Derandomized parallel repetition theorems for free games*

Ronen Shaltiel†
University of Haifa

February 3, 2011

Abstract

Raz's parallel repetition theorem [24] together with improvements of Holenstein [14] 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 [2]. 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, it 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. That is, it is possible to reduce the value from $1-\epsilon$ to $(1-\epsilon)^t$ while only multiplying the randomness complexity by $O(t)$ when $m = O(\ell)$.

Our technique uses *strong extractors* to "derandomize" a lemma of [24], and can be also used to derandomize a parallel repetition theorem of Parnafes, Raz and Wigderson [22] for *communication games* in the special case that the game is free.

*A preliminary version of this paper appeared in CCC 2010.
†This research was supported by BSF grant 2004329 and ISF grant 686/07.

1 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. Parallel repetition theorems are results that relate the hardness of solving many independent instances to that of solving a single random instance. 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 often the number of random bits used) is shorter in the derandomized version. A well known example of a direct product theorem is Yao's XOR Lemma [28] which is a "parallel repetition theorem" for circuit complexity (see [12] for a survey). Derandomized versions of variants of this lemma [11, 15, 17, 16] play a key role in Complexity Theory and Cryptography, and also 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 [24] and of the parallel repetition theorem of Parnafes, Raz and Wigderson [22] for communication games. In both settings we can only handle a subfamily of games called *"free games"*.

1.1 2-prover 1-round games

2-prover 1-round proof systems were introduced by Ben-Or, Goldwasser, Kilian and Wigderson [4]. Such proofs play an important role in Complexity Theory and Cryptography. The notion of 2P1R-games defined below was introduced to capture 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 on $(\{0,1\}^m)^2$ (that is known in advance to both players). We use the notation $(x, y) \leftarrow \mu$ to denote the experiment in which the pair (x, y) is chosen according to μ. 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 1.1. *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]

[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 preserving their value and randomness complexity) by having the referee

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 $(x_1, y_1), \ldots, (x_n, y_n)$ where each pair is sampled according to μ. The first player receives the input $(x_1, \ldots, x_n)$ and responds with an answer $(a_1, \ldots, a_n) \in (\{0,1\}^\ell)^n$. It is important to note that the rules of 2P1R-games allow each a_i to be a function of the *entire input* $(x_1, \ldots, x_n)$. Similarly, the second player receives $(y_1, \ldots, y_n)$ and responds with answers $(b_1, \ldots, b_n) \in (\{0,1\}^\ell)^n$. 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 1.2 (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*

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

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 It is natural to expect that parallel repetition of a 2P1R-game G reduces its value. Indeed, Verbitsky [27] showed that for any game G with $\mathsf{val}(G) < 1$, $\mathsf{val}(G^n)$ tends to zero as n tends to infinity. A lot of research addresses the *rate* at which the value goes down in various sub-families of games. See [6] for a survey article. This question can be naturally phrased as follows:

Question 1.1. *Let $0 < \epsilon \leq 1/2$, let G be a 2P1R-game with $\mathsf{val}(G) \leq 1 - \epsilon$ and let t be an integer. How large should n be so 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 [9] and subsequently, Lapidot and Shamir [19], and Feige [5] gave counterexamples. Specifically, there are *free* games in which $\mathsf{val}(G^2) = \mathsf{val}(G) = 1/2$. Moreover, Feige and Verbitsky [8] 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 $\mathsf{val}(G^n) \geq 1/8$.

In a celebrated result, Raz [24] proved that for every game G with $\mathsf{val}(G) \leq 1 - \epsilon$ and $\epsilon \leq 1/2$, it holds that $\mathsf{val}(G^n) \leq (1-\epsilon)^{\Omega(\frac{\epsilon^c n}{\ell})}$ where $c > 0$ is a universal constant and recall that ℓ measures the answer length of the game. Holenstein [14] simplified parts of the proof and improved the constant c from 31 to 2. In the special case that G is free, Barak, Rao, Raz, Rosen and Shaltiel [2] further improve the constant c to 1.[2] Improvements were also obtained for other special families of games such as "projection games" and games played on expander graphs. The reader is referred to [2] and the references therein for a discussion.

