Classical Interaction Cannot Replace Quantum Nonlocality

Dmitry Gavinsky
NEC Laboratories America, Inc.
4 Independence Way, Suite 200
Princeton, NJ 08540, U.S.A.

Abstract

We present a two-player communication task that can be solved by a protocol of polylogarithmic cost in the *simultaneous* message passing model with classical communication and shared entanglement, but requires exponentially more communication in the classical *interactive* model.

Our second result is a two-player nonlocality game with input length n and output of polylogarithmic length, that can be won with probability $1 - o(1)$ by players sharing polylogarithmic amount of entanglement. On the other hand, the game is lost with probability $\Omega(1)$ by players without entanglement, even if they are allowed to exchange up to k bits in interactive communication for certain $k \in \tilde{\Omega}\left(n^{1/8}\right)$.

These two results give almost the strongest possible (and the strongest known) indication of nonlocal properties of two-party entanglement.

1 Quantum vs. classical communication and games

The ultimate goal of quantum computing is to understand what advantages are offered by using the laws of quantum mechanics to build computational devices. We want to find tasks where quantum computers are considerably more efficient than classical ones.

In this paper we study quantum computation from the perspective of Communication Complexity, first defined by Yao [Y79]. Two parties, Alice and Bob, try to solve a computational problem that depends on x and y. Initially Alice knows only x and Bob knows only y; in order to solve the problem they communicate, obeying to restrictions of a specific *communication model*.

A *communication protocol* describes behavior of the participants in response to each possible input. The *cost* of a protocol is the maximum total amount of (qu)bits communicated by the parties, according to the protocol.

We say that a communication task P is solvable *with bounded error* in a given communication model by a protocol of cost $O(k)$ if for any constant $\varepsilon > 0$, there exists a corresponding protocol solving P with success probability at least $1 - \varepsilon$ and communicating $O(k)$ (qu)bits.

In order to compare the power of two communication models, one has to either prove existence of a task that can be solved more efficiently in one model than in the other, or to argue that no such task exists.

We will, in the first place, be concerned about the following models.

- *Interactive (two-way) communication* is a model where the players can interactively exchange messages till Bob decides to give an answer, based on the communication transcript and his part of input.

- *One-way communication* is a model where Alice sends a single message to Bob who has to give an answer, based on the content of the message and his part of input.

- *Simultaneous Message Passing (SMP)* is a model involving third participant, a *referee*. Here both Alice and Bob send one message each to the referee, who has to give an answer, based on the content of the received messages.

All three models can be either *classical* or *quantum*, according to the nature of communication allowed between the players. The classical versions of the models are denoted by $\mathcal{R}$, $\mathcal{R}^1$ and $\mathcal{R}^{\|}$, and the quantum versions are denoted by $\mathcal{Q}$, $\mathcal{Q}^1$ and $\mathcal{Q}^{\|}$, respectively.

It is clear that interactive communication is at least as powerful as one-way communication, and it is well-known that the former can, in fact, be much more efficient than the latter, both in quantum and in classical versions. All the same is true regarding one-way communication vs. SMP.

The model of SMP (both $\mathcal{Q}^{\|}$ and $\mathcal{R}^{\|}$) can be made stronger by allowing some *shared resource* between Alice and Bob, which can be either *shared randomness* or *shared entanglement*. The former case can be viewed as allowing Alice and Bob to follow *mixed* joint strategy, and the latter case means that they can perform quantumly-nonlocal (i.e., entangled) operation, as described by the laws of quantum mechanics. Of course, shared entanglement is always at least as helpful as shared randomness, and it has been shown ([GKRW06]) that the latter can, in fact, be much more efficient than the former (both with $\mathcal{Q}^{\|}$ and with $\mathcal{R}^{\|}$). When enhanced by shared randomness, the models are denoted as $\mathcal{Q}^{\|,pub}$ and $\mathcal{R}^{\|,pub}$, and the case of shared entanglement is addressed by $\mathcal{Q}^{\|,ent}$ and $\mathcal{R}^{\|,ent}$.

In this paper our primary concern is with separating communication models. Both for previously known results (mentioned later) and for our contribution it is the case that the first demonstrated super-polynomial separation had, in fact, been exponential.

Communication tasks can be either *functional*, meaning that there is exactly one correct answer corresponding to every possible input, or *relational*, when multiple correct answers are allowed. Functional tasks over domains forming product sets w.r.t. each player's inputs are called *total*. The three types of communication tasks form a hierarchy, if viewed as tools to separate communication model. In particular, there are known pairs of communication models that can be separated through a relational problem but are equally strong over functions, either total or partial, and there are pairs of communication models that can be separated through a partial functional problem but are conjectured to be equally strong over total functions.

For 0-error one-way and interactive protocols, separations have been demonstrated by Buhrman, Cleve and Wigderson [BCW98]. In the bounded-error setting the first separation has been given by Raz [R99], showing a problem solvable in $\mathcal{Q}$ exponentially more efficiently than in $\mathcal{R}$. Later, Buhrman, Cleve, Watrous and de Wolf [BCWW01] demonstrated an exponential separation for the SMP model. All these separations have been given for functional problems.

For one-way protocols with bounded error, the first separation has been shown by Bar-Yossef, Jayram and Kerenidis [BJK04] for a relational problem. Later, Gavinsky, Kempe, Kerenidis, Raz and de Wolf [GKKRW07] gave a similar separation for a partial functional problem.

These results show that quantum communication models can be very efficient, when compared to their classical counterparts. Recently it has been shown in [G08] that there exists

a problem that can be solved by a quantum one-way protocol more efficiently than by any classical two-way protocol. That result simultaneously subsumes the separation in [BJK04] and partially that in [R99] (in [G08] we were only able to show separation through a relational problem).

Closely related to SMP models $\mathcal{R}^{\|,pub}$ and $\mathcal{R}^{\|,ent}$ is the notion of *nonlocality games*. In this setting two players, Alice and Bob, receive two parts of input and must produce their output without communicating to each other. The difference from the SMP setting is that there is no referee (who would help the players to achieve their goal), instead there is a *verifier* who checks whether the answers from the players are good w.r.t. the given input. Every nonlocality game defines which input pairs are allowed and which pairs of answers are good for each input pair.

