Strong Parallel Repetition Theorem for Free Projection Games

Boaz Barak * Anup Rao † Ran Raz ‡ Ricky Rosen § Ronen Shaltiel ¶

April 14, 2009

Abstract

The parallel repetition theorem states that for any two provers one round game with value at most $1-\epsilon$ (for $\epsilon < 1/2$), the value of the game repeated n times in parallel is at most $(1-\epsilon^3)^{\Omega(n/\log s)}$ where s is the size of the answers set [Raz98],[Hol07]. For *Projection Games* the bound on the value of the game repeated n times in parallel was improved to $(1-\epsilon^2)^{\Omega(n)}$ [Rao08] and was shown to be tight [Raz08]. In this paper we show that if the questions are taken according to a product distribution then the value of the repeated game is at most $(1-\epsilon^2)^{\Omega(n/\log s)}$ and if in addition the game is a *Projection Game* we obtain a *strong parallel repetition* theorem, i.e., a bound of $(1-\epsilon)^{\Omega(n)}$.

1 Introduction

In a two provers one round game there are two *provers* and a *verifier*. The verifier selects randomly $(x,y) \in X \times Y$, a question for each prover, according to some distribution P_{XY} where X is the questions set of prover 1 and Y is the questions set of prover 2. Each prover knows only the question addressed to her, prover 1 knows only x and prover 2 knows only y. The provers cannot communicate during the transaction. The provers send their answers to the verifier, $a = a(x) \in A$ and $b = b(y) \in B$ where A is the answers set of the first prover and B is the answers set of the second prover. The verifier evaluates an acceptance predicate $V(x,y,a,b)$ and accepts or rejects based on the outcome of the predicate. The acceptance predicate as well as the distribution of the questions are known in advance to the provers. The provers answer the questions according to a strategy which is a pair of functions $f_a : X \to A,\ f_b : Y \to B$. The strategy of the provers is also called a protocol. If $\mathrm{P}_{XY} = \mathrm{P}_X \cdot \mathrm{P}_Y$, that is P_{XY} is a product distribution, we say that the game is a *free game*.

The *value* of the game is the maximum of the probability that the verifier accepts, where the maximum is taken over all the provers strategies. More formally, the value of the game is:

$$\max_{f_a, f_b} \mathbb{E}_{xy}\left[V\left(x, y, f_a(x), f_b(y)\right)\right]$$

*Department of Computer Science, Princeton University. Supported by NSF grants CNS-0627526, CCF-0426582 and CCF-0832797, US-Israel BSF grant 2004288 and Packard and Sloan fellowships.

†Institute for Advanced Study.

‡Faculty of mathematics and computer science, Weizmann Institute.

§Department of Computer Science, Tel-Aviv University.

¶Department of Computer Science, Haifa University.

where the expectation is taken with respect to the distribution P_{XY}.

Roughly speaking, the *n-fold parallel repetition* of a game G is a game in which the provers try to win simultaneously n copies of G and it is denoted by $G^{\otimes n}$. More precisely, the verifier sends n questions to each prover, $(x_1, x_2 \ldots, x_n)$ to prover 1 and $(y_1, y_2 \ldots, y_n)$ to prover 2 where for all i, (x_i, y_i) is distributed according to P_{XY} and is independent of the other questions. The provers generate n answers, $(a_1, a_2 \ldots, a_n)$ by prover 1 and $(b_1, b_2 \ldots, b_n)$ by prover 2. The verifier evaluates the acceptance predicate on each coordinate and accepts if and only if all the predicates accept, namely if and only if $V^{\otimes n} = \wedge_{i=1}^{n} V(x_i, y_i, a_i, b_i) = 1$. Note that the verifier treats each of the n games independently, but the provers may not; the answer of each question addressed to a prover may depend on all the questions addressed to that prover. There are examples of games where the value of the game repeated n times in parallel is strictly larger than the value of the original game to the power of n [For89], [FV02], [Raz08].

The Parallel Repetition Theorem

A series of papers deal with the nature of the value decrease of games repeated n times in parallel. The parallel repetition theorem of Raz [Raz98] states that for every game G with value at most $1 - \epsilon$ where $\epsilon < 1/2$, the value of $G^{\otimes n}$ is at most $(1 - \epsilon^{32})^{\Omega(n/\log s)}$ where s is the size of the answers support $s = |A \times B|$. In a recent elegant result, Holenstein [Hol07] improved the bound to $(1 - \epsilon^3)^{\Omega(n/\log s)}$ while simplifying the proof of [Raz98]. Subsequently, for the important special type of games known as *projection games*, Rao [Rao08] proved a bound of $(1-\epsilon^2)^{\Omega(n)}$ (for a special type of projection games known as *XOR games* such a bound was previously proven by Feige, Kindler and O'Donnell [FKO07]). Note that Rao's [Rao08] bound does not depend on the size of the answers set, s. In the general case, Feige and Verbitsky [FV02] showed that the dependency on s is tight (up to loglog factors).

Many researchers studied the problem of whether there exists a strong parallel repetition theorem in the general case or at least in some important special cases. Namely, is it the case that for a given game G of value $1 - \epsilon$, say, for $\epsilon < 1/2$, the value of $G^{\otimes n}$ is at most $(1 - \epsilon)^{\Omega(n/\log s)}$? This question was motivated by connections to hardness of approximation as well as connections to problems in geometry [FKO07], [SS07]. A recent result of Raz [Raz08] showed a counterexample for the general case, as well as for the case of projection games, unique games and XOR games. Raz [Raz08] showed that there is an example of a XOR game (thus also projection game and unique game) of value $1-\epsilon$ such that for large enough n, the value of the game is at least $(1-\epsilon^2)^{O(n)}$. For some extensions, generalization and applications see Barak, Hardt, Haviv, Rao, Regev and Steurer [BHH+08], Kindler, O'Donnell, Rao and Wigderzon [KORW08] and Alon and Klartag [AK08].

Other related results: For the special case of unique games played on expander graphs Arora, Khot, Kolla, Steurer, Tulsiani and Vishnoi [AKK+08] proved an "almost" strong parallel repetition theorem (strong up to a polylogarithmic factor). For the special case of games where the roles of the two players are symmetric and the game is played on an expander graph that contains a self loop on every vertex, Safra and Schwartz [SS07] showed that $O(1/\epsilon)$ repetitions are sufficient to reduce the value of the game from $1 - \epsilon$ to some constant.

In this paper we prove a strong parallel repetition theorem for free projection games and we improve the known bound for every free game. More precisely:

1. For every **Free game** of value $\leq (1-\epsilon)$ for $\epsilon < 1/2$, the value of $G^{\otimes n}$ is at most $(1-\epsilon^2)^{\Omega(n/\log s)}$

2. For every **Free Projection game** of value $\leq (1-\epsilon)$ for $\epsilon < 1/2$, the value of $G^{\otimes n}$ is at most $(1-\epsilon)^{\Omega(n)}$

Techniques

The main technical contribution of this paper is the ability to work throughout the whole proof with relative entropy without the need to switch to ℓ_1 norm. In previous results [Raz98], [Hol07], [Rao08] a bound on the distance between a distribution "generated by the provers' strategies" and the original distribution was derived using the relative entropy between the two distributions. This bound was then used to obtain a bound on the ℓ_1 distance between those distributions. This was done using the fact that $\|P - Q\|_1 \leq O(\sqrt{D(P\|Q)})$ where $D(P\|Q)$ is the relative entropy between P and Q. Since the bound is quadratic, there is a loss when using the ℓ_1 norm instead of using directly the relative entropy. We show that for the special case of free games one can redo the whole proof using relative entropy, without switching to ℓ_1 norm. We bound the value of a game by using our Corollary 3.4 (that might be useful for other applications). We note that since we are only considering free games, the proof is simpler than the one for general games and we do not use much of the machinery used in previous results, e.g., [Raz98], [Hol07], [Rao08].

