Derandomized parallel repetition theorems for free games

Ronen Shaltiel
Department of Computer Science University of Haifa
Haifa, Israel
Email: ronen@cs.haifa.ac.il

***Abstract*—Raz's parallel repetition theorem [21] together with improvements of Holenstein [12] shows that for any two-prover one-round game with value at most $1-\epsilon$ (for $\epsilon \leq 1/2$), the value of the game repeated n times in parallel on independent inputs is at most $(1-\epsilon)^{\Omega(\frac{\epsilon^2 n}{\ell})}$ where ℓ is the *answer length* of the game. For *free games* (which are games in which the inputs to the two players are uniform and independent) the constant 2 can be replaced with 1 by a result of Barak, Rao, Raz, Rosen and Shaltiel [1]. Consequently, $n = O(\frac{t\ell}{\epsilon})$ repetitions suffice to reduce the value of a free game from $1-\epsilon$ to $(1-\epsilon)^t$, and denoting the *input length* of the game by m, if follows that $nm = O(\frac{t\ell m}{\epsilon})$ random bits can be used to prepare n independent inputs for the parallel repetition game.**

In this paper we prove a derandomized version of the parallel repetition theorem for free games and show that $O(t(m+\ell))$ random bits can be used to generate *correlated inputs* such that the value of the parallel repetition game on these inputs has the same behavior. Thus, in terms of randomness complexity, correlated parallel repetition can reduce the value of free games at the "correct rate" when $\ell = O(m)$.

Our technique uses *strong extractors* to "derandomize" a lemma of [21], and can be also used to derandomize a parallel repetition theorem of Parnafes, Raz and Wigderson [20] for *communication games* in the special case that the game is free.

***Keywords*-Parallel repetition, Derandomization, Randomness extractors;**

I. Introduction

A fundamental question in complexity theory is to what extent is it harder to solve many independent random instances of the same problem compared to solving a single random instance. This question is sometimes referred to as the "direct product question" or "parallel repetition question" and is studied in many algorithmic settings. In cases where "parallel repetition theorems" are known, the next step is often to "derandomize" them. That is, to design a sampling procedure that uses few random bits to sample many *correlated* instances such that solving these instances is as hard as solving independent instances. When measuring complexity as a function of the input length, "derandomized parallel repetition" produces problems that are harder than "independent parallel repetition". This is because the input length (which is the number of random bits used) is shorter in the derandomized version. One famous example is Yao's XOR Lemma [25] which is a "parallel repetition theorem" for circuit complexity (see [10] for a survey). Derandomized versions of variants of this lemma [9], [13], [15], [14] play a key role in Complexity Theory and Cryptography, and provide more insight on the parallel repetition question.

In this paper we prove derandomized versions of Raz's parallel repetition theorems for 2-prover 1-round games [21] and of the parallel repetition theorem of Parnafes, Raz and Wigderson [20] for communication games. In both settings we can only handle a subfamily of games called *"free games"*.

A. 2-prover 1-round games

2-prover 1-round proof systems were introduced by Ben-Or, Goldwasser, Kilian and Wigderson [2]. Such proofs play an important role in Complexity Theory and Cryptography. The notion of 2P1R-games defined below captures the interplay between two cheating provers and an honest verifier on a *fixed* false statement and is extensively studied.

A 2P1R-game G is a game between two cooperating players. The game is administered by a referee that samples a pair of inputs $(x,y) \in (\{0,1\}^m)^2$ according to some distribution μ (that is known in advance both players). The *randomness complexity* of G denoted by $\mathsf{rand}(G)$ is the number of random coins used by the referee to sample the pair (x,y). The first player receives input x and responds with an answer $a(x) \in \{0,1\}^\ell$. The second player receives input y and responds with an answer $b(y) \in \{0,1\}^\ell$. The players cannot communicate and their goal is to satisfy a predicate $V(x,y,a,b)$ (that is known in advance to both players). The *value* of G denoted by $\mathsf{val}(G)$ is the success probability of the best strategy of the players. A formal definition follows:

Definition I.1. *A 2P1R-game G is defined by a distribution μ over $(\{0,1\}^m)^2$ and a predicate V over $(\{0,1\}^m)^2 \times (\{0,1\}^\ell)^2$. We refer to m as the* input length *and to ℓ as the* answer length. *A* strategy *Π in G is a pair $\Pi = (a,b)$ of functions $a,b : \{0,1\}^m \to \{0,1\}^\ell$ and Π* wins *on $(x,y) \in (\{0,1\}^m)^2$ if $V(x,y,a(x),b(y)) = 1$. The* value *of G denoted by $\mathsf{val}(G)$ is the maximum over all strategies Π of $\Pr_{(X,Y)\leftarrow\mu}[\Pi$ wins on $(X,Y)]$. The game is* free *if μ is the uniform distribution over $(\{0,1\}^m)^2$ and for free games* *we define* $\mathsf{rand}(G) = 2m$.[1]

Parallel repetition of 2P1R-games: The *n-fold parallel repetition* of a 2P1R-game G is a 2P1R-game G^n in which the referee samples n independent pairs $(x_1, y_1), \dots, (x_n, y_n)$ where each pair is sampled according to μ. The first player receives the input $(x_1, \dots, x_n)$ and responds with an answer $(a_1, \dots, a_n) \in (\{0,1\}^\ell)^n$. It is important to note that the rules of 2P1R-games allow each a_i to be a function of the *entire input* $(x_1, \dots, x_n)$. Similarly, the second player receives $(y_1, \dots, y_n)$ and responds with answers $(b_1, \dots, b_n) \in (\{0,1\}^\ell)^n$. The predicate V^n of game G^n checks that for every $1 \leq i \leq n$, $V(x_i, y_i, a_i, b_i) = 1$. A formal definition follows:

Definition I.2 (The n-fold repetition game G^n)**.** *For a 2P1R-game G we define a 2P1R-game G^n with input length nm and answer length $n\ell$. We think of inputs as elements in $(\{0,1\}^m)^n$ and of answers as elements in $(\{0,1\}^\ell)^n$. G^n is defined by the distribution μ^n (that is the n-fold product of μ) and the predicate*

$$V^n\big((x_1, \dots, x_n), (y_1, \dots, y_n), (a_1, \dots, a_n), (b_1, \dots, b_n)\big) \\ = \bigwedge_{1 \leq i \leq n} V(x_i, y_i, a_i, b_i).$$

Note that $\mathsf{rand}(G^n) = n \cdot \mathsf{rand}(G)$ and that G^n is free if G is free.

Reducing the value by parallel repetition: Verbitsky [24] showed that for any game G with $\mathsf{val}(G) < 1$, $\mathsf{val}(G^n)$ tends to zero as n tends to infinity. Given a game G with $\mathsf{val}(G) \leq 1 - \epsilon$ and an integer t, a natural question is how many repetitions are required to guarantee that $\mathsf{val}(G^n) \leq (1-\epsilon)^t$. One may expect that $n = t$ repetitions suffice (or more generally that $\mathsf{val}(G^n) = \mathsf{val}(G)^n$). However, Fortnow [7] and subsequently, Lapidot and Shamir [17], and Feige [3] gave counterexamples. Specifically, there are *free* games in which $\mathsf{val}(G^2) = \mathsf{val}(G) = 1/2$. Moreover, Feige and Verbitsky [6] showed that one cannot answer the question above with a number of repetitions n that depends only on ϵ and t. More specifically, that for every n there is a *free* game G such that $\mathsf{val}(G) \leq 3/4$ and yet $\mathsf{val}(G^n) \geq 1/8$. Indeed, parallel repetition of 2P1R-games is maybe the most striking example where the answer to the parallel repetition question is unintuitive and complex (and this is the case even if we only consider free games). Consequently, a lot of papers addressed the question above for various sub-families of games. See [4] for a survey article.

