Classical simulation of quantum entanglement without local hidden variables

Serge Massar,[1] Dave Bacon,[2] Nicolas J. Cerf,[3,4] and Richard Cleve[5]
[1]*Service de Physique Théorique, Université Libre de Bruxelles, CP 225, 1050 Brussels, Belgium*
[2]*Departments of Physics and Chemistry, University of California Berkeley, California 94704*
[3]*Ecole Polytechnique, CP 165, Université Libre de Bruxelles, B-1050 Bruxelles, Belgium*
[4]*Information and Computing Technologies Research Section, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109*
[5]*Department of Computer Science, University of Calgary, Calgary, Alberta, Canada T2N 1N4*
(Received 22 September 2000; published 16 April 2001)

Recent work has extended Bell's theorem by quantifying the amount of communication required to simulate entangled quantum systems with classical information. The general scenario is that a bipartite measurement is given from a set of possibilities and the goal is to find a classical scheme that reproduces exactly the correlations that arise when an actual quantum system is measured. Previous results have shown that, using local hidden variables, a finite amount of communication suffices to simulate the correlations for a Bell state. We extend this in a number of ways. First, we show that, when the communication is merely required to be finite *on average*, Bell states can be simulated *without* any local hidden variables. More generally, we show that arbitrary positive operator valued measurements on systems of n Bell states can be simulated with $O(n2^n)$ bits of communication on average (again, without local hidden variables). On the other hand, when the communication is required to be *absolutely bounded*, we show that a finite number of bits of local hidden variables is insufficient to simulate a Bell state. This latter result is based on an analysis of the nondeterministic communication complexity of the NOT-EQUAL function, which is constant in the quantum model and logarithmic in the classical model.

DOI: 10.1103/PhysRevA.63.052305 PACS number(s): 03.67.Hk, 03.67.Lx

I. INTRODUCTION

We consider how much classical communication is required to simulate the correlations exhibited by measuring entangled quantum systems. Following [1], define a *quantum measurement scenario* as a triple of the form $(|\Psi\rangle_{AB}, M_A, M_B)$, where $|\Psi\rangle_{AB}$ is an entangled bipartite quantum state, M_A is a set of measurements on the first component, and M_B is a set of measurements on the second component. The goal is to devise communication protocols that enable two separated parties, Alice and Bob, to simulate a quantum measurement scenario using classical information. The *input* to the protocol is $(x,y) \in M_A \times M_B$, and Alice receives x (but not y), while Bob receives y (but not x). Alice and Bob's *outputs* should be jointly distributed so as to exactly reproduce the probability distribution that arises if an actual quantum system in state $|\Psi\rangle_{AB}$ is measured according to (x,y). We shall refer to this problem as *classical entanglement simulation*. In [2], a related problem, dubbed *classical teleportation*, is also introduced. Here Alice is given a classical description of a quantum state $|\Psi\rangle$ and Bob is given a classical description of a quantum measurement $x \in M$. The goal is for Bob to produce data that stochastically simulates the result of applying measurement x to state $|\Psi\rangle$. As shown in [2] and discussed below, this problem is closely related to classical entanglement simulation.

The first relevant result in this topic is Bell's famous theorem [3], which implies that, when $|\Psi\rangle_{AB}$ is a Bell state, there exist (M_A, M_B) for which Alice and Bob must perform *some* (nonzero) communication in order to achieve classical entanglement simulation. More recently, Brassard, Cleve, and Tapp [1] and, independently, Steiner [4] have shown that, when $|\Psi\rangle_{AB}$ is a Bell state and M_A, M_B are each the set of all von Neumann measurements on a qubit, the simulation is possible with only a *finite* amount of classical communication between Alice and Bob.

In the protocols devised in [1] and [4], it is supposed that Alice and Bob have an infinite supply of correlated random bits (specifying real-valued parameters). Such shared random bits are generally called *local hidden variables*. The two papers differ in their technical definition of the ''finite amount of classical communication.'' In [1], the amount of communication that occurs in the protocol is *exactly* 8 bits. In contrast (a slightly generalized version of) the protocol in [4] has the property that, for any given pair of measurements $(x,y) \in M_A \times M_B$, the *average* (i.e., expected) number of bits of communication is 2.97 bits; however, the amount of communication for any particular execution of the protocol may be arbitrarily large. The result in [4] is then refined in [2], where the amount of classical communication is decreased to 1.19 bits on average for all von Neumann measurements. Also, the sets M_A and M_B are extended to include all positive-operator-valued measurements (POVM's), using 6.38 bits of communication on average. We will refer to the first kind of protocol as a *bounded communication model*, whereas the second kind will be called an *average communication model*.

Regarding the classical entanglement simulation of more than one Bell state, it is shown in [1] that the exact simulation of arbitrary von Neumann measurements on n Bell states requires $\Omega(2^n)$ bits of communication in the bounded communication model. With minor modifications to the techniques in [1,5], this $\Omega(2^n)$ lower bound also carries over to the average communication model. Also note that this result

(as well as most other results for classical simulation of entanglement) immediately applies to classical teleportation protocols. This is because any protocol for classical teleportation of an n-qubit state can be converted into one for classical entanglement simulation of n Bell states with the same amount of communication. This is accomplished by Alice first simulating (by herself) the probabilistic effect of measuring ''her'' n qubits of the n Bell states. She also computes the resulting mixture of pure states that describes ''Bob's'' n qubits. Then Alice classically teleports the state of Bob's n qubits to him. Conversely, protocols for the classical entanglement simulation can be converted into protocols for classical teleportation, at the expense of a little more communication (see [2] for details).