There is a number of known nonlocality games (cf. [CHTW04]) that can be played with higher success probability by entangled (*nonlocal*) players than by unentangled (*local*) players who share randomness, hence the name. Historically ([EPR35], [B64]), the setting of nonlocality games has been viewed as a way to demonstrate nonlocal behavior of quantum entanglement.

1.1 Our results

In this work we give even more surprising demonstration of "usefulness" of quantum mechanics in solving communication problems. We give a communication problem that is

- easy in the weakest known communication model with entanglement, $\mathcal{R}^{\|,ent}$;
- hard in (one of) the strongest classical 2-party models, $\mathcal{R}$.

Moreover, our communication problem is also easy in $\mathcal{Q}^1$, and therefore this result subsumes [G08].

Theorem 1.1. *For infinitely many $n \in \mathbb{N}$, there exists an (explicit) relation with input length $O\big(n^2 \log n\big)$ that can be efficiently solved in $\mathcal{R}^{\|,ent}$, namely there is an SMP protocol with error $o(1)$ where the two players share $O\big(\log^3 n \log\log n\big)$ EPR pairs and send classical messages of length $O\big(\log^3 n \log\log n\big)$. The same relation has communication complexity $\Omega\Big(\frac{n^{1/4}}{\log^2 n}\Big)$ in the interactive classical model ($\mathcal{R}$).*

The relation we use is a modification of a construction independently suggested by R. Cleve ([C]) and S. Massar ([B]), as a possible candidate for such separation. In [G08] a more "demanding" version of this relation was used. In particular, the communication task from [G08] is unlikely to admit an efficient solution in $\mathcal{R}^{\|,ent}$.

Our second result is a new, stronger type of a nonlocality game. In this work by *games* we mean the following setup: Two players, called Alice and Bob, receive inputs x and y and are required to produce outputs a and b, respectively. The inputs are generated according to some known probability distribution. Depending on the inputs x and y, some pairs (a, b) are viewed as "winning", while all non-winning outputs are "loosing" (technically, there is a predicate specifying which tuples (x, y, a, b) correspond to winning answers).

The goal in designing a *game strategy* is to maximize the winning probability. A strategy is called *local* if it can be carried out by players who are space-separated (i.e., a is a function of x and r_A, and b is a function of y and r_B, where r_A and r_B are mutually-independent

random strings); a strategy is called *mixed* if it is a convex combination (or "mixture") of local strategies (that is, the players have access to unlimited source of shared randomness). A strategy is called *entangled* if it can be carried out by players who are space-separated but share (unlimited amount of) quantum entanglement, and can perform any local measurements on their parts of the shared quantum state before producing their answers (Alice's measurement may depend on x and Bob's measurement may depend on y). The supremum of winning probability achievable by mixed and entangled strategies is called the *classical value* and the *entangled (or quantum) value* of the game, respectively. *Nonlocality games* are the ones with the entangled value being larger than the classical value.

We demonstrate a family of nonlocality games, where the output length is poly-logarithmic in the input length. The entangled value of our games is $1 - o(1)$, and the classical value is $1 - \Omega(1)$. Novelty of our construction is due to the fact that non-entangled players err with probability $\Omega(1)$, even if they are *allowed to use classical interaction before producing their answers*, where the allowed amount of communication is exponential in the output length (i.e., polynomially related to the input length).

Theorem 1.2. *For infinitely many $n \in \mathbb{N}$, there exists an (explicit) 2-party nonlocality game where both the players receive input of length $O\big(n^2 \log n\big)$ and are required to produce output of length $O\big(\log^3 n \log\log n\big)$. The game can be won with probability $1 - o(1)$ by players who share $O\big(\log^3 n \log\log n\big)$ bits of entanglement (EPR pairs), but any local mixed strategy would result in loss with probability $\Omega(1)$; moreover, the loss would remain $\Omega(1)$ even if the players are allowed to interactively exchange $o\Big(\frac{n^{1/4}}{\log^2 n}\Big)$ classical bits before producing their output.*

Proof. Let P be the communication problem used in Theorem 1.1, and let S be an $\mathcal{R}^{\|,ent}$-protocol for P, as promised. W.l.g., we assume that the referee's action in S is deterministic (at most $O(\log n)$ bits of randomness are required, which can be chosen by Alice and attached to her message).

Let a game G be as follows:

- Input to G is the same as input to P.
- Input distribution is any such μ that any communication protocol of cost $o\Big(\frac{n^{1/4}}{\log^2 n}\Big)$ fails to solve P with probability $\Omega(1)$ (such μ exists due to Theorem 1.1 and the Minimax theorem).
- For any (x, y), the set of valid responses in G coincides with the set of pairs of messages that would result, according to S, in a correct answer by the referee to (x, y).

Note that G is won with probability $1 - o(1)$ by two entangled players who acts according to S. On the other hand, G is lost with probability $\Omega(1)$ by any classical players exchanging $o\Big(\frac{n^{1/4}}{\log^2 n}\Big)$ bits, according to our choice of μ and definition of correct answers to G. ■

To the best of author's knowledge, this is the first example of a nonlocality game where an entangled strategy outperforms any classical one, even if the latter is "boosted" by allowing up to k bits of interaction, where k is exponential in the total length of players' answers.[1]

[1]Examples of $\Omega(n)$-party nonlocality games efficiently played with entanglement, where classical solution would require $n^{\Omega(1)}$ bits of auxiliary communication were given by Buhrman, Høyer, Massar and Röhrig in [BHMR03] and [BHMR06] – of course, the total output length is also $\Omega(n)$ in that case.

The fact that our game is a simple 2-party one makes it an appealing candidate for an actual physical experiment.

