Analytical Approach to Parallel Repetition

Irit Dinur* David Steurer†

June 10, 2014

Abstract

We propose an analytical framework for studying parallel repetition, a basic product operation for one-round two-player games. In this framework, we consider a relaxation of the value of projection games. We show that this relaxation is multiplicative with respect to parallel repetition and that it provides a good approximation to the game value. Based on this relaxation, we prove the following improved parallel repetition bound: For every projection game G with value at most ρ, the k-fold parallel repetition $G^{\otimes k}$ has value at most

$$\mathrm{val}(G^{\otimes k}) \leqslant \left(\frac{2\sqrt{\rho}}{1+\rho}\right)^{k/2}.$$

This statement implies a parallel repetition bound for projection games with low value ρ. Previously, it was not known whether parallel repetition decreases the value of such games. This result allows us to show that approximating SET COVER to within factor $(1-\varepsilon)\ln n$ is NP-hard for every $\varepsilon > 0$, strengthening Feige's quasi-NP-hardness and also building on previous work by Moshkovitz and Raz.

In this framework, we also show improved bounds for few parallel repetitions of projection games, showing that Raz's counterexample to strong parallel repetition is tight even for a small number of repetitions.

Finally, we also give a short proof for the NP-hardness of LABEL COVER$(1,\delta)$ for all $\delta > 0$, starting from the basic PCP theorem.

Keywords: parallel repetition, one-round two-player games, label cover, set cover, hardness of approximation, copositive programming, operator norms.

*Department of Computer Science and Applied Mathematics, Weizmann Institute. Part of this work was done at Microsoft Research New England and Radcliffe Institute for Advanced Study. Work supported in part by an ISF grant number 1179/09 and by an ERC-StG grant number 239985.

†Computer Science Department, Cornell University. Part of this work was done at Microsoft Research New England. Research supported in part by NSF and the Alred P. Sloan Foundation.

Contents

1 Introduction

A *one-round two-player game* G is specified by a bipartite graph with vertex sets U and V and edges decorated by constraints $\pi \subseteq \Sigma \times \Sigma$ for an alphabet Σ. The *value* of a game is the maximum, over all assignments $f\colon U \to \Sigma$ and $g\colon V \to \Sigma$, of the fraction of constraints satisfied (where a constraint π is satisfied if $(f(u), g(v)) \in \pi$)

$$\mathrm{val}(G) = \max_{f,g} \mathop{\mathbb{P}}_{u,v,\pi} \left\{ \Big(f(u), g(v) \Big) \in \pi \right\}.$$

The term one-round two-player game stems from the following scenario: A referee interacts with two players, Alice and Bob. Alice has a strategy $f\colon U \to \Sigma$, and Bob a strategy $g\colon V \to \Sigma$. A referee selects a random edge u, v in E and sends u as a question to Alice, and v as a question to Bob. Alice responds with $f(u)$ and Bob with $g(v)$. They succeed if their answers satisfy the constraint decorating the edge uv.

In the k-fold *parallel repetition* $G^{\otimes k}$, the referee selects k edges $u_1v_1, \ldots, u_kv_k$ independently from E and sends a question tuple $u_1, \ldots, u_k$ to Alice, and $v_1, \ldots, v_k$ to Bob. Each player responds with a k-tuple of answers and they succeed if their answers satisfy each of the k constraints on these edges.

Parallel repetition is a basic product operation on games, and yet its effect on the game value is far from obvious. Contrary to what one might expect, there are strategies for the repeated game that do significantly better than the naive strategy answering each of the k questions using the best single-shot strategy. Nevertheless, the celebrated parallel repetition theorem of Raz [Raz98] bounds the value of $G^{\otimes k}$ by a function of the value of G that decays exponentially with the number of repetitions. The broad impact of this theorem can be partly attributed to the general nature of parallel repetition. It is an operation that can be applied to any game without having to know almost anything about its structure. Raz's proof has since been simplified, giving stronger and sometimes tight bounds [Hol09, Rao11]. Still, there is much that is left unknown regarding the behavior of games under parallel repetition. For example, previous to this work, it was not known if repetition causes a decrease in the value for a game whose value is already small, say sub-constant. It was also not known how to bound the value of the product of just two games. Other open questions include bounding the value of games with more than two players, and bounding the value of entangled games (where the two players share quantum entanglement). The latter question has recently received some attention [DSV14, JPY14, CS13].

1.1 Our Contribution

Our main contribution is a new analytical framework for studying parallel repetitions of projection games. In this framework, we prove for any *projection game*:[1]

Theorem 1.1 (Parallel Repetition Bound)**.** *Let G be a projection game. If* $\mathrm{val}(G) \leqslant \rho$ *for some* $\rho > 0$ *then,*

$$\mathrm{val}(G^{\otimes k}) \leqslant \left(\frac{2\sqrt{\rho}}{1+\rho} \right)^{k/2}.$$

We remark that for values of $\mathrm{val}(G)$ close to 1 this theorem matches Rao's bound for projection games, with improved constants.

[1] In a projection game, for any two questions u and v to the players and any answer β of *Bob*, there exists at most one acceptable answer α for Alice.

Corollary 1.2 (Parallel Repetition for high-value games [Rao08a])**.** *For any projection game G with* $\mathrm{val}(G) \leqslant 1 - \varepsilon$,

$$\mathrm{val}(G^{\otimes k}) \leqslant \left(1 - \varepsilon^2/16\right)^k .$$

We give a particularly short proof of a (strong) parallel repetition bound for a subclass of games, namely expanding projection games. This class of games is rich enough for the main application of parallel repetition: NP-hardness of LABEL COVER with perfect completeness and soundness close to 0 (the starting point for most hardness of approximation results). See Section 3.2.

Next, we list some new results that are obtained by studying parallel repetition in this framework.

Repetition of small-value games. Our first new result is that if the initial game G has value ρ that is possibly sub-constant, then the value of the repeated game still decreases exponentially with the number of repetitions. Indeed, Theorem 1.1 for small ρ becomes,

Corollary 1.3 (Repetition of games with small value)**.** *For any projection game G with* $\mathrm{val}(G) \leqslant \rho$,

$$\mathrm{val}(G^{\otimes k}) \leqslant (4\rho)^{k/4} .$$

This bound allows us to prove NP-hardness for LABEL COVER that is better than was previously known (see Theorem 6.4) by applying our small-value parallel repetition theorem on the PCP of [MR10]. A concrete consequence is the following corollary.

Corollary 1.4 (NP-hardness for LABEL COVER)**.** *For every constant $c > 0$, given a* LABEL COVER *instance of size n with alphabet size at most n, it is NP-hard to decide if its value is* 1 *or at most* $\varepsilon = \frac{1}{(\log n)^c}$.

Hardness of SET COVER. A famous result of Uriel Feige [Fei98] is that unless $NP \subseteq DTIME(n^{O(\log\log n)})$ there is no polynomial time algorithm for approximating SET COVER to within factor $(1 - o(1)) \ln n$. Feige's reduction is slightly super-polynomial because it involves, on top of the basic PCP theorem, an application of Raz's parallel repetition theorem with $\Theta(\log\log n)$ number of repetitions. Later, Moshkovitz and Raz [MR10] constructed a stronger PCP whose parameters are closer but still not sufficient for lifting Feige's result to NP-hardness. Moshkovitz [Mos12] also generalized Feige's reduction to work from a generic projection LABEL COVER rather than the specific one that Feige was using. Our Corollary 1.4 makes the last step in this sequence of works and gives the first tight NP-hardness for approximating SET COVER.

Corollary 1.5 (Tight NP-hardness for approximating SET COVER)**.** *For every $\alpha > 0$, it is NP-hard to approximate* SET COVER *to within $(1 - \alpha) \ln n$, where n is the size of the instance. The reduction runs in time $n^{O(1/\alpha)}$.*

Unlike the previous quasi NP-hardness results for SET COVER, Corollary 1.5 rules out that approximation ratios of $(1-\alpha) \ln n$ can be achieved in time $2^{n^{o(1)}}$ (unless $\mathrm{NP} \subseteq \mathrm{TIME}(2^{n^{o(1)}})$). Together with the best known approximation algorithm for SET COVER [CKW09], we can characterize the time vs. approximation trade-off for the problem:

Corollary 1.6. *Assuming* $\mathrm{NP} \not\subseteq \mathrm{TIME}(2^{n^{o(1)}})$*, the time complexity of achieving an approximation ratio $(1 - \alpha) \ln n$ for* SET COVER *is $2^{n^{\Theta(\alpha)}}$.*

Going back to LABEL COVER, we remark that the hardness proven in Corollary 1.4 is still far from the known algorithms for LABEL COVER and it is an interesting open question to determine the correct tradeoff between ε and the alphabet size.

Few repetitions. Most parallel repetition bounds are tailored to the case that the number of repetitions k is large compared to $1/\varepsilon$ (where as usual, $val(G) = 1-\varepsilon$). For example, when the number of repetitions $k \ll 1/\varepsilon$, the bound $\mathrm{val}(G^{\otimes k}) \leqslant (1-O(\varepsilon^2))^k \approx (1-O(k\varepsilon^2))$ for projection games [Rao11] is weaker than the trivial bound $\mathrm{val}(G^{\otimes k}) \leqslant 1-\varepsilon$. The following theorem gives an improved and tight bound when $k \ll 1/\varepsilon^2$.

Theorem 1.7 (Few repetitions)**.** *Let G be a projection game with* $\mathrm{val}(G) = 1-\varepsilon$*. Then for all* $k \ll 1/\varepsilon^2$,

$$\mathrm{val}(G^{\otimes k}) \leqslant 1 - \Omega(\sqrt{k}\cdot\varepsilon)\,.$$

A relatively recent line of work [FKO07, Raz08, BHH+08, BRR+09, RR12] focused on the question of strong parallel repetition. Namely, given a game G with $\mathrm{val}(G) \leqslant 1-\varepsilon$ is it true that $\mathrm{val}(G^{\otimes k}) \leqslant (1-O(\varepsilon))^k$? If true for any unique game G, such a bound would imply a reduction from MAX CUT to unique games. However, Raz [Raz08] showed that the value of the odd cycle xor game is at least $1-O(\sqrt{k}\cdot\varepsilon)$, much larger than the $1-O(k\varepsilon)$ in strong parallel repetition. Our bound matches Raz's bound even for small values of k, thereby confirming a conjecture of Ryan O'Donnell[2] and extending the work of [FKO07] who proved such an upper bound for the odd-cycle game.

1.2 High Level Proof Overview

We associate a game G with its *label-extended graph*, by blowing up each vertex in U or in V to a cloud of $|\Sigma|$ vertices. In this bipartite graph we connect a left vertex (u,α) in the cloud of u to a right vertex (v,β) in the cloud of v iff (α,β) satisfy the constraint attached to u,v. This graph naturally gives rise to a linear operator mapping functions f on the right vertices to functions Gf on the left vertices, where $Gf(u,\alpha)$ is defined by aggregating the values of $f(v,\beta)$ over all neighboring vertices.

In this language, the product operation on games is given by the tensor product operation on the corresponding linear operators.

It turns out that a good measure for the value of the game G is its *collision value*, denoted $\|G\|$, defined by measuring how self-consistent Bob's strategy (answering questions in V) is, with respect to its projection on a random u. This value is not new and has been implicitly studied before as it describes the value of a symmetrization of the game G, for example it occurs in PCP constructions when moving from a line versus point test to a line versus line test. A simple Cauchy–Schwarz inequality shows that (see Claim 2.3)

$$\mathrm{val}(G) \leqslant \|G\| \leqslant \mathrm{val}(G)^{1/2}\,. \tag{1.1}$$

The collision value is more amenable to analysis, and indeed, our main technical theorem shows that the collision value has the following very nice property,

Theorem 1.8. *Any two projections games G and H satisfy $\|G\otimes H\| \leqslant \varphi(\|G\|)\cdot\|H\|$, where* $\varphi(x) = \frac{2\sqrt{x}}{1+x}$.

Theorem 1.1 follows directly from repeated applications of this theorem, together with the bound in (1.1). We remark that a similar statement for $\mathrm{val}(\cdot)$ instead of $\|\cdot\|$, i.e. of the form $\mathrm{val}(G\otimes H) \leqslant \varphi(\mathrm{val}(G))\cdot\mathrm{val}(H)$, is false. Feige shows [Fei91] an example of a game for which $\mathrm{val}(G\otimes G) = \mathrm{val}(G) = \frac{1}{2}$, see discussion in Appendix B. In contrast, Theorem 1.8 implies that there is no projection game G for which $\|G\otimes G\| = \|G\| < 1$.

The proof of Theorem 1.8 proceeds by looking for a parameter ρ_G for which $\|G\otimes H\| \leqslant \rho_G\cdot\|H\|$, and such that ρ_G depends only on G. One syntactic possibility is to take ρ_G to be

[2] Personal communication, 2012.

the supremum of $\frac{\|G\otimes H\|}{\|H\|}$ over all games H,[3] but from this definition it is not clear how to prove that $\rho_G \approx \mathrm{val}(G)$.

Instead, starting from this ratio we arrive at a slightly weaker relaxation that is expressed as a *generalized Rayleigh quotient*,[4] or, more precisely, as the ratio of two collision values: one for the game G and the other involving a trivial game T that provides suitable normalization. The key difference to standard linear-algebraic quantities is that we restrict this generalized Rayleigh quotient to functions that take only *nonnegative values*. Thus, intuitively one can think of the value $\mathrm{val}_+(G)$ as a "positive eigenvalue" of the game G. In order to connect the ratio $\frac{\|G\otimes H\|}{\|H\|}$ to a generalized Rayleigh quotient for G and T, we factor the operator $G\otimes H$ into two consecutive steps

$$G \otimes H = (G \otimes \mathrm{Id})(\mathrm{Id} \otimes H)\,.$$

The same factorization can be applied to the operator $T \otimes H$. This factorization allows us to magically "cancel out" the game H, and we are left with an expression that depends only on G.

The main technical component of the proof is to prove that $\mathrm{val}_+(G) \approx \mathrm{val}(G)$, and this proof has two components. The first is a rounding algorithm that extracts an assignment from a non-negative function with large Rayleigh quotient (i.e., a large ratio between the norm of Gf and the norm of Tf). In the case of expanding games, this is the only component necessary and it is obtained rather easily from the expander mixing lemma. For non-expanding games, we rely on a subtler (more "parameter-sensitive") proof, building on a variant of Cheeger's inequality in [Ste10b]. Here a non-negative function will only give a good *partial assignment*: an assignment that assigns values only to a small subset of the vertices. We then combine many partial assignments into a proper assignment using *correlated sampling*, first used in this context by [Hol09].