The present paper generalizes the above results in a number of ways. All protocols for classical entanglement simulation proposed so far apply to single Bell states, and they use an infinite number of bits of local hidden variables for the simulation. Our first result is that local hidden variables are *not* necessary in the average communication model. In particular, when $|\Psi\rangle_{AB}$ is a Bell state and M_A, M_B are each the set of all von Neumann measurements, classical entanglement simulation is possible with a constant number (less than 20) of bits of communication on average, without any local hidden variables. We also show that, when $|\Psi\rangle_{AB}$ consists of n Bell states and M_A, M_B are each the set of all POVM's, the simulation can be carried out with no local hidden variables and $O(n2^n)$ bits of communication on average. Note that this communication cost is almost optimal, due to the aforementioned lower bound of $\Omega(2^n)$.

In contrast to the above results about the the average communication model, we show that local hidden variables are *necessary* in the bounded communication model (when $|\Psi\rangle_{AB}$ is a Bell state, and M_A, M_B are all von Neumann measurements). More precisely, the simulation of $(|\Psi\rangle_{AB}, M_A, M_B)$ in the bounded communication model requires an *infinite* number of bits of local hidden variables. This follows from a connection between the quantum measurement scenario and the nondeterministic communication complexity of the NOT-EQUAL function. These results indicate that there is a fundamental difference between the absolutely bounded communication model and the model of communication with bounded expectation.

II. CASE OF A SINGLE BELL STATE

We begin by considering the case of von Neumann measurements on Bell states. Our first result, stated in theorem 1, is actually a special case of a stronger result given in theorem 3 (where the bound on the amount of communication will be decreased from 22 to 20 bits, and where the measurements can be arbitrary POVM's). The proof of theorem 1 uses the same basic approach as that of theorem 3, but since it is considerably simpler, it is presented first.

Theorem 1. For the quantum measurement scenario $(|\Psi\rangle_{AB}, M_A, M_B)$, where $|\Psi\rangle_{AB} = (1/\sqrt{2})(|00\rangle + |11\rangle)$ and where M_A, M_B are each the set of all von Neumann measurements, classical entanglement simulation is possible without any local hidden variables with a constant number (less than 22) of bits of communication on average.

Proof. We first recall Steiner's original protocol [4]. The task of the two parties, Alice and Bob, is to simulate carrying out measurements on the Bell state, $(1/\sqrt{2})(|00\rangle + |11\rangle)$ with respect to operators $R(x)$ and $R(y)$ $(x,y \in [0,1])$, where

$$R(x) = \begin{pmatrix} \cos(2\pi x) & \sin(2\pi x) \\ \sin(2\pi x) & -\cos(2\pi x) \end{pmatrix}. \tag{1}$$

In order to carry out this simulation, Alice and Bob share an infinite sequence of local hidden variables $\theta_1, \theta_2, \ldots$, which are uniformly distributed over the interval $[0,1]$. In addition, Alice has an infinite set of values $u_1, u_2, \ldots$, which are also uniformly distributed over the interval $[0, 1]$.

In order to simulate a Bell state, Alice and Bob carry out the following operations.

(i) Alice finds the smallest value $k \in \{1,2, \ldots\}$ such that $u_k \leqslant |\cos[2\pi(\theta_k - x)]|$. Then Alice sends the value of this k to Bob, and she outputs the value of $\mathrm{sgn}\{\cos[2\pi(\theta_k - x)]\}$.

(ii) After Bob receives the index k from Alice, he outputs the value of $\mathrm{sgn}\{\cos[2\pi(\theta_{k-y})]\}$.

One can verify that this protocol produces the correct statistics (namely, that Alice and Bob's outputs are random bits, correlated so as to be equal with probability $\cos^2[\pi(x-y)]$) and that the amount of communication is 1.485 bits on average.

Steiner's protocol enables Alice to effectively generate a random variable, θ, distributed according to the density function $p(\theta) = (\pi/2)|\cos[2\pi(\theta - x)]|$ and convey this value to Bob. Explicitly sending the exact value of θ requires an infinite number of bits of communication; the above method uses local hidden variables to accomplish this with a finite amount of communication.

In order to circumvent the need for local hidden variables (or an infinite amount of communication), a different approach is used. Alice generates θ herself, according to the density function $p(\theta) = (\pi/2)|\cos[2\pi(\theta - x)]|$. In most cases, only a few bits of θ suffice for Bob to be able to compute the value of $\mathrm{sgn}\{\cos[2\pi(\theta - y)]\}$. So Alice sends Bob only a few bits of θ at a time and receives a response from Bob each time as to whether or not the precision is sufficient. In the first round. Alice sends Bob the first two significant bits of θ (since one bit of precision is never sufficient for Bob). Then Bob determines whether this information unambiguously determines the value of $\mathrm{sgn}\{\cos[2\pi(\theta - y)]\}$ and indicates the answer in a bit sent to Alice. In subsequent rounds, Alice sends one additional bit of precision of θ to Bob, until Bob's response indicates that the precision is sufficient.

