

A Direct Product Theorem for One-Way Quantum Communication

Rahul Jain ^{*} Srijita Kundu [†]

Abstract

We prove a direct product theorem for the one-way entanglement-assisted quantum communication complexity of a general relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. For any $\varepsilon, \zeta > 0$ and any $k \geq 1$, we show that

$$Q_{1-(1-\varepsilon)^{\Omega(\zeta^6 k / \log |\mathcal{Z}|)}}^1(f^k) = \Omega\left(k\left(\zeta^5 \cdot Q_{\varepsilon+12\zeta}^1(f) - \log \log(1/\zeta)\right)\right),$$

where $Q_\varepsilon^1(f)$ represents the one-way entanglement-assisted quantum communication complexity of f with worst-case error ε and f^k denotes k parallel instances of f .

As far as we are aware, this is the first direct product theorem for quantum communication – direct sum theorems were previously known for one-way quantum protocols. Our techniques are inspired by the parallel repetition theorems for the entangled value of two-player non-local games, under product distributions due to Jain, Pereszlényi and Yao [JPY14], and under anchored distributions due to Bavarian, Vidick and Yuen [BVY17], as well as message-compression for quantum protocols due to Jain, Radhakrishnan and Sen [JRS05]. In particular, we show that a direct product theorem holds for the distributional one-way quantum communication complexity of f under any distribution q on $\mathcal{X} \times \mathcal{Y}$ that is anchored on one side, i.e., there exists a y^* such that $q(y^*)$ is constant and $q(x|y^*) = q(x)$ for all x . This allows us to show a direct product theorem for general distributions, since for any relation f and any distribution p on its inputs, we can define a modified relation \tilde{f} which has an anchored distribution q close to p , such that a protocol that fails with probability at most ε for \tilde{f} under q can be used to give a protocol that fails with probability at most $\varepsilon + \zeta$ for f under p .

Our techniques also work for entangled non-local games which have input distributions anchored on any one side, i.e., either there exists a y^* as previously specified, or there exists an x^* such that $q(x^*)$ is constant and $q(y|x^*) = q(y)$ for all y . In particular, we show that for any game $G = (q, \mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B}, \mathcal{V})$ where q is a distribution on $\mathcal{X} \times \mathcal{Y}$ anchored on any one side with anchoring probability ζ , then

$$\omega^*(G^k) = \left(1 - (1 - \omega^*(G))^5\right)^{\Omega\left(\frac{\zeta^2 k}{\log(|\mathcal{A}| \cdot |\mathcal{B}|)}\right)}$$

where $\omega^*(G)$ represents the entangled value of the game G . This is a generalization of the result of [BVY17], who proved a parallel repetition theorem for games anchored on both sides, i.e., where both a special x^* and a special y^* exist, and potentially a simplification of their proof.

^{*}Centre for Quantum Technologies and Department of Computer Science, National University of Singapore and MajuLab, UMI 3654, Singapore. Email: rahul@comp.nus.edu.sg

[†]Centre for Quantum Technologies, National University of Singapore, Singapore.
Email: srijita.kundu@u.nus.edu

1 Introduction

A fundamental question in complexity theory is: given k independent instances of a function or relation, does computing them require k times the amount of resources required to compute a single instance of the function or relation? Suppose solving one instance of some problem with success probability at least p requires c units of some resource. A natural way to solve k independent instances of this problem would be to solve them independently, which requires ck units of the resource. A *direct sum theorem* for this problem would state that any algorithm for solving k instances which uses $o(ck)$ units of resource has success probability at most $O(p)$. A *direct product theorem* for the problem would state that any algorithm for solving k instances that uses $o(ck)$ units of resource has success probability at most $p^{\Omega(k)}$. Hence a direct product theorem is the stronger result of the two.

In this paper, we deal with direct product theorems in the model of communication complexity. In this model, there are two parties Alice and Bob, who receive inputs x and y respectively, and wish to jointly compute a relation f . They can use local computation, public coins, and communicate with each other using classical messages, in the classical model; use local unitaries, shared entanglement, and communicate with each other using quantum messages, in the quantum model. The resource of interest is the number of bits/qubits communicated; so the parties are allowed to share an arbitrary amount of randomness or entanglement, and perform local operations of arbitrary complexity.

Direct product theorems in communication are related to *parallel repetition theorems* for *non-local games*. In a non-local game, two parties Alice and Bob are given inputs x and y respectively from some specified distribution, and without communicating with each other, they are required to give answers a and b respectively to a referee. They are considered to win the game if $V(a, b, x, y)$ holds for a specified predicate V . In the classical model, the players are allowed to share randomness, and in the quantum model they are allowed to share entanglement. A parallel repetition theorem shows that the maximum probability of winning k independent instances of a non-local game is $p^{\Omega(k)}$, if the maximum probability of winning a single instance of it is p , regardless of the amount of shared randomness or entanglement used. Direct product theorems in communication are often proved by combining techniques used to prove direct sum theorems in communication, which require message-compression, and parallel repetition theorems for games.

In classical communication complexity, there is a long line of works on direct sum and direct-product theorems including [Raz92, CSWY01, BYJKS02, Sha03, JRS03a, JRS03b, JSR08, KvdW07, BARdW08, LSv08, VW08, JKN08, JK09, HJMR10, Kla10, JY12, She12, BBCR13, BRWY13b, BRWY13a, BR14, Bra15, BW15, Jai15, JPY16, Kol16, BK18, She18]. A parallel repetition theorem for the classical value of general two-player non-local games was first shown by Raz [Raz95], and the proof was subsequently simplified by Holenstein [Hol07].

In quantum communication complexity, a direct sum theorem is known for the entanglement-assisted one-way [JSR08], *simultaneous-message-passing* (SMP), entanglement-assisted [JSR08] and unassisted models [JK09]. A strong parallel repetition theorem for the quantum value of a general two-player non-local game is not known. Parallel repetition theorems were shown for special classes of games such as XOR games [CSUU08], unique games [KRT10] and projection games [DSV15]. When the type of game is not restricted but the input distribution is, parallel repetition theorems

have been shown under product distributions [JPY14] and *anchored* distributions [BVY17, BVY15]. For general games under general distributions, the best current result is due to Yuen [Yue16], which shows that the quantum value of k parallel instances of a general game goes down polynomially in k , if the quantum value of the original game is strictly less than 1. No direct product theorems for quantum communication have so far been shown.

Using ideas from Jain, Pereszlényi and Yao [JPY14] and the message-compression scheme from Jain, Radhakrishnan and Sen [JSR08], a strong direct product theorem for one-way quantum communication under product distributions can be shown. To deal with non-product distributions, we borrow the idea of anchored distributions due to Bavarian, Vidick and Yuen [BVY17, BVY15].

1.1 Our results

Let $Q_\varepsilon^1(f)$ denote the one-way entanglement-assisted quantum communication complexity of a relation f , with worst-case error ε . Let f^k denote k parallel instances of f . Our strong direct product theorem is as follows.

Theorem 1. *For any relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, and any $\varepsilon, \zeta > 0$,*

$$Q_{1-(1-\varepsilon)\Omega(\zeta^6 k / \log |\mathcal{Z}|)}^1(f^k) = \Omega\left(k\left(\zeta^5 \cdot Q_{\varepsilon+12\zeta}^1(f) - \log \log(1/\zeta)\right)\right).$$

Let $\omega^*(G)$ represent the entangled value of a two-player non-local game G , and let G^k denote k parallel instances of G . We call a distribution q on $\mathcal{X} \times \mathcal{Y}$ *anchored on one side* with *anchoring probability* ζ if one of the following conditions holds:

- (i) There exists an $x^* \in \mathcal{X}$ such that $q(x^*) = \zeta$ and $q(y|x^*) = q(y)$ for all $y \in \mathcal{Y}$,
- (ii) There exists an $y^* \in \mathcal{Y}$ such that $q(y^*) = \zeta$ and $q(x|y^*) = q(x)$ for all $x \in \mathcal{X}$.

The game will be called *anchored on both sides* with anchoring probability ζ if both conditions hold instead.

Then our parallel repetition theorem is stated as follows.

Theorem 2. *For a two-player non-local game $G = (q, \mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B}, \mathcal{V})$ such that q is a distribution anchored on one side with anchoring probability ζ ,*

$$\omega^*(G^k) = \left(1 - (1 - \omega^*(G))^5\right)^\Omega\left(\frac{\zeta^{2k}}{\log(|\mathcal{A}| \cdot |\mathcal{B}|)}\right).$$

One can get a game anchored on one side (say the \mathcal{Y} side) from a general game in the following way: in the anchored game, the referee chooses (x, y) from the original probability distribution, and with probability ζ replaces y with a new input y^* . If Bob's input is y^* , then the referee accepts any answer from the players. In a game anchored on both sides, the referee must instead replace x with x^* and y with y^* independently with probability ζ , and accept if either Alice's input is x^* or Bob's input is y^* . It is clear that anchoring makes the game easier. In this light, a parallel repetition theorem for anchoring games can be thought of as follows: for a general game G , there exists a simple transformation taking it to another game \tilde{G} such that

1. If $\omega^*(G) = 1$, then $\omega^*(\tilde{G}^k) = 1$.
2. If $\omega^*(G) < 1$, then $\omega^*(\tilde{G}^k) = \exp(-\Omega(k))$.

The merit of our result here is that the transformation involved for anchoring on one side changes the game less than the transformation involved in anchoring it on both sides.

We note that the definition of anchoring used on [BVY17, BVY15] is more general: instead of single inputs x^*, y^* , they consider anchoring sets $\mathcal{X}^* \subseteq \mathcal{X}$ and $\mathcal{Y}^* \subseteq \mathcal{Y}$, such that $q(\mathcal{X}^*), q(\mathcal{Y}^*) \geq \zeta$, and whenever $x \in \mathcal{X}^*$ or $y \in \mathcal{Y}^*$, $q(x, y) = q(x)q(y)$. However, it appears this generalized definition is not more useful from the perspective of anchoring transformations. While our technique could go through for the one-sided version of this definition of anchoring, we do not state or prove it as such for the sake of simplicity.

Unlike in the case of communication, worst-case success probability is usually not considered for non-local games. But one could define a game $G_{\text{wc}} = (\mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B}, \mathbf{V})$ without an associated distribution, and the worst-case winning probability ω_{wc}^* of this over all inputs of this can be considered. As long as Alice and Bob are allowed to share randomness (which they are, in the quantum case), Yao's lemma [Yao79] holds just like in the case of communication, relating the worst-case winning probability to distributional winning probability. Hence, by choosing $\zeta = (1 - \omega_{\text{wc}}^*(G_{\text{wc}}))/2$ and using the same arguments as in the case of communication, Theorem 2 leads to the following corollary about the worst-case winning probability of any game.

Corollary 3. *For any two-player non-local game $G_{\text{wc}} = (\mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B}, \mathbf{V})$,*

$$\omega_{\text{wc}}^*(G_{\text{wc}}^k) = \left(1 - (1 - \omega_{\text{wc}}^*(G_{\text{wc}}))^7\right)^{\Omega\left(\frac{k}{\log(|\mathcal{A}| \cdot |\mathcal{B}|)}\right)}.$$

1.2 Proof overview

We use the information theoretic framework for parallel repetition and direct product theorems established by [Raz95] and [Hol07]. The broad idea is as follows: for a given relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, let the one-way quantum communication required to compute a single copy with constant success be c . Now consider a one-way quantum protocol \mathcal{P} for f^k which has communication $o(ck)$, in which we can condition on the success of some t coordinates. If the success probability in these t coordinates is already as small as we want, then we are done. Otherwise, we exhibit a $(t + 1)$ -th coordinate i , such that conditioned on the success on the t coordinates, the success of i in \mathcal{P} is bounded away from 1. This is done by showing that if the success probability in the t coordinates is not too small, then we can give a protocol \mathcal{P}' for f whose communication is $o(c)$ and whose success probability is constant – a contradiction.

\mathcal{P}' works by embedding its input into the i -th coordinate of a shared quantum state representing the final input, output, message and discarded registers of \mathcal{P} , conditioned on the success event in the t coordinates, which we denote by \mathcal{E} . Suppose the quantum state conditioned on \mathcal{E} , when Alice and Bob's inputs are x_i and y_i respectively at the i -th coordinates, is $|\varphi\rangle_{x_i y_i}$. On input (x_i, y_i) in \mathcal{P}' , Alice and Bob will by means of local unitaries and communication try to get the shared state close to $|\varphi\rangle_{x_i y_i}$, on which Bob can perform a measurement to get an outcome z_i . The state $|\varphi\rangle_{x_i y_i}$ is such that the resulting probability distribution $\mathbf{P}_{X_i Y_i Z_i}$ is the distribution of $X_i Y_i Z_i$ in \mathcal{P}

conditioned on success. Hence our proof mainly consists of showing how Alice and Bob can get the shared state close to $|\varphi\rangle_{x_i y_i}$.

The proof technique for a parallel repetition theorem is same, except one cannot, and need not, use communication to get the shared state $|\varphi\rangle_{x_i y_i}$ there. In order to motivate our techniques, we shall briefly describe the techniques used in [JPY14] and [BVY15] to get $|\varphi\rangle_{x_i y_i}$.

- In [JPY14] the following three states are considered: $|\varphi\rangle_{x_i}$ which is the superposition of $|\varphi\rangle_{x_i y_i}$ over the distribution of Y_i , $|\varphi\rangle_{y_i}$ which is the superposition over the distribution of X_i , and $|\varphi\rangle$ which is the superposition over both. In this setting, $X_1 \dots X_k$ are initially in product with all of Bob's registers and $Y_1 \dots Y_k$ are in product with all of Alice's registers. If the probability of \mathcal{E} is large, then conditioning on it, the following can be shown:
 1. By chain rule of mutual information, there is an X_i whose mutual information with Bob's registers in $|\varphi\rangle$ is small. Hence by Uhlmann's theorem, there exist unitaries U_{x_i} acting on Alice's registers that take $|\varphi\rangle$ close to $|\varphi\rangle_{x_i}$.
 2. Similarly, the mutual information between Y_i and Alice's registers in $|\varphi\rangle$ is small, and hence there exist unitaries U_{y_i} acting on Bob's registers that take $|\varphi\rangle$ close to $|\varphi\rangle_{y_i}$.
 3. Since U_{x_i} and U_{y_i} act on disjoint registers, using a *commuting* argument and the *monotonicity* of trace-distance under quantum-operations, $U_{x_i} \otimes U_{y_i}$ takes $|\varphi\rangle$ close to $|\varphi\rangle_{x_i y_i}$.

Alice and Bob can thus share $|\varphi\rangle$ as entanglement, and get close to $|\varphi\rangle_{x_i y_i}$ by local operations.

- In [BVY15], $X_1 \dots X_k$ are not initially in product with $Y_1 \dots Y_k$, hence they need to use what are known as *correlation-breaking variables*. For each i , correlation-breaking variables $D_i G_i$ are such that conditioned on $D_i G_i$, X_i and Y_i are independent. In particular, D_i is a uniformly distributed bit, and G_i takes values in either \mathcal{X} or \mathcal{Y} depending on whether D_i is 0 or 1, and is highly correlated with either X_i or Y_i in the respective cases. This means that conditioned on $D_i = 0$, $G_i = x^*$ with probability $\Omega(\zeta)$ and conditioned on $D_i = 1$, $G_i = y^*$ with probability $\Omega(\zeta)$.
 1. The mutual information between X_i and Bob's registers in $|\varphi\rangle$ conditioned on $D_i = 1$ and G_i is small. Further conditioning on $G_i = y^*$ (which happens with constant probability), the mutual information between X_i and Bob's registers in $|\varphi\rangle_{y^*}$ is small. Hence by Uhlmann's theorem, there exist unitaries U_{x_i} on Alice's registers, taking $|\varphi\rangle_{x^* y^*}$ close to $|\varphi\rangle_{x_i y^*}$.
 2. Similarly, the mutual information between Y_i and Alice's registers in $|\varphi\rangle$ conditioning on $D_i = 0$ and $G_i = x^*$ is small, which means there exist unitaries U_{y_i} on Bob's registers, taking $|\varphi\rangle_{x^* y^*}$ close to $|\varphi\rangle_{x^* y_i}$.
 3. Using an involved argument, it is possible to show that $U_{x_i} \otimes U_{y_i}$ takes $|\varphi\rangle_{x^* y^*}$ close to $|\varphi\rangle_{x_i y_i}$.

Alice and Bob can thus share $|\varphi\rangle_{x^* y^*}$ in this case, and get close to $|\varphi\rangle_{x_i y_i}$ by local operations.

