Classical Interaction Cannot Replace a Quantum Message

Dmitry Gavinsky
David R. Cheriton School of Computer Science and Institute for Quantum Computing
University of Waterloo

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 where by using the laws of quantum mechanics one can find a solution more efficiently than on a classical computer.

In this paper we study quantum computation from the perspective of of Communication Complexity, first defined by Yao [Y79]. Here 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 the restrictions of a specific *communication model*. In order to compare the power of two communication models one has to demonstrate a communication task that can be solved more efficiently in one model than in the other (or argue that no such task exists).

We will be mostly concerned about the following models.

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

Both of these 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 R^1 and R, and the quantum versions are denoted by Q^1 and Q, respectively.

Communication tasks can be either *functional*, corresponding to the case when for each input pair (x, y) there exists at most one correct answer, or *relational*, when multiple correct answers are allowed for the same input. Input pairs without correct answers are never offered.[1]

[1] Communication tasks with *exactly one* correct answer corresponding to each possible input pair are called *total functions*.

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 information (bits or qubits) 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 ε there exists a corresponding protocol solving P with success probability at least $1-\varepsilon$. Analogously, P is *solvable with 0-error* if the protocol refuses to answer with probability at most ε and solves P correctly whenever it produces an answer.

In this paper we will be interested in communication problems where the quantum model gives exponential savings.[2]

For 0-error one-way and interactive protocols, such problems were demonstrated by Buhrman, Cleve, and Wigderson [BCW98]. In the bounded-error setting the first exponential separation has been demonstrated by Raz [R99], who gave an example of a problem solvable in Q exponentially more efficiently than in 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 R^1.

All these separations have been demonstrated for functional problems. As of one-way protocols with bounded error, the first exponential 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 functional problem.

These result show that quantum communication can be very efficient, by establishing various settings where quantum protocols offer exponential savings over classical solutions. *But does there exist a problem that can be solved by a quantum one-way or even simultaneous protocol that is considerably more efficient than any classical two-way protocol?* The full answer to this question is not known yet.

1.1 Our result

Theorem 1.1. *For infinitely many $N \in \mathbb{N}$, there exists a 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 $\Omega\left(\frac{N^{1/8}}{\sqrt{\log N}}\right)$.*

The relation that we use for establishing this result 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

For any m being a power of 2, let $X_m \stackrel{\text{def}}{=} \mathcal{GF}_2^{\log m}$ and denote the identity element by $\bar{0}$. We will sometimes refer to subsets of X_m as elements of $\{0,1\}^m$. Define the following communication problems.

Definition 1. Let $x, y \subset X_{n^2}$, such that $|x| = n/2$ and $|y| = n$. Let $z \in X_{n^2} \setminus \{\bar{0}\}$. Then $(x,y,z) \in P_{1\times 1}^{(n)}$ if either $|x \cap y| \neq 2$ or $\langle z, a+b \rangle = 0$ where $x \cap y = \{a, b\}$. Let $\Sigma = \{\sigma_1, \dots\}$

[2] In all the examples mentioned here, the first shown super-polynomial separations were, in fact, exponential.

be a family of permutations, containing one permutation σ_i over $[i^2]$ for all i being powers of 2. Then $(x, y, z) \in P_{\Sigma}^{(n)}$ if $(\sigma_{n^2}(x), \sigma_{n^2}(y), z) \in P_{1\times 1}^{(n)}$.

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

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

Definition 3. Let $x \subset X_{n^2}$ and $y \subset X_{n^2}$, such that $|x| = n/2$ and $|y| = n$. Let $z \subset X_{n^2}$. Then $(x, y, z) \in \tilde{P}_{1\times 1}^{(n)}$ if $x \cap y = z$.

We will 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 $P_{1\times 1}^{(n)}$ 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 $\tilde{P}_{1\times 1}^{(n)}$. This intuition is, apparently, *false* for the quantum 1-way model ($P_{1\times 1}^{(n)}$ can be easily solved in Q^1 with probability $1/n$ with small error, which is unlikely to be the case for $\tilde{P}_{1\times 1}^{(n)}$). However, it is true for the model of classical 2-way communication; a "quasi-reduction" from $\tilde{P}_{1\times 1}^{(n)}$ to $P_{1\times 1}^{(n)}$ is one of the central ingredients of our lower bound proof.

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

- We claim that if there exists a protocol that solves $P^{(n)}$ with error bounded by ε then another protocol of similar cost solves $P_{\Sigma}^{(n)}$ for some Σ with probability $\Omega(1/n)$ and error $O(\varepsilon)$.
- We reduce the task of solving the problem $\tilde{P}_{1\times 1}^{(n)}$ to that of solving $P_{\Sigma}^{(n)}$.
- We show that the cost of solving $\tilde{P}_{1\times 1}^{(n)}$ with probability δ when $|x \cap y| = 2$ is $\Omega\left(n \cdot \sqrt{\delta}\right)$.
- We conclude that solving $P^{(n)}$ 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 $D(a_0) \stackrel{\text{def}}{=} \mathbf{Pr}_{a\sim D}[a = a_0]$; for $B \subseteq A$, $D(B) \stackrel{\text{def}}{=} \sum_{b\in B} D(b)$. Denote $\text{supp}(D) \stackrel{\text{def}}{=} \left\{a \in A \middle| D(a) > 0\right\}$.

We use the following notation.

$$\begin{aligned} \mathit{DISJ} &\stackrel{\text{def}}{=} \left\{(x,y)\middle|x,y\in\{0,1\}^*,\ |x|=|y|>0,\ \forall 1\le i\le |x|: x_i=0\vee y_i=0\right\}\\ \mathit{DISJ}_n &\stackrel{\text{def}}{=} \left\{(x,y)\in \mathit{DISJ}\middle|x,y\in\{0,1\}^n\right\}\\ \overline{\mathit{DISJ}} &\stackrel{\text{def}}{=} \left\{(x,y)\middle|x,y\in\{0,1\}^*,\ |x|=|y|>0,\ (x,y)\notin \mathit{DISJ}\right\}\\ \overline{\mathit{DISJ}}_n &\stackrel{\text{def}}{=} \left\{(x,y)\in \overline{\mathit{DISJ}}\middle|x,y\in\{0,1\}^n\right\} \end{aligned}$$

We use the standard notion of a *(combinatorial) rectangle*. The sides of a rectangle will always correspond to subsets of the input sets of Alice and Bob, as defined by the communication problem under consideration. We will use the same notation for an input rectangle and for the *event that 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 $x|_I \stackrel{\text{def}}{=} x\cap I$. For $B\subseteq 2^A$, let $B\|_I \stackrel{\text{def}}{=} \left\{x|_I\middle|x\in B\right\}$. For a distribution D over A, let $D|_I$ be the conditional distribution of $x\sim D$, subject to $x\in I$. For a distribution D over 2^A, let $D\|_I$ be the marginal distribution of $y\stackrel{\text{def}}{=} x|_I$, when $x\sim D$.

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 $(x,y)|_{\mathbf{Al}}\stackrel{\text{def}}{=} x$ and $(x,y)|_{\mathbf{Bo}}\stackrel{\text{def}}{=} y$. Similarly, define the operators $\|_{\mathbf{Al}}$ and $\|_{\mathbf{Bo}}$ for distributions and sets.

3.1 Definitions related to $P^{(n)}_{1\times 1}$ and $P^{(n)}$

Define the following events characterizing input to $P^{(n)}_{1\times 1}$, $P^{(n)}_{\Sigma}$ or $\tilde{P}^{(n)}_{1\times 1}$.

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., $\mathcal{X}_0\stackrel{\text{def}}{=}\cup_{n=2^i}\left\{(x,y)\in X_{n^2}\times X_{n^2}\middle|x\cap y=\emptyset\right\}$, and so forth.

