Nature and location of quantum information

Robert B. Griffiths*
Department of Physics, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213
(Received 28 March 2002; published 25 July 2002)

Quantum information is defined by applying the concepts of ordinary (Shannon) information theory to a quantum sample space consisting of a single framework or consistent family. A classical analogy for a spin-half particle and other arguments show that the infinite amount of information needed to specify a precise vector in its Hilbert space is not a measure of the information carried by a quantum entity with a d-dimensional Hilbert space; the latter is, instead, bounded by $\log_2 d$ bits (one bit per qubit). The two bits of information transmitted in dense coding are located not in one but in the correlation between two qubits, consistent with this bound. A quantum channel can be thought of as a *structure* or collection of frameworks, and the physical location of the information in the individual frameworks can be used to identify the location of the channel. Analysis of a quantum circuit used as a model of teleportation shows that the location of the channel depends upon which structure is employed; for ordinary teleportation it is not (contrary to Deutsch and Hayden) present in the two bits resulting from the Bell-basis measurement, but in correlations of these with a distant qubit. In neither teleportation nor dense coding does information travel backwards in time, nor is it transmitted by nonlocal (superluminal) influences. It is (tentatively) proposed that *all* aspects of quantum information can in principle be understood in terms of the (basically classical) behavior of information in a particular framework, along with the framework dependence of this information.

DOI: 10.1103/PhysRevA.66.012311 PACS number(s): 03.67.−a, 03.65.Ta

I. INTRODUCTION

Quantum information theory has attracted a great deal of attention, with the interest driven in no small part by potential applications to quantum cryptography and quantum computing. It represents an extension of classical information theory, as developed by Shannon and his successors, to a domain in which quantum effects are important and therefore quantum mechanics must be used as the underlying physical theory. It has been suggested [1] that quantum information generalizes classical information in a sense analogous to that in which complex numbers generalize real numbers. However, despite impressive advances in the field, the specifics of this generalization remain unclear, and various experts have presented rather different views as to the nature of quantum information; see, for example [2,3,4]. It would seem that the organizing principles and general laws which govern quantum information have yet to be discovered, or at least have yet to be recognized for what they are.

The aim of this paper is to make a clear, specific proposal about the nature of quantum information, and then apply it to the problem of the physical location, in space and time, of information in a quantum system. The proposal may or may not turn out to be correct (or interesting, or useful, etc.), but it is at least unambiguous, and can be tested in various ways. In this paper we apply it to some specific examples, including dense coding and teleportation, to illustrate what the concepts mean, and to show that they provide sensible results in certain situations which have often seemed obscure and hard to think about in physical terms, even though their mathematical description is relatively straightforward.

Our presentation begins in Sec. II with a brief discussion of classical information and its location in space and time. This shows what it is that we want to generalize to the quantum case, and provides some analogies which are helpful later in the paper. The discussion of quantum information begins in Sec. III with an analysis of how much information can be contained in, or carried by, a single spin-half particle, a single qubit. We argue that this is at most one bit, and give reasons why it is not larger despite the fact that an infinite amount of information is needed to describe (with infinite precision) the mathematical state of the qubit as an element of its Hilbert space. In addition, in Sec. III B we develop a classical analogy for a spin-half particle, one that provides what seems to be a more satisfactory intuitive picture than the fairly common mental image of a gyroscope with its axis pointed in a precise direction in space.

Since the quantitative measures of information, such as entropy, used in ordinary (classical) information theory are expressed in terms of probabilities, their extension to the quantum domain requires a consistent probabilistic formulation of quantum theory. For our purposes the traditional approach based upon outcomes of measurements is not adequate, and Sec. IV A indicates how this goal can be reached using quantum frameworks (consistent families). One consequence, Sec. IV B, is a maximum information capacity of ''one bit per qubit'' for carriers of quantum information, confirming the conclusions reached in Sec. III. The use of frameworks also allows one to assign a location to quantum information, as explained in Sec. IV C.

Dense coding is taken up in Sec. V. It has often been supposed that this is a situation in which a single qubit can transport two bits of information. However, using the tools developed in Sec. IV we show that this is not so: the two bits (and each bit independently) reside on two qubits, not one, so the maximum capacity result of Sec. IV B is respected. We also show that the information does not travel backwards

*Electronic address: rgrif@cmu.edu

in time, and argue that it is not transmitted instantaneously, by some peculiar superluminal influence violating classical causality, to a distant qubit. To be sure, there is something mysterious (i.e., nonclassical) about dense coding, and our analysis makes clear precisely what it is.

Quantum theory allows a physical system to be described by a variety of different frameworks, and the presence or absence of information (and hence its location) depends, in general, upon the framework. As an aid to thinking about this *framework dependence*, which is a novel feature of quantum mechanics with no classical analog, we introduce in Sec. VI A the notion of a *structure*, or collection of frameworks with a common algorithm for calculating probabilities. A particular example of a structure is an ideal quantum channel, discussed in Sec. VI B from the point of view of the physical location of information in the different frameworks that constitute the channel, and this is used to define the location of the channel itself.

The tools developed in Sec. VI B for locating a quantum channel are applied in Sec. VII to a particular three-qubit quantum circuit which has been used to discuss teleportation [5]. This circuit is analyzed from different points of view, i.e., by using different structures, and the location of the channel is shown to depend upon which of these one employs. Teleportation itself corresponds to a particular structure, and we identify the location of the relevant information during the ''classical'' communication process. Once again, the information does not travel backwards in time, nor are there magical long-range influences.

The concluding Sec. VIII summarizes the basic proposal of this paper in terms of two theses. The first is that the use of frameworks is an appropriate (consistent, useful, etc.) generalization of classical information to quantum mechanics: it embodies at least something of what quantum information is all about. The second, more speculative thesis is that this generalization is sufficient to cover, at least in principle, *all* of quantum information: that is, there is no additional, irreducibly-quantum information which is not ''classical'' information in a particular framework, or else the framework dependence of such information.

The only other work we are aware of which addresses in a similar way questions about the location of quantum information is that of Deutsch and Hayden [3]. Aside from one or two points of similarity, e.g., we agree that there can be no instantaneous (superluminal) transmission of information in a quantum system, their approach is completely different from ours; see the comments in the Appendix.

II. CLASSICAL INFORMATION

The morning newspaper contains an account of a speech which the President made yesterday. One can say that the newspaper *contains information* because from the shapes and positions of the symbols on the paper it is possible to infer something of what the President said. It is a *carrier* of information about a *message*, in this case the original speech. This information has a physical location in the same sense that the newspaper has a physical location. Of course this location is not unique—there are many copies of the newspaper—nor is it a mathematically precise location, since the printed words take up some space, and the message cannot be recovered from a single letter or even a single word; in some sense it resides in the correlations among the symbols.

The notions of information and its location can be made more precise by embedding them in a probabilistic or stochastic model, something which is necessary if one wants to use the concepts of classical information theory [6]. Let M be a random variable representing the different possible messages which can be produced by a source, and let N be another random variable representing different physical states of a carrier of information. If for any M that occurs with nonzero probability there is some N such that the conditional probability $\Pr(M|N)$ is 1, we shall say that the carrier contains the information produced by the source, whereas if M and N are statistically independent, the carrier does not contain this information. It is convenient to use Shannon's mutual information $I(M:N)$ [6] as a quantitative measure of the information present in N, especially when, as will be the case in most of the discussion that follows, we are interested in a small number of messages which occur with equal probability.

As a simple example, suppose Alice sends a one-bit message $M=0$ or 1 to Bob by mailing him a colored slip of paper, red (R) for 0, green (G) for 1, inside an envelope. While it is in transit the message is contained in or carried by the slip of paper, and if 0 and 1 are equally likely, the slip carries $\log_2 2=1$ bit of information. Next consider a more complicated arrangement in which the message is encoded in two slips of paper which Alice sends to Bob in separate envelopes: identical colors, RR or GG, chosen at random, stand for 0, and opposite colors, RG or GR, also chosen at random, stand for 1. In this case a single envelope contains no information about the message, since both for $M=0$ and $M=1$, R or G will be present with probability 1/2. Instead, the message is present as a *correlation* between the contents of the two envelopes. It is *delocalized* in the sense that it is not present in either envelope by itself, but can still be said to be in the region of space occupied by both envelopes—the union or their interiors, to use a set-theoretical term—even in cases in which the envelopes are far away from each other. Information is something abstract whose location need not satisfy the usual rules for the location of material objects.

A third example useful in a later discussion of an analogous quantum situation is the following: Suppose that Alice and Bob initially share slips of paper that are known to be identical in color: both R or both G. For example, they were prepared and placed in separate envelopes sent to Alice and Bob by Charlie, who flipped an honest coin to decide between RR and GG. Alice now sends a one-bit message M to Bob in the following way. For $M=0$ she sends him the slip of paper in her possession, and for $M=1$ she sends him a slip of the opposite color. Where is the information about M located while this slip is in transit from Alice to Bob? It cannot be in the slip which is in the mail, because in our stochastic model both $M=0$ and $M=1$ lead to R and G with equal probability. Nor is it, obviously, in the slip that is already in Bob's possession. Instead, it is present in the correlation between these two slips, and physically located in the

union of the two regions that they occupy.

One can imagine more complicated situations. For example, modify the previous example so that Charlie prepares RG and GR as well as RR and GG, each with probability 1/4, and sends the first slip to Alice and the second to Bob. Later, at the same time Alice sends her message to Bob by following the procedure in the preceding paragraph, Charlie mails to Bob a yellow slip of paper indicating that the earlier slips he sent were of the same color, or a blue slip indicating that the colors were opposite. While the slips from Alice and Charlie are in transit to Bob, the message is contained in the correlation between these and the slip already in Bob's possession, whereas it is not present in the individual slips or in any of the three pairs of two slips. A quantum analog of this situation will be discussed later.