1.2 Our techniques

The main result of this paper can be viewed as strengthening of [G08]. The communication task that we use in this paper is, intuitively, easier to solve than the one used in [G08], and so its analysis requires more refined lower bound tools. On the high level, the difference can be viewed as follows.

Ignoring some technical details, in both the cases the task is:

- there is an $n \times n$ table, where each cell contains 2 elements;
- every element from $\{1, \ldots, 2n^2\}$ appears in the table once;
- Alice knows the row and Bob knows the column of each element;
- for the answer the players must choose a cell and give certain witness of knowledge about both the element in it.

The intuitive motivation to ask for a witness is that *ability to provide it would effectively require from classical players to know the element of the chosen cell*, while entangled players are able to get a witness with much less communication than it would take them to learn the content of a cell.

The two versions of the task differ in *how may the players choose a cell for their answer*. In this paper we consider a version where the players can pick *any cell*. In [G08] the players had to pick a *cell from the first row*.

This difference surprisingly makes the earlier version considerably harder to solve; in particular, it is unlikely to admit an efficient solution in the model of simultaneous messages with entanglement (but still has an efficient solution in the model of quantum one-way communication, as demonstrated in [G08]). The same difference makes the earlier version less robust against lower bound techniques.

The lower bound proof in [G08] consists of several "quasi-reductions", the first of which can be (ignoring some technicalities) stated as follows.

Proposition 1. *([G08])* If there is a two-way classical solution of cost k to the communication task under consideration, then the 1×1-version of the task can be solved with probability $\Omega(1/n)$ by a protocol of similar cost.

Here by a 1×1-version we mean a modification of the task where the players are more restricted – namely, they no longer have freedom to choose a cell, but must give their answer w.r.t. a fixed (say, the first) cell of the first row. In order to prove the above statement, we used the fact that a protocol can only choose one of n cells from the first row, and therefore a "typical" cell must be chosen by the protocol with probability at least $1/n$. It turned out [G08] that a short classical communication protocol could not solve the 1×1-version with probability significantly larger than $1/n^2$, and therefore a lower bound on the classical complexity of the original problem followed.

In order to allow for an efficient solution in the model of $\mathcal{R}^{\|,ent}$, we have to give the players more freedom; namely, to let them choose any cell from the table. When we define the communication problem this way, the above proof idea no longer applies: there are n^2

cells, and therefore via using a simple pigenhole-based argument (similar to the one described above), we can only hope to argue that there exists a cell that is chosen by a valid protocol with probability at least $1/n^2$. However, a trivial classical protocol of communication cost $O(\log n)$ can solve the 1×1-version of the problem with probability $1/n^2$.

The main technical contribution of this paper is proving a statement analogous to Proposition 1 w.r.t. a relaxed version of the communication problem (where the players are allowed to choose any cell). The new proof is based on a lemma due to Harsha, Jain, McAllester and Radhakrishnan [HJAR07] that was proved via an elegant information-theoretic argument.

We had to carefully adjust the definition of our communication problem in several ways, such that the lemma from [HJAR07] became applicable (which required our problem to be defined over input distribution being "product" in a very strong sense), but the communcation problem still was easy for solving in the model of $\mathcal{R}^{\|,ent}$. As a result, the communication problem used in this paper and the $\mathcal{R}^{\|,ent}$-protocol that solves it are significantly more involved than those in [G08].

2 Preliminaries

In our analysis we use the following generalization of the standard bounded error setting. We say that a protocol solves a problem *with probability δ with error bounded by ε* if with probability at least δ the protocol produces an answer, and whenever produced, the answer is correct with probability at least $1-\varepsilon$.

Let n be a power of 2. Denote by $\bar{0}$ the (additive) identity of the vector space $\mathcal{GF}_2^{\log n}$. We will implicitly assume equivalence between $a \in [n]$ and the lexicographically a'th element of $\mathcal{GF}_2^{\log n}$.

Let $x = (x_1, \ldots, x_k)$ and $y = (y_1, \ldots, y_l)$ be tuples of sets. We will call $x_i \cap y_j$ a *cell*, denoting it by $\mathrm{cell}(i,j)$.

We will demonstrate our separation via the following communication problem.

Definition 1. (Communication problem P) *Let $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ be sequences of subsets of $[4n^2]$, where each subset is of size n. Let $z = \big((i_1, j_1, c_1), \ldots, (i_{t_n}, j_{t_n}, c_{t_n})\big)$ for $t_n \stackrel{def}{=} \lfloor \log \log n \rfloor$. Then $(x, y, z) \in P$ if it holds that*

- *For any $k_1 \neq k_2$, $(i_{k_1}, j_{k_1}) \neq (i_{k_2}, j_{k_2})$;*
- *for every $k \in [t_n]$, either $|\mathrm{cell}(i_k, j_k)| \neq 2$, or $c_k \in [4n^2] \setminus \{\bar{0}\}$ and $\langle c_k, a+b \rangle = 0$, where $x_{i_k} \cap y_{j_k} = \{a, b\}$;*
- *at least one of the following holds:*

$$|\{k \in [t_n] \,|\, |\mathrm{cell}(i_k, j_k)| = 2\}| \geq \frac{t_n}{66} \quad \text{or} \quad |\{(i,j) \in [n] \times [n] \,|\, |\mathrm{cell}(i,j)| = 2\}| < \frac{n^2}{65}.$$

The above definition is somewhat involved, as a result of "fine-tuning" the complexity of the problem to make it efficiently solvable in the model of $\mathcal{R}^{\|,ent}$, but (provably) not in $\mathcal{R}$. Intuitively, the "most relevant" part of an answer to P are the references to cells of size 2. We did not "promise" that all cells were of size 2, as we wanted to avoid, as much as we could, conditions on input and consider a problem that would be hard for $\mathcal{R}$ even w.r.t. some

product input distribution. However, it is probably not possible in the model of $\mathcal{R}^{\|,ent}$ to check efficiently whether a given cell is of size 2; so, we required that t_n different cells were listed in a valid answer in order to guarantee that with high probability there were some cells of size 2 among them.

