A parallel repetition theorem for entangled two-player one-round games under product distributions

Rahul Jain
Centre for Quantum Technologies and
Department of Computer Science
National University of Singapore
Singapore
rahul@comp.nus.edu.sg

Attila Pereszlényi
Centre for Quantum Technologies
National University of Singapore
Singapore
attila.pereszlenyi@gmail.com

Penghui Yao
Centre for Quantum Technologies
National University of Singapore
Singapore
phyao1985@gmail.com

***Abstract*—We show a *parallel repetition* theorem for the *entangled* value $\omega^*(G)$ of any *two-player one-round game* G where the questions $(x,y) \in \mathcal{X} \times \mathcal{Y}$ to Alice and Bob are drawn from a product distribution on $\mathcal{X} \times \mathcal{Y}$. We show that for the k-fold product G^k of the game G (which represents the game G played in parallel k times independently)**

$$\omega^*\left(G^k\right) = \left(1 - (1-\omega^*(G))^3\right)^{\Omega\left(\frac{k}{\log(|\mathcal{A}|\cdot|\mathcal{B}|)}\right)}$$

where $\mathcal{A}$ and $\mathcal{B}$ represent the sets from which the answers of Alice and Bob are drawn.

The arguments we use are information theoretic and are broadly on similar lines as that of Raz [1] and Holenstein [2] for classical games. The additional quantum ingredients we need, to deal with entangled games, are inspired by the work of Jain, Radhakrishnan, and Sen [3], where quantum information theoretic arguments were used to achieve message compression in quantum communication protocols.

***Index Terms*—parallel repetition theorem; two-player game; entangled value**

I. Introduction

A *two-player one-round game* G is specified by finite sets $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{A}$, and $\mathcal{B}$, a distribution μ over $\mathcal{X} \times \mathcal{Y}$, and a predicate $V : \mathcal{X} \times \mathcal{Y} \times \mathcal{A} \times \mathcal{B} \to \{0,1\}$. It is played as follows.

- The referee selects questions $(x,y) \in \mathcal{X} \times \mathcal{Y}$ according to distribution μ.
- The referee sends x to Alice and y to Bob. Alice and Bob are spatially separated, so they do not see each other's input.
- Alice chooses answer $a \in \mathcal{A}$ and sends it back to the referee. Bob chooses answer $b \in \mathcal{B}$ and sends it back to the referee.
- The referee accepts if $V(x,y,a,b) = 1$ and otherwise rejects. Alice and Bob win the game if the referee accepts.

The *value* of the game G, denoted by $\omega(G)$, is defined to be the maximum winning probability (averaged over the distribution μ) achieved by Alice and Bob.

These games have played an important and pivotal role in the study of the rich theory of *inapproximability*, leading to the development of *Probabilistically Checkable Proofs* and the famous *Unique Games Conjecture*. One of the most fundamental problems regarding this model is the so called *parallel repetition* question, which concerns the behavior of multiple copies of the game played in parallel. For the game $G = (\mu, \mathcal{X}, \mathcal{Y}, \mathcal{A}, \mathcal{B}, V)$, its k-fold product is given by $G^k = (\mu^k, \mathcal{X}^k, \mathcal{Y}^k, \mathcal{A}^k, \mathcal{B}^k, V^k)$, where $V^k(x,y,a,b) = 1$ if and only if $V(x_i, y_i, a_i, b_i) = 1$ for all $i \in [k]$. Namely, Alice and Bob play k copies of game G simultaneously, and they win iff they win all the copies. It is easily seen that $\omega(G^k) \geq \omega(G)^k$ for any game G. The equality of the two quantities, for all games, was conjectured by Ben-Or, Goldwasser, Kilian and Wigderson [4]. The conjecture was shown to be false by Fortnow [5].

However one could still expect that $\omega(G^k)$ goes down exponentially in k (asymptotically). This is referred to as the parallel repetition (also known as the *direct product*) conjecture. This was shown to be indeed true in a seminal paper by Raz [1]. Raz showed that

$$\omega\left(G^k\right) = \left(1 - (1-\omega(G))^c\right)^{\Omega\left(\frac{k}{\log(|\mathcal{A}||\mathcal{B}|)}\right)}$$

where c is a universal constant. This result, along with the the *PCP theorem* had deep consequences for the theory of inapproximability [6], [7], [8]. A series of works later exhibited improved results for general and specific games [2], [9], [10], [11], [12].

In the quantum setting, it is natural to consider the so called *entangled games* where Alice and Bob are, in addition, allowed to share a quantum state before the games starts. The questions and answers in the game remain classical. On receiving questions, Alice and Bob can generate their answers by making quantum measurements on their shared entangled state. The value of the entangled version of the game G is denoted by $\omega^*(G)$. The study of entangled games is deeply related to the foundation of quantum mechanics and that of quantum entanglement. These games have been used to give a novel interpretation to *Bell inequalities*, one of the most famous and useful methods for differentiating classical and quantum mechanics (e.g., by Clauser, Horne, Shimony and Holt [13]). Recently these games have also been studied from cryptographic motivations such as in Refs. [14], [15], [16]. Analogously to the classical case, the study of the parallel repetition question in this setting may potentially have applications in quantum complexity theory.

The parallel repetition conjecture has been shown to be true for several sub-classes of entangled games, starting with the so called *XOR games* by Cleve, Slofstra, Unger and Upadhyay [17], later generalized to *unique games* by Kempe, Regev and Toner [18] and very recently further generalized to *projection games* by Dinur, Steurer and Vidick [19] (following an analytical framework introduced by Dinur and Steurer in [20] to deal with classical projection games). For general games, Kempe and Vidick [21] (following a framework by Feige and Killian [22] for classical games) showed a parallel repetition theorem albeit with only a polynomial decay in k, in the value $\omega^*(G^k)$. In a recent work, Chailloux and Scarpa [23] showed an exponential decay in $\omega^*(G^k)$ using information theoretic arguments.

Theorem I.1 ([23])**.** *For any game $G = (\mu, \mathcal{X}, \mathcal{Y}, \mathcal{A}, \mathcal{B}, V)$, where μ is the uniform distribution on $\mathcal{X} \times \mathcal{Y}$, it holds that*

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

As a corollary, for a general distribution μ,

$$\omega^*\left(G^k\right) = \left(1-(1-\omega^*(G))^2\right)^{\Omega\left(\frac{k}{Q^4\log(Q)\log(|\mathcal{A}||\mathcal{B}|)}\right)}$$

where

$$Q = \max\left\{\left\lceil \frac{1}{\min_{x,y:\mu(x,y)\neq 0}\left\{\sqrt{\mu(x,y)}\right\}}\right\rceil, |\mathcal{X}|\cdot|\mathcal{Y}|\right\}.$$

Note that here $\omega^*(G^k)$ depends on $|\mathcal{X}|\cdot|\mathcal{Y}|$ as well, in addition to $|\mathcal{A}|\cdot|\mathcal{B}|$ (as in Raz's result). Also the value of Q can be arbitrarily large, depending on the distribution μ.

Our result

In this paper we consider the case when the distribution μ is product across $\mathcal{X} \times \mathcal{Y}$. That is, there are distributions μ_X, μ_Y on $\mathcal{X}, \mathcal{Y}$ respectively such that $\forall (x, y) \in \mathcal{X} \times \mathcal{Y} : \mu(x, y) = \mu_X(x) \cdot \mu_Y(y)$. We show the following.

Theorem I.2 (Main Result)**.** *For any game*

$$G = (\mu, \mathcal{X}, \mathcal{Y}, \mathcal{A}, \mathcal{B}, V)$$

where μ is a product distribution on $\mathcal{X} \times \mathcal{Y}$, it holds that

$$\omega^*\left(G^k\right) = \left(1-(1-\omega^*(G))^3\right)^{\Omega\left(\frac{k}{\log(|\mathcal{A}||\mathcal{B}|)}\right)}.$$

Note that the uniform distribution on $\mathcal{X} \times \mathcal{Y}$ is a product distribution and our result has no dependence on the size of $\mathcal{X} \times \mathcal{Y}$. Hence, our result implies and strengthens on the result of Chailloux and Scarpa [23] (up to the exponent of $1-\omega^*(G)$).

Our techniques

The arguments we use are information theoretic and are broadly on similar lines as that of Raz [1] and Holenstein [2] for classical games. The additional quantum ingredients we need, to deal with entangled games, are inspired by the work of Jain, Radhakrishnan, and Sen [3], where quantum information theoretic arguments were used to achieve message compression in quantum communication protocols.