In a celebrated result, Raz [21] proved that for every game G with $1/2 \leq \mathsf{val}(G) \leq 1 - \epsilon$, $\mathsf{val}(G^n) \leq (1-\epsilon)^{\Omega(\frac{\epsilon^{31} n}{\ell})}$. Holenstein [12] simplified parts of the proof and replaced the constant 31 with 2. In the special case that G is free, Barak, Rao, Raz, Rosen and Shaltiel [1] improve the constant to 1. We remark that improvements were also obtained for other special families of games such as "projection games" and games played on certain expander graphs. The reader is referred to [1] and the references therein for a discussion.

In this paper we are interested in the relationship between the randomness complexity and the value of parallel repetition and focus on free games. Summing up the discussion above we have that $n = O(\frac{t\ell}{\epsilon})$ repetitions suffice to reduce the value of any free game from $1 - \epsilon$ to $(1-\epsilon)^t$. The aforementioned example of Feige and Verbitsky also shows that the dependence of n on ℓ is tight up to log factors. It is not known whether the dependence of n on ϵ is tight for free games, however an example of Raz [22] shows that one cannot expect better dependence of n on ϵ for general games).

Reducing the randomness complexity of parallel repetition: Let G be a free game with $1/2 \leq \mathsf{val}(G) \leq 1 - \epsilon$. As explained above, to reduce the value below $(1-\epsilon)^t$, parallel repetition uses a game G^n with randomness complexity $nm = \Omega(\frac{t\ell m}{\epsilon})$. In this paper we introduce a "derandomized parallel repetition game" which achieves the same effect using randomness complexity $O(t \cdot (m + \ell))$. For $\ell = O(m)$ the randomness complexity used is $O(tm)$ which is asymptotically the same as that of a t-fold parallel repetition. In other words, in terms of randomness complexity (and for $\ell = O(m)$), the value of derandomize parallel repetition of a free game goes down at the "correct rate", that is as if $\mathsf{val}(G^n) = \mathsf{val}(G)^n$.

The derandomized game: Given a free game G, the derandomized game G^E is a free game defined given a function $E : \{0,1\}^r \times [n] \to \{0,1\}^m$. G^E has input length r and answer length $n\ell$. We denote the inputs to G^E by $(\bar{x}, \bar{y}) \in (\{0,1\}^r)^2$. The first player receives input $\bar{x} \in \{0,1\}^r$ and computes an input $(\bar{x}_1, \dots, \bar{x}_n)$ to G^n by $\bar{x}_i = E(\bar{x}, i)$. The second player uses $\bar{y}$ to compute an input $(\bar{y}_1, \dots, \bar{y}_n)$ to G^n by $\bar{y}_i = E(\bar{y}, i)$. The outcome of the game G^E is the outcome of G^n on inputs $\big((\bar{x}_1, \dots, \bar{x}_n), (\bar{y}_1, \dots, \bar{y}_n)\big)$. A formal definition follows:

Definition I.3. *Let $E : \{0,1\}^r \times [n] \to \{0,1\}^m$ be a function. For a string $\bar{x} \in \{0,1\}^r$ and $i \in [n]$ we define $\bar{x}_i = E(\bar{x}, i)$.*

Definition I.4 (Derandomized 2P1R-game)**.** *Let G be a free 2P1R-game with input length m and answer length ℓ. Let $E : \{0,1\}^r \times [n] \to \{0,1\}^m$ be a function. We define a free 2P1R-game G^E with input length r and answer length $n\ell$. The game G^E is defined by the predicate:*

[1] One can also consider games in which the input length or answer length of the two players are different. All the results in this paper also hold for such games taking m, ℓ to be the average of input lengths and answer lengths respectively. In some previous work the term "free game" is used to describe games where $(X, Y) \leftarrow \mu$ are independent but not necessarily uniformly distributed. Such games can be converted to our definition (while maintaining their value and randomness complexity) by having the referee send a pair of independent and uniformly distributed "seeds" (X', Y') using which each player privately generates his own input.

$$V^E(\bar{x}, \bar{y}, (a_1, \ldots, a_n), (b_1, \ldots, b_n))$$
$$= V^n((\bar{x}_1, \ldots, \bar{x}_n), (\bar{y}_1, \ldots, \bar{y}_n), (a_1, \ldots, a_n), (b_1, \ldots, b_n)).$$

The derandomized game has $\mathsf{rand}(G^E) = 2r$ and by using a *strong extractor* E we can achieve $\mathsf{rand}(G^E) \ll \mathsf{rand}(G^n) = 2nm$. We state our result informally below:

Theorem I.5. *(informal) Let t be an integer and let G be a free 2P1R-game with $1/2 \leq \mathsf{val}(G) \leq 1 - \epsilon$. For an appropriate choice of a strong extractor $E : \{0,1\}^r \times [n] \rightarrow \{0,1\}^m$ the derandomized game G^E satisfies $\mathsf{val}(G^E) \leq (1-\epsilon)^t$ and $\mathsf{rand}(G^E) = O(t \cdot (m + \ell)) = O(t \cdot (\mathsf{rand}(G) + \ell))$. Moreover, such an extractor can be computed in time polynomial in $m + \ell + t$.*

We define strong extractors in Section II and restate Theorem I.5 formally in Section IV.

Feige and Kilian [5] prove impossibility results for derandomizing Raz's parallel repetition theorem. Our result does not contradict theirs because of two reasons. First, their impossibility result does not apply to free games but rather to a subfamily of "constant degree games". The latter are games in which after revealing the input of one player, there are only a constant number of possible values for the input of the other player. Note that free games are very far from having this property. Second, the impossibility results of [5] rule out a much more ambitious derandomization than the one presented here. Namely, a derandomization that reduces the randomness complexity to $o(t \cdot \mathsf{rand}(G))$. Following [5] we remark that when making analogies to other settings of "derandomized parallel repetition" (for example "derandmoized versions of Yao's XOR-Lemma" [9], [13], [15], [14] or "averaging samplers" [27], [8]) one can hope to construct a derandomized game with randomness complexity $O(t + \mathsf{rand}(G))$. It is open whether it is possible to obtain randomness complexity $o(t \cdot \mathsf{rand}(G))$ for free games.

B. Communication games

Communication complexity introduced by Yao [26] considers two cooperating players who receive a pair of inputs $(x, y) \in (\{0,1\}^m)^2$ and want to compute a function $f(x, y)$. The computation is carried out using a *communication protocol* $P(x, y)$. The reader is referred to [16] for a definition of communication protocols and a comprehensive treatment of communication complexity. A communication protocol is called a c-bit communication protocol if for every input (x, y) no more than c bits are exchanged. The setup we consider below is "distributional communication complexity" where the inputs are chosen at random.