Let $\mathcal{U}^{(n)}_{1\times1}$ be the uniform distribution of input to $P^{(n)}_{1\times 1}$, $\mathcal{U}_{\mathbf{Al}}\stackrel{\text{def}}{=}\mathcal{U}^{(n)}_{1\times1}\|_{\mathbf{Al}}$ and $\mathcal{U}_{\mathbf{Bo}}\stackrel{\text{def}}{=}\mathcal{U}^{(n)}_{1\times1}\|_{\mathbf{Bo}}$.

Definition 5. For $k_1,..,k_t\in\mathbb{N}$, let $\mathcal{U}^{(n;k_1,..,k_t)}_{1\times1}\stackrel{\text{def}}{=}\mathcal{U}^{(n)}_{1\times1}|_{\mathcal{X}_{k_1}\cup..\cup\mathcal{X}_{k_t}}$ and $\mathcal{U}^{(n;k_1+)}_{1\times1}\stackrel{\text{def}}{=}\mathcal{U}^{(n)}_{1\times1}|_{\cup_{i\ge k_1}\mathcal{X}_i}$.

Definition 6. Given input set A (not necessarily a rectangle), define $\mathcal{U}_A\stackrel{\text{def}}{=}\mathcal{U}^{(n)}_{1\times1}|_A$, $\mathcal{U}^{(\mathbf{Al})}_A\stackrel{\text{def}}{=}\mathcal{U}_A\|_{\mathbf{Al}}$ and $\mathcal{U}^{(\mathbf{Bo})}_A\stackrel{\text{def}}{=}\mathcal{U}_A\|_{\mathbf{Bo}}$. Given $k_1,..,k_t\in\mathbb{N}$, let $\mathcal{U}^{(k_1,..,k_t)}_A\stackrel{\text{def}}{=}\mathcal{U}_A|_{\mathcal{X}_{k_1}\cup..\cup\mathcal{X}_{k_t}}$.

Claim 3.1. *For sufficiently large n it holds that $\mathcal{U}^{(n)}_{1\times1}(\mathcal{X}_0)\ge 1/3$, $\mathcal{U}^{(n)}_{1\times1}(\mathcal{X}_1)\ge 1/6$ and $\mathcal{U}^{(n)}_{1\times1}(\mathcal{X}_2)\ge 1/13$. On the other hand, for any $t\le n/2$ it holds that $\mathcal{U}^{(n)}_{1\times1}\left(\cup_{i\ge t}\mathcal{X}_i\right)\le\left(\frac{3}{4}\right)^t$.*

Proof of Claim 3.1. Think about choosing $(x,y)\sim\mathcal{U}^{(n)}_{1\times1}$ as selecting a random subset $y\subset X_{n^2}$, $|y|=n$, followed by selecting $n/2$ different elements for x. Under such interpretation it is clear that $\mathcal{U}^{(n)}_{1\times1}\left(\cup_{i\ge t}\mathcal{X}_i\right)\le\binom{n/2}{t}\cdot\left(\frac{n}{n^2-n/2}\right)^t$. Therefore, $\mathcal{U}^{(n)}_{1\times1}(\mathcal{X}_0)\ge 1-n/2\cdot\frac{n}{n^2-n/2}\ge\frac{1}{3}$ and $\mathcal{U}^{(n)}_{1\times1}\left(\cup_{i\ge t}\mathcal{X}_i\right)\le\left(\frac{n}{2}\right)^t\cdot\left(\frac{3}{2n}\right)^t=\left(\frac{3}{4}\right)^t$, for $n\ge 2$.

Let E_i be the event that $i \in x \cap y$. 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 $\mathcal{U}_{1\times 1}^{(n)}(\mathcal{X}_1)$ is equal to $n/2$ times the probability that the first element selected for x belongs to y and all the following are not in y. The former occurs with probability at least $1/n$ and the latter with probability not smaller than $\mathcal{U}_{1\times 1}^{(n)}(\mathcal{X}_0)$, therefore $\mathcal{U}_{1\times 1}^{(n)}(\mathcal{X}_1) \geq \frac{n}{2} \cdot \frac{1}{n} \cdot \frac{1}{3} \geq \frac{1}{6}$.

Similarly, $\mathcal{U}_{1\times 1}^{(n)}(\mathcal{X}_2) \geq \binom{n/2}{2} \cdot \frac{1}{n} \cdot \frac{n-1}{n^2-1} \cdot \mathcal{U}_{1\times 1}^{(n)}(\mathcal{X}_0) > \frac{1}{13}$ for sufficiently large n. ■*Claim 3.1*

4 Efficient protocol for $P^{(n)}$ in Q^1

We give a 1-way quantum protocol $S^{(n)}$ that receives input to $P^{(n)}$, 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 $S^{(n)}$ in parallel, thus obtaining a 0-error protocol for $P^{(n)}$ 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 $S^{(n)}$ works.

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

2. Bob measures $|\alpha\rangle$ with the $\frac{n}{4}+1$ projectors $E_i \stackrel{\text{def}}{=} \sum_{j\in y_i} |j\rangle\langle j|$ and $E_0 \stackrel{\text{def}}{=} \sum_{j\notin \cup y_i} |j\rangle\langle j|$, 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 $i > 0$,

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

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 $\left|\alpha'_{i_0}\right\rangle$. Then

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

and therefore Bob obtains a uniformly random element of $\left\{ j \in X_{2n^2} \,\middle|\, \langle j, b_1 + b_2\rangle = 0 \right\}$ 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 $P^{(n)}$ is expensive in R

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

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

Lemma 5.1. [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 5.1.[3]

Lemma 5.2. *Let n be sufficiently large and A be an input rectangle for $DISJ_n$. Let D be a product distribution w.r.t. 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 $\delta = \frac{{\alpha_1}^2}{45 \cdot 4^{{\alpha_2}^2}}$ 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 5.2. We will reduce the communication task considered in Lemma 5.1 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 r as a uniform random string of sufficient length (we will address this situation by "$r \sim \mathcal{U}$") 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 $r \sim \mathcal{U}$ then $T_r(x', y')$ is uniformly distributed over $\mathcal{X}_i \cap \text{supp}(D)$, for any $i \geq 0$. In particular, for $i \in \{0, 1\}$,

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

[3] A direct proof of Lemma 5.2 would be simpler than that of Lemma 5.1. The latter is, in the qualitative sense, a much stronger statement than what we need (that is caused by the implicit requirement in Lemma 5.1 for non-product input distribution).

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

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

therefore

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

It is clear that T_r is rectangle-invariant, so B_r-s are rectangles and we can apply Lemma 5.1.

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

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

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

Note that

$$\begin{aligned}
D(\mathcal{X}_0) &\ge \left(\frac{n - k_1(n) - k_2(n)}{n}\right)^{k_2(n)} \ge \left(1 - \frac{2\alpha_2}{\sqrt{n}}\right)^{\alpha_2\sqrt{n}} \ge \left(\frac{1}{2}\right)^{2\alpha_2{}^2} = \left(\frac{1}{4}\right)^{\alpha_2{}^2}, \\
D(\mathcal{X}_1) &\ge k_2(n) \cdot \frac{k_1(n)}{n} \cdot D(\mathcal{X}_0) \ge \frac{\alpha_1{}^2}{4^{\alpha_2{}^2}}
\end{aligned}$$

(the second inequality can be established analogously to the proof of Claim 3.1). The result follows. ■*Lemma 5.2*

5.1 Solving $P^{(n)}$ implies solving $P_\Sigma^{(n)}$

Lemma 5.3. *Assume that there exists a protocol S of cost k that solves $P^{(n)}$ with error bounded by ε. Then there exists a family of permutations Σ, such that $P_\Sigma^{(n)}$ can be solved w.r.t. $\mathcal{U}_{1\times 1}^{(n;2)}$ with probability $2/n$ with error bounded by 2ε by a protocol of cost $O(k)$.*

Proof of Lemma 5.3. We will use the integers from $[2n^2]$ to address the elements of X_{2n^2}, according to their natural ordering. Let (x,y) be an instance of $P_{1\times 1}^{(n)}$, satisfying $|x \cap y| = 2$.

