PHYSICAL REVIEW A VOLUME 60, NUMBER 4 OCTOBER 1999

Optimal manipulations with qubits: Universal-NOT gate

V. Bužek,[1,2] M. Hillery,[3] and R. F. Werner[4]
[1]*Institute of Physics, Slovak Academy of Sciences, Dúbravská cesta 9, 842 28 Bratislava, Slovakia*
[2]*Faculty of Informatics, Masaryk University, Botanická 68a, 602 00 Brno, Czech Republic*
[3]*Department of Physics and Astronomy, Hunter College, CUNY, 695 Park Avenue, New York, New York 10021*
[4]*Institut für Mathematische Physik, TU Braunschweig, Mendelssohnstraße 3, 38304 Braunschweig, Germany*
(Received 19 January 1999)

Quantum information processing is usually associated with a specific computational basis. Nevertheless, for a deeper understanding of the fundamental principles of quantum computing, it is essential to investigate what is the fidelity of *universal* operations that are basis independent; i.e., the task is to perform optimally a specific operation on a qubit or a quantum register that is in an unknown state. In this paper we introduce the universal-NOT gate that takes as an input a qubit in an arbitrary state $|\Psi\rangle$ and generates an output that is as close as possible to the orthogonal state $|\Psi^{\perp}\rangle$. [S1050-2947(99)51110-4]

PACS number(s): 03.67.Lx, 03.65.Bz

Classical information consists of bits, each of which can be either 0 or 1. Quantum information, on the other hand, consists of qubits which are two-level quantum systems (spin-1/2 particles) with one level labeled $|0\rangle$ and the other $|1\rangle$. Quantum two-level systems cannot only be in one of the two levels, but in any superposition of them as well. This fact makes the properties of quantum information quite different from that of its classical counterpart. For example, it is not possible to construct a device that will perfectly copy an *arbitrary* (unknown) state of a spin-1/2 particle [1,2], while the copying of classical information presents no difficulties. Another difference between classical and quantum information is as follows: It is not a problem to complement a classical bit, i.e., to change the value of a bit, a 0 to a 1 and vice versa. This is accomplished by a NOT gate. Complementing a qubit (i.e., inverting the state of the spin-1/2 particle), however, is another matter. The complement of a state $|\Psi\rangle$ is the state $|\Psi^{\perp}\rangle$ that is orthogonal to it. The question we want to address is: Is it possible to build a device that will take an *arbitrary* (unknown) qubit and transform it into the qubit orthogonal to it?

The best intuition for this problem is obtained by looking at the desired operation as an operation on the Poincaré sphere, which represents the set of pure states of a qubit system. Thus every state, pure or mixed, is represented by a vector in a three-dimensional space, whose components are the expectations of the three Pauli matrices. The full state space is thereby mapped onto the unit ball, whose surface represents the set of pure states. In this picture the ambiguity of choosing an overall phase for $|\Psi\rangle$ is already eliminated. The points corresponding to $|\Psi\rangle$ and $|\Psi^{\perp}\rangle$ are antipodes of each other. The desired universal-NOT (U-NOT) operation is therefore nothing but the *inversion of the Poincaré sphere*.

Note that the inversion preserves angles (related in a simple way to the scalar product $|\langle\Phi,\Psi\rangle|$ of rays), so by Wigner's theorem the ideal U-NOT operation is just implemented either by a unitary or by an antiunitary operation. Unitary operations correspond to proper rotations of the Poincaré sphere, whereas antiunitary operations correspond to orthogonal transformations with determinant -1. Clearly, the U-NOT operation is of the latter kind, and an antiunitary operator Θ (unique up to a phase) implementing it is

$$\Theta(\alpha|0\rangle+\beta|1\rangle)=\beta^{*}|0\rangle-\alpha^{*}|1\rangle. \tag{1}$$

