The Structure of Winning Strategies in Parallel Repetition Games

Irit Dinur[1] and Elazar Goldenberg[2]

[1] Weizmann Institute of Science,irit.dinur@weizmann.ac.il*
[2] Weizmann Institute of Science,elazar.goldenberg@weizmann.ac.il

Abstract. Given a function $f : X \to \Sigma$, its ℓ-wise direct product is the function $F = f^\ell : X^\ell \to \Sigma^\ell$ defined by $F(x_1, \dots, x_\ell) = (f(x_1), \dots, f(x_\ell))$. A two prover game G is a game that involves 3 participants: $V, \mathcal{A},$ and $\mathcal{B}$. V picks a random pair (x, y) and sends x to $\mathcal{A}$, and y to $\mathcal{B}$. $\mathcal{A}$ responds with $f(x)$, $\mathcal{B}$ with $g(y)$. $\mathcal{A}, \mathcal{B}$ win if $V(x, y, f(x), g(y)) = 1$. The repeated game G^ℓ is the game where $\mathcal{A}, \mathcal{B}$ get ℓ questions in a single round and each of them responds with an ℓ symbol string (this is also called the parallel repetition of the game). $\mathcal{A}, \mathcal{B}$ win if they win each of the questions.

In this work we analyze the structure of the provers that win the repeated game with non negligible probability. We would like to deduce that in such a case $\mathcal{A}, \mathcal{B}$ must have a global structure, and in particular they are close to some direct product encoding.

A similar question was studied by the authors and by Impagliazzo et. al. in the context of testing Direct Product. Their result can be be interpreted as follows: For a specific game G, if $\mathcal{A}, \mathcal{B}$ win G^ℓ with non negligible probability, then $\mathcal{A}, \mathcal{B}$ must be close to be a direct product encoding. We would like to generalize these results for any 2-prover game.

In this work we prove two main results: In the first part of the work we show that for a certain type of games, there exist $\mathcal{A}, \mathcal{B}$ that win the repeated game with non negligible probability yet are still very far from any Direct Product encoding. In contrast, in the second part of the work we show that for a certain type of games, called "miss match" games, we have the following behavior. Whenever $\mathcal{A}, \mathcal{B}$ win non negligibly then they are both close to a Direct Product strategy.

* Work supported by ISF grant 1179/09, BSF grant 2008293, and ERC starting grant 239985.

1 Introduction

Given a function $f : S \to \Sigma$ its ℓ-wise direct product is the function $f^\ell : S^\ell \to \Sigma^\ell$ defined by: $f^\ell(s_1, \ldots, s_\ell) = (f(s_1), \ldots, f(s_\ell))$. The Direct Product Testing Theorem by [DG08] and [IKW09] asserts that there exists a two query test T such that, whenever a function $F : S^\ell \to \Sigma^\ell$ passes T with non negligible probability, then F is somewhat close to an ℓ-wise direct product for some global function $f : S \to \Sigma$.

Let us describe the 2-query direct product test T. The test picks a random tuple $\mathbf{x} \in S^\ell$ and then picks another tuple $\mathbf{x}'$ as follows: For each such coordinate i with probability α $\mathbf{x}'_i = \mathbf{x}_i$, otherwise, $\mathbf{x}'_i$ is drawn uniformly at random from S. The test queries $F(\mathbf{x}), F(\mathbf{x}')$ and accepts if and only if $F(\mathbf{x}), F(\mathbf{x}')$ are consistent among the common values of $\mathbf{x}$ and $\mathbf{x}'$.

The test can be viewed as a repeated 2-prover equality game in the following way: The original game, EQ, is the game in which with probability α $\mathcal{A}, \mathcal{B}$ get the same question x and with probability $1 - \alpha$ they get two independent questions x and x'. $\mathcal{A}$ responds with $a \in \Sigma$ and $\mathcal{B}$ with $b \in \Sigma$. If $\mathcal{A}, \mathcal{B}$ get the same question the verifier checks that $a = b$ otherwise it always accepts. The repeated game EQ^ℓ- the game where the verifier picks ℓ independent pairs of questions and sends them in a single round- is exactly the test T described above. The Direct Product Testing Theorem asserts that for this specific game: Whenever the provers win with non negligible probability, then the provers' strategy has a global structure: They have a global agreement with some direct product function.

The Parallel Repetition Theorem by [Raz98] asserts that, for any 2-prover game, the value of the repeated game decreases exponentially with the number of repetitions. Thus, if the provers win the repeated game with probability above 1%, then the value of the original game is almost 1. The Parallel Repetition Theorem concludes nothing about the structure of the provers' strategy assuming they win with probability above 1%. Furthermore, it is easy to see that the value of the EQ game is 1. Therefore, the Parallel Repetition Theorem, unlike the Direct Product Testing Theorem, tells us nothing about EQ^ℓ.

This work is a bridge between the Parallel Repetition Theorem and the Direct Product Testing Theorem showing that for every 2-prover game, if $\mathcal{A}, \mathcal{B}$ win with non negligible probability, then $\mathcal{A}, \mathcal{B}$ have global structure, namely $\mathcal{A}, \mathcal{B}$ are close to a direct product encoding.

Let us introduce some of our notations: A two-prover game G is defined by a distribution $\mathcal{D}$ on questions (X, Y) and a verifier V. The verifier V picks a questions pair $(x, y) \in (X, Y)$ according to $\mathcal{D}$. Then, the verifier sends the question x to prover $\mathcal{A}$ and the question y to prover $\mathcal{B}$. The provers $\mathcal{A}, \mathcal{B}$ are not allowed to communicate with each other during the game, and $\mathcal{A}$ responds with $f(x)$, while $\mathcal{B}$ responds with $g(y)$. The players win if $V(x, y, f(x), g(y)) = 1$. The value of the game G, denoted $val(G)$ is the maximum success probability of the players.

For functions $f : X \to \Sigma_A$ and $g : Y \to \Sigma_B$ we denote by $val(G, f, g)$ the value of the game if A plays according to f and $\mathcal{B}$ according to g, i.e. $val(G, f, g) = \mathbf{E}_{(x,y)\sim\mathcal{D}} V(x, y, f(x), g(y))$. We call the pair (f, g) a perfect strategy if $val(G, f, g) = 1$.

