Quantum versus classical simultaneity in communication complexity

Dmitry Gavinsky*

March 29, 2018

Abstract

This work addresses two problems in the context of two-party communication complexity of functions. First, it concludes the line of research, which can be viewed as demonstrating qualitative advantage of quantum communication in the three most common communication "layouts": two-way interactive communication; one-way communication; simultaneous message passing (SMP). We demonstrate a functional problem $\widetilde{cEq}\text{-}neg_T$, whose communication complexity is $O\big((\log n)^2\big)$ in the quantum version of SMP and $\tilde{\Omega}(\sqrt{n})$ in the classical (randomised) version of SMP.

Second, this work contributes to understanding the power of the weakest commonly studied regime of quantum communication – SMP with quantum messages and *without shared randomness* (the latter restriction can be viewed as a somewhat artificial way of making the quantum model "as weak as possible"). Our function $\widetilde{cEq}\text{-}neg_T$ has an efficient solution in this regime as well, which means that even lacking shared randomness, quantum SMP can be exponentially stronger than its classical counterpart with shared randomness.

1 Introduction

Communication complexity is among the most interesting computational realms so far: Being one of the strongest where we can establish non-trivial (often tight) hardness statements – *lower bounds*, at the same time, it is one of the weakest that is capable to "accommodate" rather involved algorithms – *protocols.* As of today, communication complexity is one of the very few computational scenarios, where both upper and (non-speculative) lower bounds play central roles in the research.

We address two questions, related to the most basic communication complexity setting – the regime of *two parties*, solving a *functional problem.*

Two-way, one-way and SMP First, this works concludes the line of research, which can be viewed as demonstrating qualitative advantage of quantum communication over the

*Institute of Mathematics, Czech Academy of Sciences, Žitna 25, Praha 1, Czech Republic. Partially funded by the grant P202/12/G061 of GA ČR and by RVO: 67985840. Part of this work was done while visiting the Centre for Quantum Technologies at the National University of Singapore, and was partially supported by the National Research Foundation, Prime Minister's Office, Singapore and the Ministry of Education, Singapore under the Research Centres of Excellence programme and by Grant No. MOE2012-T3-1-009.

classical one in the three most commonly studied bipartite "layouts": *two-way (interactive) communication*; *one-way communication*; *simultaneous message passing (SMP).*

These models involve two players, *Alice* and *Bob*, who receive one "portion" of input each: Alice gets X and Bob gets Y. Their goal is to use the allowed type of communication (as determined by the "layout", see below) in order to compute the value of $f(X, Y)$, where f is a two-argument function defining the computational problem that the players have to solve.

- In the model of two-way communication the players can exchange messages, until one of them outputs the answer.
- In the one-way model Alice can send one message to Bob, who then produces the answer, based on this message and his portion of input.
- In the model of simultaneous message passing both Alice and Bob send one message each to the third participant – the *referee* – who has to produce the answer, based on these two messages only (unlike the players, the referee doesn't directly receive any portion of the input).

In all three regimes we say that a communication protocol *computes* a Boolean function f if for every pair (x, y) from the support of f, when the players receive $(X, Y) = (x, y)$, they output $f(x, y)$ with probability at least $2/3$. The participants are "all powerful" in terms of their local computational abilities, and the only resource considered for determining the *cost* of a protocol is the "amount of communication" that it consumes.

- When the communication model is *randomised*, the participants can send (classical) bits, the correctness condition must hold with respect to the random choices made by them and the complexity of a protocol is the (maximum) total number of bits sent during its execution.
- When the model is *quantum*, the participants can send qubits and perform arbitrary quantum measurements, the correctness condition must hold with respect to these quantum operations and the complexity of a protocol is the (maximum) total number of qubits sent during its execution.

It is known (and easy to see) that for virtually any type of communication "primitive" (i.e., classical randomised; classical deterministic; quantum; ...), the two-way layout is the most powerful, one-way is intermediate and SMP is the weakest.

Demonstrating advantage of quantum over classical communication in a weaker regime (say, one-way) could – in principle – turn out to be either less or more challenging than in a stronger one (say, interactive): While in the latter case one would have to prove a *stronger* lower bound, at the same time the communication problem being used for the separation would likely be *harder*, and therefore easier to prove a lower bound for.

The history of research seems to suggest that separating models on the "lower levels" – namely, one-way communication, and even more so SMP,– is more challenging than under the stronger setting of interactive communication. In 1999 Raz [Raz99] demonstrated a *function* that had an efficient[1] *quantum two-way protocol*, but no efficient *classical two-way protocol.* In 2004 Bar-Yossef, Jayram and Kerenidis [BYJK04] demonstrated a *relation* that had an efficient *quantum one-way protocol*, but no efficient *classical one-way protocol.* Note that

[1] We call *efficient* communication protocols of poly-logarithmic complexity.

the original separation from [Raz99] was demonstrated via a functional problem; on the other hand, the result of [BYJK04] used a relation – a more general class of problems and a *stronger* model-separating tool.[2] In the same work it has been asked whether it was possible to demonstrate similar qualitative advantage of quantum one-way communication via a functional problem, which was answered affirmatively in 2008 in a joint work with Kempe, Kerenidis, Raz and de Wolf [GKK+08].

The work [BYJK04] has also demonstrated a *relation* that had an efficient *quantum SMP protocol*, but no efficient *classical SMP protocol*, and – similarly to the one-way case – it has been left open whether there existed a functional problem, easy for quantum and hard for classical SMP.

In the meanwhile, separations "against classical two-way" have been strengthened in a sequence of works [Gav08, KR11, Gav16] that presented "super-separations": e.g., in 2010 Klartag and Regev demonstrated a *function* with an efficient *quantum one-way protocol*, but no efficient *classical two-way protocol.* On the other hand, it has remained open till now whether a function could witness quantum superiority in the case of SMP.[3]

This work answers that question and demonstrates a functional problem $\widetilde{cEq}\text{-}neg_T$, whose communication complexity is $O\big((\log n)^2\big)$ in the quantum version of SMP and $\tilde{\Omega}(\sqrt{n})$ in the classical (randomised) version of SMP.

Weakening the weak: SMP without shared randomness The second aspect of this work is related to understanding the power of, arguably, the weakest commonly studied regime of quantum communication – SMP with quantum messages and *without shared randomness.*

We will write $\mathcal{Q}^{\|}$ and $\mathcal{R}^{\|}$ for, respectively, the quantum and the classical version of SMP without shared randomness. To denote the corresponding standard counterparts – those equipped with (unlimited) shared randomness – we will write, respectively, $\mathcal{Q}^{\|,pub}$ and $\mathcal{R}^{\|,pub}$.

Both $\mathcal{Q}^{\|}$ and $\mathcal{R}^{\|}$ (i.e., the versions lacking shared randomness) can be viewed as "purposely weakened", somewhat artificial versions of SMP – as opposed to the standard $\mathcal{Q}^{\|,pub}$ and $\mathcal{R}^{\|,pub}$.[4] The families of efficiently-computable tasks in $\mathcal{Q}^{\|}$ and in $\mathcal{R}^{\|}$ are not closed with respect to mixed strategies[5], and the usual minimax principle does not hold for these models: for example, the *equality function (Eq)* has $\mathcal{R}^{\|}$-complexity $O(1)$ over any fixed input distribution, but its worst-case $\mathcal{R}^{\|}$-complexity is $\Omega(\sqrt{n})$, due to [NS96].

Von Neumann, who proved the minimax principle for the case of 2-player zero-sum games with mixed strategies in 1928, later remarked: "*As far as I can see, there could be no theory*

[2] There are known cases where a quantum communication model can be separated from a classical one via a relation, but a functional separation is provably impossible (see [Aar04, GRdW08]). In particular, [GRdW08] showed that the model of "quantum-classical SMP" – the regime where Alice could send a quantum message but Bob was classical (or vice versa) – was equal to the "fully classical" SMP for functional problems; on the other hand, a relational separation between these two models followed from [BYJK04]. As in this work we are concerned with functional separations, for us the model of SMP with both quantum players is the weakest (non-trivial) regime of quantum communication.

[3] The result in [Gav16] implied existence of a function, hard for classical SMP (and even for classical two-way protocols), but easy for the model of *quantum SMP with shared entanglement* – a significantly strengthened version of quantum SMP, where the players could share an arbitrary (input-independent) quantum state of finite dimension.

[4] Note that in the context of "Two-way, one-way and SMP" we only referred to the "natural" models $\mathcal{Q}^{\|,pub}$ and $\mathcal{R}^{\|,pub}$.

[5] $\mathcal{R}^{\|,pub}$ – the "unrestricted" randomised SMP – can be defined as the "closure" of $\mathcal{R}^{\|}$ with respect to mixed strategies, and similarly for $\mathcal{Q}^{\|,pub}$ and $\mathcal{Q}^{\|}$.

of games [...] without that theorem." The question of determining the complexity of a given communication problem can be phrased in the language of 2-player zero-sum games, and the case of SMP without shared randomness is probably the only commonly studied one that goes "without that theorem". Although we have seen some non-trivial results both in $\mathcal{Q}^{\|}$ and in $\mathcal{R}^{\|}$, these models still lack the aesthetic appeal and the cognitive depth of those obeying the minimax principle.

So, the model of SMP with quantum messages and without shared randomness ($\mathcal{Q}^{\|}$) indeed can be viewed as the weakest commonly studied quantum model in communication complexity. Prior to this work, $\mathcal{Q}^{\|}$ was known to be stronger than $\mathcal{R}^{\|}$: in 2001 Buhrman, Cleve, Watrous and de Wolf [BCWdW01] demonstrated that there existed a $\mathcal{Q}^{\|}$-protocol for the function Eq of complexity $O(\log n)$; as we already mentioned, it had been known that $\mathcal{R}^{\|}(Eq) \in \Omega(\sqrt{n})$. Till now it has remained open whether $\mathcal{Q}^{\|}$ was capable to do more than that – in particular, to solve efficiently any problem, hard for the "natural closure" of $\mathcal{R}^{\|}$, namely $\mathcal{R}^{\|,pub}$.

We show that the main communication problem studied in this work – the function $\widetilde{cEq}\text{-}neg_T$ – has an efficient protocol in $\mathcal{Q}^{\|}$ as well. Due to the same lower bound of $\tilde{\Omega}(\sqrt{n})$ on its $\mathcal{R}^{\|,pub}$-complexity, this demonstrates exponential advantage of $\mathcal{Q}^{\|}$ over $\mathcal{R}^{\|,pub}$ in solving a functional problem.

Our method The questions considered in this work have remained open for some time. We are proving $\mathcal{R}^{\|,pub}$-hardness of a communication problem that is easy for virtually any model stronger than $\mathcal{R}^{\|,pub}$ – so, the argument had to be "tuned" rather accurately in order to distinguish between $\mathcal{R}^{\|,pub}$ and "anything above it". The hardness of the communication problem also had to be "tuned" with some precision, as it had to be easy for $\mathcal{Q}^{\|}$ and hard for a "slightly weaker"[6] model $\mathcal{R}^{\|,pub}$ (in particular we could not use a problem with "worst-case hardness in spite of average-case easiness", like Eq, as $\mathcal{R}^{\|,pub}$ allowed mixed strategies).

It may be for these reasons that this work is built around several ad hoc ideas. Some of them will be informally discussed in Section 3.

2 Preliminaries

For $x \in \{0,1\}^n$ and $i \in [n] = \{1,\ldots,n\}$, we will write x_i or $x(i)$ to address the i'th bit of x (preferring "x_i" unless it may cause ambiguity). Similarly, for $S \subseteq [n]$, let both x_S and $x(S)$ denote the $|S|$-bit string, consisting of (naturally-ordered) bits of x, whose indices are in S. For a set (or a family) A, we will write $A|_i$ and $A|_S$ to address, respectively, $\left\{x_i \middle| x \in A\right\}$ and $\left\{x_S \middle| x \in A\right\}$. We will use similar notation in all cases when x can be viewed naturally as an element of $\mathcal{X}_1 \times \cdots \times \mathcal{X}_n$.

For $x, y \in \{0,1\}^n$, let $|x|$ denote the Hamming weight of x and $x \oplus y$ denote the bit-wise XOR operation.

For a (discrete) set A and $k \in \mathbb{N}$, we denote by $\mathrm{Pow}(A)$ the set of A's subsets and by $\binom{A}{k}$ the set $\left\{a \in \mathrm{Pow}(A) \middle| |a| = k\right\}$. We write "$A \triangle B$" to denote the symmetric difference between the two sets and "$A \mathbin{\dot\cup} B$" to denote the union of the two sets that must be disjoint (i.e., using this notation implies that $A \cap B = \emptyset$).

[6] In fact, incomparable: there are known examples where $\mathcal{R}^{\|,pub}$ is exponentially stronger than $\mathcal{Q}^{\|}$ for relational problems, see [GKRdW06].

We write $\mathcal{U}_A$ to denote the uniform distribution over the elements of A. Sometimes (e.g., in subscripts) we will write "$\in A$" instead of "$\sim \mathcal{U}_A$". We will sometimes emphasise that a distribution on $\{0,1\}^{2n}$ is "viewed as bipartite" (i.e., assumed to be the joint distribution of two random variables, containing n bits each) by addressing it as a *distribution on* $\{0,1\}^{n+n}$; similarly, we will write "$(X,Y) \in \{0,1\}^{n+n}$", etc.

Let S_n denote the group of permutations of $[n]$, and let $\sigma_i \in S_n$ be the i'th cyclic shift (i.e., $\sigma_i(j) = i+j$ if $i+j \le n$ and $i+j-n$ otherwise). For $x \in \{0,1\}^n$ and $\tau \in S_n$, denote by $\tau(x)$ the element of $\{0,1\}^n$, whose $\tau(i)$'th position contains x_i for each i – in particular, $\sigma_j(x)$ is the j-bit cyclic shift of x.

