JOURNAL OF MODERN OPTICS, 2000, VOL. 47, NO. 2/3, 211–232

Universal-NOT gate

V. BUŽEK,

Institute of Physics, Slovak Academy of Sciences, Dúbravská cesta 9,
842 28 Bratislava, Slovakia and Faculty of Informatics, Masaryk
Universtiy, Botanická 68a, 602 00 Brno, Czech Republic

M. HILLERY

Department of Physics and Astronomy, Hunter College, CUNY, 695
Park Avenue, New York, NY 10021, USA.

and F. WERNER

Inst. f. Mathematische Physik, TU Braunschweig, Mendelssohnstr. 3,
38304 Braunschweig, Germany

(*Received 20 March 1999; revised version received 25 June, 1999*)

Abstract. The action of a NOT gate on a classical bit results in a change of its value from a 0 to a 1 and *vice versa*. The action of the classical NOT gate is in principle perfect because with fidelity equal to unity it complements the value of a bit. The action of the quantum NOT gate in a computational basis $|0\rangle$ and $|1\rangle$ is very similar to the action of the classical NOT gate. However, a more general quantum mechanical operation which corresponds to a classical NOT gate would take a qubit in an *arbitrary* state $|\Psi\rangle$ and produce a qubit in the state $|\Psi^{\perp}\rangle$ orthogonal to $|\Psi\rangle$. This operation is anti-unitary and therefore, cannot be realized exactly. *So how well we can do?* We find a unitary transformation acting on an input qubit and some auxiliary qubits, which represent degrees of freedom of the quantum NOT gate itself, which approximately realizes the NOT operation on the state of the original qubit. We call this 'device' a *universal*-NOT gate because the size of the error it produces is independent of the input state. We show that an *optimal* U-NOT gate which has as its input N identical qubits and produces M outputs achieves a fidelity of $\mathcal{F} = (N+1)/(N+2)$, which is equal to the fidelity of estimation of the input qubits. We also show that when *a priori* information about the state of the input qubit is available, the fidelity of a quantum NOT gate can be much better than the fidelity of estimation.

1. Introduction

In order to utilize the full potential of quantum information processing, we have to understand clearly what are the 'rules of the game'. In particular, the limits within which quantum information can be manipulated have to be determined. To be specific, quantum information is represented by qubits which are two-level quantum systems with one level labelled $|0\rangle$ and the other $|1\rangle$. Qubits can be not only 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 those of its classical counterpart. For example, it is not possible to construct a device which

Journal of Modern Optics ISSN 0950–0340 print/ISSN 1362–3044 online © 2000 Taylor & Francis Ltd
http://www.tandf.co.uk/JNLS/mop.htm
http://www.taylorandfrancis.com/JNLS/mop.htm

will perfectly copy an *arbitrary* qubit [1–4] while the copying of classical information presents no difficulties. This impossibility of copying (cloning) quantum information puts fundamental limits on the amount of information extractable from finite ensembles of identically prepared quantum systems [5, 6].

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 in an *unknown* state, however, is another matter. The complement of a qubit $|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle$ is the qubit $|\Psi^{\perp}\rangle = \beta^{*}|0\rangle - \alpha^{*}|1\rangle$ which is orthogonal to it. The question is: *Is it possible to build a device which will take an arbitrary 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 Bloch (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 Bloch 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 must be implemented either by a unitary or by an anti-unitary operation. Unitary operations correspond to proper rotations of the Bloch sphere, whereas anti-unitary operations correspond to orthogonal transformations with determinant -1. Clearly, the U-NOT operation is of the latter kind, and an anti-unitary operator Θ (unique up to a phase) implementing it is

$$\Theta(|\Psi\rangle) = |\Psi^{\perp}\rangle. \tag{1}$$

The difficulty with anti-unitarily 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 anti-linear 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 [7] (see [8, 9] for a version adapted to our needs) 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 to a subsystem. Hence an ideal U-NOT gate does not exist.

The same is true, of course, for other anti-unitarily implemented operations like the complex conjugation (or equivalently the transposition) of the density matrix, which corresponds to the reflection of the sphere at the $x_2 = 0$ plane, because only the Pauli matrix σ_2 has imaginary entries. Clearly, any such operation can be represented by a U-NOT, followed by a suitable unitary rotation, and conversely. On the other hand, if we relax the 'universality' condition, the U-NOT operation may become viable: if we are promised that the elements of the density matrix (or the components of $|\Psi\rangle$) are *real*, the states lie in the $x_2 = 0$ plane so that the inversion at the centre is equivalent to a proper rotation by π around the x_2-axis.

Because we cannot design a perfect Universal-NOT gate, we would like to see how close we can come. The most straightforward approach would be just to measure optimally [10–12] the input state $|\Psi\rangle$, resulting in a description of the state in terms of classical parameters, then to perform the desired transformation on these parameters, and to prepare a new state with the transformed parameters. This idea can be applied to any transformation on density matrices, be it completely positive or anti-linearly implemented. It even applies to non-linear operations or to the reversal of the effects of noise. The only problem with this approach is that states (pure states or density matrices) can only be determined statistically. Therefore, the input cannot be a single system, but has to be a collection of, say N, identically prepared systems. The quality of the result will depend on N, and will become perfect in the limit $N \to \infty$.

One task to which this approach has been applied is cloning. The single input or 'one-shot' version of this problem requires the production of two quantum systems, both in the same state as a given (unknown) input state, and is forbidden by the No-Cloning Theorem [1, 4]. The multiple input version [13] asks, more generally, for the optimal way to increase the number N of identically prepared input systems to number $N + M$ of output systems, all in a state as close as possible to the state of the inputs. In the case of pure input states this optimization problem has been solved completely [8, 9]. The result is that the above scenario, using a statistical measurement as an intermediate step is *not* optimal. Only when either N or M goes to infinity, i.e. either there are so many inputs that the statistical measurement is very good, or else very many output systems are required, the optimal solution approaches the classical, measurement-based one.

These results suggest that while a measurement-based method will produce a good U-NOT gate in the limit of very many input systems, there might be a better way, staying completely in the quantum world, and somehow utilizing the input more coherently. However, in a sense complementing is harder than cloning: we will show (at least in the pure input case) that the classical, measurement-based method is indeed optimal. The methods for this proof are very similar to those needed for the cloning problem, and we refer to [9] for a broader explanation of this background.

In this paper we will present a detailed description of the universal NOT gate which has been introduced in [14]. The presentation is organized as follows. In section 2 we treat the simplest case, a single input U-NOT gate with a single output. The discussion here is at an elementary level, and we again show that a perfect U-NOT gate is impossible after which we proceed to the description of an approximate gate. We also explore the connection between a U-NOT gate and a quantum cloner. In section 3 we use this connection to develop a network for the U-NOT gate. The role of *a priori* knowledge about the state of the input qubit is analysed in section 4. In section 5 we present the multiple-input U-NOT gate the optimality of which is proven in section 6. We conclude the paper with section 7.

2. Single-input U-NOT gate

We start the discussion with a single-input case. We will discuss first the measurement-based scenario, when the single qubit is first optimally measured and, based on the result of the measurement, a complement is prepared. Then we will study the quantum scenario when the input qubit interacts with an additional quantum system, the gate, and as a result of the unitary evolution of the whole system, the qubit of interest, i.e. a subsystem, can be found in the state as orthogonal as possible to the input.

