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.pereszlényi@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)

\[
\omega^*(G^k) = (1 - (1 - \omega^*(G))^3)^{O\left(\frac{\log(1/\epsilon)}{\log(1/k)}\right)}
\]

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 \( \mu \) over \( \mathcal{X} \times \mathcal{Y} \), and a predicate \( V : \mathcal{X} \times \mathcal{Y} \times \mathcal{A} \times \mathcal{B} \rightarrow \{0,1\} \). It is played as follows.

- The referee selects questions \( (x,y) \in \mathcal{X} \times \mathcal{Y} \) according to distribution \( \mu \).
- 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 \( \mu \)) 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

\[
\omega(G^k) = (1 - (1 - \omega(G))^c)^{O\left(\frac{\log(1/\epsilon)}{\log(1/k)}\right)}
\]

where \( c \) is a universal constant. This result, along with 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 \textit{XOR games} by Cleve, Slofstra, Unger and
Upadhyay [17], later generalized to \textit{unique games} by Kempe,
Regev and Toner [18] and very recently further generalized
to \textit{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 1.1 (23).** For any game \(G = (\mu, \mathcal{X}, \mathcal{Y}, A, B, V)\),
where \(\mu\) is the uniform distribution on \(\mathcal{X} \times \mathcal{Y}\), it holds that
\[
\omega^*(G^k) = \left(1 - (1 - \omega^*(G))^2\right)^{\Omega\left(\frac{1}{k\log(1/|A||B||X||Y|)}\right)}.
\]
As a corollary, for a general distribution \(\mu\),
\[
\omega^*(G^k) = \left(1 - (1 - \omega^*(G))^2\right)^{\Omega\left(\frac{1}{\sqrt{Q}\log(1/|A||B|)}\right)}
\]
where
\[
Q = \max \left\{ \frac{1}{\min_{x,y: \mu(x,y) \neq 0} \left\{ \sqrt{\mu(x,y)} \right\}}, |\mathcal{X}| \cdot |\mathcal{Y}| \right\}.
\]
Note that here \(\omega^*(G^k)\) depends on \(|\mathcal{X}| \cdot |\mathcal{Y}|\) as well, in
addition to \(|A| \cdot |B|\) (as in Raz’s result). Also the value of \(Q\)
can be arbitrarily large, depending on the distribution \(\mu\).

Our result

In this paper we consider the case when the distribution \(\mu\) is
product across \(\mathcal{X} \times \mathcal{Y}\). That is, there are distributions \(\mu_X, \mu_Y\)
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 1.2 (Main Result).** For any game
\(G = (\mu, \mathcal{X}, \mathcal{Y}, A, B, V)\)
where \(\mu\) is a product distribution on \(\mathcal{X} \times \mathcal{Y}\), it holds that
\[
\omega^*(G^k) = \left(1 - (1 - \omega^*(G))^3\right)^{\Omega\left(\frac{1}{k\log(1/|A||B|)}\right)}.
\]
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 \(\mu\), 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
\[
\sigma^{XYAB} = \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} \tilde{\mu}(x, y) |xy\rangle\langle xy|^{XY} \otimes |\phi_{xy}\rangle^{AB}
\]
where \(\tilde{\mu}\) is a distribution, potentially different from \(\mu\) 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 \(\sigma^{XYAB}\).) In protocol \(\mathcal{P}\) for the single instance of \(G\),
we let Alice and Bob start with the shared pure state
\[
|\varphi\rangle = \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} \sqrt{\tilde{\mu}(x, y)} |xxyy\rangle^{XXYY} \otimes |\phi_{xy}\rangle^{AB}.
\]
Note that \(|\varphi\rangle\) is a purification of \(\sigma^{XYAB}\), where registers
\(X\) and \(Y\) 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 \(I(\tilde{X}_j : YYB)\) and \(I(\tilde{Y}_j : XXA)\) are very
small (close to 0), in \(|\varphi\rangle\). This, obviously, is only possible
when the distribution \(\mu\) is product. In addition, the distribution
of the questions in the \(j\)-th coordinate, in \(|\varphi\rangle\), remains close
to \(\mu\), in the \(\ell_1\)-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 $|\varphi_{x_j'y_j}'\rangle$ be the resulting state. If by luck it so happens that $(x', y') = (x_j', y_j')$, then they can measure the answer registers $A_j$ and $B_j$, in $|\varphi_{x_j'y_j}'\rangle$, respectively, and send the answers to the referee. However, the probability that $(x', y') = (x_j', y_j')$ can be very small and they want to get this desired outcome with probability very close to 1. We describe now how this can be achieved.