2 Preliminaries

2.1 Notations

General Notations

We denote an n-dimensional vector by a superscript n, e.g., $\phi^n = (\phi_1, \ldots, \phi_n)$ where ϕ_i is the i^{th} coordinate. The function $\log(x)$ is the logarithm base 2 of x. We use the common notation $[n]$ to denote the set $\{1, \ldots, n\}$.

Random Variables and Sets

By slightly abusing notations, we will use capital letters to denote both sets and random variables distributed over these sets, and we will use lower case letters to denote values. For example, X, Y will denote sets as well as random variables distributed over these sets, and x, y will denote values in these sets that the random variables can take. Nevertheless, it will always be clear from the context whether we are referring to sets or random variables. For a random variable Z it will be convenient in some lemmas, such as Lemma 3.7, to think of $\Pr(Z)$ as a random variable.

Random Variables and their Distributions

For a random variable X, we denote by P_X the distribution of X. For an event U we use the notation $P_{X|U}$ to denote the distribution of $X|U$, that is, the distribution of X conditioned on the event U. If Z is an additional random variable that is fixed (e.g., inside an expression where an expectation over Z is taken), we denote by $P_{X|Z}$ the distribution of X conditioned on Z. In the same way, for two (or more) random variables X, Y, we denote their joint distribution by P_{XY}, and we use the same notations as above to denote conditional distributions. For example, for an event U, we write $P_{XY|U}$ to denote the distribution of X, Y conditioned on the event U, i.e., $P_{XY|U}(x, y) = \Pr(X = x, Y = y|U)$. For two (or more) random variables X, Y with distribution P_{XY}, we use the notation P_X to denote the marginal distribution of X.

The Game G

We denote a game by G and define X to be the set of questions to prover 1, Y to be the set of questions to prover 2 and P_{XY} to be the joint distribution according to which the verifier chooses a pair of questions to the provers. We denote by A the set of answers of prover 1 and by B the set of answers of prover 2. We denote the acceptance predicate by V. A game G with acceptance predicate V and questions distribution P_{XY} is denoted by $\mathrm{G}(\mathrm{P}_{XY}, V)$. As mentioned above, we also denote by X, Y, A, B random variables distributed over X, Y, A, B respectively. X, Y will be the questions addressed to the two provers, distributed over the question sets X and Y respectively. Fixing a strategy f_a, f_b for the game G, we can also think of the answers A and B as random variables distributed over the answer sets A and B respectively.

The Game G **Repeated** n **Times**

For the game G repeated n times in parallel, $\mathrm{G}^{\otimes n} = G(\mathrm{P}_{X^nY^n}, V^{\otimes n})$, the random variable X_i denotes the question to prover 1 in coordinate i, and similarly, the random variable Y_i denotes the question to prover 2 in coordinate i. We denote by X^n the tuple $(X_1, \ldots, X_n)$ and by Y^n the tuple $(Y_1, \ldots, Y_n)$. Fixing a strategy f_a, f_b for $\mathrm{G}^{\otimes n}$, the random variable A_i denotes the answer of prover 1 in coordinate i, and similarly, the random variable B_i denotes the answer of prover 2 in coordinate i. We denote by A^n the tuple $(A_1, \ldots, A_n)$ and by B^n the tuple $(B_1, \ldots, B_n)$. It will be convenient in some lemmas to denote $X^k = (X_{n-k+1}, \ldots, X_n)$, i.e., the last k coordinates of X^n and in the same way, $Y^k = (Y_{n-k+1}, \ldots, Y_n)$, $A^k = (A_{n-k+1}, \ldots, A_n)$ and $B^k = (B_{n-k+1}, \ldots, B_n)$. We also denote $X^{n-k} = (X_1, \ldots, X_{n-k})$, i.e., the first $n-k$ coordinates of X^n, and similarly, $Y^{n-k} = (Y_1, \ldots, Y_{n-k})$. For fixed $i \in [n-k]$, we denote $X^m = (X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_{n-k})$, i.e., X^{n-k} without X_i, and similarly, $Y^m = (Y_1, \ldots, Y_{i-1}, Y_{i+1}, \ldots, Y_{n-k})$.

The Event W_i

For the game $\mathrm{G}^{\otimes n} = G(\mathrm{P}_{X^nY^n}, V^{\otimes n})$ and a strategy $f_a : X^n \to A^n, f_b : Y^n \to B^n$ we can consider the joint distribution:

$$\mathrm{P}_{X^n,Y^n,A^n,B^n}(x^n, y^n, a^n, b^n) = \begin{cases} \mathrm{P}_{X^n,Y^n}(x^n, y^n) & \text{if } a^n = f_a(x^n) \text{ and } b^n = f_b(y^n) \\ 0 & \text{otherwise} \end{cases}$$

We define the event W_i to be the event of winning the game in coordinate i, i.e., the event that the verifier accepts on coordinate i. Since the random variables A^n and B^n are functions of X^n and Y^n respectively, we can think of W_i as an event in the random variables X^n, Y^n.

2.2 Special Types of Games

Definition 2.1 (Free Games) *A game is* Free *if the distribution of the questions is a product distribution, i.e.,* $\mathrm{P}_{XY} = \mathrm{P}_X \times \mathrm{P}_Y$

Definition 2.2 (Projection Games) *A* Projection game *is a game where for each pair of questions x, y there is a function $f_{xy} : B \to A$ such that $V(x, y, a, b)$ is satisfied if and only if $f_{xy}(b) = a$.*

2.3 Entropy and Relative Entropy

Definition 2.3 (Entropy) *For a probability distribution ϕ over a sample space Ω we define the entropy of ϕ to be* $\mathrm{H}(\phi) = -\sum_{x\in\Omega}\phi(x)\log\phi(x) = -\mathbb{E}_{x\sim\phi}\log\phi(x) = \mathbb{E}_{x\sim\phi}\log\left(\frac{1}{\phi(x)}\right)$

By applying Jensen's inequality on the concave function $\log(\cdot)$ one can derive the following fact:

Fact 2.4 *For every distribution ϕ over Ω,* $\mathrm{H}(\phi) \leq \log(|\mathrm{supp}(\phi)|)$ *where*

$$\mathrm{supp}(\phi) = \{x \in \Omega | \phi(x) > 0\}$$

Definition 2.5 (Relative Entropy) *We define Relative Entropy, also called the Kullback-Leibler Divergence or simply* divergence. *Let* P *and* Q *be two probability distributions defined on the same sample space Ω. The relative entropy of* P *with respect to* Q *is:*

$$\mathrm{D}(\mathrm{P}\|\mathrm{Q}) = \sum_{x\in\Omega}\mathrm{P}(x)\log\frac{\mathrm{P}(x)}{\mathrm{Q}(x)}$$

where $0\log\frac{0}{0}$ is defined to be 0 and $p\log\frac{p}{0}$ where $p \neq 0$ is defined to be ∞.

Vaguely speaking, we could think of the relative entropy as a way to measure the information we gained by learning that a random variable is distributed according to P when apriority we thought that it was distributed according to Q. This indicates how *far* Q is from P; if we don't gain much information then the two distributions are very *close* in some sense. Note that the relative entropy is not symmetric (and therefore is not a metric).