Given the k-fold game G^k, let us condition on success on a set $\mathcal{C} \subseteq [k]$ of coordinates. If the overall success in coordinates in $\mathcal{C}$ is already as small as we want, then we are done. Otherwise, we exhibit another coordinate $j \notin \mathcal{C}$ such that the success in the j-th coordinate, even when conditioning on success in the coordinates inside $\mathcal{C}$, is bounded away from 1. Here we assume that $\omega^*(G)$ is bounded away from 1. This way the overall success keeps going down and becomes exponentially small in k, after we have identified $\Omega(k)$ such coordinates. To argue that the probability with which Alice and Bob win the j-th coordinate, conditioned on success in $\mathcal{C}$, is bounded away from 1, we show that close to this success probability can be achieved for a single instance of the game G. That is, given inputs (x', y'), drawn from μ, for a single instance of G, Alice and Bob can embed (x', y') to the j-th coordinate of G^k, conditioned on success in $\mathcal{C}$, and generate the rest of the state with good approximation. So, if the probability of success in the j-th coordinate, conditioned on success in $\mathcal{C}$, is very close to 1, there is a strategy for G with probability of success strictly larger than $\omega^*(G)$, reaching a contradicting to the definition of $\omega^*(G)$.

Suppose the global state, conditioned on success in $\mathcal{C}$, is of the form

$$\sigma^{XYAB} = \sum_{x\in\mathcal{X}^k, y\in\mathcal{Y}^k} \tilde{\mu}(x,y)\, |xy\rangle\langle xy|^{XY} \otimes |\phi_{xy}\rangle\langle\phi_{xy}|^{AB}$$

where $\tilde{\mu}$ is a distribution, potentially different from μ because of the conditioning on success. (Here we further fix the questions and answers in $\mathcal{C}$ to specific values and do not specify them in σ^{XYAB}.) In protocol $\mathcal{P}$ for the single instance of G, we let Alice and Bob start with the shared pure state

$$|\varphi\rangle = \sum_{x\in\mathcal{X}^k, y\in\mathcal{Y}^k} \sqrt{\tilde{\mu}(x,y)}\, |xxyy\rangle^{\tilde{X}X\tilde{Y}Y} \otimes |\phi_{xy}\rangle^{AB}.$$

Note that $|\varphi\rangle$ is a purification of σ^{XYAB}, where registers $\tilde{X}$ and $\tilde{Y}$ are identical to X and Y. We introduce these copies of the registers X and Y so that the marginal state in these registers remains a *classical state* and these registers can be viewed as classical registers, which is important in our arguments.

Using the chain rule for mutual information, we are able to argue that both $\mathrm{I}\left(X_j : Y\tilde{Y}B\right)$ and $\mathrm{I}\left(Y_j : X\tilde{X}A\right)$ are very small (close to 0), in $|\varphi\rangle$. This, obviously, is only possible when the distribution μ is product. In addition, the distribution of the questions in the j-th coordinate, in $|\varphi\rangle$, remains close to μ, in the ℓ_1*-distance*. In protocol $\mathcal{P}$, when Alice and Bob get questions x' and y', suppose they measure registers X_j

and Y_j, in $|\varphi\rangle$, and get x'_j and y'_j. Let $\left|\varphi_{x'_j y'_j}\right\rangle$ be the resulting state. If by luck it so happens that $(x', y') = \left(x'_j, y'_j\right)$, then they can measure the answer registers A_j and B_j, in $\left|\varphi_{x'_j y'_j}\right\rangle$, respectively, and send the answers to the referee. However, the probability that $(x', y') = \left(x'_j, y'_j\right)$ can be very small and they want to get this desired outcome with probability very close to 1. We describe next how this can be achieved.

Let $\left|\varphi_{x'_j}\right\rangle$ be the resulting state obtained after we measure register X_j (in $|\varphi\rangle$) and obtain outcome x'_j. The fact that $\mathrm{I}\left(X_j : Y\tilde{Y}B\right)$ is close to 0 implies that Bob's side of $\left|\varphi_{x'_j}\right\rangle$ is mostly independent of x'_j. By the unitary equivalence of purifications and Uhlmann's theorem, there is a unitary transformation $\mathbf{U}_{x'_j}$ that Alice can apply to take the state $|\varphi\rangle$ quite close to the state $\left|\varphi_{x'_j}\right\rangle$. Similarly, let us define $\left|\varphi_{y'_j}\right\rangle$ and again $\mathrm{I}\left(Y_j : X\tilde{X}A\right)$ being close to 0 implies that Alice's side of $\left|\varphi_{y'_j}\right\rangle$ is mostly independent of y'_j. Again, by Uhlmann's theorem, there is a unitary transformation $\mathbf{U}_{y'_j}$ that Bob can apply to take the state $|\varphi\rangle$ quite close to the state $\left|\varphi_{y'_j}\right\rangle$. Interestingly, as was argued in [3], when Alice and Bob simultaneously apply $\mathbf{U}_{x'_j}$ and $\mathbf{U}_{y'_j}$, they take $|\varphi\rangle$ quite close to the state $\left|\varphi_{x'_j y'_j}\right\rangle$! This again requires the distribution of questions to be independent across Alice and Bob.

Organization of the paper

In Section II, we present some background on information theory, as well as some useful lemmas that we will need for our proof. In Section III, we prove our main result, Theorem I.2.

II. Preliminaries

In this section we present some notations, definitions, facts, and lemmas that we will use later in our proof.

Information theory

For integer $n \geq 1$, let $[n]$ represent the set $\{1, 2, \ldots, n\}$. Let $\mathcal{X}$ and $\mathcal{Y}$ be finite sets and k be a natural number. Let $\mathcal{X}^k$ be the set $\mathcal{X} \times \cdots \times \mathcal{X}$, the cross product of $\mathcal{X}$, k times. Let μ be a probability distribution on $\mathcal{X}$. Let $\mu(x)$ represent the probability of $x \in \mathcal{X}$ according to μ. Let X be a random variable distributed according to μ. We use the same symbol to represent a random variable and its distribution whenever it is clear from the context. The expectation value of function f on $\mathcal{X}$ is defined as $\mathbb{E}_{x \leftarrow X}[f(x)] \stackrel{\text{def}}{=} \sum_{x \in \mathcal{X}} \Pr[X = x] \cdot f(x)$, where $x \leftarrow X$ means that x is drawn from the distribution of X. A quantum state (or just a state) ρ is a positive semi-definite matrix with trace equal to 1. It is pure if and only if the rank is 1. Let $|\psi\rangle$ be a unit vector. With slight abuse of notation, we use ψ to represent the state and also the density matrix $|\psi\rangle\langle\psi|$, associated with $|\psi\rangle$. A classical distribution μ can be viewed as a quantum state with diagonal entries $\mu(x)$ and non-diagonal entries 0. For two quantum states ρ and σ, $\rho \otimes \sigma$ represents the tensor product (Kronecker product) of ρ and σ. A quantum super-operator $\mathcal{E}(\cdot)$ is a completely positive and trace preserving (CPTP) linear map from states to states. Readers can refer to [24], [25], [26] for more details.

Definition II.1. For quantum states ρ and σ, the ℓ_1-distance between them is given by $\|\rho - \sigma\|_1$, where $\|X\|_1 \stackrel{\text{def}}{=} \mathrm{Tr}\sqrt{X^\dagger X}$ is the sum of the singular values of X. We say that ρ is ε-close to σ if $\|\rho - \sigma\|_1 \leq \varepsilon$.

Definition II.2. For quantum states ρ and σ, the *fidelity* between them is given by $\mathrm{F}(\rho, \sigma) \stackrel{\text{def}}{=} \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1$.

The following proposition states that the distance between two states can't be increased by quantum operations.

Proposition II.3 ([25], pages 406 and 414). *For states ρ, σ, and quantum operation $\mathcal{E}(\cdot)$, it holds that*

$$\|\mathcal{E}(\rho) - \mathcal{E}(\sigma)\|_1 \leq \|\rho - \sigma\|_1$$

and

$$\mathrm{F}(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \geq \mathrm{F}(\rho, \sigma).$$

The following proposition relates the ℓ_1-distance and the fidelity between two states.

Proposition II.4 ([25], page 416). *For quantum states ρ and σ, it holds that*

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

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

$$\begin{aligned} \||\phi\rangle\langle\phi| - |\psi\rangle\langle\psi|\|_1 &= \sqrt{1 - \mathrm{F}\left(|\phi\rangle\langle\phi|, |\psi\rangle\langle\psi|\right)^2} \\ &= \sqrt{1 - |\langle\phi|\psi\rangle|^2}. \end{aligned}$$

Let ρ^{AB} be a bipartite quantum state in registers AB. We use the same symbol to represent a quantum register and the Hilbert space associated with it. We define

$$\rho^B \stackrel{\text{def}}{=} \mathrm{Tr}_A\left(\rho^{AB}\right) \stackrel{\text{def}}{=} \sum_i (\langle i| \otimes \mathbb{1}_B)\rho^{AB}(|i\rangle \otimes \mathbb{1}_B)$$

where $\{|i\rangle\}_i$ is a basis for the Hilbert space A and $\mathbb{1}_B$ is the identity matrix in space B. The state ρ^B is referred to as the marginal state of ρ^{AB} in register B.