2.1. *Measurement-based scenario*

Here we first measure the original qubit and using the result of the measurement we manufacture an orthogonal one. In the case of a single input qubit, the *optimal* way to estimate the state, is to measure it along a randomly chosen direction in the two-dimensional Hilbert space [10–12]. Therefore, the first step in implementing the measurement-based procedure is choosing a random vector $|\eta\rangle$, where

$$|\eta\rangle = \cos(\vartheta'/2)|0\rangle + \exp(i\varphi')\sin(\vartheta'/2)|1\rangle \tag{2}$$

and measuring $|\Psi\rangle$ along it. If the result is positive, then the output is taken to be $|\eta^{\perp}\rangle$, and if negative, the output is $|\eta\rangle$. This gives an output density matrix

$$\rho^{(\mathrm{out})}(\eta) = |\langle\Psi|\eta\rangle|^2|\eta^{\perp}\rangle\langle\eta^{\perp}| + |\langle\Psi|\eta^{\perp}\rangle|^2|\eta\rangle\langle\eta|. \tag{3}$$

To get the final output density matrix one averages this over all possible choices of the measurement (i.e. over all vectors $|\eta\rangle$)

$$\rho^{(\mathrm{out})} = \frac{1}{4\pi}\int_0^{2\pi}\mathrm{d}\varphi'\int_0^{\pi}\mathrm{d}\vartheta'\,\sin\vartheta'\,\rho^{(\mathrm{out})}(\eta). \tag{4}$$

After the integration is performed we find

$$\rho^{(\mathrm{out})} = s\rho^{\perp} + \frac{1-s}{2}\mathbb{1} \tag{5}$$

where for a single input qubit we have $s = 1/3$ and $\rho^{\perp} = |\Psi^{\perp}\rangle\langle\Psi^{\perp}|$. This is the best 'U-NOT operation' performed via measurement and estimation of the original qubit. It gives for the mean fidelity of the measurement-based U-NOT gate

$$\mathcal{F} = \int \mathrm{d}\Omega_{\rho}\langle\Psi^{\perp}|\rho^{(\mathrm{out})}|\Psi^{\perp}\rangle = \frac{2}{3} \tag{6}$$

where $d\Omega_{\rho} = 1/4\pi\,\sin\vartheta\,\mathrm{d}\vartheta\,\mathrm{d}\varphi$ is the corresponding integration measure. The advantage of the measurement-based scenario is that once the input qubit(s) is measured and its state is estimated an arbitrary number of identical (approximately) complemented qubits can be produced with the same fidelity.

2.2. *Quantum scenario*

Let us now construct a *unitary* gate which complements an arbitrary qubit. As discussed in the Introduction this cannot be done perfectly, and we review this point. An approximate gate can be constructed by appending an ancilla to the input qubit and performing a unitary transformation on the larger system. As we shall see, the ancilla can be taken to be two qubits, and the whole system, the input qubit and the two additional ones, is closely related to a quantum cloner.

We first note that logical operations in quantum information processing are usually defined in a specific basis. For instance the logical NOT operation defined as

$$\mathcal{N}|0\rangle = -|1\rangle; \qquad \mathcal{N}|1\rangle = |0\rangle \tag{7}$$

generates a complement (with some phase shift) to given basis vectors $|0\rangle$ and $|1\rangle$. If it is known *a priori* that the qubit is in one of two mutually orthogonal basis states (e.g. $|0\rangle$ or $|1\rangle$) then the NOT operation can be performed via a simple unitary rotation $R(\theta)$ defined as

$$R(\theta)|0\rangle \to \cos\theta/2|0\rangle - \sin\theta/2|1\rangle;$$
$$R(\theta)|1\rangle \to \sin\theta/2|0\rangle + \cos\theta/2|1\rangle \qquad (8)$$

with a specific angle of rotation $\theta = \pi$. On the other hand, suppose that we want to construct a gate which will take a qubit in an *unknown* state, $|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle$ (we shall sometimes find it convenient to represent α and β as $\alpha = \cos\vartheta/2$ and $\beta = \exp(i\varphi)\sin\vartheta/2$) and transform it into the state orthogonal to it,

$$|\Psi^{\perp}\rangle = \beta^{*}|0\rangle - \alpha^{*}|1\rangle \qquad (9)$$

That this operation cannot be performed perfectly can be seen by considering the action such a transformation must have on the basis states

$$|0\rangle \to \exp(\mathrm{i}\phi_0)|1\rangle; \qquad |1\rangle \to \exp(\mathrm{i}\phi_1|0\rangle \qquad (10)$$

where the phases, ϕ_0 and ϕ_1 are, for the moment, undetermined. We then note that there is no choice of these phases for which the output of our gate for a general input state

$$|\Psi\rangle \to \alpha\exp(\mathrm{i}\phi_0)|1\rangle + \beta\exp(i\phi_1)|0\rangle \qquad (11)$$

will be equal to $|\Psi^{\perp}\rangle$ for all values of α and β. The problem is, of course, the complex conjugates which appear in the orthogonal state. If α and β are real, then it is possible to realize this transformation by choosing $\phi_0 = \pi$ and $\phi_1 = 0$. As we have discussed this in the Introduction the transformation $|\Psi\rangle \to |\Psi^{\perp}\rangle$ can be realized by an anti-unitary transformation, Θ, which transforms the basis vectors as given by equation (7).

Quantum mechanics limits us to unitary transformations, so there is no quantum gate which can realize the transformation Θ. Because we cannot design a *perfect* U-NOT gate, we would like to see how close we can come. This entails approximating an anti-unitary transformation on a two-dimensional Hilbert space by a unitary transformation on a larger one. The gate is assumed to have its own degrees of freedom, and its quantum state is described by a vector in a Hilbert space whose dimension is left, for now, unspecified. We shall assume that the gate is always in the same initial state, which we shall designate as $|X\rangle$, and this state is assumed to be normalized. Our gate can now be described by the transformation

$$|0\rangle|X\rangle \to |1\rangle|Q_0\rangle + |0\rangle|Y_0\rangle;$$
$$|1\rangle|X\rangle \to |0\rangle|Q_1\rangle + |1\rangle|Y_1\rangle \qquad (12)$$

where the vectors $|Q_j\rangle$, and $|Y_j\rangle$, for $j = 1,2$, are states of the gate. Note that because we need only specify four output states of the gate, we can assume that its state space is four-dimensional. Unitarity implies these vectors must satisfy the equations

$$
\begin{aligned}
\|Q_0\|^2 + \|Y_0\|^2 &= 1;\\
\|Q_1\|^2 + \|Y_1\|^2 &= 1;\\
\langle Q_0|Y_1\rangle + \langle Y_0|Q_1\rangle &= 0.
\end{aligned}
\tag{13}
$$

Using the transformation above we find that for the input state $|\Psi\rangle$, the output state is given by

$$
\begin{aligned}
\rho^{(\mathrm{out})} = {} & |0\rangle\langle 0|(|\alpha|^2\|Y_0\|^2 + |\beta|^2\|Q_1\|^2 + \alpha^*\beta\langle Y_0|Q_1\rangle + \beta^*\alpha\langle Q_1|Y_0\rangle)\\
& + |1\rangle\langle 1|(|\alpha|^2\|Q_0\|^2 + |\beta|^2\|Y_1\|^2 + \alpha^*\beta\langle Q_0|Y_1\rangle + \beta^*\alpha\langle Y_1|Q_0\rangle)\\
& + |0\rangle\langle 1|(|\alpha|^2\langle Q_0|Y_0\rangle + |\beta|^2\langle Y_1|Q_1\rangle + \alpha^*\beta\langle Q_0|Q_1\rangle + \beta^*\alpha\langle Y_1|Y_0\rangle)\\
& + |1\rangle\langle 0|(|\alpha|^2\langle Y_0|Q_0\rangle + |\beta|^2\langle Q_1|Y_1\rangle + \alpha^*\beta\langle Y_0|Y_1\rangle + \beta^*\alpha\langle Q_1|Q_0\rangle).
\end{aligned}
\tag{14}
$$

From this expression we can compute the mean fidelity given by equation (6) of the output state:

$$
\begin{aligned}
\mathcal{F} = {} & |\alpha|^4\|Q_0\|^2 + \alpha^*\beta|\alpha|^2(\langle Q_0|Y_1\rangle - \langle Q_0|Y_0\rangle) + \alpha\beta^*|\alpha|^2(\langle Y_1|Q_0\rangle - \langle Y_0|Q_0\rangle)\\
& + |\alpha|^2|\beta|^2(\|Y_0\|^2 + \|Y_1\|^2 - \langle Y_0|Y_1\rangle - \langle Y_1|Y_0\rangle) - (\alpha^*\beta)^2\langle Q_0|Q_1\rangle\\
& - (\alpha\beta^*)^2\langle Q_1|Q_0\rangle + \alpha\beta^*|\beta|^2(\langle Q_1|Y_0\rangle - \langle Q_1|Y_1\rangle)\\
& + \alpha^*\beta|\beta|^2(\langle Y_0|Q_1\rangle - \langle Y_1|Q_1\rangle) + |\beta|^4\|Q_1\|^2.
\end{aligned}
\tag{15}
$$

