Types of Quantum Information

Robert B. Griffiths*

*Department of Physics, Carnegie-Mellon University,
Pittsburgh, PA 15213, USA*

Quantum, in contrast to classical, information theory, allows for different incompatible types (or species) of information which cannot be combined with each other. Distinguishing these incompatible types is useful in understanding the role of the two classical bits in teleportation (or one bit in one-bit teleportation), for discussing decoherence in information-theoretic terms, and for giving a proper definition, in quantum terms, of "classical information." Various examples (some updating earlier work) are given of theorems which relate different incompatible kinds of information, and thus have no counterparts in classical information theory.

PACS numbers:

I. INTRODUCTION

Despite an enormous number of publications in the field of quantum information (see [1, 2] for useful introductions), neither the fundamental principles underlying the subject, nor its connection with classical information theory as developed by Shannon and his successors [3], is altogether clear. On the one hand there has been some dispute [4, 5] about whether Shannon's ideas can be applied at all in the quantum domain. On the other hand there have been suggestions that the connection with Shannon's ideas occurs only for macroscopic systems or asymptotically large N (number of transmissions, or whatever) limits, as in what is sometimes called "quantum Shannon" theory [6, 7]. The author's position is that, to the contrary, there are perfectly consistent ways of applying the basic ideas of classical information theory to small numbers (even one) of microscopic quantum systems provided attention is paid to the Hilbert space structure of quantum theory, and probabilities are introduced in a consistent fashion. And, further, that this approach has advantages in that simple systems are simpler to think about than complicated systems, so it is useful to develop some intuition as to how they behave. One goal is to understand both classical and quantum information theory in fully quantum terms, since the world is (most physicists believe) fundamentally quantum mechanical.

The basic strategy of this paper is based on the idea that quantum information comes in a variety of incompatible *types* or *species*. Each type or species refers to a certain class of (typically microscopic) mutually-compatible properties of a quantum system. As long as the discussion of information is limited to a single type, all the usual formalism and intuition provided by classical information theory apply directly to the quantum

domain. On the other hand, incompatible types of quantum information cannot be combined, as this makes no sense in the context of standard Hilbert space quantum mechanics. Since in classical information theory there is only a single type or species of information, or, equivalently, all different types are compatible with each other, the main respect in which quantum information theory needs to go beyond its classical counterpart is in relating incompatible types of information in a useful way. Various examples are given below.

In the real (i.e., quantum) world it must, of course, be the case that so-called "classical" information, as in the acronym LOCC, "local operations and *classical* communication," is describable in quantum terms. A relatively precise definition can be given as indicated in Sec. IV: classical information is a particular type of quantum information, the only one that survives under circumstances (implicitly assumed in much writing on the subject) where there is strong decoherence.

The remainder of this paper is organized as follows. The concept of quantum information types is introduced in the context of a discussion of quantum incompatibility in Sec. II. The idea is illustrated by the simple examples of one- and (standard) two-bit teleportation in Sec. III, where using quantum information types and what we call the Presence theorem helps understand why one or two bits, respectively, are needed in these protocols, or "dits" in their d-dimensional ("qudit") generalizations. Decoherence and "classical" information are the subjects of Sec. IV, which begins with a simple beam splitter example that illustrates the importance of the Exclusion theorem, derived from a more general Truncation theorem, and this sets the stage for a proper understanding of classical information in quantum terms.

Quantum information theory requires theorems that relate different types of information, and hence go beyond anything in classical information theory. Those used in Secs. III and IV and some others closely related to them are given precise formulations in Sec. V, extending earlier work in [8]. They all have the "smell" of no-cloning [9],

---

*Electronic address: rgrif@cmu.edu

but the connection is not altogether straightforward, as shown by an additional Generalized No-Cloning theorem. Proofs and some additional technical details are found in the appendices. A summary and an indication of various ways the present work needs to be extended comprise the concluding Sec. VI.

II. TYPES OF INFORMATION

Central to the following discussion will be the notion of *quantum incompatibility* [10], which can be illustrated using the familiar two-dimensional Hilbert space of a spin-half particle. Each one-dimensional subspace or ray, which is to say all complex multiples of a fixed nonzero ket $|w\rangle$, is associated with the property that a particular component of angular momentum is positive, $S_w = +1/2$ (in units of $\hbar$) for some direction w in space, e.g., $w = z$ or $w = x$ or $w = -z$, etc. The negation of the property (or proposition) $S_w = +1/2$ is the property $S_w = -1/2$, corresponding to the orthogonal complement of the ray associated with $S_w = +1/2$. In the notation commonly used in quantum information theory, $S_z = +1/2$ and $-1/2$ correspond to rays passing through (i.e., multiples of) $|0\rangle$ and $|1\rangle$, respectively, and of course these are orthogonal, $\langle 0|1\rangle = 0$.

It always makes sense to talk about the conjunction P AND Q of two properties P and Q of a classical system, such as $p > 0$ and $x < 0$ for a harmonic oscillator. The result may be a property that is always false, as when P is $p > 0$ and Q is $p < 0$, and in this case the negation (NOT P) OR (NOT Q) of P AND Q is the property $p \leq 0$ OR $p \geq 0$, which is always true. But in quantum theory it is possible to write down conjunctions, such as

$$S_x = +1/2 \text{ AND } S_z = +1/2, \tag{1}$$

which make no sense. Obviously (1) cannot correspond to any ray in the Hilbert space, since each ray is associated with $S_w = +1/2$ for some direction w, and (1) is not of this form. Could it be a proposition that is always false? Then its negation $S_x = -1/2$ OR $S_z = +1/2$ must always be true, which does not seem very plausible. Indeed, assuming that (1) and similar conjunctions are false swiftly leads to a contradiction if one follows the usual rules of logic—for details, see Sec. 4.6 of [11]. This was understood by Birkhoff and von Neumann [12], who proposed altering the rules of propositional logic to get around this difficulty. Their proposal has not been of much use for interpreting quantum mechanics, which may merely mean that we physicists are not smart enough. By contrast, if one restricts the domain of meaningful discourse so as to exclude (1) and similar things—in particular, conjunctions (AND) and disjunctions (OR) of properties corresponding to projectors that do not commute—it is possible to produce a consistent interpretation of quantum mechanics [11, 13–20] that follows the usual rules of logic (as applied to meaningful statements), and resolves all the standard paradoxes [21].

The compatible propositions $S_z = +1/2$ and $S_z = -1/2$, corresponding to mutually orthogonal projection operators, form a *quantum sample space* of mutually exclusive possibilities: their conjunction is always false, and since each is the negation of the other, one or the other is always true. This makes physical sense in that one can in principle carry out a Stern-Gerlach measurement to determine whether $S_z = +1/2$ or $-1/2$ [22]. (By contrast, there is no measurement which can determine the truth or falsity of (1), as one would expect for something that is meaningless.) Information that answers the question of whether $S_z = +1/2$ or $-1/2$ is what we shall call the $\mathcal{Z}$ *type* (or *species*) of information. Similarly, $\mathcal{X}$ information answers the question whether $S_x = +1/2$ or $-1/2$. It is incompatible with $\mathcal{Z}$ information in that there is no way in which the two can be meaningfully combined: (1) makes no sense, and asking whether $S_z = +1/2$ or $S_x = +1/2$ is equally meaningless. For a spin-half particle there is a type of information associated with each pair w and $-w$ of opposite directions in three-dimensional space, and the different species associated with distinct pairs are incompatible.

In larger Hilbert spaces a quantum sample space or type of information always corresponds to a *decomposition of the identity*, a collection of mutually orthogonal projectors $\mathcal{V} = \{V_j\}$, $V_j^\dagger = V_j = V_j^2$, that sum to the identity I. In the case of an orthonormal basis, $V_j = |v_j\rangle\langle v_j|$ and $\langle v_j|v_k\rangle = \delta_{jk}$, we also write $\mathcal{V} = \{|v_j\rangle\}$, since the meaning is obvious. Two such collections or types of information $\mathcal{V}$ and $\mathcal{W}$ are *compatible* if and only if all projectors in one commute with all projectors in the other; otherwise they are *incompatible*. The "single framework" rule of quantum reasoning [23] generalizes the example discussed above, and states that incompatible quantum descriptions (decompositions, information types) cannot be meaningfully combined.

