A lower bound on entanglement-assisted quantum communication complexity

Ashley Montanaro
Department of Computer Science, University of Bristol, Bristol, BS8 1UB, U.K.[*]

Andreas Winter
Department of Mathematics, University of Bristol, Bristol BS8 1TW, U.K.[†]

(Dated: October 19, 2006)

We prove a general lower bound on the bounded-error entanglement-assisted quantum communication complexity of Boolean functions. The bound is based on the concept that any classical or quantum protocol to evaluate a function on distributed inputs can be turned into a quantum communication protocol. As an application of this bound, we give a very simple proof of the statement that almost all Boolean functions on $n + n$ bits have linear communication complexity, even in the presence of unlimited entanglement.

I. INTRODUCTION

Consider a total Boolean function $f : \{0,1\}^n \times \{0,1\}^n \mapsto \{0,1\}$. The *quantum communication complexity* of f is defined to be the minimum number of qubits required to be transmitted between two parties (Alice and Bob) for them to compute $f(x, y)$ for any two n-bit inputs x, y, given that Alice starts out with x and Bob with y. This number is clearly upper bounded by n, but for some functions may be considerably lower. Alice and Bob may be allowed some probability of error ϵ, and may be allowed to share an entangled state before they start their protocol. We will assume that Bob has to output the result. (See [25] and [17] for excellent introductions to quantum and classical communication complexity, respectively.)

Some functions are known to have a quantum communication complexity lower than their classical communication complexity (for example, a bounded-error protocol for the disjointness function $f(x, y) = 1 \Leftrightarrow |x \wedge y| = 0$ requires $\Omega(n)$ bits of classical communication, but only $\Theta(\sqrt{n})$ qubits of quantum communication [1, 23]), but it is still open whether the quantum communication complexity of total functions can ever be exponentially smaller than the classical communication complexity. It is therefore of interest to produce lower bounds on quantum communication complexity. In this context, the model with prior entanglement is less well understood; although there are strong bounds known for some classes of functions [5, 23], there are few general lower bounds [4]. It has been shown [9, 10] that sharing entanglement may significantly reduce the communication cost of computing a partial function (where there is a promise on the input), but it is unknown whether a similar result may hold for total functions.

In this paper, we develop an elegant result of Cleve et al. that relates computation to communication. Cleve et al. showed [5] that, if Alice and Bob have access to a protocol to exactly compute the inner product function $IP(x, y) = \sum_i x_i y_i \pmod 2$, then this can be used to produce a quantum protocol that communicates Alice's input x to Bob. They used this to show that IP cannot be computed (exactly and without prior entanglement) by sending fewer than n bits from Alice to Bob. Similar results hold for the bounded-error case and with prior entanglement.

[*]Electronic address: montanar@cs.bris.ac.uk
[†]Electronic address: a.j.winter@bristol.ac.uk

We show that a weaker form of this result can be extended to *all* Boolean functions. The extension leads to the development of a new complexity measure for Boolean functions: *communication capacity*. Given a Boolean function $f(x, y)$, we define the communication capacity of f as the maximum number of bits which the execution of a protocol to compute f allows Alice to communicate to Bob (in an asymptotic sense). This is a concept which has no classical analogue, and which can be shown to give a lower bound on the quantum communication complexity of f (with or without entanglement).

The lower bound we obtain turns out to be a generalisation of a bound obtained by Klauck [15] on quantum communication complexity in the model without entanglement. The result here can thus be seen as extending Klauck's bound to the model of entanglement-assisted quantum communication, and giving it a satisfying operational interpretation. As our bound also holds for classical communication complexity, it fits into the framework of results using ideas from quantum information to say something about classical computation.

We will use the standard notation $Q_E(f)$ to denote the quantum communication complexity of f in the case where the protocol must be exact, $Q_\epsilon(f)$ the complexity where Alice and Bob are allowed to err with probability $\epsilon < 1/2$, and $Q_2(f)$ the complexity in the case where $\epsilon = 1/3$. In all three cases, Alice and Bob's initial state is separable; $Q^*_E(f)$, $Q^*_\epsilon(f)$ and $Q^*_2(f)$ will represent the equivalent quantities in the case where they are allowed to share an arbitrary initial entangled state.