For functions $f, g : \{0,1\}^n \to \mathbb{R}$, we define

$$\langle f\,,\,g\rangle \stackrel{\text{def}}{=} 2^{-n} \cdot \sum_{x\in\{0,1\}^n} f(x)\cdot g(x) = \mathop{\mathbf{E}}_{X \in \{0,1\}^n}[f(X)\cdot g(X)]$$

and $\|f\|_2 \stackrel{\text{def}}{=} \sqrt{\langle f\,,\,f\rangle}$. For $s \subseteq [n]$ and $x \in \{0,1\}^n$, let $\chi_s(x) \stackrel{\text{def}}{=} (-1)^{|x_s|}$ and $\hat{f}(s) \stackrel{\text{def}}{=} \langle f\,,\,\chi_s\rangle$. The *Fourier transform* $f \to \hat{f}$ is a *norm-preserving* linear mapping: $\|f\|_2^2 = \sum_s \hat{f}(s)^2$ (*Parseval's identity*). The vectors χ_s form an orthonormal basis of $\mathbb{R}^{2^n}$ and

$$f(x) = \sum_{s\subseteq[n]} \hat{f}(s)\cdot\chi_s(x)$$

for every $x \in \{0,1\}^n$.

Definition 1 (*small-bias spaces*)**.** *For $\varepsilon \ge 0$, we call $T \subseteq \{0,1\}^n$ an ε-bias space if*

$$\left|\mathop{\mathbf{E}}_{\tau\in T}[\chi_s(\tau)]\right| \le \varepsilon$$

for every $s \subseteq [n]$, $s \ne \emptyset$.

Being a small-bias space is a "pseudorandom property": it holds for random subsets of $\{0,1\}^n$ almost always, and there are efficient constructions.

Fact 1 *([NN93]). For $\varepsilon > 0$, an ε-bias space can be constructed deterministically in time* $\mathrm{poly}(n/\varepsilon)$. *Every pair of elements $\tau_1 \ne \tau_2$ of the constructed space satisfies $|\tau_1 \oplus \tau_2| \in \frac{n}{2} \pm o(n)$.*

Communication complexity

For an excellent survey of classical communication complexity, see [KN97]. Quantum communication models differ from their classical counterparts in two aspects: the players are allowed to send quantum messages (accordingly, the complexity is measured in *qubits*) and to perform arbitrary quantum operations locally.

Of central importance to this work is the model of *simultaneous message passing (SMP)*, where there are 3 participants: *players* Alice and Bob, and *the referee*. An SMP-protocol for computing a Boolean function $f(X,Y)$ has the following structure: Alice receives X and sends her message to the referee; at the same time, Bob receives Y and send his message to the referee; the referee uses the content of the two received messages to compute the answer. The answer is correct when it equals $f(X,Y)$ (the input is always such that $f(X,Y)$ is defined). We will consider the following variations of SMP:

1. In $\mathcal{D}^{\|}_{\mu,\varepsilon}$ (sometimes written as $\mathcal{D}^{\|}_{\varepsilon}$ if μ is irrelevant or clear from the context) the players and the referee are *deterministic*, and the answer must be correct with probability at least $1-\varepsilon$ when $(X,Y)\sim\mu$.[7]

2. In $\mathcal{R}^{\|}$ the players and the referee can use *local randomness*, and the answer must be correct with probability at least $2/3$ for every valid input.

3. $\mathcal{R}^{\|,pub}$ is similar to $\mathcal{R}^{\|}$, but the players and the referee can use *shared randomness*.

4. In $\mathcal{Q}^{\|}$ the players can send *quantum* messages and the referee can apply any quantum measurement to compute the answer, which must be correct with probability at least $2/3$ for every valid input.

Variations of equality

The communication problem that we use for our separation is a function that can be viewed as a variation of the *equality* problem.

The *equality function* (viewed as a communication problem) is the following total[8] bipartite function. Let $u\subseteq[n]$ (for technical reasons, we consider a "projected version" of equality), then

$$Eq_u : \{0,1\}^{n+n}\to\{0,1\},$$
$$Eq_u(x,y)\stackrel{\text{def}}{=}\begin{cases}1 & \text{if } x_u=y_u;\\ 0 & \text{otherwise.}\end{cases}$$

We write Eq for $Eq_{[n]}$. Define input distributions for Eq_u:

- for $a\in\{0,1\}$, let $\mu^a_{Eq_u}$ be the uniform distribution over $Eq_u^{-1}(a)$;
- let $\mu_{Eq_u}\stackrel{\text{def}}{=}\frac{1}{2}\cdot\left(\mu^0_{Eq_u}+\mu^1_{Eq_u}\right)$.

The next problem intuitively corresponds to asking whether $Eq_u(X\oplus\tau,Y)=1$ for some τ from a predetermined set $T\subseteq\{0,1\}^n$, usually of size $\mathrm{poly}(n)$ (in our analysis T will be a small-bias space).

$$Eq_u\text{-}neg_T : \{0,1\}^{n+n}\to\{0,1\},$$
$$Eq_u\text{-}neg_T(x,y)\stackrel{\text{def}}{=}\begin{cases}1 & \text{if } (x\oplus\tau)_u=y_u \text{ for some } \tau\in T;\\ 0 & \text{otherwise.}\end{cases}$$

Define input distributions for $Eq_u\text{-}neg_T$:

- for $\tau\in T$, let $\mu^\tau_{Eq_u\text{-}neg_T}$ be the distribution of (X,Y) when $(X\oplus\tau,Y)\sim\mu_{Eq_u}$;
- let $\mu_{Eq_u\text{-}neg_T}\stackrel{\text{def}}{=}\frac{1}{|T|}\cdot\sum_{\tau\in T}\mu^\tau_{Eq_u\text{-}neg_T}$.

[7] In this work we will only deal with binary-valued functions; accordingly, we always assume that $\varepsilon<1/2$.

[8] A functional problem in communication complexity is called *total* when it is supported on the product set of the players' individual sets of input.

Next we define a "noisy" (or gapped) version of $\mathit{Eq}\text{-}\mathit{neg}_T$:

$$\widetilde{\mathit{Eq}}\text{-}\mathit{neg}_T : \{0,1\}^{n+n} \to \{0,1\},$$

$$\widetilde{\mathit{Eq}}\text{-}\mathit{neg}_T(x,y) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } |x \oplus y \oplus \tau| \le \frac{6n}{15} \text{ for some } \tau \in T \\ & \quad \text{and } |x \oplus y \oplus \tau| \notin (\frac{6n}{15}, \frac{7n}{15}) \text{ for every } \tau; \\ 0 & \text{if } |x \oplus y \oplus \tau| \ge \frac{7n}{15} \text{ for every } \tau; \\ \text{undefined} & \text{otherwise.} \end{cases}$$

Intuitively, $\widetilde{\mathit{Eq}}\text{-}\mathit{neg}_T(x,y)$ "asks" whether $x+\tau$ is close to y with respect to one of the "permitted" bit-negations $\tau \in T$. The *promise* is that $x+\tau$ must be either far enough from y (at distance $\ge \frac{7n}{15}$) or close to it (at distance $\le \frac{6n}{15}$) for every $\tau \in T$ – otherwise the function is undefined.

Define input distributions for $\widetilde{\mathit{Eq}}\text{-}\mathit{neg}_T$:

- let $\mu_{\widetilde{\mathit{Eq}}\text{-}\mathit{neg}_T} \stackrel{\text{def}}{=} \frac{1}{\binom{n}{n/3}} \cdot \sum_{u \in \binom{[n]}{n/3}} \mu_{\mathit{Eq}_u\text{-}\mathit{neg}_T}$.

In our construction T will be a set, where $|\tau_1 \oplus \tau_2| \in \frac{n}{2} \pm o(n)$ for every $\tau_1 \ne \tau_2 \in T$ (cf. Fact 1). Note that in such case a pair $(X,Y) \sim \mu_{\widetilde{\mathit{Eq}}\text{-}\mathit{neg}_T}$ satisfies the promise of $\widetilde{\mathit{Eq}}\text{-}\mathit{neg}_T(X,Y)$ with probability $1 - 2^{-\Omega(n)}$. [9]

We are ready to introduce the main communication problem considered in this work – a function that can be viewed as a "cyclic version" of $\mathit{Eq}_u\text{-}\mathit{neg}_T$:

$$\widetilde{c\mathit{Eq}}\text{-}\mathit{neg}_T : \{0,1\}^{n+n} \to \{0,1\},$$

$$\widetilde{c\mathit{Eq}}\text{-}\mathit{neg}_T(x,y) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } |\sigma_j(x) \oplus y \oplus \tau| \le \frac{6n}{15} \text{ for some } \tau \in T \text{ and } j \in [n] \\ & \quad \text{and } |\sigma_j(x) \oplus y \oplus \tau| \notin (\frac{6n}{15}, \frac{7n}{15}) \text{ for every } \tau \text{ and } j; \\ 0 & \text{if } |\sigma_j(x) \oplus y \oplus \tau| \ge \frac{7n}{15} \text{ for every } \tau \text{ and } j; \\ \text{undefined} & \text{otherwise.} \end{cases}$$

The intuition behind this definition is very similar to that behind $\widetilde{\mathit{Eq}}\text{-}\mathit{neg}_T(x,y)$, but "the question" here is whether $\sigma_j(x) + \tau \approx y$ with respect to some cyclic shift σ_j and one of the bit-negations $\tau \in T$.

Define input distributions for $\widetilde{c\mathit{Eq}}\text{-}\mathit{neg}_T$:

- for $j \in [n]$, let $\mu^j_{\widetilde{c\mathit{Eq}}\text{-}\mathit{neg}_T}$ be the distribution of (X,Y) when $(\sigma_j(X), Y) \sim \mu_{\widetilde{\mathit{Eq}}\text{-}\mathit{neg}_T}$;
- let $\mu_{\widetilde{c\mathit{Eq}}\text{-}\mathit{neg}_T} \stackrel{\text{def}}{=} \frac{1}{n} \cdot \sum_{j \in [n]} \mu^j_{\widetilde{c\mathit{Eq}}\text{-}\mathit{neg}_T}$.

Like in the case of $\widetilde{\mathit{Eq}}\text{-}\mathit{neg}_T$, in our constructions T will always be a set with the "minimal distance property". In such case a pair $(X,Y) \sim \mu_{\widetilde{c\mathit{Eq}}\text{-}\mathit{neg}_T}$ satisfies the promise of $\widetilde{c\mathit{Eq}}\text{-}\mathit{neg}_T(X,Y)$ with probability $1 - 2^{-\Omega(n)}$. [10]

---

[9] In particular, under $(X,Y) \sim \mu_{\widetilde{\mathit{Eq}}\text{-}\mathit{neg}_T}$ conditioned upon $[|X \oplus Y \oplus \tau_0| \le \frac{6n}{15}]$ for some $\tau_0 \in T$, it holds that $|X \oplus Y \oplus \tau| \ge \frac{7n}{15}$ for every $\tau \ne \tau_0$ with probability $1 - 2^{-\Omega(n)}$.

[10] In particular, under $(X,Y) \sim \mu_{\widetilde{c\mathit{Eq}}\text{-}\mathit{neg}_T}$ conditioned upon $[|\sigma_{j_0}(X) \oplus Y \oplus \tau_0| \le \frac{6n}{15}]$ for some $j_0 \in [n]$ and $\tau_0 \in T$, it holds that $|\sigma_j(X) \oplus Y \oplus \tau| \ge \frac{7n}{15}$ for every $j \ne j_0$ and $\tau \ne \tau_0$ with probability $1 - 2^{-\Omega(n)}$.

3 Intuition behind the new separation

Recall that we are looking for a functional communication problem, easy for quantum but hard for classical SMP (naturally, equipped with shared randomness). The initial inspiration comes from the observation that the most obvious quantum SMP protocol for *equality with gap ($\widetilde{Eq}$)* has certain "robustness" that seems impossible to achieve in a classical protocol.

Let

$$\widetilde{Eq}(x,y) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } |x \oplus y| \le \frac{n}{5}; \\ 0 & \text{if } |x \oplus y| \ge \frac{2n}{5}; \\ \text{undefined} & \text{otherwise.} \end{cases}$$

A natural $\mathcal{Q}^{\|}$-solution to this problem would be for Alice to send $\sum_i |i\rangle|X_i\rangle$, for Bob to send $\sum_i |i\rangle|Y_i\rangle$ and for the referee to perform the swap test, thus estimating the inner product between the two messages: The probability of "passing" the test is $\frac{1}{2} + \frac{|X \oplus Y|}{2n}$, so estimating it with constant precision allows the referee to give the correct answer with constant-bounded error, thus solving the problem.[11]

Note that the same pair of messages sent by the players can be used by the referee for solving

$$\widetilde{Eq}(\pi(X) \oplus \tau, Y)$$

for *any* $\pi \in S_n$ and $\tau \in \{0,1\}^n$: Upon receiving the messages and before performing the swap test, the referee would have to apply the obvious unitary transformation to the message from Alice (namely, permuting the indices and negating some bit values).

Let $S \subset S_n$, $T \subset \{0,1\}^n$ and $|S|, |T| \in \text{poly}(n)$. Using the above intuition, we conclude that there exists an efficient quantum protocol for the problem

$$\widetilde{Eq}_{S,T}(x,y) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } |\pi(x) \oplus \tau \oplus y| \le \frac{n}{5} \text{ for some } \pi \in S \text{ and } \tau \in T; \\ 0 & \text{if } |\pi(x) \oplus \tau \oplus y| \ge \frac{2n}{5} \text{ for every } \pi \in S \text{ and } \tau \in T; \\ \text{undefined} & \text{otherwise.} \end{cases}$$

To solve it in $\mathcal{Q}^{\|}$, both Alice and Bob send $O(\log n)$ copies of their messages from the $\widetilde{Eq}$-protocol described above, which allows the referee to solve any instance of $\widetilde{Eq}(\pi(X) \oplus \tau, Y)$ with error $1/poly(n)$ (arbitrarily small). In particular, this means that he can "reuse" the messages and test $\widetilde{Eq}(\pi'(X) \oplus \tau', Y)$ for every $\pi' \in S$ and $\tau' \in T$ with polynomially-small error, thus solving the problem.