Our first requirement is that $\mathcal{F}$ be independent of α and β. It is perhaps easiest to see how this happens if we set

$$
\alpha = \cos(\vartheta/2); \qquad \beta = \exp(\mathrm{i}\varphi)\sin(\vartheta/2).
\tag{16}
$$

The terms in equation (15) are either independent of φ, proportional to (exp ± iφ) or proportional to $\exp(\pm 2\mathrm{i}\varphi)$. Each of the terms with a particular kind of φ dependence, except of course those which are independent of φ, must vanish. For example, there are two terms proportional to $\exp(\mathrm{i}\varphi)$, so their sum must vanish, but each of these terms has a different ϑ dependence, so they must vanish individually. Reasoning of this type gives us that each of the φ-dependent terms is equal to zero. This implies that

$$
\begin{aligned}
\langle Q_0|Y_1\rangle &= \langle Q_0|Y_0\rangle;\\
\langle Q_1|Y_0\rangle &= \langle Q_1|Y_1\rangle;\\
\langle Q_0|Q_1\rangle &= 0.
\end{aligned}
\tag{17}
$$

We can take the remaining terms in equation (15) and express them as functions of $\cos(\vartheta/2)$. We then demand that the terms which are not independent of ϑ vanish. This gives us that $\mathcal{F} = \|Q_1\|^2$, and that

$$
\begin{aligned}
\|Y_0\|^2 + \|Y_1\|^2 - \|Q_0\|^2 - \|Q_1\|^2 &= \langle Y_0|Y_1\rangle + \langle Y_1|Y_0\rangle;\\
\|Y_0\|^2 + \|Y_1\|^2 - 2\|Q_0\|^2 &= \langle Y_0|Y_1\rangle + \langle Y_1|Y_0\rangle.
\end{aligned}
\tag{18}
$$

These two equations imply that

$$\|Q_0\|^2 = \|Q_1\|^2 \tag{19}$$

which together with the unitarity conditions, equation (13) further imply that

$$\|Y_0\|^2 = \|Y_1\|^2 = 1 - \|Q_0\|^2. \tag{20}$$

Let us now take into account the fact that we want to maximize $\mathcal{F}$, which means that we want to minimize $\|Y_0\|^2$. If we now substitute the results of the previous two equations into equation (18) we find that

$$\|Y_0\|^2 = \tfrac{1}{2} + \tfrac{1}{4}(\langle Y_0|Y_1\rangle + \langle Y_1|Y_0\rangle). \tag{21}$$

If we define x to be

$$x = \frac{\mathrm{Re}(\langle Y_0|Y_1\rangle)}{\|Y_0\|^2} \tag{22}$$

then $-1 \leqslant x \leqslant 1$, and

$$\|Y_0\|^2 = \frac{1}{2-x}. \tag{23}$$

The minimum of the right-hand side of this equation occurs when $x = -1$. This implies that

$$\|Y_0\| = \|Y_1\| = \sqrt{\tfrac{1}{3}}; \qquad \|Q_0\| = \|Q_1\| = \sqrt{\tfrac{2}{3}} \tag{24}$$

and that $|Y_0\rangle = -|Y_1\rangle$. This last condition in conjunction with equation (17) gives us that both $|Q_0\rangle$ and $|Q_1\rangle$ are orthogonal to $|Y_0\rangle$.

Summarizing our results we find that the Hilbert space of the gate is 3-dimensional, and that the vectors appearing in transformation in equation (12) can be take to be

$$|Q_0\rangle = \sqrt{\frac{2}{3}}\begin{pmatrix}1\\0\\0\end{pmatrix}; \qquad |Q_1\rangle = \sqrt{\frac{2}{3}}\begin{pmatrix}0\\1\\0\end{pmatrix}, \qquad |Y_0\rangle = \sqrt{\frac{1}{3}}\begin{pmatrix}0\\0\\1\end{pmatrix} \tag{25}$$

and $|Y_1\rangle = -|Y_0\rangle$. This gives $\mathcal{F} = 2/3$, and the qubit output state, obtained by tracing over the gate states, is

$$\rho^{(\mathrm{out})} = \tfrac{1}{3}\rho^{\perp} + \tfrac{1}{3}\mathbb{1}. \tag{26}$$

Let us now choose a specific representation for the vectors $|Q_0\rangle$, $|Q_1\rangle$, and $|Y_0\rangle$ as vectors in the symmetric subspace of the state space of two qubits. In particular, denote these qubits as qubits b and c (qubit a being the input qubit) and let

$$\begin{aligned} |Q_0\rangle &= -\sqrt{\tfrac{2}{3}}|00\rangle_{bc};\\ |Q_1\rangle &= \sqrt{\tfrac{2}{3}}|11\rangle_{bc};\\ |Y_0\rangle &= \frac{1}{\sqrt{6}}(|01\rangle_{bc} + |10\rangle_{bc}). \end{aligned} \tag{27}$$

With this choice we find that the NOT gate acts in the following way (for convenience we assume that qubits b and c are initially in the state $|00\rangle_{bc}$)

$$|\Psi\rangle_a|0\rangle_b|0\rangle_c \longrightarrow \sqrt{\frac{2}{3}}|\Psi^\perp\rangle_a|\Psi\rangle_b|\Psi\rangle_c - \frac{1}{\sqrt{6}}|\Psi\rangle_a\left(|\Psi^\perp\rangle_b|\Psi\rangle_c + |\Psi\rangle_b|\Psi^\perp\rangle_c\right). \qquad (28)$$

If we now examine the reduced density matrixes of the individual qubits at the output we find that qubit 1 is in the state given in equation (26) and that qubits b and c are both in the state

$$\rho_j^{(\text{out})} = \tfrac{2}{3}\rho + \tfrac{1}{6}\mathbb{1}; \qquad j = b, c. \qquad (29)$$

The output states of qubits b and c are the same as the output states of the Universal Quantum Cloning Machine (UQCM) [4]. That device can also be realized with three qubits; one is the state we want to copy, which we assume to be in the state $|\Psi\rangle$, one is a qubit onto which information about the quantum state of the first is to be copied, and the third, which has been called the 'idle' qubit, represents degrees of freedom of the cloning machine. The output of that device consists of two qubits whose reduced density matrixes are given by equation (29), and the 'idle' qubit which is in the state

$$\rho^{(\text{out})} = \tfrac{1}{3}\rho^{\mathrm{T}} + \tfrac{1}{3}\mathbb{1} \qquad (30)$$

where the superscript T denotes the transpose. This means that our 3-qubit realization of the U-NOT gate is not identical to the previously developed UQCM, but is very closely related to it. There is, in fact, some flexibility in the definition of the UQCM in the choice of how the two vectors which represent the degrees of freedom of the copy machine, correspond to the states of a single qubit, and in the U-NOT gate we have made a different choice from the original realization of the UQCM. The result is a quantum machine which is both the optimal quantum copier, and the optimal U-NOT gate.

3. Quantum network for U-NOT gate

We can now use the relation between the U-NOT gate and quantum cloner to derive a network for the U-NOT gate. The network for the cloner is known [15, 16] and only minor modifications are necessary. Our network takes 3 input qubits (one for the input with the two other qubits playing the rôle of the quantum U-NOT gate) and transforms them into 3 output qubits. The input qubit is indexed by a, while the two gate qubits are indexed by b and c.

