Parallel Repetition From Fortification

Parallel Repetition From Fortification

Dana Moshkovitz
Department of Electrical Engineering and Computer Science,
Massachusetts Institute of Technology
Cambridge, MA, USA
dmoshkov@csail.mit.edu

***Abstract*—The Parallel Repetition Theorem upper-bounds the value of a repeated (tensored) two prover game in terms of the value of the base game and the number of repetitions. In this work we give a simple transformation on games – "fortification" – and show that for fortified games, the value of the repeated game decreases perfectly exponentially with the number of repetitions, up to an arbitrarily small additive error. Our proof is combinatorial and short. As corollaries, we obtain: (1) Starting from a PCP Theorem with soundness error bounded away from 1, we get a PCP with arbitrarily small constant soundness error. In particular, starting with the combinatorial PCP of Dinur, we get a combinatorial PCP with low error. The latter can be used for hardness of approximation as in the work of Håstad. (2) Starting from the work of the author and Raz, we get a projection PCP theorem with the smallest soundness error known today. The theorem yields nearly a quadratic improvement in the size compared to previous work. We then discuss the problem of derandomizing parallel repetition, and the limitations of the fortification idea in this setting. We point out a connection between the problem of derandomizing parallel repetition and the problem of composition. This connection could shed light on the so-called Projection Games Conjecture, which asks for projection PCP with minimal error.**

***Keywords*-parallel repetition; PCP; hardness of approximation; projection game; fortification;**

I. Introduction

A. The Parallel Repetition Theorem

In a two prover game $\mathcal{G}$, a verifier picks at random a pair of questions (x, y) from a specified set of possible questions, sends x to the first prover, and sends y to the second prover; the first prover replies with an answer a, and the second prover replies with an answer b; the verifier, knowing x and y, and having inspected both a and b, decides whether to accept or reject. The value the prover strategies achieve is the probability that the verifier accepts. The *value* of $\mathcal{G}$, denoted $val(\mathcal{G})$, is the maximum of this quantity over all prover strategies.

A k-repetition (tensor) of a game $\mathcal{G}$ is the game $\mathcal{G}^{\otimes k}$, in which the verifier picks at random k question pairs $(x_1, y_1), \ldots, (x_1, y_k)$; sends one prover $x_1, \ldots, x_k$, and sends the other prover $y_1, \ldots, y_k$; the first prover replies with $a_1, \ldots, a_k$, and the second prover replies with $b_1, \ldots, b_k$; the verifier checks that it would have accepted in all k tests.

A long line of work analyzes how $val(\mathcal{G}^{\otimes k})$ depends on $val(\mathcal{G})$ and k. Clearly, $val(\mathcal{G}^{\otimes k}) \geq val(\mathcal{G})^k$, since the provers can follow the same strategy in each one of the k rounds. One might guess that $val(\mathcal{G}^{\otimes k}) = val(\mathcal{G})^k$, but this turns out to be false [17], [13], [16]. In a breakthrough result, Raz [26] showed that $val(\mathcal{G}^{\otimes k})$ does exhibit an exponential decay with k when $val(\mathcal{G}) < 1$ (below, Σ_X is the set of possible answers a of the first prover, while Σ_Y is the set of possible answers b of the second prover):

Theorem 1 (Raz's Parallel Repetition Theorem [26]): There exists $W : [0,1] \to [0,1]$ such that $W(x) < 1$ for $x < 1$, and

$$val(\mathcal{G}^{\otimes k}) \leq (W(val(\mathcal{G})))^{k/\log(|\Sigma_X||\Sigma_Y|)}.$$

Interestingly, the dependence of the exponent in $|\Sigma_X|$ and $|\Sigma_Y|$ is inherent [16]. Disappointingly, the base of the exponent is quite far from $val(\mathcal{G})$. In fact, in Raz's theorem, $W(val(\mathcal{G}))$ is close to 1 even when $val(\mathcal{G})$ is close to 0! Many works simplified and improved the parameters of the Parallel Repetition Theorem for general games [19], as well as for games with a special structure, most notably *projection games* [25], [12] and *expanding projection games* [28], [12]. Before we describe the main results of those papers, let us discuss projection games and their importance.

Arguably, the most important application of the Parallel Repetition Theorem is soundness amplification for projection games. In this paper it will be convenient for us to consider the following definition of a projection game:

Definition 1.1 (Projection game): A projection game is defined by a bipartite graph $G = (X, Y, E)$, alphabets Σ_X and Σ_Y and functions $\{\pi_e : \Sigma_X \to \Sigma_Y\}_{e \in E}$, called "projections". In the game, the verifier picks uniformly at random $y \in Y$, and two edges $e = (x, y), e' = (x', y) \in E$, sends x to the first prover, and sends x' to the second prover; the first prover replies with $a \in \Sigma_X$, and the second prover replies with $a' \in \Sigma_X$; the verifier accepts if $\pi_e(a) = \pi_{e'}(a')$.

Remark 1.1: The more standard definition of projection games is as follows: the verifier picks uniformly at random an edge $(x, y) \in E$, sends x to the first prover, and sends y to the second prover; the first prover replies with $a \in \Sigma_X$, and the second prover replies with $b \in \Sigma_Y$; the verifier accepts if $\pi_e(a) = b$. Definition 1.1 is a symmetric version of this definition, and as useful to hardness of approximation

(or more). If $\mathcal{G}'$ is the game in Definition 1.1, and $\mathcal{G}$ is the game we defined here, then $val(\mathcal{G})^2 \leq val(\mathcal{G}') \leq val(\mathcal{G})$ (the first inequality follows from convexity, while the second inequality follows from a probabilistic assignment).

The PCP Theorem, in the form that is most useful for hardness of approximation, states that it is NP-hard, given a projection game $\mathcal{G}$, to distinguish the case where $val(\mathcal{G}) = 1$ from the case where $val(\mathcal{G}) \leq \varepsilon$. The parameter ε is called the *soundness error* of the PCP. Since parallel repetition of a projection game is itself a projection game, the Parallel Repetition Theorem, when applied on the basic PCP Theorem [5], [4], [3], [2], yields a projection PCP theorem with arbitrarily small soundness error. Projection PCP with low soundness error is the basis of most of the best NP-hardness of approximation results we have today. In particular, it is the basis of the hardness results in Håstad's seminal paper [18].

For projection games, the size of the game – denoted $size(\mathcal{G})$ – is the size of the graph G. An unfortunate aspect of parallel repetition is that it raises the size of the game to the k-th power. In particular, if k is super-constant, one gets a super-polynomial reduction from SAT to the repeated game, rather than an NP-hardness result. When one assumes that solving SAT on inputs of size n requires time $exp(n)$ ("The Exponential Time Hypothesis"), the reductions obtained using parallel repetition only yield time lower bounds of the form $exp(N^{1/k})$ for input size N. Due to this state of affairs, parallel repetition is used mostly for constant k.

One of the most important open problems in approximability is to construct projection games with error that is inverse polynomial in the size of the game. The author named this problem "The Projection Games Conjecture" in [21]:

Conjecture 1.1 (Projection Games Conjecture): There exists $\alpha > 0$, such that for every N and $\varepsilon \geq 1/N^{\alpha}$, it is NP-hard to distinguish, given a projection game $\mathcal{G}$ of size and alphabet size N, between the case that $val(\mathcal{G}) = 1$ and the case that $val(\mathcal{G}) \leq \varepsilon$.

One of the most notable applications of this conjecture is an NP-hardness result for approximating CLOSEST-VECTOR-PROBLEM in lattices to within polynomial factors (see [21] for a discussion of more applications). The lowest soundness error known today is $\varepsilon = 1/(\log N)^c$, for any constant $c > 0$, when N is the size of the game [12]. This is by a reduction of the author and Raz [23] from SAT on input of size n to projection games of size $N = n^{1+o(1)}$, where the soundness error is $\varepsilon = 1/(\log N)^{\beta}$ for some $\beta > 0$. When this game is repeated in parallel, the soundness error can be $\varepsilon = 1/(\log N)^c$ for any constant $c > 0$, while the size is raised to $O(c/\beta)$.

