Classical Interaction Cannot Replace a Quantum Message

Dmitry Gavinsky [*]

NEC Laboratories America, Inc.

4 Independence Way, Suite 200

Princeton, NJ 08540, U.S.A.

Abstract

We demonstrate a two-player communication problem that can be solved in the one-way quantum model by a 0-error protocol of cost $O(\log n)$ but requires exponentially more communication in the classical interactive (bounded error) model.

1 Introduction

The ultimate goal of quantum computing is to identify computational tasks that by using the laws of quantum mechanics can be solved more efficiently than on a classical computer.

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*. 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.

- *One-way communication* is the 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.

- *Interactive (two-way) communication* is the 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.

Both 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 $\mathcal{R}^1$ and $\mathcal{R}$, and the quantum versions are denoted by $\mathcal{Q}^1$ 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 sometimes be much more efficient than the latter, both in quantum and in classical versions.

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 players' inputs are called *total*.

A *communication protocol* describes behavior of Alice and Bob 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 $1 - \varepsilon$. If the protocols, in addition, either refuse to answer or succeed, then we say that the solution is *0-error*.

In this paper our primary concern is with separating communication models; more specifically, with finding communication problems that demonstrate super-polynomial advantage of quantum communication over classical one. In fact, both with the previously known examples considered below and with our own contribution it has been the case that the first shown super-polynomial separation had actually been exponential.

It is important to note that the three types of communication tasks mentioned above (relational, functional and total functional) 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; likewise, there are pairs of communication models that can be separated through a partial functional problem but are *widely conjectured* to be equally strong over total functions.

For 0-error, both 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 *simultaneous protocols*, which is a communication model even more limited than one-way. 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. But *does there exist a problem that can be solved by a quantum one-way protocol more efficiently than by any classical two-way protocol?*

1.1 Our result

Theorem 1.1. *For infinitely many $N \in \mathbb{N}$, there exists an (explicit) relation with input length N that can be solved by a 0-error one-way quantum protocol of cost $O(\log N)$ and whose complexity in the interactive classical model is $\Omega\left(\frac{N^{1/8}}{\sqrt{\log N}}\right)$.*

This statement simultaneously subsumes the separation in [BJK04] and, as our theorem speaks about a relational problem, partially that in [R99]. To obtain a similar result for a functional problem is an important open question (see Section 6 for more).

The relation we use is a modification of a communication task independently suggested by R. Cleve ([C]) and S. Massar ([B]) as a possible candidate for such separation.

Some of the intermediate steps in our proof might be of independent interest.

2 Our approach

Denote by $\bar{0}$ the additive identity of a field. For n being a power of 2, define the following communication problems.

Definition 1. Let $x, y \subset [n^2]$, such that $|x| = n/2$ and $|y| = n$. Let $z \in \mathcal{GF}_2^{2\log n} \setminus \{\bar{0}\}$. Let $\Sigma = \{\sigma_{2^{2i}}\}_{i=1}^\infty$ be a set of reversible mappings from $[2^{2i}]$ to $\mathcal{GF}_2^{2i}$. Then $(x, y, z) \in P_{1\times 1}^\Sigma$ if either $|x \cap y| \neq 2$ or $\langle z, a+b \rangle = 0$, where $\sigma_{n^2}(x \cap y) = \{a, b\}$.

Let Σ_0 be the set of reversible mappings from $[2^{2i}]$ to $\mathcal{GF}_2^{2i}$, preserving the lexicographic ordering of the elements.

Definition 2. Let $x \subset [2n^2]$, $|x| = n$. Let $y = (y_1, \ldots, y_{n/4})$ be a tuple of disjoint subsets of $[n^2]$, each of size n, such that $|x \cap y_j| = 2$ for all $1 \leq j \leq n/4$. Let $z \in \mathcal{GF}_2^{2\log n + 1} \setminus \{\bar{0}\}$ and $1 \leq i \leq n/4$, then $(x, y, (i, z)) \in P^{(n)}$ if $\langle z, a+b \rangle = 0$, where $\sigma_0(x \cap y_i) = \{a, b\}$ for some $\sigma_0 \in \Sigma_0$.

In the rest of the paper we will implicitly assume equivalence between the arguments and the corresponding values of every $\sigma_0 \in \Sigma_0$.

We will show that P is easy to solve in $\mathcal{Q}^1$ and is hard for $\mathcal{R}$. In order to prove the lower bound we will use the following modification of $P_{1\times 1}^\Sigma$.

Definition 3. Let $x, y \subset [n^2]$, such that $|x| = n/2$ and $|y| = n$. Let $z \subset [n^2]$. Then $(x, y, z) \in P_{1\times 1}^{search}$ if $x \cap y = z$.

We consider correctness probability of communication protocols w.r.t. input distribution, or the randomness of the protocol, or both. Unless stated otherwise, all available randomness is taken into account, i.e., input distribution is considered whenever it is known and protocol's randomness is considered unless the protocol under consideration is deterministic.

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 $1 - \varepsilon$.

Solving $P_{1\times 1}^\Sigma$ when $|x \cap y| = 2$ requires providing an evidence of knowledge of these elements, and intuitively should be as hard as finding them, as required by $P_{1\times 1}^{search}$, when $|x \cap y| = 2$. This intuition is most likely *false* for the quantum 1-way model (when $|x \cap y| = 2$, $P_{1\times 1}^\Sigma$ can be efficiently[1] solved in $\mathcal{Q}^1$ with probability $1/n$ with small error, which is unlikely to be the case for $P_{1\times 1}^{search}$). However, it is true for the model of classical 2-way communication; a "quasi-reduction" from $P_{1\times 1}^{search}$ to $P_{1\times 1}^\Sigma$ is one of the central ingredients of our lower bound proof.

The high-level structure of the proof is the following.

Solution to P $\implies$ Solution to $\mathbf{P_{1\times 1}^\Sigma}$ *(Lemma 5.1)* We claim that if there exists a protocol that solves P with error bounded by ε then another protocol of similar cost solves $P_{1\times 1}^\Sigma$ for some Σ with probability $\Omega(1/n)$ and error $O(\varepsilon)$.

---

[1] In the context of communication complexity, *efficient* protocols are those of polylogarithmic cost.

Solution to $\mathbf{P_{1\times1}^{\Sigma}}$ $\implies$ Solution to $\mathbf{P_{1\times1}^{search}}$ *(Theorem 5.2)* We reduce the task of solving the problem $P_{1\times1}^{search}$ to that of solving $P_{1\times1}^{\Sigma}$.

$\mathbf{P_{1\times1}^{search}}$ **is hard** *(Theorem 5.6)* We show that the cost of solving $P_{1\times1}^{search}$ with probability δ when $|x \cap y| = 2$ is $\Omega\left(n \cdot \sqrt{\delta}\right)$.

We will conclude that solving P with bounded error requires an interactive classical protocol of complexity $n^{\Omega(1)}$.

3 Notation and more

We assume basic knowledge of (classical) communication complexity ([KN97]).

We will consider only discrete probability distributions. For a set A we write $\mathcal{U}_A$ to denote the uniform distribution over the elements of A. Given a distribution D over A and some $a_0 \in A$ we denote $D(a_0) \overset{\text{def}}{=} \mathbf{Pr}_{\mathbf{X}\sim D}[\mathbf{X} = a_0]$; for $B \subseteq A$, $D(B) \overset{\text{def}}{=} \sum_{b\in B} D(b)$. Denote $\text{supp}(D) \overset{\text{def}}{=} \{a \in A \,|\, D(a) > 0\}$.

Let $\mathbf{X}$, $\mathbf{Y}$ be (discrete) random variables. We let $\mathbf{H}[\mathbf{X}]$ and $\mathbf{H}[\mathbf{X}|\mathbf{Y}]$ denote the corresponding entropy and conditional entropy. As a function of $\mathbf{Y}$, we will denote the conditional entropy by $\mathbf{H}[\mathbf{X}|\mathbf{Y} = y]$.

We will need the Chernoff bound in the following form.

Claim 3.1. *Let $\mathbf{X}_1, \ldots, \mathbf{X}_m$ be random variables, distributed independently and satisfying for some $\mu, \alpha > 0$*

$$\forall 1 \leq i \leq m: \ 0 \leq \mathbf{X}_i \leq \alpha, \ \mathbf{E}[\mathbf{X}_i] \leq \mu \ .$$

Then

$$\mathbf{Pr}\left[\frac{1}{m} \cdot \sum_{i=1}^{m} \mathbf{X}_i \geq (1 + \Omega(1)) \cdot \mu\right] \in 2^{-\Omega(\frac{m\mu}{\alpha})}.$$

We use the following notation.

$$DISJ_n \overset{\text{def}}{=} \left\{(x,y)\,|\,x,y \in \{0,1\}^n, \ \forall i \in [n]: \ x_i = 0 \vee y_i = 0\right\};$$

$$DISJ \overset{\text{def}}{=} \cup_{n\in\mathbb{N}} DISJ_n.$$

We use the standard notion of a *(combinatorial) rectangle*. The sides of considered rectangle always correspond to subsets of the input sets of Alice and Bob, as defined by the communication problem under consideration (to emphasize this, we will sometimes use the term *input rectangle*). We will use the same notation for an input rectangle and for the *event that the input belongs to the rectangle*.

Define context-sensitive "projection operators" $\cdot|.$ and $\cdot\|.$ as follows. For a discrete set A, $x \subseteq A$ and $I \subseteq A$, let $x|_I \overset{\text{def}}{=} x \cap I$. For $B \subseteq 2^A$, let $B\|_I \overset{\text{def}}{=} \{x \cap I \,|\, x \in B\}$. For a distribution D over A, let $D|_I$ be the conditional distribution of $\mathbf{X} \sim D$, subject to $\mathbf{X} \in I$. For a distribution D over 2^A, let $D\|_I$ be the marginal distribution of $\mathbf{Y} \overset{\text{def}}{=} \mathbf{X}|_I$, when $\mathbf{X} \sim D$.

We will use special notation for "one-sided" projections of input pairs. Let $(x, y) \in \mathcal{A} \times \mathcal{B}$, where $\mathcal{A}$ and $\mathcal{B}$ are input sets of Alice and Bob, respectively. Then $(x,y)|_{Alice} \overset{\text{def}}{=} x$ and $(x,y)|_{Bob} \overset{\text{def}}{=} y$. Similarly, define the operators $\|_{Alice}$ and $\|_{Bob}$ for distributions and sets.

3.1 More details on $P_{1\times1}^{\Sigma}$ and P

Define the following events characterizing input to $P_{1\times1}^{\Sigma}$ or $P_{1\times1}^{search}$.

Definition 4. For $j \in \mathbb{N}$, let $\mathcal{X}_j$ be the event that the input pair (x, y) satisfies $|x \cap y| = j$. For $i, j \in \mathbb{N}$, let $\mathcal{X}_1(i)$ and $\mathcal{X}_2(i, j)$ be, respectively, the events that $x \cap y = \{i\}$ and $x \cap y = \{i, j\}$.

We will use the same notation to address the subsets of input that give rise to these events, i.e.,

$$\mathcal{X}_0 \overset{\text{def}}{=} \underset{n=2^i}{\cup} \left\{ (x, y) \in [n^2] \times [n^2] \big| x \cap y = \emptyset \right\},$$

and so forth.

We define $\mathcal{U}_{1\times1}^{(n)}$ to be the uniform distribution of input to $P_{1\times1}^{\Sigma}$, $\mathcal{U}_{Alice} \overset{\text{def}}{=} \mathcal{U}_{1\times1}^{(n)}\|_{Alice}$ and $\mathcal{U}_{Bob} \overset{\text{def}}{=} \mathcal{U}_{1\times1}^{(n)}\|_{Bob}$.