III. TELEPORTATION

Let us see how using incompatible types of information assists in understanding why quantum teleportation uses two classical bits of information in the standard protocol [24]. It is simplest to start with a variant known as "one bit" teleportation [25], corresponding to the quantum circuit in Fig. 1(a), where the teleportation process transports the state $|\psi\rangle$ from the upper left a to the lower right b'. First, a CNOT, shown as a controlled-X (CX) gate, acts between qubits a and b, and then an S_x measurement is carried out on qubit a. In the figure this measurement is indicated by the Hadamard gate H that interchanges S_x and S_z, followed by a measurement in the standard or S_z or "computational" basis indicated by the D-like symbol. If the measurement reveals $S_x^a = -1/2$ a classical bit (dashed line labeled x) is transmitted to where it actuates a Z gate on qubit b, whereas if $S_x^a = +1/2$ nothing is done. It is an easy exercise to show that whatever initial state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ enters at a will later
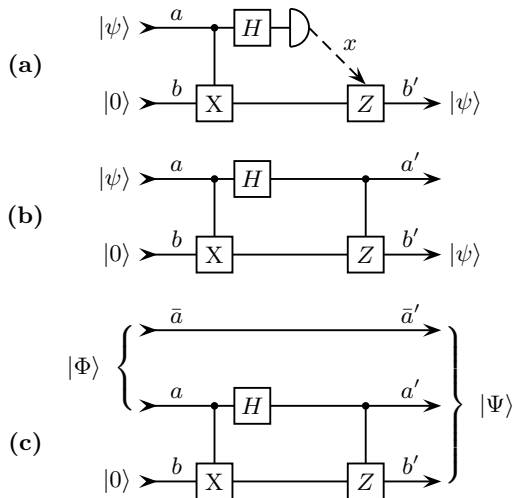
FIG. 1: (a) One bit teleportation. (b) Quantized version. (c) Channel ket construction.

reappear at b'.

In the case of $\mathcal{Z}$ information, meaning the input is $|\psi\rangle = |0\rangle$ or $|1\rangle$, corresponding to $S_z^a = +1/2$ or $-1/2$, the CX (CNOT) gate copies it from the a to the b qubit so that $S_z^b = S_z^a$, and the later Z gate has no effect, since even if it acts it only changes the phase of $|1\rangle$, leaving the ray (or projector) corresponding to $S_z^b = -1/2$ the same. Thus failing to do the measurement, or throwing away the classical bit, has no influence so far as transporting the $\mathcal{Z}$ information is concerned.

In the case of $\mathcal{X}$ information the input $|\psi\rangle$ is either $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ or $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$, corresponding to $S_x = +1/2$ or $-1/2$, and the analysis is somewhat more complicated. The CX gate maps $|\psi\rangle = |+\rangle$ into the two qubit state $|++\rangle + |--\rangle$ corresponding to $S_x^b = S_x^a$, and $|\psi\rangle = |-\rangle$ into $|+-\rangle + |-+\rangle$, $S_x^b = -S_x^a$. This means the original $\mathcal{X}$ information is not present in either qubit by itself, since the corresponding reduced density operator is $\frac{1}{2}I$, but resides in a correlation between the two. Information residing in a correlation is not in itself a quantum effect. One can, for instance, encode a classical bit $\{0_L, 1_L\}$ in two coding bits by letting 00 or 11, chosen at random, represent 0_L, and 01 or 10, again chosen at random, represent 1_L. From either coding bit alone one can extract no information about 0_L versus 1_L, but it is obviously present in the two together through their correlation. In the case under discussion the measurement in Fig. 1(a) extracts the value of S_x^a *after* the CX has acted (note that this is *not* the original $\mathcal{X}$ information), and if this is negative the Z gate applied to qubit b changes the sign of S_x^b. The net effect is that at the end of the process the value of S_x^b is exactly the same as that of S_x^a at the beginning, so the $\mathcal{X}$ information has also been successfully transmitted from a to b'.

One could continue to check what happens to other

types of information, but that is not necessary. The Presence theorem of Sec. V A says that once it is known that *two* suitably incompatible types of information, $\mathcal{Z}$ and $\mathcal{X}$ in the case at hand, are present in the output b', all other types of information about the input a are also present, so there is a perfect quantum channel from a to b', the desired result for teleportation. In summary, the transmission of two (suitably incompatible) types of information is needed to ensure that there is a good quantum channel from a to b'. The CX gate by itself transmits the $\mathcal{Z}$ type, while the later measurement and the single classical bit carrying its outcome are needed to transmit the incompatible $\mathcal{X}$ type of information.

The Presence theorem is a statement about *quantum* information discussed in fully *quantum* terms, so to apply it to the system in Fig. 1(a) one needs to understand the measurement and the "classical" bit in quantum terms. This can be done in the manner indicated in Sec. IV. But for present purposes it is convenient to avoid having to introduce the Hilbert space of a complicated macroscopic system, by "quantizing" the circuit in Fig. 1(a) so that it takes the form shown in (b), with the measurement of the a qubit following the part of the circuit where it controls (in the usual quantum sense) a Z gate. (See [26] and pp. 186f in [1] for this "trick," based on ideas in [27].) The two circuits in (a) and (b) are equivalent so far as teleportation is concerned, but the second is simpler to analyze in fully quantum terms. Indeed, the later measurement of qubit a in (b) need not be made at all, which is why it is not shown, as its outcome is not used in the protocol. (This discussion continues in the latter part of Sec. IV.)

In addition, the Presence theorem is stated in Sec. V A in the language of entangled kets, rather than in terms of the input and output of a quantum channel. One way of connecting the two is indicated in Fig. 1(c), where an auxiliary qubit $\bar{a}$ has been introduced, and

$$|\Phi\rangle = \left(\sum_j |a_j\rangle \otimes |\bar{a}_j\rangle\right)/\sqrt{2}, \tag{2}$$

where $\{|a_j\rangle\}$ and $\{|\bar{a}_j\rangle\}$ are orthonormal bases of $\mathcal{H}_a$ and $\mathcal{H}_{\bar{a}}$, is a fully-entangled state. The result is a final state $|\Psi\rangle \in \mathcal{H}_{\bar{a}} \otimes \mathcal{H}_a \otimes \mathcal{H}_b$, referred to as a "channel ket" in [8] (which see for additional details). That there is a perfect quantum channel from a to b' in Fig. 1(a) or (b) is the same as saying that all the information about qubit $\bar{a}$ is in qubit b' if one uses $|\Psi\rangle$, or equivalently the reduced density operator obtained by tracing $|\Psi\rangle\langle\Psi|$ over $\mathcal{H}_b$, to generate probabilities for correlations between the two in the usual way.

An alternative way to associate channels with kets is to use *map-state duality* [8, 28] in which an entangled ket

$$|\psi\rangle = \sum_j |a_j\rangle|\phi_j\rangle \tag{3}$$

on the tensor product of $\mathcal{H}_a$ and another space $\mathcal{H}_f$ is expanded in an orthonormal basis of $\mathcal{H}_a$, with $\{|\phi_j\rangle\}$ the

(unnormalized) expansion coefficients. One can always "transpose" $|\psi\rangle$ to an operator

$$M = \sum_j |\phi_j\rangle\langle a_j| \qquad (4)$$

mapping $\mathcal{H}_a$ to $\mathcal{H}_f$. In the particular case in which $|\psi\rangle$ is maximally entangled, which is to say $\mathrm{Tr}_f(|\psi\rangle\langle\psi|)$ is proportional to the identity I_a, the map M is, up to normalization, an isometry (a unitary operator from $\mathcal{H}_a$ to the subspace $M\mathcal{H}_a$ of $\mathcal{H}_f$), that is, $M^\dagger M = I_a$. Conversely, given a map M from $\mathcal{H}_a$ to $\mathcal{H}_f$ and an orthonormal basis of $\mathcal{H}_a$, it can be expanded in the form (4), and (3) then defines a corresponding entangled state (typically not normalized), which when M is an isometry is maximally entangled. Of course, if $|\psi\rangle$ is given, M depends on the choice of orthonormal basis $\{|a_j\rangle\}$, and vice versa, but the Presence theorem is unaffected by the basis choice. While this and the other theorems in Sec. V can be expressed either in "map" or "entanglement" language, the latter has the advantage of being more symmetrical (see remarks in Sec. VI of [8]). The idea of regarding the input and output of a quantum channel as corresponding to the tensor product of two Hilbert spaces, as suggested by the preceding discussion, is a very natural notion when using atemporal diagrams [29], and within the consistent histories approach to probabilistic time development [11], where the idea of such a tensor product goes back to Isham [30].
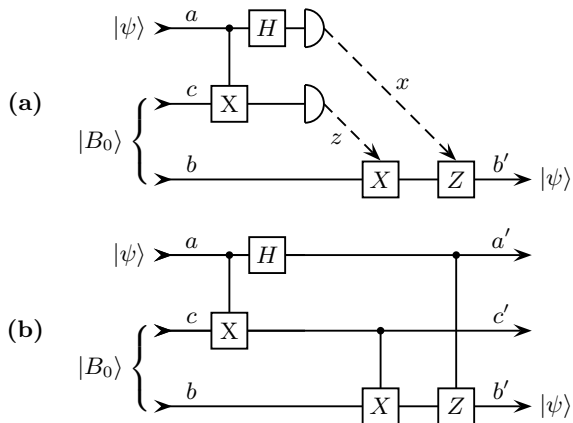

FIG. 2: (a) Standard (two bit) teleportation. (b) Quantized version.