Understanding the significance of projection games, we now turn to review what is known about their repetition. Interestingly, in the projection case $val(\mathcal{G}^{\otimes k})$ does not depend on the number of possible answers of the provers [25]. The state of the art results are as follows:

Theorem 2 (Parallel repetition of projection games [25], [12]): For any projection game $\mathcal{G}$ as in Remark[1] 1.1,

1) [25] If $val(\mathcal{G}) = 1 - \epsilon$, then
$$val(\mathcal{G}^{\otimes k}) \leq (1 - \epsilon/2)^{\Omega(\epsilon k)}.$$
2) [12] If $val(\mathcal{G}) = 1 - \epsilon$ and $\epsilon \ll 1/\sqrt{k}$, then
$$val(\mathcal{G}^{\otimes k}) \leq 1 - \Omega(\sqrt{k} \cdot \epsilon).$$
3) [12] $val(\mathcal{G}^{\otimes k}) \leq 2^k \cdot val(\mathcal{G})^{k/2}$.

The first result is best when $val(\mathcal{G})$ is a constant close to 1; the second result is best for $val(\mathcal{G})$ very close to 1; while the last result is best for the case of small $val(\mathcal{G})$ (note that in the first result the base of the exponent is about $\frac{1}{2}$ rather than $val(\mathcal{G})$ when $val(\mathcal{G})$ is very small). The second result is tight when it applies, as Raz [27] showed a unique game $\mathcal{G}$ with $val(\mathcal{G}) = 1 - \epsilon$ for which $val(\mathcal{G}^{\otimes k}) \geq 1 - O(\sqrt{k} \cdot \epsilon)$. More generally, Barak et al [6] analyze the behavior of general unique games under parallel repetition.

For projection games on expanders, Dinur and Steurer's proof is somewhat simpler than its general case [12]. More than that, Raz and Rosen [28] prove a stronger result in the expander case: if $val(\mathcal{G}) = 1 - \epsilon$ for $\epsilon < 1/2$, then $val(\mathcal{G}^{\otimes k}) \leq (1 - \epsilon)^{\Omega(k)}$.

We note that in all the aforementioned results, either explicitly or hiding in $\Omega(\cdot)$, is the fact that *not all repetitions count*. That is, in many of the k repetitions, the provers may win with probability 1 conditioned on winning other rounds. This phenomenon is known to actually occur – there are unique games with $val(\mathcal{G}^{\otimes 2}) = val(\mathcal{G})$ [13].

B. Our Contribution

Instead of exploring the subtle behavior of general projection games under repetition, in this work we engineer the games so they behave well under repetition. We present a simple combinatorial transformation on projection games, which we call "fortification". Fortification endows the game with extractor structure and ensures that certain sub-games of the game have (approximately) the same value as the global game. Fortification preserves a projection structure, while increasing $|X|$ and $|\Sigma_X|$ in a controlled way. We show that for fortified projection games $\mathcal{G}$, the value of the k-repeated game is, approximately, $val(\mathcal{G})^k$, i.e.,

$$val(\mathcal{G})^k \leq val(\mathcal{G}^{\otimes k}) \leq val(\mathcal{G})^k + err,$$

where the small additive error err can be made arbitrarily small by fortification.

[1]Using the relation between the projection games of Definition 1.1 and the projection games of Remark 1.1 (explained in Remark 1.1), Theorem 2 yields a (weaker) parallel repetition theorem for projection games as in Definition 1.1. It is quite possible that the techniques of Dinur and Steurer yield bounds as in Theorem 2 for games as in Definition 1.1 too.

In the fortified game, rather than sending the first prover a question x and the second prover a question x', the verifier sends the first prover a set of correlated questions $\{x_1, \ldots, x_t\} \ni x$, and it sends the second prover a set of correlated questions $\{x'_1, \ldots, x'_t\} \ni x'$. The provers are asked to provide answers for all t questions they got. The verifier then uses their answers to perform the test involving x and x' (note that other questions among the $2t$ typically induce no tests). The choice of the correlated questions is done using an extractor or a random walk on an expander, in a manner that was inspired by ideas in combinatorial construction of error correcting codes.

Notably, our analysis of parallel repetition is much simpler than all existing analyses. Unlike Raz's proof, our analysis does not require information theory, or clever choices of sub-games à la Razborov, nor does it require a heavy use of linear algebra and Cheeger's inequality as in the recent analysis of Dinur and Steurer for projection games.

As corollaries, we obtain:

1) Starting from a PCP Theorem with soundness error bounded away from 1 [5], [4], [3], [2], we get a PCP with arbitrarily small constant soundness error. In particular, starting with the combinatorial PCP of Dinur, we get a combinatorial PCP with low error whose analysis is combinatorial. The latter can be used for hardness of approximation as in the work of Håstad.
2) Starting from the work of the author and Raz [23], which gives a projection PCP theorem with error $1/(\log n)^{\beta}$ for *some* constant $\beta > 0$, we get a projection PCP theorem with error $1/(\log n)^{c}$ for *any* constant $c \geq 1$ (which is the lowest known today [12]). Our theorem yields nearly a quadratic improvement in the size for a given c compared to [12].

Our proof evolved from a previous work of the author [22] about soundness amplification for low degree testing. As happened several times in the past in PCP, we could transform some of the ideas from the algebraic analysis into a purely combinatorial setting.

Finally, we explore the possibility of obtaining stronger projection PCP theorems using our ideas. The bottleneck here is the large size blow-up introduced by parallel repetition, and hence the question is whether parallel repetition could be "derandomized" for appropriately fortified games. That is, whether the verifier can pick all k tests in a randomness-efficient way. While we do not know how to extend our fortification ideas to this case (we explain the difficulty in Section VI), we are able to point out an intriguing connection between the problem of derandomizing parallel repetition and the well-studied problem of *composition* of two prover games. The connection – which holds for general two prover games – is that both problems share a combinatorial hard core. Since repetition and composition constitute the two existing approaches to the Projection Games Conjecture (error reduction and alphabet reduction, respectively), the connection sheds light on the difficulty of proving the conjecture.

C. Previous Work on Combinatorial Analysis of Parallel Repetition

Feige and Kilian [15], as well as Impagliazzo, Kabanets and Wigderson [20] already gave combinatorial analyses of parallel repetition. Crucially, those parallel repetition theorems were *weaker* than what was known via other techniques, while our theorem is *stronger* than what is known via other techniques. As in the current paper, Feige and Kilian, as well as Impagliazzo, Kabanets and Wigderson, first apply a transformation on the game, and then repeat the game in parallel. The transformation differs from our fortification, and is (up to variants) as follows: The verifier picks uniformly at random either (i) "compare": edges with a common endpoint $e = (x, y), e' = (x', y) \in E$; or (ii) "confuse": independent edges $e = (x, y), e' = (x', y') \in E$. One prover is sent x and the other prover is sent x'. The provers reply $a, a' \in \Sigma_X$, respectively; In case the two edges have a common endpoint y, the verifier checks that $\pi_e(a) = \pi_{e'}(a')$. The intuition of this transformation is that in some of the rounds the provers are compared, hence for the verifier to accept with good probability, the provers are forced to a consistent strategy. The confuse rounds ensure that the consistent strategy is pervasive.

Feige and Kilian [15] show that for games $\mathcal{G}$ transformed as we described, for any $\delta > 0$ such that k is a sufficiently large polynomial in $1/\delta$ and $1/(1 - val(\mathcal{G}))$, it holds that $val(\mathcal{G}^{\otimes k}) \leq \delta$. In this theorem, the decay of the value of the game with repetition is polynomial in k, rather than exponential in k. Impagliazzo, Kabanets and Wigderson [20] prove that $val(\mathcal{G}^{\otimes k}) \leq 2^{-\Omega(\sqrt{k}/(1-val(\mathcal{G})))}$. Here the decay is exponential in $\sqrt{k}$ instead of in k.