Fact 2.6 *Let $\Phi^n = \Phi_1 \times \Phi_2 \times \cdots \times \Phi_n$ and let μ^n be any distribution over the same sample space (not necessarily a product distribution) then $\sum_{i=1}^{n}\mathrm{D}(\mu_i\|\Phi_i) \leq \mathrm{D}(\mu^n\|\Phi^n)$ thus $\mathbb{E}_{i\in[n]}\mathrm{D}(\mu_i\|\Phi_i) = \frac{1}{n}\sum_{i\in[n]}\mathrm{D}(\mu_i\|\Phi_i) \leq \frac{\mathrm{D}(\mu^n\|\Phi^n)}{n}$*

3 Our Results

We prove the following theorems:

Theorem 3.1 (Parallel Repetition For Free Games) *For every game* G *with value $1-\epsilon$ where $\epsilon < 1/2$ and $\mathrm{P}_{XY} = \mathrm{P}_X \times \mathrm{P}_Y$ (the questions are distributed according to some product distribution), the value of $\mathrm{G}^{\otimes n}$ is at most $(1-\epsilon^2/9)^{n/(18\log s+3)}$*

Theorem 3.2 (Strong Parallel Repetition For Free Projection Games) *For every projection game* G *with value $1-\epsilon$ where $\epsilon < 1/2$ and $\mathrm{P}_{XY} = \mathrm{P}_X \times \mathrm{P}_Y$ (the questions are distributed according to some product distribution), the value of $\mathrm{G}^{\otimes n}$ is at most $(1-\epsilon/9)^{(n/33)-1}$*

3.1 Technical Lemma

Lemma 3.3 *For every $0 \leq p, q \leq 1$ define binary distributions $\mathrm{P} = (p, 1-p)$ and $\mathrm{Q} = (q, 1-q)$, over $\{0,1\}$, if $\mathrm{D}(\mathrm{P}\|\mathrm{Q}) \leq \delta$ and $p < \delta$ then*

$$q \leq 4\delta$$

Proof: If $\delta \geq \frac{1}{4}$ then the statement is obviously true. For the case that $\delta < \frac{1}{4}$, assume by way of contradiction that $q > 4\delta$. Since for $q > p$, $\mathrm{D}(\mathrm{P}\|\mathrm{Q})$ is decreasing in p and increasing in q,

$$\begin{aligned} \mathrm{D}(\mathrm{P}\|\mathrm{Q}) &= p\log\frac{p}{q} + (1-p)\log\frac{1-p}{1-q} \\ &> \delta\log(\frac{\delta}{4\delta}) + (1-\delta)\log\frac{1-\delta}{1-4\delta} \\ &= -2\delta + (1-\delta)\log\left(1+\frac{3\delta}{1-4\delta}\right) \qquad (1) \end{aligned}$$

If $\delta \geq 1/7$ then $\log\left(1+\frac{3\delta}{1-4\delta}\right) \geq 1$. Thus,

$$(1) \quad \geq \quad -2\delta + (1-\delta) > \delta$$

where the last inequality follows since $\delta < 1/4$.
If $\delta < 1/7$ then $\frac{3\delta}{1-4\delta} < 1$. Using the inequality $\log_2(1+x) \geq x$ for every $0 \leq x \leq 1$ we obtain,

$$(1) \quad \geq \quad -2\delta + (1-\delta)\frac{3\delta}{1-4\delta} \geq -2\delta + 3\delta = \delta$$

where the last inequality follows since $\frac{1-\delta}{1-4\delta} > 1$. Since we obtained a contradiction in both cases, the lemma holds. ■

Corollary 3.4 *For every probability distributions* P,Q *over the same sample space* Ω *and for every* $T \subseteq \Omega$*, if* $\mathrm{D}(\mathrm{P}\|\mathrm{Q}) \leq \delta$ *and* $\mathrm{P}(T) \leq \delta$ *then* $\mathrm{Q}(T) \leq 4\delta$

Proof: Denote $p = \mathrm{P}(T)$ and $q = \mathrm{Q}(T)$ and let $\mathrm{P}' = (p, 1-p)$, $\mathrm{Q}' = (q, 1-q)$. By the data processing inequality for mutual information $\mathrm{D}(\mathrm{P}\|\mathrm{Q}) \geq \mathrm{D}(\mathrm{P}'\|\mathrm{Q}')$ and the corollary follows. ■

3.2 Main Lemmas

We now state the main lemmas for general product distribution games.
Recall that for a coordinate i, W_i is the event of the provers winning the game played in this coordinate.

Lemma 3.5 (Main Lemma For General Free Games) *Let* G *be a free game with value* $1-\epsilon$*. For any set* T *of* k *coordinates,* ($T \subseteq [n]$ *and* $|T| = k$)*, let* W *be the event of the provers winning the games in those* k *coordinates. If* $\Pr(W) \geq 2^{-\epsilon(n-k)/9+k\log s}$ *where* s *is the size of the answers set, then there is* $i \notin T$ *for which*

$$\Pr(W_i|W) \leq 1 - \frac{\epsilon}{9}$$

Lemma 3.6 (Main Lemma For Free Projection Games) *Let* G *be a free projection game with value* $1-\epsilon$*. For any set* T *of* k *coordinates,* ($T \subseteq [n]$ *and* $|T| = k$)*, let* W *be the event of the provers winning the games in those* k *coordinates. If* $\Pr(W) \geq 2^{-\epsilon(n-k)/144}$ *and* $n - k \geq (48/\epsilon)\log(8/\epsilon)$ *then there is* $i \notin T$ *for which*

$$\Pr(W_i|W) \leq 1 - \frac{\epsilon}{9}$$

In the lemmas below we assume without loss of generality that the set T of k coordinates is the set of the last k coordinates. Recall that $\mathrm{P}_{X^nY^n} = P_{XY} \times \cdots \times \mathrm{P}_{XY}$ n-times. Recall that $X^k = (X_{n-k+1}, \ldots, X_n)$, i.e., the last k coordinates of X^n and in the same way, $Y^k = (Y_{n-k+1}, \ldots, Y_n)$, $A^k = (A_{n-k+1}, \ldots, A_n)$ and $B^k = (B_{n-k+1}, \ldots, B_n)$. Recall that $X^{n-k} = (X_1, \ldots, X_{n-k})$, i.e., the first $n-k$ coordinates of X^n, and similarly, $Y^{n-k} = (Y_1, \ldots, Y_{n-k})$.

Lemma 3.7 *For any event*[1] U*, the following holds:*

$$\mathbb{E}_{X^k,Y^k,A^k|U} \mathrm{D}\left(\mathrm{P}_{X^{n-k},Y^{n-k}|X^k,Y^k,A^k,U} \| \mathrm{P}_{X^{n-k},Y^{n-k}}\right) \leq \log\left(\frac{1}{\Pr(U)}\right) + \mathbb{E}_{X^k,Y^k|U} \mathrm{H}(\mathrm{P}_{A^k|X^k,Y^k,U})$$

