Quantum versus classical simultaneity in communication complexity

Dmitry Gavinsky*

June 2, 2017

Abstract

We present a bipartite partial function $\widetilde{cEq}\text{-}neg_T$, whose communication complexity is $O\big((\log n)^2\big)$ in the model of *quantum simultaneous message passing* ($\mathcal{Q}^{\|,pub}$) and $\tilde{\Omega}(\sqrt{n})$ in the model of *randomised simultaneous message passing* ($\mathcal{R}^{\|,pub}$).

In fact, our function has a poly-logarithmic protocol even in the (restricted) model of *quantum simultaneous message passing without shared randomness* ($\mathcal{Q}^{\|}$), thus witnessing the possibility of qualitative advantage of $\mathcal{Q}^{\|}$ over $\mathcal{R}^{\|,pub}$. This can be interpreted as the strongest known – as of today – example of "super-classical" capabilities of the weakest studied model of quantum communication.

The closest previously-known result was given by Buhrman, Cleve, Watrous and de Wolf in 2001: they have shown that the equality function had a protocol of logarithmic cost in $\mathcal{Q}^{\|}$, while its complexity in the model of *classical simultaneous message passing without shared randomness* ($\mathcal{R}^{\|}$) had already been known to be in $\Omega(\sqrt{n})$. $\mathcal{R}^{\|}$ and $\mathcal{Q}^{\|}$ are "purposely weakened" implementations of simultaneous message passing in communication complexity, which are not closed with respect to mixed strategies ($\mathcal{R}^{\|,pub}$ and $\mathcal{Q}^{\|,pub}$ can be viewed as the respective "closures").

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*. This subjective judgement can be supported by the following empirical evidence: as of today, communication complexity is one of very few computational scenarios where both upper and (non-speculative) lower bounds play central roles in the research.

The regime of *functional problems with bipartite input* is the most commonly studied and, probably, the most natural case of communication complexity. The model of *simultaneous message passing (SMP)* is one of the weakest commonly studied.

*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.

The most standard quantum and classical implementations of SMP are $\mathcal{Q}^{\|,pub}$ and $\mathcal{R}^{\|,pub}$, respectively. Prior to this work it was not known whether the former was qualitatively stronger than the latter.

"Purposely weakened" $\mathcal{Q}^{\|}$ and $\mathcal{R}^{\|}$ are, probably, the weakest commonly studied quantum and randomised models, respectively. These are implementations of SMP *without shared randomness*; in particular, the families of efficiently-computable tasks for $\mathcal{Q}^{\|}$ and for $\mathcal{R}^{\|}$ are not closed with respect to mixed strategies.[1] Prior to this work, $\mathcal{Q}^{\|}$ was known to be stronger than $\mathcal{R}^{\|}$ due to [BCWdW01] [2], and it was open whether $\mathcal{Q}^{\|}$ could also "outperform" the "unrestricted" classical SMP – $\mathcal{R}^{\|,pub}$.

Our results

The two most intuitive approaches towards demonstrating the strength of quantum communication are, probably, looking for the *weakest quantum model that can outperform the strongest studied classical one* and for the *strongest classical model that can be outperformed by the weakest studied quantum one*. This work takes the latter approach and investigates the power of the "weakest commonly studied" quantum class $\mathcal{Q}^{\|}$ and its "stronger sibling" $\mathcal{Q}^{\|,pub}$.

We prove that $\mathcal{Q}^{\|,pub}$ is qualitatively stronger than $\mathcal{R}^{\|,pub}$, demonstrating a partial function $\widetilde{cEq}\text{-}neg_T$, whose $\mathcal{Q}^{\|,pub}$-complexity is $O\big((\log n)^2\big)$ and $\mathcal{R}^{\|,pub}$-complexity is $\Omega(\sqrt{n}/\log n)$. Moreover, $\widetilde{cEq}\text{-}neg_T$ has a protocol of complexity $O\big((\log n)^2\big)$ even in $\mathcal{Q}^{\|}$, which means that this "weakest studied" quantum model can be exponentially more efficient than $\mathcal{R}^{\|,pub}$ in solving a functional problem.

Proving qualitative advantage of quantum communication via separating a pair of models is an exciting activity; even more interesting probably is showing *impossibility* of separation for some models. So far, virtually no such impossibility results are known for the models considered in this work; nevertheless, many researchers may agree, for instance, that $\mathcal{Q}^{\|}$ is *unlikely* to have super-polynomial advantage over $\mathcal{R}$ with respect to a partial function, and even less so with respect to a total one. We will discuss this and related questions in Section 7, now we only mention that it *may* be the case that the results of this work bring our collection of possible separations close to completeness – which would be both aesthetically rewarding and potentially helpful in formulating plausible "non-separability" conjectures.