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

**Organization of the paper**

In Section \ref{section2} we present some background on information theory, as well as some useful lemmas that we will need for our proof. In Section \ref{section3} we prove our main result, Theorem \ref{Theorem2}.

\section{Preliminaries}

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

\subsection*{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 $\mu$ be a probability distribution on $\mathcal{X}$. Let $\mu(x)$ represent the probability of $x \in \mathcal{X}$ according to $\mu$. Let $X$ be a random variable distributed according to $\mu$. We use the symbol $x$ 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 $E_{x \sim \mathcal{X}}[f(x)] = \sum_{x \in \mathcal{X}} \mu[X = x] \cdot f(x)$, where $x \sim \mathcal{X}$ means that $x$ is drawn from the distribution of $X$. A quantum state (or just a state) $\rho$ 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 $\psi$ to represent the state and also the density matrix $|\psi\rangle\langle\psi|$, associated with $|\psi\rangle$. A classical distribution $\mu$ can be viewed as a quantum state with diagonal entries $\mu(x)$ and non-diagonal entries 0. For two quantum states $\rho$ and $\sigma$, $\rho \otimes \sigma$ represents the tensor product (Kronecker product) of $\rho$ and $\sigma$. 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 \cite{24, 25, 26} for more details.

**Definition II.1.** For quantum states $\rho$ and $\sigma$, the $\ell_1$-distance between them is given by $||\rho - \sigma||_{1}$, where $||X||_{1} \overset{\text{def}}{=} \text{Tr}\sqrt{X^\dagger X}$ is the sum of the singular values of $X$. We say that $\rho$ is $\varepsilon$-close to $\sigma$ if $||\rho - \sigma||_{1} \leq \varepsilon$.

**Definition II.2.** For quantum states $\rho$ and $\sigma$, the *fidelity* between them is given by $F(\rho, \sigma) \overset{\text{def}}{=} \sqrt{\text{Tr} \rho^\dagger \sigma}$.

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

**Proposition II.3** \cite{25}, pages 406 and 414. For states $\rho, \sigma$, and quantum operation $\mathcal{E}(\cdot)$, it holds that

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

and

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

The following proposition relates the $\ell_1$-distance and the fidelity between two states.

**Proposition II.4** \cite{25}, page 416. For quantum states $\rho$ and $\sigma$, it holds that

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

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

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

Let $\rho^{AB}$ 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

$$
\rho^B \overset{\text{def}}{=} \text{Tr}_A(\rho^{AB}) \overset{\text{def}}{=} \sum_i (|i\rangle \otimes I_B)\rho^{AB}(|i\rangle \otimes I_B)
$$

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

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

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

The **entropy** of a quantum state $\rho$ (in register $X$) is defined as $S(\rho) \overset{\text{def}}{=} -\text{Tr}\rho \log \rho$. We also let $S(X) \overset{\text{def}}{=} S(\rho)$. The **relative entropy** between quantum states $\rho$ and $\sigma$ is defined as $S(\rho||\sigma) \overset{\text{def}}{=} \text{Tr}\rho \log \rho - \text{Tr}\rho \log \sigma$. The **relative min-entropy** between them is defined as $S_\infty(\rho||\sigma) \overset{\text{def}}{=} \min \{\lambda : \rho \leq 2^\lambda \sigma\}$.

Since the logarithm is operator-monotone, $S(\rho||\sigma) \leq S_\infty(\rho||\sigma)$. 