It is not known whether the results mentioned above are tight. However, it is known that the dependence of n on ℓ in the results mentioned above is optimal up to polylogarithmic factors. This follows from the aforementioned results of [8]. It is also known that for general games the constant c has to be at least 1. This follows from Raz's "counterexample to strong parallel repetition theorems" [25]. It is open whether the constant c can be improved from 2 to 1 for general games. It is also open whether c can be improved from

send a pair of independent and uniformly distributed "seeds" (X', Y') using which each player privately generates his own input. Another standard comment is that we could have allowed the strategy Π to be randomized (either with private coins or shared coins) without affecting the value of the game.

[2] In some of the previous work the bound on $\mathsf{val}(G^n)$ is presented in the form $(1-\epsilon^{c'})^{\Omega(\frac{n}{\ell})}$ in contrast to the form $(1-\epsilon)^{\Omega(\frac{\epsilon^c n}{\ell})}$ that we use in this paper. Note that the two forms are essentially the same under the translation $c = c' - 1$.

1 to 0 for free games. We remark that for "projection games" there are matching upper and lower bounds [23, 1] and that that this is also the case for free projection games [2].

Summing up this discussion we note that parallel repetition of 2P1R-games is a 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).

The randomness complexity of parallel repetition The previous discussion is focused on the relationship between the number of repetitions and the value of the game. In this paper we are interested in the relationship between the randomness complexity and the value of the game. This question was studied by Bellare, Goldreich and Goldwasser [3] in a related context of single prover interactive proofs. In this paper we focus on free games as we do not know how to handle general games (in Section 5 we discuss what parts of our techniques apply for general games).

Let G be a free game with $\mathsf{val}(G) \leq 1-\epsilon$ for $\epsilon \leq 1/2$. By the best known parallel repetition theorems, $n = O(\frac{t\ell}{\epsilon})$ repetitions suffice to reduce the value to $(1-\epsilon)^t$. Note that the game G^n has randomness complexity

$$\mathsf{rand}(G^n) = n \cdot \mathsf{rand}(G) = 2 \cdot nm = O(\frac{t\ell m}{\epsilon}).$$

In this paper we introduce a "derandomized parallel repetition game" which achieves the same effect using randomness complexity $O(t \cdot (m+\ell))$. More precisely, we show that the referee can use $O(t \cdot (m+\ell))$ random bits to sample inputs $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$ so that when the players play the game G^n on these inputs, the value is bounded by $(1-\epsilon)^t$. For $\ell = O(m)$ the randomness complexity used is $O(tm)$ which is asymptotically the same as the randomness complexity of a t-fold parallel repetition. In other words, the value of such games can be decreased from $1-\epsilon$ to $(1-\epsilon)^t$ while only multiplying the randomness complexity a factor of $O(t)$ independent of ℓ and ϵ. We now describe the derandomized game.

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 $(\bar{x}_1, \ldots, \bar{x}_n)$ to G^n by $\bar{x}_i = E(\bar{x}, i)$. The second player uses $\bar{y}$ to compute an input $(\bar{y}_1, \ldots, \bar{y}_n)$ 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 $\big((\bar{x}_1, \ldots, \bar{x}_n), (\bar{y}_1, \ldots, \bar{y}_n)\big)$. A formal definition follows:

Definition 1.3. *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 1.4 (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*

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

Parallel repetition can be seen as a special case in which $r = nm$ and $E((x_1, \ldots, x_n), i) = x_i$ and in this case G^E coincides with the game G^n. In general, the derandomized game has $\mathsf{rand}(G^E) = 2r$ and by choosing E to be a *strong extractor* (to be defined later) with suitable parameters we can achieve $\mathsf{rand}(G^E) \ll \mathsf{rand}(G^n) = 2nm$. We state our result informally below:

Theorem 1.5. *(informal) Let t be an integer, let $0 \leq \epsilon \leq 1/2$ and let G be a free 2P1R-game with $\mathsf{val}(G) \leq 1-\epsilon$. Let $E : \{0,1\}^r \times [n] \to \{0,1\}^m$ be a strong extractor with appropriate parameters, then the derandomized game G^E satisfies $\mathsf{val}(G^E) \leq (1-\epsilon)^t$, $\mathsf{rand}(G^E) = O(t \cdot (m+\ell)) = O(t \cdot (\mathsf{rand}(G)+\ell))$ and $n = \text{poly}(m,t,\ell)$. Moreover, there is such an extractor that can be computed in time $\text{poly}(m,t,\ell)$.*

We define strong extractors in Section 2 and restate Theorem 1.5 formally in Section 4.