In a communication game G a referee samples a pair of inputs $(x, y) \in (\{0,1\}^m)^2$ according to some distribution μ (that is known to in advance both players). The *randomness complexity* of G denoted by $\mathsf{rand}(G)$ is the number of random coins used by the referee to sample the pair (x, y). The first player receives input x and the second player receives input y. The two players can run a c-bit communication protocol (where c is a parameter of the game) and their goal is to correctly compute some function $f(x, y)$ (that is known in advance to both players). A formal definition follows:

Definition I.6. *A communication game G is defined by a distribution μ over $(\{0,1\}^m)^2$, a function f over $(\{0,1\}^m)^2)$ and an integer $c \geq 0$. We refer to m as the* input length *and to c as the* communication complexity. *A* strategy *in G is a c-bit communication protocol $P(x, y)$ and P* wins *on $(x, y) \in (\{0,1\}^m)^2$ if $P(x, y) = f(x, y)$. The* value *of G denoted by $\mathsf{val}(G)$ is the maximum over all strategies P of $\Pr_{(X,Y)\leftarrow\mu}[P \text{ wins on } (X, Y)]$. The game is* free *if μ is the uniform distribution over $(\{0,1\}^m)^2$ and for free games we define $\mathsf{rand}(G) = 2m$.*

Parallel repetition of communication games: We now define the *n-fold parallel repetition* of a communication game G. Similar to 2P1R games we consider a referee that samples n independent pairs $(x_1, y_1), \ldots, (x_n, y_n)$ where each pair $(x_i, y_i) \in (\{0,1\}^m)^2$ is sampled according to μ and each player gets an n-tuple of inputs. The goal of the players is to correctly compute $f(x_i, y_i)$ for all $1 \leq i \leq n$ simultaneously. We want to define a communication game corresponding to parallel repetition of n original games. In contrast to 2P1R-games, there are subtleties as to how to formally define this concept. A natural attempt is to allow the players to use an (nc)-bit protocol on the input $((x_1, \ldots, x_n), (y_1, \ldots, y_n))$. However, Shaltiel [23] shows that with this definition there are examples where the value of the n-fold game is in fact *larger* than the value of the original game. Parnafes, Raz and Wigderson [20] suggested the following definition: In the game G^n the two players are allowed to run n c-bit communication protocols $P_1, \ldots, P_n$ "in parallel". The goal of the i'th protocol is to compute $f(x_i, y_i)$ and the input to P_i is $\big((x_1, \ldots, x_n), (y_1, \ldots, y_n)\big)$ and not just (x_i, y_i). (This model was initially suggested by Nisan, Rudich and Saks [18] in a related context of "parallel repetition of decision trees" and is called the "forest model"). Note that such a game cannot be described as a single communication game. A formal definition of G^n follows:

Definition I.7 (The n-fold repetition game G^n)**.** *For a communication game G with input length m and communication complexity c we define a game G^n. A* strategy *in G^n is a collection $\Pi = (P_1, \ldots, P_n)$ of c-bit communication protocols where each protocol receives input $\big((x_1, \ldots, x_n), (y_1, \ldots, y_n)\big) \in (\{0,1\}^{mn})^2$.* Π wins *on $\big((x_1, \ldots, x_n), (y_1, \ldots, y_n)\big) \in (\{0,1\}^{mn})^2$ if $P_i\big((x_1, \ldots, x_n), (y_1, \ldots, y_n)\big) = f(x_i, y_i)$ for every $1 \leq i \leq n$. The value of G^n denoted by $\mathsf{val}(G^n)$ is the maximum over strategies Π of*

$$\Pr_{(X_1,\ldots,X_n),(Y_1,\ldots,Y_n)\leftarrow\mu^n}[\Pi \text{ wins on } \big((X_1, \ldots, X_n), (Y_1, \ldots, Y_n)\big)].$$

Reducing the value by parallel repetition: Parnafes, Raz and Wigderson [20] proved a parallel repetition theorem for communication games. The proof is a reduction to an "enhanced version" of the Raz's parallel repetition theorem. Specifically, it follows that for a communication game G with $1/2 \leq \mathsf{val}(G) \leq 1-\epsilon$, taking $n = \frac{tc}{\epsilon^{31}}$ repetitions guarantees that $\mathsf{val}(G^n) \leq (1-\epsilon)^t$. Using the aforementioned improvements to the parallel repetition theorem the constant 31 can be reduced to 2 for general games and to 1 in free games. Note that the setting here is analogous to that in 2P1R-games with *communication complexity* c playing the role of *answer length* ℓ. (One difference is that in communication games it is unknown whether the dependence of n on c is necessary).

Reducing the randomness complexity of parallel repetition: Continuing the analogy, when we want to reduce the value of a free game from $1-\epsilon$ to $(1-\epsilon)^t$ we use a game G^n with randomness complexity $nm = \Omega(\frac{tcm}{\epsilon})$. Using our derandomized game G^E we can achieve the same effect using randomness complexity $O(tm)$. The construction of G^E is similar to that used in 2P1R-games. Namely, when given inputs $(\bar{x}, \bar{y}) \in (\{0,1\}^r)^2$ the two players use a function $E : \{0,1\}^r \times [n] \rightarrow \{0,1\}^m$ to privately compute inputs $(\bar{x}_1, \ldots, \bar{x}_n)$ and $(\bar{y}_1, \ldots, \bar{y}_n)$ for G^n and the outcome of G^E is the outcome of G^n on this pair of inputs. A formal definition follows:

Definition I.8 (Derandomized communication game)**.** *For a communication game G with input length m and communication complexity c, and a function $E : \{0,1\}^r \times [n] \rightarrow \{0,1\}^m$ we define a game G^E. A* strategy *in G^E is a collection $\Pi = (P_1, \ldots, P_n)$ of c-bit communication protocols where each protocol receives input $(\bar{x}, \bar{y}) \in (\{0,1\}^r)^2$ and Π* wins *on $(\bar{x}, \bar{y}) \in (\{0,1\}^r)^2$ if for every $1 \leq i \leq n$, $P_i(\bar{x}, \bar{y}) = f(\bar{x}_i, \bar{y}_i)$. The value of G^n denoted by $\mathsf{val}(G^n)$ is the maximum over strategies Π of $\Pr_{(\bar{x},\bar{y}) \leftarrow U_{2r}}[\Pi$ wins on $\big((X_1, \ldots, X_n), (Y_1, \ldots, Y_n)\big)]$ where U_{2r} denotes the uniform distribution over $(\{0,1\}^r)^2$.*

In this setting we prove the following theorem (that is analogous to Theorem I.5):

Theorem I.9. *(informal) Let t be an integer and let G be a free communication game with $1/2 \leq \mathsf{val}(G) \leq 1-\epsilon$. For an appropriate choice of a strong extractor $E : \{0,1\}^r \times [n] \rightarrow \{0,1\}^m$ the derandomized game G^E satisfies $\mathsf{val}(G^E) \leq (1-\epsilon)^t$ and $\mathsf{rand}(G^E) = O(tm) = O(t \cdot \mathsf{rand}(G))$. Moreover, such an extractor can be computed in time polynomial in $m+t$.*

Similar to 2P1R games, in terms of randomness complexity, the value of a free game goes down at the "correct rate" as if $\mathsf{val}(G^n) = \mathsf{val}(G)^n$. Note that in the setting of communication games the randomness complexity of G^E is independent of the communication complexity c (whereas in 2P1R-games the randomness complexity depends on the answer length ℓ). This is because a protocol with communication complexity $c = m+1$ can compute any function f on $(\{0,1\}^m)^2$). Thus, the assumption that $\mathsf{val}(G) < 1$ implies that $c \leq m$. Using our techniques for 2P1R games we can construct a game G^E with randomness complexity $O(t \cdot (m+c))$. However, by the previous discussion this is $O(tm)$.

II. Preliminaries

We use $[n]$ to denote $\{1, \ldots, n\}$. For a distribution P, $\Pr_P[T]$ denotes the probability of the event T under P, and $x \leftarrow P$ denotes the experiment in which x is chosen according to P. For an event T which receives positive probability under distribution P, $(P|T)$ denotes the distribution of P conditioned on T, namely for every event E, $\Pr_{(P|T)}[E] = \Pr_P[E|T]$. U_m denotes the uniform distribution on $\{0,1\}^m$. For a set S, U_S denotes the uniform distribution on T and $x \leftarrow S$ is a shorthand for $x \leftarrow U_S$. The *min-entropy* of a random variable X denoted $H_\infty(X)$ is the minimum of $\log(1/Pr[X=x])$ where the minimum is over all x in the support of X. The *statistical distance* between two distributions P and Q over the same domain S is the defined by $\mathsf{SD}(P;Q) = max_{T \subseteq S} |\Pr_P[T] - \Pr_Q[S]|$.