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

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

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

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

where the *conditional entropy* is defined as

$$S(Y|X)_{\rho} \overset{\text{def}}{=} \mathbb{E}_{x \sim \rho} \left[ S(\rho_x) \right].$$

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

$$I(X : Z|Y)_{\rho} \overset{\text{def}}{=} \mathbb{E}_{y \sim Y} \left[ I(X : Z|Y = y)_{\rho} \right] = S(X|Y)_{\rho} + S(Z|Y)_{\rho} - S(XZ|Y)_{\rho}.$$ 

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

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

We will need the following basic facts.

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

$$S(\rho p + (1 - p) \rho^1 \| \sigma + (1 - p) \sigma^1) \leq p \cdot S(\rho \| \sigma) + (1 - p) \cdot S(\rho^1 \| \sigma^1).$$

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

**Fact II.8.** Let

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

and

$$\rho^1 = \sum_x \mu^1(x) |x\rangle\langle x| \otimes \rho_x^1.$$ 

It holds that

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

**Fact II.9.** For quantum states $\rho^{XY}$, $\sigma^X$, and $\tau^Y$, it holds that

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

**Fact II.10 (20, 27).** For quantum states $\rho$ and $\sigma$, it holds that

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

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

The following fact is easily verified.

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

**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 $\rho$. Let $a/01$ outcome measurement be performed on register $A$ with outcome $O$. Let $\Pr[O = 1] = \rho$. Let the marginal states on register $B$ conditioned on $O = 0$ and $O = 1$ be $\rho_0$ and $\rho_1$ respectively. Then, $S_{\infty}(\rho_1 \| \rho_0) \leq \log \frac{1}{q}$.

**Proof.** It is easily seen that $\rho = q \rho_1 + (1 - q) \rho_0$. Hence $S_{\infty}(\rho_1 \| \rho_0) \leq \log \frac{1}{q}$. □

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 $\mu$ be a probability distribution on $\mathcal{X}$. Let $|\varphi\rangle = \sum_{x \in \mathcal{X}} \sqrt{\mu(x)} |x\rangle^{XX} \otimes |\psi_x\rangle^{AB}$ be a joint pure state of Alice and Bob, where registers $XXA$ are with Alice and register $B$ is with Bob. Let $I(X : B)_{\rho} \leq \varepsilon$ and $|\varphi_x\rangle = |x\rangle \otimes |\psi_x\rangle$. There exist unitary operators $\{U_x\}_{x \in \mathcal{X}}$ acting on $XXA$ such that

$$\mathbb{E}_{x \sim \mu} \left[ S(\varphi_x \| \varphi_x) - (U_x \otimes 1_B) |\varphi\rangle \langle \varphi| (U_x^* \otimes 1_B) \right] \|_1 \leq 4 \sqrt{\varepsilon}.$$ 

**Proof.** Let us denote the reduced state of Bob in $|\varphi_x\rangle$ and $|\varphi\rangle$ by $\rho_x = \text{Tr}_A(|\psi_x\rangle \langle \psi_x|)$ and $\rho = \text{Tr}_{XXA}(|\varphi\rangle \langle \varphi|)$. Using Fact II.10, it holds that

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

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

$$|\varphi_x\rangle = (U_x \otimes 1_B) |\varphi\rangle \quad |\varphi_x\rangle = (U_x \otimes 1_B) |\varphi\rangle$$

The lemma follows from the following calculation.

$$\mathbb{E}_{x \sim \mu} \left[ S(\varphi_x \| \varphi_x) - (U_x \otimes 1_B) |\varphi\rangle \langle \varphi| (U_x^* \otimes 1_B) \right]$$

$$= 2 \mathbb{E}_{x \sim \mu} \left[ \sqrt{1 - |< \varphi_x | (U_x \otimes 1_B) |\varphi\rangle|^2} \right] \leq 2 \sqrt{1 - \mathbb{E}_{x \sim \mu} \left[ |< \varphi_x | (U_x \otimes 1_B) |\varphi\rangle|^2 \right] \right]$$

$$= 2 \sqrt{1 - \mathbb{E}_{x \sim \mu} \left[ F(\rho_x, \rho) \right]^2 $$

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

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].