Feige and Kilian [7] 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 [7] 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 [7] we remark that when making analogies to other settings of "derandomized parallel repetition" (for example "derandomized versions of Yao's XOR-Lemma" [11, 15, 17, 16] or "averaging samplers" [30, 10]) 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.

1.2 Communication games

Communication complexity introduced by Yao [29] 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 [18] 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 to 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 1.6. *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 [26] 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 [22] 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 [20] 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 1.7 (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 $1\le i\le n$. The value of G^n denoted by $\mathsf{val}(G^n)$ is the maximum over strategies Π of $\Pr_{((X_1,\ldots,X_n),(Y_1,\ldots,Y_n))\leftarrow\mu^n}[\Pi$ wins on $\big((X_1,\ldots,X_n),(Y_1,\ldots,Y_n)\big)]$.*

Reducing the value by parallel repetition Parnafes, Raz and Wigderson [22] proved a parallel repetition theorem for communication games. The proof is a reduction to an "enhanced version" of Raz's parallel repetition theorem. Specifically, it follows that for $0<\epsilon\le 1/2$ and a communication game G with $\mathsf{val}(G)\le 1-\epsilon$, taking $n=O(\frac{tc}{\epsilon^{31}})$ repetitions guarantees that $\mathsf{val}(G^n)\le(1-\epsilon)^t$. 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 a 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 $E:\{0,1\}^r\times[n]\to\{0,1\}^m$ 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 1.8 (Derandomized communication game)**.** *For a communication game G with input length m and communication complexity c, and a function $E:\{0,1\}^r\times[n]\to\{0,1\}^m$ 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 $1\le i\le n$, $P_i(\bar{x},\bar{y})=f(\bar{x}_i,\bar{y}_i)$. 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 1.5):

Theorem 1.9. *(informal) Let t be an integer, let $0\le\epsilon\le 1/2$ and let G be a free communication game with $1/2\le\mathsf{val}(G)\le 1-\epsilon$. Let $E:\{0,1\}^r\times[n]\to\{0,1\}^m$ be a strong extractor with appropriate parameters, then the derandomized game G^E satisfies $\mathsf{val}(G^E)\le(1-\epsilon)^t$, $\mathsf{rand}(G^E)=O(tm)=O(t\cdot\mathsf{rand}(G))$ and $n=poly(m,t)$. Moreover, there is such an extractor that can be computed in time $poly(m,t)$.*

Similar to 2P1R games, the value of a free game goes down from $1-\epsilon$ to $(1-\epsilon)^t$ while only multiplying the randomness complexity by $O(t)$. 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 $O(t \cdot (m + c))$, and by the previous discussion this is $O(tm)$ independent of c.

2 Preliminaries

We use $[n]$ to denote $\{1, \ldots, n\}$.

2.1 Probability distributions

For a distribution P, $x \leftarrow P$ denotes the experiment in which x is chosen according to P, and $\Pr_{x \leftarrow P}[T]$ denotes the probability of event T under this experiment. We often define a probability space by explicitly specifying the random experiments and variables in the probability space and in this case we use $\Pr[T]$ to denote the probability of event T in the probability space. For a random variable Z and an event T with positive probability in the underlying probability space $(Z|T)$ denotes the probability distribution obtained by conditioning Z on T. More precisely, for a in the support of Z, $\Pr_{x \leftarrow (Z|T)}[x = a] = \Pr[Z = a|T]$. U_m denotes the uniform distribution on $\{0,1\}^m$. For a set S, U_S denotes the uniform distribution on S 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 defined by $\mathsf{SD}(P;Q) = max_{T \subseteq S} |\Pr_{x \leftarrow P}[x \in T] - \Pr_{x \leftarrow Q}[x \in T]|$.

2.2 Strong extractors

Our derandomized games make use of strong extractors [21]. Preparing for our application, the definition below is phrased in a non-standard way.

Definition 2.1 (Strong extractors [21])**.** *A function $E : \{0,1\}^r \times [n] \to \{0,1\}^m$ is a strong (k, ϵ)-extractor if for every random variable X over $\{0,1\}^r$ with $H_\infty(X) \geq k$, $\mathbb{E}_{i \leftarrow [n]}[\mathsf{SD}(E(X,i); U_m)] \leq \epsilon$.*

We stress that Definition 2.1 it is equivalent to the more 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 over $\{0,1\}^m \times [n]$.

3 Technique

Our results follow by showing that extractors can derandomize (part of) the proof of Raz's parallel repetition theorem. While we do not know how to handle general games, our techniques suffice to deramndomize Raz's parallel repetition theorem for free games.

