Parallel Repetition for Entangled k-player Games via Fast Quantum Search

Kai-Min Chung[1], Xiaodi Wu[2], and Henry Yuen[3]

1   Academia Sinica, Institute of Information Science
     128 Academia Road, Nankang, Taipei 11529, Taiwan
     kmchung@iis.sinica.edu.tw
2   MIT Center for Theoretical Physics
     77 Massachusetts Ave, Cambridge MA 02139, USA
     xiaodiwu@mit.edu
3   MIT CSAIL
     32 Vassar St, Cambridge MA 02139, USA
     hyuen@mit.edu

─── **Abstract** ───

We present two parallel repetition theorems for the entangled value of multi-player, one-round free games (games where the inputs come from a product distribution). Our first theorem shows that for a k-player free game G with entangled value $\mathrm{val}^*(G) = 1 - \epsilon$, the n-fold repetition of G has entangled value $\mathrm{val}^*(G^{\otimes n})$ at most $(1 - \epsilon^{3/2})^{\Omega(n/sk^4)}$, where s is the answer length of any player. In contrast, the best known parallel repetition theorem for the *classical* value of two-player free games is $\mathrm{val}(G^{\otimes n}) \leq (1 - \epsilon^2)^{\Omega(n/s)}$, due to Barak, et al. (RANDOM 2009). This suggests the possibility of a separation between the behavior of entangled and classical free games under parallel repetition.

Our second theorem handles the broader class of free games G where the players can output (possibly entangled) quantum states. For such games, the repeated entangled value is upper bounded by $(1-\epsilon^2)^{\Omega(n/sk^2)}$. We also show that the dependence of the exponent on k is necessary: we exhibit a k-player free game G and $n \geq 1$ such that $\mathrm{val}^*(G^{\otimes n}) \geq \mathrm{val}^*(G)^{n/k}$.

Our analysis exploits the novel connection between communication protocols and quantum parallel repetition, first explored by Chailloux and Scarpa (ICALP 2014). We demonstrate that better communication protocols yield better parallel repetition theorems: in particular, our first theorem crucially uses a quantum search protocol by Aaronson and Ambainis, which gives a quadratic Grover speed-up for distributed search problems. Finally, our results apply to a broader class of games than were previously considered before; in particular, we obtain the first parallel repetition theorem for entangled games involving more than two players, and for games involving quantum outputs.

1   Introduction

The study of multi-player one-round games has been central to both theoretical computer science and quantum information. Games have served as an indispensible tool with which to study a diverse array of topics, from the hardness of approximation to cryptography; from delegated computation to Bell inequalities; from proof systems to the monogamy of

entanglement. In particular, two-player games have received the most scrutiny. In a two-player game G, a referee samples a pair of questions (x, y) from some distribution μ, and sends question x to one player (typically named Alice), and y to the other (typically named Bob). Alice and Bob then utilize some non-communicating strategy to produce answers a and b, respectively, upon which the referee computes some predicate $V(x, y, a, b)$ to decide whether to accept or not. In this paper, we focus on the setting where Alice and Bob may utilize quantum entanglement as part of their strategy. The primary quantity of interest is the *entangled value* val*(G) of game G, which is the maximum success probability over all possible entangled strategies for the players.

Recently, there has been significant interest in the *parallel repetition* of entangled games [15, 6, 7, 13, 9]. More formally, the n-fold parallel repetition of a game G is a game $G^{\otimes n}$ where the referee will sample n independent pairs of questions $(x_1, y_1), \ldots, (x_n, y_n)$ from the distribution μ. Alice receives $(x_1, \ldots, x_n)$ and Bob receives $(y_1, \ldots, y_n)$. They produce outputs $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$, respectively, and they win only if $V(x_i, y_i, a_i, b_i) = 1$ for all i. We call each i a "coordinate" of $G^{\otimes n}$ or "repetition" of G.

Suppose we have a game G where val*$(G) = 1 - \epsilon$. Intuitively, one should expect that val*$(G^{\otimes n})$ should behave as $(1 - \epsilon)^n$. Indeed, this would be the case if the game G were played n times *sequentially*. However, there are counterexamples of games G and $n > 1$ where val*$(G^{\otimes n}) = $ val*(G) (e.g., as in [8]). Despite such counterexamples, it has been shown that the *classical* value val$(G^{\otimes n})$ (i.e. where the players are restricted to using classical strategies) of a repeated game $G^{\otimes n}$ goes down exponentially with n, for large enough n [19, 12]. This result is known as the Parallel Repetition Theorem, and is central in the study of hardness of approximation, probabilistically checkable proofs, and hardness amplification in classical theoretical computer science.

Recently, quantum analogues of the Parallel Repetition Theorem have been studied, and for certain types of games, it has been shown that the entangled game value also goes down exponentially with the number of repetitions. In particular, parallel repetition theorems have been shown for 2-player *free games* (see [6, 7, 13]) and *projection games* (see [9]). Free games are where the input distribution to the players is a product distribution (i.e. each players' questions are chosen independently of each other). Projection games are where, for each answer of one designated player, there is at most one other answer for the other player that the referee would accept.

Most relevant to this work are the results of [6, 7, 13] on free entangled games. Among them, the best parallel repetition theorem was obtained by [7], who prove that for a two-player free game G, the entangled value of the n-fold repetition is at most $(1 - \epsilon^2)^{\Omega(n/s)}$, where s is the answer length of the players. When G is also a projection game, they obtain *strong parallel repetition*: the repeated game value is at most $(1 - \epsilon)^{\Omega(n)}$. The centerpiece of their analysis is a novel connection between communication complexity and parallel repetition of games.

1.1 Our results

In this work, we further develop this connection between games and communication protocols to obtain improved parallel repetition theorems for free entangled games. We present a generic framework where one obtains parallel repetition theorems for free games by designing succinct communication protocols. The core concept we present is:

> *Better parallel repetition theorems from better communication protocols.*

The first instantiation of this concept is the following theorem:

▶ **Theorem 1.1.** *Let $k \geq 2$ be an integer. Let G be a k-player free game with entangled value* $\mathrm{val}^*(G) = 1 - \epsilon$. *Then, for* $n = \Omega(sk^4 \log(k/\epsilon)/\epsilon^{3/2})$,

$$\mathrm{val}^*(G^{\otimes n}) \leq (1 - \epsilon^{3/2})^{\Omega(n/k^4 s)}$$

where s is the output answer length of the players.

The proof of Theorem 1.1 uses a quantum communication protocol that performs a version of distributed unstructured search (i.e. searching for a 1 in a bitstring). The improvement of the base from $1 - \epsilon^2$ (as found in [7]) to $1 - \epsilon^{3/2}$ comes from the fact that the unstructured search problem on N bits can be solved by a quantum algorithm using only $O(\sqrt{N})$ queries. We discuss this in more detail in the next section.

Our second theorem handles a broader class of games, where the players can output quantum states as answers. We are able to handle this broader class of games because our framework allows general quantum communication protocols.

▶ **Theorem 1.2.** *Let $k \geq 2$ be an integer. Let G be a k-player free game, where players can output (possibly entangled) quantum states, and has entangled value* $\mathrm{val}^*(G) = 1 - \epsilon$. *Then, for all n,*

$$\mathrm{val}^*(G^{\otimes n}) \leq (1 - \epsilon^2)^{\Omega(n/k^2 s)}$$

where $s = \max_j \log(d_j)$, where d_j is the dimension of player j's output state.

Furthermore, we prove that the dependence of the exponent on the number of players k is necessary:

▶ **Theorem 1.3.** *For all $k \geq 2$, there exists a k-player free game G and $n > 1$ where* $\mathrm{val}^*(G^{\otimes n}) \geq \mathrm{val}^*(G)^{n/k}$.

To our knowledge, our results are the first to show quantum parallel repetition in the setting of games with more than 2 players.

Finally, we give a proof of parallel repetition for the *classical* value of k-player free games. While this theorem appears to be a folklore result, we were not able to find any explicit proof of it. We provide one here for the sake of completeness.

▶ **Theorem 1.4.** *Let G be a k-player free game with classical value* $\mathrm{val}(G) = 1 - \epsilon$. *Then*

$$\mathrm{val}(G^{\otimes n}) \leq (1 - \epsilon^2)^{\Omega(n/sk)},$$

where s is the output answer length of the players.

CQ Games. Our second theorem applies to a class of games that is a generalization of the traditional notion of games that involve two players and have classical inputs and outputs. In this paper we introduce the class of k-player *classical-quantum* (CQ) games, where the players receive classical inputs, apply local unitary operators to their share of an entangled state, and return some qubits to the referee. The referee then makes a measurement on the answer qubits to decide whether to accept or reject. If we restrict the players' unitaries to be permutation matrices, and the referee's measurement to be diagonal in the standard basis, then we recover the class of classical games.

We believe the model of CQ games is worth deeper investigation. One motivation for the study of CQ games comes from the recent exciting work of Fitzsimons and Vidick [11], who demonstrated an efficient reduction transforming a local Hamiltonian $H = H_1 + \cdots + H_m$ acting on n qubits to a 5-player CQ-game G_H such that approximating $\mathrm{val}^*(G_H)$ with inverse

polynomial accuracy will decide whether the ground state energy of H is a YES or NO instance of the QMA-complete problem LOCAL HAMILTONIANS. In this game, the referee sends $O(\log n)$-sized questions, and the players responds with $O(1)$-qubit states as answers. The significance of this is that it opens up the possibility of proving a "games" version of the Quantum PCP conjecture. This intriguing possibility calls for further study of the behavior of CQ games.

1.2 Parallel repetition and communication protocols

At a high level, most proofs of parallel repetition proceed via reduction. Let G be a two-player free game with verification predicate $V(x, y, a, b)$. If there were a strategy $\mathcal{S}$ for the repeated game $G^{\otimes n}$ that wins with too large probability, then one can transform $\mathcal{S}$ to a strategy $\mathcal{T}$ to play a single instance of the game G with probability larger than $\mathrm{val}^*(G)$, which would be a contradiction.

In [6, 7, 13], the reduction from a repeated game strategy to a single game strategy has two steps: (1) a "too-good" repeated game strategy $\mathcal{S}$ is converted to an *advice-based strategy* for game G, which wins with high probability. An advice-based strategy is a collection of *advice states* $\{\varphi_{xy}\}_{xy}$ so that when Alice and Bob receive inputs x and y, they happen to share the entangled state φ_{xy}, which they can measure to produce answers. Of course, this is not a valid quantum strategy for game G, but (2) using the assumption that $\mathcal{S}$ has very high winning probability, the advice-based strategy can be *rounded* to a true game strategy: Alice and Bob can apply local operations U_x and V_y, respectively, on some input-independent state φ to approximate φ_{xy}, and thus simulate the advice-based strategy (with some error).

One can construct the advice states $\{\varphi_{xy}\}_{xy}$ from $\mathcal{S}$ in different ways. Generally, the goal is to create advice states that closely mimic the joint state of the players during an actual execution of the strategy $\mathcal{S}$, *conditioned* on some event. Ideally, we would like to condition on the event that the players won all n coordinates of $G^{\otimes n}$. This would give rise to the ideal advice states: whenever the players receive an input x and y, their advice state φ_{xy} would have precisely the correct answers a and b that would allow them to win the single game G.

However, it seems impossible to argue that such an ideal advice-based strategy can be simulated by a true game strategy. The approach taken by [6, 7] is to construct advice states $\{\varphi_{xy}\}$ that have two properties: (a) the advice-based strategy succeeds with high probability for G, and (b) there is a *low-cost communication protocol* between Alice and Bob that produces φ_{xy} when they receive inputs x and y, respectively. Property (b) makes it possible to approximate the advice-based strategy using a valid quantum strategy: small communication complexity translates into small rounding error.

The communication protocol used in [6, 7] is a simple one: Alice and Bob first play the optimal strategy $\mathcal{S}$ for $G^{\otimes n}$. They receive inputs $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$, and measure a shared state $|\xi\rangle$ and obtain n-tuples of outputs $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$. Then, Alice samples a small subset of coordinates $i_1, \ldots, i_h$, and sends over her inputs and outputs in this subset to Bob, who verifies that the original game G was won in each of these coordinates. If Bob finds a i_j such that the tuple $V(x_{i_j}, y_{i_j}, a_{i_j}, b_{i_j}) = 0$, then Bob aborts. Otherwise, Bob accepts. If we condition the final state of the protocol on Bob accepting, then we have a very good proxy for the ideal advice states described above. But since the communication complexity of this protocol is small, we can round this to a valid quantum strategy.

However, the analysis of [6, 7] is tailored to simple one-way communication protocols involving classical messages. We generalize this paradigm to show that, if the advice states $\{\varphi_{xy}\}$ can be constructed using *any* communication protocol (which can be two-way, and involve quantum messages), then the advice-based strategy using $\{\varphi_{xy}\}$ can be simulated

with a true game strategy, with error that is related to the communication complexity of the protocol. This unlocks a richer toolbox for the reduction designer: one can use many more tools from communication complexity to engineer good advice states. This gives rise to the concept of "Better parallel repetition theorems from better communication protocols."