The difficulty with antiunitarily implemented symmetries is that they are not completely positive; i.e., they cannot be applied to a small system, leaving the rest of the world alone. (The tensor product of an antilinear and a linear operator is ill-defined.) Thus time reversal, perhaps the best known operation of this kind, can only be a global symmetry, but makes no sense when applied only to a subsystem. By definition, a ''gate'' is an operation applied to only a part of the world, so must be represented by a completely positive operation. By the Stinespring dilation theorem this is equivalent to saying that any gate must have a realization by coupling the given system to a larger one (some ancillas), performing a unitary operation on the large system, and subsequently restricting it to a subsystem. Hence an ideal U-NOT gate does not exist. The same is true, of course, for other antiunitarily implemented operations like the complex conjugation (or equivalently the transposition) of the density matrix.

Because we cannot design a perfect universal-NOT gate, what we would like to do is see how close we can come. At this point we can consider two scenarios. The first one is based on the measurement of input qubit(s) — using the results of an optimal measurement we can manufacture an orthogonal qubit, or any desired number of them. Obviously, the fidelity of the NOT operation in this case is equal to the fidelity of the estimation of the state of the input qubit(s). The second scenario would be to approximate an antiunitary transformation on a Hilbert space of the input qubit(s) by a unitary transformation on a larger Hilbert space that describes the input qubit(s), blank qubits that are to become the complements, and the quantum device playing the role of the gate. We demand that the gate perform equally well for any (unknown) pure input state, so it is natural to focus on *universal* gates ''U-NOT,'' i.e., gates that treat every state vector in the same way in the sense of a unitary symmetry. In what follows we shall address both scenarios.

In order to state our problem precisely, let $\mathcal{H}=\mathbf{C}^2$ denote the two-dimensional Hilbert space of a single qubit. Then the input state of N systems prepared in the pure state $|\Psi\rangle$ is the N-fold tensor product $|\Psi\rangle^{\otimes N}\in\mathcal{H}^{\otimes N}$. The corresponding density matrix is $\rho\equiv\sigma^{\otimes N}$, where $\sigma=|\Psi\rangle\langle\Psi|$ is the one-particle density matrix. An important observation is that the vectors $|\Psi\rangle^{\otimes N}$ are invariant under permutations of all N sites, i.e., they belong to the symmetric, or ''Bose'' subspace $\mathcal{H}_+^{\otimes N}\subset\mathcal{H}^{\otimes N}$. Thus as long as we consider only pure input states we can assume all the input states of the device under consideration to be density operators on $\mathcal{H}_+^{\otimes N}$. We will denote by $\mathcal{S}(\mathcal{H})$ the density operators over a Hilbert space $\mathcal{H}$. Then the U-NOT gate must be a completely positive trace preserving map $T:\mathcal{S}(\mathcal{H}_+^{\otimes N})\rightarrow\mathcal{S}(\mathcal{H})$. Our aim is to design T in such a way that for any pure one-particle state $\sigma\in\mathcal{S}(\mathcal{H})$ the output $T(\sigma^{\otimes N})$ is as close as possible to the orthogonal qubit state $\sigma^{\perp}=\mathbb{1}-\sigma$. In other words, we are trying to make the fidelity $\mathcal{F}:=\mathrm{Tr}[\sigma^{\perp}T(\sigma^{\otimes N})]=1-\Delta$ of the optimal complement with the result of the transformation T as close as possible to unity for an arbitrary input state. This corresponds to the problem of finding the minimal value of the error measure $\Delta(T)$ defined as

$$\Delta(T)=\max_{\sigma,\,\mathrm{pure}}\mathrm{Tr}[\sigma T(\sigma^{\otimes N})]. \tag{2}$$

Note that this functional Δ is completely unbiased with respect to the choice of input state. More formally, it is invariant with respect to unitary rotations (basis changes) in $\mathcal{H}$: When T is any admissible map, and U is a unitary rotation on $\mathcal{H}$, the map $T_U(\rho)=U^*T(U^{\otimes N}\rho U^{*\otimes N})U$ is also admissible, and satisfies $\Delta(T_U)=\Delta(T)$. We will show later on that one may look for optimal gates T, minimizing $\Delta(T)$, among the *universal* ones, i.e., the gates satisfying $T_U=T$ for all U. For such U-NOT gates, the maximization can be omitted from the definition (2), because the fidelity $\mathrm{Tr}[\sigma T(\sigma^{\otimes N})]$ is independent of σ.