3.1 Where extractors come in

Raz's parallel repetition theorem makes use of a simple lemma that states that if we condition i.i.d. random variables $Z_1, \ldots, Z_n$ on an event T of not too small probability, then for a randomly chosen $i \leftarrow [n]$, the conditioned random variable $(Z_i|T)$ has small statistical distance from the unconditioned variable Z_i. We state the Lemma precisely below.[3]

Lemma 3.1. *[24] Consider a probability space that consists of independent variables $Z_1, \ldots, Z_n$ where each variable is uniformly distributed over $\{0,1\}^v$. Let T be an event with $\Pr[T] \geq 2^{-\Delta}$ and let $\epsilon > 0$. If $n \geq c\Delta/\epsilon^2$ for some universal constant c then*

$$\mathbb{E}_{i \leftarrow [n]}[\mathit{SD}((Z_i|T); Z_i)] \leq \epsilon$$

It is known that Lemma 3.1 is tight in the sense that the Lemma does not hold for $n = o(\Delta/\epsilon^2)$. A key observation that we make in this paper is that one can interpret this Lemma as a strong extractor. More precisely, let $r = nv$ and identify strings $z \in \{0,1\}^r$ with tuples $(z_1, \ldots, z_n) \in (\{0,1\}^v)^n$. Let $E : \{0,1\}^r \times [n] \to \{0,1\}^v$ be the function $E(z,i) = z_i$ and assume (as in the lemma) that $n \geq c\Delta/\epsilon^2$. We notice that the lemma is equivalent to the statement that E is a strong $(r - \Delta, \epsilon)$-extractor.

We now explain this relationship more precisely. It should be noted that the proof does not explicitly make use of this relationship and that the explanation below is given mainly for intuition. We first note that Lemma 3.1 follows if $E(z,i) = z_i$ is a strong $(r - \Delta, \epsilon)$-extractor (which is the more interesting direction for the purpose of this paper as we plan to replace the use of Lemma 3.1 with some "off the shelf" extractor). For any event T with $\Pr[T] \geq 2^{-\Delta}$ in the probability space of choosing uniform $Z = (Z_1, \ldots, Z_n)$, we define a random variable $X = (Z|T)$. To bound the min-entropy of X we note that for every element a in the support of X,

$$\Pr[X = a] = \Pr[Z = a|T] \leq \frac{\Pr[Z = a]}{\Pr[T]} \leq 2^{-(r-\Delta)}.$$

Thus, $H_\infty(X) \geq r - \Delta$ and if E is a strong $(r - \Delta, \epsilon)$-extractor then

$$\mathbb{E}_{i \leftarrow [n]}[\mathsf{SD}((Z_i|T); Z_i)] = \mathbb{E}_{i \leftarrow [n]}[\mathsf{SD}(E(X,i); U_v)] \leq \epsilon.$$

For completeness we also note that Lemma 3.1 implies the fact that E is a strong $(r - \Delta, \epsilon)$-extractor. It is standard that in order to prove that a function is a strong $(r - \Delta, \epsilon)$-extractor it is sufficient to consider only distributions X that are uniformly distributed over a subset $T \subseteq \{0,1\}^r$ of size $2^{r-\Delta}$. Each such subset T is an event with $\Pr[T] \geq 2^{-\Delta}$ in the probability space of Lemma 3.1, and therefore the conclusion of the Lemma implies that E is an extractor.

The observation that Lemma 3.1 follows from strong extractors can be seen as saying that strong extractors "derandomize" Lemma 3.1. That is, given any strong $(r - \Delta, \epsilon)$-extractor $E : \{0,1\}^r \times [n] \to \{0,1\}^v$ one can sample random variables $\bar{Z}_1, \ldots, \bar{Z}_n$ by uniformly choosing a string $\bar{Z} \leftarrow \{0,1\}^r$ and setting $\bar{Z}_i = E(\bar{Z}, i)$. The (possibly correlated) random variables $\bar{Z}_1, \ldots, \bar{Z}_n$ chosen this way satisfy the guarantee in the conclusion of the Lemma. Namely, for every event T with $\Pr_{\bar{Z} \leftarrow \{0,1\}^r}[T] \geq 2^{-\Delta}$ we have that $\mathbb{E}_{i \leftarrow [n]}[\mathsf{SD}((\bar{Z}_i|T); U_v)] \leq \epsilon$. The advantage is that there exist strong extractors with which this sampling process requires only $r = v + \Delta + O(\log(1/\epsilon))$ random bits compared to the $nv = \Omega(\Delta v/\epsilon^2)$ bits used to sample independent $Z_1, \ldots, Z_n$.