Conventional "two bit" teleportation, Fig. 2, with $\sqrt{2}|B_0\rangle = |00\rangle + |11\rangle$, can be analyzed in the same way; the details are left as an exercise. The two classical bits in (a) are labeled x and z to indicate that they are essential for correct transmission of the $\mathcal{Z}$ and $\mathcal{X}$ information, respectively; throwing z away will not affect $\mathcal{X}$ information, and x is dispensable if only $\mathcal{Z}$ information is of interest. Neither classical bit, nor the two together, actually contain any information in themselves, a con-

sequence of No Splitting, Sec. VB. All the information is in correlations of the classical bits with the b qubit (more details in[31]). One classical bit is needed for each of these two incompatible types of quantum information, but that is enough, for the Presence theorem then guarantees that all other species are correctly transmitted, so one has a perfect quantum channel from a to b'. Figure 2(b) is the quantized version of (a), which is convenient for analyzing the situation in quantum terms, and one could once again introduce a channel ket $|\Psi\rangle$ (not shown), this time on four qubits, in analogy with Fig. 1(c). The Presence theorem can be applied to the channel ket, or one can regard the channel entrance as constituting its own Hilbert space. Generalization to the teleportation of a qudit (with Hilbert space of dimension $d > 2$) is straightforward: if two suitably incompatible types of information are correctly transmitted—each type requires a classical "dit"—one has a perfect quantum channel.
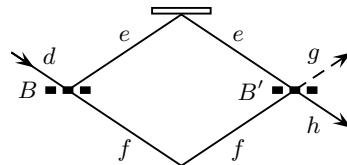
IV. DECOHERENCE AND CLASSICAL INFORMATION

FIG. 3: Interferometer with beamsplitters B and B'.

An application of types of information to a simple case of decoherence is shown in Fig. 3, where a particle (neutron or photon) enters an interferometer on path d at beamsplitter B and, because at an intermediate time it is in a coherent superposition $(|e\rangle + |f\rangle)/\sqrt{2}$, leaves the second beamsplitter B' in channel h rather than g. But if while inside the interferometer some interaction with the environment leaves a trace indicating that the particle took path e rather than f, or vice versa, the interference effect is lost, and the particle emerges with equal probability in g or h. Let $\mathcal{Z}$ be the e vs. f "which way" information, and $\mathcal{X}$ be the $(|e\rangle + |f\rangle)/\sqrt{2}$ vs. $(|e\rangle - |f\rangle)/\sqrt{2}$ "coherent" information. Decoherence, the disappearance of coherence, in this case $\mathcal{X}$ information, when $\mathcal{Z}$ information about the path resides in the environment, illustrates the Exclusion theorem of Sec. V: one type of information about $\mathcal{S}_a$ perfectly present in $\mathcal{S}_b$ means that a mutually-unbiased type is completely absent from $\mathcal{S}_c$. Two types of information $\mathcal{X}$ and $\mathcal{Z}$ are said to be mutually unbiased if they correspond to mutually unbiased orthonormal bases $\{|x_j\rangle\}$ and $\{|z_k\rangle\}$, with $|\langle x_j|z_k\rangle|^2$ equal to 1 divided by the dimension of the Hilbert space, independent of j and k.

To apply this theorem to the situation in Fig. 3, think of the particle that has just passed through the first beam splitter as system $\mathcal{S}_a$, and just before it reaches the second beam splitter B' as $\mathcal{S}_c$, while $\mathcal{S}_b$ is the environment at this second time. (See the discussion in Sec. III on why one can regard the particle at two different times as two separate systems, and how to apply the Exclusion theorem, worded in terms of entangled states, to situations with unitary time evolution.) For our purposes it suffices to model $\mathcal{S}_a$ and $\mathcal{S}_c$ using a $d = 2$ dimensional Hilbert space spanned by $|e\rangle$ and $|f\rangle$—this is analogous to focusing on the spin of a particle when its other degrees of freedom are not relevant to the analysis. The Exclusion theorem says that when the $\mathcal{Z}$ or which-way information about $\mathcal{S}_a$ is perfectly present in the environment, i.e., at the time the particle reaches the second beam splitter, the (mutually unbiased) $\mathcal{X}$ or coherence information must be perfectly absent from $\mathcal{S}_c$, i.e., from the particle itself at this later time. And in the absence of coherence all interference effects disappear: the situation after the second beam splitter is, statistically, just the same as if the particle arrived at random on path e or path f. All this is well known, and the connection between decoherence and information in the environment has been previously pointed out by Zurek and his collaborators [32–35]. The use of *types* of information, not tied to some notion of measurement [36], is our attempt to add further clarity and precision to these seminal ideas.

The situation to which the Exclusion theorem applies is that of *strong*, meaning essentially complete, decoherence. Clearly extensions are needed (Sec. VI) to cases of only partial decoherence. Nevertheless, strong decoherence is a useful idealization both because if is often a good approximation to what is realized in the laboratory (to the dismay of those who want to build quantum computers), and because it yields a precise definition of another idealization, *classical information*. Indeed, it is rather odd to find the term "classical information" floating around in technical books and articles on quantum information theory when most, even if not all, physicists believe that all physical processes in the real world are quantum mechanical, with classical physics a good approximation in appropriate circumstances, but hardly part of our fundamental understanding of nature. A good way to see how "classical" information can arise in quantum mechanics is to note that one consequence of the Truncation theorem as discussed in Sec. V B is the fact that if a particular type of information about $\mathcal{S}_a$ associated with an orthonormal basis $\{|v_j\rangle\}$ is perfectly present in $\mathcal{S}_b$ it is the *only* type of information about $\mathcal{S}_a$ which can be present in a third system $\mathcal{S}_c$ in the sense that any other species of information is parasitic upon, or controlled by, or compatible with the $\{|v_j\rangle\}$ type. Whenever *only one* type of information needs to be considered all the rules of classical information theory apply to it; conversely, "classical information" in the quantum context refers to the single dominant type of *quantum* information available in a situation of strong decoherence. Typically it is the presence of this type of information in the environment that means that other types can be ignored in systems which are not isolated from the environment. In particular, the measurements indicated in part (a) of Figs. 1 and 2 when instantiated in physical apparatus amplify a particular type of information, and the environment rapidly copies the "pointer positions," resulting in strong decoherence. To avoid the rather unwieldy task of trying to describe this amplification process and interaction with the environment in correct quantum mechanical terms, which is certainly possible in principle, it is often preferable (as noted earlier) to employ a simple quantum circuit in which the decoherence is "built in": the a qubit in Fig. 1(b), and the a and c in qubits in Fig. 2(b), are at later times good copies of the $\mathcal{Z}$ information preceding the final control gates, and since no further use is made of them, they may be regarded as carrying this type of information off into the environment.

V. THEOREMS

In this section we state and prove results used in the preceding sections, plus some additional ones that are closely related. The treatment builds upon ideas and terminology from [8], repeated here to the extent needed to make the exposition self-contained. Note in particular that $\mathcal{H}_a$ is the Hilbert space of system $\mathcal{S}_a$, $\mathcal{H}_{ab} = \mathcal{H}_a \otimes \mathcal{H}_b$ that of $\mathcal{S}_{ab}$, the systems $\mathcal{S}_a$ and $\mathcal{S}_b$ regarded as a single system, ρ_{ab} is a density operator on $\mathcal{H}_{ab}$, often traced down from that of a larger system, d_a the dimension of $\mathcal{H}_a$, and so forth. All Hilbert spaces are assumed to be of finite dimension in order to avoid technical complications.