III. SPIN HALF

A. Distinct states

The simplest quantum carrier of information is a two-level system or qubit, which for convenience we shall think of as the spin degree of freedom of a spin-half particle, although the polarization of a photon would do just as well. Suppose a polarization apparatus (polarizer) produces a particle with its z component of spin angular momentum S_z equal to $+1/2$ or $-1/2$ (in units of $\hbar$), depending upon the setting 0 or 1 of a switch M. This particle then travels through a field-free region to a Stern-Gerlach measuring apparatus where the value of S_z is measured, with outcome N equal to 0 or 1 indicated by a suitable pointer. As there is a perfect correlation between the polarization setting M and the measurement outcome N, the mutual information $I(M{:}N)$ is $\log_2 2$ or 1 bit, assuming equal *a priori* probabilities for M.

It makes sense to suppose that this 1 bit of information is carried by the spin-half particle during the time between preparation and measurement, since, for example, the perfect correlation between M and N requires that the particle actually arrive at the measuring apparatus, and can be altered if the particle is scattered or passes through a magnetic field during its trajectory. That the particle contains information about M can be confirmed using the formulation in Sec. II by working out the joint probability distribution for the random variables M and S_z,

$$\begin{aligned}\Pr(M=0,S_z=+1/2)&=1/2\\&=\Pr(M=1,S_z=-1/2),\end{aligned}\tag{1}$$

with $\Pr(M=0,S_z=-1/2)$ and $\Pr(M=1,S_z=+1/2)$ both equal to 0. Introducing probabilities that refer to a microscopic quantum system requires some care, and we shall return in Sec. IV A to the considerations which justify (1). Assuming its validity, the situation is analogous to one in which a switch setting of M equal to 0 or 1 on one piece of apparatus gives rise to two distinct macroscopic electrical signals in a cable (the analogs of $S_z=\pm 1/2$) connecting this apparatus to a second one in which the signals give rise to two different pointer positions. In such a situation one would not hesitate to say that the information about M is carried by the electrical signal, and in the same way it seems plausible that information is carried by the spin-half particle.

However, the quantum case is conceptually more complicated than its classical analog because the polarizer can prepare other components of spin angular momentum besides S_z. There might, for example, be additional switch settings $M=\overline{0}$ or $\overline{1}$ which result in a particle with $S_x=+1/2$ or $-1/2$, respectively. The vectors in the Hilbert space corresponding to $S_x=\pm 1/2$ are quite distinct from those corresponding to $S_z=\pm 1/2$, and if these were to correspond to physically distinct states of the quantum particle one might plausibly conclude that the particle could carry $\log_2 4=2$ bits of information. Indeed, there is no reason not to allow the polarizer the possibility of producing a particle with a polarization $S_w=1/2$ for w any direction in space. Since the states (or, to be more precise, the rays or one-dimensional subspaces) of the Hilbert space corresponding to distinct w are different, a single spin-half particle would, according to this point of view, be capable of containing or carrying an infinite amount of information. (This idea seems to be widespread; for a specific reference, see [2].)

There are, however, serious objections to the possibility of a spin-half particle carrying more than 1 bit of information in its spin degree of freedom. To begin with, the classical definition of $\log_2 n$ as the information (entropy) of a source of n equally likely messages requires that these messages be *distinct* from one another. From a quantum perspective, any two states of affairs which are classically (i.e., macroscopically) distinct, such as two pointer positions, always correspond to quantum states that are *orthogonal* to each other. (If the distinct states are described by two density matrices ρ and ρ', then $\rho\rho'=0$.) For spin half, the eigenstates corresponding to $S_z=+1/2$ and $-1/2$ are orthogonal, whereas eigenstates of S_z are not orthogonal to eigenstates of S_x. Therefore, $S_z=+1/2$ and $S_x=+1/2$ do not correspond to distinct states in the classical sense, and this is confirmed by the fact that there is no measurement that can distinguish them. To be sure, there might be a distinction which is physically present but is inaccessible to measurements. But this sounds a bit like a student's claim to understand a subject when he has just failed the examination, and one tends to be skeptical.

To be sure, the states $S_z=1/2$ from $S_x=1/2$, while they cannot be distinguished for an individual particle, are distinct in that they give rise to different statistical predictions for the outcomes of various measurements. Such a distinction does not, however, mean that individual particles are in some sense distinct depending on whether they were prepared with the polarizer switch setting of $M=0$ or $M=\overline{0}$, as can be seen from a classical analogy. Consider an ensemble of bolts produced by machine no. 1 that creates one defective bolt in 10, and another ensemble produced by machine no. 2 that creates one defective bolt in 20. Separate ensembles of bolts produced by the two machines can be distinguished with a probability that approaches 1 if the ensembles are sufficiently large, but the distinction cannot be made on the basis of a single bolt. Distinct probability distributions do not mean distinct bolts, and it would be somewhat strange to

claim that because the probability of a defective bolt is described by a real number p having an uncountably infinite number of possibilities, that therefore every bolt carries an infinite amount of information relative to its being or not being defective. Of course, a pure quantum state is not the same thing as a probability distribution. But when it functions as a preprobability [7] used to calculate a probability distribution, it has a somewhat similar character.

In summary, the undeniable *mathematical* difference between the states $S_z=1/2$ and $S_x=1/2$ as elements (or, better, rays) in the Hilbert space need not reflect a *physical* distinction in any sense analogous to those between two nonidentical classical messages. And if it does not, then it provides no basis for supposing that a single qubit can contain more than one bit of information.

B. Classical analogy

One source of the (mistaken, in our opinion) idea that a spin-half particle can carry an infinite amount of information is a mental image in which one thinks of it as a classical object like a gyroscope, spinning about an axis pointing in a well-defined direction in space, say the $+z$ direction if the quantum state is $S_z=+1/2$. Such mental images, or models, of quantum objects in classical terms are probably unavoidable in that quantum theory is a product of human thinking that is shaped in large part by ordinary ''classical'' experience, and they need cause no harm if one understands their limitations and realizes how they can be misleading. The gyroscope model under discussion is misleading in that the components of angular momentum in directions perpendicular to the special axis are zero, whereas one knows that the quantum operators S_x^2 and S_y^2 are identical to S_z^2, and thus thinking of S_x and S_y as equal to zero when $S_z=1/2$ cannot be correct. A modification of this model, which is still classical but a bit less misleading, has the axis of the gyroscope pointing in a random direction in space consistent with a definite positive value of S_z, but with random nonzero values of S_x and S_y. It resembles what one finds in older texts on atomic physics, where even when S_z has its maximum value, the total angular momentum is shown as a vector pointing in a direction which does not coincide with the z axis.

If one uses the first model, in which $S_z=+1/2$ corresponds to the gyroscope axis pointing along the z axis, one tends to get the idea that the quantum state contains a large amount of *orientational* information needed to specify the precise direction in which the spin is ''pointing.'' This is reinforced by the observation that a spin polarizer, a large classical apparatus, can be oriented quite precisely, combined with the plausible, but questionable, supposition that the same precise orientation is imparted to the spin-half particle. The second model, in which the orientation of the gyroscope axis is to a large extent random, suggests that, on the contrary, while the polarizer can fix one component of the angular momentum quite precisely, it does not impart a precise orientation, for the other components of angular momentum are random. That the polarizer should produce random values for the components of angular momentum perpendicular to the polarization direction is consistent with the usual belief that a measurement of one component of the spin of a spin-half particle randomizes the orthogonal components, when one remembers that such measurements are one way of producing polarized particles. In summary, while this second model remains classical, and is thus bound to be misleading in some respects, in other respects it provides a significantly better intuitive feeling for the quantum situation than does the first model.

It is possible to carry out a quantitative analysis of the information contained in a simplified version of the second model that no longer respects the requirement of rotational invariance. We suppose that a classical object has x, y, and z components of angular momentum equal to $+1/2$ or $-1/2$ in some set of units, and thus a set of eight possible states, which can be denoted by $(\pm,\pm,\pm)$; e.g., $(+,+,-)$ stands for $S_x=+1/2$, $S_y=+1/2$, $S_z=-1/2$. We assume that the polarizing apparatus in preparing a state $S_z=+1/2$ randomly perturbs the x and y components of angular momentum, so that it produces one of the four states $(\pm,\pm,+)$ with equal probability.

Now suppose that the polarizer is equipped with six switch settings allowing it to produce the states $S_x=\pm 1/2$, $S_y=\pm 1/2$, and $S_z=\pm 1/2$, where in each case the state has the specified value for that component of angular momentum, but random values for the other two. It is then clear that from the resulting ''spin configuration'' of the object, say $(+,-,+)$, one can decide unambiguously between the polarization settings $S_x=+1/2$ and $S_x=-1/2$, but not between $S_x=+1/2$ and $S_z=+1/2$. For this model one can calculate the mutual information between the six polarization settings, assumed equiprobable, and the spin state, and the result is 1 bit. That is to say, the object can carry enough information to distinguish two opposite values of one component of angular momentum, but contains no additional information about the orientation of the polarizer.

Since this model is classical it cannot clarify all the conceptual difficulties of the quantum case. However, it does suggest how a preparation procedure can impart to a single particle much less information than is required to describe the procedure (in particular, the orientation of the polarizer) itself.

IV. QUANTUM INFORMATION

A. Frameworks

While the qualitative arguments in Sec. III A and the (second) classical model in Sec. III B provide reasons for supposing that a spin-half particle cannot carry a large amount of information in its spin degree of freedom, they are obviously not substitutes for a precise discussion properly grounded in the fundamental principles of quantum theory, which we shall now begin to construct.