The questions considered in this work have remained open for some time, so one could hope that an interesting approach would be required to address them. Indeed, in this work we are demonstrating $\mathcal{R}^{\|,pub}$-hardness of a communication problem that is easy for virtually any model stronger than $\mathcal{R}^{\|,pub}$ – so, our argument had to be rather accurately "tuned" in order to be able to distinguish between $\mathcal{R}^{\|,pub}$ and "anything above it". The hardness of the communication problem itself had to be "tuned" quite precisely as well, to make it easy for $\mathcal{Q}^{\|}$ but hard for "just slightly weaker" [3] $\mathcal{R}^{\|,pub}$: e.g., we could not use a communication problem

[1] We call *efficient* communication protocols of poly-logarithmic complexity. For example, the *equality function (Eq)* has $\mathcal{R}^{\|}$-complexity $O(1)$ over any fixed input distribution, but its worst-case complexity is $\Omega(\sqrt{n})$ (thus "violating" the minimax principle). The "unrestricted" randomised SMP – $\mathcal{R}^{\|,pub}$ can be defined as the "closure" of $\mathcal{R}^{\|}$ with respect to mixed strategies, and similarly for $\mathcal{Q}^{\|,pub}$ and $\mathcal{Q}^{\|}$.

[2] In 2001, Buhrman, Cleve, Watrous and de Wolf demonstrated qualitative advantage of $\mathcal{Q}^{\|}$: they proved that there existed a $\mathcal{Q}^{\|}$-protocol for *Eq* of complexity $\Omega(\log n)$, and it was already known due to [NS96] that $\mathcal{R}^{\|}(Eq) \in \Omega(\sqrt{n})$.

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

with "worst-case hardness in spite of average-case easiness" (like *Eq*), as $\mathcal{R}^{\|,pub}$ allowed mixed strategies.

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, $\{x_i | x \in A\}$ and $\{x_S | x \in A\}$. 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 $\exp(A)$ the set of A's subsets and by $\binom{A}{k}$ the set $\{a \in \exp(A) \,|\, |a| = k\}$. We write "$A \triangle B$" to denote the symmetric difference between the two sets and "$A \cup B$" to denote the union of two set that must be disjoint (i.e., using this notation implies that $A \cap B = \emptyset$).

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 $i + j \leq n$ 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 $\langle f\,,\, g\rangle \stackrel{\text{def}}{=} 2^{-n} \cdot \sum_{x \in \{0,1\}^n} f(x) \cdot g(x) = \mathbf{E}_{X \in \{0,1\}^n}[f(X) \cdot g(X)]$. 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 $\mathbb{R}^n$ 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 $\varepsilon \geq 0$, we call $T \subseteq \{0,1\}^n$ an ε-bias space if*

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

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

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 $\tau_1 \neq \tau_2$ of the constructed space satisfy $|\tau_1 \oplus \tau_2| \in \frac{n}{2} \pm o(n)$.*

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 $(X, Y) \sim \mu$.[4]

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 partial function that can be viewed as a variation of the *equality* problem.

The *equality function* (viewed as a communication problem) is the following total 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 $\text{poly}(n)$ (in our analysis T will be a

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

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 $\mu_{Eq_u\text{-}neg_T} \stackrel{\text{def}}{=} \frac{1}{|T|} \cdot \sum_{\tau \in T} \mu^{\tau}_{Eq_u\text{-}neg_T}$.

Next we define a "noisy" (or gapped) version of $Eq\text{-}neg_T$, which is a partial function:

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

Intuitively, $\widetilde{Eq}\text{-}neg_T(x,y)$ "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 $\widetilde{Eq}\text{-}neg_T$:

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

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 $(X,Y) \sim \mu_{\widetilde{Eq}\text{-}neg_T}$ satisfies the promise of $\widetilde{Eq}\text{-}neg_T(X,Y)$ with probability $1 - 2^{-\Omega(n)}$. [5]

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

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

The intuition behind this definition is very similar to that behind $\widetilde{Eq}\text{-}neg_T(x,y)$, 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 $\widetilde{cEq}\text{-}neg_T$:

[5] In particular, under $(X,Y) \sim \mu_{\widetilde{Eq}\text{-}neg_T}$ 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 $1 - 2^{-\Omega(n)}$.

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

Like in the case of $\widetilde{Eq}\text{-}neg_T$, in our constructions T will always be a set with the "minimal distance property". In such case a pair $(X,Y) \sim \mu_{\widetilde{cEq\text{-}neg}_T}$ satisfies the promise of $\widetilde{cEq}\text{-}neg_T(X,Y)$ with probability $1-2^{-\Omega(n)}$. [6]

3 The $\mathcal{Q}^{\|}$-complexity of $\widetilde{cEq}$-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

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

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

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

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 $1/\mathrm{poly}(n\cdot|T|)$.[7] 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 $T \subseteq \{0,1\}^n$,*

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

4 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.

