Mutually exclusive aspects of information carried by physical systems: Complementarity between local and nonlocal information

Jonathan Oppenheim,[1,2] Karol Horodecki,[3] Michał Horodecki,[2] Paweł Horodecki,[4] 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*
[3]*Faculty of Mathematics, University of Gdańsk, Gdańsk, Poland*
[4]*Faculty of Applied Physics and Mathematics, Technical University of Gdańsk, 80-952 Gdańsk, Poland*
(Received 7 July 2002; published 19 August 2003)

Complex physical systems contain information which, under some well-defined processes can differentiate between local and nonlocal information. Both these fundamental aspects of information are defined *operationally*. Local information is locally accessible and allows one to perform processes, such as physical work, while nonlocal information allows one to perform processes such as teleportation. It is shown that these two kinds of information are *complementary* in the sense that two parties can either gain access to the nonlocal information or to the local information but not both. This complementarity has a form similar to that expressed by entropic uncertainty relations. For pure states, the entanglement plays the role of Planck's constant. We also find another class of complementarity relations which applies to operators and is induced when two parties can only perform local operations and communicate classical (LOCC). In particular, observables such as the parity and phase of two qubits commute but under LOCC, they are complementary observables. It is also found this complementarity is pure in the sense that it can be ''decoupled'' from the uncertainty principle. It is suggested that these complementarities represent an essential extension of Bohr's complementarity to complex (distributed) systems which are entangled.

DOI: 10.1103/PhysRevA.68.022307 PACS number(s): 03.67.−a, 03.65.Ud

I. INTRODUCTION

The problem of mutually exclusive aspects of quantum phenomena appeared together with the birth of quantum mechanics. Pauli, in his letter to Heisenberg in December 1926, wrote: ''One may view the world with the *p*-eye and one may view it with the *q*-eye but if one opens both eyes simultaneously then one get crazy.'' Soon after, Heisenberg discovered the uncertainty principle for momentum and position [1]. The next year, Bohr introduced the concept of complementarity [2], which by 1935 acquired its final form [3]. According to Bohr, there are different aspects of complementarity in quantum mechanics. In particular, observables, such as the position and momentum of the particle, are complementary—an accurate measurement of momentum will make a subsequent measurement of position yield random results—the position information is destroyed during the momentum measurement. This complementarity between incompatible observables was inherently connected with the uncertainty principle (for deeper analysis of these notions see Ref. [4]).

However, for Bohr, the term ''complementarity'' meant something more. Quoting Bohr, ''the impossibility of combining phenomena observed under different experimental arrangements into a single classical picture implies that such apparently contradictory phenomena must be regarded as complementary in the sense that taken together, they exhaust all well defined knowledge about the atomic objects'' [5].

It is not quite clear what the phrase ''taken together, they exhaust all well-defined knowledge'' means in the context of distributed systems which involve entanglement as a physical resource. Does entanglement constitute a physical feature of quantum systems which can somehow be incorporated into the principle of complementarity? Indeed, there is hope that the fusion of information theory and quantum formalism will offer a more understandable description of nature [6].

In quantum information theory, one can view information as having a central role akin to quantities such as energy in classical physics [7]. Given this central role, one therefore wonders whether information obeys complementarity principles like other physical observables. Indeed, we will see that this is the case. This further supports a view that information plays a fundamental role in quantum mechanics and provides an underlying structure. This is perhaps interesting in light of the fact that long before the discovery of quantum communication [8,9], entropic uncertainty relations [10–16] (which are generalizations of the Heisenberg and Robertson [17] inequalities) had already been introduced. Even in 1987, a *generic information paradigm* was proposed, according to which the information is a fundamental concept in the description of physical reality [18–20]. An *operational* definition of information carried by physical systems has also been introduced [21–23].

It is with the goal of incorporating information into our description of physical systems that we investigate the complementary relationships that are induced when one considers distributed systems jointly held by two parties. We earlier suggested complementarity between ''classical'' (local) information and ''quantum'' (nonlocal) information contained in quantum states [21]. In this paper, we explore two types of complementarity within the context of quantum information theory. The first type leads us to the following complementarity principle: complex quantum systems carry information, which under well-defined mutually exclusive processes manifests itself as local information or as nonlocal information. The second type applies to operators and is induced when two parties can only perform local operations and communicate classically (LOCC).

In Sec. II we will first demonstrate how to divide information into local and nonlocal parts. The method we use comes from operational considerations. Local information is locally accessible and can be measured by how many pure separable bits can be obtained from a state, while one can obtain nonlocal information by distilling singlets from the state. In Sec. III we show that the two types of information are complementary. One finds that one can exploit the nonlocal information to perform teleportation but then the ability to use the local information to perform physical work is completely destroyed. Or, one can obtain local information but then the ability to perform teleportation is completely destroyed. The process of obtaining and using either entanglement or pure local states is irreversible, and using one resource destroys the possibility of obtaining the other. This irreversibility is crucial in leading to our complementarity relations. If one distills the state to local form, then one will no longer be able to teleport. If one uses nonlocal states to perform a task, such as teleportation, then purity is necessarily destroyed and one can no longer obtain pure local states. We will show that this complementarity can be expressed as an information-theoretic bound which has the same form as an entropic uncertainty relation. For pure states, the bound has the feature that the entanglement plays the role of Planck's constant $\hbar$.

In Sec. IV we introduce a complementarity principle involving individual measurements. For example, if one has two qubits, one can measure the parity and phase. However, when each of the two parties hold one of the qubits in distant labs, they find that all parity and phase measurements will be complementary. They can measure the parity of the state or the phase of the state, but not both. To quantify this, we introduce the idea of a LOCC complementarity inequality. We also argue that two observables can be complementary without being uncertain, demonstrating that the two concepts can be decoupled. In Sec. V we conclude by considering our complementarities in the context of Bohr's complementarity and we argue that the complementarity between local and nonlocal information can be viewed as an extension of Bohr's complementarity to complex (distributed) systems. Here, *information* becomes the central concept. Finally, we raise some open questions.

II. LOCAL AND NONLOCAL INFORMATION

Consider a bipartite state ρ_{AB} composed of n qubits which could be shared between two parties, Alice and Bob [24]. Let ρ_A and ρ_B be the reduced density matrix for each party and let n_A and n_B be the number of qubits that each party holds.

The total information encoded in the state is given by

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

where $S(\rho)$ is the von Neumann entropy of a state ρ. The more we know about the state of a system, the lower is its entropy and greater the informational content of the state. The quantity I has an operational meaning. It is the number of pure qubits which can be obtained from a state, by use of Schumacher compression [25] (cf. Ref. [26]). Moreover, it was shown [22,23] to be the *unique* measure of information under some natural assumptions in the asymptotic regime. This information can be divided into local information contents $I_A = n_A - S(\rho_A)$, $I_B = n_B - S(\rho_B)$, and mutual information $I_M = S(\rho_A) + S(\rho_B) - S(\rho_{AB})$ so that

$$I = I_A + I_B + I_M. \tag{2}$$

The typically, mutual information I_M is used as a measure of the total correlations between ρ_A and ρ_B. It tells us how much information the two systems have in common.

In classical information theory [27], a classical system [28] ρ_{cl} has a mutual information, which is always smaller than the total Shannon entropy H of the state:

$$I_M(\rho_{cl}) \leq H(\rho_{cl}). \tag{3}$$

This means that the correlations are always accompanied by a lack of information about the total system: only mixed states can have nonzero correlations. Also, for two classical systems composed of n bits, the correlations cannot exceed $n/2$.

For quantum system there is no restriction like Eq. (3). Therefore, pure states can contain correlations and the mutual information can be twice as much as in the classical case. For a general n qubit state we have

$$I_M(\rho_{qu}) \leq n. \tag{4}$$

Thus, two qubits can share two bits of mutual information, as in the case of a maximally entangled state such as the singlet

$$\psi^- = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle). \tag{5}$$

There is a basic question: For the singlet, what is the meaning of the fact that the amount of mutual information is *two*? One possible answer comes from superdense coding [8]. By using a singlet, one can communicate two bits of information through one qubit [29]. It has also been argued that the additional correlations are related to negative conditional entropies [30]. We will propose a different answer to this fundamental question. We will argue that 2 is not equal to $1+1$ but rather it is equal to *either* 1 *or* 1. In other words, the two bits of mutual information can be divided into one bit of nonlocal information and one bit of local information, but these two types of information are complementary—one can retrieve the one bit of local information or the one bit of nonlocal information, but not both. In general, as we will see below, the correlations of a quantum state consist of two *complementary* parts—one local and the other nonlocal. The first attempt at quantifying quantum contents of correlations other than through entanglement is due to Zurek [31]. Quantifying classical correlations of a quantum state and the division into classical and quantum correlations was proposed in Ref. [32]. An operational proposal of quantifying different types of correlations was first proposed in Ref. [21]. Another method, using the entanglement of purification was given in

Ref. [33]. Here we will follow Ref. [21], where the division emerges from thermodynamical considerations.