In our direct product proof, since the distribution is anchored on one side, we use correlation-breaking variables that are identical to those in [BVY15] in the $D_i = 1$ case, but in the $D_i = 0$ we consider a simpler distribution where G_i is perfectly correlated with X_i . Here we also clarify what

we mean by G_i and Y_i being highly correlated when $D_i = 1$: if $G_i = y^*$, then Y_i is always y^* ; but if $G_i = y_i$ for $y_i \neq y^*$, then Y_i still takes value y^* with probability $\Omega(\zeta)$, and is y_i otherwise. The distribution of X_i conditioned on $G_i = y^*$ is the marginal distribution of X_i , while conditioned on y_i , it is the same as the distribution of X_i conditioned on $Y_i = y_i$ (potentially different from the marginal distribution of X_i). Our use of these correlation-breaking variables is quite different from that in [BVY15], however.

We note that in a communication protocol where Alice sends the message, we cannot hope to show that the mutual information between X_i and Bob's registers is small even conditioned on the correlation-breaking variables, since the final state on Bob's side includes the message from Alice, which can potentially be fully correlated with Alice's inputs. Since Bob does not communicate however, the same does not apply to him. Hence we can show the following:

1. If the message size is $o(ck)$, by chain rule of mutual information, the mutual information between X_i and Bob's registers in $|\varphi\rangle$ is $o(c)$, conditioned on $D_i = 1, G_i = y^*$. Since the distribution is anchored on Bob's side, this means that the mutual information between X_i and Bob's registers in $|\varphi\rangle_{y^*}$ is $o(c)$. Using a result from [JRS02, JSR08], then there exist projectors Π_{x_i} acting on Alice's registers, which succeed with probability $2^{-o(c)}$ on $|\varphi\rangle_{y^*}$, and on success take it close to $|\varphi\rangle_{x_i y^*}$.
2. The mutual information between Y_i and Alice's registers conditioned on $D_i = 1, G_i \neq y^*$ is small. For each value of $G_i \neq y^*$, there exist only two possible values of Y_i : y_i and y^* , and hence Alice's registers in $|\varphi\rangle_{y_i}$ and $|\varphi\rangle_{y^*}$ must be close on average. By Uhlmann's theorem, there exist unitaries U_{y_i} acting on Bob's registers, taking $|\varphi\rangle_{y^*}$ close to $|\varphi\rangle_{y_i}$.
3. Since the marginal distribution of X_i conditioned on $G_i = y_i$ is approximately the same as the marginal distribution of X_i conditioned on $Y_i = y_i$, we can show by the same argument as in [JSR08, JPY14], that conditioned on success of Π_{x_i} , $\Pi_{x_i} \otimes U_{y_i}$ takes $|\varphi\rangle_{y^*}$ close to $|\varphi\rangle_{x_i y_i}$.

Hence there is a communication protocol with prior shared entanglement which allows Alice and Bob to obtain a state close to $|\varphi\rangle_{x_i y_i}$ as a shared state on input (x_i, y_i) : Alice and Bob share $2^{o(c)}$ copies of $|\varphi\rangle_{y^*}$ as entanglement; Alice performs the Π_{x_i} measurement on all these copies, and succeeds on at least one copy with high probability. She sends the index of the copy on which she succeeds to Bob, who performs U_{y_i} on the same copy. This protocol has communication $o(c)$, since that is how many classical bits Alice needs in order to encode the index of the successful copy out of $2^{o(c)}$ copies. This completes the proof of the direct product theorem.

Our parallel repetition proof is same as above, except no communication is necessary, since there was no communication in the original protocol. Instead of a projector on Alice's registers taking $|\varphi\rangle_{y^*}$ close to $|\varphi\rangle_{x_i y^*}$, in this case we will have a unitary U_{x_i} doing it. We can argue identically to the direct product proof that there exist U_{y_i} taking $|\varphi\rangle_{y^*}$ close to $|\varphi\rangle_{y_i}$, and $U_{x_i} \otimes U_{y_i}$ takes $|\varphi\rangle_{y^*}$ close to $|\varphi\rangle_{x_i y_i}$. The last part, indicated as step 3 above, is arguably simpler in our proof compared to [BVY15].

2 Preliminaries

2.1 Probability theory

We shall denote the probability distribution of a random variable X on some set \mathcal{X} by \mathbb{P}_X . For any event \mathcal{E} on \mathcal{X} , the distribution of X conditioned on \mathcal{E} will be denoted by $\mathbb{P}_{X|\mathcal{E}}$. For joint random variables XY , $\mathbb{P}_{X|Y=y}(x)$ is the conditional distribution of X given $Y = y$; when it is clear from context which variable's value is being conditioned on, we shall often shorten this to $\mathbb{P}_{X|y}$. We shall use $\mathbb{P}_{XY}\mathbb{P}_{Z|X}$ to refer to the distribution

$$(\mathbb{P}_{XY}\mathbb{P}_{Z|X})(x, y, z) = \mathbb{P}_{XY}(x, y) \cdot \mathbb{P}_{Z|X=x}(z).$$

For two distributions \mathbb{P}_X and $\mathbb{P}_{X'}$ on the same set \mathcal{X} , the ℓ_1 distance between them is defined as

$$\|\mathbb{P}_X - \mathbb{P}_{X'}\|_1 = \sum_{x \in \mathcal{X}} |\mathbb{P}_X(x) - \mathbb{P}_{X'}(x)|.$$

Fact 4. For joint distributions \mathbb{P}_{XY} and $\mathbb{P}_{X'Y'}$ on the same sets,

$$\|\mathbb{P}_X - \mathbb{P}_{X'}\|_1 \leq \|\mathbb{P}_{XY} - \mathbb{P}_{X'Y'}\|_1.$$

Fact 5. For two distributions \mathbb{P}_X and $\mathbb{P}_{X'}$ on the same set and an event \mathcal{E} on the set,

$$|\mathbb{P}_X(\mathcal{E}) - \mathbb{P}_{X'}(\mathcal{E})| \leq \frac{1}{2} \|\mathbb{P}_X - \mathbb{P}_{X'}\|_1.$$

Fact 6. For two distributions \mathbb{P}_X and $\mathbb{P}_{X'}$ on the same set, and any joint distribution $\mathbb{P}_{XX'}$ whose marginals are \mathbb{P}_X and $\mathbb{P}_{X'}$ respectively, we have

$$\|\mathbb{P}_X - \mathbb{P}_{X'}\|_1 \leq 2\mathbb{P}_{XX'}(X \neq X').$$

Fact 7. Suppose probability distributions $\mathbb{P}_X, \mathbb{P}_{X'}$ satisfy $\|\mathbb{P}_X - \mathbb{P}_{X'}\|_1 \leq \varepsilon$, and an event \mathcal{E} satisfies $\mathbb{P}_X(\mathcal{E}) \geq \alpha$, where $\alpha > \varepsilon$. Then,

$$\|\mathbb{P}_{X|\mathcal{E}} - \mathbb{P}_{X'|\mathcal{E}}\|_1 \leq \frac{2\varepsilon}{\alpha}.$$

Proof. From Fact 5, $\alpha - \varepsilon/2 \leq \mathbb{P}_{X'}(\mathcal{E}) \leq \alpha + \varepsilon/2$. By definition, there exists an event \mathcal{E}' such that $2(\mathbb{P}_{X|\mathcal{E}}(\mathcal{E}') - \mathbb{P}_{X'|\mathcal{E}}(\mathcal{E}')) = \|\mathbb{P}_{X|\mathcal{E}} - \mathbb{P}_{X'|\mathcal{E}}\|_1$. Now, $\mathbb{P}_X(\mathcal{E} \wedge \mathcal{E}') = \mathbb{P}_X(\mathcal{E})\mathbb{P}_{X|\mathcal{E}}(\mathcal{E}') \geq \alpha\mathbb{P}_{X|\mathcal{E}}(\mathcal{E}')$. Similarly, $\mathbb{P}_{X'}(\mathcal{E} \wedge \mathcal{E}') \leq (\alpha + \varepsilon/2)\mathbb{P}_{X'|\mathcal{E}}(\mathcal{E}') \leq \alpha\mathbb{P}_{X'|\mathcal{E}}(\mathcal{E}') + \frac{1}{2}\|\mathbb{P}_X - \mathbb{P}_{X'}\|_1$.

Now,

$$\begin{aligned} \|\mathbb{P}_X - \mathbb{P}_{X'}\|_1 &\geq 2(\mathbb{P}_X(\mathcal{E} \wedge \mathcal{E}') - \mathbb{P}_{X'}(\mathcal{E} \wedge \mathcal{E}')) \\ &\geq 2\alpha(\mathbb{P}_{X|\mathcal{E}}(\mathcal{E}') - \mathbb{P}_{X'|\mathcal{E}}(\mathcal{E}')) - \|\mathbb{P}_X - \mathbb{P}_{X'}\|_1 \\ &\geq \alpha\|\mathbb{P}_{X|\mathcal{E}} - \mathbb{P}_{X'|\mathcal{E}}\|_1 - \|\mathbb{P}_X - \mathbb{P}_{X'}\|_1 \end{aligned}$$

which gives the required result. \square

Fact 8 ([BVY15], Lemma 16). Suppose XYZ are random variables satisfying $\mathbb{P}_{XY}(x, y^*) = \alpha \cdot \mathbb{P}_X(x)$ for all x . Then,

$$\|\mathbb{P}_{XYZ} - \mathbb{P}_{XY}\mathbb{P}_{Z|X, y^*}\|_1 \leq \frac{2}{\alpha} \|\mathbb{P}_{XYZ} - \mathbb{P}_{XY}\mathbb{P}_{Z|X}\|_1.$$

Corollary 9. Suppose P_{XY} and $P_{X'Y'Z'}$ are distributions such that $\|P_{XY} - P_{X'Y'}\|_1 \leq \varepsilon$, and $P(x, y^*) = \alpha \cdot P_X(x)$ for all x . Then,

$$\|P_{X'Z'|y^*} - P_{X'Z'}\|_1 \leq \frac{11}{\alpha} \|P_{X'Y'Z'} - P_{XY}P_{Z'|X'}\|_1.$$

Proof. Let $\|P_{X'Y'Z'} - P_{XY}P_{Z'|X'}\|_1 = \varepsilon$. Note that

$$\|P_{X|y^*} - P_{X'|y^*}\|_1 \leq \frac{2\varepsilon}{\alpha}$$

by Fact 7. Let $P_{XZ''}$ denote the distribution $P_{XY}P_{Z'|X'Y'}$.

$$\begin{aligned} \|P_{X'Z'} - P_{XZ''}\|_1 &= \sum_{x,z} \left| P_{X'}(x) \sum_y P_{Y'|x}(y) P_{Z'|xy}(z) - P_X(x) \sum_y P_{Y|x}(y) P_{Z'|xy}(z) \right| \\ &\leq \sum_{x,y,z} |P_{X'}(x) P_{Y'|x}(y) - P_X(x) P_{Y|x}(y)| P_{Z'|xy}(z) \\ &= \|P_{X'Y'} - P_{XY}\|_1 \leq \varepsilon. \end{aligned}$$

$$\begin{aligned} \|P_{XZ''} - P_{XY}P_{Z'|X'}\|_1 &\leq \|P_{XZ''} - P_{X'Y'Z'}\|_1 + \|P_{X'Y'Z'} - P_{XY}P_{Z'|X'}\|_1 \\ &\quad + \|P_{XY}P_{Z'|X'} - P_{XY}P_{Z''|X'}\|_1 \\ &= \|P_{XY} - P_{X'Y'}\|_1 + \|P_{X'Y'Z'} - P_{XY}P_{Z'|X'}\|_1 \\ &\quad + \sum_{x,y} P_{XY}(x,y) \|P_{Z'|x} - P_{Z''|x}\|_1 \\ &\leq 2\varepsilon + \sum_x P_X(x) \sum_{y,z} |P_{Y|x}(y) - P_{Y'|x}(y)| P_{Z'|xy}(z) \\ &\leq 2\varepsilon + \sum_{x,y} |P_X(x) P_{Y|x}(y) - P_{X'}(x) P_{Y'|x}(y)| \\ &\quad + \sum_{x,y} |P_{X'}(x) - P_X(x)| P_{Y'|x}(y) \\ &\leq 2\varepsilon + 2\|P_{XY} - P_{X'Y'}\|_1 \leq 4\varepsilon. \end{aligned}$$

Combining all this,

$$\begin{aligned} \|P_{X'Z'|y^*} - P_{X'Z'}\|_1 &\leq \|P_{X'Z'|y^*} - P_{XZ''|y^*}\|_1 + \|P_{XZ''|y^*} - P_{XZ''}\|_1 + \|P_{XZ''} - P_{X'Z'}\|_1 \\ &\leq \|P_{X|y^*} - P_{X'|y^*}\|_1 + \|P_{XZ''|y^*} - P_{XZ''}\|_1 + \|P_{XZ''} - P_{X'Z'}\|_1 \\ &\leq \frac{2\varepsilon}{\alpha} + \frac{2}{\alpha} \|P_{XZ''} - P_{XY}P_{Z'|X'}\|_1 + \varepsilon \\ &\leq \frac{2\varepsilon}{\alpha} + \frac{8\varepsilon}{\alpha} + \varepsilon \leq \frac{11\varepsilon}{\alpha}. \end{aligned}$$

where we have used Lemma 8 in the third inequality. \square

Fact 10 ([Hol07], Corollary 6). Let $P_{TU_1\dots U_kV} = P_T P_{U_1|T} P_{U_2|T} \dots P_{U_k|T} P_{V|TU_1\dots U_k}$ be a probability distribution over $\mathcal{T} \times \mathcal{U}^k \times \mathcal{V}$, and let \mathcal{E} be any event. Then,

$$\sum_{i=1}^k \|P_{TU_iV|\mathcal{E}} - P_{TV|\mathcal{E}} P_{U_i|T}\|_1 \leq \sqrt{k \left(\log(|\mathcal{V}|) + \log\left(\frac{1}{\Pr[\mathcal{E}]}\right) \right)}.$$

Definition 1 ([Hol07]). For two distributions P_{XY} and $P_{X'Y'ST}$, we say (X, Y) is $(1-\varepsilon)$ -embeddable in $(X'S, Y'T)$ if there exists a random variable R on a set \mathcal{R} independent of XY and functions $f_A : \mathcal{X} \times \mathcal{R} \rightarrow \mathcal{S}$ and $f_B : \mathcal{Y} \times \mathcal{R} \rightarrow \mathcal{T}$, such that

$$\|P_{XY f_A(X,R) f_B(X,R)} - P_{X'Y'ST}\|_1 \leq \varepsilon.$$

Fact 11 ([Hol07, JPY16]). If two distributions P_{XY} and $P_{X'Y'R'}$ satisfy

$$\|P_{X'Y'R'} - P_{XY} P_{R'|X'}\|_1 \leq \varepsilon \quad \|P_{X'Y'R'} - P_{XY} P_{R'|Y'}\|_1 \leq \varepsilon,$$

then (X, Y) is $(1 - 5\varepsilon)$ -embeddable in $(X'R', Y'R')$.¹

2.2 Quantum information

The ℓ_1 distance between two quantum states ρ and σ is given by

$$\|\rho - \sigma\|_1 = \text{Tr} \sqrt{(\rho - \sigma)^\dagger (\rho - \sigma)} = \text{Tr} |\rho - \sigma|.$$

The fidelity between two quantum states is given by

$$F(\rho, \sigma) = \|\sqrt{\rho} \sqrt{\sigma}\|_1.$$

ℓ_1 distance and fidelity are related in the following way.

Fact 12 (Fuchs-van de Graaf inequality). For any pair of quantum states ρ and σ ,

$$2(1 - F(\rho, \sigma)) \leq \|\rho - \sigma\|_1 \leq 2\sqrt{1 - F(\rho, \sigma)^2}.$$

For two pure states $|\psi\rangle$ and $|\phi\rangle$, we have

$$\| |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \|_1 = \sqrt{1 - F(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|)^2} = \sqrt{1 - |\langle\psi|\phi\rangle|^2}.$$

Fact 13 (Uhlmann's theorem). Suppose ρ and σ are mixed states on register X which are purified to $|\rho\rangle$ and $|\sigma\rangle$ on registers XY , then it holds that

$$F(\rho, \sigma) = \max_U |\langle\rho|\mathbb{1}_X \otimes U|\sigma\rangle|$$

where the maximization is over unitaries acting only on register Y .

¹This fact is equivalent to Lemma 2.11 in [JPY16], although this lemma is stated in terms of relative entropies instead of trace distances between the various distributions. In the proof of the lemma, the relative entropies are converted to the same trace distances as we consider, using Pinsker's inequality. This justifies our statement of the fact, which is tailored towards our application.