[6] In particular, under $(X,Y) \sim \mu_{\widetilde{cEq\text{-}neg}_T}$ 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 $1-2^{-\Omega(n)}$.

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

4.1 Optimistic inequalities

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

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{\mu}\left[\bigwedge_{j=1}^{m} X_j \in G_j\right] &= \prod_{i=1}^{m} \mathop{\mathbf{Pr}}_{\mu}\left[X_i \in G_i \middle| \bigwedge_{j=1}^{i-1} X_j \in G_j\right] \\
&= \prod_{i=1}^{m} \mathop{\mathbf{E}}_{(X'_1,\ldots,X'_m)\sim\mu}\left[\mathop{\mathbf{Pr}}_{\mu}\left[X_i \in G_i \middle| \bigwedge_{j=1}^{i-1} X_j = X'_j\right] \middle| \bigwedge_{j=1}^{\mathbf{i-1}} X'_j \in G_j\right] \\
&\leq \prod_{i=1}^{m} \mathop{\mathbf{E}}_{(X'_1,\ldots,X'_m)\sim\mu}\left[\mathop{\mathbf{Pr}}_{\mu}\left[X_i \in G_i \middle| \bigwedge_{j=1}^{i-1} X_j = X'_j\right] \middle| \bigwedge_{j=1}^{\mathbf{m}} X'_j \in G_j\right].
\end{aligned}$$

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

We call the inequality *optimistic*, viewing the subsets G_i as *good*, B_i-s as *bad* and interpreting the statement 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 $[\bigwedge_{j=1}^{\mathbf{i-1}} X'_j \in G_j]$ (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 $[\bigwedge_{j=1}^{\mathbf{m}} X'_j \in G_j]$.

The inequality isn't, in general, true "element-wise": one can construct an example, where

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{\mu}\left[X_{i_0} \in G_{i_0} \middle| \bigwedge_{j=1}^{i_0-1} X_j \in G_j\right] &= \mathop{\mathbf{E}}_{X'_1,\ldots,X'_m}\left[\mathop{\mathbf{Pr}}_{\mu}\left[X_{i_0} \in G_{i_0} \middle| \bigwedge_{j=1}^{i_0-1} X_j = X'_j\right] \middle| \bigwedge_{j=1}^{i_0-1} X'_j \in G_j\right] \quad (1) \\
&> \mathop{\mathbf{E}}_{X'_1,\ldots,X'_m}\left[\mathop{\mathbf{Pr}}_{\mu}\left[X_{i_0} \in G_{i_0} \middle| \bigwedge_{j=1}^{i_0-1} X_j = X'_j\right] \middle| \bigwedge_{j=1}^{m} X'_j \in G_j\right]
\end{aligned}$$

for some $i_0 \in [m]$.

Our proof of Claim 1 uses the following core observation.

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

$$\mathop{\mathbf{Pr}}_{\mu}\left[X_2 \in G_2 \middle| X_1 \in G_1\right] \leq \mathop{\mathbf{E}}_{(X'_1,X'_2)\sim\mu}\left[\mathop{\mathbf{Pr}}_{\mu}\left[X_2 \in G_2 \middle| X_1 = X'_1\right] \middle| X'_1 \in G_1, X'_2 \in G_2\right].$$

Note that Claim 2 does not contradict what we are saying in (1), it only implies that for that inequality to hold, one must take $i_0 < m$.

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

$$\mathop{\mathbf{Pr}}_{\mu}\left[X_2 \in G_2 \middle| X_1 \in G_1\right] = \frac{\mathbf{Pr}\left[X_1 \in G_1, X_2 \in G_2\right]}{\mathbf{Pr}\left[X_1 \in G_1\right]} = \frac{\sum_{c\in G_1} p_c \cdot q_c}{\sum_{c\in G_1} p_c}$$

and

$$\begin{aligned}
&\mathop{\mathbf{E}}_{(X_1',X_2')\sim\mu}\left[\mathop{\mathbf{Pr}}_{\mu}\left[X_2\in G_2\middle|X_1=X_1'\right]\middle|X_1'\in G_1,X_2'\in G_2\right]\\
&\qquad=\mathbf{Pr}\left[X_2\in G_2\middle|X_1=X_1',X_1'\in G_1,X_2'\in G_2\right]\\
&\qquad=\frac{\mathbf{Pr}\left[X_1'\in G_1,X_2\in G_2,X_2'\in G_2\middle|X_1=X_1'\right]}{\mathbf{Pr}\left[X_1'\in G_1,X_2'\in G_2\middle|X_1=X_1'\right]}=\frac{\sum_{G_1}p_c\cdot q_c\cdot q_c}{\sum_{G_1}p_c\cdot q_c}.
\end{aligned}$$