The idea is to define information operationally—*local information* can be manipulated into a locally accessible form. In other words, it is the information that can be localized by the two parties Alice and Bob and used to perform classical tasks. For example, as discussed in Ref. [21], it can be used to extract real physical work from a local heat bath, using a Szilard engine [34] or (for quantum states) a von Neumann engine [35]. Information that is locally accessible is equal to the maximum amount of work which can be drawn from a local heat bath by Alice and Bob under LOCC in units of kT, where T is the temperature of the bath and k is Boltzmann's constant. We will henceforth set $kT=1$ so that the amount of work drawn is measured in bits. One can think of the local information as the maximum amount of pure separable states which can be extracted from a state. We will therefore talk of *extracting* local information from a state, with the understanding that it could refer to extracting physical work from the information encoded in the state, or extracting a number of pure separable states. It should be mentioned here that this definition of local information is independent of the interpretation of quantum mechanics one uses (Copenhagen, Many Worlds, Bohmian, etc.).

On the other hand, *nonlocal information* is defined to be the information which can be used to perform tasks which have no classical counterpart such as teleportation and double-dense coding. One bit of nonlocal information can be used to teleport one qubit. One can think of teleportation (sending qubits) as analogous to a form of quantum logical work [7]. It implies an *operational* way of understanding local and nonlocal information. Of course, one could also simply understand the two quantities as being the number of pure local states versus the number of maximally entangled states.

Let us first look at that case where Alice and Bob are only allowed to perform local operations (LO). In this case, the amount of information I_{LO} they can obtain is

$$I_{LO}=n_A-S(\rho_A)+n_B-S(\rho_B)=I_A+I_B. \tag{6}$$

On the other hand, if we allow Alice and Bob to perform any LO and send qubits to each other through a classical channel [36], then they will be able to obtain more information from the state by exploiting correlations. Alice and Bob can then transform state ρ_{AB} into another state ρ'_{AB} such that the amount of local information is maximized. The amount of obtainable local information is

$$I_l=I_A(\rho'_A)+I_B(\rho'_B)=n-S(\rho'_A)-S(\rho'_B); \tag{7}$$

then, the difference

$$\Delta_c\equiv I_l-I_{LO} \tag{8}$$

tells us the additional information that can be obtained if the two parties are also able to perform classical communication (CC). I.e., they have access to a classical channel. Since the channel is classical, we will refer to Δ_c as the *classical deficit*.

On the other hand, if Alice and Bob had access to a quantum channel (QC) rather than to the classical channel, they would be able to localize all the information (and draw all the work from the state). The information they can obtain under these operations (LOQC) is just

$$I_{LOQC}=I=n-S(\rho_{AB}), \tag{9}$$

since Alice can just send her part of the state to Bob, who can then perform local operations on it to draw all the information. The quantity

$$\Delta\equiv I_{LOQC}-I_l \tag{10}$$

then tells us how much more information can be obtained when the channel is changed from a classical channel to a quantum channel. It is the *quantum deficit*.

It is easy to verify that the classical deficit Δ_c plus the quantum deficit Δ are equal to the total amount of correlations contained in the state. I.e.,

$$I_M=\Delta+\Delta_c. \tag{11}$$

Remarkably, for pure states, it was found that $\Delta=E_D$, where E_D is the amount of distillable entanglement contained in the state (i.e., the number of singlets per state ρ_{AB} that can be drawn under LOCC from a large number of copies of ρ_{AB}) [21]. This was also conjectured to be true for sets of states such as the "maximally correlated" state of Ref. [37]. In general, the quantum deficit Δ can be due to entanglement, as in the case of pure states, but it also appears that separable states can have a nonzero Δ, as in the case of mixtures of states which are separable, but indistinguishable [38]. States, such as Werner states [39], are believed to have $\Delta > E_D$. However, it is not yet clear whether $\Delta > E_D$ in the case of collective operations on many copies.

Under LOCC, I_l is the amount of local information that can be extracted from state ρ and used to perform physical work. E_D has the interpretation of the maximal amount of useful nonlocal information which can be extracted from state ρ_{AB}. Each bit of nonlocal information (singlet) can then be used for such tasks as the teleportation of one qubit, or super-dense coding. Here, we will take quantum work to mean teleportation of qubits but it is certainly not excluded that there are other forms of quantum work [40].

Generally, Eq. (10) divides the total informational content into a local part I_l which is locally accessible and Δ which represents quantum information destroyed by the communication via classical channel:

$$I=I_l+\Delta. \tag{12}$$

Yet, in this paper we are interested in nonlocal part of Δ which is E_D [41]. We will derive complementarity relations between E_D and I_l.

III. COMPLEMENTARITY BETWEEN LOCAL AND NONLOCAL INFORMATION

We will now show that local information and nonlocal information are complementary—one can use the local information, or the nonlocal information, but not both. The use of each type of information is an irreversible process and the information is destroyed. Before discussing the general case, it may be useful to first show how this complementarity principle plays out with a simple state such as the singlet (5). In Sec. III B we will show how this can be extended to other states. It is expressed mathematically as a basic inequality. In Sec. III C we will discuss this complementarity in more generality and show two different and useful ways that it can be expressed. Furthermore, for pure states, one can express the relationship in a particularly simple form, where the entanglement plays the role of Planck's constant $\hbar$. In Sec. III D we show that these complementarities are of the same form as the more familiar ones encountered in quantum mechanics between conjugate observables such as position and momentum. Finally, in Sec. III E we give an example of information extraction, which illustrates our complementarity principle and shows that pure states can be thought of as the counterpart to coherent states (i.e., minimum uncertainty wave packets).

A. An example

Although the mutual information of the singlet is two bits, initially, neither Alice nor Bob can obtain any information since their local-density matrices are maximally mixed. However, in Ref. [21] we showed that one bit of information can be obtained by the two parties. This can be done using the following process, which was proven to be optimal.

Alice simply sends her qubit down the classical channel. Since the channel is classical, her qubit is dephased [36]. (Alice can take the dephasing to be in the computational basis.) I.e., the channel causes the singlet to become the classically correlated state:

$$\rho = \frac{1}{2}(|01\rangle\langle 10| + |10\rangle\langle 01|). \tag{13}$$

This step is irreversible and one bit of information gets lost [initially, the state had two bits of information and now, one can easily check via Eq. (1) that it finally has one bit of information]. This information is now locally held by Bob and by performing a cnot [42], he will hold one qubit in a pure state (the other is maximally mixed). They have thus extracted one bit of local information (i.e., one local qubit).

Alice and Bob can also use measurements to distill one bit of local information. One of course must include the state of the measuring device in the calculation. The procedure is as follows.

(a) Alice uses a measuring device represented by a qubit prepared in the standard state $|0\rangle$ [43]. She performs a CNOT [42] using her original state as the control qubit and the measuring qubit as the target.

(b) The measurement qubit is now in the same state as her original bit and can be dephased (i.e., decohered) in the $|0\rangle,|1\rangle$ basis so that the information is purely "classical" (dephasing simply brings the off-diagonal elements of the density matrix to zero, destroying all quantum coherence). Again, during the dephasing process, one bit of information is irreversibly transferred into the environment and is no longer available.

(c) The measuring qubit can now be sent to Bob.

(d) Bob performs a CNOT using the measuring qubit as the control. His original qubit is now in the standard state $|0\rangle$.

(e) Bob sends the measuring qubit back to Alice.

(f) Alice resets the measuring device by performing a CNOT using her original bit as the control. Alice's state is now maximally mixed, while Bob's state is known. They have obtained one bit of information. This information can be used to extract one bit of physical work from a heat bath using a Szilard heat engine.