1.3 Related work

Already in [FL92], Feige and Lovász proposed to study parallel repetition via a relaxation of the game value. Their relaxation is defined as the optimal value of a semidefinite program. While this relaxation is multiplicative, it does not provide a good approximation for the game value.[5] In particular, the value of this relaxation can be 1, even if the game value is close to 0. The proof that the *Feige–Lovász relaxation* is multiplicative uses semidefinite programming duality (similar to [Lov79]). In contrast, we prove the multiplicativity of val_+ in a direct way.[6]

For unique two-player games, Barak et al. [BHH$^+$08] introduced a new relaxation, called *Hellinger value*, and showed that this relaxation provides a good approximation to both the game value and the value of the Feige–Lovász relaxation (see [Ste10a] for improved approximation bounds). These quantitative relationships between game value, Hellinger value, and the Feige–Lovász relaxation lead to counter-examples to "strong parallel repetition," generalizing [Raz08].

[3] Goldreich suggests to call this the "environmental value" of a game G because it bounds the value of G relative to playing in parallel with any environment H

[4] Rayleigh quotients refer to the expressions of the form $\langle f, Af\rangle/\langle f, f\rangle$ for operators A and functions f. *Generalized Rayleigh quotients* refer to expressions of the form $f \mapsto \langle f, Af\rangle/\langle f, Bf\rangle$ for operators A and B and functions f.

[5] In fact, no polynomial-time computable relaxation can provide a good approximation for the game value unless P = NP, because the game value is NP-hard to approximate.

[6] The relaxation val_+ can be defined as a convex program (in fact, a copositive program). However, it turns out that unlike for semidefinite programs, the dual objects are not closed under tensor products.

The relaxation val_+ is a natural extension of the Hellinger value to projection games (and even, general games). Our proof that val_+ satisfies the approximation property for projection games follows the approach of [BHH+08]. The proof is more involved because, unlike for unique games, val_+ is no longer easily expressed in terms of Hellinger distances. Another difference to [BHH+08] is that we need to establish the approximation property also when the game's value is close to 0. This case turns out to be related to Cheeger-type inequalities in the near-perfect expansion regime [Ste10b].

1.4 Organization

In Section 2 we describe the analytic framework in which we study games. In Section 3 we outline the main approach and give a complete and relatively simple analysis of the parallel repetition bound for expanding projection games. This proof gives a taste of our techniques in a simplified setting, and gives gap amplification for LABEL COVER, thereby proving the NP-hardness of LABEL COVER $(1,\delta)$. In Section 4 we prove the approximation property of val_+ for non-expanding games and then immediately derive Theorem 1.1 and Corollary 1.3. In Section 5 we analyze parallel repetition with few repetitions, proving Theorem 1.7. We prove Corollary 1.4 and related hardness results for LABEL COVER and SET COVER in Section 6.

2 Technique

2.1 Games and linear operators

A *two-prover game* G is specified by a bipartite graph with vertex sets U and V and edges decorated by constraints $\pi \subseteq \Sigma \times \Sigma$ for an alphabet Σ. (We allow parallel edges and edges with nonnegative weights.) The graph gives rise to a distribution on triples (u, v, π) (choose an edge of the graph with probability proportional to its weight). The marginals of this distribution define probability measures on U and V. When the underlying graph is regular, these measures on U and V are uniform. It's good to keep this case in mind because it captures all of the difficulty. We write $(v,\pi) \mid u$ to denote the distribution over edges incident to a vertex u. (Formally, this distribution is obtained by selecting a triple (u, v, π) conditioned on u)

We say that G is a *projection game* if every constraint π that appears in the game is a projection constraint, i.e., each $\beta \in \Sigma$ has at most one $\alpha \in \Sigma$ for which $(\alpha,\beta) \in \pi$. We write $\alpha \overset{\pi}{\hookleftarrow} \beta$ to denote $(\alpha,\beta) \in \pi$ for projection constraints π. If the constraint is clear from the context, we just write $\alpha \hookleftarrow \beta$.

Linear-algebra notation for games. In this work we will represent an *assignment* for Bob by a nonnegative function $f\colon V \times \Sigma \to \mathbb{R}$ such that $\sum_{\beta\in\Sigma} f(v,\beta) = 1$ for every $v \in V$. The value $f(v,\beta)$ is interpreted as the probability that Bob answers β when asked v. Similarly, an assignment for Alice is a nonnegative function $g\colon U \times \Sigma \to \mathbb{R}$ such that $\sum_\alpha g(u,\alpha) = 1$ for all $u \in U$.

Let $L(U \times \Sigma)$ be the space of real-valued functions on $U \times \Sigma$ endowed with the inner product $\langle \cdot,\cdot \rangle$,

$$\langle g, g' \rangle = \mathop{\mathbb{E}}_{u} \sum_{\alpha} g(u,\alpha) g'(u,\alpha).$$

The measure on U is the probability measure defined by the graph; the measure on Σ is the counting measure. More generally, if Ω is a measure space, $L(\Omega)$ is the space of real-valued functions on Ω endowed with the inner product defined by the measure on Ω. The inner product induces a norm with $\|g\| = \langle g, g\rangle^{1/2}$ for $g \in L(\Omega)$.

We identify a projection game G with a linear operator from $L(V \times \Sigma)$ to $L(U \times \Sigma)$, defined by

$$Gf(u,\alpha) = \mathop{\mathbb{E}}_{(v,\pi)|u} \sum_{\beta\colon \alpha \hookleftarrow \beta} f(v,\beta)\,.$$

The bilinear form $\langle \cdot, G\cdot\rangle$ measures the value of assignments for Alice and Bob.

Claim 2.1. If Alice and Bob play the projection game G according to assignments f and g, then their success probability is equal to $\langle g, Gf\rangle$,

$$\langle g, Gf\rangle = \mathop{\mathbb{E}}_{(u,v,\pi)} \sum_{\alpha \hookleftarrow \beta} g(u,\alpha) f(v,\beta)\,.$$

This setup shows that the value of the game G is the maximum of the bilinear form $\langle g, Gf\rangle$ over assignments f and g. If we were to maximize the bilinear form over all functions with unit norm (instead of assignments), the maximum value would be the largest singular value of an associated matrix.

2.2 Playing games in parallel

Let G be a projection game with vertex sets U and V and alphabet Σ. Let H be a projection game with vertex sets U' and V' and alphabet Σ'. The *direct product* $G \otimes H$ is the following game with vertex sets $U \times U'$ and $V \times V'$ and alphabet $\Sigma \times \Sigma'$: The referee chooses (u, v, π) from G and (u', v', π') from H independently. The referee sends u, u' to Alice and v, v' to Bob. Alice answers α, α' and Bob answers β, β'. The players succeed if both $\alpha \hookleftarrow \beta$ and $\alpha' \hookleftarrow \beta'$. In linear algebra notation,

Claim 2.2. Given two games $G\colon L(V \times \Sigma) \to L(U \times \Sigma)$ and $H\colon L(V' \times \Sigma) \to L(U' \times \Sigma)$, the direct product game $G \otimes H\colon L(V \times V' \times \Sigma \times \Sigma') \to L(U \times U' \times \Sigma \times \Sigma')$ is given by the tensor of the two operators G and H. More explicitly, for any $f \in L(V \times V' \times \Sigma \times \Sigma')$, the operator for $G \otimes H$ acts as follows,

$$(G \otimes H) f(u, u', \alpha, \alpha') = \mathop{\mathbb{E}}_{(v,\pi)|u} \mathop{\mathbb{E}}_{(v',\pi')|u'} \sum_{\beta\colon (\alpha,\beta)\in\pi} \sum_{\beta'\colon (\alpha',\beta')\in\pi'} f(v, v', \beta, \beta')$$

$\square$

The notation $G^{\otimes k}$ is short for $G \otimes \cdots \otimes G$ (k times).

2.3 The collision value of a game

The collision value of a projection game G is a relaxation of the value of a game that is obtained by moving from G to a symmetrized version of it.[7] The advantage of the collision value is that it allows us to eliminate one of the players (Alice) in a simple way. Let the collision value of an assignment f for Bob be $\|Gf\| = \langle Gf, Gf\rangle^{1/2}$. We define the *collision value* of G to be

$$\|G\| = \max_f \|Gf\|$$

where the maximum is over all assignments $f \in L(V \times \Sigma)$.

[7] Such a transformation is well-known in e.g. the PCP literature, albeit implicitly. Moving to a symmetrized version occurs for example in low degree tests when when converting a line versus point test to a line versus line test.

The value $\|Gf\|^2$ can be interpreted as the success probability of the following process: Choose a random u and then choose independently $(v,\pi)|u$ and $(v',\pi')|u$; choose a random label β with probability $f(v,\beta)$ and a random label β' with probability $f(v',\beta')$, accept if there is some label α such that $\alpha \overset{\pi}{\hookleftarrow} \beta$ and $\alpha \overset{\pi'}{\hookleftarrow} \beta'$ (in this case β, β' 'collide').

The collision value of a projection game is quadratically related to its value. This claim is the only place where we use that the constraints are projections.

Claim 2.3. Let G be a projection game. Then $\mathrm{val}(G) \leqslant \|G\| \leqslant \mathrm{val}(G)^{1/2}$.

Proof. Let f, g be assignments attaining the value of G. The first inequality holds because

$$\mathrm{val}(G) = \langle g, Gf \rangle \leqslant \|g\| \cdot \|Gf\| \leqslant \|Gf\| \leqslant \|G\|,$$

where we used Cauchy–Schwarz followed by the bound $\|g\| \leqslant 1$ which holds for any assignment g. For the second inequality, let f be an assignment for G such that $\|Gf\| = \|G\|$. Then,

$$\|G\|^2 = \|Gf\|^2 = \langle Gf, Gf \rangle \leqslant \max_g \langle g, Gf \rangle = \mathrm{val}(G)$$

where the inequality used the fact that if f is an assignment then $\sum_\alpha Gf(u,\alpha) \leqslant 1$ for every u, so Gf can be turned into a proper assignment $g \in L(U \times \Sigma)$ by possibly increasing some of its entries. Thus, $\langle Gf, Gf \rangle \leqslant \langle g, Gf \rangle$ because all entries of these vectors are non-negative. □

The following claim says that the collision value cannot increase if we play, in parallel with G, another game H.

Claim 2.4. Let G, H be two projection games, then $\|G \otimes H\| \leqslant \|G\|$.

Proof. This claim is very intuitive and immediate for the standard value of a game, as it is always easier to play one game rather than two, and it is similarly proven here for the collision value. Given an assignment for $G \otimes H$ we show how to derive an assignment for G that has a collision value that is at least as high. Let G be a game on question sets U and V, and let H be a game on question sets U' and V'. Let $f \in L(V \times \Sigma \times V' \times \Sigma')$ be such that $\|(G \otimes H)f\| = \|G \otimes H\|$. For each v' we can define a strategy $f_{v'}$ for G by fixing v' and summing over β', i.e. $f_{v'}(v,\beta) := \sum_{\beta'} f(v,v',\beta,\beta')$. For each $u' \in U'$ let $f_{u'}$ be an assignment for G defined by $f_{u'}(v,\beta) := \mathbb{E}_{v'|u'} f_{v'}(v,\beta)$. In words, $f_{u'}$ is the strategy obtained by averaging over $f_{v'}$ for all neighbors v' of u'. We claim that $\|G\|^2 \geqslant \mathbb{E}_{u'}\|Gf_{u'}\|^2 \geqslant \|(G \otimes H)f\|^2$ where the second inequality comes from the following 'coupling' argument: Select a random vertex uu' and then two possible neighbors v_1v_1' and v_2v_2', and then two answers $\beta_1\beta_1'$ and $\beta_2\beta_2'$ according to f. With these random choices a collision for f is if $\beta_1\beta_1'$ is consistent with $\beta_2\beta_2'$. With the same random choices a collision for $f_{u'}$ is when β_1 is consistent with β_2, an easier requirement. □

Symmetrizing the Game. An additional way to view the collision value is as the value of a constraint satisfaction problem (CSP) that is obtained by symmetrizing the game.

Definition 2.5 (Symmetrized Constraint Graph)**.** For a projection game G given by distribution μ, we define its *symmetrized constraint graph* to be the weighted graph G_{sym} on vertex set V, described as a distribution μ_{sym} given by

- Select u at random (according to μ), and then select $(v,\pi)|u$ and independently $(v',\pi')|u$.

– Output (v, v', τ) where $\tau \subseteq \Sigma \times \Sigma$ consists of all pairs (β, β') such that there exists some α such that $\alpha \overset{\pi}{\hookleftarrow} \beta$ and $\alpha \overset{\pi'}{\hookleftarrow} \beta'$.

It is standard to define the *value* of a deterministic assignment $a : V \to \Sigma$ in a constraint graph by

$$val(G_{sym}, a) = \mathop{\mathbb{P}}_{(v,v',\tau)\sim\mu_{sym}} [(a(v), a(v')) \in \tau] .$$

We extend this to randomized assignments a, i.e. when $a(v)$ is a random variable taking values in Σ. In this case the value of a in G_{sym} is defined as the expectation over the values of a. If the randomized assignment is described by a vector $f \in \mathbb{R}_{\geqslant 0}^{V\times\Sigma}$ so that $a(v) = \beta$ with probability $f(v, \beta)$, one can check that the value of this randomized assignment is equal to

$$\mathop{\mathbb{E}}_{(v,v',\tau)\sim\mu_{sym}} \sum_{(\beta,\beta')\in\tau} f(v, \beta) f(v', \beta') = \|Gf\|^2 , \tag{2.1}$$

which is the square of the collision value. Thus, the value of the CSP described by G_{sym} (which is, as usual, the maximum value over all possible assignments) is equal to $\|G\|^2$.

We note that the symmetrized game $(G \otimes H)_{sym}$ can be obtained as the natural direct product of G_{sym} and H_{sym}, giving an alternative (intuitive) proof for Claim 2.4.

2.4 Expanding Games

A game G will be called *expanding* if the constraint graph of the symmetrized game G_{sym} is an expander.

Formally let A be defined by the matrix describing the random walk in the graph underlying G_{sym}, so $A_{v_1,v_2} = \mu_{sym}(v_2|v_1)$. In words, this is the probability of the next step in the random walk landing in v_2 conditioned on being in v_1. Observe that for any two vectors $x, y \in L(V)$, $\langle x, Ay\rangle = \langle Ax, y\rangle$ where the inner product is taken as usual with respect to the measure of V, which is also the stationary measure of A. This implies that A is diagonalizable with real eigenvalues $1 = \lambda_1 \geqslant \lambda_2 \geqslant \cdots \lambda_n > -1$. Since A is stochastic the top eigenvalue is 1 and the corresponding eigenvector is the all 1 vector. We define the spectral gap of the graph to be $1 - \max(|\lambda_2|, |\lambda_n|)$.