Proof: Since $\mathrm{P}_{X^nY^n} = \mathrm{P}_{XY} \times \cdots \times \mathrm{P}_{XY}$ n-times,

$$\begin{aligned}
&\mathbb{E}_{X^k,Y^k,A^k|U} \mathrm{D}\left(\mathrm{P}_{X^{n-k},Y^{n-k}|X^k,Y^k,A^k,U} \| \mathrm{P}_{X^{n-k},Y^{n-k}}\right) \\
&= \mathbb{E}_{X^k,Y^k,A^k|U} \mathrm{D}\left(\mathrm{P}_{X^{n-k},Y^{n-k}|X^k,Y^k,A^k,U} \| \mathrm{P}_{X^{n-k},Y^{n-k}|X^k,Y^k}\right) \\
&= \mathbb{E}_{X^k,Y^k,A^k|U} \mathbb{E}_{X^{n-k},Y^{n-k}|X^k,Y^k,A^k,U} \log\left(\frac{\Pr(X^{n-k},Y^{n-k}|X^k,Y^k,A^k,U)}{\Pr(X^{n-k},Y^{n-k}|X^k,Y^k)}\right) \\
&= \mathbb{E}_{X^k,Y^k,A^k|U} \mathbb{E}_{X^{n-k},Y^{n-k}|X^k,Y^k,A^k,U} \log\left(\frac{\Pr(X^{n-k},Y^{n-k},X^k,Y^k,A^k,U)\Pr(X^k,Y^k)}{\Pr(X^k,Y^k,A^k,U)\Pr(X^{n-k},Y^{n-k},X^k,Y^k)}\right) \\
&= \mathbb{E}_{X^k,Y^k,A^k,X^{n-k},Y^{n-k}|U} \log\left(\frac{\Pr(X^{n-k},Y^{n-k},X^k,Y^k,A^k,U)}{\Pr(X^{n-k},Y^{n-k},X^k,Y^k)}\right) \\
&+ \mathbb{E}_{X^k,Y^k,A^k,X^{n-k},Y^{n-k}|U} \log\left(\frac{\Pr(X^k,Y^k)}{\Pr(X^k,Y^k,A^k,U)}\right)
\end{aligned} \tag{2}$$

Since $\Pr(X^{n-k},Y^{n-k},X^k,Y^k,A^k,U) \leq \Pr(X^{n-k},Y^{n-k},X^k,Y^k)$ the term

$$\log\left(\frac{\Pr(X^{n-k},Y^{n-k},X^k,Y^k,A^k,U)}{\Pr(X^{n-k},Y^{n-k},X^k,Y^k)}\right) \leq 0$$

Therefore,

$$\begin{aligned}
(2) &\leq \mathbb{E}_{X^k,Y^k,A^k,X^{n-k},Y^{n-k}|U} \log\left(\frac{\Pr(X^k,Y^k)}{\Pr(X^k,Y^k,A^k,U)}\right) \\
&= \mathbb{E}_{X^k,Y^k,A^k,X^{n-k},Y^{n-k}|U} \log\left(\frac{1}{\Pr(A^k,U|X^k,Y^k)}\right)
\end{aligned} \tag{3}$$

[1]We will use the lemma for events that depend only on X^k, Y^k, A^k, B^k, e.g., we will use it for the event W, see definition in Lemma 3.5

Since $\Pr(A^k = a^k, U | X^k = x^k, Y^k = y^k)$ is a function of only a^k, x^k, y^k (and not of x^{n-k}, y^{n-k}) we obtain:

$$(3) = \mathbb{E}_{X^k,Y^k,A^k|U} \log\left(\frac{1}{\Pr(U|X^k,Y^k)}\right) + \mathbb{E}_{X^k,Y^k,A^k|U} \log\left(\frac{1}{\Pr(A^k|X^k,Y^k,U)}\right) \tag{4}$$

In the same way,

$$(4) = \mathbb{E}_{X^k,Y^k|U} \log\left(\frac{1}{\Pr(U|X^k,Y^k)}\right) + \mathbb{E}_{X^k,Y^k|U}\mathbb{E}_{A^k|X^k,Y^k,U} \log\left(\frac{1}{\Pr(A^k|X^k,Y^k,U)}\right)$$

$$= \sum_{x^k,y^k \in \mathrm{supp}(\mathrm{P}_{X^k,Y^k|U})} \Pr(X^k = x^k, Y^k = y^k|U) \log\left(\frac{1}{\Pr(U|X^k = x^k, Y^k = y^k)}\right) + \mathbb{E}_{X^k,Y^k|U}\mathrm{H}(\mathrm{P}_{A^k|X^k,Y^k,U}) \tag{5}$$

By the concavity of $\log(\cdot)$,

$$\begin{aligned}
(5) &\leq \log\left(\sum_{x^k,y^k \in \mathrm{supp}(\mathrm{P}_{X^k,Y^k|U})} \frac{\Pr(X^k = x^k, Y^k = y^k|U)}{\Pr(U|X^k = x^k, Y^k = y^k)}\right) + \mathbb{E}_{X^k,Y^k|U}\mathrm{H}(\mathrm{P}_{A^k|X^k,Y^k,U}) \\
&= \log\left(\sum_{x^k,y^k \in \mathrm{supp}(\mathrm{P}_{X^k,Y^k|U})} \frac{\Pr(X^k = x^k, Y^k = y^k)}{\Pr(U)}\right) + \mathbb{E}_{X^k,Y^k|U}\mathrm{H}(\mathrm{P}_{A^k|X^k,Y^k,U}) \\
&\leq \log\left(\frac{1}{\Pr(U)}\right) + \mathbb{E}_{X^k,Y^k|U}\mathrm{H}(\mathrm{P}_{A^k|X^k,Y^k,U})
\end{aligned}$$

■

We define W to be the event that the provers win all the games in the last k coordinates and define E to be $\left\{(a^k, x^k, y^k) \in A^k \times X^k \times Y^k \;\middle|\; \Pr(A^k = a^k | X^k = x^k, Y^k = y^k) \geq 2^{-\epsilon(n-k)/16}\right\}$. The event W' is defined as $W \wedge [(A^k, X^k, Y^k) \in E]$

Proposition 3.8 *For W and W', the events defined above, the following holds:*

1. *For general games and the event W*

$$\mathbb{E}_{X^k,Y^k|W}\mathrm{H}\left(\mathrm{P}_{A^k|X^k,Y^k,W}\right) \leq k \log s \qquad \textit{[Raz98],[Hol07]}$$

2. *For projection games and the event W'*

$$\mathbb{E}_{X^k,Y^k|W'}\mathrm{H}\left(\mathrm{P}_{A^k|X^k,Y^k,W'}\right) \leq \epsilon(n-k)/16 \qquad \textit{[Rao08]}$$

Proof for general games: We use the trivial bound on the size of the support, namely, for every x^k, y^k we can bound $|\mathrm{supp}(\mathrm{P}_{A^k|X^k=x^k,Y^k=y^k,W})| \leq |\mathrm{supp}(\mathrm{P}_{A^k})| \leq s^k$ where s is the size of the answers set. Using Fact 2.4 we obtain:

$$\mathbb{E}_{X^k,Y^k|W}\mathrm{H}\left(\mathrm{P}_{A^k|X^k,Y^k,W}\right) \leq \mathbb{E}_{X^k,Y^k|W} \log(|\mathrm{supp}(\mathrm{P}_{A^k|X^k,Y^k,W})|) \leq \log s^k = k \log s$$

■

Proof for projection games: Using Fact 2.4 we can trivially bound:

$$\mathbb{E}_{X^k,Y^k|W'}\mathrm{H}\left(\mathrm{P}_{A^k|X^k,Y^k,W'}\right) \leq \mathbb{E}_{X^k,Y^k|W'}\log(|\mathrm{supp}(\mathrm{P}_{A^k|X^k,Y^k,W'})|) \tag{6}$$