Our theorems are instantiations of this mantra. At the heart of the communication protocol used in our first theorem is a variant of the Grover search algorithm. There, the players sample a random subset of coordinates $i_1, \ldots, i_h$ as before, but now they perform quantum search over the indices to find a "losing coordinate": i.e., a coordinate i_j such that $V(x_{i_j}, y_{i_j}, a_{i_j}, b_{i_j}) = 0$. The quadratic speedup of Grover's search algorithm translates into a quadratic savings in communication complexity, which is precisely what allows us to improve the base of the repeated game value from $1 - \epsilon^2$ to $1 - \epsilon^{3/2}$. For our second theorem, we take advantage of the fact that the communication protocol can be quantum, which allows us to handle games with quantum outputs.

Our use of quantum search in the protocol to generate the advice states gives a generic way to improve the reduction for arbitrary free games. However, one could also use this technique to prove *game-specific* parallel repetition theorems. That is, one could try to leverage special properties of a particular game to design a succinct communication protocol for generating advice states, and in turn, obtain a parallel repetition theorem with better parameters. Indeed, one can see this idea in the result of [7] for free projection games: by using the projection property of the game, their communication protocol avoids sending whole input and output symbols. This allows them to prove a repeated game value of $(1 - \epsilon)^{\Omega(n)}$ – note that this does not depend on the output alphabet!

1.3   Related work

We discuss how our result relates to prior results in parallel repetition, classical and quantum. Most relevant to our work are the results on free games. Jain, et al. [13] and Chailloux and Scarpa [6, 7] both proved that the entangled value of 2-player free games (with classical inputs and outputs) goes down exponentially with the number of repetitions. In particular, [7] showed for such a game G with $\mathrm{val}^*(G) = 1 - \epsilon$, we have that $\mathrm{val}^*(G^{\otimes n}) \leq (1 - \epsilon^2)^{\Omega(n/s)}$, where s is the output length of the players. They also show that, when G is also a projection game, *strong* parallel repetition holds: $\mathrm{val}^*(G^{\otimes n}) \leq (1 - \epsilon)^{\Omega(n)}$.

In a different line of work, Dinur, Steurer and Vidick show that projection games (with an arbitrary input distribution) also have an exponential decay in entangled value under parallel repetition: if G be a 2-player projection game with classical inputs and outputs, and $\mathrm{val}^*(G) = 1 - \epsilon$, then $\mathrm{val}^*(G^{\otimes n}) \leq (1 - \epsilon^{12})^{\Omega(n)}$ [9]. This result is not comparable with our work, nor with the work of [7, 13]. While [9] can handle games with arbitrary input distributions, the games need to satisfy the projection property. On the other hand, the results on free games can handle arbitrary verification predicates, but the input distributions need to be product.

There is a rich history of study of parallel repetition in classical theoretical computer science, which we will not detail here. Most relevant to us is the work of Barak, et al. [3], who showed that for 2-player free games G with classical value $1 - \epsilon$, $\mathrm{val}(G^{\otimes n}) \leq (1 - \epsilon^2)^{\Omega(n/s)}$, where s is the output length of the players. Intriguingly, it is not known whether the ϵ^2 term is tight for free games (it is known that this is necessary for classical parallel repetition of general games [20]). Our first theorem demonstrates a possible separation between classical and quantum parallel repetition; the base of our repeated game value is $1 - \epsilon^{3/2}$, rather than $1 - \epsilon^2$.

Finally, there has been little prior study of the parallel repetition of games with more

than 2 players. Buhrman et al. studied this question for non-signaling players, and showed that the non-signaling value of repeated games goes down exponentially with the number of repetitions [5]. Their parallel repetition theorem holds for games with *full support*, meaning that every possible combination of questions gets asked with positive probability; furthermore, the rate of decay also depends on the complete description of the game, not just the original game value and the number of repetitions. Arnon-Friedman et al. prove similar results for multi-player non-signaling games, but they use a new technique called de Finetti reductions [2]. Rosen also studied k-player parallel repetition in a weaker version of the non-signaling model, and demonstrated an exponential rate of decay [21].

2 Proof overviews

2.1 Overview of Theorem 1.1

Here we give a very informal outline of the proof of Theorem 1.1, for the case of two-player free games. The full proof that handles an arbitrary number of players can be found in Section 5 of the Appendix.

Let G be a two-player free game, with inputs (x, y) drawn from a product distribution $\mu = \mu_X \otimes \mu_Y$, and with verification predicate $V(x, y, a, b)$. Let $\mathcal{S}$ be an optimal strategy for the repeated game $G^{\otimes n}$, where Alice and Bob share an entangled state $|\xi\rangle$, and upon receiving a tuple of inputs $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$, Alice and Bob perform local measurements $M^{\mathbf{x}} = \{M_{\mathbf{a}}^{\mathbf{x}}\}_{\mathbf{a}}$ and $N^{\mathbf{y}} = \{N_{\mathbf{b}}^{\mathbf{y}}\}_{\mathbf{b}}$ on their respective parts of $|\xi\rangle$, to obtain answer tuples $\mathbf{a} = (a_1, \ldots, a_n)$ and $\mathbf{b} = (b_1, \ldots, b_n)$. Let $|\xi_{\mathbf{xyab}}\rangle$ be the (unnormalized) post-measurement state of $|\xi\rangle$ after making measurements $(M^{\mathbf{x}}, N^{\mathbf{y}})$, and obtaining outcomes $(\mathbf{a}, \mathbf{b})$. We assume for contradiction that $\mathrm{val}^*(G^{\otimes n}) > 2^{-\gamma n}$, for some small γ.

A naive approach. Consider the following state

$$|\theta\rangle = \frac{1}{\sqrt{\lambda}} \sum_{\mathbf{x}, \mathbf{y}} \sqrt{\mu^{\otimes n}(\mathbf{x}, \mathbf{y})} \, |\mathbf{x}\rangle \otimes |\mathbf{y}\rangle \otimes \sum_{\mathbf{a}, \mathbf{b}: V(\mathbf{x}, \mathbf{y}, \mathbf{a}, \mathbf{b}) = 1} |\xi_{\mathbf{xyab}}\rangle \otimes |\mathbf{a}\rangle \otimes |\mathbf{b}\rangle$$

where λ is a normalizing constant, $\mu^{\otimes n}$ is the input distribution for $G^{\otimes n}$, and $V(\mathbf{x}, \mathbf{y}, \mathbf{a}, \mathbf{b}) = \prod_i V(\mathbf{x}_i, \mathbf{y}_i, \mathbf{a}_i, \mathbf{b}_i)$. This would be the joint state of Alice and Bob if they received inputs $\mathbf{x}$ and $\mathbf{y}$ in coherent superposition, played strategy $\mathcal{S}$, and won the game. Here is a naive idea to use $|\theta\rangle$ in a strategy $\mathcal{T}$ for G: Alice and Bob share $|\theta\rangle$, with Alice possessing the $|\mathbf{x}\rangle$ input register, the $|\mathbf{a}\rangle$ output register, and half of $|\xi_{\mathbf{xyab}}\rangle$; Bob possesses the $|\mathbf{y}\rangle$ input register, the $|\mathbf{b}\rangle$ output register, and the other half of $|\xi_{\mathbf{xyab}}\rangle$. When they receive inputs $(x, y) \leftarrow \mu$, they each measure some fixed coordinate i of their respective input registers of $|\theta\rangle$ to obtain input symbols x' and y', respectively. Suppose that $x = x'$ and $y = y'$: their shared state has collapsed to $|\theta_{xy}\rangle$. Then measuring the ith coordinate of their output registers will yield outputs (a, b) such that $V(x, y, a, b) = 1$. Thus $\{\theta_{xy}\}$ is an excellent set of advice states – call this ensemble the *ideal advice*. However, in general, this cannot be rounded to a valid game strategy.

Advice states from Grover search. Instead, we will construct another ensemble that mimics the ideal advice, and if $\mathrm{val}^*(G^{\otimes n})$ is too large, can be rounded to a valid game strategy with small error. We construct a state $|\varphi\rangle$ in steps. First, Alice and Bob start with

$$|\psi^0\rangle = \sum_{\mathbf{x}, \mathbf{y}} \sqrt{\mu^{\otimes n}(\mathbf{x}, \mathbf{y})} \, |\mathbf{x}\rangle \otimes |\mathbf{y}\rangle \otimes \sum_{\mathbf{a}, \mathbf{b}} |\xi_{\mathbf{xyab}}\rangle \otimes |\mathbf{a}\rangle \otimes |\mathbf{b}\rangle \, .$$

Note that this state can be produced without any communication. Then, Alice and Bob engage in a short communication protocol to determine whether they've lost or won the game repeated: they need to determine whether there exists a coordinate i such that $V(\mathbf{x}_i, \mathbf{y}_i, \mathbf{a}_i, \mathbf{b}_i) = 0$ – call this a *losing coordinate*. Classically, this would require $\Omega(n)$ bits of communication, which is too large for us. Instead, Alice and Bob can perform a distributed version of Grover's algorithm to search for a losing coordinate. Although Grover's search algorithm is a quantum query algorithm, it is a standard technique to convert query algorithms into communication protocols (see [4]): Alice executes the Grover search algorithm, and whenever she has to query the ith coordinate, Alice sends the query request to Bob, who responds with $(\mathbf{y}_i, \mathbf{b}_i)$. Alice can then compute $V(\mathbf{x}_i, \mathbf{y}_i, \mathbf{a}_i, \mathbf{b}_i)$. If Alice finds a losing coordinate, she aborts the protocol. Otherwise, she accepts. Since the Grover algorithm requires $O(\sqrt{n})$ queries, this communication protocol uses $\tilde{O}(\sqrt{n})$ qubits of communication, where $\tilde{O}(\cdot)$ hides the $\log n$ bits needed for the query request, as well as the input and output lengths. This protocol is performed coherently with the $|\mathbf{x}\rangle$, $|\mathbf{y}\rangle$, $|\mathbf{a}\rangle$, and $|\mathbf{b}\rangle$ registers. Let $|\psi\rangle$ denote the final state of this protocol, and let $|\varphi\rangle$ denote $|\psi\rangle$ *conditioned* on Alice accepting.

If Grover search worked perfectly, then $|\varphi\rangle$ would be essentially the same as the naive $|\theta\rangle$ we described first. However, Grover's algorithm does not perform search perfectly, and has some error. Furthermore, when we condition on Alice not finding a losing coordinate, this error gets multiplied by $1/\mathrm{val}^*(G^{\otimes n})$. Though we are assuming $\mathrm{val}^*(G^{\otimes n})$ is "large", it is still exponentially small, and hence we require that the Grover search has exponentially small error. In our proof, we make some technical adjustments to the search protocol in order to handle this exponential blowup of the Grover error (without increasing the communication complexity to $\Omega(n)$ bits), but for the sake of exposition we will ignore this issue. For now, we can treat $|\varphi\rangle$ as a very good approximation of $|\theta\rangle$ – thus, defining $|\varphi_{xy}\rangle$ in the same way we defined $|\theta_{xy}\rangle$ yields a good ensemble of advice states $\{\varphi_{xy}\}$.

Rounding to a valid quantum strategy.   Now it remains to show that $\{\varphi_{xy}\}$ can be rounded to a valid quantum strategy. We do this by establishing two properties of the state $|\varphi\rangle$: there exists a coordinate $i \in [n]$ such that
1. Suppose we decohere (i.e. measure) the ith coordinate of the $|\mathbf{xy}\rangle$ registers of $|\varphi\rangle$, and let (X_i, Y_i) denote the random measurement outcomes. Then the distribution of (X_i, Y_i) is γ-close to μ, the input distribution of G; and
2. Let B denote the part of φ controlled by Bob. Then the quantum mutual information between X_i (after measurement) and B in φ, denoted by $I(X_i : B)_\varphi$, is at most γ. Similarly, we have $I(Y_i : A)_\varphi \leq \gamma$, where Y_i and A are defined analogously.

Property 1 follows from the fact that the distribution of $\mathbf{x}$ and $\mathbf{y}$, before conditioning, is a product distribution across coordinates i. When we condition on an event with probability λ, then on average, the distribution of the individual coordinates (x_i, y_i) are skewed by at most $\sqrt{\log(1/\lambda)/n}$ in total variation distance. This simple but useful fact is known as *Raz's Lemma* in the parallel repetition literature. Thus, if $\lambda \gg 2^{-n}$, then the input distribution of most coordinates, even after conditioning, is largely unaffected. Here, λ corresponds to the probability that Alice does not abort the protocol, which is at least $\mathrm{val}^*(G^{\otimes n})$.

Property 2 is the most interesting part of our proof. It states that Bob's part of the state φ is relatively uncorrelated with the value of the input X_i, and similarly Alice's part of φ is relatively uncorrelated with the value of Y_i. This uses the fact that our protocol has low communication complexity: intuitively, since Alice and Bob communicate at most $\tilde{O}(\sqrt{n})$ qubits in the protocol, they "learn" at most $\tilde{O}(\sqrt{n})$ bits total about each other's inputs.

Amortized over the n coordinates, this means Alice has about $1/\sqrt{n}$ bits of information about each y_i, on average, and similarly for Bob. When we condition on Alice not aborting, each player's knowledge of the other's inputs increases by at most $\log 1/\lambda$. If $\lambda > 2^{-\gamma n}$, Alice's state has $O(\gamma)$ mutual information with each y_i, on average.