Classical information theory is formulated in probabilistic terms, and hence it makes sense to try and construct a quantum counterpart using a consistent probabilistic formulation of quantum theory. Unfortunately, quantum textbooks usually introduce probabilities in terms of the outcomes of measurements, and take a more or less agnostic attitude towards

what goes on between measurements, or between the preparation of a system by some macroscopic device and its measurement by another. During this time interval the quantum system is thought of as a black box, or ''smoky dragon.'' Such an approach is obviously not adequate for addressing questions about the amount of information carried by a single qubit when it is not being measured. One might still make some progress if one could assume that the outcome of a measurement reveals a property the measured system had before the measurement took place, but it is precisely this that the traditional approach denies. Fortunately, there are other tools available for assigning probabilities in a consistent way to quantum systems without reference to measurements [8–14].

Ordinary probability theory is based upon the notion of a *sample space* of mutually exclusive possibilities, one and only one of which occurs in any given experimental trial: e.g., the two sides of a coin, or the six sides of a die. The quantum counterpart of a sample space is a *framework*, also known as a consistent family or decoherent set. The simplest example is an orthonormal basis $\{|a_j\rangle\}$ of states of the quantum Hilbert space, or, equivalently, the associated collection of projectors $P_j=|a_j\rangle\langle a_j|$ which form a decomposition of the identity,

$$I=\sum_j P_j, \quad P_j=P_j^\dagger, \quad P_jP_k=\delta_{jk}P_j. \tag{2}$$

It is important that the states be orthogonal, for only then do they correspond to physically distinct states that are mutually exclusive (see the discussion in Sec. III A). A framework or quantum sample space for a system at a single time is always associated with a decomposition of the identity $\{P_j\}$ with the projectors satisfying (2), but some of these projectors may project onto subspaces of dimension greater than 1, in which case the decomposition is not associated with a unique orthonormal basis.

For a spin-half system the orthonormal basis consisting of the eigenstates of S_z constitutes a framework, and that corresponding to the eigenstates of S_x a different framework. These two frameworks are *incompatible*: it is impossible to combine them into a single sample space, because one cannot simultaneously assign values to S_x and S_z for a spin-half particle (there is no ray in the Hilbert space corresponding to the combined values) [15]. More generally, two frameworks are incompatible when the projectors corresponding to one decomposition of the identity fail to commute with the projectors of a different decomposition. While it is possible to assign probabilities separately in different frameworks, it is impossible to combine probabilities associated with incompatible frameworks, because there is no common sample space. This is an instance of the *single framework rule* [16], which states that it is meaningless to try and combine two (or more) quantum descriptions of a system based upon incompatible frameworks. This rule is important, because when one is analyzing quantum systems it is easy to be misled into combining results in a manner which is quite acceptable in classical physics, where incompatible frameworks never arise, but which leads to quantum paradoxes.

When one is considering probabilities of events that take place at different times, as in (1), it is necessary to employ a framework or quantum sample space of *histories*, sequences of events at different times. Rules for incompatibility of frameworks and for assigning probabilities to histories in a single framework have been worked out by Gell-Mann and Hartle [14], Omnès [9,12,13], and the author [11]; for a recent, detailed formulation see Chap. 10 of [8]. These are more complicated than those for a single Hilbert space at a single time, but for the purposes of the present paper these complications, embodied in what are called ''consistency conditions'' or ''decoherence conditions,'' can for the most part be ignored; exceptions will be noted as they occur. One can also employ frameworks to discuss measurements, provided the measuring apparatus is regarded as a quantum object, part of an overall system that is described using quantum theory. The probabilities of measurement outcomes are then the same as those calculated by the rules of standard textbook quantum theory. Hence the use of frameworks is not a new version of quantum theory different from what is found in textbooks, but instead a consistent extension that allows probabilistic descriptions of microscopic as well as macroscopic systems.

The use of frameworks provides a quantum mechanical justification for (1), where the values of M refer to the polarizer switch setting at a time before the particle is produced, and $S_z=\pm 1/2$ to a later time before any measurement occurs. The assignment of probabilities in (1) is a consequence of the usual Born rule along with the assumption that $M=0$ and $M=1$ are equally probable. Another framework can be used to discuss the values of S_x rather than S_z. It is possible to incorporate all four values of M, $0,1,\overline{0},\overline{1}$—note that these are four mutually exclusive possibilities—at the earlier time in either the S_x or S_z (at the later time) framework, but the S_x and S_z frameworks are incompatible, since it makes no sense to simultaneously assign values of S_x and S_z to a spin-half particle.

B. Maximum capacity

Consider a carrier of information described quantum mechanically by a Hilbert space with a finite dimension d. Then d is the sum of the dimensions of the different subspaces corresponding to the projectors making up a decomposition of the identity, as is obvious by taking the trace of both sides of (2). Consequently, there can be at most d projectors in such a decomposition of the identity, and the corresponding sample space contains at most d mutually exclusive possibilities. For this reason, the maximum amount of information that can be carried by this system is the same as that of a classical object with d different states, namely $\log_2 d$ bits. In particular, for $d=2$, a single qubit, the maximum capacity is 1 bit; for $d=2^n$, n qubits, it is n bits. Thus a quantum system can contain at most one bit of information per qubit.

One might suppose that this limit could be exceeded by using more than one framework. However, that would contradict the single framework rule: one could not find a single sample space or a corresponding random variable (Hermitian operator) to represent the additional information. To better

understand what is involved in using different frameworks one needs to look at various examples, which is what we shall do in the following sections.

C. Location of information

We looked at various examples in Sec. II where classical information can be assigned a location in the sense that a carrier of information (newspaper, colored slip of paper) can be assigned a location. It seems plausible that the same principle applies to the quantum case: if a spin-half particle is in a particular location and its spin is appropriately correlated with some initial message, then the information is located wherever the particle is located. Quantum particles cannot be assigned a precise location, but neither can newspapers, so this is not a serious difficulty. Of greater concern is the fact that quantum theory allows the use of different incompatible frameworks for describing quantum entities, and, as we shall see, the presence or absence of information in some carrier of information can depend upon the framework used to describe it.

Let us begin with the simple case of a spin-half particle prepared by a polarizer in the state $S_z = +1/2$ and traveling through a field-free region, as in Sec. III A. If we use the S_z framework, this component of spin is correlated with the switch setting M on the polarizer, and the particle can be said to contain or carry the information as to whether M was 0 or 1. If instead one uses the S_x framework, the joint probability of $M=0$ or 1 and S_x can be calculated using the Born rule, and one finds that they are statistically independent, so in this framework there is no information about the two switch settings. (For present purposes we are ignoring $M=\overline{0}$ or $\overline{1}$.) That the information should be present in one framework and not another is not too surprising given the classical analogy (the second model) introduced in Sec. III B: If the polarizer stores information in the z component of angular momentum while leaving the x and y components undefined or random, it is not surprising that the information is later found in S_z and not in S_x. Thus in studying quantum information it is important to look at a variety of frameworks.

Consider a situation in which two spin-half particles are used to carry two bits of information. As long as each particle is polarized separately, the situation is straightforward. Suppose that polarizer A imparts to particle a a polarization of $S_z = \pm 1/2$ and polarizer B imparts to particle b a polarization of $S_x = \pm 1/2$. Then one can exhibit perfect correlation between the initial switch settings of the polarizers and the spin states of the particles by using for the latter a framework consisting of the orthonormal basis (with kets in the order $|a,b\rangle$) of

$$|0,+\rangle, \quad |0,-\rangle, \quad |1,+\rangle, \quad |1,-\rangle, \tag{3}$$

where $|0\rangle$ and $|1\rangle$ are the states with $S_z = +1/2$ and $-1/2$, respectively, and

$$|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}, \quad |-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}, \tag{4}$$

are states with $S_x = +1/2$ and $-1/2$, with a particular choice of overall phase. For this framework, information about A will be located wherever particle a is located, information about B will be at the location of particle b, and the combined information can be said to be in the union of the two regions of space occupied by the two particles.

This is entirely analogous to the corresponding classical case and would scarcely be worth mentioning except that it provides a contrast with a very different way in which two bits of information can be encoded in two spin-half particles by using the four Bell states

$$\begin{aligned} |B_{00}\rangle &= (|00\rangle + |11\rangle)/\sqrt{2}, \\ |B_{01}\rangle &= (|01\rangle + |10\rangle)/\sqrt{2}, \\ |B_{10}\rangle &= (|00\rangle - |11\rangle)/\sqrt{2}, \\ |B_{11}\rangle &= (|01\rangle - |10\rangle)/\sqrt{2}. \end{aligned} \tag{5}$$

As these form an orthonormal basis of the two spin system, they can be used as a framework of mutually exclusive possibilities. One could imagine that they are produced by a suitable apparatus, with each state corresponding to one of four settings of a switch on the apparatus. Then the information about the switch setting is located in the union of the regions occupied by the two particles. However, in contrast to the previous case, it is not possible to say that any part of this information is located in either particle by itself.

To begin with, each of the Bell states is a delocalized entity of a type which possesses no classical analog: it is impossible to ascribe properties to the individual spins when they are in such an entangled state. The reason is that a projector representing some property of particle a is necessarily of the form $\frac{1}{2}(I_a + S_{aw})$, where the subscript a labels the particle, w indicates some direction in space, and I is the identity operator on the Hilbert space. Such a projector does not commute with a projector $|B_{jk}\rangle\langle B_{jk}|$ representing one of the Bell states. To suppose that a quantum system can simultaneously possess properties corresponding to two noncommuting projectors is to abandon standard quantum mechanics in favor of a theory of hidden variables. And it violates the single-framework rule.

Even if a system has been prepared in one of the four Bell states by some apparatus, there is no reason why at a later time we must describe it using the Bell states as an orthonormal basis. We could, instead, adopt a basis in which the individual particles a and b are assigned specific properties, for example, S_{az} and S_{bz}. If one works out the corresponding probabilities (which are the same as if S_{az} and S_{bz} were measured), one will find that in this framework the difference between a preparation which produced $|B_{00}\rangle$ or $|B_{10}\rangle$ on the one hand, and $|B_{01}\rangle$ or $|B_{11}\rangle$ on the other, is present in the correlation between the value of S_{az} and S_{bz}, but not in the values of S_{az} by itself or S_{bz} by itself. Other frameworks can be used to extract other pieces of information, but in no case is any part of the information located in one of the particles by itself. The situation is thus analogous to the example of classical information encoded in two slips of paper of the same or of different colors, Sec. II, and thus not available in either slip by itself.