Definition 5. For $k_1, \ldots, k_t \in \mathbb{N}$, let

$$\mathcal{U}_{1\times1}^{(n;k_1,\ldots,k_t)} \overset{\text{def}}{=} \mathcal{U}_{1\times1}^{(n)}|_{\mathcal{X}_{k_1} \cup \cdots \cup \mathcal{X}_{k_t}},$$
$$\mathcal{U}_{1\times1}^{(n;\geq k_1)} \overset{\text{def}}{=} \mathcal{U}_{1\times1}^{(n)}|_{\cup_{i\geq k_1} \mathcal{X}_i}.$$

Definition 6. Given input set A (not necessarily a rectangle), define

$$\mathcal{U}_A \overset{\text{def}}{=} \mathcal{U}_{1\times1}^{(n)}|_A,$$
$$\mathcal{U}_A^{(Alice)} \overset{\text{def}}{=} \mathcal{U}_A\|_{Alice},$$
$$\mathcal{U}_A^{(Bob)} \overset{\text{def}}{=} \mathcal{U}_A\|_{Bob}.$$

Given $k_1, \ldots, k_t \in \mathbb{N}$, let

$$\mathcal{U}_A^{(k_1,\ldots,k_t)} \overset{\text{def}}{=} \mathcal{U}_A|_{\mathcal{X}_{k_1} \cup \cdots \cup \mathcal{X}_{k_t}}.$$

Claim 3.2. *For sufficiently large n it holds that $\mathcal{U}_{1\times1}^{(n)}(\mathcal{X}_0) \geq 1/3$, $\mathcal{U}_{1\times1}^{(n)}(\mathcal{X}_1) \geq 1/6$ and $\mathcal{U}_{1\times1}^{(n)}(\mathcal{X}_2) \geq 1/13$. On the other hand, for any $t \leq n/2$ it holds that $\mathcal{U}_{1\times1}^{(n)} (\cup_{i\geq t} \mathcal{X}_i) \leq \left(\frac{3}{4}\right)^t$.*

Proof of Claim 3.2. Think about choosing input pair $(\mathbf{X}, \mathbf{Y}) \sim \mathcal{U}_{1\times1}^{(n)}$ as selecting a random subset $\mathbf{Y} \subset [n^2]$, subject to $|\mathbf{Y}| = n$, followed by selecting $n/2$ distinct elements for $\mathbf{X}$. Under such interpretation it is clear that $\mathcal{U}_{1\times1}^{(n)} (\cup_{i\geq t} \mathcal{X}_i) \leq \binom{n/2}{t} \cdot \left(\frac{n}{n^2-n/2}\right)^t$. Therefore, $\mathcal{U}_{1\times1}^{(n)}(\mathcal{X}_0) \geq 1 - n/2 \cdot \frac{n}{n^2-n/2} \geq \frac{1}{3}$ and $\mathcal{U}_{1\times1}^{(n)} (\cup_{i\geq t} \mathcal{X}_i) \leq \left(\frac{n}{2}\right)^t \cdot \left(\frac{3}{2n}\right)^t = \left(\frac{3}{4}\right)^t$, for $n \geq 2$.

Let E_i be the event that $i \in \mathbf{X} \cap \mathbf{Y}$. It clearly follows from the symmetry between all E_i-s and from the fact that the events are mutually exclusive when conditioned upon $\mathcal{X}_1$, that $\mathcal{U}_{1\times1}^{(n)}(\mathcal{X}_1)$ is equal to $n/2$ times the probability that the first element selected for $\mathbf{X}$ belongs to $\mathbf{Y}$ and all the following are not from $\mathbf{Y}$. The former occurs with probability at least $1/n$ and the latter with probability not smaller than $\mathcal{U}_{1\times1}^{(n)}(\mathcal{X}_0)$, therefore $\mathcal{U}_{1\times1}^{(n)}(\mathcal{X}_1) \geq \frac{n}{2} \cdot \frac{1}{n} \cdot \frac{1}{3} \geq \frac{1}{6}$.

Similarly, $\mathcal{U}_{1\times1}^{(n)}(\mathcal{X}_2) \geq \binom{n/2}{2} \cdot \frac{1}{n} \cdot \frac{n-1}{n^2-1} \cdot \mathcal{U}_{1\times1}^{(n)}(\mathcal{X}_0) > \frac{1}{13}$, for sufficiently large n. $\blacksquare_{Claim\ 3.2}$

5

3.2 Size of near-monotone rectangles for $DISJ_n$

Two following lemmas can be viewed as the core of our lower bound proof, from both conceptual and technical points of view.

In his elegant lower bound proof for $DISJ$, Razborov [R92] has established the following lemma.

Lemma 3.3. [R92] *Let A be an input rectangle for $DISJ_n$, assume that $n = 4l - 1$. Let D be the following input distribution – with probability $3/4$ Alice and Bob receive two uniformly distributed disjoint subsets of $[n]$ of size l and with probability $1/4$ they receive two uniformly distributed subsets of $[n]$ of size l that share exactly one element. Then*

$$D(A \cap \mathcal{X}_1) \geq \frac{1}{135} \cdot D(A \cap \mathcal{X}_0) - 2^{-\Omega(n)}.$$

We need the following consequence of Lemma 3.3.[2]

Lemma 3.4. *Let n be sufficiently large and A be an input rectangle for $DISJ_n$. Let D be a product distribution of the two halves of the input, such that Alice receives a uniformly chosen subset of $[n]$ of size $k_1(n)$ and Bob receives a uniformly chosen subset of $[n]$ of size $k_2(n)$, where $\alpha_1 \sqrt{n} \leq k_1(n) \leq k_2(n) \leq \alpha_2 \sqrt{n}$ for some α_1, α_2. Then for $\delta = \frac{\alpha_1^2}{45 \cdot 16 \alpha_2^2}$ it holds that*

$$D(A \cap \mathcal{X}_1) \geq \delta \cdot D(A \cap \mathcal{X}_0) - 2^{-\Omega(\sqrt{n})}.$$

Proof of Lemma 3.4. We will reduce the communication task considered in Lemma 3.3 to that defined in the lemma we are proving. Address the former task by P' and the latter one by P (they both are, in fact, versions of $DISJ$, defined w.r.t. different distributions). We will use m to denote the input length to P'. The distribution of input to P' corresponding to m will be denoted by D'_m The length and the distribution of input to P will be denoted by n and D, respectively.

Let $m = 4k_1(n) - 1$. Let T_r be a transformation $(x', y') \to (x, y)$, where $r \in \{0, 1\}^*$, $x', y' \in \{0, 1\}^{[m]}$, and $x, y \in \{0, 1\}^{[n]}$. Think of $\mathbf{R} = r$ as a uniform random string of sufficient length (we will address this situation by "$\mathbf{R} \sim \mathcal{U}$") and of T as a *randomized* transformation of x' and y' only (random bits are implicitly taken from r). In order to compute $T_r(x', y')$ choose randomly and uniformly a pair (M, β) of disjoint subsets of $[n]$ of sizes m and $k_2(n) - l$, respectively (our choice of n guarantees that the latter value is not negative). Define (x, y) by $x|_M = x'$, $y|_M = y'$, $x|_{\overline{M}} = \emptyset$ and $y|_{\overline{M}} = \beta$. Note that T can be applied locally by Alice and Bob if they share public randomness (that is, x only depends on r and x' and y only depends on r and y').

We can see that (x, y) is input to $DISJ_n$ and $DISJ_n(x, y) = DISJ_m(x', y')$, so indeed T is a reduction from $DISJ_m$ to $DISJ_n$. If (x', y') comes from $\mathcal{X}_i \cap \text{supp}(D'_m)$ and $\mathbf{R} \sim \mathcal{U}$

---

[2]Our Lemma 3.4 is similar to a statement made in the original lower bound proof of $\Omega(\sqrt{n})$ for $DISJ$ by Babai, Frankl and Simon [BFS86]. They consider a product distribution similar to our D and give a lower bound on $D(A \setminus \mathcal{X}_0)$ in terms of $D(A \cap \mathcal{X}_0)$, while we need a lower bound on $D(A \cap \mathcal{X}_1)$. We found extending the approach of [BFS86] to be technically more challenging than deriving our statement from a stronger (non-product) case of Lemma 3.3.

then $T_r(x', y')$ is uniformly distributed over $\mathcal{X}_i \cap \mathrm{supp}(D)$, for any $i \geq 0$. In particular, for $i \in \{0, 1\}$,

$$\mathop{\mathbf{E}}_{\mathbf{R} \sim \mathcal{U}} \left[\mathop{\mathbf{Pr}}_{(x', y') \sim D'_m|_{\mathcal{X}_i}} \left[T_r(x', y') \in A \right] \right] = \mathop{\mathbf{Pr}}_{(x, y) \sim D} \left[(x, y) \in A | \mathcal{X}_i \right].$$

For every $r \in \{0, 1\}^*$ let $B_r \stackrel{\text{def}}{=} T_r^{-1}(A)$. It holds that

$$\mathop{\mathbf{Pr}}_{(x', y') \sim D'_m|_{\mathcal{X}_i}} \left[T_r(x', y') \in A \right] = D'_m|_{\mathcal{X}_i}(B_r) = \frac{D'_m(B_r \cap \mathcal{X}_i)}{D'_m(\mathcal{X}_i)},$$

therefore

$$\mathop{\mathbf{E}}_{\mathbf{R} \sim \mathcal{U}} \left[D'_m(B_r \cap \mathcal{X}_i) \right] = \frac{D'_m(\mathcal{X}_i)}{D(\mathcal{X}_i)} \cdot D(A \cap \mathcal{X}_i).$$

It is clear that T_r is rectangle-invariant, so B_r-s are rectangles and we can apply Lemma 3.3.

$$\begin{aligned}
-2^{-\Omega(\sqrt{n})} = -2^{-\Omega(m)} &\leq \mathop{\mathbf{E}}_{\mathbf{R} \sim \mathcal{U}} \left[D'_m(B \cap \mathcal{X}_1) - \frac{D'_m(B \cap \mathcal{X}_0)}{135} \right] \\
&= \mathop{\mathbf{E}}_{\mathbf{R} \sim \mathcal{U}} \left[D'_m(B \cap \mathcal{X}_1) \right] - \frac{1}{135} \cdot \mathop{\mathbf{E}}_{\mathbf{R} \sim \mathcal{U}} \left[D'_m(B \cap \mathcal{X}_0) \right] \\
&= \frac{D'_m(\mathcal{X}_1)}{D(\mathcal{X}_1)} \cdot D(A \cap \mathcal{X}_1) - \frac{D'_m(\mathcal{X}_0)}{135 \cdot D(\mathcal{X}_0)} \cdot D(A \cap \mathcal{X}_0).
\end{aligned}$$

Together with the facts that $D'_m(\mathcal{X}_0) = \frac{3}{4}$ and $D'_m(\mathcal{X}_1) = \frac{1}{4}$, it implies that

$$\begin{aligned}
D(A \cap \mathcal{X}_1) &\geq \frac{D(\mathcal{X}_1)}{135 \cdot D(\mathcal{X}_0)} \cdot \frac{D'_m(\mathcal{X}_0)}{D'_m(\mathcal{X}_1)} \cdot D(A \cap \mathcal{X}_0) - \frac{D(\mathcal{X}_1)}{D'_m(\mathcal{X}_1)} \cdot 2^{-\Omega(\sqrt{n})} \\
&\geq \frac{D(\mathcal{X}_1)}{45} \cdot D(A \cap \mathcal{X}_0) - 2^{-\Omega(\sqrt{n})}.
\end{aligned}$$

Note that

$$D(\mathcal{X}_0) \geq \left(\frac{n - k_1(n) - k_2(n)}{n} \right)^{k_2(n)} \geq \left(1 - \frac{2\alpha_2}{\sqrt{n}} \right)^{\alpha_2 \sqrt{n}} \geq \left(\frac{1}{2} \right)^{4\alpha_2^2} = \left(\frac{1}{16} \right)^{\alpha_2^2},$$

$$D(\mathcal{X}_1) \geq k_2(n) \cdot \frac{k_1(n)}{n} \cdot D(\mathcal{X}_0) \geq \frac{\alpha_1^2}{16^{\alpha_2^2}}$$

(the second inequality can be established analogously to the proof of Claim 3.2). The result follows. $\blacksquare$ *Lemma 3.4*

4   Efficient protocol for P in $\mathcal{Q}^1$

We give a 1-way quantum protocol S that receives input to P, communicates $O(\log n)$ qubits and either produces a correct answer or refuses to answer. For n large enough the former occurs with probability at least $\frac{1}{3}$. Therefore, for any given ε one can run $t \in O\left(\log\left(\frac{1}{\varepsilon}\right)\right)$ instances of S in parallel, thus obtaining a 0-error protocol for P with answering probability at least $1 - \varepsilon$. The communication cost of the new protocol remains in $O(\log n)$ as long as ε is a constant.

Let us see how S works.

1. Alice sends to Bob the state $|\alpha\rangle \overset{\text{def}}{=} \frac{1}{\sqrt{n}} \sum_{j \in x} |j\rangle$.

2. Bob measures $|\alpha\rangle$ with the $\frac{n}{4} + 1$ projectors $E_i \overset{\text{def}}{=} \sum_{j \in y_i} |j\rangle\langle j|$ and $E_0 \overset{\text{def}}{=} \sum_{j \notin \cup y_i} |j\rangle\langle j|$, let i_0 be the index of the outcome of the measurement and $|\alpha_{i_0}\rangle$ be the projected state. Bob applies the Hadamard transform over $\mathcal{GF}_2^{2\log n + 1}$ to $|\alpha_{i_0}\rangle$ and measures the result in the computational basis. Denote by a_{i_0} be the outcome of the measurement.

3. If $a_{i_0} = \bar{0}$ or $i_0 = 0$ then Bob refuses to answer, otherwise he outputs (i_0, a_{i_0}).

Obviously, the protocol transmits $O(\log n)$ qubits.

