Electronic Colloquium on Computational Complexity, Report No. 34 (2018)

On Symmetric Parallel Repetition : Towards Equivalence of MAX-CUT and UG

Young Kun Ko

Abstract

Unique Games Conjecture (UGC), proposed by [24], lies in the center of many inapproximability results. At the heart of UGC lies approximability of MAX-CUT, which is a special instance of Unique Game. [25, 29] showed that assuming Unique Games Conjecture, it is NP-hard to distinguish between MAX-CUT instance that has a value $1-\varepsilon$ vs. $1-\Omega(\sqrt{\varepsilon})$. [14] then showed a matching polynomial time algorithm using Semi-Definite Programming. Towards resolving UGC, it has been long conjectured that inapproximability of MAX-CUT and UGC are equivalent. Assuming the equivalence, it suffices to exhibit lower bounds on MAX-CUT towards resolving UGC.

Towards showing the equivalence of the hardness of MAX-CUT and UGC, we initiate the study of symmetric parallel repetition, which is parallel repetition without coordinate. In particular, we show that symmetric parallel repetition beats strong parallel repetition in certain regimes, that is the value decays $(1-\varepsilon^c)^{\Omega(r)}$ with $c < 2$, exhibiting the first separation between symmetric parallel repetition and usual parallel repetition. This is in sharp contrast to the usual parallel repetition as shown by [21, 30] where the best upper bound known for the value of the game $\mathcal{G}^{\otimes r}$ is $(1-\varepsilon^2/2)^{\Omega(r)}$ for projection games. The counterexample shown by [32] gives a lower bound of $\mathbf{val}(\mathcal{G}^{\otimes r}) \geq (1-\varepsilon^2)^{O(r)}$ for $r = \Omega(n^2)$ where n is the size of the graph. This also implies that the odd cycle game is not a counterexample for symmetric parallel repetition.

The main technical tool is the analysis of the Birthday Repetition in high intersection regime first introduced in [1] subsequently improved in [27]. From a technical perspective we show that (a) if the set size is slightly larger than $\sqrt{n}$, then the value decays strong exponentially (i.e. $(1-\varepsilon)^{\tilde{\Omega}(r)}$) in the expected intersection size; (b) if the set becomes large as in the whole vertex set, then the value decays strong exponentially in the number of edges that are checked by the verifier. Then we use prove a translation lemma to translate these technical results to corollaries in symmetric parallel repetition.

This exhibits a dichotomy between the usual parallel repetition and symmetric parallel repetition. In particular, it shows that the avenue of attack for showing the equivalence between the hardness of MAX-CUT and the Unique Games Conjecture using some model of repetition is still open.

ISSN 1433-8092

1 Introduction

Two Prover game ($\mathcal{G}$) is a key concept in computational complexity defined by underlying (bipartite) graph $G = (X, Y, E)$, alphabets A and B respectively for Prover 1 (Alice) and Prover 2 (Bob) and a verification function $V : X \times Y \times A \times B \to \{0, 1\}$ for the verifier. In particular, computing the *value* of two prover games i.e.

$$\mathbf{val}(\mathcal{G}) := \max_{f,g} \mathop{\mathbb{E}}_{(x,y)\in E} [V(x, y, f(x), g(y))]$$

where $f : X \to A$ and $g : Y \to B$ are Alice and Bob's respective answering strategy, is used as a starting point for the hardness of approximation for various problems. The original hardness of two prover game [4, 5] and subsequently [16] show that there is **some constant** ε such that it is NP-hard to distinguish between $\mathbf{val}(\mathcal{G}) = 1$ and $\mathbf{val}(\mathcal{G}) < 1 - \varepsilon$ instead of **any constant** or any super-constant hardness.

The key technique to amplifying the hardness of two prover game is **Parallel Repetition**, repeating the game in parallel r-times. More precisely, the verifier, instead of picking a single edge, picks r random edges from G, $(x_1, y_1), \ldots, (x_r, y_r)$, then let Alice and Bob give assignment to all r-edges. The verifier then returns $\bigwedge_{i=1}^{r} V(x_i, y_i, f_i(\vec{x}), g_i(\vec{y}))$, where $f_i(\vec{x})$ is the assignment to x_i and similarly for $g_i(\vec{y})$, and $\vec{x} = (x_1, \ldots, x_r)$ and similarly for $\vec{y}$.

At a first glance, it seems that if $\mathbf{val}(\mathcal{G}) < 1 - \varepsilon$, then the value should decay as $(1 - \varepsilon)^r$. However, this is not true, since one can correlate the answer between different coordinates and perform better than $(1 - \varepsilon)^r$. (note that f_i has the whole $\vec{x}$ as the input.)

Still Parallel Repetition does hold. [39] first showed that the value converges to 0 as r goes to infinity. A major breakthrough by Raz [31] showed first exponential decay (as in $(1 - \varepsilon^c)^{\Omega(r)}$ where c depends on the answer length) followed by subsequent simplifications and improvements by [21]. Rao improved Holenstein's proof for the case of projection games (i.e. games where Bob's assignment forces an assignment on Alice) upper-bounding the value by $(1 - \varepsilon^2/2)^{\Omega(r)}$ [30, 18]. The proof was later generalized by [9] for games with small value as well.

Then a natural question is, "can we improve this bound" In particular, can we prove a value decay of $(1-\varepsilon^c)^{\Omega(r)}$ where $c < 2$? Indeed, one might argue that achieving $c = O(1)$ suffices for many hardness amplification applications. But apart from just curiosity, this has a deep connection to the Unique Games Conjecture, first proposed by [24], which lies as one of the fundamental conjectures in the hardness of approximation. In particular, if the upper bound of $(1-\varepsilon^c)^{\Omega(r)}$ for some constant $c < 2$ holds, then NP-hardness of distinguishing between the two cases of MAXCUT: (YES) there exists a cut that cuts $1-\varepsilon$-fraction of the edges; (NO) any cut can only cut at most $1-\varepsilon^{1/c}$-fraction of the edges, implies the full Unique Games Conjecture, giving a potential attack route for resolving UGC.

To partially answer this question, since [31], parallel repetition under various special settings were studied. These include free games where Alice (Prover 1)'s input and Bob (Prover 2)'s input are independent [8], projection games on expanders [34, 38], entangled games [23, 22, 13, 19, 40], fortified games [28], multiplayer games [17]. Some of these games indeed obtain the decay rate of $(1 - \varepsilon)^{\Omega(r)}$, i.e. strong parallel repetition. For instance [34] obtains strong parallel repetition for expanding projection games. But none of these results achieve bound of $(1 - \varepsilon^c)^{\Omega(r)}$ with $c < 2$ in full generality.

If we cannot prove parallel repetition in full generality, then the next question to ask is whether there is a "hard" instance that satisfies such special conditions, or transforming any instances to an instance that obeys better parallel repetition as in [28], which are not ruled out by upper bound results. Fortification transformation introduced by [28] indeed achieves strong parallel repetition but does not preserve uniqueness mainly due to its asymmetric property. It only preserves projection property of the game. For example, we know that unique games on expander has polynomial time algorithm [3], implying that there is no hard instance that satisfies the expander condition.

However, [33] showed a game (the odd-cycle game) where the upper bound is tight. If $r = \Omega(n^2)$ where n is the size of the graph, the value of the game is at least $(1 - \varepsilon^2)^{\Omega(r)}$, showing that [30] is tight upto some constant. This implies that the statement is false in full generality.

Further generalizing this impossibility result, this has been later generalized to all "hard" unique games for SDP [7, 37] and games with low information cost protocols [10]. Specifically, [7] rules out the parallel repetition approach to the Unique Games Conjecture. It implies that any game that decays at $(1 - \varepsilon)^{\Omega(r)}$-rate will be easy under SDP.

These counterexamples to parallel repetition depend heavily on the fact that we put coordinates for each copy of the game. To avoid these technical complications, **symmetric parallel repetition** has been suggested as an alternative to parallel repetition, that is to sample sets of edges with no coordinates attached. More precisely, instead of playing r independent copies of the game, the verifier samples edges at random then sends the set of vertices $\{x_1, \ldots, x_r\}$ and $\{y_1, \ldots, y_r\}$ to Alice and Bob, instead of tuples of vertices indexed by the coordinates.[1] This is indeed at least as strong as the usual parallel repetition, since any strategy for symmetric parallel repetition can be translated to a strategy for usual parallel repetition.

However, it is not clear whether this is strictly stronger than the usual parallel repetition. Indeed the arguments from [33, 7] no longer applies to symmetric parallel repetition. Thus it might be the case that strong parallel repetition indeed holds for symmetric parallel repetition and is the way to attack UGC. But all known methods in proving parallel repetition seem to fail for symmetric parallel repetition since the proofs argue about extracting a strategy on a single coordinate conditioned on winning some subset of coordinates. This strategy becomes "too good to be true" strategy leading to a contradiction if the probability of winning some subset of coordinates is too high.

Independent of the developments in parallel repetition, there have been developments on set-based reductions on two prover games. Birthday Repetition, first introduced in [1], converts any game to a free game. In $\mathcal{G}^{k \times \ell}$, the verifier independently chooses k-sized random set from X for Alice and ℓ-sized random set from Y for Bob, and checks all the edges that lie between those two sets. [1] showed that if $\mathbf{val}(\mathcal{G}) = 1 - \varepsilon$ and $k = \ell = O(\sqrt{n}/\varepsilon)$, $\mathbf{val}(\mathcal{G}^{k \times \ell}) \approx \mathbf{val}(\mathcal{G})$. This was then used to show that free games are hard to approximate in quasi-polynomial time and has seen numerous applications in showing quasi-polytime hardnesses [12, 11, 35, 36, 26, 15]. It was later shown with improved bound in [27].

An interesting observation on Birthday Repetition is that it is the only known set-based hardness amplification. In fact, the whole reduction relies on the fact that it is set-based (Birthday Paradox on the set of n elements), since otherwise the value will not be preserved. We use the intuitions developed through series of paper on Birthday Repetition to first show tighter bounds for Birthday Repetition with larger set sizes or higher degree. Then we show that these bounds can be translated to obtain result on symmetric parallel repetition as a corollary using the translation lemma.

1.1 Our Technical Results

Instead of the original Birthday Repetition presented in [1], we redefine it with "Refined" Birthday Repetition, first observed in [15] to prove tighter bounds for the best approximate Nash Equilibria. The main difference is instead of choosing fixed sized random sets, we flip coins independently for each $x \in X$ and $y \in Y$ and consider the corresponding random sets $S \subset X$ and $T \subset Y$.

[1] There are many suggested definitions to symmetric parallel repetition. But note that this is one option since the size of the reduction is at most $n^{\Omega(r)}$, and the verification function is well defined

Though it does not give the same distribution over the challenges as in the original Birthday Repetition, "Refined" Birthday Repetition indeed makes Alice's distribution independent of Bob's distribution, thereby making the game a "free game." Yet, it is easier to analyze the value and requires a smaller subset size due to a stronger concentration property. While [1] required a set size of $\Omega(\sqrt{n/\varepsilon})$ for ε-additive approximation of the value, $\Omega(\sqrt{n\log(1/\varepsilon)})$ sized set (in expectation) suffices for "Refined" Birthday Repetition. In particular, we prove the following theorem:

Theorem 1.1 (Refined Birthday Repetition)**.** *Suppose* $k = \ell = \Omega(\sqrt{n\log(1/\varepsilon)})$*, and* $\mathbf{val}(\mathcal{G}) < 1 - \varepsilon$*. Then*

$$\mathbf{val}(\mathcal{G}^{k\times\ell}) \leq 1 - \frac{\varepsilon}{2}.$$

This not only improves the required set size in [1], but also substantially simplifies the proof. With a condition on the degree of two prover game (d), we can also obtain the following.

Theorem 1.2 (High Degree Refined Birthday Repetition)**.** *Suppose* $d = \Omega(n\log n/k)$*,* $k = \ell$ *and* $\mathbf{val}(\mathcal{G}) < 1 - \varepsilon$*. Then*

$$\mathbf{val}(\mathcal{G}^{k\times\ell}) \leq 1 - \frac{\varepsilon}{2}.$$

With Theorem 1.1 and results from [8], it is straightforward to prove a naive parallel repetition for "Refined" Birthday Repetition:

Theorem 1.3. *Suppose* $k = \ell = \Omega(\sqrt{n\log(1/\varepsilon)})$*, and* $\mathcal{G}$ *is a projection game, that is for every* $(x,y) \in E$*, there exists a function* $p_{(x,y)} : B \to A$ *such that* $V(x,y,a,b) = 1$ *iff* $p_{(x,y)}(b) = a$*. Then if* $\mathbf{val}(\mathcal{G}) < 1 - \varepsilon$

$$\mathbf{val}(\mathcal{G}^{rk\times r\ell}) \leq \mathbf{val}((\mathcal{G}^{k\times\ell})^{\otimes r}) \leq \left(1 - \frac{\varepsilon}{2}\right)^{\Omega(r)}.$$

As a quick remark, though it is not fully exponential in r, if one focuses on achieving hardness of any constant, this is a quantitatively better bound compared to [27], whose bound in this language is comparable to $(1-\varepsilon^c)^{\Omega(r^2)}$ for some constant $c > 2$ thereby achieving a simpler proof of [27]. Unfortunately Theorem 1.3 is not tight enough to obtain corollaries for symmetric parallel repetition for our application. We can further tighten the bound:

Theorem 1.4. *Suppose* $k = \ell = \Omega(\sqrt{n\log(1/\varepsilon)})$*, and* $\mathcal{G}$ *is a projection game. Further, suppose* $r \leq \min\{\sqrt{n}/2, \varepsilon n/320\}$*. Then if* $\mathbf{val}(\mathcal{G}) < 1 - \varepsilon$*,*

$$\mathbf{val}(\mathcal{G}^{rk\times r\ell}) \leq (1 - \varepsilon/2)^{\Omega(r^2)}.$$

For "high" degree regime, that is with a lower bound condition on d, we get the following

Theorem 1.5. *Suppose* $k = \ell$*, and* $\mathcal{G}$ *is a projection game, and* $d = \Omega(\frac{n\log n}{k})$*. Further, suppose* $r \leq \min\{\sqrt{n}/2, \varepsilon n/320, d/2\}$*. Then if* $\mathbf{val}(\mathcal{G}) < 1 - \varepsilon$*,*

$$\mathbf{val}(\mathcal{G}^{rk\times r\ell}) \leq (1 - \varepsilon/2)^{\Omega(r^2)}.$$

Fortunately, Theorem 1.4 and Theorem 1.5 suffice to obtain bounds for symmetric parallel repetition that beats the bounds for usual parallel repetition in some regime.

1.2 Symmetric Parallel Repetition

First, we give a general definition of Symmetric Parallel Repetition. To the best of our knowledge, this has not yet been precisely defined in prior works. Define $\mathcal{G}_{\mathcal{D}}$ as the following two prover game, where $\mathcal{D}$ is a distribution over 2^E.

- The referee pick a subset of edges $C \subseteq E$ according to $\mathcal{D}$
- Send $C_X = \{x | (x, y) \in C\}$ to Alice and $C_Y = \{y | (x, y) \in C\}$ to Bob.
- Alice and Bob returns the assignment to Z_X and Z_Y respectively.
- The referee replies with $\bigwedge_{(x,y)\in C} V(x, y, a_x, b_y)$, where a_x is assignment to x and b_y is assignment to y.

Observe that if $|C_X|, |C_Y| < r$ over $\mathcal{D}$ then the size of $\mathcal{D}$ is at most n^r.