As is usual in computational complexity, we would expect most functions to have "high" quantum communication complexity. Kremer showed [16] by a counting argument that a random function f has $Q_2(f) \geq n/2$ (and thus $Q_E(f) \geq n/2$). Buhrman and de Wolf extended Kremer's methods to show that, for all f, $Q^*_E(f) \geq (\log \mathrm{rank}(f))/2$ [4] (an equivalent result is shown in section 6.4.2 of [20]). As almost all Boolean matrices have full rank, this shows that for almost all f, $Q^*_E(f) \geq n/2$. Very recently, Gavinsky, Kempe and de Wolf [8] have shown the final remaining case: for almost all f, $Q^*_2(f) = \Omega(n)$. Their technique was to relate quantum communication protocols to quantum fingerprinting protocols, and then to show a relationship between quantum fingerprinting and some well-studied concepts from classical computational learning theory. This result was shown independently by Linial and Shraibman [18]; their paper also extends the well-known discrepancy lower bound to the model of quantum communication with entanglement.

As an application of our communication capacity technique, we reprove the result that for almost all f, $Q^*_2(f) = \Omega(n)$. The proof is of a quite different character and of (arguably) a more "quantum" nature, as it is based on showing that the entropy of almost all density matrices produced in a certain random way is high.

A. Notation

We will use M to denote the square communication matrix of f (where M_{xy} is equal to $(-1)^{f(x,y)}$). $H(v)$ will denote the Shannon entropy of a vector v, and $S(\rho)$ the von Neumann entropy of a density matrix ρ ($S(\rho) = -\mathrm{tr}(\rho \log \rho)$). All logarithms will be taken to base 2.

II. TURNING ANY DISTRIBUTED FUNCTION INTO A COMMUNICATION PROTOCOL

In this section, we will describe a protocol (which is a simple extension of the protocol in [5] for IP) that allows any protocol for evaluating a distributed function to be turned into a communication protocol. However, for some functions, the communication will be considerably more inefficient than IP allows (Alice may only be able to send $\ll n$ bits to Bob).

A. Exact protocols

Say Alice and Bob have access to a classical or quantum protocol that computes $f(x,y)$ exactly. We express this as a unitary P that performs the following action.

$$P|x\rangle_A|y\rangle_B|0\rangle_B|a\rangle_{AB} = |x\rangle_A|y\rangle_B|f(x,y)\rangle_B|a'\rangle_{AB} \tag{1}$$

where $|a\rangle$, $|a'\rangle$ are arbitrary (and possibly entangled) ancilla states shared by Alice and Bob. Note that, as P does not modify the first two registers, we may decompose it as follows:

$$P = \sum_{x,y} |x\rangle\langle x|_A \otimes |y\rangle\langle y|_B \otimes U_{xy} \tag{2}$$

for some unitary U_{xy} acting only on the last two registers. Following [5], we will turn this into a "clean" protocol P' by giving Bob an additional qubit to copy the answer into, then running the protocol backwards to uncompute the "junk" $|a'\rangle$. The steps of the clean protocol are thus

$$\begin{aligned}
&\text{(i)} && |x\rangle_A|y\rangle_B|0\rangle_B|0\rangle_B|a\rangle_{AB} \\
&\text{(ii)} \;\rightarrow\; && |x\rangle_A|y\rangle_B|f(x,y)\rangle_B|0\rangle_B|a'\rangle_{AB} \\
&\text{(iii)} \;\rightarrow\; && |x\rangle_A|y\rangle_B|f(x,y)\rangle_B|f(x,y)\rangle_B|a'\rangle_{AB} \\
&\text{(iv)} \;\rightarrow\; && |x\rangle_A|y\rangle_B|0\rangle_B|f(x,y)\rangle_B|a\rangle_{AB}
\end{aligned}$$

where now the fourth register contains the answer. Ignoring the third and fifth registers, which are the same at the beginning and the end of the protocol, we are left with the map

$$P'|x\rangle_A|y\rangle_B|0\rangle_B = |x\rangle_A|y\rangle_B|f(x,y)\rangle_B \tag{3}$$

Note that, if the original protocol P communicated a qubits from Alice to Bob and b qubits from Bob to Alice, the protocol P' requires $a+b$ qubits to be communicated in each direction. That is, P' sends as many qubits in the "forward" direction as the original protocol P sends in total. Now say Alice wants to communicate her input x to Bob using this protocol. They start with the following state, where (b_y) is an arbitrary probability distribution on Bob's inputs:

$$|\psi\rangle = |x\rangle_A \left(\sum_{y\in\{0,1\}^n} \sqrt{b_y}|y\rangle_B \right) \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)_B \tag{4}$$

