# Local Information as a Resource in Distributed Quantum Systems

Michał Horodecki,[1] Karol Horodecki,[2] Paweł Horodecki,[3] Ryszard Horodecki,[1] Jonathan Oppenheim,[1,4] Aditi Sen(De),[1] and Ujjwal Sen[1]

[1]*Institute of Theoretical Physics and Astrophysics, University of Gdańsk, Gdańsk, Poland*
[2]*Faculty of Mathematics, University of Gdańsk, Gdańsk, Poland*
[3]*Faculty of Applied Physics and Mathematics, Technical University of Gdańsk, 80-952 Gdańsk, Poland*
[4]*Racah Institute of Theoretical Physics, Hebrew University of Jerusalem, Givat Ram, Jerusalem 91904, Israel*

(Received 30 July 2002; published 12 March 2003)

A new paradigm for distributed quantum systems where *information* is a valuable resource is developed. After finding a unique measure for information, we construct a scheme for its manipulation in analogy with entanglement theory. In this scheme, instead of maximally entangled states, two parties distill *local* states. We show that, surprisingly, the main tools of entanglement theory are general enough to work in this opposite scheme. Up to plausible assumptions, we show that the amount of information that must be lost during the protocol of concentration of local information can be expressed as the relative entropy distance from some special set of states.

DOI: 10.1103/PhysRevLett.90.100402 PACS numbers: 03.65.Ud, 03.67.–a

The notion of quantum correlations is more general than entanglement [1,2]. A formal measure of quantum correlations in measurements (quantum discord) [2] was found, based on an entropylike function. Recently, the first *operational* approach to quantify quantum correlations was introduced in [3]. Subsequently, a similar approach was used to justify a physical interpretation of (the optimized) quantum discord [4]. The results of [3] were based on the idea that by using a system in a pure state one can draw work from a single heat bath. This scenario was used in the case of distributed quantum systems: Alice and Bob share a state, have local heat baths, and can use only local operations and classical communication (LOCC), to concentrate the information contained in the state, in order to draw work. The amount of work drawn by LOCC is usually smaller than that extractable if Alice and Bob can use global operations. The resulting difference denoted by the deficit $\Delta$ accounts for the part of correlations that must be lost during classical communication, thus describing purely quantum correlations. In the case of $\Delta$ for pure states, it was argued to be exactly equal to the entanglement while for mixed states it is supposed to be an independent quantity. In this context, it is clear that understanding the problem of concentration of information will provide valuable insight into the nature of quantum correlations. Yet the early development of these ideas [3,5] indicated that the proposed scenario is completely different than anything we had in quantum information theory so far. In particular, the serious difficulty (which is removed in this Letter) was that one is not even able to obtain (without additional assumptions) the value of $\Delta$ in the simplest case of a two-qubit Bell state.

In this context basic questions arise: (i) What is the connection between the above thermodynamical quantification of quantum correlations and the main concepts of quantum information theory? (ii) Can we formulate the concentration of information within a framework of manipulating resources as in entanglement theory? An even more basic question follows: (iii) Can we give up thermodynamics and formulate the problem solely in terms of quantum information?

A powerful domain of quantum information theory which is well formalized is entanglement theory where the primitive notions are only states of compound systems and a class of operations. One would like to formalize our scheme in a similar manner. An important question then follows: Can we make our scheme similar to entanglement theory? Apart from conceptual benefits, one then hopes that the powerful tools of entanglement theory can be borrowed to tackle information manipulations.

These questions, answered in this Letter, are of fundamental importance for investigations of quantum correlations of composite states. First, we do not start from an *a priori* notion of information, and we do not assume that we are interested in the usual function for information $I = N - S$ ($N$ is the number of qubits associated with a given state [6] and $S$ is its von Neumann entropy), which we used in [3] based on thermodynamical considerations. Instead we define a class of global operations over a quantum system called noisy operations (NO) and define information as whatever quantity does not increase under the class. We find that, under certain natural assumptions $N - S$ is the *unique measure of information.*