Consider the following protocol S'.

- Let $x' = \left\{n^2+1, \ldots, n^2+\frac{n}{2}\right\} \cup x$. For $1 \le j \le \frac{n}{4}-1$, let $y'_j = \left\{n^2 + j + \frac{kn}{4} \middle| 1 \le k \le n\right\}$ 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 $P^{(n)}$ (the deterministically constructed $(x', \bar{y})$ forms a correct input for $P^{(n)}$, and the action of permutations upon instances of $P^{(n)}$ is transitive). Moreover, the original problem is mapped to a uniformly random coordinate of the instance of $P^{(n)}$ that is fed into S.

Denote by E the event that S' returns an answer, by E_0 the event that S' outputs a pair (σ, z) such that $(\sigma(x), \sigma(y), z) \in P^{(n)}_{1\times 1}$, and by E_1 the event $E \setminus E_0$. By the symmetry argument, the following holds: If S returns a correct answer then E_0 occurs with probability $4/n$; if S makes a mistake then E_1 occurs with probability $4/n$. In particular, $\mathbf{Pr}\left[E\right] = 4/n$ and $\mathbf{Pr}\left[E_1\right] \le \varepsilon \cdot \mathbf{Pr}\left[E\right]$.

Let us derandomize S'. Suppose that S' uses s random bits and let r be the corresponding random variable. Let R_0 be the set of $r' \in \{0,1\}^s$, such that $\mathbf{Pr}\left[E_1 \middle| E, r = r'\right] \ge 2\varepsilon$ (here and till the end of the proof all the probabilities are taken w.r.t. $(x, y) \sim \mathcal{U}^{(n;2)}_{1\times 1}$). From the properties of S it follows that

$$\varepsilon \cdot \mathbf{Pr}\left[E\right] \ge \mathbf{Pr}\left[E_1\right] = \mathbf{Pr}\left[E\right] \cdot \mathbf{Pr}\left[E_1 \middle| E\right] \ge \mathbf{Pr}\left[r \in R_0\right] \cdot \mathbf{Pr}\left[E \middle| r \in R_0\right] \cdot 2\varepsilon,$$

which leads to

$$\frac{1}{2}\mathbf{Pr}\left[E\right] \le \mathbf{Pr}\left[E\right] - \mathbf{Pr}\left[r \in R_0\right] \cdot \mathbf{Pr}\left[E \middle| r \in R_0\right] = \mathbf{Pr}\left[r \notin R_0\right] \cdot \mathbf{Pr}\left[E \middle| r \notin R_0\right].$$

Therefore, there exists some $r_0 \in \{0,1\}^s \setminus R_0$, such that $\mathbf{Pr}\left[E \middle| r = r_0\right] \ge \mathbf{Pr}\left[E\right]/2 = 2/n$ and $\mathbf{Pr}\left[E_1 \middle| E, r = r_0\right] < 2\varepsilon$.

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 $r = r_0$, in particular, fixes the permutation $\sigma_1 \stackrel{\text{def}}{=} \sigma'_1$. Let us denote by Σ the family of σ'_1-s, obtained as a result of the described derandomization, subsequently applied to every permitted input length. We claim that S'' solves $P^{(n)}_{\Sigma}$ w.r.t. $\mathcal{U}^{(n;2)}_{1\times 1}$ 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. $\blacksquare_{\textit{Lemma 5.3}}$

5.2 Solving $\tilde{P}^{(n)}_{1\times 1}$ is as simple as solving $P^{(n)}_{\Sigma}$

We will show the following.

Theorem 5.4. *Assume that there exists a protocol of cost $k \in o(n) \cap \omega(1)$ that solves $P^{(n)}_{\Sigma}$ for some Σ w.r.t. $\mathcal{U}^{(n;2)}_{1\times 1}$ with probability $\gamma \in \omega\left(2^{-k}\right)$ and error bounded by 10^{-22}. Then $\tilde{P}^{(n)}_{1\times 1}$ can be solved w.r.t. $\mathcal{U}^{(n;2)}_{1\times 1}$ with probability $\frac{\gamma}{k^2 \cdot \log^2(n/\gamma)}$ in 0-error setting by a protocol of cost $O\left(k + \log^2(n/\gamma)\right)$.*

The proof will be done in several stages. Note that the relation $P^{(n)}_{1\times 1}$ is a special case of $P^{(n)}_{\Sigma}$, we will reason about properties of the former whenever generality of the latter is not essential.

Lemma 5.5. *Let n be sufficiently large and A be an input rectangle for $P^{(n)}_{1\times 1}$, such that $\mathcal{U}^{(n;1)}_{1\times 1}(A) \in 2^{-o(n)} \cap o(1)$. Assume that for some constant $0 < \varepsilon < 1$ and $I_0 \subseteq X_{n^2}$, $|I_0| \geq \frac{n^2}{2}$, it holds that*

$$\sum_{i \in I_0} \mathcal{U}^{(1)}_A \left(\mathcal{X}_1(i)\right) \leq \frac{\varepsilon}{10^6}.$$

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

The intuitive meaning of this lemma is that a rectangle that selectively accepts input pairs from $\mathcal{X}_1$ (mostly from $\cup_{i \notin I_0} \mathcal{X}_1(i)$) must reject pairs from $\mathcal{X}_0$ with high probability.

Proof of Lemma 5.5. 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 $\varepsilon_0 \stackrel{\text{def}}{=} \mathcal{U}^{(0,1)}_A(\mathcal{X}_0) \in \Omega(1)$, 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 $(x, y) = (x_1, x_2, y_1, y_2) \sim \mathcal{U}^{(0,1)}_A$. Assume that the values $x_1 = x_1'$ and $y_1 = y_1'$. Define the following events:

- E_1 denotes the event that $|x_1'| \leq \frac{n}{3}$ and $|y_1'| \leq \frac{2n}{3}$.
- E_2 denotes the event that $\mathbf{Pr}_{\mathcal{U}^{(0,1)}_A}\left[\mathcal{X}_0 \middle| x_1 = x_1', y_1 = y_1'\right] \geq \frac{\varepsilon_0}{2}$.
- E_3 denotes the event that

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

- E_4 denotes the event that

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

Observe that none of the events depends on the values of x_2 and y_2. Our first step will be to show that all four events hold simultaneously with non-negligible probability. This will let us apply Lemma 5.2 to many "subrectangles" of A defined over $I_0 \times I_0$, which, in turn, will lead to the desired lower bound.

The event E_1 occurs with probability $1 - 2^{-\Omega(n)}$ if $(x_1', y_1') \sim \mathcal{U}^{(n;0,1)}_{1\times 1}\|_{\overline{I_0} \times \overline{I_0}}$, due to the Chernoff bound. In our case $(x_1', y_1') \sim \mathcal{U}^{(0,1)}_A\|_{\overline{I_0} \times \overline{I_0}}$, but on the other hand, $\mathcal{U}^{(n;0,1)}_{1\times 1}(A) \in 2^{-o(n)}$, and therefore $\mathbf{Pr}_{\mathcal{U}^{(0,1)}_A}[E_1] \in 1 - o(1)$.

We know that

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

which implies that $\mathbf{Pr}_{\mathcal{U}^{(0,1)}_A}[E_2] \geq \frac{\varepsilon_0}{2}$.

