PHYSICAL REVIEW A **70**, 022312 (2004)

Probabilistic and information-theoretic interpretation of quantum evolutions

J. Oppenheim*
Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, United Kingdom and Racah Institute of Theoretical Physics, Hebrew University of Jerusalem, Givat Ram, Jerusalem 91904, Israel

B. Reznik†
Department of Physics and Astronomy, Beverly and Raymond Sackler Faculty of Exact Sciences, Tel-Aviv University, Tel Aviv 69978, Israel

(Received 23 December 2003; published 19 August 2004)

In quantum mechanics, outcomes of measurements on a state have a probabilistic interpretation while the evolution of the state is treated deterministically. Here we show that one can also treat the evolution as being probabilistic in nature and one can measure which unitary acted. In further analogy to states, one can also choose which basis of unitaries to measure. Likewise, one can give an information-theoretic interpretation to evolutions by defining the entropy of a completely positive map. This entropy gives the rate at which the informational content of the evolution can be compressed. One cannot compress this information and still have the evolution act on an unknown state, but we demonstrate a general scheme to do so probabilistically. This allows one to generalize super-dense coding to the sending of quantum information. One can also define the "interaction-entanglement" of a unitary, and concentrate this entanglement.

DOI: 10.1103/PhysRevA.70.022312 PACS number(s): 03.67.−a, 03.65.Ta, 03.65.Ud

I. INTRODUCTION

An isolated system is represented in quantum mechanics by a state vector that conveys statistical predictions for measurement outcomes and manifests phenomena such as superpositions, and entanglement. In contrast, the temporal evolution law of the state is determined by the unitary operator $U=\exp -iHt/\hbar$ where the Hamiltonian H is dictated either by external classical potentials and/or universal interactions between fields or particles. Therefore, while the state vector manifests the non-deterministic features of quantum mechanics, the temporal evolution law of an isolated system is regarded as fully deterministic. This asymmetry between the properties of states and evolutions is also maintained within the framework of quantum information theory wherein the information is carried by the state alone.

Psychologically, the asymmetry may partly result from thinking of the unitary evolution as being due to some external macroscopic device such as a large magnet, while the state might be a single electron. Thus we think of the state as being quantum and probabilistic in nature, while the unitary evolution is treated deterministically. However, we should remember that the state of the electron is also determined by some macroscopic device (a Stern-Gerlach machine say), and thus the asymmetric treatment between states and unitaries does not arise from such considerations.

In this work we examine the consequence of measurements of the evolution law and suggest that the above restricted view of quantum evolutions can be extended even within the conventional framework of quantum mechanics and quantum information. We find that features such as superposition of unitary evolutions, collapse to a certain evolution and a corresponding probability law, can in fact be attributed in a natural fashion to unitary evolutions, as well as to the more general case of nonunitary evolutions that can be described by completely positive (CP) trace preserving linear maps. Particularly, we show that a measurement of 'which evolution occurred' during a certain time interval 'collapses' the quantum evolution to a particular evolution with a probability given by a simple extension of the ordinary probability law. Our results provide an operational meaning to a formal correspondence between states and operations introduced by Jamiołkowski [1] and Choi [2].

Although the present work is aimed at extending concepts ascribed to states into the domain of evolutions, it also has applications to quantum computation in that we present methods which can be used to monitor the interactions of a quantum device without changing the physical setup of the device. There are also links between the measurement of unitaries, and syndrome detection in error correction codes.

Next, we turn to the question of whether operations have informational content in a manner analogous to quantum states. We find that one can assign a state independent entropy to an arbitrary completely positive (CP) map, and that this entropy gives the rate at which the informational contents of the map can be compressed. Here, the information content refers to the sequence of Kraus operators [3] used to implement the CP map, although other interpretations are possible. This can be considered as the equivalent of Schumacher's noiseless coding theorem for operations. A different interpretation of compression and storage of unitaries was given in Ref. [4] where for a specific known ensemble of phase gates it was shown how to store them efficiently. This can be thought of as storage of a known ensemble of evolutions [5]. Here, our compression rate is ensemble independent and generic, akin to compression of a source emitting quantum states.

*Electronic address: jono@damtp.cam.ac.uk
†Electronic address: reznik@post.tau.ac.il

We invoke a no-go theorem [5,6] for programmable quantum gates, and storage of unitaries to show that one cannot have a compressed evolution act on an unknown state, and still preserve its informational content. This is true even if one only demands approximate fidelity. This is because there are an infinite number of ways one can implement a given CP map (there are in a sense, an infinite number of *evolution ensembles*, which are unknown). We show however, a generic scheme to probabilistically act the evolution on an unknown state.

We then generalize super-dense coding to the sending of quantum information contained in unitaries. This has certain cryptographic implementations which we briefly explore.

Finally, we turn to the notion of entanglement of a unitary. A number of authors have used the formal correspondence between states and operations to investigate the entangling capabilities of unitary operations [7], (e.g., Refs. [4,8–11]). The present framework suggests the notion of *interaction entanglement* of a unitary acting on systems, and we show that this entanglement can be concentrated in a manner analogous to the concentration of states into pure entanglement. We conclude with some remarks on the interpretational issues involved with the measurement of evolutions.

II. A PROBABILISTIC INTERPRETATION OF UNITARIES

Let us first provide an operational meaning to the measurement of unitaries. We consider a system with an N-dimensional Hilbert space whose state evolves in time according to

$$|\psi(t_1)\rangle \rightarrow |\psi(t_2)\rangle = U(t_2 - t_1)|\psi(t_1)\rangle. \tag{1}$$

It is know that for any N there exists a basis of N^2 orthogonal unitary operators [12], where orthogonality is defined with respect to the trace inner product $U \cdot V \equiv \mathrm{tr}\, U^\dagger V$. Thus the unitary time evolution operator can be decomposed with respect to an orthogonal basis U_α,

$$U = \sum_{\alpha=0}^{N^2-1} C_\alpha U_\alpha, \tag{2}$$

where $U_\alpha \cdot U_\beta = N\delta_{\alpha\beta}$ and the complex amplitudes are given by $C_\alpha = (1/N) U_\alpha \cdot U(t_2 - t_1)$. The converse of the above statement is not true. A superposition of unitary operators with arbitrary amplitudes generally does not give rise to a unitary. The operators space contains nonunitary operators which can be also spanned by a unitary basis.