To upper bound the expected amount of information communicated, note that, after each round, Bob has at least a $\frac{1}{5}$ chance of having θ with sufficient precision. This is because, for any $z \in [0,1]$ and $w \in [0, \frac{1}{4}]$,

$$\int_{z-w}^{z} \frac{\pi}{2} |\cos(2\pi\theta)| d\theta > \tfrac{1}{4} \int_{z}^{z+w} \frac{\pi}{2} |\cos(2\pi\theta)| d\theta. \tag{2}$$

Thus the expected number of rounds is less than 5. Since the first round consists of 3 bits (two from Alice and one from Bob) and each subsequent round consists of 2 bits (one from

each of Alice and Bob), the expected number of bits of communication is less than 11. To simulate an arbitrary von Neumann measurement, it suffices to simulate two measurements with respect to operators of the form $R(x)$ [1]. Thus the expected amount of communication is less than 22 bits. ■

Regarding the *minimum* number of bits of communication necessary to perform classical entanglement simulation without local hidden variables as in theorem 1, it should be noted that a single run of a protocol without local hidden variables cannot succeed in general if the communication is less than 1 bit. This is because, in the case where the two measurements x and y are both in the same basis, Alice's and Bob's outputs have exactly one bit of mutual information. One can easily check that this mutual information is a lower bound on the amount of forward and backward communication that must be used to simulate entanglement (see the Appendix). Therefore, 1 bit of communication is necessary in this worst case. For other specific pairs of measurements (x,y), the mutual information is lower than 1 bit and so is the minimum amount of communication. Let us now consider the case where the measurement directions x and y are chosen at random and assumed to be isotropic (the distribution of maximum uncertainty). The average communication here is with respect to the probabilistic selection of a pair of measurements as well as the probabilistic choices made by Alice and Bob during the execution of the protocol.

Lemma 1. Let $(|\Psi\rangle_{AB}, M_A, M_B)$ be the quantum measurement scenario where $|\Psi\rangle_{AB}=(1/\sqrt{2})(|00\rangle+|11\rangle)$, and M_A, M_B are each the set of all von Neumann measurements. Suppose that a pair (x,y) is selected according to two independent uniform distributions on the surface of the Bloch sphere. Then, for any protocol in the average communication model that has no local hidden variables, the sum of the (forward and backward) communication must be at least 0.279 bits on average.

Proof. Consider the situation where Alice and Bob are each given a random measurement direction ($\vec{x}$ and $\vec{y}$) by a third person—say, Charles. As before, Bob does not know the measurement Alice is performing, and, conversely, Alice ignores Bob's measurement. We also assume that, initially, the two parties share no information. Then Charles observes the outcomes of Alice's and Bob's measurements, noted a $(=\pm 1)$ and b $(=\pm 1)$. Our goal here is to estimate the number of bits that must be communicated from Alice to Bob (and from Bob to Alice) in order for them to exactly reproduce the quantum correlations that would be observed if they shared a singlet. This quantity can be bounded from below by the amount of shared randomness that Charles observes between Alice's and Bob's outcomes, while *knowing* the measurement directions. In other words, what is relevant here is the mutual information between Alice's and Bob's outcomes a and b *conditionally* on the measurement directions x and y—that is, $I(a:b|x,y)$. (It is shown in the Appendix that the mutual information is indeed a lower bound on the number of bits that must be communicated.) For given measurement directions $\vec{x}$ and $\vec{y}$, the correlation coefficient is $r=-\vec{x}\cdot\vec{y}$, so that the joint distribution of the outcomes is $p(a,b|r)=(1+rab)/4$, with $a=\pm 1$ and $b=\pm 1$. The resulting mutual information for a given r is then equal to

$$I(a:b|r)=\frac{1+r}{2}\log_2(1+r)+\frac{1-r}{2}\log_2(1-r). \qquad (3)$$

If $\vec{x}$ and $\vec{y}$ are uniformally distributed, then the correlation coefficient is distributed as $P(r)=1/2$ in the interval $[-1,1]$. As a result, the (average) mutual information between a and b conditionally on r can be written as

$$I=\int I(a:b|r)P(r)dr=\int_{-1}^{1}(1+r)\log_2(1+r)dr$$
$$=\log_2(2/\sqrt{e}). \qquad (4)$$

Thus the amount of (forward and backward) communication that is necessary to establish this shared randomness between Alice and Bob is bounded by $C_f+C_b\geq I=0.279$ bits. ■

Note that this bound assumes that there are no initially shared local hidden variables between Alice and Bob. In a more general scenario, however, the bound in lemma 1 only measures the *total* amount of shared randomnes, possibly including prior shared randomness. In other words, I is the sum of the initial shared randomness and the communication, so it does not discriminate the random bits that are shared beforehand (the local hidden variables) from the bits that are communicated after the measurement basis are disclosed to Alice and Bob.

Now we shall show that, if the bound on the communication is changed from being constant on average to being an absolute constant, then the classical entanglement simulation without local hidden variables that occurs in theorem 1 becomes impossible to achieve. Prior to doing this, we review a relevant result from the theory of communication complexity (see [6] for an extensive review of the field). Consider the NOT-EQUAL function, F_{NE}: $\{0,1\}^n\times\{0,1\}^n\rightarrow\{0,1\}$, defined as

$$F_{\mathrm{NE}}(x,y)=\begin{cases}1 & \text{if } x\neq y\\ 0 & \text{if } x=y,\end{cases} \qquad (5)$$

and suppose that Alice and Bob are given x and y, respectively, as inputs and their goal is to evaluate $F_{\mathrm{NE}}(x,y)=0$, in the following weak sense. Bob should output a bit b that is distributed so that if $F_{\mathrm{NE}}(x,y)=0$, then $\Pr[b=1]=0$; if $F_{\mathrm{NE}}(x,y)=1$, then $\Pr[b=1]>0$. Also, assume that Alice and Bob have no *a priori* shared random information. A protocol that accomplishes this can be regarded as a *nondeterministic* protocol for the F_{NE} function. By standard techniques in communication complexity (see [6], p. 19), the following lower bound can be obtained on the amount of communication required by Alice and Bob in order to achieve this.