Since for every x^k, y^k and $a^k \in \mathrm{supp}(\mathrm{P}_{A^k|X^k=x^k,Y^k=y^k,W'})$,

$$\Pr(A^k = a^k | X^k = x^k, Y^k = y^k) \geq 2^{-\epsilon(n-k)/16},$$

there are at most $2^{\epsilon(n-k)/16}$ such a^k. Hence,

$$(6) \leq \mathbb{E}_{X^k,Y^k|W'}\log\left(2^{\epsilon(n-k)/16}\right) = \epsilon(n-k)/16$$

■

Corollary 3.9 *For the events W, W' the following holds:*

1. *For general games and the event W*

$$\mathbb{E}_{i\in[n-k]}\mathbb{E}_{X^k,Y^k,A^k|W}\mathrm{D}\left(\mathrm{P}_{X_i,Y_i|X^k,Y^k,A^k,W}\|\mathrm{P}_{X_i,Y_i}\right) \leq \frac{1}{n-k}\left(k\log s - \log(\Pr(W))\right)$$

2. *For projection games and the event W'*

$$\mathbb{E}_{i\in[n-k]}\mathbb{E}_{X^k,Y^k,A^k|W'}\mathrm{D}\left(\mathrm{P}_{X_i,Y_i|X^k,Y^k,A^k,W'}\|\mathrm{P}_{X_i,Y_i}\right) \leq \frac{1}{n-k}\left(\ \epsilon(n-k)/16 - \log\left(\Pr(W) - 2^{-\epsilon(n-k)/16}\right)\ \right)$$

(for $z < 0$ we define $\log(z) = -\infty$.)

Proof: For the general case, fixing $U = W$ in Lemma 3.7 and using the bound on $\mathbb{E}_{X^k,Y^k|W}\mathrm{H}\left(\mathrm{P}_{A^k|X^k,Y^k,W}\right)$ from Proposition 3.8 we obtain:

$$\mathbb{E}_{X^k,Y^k,A^k|W}\mathrm{D}\left(\mathrm{P}_{X^{n-k},Y^{n-k}|X^k,Y^k,A^k,W}\|\mathrm{P}_{X^{n-k},Y^{n-k}}\right) \leq k\log s - \log(\Pr(W))$$

To complete the proof apply Fact 2.6.

For the projection game case, fix $U = W'$ in Lemma 3.7 and use the bound on $\mathbb{E}_{X^k,Y^k|W'}\mathrm{H}\left(\mathrm{P}_{A^k|X^k,Y^k,W'}\right)$ from Proposition 3.8 to obtain:

$$\mathbb{E}_{X^k,Y^k,A^k|W'}\mathrm{D}\left(\mathrm{P}_{X^{n-k},Y^{n-k}|X^k,Y^k,A^k,W'}\|\mathrm{P}_{X^{n-k},Y^{n-k}}\right) \leq \epsilon(n-k)/16 - \log(\Pr(W'))$$

We bound $\Pr(W')$ in the following way:

$$\Pr(W') = \Pr(W \wedge [(A^k, X^k, Y^k) \in E]) = \Pr(W) - \Pr(W \wedge [(A^k, X^k, Y^k) \notin E])$$

We now bound the term $\Pr(W \wedge [(A^k, X^k, Y^k) \notin E])$. For every game G and strategy f_a, f_b, the probability of winning the game played with strategy f_a, f_b is

$$\mathbb{E}_{X,Y}\sum_{b\in B}\Pr(B = b|Y)\sum_{a\in A}\Pr(A = a|X)V(X, Y, a, b).$$

Recall that for every projection game G and every $x \in X, y \in Y, b \in B$ there is only one $a \in A$ for which $V(x,y,a,b) = 1$, this a is $f_{xy}(b)$ (recall that f_{xy} is the projection function, see Definition 2.2). Thus for every projection game G and strategy f_a, f_b, the probability of winning the game played according to f_a, f_b is:

$$\mathbb{E}_{XY} \sum_{(b,f_{XY}(b)) \in B \times A} \Pr(B = b|Y) \Pr(A = a|X).$$

For x^k, y^k we define $f_{x^k,y^k} : B^k \to A^k$ by $[f_{x^k,y^k}(b^k)]_i = f_{x_i,y_i}(b_i)$. We want to bound the probability of winning in the last k coordinates and that $(A^k, X^k, Y^k) \notin E$. Thus, for every x^k, y^k we want to sum $\Pr(B^k = b^k|Y^k = y^k) \Pr(A^k = a^k|X^k = x^k)$, only over $(b^k, f_{x^k,y^k}(b^k)) \in B^k \times A^k$ for which $(f_{x^k,y^k}(b^k), x^k, y^k) \notin E$. Thus

$$\begin{aligned}
&\Pr(W \wedge [(A^k, X^k, Y^k) \notin E]) \\
&= \mathbb{E}_{X^k,Y^k} \sum_{(b^k, f_{X^k,Y^k}(b^k)) \text{ s.t. } (f_{X^k,Y^k}(b^k), X^k, Y^k) \notin E} \Pr(B^k = b^k|Y^k) \Pr(A^k = f_{X^k,Y^k}(b^k)|X^k, Y^k) \\
&< 2^{-\epsilon(n-k)/16} \qquad (7)
\end{aligned}$$

where the last inequality follows since if $(a^k, x^k, y^k) \notin E$ then

$$\Pr(A^k = a^k|X^k = x^k) = \Pr(A^k = a^k|X^k = x^k, Y^k = y^k) < 2^{-\epsilon(n-k)/16}.$$

Thus $\Pr(W') > \Pr(W) - 2^{-\epsilon(n-k)/16}$. We now conclude that

$$\mathbb{E}_{X^k,Y^k,A^k|W'} \mathrm{D}\left(\mathrm{P}_{X^{n-k},Y^{n-k}|X^k,Y^k,A^k,W'} \| \mathrm{P}_{X^{n-k},Y^{n-k}}\right) \leq \epsilon(n-k)/16 - \log\left(\Pr(W) - 2^{-\epsilon(n-k)/16}\right)$$

The corollary follows by using Fact 2.6. ■

Observation 3.10 *For any product distribution $\mathrm{P}_{\alpha,\beta} = \mathrm{P}_\alpha \times \mathrm{P}_\beta$ and any event τ that is determined only by α (or only by β) $\mathrm{P}_{\alpha,\beta|\tau}$ is a product distribution*

$$\mathrm{P}_{\alpha,\beta|\tau} = \mathrm{P}_{\alpha|\tau} \times \mathrm{P}_{\beta|\tau} = \mathrm{P}_{\alpha|\tau} \times \mathrm{P}_\beta$$

(or $\mathrm{P}_{\alpha,\beta|\tau} = \mathrm{P}_\alpha \times \mathrm{P}_{\beta|\tau}$)

Proposition 3.11 *For a free game G, an event U that is determined by X^k, Y^k, A^k, B^k and for every x^k, y^k, a^k the following holds:*

$$\mathrm{P}_{X^{n-k}Y^{n-k}|X^k=x^k,Y^k=y^k,A^k=a^k,U} = \mathrm{P}_{X^{n-k}|X^k=x^k,Y^k=y^k,A^k=a^k,U} \times \mathrm{P}_{Y^{n-k}|X^k=x^k,Y^k=y^k,A^k=a^k,U}$$

That is $\mathrm{P}_{X^{n-k}Y^{n-k}|X^k=x^k,Y^k=y^k,A^k=a^k,U}$ is a product distribution.