[3] We remark that in [24] the lemma is stated for general i.i.d. variables without the additional requirement that each Z_i is uniformly distributed. Nevertheless, we can imagine that each Z_i is sampled by choosing a uniformly chosen Z'_i and setting $Z_i = g(Z'_i)$ for some function g, and the general formulation follows by applying the weaker formulation on $Z'_1, \ldots, Z'_n$.

3.2 The role of Lemma 3.1 in Raz's proof of the parallel repetition theorem

Let G be a (not necessarily free) 2P1R-game and let v denote the randomness complexity of G. The referee samples $z \leftarrow \{0,1\}^m$ and uses some function $g : \{0,1\}^v \to (\{0,1\}^m)^2$ to prepare inputs x, y to the game G by computing $(x, y) = g(z)$. In the parallel repetition game G^n the referee samples n independent variables $z_1, \ldots, z_n$ and prepares inputs for n games. Let Π be the best strategy for the players. Let W_i denote the event that Π wins on the i'th repetition and let $W = \cap W_i$ denote the event that the players win all repetitions. The goal of the parallel repetition theorem is to bound $\Pr[W]$. At a high level, the proof of the parallel repetition works as follows: Let $S \subseteq [n]$ be a set of distinct indices (initially, $S = \emptyset$) and let $T = \cap_{i \in S} W_i$. We want to add a new index i to S while preserving the invariant that $\Pr[T] \leq (1 - \epsilon/2)^{|S|}$. The theorem will then follow by applying this process sufficiently many times and noting that as $W \subseteq T$, $\Pr[W] \leq \Pr[T]$. Let $i' \notin S$ be an index that we can add. Let $S' = S \cup \{i'\}$ and $T' = \cap_{i \in S'} W_i$ be the new values for S and T if we choose to add i'. Note that $\Pr[T'] = \Pr[T] \cdot \Pr[W_{i'}|T]$. Thus, if we can find an i' such that $\Pr[W_{i'}|T] \leq (1 - \epsilon/2)$ then we can add i' to S and maintain the invariant.

In order to analyze $\Pr[W_{i'}|T]$ we need to understand the success probability of the players at index i' when the probability space is conditioned on event T. Initially, (before conditioning) we know that the players can win on every index with probability at most $1 - \epsilon$ and our hope is that there exists $i' \notin S$ on which the success probability is bounded by $1 - \epsilon/2$ even when conditioning on T. Note that conditioning on T may skew the distribution of the pair of inputs given to the players. In particular, it could be that for some i' the distribution of pairs $(x_{i'}, y_{i'})$ that the players see on repetition i' when conditioned on T is very different from the original distribution μ. This is where Lemma 3.1 comes in. It says that for a uniformly chosen $i' \in [n]$ the distribution of $z_{i'}$ is close to its initial value after conditioning which in turn means that conditioning does not significantly affect the distribution of the pair $(x_{i'}, y_{i'})$ by much.

It is tempting to use this observation to define a derandomized game that we will denote by G^E_{S} (to distinguish it from the game G^E from definition 1.4 that only applies to free games). In G^E_{S} the referee will sample $z_1, \ldots, z_n$ using a strong extractor as explained in Section 3.2. The properties of extractors can replace Lemma 3.1 and argue that for a random i', the distribution of $(x_{i'}, y_{i'})$ is not significantly affected by conditioning on T.

Unfortunately, this does not suffice to bound $\Pr[W_{i'}|T]$. This is because conditioned on T it could be the case that $x_{i'}$ and y_j may become correlated for $j \neq i$. For example, it could be that $y_j = x_{i'}$ giving the second player knowledge that he does not posses in the original game G. It may become much easier for the players to win on repetition i' when given this additional knowledge, and thus we cannot hope to bound $\Pr[W_{i'}|T]$ by $\mathsf{val}(G)$. (We remark that this phenomenon occurs in some of the examples mentioned in the introduction for parallel repetition of free games).

Indeed, Lemma 3.1 does not suffice and the proof of the parallel repetition theorem uses a much more delicate argument in order to show that on a random i' the inputs that the players see on indices different than i' do not help them to win the index i'. We do not know how to imitate this argument in the derandomized version. However, for free games and using definition 1.4 we can imitate the argument of the parallel repetition theorem (with some modifications) and bound the value of the derandomized game. It is open whether the same can be done for general games and we discuss this problem in Section 5.

