Are There Phase Transitions in Information Space?

Jonathan Oppenheim,[1,2] Michał Horodecki,[2] and Ryszard Horodecki[2]
[1]*Racah Institute of Theoretical Physics, Hebrew University of Jerusalem, Givat Ram, Jerusalem 91904, Israel*
[2]*Institute of Theoretical Physics and Astrophysics, University of Gdańsk, Gdańsk, Poland*
(Received 30 July 2002; published 10 January 2003)

The interplay between two basic quantities — quantum communication and information — is investigated. Quantum communication is an important resource for quantum states shared by two parties and is directly related to entanglement. Recently, the amount of *local information* that can be drawn from a state has been shown to be closely related to the nonlocal properties of the state. Here we consider both formation and extraction processes, and analyze informational resources as a function of quantum communication. The resulting diagrams in *information space* allow us to observe phaselike transitions when correlations become classical.

DOI: 10.1103/PhysRevLett.90.010404 PACS numbers: 03.65.Ud, 03.67.–a

Quantum communication (QC) — the sending of qubits between two parties — is a primitive concept in quantum information theory. Entanglement cannot be created without it, and conversely, entanglement between two parties can be used to communicate quantum information through teleportation [1]. The amount of quantum communication needed to prepare a state and the amount of quantum communication that a state enables one to perform are fundamental properties of states shared by two parties. This amount is identical to the entanglement cost E_c [2] and entanglement of distillation E_D [3], respectively. Perhaps surprisingly these two quantities are different [4]. There are even states which are "bound" in the sense that quantum communication is needed to create it, but nothing can be obtained from it [5]. Yet QC is a distinct notion from entanglement. For example, one may need to use a large amount of QC while creating some state of low E_c, in order to save some other resource. In the present Letter we consider such a situation. The second primitive resource of interest will be *information* which quantifies how pure a state is. The motivation comes from (both classical or quantum) thermodynamics: it is known that bits that are in a pure state can be used to extract physical work from a single heat bath [6], and conversely work is required to reset mixed states to a pure form [7,8].

For distant parties, in order to use information to perform such tasks, it must first be localized. In [9] we considered how much information (i.e., pure states) can be localized (or drawn) from a state shared between two parties. Thus far, the amount of information needed to prepare a state has not been considered, a possible exception being in [10] where it was noted that there was a thermodynamical irreversibility between preparation and measurement for ensembles of certain pure product states. However, given the central role of quantum communication and information, it would be of considerable importance to understand the interplay between these two primitive resources. In this Letter, we attempt to lay the foundation for this study by examining how much information is needed to prepare a state and how much can be extracted from it as a function of quantum communication. For a given state, this produces a curve in *information space*. The shapes of the curve fall into a number of distinctive categories which classify the state and only a small number of parameters are needed to characterize them. The curves for pure states can be calculated exactly, and they are represented by a one parameter family of lines of constant slope. The diagrams exhibit features reminiscent of thermodynamics, and phaselike transitions (cf. [11]) are observed.

An important quantity that emerges in this study is the *information surplus* Δ_f. It quantifies the additional information that is needed to prepare a state when quantum communication resources are minimized. Δ_f tells us how much information is dissipated during the formation of a state and is therefore closely related to the irreversibility of state preparation and, therefore, to the difference between the entanglement of distillation and entanglement cost. When it is zero, there is no irreversibility in entanglement manipulations. Examples of states with $\Delta_f = 0$ include pure states, and states with an optimal decomposition [3] which is locally orthogonal.

Consider two parties in distant labs who can perform local unitary operations, dephasing [12], and classical communication. It turns out to be simpler to substitute measurements with dephasing operations, since we no longer need to keep track of the informational cost of resetting the measuring device. (This cost was noted by Landauer [7] and used by Bennett [8] to exorcize Maxwell's demon.) The classical communication channel can also be thought of as a dephasing channel. Finally, we allow Alice and Bob to add noise (states which are proportional to the identity matrix) since pure noise contains no information. Note that we are interested only in accounting for resources that are "used up" during the preparation procedure. For example, a pure state which is

 0031-9007/03/90(1)/010404(4)$20.00

used and then reset to its original state at the end of the procedure does not cost anything.