In the network we will use a single-qubit NOT gate R defined by equation (7) and a two-qubit operator, the so-called controlled-NOT gate, which has as its inputs a control qubit and a target qubit. The control qubit is unaffected by the action of the gate, and if the control qubit is $|0\rangle$, the target qubit is unaffected as well. However, if the control qubit is in the $|1\rangle$ state, then a NOT operation is performed on the target qubit. The operator which implements this gate, P_{kl}, acts on the basis vectors of the two qubits as follows (k denotes the control qubit and l the target):

$$\begin{aligned} P_{kl}|0\rangle_k|0\rangle_l &= |0\rangle_k|0\rangle_l; & P_{kl}|0\rangle_k|1\rangle_l &= |0\rangle_k|1\rangle_l; \\ P_{kl}|1\rangle_k|0\rangle_l &= |1\rangle_k|1\rangle_l; & P_{kl}|1\rangle_k|1\rangle_l &= |1\rangle_k|0\rangle_l. \end{aligned} \qquad (31)$$

We can decompose the U-NOT gate network into two parts. In the first part the gate itself (qubits b and c) are prepared in a specific state $|X\rangle_{bc}^{(\mathrm{prep})}$. Then in the second part of the U-NOT network the information from the original qubit a is *redistributed* among the three qubits. That is, the action of the quantum U-NOT can be described as a sequence of two unitary transformations

$$|\Psi\rangle_a^{(\mathrm{in})}|0\rangle_b|0\rangle_c \longrightarrow |\Psi\rangle_a^{(\mathrm{in})}|X\rangle_{bc}^{(\mathrm{prep})} \longrightarrow |\Psi\rangle_{abc}^{(\mathrm{out})}. \tag{32}$$

Prior to any interaction with the input qubit we have to prepare the quantum U-NOT gate qubits (b and c) in a specific state $|X\rangle_{bc}^{(\mathrm{prep})}$

$$|X\rangle_{bc}^{(\mathrm{prep})} = C_{00}|00\rangle_{bc} + C_{01}|01\rangle_{bc} + C_{10}|10\rangle_{bc} + C_{11}|11\rangle_{bc} \tag{33}$$

with the amplitudes C_{ij}

$$C_{00} = \sqrt{\tfrac{2}{3}}; \qquad C_{01} = C_{10} = \frac{1}{\sqrt{6}}; \qquad C_{11} = 0. \tag{34}$$

The state (33) can be prepared with the help of a simple network [17] which is shown in the left block in figure 1.

Once the qubits of the quantum U-NOT gate are properly prepared then the complementing of the initial state $|\Psi\rangle_a^{(\mathrm{in})}$ of the original qubit can be performed by a sequence of four controlled-NOT operations followed by a single NOT operation performed on the qubit a (see figure 1)

$$|\Psi\rangle_{abc}^{(\mathrm{out})} = \mathcal{N}_a P_{ca} P_{ba} P_{ac} P_{a_b} |\Psi\rangle_a^{(\mathrm{in})} |X\rangle_{bc}^{(\mathrm{prep})}. \tag{35}$$

When this operation is combined with the preparation stage, we find that the basis states of the original qubit (a) are transformed as

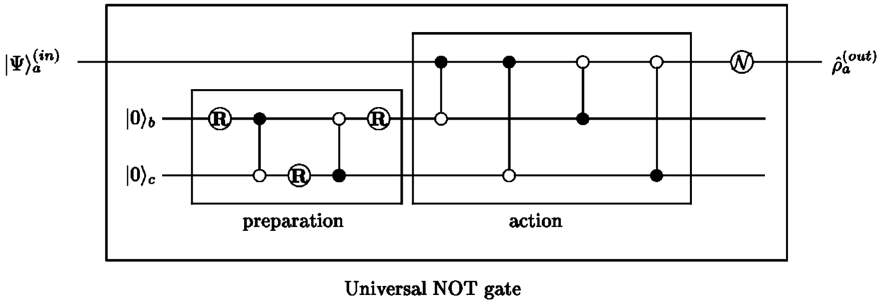

Figure 1. A graphical representation of a logical network for the U-NOT gate corresponding to the unitary transformation given by equation (35). The logical operation of the controlled NOT P_{kl} defined by equation (31) has as its input a control qubit (denoted as ●) and a target qubit (denoted as ○). The single-qubit rotation R and the NOT operation $\mathcal{N}$ are defined by equations (8) and (7), respectively. We see that the action of the U-NOT gate is represented by a sequence of four controlled NOT gates via which the information from the original qubit is distributed in the network. This distribution of information depends very much on the character of preparation of the U-NOT gate, i.e. on the state $|X\rangle_{bc}^{(\mathrm{prep})}$ (33) of the ancilla qubits b and c. When these two qubits are prepared in the state (33) with the probability amplitudes (34) then the qubit a is the one which at the output is in the complemented state (29).

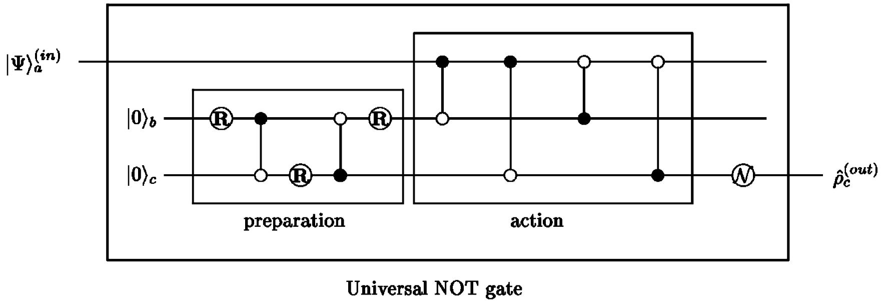

Figure 2. A graphical representation of a logical network for the U-NOT gate corresponding to the unitary transformation given by equation (38). Here the network is essentially the same as in figure 1 except that the U-NOT gate is initially prepared in the state (33) with the probability amplitudes given by equation (37) which result in the fact that at the output of the U-NOT gate the qubit c is complemented. In other words the flow of information in the U-NOT gate can be controlled by a specific preparation of this gate.

$$\begin{aligned} |0\rangle_a|00\rangle_{bc} &\to -\sqrt{\tfrac{2}{3}}|1\rangle_a|00\rangle_{bc} + \sqrt{\tfrac{1}{6}}|0\rangle_a(|01\rangle_{bc} + |10\rangle_{bc}); \\ |1\rangle_a|00\rangle_{bc} &\to \sqrt{\tfrac{2}{3}}|0\rangle_a|11\rangle_{bc} - \sqrt{\tfrac{1}{6}}|1\rangle_a(|01\rangle_{bc} + |10\rangle_{bc}) \end{aligned} \qquad (36)$$

which is equivalent to the U-NOT transformation given by equation (28). From our previous discussion we then conclude that the network shown in figure 1 can be used as the U-NOT gate as well as the universal cloning machine.

We note that if the amplitudes of the state $|X\rangle_{bc}^{prep}$ given by equation (33) are chosen to be

$$C_{00} = \sqrt{\tfrac{2}{3}}; \qquad C_{01} = C_{11} = \frac{1}{\sqrt{6}}; \qquad C_{10} = 0 \qquad (37)$$

then at the output of the network similar to the one given by equation (35) except with the NOT operation $\mathcal{N}_c$ applied to the qubit c, that is,

$$|\Psi\rangle_{abc}^{(\text{out})} = \mathcal{N}_c P_{ca} P_{ba} P_{ac} P_{ab} |\Psi\rangle_a^{(\text{in})} |X\rangle_{bc}^{(\text{prep})} \qquad (38)$$

we find the qubit c to be in the 'orthogonal' state (26) while the qubits a and b play the rôle of clones (see figure 2). This example illustrates the fact that using different preparations of the U-NOT gate (i.e. using different states $|X\rangle_{bc}^{(\text{prep})}$) we can control the flow of information in the network. In what follows we will use the convention that the qubit c will be the complement of the input qubit a.

3.1. *Multiple complements*

Here we present a generalization of the network (38) to the case when out of a single input qubit we wish to produce a set of M qubits in states which are as orthogonal as possible to the input. In order to generate these qubits we need a quantum U-NOT gate composed of $2M$ qubits. We can assume that initially all these qubits are prepared in the state $|0\rangle$, i.e. they are described by the state vector $|0\rangle_b^{\otimes M}|0\rangle_c^{\otimes M}$. Then using a unitary transformation these U-NOT gate qubits are prepared in the state

$$|X\rangle_{bc}^{(\mathrm{prep})} = \sum_{j=0}^{M}\left[e_j|\{(M+1)\cdot 0; j\cdot 1\}\rangle_b + f_j|\{(M-j+1)\cdot 0;(j-1)\cdot 1\}\rangle_b\right] \times |\{(M-j)\cdot 0; j\cdot 1\}\rangle_c \tag{39}$$

where $|\{A\cdot 0; B\cdot 1\}\rangle$ denotes a symmetric state with A qubits in the state $|0\rangle$ and B qubits in the state $|1\rangle$. The amplitudes e_j and f_j given by the expressions

$$e_j = \gamma_j^{(1,M)}\left(\frac{M+1-j}{M+1}\right)^{1/2}; \qquad f_j = \gamma_j^{(1,M)}\left(\frac{j}{M+1}\right) \tag{40}$$

with

$$\gamma_j^{(1,M)} = (-1)^j\left[\frac{2(M+1-j)}{(M+2)(M+1)}\right]^{1/2} \tag{41}$$

(this notation will become clear later in section 6). With the ancilla (i.e. the U-NOT gate) prepared in the state (39) the network for the 1 → M U-NOT gate can be expressed in the form

$$|\Psi\rangle_{abc}^{(\mathrm{out})} = \vec{\mathcal{N}}_c\vec{P}_{ca}\vec{P}_{ba}\vec{P}_{ac}\vec{P}_{ab}|\Psi\rangle_a^{(\mathrm{in})}|X\rangle_{bc}^{(\mathrm{prep})} \tag{42}$$

where

$$\vec{\mathcal{N}}_c = \prod_{j=1}^{M}\mathcal{N}_{c_j} \tag{43}$$

describes the action of the NOT gates on the qubits c and

$$\vec{P}_{ac} = \prod_{j=1}^{M}P_{ac_j} \tag{44}$$

describes the sequence of M controlled-NOT gates with the qubit a being the control and the qubits c_j being targets (see figure 3).