3.3 Extractors and averaging samplers

We use extractors to sample correlated random variables $Z_1, \ldots, Z_n$ with certain properties (as explained in Section 3.2). It was observed by Zuckerman [30] that the sample space that we use is an *averaging sampler*. More precisely, that choosing $\bar{Z} \leftarrow U_r$ and applying an extractor $E : \{0,1\}^r \times [n] \to \{0,1\}^v$ to generate

$\bar{Z}_1, \ldots, \bar{Z}_n$ by $\bar{Z}_i = E(\bar{Z}, i)$ produces a sample space with the property that for every set $A \subseteq \{0,1\}^v$, 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 $Z_1, \ldots, Z_n$. The reader is referred to Goldreich's survey [10] for more details on averaging samplers. Averaging samplers are often useful in direct product theorems and in some sense averaging samplers (or more precisely "hitters") are necessary to achieve derandomized parallel repetition theorems.[4] The derandomization of this paper does not argue using averaging samplers or hitters. Instead, we use a seemingly different property of the sample space $\bar{Z}_1, \ldots, \bar{Z}_n$ which may be useful in other settings.

4 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 1.3,1.4,1.8 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 4.1 (main theorem)**.** *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 $\mathsf{val}(G) \leq 1 - \epsilon$, 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 $\mathsf{val}(G) \leq 1-\epsilon$, input length m and communication complexity c, and $\Delta = t(2m + c + 1) + \log(1/\epsilon) + 2$ then $\mathsf{val}(G^E) \leq (1 - \frac{\epsilon}{2})^t$.*

Theorem 4.1 formalizes the statement of both informal theorems (Theorems 1.5 and Theorem 1.9) stated in the introduction. Below we explain that the parameters guaranteed by the two informal theorems indeed follow by plugging known explicit constructions of extractors.

We start by observing that some of the quantities in Theorem 4.1 can be simplified 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 4.1 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 in Theorem 4.1. In summary, for 2P1R games we can set $\Delta = O(t(m + \ell))$ and for communication games we can set $\Delta = O(tm)$. We now consider specific choices of extractors.

Parallel repetition as a strong extractor One possible choice for E is "independent parallel repetition". Namely $r = nm$ and for $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_n) \in (\{0,1\}^m)^n \cong \{0,1\}^r$ we define $E((\bar{x}_1, \ldots, \bar{x}_n), i) = \bar{x}_i$. By

[4] More precisely, let G be a free game in which whether or not the players win depends only on whether the input of the first player lands in some set A. To derandomize the parallel repetition G^n of such a game we must use a sample space in which the n inputs $x_1, \ldots, x_n$ of the first player have the property that with high probability there exist $i \in [n]$ such that $x_i \in A$ (which is precisely the guarantee of hitters).

Lemma 3.1, E is a $(r-\Delta,\epsilon)$-extractor for $n=O(\Delta/\epsilon^2)$. Plugging this extractor into Theorem 4.1 gives a proof of the parallel repetition theorem for free games.[5]

Using strong extractors to obtain the parameters guaranteed in Theorems 1.5,1.9 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 $E:\{0,1\}^r\times[n]\to\{0,1\}^m$ 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 $n=\text{poly}(\Delta/\epsilon)$ [30, 13]. (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 4.1 and using the simplifications explained above gives the parameters guaranteed in Theorems 1.5,1.9. More specifically, when starting with a free game G with $\mathsf{val}(G)\leq 1-\epsilon$ we construct a game G^E with $\mathsf{val}(G^E)\leq(1-\epsilon/2)^t$. For a 2P1R-game G, the randomness complexity of G^E is $\mathsf{rand}(G^E)=O(t(m+\ell))$ and it uses $n=\text{poly}(t,m,\ell)$ repetitions. This should be compared to independent parallel repetition that uses $n=O(t\ell/\epsilon)$ repetitions and randomness complexity $\mathsf{rand}(G^n)=O(tm\ell/\epsilon)$ 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 $n=\text{poly}(t,m)$ repetitions. This should be compared to independent parallel repetition that uses $n=O(tc/\epsilon)$ repetitions and randomness complexity $\mathsf{rand}(G^n)=O(tmc/\epsilon)$ for the same goal.