Measurement-based scenario. An estimation device by definition takes an input state $\rho\in\mathcal{S}(\mathcal{H}_+^{\otimes N})$ and produces, in every single experiment, an ''estimated pure state'' $\sigma\in\mathcal{S}(\mathcal{H})$. As in any quantum measurement this will not always be the same σ, even with the same input state ρ, but a random quantity. The estimation device is therefore described completely by the probability distribution of pure states it produces for every given input. Still simpler, we will characterize it by the corresponding probability density with respect to the unique normalized measure on the pure states (denoted ''$d\Phi$'' in integrals), which is also invariant under unitary rotations (for more details see Ref. [3]). For an input state $\rho\in\mathcal{S}(\mathcal{H}_+^{\otimes N})$, the value of this probability density at the pure state $|\Phi\rangle$ is

$$p(\Phi,\rho)=(N+1)\langle\Phi^{\otimes N},\rho\,\Phi^{\otimes N}\rangle. \tag{3}$$

To check the normalization, note that $\int d\Phi\,p(\Phi,\rho)=\mathrm{Tr}[X\rho]$ for a suitable operator X, because the integral depends linearly on ρ. By unitary invariance of the measure ''$d\Phi$'' this operator commutes with all unitary rotations of the form $U^{\otimes N}$, and since these operators, restricted to $\mathcal{H}_+^{\otimes N}$, form an irreducible representation of the unitary group of $\mathcal{H}$ [for $d=2$, it is just the spin-$N/2$ irreducible representation of SU(2)], the operator X is a multiple of the identity. To determine the factor, one inserts $\rho=\mathbb{1}$, and uses the normalization of ''$d\Phi$'' to verify that $X=1$.

Note that the density (3) is proportional to $|\langle\Phi,\Psi\rangle|^{2N}$, when $\rho=|\Psi^{\otimes N}\rangle\langle\Psi^{\otimes N}|$ is the typical input to such a device: N systems prepared in the same pure state $|\Psi\rangle$. In that case the probability density is clearly peaked sharply at states $|\Phi\rangle$ that are equal to $|\Psi\rangle$ up to a phase.

Suppose now that we combine the state estimation with the preparation of a new state, which is some function of the estimated state. The overall result will then be the integral of the state valued function with respect to the probability distribution just determined. In the case at hand the desired function is $f(\Phi)=(\mathbb{1}-|\Phi\rangle\langle\Phi|)$. So the result of the whole measurement-based (''classical'') scheme is

$$\sigma^{(out)}=T(\rho)=\int d\Phi\,p(\Phi,\rho)(\mathbb{1}-|\Phi\rangle\langle\Phi|). \tag{4}$$

The fidelity required for the computation of Δ from Eq. (2) is then equal to (see also [3])

$$\Delta=(N+1)\int d\Phi|\langle\Phi,\Psi\rangle|^{2N}(1-|\langle\Phi,\Psi\rangle|^2)=\frac{1}{N+2}, \tag{5}$$

where we have used that the two integrals have exactly the same form (differing only in the choice of N), and that the first integral is just the normalization integral. Since this expression does not depend on σ, we can drop the maximization in the definition (2) of Δ, and find $\Delta(T)=1/(N+2)$, from which we find that the fidelity of the creation of a complement to the original state σ is $\mathcal{F}=(N+1)/(N+2)$. Finally, we note that the result of the operation (4) can be expressed in the form