Consider now the information extraction process of [9]. If the two parties have access to a quantum channel and share a state ϱ_{AB}, they can extract all the information from the state

$$I = n - S(\varrho_{AB}), \tag{1}$$

where n is the number of qubits of the state, and $S(\varrho_{AB})$ is its von Neumann entropy. Put another way, the state can be compressed, leaving I pure qubits. However, if two parties do not have access to a quantum channel and can perform only local operations and communicate classically (LOCC), then, in general, they will be able to draw less local information from the state. In [9] we defined the notion of the *deficit* Δ to quantify the information that can no longer be drawn under LOCC. For pure states, it was proven that Δ was equal to the amount of distillable entanglement in the state.

Let us now turn to formation processes and define $\Delta_f(Q)$ as follows. Given an amount of quantum communication Q, the amount of information (pure states) needed to prepare the state ϱ_{AB} under LOCC is given by $I_f(Q)$. Clearly at least E_c bits of quantum communication are necessary. In general, $I_f(Q)$ will be greater than the information content I. The surplus is then

$$\Delta_f(Q) = I_f(Q) - I. \tag{2}$$

The two end points are of particular interest, i.e., $\Delta_f \equiv \Delta_f(E_c)$, where quantum communication is minimized, and $\Delta_f(E_r) = 0$, where we use the quantum channel enough times that $I_f(E_r) = I$. Clearly $E_c \leq E_r \leq \min\{S(\rho_A), S(\rho_B)\}$, where ρ_A is obtained by tracing out on Bob's system. This rough bound is obtained by noting

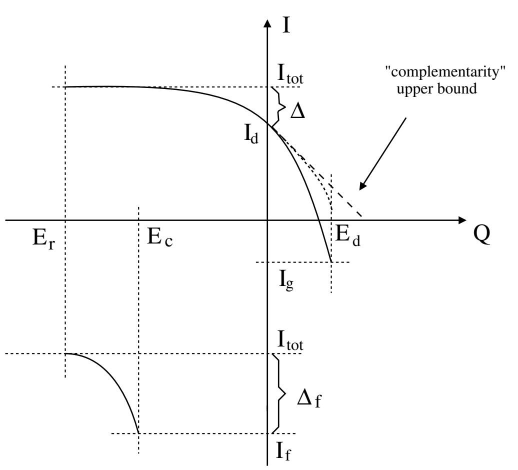

FIG. 1. Formation and extraction curves for a generic mixed state. The short-dashed line represents the variant where information can be extracted from the "garbage" left after entanglement distillation ($I_g > 0$). In general, the curves need not be smooth. The formation curve is in the lower left quadrant.

that at a minimum, Alice or Bob can prepare the entire system locally and then send it to the other party through the quantum channel (after compressing it). We obtain a tight bound later in this Letter.

The general procedure for state preparation is that Alice uses a *classical instruction set* (ancilla in a classical state) with probability distribution matching that of the decomposition which is optimal for a given Q. Since the instruction set contains classical information, it can be passed back and forth between Alice and Bob. Additionally they need n pure standard states. The pure states are then correlated with the ancilla and then sent. The ancilla need not be altered by this procedure and can be reset and then reused, and so at worse we have $I_f \equiv I_f(E_c) \leq n$ and $\Delta_f \leq S(\varrho_{AB})$. Shortly we describe how one can do better by extracting information from correlations between the ancilla and the state.

The pairs $(Q, I_f(Q))$ form curves in *information space*. In Fig. 1 we show a typical curve which we now explain. Since we compare the formation curves to extraction curves, we adopt the convention that $I_f(Q)$ and Q are plotted on the negative axis since we are using up resources. It can be shown that $I_f(Q)$ is concave, monotonic, and continuous. To prove concavity, we take the limit of many copies of the state ϱ_{AB}. Then given any two protocols, we can always toss a coin weighted with probabilities p and $1-p$ and perform one of the protocols with this probability. There will always be a protocol which is at least as good as this. Monotonicity is obvious (additional quantum communication can only help), and continuity follows from monotonicity, and the existence of the probabilistic protocol.

The probabilistic protocol can be drawn as a straight line between the points $(E_r, I_f(E_r))$ and $(E_c, I_f(E_c))$. There may, however, exist a protocol which has a lower $I_f(Q)$ than this straight line; i.e., the curve $I_f(Q)$ satisfies

$$I_f(Q) \leq I + \frac{I(E_r) - I_f(E_c)}{E_c - E_r}(Q - E_r). \tag{3}$$

Let us now look at extraction processes. The idea is that we draw both local information (pure separable states) and distill singlets. The singlets allow one to perform teleportations so that we are, in fact, extracting the potential to use a quantum channel. We can also consider the case where we use the quantum channel to assist in the information extraction process. We can therefore write the extractable information $I_l(Q)$ as a function of Q. When Q is positive, we distill singlets at the same time as drawing information, and when Q is negative, we are using the quantum channel Q times to assist in the extraction (see also Fig. 1).