Definition II.1 (Strong extractors [19])**.** *A function $E : \{0,1\}^r \times [n] \rightarrow \{0,1\}^m$ is a (k,ϵ)-strong extractor if for every random variable X over $\{0,1\}^r$ with $H_\infty(X) \geq k$, $\mathbb{E}_{i \leftarrow [n]}[\mathsf{SD}(E(X,i); U_m)] \leq \epsilon$.*

Preparing for our application, the definition above is phrased in a non-standard way. However, it is equivalent to the standard definition which requires that the distribution $(E(X,I), I)$ where $I \leftarrow [n]$ is of statistical distance at most ϵ from the uniform distribution $U_{\{0,1\}^m \times [n]}$.

III. Technique

Raz's parallel repetition theorem makes use of the following lemma.

Lemma III.1. *[21] Let $Z_1, \ldots, Z_n$ be independent random variables where each Z_i is uniformly distributed. Let T be an event with $\Pr[T] \geq 2^{-\Delta}$. Then $\mathbb{E}_{i \leftarrow [n]}[\mathsf{SD}((Z_i|T); Z_i)] \leq \epsilon$ for $\epsilon = O(\sqrt{\Delta/n})$.*[2]

The lemma says that for given ϵ, Δ, setting $n = O(\Delta/\epsilon^2)$ we have that for a random $i \in [n]$, the distribution of Z_i conditioned on T is within distance ϵ of the original distribution of Z_i (which is uniform). In the proof of the parallel repetition theorem we consider the following setup: The referee of G^n chooses independent $Z_1, \ldots, Z_n$ where each is a uniformly distributed string of length $\mathsf{rand}(G)$ and

[2] In [21] the lemma is stated for general i.i.d. variables without the additional requirement that each Z_i is uniform. We can imagine that each Z_i is sampled by choosing a uniform Z'_i and setting $Z_i = g(Z'_i)$ for some function g, and the general formulation follows by applying the weaker formulation on $Z'_1, \ldots, Z'_m$.

uses some function $g(z) = (x, y)$ to generate the n pairs of inputs $(X_1, Y_1), \dots, (X_n, Y_n)$ to G^n by $(X_i, Y_i) = g(Z_i)$. In intermediate steps, the proof considers conditioning on an event T. (We do not explain the role of T here as we elaborate on the structure of the proof in Section IV-A). Lemma III.1 says that for a random $i \in [n]$, the distribution $(Z_i|T)$ is close to uniform. This means that conditioned on T the inputs (X_i, Y_i) to the i'th repetition are distributed essentially like a pair of inputs from G. This property plays an important part in the proof and note that we need to invest $n \cdot \mathsf{rand}(G) = O(\mathsf{rand}(G) \cdot \Delta/\epsilon^2)$ random bits to generate $Z_1, \dots, Z_n$ with this property.

A key observation that we make in this paper is that this property can be achieved using correlated $Z_1, \dots, Z_n$ that are generated from much fewer random bits. Specifically, let $E : \{0,1\}^r \to \{0,1\}^{\mathsf{rand}(G)}$ be an $(r - \Delta, \epsilon)$-strong extractor. Such an E exists for $r = \Delta + \mathsf{rand}(G) + O(\log(1/\epsilon)$. We sample a uniform $\bar{Z} \in \{0,1\}^r$ and set $\bar{Z}_i = E(\bar{Z}, i)$. For T such that $\Pr[\bar{Z} \in T] \geq 2^{-\Delta}$, the distribution $(\bar{Z}|T)$ has min-entropy at least $r - \Delta$ and therefore by Definition II.1 we get

$$\mathbb{E}_{i \leftarrow [n]}[\mathsf{SD}((\bar{Z}_i|T); U_{\mathsf{rand}(G)})] \leq \epsilon$$

and obtain the behavior of the lemma using only $\mathsf{rand}(G) + \Delta + O(\log(1/\epsilon))$ random bits.

The discussion above explains the main idea behind our derandomization and note that so far we did not need to assume that the original game G is free. Specifically, we can define a derandomized game G^E_{S} in which a referee chooses $\bar{Z}_1, \dots, \bar{Z}_n$ by the process described above, computes $(\bar{X}_i, \bar{Y}_i) = g(\bar{Z}_i)$ and the players play G^n on inputs $(\bar{X}_1, \dots, \bar{X}_n)$ and $(\bar{Y}_1, \dots, \bar{Y}_n)$. The definition of G^E_{S} makes sense for general games and we can try and bound the value of this game. Unfortunately, there are complications in implementing this idea and making the proof go through. We are able to solve the complications for the special case of free games when using the construction G^E from Definition I.4. Note that the construction of G^E is different than G^E_{S} and applies the extractor twice on separate inputs to generate the inputs of the two players. We discuss some direction towards analyzing G^E_{S} for general games in Section V.

Extractors and averaging samplers: The sample space that we use in this paper, namely $\bar{Z} \to (\bar{Z}_1, \dots, \bar{Z}_n)$ where $\bar{Z}_i = E(\bar{Z}, i)$ for an extractor E, is often used in derandomization. Indeed, Zuckerman [27] (see also Goldreich's survey [8]) observed that this construction gives an "averaging sampler". Namely that for every set A, the random variable $|\{i : \bar{Z}_i \in A\}|$ is with high probability close to the expectation of $|\{i : Z_i \in A\}|$ for independently chosen $Z_1, \dots, Z_n$. The derandomization of this paper uses a seemingly different property of this sample space which may be useful in other settings.

IV. A DERANDOMIZED PARALLEL REPETITION THEOREM FOR FREE GAMES

In this section we state and prove our main results. Our approach for 2P1R-games and communication games is very similar and therefore within this section we will refer to both as "games" and mention the precise type of the game (2P1R-game or communication game) only when it makes a difference.

When given a free game G with input length m and a function $E : \{0,1\}^r \times [n] \to \{0,1\}^m$ we use Definitions I.3,I.4,I.8 to consider the game G^E. The following theorem (that is the main technical contribution of this paper) bounds the value of G^E in case E is a strong extractor with suitable parameters.

Theorem IV.1. *Let $0 \leq \epsilon \leq 1$, let $t \geq 0$ be an integer and let $E : \{0,1\}^r \times [n] \to \{0,1\}^m$ be a strong $(r - \Delta, \epsilon/8)$-extractor.*

- *If G is a free 2P1R-game with input length m and answer length ℓ, and $\Delta = t(2m+2\ell+1) + \log(1/\epsilon) + 2$ then $\mathsf{val}(G^E) \leq (1 - \frac{\epsilon}{2})^t$.*
- *If G is a free communication game with input length m and communication complexity c, and $\Delta = t(2m + c + 1) + \log(1/\epsilon) + 2$ then $\mathsf{val}(G^E) \leq (1 - \frac{\epsilon}{2})^t$.*

We can simplify some of the quantities above if we are less picky: Note that the theorem is trivial when $\epsilon = 0$ and $\mathsf{val}(G) = 1$ and so we can assume that $\epsilon > 0$. In a free game G with input length m we have $\mathsf{rand}(G) = 2m$ and therefore if $\mathsf{val}(G) \leq 1 - \epsilon < 1$ then $\epsilon \geq 2^{-2m}$. Thus, the term $\log(1/\epsilon)$ in Theorem IV.1 can be replaced by $2m$. In the case of communication games, a game with communication complexity $c = m + 1$ has value 1 (as any function can be computed with communication complexity $c = m + 1$). Therefore, the assumption that $\mathsf{val}(G) \leq 1 - \epsilon < 1$, implies $c \leq m$ and we can replace c with m in the definition of Δ.