After the first measurement, if $i_0 = 0$ then Bob refuses to answer, otherwise the register remains in the state $|\alpha_{i_0}\rangle = \frac{1}{\sqrt{2}} \sum_{j \in x \cap y_{i_0}} |j\rangle$. Denote by p_i the probability that $i_0 = i$. Then for $i > 0$,

$$p_i = \text{tr}\left(|\alpha\rangle\langle\alpha| \cdot E_i\right) = \frac{1}{n} \cdot \text{tr}\left(\left(\sum_{\substack{j,k \in x \\ j \neq k}} |j\rangle\langle k| + \sum_{j \in x} |j\rangle\langle j|\right) \cdot \sum_{j \in y_i} |j\rangle\langle j|\right) = \frac{|x \cap y_i|}{n} = \frac{2}{n},$$

and consequently, $p_0 = 1 - \sum_{i>0} p_i = 1/2$.

Assume that $i_0 \neq 0$. Bob applies the Hadamard transform to the state $|\alpha_{i_0}\rangle = \frac{|b_1\rangle + |b_2\rangle}{\sqrt{2}}$ where $x \cap y_{i_0} = \{b_1, b_2\}$, denote the outcome by $|\alpha'_{i_0}\rangle$. Then

$$|\alpha'_{i_0}\rangle = \frac{1}{2n} \cdot \sum_{j \in [2n^2]} \left((-1)^{\langle j, b_1\rangle} + (-1)^{\langle j, b_2\rangle}\right) |j\rangle = \frac{1}{n} \cdot \sum_{\langle j, b_1 + b_2\rangle = 0} \pm |j\rangle,$$

and therefore Bob obtains a uniformly random element of

$$\left\{j \in [2n^2] \mid \langle j, b_1 + b_2\rangle = 0\right\},$$

as the outcome of his second measurement.

If $a_{i_0} = \bar{0}$ then Bob refuses to answer, otherwise he returns a pair (i_0, a_{i_0}) that satisfies the requirement. The latter occurs with probability $1 - o(1)$, conditioned on $i_0 \neq 0$. So, the protocol is successful with probability $\frac{1}{2} - o(1) > \frac{1}{3}$, for sufficiently large n.

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

We will establish a lower bound of $\frac{n^{1/4}}{\sqrt{\log n}}$ for the 2-way classical communication complexity of P. We will always assume this model of communication, unless stated otherwise.

As outlined in Section 2, we will first prove that solving P implies solving $P_{1 \times 1}^{\Sigma}$, then that solving $P_{1 \times 1}^{search}$ is as simple as solving $P_{1 \times 1}^{\Sigma}$, and finally that solving $P_{1 \times 1}^{search}$ is expensive.

5.1 Solving P implies solving $P_{1 \times 1}^{\Sigma}$

Lemma 5.1. *Assume that there exists a (possibly, randomized) protocol S of cost k that solves P with error bounded by ε. Then there exists Σ, such that $P_{1 \times 1}^{\Sigma}$ can be solved w.r.t. $\mathcal{U}_{1 \times 1}^{(n;2)}$ with probability $2/n$ with error bounded by 2ε by a deterministic protocol of cost k.*

Proof of Lemma 5.1. Let (x, y) be an instance of $P_{1 \times 1}^{\Sigma}$, satisfying $|x \cap y| = 2$ (recall that x and y are subsets of $[n^2]$, $|x| = n/2$ and $|y| = n$). Consider the following protocol S'.

- Let $x' = \{n^2 + 1, \ldots, n^2 + \frac{n}{2}\} \cup x$. For $1 \le j \le \frac{n}{4} - 1$, let $y'_j = \{n^2 + j + \frac{kn}{4} | 1 \le k \le n\}$ and $\bar{y} = (y, y'_1, \ldots, y'_{\frac{n}{4}-1})$.

- Using public randomness, choose random permutations: σ_1 over $[2n^2]$ and σ_2 over $[\frac{n}{4}]$.

- Run the protocol S over $\sigma_1(x', (\bar{y}_{\sigma_2(1)}, \ldots, \bar{y}_{\sigma_2(n/4)}))$; let (i, z) be the response by S.

- If $\sigma_2(1) = i$ then output (σ_1, z), otherwise refuse to answer.

This protocol maps the given pair (x, y) to a uniformly random instance of P (the deterministically constructed $(x', \bar{y})$ forms a correct input for P, and the action of permutations upon instances of P is transitive). Moreover, the original problem is mapped to a uniformly random coordinate of the instance of P that is fed into S.

Let $(\mathbf{X}, \mathbf{Y}) = (x, y)$ and $(\mathbf{X}, \mathbf{Y}) \sim \mathcal{U}_{1 \times 1}^{(n;2)}$. Denote by $\mathcal{E}$ the event that S' returns an answer, by $\mathcal{E}_0$ the event that S' outputs a pair (σ, z) such that $z \ne \bar{0}$ and $\langle z, a + b \rangle = 0$, where $\sigma(x \cap y) = \{a, b\}$, and by $\mathcal{E}_1$ the event $\mathcal{E} \setminus \mathcal{E}_0$. By the symmetry argument, the following holds: If S returns a correct answer then $\mathcal{E}_0$ occurs with probability $4/n$; if S makes a mistake then $\mathcal{E}_1$ occurs with probability $4/n$. In particular, $\mathbf{Pr}[\mathcal{E}] = 4/n$ and $\mathbf{Pr}[\mathcal{E}_1] \le \varepsilon \cdot \mathbf{Pr}[\mathcal{E}]$.

Let us derandomize S'. Suppose that S' uses s random bits and let $\mathbf{R}$ be the corresponding random variable. Let R_0 be the set of $r \in \{0, 1\}^s$, such that $\mathbf{Pr}[\mathcal{E}_1 | \mathcal{E}, \mathbf{R} = r] \ge 2\varepsilon$. From the properties of S it follows that

$$\varepsilon \cdot \mathbf{Pr}[\mathcal{E}] \ge \mathbf{Pr}[\mathcal{E}_1] = \mathbf{Pr}[\mathcal{E}] \cdot \mathbf{Pr}[\mathcal{E}_1 | \mathcal{E}] \ge \mathbf{Pr}[\mathbf{R} \in R_0] \cdot \mathbf{Pr}[\mathcal{E} | \mathbf{R} \in R_0] \cdot 2\varepsilon,$$

which leads to

$$\frac{1}{2} \mathbf{Pr}[\mathcal{E}] \le \mathbf{Pr}[\mathcal{E}] - \mathbf{Pr}[\mathbf{R} \in R_0] \cdot \mathbf{Pr}[\mathcal{E} | \mathbf{R} \in R_0] = \mathbf{Pr}[\mathbf{R} \notin R_0] \cdot \mathbf{Pr}[\mathcal{E} | \mathbf{R} \notin R_0].$$

Therefore, there exists some $r_0 \notin R_0$, such that $\mathbf{Pr}[\mathcal{E} | \mathbf{R} = r_0] \ge \mathbf{Pr}[\mathcal{E}]/2 = 2/n$ and $\mathbf{Pr}[\mathcal{E}_1 | \mathcal{E}, \mathbf{R} = r_0] < 2\varepsilon$.

Define a deterministic protocol S'', which is similar to S' but uses r_0 instead of the random string and outputs only z. Observe that fixing $\mathbf{R} = r_0$, in particular, fixes the mapping $\sigma_1 \stackrel{\text{def}}{=} \sigma'_1$. Let Σ consist of σ'_1-s, obtained as a result of the described derandomization, subsequently applied to every permitted input length. We claim that S'' solves $P_{1 \times 1}^{\Sigma}$ w.r.t. $\mathcal{U}_{1 \times 1}^{(n;2)}$ with probability at least $2/n$ with error bounded by 2ε – this follows from the aforementioned properties of S' and the definition of Σ. The complexity of S'' is k, as pre- and post-processing are performed locally. $\blacksquare$ *Lemma 5.1*

5.2 Solving $P_{1 \times 1}^{\Sigma}$ implies solving $P_{1 \times 1}^{search}$

We will show the following.

Theorem 5.2. *Assume that there exists a deterministic protocol of cost $k \in o(n) \cap \omega(1)$ that solves $P_{1 \times 1}^{\Sigma}$ for some Σ w.r.t. $\mathcal{U}_{1 \times 1}^{(n;2)}$ with probability $\gamma \in \omega(2^{-k})$ and error bounded by 10^{-22}. Then $P_{1 \times 1}^{search}$ can be solved w.r.t. $\mathcal{U}_{1 \times 1}^{(n;2)}$ with probability $\frac{\gamma}{k^2 \cdot \log^2(n/\gamma)}$ with error 0 by a public coin protocol of cost $O(k + \log^2(n/\gamma))$.*

The proof will be done in several stages.

Lemma 5.3. *Let n be sufficiently large and A be an input rectangle for $P_{1\times 1}^{\Sigma}$, such that $\mathcal{U}_{1\times 1}^{(n;1)}(A) \in 2^{-o(n)} \cap o(1)$. Assume that for some constant $0 < \varepsilon < 1$ and $I_0 \subseteq [n^2]$, $|I_0| \geq \frac{n^2}{2}$, it holds that*

$$\sum_{i \in I_0} \mathcal{U}_A^{(1)}\left(\mathcal{X}_1(i)\right) \leq \frac{8\varepsilon}{10^7}.$$

Then $\mathcal{U}_A^{(0,1)}(\mathcal{X}_0) < \varepsilon$.

The meaning of the lemma is that a rectangle accepting input pairs from $\mathcal{X}_1$, but mostly those that intersect *not over* I_0, must reject, with high probability, pairs from $\mathcal{X}_0$.

Proof of Lemma 5.3. In this proof we will casually view input pairs (x, y) as 4-tuples (x_1, x_2, y_1, y_2), where $x|_{\overline{I_0}} = x_1$, $x|_{I_0} = x_2$, $y|_{\overline{I_0}} = y_1$, $y|_{I_0} = y_2$.

Let $\varepsilon_0 \stackrel{\text{def}}{=} \mathcal{U}_A^{(0,1)}(\mathcal{X}_0) \in \Omega(1)$, in terms of this value we will derive a lower bound on the probability that a uniformly chosen $\mathcal{X}_1$-instance from A intersects over I_0.

Let $(\mathbf{X}, \mathbf{Y}) = (\mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}_1, \mathbf{Y}_2) \sim \mathcal{U}_A^{(0,1)}$. Let x_1 and y_1 be the values taken by $\mathbf{X}_1$ and $\mathbf{Y}_1$, we define the following events characterizing these values (note that the events do not depend on the values of $\mathbf{X}_2$ and $\mathbf{Y}_2$):

- $\mathcal{E}_1$ denotes the event that $|x_1| \leq \frac{n}{3}$ and $|y_1| \leq \frac{2n}{3}$.

- $\mathcal{E}_2$ denotes the event that $\mathbf{Pr}_{\mathcal{U}_A^{(0,1)}}\left[\mathcal{X}_0 \big| \mathbf{X}_1 = x_1, \mathbf{Y}_1 = y_1\right] \geq \frac{\varepsilon_0}{2}$. Observe that $\mathcal{E}_2$ implies that $(x_1, y_1) \in \text{supp}\left(\mathcal{U}_A^{(0)}\|_{\overline{I_0} \times \overline{I_0}}\right)$.

- $\mathcal{E}_3$ denotes the event that either $\neg \mathcal{E}_2$ or $\mathcal{E}_2$ and

$$\mathbf{H}\left[\mathcal{U}_A^{(0)} \Big| \mathbf{X}_1 = x_1, \mathbf{Y}_1 = y_1\right] \geq \mathbf{H}\left[\mathcal{U}_{1\times 1}^{(n;0)}\|_{I_0 \times I_0}\right] - \left(\frac{8}{\varepsilon_0} + 1\right) \cdot \log\left(\frac{1}{\mathcal{U}_{1\times 1}^{(n;0)}(A)}\right).$$

- $\mathcal{E}_4$ denotes the event that either $\neg \mathcal{E}_2$ or $\mathcal{E}_2$ and

$$\mathbf{H}\left[\mathcal{U}_{1\times 1}^{(n;0)} \Big| \mathbf{X}_1 = x_1, \mathbf{Y}_1 = y_1\right] \leq \mathbf{H}\left[\mathcal{U}_{1\times 1}^{(n;0)}\|_{I_0 \times I_0}\right] + \log\left(\frac{8}{\varepsilon_0 \cdot \mathcal{U}_{1\times 1}^{(n;0,1)}(A)}\right).$$

Our first step will be to show that all four events hold simultaneously with non-negligible probability. This will let us apply Lemma 3.4 to many "subrectangles" of A defined over $I_0 \times I_0$, which, in turn, will lead to the desired lower bound.