Unless stated otherwise, we will assume that input to P satisfies the following additional promises:

- at least $\frac{n^2}{65}$ cells are of size 2;
- $\max\left\{|\{i \,|\, a \in x_i \text{ or } a \in y_i\}| \;\middle|\; a \in [4n^2]\right\} \leq 9 \log n$;
- $\sum_{i,j\in[n]} |\text{cell}(i,j)| \leq 2n^2$.

We denote by $\mathcal{U}_n$ the uniform distribution of such input.

These promises will not be an obstacle in proving a lower bound, as they are satisfied with probability $1-o(1)$ by input drawn from the uniform distribution over all possible pairs of n-tuples of n-element subsets of $[4n^2]$ – denote this distribution by $\mathcal{U}'_n$. Observe that $\mathcal{U}'_n$ is a product distribution not only with respect to the two sides of input, but also every x_i or y_i is chosen independently from the rest of the input (this will play a crucial role later, when we prove a lower bound).

Observation 2.1. *$\mathcal{U}'_n$ produces an instance outside the support of $\mathcal{U}_n$ with probability $o(1)$.*

Proof. Let n be sufficiently large. First, observe that

$$\mathbf{Pr}\left[\text{size of a given cell is 2}\right] \geq \binom{n}{2} \cdot \left(\frac{4n^2-2n+3}{4n^2}\right)^{n-2} \cdot \left(\frac{n-1}{4n^2}\right)^2 > \frac{1}{65},$$

where the estimate corresponds to first fixing arbitrarily the n row elements, followed by sequentially choosing the n column elements, exactly 2 of which should come from the set of row elements. Second,

$$\mathbf{E}\left[\text{cell size}\right] = 4n^2 \cdot \mathbf{Pr}\left[\text{a given element belongs to a given cell}\right] = 4n^2 \cdot \left(\frac{n}{4n^2}\right)^2 = \frac{1}{4}.$$

The statement follows by standard tail-bounding methods. ■

We will give an efficient $\mathcal{R}^{\|,ent}$-protocol that correctly solves P with probability $1-o(1)$, for *any* input from the support of $\mathcal{U}_n$. Then we will show that any short deterministic $\mathcal{R}$-protocol fails with probability $\Omega(1)$, when the input is drawn from $\mathcal{U}'_n$. By Observation 2.1 and the Minimax theorem, that will imply Theorem 1.1.

3 Efficient protocol for P in $\mathcal{R}^{\|,ent}$

Let us construct an SMP protocol that uses entanglement and sends $O\big(\log^3 n \log\log n\big)$ classical bits that allow the referee to solve P with probability $1-o(1)$, as long as the input comes from the support of $\mathcal{U}_n$.

For any tuple of sets $x = (x_1, \ldots, x_k)$, such that $\cup x_i \subseteq [m]$ for some m being a power of 2, we define a measurement Π_x acting on $\log m$ qubits, as follows.

Let $\alpha_x \stackrel{\text{def}}{=} \max\{|\{j \,|\, i \in x_j\}| \,|\, i \in [m]\}$. Intuitively, in defining Π_x we would like to measure $\log m$ qubits with the $k+1$ projectors $\tilde{E}_i^{(x)} \stackrel{\text{def}}{=} \sum_{j \in x_i} |j\rangle\langle j|$ and $\tilde{E}_0^{(x)} \stackrel{\text{def}}{=} \sum_{j \notin \cup x_i} |j\rangle\langle j|$. However, if $\alpha_x > 1$ then the above collection of projectors does not constitute a valid measurement, as $\sum E_i^{(x)} \not\equiv I$. To handle that, define:

$$\Pi_x \stackrel{\text{def}}{=} \left(E_0^{(x)}, \ldots, E_k^{(x)}\right);\ E_i^{(x)} \stackrel{\text{def}}{=} \frac{1}{\alpha_x}\tilde{E}_i^{(x)} \text{ for } 1 \le i \le k;\ E_0^{(x)} \stackrel{\text{def}}{=} I - \sum E_i^{(x)}.$$

Defined this way, Π_x is a valid quantum measurement.[2]

Consider the following protocol S.

1. Alice and Bob share the entangled state $\frac{1}{2n}\sum_{t \in [4n^2]} \underbrace{|t\rangle}_{\text{Alice}} \underbrace{|t\rangle}_{\text{Bob}}$.
2. Alice measures her part of the shared state with Π_x and Bob measures his with Π_y.
3. Alice and Bob both apply the Hadamard transform to their halves of the shared state.
4. Alice and Bob both measure their parts of the state in the computational basis.
5. Alice and Bob send the outcomes of their measurements to the referee (i.e., the referee receives the outcomes of 4 measurements).
6. Having received (i,k) from Alice and (j,l) from Bob, the referee outputs $(i,j,k+l)$.

Obviously, S does not solve P, though we will see how it can be used as a "stepping stone" for the desired protocol.

Let us analyze the action of S when the input comes from the support of $\mathcal{U}_n$. Assume that $i = i_0 \neq 0$ and $j = j_0 \neq 0$. This means that in step 2 of S, Alice's measurement has resulted in $E_{i_0}^{(x)}$ and Bob's in $E_{j_0}^{(y)}$. Then the shared state turned into

$$\frac{1}{\sqrt{|\text{cell}(i_0, j_0)|}} \sum_{t \in \text{cell}(i_0, j_0)} |t\rangle\,|t\rangle$$

(note that unless $\text{cell}(i_0, j_0)$ contains less than 2 elements, its content is not locally accessible by the players, neither at this stage nor later in the protocol).

Assume that $\text{cell}(i_0, j_0)$ contained exactly 2 elements, a and b. Then after step 3 of the protocol the shared state became (ignoring normalization)

$$\sum_{k,l} \left((-1)^{\langle k+l, a\rangle} + (-1)^{\langle k+l, b\rangle}\right) |k\rangle\,|l\rangle\,,$$

and $|k\rangle\,|l\rangle$ has non-zero amplitude if and only if $\langle k+l, a\rangle = \langle k+l, b\rangle$, which is equivalent to $\langle k+l, a+b\rangle = 0$. Moreover, each pair k_0, l_0 satisfying this condition is equally likely to become the measurement outcome in step 4; in particular, the value of $t_0 \stackrel{\text{def}}{=} k_0 + l_0$ is drawn from uniform distribution whose support is $\{t \,|\, \langle t, a+b\rangle = 0\}$.