We want to see that

$$\frac{\sum_c p_c\cdot q_c}{\sum_c p_c}\le\frac{\sum_c p_c\cdot q_c\cdot q_c}{\sum_c p_c\cdot q_c}.$$

Let us define three $|G_1|$-dimensional vectors (indexed by $c\in G_1$):

$$\forall c\in G_1:\quad a_c\stackrel{\text{def}}{=}\frac{p_c}{\sum_{c'}p_{c'}},\quad b_c\stackrel{\text{def}}{=}\frac{p_c\cdot q_c}{\sum_{c'}p_{c'}\cdot q_{c'}},\quad v_c\stackrel{\text{def}}{=}\frac{b_c}{a_c}.$$

Note that the vectors (a_c) and (b_c) are probability distributions.

As $v_c\sim q_c$, we want to see that

$$\sum_c a_c\cdot v_c\le\sum_c b_c\cdot v_c.$$

As $\sum_c a_c\cdot v_c=\sum_c b_c=1$, we want to see that

$$\sum_c b_c\cdot v_c=\sum_c b_c\cdot\frac{b_c}{a_c}\ge 1.$$

That follows from a repeated application of Jensen's inequality with respect to the *concave* function $\log(\cdot)$:

$$\log\left(\sum_c b_c\cdot\frac{b_c}{a_c}\right)\ge\sum_c b_c\cdot\log\frac{b_c}{a_c}=-\sum_c b_c\cdot\log\frac{a_c}{b_c}\ge-\log\left(\sum_c b_c\cdot\frac{a_c}{b_c}\right)=0.$$

■*Claim 2*

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

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{\nu}\left[Y_1\in\mathcal{G}_1,Y_2\in\mathcal{G}_2\right]&=\mathbf{Pr}\left[Y_1\in\mathcal{G}_1\right]\cdot\mathbf{Pr}\left[Y_2\in\mathcal{G}_2\middle|Y_1\in\mathcal{G}_1\right]\qquad(2)\\
&\le\mathbf{Pr}\left[Y_1\in\mathcal{G}_1\right]\cdot\mathop{\mathbf{E}}_{(Y_1',Y_2')\sim\nu}\left[\mathbf{Pr}\left[Y_2\in\mathcal{G}_2\middle|Y_1=Y_1'\right]\middle|Y_1'\in\mathcal{G}_1,Y_2'\in\mathcal{G}_2\right],
\end{aligned}$$

as follows from Claim 2.

Now the result follows by induction:

$$\mathop{\mathbf{Pr}}_{\mu}\left[\bigwedge_{j=1}^{m}X_j\in G_j\right]=\mathbf{Pr}\left[\bigwedge_{j=1}^{m-1}X_j\in G_j\right]\cdot\mathbf{Pr}\left[X_m\in G_m\middle|\bigwedge_{j=1}^{m-1}X_j\in G_j\right]$$

$$
\begin{aligned}
&\leq \mathbf{Pr}\left[\bigwedge_{j=1}^{m-1} X_j \in G_j\right] \\
&\quad \cdot \mathop{\mathbf{E}}_{(X'_1,\ldots,X'_m)\sim\mu}\left[\mathbf{Pr}\left[X_m \in G_m \middle| \bigwedge_{j=1}^{m-1} X_j = X'_j\right] \middle| \bigwedge_{j=1}^{m} X'_j \in G_j\right] \\
&\cdots \\
&\leq \mathbf{Pr}\left[\bigwedge_{j=1}^{k} X_j \in G_j\right] \\
&\quad \cdot \prod_{i=k+1}^{m} \mathop{\mathbf{E}}_{(X'_1,\ldots,X'_m)\sim\mu}\left[\mathop{\mathbf{Pr}}_{\mu}\left[X_i \in G_i \middle| \bigwedge_{j=1}^{i-1} X_j = X'_j\right] \middle| \bigwedge_{j=1}^{m} X'_j \in G_j\right] \\
&= \mathbf{Pr}\left[\bigwedge_{j=1}^{k-1} X_j \in G_j\right] \cdot \mathbf{Pr}\left[X_k \in G_k \middle| \bigwedge_{j=1}^{k-1} X_j \in G_j\right] \\
&\quad \cdot \prod_{i=k+1}^{m} \mathop{\mathbf{E}}_{(X'_1,\ldots,X'_m)\sim\mu}\left[\mathop{\mathbf{Pr}}_{\mu}\left[X_i \in G_i \middle| \bigwedge_{j=1}^{i-1} X_j = X'_j\right] \middle| \bigwedge_{j=1}^{m} X'_j \in G_j\right] \\
&\leq \mathbf{Pr}\left[\bigwedge_{j=1}^{k-1} X_j \in G_j\right] \\
&\quad \cdot \prod_{i=k}^{m} \mathop{\mathbf{E}}_{(X'_1,\ldots,X'_m)\sim\mu}\left[\mathop{\mathbf{Pr}}_{\mu}\left[X_i \in G_i \middle| \bigwedge_{j=1}^{i-1} X_j = X'_j\right] \middle| \bigwedge_{j=1}^{m} X'_j \in G_j\right] \\
&\cdots \\
&\leq \prod_{i=1}^{m} \mathop{\mathbf{E}}_{(X'_1,\ldots,X'_m)\sim\mu}\left[\mathop{\mathbf{Pr}}_{\mu}\left[X_i \in G_i \middle| \bigwedge_{j=1}^{i-1} X_j = X'_j\right] \middle| \bigwedge_{j=1}^{m} X'_j \in G_j\right],
\end{aligned}
$$