It can be shown that the network (42) is equivalent to the unitary transformation of the basis vectors of the original qubit a

$$|0\rangle_a|X\rangle_{bc}^{(\mathrm{prep})} \longrightarrow \sum_{j=0}^{M}\gamma_j^{(1,M)}|\{(M+1-j)\cdot 0; j\cdot 0^{\perp}\}\rangle_{ab}|\{(M-j)\cdot 0^{\perp}; j\cdot 0\}\rangle_c;$$

$$|1\rangle_a|X\rangle_{bc}^{(\mathrm{prep})} \longrightarrow \sum_{j=0}^{M}\gamma_j^{(1,M)}|\{(M+1-j)\cdot 1; j\cdot 1^{\perp}\}\rangle_{ab}|\{(M-j)\cdot 1^{\perp}; j\cdot 1\}\rangle_c \tag{45}$$

where the coefficients $\gamma_j^{(1,M)}$ are given by equation (41). We have also used the notation $|1^{\perp}\rangle = |0\rangle$ and $|0^{\perp}\rangle = -|1\rangle$.

Using the transformation (45) we can check that all qubits c (there are M of them) at the output of the U-NOT gate are in the state

$$\rho_{c_j}^{(\mathrm{out})} = \tfrac{1}{3}\rho^{\perp} + \tfrac{1}{3}\mathbb{1}; \qquad j = 1,\ldots,M \tag{46}$$

i.e. that they all are the best possible complements of the original qubit.

It is important to note that the fidelity of the gate does not depend on the number of complements produced, i.e. it is independent of M. In addition, these

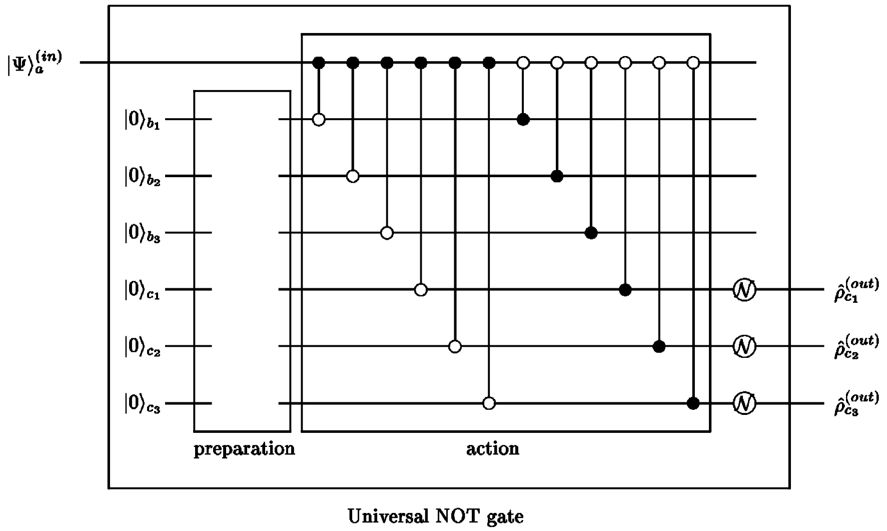

Figure 3. A graphical representation of a logical network for the U-NOT gate corresponding to the unitary transformation given by equation (45) when out of a single input qubit M complements are generated. In this particular case we consider $M = 3$. The structure of the network is very simple and the flow of information is controlled via the preparation of the gate qubits b_k and c_j.

qubits are pairwise separable, i.e. the two-qubit density operators $\rho_{c_i c_j}^{(\mathrm{out})}$ $(i \neq j)$ are separable.

To check this property, we first recall that a density operator of two subsystems is inseparable if it *cannot* be written as a convex sum

$$\rho_{xy} = \sum_m w^{(m)} \rho_x^{(m)} \otimes \rho_y^{(m)}. \tag{47}$$

Inseparability is one of the most fundamental quantum phenomena. It is required for a violation of Bell's inequality (to be specific, a separable system always satisfies Bell's inequality, but the contrary is not necessarily true). Distant parties cannot prepare an inseparable state from a separable one if they only use local operations and classical communication. In the case of two spin-1/2 particles we can utilize the Peres–Horodecki theorem [18, 19] which states that the positivity of the partial transposition of a state is *necessary* and *sufficient* for its separability.

Using the transformation (45) we can find the density operator $\rho_{c_i c_j}^{(\mathrm{out})}$ (here $i, j = 1, \ldots, M$ and $i \neq j$) describing an arbitrary state of two complemented qubits at the output of the U-NOT gate. From here we straightforwardly obtain eigenvalues $\vec{E} = \{E_1, E_2, E_3, E_4\}$ of the partially transposed density operator $(\rho_{c_i c_j}^{(\mathrm{out})})^{\mathrm{T}_2}$. These eigenvalues are input-state independent

$$\vec{E} = \left\{ \frac{1}{6}, \frac{1}{6}, \frac{1}{3} + \frac{\sqrt{2}}{6}, \frac{1}{3} - \frac{\sqrt{2}}{6} \right\} \tag{48}$$

and they do not depend on the number M of complemented qubits. Moreover, these eigenvalues are positive, which means that pairs of the complemented qubits at the output of the U-NOT gate are *not* quantum-mechanically entangled, i.e.

they are separable. We can conclude that from this point of view the U-NOT gate behaves as a classical device.

4. Role of *a priori* information

It follows from our previous results that there is no advantage in using the unitary device if one is trying to complement an arbitrary qubit in a pure state. However, if one is faced with mixed states, or if one has some information about the input state, a unitary device will do better. In this section we will study these two cases.

4.1. *Complementing mixed states*

The first thing to decide is what the complement of a mixed state of a qubit should be. The density matrix represents an ensemble of qubits, and so a natural way of producing an ensemble which is the complement of the first is to take each qubit in the original ensemble and replace it by a qubit in the orthogonal state. This corresponds to defining the complement, $\rho^{\perp}$, of a single qubit density matrix, ρ, to be

$$\rho^{\perp} = \Theta\rho\Theta^{-1} \tag{49}$$

where Θ is the anti-unitary operator given by equation (1). If we now send a qubit in the state ρ into the U-NOT gate, the reduced density matrix of the output will be

$$\rho^{(\text{out})} = \tfrac{1}{3}\rho^{\perp} + \tfrac{1}{3}\mathbb{1} \tag{50}$$

just as in the case of a pure state.

Developing a measurement-based strategy with which to compare this is now much harder. It is more difficult to estimate a mixed state than a pure state (which is due to the fact, that for a given quantum system a space of mixed states is much larger than a space of pure states). The best estimate of a density operator based on an optimal measurement of a single qubit prepared in a mixed state is given by the expression [20]

$$\rho^{(\text{est})} = \tfrac{1}{5}\rho + \tfrac{2}{5}\mathbb{1}. \tag{51}$$

Consequently, the estimation-based strategy for constructing the complement would on average result in the density operator

$$\rho^{(\text{out})} = \tfrac{1}{5}\rho^{\perp} + \tfrac{2}{5}\mathbb{1}. \tag{52}$$

This illustrates the fact that in the case of mixed states, the U-NOT gate is superior to a strategy based on measuring the qubit.