Note that this state is separable (so we do not *require* entanglement to execute the communication protocol). After executing the clean protocol for f, they are left with

$$\begin{aligned}
P'|\psi\rangle &= |x\rangle_A \left(\frac{1}{\sqrt{2}} \sum_{y\in\{0,1\}^n} \sqrt{b_y}|y\rangle_B \left(|f(x,y)\rangle - |1-f(x,y)\rangle\right)_B \right) && (5) \\
&= |x\rangle_A \left(\sum_{y\in\{0,1\}^n} (-1)^{f(x,y)} \sqrt{b_y}|y\rangle_B \right) \frac{1}{\sqrt{2}} \left(|0\rangle - |1\rangle\right)_B && (6)
\end{aligned}$$

Ignoring the registers that remain the same throughout, Bob has the following state at the end of the protocol.

$$|\psi_x\rangle = \sum_{y\in\{0,1\}^n} (-1)^{f(x,y)} \sqrt{b_y}|y\rangle \tag{7}$$

This state provides some information about Alice's bit string x. If $\langle\psi_x|\psi_{x'}\rangle = 0$ for all $x' \neq x$ (as is the case with the protocol of [5] for IP, where Bob uses the uniform distribution on his inputs) then Bob can determine x with certainty and hence has received n bits from Alice. If this is not the case, then we can still quantify precisely how much information can be transmitted. The protocol is equivalent to Alice encoding the classical bit-string x as a state $|\psi_x\rangle$, and co-operating with Bob to send it to him. Say Alice uses a distribution (a_x) on her inputs. Then the ensemble representing what Bob eventually receives is

$$\rho = \sum_{x\in\{0,1\}^n} a_x |\psi_x\rangle\langle\psi_x| \tag{8}$$

By Holevo's theorem [12], the entropy $S(\rho)$ describes the maximum number of bits of classical information about x available to Bob by measuring ρ. And, by the Holevo-Schumacher-Westmoreland channel coding theorem for a channel with pure signal states [11], Alice and Bob can achieve this bound (in an asymptotic sense) using block coding!

Therefore, the ability to compute f exactly can be used to transmit $S(\rho)$ bits of information through a quantum channel, even though this does not hold if Alice and Bob are restricted to a classical channel. We thus define the *communication capacity* of a Boolean function f as the maximum over all probability distributions (a_x) (on Alice's inputs) and (b_y) (on Bob's inputs) of

$$S\left(\sum_{x\in\{0,1\}^n} a_x |\psi_x\rangle\langle\psi_x|\right) \text{ , where } |\psi_x\rangle = \sum_{y\in\{0,1\}^n} (-1)^{f(x,y)}\sqrt{b_y}|y\rangle \tag{9}$$

B. Bounded error protocols

In the case where Alice and Bob have access to a protocol computing f with some probability of error, Bob will not have the state $|\psi_x\rangle$ at the end of the protocol, but rather some approximation $|\psi_x^\epsilon\rangle$. We will now show that, if the error probability is small, this is in fact still sufficient to communicate a significant amount of information from Alice to Bob. As before, Alice will use a distribution (a_x) on her inputs, and Bob a distribution (b_y).

Say Alice and Bob are using a protocol P^ϵ that computes f with probability of error ϵ, where $\epsilon < 1/2$. As before, the $|x\rangle$ and $|y\rangle$ registers will be unchanged by this protocol, so we can write

$$P^\epsilon = \sum_{x,y} |x\rangle\langle x|_A \otimes |y\rangle\langle y|_B \otimes U_{xy}^\epsilon \tag{10}$$

Now let us run the protocol on the same starting state $|\psi\rangle$ as in the previous section.