where the inequalities are applications of (2). ■*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) $\mathcal{G}_i \cup \mathcal{B}_i$ and for all $i_1 < i_2$ it holds that $\mathcal{G}_{i_2}([i_1]) \subseteq \mathcal{G}_{i_1}$. Let μ be the joint distribution of $(X_j)_1^m$, then*

$$
\mathop{\mathbf{Pr}}_{\mu}\left[(X_j)_1^m \in \mathcal{G}_m\right] \leq \prod_{i=1}^{m} \mathop{\mathbf{E}}_{(X'_j)_1^m\sim\mu}\left[\mathop{\mathbf{Pr}}_{(X_j)_1^m\sim\mu}\left[(X_j)_1^i \in \mathcal{G}_i \middle| (X_j)_1^{i-1} = (X'_j)_1^{i-1}\right] \middle| (X'_j)_1^m \in \mathcal{G}_m\right].
$$

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

$$Q_i = \begin{cases} 1 & \text{if } (X_j)_1^i \in \mathcal{G}_i; \\ 0 & \text{otherwise.} \end{cases}$$

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

4.2 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

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

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 $X \sim \mathcal{U}_{\{0,1\}}$.[8] Note that the above statement demonstrates qualitative difference between this quantity and the standard notion of *confidence*:

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

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

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

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

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

as required. ■*Claim 3*

[8] 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'.

5 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 $\alpha : \{0,1\}^n \to \exp(\{0,1\}^n)$ be the "neighbourhood function" corresponding to $Al(\cdot)$ – i.e., $\alpha(x) \stackrel{\text{def}}{=} \{x' | Al(x') = Al(x)\}$;
- 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 even more symmetric criterion (cf. Lemma 6) and then to the desired complexity lower bound (cf. Corollary 4).

5.1 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 $X = Y \in_R \{0,1\}^n$, the pair of subsets $(\alpha(X), \beta(Y))$ will (typically) be such that $[X = Y]$ is "likely", given the messages – namely,

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

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 $\left\{x' \in \{0,1\}^n \middle| x'_u = z\right\}$;
- denote by $\alpha^*(z)$ the *distribution* over $\exp(\{0,1\}^n)$, corresponding to $\left\{x' \middle| Al(x') = m_0\right\}$ 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 $\left\{x' \in \{0,1\}^n \middle| x'_u = z\right\}$);
- 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*

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

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

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

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

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 $(X,Y) \in A \times B$ conditioned on $[X_u = Y_u]$.

Proof. By the definition of good rectangles,

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

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

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

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

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

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

Therefore,

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

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

So,

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

as required. ■*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 $\exp(\{0,1\}^n)$-valued variables $(\mathcal{A}, \mathcal{B})$ be distributed as $(\alpha^*(Z), \beta^*(Z))$.
- Let Z' be distributed as X_u when $(X, Y) \circledcirc \mathcal{A} \times \mathcal{B}$ 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.*

In other words, the variables Z and Z' are "siblings" indeed.

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

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

On the other hand,

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

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

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

Therefore,

$$\begin{aligned}\mathbf{Pr}\left[[(a,b)]\wedge Z=z_0\right] &= \mathbf{Pr}\left[Z=z_0\right]\cdot\mathbf{Pr}\left[(a,b)\middle|Z=z_0\right] \qquad (3)\\ &= \mathop{\mathbf{Pr}}_{X\Subset a}\left[X_u=z_0\right]\cdot\mathop{\mathbf{Pr}}_{Y\Subset b}\left[Y_u=z_0\right]\cdot|a|\cdot|b|\cdot 2^{|u|-2n}\end{aligned}$$

and

$$\begin{aligned}\mathbf{Pr}\left[(a,b)\right] &= \sum_z \mathbf{Pr}\left[Z=z\right]\cdot\mathop{\mathbf{Pr}}_{X\Subset a}\left[X_u=z\right]\cdot\mathop{\mathbf{Pr}}_{Y\Subset b}\left[Y_u=z\right]\cdot|a|\cdot|b|\cdot 2^{2|u|-2n} \qquad (4)\\ &= \mathop{\mathbf{Pr}}_{\substack{X\Subset a\\ Y\Subset b}}\left[X_u=Y_u\right]\cdot|a|\cdot|b|\cdot 2^{|u|-2n}.\end{aligned}$$