A game G is said to be *c-expanding* if the spectral gap of the Markov chain A corresponding to G_{sym} is at least c.

2.5 A family of trivial games

We will consider parameters of games (meant to approximate the game value) that compare the behavior of a game to the behavior of certain trivial games. As games they are not very interesting, but their importance will be for normalization.

Let T be the following projection game with the same vertex set V on both sides and alphabet Σ: The referee chooses a vertex v from V at random (according to the measure on V). The referee sends v to both Alice and Bob. The players succeed if Alice answers with $1 \in \Sigma$, regardless of Bob's answer. The operator of this game acts on $L(V \times \Sigma)$ as follows,

$$Tf(v, \alpha) = \begin{cases} \sum_\beta f(v, \beta) & \text{if } \alpha = 1, \\ 0 & \text{otherwise.} \end{cases}$$

We consider a related trivial game T_v with vertex sets $\{v\}$ and V and alphabet Σ. The operator maps $L(V \times \Sigma)$ to $L(\{v\} \times \Sigma)$ as follows: $T_v f(v, \alpha) = Tf(v, \alpha)$ for $\alpha \in \Sigma$. The operator T_v "removes" the part of f that assigns values to questions other than v.

While both T and T_v have the same value 1 for every assignment, they behave differently when considering a product game $G \otimes H$ and an assignment f for it. In particular, the norm of $(T \otimes H)f$ may differ significantly from the norms of $(T_v \otimes H)f$, and this difference will be important.

3 The Basic Approach

We will prove parallel repetition bounds for the collision value of projection games. Since the collision value is quadratically related to the usual value, due to Claim 2.3, these bounds imply the same parallel repetition bounds for the usual value (up to a factor of 2 in the number of repetitions). We state again our main theorem,

Theorem 1.8. *Any two projections games G and H satisfy $\|G \otimes H\| \leqslant \varphi(\|G\|) \cdot \|H\|$, where $\varphi(x) = \frac{2\sqrt{x}}{1+x}$.*

Before describing the proof, let us show how this theorem implies parallel repetition bounds. By repeated application of the theorem,

$$\|G^{\otimes k}\| \leqslant \varphi(\|G\|) \cdot \|G^{\otimes k-1}\| \leqslant \ldots \leqslant \varphi(\|G\|)^{k-1} \cdot \|G\| \leqslant \varphi(\|G\|)^k ,$$

On the interval $[0,1]$, the function φ satisfies $\varphi(x) \leqslant 2\sqrt{x}$ and $\varphi(1-\varepsilon) \leqslant 1 - \varepsilon^2/8$. The first bound implies a new parallel repetition bound for games with value close to 0. The second bound improves constant factors in the previous best bound for projection games [Rao08b].

3.1 Multiplicative game parameters

Let G be a projection game with vertex sets U and V and alphabet Σ. We are looking for a parameter ρ_G of G (namely, a function that assigns a non-negative value to a game) that approximates the value of the game and such that for every projection game H,

$$\|G \otimes H\| \leqslant \rho_G \cdot \|H\| . \tag{3.1}$$

Dividing both sides by $\|H\|$, we want a parameter of G that upper bounds the ratio $\|G \otimes H\|/\|H\|$ for all projection games H. The smallest such parameter is simply

$$\rho_G = \sup_H \frac{\|G \otimes H\|}{\|H\|} . \tag{3.2}$$

An intuitive interpretation of this value is that it is a kind of "parallel value" of G, in that it measures the relative decrease in value caused by playing G in parallel with any game H, compared to playing only H. Clearly $\|G\| \leqslant \rho_G$, but the question is whether $\rho_G \approx \|G\|$. We will show that this is the case through another game parameter, $\mathrm{val}_+(G)$, such that $\mathrm{val}_+(G) \geqslant \rho_G \geqslant \|G\|$ and such that $\mathrm{val}_+(G) \approx \|G\|$. First we introduce the game parameter $\lambda_+(G)$ which is not a good enough approximation for ρ_G (although it suffices when G is expanding), and then we refine it to obtain the final game parameter, $\mathrm{val}_+(G)$.

To lead up to the definition of $\lambda_+(G)$ and then $\mathrm{val}_+(G)$, let us assume that we know that $\rho_G > \rho$ for some fixed $\rho > 0$. This means that there is some specific projection game H for which $\frac{\|G \otimes H\|}{\|H\|} > \rho$. Let H be a projection game with vertex sets U' and V' and alphabet Σ', and let f be an optimal assignment for $G \otimes H$, so that $\|(G \otimes H)f\| = \|G \otimes H\|$. We can view f also as an assignment for the game $T \otimes H$, where T is the trivial game from Section 2.5.

The assignment f satisfies $\|(T \otimes H)f\| \leqslant \|T \otimes H\| \leqslant \|H\|$ (the second inequality is by Claim 2.4). So

$$\rho < \frac{\|G \otimes H\|}{\|H\|} \leqslant \frac{\|(G \otimes H)f\|}{\|(T \otimes H)f\|}\,.$$

We would like to "cancel out" H, so as to be left with a quantity that depends only on G and not on H. To this end, we consider the factorizations $G \otimes H = (G \otimes \mathrm{Id})(\mathrm{Id} \otimes H)$ and $T \otimes H = (T \otimes \mathrm{Id})(\mathrm{Id} \otimes H)$, where Id is the identity operator on the appropriate space. Intuitively, this factorization corresponds to a two step process of first applying the operator $\mathrm{Id} \otimes H$ on f to get h and then applying either $G \otimes \mathrm{Id}$ or $T \otimes \mathrm{Id}$.

$$\begin{tikzcd}
(G \otimes \mathrm{Id})f & & \\
 & h \arrow[ul, "G \otimes \mathrm{Id}"'] \arrow[dl, "T \otimes \mathrm{Id}"] & f \arrow[l, "\mathrm{Id} \otimes H"'] \\
(T \otimes \mathrm{Id})f & &
\end{tikzcd}$$

If we let $h = (\mathrm{Id} \otimes H)f$, then

$$\rho < \frac{\|G \otimes H\|}{\|H\|} \leqslant \frac{\|(G \otimes H)f\|}{\|(T \otimes H)f\|} = \frac{\|(G \otimes \mathrm{Id})h\|}{\|(T \otimes \mathrm{Id})h\|}\,. \tag{3.3}$$

By maximizing the right-most quantity in (3.3) over all nonnegative functions h, we get a value that does not depend on the game H so it can serve as a game parameter that is possibly easier to relate to the value of G algorithmically. Observe that there is a slight implicit dependence on (the dimensions of) H since the Id operator is defined on the same space as that of H. It turns out though that the extra dimensions here are unnecessary and the maximum is attained already for one dimensional functions h, so the identity operator Id can be removed altogether. This leads to the following simplified definition of a game parameter

Definition 3.1. For any projection game G let

$$\lambda_+(G) = \max_{h \geqslant 0} \frac{\|Gh\|}{\|Th\|}\,.$$

Theorem 3.2. *Any two projection games G and H satisfy $\|G \otimes H\| \leqslant \lambda_+(G) \cdot \|H\|$. Therefore, $\lambda_+(G) \geqslant \rho_G$.*

Before proving this, we mention again that in general $\lambda_+(G) \not\approx \|G\|$ which is why we later make a refined definition $\mathrm{val}_+(G)$.

Proof. By (3.3), there exists a non-negative function $h \in L(V \times \Sigma \times V' \times \Sigma')$ such that $\|G \otimes H\|/\|H\| \leqslant \|(G \otimes \mathrm{Id})h\|/\|(T \otimes \mathrm{Id})h\|$. We can view h as a matrix with each column belonging to $L(V \times \Sigma)$ - the input space for G and T. Then, $(G \otimes \mathrm{Id})h$ is the matrix obtained by applying G to the columns of the matrix h, and $(T \otimes \mathrm{Id})h$ is the matrix obtained by applying T to the columns of h. Next, we expand the squared norms of these matrices column-by-column,

$$\frac{\|(G \otimes \mathrm{Id})h\|^2}{\|(T \otimes \mathrm{Id})h\|^2} = \frac{\mathbb{E}_j \|Gh_j\|^2}{\mathbb{E}_j \|Th_j\|^2}$$

where j runs over the columns (this happens to be $j \in U' \times \Sigma'$ but it is not important at this stage). An averaging argument implies that there is some j^* for which $\frac{\|Gh_{j^*}\|^2}{\|Th_{j^*}\|^2}$ is at least as large as the ratio of the averages. Removing the squares from both sides, we get

$$\frac{\|G \otimes H\|}{\|H\|} \leqslant \frac{\|Gh_{j^*}\|}{\|Th_{j^*}\|} \leqslant \lambda_+(G) \qquad \square$$

In the next subsection we will show that for expanding games G, $\lambda_+(G) \approx \|G\|$, thus proving Theorem 1.8 for the special case of expanding games. For non-expanding games, $\lambda_+(G)$ is not a good approximation to $\|G\|$ and a more refined argument is called for as follows. Instead of comparing the value of $G \otimes H$ to $T \otimes H$, we compare it to the collection $T_v \otimes H$ for all $v \in V$ (defined in Section 2.5). We observe that the inequality (3.3) also holds with T replaced by T_v, so that for all $v \in V$

$$\rho < \frac{\|G \otimes H\|}{\|H\|} \leqslant \frac{\|(G \otimes H)f\|}{\|(T_v \otimes H)f\|} = \frac{\|(G \otimes \mathrm{Id})h\|}{\|(T_v \otimes \mathrm{Id})h\|} . \tag{3.4}$$

Now, by maximizing the right hand side over all measure spaces $L(\Omega)$ for Id and over all non-negative functions $h \in L(V \times \Sigma \times \Omega)$, we finally arrive at the game parameter

$$\mathrm{val}_+(G) \stackrel{\mathrm{def}}{=} \sup_{\Omega} \max_{h \geqslant 0} \frac{\|(G \otimes \mathrm{Id})h\|}{\max_v \|(T_v \otimes \mathrm{Id})h\|} , \tag{3.5}$$

where the Id operator is defined on the measure space $L(\Omega)$ and we are taking the supremum over all finite dimensional measure spaces Ω. It turns out that the supremum is attained for finite spaces—polynomial in the size of G.

Theorem 3.3. *Any two projection games G and H satisfy $\|G \otimes H\| \leqslant \mathrm{val}_+(G) \cdot \|H\|$.*

Proof. We essentially repeat the proof for $\lambda_+(\cdot)$. Let H be any projection game, and let f be an optimal assignment for $G \otimes H$. For every question v we compare $\|(G \otimes H)f\|$ to $\|(T_v \otimes H)f\|$ for T_v the trivial operator from Section 2.5. Since $\|(T_v \otimes H)f\| \leqslant \|T_v \otimes H\| \leqslant \|H\|$ we get,

$$\begin{aligned} \frac{\|G \otimes H\|}{\|H\|} \leqslant \frac{\|(G \otimes H)f\|}{\max_v \|(T_v \otimes H)f\|} &\leqslant \max_{f' \geqslant 0} \frac{\|(G \otimes H)f'\|}{\max_v \|(T_v \otimes H)f'\|} \\ &\leqslant \max_{h \geqslant 0} \frac{\|(G \otimes \mathrm{Id})h\|}{\max_v \|(T_v \otimes \mathrm{Id})h\|} = \mathrm{val}_+(G) . \end{aligned}$$

□

The advantage of val_+ over λ_+ will be seen in the next sections, when we show that this value is a good approximation of value of the game G.

3.2 Approximation bound for expanding projection games

In this section we show that if G is an expanding game then $\lambda_+(G) \approx \|G\|$. (Recall from Section 2.4 that a game is called γ-expanding if the graph underlying G_{sym} is has spectral gap at least γ.) By definition, $\|G\| \leqslant \lambda_+(G)$. The interesting direction is

Theorem 3.4. *Let G be a γ-expanding projection game for some $\gamma > 0$. Suppose $\lambda_+(G) > 1 - \varepsilon$. Then, $\|G\| > 1 - O(\varepsilon/\gamma)$.*

Proof. We may assume ε/γ is sufficiently small, say $\varepsilon/\gamma \leqslant 1/6$, for otherwise the theorem statement is trivially true. Let $f \in L(V \times \Sigma)$ be nonnegative, such that

$$\frac{\|Gf\|}{\|Tf\|} > 1 - \varepsilon ,$$

witnessing the fact that $\lambda_+(G) > 1 - \varepsilon$. First, we claim that without loss of generality we may assume that f is *deterministic*, i.e., for every vertex v, there is at most one label β

such that $f(v, \beta) > 0$. (This fact is related to the fact that randomized strategies can be converted to deterministic ones without decreasing the value.) We can write a general nonnegative function f as a convex combination of deterministic functions f' using the following sampling procedure: For every vertex v independently, choose a label β_v from the distribution that gives probability $\frac{f(v,\beta')}{\sum_\beta f(v,\beta)}$ to label $\beta' \in \Sigma$. Set $f'(v, \beta_v) = \sum_\beta f(v, \beta)$ and $f'(v, \beta') = 0$ for $\beta' \neq \beta_v$. The functions f' are deterministic and satisfy $\mathbb{E} f' = f$. By convexity of the function $f \mapsto \|Gf\|$ (being a linear map composed with a convex norm), we have $\mathbb{E}\|Gf'\| \geqslant \|G \mathbb{E} f'\| = \|Gf\|$. So, there must be some f' for which $\|Gf'\| \geqslant \|Gf\|$. By construction, $Tf(v) = \sum_\beta f(v, \beta) = Tf'(v)$, so $\frac{\|Gf'\|}{\|Tf'\|} \geqslant \frac{\|Gf\|}{\|Tf\|}$. (We remark that this derandomization step would fail without the premise that f is nonnegative.)