$$\text{(i)} \qquad |x\rangle_A \left(\frac{1}{\sqrt{2}} \sum_{y\in\{0,1\}^n} \sqrt{b_y}|y\rangle_B |0\rangle_B (|0\rangle - |1\rangle)_B\right) |a\rangle_{AB}$$

$$\text{(ii)} \ \rightarrow \ |x\rangle_A \left(\frac{1}{\sqrt{2}} \sum_{y\in\{0,1\}^n} \sqrt{b_y}|y\rangle_B (\alpha_{xy}|0\rangle + \beta_{xy}|1\rangle)_B (|0\rangle - |1\rangle)_B\right) |a'\rangle_{AB}$$

where the effect of U_{xy}^ϵ on the "answer" qubit has been decomposed into α_{xy} and β_{xy} components. If $f(x,y) = 0$, then $|\alpha_{xy}|^2 \geq 1-\epsilon$, and thus (by unitarity) $|\beta_{xy}|^2 \leq \epsilon$; if $f(x,y) = 1$, $|\beta_{xy}|^2 \geq 1-\epsilon$ and $|\alpha_{xy}|^2 \leq \epsilon$. The ancilla register is still completely arbitrary, and in particular may be entangled with any of the other registers. Continuing the protocol, we have

$$\text{(iii)} \to |x\rangle_A \left(\frac{1}{\sqrt{2}} \sum_{y\in\{0,1\}^n} \sqrt{b_y}|y\rangle_B(\alpha_{xy}|0\rangle|0\rangle - \alpha_{xy}|0\rangle|1\rangle - \beta_{xy}|1\rangle|0\rangle + \beta_{xy}|1\rangle|1\rangle)_B \right) |a'\rangle_{AB} \tag{11}$$

$$\begin{aligned}\text{(iv)} \to |x\rangle_A \Bigg(\frac{1}{\sqrt{2}} \sum_{y\in\{0,1\}^n} & \sqrt{b_y}|y\rangle_B(\alpha_{xy}(\alpha^*_{xy}|0\rangle + \gamma^*_{xy}|1\rangle))|0\rangle - \alpha_{xy}(\alpha^*_{xy}|0\rangle + \gamma^*_{xy}|1\rangle))|1\rangle \\ & - \beta_{xy}(\beta^*_{xy}|0\rangle + \delta^*_{xy}|1\rangle))|0\rangle + \beta_{xy}(\beta^*_{xy}|0\rangle + \delta^*_{xy}|1\rangle))|1\rangle)_B \Bigg) |a\rangle_{AB}\end{aligned} \tag{12}$$

$$= |x\rangle_A \left(\frac{1}{\sqrt{2}} \sum_{y\in\{0,1\}^n} \sqrt{b_y}|y\rangle_B((\alpha_{xy}\alpha^*_{xy} - \beta_{xy}\beta^*_{xy})|0\rangle + (\alpha_{xy}\gamma^*_{xy} - \beta_{xy}\delta^*_{xy})|1\rangle)_B(|0\rangle - |1\rangle)_B \right)|a\rangle_{AB} \tag{13}$$

where we introduce γ^*_{xy} and δ^*_{xy} as arbitrary elements of $(U^\epsilon_{xy})^\dagger$, subject only to the constraint that U^ϵ_{xy} be unitary. We may now remove registers that end the protocol unchanged and rewrite Bob's final state as

$$|\psi^\epsilon_x\rangle = \sum_{y\in\{0,1\}^n} \sqrt{b_y}|y\rangle \left((|\alpha_{xy}|^2 - |\beta_{xy}|^2)|0\rangle + (\alpha_{xy}\gamma^*_{xy} - \beta_{xy}\delta^*_{xy})|1\rangle \right) \tag{14}$$

Now, if $f(x,y) = 0$, then $|\alpha_{xy}|^2 - |\beta_{xy}|^2 \geq 1 - 2\epsilon > 0$, whereas if $f(x,y) = 1$, $|\alpha_{xy}|^2 - |\beta_{xy}|^2 \leq 2\epsilon - 1 < 0$. We may therefore write

$$|\psi^\epsilon_x\rangle = \sum_{y\in\{0,1\}^n} \sqrt{b_y}|y\rangle \left((-1)^{f(x,y)} \cos\theta_{xy}|0\rangle + e^{i\phi_{xy}} \sin\theta_{xy}|1\rangle \right) \tag{15}$$

where θ_{xy} is real with $\cos\theta_{xy} \geq 1 - 2\epsilon$, and ϕ_{xy} is an arbitrary phase. Crucially, the form of these states is quite restricted and close to the original $|\psi_x\rangle$. In fact, it is clear that

$$|(\langle\psi_x|\langle 0|)|\psi^\epsilon_x\rangle|^2 \geq (1-2\epsilon)^2 \tag{16}$$