On the other hand,

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

where the last two inequalities follow from (4) and (3), respectively. ■*Lemma 2*

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

Lemma 3.

$$\mathbf{Pr}\left[\mathcal{A}\times\mathcal{B} \text{ is a good rectangle}\right] > 1-2\varepsilon-2\sqrt{\varepsilon}.$$

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

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

Finally,

$$\underset{(X,Y)\in\{0,1\}^{n+n}}{\mathbf{Pr}}\left[X_u = Y_u \middle| (a,b)\right] = \frac{\mathbf{Pr}\left[(a,b)\middle| X_u = Y_u\right]}{\mathbf{Pr}\left[(a,b)\right]} \cdot \mathbf{Pr}\left[X_u = Y_u\right] \tag{5}$$
$$> \frac{1}{\frac{1}{2^{|u|}} + \frac{\sqrt{\varepsilon}}{1-\sqrt{\varepsilon}}} \cdot \frac{1}{2^{|u|}} \geq \frac{1}{4\sqrt{\varepsilon} + 2^{-|u|}} \cdot \frac{1}{2^{|u|}},$$

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*

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

Proof. We analyse the quantity

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

On the one hand,

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

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

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

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

5.2 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 $|T_u| = |T|$.[9]

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,

$$\sum_{i=1}^{|u|}\widehat{\Delta^{u,i}_{\alpha}}(\emptyset)\cdot\widehat{\Delta^{u,i}_{\beta}}(\emptyset) > \frac{1}{4}\cdot\ln\left(\frac{1}{4\sqrt{\varepsilon_T}+2^{-|u|}}\right) - \left(2\sqrt{\varepsilon_T}+\frac{\delta}{4}\right)\cdot|u|. \tag{6}$$

[9] 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 $\widehat{\Delta_\alpha^{u,i}}(\emptyset)$.

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

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

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

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

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

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[10], followed by mutually-independent selecting $X\Subset\mathcal{A}$ and $X'\Subset\mathcal{A}$. Accordingly,

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

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 $X_u(i)$:

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

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

Applying similar reasoning to $\widehat{\Delta_{\beta}^{u,i}}(\emptyset)$ and plugging into (6) leads to

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

By monotonicity of mutual information,

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

as required. ■*Lemma 4*

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

Lemma 5. *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 $\widetilde{Eq}$-$neg_T(X, Y)$ in $\mathcal{D}^{\|}_{\mu_{\widetilde{Eq}\text{-}neg_T}, \varepsilon}$ satisfies*

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

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,

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

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

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

and therefore for sufficiently large n,

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

By monotonicity of mutual information,

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

as required. ■*Lemma 5*

5.4 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*

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

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

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 $j \in [n]$, 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

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

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

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

where $u_1 \in \binom{[n] \setminus \{i_1\}}{2n/3}$ and $u_2 \in \binom{[n] \setminus \{i_2\}}{2n/3}$, as required. ■*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 $\mathrm{poly}(i)$, such that for the corresponding $\widetilde{cEq}$-neg_T it holds that*

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

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 $\mathrm{poly}(n/\delta)$ (as guaranteed by Fact 1) and $\varepsilon \in \frac{1}{\mathrm{poly}(n)}$ 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

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

Without loss of generality, assume that

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

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

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

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

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

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. ■*Corollary 4*

6 Summary of the new separation

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*

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

6.1 Intuition behind our choice of communication problem

Recall that we were looking for a functional communication problem, easy for $\mathcal{Q}^{\|,pub}$ (even better, for $\mathcal{Q}^{\|}$) but hard for $\mathcal{R}^{\|,pub}$. The initial inspiration came from the observation that the natural quantum SMP protocol for *equality with gap ($\widetilde{Eq}$)* had certain "robustness" that seemed impossible to have 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]

[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 $1/2 - \Omega(1)$.

Note that the 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).

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

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

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. The main communication task considered in this work – $\widetilde{cEq}\text{-}neg_T$ – 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.

6.2 Intuition behind our lower-bound argument

Ignoring many technical details, our lower-bound argument can be informally outlined as follows.