A. Presence

Theorem (Presence). *Let $\mathcal{S}_a$ and $\mathcal{S}_b$ be two quantum systems with Hilbert spaces $\mathcal{H}_a$ and $\mathcal{H}_b$, and $\mathcal{V} = \{V_j\}$ and $\mathcal{W} = \{W_k\}$ two strongly incompatible projective decompositions of the identity I_a. If both the $\mathcal{V}$ and the $\mathcal{W}$ information is perfectly present in $\mathcal{S}_b$ for a density operator ρ on $\mathcal{H}_{ab}$ (possibly a pure state), then* all *types of information about $\mathcal{S}_a$ are perfectly present in $\mathcal{S}_b$.*

The terms are to be understood as follows. The density operator ρ on $\mathcal{H}_{ab}$ or pure state $|\Psi\rangle \in \mathcal{H}_{ab}$ will be called a *pre-probability* using the terminology of Ch. 9 of [11], because it can be used to generate probabilities once a quantum sample space—an orthonormal basis of $\mathcal{H}_{ab}$ or a decomposition of the identity I_{ab}—has been specified, following the usual rule that the probability associated with a projector P is

$$\Pr(P) = \langle P \rangle = \mathrm{Tr}(P\rho) = \langle \Psi | P | \Psi \rangle, \qquad (5)$$

with $\rho = |\Psi\rangle\langle\Psi|$ for a pure state. For example, ρ as a density *operator* is represented by different *matrices* if different orthonormal bases are chosen. The diagonal elements

of one of these matrices form a probability distribution associated with the corresponding basis (or type of information), whereas the single density operator giving rise to the different distributions is the pre-probability.

The $\mathcal{V}$ type of information is *perfectly present* in $\mathcal{S}_b$ for a given pre-probability (Sec. III C of [8]) when the unnormalized conditional density operators

$$\rho_{bj} = \text{Tr}_a \left(V_j \rho \right) \tag{6}$$

on $\mathcal{H}_b$ are mutually orthogonal, i.e.,

$$\rho_{bj} \, \rho_{bk} = 0 \text{ for } j \neq k. \tag{7}$$

In the language of measurements, if one thinks of carrying out a projective measurement on $\mathcal{H}_a$ corresponding to $\{V_j\}$, then there is a corresponding decomposition $\{T_k\}$ of I_b such that the measurement outcomes are in one-to-one correspondence. An analogous definition applies to $\mathcal{W}$ information. The conclusion of the theorem, that all species of information about $\mathcal{S}_a$ are perfectly present in $\mathcal{S}_b$, conveniently abbreviated to "all information about $\mathcal{S}_a$ is in $\mathcal{S}_b$," means that for any decomposition of the identity, in particular for any orthonormal basis of $\mathcal{H}_a$, that kind of information is perfectly present in $\mathcal{S}_b$ in the sense just discussed. When the pre-probability is a pure state $|\Psi\rangle$, this implies it is maximally entangled, i.e., $\rho_a = \text{Tr}_b(|\Psi\rangle\langle\Psi|)$ is proportional to the identity I_a. For a general ρ a similar but more complicated result obtains—see theorem 3(ii) in [8]—and once again $\rho_a = \text{Tr}_b(\rho)$ is proportional to I_a.

The decompositions $\mathcal{V}$ and $\mathcal{W}$ are said to be *strongly incompatible* (Sec. IV of [8]) when the only projector P that commutes with *every* V_j and *every* W_k is either $P = 0$ or $P = I_a$. While concise, this definition is not very intuitive. In the case of orthonormal bases $\mathcal{V} = \{|v_j\rangle\}$ and $\mathcal{W} = \{|w_k\rangle\}$ one can use a somewhat simpler definition. Construct a graph containing $2d_a$ nodes, one for each $|v_j\rangle$ and one for each $|w_k\rangle$. Whenever the inner product $\langle v_j|w_k\rangle$ is nonzero, draw an edge between the corresponding nodes. Then $\mathcal{V}$ and $\mathcal{W}$ (i.e., the corresponding collections of projectors) are strongly incompatible if and only if this graph is *connected*. The proof is at the end of App. A. In the case of two mutually unbiased bases, as in Sec. IV, every $\{|v_j\rangle\}$ is connected to every $\{|w_k\rangle\}$ node, so connectivity of the graph is obvious. It is equally obvious for two bases in which $\langle v_j|w_k\rangle$ is never zero. However, strong incompatibility can still hold if some of the $\langle v_j|w_k\rangle$ are zero, provided the graph remains connected.

The proof of the Presence theorem, extending a weaker theorem in [8], is in App. A

B. Truncation, Exclusion, No Splitting, Somewhere

A series of useful "all-or-nothing" results about information in three systems $\mathcal{S}_a$, $\mathcal{S}_b$, and $\mathcal{S}_c$ begins with:

Theorem (Truncation). *Let $\mathcal{S}_a$, $\mathcal{S}_b$ and $\mathcal{S}_c$ be three quantum systems, and suppose that for some decomposition $\mathcal{V} = \{V_j\}$ of I_a and for some density operator ρ on $\mathcal{H}_{abc}$ all the $\mathcal{V}$ information about $\mathcal{S}_a$ is present in $\mathcal{S}_b$. Then any other type of information $\mathcal{W} = \{W_k\}$ about $\mathcal{S}_a$ will be "truncated" (or "censored") in the sense that*

$$\rho_{ac} = \sum_j V_j \rho_{ac} V_j, \tag{8}$$

that is, ρ_{ac}, the partial trace of ρ over $\mathcal{H}_b$, commutes with all the V_j. (Note that V_j is here understood as $V_j \otimes I_c$ on $\mathcal{H}_{ac}$.) Equivalently, all correlations between $\mathcal{S}_a$ and the third system $\mathcal{S}_c$ satisfy

$$\langle AC \rangle = \langle \bar{A} C \rangle \tag{9}$$

for any operators A and C on $\mathcal{H}_a$ ad $\mathcal{H}_c$, respectively (one could write $A \otimes C$ in place of AC), with

$$\bar{A} = \sum_j V_j A V_j. \tag{10}$$

the truncated version of the operator A, and $\langle \ \rangle$ the average taken with respect to ρ, as in (5).

This theorem is closely related to, but not the same as, theorem 6(i) in [8], and its proof is in App. B. Since any operator A can be written as $A = \sum_{jk} V_j A V_k$, its truncated version $\bar{A}$ is obtained by throwing away the off-diagonal blocks. To understand the implications of the theorem it helps to consider the case in which $\mathcal{V} = \{|v_j\rangle\}$ is an orthonormal basis of $\mathcal{H}_a$, so that

$$\bar{A} = \sum_j \langle v_j|A|v_j\rangle|v_j\rangle\langle v_j| \tag{11}$$

is diagonal in this basis, meaning that all correlations between A and C can be computed from the correlations $\langle V_j C \rangle$, that is from $\mathcal{V}$ information about $\mathcal{S}_a$ in $\mathcal{S}_c$. Equivalently,

$$\rho_{ac} = \sum_j |v_j\rangle\langle v_j| \otimes \Gamma_j, \tag{12}$$

where the Γ_j are operators on $\mathcal{H}_c$. All other information about $\mathcal{S}_a$ in $\mathcal{S}_c$, of whatever kind, is then "parasitic upon," "truncated by," or "censored relative to" the $\mathcal{V}$ information. When the V_j projectors have rank greater than 1 the truncation or censorship is less extreme, but it remains true that the only sort of information about $\mathcal{S}_a$ allowed in $\mathcal{S}_c$ is represented by $\bar{A}$-type operators which are compatible with $\mathcal{V}$ in the sense of commuting with every V_j or, equivalently, ρ_{ac} commutes with every V_j.

The situation is particularly clear if there is another basis $\mathcal{W} = \{|w_k\rangle\}$ which is mutually unbiased with respect to $\mathcal{V} = \{|v_j\rangle\}$, i.e., $|\langle w_k|v_j\rangle|^2 = 1/d_a$ for all j and k. In that case the truncated projectors $\bar{W}_k$ are not only diagonal in the $\mathcal{V}$ representation, but are all equal to I_a/d_a independent of k, proving the next theorem (which is the same as theorem 7(ii) in [8]):

Theorem (Exclusion). *Let $\mathcal{S}_a$, $\mathcal{S}_b$ and $\mathcal{S}_c$ be three quantum systems, and $\mathcal{V} = \{|v_j\rangle\}$ and $\mathcal{W} = \{|w_k\rangle\}$ two mutually unbiased orthonormal bases of $\mathcal{H}_a$. Then if the $\mathcal{V}$ information about $\mathcal{S}_a$ is perfectly present in $\mathcal{S}_b$, the $\mathcal{W}$ information about $\mathcal{S}_a$ is perfectly absent from $\mathcal{S}_c$.*

The *perfect absence* of some type of information can be defined using reduced density operators, as in (6), but now they are required to be the same up to a multiplicative constant. That is, the $\mathcal{W}$ or $\{W_k\}$ information about $\mathcal{S}_a$ is perfectly absent from $\mathcal{S}_c$ if and only if for every k

$$\mathrm{Tr}_a\left(W_k \rho_{ac}\right) = p_k \rho_c, \tag{13}$$

where $\rho_c = \mathrm{Tr}_a(\rho_{ac})$, ρ_{ac} is the density operator (pre-probability) on $\mathcal{H}_{ac}$, and the p_k are nonnegative numbers. One can think of this in terms of measurements as saying that when a projective measurement corresponding to $\{W_k\}$ is made on $\mathcal{S}_a$, the probability of any measurement on $\mathcal{S}_c$ conditioned on the outcome k will be independent of k.