\textbf{Lemma II.15 (E).} Let $\mu$ be a probability distribution over $\mathcal{X} \times \mathcal{Y}$. Let $\mu_X$ and $\mu_Y$ be the marginals of $\mu$ on $\mathcal{X}$ and $\mathcal{Y}$. Let

$$|\varphi\rangle \triangleq \sum_{x \in X, y \in Y} \sqrt{\mu(x, y)} |xxy\rangle \langle xy| \otimes |\psi_{x,y}\rangle^{AB}$$

be a joint pure state of Alice and Bob, where registers $\hat{X}A$ belong to Alice and registers $\hat{Y}B$ belong to Bob. Let $I(X : BY|Y) \leq \varepsilon$ and $I(Y : AX|X) \leq \varepsilon$.

Let $|\varphi_{x,y}\rangle \triangleq |xxy\rangle \otimes |\psi_{x,y}\rangle$. There exist unitary operators $\{U_x\}_{x \in X}$ on $\hat{X}A$ and $\{V_y\}_{y \in Y}$ on $\hat{Y}B$ such that

$$E_{\langle x,y \rangle \rightarrow \mu} \left[ \| |\varphi_{x,y}\rangle \langle \varphi_{x,y}| - (U_x \otimes V_y) |\varphi\rangle \langle \varphi| (U_x^* \otimes V_y^*) \|_1 \right] \leq 8\sqrt{\varepsilon} + 2 \| \mu - \mu_X \otimes \mu_Y \|_1.$$  

\textbf{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 $\{U_x\}_{x \in X}$ and $\{V_y\}_{y \in Y}$ such that

$$E_{x \rightarrow \mu_X} \left[ \| |\varphi_x\rangle \langle \varphi_x| - (I_A \otimes V_y) |\varphi\rangle \langle \varphi| (I_A \otimes V_y^*) \|_1 \right] \leq 4\sqrt{\varepsilon},$$

and

$$E_{y \rightarrow \mu_Y} \left[ \| |\varphi_y\rangle \langle \varphi_y| - (U_x \otimes I_B) |\varphi\rangle \langle \varphi| (U_x^* \otimes I_B) \|_1 \right] \leq 4\sqrt{\varepsilon}.$$  

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 $U_x$ doesn’t change the $\ell_1$-distance, and the first term in Eq. (3) follows from Proposition II.3 with the superoperator that corresponds to measuring $Y$ in the standard basis and storing the outcome in a new register. The lemma follows from the following calculation.

$$E_{\langle x,y \rangle \rightarrow \mu} \left[ \| |\varphi_{x,y}\rangle \langle \varphi_{x,y}| - (U_x \otimes V_y) |\varphi\rangle \langle \varphi| (U_x^* \otimes V_y^*) \|_1 \right]$$

\begin{align*}
&= \left[ E_{\langle x,y \rangle \rightarrow \mu} \left[ |xy\rangle \langle xy| \otimes |\varphi_{x,y}\rangle \langle \varphi_{x,y}| - (U_x \otimes V_y) |\varphi\rangle \langle \varphi| (U_x^* \otimes V_y^*) \right] \right]_{1} \\
&\leq \left[ E_{\langle x,y \rangle \rightarrow \mu} \left[ |xy\rangle \langle xy| \otimes |\varphi_{x,y}\rangle \langle \varphi_{x,y}| - (U_x \otimes V_y) |\varphi\rangle \langle \varphi| (U_x^* \otimes V_y^*) \right] \right]_{1} \\
&\leq \left[ \| \mu - \mu_X \otimes \mu_Y \|_1 \right].
\end{align*}  

where the first inequality follows from the triangle inequality and at the last inequality we used Eq. (4) and Fact II.12. ☐