Under this definition, we prove the following translation lemma which translates results on Birthday Repetition to results on symmetric parallel repetition:

Lemma 1.6 (Translation Lemma). *If the degree of the original Two Prover game $d = \Omega(n \log n/k)$, with $k = \ell$ then there exists $\mathcal{D}_{k,\ell}$ that is supported only on sets of size at most $O(k)$ such that*

$$|\mathbf{val}(\mathcal{G}_{\mathcal{D}_{k,\ell}}) - \mathbf{val}(\mathcal{G}^{k\times\ell})| = o(1)$$

Combining theorems from Section 1.1 and Lemma 1.6, we obtain the main theorem on symmetric parallel repetition as a corollary:

Theorem 1.7 (Main). *Let $\mathcal{G}$ be a projection game where the underlying constraint graph is a d-regular graph. Let $\mathbf{val}(\mathcal{G}) < 1 - \varepsilon$. Set $k = \ell = 4n \log n/d$. If $k = \ell = \Omega(\sqrt{n \log(1/\varepsilon)})$ and $\frac{k^2}{n \log(1/\varepsilon)} \leq \min\{\sqrt{n}/2, \varepsilon n/320\}$ then there exists a distribution on $\mathcal{D}_{k,\ell}$ on 2^E that is only supported on the sets of size $O(k)$ such that*

$$\mathbf{val}(\mathcal{G}_{\mathcal{D}_{k,\ell}}) \leq (1 - \varepsilon/2)^{\Omega(\frac{k^2}{n \log(1/\varepsilon)})}.$$

Else, for any $r > 0$ that satisfies $r \leq \min\{\sqrt{n}/2, \varepsilon n/320, d/2\}$, there exists a distribution on $\mathcal{D}_{rk,r\ell}$ on 2^E that is only supported on the sets of size $O(rk)$ such that

$$\mathbf{val}(\mathcal{G}_{\mathcal{D}_{rk,r\ell}}) \leq (1 - \varepsilon/2)^{\Omega(r^2)}.$$

As a complementary result for amplifying the gaps with imperfect completeness (for example MAXCUT with completeness $1 - \varepsilon$ and soundness $1 - \delta$), we also prove the following lower bound on the value, which along with Lemma 1.6, gives a lower bound on $\mathbf{val}(\mathcal{G}_{\mathcal{D}_{k,\ell}})$.

Theorem 1.8 (Lower Bound). *If $\mathcal{G}$ is a projection game where the underlying constraint graph is d-regular graph. Let $\mathbf{val}(\mathcal{G}) > 1 - \varepsilon$, $k, \ell = o(n)$ and $\frac{\varepsilon d k \ell}{n} \leq 1/8$. Then*

$$\mathbf{val}(\mathcal{G}^{k\times\ell}) \geq e^{-O(\frac{\varepsilon d k \ell}{n})}.$$

Unfortunately Theorem 1.4, Theorem 1.5 and therefore Theorem 1.7 does not handle the case where r is too large. But observe that if r becomes too large, then $rk = r\ell = n$. In that case, it essentially becomes a game with only two vertices, one for Alice and one for Bob, but with exponentially sized alphabet. The analysis becomes extremely simple in that regime and does not require the underlying game to be a projection game. The analysis in that regime gives the following theorem as a corollary:

Theorem 1.9 (High repetition regime)**.** *Let $\mathcal{G}$ be a two prover game (not necessarily a projection game) with $\mathbf{val}(\mathcal{G}) < 1-\varepsilon$. If $r = \Omega(n\log(1/\varepsilon))$ then there exists $\mathcal{D}_r$ on 2^E that is only supported on edge sets of size at most $2r$ such that $\mathbf{val}(\mathcal{G}_{\mathcal{D}_r}) = (1-\varepsilon)^{\Omega(r)}$. Similarly, if $\mathbf{val}(\mathcal{G}) > 1-\varepsilon$, then $\mathbf{val}(\mathcal{G}_{\mathcal{D}_r}) = (1-\varepsilon)^{O(r)}$*

Theorem 1.9 exhibits an interesting dichotomy with the usual parallel repetition. Recall the bound in [33] in the case of projection games. For usual parallel repetition $\lim_{r\to\infty}\log(\mathbf{val}(\mathcal{G}^{\otimes r}))/r = \Theta(-\varepsilon^2)$ (and this is tight by [30]), while for symmetric parallel repetition $\lim_{r\to\infty}\log(\mathbf{val}(\mathcal{G}_{\mathcal{D}_r}))/r \leq -\Theta(\varepsilon)$ from Theorem 1.9.

1.3 Further Directions

Connections to the Unique Games Conjecture Recall that the Unique Games Conjecture is phrased as following.

Conjecture 1.10 (Unique Games Conjecture (UGC), [24])**.** *For every $\varepsilon, \delta > 0$, there exists $n = n(\varepsilon, \delta)$, such that there exists unique game instance $\mathcal{G}$ such that it is NP-hard to distinguish between (YES) $\mathbf{val}(\mathcal{G}) \geq 1-\varepsilon$; (NO) $\mathbf{val}(\mathcal{G}) \leq \delta$.*

Rao's parallel repetition theorem for projection games [30] implies that if it is "hard" to distinguish between two instances where at least $1-\varepsilon$ fraction of the edges are cut (completeness) and at most $1-\sqrt{\varepsilon}$ fraction of the edges are cut (soundness), which we abbreviate as $(1-\varepsilon, 1-\sqrt{\varepsilon})$ MAX-CUT, then one can amplify the gap to obtain hardness result for the Unique Games with a blow up size of $n^{1/\varepsilon}$. By applying parallel repetition $1/\varepsilon$-times, completeness and soundness both becomes some arbitrary constants, as in UGC.

In particular, the following approach has been suggested to show that there is no polynomial time algorithm for Unique Games under Exponential Time Hypothesis[2] : (i) reduction from 3SAT to $(1-\varepsilon, 1-\sqrt{\varepsilon})$ MAX-CUT of size $s(n)$; (ii) apply parallel repetition to obtain hardness for any constant (c, s) which blows up the instance size to $s(n)^{1/\varepsilon}$. In particular, note that if one have $s(n) = 2^{O(\sqrt{n})}$, $\varepsilon = \omega(1/\sqrt{n})$, it then implies that there is no polynomial time algorithm for the Unique Games Conjecture.

Unfortunately, this proof outline has been undermined by impossibility results in both (i) and (ii). [14] showed a polynomial time algorithm based on Semi Definite Programming for $(1-\varepsilon, 1-\sqrt{\varepsilon})$ MAX-CUT. Therefore, there is no base "hard" instance that works for (i). Then a natural question is whether we can improve (ii), that is whether we can obtain $(1-\varepsilon^c)^r$ with $c < 2$. By improving [30], one can hope to relax the condition for (i), thereby bypassing the upper bound result by [14]. Whether such bound is achievable has been addressed by [33], and generalized in [7, 37] and [10]. The bottom line is that **games that have improved bounds in (ii) cannot be a hard instance to begin with if we are using the usual parallel repetition**.

The main motivation for proving bounds for symmetric parallel repetition is to avoid these impossibility results. By improving bounds on (ii), one can have relaxed conditions for (i), bypassing [14], and instead requiring hardness result for $(1-\varepsilon, 1-\varepsilon^\alpha)$ MAX-CUT with $\alpha < 1/2$.

Our result on symmetric parallel repetition shows that resolving the following conjectures would result in runtime lower bound for Unique Games, that is there is no polynomial time algorithm for Unique Games assuming Exponential Time Hypothesis as further elaborated in Section B.

Conjecture 1.11. *There exists a reduction for some $2/3 < c < 3/4$ from 3SAT (of size n) to MAX-CUT $(1-N^{-c}, 1-N^{1-2c})$ of size N where $N = O(n^{1/c})$ with degree $d = \Theta(N^{1-c}\log N)$.*

[2] Any algorithm for 3SAT requires $2^{\Omega(n)}$-time

Conjecture 1.12. *There exists a reduction from 3SAT (of size n) to MAX-CUT* $(1-d^{-1}N^{1-2c}, 1-d^{-2}N^{2-2c})$ *of size N where* $N = O(n^{1/c})$ *with degree* $d = \tilde{\Omega}(N^{1-\frac{2c}{3}})$ *for some* $c < 3/4$.

Unfortunately, it has been pointed out that these conjectures are false due to upper bound results in low ε regimes. If ε is too small (as in less than $1/\text{poly}\log(N)$), then the MAX-CUT instance becomes easy to begin with. Therefore, to start from a conjecture that is not known to be false, we need improvements in the following explicit two directions.

Extension to Low Repetition Regime Analyzing the curve $\log(\mathbf{val}(\mathcal{G}^{\tilde{\otimes}r}))/r$ in terms of r is a fundamental open question. Via limit the lower bound of $-\Omega(\varepsilon)$ holds while [30] the upper bound of $-\Omega(\varepsilon^2)$ exists for all r, since symmetric parallel repetition is by definition stronger than the usual parallel repetition. The question is the speed of convergence to $-\Omega(\varepsilon)$. For practical purposes, r needs to be much smaller than n. Current bounds from Theorem 1.9 do not say anything about r much smaller than n. In particular, can we show $\mathbf{val}(\mathcal{G}^{\tilde{\otimes}2r}) \geq \mathbf{val}(\mathcal{G}^{\tilde{\otimes}r})^2$? Note that this is trivial for usual parallel repetition. We suspect that the graph expansion must come into the picture, since the statement is false for disjoint copies of odd cycle game. This can also be a candidate approach to improve bounds for symmetric repetition under constant d.

Extension to Constant Degree Regime One caveat to our results is that we require $d = \frac{n \log n}{k}$, instead of $d = O(1)$, as in many hardness results. Intuitively, higher degree regime should be easier than constant degree regime due to subsampling lemmas [2, 6]. These subsampling lemmas give randomized algorithms for distinguishing between YES and NO case that run in time $n^{O(n/d)}$. Since $d = o(n)$, this only gives super-polynomial time algorithm. Thus, there is still hope for obtaining a hard MAX-CUT instance with $d = o(n)$. But we suspect that it would be easier to obtain hard instances with $d = O(1)$. It would be nice to extend results on symmetric repetition to $d = O(1)$ towards resolving the Unique Games Conjecture.

2 Preliminary

2.1 Two Prover Game and Parallel Repetition

In this section, we formally define Two Prover One Round game and its repetition. Two Prover game consists of one verifier and two provers, Alice and Bob. The verifier draws (x, y) from some distribution over $\mathcal{X} \times \mathcal{Y}$ where X is the question set for Alice and Y the question set for Bob. Without loss of generality, one could view it as a bipartite graph where left set of vertices is X and right set of vertices is Y, and the distribution is a uniform distribution over some edge set E. Alice and Bob, depending on their respective input x and y answers $a \in A$ and $b \in B$, i.e. $a = f(x)$ and $b = g(y)$. Verifier then checks via a verification function $V : X \times Y \times A \times B \to \{0, 1\}$ which checks whether a and b are correct assignments for the edge (x, y). Value of the game $\mathcal{G} = (X, Y, V, E)$ is then defined as

$$\mathbf{val}(\mathcal{G}) = \max_{f,g} \frac{1}{|E|} \sum_{(x,y)\in E} V(x, y, f(x), g(y))$$

For two prover games, we are interested projection games and unique games. $\mathcal{G}$ is a *projection game* if for each $(x, y) \in E$, there exists a function $p_{(x,y)} : A \to B$ such that $V(x, y, a, b) = 1$ iff $p_{(x,y)}(a) = b$. Furthermore, we say $\mathcal{G}$ is a *unique game* if all $p_{(x,y)}$'s are permutations.

We formally define what it means to repeat the game in parallel r-times:

Definition 2.1 (Parallel Repetition)**.** *Let $\mathcal{G} = (X, Y, V, E)$ be a two-prover game. Then define $\mathcal{G}^{\otimes r}$ as the following game. Pick r random edges : $e_1 = (x_1, y_1), \ldots, e_r = (x_r, y_r)$. The referee sends $\vec{x} = (x_1, \ldots, x_r)$ to Alice and $\vec{y} = (y_1, \ldots, y_r)$ to Bob. Alice answers $\vec{a} = f(\vec{x})$ and Bob answers $\vec{b} = g(\vec{y})$. The referee replies with $\bigwedge_{i=1}^{r} V(x_i, y_i, a_i, b_i)$.*

One weakness of the parallel repetition is that the verification **is coordinate dependent**. Alice and Bob can somehow exploit this structure. In particular, [10] asked about the amortized behavior of the game and showed that this in tied to "how much information" you need to win one copy. [32] indeed exhibited such game. [7] and [37] extended [32] to show that only games that are "easy" in terms of Semi-definite programming obeys strong parallel repetition, ruling out attempts to prove the Unique Games Conjecture via amplifying the gap in MAX-CUT.

To overcome these impossibility results, we introduce a new notion of generalized parallel repetition which is set-based and coordinate free:

Definition 2.2 (Symmetric Parallel Repetition)**.** *Define $\mathcal{G}_{\mathcal{D}}$ as the following two prover game, where $\mathcal{D}$ is a distribution over $2^X \times 2^Y$ where $\mathsf{Supp}(\mathcal{D}) \subseteq 2^E$ and can be efficiently sampled.*

- *The referee pick a subset of edges $C \subseteq E$ according to $\mathcal{D}$*
- *Send $C_X = \{x | (x, y) \in C\}$ to Alice and $C_Y = \{y | (x, y) \in C\}$ to Bob.*
- *Alice and Bob returns the assignment to Z_X and Z_Y respectively.*
- *The referee replies with $\bigwedge_{(x,y) \in C} V(x, y, a_x, b_y)$, where a_x is assignment to x and b_y is assignment to y.*

It is easy to check that the uniqueness and projection property of the game is preserved under this transformation. With the two prover game defined, we can also define $\mathbf{val}(\mathcal{G}_{\mathcal{D}})$. Observe that the size of $\mathcal{G}_{\mathcal{D}}$ depends on $\mathcal{D}$, in particular $\mathsf{Supp}(\mathcal{D})$. If for all $C \in \mathsf{Supp}(\mathcal{D})$, $|C| \leq r$, or similarly $|C_X|, |C_Y| \leq r$ then the size of $\mathcal{G}_{\mathcal{D}}$ is $O(n^r)$. With Note that for "strong parallel repetition" to hold, one needs to exhibit that there exists $\mathcal{D}$ such that $\mathbf{val}(\mathcal{G}_{\mathcal{D}}) = \mathbf{val}(\mathcal{G})^{\Theta(r)}$.

Remark 2.3 ("Folklore" symmetric parallel repetition)**.** *In "folklore" definition of symmetric parallel repetition, $\mathcal{D}$ is a distribution where C is distributed uniformly over r-sized subsets of E.*

Remark 2.4 (Check all edges)**.** *We can strengthen the verification function without any penalty on the size by following: instead of $\bigwedge_{(x,y) \in C} V(x, y, a_x, b_y)$, check $\bigwedge_{(x,y) \in E \cap (C_X \times C_Y)} V(x, y, a_x, b_y)$. Since $C \subset E \cap (C_X \times C_Y)$, the later verification is stronger. But for our purpose, these two notions are equivalent.*

It has been long conjectured that if $\mathbf{val}(\mathcal{G}) < 1 - \varepsilon$ and $\mathcal{G}$ is a projection game, then $\mathbf{val}(\mathcal{G}_{\mathcal{D}}) \leq (1 - \varepsilon^c)^{\Omega(r)}$ where $c < 2$ for some $\mathcal{D}$ where for all $C \in \mathsf{Supp}(\mathcal{D})$, $|C| \leq r$, beating the bound of [30]. Also odd cycle game in [32] fails to be a counterexample for such repetition. We show that Theorem 1.9 implies that this is indeed the case for large r.

2.2 Birthday Repetition

To analyze the behavior of symmetric parallel repetition, we introduce the Birthday Repetition. This was first introduced in [1] as a tool to transform any general game into a free-game, that is a game where $E = X \times Y$, or Alice's input x and Bob's input y are independent. (It is not hard to see that these two descriptions are equivalent without loss of generality)

Definition 2.5 ('Original' Birthday Repetition [1])**.** *Define $\mathcal{G}^{k\times\ell}$ as the following game. The referee chooses k-sized random subset $S \subset X$, and ℓ-sized random subset $T \subset Y$. The referee sends S to Alice, T to Bob. Alice and Bob gives assignment to S and T respectively. The referee checks all edges in $E(S,T) := E \cap (S \times T)$.*

Note that $\mathcal{G}^{k\times\ell}$ is a free-game by design, since S and T are chosen independently at random. Also note that this is a set-based repetition of the game by design. [1] showed that if $k = \ell = O(\sqrt{n/\varepsilon})$, $\mathbf{val}(\mathcal{G}^{k\times\ell})$ is indeed upper-bounded by $\mathbf{val}(\mathcal{G})$. [27] asked the behavior of $\mathbf{val}(\mathcal{G}^{k\times\ell})$ if k and ℓ becomes larger than $O(\sqrt{n/\varepsilon})$, and indeed showed that the value decays exponentially $(1-\varepsilon^c)^{\Omega(r)}$ where c is some constant and $r = k\ell/n$.

With an abuse of notation, we redefine $\mathcal{G}^{k\times\ell}$ as following, which was first observed in [15] as a useful replacement for the original Birthday Repetition due to its concentration property.