Let us see that E_3 occurs with high probability. Observe that

$$\begin{aligned} \mathbf{H}\left[\mathcal{U}_A^{(0)}\right] &= \mathbf{H}\left[\mathcal{U}_A^{(0)}\|_{\overline{I_0}\times\overline{I_0}}\right] + \underset{\mathcal{U}_A^{(0)}}{\mathbf{H}}\left[x_2,y_2\middle|x_1,y_1\right]; \\ \mathbf{H}\left[\mathcal{U}_{1\times1}^{(n;0)}\right] &= \mathbf{H}\left[\mathcal{U}_{1\times1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}\right] + \underset{\mathcal{U}_{1\times1}^{(n;0)}}{\mathbf{H}}\left[x_2,y_2\middle|x_1,y_1\right] \\ &= \mathbf{H}\left[\mathcal{U}_{1\times1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}\right] + \mathbf{H}\left[\mathcal{U}_{1\times1}^{(n;0)}\|_{I_0\times I_0}\right], \end{aligned} \tag{1}$$

where the last equality follows from the fact that $\mathcal{U}_{1\times1}^{(n;0)}$ is a product distribution w.r.t. the marginal projections considered above. Moreover, $\mathcal{U}_{1\times1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}$ and $\mathcal{U}_{1\times1}^{(n;0)}\|_{I_0\times I_0}$ are uniform over their supports, and therefore

$$\mathbf{H}\left[\mathcal{U}_A^{(0)}\|_{\overline{I_0}\times\overline{I_0}}\right] \le \mathbf{H}\left[\mathcal{U}_{1\times1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}\right] \tag{2}$$

and

$$\underset{\mathcal{U}_A^{(0)}\|_{I_0\times I_0}}{\mathbf{H}}\left[x_2,y_2\middle|x_1=x_1',y_1=y_1'\right] \le \mathbf{H}\left[\mathcal{U}_{1\times1}^{(n;0)}\|_{I_0\times I_0}\right], \tag{3}$$

for any (x_1',y_1') in the support of $\mathcal{U}_A^{(0,1)}\|_{\overline{I_0}\times\overline{I_0}}$. In particular, (2) and (1) imply that

$$\underset{\mathcal{U}_A^{(0)}}{\mathbf{H}}\left[x_2,y_2\middle|x_1,y_1\right] \ge \mathbf{H}\left[\mathcal{U}_{1\times1}^{(n;0)}\|_{I_0\times I_0}\right] - \left(\mathbf{H}\left[\mathcal{U}_{1\times1}^{(n;0)}\right] - \mathbf{H}\left[\mathcal{U}_A^{(0)}\right]\right). \tag{4}$$

Observe that both $\mathcal{U}_{1\times1}^{(n;0)}$ and $\mathcal{U}_A^{(0)}$ are uniform over their supports; moreover, the latter support is a subset of the former. This leads to

$$\mathbf{H}\left[\mathcal{U}_{1\times1}^{(n;0)}\right] - \mathbf{H}\left[\mathcal{U}_A^{(0)}\right] = \log\left(\frac{\left|\mathrm{supp}\,\mathcal{U}_{1\times1}^{(n;0)}\right|}{\left|\mathrm{supp}\,\mathcal{U}_A^{(0)}\right|}\right) = \log\left(\frac{1}{\mathcal{U}_{1\times1}^{(n;0)}(A)}\right),$$

that, together with (4), gives

$$\mathbf{E}\left[\mathbf{H}\left[\mathcal{U}_A^{(0)}\middle|x_1=x_1',y_1=y_1'\right]\right] \ge \mathbf{H}\left[\mathcal{U}_{1\times1}^{(n;0)}\|_{I_0\times I_0}\right] - \log\left(\frac{1}{\mathcal{U}_{1\times1}^{(n;0)}(A)}\right),$$

where the expectancy is taken w.r.t. $(x_1',y_1') \sim \mathcal{U}_A^{(0,1)}\|_{\overline{I_0}\times\overline{I_0}}$. Together with (3), this implies that $\mathbf{Pr}_{\mathcal{U}_A^{(0,1)}}\left[E_3\right] \ge 1 - \frac{\varepsilon_0}{8}$.

Let us denote by G the set of pairs (x_1',y_1') that falsify the condition of E_4. Then, starting from (1), we get

$$\begin{aligned} &\mathbf{H}\left[\mathcal{U}_{1\times1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}\right] + \mathbf{H}\left[\mathcal{U}_{1\times1}^{(n;0)}\|_{I_0\times I_0}\right] = \mathbf{H}\left[\mathcal{U}_{1\times1}^{(n;0)}\right] \ge \mathbf{H}\left[\mathcal{U}_{1\times1}^{(n;0)}\middle|(x_1,y_1)\in G\right] \\ &= \mathbf{H}\left[\left(\mathcal{U}_{1\times1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}\right)\Big|_G\right] + \underset{\mathcal{U}_{1\times1}^{(n;0)};(x_1,y_1)\in G}{\mathbf{H}}\left[x_2,y_2\middle|x_1,y_1\right] \\ &\ge \mathbf{H}\left[\left(\mathcal{U}_{1\times1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}\right)\Big|_G\right] + \mathbf{H}\left[\mathcal{U}_{1\times1}^{(n;0)}\|_{I_0\times I_0}\right] + \log\left(\frac{8}{\varepsilon_0\cdot\mathcal{U}_{1\times1}^{(n;0,1)}(A)}\right), \end{aligned}$$

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

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

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

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

This leads to the conclusion that $\mathcal{U}_{1\times 1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}(G) \le \frac{\varepsilon_0}{8}\cdot\mathcal{U}_{1\times 1}^{(n;0,1)}(A)$. Note that G, by definition, consists exclusively of disjoint pairs. But it is easy to see that modulo the condition that $x_1'\cap y_1' = \emptyset$, the distribution $\mathcal{U}_{1\times 1}^{(n;0)}\|_{\overline{I_0}\times\overline{I_0}}$ is identical to $\mathcal{U}_{1\times 1}^{(n;0,1)}\|_{\overline{I_0}\times\overline{I_0}}$, and therefore $\mathcal{U}_{1\times 1}^{(n;0,1)}\|_{\overline{I_0}\times\overline{I_0}}(G) \le \frac{\varepsilon_0}{8}\cdot\mathcal{U}_{1\times 1}^{(n;0,1)}(A)$. This gives us that

$$\mathop{\mathbf{Pr}}_{\mathcal{U}_A^{(0,1)}}[E_4] = 1 - \mathcal{U}_A^{(0,1)}\|_{\overline{I_0}\times\overline{I_0}}(G) \ge 1 - \frac{\varepsilon_0}{8}.$$

For n sufficiently large, the events E_1, E_3 and E_4 simultaneously hold with probability at least $1-\frac{\varepsilon_0}{4}-o(1) > 1-\frac{\varepsilon_0}{3}$ and E_2 holds with probability at least $\frac{\varepsilon_0}{2}$. The event $E \stackrel{\text{def}}{=} E_1\cap E_2\cap E_3\cap E_4$ holds with probability at least $\frac{\varepsilon_0}{6}$, w.r.t. $(x_1',y_1')\sim\mathcal{U}_A^{(0,1)}\|_{\overline{I_0}\times\overline{I_0}}$.

It remains to apply Lemma 5.2 to the rectangles $A_{x_1',y_1'} \stackrel{\text{def}}{=} \left\{(x_2,y_2)\middle|(x_1',x_2,y_1',y_2)\in A\right\}$, where E holds w.r.t. x_1' and y_1'. Let us view $A_{x_1',y_1'}$ as an input rectangle for $DISJ_{|I_0|}$. Denote input to $DISJ_{|I_0|}$ by (x_2,y_2). Define D to be the distribution obtained by independently choosing x_2 and y_2 as subsets of I_0 of sizes $\frac{n}{2}-|x_1'|$ and $n-|y_1'|$, respectively. As follows from E_1, $\frac{n}{6}\le|x_2|\le\frac{n}{2}$ and $\frac{n}{3}\le|y_2|\le n$. The fact that $x_1'\cap y_1'=\emptyset$ (as implied by E_2) means that the restrictions of D to the cases of $\mathcal{X}_1$ and $\mathcal{X}_0$ are isomorphic (in the obvious sense) to the corresponding restrictions of $\mathcal{U}_{1\times 1}^{(n)}$, conditioned upon $x_1=x_1', y_1=y_1'$. Moreover, the same isomorphism maps A to $A_{x_1',y_1'}$.