This intuition is formalized by leveraging the beautiful result of Nayak and Salzman that gives limits on the ability of entanglement-assisted quantum communication protocols to transmit classical messages [17]. More specifically, consider a general two-way quantum communication protocol between Alice and Bob, who may start with some shared entangled state. Suppose that Alice is given a uniformly random m-bit message X at the beginning of the protocol. If T qubits are exchanged between Alice and Bob over the course of the protocol, Bob can only guess Alice's input X with probability at most $2^{2T}/2^m$. Equivalently, the mutual information between Bob's final state and X is at most $2T$. Applying the Nayak-Salzman theorem to our setting, and using what we call *Quantum Raz's Lemma*[1], we can conclude that on average, $I(Y_i : A)_\varphi = I(X_i : B)_\varphi = \frac{1}{n}\left(\tilde{O}(\sqrt{n}) + \log 1/\lambda\right) = O(\gamma)$.

Once we have established Property 1 and 2, then the *Quantum Strategy Rounding Lemma* (which can be found in both [6, 13]) then gives that there exists a unitaries $\{U_x\}_x$ and $\{V_y\}_y$ for Alice and Bob, respectively, so that $U_x \otimes V_y |\varphi\rangle \approx |\varphi_{xy}\rangle$. Thus we have a valid quantum strategy for G: on input (x, y), Alice and Bob locally apply unitaries U_x and V_y to their shared state φ, and obtain something close to the advice state φ_{xy}, which they can use to win game G with probability close to 1. For sufficiently small γ, this will be greater than $\text{val}^*(G)$, a contradiction. Thus, $\text{val}^*(G^{\otimes n}) \leq 2^{-\gamma n}$. This concludes the proof outline.

Other technical considerations. While this discussion has been very informal, it captures the conceptual arguments that are required by our analysis. There are many technical details that are handled by the full proof in the Appendix: for example, in order to make the error in the Grover search exponentially small, we increase the communication complexity of the protocol to $\tilde{O}(\log(1/\lambda)/\sqrt{\epsilon})$. If $\lambda < 2^{-\epsilon^{3/2}n}$, where $\text{val}^*(G) = 1 - \epsilon$, then the communication complexity is at most $\tilde{O}(\epsilon n)$, which is still small enough for use in Quantum Raz's Lemma. Another issue is that the communication protocol described requires that Bob transmit his input symbols y_i, which would incur a dependence on the input alphabet size. Through a modification of the protocol and the analysis, we are able to avoid this dependence. Finally, instead of using Grover's algorithm exactly, we use the 3-dimensional search algorithm of Aaronson and Ambainis [1], which performs quantum search in a "spatially local" way. When converted to a communication protocol, the parties no longer need to incur a $\log n$-qubit overhead per round simply to transmit a query request.

The arguments above are not specific to two-players. We prove our theorem for the general k-player case. An important part of this is the k-player generalization of the Quantum Strategy Rounding lemma of [6, 13], which we prove in Lemma 4.3.

2.2 Overview of Theorem 1.2

Theorem 1.2 shows parallel repetition for the entangled value of k-player free CQ games: these are games where the players may produce quantum states as answers. Instead of

---

[1] We make a quick remark about Quantum Raz's Lemma. The ingredients of Quantum Raz's Lemma can be found in various forms in [6, 7, 13], but we find it conceptually advantageous to consolidate these ingredients into a single Lemma that is used as a black box. The benefit of this consolidation is that the overarching structure of the proofs of Theorems 1.1 and 1.2 are the same – really, the only essential difference is the communication protocol!

making measurements on their share of the entangled state, players will apply local unitaries (that depend on their inputs), and transmit a number of qubits to the referee. The referee will then perform some joint verification measurement on all the answer qubits to decide whether to accept or not.

Similarly to the proof of Theorem 1.1, we will use a low-cost communication protocol to design advice states for the repeated-game-to-single-game reduction. However, we were not able to use the distributed Grover search technique. Instead, our low-cost communication protocol performs the following (in the two-player setting): Alice will send Bob her inputs and answer qubits corresponding to coordinates in a small subset $C \subseteq [n]$. Bob will then perform the referee's verification measurement on Alice's inputs and outputs, and his own inputs and outputs, to determine whether they won all the coordinates in the subset C. Having determined whether they won or not, Bob will return Alice's message back to her. If C is sufficiently small, then the communication cost of this task is small.

This simple checking protocol is similar to the checking protocol of [7]. However, in our protocol, Alice and Bob exchange quantum messages, and it is a two-way protocol (because Bob has to return Alice's message back to her). The theorem of [17] again allows us to show that $I(X_i : B)_\varphi$ and $I(Y_i : A)_\varphi$ are small, where φ is the advice state that arises from this communication protocol, and thus we can apply quantum strategy rounding as before.

For the proofs of Theorems 1.2, 1.3, and 1.4, we refer the reader to the full version of our paper [8].

Outline.   In Section 3, we list the quantum information theoretic facts we'll need, as well as prove a few useful technical lemmas (including Quantum Raz's Lemma). In Section 4, we prove our k-player Quantum Strategy Rounding lemma. We prove Theorem 1.1 in Section 5.

3   Preliminaries

We assume familiarity with the basics of quantum information and computation. For a comprehensive reference, we refer the reader to [18, 22]. For a pure state $|\psi\rangle$, we will let ψ denote the density matrix $|\psi\rangle\langle\psi|$. If $|\psi\rangle^{AB}$ is a bipartite state, then ψ^A will be the reduced density matrix of ψ^{AB} on space A. A density matrix ρ^{XA} is a *classical-quantum* (CQ) state if $\rho^{XA} = \sum_x p(x) |x\rangle\langle x| \otimes \rho_x^A$, where $p(x)$ is a probability distribution and ρ_x^A is an arbitrary density matrix on space A. For a probability distribution μ, $x \leftarrow \mu$ indicates x is drawn from μ. For a classical state $\rho^X = \sum_x \mu(x) |x\rangle\langle x|$, we write $x \leftarrow \rho^X$ to denote $x \leftarrow \mu$. We let id denote the identity matrix.

3.1 k-player games

We give a formal definition of games, where the inputs and outputs are classical (in the full version of this paper, we give a more general definition of *CQ games*, where the outputs may be quantum [8]). A k-player game is a tuple $G = (\mathcal{X}, \mathcal{A}, \mu, V)$, where:
1. $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_k$ with each $\mathcal{X}_j$ a finite alphabet,
2. $\mathcal{A} = \mathcal{A}_1 \times \cdots \mathcal{A}_k$ with each $\mathcal{A}_j$ a finite alphabet,
3. μ is a distribution over $\mathcal{X}$,
4. $V : \mathcal{X} \times \mathcal{A} \to \{0, 1\}$ is the verification predicate.
In a k-player G, a referee samples an input $x = (x_1, \ldots, x_k)$ from μ, and sends x_j to player j. The players produce a vector of outputs $a = (a_1, \ldots, a_k)$ (where the jth player outputs symbol a_j, and the referee accepts if $V(x, a) = 1$.

We say a game is *free* if $\mu = \mu_1 \otimes \cdots \otimes \mu_k$, where μ_j is a distribution over $\mathcal{X}_j$ (i.e. μ is a product distribution). A *quantum strategy* for G is a shared state $|\xi\rangle^E$ (where E are k-partite spaces split between the k players), and for each player j a set of measurements $\{M^{j,x_j}\}_{x_j \in \mathcal{X}_j}$ (with each M^{j,x_j} being a set of POVM elements $\{M^{j,x_j}_{a_j}\}_{a_j \in \mathcal{A}_j}$) which act on the space E_j. On input x_j, player j measures the E_j register of $|\xi\rangle$ using measurement M^{j,x_j}, and obtains an outcome a_j, which is then sent to the referee. The entangled value of a game G is defined as the maximum probability a referee will accept over all possible (finite-dimensional) quantum strategies for k players:

$$\mathrm{val}^*(G) = \max_{|\xi\rangle, \{\{M^{j,x_j}\}_{x_j}\}_j} \mathop{\mathbb{E}}_{x \leftarrow \mu} \left[\sum_{a \in \mathcal{A} : V(x,a)=1} \mathrm{tr}\left(M^{1,x_1}_{a_1} \otimes \cdots \otimes M^{k,x_k}_{a_k} \xi \right) \right].$$

The n-fold repetition of a game $G = (\mathcal{X}, \mathcal{A}, \mu, V)$ is denoted by $G^{\otimes n} = (\mathcal{X}^n, \mathcal{A}^n, \mu^{\otimes n}, V^n)$, where: $\mu^{\otimes n}$ is the product distribution over n independent copies of $\mathcal{X}$, and $V^n(\vec{x}, \vec{a}) := \prod_i V(\vec{x}_i, \vec{a}_i)$, with $\vec{x} \in \mathcal{X}^n$ and $\vec{a} \in \mathcal{A}^n$.

3.2 Properties of the squared Bures metric

For two positive semidefinite operators ρ, σ, let the fidelity between ρ and σ be denoted by $F(\rho, \sigma) := \mathrm{tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}}$. The fidelity distance measure has the well-known property that for pure states $|\psi\rangle$ and $|\varphi\rangle$, $F(\psi, \varphi) = |\langle\psi|\varphi\rangle|$. Furthermore, when ρ and σ are classical probability distributions in the same basis (i.e. $\rho = \sum_i p_i |i\rangle\langle i|$ and $\sigma = \sum_i q_i |i\rangle\langle i|$), then $F(\rho, \sigma) = \sum_i \sqrt{p_i q_i}$.

The fidelity distance measure is not a metric on the space of positive semidefinite operators. For one, it does not satisfy a triangle inequality. However, one can convert fidelity into other measures that are metrics. One such measure is the *Bures metric*, defined as $\mathcal{B}(\rho, \sigma) := \sqrt{1 - F(\rho, \sigma)}$. In this paper, we will use the *squared Bures metric*, denoted by $K(\rho, \sigma) := \mathcal{B}(\rho, \sigma)^2$, as the primary distance measure between quantum states. It satisfies many pleasant properties, including the following:

▶ **Fact 3.1** (Triangle inequality). *Let $n \geq 2$ and let $\rho_1, \ldots, \rho_{n+1}$ be density matrices. Then*

$$K(\rho_1, \rho_{n+1}) \leq n \sum_i K(\rho_i, \rho_{i+1}).$$

Proof. We adapt the proof from [7]. For $i \in [n]$ let $\alpha_i = \arccos(F(\rho_i, \rho_{i+1}))$. Let $\alpha = \arccos(F(\rho_1, \rho_{n+1}))$. Then, since $\arccos(F(\cdot, \cdot))$ is a distance measure for quantum states, we have $\alpha \leq \sum_i \alpha_i$. Then we have

$$K(\rho_1, \rho_{n+1}) = 1 - \cos(\alpha) \leq n^2(1 - \cos(\alpha/n)) \leq n \sum_i (1 - \cos(\alpha_i)) = n \sum_{i=1}^n K(\rho_i, \rho_{i+1}).$$

◀

▶ **Fact 3.2** (Contractivity under quantum operations). *Let $\mathcal{E}$ be a quantum operation, and let ρ and σ be density matrices. Then $K(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \leq K(\rho, \sigma)$.*

▶ **Fact 3.3** (Unitary invariance). *Let U be unitary, and let ρ and σ be density matrices. Then $K(U\rho U^\dagger, U\sigma U^\dagger) = K(\rho, \sigma)$.*

▶ **Fact 3.4** (Convexity). *Let $\{A_i\}$ and $\{B_i\}$ be finite collections of positive semidefinite operators, and let $\{p_i\}$ be a probability distribution. Then $K(\sum_i p_i A_i \sum_i p_i B_i) \leq \sum_i p_i K(A_i, B_i)$.*

▶ **Fact 3.5.** *Let $\{A_i\}$ and $\{B_i\}$ be finite collections of positive semidefinite operators, and let $\{p_i\}$ be a probability distribution. Then $K(\sum_i p_i |i\rangle\langle i| \otimes A_i, \sum_i p_i |i\rangle\langle i| \otimes B_i) = \sum_i p_i K(A_i, B_i)$.*

3.3   Quantum information theory

For two positive semidefinite operators ρ, σ, the relative entropy $S(\rho\|\sigma)$ is defined to be $\mathrm{tr}(\rho(\log\rho - \log\sigma))$. The relative min-entropy $S_\infty(\rho\|\sigma)$ is defined as $\min\{\lambda : \rho \preceq 2^\lambda\sigma\}$. The entropy of ρ is denoted by $H(\rho) := -\mathrm{tr}(\rho\log\rho)$. For a tripartite state ρ^{ABC}, the conditional mutual information $H(A|B)_\rho$ is defined as $H(\rho^{AB}) - H(\rho^B)$. Let ρ^{AB} be a bipartite state. Then the mutual information $I(A:B)_\rho$ is defined as $H(A)_\rho - H(A|B)_\rho$. An equivalent definition is $I(A:B)_\rho = S(\rho^{AB}\|\rho^A \otimes \rho^B)$.