Thus we can assume that for every vertex v, there is some label β_v such that $f(v, \beta') = 0$ for all $\beta' \neq \beta_v$. We may also assume that $\|Tf\| = 1$ (because we can scale f). Since f is deterministic, we can simplify the quadratic form $\|Gf\|^2$,

$$(1 - \varepsilon)^2 \leqslant \|Gf\|^2 = \mathop{\mathbb{E}}_{(v,v')} f(v, \beta_v) f(v', \beta_{v'}) \cdot Q_{v,v'}, \tag{3.6}$$

where the pair (v, v') is distributed according to the edge distribution of the symmetrized game G_{sym} and $Q_{v,v'} = \mathbb{P}_{\tau|(v,v')}\{(\beta_v, \beta_{v'}) \in \tau\}$. Since we can view $b(v) := \beta_v$ as an assignment for G_{sym}, we can lower bound the value of the symmetrized game in terms of $Q_{v,v'}$,

$$\|G\|^2 = \mathrm{val}(G_{sym}) \geqslant \mathrm{val}(G_{sym}, b) = \mathop{\mathbb{E}}_{(v,v')} Q_{v,v'}\,. \tag{3.7}$$

To prove the theorem, we will argue that the right hand sides of (3.6) and (3.7) are close. The key step toward this goal is to show that $f(v, \beta_v) \approx 1$ for a typical vertex v. A priori $f(v, \beta_v)$ can be very high for some vertices, and very small for others, while maintaining $\mathbb{E}_v f(v, \beta_v)^2 = 1$. However, the expansion of G will rule this case out.

Denote $g(v) = f(v, \beta_v)$ and let A be the matrix corresponding to the Markov chain on the graph underlying G_{sym} and let $L = Id - A$. The smallest eigenvalue of L is 0 corresponding to the constant functions, and the second smallest eigenvalue is at least $\gamma > 0$ because of the expansion of G. Now, using inner products with respect to the natural measure on V,

$$\langle g, Lg \rangle = \mathop{\mathbb{E}}_{v} f(v, \beta_v)^2 - \mathop{\mathbb{E}}_{(v,v')} f(v, \beta_v) f(v', \beta_{v'}) \leqslant 1 - (1 - \varepsilon)^2 \leqslant 2\varepsilon\,.$$

On the other hand if we write $g = \bar{f} + g^\perp$ for $\bar{f} = \mathbb{E}_v f(v, \beta_v) = \mathbb{E}_v g(v)$, we have $\langle g^\perp, \bar{f} \rangle = 0$ and

$$\langle g, Lg \rangle = \langle \bar{f}, L\bar{f} \rangle + \langle g^\perp, Lg^\perp \rangle = 0 + \langle g^\perp, Lg^\perp \rangle \geqslant \gamma \|g^\perp\|^2\,.$$

Combining the above we get $\mathbb{E}_v (f(v, \beta_v) - \bar{f})^2 = \|g^\perp\|^2 \leqslant 2\varepsilon/\gamma$ which means that $g \approx \bar{f} \approx 1$. The following bound concludes the proof of theorem,

$$\begin{aligned}
1 - \|G\|^2 &\leqslant \mathop{\mathbb{E}}_{(v,v')} (1 - Q_{v,v'}) \quad \text{(using (3.7))} \\
&\leqslant \mathop{\mathbb{E}}_{(v,v')} (1 - Q_{v,v'}) \cdot 9\Big(\big(f(v, \beta_v) - \bar{f}\big)^2 + \big(f(v, \beta_v) - \bar{f}\big)^2 + f(v, \beta_v) f(v', \beta_{v'})\Big) \\
&\leqslant 36\varepsilon/\gamma + 9 \cdot \mathop{\mathbb{E}}_{(v,v')} (1 - Q_{v,v'}) \cdot f(v, \beta_v) f(v', \beta_{v'}) \\
&\leqslant 36\varepsilon/\gamma + 18\varepsilon\,.
\end{aligned}$$

The second step uses that all nonnegative numbers a and b satisfy the inequality $1 \leqslant 9ab + 9(a - \bar{f})^2 + 9(b - \bar{f})^2$ (using $\bar{f} \geqslant 2/3$). To verify this inequality we will do a case distinction based on whether a or b are smaller than $1/3$ or not. If one of a or b is smaller than 1/3, then one of the last two terms contributes at least 1 because $\bar{f} \geqslant 2/3$. On the other hand, if both a and b are at least 1/3, then the first term contributes at least 1. The third step uses the $f(v, \beta_v)$ is close to the constant function $\bar{f} \cdot \mathbb{1}$. The fourth step uses (3.6) and the fact that $\mathbb{E}_{(v,v')} f(v, \beta_v) f(v', \beta_{v'}) \leqslant \mathbb{E}_v f(v, \beta_v)^2 = 1$. $\square$

3.3 Short proof for the hardness of LABEL COVER

LABEL COVER$(1, \delta)$ is the gap problem of deciding if the value of a given projection game is 1 or at most δ. The results of this section suffice to give the following hardness of LABEL COVER, assuming the PCP theorem. This result is a starting point for many hardness-of-approximation results.

Theorem. LABEL COVER$(1, \delta)$ *is NP-hard for all* $\delta > 0$.

Let us sketch a proof of this. The PCP theorem [AS98, ALM+98] directly implies that LABEL COVER$(1, 1-\varepsilon)$ is NP-hard for some constant $\varepsilon > 0$. Let G be an instance of LABEL COVER$(1, 1-\varepsilon)$. We can assume wlog that G is expanding, see Claim A.1. We claim that $G^{\otimes k}$ for $k = O(\frac{\log 1/\delta}{\varepsilon})$ has the required properties. If $\mathrm{val}(G) = 1$ clearly $\mathrm{val}(G^{\otimes k}) = 1$. If $\mathrm{val}(G) < 1 - \varepsilon$, then

$$\mathrm{val}(G^{\otimes k}) \leqslant \|G^{\otimes k}\| \leqslant \lambda_+(G)^k \leqslant (1 - \Omega(\varepsilon))^k \leqslant \delta$$

where the second inequality is due to repeated applications of Theorem 3.2, and the third inequality is due to Theorem 3.4. $\square$

4 Approximation bound for general projection games

Theorem 4.1. *Let G be a game with* $\mathrm{val}_+(G)^2 > \rho$, *then*

$$\mathrm{val}(G) \geqslant \|G\|^2 > \frac{1 - \sqrt{1-\rho^2}}{1 + \sqrt{1-\rho^2}}\,.$$

Contrapositively, if $\|G\|^2 < \delta$, *then*

$$\mathrm{val}_+(G)^2 < \frac{2\sqrt{\delta}}{1+\delta}\,.$$

In particular, if $\|G\|^2$ *is small then the above bound becomes* $\mathrm{val}_+(G)^2 \leqslant 2\|G\|$; *and if* $\|G\|^2 < 1 - \varepsilon$ *then* $\mathrm{val}_+(G)^2 < 1 - \varepsilon^2/8$.

Let G be a projection game with vertex set U, V and alphabet Σ. Our assumption that $\mathrm{val}_+(G)^2 > \rho$ implies the existence of a measure space Ω and a non-negative function $f \in L(V \times \Sigma \times \Omega)$ such that

$$\|(G \otimes \mathrm{Id}_\Omega) f\|^2 > \rho \max_v \|(T_v \otimes \mathrm{Id}_\Omega) f\|^2$$

(In this section we denote by Id_Ω, rather than Id, the identity operator on the space $L(\Omega)$, to emphasize the space on which it is operating.)

Without loss of generality, by rescaling, assume that $f \leqslant 1$. Since the integration over Ω occurs on both sides of the inequality, we can rescale the measure on Ω without changing the inequality, so that $\max_v \|(T_v \otimes \mathrm{Id}_\Omega) f\| = 1$.

We also claim that for each $\omega \in \Omega$, the slice of f restricted to ω, can be assumed to be a "deterministic fractional assignment", i.e., that for every v, ω, there is at most one β such that $f(v, \beta, \omega) > 0$. The reason is, just as we have done in the proof of Theorem 3.4, that any slice f_ω of f can be written a convex combination of deterministic fractional assignments, $f_\omega = \mathbb{E} f'$ such that $T_v f' = T_v f_\omega$ for all v and all f'. By convexity $\mathbb{E}\|Gf'\| \geqslant \|G \mathbb{E} f'\| = \|G f_\omega\|$ so there must be some f' for which $\|Gf'\|$ is as high as the average.

Proof Overview. The proof is by an algorithm that extracts from f an assignment for G. We have seen that for expanding games G there is always a single element $\omega \in \Omega$ such that the slice f_ω of f restricted to ω is already "good enough", in that it can be used to derive a good assignment for G (see Theorem 3.4). When G is not expanding this is not true because each f_ω can potentially concentrate its mass on a different part of G. For example, imagine that G is made of two equal-sized but disconnected sub-games G_1 and G_2, with optimal assignments f_1, f_2 respectively, and let f be a vector assignment defined as follows. For every v, β we will set $f(v, \beta, \omega_1) = f_1(v, \beta)$ if v is in G_1 and $f(v, \beta, \omega_2) = f_2(v, \beta)$ if v is in G_2. Everywhere else we set f to 0. The quality of f will be proportional to the average quality of f_1, f_2, yet there is no single ω from which an assignment for G can be derived. Our algorithm, therefore, will have to construct an assignment for G by combining between the assignments derived from different components (ω's) of f.

Conceptually the algorithm has two steps. In the first step we convert each f_ω to a partial assignment, namely, a 0/1 function that might be non-zero only on a small portion of G. This is done by randomized rounding. We use Cheeger-type arguments to show that on average over ω the quality of this partial assignment is not too far from ρ. This is done in Lemma 4.2 and is analogous to but more subtle than Theorem 3.4.

The second step, done in Lemma 4.3, is to combine the different partial assignments into one global assignment. For this step to work we must ensure that the different partial assignments cover the entire game in a uniform way. Otherwise one may worry that all of the partial assignments are concentrated on the same small part of G. Indeed, that would have been a problem had we defined $\mathrm{val}_+(G)$ to be $\sup_{\Omega, f} \frac{\|(G \otimes \mathrm{Id}_\Omega) f\|^2}{\|(T \otimes \mathrm{Id}_\Omega) f\|^2}$. Instead, the denominator in the definition of val_+ is the maximum over v of $\|(T_v \otimes \mathrm{Id}_\Omega) f\|^2$. This essentially forces sufficient mass to be placed on each vertex v in G, so G is uniformly covered by the collection of partial assignments. The second step is known as correlated sampling, introduced to this context in [Hol09], because when viewed as a protocol between two players, the two players will choose ω using shared randomness, and then each player will answer his question according to the partial assignment derived from f_ω. (To be more accurate, shared randomness is also used in the first step, for deciding the rounding threshold through which a partial assignment is derived).

Lemma 4.2 (Threshold Rounding)**.** *There exists a* 0/1*-valued function* $f' \in L(V \times \Sigma \times \Omega \times [0,1])$*, such that for* $\psi(x) = 1 - (1 - x^2)^{1/2}$*,*

$$\|(G \otimes \mathrm{Id}_\Omega \otimes \mathrm{Id}_{[0,1]}) f'\|^2 > \psi(\rho)$$

and yet

$$\max_v \|(T_v \otimes \mathrm{Id}_\Omega \otimes \mathrm{Id}_{[0,1]}) f'\| \leqslant 1\,.$$

Proof. Choose $f' : (V \times \Sigma) \times \Omega \times [0,1]$ as

$$f'(v, \beta, \omega, \tau) = \begin{cases} 1 & \text{if } f(v, \beta, \omega)^2 > \tau\,, \\ 0 & \text{otherwise.} \end{cases}$$

Notation: We will use shorthand $f'_{\omega,\tau}$ to mean the function in $L(V \times \Sigma)$ which is the slice of f' defined by $f'_{\omega,\tau}(v,\beta) = f'(v,\beta,\omega,\tau)$ and similarly $f'_{v,\beta}$ is a function of τ and ω.

Since the slices f_ω are deterministic fractional assignments, the slices $f'_{\omega,\tau}$ are also deterministic fractional assignments. (Actually, they are 0/1-valued partial assignments.) First observe that for every v,β,ω, we have $\mathbb{E}_\tau f'(v,\beta,\omega,\tau)^2 = f(v,\beta,\omega)^2$. Thus,

$$\|(T_v \otimes \mathrm{Id}_{\Omega\times[0,1]})f'\|^2 = \sum_\beta \|f'_{v,\beta}\|^2 = \sum_\beta \|f_{v,\beta}\|^2 = \|(T_v \otimes \mathrm{Id}_\Omega)f\|^2 \leqslant 1\,,$$

where the first and third equalities hold because f (and f') assign a non-zero value to at most one β for every v,ω (and τ). It remains to show that

$$\|(G \otimes \mathrm{Id}_\Omega \otimes \mathrm{Id}_{[0,1]})f'\|^2 = \mathop{\mathbb{E}}_{\tau\sim[0,1]} \mathop{\mathbb{E}}_{\omega\sim\Omega} \|f'_{\tau,\omega}\|^2 > \psi(\rho)\,.$$

Let $\rho_\omega \in [0,1]$ be such that $\|Gf_\omega\|^2 = \rho_\omega\|f_\omega\|^2$. We will first show that for all $\omega \in \Omega$,

$$\mathop{\mathbb{E}}_{\tau\sim[0,1]} \|Gf'_{\omega,\tau}\|^2 \geqslant \psi(\rho_\omega)\|f_\omega\|^2 \tag{4.1}$$

We will deduce the desired bound on $\|(G \otimes \mathrm{Id}_{\Omega\times[0,1]})f'\|$ by integrating (4.1) over Ω and using the convexity of ψ on $[0,1]$.

Let $x_\omega \in \Sigma^V$ be an assignment that is consistent with the fractional assignment f_ω (so that $f_\omega(v,\beta) = 0$ for all $\beta \neq x_{\omega,v}$). Note that also $f'_{\omega,\tau}(v,\beta) = 0$ whenever $\beta \neq x_{\omega,v}$, so

$$\begin{aligned}\|Gf'_{\omega,\tau}\|^2 &= \mathop{\mathbb{E}}_u \sum_\alpha \mathop{\mathbb{E}}_{(v_1,\pi_1),(v_2,\pi_2)|u} \sum_{\alpha \overset{\pi_1}{\leftarrow} \beta_1, \alpha \overset{\pi_2}{\leftarrow} \beta_2} f'_{\omega,\tau}(v_1,\beta_1) \cdot f'_{\omega,\tau}(v_2,\beta_2)\,, \\ &= \mathop{\mathbb{E}}_{(v_1,v_2,\pi)} \pi(x_{\omega,v_1}, x_{\omega,v_2}) f'_{\omega,\tau}(v_1, x_{\omega,v_1}) \cdot f'_{\omega,\tau}(v_2, x_{\omega,v_2}) \end{aligned} \tag{4.2}$$

where in the remainder of this proof (v_1,v_2,π) will always be distributed by choosing u at random and then $(v_1,\pi_1),(v_2,\pi_2)|u$; and the constraint $\pi : \Sigma\times\Sigma \to \{0,1\}$ is 1 on pairs β_1,β_2 for which there is some α such that $\alpha \overset{\pi_1}{\leftarrow} \beta_1$ and $\alpha \overset{\pi_2}{\leftarrow} \beta_2$.