D. Previous Work on Derandomizing Parallel Repetition

For simplicity, let us continue to denote the repeated game $\mathcal{G}^{\otimes k}$, with the understanding that the k questions to the provers may be correlated. Feige and Kilian [14] showed that in the derandomized case, for $val(\mathcal{G}^{\otimes k}) \leq \delta$, it must be the case that the degrees in $\mathcal{G}$'s graph are at least $\approx 1/\delta$ (under an assumption on $\mathcal{G}$ they call *softness*, which indeed holds in the cases of interest). The degrees in the graph correspond to the uncertainty each prover has with respect to the questions of the other prover. For any two prover game in which each of the verifier's tests is satisfiable, if the graph is bi-regular, and one of the sides has degree D, then the value of the game is at least $1/D$. The interesting feature of Feige and Kilian's result is that they relate the value of $\mathcal{G}^{\otimes k}$ to the degree in $\mathcal{G}$. Taking this restriction into account, one might hope for a derandomized parallel repetition where $val(\mathcal{G}^{\otimes k}) \leq \delta$ and $size(\mathcal{G}^{\otimes k}) = size(\mathcal{G}) \cdot (1/\delta)^{O(k)}$. If such a derandomization had been available, it would have given

projection PCP with soundness error $\delta = 2^{-(\log n)^{\beta}}$ for some constant $\beta > 0$.

However, so far there has been little progress even on suggesting candidate games $\mathcal{G}$ with a derandomization $\mathcal{G}^{\otimes k}$. The two exceptions have been *free games* [31] and *linear games* [11]. A free game is a game in which the questions of the two provers are independent. A linear game is a game in which the questions correspond to points in a linear space, and the verifier's tests correspond to linear sub-spaces. For free games, Shaltiel [31] analyzed repetition where the dependence between the randomness the verifier needs in order to reach a given target $val(\mathcal{G}^{\otimes k}) \leq \delta$, and the number of possible answers of the provers, is improved (recall that for general two prover games $val(\mathcal{G}^{\otimes k})$ depends on the number of possible answers of the provers). The size of the game in Shaltiel's theorem is still $(size(\mathcal{G}))^{\Omega(\log(1/\delta))}$. For linear games, Dinur and Meir [11] analyzed derandomized repetition, but where the soundness error does not decrease exponentially. For both types of games, known transformations from general games incur a large blow-up in the size (for free games [1]) or in the soundness error (for linear games [11]). In fact, for free games it was proved that the size blow-up is inherent [1]. Hence, neither free games nor linear games seem useful for making further progress toward the Projection Games Conjecture.

II. Preliminaries

Let $Dist$ be a distribution over a space X. The *entropy* in the distribution, denoted $H(Dist)$, is $\sum_{x \in X} Dist(x) \log(1/Dist(x))$. We say that $Dist$ has *min-entropy* at least k, and denote $H_{\infty}(Dist) \geq k$, if no $x \in X$ has probability higher than 2^{-k}. If a distribution is uniform over a set $S \subseteq X$ (a "flat" distribution; we'll also refer to S as an event), then it has min-entropy $\log |S|$. It is known that any distribution with min-entropy k can be viewed as a convex combination $\{S_i\}_i$ where for every i, it holds that $\Pr[S_i] = |S_i| / |X| \geq 2^{-k}$.

A (δ, ε)-extractor is a bi-regular bipartite graph $H = (X, Y, E)$, such that for every distribution $Dist$ over X with min-entropy at least $\log(\delta |X|)$, the distribution on Y obtained by picking x according to $Dist$ and picking a uniformly random neighbor $y \in Y$ of x is ε-close to uniform over Y in statistical distance.

The following extractor construction follows from expander random walk:

Lemma 2.1 (Extractor construction [30]): For any $\delta, \varepsilon > 0$, there exist (δ, ε)-extractors $G = (X, Y, E)$ such that $|X| = O(|Y| / \delta)$ and each vertex in X has degree $D = O(\log(1/\delta) \cdot (1/\varepsilon)^2)$. Moreover, there exist explicit constructions achieving $|X| = O(|Y| / \delta)$ and $D = D(\delta, \varepsilon) = \exp(poly \log\log(1/\delta)) \cdot (1/\varepsilon)^2$.

A two prover game $\mathcal{G}$ is defined by a set X of questions to the first prover, a set Y of questions to the second prover, an alphabet Σ_X for the answers of the first prover, and an alphabet Σ_Y for the answers of the second prover. In addition, there is a distribution μ over question pairs $X \times Y$, and a predicate $V \subseteq X \times Y \times \Sigma_X \times \Sigma_Y$. The verifier picks (x, y) from μ sends x to the first prover and y to the second prover; receives $a \in \Sigma_X$ from the first prover and $b \in \Sigma_Y$ from the second prover; then accepts or reject based on $V(x, y, a, b)$.

One often considers the bipartite graph associated with $\mathcal{G}$. This is the graph $G = (X, Y, E)$ on vertex sets X and Y, where the edges are the question pairs (x, y) with non-zero probability in μ. When one refers to degrees in $\mathcal{G}$, the intention is degrees in G. Typically, and by default in this paper, μ is uniform over E, and a question pair (x, y) from μ is such that x is uniform over X, while y is uniform over Y. The value achieved by certain prover strategies is the probability that the verifier accepts. The value of $\mathcal{G}$, denoted $val(\mathcal{G})$, is the maximum of this quantity over all prover strategies. The size of $\mathcal{G}$, denoted $size(\mathcal{G})$, is $|X| + |Y| + |E|$. The randomness of the verifier is $\log |E|$.

III. Fortification

If $\mathcal{G}'$ is a sub-game of a game $\mathcal{G}$ obtained by picking only a subset of the possible question pairs of the verifier, then the value of $\mathcal{G}'$ can be much higher than the value of the original game $\mathcal{G}$. Fortified games $\mathcal{G}$ are such that certain large sub-games $\mathcal{G}'$ of $\mathcal{G}$ have $val(\mathcal{G}') \approx val(\mathcal{G})$. The largeness is with respect to the upper bound $val(\mathcal{G}^{\otimes k}) \leq \delta$ we wish to obtain. Note that the requirement that *every* sub-game $\mathcal{G}'$ of fraction δ in $\mathcal{G}$ has $val(\mathcal{G}') < 1$ is equivalent to saying that $val(\mathcal{G}) < \delta$. Hence, it is important that we focus on a family of large sub-games, rather than on all large sub-games.

Specifically, we focus on *rectangular* sub-games, defined as follows: If S is an event depending on the question to the first prover, and T is an event depending on the question to the second prover, then the rectangular sub-game $\mathcal{G}_{|S \times T}$ is the game $\mathcal{G}$ conditioned on the questions to the provers satisfying S and T, respectively. That is, the verifier picks at random $y \in Y$, $x, x' \in X$ such that $(x, y), (x', y) \in E$, conditioned on $x \in S$, $x' \in T$. It then performs the test as before. We say that the rectangular game is δ-large if $\Pr[S], \Pr[T] \geq \delta$.

We further extend the definition to convex combinations over events $\{S_i \times T_i\}_i$. Here we first pick $S_i \times T_i$ from the combination, then consider the relevant sub-game. The value of the prover strategies in the game is the convex combination of the values of the prover strategies in the sub-games. We say that such a sub-game is δ-large if for all i, we have $\Pr[S_i], \Pr[T_i] \geq \delta$.

We define the fortified value of a game as follows:

Definition 3.1: The *δ-fortified value* of a game $\mathcal{G}$, denoted $val_{\delta}(\mathcal{G})$, is the maximum of $val(\mathcal{G}_{|\{S_i \times T_i\}_i})$ over all δ-large convex combinations $\{S_i \times T_i\}_i$.

We say that a projection game $\mathcal{G}$ on a (δ, ε)-extractor is (δ, ε)-*fortified* if δ-large rectangular sub-games have value at

most $val(\mathcal{G}) + \varepsilon$. We show that every projection game can be fortified easily. Fortification does not increase the size or the alphabets of the game too much. Fortification does not change the value of the game, only makes sure that the value of large rectangular sub-games is similar to the value of the overall game.

Our fortification lemma assumes that the bipartite graph underlying the projection game is bi-regular. Projection games on general graphs can be transformed to bi-regular using transformations of [23]. The first transformation regularizes the Y side, so each Y vertex has a small degree:

Lemma 3.1 (Y-degree reduction [23]): For any $\eta > 0$, any projection game $\mathcal{G}$ can be efficiently transformed to a new projection game $\mathcal{G}'$ on a graph (X, Y, E) that is Y-regular with degree $poly(1/\eta)$, where $size(\mathcal{G}') \leq size(\mathcal{G}) \cdot poly(1/\eta)$ and $val(\mathcal{G}') \leq val(\mathcal{G}) + \eta$ (the alphabets are unchanged).

