Consider now the information extraction process of a physical system. The extraction process, as before, can be decomposed into two nested processes: the information transmission process and the information extraction process. The information transmission process is the process by which information is transmitted from one system to another. The information extraction process is the process by which information is extracted from a system.

If we assume that the information transmission process is a linear process, then the information extraction process can be described as a linear transformation. This linear transformation is a function of the input information and the output information. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.

In this paper, we will consider the case where the information transmission process is a linear process. We will assume that the information extraction process is a linear process as well. The output information is a linear combination of the input information and the parameters of the linear transformation.

The output information is a function of the input information and the parameters of the linear transformation. The output information is a linear combination of the input information and the parameters of the linear transformation.
where $n$ is the number of qubits of the state, and $S(\rho_{AB})$ 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 only perform local operations and communicate classically (LOCC), then in general, they will be able to draw less local information from the state. In we defined the notion of the deficit $\Delta$ to quantify the information that can no longer be drawn under LOCC. For pure states, it was proven that $\Delta$ 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 $\rho_{AB}$ 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

$$\Delta_f(Q) = I_f(Q) - I . \quad (2)$$

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 $\rho_A$ is obtained by tracing out on Bob’s system. This rough bound is obtained by noting 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 will obtain a tight bound later in this paper.

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 so that at maximum we have $I_f \equiv I_f(E_c) \leq n$ and $\Delta_f \leq S(\rho_{AB})$. We will shortly 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 Figure 1 we show a typical curve which we now explain. Since we will be comparing the formation curves to extraction curves, we will adopt the convention that $I_f(Q)$ and $Q$ will be plotted on the negative axis since we are using up resources. It can be shown that $I_f(Q)$ is concave, monotonically continuous. To prove concavity, we take the limit of many copies of the state $\rho_{AB}$. 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

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

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 $I_e(Q)$ 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 Figure 4).

There appear to be at least three special points on the curve. The first, is the point $I_1 \equiv I(0)$. This was considered in $\mathcal{I}$ when we draw maximal local information without the aid of a quantum channel. Another special point is the usual entanglement distillation procedure $I_g = I(E_B)$. 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, $E \equiv I(E_r)$ 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, $h(Q)$ 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 $\Delta$.

$$I + Q \leq h$$

(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 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 Figure 4a, we plot the region occupied by pure states. For extraction processes, pure states saturate the bound of Eq. (4).

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. The converse is not true, at least for single copies, as there are separable states such as those of $\Delta \neq 0$, and $\Delta \neq 0$.

It therefore appears that $\Delta_f$ 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 $\Delta$ can be thought of as the bound information which cannot be extracted under LOCC.

Figure 4a–d shows 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. 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 $\Delta_f$. Consider the notion of LOCC-orthogonality (cf. \textit{S}). One says that $\varphi$ is LOCC-orthogonal, if by LOCC Alice and Bob can transformed $\sum_i \rho_i |i\rangle \langle i| \otimes |\tilde{i}\rangle \langle \tilde{i}|$ into $|i\rangle \langle i| \otimes \sum_i \rho_i \tilde{i} \langle \tilde{i}|$ and vice versa; $|\tilde{i}\rangle$ is the basis of Alice's ancilla. In other words, Alice and Bob are able to correlate the state $\varphi_i$ to orthogonal states of a local ancilla as well as reset the correlations. Consider a state $\varphi_{AB}$ that can be LOCC-decomposed, i.e., it is a mixture of LOCC-orthogonal states $\varphi = \sum_i \rho_i \varphi_i$. The decomposition suggests a scheme for reversible formation of $\varphi$. Alice prepares locally the state $\varphi_{A'} = \sum_i \rho_i |i\rangle \langle i| \otimes |\tilde{i}\rangle \langle \tilde{i}|$. This costs $n_{A'A'} = S(\varphi_{A'A'})$ bits of information. Conditioned on $i$, Alice compresses the halves of $\varphi_i$, and sends them to Bob via a quantum channel. This costs $\sum_i \rho_i S(\varphi_i)$ qubits of quantum communication. Then, since the $\varphi_i$ are LOCC-orthogonal, Alice and Bob can reset the ancilla, and return $n_{A'}$ bits. One then finds, in this protocol, formation costs $n_{A'A'} = S(\varphi_{A'A'})$ bits, hence it is reversible. Consequently $E_r(\varphi_{AB}) \leq \sum_i \rho_i S(\varphi_i)$, hence

$$E_r(\varphi_{AB}) \leq \inf_{X} \sum_i \rho_i S(\tilde{\varphi}_i), \quad X = A, B$$

(5)

where the infimum runs over all LOCC-orthogonal decompositions of $\varphi_{AB}$.