Fact 14. For a quantum channel \mathcal{E} and states ρ and σ ,

$$\|\mathcal{E}(\rho) - \mathcal{E}(\sigma)\|_1 \leq \|\rho - \sigma\|_1 \quad \text{and} \quad F(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \geq F(\rho, \sigma).$$

The entropy of a quantum state ρ on a register Z is given by

$$S(\rho) = -\text{Tr}(\rho \log \rho).$$

The relative entropy between two states ρ and σ of the same dimensions is given by

$$S(\rho\|\sigma) = \text{Tr}(\rho \log \rho) - \text{Tr}(\rho \log \sigma).$$

The relative min-entropy between ρ and σ is defined as

$$S_\infty(\rho\|\sigma) = \min\{\lambda : \rho \leq 2^\lambda \sigma\}.$$

It is easy to see that $S(\rho\|\sigma)$ and $S_\infty(\rho\|\sigma)$ only take finite values when the support of ρ is contained in the support of σ . Moreover, clearly $0 \leq S(\rho\|\sigma) \leq S_\infty(\rho\|\sigma)$ for all ρ and σ .

The ε -smooth relative min-entropy between ρ and σ is defined as

$$S_\infty^\varepsilon(\rho\|\sigma) = \inf_{\rho': \|\rho - \rho'\|_1 \leq \varepsilon} S(\rho'\|\sigma).$$

$S_\infty^\varepsilon(\rho\|\sigma)$ can take a finite value even if the support of ρ is not contained in the support of σ , for example if ρ is ε -close to a state contained within the support of σ . $S_\infty(\rho\|\sigma)$ cannot be upper bounded by $S(\rho\|\sigma)$, but $S_\infty^\varepsilon(\rho\|\sigma)$ can be, due to the Quantum Substate Theorem.

Fact 15 (Quantum Substate Theorem, [JRS09, JN12]). For any two states ρ and σ such that the support of ρ is contained in the support of σ , and any $\varepsilon > 0$,

$$S_\infty^\varepsilon(\rho\|\sigma) \leq \frac{4S(\rho\|\sigma)}{\varepsilon^2} + \log\left(\frac{1}{1 - \varepsilon^2/4}\right).$$

Fact 16 (Pinsker's Inequality). For any two states ρ and σ , $\|\rho - \sigma\|_1 \leq \sqrt{4S(\rho\|\sigma)}$.

Fact 17. If $\sigma = \varepsilon\rho + (1 - \varepsilon)\rho'$, then $S_\infty(\rho\|\sigma) \leq \log(1/\varepsilon)$.

Fact 18. For any three quantum states ρ, σ, φ such that $\text{supp}(\rho) \subseteq \text{supp}(\varphi) \subseteq \text{supp}(\sigma)$,

$$S_\infty(\rho\|\sigma) \leq S_\infty(\rho\|\varphi) + S_\infty(\varphi\|\sigma).$$

Fact 19. For any unitary U , $S_\infty(U\rho U^\dagger\|U\sigma U^\dagger) = S_\infty(\rho\|\sigma)$.

A state of the form

$$\rho_{XY} = \sum_x P_X(x) |x\rangle\langle x|_X \otimes \rho_{Y|x}$$

is called a CQ (classical-quantum) state, with X being the classical register and Y being quantum. We shall use X to refer to both the classical register and the classical random variable with the associated distribution. As in the classical case, here we are using $\rho_{Y|x}$ to denote the state of the register Y conditioned on $X = x$, or in other words the state of the register Y when a measurement

is done on the X register and the outcome is x . Hence $\rho_{XY|x} = |x\rangle\langle x|_X \otimes \rho_{Y|x}$. When the registers are clear from context we shall often write simply ρ_x .

The mutual information between Y and Z with respect to a state ρ on YZ is defined as

$$I(Y : Z)_\rho = S(\rho_{YZ} \| \rho_Y \otimes \rho_Z).$$

The ε -smooth max-information between Y and Z with respect to ρ is defined as

$$I_{\max}^\varepsilon(Y : Z)_\rho = \inf_{\sigma_Z} S_\infty^\varepsilon(\rho_{YZ} \| \rho_Y \otimes \sigma_Z).$$

The conditional mutual information between Y and Z conditioned on a classical register X , is defined as

$$I(Y : Z|X) = \mathbb{E}_{P_X} [I(Y : Z)_{\rho_x}].$$

Mutual information can be seen to satisfy the chain rule

$$I(XY : Z)_\rho = I(X : Z)_\rho + I(Y : Z|X)_\rho.$$

Fact 20 ([BCR11], Lemma B.7). *For any quantum state ρ_{YZ} ,*

$$\inf_{\sigma_Z} S_\infty(\rho_{YZ} \| \rho_Y \otimes \sigma_Z) \leq 2 \min\{\log |\mathcal{Y}|, \log |\mathcal{Z}|\}.$$

Fact 21. *For CQ states*

$$\rho_{XY} = \sum_x P_X(x) |x\rangle\langle x|_X \otimes \rho_{Y|x} \quad \sigma_{XY} = \sum_x P_{X'}(x) |x\rangle\langle x|_X \otimes \sigma_{Y|x},$$

their relative entropy is given by

$$S(\rho_{XY} \| \sigma_{XY}) = S(P_X \| P_{X'}) + \mathbb{E}_{P_X} [S(\rho_{Y|x} \| \sigma_{Y|x})].$$

Fact 22. *Suppose σ_{XYZ} and ρ_{XYZ} are CQ states defined as follows*

$$\sigma_{XYZ} = \sum_{x,y} P_{XY}(x,y) |x,y\rangle\langle x,y| \otimes \sigma_{Z|xy} \quad \rho_{XYZ} = \sum_{x,y} P_{X'Y'}(x,y) |x,y\rangle\langle x,y| \otimes \sigma_{Z|xy},$$

where $\|P_{XY} - P_{X'Y'}\|_1 \leq \delta < \frac{1}{2}$. Let $I(Y : Z|X)_\sigma \leq c$. Then, for any $\delta < \varepsilon < \frac{1}{2}$,

$$P_{X'} \left(I_{\max}^{\varepsilon+7\delta/\varepsilon}(Y : Z)_{\rho_x} > \frac{4c+1}{\varepsilon^3} \right) \leq 2\varepsilon + \frac{\delta}{2}.$$

Proof. Let Good_1 denote the set of x such that $I(Y : Z)_{\sigma_x} \leq c/\varepsilon$. Due to Markov's inequality, $I(Y : Z|X)_\rho \leq c$ implies $P_X(\text{Good}_1) \geq 1 - \varepsilon$. By Quantum Substate Theorem, for each $x \in \text{Good}_1$, there exist a $\sigma'_{YZ|x}$ such that $\|\sigma_{YZ|x} - \sigma'_{YZ|x}\|_1 \leq \varepsilon$, and a $\theta_{Z|x}$ such that

$$S_\infty(\sigma'_{YZ|x} \| \sigma_{Y|x} \otimes \theta_{Z|x}) \leq \frac{4c}{\varepsilon^3} + \log \left(\frac{1}{1 - \varepsilon^2/4} \right) = k \text{ (say).}$$

Clearly $\sigma_{Y|z} = \sum_y P_{Y|X=x}(y)|y\rangle\langle y|_Y$ and $\rho_{Y|x} = \sum_y P_{Y'|X'=x}(y)|y\rangle\langle y|_Y$. Let

$$\mathcal{E}_x = \left\{ y : \log \frac{P_{Y|x}(y)}{P_{Y'|x}(y)} \leq 1 \right\} \quad \mathcal{E} = \left\{ (x, y) : \log \frac{P_{Y|x}(y)}{P_{Y'|x}(y)} \leq 1 \right\}.$$

We observe that

$$\begin{aligned} P_{XY}(\mathcal{E}^c) &\leq \delta/2 + P_{X'Y'}(\mathcal{E}^c) \\ &\leq \delta/2 + \frac{1}{2} \sum_{(x,y) \notin \mathcal{E}} P_{X'}(x) P_{Y|x}(y) \\ &\leq \delta/2 + P_{XY}(\mathcal{E}^c)/2 + \frac{1}{2} \sum_{(x,y) \notin \mathcal{E}} |P_{X'}(x) - P_X(x)| P_{Y|x}(y) \\ &\leq \delta/2 + P_{XY}(\mathcal{E}^c)/2 + \frac{1}{2} \sum_x |P_{X'}(x) - P_X(x)| \\ &\leq \delta + P_{XY}(\mathcal{E}^c)/2, \end{aligned}$$

which gives us $P_{XY}(\mathcal{E}^c) = \mathbb{E}_{P_X} P_Y(\mathcal{E}_x^c) \leq 2\delta$. Let Good_2 denote the set of x such that $P(\mathcal{E}_x^c) \leq 2\delta/\varepsilon$. By Markov's inequality, $P_X(\text{Good}_2) \geq 1 - \varepsilon$.

Let Π_x denote the projector on Y that projects to the subset \mathcal{E}_x . By definition, $\Pi_x \sigma_{Y|x} \Pi_x \leq 2\rho_{Y|x}$. Now $\sigma'_{YZ|x} \leq 2^k(\sigma_{Y|x} \otimes \theta_{Z|x})$ implies

$$\Pi_x \sigma'_{Y|x} \Pi_x \leq 2^k \Pi_x (\sigma_{Y|x} \otimes \theta_{Z|x}) \Pi_x \leq 2^{k+1} \rho_{Y|x} \otimes \theta_{Z|x}.$$

Let $\rho'_{YZ|x}$ denote $\Pi_x \sigma'_{YZ|x} \Pi_x / \text{Tr}(\Pi_x \sigma'_{YZ|x} \Pi_x)$. We note

$$\begin{aligned} \text{Tr}(\Pi_x \sigma'_{YZ|x} \Pi_x) &\geq \text{Tr}(\Pi_x \sigma_{YZ|x} \Pi_x) - \text{Tr}(\Pi_x |\sigma'_{YZ|x} - \sigma_{YZ|x}| \Pi_x) \\ &\geq P_Y(\mathcal{E}_x) - \|\sigma'_{YZ|x} - \sigma_{YZ|x}\|_1 \\ &\geq 1 - 2\delta/\varepsilon - \delta \end{aligned}$$

for $x \in \text{Good}_1 \cap \text{Good}_2$. Hence for such x ,

$$S_\infty(\rho'_{YZ|x} \|\rho_{Y|x} \otimes \theta_{Z|x}) \leq k + 1 + \log \left(\frac{1}{1 - 2\delta/\varepsilon - \delta} \right) \leq \frac{4c + 1}{\varepsilon^3}.$$

Also for these x ,

$$\|\rho'_{YZ|x} - \rho_{YZ|x}\|_1 \leq \|\rho'_{YZ|x} - \sigma'_{YZ|x}\|_1 + \|\sigma'_{YZ|x} - \sigma_{YZ|x}\|_1 + \|\sigma_{YZ|x} - \rho_{YZ|x}\|_1 \leq \varepsilon + \frac{7\delta}{\varepsilon}$$

which gives us

$$I_{\max}^{\varepsilon+7\delta/\varepsilon}(Y : Z)_{\rho_x} \leq \frac{4c + 1}{\varepsilon^3}.$$

We know $P_X(\text{Good}_1 \cap \text{Good}_2) \geq 1 - 2\varepsilon$. Hence $P_{X'}(\text{Good}_1 \cap \text{Good}_2) \geq 1 - 2\varepsilon - \delta/2$, which gives us the desired result. \square

Fact 23 (Quantum Raz’s Lemma, [BVY15]). *Let ρ_{XY} and σ_{XY} be two CQ states with $X = X_1 \dots X_k$ being classical, and σ being product across all registers. Then,*

$$\sum_{i=1}^k I(X_i : Y)_\rho \leq S(\rho_{XY} \parallel \sigma_{XY}).$$

Fact 24 ([JRS05], Lemma 2). *Suppose the state*

$$|\sigma\rangle_{X\tilde{X}AB} = \sum_x \sqrt{P_X(x)} |xx\rangle_{X\tilde{X}} |\sigma\rangle_{AB|x}$$

satisfies $I_{\max}^\delta(X : B)_\sigma \leq k$ for some $\delta > 0$. Then there is a family of measurement operators $\{\Pi_x\}_x$ acting only on $X\tilde{X}A$ such that:

- (i) *Each Π_x succeeds with probability $\alpha = 2^{-k/\delta}$ on $|\sigma\rangle_{X\tilde{X}AB}$,*
- (ii) *$(\Pi_x \otimes \mathbb{1}_B)|\sigma\rangle\langle\sigma|(\Pi_x \otimes \mathbb{1}_B)$ is of the form $|xx\rangle\langle xx| \otimes \rho_x$, for some state ρ_x on AB , and*

$$\mathbb{E}_{P_X} \left\| \frac{1}{\alpha} (\Pi_x \otimes \mathbb{1}_B) |\sigma\rangle\langle\sigma|_{X\tilde{X}AB} (\Pi_x \otimes \mathbb{1}_B) - |xx\rangle\langle xx|_{X\tilde{X}} \otimes |\sigma\rangle\langle\sigma|_{AB|x} \right\|_1 \leq \delta.^2$$

2.3 Quantum communication & entangled games

We briefly describe a quantum communication protocol \mathcal{P} for computing a relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, between two parties Alice and Bob sharing prior entanglement, with inputs x and y respectively.

In each round, either Alice or Bob will apply a unitary on their classical input register, along with the quantum register they received as a message from the other party in the last round, and memory registers they may have kept from previous rounds; after the unitary they will keep some registers as memory and send the rest to the other party as the message for that round. We can always assume that players make ‘safe’ copies of their inputs using CNOT gates in such protocols, so that the input registers come out as is after each round. We also note that though in general we need not consider shared classical randomness in quantum communication protocols, protocols with shared randomness fall under the shared entanglement framework we have described. This is because shared randomness can be obtained by sharing entanglement and then both parties measuring in the same basis.

In a one-way, i.e., a single round protocol, the memory from previous rounds is replaced by Alice’s (who we consider to be sending the single message) part of the shared entangled state, and any register she does not send as a message is simply discarded. After Alice’s message, Bob performs a projective measurement on his input register, his part of the shared entanglement, and Alice’s message, and gives the outcome of this measurement as the output of the protocol, which we

²The version of this fact stated here is slightly different from the original statement in [JRS05], in order to suit our application. In the original statement, $I(X : B)$ is used instead of $I_{\max}^\delta(X : B)$, and the superposition state lacks the \tilde{X} register. However, in the proof of the fact in [JRS05], $I(X : B)$ is converted to $I_{\max}^\delta(X : B)$ anyway, so the first change makes no difference. The second change also makes no difference as the same projector that takes the superposition state without the \tilde{X} register to $|x\rangle\langle x| \otimes |\sigma\rangle\langle\sigma|_{AB|x}$ takes the superposition state with the \tilde{X} register to $|xx\rangle\langle xx| \otimes |\sigma\rangle\langle\sigma|_{AB|x}$.

shall denote by $\mathcal{P}(x, y)$. We can of course think of this measurement as Bob performing a unitary on the three registers, and then doing a measurement in the computational basis on some $\log |\mathcal{Z}|$ qubits which are designated for the output.

Definition 2. The one-way entanglement-assisted quantum communication complexity, with error $0 < \varepsilon < 1$, of a relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, denoted by $Q_\varepsilon^1(f)$, is the minimum message size, i.e., number of qubits sent, in a one-way entanglement-assisted quantum protocol \mathcal{P} such that for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$,

$$\Pr[\mathcal{P}(x, y) \in f(x, y)] \geq 1 - \varepsilon,$$

where the probability is taken over the inherent randomness in the protocol.

Definition 3. For a probability distribution p on $\mathcal{X} \times \mathcal{Y}$, the distributional one-way entanglement-assisted quantum communication complexity of a relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, with error $0 < \varepsilon < 1$ with respect to p , is defined as the minimum message size of a one-way entanglement-assisted quantum protocol \mathcal{P} such that

$$\Pr[\mathcal{P}(x, y) \in f(x, y)] \geq 1 - \varepsilon,$$

where the probability is taken over the distribution p on (x, y) as well as the inherent randomness in the protocol.