There appear to be at least three special points on the curve. The first is the point $I_l \equiv I_l(0)$. This is considered in [9] when we draw maximal local information without the aid of a quantum channel. Another special point is the usual entanglement distillation procedure $I_g = I_l(E_D)$.

The quantity I_g is the amount of local information extractable from the "garbage" left over from distillation. I_g can be negative as information may need to be added to the system in order to distill all the available entanglement. Finally, $I = I_l(E_r)$ is the point where we use the quantum channel enough times that we can extract all the available information. This is the same number of times that the quantum channel is needed to prepare the state without any information surplus since both procedures are now reversible.

Just as with the formation curve, $I_l(Q)$ is convex, continuous, and monotonic. For $Q \geq 0$ there is an upper bound on the extraction curve due to the classical/quantum complementarity of [13].

$$I + Q \leq I_l. \tag{4}$$

It arises because one bit of local information can be drawn from each distilled singlet, destroying the entanglement. One might suppose that the complementarity relation (4) can be extended into the region $Q < 0$. Perhaps surprisingly, this is not the case, and we have found that a small amount of quantum communication can free up a large amount of information. In Fig. 2(a) we plot the region occupied by pure states. For extraction processes, pure states saturate the bound of Eq. (4) [13]. For formation processes they are represented as points.

In general, if $\Delta_f = 0$ then $E_c = E_D$. Examples include mixtures of locally orthogonal pure states[14]. The converse is not true, at least for single copies, as there are separable states such as those of [10] which have $\Delta_f \neq 0$ and $\Delta \neq 0$.

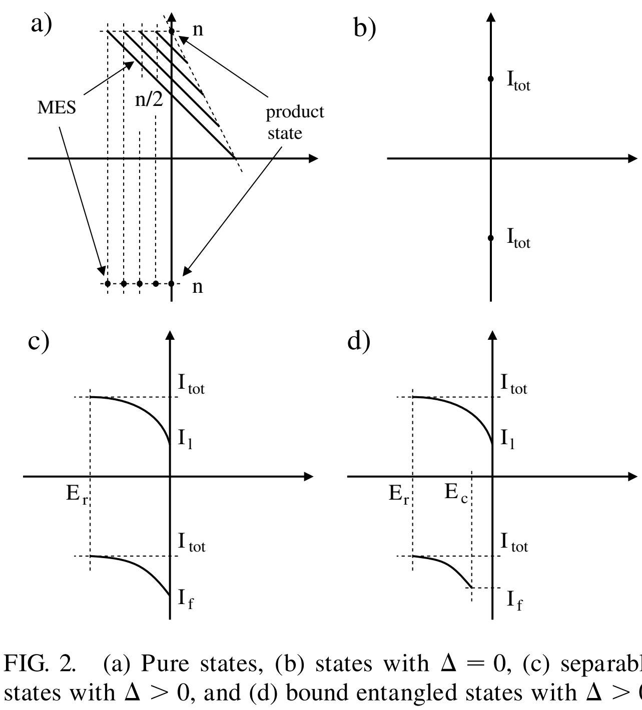

FIG. 2. (a) Pure states, (b) states with $\Delta = 0$, (c) separable states with $\Delta > 0$, and (d) bound entangled states with $\Delta > 0$.

It therefore appears that Δ_f is not a function of the entropy of the state, or of the entanglement, but rather, shows how chaotic the quantum correlations are. It can also be thought of as the information that is dissipated during a process, while Δ can be thought of as the *bound information* which cannot be extracted under LOCC. Figures 2(b)–2(d) show the curves for some different types of states. It is interesting the extent to which one can classify the different states just by examining the diagrams in information space.

The quantities we are discussing have (direct or metaphoric) connections with thermodynamics. Local information can be used to draw physical work, and quantum communication has been likened to quantum logical work [14]. One is therefore tempted to investigate whether there can be some effects similar to phase transitions. Indeed, we will demonstrate such an effect for a family of mixed states where the transition is of second order, in that the derivative of our curves will behave in a discontinuous way.