Lemma 5.2 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 $\delta = \frac{{\alpha_1}^2}{45\cdot 4^{{\alpha_2}^2}} \ge \frac{1}{25920}$,

$$D(A_{x_1',y_1'}\cap\mathcal{X}_1) \ge \delta\cdot D(A_{x_1',y_1'}\cap\mathcal{X}_0) - 2^{-\Omega\left(\sqrt{|I_0|}\right)} \ge \frac{D(A_{x_1',y_1'}\cap\mathcal{X}_0)}{25920} - 2^{-\Omega(n)}. \qquad (5)$$

Let $D_A \stackrel{\text{def}}{=} D|_{A_{x_1',y_1'}}$ and $D_0 \stackrel{\text{def}}{=} D|_{\mathcal{X}_0}$. Events E_3 and E_4 together mean that for n sufficiently large (recall that $\mathcal{U}_{1\times 1}^{(n;0)}(A)\in o(1)$),

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

where $\Delta \in O\left(\log\left(\frac{1}{\mathcal{U}_{1\times 1}^{(n;0)}(A)}\right)\right)$ (because $\mathcal{U}_{1\times 1}^{(n;0,1)}(A) \geq \mathcal{U}_{1\times 1}^{(n;0,1)}(\mathcal{X}_0)\cdot \mathcal{U}_{1\times 1}^{(n;0)}(A)$). That can be restated as

$$\mathbf{H}\left[D_A|_{\mathcal{X}_0}\right] \geq \mathbf{H}\left[D_0\right] - \Delta,$$

and again, the both arguments of $\mathbf{H}[\cdot]$ are uniform distributions, one support being a subset of the other, which 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 $\mathcal{U}_{1\times 1}^{(n;1)}(A) \in 2^{-o(n)}$. On the other hand, $\mathcal{U}_{1\times 1}^{(n;0,1)}(\mathcal{X}_0 \cup A) = \mathcal{U}_{1\times 1}^{(n;0,1)}(A)\cdot \mathcal{U}_A^{(0,1)}(\mathcal{X}_0) = \mathcal{U}_{1\times 1}^{(n;0,1)}(\mathcal{X}_0)\cdot \mathcal{U}_{1\times 1}^{(n;0)}(A)$ implies that

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

and therefore $D_0(A_{x_1',y_1'}) \in 2^{-o(n)}$. The fact that E_2 holds implies that

$$D(A_{x_1',y_1'} \cap \mathcal{X}_0) \geq \frac{\varepsilon_0}{2}\cdot D_0(A_{x_1',y_1'} \cap \mathcal{X}_0) = \frac{\varepsilon_0}{2}\cdot D_0(A_{x_1',y_1'}) \in 2^{-o(n)},$$

where the containment follows from (6) and (7). This means that for sufficiently large n, (5) leads to $D(A_{x_1',y_1'} \cap \mathcal{X}_1) > D(A_{x_1',y_1'} \cap \mathcal{X}_0)/25921$, implying $\mathbf{Pr}_{D_A}\left[\mathcal{X}_1 \middle| \mathcal{X}_0 \cup \mathcal{X}_1\right] > \frac{1}{25921}$.

We conclude:

$$\sum_{i\in I_0} \mathcal{U}_A^{(1)}\left(\mathcal{X}_1(i)\right) \geq \sum_{i\in I_0} \mathcal{U}_A^{(0,1)}\left(\mathcal{X}_1(i)\right) \geq \mathop{\mathbf{Pr}}_{\mathcal{U}_A^{(0,1)}}[E]\cdot \mathop{\mathbf{Pr}}_{D_A}\left[\mathcal{X}_1 \middle| \mathcal{X}_0 \cup \mathcal{X}_1\right] > \frac{\varepsilon_0}{10^6},$$

as required. ■*Lemma 5.5*

We will need the following corollary, that extends the statement of Lemma 5.5 to the case of rectangles, selectively accepting instances of $\mathcal{X}_2$.

Corollary 5.6. *Let n be sufficiently large and A be an input rectangle for $P_{1\times 1}^{(n)}$, such that $\mathcal{U}_{1\times 1}^{(n;2)}(A) \in 2^{-o(n)}$. Let $\left(I_0^{(i)}\right)_{i\in X_{n^2}}$ be a family of subsets of X_{n^2}, such that for every $i \in X_{n^2}$ it holds that $i \notin I_0^{(i)}$ and $|I_0^{(i)}| \geq \frac{n^2}{2}$, and for every $i,j \in X_{n^2}$ it holds that $i \in I_0^{(j)}$ if and only if $j \in I_0^{(i)}$. If A satisfies that*

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

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

Proof of Corollary 5.6. We will show that $\mathcal{U}_A^{(0,1,2)}(\mathcal{X}_1) \leq \frac{\varepsilon}{1441}$ and $\mathcal{U}_A^{(0,1,2)}(\mathcal{X}_0) \leq \frac{\varepsilon}{2}$.

Define $A_i \stackrel{\text{def}}{=} \left\{(x,y) \in A \middle| i \in x \cap y\right\}$ for each $i \in X_{n^2}$. Let D be the probability distribution over X_{n^2} defined by $D(i) = \frac{1}{2}\mathcal{U}_A^{(2)}(A_i)$, then choosing $(x,y) \sim \mathcal{U}_A^{(2)}$ can be viewed as first choosing $i \sim D$, followed by $(x,y) \sim \mathcal{U}^{(2)}_{A_i}$. The main condition of the corollary can now be expressed as

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

Let

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

$$I_2 \stackrel{\text{def}}{=} \left\{ i \in X_{n^2} \middle| \sum_{j\in I_0^{(i)}} \mathcal{U}^{(2)}_{A_i}(\mathcal{X}_2(i,j)) > \frac{\varepsilon}{10^{13}} \right\}.$$

Then

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

and

$$D(I_2) < \frac{\varepsilon}{10^7} \Rightarrow \sum_{i\in I_2} \mathcal{U}^{(2)}_{A}(A_i) < \frac{2\varepsilon}{10^7}.$$

That is,

$$\sum_{i\in I_1\cup I_2} \mathcal{U}^{(2)}_{A}(A_i) < \frac{3\varepsilon}{10^7}. \tag{8}$$

For any $i_0 \in X_{n^2}$, we treat A_{i_0} as an input rectangle for $P^{(n-1)}_{1\times 1}$, defined over $X_{n^2} \setminus \{i_0\}$.[4] For $i_0 \in X_{n^2} \setminus I_1 \setminus I_2$, it holds that $\mathcal{U}^{(n;2)}_{1\times 1}(A_{j_0}) \ge \frac{\varepsilon}{10^7 \cdot n^2} \cdot \mathcal{U}^{(n;2)}_{1\times 1}(A) \in 2^{-o(n)}$. The properties of $I_0^{(i_0)}$ and the fact that $i_0 \notin I_2$ allow us to apply Lemma 5.5, concluding that

$$\mathcal{U}^{(1,2)}_{A_{i_0}}(\mathcal{X}_1) < \frac{\varepsilon}{10^7}. \tag{9}$$

For every $i_0 \in I_1 \cup I_2$ we, on the other hand, apply Lemma 5.2 to A_{i_0}. Then for $\delta = \frac{1}{720}$,

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

Clearly,

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

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

$$\begin{aligned}\sum_{i\in I_1\cup I_2} \mathcal{U}^{(1,2)}_{A}(A_i \cap \mathcal{X}_1) &\le \frac{1}{\delta} \cdot \sum_{i\in I_1\cup I_2} \mathcal{U}^{(1,2)}_{A}(A_i \cap \mathcal{X}_2) + 2^{2\log n - \Omega(n) + o(n)} \\ &\le \frac{1}{\delta} \cdot \sum_{i\in I_1\cup I_2} \mathcal{U}^{(2)}_{A}(A_i) + 2^{-\Omega(n)} \le \frac{3\varepsilon}{\delta \cdot 10^7} + 2^{-\Omega(n)},\end{aligned} \tag{12}$$

[4] Strictly speaking, this violates our requirement that n is a power of 2 and slightly affects the Hamming waits of x and y as functions of n, though the former is irrelevant for the present context and the influence of the latter is negligible for sufficiently large n. We allow this abuse to keep the notation simple.

as follows from (8).