▶ **Fact 3.6** ([14]). *Let ρ and σ be density matrices. Then $S(\rho\|\sigma) \geq K(\rho, \sigma)$.*

▶ **Fact 3.7** ([14]). *Let μ be a probability distribution on $\mathcal{X}$. Let $\rho = \sum_{x\in X} \mu_x |x\rangle\langle x| \otimes \rho_x^A$. Then $I(X:A)_\rho = \mathbb{E}_{x\leftarrow\mu}[S(\rho_x\|\rho)]$.*

▶ **Fact 3.8** ([13], Fact II.11). *Let ρ^{XY} and σ^{XY} be quantum states. Then $S(\rho^{XY}\|\sigma^{XY}) \geq S(\rho^X\|\sigma^X)$.*

▶ **Fact 3.9.** *Let ρ^{XY} and $\sigma^{XY} = \sigma^X \otimes \sigma^Y$ be quantum states. Then $S(\rho^{XY}\|\sigma^{XY}) \geq S(\rho^X\|\sigma^X) + S(\rho^Y\|\sigma^Y)$.*

▶ **Fact 3.10** ([13], Fact II.8). *Let $\rho = \sum_x \mu(x) |x\rangle\langle x| \otimes \rho_x$, and $\rho^1 = \sum_x \mu^1(x) |x\rangle\langle x| \otimes \rho_x^1$. Then $S(\rho^1\|\rho) = S(\mu^1\|\mu) + \mathbb{E}_{x\leftarrow\mu^1}\left[S(\rho_x^1\|\rho_x)\right]$.*

▶ **Fact 3.11** ([13], Lemma II.13). *Let $\rho = p\rho_0 + (1-p)\rho_1$. Then $S_\infty(\rho_0\|\rho) \leq \log 1/p$.*

▶ **Fact 3.12.** *Let ρ^{AB} and σ^{AB} be density matrices. Then $S_\infty(\rho^{AB}\|\sigma^{AB}) \geq S_\infty(\rho^A\|\sigma^B)$.*

▶ **Fact 3.13.** *Let ρ, σ, and τ be density matrices such that $S_\infty(\rho\|\sigma) \leq \lambda_1$ and $S_\infty(\sigma\|\tau) \leq \lambda_2$. Then $S_\infty(\rho\|\tau) \leq \lambda_1 + \lambda_2$.*

▶ **Fact 3.14.** *Let ρ, σ, and τ be density matrices such that $S(\rho\|\sigma) \leq \lambda_1$ and $S_\infty(\sigma\|\tau) \leq \lambda_2$. Then $S_\infty(\rho\|\tau) \leq \lambda_1 + \lambda_2$.*

Proof. $S_\infty(\sigma\|\tau) = \lambda_2$ implies that $2^{-\lambda_2}\sigma \preceq \tau$. Then,

$$
\begin{aligned}
S(\rho\|\tau) &= \mathrm{tr}(\rho(\log\rho - \log\tau)) \\
&\leq \mathrm{tr}(\rho(\log\rho - \log 2^{-\lambda_2}\sigma)) \\
&\leq \mathrm{tr}(\rho(\log\rho - (-\lambda_2)\mathsf{id} - \log\sigma)) \\
&\leq \lambda_2 + \mathrm{tr}(\rho(\log\rho - \log\sigma)) \\
&= \lambda_1 + \lambda_2.
\end{aligned}
$$

◀

3.4   Some technical lemmas

The following lemma is due to [3]:

▶ **Lemma 3.15** ([3], Lemma 3.3). *Let $P = (p, 1-p)$ and $Q = (q, 1-q)$ be binary distributions. If $S(P\|Q) \leq \delta$, and $p < \delta$, then $q \leq 4\delta$.*

The following adapts Lemma 3.15 to use the distance measure K instead:

▶ **Lemma 3.16.** *Let $P = (p, 1-p)$ and $Q = (q, 1-q)$ be binary distributions. If $K(P, Q) \leq \delta$, and $p < \delta$, then $q \leq 9\delta$.*

Proof. If $q \leq p$, then we are done. Assume otherwise. We have that $\delta \geq K(P, Q) \geq (1 - F(P, Q)^2)/2$, because $0 \leq F(P, Q) \leq 1$. Then,

$$\begin{aligned} F(P, Q)^2 &= (\sqrt{pq} + \sqrt{(1-p)(1-q)})^2 \\ &= pq + 1 - p - q + pq + 2\sqrt{pq(1-p)(1-q)}, \end{aligned}$$

and thus

$$\begin{aligned} 2\delta &\geq p + q - 2pq - 2\sqrt{pq(1-p)(1-q)} \\ &\geq p + q - 2pq - 2\sqrt{pq} \\ &= (\sqrt{p} - \sqrt{q})^2 - 2pq \\ &\geq (\sqrt{p} - \sqrt{q})^2 - 2\delta, \end{aligned}$$

where in the last line we used the assumption that $p \leq \delta$. Then $2\sqrt{\delta} \geq |\sqrt{p} - \sqrt{q}|$. Either $q \leq p$, in which case $q \leq \delta$, or $q \geq p$, in which case $\sqrt{q} \leq 2\sqrt{\delta} + \sqrt{p} \leq 3\sqrt{\delta}$, so $q \leq 9\delta$. ◀

Finally, we prove a quantum analogue of Raz's Lemma, which is the central tool behind many information-theoretic proofs of parallel repetition theorems [19, 12, 3]:

▶ **Lemma 3.17** (Quantum Raz's Lemma). *Let $\psi^{XA} = \left(\sum_x \mu(x) |x\rangle\langle x|^X\right) \otimes \psi^A$ be a CQ-state, classical on X and quantum on A, where X is n-partite. Furthermore, suppose that $\mu(x) = \prod \mu_i(x_i)$. Let $\varphi^{XA} = \sum_x \sigma(x) |x\rangle\langle x|^X \otimes \varphi_x^A$ be such that $S(\varphi\|\psi) \leq t$. Then,*

$$\sum_i I(X_i : A)_\varphi \leq 2t.$$

Proof. First observe the following manipulations:

$$\begin{aligned} t &\geq S(\varphi^{XA}\|\psi^{XA}) \\ &= S(\varphi^{XA}\|\psi^X \otimes \psi^A) \\ &\geq S(\varphi^{XA}\|\varphi^X \otimes \varphi^A) \\ &= I(X : A)_\varphi \\ &= H(X)_\varphi - H(X|A)_\varphi \\ &\geq H(X)_\varphi - \sum_i H(X_i|A)_\varphi. \end{aligned}$$

We focus on $H(X)_\varphi$ now. Using that relative entropy is always non-negative:

$$\begin{aligned} -H(X)_\varphi + \sum_i H(X_i)_\varphi &\leq -H(X)_\varphi + \sum_i S(\varphi^{X_i}\|\psi^{X_i}) + H(X_i)_\varphi \\ &= -H(X)_\varphi - \sum_i \text{tr}(\varphi^{X_i} \log \psi^{X_i}) \\ &= -H(X)_\varphi - \text{tr}(\varphi^X \log \psi^X) \\ &= S(\varphi^X\|\psi^X) \\ &\leq t. \end{aligned}$$

Continuing, we have

$$t \geq -t + \sum_i H(X_i)_\varphi - H(X_i|A)_\varphi = -t + \sum_i I(X_i : A)_\varphi.$$

◀

4   Quantum strategy rounding

In this section we prove our k-player Quantum Strategy Rounding lemma, generalizing the technique of [7, 13].

▶ **Lemma 4.1** ([13]). *Let μ be a probability distribution on $\mathcal{X}$. Let*

$$|\varphi\rangle := \sum_{x \in \mathcal{X}} \sqrt{\mu(x)} \, |xx\rangle^{XX'} \otimes |\varphi_x\rangle^{AB} \, .$$

Let $|\varphi_x\rangle := |xx\rangle^{XX'} \otimes |\varphi_x\rangle^{AB}$. Then there exists unitary operators $\{U_x\}_{x \in \mathcal{X}}$ acting on $XX'A$ such that

$$\mathop{\mathbb{E}}_{x \leftarrow \mu} \left[K(\varphi_x, U_x \varphi \, U_x^\dagger) \right] \leq I(X : B)_\varphi.$$

Proof. We follow the proof in [13]. Denote the reduced states of Bob by $\rho_x := \mathrm{tr}_{XX'A}(\varphi_x)$ and $\rho := \mathrm{tr}_{XX'A}(\varphi)$. By Facts 3.6 and 3.7, we get that

$$I(X : B)_\varphi = \mathop{\mathbb{E}}_{x \leftarrow \mu}[S(\rho_x \| \rho)] \geq \mathop{\mathbb{E}}_{x \leftarrow \mu}[K(\rho_x, \rho)].$$

By Uhlmann's Theorem, for each $x \in \mathcal{X}$ there exists U_x such that $|\langle \varphi_x | (U_x \otimes \mathsf{id}_B) |\varphi\rangle| = F(\rho_x, \rho)$. Furthermore, this is equal to $F(\varphi_x, U_x \otimes \mathsf{id}_B \, \varphi \, U_x^\dagger \otimes \mathsf{id}_B)$. We thus obtain the claim. ◀

▶ **Lemma 4.2.** *Let $\{|\varphi_a\rangle\}_{a \in \mathcal{A}}$ be a finite collection of pure states. Let μ and τ be probability distributions over $\mathcal{A}$ such that $S(\mu \| \tau) \leq \epsilon$. Then*

$$K\left(\mathop{\mathbb{E}}_{a \leftarrow \mu}[|\varphi_a\rangle\langle\varphi_a|], \mathop{\mathbb{E}}_{a \leftarrow \tau}[|\varphi_a\rangle\langle\varphi_a|] \right) \leq \epsilon.$$

Proof. Consider the states

$$|\psi^\mu\rangle = \sum_{a \in \mathcal{A}} \sqrt{\mu_a} \, |aa\rangle^{AA'} \otimes |\varphi_a\rangle$$

and

$$|\psi^\tau\rangle = \sum_{a \in \mathcal{A}} \sqrt{\tau_a} \, |aa\rangle^{AA'} \otimes |\varphi_a\rangle \, .$$

Let $\rho^\mu = \mathrm{tr}_{A'}(|\psi^\mu\rangle\langle\psi^\mu|)$ and $\rho^\tau = \mathrm{tr}_{A'}(|\psi^\tau\rangle\langle\psi^\tau|)$. Then notice that $\mathbb{E}_{a \leftarrow \mu}[|\varphi_a\rangle\langle\varphi_a|] = \mathrm{tr}_{AA'}(\rho^\mu)$ and $\mathbb{E}_{a \leftarrow \tau}[|\varphi_a\rangle\langle\varphi_a|] = \mathrm{tr}_{AA'}(\rho^\tau)$, respectively. We then have that, considering the partial trace as a quantum operation, $K(\mathbb{E}_{a \leftarrow \mu}[|\varphi_a\rangle\langle\varphi_a|], \mathbb{E}_{a \leftarrow \tau}[|\varphi_a\rangle\langle\varphi_a|]) \leq K(\rho^\mu, \rho^\tau)$. By Uhlmann's Theorem, this is at most $1 - |\langle\psi^\mu|\psi^\tau\rangle| = 1 - \sum_{a \in \mathcal{A}} \sqrt{\mu_a \tau_a} = K(\mu, \tau)$. By Fact 3.6, this is at most $S(\mu \| \tau) \leq \epsilon$. ◀

▶ **Lemma 4.3** (Quantum strategy rounding). *Let $k \geq 1$. Let μ be a probability distribution over $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_k$, where the $\mathcal{X}_i$ are finite alphabets. Let*

$$|\varphi\rangle := \sum_{x \in \mathcal{X}} \sqrt{\mu(x)} \, |xx\rangle^{XX'} \otimes |\varphi_x\rangle^{AB}$$

where $X = X_1 \cdots X_k$, $X' = X'_1 \cdots X'_k$, and $A = A_1 \cdots A_k$ are k-partite registers. Then for all $i \in [k]$ there exist operators $\{U^i_a\}_{a \in \mathcal{X}_i}$ acting on $X_i X'_i A_i$ such that

$$\mathop{\mathbb{E}}_{x \leftarrow \mu} \left[K\left(\varphi_x, (U^1_{x_1} \otimes \cdots \otimes U^k_{x_k}) \varphi (U^{1,\dagger}_{x_1} \otimes \cdots \otimes U^{k,\dagger}_{x_k}) \right) \right] \leq 4k \sum_i I(X_i : X_{-i} X'_{-i} A_{-i} B)_\varphi,$$

where A_{-i}, X_{-i}, and X'_{-i} denote the A, X, and X' registers excluding the ith coordinate, respectively, and for all $x \in \mathcal{X}$, $|\varphi_x\rangle := |xx\rangle \otimes |\varphi_x\rangle$.