4.2. *Restricted input ensembles*

As was noted in the Introduction, 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 NOT gate. A measurement-based strategy in this case does not do as well. If we measure $|\Psi\rangle$ along a random direction

$$|\eta'\rangle = \cos\vartheta'|0\rangle + \sin\vartheta'|1\rangle \tag{53}$$

where $0 \leqslant \vartheta' < 2\pi$, compute $\rho^{(\mathrm{out})}(\eta')$ as in Equation (4), and then average over ϑ', we obtain

$$\rho^{(\mathrm{out})} = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\vartheta\, \rho^{(\mathrm{out})}(\vartheta) = \tfrac{1}{2}\rho^{\perp} + \tfrac{1}{4}\mathbb{1}, \tag{54}$$

giving a fidelity of 3/4 rather than one.

We now want to examine the situation in which we have some knowledge about the phases of α and β, but not complete information. Let α and β be given in terms of the angles ϑ and φ. We shall assume that the probability distribution, $p(\vartheta,\varphi)$, is normalized to unity on the Bloch sphere. Note that $p(\vartheta,\varphi) = 1$ corresponds to the case where nothing about the input qubit is known. We want to design a U-NOT gate which makes use of any extra *a priori* information to improve the fidelity between its output and the ideal output.

We shall consider a very simple gate, the one specified in equation (10) which generates the state $|\Psi\rangle^{(\mathrm{out})}$ given by equation (11). In this case the average fidelity of the gate for the input ensemble specified by $p(\vartheta,\varphi)$ is given by

$$\begin{aligned}\mathcal{F} &= \frac{1}{4\pi}\int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \mathrm{d}\vartheta \sin\vartheta\, p(\vartheta,\varphi) |\langle \Psi^{\perp} | \Psi^{(\mathrm{out})}\rangle|^2 \\ &= \tfrac{1}{2} + \frac{1}{8\pi}\int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \mathrm{d}\vartheta \sin\vartheta\, p(\vartheta,\varphi) F(\vartheta,\Delta\varphi)\end{aligned} \tag{55}$$

where $F(\vartheta,\Delta\varphi) = [\cos^2\vartheta - \sin^2\vartheta\cos(2\varphi - \Delta\varphi)]$ and $\Delta\varphi = \varphi_1 - \varphi_0$. For a given input ensemble $\Delta\varphi$ is chosen to maximize $\mathcal{F}$. For example, if φ is fixed at some value $\varphi = \varphi_p$, i.e. $p(\vartheta,\varphi) = \delta(\varphi - \varphi_p)h(\vartheta)$, where $h(\vartheta)$ is the probability distribution for ϑ, then we can choose $\Delta\varphi = \pi - 2\varphi_p$, and the average fidelity will be one. If nothing is known about the input states we find that $\mathcal{F} = 2/3$, the same value achieved by the universal device. Finally, let us look at the case where the probability distribution does not depend on ϑ (i.e. all values of ϑ are equally probable). Then $p(\vartheta,\varphi) \rightarrow p(\varphi)$, where $1/(2\pi)\int_0^{2\pi} d\varphi\, p(\varphi) = 1$. In this case we find

$$\mathcal{F} = \tfrac{2}{3} + \frac{u}{3} \tag{56}$$

where

$$u = \sup \int_0^{2\pi} \frac{d\varphi}{2\pi} p(\varphi)\cos(2\varphi - \Delta\varphi) \tag{57}$$

and the supremum is taken over all angles $\Delta\varphi$.

Let us now consider the measurement-based strategy in more detail so that, eventually, we can compare the fidelities it gives to those of the gate given above. As before, we choose an input qubit from our ensemble, measure it along a direction $|\eta\rangle$ [see equation (2)], and output $|\eta^{\perp}\rangle$ if the result is positive and $|\eta\rangle$ if it is negative. Whereas earlier we chose $|\eta\rangle$ to be a completely random state (any state on the Bloch sphere being equally likely), now we choose its angles ϑ' and φ' according to a probability distribution $q(\vartheta',\varphi')$, which will be chosen to optimize the average fidelity. The average fidelity for this scheme is given by

$$\mathcal{F} = \int \mathrm{d}\Omega \int \mathrm{d}\Omega'\, p(\vartheta,\varphi)\, q(\vartheta',\varphi')\, \langle \Psi^{\perp} | \rho(\eta) | \Psi^{\perp}\rangle \tag{58}$$

where both integrations are over the Bloch sphere with the invariant integration measure $\mathrm{d}\Omega = \mathrm{d}\varphi\,\mathrm{d}\vartheta\sin\vartheta/4\pi$. Inserting the explicit expressions for $\rho(\eta)$ and $|\Psi^{\perp}\rangle$ and simplifying the expression we find

$$\mathcal{F} = \int \mathrm{d}\Omega \int \mathrm{d}\Omega'\, p(\vartheta,\varphi)\, q(\vartheta',\varphi')\, f(\vartheta,\varphi,\vartheta',\varphi') \tag{59}$$

where

$$f(\vartheta,\varphi,\vartheta',\varphi') = \frac{[\sin\vartheta\sin\vartheta'\cos(\varphi-\varphi') + \cos\vartheta\cos\vartheta']^2}{2}. \tag{60}$$

The fidelity is maximized if $q(\vartheta',\varphi')$ is chosen to be a delta function at the values of ϑ' and φ' where $\int \mathrm{d}\Omega\, pf$ is a maximum. This implies that the maximum value of the fidelity is given by

$$\mathcal{F} = \sup \int \mathrm{d}\Omega\, p(\vartheta,\varphi)\, f(\vartheta,\varphi,\vartheta',\varphi') \tag{61}$$

where the supremum is taken over all values of ϑ' and φ'.

We now use this result to find the optimal fidelities from the measurement-based strategy in the two cases considered above. If $p(\vartheta,\varphi) = \delta(\varphi - \varphi_p)h(\vartheta)$, then the fidelity given in equation (61), while it can be one for special choices of $h(\vartheta)$, will in general be less than one. For the unitary gate, however, it is one for any choice of h. In the second case, $p(\vartheta,\varphi) \to p(\varphi)$, we find that the fidelity from equation (61) is given by

$$\mathcal{F} = \tfrac{2}{3} + \frac{u}{6} \tag{62}$$

which is, again, less than that resulting from the unitary gate. Therefore, we can conclude that while a unitary NOT gate does not do better than a measurement-based strategy in the case in which we have no information about the input qubits, this is not always true if we do have information about the input ensemble. In that case, there are situations in which the unitary gate will perform better.

5. Multiple-input U-NOT gate

Let us now suppose that instead of one input qubit in the state $|\Psi\rangle$, we have N. This, clearly, will allow us to produce approximate complements which are closer to $|\Psi^{\perp}\rangle$ than if we have only one input qubit. We shall first find the fidelity which is produced by the measurement-based strategy and then find a unitary gate which leads to the same fidelity. Finally we shall prove that this is the best one can do. This last point requires a more mathematical treatment, so that we shall begin by restating exactly what we wish to do.

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 $\sigma \equiv \rho^{\otimes N}$, where $\rho = |\Psi\rangle\langle\Psi|$ 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 $\mathcal{H}_+^{\otimes N} \otimes \mathcal{H}^{\otimes N}$. 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 $\rho \in \mathcal{S}(\mathcal{H})$ the output $T(\rho^{\otimes N})$ is as close as possible to the orthogonal qubit state $\rho^{\perp} = \mathbb{1} - \rho$. In other words, we are trying to make the fidelity $\mathcal{F} := \mathrm{Tr}[\rho^{\perp} T(\rho^{\otimes N})] = 1 - \Delta$ 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

$$\Delta(T) = \max_{\rho,\mathrm{pure}} \mathrm{Tr}\left[\rho \mathrm{T}(\rho^{\otimes N})\right]. \tag{63}$$

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 on $\mathcal{H}$, the map $T_U(\sigma) = U^* T(U^{\otimes N} \sigma U^{*\otimes N}) U$ is also admissible, and satisfies $\Delta(T_U) = \Delta(T)$. We will show later that one may look for optimal gates T, minimizing $\Delta(T)$, among the *universal* gates, i.e. those satisfying $T_U = T$ for all U. For such U-NOT gates, the maximization can be omitted from the definition (63), because the fidelity $\mathrm{Tr}[\rho \mathrm{T}(\rho^{\otimes N})]$ is independent of ρ.