Definition 2.6 ("Refined" Birthday Repetition [15])**.** *Define $\mathcal{G}^{k\times\ell}$ as the following game with a parameter $\delta > 0$. The referee picks a random subset S from X, picked via following process. Each variable $i \in X$ is added to S with probability $1 - e^{-k/|X|}$. Similarly, the referee picks T from Y by adding each $j \in Y$ with probaiblity $1 - e^{-\ell/|Y|}$. If $|S| < (1+\delta)\,\mathbb{E}[|S|]$ and $|T| < (1+\delta)\,\mathbb{E}[|T|]$, then referee sends S to Alice, T to Bob and checks $E(S,T)$.*

'Refined' Birthday Repetition could be viewed as having different independent distribution for S and T for $S \in 2^X$ and $T \in 2^Y$. In [1] definition of Birthday Repetition, the distribution is uniform over $S \in 2^X$ and $T \in 2^Y$ such that $|S| = k$ and $|T| = \ell$, while in our new definition, each element in $x \in X$ and $y \in Y$ are randomly picked to be included in S and T respectively. This new distribution simplifies proof steps in [1] and [12]. Indeed, in [15], the probability is $k/|X|$ and $\ell/|Y|$ respectively. But it is easy to check that these two probabilities are the same upto constant factor. This change is essentially made to make the analysis in Section E simpler.

3 Main Results

First, we reprove the main lower bound in [1] through "Refined" Birthday Repetition.

Theorem 1.1. *If $\mathbf{val}(\mathcal{G}) < 1 - 2\varepsilon$, then $\mathbf{val}(\mathcal{G}^{k\times\ell}) < 1 - \varepsilon$ for $k = \ell = \Omega(\sqrt{n\log(1/\varepsilon)})$.*

For high degree regime, we prove

Theorem 1.2. *Suppose $d = \Omega(n \log n/k)$, $k = \ell$ and $\mathbf{val}(\mathcal{G}) < 1 - \varepsilon$. Then*

$$\mathbf{val}(\mathcal{G}^{k\times\ell}) \le 1 - \frac{\varepsilon}{2}.$$

The proof of Theorem 1.1 and Theorem 1.2 is simple and straightforward. We append the full proof in Section F.

3.1 Repeated Birthday Repetition

With Theorem 1.1, we are ready to restate two technical theorems about repeated Birthday Repetition.

Theorem 1.3. *Suppose $k = \ell = \Omega(\sqrt{n\log(1/\varepsilon)})$, and $\mathcal{G}$ is a projection game. Then if $\mathbf{val}(\mathcal{G}) < 1 - \varepsilon$*

$$\mathbf{val}(\mathcal{G}^{rk\times r\ell}) \le \mathbf{val}((\mathcal{G}^{k\times\ell})^{\otimes r}) \le \left(1 - \frac{\varepsilon}{2}\right)^{\Omega(r)}.$$

The proof of Theorem 1.3 is rather immediate from [8] which we attach in Section F. The main observation is that *block-projection property* is preserved in Birthday Repetition. It it clear from the definition that even if $\mathcal{G}$ is a projection game, $\mathcal{G}^{k\times\ell}$ is not necessarily a projection game. The main observation is that the satisfying assignments still have a projection function on *some* coordinates. In fact, the pair S and T are satisfied if and only if for all $(x, y) \in E(S, T)$ $p_{(x,y)}(b_y) = a_x$ where b_y is the assignment to $y \in T$ and a_x is the assignment to $x \in S$. This suffices for modifying the proof of [8] to show that if the underlying game is a projection game, repeating the game r times in parallel obeys strong parallel repetition.

But note that repeating the game r-times will make the input set size S and T bigger by a factor of r, while the expected number of edges that will be checked by the verifier actually increase by a factor of r^2. Thus the expected number of edges grows quadratically in the number of repetition. Indeed this might suffice for many applications. Unfortunately, we need the exponent to be linear in the expected number of edges checked by the verifier to have desired consequences in symmetric parallel repetition. But we can improve the analysis to obtain a tighter bound.

Theorem 1.4. *Suppose* $k = \ell = \Omega(\sqrt{n\log(1/\varepsilon)})$*, and* $\mathcal{G}$ *is a projection game. Further, suppose* $r \le \min\{\sqrt{n}/2, \varepsilon n/320\}$*. Then if* $\mathbf{val}(\mathcal{G}) < 1 - \varepsilon$*,*

$$\mathbf{val}(\mathcal{G}^{rk\times r\ell}) \le (1 - \varepsilon/2)^{\Omega(r^2)}.$$

We avoid this square loss by the following observation: if we repeat r-copies in parallel, we get a tuple of length r, where each entry is set S_i and T_i for Alice and Bob respectively with $i \in [r]$. For each i, we are expected to check an edge with at least $(1 - \varepsilon)$-probability, since S_i and T_i have an edge between them with at least $(1-\varepsilon)$-probability. If we consider the ordering given by the index i, and only compare (S_i, T_i) pairs, this corresponds to 1.3. But in the actual Birthday Repetition, since it **has no ordering** inside the set, **we also check** (S_i, T_j) **pairs** as well where $i \neq j$.

We wish to argue how these pairs lowers the value of the game. The main technical challenge is that (S_i, T_j)-pairs are not r^2 independently chosen sets. For instance, (S_i, T_j) pair and $(S_{i'}, T_{j'})$ pair are no longer independently distributed if $i = i'$ or $j = j'$ while it is crucial that each coordinate is distributed independently at random in all parallel repetition proofs.

More precisely, in parallel repetition proofs initiated by [31], we condition on a subset of coordinates (the assignment and challenges) to extract too good to be true strategy. In our setting, we would like to condition on $\Omega(r^2)$-pairs, since the size of the subset corresponds to the exponent of the value of the game for parallel repetition proofs. But naively conditioning on the challenges is problematic for our instance. In an extreme case, for example, if we condition on all (S_i, T_i) for $i \in [r]$, all S_i and T_i's become fixed for $i \in [r]$. Here, we only conditioned on r-pairs, but this already fix the edges that will be checked between S_i and T_j is already determined from the conditioning. We cannot therefore extract any too good to be true single pair strategy.

Instead, we exploit the property of Birthday Repetition. We condition on an edge that will be checked by the (S_i, T_j) pair, that is the event $x_{(i,j)} \in S_i$ and $y_{(i,j)} \in T_j$ where $(x_{(i,j)}, y_{(i,j)}) \in E$. For $(i, j) \in [r] \times [r]$, we do not condition on the whole (S_i, T_j) but **only on the edge that will be conditioned on being checked**, and **the assignment to such** $x_{(i,j)}$ **and** $y_{(i,j)}$ (i.e. this is where we use the inherent structure of the game created by Birthday Repetition). Indeed even conditioning on these events will tilt the distribution, since S_i and T_j must contain those vertices as an element, thereby tilting the distributions on the other entries of the grids. But we show that this cannot tilt the distribution too much in terms of the divergence, if the edges are chosen carefully.

Now observe that in order to win the whole game in the Birthday Repetition, Alice and Bob must win all the edges that are conditioned to be in the sets. That is, satisfying the conditioned edges is a sub-event of winning all the edges. Since the probability of satisfying all the edges that are conditioned is low, the probability of winning the whole Birthday Repetition must be low as well. We append the full proof in Section F. Also with an analogous, we can also prove the following for high degree regime.

Theorem 1.5. *Suppose $k = \ell$, and $\mathcal{G}$ is a projection game, and $d = \Omega(\frac{n \log n}{k})$. Further, suppose $r \leq \min\{\sqrt{n}/2, \varepsilon n/320, d/2\}$. Then if* $\mathbf{val}(\mathcal{G}) < 1 - \varepsilon$,

$$\mathbf{val}(\mathcal{G}^{rk \times r\ell}) \leq (1 - \varepsilon/2)^{\Omega(r^2)}.$$

3.2 Translation Lemma and Corollaries

Lemma 1.6. *If the degree of the original Two Prover game $d \geq (1+\delta)n \log n/k$ with $\delta = \Omega(1/\log n)$ and $k = \ell$, then*

$$|\mathbf{val}(\mathcal{G}_{\mathcal{D}_{k,\ell}}) - \mathbf{val}(\mathcal{G}^{k \times \ell})| = o(1).$$

The main difference between the Birthday Repetition and the symmetric parallel repetition is that while for symmetric parallel repetition, all the vertices included in the set are checked by the verifier, it is not the case for the Birthday Repetition. There are "dummy" vertices that are not checked, thus it is not clear how well the strategy for the symmetric parallel repetition will perform in the Birthday Repetition and vice versa.

We translate the strategy for the symmetric parallel repetition to a strategy for the Birthday Repetition. The main observation is that if d, the degree of the graph, is large enough compared to k and ℓ, the expected size of the input sets, all of the vertices in the set even for the Birthday Repetition must be checked by the verifier. Conditioned on all the vertices being used, the game essentially becomes equivalent to symmetric parallel repetition. Then we obtain bound on the symmetric parallel repetition as a corollary of the bound on the Birthday Repetition. We attach the full proof in Section G.

Now combining Theorem 1.4 and Lemma 1.6, we get the following corollaries on the symmetric repetition.

Theorem 1.7. *Let $\mathcal{G}$ be a projection game where the underlying constraint graph is a d-regular graph. Let* $\mathbf{val}(\mathcal{G}) < 1 - \varepsilon$. *Set $k = \ell = 4n \log n/d$. If $k = \ell = \Omega(\sqrt{n \log(1/\varepsilon)})$ and $\frac{k^2}{n \log(1/\varepsilon)} \leq \min\{\sqrt{n}/2, \varepsilon n/320\}$ then there exists a distribution on $\mathcal{D}_{k,\ell}$ on 2^E that is only supported on the sets of size $O(k)$ such that*

$$\mathbf{val}(\mathcal{G}_{\mathcal{D}_{k,\ell}}) \leq (1 - \varepsilon/2)^{\Omega(\frac{k^2}{n \log(1/\varepsilon)})}.$$

Else, for any $r > 0$ that satisfies $r \leq \min\{\sqrt{n}/2, \varepsilon n/320, d/2\}$, there exists a distribution on $\mathcal{D}_{rk,r\ell}$ on 2^E that is only supported on the sets of size $O(rk)$ such that

$$\mathbf{val}(\mathcal{G}_{\mathcal{D}_{rk,r\ell}}) \leq (1 - \varepsilon/2)^{\Omega(r^2)}.$$

3.3 High Repetition Regime

Recall the high repetition regime version of the theorem.

Theorem 1.9. *Let $\mathcal{G}$ be a two prover game (not necessarily a projection game) with* $\mathbf{val}(\mathcal{G}) < 1 - \varepsilon$. *If $r = \Omega(n \log(1/\varepsilon))$ then there exists $\mathcal{D}_r$ on 2^E that is only supported on edge sets of size at most $2r$ such that* $\mathbf{val}(\mathcal{G}_{\mathcal{D}_r}) = (1 - \varepsilon)^{\Omega(r)}$. *Similarly, if* $\mathbf{val}(\mathcal{G}) > 1 - \varepsilon$, *then* $\mathbf{val}(\mathcal{G}_{\mathcal{D}_r}) = (1 - \varepsilon)^{O(r)}$

We append the full proof in Section H. The main idea is to take $\mathcal{D}_r$ as the following distribution on 2^E. We choose the set of edges C by picking each $e \in E$ with probability $1 - e^{-r/|E|}$. Indeed, it might be the case that the number of edges picked might be too large. We prune the cases where $|C| > 2r$, and call such distribution $\overline{\mathcal{D}}_r$. This distribution will satisfy the condition in Theorem 1.9.

Note that if the set size becomes exactly n, the usual Birthday Repetition will give a value 0 game, since the verifier checks all the edges, unless the original game was completely satisfiable. Suppose the verifier checks r random edges instead. It is easy to check that the value of such game is at most $(1-\varepsilon)^r$ because the input does not give any information. We show that if sufficiently many edges are picked, a naive guess by Alice and Bob will guess the actual vertices that are checked by the verifier correctly (note that the entropy converges to 0). And if they guess correctly, the setting becomes exactly the same as symmetric parallel repetition. We show that conditioned on guessing the actual inputs correctly, the strong parallel repetition holds as well. Using an analogous arguments from Lemma 1.6, we show bounds for $\mathbf{val}(\mathcal{G}_{\mathcal{D}_r})$.

The interesting consequence of this theorem is a dichotomy between symmetric parallel repetition and the conventional parallel repetition which can be summarized as the following table.

	Repetition Model	Repeated Value
Projection Game [30, 33]	$\mathcal{G}^{\otimes r}$	$(1-\varepsilon^2)^{\Theta(r)}$
General Game [21, 9, 20]	$\mathcal{G}^{\otimes r}$	$(1-\varepsilon^3)^{r/\log(\lvert A\rvert+\lvert B\rvert)}$
Theorem 1.9	$\mathcal{G}_{\mathcal{D}_r}$	$(1-\varepsilon)^{\Theta(r)}$

Table 1: Amortized Values

Not only $\mathcal{G}_{\mathcal{D}_r}$ improves the bound for the projection case, it also removes the dependence on the alphabet size in the general case as seen in [21, 9, 20]. It would be interesting to see whether such bound holds for smaller r, instead of $r = \Omega(n \log(1/\varepsilon))$.

4 Acknowledgement

The author would like to thank Mark Braverman and Ankit Garg for their useful discussions.

References

[1] Scott Aaronson, Russell Impagliazzo, and Dana Moshkovitz. Am with multiple merlins. In *2014 IEEE 29th Conference on Computational Complexity (CCC)*, pages 44–55. IEEE, 2014.

[2] Noga Alon, W Fernandez De La Vega, Ravi Kannan, and Marek Karpinski. Random sampling and approximation of max-csps. *Journal of computer and system sciences*, 67(2):212–243, 2003.

[3] Sanjeev Arora, Subhash A Khot, Alexandra Kolla, David Steurer, Madhur Tulsiani, and Nisheeth K Vishnoi. Unique games on expanding constraint graphs are easy. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 21–28. ACM, 2008.

[4] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM*, 45, 1998.

[5] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of NP. *Journal of the ACM*, 45(1):70–122, January 1998.

[6] Boaz Barak, Moritz Hardt, Thomas Holenstein, and David Steurer. Subsampling mathematical relaxations and average-case complexity. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, pages 512–531. Society for Industrial and Applied Mathematics, 2011.

[7] Boaz Barak, Ishay Haviv, Moritz Hardt, Anup Rao, Oded Regev, and David Steurer. Rounding parallel repetitions of unique games. In *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science*. IEEE Computer Society, 2008.

[8] Boaz Barak, Anup Rao, Ran Raz, Ricky Rosen, and Ronen Shaltiel. Strong parallel repetition theorem for free projection games. *RANDOM*, 2009.

[9] Mark Braverman and Ankit Garg. Small value parallel repetition for general games. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, STOC '15, pages 335–340, New York, NY, USA, 2015. ACM. URL: `http://doi.acm.org/10.1145/2746539.2746565`, `doi:10.1145/2746539.2746565`.

[10] Mark Braverman and Young Kun Ko. Information value of the game. *Manuscript*, 2016.

[11] Mark Braverman, Young Kun Ko, Aviad Rubinstein, and Omri Weinstein. Eth hardness for densest-k-subgraph with perfect completeness. *arXiv preprint arXiv:1504.08352*, 2015.

[12] Mark Braverman, Young Kun Ko, and Omri Weinstein. Approximating the best nash equilibrium in no (log n)-time breaks the exponential time hypothesis. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 970–982. SIAM, 2015.

[13] André Chailloux and Giannicola Scarpa. Parallel repetition of entangled games with exponential decay via the superposed information cost. *41st International Colloquium on Automata, Languages and Programming*, 2014.

[14] Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Near-optimal algorithms for unique games. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, 2006.

[15] Yu Cheng and Young Kun Ko. Finding best mixture of nash is hard. *Manuscript*, 2017.

[16] Irit Dinur. The pcp theorem by gap amplification. *Journal of the ACM (JACM)*, 54(3):12, 2007.

[17] Irit Dinur, Prahladh Harsha, Rakesh Venkat, and Henry Yuen. Multiplayer parallel repetition for expander games. *arXiv preprint arXiv:1610.08349*, 2016.

[18] Irit Dinur and David Steurer. Analytical approach to parallel repetition. *46th Annual Symposium on the Theory of Computing*, 2014.