This process, though optimal, only extracts one bit of local information, even though two bits of information could be extracted by someone who is not constrained by LOCC. However, the singlet also has one bit of nonlocal information, which can be used to teleport a single qubit. In this case, the ability to obtain local information will be lost. If Alice wishes to teleport state $\psi_{A'}$ using a singlet, the total initial state is

$$\psi_{A'} \otimes \psi_{AB}^{-}. \tag{14}$$

The final state (after Alice resets her measuring device) is

$$\frac{1}{4} I_{A'A} \otimes \psi_B. \tag{15}$$

Thus, the state (excluding the teleported state $\psi_{A'}$) is now maximally mixed, containing no local information. One might think that there could be some other, more sophisticated protocol that allows one to teleport a qubit in such a way that the final state will not be maximally mixed. However, this is not the case. All perfect fidelity teleportation schemes were considered by Werner [39], who showed that essentially the standard teleportation protocol is unique.

We, therefore, see that for the singlet, there appears to be a complementarity between teleportation and local information—one must choose which one to obtain, and one bit of information gets destroyed. The example of the singlet leads to the following general procedure and result.

B. The basic inequality

For a given state, one can use a particular process to distill singlets (nonlocal information). For this process, the amount of singlets need not be optimal (i.e., can be less than E_D). Similarly, the classical correlations can be exploited to obtain local information under a process which need not be optimal (i.e., can be less than I_l). After distilling singlets from a state, one can then use the rest of the state to gain local information and vice versa. More generally, consider state $\rho^{\otimes n}$ and process $\mathcal{P}$, according to which one extracts nI_l bits of local information and teleports some amount of qubits nQ_D. It turns out that each such process $\mathcal{P}$ satisfies the following inequality:

$$Q_D(\mathcal{P}) + I_l(\mathcal{P}) \leq I_l, \tag{16}$$

where now Q_D and I_l denote number of qubits and bits per input pair, respectively.

We know from quantum channel theory [44] that Q_D is equivalent to the number of singlets ($E_D = Q_D$) that must be used to transmit qubits faithfully. Then, the above inequality implies that the amount of local information I_l plus nonlocal information E_D, which can be drawn from state $\rho^{\otimes n}$ in the asymptotic limit under any process $\mathcal{P}$, cannot exceed the optimal amount of local information that can be drawn under LOCC.

To see this, we will demonstrate that if there is a process $\mathcal{P}$ such that the bound of Eq. (16) is violated, then there must exist a process $\mathcal{P}'$ that would enable us to draw a greater amount of classical work than the optimal amount I_l. Process $\mathcal{P}'$ is as follows: we first apply process $\mathcal{P}$ to draw an amount $I_l(\mathcal{P})$ of nonlocal information (pure separable states) and $E_D(\mathcal{P})$ of singlets from state ρ_{AB} (we don't perform teleportations yet).

Alice and Bob can then obtain more local information by converting the $E_D(\mathcal{P})$ singlets into $E_D(\mathcal{P})$ pure separable states using the optimal procedure described above to convert each singlet into one local pure state. Using this process, Alice and Bob can draw $I_l(\mathcal{P}) + E_D(\mathcal{P})$ bits of local information from state ρ_{AB}. Since I_l is the optimal amount of information, the bound given by Eq. (16) follows.

Equation (16) shows that there is a *trade off* between two different processes: if we define goal (i) as having one bit of local information on either site and goal (ii) as sending one bit of nonlocal information from one site to another, then only one of the goals can be reached. This represents the trade off. However, there is more going on here than a trade off. Namely, reaching (i) *irreversibly destroys* the possibility of access to (ii). This is what corresponds to *complementarity*. All this can be seen easily in the scenario before the teleportation process: Alice and Bob share a singlet and Alice has an unknown qubit. The latter does not change the balance because as an additional resource, it must be counted in both the input and the output. Then, to achieve (i) we can only spend the singlet which can finally lead to one classical bit according to the result of Ref. [21]. This *destroys* all the quantum correlations and consequently, the possibility to reach goal (ii). If, on the other hand, Alice and Bob decide to teleport, then goal (ii) is reached, but finally Alice's state is completely mixed (Bob's qubit is in an *unknown* pure qubit that does not enter the balance), so local information has been irreversibly destroyed to enable us to obtain (i).

It is worthwhile to compare Eq. (16) with Eq. (12) since in a number of cases [21], $\Delta = E_D$. In this case, Eq. (12) gives $E_D + I_l \leqslant I$. One can see that the optimal amount of distillable nonlocal information plus the amount of distillable local information is in general much greater than the amount that is actually extracted because of the complementarity between the two. There is an irreversible process which destroys our ability to obtain one kind of information, if the other kind is obtained.

C. Complementarity between local and nonlocal information expressed in terms of entropies

Although Eq. (16) essentially expresses the complementarity between local and nonlocal information, it is not in a form usually associated with uncertainty relations. We will therefore re-express our bound in two different ways. First, we will rewrite it as an informational bound, where the right-hand side is a constant, as opposed to something which depends on the state. We will also re-express it as a bound on entropies. In this case, the right-hand side is related to the entanglement of a state. Both these bounds have a form like those associated with complementary observables.

Let us first reexpress Eq. (16) so that the right-hand side is independent of the particular state chosen. To do this, we write it in terms of the local information which is extractable from correlations (as opposed to the local informational content). Defining the information which can be extracted from correlations as $I_{cor}(\mathcal{P},\rho_{AB}) \equiv I_l(\mathcal{P},\rho_{AB}) - I_{LO}$, we can subtract I_{LO} from both sides of Eq. (16) and use Eq. (8) to give

$$E_D(\mathcal{P},\rho_{AB}) + I_{cor}(\mathcal{P},\rho_{AB}) \leqslant \Delta_c. \tag{17}$$

Under some assumptions in Ref. [21], we have proved that $I_l \leqslant n - S_X$, $X = A,B$. We believe that this is true in general. Since $\Delta_c = n - S_A - S_B - I_l$, we then would obtain $\Delta_c \leqslant \min(S_A,S_B) \leqslant n/2$, so that Eq. (17) would take the form

$$E_D(\mathcal{P},\rho_{AB}) + I_{cor}(\mathcal{P},\rho_{AB}) \leqslant n/2. \tag{18}$$

This bound is the tightest bound one can have, which is state independent, as it is saturated by maximally entangled states, since for two qubit states we have $E_D(\mathcal{P},\rho_{AB}) + I_{cor}(\mathcal{P},\rho_{AB}) \leqslant 1$.

We can also reexpress the complementarity relation in terms of entropies, which will also be useful in relating our complementarity to the ones usually encountered in quantum mechanics. We, therefore, rewrite Eq. (16) in the following form:

$$H_{LOCC}(\mathcal{P}) + H_B(\mathcal{P}) \geqslant n + E_f - I_l, \tag{19}$$

where $H_{LOCC}(\mathcal{P})$ is defined, in analogy with Eq. (1), through

$$I_{LOCC}(\mathcal{P}) \equiv n - H_{LOCC}(\mathcal{P}). \tag{20}$$

Quantity $H_{LOCC}(\mathcal{P})$ can be thought of as the Shannon entropy, as Alice and Bob would perceive it during the local information localizing procedure [45]. In fact, it has been advocated [46] that entropy should always be defined with respect to one's measuring apparatus and how they can be used. For example, the coarse-grained entropy is defined with respect to detectors that can only probe with a finite resolution. Here, the measuring devices of Alice and Bob are restricted to LOCC operations.

Also in analogy with Eq. (1), $H_B(\mathcal{P})$ is defined through

$$E_D(\mathcal{P}) \equiv E_f - H_B(\mathcal{P}). \tag{21}$$

Instead of n which is the number of qubits needed to create state ρ_{AB}, one ought to define $H_B(\mathcal{P})$ with respect to E_f—the number of singlets needed to create the state under LOCC (called the entanglement of formation). The definition of $H_B(\mathcal{P})$ simply reflects the fact that not all the entanglement can be distilled to perform teleportations—there is "bound entanglement" [47]. Here, since the process is not necessarily optimal, $H_B(\mathcal{P})$ can be less than the bound entanglement. The relationship between bound entanglement and entropy (or heat) was discussed in Ref. [48].