Derandomized parallel repetition of games with value approaching zero Theorem 4.1 is tailored to handle games with value approaching 1. We remark that we can also tailor it for games with value approaching zero. 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 $\mathsf{val}(G^t)\leq(2\epsilon)^t$.

4.1 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 4.1 are met and consider a probability space consisting of two independent random variables $\bar{X},\bar{Y}$ that are uniformly distributed over $\{0,1\}^r$. Note that events in this probability space correspond to subsets $T\subseteq(\{0,1\}^r)^2$.

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 $a,b:\{0,1\}^r\to(\{0,1\}^\ell)^n$ 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 $\Pr[W_{[n]}]\leq(1-\frac{\epsilon}{2})^t$.

4.1.1 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

[5] We remark that the the proof of [2] 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 4.1. Loosely speaking, the proof of [2] 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.

$\mathsf{val}(G) = 1/2$ and strategies with $\Pr[W_1] = 1/2$, but $\Pr[W_2|W_1] = 1$ [9, 19, 5]. We follow the strategy suggested in [24] and prove the following lemma.

Lemma 4.2. *Let $S \subseteq [n]$ with $|S| \leq t$ and $\Pr[W_S] \geq (1 - \frac{\epsilon}{2})^t$. There exists $i \notin S$ such that $\Pr[W_i|W_S] \leq 1 - \frac{\epsilon}{2}$.*

Proof of Theorem 4.1 using Lemma 4.2 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 $\Pr[W_S] \leq (1 - \frac{\epsilon}{2})^k$. At each step, if $\Pr[W_S] \leq (1 - \frac{\epsilon}{2})^t$ then we are done. Otherwise, Lemma 4.2 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 $\Pr[W_S] \leq (1 - \frac{\epsilon}{2})^t$ and we are done.
In the remainder of the section we prove Lemma 4.2.

4.1.2 The value of conditioned games

Lemma 4.2 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 4.3 (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 $(\{0,1\}^m)^2$.*

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 inputs (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 $\Pr[W_i|T] \leq \mathsf{val}(G) + \text{error}(T, i)$. However, this is not true in general. The reason is that when the players play the conditioned game, they are not forced to 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 [5, 8] 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 4.4 (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 4.5. *If T is a rectangle then for every $i \in [n]$, $\Pr[W_i|T] \leq \mathsf{val}(G) + \text{error}(T, i)$.*

Proof. We show how to use the strategy Π^E in G^E to define a strategy Π in G that wins with probability $\Pr[W_i|T] - \text{error}(T,i)$. 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 $(\bar{x},\bar{y}) \leftarrow ((X',Y')|X'_i = x, Y'_i = y)$. (Note that here we are critically using the fact that T is a rectangle. Recall that the first player does not receive y and the second player does not receive x. Thus, if the random variables X', Y' are correlated, the two players may not be able jointly sample from $((X',Y')|X'_i = x, Y'_i = y)$ without communicating).
- 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 $\text{error}(T,i)$. 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 $\text{error}(T,i)$ 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 $\text{error}(T,i)$. □

4.1.3 The role of extractors

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

Lemma 4.6. *If T is a rectangle with deficiency Δ then $\mathbb{E}_{i\leftarrow[n]}[\text{error}(T,i)] \leq \frac{\epsilon}{4}$*

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

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

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 4.5 and Lemma 4.6 gives that there exists an i such that $\Pr[W_i|T] \leq \mathsf{val}(G) + \epsilon/4 \leq 1 - \epsilon/2$.

4.1.4 Proof of Lemma 4.2

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 4.2 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 4.1).

Lemma 4.7. *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 4.7 appears in Section 4.1.5. We are now ready to prove Lemma 4.2 and conclude the proof of Theorem 4.1.

Proof. (of Lemma 4.2) Given a set S that satisfies the requirements of Lemma 4.2 we can apply Lemma 4.7 and let $T_0, \ldots, T_L$ be the events that are guaranteed by Lemma 4.7. We first use Lemma 4.5 to estimate $\Pr[W_i|W_S]$ for a fixed $i \in [n]$.

$$\begin{aligned}
\Pr[W_i|W_S] &= \sum_{0 \leq j \leq L} \Pr[W_i|T_j] \cdot \Pr[T_j|W_S] \\
&\leq \Pr[T_0|W_S] + \sum_{1 \leq j \leq L} \Pr[T_j|W_S] \cdot (\mathsf{val}(G) + \mathrm{error}(T_j, i)) \\
&\leq \frac{\epsilon}{4} + \mathsf{val}(G) + \sum_{1 \leq j \leq L} \Pr[T_j|W_S] \cdot \mathrm{error}(T_j, i)
\end{aligned}$$