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

[20] Uriel Feige and Oleg Verbitsky. Error reduction by parallel repetition–A negative result. *Combinatorica*, 22, 2002.

[21] Thomas Holenstein. Parallel repetition: Simplifications and the no-signaling case. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, 2007.

[22] 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.

[23] Julia Kempe and Thomas Vidick. Parallel repetition of entangled games. *43rd annual ACM symposium on Theory of computing*, 2011.

[24] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing*, pages 767–775, 2002. URL: `http://doi.acm.org/10.1145/509907.510017`.

[25] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for max-cut and other 2-variable CSPs? In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, 2004.

[26] Pasin Manurangsi. Almost-polynomial ratio eth-hardness of approximating densest k-subgraph. *arXiv preprint arXiv:1611.05991*, 2016.

[27] Pasin Manurangsi and Prasad Raghavendra. A birthday repetition theorem and complexity of approximating dense csps. 07 2016. URL: `https://arxiv.org/abs/1607.02986`, `arXiv:1607.02986`.

[28] Dana Moshkovitz. Parallel repetition of fortified games. *Electronic Colloquium on Computational Complexity (ECCC)*, 2014.

[29] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low in.uences invariance and optimality. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, pages 21–30. IEEE Computer Society, 2005. URL: `http://doi.ieeecomputersociety.org/10.1109/SFCS.2005.53`.

[30] Anup Rao. Parallel repetition in projection games and a concentration bound. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, 2008.

[31] Ran Raz. A parallel repetition theorem. *SIAM Journal on Computing*, 27(3):763–803, June 1998. Prelim version in STOC '95.

[32] Ran Raz. A counterexample to strong parallel repetition. In *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science*. IEEE Computer Society, 2008.

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

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

[35] Aviad Rubinstein. Eth-hardness for signaling in symmetric zero-sum games. *arXiv preprint arXiv:1510.04991*, 2015.

[36] Aviad Rubinstein. Detecting communities is hard, and counting them is even harder. *arXiv preprint arXiv:1611.08326*, 2016.

[37] David Steurer. Improved rounding for parallel repeated unique games. In *Proceedings of the 13th International Conference on Approximation, and 14 the International Conference on Randomization, and Combinatorial Optimization: Algorithms and Techniques*, APPROX/RANDOM'10, pages 724–737, Berlin, Heidelberg, 2010. Springer-Verlag. URL: `http://dl.acm.org/citation.cfm?id=1886521.1886577`.

[38] Madhur Tulsiani, John Wright, and Yuan Zhou. Optimal strong parallel repetition for projection games on low threshold rank graphs. *ICALP*, 2014.

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

[40] Henry Yuen. A parallel repetition theorem for all entangled games. 04 2016. URL: `https://arxiv.org/abs/1604.04340`, `arXiv:1604.04340`.

A Information Theory

In this section, we provide backgrounds on information theory that will be used to prove main results but not necessary to understand the statement of the result. We remark that throughout the paper, log is of base 2 and ln is of base e.

First we describe the notion of amount of information needed to describe a random variable.

Definition A.1. *Let X be a random variable. Then the entropy of X is*

$$H(X) = \sum_{x} \Pr[X = x] \cdot \log \frac{1}{\Pr[X = x]}$$

Note that the following naive upper bound on the entropy holds.

Fact A.2. *Let $\mathcal{X}$ be t he support set of X. Then*

$$H(X) \leq \log |\mathcal{X}|$$

In particular, this implies that the uniform distribution has the highest entropy.

Since we will be mainly talking about X on different distributions, we introduce the following notion of distance between two different distributions.

Definition A.3 (Kullback-Leiber Divergence)**.** *Given two probability distributions μ_1 and μ_2 on the same sample space Ω such that $(\forall \omega \in \Omega)(\mu_2(\omega) = 0 \Rightarrow \mu_1(\omega) = 0)$, the* Kullback-Leibler Divergence *between is defined as (also known as relative entropy)*

$$D(\mu_1 || \mu_2) = \sum_{\omega \in \Omega} \mu_1(\omega) \log \frac{\mu_1(\omega)}{\mu_2(\omega)}.$$

One nice property of the divergence is the chain rule:

Fact A.4 (Chain Rule)**.** *Consider two distributions $P(x, y)$ and $Q(x, y)$. Then*

$$D(P(x, y) || Q(x, y)) = D(P(x) || Q(x)) + \mathop{\mathbb{E}}_{x \sim P} [D(P(y|x) || Q(y|x))]$$

We will use the following corollary of Fact A.4.

Fact A.5. *Consider two distributions P and Q, where Q is a product distribution, that is $Q(x_1, \ldots, x_n) = Q(x_1) \ldots Q(x_n)$. Then*

$$D(P(x_1, \ldots, x_n) || Q(x_1, \ldots, x_n)) \geq \sum_{i=1}^{n} D(P(x_i) || Q(x_i))$$

We also use the following fact from [8].

Fact A.6 (Corollary 3.4 of [8])**.** *Let P and Q be probability distribution. If $D(P||Q) < \delta$ and $P(T) < \delta$, then $Q(T) < 4\delta$.*

B Implications for the Unique Games Conjecture

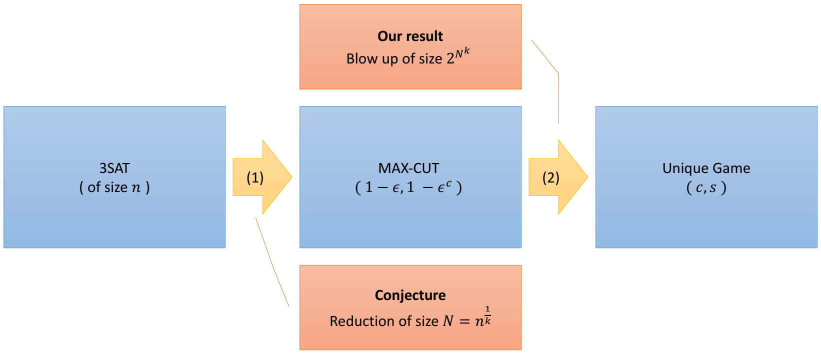
Recall that the Unique Games Conjecture is stated as following.

Conjecture B.1 (Unique Games Conjecture, [24])**.** *For every $\varepsilon, \delta > 0$, there exists $n = n(\varepsilon, \delta)$, such that there exists a unique game instance $\mathcal{G}$ such that it is NP-hard to distinguish between (YES)* $\mathbf{val}(\mathcal{G}) \geq 1 - \varepsilon$*; (NO)* $\mathbf{val}(\mathcal{G}) \leq \delta$*.*

Instead of trying to prove the full conjecture, we instead focus on the runtime lower bound.

Conjecture B.2 (Weak Unique Games Conjecture)**.** *For every $\varepsilon, \delta > 0$, there exists $n = n(\varepsilon, \delta)$, such that there exists a unique game instance $\mathcal{G}$ such that there is* **no polynomial time algorithm** *that distinguishes between (YES)* $\mathbf{val}(\mathcal{G}) \geq 1 - \varepsilon$*; (NO)* $\mathbf{val}(\mathcal{G}) \leq \delta$*.*

Towards showing runtime lower bound (assuming Exponential Time Hypothesis) the following approach has been suggested.



Our result
Blow up of size 2^{N^k}
3SAT
(of size n)
(1)
MAX-CUT
($1-\epsilon, 1 - \epsilon^c$)
(2)
Unique Game
(c, s)
Conjecture
Reduction of size $N = n^{\frac{1}{k}}$

Figure 1: Overview of an approach towards Conjecture B.2

In this section, we show how our technical results imply bounds on the second part of this approach, that is amplifying the MAX-CUT hardness to achieve required hardness for Unique Games, that is $(1 - \varepsilon, \delta)$ for any constant $\varepsilon, \delta > 0$.

For completeness part of the reduction, we prove the following theorem whose proof we attach in Section G.

Theorem 1.8. *If $\mathcal{G}$ is a projection game where the underlying constraint graph is d-regular graph. Let* $\mathbf{val}(\mathcal{G}) > 1 - \varepsilon$*, $k, \ell = o(n)$ and $\frac{\varepsilon d k \ell}{n} \leq 1/8$. Then*

$$\mathbf{val}(\mathcal{G}_{\mathcal{D}_{k,\ell}}) \geq e^{-O\left(\frac{\varepsilon d k \ell}{n}\right)}.$$

We show two theorems that follow immediately from our technical results.

Theorem B.3. *For $2/3 < c < 3/4$, there exists a $2^{\tilde{O}(n^c)}$-sized reduction from MAX-CUT $(1 - n^{-c} \log^{-1} n, 1 - n^{1-2c} \log n)$ to Unique Game $(1 - \varepsilon, \delta)$ for any constant ε and δ.*

Proof. Let $\mathcal{G}$ be the original MAX-CUT instance. Then consider $\mathcal{G}_{\mathcal{D}_{k,\ell}}$ and its value with $k = \ell = n^c$.

For completeness, we use Theorem 1.8. Recall that $d = \frac{\alpha n \log n}{n^c}$ for some constant α. In (YES) case,

$$\mathbf{val}(\mathcal{G}_{\mathcal{D}_{k,\ell}}) \geq e^{-O\left(\frac{n^{-c}(\log^{-1} n)(dk\ell)}{n}\right)} = e^{-O(1)}$$

For soundness, we use Theorem 1.7. In (NO) case, since $c < 3/4$,

$$\mathbf{val}(\mathcal{G}_{\mathcal{D}_{k,\ell}}) \leq (1 - n^{1-2c} \log n)^{\Omega(\frac{n^{2c}}{(2c-1)n \log n})} = e^{-\Omega(\frac{1}{2c-1})} \leq e^{-\Omega(1)}$$

where the last inequality holds since we assume $c > 2/3$. □

Theorem B.4. *There exists a $2^{\tilde{O}(n^c)}$-sized reduction from MAX-CUT $(1 - \frac{1}{dn^{2c-1}}, 1 - \frac{n^{2-2c} \log^2 n}{d^2})$ to Unique Game $(1 - \varepsilon, \varepsilon)$ for any constant ε if $d < n^{4/3-c} \log n$.*

Proof. The proof is similar to that of Theorem B.3, but instead uses the second part of Theorem 1.7. Note that the completeness follows from design. For soundness,

$$\mathbf{val}(\mathcal{G}_{\mathcal{D}_{k,\ell}}) \leq (1 - \frac{n^{2-2c} \log^2 n}{d^2})^{\Omega(\left(\frac{dn^c}{4n \log n}\right)^2)} \leq e^{-\Omega(1)}$$

where the inequality holds since we assume $d < n^{4/3-c} \log n$. □

C Proof of Theorem 1.1

In this section, we simplify and improve the parameter given in lower bound in [1] using 'Refined' Birthday Repetition which first appeared in [15]. In particular, we prove Theorem 1.1

Claim C.1. *Consider 'Refined' Birthday Repetition $\mathcal{G}^{k \times \ell}$. If $k = \ell = \Omega\left(\sqrt{n \log(1/\varepsilon)}\right)$. then $\Pr_{S,T}[E \cap (S \times T) = \emptyset] < \varepsilon$*

Proof.

$$\begin{aligned}\Pr_{S,T}[E \cap (S \times T) = \emptyset] &= \mathop{\mathbb{E}}_{S}\left[\Pr_{T}[E \cap (S \times T) = \emptyset]\right] \leq \mathop{\mathbb{E}}_{S}\left[(e^{-\ell/|Y|})^{|\mathcal{N}(S)|}\right] + e^{-\Theta(\delta^2)\sqrt{n}} \\ &\leq \mathop{\mathbb{E}}_{S}\left[e^{-\ell \cdot |S|/|Y|}\right] + e^{-\Theta(\delta^2)\sqrt{n}} \leq e^{-\ell(1-\delta)\,\mathbb{E}_S[|S|]/|Y|} + e^{-\Theta(\delta^2)\sqrt{n}}\end{aligned}$$

The first inequality holds by ℓ_1 norm bound from conditioning on the size of S and T. The second inequality holds by assuming that any set of size k on one side has at least k neighbors which we can assume without loss of generality, and the third inequality is the assumption on the size of S. Now we bound $\mathbb{E}_S[|S|]$:

$$\mathop{\mathbb{E}}_{S}[|S|] = |X| \cdot (1 - e^{-k/|X|}) \geq |X| \cdot \Omega(k/|X|) = \Omega(k) = k/\alpha$$

for some constant α. Plugging this bound and setting $k = \ell = \sqrt{\frac{2\alpha n \log(1/\varepsilon)}{1-\delta}}$, we have

$$\Pr_{S,T}[E \cap (S \times T) = \emptyset] \leq e^{-\frac{(1-\delta)k\ell}{\alpha|Y|}} = e^{-\frac{\alpha n \log(1/\varepsilon)}{\alpha n}} = \varepsilon.$$

□

Now we restate and prove Theorem 1.1.

Theorem 1.1. *If* $\mathbf{val}(\mathcal{G}) < 1 - 2\varepsilon$*, then* $\mathbf{val}(\mathcal{G}^{k\times\ell}) < 1 - \varepsilon$ *for* $k = \ell = \Omega(\sqrt{n\log(1/\varepsilon)})$.

Proof. We prove by decomposing $\Pr_{S,T}[\text{ satisfy } (S,T)]$:

$$\begin{aligned}\Pr_{S,T}[\text{ satisfy } (S,T)] &= \Pr_{S,T}[E\cap(S\times T)=\emptyset]\cdot\Pr_{S,T}[\text{ satisfy } (S,T)\mid E\cap(S\times T)=\emptyset] \\ &\quad + \Pr_{S,T}[E\cap(S\times T)\neq\emptyset]\cdot\Pr_{S,T}[\text{ satisfy } (S,T)\mid E\cap(S\times T)\neq\emptyset] \\ &\le \varepsilon\cdot 1 + \Pr_{S,T}[\text{ satisfy } (S,T)\mid E\cap(S\times T)\neq\emptyset] \\ &\le \varepsilon + (1-2\varepsilon) = 1-\varepsilon.\end{aligned}$$

The first inequality holds by Claim C.1. Now the second probability holds by our assumption on $\mathbf{val}(\mathcal{G})$, that is

$$\Pr_{S,T}[\text{ satisfy } (S,T)\mid E\cap(S\times T)\neq\emptyset] \le \mathbf{val}(\mathcal{G}) < 1-2\varepsilon.$$

since the distribution over a single edge in $E\cap(S\times T)$ is uniform. □

It is noteworthy that $k = \ell = \Omega(\sqrt{n\log(1/\varepsilon)})$ which is smaller than $\Omega(\sqrt{n/\varepsilon})$ as in [1].

D Proof of Theorem 1.2

In this section, we prove Theorem 1.2, which is Birthday Repetition in high degree regime. It is noteworthy that there is no lower bound restriction on k and ℓ unlike in Theorem 1.1.

