channels [5–8]. Recall that the use of entanglement [9] (and even nonsignaling correlations [10]) cannot increase the transmission rate if we only demand that the error rate goes to zero in the large block length limit: it remains equal to the normal Shannon capacity [11].

We briefly review classical zero-error coding, then we show how to construct classical channels where entanglement can increase the one-shot zero-error capacity. We then discuss the relationship of entanglement-assisted zero-error coding to "pseudotelepathy" games. After that, we upper bound this entanglement assistance by considering generalized nonsignaling correlations, giving a simple formula for the nonsignaling assisted zero-error capacity of any channel. This turns out to have an interesting relationship to classical results of Shannon from his original paper [3] on zero-error capacities.

Two input symbols of a channel are confusable if the corresponding distributions on output symbols overlap. Shannon introduced the confusability graph $G(\mathcal{N})$ of a classical channel $\mathcal{N}$: Its vertices are the set of input symbols and they are joined if and only if they are confusable. Classically, a zero-error code is a set of nonconfusable inputs. The one-shot zero-error capacity $c_0(\mathcal{N})$ of a channel $\mathcal{N}$ is simply the maximum size of such a set. In the language of graph theory, a maximum nonconfusable set of inputs is a maximum independent set of the confusability graph, and when Bob receives a channel output, the possible inputs are a clique in the confusability graph. A channel has no unassisted zero-error capacity if and only if its confusability graph is complete, i.e., all vertices are connected.

It is also useful to define the hypergraph of a channel: A hypergraph is just a set S (the vertices) and a set of subsets of S called the hyperedges. The hypergraph of a channel $\mathcal{N}$ has the set of inputs as vertices and one hyperedge for each of the outputs, which contains all the inputs that have a nonzero probability of causing that output; we denote it $H(\mathcal{N})$.

In this work we deal with correlations (bipartite conditional probability distributions) in the classes SR, SE, and NS: Correlations belong to SR if and only if they can be obtained using (classical) shared randomness (and local operations), to SE (shared entanglement) if and only if they can be realized by local operations on a shared quantum state, and to NS if and only if the correlation is nonsignaling (meaning that the marginal distribution on the output of each party is independent of the other party's input). Each class in this list strictly contains the previous one. We denote the maximum number of messages which can be



$1 : (1,0,0,0)$	$2 : (0,1,0,0)$	$3 : (0,0,1,0)$	$4 : (0,0,0,1)$
$5 : (0,1,1,0)$	$6 : (1,0,0,-1)$	$7 : (1,0,0,1)$	$8 : (0,1,-1,0)$
$9 : (1,1,1,1)$	$10 : (1,-1,1,-1)$	$11 : (1,-1,-1,1)$	$12 : (1,1,-1,-1)$
$13 : (1,-1,0,0)$	$14 : (1,1,0,0)$	$15 : (0,0,1,1)$	$16 : (0,0,1,-1)$
$17 : (-1,1,1,1)$	$18 : (1,1,1,-1)$	$19 : (1,-1,1,1)$	$20 : (1,1,-1,1)$
$21 : (1,0,1,0)$	$22 : (0,1,0,1)$	$23 : (1,0,-1,0)$	$24 : (0,1,0,-1)$

FIG. 1. A KS basis set of six bases for $\mathbb{C}^4$ is tabulated at the bottom of the figure, one basis per row. The vectors are presented as 4-tuples labeled by a number. The diagram represents a channel $\mathcal{N}$ with an input symbol for each vector in the set. It has an output symbol for each gray loop: on input x the output is drawn uniformly at random from those corresponding to the three loops which contain that x. Inputs are confusable if and only if corresponding vectors are orthogonal, so by Theorem 2, $c_0(\mathcal{N}) < 6$ (in fact it is 5), but $c_{\mathrm{SE}}(\mathcal{N}) \geq 6$. It is interesting to note that to send one of six symbols (with equal prior probabilities) by a single use of $\mathcal{N}$, the best unassisted code has error probability 1/18.

sent without error by a single use of $\mathcal{N}$ when any correlation in class Ω can be used by $c_\Omega(\mathcal{N})$. The corresponding limiting rate to send zero-error bits is $C_\Omega(\mathcal{N}) := \lim_{n\to\infty} \frac{1}{n} \log c_\Omega(\mathcal{N}^{\otimes n})$. A simple convexity argument shows that shared randomness between sender and receiver cannot help, so $c_{\mathrm{SR}}(\mathcal{N}) = c_0(\mathcal{N})$ for all channels. In contrast, we will next show how to construct channels $\mathcal{N}$ for which the number of messages which can be sent perfectly using entanglement, $c_{\mathrm{SE}}(\mathcal{N})$, is greater than $c_0(\mathcal{N})$.