To this end we need to know more about E_r and I_f. Consider the notion of *LOCC orthogonality* (cf. [14]). One says that ϱ^i is LOCC orthogonal if by LOCC Alice and Bob can transform $\sum_i p_i |i\rangle_{A'}\langle i| \varrho^i_{AB}$ into $|0\rangle_{A'}\langle 0| \otimes \sum_i p_i \varrho^i_{AB}$ and vice versa; $|i\rangle_{A'}$ is the basis of Alice's ancilla. In other words, Alice and Bob are able to correlate the state ϱ_i to orthogonal states of a local ancilla as well as reset the correlations. Consider a state ϱ_{AB} that can be LOCC decomposed; i.e., it is a mixture of LOCC-orthogonal states $\varrho = \sum_i p_i \varrho_i$. The decomposition suggests a scheme for reversible formation of ϱ. Alice prepares locally the state $\varrho_{A'AB} = \sum_i p_i |i\rangle_{A'}\langle i| \varrho^i_{AB}$. This costs $n_{A'AB} - S(\varrho_{A'AB})$ bits of information. Conditioned on i, Alice compresses the halves of ϱ_i and sends them to Bob via a quantum channel. This costs $\sum_i p_i S(\varrho_B)$ qubits of quantum communication. Then, since the ϱ_i are LOCC orthogonal, Alice and Bob can reset the ancilla and return $n_{A'}$ bits. One then finds, in this protocol, formation costs $n_{AB} - S(\varrho_{AB})$ bits; hence it is reversible. Consequently, $E_r \leq \sum_i p_i S(\varrho_B)$; hence

$$E_r(\varrho_{AB}) \leq \inf \min_X \sum_i p_i S(\varrho^i_X), \qquad X = A, B, \tag{5}$$

where the infimum runs over all LOCC-orthogonal decompositions of ϱ_{AB}.

We can also estimate I_f by observing that the optimal decomposition for entanglement cost is compatible with LOCC-orthogonal decompositions; i.e., it is of the form $\{p_i q_{ij}, \psi_{ij}\}$, where $\sum_j q_{ij} |\psi_{ij}\rangle\langle\psi_{ij}| = \varrho_i$. Now, Alice prepares locally the state $\varrho_{A'A''AB} = \sum_i p_i q_{ij} |i\rangle_{A'}\langle i| \otimes |j\rangle_{A''}\langle j| \otimes |\psi_{ij}\rangle_{AB}\langle\psi_{ij}|$. Conditional on ij, Alice compresses the halves of ψ_{ij}'s and sends them to Bob. This costs on average E_c qubits of communication. So far Alice borrowed $n_{A'A''AB} - S(\varrho_{A'A''AB})$ bits. Alice and Bob then reset and return ancilla A' (this is possible due to LOCC orthogonality of ϱ_i) and also return ancilla A'' without resetting. The amount of bits used is $n_{AB} - [S(\varrho_{AB}) - \sum_i p_i S(\varrho_i)]$, giving

$$\Delta_f \leq \inf \sum_i p_i S(\varrho_i) \leq S(\varrho), \tag{6}$$

where, again, the infimum runs over the same set of decompositions as in Eq. (5) providing a connection between Δ_f and E_r. In the procedure above, collective operations were used only in the compression stage. In such a regime the above bounds are tight. There is a question, whether by some sophisticated collective scheme, one can do better. We conjecture that it is not the case, supported by the remarkable result of [15]. The authors show that an ensemble of nonorthogonal states cannot be compressed to less than $S(\varrho)$ qubits even at the expense of classical communication. In our case orthogonality is replaced by LOCC orthogonality, and classical communication by resetting. We thus assume equality in Eqs. (5) and (6). Thus for a state that is not LOCC decomposable (this holds for all two qubit states that do not have a product eigenbasis) we have $\Delta_f = S(\varrho_{AB})$, $E_r = \min\{S(\varrho_A), S(\varrho_B)\}$.

Having fixed two extremal points of our curves, let us see if there is a protocol which is better than the probabilistic one (a straight line on the diagram). We need to find some intermediate protocol which is cheap in both resources. The protocol is suggested by the decomposition $\varrho = \sum_i p_i \varrho_i$, where ϱ_i are themselves LOCC-orthogonal mixtures of *pure* states. Thus Alice can share with Bob each ϱ_i at a communication cost of $Q = \sum_i p_i E_c(\varrho_i)$. If the states ϱ_i are not LOCC orthogonal, Alice cannot reset the instruction set, so that the information cost is $I = n - \sum_i p_i S(\varrho_i)$. We now show by example that this may be a very cheap scenario. Consider

$$\varrho = p|\psi_+\rangle\langle\psi_+| + (1-p)|\psi_-\rangle\langle\psi_-|, \qquad p \in [0, \tfrac{1}{2}] \tag{7}$$