Theorem IV.1 is tailored to handle games with value approaching 1. We can also tailor it for games with value approaching 0. Specifically, if we assume that $\mathsf{val}(G) \leq \epsilon$ and replace the term "1" in the definition of Δ with $\log(1/\epsilon)$ then the proof gives that $\mathsf{val}(G^t) \leq (2\epsilon)^t$.

Parallel repetition as a strong extractor: One possible choice for E is "independent parallel repetition". Namely $r = nm$ and for $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n) \in (\{0,1\}^m)^n \cong \{0,1\}^r$ we define $E((\bar{x}_1, \dots, \bar{x}_n), i) = \bar{x}_i$. By Lemma III.1, E is a $(r - \Delta, \epsilon)$-extractor for $n = O(\Delta/\epsilon^2)$. Plugging this extractor into Theorem IV.1 gives a proof of the parallel repetition theorem for free games.[3]

[3] We remark that the the proof of [1] for free 2P1R-games uses a smaller number $n = O(t\ell/\epsilon)$ of repetitions, compared to $n = O(t(\ell+m)/\epsilon^2)$ that are obtained using Theorem IV.1. Loosely speaking, the proof of [1] exploits some additional properties of independent repetitions. These properties can be abstracted and incorporated into our framework. We avoid this as this does not help in reducing the randomness complexity.

Using strong extractors to obtain the parameters guaranteed in Theorems I.5,I.9: We can reduce the randomness complexity of G^E by plugging in better extractors. Specifically, by the probabilistic method there exist $(r-\Delta,\epsilon)$-extractors $E : \{0,1\}^r \times [n] \rightarrow \{0,1\}^m$ with $r = \Delta + m + O(\log(1/\epsilon))$ and $n = O(\Delta/\epsilon^2)$. Recent explicit (that is polynomial time computable) constructions of extractors come close to these parameters and achieve $r = O(\Delta + m + \log(1/\epsilon))$ and $n = \mathrm{poly}(\Delta/\epsilon)$ [27], [11]. (We can say more about some of the constants hidden in the last statement, but this is insignificant for our final results). Plugging these extractors into Theorem IV.1 and using the simplifications explained above gives the parameters guaranteed in Theorems I.5,I.9. More specifically, when starting with a free game G with $\mathsf{val}(G) \leq 1-\epsilon$ we construct a game G^E with $\mathsf{val}(G) \leq (1-\epsilon/2)^t$. For a 2P1R-game G, the randomness complexity of G^E is $\mathsf{rand}(G^E) = O(t(m+\ell))$ and it uses $n = \mathrm{poly}(t,m,\ell)$ repetitions. This should be compared to independent parallel repetition that uses $n = O(t\ell/\epsilon)$ repetitions and randomness complexity $\mathsf{rand}(G^n) = O(mt\ell/\epsilon)$ for the same goal. For a communication game G, the randomness complexity of G^E is $\mathsf{rand}(G^E) = O(tm)$ and it uses $n = \mathrm{poly}(t,m)$ repetitions. This should be compared to independent parallel repetition that uses $n = O(tc/\epsilon)$ repetitions and randomness complexity $\mathsf{rand}(G^n) = O(mtc/\epsilon)$ for the same goal.

A. The analysis

We are given a free game G with $\mathsf{val}(G) \leq 1-\epsilon$. Throughout the section we assume that the conditions of Theorem IV.1 are met. We consider a probability space consisting of two independent random variables $\bar{X}, \bar{Y}$ that are uniformly distributed over $\{0,1\}^r$. Let Π^E be some strategy of the two players in G^E. For $i \in [n]$, let W_i denote the event "Π^E wins the i'th repetition in G^E". More formally, for 2P1R-games Π^E consists of two functions $a, b : \{0,1\}^r \rightarrow (\{0,1\}^\ell)^n$ and $W_i = \left\{V(\bar{X}_i, \bar{Y}_i, a(\bar{X})_i, b(\bar{Y})_i) = 1\right\}$. For communication games the strategy Π^E consists of n c-bit communication protocols $(P_1, \ldots, P_n)$ and $W_i = \left\{P_i(\bar{X}, \bar{Y}) = f(\bar{X}_i, \bar{Y}_i)\right\}$. For $S \subseteq [n]$ let $W_S = \cap_{i \in S} W_i$. Our goal is to show that $\Pr[W_{[n]}] \leq (1-\frac{\epsilon}{2})^t$.

1) The high level strategy: A natural approach to bound $\Pr[W_{[n]}] = \Pr[W_1] \cdot \Pr[W_2|W_1] \cdot \ldots \cdot \Pr[W_n|W_1 \cap \ldots \cap W_{n-1}]$ is to try and bound each of the terms by $1-\frac{\epsilon}{2}$. However, as noted in the introduction there are counterexamples to this approach in the case of 2P1R games. Specifically, there are examples of free 2P1R-games with $\mathsf{val}(G) = 1/2$ and strategies with $\Pr[W_1] = 1/2$, but $\Pr[W_2|W_1] = 1$. We follows the strategy suggested in [21] and prove the following Lemma.

Lemma IV.2. *Let $S \subseteq [n]$ with $|S| \leq t$ and $\Pr[W_S] \geq (1-\frac{\epsilon}{2})^t$ then there exists $i \notin S$ such that $\Pr[W_i|W_S] \leq 1-\frac{\epsilon}{2}$.*

Proof of Theorem IV.1 using Lemma IV.2: Note that for every set $S \subseteq [n]$, $\Pr[W_{[n]}] \leq \Pr[W_S]$. Thus, it suffices to find an S with $\Pr[W_S] \leq (1-\frac{\epsilon}{2})^t$. We show the existence of such a set by the following iterative process: We start with $S = \emptyset$, $k = 0$ and maintain the invariant that S is of size k with $\Pr[W_S] \leq (1-\epsilon/2)^k$. A each step, if $\Pr[W_S] \leq (1-\frac{\epsilon}{2})^t$ then we are done. Otherwise, Lemma IV.2 gives an $i \notin S$ such that $\Pr[W_i|W_S] \leq 1-\frac{\epsilon}{2}$. This implies that

$$\Pr[W_{S\cup\{i\}}] = \Pr[W_S] \cdot \Pr[W_i|W_S]$$
$$\leq (1-\frac{\epsilon}{2})^k \cdot (1-\frac{\epsilon}{2}) = (1-\frac{\epsilon}{2})^{k+1}$$

Thus, adding i to S maintains the invariant. If we did not stop in the first t steps then $\Pr[W_S] \leq (1-\frac{\epsilon}{2})^t$.
In the remainder of the section we prove Lemma IV.2.

2) The value of conditioned games: Lemma IV.2 considers a "conditioned game" in which the players receive the inputs $(\bar{X}, \bar{Y})$ conditioned on an event $T = W_S$ and their goal is to win the i'th repetition. We will try to understand such games for arbitrary events T and $i \in [n]$. We want to know when is the value of such games bounded by the value of the original game G. This motivates the following definition.

Definition IV.3 (error of a conditioned game)**.** *Let T be an event, $i \in [n]$ and let $(X', Y') = ((\bar{X}, \bar{Y})|T)$. We define error$(T,i)$ to be the statistical distance between (X'_i, Y'_i) and the uniform distribution over $(\{0,1\}^r)^2$.*

If error(T,i) is large then conditioned on T, the pair $(\bar{X}_i, \bar{Y}_i)$ has a significantly different distribution than a pair of questions (X, Y) in the original game G. It may be the case that G becomes easy to win under this distribution and we cannot hope to approximate $\Pr[W_i|T]$ by $\mathsf{val}(G)$.