Definition II.5. We say that a pure state $|\psi\rangle \in A \otimes B$ is a purification of some state ρ if $\mathrm{Tr}_A(|\psi\rangle\langle\psi|) = \rho$.

Theorem II.6 (Uhlmann's theorem). *Given quantum states ρ, σ, and a purification $|\psi\rangle$ of ρ, it holds that $\mathrm{F}(\rho, \sigma) = \max_{|\phi\rangle} |\langle\phi|\psi\rangle|$, where the maximum is taken over all purifications of σ.*

The *entropy* of a quantum state ρ (in register X) is defined as $\mathrm{S}(\rho) \stackrel{\text{def}}{=} -\mathrm{Tr}\rho\log\rho$. We also let $\mathrm{S}\left(X\right)_\rho$ represent $\mathrm{S}(\rho)$. The *relative entropy* between quantum states ρ and σ is defined as $\mathrm{S}(\rho\|\sigma) \stackrel{\text{def}}{=} \mathrm{Tr}\rho\log\rho - \mathrm{Tr}\rho\log\sigma$. The *relative min-entropy* between them is defined as $\mathrm{S}_\infty(\rho\|\sigma) \stackrel{\text{def}}{=} \min\left\{\lambda : \rho \leq 2^\lambda\sigma\right\}$. Since the logarithm is operator-monotone, $\mathrm{S}(\rho\|\sigma) \leq \mathrm{S}_\infty(\rho\|\sigma)$.

Let ρ^{XY} be a quantum state in space $X \otimes Y$. The *mutual information* between registers X and Y is defined to be

$$\mathrm{I}(X:Y)_\rho \stackrel{\text{def}}{=} \mathrm{S}(X)_\rho + \mathrm{S}(Y)_\rho - \mathrm{S}(XY)_\rho .$$

It is easy to see that $\mathrm{I}(X:Y)_\rho = \mathrm{S}\left(\rho^{XY} \middle\| \rho^X \otimes \rho^Y\right)$. If X is a classical register, namely $\rho = \sum_x \mu(x) \left|x\right\rangle\!\left\langle x\right| \otimes \rho_x$, where μ is a probability distribution over X, then

$$\begin{aligned}\mathrm{I}(X:Y)_\rho &= \mathrm{S}(Y)_\rho - \mathrm{S}(Y|X)_\rho \\ &= \mathrm{S}\left(\sum_x \mu(x)\rho_x\right) - \sum_x \mu(x)\mathrm{S}(\rho_x)\end{aligned}$$

where the *conditional entropy* is defined as

$$\mathrm{S}(Y|X)_\rho \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{x\leftarrow\mu}[\mathrm{S}(\rho_x)] .$$

Let ρ^{XYZ} be a quantum state with Y being a classical register. The mutual information between X and Z, conditioned on Y, is defined as

$$\begin{aligned}\mathrm{I}(X:Z|Y)_\rho &\stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{y\leftarrow Y}\left[\mathrm{I}(X:Z|Y=y)_\rho\right] \\ &= \mathrm{S}(X|Y)_\rho + \mathrm{S}(Z|Y)_\rho - \mathrm{S}(XZ|Y)_\rho .\end{aligned}$$

The following *chain rule* for mutual information follows easily from the definitions, when Y is a classical register.

$$\mathrm{I}(X:YZ)_\rho = \mathrm{I}(X:Y)_\rho + \mathrm{I}(X:Z|Y)_\rho .$$

We will need the following basic facts.

Fact II.7. The relative entropy is jointly convex in its arguments. That is, for quantum states ρ, ρ^1, σ, and σ^1, and $p \in [0,1]$,

$$\begin{aligned}&\mathrm{S}\left(p\rho + (1-p)\rho^1 \middle\| p\sigma + (1-p)\sigma^1\right) \\ &\quad \leq p \cdot \mathrm{S}(\rho\|\sigma) + (1-p)\cdot \mathrm{S}\left(\rho^1 \middle\| \sigma^1\right) .\end{aligned}$$

We have the following chain rule for the relative-entropy.

Fact II.8. Let

$$\rho = \sum_x \mu(x) \left|x\right\rangle\!\left\langle x\right| \otimes \rho_x$$

and

$$\rho^1 = \sum_x \mu^1(x) \left|x\right\rangle\!\left\langle x\right| \otimes \rho^1_x .$$

It holds that

$$\mathrm{S}\left(\rho^1 \middle\| \rho\right) = \mathrm{S}\left(\mu^1 \middle\| \mu\right) + \mathop{\mathbb{E}}_{x\leftarrow\mu^1}\left[\mathrm{S}\left(\rho^1_x \middle\| \rho_x\right)\right] .$$

Fact II.9. For quantum states ρ^{XY}, σ^X, and τ^Y, it holds that

$$\mathrm{S}\left(\rho^{XY} \middle\| \sigma^X \otimes \tau^Y\right) \geq \mathrm{S}\left(\rho^{XY} \middle\| \rho^X \otimes \rho^Y\right) = \mathrm{I}(X:Y)_\rho .$$

Fact II.10 ([26], [27])**.** For quantum states ρ and σ, it holds that

$$\|\rho - \sigma\|_1 \leq \sqrt{\mathrm{S}(\rho\|\sigma)} \quad \text{and} \quad 1 - \mathrm{F}(\rho,\sigma) \leq \mathrm{S}(\rho\|\sigma) .$$

Fact II.11. The relative entropy is non-increasing when subsystems are considered. Let ρ^{XY} and σ^{XY} be quantum states, then $\mathrm{S}\left(\rho^{XY} \middle\| \sigma^{XY}\right) \geq \mathrm{S}\left(\rho^X \middle\| \sigma^X\right)$.

The following fact is easily verified.

Fact II.12. Let $0 < \varepsilon, \varepsilon' < 1$, $0 < c$, μ and μ' be probability distributions on a set $\mathcal{X}$, and $f : \mathcal{X} \to [0,c]$ be a function. If $\mathbb{E}_{x\leftarrow\mu}[f(x)] \leq \varepsilon$ and $\|\mu - \mu'\|_1 \leq \varepsilon'$ then $\mathbb{E}_{x\leftarrow\mu'}[f(x)] \leq \varepsilon + c\varepsilon'$.

Useful lemmas

Here we state and prove some lemmas that we will use later.

Lemma II.13. *Let $|\psi\rangle^{AB}$ be a bipartite pure state with the marginal state on register B being ρ. Let a $0/1$ outcome measurement be performed on register A with outcome O. Let $\Pr[O=1] = q$. Let the marginal states on register B conditioned on $O=0$ and $O=1$ be ρ_0 and ρ_1 respectively. Then, $\mathrm{S}_\infty(\rho_1\|\rho) \leq \log\frac{1}{q}$.*

Proof. It is easily seen that $\rho = q\rho_1 + (1-q)\rho_0$. Hence $\mathrm{S}_\infty(\rho_1\|\rho) \leq \log\frac{1}{q}$. ☐

The following lemma states that when the concerned mutual information is small, then a measurement on Alice's side can be simulated by a unitary operation on Alice's side.

Lemma II.14. *Let μ be a probability distribution on $\mathcal{X}$. Let*

$$|\varphi\rangle \stackrel{\text{def}}{=} \sum_{x\in\mathcal{X}} \sqrt{\mu(x)}\, |xx\rangle^{\tilde{X}X} \otimes |\psi_x\rangle^{AB}$$

be a joint pure state of Alice and Bob, where registers $\tilde{X}XA$ are with Alice and register B is with Bob. Let $\mathrm{I}(X:B)_\varphi \leq \varepsilon$ and $|\varphi_x\rangle \stackrel{\text{def}}{=} |xx\rangle \otimes |\psi_x\rangle$. There exist unitary operators $\{\mathbf{U}_x\}_{x\in\mathcal{X}}$ acting on $\tilde{X}XA$ such that

$$\mathop{\mathbb{E}}_{x\leftarrow\mu}[\| |\varphi_x\rangle\!\langle\varphi_x| - (\mathbf{U}_x \otimes \mathbb{1}_B)\, |\varphi\rangle\!\langle\varphi|\, (\mathbf{U}^*_x \otimes \mathbb{1}_B)\|_1] \leq 4\sqrt{\varepsilon}.$$

Proof. Let us denote the reduced state of Bob in $|\varphi_x\rangle$ and $|\varphi\rangle$ by

$$\rho_x \stackrel{\text{def}}{=} \mathrm{Tr}_A(|\psi_x\rangle\!\langle\psi_x|) \quad \text{and} \quad \rho \stackrel{\text{def}}{=} \mathrm{Tr}_{\tilde{X}XA}(|\varphi\rangle\!\langle\varphi|) .$$