Claim D.1. *Consider 'Refined' Birthday Repetition* $\mathcal{G}^{k\times\ell}$*. If* $k = \ell$ *and* $d = \Omega(n\log n/k)$*, then* $\Pr_{S,T}[E\cap(S\times T)=\emptyset] < \varepsilon$

Proof.

$$\begin{aligned}\Pr_{S,T}[E\cap(S\times T)=\emptyset] &= \mathop{\mathbb{E}}_{S}\left[\Pr_{T}[\bigwedge_{i\in S} E\cap(\{i\}\times T)=\emptyset]\right] \\ &\le \mathop{\mathbb{E}}_{S}\left[\Pr_{T}[E\cap(\{i\}\times T)=\emptyset]\right] \le (e^{-\ell/n})^d + e^{-\Theta(\delta k)} \le \varepsilon\end{aligned}$$

The second inequality holds by ℓ_1 norm bound from conditioning on the size of S and T. The last inequality holds by setting $\delta = \Omega(\log(1/\varepsilon))$.[3] □

Theorem 1.2. *Suppose* $d = \Omega(n\log n/k)$*,* $k = \ell$ *and* $\mathbf{val}(\mathcal{G}) < 1-\varepsilon$*. Then*

$$\mathbf{val}(\mathcal{G}^{k\times\ell}) \le 1 - \frac{\varepsilon}{2}.$$

Proof. Again we prove by decomposing $\Pr_{S,T}[\text{ satisfy } (S,T)]$:

$$\begin{aligned}\Pr_{S,T}[\text{ satisfy } (S,T)] &= \Pr_{S,T}[E\cap(S\times T)=\emptyset]\cdot\Pr_{S,T}[\text{ satisfy } (S,T)\mid E\cap(S\times T)=\emptyset] \\ &\quad + \Pr_{S,T}[E\cap(S\times T)\neq\emptyset]\cdot\Pr_{S,T}[\text{ satisfy } (S,T)\mid E\cap(S\times T)\neq\emptyset]\end{aligned}$$

[3] Note that $\delta = O(\log n)$ since the number of edges is at most n^2.

$$
\begin{aligned}
&\leq \varepsilon \cdot 1 + \Pr_{S,T}[\text{ satisfy } (S,T) \mid E \cap (S \times T) \neq \emptyset] \\
&\leq \varepsilon + (1 - 2\varepsilon) = 1 - \varepsilon.
\end{aligned}
$$

The first inequality holds by Claim D.1. Now the second probability holds by our assumption on $\mathbf{val}(\mathcal{G})$, that is

$$\Pr_{S,T}[\text{ satisfy } (S,T) \mid E \cap (S \times T) \neq \emptyset] \leq \mathbf{val}(\mathcal{G}) < 1 - 2\varepsilon.$$

since the distribution over a single edge in $E \cap (S \times T)$ is uniform. □

E Proof of Theorem 1.3

The main technical challenge is that Birthday Repetition does not preserve the projection property of the game, i.e. Alice's answer does not no longer force a single correct answer to Bob. One might consider using parallel repetition for general games. However, note that Birthday Repetition blows up the size of the alphabet as well. In particular, since $k = \ell = \Omega(\sqrt{n \log(1/\varepsilon)})$, the alphabet size also becomes $s^{\Omega(\sqrt{n \log(1/\varepsilon)})}$ where s is the alphabet size of the original projection game.

However, we note that **the edge and vertices that actually matter** preserves the projection property, i.e. it maintains "block projection property". It can be formally stated as following.

Definition E.1 (Block Projection Game)**.** *$\mathcal{G}$ is a block projection game if for every $(x, y) \in E$, there exists a partition over A and B, say $\mathcal{P}_A = \{A_1, \ldots, A_p\}$ and $\mathcal{P}_B = \{B_1, \ldots B_p\}$ and $p_{(x,y)}$: $\mathcal{P}_B \to \mathcal{P}_A$ such that $V(x, y, a, b) = 1$ iff $p_{(x,y)}(\mathcal{P}_B(b)) = \mathcal{P}_A(a)$ where $\mathcal{P}_A(a)$ is A_i such that $a \in A_i$ similarly for $\mathcal{P}_B(b)$.*

Being a block projection game suffices for the proof of parallel repetition for projection game to avoid $\log s$ factor in the exponent. It is easy to observe that Birthday Repetition preserves "block projection," though it does not preserve projection. Let $c_a(S,T)$ denote the partition on A with the input (S,T). Similarly let $c_b(S,T)$ denote the partition on B. With an abuse of notation, we denote $a_{c_a(S,T)}$ as the assignment by Alice on $c_a(S,T)$ and $A_{c_a(S,T)}$ as the corresponding random variable. We similarly define $b_{c_b(S,T)}$ and $B_{c_b(S,T)}$. We prove the following claim which replaces Proposition 3.8 and Corollary 3.9 of [8].

Claim E.2. *Let W be the event where Alice and Bob win all the games in the last k coordinates. Define E as the set of blocks $((b_{c_b(S_1,T_1)}, \ldots, b_{c_b(S_k,T_k)}), (S_1, \ldots, S_k), (T_1, \ldots, T_k))$ such that $\Pr[(b_{c_b(S_1,T_1)}, \ldots, b_{c_b(S_k,T_k)}) \mid \mathcal{S}^k = S^k, \mathcal{T}^k = T^k] \geq 2^{-\varepsilon(n-k)/16}$. With an abuse of notation, define E as the event that lies in the set. Then $\Pr[W'] > \Pr[W] - 2^{-\varepsilon(r-k)/16}$ where $W' = W \wedge E$*

Proof. We generalize the argument in Corollary 3.9 of [8]. Note that we can rewrite the winning probability for a single copy of Birthday Repetition as

$$\Pr[W] = \mathop{\mathbb{E}}_{S,T}\left[\sum_{b_{c_b(S,T)}} \Pr[A_{c_a(S,T)} = p_{(S,T)}(b_{c_b(S,T)}) | S] \cdot \Pr[B_{c_b(S,T)} = b_{c_b(S,T)} | T]\right].$$

where $p_{S,T}$ refers to the projection function on the set (S,T). Similarly, for repeated game, we can rewrite as

$$\Pr[W] = \mathop{\mathbb{E}}_{S^r,T^r} \sum_{\vec{b} \in B^{\otimes r}} \Pr[B^1_{c_b(S_1,T_1)} = b^1_{c_b(S_1,T_1)}, \ldots, B^r_{c_b(S_r,T_r)} = b^r_{c_b(S_r,T_r)} | T^r]$$

$$\cdot \Pr[A^1_{c_a(S_1,T_1)} = p_{(S_1,T_1)}(b^1_{c_b(S_1,T_1)}), \ldots, A^r_{c_b(S_r,T_r)} = p_{(S_r,T_r)}(b^r_{c_b(S_r,T_r)})|S^r]$$

where A^k_s refers to the assignment on k-th copy of the game. Now for $\overline{E}$, we can write the probability as

$$\begin{aligned}
\Pr[W \wedge \overline{E}] = & \mathop{\mathbb{E}}_{S^r,T^r} \sum_{\vec{b} \in B^{\otimes r}} \Pr[B^1_{c_b(S_1,T_1)} = b^1_{c_b(S_1,T_1)}, \ldots, B^r_{c_b(S_r,T_r)} = b^r_{c_b(S_r,T_r)}|T^r] \\
& \cdot \Pr[A^1_{c_a(S_1,T_1)} = p_{(S_1,T_1)}(b^1_{c_b(S_1,T_1)}), \ldots, A^r_{c_b(S_r,T_r)} = p_{(S_r,T_r)}(b^r_{c_b(S_r,T_r)})|S^r] \\
< & \mathop{\mathbb{E}}_{S^r,T^r} \left[\sum_{\vec{b} \in B^{\otimes r}} \Pr[B^1_{c_b(S_1,T_1)} = b_1, \ldots, B^r_{c_b(S_r,T_r)} = b_r|T^r] \cdot 2^{-\varepsilon(r-k)/16} \right] \\
\leq & \; 2^{-\varepsilon(r-k)/16}
\end{aligned}$$

First equality holds since Alice's answer only depends on S^r and vice versa. The inequality follows from the property of E. □

Remark E.3. *Proposition 3.8 and Corollary 3.9 of [8] is the only place where projection property is used. Thus Claim E.2 shows that block projection property suffices for strong parallel repetition. Thus if $\mathcal{G}$ is a free, block projection game with $\mathbf{val}(\mathcal{G}) < 1 - \varepsilon$ then $\mathbf{val}(\mathcal{G}^{\otimes r}) \leq (1-\varepsilon)^{\Omega(r)}$.*

Proof of Theorem 1.3. First we show $\mathbf{val}(\mathcal{G}^{rk \times r\ell}) \leq \mathbf{val}((\mathcal{G}^{k \times \ell})^{\otimes r})$. Consider answering $(\mathcal{G}^{k \times \ell})^{\otimes r}$ in a following manner: Take union of $S_0 = \bigcup_{i=1}^r S_i$. Similarly take union of $T_0 = \bigcup_{i=1}^r T_i$. Then take the strategy from $\mathcal{G}^{rk \times r\ell}$ to answer $(\mathcal{G}^{k \times \ell})^{\otimes r}$. Then note that the distributions match since

$$1 - (1 - (1 - e^{-k/n}))^r = 1 - e^{-rk/n}$$

Thus $\mathbf{val}(\mathcal{G}^{rk \times r\ell}) \leq \mathbf{val}((\mathcal{G}^{k \times \ell})^{\otimes r})$.

The second inequality holds by Remark E.3 along with $\mathbf{val}(\mathcal{G}^{k \times \ell}) \leq 1 - \varepsilon/2$ via Theorem 1.1 Thus we have $\mathbf{val}((\mathcal{G}^{k \times \ell})^{\otimes r}) \leq (1 - \varepsilon/2)^{\Omega(r)}$ □

Corollary E.4. *If $k = \ell > \Omega(\sqrt{n \log(1/\varepsilon)})$,*

$$\mathbf{val}(\mathcal{G}^{k \times \ell}) \leq (1-\varepsilon)^{\Omega(\sqrt{\frac{k\ell}{n \log(1/\varepsilon)}})}$$

Proof. Let $k = \ell = r \cdot \sqrt{n \log(1/\varepsilon)}$. Then $r = \sqrt{\frac{k\ell}{n \log(1/\varepsilon)}}$. Then apply Theorem 1.3. □

This is practically stronger than [27], since to achieve any constant gap, we require $\sqrt{\frac{k\ell}{n \log(1/\varepsilon)}} = \Omega(1/\varepsilon)$, i.e. $k = \ell = \sqrt{n \log(1/\varepsilon)/\varepsilon}$ instead of $\frac{k\ell}{n} = \Omega(1/\varepsilon^c)$ where $c > 6$ for [27]. But we can further improve this by coming up with the sub-distribution that the verifier checks in the birthday repetition.

Also for high degree regime, we get the following corollary

Corollary E.5. *If $k = \ell$, $d = \Omega(\frac{n \log n}{K})$ and $\mathbf{val}(\mathcal{G}) < 1 - \varepsilon$ then*

$$\mathbf{val}(\mathcal{G}^{k \times \ell}) \leq (1-\varepsilon)^{\Omega(k/K)}$$

Proof. Let $r = k/K$. Note that $\mathbf{val}(\mathcal{G}^{K \times K}) \leq 1 - \varepsilon/2$ Then apply Theorem 1.3. □

F Proof of Theorem 1.4

In this section, we prove lemmas necessary towards proving $(1-\varepsilon)^{r^2}$ upper bound for the value of symmetric parallel repetition. Towards this goal, first we improve the bound in Section E from $(1-\varepsilon)^{\Omega(r)}$ to $(1-\varepsilon)^{\Omega(r^2)}$ where $r = \sqrt{\frac{k\ell}{n\log(1/\varepsilon)}}$ for the low degree regime. Using similar technique we also improve the bound for $d = \Omega(\frac{n\log n}{K})$ (i.e. the high degree regime) where $K = \Omega(\log n)$ to $(1-\varepsilon)^{(k/K)^2}$.

F.1 Overview

Recall that towards proving Theorem 1.3, for $k = \ell > \Omega(\sqrt{n\log(1/\varepsilon)})$, we viewed it as a parallel repetition r-times, that is $S_1, \dots S_r$ and $T_1, \dots T_r$. But this is not tight since we only check the clauses between S_i, T_i but not S_i, T_j where $i \neq j$. To analyze the contribution of S_i, T_j's, instead of viewing it as r-copies of (S_i, T_i), we view it as a $r \times r$ grid where each row corresponds to S_i and each column corresponds to T_j. For each $(i,j) \in [r] \times [r]$, the verifier checks (S_i, T_j) pair.

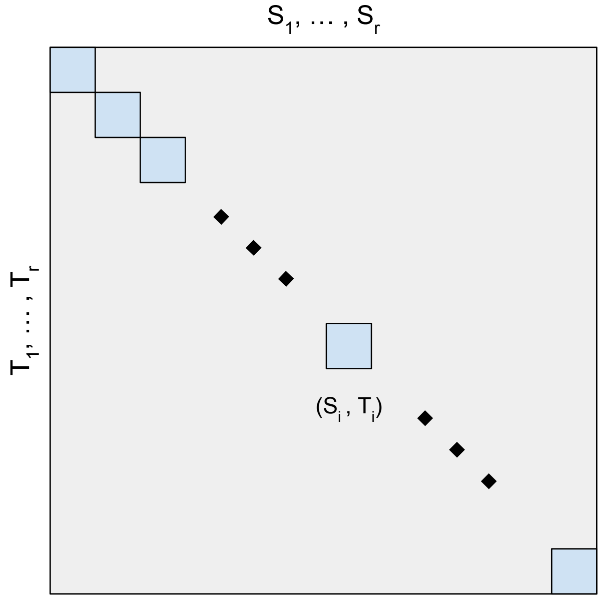


Figure 2: Overview of Theorem 1.3 : the verifier only checks blue grids

Unlike in the usual parallel repetition, challenges in each cell are not picked independently at random anymore. For instance, S_i, T_j and $S_{i'}, T_{j'}$ are not independent if $i = i'$ or $j = j'$, since we are recycling S_i and T_j's. Suppose we follow the approaches in proving the usual parallel repetition i.e. [31], which conditions on a subset of coordinates. Then conditioning on any linear number of S_i, T_j's makes the divergence between the original distribution huge. In an extreme setting, consider conditioning on all S_i, T_i's for $i \in [r]$. Note that we only conditioned on $1/r$-fraction of the coordinates, but this already fixes all the challenges.

Instead of conditioning on the actual challenges and the correct block, we actually exploit further the fact that these games are formed by Birthday Repetition, which indeed is not the case for [8]. We say **conditioning on a grid** (i,j) as conditioning with an edge that would be checked between S_i, T_j pair. In other words, instead of fixing whole S_i and T_j, we just condition $(x,y) \in E(S_i, T_j)$

or equivalently $x \in S_i$ and $y \in T_j$. Indeed, we must condition on the assignments as well. In case of the assignments, we **only consider the assignments to x and y**, which is along the lines of the proof of Theorem 1.3.

For high degree regime, the proof is essentially the same. But instead of breaking the set size down to $\mathbb{E}\left[|S_i|\right] = \sqrt{n \log(1/\varepsilon)}$, we break it into $\mathbb{E}\left[|S_i|\right] = K$ if $d = \Omega(\frac{n \log n}{K})$.

Notation Let R denote the subset of coordinates i.e. $R \subset [r] \times [r]$. $X_{(i,j)}, Y_{(i,j)}$ denote the elements in X and Y that is conditioned in (i,j). $B_{(i,j)}$ denote the assignment on $Y_{(i,j)}$ and similarly $A_{(i,j)}$ for $X_{(i,j)}$. Let R' be the set of edges and the block for Alice on such R, that is $B_{(i,j)\in R}, \left\{X_{(i,j)}, Y_{(i,j)}\right\}_{(i,j)\in R}$. Let W be defined as the event of winning all the cells in R. That is for all $(i,j) \in R$, $p_{x_{(i,j)},y_{(i,j)}}(b_{(i,j)}) = a_{(i,j)}$. Let W' be defined as $W \wedge L$ where L is the event where $\Pr\left[B_{(i,j)\in R} = b_{(i,j)\in R} \mid \left\{X_{(i,j)}, Y_{(i,j)}\right\}_{(i,j)\in R}\right] \geq 2^{-\varepsilon(r^2-|R|)/80}$. $W_{(i,j)}$ refers to the event of winning the cell (i,j) that is satisfying the edge $(X_{(i,j}, Y_{(i,j)})$. $\mathcal{U}$ denotes the uniform distribution over the edges in the original projection game. P_X denotes the distribution of random variable X when the underlying distribution is P.

F.2 Distribution

In this section, we define a distribution on the grids such that each grid (i,j) and its entry $X_{(i,j)}, Y_{(i,j)}$ corresponds to the challenge that will be checked by S_i, T_j. Note that this is equivalent to conditioning $X_{(i,j)} \in S_i$ and $Y_{(i,j)} \in S_i$. Also observe that in order to win all pairs, one must then satisfy all $(X_{(i,j)}, Y_{(i,j)})$.

For all $i, j \in [r]$, S_i and T_j are picked randomly in the same procedure as in Section E. That is each $x \in X$ is in S_i with probability $1 - e^{-k/n}$, similarly for T_j. Now if we pick $X_{(i,j)}, Y_{(i,j)}$ according to $\mathcal{U}$ but with $\Pr[X_{(i,j)}, Y_{(i,j)} = \emptyset] = \Pr_{S_i,T_j}\left[E(S_i, T_j) = \emptyset\right]$, we have that

$$\mathop{\mathbb{E}}_{X_{(i,j)},Y_{(i,j)}}\left[\mathcal{D}_{S_i,T_j | X_{(i,j)}\in S_i, Y_{(i,j)}\in T_j}\right] = \mathcal{D}_{S_i,T_j}.$$

where $\mathcal{D}_{S_i,T_j}$ conditioned on $X_{(i,j)}, Y_{(i,j)}$ is $\mathcal{D}_{S_i,T_j | X_{(i,j)}\in S_i, Y_{(i,j)}\in T_j}$.