The repeated game G^ℓ is the the game where V samples ℓ independent questions: $(x_1, y_1), \ldots, (x_\ell, y_\ell)$ each is distributed according to $\mathcal{D}$. The verifier sends $\mathbf{x} = (x_1, \ldots, x_\ell)$ to $\mathcal{A}$ and $\mathbf{y} = (y_1, \ldots, y_\ell)$ to $\mathcal{B}$. Each prover responds with ℓ answers. The provers win if they win each of the ℓ coordinates. A *projection game* is a game in which the predicate V has a special structure- every pair (x, y) defines a function $\Pi_{x,y} : \Sigma_A \to \Sigma_B$, and $V(x, y, a, b)$ is satisfied iff $\Pi_{x,y}(f(x)) = g(y)$.

As mentioned earlier, the Parallel Repetition Theorem by [Raz98] bounds the value of the repeated game. Roughly speaking, it says that for every game G, if $val(G) < 1-\varepsilon$, then $val(G^\ell) < (1-\varepsilon')^\ell$ (where ε' depends on ε and on the length of the answer in G).

How would honest verifiers $\mathcal{A}, \mathcal{B}$ play in order to win the repeated game? They choose a pair of perfect strategies (f, g). $\mathcal{A}$, upon receiving $(x_1, \ldots, x_\ell)$, answers with $(f(x_1), \ldots, f(x_\ell))$ while $\mathcal{B}$ answers with $(g(y_1), \ldots, g(y_\ell))$.**(Irit:** In fact, $\mathcal{A}, \mathcal{B}$ can choose ℓ pairs of perfect strategies $(f_1, g_1), \ldots, (f_\ell, g_\ell)$ and $\mathcal{A}$ answers with $(f_1(x_1), \ldots, f_\ell(x_\ell))$ while $\mathcal{B}$ answers with $(g_1(y_1), \ldots, g_\ell(y_\ell))$ and still win with probability 1. We call such strategies $\mathcal{A}, \mathcal{B}$ direct product strategies and denote them by $\prod f_i$ and $\prod g_i$. **)**

In this work, we consider the case where the provers win the repeated game with non negligible probability. We would like to deduce a **structure** for the provers' strategies. Ideally, such strategies are approximately direct product strategies, in other words, global structure. Let us call this the **Global Structure Hypothesis**.

(Elazar: Without loss of generality we focus only on non trivial games, i.e. games in which for every questions pair (x, y) there exists a pair of answers (a, b) such that $V(x, y, a, b) = 0$. Otherwise, if the verifier always accepts, then it is trivial that we cannot expect of $\mathcal{A}, \mathcal{B}$ being structured, since every $\mathcal{A}, \mathcal{B}$ win with probability 1. **)**

Results. Our first result is that the Global Structure Hypothesis does not hold in general even for non trivial games. We show games, for which there exists a strategy for $\mathcal{A}, \mathcal{B}$ that is extremely far from any direct product strategy (i.e. has no global structure) while attaining constant winning probability . We conclude, (perhaps surprisingly), that high success probability does not imply global structure.

Our main negative result shows that the Global Structure Hypothesis fails for any constant degree game [3] that has a large number of perfect strategies that are pairwise far apart:

Theorem 1 (Anti Structural Theorem- Informal Statement). *There exists a non-trivial constant degree game G, and constant α such that for every*

[3] the degree is the maximal number of neighbors of a certain question

ℓ: There exist strategies $\mathcal{A}, \mathcal{B}$ such that the maximal agreement between $\mathcal{A}$ and $\prod f_i$ for any $\prod f_i$ is at most $2^{-\omega(\ell)}$, and similarly for $\mathcal{B}$. Yet, $\mathcal{A}, \mathcal{B}$ win with probability α.

We extend Theorem 1 for games with unbounded degree, and also for the so-called "permuting" verifiers that permute the questions (these were called "clever" in [FK95]). For details see Section 3.

In the second part of the work we show, as our second result, that in contrast to Theorem 1 the Global Structure Hypothesis is true for a certain type of games called "miss-match" games. These games were first studied in [FK94].

Given a 2-prover game G its repeated "miss-match" game, denoted by $G^{m,\ell}$, $0 < m < \ell$, is as follows: The verifier chooses m coordinates, on each such coordinate it samples a pair (x, y) according to the distribution of G, these are called the *match* coordinates. As for the rest of the coordinates, the so called *miss* coordinates, the verifier picks $x \in X$ and $y \in Y$ independently uniformly at random. The provers answer with an ℓ symbols string and they win the game if they win each of the match coordinates.[4]

(Irit: Let us make the following simplifying assumption regarding the strategies. We assume that the answers of each player on a given tuple of questions $(x_1, \ldots, x_\ell)$ depend on the set $\{x_1, \ldots, x_\ell\}$ of questions but not on their order. This allows us to focus only on strategies f^ℓ rather than $\prod f_i$ and saves some technical complications. We believe that our results can be directly extended to the general (ordered) case as in say [DG08], although we have not fully checked the details.**)**

We first show that for every projection game G, if the provers win $G^{m,\ell}$ with non negligible probability ε, then $\mathcal{B}$ plays according to a direct product strategy: We show that there exists a small $(poly(1/\varepsilon))$ list of functions $g_1, \ldots, g_t : Y \to \Sigma_B$ such that $\mathcal{B}$ agrees non-negligibly with g_i^ℓ for each i. Furthermore, we show that essentially the only way $\mathcal{A}, \mathcal{B}$ win is whenever $\mathcal{B}(\mathbf{y}) \overset{g^{\ell}}{\approx}_i (\mathbf{y})$ where g_i is some function from the list.

Theorem 2 (Informal Statement). *Let G be a projection game. Assume $\mathcal{A}, \mathcal{B}$ win $G^{m,\ell}$ with probability $\varepsilon > \ell^{-\Omega(1)}$, then there exists a small list of t functions $g_1, \ldots g_t : Y \to \Sigma_A$ such that:*

- *For each $i \in [t]$: $\Pr_{\mathbf{y}}[\mathcal{B}(\mathbf{y}) \approx g_i^\ell(\mathbf{y})] > \varepsilon'$, where $\varepsilon' = poly(\varepsilon)$.*
- *$\Pr[\exists i$ s.t. $\mathcal{B}(\mathbf{y}) \approx g_i^\ell(\mathbf{y}) | \mathcal{A}, \mathcal{B}$ win$] \geq 1 - o(1)$.*

The proof resembles [DG08] and [IKW09] and appears in Section 4.1.