We conclude that for sufficiently large n,

$$\mathcal{U}_A^{(0,1,2)}(\mathcal{X}_1) \le \mathcal{U}_A^{(1,2)}(\mathcal{X}_1) = \sum_{i\in X_{n^2}} \mathcal{U}_A^{(1,2)}(A_i \cap \mathcal{X}_1)$$
$$= \sum_{i\in I_1\cup I_2} \mathcal{U}_A^{(1,2)}(A_i \cap \mathcal{X}_1) + \sum_{i\notin I_1\cup I_2} \mathcal{U}_A^{(1,2)}(A_i)\cdot \mathcal{U}_{A_{i_0}}^{(1,2)}(\mathcal{X}_1)$$
$$< \frac{3\varepsilon}{\delta\cdot 10^7} + 2^{-\Omega(n)} + \frac{\varepsilon}{10^7} < \frac{\varepsilon}{1441},$$

as follows from (9) and (12).

We apply Lemma 5.2 one more time. For the same value of δ it holds that

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

Like in (11),

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

and therefore for sufficiently large n,

$$\mathcal{U}_A^{(0,1,2)}(\mathcal{X}_0) \le \frac{1}{\delta}\cdot \mathcal{U}_A^{(0,1,2)}(\mathcal{X}_1) + 2^{-\Omega(n)} < \frac{720\cdot\varepsilon}{1441} + 2^{-\Omega(n)} < \frac{\varepsilon}{2}.$$

The result follows. ■ *Corollary 5.6*

The next lemma will be the last preparation step before we prove the main result of this section.

Lemma 5.7. *Let n be sufficiently large and A be an input rectangle for $P_{1\times 1}^{(n)}$, such that $\mathcal{U}_A(\mathcal{X}_0\cup\mathcal{X}_1) \le \frac{1}{6}$ and for $\delta > 0$*

$$\mathop{\mathbf{Pr}}_{y\sim \mathcal{U}_A^{(Bo)}}\left[\exists\,\{a\ne b\}\subset y:\ \mathop{\mathbf{Pr}}_{x\sim \mathcal{U}_A^{(Al)}}\left[\{a,b\}\subset x\right]\ge \delta\right] \le \frac{1}{3}.$$

Then $\mathcal{U}_{1\times 1}^{(n)}(A) \in 2^{-\Omega\left(\frac{1}{\sqrt{\delta}}\right)}$.

Proof of Lemma 5.7. Let B be the set of $y\in X_{n^2}$, such that

$$\mathop{\mathbf{Pr}}_{x\sim \mathcal{U}_A^{(\mathbf{Al})}}\left[|x\cap y|<2\right] \le \frac{1}{3} \tag{13}$$

and

$$\forall\,\{a\ne b\}\subset y:\ \mathop{\mathbf{Pr}}_{x\sim \mathcal{U}_A^{(\mathbf{Al})}}\left[\{a,b\}\subset x\right] < \delta. \tag{14}$$

If we choose $y'\sim \mathcal{U}_A^{(\mathbf{Bo})}$ then (13) holds with probability at least $\frac{1}{2}$ and (14) holds with probability at least $\frac{2}{3}$, therefore $|B\cap A\|_{\mathbf{Bo}}| \ge \frac{1}{6}\,|A\|_{\mathbf{Bo}}|$. Denote $A' \stackrel{\text{def}}{=} A\|_{\mathbf{Al}}\times B$, then $\mathcal{U}_{1\times 1}^{(n)}(A') \ge \frac{1}{6}\mathcal{U}_{1\times 1}^{(n)}(A)$.

For $\{a \neq b\} \subset X_{n^2}$, let $p_a \stackrel{\text{def}}{=} \mathbf{Pr}_{x \sim \mathcal{U}_A^{(\mathbf{Al})}} [a \in x]$ and $p_b^{(a)} \stackrel{\text{def}}{=} \mathbf{Pr}_{x \sim \mathcal{U}_A^{(\mathbf{Al})}} \left[b \in x \middle| a \in x\right]$. Condition (14) holds only if

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

Let $a_0 \in y$ be the lexicographically first value satisfying $p_{a_0} = \max_{i \in y} \{p_i\}$. Think about the process of choosing $y \sim \mathcal{U}_{\mathbf{Bo}}$ as first choosing a_0 and then the rest of the elements. We will see that conditions (13) and (15) are not likely to hold simultaneously.

First let us consider the situation when

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

Since $\mathbf{Pr}\left[|x \cap y| \geq 1 \middle| x \in A \| _{\mathbf{Al}}\right] \geq \frac{2}{3}$ can occur only if $\sum_{a \in y} p_a \geq \frac{2}{3}$, the probability that (13) and (16) hold is upper bounded by the probability that

$$\sum_{a \in y} p'_a \geq \frac{2}{3}, \tag{17}$$

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

Let $Z_1, .., Z_n$ be the elements of y and denote $W_i \stackrel{\text{def}}{=} p'_{Z_i}$. We want to use Chernoff bound in order to limit from above the value of $\sum_{i=1}^n W_i$. Strictly speaking, the variables W_i are not independent (because all Z_i-s are different), but their dependence is relatively small, which makes it possible to apply Chernoff bound using the "worst case" estimation of the variables' mean values. Note that for n large enough and any $1 \leq i_0 \leq n$ it holds that $W_{i_0} \leq \sqrt{\delta}$ and $\mathbf{E}[W_{i_0}] \leq \frac{|x|}{|X_{n^2}| - |y|} = \frac{n/2}{n^2 - n} < \frac{3}{5n}$, where the mean value is computed w.r.t. repeated "experiments", for the fixed i_0. Based on Chernoff bound, we conclude that

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

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

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

Since $\mathbf{Pr}\left[|x \cap y| \geq 2 \middle| x \in A \| _{\mathbf{Al}}, a_0 \in y\right] \geq \frac{2}{3}$ can occur only if $\sum_{b \in y \setminus \{a_0\}} p_b^{(a_0)} \geq \frac{2}{3}$, the probability that (13) and (19) hold is upper bounded by the probability that

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

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

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

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

as required. ■*Lemma 5.7*

We are ready for the

Proof of Theorem 5.4. Let S be a deterministic protocol of cost k solving $P_{\Sigma}^{(n)}$ for some Σ w.r.t. $\mathcal{U}_{1\times 1}^{(n;2)}$ with probability γ and error bounded by $\frac{1}{10^{22}}$.

We will call a rectangle A *δ-labeled* if

$$\underset{y\sim\mathcal{U}_A^{(\mathbf{Bo})}}{\mathbf{Pr}}\left[\exists\,\{a\neq b\}\subset y:\underset{x\sim\mathcal{U}_A^{(\mathbf{Al})}}{\mathbf{Pr}}\left[\{a,b\}\subset x\right]\geq\delta\right]>\frac{1}{3}.$$

Observe that Lemma 5.7 guarantees that if $\mathcal{U}_A(\mathcal{X}_0\cup\mathcal{X}_1)\leq\frac{1}{6}$ and $\mathcal{U}_{1\times 1}^{(n)}(A)\geq 2^{-\Omega(k)}$ then there exists a function $\delta(k)\in\Omega\left(\frac{1}{k^2}\right)$, such that A is $\delta(k)$-labeled. Fix any 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 $P_{\Sigma}^{(n)}$ with probability at least $1-\frac{2}{10^{22}}$ w.r.t. $\mathcal{U}_A^{(2)}$. It follows from the properties of S that $(x,y)\sim\mathcal{U}_{1\times 1}^{(n;2)}$ does not belong to a latent rectangle with probability at least $\frac{\gamma}{2}$ (at leas 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 the allowed $\frac{1}{10^{22}}$). On the other hand, with probability at least $1-\frac{\gamma}{4}$ it happens that $(x,y)\sim\mathcal{U}_{1\times 1}^{(n;2)}$ falls into a rectangle A satisfying $\mathcal{U}_{1\times 1}^{(n;2)}(A)\geq\frac{\gamma}{2^{k+2}}$. Note that for any such A it holds that $\mathcal{U}_{1\times 1}^{(n)}(A)\geq\mathcal{U}_{1\times 1}^{(n)}(\mathcal{X}_2)\cdot\frac{\gamma}{2^{k+2}}\geq 2^{-2k}$ for n large enough (recall that $k\in\omega(1)$).