Now we are ready to further define $\mathcal{D}_{X_{(i,j)},Y_{(i,j)} | \left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R}}$ as a uniform distribution over the edges with the following constraint: $X_{(i_0,j_0)} \neq X_{(i_1,j_1)}$ if $i_0 = i_1$; and $Y_{(i_0,j_0)} \neq Y_{(i_1,j_1)}$ if $j_0 = j_1$. The probability of null is defined as

$$\Pr_{\mathcal{D}}[X_{(i,j)}, Y_{(i,j)} = \emptyset | \left\{X_{(i,j)}, Y_{(i,j)}\right\}_{(i,j)\in R}] := \Pr_{S_i,T_j}\left[E(S_i \backslash S^i, T_j \backslash T^j) = \emptyset | \left\{X_{(i,j)}, Y_{(i,j)}\right\}_{(i,j)\in R}\right]$$

where $S^{i_0} := \left\{\left\{X_{(i,j)}\right\}_{(i,j)\in R} | i = i_0\right\}$ and $S^{j_0} := \left\{\left\{X_{(i,j)}\right\}_{(i,j)\in R} | j = j_0\right\}$. Intuitively, we would like to select "fresh" edges to condition on, where "fresh" edges are defined as the set of edges that has none of its vertices conditioned to be in S_i or T_j. Then one can verify that

$$\mathop{\mathbb{E}}_{\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R}}\left[\mathcal{D}_{S_i,T_j | \left\{X_{(i,j)}\in S_i, Y_{(i,j)}\in T_j\right\}_{(i,j)\in R}}\right] = \mathcal{D}_{S_i,T_j}.$$

Recall that $\mathcal{D}_{S_i,T_j} = \mathcal{D}_{S_i} \times \mathcal{D}_{T_j}$ by definition. Now we show that for this distribution, conditioning maintains the product distribution between Alice and Bob's input. This is necessary later to extract a strategy for a single copy of Birthday Repetition and obtain a contradiction.

First we remark the following observation on product distribution:

Observation F.1 (Observation 3.10 in [8]). *Let γ be event solely determined by a. Then if $P_{a,b} = P_a \times P_b$ then $P_{a,b|\gamma} = P_{a|\gamma} \times P_{b|\gamma} = P_{a|\gamma} \times P_b$.*

Now we are ready to show that $\mathcal{D}_{\{S_i\}_{i\in[r]},\{T_j\}_{j\in[r]}|R',U}$ is indeed a product distribution.

Proposition F.2 (Product Distribution). *Let U be the event that is determined solely by $R' = \{X_{(i,j)}, Y_{(i,j)}, A_{(i,j)}\}_{(i,j)\in R}$ and $B_{(i,j)\in R}$. Then for every R' and $(i,j) \notin R$,*

$$\mathcal{D}_{\{S_i\}_{i\in[r]},\{T_j\}_{j\in[r]}|R',U} = \mathcal{D}_{\{S_i\}_{i\in[r]}|R',U} \times \mathcal{D}_{\{T_j\}_{j\in[r]}|R',U}$$

Proof. By the definition of the distribution, we have

$$\mathcal{D}_{\{S_i\}_{i\in[r]},\{T_j\}_{j\in[r]}|\{X_{(i,j)},Y_{(i,j)}\}_{(i,j)\in R}} = \mathcal{D}_{\{S_i\}_{i\in[r]}|\{X_{(i,j)},Y_{(i,j)}\}_{(i,j)\in R}} \times \mathcal{D}_{\{T_j\}_{j\in[r]}|\{X_{(i,j)},Y_{(i,j)}\}_{(i,j)\in R}}.$$

Note that $\left\{A_{(i,j)}\right\}_{(i,j)\in R}$ only depends on S_i for $i \in [r]$ since it is a two-prover game and Alice's answer only depends on Alice's input. Applying Observation F.1, we get full R':

$$\mathcal{D}_{\{S_i\}_{i\in[r]},\{T_j\}_{j\in[r]}|R'} = \mathcal{D}_{\{S_i\}_{i\in[r]}|R'} \times \mathcal{D}_{\{T_j\}_{j\in[r]}|R'}.$$

Now U only depends on R' and $\left\{B_{(i,j)}\right\}_{(i,j)\in R}$, which only depends on T_j for $j \in [r]$. Applying Observation F.1 again we get

$$\mathcal{D}_{\{S_i\}_{i\in[r]},\{T_j\}_{j\in[r]}|R',U} = \mathcal{D}_{\{S_i\}_{i\in[r]}|R',U} \times \mathcal{D}_{\{T_j\}_{j\in[r]}|R',U}.$$

□

We bound the null probability for each grid $(i,j) \in [r] \times [r]$, that is $\Pr[X_{(i,j)}, Y_{(i,j)} = \emptyset]$ under any conditioning.

Proposition F.3 (Null Probability). *For any $R \subset r \times r$, if $r < \sqrt{n}/2$ then*

$$\Pr_{\mathcal{D}}\left[X_{(i,j)}, Y_{(i,j)} = \emptyset \mid \left\{X_{(i,j)}, Y_{(i,j)}\right\}_{(i,j)\in R}\right] \le \varepsilon/160$$

Proof. We can bound the probability as in Claim C.1:

$$\begin{aligned}\Pr_{\mathcal{D}}\left[X_{(i,j)}, Y_{(i,j)} = \emptyset \mid \left\{X_{(i,j)}, Y_{(i,j)}\right\}_{(i,j)\in R}\right] &\le \mathop{\mathbb{E}}_{S_i}\left[e^{-\frac{\ell|\mathcal{N}(S_i\setminus S^i)\setminus T^j|}{n}}\right] \\ \le \mathop{\mathbb{E}}_{S_i}\left[e^{-\frac{\ell(|S_i|-2r)}{n}}\right] \le e^{-\frac{\ell((1-\delta)\mathbb{E}_{S_i}[|S_i|]-2r)}{n}} &+ e^{-\Theta(\delta^2)k} < \varepsilon/160\end{aligned}$$

where the first inequality follows from Chernoff bound and we lose $2r$ since we remove at most $2r$ vertices from participating and the last bound holds by our assumption on r and setting $k = \ell$ with an appropriately large constant factor. □

For high degree regime, we instead use the following proposition.

Proposition F.4 (Null Probability for High Degree). *For any $R \subset r \times r$, if $r < d/2$ then*

$$\Pr_{\mathcal{D}}\left[X_{(i,j)}, Y_{(i,j)} = \emptyset \mid \left\{X_{(i,j)}, Y_{(i,j)}\right\}_{(i,j)\in R}\right] \le \varepsilon/160$$

Proof. We can bound the probability as in Claim D.1 :

$$\begin{aligned}&\Pr_{\mathcal{D}}\left[X_{(i,j)}, Y_{(i,j)} = \emptyset \mid \left\{X_{(i,j)}, Y_{(i,j)}\right\}_{(i,j)\in R}\right] \le \mathop{\mathbb{E}}_{S_i}\left[e^{-\frac{\ell|\mathcal{N}(S_i\backslash S^i)\backslash T^j|}{n}}\right]\\ &\le \mathop{\mathbb{E}}_{S_i}\left[e^{-\frac{\ell(d-r)}{n}}|S_i\backslash S^i \neq \emptyset\right] + \Pr[S_i\backslash S^i = \emptyset] \le e^{-\frac{\ell(d-r)}{n}} + e^{-\frac{k(n-r)}{n}} = 1/\mathrm{poly}(n) < \varepsilon/160.\end{aligned}$$

where the first inequality follows from Chernoff bound. Now since $\frac{\ell(d-r)}{n} = \Omega(\log n)$, and $k = \ell = \Omega(\log n)$, we have bound. □

Remark F.5. *Note that the distribution over S and T's are product distributions after conditioning on R' and U for $\mathcal{D}$. This property is necessary to extract a matching protocol for free game on distribution $\mathcal{D}_{S_i,T_j|R',U}$ for any fixed R' via Proposition F.2. By the property of Birthday Repetition, the probability of winning free game under distribution $\mathcal{D}_{S_i,T_j|R',U}$ is at most the probability of strategy satisfying edge sampled from $\mathcal{D}_{X_{(i,j)},Y_{(i,j)}|R',U}$ since winning the birthday repetition is a sub-event.*

F.3 Dependency Breaking Lemmas

First we show that the modified version of Lemma 3.7 of [8] holds. This shows that the divergence between the uniform distribution over the edges and distribution conditioned on R' and event U must be small. This holds for both low degree and high degree regime.

Lemma F.6. *For any event U and $R \subset [r]\times[r]$ with $|R| = s$. Also let $\mathcal{U}$ be the uniform distribution over the edges with the following property where*

$$\Pr_{\mathcal{U}}[X_{(i,j)}, Y_{(i,j)} = \emptyset] := \max_{R,\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R}} \Pr_{\mathcal{D}}\left[X_{(i,j)}, Y_{(i,j)} = \emptyset \mid \left\{X_{(i,j)}, Y_{(i,j)}\right\}_{(i,j)\in R}\right]$$

that is it is a uniform distribution with some probability of null. Then

$$\begin{aligned}&\mathop{\mathbb{E}}_{\mathcal{D}_{\left\{X_{(i,j)},Y_{(i,j)},B_{(i,j)}\right\}_{(i,j)\in R}|U}}\left[D\left(\mathcal{D}_{\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\notin R}|\left\{X_{(i,j)},Y_{(i,j)},B_{(i,j)}\right\}_{(i,j)\in R},U}||\mathcal{U}^{\otimes(r^2-s)}\right)\right]\\ &\quad\le \log\left(\frac{1}{\Pr_{\mathcal{D}}[U]}\right) + \mathop{\mathbb{E}}_{\mathcal{D}_{\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R}|U}}\left[H(\mathcal{D}_{B_{(i,j)\in R}|\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R},U})\right]\\ &\qquad + (r^2-s)\log\left(\frac{1}{1-\frac{2r}{n}}\cdot\frac{1}{1-\frac{\varepsilon}{160}}\right)\end{aligned}$$

Proof. We can expand the divergence expression as

$$\begin{aligned}&\mathop{\mathbb{E}}_{\mathcal{D}_{\left\{X_{(i,j)},Y_{(i,j)},B_{(i,j)}\right\}_{(i,j)\in R}|U}}\left[D\left(\mathcal{D}_{\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\notin R}|\left\{X_{(i,j)},Y_{(i,j)},B_{(i,j)}\right\}_{(i,j)\in R},U}||\mathcal{U}^{\otimes(r-k)}\right)\right]\\ &= \mathop{\mathbb{E}}_{\mathcal{D}_{X,Y,\left\{B_{(i,j)}\right\}_{(i,j)\in R}|U}} \log\left(\frac{\Pr_{\mathcal{D}}\left[\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\notin R} \mid \left\{X_{(i,j)},Y_{(i,j)},B_{(i,j)}\right\}_{(i,j)\in R},U\right]}{\Pr_{\mathcal{U}^{\otimes(r^2-s)}}\left[\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\notin R}\right]}\right)\\ &= \mathop{\mathbb{E}}_{\mathcal{D}_{X,Y,\left\{B_{(i,j)}\right\}_{(i,j)\in R}|U}} \log\left(\frac{\Pr_{\mathcal{D}}\left[\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\notin R} \mid \left\{X_{(i,j)},Y_{(i,j)},B_{(i,j)}\right\}_{(i,j)\in R},U\right]}{\Pr_{\mathcal{D}}\left[\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\notin R} \mid \left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R}\right]}\right) \quad (1)\end{aligned}$$

$$+ \mathop{\mathbb{E}}_{\mathcal{D}_{X,Y,\left\{B_{(i,j)}\right\}_{(i,j)\in R}|U}} \log\left(\frac{\Pr_{\mathcal{D}}\left[\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\notin R} \mid \left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R}\right]}{\Pr_{\mathcal{U}^{\otimes(r^2-s)}}\left[\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\notin R} \mid \left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R}\right]}\right) \tag{2}$$

First we bound (2). Observe that

$$\frac{\Pr_{\mathcal{D}}\left[X_{(i,j)},Y_{(i,j)} \mid \left\{X_{(i',j')},Y_{(i',j')}\right\}_{(i',j')\neq(i,j)}\right]}{\Pr_{\mathcal{U}}\left[X_{(i,j)},Y_{(i,j)}\right]} \leq \frac{1}{1-\frac{2r}{n}}\cdot\frac{1}{1-\frac{\varepsilon}{160}}.$$

If $X_{(i,j)},Y_{(i,j)}$ is a pair that is prohibited by the choice of $\left\{X_{(i',j')},Y_{(i',j')}\right\}_{(i',j')\neq(i,j)}$, note that the probability is 0, therefore the above bound holds. If $X_{(i,j)},Y_{(i,j)}$ is null, the ratio must be less than 1, thus the above bound holds as well.

Now suppose $X_{(i,j)},Y_{(i,j)}$ is an edge that is not prohibited. Recall that $X_{(i,j)},Y_{(i,j)}$ is uniform over the entries that are not prohibited by the entries in the same row or column. This removes at most $2dr$ edges from the support. Rest of the mass is evenly distributed among other edges, conditioned on not being null. We can formally then bound the term as

$$\begin{aligned}
&\frac{\Pr_{\mathcal{D}}\left[X_{(i,j)},Y_{(i,j)} \mid \left\{X_{(i',j')},Y_{(i',j')}\right\}_{(i',j')\neq(i,j)}\right]}{\Pr_{\mathcal{U}}\left[X_{(i,j)},Y_{(i,j)}\right]}\\
&= \frac{\Pr_{\mathcal{D}}\left[X_{(i,j)},Y_{(i,j)} \mid X_{(i,j)},Y_{(i,j)}\neq\emptyset, \left\{X_{(i',j')},Y_{(i',j')}\right\}_{(i',j')\neq(i,j)}\right]}{\Pr_{\mathcal{U}}\left[X_{(i,j)},Y_{(i,j)} \mid X_{(i,j)},Y_{(i,j)}\neq\emptyset\right]}\\
&\cdot \frac{\Pr_{\mathcal{D}}\left[X_{(i,j)},Y_{(i,j)}\neq\emptyset \mid \left\{X_{(i',j')},Y_{(i',j')}\right\}_{(i',j')\neq(i,j)}\right]}{\Pr_{\mathcal{U}}\left[X_{(i,j)},Y_{(i,j)}\neq\emptyset\right]} \leq \frac{1}{1-\frac{2r}{n}}\cdot\frac{1}{1-\frac{\varepsilon}{160}}.
\end{aligned}$$

where the last inequality holds by Proposition F.3 (Proposition F.4 for the high degree case) and the observation on the size of the support. We get

$$\frac{\Pr_{\mathcal{D}}\left[X_{(i,j)},Y_{(i,j)} \mid \left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R}\right]}{\Pr_{\mathcal{U}}\left[X_{(i,j)},Y_{(i,j)}\right]} \leq \frac{1}{1-\frac{2r}{n}}\cdot\frac{1}{1-\frac{\varepsilon}{160}}. \tag{3}$$

Applying (3) exactly r^2-s times to (2), we get

$$\log\left(\frac{\Pr_{\mathcal{D}}\left[\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\notin R} \mid \left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R}\right]}{\Pr_{\mathcal{U}^{\otimes(r^2-s)}}\left[\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\notin R}\right]}\right) \leq (r^2-s)\log\left(\frac{1}{1-\frac{2r}{n}}\frac{1}{1-\frac{\varepsilon}{160}}\right).$$

Bounding (1) follows from Lemma 3.7 from [8] which we add for the completeness of the proof.