Proof. For $i \in [k]$, let $\nu_i = \mu_1 \otimes \cdots \otimes \mu_i \otimes \mu_{>i}$, where μ_j denotes the marginal distribution of μ on coordinate j, and $\mu_{>i}$ denotes the marginal distribution of μ on coordinates $i+1, \ldots, k$. For $x \in \mathcal{X}$, for all $S \subseteq [k]$, let x_S denote the coordinates of x that are in S. Therefore, $x_{\leq i} = x_{1 \ldots i}$, and $x_{>i} = x_{i+1 \ldots k}$, etc. For all $i \in [k]$ and $x \in \mathcal{X}$, define

$$|\varphi_{x_{>i}}\rangle := |x_{>i} x_{>i}\rangle^{X_{>i} X'_{>i}} \otimes \left(\sum_{x_{\leq i}} \sqrt{\mu(x_{\leq i} | x_{>i})} |x_{\leq i} x_{\leq i}\rangle^{X_{\leq i} X'_{\leq i}} \otimes |\varphi_x\rangle^{AB} \right)$$

and

$$|\varphi_{x_i}\rangle := |x_i x_i\rangle^{X_i X'_i} \otimes \left(\sum_{x_{-i}} \sqrt{\mu(x_{-i} | x_i)} |x_{-i} x_{-i}\rangle^{X_{-i} X'_{-i}} \otimes |\varphi_x\rangle^{AB} \right).$$

Note that for all i, $|\varphi\rangle = \sum_{x_i} \sqrt{\mu_i(x_i)} |\varphi_{x_i}\rangle$. Then by Lemma 4.1, we get that there exists unitaries $\{U^i_u\}_{u \in \mathcal{X}_i}$ acting on $X_i X'_i A_i$ such that

$$\mathop{\mathbb{E}}_{x_i \leftarrow \mu_i} [K(\varphi_{x_i}, \mathcal{U}^i_{x_i}(\varphi))] \leq I(X_i : X_{-i} X'_{-i} A_{-i} B)_\varphi,$$

where $\mathcal{U}^i_{x_i}$ is the CP map $\sigma \mapsto U^i_{x_i} \sigma (U^i_{x_i})^\dagger$. Define $|\tilde{\varphi}_{x_{>i}}\rangle = |x_{>i}\rangle^{X''_{>i}} \otimes |\varphi_{x_{>i}}\rangle$, $|\tilde{\varphi}_{x_i}\rangle = |x_i\rangle^{X''_i} \otimes |\varphi_{x_i}\rangle$, and $|\tilde{\varphi}_x\rangle = |x\rangle^{X''} \otimes |\varphi_x\rangle$. For notational convenience, let $\epsilon_i = I(X_i : X_{-i} X'_{-i} A_{-i} B)_\varphi$, and let x, x_i and $x_{>i}$ denote the pure states $|x\rangle\langle x|$, $|x_i\rangle\langle x_i|$, and $|x_{>i}\rangle\langle x_{>i}|$ respectively.

Define the following states: $\rho_0 = \mathbb{E}_{x \leftarrow \mu}[x \otimes \varphi_x] = \mathbb{E}_{x \leftarrow \mu}[\tilde{\varphi}_x]$, and for all $i \in [k]$, $\rho_i = \mathbb{E}_{x \leftarrow \nu_i}[x \otimes \mathcal{U}_{x_{\leq i}}(\varphi_{x_{>i}})]$, where $\mathcal{U}_{x_{\leq i}}$ denotes the CP map $\sigma \mapsto \left(\bigotimes_{j \leq i} U^j_{x_j} \right) \sigma \left(\bigotimes_{j \leq i} U^j_{x_j} \right)^\dagger$. Then by the triangle inequality for the squared Bures metric (Fact 3.1),

$$K(\rho_0, \rho_n) \leq k \sum_{i=0}^{k-1} K(\rho_i, \rho_{i+1}).$$

We upper bound each term $K(\rho_i, \rho_{i+1})$:

$$K\left(\mathop{\mathbb{E}}_{x \leftarrow \nu_i}[x \otimes \mathcal{U}_{x_{\leq i}}(\varphi_{x_{>i}})], \mathop{\mathbb{E}}_{x \leftarrow \nu_{i+1}}[x \otimes \mathcal{U}_{x_{\leq i+1}}(\varphi_{x_{>i+1}})] \right) \leq$$

$$\mathop{\mathbb{E}}_{x_{\leq i} \leftarrow \otimes \mu_i} K\left(\mathcal{U}_{x_{\leq i}} \left(\mathop{\mathbb{E}}_{x_{>i} \leftarrow \mu_{>i}}[x_{>i} \otimes \varphi_{x_{>i}}] \right), \mathcal{U}_{x_{\leq i}} \left(\mathop{\mathbb{E}}_{x_{>i} \leftarrow \mu_{i+1} \otimes \mu_{>i+1}}[x_{>i} \otimes \mathcal{U}^{i+1}_{x_{i+1}}(\varphi_{x_{>i+1}})] \right) \right)$$

$$= K\left(\mathop{\mathbb{E}}_{x_{>i} \leftarrow \mu_{>i}}[x_{>i} \otimes \varphi_{x_{>i}}], \mathop{\mathbb{E}}_{x_{>i} \leftarrow \mu_{i+1} \otimes \mu_{>i+1}}[x_{>i} \otimes \mathcal{U}^{i+1}_{x_{i+1}}(\varphi_{x_{>i+1}})] \right)$$

$$= K\left(\mathop{\mathbb{E}}_{x_{>i} \leftarrow \mu_{>i}}[\tilde{\varphi}_{x_{>i}}], \mathop{\mathbb{E}}_{x_{i+1} \leftarrow \mu_{i+1}} \left[x_{i+1} \otimes \mathcal{U}^{i+1}_{x_{i+1}} \left(\mathop{\mathbb{E}}_{x_{>i+1} \leftarrow \mu_{>i+1}}[\tilde{\varphi}_{x_{>i+1}}] \right) \right] \right)$$

The second and third lines follow from the convexity and unitary invariance of the squared Bures metric, respectively (Facts 3.4 and Fact 3.3). Consider the operation $\mathcal{E}$ that measures the

registers $X_{>i+1} = X_{i+2} \dots X_k$ in the standard basis, and copies the outcomes into new registers $X''_{>i+1}$. Then $\mathcal{E}(\mathbb{E}_{x_{i+1} \leftarrow \mu_{i+1}}[\tilde{\varphi}_{x_{i+1}}]) = \mathbb{E}_{x_{>i} \leftarrow \mu_{>i}}[\tilde{\varphi}_{x_{>i}}]$ and $\mathcal{E}(\varphi) = \mathbb{E}_{x_{>i+1} \leftarrow \mu_{>i+1}}[\tilde{\varphi}_{x_{>i+1}}]$. Then since $\mathcal{E}$ commutes with $\mathcal{U}^{i+1}_{x_{i+1}}$ and doesn't act on the X''_{i+1} register, we have that the line above is equal to

$$
= K\left(\mathcal{E}\left(\mathop{\mathbb{E}}_{x_{i+1} \leftarrow \mu_{i+1}}[\tilde{\varphi}_{x_{i+1}}]\right), \mathcal{E}\left(\mathop{\mathbb{E}}_{x_{i+1} \leftarrow \mu_{i+1}}[x_{i+1} \otimes \mathcal{U}^{i+1}_{x_{i+1}}(\varphi)]\right)\right)
$$

$$
\leq K\left(\mathop{\mathbb{E}}_{x_{i+1} \leftarrow \mu_{i+1}}[\tilde{\varphi}_{x_{i+1}}], \mathop{\mathbb{E}}_{x_{i+1} \leftarrow \mu_{i+1}}[x_{i+1} \otimes \mathcal{U}^{i+1}_{x_{i+1}}(\varphi)]\right)
$$

$$
= \mathop{\mathbb{E}}_{x_{i+1} \leftarrow \mu_{i+1}} K\left(\varphi_{x_{i+1}}, \mathcal{U}^{i+1}_{x_{i+1}}(\varphi)\right)
$$

$$
\leq \epsilon_{i+1}.
$$

To complete the proof, we use the triangle inequality once more:

$$
\mathop{\mathbb{E}}_{x \leftarrow \mu} K(\varphi_x, \mathcal{U}_x(\varphi)) = K\left(\mathop{\mathbb{E}}_{x \leftarrow \mu}[\tilde{\varphi}_x], \mathop{\mathbb{E}}_{x \leftarrow \mu}[x \otimes \mathcal{U}_x(\varphi)]\right)
$$

$$
\leq 2K\left(\mathop{\mathbb{E}}_{x \leftarrow \mu}[\tilde{\varphi}_x], \mathop{\mathbb{E}}_{x \leftarrow \nu_k}[x \otimes \mathcal{U}_x(\varphi)]\right) +
$$

$$
2K\left(\mathop{\mathbb{E}}_{x \leftarrow \nu_k}[x \otimes \mathcal{U}_x(\varphi)], \mathop{\mathbb{E}}_{x \leftarrow \mu}[x \otimes \mathcal{U}_x(\varphi)]\right)
$$

$$
\leq 2k \sum_i \epsilon_i + 2k \sum_i \epsilon_i
$$

$$
\leq 4k \sum_i \epsilon_i.
$$

where $\mathcal{U}_x$ is the composition of $\mathcal{U}^i_{x_i}$ for all $i \in [k]$. Here we used Lemma 4.2 in the second line, and the fact that $S(\mu \| \nu_k) = I(X_1 : X_2 : \dots : X_k)_\mu$, which is the multipartite mutual information between the coordinates of X. It is a known fact (see, e.g., [23]) that the multipartite mutual information can be written in terms of the (standard) bipartite mutual information like so:

$$
I(X_1 : X_2 : \dots : X_k)_\mu \leq I(X_1 : X_2)_\mu + I(X_1 X_2 : X_3)_\mu + \dots + I(X_1 X_2 \cdots X_{k-1} : X_k)_\mu,
$$

but by the data processing inequality, we have that for all i, $I(X_1 \cdots X_{i-1} : X_i)_\mu \leq I(X_{-i} : X_i)_\mu \leq I(X_i : X_{-i} X'_{-i} A_{-i} B)_\varphi = \epsilon_i$. ◀

5   Parallel repetition using fast quantum search

Notation.   Let $G = (\mathcal{X}, \mathcal{A}, \mu, V)$ be a k-player free game. In what follows, we will think of $x \in \mathcal{X}^n$ as $n \times k$ matrices, where the ith row indicates the inputs of all k players in the ith coordinate, and the jth column indicates the inputs of the jth player. Thus $x(i, \cdot)$ indicates the ith row of x, and $x(\cdot, j)$ indicates the jth column. When we write x_S for some subset $S \subseteq [n]$, we mean the submatrix of x consisting of the rows indexed by $i \in S$.

Let X be an $n \times k$-partite register. Then we will also format X as a $n \times k$ matrix, so $X_{(i,\cdot)}$ and $X_{(\cdot,j)}$ have the natural meaning. For a subset $S \subseteq [n]$, X_S denotes the registers corresponding to the rows of X indexed by S. For an index j, $X_{(S,j)}$ denotes the jth column of the rows indexed by S. $X_{(S,-j)}$ denotes the submatrix of X corresponding to rows indexed by $i \in S$, and all columns except for the jth one.

We make the following observation, which will be useful for us in our analysis: without loss of generality, we can restrict our attention to free games whose input distribution is the uniform distribution over some alphabet. Let $G = (\mathcal{X}, \mathcal{A}, \mu, V)$ be a k-player free game. Write $\mu = \mu_1 \otimes \cdots \otimes \mu_k$, where μ_j is a distribution over the alphabet $\mathcal{X}_j$. Fix an $\gamma > 0$. For each i, there exists an alphabet $\mathcal{X}'_j$ and a map $f_j : \mathcal{X}'_j \to \mathcal{X}_j$ such that the random variable $X'_j = f_j(U_j)$ (where U_j is a uniformly random element from $\mathcal{X}'_j$) is γ/k-close in total variation distance to being distributed according to μ_j – and hence the distribution of $(f_1(U_1), \ldots, f_k(U_k))$ is at most γ-far from μ. Thus, we can "simulate" the game G with another game $G' = (\mathcal{X}', \mathcal{A}, U, V')$, where $\mathcal{X}' = \mathcal{X}'_1 \times \cdots \mathcal{X}'_k$, U is the uniform distribution on $\mathcal{X}'$, and $V' : \mathcal{X}' \times \mathcal{A} \to \{0,1\}$ is the map $(x', a) \to V(\langle f_1(x'_1), \ldots, f_k(x'_k)\rangle, a)$.

▶ **Claim 5.1.** $\mathrm{val}^*(G') = \mathrm{val}^*(G) \pm \gamma$.

Proof. Consider the optimal strategy for G. Then a strategy for G' is the following: player j, on input $u'_j \in \mathcal{X}'_j$, computes $u_j = f_j(u'_j)$, and performs the strategy she would've done for G. The input distribution, from the point of view of the strategy for G, is at most γ-far from the original input distribution μ. Thus the winning probability is at least $\mathrm{val}^*(G) - \gamma$.

Now consider the optimal strategy for G'. Then a strategy for G is the following: player j, on input $u_j \in \mathcal{X}_j$, computes a uniformly random preimage $u'_j \in f_j^{-1}(u_j)$, and performs the strategy she would've done for G'. The input distribution, from the point of view of the strategy for G', is at most γ-far from the uniform distribution U. Thus the winning probability is at least $\mathrm{val}^*(G') - \gamma$. ◀

Furthermore, this simulation "commutes" with parallel repetition, in that $\mathrm{val}^*((G')^{\otimes n}) = \mathrm{val}^*(G^{\otimes n}) \pm \gamma n$. We can make γ arbitrarily small, at the cost of (potentially) increasing the input alphabet size, so that the behavior of the simulation G' is essentially the same as the original game G. However, since our theorems do not depend on the input alphabet size, we will treat γ as infinitesimally small, and hence neglect it.