Set $\rho^\epsilon = \sum_{x\in\{0,1\}^n} a_x|\psi^\epsilon_x\rangle\langle\psi^\epsilon_x|$. We will compare this to the state $\rho' = \sum_{x\in\{0,1\}^n} a_x|\psi_x\rangle|0\rangle\langle\psi_x|\langle 0|$ (where of course $S(\rho') = S(\rho)$). We have

$$\|\rho' - \rho^\epsilon\|_1 \leq 2\sqrt{1-(1-2\epsilon)^2} \leq 4\sqrt{\epsilon} \tag{17}$$

We will use Fannes' inequality [6] to show that $S(\rho^\epsilon) \approx S(\rho)$. Define the function

$$\eta_0(x) = \begin{cases} -x\log x & \text{for } x \leq 1/e \\ 1/e\log e & \text{for } x > 1/e \end{cases} \tag{18}$$

Then Fannes' inequality gives that

$$S(\rho^\epsilon) \geq S(\rho) - 4\sqrt{\epsilon}n - \log\eta_0(4\sqrt{\epsilon}) \tag{19}$$

C. Communication complexity lower bounds from communication capacity

A lower bound for the communication capacity of a function f can be written down in terms of its communication matrix M as follows. As before, set

$$\rho = \sum_{x\in\{0,1\}^n} a_x|\psi_x\rangle\langle\psi_x| \text{ for } |\psi_x\rangle = \sum_{y\in\{0,1\}^n} (-1)^{f(x,y)}\sqrt{b_y}|y\rangle \tag{20}$$

for arbitrary probability distributions (a_x), (b_y) on Alice and Bob's inputs. Define the rescaled Gram matrix G as $G_{ij} = \sqrt{a_i}\sqrt{a_j}\langle\psi_i|\psi_j\rangle$. Now it is known [14] that G will have the same eigenvalues as ρ, and thus the same entropy. But it can easily be verified that

$$G = (AMB)(AMB)^\dagger \tag{21}$$

where A and B are diagonal matrices with $A_{ii} = \sqrt{a_i}$, $B_{ii} = \sqrt{b_i}$. So the eigenvalues of G are simply the singular values squared of AMB. We may thus write

$$S(\rho) = H(\sigma^2(AMB)) \tag{22}$$

where $\sigma^2(M)$ denotes the vector containing the squared singular values of a matrix M. We can now produce lower bounds on the quantum communication complexity of f by appealing to the result of Nayak and Salzman [19] which states that, if Alice wishes to transmit n bits to Bob over a quantum channel with probability of success p, Alice must send $m \geq \frac{1}{2}\left(n - \log\frac{1}{p}\right)$ bits to Bob. If they are not allowed to share prior entanglement, the factor of $1/2$ vanishes. This immediately gives a lower bound on the exact quantum communication complexity of f, as lower bounds on the forward communication required for the "clean" protocols that we use translate into lower bounds on the total amount of communication needed for any communication protocol.

In the bounded-error case, we can still use the Nayak-Salzman result. Consider a block coding scheme with blocks of length k where each letter $|\psi_x^\epsilon\rangle$ is produced by one use of f, as in the previous section. By [11] there exists such a scheme that transmits $kS(\rho^\epsilon) - o(k)$ bits of information with k uses of f, as $k \to \infty$, and probability of success $p \to 1$. A lower bound on the bounded-error quantum communication complexity of f follows immediately:

$$mk \geq \frac{1}{2}(kS(\rho^\epsilon) - o(k) - o(1)), \tag{23}$$

hence, after taking the limit $k \to \infty$, $p \to 1$, we find $m \geq \frac{1}{2}S(\rho^\epsilon)$.

In order to reduce the error probability ϵ to $O(1/n^2)$ (to remove the additive term linear in n in inequality (19)), it is sufficient to repeat the original protocol $O(\log n)$ times and take a majority vote [16]. Alternatively, using (19) directly gives a better bound for functions for which $S(\rho)$ is linear in n. We thus have the following theorem.