For every pair v_1,β_1 and v_2,β_2,

$$\mathop{\mathbb{E}}_{\tau\sim[0,1]} f'_{\omega,\tau}(v_1,\beta_1) \cdot f'_{\omega,\tau}(v_2,\beta_2) = \min\left\{f_\omega(v_1,\beta_1)^2, f_\omega(v_2,\beta_2)^2\right\} \tag{4.3}$$

Now, combining (4.2) and (4.3),

$$\mathop{\mathbb{E}}_\tau \|Gf'_{\omega,\tau}\|^2 = \mathop{\mathbb{E}}_{(v_1,v_2,\pi)} \pi(x_{\omega,v_1}, x_{\omega,v_2}) \cdot \min\left\{f_\omega(v_1,x_{\omega,v_1})^2, f_\omega(v_2,x_{\omega,v_2})^2\right\}\,. \tag{4.4}$$

Again using that f_ω is a fractional assignment, we can express $\|Gf_\omega\|^2$ in a similar way,

$$\|Gf_\omega\|^2 = \mathop{\mathbb{E}}_{(v_1,v_2,\pi)} \pi(x_{\omega,v_1}, x_{\omega,v_2}) \cdot f_\omega(v_1,x_{\omega,v_1}) f_\omega(v_2,x_{\omega,v_2}) \tag{4.5}$$

At this point we will use the following simple inequality (see Corollary 4.5 toward the end of this subsection):

> *Let A,B,Z be jointly-distributed random variables, such that A,B take only nonnegative values and Z is 0/1-valued. Then, $\mathbb{E}\, Z \min\{A,B\} \geqslant \psi(\rho_\omega)\, \mathbb{E}\, \frac{1}{2}(A+B)$ holds as long as $\mathbb{E}\, Z\sqrt{AB} \geqslant \rho_\omega\, \mathbb{E}\, \frac{1}{2}(A+B)$.*

We will instantiate the inequality with $Z = \pi(x_{\omega,v_1}, x_{\omega,v_2})$, $A = f_\omega(v_1, x_{v_1})^2$, and $B = f_\omega(v_2, x_{v_2})^2$ (with (v_1, v_2, π) drawn as above). With this setup, $\mathbb{E}\,\frac{1}{2}(A+B) = \|f_\omega\|^2$. Furthermore, $\mathbb{E}\,Z\min\{A,B\}$ corresponds to the right-hand side of (4.4) and $\mathbb{E}\,Z\sqrt{AB}$ corresponds to the right-hand side of (4.5). Thus, $\|Gf_\omega\|^2 \geqslant \rho_\omega\|f_\omega\|^2$ means that the condition above is satisfied, and we get the desired conclusion, $\mathbb{E}_\tau\|Gf'_{\omega,\tau}\|^2 \geqslant \psi(\rho_\omega)\|f_\omega\|^2$ as required in (4.1).

Finally,

$$\|(G \otimes I_{\Omega\times[0,1]})f'\|^2 = \int_\Omega \mathop{\mathbb{E}}_\tau \|Gf'_{\omega,\tau}\|^2\, d\omega \geqslant \int_\Omega \psi(\rho_\omega)\cdot\|f_\omega\|^2\, d\omega \geqslant \psi(\rho).$$

For the last step, we use the convexity of ψ and that $\rho = \int_\Omega \rho_\omega\|f_\omega\|^2\, d\omega$ and $\|f\|^2 = \int_\Omega \|f_\omega\|^2\, d\omega \leqslant 1$. □

Let $\Omega' = \Omega \times [0,1]$. We now have a 0/1 function $f' \in L(V \times \Sigma \times \Omega')$ such that

$$\|(G\otimes \mathrm{Id}_{\Omega'})f'\|^2 \geqslant \psi(\rho) \qquad \text{and} \qquad \max_v \|(T_v \otimes \mathrm{Id}_{\Omega'})f'\|^2 = 1\,.$$

The next step is to combine the different components of f' into one assignment for G.

Lemma 4.3 (Correlated Sampling)**.** *There exists an assignment x for G with value at least $\frac{1-\gamma}{1+\gamma}$ for $1-\gamma = \|(G\otimes \mathrm{Id}_{\Omega'})f'\|^2 \geqslant \psi(\rho)$.*

Proof. We may assume $\|(T_v\otimes \mathrm{Id})f'\|^2 = 1$ for all $v \in V$. (We can arrange this condition to hold by adding additional points to Ω', one for each vertex in V, and extending f' in a suitable way.)

Rescale Ω' to a probability measure. Let λ be the scaling factor, so that $\|(G\otimes I_{\Omega'})f'\|^2 = (1-\gamma)\lambda$ (after rescaling). Then, f' also satisfies $\|(T_v\otimes \mathrm{Id})f'\|^2 = \lambda$ for all $v \in V$.

For every $\omega \in \Omega'$, the slice f'_ω is a partial assignment (in the sense that it uniquely assigns a label to a subset of vertices). We will construct (jointly-distributed) random variables $\{X_v\}_{v\in V}$, taking values in Σ, by combining the partial assignments f'_ω in a probabilistic way. (The sampling procedure we describe corresponds to correlated sampling applied to the distributions $\Omega'_v = \{\omega \in \Omega' \mid f'_\omega(v,\beta) > 1\}$.)

Let $\{\omega(n)\}_{n\in\mathbb{N}}$ be an infinite sequence of independent samples from Ω' and let $f'^{(n)} = f'_{\omega_n}$ be the corresponding slices of f'. Let $R(v)$ be the smallest number r such that the partial assignment $f'_{\omega(r)}$ assigns a label to v, i.e. such that $f'_{\omega(r)}(v,\beta) > 0$ for some (unique) $\beta \in \Sigma$. We define X_v to be the label assigned to v by this procedure, so that $f'_{\omega(R(v))}(v, X_v) = 1$.

In this randomized process, the probability that the random assignment X satisfies a constraint (v_1, v_2, π) is bounded from below by the probability that $R(v_1) = R(v_2)$ and the partial assignment $f'_{\omega(R(v_1))} = f'_{\omega(R(v_2))}$ assigns consistent values to v_1, v_2. The probability of this event is equal to the probability that the partial assignment f'_ω satisfies the constraint (v_1, v_2, π) conditioned on the event that f'_ω assigns a label to either v_1 or v_2. Therefore,

$$\begin{aligned}
\mathop{\mathbb{P}}_X\{\pi(X_{v_1}, X_{v_2}) = 1\} &\geqslant \mathop{\mathbb{P}}_X\{\pi(X_{v_1}, X_{v_2}) = 1 \wedge R(v_1) = R(v_2)\} \\
&= \mathop{\mathbb{P}}_{\omega\sim\Omega'}\{\pi(x_{\omega,v_1}, x_{\omega,v_2}) = 1 \mid f_\omega(v_1, x_{\omega,v_1}) = 1 \vee f_\omega(v_2, x_{\omega,v_2}) = 1\} \\
&= \frac{\mathbb{E}_{\omega\sim\Omega'}\,\pi(x_{\omega,v_1}, x_{\omega,v_2})\min\{f'_\omega(v_1, x_{\omega,v_1}), f'_\omega(v_2, x_{\omega,v_2})\}}{\mathbb{E}_{\omega\sim\Omega'}\max\{f'_\omega(v_1, x_{\omega,v_1}), f'_\omega(v_2, x_{\omega,v_2})\}}\,. \qquad (4.6)
\end{aligned}$$

(Here, $x_\omega \in \Sigma^W$ denotes any assignment consistent with the partial assignment f'_ω for $\omega \in \Omega'$.) At this point, we will use the following simple inequality (see Lemma 4.6 toward the end of this subsection):

Let A, B, Z be jointly-distributed random variables such that A, B are nonnegative-valued and Z is $0/1$-valued. Then, $\mathbb{E}\, Z \cdot \min\{A, B\} \geqslant (\frac{1-\gamma'}{1+\gamma'})\,\mathbb{E}\max\{A, B\}$ as long as $\mathbb{E}\, Z \cdot \min\{A, B\} \geqslant (1-\gamma')\,\mathbb{E}\,\frac{1}{2}(A+B)$.

We instantiate this inequality with $A = f'_\omega(v_1, x_{\omega,v_1})$, $B = f'_\omega(v_2, x_{\omega,v_2})$, $Z = \pi(x_{\omega,v_1}, x_{\omega,v_2})$ for $\omega \sim \Omega'$, and $\gamma' = \gamma_{v_1,v_2,\pi}$, where

$$\mathop{\mathbb{E}}_{\omega} \pi(x_{\omega,v_1}, x_{\omega,v_2}) f'_\omega(v_1, x_{\omega,v_1}) f'_\omega(v_2, x_{\omega,v_2}) = (1-\gamma_{v_1,v_2,\pi})\lambda.$$

The condition of the A, B, Z-inequality corresponds to the condition $\mathbb{E}_\omega\, \pi(x_{\omega,v_1}, x_{\omega,v_2}) \min\{f'_\omega(v_1, x_{\omega,v_1}), f'_\omega(v_2, x_{\omega,v_2})\} \geqslant (1-\gamma_{v_1,v_2,\pi})\lambda$, because f' is $0/1$-valued and $\mathbb{E}_\omega\, f'_\omega(v, x_{\omega,v})^2 = \|(T \otimes \mathrm{Id}) f'\|^2 = \lambda$. The conclusion of the A, B, Z-inequality shows that the right-hand side of (4.6) is bounded from below by $1-\gamma_{v_1,v_2,\pi}/1+\gamma_{v_1,v_2,\pi}$.

Hence, by convexity of $\varphi'(x) = (1-x)/(1+x)$,

$$\mathop{\mathbb{E}}_{X} \mathrm{val}(G; X) = \mathop{\mathbb{E}}_{(v_1,v_2,\pi)\sim G} \mathop{\mathbb{P}}_{X}\{\pi(X_{v_1}, X_{v_2}) = 1\} \geqslant \mathop{\mathbb{E}}_{(v_1,v_2,\pi)\sim G} \varphi'(\gamma_{v_1,v_2,\pi}) \geqslant \varphi'\Big(\mathop{\mathbb{E}}_{(v_1,v_2,\pi)\sim G} \gamma_{v_1,v_2,\pi}\Big).$$

It remains to lowerbound the expectation of $\gamma_{v_1,v_2,\pi}$ over $(v_1, v_2, \pi) \sim G$.

$$\begin{aligned}\mathop{\mathbb{E}}_{(v_1,v_2,\pi)\sim G} 1-\gamma_{v_1,v_2,\pi} &= \mathop{\mathbb{E}}_{(v_1,v_2,\pi)\sim G} \tfrac{1}{\lambda} \mathop{\mathbb{E}}_{\omega} \pi(x_{\omega,v_1}, x_{\omega,v_2}) f'_\omega(v_1, x_{\omega,v_1}) f'_\omega(v_2, x_{\omega,v_2}) \\ &= \tfrac{1}{\lambda}\|(G \otimes I_{\Omega'}) f'\|^2 = 1-\gamma\,. \end{aligned}$$

□

Some inequalities. The following lemma show that if the expected geometric average of two random variable is close to their expected arithmetic average, then the expected minimum of the two variables is also close to the expected arithmetic average. A similar lemma is also used in proofs of Cheeger's inequality.

Lemma 4.4. *Let A, B be jointly-distributed random variables, taking nonnegative values. If $\mathbb{E}\sqrt{AB} = \rho\,\mathbb{E}\,\frac{1}{2}(A+B)$, then*

$$\mathbb{E}\min\{A, B\} \geqslant \varphi(\rho) \cdot \mathbb{E}\,\tfrac{1}{2}(A+B)\,,$$

for $\varphi(x) = 1 - \sqrt{1-x^2}$.

Proof. Since $\min\{A, B\} = \frac{1}{2}(A+B) - \frac{1}{2}|A-B|$, it is enough to lower bound

$$\begin{aligned}\mathbb{E}\,\tfrac{1}{2}|A-B| &= \mathbb{E}\,\tfrac{1}{2}\left|A^{1/2}-B^{1/2}\right|\left|A^{1/2}+B^{1/2}\right| \\ &\leqslant \Big(\mathbb{E}\,\tfrac{1}{2}\left(A^{1/2}-B^{1/2}\right)^2 \cdot \mathbb{E}\,\tfrac{1}{2}\left(A^{1/2}+B^{1/2}\right)^2\Big)^{1/2} \\ &= \Big(\Big(\mathbb{E}\,\tfrac{1}{2}(A+B) - \mathbb{E}\sqrt{AB}\Big)\cdot\Big(\mathbb{E}\,\tfrac{1}{2}(A+B) + \mathbb{E}\sqrt{AB}\Big)\Big)^{1/2} \\ &= \sqrt{1-\rho^2} \cdot \mathbb{E}\,\tfrac{1}{2}(A+B)\,.\end{aligned}$$

The second step uses Cauchy–Schwarz. □

The following corollary will be useful for us to prove Cheeger's inequality for two-player games.

Corollary 4.5. *Let A, B be as before. Let Z be a $0/1$-valued random variable, jointly distributed with A and B. If $\mathbb{E}\, Z \cdot \sqrt{AB} = \rho\,\mathbb{E}\,\frac{1}{2}(A+B)$, then $\mathbb{E}\, Z \cdot \min\{A, B\} \geqslant \varphi(\rho)\,\mathbb{E}\,\frac{1}{2}(A+B)$, for φ as before.*

Proof. The corollary follows from the convexity of φ. For notational simplicity, assume $\mathbb{E}\,\frac{1}{2}(A+B) = 1$ (by scaling). Let $\lambda = \mathbb{E}\,Z \cdot \frac{1}{2}(A+B)$. Write ρ as a convex combination of 0 and a number ρ' such that $\rho = \rho' \cdot \lambda + 0 \cdot (1-\lambda)$. Since $\mathbb{E}\,Z\sqrt{AB} = \rho' \cdot \lambda$, Lemma 4.4 implies $\mathbb{E}\,Z \min\{A,B\} \geqslant \varphi(\rho') \cdot \lambda$. The convexity of φ implies $\varphi(\rho) \leqslant \lambda \cdot \varphi(\rho') + (1-\lambda)\varphi(0) = \lambda \cdot \varphi(\rho')$. □

Lemma 4.6. *Let A, B, Z be jointly-distributed random variables as before (A, B taking nonnegative values and Z taking $0/1$ values). If $\mathbb{E}\,Z \cdot \min\{A,B\} = (1-\gamma)\,\mathbb{E}\,\frac{1}{2}(A+B)$, then*

$$\frac{\mathbb{E}\,Z \cdot \min\{A,B\}}{\mathbb{E}\max\{A,B\}} \geqslant \frac{1-\gamma}{1+\gamma}\,.$$

Proof. For simplicity, assume $\mathbb{E}\,\frac{1}{2}(A+B) = 1$ (by scaling). Since $\min\{A,B\} = \frac{1}{2}(A+B) - \frac{1}{2}|A-B|$, we get $1-\gamma \leqslant \mathbb{E}\min\{A,B\} = 1 - \mathbb{E}\,\frac{1}{2}|A-B|$, which means that $\mathbb{E}\,\frac{1}{2}|A-B| \leqslant \gamma$. Since $\max\{A,B\} = \frac{1}{2}(A+B) + \frac{1}{2}|A-B|$, it follows that

$$\frac{\mathbb{E}\,Z \cdot \min\{A,B\}}{\mathbb{E}\max\{A,B\}} = \frac{1-\gamma}{1 + \mathbb{E}\,\frac{1}{2}|A-B|} \geqslant \frac{1-\gamma}{1+\gamma}\,.$$

□

4.1 Parallel repetition bound for general projection games

Let us now prove Theorem 1.1 and Corollary 1.3 and Corollary 1.2 together:

Theorem 4.7. *Let G be a projection game with $\mathrm{val}(G) \leqslant \delta$. Then,*

$$\mathrm{val}(G^{\otimes k}) \leqslant \left(\frac{2\delta^{1/2}}{1+\delta}\right)^{k/2}.$$

In particular, $\mathrm{val}(G^{\otimes k}) \leqslant (1-\varepsilon^2/16)^k$ for $\delta = 1-\varepsilon$ and $\mathrm{val}(G^{\otimes k}) \leqslant (4\delta)^{k/4}$.