Below we will need the notion of the perfect absence of *all* types of information about $\mathcal{S}_a$ from $\mathcal{S}_c$, conveniently abbreviated to "no information about $\mathcal{S}_a$ is in $\mathcal{S}_c$." This is equivalent to $\rho_{ac} = \rho_a \otimes \rho_c$, or to $|\Psi\rangle = |\alpha\rangle \otimes |\gamma\rangle$ for a pure state, theorem 1(iii) of [8]. As the relationship is obviously symmetrical, one can also say that $\mathcal{S}_a$ and $\mathcal{S}_c$ are *uncorrelated*.

An important corollary of the Exclusion theorem is:

Theorem (No Splitting). *Let $\mathcal{S}_a$, $\mathcal{S}_b$ and $\mathcal{S}_c$ be three quantum systems. If all types of information about $\mathcal{S}_a$ are perfectly present in $\mathcal{S}_b$, then all types will be perfectly absent from $\mathcal{S}_c$. That is, if all information about $\mathcal{S}_a$ is in $\mathcal{S}_b$, no information about $\mathcal{S}_a$ is in $\mathcal{S}_c$.*

This is theorem 8(i) in [8]. It follows at once from the Exclusion theorem, because to show the absence of some species of information about $\mathcal{S}_a$ in $\mathcal{S}_c$ it suffices to consider orthonormal bases, and for each of these we know that there is at least one mutually unbiased basis for which all the corresponding information is, by hypothesis, perfectly present in $\mathcal{S}_b$. The No Splitting theorem has lots of applications. For example, in either one or two bit teleportation after the final corrections have been made, there is no information about the input state $|\psi\rangle$ remaining in the environment treated as a quantum system, and since copies of the classical bits x and y used to complete the protocol can remain in the environment, it is evident that they, as has often been observed, can contain no information about the input: their probabilities cannot depend upon $|\psi\rangle$. In the case of quantum codes the presence of the encoded information in some subset of the coding bits (which is what makes error correction possible) means its absence from the complementary subset of coding bits, and this can provide additional intuition about the coding process [8].

Is there a converse to the Exclusion theorem which says that if the $\mathcal{W}$ information about $\mathcal{S}_a$ is perfectly absent from $\mathcal{S}_c$, then that associated with any mutually unbiased basis $\mathcal{V}$ must be present in $\mathcal{S}_b$? No, not even if one knows that all information about $\mathcal{S}_a$ is present in the combined system of $\mathcal{S}_{bc}$; see the end of App. B for a counterexample. There is, on the other hand, a partial converse of the No Splitting theorem:

Theorem (Somewhere). *If for a pure state pre-probability $|\Psi\rangle$ on $\mathcal{H}_{abc}$ it is the case that all the information about $\mathcal{S}_a$ is in the combined system $\mathcal{S}_{bc}$, and none of it is in $\mathcal{S}_c$, then it is all in $\mathcal{S}_b$.*

The name "Somewhere" comes from the idea that if we know that an object is in one of two rooms and it is not in the second, it has to be in the first: it must be *somewhere*. However, information is very different from a lost child, as it can be present in *correlations* between two systems, while not being available in either system by itself. See, for example, the discussion of $\mathcal{X}$ information in one bit teleportation in Sec. III. Consequently, the Somewhere theorem is a decidedly quantum mechanical result. Also, it fails (in general) if the pure state is replaced by a density operator. The theorem itself is proved in [8] as theorem 8(ii), where one will also find an application to quantum codes.

C. Absence

Given the Presence theorem one might anticipate a similar Absence theorem. It comes in two versions:

Theorem (Absence). *Let $\mathcal{S}_a$ and $\mathcal{S}_b$ be two quantum systems.*

i) Simple version. If the pre-probability is a pure state $|\Psi\rangle$ on $\mathcal{H}_a \otimes \mathcal{H}_b$ and the information associated with a single orthonormal basis $\{|v_j\rangle\}$ of $\mathcal{H}_a$ is completely absent from $\mathcal{H}_b$, then $|\Psi\rangle$ is a product state of the form $|a\rangle \otimes |b\rangle$, so there is no information about $\mathcal{S}_a$ in $\mathcal{S}_b$ or vice versa; the two are uncorrelated.

ii) Complicated version. Let the pre-probability be a general density operator on $\mathcal{H}_a \otimes \mathcal{H}_b$, and let $\{V^{(m)}\}$ be a collection of decompositions of the identity I_a of $\mathcal{H}_a$,

$$I_a = \sum_j V_j^{(m)} \tag{14}$$

for each m, where the $V_j^{(m)}$ for different j are the projectors belonging to $V^{(m)}$. If the collection $\{V_j^{(m)}\}$ for all m and all j spans the space $\hat{\mathcal{H}}_a$ of operators on $\mathcal{H}_a$, and if each species $V^{(m)}$ of information about $\mathcal{S}_a$ is completely absent from $\mathcal{S}_b$, then

$$\rho = \rho_a \otimes \rho_b, \tag{15}$$

so there is no information about $\mathcal{S}_a$ in $\mathcal{S}_b$ or vice versa; the two are uncorrelated.

The proof of both versions is given in App. C. For version (ii) the conditions are definitely more complicated than for the Presence theorem: if $\mathcal{H}_a$ has dimension d_a, one needs to check not two but at least $d_a + 1$ orthonormal bases (see end of App. C) in order to be sure that all information is absent. For instance, if $\mathcal{S}_a$ is a qubit, $d_a =$

2, it suffices to check that the $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ types of information are absent, but two out of the three is not enough, as shown in the example at the end of App. B.

D. No Cloning

One might suspect that the No-Splitting theorem is the same as, or at least closely related to, the well-known no cloning result [9]. However, the two seem to be different, since neither the conditions nor the consequences of no-cloning are expressed in terms of types of information as used here. The following theorem is the closest we have been able to come in finding a connection between the two.

Theorem (Generalized No Cloning). *Let M be an isometry from $\mathcal{H}_a$ to $\mathcal{H}_{bc}$, and $\{|\alpha_j\rangle\}$, with j lying in a finite index set J, a collection of normalized kets on $\mathcal{H}_a$ with the property that the pairs (j, k) for which the inner product $\langle\alpha_j|\alpha_k\rangle$ is nonzero when treated as edges produce a connected graph on the set J. Assume that for each $|\alpha_j\rangle$ its image under M is a product state*

$$M|\alpha_j\rangle = |\beta_j\rangle \otimes |\gamma_j\rangle, \tag{16}$$

where both $|\beta_j\rangle$ and $|\gamma_j\rangle$ are normalized, and that

$$|\langle\alpha_j|\alpha_k\rangle| = |\langle\beta_j|\beta_k\rangle| \tag{17}$$

whenever the left side is nonzero.

Under these conditions M restricted to the subspace $\mathcal{G}_a$ spanned by the $\{|\alpha_j\rangle\}$ is of the form

$$M|a\rangle = (U|a\rangle) \otimes |\gamma_1\rangle, \tag{18}$$

where U is a unitary map of $\mathcal{G}_a$ onto the subspace $\mathcal{G}_b$ of $\mathcal{H}_b$ spanned by the $\{|\beta_j\rangle\}$, and $|\gamma_1\rangle$ is a fixed ket in $\mathcal{H}_c$.