Lemma 2. Any nondeterministic *classical* protocol for computing the function F_{NE} requires at least $\log_2(n)$ bits of communication.

The above lemma will be used to prove the following theorem.

Theorem 2. For the quantum measurement scenario $(|\Psi\rangle_{AB}, M_A, M_B)$, where $|\Psi\rangle_{AB}=(1/\sqrt{2})(|00\rangle+|11\rangle)$ and where M_A, M_B are each of the set of all von Neumann measurements, classical entanglement simulation is impossible if the number of bits of communication is absolutely bounded by a constant and the number of bits of local hidden variables is also a finite constant.

Proof. We will show that any protocol for classical entanglement simulation that uses a constant number of bits of communication (in the absolute sense) and a constant number of bits of local hidden variables can be converted into a nondeterministic protocol for F_{NE} with a constant amount of communication (independent of n), thereby contradicting lemma 2. First, note that a finite amount of prior shared randomness can always be simulated by a finite amount of communication at the start of the protocol (to establish the shared randomness). Thus we can suppose, without loss of generality, that the protocol for classical entanglement simulation uses no prior shared randomness.

Consider the restricted set of measurements, where M_A and M_B each consist of all measurements with respect to the operators of the form $R(x/2^n)$ [R is defined in Eq. (1)], where $x\in\{0,1\}^n$ is an n-bit binary number. Note that if the protocol for entanglement simulation is given x and y as inputs, then the resulting output bits of Alice and Bob, call them a and b, satisfy $\Pr[a=b]=\cos^2[\pi(x-y)/2^n]$. It follows that $\Pr[a\neq b]=0$ if $x=y$ and $\Pr[a\neq b]>0$ if $x\neq y$. Therefore, if at the end of the protocol Alice sends her bit a to Bob (increasing the communication cost of the protocol by one bit) and then Bob outputs $a\oplus b$, the result is a nondeterministic protocol for computing F_{NE} with a constant number of bits of communication (independent of n), contradicting lemma 2.■

In contrast to lemma 2, we note the following.

Lemma 3. There is a nondeterministic quantum protocol for F_{NE} where the communication cost is exactly one qubit.

Proof. The idea is for Alice to create the state $\cos(\pi x/2^n)|0\rangle+\sin(\pi x/2^n)|1\rangle$ and send it to Bob. Then Bob measures with respect to the operator $R(y)$. It is straightforward to calculate that the outcome of Bob's measurement b satisfies $\Pr[b=1]=0$ if $x=y$ and $\Pr[b=1]>0$ if $x\neq y$. ■

Comparing lemmas 2 and 3, there is a 1 vs $\log_2(n)$ quantum vs classical gap for the nondeterministic communication complexity of F_{NE}. This is noteworthy since it is a case where even an exponential increase in the amount of communication permitted is not sufficient for a classical protocol to simulate a quantum protocol. See [7] for other results about nondeterministic communication complexity in the quantum case.

III. CASE OF SEVERAL BELL STATES

In this section, we shall exhibit a protocol that generalizes theorem 1 as follows.

Theorem 3. Suppose that Alice and Bob must simulate the classical teleportation of a state belonging to a 2^n-dimensional Hilbert space and suppose that Bob must carry out an arbitrary POVM. Or suppose that Alice and Bob must simulate carrying out arbitrary POVM's on ebits (that is, an entangled state belonging to the tensor product of two 2^n-dimensional Hilbert spaces). Both simulations can be realized by communicating on average fewer than $(3n+6)2^n+2$ bits. Specifically, we shall exhibit a protocol in which alice sends Bob on average fewer than $(3n+6)2^n$ bits and Bob sends Alice on average fewer than 2 bits of communication.

We now proceed with a general proof of theorem 3. We start with a discussion of how much classical communication is necessary for approximate simulation of quantum communication.

Lemma 4. Consider the problem of classical teleportation in which Alice is given a quantum state $|\Psi\rangle$ belonging to a 2^n-dimensional Hilbert space (i.e., Alice is given n qubits) and Bob is given an arbitrary POVM $x=\{B_l\}$ where B_l are the POVM elements. Suppose Alice sends Bob $(m+1)2^{n+1}$ (with $m\geqslant n/2$) bits of classical information about state $|\Psi\rangle$. With this classical information Bob can calculate an approximation $P^m(l)$ to the true probability $P(l)=\langle\Psi|B_l|\Psi\rangle$ that his measurement yields outcome l. Alice can choose the bits she sends to Bob, and Bob can use an algorithm, such that the approximate probabilities sum to 1 [$\Sigma_l P^m(l)=1$] and satisfy the constraint

$$|P(l)-P^m(l)|\leqslant\alpha^m\mathrm{Tr}(B_l), \tag{6}$$

where $\mathrm{Tr}(B_l)$ is the trace of the POVM element B_l and α^m is bounded by

$$\alpha^m<2^{n/2-m+1}. \tag{7}$$

An equivalent formulation of Eq. (6) is that the information provided by Alice enables Bob to define two bounds

$$P^m_{\min}(l)=\max\{0,P^m(l)-\alpha^m\mathrm{Tr}(B_l)\}, \tag{8}$$

$$P^m_{\max}(l)=\min\{1,P^m(l)+\alpha^m\mathrm{Tr}(B_l)\}, \tag{9}$$

such that he knows with certainty that $P(l)$ belongs to the interval

$$P(l)\in[P^m_{\min}(l),P^m_{\max}(l)]. \tag{10}$$

This interval has the property that as m increases the interval shrinks:

$$0\leqslant P^m_{\min}(l)\leqslant P^{m+1}_{\min}(l), \tag{11}$$

$$1\geqslant P^m_{\max}(l)\geqslant P^{m+1}_{\max}(l). \tag{12}$$

Proof. Let us choose an arbitrary basis $|j\rangle$ of the Hilbert space. This basis is known to both Alice and Bob. In this basis, the state $|\Psi\rangle$ can be written as

$$|\Psi\rangle=\sum_{j=1}^{2^n}[X(j)+iY(j)]|j\rangle, \tag{13}$$

where $X(j)$ and $Y(j)$ are real numbers. We can write them as

$$X(j)=(-1)^{x_0(j)}\sum_{r=1}^{\infty} x_r(j)2^{-r},$$
$$Y(j)=(-1)^{y_0(j)}\sum_{r=1}^{\infty} y_r(j)2^{-r}, \tag{14}$$

were $x_r(j),y_r(j)\in\{0,1\}$.

We shall suppose that the $(m+1)2^{n+1}$ bits of information about $|\Psi\rangle$ sent by Alice are the values of $x_r(j),y_r(j)$ for all j and for $0\leqslant r\leqslant m$. Bob then knows the coefficients $X(j)$, $Y(j)$ with finite precision. Denote the part of $X(j)$ and $Y(j)$ which is known to Bob by

$$X^m(j)=(-1)^{x_0(j)}\left(\sum_{r=1}^{m} x_r(j)2^{-r}+2^{-m-1}\right),$$
$$Y^m(j)=(-1)^{y_0(j)}\left(\sum_{r=1}^{m} y_r(j)2^{-r}+2^{-m-1}\right). \tag{15}$$

We then have

$$|X(j)-X^m(j)|\leqslant 2^{-m-1},$$
$$|Y(j)-Y^m(j)|\leqslant 2^{-m-1}. \tag{16}$$

Denote Bob's estimate of the state Ψ by

$$|\Psi^m\rangle=\sum_{j=1}^{2^n}[X^m(j)+iY^m(j)]|j\rangle. \tag{17}$$

We can write the true state as

$$|\Psi\rangle=|\Psi^m\rangle+|\Delta\Psi^m\rangle. \tag{18}$$

Bob's uncertainty can be measured by

$$\langle\Delta\Psi^m|\Delta\Psi^m\rangle=\sum_{j=1}^{2^n}[X(j)-X^m(j)]^2+[Y(j)-Y^m(j)]^2$$
$$\leqslant 2^{n-2m-1}. \tag{19}$$

For this inequality to be informative, it is necessary that $m\geqslant n/2$.

Bob's estimate for the probability $P(l)$ is

$$P^m(l)=\langle\Psi^m|B_l|\Psi^m\rangle. \tag{20}$$

Let us write Bob's POVM elements as

$$B_l=\mathrm{Tr}(B_l)|\tilde{\beta}_l\rangle\langle\tilde{\beta}_l|, \tag{21}$$

where $|\tilde{\beta}_l\rangle$ is a normalized state. We then have

$$P(l)-P^m(l)=\mathrm{Tr}(B_l)[2\,\mathrm{Re}\langle\Psi^m|\tilde{\beta}_l\rangle\langle\tilde{\beta}_l|\Delta\Psi^m\rangle$$
$$+|\langle\tilde{\beta}_l|\Delta\Psi^m\rangle|^2], \tag{22}$$

which we can bound by

$$|P(l)-P^m(l)|\leqslant\mathrm{Tr}(B_l)[2|\langle\Psi^m|\tilde{\beta}_l\rangle||\langle\tilde{\beta}_l|\Delta\Psi^m\rangle|$$
$$+|\langle\tilde{\beta}_l|\Delta\Psi^m\rangle|^2]$$
$$\leqslant\mathrm{Tr}(B_l)[2|\langle\tilde{\beta}_l|\Delta\Psi^m\rangle|+|\langle\tilde{\beta}_l|\Delta\Psi^m\rangle|^2]$$
$$\leqslant\mathrm{Tr}(B_l)[2\sqrt{\langle\Delta\Psi^m|\Delta\Psi^m\rangle}+\langle\Delta\Psi^m|\Delta\Psi^m\rangle]$$
$$\leqslant\mathrm{Tr}(B_l)[2^{n/2-m+1/2}+2^{n-2m-1}]$$
$$\leqslant\mathrm{Tr}(B_l)\alpha^m, \tag{23}$$

where we have used that $2^{n/2-m+1/2}+2^{n-2m-1}<\alpha^m$ if $m\geqslant n/2$. This proves Eq. (6)

Let us now prove the monotonicity properties (11) and (12). To this end we compute the difference between successive estimates $|P^m(l)-P^{m+1}(l)|$. Define the quantities $\delta X^{m+1}(j)$, $\delta Y^{m+1}(j)$ by

$$\delta X^{m+1}(j)=X^m(j)-X^{m+1}(j),$$
$$\delta Y^{m+1}(j)=Y^m(j)-Y^{m+1}(j). \tag{24}$$

We have

$$|\delta X^{m+1}(j)|\leqslant 2^{-m-2},\quad |\delta Y^{m+1}(j)|\leqslant 2^{-m-2}. \tag{25}$$

Then we define the difference of estimated state for two successive values of m:

$$|\delta\Psi^m\rangle=|\Psi^m\rangle-|\Psi^{m+1}\rangle=\sum_{j=1}^{2^n}[\delta X^{m+1}(j)+i\,\delta Y^{m+1}(j)]|j\rangle. \tag{26}$$