Note that Theorem 2 only discusses $\mathcal{B}$'s strategy. It turns out that deducing a similar result for $\mathcal{A}$ is more subtle, and is only true if G is **smooth** enough. This smoothness parameter, first defined by[HK04], is as follows:

[4] Alternatively, we can define "miss-match" as follows: Given a game G, we define $mm - G$ as the game that with probability $\alpha = (m/\ell)$ the verifier plays the original game G, and with probability $1 - \alpha$ it picks two independent questions and always accept. The repeated game $(mm - G)^\ell$ is very similar to $G^{m,\ell}$.

Definition 1. *A projection game G is called α-smooth if for every $x \in X$ and distinct answers $a, a' \in \Sigma_A$, we have:* $\Pr_y[\Pi_{x,y}(a) = \Pi_{x,y}(a')] < 1 - \alpha$*, where y is a random neighbor of x.*

Assuming the game is sufficiently smooth, we show an analog of Theorem 2, namely: we show that whenever $\mathcal{A}, \mathcal{B}$ win $G^{m,\ell}$ with non negligible probability, then there exists a short list of functions pairs $(f_1, g_1), \ldots, (f_s, g_s)$ such that: $\mathcal{A}, \mathcal{B}$ agree with f_i^ℓ, g_i^ℓ non-negligibly, and $val(G, f_i, g_i)$ is close to 1. We also prove that if $\mathcal{B}$ plays on $\mathbf{y}$ according to g_i^ℓ, and $\mathcal{A}$ does not play according to f_i^ℓ, or vice versus, then with high probability $\mathcal{A}, \mathcal{B}$ lose. Combining with Theorem 2 we get that there exists a small list of functions pairs (f_i, g_i), such that the only way to win the repeated game is whenever $\mathcal{A}$ plays according to direct product of f_i^ℓ while $\mathcal{B}$ plays according to g_i^ℓ. Thus, we fully explain the high winning probability of the provers through a direct product structure of their strategies.

Theorem 3 (Informal Statement). *Let G be a an α smooth projection game (where α is a constant). Assume $\mathcal{A}, \mathcal{B}$ win $G^{m,\ell}$ with probability $\varepsilon > \ell^{-\Omega(1)}$, then there exists a small list of s pairs of functions $(f_1, g_1), \ldots (f_s, g_s)$ such that:*

- $f_i : X \to \Sigma_A$, $g_i : Y \to \Sigma_B$ *and:* $val(G, f_i, g_i) > 1 - o(1)$.
- *Let (A, B) be a random pair of questions, then:*

$$\Pr[\exists i \ s.t. \ \mathcal{A}(\mathbf{x}) \approx f_i^\ell(\mathbf{x}) \ and \ \mathcal{B}(\mathbf{y}) \not\approx g_i^\ell(\mathbf{y}) | \mathcal{A}, \mathcal{B} \ win\ ] > 1 - o(1).$$

The smoothness property is essential for Theorem3. Theorem 4 shows a game that is not smooth enough, for which there exist strategies $\mathcal{A}, \mathcal{B}$ that win the game with probability 1, yet $\mathcal{A}$ is unstructured.

Theorem 4 (Informal Statement).

There exists a projection game G, such that for every ℓ there exist strategies $\mathcal{A}, \mathcal{B}$ such that the maximal agreement between $\mathcal{A}$ and f^ℓ for any f is at most $2^{-\omega(\ell)}$. Yet, $\mathcal{A}, \mathcal{B}$ win $G^{m,\ell}$ with probability 1.

Additional Motivation and Context. The study of structure of winning strategies, aside from being an interesting generalization of the direct product testing question, has also some additional motivation coming from PCP constructions.

In recent years, stronger variants of PCPs called PCPPs [BSGH$^+$06] or assignment testers [DR06] and more recently dPCPs [DH09] have been introduced. These are constructs that are similar to PCPs but are stronger, and much more useful in composition. Without getting into the details, let us say that the main difference between these objects and regular PCPs lies in the soundness criterion. The difference is closely related to the difference between just knowing that the soundness error of repeated games is small (this only gives a PCP), and between being able to say that strategies that have non-negligible winning probability must be structured as direct products (such a result will give you the stronger object, i.e., a dPCP or a PCPP). Whereas the former is already given by the parallel repetition theorem of [Raz98], the later is the content of this work.

In fact, our structure result (Theorem 3) can be used in order to show that a parallel repetition of a dPCP is a dPCP with amplified soundness. However, since this has already been done (with better parameter setting) in [DM10], we do not work out the details here.

Future Work. In this work we deal with several types of games and repetitions. We show that for part of them, such as $G^{m,\ell}$ the Global Structure Hypothesis holds. Contrary, we show hat for other types of games, such as constant degree games with many perfect strategies, the hypothesis fails. It would be interesting to characterize the types of games and repetitions for which the hypothesis holds.

Organization of the Paper. Subsection 2.1 shows the Direct Product Lemma which is the basis for our approach. In section 3 prove Theorem 1. Finally, in Section 4 we prove Theorem 2 and Theorem 3.

2 Preliminaries

In this work we deal with several kinds of repetitions: Repetition where the provers gets ordered tuples, sets and multisets.