Following the discussion above, one may expect that $\Pr[W_i|T] \leq \mathsf{val}(G) + \mathrm{error}(T,i)$. However, this is not true in general. The reason is that when the players play the conditioned game, they are not forced play as a function of $(\bar{x}_i, \bar{y}_i)$ and are allowed to use all of $(\bar{x}, \bar{y})$. It could be the case that conditioned on T, $(\bar{X}_i, \bar{Y}_i)$ are uniformly distributed and independent, and yet $\bar{X}$ is correlated with $\bar{Y}_i$. The scenario above gives the player holding $\bar{X}$ information about $\bar{Y}_i$ that he does not receive in the original game. For example, it could be the case that $\bar{X}_j = \bar{Y}_i$ for some $j \neq i$ and then the player holding $\bar{X}$ knows the input $\bar{Y}_i$ of the other player. We stress that this scenario actually happens in the "counterexamples" of [3], [6] mentioned in the introduction. Nevertheless, the problem above is avoided in the case where $\bar{X}, \bar{Y}$ are independent conditioned on T. This leads to the following definition and lemma.

Definition IV.4 (Rectangles)**.** *Let $T \subseteq (\{0,1\}^r)^2$ be an event. We say that T is a rectangle if there exist $T_1, T_2 \in \{0,1\}^r$ such that $T = T_1 \times T_2$. We say that a rectangle T has deficiency Δ if $|T_1| \geq 2^{r-\Delta}$ and $|T_2| \geq 2^{r-\Delta}$.*

Lemma IV.5. *If T is a rectangle then for every $i \in [n]$, $\Pr[W_i|T] \leq \mathsf{val}(G) + error(T,i)$.*

Proof: We show how to use the strategy Π^E in G^E to define a strategy Π in G that wins with probability $\Pr[W_i|T] - \mathrm{error}(T,i)$. The Lemma will follow as the latter probability is bounded from above by $\mathsf{val}(G)$.

Let $(X',Y') = ((\bar{X},\bar{Y})|T)$. As T is a rectangle we have that X',Y' are independent. We will construct a strategy Π for G in which the players are randomized and use private coins. This strategy can be converted into a standard (deterministic) strategy by fixing the coins of the players to the best possible choice and this transformation does not reduce the success probability. The strategy Π receives a pair of inputs $(x,y) \in (\{0,1\}^m)^2$ for G and works as follows:

- The first player samples $\bar{x} \leftarrow (X'|X'_i = x)$ and the second player samples $\bar{y} \leftarrow (Y'|Y'_i = y)$. Note that as X',Y' are independent, the distribution $(X'|X'_i = x) = (X'|X'_i = x, Y'_i = y)$ and similarly $(Y'|Y'_i = y) = (Y'|Y'_i = y, X'_i = x)$. Thus, this sampling process can be described as choosing $(\bar{x},\bar{y}) \leftarrow ((X',Y')|X'_i = x, Y'_i = y)$.
- The two players simulate the strategy Π^E on the pair $(\bar{x},\bar{y}) \in (\{0,1\}^r)^2$ and use the simulation to determine their actions on (x,y) by "restricting" the strategy Π^E to the i'th repetition. Specifically, if G is a 2P1R-game then given $(\bar{x},\bar{y})$ the strategy Π^E defines answers $(a_1,\ldots,a_n)$ and $(b_1,\ldots,b_n)$ and the strategy Π will output answers a_i and b_i on (x,y). If G is a communication game then the strategy Π^E applies the communication protocols $P_1,\ldots,P_n$ on inputs $(\bar{x},\bar{y})$ and the strategy Π given (x,y) applies the protocol $P_i(\bar{x},\bar{y})$ and uses its output. (We remark that the fact that restricting Π^E induces a strategy for G follows because the choice of the "forest model" in the definition of G^E).

Let $(\hat{X},\hat{Y})$ be the distribution of $\bar{x},\bar{y}$ induced by applying the strategy Π to $(x,y) \leftarrow U_{2r}$. We claim that

$$\mathsf{SD}((\hat{X},\hat{Y}),(X',Y')) \leq \mathsf{SD}((X'_i,Y'_i),U_{2r})$$

and recall that the latter expression is the definition of $\mathrm{error}(T,i)$. The inequality follows because the distribution (X',Y') can also be described as applying the strategy Π to $(x,y) \leftarrow (X'_i,Y'_i)$. This means that the distribution $(\hat{X},\hat{Y})$ of the pair $(\bar{x},\bar{y})$ obtained when playing the strategy Π in G has distance at most $\mathrm{error}(T,i)$ from the distribution (X',Y') obtained when playing the strategy Π^E in G^E conditioned on T. In particular, $\Pr[W_i|T]$ and the success probability of the strategy Π in G differ by at most $\mathrm{error}(T,i)$. ■

3) The role of extractors: We have that for a rectangle T and $i \in [n]$, $\Pr[W_i|T] \leq \mathsf{val}(G) + \mathrm{error}(T,i)$. The use of extractors guarantees that if the rectangle is not too small then for a random $i \leftarrow [n]$, $\mathrm{error}(T,i)$ is small.

Lemma IV.6. *If T is a rectangle with deficiency Δ then $\mathbb{E}_{i \leftarrow [n]}[error(T,i)] \leq \frac{\epsilon}{4}$*

Proof: Let $(X',Y') = ((\bar{X},\bar{Y})|T)$. Recall that $X'_i = E(X',i)$ and $Y'_i = E(Y',i)$ and thus

$$\mathrm{error}(T,i) = \mathsf{SD}\big((E(X',i),E(Y',i));U_{2r}\big)$$

Our goal is to show that:

$$\mathbb{E}_{i \leftarrow [n]}[\mathsf{SD}\big((E(X',i),E(Y',i));U_{2r}\big)] \leq \frac{\epsilon}{4}$$

As T is a rectangle with deficiency Δ we have that X',Y' are independent and $H_\infty(X'),H_\infty(Y') \geq r - \Delta$. As E is a strong $(r-\Delta,\epsilon/8)$-extractor we have that:

$$\mathbb{E}_{i \leftarrow [n]}[\mathsf{SD}\big(E(X',i);U_r\big)] \leq \frac{\epsilon}{8}$$

$$\mathbb{E}_{i \leftarrow [n]}[\mathsf{SD}\big(E(Y',i);U_r\big)] \leq \frac{\epsilon}{8}$$

For a fixed $i \in [n]$ the variables $E(X',i),E(Y',i)$ are independent and therefore their joint distance from the uniform distribution is the sum of the individual distances. That is,

$$\mathsf{SD}\big((E(X',i),E(Y',i));U_{2r}\big)$$
$$\leq \mathsf{SD}(E(X',i);U_r) + \mathsf{SD}(E(Y',i);U_r).$$

The claim follows by the taking the expectation over $i \leftarrow [n]$ and using the linearity of expectation. ■

In particular, for a rectangle T with deficiency Δ combining Lemma IV.5 and Lemma IV.6 gives that there exists an i such that $\Pr[W_i|T] \leq \mathsf{val}(G) + \epsilon/4 \leq 1 - \epsilon/2$.

4) Proof of Lemma IV.2: We have developed machinery that for a rectangle T with deficiency Δ allows us to find an i such that $\Pr[W_i|T] \leq 1 - \epsilon/2$. To prove Lemma IV.2 we need to handle events of the form W_S which may not be rectangles. The following Lemma shows that each such event W_S is essentially a disjoint union of rectangles with deficiency Δ. This holds both for 2P1R-games and communication games (using the appropriate choice of Δ in Theorem IV.1).

Lemma IV.7. *Let $S \subseteq [n]$ such that $|S| \leq t$ and $\Pr[W_S] \geq (1-\frac{\epsilon}{2})^t$. There exist disjoint events $T_0,\ldots,T_L$ such that:*

- $\cup_{0 \leq j \leq L} T_j = W_S$.
- $\Pr[T_0|W_S] \leq \epsilon/4$.
- *For every $1 \leq j \leq L$, T_j is a rectangle of deficiency Δ.*

The proof of Lemma IV.7 appears in Section IV-A5. We are now ready to prove Lemma IV.2 and conclude the proof of Theorem IV.1.