One can see that the main communication task studied here – $\widetilde{cEq}\text{-}neg_T$ – is an instance of $\widetilde{Eq}_{S,T}$ with different constants, S being the set of cyclic bit-shifts and T being a small-bias space.

3.1 Roadway to the lower bound

Ignoring some technical details, our lower-bound argument "against $\mathcal{R}^{\|,pub}$" can be informally outlined as follows.

[11] For simplicity, in this informal overview we do not "normalise" quantum states and only require that a protocol solves a Boolean problem with error $1/2 - \Omega(1)$. The definition that we are making in Section 3 are not used elsewhere.

First of all, we need a convenient characterisation of efficient protocols for $\widetilde{Eq}$. It will be based on the observation that if random input satisfying $X \approx Y$ is given to an $\mathcal{R}^{\|,pub}$-protocol for $\widetilde{Eq}(X,Y)$, then the two messages received by the referee are likely to "witness" that fact. After some technical manipulations, this idea will lead to

$$\mathop{\mathbf{E}}_{i}\left[\Delta_{\alpha}^{i}\cdot\Delta_{\beta}^{i}\right]\in\Omega\left(\frac{1}{n}\right), \tag{1}$$

where Δ_{α}^{i} is the "bias" of the referee's knowledge about X_i, gained from Alice's message $Al(X)$, and Δ_{β}^{i} is defined similarly with respect to Y and Bob's message $Bo(Y)$.

Next we take T into account. We will use its small-bias properties to conclude that a protocol for $\widetilde{Eq}\text{-}neg_T(X,Y)$ must satisfy

$$\mathop{\mathbf{E}}_{i}\left[\mathbf{I}\left[X_i : Al(X)\right]\cdot\mathbf{I}\left[Y_i : Bo(Y)\right]\right]\in\Omega\left(\frac{1}{n}\right). \tag{2}$$

The bound in (2) is significantly stronger than that in (1): Both X_i and Y_i are uniformly-random bits, so "bias" $\gamma > 0$ in the referee's knowledge, say, about X_i corresponds to $\Theta\left(\gamma^2\right)$ bits of information. The "quadratic improvement" from (1) to (2) captures the "added hardness" in the transition from $\widetilde{Eq}$ to $\widetilde{Eq}\text{-}neg_T$ – at least, from the point of view of our analysis.

Finally, we add cyclic shifts in order to "disconnect" $\mathbf{I}\left[X_i : Al(X)\right]$ from $\mathbf{I}\left[Y_i : Bo(Y)\right]$. We will show that any protocol for $\widetilde{cEq}\text{-}neg_T(X,Y)$ must satisfy

$$\mathop{\mathbf{E}}_{i}\left[\mathbf{I}\left[X_i : Al(X)\right]\right]\cdot\mathop{\mathbf{E}}_{j}\left[\mathbf{I}\left[Y_j : Bo(Y)\right]\right]\in\Omega\left(\frac{1}{n}\right), \tag{3}$$

and this gives the desired lower bound, as at least one of $\mathbf{E}_i\left[\mathbf{I}\left[X_i : Al(X)\right]\right]$ and $\mathbf{E}_j\left[\mathbf{I}\left[Y_j : Bo(Y)\right]\right]$ must be $\Omega(1/\sqrt{n})$ in order to satisfy (3).

4 The $\mathcal{Q}^{\|}$-complexity of $\widetilde{cEq}\text{-}\boldsymbol{neg}_T$ – an upper bound

Let $\tau_0 \in T$, $j_0 \in [n]$ and consider the following protocol:

- Alice sends $|\phi_{Al}\rangle = \frac{1}{\sqrt{n}}\cdot\sum_{i=1}^{n}|i\rangle|X_i\rangle$;
- Bob sends $|\phi_{Bo}\rangle = \frac{1}{\sqrt{n}}\cdot\sum_{i=1}^{n}|i\rangle|Y_i\rangle$;
- the referee transforms $|\phi_{Al}\rangle$ to

$$\left|\phi'_{Al}\right\rangle = \frac{1}{\sqrt{n}}\cdot\sum_{i=1}^{n}|\sigma_{j_0}(i)\rangle|X_i\oplus\tau_0(\sigma_{j_0}(i))\rangle;$$

- the referee applies the swap test to the received messages and outputs "1" with probability

$$\frac{1+|\langle\phi'_{Al}|\phi_{Bo}\rangle|^2}{2}=\frac{1}{2}+\frac{|\sigma_{j_0}(X)\oplus\tau_0\oplus Y|}{2n}.$$

Running this protocol once allows to estimate $\frac{|\sigma_{j_0}(X)\oplus\tau_0\oplus Y|}{n}$ with constant expected accuracy; running it in parallel $O(1)$ times allows estimating that value with arbitrarily small constant expected accuracy; running it in parallel $O(\log n)$ times allows estimating with arbitrarily small constant accuracy and arbitrarily high confidence $1 - {}^1\!/\!_{\mathrm{poly}(n)}$. Accordingly, $O(\log n)$ parallel copies of the above protocol can distinguish between $|\sigma_{j_0}(X) \oplus Y \oplus \tau_0| \le \frac{6n}{15}$ and $|\sigma_{j_0}(X) \oplus Y \oplus \tau_0| \ge \frac{7n}{15}$ with arbitrary small error ${}^1\!/\!_{\mathrm{poly}(n)}$.

As only the action of the referee depends on τ_0 and j_0, the players' messages can be "reused" for distinguishing between $|\sigma_j(X) \oplus Y \oplus \tau| \le \frac{6n}{15}$ and $|\sigma_j(X) \oplus Y \oplus \tau| \ge \frac{7n}{15}$ for every $\tau \in T$ and $j \in [n]$ (note that this gap is always guaranteed by the definition, for both $\widetilde{cEq}\text{-}neg_T^{-1}(0)$ and $\widetilde{cEq}\text{-}neg_T^{-1}(1)$). To achieve that, the referee performs every single estimation with sufficiently small error ${}^1\!/\!_{\mathrm{poly}(n \cdot |T|)}$.[12] Therefore, $O(\log n + \log |T|)$ parallel copies of the above protocol can solve $\widetilde{cEq}\text{-}neg_T(X, Y)$ in $\mathcal{Q}^{\|}$.

Corollary 1. *For every $T \subseteq \{0,1\}^n$,*

$$\mathcal{Q}^{\|}(\widetilde{cEq}\text{-}neg_T) \in O\big((\log n)^2 + \log n \cdot \log |T|\big).$$

5 A probabilistic interlude

Here we prove several claims addressing the behaviour of non-independent random variables. The statements are rather intuitive, though we are not aware of previously published proofs.

5.1 Optimistic inequalities

Claim 1 (*Optimistic chain inequality*)**.** *Let $X_1, \ldots, X_m$ be random variables, where each X_i is supported on (finite) $G_i \dot\cup B_i$. Let μ denote the joint distribution of $X = (X_1, \ldots, X_m)$, then*

$$\begin{aligned} \Pr_{X\sim\mu}\left[\bigwedge_{j=1}^{m} X_j \in G_j\right] &= \prod_{i=1}^{m} \Pr_{X\sim\mu}\left[X_i \in G_i \,\middle|\, \bigwedge_{j=1}^{i-1} X_j \in G_j\right] \\ &= \prod_{i=1}^{m} \mathop{\mathbf{E}}_{X'\sim\mu}\left[\Pr_{X\sim\mu}\left[X_i \in G_i \,\middle|\, \bigwedge_{j=1}^{i-1} X_j = X'_j\right] \,\middle|\, \bigwedge_{j=1}^{\mathbf{i-1}} X'_j \in G_j\right] \\ &\le \prod_{i=1}^{m} \mathop{\mathbf{E}}_{X'\sim\mu}\left[\Pr_{X\sim\mu}\left[X_i \in G_i \,\middle|\, \bigwedge_{j=1}^{i-1} X_j = X'_j\right] \,\middle|\, \bigwedge_{j=1}^{\mathbf{m}} X'_j \in G_j\right], \end{aligned} \tag{4}$$

where $X' = (X'_1, \ldots, X'_m)$ and X are independent from one another, unless conditioned explicitly. Moreover,

$$\begin{aligned} &\log\left(\Pr_{X\sim\mu}\left[\bigwedge_{j=1}^{m} X_j \in G_j\right]\right) \\ &\qquad \le \sum_{i=1}^{m} \mathop{\mathbf{E}}_{X'\sim\mu}\left[\log\left(\Pr_{X\sim\mu}\left[X_i \in G_i \,\middle|\, \bigwedge_{j=1}^{i-1} X_j = X'_j\right]\right) \,\middle|\, \bigwedge_{j=1}^{m} X'_j \in G_j\right]. \end{aligned} \tag{5}$$

The equalities in (4) correspond to the standard "chain" decomposition (included here for convenience). In comparison to the standard decomposition, the inequality offers more symmetric upper bound on $\mathbf{Pr}[\wedge X_j \in G_j]$ at the sacrifice of tightness.

[12] E.g., see Lemma 2 in [Aar04].

We call the above claim *optimistic*, viewing the subsets G_i as *good*, B_i-s as *bad* and interpreting the statement of (4) as saying that *the estimated probability of m good outcomes doesn't decrease as a result of making the estimation "optimistically biased"*: instead of conditioning the expectation on $\left[\bigwedge_{j=1}^{i-1} X'_j \in G_j\right]$ (which would give the actual probability of all good outcomes), the right-hand side of the above inequality uses more "good-oriented" (and more restricting) condition $\left[\bigwedge_{j=1}^{m} X'_j \in G_j\right]$.

Moreover, the right-hand side of (5) *is likely to have grown as a result of making the expectations "optimistically biased"*: due to the strict concavity of log, the statement wouldn't hold if the condition $\left[\bigwedge_{j=1}^{m} X'_j \in G_j\right]$ were replaced by $\left[\bigwedge_{j=1}^{i-1} X'_j \in G_j\right]$, unless the quantities under the expectations are constant (that is, unless every event $[X_i \in G_i]$ is independent from the values of $X_1, \ldots, X_{i-1}$, subject to $\left[\bigwedge_{j=1}^{i-1} X_j \in G_j\right]$).

Note also that the inequality in (4) isn't, in general, true "element-wise": one can construct an example, where for some $i_0 \in [m]$:

$$\Pr_{X\sim\mu}\left[X_{i_0} \in G_{i_0} \middle| \bigwedge_{j=1}^{i_0-1} X_j \in G_j\right] = \mathop{\mathbf{E}}_{X'\sim\mu}\left[\Pr_{X\sim\mu}\left[X_{i_0} \in G_{i_0} \middle| \bigwedge_{j=1}^{i_0-1} X_j = X'_j\right] \middle| \bigwedge_{j=1}^{i_0-1} X'_j \in G_j\right]$$
$$> \mathop{\mathbf{E}}_{X'\sim\mu}\left[\Pr_{X\sim\mu}\left[X_{i_0} \in G_{i_0} \middle| \bigwedge_{j=1}^{i_0-1} X_j = X'_j\right] \middle| \bigwedge_{j=1}^{m} X'_j \in G_j\right].$$

The following statement implies that one must take $i_0 < m$ for the above inequality to hold.

Claim 2 (*Optimistic conditioning*)**.** *Let X_1 and X_2 be random variables, where X_i is supported on (finite) $G_i \uplus B_i$ and μ is the joint distribution of $X = (X_1, X_2)$, then*

$$\log\left(\Pr_{X\sim\mu}\left[X_2 \in G_2 \middle| X_1 \in G_1\right]\right) = \log\left(\mathop{\mathbf{E}}_{X'\sim\mu}\left[\Pr_{X\sim\mu}\left[X_2 \in G_2 \middle| X_1 = X'_1\right] \middle| X'_1 \in G_1\right]\right)$$
$$= \mathop{\mathbf{E}}_{X'\sim\mu}\left[\log\left(\Pr_{X\sim\mu}\left[X_2 \in G_2 \middle| X_1 = X'_1\right]\right) \middle| X'_1 \in G_1, X'_2 \in G_2\right] - d_{KL}(\beta \| \alpha)$$
$$\leq \mathop{\mathbf{E}}_{X'\sim\mu}\left[\log\left(\Pr_{X\sim\mu}\left[X_2 \in G_2 \middle| X_1 = X'_1\right]\right) \middle| X'_1 \in G_1, X'_2 \in G_2\right],$$

where $X' = (X'_1, X'_2)$ is independent from X (unless conditioned explicitly), and α and β denote the distributions of X_1, conditioned, respectively, on $[X_1 \in G_1]$ and on $[X_1 \in G_1, X_2 \in G_2]$.

The statement of Claim 2, similarly to (5), witnesses the qualitative "benefit" of optimistic conditioning: since log is strictly concave, whenever $[X_2 \in G_2]$ is not independent from X_1 (subject to $[X_1 \in G_1]$), the above inequality wouldn't hold if the expectation were not a subject to $[X'_2 \in G_2]$.

Proof of Claim 2. For every $c \in G_1$, let $p_c \stackrel{\text{def}}{=} \mathbf{Pr}[X_1 = c]$ and $q_c \stackrel{\text{def}}{=} \mathbf{Pr}\left[X_2 \in G_2 \middle| X_1 = c\right]$.