The second transformation switches between the Y and the X side. The idea is that each assignment to a vertex $y \in Y$ now contains assignments to all the neighbors of y, such that the assignments to the neighbors agree on their projection to y:

Lemma 3.2 (Switching sides [23]): Any projection game $\mathcal{G}$ on a graph $G = (X, Y, E)$ with Y-degree D and alphabets Σ_X, Σ_Y can be transformed into a projection game $\mathcal{G}'$ on a graph $G' = (Y, X, E)$ and alphabets Σ_X^D, Σ_X, where $val(\mathcal{G}') = val(\mathcal{G})$.

By applying Y-degree reduction, switching sides, and Y-degree reduction again, we obtain a projection game on a bi-regular graph that has approximately the same value as the original game.

Having gotten bi-regularity out of the way, let us describe the fortification transformation:

Lemma 3.3 (Fortification): For any $\varepsilon, \delta > 0$, a projection game $\mathcal{G}$ on a bi-regular graph $G = (X, Y, E)$, with alphabets Σ_X, Σ_Y, and projections $\{\pi_e\}_{e \in E}$ can be efficiently converted to a game $\mathcal{G}^*$ on a graph $G^* = (X^*, Y, E^*)$ with alphabets Σ_X^D, Σ_Y, and projections $\{\pi_e^*\}_{e \in E^*}$, such that

1) G^* is a (δ, ε)-extractor.
2) $D = D(\delta, \varepsilon)$ as in Lemma 2.1.
3) The size of G^* is linear in the size of G, $1/\delta$, $poly(1/\varepsilon)$.
4) $val(\mathcal{G}^*) = val(\mathcal{G})$.
5) $val_\delta(\mathcal{G}^*) \leq val(\mathcal{G}) + 2\varepsilon$.

Proof: Let $H = (X^*, X, E_H)$ be a (δ, ε)-extractor. By Lemma 2.1, such can be constructed so $|X^*| = poly(|X|, 1/\delta)$ and each vertex in X^* has $D = D(\delta, \varepsilon)$ neighbors in X. Let E^* contain an edge $e^* = (x^*, y)$ for every pair $(x^*, x) \in E_H$ and $e = (x, y) \in E$. An assignment $\vec{a}$ to x^* consists of assignments to all D neighbors of x^* in H, and in particular some $a(x)$ to x. The projection on the edge e^* is $\pi_{e^*}^*(\vec{a}) = \pi_e(a(x))$. Note that G^* is a (δ, ε)-extractor, and that $size(\mathcal{G}^*)$ is $O(size(\mathcal{G}) poly(1/\varepsilon)/\delta)$. Consider the game $\mathcal{G}^*$ associated with the graph G^*, alphabets Σ_X^D, Σ_Y and projections $\{\pi_{e^*}^*\}_{e^* \in E^*}$. In this game, the verifier picks uniformly at random $y \in Y$ and $x^*, (x^*)' \in X^*$ such that $e^* = (x^*, y) \in E^*$ and $(e^*)' = ((x^*)', y) \in E^*$. Upon receipt of answers $\vec{a}, (\vec{a})' \in \Sigma_X^D$, the verifier checks that $\pi_{e^*}^*(\vec{a}) = \pi_{(e^*)'}^*((\vec{a})')$.

We have $val(\mathcal{G}) \leq val(\mathcal{G}^*)$, since any strategy $a : X \to \Sigma_X$ for $\mathcal{G}$ induces a strategy for $\mathcal{G}^*$ achieving the same value: given $x^* \in X^*$, the answer $\vec{a}$ of the prover assigns every neighbor $x \in X$ of x^* in H the answer $a(x)$. Moreover, $val(\mathcal{G}^*) \leq val(\mathcal{G})$, since every strategy in $\mathcal{G}^*$ induces a randomized strategy in $\mathcal{G}$ achieving the same value in expectation (and hence there exists a strategy for $\mathcal{G}$ achieving this value): given $x \in X$, the prover picks at random a neighbor $x^* \in X^*$ of x in H, and responds according to the strategy for x^*.

Let $\{S_i \times T_i\}_i$ be a convex combination of events, where for all i, the event S_i depends only on x^*, the event T_i depends only on $(x^*)'$ and $\Pr[S_i], \Pr[T_i] \geq \delta$. We'd like to prove that $val(\mathcal{G}^*)_{|\{S_i \times T_i\}_i} \leq val(\mathcal{G}) + 2\varepsilon$. Select at random i, and $y \in Y$, $x^*, (x^*)' \in X^*$, conditioned on the events S_i and T_i. Let $x, x' \in X$ be the vertices for which $(x, y), (x', y) \in E$, while $(x^*, x), ((x^*)', x') \in E_H$. By the extractor property of H, the vertices x and x' are each ε-close to uniform over X. The claim that $val_\delta(\mathcal{G}^*) \leq val(\mathcal{G}) + 2\varepsilon$ follows from the definition of $\mathcal{G}$ and $\mathcal{G}^*$. ∎

Fortification preserves projection, but does not preserve uniqueness. Indeed, due to the works [26], [6], we do not expect to prove a strong parallel repetition for unique games.

We wish to emphasize that *not every projection game on extractors is fortified*. Indeed, if we take any projection game on extractors and change the projections on edges touching δ fraction of the vertices so they are trivially satisfied, we hardly change the value of the game, but we make sure that the game is not fortified.

Fortification increases the size by a factor $O(1/\delta)$, where we fortify against sub-games of fraction δ. When repeating the game for k rounds, the size increases by a factor $\approx (1/\delta)^k$. However, due to fortification, $val(\mathcal{G}^{\otimes k})$ decreases exponentially with k, rather than with $k/2$. Hence, to reach a target $val(\mathcal{G}^{\otimes k}) \leq \delta$ previous methods required twice as many rounds k as we do, and thus the right comparison is between size $\gg (size(\mathcal{G}))^{2k}$ for previous methods and size $\approx (size(\mathcal{G})/\delta)^k$ for us. Since typically $size(\mathcal{G})$ is much larger than $1/\delta$, our method yields better size than before.

Fortification also raises the size of the alphabet Σ_X to a power $D = D(\delta, \varepsilon)$. This price is quite tolerable since in order to reach a target $val(\mathcal{G}^{\otimes k}) \leq \delta$, we take $k = \Theta(\log(1/\delta))$, and in repetition, Σ_X is raised to a power k anyway ($1/\varepsilon$ is typically smaller than, or comparable to, $\log(1/\delta)$). Moreover, there is a hope that the large alphabet due to fortification could be re-used for the repeated tests.

IV. A Parallel Repetition Theorem

In this work we suggest to prove parallel repetition theorems assuming that the underlying game is fortified:

Theorem 3 (Parallel repetition): For any $\varepsilon, \delta > 0$, if $\mathcal{G}$ is a projection game on a (δ, ε)-extractor where $\delta \leq \varepsilon^4 \left|\Sigma_Y\right|^{-(k-1)}$, then there exists $err = O(k\varepsilon)$ for which

$$val(\mathcal{G}^{\otimes k}) \leq (val_{\delta/\varepsilon^3}(\mathcal{G}) + err)^k.$$

Theorem 3 relates the value of the repeated game $\mathcal{G}^{\otimes k}$ to the fortified value of the original game. If $\mathcal{G}$ is fortified, then its fortified value is approximately $val(\mathcal{G})$. The error term is $err = O(k\varepsilon)$, which means that we should pick $\varepsilon \ll 1/k$. We also need $\delta \approx \left|\Sigma_Y\right|^{-k}$. Existing constructions of projection games $\mathcal{G}$ with $val(\mathcal{G}) \leq \varepsilon$ have $|\Sigma_Y| = poly(1/\varepsilon)$ [23]. Moreover, there is a simple transformation of Dinur and Harsha [10] based on code concatenation that decreases Σ_Y to $poly(1/\varepsilon)$ for any projection game:

Lemma 4.1 (Σ_Y reduction [10]): For any $\eta > 0$, any projection game $\mathcal{G}$ with $val(\mathcal{G}) \leq \varepsilon$ and alphabet Σ_Y can be efficiently transformed to a new projection game $\mathcal{G}'$ where $size(\mathcal{G}') \leq size(\mathcal{G}) \cdot \log|\Sigma_Y| \cdot poly(1/\eta)$, $val(\mathcal{G}') \leq \varepsilon + O(\eta)$ and the new $|\Sigma_Y|$ is $poly(1/\eta)$.