$$\sigma^{(out)}=s_{_N}\sigma^{\perp}+\frac{1-s_{_N}}{2}\mathbb{1}, \tag{6}$$

with the ''scaling'' parameter $s_{_N}=N/(N+2)$. From here it is seen that, in the limit $N\rightarrow\infty$, a perfect estimation of the input state can be performed, and, consequently, the perfect complement can be generated. For finite N the mean fidelity is always smaller than unity. The advantage of the measurement-based scenario is that once the input qubit(s) is measured and its state estimated an arbitrary number M of identical (approximately) complemented qubits can be produced with the same fidelity, simply by replacing the output function $f(\Phi)=(\mathbb{1}-|\Phi\rangle\langle\Phi|)$ by $f_M(\Phi)=(\mathbb{1}-|\Phi\rangle\langle\Phi|)^{\otimes M}$.

Quantum scenario: U-NOT gate. Let us assume we have N input qubits in an unknown state $|\Psi\rangle$ and we are looking for a transformation that generates M qubits at the output in a state as close as possible to the orthogonal state $|\Psi^{\perp}\rangle$. The universality of the proposed transformation has to guarantee that an arbitrary input state is complemented with the same fidelity. If we want to generate M approximately complemented qubits at the output, the U-NOT gate has to be represented by $2M$ qubits (irrespective of the number N of input qubits), M of which will only serve as ancilla, and M of which become the output complements. We will indicate these subsystems by subscripts a (input), b (ancilla), and c

(prospective output). The U-NOT gate transformation U_{NM} acts on the tensor product of all three systems. The gate is always prepared in some state $|X\rangle_{bc}$, independently of the input state $|\Psi\rangle$. The transformation is determined by the following explicit expression, valid for every unit vector $|\Psi\rangle \in \mathcal{H}$:

$$U_{NM}|N\Psi\rangle_a \otimes |X\rangle_{bc} = \sum_{j=0}^{M} \gamma_j |X_j(\Psi)\rangle_{ab} \otimes |\{(M-j)\Psi^{\perp}; j\Psi\}\rangle_c ,$$

$$\gamma_j = (-1)^j \binom{N+M-j}{N}^{1/2} \binom{N+M+1}{M}^{-1/2}, \qquad (7)$$

where $|N\Psi\rangle_a = |\Psi\rangle^{\otimes N}$ is the input state consisting of N qubits in the same state $|\Psi\rangle$. On the right-hand side of Eq. (7) $|\{(M-j)\Psi^{\perp}; j\Psi\}\rangle_c$ denotes symmetric and normalized states with $(M-j)$ qubits in the complemented (orthogonal) state $|\Psi^{\perp}\rangle$ and j qubits in the original state $|\Psi\rangle$. Similarly, the vectors $|X_j(\Psi)\rangle_{ab}$ consist of $N+M$ qubits, and are given explicitly by

$$|X_j(\Psi)\rangle_{ab} = |\{(N+M-j)\Psi; j\Psi^{\perp}\}\rangle_{ab} . \qquad (8)$$

Here the coefficients γ_j were chosen so that the scalar product of the right-hand side with a similar vector written out for $|\Phi\rangle$ becomes $\langle \Psi, \Phi\rangle^N$. This implies at the same time that U_{NM} is linear and that it is unitary after suitable extension to the orthogonal complement of the vector $|X\rangle_{bc}$.

Each of the M qubits at the output of the U-NOT gate is described by the density operator (6) with $s_{_N} = N/(N+2)$, *irrespective* of the number of complements produced. The fidelity of the U-NOT gate depends only on the number of inputs. This means that this U-NOT gate can be thought of as producing an approximate complement and then cloning it, with the quality of the cloning independent of the number of clones produced. The universality of the transformation is directly seen from the ''scaled'' form of the output operator (6).