Proof. By applying Theorem 3.3 repeatedly, $\mathrm{val}(G^{\otimes k}) \leqslant \|G^{\otimes k}\| \leqslant \mathrm{val}_+(G)^k$. The game G satisfies $\|G\|^2 \leqslant \mathrm{val}(G) \leqslant \delta$. Thus, by Theorem 4.1, $\mathrm{val}+(G)^2 \leqslant \frac{2\delta^{1/2}}{1+\delta}$. Therefore, $\mathrm{val}(G) \leqslant (\frac{2\delta^{1/2}}{1+\delta})^{k/2}$ □

5 Few Repetitions — Proof of Theorem 1.7

The following theorem will allow us to prove a tight bound on the value of projection games after few parallel-repetitions (Theorem 1.7).

Theorem 5.1. *Let G and H be two projection games. Suppose $\|H\|^2 \leqslant 1-\gamma$ but $\|G \otimes H\|^2 \geqslant 1-\eta-\gamma$ (with γ small enough, say $\gamma \leqslant 1/100$). Then,*

$$\|G\|^2 \geqslant 1 - O\Big(\eta + \sqrt{\gamma\eta}\Big).$$

Let us first explain how the bound above improves over the bounds in the previous sections. In the notation of the above theorem, Theorem 3.3 implies that $\mathrm{val}_+(G)^2 = \frac{\|G \otimes H\|^2}{\|H\|^2} \geqslant 1 - O(\eta)$. Thus, $\|G\|^2 \geqslant 1 - O(\sqrt{\eta})$ by Theorem 4.1. We see that this bound is worse than the above bound whenever γ is close to 0.

Before proving the theorem, we will show how it implies Theorem 1.7 (parallel-repetition bound for few repetitions).

Proof of Theorem 1.7. Let us first reformulate Theorem 5.1. (The above formulation reflects our proof strategy for Theorem 5.1 but for our current application, the following formulation is more convenient.) Let G and H be two projection games. Suppose $\|G\|^2 \leqslant 1-\varepsilon$ and $\|H\|^2 \leqslant 1 - t\varepsilon$ for some $\varepsilon > 0$ and $1 \leqslant t \ll 1/\varepsilon$. Then, Theorem 5.1 shows that

$$\|G \otimes H\|^2 \leqslant 1 - \left(t + \Omega\left(\tfrac{1}{t}\right)\right) \cdot \varepsilon$$

(Here, we use that for $\gamma = t\varepsilon$ and $\eta = \Omega(1/t)\varepsilon$, we have $\eta + \sqrt{\eta\gamma} = \Omega(\varepsilon)$.) From this bound, we can prove by induction on k that $\|G^{\otimes k}\|^2 \leqslant 1 - \Omega(k)^{1/2} \cdot \varepsilon$ for $k \ll 1/\varepsilon^2$. Concretely, let $\{t(k)\}_{k\in\mathbb{N}}$ be the sequence such that $\|G^{\otimes k}\|^2 = 1 - t(k)\cdot\varepsilon$. The above formulation of Theorem 5.1 implies $t(k+1) \geqslant t(k) + \Omega(1/t(k))$ (as long as $t(k) \ll 1/\varepsilon$). Since the sequence increases monotonically, it follows that $t(k+1)^2 \geqslant t(k)^2 + \Omega(1)$ (by multiplying the recurrence relation with $t(k+1)$ on both sides). Hence, as long as $t(k) \ll 1/\varepsilon$, we have $t(k)^2 = \Omega(k)$ and $t(k) = \Omega(k)^{1/2}$, as desired. □

The proof of Theorem 5.1 follows a similar structure as the proofs in Section 3 and Section 4. If $\|H\|^2 \leqslant 1-\gamma$ and $\|G\|^2 \geqslant 1-\eta-\gamma$, then following the proof of Theorem 3.3 we can construct a function for G that satisfies certain properties. (In particular, the function certifies $\mathrm{val}_+(G)^2 \geqslant 1 - O(\eta)$). The challenge is to construct an assignment for G from this function. This construction is the content of the following two lemmas, Lemma 5.2 and Lemma 5.3.

In the following G will be the "square" of a projection game G_0 (formally, the operator $G = G_0^T G_0$). Let $\mathtt{Assign}_{W,\Sigma}(\Omega)$ be the set of nonnegative functions $f \in L(W \times \Sigma \times \Omega)$ such that every slice f_ω is deterministic, i.e., for every $v \in W$, there exists at most one label β with $f_\omega(v,\beta) > 0$. We write $f_{v,*}$ to denote the function $(T_v \otimes \mathrm{Id}_\Omega) f$.

Lemma 5.2. *Let $f \in \mathtt{Assign}_{W,\Sigma}(\Omega)$ be a $[0,1]$-valued function with $\|f_{u,*}\|^2 \leqslant 1-\gamma$ for all $u \in W$ and $\|f\|_1 = 1$. Suppose every slice of f is a deterministic fractional assignment. Suppose $\langle f, (G \otimes I_\Omega) f\rangle \geqslant 1 - \eta - \gamma$ and that γ is small enough (say, $\gamma < 1/100$). Then, there exists a $0/1$-valued function $f' \in \mathtt{Assign}_{W,\Sigma}(\Omega')$ with*

$$\langle f', (G \otimes I_{\Omega'}) f'\rangle \geqslant \left(1 - O(\eta + \sqrt{\eta\gamma})\right)\|f'\|^2\,,$$

and $\|f'_{u,*}\|^2 \geqslant 1/3$ for all but an $O(\eta)$ fraction of vertices in W.

Proof. Since $\|f_{u,*}\|^2 \leqslant 1-\gamma$ for all $u \in W$, the condition $\langle f, (G\otimes I_\Omega) f\rangle \geqslant 1-\eta-\gamma$ implies that $\|f_{u,*}\|^2 \geqslant 0.9$ for all but $O(\eta)$ fraction of vertices $u \in W$.

Let $\Omega' = \Omega \times [1/10, 9/10]$ (with the uniform measure on the interval $[1/10, 9/10]$). Let $f' \colon W \times \Sigma \times \Omega' \to \{0,1\}$ be such that for all $u \in W$, $\alpha \in \Sigma$, $\omega \in \Omega$, $\tau \in [1/10, 9/10]$,

$$f'_{\omega,\tau}(u,\alpha) = \begin{cases} 1 & \text{if } f_\omega(u,\alpha)^2 > \tau\,, \\ 0 & \text{otherwise.} \end{cases}$$

It follows that $\|f'_{u,*}\|^2 \geqslant 1/3$ for all but $O(\eta)$ vertices $u \in W$ (using that $\|f_{u,*}\|^2 \geqslant 0.9$ for all but $O(\eta)$ vertices). We see that $f_\omega(u,\alpha)^2 - 1/10 \leqslant \mathbb{E}_\tau f'_{\omega,\tau}(u,\alpha)^2 \leqslant 10/9\, f_\omega(u,\alpha)^2$. Let $B_\omega \colon W \times \Sigma \to \mathbb{R}$ be the $0/1$ indicator function of the event $\{f(u,\alpha)^2 \in [1/10, 9/10]\}$. Then,

$$\begin{aligned} \mathop{\mathbb{E}}_{\tau} \left(f'_{\omega,\tau}(u,\alpha) - f'_{\omega,\tau}(u',\alpha')\right)^2 &\leqslant 2 \cdot \left(f_\omega(u,\alpha) - f_\omega(u',\alpha')\right)^2 \\ &\quad + 2\left(B_\omega(u,\alpha) + B_\omega(u',\alpha')\right) \cdot \left|f_\omega(u,\alpha)^2 - f_\omega(u',\alpha')^2\right| \qquad (5.1) \end{aligned}$$

We can see that the left-hand side is always at most $2|f_\omega(u,\alpha)^2 - f_\omega(u',\alpha')^2|$. Hence, the inequality holds if $B_\omega(u,\alpha) + B_\omega(u',\alpha') \geqslant 1$. Otherwise, if $B_\omega(u,\alpha) + B_\omega(u',\alpha') = 0$, then the left-hand side is either 0 or 1. In both cases, the left-hand side is bounded by $2(f_\omega(u,\alpha) - f_\omega(u',\alpha'))^2$.

For every $\omega \in \Omega$, let $x_\omega \in \Sigma^V$ be an assignment consistent with the fractional assignment f_ω. Let $h_\omega \colon W \to [0,1]$ be the function $h_\omega(u) = f_\omega(u, x_{\omega,u})$. Similarly, $h'_{\omega,\tau}(u) = f'_{\omega,\tau}(u, x_{\omega,u})$. Let G_ω be the linear operator on functions on W, so that for every $g \colon W \to \mathbb{R}$

$$\langle g, G_\omega g\rangle = \mathop{\mathbb{E}}_{(u,v,\pi)\sim G} \pi(x_{\omega,u}, x_{\omega,v}) \cdot g(u) \cdot g(v)\,.$$

(As a graph, G_ω corresponds to the constraint graph of G, but with some edges deleted.) Let L_ω be the corresponding Laplacian, so that for all $g \colon W \to \mathbb{R}$

$$\langle g, L_\omega g\rangle = \mathop{\mathbb{E}}_{(u,v,\pi)\sim G} \pi(x_{\omega,u}, x_{\omega,v}) \cdot \tfrac{1}{2}\Big(g(u) - g(v)\Big)^2\,.$$

With these definitions, $\langle f_\omega, G f_\omega\rangle = \langle h_\omega, G_\omega h_\omega\rangle$ and $\langle f'_{\omega,\tau}, G f'_{\omega,\tau}\rangle = \langle h'_{\omega,\tau}, G_\omega h'_{\omega,\tau}\rangle$. We also use the short-hand, $B_\omega(u) = B_\omega(u, x_{\omega,u})$. With this setup, we can relate $\mathbb{E}_\tau \langle h'_{\omega,\tau}, L_\omega, h'_{\omega,\tau}\rangle$ and $\langle h_\omega, L_\omega, h_\omega\rangle$,

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\tau} \langle h'_{\omega,\tau}, L_\omega h'_{\omega,\tau}\rangle \\
&\leqslant 2\langle h_\omega, L_\omega h_\omega\rangle + \mathop{\mathbb{E}}_{(u,v,\pi)\sim G} \Big(B_\omega(u) + B_\omega(v)\Big) \cdot \pi(x_{\omega,u}, x_{\omega,v}) \cdot \Big|h_\omega(u)^2 - h_\omega(v)^2\Big| \quad \text{(by (5.1))} \\
&\leqslant 2\langle h_\omega, L_\omega, h_\omega\rangle + 10\langle h_\omega, L_\omega h_\omega\rangle^{1/2} \cdot \|B_\omega\| \quad \text{(using Cauchy–Schwarz)}\,.
\end{aligned}$$

Let M_ω be the linear operator on functions on W with the following quadratic form,

$$\langle h_\omega, M_\omega h_\omega\rangle = \mathop{\mathbb{E}}_{(u,v,\pi)\sim G} (1 - \pi(x_u, x_v)) \tfrac{1}{2}\Big(h_\omega(u)^2 + h_\omega(v)^2\Big)$$

Let $L'_\omega = L_\omega + M_\omega$. The following identity among these operators holds

$$\langle h_\omega, G_\omega h_\omega\rangle = \|h_\omega\|^2 - \langle h_\omega, L'_\omega h_\omega\rangle = \|h_\omega\|^2 - \langle h_\omega, L_\omega h_\omega\rangle - \langle h_\omega, M_\omega h_\omega\rangle. \tag{5.2}$$

Let $\gamma_\omega = \|h_\omega\|_1 - \|h_\omega\|^2$, and let $\eta_\omega = \langle h_\omega, L'_\omega h_\omega\rangle$. We see that $\mathbb{E}_\tau \langle h'_{\omega,\tau}, M_\omega h'_{\omega,\tau}\rangle \leqslant 2\langle h_\omega, M_\omega h_\omega\rangle$. Thus, $\mathbb{E}_\tau \langle h'_{\omega,\tau}, L'_\omega h'_{\omega,\tau}\rangle \leqslant O(\eta_\omega + \sqrt{\eta_\omega} \cdot \|B_\omega\|)$ (using (5.1)).