The event $\mathcal{E}_1$ occurs with probability $1 - 2^{-\Omega(n)}$ if $(\mathbf{X}_1, \mathbf{Y}_1) \sim \mathcal{U}_{1\times 1}^{(n;0,1)}\|_{\overline{I_0} \times \overline{I_0}}$, due to the Chernoff bound (Claim 3.1). In our case $(\mathbf{X}_1, \mathbf{Y}_1) \sim \mathcal{U}_A^{(0,1)}\|_{\overline{I_0} \times \overline{I_0}}$, but on the other hand, $\mathcal{U}_{1\times 1}^{(n;0,1)}(A) \in 2^{-o(n)}$, and therefore $\mathbf{Pr}_{\mathcal{U}_A^{(0,1)}}[\mathcal{E}_1] \in 1 - o(1)$.

We know that

$$\varepsilon_0 = \mathcal{U}_A^{(0,1)}(\mathcal{X}_0) = \mathop{\mathbf{E}}_{(x_1, y_1) \sim \mathcal{U}_A^{(0,1)}\|_{\overline{I_0} \times \overline{I_0}}}\left[\mathop{\mathbf{Pr}}_{\mathcal{U}_A^{(0,1)}}\left[\mathcal{X}_0 \big| \mathbf{X}_1 = x_1, \mathbf{Y}_1 = y_1\right]\right],$$

which implies that $\mathbf{Pr}_{\mathcal{U}_A^{(0,1)}}[\mathcal{E}_2] \geq \frac{\varepsilon_0}{2}$.

Let us see that $\mathcal{E}_3$ occurs with high probability. Observe that by the chain rule,

$$
\begin{aligned}
\mathbf{H}\left[\mathcal{U}_A^{(0)}\right] &= \mathbf{H}\left[\mathcal{U}_A^{(0)}\|_{\overline{I_0}\times\overline{I_0}}\right] + \underset{\mathcal{U}_A^{(0)}}{\mathbf{H}}\left[\mathbf{X}_2,\mathbf{Y}_2|\mathbf{X}_1,\mathbf{Y}_1\right], \\
\mathbf{H}\left[\mathcal{U}_{1\times 1}^{(n;0)}\right] &= \mathbf{H}\left[\mathcal{U}_{1\times 1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}\right] + \mathbf{H}\left[\mathcal{U}_{1\times 1}^{(n;0)}\|_{I_0\times I_0}\right],
\end{aligned}
\tag{1}
$$

where the last equality follows from the fact that $\mathcal{U}_{1\times 1}^{(n;0)}$ is a product distribution of its two marginal projections, as appear on the right-hand side. Moreover, these projections are uniform over their supports, and therefore

$$
\mathbf{H}\left[\mathcal{U}_{1\times 1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}\right] \geq \mathbf{H}\left[\mathcal{U}_A^{(0)}\|_{\overline{I_0}\times\overline{I_0}}\right]
\tag{2}
$$

and for any (x_1,y_1) in the support of $\mathcal{U}_A^{(0)}\|_{\overline{I_0}\times\overline{I_0}}$,

$$
\underset{\mathcal{U}_A^{(0)}}{\mathbf{H}}\left[\mathbf{X}_2,\mathbf{Y}_2|\mathbf{X}_1=x_1,\mathbf{Y}_1=y_1\right] \leq \mathbf{H}\left[\mathcal{U}_{1\times 1}^{(n;0)}\|_{I_0\times I_0}\right].
\tag{3}
$$

In particular, (2) and (1) imply that

$$
\mathbf{H}\left[\mathcal{U}_{1\times 1}^{(n;0)}\right] - \mathbf{H}\left[\mathcal{U}_A^{(0)}\right] \geq \mathbf{H}\left[\mathcal{U}_{1\times 1}^{(n;0)}\|_{I_0\times I_0}\right] - \underset{\mathcal{U}_A^{(0)}}{\mathbf{H}}\left[\mathbf{X}_2,\mathbf{Y}_2|\mathbf{X}_1,\mathbf{Y}_1\right].
\tag{4}
$$

Observe that both $\mathcal{U}_{1\times 1}^{(n;0)}$ and $\mathcal{U}_A^{(0)}$ are uniform over their supports; moreover, the latter support is a subset of the former. This leads to

$$
\mathbf{H}\left[\mathcal{U}_{1\times 1}^{(n;0)}\right] - \mathbf{H}\left[\mathcal{U}_A^{(0)}\right] = \log\left(\frac{\left|\operatorname{supp}\left(\mathcal{U}_{1\times 1}^{(n;0)}\right)\right|}{\left|\operatorname{supp}\left(\mathcal{U}_A^{(0)}\right)\right|}\right) = \log\left(\frac{1}{\mathcal{U}_{1\times 1}^{(n;0)}(A)}\right),
$$

that, together with (4), gives us

$$
\underset{\mathcal{U}_A^{(0)}}{\mathbf{H}}\left[\mathbf{X}_2,\mathbf{Y}_2|\mathbf{X}_1,\mathbf{Y}_1\right] \geq \mathbf{H}\left[\mathcal{U}_{1\times 1}^{(n;0)}\|_{I_0\times I_0}\right] - \log\left(\frac{1}{\mathcal{U}_{1\times 1}^{(n;0)}(A)}\right).
$$

Together with (3) this implies, by the Markov inequality, that $\mathbf{Pr}_{\mathcal{U}_A^{(0,1)}}[\mathcal{E}_3] \geq 1 - \frac{\varepsilon_0}{8}$.

Let us denote by G the set of pairs (x_1,y_1) that falsify the condition of $\mathcal{E}_4$. Then, starting from (1), we get

$$
\begin{aligned}
\mathbf{H}\left[\mathcal{U}_{1\times 1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}\right] + \mathbf{H}\left[\mathcal{U}_{1\times 1}^{(n;0)}\|_{I_0\times I_0}\right] &= \mathbf{H}\left[\mathcal{U}_{1\times 1}^{(n;0)}\right] \geq \mathbf{H}\left[\mathcal{U}_{1\times 1}^{(n;0)}\Big|(\mathbf{X}_1,\mathbf{Y}_1)\in G\right] \\
&= \mathbf{H}\left[\left(\mathcal{U}_{1\times 1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}\right)\Big|_G\right] + \underset{\mathcal{U}_{1\times 1}^{(n;0)}|(\mathbf{X_1},\mathbf{Y_1})\in\mathbf{G}}{\mathbf{H}}\left[\mathbf{X}_2,\mathbf{Y}_2|\mathbf{X}_1,\mathbf{Y}_1\right] \\
&\geq \mathbf{H}\left[\left(\mathcal{U}_{1\times 1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}\right)\Big|_G\right] + \mathbf{H}\left[\mathcal{U}_{1\times 1}^{(n;0)}\|_{I_0\times I_0}\right] + \log\left(\frac{8}{\varepsilon_0\cdot\mathcal{U}_{1\times 1}^{(n;0,1)}(A)}\right),
\end{aligned}
$$

where the last inequality is implied by the definition of G. Therefore,

$$\mathbf{H}\left[\left(\mathcal{U}_{1\times1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}\right)\Big|_G\right] \leq \mathbf{H}\left[\mathcal{U}_{1\times1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}\right] - \log\left(\frac{8}{\varepsilon_0 \cdot \mathcal{U}_{1\times1}^{(n;0,1)}(A)}\right).$$

Both arguments of $\mathbf{H}\left[\cdot\right]$ in the last inequality are uniform distributions over their supports, one being a subset of the other, which gives us

$$\log\left(\frac{1}{\mathcal{U}_{1\times1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}(G)}\right) = \mathbf{H}\left[\mathcal{U}_{1\times1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}\right] - \mathbf{H}\left[\left(\mathcal{U}_{1\times1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}\right)\Big|_G\right]$$

$$\geq \log\left(\frac{8}{\varepsilon_0 \cdot \mathcal{U}_{1\times1}^{(n;0,1)}(A)}\right).$$

This leads to $\mathcal{U}_{1\times1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}(G) \leq \frac{\varepsilon_0}{8} \cdot \mathcal{U}_{1\times1}^{(n;0,1)}(A)$. Note that G by definition implies $\mathcal{E}_2$, thus consists exclusively of disjoint pairs, and therefore $\mathcal{U}_{1\times1}^{(n;0,1)}\|_{\overline{I_0}\times\overline{I_0}}(G) \leq \mathcal{U}_{1\times1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}(G) \leq \frac{\varepsilon_0}{8} \cdot \mathcal{U}_{1\times1}^{(n;0,1)}(A)$. Therefore,

$$\Pr_{\mathcal{U}_A^{(0,1)}}\left[\mathcal{E}_4\right] = 1 - \mathcal{U}_A^{(0,1)}\|_{\overline{I_0}\times\overline{I_0}}(G) \geq 1 - \frac{\varepsilon_0}{8}.$$

For n sufficiently large, the events $\mathcal{E}_1$, $\mathcal{E}_3$ and $\mathcal{E}_4$ simultaneously hold with probability at least $1 - \frac{\varepsilon_0}{4} - o\left(1\right) > 1 - \frac{\varepsilon_0}{3}$, and $\mathcal{E}_2$ holds with probability at least $\frac{\varepsilon_0}{2}$. The event $\mathcal{E} \stackrel{\text{def}}{=} \mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3 \cap \mathcal{E}_4$ therefore holds with probability at least $\frac{\varepsilon_0}{6}$ when $(\mathbf{X}_1, \mathbf{Y}_1) \sim \mathcal{U}_A^{(0,1)}\|_{\overline{I_0}\times\overline{I_0}}$.

If $\mathcal{E}$ holds w.r.t. $\mathbf{X}_1 = x_1'$ and $\mathbf{Y}_1 = y_1'$ then we can apply Lemma 3.4 to the rectangle $A_{x_1',y_1'} \stackrel{\text{def}}{=} \left\{(x_2,y_2)\big|(x_1',x_2,y_1',y_2) \in A\right\}$, as follows. Let us view $A_{x_1',y_1'}$ as an input rectangle for $DISJ_{|I_0|}$. Denote input to $DISJ_{|I_0|}$ by $(\mathbf{X}_2, \mathbf{Y}_2)$. Define D to be the distribution obtained by independently choosing $\mathbf{X}_2$ and $\mathbf{Y}_2$ as subsets of I_0 of sizes $\frac{n}{2} - |x_1'|$ and $n - |y_1'|$, respectively. As follows from $\mathcal{E}_1$, $\frac{n}{6} \leq |\mathbf{X}_2| \leq \frac{n}{2}$ and $\frac{n}{3} \leq |\mathbf{Y}_2| \leq n$.

Observe that the mapping $M : I_0 \times I_0 \to [n^2] \times [n^2]$ defined as

$$M(x_2,y_2) \stackrel{\text{def}}{=} (x_1' \cup x_2, \ y_1' \cup y_2)$$

transforms D into $\mathcal{U}_{1\times1}^{(n)}$, conditioned upon $\mathbf{X}_1 = x_1', \mathbf{Y}_1 = y_1'$. The fact that $x_1' \cap y_1' = \emptyset$ (as implied by $\mathcal{E}_2$) means that the M transforms D into $\mathcal{U}_{1\times1}^{(n)}|_{\mathbf{X}_1=x_1',\mathbf{Y}_1=y_1'}$ also when the both distributions are conditioned upon some $\mathcal{X}_j$, for any valid j. Note also that M maps A to $A_{x_1',y_1'}$. In what follows we will be implicitly assuming equivalence between the arguments and the corresponding values of M, whenever necessary.