Proof: (of Lemma IV.2) Given a set S that satisfies the requirements of Lemma IV.2 we can apply Lemma IV.7 and let $T_0,\ldots,T_L$ be the events that are guaranteed by Lemma IV.7. We first use Lemma IV.5 to estimate $\Pr[W_i|W_S]$ for a fixed $i \in [n]$.

$$\Pr[W_i|W_S] = \textstyle\sum_{0\leq j\leq L}\Pr[W_i|T_j]\cdot\Pr[T_j|W_S]$$

$$\leq \Pr[T_0|W_S] + \textstyle\sum_{1\leq j\leq L}\Pr[T_j|W_S]\cdot(\mathsf{val}(G)+\text{error}(T_j,i))$$

$$\leq \tfrac{\epsilon}{4} + \mathsf{val}(G) + \textstyle\sum_{1\leq j\leq L}\Pr[T_j|W_S]\cdot\text{error}(T_j,i).$$

We now use the bound above, Lemma IV.6 and the linearity of expectation to estimate $\mathbb{E}_{i\leftarrow[n]}\big[\Pr[W_i|W_S]\big]$.

$$\mathbb{E}_{i\leftarrow[n]}\big[\Pr[W_i|W_S]\big]$$

$$\leq \tfrac{\epsilon}{4} + \mathsf{val}(G) + \textstyle\sum_{1\leq j\leq L}\Pr[T_j|W_S]\cdot\mathbb{E}_{i\leftarrow[n]}[\text{error}(T_j,i)]$$

$$\leq \tfrac{\epsilon}{4} + \mathsf{val}(G) + \textstyle\sum_{1\leq j\leq L}\Pr[T_j|W_S]\cdot\tfrac{\epsilon}{4}$$

$$\leq \tfrac{\epsilon}{2} + \mathsf{val}(G)$$

Therefore, there exists $i\in[n]$ such that $\Pr[W_i|W_S]\leq\frac{\epsilon}{2}+\mathsf{val}(G)\leq 1-\frac{\epsilon}{2}$ and note that such an i must satisfy $i\notin S$. ■

5) Proof of Lemma IV.7: The proof uses the same outline as the initial proof of Raz. Let $k=|S|$. Throughout this proof we use the following notation: Given a sequence $R_1,\ldots,R_n$ of random variables we define $R_S=(R_i)_{i\in S}$ (the concatenation of R_i for $i\in S$).

For every possible value $\hat{x}$ of $\bar{X}_S$ we define the event $E_1^{\hat{x}}=\{\bar{X}_S=\hat{x}\}$. Similarly, for every possible value $\hat{y}$ of $\bar{Y}_S$ we define the event $E_2^{\hat{y}}=\{\bar{Y}_S=\hat{y}\}$. We also define the event $E^{\hat{x},\hat{y}}=E_1^{\hat{x}}\cap E_2^{\hat{y}}$. Note that the latter event is a rectangle by definition. Conditioning on such an event fixes all the input pairs in S. There are at most $2^{2km}\leq 2^{2tm}$ such events.

At this point, we distinguish between the case that G is a 2P1R-game and the case that G is a communication game.

The case of 2P1R-games: Given inputs $\bar{X},\bar{Y}$ the strategy $\Pi^E=(a,b)$ defines answers $(\bar{A},\bar{B})\in(\{0,1\}^{n\ell})^2$ by $\bar{A}=a(\bar{X})$ and $\bar{B}=a(\bar{Y})$. For every possible value $\hat{a}$ of $\bar{A}_S$ we define the event $F_1^{\hat{a}}=\{\bar{A}_S=\hat{a}\}$. For every possible value $\hat{b}$ of $\bar{B}_S$ we define the event $F_2^{\hat{b}}=\left\{\bar{B}_S=\hat{b}\right\}$. We also define the event $F^{\hat{a},\hat{b}}=F_1^{\hat{a}}\cap F_2^{\hat{b}}$. Note that the latter event is a rectangle because $\bar{A}$ is a function of $\bar{X}$ and $\bar{B}$ is a function of $\bar{Y}$. Conditioning on such an event fixes the answers of the repetitions in S and there are at most $2^{2k\ell}\leq 2^{2t\ell}$ such events.

For every $\hat{x},\hat{y},\hat{a},\hat{b}$ we define the event

$$T^{\hat{x},\hat{y},\hat{a},\hat{b}}=E^{\hat{x},\hat{y}}\cap F^{\hat{a},\hat{b}}$$

and note that it is a rectangle as the intersection of rectangles is a rectangle. Furthermore, conditioning on this event determines the outcome of the repetitions in S. We define $p=t(2m+2\ell)$ so that the number of such events is bounded by 2^p.

The case of communication games: Given inputs $\bar{X},\bar{Y}$ the strategy Π^E consists of n communication protocols $P_1,\ldots,P_n$ each over input pair $(\bar{X},\bar{Y})$. For every such protocol let $\bar{Q}_i$ denote the transcript of the protocol $P_i(\bar{X},\bar{Y})$ (that is the concatenation of all exchanged messages). For every possible value $\hat{q}$ of $\bar{Q}_S$ we define the event $F^{\hat{q}}=\{\bar{Q}_S=\hat{q}\}$. Note that this event is a rectangle by properties of communication protocols. More precisely, for every i and every possible transcript $q\in\{0,1\}^c$ of the protocol P_i, the set of inputs $(\bar{X},\bar{Y})$ on which the transcript $\bar{Q}_i=q$ is a rectangle. Conditioning on such an event fixes the transcripts of the protocols in S and there are at most $2^{kc}\leq 2^{tc}$ such events.

For every $\hat{x},\hat{y},\hat{q}$ we define the event

$$T^{\hat{x},\hat{y},\hat{q}}=E^{\hat{x},\hat{y}}\cap F^{\hat{q}}$$

and note that it is a rectangle as the intersection of rectangles is a rectangle. Furthermore, conditioning on this event determines the outcome of the repetitions in S. We define $p=t(2m+c)$ so that the number of such events is bounded by 2^p.

Continuing the proof in both cases: In both cases, we have a partition of the probability space to at most 2^p disjoint events. Furthermore, conditioning on each such event completely describes the outcome of the repetitions in S. In particular, such an event determines whether or not W_S occurs. More formally, each such event is either contained in W_S or disjoint to W_S. Let Γ denote the set of all such events that are contained in W_S. We have that

$$W_S=\bigcup_{T\in\Gamma}T.$$

At this point, we expressed W_S as a disjoint union of rectangles. However, some of these rectangles may not have deficiency Δ. Let T_0 be the union of all rectangles that do not have deficiency Δ and let $T_1,\ldots,T_\ell$ denote all the rectangles in Γ that have deficiency Δ. Indeed, $W_S=\cup_{0\leq j\leq L}T_j$ and for $j\geq 1$, T_j is a rectangle with deficiency Δ.

It is left to bound $\Pr[T_0|W_S]$. Every rectangle T that does not have deficiency Δ satisfies $\Pr[T]\leq 2^{-\Delta}$. We have that T_0 contains at most 2^p such rectangles and therefore

$$\Pr[T_0|W_S]=\frac{\Pr[T_0]}{\Pr[W_S]}\leq\frac{2^p\cdot 2^{-\Delta}}{(1-\frac{\epsilon}{2})^t}\leq\frac{\epsilon}{4}$$

where the last inequality follows by our choice of Δ and the guarantee that $\epsilon\leq 1$.