Can the formal expansion (2) be given a general physical interpretation? It has been shown that under certain conditions, a superposition of unitary evolutions that gives rise to another unitary, can be produced by post-selecting an ancillary system that interacts weakly with our system [13]. In the present work we propose another approach. We shall show that for any chosen basis, we can measure which unitary evolution U_α the system evolved under. The outcome of such a measurement has probability $\mathrm{Prob}(U_\alpha) = |C_\alpha|^2$. More formally, we have the following.

(1) Observables and eigenvalues: To each orthogonal basis of unitary operators, we can find an observable $A(t_2, t_1)$ that assigns to each unitary U_α a distinct real eigenvalue λ_α through the eigenvalue equation,

$$T{:}A(t_2; t_1) U_\alpha = \lambda_\alpha U_\alpha. \tag{3}$$

The operator $A(t_2, t_1)$ describes temporal correlations. It is constructed as a linear combination of bilinear products of operators at each of the two instances t_2 and t_1. The symbol T: denotes temporal ordering. For instance if $A(t_2; t_1) = \beta A(t_2) B(t_1) + \gamma C(t_2) D(t_1)$ then $T{:}A(t_2; t_1) U_\alpha(t_2 - t_1) = \beta A U_\alpha B + \gamma C U_\alpha D$. Since $\lambda_\mu = \beta A U_\alpha B U_\alpha^\dagger + \gamma C U_\alpha D U_\alpha^\dagger$ is a constant, the operator $A(t_2, t_1)$ describes constant of motion with respect to each of the basis elements. The eigenvalues are thus state independent.

(2) Probability law and measurements: The outcome of a measurement of $A(t_2, t_1)$ is one of eigenvalues λ_α with a probability given by

$$\mathrm{Prob}(\lambda_\alpha) = |C_\alpha|^2. \tag{4}$$

(3) Reduction of U (collapse): A measurement with an outcome λ_α leads to a collapse (effectively or truly depending to the readers preferred interpretation) of the superposition (2) according to

$$U(t_2 - t_1)|\psi\rangle \rightarrow U_\alpha|\psi\rangle. \tag{5}$$

We interpret (1)–(3) as specifying criteria for a measurement that detects which particular transformation U_α in the superposition (2) has been realized on the system with *a priori* probability $\mathrm{Prob}(U_\alpha) = |C_\alpha|^2$.

It should be emphasized that (1)–(3) are independent of the initial state of the system and hence can be interpreted as a measurement of a term in the superposition (2). The initial state of the system is here arbitrary, hence includes the case of a unitary acting (locally) on a part of an entangled state. As a consequence, the present measurement of the unitary transformation *does not reduce* the entanglement of the system.

Although here we will show an operational correspondence between the measurement of states and the measurement of unitaries, it must be emphasized that there are important differences. One interesting result is that one can perfectly distinguish between two unitaries which are not orthogonal [14] if one can act each unitary many times on a state. In contrast, for any finite number of copies of two nonorthogonal states, there is a probability of error.

We now proceed to prove the above three statements. To begin with, we consider some simple properties of a general given basis of orthogonal unitary operators $\{U_\alpha\}$ of Eq. (2). Clearly the set $\{\mathrm{I}, U_i'; i = 1, \ldots, d-1\}$ where $U_i' = U_0^\dagger U_i$ is also orthogonal under partial trace. Since U_i are traceless orthogonal operators (they are orthogonal to I), all sets of unitary orthogonal basis can be expressed as a product of an arbitrary fixed unitary U_0 with some traceless unitary orthogonal basis.

We explicitly consider the $N = 2$ case—generalizing our results to higher dimensional Hilbert spaces (including the infinite dimensional case) is straightforward and described in the Appendix. The general structure of the basis is given to be

$$U_\alpha = U_0 \sigma_\alpha, \tag{6}$$

where $\sigma_\alpha = (1, \sigma_i)$ with $i = x, y, z$ the Pauli matrices.

The observables corresponding to the measurement of U_α can then be chosen as

$$A_i(t_2, t_1) = [U_0 \sigma_i U_0^\dagger]_{t_2} [\sigma_i]_{t_1}. \tag{7}$$

Replacing into (3),

$$T{:}A_i(t_2, t_1) U_\alpha = [U_0 \sigma_i U_0^\dagger] U_\alpha [\sigma_i] = \lambda_{i\alpha} U_\alpha, \tag{8}$$

and using (6), we get

$$\lambda_{i\alpha} = [U_0 \sigma_i U_0^\dagger] U_\alpha \sigma_i U_\alpha^\dagger \tag{9}$$

$$= U_0 \sigma_i U_0^\dagger U_0 \sigma_\alpha \sigma_i \sigma_\alpha U_0^\dagger \tag{10}$$

$$= \pm 1. \tag{11}$$

Since we need to resolve between four basis elements, it is sufficient to consider a pair of operators, say, $A_i(t_2; t_1)$ with $i = z, x$.

Next demonstrate (2) by explicit construction of a measurement. One possibility is to have the unitary act on half of a maximally entangled state. Each orthogonal unitary in a basis of unitaries would then produce an orthogonal maximally entangled state and one could then perform a measurement on the state to determine which unitary acted. This has the disadvantage that one cannot use this method for an evolution acting on a particular physical system in an unknown state. We therefore propose to observe the operator $A(t_2, t_1)$ by coupling twice with the system in a manner which preserves the state. A method for measuring sums of operators as $\sigma(t_2) + \sigma(t_1)$ has been suggested [15] and used to demonstrate teleportation [16]. We employ a similar method using a pair of ancillary two-level particles taken initially in the state $(|0\rangle + |1\rangle)(\tilde{0}\rangle + |\tilde{1}\rangle)/2$. We assume a vanishing free Hamiltonian for the ancillary particles.

The ancilla and the system then interact twice, first at $t = t_1$ and then at $t = t_2$, and then the ancilla is measured. To specify the interaction between the system and ancilla we define the controlled Pauli

$$V_i = |0\rangle\langle 0| + |1\rangle\langle 1| \sigma_i, \tag{12}$$

where σ_i acts on the system, and similarly we denote by $\tilde{V}_i$ the same interaction between the system and the second ancilla. We further assume that the interactions are nearly impulsive: the duration Δt required to apply $V_i \tilde{V}_i$ is much shorter than $t_2 - t_1$, hence the correction due to the free evolution can be neglected while we apply the interactions.