Our definitions help elucidate the strong parallels between entanglement and local information. n separable pure states enable one to perform n bits of physical work, while E_f singlets allow one to perform E_f bits of quantum work such as teleportation. To create a state ρ_{AB}, Alice and Bob will also need to use n pure separable states but they will also need E_f singlets. The entropy $H_{LOCC}(\mathcal{P})$ prevents Alice and Bob from extracting the full n bits of local information, while the bound entanglement $H_B(\mathcal{P})$ prevents them from extracting the full E_f bits of nonlocal information.

The information-theoretic version of our complementarity relation takes a particularly simple form for pure states. For pure states, it was shown in [21] that $I_l = n - E_D = n - E_f$. We therefore have

$$H_{LOCC}(\mathcal{P},\psi) + H_B(\mathcal{P},\psi) \geqslant 2E_D(\psi). \tag{22}$$

D. Informational complementarity compared with entropic uncertainty relations

Although the relations given above may seem unfamiliar, they actually have a logical structure similar to the usual complementarity principle between noncommuting observables such as x and p.

The reason that Eqs. (16), (19), and (22) do not immediately strike one as being like the usual complementarity relationship, is because we are used to seeing them written like a Heisenberg uncertainty principle, such as

$$\Delta x \Delta p \geqslant \hbar, \tag{23}$$

or for general operators M, N, the Robertson inequality [17]

$$\Delta M \Delta N \geqslant \langle \psi [M,N] \psi \rangle. \tag{24}$$

However, it is now recognized that these inequalities can be better expressed as relationships between entropies. This approach to the uncertainty principle was begun by Białynicki-Birula and Mycielski [10] and later advocated by Deutsch [11], who was dissatisfied with the fact that the bound on the right-hand side of Eq. (24) is not a constant but instead, depends on the state. His bound was improved by Partovi [12], Kraus [13], and Maassen and Uffink [14]. The latter bound can be written as

$$H_{\hat{M}}(\psi) + H_{\hat{N}}(\psi) \geqslant -2\ln(\sup|\langle m|n\rangle|), \tag{25}$$

where m and n are the eigenstates of two operators $\hat{M}$ and $\hat{N}$, and entropies $H_{\hat{M}}$ and $H_{\hat{N}}$ of state ψ are the usual Shannon entropies defined, for example, by

$$H_{\hat{M}}(\psi) = -\sum_m |\langle m|\psi\rangle|^2 \ln|\langle m|\psi\rangle|^2. \tag{26}$$

That an uncertainty principle can be written in such a form makes intuitive sense because having a larger Shannon entropy in a certain basis corresponds to a larger uncertainty in measurements in that basis. The entropic uncertainty bound, therefore, expresses the trade off between the spread of measurement results in one basis (as given by the Shannon entropy in that basis), versus the spread of measurements in another basis. We, therefore, see that our complementarity principle is closely related to the more familiar one encountered in quantum mechanics.

For position and momentum, the Partovi bound takes the form

$$H_x(\psi) + H_p(\psi) \geqslant 2\ln[2/(1+\delta x\,\delta p/2\pi\hbar)] \tag{27}$$

for small values of $\delta x\,\delta p/2\pi\hbar$, where δx and δp are the resolution of the detector (i.e., phase space is divided into cells).

Comparing this to Eq. (22), we see that for pure states, the entanglement plays a role analogous to Planck's constant $\hbar$. The difference of course is that $\hbar$ is a constant that is independent of the state. In our case, fixing the amount of entanglement in the allowable states is equivalent to fixing $\hbar$ and the detector resolution. The right-hand side only depends on the amount of entanglement of the state and not on any other properties. It is the addition of entanglement into the system, which acts like $\hbar$ and creates this complementarity.

Our informational complementarity principle, expressed by Eq. (18), does have the appealing feature that the right-hand side is completely independent of the state. It has the form of the informational bound derived by Hall [15] for complementary observables, which is given by

$$I_{\hat{M}} + I_{\hat{N}} \leqslant \log_2 d, \tag{28}$$

where $I_{\hat{M}}$ and $I_{\hat{N}}$ gives the amount of information obtainable from a measurement of complementary observables $\hat{M}$ and $\hat{M}$, and d is the dimension of the Hilbert space. The similarity between this equation and Eq. (18) is striking.

E. Example: Drawing local and nonlocal information from pure states

We will now consider a protocol $\mathcal{P}$ on pure state, where both local and nonlocal information is extracted optimally. It will be used to show the balance between local and nonlocal information. We will also see that pure states can be thought of as being analogous to coherent states.

Essentially, the procedure is that Alice will perform a measurement which determines how much entanglement is available. Depending on the result of the measurement, the parties can choose whether they want to extract nonlocal information or local information. For example, they may choose to extract nonlocal information when they find a lot of entanglement (i.e., more than the average optimal amount E_D) and extract local information when there is a small amount of entanglement (since in this case, they can extract more local information than the optimal amount I_l).

The scenario is similar to the concentration of entanglement scheme of Ref. [49]. Alice and Bob share n pairs of a pure state $\psi_{AB} = a|00\rangle + b|11\rangle$. Alice performs a measurement with $n+1$ outcomes. As a result, Alice and Bob share a maximally entangled state with Schmidt rank $d_k = \binom{n}{k}$ with probability $p_k=\binom{n}{k}a^{2k}b^{2(n-k)}$, $k=0,\ldots,n$. The singlet is "diluted" into all 2n qubits. However, it is not a maximally entangled state of those $2n$ qubits, so that it can be swapped into a smaller number $\ln d_k$ of qubit pairs. Then, the remaining pairs will be in product states.

Each process $\rho\rightarrow\{p_k,\rho_k\}$, after which information I_k is extracted from ρ_k with probability p_k, provides $I_l=\Sigma_k p_k I_k - H(\{p\})$ of information. The Shannon entropy $H(\{p\})$ of distribution $\{p_k\}$ equals the cost of erasure of information, which allows Alice and Bob to work with an ensemble of ρ_k's [50]. Thus, in our example, Alice and Bob have to put $I_{er}=H(\{p\})$ of erasure to pay for the next part of the scheme in which they draw the $\Sigma_k p_k I_k$ amount of information.

In our protocol, Alice and Bob will decide whether to extract entanglement or local information based on the result of Alice's measurement (i.e., what value of k she measures). They divide the outcomes of k into two sets K_q and K_l.

If they obtain outcome $k\epsilon K_q$, they (i) concentrate the diluted singlet, obtaining on average

$$E_D(\mathcal{P})=\sum_{k\epsilon K_q} p_k \log_2 d_k \tag{29}$$

singlets; (ii) draw local information from the rest of qubits, obtaining, on an average,

$$I_{l_1}(\mathcal{P})=\sum_{k\epsilon K_q} p_k(2n-2\log_2 d_k). \tag{30}$$

If instead, they obtained the outcomes with $k\epsilon K_l$, they (iii) draw local information directly from the state with the average result

$$I_{l_2}(\mathcal{P})=\sum_{k\epsilon K_l} p_k(2n-\log_2 d_k). \tag{31}$$

Summing up all the information drawn from the system, we have

$$I(\mathcal{P})=E_D(\mathcal{P})+I_{l_1}(\mathcal{P})+I_{l_2}(\mathcal{P})-I_{er}, \tag{32}$$

which gives

$$I(\mathcal{P})=2n-\sum_{k=0}^{n} p_k \log_2 d_k - H(p). \tag{33}$$

Passing to intensive quantities $\tilde{i}\equiv I(\mathcal{P})/n$ is asymptotically equal to the maximum possible amount of local information per pair, which can be obtained starting with $|\psi_{AB}\rangle\langle\psi_{AB}|^{\otimes n}$, namely, $\tilde{i}\approx 2-S_A$, where S_A is the entropy of reduction of ψ_{AB}. Indeed, the last term in Eq. (33) is of order of $\ln n$, so that its contribution vanishes in the asymptotic limit. Thus, the complementarity inequality (16) is saturated. It follows that for pure states, it is possible to obtain partially quantum and partially local information without any loss. However, for mixed states, it is rather unlikely that an optimal protocol which saturates the inequality would exist.

Since the bound is saturated and Alice and Bob may obtain any amount of either $I_l(\mathcal{P})$ or $E_D(\mathcal{P})$ (up to their maximal value), we can therefore think of pure states as being analogous to coherent states, i.e., minimally uncertain states. Maximally entangled states are also the only ones which also saturate the constant bound of Eq. (18).