Fact 25 (Yao's lemma, [Yao79]). *For any $0 < \varepsilon < 1$, and any relation f , $Q_\varepsilon^1(f) = \max_p Q_{p, \varepsilon}^1(f)$.*

A two-player non-local game G is described as $(q, \mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B}, \mathbf{V})$ where q is a distribution over the input set $\mathcal{X} \times \mathcal{Y}$, $\mathcal{A} \times \mathcal{B}$ is the output set, and $\mathbf{V} : \mathcal{X} \times \mathcal{Y} \times \mathcal{A} \times \mathcal{B} \rightarrow \{0, 1\}$ is a predicate. It is played as follows: a referee selects inputs (x, y) according to q , sends x to Alice and y to Bob. If Alice and Bob are allowed to share entanglement, they perform measurements on their respective halves of the entangled state along with their respective input registers (which we model as performing unitaries and then measuring in the computational basis on some $\log |\mathcal{A}|$ and $\log |\mathcal{B}|$ qubits designated for outputs respectively), and send their outputs (a, b) back to the referee. The referee accepts and Alice and Bob win the game iff $\mathbf{V}(x, y, a, b) = 1$.

Definition 4. The entangled value of a game $G = (q, \mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B}, \mathbf{V})$, denoted by $\omega^*(G)$, is the maximum winning probability of Alice and Bob, averaged over the distribution q as well as inherent randomness in the strategy, over all shared entanglement strategies for G .

3 Proof of direct product theorem

In this section, we prove Theorem 1, whose statement we recall below.

Theorem 1. *For any relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, and any $\varepsilon, \zeta > 0$,*

$$Q_{1-(1-\varepsilon)\Omega(\zeta^6 k / \log |\mathcal{Z}|)}^1(f^k) = \Omega\left(k\left(\zeta^5 \cdot Q_{\varepsilon+12\zeta}^1(f) - \log \log(1/\zeta)\right)\right).$$

Let p be the hard distribution on $\mathcal{X} \times \mathcal{Y}$ for $Q_{\varepsilon+12\zeta}^1(f)$ from Yao's lemma, i.e., $Q_{\varepsilon+12\zeta}^1(f) = Q_{p, \varepsilon+12\zeta}^1(f)$. Consider the relation $\tilde{f} \subseteq \mathcal{X} \times (\mathcal{Y} \cup \{y^*\}) \times \mathcal{Z}$ which is the same as f on $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ and additionally,

$$(x, y^*, z) \in \tilde{f} \quad \forall x \in \mathcal{X}, \forall z \in \mathcal{Z}.$$

We can think of p as a distribution on $\mathcal{X} \times (\mathcal{Y} \cup \{y^*\})$ as well, which has $p(y^*) = 0$. Clearly,

$$\mathbb{Q}_{p,\gamma}^1(\tilde{f}) = \mathbb{Q}_{p,\gamma}^1(f) \quad (1)$$

for any error γ , since p has no support on the extra inputs on which \tilde{f} is defined. We also note that

$$\mathbb{Q}_{\gamma}^1(f^k) \geq \mathbb{Q}_{\gamma}^1(\tilde{f}^k) \quad (2)$$

for any γ . This is because any protocol for f^k is also a protocol for \tilde{f}^k : on the indices where Bob's input is y^* instead of an element of \mathcal{Y} , he pretends he has gotten an input from \mathcal{Y} , runs the protocol with this input and gives the answer accordingly. This gives a correct output if the original protocol gives a correct output, since any output is correct when Bob's input in y^* .

For a distribution q related to p , we shall show that

$$\mathbb{Q}_{q^k, 1-(1-\varepsilon)\Omega(\zeta^6 k / \log |Z|)}^1(\tilde{f}^k) \geq \frac{\zeta^5 k}{300} \cdot \mathbb{Q}_{p, \varepsilon+12\zeta}^1(\tilde{f}) - k \log \log \left(\frac{24}{5\zeta} \right). \quad (3)$$

Since $\mathbb{Q}_{\gamma}^1(\tilde{f}^k) \geq \mathbb{Q}_{q^k, \gamma}^1(\tilde{f}^k)$, (1), (2) and (3) imply the theorem. The distribution q is defined as follows

$$\begin{aligned} q(x, y) &= (1 - \zeta) \cdot p(x, y) \quad \forall x \in \mathcal{X}, y \in \mathcal{Y} \\ q(x, y^*) &= \zeta \cdot p(x) \quad \forall x \in \mathcal{X}. \end{aligned}$$

Clearly, $q(x, y^*) = q(x)q(y^*)$ for all x , and

$$\|p(x, y) - q(x, y)\|_1 \leq 2\zeta. \quad (4)$$

Following [BVY15], for each $i \in [k]$, we shall define a joint distribution $\mathbb{P}_{X_i Y_i D_i G_i}$, where the marginal on $X_i Y_i$ is $q(x, y)$, and $D_i G_i$ are correlation-breaking variables such that conditioned on $D_i G_i = d_i g_i$, X_i and Y_i are independent. Each $X_i Y_i D_i G_i$ is distributed independently of the rest. Each D_i is distributed uniformly in $\{0, 1\}$. Depending on the value of D_i , G_i is distributed in the following way:

$$G_i = \begin{cases} x & \text{w.p. } p(x) & \text{if } D_i = 0 \\ y^* & \text{w.p. } 1 - (1 - \zeta)^{2/3} & \text{if } D_i = 1 \\ y & \text{w.p. } (1 - \zeta)^{2/3} \cdot p(y) & \text{if } D_i = 1 \end{cases}$$

Now depending on the value of $D_i G_i$, $X_i Y_i$ is distributed in the following way:

$$X_i Y_i = \begin{cases} (x, y^*) & \text{w.p. } \zeta & \text{if } D_i = 0, G_i = x \\ (x, y) & \text{w.p. } (1 - \zeta) \cdot p(y|x) & \text{if } D_i = 0, G_i = x \\ (x, y^*) & \text{w.p. } p(x) & \text{if } D_i = 1, G_i = y^* \\ (x, y^*) & \text{w.p. } (1 - (1 - \zeta)^{1/3}) \cdot p(x|y) & \text{if } D_i = 1, G_i = y \\ (x, y) & \text{w.p. } (1 - \zeta)^{1/3} \cdot p(x|y) & \text{if } D_i = 1, G_i = y. \end{cases}$$

The following lemma is similar to Claim 18 from [BVY15]; we provide a proof for completeness.

Lemma 26. *For all $(x, y) \in \mathcal{X} \times (\mathcal{Y} \cup \{y^*\})$, $\mathbb{P}_{X_i Y_i}(x, y) = q(x, y)$.*

Proof. It is trivial to see that $\mathbb{P}_{G_i Y_i | D_i=0}(x, y) = \mathbb{P}_{X_i Y_i | D_i=0}(x, y) = q(x, y)$, since $G_i = X_i$ conditioned on $D_i = 0$. We now prove the $D_i = 1$ case. First consider a $y \in \mathcal{Y}$. Y_i can only take value y if G_i takes value y . Hence,

$$\begin{aligned} \mathbb{P}_{X_i Y_i | D_i=1}(x, y) &= \mathbb{P}_{G_i | D_i=1}(y) \cdot \mathbb{P}_{X_i Y_i | D_i=1, G_i=y}(x, y) \\ &= (1 - \zeta)^{2/3} p(y) \cdot (1 - \zeta)^{1/3} p(x|y) \\ &= (1 - \zeta) \cdot p(x, y) = q(x, y). \end{aligned}$$

On the other hand, Y_i can take value y^* when $G_i = y^*$ or when $G_i = y$ for any $y \in \mathcal{Y}$. Hence,

$$\begin{aligned} \mathbb{P}_{X_i Y_i | D_i=1}(x, y^*) &= \mathbb{P}_{G_i | D_i=1}(y^*) \cdot \mathbb{P}_{X_i Y_i | D_i=1, G_i=y^*}(x, y^*) + \sum_{y \in \mathcal{Y}} \mathbb{P}_{G_i | D_i=1}(y) \cdot \mathbb{P}_{X_i Y_i | D_i=1, G_i=y}(x, y^*) \\ &= \left(1 - (1 - \zeta)^{2/3}\right) \cdot p(x) + (1 - \zeta)^{2/3} \left(1 - (1 - \zeta)^{1/3}\right) \sum_{y \in \mathcal{Y}} p(y) \cdot p(x|y) \\ &= \left(1 - (1 - \zeta)^{2/3}\right) \cdot p(x) + \left((1 - \zeta)^{2/3} - (1 - \zeta)\right) \cdot p(x) \\ &= \zeta \cdot p(x) = q(x, y^*). \end{aligned} \quad \square$$

In particular the lemma means $\mathbb{P}_{X_i Y_i}(x, y^*) = \mathbb{P}_{X_i}(x) \mathbb{P}_{Y_i}(y^*)$. We also note

$$\mathbb{P}_{Y_i G_i | D_i=1}(Y_i \neq G_i) = (1 - \zeta)^{2/3} (1 - (1 - \zeta)^{1/3}) \leq 1 - 2\zeta/3 - 1 + \zeta = \zeta/3. \quad (5)$$

To prove (3), let \mathcal{P} be any quantum one-way protocol between Alice and Bob, for $\tilde{f}^k \subseteq \mathcal{X}^k \times (\mathcal{Y} \cup \{y^*\})^k \times \mathcal{Z}^k$. \mathcal{P} is depicted in Figure 1. Alice and Bob's inputs are in registers $X = X_1 \dots X_k$ and $Y = Y_1 \dots Y_k$, and they share an entangled pure state uncorrelated with the inputs on registers $E_A E_B$, with Alice holding E_A and Bob holding E_B . Alice applies a unitary V^{Alice} on $X E_A$, to get the message register M , and the register A to be discarded. We shall use $|\theta\rangle_{AME_B|x}$ to refer to the pure state in AME_B in the protocol after Alice's unitary, for inputs xy ($|\theta\rangle_x$ only depends on y via x). When Alice and Bob's inputs are distributed according to \mathbb{P}_{XY} , the state of the protocol after Alice's message, will be given by the following CQ state:

$$\theta_{XYAME_B} = \sum_{xy} \mathbb{P}_{XY}(xy) |xy\rangle \langle xy|_{XY} \otimes |\theta\rangle \langle \theta|_{AME_B|x}.$$

We shall also consider the following purification of it, with the purifying registers \tilde{X} and \tilde{Y} :

$$|\theta\rangle_{X\tilde{X}Y\tilde{Y}AME_B} = \sum_{xy} \sqrt{\mathbb{P}_{XY}(xy)} |xxyy\rangle_{X\tilde{X}Y\tilde{Y}} |\theta\rangle_{AME_B|x}.$$

After receiving Alice's message, Bob applies a unitary V^{Bob} to YME_B , after which ME_B gets converted to BZ , where $Z = Z_1 \dots Z_k$ are the answer registers. We shall use $|\rho\rangle_{X\tilde{X}Y\tilde{Y}ABZ}$ to refer to $|\theta\rangle_{X\tilde{X}Y\tilde{Y}AME_B}$ after V^{Bob} . We shall use \mathbb{P}_{XYDGZ} to refer to the joint distribution of these variables in $|\rho\rangle$, where the Z distribution is obtained by measuring the Z register in the computational basis.

We shall show that if the communication cost of \mathcal{P} is $< \frac{\zeta^5 k}{300} \cdot \mathbb{Q}_{p, \varepsilon+12\zeta}^1(\tilde{f}) - k \log \log(24/5\zeta)$, then the success probability of \mathcal{P} is $(1 - \varepsilon)^{\Omega(\zeta^6 k / \log |Z|)}$. This is implied by the following claim, which the rest of the proof will show.

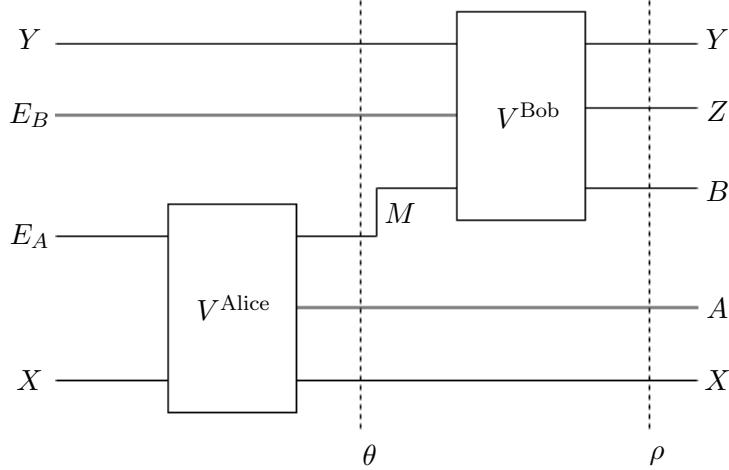


Figure 1: One-way quantum protocol \mathcal{P}

Lemma 27. Let $\delta = \frac{\zeta^6}{1440000}$ and $\delta' = \frac{\zeta^6}{1440000 \log |\mathcal{Z}|}$. For $i \in [k]$, let T_i be the random variable which takes value 1 if \mathcal{P} computes $f(X_i, Y_i)$ correctly, and value 0 otherwise. If the communication cost of \mathcal{P} is $< \frac{\zeta^5 k}{300} \cdot \mathbf{Q}_{p, \varepsilon + 12\zeta}^1(\tilde{f}) - k \log \log(24/5\zeta)$, then there exist $\lfloor \delta' k \rfloor$ coordinates $\{i_1, \dots, i_{\lfloor \delta' k \rfloor}\} \subseteq [k]$, such that for all $1 \leq r \leq \lfloor \delta' k \rfloor - 1$, at least one of the following two conditions holds

- (i) $\Pr \left[\prod_{j=1}^r T_{i_j} = 1 \right] \leq (1 - \varepsilon)^{\delta k}$
- (ii) $\Pr \left[T_{i_{r+1}} = 1 \mid \prod_{j=1}^r T_{i_j} = 1 \right] \leq 1 - \varepsilon.$

Lemma 27 can be proved inductively. Suppose we have already identified $1 \leq t \leq \lfloor \delta' k \rfloor$ coordinates in $C = \{i_1, \dots, i_t\}$, such that for all $1 \leq r \leq t - 1$, $\Pr \left[T_{i_{r+1}} = 1 \mid \prod_{j=1}^r T_{i_j} = 1 \right] \leq 1 - \varepsilon.$ Let \mathcal{E} refer to the event $\prod_{i \in C} T_i = 1.$ If $\Pr[\mathcal{E}] \leq (1 - \varepsilon)^{\delta k},$ then we are already done. If not, then we shall show how to identify the $(t+1)$ -th coordinate i such that $\Pr [T_i = 1 \mid \mathcal{E}] \leq 1 - \varepsilon.$ The process of identifying the first coordinate is also similar, except in that case the conditioning event is empty. Since we only use the lower bound $(1 - \varepsilon)^{\delta k}$ on the probability of the conditioning event in our proof, the proof goes through for that case as well.

We shall use the state $|\varphi\rangle,$ which is $|\rho\rangle_{X\tilde{X}Y\tilde{Y}ABZ}$ conditioned on $\mathcal{E},$ for the proof of Lemma 27. For any value $DG = dg,$ $|\varphi\rangle_{X\tilde{X}Y\tilde{Y}ABZ|dg}$ is defined as:

$$|\varphi\rangle_{X\tilde{X}Y\tilde{Y}ABZ|dg} = \frac{1}{\sqrt{\gamma_{dg}}} \sum_{xy} \sqrt{\mathbf{P}_{XY|dg}(xy)} |xyxy\rangle_{X\tilde{X}Y\tilde{Y}} \otimes \sum_{z_C: (x_C, y_C, z_C) \in \tilde{f}^t} |z_C\rangle_{Z_C} |\tilde{\varphi}\rangle_{ABZ_{\bar{C}}|xyz_C}.$$

Here $|\tilde{\varphi}\rangle_{xyz_C}$ is a subnormalized state with $\| |\tilde{\varphi}\rangle_{ABZ_{\bar{C}}|xyz_C} \|_2^2 = \mathbf{P}_{Z_C|xy}(z_C).$ The overall normalization factor γ_{dg} is the probability of \mathcal{E} conditioned on $dg,$ and satisfies

$$\sum_{dg} \mathbf{P}_{DG}(dg) \cdot \gamma_{dg} = \Pr[\mathcal{E}].$$

It is clear that the distribution of XYZ in $|\varphi\rangle_{X\tilde{X}Y\tilde{Y}ABZ|dg}$ is $P_{XYZ|\mathcal{E},dg}$. Note that we are using the notation $|\varphi\rangle_{dg}$ without explicitly considering registers DG on which a measurement is done to obtain $|\varphi\rangle_{dg}$. We shall also sometimes use $|\varphi\rangle_{d_{-i}g_{-i}}$ in which the xy distributions are conditioned on $d_{-i}g_{-i}$ instead, which changes the normalization factor to some $\gamma_{d_{-i}g_{-i}}$, everything else remaining the same. $\varphi_{x_i y_i d_{-i} g_{-i}}$ refers as usual to the state obtained when a measurement done on the $X_i Y_i$ registers (which are actually present in $|\varphi\rangle$) in $|\varphi\rangle_{d_{-i}g_{-i}}$. For $i \notin \bar{C}$, we shall use the states $|\varphi\rangle_{X_{\bar{C}}\tilde{X}_{\bar{C}}Y_{\bar{C}}\tilde{Y}_{\bar{C}}ABZ_{\bar{C}}|x_i y_i x_C y_C z_C d_{-i} g_{-i}}$ in our proof, which we note are pure states.