▶ **Theorem 5.2.** *Let $G = (\mathcal{X}, \mathcal{A}, \mu, V)$ be a k-player free game with classical outputs and classical verification predicate $V : \mathcal{X} \times \mathcal{A} \to \{0,1\}$. Suppose that $\mathrm{val}^*(G) = 1 - \epsilon$. Let $s = \max_j \log |\mathcal{A}_j|$. Then, for all $n > k^4 s \log(k^2/\epsilon)/\epsilon^{3/2}$, we have that*

$$\mathrm{val}^*(G^{\otimes n}) \le (1 - \epsilon^{3/2})^{\Omega(n/k^4 s)}.$$

Proof. Because of Claim 5.1, it is without loss of generality to assume that the input distribution μ is the uniform distribution – the following analysis can be performed on a simulation of G, which will still bound the repeated game value of G.

Let n be some integer greater than $k^4 s \log(k^2/\epsilon)/\epsilon^{3/2}$, and consider an optimal entangled strategy for $G^{\otimes n}$, and let 2^{-t} denote its winning probability. Suppose for contradiction that $t \le c\epsilon^{3/2} n/(k^4 s)$ for some universal constant c. Using this strategy, we will construct the following state

$$|\varphi\rangle^{XX'EA} := \sum_{x \in \mathcal{X}^n} \sqrt{\nu(x)} |xx\rangle^{XX'} \otimes |\varphi_x\rangle^{EA},$$

where $\nu(x)$ is a probability distribution over $\mathcal{X}^n$, X, X', A are $n \times k$-partite registers, and E is a k-partite register. We will show that exists a coordinate $i \in [n]$, and $\delta < \epsilon/32k^2$ satisfying the following properties:

1. Measuring the $X_{(i,\cdot)}A_{(i,\cdot)}$ register of φ yields a tuple $(x_{(i,\cdot)}, a_{(i,\cdot)})$ that satisfies $V(x_{(i,\cdot)}, a_{(i,\cdot)}) = 1$ with probability at least $1 - \epsilon/8$;
2. $S(\varphi^{X_{(i,\cdot)}} \| \mu) \le \delta$.
3. For all $j \in [k]$, $I(X_{(i,j)} : Z_{-j})_\varphi \le \delta$, where $Z_{-j} = X'_{(\cdot,-j)} X_{(\cdot,-j)} E_{-j} A_{(\cdot,-j)}$.

For now, we assume the existence of such a state $|\varphi\rangle$; we will construct it in Lemma 5.3. We use Lemma 4.3 on the state φ to obtain for each player j a set of unitaries $\{U_u^j\}_{u \in \mathcal{X}_j}$ acting on $X_{(\cdot,j)} X'_{(\cdot,j)} E_j A_{(\cdot,j)}$ such that

$$\mathop{\mathbb{E}}_{x_{(i,\cdot)} \leftarrow \varphi^{X_{(i,\cdot)}}} \left[K\left(\varphi_{x_{(i,\cdot)}}, \mathcal{U}_{x_{(i,\cdot)}}(\varphi)\right)\right] \leq 4k \sum_j I(X_{(i,j)} : Z_{-j})_\varphi \leq 4k^2\delta,$$

where we let $U_{x_{(i,\cdot)}} = \bigotimes_j U_{x_{(i,j)}}^j$, and let $\mathcal{U}_{x_{(i,\cdot)}}$ be the CP map $\varphi \mapsto U_{x_{(i,\cdot)}} \varphi U_{x_{(i,\cdot)}}^\dagger$. The state $\varphi_{x_{(i,\cdot)}}$ denotes φ conditioned on $X_{(i,\cdot)} = x_{(i,\cdot)}$.

We now describe a protocol for the k players to play game G. The players receive $u \in \mathcal{X}$, drawn from the product distribution μ. Player j receives $u_j \in \mathcal{X}_j$. The players share the state φ, where player j has access to the $X_{(\cdot,j)} X'_{(\cdot,j)} E_j A_{(\cdot,j)}$ registers.

Protocol A

Input: $u \in \mathcal{X}$. Player j receives u_j.
Preshared entanglement: φ
Strategy for player j:
1. Apply the local unitary $U_{u_j}^j$ on the $X_{(\cdot,j)} X'_{(\cdot,j)} E_j A_{(\cdot,j)}$ registers of φ.
2. Output the $A_{(i,j)}$ part of φ.

Slightly overloading notation, we let V_u^i denote the projector $\sum_{a \in \mathcal{A}: V(u,a)=1} |a\rangle\langle a|$ that acts on the $A_{(i,\cdot)}$ registers. Let κ denote the winning probability of Protocol A. This is equal to

$$\kappa = \mathop{\mathbb{E}}_{u \leftarrow \mu} \left\| V_u^i U_u |\varphi\rangle \right\|^2$$
$$\geq \mathop{\mathbb{E}}_{u \leftarrow \varphi^{X_{(i,\cdot)}}} \left\| V_u^i U_u |\varphi\rangle \right\|^2 - 4\delta.$$

where we use property **(B)** and appeal to Lemma 3.15. Let

$$\tau := \mathop{\mathbb{E}}_{u \leftarrow \varphi^{X_{(i,\cdot)}}} \left\| V_u^i U_u |\varphi\rangle \right\|^2.$$

For every $i \in [n]$, $u \in \mathcal{X}$, define the quantum operation $\mathcal{E}_{i,u}$ that, given a state φ, measures the $A_{(i,\cdot)}$ registers using V_u^i measurement, and outputs a classical binary random variable F indicating the verification measurement outcome (outcome 1 corresponds to "accept" and outcome 0 corresponds to "reject"). Let

$$F_0 = \mathop{\mathbb{E}}_{x_{(i,\cdot)} \leftarrow \varphi^{X_{(i,\cdot)}}} \mathcal{E}_{i,x_{(i,\cdot)}} \left(\varphi_{x_{(i,\cdot)}}\right) \qquad \text{and} \qquad F_1 = \mathop{\mathbb{E}}_{u \leftarrow \varphi^{X_{(i,\cdot)}}} \mathcal{E}_{i,u}\left(\mathcal{U}_u(\varphi)\right).$$

Note that $\Pr(F_0 = 1) \geq 1 - \epsilon/8$ by our assumption on φ, and $\Pr(F_1 = 1) = \tau$. Then,

$$K(F_0, F_1) = K\left(\mathop{\mathbb{E}}_{x_{(i,\cdot)} \leftarrow \varphi^{X_{(i,\cdot)}}} \mathcal{E}_{i,x_{(i,\cdot)}} \left(\varphi_{x_{(i,\cdot)}}\right), \mathop{\mathbb{E}}_{u \leftarrow \varphi^{X_{(i,\cdot)}}} \mathcal{E}_{i,u}\left(\mathcal{U}_u(\varphi)\right) \right)$$
$$\leq \mathop{\mathbb{E}}_{x_{(i,\cdot)} \leftarrow \varphi^{X_{(i,\cdot)}}} K\left(\mathcal{E}_{i,x_{(i,\cdot)}} \left(\varphi_{x_{(i,\cdot)}}\right), \mathcal{E}_{i,x_{(i,\cdot)}} \left(\mathcal{U}_{x_{(i,\cdot)}}(\varphi)\right) \right) \qquad \text{(Fact 3.4)}$$
$$\leq \mathop{\mathbb{E}}_{x_{(i,\cdot)} \leftarrow \varphi^{X_{(i,\cdot)}}} K\left(\varphi_{x_{(i,\cdot)}}, \mathcal{U}_{x_{(i,\cdot)}}(\varphi) \right) \qquad \text{(Fact 3.2)}$$
$$\leq 4k^2\delta.$$

By our assumption on δ, this is at most $K(F_0, F_1) \leq \epsilon/8$. By Lemma 3.16, $\Pr(F_1 = 1) \geq 1 - \epsilon/8 - \epsilon/2$. Thus $\kappa \geq 1 - 3\epsilon/4$. But notice that Protocol A is a valid strategy for the game G; thus we have produced a strategy for game G that wins with probability strictly greater than $1 - \epsilon$, a contradiction. Thus, it must be at $t = \Omega(\epsilon^{3/2} n/(k^4 s))$, which establishes the theorem. ◀

5.1 Construction of φ

▶ **Lemma 5.3.** *There exists a state $|\varphi\rangle$, and a coordinate $i \in [n]$ satisfying properties **(A)**, **(B)**, and **(C)**.*

Proof. Set $\epsilon' = \epsilon/32$, $\eta = 2^{-t}\epsilon/32k^2$, and $h = c' \log(1/\eta)/\epsilon'$ for some constant c'. We have $h = (32c'/\epsilon)(t + \log(32k^2/\epsilon))$, and by our assumptions on t and n, this is at most $n/2$.

Suppose there was a strategy to win the repeated game $G^{\otimes n}$ with probability 2^{-t}, involving a shared state $|\xi\rangle^E$ (where E is a k-partite state register) and measurements $\{M_a^{j,x_j}\}$ for the players, respectively. That is, player j, on input $x_j \in \mathcal{X}_j^n$, applies the measurement with POVM elements $\{M_a^{j,x_j}\}$ and reports the outcome.

We will build the state φ in steps. Consider the initial state

$$\left|\psi^0\right\rangle := \sum_{x \in \mathcal{X}^n} \sqrt{\mu^{\otimes n}(x)} |xx\rangle^{XX'} \otimes \sum_{a \in \mathcal{A}^n} |\xi_{xa}\rangle^E \otimes |a\rangle^A$$

where $|\xi_{xa}\rangle = \left(\bigotimes_j \sqrt{M_{a_j}^{j,x_j}}\right) |\xi\rangle$ (which is a subnormalized state), and $\mu^{\otimes n}(x)$ is the probability distribution associated with the repeated game $G^{\otimes n}$. For every set $C \subset [n]$, and every fixing of the inputs x_C to the coordinates indexed by C, define the state $\left|\psi_{C,x_C}^0\right\rangle$ to be $\left|\psi^0\right\rangle$ conditioned on $X_C = x_C$.

Now consider the following k-player communication protocol: for every set $C \subset [n]$ and every x_C, the players share the entangled state $\left|\psi_{C,x_C}^0\right\rangle$, where player j has access to the registers $X_{(\cdot,j)}X'_{(\cdot,j)}E_j A_{(\cdot,j)}$. Using shared randomness, the players sample h independent and uniformly random coordinates $C = \{i_1, \ldots, i_h\} \subset [n]$, and sample x_C from the marginal distribution of $\mu^{\otimes n}$ on the subset C. For the remainder of the protocol, the players perform all their operations on the shared state $\left|\psi_{C,x_C}^0\right\rangle$.

In the next phase of the protocol, the k players communicate qubits to each other to determine whether they have won or lost the parallel repeated game $G^{\otimes n}$. In particular, they run a protocol to search for a coordinate $i \in C$ such that $V(x_{(i,\cdot)}, a_{(i,\cdot)}) = 0$, if it exists – call such a coordinate a losing coordinate. The state $\left|\psi_{C,x_C}^0\right\rangle$ becomes transformed to

$$|\psi_{C,x_C}\rangle^{XX'EAR} := \sum_{x \in \mathcal{X}^n} \sqrt{\mu^{\otimes n}(x|x_C)} |xx\rangle^{XX'} \otimes \sum_{a \in \mathcal{A}^n} |\xi'_{Cxa}\rangle^E \otimes |a\rangle^A \otimes (\alpha_{Cxa} |1\rangle + \beta_{Cxa} |0\rangle)^R$$

where $\mu^{\otimes n}(x|x_C)$ is probability of x conditioned on x_C, and $|\xi'_{Cxa}\rangle = |\xi_{xa}\rangle \otimes |w_{Cxa}\rangle$ with $|w_{Cxa}\rangle$ denoting the workspace qubits that are used during the protocol. The coefficients α_{Cxa} and β_{Cxa} denote the amplitude that the search protocol places on the flags "No losing coordinates" and "Exists a losing coordinate" respectively.

For now, we will abstract away from the particulars of this communication protocol and defer the details of it until later. The only things we will use about this search protocol is the following:

1. The search protocol is run conditioned on C, and the XA registers;
2. At most $T = O(\sqrt{1/\epsilon'} \log(1/\eta) \log |\mathcal{A}|)$ qubits in total are exchanged between all parties, where $\mathcal{A}$ is the output alphabet in game G;

3. For every fixing of $(x, a) \in \mathcal{X}^n \times \mathcal{A}^n$, if there are no coordinates $i \in [n]$ such that $V(x_{(i,\cdot)}, a_{(i,\cdot)}) = 0$, then the search procedure reports "No losing coordinates" with certainty; and

4. If there are at least an $\epsilon' n$ bad coordinates, then the search procedure reports "No losing coordinates" with probability at most η (over the quantum randomness of the protocol, as well as over the choice of C). In other words, for tuples $(x, a) \in \mathcal{X}^n \times \mathcal{A}^n$ with $\mathbb{E}_i[V(x_{(i,\cdot)}, a_{(i,\cdot)})] < 1 - \epsilon'$,

$$\sum_C p(C) |\alpha_{Cxa}|^2 \le \eta,$$

where $p(C)$ is the distribution that samples h independent and uniformly random coordinates from $[n]$.