The proof of Theorem 3 is in Section V. Using fortification as in Lemma 3.3 and our parallel repetition theorem in Theorem 3, we obtain a new combinatorial soundness amplification technique for two prover games. We demonstrate the usage of this technique for two purposes: the first is a purely combinatorial projection PCP with low error which suffices for Håstad's hardness of approximation results [18]; the second is a projection PCP with the lowest error known today. The second is stronger than the first, and has applications to hardness of approximation beyond those of the first (e.g., for tight NP-hardness of approximating Set-Cover [21], [12]). For the second, we achieve nearly a quadratic improvement in the size compared to previous work [12].

Our first corollary is a projection PCP Theorem with arbitrarily small constant error, as was known by applying Raz's analysis [26] on the PCP Theorem [3], [2], but with a combinatorial proof from beginning to end. The proof relies on Dinur's combinatorial PCP theorem [9], and the fact that any two prover game $\mathcal{G}$ can be transformed into a projection game $\mathcal{G}'$, such that $val(\mathcal{G}) = 1 \Rightarrow val(\mathcal{G}') = 1$, while $val(\mathcal{G}) < 1 \Rightarrow val(\mathcal{G}') < 1$.

Corollary 4.2 (Combinatorial PCP with low error): For any $\alpha > 0$, it is NP-hard to distinguish, given a projection game $\mathcal{G}$, between the case where $val(\mathcal{G}) = 1$ and the case where $val(\mathcal{G}) \leq \alpha$.

Our second corollary is a projection PCP Theorem with arbitrarily small poly-logarithmic error. This result was previously obtained by applying Dinur and Steurer's parallel repetition theorem [12] on the low error projection games of the author and Raz [23], however here we get nearly a quadratic improvement in the size compared to the Dinur-Steurer result:

Corollary 4.3 (Sub-constant error projection PCP): There exists $\beta' > 0$, such that for any constant $c \geq 1$, there is a reduction from Sat of size n to a projection game $\mathcal{G}$ of size $n^{(1+o(1))c/\beta'}$, such that satisfiable instances are mapped to $\mathcal{G}$ with $val(\mathcal{G}) = 1$, while unsatisfiable instances are mapped to $\mathcal{G}$ with $val(\mathcal{G}) \leq O(1/(\log n)^c)$.

V. Proof of Parallel Repetition Theorem

Set $\hat{\varepsilon} = val_{\delta/\varepsilon^3}(\mathcal{G}) + ck\varepsilon$ where c is a sufficiently large constant. We assume a strategy of the provers in $\mathcal{G}^{\otimes k}$ that achieves value larger than $\hat{\varepsilon}^k$, and wish to arrive at a contradiction. If for each round $i = 1, \ldots, k$ the provers fix strategies $a_i : X \to \Sigma_X$ and $a'_i : X \to \Sigma_X$ that depend only on the questions of the i-th round, then the value they achieve is at most $val(\mathcal{G})^k$ by definition. However, the provers may answer the questions in each round based also on the questions to the other rounds. For example, suppose that the provers win in the first round iff there is a certain relation between their questions in the second round. Then, a-priori, it is possible that they win the first round with probability $val(\mathcal{G})$, and conditioned on winning the first round, win the second round with probability much larger than $val(\mathcal{G})$, since the second round is effectively played in a sub-game of the base game that potentially can be won with higher probability. We show that thanks to fortification this cannot happen.

We first identify a list $I \subset \{1, \ldots, k\}$ of *influential rounds*, such that conditioned on winning them, the provers win any other round with probability larger than $\hat{\varepsilon}$. The intuition is that these are the rounds where the provers try to make gains that will help them win other rounds better than expected. Note that the list cannot contain all rounds – as otherwise the total probability of winning all k rounds would have been too small. For $i \in \{1, \ldots, k\}$, Let W_i be the event that the provers win the i-th round. For $I \subseteq \{1, \ldots, k\}$, let W_I be the event that the provers win all the rounds in I.

Lemma 5.1 (Influential rounds): There exists $I \subseteq \{1, \ldots, k\}$, $l \doteq |I| < k$, such that for every $i \in \{1, \ldots, k\} - I$, it holds that $\Pr[W_i | W_I] > \hat{\varepsilon}$.

Proof: Construct I as follows: start with $I = \phi$, and while there is still $i \in \{1, \ldots, k\}$ such that $\Pr[W_i | W_I] \leq \hat{\varepsilon}$, add i to I.

By construction, for every $i \in \{1, \ldots, k\} - I$, it holds that $\Pr[W_i | W_I] > \hat{\varepsilon}$. We claim that at each step $\Pr[W_I] \leq \hat{\varepsilon}^{|I|}$. This is certainly true when $|I| = 0$. Moreover, if it is true for I, it continues to be true if we decide to insert i to I, as

$$\Pr\left[W_{I \cup \{i\}}\right] = \Pr[W_i | W_I] \cdot \Pr[W_I] \leq \hat{\varepsilon} \cdot \hat{\varepsilon}^{|I|} = \hat{\varepsilon}^{|I \cup \{i\}|}.$$

Since $\hat{\varepsilon}^{|I|} \geq \Pr[W_I] \geq \Pr[W_{1..k}] > \hat{\varepsilon}^k$, necessarily $|I| < k$. ∎

Let $W = W_I$ be the event that the provers win all l influential rounds. That is, if the verifier picks

edges $e_1,\ldots,e_k,e'_1,\ldots,e'_k \in E$, and the provers reply $a_1,\ldots,a_k,a'_1,\ldots,a'_k \in \Sigma_X$, then $\pi_{e_j}(a_j) = \pi_{e'_j}(a'_j)$ for all $j \in I$. Note that this event may depend on the questions in all rounds, and not just on the questions in rounds $j \in I$. By Lemma 5.1, conditioned on W, the provers win each of the other rounds with probability larger than $\hat{\varepsilon}$. We will argue that this cannot happen.

Consider a fixing of the questions to the provers in the influential rounds,

$$\{y_j\}_{j\in I} \subseteq Y,\ \{x_j\}_{j\in I} \subseteq X,\ \{x'_j\}_{j\in I} \subseteq X.$$

Let $(\mathcal{G}^{\otimes k})'$ be the sub-game associated with this fixing. Let W' be the event of winning all l influential rounds in $(\mathcal{G}^{\otimes k})'$. We further partition $(\mathcal{G}^{\otimes k})'$ into sub-games: There is a sub-game per choice of l labels from Σ_Y for $\{y_j\}_{j\in I}$ of the influential rounds. If the labels are denoted $\{\sigma_j\}_{j\in I} \subseteq \Sigma_Y$, then $S_{\vec{\sigma}} \subseteq X^k$ contains all the questions $(x_1,\ldots,x_k)$ to the first prover such that the answers of the first prover $a_1,\ldots,a_k$ agree with the choice, i.e., for all $j \in I$, we have $\pi_{e_j}(a_j) = \sigma_j$. Similarly, $T_{\vec{\sigma}} \subseteq X^k$ contains all questions $(x'_1,\ldots,x'_k)$ to the second prover such that the answers of the second prover $a'_1,\ldots,a'_k$ agree with the choice, i.e., for all $j \in I$, we have $\pi_{e'_j}(a'_j) = \sigma_j$. For every $\vec{\sigma}$, in the sub-game $S_{\vec{\sigma}} \times T_{\vec{\sigma}}$ it holds that the provers win the influential rounds. Moreover, whenever the provers win the influential rounds, there is $\vec{\sigma}$ such that they land in the sub-game $S_{\vec{\sigma}} \times T_{\vec{\sigma}}$. Note that there are only $|\Sigma_Y|^l$ sub-games $S_{\vec{\sigma}} \times T_{\vec{\sigma}}$. Hence, for any $0 < \delta' < 1$, the probability of landing in sub-games $S_{\vec{\sigma}} \times T_{\vec{\sigma}}$ where $S_{\vec{\sigma}}$ or $T_{\vec{\sigma}}$ have probability at most δ' is at most $\delta' |\Sigma_Y|^l$.