IV. COMPLEMENTARITY OF OBSERVABLES INDUCED BY LOCC

In the preceding sections, we introduced a complementarity principle between local and nonlocal information. Physically, it referred to general processes, rather than any particular implementation. It would therefore be useful to see if there is a complementarity principle which just refers to general measurements or operations. Indeed, we will find that when the implementation of a measurement is restricted to LOCC, it induces a type set of complementarities. One can generalize this further and consider complementarities when one is restricted to other classes of operations.

In Sec. IV A we will demonstrate that the parity and phase operator, which normally commute in quantum mechanics, no longer commute under LOCC. In Sec. IV B we will look at measurements that distinguish between the orthogonal states of Ref. [38].

A. The parity and phase observable

Consider two observables $\Sigma_x=\sigma_x^{(A)}\otimes\sigma_x^{(B)}$ and $\Sigma_z=\sigma_z^{(A)}\otimes\sigma_z^{(B)}$, where σ_x and σ_z are the usual Pauli spin matrices and the superscript refers to which subsystem it acts upon. They commute:

$$[\Sigma_x,\Sigma_z]=0, \tag{34}$$

and their eigenbasis is the Bell basis

$$\psi^{\pm}=|01\rangle\pm|10\rangle,$$

$$\Phi^{\pm}=|00\rangle\pm|11\rangle. \tag{35}$$

Σ_z measures the parity bit and will therefore, distinguish between the ψ and Φ eigenstates, while Σ_x measures the phase bit and will distinguish between the eigenstates that have a $+$ as the relative phase or a $-$. E.g., if one finds $\Sigma_z=0$ and $\Sigma_x=0$, then we have a singlet.

However, if Alice and Bob are restricted to LOCC operations and want to measure such observables on their shared system, it is impossible. Indeed, to measure Σ_x, Alice and Bob must separately measure σ_x, while to measure Σ_z they have to measure separately σ_z. Clearly, since σ_x does not commute with σ_z, they cannot measure both the parity and the phase. One might suspect that there could exist some complicated LOCC protocol that measures them jointly, somehow avoiding measuring directly local noncommuting observables. Later, we will show by a simple argument that this is impossible in general by any LOCC operation. However, here we would like to grasp the rough idea of the difference between the global and LOCC measurement.

To this end, note that in the distant labs case, Alice and Bob measure too much. Indeed, measuring Σ_x globally gives one bit of information (phase bit) because Σ_x has only two eigenvalues. In contrast, measuring the phase locally (by having Alice and Bob measure σ_x on their subsystem), the two parties will acquire two bits of information. Thus, in local measurements, the measurement is nondegenerate while Σ_x and Σ_z are degenerate. In fact, a local determination of parity and phase *must* acquire two bits of information.

We can simulate a global measurement of parity or phase by using the local operators

$$\Sigma_z^{LOCC} = \sigma_z^{(A)} + \alpha_z \sigma_z^{(B)}, \tag{36}$$

$$\Sigma_x^{LOCC} = \sigma_x^{(A)} + \alpha_x \sigma_x^{(B)}, \tag{37}$$

where parameters α act to break the degeneracy in operators Σ_x and Σ_z. Then, these local measurements of parity and phase no longer commute:

$$[\Sigma_x^{LOCC}, \Sigma_z^{LOCC}] = -i(\sigma_y^{(A)} + \alpha \sigma_y^{(B)}), \tag{38}$$

where we have redefined constants α. Thus, this measurement of parity and phase cannot be jointly measured under LOCC.

Here, we have only given one possible local realization of the parity and phase measurement. It therefore may be possible that one can find a clever procedure, perhaps involving positive operator valued measured (POVM's), such that the parity and phase measurements commute. This, however, cannot be the case.

Let us imagine that there exists local implementations of Σ_z and Σ_x, which are jointly measurable (i.e., commute). Then, we would be able to use these locally implementable operators to distinguish between the four Bell states. This is, however, impossible, as shown in Ref. [51]. Naively, this is because if one could distinguish between the Bell states, then one can produce entanglement from the identity state (which is separable). This would contradict the fact that entanglement cannot be created under LOCC. In fact, the problem is more subtle. One could, in principle, be able to distinguish the Bell states but in so doing, the entanglement could be destroyed. Indeed, in the case of two entangled states, one can distinguish them by LOCC [52]. For this case, the entanglement is necessarily destroyed during the measurement. However, distinguishing between the four Bell states would lead to entanglement creation under LOCC [51]. We therefore see that the parity and phase cannot be jointly measurable. In fact, parity and phase must be complementary, since if one were able to measure the parity and get even partial knowledge of the phase, then one would be able to create entanglement.

It is also interesting to ask how much entanglement is needed in order to implement the parity and phase operators in such a way that they commute. The answer is one bit of entanglement. To see this, we note that one bit of entanglement is clearly sufficient since Alice can use a singlet to teleport her qubit to Bob, who can then measure the parity and phase. One bit of entanglement must also be necessary, since measuring parity and phase under LOCC allows one to create one bit of entanglement [51]. Since one cannot create entanglement under LOCC, one must use up at least one bit of entanglement to make the measurement. It is therefore rather interesting that if we act the commutator of Eq. (38) on a separable state then we can get an entangled state, which is maximally entangled for $\alpha = 1$.

Finally, it is worth asking whether one can find other examples for $2 \otimes 2$ systems. In other words, are there other observables that commute globally but do not commute under LOCC. It appears that the number of examples is very limited. For pairs of product observables of the form $A \otimes B$, there is (up to local unitary transformations) only one other pair of operators which commute globally, namely, Σ_x^{LOCC} and Σ_z^{LOCC}.

The proof of this result is contained in the Appendix to this paper. For now, we simply state the result.

Proposition. If for some products of two qubit observables $[A \otimes B, C \otimes D] = 0$ then up to unitary product transformation $U_1 \otimes U_2$ and constant factor, one has

$$A = B = \hat{\sigma}_z, \tag{39}$$

$$C = D = \hat{\sigma}_x. \tag{40}$$

B. Distinguishing separable states

In Ref. [38], a set of nine states which are orthogonal and separable were presented. It was then proven, that although they are orthogonal, they cannot be distinguished under LOCC. These states (often referred to as "sausage states") therefore exhibit a form of nonlocality without entanglement. One can however distinguish between some of the states. There are therefore operators which can be constructed using the sausage states as basis states. As an example, we will construct two such operators, which although they commute globally, and are implementable locally, do not commute under LOCC.

The nine sausage states are (using a different numbering scheme from Ref. [38] for convenience)

$$\begin{array}{lll} & A & B \\ \psi_1 = & |0+1\rangle & |2\rangle \\ \psi_2 = & |0-1\rangle & |2\rangle \\ \psi_3 = & |0\rangle & |0+1\rangle \\ \psi_4 = & |0\rangle & |0-1\rangle \\ \psi_5 = & |1+2\rangle & |0\rangle \\ \psi_6 = & |1-2\rangle & |0\rangle \\ \psi_7 = & |1\rangle & |1\rangle \\ \psi_8 = & |2\rangle & |1+2\rangle \\ \psi_9 = & |2\rangle & |1-2\rangle. \end{array} \tag{41}$$

Now we can construct an operator O_1 which has as its eigenstates, the first seven states with seven different, nonzero eigenvalues and remaining two eigenstates with zero eigenvalues. We can also construct an operator O_2 which has ψ_8 and ψ_9 as eigenstates with two different, nonzero eigenvalues and remaining seven eigenstates with zero eigenvalues. These operators clearly commute globally since all the ψ_i are orthogonal. However, under LOCC, they clearly cannot commute. If they did, one could measure O_1 and O_2 simultaneously, and therefore, distinguish between all nine sausage states, in violation of the indistinguishability proof given in Ref. [38].