Theorem II.1. *Let $f : \{0,1\}^n \times \{0,1\}^n \mapsto \{0,1\}$ be a total Boolean function with communication matrix M. Then, for any non-negative diagonal matrices A and B with $\|A\|_2 = \|B\|_2 = 1$,*

$$Q_E(f) \geq H(\sigma^2(AMB)) \tag{24}$$

$$Q_E^*(f) \geq \frac{1}{2}H(\sigma^2(AMB)) \tag{25}$$

$$Q_\epsilon(f) \geq \begin{cases} \Omega(H(\sigma^2(AMB))/\log n) \\ H(\sigma^2(AMB)) - 4\sqrt{\epsilon}n - \log\eta_0(4\sqrt{\epsilon}) \end{cases} \tag{26}$$

$$Q_\epsilon^*(f) \geq \begin{cases} \Omega(H(\sigma^2(AMB))/\log n) \\ \frac{1}{2}(H(\sigma^2(AMB)) - 4\sqrt{\epsilon}n - \log\eta_0(4\sqrt{\epsilon})) \end{cases} \tag{27}$$

where $\eta_0(x)$ is defined as in equation (18).

If we use the uniform distribution on Alice and Bob's inputs, then $AMB = M/2^n$. In the case of the models without entanglement, Klauck obtained this specialised result via a different method [15]. This theorem can thus be seen as simultaneously extending Klauck's work to the model with entanglement, generalising it, and giving it an operational interpretation. The special case of the uniform distribution was also used by Cleve et al. [5] to prove their lower bound on the communication complexity of IP.

III. RÉNYI ENTROPIC BOUNDS ON COMMUNICATION CAPACITY

A disadvantage of the von Neumann entropy $S(\rho)$ is the difficulty involved in its computation. The *second Rényi entropy* $S_2(\rho)$ [24] provides an easily computable lower bound on $S(\rho)$. $S_2(\rho)$ is defined as

$$S_2(\rho) = -\log \mathrm{tr}(\rho^2) = -\log \sum_{i,j} |\rho_{ij}|^2 \tag{28}$$

and we have the fundamental property that $S_2(\rho) \leq S(\rho)$. The Rényi entropy also obeys the bounds $0 \leq S_2(\rho) \leq n$. As with the von Neumann entropy, the Rényi entropy is a function only of the eigenvalues of ρ, so the Rényi entropy of the density matrix corresponding to an ensemble of equiprobable states is the same as that of the rescaled Gram matrix corresponding to these states. We can use this to write down a formula for the Rényi entropy of a density matrix ρ corresponding to the communication matrix M of a function (as in the previous section, specialising to the uniform distribution on Alice and Bob's inputs), which gives a lower bound on its communication capacity and thus its entanglement-assisted communication complexity.

$$\begin{aligned} S_2(\rho) &= -\log \mathrm{tr}\left(\frac{1}{2^{4n}}(MM^\dagger)^2\right) && (29)\\ &= 4n - \log\left(\sum_{i,j}\left(\sum_k M_{ik}M_{jk}\right)^2\right) && (30)\\ &= 4n - \log\left(\sum_{i,j,k,l} M_{ik}M_{jk}M_{il}M_{jl}\right) && (31) \end{aligned}$$

Rényi entropic arguments have previously been used in a different way by van Dam and Hayden [7] to put lower bounds on quantum communication complexity.

IV. THE QUANTUM COMMUNICATION COMPLEXITY OF A RANDOM FUNCTION

In this section, we will show a lower bound on the communication capacity – and thus the quantum communication complexity – of a random function (one which takes the value 0 or 1 on each possible input with equal probability). Define the state ρ as

$$\rho = \frac{1}{2^n}\sum_{k\in\{0,1\}^n} |\psi_k\rangle\langle\psi_k|, \text{ where } |\psi_k\rangle = \frac{1}{\sqrt{2^n}}\sum_{i\in\{0,1\}^n} (-1)^{a^k_{i+1}}|i\rangle \tag{32}$$

where a^k is a randomly generated 2^n-bit string, and a^k_i represents the i'th bit of a^k. We will show that the Rényi entropy $S_2(\rho)$ is high for almost all ρ.

Theorem IV.1. $\Pr\left[S_2(\rho) < (1-\delta)n\right] \leq e^{-(2^{\delta n}-1)^2/2}$.