Entanglement-assisted zero-error communication.—Given a classical channel $\mathcal{N}$ from Alice and Bob, with inputs X and outputs Y, how might they make use of entanglement to increase the number of messages which can be sent? Suppose that Alice wants to send one of q messages to Bob without error and that their entangled shared system is in state ρ_{AB}. She will perform some operations on her side of the entangled system, and conditioned on the outcomes of any classical measurements that she does, and on the message m that she wants to send, choose some input to $\mathcal{N}$. All of this can be represented by saying that she chooses one of q generalized measurements according to m, each with $|X|$ outcomes, to perform on her side of the state, and then uses the outcome k as input to $\mathcal{N}$. Since the residual state on Alice's side is irrelevant to Bob's ability to decode the message, the encoding is fully specified by the positive operator valued measures $\{E_1^{(m)}, \ldots, E_k^{(m)}\}$ for $m \in [q] := \{1, \ldots, q\}$ corresponding to the q different generalized measurements.

If Alice sends message m, then with probability $p_k^{(m)}$, Alice inputs k and the residual state of Bob's system is $\rho_k^{(m)} = (\mathrm{Tr}_A E_k^{(m)} \otimes \mathbb{1}\rho)/p_k^{(m)}$. Letting $\beta_k^{(m)} := p_k^{(m)} \rho_k^{(m)}$, for all messages m: $\sum_k \beta_k^{(m)} = \mathrm{Tr}_A \rho_{AB} =: \rho_B$, reflecting the fact that without information from the classical channel, Bob has no idea which message Alice sent (i.e., causality). Conversely, any set of positive operators $\beta_k^{(m)}$ which satisfies this condition for some ρ_B can be realized by a suitable choice of ρ_{AB} and generalized measurements. Now, including the state of the channel output (we label the system C) as well as his half of the entangled system, Bob's state after receiving the channel output $y \in Y$ is $\sigma_m := \sum_{x\in X, y\in Y} \mathcal{N}(y|x)|y\rangle\langle y|_C \otimes \beta_x^{(m)}$. The encoding works if and only if Bob can distinguish perfectly between all the σ_m, i.e., for all m, $m' \in [q]$: $0 = \mathrm{Tr}\sigma_m\sigma_{m'} = \sum_{x,x'\in X \text{ confusable}} [\sum_y \mathcal{N}(y|x)\mathcal{N}(y|x')] \mathrm{Tr}\beta_x^{(m)}\beta_{x'}^{(m')}$. We therefore have the following.

Theorem 1. For any channel $\mathcal{N}$ with inputs X and outputs Y, $c_{\mathrm{SE}}(\mathcal{N}) = q(G(\mathcal{N}))$, where $q(G(\mathcal{N}))$ is the maximum integer q such that there exists a density matrix ρ_B and positive semidefinite operators $\beta_x^{(m)}$ for all $m \in [q]$, $x \in X$, on some Hilbert space such that for all m, $\sum_{x\in X} \beta_x^{(m)} = \rho_B$, and

$$\forall\ m \neq m' \quad \forall \text{ confusable } x, x' \qquad \mathrm{Tr}\beta_x^{(m)}\beta_{x'}^{(m')} = 0.$$

In particular, $c_{\mathrm{SE}}(\mathcal{N})$ depends only on $G(\mathcal{N})$. ■

In light of this fact, it is clear that if a channel has no unassisted zero-error capacity, entanglement cannot change this. Otherwise, entanglement would allow perfect communication over the completely noisy channel, in violation of causality!

However, there are some channels for which $c_{\mathrm{SE}} > c_0 > 0$. Examples of such channels can be constructed from proofs of the Kochen-Specker (KS) theorem [12]: We call a family $\{B_m\}_{m=1}^q$ of complete orthogonal bases B_m of $\mathbb{C}^d$ a KS basis set if it is impossible to select one vector from each basis such that no two are orthogonal. That such sets exist is a corollary of the KS theorem [12].