Set $\delta' = \delta/\varepsilon^3$. For the remainder of the analysis, we focus on a choice of $S_{\vec{\sigma}}$ and $T_{\vec{\sigma}}$ whose probabilities are at least δ'. Let $(\mathcal{G}^{\otimes k})''$ be the sub-game after the additional conditioning in $S_{\vec{\sigma}}$ and $T_{\vec{\sigma}}$. Let $\hat{k} = k - l$, and denote $[\hat{k}] = \{1,\ldots,k\} - I$. Note that effectively $(\mathcal{G}^{\otimes k})''$ has only $\hat{k}$ rounds. For every $i \in [\hat{k}]$, define the game $\hat{\mathcal{G}}_i$ as the restriction of the game $(\mathcal{G}^{\otimes k})''$ to the i-th round, where the provers are given their questions in all $\hat{k}$ rounds, but are tested only on their answers in the i-th round. Consider the marginals of $S_{\vec{\sigma}}$ and $T_{\vec{\sigma}}$ corresponding to the i-th question, and let $\mathcal{G}_i$ denote the sub-game of $\mathcal{G}$ corresponding to those marginals. Note that this sub-game is δ-large.

In Claim 5.2 we use the independence between the rounds and the extractor structure of G to argue that, no matter what was the fixing of questions and Σ_Y-labels for the influential rounds, a strategy for $\hat{\mathcal{G}}_i$ can be used to derive a strategy for $\mathcal{G}_i$ whose value is at least $val(\hat{\mathcal{G}}_i) - O(k\varepsilon)$. By definition, we have $val(\mathcal{G}_i) \leq val_{\delta'}(\mathcal{G})$, and hence $val(\hat{\mathcal{G}}_i) \leq val_{\delta'}(\mathcal{G}) + O(k\varepsilon)$. On the other hand, from Lemma 5.1, if we take expectation of $val(\hat{\mathcal{G}}_i)$ over all fixing of questions and Σ_Y-labels for the influential rounds (each weighted according to its probability in $\mathcal{G}^{\otimes k}$ with the provers strategy we fixed),

$$\mathbf{E}\left[val(\hat{\mathcal{G}}_i)\right] > \hat{\varepsilon} - \delta' |\Sigma_Y|^l \geq \hat{\varepsilon} - \varepsilon.$$

Since the left hand side is upper bounded by $val_{\delta'}(\mathcal{G}) + O(k\varepsilon)$, we get a contradiction (recall the definition of $\hat{\varepsilon}$). The heart of our analysis is the following claim; a discussion comparing our ideas to those of Raz can be found subsequently.

Claim 5.2 (One round approximation): There is $err = O(k\varepsilon)$, such that for every fixing of questions and Σ_Y labels to the influential rounds (captured by events $S_{\vec{\sigma}}$, $T_{\vec{\sigma}}$), for every $i \in [\hat{k}]$,

$$val(\hat{\mathcal{G}}_i) \leq val(\mathcal{G}_i) + err.$$

Proof: We consider the event $y_i = y$ for each $y \in Y$, and relate the provers winning in $\hat{\mathcal{G}}_i$, the i-th round of the repeated game $\mathcal{G}^{\otimes k}$, to their winning in $\mathcal{G}_i$, the corresponding sub-game of $\mathcal{G}$.

For every $y \in Y$, consider the bipartite graph $(G^{\otimes k})_{y,i}$ whose vertices consist of all $\vec{x} = (x_1,\ldots,x_k) \in X^k$ such that $\{x_j\}_{j\in I}$ is as fixed and $(x_i, y) \in E$, and all $\vec{y} = (y_1,\ldots,y_k) \in Y^k$ such that $\{y_j\}_{j\in I}$ is as fixed and $y_i = y$. There is an edge between $\vec{x}$ and $\vec{y}$ if $e_j = (x_j, y_j) \in E$ for all $1 \leq j \leq k$. Denote $(G^{\otimes k})_{y,i} = ((X^k)_{y,i}, (Y^k)_{y,i}, (E^k)_{y,i})$. Since G is a (δ,ε) extractor, the product graph $(G^{\otimes k})_{y,i}$ is a $(\delta'' = \delta/\varepsilon, \varepsilon'' = 2k\varepsilon)$-extractor [7].

Let $S_y \subseteq (X^k)_{y,i}$ be those vertices $\vec{x} \in (X^k)_{y,i}$ in $S_{\vec{\sigma}}$. Let $T_y \subseteq (X^k)_{y,i}$ be those vertices $\vec{x} \in (X^k)_{y,i}$ in $T_{\vec{\sigma}}$. Partition the vertices $\vec{x} \in S_y$ according to the assignments to y: For every $b \in \Sigma_Y$, let $S_{y,b} \subseteq S_y$ consist of those $\vec{x}$ for which prover 1 assigns label $a_i \in \Sigma_X$ to x_i and $\pi_{e_i}(a_i) = b$. Partition the vertices $\vec{x} \in T_y$ according to the assignments to y: For every $b \in \Sigma_Y$, let $T_{y,b} \subseteq T_y$ consist of those $\vec{x}$ for which prover 2 assigns label $a_i \in \Sigma_X$ to x_i and $\pi_{e_i}(a_i) = b$.

Focus on $y \in Y$ such that $|S_y| \geq \varepsilon \mathbf{E}_y[|S_y|]$ and $|T_y| \geq \varepsilon \mathbf{E}_y[|T_y|]$. The probability other y's are selected as y_i in $\hat{\mathcal{G}}_i$ is at most ε. Focus on $b \in \Sigma_Y$ such that $|S_{y,b}| \geq \varepsilon |S_y|$ and $|T_{y,b}| \geq \varepsilon |T_y|$. The contribution to winning $\hat{\mathcal{G}}_i$ from b's that do not satisfy this is at most ε. We have $|S_{y,b}| \geq \varepsilon |S_y| \geq \varepsilon \cdot \varepsilon \sum_y |S_y| \geq \varepsilon^2 \delta' \left|(X^k)_{y,i}\right|$ and similarly, $|T_{y,b}| \geq \varepsilon^2 \delta' \left|(X^k)_{y,i}\right|$. Since $\varepsilon^2\delta' \geq \delta''$, we can apply the extractor property of $(G^{\otimes k})_{y,i}$ and get that the probability distribution of $\vec{y}$ conditioned on $\vec{x} \in S_{y,b}$ is ε''-close to uniform over $(Y^k)_{y,i}$.

Consider the event that when picking uniformly at random $\vec{y} \in (Y^k)_{y,i}$ and $\vec{x} \in S_y$, $\vec{x}' \in T_y$ such that $(\vec{x},\vec{y}), (\vec{x}',\vec{y}) \in (E^k)_{y,i}$, it holds that $\vec{x}' \in T_{y,b}$. Note that by bi-regularity, $\vec{x}'$ is uniform in T_y, and hence this event happens with probability $|T_{y,b}| / |T_y|$. Similarly, $\Pr[\vec{x} \in S_{y,b}] = |S_{y,b}| / |S_y|$. By the extractor property,

$$\Pr[\vec{x}' \in T_{y,b} | \vec{x} \in S_{y,b}] \leq \frac{|T_{y,b}|}{|T_y|} + \varepsilon''.$$

Hence,

$$\Pr[\vec{x}' \in T_{y,b} \wedge \vec{x} \in S_{y,b}] \leq \frac{|S_{y,b}|}{|S_y|} \cdot \frac{|T_{y,b}|}{|T_y|} + \varepsilon''.$$