Then we pass to distributed quantum systems. In this context, a natural class of operations is LOCC [7]. To tackle information, we restrict this class, obtaining N(oisy)LOCC. It differs from LOCC in that only maximally mixed local ancillas can be added for free. We show that the task of concentration of information is a dual scheme to the entanglement distillation scheme. Instead of maximally entangled states, Alice and Bob have to draw *local states.* In entanglement theory, pure local states were a free resource; in our scheme, they represent a useful resource: *localized information.* The schemes are

  

then quite opposite. Surprisingly, however, we show that the main tools of entanglement theory can be applied to the problem of concentration of information. This is due to the fact that both paradigms are based on the general scheme of state transformations under a given class of operations. Using the technique of monotones [8,9] and Rains semidefinite program [10] we obtain rigorously that for pure states, the amount of localizable information $I_l$ is equal to $N - E$, where $E$ is the entanglement. This means that the quantum correlations defined by the deficit $\Delta = (N - S) - I_l$ for pure states are equal to its entanglement. Also up to plausible assumptions we obtain that $\Delta$ is the relative entropy distance from some special set of states (smaller than the set of separable states; cf. [11]).

*Unique measure of information.*—Let us begin by looking at the class of allowable operations for a single system. We use *noisy operations* [12]. We imagine that the Universe is mostly filled with qubits in maximally mixed states. This is reminiscent of the heat bath in thermodynamics. Then a qubit in a nonmaximally mixed state is valuable. Such qubits cannot be brought in for free, but *only manipulated.* Accordingly we allow (a) unitary operations, (b) adding an ancilla in the maximally mixed state, and (c) rejecting (tracing out) part of the system. Operations (a) and (b) are reversible while (c) is not. We do not consider measurement as a separate operation, as the latter involves implicitly a device with memory initially in a pure state. Rather we count the device qubits explicitly, i.e., treat it as part of the system. The measurement is then a unitary transformation. We can now ask, under NO, what is the number of pure states which can be distilled from a state $\varrho$. It turns out that the number of pure states which can be distilled from a $N$-qubit state $\varrho$ under NO is the information $I(\varrho) = N - S(\varrho)$, where $S(\varrho) = -\mathrm{Tr}\varrho\log\varrho$ is the von Neumann entropy and that this process is reversible. The details of the proof of this will be presented in [13], so here, we just sketch an outline. The method of distilling pure states is essentially Schumacher compression [14] (cf. [15] and references therein). In the original compression scheme, the pure qubits were rejected, while the qubits carrying the compressed state were kept, as the latter was the signal. Here, we do the *converse:* keep the pure part (the mixed state is not interpreted as signal, but as noise, as in [16]). Therefore it is cooling rather than compression [17].

One can also show that the converse protocol is possible: i.e., to create $n$ copies of state $\varrho$ one needs $n(N - S(\varrho))$ pure qubits and $nS(\varrho)$ qubits of noise. (This is somewhat similar to the much more advanced reverse second Shannon theorem problem [18].)

It follows that in the asymptotic regime (many copies), apart from the maximally mixed state, any state can be reversibly converted into any other, at a rate fixed by the entropy of the state. Then, as with the Carnot efficiency (cf. [19]) there is a unique function that is monotonic under transitions possible by NO (up to a factor and additive constant). The function we can call *information,* as everybody can agree that information, whatever it is, cannot increase under NO. We fix the free parameters by requiring that for pure $N$-qubit states $I = N$. *We have thus obtained that I is the unique measure of information contained in quantum systems:* Information cannot be created; it can only be manipulated or lost.

*Concentration of information to a local form.*—Now we are in a position to pass to distributed systems shared between Alice and Bob (for simplicity we deal with bipartite systems). We now restrict the class of operations (called NLOCC) allowing (i) local noisy operations (ii) sending qubits down the dephasing channel which is equivalent to classical communication. The only difference between LOCC and the present NLOCC class is that in the former, local ancillas are free while in the latter, only maximally mixed ancilla are free [20].