Now, O_2 can easily be implemented under LOCC. Bob can simply use projectors $|1+2\rangle$ and $|1-2\rangle$ to implement O_2. These projectors distinguish between ψ_8 and ψ_9. O_1 can also be implemented under LOCC, although some effort is needed. Consider, for example, an implementation O_1' which instead, has the following orthogonal eigenbasis.

$$\begin{array}{lll} & A & B \\ \psi_1= & |0+1\rangle & |2\rangle \\ \psi_2= & |0-1\rangle & |2\rangle \\ \psi_3= & |0\rangle & |0+1\rangle \\ \psi_4= & |0\rangle & |0-1\rangle \\ \psi_5= & |1+2\rangle & |0\rangle \\ \psi_6= & |1-2\rangle & |0\rangle \\ \psi_7= & |1\rangle & |1\rangle \\ \psi_{10}= & |2\rangle & |2\rangle \\ \psi_{11}= & |2\rangle & |1\rangle. \end{array} \tag{42}$$

The first seven eigenstates are identical to the eigenstates of O_1, and so O_1' is an implementation of O_1. Furthermore, O_1' can be implemented under LOCC using a sequence of von Neumann projection measurements, which was given in Ref. [38]. The detailed procedure is contained in Appendix B.

The difficulty is that while eigenvectors ψ_1–ψ_7 commute with O_2', the projectors onto ψ_{10} and ψ_{11} do not. If we write O_1' and O_2' as

$$O_1'=O_1+|2\rangle\langle 2|\otimes\sigma_z^{(B)}, \quad O_2'=|2\rangle\langle 2|\otimes\sigma_x^{(B)}, \tag{43}$$

where we once again use the Pauli matrices, this time written in the $|1\rangle$, $|2\rangle$ basis, then we find that while $[O_1,O_2]=0$, we have

$$[O_1',O_2']=-i|2\rangle\langle 2|\otimes\sigma_y^{(B)}. \tag{44}$$

Unlike the case of the parity-phase commutator, this commutator, operating on a separable state, cannot create entanglement. This may be related to the fact that for parity and phase, the eigenstates are entangled, while for O_1 and O_2, the eigenstates are separable.

C. LOCC complementarity inequalities

Although we have calculated the commutator for the LOCC implementation of some observables, it would be desirable to have a general expression for the LOCC commutator. However, the commutation relations have a disadvantage in that they also depend on the spectrum of eigenvalues, and therefore, do not capture the essence of complementarity. For a measurement, the eigenvalues are merely labels and the complementarity relation should not depend on them. The *entropic inequalities* which we discussed in Sec. III D are thus more appropriate. We will, therefore, quantify the extent to which LOCC induces complementarity by use of such inequalities. Recall the entropic uncertainty relation of Eq. (25):

$$H_M(\psi)+H_N(\psi)\geqslant-2\ln(\sup|\langle m|n\rangle|). \tag{45}$$

We can then define the LOCC entropic inequality as

$$\begin{aligned}[H_M(\psi)+H_N(\psi)]_{LOCC}&=min[H_{M'}(\psi)+H_{N'}(\psi)]\\ &\geqslant-2\ln(\sup|\langle m'|n'\rangle|),\end{aligned} \tag{46}$$

where the minimum is taken over all LOCC implementations N' and M' of operators N and M. For the parity and phase implementations of Eqs. (36) and (37), the right-hand side gives 2 and is independent of α (i.e., the eigenvalues).

Such a definition is still not ideal, as the entropic inequalities suffer from a problem which also plagues the Robertson inequality. Namely, two observables which share a common eigenvector will be said to "commute" even though they do not commute on a subspace of the total Hilbert space. This is an issue which is not particular to defining LOCC implementations, but exists, in general, with both the ordinary commutator and entropic inequalities. A solution exists, using a variation of the information theoretic bounds of Ref. [15] and we hope to address this point in the future.

Leaving such issues aside, the LOCC inequality can also be generalized to any class of operations. If we consider the set of all allowable operations A and a restricted subset of these $R\subseteq A$, then we can define a restricted inequality much in the way we have done here.

There is, however, one key difference between the interpretation of this inequality and the type of complementarity we are familiar with. Usually, complementarity and uncertainty are linked together. If two observables cannot both be measured on the same state (complementarity), then one necessarily finds uncertainty. Namely, if one measures one observable on half of an identically prepared ensemble and one measures the other observable on the other half of the ensemble, then one necessarily finds a dispersion in the results of the two measurements. One can now ask: does LOCC complementarity imply the uncertainty principle? It seems that the answer is *no*. Recall that a singlet, for example, has definite parity and phase. If Alice and Bob are given an ensemble of singlets and measure the phase on half the ensemble, they will always get the same result $(-)$. If they measure the parity on the other half of the ensemble, then also they will always get a definite result (0). There is nothing uncertain about what the outcome of a measurement of parity *or* phase will be. However, if Alice and Bob are given an unknown state (perhaps a singlet) then their measurement of phase will completely destroy their ability to determine what the parity is, and vice versa.

Thus, one may conclude that here, complementarity is decoupled from uncertainty. The main reason would be that the measurement does not prepare the system in the eigenstate of the observable we are trying to measure. Usually, the von Neumann postulate holds—after a measurement, the state is in an eigenstate of the observable. Therefore, it was hard to distinguish between complementarity and uncertainty (see Ref. [4]).

Note, however, that one can also phrase the uncertainty principle as meaning that one cannot prepare a state that would have a definite value of both parity and phase. Now, one can observe some asymmetry in our argument: we allow Alice and Bob to use only a restricted class of operations (LOCC), hence they cannot measure both phase and parity and complementarity emerges. Yet, we say uncertainty does not emerge because uncertainty-free states (e.g., the singlet) can be prepared. However, to prepare a singlet, one needs global unitary operations (or entanglement). Thus, in this case we allow for unrestricted operations. Now, if we will insist that the preparation of the state should also be done by LOCC, then there will not exist a state that is uncertainty-free and we will have uncertainty. Of course, the inequalities expressing this uncertainty will be obeyed only by states that can be prepared by LOCC (i.e., by separable states). We therefore have that for some states, one can decouple the uncertainty principle from complementarity while for LOCC-prepared states, the uncertainty principle and complementarity are linked.

V. DISCUSSION AND CONCLUSION

The concept of complementarity was introduced by Bohr before the discovery of quantum entanglement by Einstein, Podolsky, and Rosen [53] and Schrödinger [54], which is at the root of quantum communication. In fact, Bohr's complementarity concerns mutually exclusive quantum phenomena associated with a single system and observed under different experimental arrangements [55]. One can argue that this does not exhaust all complementary aspects of quantum phenomena as we often deal with complex systems that involve entanglement.

As noted in the Introduction, for Bohr, complementary observables "taken together" necessarily "exhausted all well-defined knowledge" of the system. It is natural to suppose that the best defined "knowledge" (information) is the one defined *operationally*. Then, information is necessarily either nonlocal or local. Having so defined the notions of local and nonlocal information, we have found that they are complementary under apparently contradictory well-defined processes: (i) extract local information to perform physical work, and (ii) extract nonlocal information to perform useful logical quantum work (teleportation). This complementarity is mathematically expressed by Eq. (16) or (18). However, it can be formulated in the spirit for Bohr's principle as follows: Complex quantum systems carry information that under well-defined mutually exclusive processes manifests itself as a local information or as nonlocal information. In this sense, the above principle can be viewed as a consistent extension of Bohr complementarity to complex (distributed) systems.

In particular, for the singlet state, it follows that under the above mutually exclusive process we can extract one bit of nonlocal information *or* one bit of locally accessible information. They are complementary in a sense that "taken together" they exhaust all well-defined mutual information of the singlet, which amounts to two bits.

In our approach, it is natural to view information as being a fundamental physical quantity. It is more convenient to think of information unfettered from the subject whose knowledge is usually represented by the information. If one maintains that the information encoded in quantum states represents "states of belief" (e.g., Ref. [56] in relation to the Copenhagen interpretation), then describing operations, such as our complementarity or teleportation or the workings of a quantum computer, becomes more uncomfortable.

There is a natural interpretation of information within the *generic information paradigm* (GIP) [18,19]. This was introduced for two basic reasons: (i) to overcome Bohr's quantum-classical dichotomy and (ii) to provide an ontological basis for quantum formalism. It implies, in particular, an informational interpretation of the quantum states, according to which their information content is *isomorphic* to information carried by real partial information fields. It is the information accessible, in principle, under well-defined physical situations. Then, the basic features "nonlocality" and "locality" of information encoded in quantum states reflects the double (*hylemorphic*) nature of *partial fields*, which inherently links two fundamental levels of reality: *logical* due to potential fields of alternatives and *physical* due to field of activities (events).