We stress that the fidelity of the U-NOT gate (7) is exactly the same as in the measurement-based scenario. Moreover, it also behaves as a classical (measurement-based) gate in a sense that it can generate an arbitrary number of complements with the same fidelity. We have also checked that these cloned complements are pairwise separable.

The $N+M$ qubits at the output of the gate that do not represent the complements are described individually in the state by the density operator

$$\sigma_j^{(out)} = s\sigma + \frac{1-s}{2}\mathbb{1}, \quad j = 1, \ldots, N+M, \qquad (9)$$

with the scaling factor $s = N/(N+2) + 2N/[(N+M)(N+2)]$; i.e., these qubits are the *clones* of the original state with a fidelity of cloning larger than the fidelity of estimation. This fidelity depends on the number M of clones produced out of the N originals, and in the limit $M \to \infty$ the fidelity of cloning becomes equal to the fidelity of estimation. These qubits represent the output of the *optimal* $N \to N+M$ cloner introduced by Gisin and Massar [4] and also discussed by Bruss *et al.* [5]. This means that the U-NOT gate, as presented by the transformation in Eq. (7), serves also as a universal cloning machine.

At this point the question arises whether the transformation (7) represents the *optimal* U-NOT gate via the quantum scenario. If this is so, then it would mean that the measurement-based and the quantum scenarios realize the U-NOT gate with the same fidelity.

Theorem. Let $\mathcal{H}$ be a Hilbert space of dimension $d=2$. Then among all completely positive trace preserving maps $T: \mathcal{S}(\mathcal{H}_+^{\otimes N}) \to \mathcal{S}(\mathcal{H})$, the measurement-based U-NOT scenario (4) attains the smallest possible value of the error measure defined by Eq. (2), namely $\Delta(T) = 1/(N+2)$.

We have already shown [see Eq. (5)] that for the measurement-based strategy the error Δ attains the minimal value $1/(N+2)$. The more difficult part, however, is to show that no other scheme (i.e., quantum scenario) can do better. Here we will largely follow the arguments in [6,7].

Recall first the rotation invariance of the functional Δ, noted after Eq. (2). Moreover, Δ is defined as the maximum of a collection of linear functions in T, and is therefore convex. Putting these observations together we get

$$\Delta(\hat{T}) \leqslant \int dU\, \Delta(T_U) = \Delta(T), \qquad (10)$$

where $\hat{T} = \int dU\, T_U$ is the average of the rotated operators T_U with respect to the Haar measure on the unitary group. Thus $\hat{T}$ is at least as good as T, and is a *universal*-NOT gate ($\hat{T}_U = \hat{T}$). Without loss we will therefore assume from now on that $T_U = T$ for all U.

An advantage of this step is that a very explicit general form for universal operations is known from the ''covariant form'' of the Stinespring dilation theorem (see [7] for a version adapted to our needs). The form of T is further simplified in our case by the fact that both representations involved are irreducible: the defining representation of SU(2) on $\mathcal{H}$, and the representation by the operators $U^{\otimes N}$ restricted to the symmetric subspace $\mathcal{H}_+^{\otimes N}$. Then T can be represented as a convex combination $T = \Sigma_j \lambda_j T_j$, with $\lambda_j \geqslant 0$, $\Sigma_j \lambda_j = 1$, and T_j universal gates in their own right, but of an even simpler form. The universality of T already implies that the maximum can be omitted from the definition (2) of Δ, because the fidelity no longer depends on the pure state chosen. In a convex combination of universal operators T_j we therefore get

$$\Delta(T) = \sum_j \lambda_j \Delta(T_j). \qquad (11)$$

Minimizing this expression is obviously equivalent to minimizing with respect to the discrete parameter j.