Lemma 27 will be proved with the help of the following lemma, whose proof we give later.

Lemma 28. *If $\Pr[\mathcal{E}] \geq (1 - \varepsilon)^{\delta k}$, then there exist a coordinate $i \in \bar{C}$, a random variable $R_i = X_C Y_C Z_C D_{-i} G_{-i}$ and for each $R_i = r_i$ a state $|\varphi'\rangle_{X_{\bar{C}}\tilde{X}_{\bar{C}}Y_{\bar{C}}\tilde{Y}_{\bar{C}}ABZ_{\bar{C}}|y^* r_i}$ such that the following conditions hold:*

$$(i) \quad \|\mathbb{P}_{X_i Y_i R_i | \mathcal{E}} - \mathbb{P}_{X_i Y_i} \mathbb{P}_{R_i | \mathcal{E}, X_i}\|_1 \leq \frac{7\zeta}{120}$$

$$(ii) \quad \|\mathbb{P}_{X_i Y_i R_i | \mathcal{E}} - \mathbb{P}_{X_i Y_i} \mathbb{P}_{R_i | \mathcal{E}, Y_i}\|_1 \leq \frac{7\zeta}{120}.$$

There exist projectors $\{\Pi_{x_i r_i}\}_{x_i r_i}$ acting only on registers $X_{\bar{C}}\tilde{X}_{\bar{C}}A$ and unitaries $\{U_{y_i r_i}\}_{y_i r_i}$ acting only on $Y_{\bar{C}}\tilde{Y}_{\bar{C}}BZ_{\bar{C}}$, such that each $\Pi_{x_i r_i}$ succeeds on $|\varphi'\rangle_{r_i}$ with probability $\alpha_{r_i} = 2^{-c' r_i}$, and

$$(iii) \quad \mathbb{E}_{\mathbb{P}_{R_i | \mathcal{E}}} c'_{r_i} \leq \frac{300c}{\zeta^5}$$

$$(iv) \quad \mathbb{E}_{\mathbb{P}_{X_i Y_i R_i | \mathcal{E}}} \left\| \frac{1}{\alpha_{r_i}} (\Pi_{x_i r_i} \otimes U_{y_i r_i}) |\varphi'\rangle \langle \varphi'|_{y^* r_i} (\Pi_{x_i r_i} \otimes U_{y_i r_i}^\dagger) - |\varphi\rangle \langle \varphi|_{x_i y_i r_i} \right\|_1 \leq 21\zeta.$$

Proof of Lemma 27. We give a one-way quantum protocol \mathcal{P}' for \tilde{f} , whose inputs are distributed according to $\mathbb{P}_{X_i Y_i}$, i.e., q , by embedding Alice and Bob's inputs into the i -th coordinate of $|\varphi\rangle_{x_i y_i r_i}$, as follows:

- Alice and Bob have r according to the distribution required by Fact 11 as shared randomness, and $2^{300c/\zeta^5} \log(24/5\zeta)$ copies of $|\varphi'\rangle_{y^* r_i}$ as shared entanglement, with Alice holding registers $X_{\bar{C}}\tilde{X}_{\bar{C}}A$ and Bob holding registers $Y_{\bar{C}}\tilde{Y}_{\bar{C}}BZ_{\bar{C}}$ of each copy.
- On input (x_i, y_i) from $\mathbb{P}_{X_i Y_i}$, using items (i), (ii) of Lemma 28, their shared randomness, and the protocol from Fact 11, Alice and Bob generate random variables $R_i^{\text{Alice}} R_i^{\text{Bob}}$ such that

$$\|\mathbb{P}_{X_i Y_i R_i^{\text{Alice}} R_i^{\text{Bob}}} - \mathbb{P}_{X_i Y_i R_i R_i | \mathcal{E}}\|_1 \leq \frac{7\zeta}{24}.$$

where $R_i R_i$ denotes two perfectly correlated copies of R_i in $\mathbb{P}_{X_i Y_i R_i R_i | \mathcal{E}}$.

- Alice applies the $\{\Pi_{x_i r_i^A}, \mathbb{1} - \Pi_{x_i r_i^A}\}$ measurement according to her input and R_i^{Alice} on her registers for each copy of the shared entangled state. If the $\Pi_{x_i r_i^A}$ measurement does not succeed on any copy, then she aborts. Otherwise, she sends to Bob a $(\frac{300c}{\zeta^5} + \log \log(24/5\zeta))$ -bit message indicating an index where $\Pi_{x_i r_i^A}$ measurement succeeded.

- Bob applies the unitary $U_{y_i r_i^B}$ according to his input and R_i^{Bob} on the copy of the shared entangled state whose index Alice has sent, and measures the Z_i register of the resulting state to give her output.

To analyze the success of this protocol, first note that

$$\mathbb{E}_{\mathbb{P}_{X_i Y_i R_i | \mathcal{E}}} \Pr[\text{Result of } Z_i \text{ measurement on } |\varphi\rangle_{x_i y_i r_i} \in \tilde{f}(x_i, y_i)] = \Pr[T_i = 1 | \mathcal{E}].$$

Let us first assume Alice and Bob have (x_i, y_i, r_i^A, r_i^B) distributed exactly according to $\mathbb{P}_{X_i Y_i R_i | \mathcal{E}}$ – we shall denote both r_i^A and r_i^B by r_i in this case. Alice aborts the protocol if none of her measurements succeed. On expectation, this happens with probability

$$\mathbb{E}_{\mathbb{P}_{R_i | \mathcal{E}}} \left[(1 - 2^{-c'_{r_i}})^{300c/\zeta^5 \log(24/7\zeta)} \right] \leq \left(1 - 2^{-\mathbb{E}_{\mathbb{P}_{R_i | \mathcal{E}}} c'_{r_i}} \right)^{300c/\zeta^5 \log(24/5\zeta)} \leq \frac{5\zeta}{24}$$

from (iii). If Alice does not abort, then Alice and Bob's state after Bob's unitary is $\frac{1}{\sqrt{\alpha_{r_i}}} (\Pi_{x_i r_i} \otimes U_{y_i r_i} |\varphi'\rangle_{y^* r_i})$. From (iv), the expected probability of the Z_i measurement on this state giving an answer $\in \tilde{f}(x_i, y_i)$ is at least $\Pr[T_i = 1 | \mathcal{E}] - \frac{21\zeta}{2}$. Hence, if Alice and Bob had (x_i, y_i, r_i^A, r_i^B) distributed according to $\mathbb{P}_{X_i Y_i R_i | \mathcal{E}}$, then their expected success probability would have been at least $\Pr[T_i = 1 | \mathcal{E}] - \frac{21\zeta}{2} - \frac{5\zeta}{24}$. Since Alice and Bob have (x_i, y_i, r_i^A, r_i^B) according to $\mathbb{P}_{X_i Y_i R_i^{\text{Alice}} R_i^{\text{Bob}}}$ instead, their expected success probability is at least

$$\Pr[T_i = 1 | \mathcal{E}] - \frac{21\zeta}{2} - \frac{5\zeta}{24} - \frac{7\zeta}{24} \geq \Pr[T_i = 1 | \mathcal{E}] - 11\zeta.$$

Since $\|q(x, y) - p(x, y)\|_1 \leq 2\zeta$, when the same protocol is run on $X_i Y_i$ distributed according to p instead, it must succeed with probability at least $\Pr[T_i = 1 | \mathcal{E}] - 12\zeta$. Since the communication in \mathcal{P}' is at most $(\frac{300c}{\zeta^5} + \log \log(24/5\zeta)) < Q_{p, \varepsilon + 12\zeta}^1(\tilde{f})$, $\Pr[T_i = 1 | \mathcal{E}] \geq 1 - \varepsilon$ gives the error probability of \mathcal{P}' to be $\leq \varepsilon + 12\zeta$, which is a contradiction. Hence we must have $\Pr[T_i = 1 | \mathcal{E}] \leq 1 - \varepsilon$. The desired result thus follows by setting $i_{t+1} = i$. \square

Proof of Lemma 28. Applying Fact 10 with T and V being trivial and $U_i = X_i Y_i D_i G_i$ for $i \in \bar{C}$, we get,

$$\mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i Y_i D_i G_i | \mathcal{E}} - \mathbb{P}_{X_i Y_i D_i G_i}\|_1 \leq \frac{1}{k-t} \sqrt{k \cdot \log((1-\varepsilon)^{\delta k})} \leq \sqrt{2\delta}. \quad (6)$$

In particular, due to (5), this means

$$\mathbb{E}_{i \in \bar{C}} \mathbb{P}_{Y_i G_i | \mathcal{E}, D_i=1}(Y_i = G_i) \geq 1 - \zeta/3 - \sqrt{2\delta}. \quad (7)$$

And since $\mathbb{P}_{G_i | D_i=1}(y^*) = 1 - (1-\zeta)^{2/3}$, $\mathbb{P}_{Y_i | D_i=1, G_i=y}(y_i) = (1-\zeta)^{1/3}$ for $y_i \in \mathcal{Y}$, we have

$$\zeta + \sqrt{2\delta} \geq 1 - (1-\zeta)^{2/3} + \sqrt{2\delta} \geq \mathbb{E}_{i \in \bar{C}} \mathbb{P}_{G_i | \mathcal{E}, D_i=1}(y^*) \geq 1 - (1-\zeta)^{2/3} - \sqrt{2\delta} \geq 2\zeta/3 - \sqrt{2\delta} \quad (8)$$

$$(1 - \zeta/3 + \sqrt{2\delta}) \cdot \mathbb{E}_{i \in \bar{C}} \mathbb{P}_{G_i | \mathcal{E}, D_i=1}(y_i) \geq \mathbb{E}_{i \in \bar{C}} \mathbb{P}_{Y_i G_i | \mathcal{E}, D_i=1}(y_i, y_i) \geq (1 - \zeta - \sqrt{2\delta}) \cdot \mathbb{E}_{i \in \bar{C}} \mathbb{P}_{G_i | \mathcal{E}, D_i=1}(y_i). \quad (9)$$

Fact 10 can again be applied with $U_i = X_i Y_i$, $T = X_C Y_C D G$ and $V = Z_C$. Let $\delta_1 = \delta + \delta' \log |\mathcal{Z}| = \frac{\zeta^6}{720000}$. Then we have,

$$\begin{aligned}
\sqrt{2\delta_1} &\geq \mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i Y_i X_C Y_C Z_C D G | \mathcal{E}} - \mathbb{P}_{X_C Y_C Z_C D G | \mathcal{E}} \mathbb{P}_{X_i Y_i | X_C Y_C D G}\|_1 \\
&= \mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i Y_i X_C Y_C Z_C D G | \mathcal{E}} - \mathbb{P}_{X_C Y_C Z_C D G | \mathcal{E}} \mathbb{P}_{X_i Y_i | D_i G_i}\|_1 \\
&= \mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i Y_i D_i G_i R_i | \mathcal{E}} - \mathbb{P}_{D_i G_i R_i | \mathcal{E}} \mathbb{P}_{X_i Y_i | D_i G_i}\|_1.
\end{aligned} \tag{10}$$

We note that D_i takes value uniformly in $\{0, 1\}$ even conditioned on \mathcal{E} . Hence from (10),

$$\begin{aligned}
\sqrt{2\delta_1} &\geq \frac{1}{2} \mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i Y_i G_i R_i | \mathcal{E}, D_i=0} - \mathbb{P}_{G_i R_i | \mathcal{E}, D_i=0} \mathbb{P}_{X_i Y_i | G_i, D_i=0}\|_1 \\
&= \frac{1}{2} \mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i Y_i R_i | \mathcal{E}} - \mathbb{P}_{X_i R_i | \mathcal{E}} \mathbb{P}_{Y_i | X_i}\|_1
\end{aligned}$$

where we have used the fact that $X_i = G_i$ conditioned on $D_i = 0$. Combining this with the fact that $\mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i | \mathcal{E}} - \mathbb{P}_{X_i}\|_1 \leq \sqrt{2\delta}$, we have,

$$\mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i Y_i R_i | \mathcal{E}} - \mathbb{P}_{X_i Y_i} \mathbb{P}_{R_i | \mathcal{E}, X_i}\|_1 \leq 3\sqrt{2\delta_1} < \frac{7\zeta^3}{600}. \tag{11}$$

Due to Corollary 9 we also have from (11),

$$\mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i R_i | \mathcal{E}, y^*} - \mathbb{P}_{X_i R_i | \mathcal{E}}\|_1 \leq \frac{33\sqrt{2\delta_1}}{\zeta}. \tag{12}$$

Let \mathcal{F}_i denote the event $Y_i = G_i$. We know $\mathbb{E}_{i \in \bar{C}} \mathbb{P}_{X_i Y_i G_i | D_i=1}(\mathcal{F}_i) \geq 1 - \zeta/3 - \sqrt{2\delta}$, from (7). Hence, using Fact 7,

$$\begin{aligned}
\mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i Y_i R_i | \mathcal{E}} - \mathbb{P}_{Y_i R_i | \mathcal{E}} \mathbb{P}_{X_i | Y_i}\|_1 &= \mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i Y_i G_i R_i | \mathcal{E}, D_i=1, \mathcal{F}_i} - \mathbb{P}_{G_i R_i | \mathcal{E}, D_i=1, \mathcal{F}_i} \mathbb{P}_{X_i Y_i | G_i, D_i=1, \mathcal{F}_i}\|_1 \\
&\leq 6 \mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i Y_i D_i R_i | \mathcal{E}, G_i=1} - \mathbb{P}_{G_i R_i | \mathcal{E}, D_i=1} \mathbb{P}_{X_i Y_i | G_i, D_i=1}\|_1 \leq 6\sqrt{2\delta_1}.
\end{aligned}$$

Using $\mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{Y_i | \mathcal{E}} - \mathbb{P}_{Y_i}\|_1 \leq \sqrt{2\delta}$, we have as before,

$$\mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i Y_i R_i | \mathcal{E}} - \mathbb{P}_{X_i Y_i} \mathbb{P}_{R_i | \mathcal{E}, Y_i}\|_1 \leq 7\sqrt{2\delta_1} = \frac{7\zeta^3}{600}. \tag{13}$$

Let M be ck qubits. By Fact 20, for any value $DG = dg$, there exists some state $\sigma_{M|dg}$ such that

$$S_\infty(\theta_{XY\tilde{Y}E_B M|dg} \|\theta_{XY\tilde{Y}E_B|dg} \otimes \sigma_{M|dg}) \leq 2ck.$$

By Fact 19 we have,

$$S_\infty\left(\rho_{XY\tilde{Y}BZ|dg} \|V^{\text{Bob}}(\theta_{XY\tilde{Y}E_B|dg} \otimes \sigma_{M|dg})(V^{\text{Bob}})^\dagger\right) \leq 2ck.$$

Let $\psi_{X_C Y_C \tilde{Y}_C B Z_C | dg} = \text{Tr}_{Z_C}(V^{\text{Bob}}(\theta_{XYE_B|dg} \otimes \sigma_{M|X_C Y_C dg})(V^{\text{Bob}})^\dagger)$. Note that $\theta_{XY\tilde{Y}E_B|dg} \otimes \sigma_{M|dg}$ is product across X and the other registers, and V^{Bob} does not act on X . Hence $\psi_{X_C Y_C \tilde{Y}_C B Z_C | dg}$ is

also product across X and the other registers, and moreover, all the X_i -s are in product with each other as well. We have,

$$\mathbb{S}_\infty \left(\rho_{XY\tilde{Y}BZ_{\bar{C}}|dg} \|\psi_{XY\tilde{Y}BZ_{\bar{C}}|dg} \right) \leq 2ck.$$