We now apply the following sequence:

$$(U_0 \tilde{V}_x V_z U_0^\dagger) U(t_2 - t_1) (\tilde{V}_x V_z). \tag{13}$$

The measurement interaction acts twice at $t = t_1$ and $t = t_2$, while at intermediate times the system evolves freely. The resulting total state becomes

$$\frac{1}{2} \sum_\alpha C_\alpha (|\tilde{0}\rangle + \lambda_{x\alpha} |\tilde{1}\rangle)(|0\rangle + \lambda_{z\alpha} |1\rangle U_\alpha) |\psi\rangle. \tag{14}$$

Finally, using the notation $|\alpha\rangle = \{\uparrow_z \uparrow_z, \uparrow_z \downarrow_z, \downarrow_z \downarrow_z, \downarrow_z \uparrow_z\}$ where $\uparrow_z, \downarrow_z = (|0\rangle \pm |1\rangle)/\sqrt{2}$, we can perform a projective measurement in the $|\alpha\rangle$ basis on the ancilla. The final total state of the system and the two spins and the effect of the measurement can be expressed as

$$\sum_{\alpha=0}^{3} C_\alpha |\alpha\rangle U_\alpha |\psi\rangle \to U_{\alpha_0} |\psi\rangle, \tag{15}$$

thus demonstrating the notion of collapse (3) to α_0, a particular α. One could also interpret the above as instead being a collapse induced by the interaction.

We have used here the standard probability interpretation with respect to a measurement of the final ancillary basis $|\alpha\rangle$. Since the probability to find α_0 is given by $|C_{\alpha_0}|^2$, this demonstrates (3) and (4) for the present two-dimensional case. One can verify that the above procedure effectively moves the information contained in the state onto the ancilla, while having the unitary act on half a maximally entangled state. The information of the state (with the action of the unitary) is then transferred back from the ancilla to the original system. However, the physical particle that is the system is not actually swapped, allowing one to use such a measurement without changing the particular system. For example, one could use this to detect the noise in an ion-trap quantum computer while still preserving the information of the state and the setup of the experiment. The measurement procedure gives a generic way to transfer the state of a system onto another system without performing a physical swap.

An important point is that the measurement of which unitary is independent of the state that the unitary acts on. This allows one to distinguish between unitaries which when acting on certain initial states would not lead to orthogonal final states.

To exemplify our result, consider the evolution of a spin in a magnetic field with $U = \exp(-iB\sigma_z t) = \cos(Bt) 1 - i \sin(Bt) \sigma_z$. If we select to measure in the basis $(U, U\sigma_x, U\sigma_y, U\sigma_z)$, we will find $\mathrm{Prob}(U) = 1$ and $\mathrm{Prob}(U\sigma_i) = 0$. Therefore, in this case we verified with certainty that the evolution is $U(t)$ without causing any disturbance. On the other hand, if we choose to measure in the basis $(1, \sigma_i)$, we will reduce the evolution to $1|\Psi\rangle$ with probability $\cos^2(Bt)$, or to $\sigma_z |\psi\rangle$ with probability $\sin^2(Bt)$. More generally, in a d-dimensional space, we can distinguish with certainty between d^2 orthogonal unitary operators.

What is more, we are able to distinguish between unitaries which do not themselves commute. This is because each element of the basis gives orthogonal outcomes on maximally entangled states. There is however an uncertainty principle between different possible measurements of which unitary given by the uncertainty between two two-time operators $A(t_2, t_1)$ and $A'(t_2, t_1)$. This uncertainty can be compactly expressed as an entropic relation [17,18]. Consider two measurements of a basis $\{U_\alpha\}$ and $\{U_{\alpha'}\}$, which can be described in terms of the orthogonal traceless set $\{\mathrm{I}, U_i\}$ and a unitary

U_0 and U_0' for the α and α' basis, respectively. Then, for all outcomes of measurements with probability p_α and $p_{al'}$ one finds the uncertainty relation

$$H_\alpha + H_{\alpha'} \geqslant -2 \ln \mathrm{Tr}(U_0^\dagger U_0'), \tag{16}$$

with H_α and H_α the Shannon entropy of the measurement outcomes, i.e., $H_\alpha = -p_\alpha \log p_\alpha$. The relationship has the appeal that it depends only on the chosen basis of unitaries, and not on the particular unitary being measured or the state that the unitary acts on. It can be derived by noting that the measurement procedure for two different basis is identical until t_2 when one essentially makes two different projective measurements in some maximally entangled basis given by $\mathrm{I} \otimes U_\alpha \psi^+$ or $\mathrm{I} \otimes U_{\alpha'} \psi^+$. Then, from [19], one has the uncertainty relation

$$H_\alpha + H_{\alpha'} \geqslant -2lnc \tag{17}$$

with $c = \max_{\alpha,\alpha'} |\langle \psi^+ \mathrm{I} \otimes U_\alpha^\dagger | \mathrm{I} \otimes U_{\alpha'} \psi^+ \rangle|$. The relationship Eq. (16) then follows.

Finally, we comment that the above measurement procedure can be extended to nonunitary orthogonal operators, which may be also used as a basis. Such a nonunitary basis can be obtained by the transformation $A_\mu = \Sigma_\nu K_{\mu\nu} U_\nu$, where K is a $N^2 \times N^2$ dimensional unitary matrix. The operators A_μ are generally not unitary, but are orthogonal with respect to the trace inner product. Thus, we can distinguish between the elements A_μ using the procedure used above.

III. AN INFORMATION THEORETIC INTERPRETATION OF EVOLUTION

Having shown that the correspondence between unitaries and states has an operational meaning in terms of probability amplitudes, we now turn to the question of whether there is an information theoretic interpretation to unitary operations. An informational interpretation of quantum states was given by Schumacher's noiseless coding theorem [20] (cf. Refs. [21,22]). We will now see that a similar interpretation can be given to unitary operations. Instead of considering a pure unitary, we consider an arbitrary completely positive (CP) map $\mathcal{E}(\rho)$. We will see that one can define an entropy for the CP map which only depends on the map, and not on how it is implemented, nor on what state it acts, and that this has an interpretation of the rate at which the informational contents of the map can be compressed. It is also equal to the maximum classical information which the map can transfer. The entropy production that a CP map produces on particular states was considered in Ref. [23]. We will further prove two theorems showing that while the information can be stored and compressed, it is impossible to later act it on an unknown state, or even a known state chosen after the information has been stored.