Using Fact II.10, it holds that

$$\varepsilon \geq \mathrm{I}(X:B) = \mathop{\mathbb{E}}_{x\leftarrow\mu}[\mathrm{S}(\rho_x\|\rho)] \geq 1 - \mathop{\mathbb{E}}_{x\leftarrow\mu}[\mathrm{F}(\rho_x,\rho)] .$$

By the unitary equivalence of purifications and Theorem II.6, there exists a $\mathbf{U}_x$ for each $x \in \mathcal{X}$ such that

$$|\langle\varphi_x|\, (\mathbf{U}_x \otimes \mathbb{1}_B)\, |\varphi\rangle| = \mathrm{F}(\rho_x,\rho).$$

The lemma follows from the following calculation.

$$\begin{aligned}&\mathop{\mathbb{E}}_{x\leftarrow\mu}[\| |\varphi_x\rangle\!\langle\varphi_x| - (\mathbf{U}_x \otimes \mathbb{1}_B)\, |\varphi\rangle\!\langle\varphi|\, (\mathbf{U}^*_x \otimes \mathbb{1}_B)\|_1] \\ &\quad = 2\mathop{\mathbb{E}}_{x\leftarrow\mu}\left[\sqrt{1 - |\langle\varphi_x|\, (\mathbf{U}_x \otimes \mathbb{1}_B)\, |\varphi\rangle|^2}\right] \qquad (4) \\ &\quad \leq 2\sqrt{1 - \mathop{\mathbb{E}}_{x\leftarrow\mu}[|\langle\varphi_x|\, (\mathbf{U}_x \otimes \mathbb{1}_B)\, |\varphi\rangle|]^2} \qquad (5) \\ &\quad = 2\sqrt{1 - \mathop{\mathbb{E}}_{x\leftarrow\mu}[\mathrm{F}(\rho_x,\rho)]^2} \\ &\quad \leq 4\sqrt{\varepsilon}.\end{aligned}$$

where Eq. (4) follows from Proposition II.4 and at Eq. (5) we used the concavity of the function $\sqrt{1-\alpha^2}$. $\square$

The following is a generalization of the above lemma that states that when the concerned mutual informations are small then the simultaneous measurements on Alice's and Bob's side can be simulated by unitary operations on Alice's and Bob's side. It is a special case of a more general result in Ref. [3].

Lemma II.15 ([3])**.** *Let μ be a probability distribution over $\mathcal{X} \times \mathcal{Y}$. Let μ_X and μ_Y be the marginals of μ on $\mathcal{X}$ and $\mathcal{Y}$. Let*

$$|\varphi\rangle \stackrel{\text{def}}{=} \sum_{x\in\mathcal{X}, y\in\mathcal{Y}} \sqrt{\mu(x,y)}\, |xxyy\rangle^{\tilde{X}X\tilde{Y}Y} \otimes |\psi_{x,y}\rangle^{AB}$$

be a joint pure state of Alice and Bob, where registers $\tilde{X}XA$ belong to Alice and registers $\tilde{Y}YB$ belong to Bob. Let

$$\mathrm{I}\Big(X : BY\tilde{Y}\Big)_{\varphi} \leq \varepsilon \quad \text{and} \quad \mathrm{I}\Big(Y : AX\tilde{X}\Big)_{\varphi} \leq \varepsilon.$$

Let $|\varphi_{x,y}\rangle \stackrel{\text{def}}{=} |xxyy\rangle \otimes |\psi_{x,y}\rangle$. There exist unitary operators $\{\mathbf{U}_x\}_{x\in\mathcal{X}}$ on $\tilde{X}XA$ and $\{\mathbf{V}_y\}_{y\in\mathcal{Y}}$ on $\tilde{Y}YB$ such that

$$\mathop{\mathbb{E}}_{(x,y)\leftarrow\mu}\Big[\big\||\varphi_{x,y}\rangle\langle\varphi_{x,y}| - (\mathbf{U}_x\otimes\mathbf{V}_y)\,|\varphi\rangle\langle\varphi|\,\big(\mathbf{U}_x^*\otimes\mathbf{V}_y^*\big)\big\|_1\Big] \leq 8\sqrt{\varepsilon} + 2\,\|\mu - \mu_X\otimes\mu_Y\|_1 .$$

Proof. Let $|\varphi_x\rangle$ be the state obtained when we measure register X in $|\varphi\rangle$ and obtain x. Similarly let $|\varphi_y\rangle$ be the state obtained when we measure register Y in $|\varphi\rangle$ and obtain y. By Lemma II.14, there exist unitary operators $\{\mathbf{U}_x\}_{x\in\mathcal{X}}$ and $\{\mathbf{V}_y\}_{y\in\mathcal{Y}}$ such that

$$\mathop{\mathbb{E}}_{x\leftarrow\mu_X}\left[\||\varphi_x\rangle\langle\varphi_x| - (\mathbf{U}_x\otimes\mathbb{1}_B)\,|\varphi\rangle\langle\varphi|\,(\mathbf{U}_x^*\otimes\mathbb{1}_B)\|_1\right] \leq 4\sqrt{\varepsilon}$$

and

$$\mathop{\mathbb{E}}_{y\leftarrow\mu_Y}\Big[\big\||\varphi_y\rangle\langle\varphi_y| - (\mathbb{1}_A\otimes\mathbf{V}_y)\,|\varphi\rangle\langle\varphi|\,\big(\mathbb{1}_A\otimes\mathbf{V}_y^*\big)\big\|_1\Big] \leq 4\sqrt{\varepsilon}.$$

Using the above, we get the bound of Eq. (3) from the calculation that is on the bottom of this page. Equation (1) follows from the triangle inequality, the second term in Eq. (2) is because $\mathbf{U}_x$ doesn't change the ℓ_1-distance, and the first term in Eq. (2) follows from Proposition II.3 with the super-operator that corresponds to measuring Y in the standard basis and storing the outcome in a new register. The lemma follows from the following calculation.

$$\begin{aligned}
&\mathop{\mathbb{E}}_{(x,y)\leftarrow\mu}\Big[\big\||\varphi_{x,y}\rangle\langle\varphi_{x,y}| - (\mathbf{U}_x\otimes\mathbf{V}_y)\,|\varphi\rangle\langle\varphi|\,\big(\mathbf{U}_x^*\otimes\mathbf{V}_y^*\big)\big\|_1\Big] \\
&= \Bigg\| \mathop{\mathbb{E}}_{(x,y)\leftarrow\mu}\Big[\,|xy\rangle\langle xy|\otimes|\varphi_{x,y}\rangle\langle\varphi_{x,y}| \\
&\qquad - |xy\rangle\langle xy|\otimes(\mathbf{U}_x\otimes\mathbf{V}_y)\,|\varphi\rangle\langle\varphi|\,\big(\mathbf{U}_x^*\otimes\mathbf{V}_y^*\big)\,\Big]\Bigg\|_1 \\
&\leq \Bigg\| \mathop{\mathbb{E}}_{(x,y)\leftarrow\mu}\left[|xy\rangle\langle xy|\otimes|\varphi_{x,y}\rangle\langle\varphi_{x,y}|\right] \\
&\qquad - \mathop{\mathbb{E}}_{(x,y)\leftarrow\mu_X\otimes\mu_Y}\Big[\,|xy\rangle\langle xy| \\
&\qquad\qquad \otimes(\mathbf{U}_x\otimes\mathbf{V}_y)\,|\varphi\rangle\langle\varphi|\,\big(\mathbf{U}_x^*\otimes\mathbf{V}_y^*\big)\,\Big]\Bigg\|_1 \\
&\quad + \Bigg\| \mathop{\mathbb{E}}_{(x,y)\leftarrow\mu_X\otimes\mu_Y}\Big[\,|xy\rangle\langle xy| \\
&\qquad\qquad \otimes(\mathbf{U}_x\otimes\mathbf{V}_y)\,|\varphi\rangle\langle\varphi|\,\big(\mathbf{U}_x^*\otimes\mathbf{V}_y^*\big)\,\Big] \\
&\qquad - \mathop{\mathbb{E}}_{(x,y)\leftarrow\mu}\Big[\,|xy\rangle\langle xy| \\
&\qquad\qquad \otimes(\mathbf{U}_x\otimes\mathbf{V}_y)\,|\varphi\rangle\langle\varphi|\,\big(\mathbf{U}_x^*\otimes\mathbf{V}_y^*\big)\,\Big]\Bigg\|_1 \\
&\leq 8\sqrt{\varepsilon} + 2\,\|\mu-\mu_X\otimes\mu_Y\|_1
\end{aligned}$$

where the first inequality follows from the triangle inequality and at the last inequality we used Eq. (3) and Fact II.12. $\square$