Next, we claim that $\|B_\omega\|^2 = O(\gamma_\omega)$. On the whole domain W, we have $B_\omega \leqslant 100(1 - h_\omega)h_\omega = 100(h_\omega - h_\omega^2)$. So, $\|B_\omega\|^2 \leqslant 100(\|h_\omega\|_1 - \|h_\omega\|^2) = 100\gamma_\omega$. Hence,

$$\mathop{\mathbb{E}}_{\tau} \langle h'_{\omega,\tau}, L'_\omega h'_{\omega,\tau}\rangle \leqslant O(\eta_\omega + \sqrt{\eta_\omega \gamma_\omega}). \tag{5.3}$$

Using (5.2) and the relation between G and G_ω, we can integrate (5.3) over Ω,

$$\begin{aligned}
\langle f', (G \otimes I_{\Omega'}) f'\rangle = \int_\Omega \mathop{\mathbb{E}}_{\tau} \langle h'_{\omega,\tau}, G h'_{\omega,\tau}\rangle\, d\omega &= \|f'\|^2 - O(1) \int_\Omega \eta_\omega + \sqrt{\eta_\omega \gamma_\omega}\, d\omega \\
&= \|f'\|^2 - O(\eta + \sqrt{\eta\gamma})\,.
\end{aligned}$$

The last step uses that $\int_\omega \eta_\omega = \|f\|^2 - \langle f, (G \otimes I_\Omega) f\rangle \leqslant \eta$ and $\int \gamma_\omega = \|f\|_1 - \|f\|^2 \leqslant \gamma$, as well as Cauchy–Schwarz. Since $\|f'\|^2 \geqslant 0.9 - \eta - \gamma$ (using that $\|f\| \geqslant 1 - \eta - \gamma$), we see that $\langle f', (G \otimes I_{\Omega'}) f'\rangle \geqslant (1 - O(\eta + \sqrt{\eta\gamma}))\|f'\|^2$. $\square$

Lemma 5.3. *Let $f \in \mathtt{Assign}_{W,\Sigma}(\Omega)$ be a 0/1-valued function with $\|f_{u,*}\| \geqslant 1/3$ for all but $O(\varepsilon)$ vertices, and $\langle f, (G \otimes I_\Omega) f\rangle = (1 - \varepsilon)\|f\|^2$. Then,*

$$\mathrm{val}(G) \geqslant 1 - O(\varepsilon)\,.$$

Proof. Rescale Ω so that it becomes a probability measure. Let λ be the scaling factor so that $1/3\lambda \leqslant \|f_{u,*}\|^2 \leqslant \lambda$ for all $u \in W$ (after rescaling of Ω). Following the analysis in Lemma 4.3, there exists an assignment for G with value at least $\mathrm{val}(G) \geqslant (1-\varepsilon')/(1+\varepsilon')$ for

$$\varepsilon' = \mathop{\mathbb{E}}_{(u,v,\pi)\sim G} 1 - \tfrac{2}{\|f_{u,*}\|^2+\|f_{v,*}\|^2} \cdot \mathop{\mathbb{E}}_{\omega} \pi(x_{\omega,u}, x_{\omega,v}) \cdot f_\omega(u, x_{\omega,u}) f_\omega(u, x_{\omega,v})\,.$$

Here, $x_\omega \in \Sigma^W$ is a consistent assignment with f_ω for all $\omega \in \Omega$. We are to show $\varepsilon' = O(\varepsilon)$. To ease notation, let use define two jointly-distributed random variables, for $(u,v,\pi) \sim G$,

$$\Delta \cdot \|f\|^2 = \tfrac{1}{2}(\|f_{u,*}\|^2 + \|f_{v,*}\|^2)$$
$$\Gamma \cdot \|f\|^2 = \tfrac{1}{2}(\|f_{u,*}\|^2 + \|f_{v,*}\|^2) - \mathop{\mathbb{E}}_{\omega} \pi(x_{\omega,u}, x_{\omega,v}) \cdot f_\omega(u, x_{\omega,u}) f_\omega(u, x_{\omega,v})$$

In expectation, $\mathbb{E}\,\Gamma = \varepsilon$. Also, $0 \leqslant \Gamma \leqslant \Delta$ with probability 1. Furthermore, $\mathbb{P}\{\Delta < 1/3\} = O(\varepsilon)$. We can express ε' in terms of these variables,

$$\varepsilon' = \mathbb{E}\,1 - \tfrac{1}{\Delta} \cdot (\Delta - \Gamma) = \mathbb{E}\,\tfrac{1}{\Delta}\Gamma\,.$$

Let B (for bad) be the $0/1$-indicator variable of the event $\{\Delta < -2/3\}$ (which is the same as $\{1/(1+\Delta) > 3\}$). Since this event happens with probability at most $O(\varepsilon)$, we have $\mathbb{E}\,B = O(\varepsilon)$. Then, we can bound ε' as

$$\varepsilon' = \mathbb{E}\,\tfrac{1}{\Delta}\Gamma = \mathbb{E}\,B \cdot \underbrace{\tfrac{1}{\Delta}\Gamma}_{\leqslant 1} + \underbrace{(1-B) \cdot \tfrac{1}{\Delta}}_{\leqslant 3} \cdot \Gamma \leqslant \mathbb{E}\,B + 3\,\mathbb{E}\,\Gamma = O(\varepsilon)\,. \qquad \square$$

Proof of Theorem 5.1. Let G and H be two projection games. Suppose G has vertex sets U and V and alphabet Σ. Suppose $\|H\|^2 \leqslant 1-\gamma$ but $\|G \otimes H\|^2 \geqslant 1-\eta-\gamma$ (with γ small enough, say $\gamma \leqslant 1/100$). We are to show $\|G\|^2 \geqslant 1 - O(\eta + \sqrt{\gamma\eta})$.

Let f be an optimal assignment for $G \otimes H$ so that $\|(G \otimes H)f\|^2 \geqslant 1-\eta-\gamma$. Consider the function $h = (\mathrm{Id} \otimes H)f$. This function satisfies $\|(G \otimes \mathrm{Id})h\|^2 \geqslant 1-\eta-\gamma$. Since $\|H\|^2 \leqslant 1-\gamma$, we know that $\|(T_v \otimes \mathrm{Id})h\|^2 \leqslant 1-\gamma$ for all vertices $v \in V$. As before (by a convexity argument), we can derandomize h so that $h \in \mathtt{Assign}_{V,\Sigma}(\Omega)$. Since $0 \leqslant h \leqslant 1$, we can apply Lemma 5.2 to the operator $G' = G^T G$ and the function h to obtain a $0/1$-valued function $h' \in \mathtt{Assign}_{V,\Sigma}(\Omega')$ such that $\|(G \otimes \mathrm{Id})h'\| \geqslant 1 - O(\eta + \sqrt{\eta\gamma})\|h'\|^2$ and $\|h'_{v,*}\|^2 \geqslant 1/3$ for all but an $O(\eta)$ fraction of vertices in V. (The transposition in G^T is with respect to the underlying inner product so that $\langle p, Gq\rangle = \langle G^T p, q\rangle$ for all functions p and q.) By Lemma 5.3, we can use this function to get an assignment for G' of value at least $1 - O(\eta + \sqrt{\eta\gamma})$. The value of an assignment in G' corresponds exactly to the collision value of the assignment in G. Therefore, $\|G\|^2 \geqslant 1 - O(\eta + \sqrt{\gamma\eta})$. $\square$

6 Inapproximability Results

6.1 LABEL COVER

In this section we prove a new hardness result for LABEL COVER (Theorem 6.4 below), and derive Corollary 1.4.

Definition 6.1 (LABEL COVER). Let $\varepsilon : \mathbb{N} \to [0,1]$ and let $s : \mathbb{N} \to \mathbb{N}$ be functions. We define the LABEL COVER$_s(1,\varepsilon)$ problem to be the problem of deciding if an instance of LABEL COVER of size n and alphabet size at most $s(n)$ has value 1 or at most $\varepsilon(n)$.

When we refer to a reduction from 3SAT to LABEL COVER$_s(1, \varepsilon)$, we mean that satisfiable 3SAT instances are mapped to LABEL COVER instances with value 1, and unsatisfiable 3SAT instances are mapped to LABEL COVER instances with value at most ε.

Let us begin by reviewing the known results. We mentioned in Section 3.3 that the PCP theorem [AS98, ALM$^+$98] implies that LABEL COVER$_s(1, 1-\delta)$ for some constant s and (small enough) $\delta > 0$. Parallel repetition applied to this instance k times implies, via Raz's bound that LABEL COVER$_{a^k}(1, \beta^k)$ is NP-hard, for some $\beta < 1$ and some constant $a > 1$. So, taking $k = O(\log 1/\varepsilon)$ will imply a soundness of ε, with an alphabet of size $s = \text{poly}(1/\varepsilon) = a^k$. (This is proved in Section 3.3).

Theorem 6.2 (PCP theorem followed by Raz's theorem)**.** *There are some absolute constants $a > 1$ and $0 < \beta < 1$ such that for every $k \in \mathbb{N}$ there is a reduction that takes instances of 3SAT of size n to instances of* LABEL COVER$_s(1, \varepsilon)$ *of size $n^{O(k)}$, such that $s \leqslant a^k$ and $\varepsilon \leqslant \beta^k$, and in particular, $s \leqslant \text{poly}(1/\varepsilon)$.*

In fact, one can take $k = \omega(1)$ in the above and get a reduction from the original LABEL COVER to LABEL COVER$_s(1, \varepsilon)$ still with $s = \text{poly}(1/\varepsilon) = a^k$ but now the size of the instance grows to be n^k. Setting $k = \log\log n$ or $\log n$ is still often considered reasonable, and yields quasi NP hardness results, namely, hardness results under quasi-polynomial time reductions.

For strictly polynomial-time reductions, the work of Moshkovitz and Raz [MR10] gives hardness for LABEL COVER with sub-constant value of ε,

Theorem 6.3 (Theorem 10 in [MR10])**.** *There exists some constant $c > 0$ such that the following holds. For every $\varepsilon : \mathbb{N} \to [0, 1]$ there is a reduction taking 3SAT instances of size n to* LABEL COVER$_s(1, \varepsilon)$ *instances of size $n^{1+o(1)} \cdot \text{poly}(1/\varepsilon)$ such that $s \leqslant \exp(1/\varepsilon^c)$.*

By applying parallel repetition to an instance of LABEL COVER from the above theorem, and using the bound of Corollary 1.3, we get,

Theorem 6.4 (New NP-hardness for LABEL COVER)**.** *For every constant $\alpha > 0$ the following holds. For every $\varepsilon : \mathbb{N} \to [0, 1]$ there is a reduction taking 3SAT instances of size n to* LABEL COVER$_s(1, \varepsilon)$ *instances of size $n^{O(1)} \cdot \text{poly}(1/\varepsilon)$ such that $s \leqslant \exp(1/\varepsilon^\alpha)$.*

The improvement of this theorem compared to Theorem 6.3 is in the for-all quantifier over α.

Proof. Assume $\varepsilon = o(1)$ otherwise Theorem 6.2 can be applied with $k = O(\log 1/\varepsilon)$ resulting in a much smaller bound on s. The reduction is as follows. Starting with a 3SAT instance of size n, let G be the LABEL COVER$_{s_1}(1, \varepsilon_1)$ instance output of the reduction from Theorem 6.3 with $\varepsilon_1 = (3c/\alpha \cdot \varepsilon^\alpha)^{1/c}$, and output $G^{\otimes k}$ for $k = 3c/\alpha$.

The resulting instance $G^{\otimes k}$ has size $n^{O(1)}$, alphabet size $s = (s_1)^k$ and the soundness is at most $(4\varepsilon_1)^{k/2} \leqslant (4 \cdot (\frac{3c}{\alpha})^{1/c} \varepsilon^{\alpha/c})^{3c/2\alpha} \leqslant \varepsilon$ by Corollary 1.3 (and assuming $\varepsilon = o(1)$). Finally, plugging in the bound for s_1 and the value of ε_1,

$$s = (s_1)^k \leqslant \exp(k/\varepsilon_1^c) = \exp(1/\varepsilon^\alpha)\,.$$

□

Finally, Corollary 1.4 follows immediately from the above theorem, by taking $\varepsilon = (\log n)^{-c}$ and choosing $\alpha < 1/c$ so that the alphabet size is bounded by $s \leqslant \exp(1/\varepsilon^\alpha) \leqslant n$. We also remark that the resulting instance is regular, due to the regularity of the [MR10] instance.

6.2 SET COVER

In this section we prove Corollary 1.5. First, a brief background. Feige [Fei98] proved (extending [LY94]) that SET COVER is hard to approximate to within factor $(1 - o(1)) \ln n$, by showing a quasi-polynomial time reduction from 3SAT. His reduction has two components: a multi-prover verification protocol, and a set-partitioning gadget. Moshkovitz [Mos12] shows that the multi-prover protocol can be replaced by a LABEL COVER instance with an *agreement soundness* property that she defines. She also shows how to obtain such an instance starting with a standard LABEL COVER instance. Corollary 1.5 follows by instantiating this chain of reductions starting with a LABEL COVER instance from Theorem 6.4. Details follow.

Let $\alpha > 0$ and let G be an instance of LABEL COVER as per Corollary 1.4, with soundness $\varepsilon < a^2/(\log n)^4$ for $a = O(\alpha^5)$, and with $|\Sigma| \leqslant n$. This clearly follows by setting $c = 5$ in the corollary, and we can also assume that G is regular. Using a reduction of Moshkovitz (Lemma 2.2 in [Mos12]) we construct in polynomial-time a LABEL COVER instance G_1 such that the Alice degree of G_1 is $D = \Theta(1/\alpha)$, and such that, setting $\varepsilon_1 = a/(\log n)^2$,

- If $\mathrm{val}(G) = 1$ then $\mathrm{val}(G_1) = 1$
- If $\mathrm{val}(G) < \varepsilon_1^2 = a^2/(\log n)^4$ then every assignment for G_1 has the following *agreement soundness* property[8]. For at least $1 - \varepsilon_1$ fraction of the vertices $u \in U_1$ the neighbors of u project to D distinct values (i.e., they completely disagree).

While a hardness for SET COVER is proven in [Mos12], it is unfortunately proven under a stronger conjecture than our Corollary 1.4. For completeness, we repeat the argument. Let us recall the gadget used in [Fei98] (see Definition 3.1 in [Fei98])

Definition 6.5 (Partition Systems)**.** A partition system $B(m, L, k, d)$ has the following properties:

- m: There is a ground set $[m]$. (We will use $m = n_1^D$, where n_1 is the size of G_1).
- L: There is a collection of L distinct partitions, $p_1, \ldots, p_L$. (We will use $L = |\Sigma| \leqslant n$).
- k: For $1 \leqslant i \leqslant L$, partition p_i is a collection of k disjoint subsets of $[m]$ whose union is $[m]$. (We will use $k = D = O(1/\alpha)$).

  Let us denote $p_i(j)$ the j-th set in the partition p_i.
- d: Any cover of $[m]$ by subsets that appear in pairwise different partitions requires at least d subsets. (We will use $d = k \cdot (1 - \frac{2}{D}) \ln m$).

Such a gadget is explicitly constructed in time linear in m and termed "anti-universal sets" in [NSS95]). The intention is that the gadget can be covered in the yes case by k sets, and in the no case by at least $d = k \cdot (1 - \frac{2}{D}) \ln m$ sets.