In entanglement theory the valuable resource was the maximally entangled state $\psi_+^d = 1/\sqrt{d}\sum_{i=1}^d |i\rangle|i\rangle$ and one asks the following question: *How many maximally entangled states can be distilled from $\varrho$ with help from LOCC operations?* Instead, we propose to care about *local information.* Consider a bipartite system with information contents $N - S(\varrho)$. The main question follows [3]: *How much of the information can be concentrated to local form by NLOCC?.* Since we know that local information is equivalent to local pure states, the question can be converted into *how many pure product states can one distill from the state $\varrho$ under the NLOCC class?* From the formal point of view this happens to be very similar to the problem of entanglement distillation. The key difference is that instead of a maximally entangled state, we want to distill states that were regarded as useless so far. (The character of the distilled resource naturally imposes the use NLOCC instead of LOCC.) Surprisingly, this allows us to quantify quantum correlations: the amount of information that cannot be concentrated must be quantum, because it is destroyed by the dephasing channel. Thus we define the classical (i.e., local) information contents $I_l$ of the state $\varrho$, as the optimal rate $2m/n$, where $n$ is the number of copies in the state $\varrho$, $m$ is the number of two-qubit pairs output into pure product states (we put the factor 2 since we count local information in bits). The contents of quantum correlations $\Delta$ are then defined as the difference: $\Delta = I - I_l$. The task of evaluating $I_l$ (hence also $\Delta$) turned out to be surprisingly hard. Even for the state $\psi_+^d$, we were not able to prove rigorously that $I_l = \log d$ [3], which is intuitively obvious.

Below, we find that since $I_l$ is some conversion rate under NLOCC, the following powerful tools of entanglement theory (with suitable modifications) can be applied: enlarging the class of operations, Rains approach [10], and the concept of monotones [8,9].

First of all, NLOCC operations (like LOCC ones) are hard to deal with from a mathematical point of view. In entanglement theory, one considers the greater classes of operations [10,21]: separable operations and so-called positive partial transpose (PPT) operations

[10]. To obtain analogous classes, we need only to add a condition of *preserving maximally mixed state* (PMM). Thus we can optimize the rate over PMM versions of the above classes, which is a great simplification and leads to useful bounds for $I_l$.

In the Rains approach, one assumes a fixed rate of conversion from one state to another and evaluates the fidelity of the conversion. The rate is attainable if the fidelity can approach 1 in the limit of many input copies. For simplicity we work with two-qubit states. We fix the rate $r$, which means that for $n$ input copies we obtain $m = nr/2$ output pairs. The pairs are in a final joint state $\varrho' = \Lambda(\varrho^{\otimes n})$, where $\Lambda$ is the optimal NLOCC operation. We want to estimate the fidelity between $\varrho'$ and the $m$ copies of the two-qubit state; we would like to have (e.g., $P_{00} = |00\rangle\langle 00|$) for fixed $n$

$$F = \mathrm{Tr}[P_{00}^{\otimes m}\Lambda(\varrho^{\otimes n})]. \tag{1}$$

To evaluate $F$ one can use techniques developed in [10]. One rewrites (1) in the form $F = \mathrm{Tr}[\Lambda^\dagger(P_{00}^{\otimes m})\varrho^{\otimes n}] \equiv \mathrm{Tr}\Pi\varrho^{\otimes n}$, where $\Lambda^\dagger$ is the dual map [22].

Let us find constraints for $\Pi$. One can check that since $\Lambda$ is trace preserving, then $0 \le \Pi \le I$. Since $\Lambda$ preserves a maximally mixed state [i.e., $\Lambda(I/d_{\mathrm{in}}) = I/d_{\mathrm{out}}$), one finds that $\frac{d_{\mathrm{in}}}{d_{\mathrm{out}}}\Lambda^\dagger$ is trace preserving; in our case $d_{\mathrm{in}} = 2^{2n}$, $d_{\mathrm{out}} = 2^{2m}$; hence $\mathrm{Tr}\Pi = 2^{n(2-r)} \equiv K$, so that $\Pi/K$ is a state. Since $\Lambda$ is NLOCC, so is $\Lambda^\dagger$. Then, since $P_{00}^{\otimes m}$ is separable, so is $\Pi/K$. This constrained optimization problem can be treated using the duality concept in semidefinite programming as in [10]. We will do so elsewhere. Here we prove the following bound:

Proposition 1: *For any bipartite state $\varrho_{AB}$ of $N = N_A + N_B$ qubits, we have*

$$I_l \le N - S_\infty(\varrho_X), \qquad X = A, B \tag{2}$$

*[or, equivalently, $\Delta \ge S(\varrho) - S_\infty(\varrho_X)$], where $S_\infty = -\log\lambda_{\max}$, with $\lambda_{\max}$ being the largest eigenvalue of $\varrho_X$.*

Remark: Since usually $S_\infty < S$, this bound is weaker than the one proven under an assumption in [3]. However for the maximally entangled state it is tight.