III. Proof of the main result

Let a game $G = (\mu, \mathcal{X}, \mathcal{Y}, \mathcal{A}, \mathcal{B}, V)$ be given. We assume that the distribution $\mu = \mu_X\otimes\mu_Y$ is product across $\mathcal{X}$ and

$$\begin{aligned}
&\Bigg\| \mathop{\mathbb{E}}_{(x,y)\leftarrow\mu}\left[|xy\rangle\langle xy|\otimes|\varphi_{x,y}\rangle\langle\varphi_{x,y}|\right] - \mathop{\mathbb{E}}_{(x,y)\leftarrow\mu_X\otimes\mu_Y}\Big[|xy\rangle\langle xy|\otimes(\mathbf{U}_x\otimes\mathbf{V}_y)\,|\varphi\rangle\langle\varphi|\,\big(\mathbf{U}_x^*\otimes\mathbf{V}_y^*\big)\Big]\Bigg\|_1 \\
&\leq \Bigg\| \mathop{\mathbb{E}}_{(x,y)\leftarrow\mu}\left[|xy\rangle\langle xy|\otimes|\varphi_{x,y}\rangle\langle\varphi_{x,y}|\right] - \mathop{\mathbb{E}}_{(x,y)\leftarrow\mu_X\otimes\mu_Y}\left[|xy\rangle\langle xy|\otimes(\mathbf{U}_x\otimes\mathbb{1}_B)\,|\varphi_y\rangle\langle\varphi_y|\,(\mathbf{U}_x^*\otimes\mathbb{1}_B)\right]\Bigg\|_1 \\
&\quad + \Bigg\| \mathop{\mathbb{E}}_{(x,y)\leftarrow\mu_X\otimes\mu_Y}\Big[|xy\rangle\langle xy|\otimes(\mathbf{U}_x\otimes\mathbb{1}_B)\,|\varphi_y\rangle\langle\varphi_y|\,(\mathbf{U}_x^*\otimes\mathbb{1}_B) - |xy\rangle\langle xy|\otimes(\mathbf{U}_x\otimes\mathbf{V}_y)\,|\varphi\rangle\langle\varphi|\,\big(\mathbf{U}_x^*\otimes\mathbf{V}_y^*\big)\Big]\Bigg\|_1 \qquad (1) \\
&\leq \Bigg\| \mathop{\mathbb{E}}_{x\leftarrow\mu_X}\left[|x\rangle\langle x|\otimes|\varphi_x\rangle\langle\varphi_x| - |x\rangle\langle x|\otimes(\mathbf{U}_x\otimes\mathbb{1}_B)\,|\varphi\rangle\langle\varphi|\,(\mathbf{U}_x^*\otimes\mathbb{1}_B)\right]\Bigg\|_1 \\
&\quad + \Bigg\| \mathop{\mathbb{E}}_{xy\leftarrow\mu_X\otimes\mu_Y}\Big[|xy\rangle\langle xy|\otimes|\varphi_y\rangle\langle\varphi_y| - |xy\rangle\langle xy|\otimes(\mathbb{1}_A\otimes\mathbf{V}_y)\,|\varphi\rangle\langle\varphi|\,\big(\mathbb{1}_A\otimes\mathbf{V}_y^*\big)\Big]\Bigg\|_1 \qquad (2) \\
&= \mathop{\mathbb{E}}_{x\leftarrow\mu_X}\left[\||\varphi_x\rangle\langle\varphi_x| - (\mathbf{U}_x\otimes\mathbb{1}_B)\,|\varphi\rangle\langle\varphi|\,(\mathbf{U}_x^*\otimes\mathbb{1}_B)\|_1\right] \\
&\quad + \mathop{\mathbb{E}}_{y\leftarrow\mu_Y}\Big[\big\||\varphi_y\rangle\langle\varphi_y| - (\mathbb{1}_A\otimes\mathbf{V}_y)\,|\varphi\rangle\langle\varphi|\,\big(\mathbb{1}_A\otimes\mathbf{V}_y^*\big)\big\|_1\Big] \\
&\leq 8\sqrt{\varepsilon} \qquad (3)
\end{aligned}$$

$\mathcal{Y}$. Before the game starts, Alice and Bob share a pure state on the registers $AE'_ABE'_B$, where A and B are used to store the answers for Alice and Bob, respectively. After getting the inputs, Alice and Bob perform unitary operations independently and then they measure registers A and B. The outcomes of the measurements are sent to the referee. Now, let's consider the game G^k. Let $x = x_1 \dots x_k \in \mathcal{X}^k$, $y = y_1 \dots y_k \in \mathcal{Y}^k$, $a = a_1 \dots a_k \in \mathcal{A}^k$, and $b = b_1 \dots b_k \in \mathcal{B}^k$. To make notations short, we denote $\mu(x, y) = \prod_i \mu(x_i, y_i)$ and $V(x, y, a, b) = \prod_i V(x_i, y_i, a_i, b_i)$, whenever it is clear from the context. Let $\mathcal{C} \subseteq [k]$ and let $\bar{\mathcal{C}}$ represent its *complement* in $[k]$. Let $x_{\mathcal{C}}$ represent the substring of x corresponding to the indices in $\mathcal{C}$. (Similarly, we will use $y_{\mathcal{C}}, a_{\mathcal{C}}, b_{\mathcal{C}}$.) Let's define

$$|\theta\rangle \stackrel{\text{def}}{=} \sum_{x,y} \sqrt{\mu(x,y)}\, |xxyy\rangle^{\tilde{X}X\tilde{Y}Y} \otimes \sum_{a_{\mathcal{C}}b_{\mathcal{C}}} |a_{\mathcal{C}}b_{\mathcal{C}}\rangle^{A_{\mathcal{C}}B_{\mathcal{C}}} \otimes |\gamma_{xya_{\mathcal{C}}b_{\mathcal{C}}}\rangle^{E_AE_B}$$

where $E_A \stackrel{\text{def}}{=} E'_A A_{\bar{\mathcal{C}}}$, $E_B \stackrel{\text{def}}{=} E'_B B_{\bar{\mathcal{C}}}$, and $\sum_{a_{\mathcal{C}}b_{\mathcal{C}}} |a_{\mathcal{C}}b_{\mathcal{C}}\rangle \otimes |\gamma_{xya_{\mathcal{C}}b_{\mathcal{C}}}\rangle$ is the shared state after Alice and Bob performed their unitary operations corresponding to questions x and y. (Note that $|\gamma_{xya_{\mathcal{C}}b_{\mathcal{C}}}\rangle$ is unnormalized.) Consider the state

$$|\varphi\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{q}} \sum_{x,y} \sqrt{\mu(x,y)}\, |xxyy\rangle^{\tilde{X}X\tilde{Y}Y} \otimes \sum_{a_{\mathcal{C}}b_{\mathcal{C}}: V(x_{\mathcal{C}}, y_{\mathcal{C}}, a_{\mathcal{C}}, b_{\mathcal{C}})=1} |a_{\mathcal{C}}b_{\mathcal{C}}\rangle^{A_{\mathcal{C}}B_{\mathcal{C}}} \otimes |\gamma_{xya_{\mathcal{C}}b_{\mathcal{C}}}\rangle^{E_AE_B}$$

where normalizer q is the probability of success on $\mathcal{C}$.

Lemma III.1.

$$\mathop{\mathbb{E}}_{x_{\mathcal{C}}y_{\mathcal{C}}a_{\mathcal{C}}b_{\mathcal{C}} \leftarrow \varphi^{X_{\mathcal{C}}Y_{\mathcal{C}}A_{\mathcal{C}}B_{\mathcal{C}}}} \left[\mathrm{S}\left(\varphi_{x_{\mathcal{C}}y_{\mathcal{C}}a_{\mathcal{C}}b_{\mathcal{C}}}^{\tilde{X}_{\bar{\mathcal{C}}}\tilde{Y}_{\bar{\mathcal{C}}}XYE_AE_B} \middle\| \theta_{x_{\mathcal{C}}y_{\mathcal{C}}}^{\tilde{X}_{\bar{\mathcal{C}}}\tilde{Y}_{\bar{\mathcal{C}}}XYE_AE_B} \right) \right] \le -\log q + |\mathcal{C}| \cdot \log(|\mathcal{A}| \cdot |\mathcal{B}|).$$

Proof. Note that, by Lemma II.13,

$$\mathrm{S}_\infty\left(\varphi^{\tilde{X}_{\bar{\mathcal{C}}}\tilde{Y}_{\bar{\mathcal{C}}}XYE_AE_B} \middle\| \theta^{\tilde{X}_{\bar{\mathcal{C}}}\tilde{Y}_{\bar{\mathcal{C}}}XYE_AE_B} \right) \le -\log q.$$