To sum up, if S produces a triple (i_0, j_0, t_0), such that $i_0 \neq 0$, $j_0 \neq 0$ and $\text{cell}(i_0, j_0) = \{a, b\}$ for some $a \neq b$, then

[2] Although Π_x is not, in general, a projection, it is a measurement of a very special kind – a *POVM with commuting elements*.

- $t_0 \neq \bar{0}$ with probability $1 - o(1)$, and
- $\langle t, a + b \rangle = 0$ with certainty.

It remains to analyze the probability that $|\mathrm{cell}(i,j)| = 2$. The joint measurement taken by the parties in step 2 can be written as

$$\Pi_{x,y} \stackrel{\mathrm{def}}{=} \left\{ \frac{1}{\alpha_x \alpha_y} \tilde{E}_{i,j}^{(x,y)} \right\}_{i,j=1}^{n} \cup \left\{ E_0^{(x,y)} \right\},$$

where $\tilde{E}_{i,j}^{(x,y)} \stackrel{\mathrm{def}}{=} \left(\sum_{t_1 \in x_i} |t_1\rangle\langle t_1| \right) \otimes \left(\sum_{t_2 \in y_j} |t_2\rangle\langle t_2| \right)$ and $E_0^{(x,y)} \stackrel{\mathrm{def}}{=} \left(I - \sum E_{i,j}^{(x,y)} \right)$ (the latter corresponds to the case when S returns "$i = 0$" or "$j = 0$").

Let $E_{i,j}^{(x,y)} \stackrel{\mathrm{def}}{=} \frac{1}{\alpha_x \alpha_y} \tilde{E}_{i,j}^{(x,y)}$. Then for any $i_0 \neq 0$ and $j_0 \neq 0$,

$$\mathbf{Pr}\left[i = i_0, j = j_0\right] = \mathrm{tr}\left(\frac{1}{2n} \sum_{s_1 \in [4n^2]} |s_1\rangle\langle s_1| \cdot E_{i_0,j_0}^{(x,y)} \cdot \frac{1}{2n} \sum_{s_2 \in [4n^2]} |s_2\rangle\langle s_2| \right) = \frac{|\mathrm{cell}(i_0, j_0)|}{4n^2 \alpha_x \alpha_y}.$$

In particular,

$$\mathbf{Pr}\left[|\mathrm{cell}(i,j)| = 2 \middle| i \neq 0, j \neq 0 \right] = \frac{2 \cdot |\{(i',j') : |\mathrm{cell}(i',j')| = 2\}|}{\sum_{i',j' \in [n]} |\mathrm{cell}(i',j')|} \geq \frac{1}{65} \tag{1}$$

and

$$\begin{aligned} \mathbf{Pr}\left[i \neq 0, j \neq 0\right] \geq \mathbf{Pr}\left[i \neq 0, j \neq 0, |\mathrm{cell}(i,j)| = 2\right] &= \frac{2 \cdot |\{(i',j') : |\mathrm{cell}(i',j')| = 2\}|}{4n^2 \alpha_x \alpha_y} \\ &\geq \frac{1}{260 \alpha_x \alpha_y} \geq \frac{1}{21060 \log^2 n}, \end{aligned} \tag{2}$$

where all the inequalities follow from the fact that input belongs to the support of $\mathcal{U}_n$.

We are ready to construct a protocol for solving P. Simply, repeat in parallel S sufficiently many times in order to guarantee that, with probability $1 - o(1)$, at least t_n instances produce triples (i', j', t') referring to pairwise different cells, consisting exclusively of non-zero elements (i.e., $i', j' \neq 0$ and $t' \neq \bar{0}$). If that happens – our protocol outputs t_n such triples (randomly chosen from the obtained ones). From (2), (1), and an observation that all valid non-zero triples are equally likely to show up, it is clear that $O\big(\log^2 n \log\log n\big)$ parallel copies of S are sufficient. Such protocol solves P with probability $1 - o(1)$, and its communication complexity is $O\big(\log^3 n \log\log n\big)$.

4 Solving P is expensive in $\mathcal{R}$

Now we show that solving P in $\mathcal{R}$ requires a protocol of cost $\Omega\Big(\frac{n^{1/4}}{\log^2 n}\Big)$. Unless stated otherwise, all the following statements are made w.r.t. the model $\mathcal{R}$.

Let S be a protocol for P of cost k that makes an error with probability at most ε (a sufficiently small constant). Then there exists a deterministic protocol S' that errs with probability at most 2ε w.r.t. $\mathcal{U}'_n$, as follows from Observation 2.1 and the Minimax theorem.

Let $\mathrm{infl}(x,y,i,j,c)$ be the predicate indicating that when (x,y) is the input to S', then $|\mathrm{cell}(i,j)| = 2$ and (i,j,c) is a part of the output by S'. Define $\mathrm{infl}(x,y,i,j)$ to be the predicate that is satisfied when there exists c satisfying $\mathrm{infl}(x,y,i,j,c)$. Intuitively, $\mathrm{infl}(x,y,i,j)$ means that $\mathrm{cell}(i,j)$ plays an important role in the output of S', when the input is (x,y). For every satisfying assignment to $\mathrm{infl}(x,y,i,j)$, denote by $c_{i,j}^{x,y}$ the (unique) value that satisfies $\mathrm{infl}(x,y,i,j,c_{i,j}^{x,y})$ (we can assume w.l.g. that S' always outputs a list of t_n different cells, as required by P).

For the purpose of analyzing P we will consider the following communication problem.