Proof: By applying Observation 3.10 three times on the events $X^k = x^k, Y^k = y^k, A^k = a^k$, we obtain that

$$\mathrm{P}_{X^{n-k}Y^{n-k}|X^k=x^k,Y^k=y^k,A^k=a^k}$$

is a product distribution. Since after we fixed x^k, y^k, a^k, the event U only depends on B^k, which is only a function of Y^{n-k}, we can apply Observation 3.10 one more time to obtain the proposition. ■

Corollary 3.12 *For a free game* G*, any event* U *that is determined by* X^k, Y^k, A^k, B^k *and for every* x^k, y^k, a^k, x, y *and every* $i \in [n-k]$ *the following holds:*

$$\begin{aligned}&\mathrm{P}_{X^{n-k}Y^{n-k}|X^k=x^k,Y^k=y^k,A^k=a^k,U,X_i=x,Y_i=y}\\&= \mathrm{P}_{X^{n-k}|X^k=x^k,Y^k=y^k,A^k=a^k,U,X_i=x} \times \mathrm{P}_{Y^{n-k}|X^k=x^k,Y^k=y^k,A^k=a^k,U,Y_i=y}\end{aligned}$$

Proof: From Proposition 3.11 we obtain that:

$$\mathrm{P}_{X^{n-k}Y^{n-k}|X^k=x^k,Y^k=y^k,A^k=a^k,U}$$

is a product distribution

$$\begin{aligned}&\mathrm{P}_{X^{n-k}Y^{n-k}|X^k=x^k,Y^k=y^k,A^k=a^k,U}\\&= \mathrm{P}_{X^{n-k}|X^k=x^k,Y^k=y^k,A^k=a^k,U} \times \mathrm{P}_{Y^{n-k}|X^k=x^k,Y^k=y^k,A^k=a^k,U}\end{aligned}$$

Applying Observation 3.10 on the event $X_i = x$ we obtain that

$$\begin{aligned}&\mathrm{P}_{X^{n-k}Y^{n-k}|X^k=x^k,Y^k=y^k,A^k=a^k,U,X_i=x}\\&= \mathrm{P}_{X^{n-k}|X^k=x^k,Y^k=y^k,A^k=a^k,U,X_i=x} \times \mathrm{P}_{Y^{n-k}|X^k=x^k,Y^k=y^k,A^k=a^k,U}\end{aligned}$$

Applying Observation 3.10 on the event $Y_i = y$ we obtain that

$$\begin{aligned}&\mathrm{P}_{X^{n-k}Y^{n-k}|X^k=x^k,Y^k=y^k,A^k=a^k,U,X_i=x,Y_i=y}\\&= \mathrm{P}_{X^{n-k}|X^k=x^k,Y^k=y^k,A^k=a^k,U,X_i=x} \times \mathrm{P}_{Y^{n-k}|X^k=x^k,Y^k=y^k,A^k=a^k,U,Y_i=y}\end{aligned}$$

■

Recall that for fixed $i \in [n-k]$, we denote $X^m = (X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_{n-k})$, i.e., X^{n-k} without X_i, and similarly, $Y^m = (Y_1, \ldots, Y_{i-1}, Y_{i+1}, \ldots, Y_{n-k})$.

Proof of Lemma 3.5 and Lemma 3.6: For both $U = W$ and $U = W'$ and for every x^k, y^k, a^k and $i \in [n-k]$, we will use a strategy for the game $G(\mathrm{P}_{X^n,Y^n}, V^{\otimes n})$ to obtain a strategy for the game $G(\mathrm{P}_{X_iY_i|X^k=x^k,Y^k=y^k,A^k=a^k,U}, V)$. Fix any strategy, f_a, f_b, for the game $\mathrm{G}(\mathrm{P}_{X^nY^n}, V^{\otimes n})$, and apply the following to obtain a strategy for $G(\mathrm{P}_{X_iY_i|X^k=x^k,Y^k=y^k,A^k=a^k,U}, V)$:

Algorithm 3.13 *Protocol for* $G(\mathrm{P}_{X_iY_i|X^k=x^k,Y^k=y^k,A^k=a^k,U}, V)$ *for fixed* x^k, y^k, a^k, i

1. *When the game starts, prover* 1 *receives a question* x *and prover* 2 *receives a question* y *according to* $\mathrm{P}_{X_iY_i|X^k=x^k,Y^k=y^k,A^k=a^k,U}$*. Define* $X_i = x, Y_i = y$ *(the provers will play this game in coordinate* i*).*

2. *Prover* 1 *randomly chooses*
   $x^m = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{n-k})$ *according to* $\mathrm{P}_{X^{n-k}|X^k=x^k,Y^k=y^k,A^k=a^k,U,X_i=x}$
   and Prover 2 randomly chooses
   $y^m = (y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_{n-k})$ *according to* $\mathrm{P}_{Y^{n-k}|X^k=x^k,Y^k=y^k,A^k=a^k,U,Y_i=y}$

3. *Prover* 1 *answers* $[f_a(x^n)]_i$ *and prove* 2 *answers* $[f_b(y^n)]_i$*.*

Remark 1 *Notice that in step 2, since both events $U = W$ and $U = W'$ are determined by X^k, Y^k, A^k, B^k, the joint distribution of x^m, y^m is $\mathrm{P}_{X^m,Y^m|X^k=x^k,Y^k=y^k,A^k=a^k,X_i=x,Y_i=y,U}$ which follows from Corollary 3.12.*

Remark 2 *Notice that since Remark 1 holds, the probability of winning the game*

$$G(\mathrm{P}_{X_iY_i|X^k=x^k,Y^k=y^k,A^k=a^k,U}, V)$$

is exactly

$$\Pr(W_i|X^k = x^k, Y^k = y^k, A^k = a^k, U).$$

Remark 3 *Notice that this is a randomized algorithm. However, it is well known that since any randomized algorithm is a convex combination of deterministic algorithms, there is a deterministic algorithm that achieves the same value as the randomized algorithm. Namely, there is a deterministic protocol for which the probability of winning the game*

$$G(\mathrm{P}_{X_iY_i|X^k=x^k,Y^k=y^k,A^k=a^k,U}, V)$$

is exactly

$$\Pr(W_i|X^k = x^k, Y^k = y^k, A^k = a^k, U).$$

Using this remark we will think of this algorithm as a deterministic algorithm.

Proof for General Games

By Corollary 3.9 for a fixed strategy f_a, f_b for $\mathrm{G}(\mathrm{P}_{X^nY^n}, V^{\otimes n})$,

$$\mathbb{E}_{i\in[n-k]}\mathbb{E}_{X^k,Y^k,A^k|W}\mathrm{D}\left(\mathrm{P}_{X_i,Y_i|X^k,Y^k,A^k,W}\|\mathrm{P}_{X_i,Y_i}\right) \leq \frac{1}{n-k}\left(k\log s - \log(\Pr(W))\right)$$

By the assumption in the lemma, $\Pr(W) \geq 2^{-\epsilon(n-k)/9+k\log s}$. Therefore, it follows that:

$$\mathbb{E}_{i\in[n-k]}\mathbb{E}_{X^k,Y^k,A^k|W}\mathrm{D}\left(\mathrm{P}_{X_i,Y_i|X^k,Y^k,A^k,W}\|\mathrm{P}_{X_i,Y_i}\right) \leq \epsilon/9$$