Proof. We have

$$S_2(\rho) = 4n - \log\left(\sum_{i,j}\left(\sum_k M_{ik}M_{jk}\right)^2\right) \tag{33}$$

$$= 4n - \log\left(\sum_i\left(\sum_k (M_{ik})^2\right)^2 + \sum_{i\neq j}\left(\sum_k M_{ik}M_{jk}\right)^2\right) \tag{34}$$

$$= 4n - \log\left(N^3 + T\right) \tag{35}$$

where we define $N = 2^n$ and $T = \sum_{i\neq j}\left(\sum_k M_{ik}M_{jk}\right)^2$. It is then clear that

$$\Pr\left[S_2(\rho) < (1-\delta)n\right] = \Pr\left[T > N^3(N^\delta - 1)\right] \tag{36}$$

Each term in the inner sum in T (the sum over k) is independent and picked uniformly at random from $\{-1, 1\}$. We will now produce a tail bound for T using "Bernstein's trick" (see Appendix A of [3]): from Markov's inequality we have

$$\Pr\left[T > a\right] < \mathbb{E}(e^{\lambda T})/e^{\lambda a} < \mathbb{E}(e^{\lambda X_{11}})^{N^2}/e^{\lambda a} \tag{37}$$

where we define $X_{ij} = \left(\sum_k M_{ik}M_{jk}\right)^2$: each X_{ij} is independent and identically distributed, so T is the sum of $N(N-1) < N^2$ copies of X_{11}. It remains to calculate $\mathbb{E}(e^{\lambda X_{11}})$. This can be written out explicitly as follows.

$$\mathbb{E}(e^{\lambda X_{11}}) = \frac{1}{2^N}\sum_{k=0}^{N}\binom{N}{k}e^{\lambda(N-2k)^2} \tag{38}$$

It is then straightforward to see (using an inequality from [3]) that the following series of inequalities holds.

$$\mathbb{E}(e^{\lambda X_{11}}) \leq \frac{1}{2^N}\sum_{k=0}^{N}\binom{N}{k}\left(e^{\lambda(N-2k)^2} + e^{-\lambda(N-2k)^2}\right) \leq \frac{1}{2^{N-1}}\sum_{k=0}^{N}\binom{N}{k}e^{\lambda^2(N-2k)^4/2} \tag{39}$$

$$\leq \frac{1}{2^{N-1}}\sum_{k=0}^{N}\binom{N}{k}e^{\lambda^2 N^4/2} = 2e^{\lambda^2 N^4/2} \tag{40}$$

Inserting this in eqn (37), and minimising over λ, gives

$$\Pr\left[T > a\right] < 2e^{-a^2/2N^6} \tag{41}$$

and substituting $a = N^3(N^\delta - 1)$ gives the required result. □

In particular, putting $\delta = 1/2$ gives that $\Pr\left[S_2(\rho) < n/2\right] \leq 2e^{-(\sqrt{N}-1)^2/2}$, which is doubly exponentially small in n. As ρ corresponds to the communication matrix of a random function, Theorem II.1 immediately gives the result that the entanglement-assisted quantum communication complexity of almost all functions is $\Omega(n)$.

V. DISCUSSION AND OPEN PROBLEMS

We have shown that the implementation of any distributed computation between Alice and Bob entails the ability to communicate from one user to the other. This communication capacity of a

Boolean function of two arguments is naturally a lower bound on the communication complexity to compute that function, and we have proved corresponding lower bounds, even in the presence of arbitrary entanglement.

These bounds show that random functions of two n-bit strings mostly have communication complexity close to n. However, in general it has to be noted that our bounds are not that good: an example is provided by the set-disjointness problem, where Alice and Bob want to determine if their strings x and y have a position where they are both 1. It is known that the quantum communication complexity of this function is $\Theta(\sqrt{n})$ [1, 23]. On the other hand, the entropy in our main theorem was already computed for this case in [2], and it is only $O(\log n)$. Thus, not quite surprisingly, the ability of a function to let Alice communicate to Bob is not the same as the communication cost of implementing this computation.

Looking again at our main theorem, we are left with one interesting question: is the logarithmic factor that we lose in the bounded error model really necessary? It appears to be a technicality, since we need to boost the success probability to apply Fannes' inequality, but we were unable to determine if it is just that or if there are cases in which the lower bound is tight.

Acknowledgements

AM would like to thank Richard Jozsa for careful reading and comments on this manuscript, and Tony Short and Aram Harrow for helpful discussions. We thank Ronald de Wolf for pointing out references [18] and [20]. AW acknowledges support via the EC project QAP, as well as from the U.K. EPSRC. He also gratefully notes the hospitality of the Perimeter Institute for Theoretical Physics in Waterloo, Ontario, where part of this work was done.