The difference between successive estimates decreases as

$$\langle\delta\Psi^{m+1}|\delta\Psi^{m+1}\rangle=\sum_{j=1}^{2^n}[\delta X^{m+1}(j)]^2+[\delta Y^{m+1}(j)]^2$$
$$\leqslant 2^{n-2m-3}. \tag{27}$$

Finally, we have

$$|P^m(l)-P^{m+1}(l)|=\mathrm{Tr}(B_l)|2\,\mathrm{Re}\langle\Psi^m|\tilde{\beta}_l\rangle\langle\tilde{\beta}_l|\delta\Psi^{m+1}\rangle$$
$$+|\langle\tilde{\beta}_l|\delta\Psi^{m+1}\rangle|^2|$$
$$\leqslant\mathrm{Tr}(B_l)[2\sqrt{\langle\delta\Psi^{m+1}|\delta\Psi^{m+1}\rangle}$$
$$+\langle\delta\Psi^{m+1}|\delta\Psi^{m+1}\rangle]$$
$$\leqslant\mathrm{Tr}(B_l)[2^{n/2-m-1/2}+2^{n-2m-3}]$$
$$\leqslant\mathrm{Tr}(B_l)\alpha^{m+1}\quad(\text{if } m\geqslant n/2-1), \tag{28}$$

which together with the definitions (8) and (9) implies Eqs. (11) and (12). ■

We now turn to the proof of theorem 3. The two complications with respect to theorem 1 are that the state is described by many parameters and not one angle x and that Bob may have more than two outcomes between which to

choose since his POVM may have more than two outcomes. These two complications lead to the more intricate protocol given below.

Proof of theorem. Note that the second part of the theorem (dealing with simulating measurements on ebits) follows directly from the first part of the theorem (dealing with the simulating the transmission of qubits) in view of the relationships between classical teleportation and classical entanglement simulation (discussed at the end of Sec. I). Hence we consider only the first part dealing with the simulation of quantum communication.

The protocol used by Alice and Bob consists of a series of rounds which we label by K. Alice's role during each round is simple to describe. She starts the round by sending Bob some information about the state $|\Psi\rangle$. Specifically, during the first round ($K=1$) this information consists of the values of the coefficients $x_r(j), y_r(j)$ defined in Eqs. (14) for $r=0,\ldots,3n/2+2$ and all values of j ($j=1,\ldots,2^n$). During the next rounds ($K=2,3,\ldots$), this information consists of the values of the coefficients $x_{3n/2+K+1}(j), y_{3n/2+K+1}(j)$ for $j=1,\ldots,2^n$.

Upon receiving this information, Bob will carry out a computation (which we describe below) and reaches one of two conclusions. One possibility is that he is able to choose an outcome l for his measurement. The second possibility is that he is unable to choose an outcome l in which case he needs more information about $|\Psi\rangle$. Thus the end of the round consists of Bob sending Alice one bit telling her whether or not he needs more information about $|\Psi\rangle$. If Bob does not need more information, then the protocol terminates since Bob has chosen an outcome for his measurement. If Bob needs more information, then they both increments K by 1 and the next round starts.

The reason why the first round differs slightly from the next rounds is that Bob needs a large amount of initial information before he can start trying to choose an outcome. If this first try does not succeed, then only small additional amounts of information are necessary for Bob to try again to choose an outcome. Mathematically, the necessity for the large amount of initial information is expressed in Eqs. (31) and (38) below which are nontrivial inequalities only when a sufficient amount of information has been transmitted by Alice to Bob.

We now describe the computation carried out by Bob. Recall that with the information sent to him by Alice, Bob can construct the approximation $|\Psi^{3n/2+K+1}\rangle$ to the true state $|\Psi\rangle$ [defined in Eq. (17)]. Using this approximate state, he knows that the true probability $P(l)$ of outcome l is comprised between $P_{\min}^{3n/2+K+1}(l)$ and $P_{\max}^{3n/2+K+1}(l)$; see Eq. (10). It is convenient for Bob to reexpress this approximate knowledge of the true probabilities in terms of set of subintervals $I^K(l), R^K$ of the unit interval [0, 1]. Bob's strategy will then be simply expressed in terms of these intervals.

To define these subintervals we introduce the following notations:

$$\Delta^m(l)=P_{\min}^m(l)-P_{\min}^{m-1}(l) \tag{29}$$

(note that $\Delta^m(l)\geqslant 0$ [see Eq. (11)]) and

$$T^m=\sum_{l=1}^{L} P_{\min}^m(l), \tag{30}$$

where L is the number of outcomes of Bob's POVM $\{B_l\}$, $l=1,\ldots,L$. We have the following property:

$$T^m \geqslant \sum_{l=1}^{L} P^m(l)-\alpha^m \mathrm{Tr}(B_l)=1-\alpha^m \sum_{l=1}^{L} \mathrm{Tr}(B_l)=1-\alpha^m 2^n$$
$$\geqslant 1-2^{3n/2-m+1} \quad (\text{if } m\geqslant n/2+1), \tag{31}$$

which follow from Eqs. (8) and (7).

The subintervals are defined for $K=1$ by

$$I^1(1)=[0,P_{\min}^{3n/2+2}(1)[,$$

$$I^1(l)=\left[\sum_{l'=1}^{l-1} P_{\min}^{3n/2+2}(l'), \sum_{l'=1}^{l} P_{\min}^{3n/2+2}(l')\right[, \tag{32}$$

and for $K=2,3,\ldots,\infty$ by

$$I^K(1)=[T^{3n/2+K}, T^{3n/2+K}+\Delta^{3n/2+K+1}(1)[,$$

$$I^K(l)=\left[T^{3n/2+K}+\sum_{l'=1}^{l-1} \Delta^{3n/2+K+1}(l'),\right.$$

$$\left. T^{3n/2+K}+\sum_{l'=1}^{l} \Delta^{3n/2+K+1}(l')\right[. \tag{33}$$

We also define the subintervals

$$R^K=[T^{3n/2+K+1},\ 1[, \quad K=1,2,\ldots,\infty. \tag{34}$$

These subintervals have several properties which follow directly from Eqs. (6), (8), (11), (10), and (31).