We write the general form of the extremal gates T_j in terms of expectation values of the output state for an observable X on $\mathcal{H}$:

$$\mathrm{Tr}[T(\rho)X] = \mathrm{Tr}[\rho V^*(X \otimes \mathbb{1})V], \qquad (12)$$

where $V: \mathcal{H}_+^{\otimes N} \to \mathcal{H} \otimes \mathbf{C}^{2j+1}$ is an isometry intertwining the respective representations of SU(2), namely the restriction of the operators $U^{\otimes N}$ to $\mathcal{H}_+^{\otimes N}$ (which has spin $N/2$), on the one

hand, and the tensor product of the defining representation (spin-1/2) with the irreducible spin-j representation. By the triangle inequality for Clebsch-Gordan reduction, this implies $j=(N/2)\pm(1/2)$, so only two terms appear in the decomposition of T. It remains to compute $\Delta(T_j)$ for these two values. Omitting the details of the calculations (these follow closely the arguments presented in Ref. [7]), we find that

$$\Delta(T)=\begin{cases}1 & \text{for } j=N/2+\frac{1}{2},\\ 1/(N+2) & \text{for } j=N/2-\frac{1}{2}.\end{cases} \tag{13}$$

The first value corresponds to getting the state σ from a set of N copies of σ. The fidelity 1 is expected for this trivial task, because taking any one of the copies will do perfectly. On the other hand, the second value is the minimal error in the *optimal* U-NOT gate, which we were looking for. This clearly coincides with the value (5), so the theorem is proved.

Role of a priori knowledge. If the input state $|\Psi\rangle=\alpha|0\rangle+\beta|1\rangle$ is restricted to the case where the coefficients α and β are real, then it is possible to construct a perfect quantum-NOT gate. A measurement-based strategy in this case does not do as well. Specifically, the mean fidelity of an optimal estimation in the present case increases as a function of input qubits as (see [3])

$$\mathcal{F}=\frac{1}{2}+\frac{1}{2^{N+1}}\sum_{j=0}^{N-1}\sqrt{\binom{N}{j}\binom{N}{j+1}},$$

and it attains a value equal to unity only in the limit $N\to\infty$. This means that with *a priori* knowledge of the set of inputs, the quantum-NOT gate can perform better than the measurement-based strategy.

Summarizing our conclusions, we have shown that there is another difference between classical and quantum information: classical bits can be complemented, while arbitrary qubits cannot. It is, nonetheless, possible to construct approximate quantum-complementing devices, the quality of whose output is independent of the state of their input. These devices we called U-NOT gates. They are closely related to quantum cloners, and exploiting this connection it is possible to find an explicit transformation for an N-qubit input and M-qubit output U-NOT gate. When there is no *a priori* information available about the state of input qubits then these U-NOT gates do not do better than a measurement-based strategy. On the other hand, as we have shown, partial *a priori* information can dramatically improve the performance of the U-NOT gate.

Note added. Recently, a very interesting work by Gisin and Popescu [8] has been posted in the LANL e-print archive. These authors have introduced the so-called spin-flip operators in their analysis of encoding of quantum information into pairs of spins. The spin-flip operator is in fact equivalent to the one-input qubit realization of our U-NOT gate.

This work was in part supported by the Royal Society and by the Slovak Academy of Sciences. V.B. and R.F.W. thank the Benasque Center for Physics where part of this work was carried out.

[1] W. Wootters and W. Zurek, Nature (London) **299**, 802 (1982).
[2] V. Bužek and M. Hillery, Phys. Rev. Lett. **81**, 5003 (1998).
[3] R. Derka, V. Bužek, and A. Ekert, Phys. Rev. Lett. **80**, 1571 (1998).
[4] N. Gisin and S. Massar, Phys. Rev. Lett. **79**, 2153 (1997).
[5] D. Bruss, A. Ekert, and C. Macchiavello, Phys. Rev. Lett. **81**, 2598 (1998).
[6] R.F. Werner, Phys. Rev. A **58**, 1827 (1998).
[7] M. Keyl and R.F. Werner, e-print quant-ph/9807010.
[8] N. Gisin and S. Popescu, e-print quant-ph/9901072.