III. PROOF OF THE MAIN RESULT

Let a game $G = (\mu, \mathcal{X}, \mathcal{Y}, A, B, V)$ be given. We assume that the distribution $\mu = \mu_X \otimes \mu_Y$ is product across $\mathcal{X}$ and
\(Y\). Before the game starts, Alice and Bob share a pure state on the registers \(A_E B E_B^r\), 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 \(x = x_1 \ldots x_k \in X^k\), \(y = y_1 \ldots y_k \in Y^k\), \(a = a_1 \ldots a_k \in A^k\), and \(b = b_1 \ldots b_k \in B^k\). 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 \(\mathcal{C}\) represent its complement in \([k]\). Let \(x_C\) represent the substring of \(x\) corresponding to the indices in \(\mathcal{C}\). (Similarly, we will use \(y_C\).)

Let \(E_A = E_A|C\), \(E_B = E_B|C\), and \(a_c b_c \in \mathcal{AC} \otimes |\gamma_{xyacb_c}\rangle\) is the shared state after Alice and Bob performed their unitary operations corresponding to questions \(x\) and \(y\). (Note that \(|\gamma_{xyacb_c}\rangle\) is unnormalized.) Consider the state

\[
|\varphi\rangle = \frac{1}{\sqrt{q}} \sum_{x,y} \sqrt{\mu(x, y)} |xxyy\rangle^X \overline{Y} Y \\
\otimes \sum_{a_c b_c : V(x_C, y_C, a_c, b_c) = 1} |a_c b_c\rangle^A B_c \otimes |\gamma_{xyacb_c}\rangle^E A E_B
\]

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

**Lemma III.1.**

\[
\mathbb{E}_{x_C y_C a_c b_c \sim \varphi x_Y c a \mathcal{C}} \left[ S\left( |\varphi_{x_C Y c a_c b_c}\rangle\langle \varphi_{x_C Y c a_c b_c}| \right) \right] \\
\leq -\log q + |C| \cdot \log(\|A\| \cdot |B|) \]  

\(\mathbb{E}_{x_C y_C a_c b_c \sim \varphi x_Y c a \mathcal{C}} \left[ S\left( |\varphi_{x_C Y c a_c b_c}\rangle\langle \varphi_{x_C Y c a_c b_c}| \right) \right] \\
\leq -\log q + |C| \cdot \log(\|A\| \cdot |B|) \]

**Proof.** Note that, by Lemma [II.13]

\[
S(\varphi) = S(\varphi_{x_C Y c a_c b_c}) \\
\leq -\log q
\]

Let \(p(a_c, b_c)\) be the probability of obtaining \((a_c, b_c)\) when measuring registers \((A_c, B_c)\) in \(|\varphi\rangle\). Consider,

\[
\mathbb{E}_{a_c b_c \sim \varphi a_c b_c} \left[ S(\varphi_{a_c b_c}) \right] \\
\leq \mathbb{E}_{a_c b_c \sim \varphi a_c b_c} \left[ S(\varphi_{a_c b_c}) \right] \\
= -\log q + S(\varphi_{a_c b_c}) \]

Now,

\[
-\log q + |C| \cdot \log(\|A\| \cdot |B|) \\
\geq \mathbb{E}_{a_c b_c \sim \varphi a_c b_c} \left[ S(\varphi_{a_c b_c}) \right] \\
\geq \mathbb{E}_{a_c b_c \sim \varphi a_c b_c} \left[ S(\varphi_{a_c b_c}) \right] \\
\geq \mathbb{E}_{x_C y_C a_c b_c \sim \varphi x_Y c a_c b_c} \left[ S(\varphi_{x_C Y c a_c b_c}) \right] \\
\geq \mathbb{E}_{x_C y_C a_c b_c \sim \varphi x_Y c a_c b_c} \left[ S(\varphi_{x_C Y c a_c b_c}) \right]
\]

where the last inequality follows from Fact [II.3]