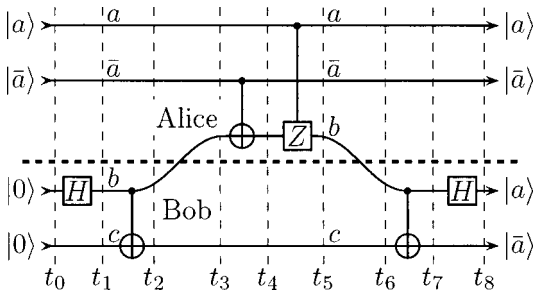

FIG. 1. Quantum circuit for dense coding.

V. DENSE CODING

A. Circuit and unitary time development

As noted in Sec. I, the phenomenon of dense (or superdense) coding has been interpreted as indicating that a single qubit can somehow contain two bits of information. If true, this would contradict the maximum capacity result of Sec. IV B, and both for this reason and also because it provides an instructive example of using frameworks to locate information, dense coding is worth examining in some detail.

A quantum circuit corresponding to this process is shown in Fig. 1, with the usual convention that time increases from left to right. Each of the two qubits a and $\bar{a}$ in Alice's laboratory, above the dashed line, is initially in a state $|0\rangle$ or $|1\rangle$, and together they constitute an information source with four possible messages. In Bob's laboratory, below the dashed line, two qubits b and c, initially both in the $|0\rangle$ state, are entangled by applying a Hadamard gate H to $|b=0\rangle$, producing $(|0\rangle+|1\rangle)/\sqrt{2}$, followed by a controlled-NOT in which the c qubit is flipped from $|0\rangle$ to $|1\rangle$, or vice versa, if and only if the b qubit is in the state $|1\rangle$, where the notation follows [17]. These two unitary transformations place the two qubits b and c in the Bell state $|B_{00}\rangle$, see (5). At this point the b qubit is transported to Alice's laboratory and the information in the a and $\bar{a}$ qubits is written into it using a controlled-NOT and a controlled-Z gate. The latter multiplies the total state by -1 if and only if both a and b are in the state $|1\rangle$. After this the b qubit is returned to Bob's laboratory where it and the c qubit undergo the inverse of the operation that initially entangled them. The overall unitary time evolution results in the b qubit being in the same state, $|0\rangle$ or $|1\rangle$, as the a qubit, and c in the same state as $\bar{a}$, as indicated in the figure. (Note that if a is in a superposition of $|0\rangle$ and $|1\rangle$, this same state will *not* emerge at the final time at either of the points labeled $|a\rangle$ in the diagram, and the same is true of $\bar{a}$.) The term "dense coding" refers to the idea, which at first glance seems intuitively plausible, that the single qubit b on its way back to Bob's laboratory contains the information that was initially in both a and $\bar{a}$, which is to say two bits of information, and it is this idea that we shall investigate using various frameworks.

B. Different frameworks

Let us begin with the framework $\mathcal{F}_1$ corresponding to the unitary time development of the initial states discussed above. One then finds that during the time interval from t_5 to t_6 in Fig. 1 the qubits b and c are in one of the entangled Bell states defined in (5), which state depends on the values of a and $\bar{a}$. But in this situation, as discussed in the Sec. IV C, the information cannot be said to be present in qubit b, instead it is in a sort of quantum correlation between b and c. Consequently, if we use $\mathcal{F}_1$ there is no violation of the maximum capacity result of Sec. IV B, because 2 bits of information are carried by a Hilbert space of dimension $2\times 2=4$ representing both qubits.

The reader may find this result, even though formally correct, to be intuitively troubling in that it seems as if the information which at times before t_3 was entirely in Alice's laboratory has somehow managed to "jump," so that at time t_5 it is at least partly in Bob's laboratory in qubit b, despite the fact that this qubit has not interacted with either a or $\bar{a}$. Have we uncovered some mysterious long-range interaction which can influence b despite the fact that it is far away from Alice's laboratory?

To see that there is nothing particularly odd about information suddenly appearing in this manner in a delocalized correlation with a distant object, it is useful to analyze the circuit in Fig. 1 using a different framework $\mathcal{F}_2$ in which we suppose that at t_2 and all later times each qubit is described using the S_z or computational basis, $|0\rangle$ or $|1\rangle$, which for this discussion we shall denote by 0 or 1 without using the ket symbol. (The state of b at t_8 is an exception, see below.) Given the initial states shown in the figure, it follows from the Born rule that in this framework at time t_2 either $b=c=0$ or $b=c=1$, with equal probability. The same is true at t_3, but between t_3 and t_4 the b qubit will be flipped from 0 to 1 or from 1 to 0 if $\bar{a}=1$, or remain the same if $\bar{a}=0$. As a consequence, at t_4 the information initially present in $\bar{a}$ will be present in the *correlation* between b and c: if $\bar{a}=0$, then $b=c$, and if $\bar{a}=1$, then $b\neq c$. Notice that this information is *not* present in either b or c separately, for at t_4 each is 0 or 1 with a probability of 1/2, the same as at t_3. Once again we have a situation in which information "suddenly" (between t_3 and t_4) appears as a correlation with a distant object which has taken no part in any interaction. However, the situation is precisely parallel with the classical analogy discussed in Sec. II, in which Alice and Bob share two slips of paper which are both red or both green, and there are obviously no nonlocal influences or other sorts of magic. And if such an information "jump" is unproblematic in the classical case, and thus in framework $\mathcal{F}_2$, there seems to be no reason to be concerned about it in framework $\mathcal{F}_1$.

Now let us continue the description using framework $\mathcal{F}_2$ to times later than t_4 in Fig. 1. Between t_4 and t_5 the b qubit is left unchanged, for even if $a=1$ the only effect of this is to change the sign of $|b=1\rangle$, but this has no effect upon the projector $|1\rangle\langle 1|$. (This is perhaps clearer if one employs the histories formalism, see the next paragraph.) Nor is there, obviously, any change between t_5 and t_6. The next controlled-NOT gate has the effect that c is equal to 0 or 1 at t_7 depending upon whether $b=c$ or $b\neq c$ at t_6. Consequently, the information originally contained in $\bar{a}$, which during the time interval from t_4 to t_6 is present as a correlation between b and c, emerges at the end in c, as indicated in the

figure. The final Hadamard gate changes qubit b from $|0\rangle$ to $|+\rangle$ or $|1\rangle$ to $|-\rangle$ as the case may be; in $\mathcal{F}_2$ one cannot assign $|0\rangle$ or $|1\rangle$ at t_8 to b without violating consistency (decoherence) conditions [18].

The somewhat informal discussion of framework $\mathcal{F}_2$ in the previous paragraphs can be given a formal justification using the consistent histories formalism, where the notation is that of [8]. Consider a consistent family with four initial states

$$|\Psi_0^{a\bar{a}}\rangle = |a,\bar{a},b=0,c=0\rangle, \tag{6}$$

where a and $\bar{a}$ are either 0 or 1, and a set of histories of the form

$$[\Psi_0^{a\bar{a}}]\odot F_2\odot F_4\odot F_7, \tag{7}$$

where the subscripts refer to the times shown in Fig. 1. (One could include additional events at additional times, but this selection suffices for displaying the essential ideas.) In (7) and below we use the abbreviation $[\Phi]$ for the projector $|\Phi\rangle\langle\Phi|$. For each possible initial state there are two histories

$$[\Psi_0^{a\bar{a}}]\odot[a,\bar{a},0,0]_2\odot[a,\bar{a},\bar{a},0]_4 \quad \odot[a,\bar{a},\bar{a},\bar{a}]_7,$$

$$[\Psi_0^{a\bar{a}}]\odot[a,\bar{a},1,1]_2\odot[a,\bar{a},1-\bar{a},1]_4\odot[a,\bar{a},1-\bar{a},\bar{a}]_7 \tag{8}$$

with nonzero probabilities. Here subscripts have been added to the projectors to identify the time. For a given a and $\bar{a}$ the two nonzero probabilities are equal to each other; they can be calculated using the methods discussed in [11] or in Chaps. 10 and 11 of [8]. From these one can see that at t_4 information corresponding to the value of $\bar{a}$ is, indeed, present in the correlation between the last two qubits b and c, but not in either of these separately, because one does not know which of the histories in (8) is the one that actually occurs.

Finally, let us consider a third framework $\mathcal{F}_3$ which is similar to $\mathcal{F}_2$ except that for times in the interval $t_2 \leqslant t \leqslant t_6$ we use the S_x basis for qubits b and c, i.e., the states $|+\rangle$ and $|-\rangle$ defined in (4), while retaining $|0\rangle$ and $|1\rangle$ for a and $\bar{a}$. Since b and c evolve unitarily to

$$|B_{00}\rangle = (|00\rangle + |11\rangle)/\sqrt{2} = (|++\rangle + |--\rangle)/\sqrt{2} \tag{9}$$

at time t_2, by thinking of this state as a preprobability [7], or by invoking the Born rule, one sees that in framework $\mathcal{F}_2$, $b=c=+$ and $b=c=-$ occur with equal probability. The controlled-NOT between t_3 and t_4 has no effect upon b in this basis, whereas the controlled-Z between t_4 and t_5 interchanges $+$ and $-$ if $a=1$, or leaves b the same if $a=0$. Thus in the time interval between t_5 and t_6 the information originally present in a is found in the correlation $b=c=+$ or $-$ for $a=0$, and $b\neq c$ for $a=1$. The situation is thus analogous to what we found using $\mathcal{F}_2$, but with the roles of a and $\bar{a}$ interchanged. The controlled-NOT between t_6 and t_7 maps $b=c$ into $b=+$, and $b\neq c$ into $b=-$, while the final Hadamard gate maps these to $b=0$ and $b=1$, respectively, so that at the end b is identical to a. On the other hand, c continues to have the same value, $+$ or $-$, at t_7 and t_8 as it had earlier; ascribing a value of 0 or 1 to it at these times would violate the consistency conditions [18].