Using Facts 21 and 18, this gives us

$$\begin{aligned} & \mathbb{E}_{\mathbb{P}_{X_C Y_C Z_C DG|\mathcal{E}}} \left[\mathbb{S} \left(\varphi_{X_{\bar{C}} Y_{\bar{C}} \tilde{Y}_{\bar{C}} BZ_{\bar{C}}|x_C y_C z_C dg} \|\psi_{X_{\bar{C}} Y_{\bar{C}} \tilde{Y}_{\bar{C}} BZ_{\bar{C}}|x_C y_C dg} \right) \right] \\ & \leq \mathbb{E}_{\mathbb{P}_{Z_C DG|\mathcal{E}}} \left[\mathbb{S} \left(\varphi_{XY\tilde{Y}BZ_{\bar{C}}|z_C dg} \|\psi_{XY\tilde{Y}BZ_{\bar{C}}|dg} \right) \right] \\ & \leq \mathbb{E}_{\mathbb{P}_{Z_C DG|\mathcal{E}}} \left[\mathbb{S}_\infty \left(\varphi_{XY\tilde{Y}BZ_{\bar{C}}|z_C dg} \|\psi_{XY\tilde{Y}BZ_{\bar{C}}|dg} \right) \right] \\ & \leq \mathbb{E}_{\mathbb{P}_{Z_C DG|\mathcal{E}}} \left[\mathbb{S}_\infty \left(\varphi_{XY\tilde{Y}BZ_{\bar{C}}|z_C dg} \|\varphi_{XY\tilde{Y}BZ_{\bar{C}}|dg} \right) \right. \\ & \quad \left. + \mathbb{S}_\infty \left(\varphi_{XY\tilde{Y}BZ_{\bar{C}}|dg} \|\rho_{XY\tilde{Y}BZ_{\bar{C}}|dg} \right) + \mathbb{S}_\infty \left(\rho_{XY\tilde{Y}BZ_{\bar{C}}|dg} \|\psi_{XY\tilde{Y}BZ_{\bar{C}}|dg} \right) \right] \\ & \leq \mathbb{E}_{\mathbb{P}_{Z_C DG|\mathcal{E}}} \left[\log(1/\mathbb{P}_{Z_C|\mathcal{E}}(z_C)) + \log(1/\Pr[\mathcal{E}]) + 2ck \right] \\ & \leq |C| \log |Z| + \delta k + 2ck \leq (\delta_1 + 2c)k. \end{aligned}$$

By Quantum Raz's Lemma,

$$\begin{aligned} 4c + 2\delta_1 & \geq \mathbb{E}_{i \in \bar{C}} \mathbb{E}_{\mathbb{P}_{X_C Y_C Z_C DG|\mathcal{E}}} \mathbb{I}(X_i : Y_{\bar{C}} \tilde{Y}_{\bar{C}} BZ_{\bar{C}})_{\varphi_{x_C y_C z_C dg}} \\ & = \mathbb{E}_{i \in \bar{C}} \mathbb{E}_{\mathbb{P}_{D_i G_i R_i|\mathcal{E}}} \mathbb{I}(X_i : Y_{\bar{C}} \tilde{Y}_{\bar{C}} BZ_{\bar{C}})_{\varphi_{d_i g_i r_i}} \\ & \geq \mathbb{E}_{i \in \bar{C}} \frac{1}{2} \mathbb{P}_{G_i|\mathcal{E}, D_i=1}(y^*) \mathbb{E}_{\mathbb{P}_{R_i|\mathcal{E}, D_i=1, G_i=y^*}} \mathbb{I}(X_i : Y_{\bar{C}} \tilde{Y}_{\bar{C}} BZ_{\bar{C}})_{\varphi_{r_i, D_i=1, G_i=y^*}} \\ & \geq \mathbb{E}_{i \in \bar{C}} \frac{1}{2} (2\zeta/3 - \sqrt{2\delta}) \mathbb{E}_{\mathbb{P}_{R_i|\mathcal{E}, D_i=1, G_i=y^*}} \mathbb{I}(X_i : Y_{\bar{C}} \tilde{Y}_{\bar{C}} BZ_{\bar{C}})_{\varphi_{r_i, D_i=1, G_i=y^*}} \end{aligned} \quad (14)$$

where we have used (8) in the last inequality.

Note that $\varphi_{X_{\bar{C}} \tilde{X}_{\bar{C}} Y_{\bar{C}} \tilde{Y}_{\bar{C}} ABZ_{\bar{C}}|x_i r_i, D_i=1, G_i=y^*}$ is the same state as $\varphi_{X_{\bar{C}} \tilde{X}_{\bar{C}} Y_{\bar{C}} \tilde{Y}_{\bar{C}} ABZ_{\bar{C}}|x_i y^* r_i}$, where the value of Y_i is being conditioned on, instead of G_i . $|\varphi\rangle_{r_i, D_i=1, G_i=y^*}$ is the superposition over X_i of $|\varphi\rangle_{x_i r_i, D_i=1, G_i=y^*}$, with the X_i distribution being $\mathbb{P}_{X_i|\mathcal{E}, r_i, D_i=1, G_i=y^*}$. The only difference between $|\varphi\rangle_{y^* r_i}$ and $|\varphi\rangle_{r_i, D_i=1, G_i=y^*}$ is the X_i distribution, which in the former is $\mathbb{P}_{X_i|\mathcal{E}, y^* r_i}$ instead. We shall refer to $|\varphi\rangle_{r_i, D_i=1, G_i=y^*}$ as simply $|\varphi\rangle_{r_i, 1, y^*}$ as now on – note that there is no ambiguity between this and $|\varphi\rangle_{y^* r_i}$. The same goes for the distributions $\mathbb{P}_{X_i R_i|\mathcal{E}, 1, y^*}$ and $\mathbb{P}_{X_i R_i|\mathcal{E}, y^*}$.

$\mathbb{P}_{X_i|1, y^*}$ is the same distribution as $\mathbb{P}_{X_i|y^*}$ and $\mathbb{P}_{R_i|\mathcal{E}, x_i, 1, y^*}$ is the same distribution as $\mathbb{P}_{R_i|\mathcal{E}, x_i y^*}$ for any x_i . Hence,

$$\begin{aligned} \mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i R_i|\mathcal{E}, y^*} - \mathbb{P}_{X_i R_i|\mathcal{E}, 1, y^*}\|_1 & \leq \mathbb{E}_{i \in \bar{C}} \left[\|\mathbb{P}_{X_i R_i|\mathcal{E}, y^*} - \mathbb{P}_{X_i|y^*} \mathbb{P}_{R_i|\mathcal{E}, X_i, y^*}\|_1 \right. \\ & \quad \left. + \|(\mathbb{P}_{X_i|1, y^*} - \mathbb{P}_{X_i|\mathcal{E}, 1, y^*}) \mathbb{P}_{R_i|\mathcal{E}, X_i, y^*}\|_1 \right] \\ & \leq \mathbb{E}_{i \in \bar{C}} \left[\frac{\|\mathbb{P}_{X_i R_i|\mathcal{E}} - \mathbb{P}_{X_i} \mathbb{P}_{R_i|\mathcal{E}, X_i}\|_1}{2\zeta/3 - \sqrt{2\delta}} + \frac{\|\mathbb{P}_{X_i|\mathcal{E}} - \mathbb{P}_{X_i}\|_1}{2\zeta/3 - \sqrt{2\delta}} \right] \end{aligned}$$

$$\leq \frac{7\sqrt{2\delta_1}}{\zeta}$$

where we have used (8) in the second inequality. Using the above computation and (12), we get,

$$\mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i R_i | \mathcal{E}} - \mathbb{P}_{X_i R_i | \mathcal{E}, 1, y^*}\|_1 \leq \frac{40\sqrt{2\delta_1}}{\zeta}.$$

Let $|\varphi'\rangle_{y^* r_i}$ denote the pure state where the distribution of X_i is unconditioned on $Y_i = y^*$, but everything else is conditioned. From (14) and Fact 22, we then have that,

$$\mathbb{E}_{i \in \bar{C}} \mathbb{P}_{R_i | \mathcal{E}} \left(I_{\max}^{\zeta + 280\sqrt{2\delta_1}/\zeta^2} (X_i : Y_{\bar{C}} \tilde{Y}_{\bar{C}} B Z_{\bar{C}})_{\varphi'_{y^* r_i}} > \frac{28(2c + \delta_1) + 1}{\zeta^4} \right) \leq 2\zeta + \frac{20\sqrt{2\delta_1}}{\zeta}.$$

Hence by Fact 24, there exist projectors $\Pi_{x_i r_i}$ acting on registers $X_{\bar{C}} \tilde{X}_{\bar{C}} A$, such that $\Pi_{x_i r_i}$ succeeds with probability $\alpha_{r_i} = 2^{-c'_{r_i}}$ on $|\varphi'\rangle_{X_{\bar{C}} \tilde{X}_{\bar{C}} Y_{\bar{C}} \tilde{Y}_{\bar{C}} A B Z_{\bar{C}} | y^* r_i}$, where

$$\mathbb{E}_{i \in \bar{C}} \mathbb{E}_{\mathbb{P}_{R_i | \mathcal{E}}} c'_{r_i} \leq \frac{1}{\zeta} \cdot \frac{28(2c + \delta_1) + 1}{\zeta^4} \leq \frac{60c}{\zeta^5} \quad (15)$$

$$\mathbb{E}_{i \in \bar{C}} \mathbb{E}_{\mathbb{P}_{X_i R_i | \mathcal{E}}} \left\| \frac{1}{\alpha_{r_i}} (\Pi_{x_i r_i} \otimes \mathbb{1}) |\varphi'\rangle \langle \varphi' |_{y^* r_i} (\Pi_{x_i r_i} \otimes \mathbb{1}) - |\varphi\rangle \langle \varphi |_{x_i y^* r_i} \right\|_1 \leq 3\zeta + \frac{300\sqrt{2\delta_1}}{\zeta^2} \leq \frac{7\zeta}{2}. \quad (16)$$

By similar arguments as the ones leading to (14) on Bob's side (except the first step where we consider the information due to the message sent by Alice to Bob, which does not apply here), we can also upper bound $\mathbb{E}_{\mathbb{P}_{X_C Y_C Z_C D_G | \mathcal{E}}} \left[\mathbb{S} \left(\varphi_{Y_{\bar{C}} X_{\bar{C}} \tilde{X}_{\bar{C}} A | x_C y_C z_C d_G} \parallel \rho_{Y_{\bar{C}} X_{\bar{C}} \tilde{X}_{\bar{C}} A | x_C y_C d_G} \right) \right]$. Hence by Raz's lemma again,

$$\begin{aligned} 2\delta_1 &\geq \mathbb{E}_{i \in \bar{C}} \mathbb{E}_{\mathbb{P}_{D_i G_i R_i | \mathcal{E}}} I(Y_i : X_{\bar{C}} \tilde{X}_{\bar{C}} A)_{\varphi_{d_i g_i r_i}} \\ &\geq \mathbb{E}_{i \in \bar{C}} \frac{1}{2} (1 - \zeta - \sqrt{2\delta}) \mathbb{E}_{\mathbb{P}_{R_i G_i | \mathcal{E}, D_i=1, G_i \neq y^*}} I(Y_i : X_{\bar{C}} \tilde{X}_{\bar{C}} A)_{\varphi_{r_i, D_i=1, g_i}} \\ &= \mathbb{E}_{i \in \bar{C}} \frac{1}{2} (1 - \zeta - \sqrt{2\delta}) \mathbb{E}_{\mathbb{P}_{R_i G_i Y_i | \mathcal{E}, D_i=1, G_i \neq y^*}} \left[\mathbb{S} \left(\varphi_{X_{\bar{C}} \tilde{X}_{\bar{C}} A | y_i, D_i=1, g_i} \parallel \varphi_{X_{\bar{C}} \tilde{X}_{\bar{C}} A | D_i=1, g_i} \right) \right] \\ &\geq \mathbb{E}_{i \in \bar{C}} \frac{1}{2} (1 - \zeta - \sqrt{2\delta}) \sum_{y_i \in \mathcal{Y}} \mathbb{E}_{\mathbb{P}_{R_i | \mathcal{E}, D_i=1, G_i=y_i}} \mathbb{P}_{G_i | \mathcal{E}, D_i=1}(y_i) \\ &\quad \left[(1 - \zeta - \sqrt{2\delta}) \|\varphi_{X_{\bar{C}} \tilde{X}_{\bar{C}} A | y_i, r_i, D_i=1, G_i=y_i} - \varphi_{X_{\bar{C}} \tilde{X}_{\bar{C}} A | r_i, D_i=1, G_i=y_i}\|_1^2 \right. \\ &\quad \left. + (\zeta/3 - \sqrt{2\delta}) \|\varphi_{X_{\bar{C}} \tilde{X}_{\bar{C}} A | y^*, r_i, D_i=1, G_i=y_i} - \varphi_{X_{\bar{C}} \tilde{X}_{\bar{C}} A | r_i, D_i=1, G_i=y_i}\|_1^2 \right]. \end{aligned}$$

where we have used (9) and Pinsker's inequality in the last line. Hence by triangle inequality we have,

$$\mathbb{E}_{i \in \bar{C}} \sum_{y_i \in \mathcal{Y}} \mathbb{E}_{\mathbb{P}_{R_i | \mathcal{E}, 1, y_i}} \mathbb{P}_{G_i | \mathcal{E}, 1}(y_i) \|\varphi_{X_{\bar{C}} \tilde{X}_{\bar{C}} A | y_i r_i, 1, y_i} - \varphi_{X_{\bar{C}} \tilde{X}_{\bar{C}} A | y^* r_i, 1, y_i}\|_1^2 \leq \frac{32\delta_1}{\zeta}.$$

We note that $\varphi_{X_{\bar{C}}\tilde{X}_{\bar{C}}Y_{\bar{C}}\tilde{Y}_{\bar{C}}ABZ_{\bar{C}}|y_i r_i, 1, y_i}$ and $\varphi_{X_{\bar{C}}\tilde{X}_{\bar{C}}Y_{\bar{C}}\tilde{Y}_{\bar{C}}ABZ_{\bar{C}}|y^* r_i, 1, y_i}$ are pure states. Hence, using the Fuchs-van de Graaf inequality and Uhlmann's theorem, there exist unitaries $U_{y_i r_i}$ acting only on $Y_{\bar{C}}\tilde{Y}_{\bar{C}}BZ_{\bar{C}}$ such that

$$\mathbb{E}_{i \in \bar{C}} \sum_{y_i \in \mathcal{Y}} \mathbb{E}_{P_{R_i|\mathcal{E}, 1, y_i}} P_{G_i|\mathcal{E}, 1}(y_i) \|\varphi\rangle\langle\varphi|_{y_i r_i, 1, y_i} - (\mathbb{1} \otimes U_{y_i r_i})|\varphi\rangle\langle\varphi|_{y^* r_i, 1, y_i} (\mathbb{1} \otimes U_{y_i r_i}^\dagger)\|_1 \leq \left(\frac{32\delta_1}{\zeta}\right)^{1/4}. \quad (17)$$

Now consider the superoperator \mathcal{O}_{X_i} that measures the register X_i and writes it in a different register.

$$\begin{aligned} \mathcal{O}_{X_i}(|\varphi\rangle\langle\varphi|_{y_i r_i, 1, y_i}) &= \sum_{x_i} P_{X_i|\mathcal{E}, y_i r_i, D_i=1, G_i=y_i}(x_i) |x_i\rangle\langle x_i| \otimes |\varphi\rangle\langle\varphi|_{x_i y_i r_i, 1, y_i} \\ &= \sum_{x_i} P_{X_i|\mathcal{E}, y_i r_i, D_i=1, G_i=y_i}(x_i) |x_i\rangle\langle x_i| \otimes |\varphi\rangle\langle\varphi|_{x_i y_i r_i} \\ \mathcal{O}_{X_i}(|\varphi\rangle\langle\varphi|_{y^* r_i, 1, y_i}) &= \sum_{x_i} P_{X_i|\mathcal{E}, y^* r_i, D_i=1, G_i=y_i}(x_i) |x_i\rangle\langle x_i| \otimes |\varphi\rangle\langle\varphi|_{x_i y^* r_i} \end{aligned}$$

where we have made the observation that $|\varphi\rangle\langle\varphi|_{x_i y_i r_i, 1, y_i}$ and $|\varphi\rangle\langle\varphi|_{x_i y^* r_i, 1, y_i}$ are the same states as $|\varphi\rangle\langle\varphi|_{x_i y_i r_i}$ and $|\varphi\rangle\langle\varphi|_{x_i y^* r_i}$. By Fact 10 we can get,

$$\mathbb{E}_{i \in \bar{C}} \|P_{X_i G_i R_i|\mathcal{E}, 1} - P_{G_i R_i|\mathcal{E}, 1} P_{X_i|1, G_i}\|_1 \leq 2\sqrt{2\delta_1}.$$