Let $p(a_{\mathcal{C}}, b_{\mathcal{C}})$ be the probability of obtaining $(a_{\mathcal{C}}, b_{\mathcal{C}})$ when measuring registers $(A_{\mathcal{C}}, B_{\mathcal{C}})$ in $|\varphi\rangle$. Consider,

$$\begin{aligned}
&\mathop{\mathbb{E}}_{a_{\mathcal{C}}b_{\mathcal{C}} \leftarrow \varphi^{A_{\mathcal{C}}B_{\mathcal{C}}}} \left[\mathrm{S}_\infty\left(\varphi_{a_{\mathcal{C}}b_{\mathcal{C}}}^{\tilde{X}_{\bar{\mathcal{C}}}\tilde{Y}_{\bar{\mathcal{C}}}XYE_AE_B} \middle\| \theta^{\tilde{X}_{\bar{\mathcal{C}}}\tilde{Y}_{\bar{\mathcal{C}}}XYE_AE_B} \right) \right] \\
&\le \mathop{\mathbb{E}}_{a_{\mathcal{C}}b_{\mathcal{C}} \leftarrow \varphi^{A_{\mathcal{C}}B_{\mathcal{C}}}} \left[\mathrm{S}_\infty\left(\varphi_{a_{\mathcal{C}}b_{\mathcal{C}}}^{\tilde{X}_{\bar{\mathcal{C}}}\tilde{Y}_{\bar{\mathcal{C}}}XYE_AE_B} \middle\| \varphi^{\tilde{X}_{\bar{\mathcal{C}}}\tilde{Y}_{\bar{\mathcal{C}}}XYE_AE_B} \right) + \mathrm{S}_\infty\left(\varphi^{\tilde{X}_{\bar{\mathcal{C}}}\tilde{Y}_{\bar{\mathcal{C}}}XYE_AE_B} \middle\| \theta^{\tilde{X}_{\bar{\mathcal{C}}}\tilde{Y}_{\bar{\mathcal{C}}}XYE_AE_B} \right) \right] \\
&\le \mathop{\mathbb{E}}_{a_{\mathcal{C}}b_{\mathcal{C}} \leftarrow \varphi^{A_{\mathcal{C}}B_{\mathcal{C}}}} \left[-\log p(a_{\mathcal{C}}, b_{\mathcal{C}}) - \log q \right] \\
&= -\log q + \mathrm{S}\left(\varphi^{A_{\mathcal{C}}B_{\mathcal{C}}}\right) \\
&\le -\log q + |\mathcal{C}| \cdot (\log|\mathcal{A}| + \log|\mathcal{B}|).
\end{aligned}$$

Now,

$$\begin{aligned}
&-\log q + |\mathcal{C}| \cdot \log(|\mathcal{A}| \cdot |\mathcal{B}|) \\
&\ge \mathop{\mathbb{E}}_{a_{\mathcal{C}}b_{\mathcal{C}} \leftarrow \varphi^{A_{\mathcal{C}}B_{\mathcal{C}}}} \left[\mathrm{S}_\infty\left(\varphi_{a_{\mathcal{C}}b_{\mathcal{C}}}^{\tilde{X}_{\bar{\mathcal{C}}}\tilde{Y}_{\bar{\mathcal{C}}}XYE_AE_B} \middle\| \theta^{\tilde{X}_{\bar{\mathcal{C}}}\tilde{Y}_{\bar{\mathcal{C}}}XYE_AE_B} \right) \right] \\
&\ge \mathop{\mathbb{E}}_{a_{\mathcal{C}}b_{\mathcal{C}} \leftarrow \varphi^{A_{\mathcal{C}}B_{\mathcal{C}}}} \left[\mathrm{S}\left(\varphi_{a_{\mathcal{C}}b_{\mathcal{C}}}^{\tilde{X}_{\bar{\mathcal{C}}}\tilde{Y}_{\bar{\mathcal{C}}}XYE_AE_B} \middle\| \theta^{\tilde{X}_{\bar{\mathcal{C}}}\tilde{Y}_{\bar{\mathcal{C}}}XYE_AE_B} \right) \right] \\
&\ge \mathop{\mathbb{E}}_{\substack{x_{\mathcal{C}}y_{\mathcal{C}}a_{\mathcal{C}}b_{\mathcal{C}} \\ \leftarrow \varphi^{X_{\mathcal{C}}Y_{\mathcal{C}}A_{\mathcal{C}}B_{\mathcal{C}}}}} \left[\mathrm{S}\left(\varphi_{x_{\mathcal{C}}y_{\mathcal{C}}a_{\mathcal{C}}b_{\mathcal{C}}}^{\tilde{X}_{\bar{\mathcal{C}}}\tilde{Y}_{\bar{\mathcal{C}}}XYE_AE_B} \middle\| \theta_{x_{\mathcal{C}}y_{\mathcal{C}}}^{\tilde{X}_{\bar{\mathcal{C}}}\tilde{Y}_{\bar{\mathcal{C}}}XYE_AE_B} \right) \right]
\end{aligned}$$

where the last inequality follows from Fact II.8. □

For each $i \in [k]$, let us define a binary random variable $T_i \in \{0, 1\}$, which indicates success in the i-th repetition. That is, $T_i = V(X_i, Y_i, A_i, B_i)$. Our main theorem will follow from the following lemma.

Lemma III.2. *Let $0.1 > \delta_1, \delta_2, \delta_3 > 0$ such that $\delta_3 = \delta_2 + \delta_1 \cdot \log(|\mathcal{A}| \cdot |\mathcal{B}|)$. Let $k' \stackrel{\text{def}}{=} \lfloor \delta_1 k \rfloor$. For any quantum strategy for the k-fold game G^k, there exists a set $\{i_1, \dots, i_{k'}\}$, such that for each $1 \le r \le k' - 1$, either*

$$\Pr\left[T^{(r)} = 1\right] \le 2^{-\delta_2 k}$$

or

$$\Pr\left[T_{i_{r+1}} = 1 \middle| T^{(r)} = 1\right] \le \omega^*(G) + 12\sqrt{10\delta_3}$$

where $T^{(r)} \stackrel{\text{def}}{=} \prod_{j=1}^{r} T_{i_j}$.

Proof. In the following, we assume that $1 \le r < k'$. However, the same argument also works when $r = 0$, i.e., for identifying the first coordinate, which we skip for the sake of avoiding repetition. Suppose that we have already identified r coordinates $i_1, \dots, i_r$ satisfying that

$$\Pr[T_{i_1} = 1] \le \omega^*(G) + 12\sqrt{10\delta_3}$$

and

$$\Pr\left[T_{i_{j+1}} = 1 \middle| T^{(j)} = 1\right] \le \omega^*(G) + 12\sqrt{10\delta_3}$$

for $1 \le j \le r-1$. If $\Pr\left[T^{(r)} = 1\right] \le 2^{-\delta_2 k}$ then we are done, so from now on, we assume that $\Pr\left[T^{(r)} = 1\right] > 2^{-\delta_2 k}$. Let $\mathcal{C} \stackrel{\text{def}}{=} \{i_1, \dots, i_r\}$. To simplify notations, let $\tilde{A} \stackrel{\text{def}}{=} \tilde{X}_{\bar{\mathcal{C}}} X E_A$, $\tilde{B} \stackrel{\text{def}}{=} \tilde{Y}_{\bar{\mathcal{C}}} Y E_B$, and $R_i \stackrel{\text{def}}{=} X_{\mathcal{C}} Y_{\mathcal{C}} X_{<i} Y_{<i} A_{\mathcal{C}} B_{\mathcal{C}}$. For coordinate i, let $|\varphi_{x_{<i}y_{<i}}\rangle$ be the pure state that results when we measure registers $X_{<i}Y_{<i}$ (i.e., registers $X_1, \dots, X_{i-1}, Y_1, \dots, Y_{i-1}$) in $|\varphi\rangle$ and get outcome $x_{<i}y_{<i}$. We argue now that for a typical coordinate outside $\mathcal{C}$, the distribution of questions is close to μ in the state φ. We also prove that, for this coordinate, the questions and R_i are almost