(i) The intervals $I^K(l)$ ($K=1,\ldots,\infty$ and $l=1,\ldots,L$) are disjoint.

(ii) The intervals $I^{K'}(l)$ ($K'=1,\ldots,K$ and $l=1,\ldots,L$) and the interval R^K are disjoint.

(iii) The intervals $I^{K'}(l)$ ($K'=K+1,\ldots,\infty$ and $l=1,\ldots,L$) all belong to the interval R^K.

(iv) The union of the intervals $I^{K'}(l)$ ($K'=1,\ldots,K$ and $l=1,\ldots,L$) and of R^K is the unit interval:

$$\left(\bigcup_{K'=1}^{K}\bigcup_{l=1}^{L} I^{K'}(l)\right)\cup R^K=[0,1[. \tag{35}$$

(v) The union of the intervals $I^{K+1}(l)$ (K fixed and $l=1,\ldots,L$) and of R^{K+1} is the interval R^K:

$$\left(\bigcup_{l=1}^{L} I^{K+1}(l)\right)\cup R^{K+1}=R^K. \tag{36}$$

(vi) The union of all the intervals $I^K(l)$ ($K=1,\ldots,\infty$ and $l=1,\ldots,L$) is the unit interval

$$\bigcup_{K=1}^{\infty}\bigcup_{l=1}^{L} I^K(l) = [0,1[. \tag{37}$$

(vii) The length of the interval R^K is

$$\mu(R^K) = 1 - T^{3n/2+K+1} \leqslant 2^{-K}. \tag{38}$$

(viii) The length of the union of the intervals $I^K(l)$ ($K=1,\ldots,\infty$ and l fixed) is $P(l)$:

$$\mu\left(\bigcup_{K=1}^{\infty} I^K(l)\right) = \sum_{K=1}^{\infty} \mu(I^K(l)) = P(l). \tag{39}$$

Bob's strategy is now simple to describe. Initially, before Alice sends him any information, he chooses a random number r uniformly distributed in the interval $[0,1[$. He then carries out the following operations at each round.

Bob's strategy. At round K, he checks whether r belongs to $I^K(l)$. If so, he outputs outcome l and tells Alice he does not need any more information. On the other hand, if at round K,r belongs to R^K, he tells Alice he needs more information.

Because of properties (i)–(v), Bob is sure that at round K, r will belong to one of the intervals $I^K(l)$ or to R^K. Hence the strategy described above is well defined. Furthermore, in view of properties (i) and (vi), r belongs to one and only one interval $I^K(l)$: hence, the protocol will eventually terminate.

To calculate the probability that Bob outputs outcome l, note that this occurs if and only if r belongs to one of the intervals $I^K(l)$ ($K=1,\ldots,\infty$ and l fixed). The probability that Bob outputs outcome l is therefore equal to $\mu(\cup_{K=1}^{\infty} I^K(l))$. From property (viii) this is equal to $P(l) = \langle\Psi|B_l|\Psi\rangle$, as required.

Finally, we compute the mean amount of communication required by the above protocol. Note that the first round always occurs. The amount of communication during this round, denoted C^1, consists of $(3n/2+2)2^{n+1}$ bits of communication sent by Alice to Bob [namely, the values of the coefficients $x_m(j), y_m(j)$, $j=1,\ldots,2^n$, $m=1,\ldots,3n/2+2$] and of 1 bit of communication sent by Bob to Alice (telling her whether he could choose an outcome or not).

The subsequent rounds $K \geqslant 2$ do not always occur. Round K only occurs if Bob was not able to choose an outcome before round K—that is, if r does not belong to any of the intervals $I^{K'}(l)$ ($K'=1,\ldots,K-1$ and $l=1,\ldots,L$). Using property (iv) this can be reexpressed as the fact that round K occurs if and only if r belongs to R^{K-1}. The probability that round K occurs is therefore

$$P(\text{round } K \text{ occurs}) = \mu(R^{K-1}) = 1 - T^{3n/2+K} \leqslant 2^{-K+1}. \tag{40}$$

During rounds $K \geqslant 2$, a certain amount of communication occurs, always the same, denoted C'. This consists of 2^{n+1} bits sent by Alice to Bob [namely, the values of the coefficients $x_{3n/2+K+1}(j), y_{3n/2+K+1}(j)$, $j=1,\ldots,2^n$] and of 1 bit of communication sent by Bob to Alice (telling her whether he could choose an outcome or not).

The average amount of communication is therefore

$$\overline{C} = C^1 + \sum_{K=2}^{\infty} P(\text{round } K \text{ occurs})\ C' \leqslant C^1 + C' \sum_{K=2}^{\infty} 2^{-K+1} = C^1 + C'. \tag{41}$$

The average amount of communication therefore consists of fewer than $(3n+6)2^n$ bits sent by Alice to Bob and fewer than 2 bits sent by Bob to Alice. ■

IV. CONCLUSION

We have shown in theorem 2 that perfect classical simulation of quantum communication and entanglement is impossible if the amount of communication is bounded and the two parties share a finite number of random bits. Indeed, with bounded communication and finite prior shared randomness, only approximate simulations of quantum communication are possible.