Call a rectangle A *good* if it is not latent and $\mathcal{U}_{1\times 1}^{(n)}(A)\geq 2^{-2k}$. It holds that $(x,y)\sim\mathcal{U}_{1\times 1}^{(n;2)}$ falls into a good rectangle with probability at least $\frac{\gamma}{2}-\frac{\gamma}{4}=\frac{\gamma}{4}$. Consequently, $(x,y)\sim\mathcal{U}_{1\times 1}^{(n;2+)}$ falls into a good rectangle with probability at least $\mathcal{U}_{1\times 1}^{(n)}(\mathcal{X}_2)\cdot\frac{\gamma}{4}\geq\frac{\gamma}{52}$.

Any good A is $\delta(k)$-labeled. It follows from the fact that there exists some $z_A\in X_{n^2}\setminus\{0\}$, such that

$$1-\frac{2}{10^{22}}\leq\underset{\mathcal{U}_A^{(2)}}{\mathbf{Pr}}\left[(x,y,z_A)\in P_{\Sigma}^{(n)}\right]=\underset{\mathcal{U}_A^{(2)}}{\mathbf{Pr}}\left[\langle z_A,\sigma_{n^2}(a)+\sigma_{n^2}(b)\rangle=0\right],$$

where $x\cap y=\{a,b\}$ and $\sigma_{n^2}\in\Sigma$. If we define $I_0^{(a)}\stackrel{\text{def}}{=}\left\{b\in X_{n^2}\,\middle|\,\langle z_A,\sigma_{n^2}(a)+\sigma_{n^2}(b)\rangle=1\right\}$ that will satisfy the requirement of Corollary 5.6 for $\varepsilon=\frac{1}{6}$, therefore it holds that $\mathcal{U}_A(\mathcal{X}_0\cup\mathcal{X}_1)\leq\mathcal{U}_A^{(0,1,2)}(\mathcal{X}_0\cup\mathcal{X}_1)<\frac{1}{6}$. As $\mathcal{U}_{1\times 1}^{(n)}(A)\geq 2^{-2k}$, we can apply the contrapositive of Lemma 5.7 (as suggested in the beginning of the proof), which guarantees that A is $\delta(k)$-labeled.

Let us construct a protocol satisfying the promise of our theorem. We will use an efficient randomized mapping of any $(x,y)\in\mathcal{X}_2$ to $(x',y')\sim\mathcal{U}_{1\times 1}^{(n;2+)}$, then feed (x',y') to the original protocol S, hoping that the pair will fall into a $\delta(k)$-labeled rectangle. Let D be the distribution over $[n]$ satisfying $D(j)\stackrel{\text{def}}{=}\mathcal{U}_{1\times 1}^{(n;2+)}(\mathcal{X}_j)$. Consider the following protocol S'.

1. Alice chooses $j_0\sim D$. 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 $(x_1,..,x_{j_0})$.

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

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

4. Alice and Bob use public randomness to choose a random permutation ρ over the elements of X_{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 $\{a \neq b\} \subset y$, such that $\mathbf{Pr}_{x \sim \mathcal{U}_A^{(\mathbf{Al})}} [\{a, b\} \subset x] \geq \delta(k)$, 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 those two elements. Otherwise the protocol returns no answer.

It is clear that the protocol is 0-error and its communication cost is $O(k + j_0 \cdot \log n) \subseteq O\left(k + \log^2(n/\gamma)\right)$. 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 behavior of S'':

- E_1 is the event that in the stage 5 of S'' a pair (a', b') has been chosen and $\left\{\rho^{-1}(a'), \rho^{-1}(b')\right\} \subseteq \tilde{x} \cap y$.

- E_2 is the event that $j_0 \leq 3 \log\left(\frac{312}{\gamma \cdot \delta(k)}\right)$ and E_1 occurs.

- E_3 is the event that $j_0 \leq 3 \log\left(\frac{312}{\gamma \cdot \delta(k)}\right)$, (a', b') has been chosen and $\left\{\rho^{-1}(a'), \rho^{-1}(b')\right\} \subseteq x \cap y$.

Obviously, the probability that S' is successful is equal to the probability that E_3 occurs.

Event E_1 occurs if $\Big[$ $(\rho(\tilde{x}), \rho(y))$ belongs to a $\delta(k)$-labeled rectangle $\Big]$ and $\Big[$ for some $\{a', b'\} \subset y$ it holds that $\mathbf{Pr}_{x \sim \mathcal{U}_A^{(\mathbf{Al})}} [\{a', b'\} \subset x] \geq \delta(k)$ $\Big]$ and $\Big[$ $\{a', b'\} \subset x$ $\Big]$, denote these events by $E_1^{(1)}$, $E_1^{(2)}$ and $E_1^{(3)}$, respectively. Note that since ρ is a uniformly random permutation and $j_0 \sim D$, it holds that $(\rho(\tilde{x}), \rho(y)) \sim \mathcal{U}_{1\times 1}^{(n;2+)}$, and so $\mathbf{Pr}\left[E_1^{(1)}\right] \geq \frac{\gamma}{52}$. By the definition of a $\delta(k)$-labeled rectangle, $\mathbf{Pr}\left[E_1^{(2)} \middle| E_1^{(1)}\right] \geq \frac{1}{3}$. Clearly, $\mathbf{Pr}\left[E_1^{(3)} \middle| E_1^{(3)}\right] \geq \delta(k)$. Therefore, $\mathbf{Pr}[E_1] \geq \frac{\gamma \cdot \delta(k)}{156}$.

Event E_2 occurs if E_1 occurs and $j_0 \leq 3 \log\left(\frac{312}{\gamma \cdot \delta(k)}\right)$, therefore

$$\mathbf{Pr}[E_2] \geq \frac{\gamma \cdot \delta(k)}{156} - \mathbf{Pr}_D\left[j_0 > 3 \log\left(\frac{312}{\gamma \cdot \delta(k)}\right)\right] \geq \frac{\gamma \cdot \delta(k)}{312},$$

where the second inequality follows from Claim 3.1.

Finally, E_3 occurs if E_2 occurs and the points $\rho^{-1}(a')$ and $\rho^{-1}(b')$ belong to $x \cap y$. Given j_0, the randomized mapping of (x, y) to $(\rho(\tilde{x}), \rho(y))$ produces a uniformly random instance according to $\mathcal{U}_{1\times 1}^{(n;2+)}(\mathcal{X}_j)$. Moreover, the two elements of $x \cap y$ are mapped to uniformly random elements of $\rho(\tilde{x}) \cap \rho(y)$. Therefore, the probability that $\{\rho^{-1}(a'), \rho^{-1}(b')\} = x \cap y$ is equal to $1 \Big/ \binom{j_0}{2} \geq \frac{1}{j_0^2}$. But E_2 guarantees that $j_0 \leq 3 \log\left(\frac{312}{\gamma \cdot \delta(k)}\right)$, and so $\mathbf{Pr}\left[E_3 \middle| E_2\right] \in \Omega\left(1 \Big/ \log^2\left(\frac{1}{\gamma \cdot \delta(k)}\right)\right)$. Given that $\delta(k) \in \Omega\left(\frac{1}{k^2}\right)$, we obtain $\mathbf{Pr}\left[E_3\right] \in \Omega\left(\frac{\gamma}{k^2 \cdot \log^2(n/\gamma)}\right)$.