The proof is in App. D. The connection with no-cloning, not obvious given the somewhat abstract statement of the theorem, is the following. Suppose j takes on just two values 1 and 2, the states $|\alpha_1\rangle$ and $|\alpha_2\rangle$ are linearly independent, and $\langle\alpha_1|\alpha_2\rangle \neq 0$. Imagine these are two states to be cloned, and the isometry M (which can be replaced with a unitary acting on the tensor product of $\mathcal{H}_a$ and an additional space $\mathcal{H}_s$ initially in a state $|s_0\rangle$) is supposed to carry out the cloning process. If $|\beta_1\rangle$ and $|\beta_2\rangle$ are good copies up to some unitary transformation of $\mathcal{H}_b$, their inner product must equal $\langle\alpha_1|\alpha_2\rangle$ apart from an unimportant phase. As the conditions of the theorem are fulfilled—the graph consists of two nodes joined with the edge (1,2)—it follows that M is not only unable to produce additional copies in $\mathcal{G}_c$, but in fact there is no information at all in $\mathcal{G}_c$ which would allow distinguishing the states $|\alpha_1\rangle$ and $|\alpha_2\rangle$. Thus at least for the subspace $\mathcal{G}_a$ (which could be all of $\mathcal{H}_a$ if the span of $\{|\alpha_j\rangle\}$ is large enough) one arrives at the same conclusion as with the No Splitting theorem, but using somewhat different hypotheses.

VI. CONCLUSION

Identifying types or species of quantum information and noting when they are compatible (i.e., the projectors commute) or incompatible looks like a promising approach to the foundations of quantum information for the following reasons. First, it allows a more intuitive, as well as a fully consistent, approach to quantum probabilities at the microscopic level, in contrast to the usual textbook approach, with its preparations, measurements, and "great smoky dragon" [37], long known to provide an awkward and difficult (and internally inconsistent [38, 39]) way of thinking about the quantum world, however effective it may be as a calculational tool for the final outcomes of measurements. Second, the ideas of classical information theory [3] are directly applicable to quantum systems as long as one restricts oneself to a *single* type of quantum information, or to two or more *compatible* types (which can then be combined to form a single type), because there is a properly defined sample space on which probabilities of *quantum* events and processes, and their correlations, satisfy the standard rules of probability theory, which are fundamental to the structure of information theory as developed by Shannon and his successors. Note that it is not necessary to restrict oneself to macroscopic systems or asymptotically large N (number of transmissions, or whatever) limits.

Third, the existence of *different* incompatible species of quantum information is at the heart of the objections raised in [4] to extending Shannon's theory to the quantum domain. Recognizing the role of quantum incompatibility and using different information types gets around these problems and allows a fully consistent formulation of the microscopic statistical correlations needed to properly begin the "quantization" of classical information theory. Fourth, one of the principal ways quantum information goes beyond its classical counterpart is in its discussion of how incompatible types of information relate to, or so-to-speak constrain, each other for a given setup, or quantum circuit, or entangled state. The Presence, Truncation, and Absence theorems and their various corollaries in Sec. V clearly do not belong to the domain of classical information, since their very formulation requires reference to noncommuting operators, the hallmark of "quantum" effects.

The approach presented here provides, we believe, new intuitive insight into the processes of teleportation and decoherence, and into how "classical" information can be consistently described as a quantum phenomenon. It is obviously incomplete in two respects. First, the theorems of Sec. V are of the "all or nothing" variety: they apply to extreme situations in which information is either completely present or completely absent. Obviously it would be valuable to have quantitative extensions of these theorems, presumably in the form of inequalities, that apply to situations where information of different kinds is partially present or absent. Finding suitable information measures and proving appropriate bounds looks like a

challenging problem, but one that needs to be addressed given that one is often interested in situations where there is noise, so different types of quantum information will be degraded in different ways. There are, of course, many inequalities involving quantum information in the published literature, and some of them, such as those of Hall [40, 41], look as if they can be reformulated to apply to different species of information as discussed here.

Second, the examples and theorems given in this paper (and their extensions beyond "all or nothing" noted above) need to be generalized to cases in which microscopic quantum properties are considered at a large number of successive times, as in the case of "quantum jumps" [42–44]. For this purpose it is likely that the full machinery of quantum histories [11] will be needed in order to provide consistent probabilistic descriptions without having to invoke the awkward concepts of macroscopic "preparation" and "measurement," which are obviously not a fundamental part of microscopic quantum mechanics.

Acknowledgments

APPENDIX A: PRESENCE THEOREM

The Presence theorem of Sec. V is an extension of theorem 4 of [8], where it was shown to hold when ρ is a pure state on $\mathcal{H}_{ab}$. Our task here will be to remove that restriction and prove it for a general density operator $\rho = \rho_{ab}$, where subscripts have been added to avoid any confusion in the following argument. To that end first "purify" ρ_{ab} by introducing an auxiliary system $\mathcal{H}_c$ and a ket

$$|\Phi\rangle = \sum_q \sqrt{p_q}|\psi_q\rangle \otimes |c_q\rangle \tag{A1}$$

on $\mathcal{H}_{abe}$, with $\{|c_q\rangle\}$ an orthonormal basis of $\mathcal{H}_c$, and coefficients $\{|\psi_q\rangle\}$ chosen so that

$$\rho_{ab} = \mathrm{Tr}_c(|\Phi\rangle\langle\Phi|) = \sum_q p_q|\psi_q\rangle\langle\psi_q|. \tag{A2}$$

Lemma *Suppose that the $\mathcal{V} = \{V_j\}$ information about $\mathcal{S}_a$ is perfectly present in $\mathcal{S}_b$ for ρ_{ab}. Then that is also true for every $|\psi_q\rangle$ in (A1) for which $p_q > 0$.*

To show this, insert (A2) in place of ρ on the right side of (6), so each ρ_{bj} is a sum over q of positive operators ρ_{bqj}. Then in order that (7) hold it is necessary (and obviously sufficient) that $\rho_{bqj}\rho_{bq'k} = 0$, for all q and q', whenever $j \neq k$. Setting $q' = q$ gives the desired result.

Now apply the lemma to both the $\mathcal{V}$ and the strongly incompatible $\mathcal{W}$ type of information. Since for one of the pure states $|\psi_q\rangle$ on $\mathcal{H}_{ab}$ both types of information about $\mathcal{S}_a$ are in $\mathcal{S}_b$, theorem 4 of [8] tells us that for this pure state *all* information about $\mathcal{S}_a$ is present in $\mathcal{S}_b$. This implies (and is implied by, see theorem 3 (i) of [8]) that $|\psi_q\rangle$ is maximally entangled, or

$$\mathrm{Tr}_b(|\psi_q\rangle\langle\psi_q|) = I_a/d_a. \tag{A3}$$

Next argue that the $\mathcal{C} = \{|c_q\rangle\langle c_q|\}$ type of information about $\mathcal{S}_c$ is absent from $\mathcal{S}_a$ by showing that the conditional density operators

$$\mathrm{Tr}_c(\rho_{ac}|c_q\rangle\langle c_q|) = p_q(I_a/d_a) \tag{A4}$$

depend on $|c_q\rangle$ only through the numerical factor p_q; see (13) with systems appropriately renamed. Indeed, (A4) follows from (A1) when one replaces ρ_{ac} with $\mathrm{Tr}_b(|\Phi\rangle\langle\Phi|)$ and interchanges the order of partial traces. The same argument applies if we use any other choice of orthonormal basis of $\mathcal{H}_c$ for the expansion (A1): as is well known, changing that basis does not alter the density operator $\rho = \rho_{ab}$ we began with, but simply expresses it in terms of a different ensemble. Consequently, all types of information about $\mathcal{S}_c$ are absent from $\mathcal{S}_a$, which is to say $\rho_{ac} = \rho_a \otimes \rho_c$, theorem 1 (iii) of [8], and thus all types of information about $\mathcal{S}_a$ are also absent from $\mathcal{S}_c$.

One more step is needed. The presence of $\mathcal{V}$ information about $\mathcal{S}_a$ in $\mathcal{S}_b$ means it is also present in $\mathcal{S}_{bc}$. (This is intuitively obvious, but can also be shown formally from the definition in (35) of [8], where one simply replaces B^k with $B^k \otimes I_c$. Or one can use the definition in (7) of the present paper, along with the fact that if P and Q are positive operators on $\mathcal{H}_e \otimes \mathcal{H}_f$ with $PQ \neq 0$, then $P_eQ_e \neq 0$, where P_e and Q_e are partial traces over $\mathcal{H}_f$—use the spectral representations and take traces.) Of course the same is true of the $\mathcal{W}$ information. Hence, applying theorem 4 of [8] to the bipartite system consisting of $\mathcal{S}_a$ on the one hand and $\mathcal{S}_{bc}$ on the other, with $|\Phi\rangle$ an entangled ket on $\mathcal{H}_a \otimes \mathcal{H}_{bc}$, we conclude, from the presence of two strongly incompatible types of information, that all the information about $\mathcal{S}_a$ is present in the combined system $\mathcal{S}_{bc}$. From this and the absence of all information about $\mathcal{S}_a$ from $\mathcal{S}_c$ demonstrated earlier, it follows from theorem 8 (iii) of [8], using $|\Phi\rangle$ as a pure state pre-probability on $\mathcal{H}_{abc}$, that all information about $\mathcal{S}_a$ is in $\mathcal{S}_b$. This completes the proof of the theorem.