In terms of the location of information, the $\mathcal{F}_3$ framework leads to a conclusion parallel to the one obtained using $\mathcal{F}_2$: during the time interval from t_5 to t_6 the information about a is located in a correlation between b and c in a manner which is, once again, quite analogous to the classical situation considered in Sec. II. In particular, one cannot say that the information is present in b by itself any more than that it is in c by itself.

C. Discussion

It should be noted that the three frameworks $\mathcal{F}_1$, $\mathcal{F}_2$, and $\mathcal{F}_3$ are incompatible, as is immediately obvious from the fact that they use noncommuting projectors to describe the b and c qubits at t_2 and at each later time. Thus the descriptions provided by these frameworks cannot be combined into a single quantum description; this would be just as meaningless as assigning values to both S_x and S_z for a spin-half particle. Nonetheless, they provide qualitatively similar answers to the question of the location of the information transmitted from Alice to Bob: it is not to be found in the b qubit alone during the time interval between t_5 and t_6, but instead in some sort of correlation between b and c. In $\mathcal{F}_1$ this ''nonlocality'' arises from the delocalized nature of entangled quantum states, which has no classical analog, whereas in $\mathcal{F}_2$ and $\mathcal{F}_3$ it is rather more like the classical nonlocality discussed in Sec. II. In addition, in $\mathcal{F}_2$ and $\mathcal{F}_3$, half of the information—a different half in the two cases—has completely disappeared by the time $t=t_5$, and does not reemerge. That information should be ''invisible'' in certain frameworks is not by itself too surprising, since there is a classical analog, Sec. III B: the x component of angular momentum of a spinning object reveals no information about the z component. However, the inability to combine these incompatible frameworks into a single description is very much a quantum effect.

Note that in no case is the maximum capacity limit of Sec. IV B exceeded, because with each of these frameworks the information is carried jointly by b and c, and not by b alone. One might worry that there could be some other framework in which all of the information is found in b, but one can easily show that this is impossible by calculating the reduced density matrix for b at time t_5 using unitary time development from any of the initial states, and verifying that it is 1/2 times the identity operator whatever may be the initial choice of a and $\bar{a}$. This density matrix functions as a preprobability [19], i.e., it can be used to assign probabilities for any basis in the Hilbert space of b, and the corresponding probabilities are thus independent of the values of a and $\bar{a}$.

Difficulties in understanding dense coding have led to the suggestion [20] that somehow some of the information present in a and $\bar{a}$ travels backwards in time in order to reach Bob's laboratory. However, it is easy to show that given the initial states shown in Fig. 1, at any time $t \leqslant t_3$ any random variable (i.e., any Hermitian operator) on the Hilbert space of the two qubits b and c is uncorrelated with a and $\bar{a}$. Thus the

information does not travel backwards in time, except, perhaps, in some metaphorical sense, and then one would expect this to be equally true in the corresponding classical case discussed in Sec. II.

Nonetheless there *is* something very odd, that is to say nonclassical, about dense coding. In the classical analog, while it is possible to encode information in correlations between two carriers of information, each with a maximum capacity of 1 bit, one can insert at most a single bit of information into this system by using interactions affecting *only one* of the two carriers. In the quantum case the situation is different, for starting with a single Bell state one can generate any of the other three by applying an appropriate unitary interaction to just *one* of the qubits, as noted in [21]. In this respect quantum theory is very different from classical physics, and one has a genuine quantum mystery. But it is a mystery fully compatible with the information bound in Sec. IV B, and with the absence of mysterious nonlocal influences which can transfer information instantly between separated (thus noninteracting) systems.

VI. STRUCTURES AND CHANNELS

A. Structures

Thus far we have discussed information in the quantum context in terms of frameworks. As long as a single framework is employed to describe a quantum system, the usual rules which govern classical probabilities and ordinary (classical) information theory apply, so in this sense there is nothing new. However, in quantum theory, in contrast to classical physics, one has the possibility of using many different incompatible frameworks, and each framework has a distinct probabilistic structure. This *framework dependence* is something with no classical analog, and gives rise to problems of a sort not encountered in other uses of information theory. Examples of framework dependence have already come up in the discussion of dense coding in Sec. V.

For discussions of framework dependence it is useful to introduce the concept of a *structure*: a collection of frameworks, typically mutually incompatible, along with a common algorithm for assigning probabilities in the different frameworks. This is not a precise definition, for the concept is still under development, but a few examples will serve to indicate what we have in mind. A particular density matrix can be used to assign probabilities to the elements of any orthonormal basis of the Hilbert space on which it is defined, and in this case the bases are the different frameworks that constitute the structure, while the density matrix provides the probability algorithm. As another example, suppose Eve is contemplating how best to extract information about the common key which Alice and Bob are in the process of establishing using the BB84 quantum cryptographic protocol [22]. She has to design an apparatus to intercept and in some sense copy a qubit on the way from Alice to Bob, and in working out the design needs to compute what will happen using either an S_x or S_z framework for the initial state of the qubit, as well as various alternatives at later times. The dynamical properties of the eavesdropping device provide the algorithm for calculating probabilities for the different scenarios, each corresponding to some particular framework.

B. Channels

Perhaps the simplest example of a structure involving histories of a quantum system, thus properties at more than one time, is a *quantum channel*. Typically, one is interested in some basis for the Hilbert space representing the input to the channel, which is tensored to a second Hilbert space representing the environment, and then another (possibly the same) basis for the first space at a later time. A particular framework corresponds to a particular choice of the initial and final bases, the structure consists of all frameworks of this sort, and unitary time evolution of the total system (channel plus environment) along with the Born rule is used to calculate probabilities in each framework.

Let $\mathcal{A}$ be the input of the channel, $\mathcal{E}$ the environment, and $\mathcal{H}=\mathcal{A}\otimes\mathcal{E}$ the total Hilbert space. Suppose that at t_0 the environment is always in a fixed state $|e_0\rangle$, while the initial state $|a\rangle$ of the channel is one of a set of states $\{|p^m\rangle\}$, $m=1,2,\ldots$, which form an orthonormal basis of $\mathcal{A}$. At a time $t_1>t_0$, define

$$|\Psi_1^m\rangle = T(t_1,t_0)(|p^m\rangle|e_0\rangle), \tag{10}$$

where $T(t',t)$ is the unitary time development operator, equal to $\exp[-i(t'-t)H/\hbar]$ for a time-independent Hamiltonian H.

Next consider a framework consisting of a family of histories, each of which begins with some initial state $|p^m\rangle|e_0\rangle$ at t_0, and ends with one of the events associated with a decomposition of the identity

$$I=\overline{P}+\sum_m P^m \tag{11}$$

in orthogonal projectors, where $\overline{P}$ might be zero, and the $\{P^m\}$ have the property that

$$P^m|\Psi_1^{m'}\rangle = \delta_{mm'}|\Psi_1^m\rangle. \tag{12}$$

As a consequence, information about the basis states $\{|p^m\rangle\}$ at t_0 is at time t_1 contained in the collection of projectors $\{P^m\}$.

Suppose that the $\{P^m\}$ have been chosen in such a way that for some factorization $\mathcal{B}\otimes\mathcal{F}$ of $\mathcal{H}$, which could but need not be the same as $\mathcal{A}\otimes\mathcal{E}$, it is the case that each projector P^m is *on* $\mathcal{B}$ in the sense that it is an operator of the form $P^m\otimes I$ (which means $\overline{P}$ is also of this form). Then we can say that the information about $\{|p^m\rangle\}$ is *in* $\mathcal{B}$ (or "on" or "carried by" $\mathcal{B}$). Typically $\mathcal{B}$ will refer to some subsystem of the total system described by $\mathcal{H}$, and if this subsystem is located in a particular region of space, then the information can be said to be located in that region.

One can always find projectors such that (12) is satisfied, for example,

$$P^m = |\Psi_1^m\rangle\langle\Psi_1^m|. \tag{13}$$

However, this choice is so precise that $\mathcal{B}$ will be equal to $\mathcal{H}$, and to find more interesting cases in terms of location of information one needs to choose coarser projectors. Some examples will be found in Sec. VII B, or the reader may wish to think about the case in which the time transformation is trivial, $T(t',t)=I$, and $\mathcal{B}=\mathcal{A}$.

In the case of a quantum channel one is usually not interested in just a single orthonormal basis of $\mathcal{A}$ at t_0, but instead all possible orthonormal bases. Let $\{|p_\lambda^m\rangle, m=1,2,\ldots\}$ be the collection of vectors forming the orthonormal basis λ. Here λ can be any convenient label or set of parameters used to label the bases, but it should be chosen so that it does not distinguish bases which differ only through relabeling the basis vectors, or multiplying them by phase factors: in other words, it should label distinct decompositions of the identity into pure states. For example, for a spin-half particle or a single qubit, with $|0\rangle$ and $|1\rangle$ the eigenstates of S_z with eigenvalues $+1/2$ and $-1/2$, define

$$|p_\lambda^m\rangle=\alpha_\lambda^m|0\rangle+\beta_\lambda^m|1\rangle, \tag{14}$$

with $m=0$ or 1, where

$$\alpha_\lambda^0=\frac{1}{\sqrt{1+|\lambda|^2}}, \quad \beta_\lambda^0=\frac{\lambda}{\sqrt{1+|\lambda|^2}},$$

$$\alpha_\lambda^1=\frac{\lambda^*}{\sqrt{1+|\lambda|^2}}, \quad \beta_\lambda^1=\frac{-1}{\sqrt{1+|\lambda|^2}}, \tag{15}$$

for λ a complex number satisfying $|\lambda|<1$, or $\lambda=e^{i\phi}$ with $0\leqslant\phi<\pi$. In particular, $\lambda=0$ is the S_z or computational basis, $\lambda=1$ the S_x basis, and $\lambda=i$ the S_y basis in the familiar Bloch sphere picture.