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 \(0 > \delta_1, \delta_2, \delta_3 > 0\) such that \(\delta_3 = \delta_2 + \delta_1 \cdot \log(\|A\| \cdot |B|)\). Let \(k' = \lceil \delta_2 k \rceil\). For any quantum strategy for the \(k\)-fold game \(G^k\), there exists a set \(\{i_1, \ldots, i_k\}\), such that for each \(1 \leq r \leq k'\), either

\[
\Pr[T(r) = 1] \leq 2^{-\delta_k k}
\]

or

\[
\Pr[T_{i_1} = 1 | T(r) = 1] \leq \omega'(G) + 12 \sqrt{10 \delta_3}
\]

where \(T(r) \equiv \prod_{j=1}^r T_{i_j}\).

**Proof.** In the following, we assume that \(1 \leq r \leq k'\). 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 \(i_1, \ldots, i_r\) satisfying that

\[
\Pr[T_{i_1} = 1] \leq \omega'(G) + 12 \sqrt{10 \delta_3}
\]

and

\[
\Pr[T_{i_1} = 1 | T(j) = 1] \leq \omega'(G) + 12 \sqrt{10 \delta_3}
\]

for \(1 \leq j \leq r - 1\). If \(\Pr[T(r) = 1] \leq 2^{-\delta_k k}\) then we are done, so from now on, we assume that \(\Pr[T(r) = 1] > 2^{-\delta_k k}\).

Let \(\mathcal{C} \equiv \{i_1, \ldots, i_r\}\). To simplify notations, let \(\hat{A} \equiv \hat{X}_{i_1} \cdots \hat{X}_{i_r} E_A\), \(\hat{B} \equiv \hat{Y}_{i_1} \cdots \hat{Y}_{i_r} E_B\), and \(R_I \equiv X_{i_1} \cdots X_{i_r} Y_{i_1} \cdots Y_{i_r} A_c B_c\). For coordinate \(i\), let \(|\phi_{x_C Y c a_c b_c}\rangle\) be the pure state that results when we measure registers \(X_{i} Y_{i} \) (i.e., registers \(X_1, \ldots, X_{i-1}, Y_1, \ldots, Y_{i-1}\) in \(|\phi\rangle\) 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 \(\mu\) in the state \(\phi\). We also prove that, for this coordinate, the questions and \(R_I\) are almost
independent. From Lemma \[\text{III.1}\] we get that
\[
\delta_3 k \geq \delta_2 k + |C| \cdot \log(|A| \cdot |B|)
\]
\[
\geq \mathbb{E}_{x \in \mathcal{X}, y \in \mathcal{Y}, \theta \in \mathcal{C}} \left[ S\left( \theta_{xyc}^{xy} \left\| \theta_{xyc}^{xy} \right\| \right) \right] \geq 1 \left( X : \hat{B} \bigg| R_1 \right) \varphi
\]
where \( \delta_3 \) follows from Lemma \[\text{III.1}\] and the fact that \( \theta_{xyc}^{xy} = \theta_{xyc}^{xy} \otimes \hat{\theta}_{xyc}^{xy} \). Equations \[\text{15}\] and \[\text{16}\] follow from the chain rule for the mutual information and at Eq. \[\text{16}\] we also used the observation that \( \hat{B} \) contains register \( Y \). Similarly to the above, for Bob’s questions we have
\[
\delta_3 k \geq \sum_{i \in \mathcal{C}} I \left( Y_i : \hat{A} \big| R_1 \right) \varphi
\]
From Eqs. \[\text{8}\], \[\text{13}\], \[\text{16}\] and \[\text{17}\] and using standard application of Markov’s inequality, we get that there exists a coordinate \( j \not\in C \) such that
\[
\mathbb{E}_{r_j \sim \varphi_{r_j}} \left[ S\left( \varphi_{r_j}^{Y_j} \bigg| \varphi_{X_j}^{Y_j} \right) \right] \leq \frac{5\delta_3}{1 - \delta_1} \leq 10\delta_3
\]
\[
\mathbb{E}_{r_j \sim \varphi_{r_j}} \left[ \varphi_{r_j}^{Y_j} \right] \left\| \varphi_{r_j}^{Y_j} - \varphi_{X_j}^{Y_j} \right\| \leq \sqrt{\frac{5\delta_3}{1 - \delta_1}} \leq \sqrt{10\delta_3}
\]
\[
I \left( X_j : \hat{B} \bigg| R_j \right) \varphi \leq \frac{5\delta_3}{1 - \delta_1} \leq 10\delta_3
\]
\[
I \left( Y_j : \hat{A} \bigg| R_j \right) \varphi \leq \frac{5\delta_3}{1 - \delta_1} \leq 10\delta_3
\]
Let \( \varphi_{r_j}^{Y_j} \) be the pure state that we get when we measure register \( R_j \) in \( \varphi \) and get outcome \( r_j \).
Suppose that there exists a protocol \( \mathcal{P}_1 \) for \( C^k \) which wins all coordinates in \( C \) with probability greater than \( 2^{-b_2k} \). Moreover, conditioning on success on all coordinates in \( C \), the probability it wins the game in the \( j \)-th coordinate is \( \omega \).