Next we will show that two orthonormal bases $\mathcal{V} = \{|v_j\rangle\}$ and $\mathcal{W} = \{|w_k\rangle\}$ are strongly incompatible if and only if the graph described in Sec. V A is connected. First, assume $\mathcal{V}$ and $\mathcal{W}$ are *not* strongly incompatible, so there is a projector P that commutes with every $|v_j\rangle\langle v_j|$ and every $|w_k\rangle\langle w_k|$, and is neither 0 nor I, so projects onto a proper subspace $\mathcal{P}$ of the Hilbert space. Then some subset of the $\{|v_j\rangle\}$ are inside $\mathcal{P}$ and the rest are in its orthogonal complement $\mathcal{P}^\perp$, for otherwise $\mathcal{P}$ would not commute with them, and the same is true of the $\{|w_k\rangle\}$. Evidently there can be no nonzero $\langle v_j|w_k\rangle$ for a

$|v_j\rangle$ (or $|w_k\rangle$) in $\mathcal{P}$ and a $|w_k\rangle$ (or $|v_j\rangle$) in $\mathcal{P}^\perp$. Consequently the graph cannot be connected, as it has at least two components, one for $\mathcal{P}$ and one for $\mathcal{P}^\perp$ vertices.

For the converse, assume the graph is not connected, and renumber the vertices so that the $\{|v_j\rangle\}$ with $1 \le j \le m < d_a$ are all the v vertices in one connected component $\mathcal{C}$ of the graph, with the others in its complement. The projector

$$P = \sum_{j=1}^{m} |v_j\rangle\langle v_j| \qquad (A5)$$

obviously commutes with all the $\{|v_j\rangle\}$, and in addition commutes with every $|w_k\rangle\langle w_k|$ when $|w_k\rangle$ is *not* in $\mathcal{C}$, as otherwise there would be an edge from some $|v_j\rangle$ with $j \le m$ to this $|w_k\rangle$. Now apply the same argument with $I - P$ in place of P to show that $I - P$ commutes with every $|w_k\rangle\langle w_k|$ when $|w_k\rangle$ *is* in $\mathcal{C}$. Since P and $I - P$ commute with the same things, we have shown that P commutes with all the $|w_k\rangle\langle w_k|$ as well as the $|v_j\rangle\langle v_j|$. As P is neither 0 nor I, $\mathcal{V}$ and $\mathcal{W}$ are not strongly incompatible.

APPENDIX B: TRUNCATION THEOREM AND RELATED

If the $\mathcal{V} = \{V_j\}$ information about $\mathcal{S}_a$ is perfectly present in $\mathcal{S}_b$, there is, as noted following (7), a decomposition $\{T_k\}$ of I_b such that

$$\langle V_j T_k \rangle = \delta_{jk}\langle V_j \rangle. \qquad (B1)$$

(Note that $V_j T_k$ is $V_j \otimes T_k \otimes I_c$.) The equivalence is shown in [8], where in fact (B1) is the primary definition. If the pre-probability defining $\langle \rangle$ is a pure state $|\Psi\rangle \in \mathcal{H}_{abc}$, (B1) implies that

$$V_j T_k |\Psi\rangle = 0 \text{ for } j \ne k. \qquad (B2)$$

As a consequence, and using the fact that $I_a = \sum_j V_j$, $I_b = \sum_k T_k$, we have

$$V_j |\Psi\rangle = V_j T_j |\Psi\rangle = T_j |\Psi\rangle, \qquad (B3)$$

and

$$|\Psi\rangle = \sum_j V_j T_j |\Psi\rangle = \sum_j V_j |\Psi\rangle. \qquad (B4)$$

Therefore it follows that

$$\langle AC \rangle = \langle\Psi|\Big[\sum_{jk} V_j T_j A C V_k T_k\Big]|\Psi\rangle = \langle\Psi|\Big[\sum_j V_j A V_j C\Big]|\Psi\rangle, \qquad (B5)$$

which is (9) when the pre-probability is a pure state. In (B5) we have used the fact that T_j commutes with AC, $T_j T_k = \delta_{jk} T_j$, and (B4).

To extend the argument to a general density operator ρ on $\mathcal{H}_{abc}$, introduce a fictitious system $\mathcal{H}_r$, purify ρ to a ket $|\Psi\rangle \in \mathcal{H}_{abcr}$, and apply (B5) to the three part system

consisting of $\mathcal{S}_a$, $\mathcal{S}_b$, and in place of $\mathcal{S}_c$ the combined system $\mathcal{S}_{cr}$. The significance of V_j, T_j, and A is the same as before, while C can be replaced by any operator on $\mathcal{H}_{cr}$. If in particular we use $C \otimes I_r$, the result is (9). The equivalence of (8) and (9) is a straightforward exercise when one notes that $\langle AC \rangle = \text{Tr}(AC\rho_{ac})$, and that (9) holds for all A and C (operating on $\mathcal{H}_a$ and $\mathcal{H}_c$). This completes the proof.

The following example shows that the Exclusion theorem does not possess a simple converse of the type mentioned in Sec. V B. The entangled state

$$2|\Psi\rangle = |000\rangle + |011\rangle + |100\rangle - |111\rangle \qquad (B6)$$

on $\mathcal{H}_{abc}$, with qubits in the order $|abc\rangle$, has the property that all information about $\mathcal{S}_a$ is present in the combined system $\mathcal{S}_{bc}$, the $\mathcal{X}$ information about $\mathcal{S}_a$ is perfectly present in both $\mathcal{S}_b$ and in $\mathcal{S}_c$, whereas both the $\mathcal{Y}$ (basis $\{(|0\rangle \pm i|1\rangle)/\sqrt{2}\}$) and the mutually unbiased $\mathcal{Z}$ information about $\mathcal{S}_a$ are perfectly absent from both $\mathcal{S}_b$ and from $\mathcal{S}_c$. Perhaps the easiest way to check this is to expand $|\Psi\rangle$ in the $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ bases of $\mathcal{H}_a$ in turn, and look at the coefficients in $\mathcal{H}_{bc}$. That all the information about $\mathcal{S}_a$ is in $\mathcal{S}_{bc}$ follows from the observation that any one of these expansions (and therefore all three) is in Schmidt form with Schmidt coefficients of equal magnitude, so theorem 3(i) of [8] applies. Note that we have an example in which if $\mathcal{S}_a$ and $\mathcal{S}_b$ are considered two parts of a bipartite system with pre-probability given by the density operator $\text{Tr}_c(|\Psi\rangle\langle\Psi|)$, one would be mistaken to suppose that the absence of the two mutually unbiased types of information $\mathcal{Y}$ and $\mathcal{Z}$ about $\mathcal{S}_a$ from $\mathcal{S}_b$ implied the complete absence of all information. This confirms the remarks at the end of Sec. V C.

APPENDIX C: ABSENCE THEOREMS

For part (i), write $|\Psi\rangle$ in the form

$$|\Psi\rangle = \sum_j |v_j\rangle \otimes |\beta_j\rangle, \qquad (C1)$$

where the $|\beta_j\rangle$ are expansion coefficients. The fact that the $\{|v_j\rangle\}$ or $\{|v_j\rangle\langle v_j|\}$ information about $\mathcal{S}_a$ is absent from $\mathcal{S}_b$ means that the $|\beta_j\rangle$ are all proportional to one another, thus multiples of $|\beta_1\rangle$, assuming it is nonzero. Inserting $|\beta_j\rangle = c_j|\beta_1\rangle$ in (C1) shows that $|\Psi\rangle = |a\rangle \otimes |\beta_1\rangle$ is a product state, so no information about $\mathcal{S}_a$ is in $\mathcal{S}_b$ or vice versa.