Then

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{X\sim\mu}[X_1 \in G_1] &= \sum_{d\in G_1} p_d,\\
\alpha(c) = \mathop{\mathbf{Pr}}_{X\sim\mu}\big[X_1 = c \big| X_1 \in G_1\big] &= \frac{p_c}{\sum_{d\in G_1} p_d},
\end{aligned}$$

and

$$\begin{aligned}
&\mathop{\mathbf{Pr}}_{X\sim\mu}[X_1 = c, X_2 \in G_2] = p_c q_c,\\
&\mathop{\mathbf{Pr}}_{X\sim\mu}[X_1 \in G_1, X_2 \in G_2] = \sum_{d\in G_1} p_d q_d,\\
&\beta(c) = \mathop{\mathbf{Pr}}_{X\sim\mu}\big[X_1 = c \big| X_1 \in G_1, X_2 \in G_2\big] = \frac{p_c q_c}{\sum_{d\in G_1} p_d q_d}.
\end{aligned}$$

Let

$$k \stackrel{\text{def}}{=} \frac{\sum_{d\in G_1} p_d q_d}{\sum_{d\in G_1} p_d} \equiv \frac{\alpha(d)}{\beta(d)} \cdot q_d,$$

then

$$\begin{aligned}
\log\Big(\mathop{\mathbf{Pr}}_{X\sim\mu}\big[X_2 \in G_2 \big| X_1 \in G_1\big]\Big) &= \log\left(\sum_{d\in G_1} \alpha(d)\cdot q_d\right) = \log\left(\sum_{d\in G_1} k\cdot\beta(d)\right) = \log(k)\\
&= \sum_{d\in G_1} \beta(d)\cdot\log\left(k\cdot\frac{\beta(d)}{\alpha(d)}\right) - \sum_{d\in G_1}\beta(d)\cdot\log\left(\frac{\beta(d)}{\alpha(d)}\right)\\
&= \sum_{d\in G_1} \beta(d)\cdot\log(q_d) - d_{KL}\big(\beta\big\|\alpha\big)\\
= \mathop{\mathbf{E}}_{X'\sim\mu}\Big[\log\Big(\mathop{\mathbf{Pr}}_{X\sim\mu}\big[X_2 \in G_2 \big| X_1 = X_1'\big]\Big)\Big| X_1' \in G_1, X_2' \in G_2\Big] &- d_{KL}\big(\beta\big\|\alpha\big),
\end{aligned}$$

as required (the stated inequality follows from the non-negativity of relative entropy). $\blacksquare_{\textit{Claim 2}}$

Proof of Claim 1. Let us first consider the case of two variables $(Y_1, Y_2) \sim \nu$, supported, respectively, on $\mathcal{G}_1 \uplus \mathcal{B}_1$ and $\mathcal{G}_2 \uplus \mathcal{B}_2$:

$$\begin{aligned}
&\log\Big(\mathop{\mathbf{Pr}}_{\nu}[Y_1 \in \mathcal{G}_1, Y_2 \in \mathcal{G}_2]\Big) \qquad (6)\\
&\quad= \log\Big(\mathbf{Pr}[Y_1 \in \mathcal{G}_1]\Big) + \log\Big(\mathbf{Pr}\big[Y_2 \in \mathcal{G}_2 \big| Y_1 \in \mathcal{G}_1\big]\Big)\\
&\quad\le \log\Big(\mathbf{Pr}[Y_1 \in \mathcal{G}_1]\Big) + \mathop{\mathbf{E}}_{(Y_1', Y_2')\sim\nu}\Big[\log\Big(\mathbf{Pr}\big[Y_2 \in \mathcal{G}_2 \big| Y_1 = Y_1'\big]\Big)\Big| Y_1' \in \mathcal{G}_1, Y_2' \in \mathcal{G}_2\Big],
\end{aligned}$$

as follows from Claim 2.

Let μ' denote the distribution of $(X_1, \dots, X_m)$, conditioned upon $[\bigwedge_{j=1}^m X_j \in G_j]$.

Inequality (5) follows by induction. For the base case, note that

$$\log\left(\mathop{\mathbf{Pr}}_{X\sim\mu}\left[\bigwedge_{j=1}^{m} X_j \in G_j\right]\right)$$

$$\le \log\left(\Pr_{X\sim\mu}[X_1 \in G_1]\right) + \mathop{\mathbf{E}}_{X'\sim\mu}\left[\log\left(\Pr_{X\sim\mu}\left[\bigwedge_{j=2}^{m} X_j \in G_j \,\middle|\, X_1 = X_1'\right]\right) \,\middle|\, \bigwedge_{j=1}^{m} X_j' \in G_j\right]$$
$$= \log\left(\Pr_{X\sim\mu}[X_1 \in G_1]\right) + \mathop{\mathbf{E}}_{X'\sim\mu'}\left[\log\left(\Pr_{X\sim\mu}\left[\bigwedge_{j=2}^{m} X_j \in G_j \,\middle|\, X_1 = X_1'\right]\right)\right],$$

as a direct application of (6).

For the induction step, assume that

$$\log\left(\Pr_{X\sim\mu}\left[\bigwedge_{j=1}^{m} X_j \in G_j\right]\right) \le \mathop{\mathbf{E}}_{X'\sim\mu'}\left[\sum_{i=1}^{k}\log\left(\Pr_{X\sim\mu}\left[X_i \in G_i \,\middle|\, \bigwedge_{j=1}^{i-1} X_j = X_j'\right]\right)\right] \tag{7}$$
$$+ \mathop{\mathbf{E}}_{X'\sim\mu'}\left[\underbrace{\log\left(\Pr_{X\sim\mu}\left[\bigwedge_{j=k+1}^{m} X_j \in G_j \,\middle|\, \bigwedge_{j=1}^{k} X_j = X_j'\right]\right)}_{\circledast}\right]$$

holds for some $k \ge 1$. Denote by $(x_j')_{j=1}^{k}$ the value taken by $(X_j')_{j=1}^{k}$ and by ν the distribution of $(X_j)_{j=k+1}^{m}$ when $(X_j)_{j=1}^{m} \sim \mu$ and $\bigwedge_{j=1}^{k} X_j = x_j'$.

Let us inspect $\circledast$:

$$\log\left(\Pr_{X\sim\mu}\left[\bigwedge_{j=k+1}^{m} X_j \in G_j \,\middle|\, \bigwedge_{j=1}^{k} X_j = X_j'\right]\right) = \log\left(\Pr_{X\sim\nu}\left[\bigwedge_{j=k+1}^{m} X_j \in G_j\right]\right)$$
$$\le \log\left(\Pr_{X\sim\nu}[X_{k+1} \in G_{k+1}]\right)$$
$$+ \mathop{\mathbf{E}}_{X''\sim\nu}\left[\log\left(\Pr_{X\sim\nu}\left[\bigwedge_{j=k+2}^{m} X_j \in G_j \,\middle|\, X_{k+1} = X_{k+1}''\right]\right) \,\middle|\, \bigwedge_{j=k+1}^{m} X_j'' \in G_j\right],$$

where "$X \sim \nu$" stands for $(X_j)_{k+1}^{m} \sim \nu$, $X'' = (X_{k+1}'', \ldots, X_m'')$ and the inequality is an application of (6). Accordingly,

$$\mathop{\mathbf{E}}_{X'\sim\mu'}\left[\underbrace{\log\left(\Pr_{X\sim\mu}\left[\bigwedge_{j=k+1}^{m} X_j \in G_j \,\middle|\, \bigwedge_{j=1}^{k} X_j = X_j'\right]\right)}_{\circledast}\right]$$
$$\le \mathop{\mathbf{E}}_{X'\sim\mu'}\left[\log\left(\Pr_{X\sim\nu}[X_{k+1} \in G_{k+1}]\right)\right]$$
$$+ \mathop{\mathbf{E}}_{\substack{X'\sim\mu' \\ X''\sim\nu}}\left[\log\left(\Pr_{X\sim\nu}\left[\bigwedge_{j=k+2}^{m} X_j \in G_j \,\middle|\, X_{k+1} = X_{k+1}''\right]\right) \,\middle|\, \bigwedge_{j=k+1}^{m} X_j'' \in G_j\right]$$
$$= \mathop{\mathbf{E}}_{X'\sim\mu'}\left[\log\left(\Pr_{X\sim\mu}\left[X_{k+1} \in G_{k+1} \,\middle|\, \bigwedge_{j=1}^{k} X_j = X_j'\right]\right)\right]$$

$$+ \underset{X''\sim\mu'}{\mathbf{E}}\left[\log\left(\underset{X\sim\mu}{\mathbf{Pr}}\left[\bigwedge_{j=k+2}^{m} X_j \in G_j \middle| \bigwedge_{j=1}^{k+1} X_j = X''_j\right]\right)\right],$$

as follows from the definition of ν – namely, from the facts that choosing $X \sim \nu$ is the same as choosing $X \sim \mu$, subject to $\left[\bigwedge_{j=1}^{k} X_j = X'_j\right]$, and that the distribution of X'' resulting from $X' \sim \mu'$ and $X'' \sim \nu$, subject to $\left[\bigwedge_{j=k+1}^{m} X''_j \in G_j\right]$, is simply μ'. Substituting this to (7) gives

$$\begin{aligned}\log\left(\underset{X\sim\mu}{\mathbf{Pr}}\left[\bigwedge_{j=1}^{m} X_j \in G_j\right]\right) &\leq \underset{X'\sim\mu'}{\mathbf{E}}\left[\sum_{i=1}^{k}\log\left(\underset{X\sim\mu}{\mathbf{Pr}}\left[X_i \in G_i \middle| \bigwedge_{j=1}^{i-1} X_j = X'_j\right]\right)\right] \\ &\quad + \underset{X'\sim\mu'}{\mathbf{E}}\left[\log\left(\underset{X\sim\mu}{\mathbf{Pr}}\left[X_{k+1} \in G_{k+1} \middle| \bigwedge_{j=1}^{k} X_j = X'_j\right]\right)\right] \\ &\quad + \underset{X'\sim\mu'}{\mathbf{E}}\left[\log\left(\underset{X\sim\mu}{\mathbf{Pr}}\left[\bigwedge_{j=k+2}^{m} X_j \in G_j \middle| \bigwedge_{j=1}^{k+1} X_j = X'_j\right]\right)\right] \\ &= \underset{X'\sim\mu'}{\mathbf{E}}\left[\sum_{i=1}^{k+1}\log\left(\underset{X\sim\mu}{\mathbf{Pr}}\left[X_i \in G_i \middle| \bigwedge_{j=1}^{i-1} X_j = X'_j\right]\right)\right] \\ &\quad + \underset{X'\sim\mu'}{\mathbf{E}}\left[\log\left(\underset{X\sim\mu}{\mathbf{Pr}}\left[\bigwedge_{j=k+2}^{m} X_j \in G_j \middle| \bigwedge_{j=1}^{k+1} X_j = X'_j\right]\right)\right],\end{aligned}$$

thus completing the induction step; for $k = m-1$ the above reads:

$$\log\left(\underset{X\sim\mu}{\mathbf{Pr}}\left[\bigwedge_{j=1}^{m} X_j \in G_j\right]\right) \leq \underset{X'\sim\mu'}{\mathbf{E}}\left[\sum_{i=1}^{m}\log\left(\underset{X\sim\mu}{\mathbf{Pr}}\left[X_i \in G_i \middle| \bigwedge_{j=1}^{i-1} X_j = X'_j\right]\right)\right],$$

which is precisely (5), and (4) follows from it by the concavity of log. $\blacksquare$ *Claim 1*

As a side note, we give the following generalisation, where the i'th "goodness criterion" may depend not only on the value taken by X_i, but also on the values of $X_1, \ldots, X_{i-1}$, as long as the condition is "monotone non-increasing" (e.g., the value of (X_1, X_2) cannot be good when that of X_1 is bad).

Corollary 2. *Let $X_1, \ldots, X_m$ be random variables, so that for each $i \in [n]$ the tuple $(X_j)_1^i$ is supported on (finite) $\mathcal{G}_i \,\dot\cup\, \mathcal{B}_i$ and for all $i_1 < i_2$ it holds that $\mathcal{G}_{i_2}([i_1]) \subseteq \mathcal{G}_{i_1}$. Let μ denote the joint distribution of $X = (X_1, \ldots, X_m)$, then*

$$\underset{X\sim\mu}{\mathbf{Pr}}[X \in \mathcal{G}_m] \leq \prod_{i=1}^{m} \underset{X'\sim\mu}{\mathbf{E}}\left[\underset{X\sim\mu}{\mathbf{Pr}}\left[(X_j)_{j=1}^{i} \in \mathcal{G}_i \middle| \bigwedge_{j=1}^{i-1} X_j = X'_j\right] \middle| X' \in \mathcal{G}_m\right],$$

where $X' = (X'_1, \ldots, X'_m)$ and X are independent from one another, unless conditioned explicitly.