We start by showing that the interpretation of unitaries described in the preceding section, can be extended to other positive operators. Namely, we can expand an arbitrary CP map in terms of Kraus operators M_i [3],

$$\mathcal{E}(\rho) = \sum_i M_i \rho M_i^\dagger. \tag{18}$$

One usually thinks of the Kraus representation as being a formal representation of a CP map. Here, the aim is to find a physical and informational interpretation. The operator-sum decomposition is not unique, but it can be shown [3,24] that all other decompositions have Kraus operators N_j related by a unitary transformation $N_j = U_{ij} M_i$. The operator-sum decomposition may therefore be thought of as being analogous to a density matrix. In particular, it can be shown [23] that for a given state ρ, there exists a diagonal representation, such that

$$\mathrm{tr}\, M_\mu \rho M_\nu^\dagger = 0 \quad \text{for } \mu \neq \nu. \tag{19}$$

If ρ is taken to be the maximally entangled state ψ^+, with the Kraus operators acting on half of it, then one sees that the M_μ are orthogonal under the trace inner product as with the orthogonal unitaries (or the nonunitary set A_μ) considered in the preceding section. One therefore has that

$$|\tilde{\psi}_\mu\rangle = M_\mu |\psi^+\rangle \tag{20}$$

are orthogonal states (unnormalized). The normalized states we call $|\psi_\mu\rangle$. After the CP map has acted on half the ψ^+, we are left with a density matrix given by

$$W = \sum p_\mu |\psi_\mu\rangle\langle\psi_\mu|. \tag{21}$$

One way to think of how the CP map arises is to consider a unitary which acts not only on ρ, but also on the system plus an ancilla $|0_C\rangle$ (so-called Stinespring dilation), namely,

$$\mathcal{E}(\rho) = \mathrm{tr}_C\, U(\rho |0_C\rangle\langle 0_C|). \tag{22}$$

After considering such a global unitary, one can take the ancilla to be with a third party (who we will call Charlie), who is considered to be the source C of the CP map.

We then define the *entropy of a CP map* as

$$S_\mathcal{E} = -\sum p_\mu \ln(p_\mu) \tag{23}$$

and show that it gives the rate at which one can noiselessly compress the informational content of the CP map. By information, we mean something analogous to the informational content of a state under compression. In the case of states, the compression is done without knowing the ensemble, and after decompression, one can verify that one faithfully obtained some series of states by having the source read out each state that was sent. One then performs a measurement on the decompressed states to verify fidelity.

Here, in analogy with ensembles of states, we have choices of the Kraus representation M_i. We can therefore verify that all the information of the CP map has been faithfully stored under the following test: Charlie performs a measurement on the ancilla in an arbitrary basis. We will see that choosing the basis is the equivalent of choosing some Kraus representation (like choosing the ensemble). Charlie's result is in one to one correspondence with a particular M_i,

and we can verify that indeed this M_i acted on our state. We thus have a correspondence between the informational content of states, and that of operations.

To see this we consider a measurement on the ancilla C in the basis $|i_C\rangle$ after the unitary U of Eq. (22) has acted on the ancilla and ψ^+. This then selects the M_i via

$$M_i|\psi^+\rangle = \langle i_C|U|\psi^+\rangle \otimes |0_C\rangle. \tag{24}$$

Therefore if given the value of i from the source, one can verify that M_i did indeed act. Furthermore, for an ensemble which is made up of orthogonal M_i, we can distinguish them without being given the value of i. Note that the particular form of the M_i is dependent on the state acted upon, although the CP map itself is state independent.

That $S_\mathcal{E}$ qubits are necessary and sufficient to store this information is straightforward. The rate can be achieved for large n, simply by having the source perform each unitary on a maximally entangled state ψ^+, creating the density matrix given by

$$\varrho = \sum p_\mu |\psi_\mu\rangle\langle\psi_\mu|. \tag{25}$$

From Shannon's noiseless coding theorem, the state with density matrix ϱ can be compressed at a rate of $S_\mathcal{E}+\epsilon$ with ϵ as small as desired in the limit of large n. The encoding clearly preserves the informational content as described above.

That this rate is optimal, can be seen from the fact that the encoding must work for all ensembles, and in particular we could choose the ensemble to be the set of orthogonal operators M_μ. A better compression rate would then imply a violation of the Holevo bound.

A particular example of the above scheme are CP maps which correspond to unitaries applied probabilistically. We imagine that a sequence of unitaries are performed by a source C, and that while we do not know what unitaries are being performed, nor from what ensemble the unitaries are being drawn from, we do know the CP map that the source performs. Again, this is in analogy with knowing the density matrix of a source which emits states. That is, one images a sequence of unitaries performed on the state, where the unitaries are chosen from some *unknown* ensemble $\zeta=\{p_i,U_i\}$ (the U_i need not be orthogonal), and we wish to compress a particular sequence X of n draws from this ensemble. All we are given is a Kraus representation of the CP map. Using the method above, the sequence of U_i can be compressed at a rate $S_\mathcal{E}$, and one can indeed verify whether any particular sequence X of unitaries was performed.

One might hope that the information concerning the sequence of positive operators could be encoded and decoded in such a way that a recipient can act the map on an unknown state given after the encoding. We will see that this is impossible for an arbitrary ensemble even if only approximate fidelity is demanded.

This result is easily extended to the case of Kraus operators. Consider an unknown sequence of Kraus operators $X=M_1M_2M_3\cdots M_n$ and similarly X', and a distance measure $D(X,X')=\mathrm{tr}(|X-X'|)$. A given protocol aims to act X on an unknown set of states ψ, generating the state Ψ, call the error rate of a given protocol $\epsilon=|\langle\psi|M_1|\psi\rangle M_2|\psi\rangle M_3|\psi\rangle\cdots|^2$.