Remark IV.8. *Lemma IV.7 shows that we can split the set W_S into "relatively large" rectangles. The proof partitions W_S into many rectangles and as a result the average size of a rectangle may be small. The number of rectangles depends on ℓ in the case of 2P1R-games, and on c in the case of communication games. For some games G it may be possible to use fewer rectangles and improve the parameters. This idea was used in [20] to prove versions of the parallel* *repetition theorem that replace the answer length ℓ (or communication complexity c) with other parameters of the game. This idea can also be applied in our setting. However, the low level details are different.*

V. Discussion and Open problems

We believe that recasting the proof of the parallel repetition theorem as using strong extractors gives insight on the structure of the overall argument. It is plausible that the same high level idea can be used to derandomize parallel repetition in other settings.

A natural open problem is to extend our results to general games. It may be easier to start with sub-families of games such as "projection games". We now try to explain which parts of the proof of Theorem IV.1 extend to general games. The presentation of our construction G^E is tailored to free games. In the case of general games it makes sense to use the construction G^E_{S} outlined in Section III. Namely, the referee chooses a uniform string $\bar{Z} \in \{0,1\}^r$ and uses it to generate variables $\bar{Z}_1, \ldots, \bar{Z}_n \in \{0,1\}^{\mathsf{rand}(G)}$ by $\bar{Z}_i = E(\bar{Z}, i)$ where $E : \{0,1\}^r \times [n] \rightarrow \{0,1\}^{\mathsf{rand}(G)}$ is a strong $(r - \Delta, \epsilon/8)$-extractor. For each i the referee prepares the pair of inputs $(\bar{X}_i, \bar{Y}_i)$ by applying the sampling procedure $g(z) = (x, y)$ of the game G on $\bar{Z}_i$. As a sanity check, note that standard parallel repetition can be expressed as G^E_{S} where $E((\bar{Z}_1, \ldots, \bar{Z}_n), i) = \bar{Z}_i$.

When considering G^E_{S} we also need to reconsider our notion of rectangles. We say that an event T is a rectangle if $T = T_1 \cap T_2$ where T_1 is determined by $\bar{X}_1, \ldots, \bar{X}_n$ and T_2 is determined by $\bar{Y}_1, \ldots, \bar{Y}_n$. Some parts of the proof of Theorem IV.1 work for general games with these modifications. Specifically, Lemma IV.6,IV.7 follow exactly as stated.

The difficulty is in extending Lemma IV.5. Our proof for free games can be seen as solving this problem using a specific choice of extractor (which in turn leads to the definition of G^E). The proof of the parallel repetition theorem for general games can be viewed as using a weaker formulation of Lemma IV.5 in which the conclusion is only guaranteed for a rectangle with deficiency Δ and a random i. The proof of the latter statement uses once again Lemma III.1 and this suggests that it may be possible to derandomize it using strong extractors. However, it seems that these extractors will need to have additional properties to make the argument go through.

Acknowledgment

This research was supported by BSF grant 2004329 and ISF grant 686/07. I am grateful to Avi Wigderson for introducing me to this problem and for many discussions.

References

[1] Boaz Barak, Anup Rao, Ran Raz, Ricky Rosen, and Ronen Shaltiel. Strong parallel repetition theorem for free projection games. In *APPROX-RANDOM*, volume 5687 of *Lecture Notes in Computer Science*, pages 352–365. Springer, 2009.

[2] Michael Ben-Or, Shafi Goldwasser, Joe Kilian, and Avi Wigderson. Multi-prover interactive proofs: How to remove intractability assumptions. In *STOC*, pages 113–131. ACM, 1988.

[3] Uriel Feige. On the success probability of the two provers in one-round proof systems. In *Structure in Complexity Theory Conference*, pages 116–123, 1991.

[4] Uriel Feige. Error reduction by parallel repetition-the state of the art. Technical report, Weizmann Institute, Jerusalem, Israel, Israel, 1995.

[5] Uriel Feige and Joe Kilian. Impossibility results for recycling random bits in two-prover proof systems. In *STOC*, pages 457–468. ACM, 1995.

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

[7] Lance Fortnow, John Rompel, and Michael Sipser. Errata for on the power of multi-prover interactive protocols. In *Structure in Complexity Theory Conference*, pages 318–319, 1990.

[8] Oded Goldreich. A sample of samplers - a computational perspective on sampling (survey). *Electronic Colloquium on Computational Complexity (ECCC)*, 4(20), 1997.

[9] Oded Goldreich, Russell Impagliazzo, Leonid Levin, Venkatesan Ramarathanan, and David Zuckerman. Security preserving amplification of hardness. In *FOCS*, pages 318–326, 1990.

[10] Oded Goldreich, Noam Nisan, and Avi Wigderson. On yao's xor-lemma. *Electronic Colloquium on Computational Complexity (ECCC)*, 2(50), 1995.

[11] Venkatesan Guruswami, Christopher Umans, and Salil P. Vadhan. Unbalanced expanders and randomness extractors from parvaresh–vardy codes. *J. ACM*, 56(4), 2009.

[12] Thomas Holenstein. Parallel repetition: simplifications and the no-signaling case. In David S. Johnson and Uriel Feige, editors, *STOC*, pages 411–419. ACM, 2007.

[13] Russell Impagliazzo. Hard-core distributions for somewhat hard problems. In *FOCS*, pages 538–545, 1995.

[14] Russell Impagliazzo, Ragesh Jaiswal, Valentine Kabanets, and Avi Wigderson. Uniform direct product theorems: simplified, optimized, and derandomized. In Richard E. Ladner and Cynthia Dwork, editors, *STOC*, pages 579–588. ACM, 2008.

[15] Russell Impagliazzo and Avi Wigderson. $P = BPP$ if E requires exponential circuits: Derandomizing the xor lemma. In *STOC*, pages 220–229, 1997.

[16] Eyal Kushilevitz and Noam Nisan. *Communication Complexity*. Cambridge University Press, 1997.

[17] Dror Lapidot and Adi Shamir. A one-round, two-prover, zero-knowledge protocol for np. *Combinatorica*, 15(2):204–214, 1995.

[18] Noam Nisan, Steven Rudich, and Michael E. Saks. Products and help bits in decision trees. *SIAM J. Comput.*, 28(3):1035–1050, 1999.

[19] Noam Nisan and David Zuckerman. Randomness is linear in space. *J. Comput. Syst. Sci.*, 52(1):43–52, 1996.

[20] Itzhak Parnafes, Ran Raz, and Avi Wigderson. Direct product results and the gcd problem, in old and new communication models. In *ACM Symposium on Theory of Computing*, pages 363–372, 1997.

[21] Ran Raz. A parallel repetition theorem. *SIAM J. Comput.*, 27(3):763–803, 1998.

[22] Ran Raz. A counterexample to strong parallel repetition. In *FOCS*, pages 369–373. IEEE Computer Society, 2008.

[23] Ronen Shaltiel. Towards proving strong direct product theorems. *Computational Complexity*, 12(1-2):1–22, 2003.

[24] Oleg Verbitsky. Towards the parallel repetition conjecture. In *Structure in Complexity Theory Conference*, pages 304–307, 1994.

[25] Andrew C. Yao. Theory and applications of trapdoor functions (extended abstract). In *23rd Annual Symposium on Foundations of Computer Science*, pages 80–91, Chicago, Illinois, 3–5 November 1982. IEEE.

[26] Andrew Chi-Chih Yao. Some complexity questions related to distributive computing. In *Conference Record of the Eleventh Annual ACM Symposium on Theory of Computing, 30 April-2 May, 1979, Atlanta, Georgia, USA*, pages 209–213, 1979.

[27] David Zuckerman. Randomness-optimal oblivious sampling. *Random Struct. Algorithms*, 11(4):345–367, 1997.