Lemma 3.4 can be applied to $A_{x_1',y_1'}$ w.r.t. the distribution D by choosing $\alpha_1 = \frac{n}{6\sqrt{|I_0|}}$ and $\alpha_2 = \frac{n}{\sqrt{|I_0|}}$. The conclusion is that for $\delta = \frac{\alpha_1{}^2}{45\cdot16\alpha_2{}^2} \geq \frac{1}{207360}$,

$$D(A_{x_1',y_1'} \cap \mathcal{X}_1) \geq \delta \cdot D(A_{x_1',y_1'} \cap \mathcal{X}_0) - 2^{-\Omega\left(\sqrt{|I_0|}\right)} \geq \frac{D(A_{x_1',y_1'} \cap \mathcal{X}_0)}{207360} - 2^{-\Omega(n)}. \quad (5)$$

Let $D_A \stackrel{\text{def}}{=} D|_{A_{x_1',y_1'}}$ and $D_0 \stackrel{\text{def}}{=} D|_{\mathcal{X}_0}$. Events $\mathcal{E}_3$ and $\mathcal{E}_4$ together mean that for n sufficiently large (recall that $\mathcal{U}_{1\times 1}^{(n;0)}(A) \in o(1)$),

$$\mathbf{H}\left[\mathcal{U}_A^{(0)}\Big|\mathbf{X}_1 = x_1', \mathbf{Y}_1 = y_1'\right] \geq \mathbf{H}\left[\mathcal{U}_{1\times 1}^{(n;0)}\Big|\mathbf{X}_1 = x_1', \mathbf{Y}_1 = y_1'\right] - \Delta,$$

where $\Delta \in O\left(\log\left(\frac{1}{\mathcal{U}_{1\times 1}^{(n;0)}(A)}\right)\right)$, as follows from $\mathcal{U}_{1\times 1}^{(n;0,1)}(A) \geq \mathcal{U}_{1\times 1}^{(n;0,1)}(\mathcal{X}_0) \cdot \mathcal{U}_{1\times 1}^{(n;0)}(A) \in \Omega\left(\mathcal{U}_{1\times 1}^{(n;0)}(A)\right)$. That can be restated as $\mathbf{H}\left[D_A|_{\mathcal{X}_0}\right] \geq \mathbf{H}\left[D_0\right] - \Delta$, and again, since the both arguments of $\mathbf{H}\left[\cdot\right]$ are uniform distributions, one's support being a subset of the other's, this leads to

$$D_0(A_{x_1',y_1'}) \geq 2^{-\Delta} = \left(\mathcal{U}_{1\times 1}^{(n;0)}(A)\right)^{O(1)}. \tag{6}$$

We know that $\mathcal{U}_{1\times 1}^{(n;1)}(A) \in 2^{-o(n)}$. Therefore,

$$\mathcal{U}_{1\times 1}^{(n;0,1)}(\mathcal{X}_0 \cup A) = \mathcal{U}_{1\times 1}^{(n;0,1)}(A) \cdot \mathcal{U}_A^{(0,1)}(\mathcal{X}_0) = \mathcal{U}_{1\times 1}^{(n;0,1)}(\mathcal{X}_0) \cdot \mathcal{U}_{1\times 1}^{(n;0)}(A)$$

together with $\mathcal{U}_{1\times 1}^{(n;0,1)}(\mathcal{X}_0) \in \Omega(1)$ imply

$$\mathcal{U}_{1\times 1}^{(n;0)}(A) \in \Omega\left(\mathcal{U}_{1\times 1}^{(n;0,1)}(A)\right) \subseteq \Omega\left(\mathcal{U}_{1\times 1}^{(n;1)}(A)\right) \subseteq 2^{-o(n)}, \tag{7}$$

and therefore $D_0(A_{x_1',y_1'}) \in 2^{-o(n)}$. By the definition of D it is easy to see that $D(\mathcal{X}_0) \in \Omega(1)$, and therefore

$$D(A_{x_1',y_1'} \cap \mathcal{X}_0) = D(\mathcal{X}_0) \cdot D_0(A_{x_1',y_1'}) \in 2^{-o(n)}.$$

This means that for sufficiently large n, (5) leads to $D(A_{x_1',y_1'} \cap \mathcal{X}_1) > D(A_{x_1',y_1'} \cap \mathcal{X}_0)/207361$, implying $\mathbf{Pr}_{D_A}\left[\mathcal{X}_1\big|\mathcal{X}_0 \cup \mathcal{X}_1\right] > \frac{1}{207361}$.

We know that $\mathcal{E}$ occurs with probability at least $\frac{\varepsilon_0}{6}$, and thus

$$\sum_{i \in I_0} \mathcal{U}_A^{(1)}\left(\mathcal{X}_1(i)\right) \geq \sum_{i \in I_0} \mathcal{U}_A^{(0,1)}\left(\mathcal{X}_1(i)\right) \geq \Pr_{\mathcal{U}_A^{(0,1)}|_{\overline{I_0} \times \overline{I_0}}}[\mathcal{E}] \cdot \mathbf{Pr}^*[\mathcal{X}_1|\mathcal{X}_0 \cup \mathcal{X}_1] > \frac{8\varepsilon_0}{10^7},$$

where $\mathbf{Pr}^*[\mathcal{X}_1|\mathcal{X}_0 \cup \mathcal{X}_1]$ denotes the maximum possible value of $\mathbf{Pr}\left[\mathcal{X}_1\big|\mathcal{X}_0 \cup \mathcal{X}_1\right]$, taken over D_A that is defined as above, over some pair x_1', y_1' for which $\mathcal{E}$ holds. The result follows.

$$\blacksquare {}_{Lemma\ 5.3}$$

We will need the following extension of Lemma 5.3 to the case of rectangles, selectively accepting instances of $\mathcal{X}_2$.

Lemma 5.4. *Let n be sufficiently large and A be an input rectangle for $P_{1\times 1}^{\Sigma}$, such that $\mathcal{U}_{1\times 1}^{(n;2)}(A) \in 2^{-o(n)}$. Let $\left\{I_0^{(i)}\right\}_{i \in [n^2]}$ be a family of subsets of $[n^2]$, such that for every $i, j \in [n^2]$ it holds that $i \notin I_0^{(i)}$, $|I_0^{(i)}| \geq \frac{n^2}{2}$, and $i \in I_0^{(j)}$ if and only if $j \in I_0^{(i)}$. If A satisfies that*

$$\frac{1}{2}\sum_{\substack{i \in [n^2] \\ j \in I_0^{(i)}}} \mathcal{U}_A^{(2)}\left(\mathcal{X}_2(i,j)\right) \leq \frac{\varepsilon^2}{10^{20}},$$

then $\mathcal{U}_A^{(0,1,2)}(\mathcal{X}_0 \cup \mathcal{X}_1) < \varepsilon$.

An alternative way to look at $\left\{ I_0^{(i)} \right\}$ would be to say that $I_0^{(i)}$ is the set of neighbors of i in an undirected graph without self-loops over n^2 vertices, of degree at least $n^2/2$ each. Call that graph Γ, then the requirement of the lemma is that a randomly chosen pair of vertices $\{i,j\} = \mathbf{X} \cap \mathbf{Y}$ when $(\mathbf{X}, \mathbf{Y}) \sim \mathcal{U}_A^{(2)}$ is unlikely to be connected in Γ. If the condition is met (i.e., such Γ can be found), then A cannot be large.

Proof of Lemma 5.4. We will show that $\mathcal{U}_A^{(0,1,2)}(\mathcal{X}_1) < \frac{\varepsilon}{2883}$ and $\mathcal{U}_A^{(0,1,2)}(\mathcal{X}_0) < \frac{2881}{2883}\varepsilon$.

Define $A_i \stackrel{\text{def}}{=} \left\{ (x,y) \in A \middle| i \in x \cap y \right\}$ for each $i \in [n^2]$. Let D be the probability distribution over $[n^2]$ defined by $D(i) = \frac{1}{2}\mathcal{U}_A^{(2)}(A_i)$. Choosing $(\mathbf{X}, \mathbf{Y}) \sim \mathcal{U}_A^{(2)}$ can be viewed as first choosing $i \sim D$, followed by $(\mathbf{X}, \mathbf{Y}) \sim \mathcal{U}_{A_i}^{(2)}$, and our assumptions guarantee that

$$\mathop{\mathbf{E}}_{i \sim D} \left[\sum_{j \in I_0^{(i)}} \mathcal{U}_{A_i}^{(2)}(\mathcal{X}_2(i,j)) \right] \le \frac{\varepsilon^2}{10^{20}}.$$

Let

$$I_1 \stackrel{\text{def}}{=} \left\{ i \in [n^2] \middle| \mathcal{U}_{1 \times 1}^{(n;2)}(A_i) < \frac{\varepsilon}{10^7 \cdot n^2} \cdot \mathcal{U}_{1 \times 1}^{(n;2)}(A) \right\},$$

$$I_2 \stackrel{\text{def}}{=} \left\{ i \in [n^2] \middle| \sum_{j \in I_0^{(i)}} \mathcal{U}_{A_i}^{(2)}(\mathcal{X}_2(i,j)) > \frac{\varepsilon}{10^{12}} \right\}.$$

Then

$$\sum_{i \in I_1} \mathcal{U}_{1 \times 1}^{(n;2)}(A_i) < \frac{\varepsilon}{10^7} \cdot \mathcal{U}_{1 \times 1}^{(n;2)}(A) \Rightarrow \sum_{i \in I_1} \mathcal{U}_A^{(2)}(A_i) < \frac{\varepsilon}{10^7},$$

and by Markov inequality,

$$D(I_2) < \frac{\varepsilon}{10^8} \Rightarrow \sum_{i \in I_2} \mathcal{U}_A^{(2)}(A_i) < \frac{2\varepsilon}{10^8}.$$

That is,

$$\sum_{i \in I_1 \cup I_2} \mathcal{U}_A^{(2)}(A_i) < \frac{1.2\,\varepsilon}{10^7}. \tag{8}$$

For any $i_0 \in [n^2]$, we view A_{i_0} as an input rectangle for $P_{1 \times 1}^\Sigma$, defined over $[n^2] \setminus \{i_0\}$.[3] Assume $i_0 \in [n^2] \setminus I_1 \setminus I_2$, then it holds that $\mathcal{U}_{1 \times 1}^{(n;2)}(A_{j_0}) \ge \frac{\varepsilon}{10^7 \cdot n^2} \cdot \mathcal{U}_{1 \times 1}^{(n;2)}(A) \in 2^{-o(n)}$. The properties of $I_0^{(i_0)}$ and the fact that $i_0 \notin I_2$ allow us to apply Lemma 5.3, concluding that

$$\mathcal{U}_{A_{i_0}}^{(1,2)}(\mathcal{X}_1) < \frac{\varepsilon}{8 \cdot 10^5}. \tag{9}$$

---

[3] Strictly speaking, this violates our requirement that n is a power of 2 and slightly affects the Hamming weights of x and y as functions of n. However, the former is irrelevant in this context and the influence of the latter is negligible for sufficiently large n, so we allow this abuse to keep the notation simple. Note also that we keep counting the size of $x \cap y$ when using the $\mathcal{X}_j$-notation according to the original definition of the communication task, i.e., w.r.t. $P_{1 \times 1}^\Sigma$ defined over $[n^2]$.

For every $i_0 \in I_1 \cup I_2$ we, on the other hand, apply Lemma 3.4 to A_{i_0} (replacing $\mathcal{X}_0$ by $\mathcal{X}_1$ and $\mathcal{X}_1$ by $\mathcal{X}_2$, due to the guaranteed $i_0 \in (x, y)$ for every $(x, y) \in A_{i_0}$). Then for $\delta = \frac{1}{2880}$ (we use $\alpha_1 = 1/2$ and $\alpha_2 = 1$, in accordance with the definition of $P_{1 \times 1}^{\Sigma}$),

$$\sum_{i \in I_1 \cup I_2} \mathcal{U}_{1 \times 1}^{(n)}(A_i \cap \mathcal{X}_1) \leq \frac{1}{\delta} \cdot \sum_{i \in I_1 \cup I_2} \mathcal{U}_{1 \times 1}^{(n)}(A_i \cap \mathcal{X}_2) + n^2 \cdot 2^{-\Omega(n)}. \tag{10}$$

Clearly,

$$\mathcal{U}_{1 \times 1}^{(n)}\left((\mathcal{X}_1 \cup \mathcal{X}_2) \cap A\right) \geq \mathcal{U}_{1 \times 1}^{(n)}(\mathcal{X}_2 \cap A) = \mathcal{U}_{1 \times 1}^{(n)}(\mathcal{X}_2) \cdot \mathcal{U}_{1 \times 1}^{(n;2)}(A) \in 2^{-o(n)}, \tag{11}$$

and dividing (10) by $\mathcal{U}_{1 \times 1}^{(n)}\left((\mathcal{X}_1 \cup \mathcal{X}_2) \cap A\right)$ gives

$$\begin{aligned}
\sum_{i \in I_1 \cup I_2} \mathcal{U}_A^{(1,2)}(A_i \cap \mathcal{X}_1) &\leq \frac{1}{\delta} \cdot \sum_{i \in I_1 \cup I_2} \mathcal{U}_A^{(1,2)}(A_i \cap \mathcal{X}_2) + 2^{-\Omega(n)} \\
&\leq \frac{1}{\delta} \cdot \sum_{i \in I_1 \cup I_2} \mathcal{U}_A^{(2)}(A_i) + 2^{-\Omega(n)} \leq \frac{1.2\,\varepsilon}{\delta \cdot 10^7} + 2^{-\Omega(n)},
\end{aligned} \tag{12}$$

as follows from (8).

We conclude that for sufficiently large n,

$$\begin{aligned}
\mathcal{U}_A^{(0,1,2)}(\mathcal{X}_1) \leq \mathcal{U}_A^{(1,2)}(\mathcal{X}_1) &= \sum_{i \in [n^2]} \mathcal{U}_A^{(1,2)}(A_i \cap \mathcal{X}_1) \\
&= \sum_{i \in I_1 \cup I_2} \mathcal{U}_A^{(1,2)}(A_i \cap \mathcal{X}_1) + \sum_{i \notin I_1 \cup I_2} \mathcal{U}_A^{(1,2)}(A_i) \cdot \mathcal{U}_{A_{i_0}}^{(1,2)}(\mathcal{X}_1) \\
&< \frac{1.2\,\varepsilon}{\delta \cdot 10^7} + 2^{-\Omega(n)} + \frac{\varepsilon}{8 \cdot 10^5} < \frac{\varepsilon}{2883},
\end{aligned}$$

as follows from (9) and (12).