Theorem 1: Given X,X' drawn from any operator-sum decomposition of the CP map $\mathcal{E}(\cdot)$, and any encoding $A(\mathcal{E}(\cdot))$ which maps sequences X,X' to states τ_x,τ_x', and decoding algorithm $B(A(\mathcal{E}(\cdot)),\psi)$ which maps τ_x to $|\Psi\rangle$ close to $M_1|\psi\rangle M_2|\psi\rangle M_3|\psi\rangle\ldots$ with error rate ϵ. Then if $|\psi\rangle$ is an arbitrary unknown state chosen after encoding, $\mathrm{tr}|\tau_x\tau_x'| \leqslant O(\sqrt{\epsilon})/D(X,X')$.

We can invoke a no-go theorem for programmable unitary gates [6], extended to the approximate case in Ref. [5]. We refer the reader to Ref. [5] for the full proof of the used result, and just give the no-go theorem in the exact case, using the fact that the encoding must be unitary. The decoding takes as input, a state $|\psi\rangle^{\otimes n}$, and the encoding of the map realization τ_x. Let us first take τ_x to be a pure state $|x\rangle$ (our proof will extend to any mixed state τ_x by the linearity of quantum mechanics). The decoding then takes this input and produces the sequence $|Y\rangle=M_1|\psi\rangle M_2|\psi\rangle M_3|\psi\rangle\cdots$ and some ancilla $|\chi_x\rangle$. The ancilla cannot depend on ψ for coherence to be preserved. We can imagine the encoding being performed on another sequence X', encoded in $|x'\rangle$, and producing a sequence $|Y'\rangle=M_1'|\psi\rangle M_2'|\psi\rangle M_3'|\psi\rangle\cdots$, and ancilla $|\chi_{x'}\rangle$. Then, since the decoding must be unitary it must preserve the inner product of any two inputs,

$$\langle x|x'\rangle = \langle\chi_x|\chi_{x'}\rangle\langle Y|Y'\rangle. \tag{26}$$

Since neither $\langle x|x'\rangle$ nor $\langle\chi_x|\chi_{x'}\rangle$ can depend on ψ it follows that either $\langle x|x'\rangle=\langle\chi_x|\chi_{x'}\rangle=0$ or $\langle Y|Y'\rangle$ cannot depend on ψ. The latter can only occur if $X=X'$, therefore, if the encoding/decoding is to work for different possible inputs we require $\langle x|x'\rangle=0$. That is, an orthogonal state must be chosen for each possible sequence, and the size of the encoded state must then be as large as the number of possible sequences. Since there are an arbitrarily large number of possible ensembles which implement a given CP map, it follows that the size of the encoded state must be infinite. In essence, the size of the program grows with the size of the ensemble.

It is not clear if one can do better if the state is known to the decoder.

It is perhaps amusing that there is an infinite discontinuity which occurs if all M_i are identical and perfect fidelity is required. One can imagine a CP map which can be decomposed into two orthogonal unitaries U_1 and U_2 and that one is applied with probability $1-\epsilon$, and the other with probability ϵ. There is an infinite discontinuity in that the number of possible Kraus representations goes from infinity to one. The same discontinuity exists for ensembles of density matrices. There is therefore potentially something special about pure states and pure unitaries. This discontinuity only exists if one demands perfect fidelity of the decoding, therefore it is unclear what the interpretation of this observation is. The above has the flavor of a phase transition (cf. Refs. [25,26]).

One can now ask whether one can perhaps act the compressed evolution on an unknown state probabilistically. Indeed, for the case of a stored phase gate of the form $U(\alpha)=\exp(i\alpha\sigma_z)$ one can act the stored gate on an unknown state

with probability 1/2 [5]. We now generalize this to arbitrary unitaries and Kraus operators.

Consider an unknown state ψ and evolution M_i stored in state ψ_i. We then perform the unitary

$$V = \sum_\mu P_\mu M_\mu, \tag{27}$$

where P_μ are projectors onto the orthogonal states ψ_μ which are eigenkets of ϱ defined via Eq. (25). The stored evolution can be expanded in terms of the orthogonal set of Kraus operators as

$$M_i = \sum c_{i\mu} M_\mu. \tag{28}$$

We then cause V to act on the stored evolution and the unknown state

$$V|\psi_i\rangle \otimes |\psi\rangle = \sum c_{i\mu} \psi_\mu \otimes |M_\mu \psi\rangle. \tag{29}$$

We then measure the state which was storing the unitary, in a basic complementary to ψ_μ. For example, we can measure using projectors onto $\psi_{\mu'}$ with $\langle \psi_{\mu'} | \psi_\mu \rangle = \pm 1/\sqrt{d}$. Then, with probability $1/d$ we will have succeeded in performing the correct Kraus operator.

IV. SUPER-DENSE CODING OF UNITARIES

The preceding section therefore gives an informational interpretation of evolutions. In fact, one can regard the entropy of Eq. (23) as representing the maximum amount of information that the evolution can transfer from an environment or source to a state. This leads to a natural generalization of super-dense coding where the information that is conveyed is not classical bits, but rather, pure quantum information.

One can imagine that two parties (Alice and Bob) share a maximally entangled state, and that Alice has access to a source C of random unitaries which acts on her half of the singlet. Alternatively, Alice might apply unitaries conditional on quantum states, or might apply the unitaries herself according to some classical probability distribution. The action of the unitaries will produce a sequence of maximally entangled states shared between Alice and Bob. After Alice sends her half of the singlet to Bob, he will obtain all the quantum information about the unitary. Since the basis of qubit unitaries is 2^n larger than the basis for qubit states, this can therefore be viewed as a "quantum" version of the classical communication sent in super-dense coding. In the case of super-dense coding, Alice chooses from four orthogonal unitaries and applies them to her half singlet and sends. Here, one allows arbitrary superpositions of the orthogonal unitaries to be applied. What is more, the information that is sent can be sent blindly. Alice need not know which unitaries are being applied by the source C. If she first tried to know which unitaries were being applied by the source, she would of course, destroy the quantum state.

An alternative generalization of super-dense coding has been independently proposed in Ref. [27]. There, it was shown that in large dimensions, using singlets and shared randomness, Alice could send known quantum states using only half as many qubits.