First of all, we needed a convenient characterisation of efficient protocols for $\widetilde{Eq}$. In our proof it was based on the observation that if random input satisfying $X \approx Y$ was given to a 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 leads to

$$\mathop{\mathbf{E}}_{i}\left[\Delta_\alpha^i \cdot \Delta_\beta^i\right] \in \Omega\left(\frac{1}{n}\right), \tag{7}$$

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

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

$$\mathop{\mathbf{E}}_{i}\left[\mathbf{I}\big[X_i : Al(X)\big] \cdot \mathbf{I}\big[Y_i : Bo(Y)\big]\right] \in \Omega\left(\frac{1}{n}\right). \tag{8}$$

The bound in (8) is significantly stronger than that in (7): 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 $\Theta\big(\gamma^2\big)$ bits of information. The "quadratic improvement" from (7) to (8) 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, which allow us to "disconnect" $\mathbf{I}\big[X_i : Al(X)\big]$ from $\mathbf{I}\big[Y_i : Bo(Y)\big]$. We show that any protocol for $\widetilde{cEq}\text{-}neg_T(X,Y)$ must satisfy

$$\mathop{\mathbf{E}}_{i}\Big[\mathbf{I}\big[X_i : Al(X)\big]\Big] \cdot \mathop{\mathbf{E}}_{j}\Big[\mathbf{I}\big[Y_j : Bo(Y)\big]\Big] \in \Omega\left(\frac{1}{n}\right), \tag{9}$$

and this gives the desired lower bound, as at least one of $\mathbf{E}_i\Big[\mathbf{I}\big[X_i : Al(X)\big]\Big]$ and $\mathbf{E}_j\Big[\mathbf{I}\big[Y_j : Bo(Y)\big]\Big]$ must be $\Omega(1/\sqrt{n})$ in order to satisfy (9).

7 The landscape of quantum superiority

Somewhat surprisingly, a "practical" concept of *quantum computers* has inspired a number of interesting and non-trivial theoretical questions in communication complexity: In 1997 (the year when [KN97] was published) many researchers would probably have agreed that we "mostly understood" the bipartite communication complexity in the "range" up to $\mathcal{R}$ – the model of interactive randomised communication. Today it is clear that there exists a noticeable gap in our understanding; while this gap has been "witnessed" by questions in the context of quantum communication, there is no reason to believe that the underlying lack of understanding is "inherently quantum" – it is more plausible to assume that we miss something important and fundamental about communication complexity per se.

One of the most common questions asked in the context of quantum communication is "*Can quantum outperform classical?*" – formally, for which pairs of quantum and classical communication complexity classes the former is super-polynomially[12] more efficient than the latter in solving a specific problem.

There are three main *types* of communication problems used for separations: *total functions*, *partial functions* and *relations*. Partial functions are a special case of relations and a generalisation of total functions; accordingly, these types form a "hierarchy": separations via total functions are the strongest ("most convincing"), separations via relations are the weakest (and usually the easiest to analyse) and separations via partial functions are "in between". There are known cases where a quantum communication complexity class can be separated from a classical one via a relation, while a functional separation is provably impossible (see [Aar04, GRdW08]).

The history of (super-polynomial) separations that showed advantage of quantum communication can be briefly outlined as follows.

- In 1999 Raz [Raz99] demonstrated a *partial function* that had an efficient *quantum two-way protocol*, but no efficient *classical two-way protocol*.
- In 2001 Buhrman, Cleve, Watrous and de Wolf [BCWdW01] demonstrated a *total function* (namely, the equality) that had an efficient *quantum simultaneous-messages protocol without shared randomness*, but no efficient *classical simultaneous-messages protocol without shared randomness*.
- In 2004 Bar-Yossef, Jayram and Kerenidis [BYJK04] demonstrated a *relation* that had an efficient *quantum simultaneous-messages protocol without shared randomness*, but no efficient *classical one-way protocol*.

[12] All known super-polynomial separations are, in fact, exponential.

- In 2008 in a joint work with Kempe, Kerenidis, Raz and de Wolf [GKK+08] a *partial function* was demonstrated, that had an efficient *quantum one-way protocol*, but no efficient *classical one-way protocol*.
- In 2008 a *relation* was demonstrated [Gav08] with an efficient *quantum one-way protocol*, but no efficient *classical two-way protocol*.
- In 2010 Klartag and Regev [KR11] demonstrated a *partial function* with an efficient *quantum one-way protocol*, but no efficient *classical two-way protocol*.
- In 2016 a *partial function* was demonstrated [Gav16] with an efficient *quantum simultaneous-messages protocol with entanglement*, but no efficient *classical two-way protocol*.
- This work presents a *partial function* with an efficient *quantum simultaneous-messages protocol without shared randomness*, but no efficient *classical simultaneous-messages protocol with shared randomness*.

Is it the case that (almost) "everything separable" has already been discovered – in other words, that for the (most of) quantum-classical pairs where we do not yet have an example of "quantum superiority", such examples do not exist? Our current knowledge of "limitations to separability" is very limited (in particular, virtually nothing is known regarding the communication models considered in this work).