Proof. For every $i \in [m]$, let Y_i be a random variable that takes value (X_i, Q_i), where

$$Q_i = \begin{cases} 1 & \text{if } (X_j)_{j=1}^{i} \in \mathcal{G}_i; \\ 0 & \text{otherwise.} \end{cases}$$

Let $G_i \stackrel{\text{def}}{=} \text{supp}(X_i) \times \{1\}$ and apply Claim 1 to the case of random variables Y_i and "good" sets G_i. ■

5.2 Confidence-weighted accuracy of Boolean prediction

Claim 3. *Let X and Y be random variables, X being supported on $\{0,1\}$. Then*

$$\mathop{\mathbf{E}}_{X',Y'}\left[\mathop{\mathbf{Pr}}_{X,Y}\left[X = X' \middle| Y = Y'\right]\right] - \frac{1}{2} = 2 \cdot \mathop{\mathbf{E}}_{Y=y}\left[\left(\mathop{\mathbf{E}}_{X}\left[X \middle| Y = y\right] - \frac{1}{2}\right)^2\right],$$

where (X', Y') are distributed identically to (X, Y). In particular, if $X \sim \mathcal{U}_{\{0,1\}}$, then

$$\mathop{\mathbf{E}}_{X',Y'}\left[\mathop{\mathbf{Pr}}_{X,Y}\left[X = X' \middle| Y = Y'\right] - \frac{1}{2}\right] \in \Theta\big(\mathbf{I}\left[X : Y\right]\big).$$

Intuitively, if we view X as "unknown", Y as "known" and try to predict the former using the latter, then the expectation of $\mathbf{Pr}\left[X = X' \middle| Y = Y'\right] - 1/2$ can be interpreted as *confidence-weighted accuracy* when $X \sim \mathcal{U}_{\{0,1\}}$.[13] Note that the above statement demonstrates qualitative difference between this quantity and the standard notion of *confidence*:

$$\mathop{\mathbf{E}}_{Y=y}\left[\left|\mathop{\mathbf{E}}_{X}\left[X \middle| Y = y\right] - \frac{1}{2}\right|\right] = \mathop{\mathbf{E}}_{X',Y'}\left[\left|\mathop{\mathbf{Pr}}_{X,Y}\left[X = X' \middle| Y = Y'\right] - \frac{1}{2}\right|\right] \in \Theta\left(\sqrt{\mathbf{I}\left[X : Y\right]}\right).$$

Proof of Claim 3. Let $g(y) \stackrel{\text{def}}{=} \mathbf{Pr}\left[X = 0 \middle| Y = y\right]$ for every $y \in \text{supp}(Y)$, then

$$\begin{aligned}
&\mathop{\mathbf{E}}_{X',Y'}\left[\mathop{\mathbf{Pr}}_{X,Y}\left[X = X' \middle| Y = Y'\right]\right] \\
&\quad = \mathop{\mathbf{E}}_{Y'=y'}\left[\mathbf{Pr}\left[X' = 0 \middle| Y' = y'\right] \cdot g(y') + \mathbf{Pr}\left[X' = 1 \middle| Y' = y'\right] \cdot (1 - g(y'))\right] \\
&\quad = \mathop{\mathbf{E}}_{Y'}\left[g(Y') \cdot g(Y') + (1 - g(Y')) \cdot (1 - g(Y'))\right] \\
&\quad = 2 \cdot \mathop{\mathbf{E}}_{Y}\left[\left(g(Y) - \frac{1}{2}\right)^2\right] + \frac{1}{2} = 2 \cdot \mathop{\mathbf{E}}_{Y=y}\left[\left(\mathop{\mathbf{E}}_{X}\left[X \middle| Y = y\right] - \frac{1}{2}\right)^2\right] + \frac{1}{2}.
\end{aligned}$$

If $X \sim \mathcal{U}_{\{0,1\}}$, then

$$H(X) - H\left(X \middle| Y = y\right) = 1 - H\left(X \middle| Y = y\right) \in \Theta\left(\left(\mathbf{E}\left[X \middle| Y = y\right] - \frac{1}{2}\right)^2\right)$$

for every $y \in \text{supp}(Y)$, and therefore,

$$\mathbf{I}\left[X : Y\right] = \mathop{\mathbf{E}}_{Y=y}\left[H(X) - H\left(X \middle| Y = y\right)\right] \in \Theta\left(\mathop{\mathbf{E}}_{Y=y}\left[\left(\mathop{\mathbf{E}}_{X}\left[X \middle| Y = y\right] - \frac{1}{2}\right)^2\right]\right),$$

as required. ■*Claim 3*

[13] Interpret the pair (X', Y') as the "actual outcome" of the experiment, then $\mathbf{Pr}\left[X = X' \middle| Y = Y'\right]$ measures "how likely" the value of X' was, conditioned upon the value of Y'.

6 The $\mathcal{R}^{\|,pub}$-complexity of $\widetilde{cEq}$-neg_T – a lower bound

Notation (*protocols in* $\mathcal{D}^{\|}_{\varepsilon}$)**.** Let $\mathcal{P}$ be a protocol in $\mathcal{D}^{\|}_{\varepsilon}$, where both Alice and Bob send r bits to the referee.

- Let $Al : \{0,1\}^n \to \{0,1\}^r$ be the "message function" of Alice, according to $\mathcal{P}$ – i.e., $Al(x)$ is sent when she receives input x;
- let $\alpha : \{0,1\}^n \to \mathrm{Pow}(\{0,1\}^n)$ be the "neighbourhood function" corresponding to $Al(\cdot)$ – i.e., $\alpha(x) \stackrel{\text{def}}{=} \{x' \big| Al(x') = Al(x)\}$;
- define $Bo(y)$ and $\beta(y)$ similarly.

Note that $\alpha(\cdot)$ and $\beta(\cdot)$ naturally correspond to *partitions* of, respectively, Alice's and Bob's input spaces: every possible message sent by a player corresponds to an element of his partition, which is the set of input values corresponding to this message. These partitions are fully determined by the message functions $Al(\cdot)$ and $Bo(\cdot)$ and, in some sense, they reveal "all that matters" about a protocol in $\mathcal{D}^{\|}_{\mu,\varepsilon}$, as we can always consider (in the context of lower bounds, *assume*) an "optimal" referee – the one who outputs a most likely guess regarding $f(X,Y)$ with respect to μ, given the messages $Al(X)$ and $Bo(Y)$ from the players.

To analyse the complexity of $\widetilde{cEq}$-neg_T, we reason as follows.

- We identify a useful property of all sufficiently accurate protocols for Eq_u (cf. Corollary 3).
- We consider protocols for Eq_u-neg_T and see that a more rigid form of the above characterisation must hold if T is a so-called "small-bias space" (cf. Lemma 4).
- We view $\widetilde{Eq}$-neg_T as "Eq_u-neg_T on a random subset u" – accordingly, a protocol for $\widetilde{Eq}$-neg_T must satisfy the above characterisation with respect to "random projections", which leads to a more symmetric criterion (cf. Lemma 5).
- We observe that a protocol for $\widetilde{cEq}$-neg_T must, in a sense, simultaneously solve n "rotated instances" of $\widetilde{Eq}$-neg_T – therefore, such a protocol must satisfy the n "rotated versions" of the above characterisation, which in turn leads to an even more symmetric criterion (cf. Lemma 6) and then to the desired complexity lower bound (cf. Corollary 4).

6.1 Characterising protocols for Eq_u

To characterise protocols that solve the equality problem, we use the following idea: Suppose for simplicity that $u = [n]$ (i.e., the protocol solves the standard Eq). If the partitions of $\{0,1\}^n$ defined by $\alpha(\cdot)$ and $\beta(\cdot)$ are suitable for solving Eq, then with respect to $X = Y \Subset \{0,1\}^n$, the pair of subsets $(\alpha(X), \beta(Y))$ will (typically) be such that $[X = Y]$ is "likely", given the messages – namely,

$$\mathop{\mathbf{Pr}}_{(X',Y') \Subset \alpha(X) \times \beta(X)} \big[X' = Y'\big] \gg \mathop{\mathbf{Pr}}_{(X',Y') \Subset \{0,1\}^{n+n}} \big[X' = Y'\big] = \frac{1}{2^n}.$$

Applying the optimistic chain inequality (Claim 1) with respect to the event $[X' = Y'] = [\bigwedge_i X'_i = Y'_i]$ and integrating over the rectangles of the form $\alpha(x) \times \beta(x)$ will lead to a convenient protocol characterisation.

Notation (*protocols for* Eq_u)**.** Fix some $u \subseteq [n]$ and let $\mathcal{P}$ be a protocol that solves Eq_u in $\mathcal{D}^{\|}_{\mu_{Eq_u},\varepsilon}$. In addition to $Al(\cdot)$, $Bo(\cdot)$, $\alpha(\cdot)$ and $\beta(\cdot)$ defined earlier, we will use the following variations: Let $z \in \{0,1\}^{|u|}$, then

- denote by $Al^*(z)$ the *distribution* over $\{0,1\}^r$, corresponding to $Al(X')$ when X' is chosen uniformly at random from $\{x' \in \{0,1\}^n \big| x'_u = z\}$;
- denote by $\alpha^*(z)$ the *distribution* over $\mathrm{Pow}(\{0,1\}^n)$, corresponding to $\{x' \big| Al(x') = m_0\}$ when m_0 is the value taken by $M \sim Al^*(z)$ (alternatively, $\alpha^*(z)$ can be defined as the distribution of $\alpha(X')$ when X' is chosen uniformly at random from $\{x' \in \{0,1\}^n \big| x'_u = z\}$);
- define $Bo^*(z)$ and $\beta^*(z)$ similarly.

We will argue that the following type of objects are, in a sense, "typical for $\mathcal{P}$" (that will be the technical core of our characterisation).

Definition 2 (*good rectangles*)**.** *Let $A, B \subseteq \{0,1\}^n$. We call the rectangle $A \times B \subseteq \{0,1\}^{n+n}$ good if*

$$\mathop{\mathbf{Pr}}_{(X',Y') \Subset A \times B}\left[X'_u = Y'_u\right] \geq \frac{1}{4\sqrt{\varepsilon} + 2^{-|u|}} \cdot \frac{1}{2^{|u|}}.$$

Our first step in this part is characterising good rectangles in a technically-convenient manner. We need the following.

Notation (*delta-properties of sets and partitions*)**.** Let $W \subseteq \{0,1\}^n$, $i \in [|u|]$ and $z \in \{0,1\}^{|u|}$. Then

$$\delta_W^{u,i}(z) \stackrel{\text{def}}{=} \mathop{\mathbf{Pr}}_{X \Subset W}\left[X_u(i) = z_i \big| X_u([i-1]) = z_{[i-1]}\right] - \frac{1}{2},$$

$$\Delta_\alpha^{u,i}(z) \stackrel{\text{def}}{=} \mathop{\mathbf{Pr}}_{X \Subset \alpha^*(z)}\left[X_u(i) = z_i \big| X_u([i-1]) = z_{[i-1]}\right] - \frac{1}{2} \quad \left\{ = \mathop{\mathbf{E}}_{A \sim \alpha^*(z)}\left[\delta_A^{u,i}(z)\right] \right\},$$

and similarly for $\Delta_\beta^{u,i}(z)$.

Lemma 1. *Let $A, B \subseteq \{0,1\}^n$. If the rectangle $A \times B$ is good, then*

$$\mathop{\mathbf{E}}_{Z}\left[\sum_{i=1}^{|u|} \delta_A^{u,i}(Z) \cdot \delta_B^{u,i}(Z)\right] \geq \frac{1}{4} \cdot \ln\left(\frac{1}{4\sqrt{\varepsilon} + 2^{-|u|}}\right),$$

where Z is distributed as X_u when $(X,Y) \Subset A \times B$ conditioned on $[X_u = Y_u]$.

Proof. By the definition of good rectangles,

$$\begin{aligned}
\frac{1}{4\sqrt{\varepsilon} + 2^{-|u|}} \cdot \frac{1}{2^{|u|}} &\leq \mathop{\mathbf{Pr}}_{(X',Y') \Subset A \times B}\left[X'_u = Y'_u\right] = \mathbf{Pr}\left[\bigwedge_{j=1}^{|u|} X'_u(i) = Y'_u(i)\right] \\
&\leq \prod_{i=1}^{|u|} \mathop{\mathbf{E}}_{(X',Y') \Subset A \times B}\left[\circledast \, \big| X'_u = Y'_u\right] \\
&= \prod_{i=1}^{|u|} \mathop{\mathbf{E}}_{Z}\left[\mathop{\mathbf{Pr}}_{(X,Y) \Subset A \times B}\left[X_u(i) = Y_u(i) \big| X_u([i-1]) = Y_u([i-1]) = Z_{[i-1]}\right]\right],
\end{aligned}$$

where the second inequality is the optimistic chain (Claim 1), $\circledast$ stands for

$$\mathop{\mathbf{Pr}}_{(X,Y) \Subset A \times B}\left[X_u(i) = Y_u(i) \big| X_u([i-1]) = X'_u([i-1]),\, Y_u([i-1]) = Y'_u([i-1])\right]$$

and Z is distributed as X'_u when $(X', Y') \in A \times B$ conditioned on $[X'_u = Y'_u]$.

On the other hand, for every $i \in [|u|]$ and $z \in \{0,1\}^{|u|}$:

$$\begin{aligned}
&\mathop{\mathbf{Pr}}_{(X,Y)\in A\times B}\big[X_u(i) = Y_u(i) \big| X_u([i-1]) = Y_u([i-1]) = z_{[i-1]}\big] \\
&= \mathbf{Pr}\big[X_u(i) = Y_u(i) = z_i \big| X_u([i-1]) = Y_u([i-1]) = z_{[i-1]}\big] \\
&\quad + \mathbf{Pr}\big[X_u(i) = Y_u(i) = 1 - z_i \big| X_u([i-1]) = Y_u([i-1]) = z_{[i-1]}\big] \\
&= \mathop{\mathbf{Pr}}_{X\in A}\big[X_u(i) = z_i \big| X_u([i-1]) = z_{[i-1]}\big] \cdot \mathop{\mathbf{Pr}}_{Y\in B}\big[Y_u(i) = z_i \big| Y_u([i-1]) = z_{[i-1]}\big] \\
&\quad + \mathop{\mathbf{Pr}}_{X\in A}\big[X_u(i) = 1 - z_i \big| X_u([i-1]) = z_{[i-1]}\big] \cdot \mathop{\mathbf{Pr}}_{Y\in B}\big[Y_u(i) = 1 - z_i \big| Y_u([i-1]) = z_{[i-1]}\big] \\
&= \frac{1}{2} + 2 \cdot \left(\mathop{\mathbf{Pr}}_{X\in A}\big[X_u(i) = z_i \big| X_u([i-1]) = z_{[i-1]}\big] - \frac{1}{2} \right) \\
&\qquad\qquad \cdot \left(\mathop{\mathbf{Pr}}_{Y\in B}\big[Y_u(i) = z_i \big| Y_u([i-1]) = z_{[i-1]}\big] - \frac{1}{2} \right) \\
&= \frac{1}{2} + 2 \cdot \delta_A^{u,i}(z) \cdot \delta_B^{u,i}(z).
\end{aligned}$$

Therefore,

$$\frac{1}{4\sqrt{\varepsilon} + 2^{-|u|}} \cdot \frac{1}{2^{|u|}} \le \prod_{i=1}^{|u|} \left(\frac{1}{2} + 2 \cdot \mathop{\mathbf{E}}_{Z}\left[\delta_A^{u,i}(Z) \cdot \delta_B^{u,i}(Z)\right] \right),$$

where Z is distributed as X_u when $(X, Y) \in A \times B$ conditioned on $[X_u = Y_u]$.