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 5.4*

5.3 Solving $\tilde{P}_{1\times 1}^{(n)}$ is expensive

It is not hard to see that a protocol of communication cost k can solve $DISJ_n$ only with probability $O\left(\frac{k}{n}\right)$. In this section we will prove the following generalization of this statement.[5]

Theorem 5.8. *Let $t \in o\left(\sqrt{n}\right)$, then any 0-error protocol of cost $k \in \Omega\left(t \log n\right)$ solving $\tilde{P}_{1\times 1}^{(n)}$ w.r.t. $\mathcal{U}_{1\times 1}^{(n;t)}$ can succeed with probability $O\left(\left(\frac{kt}{n}\right)^t\right)$.*

Proof of Theorem 5.8. Let S be a 0-error protocol of cost k solving $\tilde{P}_{1\times 1}^{(n)}$ w.r.t. $\mathcal{U}_{1\times 1}^{(n;t)}$ with probability $p_t^{(t)}$. Let us define $p_i^{(t)}$ for $i > t$ to be the probability that S outputs t elements from $x \cap y$ when $(x, y) \sim \mathcal{U}_{1\times 1}^{(n;i)}$.

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 sufficiently large such that $t + \frac{n}{2ck} < \frac{n}{2}$. If for any i, $t \leq i \leq t + \frac{n}{2ck}$, it holds that $p_i^{(t)} \leq \left(\frac{k}{n}\right)^t$ then let i_0 be the smallest value like that and $p_t^{(t)} \leq \left(1 + \frac{ck}{n}\right)^{i_0 - t} p_{i_0}^{(t)} \in O\left(\left(\frac{k}{n}\right)^t\right)$. Otherwise

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

as required.

Now we prove the proposition. Let $i_0 \geq t$ 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 $p_{i_0}^{(t)} \leq \left(1 + \frac{ck}{n}\right)\left(1 - \frac{t}{i+1}\right) p_{i+1}^{(t)}$. Let $m \stackrel{\text{def}}{=} n^2 - i_0$, consider the following public coin protocol S' running on input (x', y'), such that $x' \subset [m]$, $|x'| = n/2 - i_0$, $y' \subset [m]$, $|y'| = n - i_0$.

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.

[5] We believe that this theorem might be of independent interest.

Let us assume that we know that either $(x',y') \in \mathcal{X}_0$ or $(x',y') \in \mathcal{X}_1$ and our goal is to distinguish the two cases. In the first case 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 $p_{i_0}^{(t)}$. In the second case 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 $p_{i_0+1}^{(t)} \cdot \binom{i_0}{t} \Big/ \binom{i_0+1}{t} = \left(1-\frac{t}{i_0+1}\right) \cdot p_{i_0+1}^{(t)}$. We know that the former is higher than the latter, and so if S' outputs $\mathbf{0}$ that can be viewed as an argument towards $(x',y') \in \mathcal{X}_0$.

Define D as the uniform distribution over the domain of S'. Then $D(\mathcal{X}_0) \geq \frac{1}{3}$ (by analogy to Claim 3.1), and the whole situation satisfies the requirements of Lemma 5.2 for $\alpha_1 = \frac{1}{4}$ and $\alpha_2 = 1$. The lemma implies that for $\delta = \frac{1}{3000}$, 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 and outputs $\mathbf{0}$ if all the instantiations of S' return $\mathbf{0}$ (otherwise S'_l refuses to answer). Denote by E_0 the event that S'_l outputs $\mathbf{0}$. If $(x',y') \in \mathcal{X}_0$ then E_0 occurs with probability $\left(p_{i_0}^{(t)}\right)^l$, if $(x',y') \in \mathcal{X}_1$ then E_0 occurs with probability $\left(\left(1-\frac{t}{i_0+1}\right) \cdot p_{i_0+1}^{(t)}\right)^l$. Therefore,

$$\Pr_D\left[\mathcal{X}_0 \text{ and } E_0\right] \geq \frac{1}{3} \cdot \left(p_{i_0}^{(t)}\right)^l$$

and

$$\Pr_D\left[\mathcal{X}_1 \text{ and } E_0\right] \leq \left(\left(1-\frac{t}{i_0+1}\right) \cdot p_{i_0+1}^{(t)}\right)^l.$$

Suppose that S'_l uses s uniformly distributed random bits. For any $r \in \{0,1\}^s$, let $S'_l(r)$ be the deterministic protocol obtained from S'_l by using the bits of r instead of the random bits. Note that $S'_l(r)$ is a protocol of communication cost kl, therefore it partitions the domain into rectangles $A_1^{(r)}, \ldots, A_{2^{kl}}^{(r)}$. Let B be the set of all $A_i^{(r)}$-s on which $S'_l(r)$ outputs $\mathbf{0}$.

Let

$$\beta(l) \stackrel{\text{def}}{=} \frac{1}{3} \cdot \left(\frac{p_{i_0}^{(t)}}{\left(1-\frac{t}{i_0+1}\right) \cdot p_{i_0+1}^{(t)}}\right)^l,$$

then

$$\frac{1}{2^s} \cdot \sum_{A \in B} D\left(A \cap \mathcal{X}_0\right) = \Pr_D\left[\mathcal{X}_0 \text{ and } E_0\right] \geq \beta(l) \cdot \Pr_D\left[\mathcal{X}_1 \text{ and } E_0\right] = \frac{\beta(l)}{2^s} \cdot \sum_{A \in B} D\left(A \cap \mathcal{X}_1\right).$$

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

$$\mu \geq \frac{1}{2^{kl}} \cdot \Pr_D [\mathcal{X}_0 \text{ and } E_0] \geq \frac{\left(p_{i_0}^{(t)}\right)^l}{3 \cdot 2^{kl}} > \frac{k^{tl}}{3 \cdot 2^{kl} \cdot n^{tl}} > 2^{-kl-tl\log n-2},$$

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

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

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

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

Recall that $k \in \Omega(t\log n)$, so there exists an absolute constant c_1 that guarantees that $2^{l(k+t\log n)+3-c_0 \cdot n} < \frac{\delta}{2}$ 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 5.8*

5.4 Lower bound on the classical 2-way communication complexity of $P^{(n)}$

Claim 5.9. *Solving $P^{(n)}$ 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 Claim 5.9. Assume that a protocol S of communication cost $k \in o(n)$ solves $P^{(n)}$ with error bounded by $\frac{1}{2 \cdot 10^{22}}$.

Then Lemma 5.3 implies that there exists a protocol S' of communication cost $O(k)$ that solves $P_{\Sigma}^{(n)}$ for some Σ w.r.t. $\mathcal{U}_{1\times 1}^{(n;2)}$ with probability $\frac{1}{2n}$ and error bounded by $\frac{1}{10^{22}}$.

By Theorem 5.4 there exists a protocol S'' of communication cost $O\left(k + \log^2(n)\right)$ solving $\tilde{P}_{1\times 1}^{(n)}$ in 0-error setting w.r.t. $\mathcal{U}_{1\times 1}^{(n;2)}$ with probability $\Omega\left(\frac{1}{nk^2\log^2(n)}\right)$.

Choose $t = 2$, Theorem 5.8 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. ■ *Claim 5.9*

6 Conclusions and further work.

The protocol described in Section 4 together with Claim 5.9 imply Theorem 1.1.

It would be interesting to strengthen this result. Is it possible to find a *functional* problem that requires exponentially more expensive protocol in R than in Q^1? How about simultaneous protocols?

In other words, give a separation that would logically imply as many results mentioned in the Introduction as possible.

Acknowledgments

This work has started from Richard Cleve's sharing with me his conjecture and Harry Buhrman's letting me know about the conjecture made by Serge Massar.

I would like to thank Alexander Razborov for finding a mistake in a earlier draft.

References

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

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

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

[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, pp. 128-137*, 2004.

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

[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*, 2007.

[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), pp. 385-390*, 1992.

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

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