Definition 2. (Communication problem $P_{1\times 1}$) *Let $x,y \subset [4n^2]$, such that $|x| = |y| = n$ and $|x \cap y| = 2$. Let σ be a permutation over $[4n^2]$ and $c \in [4n^2] \setminus \{\bar{0}\}$. Let $x \cap y = \{a,b\}$, then $P_{1\times 1}$ allows two types of answers, as follows:*

- $(x,y,c) \in P_{1\times 1}$ *if* $\langle c, a+b\rangle = 0$;
- $(x,y,(\sigma,c)) \in P_{1\times 1}$ *if* $\langle c, \sigma(a)+\sigma(b)\rangle = 0$.

Let $\mathcal{U}^n 1\times 1$ be the uniform distribution of valid inputs to $P_{1\times 1}$. Note that unless stated otherwise, we always assume that the input is distributed according to $\mathcal{U}'_n$.

Denote:

$$p_{i,j} \stackrel{\text{def}}{=} \mathop{\mathbf{Pr}}_{x,y}\left[\mathrm{infl}(x,y,i,j) \text{ and } (x_i,y_j,c_{i,j}^{x,y}) \in P_{1\times 1}\right],$$
$$q_{i,j} \stackrel{\text{def}}{=} \mathop{\mathbf{Pr}}_{x,y}\left[\mathrm{infl}(x,y,i,j) \text{ and } (x_i,y_j,c_{i,j}^{x,y}) \notin P_{1\times 1}\right],$$
$$r_{i,j} \stackrel{\text{def}}{=} p_{i,j} + q_{i,j} = \mathbf{Pr}\left[\mathrm{infl}(x,y,i,j)\right].$$

Note that the correctness assumptions for S' imply that (assuming $\varepsilon \le 1/3$)

$$\sum_{1\le i,j\le n} p_{i,j} \ge \frac{(1-2\varepsilon)t_n}{66} > \frac{t_n}{200} \quad \text{and} \quad \sum_{1\le i,j\le n} q_{i,j} \le 2\varepsilon\cdot t_n. \tag{3}$$

Denote $A \stackrel{\text{def}}{=} \left\{(i,j) \;\middle|\; \mathbf{Pr}\left[(x_i,y_j,c_{i,j}^{x,y}) \in P_{1\times 1} \middle| \mathrm{infl}(x,y,i,j)\right] \ge 1 - 800\varepsilon\right\}$.

Claim 4.1. $\sum_{(i,j)\in A} r_{i,j} > \frac{t_n}{400}$.

Proof. From (3),

$$\sum_{1\le i,j\le n} p_{i,j} > \frac{t_n}{200} \ge \frac{t_n}{400} + \frac{1}{800\varepsilon}\sum_{1\le i,j\le n} q_{i,j},$$

and $\frac{t_n}{400} < \sum\left(p_{i,j} - \frac{q_{i,j}}{800\varepsilon}\right)$. For $i,j \in [n]$, let

$$p'_{i,j} \stackrel{\text{def}}{=} \mathbf{Pr}\left[(x_i,y_j,c_{i,j}^{x,y}) \in P_{1\times 1} \middle| \mathrm{infl}(x,y,i,j)\right] = \frac{p_{i,j}}{r_{i,j}},$$
$$q'_{i,j} \stackrel{\text{def}}{=} \mathbf{Pr}\left[(x_i,y_j,c_{i,j}^{x,y}) \notin P_{1\times 1} \middle| \mathrm{infl}(x,y,i,j)\right] = \frac{q_{i,j}}{r_{i,j}}.$$

Observe that $p'_{i,j} + q'_{i,j} = 1$ and $A = \left\{(i,j) \;\middle|\; q'_{i,j} \le 800\varepsilon\right\}$. Therefore

$$\frac{t_n}{400} < \sum_{1\le i,j\le n} r_{i,j}\left(p'_{i,j} - \frac{q'_{i,j}}{800\varepsilon}\right) \le \sum_{\substack{i,j:\\ q'_{i,j}\le 800\varepsilon p'_{i,j}}} r_{i,j} \le \sum_{(i,j)\in A} r_{i,j},$$

as wanted. ∎

We will need the following lemma, based on [HJAR07].

Lemma 4.2. ([HJAR07], Lemma 5.3, reformulated) *Let $\delta_1, \delta_2 > 0$, $k_1, k_2 \in \mathbb{N}$, such that a communication problem P is solvable w.r.t. distribution D by a k_1-round private coin protocol which produces an answer with probability at least δ_1, and a produced answer is correct with probability at least δ_2. Assume also that the mutual information between the transcript of the protocol and its inputs is at most k_2 bits, with probability $1 - o(\delta_1)$.*

Then P is solvable w.r.t. D by a public coin protocol of communication cost $O(k_1 + k_2)$ which produces an answer with probability at least $\delta_1/2$ and any produced answer is correct with probability at least $\delta_2 - o(1)$.

The statement in [HJAR07] is made in different terms, for completeness we outline a proof for our version of it.

Proof. Let S be the protocol guaranteed by the lemma requirement, assume that the input (x, y) is produced according to the distribution D. Let J be a random variable representing the amount of mutual information between the transcript of S and the input.

It is shown in [HJAR07] how to build a public coin protocol S' with the following properties:

- For some absolute constant β, communication cost of S' is upper-bounded by $\beta k_1 + 6J$.
- S' refuses to answer with probability at most $1/3$ (denote this event by E_1). Otherwise it produces an answer, distributed according to the same distribution as the answer of S for the given input.

Even if E_1 does not occur no answer may be returned, as S is allowed to have positive probability to refuse to answer. We will denote this second sort of refusal to answer by E_2 (note that this event is mutually exclusive with E_1). It is clear though that an answer is produced (i.e., neither E_1 nor E_2 occurs) with probability at least $\frac{2\delta_1}{3}$, and if that happens then the answer is correct with probability at least δ_2.