Let us in addition allow many different times: $t_0<t_1<t_2<\cdots<t_f$. For each basis λ and each time t_j choose a decomposition of the identity

$$I=\overline{P}_{\lambda j}+\sum_m P_{\lambda j}^m, \tag{16}$$

such that

$$P_{\lambda j}^m|\Psi_{\lambda j}^{m'}\rangle=\delta_{mm'}|\Psi_{\lambda j}^m\rangle, \tag{17}$$

where

$$|\Psi_{\lambda j}^m\rangle=T(t_j,t_0)(|p_\lambda^m\rangle|e_0\rangle). \tag{18}$$

Next suppose that at time t_j there is a factorization $\mathcal{B}_j\otimes\mathcal{F}_j$ of $\mathcal{H}$ such that for all λ and every m, $P_{\lambda j}^m$ is *on* $\mathcal{B}_j$. Then since for any λ the information about $\{|p_\lambda^m\rangle\}$ is in $\mathcal{B}_j$, it seems reasonable to say that the channel itself is *in* (or "on" or "carried by") $\mathcal{B}_j$ at time t_j. Let us adopt this as a tentative definition that deserves further exploration to see if it is reasonable and useful. That exploration begins in Sec. VII.

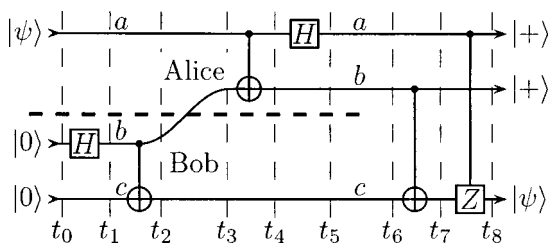

FIG. 2. Quantum circuit for teleportation.

Note that for a fixed λ the set of histories

$$[p_\lambda^m]\odot P_{\lambda 1}^m\odot P_{\lambda 2}^m\odot\cdots\odot P_{\lambda f}^m, \tag{19}$$

with $m=1,2,\ldots$, form the support of a consistent family or framework $\mathcal{F}_\lambda$ in the notation of [8]. Consistency is easily established using the method of chain kets [23]. Note that $\mathcal{F}_\lambda$ and $\mathcal{F}_{\lambda'}$ are incompatible for $\lambda\neq\lambda'$, so it is not correct to think of information about $\{|p_\lambda^m\rangle\}$ and $\{|p_{\lambda'}^m\rangle\}$ as simultaneously present in $\mathcal{B}_j$ even when the channel itself, viewed as a structure, can be said to be present in $\mathcal{B}_j$.

VII. TELEPORTATION CIRCUIT

A. Circuit and unitary time development

The circuit shown in Fig. 2, with labels chosen to facilitate comparison with Fig. 1, was first proposed (in a slightly different notation) as a model for teleportation in [5]; also see p. 187 of [17]. The unknown state $|\psi\rangle$ to be teleported is initially in qubit a, whereas b and c at time t_2 are in the Bell state $|B_{00}\rangle$, see (5). If qubits a and b are in one of the Bell states $|B_{jk}\rangle$ at t_3, the two gates which follow will put them in a product state $|ab\rangle=|jk\rangle$. Consequently, a measurement of both a and b at time t_5 in the computational basis is equivalent to measuring the pair of qubits a,b at t_3 in the Bell basis [24]. In the standard teleportation scenario Alice sends the results of such measurements over a classical channel to Bob, who carries out appropriate unitary operations on qubit c. There are no measurements in the circuit in Fig. 2 and the final unitary operations on c are carried out "automatically" by the last two gates. Thus this circuit does not actually represent teleportation, though it can be viewed in a particular way using an appropriate structure, Sec. VII D below, which makes it essentially equivalent to teleportation. However, it is also of interest as a nontrivial example of an ideal quantum channel, and we shall begin by studying it from this point of view, in particular addressing the question of the location of information in various frameworks.

The circuit in Fig. 2 constitutes a quantum channel in the sense defined in Sec. VI B, where qubits b and c in the state $|e_0\rangle=|00\rangle$ at t_0 are the environment, while the channel input can be thought of as being in a state

$$|a\rangle=\alpha|0\rangle+\beta|1\rangle, \quad |\alpha|^2+|\beta|^2=1. \tag{20}$$

Unitary time development will then result in a state $|\Psi_k\rangle$ for the total system at time t_k, with, in particular,

$$|\Psi_2\rangle=(\alpha|0\rangle+\beta|1\rangle)\otimes(|00\rangle+|11\rangle)/\sqrt{2}, \tag{21}$$

$$\begin{aligned}|\Psi_5\rangle&=(1/\sqrt{2})[(\alpha|+\rangle+\beta|-\rangle)\otimes|++\rangle\\&\quad+(\alpha|+\rangle-\beta|-\rangle)\otimes|--\rangle]\\&=\tfrac{1}{2}[|00\rangle\otimes(\alpha|0\rangle+\beta|1\rangle)+|01\rangle\otimes(\beta|0\rangle+\alpha|1\rangle)+|10\rangle\\&\quad\otimes(\alpha|0\rangle-\beta|1\rangle)+|11\rangle\otimes(-\beta|0\rangle+\alpha|1\rangle)],\end{aligned} \tag{22}$$

$$|\Psi_8\rangle=|++\rangle\otimes(\alpha|0\rangle+\beta|0\rangle). \tag{23}$$

Here $|+\rangle$ and $|-\rangle$ are the S_x basis states defined in (4), the qubits are in the order $|abc\rangle$, and the tensor product symbol $\otimes$ has been inserted at various points for clarity.

B. Location of the channel

Let us now consider a particular basis λ for the a qubit at t_0, with the basis states given by $|a\rangle=|p_\lambda^m\rangle$, (14), with coefficients defined in (15). Then it is easy to construct a framework in which at t_3 the information about the initial states is found in qubit a using the criteria of Sec. VI B. In particular, the decomposition of the identity (16) is of the form

$$I=[p_\lambda^0]\otimes I\otimes I+[p_\lambda^1]\otimes I\otimes I \tag{24}$$

with projectors which are on qubit a, and one can check that the conditions (17) are satisfied. As this is true for any choice of basis λ, we can, using the definition given in Sec. VI B, say that the channel itself is *in* qubit a at t_3, and of course the same argument will work at t_1 and t_2. This agrees with one's intuition in that up to time t_3 the a qubit has not interacted with the other qubits, so one would expect that any information initially present in any basis, and thus the channel itself, would continue to be present in qubit a. Note, by the way, that for all $t\leqslant t_3$ and for any choice of initial basis λ, there is no information about the initial state of a in qubits b or c or the combination of the two—the situation is analogous to the case of dense coding, see the discussion in Sec. V—so once again there is no question of information in some way traveling backwards in time through either b or c [25].

Next let us consider the situation at time t_5. First, it is easy to show that except for special choices of λ the information about the initial state is not present (or at least not completely present) in qubit a by itself, and therefore the quantum channel cannot be said to be present in a. However, it is present in the *pair* a,b. This can be seen most easily by noting that in the first expression for $|\Psi_5\rangle$ in (22), the $|+\rangle$ and $|-\rangle$ states for qubit b occur in conjunction with $\alpha|+\rangle+|\beta|-\rangle$ and $\alpha|+\rangle-|\beta|-\rangle$ for qubit a, and the latter are, in turn, obtained by applying the unitary operators H and HZ, respectively, to the state $\alpha|0\rangle+\beta|1\rangle$. With this as a hint one can work out the projectors for which (17) will be satisfied; they are

$$P_{\lambda 5}^m=H[p_\lambda^m]H\otimes[+]\otimes I+HZ[p_\lambda^m]ZH\otimes[-]\otimes I \tag{25}$$

for $m=0$ and 1, and are obviously located in the pair a,b. As this is the case for any initial basis λ, it shows that the channel itself is located in (or on) the pair of qubits a and b.

A similar analysis shows that the same is true at time t_4, and this is not surprising, since between t_4 and t_5 the only thing that happens is a unitary transformation on a, and while this can change the form in which the information is located on the two qubits, it cannot affect its presence. Indeed, the same sort of reasoning shows that one should not be surprised that the information is in a,b at t_4, since it is already in a and therefore a,b at t_3, and between t_3 and t_4 these qubits interact with each other, but with nothing else. But is it really true that the information is in the a,b pair at t_3 given that it is in a? Yes, because the projectors used to establish the latter are given in (24), and the projector $[p_\lambda^m]\otimes I\otimes I$ on a is also a projector $([p_\lambda^m]\otimes I)\otimes I$ on a,b. Note that the argument by which the information initially in a at t_3 should be present in the pair a,b at t_4 depends upon the fact that the quantum gates in the circuit represent unitary, and thus reversible, operations; one would not reach the same conclusion if they represented stochastic processes. Indeed if one adopts, as is quite possible, frameworks in which some gates produce random effects, information can disappear. (See the first example of Sec. VII C, where the action of the Hadamard gate acting on qubit a destroys information.)

Another way of seeing that the channel must be in a,b at t_5 is to note that it can be extracted and placed once again in qubit a by replacing the part of the circuit following t_6 in Fig. 2 with a different set of gates acting only on a,b. One choice is simply to invert the unitary operation produced earlier on a,b between t_3 and t_5 by applying a second Hadamard gate to a, followed by another controlled-NOT between a (control) and b.