So,

$$\begin{aligned}
\ln\left(\frac{1}{4\sqrt{\varepsilon} + 2^{-|u|}} \cdot \frac{1}{2^{|u|}} \right) &\le \sum_{i=1}^{|u|} \left(\ln\left(\frac{1}{2}\right) + \ln\left(1 + 4 \cdot \mathop{\mathbf{E}}_{Z}\left[\delta_A^{u,i}(Z) \cdot \delta_B^{u,i}(Z)\right]\right) \right) \\
&\le |u| \cdot \ln\left(\frac{1}{2}\right) + 4 \cdot \sum_{i=1}^{|u|} \mathop{\mathbf{E}}_{Z}\left[\delta_A^{u,i}(Z) \cdot \delta_B^{u,i}(Z)\right],
\end{aligned}$$

as required. $\blacksquare_{\textit{Lemma 1}}$

Next we will "look inside" $\mathcal{P}$, for which we need the following.

Notation (*random variables corresponding to* $[X_u = Y_u]$)**.** Define:

- Let $Z \sim \mathcal{U}_{\{0,1\}^{|u|}}$.
- Let the pair of $\mathrm{Pow}(\{0,1\}^n)$-valued variables $(\mathcal{A}, \mathcal{B})$ be distributed as $(\alpha^*(Z), \beta^*(Z))$.
- Let Z' be distributed as X_u when $(X, Y) \in \mathcal{A} \times \mathcal{B}$ conditioned on $[X_u = Y_u]$.

Intuitively, the variable Z corresponds to sampling the protocol input from $\mu^{1}_{Eq_u}$: think of it as drawing uniformly-random (X, Y), subject to $X_u = Y_u = Z$. Then the rectangle $\mathcal{A} \times \mathcal{B}$ can be viewed as the knowledge that the referee obtains from the players' messages regarding the input pair. View Z' as a "sibling of Z", used in the proof for technical reasons.

Note two Markov chains that correspond to these random variables:

$$\mathcal{A} \leftrightarrow Z \leftrightarrow \mathcal{B} \quad \text{and} \quad Z \leftrightarrow (\mathcal{A}, \mathcal{B}) \leftrightarrow Z'.$$

We claim that the latter chain is *symmetric* in the following sense:

Lemma 2. *The marginal distributions of* $((\mathcal{A},\mathcal{B}),Z)$ *and of* $((\mathcal{A},\mathcal{B}),Z')$ *are the same.*

Proof. Let $(a,b) \in \mathrm{supp}(\mathcal{A},\mathcal{B})$ and denote by $[(a,b)]$ the event that $(\mathcal{A},\mathcal{B}) = (a,b)$, by $[a]$ the event that $\mathcal{A} = a$ and by $[b]$ the event that $\mathcal{B} = b$. Let $z_0 \in \{0,1\}^{|u|}$, then

$$\begin{aligned}
\mathbf{Pr}\big[(a,b)\big|Z = z_0\big] &= \mathbf{Pr}\big[a\big|Z = z_0\big] \cdot \mathbf{Pr}\big[b\big|Z = z_0\big] \\
&= \mathbf{Pr}\big[Z = z_0\big|a\big] \cdot \frac{\mathbf{Pr}[a]}{\mathbf{Pr}[Z = z_0]} \cdot \mathbf{Pr}\big[Z = z_0\big|b\big] \cdot \frac{\mathbf{Pr}[b]}{\mathbf{Pr}[Z = z_0]} \\
&= \mathbf{Pr}\big[Z = z_0\big|a\big] \cdot \mathbf{Pr}[a] \cdot \mathbf{Pr}\big[Z = z_0\big|b\big] \cdot \mathbf{Pr}[b] \cdot 2^{2|u|}.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\mathbf{Pr}[a] &= \mathop{\mathbf{Pr}}_{Z \Subset \{0,1\}^{|u|}}[\alpha^*(Z) = a] = \mathop{\mathbf{Pr}}_{X \Subset \{0,1\}^n}[\alpha(X) = a] = \frac{|a|}{2^n}, \\
\mathbf{Pr}\big[Z = z_0\big|a\big] &= \mathbf{Pr}\big[Z = z_0\big|\alpha^*(Z) = a\big] = \mathbf{Pr}\big[X_u = z_0\big|\alpha(X) = a\big] = \mathop{\mathbf{Pr}}_{X \Subset a}[X_u = z_0],
\end{aligned}$$

and similarly for $\mathbf{Pr}[b]$ and $\mathbf{Pr}\big[Z = z_0\big|b\big]$. Accordingly,

$$\mathbf{Pr}\big[(a,b)\big|Z = z_0\big] = \mathop{\mathbf{Pr}}_{X \Subset a}[X_u = z_0] \cdot \mathop{\mathbf{Pr}}_{Y \Subset b}[Y_u = z_0] \cdot |a| \cdot |b| \cdot 2^{2|u|-2n}.$$

Therefore,

$$\begin{aligned}
\mathbf{Pr}[[(a,b)] \wedge Z = z_0] &= \mathbf{Pr}[Z = z_0] \cdot \mathbf{Pr}\big[(a,b)\big|Z = z_0\big] \\
&= \mathop{\mathbf{Pr}}_{X \Subset a}[X_u = z_0] \cdot \mathop{\mathbf{Pr}}_{Y \Subset b}[Y_u = z_0] \cdot |a| \cdot |b| \cdot 2^{|u|-2n}
\end{aligned} \tag{8}$$

and

$$\begin{aligned}
\mathbf{Pr}[(a,b)] &= \sum_z \mathbf{Pr}[Z = z] \cdot \mathop{\mathbf{Pr}}_{X \Subset a}[X_u = z] \cdot \mathop{\mathbf{Pr}}_{Y \Subset b}[Y_u = z] \cdot |a| \cdot |b| \cdot 2^{2|u|-2n} \\
&= \mathop{\mathbf{Pr}}_{\substack{X \Subset a \\ Y \Subset b}}[X_u = Y_u] \cdot |a| \cdot |b| \cdot 2^{|u|-2n}.
\end{aligned} \tag{9}$$

On the other hand,

$$\begin{aligned}
\mathbf{Pr}\big[[(a,b)] \wedge Z' = z_0\big] &= \mathbf{Pr}\big[Z' = z_0\big|(a,b)\big] \cdot \mathbf{Pr}[(a,b)] \\
&= \frac{\mathbf{Pr}_{X \Subset a}[X_u = z_0] \cdot \mathbf{Pr}_{Y \Subset b}[Y_u = z_0]}{\mathbf{Pr}_{\substack{X \Subset a \\ Y \Subset b}}[X_u = Y_u]} \cdot \mathbf{Pr}[(a,b)] \\
&= \mathop{\mathbf{Pr}}_{X \Subset a}[X_u = z_0] \cdot \mathop{\mathbf{Pr}}_{Y \Subset b}[Y_u = z_0] \cdot |a| \cdot |b| \cdot 2^{|u|-2n} \\
&= \mathbf{Pr}[[(a,b)] \wedge Z = z_0],
\end{aligned}$$

where the last two inequalities follow from (9) and (8), respectively. $\blacksquare$ *Lemma 2*

Our characterisation of $\mathcal{P}$ will be based on the following structural observation.

Lemma 3.

$$\mathop{\mathbf{Pr}}_{\mathcal{A},\mathcal{B}}[\mathcal{A}\times\mathcal{B} \textit{ is a good rectangle}] > 1 - 2\varepsilon - 2\sqrt{\varepsilon}.$$

Proof. Let $(a,b) \in \{0,1\}^{r+r}$ be a pair of players' messages and

$$err(a,b) \stackrel{\text{def}}{=} \mathop{\mathbf{Pr}}_{(X,Y)\sim\mu_{Eq_u}}\left[\mathcal{P}(X,Y) \text{ makes an error}\middle| Al(X) = a, Bo(Y) = b\right].$$

By the correctness assumption,

$$\mathop{\mathbf{Pr}}_{(X,Y)\sim\mu_{Eq_u}}\left[err(Al(X), Bo(Y)) > \sqrt{\varepsilon}\right] < \sqrt{\varepsilon}.$$

Call a pair of messages $(a,b) \in \{0,1\}^{r+r}$ *bad* if $err(a,b) > \sqrt{\varepsilon}$ and *good* otherwise.

Recall that μ_{Eq_u} is the "uniform mixture" of $\mu^{0}_{Eq_u}$ and $\mu^{1}_{Eq_u}$. Accordingly, from the correctness assumption it follows that with respect to $(X,Y) \sim \mu^{1}_{Eq_u}$,

- $\mathcal{P}$ accepts (that is, produces output "1") with probability at least $1 - 2\varepsilon$;
- $(Al(X), Bo(Y))$ is a bad message with probability at most $2\sqrt{\varepsilon}$.

Note that sampling $(Al(X), Bo(Y))$ when $(X,Y) \sim \mu^{1}_{Eq_u}$ is the same as sampling $(Al^*(Z), Bo^*(Z))$ when $Z \sim \mathcal{U}_{\{0,1\}^{|u|}}$ – therefore, $(Al^*(Z), Bo^*(Z))$ is a good pair of messages accepted by the referee with probability at least $1 - 2\varepsilon - 2\sqrt{\varepsilon}$.

We will see next that a good pair of messages accepted by the referee defines a good rectangle; this will imply the lemma, as the rectangle corresponding to the pair of messages $(Al^*(Z), Bo^*(Z))$ under $Z \sim \mathcal{U}_{\{0,1\}^{|u|}}$ is distributed the same way as $\mathcal{A}\times\mathcal{B}$.

Suppose that (a,b) is a good pair of messages accepted by the referee and let $[(a,b)]$ denote the event $[(Al^*(Z), Bo^*(Z)) = (a,b)]$. Then

$$\begin{aligned}\mathop{\mathbf{Pr}}_{(X,Y)\in\{0,1\}^{n+n}}\left[(a,b)\middle|X_u \neq Y_u\right] &= \mathop{\mathbf{Pr}}_{\mu_{Eq_u}}\left[(a,b)\middle|X_u \neq Y_u\right]\\ &= \mathop{\mathbf{Pr}}_{\mu_{Eq_u}}\left[X_u \neq Y_u\middle|(a,b)\right]\cdot\frac{\mathbf{Pr}_{\mu_{Eq_u}}[(a,b)]}{\mathbf{Pr}_{\mu_{Eq_u}}[X_u \neq Y_u]}\\ &< 2\sqrt{\varepsilon}\cdot\mathop{\mathbf{Pr}}_{\mu_{Eq_u}}[(a,b)],\end{aligned}$$

as $\mathbf{Pr}_{\mu_{Eq_u}}[X_u \neq Y_u] < 1/2$. Similarly,

$$\mathop{\mathbf{Pr}}_{(X,Y)\in\{0,1\}^{n+n}}\left[(a,b)\middle|X_u = Y_u\right] > 2(1-\sqrt{\varepsilon})\cdot\mathop{\mathbf{Pr}}_{\mu_{Eq_u}}[(a,b)].$$

So,

$$\mathop{\mathbf{Pr}}_{(X,Y)\in\{0,1\}^{n+n}}\left[(a,b)\middle|X_u \neq Y_u\right] < \frac{\sqrt{\varepsilon}}{1-\sqrt{\varepsilon}}\cdot\mathop{\mathbf{Pr}}_{(X,Y)\in\{0,1\}^{n+n}}\left[(a,b)\middle|X_u = Y_u\right]$$

and

$$\begin{aligned}\mathop{\mathbf{Pr}}_{(X,Y)\in\{0,1\}^{n+n}}[(a,b)] &\leq \mathbf{Pr}[X_u = Y_u]\cdot\mathbf{Pr}\left[(a,b)\middle|X_u = Y_u\right] + \mathbf{Pr}\left[(a,b)\middle|X_u \neq Y_u\right]\\ &< \left(\frac{1}{2^{|u|}} + \frac{\sqrt{\varepsilon}}{1-\sqrt{\varepsilon}}\right)\cdot\mathbf{Pr}\left[(a,b)\middle|X_u = Y_u\right].\end{aligned}$$

Finally,

$$\Pr_{(X,Y)\Subset\{0,1\}^{n+n}}\big[X_u = Y_u \big| (a,b)\big] = \frac{\mathbf{Pr}\big[(a,b)\big|X_u = Y_u\big]}{\mathbf{Pr}\big[(a,b)\big]} \cdot \mathbf{Pr}[X_u = Y_u] \tag{10}$$
$$> \frac{1}{\frac{1}{2^{|u|}} + \frac{\sqrt{\varepsilon}}{1-\sqrt{\varepsilon}}} \cdot \frac{1}{2^{|u|}} \geq \frac{1}{4\sqrt{\varepsilon} + 2^{-|u|}} \cdot \frac{1}{2^{|u|}},$$

as $\varepsilon < 1/2$. The result follows. ■*Lemma 3*

We are ready for the main statement of this part.