$$\begin{aligned}
(1) = &\mathop{\mathbb{E}}_{\mathcal{D}_{X,Y,\left\{B_{(i,j)}\right\}_{(i,j)\in R}|U}} \log\left(\frac{\Pr_{\mathcal{D}}\left[\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\notin R}, \left\{X_{(i,j)},Y_{(i,j)},B_{(i,j)}\right\}_{(i,j)\in R}, U\right]}{\Pr_{\mathcal{D}}\left[\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\notin R}, \left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R}\right]}\right)\\
&+ \mathop{\mathbb{E}}_{\mathcal{D}_{X,Y,\left\{B_{(i,j)}\right\}_{(i,j)\in R}|U}} \log\left(\frac{\Pr_{\mathcal{D}}\left[\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R}\right]}{\Pr_{\mathcal{D}}\left[\left\{X_{(i,j)},Y_{(i,j)},B_{(i,j)}\right\}_{(i,j)\in R}, U\right]}\right)
\end{aligned}$$

$$
\begin{aligned}
&\leq \mathop{\mathbb{E}}_{\mathcal{D}_{X,Y,\left\{B_{(i,j)}\right\}_{(i,j)\in R}|U}} \log\left(\frac{\Pr_{\mathcal{D}}\left[\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R}\right]}{\Pr_{\mathcal{D}}\left[\left\{X_{(i,j)},Y_{(i,j)},B_{(i,j)}\right\}_{(i,j)\in R},U\right]}\right)\\
&= \mathop{\mathbb{E}}_{\mathcal{D}_{X,Y,\left\{B_{(i,j)}\right\}_{(i,j)\in R}|U}} \log\left(\frac{1}{\Pr_{\mathcal{D}}\left[\left\{B_{(i,j)}\right\}_{(i,j)\in R},U|\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R}\right]}\right)\\
&= \mathop{\mathbb{E}}_{\mathcal{D}_{\left\{X_{(i,j)},Y_{(i,j)},B_{(i,j)}\right\}_{(i,j)\in R}|U}} \log\left(\frac{1}{\Pr_{\mathcal{D}}\left[\left\{B_{(i,j)}\right\}_{(i,j)\in R},U|\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R}\right]}\right) \qquad (4)
\end{aligned}
$$

Now we can decompose (4) as

$$
\begin{aligned}
(4) = &\mathop{\mathbb{E}}_{\mathcal{D}_{\left\{X_{(i,j)},Y_{(i,j)},B_{(i,j)}\right\}_{(i,j)\in R}|U}} \log\left(\frac{1}{\Pr_{\mathcal{D}}\left[U|\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R}\right]}\right)\\
&+ \mathop{\mathbb{E}}_{\mathcal{D}_{\left\{X_{(i,j)},Y_{(i,j)},B_{(i,j)}\right\}_{(i,j)\in R}|U}} \log\left(\frac{1}{\Pr_{\mathcal{D}}\left[\left\{B_{(i,j)}\right\}_{(i,j)\in R}|\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R},U\right]}\right)\\
&\leq \log\left(\frac{1}{\Pr_{\mathcal{D}}[U]}\right) + \mathop{\mathbb{E}}_{\mathcal{D}_{\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R}|U}} H\left(\mathcal{D}_{\left\{B_{(i,j)}\right\}_{(i,j)\in R}|\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R},U}\right)
\end{aligned}
$$

where the inequality follows from the concavity of log. Therefore, we have

$$
(1) \leq \log\left(\frac{1}{\Pr_{\mathcal{D}}[U]}\right) + \mathop{\mathbb{E}}_{\mathcal{D}_{\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R}|U}} \left[H(\mathcal{D}_{B_{(i,j)\in R}|\left\{X_{(i,j)},Y_{(i,j)}\right\}_{(i,j)\in R},U})\right]
$$

Combining the bound for (1) and (2) completes the proof. □

We then proceed to bound $\mathbb{E}_{\mathcal{D}_{\left\{X_i,Y_j\right\}_{(i,j)\in R}|W'}} H\left(\mathcal{D}_{B_{(i,j)\in R}|\{X_i,Y_j\}_{(i,j)\in R},W'}\right)$.

Proposition F.7.

$$
\mathop{\mathbb{E}}_{\mathcal{D}_{\left\{X_i,Y_j\right\}_{(i,j)\in R}|W'}} H\left(\mathcal{D}_{B_{(i,j)\in R}|\{X_i,Y_j\}_{(i,j)\in R},W'}\right) \leq \varepsilon(r^2-s)/80
$$

Proof. This follows directly the argument of [8] and the definition of W'. Recall that

$$
\left|\mathsf{Supp}\left(\mathcal{D}_{B_{(i,j)\in R}|\{X_i,Y_j\}_{(i,j)\in R},W'}\right)\right| \leq 2^{\varepsilon(r^2-s)/80}
$$

by the definition of W'. The bound on the entropy term then follows since

$$
H\left(\mathcal{D}_{B_{(i,j)\in R}|\{X_i,Y_j\}_{(i,j)\in R},W'}\right) \leq \log\left|\mathsf{Supp}\left(\mathcal{D}_{B_{(i,j)\in R}|\{X_i,Y_j\}_{(i,j)\in R},W'}\right)\right|.
$$

□

Corollary F.8. *For R' and W' where $|R| = s$, if $r < 4n$,*

$$
\begin{aligned}
&\mathop{\mathbb{E}}_{(i,j)\notin R}\mathop{\mathbb{E}}_{\mathcal{D}_{R'|W'}}\left[D\left(\mathcal{D}_{X_i,Y_j|R',W'}||\mathcal{U}\right)\right]\\
&\quad\leq \frac{1}{r^2-s}\left(\varepsilon(r^2-s)/80 - \log\left(\Pr_{\mathcal{D}}[W] - 2^{-\varepsilon(r^2-s)/80}\right)\right) + \frac{4r}{n} + \varepsilon/80
\end{aligned}
$$

Proof. By taking the expectation over the remaining grids to Lemma F.6 and setting $U = W'$ with Proposition F.7, the following bound then follows from Fact A.5 and the property $\log \frac{1}{1-x} < 2x$ if $x < 1/2$.

$$\mathop{\mathbb{E}}_{(i,j)\notin R}\mathop{\mathbb{E}}_{\mathcal{D}_{R'|W'}}\left[D\left(\mathcal{D}_{X_i,Y_j|R',W'}||\mathcal{U}\right)\right]$$
$$\leq \frac{1}{r^2-s}\left(\varepsilon(r^2-s)/80 - \log\left(\Pr_{\mathcal{D}}[W']\right)\right) + \frac{4r}{n} + \varepsilon/80$$

Then observe that $\Pr_{\mathcal{D}}[W'] \geq \Pr_{\mathcal{D}}[W] - 2^{-\varepsilon(r^2-s)/80}$ since $\Pr_{\mathcal{D}}[W \wedge \neg L]$ is bounded by $2^{-\varepsilon(r^2-s)/80}$ from definition of L. Plugging in the bound for $\Pr_{\mathcal{D}}[W']$ we get the desired bound for $\Pr_{\mathcal{D}}[W']$. □

F.4 Proof of Theorem 1.4

Now we combine lemmas and propositions from Section F.2 and Section F.3 to prove the main lemma of this section:

Lemma F.9. *If* $\Pr_{\mathcal{D}}[W] \geq 2^{-\varepsilon(r^2-s)/320}$, $r^2 - s \geq (320/\varepsilon)\log(320/\varepsilon)$, $r < \min\{\sqrt{n}/2, \varepsilon n/320\}$ *(for high degree,* $r < d/2$ *as well) and* $\mathbf{val}(\mathcal{G}) < 1 - 2\varepsilon$*, then there exists* $(i,j) \notin R$ *such that*

$$\Pr_{\mathcal{D}}[W_{(i,j)}|W] \leq 1 - \frac{\varepsilon}{160} \tag{5}$$

Proof. First, we show that $\Pr_{\mathcal{D}}[W_{(i,j)}|W'] < 1 - \varepsilon/80$ for some coordinate (i,j). Assume by contradiction for any coordinate (i,j), $\Pr_{\mathcal{D}}[W_{(i,j)}|W'] \geq 1 - \varepsilon/80$. Then observe that we can rewrite $\Pr_{\mathcal{D}}[W_{(i,j)}|W']$ as

$$\Pr_{\mathcal{D}}[W_{(i,j)}|W'] = \mathop{\mathbb{E}}_{\mathcal{D}_{R'|W'}} \Pr[W_{(i,j)}|R', W'] \geq 1 - \varepsilon/80.$$

An equivalent assumption is

$$\mathop{\mathbb{E}}_{\mathcal{D}_{R'|W'}} \Pr[\neg W_{(i,j)}|R', W'] < \varepsilon/80$$

Recall that there exists a matching strategy for S_i, T_j via Proposition F.2, which indeed gives an assignment for $X_{(i,j)}, Y_{(i,j)}$ conditioned on R', W'. $X_{(i,j)}, Y_{(i,j)}$ is indeed checked by definition of the Birthday Repetition. Now from Corollary F.8,

$$\mathop{\mathbb{E}}_{(i,j)\notin R}\mathop{\mathbb{E}}_{R'|W'}\left[D\left(\mathcal{D}_{X_i,Y_j|R',W'}||\mathcal{U}\right)\right] \leq \frac{1}{r^2-s}\left(\varepsilon(r^2-s)/80 - \log\left(\Pr_{\mathcal{D}}[W] - 2^{-\varepsilon(r^2-s)/80}\right)\right)$$
$$+ \frac{4r}{n} + \frac{\varepsilon}{80} \leq \varepsilon/80 + \frac{4r}{n} - \frac{\log\left(\Pr_{\mathcal{D}}[W] - 2^{-\varepsilon(r^2-s)/80}\right)}{r^2-s} + \frac{\varepsilon}{80}$$
$$\leq \varepsilon/80 + \varepsilon/80 + \frac{4r}{n} + \varepsilon/80 \leq \varepsilon/20$$

where the second to last inequality holds by our assumption on r that is

$$\Pr_{\mathcal{D}}[W] - 2^{-\varepsilon(r^2-s)/80} \geq 2^{-\varepsilon(r^2-s)/160} - 2^{-\varepsilon(r^2-s)/80} = 2^{-\varepsilon(r^2-s)/80}\left(2^{\varepsilon(r^2-s)/160} - 1\right) \geq 2^{-\varepsilon(r^2-s)/80}$$

and the last inequality holds by our assumption $r < \varepsilon n/320$. Thus we have

$$\mathop{\mathbb{E}}_{(i,j)\notin R}\mathop{\mathbb{E}}_{R'|W'}\left[D\left(\mathcal{D}_{X_i,Y_j|R',W'}||\mathcal{U}\right)\right] \leq \varepsilon/20$$

By Markov argument, there exists a fixing of R' and (i,j) such that satisfies all of following.

$$D\left(\mathcal{D}_{X_{(i,j)},Y_{(i,j)}|R',W'}||\mathcal{U}\right) \qquad \leq \varepsilon/4 \tag{6}$$

$$\Pr_{\mathcal{D}}[\neg W_{(i,j)}|R',W'] \qquad \leq \varepsilon/16 \tag{7}$$

Combining these conditions with Fact A.6, one can win a copy of the original game under $\mathcal{U}$ with probability $\geq 1-\varepsilon$ which is indeed a contradiction. $\mathcal{U}$ is null with at most 1/2 probability by Proposition F.3 (for high degree case Proposition F.4) while by assumption $\mathbf{val}(\mathcal{G}) < 1-2\varepsilon$.

Now we have that there exists (i,j) such that

$$\Pr_{\mathcal{D}}[W_{(i,j)}|W'] = \mathop{\mathbb{E}}_{\mathcal{D}_{R'|W'}} \Pr[W_{(i,j)}|R',W'] < 1-\varepsilon/80$$

while we want to bound $\Pr_{\mathcal{D}}[W_{(i,j)}|W]$. This is implied by the above inequality. Since W' is a subset event of W

$$\begin{aligned}
&\Pr_{\mathcal{D}}[W_{(i,j)}|W] \leq \Pr_{\mathcal{D}}[W_{(i,j)}|W'] + \Pr[\neg W'|W] \\
&\leq \Pr_{\mathcal{D}}[W_{(i,j)}|W'] + 2^{-\varepsilon(r^2-s)/80}/\Pr_{\mathcal{D}}[W] \leq 1-\varepsilon/80 + 2^{-\frac{\varepsilon(r^2-s)}{80}+\frac{\varepsilon(r^2-s)}{320}} \\
&= 1-\varepsilon/80 + 2^{-\frac{\varepsilon(r^2-s)}{320}} \leq 1-\varepsilon/320
\end{aligned}$$

where the last bound holds by our assumption on r, that is $(r^2-s) \geq 320/\varepsilon \log(320/\varepsilon)$.

□

Now we apply Lemma F.9 recursively to prove Theorem 1.4 which is our main technical theorem, which we restate for readability.

Theorem 1.4. *Suppose $k = \ell = \Omega(\sqrt{n\log(1/\varepsilon)})$, and $\mathcal{G}$ is a projection game. Further, suppose $r \leq \min\{\sqrt{n}/2, \varepsilon n/320\}$. Then if $\mathbf{val}(\mathcal{G}) < 1-\varepsilon$,*

$$\mathbf{val}(\mathcal{G}^{rk\times r\ell}) \leq (1-\varepsilon/2)^{\Omega(r^2)}.$$

Proof of Theorem 1.4. Let R be the set of coordinates of size k and let W be the event of winning on all the coordinates in R. Suppose $r^2 \geq \frac{1600\log(320/\varepsilon)}{\varepsilon}$, and $k < r^2/5$. Then note that $r^2 - s > \frac{320\log(320/\varepsilon)}{\varepsilon}$, thus we can apply Lemma F.9. We prove by induction on k, that is for every k, there exists a subset R of size k such that

$$\Pr_{\mathcal{D}}[W] \leq (1-\varepsilon/320)^k$$

If $k=0$, the statement holds trivially. Suppose it is true for k, we show then it is also true for $k+1$ if $k \leq r^2/5-1$. If $\Pr_{\mathcal{D}}[W] \leq (1-\varepsilon/320)^{k+1}$, we are done. Suppose otherwise, that is

$$\Pr_{\mathcal{D}}[W] \geq (1-\varepsilon/320)^{k+1} \geq 2^{-\varepsilon(k+1)/160} \geq 2^{-\varepsilon(r^2-k)/320}.$$

where the last inequality holds by our assumption on r and k. Then we apply Lemma F.9 to add a coordinate to R. By Lemma F.9, there exists (i,j) such that $\Pr_{\mathcal{D}}[W_{(i,j)}|W] < 1-\varepsilon/320$. If we add (i,j) to R, then note that

$$\Pr_{\mathcal{D}}[W \wedge W_{(i,j)}] = \Pr_{\mathcal{D}}[W]\cdot\Pr_{\mathcal{D}}[W_{(i,j)}|W] \leq (1-\varepsilon/320)^k \cdot (1-\varepsilon/320) \leq (1-\varepsilon/320)^{k+1}.$$

Thus the probability of $\bigwedge_{(i,j)\in[r]\times[r]} W_{(i,j)}$ is indeed bounded by $2^{-\Omega(\varepsilon r^2)}$.[4] Since this is a sub-event of winning S_i and T_j pairs for all $(i,j) \in [r] \times [r]$, the probability of winning all S_i and T_j's is upper-bounded by $2^{-\Omega(\varepsilon r^2)}$. □

Recall that high degree regime theorem can be written as

Theorem 1.5. *Suppose $k = \ell$, and $\mathcal{G}$ is a projection game, and $d = \Omega(\frac{n \log n}{k})$. Further, suppose $r \le \min\{\sqrt{n}/2, \varepsilon n/320, d/2\}$. Then if $\mathbf{val}(\mathcal{G}) < 1 - \varepsilon$,*

$$\mathbf{val}(\mathcal{G}^{rk\times r\ell}) \le (1-\varepsilon/2)^{\Omega(r^2)}.$$

We leave the proof to the reader since Lemma F.9 applies to the high degree regime as well, that is $d = \Omega(\frac{n \log n}{k})$.

Indeed a natural question to ask is whether the statement is true for small r, since we assumed $r^2 \ge \frac{1600 \log(320/\varepsilon)}{\varepsilon}$ in the proof. [30] pointed out that the parallel repetition for large r should imply for small r with the same parameters. Now we give an analogous argument on the grid. Suppose the statement is false for some small r, r_0. That is there exists a strategy that obtains the value of $2^{-o(\varepsilon r_0^2)}$. Then one can copy this strategy multiple times for a bigger sized grid, which achieves the value of $2^{-o(\varepsilon r^2)}$. This is indeed a contradiction.

G Proof of Lemma 1.6

In this section, we prove Lemma 1.6 which translates bounds for the Birthday Repetition to bounds for symmetric parallel repetition. In particular, we exhibit bound for $\mathbf{val}(\mathcal{G}'_{\mathcal{D}_{k,\ell}})$ via bound on $\mathbf{val}(\mathcal{G}^{k\times\ell})$.