In addition to the complementarity between local and nonlocal information, we also found a complementarity that gets induced between operators implemented under LOCC. For example, we have shown that the parity and phase operators which normally commute in quantum mechanics are complementary under LOCC. We suggested using entropic uncertainty relations to quantify the degree to which observables are LOCC complementary. Such a relation is given in Eq. (46). How to interpret this quantity is an interesting open question. We further saw that this complementarity allows one to conceptually distinguish between the uncertainty principle and complementarity. It was argued that there is a sense in which the two concepts become decoupled.

It is also interesting that in order to implement both parity and phase measurements in such a way that they commute, we need one bit of entanglement. It therefore might be interesting to ask how much entanglement would be needed, such that one can jointly measure two observables. Quantifying this "entanglement assisted commutator" might help answer some of the questions raised here.

It also would be interesting to relate the complementarity principle between operators and between local and nonlocal information. The latter involves comparisons between two types of restricted operations (LO and LOCC), while the complementarity principle between operators only involves LOCC. However, both seem to involve the notion of entanglement. One possible direction is to note that the local information is believed to be equal to the right-hand side of the LOCC complementarity relation Eq. (46) for the optimal complete set of LOCC implementable observables [22].

In this paper, the part of nonlocal quantum information which was discussed was entanglement. However, the quantum deficit Δ is also nonzero for unentangled states (at least for a finite number of copies). It would, therefore, be of interest to also consider the case of date hiding [57], where Alice and Bob are essentially unable to obtain the local information encoded in a state. We believe this is related to the complementarity discussed here. Indeed, it appears that these notions of complementarity have many wide ranging applications.

Finally, the above results support the view that quantum states carry two complementary kinds of information, the local information which is locally accessible and nonlocal information which can be used for such tasks as teleportation (see, in this context, Ref. [7]). This complementarity lies at the foundations of quantum mechanics more deeply than it might seem. We believe that complementarity, in general, is a fundamental and intrinsic feature of information carried by physical systems which cannot be derived from any probabilistic models.

ACKNOWLEDGMENTS

This work was supported by EU grant EQUIP, Contract No. IST-1999-11053. M.H., P.H., and R.H. acknowledge the support of KBN Grant No. 2 P03B 103 16. We thank Jozef Gruska for discussions on trade-offs and Benjamin Schumacher for discussions on informational uncertainty relations. M.H. would like to thank Viacheslav Belavkin and Berthold-Georg Englert for discussions on uncertainty principles. J.O. would like to thank Jacob Bekenstein for interesting discussions and the Vergelle Institute for their hospitality while part of this work was being completed. He acknowledges Grant No. 129/00-1 of the Israel Science Foundation.

APPENDIX A: PROOF OF PROPOSITION 1

Let us provide a simple lemma first.

Lemma. If $X\otimes Y=R\otimes S$ for some operators X,Y,R,S then $X=\alpha R$, $Y=\alpha^{-1}S$ for some nonzero number α.

Proof of the above lemma is immediate. Without loss of generality we can consider X,Y to be of full rank and utilize their inverses (otherwise they are pseudoinverses) getting $I\otimes I=X^{-1}R\otimes Y^{-1}S$. Comparing the eigenvectors of both sides of the latter formula gives $Y^{-1}S=\alpha I$, $X^{-1}R=\alpha^{-1}I$, concluding the proof of the lemma.

Now we shall provide the simple proof of the following.

Proposition. If for some products of two qubit observables $[A\otimes B,C\otimes D]=0$ holds and one excludes trivial cases ($AC=0$, $BD=0$ or $[A,C]=[B,D]=0$) then up to unitary product transformations $U_1\otimes U_2$ and constant factor, one has

$$A=B=\hat{\sigma}_z, \tag{A1}$$

$$C=D=\hat{\sigma}_x. \tag{A2}$$

Proof. By adding and subtracting term $CA\otimes BD$, it is immediate that vanishing of the commutator from the Proposition is equivalent to

$$[A,C]\otimes BD=CA\otimes[D,B]. \tag{A3}$$

Applying the lemma to the above, we get

$$[A,C]=\alpha CA, \quad [D,B]=\alpha BD. \tag{A4}$$

Now, for two qubits we put $A=aI+\vec{a}\vec{\sigma}$, $B=bI+\vec{b}\vec{\sigma}$, $C=cI+\vec{d}\vec{\sigma}$, and $D=dI+\vec{d}\vec{\sigma}$. We then perform a simple calculation, taking into account the fact that (because of linear independence of I, σ_x, σ_y, and σ_z) absence of I on one side implies the same for the other side. This gives two equations:

$$\alpha[\vec{a}\vec{\sigma},\vec{b}\vec{\sigma}]=(a\vec{a}+c\vec{c})\vec{\sigma}+i\vec{a}\times\vec{c}\vec{\sigma},$$

$$\alpha[\vec{b}\vec{\sigma},\vec{d}\vec{\sigma}]=(b\vec{b}+d\vec{d})\vec{\sigma}+i\vec{b}\times\vec{d}\vec{\sigma}. \tag{A5}$$

Calculating the left-hand side (LHS) for both sides and using the linear independence of Pauli matrices, we finally get

$$a\vec{a}+c\vec{c}+i(1-2\alpha)\vec{a}\times\vec{c}=\vec{0},$$

$$b\vec{b}+d\vec{d}+i(1-2\alpha)\vec{b}\times\vec{d}=\vec{0}. \tag{A6}$$

Now, let us observe that (i) $\vec{a}\times\vec{c}\neq\vec{0}$ and hence, also (ii) $\vec{a}\neq\vec{0}$, $\vec{c}\neq\vec{0}$. Indeed, if $\vec{a}\times\vec{c}=\vec{0}$ then $[A,C]=0$ and consequently [see Eq. (A3)] either $AC=0$ (which is trivial because then, $A=A^\dagger$, $C=C^\dagger$ and also $CA=0$ and $[A\otimes B,C\otimes D]$ vanishes automatically) or $AC\neq 0$ but then [again because of Eq. (A3)] also $[B,D]=0$, which would be trivial again.

Because of (i) and (ii), the LHS of the first line of Eq. (A6) is a linear combination of three *nonzero* and *linearly independent* vectors $\vec{a}$, $\vec{c}$, $\vec{a}\times\vec{c}$, so all the coefficients in the combination must vanish, giving, in particular, $a=c=0$. In a similar way we get $b=d=0$. This simplifies our observables: $A=\vec{a}\vec{\sigma}$, $B=\vec{b}\vec{\sigma}$, $C=\vec{d}\vec{\sigma}$, and $D=\vec{d}\vec{\sigma}$. Putting them again into Eq. (A4), we immediately get

$$(\vec{a}\times\vec{c}\vec{\sigma})\otimes(\vec{b}\vec{d}I+i\vec{b}\times\vec{d}\vec{\sigma})=(\vec{a}\vec{c}I+i\vec{c}\times\vec{a}\vec{\sigma})\otimes(\vec{b}\times\vec{d}\vec{\sigma}), \tag{A7}$$

which, for nonzero $\vec{a}\times\vec{c}$ and $\vec{b}\times\vec{d}$ is satisfied iff $\vec{a}\vec{c}=\vec{b}\vec{d}=0$. We can put $\vec{a}=\vec{b}=\hat{z}$ since we can always choose such a local basis for Alice and Bob. We then have $\vec{c}=\vec{d}=\hat{x}$ (again using our choice of label for the direction orthogonal to $\hat{z}$). This concludes the proof of the Proposition.

APPENDIX B: MEASUREMENT OF SEVEN SAUSAGE STATES UNDER LOCC

Here we show how to implement O_1' using a *ping-pong* process between Alice and Bob. Essentially, the procedure is:

(b1) Bob first does a projection on $|2\rangle$ and communicates his result to Alice.

(a1) If his result is positive, then Alice can project onto the three states $|0+1\rangle$, $|0-1\rangle$ which will distinguish between ψ_1 and ψ_2. However, if she finds neither ψ_1 or ψ_2, she will know that the state is ψ_{10}, which is, in some sense, superfluous information which she would not get if she were measuring O_1 globally.

(a1′) If Bob's first projection yielded a negative result, then Alice, instead, projects onto the $|0\rangle$ state and communicates her result to Bob.