independent. From Lemma III.1, we get that

$$\begin{aligned}
\delta_3 k &\geq \delta_2 k + |\mathcal{C}| \cdot \log(|\mathcal{A}| \cdot |\mathcal{B}|) \\
&\geq \mathop{\mathbb{E}}_{\substack{x_\mathcal{C} y_\mathcal{C} a_\mathcal{C} b_\mathcal{C} \\ \leftarrow \varphi^{X_\mathcal{C} Y_\mathcal{C} A_\mathcal{C} B_\mathcal{C}}}} \left[\mathrm{S}\left(\varphi^{\tilde{X}_{\bar{\mathcal{C}}} \tilde{Y}_{\bar{\mathcal{C}}} X Y E_A E_B}_{x_\mathcal{C} y_\mathcal{C} a_\mathcal{C} b_\mathcal{C}} \middle\| \theta^{\tilde{X}_{\bar{\mathcal{C}}} \tilde{Y}_{\bar{\mathcal{C}}} X Y E_A E_B}_{x_\mathcal{C} y_\mathcal{C}} \right) \right] \qquad (6) \\
&\geq \mathop{\mathbb{E}}_{x_\mathcal{C} y_\mathcal{C} a_\mathcal{C} b_\mathcal{C} \leftarrow \varphi^{X_\mathcal{C} Y_\mathcal{C} A_\mathcal{C} B_\mathcal{C}}} \left[\mathrm{S}\left(\varphi^{XY}_{x_\mathcal{C} y_\mathcal{C} a_\mathcal{C} b_\mathcal{C}} \middle\| \theta^{XY}_{x_\mathcal{C} y_\mathcal{C}} \right) \right] \qquad (7) \\
&= \sum_{i \notin \mathcal{C}} \mathop{\mathbb{E}}_{r_i \leftarrow \varphi^{R_i}} \left[\mathrm{S}\left(\varphi^{X_i Y_i}_{r_i} \middle\| \theta^{X_i Y_i} \right) \right] \qquad (8) \\
&= \sum_{i \notin \mathcal{C}} \mathrm{S}\left(\varphi^{X_i Y_i R_i} \middle\| \varphi^{R_i} \otimes \theta^{X_i Y_i} \right) \qquad (9) \\
&\geq \sum_{i \notin \mathcal{C}} \mathrm{S}\left(\varphi^{X_i Y_i R_i} \middle\| \varphi^{R_i} \otimes \varphi^{X_i Y_i} \right) \qquad (10) \\
&\geq \sum_{i \notin \mathcal{C}} \mathop{\mathbb{E}}_{x_i y_i \leftarrow \varphi^{X_i Y_i}} \left[\mathrm{S}\left(\varphi^{R_i}_{x_i y_i} \middle\| \varphi^{R_i} \right) \right] \qquad (11) \\
&\geq \sum_{i \notin \mathcal{C}} \mathop{\mathbb{E}}_{x_i y_i \leftarrow \varphi^{X_i Y_i}} \left[\left\| \varphi^{R_i}_{x_i y_i} - \varphi^{R_i} \right\|_1^2 \right] \qquad (12) \\
&\geq \sum_{i \notin \mathcal{C}} \left(\mathop{\mathbb{E}}_{x_i y_i \leftarrow \varphi^{X_i Y_i}} \left[\left\| \varphi^{R_i}_{x_i y_i} - \varphi^{R_i} \right\|_1 \right] \right)^2 \qquad (13)
\end{aligned}$$

where Eq. (6) follows from Lemma III.1, Eq. (7) follows from Fact II.11, Eqs. (8), (9) and (11) follow from Fact II.8, Eq. (10) follows from Fact II.9, Eq. (12) follows from Fact II.10, and Eq. (13) follows from the convexity of the function α^2. Next, we argue that for a typical coordinate outside $\mathcal{C}$, the information between Alice's questions and Bob's registers is small in $|\varphi\rangle$. Again, from Lemma III.1 and Fact II.11, we get that

$$\begin{aligned}
\delta_3 k &\geq \mathop{\mathbb{E}}_{x_\mathcal{C} y_\mathcal{C} a_\mathcal{C} b_\mathcal{C} \leftarrow \varphi^{X_\mathcal{C} Y_\mathcal{C} A_\mathcal{C} B_\mathcal{C}}} \left[\mathrm{S}\left(\varphi^{X \tilde{B}}_{x_\mathcal{C} y_\mathcal{C} a_\mathcal{C} b_\mathcal{C}} \middle\| \theta^{X \tilde{B}}_{x_\mathcal{C} y_\mathcal{C}} \right) \right] \\
&\geq \mathrm{I}\left(X : \tilde{B} \middle| R_1 \right)_\varphi \qquad (14) \\
&\geq \sum_{i \notin \mathcal{C}} \mathrm{I}\left(X_i : \tilde{B} \middle| R_1 X_{<i} \right)_\varphi \qquad (15) \\
&\geq \sum_{i \notin \mathcal{C}} \mathrm{I}\left(X_i : \tilde{B} \middle| R_i \right)_\varphi \qquad (16)
\end{aligned}$$

where at Eq. (14) we used Fact II.9 and the fact that $\theta^{X\tilde{B}}_{x_\mathcal{C} y_\mathcal{C}} = \theta^{X}_{x_\mathcal{C} y_\mathcal{C}} \otimes \theta^{\tilde{B}}_{x_\mathcal{C} y_\mathcal{C}}$. Equations (15) and (16) follow from the chain rule for the mutual information and at Eq. (16) we also used the observation that $\tilde{B}$ contains register Y. Similarly to the above, for Bob's questions we have

$$\delta_3 k \geq \sum_{i \notin \mathcal{C}} \mathrm{I}\left(Y_i : \tilde{A} \middle| R_i \right)_\varphi . \qquad (17)$$

From Eqs. (8), (13), (16) and (17) and using standard application of Markov's inequality, we get that there exists a coordinate $j \notin \mathcal{C}$ such that

$$\mathop{\mathbb{E}}_{r_j \leftarrow \varphi^{R_j}} \left[\mathrm{S}\left(\varphi^{X_j Y_j}_{r_j} \middle\| \theta^{X_j Y_j} \right) \right] \leq \frac{5\delta_3}{1-\delta_1} \leq 10\delta_3 \qquad (18)$$

$$\mathop{\mathbb{E}}_{x_j y_j \leftarrow \varphi^{X_j Y_j}} \left[\left\| \varphi^{R_j}_{x_j y_j} - \varphi^{R_j} \right\|_1 \right] \leq \sqrt{\frac{5\delta_3}{1-\delta_1}} \leq \sqrt{10\delta_3} \qquad (19)$$

$$\mathrm{I}\left(X_j : \tilde{B} \middle| R_j \right)_\varphi \leq \frac{5\delta_3}{1-\delta_1} \leq 10\delta_3 \qquad (20)$$

$$\mathrm{I}\left(Y_j : \tilde{A} \middle| R_j \right)_\varphi \leq \frac{5\delta_3}{1-\delta_1} \leq 10\delta_3 . \qquad (21)$$

Let $\left|\varphi_{r_j}\right\rangle$ be the pure state that we get when we measure register R_j in $|\varphi\rangle$ and get outcome r_j.

Suppose that there exists a protocol $\mathcal{P}_0$ for G^k which wins all coordinates in $\mathcal{C}$ with probability greater than $2^{-\delta_2 k}$. Moreover, conditioning on success on all coordinates in $\mathcal{C}$, the probability it wins the game in the j-th coordinate is ω.

- Let us construct a new protocol $\mathcal{P}_1$, that starts with the joint state $\varphi^{X_j Y_j R_j E_A E_B}$, where $X_j E_A$ and $Y_j E_B$ are given to Alice and Bob, respectively, and R_j is shared between them. From our assumption, the probability that Alice and Bob win the game in the j-th coordinate is ω.
- Let us consider a new protocol $\mathcal{P}_2$, where Alice and Bob are given questions $(x_j, y_j) \leftarrow \varphi^{X_j Y_j}$ and they share $r_j \leftarrow \varphi^{R_j}_{x_j y_j}$ as public coins. By Lemma II.15, they are able to create a joint state that is close to the starting state of $\mathcal{P}_1$ by sharing $\left|\varphi_{r_j}\right\rangle$ and applying local unitary operations. More concretely, Eqs. (20) and (21) show the conditions for the mutual informations required by Lemma II.15. From Eq. (18), we can get

$$\begin{aligned}
10\delta_3 &\geq \mathop{\mathbb{E}}_{r_j \leftarrow \varphi^{R_j}} \left[\mathrm{S}\left(\varphi^{X_j Y_j}_{r_j} \middle\| \theta^{X_j Y_j} \right) \right] \\
&\geq \mathop{\mathbb{E}}_{r_j \leftarrow \varphi^{R_j}} \left[\mathrm{S}\left(\varphi^{X_j Y_j}_{r_j} \middle\| \varphi^{X_j}_{r_j} \otimes \varphi^{Y_j}_{r_j} \right) \right] \qquad (22) \\
&\geq \mathop{\mathbb{E}}_{r_j \leftarrow \varphi^{R_j}} \left[\left\| \varphi^{X_j Y_j}_{r_j} - \varphi^{X_j}_{r_j} \otimes \varphi^{Y_j}_{r_j} \right\|_1^2 \right] \qquad (23) \\
&\geq \left(\mathop{\mathbb{E}}_{r_j \leftarrow \varphi^{R_j}} \left[\left\| \varphi^{X_j Y_j}_{r_j} - \varphi^{X_j}_{r_j} \otimes \varphi^{Y_j}_{r_j} \right\|_1 \right] \right)^2 \qquad (24)
\end{aligned}$$

  where Eq. (22) follows from Fact II.9, Eq. (23) follows from Fact II.10, and at Eq. (24) we used the convexity of the function α^2. This implies

$$\mathop{\mathbb{E}}_{r_j \leftarrow \varphi^{R_j}} \left[\left\| \varphi^{X_j Y_j}_{r_j} - \varphi^{X_j}_{r_j} \otimes \varphi^{Y_j}_{r_j} \right\|_1 \right] \leq \sqrt{10\delta_3}.$$

  Thus, using the above and Lemma II.15, we conclude that they can win the game with probability at least $\omega - 10\sqrt{10\delta_3}$.
- Let us construct a new protocol $\mathcal{P}_3$, where Alice and Bob are given questions $(x_j, y_j) \leftarrow \varphi^{X_j Y_j}$. They share public coins $r_j \leftarrow \varphi^{R_j}$ and execute the same strategy as in $\mathcal{P}_2$. By Eq. (19), the probability that they win the game is at least $\omega - 11\sqrt{10\delta_3}$.
- Let us consider a new protocol $\mathcal{P}_4$, where Alice and Bob are given questions $(x, y) \leftarrow \mu$ and they execute the

same strategy as in $\mathcal{P}_3$. By Eq. (18) and Fact II.10, the probability that they win the game is at least $\omega - 12\sqrt{10\delta_3}$. Note that $\mathcal{P}_4$ is a strategy for game G under distribution μ. This means that $\omega - 12\sqrt{10\delta_3} \leq \omega^*(G)$.