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 $\{\rho_i, \psi_{ij}\}$, where $\sum_j \psi_{ij}^* \psi_{ij} = \delta_i$. Now, Alice prepares locally the state $\varphi_{A'A'} = \sum_i \rho_i \psi_{ij} \hat{n}_i \otimes |\tilde{i}\rangle \langle \tilde{i}|$. Conditioned on $i$, Alice compresses the halves of $\psi_{ij}$'s and sends them to Bob. This costs on average $E_r(\varphi_i)$ qubits of communication. So far Alice borrowed $n_{A'A'} = S(\varphi_{A'A'})$ bits. Alice and Bob then reset and return ancilla $A'$ (this is possible due to LOCC-orthogonality of $\psi_i$) and also return ancilla $A''$ without resetting. The amount of bits used is $n_{A'B''} = S(\varphi_{AB}) - \sum_i \rho_i S(\psi_i)$, giving

$$\Delta_f \leq \inf \sum_i \rho_i S(\psi_i) \leq S(\varphi)$$

(6)
where, again, the infimum runs over the same set of decompositions as in Eq. 4, providing a connection between $\Delta_f$ and $E_p$. 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 10. The authors show that an ensemble of nonorthogonal states cannot be compressed to less than $S(\rho) \geq$ 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. 7, 8. 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 $\Delta_f = S(\rho_{A,B})$, $E_p = \min \{ S(\rho_A), S(\rho_B) \}$.

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 $\rho = \sum_i p_i \rho_i$, where $\rho_i$ are themselves LOCC-orthogonal mixtures of pure states. Thus Alice can share with Bob each $\rho_i$ at a communication cost of $Q = \sum_i p_i E_c(\rho_i)$. If the states $\rho_i$ are not LOCC-orthogonal, Alice cannot reset the instruction set, so that the information cost is $I = n - \sum_i p_i S(\rho_i)$. We will now show by example, that this may be a very cheap scenario. Consider

$$e = p|\psi_{+}\rangle\langle\psi_{+}| + (1 - p)|\psi_{-}\rangle\langle\psi_{-}|, \quad p \in [0, \frac{1}{2}] \tag{7}$$

with $\psi_{\pm} = \frac{1}{\sqrt{2}}(|0\rangle \pm i|1\rangle)$. When $p \neq 0$ we have $E_p = 1$, $I_f = 2$, $E_c = H(\frac{1}{2} + \sqrt{p(1 - p)})$ 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 $\rho = 2p \rho_0 + (1 - 2p)|\psi_{+}\rangle\langle\psi_{+}|$, where $\rho_0$ is an equal mixture of LOCC-orthogonal states $|0\rangle$ and $|1\rangle$. The intermediate point is then $(1 - 2p, 2 - 2p)$. A family of diagrams with changing parameter $p$ is plotted in Fig. 9. The derivative $\chi(Q) = \frac{\partial Q}{\partial p}$ 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) and states with $\Delta \neq 0$ which contain quantum correlations. It would be interesting to explore these transitions and diagrams further, and also the tradeoff between information and quantum communication. To this end, the quantity $\Delta_f(Q) + Q - E_c$ appears to express this tradeoff. Finally, we hope that the presented approach may clarify an intriguing notion in quantum information theory, known as the thermodynamics of entanglement 10, 11, 12.

Acknowledgments: This work supported by EU grant EQUIP, Contract IST-1999-11053. J.O. acknowledges Lady Davis, and grant 129/00-1 of the ISF. M.H. and R.H. thank C. Tsallis for discussions.

---