Assume by way of contradiction that for all $i \in [n-k]$, $\Pr(W_i|W) > 1 - \epsilon/9$. Notice that since

$$\Pr(W_i|W) = \mathbb{E}_{X^k,Y^k,A^k|W}\Pr(W_i|X^k, Y^k, A^k, W),$$

an equivalent assumption is that for all $i \in [n-k]$,

$$\mathbb{E}_{X^k,Y^k,A^k|W}\Pr(\neg W_i|X^k, Y^k, A^k, W) < \epsilon/9.$$

By a simple averaging argument, there are x^k, y^k, a^k and $i \in [n-k]$ for which both equations hold:

$$\mathrm{D}\left(\mathrm{P}_{X_i,Y_i|X^k=x^k,Y^k=y^k,A^k=a^k,W}\|\mathrm{P}_{X_i,Y_i}\right) \leq \epsilon/4 \tag{8}$$

$$\Pr(\neg W_i|X^k = x^k, Y^k = y^k, A^k = a^k, W) < \epsilon/4 \tag{9}$$

For the strategy f_a, f_b and for x^k, y^k, a^k, i for which both Equation (8) and Equation (9) hold consider the protocol suggested in Algorithm 3.13. Recall that by Remark 3 there is a deterministic protocol for which the provers win on coordinate i with probability

$$\Pr(W_i | X^k = x^k, Y^k = y^k, A^k = a^k, W).$$

Denote this deterministic protocol by h_a, h_b. For h_a, h_b, denote by R the set of all questions on which the provers err when playing according to this protocol. By the assumption in Equation (9)

$$\mathrm{P}_{X_i,Y_i|X^k=x^k,Y^k=y^k,A^k=a^k,W}(R) < \epsilon/4. \tag{10}$$

Combining Equation (10) with Equation (8), we can apply Corollary 3.4 to obtain $\mathrm{P}_{X_i,Y_i}(R) < \epsilon$. The provers can play h_a, h_b as a strategy for $\mathrm{G}(\mathrm{P}_{X_i,Y_i}, V)$ and err only on questions in R. Since $\mathrm{P}_{X_i,Y_i}(R) < \epsilon$, the value of $\mathrm{G}(\mathrm{P}_{X_i,Y_i}, V) > 1-\epsilon$ and since $\mathrm{P}_{X_i,Y_i} = \mathrm{P}_{XY}$, the value of $\mathrm{G}(\mathrm{P}_{XY}, V) > 1-\epsilon$ which is a contradiction.

Proof for Projection Games

The proof is very similar to the general case. From Corollary 3.9 we obtain:

$$\mathbb{E}_{i\in[n-k]}\mathbb{E}_{X^k,Y^k,A^k|W'}\mathrm{D}\left(\mathrm{P}_{X_i,Y_i|X^k,Y^k,A^k,W'} \| \mathrm{P}_{X_i,Y_i}\right) \tag{11}$$

$$\leq \frac{1}{n-k}\left(\epsilon(n-k)/16 - \log\left(\Pr(W) - 2^{-\epsilon(n-k)/16}\right) \right) \tag{12}$$

By the assumption in the lemma, $\Pr(W) \geq 2^{-\epsilon(n-k)/144}$ thus,

$$\begin{aligned}
&\mathbb{E}_{i\in[n-k]}\mathbb{E}_{X^k,Y^k,A^k|W'}\mathrm{D}\left(\mathrm{P}_{X_i,Y_i|X^k,Y^k,A^k,W'} \| \mathrm{P}_{X_i,Y_i}\right) \\
&\leq \epsilon/16 - \frac{1}{n-k}\log\left(2^{-\epsilon(n-k)/144} - 2^{-\epsilon(n-k)/16}\right) \\
&= \epsilon/16 - \frac{1}{n-k}\log\left(2^{-\epsilon(n-k)/16}\left(2^{\epsilon(n-k)/18} - 1\right)\right) \\
&= \epsilon/16 + \epsilon/16 - \frac{1}{n-k}\log\left(2^{\epsilon(n-k)/18} - 1\right) \\
&\leq \epsilon/8
\end{aligned} \tag{13}$$

where the last inequality is due to the bound on $n-k$. Assume by way of contradiction that for all $i \in [n-k]$, $\Pr(W_i|W') > 1-\epsilon/8$. Notice that since

$$\Pr(W_i|W') = \mathbb{E}_{X^k,Y^k,A^k|W'} \Pr(W_i|X^k, Y^k, A^k, W'),$$

an equivalent assumption is that for all $i \in [n-k]$,

$$\mathbb{E}_{X^k,Y^k,A^k|W'} \Pr(\neg W_i|X^k, Y^k, A^k, W') < \epsilon/8.$$

By a simple averaging argument, there are x^k, y^k, a^k and $i \in [n-k]$ for which both equations hold:

$$\mathrm{D}\left(\mathrm{P}_{X_i,Y_i|X^k=x^k,Y^k=y^k,A^k=a^k,W'} \| \mathrm{P}_{X_i,Y_i}\right) \leq \epsilon/4 \tag{14}$$

$$\Pr(\neg W_i|X^k = x^k, Y^k = y^k, A^k = a^k, W') < \epsilon/4 \tag{15}$$

For the strategy f_a, f_b, and for x^k, y^k, a^k, i for which both Equation (14) and Equation (15) hold consider the protocol suggested in Algorithm 3.13. Recall that by Remark 3 there is a deterministic protocol for which the provers win on coordinate i with probability

$$\Pr(W_i|X^k = x^k, Y^k = y^k, A^k = a^k, W').$$

Denote this deterministic protocol by h_a, h_b. For h_a, h_b, denote by R the set of all questions on which the provers err when playing according to this protocol. By our assumption

$$\mathrm{P}_{X_i,Y_i|X^k=x^k,Y^k=y^k,A^k=a^k,W'}(R) < \epsilon/4. \tag{16}$$

Combining Equation (16) with Equation (14), we can apply Corollary 3.4 to obtain $\mathrm{P}_{X_i,Y_i}(R) < \epsilon$. The provers can play h_a, h_b as a strategy for $\mathrm{G}(\mathrm{P}_{X_i,Y_i}, V)$ and err only on questions in R. Since $\mathrm{P}_{X_i,Y_i}(R) < \epsilon$, the value of $\mathrm{G}(\mathrm{P}_{X_i,Y_i}, V) > 1-\epsilon$. Since $\mathrm{P}_{X_i,Y_i} = \mathrm{P}_{XY}$ the value of $\mathrm{G}(\mathrm{P}_{XY}, V) > 1-\epsilon$ which is a contradiction.

We showed that there is $i \in [n-k]$ for which

$$\Pr(W_i|W') \leq 1 - \epsilon/8$$

but we need to show that there is $i \in [n-k]$ for which $\Pr(W_i|W) \leq 1 - \epsilon/9$. This is done in the following way: Since $W' \subseteq W$

$$\Pr(W_i|W) = \Pr(W_i|W')\Pr(W'|W) + \Pr(W_i|\neg W')\Pr(\neg W'|W) \leq \Pr(W_i|W') + \Pr(\neg W'|W).$$