- Let us construct a new protocol \( \mathcal{P}_1 \), that starts with the joint state \( \varphi_{x_i}^{Y_i} \mathcal{E}_A \mathcal{E}_B \), where \( X_j \mathcal{E}_A \) and \( Y_j \mathcal{E}_B \) 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 \( \omega \).

- Let us consider a new protocol \( \mathcal{P}_2 \), where Alice and Bob are given questions \( (x_j, y_j) \sim \varphi_{x_j}^{Y_j} \) and they share \( r_j \sim \varphi_{r_j}^{Y_j} \) as public coins. By Lemma \[\text{II.15}\] they are able to create a joint state that is close to the starting state of \( \mathcal{P}_1 \) by sharing \( \varphi_{r_j}^{Y_j} \) and applying local unitary operations. More concretely, Eqs. \[\text{20}\] and \[\text{21}\] show the conditions for the mutual informations required by Lemma \[\text{II.15}\]. From Eq. \[\text{18}\], we can get
\[
10\delta_3 \geq \mathbb{E}_{r_j \sim \varphi_{r_j}} \left[ S\left( \varphi_{r_j}^{Y_j} \bigg| \varphi_{X_j}^{Y_j} \right) \right]
\]
where Eq. \[\text{14}\] follows from Fact \[\text{II.7}\] and at Eq. \[\text{16}\] we also used the convexity of the function \( \alpha^2 \). This implies
\[
\mathbb{E}_{r_j \sim \varphi_{r_j}} \left[ \left\| \varphi_{r_j}^{Y_j} - \varphi_{r_j}^{Y_j} \right\| \right] \leq \sqrt{10\delta_3}
\]

\[
\text{Thus, using the above and Lemma \[\text{II.15}\], 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 \( (x, y) \sim \varphi_{x}^{Y_j} \). They share public coins \( r_j \sim \varphi_{r_j}^{Y_j} \) and execute the same strategy as in \( \mathcal{P}_2 \).
By Eq. \[\text{19}\], 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 \( (x, y) \sim \mu \) and they execute the
same strategy as in $\rho_3$. By Eq. (18) and Fact II.10, the probability that they win the game is at least $\sqrt{\omega - 12\sqrt{10}\delta_3}$. Note that $\rho_3$ is a strategy for game $G$ under distribution $\mu$. This means that $\sqrt{\omega - 12\sqrt{10}\delta_3} \leq \omega^*(G)$.

We conclude the lemma. □

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

**Theorem 1.2** Let $\varepsilon > 0$. Given a game $G$ with value $\omega^*(G) \leq 1 - \varepsilon$, it holds that

$$
\omega^*(G^k) \leq \left(1 - \frac{\varepsilon}{2}\right) \left(\frac{2^{12000(\log |A| + \log |B|)}}{2^{12000(\log |A| + \log |B|)}} \right).
$$

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

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