When the provers get ordered tuples, then we see them as **tuple oracles**: $\mathcal{A}$ gets a tuple $\mathbf{x} = (x_1, \ldots, x_\ell) \in X^\ell$ and responds with $\mathcal{A}(\mathbf{x}) \in \Sigma_A^\ell$, and $\mathcal{B}$ gets a tuple $\mathbf{y} = (y_1, \ldots, y_\ell) \in Y^\ell$ and responds with $\mathcal{B}(\mathbf{y}) \in \Sigma_B^\ell$.**(Elazar:** Let us define the product encoding of functions $(f_1, \ldots, f_\ell)$, $f_i : S \to \Sigma$, to be a tuple oracle, $\prod f_i$, assigning for every tuple $(s_1, \ldots, s_\ell)$ the value $(f_1(s_1), \ldots, f_\ell(s_\ell))$. In the case where $f_1 = \ldots = f_\ell = f$ we denote $\prod f_i$ by f^ℓ. **)**

When the provers get multi-sets, then we see them as **multi-set oracles**: $\mathcal{A}$ gets a multi-set $A = \{x_1, \ldots, x_\ell\}$ and responds with $\mathcal{A}(A)$which is a function $A \to \Sigma_A$. $\mathcal{B}$ gets a multi-set $B = \{y_1, \ldots, y_\ell\}$ and responds with $\mathcal{B}(B)$ which is a function $B \to \Sigma_B$. Let us define the ℓ multi-set direct product encoding of a function $f : S \to \Sigma$ to be a multi-set oracle, f^ℓ, assigning for every $T \subset S$ of cardinality ℓ the restriction of f to T.

When the provers get sets, then see them as a **sets oracles**: The definitions are identical to multi-sets oracles besides that in this case the provers gets sets rather than multi-sets.

For a function $f : S \to \Sigma$, and $T \subset S$ we denote by f_T the restriction of f to T. The definition of the support of f is important in our discussion:

Definition 2. *For two vectors* $\mathbf{v}, \mathbf{w}$ *in some alphabet* Σ^ℓ *we write* $\mathbf{v} \overset{\rho}{\approx} \mathbf{w}$ *to denote* $\Pr_{i\in[\ell]}[\mathbf{v}_i = \mathbf{w}_i] \geq 1-\rho$ *and* $\mathbf{v} \overset{\rho}{\not\approx} \mathbf{w}$ *to denote* $\Pr_{i\in[\ell]}[\mathbf{v}_i \neq \mathbf{w}_i] \geq \rho$.

For two function $f, g : T \to \Sigma$ *we write:* $f \overset{\rho}{\approx} g$ *to denote* $\Pr_{t\in T}[f(t) = g(t)] \geq 1-\rho$ *and* $f \overset{\rho}{\not\approx} g$ *to denote* $\Pr_{t\in T}[f(t) \neq g(t)] \geq \rho$.

For a tuple oracle F *and* $f : S \to \Sigma$ *the* $\rho-$*support denoted by* $\mathrm{supp}_\rho^F(f)$ *defined as follows:* $\mathrm{supp}_\rho^F(f) = \{\mathbf{s} \in S^\ell \mid F(\mathbf{s}) \overset{\rho}{\approx} f^\ell(\mathbf{s})\}$.

For a multi-set oracle F and $f : S \to \Sigma$ the $\rho-$support denoted by $\mathrm{supp}_\rho^F(f)$ *defined as follows:* $\mathrm{supp}_\rho^F(f) = \{A \subset S \mid |A| = \ell \text{ and } F(A) \overset{\rho}{\approx} f^\ell(A)\}$.

Now we would like to introduce "miss match" games in these settings:

Definition 3. ***"Miss-Match" Games:*** *Let G be a game, let ℓ, m be integers $0 < m < \ell$, then we define the miss-match, $G^{m,\ell}$, as follows:*

1. *The verifier picks m pairs (x_i, y_i) where each pair is selected independently according to $\mathcal{D}$. The verifier defines a multiset $A' = \{x_1, \dots x_m\}$ and $B' = \{y_1, \dots, y_m\}$. These are the match elements, each pair (x_i, y_i) is called a match pair and A', B' are called the match questions.*
2. *The verifier picks $\ell - m$ additional pairs (x_j, y_j), where x_j, y_j are chosen independently at random from X, Y (respectively). The verifier defines multisets $A'' = \{x_{m+1}, \dots, x_\ell\}$ and $B'' = \{y_{m+1}, \dots, y_\ell\}$. These are the confuse elements.*
3. *V sends $A = A' \cup A''$ to $\mathcal{A}$, and $B = B' \cup B''$ to $\mathcal{B}$.*
4. *$\mathcal{A}$ responds with $\mathcal{A}(A) : A \to \Sigma_A$, and $\mathcal{B}$ responds with $\mathcal{B}(B) : B \to \Sigma_B$ ($\mathcal{A}, \mathcal{B}$ are multiset-oracles). The provers win G^ℓ if they win each of the match elements, i.e. for every match pair (x, y) we have:*

$$V(x, y, \mathcal{A}(A)_x, \mathcal{B}(B)_y) = 1.$$

(Irit: We note that in our definitions of "miss-match" games each prover gets a multiset from the verifier, as opposed to an ℓ-tuple. Thus, the provers are multiset oracles. Alternatively, one can define miss-match games such that the verifier sends tuples, and performs a random shuffle on the coordinates. This turns to be equivalent to the case of sets. A similar reduction was done in [DG08].)

2.1 Testing Direct Product

We now turn to describe the Direct Product Testing Lemma as in [DG08] and in [IKW09]. Let F be a ℓ set oracle that works over a set X. The goal is to test whether F is close to a direct product encoding- i.e. whether there exists f such that F is the direct product encoding of f. A two queries test that resemble the "miss match" game is used. The test chooses a random subset A and a random subset B as follows: A and B share m elements in common. As for the rest elements of B the test picks $\ell - m$ random elements from X. Then the test checks for consistency among $F(A)$ and $F(B)$ i.e. for each common element x it verifies that $F(A)_x = F(B)_x$.

The following definition is quoted from [DG08].

Definition 4. *Let $\mathcal{B}$ a ℓ set oracle that works over a set Y. Let $B' \subset Y$ of cardinality m. We call B' ε-alive if there exists $b' : B' \to \Sigma_B$ such that:*

$$\Pr_{B \supset B'}[\mathcal{B}(B)_{B'} = b'] \geq \varepsilon$$

Such an answer b' is called a live answer for B'.

Now we are ready to state Theorem 3.14 from [IKW09]. This is a local to global Lemma that claims that as long as there exist many live sets (the local property), then this implies an existence of a direct product function with a large support (the global property).