As with super-dense coding, the sending of the arbitrary unitary is cryptographically secure, in that an eavesdropper, located between Alice and Bob, obtains no information about which unitary was applied (neither can Alice learn which unitary was applied, as long as Bob holds the other half of the used singlet). One may therefore regard this as a one way private quantum channel [28,29] which uses a resource of one ebit per 2 qubits of sent information rather than 2 cbits for each qubit (although see the key-recycling results of Ref. [30,31]), or 2 ebits [31] per qubit.

V. ENTANGLEMENT AND CONCENTRATION OF UNITARIES

Does the notion of entanglement extend to the case of evolutions? Consider a unitary interaction that acts on a pair of systems. Clearly, the combined evolution operator can be expanded in terms of the unitary basis operators of each system in the general form

$$U^{(\mathrm{I,II})} = \sum C_{\mu\nu} U_\mu^{(\mathrm{I})} \otimes V_\nu^{(\mathrm{II})}, \tag{30}$$

where $U_\mu^{(\mathrm{I})}$ and $V_\nu^{(\mathrm{I})}$ are the local orthogonal unitary basis. Likewise the familiar entanglement bipartite correlations are recovered for interactions.

In the sense of a passive transformation we can re-express the general state by performing the transformation $A_\mu = \Sigma_\alpha K_{\mu\alpha} U_\alpha$ and $B_\nu = \Sigma_\beta Y_{\nu\beta} V_\beta$, such that $KCY = D$ is diagonalized in the new orthogonal basis with eigenvalues d_μ. Hence a Schmidt form can be written also for unitary interactions

$$\tilde{U}^{(\mathrm{I,II})} = \sum d_\mu A_\mu^{(\mathrm{I})} \otimes \tilde{B}_\mu^{(\mathrm{II})}. \tag{31}$$

The operators $A_\mu^{(\mathrm{I})}$ and $B_\mu^{(\mathrm{II})}$ in the above decomposition are generally not unitary, however they maintain orthogonality under the trace inner product. Hence, we can apply the same procedure, described in Sec. II, to measure which operator has acted on each side of the bipartite system. The probability to find a certain operator A_μ (or B_μ if the measurement takes place at side II) is then given by $|D_\mu|^2$. As consequence of (3) a measurement of say system I, will lead to a collapse of the sum to a single term in analogy with pure state entanglement of states. There is then a one-to-one correlation between the results of the measurement of which operator has acted on system I and II.

We can now quantify the entanglement of the interaction by computing the entropy of the probabilities, $-\Sigma |d_\mu|^2 \log |d_\mu|^2$, in this diagonal basis. To justify this choice we demonstrate a concentration procedure for n identical nonmaximal bipartite interactions. We emphasize that we now consider a concentration process that is independent of the nature of the state $\rho(\mathrm{I,II})$, on which the unitary $U^{(\mathrm{I,II})}$ acts.

Suppose that we operate n times the same bipartite interaction

$$[\alpha I^{(\mathrm{I})} \otimes I^{(\mathrm{II})} + \beta \sigma_x^{(\mathrm{I})} \otimes \sigma_x^{(\mathrm{II})}]^{\otimes n}. \tag{32}$$

We would like now to concentrate this "nonmaximal interaction" to a sum of terms with equal coefficients. Recalling that in the state concentration scheme one employs a collective measurement of the operator $J_z^{(\mathrm{I})}=\Sigma_i\sigma_{zi}^{\mathrm{I}}$, we shall now consider a measurement of the temporal collective correlation $\Delta J_z^{(\mathrm{I})}(t_2,t_1)=J_z^{(\mathrm{I})}(t_2)-J_z^{(\mathrm{I})}(t_1)$ (more generally, when we have a large number of terms one has to measure more temporal correlations). The equation $T:J(t_2,t_1)U_i=\lambda U_i$, has solutions with eigenvalues $\lambda=(-n,-n+2,\ldots,n)$. The relevant eigenoperators corresponding to U_i have the structure of a sum of terms, where each of the terms is given by products of unit operators and Pauli operators. The total number of Pauli operators is identical in all terms and determined by the eigenvalue λ. The coefficients of the terms need not be identical hence U above is generally degenerate. Nevertheless, in our particular case, a straightforward calculation shows that a measurement of the operator $\Delta J_z(t_2,t_1)$, that may be performed on subsystem I or II, collapses (32) to the operator

$$C_U=[(I_1^{(\mathrm{I})}\cdots I_k^{(\mathrm{I})}\sigma_{k+1}(I)\cdots\sigma_n^{(\mathrm{I})})(I_1^{(\mathrm{II})}\cdots\sigma_n^{(\mathrm{II})})+\cdots]. \tag{33}$$

Notice that the terms in the square brackets above are now all equally weighted, and their number is determined by the measurement outcome. The probability to collapse into a particular value of operator is given by $\alpha^{2k}\beta^{2(N-k)}$. Therefore, in complete analogy to the case of pure state concentration, the expected number of equally weighted terms in C_U, peaks in the limit of large n around $2^{nS(d_\mu)}$, where $S(d_\mu)$ is the Shannon entropy. Notice that in general the operator C_U is not unitary. Nevertheless, its entangling capability power is equivalent to n controlled-not interactions: it can convert n nonentangled pairs into a block with 2^{nS} equally weighted terms which is maximally entangled. However, unlike the case of state concentration, C_U cannot be further factored by means of local operations to a product of bipartite maximally entangled unitaries.

The above result suggests a notion of bipartite *interaction entanglement*, S_U, which is a straightforward extension of ordinary entanglement,

$$S_U=-\sum_\mu |d_\mu|^2\ln|d_\mu|^2. \tag{34}$$

This definition is in complete harmony with the entropy defined previously for a CP map. Therefore, given a bipartite unitary interaction, the entanglement entropy of the interaction corresponds locally to the entanglement of the locally generated CP map. This can be seen by noticing that the operators A_μ (B_μ) in the Schmidt decomposition (31) are in fact then the same Kraus operators that appear in the sum representation of the CP map which act on system I (II).