For all C, x_C define $|\varphi_{C,x_C}\rangle$ to be $|\psi_{C,x_C}\rangle$ conditioned on measuring 1 in the R register:

$$|\varphi_{C,x_C}\rangle^{XX'EAR} := \frac{1}{\sqrt{\lambda_{C,x_C}}} \sum_{x \in \mathcal{X}^n} \sqrt{\mu^{\otimes n}(x|x_C)} |xx\rangle^{XX'} \otimes \sum_{a \in \mathcal{A}^n} |\xi'_{Cxa}\rangle^E \otimes |a\rangle^A \otimes (\alpha_{Cxa} |1\rangle^R)$$

where λ_{C,x_C} is for normalization. In the case that $\lambda_{C,x_C} = 0$ (meaning that we were trying to normalize the 0 state), we leave the state undefined. Let $\psi^{CX_C}(C, x_C) = p(C)\mu^C(x_C)$ denote the joint probability distribution of the shared random variables C and X_C, before conditioning. Let $\varphi^{CX_C}(C, x_C) = p(C)\mu^C(x_C)\lambda_{C,x_C}/\lambda$ denote the joint distribution conditioned on $R = 1$, where $\lambda = \sum_{C,x_C} \psi^{CX_C}(C, x_C)\lambda_{C,x_C}$.

(A) A random coordinate of φ wins with high probability. Let

$$\rho = \mathop{\mathbb{E}}_{C,x_C \leftarrow \psi^{CX_C}} [|C\rangle\langle C| \otimes |x_C\rangle\langle x_C| \otimes \psi_{C,x_C}]$$

and

$$\sigma = \mathop{\mathbb{E}}_{C,x_C \leftarrow \varphi^{CX_C}} [|C\rangle\langle C| \otimes |x_C\rangle\langle x_C| \otimes \varphi_{C,x_C}].$$

Observe that σ is the post-measurement state of ρ after measuring $|1\rangle$ in the R register. Let $\mathcal{E}$ denote the quantum operation on that, (1) measures the C register, (2) chooses a uniformly random $i \notin C$, (3) measures $X_{(i,\cdot)}$ register, and (4) then conditioned on $X_{(i,\cdot)} = x_{(i,\cdot)}$, performs the binary verification measurement $V^i_{x_{(i,\cdot)}}$ defined in the previous section, setting an auxiliary register Q to $|1\rangle$ if the measurement accepts, $|0\rangle$ if it rejects. We wish to argue that the probability that a measurement of the Q register of $\mathcal{E}(\sigma)$ yields 1 with high probability. This probability is equivalent to the probability the following process succeeds: first, measure the XA registers of σ to obtain a tuple (x, a). Then, measure the C register. Finally, select a random index $i \notin C$, and we succeed if $V(x_{(i,\cdot)}, a_{(i,\cdot)}) = 1$.

In this alternative process, the probability that we measure (x, a) in σ such that $\mathbb{E}_{i \in [n]}[V(x_{(i,\cdot)}, a_{(i,\cdot)})] < 1 - \epsilon'$ (call such (x, a)'s "bad") is equal to

$$\frac{1}{\lambda} \sum_{(x,a) \text{ bad}} \mathop{\Pr}_\rho(x, a) \sum_C p(C) |\alpha_{Cxa}|^2$$

where $\Pr_\rho(x, a)$ is the probability of measuring measuring (x, a) in ρ. By our assumption on the communication protocol, this is at most η/λ. Since the players' strategy wins the repeated game $G^{\otimes n}$ with probability 2^{-t}, we have that $\lambda \ge 2^{-t}$. Thus the probability of measuring a bad (x, a) is at most $2^t \eta$.

Now suppose we measure (x, a) such that $\mathbb{E}_{i \in [n]}[V(x_{(i,\cdot)}, a_{(i,\cdot)})] \ge 1 - \epsilon'$. Then, for any C, a random $i \notin C$ loses with probability at most $\epsilon' n/(n - |C|) \le \epsilon' n/(n - h) \le \epsilon/16$. Thus, the probability that the Q register of $\mathcal{E}(\sigma)$ yields 0 is at most $2^t \eta + \epsilon/16$.

(B) Coordinate input distributions are mostly unaffected. By Fact 3.11, since $\sigma \preceq 2^\lambda \rho$, we have

$$
\begin{aligned}
\log 1/\lambda &\geq S_\infty(\sigma\|\rho) \\
&\geq S(\sigma\|\rho) \\
&\geq \mathop{\mathbb{E}}_{C,x_C \leftarrow \varphi^{CX_C}} S(\varphi_{C,x_C}^{XX'EA}\|\psi_{C,x_C}^{XX'EA}),
\end{aligned}
\tag{5.1}
$$

where in the last line we used Fact 3.10. Using Facts 3.8 and 3.9, we obtain that

$$
\begin{aligned}
\log 1/\lambda &\geq \mathop{\mathbb{E}}_{C,x_C \leftarrow \varphi^{CX_C}} S(\varphi_{C,x_C}^{X}\|\psi_{C,x_C}^{X}) \\
&\geq \mathop{\mathbb{E}}_{C,x_C \leftarrow \varphi^{CX_C}} \sum_{i \notin C} S(\varphi_{C,x_C}^{X_{(i,\cdot)}}\|\psi_{C,x_C}^{X_{(i,\cdot)}}) \\
&= \mathop{\mathbb{E}}_{C,x_C \leftarrow \varphi^{CX_C}} \sum_{i \notin C} S(\varphi_{C,x_C}^{X_{(i,\cdot)}}\|\mu).
\end{aligned}
$$

(C) Mutual information is small.

▶ **Claim 5.4.** *Fix a $j \in [k]$, and fix a C, x_C. There exists a state $\sigma_{C,x_C}^{Z_{-j}}$ such that*

$$
S_\infty(\psi_{C,x_C}^{X_{(\cdot,j)}Z_{-j}}\|\psi_{C,x_C}^{X_{(\cdot,j)}} \otimes \sigma_{C,x_C}^{Z_{-j}}) \leq 2T,
$$

where $Z_{-j} = X_{(\cdot,-j)}X'_{(\cdot,-j)}E_{-j}A_{(\cdot,-j)}$.

We defer the proof of this claim for later, and will assume it for now. Line (5.1) with Fact 3.8 implies that for all j, $\mathbb{E}_{C,x_C \leftarrow \varphi^{CX_C}} S(\varphi_{C,x_C}^{X_{(\cdot,j)}Z_{-j}}\|\psi_{C,x_C}^{X_{(\cdot,j)}Z_{-j}}) \leq \log 1/\lambda$. Using Fact 3.14 with Claim 5.4, we get that for all j, there exists a $\sigma_{C,x_C}^{Z_{-j}}$ such that

$$
\mathop{\mathbb{E}}_{C,x_C \leftarrow \varphi^{C,x_C}} S(\varphi_{C,x_C}^{X_{(\cdot,j)}Z_{-j}}\|\psi_{C,x_C}^{X_{(\cdot,j)}} \otimes \sigma_{C,x_C}^{Z_{-j}}) \leq 2T + \log 1/\lambda.
$$

Using Quantum Raz's Lemma, we get

$$
\mathop{\mathbb{E}}_{C,x_C \leftarrow \varphi^{C,x_C}} \mathop{\mathbb{E}}_{i \in [n]} I(X_{(i,j)} : Z_{-j})_{\varphi_{C,x_C}} \leq 2(\log 1/\lambda + 2T)/n.
$$

By Markov's inequality, we have that there exists a $C, x_C, i \notin C$ such that
1. Measuring the $X_{(i,\cdot)}A_{(i,\cdot)}$ register of φ_{C,x_C} yields a tuple $(x_{(i,\cdot)}, a_{(i,\cdot)})$ that satisfies $V(x_{(i,\cdot)}, a_{(i,\cdot)}) = 1$ with probability at least $1 - \epsilon/8$.
2. $S(\varphi_{C,x_C}^{X_{(i,\cdot)}}\|\mu) \leq 32t/n$.
3. For all $j \in [k]$, $I(X_{(i,j)} : Z_{-j})_{\varphi_{C,x_C}} \leq 64k(t + 2T)/n$.

Let $\delta = 64k(t + 2T)/n$. Let s be the maximum number of qubits output by any one player in game G, so $ks \geq \log |\mathcal{A}|$. Then the total communication is $T = O(ks(t + \log k^2/\epsilon)/\sqrt{\epsilon})$. Then, if $t \leq c\epsilon^{3/2}n/(k^4 s)$ for some universal constant c, we have $\delta \leq \epsilon/32k^2$. Let $\varphi = \varphi_{C,x_C}$. This yields the state and coordinate i required. ◀

5.2 The search protocol

Next, we detail the search protocol used to construct $|\psi\rangle$ and $|\varphi\rangle$. We describe the protocol for a two-player game G; the extension to k parties is straightforward.

Let $G = (\mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B}, \mu, V)$ be a two-player free game, where $\mathcal{X}$ and $\mathcal{Y}$ are Alice and Bob's input alphabets, respectively, and $\mathcal{A}$ and $\mathcal{B}$ are their output alphabets. Consider the optimal strategy for $G^{\otimes n}$, where there is a shared state $|\xi\rangle^{E_A E_B}$ where on input $(x, y) \in \mathcal{X}^n \times \mathcal{Y}^n$, Alice and Bob apply measurements $\{M_a^x\}_{a \in \mathcal{A}^n}$ and $\{N_b^y\}_{b \in \mathcal{B}^n}$ respectively on their share of $|\xi\rangle$.

At the start of the search protocol, a multiset $C = \{i_1, \ldots, i_h\}$, $x_C \in \mathcal{X}^C$, and $y_C \in \mathcal{Y}^C$ are publically visible to Alice and Bob. They are both given the state

$$\left|\psi_{C,x_C,y_C}^0\right\rangle = \sum_{x \in \mathcal{X}^n, y \in \mathcal{Y}^n} \sqrt{\mu^{\otimes n}(x, y | x_C, y_C)} |xxyy\rangle^{XX'YY'} |\xi\rangle^{E_A E_B} |0\rangle^R$$

where $\mu^{\otimes n}(x, y | x_C, y_C)$ is the distribution of (x, y) conditioned on x_C, y_C. Alice has access to registers $XX'E_A R$, and Bob has access to registers $E_B YY'$.

Then, Alice and Bob apply their measurements from the optimal strategy, controlled on the X and Y registers, respectively, to obtain

$$\left|\psi_{C,x_C,y_C}^1\right\rangle = \sum_{x \in \mathcal{X}^n, y \in \mathcal{Y}^n} \sqrt{\mu^{\otimes n}(x, y | x_C, y_C)} |xxyy\rangle \sum_{a \in \mathcal{A}^n, b \in \mathcal{B}^n} |\xi_{xyab}\rangle |ab\rangle |0\rangle$$

where $|\xi_{xyab}\rangle = (\sqrt{M_a^x} \otimes \sqrt{N_b^y}) |\xi\rangle$.

Alice and Bob then run a distributed search protocol controlled on the $XYAB$ registers. Fix (x, y, a, b). The protocol proceeds as follows: Alice and Bob divide the multiset C into groups $D_1, \ldots, D_q$, each of size $m = \lceil 1/\epsilon' \rceil$. For each $\ell = 1, \ldots, q$, Alice and Bob perform a distributed version of the Aaronson-Ambainis 3-dimensional search algorithm [1] to determine whether there is a coordinate D_ℓ contains a losing coordinate – i.e., a coordinate $i \in D_\ell$ such that $V(x_i, y_i, a_i, b_i) = 0$.

The search protocol for a group D_ℓ works as follows. Whenever the Aaronson-Ambainis algorithm is in the state $\sum_i \gamma_{i,z} |i, z\rangle$, where $|i\rangle$ corresponds to an index in D_ℓ, and $|z\rangle$ is a qubit indicating whether a marked item has been found, the joint state between Alice and Bob will be $\sum_i \gamma_{i,z} |i\rangle \otimes |z\rangle \otimes |i\rangle$, where Alice holds the first $|i\rangle$ and $|z\rangle$, and Bob holds the second $|i\rangle$. Thus, Alice and Bob query locations are "synchronized". When Aaronson-Ambainis algorithm has to perform a query controlled on $|i\rangle$, Bob sends the qubit containing $|b_i\rangle$. Alice, controlled on $|b_i\rangle$, performs $|z\rangle \mapsto |z \oplus V(x_i, y_i, a_i, b_i) \oplus 1\rangle$ – note that Alice can perform this, because in addition to x_i, a_i, and b_i, she also has access to y_i because y_C is public. We perform an additional XOR with 1 because a "marked item" for the search algorithm corresponds to a *losing* coordinate. Alice then sends back $|b_i\rangle$ to Bob. The other non-query transformations of the Aaronson-Ambainis algorithm are handled as in the the protocol described in [1]. Each step of the algorithm incurs at most $O(\log |\mathcal{B}|)$ qubits of communication, and there are $O(\sqrt{m})$ steps, resulting in $O(\sqrt{m} \log |\mathcal{B}|)$ qubits of total communication. If D_ℓ contains a losing coordinate, then this protocol will succeed in finding one with probability at least $2/3$.