Theorem 5 (Direct Product Testing:). *There exists $\ell_0 \in \mathbb{N}$ and $c > 0$ such that for every $\ell > \ell_0$: Let $\mathcal{B}$ be a ℓ set-oracle such that*

$$\Pr_{B' \subset Y || B'| = \sqrt{\ell}}[B' \text{ is } \varepsilon/2\text{-alive}] \geq \varepsilon/2,$$

where $\varepsilon \geq 1/\sqrt{\ell}$. Then, there exists a function $g : Y \to \Sigma_B$ such that $\mathcal{B}(B) \overset{\rho}{\approx} g^\ell(B)$ for at least $\Omega(\varepsilon^6)$ of the $B \in \binom{Y}{\ell}$, where $\rho \leq \ell^{-c}$.

3 Negative Results

In this section we prove Theorem 1 showing that, for any constant degree game G with many perfect strategies, $\mathcal{A}, \mathcal{B}$ can win G^ℓ with constant probability and still be very far from any **(Elazar:** generalized product) strategy. We extend Theorem 1 for games of of non-constant degree in Theorem 6. Theorem 7 extends Theorem 1 to handle "Permuting Verifiers".

For a game G we define a bipartite weighted graph, where $L = X$, $R = Y$ and $w_{x,y} = \Pr_{\mathcal{D}}[y|x]$. The game is called d regular if the degree of every left node is d, the degree of every right node is $d\,|X|\,/\,|Y|$, and $w_{x,y} = 1/d$ for every adjacent x and y. d is called the degree of the game. Another property that we take into consideration is the the rate between the cardinalities X and Y. We denote by r the ratio $|X|\,/\,|Y|$, and without loss of generality we assume $r > 1$.

3.1 Proof of Theorem 1

In this section we prove Theorem 1. We define $P_A = \{\prod f_i | f_i : X \to \Sigma_A\}$, and $P_B = \{\prod g_i | g_i : Y \to \Sigma_B\}$. For two functions $F, G : S^\ell \to \Sigma^\ell$ we define their relaxed Hamming distance with parameter γ as: $dist_\gamma(F, G) = \Pr_{\mathbf{s} \in S^\ell}[F(\mathbf{s}) \overset{\gamma}{\not\approx} G(\mathbf{s})]$. Let us first state Theorem 1 formally:

Theorem 1 (Formal Statement) *For every constants $d > 1$ and $0 < \gamma < 1/20$ there exists a non-trivial constant degree d game G, and tuples-oracles strategies $\mathcal{A}, \mathcal{B}$ such that $dist_\gamma(\mathcal{A}, P_A) \geq 1 - (1/\,|Y| + 2^{-\omega(\ell)})$, and $dist_\gamma(\mathcal{B}, P_B) \geq 1 - (1/\,|Y| + 2^{-\omega(\ell)})$. Yet, $\mathcal{A}, \mathcal{B}$ win G^ℓ with probability at least $1/d$.*

The theorem holds for any constant degree d game G, for which there exists a large list of $t = |Y|$ pairs of perfect strategies $(f_1, g_1), \ldots, (f_t, g_t)$ that satisfy: For $i \neq j : dist(f_i, f_j)$, and $dist(g_i, g_j)$ are both greater than 10γ.

The requirement for the distance between the pairs prevents the case where all the perfect strategies have a small relative distance. In such a case all of the above functions pairs (f_i, g_i) could be clustered into a single function pair (f, g) for which: $\mathcal{A}(\mathbf{x}) \approx f^\ell$ and $\mathcal{B}(\mathbf{y}) \approx g^\ell$. Such a behavior can still be viewed as a direct product structure for $\mathcal{A}, \mathcal{B}$.

Proof. The strategies of $\mathcal{A}, \mathcal{B}$ are based on the following combinatorial claim:

Claim. Let $G = (V, E)$ be a bipartite (c, d) regular graph (the left degree is c, and the right degree is d), and assume wlog $c \leq d$. Then there exists a subgraph $G' = (V, E')$ such that G' is $(1, d/c)$ regular.

Due to space limitations the proof is omitted and can found in the full version of the paper.

Let us present the strategies $\mathcal{A}, \mathcal{B}$: As a first step $\mathcal{A}, \mathcal{B}$ match for every $y \in Y$ a pair (f_i, g_i) from the list, so we associate the strategies list with the set Y and we write (f_y, g_y). Then they choose a subgraph G' as in claim 3.1.

$\mathcal{B}$ decides according to value of the first coordinate $\mathbf{y}_1$- i.e. given $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_\ell)$, $\mathcal{B}(\mathbf{y}) = g^\ell_{\mathbf{y}_1}(\mathbf{y})$.

$\mathcal{A}$ strategy is similar, it is also based just on the value of the first coordinate $\mathbf{x}_1$: Given $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_\ell)$, $\mathcal{A}(\mathbf{x}) = f^\ell_{N(\mathbf{x}_1)}(\mathbf{x})$ where $N(\mathbf{x}_1)$ is the vertex y such that $(\mathbf{x}_1, y) \in E'$.

We now turn to prove the success probability of the proves, and the distance between $\mathcal{A}, \mathcal{B}$ and any generalized product strategy.

Note that if $\mathbf{y}_1 = N(\mathbf{x}_1)$, then $\mathcal{A}, \mathcal{B}$ win the game, since they are playing according to $f_{\mathbf{y}_1}, g_{\mathbf{y}_1}$, which is a perfect strategy.

What is the probability that indeed $\mathbf{y}_1 = N(\mathbf{x}_1)$? Note that we care only about the values of the first coordinate. Once $\mathbf{x}_1$ is fixed, the probability that $\mathbf{y}_1 = N(\mathbf{x}_1)$ is exactly $1/d$. Therefore, the winning probability is $1/d$.

(Elazar: The analysis was changed**)** What is the distance between $\mathcal{B}$ and any product $\prod g_i$?

Let $\prod g_i$ be a product strategy, we divide the proof into cases: The case where for every $y \in Y$ it holds that $dist(g_i, g_y) > 5\gamma$ for at least $1/4$ fraction of the g_i, and the case where there exists $y \in Y$ such that $dist(g_i, g_y) \leq 5\gamma$ for at least $3/4$ fraction of the g_i. Note, that since $dist(g_y, g_{y'}) > 10\gamma$ for $y \neq y' \in Y$, then every function g agrees with at most a single function g_y on more than $1 - 5\gamma$ fraction of the domain, and in particular for every i there can be only a single y with $dist(g_i, g_y) \leq 5\gamma$.