Construction of SET COVER instance. Finally, the SET COVER instance will have a ground set $U_1 \times [m]$ consisting of $|U_1|$ partition-system gadgets. For every $v \in V_1$ and value $\beta \in \Sigma$ there will be a set $S_{v,\beta}$ in the SET COVER instance. Denoting by D_V the number of neighbors of v in G_1, the set $S_{v,\beta}$ will be the union of D_V sets, one for each neighbor u of v. We arbitrarily enumerate the neighbors of each $u \in U_1$ with numbers from 1 to D, and then denote by $j_{uv} \in [D]$ the number on the edge from u to v. With this notation, $S_{v,\beta}$ will be the union of the sets $\{u\} \times p_\alpha(j)$ where j is the number of the edge uv and $\alpha = \pi_{uv}(\beta)$:

$$S_{v,\beta} = \bigcup_{u \sim v} \{u\} \times (p_{\pi_{uv}(\beta)}(j_{uv}))\,.$$

[8] In [Mos12] the projections go from Alice to Bob, while ours go from Bob to Alice.

Completeness. It is easy to see that if f, g is a satisfying assignment for G_1 then by taking the sets $S_{v,f(v)}$ the gadget corresponding to u is covered by the k sets in the partition corresponding to $\alpha = g(u)$. In total the set cover has size $|V_1| = D|U_1|$.

Soundness. In the no case, we claim that every set cover has size at least $Z = (1 - \frac{4}{D}) \ln m \cdot |V_1|$. Assume otherwise, and let s_u be the number of sets in the cover that touch $\{u\} \times [m]$. By

$$\sum_u s_u = Z \cdot D_V ,$$

at least $\frac{2}{D}$ fraction of the vertices $u \in U_1$ have $s_u < \ell \stackrel{\text{def}}{=} (1 - \frac{2}{D}) D \ln m$, and we call such u's *good*. Define a randomized assignment $f : V \to \Sigma$ by selecting for each v a value β at random from the set of β's for which $S_{v,\beta}$ belongs to the set cover (or, if no such set exists output a random β).

First, by the property of the gadget, since the set cover must cover each $\{u\} \times [m]$, if u is good there must be two neighbors $v_1, v_2 \sim u$ such that the sets S_{v_1,β_1} and S_{v_2,β_2} are in the cover, and such that $\pi_{uv_1}(\beta_1) = \alpha = \pi_{uv_2}(\beta_2)$ (both $p_\alpha(j_{uv}) \subset S_{v_1,\beta_1}$ and $p_\alpha(j_{uv_2}) \subset S_{v_2,\beta_2}$).

Next, observe that if u is good then each neighbor v of u has at most $\ell = (1 - \frac{2}{D}) D \ln m$ values of β for which $S_{v,\beta}$ is in the cover, because each such set is counted in s_u.

So the probability that $f(v_1) = \beta_1$ and $f(v_2) = \beta_2$ is at least $\frac{2}{D} \cdot \frac{1}{\ell^2} > \varepsilon_1$ and this contradicts the agreement soundness property of G_1 as long as $\varepsilon_1 < a/(\log n)^2$, for $a < O(1/D^5) = O(\alpha^5)$.

Size. The size of the setcover instance is at most $m = n_1^D$ times n_1, so $\ln(n_1^{D+1}) = (D+1) \ln n_1 = (1 + \frac{1}{D}) \ln m$. So by choosing the appropriate constant relation between $1/D$ and α we get a factor of $(1 - \alpha) \ln N$ hardness for approximating a SET COVER instance of size N.

7 Conclusions

In many contexts, tight parallel repetition bounds are still open, for example, general[9] (non-projection) two-player games, XOR games, and games with more than two parties.

It is an interesting question whether the analytical approach in this work can give improved parallel repetition bounds for these cases. Recently, the authors together with Vidick gave a bound on entangled games (where the two players share entanglement), following the approach of this work.

A more open-ended question is whether analytical approaches can complement or replace information-theoretic approaches in other contexts (for example, in communication complexity) leading to new bounds or simpler proofs.

Acknowledgments

We thank Boaz Barak, Ryan O'Donnell, Ran Raz, and Oded Regev for insightful discussions and comments about this work.

[9] An earlier version of this manuscript erroneously claimed a reduction from general constraints to projection constraints. We are thankful to Ran Raz for pointing out this error.

References

[ALM+98] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy, *Proof verification and intractability of approximation problems*, Journal of the ACM **45** (1998), no. 3, 501–555. 15, 24

[AS98] S. Arora and S. Safra, *Probabilistic checking of proofs: A new characterization of NP*, Journal of the ACM **45** (1998), no. 1, 70–122. 15, 24

[BHH+08] Boaz Barak, Moritz Hardt, Ishay Haviv, Anup Rao, Oded Regev, and David Steurer, *Rounding parallel repetitions of unique games*, FOCS, 2008, pp. 374–383. 5, 6, 7

[BRR+09] Boaz Barak, Anup Rao, Ran Raz, Ricky Rosen, and Ronen Shaltiel, *Strong parallel repetition theorem for free projection games*, APPROX-RANDOM, 2009, pp. 352–365. 5

[CKW09] Marek Cygan, Lukasz Kowalik, and Mateusz Wykurz, *Exponential-time approximation of weighted set cover*, Inf. Process. Lett. **109** (2009), no. 16, 957–961. 4

[CS13] André Chailloux and Giannicola Scarpa, *Parallel repetition of entangled games with exponential decay via the superposed information cost*, arXiv:1310.7787, 2013. 3

[DSV14] Irit Dinur, David Steurer, and Thomas Vidick, *A parallel repetition theorem for entangled projection games*, IEEE Conference on Computational Complexity, 2014, arXiv:1310.4113. 3

[Fei91] Uriel Feige, *On the success probability of the two provers in one-round proof systems*, Structure in Complexity Theory Conference, IEEE Computer Society, 1991, pp. 116–123. 5, 29

[Fei98] U. Feige, *A threshold of ln n for approximating set cover*, Journal of the ACM **45** (1998), no. 4, 634–652. 4, 25

[FKO07] Uriel Feige, Guy Kindler, and Ryan O'Donnell, *Understanding parallel repetition requires understanding foams*, IEEE Conference on Computational Complexity, 2007, pp. 179–192. 5

[FL92] Uriel Feige and László Lovász, *Two-prover one-round proof systems: Their power and their problems (extended abstract)*, STOC, 1992, pp. 733–744. 6

[Hol09] Thomas Holenstein, *Parallel repetition: Simplification and the no-signaling case*, Theory of Computing **5** (2009), no. 1, 141–172. 3, 6, 16

[JPY14] Rahul Jain, Attila Pereszlényi, and Penghui Yao, *A parallel repetition theorem for entangled two-player one-round games under product distributions*, IEEE Conference on Computational Complexity, 2014. 3

[Lov79] László Lovász, *On the shannon capacity of a graph*, IEEE Transactions on Information Theory **25** (1979), no. 1, 1–7. 6

[LY94] Carsten Lund and Mihalis Yannakakis, *On the hardness of approximating minimization problems*, Journal of the ACM **41** (1994), no. 5, 960–981. 25

[Mos12] Dana Moshkovitz, *The projection games conjecture and the NP-hardness of ln n-approximating set-cover*, APPROX-RANDOM, 2012, pp. 276–287. 4, 25

[MR10] Dana Moshkovitz and Ran Raz, *Two-query PCP with subconstant error*, J. ACM **57** (2010), no. 5. 4, 24

[NSS95] Moni Naor, Leonard J. Schulman, and Aravind Srinivasan, *Splitters and near-optimal derandomization*, FOCS, 1995, pp. 182–191. 25

[Rao08a] Anup Rao, *Parallel repetition in projection games and a concentration bound*, STOC, 2008, pp. 1–10. 4

[Rao08b] ———, *Parallel repetition in projection games and a concentration bound*, Electronic Colloquium on Computational Complexity (ECCC) **15** (2008), no. 013. 11

[Rao11] ———, *Parallel repetition in projection games and a concentration bound*, SIAM J. Comput. **40** (2011), no. 6, 1871–1891. 3, 5

[Raz98] Ran Raz, *A parallel repetition theorem*, SIAM J. Comput. **27** (1998), no. 3, 763–803. 3

[Raz08] ———, *A counterexample to strong parallel repetition*, FOCS, 2008, pp. 369–373. 5, 6

[RR12] Ran Raz and Ricky Rosen, *A strong parallel repetition theorem for projection games on expanders*, IEEE Conference on Computational Complexity, 2012, pp. 247–257. 5

[Ste10a] David Steurer, *Improved rounding for parallel repeated unique games*, APPROX-RANDOM, 2010, pp. 724–737. 6

[Ste10b] ———, *Subexponential algorithms for d-to-1 two-prover games and for certifying almost perfect expansion*, Manuscript, available from the author's website., 2010. 6, 7

A Additional Proofs

Reduction to expanding games.

Claim A.1. There is a constant $\gamma > 0$ and a simple transformation from a projection game G to a projection game G' such that $\mathrm{val}(G') = \frac{1}{2} + \frac{\mathrm{val}(G)}{2}$ and such that G' has eigenvalue gap $\geqslant \gamma$. If G is (c, d)-regular (i.e. the degree of any $u \in U$ is c and the degree of any $v \in V$ is d) then G' can be made $(c, 2d)$-regular.

Proof. Let U, V be the vertices of G, and let μ be the associated distribution of G. Define $U' = U \cup \{u_0\}$, and a new distribution μ' that is with probability $1/2$ equal to μ and with probability $1/2$ equal to the product distribution given by $\{u_0\} \times \mu_V \times \{\pi_0\}$ where π_0 is a 'trivial' constraint that accepts when Alice answers 1 and regardless of Bob's answer. The claim about the value is easily verified. As for the eigenvalue gap, it follows because the symmetrized game corresponding to $\{u_0\} \times \mu_V \times \{\pi_0\}$ is the complete graph, with eigenvalue gap 1.

If G is (c, d) regular and we want G' to also be $(c, 2d)$-regular, then instead of adding one vertex u_0 we add a set U_0 of $|U|$ new vertices, so $U' = U \cup U_0$, and place a (c, d)-regular graph G_0 between U_0 and V, accompanied with π_0 constraints as above. We are free to choose the structure of G_0, and if we take it to be such that its symmetrized graph has eigenvalue gap 2γ, then the eigenvalue gap of G' will be at least γ. Clearly the distribution μ' is uniform on U' which means that with probability $1/2$ it chooses a trivial constraint, and with probability $1/2$ a G constraint. Hence, $\mathrm{val}(G') = \frac{1}{2} + \frac{1}{2}\mathrm{val}(G)$. $\square$

B Feige's Game

Uri Feige [Fei91] describes a game G (which he calls NA for non-interactive agreement) for which $\mathrm{val}(G) = \mathrm{val}(G \otimes G) = \frac{1}{2}$. This example rules out results of the form $\mathrm{val}(G \otimes H) \leqslant \mathit{val}'(G) \cdot \mathrm{val}(H)$, for a useful approximation val' of val thus perhaps discouraging the search for game-relaxations that are multiplicative. Our Theorem 1.8 sidesteps this limitation by proving a multiplicative relation for $\|G\|$ and not for $\mathrm{val}(G)$. We prove, in Theorem 1.8 that $\|G \otimes H\| \leqslant \mathrm{val}_+(G) \cdot \|H\|$ and that the relaxation $\mathrm{val}_+(G)$ is bounded away from 1 for any game G whose value is bounded away from 1.

Our theorem implies that there is no projection game (including Feige's NA game) for which $\|G \otimes G\| = \|G\|$. Indeed we calculate below and show that for Feige's game, $\|G \otimes G\| = \|G\|^2 = 1/2$.

Feige's game is defined over question set $U = \{0, 1\} = V$ where each pair of questions $(u, v) \in \{0, 1\}^2$ is equally likely (this is called a free game). The players win if they agree on a player and her question, namely, the alphabet is $\Sigma = \{A0, A1, B0, B1\}$ and for each question pair $(u, v) \in \{0, 1\}^2$, there are two acceptable pairs of answers: (Au, Au) and (Bv, Bv) (where u, v should be replaced by the bit value 0 or 1).

Is this a projection game? It is true that every answer for Bob leaves *at most* one possible accepting answer for Alice, but note that some answers leave no possible accepting answer. Nevertheless, the inequality between the value and the collision value established in Claim 2.3 holds for such games.

Claim B.1. $\|G\|^2 = 1/2$.

Proof. It is easy to see that $\|G\|^2 = \frac{1}{2}$, by having Bob always reply with $A0$, i.e., setting Bob's strategy to be $f(0, A0) = 1, f(1, A0) = 1$, and then

$$\|Gf\|^2 = 1/2 \sum_{\alpha \in \Sigma} Gf(0, \alpha)^2 + 1/2 \sum_{\alpha \in \Sigma} Gf(1, \alpha)^2 = 1/2 \cdot 1 + 1/2 \cdot 0 = 1/2\,.$$

□

Feige's strategy achieving $\mathrm{val}(G \otimes G) = 1/2$ is one where Bob hopes that his second question is equal to Alice's first question, and thus answers a question $v_1 v_2$ by $(Av_2, Bv_2) \in \Sigma^2$.

Claim B.2. Let f be Feige's strategy attaining $\mathrm{val}(G \otimes G) = 1/2$ (defined as $f(v_1 v_2, Av_2 Bv_2) = 1$ for each $v_1 v_2$). Then $\|G \otimes Gf\|^2 = \frac{1}{4}$.

Proof. Note that $\mathrm{val}(G \otimes G, f) = 1/2$ because with probability $1/2$ $v_2 = u_1$ and if Alice takes a symmetrical strategy $g(u_1 u_2, Au_1 Bu_1) = 1$ then the players win. Let us analyze $\|G \otimes Gf\|^2$ for this f. Gf looks similar for all question tuples $u_1 u_2 \in \{00, 01, 10, 11\}$ that Alice might see. For example, suppose $u_1 u_2 = 00$. Of the four possible tuples $v_1 v_2$ that Bob may have received, only $v_1 v_2 = 00$ and $v_1 v_2 = 10$ will contribute to $Gf(00, A0B0)$. Indeed

$$Gf(00, A0B0) = \frac{f(00, A0B0) + f(01, A0B0) + f(10, A0B0) + f(11, A0B0)}{4} = \frac{1 + 0 + 1 + 0}{4} = 1/2$$

For any other $\alpha \in \Sigma^2$, $Gf(00, \alpha) = 0$, because the answer $A1B1$ is not acceptable when Alice's questions are $u_1 u_2 = 00$. Thus, we get $\|G \otimes Gf\|^2 = \frac{1}{4} \cdot 4 \cdot (\frac{1}{4})^2 = \frac{1}{4}$. □

One would guess also that $\|G \otimes G\|^2 = \frac{1}{4}$ but we have not analyzed this.