We conclude the lemma. $\square$

We can now prove our main result. We restate it here for convenience.

Theorem I.2. *Let* $\varepsilon > 0$. *Given a game* G *with value* $\omega^*(G) \leq 1 - \varepsilon$, *it holds that*

$$\begin{aligned} \omega^*\left(G^k\right) &\leq \left(1 - \frac{\varepsilon}{2}\right)^{\frac{\varepsilon^2 k}{12000(\log|\mathcal{A}| + \log|\mathcal{B}|)}} \\ &= \left(1 - \varepsilon^3\right)^{\Omega\left(\frac{k}{\log|\mathcal{A}| + \log|\mathcal{B}|}\right)}. \end{aligned}$$

Proof. We set $\delta_1 = \frac{\varepsilon^2}{12000(\log|\mathcal{A}| + \log|\mathcal{B}|)}$, $\delta_2 = \frac{\varepsilon^2}{12000}$, and $\delta_3 = \frac{\varepsilon^2}{6000}$. Given any strategy for G^k, using Lemma III.2, either $\omega^*\left(G^k\right) \leq 2^{-\delta_2 k}$, or there are $\lfloor \delta_1 k \rfloor$ coordinates $\left\{i_1, \ldots, i_{\lfloor \delta_1 k \rfloor}\right\}$ such that the probability Alice and Bob win the i_j-th coordinate, conditioning on success on all the previous coordinates, is at most $1 - \varepsilon/2$. This finishes the proof of the theorem. $\square$

Acknowledgements

We thank Oded Regev, Thomas Vidick, and anonymous referees for useful comments. This work is supported by the Singapore Ministry of Education Tier 3 Grant and the Core Grants of the Center for Quantum Technologies, Singapore.

References

[1] R. Raz, "A parallel repetition theorem," in *Proceedings of the twenty-seventh annual ACM Symposium on Theory of Computing*, ser. STOC '95, New York, NY, USA, 1995, pp. 447–456. [Online]. Available: http://doi.acm.org/10.1145/225058.225181

[2] T. Holenstein, "Parallel repetition: simplifications and the no-signaling case," in *Proceedings of the thirty-ninth annual ACM Symposium on Theory of Computing*, ser. STOC '07, New York, NY, USA, 2007, pp. 411–419. [Online]. Available: http://doi.acm.org/10.1145/1250790.1250852

[3] R. Jain, J. Radhakrishnan, and P. Sen, "Optimal direct sum and privacy trade-off results for quantum and classical communication complexity," *CoRR*, vol. abs/0807.1267, 2008.

[4] M. Ben-Or, S. Goldwasser, J. Kilian, and A. Wigderson, "Multi-prover interactive proofs: How to remove intractability assumptions," in *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing*, ser. STOC '88, New York, NY, USA, 1988, pp. 113–131. [Online]. Available: http://doi.acm.org/10.1145/62212.62223

[5] L. Fortnow, "Complexity-theoretic aspects of interactive proof systems," Ph.D. dissertation, Massachusetts Institute of Technology, 1989.

[6] S. Arora and S. Safra, "Probabilistic checking of proofs: A new characterization of NP," *Journal of the ACM*, vol. 45, no. 1, pp. 70–122, Jan. 1998. [Online]. Available: http://doi.acm.org/10.1145/273865.273901

[7] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy, "Proof verification and the hardness of approximation problems," *Journal of the ACM*, vol. 45, no. 3, pp. 501–555, May 1998. [Online]. Available: http://doi.acm.org/10.1145/278298.278306

[8] I. Dinur, "The PCP theorem by gap amplification," *Journal of the ACM*, vol. 54, no. 3, Jun. 2007. [Online]. Available: http://doi.acm.org/10.1145/1236457.1236459

[9] A. Rao, "Parallel repetition in projection games and a concentration bound," in *Proceedings of the 40th annual ACM Symposium on Theory of Computing*, ser. STOC '08, New York, NY, USA, 2008, pp. 1–10. [Online]. Available: http://doi.acm.org/10.1145/1374376.1374378

[10] R. Raz, "A counterexample to strong parallel repetition," in *Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science*, Washington, DC, USA, 2008, pp. 369–373. [Online]. Available: http://portal.acm.org/citation.cfm?id=1470582.1470676

[11] B. Barak, A. Rao, R. Raz, R. Rosen, and R. Shaltiel, "Strong parallel repetition theorem for free projection games," in *Proceedings of the 12th International Workshop and 13th International Workshop on Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, ser. APPROX '09 / RANDOM '09. Berlin, Heidelberg: Springer-Verlag, 2009, pp. 352–365. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-03685-9_27

[12] R. Raz and R. Rosen, "A strong parallel repetition theorem for projection games on expanders," *2013 IEEE Conference on Computational Complexity*, vol. 0, pp. 247–257, 2012.

[13] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, "Proposed experiment to test local hidden-variable theories," *Phys. Rev. Lett.*, vol. 23, pp. 880–884, Oct. 1969. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevLett.23.880

[14] E. Hänggi and R. Renner, "Device-independent quantum key distribution with commuting measurements," *CoRR*, vol. abs/1009.1833, 2010.

[15] M. Tomamichel, S. Fehr, J. Kaniewski, and S. Wehner, "A monogamy-of-entanglement game with applications to device-independent quantum cryptography," *New Journal of Physics*, vol. 15, 2013.

[16] L. Masanes, S. Pironio, and A. Acín, "Secure device-independent quantum key distribution with causally independent measurement devices," *Nature Communications*, vol. 2, 2011.

[17] R. Cleve, W. Slofstra, F. Unger, and S. Upadhyay, "Perfect parallel repetition theorem for quantum xor proof systems," *Comput. Complex.*, vol. 17, no. 2, pp. 282–299, May 2008. [Online]. Available: http://dx.doi.org/10.1007/s00037-008-0250-4

[18] J. Kempe, O. Regev, and B. Toner, "Unique games with entangled provers are easy," *SIAM Journal on Computing*, vol. 39, pp. 3207–3229, Jul. 2010. [Online]. Available: http://dx.doi.org/10.1137/090772885

[19] I. Dinur, D. Steurer, and T. Vidick, "A parallel repetition theorem for entangled projection games," *CoRR*, vol. abs/1310.4113, 2013.

[20] I. Dinur and D. Steurer, "Analytical approach to parallel repetition," *CoRR*, vol. abs/1305.1979, 2013.

[21] J. Kempe and T. Vidick, "Parallel repetition of entangled games," in *Proceedings of the 43rd Annual ACM Symposium on Theory of Computing*, ser. STOC '11, New York, NY, USA, 2011, pp. 353–362. [Online]. Available: http://doi.acm.org/10.1145/1993636.1993684

[22] U. Feige and J. Kilian, "Two-prover protocols—low error at affordable rates," *SIAM Journal on Computing*, vol. 30, no. 1, pp. 324–346, Apr. 2000. [Online]. Available: http://dx.doi.org/10.1137/S0097539797325375

[23] A. Chailloux and G. Scarpa, "Parallel repetition of entangled games with exponential decay via the superposed information cost," *CoRR*, vol. abs/1310.7787, 2013.

[24] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, ser. Wiley Series in Telecommunications. New York, NY, USA: John Wiley & Sons, 1991.

[25] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge, UK: Cambridge University Press, 2000.

[26] J. Watrous, "Theory of Quantum Information, lecture notes," 2011. [Online]. Available: https://cs.uwaterloo.ca/~watrous/LectureNotes.html

[27] R. Jain, J. Radhakrishnan, and P. 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*, ser. FOCS '03, Washington, DC, USA, 2003, pp. 220–229. [Online]. Available: http://dl.acm.org/citation.cfm?id=946243.946331