Proof: We prove for two-qubit states (generalization to higher dimensions is obvious). We note that $\frac{1}{K}\Pi$ is a separable state. Then $F \le K\sup_\sigma \mathrm{Tr}\sigma\varrho^{\otimes n}$, where sup is taken over separable $\sigma$, which can be taken pure. Consequently, $F \le K\sup_{\psi\otimes\phi}\langle\psi\otimes\phi|\varrho^{\otimes n}|\psi\otimes\phi\rangle \le K\sup_\psi\langle\psi|\varrho_X^{\otimes n}|\psi\rangle = K\lambda_{\max}^n$. Since we require $F \to 1$ for large $n$, we must have $r \le 2 + \log\lambda_{\max}$ which gives the expected bound.

In [3] we argued that $I_l(\psi) = N - S_X(\psi)$ which is equivalent to $\Delta(\psi) = S_X(\psi)$, where $S_X(\psi)$ is either of the reductions of $\psi$ and is equal to its entanglement $E(\psi)$. Here we prove it rigorously. Following proposition 1 one has $I_l(\psi_+^d) \le \log d$. This bound is attainable [3]. Since $N(\psi_+^d) = \log d^2$ the rigorous proof for maximally entangled states is complete. To extend it to other pure states, we use a version of the entanglement dilution scheme [23], in which out of $nS_A(\psi)$ shared pairs $\psi_+^2$ one gets $n$ pairs in state $\psi$ by use of $O(\sqrt{n})$ bits of communication [24]. Because of asymptotically negligible communication cost, entropy production is negligible. Thus, to create $\psi^{\otimes n}$ one needs $m = nS_A$ copies of $\psi_+^2$ (which occupy $2m$ qubits) plus $Nn - 2m$ pure local qubits. Now if from $\psi$ one could draw more than $N - S_A$ per pair, then one could draw more local bits than $m$ from the maximally entangled state, converting them first to $\psi$ and then drawing local bits from $\psi$. On the other hand, there is an obvious protocol to get $N - S_A$ local bits: Alice sends her half to Bob via the dephasing channel [3]. Thus $I_l(\psi) = N - S_A(\psi)$ or, equivalently, $\Delta(\psi) = S_A(\psi)$, which concludes the proof, since $S_A(\psi) = S_B(\psi)$. For the tripartite GHZ (Greenberger-Horne-Zeilinger) state, one can likewise show that $\Delta = 1$. The GHZ state is $(|000\rangle + |111\rangle)/\sqrt{2}$.

We now proceed to estimate $I_l$ from below. To this end we need the notion of *implementable product basis* (IPB). It is a basis that can be achieved from the standard basis $|i\rangle_A|j\rangle_B$ by means of reversible NLOCC operations. (Of course, since we want to transfer pure states into pure ones, local noise is here not needed.) IPB is always distinguishable by LOCC. Hence the basis given in [1] is not IPB. Most likely the converse is also true. It is obvious that a state with an IPB eigenbasis (call it an IPB state) has $\Delta = 0$. There is also a natural scheme of information concentration for other states: one fixes some IPB (let it be $\mathcal{B}$) and dephases a given state $\varrho$ in this basis which, by definition, is possible by NLOCC. The final state $\varrho'$ has $\Delta = 0$, so one can draw $I_l(\varrho') = N - S(\varrho')$ local pure qubits. However, $S(\varrho') = H(\varrho, \mathcal{B})$ (Shannon entropy of $\varrho$ in basis $\mathcal{B}$). Thus we have the bound $I_l(\varrho) \ge N - \inf_{\mathcal{B}\in IPB} H(\varrho, \mathcal{B})$.