It is immediately evident from the first expression for $|\Psi_5\rangle$ in (22), and also from the second expression if one multiplies out the products, that this state is unchanged if one interchanges b and c. Consequently, the argument based on (17) for the presence of the channel in a,b can also be used to show that it is present in a,c: one only needs to interchange the second and third projectors in each of the terms on the right side of (25). Although the presence of the channel in a,c is not as intuitively obvious as in the case of a,b, one can still check that it is correct by once again replacing that part of the circuit in Fig. 2 which follows t_6 with a suitable set of gates involving only qubits a and c in such a way that the channel reemerges in a (or, if one prefers, in c); this construction is left as an exercise.

But how can it be that the channel is present in a,b and also present in a,c, but not present in a alone? Here one needs to remember that information is an abstract entity, not a physical object, and the rules for its location do not have to satisfy the axioms of set theory. If there is a mystery here, it

is not a quantum mystery, for the same thing occurs in the following classical example. Let 0 be encoded as RRR or GGG, with equal probability, on three colored slips of paper abc, and let 1 be encoded as RGG or GRR, again with equal probability. The information which distinguishes 0 from 1 is not present in any of the slips of paper separately, as in both cases R and G occur with equal probability. However, it can be obtained by comparing a with b or a with c, whereas comparing b with c tells one nothing.

There is, nonetheless, at t_5 a nonclassical effect in that for most choices of the initial basis λ (the exceptions are $\lambda=0$ and $|\lambda|=1$), the framework needed to exhibit the presence of information in the pair a,b is incompatible with that needed for the pair a,c. Consequently, while it makes sense to say that the *channel* is in a,b, or that it is in b,c, it does not make sense to say that it is both in a,b *and* in a,c.

C. Alternative frameworks

Just as in the case of dense coding, Fig. 1, additional insight into the circuit in Fig. 2 can be obtained by considering what happens in some special frameworks. In the first of these the S_z or computational basis is employed for all three qubits at all times. This resembles framework $\mathcal{F}_2$ in the case of dense coding, Sec. V B, so it will suffice to note only the main points. At t_2 (and therefore t_3) the b and c qubits are either both 0 or both 1, with equal probability, and thus at t_4 the information about whether a was 0 or 1 at t_0 resides in correlations between b and c—they are equal for $a=0$, and different if $a=1$—as well as in qubit a itself. But at t_5 the Hadamard gate has randomized a, and it is uncorrelated with its earlier value, whereas that information is still present in the pair b,c. In particular, at t_5 the information is not present in the pair a,b, nor is it present in a,c. This does not contradict the result of Sec. VII B, for the presence or absence of information in some particular location depends upon the framework used to describe the quantum system, and the one under discussion is incompatible with the frameworks needed to exhibit the same information in a,b or in a,c. The presence of information in correlations between b and c is limited to the $\lambda=0$ or S_z basis for qubit a at t_0; for all other λ the information is, at best, partially present. Next, the controlled-NOT between t_6 and t_7 writes the information from the b,c correlation into qubit c, while in this framework the final controlled-Z gate has no effect, as in the analogous situation in Fig. 1.

The second framework we will consider employs the $\lambda=1$ or S_x basis for a at t_0, and also for this and the other qubits at all later times, except that the S_z (computational) basis is used for a at times $t\geqslant t_5$. The situation resembles framework $\mathcal{F}_3$ in the case of dense coding, Sec. V B, in that at t_2 (and t_3) the b,c pair is with probability 1/2 in the state $++$ and with probability 1/2 in $--$. The action of the controlled-NOT gate acting just after t_3 is to leave the a qubit ($+$ or $-$) unchanged if $b=+$, or flip it from $+$ to $-$ or vice versa if $b=-$; in either case b remains unchanged. Consequently, at t_4 the information $a=+$ or $-$ originally present at t_0 is no longer present in a by itself, but in a correlation between a and b, or, equivalently, between a and c. The situation is very similar to the classical analogy discussed above in Sec. VII B. The Hadamard gate changes a from $+$ to 0 or $-$ to 1 between t_4 and t_5, so if we use the S_z basis for a at t_5 and later times the initial information about a at t_0 continues to be present in the correlations between a and c. The next controlled-NOT gate in Fig. 2 changes the b qubit to $+$, so it no longer contains any information, and the final controlled-Z transfers the information in the correlation between a and c to c, where it is located at t_8.

D. Teleportation

As noted in Sec. VII A, in the usual teleportation scenario Alice measures qubits a and b in Fig. 2 in the S_z or computational basis at t_5 and communicates the results to Bob over a classical channel. The same result can be obtained, following the proposal in [5] but using somewhat different language, by employing a structure of frameworks in which qubits a and b are in one of the S_z eigenstates for $t\geqslant t_5$, but (in contrast to the first example in Sec. VII C) c is not. Instead, at t_5 one uses an orthonormal basis

$$|00p_\lambda^m\rangle, \quad |01q_\lambda^m\rangle, \quad |10r_\lambda^m\rangle, \quad |00s_\lambda^m\rangle, \tag{26}$$

for the system of three qubits, where m is 0 or 1, $|p_\lambda^m\rangle$ is defined in (14), and the states $|q_\lambda^m\rangle$, $|r_\lambda^m\rangle$, $|s_\lambda^m\rangle$ are obtained from $|p_\lambda^m\rangle$ by applying the unitary transformations X, Z, and XZ, respectively. This basis consists of product states, but the c states are ''contextual'' or ''dependent'' in the notation of [8], Chap. 14. (The term ''nonlocal'' employed in [26] is, in our opinion, somewhat misleading.)

If we use the λ basis for a at t_0, the basis states in (26) at t_5, and the images of the states in (26) under unitary time development at later times, the result is a framework in which the initial information, $m=0$ versus $m=1$ in the λ basis at t_0, is contained in correlations between all three qubits at time t_5 and t_6, in correlations between a and c at t_7, and in qubit c at t_8. That all three qubits are needed at t_5 does not contradict the fact, Sec. VII B, that for each λ there is an alternative framework in which the information is contained in qubits a and b; it is merely one more example of the dependence of the location of information upon the choice of framework. In the same way, the location of a quantum channel depends upon the structure employed for describing it: If one uses the structure under discussion, the collection of frameworks that employ (26) at t_5, then at t_5 the channel is located in or expressed as correlations among all three qubits, and not in any smaller subset. Indeed, if one uses the S_z basis for a and b at t_5 there is *no* information whatsoever in these two qubits about the initial state for *any* choice of λ; their values are completely random, in accord with the usual statement that Alice's measurements tell her nothing about the nature of the state that is being teleported.

Note that by assuming that the two qubits a and b are sent from Alice to Bob through a quantum channel, as suggested by Fig. 2, but at the same time requiring that they be described in the S_z basis for all frameworks in the structure, one arrives at a situation whose end result is just the same as

if Alice had measured a and b in the S_z basis and the results had been sent to Bob over a classical channel in order to actuate the final two gates acting on c. The circuit in Fig. 2 is not an improvement over the original teleportation scheme in a commercial sense, for one of the motivations for the latter was to remove any need for a quantum channel between Alice and Bob once the initial entangled state has been established, i.e., at any time after t_3. However, both ''measurement'' and ''classical communication'' are somewhat complicated ideas, and replacing them with alternative notions which apply to a simple quantum circuit without the need to invoke external apparatus might lead to conceptual simplifications in other applications of quantum information theory.

VIII. CONCLUSION

This paper proposes a way in which to view, or analyze, or think about quantum information, and, as indicated towards the end of Sec. I, it is convenient to summarize it in terms of two theses, one of which seems well founded, though it may in the end turn out to be defective, while the other is more speculative. The first and less controversial thesis is that a framework or consistent family, Sec. IV, provides a natural means of extending classical information theory, both its formalism and the associated intuition, into the quantum domain. This is because a framework provides a consistent probabilistic description of a quantum system in which probabilities have all their usual properties (they are positive, sum to 1, etc.) and their usual intuitive interpretation. The only difference is that now they apply to quantum properties described by a quantum Hilbert space. There is no other consistent way of embedding standard probability theory in standard quantum mechanics (using a Hilbert space without hidden variables) known at the present time, aside from the well-known approach based on measurements, whose inadequacy for the purposes of this paper was pointed out in Sec. IV A. In any case, all the probabilities for measurement outcomes which one can obtain through a valid application of the (not always very clear) rules in the textbooks are also consequences of the correct application of histories methods using frameworks, so the former approach is subsumed under the latter.

The example of dense coding, Sec. V, shows that classical ideas about information of the sort discussed in Sec. II can be translated in a fairly natural way into the quantum domain, and the results are physically sensible as long as one adheres to the single framework rule, something which is in any case necessary for a consistent interpretation of quantum mechanics; see the discussion in Sec. 4.6 of [8]. That dense coding is able to transport two bits of information has nothing to do with mysterious long-range influences, which are in any case absent from quantum theory when properly formulated [27], or with information traveling backwards in time, and it is fully consistent with the condition on the maximum capacity of a quantum carrier of information (no more than one bit per qubit) stated in Sec. IV B. Its peculiarity arises from the fact that the tensor product of two quantum Hilbert spaces provides a very different description of a physical system than does the Kronecker product of two classical phase spaces. The analysis of the teleportation circuit in Sec. VII using a variety of frameworks provides additional examples of how this approach can answer questions about the physical location of information and its flow from one region in space to another. In addition, it is possible to argue that many of the ideas in the published literature having to do with ''classical'' information in the quantum context, such as the classical capacity of a quantum channel, have a quite natural formulation in terms of frameworks.