The overall probability of winning $\hat{\mathcal{G}}_i$, accounting also for y's and b's where $S_{y,b}$ or $T_{y,b}$ are not as above, is at most

$$\mathop{\mathbf{E}}_{y\in Y}\left[\sum_{b\in\Sigma_Y}\frac{|S_{y,b}|}{|S_y|}\cdot\frac{|T_{y,b}|}{|T_y|}\right]+O(\varepsilon''+\varepsilon) \tag{1}$$

Next we use this assertion to devise a successful strategy for the sub-game $\mathcal{G}_i$ of $\mathcal{G}$. In $\mathcal{G}_i$ the verifier picks $y \in Y$ uniformly at random, and then $\vec{x} \in S_y$, $\vec{x}' \in T_y$. The verifier sends the provers x and x', which are the i-th coordinates of $\vec{x}$ and $\vec{x}'$, respectively. Upon receiving answers a and a' from the provers, the verifier checks that the answers agree on y, i.e., $\pi_{(x,y)}(a) = \pi_{(x',y)}(a')$. We consider the strategy of the provers where the first prover picks uniformly $\vec{x} \in S_{\vec{\sigma}}$ with x in the i-th coordinate, and the second prover picks uniformly $\vec{x}' \in T_{\vec{\sigma}}$ with x' in the i-th coordinate. Each prover then responds with the i-th answer of its repeated strategy. Note that the provers do not necessarily guess the same $\vec{x}$ and $\vec{x}'$ that the verifier used to generate x and x', however, no matter which $y \in Y$ (unknown to the provers!) the verifier chose, we have that the $\vec{x}$ chosen by the first prover is uniform in S_y, and the $\vec{x}'$ chosen by the second prover is uniform in T_y. Therefore, the probability that the strategy we defined succeeds in $\mathcal{G}_i$ is precisely:

$$\mathop{\mathbf{E}}_{y\in Y}\left[\sum_{b\in\Sigma_Y}\frac{|S_{y,b}|}{|S_y|}\cdot\frac{|T_{y,b}|}{|T_y|}\right] \tag{2}$$

The lemma follows from (1) and (2). ∎

It is interesting to contrast our proof with that of Raz [26]. In Raz's proof, given questions $x, x' \in X$ in the original game $\mathcal{G}$, the provers appeal to the strategy in $\hat{\mathcal{G}}_i$ by coordinating questions to the rounds other than i, and playing the repeated strategy where x and x' are in the i-th round. Coordinating the questions in the remaining rounds is quite difficult because of the various correlations between the questions in $\hat{\mathcal{G}}_i$, and this is where the trick of Razborov [29] comes in. However, once this is achieved, Raz can directly relate $\mathcal{G}$ to $\hat{\mathcal{G}}_i$, and success in the latter corresponds to success in the former. We, on the other hand, take a different route. We argue that after the conditioning in the questions and Σ_Y-labels of the influential rounds, the provers in fact have a successful *global* strategy for the i-th round! This follows since if the provers have different answers for a question y for different settings of the questions in the other rounds, then the extractor guarantees that the repeated verifier detects inconsistency.

VI. Derandomized Parallel Repetition, Two Rounds and Composition

A natural question is whether it is possible to apply our parallel repetition and fortification ideas in order to obtain a projection PCP with soundness error lower than $1/poly\log n$. To obtain such a low error we can no longer apply parallel repetition with k independent rounds. The reason is that this requires a super-constant k, for which parallel repetition blows-up the size to n^k. A natural idea is to use k correlated rounds; an idea often referred to as "derandomizing parallel repetition". In this section we explain the difficulty in "fortifying" in the derandomized setting. Moreover, we relate the problem of derandomizing parallel repetition to a different well-studied problem in PCP; that of *composition*. While we continue to use our notation from the previous part of the paper, everything in this part of the paper holds for general two prover games.

A. Correlation and Fortification

We start with explaining what breaks down in the analysis in Section V when considering the correlated case. In Section V we fix questions in the influential rounds and relate the game in the remaining rounds to $\mathcal{G}$. This approach fails in the correlated case, as the questions in the remaining rounds are likely to be extremely far from uniform in $\mathcal{G}$ after such a fixing. The fixing was used in order to prevent conditioning in W from introducing dependencies between the questions of the provers beyond those captured by the graph of $\mathcal{G}^{\otimes k}$. The latter is what allowed us to fortify only against rectangular sub-games.

When conditioning on an event W that arbitrarily depends on the questions to both provers, fortification with respect to rectangular sub-games is no longer sufficient, nor is fortifying a single round without taking others into account. A natural generalization of fortfication is with respect to general large sub-games of $\mathcal{G}^{\otimes k}$. However, the condition that *any* (non-rectangular) sub-game of fraction at least δ in $\mathcal{G}^{\otimes k}$ has value smaller than 1 is equivalent to the statement that $val(\mathcal{G}^{\otimes k}) < \delta$, which is precisely what we try to prove! Interestingly, the value of a *random* large sub-game of $\mathcal{G}^{\otimes k}$ (indeed, of any game with value sufficiently smaller than 1) does have value smaller than 1 with high probability [8]. However, since the provers are adversarial, this does not constitute a useful fortification. An intriguing open problem following this work is to define fortification that is both easy to analyze and useful for the correlated case.

B. Degrees and Two Rounds

The degrees in the graph associated with a two prover game $\mathcal{G}$ correspond to the uncertainty each prover has with respect to the questions of the other prover. We know that the degrees have to be at least $1/\delta$ to allow for $val(\mathcal{G}) \leq \delta$ (assuming that each test of the verifier can be satisfied by itself). In fact, Feige and Kilian [14] show that in the randomness-efficient case, for $val(\mathcal{G}^{\otimes k}) \leq \delta$ it must be the case that the degrees in $\mathcal{G}$'s graph are at least $\approx 1/\delta$ (under an assumption on $\mathcal{G}$ they call *softness*, which indeed holds in the cases of interest). In this section we relax the problem of derandomizing parallel repetition to a corresponding combinatorial problem about degrees in graphs. We call the combinatorial problem the Two Rounds problem. We

then relate the Two Rounds problem to the well-studied problem of composition of two prover games, and show that any efficient composition scheme yields a solution to the problem.

We remark that large degrees are a necessary, but not a sufficient, condition for small value. In fact, in general two prover games one can increase the degree artificially, and without decreasing the value, by duplicating questions. Interestingly, for projection games on (δ, ε)-extractors $G = (X, Y, E)$, the degree in Y is necessarily large $\approx 1/\delta$, while the degree in X cannot be artificially increased by duplicating Y vertices due to the extractor property. This observation supports the intuition that large degrees are "morally" a sufficient condition for low value, at least in cases of interest.

Definition 6.1 (Two Rounds): Given two projection games $\mathcal{G}$, $\mathcal{H}$ on bi-regular graphs $G = (X, Y, E)$ and $H = (X', Y', E')$, respectively, where the degrees are at least d, we say that a distribution over pairs $(\vec{x}, \vec{y})$, where $\vec{x} = (x, x') \in X \times X'$ and $\vec{y} = (y, y') \in Y \times Y'$, yields "two rounds" of $\mathcal{G}$ and $\mathcal{H}$ for parameter d, if:

- (x, y) is a uniformly distributed edge in E, while (x', y') is a uniformly distributed edge in E'.
- For any fixing of x, x' (similarly, y, y'), we have $H(y|y'), H(y'|y) \geq \log d$ (similarly, $H(x|x'), H(x'|x) \geq \log d$).

Note that picking two *independent* uniform edges $(x, y) \in E$ and $(x', y') \in E'$ yields two rounds. The challenge is to pick two rounds using less randomness. Ideally, one could hope to use $\log |E| + O(\log D)$ randomness, when D is the maximal degree in the graph, since given (x, x') (similarly, given (y, y')), there are at most D^2 alternatives for (y, y') (respectively, (x, x')).

A randomness-efficient solution to the Two Rounds problem for games $\mathcal{G}$ and $\mathcal{H}$ yields a candidate construction for a derandomized 2-round parallel repetition, where the tensored games are $\mathcal{G}$ and $\mathcal{H}$: the first prover gets questions x, x', while the second prover gets questions y, y'; the first prover answers $a, a' \in \Sigma_X$, while the second prover answers $b, b' \in \Sigma_Y$; the verifier checks that $\pi_e(a) = b$ and $\pi_{e'}(a') = b'$ (for simplicity, in this part of the paper we consider the more standard definition of projection games; see Remark 1.1). The definition of two rounds guarantees that a prover who knows both x and x', even if it has information on y (e.g., by virtue of conditioning on an event W), has a lot of uncertainty about y'. The same goes for the other question and the other prover.