Theorem 2. For any KS basis set $Z = \{B_m\}_{m=1}^q$ in $\mathbb{C}^d$ consisting of q orthogonal bases, one can construct a classical channel $\mathcal{N}$ with $c_0(\mathcal{N}) < q$ and $c_{\mathrm{SE}}(\mathcal{N}) \geq q$.

Proof. Let us write $B_m = \{\psi_{m1}, \ldots, \psi_{md}\}$. We can construct a channel $\mathcal{N}_Z$ with inputs in $[q] \times [d]$ such that a pair of inputs (m, j), (m', j') are confusable if and only if the corresponding vectors ψ_{mj} and $\psi_{m'j'}$ are orthogonal. (In general there are many ways to do this, and any one will do. For instance, one can add an output symbol for each orthogonal pair which can be activated by both inputs in that pair but no others.) $G(\mathcal{N})$ has an edge between inputs if and only if the corresponding vectors are orthogonal. As such, the vertices of G can be partitioned into q cliques of size d, corresponding to the q bases of Z, so the independence number of G is certainly no larger than q. If there was an independent set of size q in G it would have to have exactly one vertex in each of the q cliques, but this would select one vector in each of the q bases such that no two are orthogonal, contradicting the assumption on Z. Therefore, $c_0(\mathcal{N}) < q$.

To send q messages using entanglement, Alice and Bob can use a maximally entangled state of rank d: to send m, Alice measures her side of the state in the bases B_m and obtains the outcome j (at random). She inputs (m, j) to the channel. Bob's output tells him that Alice's input was in some particular mutually confusable subset, but by construction, these inputs correspond to mutually orthogonal residual states of his subsystem, so he can perform a projective measurement to determine precisely which input Alice made to the classical channel, and hence which of the q messages she chose to send, with certainty. ■

In Fig. 1 we give an example of a KS basis set derived from a proof of the KS theorem due to Peres [13].

Relationship to pseudotelepathy games.—This increase of the one-shot zero-error capacity is an example of performing a classical task without error using entanglement, which becomes impossible without the entanglement. This phenomenon might sound familiar to those who have encountered pseudotelepathy games (hereafter PT games) [14]. The difference is that in these games Alice and Bob are not allowed to communicate with each other at all, but instead communicate with a verifier who sends them questions and then decides whether or not they win the game based on their replies.

To be precise, in this context a "game" $\mathfrak{g}$ consists of questions a and b (drawn according to a fixed distribution $p(a, b)$) to Alice and Bob, respectively, who reply with answers α and β. These are accepted with probability $A(a, b, \alpha, \beta)$, A also being a fixed distribution. The probability of acceptance ("winning") is given by

$$\mathfrak{g}(s) := \sum_{a,b,\alpha,\beta} A(a, b, \alpha, \beta)p(a, b)s(\alpha, \beta|a, b),$$

where the strategy $s(r|q)$ is a correlation describing the responses r of the provers to questions q. Note that $\mathfrak{g}(s)$ is a linear function of s. We call the strategy s "perfect" (for the game $\mathfrak{g}$) if and only if $\mathfrak{g}(s) = 1$. Typically we are interested in the best winning probability which can be achieved if the strategy is restricted to some class of correlations like NS or SE. A PT game is a game $\mathfrak{g}$ which can be won with certainty by a strategy in SE but cannot be won with certainty by any strategy in SR.

Proposition 3. For any channel $\mathcal{N}$ with inputs X and outputs Y, and integer n, there exists a natural game $\mathfrak{g}$ such that $\mathfrak{g}$ has a perfect strategy in the class of correlations Ω if and only if $c_\Omega \geq n$.

Proof. In the game $\mathfrak{g}$, the verifier sends Alice $m \in [n]$ and Bob $y \in Y$ drawn independently and uniformly at random. Alice sends back an answer $x \in X$ and Bob replies with $\hat{m} \in [n]$. If $\mathcal{N}(y|x) > 0$, they win the game if and only if $m = \hat{m}$. Otherwise, they always win the game. A strategy s is perfect for this game if and only if $\sum_{x,y} \mathcal{N}(y|x)s(x, \hat{m}|m, y) = \delta_{m\hat{m}}$. Therefore, there is a perfect strategy for $\mathfrak{g}$ in Ω if and only if $c_\Omega(\mathcal{N}) \geq m$. ■

This means that, in order to give an advantage for zero-error coding over SR, a correlation in SE must also be able to win a particular PT game with certainty (and hence sit on the boundary of the nonsignaling polytope).