We apply Lemma 3.4 one more time. For the same value of δ it holds that

$$\mathcal{U}_{1 \times 1}^{(n)}(A \cap \mathcal{X}_0) \leq \frac{1}{\delta} \cdot \mathcal{U}_{1 \times 1}^{(n)}(A \cap \mathcal{X}_1) + 2^{-\Omega(n)}.$$

Like in (11),

$$\mathcal{U}_{1 \times 1}^{(n)}\left((\mathcal{X}_0 \cup \mathcal{X}_1 \cup \mathcal{X}_2) \cap A\right) \geq \mathcal{U}_{1 \times 1}^{(n)}(\mathcal{X}_2 \cap A) \in 2^{-o(n)},$$

and therefore for sufficiently large n,

$$\mathcal{U}_A^{(0,1,2)}(\mathcal{X}_0) \leq \frac{1}{\delta} \cdot \mathcal{U}_A^{(0,1,2)}(\mathcal{X}_1) + 2^{-\Omega(n)} < \frac{2880}{2883}\varepsilon + 2^{-\Omega(n)} < \frac{2881}{2883}\varepsilon,$$

as required. ∎ *Lemma 5.4*

In order to state our next lemma we need the following definition.

Definition 7. We call a rectangle A *δ-labeled* if

$$\Pr_{\mathbf{Y} \sim \mathcal{U}_A^{(Bob)}}\left[\exists\, a, b \in \mathbf{Y},\, a \neq b : \Pr_{\mathbf{X} \sim \mathcal{U}_A^{(Alice)}}[\{a, b\} \subset \mathbf{X}] \geq \delta\right] > \frac{1}{3}.$$

15

The following lemma (which is our final step towards Theorem 5.2) claims, informally, that if a big rectangle A accepts instances of $\mathcal{X}_2$ while rejecting those of $\mathcal{X}_0$ and $\mathcal{X}_1$, then there can be only limited uncertainty regarding the content of $\mathbf{X} \cap \mathbf{Y}$, for a randomly chosen pair $(\mathbf{X}, \mathbf{Y}) \in A$.

Lemma 5.5. *Let n be sufficiently large and A be an input rectangle for $P^{\Sigma}_{1 \times 1}$, such that* $\mathbf{Pr}_{\mathcal{U}_A}[|\mathbf{X} \cap \mathbf{Y}| < 2] \leq \frac{1}{6}$ *and A is not δ-labeled for some $\delta > 0$. Then $\mathcal{U}^{(n)}_{1 \times 1}(A) \in 2^{-\Omega\left(\frac{1}{\sqrt{\delta}}\right)}$.*

Proof of Lemma 5.5. Let $(\mathbf{X}, \mathbf{Y})$ be the input variables. Let B be the set of $y \subset [n^2]$, such that

$$\mathbf{Pr}_{\mathbf{X} \sim \mathcal{U}^{(Alice)}_A} [|\mathbf{X} \cap y| < 2] \leq \frac{1}{3} \tag{13}$$

and

$$\forall a, b \in y,\ a \neq b: \quad \mathbf{Pr}_{\mathbf{X} \sim \mathcal{U}^{(Alice)}_A} [\{a, b\} \subset \mathbf{X}] < \delta. \tag{14}$$

If we choose $y = \mathbf{Y} \sim \mathcal{U}^{(Bob)}_A$ then (13) holds with probability at least $\frac{1}{2}$ and (14) holds with probability at least $\frac{2}{3}$ (A is not δ-labeled); therefore, $\mathcal{U}^{(n)}_{1 \times 1}(A') \geq \frac{1}{6}\mathcal{U}^{(n)}_{1 \times 1}(A)$ for $A' \overset{\text{def}}{=} A\|_{Alice} \times B$.

For $a, b \in [n^2]$, $a \neq b$, let $p_a \overset{\text{def}}{=} \mathbf{Pr}_{\mathbf{X} \sim \mathcal{U}^{(Alice)}_A} [a \in \mathbf{X}]$ and $p^{(a)}_b \overset{\text{def}}{=} \mathbf{Pr}_{\mathbf{X} \sim \mathcal{U}^{(Alice)}_A} [b \in \mathbf{X} | a \in \mathbf{X}]$. Condition (14) implies that

$$\forall a \in y: \left(p_a \geq \sqrt{\delta} \Rightarrow \forall b \in y \setminus \{a\}: p^{(a)}_b < \sqrt{\delta} \right). \tag{15}$$

We will see that both (13) and (15) are not likely to hold simultaneously for a random $y = \mathbf{Y} \sim \mathcal{U}_{Bob}$. Let $a_0(y)$ denote the lexicographically first value $i_0 \in y$, satisfying $p_{i_0} = \max_{i \in y} \{p_i\}$.

First let us consider the situation when $p_{a_0(\mathbf{Y})} < \sqrt{\delta}$, i.e.,

$$\forall a \in y: p_a < \sqrt{\delta}. \tag{16}$$

Since (13) implies $\sum_{a \in \mathbf{Y}} p_a \geq \frac{4}{3}$, the probability of both (13) and (16) holding simultaneously w.r.t. $y = \mathbf{Y}$ is upper bounded by

$$\mathbf{Pr}\left[\sum_{a \in \mathbf{Y}} p'_a \geq \frac{4}{3} \right], \tag{17}$$

where $p'_a \overset{\text{def}}{=} \begin{cases} p_a & \text{if } p_a < \sqrt{\delta} \\ 0 & \text{otherwise} \end{cases}$.

Let $Z_1, \ldots, Z_n$ be the elements of $\mathbf{Y}$ and denote $W_i \overset{\text{def}}{=} p'_{Z_i}$. We want to use Chernoff bound in order to limit from above the value of $\sum_{i=1}^n W_i$. Though the variables W_i are not independent (because all Z_i-s must be different), it is possible to apply Chernoff bound using the "worst case" estimation of the variables' conditional mean values. Formally, in order to obtain the lower bound we analyze a relaxation of the original experiment, where all W_i-s are independent but distributed according to the worst scenario, resulting from conditioning

16

upon the values of $\{W_j | j \neq i\}$. Note that for each $1 \leq i \leq n$ it holds that $W_i \leq \sqrt{\delta}$ and $\mathbf{E}[W_i] \leq \frac{|x|}{|[n^2]| - |y|} = \frac{n/2}{n^2 - n} < \frac{3}{5n}$, even if the expectation is conditioned upon some values of $\{W_j | j \neq i\}$. Based on Chernoff bound (Claim 3.1), we conclude that

$$\Pr_{\mathbf{Y} \sim \mathcal{U}_{Bob}} \left[\sum_{a \in \mathbf{Y}} p'_a \geq \frac{4}{3} \right] \in 2^{-\Omega\left(\frac{1}{\sqrt{\delta}}\right)}. \tag{18}$$

Now consider the other choice left by (15), namely

$$p_{a_0} \geq \sqrt{\delta} \quad \text{and} \quad \forall b \in \mathbf{Y}, b \neq a_0 : p_b^{(a_0)} < \sqrt{\delta}, \tag{19}$$

where a_0 stands for $a_0(\mathbf{Y})$. Let $\mathbf{Y} = y$; since $\Pr_{\mathcal{U}_A^{(Alice)}}[|\mathbf{X} \cap y| \geq 2] \geq \frac{2}{3}$ implies

$$\sum_{b \in y \setminus \{a_0\}} p_b^{(a_0)} \geq \frac{2}{3},$$

the probability that (13) and (19) hold is upper bounded by the probability that

$$\sum_{b \in y \setminus \{a_0\}} p_b^{(a_0)'} \geq \frac{2}{3}, \tag{20}$$

where $p_b^{(a_0)'} \stackrel{\text{def}}{=} \begin{cases} p_b^{(a_0)} & \text{if } p_b^{(a_0)} < \sqrt{\delta} \\ 0 & \text{otherwise} \end{cases}$.

Like in the case of (17), Chernoff bound (Claim 3.1) implies that (20) holds with probability $2^{-\Omega\left(\frac{1}{\sqrt{\delta}}\right)}$. Therefore,

$$\mathcal{U}_{1\times1}^{(n)}(A) \leq 6 \cdot \mathcal{U}_{1\times1}^{(n)}(A') \leq 6 \cdot \Pr_{\mathbf{Y} \sim \mathcal{U}_{Bob}}[\mathbf{Y} \in B] \in 2^{-\Omega\left(\frac{1}{\sqrt{\delta}}\right)},$$

as required. ∎ *Lemma 5.5*

Recall

Theorem 5.2. *Assume that there exists a deterministic protocol of cost $k \in o(n) \cap \omega(1)$ that solves $P_{1\times1}^{\Sigma}$ for some Σ w.r.t. $\mathcal{U}_{1\times1}^{(n;2)}$ with probability $\gamma \in \omega(2^{-k})$ and error bounded by 10^{-22}. Then $P_{1\times1}^{search}$ can be solved w.r.t. $\mathcal{U}_{1\times1}^{(n;2)}$ with probability $\frac{\gamma}{k^2 \cdot \log^2(n/\gamma)}$ with error 0 by a public coin protocol of cost $O\left(k + \log^2(n/\gamma)\right)$.*

Now we have all that is required to prove it. In the proof we, essentially, argue that a protocol solving $P_{1\times1}^{\Sigma}$ w.r.t. $\mathcal{U}_{1\times1}^{(n;2)}$ must give rise to "typical" rectangles satisfying the requirements of Lemma 5.4, which lets us apply the contrapositive of Lemma 5.5 and conclude that typical rectangles are δ-labeled. From the definition of δ-labeled, the pair $\{a, b\}$ chosen w.r.t. $\mathbf{Y}$ has good chances to equal $\mathbf{X} \cap \mathbf{Y}$, if the input belongs to that rectangle. The last observation leads to a protocol for solving $P_{1\times1}^{search}$.

Proof of Theorem 5.2. Let S be a deterministic protocol of cost k solving $P^{\Sigma}_{1\times 1}$ for some Σ w.r.t. $\mathcal{U}^{(n;2)}_{1\times 1}$ with probability γ and error bounded by 10^{-22}.

Let A be an input rectangle; observe that Lemma 5.5 guarantees that if $\mathcal{U}_A(\mathcal{X}_0 \cup \mathcal{X}_1) \leq \frac{1}{6}$ and $\mathcal{U}^{(n)}_{1\times 1}(A) \geq 2^{-\Omega(k)}$ then there exists a function $\delta(k) \in \Omega\left(\frac{1}{k^2}\right)$, such that A is $\delta(k)$-labeled. Fix such $\delta(k)$ for the rest of the proof.

Consider the rectangles defined by S. We will call a rectangle A *latent* if it is not possible to define an answer that would solve $P^{\Sigma}_{1\times 1}$ with probability at least $1 - \frac{2}{10^{22}}$ w.r.t. $\mathcal{U}^{(2)}_A$. As follows from the accuracy of S, $(\mathbf{X}, \mathbf{Y}) \sim \mathcal{U}^{(n;2)}_{1\times 1}$ does not belong to a latent rectangle with probability at least $\frac{\gamma}{2}$ (at least half of all pairs $(x,y) \in \mathcal{X}_2$ for which S produces an answer belong to non-latent rectangles, since otherwise the error of S would be greater than 10^{-22}). On the other hand, with probability at least $1 - \frac{\gamma}{4}$ it happens that $(\mathbf{X}, \mathbf{Y}) \sim \mathcal{U}^{(n;2)}_{1\times 1}$ falls into a rectangle A satisfying $\mathcal{U}^{(n;2)}_{1\times 1}(A) \geq \frac{\gamma}{2^{k+2}}$. Note that for any such A it holds that $\mathcal{U}^{(n;2)}_{1\times 1}(A) \geq 2^{-2k}$ and $\mathcal{U}^{(n)}_{1\times 1}(A) \geq \mathcal{U}^{(n)}_{1\times 1}(\mathcal{X}_2) \cdot \frac{\gamma}{2^{k+2}} \geq 2^{-2k}$ (if n is large enough).