(b2) If her projection found state $|0\rangle$ then Bob projects onto $|0+1\rangle$ and $|0-1\rangle$, which distinguishes between ψ_4 and ψ_5.

(b2′) If her result was negative, Bob projects onto $|0\rangle$ and $|1\rangle$ and communicates the result to Alice.

(a2) Alice can then make the final orthogonal projection, either onto $|1+2\rangle$ and $|1-2\rangle$, or onto $|1\rangle$ and $|2\rangle$, depending on Bob's result. This distinguishes between ψ_5, ψ_6, and ψ_7, as desired, but it also singles out ψ_{11}, which is again, surplus information which is not required to implement O_1.

[1] W. Heisenberg, Z. Phys. **43**, 172 (1927).

[2] N. Bohr, Nature (London) **121**, 580 (1928).

[3] N. Bohr, Phys. Rev. **48**, 696 (1935).

[4] P. Bush, M. Grabowski, and P.J. Lahti, *Operational Quantum Physics* (Springer, New York, 1995).

[5] N. Bohr, *Essays 1958–1962 on Atomic Physics and Human Knowledge*, Vol. 121 (Interscience, New York, 1963).

[6] M. Horodecki and R. Horodecki, Phys. Lett. A **244**, 473 (1998).

[7] R. Horodecki, M. Horodecki, and P. Horodecki, Phys. Rev. A **63**, 022310 (2001).

[8] C. Bennett and S. Wiesner, Phys. Rev. Lett. **69**, 20 (1992).

[9] C.H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres, and W.K. Wootters, Phys. Rev. Lett. **70**, 1895 (1983).

[10] I. Białynicki-Birula and J. Mycielski, Commun. Math. Phys. **44**, 129 (1975).

[11] D. Deutsch, Phys. Rev. Lett. **50**, 631 (1983).

[12] M. Partovi, Phys. Rev. Lett. **50**, 1883 (1983).

[13] K. Kraus, Phys. Rev. D **35**, 3070 (1987).

[14] H. Maassen and J. Uffink, Phys. Rev. Lett. **60**, 1103 (1988).

[15] M. Hall, Phys. Rev. Lett. **74**, 3307 (1995).

[16] M. Grabowski, Phys. Lett. A **124**, 19 (1987).

[17] H. Robertson, Phys. Rev. **34**, 163 (1929).

[18] R. Horodecki, in *Proceedings of International Conference on Problems in Quantum Physics: Gdańsk'87*, edited by L. Kostro, A. Posiewnik, J. Pykacz, and M. Zukowski (World Scientific, Singapore, 1988).

[19] R. Horodecki, in *Proceeding of International Conference on Problems in Quantum Physics: Gdańsk'89*, edited by J. Mizerski, A. Posiewnik, J. Pykacz, and M. Zukowski (World Scientific, Singapore, 1990).

[20] R. Horodecki, Ann. Phys. (Leipzig) **48**, 479 (1991).

[21] J. Oppenheim, M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **89**, 180402 (2002).

[22] M. Horodecki, K. Horodecki, P. Horodecki, R. Horodecki, J. Oppenheim, A. Sen(De), and U. Sen, Phys. Rev. Lett. **90**, 100402 (2003).

[23] M. Horodecki, P. Horodecki, and J. Oppenheim, Phys. Rev. A **67**, 062104 (2003).

[24] Although we will only discuss bipartite states, the analysis presented here applies equally well for multipartite states. We also work with states made up of qubits but one could just as well consider more general states.

[25] B. Schumacher, Phys. Rev. A **51**, 2738 (1995).

[26] M. Ohya and D. Petz, *Quantum Entropy and Its Use* (Springer, New York, 1993).

[27] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).

[28] By classical system we mean a *d*-level system, the states of which are probability distributions over *d* items.

[29] More generally, it was shown that the mutual information I_M describes the capacity of noisy channel assisted by entangled states. See H. Bennett, P.W. Shor, J.A. Smolin, and A.V. Thapliyal, Phys. Rev. Lett. **83**, 3081 (1999).

[30] N. Cerf and C. Adami, Phys. Rev. Lett. **79**, 5194 (1997).

[31] H. Ollivier and W.H. Żurek, e-print quant-ph/0105072.

[32] L. Henderson and V. Vedral, J. Phys. A **34**, 6899 (2001).

[33] B.M. Terhal, M. Horodecki, D.P.D. Vincenzo, and D. Leung, e-print quant-ph/0202044.

[34] L. Szilard, Z. Phys. **53**, 840 (1929).

[35] J. von Neumann, *Mathematische Grundlagen der Quantenmachanik* (Springer, Berlin, 1932), translated into English and reprinted in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. Zurek (Princeton University Press, Princeton, 1983).

[36] We can think of the classical channel as a dephasing channel which transforms state ρ into a state $\rho' = \Sigma_i P_i \rho P_i$, where P_i are orthogonal projection operators which are chosen by Alice and Bob.

[37] E. Rains, Phys. Rev. A **60**, 179 (1999).

[38] C.H. Bennett, D.P. DiVincenzo, C.A. Fuchs, T. Mor, E. Rains, P.W. Shor, J. Smolin, and W.K. Wootters, Phys. Rev. A **59**, 1070 (1999).

[39] R. Werner, e-print quant-ph/0003070.

[40] For example, data hiding [57] might also be considered to be a form of quantum work.

[41] The rest, $\Delta - E_D$ is still the quantum part of correlations, yet it perhaps should not be called "nonclassical." Consider, for example, a separable state for which $\Delta \neq 0$. Such states can be created by LOCC, hence it is unclear how to understand its nonlocality. In Ref. [38] it was understood as follows: "One can do better, if the particles are together." However, since only LOCC is allowed this may be interpreted that we do not have classical correlations of classical local properties. Yet, we can still imagine that we have classical correlations between *local quantum properties*. The latter cannot be real, hence one cannot handle them by LOCC, yet nonlocality is not involved.

[42] A CNOT, or *controlled not* gate, flips the state of a target qubit if the control qubit is in the $|1\rangle$ state, and does nothing otherwise.

[43] To ensure that we are actually extracting information from the state and not the measuring device, one needs to reset the measuring device after the entire procedure.

[44] C.H. Bennett, D.P. DiVincenzo, J. Smolin, and W.K. Wootters, Phys. Rev. A **54**, 3824 (1996).
[45] After this paper had been submitted, we were able to show [22], subject to a reasonable assumption, that $H_{LOCC}(\mathcal{P})$ is in fact the Shannon entropy in a particular basis.
[46] This operational approach has been advocated by W. Unruh (private communication), and is also contained in the work of Janes and Gibbs [E.T. Janes, in *Maximum Entropy and Bayesian Methods*, edited by C.R. Smith, G.J. Erickson, and P.O. Neudorfer (Kluwer Academic Publishers, Dordrecht, Holland, 1992)].
[47] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **80**, 5239 (1998).
[48] P. Horodecki, M. Horodecki, and R. Horodecki, Acta Phys. Slov. **48**, 144 (1998).
[49] C.H. Bennett, H. Bernstein, S. Popescu, and B. Schumacher, Phys. Rev. A **53**, 2046 (1996).
[50] R. Landauer, IBM J. Res. Dev. **5**, 183 (1961).
[51] S. Ghosh, G. Kar, A. Roy, A. Sen(De), and U. Sen, Phys. Rev. Lett. **87**, 277902 (2001).
[52] J. Walgate, A. Short, L. Hardy, and V. Vedral, Phys. Rev. Lett. **85**, 4972 (2000).
[53] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[54] E. Schrödinger, Naturwissenschaften **23**, 807 (1935).
[55] Consider, for example, the recent two-slit experiment [M.O. Scully, B.-G. Englert, and H. Walther, Nature (London) **351**, 111 (1991)] involving rubidium atoms where the "wave-particle" complementarity is due to entanglement of atoms with micromasser cavities. One could view this as giving a more fundamental role to complementarity versus the uncertainty principle. [See, in this context, S. Durr, T. Nonn, and G. Rempe, *ibid.* **395**, 33 (1998); see also Ref. [58]].
[56] C.M. Caves, C.A. Fuchs, and R. Schack, Phys. Rev. A **66**, 062111 (2002).
[57] D.P. DiVincenzo, D.W. Leung, and B.M. Terhal, e-print quant-ph/0103098.
[58] O. Pessoa, Phys. Essays **13**, 50 (2000).