Further support for the first thesis comes from the following observation. Quantum physicists (with some notable exceptions) believe that the macroscopic, classical world is also governed by quantum principles at a fundamental level, so that classical mechanics is a particular limiting case of quantum mechanics; see [28] for one expression of this widespread faith. The histories approach provides a program, not yet complete but very plausible [29], for understanding classical physics in precisely this way, by the use of quasiclassical frameworks. The first thesis states, in essence, that quantum information theory consists in applying the ideas of classical information theory to quantum processes described by a single (quasiclassical) framework. We can then ask what the implications of quantum information theory, as defined in this way, are for the classical or macroscopic world. The answer is quite clear: because the classical world can be described quantum mechanically using a single framework, quantum information theory in the classical context is automatically the same as standard classical information theory. While such a ''correspondence principle'' is obviously not a proof that quantum information can be correctly formulated in the way we are proposing, it does provide some support.

We now come to the second thesis, which is much more speculative. It is that *all* of quantum information theory is, ultimately, just a matter of applying ''classical'' information theory in different frameworks to a quantum system, and paying attention to framework dependence. That is, there is no special form of ''quantum'' information lying outside the purview encompassed by the first thesis. To put it in another way, the second thesis claims that the concept of information in a particular framework, together with a consideration of various collections of frameworks (thus structures, in the sense defined in Sec. VI A), is an adequate tool for formulating the various problems, such as entanglement, cryptography, and the capacity of quantum channels, which nowadays constitute the central concerns of quantum-information theory. Being able to formulate the problems in this way does not mean that they will be easy to solve, for, as anyone working in the field is well aware, there are a host of formidable technical difficulties confronting attempts to do calculations even in systems as simple as two qubits. Nonetheless, being able to formulate the different problems from this perspective, if it is possible, could provide a certain unity or coherence to quantum information theory, something which has been lacking up to now.

The first ''test case'' for the second thesis is the analysis of the teleportation circuit in Sec. VII, carried out using various different frameworks. In essence the circuit represents a particular form of quantum channel, thus capable of transmitting quantum information, whatever that may be. Is there

anything essential missing from our discussion based upon structures consisting of collections of frameworks, each using a different basis for the channel input? Is there any physical question which cannot be addressed in this manner? It is clear that there are various interesting questions that have not yet been addressed from this point of view, such as why it is that one needs two ''classical'' bits in the standard teleportation process, and it may turn out that the frameworks approach lacks adequate concepts to deal with this and similar issues. Only further research will show whether the use of frameworks can provide a reasonably complete understanding of teleportation and of other significant problems of quantum information theory. While the analysis of the teleportation circuit represents an encouraging first step, much remains to be done.

Thus I present the second thesis not as something for which there are strong and compelling arguments, but rather as an idea worth exploring. Even if partly successful, it could prove to be a significant advance in our understanding of quantum systems in information-theoretic terms, and if it fails, this itself could be the source of interesting insights.

ACKNOWLEDGMENTS

The research described here was supported by National Science Foundation Grant No. PHY 99-00755.

APPENDIX: DEUTSCH AND HAYDEN

Deutsch and Hayden in [3] use the following approach to describe quantum information in a collection of n qubits undergoing unitary time evolution, where we use a slightly different notation from theirs. Let σ_j^α be the jth Pauli spin operator for qubit α, where $j=x,y,$ or z, and $\alpha=1,2,\ldots n$, and let

$$\hat{\sigma}_j^\alpha(t)=T(0,t)\sigma_j^\alpha T(t,0), \tag{A1}$$

with T the time development operator in (10), be the corresponding Heisenberg operator at a reference time $t=0$, as a function of the (physical) time $t>0$. Deutsch and Hayden adopt the view that the three operators $\{\hat{\sigma}_j^\alpha\}$ for $j=x,y,z$ constitute a quantum description of qubit α at time t, and the collection of all such operators, for $\alpha=1,2,\ldots n$, provides a complete description of the entire system at time t. Note that the $\{\hat{\sigma}_j^\alpha\}$ are (in general rather complicated) operators on the full 2^n-dimensional Hilbert space. Probabilities are expressed, as is usual in the Heisenberg picture, by using expectation values of operators of the form (A1), products of these for different α, and sums of such products, in a suitable initial state, which is chosen to be $|0\rangle\otimes|0\rangle\cdots|0\rangle$.

The unitary time transformation T may depend on a real parameter θ. For example, during the interval from t_1 to $t_1+\Delta t$ qubit 5 may pass through a unitary gate corresponding to a rotation of the Bloch sphere by an angle θ around some axis. As a consequence, at later times some of the $\hat{\sigma}_j^\alpha$, in particular the $\hat{\sigma}_j^5$, but also Heisenberg operators for other qubits if they have interacted with this one after the time $t_1+\Delta t$, will depend upon θ. *Information about* θ is, by definition, said to be located *in* the qubits whose Heisenberg operators depend upon θ. Even if at time t one or more components of $\hat{\sigma}_j^\alpha(t)$ for a given α depend upon θ, it is possible that the probabilities associated with any measurement carried out on this qubit at time t will be independent of θ; see [3] for examples. In such a case, while the information is present in this qubit, it is *inaccessible*. The same ideas apply to cases in which T depends upon additional parameters ϕ, χ, etc. Information defined in this way has a definite location and flows from place to place through interactions among the qubits: there is never any instantaneous or superluminal transfer between noninteracting subsystems.

When applied to the examples considered in the present paper, the Deutsch and Hayden prescription would say that in the case of dense coding, the two bits of information sent by Alice to Bob are present but inaccessible in the qubit which she sends him, during the interval between t_5 and t_6 in Fig. 1, rather than, as we argued in Sec. V, in the correlations between this qubit and the one already in Bob's possession. In the case of teleportation, the information initially present in the unknown state $|\psi\rangle$ of qubit a in Fig. 2 is in the (classical) two-bit signal sent from Alice to Bob after her measurement, rather than in the correlation between these and the qubit in Bob's possession, as we argued in Sec. VII. There are various other differences between the conclusions reached by Deutsch and Hayden and those in the present paper. One of particular interest, as it involves a purely ''classical'' situation, is the third example in Sec. II, where they would argue that the information which Alice sends to Bob is present (but, once again, inaccessible) in the single slip of paper she sends to him, whereas we argued that it is not present there, but instead in a statistical correlation between this and a different slip of paper.

The differences between the Deutsch and Hayden approach and that found in the present paper reflect two quite distinct approaches to defining ''quantum information.'' The first is based on a notion of information as reflecting causal influences, whereas the second uses statistical correlations. The latter seems closer to the perspective of classical information theory as developed by Shannon and his successors, but of course this does not imply that it is the correct, much less the unique approach to use when dealing with quantum systems. The one point at which the two approaches agree is in affirming that, just as in the case of classical information, quantum information cannot be transmitted instantly between noninteracting systems; for more on this from the perspective of the present paper, see Chaps. 23 and 24 of [8]. However, this agreement is only superficial, since Deutsch and Hayden assert that a quantum description of the state of each individual qubit is possible even when the system as a whole is in an entangled state, in contrast to the discussion of Bell states found in Sec. IV C above. The reader is invited to compare these approaches and make up his own mind about their virtues and vices, preferably after a careful reading of [3], since the summary given above is necessarily very brief, and the original paper contains several helpful examples along with much more detail.

[1] C.H. Bennett and P.W. Shor, IEEE Trans. Inf. Theory **44**, 2724 (1998).
[2] R. Jozsa, in *Introduction to Quantum Computation and Information*, edited by H.-K. Lo, S. Popescu, and T. Spiller (World Scientific, Singapore, 1998), pp. 49–75.
[3] D. Deutsch and P. Hayden, Proc. R. Soc. London, Ser. A **456**, 1759 (2000).
[4] R. F. Werner, in *Quantum Information: An Introduction to Basic Theoretical Concepts and Experiments*, edited by G. Alber (Springer Verlag, Berlin, 2001), pp. 14–57.
[5] G. Brassard, S.L. Braunstein, and R. Cleve, Physica D **120**, 43 (1998).
[6] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
[7] See Sec. 9.4 of [8].
[8] R. B. Griffiths, *Consistent Quantum Theory* (Cambridge University Press, Cambridge, U.K., 2002).
[9] R. Omnès, *Understanding Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1999).
[10] R.B. Griffiths, Phys. Rev. A **57**, 1604 (1998).
[11] R.B. Griffiths, Phys. Rev. A **54**, 2759 (1996).
[12] R. Omnès, *The Interpretation of Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1994).
[13] R. Omnès, Rev. Mod. Phys. **64**, 339 (1992).
[14] M. Gell-Mann and J.B. Hartle, Phys. Rev. D **47**, 3345 (1993).
[15] See Sec. IV of [10] or Sec. 4.6 of [8].
[16] See p. 163 of [12], where a ''consistent logic'' is what we call a ''framework,'' Sec. III of [11], and Sec. 16.1 of [8].
[17] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, U.K., 2000).
[18] See, e.g., Chap. 10 of [8].
[19] See Sec. 15.2 of [8].
[20] Ascribed to B. Schumacher in [21].
[21] C.H. Bennett and S.J. Wiesner, Phys. Rev. Lett. **69**, 2881 (1992).
[22] C. H. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing, Bangalore, India* (IEEE, New York, 1984), pp. 175–179.
[23] See Sec. 11.6 of [8].
[24] This equivalence is most easily understood by noting that ideal measurements always reveal a preexisting state of affairs if one adopts an appropriate framework, see the discussion in Chap. 17 of [8], and by employing for the circuit in Fig. 2 a framework in which the pair a,b is in one of the Bell states at t_3 and in one of the computational basis states at t_5.
[25] Contrary to the claim on p. 58 of [2].
[26] C.H. Bennett *et al.*, Phys. Rev. A **59**, 1070 (1999).
[27] See the discussion in Chaps. 23 and 24 of [8].
[28] F. Wilczek, Phys. Today **53**(6), 11 (2000).
[29] See [14], Chaps. 16–19 of [9], and Chap. 26 of [8].