Call a rectangle A *good* if it is not latent, $\mathcal{U}^{(n;2)}_{1\times 1}(A) \geq 2^{-2k}$ and $\mathcal{U}^{(n)}_{1\times 1}(A) \geq 2^{-2k}$. As shown above, $(\mathbf{X}, \mathbf{Y}) \sim \mathcal{U}^{(n;2)}_{1\times 1}$ belongs to a good rectangle with probability at least $\frac{\gamma}{2} - \frac{\gamma}{4} = \frac{\gamma}{4}$. Consequently, $(\mathbf{X}, \mathbf{Y}) \sim \mathcal{U}^{(n;\geq 2)}_{1\times 1}$ falls into a good rectangle with probability at least $\mathcal{U}^{(n)}_{1\times 1}(\mathcal{X}_2) \cdot \frac{\gamma}{4} \geq \frac{\gamma}{52}$.

We claim that any good A is $\delta(k)$-labeled. This follows from the fact that there exists some $z_A \in [n^2] \setminus \{0\}$, such that

$$1 - \frac{2}{10^{22}} \leq \Pr_{\mathcal{U}^{(2)}_A}\left[(\mathbf{X}, \mathbf{Y}, z_A) \in P^{\Sigma}_{1\times 1}\right] = \Pr_{\mathcal{U}^{(2)}_A}\left[\langle z_A, \sigma_{n^2}(a) + \sigma_{n^2}(b)\rangle = 0\right],$$

where $\mathbf{X} \cap \mathbf{Y} = \{a, b\}$ and $\sigma_{n^2} \in \Sigma$. If we define $I^{(a)}_0 \overset{\text{def}}{=} \left\{b \in [n^2] \,\middle|\, \langle z_A, \sigma_{n^2}(a) + \sigma_{n^2}(b)\rangle = 1\right\}$ that will, w.r.t. A, satisfy the requirement of Lemma 5.4 for $\varepsilon = \frac{1}{6}$. Therefore it holds that $\mathcal{U}_A(\mathcal{X}_0 \cup \mathcal{X}_1) \leq \mathcal{U}^{(0,1,2)}_A(\mathcal{X}_0 \cup \mathcal{X}_1) < \frac{1}{6}$. As $\mathcal{U}^{(n)}_{1\times 1}(A) \geq 2^{-2k}$, we can apply the contrapositive of Lemma 5.5, as suggested in the beginning of the proof, which leads to the conclusion that A is $\delta(k)$-labeled.

We are ready to construct a protocol, as promised by the statement we are proving. The idea is to first map the input $(x, y) \in \mathcal{X}_2$ to $(x', y') = (\mathbf{X}', \mathbf{Y}') \sim \mathcal{U}^{(n;\geq 2)}_{1\times 1}$, then to feed (x', y') to the original protocol S, hoping that the pair will fall into a $\delta(k)$-labeled rectangle. If that occurs, we have a good candidate for the correct answer, as guaranteed by the fact that A is $\delta(k)$-labeled. Validity of the guess is easy to verify by a 2-way protocol.

Let D be the distribution over $[n]$ defined by $D(j) \overset{\text{def}}{=} \mathcal{U}^{(n;\geq 2)}_{1\times 1}(\mathcal{X}_j)$. Consider the following protocol S'.

1. Alice and Bob use public randomness to choose $j_0 \sim D$. If $j_0 > 3\log\left(\frac{312}{\gamma \cdot \delta(k)}\right)$ then the protocol stops and returns no answer. Otherwise Alice sends to Bob j_0 lexicographically first elements from x, denoted by $(x_1, \ldots, x_{j_0})$.

2. Bob sends to Alice any two indices i_1 and i_2, such that $I_x \overset{\text{def}}{=} \{x_i\}^{j_0}_{i=1} \setminus \{x_{i_1}, x_{i_2}\}$ and y are disjoint, followed by j_0 lexicographically first elements from y, denoted by $(y_1, \ldots, y_{j_0})$.

3. Let i_3 and i_4 be any two indices, such that $I_y \overset{\text{def}}{=} \{y_i\}_{i=1}^{j_0} \setminus \{y_{i_3}, y_{i_4}\}$ and x are disjoint, denote $\tilde{x} \overset{\text{def}}{=} (x \cup I_y) \setminus I_x$.

4. Alice and Bob use public randomness to choose a random permutation ρ over the elements of $[n^2]$.

5. Alice and Bob run the protocol S on the input $(\rho(\tilde{x}), \rho(y))$. Let A be the rectangle defined by S, where $(\rho(\tilde{x}), \rho(y))$ belongs. If there exists no pair $a, b \in y$, $a \neq b$, such that $\mathbf{Pr}\left[\{a, b\} \subset \mathbf{X}\right] \geq \delta(k)$ when $\mathbf{X} \sim \mathcal{U}_A^{(Alice)}$, then the protocol stops and returns no answer; otherwise let (a', b') be any such pair.

6. If $\left\{\rho^{-1}(a'), \rho^{-1}(b')\right\} \subseteq x \cap y$ then the protocol outputs these two elements. Otherwise the protocol returns no answer.

It is clear that the protocol is 0-error and its communication cost is $O\left(k + j_0 \cdot \log n\right) \subseteq O\left(k + \log^2(n/\gamma)\right)$. Let us calculate the probability that an answer is produced.

Consider an "idealized" protocol S'', similar to S' but having no halting condition in stage 1 (i.e., S'' continues to run regardless of the value of j_0). Define the following events characterizing the behavior of S'':

- $\mathcal{E}_1$ is the event that $(\rho(\tilde{x}), \rho(y))$ belongs to a $\delta(k)$-labeled rectangle A.

- $\mathcal{E}_2$ is the event that at step 5 a pair (a', b') has been chosen.

- $\mathcal{E}_3$ is the event that $\mathcal{E}_2$ occurs and $\{a', b'\} \subset x$.

- $\mathcal{E}_4$ is the event that $\mathcal{E}_3$ occurs and $j_0 \leq 3\log\left(\frac{312}{\gamma \cdot \delta(k)}\right)$.

- $\mathcal{E}_5$ is the event that $\mathcal{E}_4$ occurs and $\left\{\rho^{-1}(a'), \rho^{-1}(b')\right\} = x \cap y$.

Clearly, the probability that S' is successful is equal to the probability that $\mathcal{E}_5$ occurs.

Note that since ρ is a uniformly random permutation and $j_0 \sim D$, it holds that $(\rho(\tilde{x}), \rho(y)) \sim \mathcal{U}_{1 \times 1}^{(n; \geq 2)}$, and so $\mathbf{Pr}\left[\mathcal{E}_1\right] \geq \frac{\gamma}{52}$. By the definition of a $\delta(k)$-labeled rectangle, $\mathbf{Pr}\left[\mathcal{E}_2 \middle| \mathcal{E}_1\right] \geq \frac{1}{3}$ and $\mathbf{Pr}\left[\mathcal{E}_3 \middle| \mathcal{E}_2\right] \geq \delta(k)$, so $\mathbf{Pr}\left[\mathcal{E}_3\right] \geq \frac{\gamma \cdot \delta(k)}{156}$.

Event $\mathcal{E}_4$ occurs if $\mathcal{E}_3$ occurs and $j_0 \leq 3\log\left(\frac{312}{\gamma \cdot \delta(k)}\right)$, therefore

$$\mathbf{Pr}\left[\mathcal{E}_4\right] \geq \frac{\gamma \cdot \delta(k)}{156} - \underset{D}{\mathbf{Pr}}\left[j_0 > 3\log\left(\frac{312}{\gamma \cdot \delta(k)}\right)\right] \geq \frac{\gamma \cdot \delta(k)}{312},$$

where the second inequality follows from Claim 3.2.

Finally, $\mathcal{E}_5$ occurs if $\mathcal{E}_4$ occurs and the points $\rho^{-1}(a')$ and $\rho^{-1}(b')$ belong to $x \cap y$. Given j_0, our mapping of (x, y) to $(\rho(\tilde{x}), \rho(y))$ produces a random instance drawn from $\mathcal{U}_{1 \times 1}^{(n; \geq 2)}(\mathcal{X}_j)$. Moreover, the two elements of $x \cap y$ are mapped to a uniformly random pair inside $\rho(\tilde{x}) \cap \rho(y)$, even if we condition upon $\mathcal{E}_4$ (the pair $(\rho(\tilde{x}), \rho(y))$ is the input to S at step 5, and it reveals no additional information about $\rho(x \cap y)$ inside $\rho(\tilde{x} \cap y)$). The conditional probability that $\left\{\rho^{-1}(a'), \rho^{-1}(b')\right\} = x \cap y$ is equal to $1 \middle/ \binom{j_0}{2} \geq \frac{1}{j_0^2}$, and

$$\mathbf{Pr}\left[\mathcal{E}_5\right] = \mathbf{Pr}\left[\mathcal{E}_4\right] \cdot \mathbf{Pr}\left[\mathcal{E}_5 \middle| \mathcal{E}_4\right] \geq \frac{\gamma \cdot \delta(k)}{312 \cdot j_0^2} \in \Omega\left(\frac{\gamma}{k^2 \cdot \log^2(n/\gamma)}\right),$$

as follows from $\delta(k) \in \Omega\left(1/k^2\right)$.

The protocol S' is 0-error, so we can repeat it several times in order to get an answer with probability at least $\frac{\gamma}{k^2 \cdot \log^2(n/\gamma)}$. ∎ *Theorem 5.2*

5.3 Solving $P_{1 \times 1}^{search}$ is expensive

It is not hard to see that a protocol of communication cost k can solve $P_{1 \times 1}^{search}$ w.r.t. $\mathcal{U}_{1 \times 1}^{(n;1)}$ only with probability $O\left(\frac{k}{n}\right)$. We prove the following generalization of this statement.

Theorem 5.6. *Let $t \in o\left(\sqrt{n}\right)$, then any 0-error public coin protocol of cost $k \in \Omega\left(t \log n\right)$ solving $P_{1 \times 1}^{search}$ w.r.t. $\mathcal{U}_{1 \times 1}^{(n;t)}$ can succeed with probability $\left(O\left(\frac{kt}{n}\right)\right)^t$.*

This is a direct product theorem, because its statement can be rephrased as one about solving t independent instances of $P_{1 \times 1}^{search}$ w.r.t. $\mathcal{U}_{1 \times 1}^{(n;1)}$. There are known direct product results that apply to problems like $DISJ$ and $P_{1 \times 1}^{search}$ (e.g., [JKN08], and references therein). However, two obstacles prevent us from using them: on the one hand, those results mostly apply to the case of product input distributions; on the other hand, their statements have "$\Omega(t)$" in the exponent of the guaranteed upper bound on success probability, and that is not sufficient to us. In Section 6 we pose some related open questions.

In this paper we will only make use of the case corresponding to $t = 2$, though we prove the theorem in full generality, as it might be of independent interest.

Proof of Theorem 5.6. Let S be a 0-error protocol of cost k solving $P_{1 \times 1}^{search}$ w.r.t. $\mathcal{U}_{1 \times 1}^{(n;t)}$ with probability $p_t^{(t)}$. For $i > t$, let $p_i^{(t)}$ be the probability that S outputs t elements from $x \cap y$ when $(x, y) = (\mathbf{X}, \mathbf{Y}) \sim \mathcal{U}_{1 \times 1}^{(n;i)}$.

Proposition. *There exists an absolute constant c such that for $t \leq i \leq \frac{n}{2}$ it holds that*

$$p_i^{(t)} \leq \max\left\{\left(\frac{k}{n}\right)^t, \left(1 + \frac{ck}{n}\right) \cdot \left(1 - \frac{t}{i+1}\right) \cdot p_{i+1}^{(t)}\right\}.$$

The proposition implies the theorem, as follows. Let n be such that $t + \left\lfloor\frac{n}{3ck}\right\rfloor < \frac{n}{2}$. If for any $i \in \left\{t, \ldots, t + \left\lfloor\frac{n}{3ck}\right\rfloor\right\}$ it holds that $p_i^{(t)} \leq \left(\frac{k}{n}\right)^t$, let i_0 be the smallest value like this, then

$$p_t^{(t)} \leq \left(1 + \frac{ck}{n}\right)^{i_0 - t} p_{i_0}^{(t)} \in O\left(\left(\frac{k}{n}\right)^t\right).$$

Otherwise,

$$p_t^{(t)} \leq \left(1 + \frac{ck}{n}\right)^{\left\lfloor\frac{n}{3ck}\right\rfloor} \cdot \prod_{i=t}^{t+\left\lfloor\frac{n}{3ck}\right\rfloor-1} \frac{i+1-t}{i+1} \cdot p_{t+\left\lfloor\frac{n}{3ck}\right\rfloor}^{(t)} \leq 2 \frac{\prod_{i=1}^{t} i}{\prod_{j=\left\lfloor\frac{n}{3ck}\right\rfloor+1}^{t+\left\lfloor\frac{n}{3ck}\right\rfloor} j} \in O\left(\frac{kt}{n}\right)^t.$$

Let us prove the proposition. Let $i_0 \in \left\{t, \ldots, \frac{n}{2}\right\}$ be such that $p_{i_0}^{(t)} > \left(1 - \frac{t}{i_0+1}\right) p_{i_0+1}^{(t)}$ and $p_{i_0}^{(t)} > \left(\frac{k}{n}\right)^t$, our goal is to show that $p_{i_0}^{(t)} \leq \left(1 + \frac{ck}{n}\right)\left(1 - \frac{t}{i+1}\right) p_{i+1}^{(t)}$ for some fixed c.