Let $\mathcal{D}_{k,\ell}$, the distribution on $2^E \subset 2^X \times 2^Y$ parameterized by k and ℓ defined as the distribution on (S,T) formed by the following process. (which is induced from the Birthday Repetition)

- The referee selects S by adding each $x \in X$ with probability $1 - e^{-k/n}$ and selects T by adding each $y \in Y$ with probability $1 - e^{-\ell/n}$.
- Consider $E(S,T)$. Let $S_0 := \{x | (x,y) \in E(S,T)\}$ and $T_0 := \{y | (x,y) \in E(S,T)\}$.
- If $S = S_0$ and $T = T_0$, that there exists a matching between S and T in $E(S,T)$, then the referee sends S to Alice and T to Bob. The referee selects $E(S,T)$ as C. Otherwise, repeat the above process.

Lemma 1.6. *If the degree of the original Two Prover game $d \ge (1+\delta) n \log n / k$ with $\delta = \Omega(1/\log n)$ and $k = \ell$, then*

$$|\mathbf{val}(\mathcal{G}_{\mathcal{D}_{k,\ell}}) - \mathbf{val}(\mathcal{G}^{k\times\ell})| = o(1)$$

Proof. Let $\mathcal{U}_{k,\ell}$ denote the distribution on 2^X and 2^Y in $\mathcal{G}^{k\times\ell}$. We give a protocol to transform the strategy for $\mathcal{G}_{\mathcal{D}_{k,\ell}}$ into a strategy for $\mathcal{G}^{k\times\ell}$. Alice, given S, and Bob, given T, simply look up strategy for S and T respectively in $\mathcal{G}_{\mathcal{D}_{k,\ell}}$ then give the answer.

For such strategy, it suffices to bound $\|\mathcal{D}_{k,\ell} - \mathcal{U}_{k,\ell}\|_1$. Then observe that

$$\|\mathcal{D}_{k,\ell} - \mathcal{U}_{k,\ell}\|_1 = 2 \cdot \Pr[S \ne S_0 \vee T \ne T_0]$$

[4] This property is used in the proof of Theorem 1.7. But for the purpose of proving result for Birthday Repetition, this observation is not necessary.

since $\mathcal{D}_{k,\ell}$ is simply $\mathcal{U}_{k,\ell}$ conditioned on $S = S_0$ and $T = T_0$. Then

$$\Pr[x \notin S_0] = \left(e^{-\ell/n}\right)^d = e^{-\ell d/n} = 1/n^{1+\delta}$$

From union bound, since $|S|, |T| \leq O(k)$, we have

$$\Pr[S \neq S_0] \leq |S| \cdot e^{-\ell d/n} \leq 1/n^{\delta}$$
$$\Pr[T \neq T_0] \leq |T| \cdot e^{-kd/n} \leq 1/n^{\delta}.$$

since $|S|, |T| \leq n$. Combining the bounds, we get

$$\|\mathcal{D}_{k,\ell} - \mathcal{U}_{k,\ell}\|_1 \leq o(1).$$

□

Indeed we have not used the fact that the actual chosen sets (S, T) must be of size $O(k)$ and $O(\ell)$ respectively. However, this does not change the proof by much since ℓ_1 norm difference from pruning is at most $O(e^{-\Omega(k)})$ if $k = \ell$. Since $k = \ell = \Omega(\log n)$, again this ℓ_1 norm is at most $1/\text{poly}(n)$. Thus, $\Pr[S \neq S_0 \vee T \neq T_0]$ under the pruned distribution is again at most $o(1)$, and the size for $\mathcal{G}_{\mathcal{D}_{k,\ell}}$ is at most $n^{O(k)}$.

Lemma 1.6 along with Theorem 1.4 (or Theorem 1.5) gives the desired bound on $\mathbf{val}(\mathcal{G}_{\mathcal{D}_{k,\ell}})$ as a direct corollary, which we restate for completeness.

Theorem 1.7. *Let $\mathcal{G}$ be a projection game where the underlying constraint graph is a d-regular graph. Let $\mathbf{val}(\mathcal{G}) < 1 - \varepsilon$. Set $k = \ell = 4n \log n/d$. If $k = \ell = \Omega(\sqrt{n \log(1/\varepsilon)})$ and $\frac{k^2}{n \log(1/\varepsilon)} \leq \min\{\sqrt{n}/2, \varepsilon n/320\}$ are satisfied then there exists a distribution on $\mathcal{D}_{k,\ell}$ on 2^E that is only supported on the set of size $O(k)$ such that*

$$\mathbf{val}(\mathcal{G}'_{\mathcal{D}_{k,\ell}}) \leq (1 - \varepsilon/2)^{\Omega(\frac{k^2}{n \log(1/\varepsilon)})}.$$

Else, for any $r > 0$ that satisfies $r \leq \min\{\sqrt{n}/2, \varepsilon n/320, d/2\}$, there exists a distribution on $\mathcal{D}_{rk,r\ell}$ on 2^E that is only supported on the set of size $O(rk)$ such that

$$\mathbf{val}(\mathcal{G}'_{\mathcal{D}_{k,\ell}}) \leq (1 - \varepsilon/2)^{\Omega(r^2)}.$$

We leave the proof for the reader, since the proof follows from rearranging the variables.

Theorem G.1 (Lower Bound)**.** *Let $\mathcal{G}$ be a projection game where the underlying constraint graph is d-regular graph. Let $\mathbf{val}(\mathcal{G}) > 1 - \varepsilon$, $k, \ell = o(n)$ and $\frac{\varepsilon d k \ell}{n} \leq 1/8$. Then*

$$\mathbf{val}(\mathcal{G}^{k \times \ell}) \geq e^{-O(\frac{\varepsilon d k \ell}{n})}.$$

Proof. We prove by union bound. Consider the strategy for $\mathcal{G}$ that wins with probability $> 1 - \varepsilon$ say $f : X \to A$ and $g : Y \to B$. Consider its natural extension in $G^{k \times \ell}$, which is Alice, when given S as an input, for each $x \in S$, answers $f(x)$, similarly for Bob with g. Note that this strategy wins whenever $E(S, T)$ does not contain an edge that is not satisfied by f and g.

Then recall that the probability of picking an edge is $(1 - e^{-k/n})(1 - e^{-\ell/n}) \leq 4k\ell/n^2$. Applying union bound over all edges that are not satisfied, the probability of picking an edge that is not satisfied by f and g is

$$\varepsilon|E| \cdot 4k\ell/n^2 \leq \frac{4\varepsilon d k \ell}{n}$$

Thus we have a strategy for $G^{k \times \ell}$ that wins with probability $1 - \frac{4\varepsilon d k \ell}{n} \geq e^{-\frac{8\varepsilon d k \ell}{n}}$ where the inequality holds from our assumption on $\frac{\varepsilon d k \ell}{n} \leq 1/8$.

□

H Proof of Theorem 1.9

Let $\mathcal{D}_r$ be defined as the following distribution on 2^E. We choose the set of edges C by picking each $e \in E$ with probability $1 - e^{-r/|E|}$. Indeed, the size of the reduction may not be $n^{O(r)}$, since $|C| > r$. But we avoid this issue by conditioning on the cases where $|C| < 2r$, which indeed contains most of the mass due to Chernoff Bound. We denote the "pruned" distribution as $\overline{\mathcal{D}}_r$, though the difference in the value of the game in terms of ℓ_1 norm is at most $2 \cdot e^{-\Omega(r)}$.

In the rest of this section, we prove Theorem 1.9 without using complicated machineries developed in Section F for the Birthday Repetition. The main observation is to analyze the behavior of the game when the set S and T becomes large so that $|S| = |T| = n$. It is easy to check that if $\mathbf{val}(\mathcal{G}) < 1$, the value of this new game becomes 0, since the verifier will check all the edges. Instead, suppose that the verifier does not check all the edges, but chooses random edges in E according to some distribution. In particular, suppose the referee picks each edge with probability $1 - e^{-r/dn}$ and accepts iff all picked edges are satisfied, that is $\mathcal{D}_r$. But the difference with $\mathcal{G}_{\mathcal{D}_r}$ is that while Alice and Bob gets $C_X = \{x | (x, y) \in C\}$ and respectively $C_Y = \{y | (x, y) \in C\}$ as input for $\mathcal{G}_{\mathcal{D}_r}$, this is not case for this new game. Denote this new prover game as $\mathcal{G}^r_{all}$. Also note that we do not assume $\mathcal{G}$ to be a projection game.

Remark H.1. *Indeed $\mathcal{G}^r_{all}$ is not useful for practical purposes. Though the number of vertex is 2 in the bipartite graph, the alphabet size blows up exponentially to A^n and B^n, thus the size of the game is $2^{\Omega(n)}$ if the original game had constant sized alphabet. Also it does not preserve uniqueness nor projection.*

We show upper bound for $\mathbf{val}(\mathcal{G}^r_{all})$ which is relatively straightforward.

Claim H.2. *If $\mathbf{val}(\mathcal{G}) < 1 - \varepsilon$, then $\mathbf{val}(\mathcal{G}^r_{all}) < 2^{-\Omega(\varepsilon r)}$. Similarly, if $\mathbf{val}(\mathcal{G}) > 1 - \varepsilon$, then $\mathbf{val}(\mathcal{G}^r_{all}) = 2^{-O(\varepsilon r)}$*

Proof. Let a be the strategy for Alice and b be the strategy for Bob. Indeed a and b could be randomized. However, note that there is always a deterministic strategy that performs at least as good as randomized strategy. Thus without loss of generality, let a and b be a deterministic strategy. Note that our assumption on the game implies that a and b can satisfy at most $(1 - \varepsilon)$-fraction of the edges. Thus the fraction of the edges that are not satisfied is at least ε. Denote such edges as E_{bad}. The referee accepts iff all no edges from E_{bad} is picked, the probability of which is

$$\left(1 - (1 - e^{-r/dn})\right)^{|E_{bad}|} \leq e^{-\varepsilon r} = 2^{-\Omega(\varepsilon r)}$$

For the second part of the claim, if $\mathbf{val}(\mathcal{G}) > 1 - \varepsilon$,

$$\left(1 - (1 - e^{-r/dn})\right)^{|E_{bad}|} \geq e^{-\varepsilon r} = 2^{-O(\varepsilon r)}$$

□

Now we transform the bound for $\mathbf{val}(\mathcal{G}^r_{all})$ to a bound for $\mathbf{val}(\mathcal{G}_{\mathcal{D}_r})$. First we prove the following simple claim that has been stated before:

Claim H.3.

$$\|\mathcal{D}_r - \overline{\mathcal{D}}_r\|_1 = e^{-\Theta(r)}.$$

Proof. Recall that $\overline{\mathcal{D}}_r$ is simply $\mathcal{D}_r$ conditioned on the set of having size $< 2r$. Thus

$$\|\mathcal{D}_r - \overline{\mathcal{D}}_r\|_1 \leq 2 \Pr_{c \sim \mathcal{D}_r}[|c| > 2r] \leq 2e^{-\Theta((1-e^{-r/dn})dn)} \leq 2e^{-\Theta(r)}$$

where the bound follows from Chernoff Bound. □

We can without loss of generality argue on $\mathcal{D}_r$ since the value only differs by $2e^{-\Theta(r)}$.

Lemma H.4.

$$\mathbf{val}(\mathcal{G}_{\mathcal{D}_r}) < 2^{-\Omega(\varepsilon r)} \tag{8}$$

for $r = \Omega(n \log(1/\varepsilon))$.

Proof. We translate the strategy for $\mathcal{G}_{\mathcal{D}_r}$ to strategy for $\mathcal{G}^r_{all}$. Let L denote the random binary string in $\{0,1\}^X$ such that $l_x = 1$ iff x is checked by the referee, 0 otherwise. Define R as random binary string in $\{0,1\}^Y$ similarly on Bob's side.

$$H(L) = \sum_{x \in X} H(L_x) = \sum_{x \in X} H\left(\left(e^{-\frac{r}{dn}}\right)^d\right) = nH\left(e^{-r/n}\right)$$

where first equality holds since L_x's are all independent. Similarly,

$$H(R) = nH\left(2^{-\Theta(r/n)}\right)$$

Consider following strategy by Alice and Bob. Alice and Bob, pick L and R independently at random according to the same distribution as the verifier. Compare the guessed copy with actual L and R picked by the verifier. Then let E denote the event where Alice and Bob guess L and R correctly. Note that under this protocol,

$$\Pr[E] = \sum_{l,r} \Pr[L = l, R = r] \cdot \Pr[L = l] \cdot \Pr[R = r] \geq \sum_{l,r} \Pr[L = l, R = r]^3$$
$$\geq 2^{-3H(L,R)} \geq 2^{-3(H(L)+H(R))}$$

Suppose they guessed correctly. Now note that conditioned on E, the distribution becomes skewed. In particular,

$$\Pr[L_x = 0|E] = \frac{\Pr[L_x = 0]^2}{\Pr[L_x = 1]^2 + \Pr[L_x = 0]^2}$$
$$= \frac{\Pr[L_x = 0]^2}{(1 - \Pr[L_x = 0])^2 + \Pr[L_x = 0]^2} = \frac{\Pr[L_x = 0]^2}{1 - 2\Pr[L_x = 0] + 2\Pr[L_x = 0]^2}$$

And same for Bob's side as well. Let r' be such that $\left(e^{-\frac{r'}{dn}}\right)^d = \Pr[L_x = 0|E]$. Note that $r' = \Theta(r)$.

Then now we apply the strategy for $\mathcal{G}_{\mathcal{D}_{r'}}$. Note that the distribution of the input exactly matches by our choice or r'. Thus

$$\mathbf{val}(\mathcal{G}^r_{all}) \geq \mathbf{val}(\mathcal{G}_{\mathcal{D}_{r'}}) \cdot \Pr[E] \geq \mathbf{val}(\mathcal{G}_{\mathcal{D}_{r'}}) \cdot 2^{-6nH\left(2^{-\Theta(r/n)}\right)} \tag{9}$$

Combining with Claim H.2 we have

$$\mathbf{val}(\mathcal{G}_{\mathcal{D}_{r'}}) < 2^{-\Omega(\varepsilon r) + \Theta(r \cdot 2^{-\Theta(r/n)})} \tag{10}$$

Choosing $r = \Omega(n \log(1/\varepsilon))$, we get

$$\mathbf{val}(\mathcal{G}_{\mathcal{D}_{r'}}) < 2^{-\Omega(\varepsilon r)} = 2^{-\Omega(\varepsilon r')} \tag{11}$$

□

The fact that each edge is picked independently is crucial in bounding the probability $L_x = 0$ conditioned on E. It would be interesting to show similar bound for picking an edge r-times at random.

Lemma H.5. *If* $\mathbf{val}(\mathcal{G}) > 1 - \varepsilon$ *then*

$$\mathbf{val}(\mathcal{G}_{\mathcal{D}_r}) \geq 2^{-O(\varepsilon r)}$$

Proof. Recall the proof of Claim H.2. Suppose Alice and Bob ignores the inputs and just follows strategy for $\mathcal{G}^r_{all}$. Then the probability that the referee does not choose any bad edge is

$$\left(1 - (1 - e^{-r/dn})\right)^{|E_{bad}|} \geq e^{-\varepsilon r} = 2^{-O(\varepsilon r)}$$

which proves our claim. □

Proof of Theorem 1.9. We combine all the claims to prove the theorem. First note that $\mathbf{val}(\mathcal{G}_{\overline{\mathcal{D}}_r})$ and $\mathbf{val}(\mathcal{G}_{\mathcal{D}_r})$ can deviate by at most $e^{-\Theta(\delta^2 r)}$ due to ℓ_1 norm bound in the underlying distribution.

$$\mathbf{val}(\mathcal{G}_{\overline{\mathcal{D}}_r}) \leq \mathbf{val}(\mathcal{G}_{\mathcal{D}_r}) + e^{-\Theta(\delta^2 r)} \leq 2^{-\Omega(\varepsilon r)} + 2e^{-\Theta(r)}$$

Having since $r = \Omega(1/\varepsilon)$ we get the desired inequality. And similarly

$$\mathbf{val}(\mathcal{G}_{\overline{\mathcal{D}}_r}) \geq \mathbf{val}(\mathcal{G}_{\mathcal{D}_r}) - e^{-\Theta(\delta^2 r)} \geq 2^{-O(\varepsilon r)} - 2e^{-\Theta(r)} = 2^{-O(\varepsilon r)}.$$

□

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il