5.1. *Measurement-based scenario*

An estimation device by definition takes an input state $\sigma \in \mathcal{S}(\mathcal{H}_+^{\otimes N})$ and produces, on every single experiment, an 'estimated pure state' $\rho \in \mathcal{S}(\mathcal{H})$. As in any quantum measurement this will not always be the same ρ, 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 '$\mathrm{d}\Phi$' in integrals), which is also invariant under unitary rotations. For an input state $\sigma \in \mathcal{S}(\mathcal{H}_+^{\otimes N})$, the value of this probability density at the pure state $|\Phi\rangle$ is

$$p(\Phi, \sigma) = (N + 1)\langle \Phi^{\otimes N}, \sigma \Phi^{\otimes N}\rangle. \tag{64}$$

To check the normalization, note that $\int \mathrm{d}\Phi\, p(\Phi, \sigma) = \mathrm{Tr}[X\sigma]$ for a suitable operator X, because the integral depends linearly on σ. By unitary invariance of the measure '$\mathrm{d}\Phi$' this operator commutes with all unitaries 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 $\sigma = \mathbb{1}$, and uses the normalization of '$\mathrm{d}\Phi$' to verify that $X = \mathbb{1}$.

Note that the density (64) is proportional to $|\langle \Phi, \Psi\rangle|^{2N}$, when $\sigma = |\Psi^{\otimes N}\rangle\langle\Psi^{\otimes N}|$ 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 $|\Phi\rangle$ which 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

$$\rho^{(\mathrm{est})} = T(\sigma) = \int \mathrm{d}\Phi \; p(\Phi,\sigma)\,(\mathbb{1} - |\Phi\rangle\langle\Phi|). \tag{65}$$

The fidelity required for the computation of Δ from equation (63) is then equal to (see also [12])

$$\Delta = (N+1)\int \mathrm{d}\Phi\, |\langle\Phi,\Psi\rangle|^{2N}(1 - |\langle\Phi,\Psi\rangle|^2) = \frac{1}{N+2} \tag{66}$$

where we have used the conditions 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 ρ, we can drop the maximization in the definition (63) of Δ, and find $\Delta(T) = 1/(N+2)$, from which we find that the fidelity of creation of a complement to the original state ρ is

$$\mathcal{F} = \frac{N+1}{N+2} \tag{67}$$

Finally we note that the result of the operation (65) can be expressed in the form

$$\rho^{(\mathrm{out})} = s_N \rho^{\perp} + \frac{1-s_N}{2}\mathbb{1}, \tag{68}$$

with the 'scaling' parameter $s_N = N/(N+2)$. From here it is seen that in the limit $N \to \infty$, 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 (are) 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 $f_M(\Phi) = (\mathbb{1} - |\Phi\rangle\langle\Phi|)^{\otimes M}$.

5.2. *Quantum scenario*

Let us now present a transformation which produces complements whose fidelity is the same as those produced by the measurement-based method. Assume we have N input qubits in an unknown state $|\Psi\rangle$ and we are looking for a transformation which 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 all input states are 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 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 $|X\rangle_{\mathrm{bc}}$, independently of the input state $|\Psi\rangle$. The transformation is determined by the following explicit expression, valid for every unit vector $|\Psi\rangle \in \mathcal{H}$:

$$U_{NM}|N\Psi\rangle_{\mathrm{a}} \otimes |X\rangle_{\mathrm{bc}} = \sum_{j=0}^{M} \gamma_j^{(N,M)} |X_j(\Psi)\rangle_{\mathrm{ab}} \otimes |\{(M-j)\Psi^{\perp}; j\Psi\}\rangle_{\mathrm{c}} \tag{69}$$

with

$$\gamma_j^{(N,M)} = (-1)^j \binom{N+M-j}{N}^{1/2} \binom{N+M+1}{M}^{-1/2} \tag{70}$$

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 equation (69) $|\{(M-j)\Psi^{\perp};j\Psi\}\rangle_c$ 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

$$|X_j(\Psi)\rangle_{ab} = |\{(N+M-j)\Psi;j\Psi^{\perp}\}\rangle_{ab}. \tag{71}$$

Note that with this choice of the coefficients $\gamma_j^{(N,M)}$, the scalar product of the right hand side with a similar vector, with Ψ replaced by Φ, becomes $\langle\Psi,\Phi\rangle^N$. This is consistent with the unitarity of the operator U_{NM}.

Each of the M qubits at the output of the U-NOT gate is described by the density operator (68) 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 (68).

We stress that the fidelity of the U-NOT gate (69) 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 which do not represent the complements are individually in the state described by the density operator

$$\rho_j^{(\text{out})} = s\rho + \frac{1-s}{2}\mathbb{1}, \qquad j = 1,\ldots,N+M, \tag{72}$$

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 [13]. This means that the U-NOT gate as presented by the transformation in equation (69) serves also as a universal cloning machine.

6. Optimality of U-NOT gate

At this point the question arises whether the transformation (69) represents the *optimal* U-NOT gate via 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 equation* (65) *attains the smallest possible value of the error measure defined by equation* (63), *namely* $\Delta(T) = 1/(N+2)$.

We have already shown [see equation (66)] that for the measurement-based strategy the error Δ attains the 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 [9].

Note first that the functional Δ is invariant with respect to unitary rotations (basis changes) in $\mathcal{H}$: when T is any admissible map, and U is a unitary on $\mathcal{H}$, the map $T_U(\sigma) = U^* T(U^{\otimes N} \sigma U^{*\otimes N}) U$ is also admissible, and satisfies $\Delta(T_U) = \Delta(T)$. Moreover, the functional Δ is defined as the maximum of a collection of linear functions in T, and is therefore convex. Putting these observations together we get

$$\Delta(T) \leqslant \int \mathrm{d}U \, \Delta(T_U) = \Delta(T_U) \tag{73}$$

where $T = \int \mathrm{d}U \mathrm{T}_U$ is the average of the rotated operators T_U with respect to the Haar measure on the unitary group. Thus T is at least as good as T_U, and has the additional 'covariance property' $T_U = T$. Without loss we can therefore assume from now on that $T_U = T$ for all U.

An advantage of this assumption is that a very explicit general form for such covariant operations is known by a variant of the Stinespring Dilation Theorem (see [9] 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 discrete convex combination $T = \sum_j \lambda_j T_j$, with $\lambda_j \geqslant 0, \sum_j \lambda_j = 1$, and T_j admissible and covariant maps in their own right, but of an even simpler form. Covariance of T already implies that the maximum can be omitted from the definition (63) of Δ, because the fidelity no longer depends on the pure state chosen. In a convex combination of covariant operators we therefore get

$$\Delta(T) = \sum_j \lambda_j \Delta(T_j). \tag{74}$$

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

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

$$\mathrm{Tr}[T(\sigma)X] = \mathrm{Tr}[\sigma V^*(X \otimes \mathbb{1})V] \tag{75}$$

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.

The basic idea is to use the intertwining property of the isometry V for the generators S_α, J_α, and $L_\alpha, \alpha = 1,2,3$ of the SU(2)-representations on $\mathcal{H}, \mathbf{C}^{2j+1}$ and $\mathcal{H}_+^{\otimes N}$, respectively. We will show that

$$V^*(S_\alpha \otimes \mathbb{1}_j)V = \mu_j L_\alpha, \tag{76}$$

where μ_j is some constant depending on the choice of j. That such a constant exists is clear from the fact that the left hand side of this equation is a vector operator (with components labeled by $\alpha = 1,2,3$), and the only vector operators in an irreducible representation of SU(2) are multiples of angular momentum (in this case L_α). The constant μ_j can be expressed in terms of a 6j-symbol, but can also be calculated in an elementary way by using the intertwining property, $VL_\alpha = (S_\alpha \otimes \mathbb{1} + \mathbb{1} \otimes J_\alpha)V$ and the fact that the angular momentum squares $\mathbf{J}^2 = \sum_\alpha J_\alpha^2 = j(j+1)$, $\mathbf{S}^2 = 3/4$, and $\mathbf{L}^2 = N/2(N/2+1)$ are multiples of the identity in the irreducible representations involved, and can be treated as scalars:

$$\mu_j \mathbf{L}^2 = \sum_\alpha V^*(S_\alpha \otimes \mathbb{1}_j)VL_\alpha = \mathbf{S}^2 + \sum_\alpha V^*(S_\alpha \otimes J_\alpha)V. \tag{77}$$