Let $m \stackrel{\text{def}}{=} n^2 - i_0$, define D as the uniform distribution over pairs (x', y') such that $x' \subset [m]$, $|x'| = n/2 - i_0$, $y' \subset [m]$ and $|y'| = n - i_0$. Assume we know that $(x', y') \in \text{supp}(D)$ belongs to either $\mathcal{X}_0$ or $\mathcal{X}_1$, and want to distinguish the two cases. Consider the following public coin protocol S', running on (x', y').

1. Let $x_0' \stackrel{\text{def}}{=} x' \cup \{j\}_{j=m+1}^{n^2}$ and $y_0' \stackrel{\text{def}}{=} y' \cup \{j\}_{j=m+1}^{n^2}$. Alice and Bob use public randomness to choose a random permutation ρ over the elements of $[n^2]$.

2. Alice and Bob run the protocol S on the input $(\rho(x_0'), \rho(y_0'))$. If S does not outputs t elements then S' refuses to answer. Otherwise if the t produced elements belong to $\rho(\{j | m < j \leq n^2\})$ then S' outputs $\mathbf{0}$, else S' refuses to answer.

If $(x', y') \in \mathcal{X}_0$ then the pair $(\rho(x_0'), \rho(y_0'))$ is distributed according to $\mathcal{U}_{1 \times 1}^{(n;i_0)}$ and S' outputs $\mathbf{0}$ with probability $p_{i_0}^{(t)}$. If $(x', y') \in \mathcal{X}_1$ then the pair $(\rho(x_0'), \rho(y_0'))$ is distributed according to $\mathcal{U}_{1 \times 1}^{(n;i_0+1)}$ and S' outputs $\mathbf{0}$ with probability $p_{i_0+1}^{(t)} \cdot \binom{i_0}{t} / \binom{i_0+1}{t} = \left(1 - \frac{t}{i_0+1}\right) \cdot p_{i_0+1}^{(t)}$. We know that the former probability is higher than the latter, and so if S' outputs $\mathbf{0}$ that can be viewed as an argument towards $(x', y') \in \mathcal{X}_0$.

Note that $D(\mathcal{X}_0) \geq \frac{1}{3}$ (by analogy to Claim 3.2), and we can apply Lemma 3.4 with $\alpha_1 = \frac{1}{4}$ and $\alpha_2 = 1$. The lemma implies that for $\delta = \frac{1}{11520}$, some absolute constant c_0 and any rectangle A it holds that

$$D(A \cap \mathcal{X}_1) \geq \delta \cdot D(A \cap \mathcal{X}_0) - 2^{-c_0 \cdot n}. \tag{21}$$

Let $l \in \mathbb{N}$ and S_l' be a protocol that runs S' as a subroutine l times (each time using independent random bit), and outputs $\mathbf{0}$ if all the instantiations of S' return $\mathbf{0}$ (otherwise S_l' refuses to answer). Denote by $\mathcal{E}_0$ the event that S_l' outputs $\mathbf{0}$. If $(x', y') \in \mathcal{X}_0$ then $\mathcal{E}_0$ occurs with probability $\left(p_{i_0}^{(t)}\right)^l$, if $(x', y') \in \mathcal{X}_1$ then $\mathcal{E}_0$ occurs with probability $\left(\left(1 - \frac{t}{i_0+1}\right) \cdot p_{i_0+1}^{(t)}\right)^l$. Therefore, w.r.t. uniformly random bits used by S_l', we expect that

$$\mathbf{Pr}_D[\mathcal{X}_0 \text{ and } \mathcal{E}_0] \geq \frac{1}{3} \cdot \left(p_{i_0}^{(t)}\right)^l$$

and

$$\mathbf{Pr}_D[\mathcal{X}_1 \text{ and } \mathcal{E}_0] \leq \left(\left(1 - \frac{t}{i_0+1}\right) \cdot p_{i_0+1}^{(t)}\right)^l.$$

Assume that S_l' uses s random bits, for any $r \in \{0,1\}^s$ let $S_l'(r)$ be the deterministic protocol obtained from S_l' by using r instead of the random bits. Because $S_l'(r)$ is a protocol of communication cost kl, it partitions the domain into rectangles $A_1^{(r)}, \ldots, A_{2^{kl}}^{(r)}$. Let B consist of all $A_i^{(r)}$-s for $r \in \{0,1\}^s$, on which the corresponding $S_l'(r)$ outputs $\mathbf{0}$. We denote

$$\beta(l) \stackrel{\text{def}}{=} \frac{1}{3} \cdot \left(\frac{p_{i_0}^{(t)}}{\left(1 - \frac{t}{i_0+1}\right) \cdot p_{i_0+1}^{(t)}}\right)^l \leq \frac{\mathbf{Pr}_D[\mathcal{X}_0 \text{ and } \mathcal{E}_0]}{\mathbf{Pr}_D[\mathcal{X}_1 \text{ and } \mathcal{E}_0]},$$

then

$$\frac{1}{2^s} \cdot \sum_{A \in B} D(A \cap \mathcal{X}_0) = \mathbf{Pr}_D[\mathcal{X}_0 \text{ and } \mathcal{E}_0] \geq \beta(l) \cdot \mathbf{Pr}_D[\mathcal{X}_1 \text{ and } \mathcal{E}_0] = \frac{\beta(l)}{2^s} \cdot \sum_{A \in B} D(A \cap \mathcal{X}_1).$$

Let $\mu \stackrel{\text{def}}{=} \mathbf{E}_{A \in B} \left[D \left(A \cap \mathcal{X}_0 \right) \right]$ and $B' \stackrel{\text{def}}{=} \left\{ A \in B \middle| D \left(A \cap \mathcal{X}_0 \right) \geq \frac{\mu}{2} \right\}$. Then

$$\sum_{A \in B'} D \left(A \cap \mathcal{X}_0 \right) \geq \frac{1}{2} \sum_{A \in B} D \left(A \cap \mathcal{X}_0 \right) \geq \frac{\beta(l)}{2} \sum_{A \in B} D \left(A \cap \mathcal{X}_1 \right) \geq \frac{\beta(l)}{2} \sum_{A \in B'} D \left(A \cap \mathcal{X}_1 \right),$$

and there exists $A_0 \in B'$ satisfying

$$\frac{2}{\beta(l)} D \left(A_0 \cap \mathcal{X}_0 \right) \geq D \left(A_0 \cap \mathcal{X}_1 \right).$$

It holds that

$$\mu \geq \frac{1}{2^{kl}} \cdot \Pr_D \left[\mathcal{X}_0 \text{ and } \mathcal{E}_0 \right] \geq \frac{\left(p_{i_0}^{(t)} \right)^l}{3 \cdot 2^{kl}} > \frac{k^{tl}}{3 \cdot 2^{kl} \cdot n^{tl}} > 2^{-kl - tl \log n - 2},$$

and $D \left(A_0 \cap \mathcal{X}_0 \right) \geq \frac{\mu}{2} > 2^{-kl - tl \log n - 3}$. So, (21) leads to

$$\frac{2}{\beta(l)} \cdot D \left(A_0 \cap \mathcal{X}_0 \right) \geq D \left(A_0 \cap \mathcal{X}_1 \right) \geq \delta \cdot D(A_0 \cap \mathcal{X}_0) - 2^{-c_0 \cdot n};$$

$$2^{-c_0 \cdot n} \geq \left(\delta - \frac{2}{\beta(l)} \right) \cdot D(A_0 \cap \mathcal{X}_0) \geq \left(\delta - \frac{2}{\beta(l)} \right) \cdot 2^{-kl - tl \log n - 3};$$

$$\delta - \frac{2}{\beta(l)} \leq 2^{l(k + t \log n) + 3 - c_0 \cdot n}.$$

Recall that $k \in \Omega \left(t \log n \right)$, so there exists an absolute constant c_1 that guarantees that the right-hand side of the last inequality is less than $\frac{\delta}{2}$, as long as $l \leq \frac{c_1 n}{k}$. Consequently,

$$\frac{4}{\delta} \geq \beta \left(\frac{c_1 n}{k} \right) = \frac{1}{3} \cdot \left(\frac{p_{i_0}^{(t)}}{\left(1 - \frac{t}{i_0 + 1} \right) \cdot p_{i_0+1}^{(t)}} \right)^{\frac{c_1 n}{k}},$$

which implies that for some absolute constant c,

$$\left(\frac{p_{i_0}^{(t)}}{\left(1 - \frac{t}{i_0 + 1} \right) \cdot p_{i_0+1}^{(t)}} \right)^{\frac{n}{k}} \leq c \Rightarrow \frac{p_{i_0}^{(t)}}{\left(1 - \frac{t}{i_0 + 1} \right) \cdot p_{i_0+1}^{(t)}} \leq 1 + \frac{ck}{n},$$

as required.   ■ *Theorem 5.6*

5.4   Lower bound on the classical 2-way communication complexity of P

Theorem 5.7. *Solving P in the classical 2-way setting with bounded error requires a protocol of cost $\Omega \left(\frac{n^{1/4}}{\sqrt{\log n}} \right)$.*

Proof of Theorem 5.7. Assume that a protocol S of communication cost $k \in o \left(n \right)$ solves P with error bounded by $\frac{1}{2 \cdot 10^{22}}$.

Then Lemma 5.1 implies that there exists a protocol S' of communication cost $O \left(k \right)$ that solves $P_{1 \times 1}^{\Sigma}$ for some Σ w.r.t. $\mathcal{U}_{1 \times 1}^{(n;2)}$ with probability $\frac{2}{n}$ and error bounded by $\frac{1}{10^{22}}$.

By Theorem 5.2 there exists a protocol S'' of communication cost $O\left(k + \log^2(n)\right)$ solving $P_{1\times 1}^{search}$ in 0-error setting w.r.t. $\mathcal{U}_{1\times 1}^{(n;2)}$ with probability $\frac{2}{nk^2\log^2(n)}$.

Choose $t = 2$, Theorem 5.6 implies that S'' can succeed only with probability $O\left(\frac{k^2+\log^4(n)}{n^2}\right)$, therefore $k \in \Omega\left(\frac{n^{1/4}}{\sqrt{\log n}}\right)$, as required. $\blacksquare$ *Theorem 5.7*

6 Conclusions and further work

The protocol described in Section 4 together with Theorem 5.7 imply Theorem 1.1.

It would be interesting to strengthen this result. Is it possible to find a *functional* problem that requires exponentially more communication in $\mathcal{R}$ than in $\mathcal{Q}^1$? Raz [R99] constructs a *partial* function which is *complete*, in a natural and well-defined sense, for quantum one-way communication. However, it is yet unclear what the classical complexity of Raz's function is.

It seems plausible that every *total* function with an efficient one-way quantum protocol admits an efficient classical protocol (maybe, even one-way). Validity of this conjecture is a very important, well-known open problem.

What can be claimed about $\mathcal{R}$-complexity of communication problems with efficient quantum simultaneous protocols, either with or without shared entanglement?

As we have mentioned before, our Theorem 5.6 is a direct product statement and can be compared to other known direct product theorems, like that by Jain, Klauck and Nayak [JKN08] and earlier ones. On the one hand, our statement is more rigorous, in the sense that it has plain "t" in the exponent of the guaranteed upper bound on the success probability of solving t instances of the original problem, as opposed to "$\Omega(t)$" in the earlier works. On the other hand, our theorem applies (or trivially generalizes) only to a restricted family of communication problems (those with structure similar to $DISJ$), as opposed to the result of [JKN08], which speaks about *any* communication problem.

Apparently, our technique can be applied to a wider class of communication problems, and, on the other hand, the approach taken in [JKN08] can give a more rigorous statement in terms of t. It would be interesting to analyze these two possibilities in order to give a more unified theory of direct product statements in communication complexity.

Acknowledgments

References

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

[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.

[BFS86]   L. Babai, P. Frankl and J. Simon. Complexity classes in communication complexity theory. *Proceedings of the 27th Annual Symposium on Foundations of Computer Science, pages 337-347*, 1986.

[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.

[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.

[JKN08]   R. Jain, H. Klauck and A. Nayak. Direct Product Theorems for Classical Communication Complexity via Subdistribution Bounds. *Proceedings of the 40th Symposium on Theory of Computing*, 2008.

[KN97]   E. Kushilevitz and N. Nisan. Communication Complexity. *Cambridge University Press*, 1997.

[R92]   A. Razborov. On the Distributional Complexity of Disjointness. *Theoretical Computer Science 106(2), pages 385-390*, 1992.

[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.