Corollary 3. *Let $\mathcal{P}$ be a protocol that solves Eq_u in $\mathcal{D}^{\|}_{\mu_{Eq_u},\varepsilon}$, with $\Delta^{u,i}_\alpha$ and $\Delta^{u,i}_\beta$ as defined earlier. Then*

$$\sum_{i=1}^{|u|}\Big\langle \Delta^{u,i}_\alpha\,,\,\Delta^{u,i}_\beta\Big\rangle > \frac{1}{4}\cdot\ln\bigg(\frac{1}{4\sqrt{\varepsilon}+2^{-|u|}}\bigg) - 2\sqrt{\varepsilon}\cdot|u|.$$

Proof. We analyse the quantity

$$\mathop{\mathbf{E}}_{(\mathcal{A},\mathcal{B}),\,Z'}\left[\sum_{i=1}^{|u|}\delta^{u,i}_{\mathcal{A}}(Z')\cdot\delta^{u,i}_{\mathcal{B}}(Z')\right].$$

On the one hand,

$$\begin{aligned}
&\mathop{\mathbf{E}}_{(\mathcal{A},\mathcal{B}),\,Z'}\left[\sum_{i=1}^{|u|}\delta^{u,i}_{\mathcal{A}}(Z')\cdot\delta^{u,i}_{\mathcal{B}}(Z')\right]\\
&\quad\geq \mathbf{Pr}[\mathcal{A}\times\mathcal{B}\text{ is a good rectangle}]\cdot\frac{1}{4}\cdot\ln\bigg(\frac{1}{4\sqrt{\varepsilon}+2^{-|u|}}\bigg)\\
&\qquad\quad+\Big(1-\mathbf{Pr}[\mathcal{A}\times\mathcal{B}\text{ is a good rectangle}]\Big)\cdot\min_{A,B,z}\left\{\sum_{i=1}^{|u|}\delta^{u,i}_{A}(z)\cdot\delta^{u,i}_{B}(z)\right\}\\
&\quad>\bigg(\frac{1}{4}-\sqrt{\varepsilon}\bigg)\cdot\ln\bigg(\frac{1}{4\sqrt{\varepsilon}+2^{-|u|}}\bigg)-\sqrt{\varepsilon}\cdot|u|\\
&\quad\geq\frac{1}{4}\cdot\ln\bigg(\frac{1}{4\sqrt{\varepsilon}+2^{-|u|}}\bigg)-2\sqrt{\varepsilon}\cdot|u|,
\end{aligned}$$

where the first inequality is Lemma 1 and the second one is Lemma 3. On the other,

$$\begin{aligned}
\mathop{\mathbf{E}}_{(\mathcal{A},\mathcal{B}),\,Z'}\left[\sum_{i=1}^{|u|}\delta^{u,i}_{\mathcal{A}}(Z')\cdot\delta^{u,i}_{\mathcal{B}}(Z')\right] &= \mathop{\mathbf{E}}_{Z,\,(\mathcal{A},\mathcal{B})}\left[\sum_{i=1}^{|u|}\delta^{u,i}_{\mathcal{A}}(Z)\cdot\delta^{u,i}_{\mathcal{B}}(Z)\right]\\
&=\sum_{i=1}^{|u|}\mathop{\mathbf{E}}_{Z\Subset\{0,1\}^{|u|}}\left[\bigg(\mathop{\mathbf{E}}_{A\sim\alpha^*(Z)}\Big[\delta^{u,i}_{A}(Z)\Big]\bigg)\cdot\bigg(\mathop{\mathbf{E}}_{B\sim\beta^*(Z)}\Big[\delta^{u,i}_{B}(Z)\Big]\bigg)\right]\\
&=\sum_{i=1}^{|u|}\mathop{\mathbf{E}}_{Z\Subset\{0,1\}^{|u|}}\Big[\Delta^{u,i}_\alpha(Z)\cdot\Delta^{u,i}_\beta(Z)\Big] = \sum_{i=1}^{|u|}\Big\langle\Delta^{u,i}_\alpha\,,\,\Delta^{u,i}_\beta\Big\rangle,
\end{aligned}$$

where the first equality is Lemma 2. ■*Corollary 3*

6.2 Characterising protocols for Eq_u-neg_T

Lemma 4. *Let T be a δ-biased space for some $\delta > 0$ and assume that $\mathcal{P}$ solves Eq_u-$neg_T(X,Y)$ in $\mathcal{D}^{\|}_{\mu_{Eq_u\text{-}neg_T},\varepsilon}$. Then*

$$\sum_{i\in u} \mathop{\mathbf{I}}_{X\in\{0,1\}^n}\left[X_i : X_{u\setminus\{i\}}, Al(X)\right] \cdot \mathop{\mathbf{I}}_{Y\in\{0,1\}^n}\left[Y_i : Y_{u\setminus\{i\}}, Bo(Y)\right]$$
$$\in \Omega\left(\ln\left(\frac{1}{|T|\cdot(\varepsilon+2^{-|u|})}\right)\right) - O\left(\left(\sqrt{|T|\cdot\varepsilon}+\delta\right)\cdot|u|\right).$$

Proof. From the definition of $\mu_{Eq_u\text{-}neg_T}$ and the correctness assumption it follows that for any $\tau \in T$, if $(X+\tau, Y) \sim \mu_{Eq_u}$, then $\mathcal{P}$ solves $Eq_u(X+\tau,Y)$ with error at most

$$\varepsilon_T \stackrel{\text{def}}{=} |T|\cdot(\varepsilon+2^{-|u|}).$$

Let $T_u \stackrel{\text{def}}{=} \left\{\tau' \middle| \tau'_u \in T|_u, \tau'_{[n]\setminus u} = \bar{0}\right\}$ – in other words, T_u contains the elements of T with bits outside u set to 0. To keep the notation simple, assume that $|T_u| = |T|$.[14]

Observe that for any $\tau \in T$ and the corresponding $\tau' \in T_u$, it holds that $Eq_u(X+\tau,Y) \equiv Eq_u(X+\tau',Y)$ and $(X+\tau,Y) \sim \mu_{Eq_u}$ whenever $(X+\tau',Y)\sim\mu_{Eq_u}$. Accordingly, $\mathcal{P}$ solves $Eq_u(X+\tau',Y)$ when $(X+\tau',Y)\sim\mu_{Eq_u}$ with error at most ε_T.

Corollary 3 implies that

$$\mathop{\mathbf{E}}_{\tau'\in T_u}\left[\sum_{i=1}^{|u|}\left\langle \Delta^{u,i}_{\alpha,\tau'}\, ,\, \Delta^{u,i}_{\beta}\right\rangle\right] > \frac{1}{4}\cdot\ln\left(\frac{1}{4\sqrt{\varepsilon_T}+2^{-|u|}}\right) - 2\sqrt{\varepsilon_T}\cdot|u|$$

for $\Delta^{u,i}_{\alpha,\tau'}(z) \stackrel{\text{def}}{=} \Delta^{u,i}_{\alpha}(z\oplus\tau'_u)$ for every $z\in\{0,1\}^{|u|}$ and $\tau'\in T_u$. For any $i \in [|u|]$:

$$\begin{aligned}
\mathop{\mathbf{E}}_{\tau'\in T_u}\left[\left\langle \Delta^{u,i}_{\alpha,\tau'}\, ,\, \Delta^{u,i}_{\beta}\right\rangle\right] &= \mathop{\mathbf{E}}_{\tau'}\left[\sum_{s\subset[|u|]} \widehat{\Delta^{u,i}_{\alpha,\tau'}}(s)\cdot\widehat{\Delta^{u,i}_{\beta}}(s)\right]\\
&= \sum_{s\subset[|u|]}\mathop{\mathbf{E}}_{\tau'}\left[\widehat{\Delta^{u,i}_{\alpha}}(s)\cdot\chi_s(\tau'_u)\cdot\widehat{\Delta^{u,i}_{\beta}}(s)\right]\\
&= \sum_{s\subset[|u|]}\left(\widehat{\Delta^{u,i}_{\alpha}}(s)\cdot\widehat{\Delta^{u,i}_{\beta}}(s)\cdot\mathop{\mathbf{E}}_{\tau'}\left[\chi_s(\tau'_u)\right]\right)\\
&\le \widehat{\Delta^{u,i}_{\alpha}}(\emptyset)\cdot\widehat{\Delta^{u,i}_{\beta}}(\emptyset) + \frac{1}{4}\cdot\max_{s\neq\emptyset}\left\{\mathop{\mathbf{E}}_{\tau'}\left[\chi_s(\tau'_u)\right]\right\}\\
&\le \widehat{\Delta^{u,i}_{\alpha}}(\emptyset)\cdot\widehat{\Delta^{u,i}_{\beta}}(\emptyset) + \frac{\delta}{4},
\end{aligned}$$

as $\left|\Delta^{u,i}_{\alpha}(z)\right|, \left|\Delta^{u,i}_{\beta}(z)\right| \le 1/2$ always and $T_u|_u \subseteq \{0,1\}^{|u|}$ is necessarily a δ-biased space. So,

$$\sum_{i=1}^{|u|}\widehat{\Delta^{u,i}_{\alpha}}(\emptyset)\cdot\widehat{\Delta^{u,i}_{\beta}}(\emptyset) > \frac{1}{4}\cdot\ln\left(\frac{1}{4\sqrt{\varepsilon_T}+2^{-|u|}}\right) - \left(2\sqrt{\varepsilon_T}+\frac{\delta}{4}\right)\cdot|u|. \tag{11}$$

[14] This assumption does not cause loss of generality: without it we would view T_u as a “multiset”.

Let us take a closer look at $\widehat{\Delta_\alpha^{u,i}}(\emptyset)$.

$$\begin{aligned}
\widehat{\Delta_\alpha^{u,i}}(\emptyset) &= \mathop{\mathbf{E}}_{Z \in \{0,1\}^{|u|}} \left[\Delta_\alpha^{u,i}(Z)\right] \\
&= \mathop{\mathbf{E}}_{Z}\left[\mathop{\mathbf{Pr}}_{X \in \alpha^*(Z)}\left[X_u(i) = Z_i \middle| X_u([i-1]) = Z_{[i-1]}\right] - \frac{1}{2}\right] \\
&= \mathop{\mathbf{E}}_{\substack{Z \\ \mathcal{A} \sim \alpha^*(Z)}}\left[\mathop{\mathbf{Pr}}_{X \in \mathcal{A}}\left[X_u(i) = Z_i \middle| X_u([i-1]) = Z_{[i-1]}\right] - \frac{1}{2}\right].
\end{aligned}$$

From the definition of $\alpha^*(\cdot)$ it is clear that the "chain"

$$Z \in \{0,1\}^{|u|} \;\to\; \mathcal{A} \sim \alpha^*(Z) \;\to\; X \in \mathcal{A}$$

results in the same distribution of $(Z, \mathcal{A}, X)$ as

$$X \in \{0,1\}^n \;\to\; \mathcal{A} = \alpha(X) \;\to\; X' \in \mathcal{A} \;\to\; Z = X'_u.$$

Therefore,

$$\widehat{\Delta_\alpha^{u,i}}(\emptyset) = \mathop{\mathbf{E}}_{X \in \{0,1\}^n}\left[\mathop{\mathbf{Pr}}_{X' \in \alpha(X)}\left[X_u(i) = X'_u(i) \middle| X_u([i-1]) = X'_u([i-1])\right] - \frac{1}{2}\right].$$

Moreover, the marginal distributions of $(\mathcal{A}, X)$ and of $(\mathcal{A}, X')$ are the same: we can sample $(X, \mathcal{A}, X')$ by first drawing $\mathcal{A}$ according to its distribution[15], followed by mutually-independent selecting $X \in \mathcal{A}$ and $X' \in \mathcal{A}$. Accordingly,

$$\begin{aligned}
\widehat{\Delta_\alpha^{u,i}}(\emptyset) &= \mathop{\mathbf{E}}_{\mathcal{A}}\left[\mathop{\mathbf{Pr}}_{\substack{X \in \mathcal{A} \\ X' \in \mathcal{A}}}\left[X_u(i) = X'_u(i) \middle| X_u([i-1]) = X'_u([i-1])\right] - \frac{1}{2}\right] \\
&= \mathop{\mathbf{E}}_{\substack{\mathcal{A} \\ X' \in \mathcal{A}}}\left[\mathop{\mathbf{Pr}}_{X \in \mathcal{A}}\left[X_u(i) = X'_u(i) \middle| X_u([i-1]) = X'_u([i-1])\right] - \frac{1}{2}\right] \\
&= \mathop{\mathbf{E}}_{\substack{\mathcal{A}' \\ X' \in \mathcal{A}'}}\left[\mathop{\mathbf{Pr}}_{\substack{\mathcal{A} \\ X \in \mathcal{A}}}\left[X_u(i) = X'_u(i) \middle| X_u([i-1]) = X'_u([i-1]), \mathcal{A} = \mathcal{A}'\right] - \frac{1}{2}\right],
\end{aligned}$$

where $\mathcal{A}'$ is distributed identically to $\mathcal{A}$.

Finally, let us denote $W = (\mathcal{A}, X_u([i-1]))$ and $W' = (\mathcal{A}', X'_u([i-1]))$, and apply Claim 3 with respect to W and $X_u(i)$:

$$\begin{aligned}
\widehat{\Delta_\alpha^{u,i}}(\emptyset) &= \mathop{\mathbf{E}}_{W', X'_u(i)}\left[\mathop{\mathbf{Pr}}_{W, X_u(i)}\left[X_u(i) = X'_u(i) \middle| W = W'\right] - \frac{1}{2}\right] \\
&\in \Theta\big(\mathbf{I}\left[X_u(i) : W\right]\big) \\
&= \Theta\big(\mathbf{I}\left[X_u(i) : \alpha(X), X_u([i-1])\right]\big) \\
&= \Theta\left(\mathop{\mathbf{I}}_{X \in \{0,1\}^n}\left[X_u(i) : Al(X), X_u([i-1])\right]\right).
\end{aligned}$$

[15] namely, the distribution where the probability of $\mathcal{A} = a$ is proportional to $|a|$

Applying similar reasoning to $\widehat{\Delta^{u,i}_{\beta}}(\emptyset)$ and plugging into (11) leads to

$$\sum_{i=1}^{|u|} \mathop{\mathbf{I}}_{X \in \{0,1\}^n} \big[X_u(i) : Al(X), X_u([i-1])\big] \cdot \mathop{\mathbf{I}}_{Y \in \{0,1\}^n} \big[Y_u(i) : Bo(Y), Y_u([i-1])\big]$$
$$\in \Omega\left(\ln\left(\frac{1}{\varepsilon_T + 2^{-|u|}}\right)\right) - O((\sqrt{\varepsilon_T} + \delta) \cdot |u|).$$

By monotonicity of mutual information,

$$\sum_{i \in u} \mathop{\mathbf{I}}_{X \in \{0,1\}^n} \big[X_i : Al(X), X_{u \setminus \{i\}}\big] \cdot \mathop{\mathbf{I}}_{Y \in \{0,1\}^n} \big[Y_i : Bo(Y), Y_{u \setminus \{i\}}\big]$$
$$\in \Omega\left(\ln\left(\frac{1}{\varepsilon_T + 2^{-|u|}}\right)\right) - O((\sqrt{\varepsilon_T} + \delta) \cdot |u|),$$

as required. $\blacksquare_{\textit{Lemma 4}}$

6.3 Characterising protocols for $\widetilde{Eq}$-neg_T

Lemma 5. *For sufficiently large n, some $\delta \in \Theta\big(\frac{1}{n}\big)$, any δ-biased space T of size $2^{o(n)}$ and some $\varepsilon \in \Theta\Big(\frac{1}{|T| \cdot n^2}\Big)$, any protocol $\mathcal{P}$ that solves $\widetilde{Eq}$-$neg_T(X,Y)$ in $\mathcal{D}^{||}_{\mu_{\widetilde{Eq}\text{-}neg_T},\varepsilon}$ satisfies*

$$\sum_{i=1}^{n} \mathop{\mathbf{E}}_{u_1}\left[\mathop{\mathbf{I}}_{X \in \{0,1\}^n} \big[X_i : X_{u_1}, Al(X)\big]\right] \cdot \mathop{\mathbf{E}}_{u_2}\left[\mathop{\mathbf{I}}_{Y \in \{0,1\}^n} \big[Y_i : Y_{u_2}, Bo(Y)\big]\right] > 1,$$

where $u_1, u_2 \in \binom{[n]\setminus\{i\}}{2n/3}$.