Assume we are in the first case:

$$\Pr_{\mathbf{y}}[\mathcal{B}(\mathbf{y}) \overset{\gamma}{\approx} \prod g_i(\mathbf{y})] = \Pr_{y_1}[\Pr_{y_2\ldots,y_\ell}[\mathcal{B}(\mathbf{y}) \overset{\gamma}{\approx} \prod g_i(\mathbf{y})]] = \Pr_{y_1}[\Pr_{y_2\ldots,y_\ell}[g^\ell_{y_1}(\mathbf{y}) \overset{\gamma}{\approx} \prod g_i(\mathbf{y})]]$$

Now, we can use Chernoff inequality to deduce that $\Pr_{y_2\ldots,y_\ell}[g^\ell_{y_1}(\mathbf{y}) \overset{\gamma}{\approx} \prod g_i(\mathbf{y})] < 2^{-\omega(\ell)}$ (the expected number of coordinates on which there is an inequality is at least $5\gamma/4$), so we get that in the first case: $dist(\mathcal{B}, \prod g_i) > 1 - 2^{-\omega(\ell)}$.

As for the second case, where we assume that $\prod g_i$ is close for some function g_y, then:

$$\begin{aligned}\Pr_{\mathbf{y}}[\mathcal{B}(\mathbf{y}) \overset{\gamma}{\approx} \prod g_i(\mathbf{y})] &= \Pr[y_1 = y] \Pr_{y_2\ldots,y_\ell}[g^\ell_y(\mathbf{y}) \overset{\gamma}{\approx} \prod g_i(\mathbf{y})] + \Pr[y_1 \neq y] \Pr_{y_2\ldots,y_\ell}[g^\ell_{y_1}(\mathbf{y}) \overset{\gamma}{\approx} \prod g_i(\mathbf{y})] \\ &\leq 1/|Y| + 2^{-\omega(\ell)}\end{aligned}$$

We get that in this case $dist_\gamma(\mathcal{B}, \prod g_i) \geq 1 - (1/|Y| + 2^{\omega(\ell)})$, and we are done. The analysis for $\mathcal{A}$ is similar.

One may think that Theorem 1 is true just for constant degree game. However, in Theorem 6 we extend Theorem 1 for a certain non-constant game:

Theorem 6. *For every constant $d > 1, 0 < \gamma < 1/8$ there exists a non-trivial non-constant degree $\tilde{d}$ game $\tilde{G}$, and tuple-oracles strategies $\mathcal{A}, \mathcal{B}$ such that $dist_\gamma(\mathcal{A}, P_A) \geq 1 - (\frac{\tilde{d}}{d|Y|} + 2^{-\omega(\ell)})$, and $dist_\gamma(\mathcal{B}, P_B) \geq 1 - (\frac{\tilde{d}}{d|Y|} + 2^{-\omega(\ell)})$. Yet, $\mathcal{A}, \mathcal{B}$ win $\tilde{G}^\ell$ with probability $1/d$.*

The proof of Theorem 6 can be found in the full version of the paper.

Our next negative result is Theorem 7 that extends for "Permuting Verifiers". In this case we would like to view the provers as multi-sets oracles. [FK95] studied this type of verifiers and called them "Clever Verifiers". Let us first introduce them:

Definition 5 (Permuting Verifiers:). *The verifier selects ℓ pairs of questions $(x_1, y_1), \ldots, (x_\ell, y_\ell)$ Each is pair is drawn independently according to the distribution of G. V sends $A = \{x_1, \ldots, x_\ell\}$ to $\mathcal{A}$ (note that A is a multi-set), and $B = \{y_1, \ldots, y_\ell\}$ to $\mathcal{B}$. $\mathcal{A}$ answers with $\mathcal{A}(A)$, and $\mathcal{B}$ with $\mathcal{B}(B)$. The verifier accepts if for every i: $V(x_i, y_i, \mathcal{A}(A)_{x_i}, \mathcal{B}(B)_{y_i}) = 1$.*

We define $DP_A = \{f^\ell | f : X \to \Sigma_A\}$, and $DP_B = \{g^\ell | g : Y \to \Sigma_B\}$. Now let us state Theorem 7 formally:

Theorem 7. *For every constants $d > 1$ and $0 < \gamma < 1/8$, there exists a non-trivial constant degree d game G a constant c, and multiset-oracles strategies $\mathcal{A}, \mathcal{B}$ such that $dist_\gamma(\mathcal{A}, DP_A) \geq 1 - O(\ell/|Y| + 2^{-\omega(\ell)})$, and $dist_\gamma(\mathcal{B}, DP_B) \geq 1 - O(\ell/|Y| + 2^{-\omega(\ell)})$ Yet, $\mathcal{A}, \mathcal{B}$ win G^ℓ against "Permuting Verifier" with probability at least c/d.*

The proof of Theorem 7 can be found in the full version of the paper.

Now we would like to extend Theorem 7 to "miss-match" games. The result for "miss-match" game is weaker: $\mathcal{A}, \mathcal{B}$ can be unstructured and win the the game only with probability $\Omega(\frac{m}{d\ell})$ (and not $1/d$ as before). We address here that if $\mathcal{A}, \mathcal{B}$ win the game with probability $\gg m/\ell$ then we can prove that such a behavior is impossible, see section 4 for details.

Claim. For every constants $d > 1$ and $0 < \gamma < 1/8$, there exists a constant degree d game G, a constant c, and multiset-oracle strategies $\mathcal{A}, \mathcal{B}$ such that $dist_\gamma(\mathcal{A}, DP_A) \geq 1 - O(\ell/|Y| + 2^{-\omega(\ell)})$, and $dist_\gamma(\mathcal{B}, DP_B) \geq 1 - O(\ell/|Y| + 2^{-\omega(\ell)})$. Yet, $\mathcal{A}, \mathcal{B}$ win $G^{m,\ell}$ with probability at least $\frac{cm}{d\ell}$.

The proof of Claim 3.1 can be found in the full version of the paper.