Since one has $\inf_{\mathcal{B}\in IPB} H(\varrho, \mathcal{B}) = S(\varrho) + \inf S(\varrho|\sigma)$ [25], where infimum is taken over IPB states $\sigma$, and $S(\varrho|\sigma) = \mathrm{Tr}\varrho\log\varrho - \mathrm{Tr}\varrho\log\sigma$, and since it may be better to operate collectively on $\varrho^{\otimes n}$ one obtains

Proposition 2: *For $N$ qubit bipartite state $\varrho$ one has*

$$I_l \ge N - S(\varrho) - \mathcal{D}^\infty(\varrho) \tag{3}$$

*[or, equivalently, $\Delta \le \mathcal{D}^\infty(\varrho)$], where $\mathcal{D}^\infty(\varrho)$ is regularized relative entropy distance from the set of IPB states $\mathcal{D} = \inf_{\sigma\in IPB} S(\varrho|\sigma)$ [26].*

Now we make an attempt to show equality by using the method of monotones. As shown in [27,28] (see also [9]) if a function $M(\varrho)$ is (a) nonincreasing under a class of operations (b) asymptotically continuous, then its regularization satisfies

$$M^\infty(\sigma)R(\varrho \to \sigma) \le M^\infty(\varrho), \tag{4}$$

where $R$ is the optimal transition rate. As $M$ we take $N - S(\varrho) - \mathcal{D}(\varrho)$. $S(\varrho)$ satisfies (b) by Fannes inequality [29], while by [30] also $\mathcal{D}$ satisfies (b) [31]. In our case $\sigma$ is the two-qubit product state $\sigma = P_{00}$ so that $M(\sigma) = M^\infty(\sigma) = 2$. Since we intend to count not

product states, but bits, we have $I_l = 2R$. The inequality then reads $I_l(\varrho) \leq M^{\infty}(\varrho)$. Now if $M$ is monotonic [satisfies (a)], then combining proposition 2 with the inequality we would know $I_l$ *exactly*. Let us then check monotonicity. To this end we use the formula $M(\varrho) = N - \inf_{\mathcal{B}\in IPB} H(\varrho, \mathcal{B})$ and consider sending down the dephasing channel in two stages: (i) local dephasing of part of the system (ii) sending the dephased part via an ideal quantum channel. It is easy to see that operation (ii) as well as local unitary transformations does not change $M$ (it changes one IPB into other one). From the fact that a product of an IPB with some local basis is again an IPB it follows that a partial trace does not increase $M$. We are not able to prove the monotonicity of $M$ under adding noise (i.e., maximally mixed states) and local dephasing. Concerning noise, it is rather unlikely that it can increase $M$, and, moreover, most likely it is not needed in the concentration task at all (basically noise is needed when we want to create more mixed states out of less mixed ones). However, the question of whether dephasing can increase $M$ remains elusive. This is very closely related to the assumption we made in [3]. There we would obtain some improvement of the bound of proposition 1. Here if the assumption were true we would obtain the exact formula for $\Delta$. The problem can be formulated as follows: Can Alice in the optimal information concentrating protocol partially dephase her system, and at some later stage, dephase again, so that two dephasings do not commute? Since noncommuting measurements destroy information, we believe the answer is "no." This justifies the following conjecture:

Conjecture: *For N qubit bipartite state, the amount of concentratable information is*

$$I_l = N - S(\varrho) - \mathcal{D}^{\infty}(\varrho). \tag{5}$$

Thus we would have that $\Delta = \mathcal{D}^{\infty}(\varrho)$; i.e., $\Delta$ is the (regularized) relative entropy distance from IPB states. The information $I_l$ is then given in terms of a Shannon entropy, an idea which arose naturally in [5].

In conclusion, we have derived the notion of information from general principles. Then we have shown that local information $I_l$ can serve as a resource in distributed systems in an analogous way to entanglement theory. We were then able to use powerful tools developed in the latter, with suitable modifications. We obtained bounds for the amount of information that can be concentrated to local form ($I_l$) and the deficit $\Delta$. Under some assumption we argued that the deficit $\Delta$ is the relative entropy distance from the set of states having $\Delta = 0$. Since the deficit is a measure of quantum correlations based on thermodynamics, the present program can be viewed as a bridge that links quantum correlation theory and quantum thermodynamics in a systematic way. In particular, there is hope to explain nonlocality without entanglement in a thermodynamical manner, as nontrivial states diagonal in the basis of Ref. [1] may have $\Delta > 0$.