Hence, for any value $Y_i = y_i$,

$$\begin{aligned} \mathbb{E}_{i \in \bar{C}} \|P_{X_i G_i R_i|\mathcal{E}, 1} - P_{G_i R_i|\mathcal{E}, 1} P_{X_i|\mathcal{E}, y_i, 1, G_i R_i}\|_1 &\leq \mathbb{E}_{i \in \bar{C}} \left[\|P_{X_i G_i R_i|\mathcal{E}, 1} - P_{G_i R_i|\mathcal{E}, 1} P_{X_i|y_i, 1, G_i}\|_1 \right. \\ &\quad \left. + \|P_{G_i R_i|\mathcal{E}, 1} (P_{X_i|y_i, 1, G_i} - P_{X_i|\mathcal{E}, y_i, 1, G_i R_i})\|_1 \right] \\ &\leq \mathbb{E}_{i \in \bar{C}} \left[\|P_{X_i G_i R_i|\mathcal{E}, 1} - P_{G_i R_i|\mathcal{E}, 1} P_{X_i|1, G_i}\|_1 \right. \\ &\quad \left. + \frac{2}{\zeta/3 - \sqrt{2\delta}} \|P_{X_i G_i R_i|\mathcal{E}, 1} - P_{G_i R_i|\mathcal{E}, 1} P_{X_i|1, G_i}\|_1 \right] \\ &\leq \frac{8\sqrt{2\delta_1}}{\zeta} \end{aligned}$$

where we have used the fact that for any value $G_i = g_i$, we must have $P_{Y_i|1, g_i}(y_i) \geq \zeta/3 - \sqrt{2\delta}$. Finally,

$$\mathbb{E}_{i \in \bar{C}} \|P_{X_i G_i R_i|\mathcal{E}, 1} - P_{X_i Y_i R_i|\mathcal{E}, 1}\|_1 \leq 2P_{Y_i G_i|\mathcal{E}, 1}(Y_i \neq G_i) \leq \zeta/3 + \sqrt{2\delta}.$$

Observing that $P_{X_i Y_i R_i|\mathcal{E}, 1}$ is the same as $P_{X_i Y_i R_i|\mathcal{E}}$ we get,

$$\mathbb{E}_{i \in \bar{C}} \|P_{X_i Y_i R_i|\mathcal{E}} - P_{G_i R_i|\mathcal{E}, 1} P_{X_i|\mathcal{E}, y_i, 1, G_i R_i}\|_1 \leq \frac{8\sqrt{2\delta_1}}{\zeta} + \frac{\zeta}{3} + \sqrt{2\delta}.$$

Using this and (17) we get,

$$\mathbb{E}_{i \in \bar{C}} \mathbb{E}_{P_{X_i Y_i R_i|\mathcal{E}}} \|\varphi\rangle\langle\varphi|_{x_i y_i r_i} - (\mathbb{1} \otimes U_{y_i r_i})|\varphi\rangle\langle\varphi|_{x_i y^* r_i} (\mathbb{1} \otimes U_{y_i r_i}^\dagger)\|_1$$

$$\begin{aligned}
&\leq \mathbb{E}_{i \in \bar{C}} \left[\left\| \mathbb{P}_{X_i Y_i R_i | \mathcal{E}} - \mathbb{P}_{G_i R_i | \mathcal{E}, 1} \mathbb{P}_{X_i | \mathcal{E}, y_i, 1, G_i R_i} \right\|_1 + \left\| \mathbb{P}_{X_i Y_i R_i | \mathcal{E}} - \mathbb{P}_{G_i R_i | \mathcal{E}, 1} \mathbb{P}_{X_i | \mathcal{E}, y^*, 1, G_i R_i} \right\|_1 \right. \\
&\quad \left. + \mathbb{E}_{\mathbb{P}_{G_i R_i | \mathcal{E}, 1}} \left\| \mathbb{E}_{\mathbb{P}_{X_i | \mathcal{E}, y_i, r_i, 1, y_i}} |x_i\rangle\langle x_i| \otimes |\varphi\rangle\langle\varphi|_{x_i y_i r_i} - \mathbb{E}_{\mathbb{P}_{X_i | \mathcal{E}, y^*, r_i, 1, y_i}} \mathbb{1} \otimes U_{y_i r_i} |\varphi\rangle\langle\varphi|_{x_i y^* r_i} \mathbb{1} \otimes U_{y_i r_i}^\dagger \right\|_1 \right] \\
&= \frac{16\sqrt{2\delta_1}}{\zeta} + \frac{2\zeta}{3} + 2\sqrt{2\delta} + \left(\frac{32\delta_1}{\zeta} \right)^{1/4} < \frac{7\zeta}{10} \tag{18}
\end{aligned}$$

where we have bounded the last term in the first inequality by applying Fact 14 on (17) with \mathcal{O}_{X_i} . Notice that we have also removed the conditioning $G_i \neq y^*$, since for $G_i = y^*$, the corresponding states are both $|\varphi\rangle_{x_i y^* r_i}$.

From (16) and (18) we get,

$$\begin{aligned}
&\mathbb{E}_{i \in \bar{C}} \mathbb{E}_{\mathbb{P}_{X_i Y_i R_i | \mathcal{E}}} \left\| \frac{1}{\alpha_{r_i}} (\mathbb{P}_{x_i r_i} \otimes U_{y_i r_i}) |\varphi'\rangle\langle\varphi'|_{y^* r_i} (\mathbb{P}_{x_i r_i} \otimes U_{y_i r_i}^\dagger) - |\varphi\rangle\langle\varphi|_{x_i y_i r_i} \right\|_1 \\
&\leq \mathbb{E}_{i \in \bar{C}} \mathbb{E}_{\mathbb{P}_{X_i Y_i R_i | \mathcal{E}}} \left[\left\| \frac{1}{\alpha_{r_i}} (\mathbb{P}_{x_i r_i} \otimes U_{y_i r_i}) |\varphi'\rangle\langle\varphi'|_{y^* r_i} (\mathbb{P}_{x_i r_i} \otimes U_{y_i r_i}^\dagger) - (\mathbb{1} \otimes U_{y_i r_i}) |\varphi\rangle\langle\varphi|_{x_i y^* r_i} (\mathbb{1} \otimes U_{y_i r_i}^\dagger) \right\|_1 \right. \\
&\quad \left. + \left\| (\mathbb{1} \otimes U_{y_i r_i}) |\varphi\rangle\langle\varphi|_{x_i y^* r_i} (\mathbb{1} \otimes U_{y_i r_i}^\dagger) - |\varphi\rangle\langle\varphi|_{x_i y_i r_i} \right\|_1 \right] \\
&= \mathbb{E}_{i \in \bar{C}} \mathbb{E}_{\mathbb{P}_{X_i Y_i R_i | \mathcal{E}}} \left[\left\| \frac{1}{\alpha_{r_i}} (\mathbb{P}_{x_i r_i} \otimes \mathbb{1}) |\varphi'\rangle\langle\varphi'|_{y^* r_i} (\mathbb{P}_{x_i r_i} \otimes \mathbb{1}) - |\varphi\rangle\langle\varphi|_{x_i y^* r_i} \right\|_1 \right. \\
&\quad \left. + \left\| (\mathbb{1} \otimes U_{y_i r_i}) |\varphi\rangle\langle\varphi|_{x_i y^* r_i} (\mathbb{1} \otimes U_{y_i r_i}^\dagger) - |\varphi\rangle\langle\varphi|_{x_i y_i r_i} \right\|_1 \right] \\
&\leq \frac{7\zeta}{2} + \frac{7\zeta}{10} = \frac{21\zeta}{5}. \tag{19}
\end{aligned}$$

Using Markov's inequality on (11), (13), (15) and (19), we get an index $i \in \bar{C}$ such that the conditions (i)-(iv) for Lemma 28 hold. \square

4 Proof of parallel repetition theorem

In this section we prove Theorem 2, whose statement is recalled below.

Theorem 2. *For a two-player non-local game $G = (q, \mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B}, \mathbb{V})$ such that q is a distribution anchored on one side with anchoring probability ζ ,*

$$\omega^*(G^k) = (1 - (1 - \omega^*(G))^5)^{\Omega\left(\frac{\zeta^{2k}}{\log(|\mathcal{A}| \cdot |\mathcal{B}|)}\right)}.$$

The proof of this theorem is very similar to that of the direct product theorem, so we shall only highlight points of difference. Whereas in the communication case, we started with an arbitrary distribution p and defined distribution q anchored on one side close to p , here we start with an already anchored distribution. To preserve similarity with the direct product proof, we shall

consider q to be anchored on the \mathcal{Y} side here as well, but the proof goes through analogously for a distribution anchored on the \mathcal{X} side. We define the correlation-breaking variables and the joint distribution \mathbb{P}_{XYDG} exactly as before.³

We consider an entangled strategy \mathcal{S} for G^k , where Alice and Bob, with input registers $X = X_1 \dots X_k$ and $Y = Y_1 \dots Y_k$, initially share an entangled state, and perform unitaries V^{Alice} and V^{Bob} respectively on their parts of the entangled state and their input registers. As before, conditioned on any value $DG = dg$, we define the following pure state representing \mathcal{S} after these unitaries:

$$|\theta\rangle_{X\tilde{X}Y\tilde{Y}ABE'_AE'_B|dg} = \sum_{xy} \sqrt{\mathbb{P}_{XY|dg}(xy)} |xxyy\rangle_{X\tilde{X}Y\tilde{Y}} \otimes |\theta\rangle_{ABE'_AE'_B|xy}$$

where AB are the answer registers which are measured in the computational basis by Alice and Bob to obtain their answers (a, b) , and $E'_AE'_B$ are some additional registers which are discarded. We shall use $\mathbb{P}_{XYAB|dg}$ to denote the distribution of $XYAB$ in $|\theta\rangle_{dg}$; \mathbb{P}_{XYDGAB} is obtained by averaging over dg .

Let the winning probability of $\omega^*(G)$ be $1 - 5\varepsilon$ for an appropriate ε . We shall prove the following lemma, which is analogous to the direct product case. It is clear that the lemma implies

$$\omega^*(G^k) \leq (1 - \varepsilon)^{\frac{\zeta^2 \varepsilon^4 k}{\log(|\mathcal{A}| \cdot |\mathcal{B}|)}} = (1 - (1 - \omega^*(G))^5)^{\Omega\left(\frac{\zeta^2 k}{\log(|\mathcal{A}| \cdot |\mathcal{B}|)}\right)}.$$

Lemma 29. *Let $\delta = \frac{\zeta^2 \varepsilon^4}{1440000}$ and $\delta' = \frac{\zeta^2 \varepsilon^4}{1440000 \log(|\mathcal{A}| \cdot |\mathcal{B}|)}$. For $i \in [k]$, let T_i denote the random variable $\mathbb{V}(X_i, Y_i, A_i, B_i)$, where $X_i Y_i A_i B_i$ are according to \mathbb{P}_{XYAB} . Then there exist $[\delta'k]$ coordinates $\{i_1, \dots, i_{[\delta'k]}\} \subseteq [k]$, such that for all $1 \leq r \leq [\delta'k] - 1$, at least one of the conditions holds*

- (i) $\Pr \left[\prod_{j=1}^r T_{i_j} = 1 \right] \leq (1 - \varepsilon)^{\delta k}$
- (ii) $\Pr \left[T_{i_{r+1}} = 1 \mid \prod_{j=1}^r T_{i_j} = 1 \right] \leq 1 - \varepsilon.$

As before, we shall consider that we have identified a set of coordinates $C = \{i_1, \dots, i_t\}$ such that for all $1 \leq r \leq t-1$, $\Pr \left[T_{i_{r+1}} = 1 \mid \prod_{j=1}^r T_{i_j} = 1 \right] \leq 1 - \varepsilon$ and $\Pr[\mathcal{E}] = \Pr \left[\prod_{j=1}^t T_{i_j} = 1 \right] \geq (1 - \varepsilon)^{\delta k}$, and identify a $(t+1)$ -th coordinate i . Let E_A and E_B to denote $A_{\bar{C}} E'_A$ and $B_{\bar{C}} E'_B$ respectively. We define the following state, which is $|\theta\rangle_{dg}$ conditioned on success in C :

$$\begin{aligned} & |\varphi\rangle_{X\tilde{X}Y\tilde{Y}A_C B_C B E_A E_B | dg} \\ &= \frac{1}{\sqrt{\gamma_{dg}}} \sum_{xy} \sqrt{\mathbb{P}_{XY|dg}(xy)} |xxyy\rangle_{X\tilde{X}Y\tilde{Y}} \otimes \sum_{a_C b_C: \mathbb{V}^t(x_C, y_C, a_C, b_C)=1} |a_C b_C\rangle_{A_C B_C} |\tilde{\varphi}\rangle_{E_A E_B | xy a_C b_C}. \end{aligned}$$

Here $|\tilde{\varphi}\rangle_{E_A E_B | xy a_C b_C}$ is a subnormalized state satisfying $\| |\tilde{\varphi}\rangle_{E_A E_B | xy a_C b_C} \|_2^2 = \mathbb{P}_{A_C B_C | xy}(a_C b_C)$.

The following lemma is the analog of Lemma 28, which we shall use to prove Lemma 29.

³The definition of $\mathbb{P}_{X_i Y_i D_i G_i}$ in the previous section makes references to $p(x, y)$. Since there is no p in the present case, $p(x, y)$ can simply be replaced by $q(x, y | y \neq y^*)$.

Lemma 30. *If $\Pr[\mathcal{E}] \geq (1 - \varepsilon)^{\delta k}$, then there exist a coordinate $i \in \bar{C}$, a random variable $R_i = X_C Y_C A_C B_C D_{-i} G_{-i}$, such that the following conditions hold:*

$$(i) \quad \|\mathbb{P}_{X_i Y_i R_i | \mathcal{E}} - \mathbb{P}_{X_i Y_i} \mathbb{P}_{R_i | \mathcal{E}, X_i}\|_1 \leq \frac{7\varepsilon}{150}$$

$$(ii) \quad \|\mathbb{P}_{X_i Y_i R_i | \mathcal{E}} - \mathbb{P}_{X_i Y_i} \mathbb{P}_{R_i | \mathcal{E}, Y_i}\|_1 \leq \frac{7\varepsilon}{150}$$

(iii) *There exist unitaries $\{U_{x_i r_i}\}_{x_i r_i}$ and $\{U_{y_i r_i}\}_{y_i r_i}$ respectively acting only on $X_{\bar{C}} \tilde{X}_{\bar{C}} E_A$ and $Y_{\bar{C}} \tilde{Y}_{\bar{C}} E_B$, such that*

$$\mathbb{E}_{\mathbb{P}_{X_i Y_i R_i | \mathcal{E}}} \left\| (U_{x_i r_i} \otimes U_{y_i r_i}) |\varphi\rangle \langle \varphi|_{y^* r_i} (U_{x_i r_i}^\dagger \otimes U_{y_i r_i}^\dagger) - |\varphi\rangle \langle \varphi|_{x_i y_i r_i} \right\|_1 \leq \frac{36\varepsilon}{5}.$$

It is easy to see how this lemma implies Lemma 29. As in the direct product case, Alice and Bob share $|\varphi\rangle_{y^* r_i}$ as entanglement – though in this case only one copy, as well as classical randomness with which they can produce $R_i^{\text{Alice}} R_i^{\text{Bob}}$ satisfying

$$\|\mathbb{P}_{X_i Y_i R_i^{\text{Alice}} R_i^{\text{Bob}}} - \mathbb{P}_{X_i Y_i R_i R_i | \mathcal{E}}\|_1 \leq \frac{7\varepsilon}{30}.$$

Alice and Bob apply $U_{x_i r_i^A}$ and $U_{y_i r_i^B}$ according to their inputs and R_i^{Alice} and R_i^{Bob} respectively, on their registers registers E_A and E_B of $|\varphi\rangle_{y^* r_i}$. They then measure in the computational basis on the $A_i B_i$ registers of resulting state, to give their outcomes (a_i, b_i) . $\Pr[T_i = 1 | \mathcal{E}] \geq 1 - \varepsilon$ implies that the resulting strategy for G has success probability $> (1 - 5\varepsilon)$, a contradiction which lets us identify i as the $(t + 1)$ -th coordinate.

The rest of the proof will be dedicated to showing Lemma 30.