The analog of a maximal entangled state is in our case given by $(1/\sqrt{2})(I\otimes I+i\sigma_x\otimes\sigma_x)$, which is equivalent to a controlled-not (up to additional local rotations). We can now compare the proposed notion of interaction-entanglement with that of entanglement capability of an interaction [32]. The latter is defined by maximizing the amount of state-entanglement that an interaction produces by acting on a particularly chosen state. Clearly the two notions differ. Interaction entanglement does not depend on the nature of the initial state, while the entanglement capability clearly does. Furthermore, in general the numerical value of entanglement capability is larger than the interaction entanglement because one optimizes the entanglement gain over the initial states. In contrast, the interaction entanglement, as well as the CP map entropy, are independent of the entanglement content of the state.

VI. CONCLUSION

The focus of this paper has been on giving an operational interpretation to the formal correspondence between operators and states, and enlarging our view of the probabilistic interpretation of quantum mechanics. We have seen that one can treat operations in a similar manner as one treats states. By making a single measurement one is able to say which operation acted on a state. The probability of the result of this measurement is given by a simple extension of the usual probability laws of quantum mechanics, and is independent of the state that gets acted on. The results follow from the ordinary laws of quantum mechanics, and yield interesting interpretational issues. While the probabilistic interpretation and collapse can be formulated in analogy to that of quantum states, it remains to be seen to what extent can we truly interpret the expansion of U as a sum over unitary evolutions as a quantum superposition of evolutions. One could object for instance to this interpretation by arguing that while the final outcome of the measurement is indeed a collapse to a single effective evolution U_μ, the evolution of the system in between the two intervention times, t_1 and t_2, is in fact not described by the resulting U_μ. Thus we have not collapsed to a single unitary but only to an effectively equivalent unitary. Such questions do not bother us for the case of a single time measurement, and it is not clear how to interpret such questions for the two-time measurements considered here. For clarity we use the phrase "which unitary acted" rather than "which path the state took."

It also remains to be studied in what respects the probabilistic interpretation of evolutions differs from the conventional interpretation. One such important difference is that while nonorthogonal states cannot be distinguished with certainty, nonorthogonal evolutions can given sufficiently many instances of the unitary. Understanding how this can be incorporated in a rigorous probabilistic formalism is a potentially rich area of research. It has also been advocated [33,34] that unitaries should have an interpretation similar to states since the unitary can be controlled by a quantum state (e.g., a cnot where the control bit is half a singlet). In the present work, the probabilistic nature of unitaries arises for macroscopic sources of unitaries, thus it may be interesting to understand the interplay between the two effects and their interpretations.

The information theoretic nature of evolutions has also been explored, and we have given an information theoretic interpretation to CP maps, using the idea of compression of their informational contents. For arbitrary realizations of a given CP map, we found that it was possible to compress the

map, and act it probabilistically on an unknown system. It would be interesting to explore whether one could act it on a known state given after compression. A generalization of superdense coding was also introduced. With regard to our entanglement concentration scheme, we have not yet touched on possible analogies for dilution for the case of interaction entanglement. This leaves open the question whether the proposed measure of interaction entanglement is a reversible quantity.

ACKNOWLEDGMENTS

We are grateful to Yakir Aharanov, Ignacio Cirac, Daniel Oi, Roberta Rodriquez, and Lev Vaidman for interesting discussions. J.O. acknowledges the support of the Lady Davis Trust, and ISF Grant No. 129/00-1 as well as funding by project PROSECCO (IST-2001-39227) of the IST-FET program of the EC and a grant from the Cambridge-MIT Institute. B.R. acknowledges the support of ISF Grant No. 62/01-1. This research was conducted during the Banasque session on Quantum Information and Communication, and we thank the town and the organizers for their hospitality.

APPENDIX

In this appendix we demonstrate our probabilistic interpretation and measurement scheme for the general d-dimensional case. Let us denote the orthogonal basis as $U_{\mu\nu}$ where the two indices take the values $\mu,\nu=0,\ldots,d-1$. Then

$$U_{\mu\nu}=U_0\sigma_{\mu\nu}, \tag{A1}$$

where $\sigma_{\mu\nu}$ are d^2 traceless orthogonal unitary operators. We will consider first the simple case where

$$\sigma_{\mu\nu}=(Z)^{\mu}(X)^{\nu}, \tag{A2}$$

where the operators [35]

$$Z=\sum_{j=0}^{N-1}\zeta^{j}|j\rangle\langle j| \tag{A3}$$

and

$$X=\sum_{j=0}^{N-1}|(j+1)\bmod N\rangle\langle j| \tag{A4}$$

are operators satisfying $Z^N=X^N=1$ and $ZX=\zeta XZ$, where $\zeta=\exp(2\pi i/N)$. We notice that for $N=2$, $Z\to\sigma_z$, and $X\to\sigma_x$ and regain our previous construction using Pauli operators. Thus Z and X play the role of generalized phase flip and bit flip operators.

The extension of the eigenoperators is then given by

$$A_X(t_2;t_1)=(U_0XU_0)_{t_2}(X)_{t_1}+\text{H.c.}, \tag{A5}$$

$$A_Z(t_2;t_1)=(U_0ZU_0)_{t_2}(Z)_{t_1}+\text{H.c.}, \tag{A6}$$

As we will shortly see, A_X and A_Z ascribe the value of the first and second indexes of a single element of the unitary basis $U_{\mu\nu}$.

To perform the measurement we employ now a pair of N-level ancillary systems in the initial state $\Sigma|\tilde{\alpha}\rangle\Sigma|\beta\rangle$. The interaction operator can be expressed as

$$V_Z=\sum_{\alpha=0}^{N-1}|\alpha\rangle\langle\alpha|Z^{\alpha} \tag{A7}$$

and similarly we define $\tilde{V}_X$. The sequence of interaction at t_1, free evolution, and interaction at t_2 then reads

$$(\tilde{V}_XV_Z)\Big(\sum C_{\mu\nu}\sigma_{\mu\nu}\Big)(\tilde{V}_XV_Z), \tag{A8}$$

where for the simplicity of presentation we have dropped the U_0 factor. Acting on the total state we obtain

$$\sum_{\mu\nu}C_{\mu\nu}\Bigg[\sum_{\alpha}X^{\alpha}\sigma_{\mu\nu}X^{\alpha}\sigma_{\mu\nu}^{\dagger}|\tilde{\alpha}\rangle\sum_{\beta}Z^{\beta}\sigma_{\mu\nu}Z^{\beta}\sigma_{\mu\nu}^{\dagger}|\beta\rangle\Bigg]\sigma_{\mu\nu}|\psi\rangle. \tag{A9}$$