This work is supported by EU grant EQUIP, IST-1999-11053 and by Grant No. 129/00-1 of the ISF.

[1] C. H. Bennett *et al.*, Phys. Rev. A **59**, 1070 (1999).
[2] W. H. Żurek, Ann. Phys. (Leipzig) **9**, 5 (2000); H. Ollivier and W. H. Żurek, Phys. Rev. Lett. **88**, 17901 (2002).
[3] J. Oppenheim *et al.*, Phys. Rev. Lett. **89**, 180402 (2002).
[4] W. H. Żurek, quant-ph/0202123.
[5] J. Oppenheim *et al.*, quant-ph/0207025
[6] For $\varrho$ on $C^m$ one has $N(\varrho) = \log m \equiv \log_2 m$.
[7] C. H. Bennett *et al.*, Phys. Rev. A **54**, 3824 (1996).
[8] G. Vidal, J. Mod. Opt. **47**, 355 (2000).
[9] M. Horodecki *et al.*, Phys. Rev. Lett. **84**, 2014 (2000).
[10] E. Rains, quant-ph/0008047; see also Phys. Rev. A, **60**, 179 (1999); an operation $\Lambda$ is PPT if $\Gamma\Lambda\Gamma$ is a completely positive map where $\Gamma$ is the partial transpose.
[11] V. Vedral, quant-ph/9903049.
[12] Such maps were considered from a formal point of view in K. Życzkowski, report, 2001 (unpublished).
[13] M. Horodecki *et al.*, quant-ph/0212019.
[14] B. Schumacher, Phys. Rev. A **51**, 2738 (1995).
[15] D. Petz and M. Mosonyi, quant-ph/9912103.
[16] L. J. Schulman and U. Vazirani, quant-ph/9804060.
[17] D. Janzing *et al.*, quant-ph/0002048.
[18] C. H. Bennett *et al.*, quant-ph/0106052.
[19] S. Popescu and D. Rohrlich, Phys. Rev. A **56**, R3319 (1997).
[20] Actually in Ref. [3] we used C(losed)LOCC where one could borrow only the information in ancillas. Here, we consider a given state on a larger Hilbert space, which means that ancillas are added at the very beginning.
[21] V. Vedral and M. B. Plenio, Phys. Rev. A **57**, 1619 (1998).
[22] The dual map to $\Lambda$ is defined as a unique map $\Lambda^{\dagger}$ satisfying $\mathrm{Tr}A^{\dagger}\Lambda(B) = \mathrm{Tr}\Lambda^{\dagger}(A)^{\dagger}B$. For a completely positive map $\Lambda = \sum_i V_i(\cdot)V_i^{\dagger}$ we have $\Lambda^{\dagger} = \sum_i V_i^{\dagger}(\cdot)V_i$.
[23] C. H. Bennett *et al.*, Phys. Rev. A **53**, 2046 (1996).
[24] H.-K. Lo and S. Popescu, Phys. Rev. Lett. **83**, 1459 (1999).
[25] To see this, note that if $\varrho'$ stands for $\varrho$ dephased in *fixed* IPB basis $\mathcal{B}$ then the infimum of $S(\varrho'|\sigma_{\mathcal{B}})$ over all $\sigma_{\mathcal{B}}$s diagonal in $\mathcal{B}$ gives $\sigma_{\mathcal{B}} = \varrho'$ and $S(\varrho') = inf_{\sigma_{\mathcal{B}}}[-\mathrm{Tr}(\varrho' \log \sigma_{\mathcal{B}})]$.
[26] The regularization of $M(\varrho)$ is $M^{\infty}(\varrho) = \lim_{n\to\infty} \frac{1}{n} M(\varrho^{\otimes n})$.
[27] M. Donald *et al.*, J. Math. Phys. (N.Y.) **43**, 4252 (2002).
[28] M. Horodecki, Quantum Inf. Comput. **1**, 3 (2001).
[29] M. Fannes, Commun. Math. Phys. **31**, 291 (1973).
[30] M. Donald and M. Horodecki, Phys. Lett. A **264**, 257 (2001).
[31] The set of IPB states is not convex as assumed in Ref. [30]. However, it is enough that the set is starlike, so that any IPB state can be connected with a maximally mixed state by a straight line that is inside the set.