Proof of Lemma 30. We can prove

$$\mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i Y_i R_i | \mathcal{E}} - \mathbb{P}_{X_i Y_i} \mathbb{P}_{R_i | \mathcal{E}, X_i}\|_1 \leq \frac{7\varepsilon}{600} \quad (20)$$

$$\mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i Y_i R_i | \mathcal{E}} - \mathbb{P}_{X_i Y_i} \mathbb{P}_{R_i | \mathcal{E}, Y_i}\|_1 \leq \frac{7\varepsilon}{600} \quad (21)$$

$$\mathbb{E}_{i \in \bar{C}} \mathbb{E}_{\mathbb{P}_{X_i Y_i R_i | \mathcal{E}}} \left\| |\varphi\rangle \langle \varphi|_{x_i y_i r_i} - (\mathbb{1} \otimes U_{y_i r_i}) |\varphi\rangle \langle \varphi|_{x_i y^* r_i} (\mathbb{1} \otimes U_{y_i r_i}^\dagger) \right\|_1 \leq \frac{4\varepsilon}{5} \quad (22)$$

exactly the same way as in the direct product case, except conditioning on z_C is replaced by conditioning on $a_C b_C$, which leads to the factor of $\log(|\mathcal{A}| \cdot |\mathcal{B}|)$. The rest of the proof will hence be spent getting Alice's unitaries $U_{x_i r_i}$.

Letting $\delta_1 = \delta + \delta' \log(|\mathcal{A}| \cdot |\mathcal{B}|)$, the following is derived analogously to the direct product case, except for the extra factor in the mutual information bound due to communication:

$$\mathbb{E}_{i \in \bar{C}} \mathbb{E}_{R_i | \mathcal{E}, D_i=1, G_i=y^*} \mathbb{I}(X_i : Y_{\bar{C}} \tilde{Y}_{\bar{C}} E_B)_{\varphi_{r_i, D_i=1, G_i=y^*}} \leq \frac{10\delta_1}{\zeta} \quad (23)$$

$$\mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i R_i | \mathcal{E}, y^*} - \mathbb{P}_{X_i R_i | \mathcal{E}, 1, y^*}\|_1 \leq \frac{7\sqrt{2\delta_1}}{\zeta} \quad (24)$$

$$\mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i R_i | \mathcal{E}} - \mathbb{P}_{X_i R_i | \mathcal{E}, 1, y^*}\|_1 \leq \frac{40\sqrt{2\delta_1}}{\zeta}. \quad (25)$$

From (23), by applying Pinsker's inequality, we get,

$$\mathbb{E}_{i \in \bar{C}} \mathbb{E}_{\mathbb{P}_{X_i R_i | \mathcal{E}, 1, y^*}} \|\varphi_{Y_{\bar{C}} \tilde{Y}_{\bar{C}} E_B | x_i r_i, 1, y^*} - \varphi_{Y_{\bar{C}} \tilde{Y}_{\bar{C}} E_B | r_i, 1, y^*}\|_1 \leq \left(\frac{10\delta_1}{\zeta}\right)^{1/2}$$

Note that $\varphi_{Y_{\bar{C}} \tilde{Y}_{\bar{C}} E_B | x_i r_i, 1, y^*}$ is the same state as $\varphi_{Y_{\bar{C}} \tilde{Y}_{\bar{C}} E_B | x_i y^* r_i}$. But $\varphi_{Y_{\bar{C}} \tilde{Y}_{\bar{C}} E_B | r_i, 1, y^*}$ is not the same state as $\varphi_{Y_{\bar{C}} \tilde{Y}_{\bar{C}} E_B | y^* r_i}$, due to the averaging over X_i being done with respect to $\mathbb{P}_{X_i | \mathcal{E}, r_i, 1, y^*}$ in one, and with respect to $\mathbb{P}_{X_i | \mathcal{E}, y^* r_i}$ in the other. However, due to (24) we can say,

$$\begin{aligned} & \mathbb{E}_{i \in \bar{C}} \mathbb{E}_{\mathbb{P}_{X_i R_i | \mathcal{E}, 1, y^*}} \|\varphi_{Y_{\bar{C}} \tilde{Y}_{\bar{C}} E_B | x_i y^* r_i} - \varphi_{Y_{\bar{C}} \tilde{Y}_{\bar{C}} E_B | y^* r_i}\|_1 \\ & \leq \left(\frac{10\delta_1}{\zeta}\right)^{1/2} + \mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i R_i | \mathcal{E}, 1, y^*} - \mathbb{P}_{R_i | \mathcal{E}, 1, y^*} \mathbb{P}_{X_i | \mathcal{E}, R_i, y^*}\|_1 \\ & \leq \left(\frac{10\delta_1}{\zeta}\right)^{1/2} + \mathbb{E}_{i \in \bar{C}} \|\mathbb{P}_{X_i R_i | \mathcal{E}, y^*} - \mathbb{P}_{X_i R_i | \mathcal{E}, 1, y^*}\|_1 \\ & \leq \frac{2\sqrt{108\delta_1}}{\zeta}. \end{aligned}$$

Since $|\varphi\rangle_{X_{\bar{C}} \tilde{X}_{\bar{C}} Y_{\bar{C}} \tilde{Y}_{\bar{C}} E_A E_B | y^* r_i}$ is a purification of $\varphi_{Y_{\bar{C}} \tilde{Y}_{\bar{C}} E_B | y^* r_i}$ and $|\varphi\rangle_{X_{\bar{C}} \tilde{X}_{\bar{C}} Y_{\bar{C}} \tilde{Y}_{\bar{C}} E_A E_B | x_i y^* r_i}$ is a purification of $\varphi_{Y_{\bar{C}} \tilde{Y}_{\bar{C}} E_B | x_i y^* r_i}$, by the Fuchs-van de Graaf inequality and Uhlmann's theorem we can say that there exist unitaries $U_{x_i r_i}$ on $X_{\bar{C}} \tilde{X}_{\bar{C}} E_A$ such that

$$\mathbb{E}_{i \in \bar{C}} \mathbb{E}_{\mathbb{P}_{X_i R_i | \mathcal{E}, 1, y^*}} \|\langle \varphi | \langle \varphi |_{x_i y^* r_i} - (U_{x_i r_i} \otimes \mathbb{1}) | \varphi \rangle \langle \varphi |_{y^* r_i} (U_{x_i r_i}^\dagger \otimes \mathbb{1}) \|_1 \leq \left(\frac{2\sqrt{108\delta_1}}{\zeta}\right)^{1/2}$$

and by (25) again,

$$\begin{aligned} \mathbb{E}_{i \in \bar{C}} \mathbb{E}_{\mathbb{P}_{X_i R_i | \mathcal{E}}} \|\langle \varphi | \langle \varphi |_{x_i y^* r_i} - (U_{x_i r_i} \otimes \mathbb{1}) | \varphi \rangle \langle \varphi |_{y^* r_i} (U_{x_i r_i}^\dagger \otimes \mathbb{1}) \|_1 & \leq \left(\frac{2\sqrt{108\delta_1}}{\zeta}\right)^{1/2} + \frac{40\sqrt{2\delta_1}}{\zeta} \\ & \leq 2 \left(\frac{10800\delta_1}{\zeta^2}\right)^{1/4} \\ & \leq \varepsilon. \end{aligned} \quad (26)$$

Combining (26) and (22) we get,

$$\mathbb{E}_{i \in \bar{C}} \mathbb{E}_{\mathbb{P}_{X_i Y_i R_i | \mathcal{E}}} \left\| (U_{x_i r_i} \otimes U_{y_i r_i}) | \varphi \rangle \langle \varphi |_{y^* r_i} (U_{x_i r_i}^\dagger \otimes U_{y_i r_i}^\dagger) - | \varphi \rangle \langle \varphi |_{x_i y_i r_i} \right\|_1 \leq \frac{9\varepsilon}{5}.$$

The result then follows by Markov's inequality. \square

Acknowledgements

This work is supported by the National Research Foundation, including under NRF RF Award No. NRF-NRFF2013-13, the Prime Minister's Office, Singapore and the Ministry of Education, Singapore, under the Research Centres of Excellence program and by Grant No. MOE2012-T3-1-009 and in part by the NRF2017-NRF-ANR004 *VanQuTe* Grant.

References

- [BARdW08] Avraham Ben-Aroya, Oded Regev, and Ronald de Wolf. A Hypercontractive Inequality for Matrix-Valued Functions with Applications to Quantum Computing and LDCs. In *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science, FOCS '08*, pages 477–486, 2008.
- [BBCR13] Boaz Barak, Mark Braverman, Xi Chen, and Anup Rao. How to Compress Interactive Communication. *SIAM Journal on Computing*, 42(3):1327–1363, 2013.
- [BCR11] Mario Berta, Matthias Christandl, and Renato Renner. The Quantum Reverse Shannon Theorem Based on One-Shot Information Theory. *Communications in Mathematical Physics*, 306(3):579–615, 2011.
- [BK18] Mark Braverman and Gillat Kol. Interactive Compression to External Information. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC '18*, page 964977, 2018.
- [BR14] Mark Braverman and Anup Rao. Information Equals Amortized Communication. *IEEE Transactions on Information Theory*, 60(10):6058–6069, 2014.
- [Bra15] Mark Braverman. Interactive information complexity. *SIAM Journal on Computing*, 44(6):1698–1739, 2015.
- [BRWY13a] Mark Braverman, Anup Rao, Omri Weinstein, and Amir Yehudayoff. Direct Product via Round-Preserving Compression. In *Automata, Languages, and Programming*, volume 7965 of *Lecture Notes in Computer Science*, pages 232–243. 2013.
- [BRWY13b] Mark Braverman, Anup Rao, Omri Weinstein, and Amir Yehudayoff. Direct Products in Communication Complexity. In *Proceedings of the 54th Annual IEEE Symposium on Foundations of Computer Science, FOCS '13*, pages 746–755, 2013.
- [BVY15] Mohammad Bavarian, Thomas Vidick, and Henry Yuen. Anchoring Games for Parallel Repetition. <https://arxiv.org/abs/1509.07466>, 2015.
- [BVY17] Mohammad Bavarian, Thomas Vidick, and Henry Yuen. Hardness Amplification for Entangled Games via Anchoring. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC '17*, page 303316, 2017.
- [BW15] Mark Braverman and Omri Weinstein. An Interactive Information Odometer and Applications. In *Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing, STOC '15*, page 341350, 2015.
- [BYJKS02] Ziv Bar-Yossef, T. S. Jayram, Ravi Kumar, and D. Sivakumar. An Information Statistics Approach to Data Stream and Communication Complexity. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science, FOCS '02*, pages 209–218, 2002.
- [CSUU08] Richard Cleve, William Slofstra, Falk Unger, and Sarvagya Upadhyay. Perfect Parallel Repetition Theorem for Quantum XOR Proof Systems. *Computational Complexity*, 17(2):282–299, 2008.

- [CSWY01] Amit Chakrabarti, Yaoyun Shi, Anthony Wirth, and Andrew Yao. Informational Complexity and the Direct Sum Problem for Simultaneous Message Complexity. In *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science, FOCS '01*, pages 270–278, 2001.
- [DSV15] Irit Dinur, David Steurer, and Thomas Vidick. A Parallel Repetition Theorem for Entangled Projection Games. *Computational Complexity*, 24(2):201254, 2015.
- [HJMR10] Prahladh Harsha, Rahul Jain, David McAllester, and Jaikumar Radhakrishnan. The Communication Complexity of Correlation. *IEEE Transactions on Information Theory*, 56(1):438–449, 2010.
- [Hol07] Thomas Holenstein. Parallel Repetition: Simplifications and the No-Signaling Case. In *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing, STOC '07*, page 411419, 2007.
- [Jai15] Rahul Jain. New Strong Direct Product Results in Communication Complexity. *Journal of the ACM*, 62(3), 2015.
- [JK09] Rahul Jain and Hartmut Klauck. New Results in the Simultaneous Message Passing Model via Information Theoretic Techniques. In *Proceedings of the 24th Annual IEEE Conference on Computational Complexity, CCC '09*, pages 369–378, 2009.
- [JKN08] Rahul Jain, Hartmut Klauck, and Ashwin Nayak. Direct Product Theorems for Classical Communication Complexity via Subdistribution Bounds: Extended Abstract. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing, STOC '08*, pages 599–608, 2008.
- [JN12] Rahul Jain and Ashwin Nayak. Short Proofs of the Quantum Substate Theorem. *IEEE Transactions on Information Theory*, 58(6):3664–3669, 2012.
- [JPY14] Rahul Jain, Attila Pereszlényi, and Penghui Yao. A Parallel Repetition Theorem for Entangled Two-Player One-Round Games under Product Distributions. In *2014 IEEE 29th Conference on Computational Complexity (CCC '14)*, pages 209–216, 2014.
- [JPY16] Rahul Jain, Attila Pereszlényi, and Penghui Yao. A Direct Product Theorem for Two-Party Bounded-Round Public-Coin Communication Complexity. *Algorithmica*, 76(3):720748, 2016.
- [JRS02] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. The Quantum Communication Complexity of the Pointer Chasing Problem: The Bit Version. In *FSTTCS 2002: Foundations of Software Technology and Theoretical Computer Science*, volume 2556 of *Lecture Notes in Computer Science*, pages 218–229, 2002.
- [JRS03a] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. A Direct Sum Theorem in Communication Complexity via Message Compression. In *Automata, Languages and Programming*, volume 2719 of *Lecture Notes in Computer Science*, pages 300–315. 2003.

- [JRS03b] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. A Lower Bound for the Bounded Round Quantum Communication Complexity of Set Disjointness. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science, FOCS '03*, pages 220–229, 2003.
- [JRS05] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. Prior Entanglement, Message Compression and Privacy in Quantum Communication. In *20th Annual IEEE Conference on Computational Complexity (CCC '05)*, pages 285–296, 2005.
- [JRS09] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. A Property of Quantum Relative Entropy with an Application to Privacy in Quantum Communication. *Journal of the ACM*, 56(6), 2009.
- [JSR08] Rahul Jain, Pranab Sen, and Jaikumar Radhakrishnan. Optimal Direct Sum and Privacy Trade-off Results for Quantum and Classical Communication Complexity. <http://arxiv.org/abs/0807.1267>, 2008.
- [JY12] Rahul Jain and Penghui Yao. A Strong Direct Product Theorem in Terms of the Smooth Rectangle Bound. <http://arxiv.org/abs/1209.0263>, 2012.
- [Kla10] Hartmut Klauck. A Strong Direct Product Theorem for Disjointness. In *Proceedings of the 42nd ACM Symposium on Theory of Computing, STOC '10*, pages 77–86, 2010.
- [Kol16] Gillat Kol. Interactive Compression for Product Distributions. In *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing, STOC '16*, page 987998, 2016.
- [KRT10] Julia Kempe, Oded Regev, and Ben Toner. Unique Games with Entangled Provers are Easy. *SIAM Journal on Computing*, 39(7):3207–3229, 2010.
- [KvdW07] Hartmut Klauck, Robert Špalek, and Ronald de Wolf. Quantum and Classical Strong Direct Product Theorems and Optimal Time-Space Tradeoffs. *SIAM Journal on Computing*, 36(5):1472–1493, 2007.
- [LSv08] Troy Lee, Adi Shraibman, and Robert Špalek. A Direct Product Theorem for Discrepancy. In *Proceedings of the 23rd Annual IEEE Conference on Computational Complexity, CCC '08*, pages 71–80, 2008.
- [Raz92] Alexander A. Razborov. On the Distributional Complexity of Disjointness. *Theoretical Computer Science*, 106(2):385–390, 1992.
- [Raz95] Ran Raz. A Parallel Repetition Theorem. In *Proceedings of the Twenty-Seventh Annual ACM Symposium on Theory of Computing*, page 447456, 1995.
- [Sha03] Ronen Shaltiel. Towards Proving Strong direct Product Theorems. *Computational Complexity*, 12(1-2):1–22, 2003.
- [She12] Alexander A. Sherstov. Strong Direct Product Theorems for Quantum Communication and Query Complexity. *SIAM Journal on Computing*, 41(5):1122–1165, 2012.

- [She18] Alexander A. Sherstov. Compressing Interactive Communication Under Product Distributions. *SIAM Journal on Computing*, 47(2):367–419, 2018.
- [VW08] Emanuele Viola and Avi Wigderson. Norms, XOR Lemmas, and Lower Bounds for Polynomials and Protocols. *Theory of Computing*, 4(7):137–168, 2008.
- [Yao79] Andrew C.-C. Yao. Some complexity questions related to distributive computing (preliminary report). In *Proceedings of the 11th Annual ACM Symposium on Theory of Computing*, STOC '79, pages 209–213, 1979.
- [Yue16] Henry Yuen. A Parallel Repetition Theorem for All Entangled Games. In *43rd International Colloquium on Automata, Languages, and Programming (ICALP '16)*, volume 55 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 77:1–77:13, 2016.