The sum on the right hand side can be obtained as the mixed term of a square, namely as

$$\frac{1}{2}\left(\sum_\alpha V^*(S_\alpha \otimes \mathbb{1} + \mathbb{1} \otimes J_\alpha)^2 V - \mathbf{S}^2 - \mathbf{J}^2\right) = (\mathbf{L}^2 - \mathbf{S}^2 - \mathbf{J}^2). \tag{78}$$

Combining these equations we find

$$\mu_j = \frac{1}{2} + \frac{\mathbf{S}^2 - \mathbf{J}^2}{2\mathbf{L}^2} = \begin{cases} \dfrac{1}{N} & \text{for } j = \dfrac{N}{2} + \dfrac{1}{2} \\[2ex] \dfrac{-1}{N+2} & \text{for } j = \dfrac{N}{2} - \dfrac{1}{2}. \end{cases} \tag{79}$$

We combine equations (75) and (76) to get the error quantity Δ from equation (63), with the pure one-particle density matrix $\rho = \frac{1}{2}\mathbb{1} + S_3$:

$$\Delta(T) = \mathrm{Tr}[V^*(\rho \otimes \mathbb{1})V\rho^{\otimes N}] = \frac{1}{2}(1 + N\mu_j). \tag{80}$$

With equation (79) we find

$$\Delta(T) = \begin{cases} 1 & \text{for } j = \dfrac{N}{2} + \dfrac{1}{2} \\[2ex] \dfrac{1}{N+2} & \text{for } j = \dfrac{N}{2} - \dfrac{1}{2}. \end{cases} \tag{81}$$

The first value is the largest possible fidelity for getting the state ρ from a set of N copies of ρ. 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 fidelity, which we were looking for. This clearly coincides with the value (66), so the Theorem is proved.

The Theorem as it stands concerns the task of producing just one particle in the U-NOT state of the input. From the results of the previous section we see that it is valid also in the case of many outputs. We see that the maximum fidelity is achieved by the classical process via estimation: in equation (65) we just have to replace the output state $(\mathbb{1} - |\Phi\rangle\langle\Phi|)$ by the desired tensor power. Hence once again the optimum is achieved by the scheme based on classical estimation. Incidentally, this shows that the multiple outputs from such a device are completely unentangled, although they may be correlated.

7. Concluding remarks

Our results clearly indicate yet another fundamental difference between classical and quantum information. Classical bit can be complemented perfectly, while as we have shown there are fundamental limits on the fidelity of the universal NOT gate. Specifically, the fidelity of the optimal universal NOT gate with N input qubits is equal to the fidelity of the optimal measurement performed on the input qubits. In this paper we have addressed the problem of how anti-unitary operation can be realized on qubits. Our results can be generalized for higher dimensions. We will report on this elsewhere.

At this point we should note that just after the appearance of our first report of these results [14], a very interesting work by Gisin and Popescu [21] appeared in which the concept of a spin-flip operation on a single qubit, which is essentially equivalent to the U-NOT gate with a single input, was introduced. It has been shown that the fidelity of the spin–flip machine, equal to 2/3, is optimal using the methods presented by Bechmann-Pasquinucci and Gisin [22].

Finally, it would be a real experimental challenge to construct a U-NOT gate and the related universal quantum cloning machine. In fact, a hint as to how to construct a U-NOT gate for polarization states of photons can be found in a short note by Mandel on cloning via stimulated emission [2]. This idea has recently beeen revived by Simon *et al.* [23], who have proposed an experiment in which the U-NOT gate and the universal quantum cloner are realized with the help of stimulated emission. Specifically, the input qubits are represented by polarization states of photons. The cloning is then realized via stimulated emission in an inverted medium. To make the cloning universal, i.e. input-state independent, the initial state of the inverted medium, and the interaction Hamiltonian with the electromagnetic field, have to be invariant under general polarization transformations so that the medium can emit photons of an arbitrary polarization with the same probability. That is, if a photon enters such a medium, it stimulates the emission of photons of the same polarization. Simultaneously, in addition to these photons, photons with the orthogonal polarization are also emitted. As shown by Simon *et al.* [23] these photons represent the corresponding output of the U-NOT gate with the optimal fidelity.

Acknowledgments

We thank Nicolas Gisin for his kind comments and the correspondence. We also thank Christoph Simon, Gregor Weihs, and Anton Zeilinger for very interesting discussions as well as for sending us their manuscript [23] prior to publication. This work was supported by the Slovak Academy of Sciences, the grant agency GACR (project 201/98/0369), and by the Royal Society. One of us (V.B.) would gratefully like to acknowledge the support and hospitality of the Isaac Newton Institute for Mathematical Sciences.

References

[1] Wootters, W., and Zurek, W. H., 1982, *Nature* **299,** 802.

[2] Mandel, L., 1983, *Nature*, **304,** 188.

[3] Dieks, D., 1982, *Phys. Lett. A*, **92,** 271.

[4] Bužek, V., and Hillery, M., 1996, *Phys. Rev. A*, **54,** 1844; see also Bužek, V., and Hillery, M., 1998, *Phys. Rev. Lett.*, **81,** 5003.

[5] Gisin, N., and Massar, S., 1997, *Phys. Rev. Lett.* **79,** 2153.

[6] Bruss, D., Ekert, A., and Macchiavello, C., *Phys. Rev. Lett.*, **81,** 2598.

[7] Stinespring, W. F., 1955, *Proc. Am. math. Soc.*, **6,** 211; Evans, D. E., and Lewis, J. T., 1977, *Dilations of Irreversible Evolutions in Algebraic Quantum Theory*, Communications of the Dublin Institute of Advanced Studies, Series A (Theoretical Physics), No. 24; Dublin, DIAS.

[8] Werner, R. F., 1998, *Phys. Rev. A*, **58,** 1827.

[9] Keyl, M., and Werner, R. F., 1998, *Optimal cloning of pure states, judging single clones*, Los Alamos e-print archive *quant-ph*/9807010.

[10] Holevo, A., 1982, *Probabilistic and Statistical Aspects of Quantum Theory* (Amsterdam: North Holland)

[11] Massar, S., and Popescu, S., 1995, *Phys. Rev. Lett.*, **74,** 1259.

[12] Derka, R., Bužek, V., and Ekert, A., 1998, *Phys. Rev. Lett.*, **80,** 1571.

[13] Gisin, N., and Massar, S., 1997, *Phys. Rev. Lett.* **79,** 2153.

[14] Bužek, V., Hillery, M., and Werner, R. F., 1999, *Optimal manipulations with qubits: Universal NOT gate*, Los Alamos e-print archive *quant-ph*/9901053.

[15] Bužek V., Braunstein, S., Hillery, M., and Bruss, D., 1997, *Phys. Rev. A*, **56,** 3446.

[16] Bužek, V., Hillery, M., and Knight, P.L., 1998, *Fort. der Physik*, **46,** 521.

[17] Barenco, A., Bennett, C., Cleve, R., DiVincenzo, D., Margolus, M., Shor, P., Sleator, T., Smolin, J., and Weinfurter, H., 1995, *Phys. Rev. A*, **52,** 3457.

[18] Peres, A., 1996, *Phys. Rev. Lett.*, **77,** 1423.

[19] Horodecki, M., Horodecki, P., and Horodecki, R., 1997, *Phys. Lett. A*, **223,** 1.

[20] Bužek, V., Derka, R., Adam, G., and Knight, P. L., 1998, *Ann. Phys. (N.Y.)*, **266,** 454.

[21] Gisin, N., and Popescu, S., 1999, *Spin flips and quantum information for anti-parallel spins*, Los Alamos e-print archive *quant-ph*/9901072.

[22] Bechmann-Pasquinucci, H., and Gisin, N., 1998, *Incoherent and Coherent Eavesdropping in the 6-state Protocol of Quantum Cryptography*, Los Alamos e-print archive *quant-ph*/9907041.

[23] Simon, C., Weihs, G., and Zeilinger, A., 1999, *Optimal quantum cloning and universal NOT without quantum gates, J. Mod. Opt.*, **47,** 233.