Nonsignaling assisted zero-error capacity and exact simulation.—While all correlations which can be realized by measurements on entangled states are nonsignaling, the converse is not true, as in the case of the Popescu-Rohrlich box [15]. Consequently, we can study nonsignaling assisted protocols to find upper bounds for entanglement assistance, but this study also leads to a beautifully simple theory of nonsignaling assisted zero-error communication.

Recalling the definition of a hypergraph, the fractional-packing number $\alpha^*(H)$ of a hypergraph H [16] on vertices X is the maximum value of $\sum_{x\in X} v(x)$ where $v:X \to [0, 1]$ weights the vertices subject to the constraint that for all hyperedges S of H, $\sum_{x\in S} v(x) \leq 1$.

Theorem 4. For a classical channel $\mathcal{N}$ with hypergraph $H(\mathcal{N})$,

$$c_{\mathrm{NS}}(\mathcal{N}) = \lfloor \alpha^*(H(\mathcal{N}))\rfloor,$$

where $\alpha^*(H(\mathcal{N}))$ is the fractional-packing number of $H(\mathcal{N})$.

Furthermore, since the function α^* is multiplicative, in the sense that $\alpha^*(H(\mathcal{N}_1 \otimes \mathcal{N}_2)) = \alpha^*(H(\mathcal{N}_1)) \times \alpha^*(H(\mathcal{N}_2))$, the NS-assisted zero-error capacity of $\mathcal{N}$ is

$$C_{\mathrm{NS}}(\mathcal{N}) = \log\alpha^*(H(\mathcal{N})),$$

which is additive: $C_{\mathrm{NS}}(\mathcal{N}_1 \otimes \mathcal{N}_2) = C_{\mathrm{NS}}(\mathcal{N}_1) + C_{\mathrm{NS}}(\mathcal{N}_2)$.

To get the best upper bounds on entanglement-assisted zero-error communication using this result, we should minimize over all hypergraphs with the same confusability graph G as the channel in question, because c_{SE} depends only on G (see Theorem 1).

The proof of Theorem 4 is given in [10]. With one interesting proviso: the nonsignaling assisted zero-error capacity $C_{\mathrm{NS}}(\mathcal{N})$ is the same as the feedback-assisted zero-error capacity of the channel $C_{0\mathrm{F}}(\mathcal{N})$, as derived by Shannon in his seminal paper [3]. The proviso applies only when the unassisted zero-error capacity is zero: Then C_{NS} can be positive, whereas $C_{0\mathrm{F}}$ is always zero.

Channel simulation and reversibility.—One can also consider the "reverse" problem to zero-error coding [10], and ask what is the minimum identity channel needed, given correlations in Ω, to simulate one (or more) uses of some noisy channel $\mathcal{N}$ exactly (in the sense of exactly reproducing the conditional probability distribution of outputs given inputs). We denote this minimum required number of messages by $k_\Omega(\mathcal{N})$, and the Ω-assisted simulation cost of $\mathcal{N}$ by $K_\Omega(\mathcal{N}) := \lim_{n\to\infty}\frac{1}{n}\log k_{\mathrm{NS}}(\mathcal{N}^{\otimes n})$. Again, the structure of the set of all nonsignaling correlations results in a very simple formula for $k_{\mathrm{NS}}(\mathcal{N})$: For any channel $\mathcal{N}$ with inputs X and outputs Y, $k_{\mathrm{NS}}(\mathcal{N}) = \lceil\sum_y \max_x \mathcal{N}(y|x)\rceil$, and since the sum here is multiplicative under tensor products of the channel matrix, $K_{\mathrm{NS}}(\mathcal{N}) = \log[\sum_y \max_x \mathcal{N}(y|x)]$.

While we have found examples showing an arbitrarily large gap between $k_{\mathrm{NS}}(\mathcal{N})$ and $k_{\mathrm{SR}}(\mathcal{N})$, the gap disappears in the limit of many channel uses: $K_{\mathrm{SR}}(\mathcal{N}) = K_{\mathrm{SE}}(\mathcal{N}) = K_{\mathrm{NS}}(\mathcal{N})$ [10].