To prove (ii) we employ an orthonormal basis $\{Q_r\}$, $0 \le r \le d_a^2 - 1$ of the space $\hat{\mathcal{H}}_a$ of linear operators on $\mathcal{H}_a$, in the sense that

$$\langle Q_r, Q_s \rangle := (1/d_a)\text{Tr}_a(Q_r^\dagger Q_s) = \delta_{rs}, \qquad (C2)$$

with $Q_0 = I_a$, and thus $\text{Tr}_a(Q_r) = 0$ for $r > 0$. Expand ρ as

$$\rho = (1/d_a)\sum_r Q_r \otimes B_r, \qquad (C3)$$

where, see (C2), the expansion coefficients are given by

$$B_r = \mathrm{Tr}_a(Q_r \rho), \qquad \text{(C4)}$$

with

$$B_0 = \mathrm{Tr}_a(\rho) = \rho_b. \qquad \text{(C5)}$$

The absence from $\mathcal{S}_b$ of each species $V^{(m)} = \{V_j^{(m)}\}$ of information about $\mathcal{S}_a$ means that, see (13),

$$\mathrm{Tr}_a(V_j^{(m)}\rho) = p_{jm}\rho_b = p_{jm}B_0, \qquad \text{(C6)}$$

where the p_{jm} are nonnegative constants. By hypothesis, the collection $\{V_j^{(m)}\}$ for all m and all j spans the operator space $\hat{\mathcal{H}}_a$, so any Q_r can be written as a sum

$$Q_r = \sum_{jm} c_{rjm} V_j^{(m)}, \qquad \text{(C7)}$$

with suitable coefficients c_{rjm}. (These may not be unique, but that does not matter.) Insert (C7) in (C4) and use (C6) to conclude that every B_r is a multiple of B_0, and therefore $\rho = \rho_a \otimes B_0$ is a product.

The need for at least $d_a + 1$ orthonormal bases of $\mathcal{H}_a$ in order to check that all information about $\mathcal{S}_a$ is absent from $\mathcal{S}_b$ can be seen in the following way. Each basis of $\mathcal{H}_a$ gives rise to d_a orthogonal, and hence linearly independent, operators in the d_a^2-dimensional space $\hat{\mathcal{H}}_a$. But

these d_a projectors sum to the identity I for each such basis, and therefore ν such bases will give rise to at most $\nu(d_a - 1) + 1$ linearly independent operators, which is d_a^2 when $\nu = d_a + 1$.

APPENDIX D: GENERALIZED NO CLONING

From (16) one sees that for every j and k in J,

$$|\langle\gamma_j|\gamma_k\rangle| \cdot |\langle\beta_j|\beta_k\rangle| = |\langle\alpha_j|M^\dagger M|\alpha_k\rangle| = |\langle\alpha_j|\alpha_k\rangle|, \quad \text{(D1)}$$

and therefore, in view of (17),

$$|\langle\gamma_j|\gamma_k\rangle| = 1 \qquad \text{(D2)}$$

whenever $j \neq k$ and $|\langle\alpha_j|\alpha_k\rangle| \neq 0$. The pairs (j, k) for which (D2) holds form, by hypothesis, a connected graph on J, which means that the normalized kets $|\gamma_j\rangle$ are identical apart from phase factors, so each is a multiple of just one of them; let us say $|\gamma_j\rangle = e^{i\phi_j}|\gamma_1\rangle$. Replace the right side of (16) with $|\beta_j'\rangle \otimes |\gamma_1\rangle$, where $|\beta_j'\rangle = e^{i\phi_j}|\beta_j\rangle$, and define the linear operator U on $\mathcal{G}_a$ so that

$$U|\alpha_j\rangle = |\beta_j'\rangle. \qquad \text{(D3)}$$

Since $|\langle\beta_j'|\beta_k'\rangle| = |\langle\beta_j|\beta_k\rangle| = |\langle\alpha_j|\alpha_k\rangle|$, U is unitary as a map from $\mathcal{G}_a$ to $\mathcal{G}_b$.

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[2] M. Keyl, Phys. Rep. **369**, 431 (2002), quant-ph/0202122.
[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
[4] Č. Brukner and A. Zeilinger, Phys. Rev. A **63**, 022113 (2001).
[5] C. G. Timpson, Stud. Hist. Phil. Mod. Phys. **34**, 441 (2003).
[6] I. Devetak, A. Harrow, and A. Winter (2005), quant-ph/0512015.
[7] I. Devetak and J. Yard (2006), quant-ph/0612050.
[8] R. B. Griffiths, Phys. Rev. A **71**, 042337 (2005), quant-ph/0409106.
[9] W. K. Wootters and W. H. Zurek, Nature **299**, 802 (1982).
[10] For a more extended discussion of this topic, see Sec. 4.6 and various examples in [11], or Sec. III of [45].
[11] R. B. Griffiths, *Consistent Quantum Theory* (Cambridge University Press, Cambridge, U.K., 2002), http://quantum.phys.cmu.edu/CQT/.
[12] G. Birkhoff and J. von Neumann, Annals of Math. **37**, 823 (1936).
[13] R. B. Griffiths, J. Stat. Phys. **36**, 219 (1984).
[14] R. Omnès, J. Stat. Phys. **53**, 893 (1988).
[15] M. Gell-Mann and J. B. Hartle, in *Complexity, Entropy and the Physics of Information.*, edited by W. H. Zurek (Addison-Wesley, 1990), pp. 425–458.
[16] M. Gell-Mann and J. B. Hartle, Phys. Rev. D **47**, 3345 (1993).
[17] R. Omnès, *The Interpretation of Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1994).
[18] R. B. Griffiths, Phys. Rev. A **54**, 2759 (1996).
[19] R. B. Griffiths, Phys. Rev. A **57**, 1604 (1998), quant-ph/9708028.
[20] R. Omnès, *Understanding Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1999).
[21] A number of paradoxes are discussed in Chs. 19 to 25 of [11].
[22] See the discussion of measurements in [11] and [45].
[23] See p. 163 Rule 4 in [17], Sec. III of [18], Sec. II of [19], Sec. 16.1 of [11].
[24] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).
[25] X. Zhou, D. W. Leung, and I. L. Chuang, Phys. Rev. A **62**, 052316 (2000).
[26] G. Brassard, S. L. Braunstein, and R. Cleve, Physica D **120**, 43 (1998).
[27] R. B. Griffiths and C.-S. Niu, Phys. Rev. Lett. **76**, 3228 (1996).
[28] K. Życzkowski and I. Bengtsson, Open Syst. Inf. Dyn. **11**, 3 (2004), quant-ph/0401119.

[29] R. B. Griffiths, S. Wu, L. Yu, and S. M. Cohen, Phys. Rev. A **73**, 052309 (2006), quant-ph/0507215.

[30] C. J. Isham, J. Math. Phys. **35**, 2157 (1994).

[31] R. B. Griffiths, Phys. Rev. A **66**, 012311 (2002), quant-ph/0203058.

[32] W. H. Zurek, Rev. Mod. Phys. **75**, 715 (2003).

[33] H. Ollivier, D. Poulin, and W. H. Zurek, Phys. Rev. Lett. **93**, 220401 (2004), quant-ph/0307229.

[34] R. Blume-Kohout and W. H. Zurek, Found. Phys. **35**, 1857 (2005), quant-ph/0408147.

[35] H. Ollivier, D. Poulin, and W. H. Zurek, Phys. Rev. A **72**, 042113 (2005), quant-ph/0408125.

[36] The claim that decoherence by itself solves the measurement problem has, in our opinion, been overstated; see the clear critique in [46]. Instead, the usual treatments of decoherence suffer from the same internal inconsistency that constitutes the infamous "measurement" problem of quantum foundations, as discussed by Wigner [38] and Mittelstaedt [39] among others. Solving the measurement problem requires that one introduce probabilities in quantum theory in a consistent manner, as discussed in [11], and then decoherence can be consistently discussed as a quantum mechanical process: see [15, 16, 20, 47, 48] and Ch. 26 of [11].

[37] J. A. Wheeler, Am. J. Phys. **51**, 398 (1983).

[38] E. P. Wigner, Am. J. Phys. **31**, 6 (1963).

[39] P. Mittelstaedt, *The Interpretation of Quantum Mechanics and the Measurement Process* (Cambridge, Cambridge, U.K., 1998).

[40] M. J. W. Hall, Phys. Rev. Lett. **74**, 3307 (1995).

[41] M. J. W. Hall, Phys. Rev. A **55**, 100 (1997).

[42] M. B. Plenio and P. L. Knight, Rev. Mod. Phys. **70**, 101 (1998).

[43] H. M. Wiseman and G. E. Toombes, Phys. Rev. A **60**, 2474 (1999).

[44] A. Sudbery, Stud. Hist. Phil. Mod. Phys. **33**, 387 (2002).

[45] R. B. Griffiths (2006), quant-ph/0612065.

[46] S. L. Adler, Stud. Hist. Phil. Mod. Phys. **34**, 135 (2003).

[47] R. Omnès, Phys. Rev. A **56**, 3383 (1997).

[48] R. Omnès, Brazilian J. Phys. **35**, 207 (2005).