---

[1] S. Aaronson, A. Ambainis. Quantum search of spatial regions. In *Proceedings of the 44th IEEE Symposium on Foundations of Computer Science (FOCS'03)*, p200, `quant-ph/0303041`, 2003.

[2] A. Ambainis, L. J. Schulman, A. Ta-Shma, U. Vazirani, A. Wigderson. The Quantum Communication Complexity of Sampling. *SIAM J. Comput.* 32, pp. 1570-1585, 2003.

[3] N. Alon, J. Spencer. The probabilistic method. Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley, New York, 2000.

[4] H. Buhrman, R. de Wolf. Communication complexity lower bounds by polynomials. *16th Annual IEEE Conference on Computational Complexity (CCC'01)*, p. 0120, `cs.CC/9910010`, 2001.

[5] R. Cleve, W. van Dam, M. Nielsen, A. Tapp. Quantum entanglement and the communication complexity of the inner product function. *Selected papers from the First NASA International Conference on Quantum Computing and Quantum Communications*, p.61-74, February 17-20, `quant-ph/9708019`, 1998.

[6] M. Fannes. A continuity property of the entropy density for spin lattice systems. *Commun. Math. Phys.* 31, pp. 291-294 (1973).

[7] W. van Dam, P. Hayden. Renyi-entropic bounds on quantum communication. `quant-ph/0204093`, 2002.

[8] D. Gavinsky, J. Kempe, R. de Wolf. Strengths and weaknesses of quantum fingerprinting. `quant-ph/0603173`, 2006.

[9] D. Gavinsky, J. Kempe, R. de Wolf. Exponential separation of quantum and classical one-way communication complexity for a Boolean function. `quant-ph/0607174`, 2006.

[10] D. Gavinsky. On the role of shared entanglement. `quant-ph/0604052`, 2006.

[11] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, W. Wootters. Classical information capacity of a quantum channel. *Phys. Rev. A*, vol. 54, no. 3, pp. 1869-1876, 1996.

[12] A. S. Holevo. Bounds for the quantity of information transmittable by a quantum communications channel. *Problemy Peredachi Informatsii*, vol. 9, no. 3, pp. 3-11, 1973. English translation *Problems of Information Transmission*, vol. 9, pp. 177-183, 1973.

[13] R. A. Horn, C. Johnson. Matrix analysis. Cambridge University Press, Cambridge, 1996.
[14] R. Jozsa, J. Schlienz. Distinguishability of states and von Neumann entropy. *Phys. Rev. A* 62 012301, `quant-ph/9911009`, 2000.
[15] H. Klauck. Lower bounds for quantum communication complexity. In *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science (FOCS'01)*, pp. 288297, `quant-ph/0106160`, 2001.
[16] I. Kremer. Quantum communication. Master's thesis, Hebrew University, 1995.
[17] E. Kushilevitz, N. Nisan. Communication complexity. Cambridge University Press, Cambridge, 1997.
[18] N. Linial, A. Shraibman. Lower bounds in communication complexity based on factorization norms. Manuscript, `http://www.cs.huji.ac.il/~nati/PAPERS/quant_cc.pdf`, 2006.
[19] A. Nayak, J. Salzman. On communication over an entanglement-assisted quantum channel. *Proceedings of the 34th ACM Symposium on the Theory of Computing (STOC'02)*, `quant-ph/0206122`, 2002.
[20] M. A. Nielsen. Quantum information theory. PhD thesis, University of New Mexico, Albuquerque, `quant-ph/0011036`, 1998.
[21] M. A. Nielsen, I. L. Chuang. Quantum computation and quantum information. Cambridge University Press, Cambridge, 2000.
[22] R. Raz. Fourier analysis for probabilistic communication complexity. *Computational Complexity*, vol. 5, pp. 205-221, 1995.
[23] A. A. Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya of the Russian Academy of Science*, vol. 67, pp. 159-176, `quant-ph/0204025`, 2003.
[24] A. Rényi. Probability theory. North-Holland, Amsterdam, 1970.
[25] R. de Wolf. Quantum communication and complexity. Theoretical Computer Science 287(1), pp. 337353, 2002.