Curiously, a kind of combinatorial zero-error reversibility exists when nonsignaling correlations are freely available: For a given channel hypergraph H, the NS-assisted zero-error capacity of channels with hypergraph H is equal to the infimum of the NS-assisted simulation cost for channels with hypergraph H [10], in analogy to the direct and reverse Shannon theorems [9,11].

Conclusion.—We have shown that entanglement can sometimes be used to increase the number of classical messages which can be sent perfectly over classical channels. To upper bound this quantum advantage, we have given a simple formula for the nonsignaling assisted capacity as a linear program. These discoveries present many new questions: First, can entanglement improve the asymptotic zero-error capacity, compared to no assistance, as we have seen NS correlations can? More generally, can we find a simple expression for the entanglement-assisted zero-error capacity in the one-shot or asymptotic case? Note that while the best general upper bound known on C_0 is given by Lovász's famous ϑ function [17], it was very recently found (indeed prompted by our Theorem 2) that ϑ is still an upper bound on C_{SE} [18,19]. Can we find simpler, less contrived, examples of channels where $c_{\mathrm{SE}} > c_0$? In another direction, the relationship between KS theorems and PT games has been studied in [20]. We found connections between the entanglement-assisted zero-error phenomenon and both of these topics, but left open the development of a fuller understanding of the relationships among the three.

We would like to thank Nicolas Brunner, Runyao Duan, Tsuyoshi Ito, Ashley Montanaro, Marcin Pawłowski, Paul Skrzypczyk, and Stephanie Wehner for discussions. T. S. C. is supported by a Leverhulme early-career grant and the EC project "QAP" (Contract No. IST-2005-15848). D. L. was funded by CRC, CFI, ORF, CIFAR, NSERC, and QuantumWorks. W. M. acknowledges the support of NSERC and QuantumWorks. A. W. is supported by the EC, the U.K. EPSRC, the Royal Society, and a Philip Leverhulme grant. The CQT is funded by the Singapore MoE and the NRF as part of the Research Centres of Excellence programme. We are grateful for the hospitality of the Kavli Institute for Theoretical Physics at UCSB, where a large part of this research was performed. This research was supported in part by the NSF under Grant No. PHY05-51164.

*will@northala.net

[1] J. S. Bell, Physics **1**, 195 (1964).

[2] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).

[3] C. E. Shannon, IRE Trans. Inf. Theory **2**, 8 (1956).

[4] J. Körner and A. Orlitsky, IEEE Trans. Inf. Theory **44**, 2207 (1998).

[5] R. A. C. Medeiros, R. Alleaume, G. Cohen, and F. M. de Assis, arXiv:quant-ph/0611042.

[6] R. Duan, arXiv:0906.2527.

[7] T. S. Cubitt, J. Chen, and A. W. Harrow, arXiv:0906.2547.

[8] T. S. Cubitt and G. Smith, arXiv:0912.2737.

[9] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, IEEE Trans. Inf. Theory **48**, 2637 (2002).

[10] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter, arXiv:1003.3195.

[11] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948); **27**, 623 (1948).

[12] S. Kochen and E. P. Specker, J. Math. Mech. **17**, 59 (1967).

[13] A. Peres, J. Phys. A **24**, L175 (1991).

[14] G. Brassard, R. Cleve, and A. Tapp, Phys. Rev. Lett. **83**, 1874 (1999).

[15] S. Popescu and D. Rohrlich, Found. Phys. **24**, 379 (1994).

[16] See, for example, A. Schrijver, *Combinatorial Optimization: Polyhedra and Efficiency* (Springer-Verlag, Berlin, 2003), Vol. 1, p. 1429, where it is called the "fractional stable set number" of a hypergraph.

[17] L. Lovász, IEEE Trans. Inf. Theory **25**, 1 (1979).

[18] S. Beigi, arXiv:1002.2488.

[19] R. Duan, S. Severini, and A. Winter, arXiv:1002.2514.

[20] R. Renner and S. Wolf, in *Proceedings of the International Symposium on Information Theory, Chicago, 2004* (IEEE, Piscataway, NJ, 2004), p. 322.