with $\psi_\pm = 1/\sqrt{2}(|00\rangle \pm |11\rangle)$. When $p \neq 0$ we have $E_r = 1$, $I_f = 2$, $E_c = H(\frac{1}{2} + \sqrt{p(1-p)})$ [4], where $H(x) = -x \log x - (1-x)\log(1-x)$ is the binary entropy; thus our extreme points are $(1, 2 - H(p))$ and $(E_c, 2)$. For $p = 0$ the state has $\Delta = 0$; hence the formation curve is just a point. We can, however, plot it as a line $I = 1$ (increasing Q will not change I). Now, we decompose the state as $\varrho = 2p\varrho_s + (1-2p)|\psi_-\rangle\langle\psi_-|$, where ϱ_s is an equal mixture of LOCC-orthogonal states $|00\rangle$ and $|11\rangle$. The intermediate point is then $(1-2p, 2-2p)$. A family of diagrams with changing parameter p is plotted in Fig. 3. The derivative $\chi(Q) = \frac{\partial Q}{\partial I}$ has a singularity at $p = 1/2$. Thus we have something analogous to a second order phase transition. The quantity $\chi(Q)$ may be analogous to a quantity such as the magnetic susceptibility. The transition is between states having $\Delta = 0$ (classically correlated) [9] and states with $\Delta \neq 0$ which contain quantum correlations. It would be interesting to explore these transitions and diagrams further, and also the trade-off between information and quantum communication. To this end, the quantity $\Delta_f(Q) + Q - E_c$ appears to express this trade-off. Finally, we hope that the presented approach may clarify an intriguing notion in quantum information theory, known as the *thermodynamics of entanglement* [14,16,17].

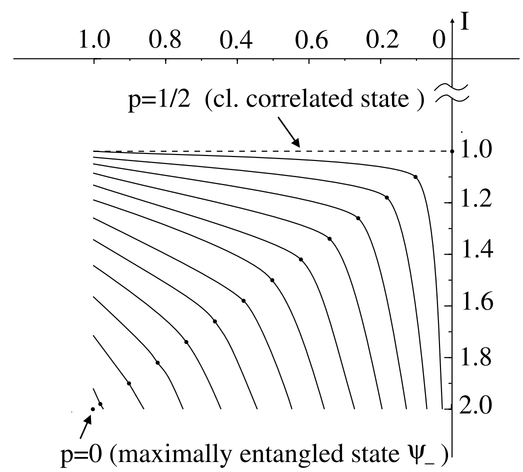

FIG. 3. "Phase transition" in the family of states of Eq. (7).

This work was supported by EU Grant EQUIP, Contract No. IST-1999-11053. J. O. acknowledges Lady Davis, and Grant No. 129/00-1 of the ISF. M. H. and R. H. thank C. Tsallis for discussions.

[1] C. H. Bennett *et al.*, Phys. Rev. Lett. **70**, 1895 (1993).
[2] P. Hayden, M. Horodecki, and B. Terhal, J. Phys. A **34**, 6891 (2001).
[3] C. H. Bennett *et al.*, Phys. Rev. A **53**, 2046 (1996).
[4] G. Vidal, W. Dür, and J. I. Cirac, Phys. Rev. Lett. **89**, 027901 (2002).
[5] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **80**, 5239 (1998).
[6] L. Szilard, Z. Phys. **53**, 840 (1929).
[7] R. Landauer, IBM J. Res. Dev. **5**, 183 (1961).
[8] C. H. Bennett, Int. J. Theor. Phys. **21**, 905 (1982).
[9] J. Oppenheim, M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **89**, 180402 (2002).
[10] C. H. Bennett *et al.*, Phys. Rev. A **59**, 1070 (1999).
[11] D. Aharonov, quant-ph/9910081.
[12] Dephasing can be written as $\sum_i P_i \rho_{AB} P_i$, where the P_i are orthogonal local projection operators.
[13] J. Oppenheim, K. Horodecki, M. Horodecki, P. Horodecki, and R. Horodecki, quant-ph/0207025.
[14] P. Horodecki, M. Horodecki, and R. Horodecki, Acta Phys. Slovaca **48**, 144 (1998).
[15] P. Hayden, R. Jozsa, and A. Winter, quant-ph/0204038
[16] S. Popescu and D. Rohrlich, Phys. Rev. A **56**, R3319 (1997).
[17] M. Horodecki, J. Oppenheim, and R. Horodecki, quant-ph/0207177.