We construct a protocol S'', similar to S' but equipped with the following "cost control mechanism". If the total number of communicated bits exceeds $\beta k_1 + 6k_2$ then S'' halts and refuses to answer. The probability that this mechanism actually stops a single run of the protocol is upper-bounded by the probability that $J > k_2$, which is the probability that S exposes more than k_2 bits of information about its inputs. The latter is in $o(\delta_1)$ by the assumption, and therefore S'' returns an answer with probability $\frac{2\delta_1}{3} - o(\delta_1) > \frac{\delta_1}{2}$ (for sufficiently long inputs). On the other hand, whenever an answer is returned it is correct with probability at least $\delta_2 - \frac{o(\delta_1)}{\delta_1} = \delta_2 - o(1)$, as required. ■

The next lemma is from [G08].

Lemma 4.3. ([G08]) *For some absolute constant δ, any public coin $\mathcal{R}$-protocol of communication cost k solving $P_{1\times 1}$ with error bounded by δ returns an answer with probability $O\Big(\frac{k^4 \log^2 n + k^2 \log^6 n}{n^2}\Big)$.*

In the definition of $P_{1\times 1}$'s analogue in [G08] (denoted there by $P^{\Sigma}_{1\times 1}$) different constants are used, but Lemma 4.3 follows from [G08] by a straightforward reduction argument. Besides, the definition of $P^{\Sigma}_{1\times 1}$ requires that the permutation σ is arbitrary but fixed, instead

of being a part of the output, as allowed by our $P_{1\times 1}$; however, all the statements made in [G08] regarding $P^{\Sigma}_{1\times 1}$ are also valid w.r.t. $P_{1\times 1}$. [3]

Let T be a random variable corresponding to the transcript of S' (observe that its value is uniquely determined by the input, as the protocol is deterministic). We will consider the mutual information between T and parts of the input. For $i \in [n]$, let $s_i^{(x)}$ and $s_i^{(y)}$ be the mutual information between T and, respectively, x_i and y_i. Let $s_{i,j} \stackrel{\text{def}}{=} s_i^{(x)} + s_j^{(y)}$. Since $\mathcal{U}'_n$ chooses all x_i's and y_i's in a mutually independent manner, it holds that

$$\sum_{i,j\in[n]} s_{i,j} = n \cdot \left(\sum_{i\in[n]} s_i^{(x)} + \sum_{i\in[n]} s_i^{(y)} \right) \leq nk.$$

Claim 4.4. *For some absolute constant α and any $(i,j) \in A$, it holds that $p_{i,j} \leq \max\left\{\frac{1}{n^3}, \frac{\alpha \cdot s_{i,j} \cdot k^3 \log^8 n}{n^2}\right\}$.*

Before we prove it, let us see how the claim leads to the desired lower bound. From (3),

$$\Omega(t_n) \ni \sum_{(i,j)\in A} p_{i,j} \leq \frac{1}{n} + \sum_{\substack{(i,j)\in A \\ p_{i,j} > 1/n^3}} p_{i,j} \leq \frac{1}{n} + \frac{\alpha k^3 \log^8 n}{n^2} \cdot \sum_{(i,j)\in A} s_{i,j} \leq \frac{1 + \alpha k^4 \log^8 n}{n},$$

where $\Omega(t_n) = \Omega(\log\log n)$. Therefore, $k \in \Omega\left(\frac{n^{1/4}}{\log^2 n}\right)$.

Proof of Claim 4.4. W.l.g., let $(1,1) \in A$ and $p_{1,1} > \frac{1}{n^3}$, we want to show that for some absolute constant α, it holds that $p_{1,1} \leq \frac{\alpha \cdot s_{1,1} \cdot k^3 \log^8 n}{n^2}$. Let q' be the value defined similarly to $s_{1,1}$, but with input distribution conditioned upon $[|\text{cell}(1,1)| = 2]$. Because the condition holds with probability bounded below by a positive constant (by analogy to Observation 2.1), $q' \in O(s_{1,1})$. Therefore, it will be sufficient to show that $p_{1,1} \in O\left(\frac{q' \cdot k^3 \log^8 n}{n^2}\right)$.

We will use S' as a protocol for solving $P_{1\times 1}$ by embedding its instance into coordinate $(1,1)$ of P, and the rest of input will play the role of private randomness. That is, in order to solve $P_{1\times 1}$ with input (x', y'), we will run S' over the pair $\big((x', r_1, \ldots, r_{n-1})(y', r_n, \ldots, r_{2n-2})\big)$, where $r_1, \ldots, r_{n-1}$ are independent random subsets of $[4n^2]$ of size n, chosen by Alice, and $r_n, \ldots, r_{n-2}$ are similarly chosen by Bob. Note that our embedding requires only private randomness and does not cost any communication.

Let $l \stackrel{\text{def}}{=} \left\lceil \frac{k}{q'} \right\rceil$, and denote by S'_l a private coin protocol that receives l instances of $P_{1\times 1}$ and runs l independent copies of S' in parallel in a mutually independent manner, one for each instance of $P_{1\times 1}$ embedded into P as described. After that S'_l outputs (j, c_j) for a random j, such that the j'th copy of emulated S' contains $(1,1,c_j)$ among its answers; if no such j exists then S'_l returns no answer. The following properties of S'_l with respect to l input pairs independently chosen according to $\mathcal{U}^n 1 \times 1$ are easy to verify:

- If an answer is returned then c_j is a right answer to the j'th instance of $P_{1\times 1}$ with probability at least $1 - 800\varepsilon$ (this follows from $(1,1) \in A$).

[3] The lower bound proof in [G08] argues that a big combinatorial rectangle cannot be consistent with *any* answer w.r.t. *any* permutation σ, and therefore it makes no difference whether σ is fixed "once forever" or "per rectangle".

- The probability that an answer is returned is $1-(1-r_{1,1})^l \geq \min\left\{\frac{1}{2}, \frac{lp_{1,1}}{2}\right\}$.
- The protocol makes at most k iterations.
- The probability that the transcript of S'_l contains more than $3q'l\sqrt{\log n}$ bits of information about the input is at most $\frac{1}{n^4}$ (this follows from the Chernoff bound and mutual independence of the input instances).