If for at least one ℓ, Alice and Bob find a losing coordinate for G_ℓ, Alice sets the R register to 0; otherwise, it sets it to 1. Thus the total amount of communication of this protocol is $T = O(q\sqrt{m} \log |\mathcal{B}|) = O(\sqrt{1/\epsilon'} \log 1/\eta \log |\mathcal{B}|)$. The final state of the protocol looks like

$$\left|\psi_{C,x_C,y_C}\right\rangle = \sum_{x,y} \sqrt{\mu^{\otimes n}(x, y | x_C, y_C)} |xxyy\rangle \sum_{a,b} |\xi'_{xyab}\rangle |ab\rangle (\alpha_{Cxyab} |1\rangle + \beta_{Cxyab} |0\rangle),$$

where $\left|\xi'_{xyab}\right\rangle = |\xi_{xyab}\rangle \otimes |w_{Cxyab}\rangle$ with $|w_{Cxyab}\rangle$ denoting the workspace qubits of the two players.

Fix a setting of the registers $XYAB = (x, y, a, b)$. Suppose there was no $i \in [n]$ such that $V(x_i, y_i, a_i, b_i) = 0$. Then the search algorithm will never find a losing coordinate in any of the G_ℓ's, so for all C, the we have $\beta_{Cxyab} = 0$. On the other hand, suppose there were at least $\epsilon'n$ losing coordinates. We analyze, for a fixed (x, y, a, b), the error quantity $\sum_C p(C) |\alpha_{Cxyab}|^2$. We can write $p(C) = \prod_\ell p(D_\ell)$, because each index in C is chosen uniformly and independently at random. Furthermore, we can decompose $|\alpha_{Cxyab}|^2 = \prod_\ell |\alpha_{D_\ell xyab}|^2$, where $\alpha_{D_\ell xyab}$ is the probability amplitude that the Aaronson-Ambainis protocol does not find a losing coordinate in D_ℓ. Thus the error quantity can be written as $\prod_\ell \sum_{D_\ell} p(D_\ell)|\alpha_{D_\ell xyab}|^2 = (\sum_D p(D)|\alpha_{Dxyab}|^2)^q$. Each D_ℓ independently has at least $1 - (1 - \epsilon')^m \geq 1 - 1/e$ probability of containing a losing coordinate. When D_ℓ has a losing coordinate, the Aaronson-Ambainis search protocol will succeed in finding it with probability at least $2/3$. Thus the error quantity is at most

$$\big(\Pr(D \text{ contains losing coordinate}) \cdot (1/3) + \Pr(D \text{ does not have losing coordinate}) \big)^q$$
$$\leq (1/3 + 1/e)^q$$
$$= \exp(-\Omega(q)) = \eta.$$

This establishes the requisite properties of the search protocol in the case of $k = 2$.

The extension to general k parties is straightforward. At the beginning of the protocol, a multiset $C = D_1 \cdots D_q$ and inputs x_C are publically visible to all players. They start with an analogous initial state $|\psi_{C,x_C}\rangle$, where each player j has access to registers $X_{(\cdot,j)}X'_{(\cdot,j)}E_jA_{(\cdot,j)}$; player 1 also has access to register R. They perform the distributed Aaronson-Ambainis protocol independently on all D_ℓ. There are $k - 1$ communication channels, one between the first player and all the other players. Whenever a query is to be made, player $j \in \{2, \ldots, k\}$ sends her answer symbol $a_{(i,j)}$ the first player, who then computes $V(x_{(i,\cdot)}, a_{(i,\cdot)})$. The other non-query transformations of the algorithm are also easily extended to the multiplayer case. The total communication is $T = O(q\sqrt{m} \log |\mathcal{A}|) = O(\sqrt{1/\epsilon'} \log 1/\eta \log |\mathcal{A}|)$, where $\mathcal{A}$ is the output alphabet for all k players.

Proof of Claim 5.4. Fix a C, x_C. Fix a player $j \in [k]$. Take the start state ψ_{C,x_C}^0 defined above (extended appropriately to k players), and trace out the $X'_{(\overline{C},j)}$ register: $\theta_{C,x_C}^0 = \mathrm{tr}_{X'_{(\overline{C},j)}}(\psi_{C,x_C}^0)$. Since $\mu^{\otimes n}$ is a product distribution across players and also across game coordinates, we have that $\theta_{C,x_C}^0 = U^{X_{(\overline{C},j)}} \otimes \phi_{C,x_C}^0$ where $U^{X_{(\overline{C},j)}}$ is the maximally mixed state for the register $X_{(\overline{C},j)}$, and

$$\big|\phi_{C,x_C}^0\big\rangle = |x_C\rangle^{X_C X'_C} \sum_{x_{(\cdot,-j)}} \sqrt{\mu_{-j}^{\otimes n}(x_{(\cdot,-j)}|x_C)} \big|x_{(\cdot,-j)}x_{(\cdot,-j)}\big\rangle^{X_{(\overline{C},-j)} X'_{(\overline{C},-j)}} |\xi\rangle^E |0\rangle^R$$

where $\mu_{-j}^{\otimes n}$ denotes the marginal distribution of $\mu^{\otimes n}$ on all players inputs, except for the jth player. Here, we used the simplifying assumption that μ is the uniform distribution. The search protocol described above never interacts with the $X'_{\overline{C}}$ registers. Thus, we can view the protocol as the jth player receiving a uniformly random input drawn from, $U^{X_{(\overline{C},j)}}$, and shares an entangled state ϕ_{C,x_C}^0 with players $[k] - \{j\}$. The rest of the protocol is some two-way communication between player j and every one else. ◄

We now wish to analyze the min-entropy of player j's input register $X_{\overline{C},j}$ relative to the state of all other players. We appeal to the beautiful result of Nayak and Salzman [17], whose theorem statement we reproduce here:

▶ **Theorem 5.5** ([17]). *Consider a communication protocol, without prior entanglement, where Alice receives a uniformly random n-bit input X, and interacts with Bob over a quantum communication channel. Let ψ^{XB} be the final joint state of Alice's input X and Bob's state in the protocol. Then, for any measurement strategy $\{M_x\}_x$ that Bob applies to his own state, the probability that Bob guesses Alice's input X correct is at most $2^{2m_A}/2^n$, where m_A is the number of qubits sent from Alice to Bob over the course of the protocol.*

We now rephrase their theorem to use relative min-entropy instead of guessing probabilities. Let α be the optimal guessing probability for Bob. Then, the *quantum conditional min-entropy* $H_{\min}(X|B)_\psi$ is defined to be $-\log\alpha$. However, by SDP duality, we have the alternative characterization that $H_{\min}(X|B)_\psi = -\inf_{\sigma^B} S_\infty(\psi^{XB}\|\mathsf{id}^X \otimes \sigma^B)$ [16]. Let σ^B be a state achieving this infimum. Then $\log\alpha = S_\infty(\psi^{XB}\|\mathsf{id}^X \otimes \sigma^B) = S_\infty(\psi^{XB}\|\frac{1}{2^n}\mathsf{id}^X \otimes \sigma^B) - n$. By the theorem of Nayak and Salzman, $\log\alpha \le 2m_A - n$, so $S_\infty(\psi^{XB}\|\frac{1}{2^n}\mathsf{id}^X \otimes \sigma^B) = S_\infty(\psi^{XB}\|\psi^X \otimes \sigma^B) \le 2m_A$, where we used the fact that ψ^X is the uniform distribution.

To apply this theorem to our setting, we can treat player j as "Alice" and the rest of the players as "Bob". Alice exchanges at most T qubits with Bob. The crucial component of the Nayak-Salzman theorem is that Bob's probability of guessing does not depend on how many qubits he sent to Alice! Thus, we can imagine that in the beginning of the protocol he sent the E_j register of the shared entangled state ϕ^0_{C,x_C} to Alice first. We have that there exists a state $\sigma^{Z_{-j}}_{C,x_C}$ such that

$$2T \ge S_\infty(\psi^{X_{(\cdot,j)}Z_{-j}}_{C,x_C}\|\psi^{X_{(\cdot,j)}}_{C,x_C} \otimes \sigma^{Z_{-j}}_{C,x_C}).$$

6 Open problems

We conclude with a variety of open problems.

1. Is it possible to extend the Grover search analysis to handle CQ games?

2. Is *strong parallel repetition* possible with the entangled value of free games? In other words, can the base of $1 - \epsilon^{3/2}$ of Theorem 1.1 be improved to $1 - \epsilon$?

3. Is the base of $1 - \epsilon^2$ for the repeated *classical* value of free games tight? If so, this would mean that there is a separation of classical and quantum parallel repetition for free games.

4. It was shown by [10] that the dependence on the output alphabet size, for classical parallel repetition, is necessary – even for free games. However, Holenstein showed the repeated game value for non-signaling games has no such alphabet dependence [12]. Is this dependence necessary for the quantum case?

5. Can we identify an interesting class of games for which we can prove improved parallel repetition theorems, by designing efficient communication protocols to generate advice states?

6. The mantra, "Better parallel repetition theorems from better communication protocols," suggests an intriguing connection between games and communication protocols. Although games are protocols that forbid communication between the players, one can define the *communication complexity of a game* as the minimum communication needed for the players to determine whether they have won or lost the game. Our mantra suggests a relationship between the value and communication complexity of a game. What is the nature of this relationship?

7. Can one use these techniques to prove parallel repetition for entangled games with an arbitrary input distribution?

--- **References** ---

1   Scott Aaronson and Andris Ambainis. Quantum search of spatial regions. In *Foundations of Computer Science, 2003. Proceedings. 44th Annual IEEE Symposium on*, pages 200–209. IEEE, 2003.

2   Rotem Arnon-Friedman, Renato Renner, and Thomas Vidick. Non-signalling parallel repetition using de finetti reductions. *arXiv preprint arXiv:1411.1582*, 2014.

3   Boaz Barak, Anup Rao, Ran Raz, Ricky Rosen, and Ronen Shaltiel. Strong parallel repetition theorem for free projection games. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 352–365. Springer, 2009.

4   Harry Buhrman, Richard Cleve, and Avi Wigderson. Quantum vs. classical communication and computation. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, pages 63–68. ACM, 1998.

5   Harry Buhrman, Serge Fehr, and Christian Schaffner. On the parallel repetition of multiplayer games: The no-signaling case. *arXiv preprint arXiv:1312.7455*, 2013.

6   André Chailloux and Giannicola Scarpa. Parallel repetition of entangled games with exponential decay via the superposed information cost. In Javier Esparza, Pierre Fraigniaud, Thore Husfeldt, and Elias Koutsoupias, editors, *Automata, Languages, and Programming – 41st International Colloquium, ICALP 2014, Copenhagen, Denmark, July 8-11, 2014, Proceedings, Part I*, volume 8572 of *Lecture Notes in Computer Science*, pages 296–307. Springer, 2014.

7   André Chailloux and Giannicola Scarpa. Parallel repetition of free entangled games: Simplification and improvements. *arXiv preprint arXiv:1410.4397*, 2014.

8   Kai-Min Chung, Xiaodi Wu, and Henry Yuen. Parallel repetition for entangled k-player games via fast quantum search. *arXiv preprint arXiv:1501.00033*, 2015.

9   Irit Dinur, David Steurer, and Thomas Vidick. A parallel repetition theorem for entangled projection games. In *IEEE 29th Conference on Computational Complexity, CCC 2014, Vancouver, BC, Canada, June 11-13, 2014*, pages 197–208, 2014.

10   Uriel Feige and Oleg Verbitsky. Error reduction by parallel repetition: a negative result. *Combinatorica*, 22(4):461–478, 2002.

11   Joseph Fitzsimons and Thomas Vidick. A multiprover interactive proof system for the local hamiltonian problem. *arXiv preprint arXiv:1409.0260*, 2014.

12   Thomas Holenstein. Parallel repetition: simplifications and the no-signaling case. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, pages 411–419. ACM, 2007.

13   Rahul Jain, Attila Pereszlényi, and Penghui Yao. A parallel repetition theorem for entangled two-player one-round games under product distributions. In *IEEE 29th Conference on Computational Complexity, CCC 2014, Vancouver, BC, Canada, June 11-13, 2014*, pages 209–216, 2014.

14   Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. A lower bound for the bounded round quantum communication complexity of set disjointness. In *Foundations of Computer Science, 2003. Proceedings. 44th Annual IEEE Symposium on*, pages 220–229. IEEE, 2003.

15   Julia Kempe and Thomas Vidick. Parallel repetition of entangled games. In *Proceedings of the 43rd Annual ACM Symposium on Theory of Computing*, pages 353–362. ACM, 2011.

16   Robert König, Renato Renner, and Christian Schaffner. The operational meaning of min- and max-entropy. *IEEE Transactions on Information Theory*, 55(9):4337–4347, 2009.

17   Ashwin Nayak and Julia Salzman. Limits on the ability of quantum states to convey classical messages. *Journal of the ACM (JACM)*, 53(1):184–206, 2006.

18   Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information.* Cambridge university press, 2010.

19   Ran Raz. A parallel repetition theorem. *SIAM Journal on Computing*, 27(3):763–803, 1998.

20   Ran Raz. A counterexample to strong parallel repetition. *SIAM Journal on Computing*, 40(3):771–777, 2011.

21   Ricky Rosen. A k-provers parallel repetition theorem for a version of no-signaling model. *Discrete Math., Alg. and Appl.*, 2(4):457–468, 2010.

22   Mark M. Wilde. *Quantum information theory.* Cambridge University Press, 2013.

23   Dong Yang, Karol Horodecki, Michal Horodecki, Pawel Horodecki, Jonathan Oppenheim, and Wei Song. Squashed entanglement for multipartite states and entanglement measures based on the mixed convex roof. *IEEE Transactions on Information Theory*, 55(7):3375–3387, 2009.