Thus for all $i \in [n-k]$,

$$\Pr(W_i|W) \leq \Pr(W_i|W') + \Pr((A^k, X^k, Y^k) \notin E|W).$$

Since $\Pr((A^k, X^k, Y^k) \notin E|W) = \Pr(W \wedge [(A^k, X^k, Y^k) \notin E])/\Pr(W)$ we can use the bound in Equation (7), $\Pr(W \wedge [(A^k, X^k, Y^k) \notin E]) < 2^{-\epsilon(n-k)/16}$ and obtain that

$$\Pr(W_i|W) \leq \Pr(W_i|W') + 2^{-\epsilon(n-k)/16}/\Pr(W).$$

Therefore:

$$\begin{aligned}
&\Pr(W_i|W) \leq 1 - \epsilon/8 + 2^{-\epsilon(n-k)/16}/2^{-\epsilon(n-k)/144} \\
&\leq 1 - \epsilon/8 + 2^{-\epsilon(n-k)/18} \\
&\leq 1 - \epsilon/9
\end{aligned}$$

where the last inequality follows from the bound on $n-k$ ∎

Proof Of Theorem 3.1: We first show by induction that for every $k \leq \frac{\epsilon n}{18\log s+3}$ there is a set $T \subseteq [n]$ of k coordinates ($|T| = k$) for which $\Pr(W) \leq (1-\epsilon/9)^k$ where W is the event of winning on all the coordinates in T. For $k = 0$ the statement trivially holds. Assume by induction that there is a set T of size k for which $\Pr(W) \leq (1-\epsilon/9)^k$. If $\Pr(W) \leq (1-\epsilon/9)^{k+1}$ then we are done. Otherwise

$$\Pr(W) > (1-\epsilon/9)^{k+1} \geq 2^{-\epsilon(k+1)/4.5}$$

where we used the inequality $(1-x) \geq 2^{-2x}$ for $0 \leq x \leq 1/2$. In order to use Lemma 3.5 we need to make sure that $\Pr(W) \geq 2^{-\epsilon(n-k)/9+k\log s}$. It is enough to show that

$$2^{-\epsilon(k+1)/4.5} \geq 2^{-\epsilon(n-k)/9+k\log s}$$

or alternatively,

$$\epsilon(k+1)/4.5 \leq \epsilon(n-k)/9 - k\log s$$

After rearranging we obtain

$$k \leq \frac{\epsilon n - 2\epsilon}{9\log s + 3\epsilon}.$$

For $n > 2$ and $\epsilon \leq 1/2$ it is enough that[2]

$$k \leq \frac{\epsilon n}{18\log s + 3}.$$

Thus, for $k \leq \frac{\epsilon n}{18\log s+3}$ we can apply Lemma 3.5 to obtain that there is $i \notin T$ for which $\Pr(W_i|W) \leq 1-\epsilon/9$ therefore,

$$\Pr(W_i \wedge W) = \Pr(W)\cdot\Pr(W_i|W) \leq (1-\epsilon/9)^k(1-\epsilon/9) = (1-\epsilon/9)^{k+1}$$

To complete the proof, set $k = \frac{\epsilon n}{18\log s+3}$ then as we showed, there is a set $T \subseteq [n], |T| = k$ for which:

$$\Pr(W_1 \wedge \ldots \wedge W_n) \leq \Pr(\bigwedge_{i\in T} W_i) \leq (1-\epsilon/9)^{\epsilon n/(18\log s+3)} \leq (1-\epsilon^2/9)^{n/(18\log s+3)}.$$

where the last inequality follows by the use of the inequality $(1-x)^y \leq 1-xy$ for every $0 \leq y \leq 1$ and $x \leq 1$ ■

Proof Of Theorem 3.2: For the case of $n \geq (50/\epsilon)\log(8/\epsilon)$, the proof is very similar to the last theorem: We first show by induction, for every $k \leq (n/33)-1$ there is a set $T \subseteq [n]$ of k coordinates ($|T| = k$) for which $\Pr(W) \leq (1-\epsilon/9)^k$ where the event W is winning on all the coordinates in T. For $k = 0$ the statement trivially holds. Assume by induction that there is a set T of size k for which $\Pr(W) \leq (1-\epsilon/9)^k$. If $\Pr(W) \leq (1-\epsilon/9)^{k+1}$ then we are done, else

$$\Pr(W) \geq (1-\epsilon/9)^{k+1} \geq 2^{-\epsilon(k+1)/4.5}.$$

In order to use Lemma 3.6 we need to make sure that

$$\Pr(W) \geq 2^{-\epsilon(n-k)/144}$$

and that

$$n-k \geq (48/\epsilon)\log(8/\epsilon)$$

Since $k \leq (n/33)-1$,

$$\text{if } \Pr(W) \geq 2^{-\epsilon(k+1)/4.5} \text{ then } \Pr(W) \geq 2^{-\epsilon(n-k)/144}$$

Since $k \leq (n/33)-1$ then $n-k \geq 32n/33+1$. Since $n \geq 50/\epsilon\log(8/\epsilon)$ then

$$32n/33 + 1 \geq (48/\epsilon)\log(8/\epsilon) + 1.$$

Therefore,

$$n-k \geq (48/\epsilon)\log(8/\epsilon)$$

[2]We may assume that $n > 2$ since for $n \leq 2$ the theorem trivially holds. We also assume that the game is not trivial, i.e., the value of the game is not 0 or 1, thus $s > 1$.

Now we can apply Lemma 3.6 to obtain that there is $i \notin T$ for which $\Pr(W_i|W) \leq 1 - \epsilon/9$. Therefore,

$$\Pr(W_i \wedge W) = \Pr(W) \cdot \Pr(W_i|W) \leq (1 - \epsilon/9)^k (1 - \epsilon/9) = (1 - \epsilon/9)^{k+1}$$

For $k = (n/33) - 1$ there is a set $T \subseteq [n], |T| = k$ for which:

$$\Pr(W_1 \wedge \ldots \wedge W_n) \leq \Pr(\bigwedge_{i \in T} W_i) \leq (1 - \epsilon/9)^{(n/33)-1}$$

For the case of $n < (50/\epsilon)\log(8/\epsilon)$, as suggested in [Rao08], it can be shown that if the theorem was false for small n it would not hold for big n. If there was a strategy with success probability greater than $(1 - \epsilon/9)^{(n/33)-1}$ then for the same game played on $m \cdot n$ coordinates the success probability was at least $(1 - \epsilon/9)^{m((n/33)-1)}$ and for large enough m, this yield a contradiction. ■

References

[AK08] N. Alon and B. Klartag. Economical toric spines via cheeger's inequality. 2008. Manuscript.

[AKK+08] S. Arora, S. A. Khot, A. Kolla, D. Steurer, M. Tulsiani, and N. K. Vishnoi. Unique games on expanding constraint graphs are easy: extended abstract. In *STOC*, pages 21–28. 2008.

[BHH+08] B. Barak, M. Hardt, I. Haviv, A. Rao, O. Regev, and D. Steurer. Rounding parallel repetitions of unique games. In *FOCS*, pages 374–383. 2008.

[FKO07] U. Feige, G. Kindler, and R. O'Donnell. Understanding parallel repetition requires understanding foams. In *IEEE Conference on Computational Complexity*, pages 179–192. 2007.

[For89] L. J. Fortnow. Complexity - theoretic aspects of interactive proof systems. Technical Report MIT-LCS//MIT/LCS/TR-447, Department of Mathematics, Massachusetts Institute of Technology, 1989.

[FV02] U. Feige and O. Verbitsky. Error reduction by parallel repetition—a negative result. *Combinatorica*, 22(4):461–478, 2002.

[Hol07] T. Holenstein. Parallel repetition: simplifications and the no-signaling case. In *STOC'07*. 2007.

[KORW08] G. Kindler, R. O'Donnell, A. Rao, and A. Wigderson. Spherical cubes and rounding in high dimensions. In *FOCS*, pages 189–198. 2008.

[Rao08] A. Rao. Parallel repetition in projection games and a concentration bound. In *STOC*. 2008.

[Raz98] R. Raz. A parallel repetition theorem. *SIAM J. Comput.*, 27(3):763–803 (electronic), 1998. ISSN 0097-5397.

[Raz08] R. Raz. A counterexample to strong parallel repetition. In *FOCS*. 2008.

[SS07] S. Safra and O. Schwartz. On Parallel-Repetition, Unique-Game and Max-Cut, 2007. Manuscript.