We view S'_l as a protocol for solving the following distributional problem: *the input consists of l independent instances of $P_{1\times 1}$, each chosen according to $\mathcal{U}^n 1\times 1$, and to solve the problem one should provide a correct answer to any given instance of $P_{1\times 1}$.* Note that the probability that S'_l returns an answer is $\Omega\big(1/n^3\big)$ (as we assume that $p_{1,1} > \frac{1}{n^3}$), and therefore we can use Lemma 4.2. The lemma implies that there exists a public coin protocol S'' of cost $O(k\log n)$ that receives l instances of $P_{1\times 1}$ and returns an answer to one of them with probability at least $\min\left\{\frac{1}{4}, \frac{kp_{1,1}}{4q'}\right\}$, such that if an answer is returned it is correct with probability at least $1-801\varepsilon$.

So far we have refrained from using using public randomness in our constructions, in order to be able to use Lemma 4.2 (allowing private randomness only). Now we are going to use public randomness in order to construct a protocol S''' for solving a single instance of $P_{1\times 1}$.

If we apply a uniformly random permutation over $[4n^2]$ to any (x,y) from the support of $\mathcal{U}^n 1\times 1$, we obtain a new pair (x',y'), drawn from $\mathcal{U}^n 1\times 1$. Our S''' will use public randomness to choose l uniformly random permutations $\sigma_1,\ldots,\sigma_l$, and apply them to its input (x,y) to obtain l new instances $(\sigma_1(x),\sigma_1(y)),\ldots,(\sigma_l(x),\sigma_l(y))$. Then S''' will feed these l instances to S''; if S'' outputs (j,c_j) then S''' will return (σ_j,c_j), otherwise it will refuse to answer.

Like S'', our S''' returns an answer with probability at least $\min\left\{\frac{1}{4}, \frac{kp_{1,1}}{4q'}\right\}$, and if an answer is returned it is correct with probability at least $1-801\varepsilon$. Unlike its predecessor, S''' deals with a single input instance of $P_{1\times 1}$, and therefore for sufficiently small ε Lemma 4.3 guarantees that

$$\frac{kp_{1,1}}{4q'} \in O\left(\frac{k^4\log^6 n + k^2\log^8 n}{n^2}\right),$$

which leads to $p_{1,1}\in O\Big(\frac{q'k^3\log^8 n}{n^2}\Big)$, as required. ■*Claim 4.4*

5 Open questions

- Can shared entanglement be stronger than *quantum* communication? In fact, it is wide-open how to compare any model with entanglement to either $\mathcal{Q}$ or $\mathcal{Q}^1$, due to the fact that we do not know how much entanglement can be needed to solve a communication problem of given input length.
- Is it possible to find a *partial functional* problem that requires exponentially more communication in $\mathcal{R}$ than in $\mathcal{R}^{\|,ent}$? How about a *total function*?
- Is it possible to find a problem that requires exponentially more communication in $\mathcal{R}$ than in $\mathcal{Q}^{\|}$?

References

[B] H. Buhrman - *Personal communication*, 2006.

[B64] J. Bell. On the Einstein-Podolsky-Rosen Paradox. *Physics 1(3), pages 195-200*, 1964.

[BCW98] H. Buhrman, R. Cleve and A. Wigderson. Quantum vs. Classical Communication and Computation. *Proceedings of the 30th Symposium on Theory of Computing, pages 63-68*, 1998.

[BCWW01] H. Buhrman, R. Cleve, J. Watrous and R. de Wolf. Quantum Fingerprinting. *Physical Review Letters 87(16), article 167902*, 2001.

[BHMR03] H. Buhrman, P. Høyer, S. Massar and H. Röhrig. Combinatorics and Quantum Nonlocality. *Physical Review Letters 91, article 047903*, 2003.

[BHMR06] H. Buhrman, P. Høyer, S. Massar and H. Röhrig. Multipartite Nonlocal Quantum Correlations Resistant to Imperfections. *Physical Review Letters A 73, article 012321*, 2006.

[BJK04] Z. Bar-Yossef, T. S. Jayram and I. Kerenidis. Exponential Separation of Quantum and Classical One-Way Communication Complexity. *Proceedings of 36th Symposium on Theory of Computing, pages 128-137*, 2004.

[C] R. Cleve - *Personal communication*, 2005.

[CHTW04] R. Cleve, P. Høyer, B. Toner and J. Watrous. Consequences and Limits of Nonlocal Strategies. *Proceedings of the 19th IEEE Conference on Computational Complexity, pages 236-249*, 2004.

[EPR35] A. Einstein, B. Podolsky and N. Rosen. Can Quantum-Mechanical Description of Physical Reality be Considered Complete? *Physical Review 41, pages 777-780*, 1935.

[G08] D. Gavinsky. Classical Interaction Cannot Replace a Quantum Message. *Proceedings of the 40th Symposium on Theory of Computing, pages 95-102*, 2008.

[GKKRW07] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz and R. de Wolf. Exponential Separations for One-Way Quantum Communication Complexity, with Applications to Cryptography. *Proceedings of the 39th Symposium on Theory of Computing, pages 516-525*, 2007.

[GKRW06] D. Gavinsky, J. Kempe, O. Regev and R. de Wolf. Bounded-error Quantum State Identification and Exponential Separations in Communication Complexity. *Proceedings of the 38th Symposium on Theory of Computing, pages 594-603*, 2006.

[HJAR07] P. Harsha, R. Jain, D. McAllester and J. Radhakrishnan. The communication complexity of correlation. *Proceedings of the 22nd IEEE Conference on Computational Complexity, pages 10-23*, 2007.

[R99] R. Raz. Exponential Separation of Quantum and Classical Communication Complexity. *Proceedings of the 31st Symposium on Theory of Computing, pages 358-367*, 1999.

[Y79] A. C-C. Yao. Some Complexity Questions Related to Distributed Computing. *Proceedings of the 11th Symposium on Theory of Computing, pages 209-213*, 1979.