However, if we give the parties something slightly more powerful than finite communication, then perfect simulation becomes possible. One possibility is for Alice and Bob to have an *a priori* supply of an infinite number of shared random bits as in [1] (this could, for instance, be established by having an infinite conversation prior to the start of the simulation protocol itself). A second possibility, considered in theorem 3, is for Alice and Bob to share no prior randomness and to require that the amount of communication is only finite on average. In this case, the amount of communication varies from one simulation to another, and can sometimes be arbitrarily large.

In all cases, as the number n of qubits or ebits that must be simulated increases, the amount of classical communication required grows exponentially with n.

ACKNOWLEDGMENTS

Part of this work was completed at the workshop on Complexity, Computation and the Physics of Information, Isaac Newton Institute, Cambridge, UK. We would like to thank Andreas Winter for help with the Appendix. S.M. and N.C. acknowledge financial support from the European Science Foundation and from European Union project EQUIP (Contract No. IST-1999-11063). R.C. is supported in part by Canada's NSERC. D.B. is supported by the U.S. Army Research Office under Contract No. DAAG55-98-1-0371.

APPENDIX: SHARED RANDOMNESS ACHIEVED BY COMMUNICATION

In this appendix, we sketch a proof that the mutual information I between Alice and Bob's outputs (if they share no prior randomness) is bounded from above by the total number of bits exchanged in an arbitrary number of rounds of two-way communication. Assume that, initially, Alice and Bob have each a random variable denoted, respectively, as A_0 and B_0 (this represents a local source of randomness), but they share no information, i.e., $I(A_0:B_0)=0$. Then Alice and Bob communicate via an arbitrary number of rounds of two-way communication. The first round consists of Alice send-

ing B_1 to Bob, followed by Bob sending A_1 to Alice. So B_1 is a function of A_0, while A_1 is a function of B_0 and B_1. In general, the ith round consists in Bob receiving B_i followed by Alice receiving A_i. Again, B_i is a function of $A_0,\ldots,A_{i-1}$, and A_i is a function of $B_0,\ldots,B_i$. Assume that this protocol terminates after N rounds. Alice then outputs $X=X(A_0,\ldots,A_N)$ which is a function of all the information Alice has, and similarly Bob outputs $Y=Y(B_0,\ldots,B_N)$.

We first note that the data processing inequality implies that

$$I(X:Y)\leqslant I(A_0,\ldots,A_N:B_0,\ldots,B_N). \tag{A1}$$

We now bound the right-hand side (RHS) of this equation by

$$\begin{aligned} &I(A_0,\ldots,A_N:B_0,\ldots,B_N)\\ &\quad=I(A_0:B_0)+H(A_1,\ldots,A_N|A_0)+H(B_1,\ldots,B_N|B_0)\\ &\qquad-H(A_1,\ldots,A_N,B_1,\ldots,B_N|A_0,B_0). \end{aligned} \tag{A2}$$

The first term of the RHS of Eq. (A2) is zero since there is no initial shared randomness. The second term of the RHS of Eq. (A2) measures the amount of randomness received by Alice during the N rounds in addition to the initial randomness A_0. It is simply bounded from above by the number of bits of backward communication C_b since

$$\begin{aligned} H(A_1,\ldots,A_N|A_0)\leqslant H(A_1,\ldots,A_N)\leqslant H(A_1)+\cdots+H(A_N)\\ =C_b. \end{aligned} \tag{A3}$$

Similarly, the third term on the RHS of Eq. (A2) is bounded from above by the number of bits of forward communication C_f. Finally, using the chain rule for entropies, the fourth term on the RHS of Eq. (A2) can be reexpressed as

$$\begin{aligned} &H(A_1,B_1|A_0,B_0)+H(A_2,B_2|A_0,A_1,B_0,B_1)+\cdots\\ &+H(A_N,B_N|A_0,\ldots,A_{N-1},B_0,\ldots,B_{N-1}). \end{aligned} \tag{A4}$$

The ith term in this sum can be written as

$$\begin{aligned} &H(A_i,B_i|A_0,\ldots A_{i-1},B_0,\ldots B_{i-1})\\ &\quad=H(B_i|A_0,\ldots,A_{i-1},B_0,\ldots,B_{i-1})\\ &\qquad+H(A_i|A_0,\ldots,A_{i-1},B_0,\ldots,B_i). \end{aligned}$$

These two conditional entropies vanish since B_i depends on $A_0,\ldots,A_{i-1}$, and A_i depends on $B_0,\ldots,B_i$. Thus the fourth term on the RHS of Eq. (A2) is zero. As a consequence, we have

$$I(X:Y)\leqslant C_f+C_b, \tag{A5}$$

as asserted above.

[1] G. Brassard, R. Cleve, and A. Tapp, Phys. Rev. Lett. **83**, 1874 (1999).

[2] N. Cerf, N. Gisin, and S. Massar, Phys. Rev. Lett. **84**, 2521 (2000).

[3] J. S. Bell, Physics (Long Island City, N.Y.) **1**, 195 (1964).

[4] M. Steiner, Phys. Lett. A **270**, 239 (2000).

[5] H. Buhrman, R. Cleve, and A. Wigderson, in *Proceedings of the 30th Annual ACM Symposium on Theory of Computing* (ACM, New York, 1998), p. 63; also available as e-print quant-ph/9702040.

[6] E. Kushilevitz and N. Nisan, *Communication Complexity* (Cambridge University Press, Cambridge, England, 1998).

[7] R. de Wolf, in *Proceedings of the 15th Annual Conference on Computational Complexity* (IEEE, Los Alamitos, 2000), p. 271; also available as e-print cs.CC/0001014.