To summarise what is known and what is still missing, let us consider the three "canonical" randomised models: two-way ($\mathcal{R}$), one-way ($\mathcal{R}^1$) and SMP ($\mathcal{R}^{\|,pub}$), and add to our picture the "purposely weakened" SMP ($\mathcal{R}^{\|}$). We are interested in their "strength relationship" with the quantum counterparts – both the "closest" (e.g., $\mathcal{R}$ vs. $\mathcal{Q}$) and "topologically weaker" (e.g., $\mathcal{R}$ vs. $\mathcal{Q}^{\|}$).

- If we allow *partial functions* and only consider the *closest pairs*, then our knowledge is now complete:

$$\begin{array}{ccccccc} \mathcal{R}^{\|} & < & \mathcal{R}^{\|,pub} & < & \mathcal{R}^{1} & < & \mathcal{R} \\ \wedge & & \wedge & & \wedge & & \wedge \\ \mathcal{Q}^{\|} & < & \mathcal{Q}^{\|,pub} & < & \mathcal{Q}^{1} & < & \mathcal{Q} \end{array}$$

  (we have just seen that $\mathcal{Q}^{\|,pub} > \mathcal{R}^{\|,pub}$, the rest has been known for some time).
- As for *"diagonal" relationship* via *partial functions*, it has been known that $\mathcal{Q}^1$ could be stronger than $\mathcal{R}$ and we have just seen that $\mathcal{Q}^{\|}$ could be stronger than $\mathcal{R}^{\|,pub}$.

  Question 1. *Can some of* $\left\{\mathcal{Q}^{\|}, \mathcal{Q}^{\|,pub}\right\}$ *be stronger than some of* $\left\{\mathcal{R}^1, \mathcal{R}\right\}$ *with respect to a partial function?*

- If we allow *relational problems*, then one additional *"diagonal"* separation is known: $\mathcal{Q}^{\|}$ can be stronger than $\mathcal{R}^1$.

  Question 2. *Can* $\mathcal{Q}^{\|}$ *or* $\mathcal{Q}^{\|,pub}$ *be stronger than* $\mathcal{R}$ *with respect to a relation?*

  As we already mentioned, looking for *the weakest quantum model that can outperform* $\mathcal{R}$ and for *the strongest classical model that can be outperformed by* $\mathcal{Q}^{\|}$ are, probably, the two most natural approaches towards understanding the strength and the limits of quantum communication. Ultimately, we would like the two approaches to "meet" – that is, to find a communication problem (even a relational one), *easy for* $\mathcal{Q}^{\|}$ *but hard for* $\mathcal{R}$.

- If we only allow *total functions*, then our current understanding of the *closest pairs* is just a step away from the ideal ignorance:

$$\begin{array}{ccccccc} \mathcal{R}^{\|} & < & \mathcal{R}^{\|,pub} & < & \mathcal{R}^{1} & < & \mathcal{R} \\ \wedge & & ? & & ? & & ? \\ \mathcal{Q}^{\|} & < & \mathcal{Q}^{\|,pub} & < & \mathcal{Q}^{1} & < & \mathcal{Q} \end{array}$$

  Question 3. *In the case of total functions, can $\mathcal{Q}^{\|,pub}$ be stronger than $\mathcal{R}^{\|,pub}$? How about $\mathcal{Q}^{1}$ vs. $\mathcal{R}^{1}$? $\mathcal{Q}$ vs. $\mathcal{R}$?*

- In the context of *"diagonal" relationship* via *total functions* we know nothing.

  Question 4. *In the case of total functions, can $\mathcal{Q}^{\|}$, $\mathcal{Q}^{\|,pub}$ or $\mathcal{Q}^{1}$ be stronger than $\mathcal{R}$? Can $\mathcal{Q}^{\|}$ be stronger than $\mathcal{R}^{\|,pub}$ or $\mathcal{R}^{1}$?*

Lastly, we would like to mention

Question 5. *What is the complexity of our $\widetilde{cEq}$-neg_T in the model of* classical SMP with shared entanglement *($\mathcal{R}^{\|,ent}$)?*

If it has an efficient solution, that would imply a functional separation between $\mathcal{R}^{\|,ent}$ and $\mathcal{R}^{\|,pub}$, which we do not have yet (a relational separation is known); if, on the other hand, $\widetilde{cEq}$-neg_T is hard for $\mathcal{R}^{\|,ent}$, that would imply the possibility of qualitative advantage of $\mathcal{Q}^{\|}$ over $\mathcal{R}^{\|,ent}$, which is currently not known even for relational problems.

Acknowledgements

I am very grateful to Pavel Pudlák, Ronald de Wolf and Thomas Vidick for helpful discussions at various stages of this work. Ronald's comments allowed me to improve the presentation significantly.

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.