The main point is that the operators $X^{\alpha}\sigma_{\mu\nu}X^{\alpha}\sigma_{\mu\nu}^{\dagger}$, and $Z^{\beta}\sigma_{\mu\nu}Z^{\beta}\sigma_{\mu\nu}^{\dagger}$ are constants of motion. Using the commutation relation of X and Z we finally get

$$\sum_{\mu\nu}C_{\mu\nu}\Bigg[\sum_{\alpha}\zeta^{\alpha\mu}|\tilde{\alpha}\rangle\sum_{\beta}\zeta^{\beta\mu}|\beta\rangle\Bigg]\sigma_{\mu\nu}|\psi\rangle \equiv\sum_{\mu\nu}C_{\mu_{\nu}}|\tilde{\phi}_{\mu}\rangle|\phi_{\nu}\rangle\sigma_{\mu\nu}|\psi\rangle, \tag{A10}$$

where the ancilla states $\tilde{\phi}_{\mu}$ and ϕ_{ν} are orthogonal, hence a measurement at $t=t_2$ will indeed collapse the sum to a single term with a probability $|C_{\mu\nu}|^2$, and leave only the unitary $\sigma_{\mu\nu}$.

More generally, it is known that for $d\geqslant 3$ there are different inequivalent unitary basis [12]. However there exists a one-to-one correspondence between the unitary basis and the

basis of maximally entangled states [36]. Since for the latter we can always identify an observable which distinguishes between the basis elements, a corresponding observable can be constructed for an arbitrary unitary basis.

We finally note, that the generalization of the particular construction above to the case of a continuous Hilbert space is straightforward. In this case $\sigma_{\mu\nu} \rightarrow \sigma_{x_0 p_0} = T_{x_0} T_{p_0}$, where

$$T_{x_0} = \int dx |x + x_0\rangle\langle x|, \tag{A11}$$

$$T_{p_0} = \int dx\, e^{ixp_0} |x\rangle\langle x|. \tag{A12}$$

The general set of orthogonal unitary operators is then $U_{x_0 p_0} = U_0(x,p)\sigma_{x_0 p_0}$, where x_0 and p_0 are continuous real numbers,

$$C_{x_0 p_0} = \int dx\, e^{ixp_0} \langle x + x_0 | U | x\rangle. \tag{A13}$$

The amplitude of a basis element for a general U has then a form similar to the Wigner distribution.

[1] A. Jamiolkowski, Rep. Math. Phys. **3**, 275 (1972).
[2] M. Choi, Linear Algebr. Appl. **10**, 285 (1975).
[3] K. Helkwig and K. Kraus, Commun. Math. Phys. **16**, 142 (1970).
[4] W. Dur and I. Cirac, Phys. Rev. A **64**, 012317 (2001).
[5] G. Vidal, L. Masanes, and J. Cirac, Phys. Rev. Lett. **88**, 047905 (2002).
[6] M. Nielson and I. Chuang, Phys. Rev. Lett. **79**, 321 (1997).
[7] D. Collins, N. Linden, and S. Popescu, Phys. Rev. A **64**, 032302 (2001).
[8] J. I. Cirac, W. Dur, B. Kraus, and M. Lewnstein, Phys. Rev. Lett. **86**, 544 (2001).
[9] P. Zanardi, C. Zalka, and L. Faoro, Phys. Rev. A **62**, 030301(R) (2000).
[10] P. Zanardi, Phys. Rev. A **63**, 040304(R) (2001).
[11] X. Wang and P. Zanardi, Phys. Rev. A **66**, 044303 (2002).
[12] R. Werner, e-print quant-ph/0003070.
[13] Y. Aharonov, J. Anandan, S. Popescu, and L. Vaidman, Phys. Rev. Lett. **64**, 2965 (1990).
[14] A. Acin, Phys. Rev. Lett. **87**, 177901 (2001).
[15] Y. Aharonov and D. Albert, Phys. Rev. D **29**, 223 (1984).
[16] L. Vaidman, Phys. Rev. A **49**, 1473 (1994).
[17] J. M. I. Bialynicki-Birula, Commun. Math. Phys. **44**, 129 (1975).
[18] D. Deutsch, Phys. Rev. Lett. **50**, 631 (1983).
[19] H. Maassen and J. Uffink, Phys. Rev. Lett. **60**, 1103 (1988).
[20] B. Schumacher, Phys. Rev. A **51**, 2738 (1995).
[21] D. Petz and M. Mosonyi, J. Math. Phys. **42**, 4857 (2001).
[22] F. Hiai and D. Petz, Commun. Math. Phys. **143**, 99 (1991).
[23] B. Schumacker, Phys. Rev. A **54**, 4707 (1996).
[24] K. Kraus, *States, Effects and Operations: Fundemental Notions in Quantum Theory* (Springer-Verlag, Berlin, 1983).
[25] D. Aharonov, e-print quant-ph/9910081.
[26] J. Oppenheim, M. Horodecki, and R. Horodecki, Phys. Rev. Lett. **90**, 010404 (2003).
[27] A. Harrow, P. Hayden, and D. Leung, e-print quant-ph/0307221.
[28] P. Boykin and V. Roychowdhury, e-print quant-ph/0003059.
[29] M. Mosca, A. Tapp, and R. de Wolf, e-print quant-ph/0003101.
[30] J. Oppenheim and M. Horodecki, e-print quant-ph/0306161.
[31] D. Leung, Quantum Inf. Comput. **2**, 13 (2001).
[32] W. Dur, G. Vidal, J. I. Cirac, N. Linden, and S. Popescu, Phys. Rev. Lett. **87**, 137901 (2001).
[33] C. Fuchs, e-print quant-ph/0205039.
[34] C. Fuchs, R. Schack, and P. Scudo, e-print quant-ph/0307198.
[35] J. Schwinger, Proc. Natl. Acad. Sci. U.S.A. **46**, 570 (1960).
[36] K. Vollbrecht and R. Werner, e-print quant-ph/9910064.