4 Positive Results: "Miss-Match" Games

In this section we show that, unlike general games, "miss match" games have the following property: If $\mathcal{A}, \mathcal{B}$ win "miss match" games with non negligible probability, then there exists a small list of pairs $(f_1, g_1), (f_2, g_2), \ldots$ such that $Val(G, f_i, g_i) \approx 1$ and: If $\mathcal{A}, \mathcal{B}$ win then $\mathcal{A}(A) \approx f_i^\ell(A)$ and $\mathcal{B}(B) \approx g_i^\ell(B)$ for some pair from the list, except with negligible probability.

We first prove Theorem 2. The theorem asserts the above only for $\mathcal{B}$, namely: if the provers win $G^{m,\ell}$ with non negligible probability ε, then $\mathcal{B}$ plays according to a direct product strategy.

It turns out that deducing a similar result for $\mathcal{A}$ is more subtle, and depends on the **smoothness** of the game (see Definition 1). Assuming the game is sufficiently smooth, we obtain in Theorem 3 the desired result claimed above.

We also address the question of whether smoothness is essential for direct product behavior. In subsection 4.3 we show that it is essential. In Theorem 4 we show a game that is not smooth such that $G^{m,\ell}$ can be won with probability 1 and still $\mathcal{A}$ is far from being a direct product strategy.

4.1 Direct Product structure for $\mathcal{B}$

In this section we prove Theorem 2, let us state it formally:

Theorem 2 (Formal Statement) *There exists $\ell_0 \in \mathbb{N}$ and $c > 0$ such that for every $\ell > \ell_0$ the following holds. Let G be a projection game, and let $m = \sqrt{\ell}$, $\varepsilon_0 = 2\sqrt{m/\ell}$ and $\delta = \sqrt{\varepsilon_0}$.*

Assume $\mathcal{A}, \mathcal{B}$ win $G^{m,\ell}$ with probability $\varepsilon > \sqrt{\varepsilon_0}$, then there exists a list of $t = O(1/(\delta \cdot \varepsilon)^6)$ functions $g_1, \ldots g_t : Y \to \Sigma_A$ such that:

- *For each $i \in [t]$: $\Pr_B[\mathcal{B}(B) \overset{\rho}{\approx} g_i^\ell(B)] > \Omega((\delta \cdot \varepsilon)^6)$, where $\rho = \ell^{-c}$.*
- $\Pr[\exists i \text{ } s.t. \text{ } \mathcal{B}(B) \overset{\rho}{\approx} g_i^\ell(B) | \mathcal{A}, \mathcal{B} \text{ } win \text{ }] \geq 1 - \delta$

Before we proceed with the proof, let us make a few remarks:

- The theorem concludes that on many Bs, $\mathcal{B}(B) \overset{\rho}{\approx} g^\ell(B)$ rather than $\mathcal{B}(B) = g^\ell(B)$. This weaker conclusion is inherent as seen by the following example. Take $\mathcal{B} = g^\ell$ and then change each $\mathcal{B}(B)$ arbitrarily in fewer than ℓ/m of the coordinates. With high probability the verifier would not notice the difference between $\mathcal{B}$ and g^ℓ, yet $\mathcal{B}$ is only close to g^ℓ in the above sense.
- We would like to address the relation between m and ℓ and the value of ε in Theorem 2. We have already proved Claim 3.1 that asserts that $\mathcal{A}, \mathcal{B}$ can be far away from direct product encoding and still win $G^{m,\ell}$ with probability $\Omega(\frac{m}{d\ell})$. This enforces two constraints regarding our choice of parameters: First, we need that the winning probability ε would be greater than m/ℓ. Indeed, we prove our theorem for values of ε that are bigger than $\sqrt[4]{m/\ell}$. Second, we must choose $m \ll \ell$, and in this work we focus on $m = \sqrt{\ell}$. We leave the study of the entire range of m, ℓ for future work (We mention that this is an open question even in the Direct Product Testing settings see [GS00], [DR06], [DG08] and [IKW09]).

– We work in the settings where $\ell \ll |Y|$ and in particular $\ell < \sqrt[6]{|Y|}$. This enables us an easy transition between sets and multi-sets.

In order to prove Theorem 2, we first show if $\mathcal{A}, \mathcal{B}$ win then there exists at least one function $g : Y \to \Sigma_B$ such that $\mathcal{B}(B) \approx g^\ell(B)$ on a non negligible part of the domain.

Lemma 1. *There exist $\ell_0 \in \mathbb{N}$, and $c > 0$ such that for every $\ell > \ell_0$ the following holds. Let G be a projection game, and let $\varepsilon_0 = 2\sqrt{\frac{m}{\ell}}$.*

Assume $\mathcal{A}, \mathcal{B}$ win $G^{m,\ell}$ where $m = \sqrt{\ell}$, with probability $\varepsilon > \varepsilon_0$, then there exists a function $g : Y \to \Sigma_B$ such that for at least $\Omega(\varepsilon^6)$ of the ℓ multi-sets B, we have $\mathcal{B}(B) \overset{\rho}{\approx} g^\ell(B)$, where $\rho = \ell^{-c}$.

The proof of Theorem 2 and Lemma 1 can be found in the full version of the paper.

4.2 Direct Product Structure for $\mathcal{A}$

In Section 4.1 we show that for every projection game G, whenever $\mathcal{A}, \mathcal{B}$ win $G^{m,\ell}$ with non-negligible probability, then $\mathcal{B}$'s strategy has a direct product structure. However, we have not involved $\mathcal{A}$ strategy at all. In this section we deduce a similar behavior for $\mathcal{A}$ for smooth games. Let us state Theorem 3 formally:

Theorem 3 (Formal Statement) *There exist $\ell_0 \in \mathbb{N}$, $0 < \alpha < 1$ and $c > 0$ such that for every $\ell > \ell_0$ the following holds. Let G be an α-smooth projection game, and let $\rho = \ell^{-c}, \varepsilon_0 = 2\sqrt{\frac{m}{\ell}}$ and $\delta = \sqrt{\varepsilon_0}$.*

Assume $\mathcal{A}, \mathcal{B}$ win $G^{m,\ell}$, with $m = \sqrt{\ell}$, with probability $\varepsilon > \sqrt{\varepsilon_0}$, then there exists a list of $s = O(1/(\delta\varepsilon)^6)$ pairs of functions $(f_1, g_1), \ldots (f_s, g_s)$ such that:

- *$f_i : X \to \Sigma_A$, $g_i : Y \to \Sigma_B$ and: $val(G, f_i, g_i) > 1 - 10\rho/\alpha$.*
- *Let (A, B) be a random pair of questions. Define the following events:*
  - *$B_1 := B \notin \cup_{i\in[s]}\mathrm{supp}_\rho(g_i)$*
  - *$B_2 := \exists i \in [s]$ s.t. $B \in \mathrm{supp}_\rho(g_i)$ while $A \notin \mathrm{supp}_{6\rho/\alpha}(f_i))$.*
  - *$B_3 := \exists i \in [s]$ s.t. $A \in \mathrm{supp}_{6\rho/\alpha}(f_i))$ while $B \notin \mathrm{supp}_{40\rho/\alpha}(g_i)$.*

  Then:
$$\Pr[\mathcal{A}, \mathcal{B} \text{ win } | B_1 \text{ or } B_2 \text{ or } B_3] < \delta + O(\exp^{-\Omega(\rho^2 m)}).$$

In order to prove Theorem 3 we use the following three lemmas:

Lemma 2. *There exist $\ell_0 \in \mathbb{N}$, $0 < \alpha < 1$ and $c > 0$ such that for every $\ell > \ell_0$ the following holds. Let G be an α-smooth projection game, $g : Y \to \Sigma_B$ and $\rho = \ell^{-c}$. Let $f : X \to \Sigma_A$ be a function that maximizes $val(G, f, g)$, then: If $B \in \mathrm{supp}_\rho(g)$ while $A \notin \mathrm{supp}_{6\rho/\alpha}(f)$. Then $\mathcal{A}, \mathcal{B}$ win with probability at most $3\exp^{-\Omega(\rho^2 m)}$.*

Lemma 3. *There exist $\ell_0 \in \mathbb{N}$, $0 < \alpha < 1$ and $c > 0$ such that for every $\ell > \ell_0$ the following holds. Let G be an α-smooth projection game, $f : X \to \Sigma_A$ and $\rho = \ell^{-c}$. Let $g : Y \to \Sigma_B$ be a function such that $val(G, f, g) > 1 - 10\rho/\alpha$, then: If $A \in \mathrm{supp}_{6\rho/\alpha}(f)$ while $B \notin \mathrm{supp}_{40\rho/\alpha}(g)$. Then $\mathcal{A}, \mathcal{B}$ win with probability at most $4\exp^{-\Omega(\rho^2 m)}$.*

Lemma 4. *There exist $\ell_0 \in \mathbb{N}$, $0 < \alpha < 1$ and $c > 0$ such that for every $\ell > \ell_0$ the following holds. Let G be an α-smooth projection game, $g : Y \to \Sigma_B$ and $\rho = \ell^{-c}$. Let $f : X \to \Sigma_A$ be a function that maximizes $val(G, f, g)$, then: If $val(G, f, g) < 1 - 10\rho/\alpha$, and assuming $B \in \mathrm{supp}_\rho(g)$ and $A \in \mathrm{supp}_{6\rho/\alpha}(f)$ then $\mathcal{A}, \mathcal{B}$ win with probability at most $3\exp^{-\Omega(\rho^2 m)}$.*

The proofs of Theorem 3, Lemma 2, Lemma 3 and Lemma 4 can be found in the full version of the paper.

4.3 The Smoothness is Essential

In this section we show that the smoothness property is crucial. We show the existence of a game G that is not smooth, such that $G^{m,\ell}$ has perfect strategies $\mathcal{A}, \mathcal{B}$ and $\mathcal{A}$ is far from being a direct product strategy. Let us state Theorem 4 formally:

Theorem 4 (Formal Statement) *There exists a projection game G, such that for every ℓ and $0 < m < \ell$: There exist multiset oracles $\mathcal{A}, \mathcal{B}$ such that for every $f : X \to \Sigma_A$: $dist_{1/2}(A, f^\ell) > 1 - 2^{-\omega(\ell)}$. Yet, $\mathcal{A}, \mathcal{B}$ win $G^{m,\ell}$ with probability 1.*

The proof of Theorem 4 can be found in the full version of the paper.

References

[BSGH⁺06] Eli Ben-Sasson, Oded Goldreich, Prahladh Harsha, Madhu Sudan, and Salil P. Vadhan, *Robust pcps of proximity, shorter pcps, and applications to coding*, SIAM J. Comput. **36** (2006), no. 4, 889–974.

[DG08] Irit Dinur and Elazar Goldenberg, *Locally testing direct product in the low error range*, FOCS '08: Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science (Washington, DC, USA), IEEE Computer Society, 2008, pp. 613–622.

[DH09] Irit Dinur and Prahladh Harsha, *Composition of low-error 2-query pcps using decodable pcps*, FOCS, 2009, pp. 472–481.

[DM10] Irit Dinur and Or Meir, *Derandomized parallel repetition of structured pcps*, CoRR **abs/1002.1606** (2010).

[DR06] Irit Dinur and Omer Reingold, *Assignment testers: Towards a combinatorial proof of the pcp theorem*, SIAM J. Comput. **36** (2006), no. 4, 975–1024.

[FK94] Uriel Feige and Joe Kilian, *Two prover protocols: low error at affordable rates*, STOC, 1994, pp. 172–183.

[FK95] ———, *Impossibility results for recycling random bits in two-prover proof systems*, STOC, 1995, pp. 457–468.

[GS00] Oded Goldreich and Shmuel Safra, *A combinatorial consistency lemma with application to proving the pcp theorem*, SIAM J. Comput. **29** (2000), no. 4, 1132–1154.

[HK04] Jonas Holmerin and Subhash Khot, *A new pcp outer verifier with applications to homogeneous linear equations and max-bisection*, STOC, 2004, pp. 11–20.

[IKW09] Russell Impagliazzo, Valentine Kabanets, and Avi Wigderson, *New direct-product testers and 2-query pcps*, STOC '09: Proceedings of the 41st annual ACM symposium on Theory of computing (New York, NY, USA), ACM, 2009, pp. 131–140.

[Raz98] Ran Raz, *A parallel repetition theorem*, SIAM Journal on Computing **27** (1998), 763–803.