The hope is that there are randomness-efficient two rounds – and, more generally, k rounds – for "interesting" games $\mathcal{G}$ and $\mathcal{H}$, namely, ones whose value is NP-hard to approximate. Ideally, such a derandomized parallel repetition scheme would yield $val(\mathcal{G}^{\otimes k}) \leq \delta$ when the size of $\mathcal{G}^{\otimes k}$ is $size(\mathcal{G}) \cdot D^{O(k)}$ for $D, d = \Theta(1/\delta)$. In other words, this would give a projection PCP with soundness error $\delta = 2^{-(\log n)^{\beta}}$ for some constant $\beta > 0$, i.e., exponentially smaller than what we know today.

We will briefly explain the relation between the Two Rounds problem and the composition problem. More details can be found in the full version of this work. The goal in the composition problem is to take an *outer* game with large alphabets, as well as small *inner* games with small alphabets, and compose them into a single game with small alphabets. This is similar to concatenation for codes, where one combines an outer code with large alphabet and inner codes over a small alphabet to get a single code over the small alphabet. The idea of composition is to simulate a test of the outer game using a test of an inner game, since the latter only requires small alphabet. In the notation of the Two Rounds problem, the edge (x, y) corresponds to an edge of the outer game, while the edge (x', y') corresponds to an edge of the inner game. To allow composition, given questions x, x' from both the outer and inner games, the prover should not gain much advantage in guessing either y or y'. The same should hold given questions y, y'. This connection between composition and the Two Rounds problem leads to the following understanding: The difference between composition and repetition is that in composition the second round comes to replace the first round, while in repetition the second round is in addition to the first round.

Let us denote the "input size" for the inner game by $n' = \log(|\Sigma_X| + |\Sigma_Y|)$. As it stands now, there are constructions of inner games based on the Hadamard code and based on the long code (these are variants of standard constructions as in [9] and [24]). Hadamard-based constructions have $|X'|, |Y'| = poly(\exp(n'))$, while the long code-based constructions have $|X'|, |Y'| = \exp(\exp(n'))$. Both have alphabets that are of size polynomial in $1/\varepsilon$. The utility of these constructions follows from the asymmetry between the outer and inner games, which is not present between repetitions of the same game. There are also constructions that have size polynomial in n' [23], [11], alas, they have a large alphabet $|\Sigma_{X'}| = \exp(poly(1/\varepsilon))$. To improve on the current state of the art in PCP one has to design inner games for the case of $n' \geq poly \log n$, whose alphabets are of size $\ll \exp(poly(1/\varepsilon))$ (desirably, $poly(1/\varepsilon)$).

The connection between composition and repetition via Two Rounds highlights a crucial barrier toward PCPs with lower error that occurs when designing either repetition or composition schemes. The connection might also lead to new schemes for either problem.

ACKNOWLEDGEMENTS

I am thankful to Ran Raz for discussions, and to Henry Yuen and an anonymous reviewer for a careful reading of the paper.

References

[1] S. Aaronson, D. Moshkovitz, and R. Impagliazzo. AM with multiple merlins. In *Computational Complexity Conference*, 2014.

[2] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM*, 45(3):501–555, 1998.

[3] S. Arora and S. Safra. Probabilistic checking of proofs: a new characterization of NP. *Journal of the ACM*, 45(1):70–122, 1998.

[4] L. Babai, L. Fortnow, L. A. Levin, and M. Szegedy. Checking computations in polylogarithmic time. In *Proc. 23rd ACM Symp. on Theory of Computing*, pages 21–32, 1991.

[5] L. Babai, L. Fortnow, and C. Lund. Nondeterministic exponential time has two-prover interactive protocols. *Computational Complexity*, 1:3–40, 1991.

[6] B. Barak, M. Hardt, I. Haviv, A. Rao, O. Regev, and D. Steurer. Rounding parallel repetitions of unique games. In *Proc. 49th IEEE Symp. on Foundations of Computer Science*, pages 374–383, 2008.

[7] M. R. Capalbo, O. Reingold, S. P. Vadhan, and A. Wigderson. Randomness conductors and constant-degree lossless expanders. In *IEEE Conference on Computational Complexity*, page 15, 2002.

[8] M. Dinitz, G. Kortsarz, and R. Raz. Label cover instances with large girth and the hardness of approximating basic k-spanner. In *ICALP*, pages 290–301, 2012.

[9] I. Dinur. The PCP theorem by gap amplification. *Journal of the ACM*, 54(3):12, 2007.

[10] I. Dinur and P. Harsha. Composition of low-error 2-query PCPs using decodable PCPs. In *Proc. 50th IEEE Symp. on Foundations of Computer Science*, pages 472–481, 2009.

[11] I. Dinur and O. Meir. Derandomized parallel repetition via structured PCPs. *Computational Complexity*, 20(2):207–327, 2011.

[12] I. Dinur and D. Steurer. Analytical approach to parallel repetition. In *Proc. 46th ACM Symp. on Theory of Computing*, 2014.

[13] U. Feige. On the success probability of the two provers in one round proof systems. In *Proc. of 6th IEEE Symposium on Structure in Complexity Theory*, pages 116–123, 1991.

[14] U. Feige and J. Kilian. Impossibility results for recycling random bits in two-prover proof systems. In *Proc. 27th ACM Symp. on Theory of Computing*, pages 457–468, 1995.

[15] U. Feige and J. Kilian. Two-prover protocols - low error at affordable rates. *SIAM Journal on Computing*, 30(1):324–346, 2000.

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

[17] L. Fortnow, J. Rompel, and M. Sipser. Errata for on the power of multi-prover interactive protocols. In *Structure in Complexity Theory Conference*, pages 318–319, 1990.

[18] J. Håstad. Some optimal inapproximability results. *Journal of the ACM*, 48(4):798–859, 2001.

[19] T. Holenstein. Parallel repetition: Simplification and the no-signaling case. *Theory of Computing*, 5(1):141–172, 2009.

[20] R. Impagliazzo, V. Kabanets, and A. Wigderson. New direct-product testers and 2-query PCPs. *SIAM Journal on Computing*, 41(6):1722–1768, 2012.

[21] D. Moshkovitz. The projection games conjecture and the NP-hardness of $\ln n$-approximating set-cover. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques - 15th International Workshop, APPROX 2012*, volume 7408, pages 276–287, 2012.

[22] D. Moshkovitz. An approach to the sliding scale conjecture via parallel repetition for low degree testing. Technical Report 30, ECCC, 2014.

[23] D. Moshkovitz and R. Raz. Two query PCP with sub-constant error. *Journal of the ACM*, 57(5), 2010.

[24] J. Radhakrishnan and M. Sudan. On Dinur's proof of the PCP theorem. *Bulletin of the AMS*, 44(1):19–61, 2007.

[25] A. Rao. Parallel repetition in projection games and a concentration bound. *SIAM Journal on Computing*, 40(6):1871–1891, 2011.

[26] R. Raz. A parallel repetition theorem. In *SIAM Journal on Computing*, volume 27, pages 763–803, 1998.

[27] R. Raz. A counterexample to strong parallel repetition. *SIAM Journal on Computing*, 40(3):771–777, 2011.

[28] R. Raz and R. Rosen. A strong parallel repetition theorem for projection games on expanders. In *IEEE Conference on Computational Complexity*, pages 247–257, 2012.

[29] A. A. Razborov. On the distributional complexity of disjointness. *Theor. Comput. Sci.*, 106(2):385–390, 1992.

[30] O. Reingold, S. P. Vadhan, and A. Wigderson. Entropy waves, the zig-zag graph product, and new constant-degree expanders and extractors. *Annals of Mathematics*, 155(1):157–187, 2002.

[31] R. Shaltiel. Derandomized parallel repetition theorems for free games. *computational complexity*, 22(3):565–594, 2013.