Proof. Suppose that a protocol solves $\widetilde{Eq}$-neg_T with respect to $\mu_{\widetilde{Eq}\text{-}neg_T}$ with error at most ε', and let ε'_u be the error that the same protocol makes in solving Eq_u-neg_T with respect to $\mu_{Eq_u\text{-}neg_T}$. By the definition of the input distributions,

$$\mathop{\mathbf{E}}_{u \in \binom{[n]}{n/3}} \big[\varepsilon'_u\big] \le \varepsilon' + 2^{-\Omega(n)}.$$

From Lemma 4, there exist choices of ε and δ in the range given by our statement, so that

$$\mathop{\mathbf{E}}_{u \in \binom{[n]}{n/3}}\left[\sum_{i \in u} \mathop{\mathbf{I}}_{X \in \{0,1\}^n} \big[X_i : Al(X), X_{u \setminus \{i\}}\big] \cdot \mathop{\mathbf{I}}_{Y \in \{0,1\}^n} \big[Y_i : Bo(Y), Y_{u \setminus \{i\}}\big]\right] \ge 2,$$

and therefore for sufficiently large n,

$$\mathop{\mathbf{E}}_{u_1, u_2 \in \binom{[n]}{2n/3}}\left[\sum_{i \in u_1 \cap u_2} \mathop{\mathbf{I}}_{X} \big[X_i : Al(X), X_{u_1 \cap u_2 \setminus \{i\}}\big] \cdot \mathop{\mathbf{I}}_{Y} \big[Y_i : Bo(Y), Y_{u_1 \cap u_2 \setminus \{i\}}\big]\right]$$
$$\ge \mathop{\mathbf{Pr}}_{u_1, u_2 \in \binom{[n]}{2n/3}} \big[|u_1 \cap u_2| \ge n/3\big] \cdot 2 > 1.$$

By the monotonicity of mutual information,

$$
\begin{aligned}
1 &< \mathop{\mathbf{E}}_{u_1,u_2 \Subset \binom{[n]}{2n/3}} \left[\sum_{i \in u_1 \cap u_2} \mathop{\mathbf{I}}_{X}\left[X_i : Al(X), X_{u_1 \setminus \{i\}}\right] \cdot \mathop{\mathbf{I}}_{Y}\left[Y_i : Bo(Y), Y_{u_2 \setminus \{i\}}\right] \right] \\
&\le \sum_{i=1}^{n} \mathop{\mathbf{E}}_{u_1,u_2 \Subset \binom{[n]\setminus\{i\}}{2n/3}} \left[\mathop{\mathbf{I}}_{X}\left[X_i : Al(X), X_{u_1}\right] \cdot \mathop{\mathbf{I}}_{Y}\left[Y_i : Bo(Y), Y_{u_2}\right] \right],
\end{aligned}
$$

as required. $\blacksquare_{\textit{Lemma 5}}$

6.4 Characterising protocols for $\widetilde{cEq}$-neg_T

Lemma 6. *For sufficiently large n, some $\delta \in \Theta\left(\frac{1}{n}\right)$, any δ-biased space T of size $2^{o(n)}$ and some $\varepsilon \in \Theta\left(\frac{1}{|T|\cdot n^2}\right)$, any protocol $\mathcal{P}$ that solves $\widetilde{cEq}$-$neg_T(X,Y)$ in $\mathcal{D}^{\|}_{\mu_{\widetilde{cEq}\text{-}neg_T},\varepsilon}$ satisfies*

$$
\mathop{\mathbf{E}}_{i_1,u_1} \left[\mathop{\mathbf{I}}_{X \Subset \{0,1\}^n} \left[X_{i_1} : X_{u_1}, Al(X)\right] \right] \cdot \mathop{\mathbf{E}}_{i_2,u_2} \left[\mathop{\mathbf{I}}_{Y \Subset \{0,1\}^n} \left[Y_{i_2} : Y_{u_2}, Bo(Y)\right] \right] > \frac{1}{2n},
$$

where $i_1, i_2 \Subset [n]$, $u_1 \Subset \binom{[n]\setminus\{i_1\}}{2n/3}$ and $u_2 \Subset \binom{[n]\setminus\{i_2\}}{2n/3}$.

Proof. Suppose that a protocol solves $\widetilde{cEq}$-neg_T with respect to $\mu_{\widetilde{cEq}\text{-}neg_T}$ with error at most ε'. By the definition of the input distributions, with probability at least $1/2$ with respect to $j \Subset [n]$, the same protocol solves $\widetilde{Eq}$-$neg_T(X, \sigma_j(Y))$ with respect to $(X, \sigma_j(Y)) \sim \mu_{\widetilde{Eq}\text{-}neg_T}$ with error at most $2\varepsilon' + 2^{-\Omega(n)}$. Accordingly, Lemma 5 implies that there exist choices of ε and δ in the range given by our statement, so that

$$
\mathop{\mathbf{E}}_{j \in [n]} \left[\sum_{i=1}^{n} \mathop{\mathbf{E}}_{u_1} \left[\mathop{\mathbf{I}}_{X \Subset \{0,1\}^n} \left[X_{\sigma_j(i)} : X_{\sigma_j(u_1)}, Al(X)\right] \right] \cdot \mathop{\mathbf{E}}_{u_2} \left[\mathop{\mathbf{I}}_{Y \Subset \{0,1\}^n} \left[Y_i : Y_{u_2}, Bo(Y)\right] \right] \right] > \frac{1}{2},
$$

where $u_1, u_2 \Subset \binom{[n]\setminus\{i\}}{2n/3}$. That is,

$$
\frac{1}{2n} < \mathop{\mathbf{E}}_{i_1,i_2 \Subset [n]} \left[\mathop{\mathbf{E}}_{u_1} \left[\mathop{\mathbf{I}}_{X}\left[X_{i_1} : X_{u_1}, Al(X)\right] \right] \cdot \mathop{\mathbf{E}}_{u_2} \left[\mathop{\mathbf{I}}_{Y}\left[Y_{i_2} : Y_{u_2}, Bo(Y)\right] \right] \right],
$$

where $u_1 \Subset \binom{[n]\setminus\{i_1\}}{2n/3}$ and $u_2 \Subset \binom{[n]\setminus\{i_2\}}{2n/3}$, as required. $\blacksquare_{\textit{Lemma 6}}$

Corollary 4. *There exists a family $\mathcal{T} = T_1, T_2, \ldots$, where every $T_i \subseteq \{0,1\}^i$ can be constructed deterministically in time* poly(i)*, such that for the corresponding $\widetilde{cEq}$-neg_T it holds that*

$$
\mathcal{R}^{\|,pub}(\widetilde{cEq}\text{-}neg_T) \ \ge\ \mathcal{D}^{\|}_{\mu_{\widetilde{cEq}\text{-}neg_T},\frac{1}{3}}(\widetilde{cEq}\text{-}neg_T) \ \in\ \Omega\left(\frac{\sqrt{n}}{\log n}\right).
$$

Proof. Let n be sufficiently large, $\delta \in \Theta\left(\frac{1}{n}\right)$ be sufficiently small, T be a δ-biased space of size poly(n/δ) (as guaranteed by Fact 1) and $\varepsilon \in \frac{1}{\text{poly}(n)}$ be sufficiently small, so that Lemma 6 guarantees that for any protocol $\mathcal{P}$ solving $\widetilde{cEq}$-neg_T in $\mathcal{D}^{\|}_{\mu_{\widetilde{cEq}\text{-}neg_T},\varepsilon}$ it holds that

$$
\mathop{\mathbf{E}}_{i_1,u_1} \left[\mathop{\mathbf{I}}_{X \Subset \{0,1\}^n} \left[X_{i_1} : X_{u_1}, Al(X)\right] \right] \cdot \mathop{\mathbf{E}}_{i_2,u_2} \left[\mathop{\mathbf{I}}_{Y \Subset \{0,1\}^n} \left[Y_{i_2} : Y_{u_2}, Bo(Y)\right] \right] > \frac{1}{2n}.
$$

Without loss of generality, assume that

$$\mathop{\mathbf{E}}_{i_1,u_1}\left[\mathop{\mathbf{I}}_{X\in\{0,1\}^n}\big[X_{i_1} : X_{u_1}, Al(X)\big]\right] > \frac{1}{2\sqrt{n}}$$

for $i_1 \in [n]$ and $u_1 \in \binom{[n]\setminus\{i_1\}}{2n/3}$, then

$$\exists\, u \in \binom{[n]}{2n/3} : \sum_{i\notin u} \mathop{\mathbf{I}}_{X\in\{0,1\}^n}\big[X_i : X_u, Al(X)\big] > \frac{n}{3}\cdot\frac{1}{2\sqrt{n}},$$

and therefore the complexity of $\mathcal{P}$ is at least

$$\mathop{\mathbf{I}}_{X\in\{0,1\}^n}\big[Al(X) : X\big|X_u\big] > \frac{\sqrt{n}}{6}.$$

If, on the other hand, a protocol solves $\widetilde{cEq}\text{-}neg_T$ in $\mathcal{D}^{\|}_{\mu_{\widetilde{cEq}\text{-}neg_T},\frac{1}{3}}$, then repeated k times in parallel for a sufficient $k \in O(\log n)$, it would solve $\widetilde{cEq}\text{-}neg_T$ with error at most ε. $\blacksquare_{\textit{Corollary 4}}$

7 Conclusion

From Corollaries 4 and 1:

Corollary 5. *There exists a family $\mathcal{T} = T_1, T_2, \ldots$, where every $T_i \subseteq \{0,1\}^i$ can be constructed deterministically in time* poly(i) *and for the corresponding $\widetilde{cEq}\text{-}neg_T$ it holds that*

$$\mathcal{Q}^{\|}(\widetilde{cEq}\text{-}neg_T) \in O\big((\log n)^2\big) \quad \textit{and} \quad \mathcal{R}^{\|,pub}(\widetilde{cEq}\text{-}neg_T) \in \Omega\left(\frac{\sqrt{n}}{\log n}\right).$$

Acknowledgements

I am very grateful to Pavel Pudlák, Ronald de Wolf and Thomas Vidick for helpful discussions at various stages of this work. I am especially thankful to Alexander Razborov for finding a mistake in the initial version of the proof of Claim 1.

References

[Aar04] S. Aaronson. Limitations of Quantum Advice and One-Way Communication. *Proceedings of the 19th IEEE Conference on Computational Complexity*, pages 320–332, 2004.

[BCWdW01] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf. Quantum Fingerprinting. *Physical Review Letters 87(16)*, 167902, 2001.

[BYJK04] 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.

[Gav08] D. Gavinsky. Classical Interaction Cannot Replace a Quantum Message. *Proceedings of the 40th Symposium on Theory of Computing*, pages 95–102, 2008.

[Gav16] D. Gavinsky. Entangled Simultaneity Versus Classical Interactivity in Communication Complexity. *Proceedings of the 48th Symposium on Theory of Computing*, pages 877–884, 2016.

[GKK+08] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf. Exponential Separations for One-Way Quantum Communication Complexity, with Applications to Cryptography. *SIAM Journal on Computing 38(5)*, pages 1695–1708, 2008.

[GKRdW06] D. Gavinsky, J. Kempe, O. Regev, and R. de Wolf. Bounded-Error Quantum State Identification and Exponential Separations in Communication Complexity. *Proceedings of the 38th Symposium on Theory of Computing*, pages 594–603, 2006.

[GRdW08] D. Gavinsky, O. Regev, and R. de Wolf. Simultaneous Communication Protocols with Quantum and Classical Messages. *Chicago Journal of Theoretical Computer Science*, 7, 2008.

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

[KR11] B. Klartag and O. Regev. Quantum One-Way Communication Can Be Exponentially Stronger than Classical Communication. *Proceedings of the 43st Symposium on Theory of Computing*, pages 31–40, 2011.

[NN93] J. Naor and M. Naor. Small-Bias Probability Spaces: Efficient Constructions and Applications. *SIAM Journal on Computing 22(4)*, pages 838–856, 1993.

[NS96] I. Newman and M. Szegedy. Public vs. Private Coin Flips in One Round Communication Games. *Proceedings of the 28th Symposium on Theory of Computing*, pages 561–570, 1996.

[Raz99] R. Raz. Exponential Separation of Quantum and Classical Communication Complexity. *Proceedings of the 31st Symposium on Theory of Computing*, pages 358–367, 1999.