We now use the bound above, Lemma 4.6 and the linearity of expectation to estimate $\mathbb{E}_{i\leftarrow[n]}\big[\Pr[W_i|W_S]\big]$.

$$\begin{aligned}\mathbb{E}_{i\leftarrow[n]}\big[\Pr[W_i|W_S]\big] &\leq \frac{\epsilon}{4} + \mathsf{val}(G) + \sum_{1\leq j\leq L}\Pr[T_j|W_S]\cdot\mathbb{E}_{i\leftarrow[n]}[\text{error}(T_j,i)]\\ &\leq \frac{\epsilon}{4} + \mathsf{val}(G) + \sum_{1\leq j\leq L}\Pr[T_j|W_S]\cdot\frac{\epsilon}{4}\\ &\leq \frac{\epsilon}{2} + \mathsf{val}(G)\end{aligned}$$

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 $i\notin S$. □

4.1.5 Proof of Lemma 4.7

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 $E_1^{\hat{x}}=\big\{\bar{X}_S=\hat{x}\big\}$. Similarly, for every possible value $\hat{y}$ of $\bar{Y}_S$ we define the event $E_2^{\hat{y}}=\big\{\bar{Y}_S=\hat{y}\big\}$. 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 $F_1^{\hat{a}}=\big\{\bar{A}_S=\hat{a}\big\}$. For every possible value $\hat{b}$ of $\bar{B}_S$ we define the event $F_2^{\hat{b}}=\Big\{\bar{B}_S=\hat{b}\Big\}$. 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 $F^{\hat{q}}=\big\{\bar{Q}_S=\hat{q}\big\}$. 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 $T_1, \ldots, T_L$ 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 4.8. *Lemma 4.7 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 [22] 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.*

5 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 subfamilies of games such as "projection games".

We now explain which parts of the proof of Theorem 4.1 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 3. 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 $E : \{0,1\}^r \times [n] \to \{0,1\}^{\mathsf{rand}(G)}$ 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_S^E 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 4.1 work for general games. Specifically, Lemma 4.6 and Lemma 4.7 follow exactly as stated with the modified definitions explained above.

The difficulty is in extending Lemma 4.5. 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 4.5 in which the conclusion is only guaranteed for a rectangle with deficiency Δ and a random i. The original proof of the latter statement also relies on Lemma 3.1 and this suggests that it may be possible to derandomize it using strong extractors. However, it seems to us that these extractors will need to have many additional properties to make the argument go through.

Acknowledgements

I am grateful to Avi Wigderson for introducing me to this problem and for many discussions.

References

[1] Boaz Barak, Moritz Hardt, Ishay Haviv, Anup Rao, Oded Regev, and David Steurer. Rounding parallel repetitions of unique games. In *FOCS*, pages 374–383, 2008.

[2] Boaz Barak, Anup Rao, Ran Raz, Ricky Rosen, and Ronen Shaltiel. Strong parallel repetition theorem for free projection games. In *APPROX-RANDOM*, pages 352–365, 2009.

[3] Mihir Bellare, Oded Goldreich, and Shafi Goldwasser. Randomness in interactive proofs. *Computational Complexity*, 3:319–354, 1993.

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

[5] 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.

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

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

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

[9] 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.

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

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

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

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

[14] Thomas Holenstein. Parallel repetition: simplifications and the no-signaling case. In *STOC*, pages 411–419, 2007.

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

[16] Russell Impagliazzo, Ragesh Jaiswal, Valentine Kabanets, and Avi Wigderson. Uniform direct product theorems: simplified, optimized, and derandomized. In *STOC*, pages 579–588, 2008.

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

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

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

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

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

[22] Itzhak Parnafes, Ran Raz, and Avi Wigderson. Direct product results and the gcd problem, in old and new communication models. In *STOC*, pages 363–372, 1997.

[23] Anup Rao. Parallel repetition in projection games and a concentration bound. In *STOC*, pages 1–10, 2008.

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

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

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

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

[28] Andrew C. Yao. Theory and applications of trapdoor functions (extended abstract). In *FOCS*, pages 80–91, 1982.

[29] Andrew Chi-Chih Yao. Some complexity questions related to distributive computing. In *STOC*, pages 209–213, 1979.

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