On Fortification of Projection Games

Amey Bhangale* Ramprasad Saptharishi† Girish Varma‡ Rakesh Venkat§

May 25, 2015

Abstract

A recent result of Moshkovitz [Mos14] presented an ingenious method to provide a completely elementary proof of the *Parallel Repetition Theorem* for certain projection games via a construction called *fortification*. However, the construction used in [Mos14] to fortify arbitrary label cover instances using an arbitrary extractor is insufficient to prove parallel repetition. In this paper, we provide a fix by using a stronger graph that we call *fortifiers*. Fortifiers are graphs that have both ℓ_1 and ℓ_2 guarantees on induced distributions from large subsets.

We then show that an expander with sufficient spectral gap, or a bi-regular extractor with stronger parameters (the latter is also the construction used in an independent update [Mos15] of [Mos14] with an alternate argument), is a good fortifier. We also show that using a fortifier (in particular ℓ_2 guarantees) is necessary for obtaining the robustness required for fortification.

1 Introduction

Label-cover and general two-prover games

A label cover instance is specified by a bipartite graph $G = ((X, Y), E)$, a pair of alphabets Σ_X and Σ_Y and a set of constraints $\psi_e : \Sigma_X \to \Sigma_Y$ on each edge $e \in E$. The goal is to label the vertices of X and Y using labels from Σ_X and Σ_Y so as to satisfy as many constraints are possible.

This problem is often viewed as a two-prover game. The verifier picks an edge (x, y) at random and sends x to the first prover and y to the second prover. They are to return a label of the vertex that they received, and the verifier accepts if the labels they returned are consistent with the constraint $\psi_{(x,y)}$. The value of this game G, denoted by $\mathrm{val}(G)$, is given by the acceptance probability of the verifier maximized over all possible strategies of the provers. These are also called *projection games* as the constraints are functions from Σ_X to Σ_Y. They are called *general games* if the constraint on each edge is an arbitrary relation $\psi_{(x,y)} \subseteq \Sigma_X \times \Sigma_Y$.

These two notions are equivalent in the sense that $\mathrm{val}(G)$ is exactly equal to the maximum fraction of constraints that can be satisfied by any labelling.

*Rutgers University. Research supported by the NSF grant CCF-1253886. amey.bhangale@rutgers.edu

†Tel Aviv University. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement number 257575. ramprasad@cmi.ac.in

‡Tata Institute of Fundamental Research, Mumbai. Supported by the Google India Fellowship in Algorithms. girishrv@tifr.res.in

§Tata Institute of Fundamental Research, Mumbai. rakesh@tifr.res.in

This problem is central to the PCP Theorem [AS98, ALM+98] and almost all inapproximability results that stem from it. The (Strong) PCP Theorem can be rephrased as stating that for every $\varepsilon > 0$, it is NP-hard to distinguish whether a given label cover instance has $\mathrm{val}(G) = 1$ or $\mathrm{val}(G) < \varepsilon$. An important step is a way to transform instances with $\mathrm{val}(G) < 1 - \varepsilon$ to instances G' with $\mathrm{val}(G') < \varepsilon$. This is usually achieved via the *Parallel Repetition Theorem*.

Parallel Repetition

The k-fold repetition of a game G, denoted by G^k, is the following natural definition — the verifier picks k edges $(x_1, y_1), \cdots, (x_k, y_k)$ from E uniformly and independently, sends $(x_1, \ldots, x_k)$ and $(y_1, \ldots, y_k)$ to the provers respectively, and accepts if the labels returned by them are consistent on each of the k edges.

If $\mathrm{val}(G) = 1$ to start with then $\mathrm{val}(G^k)$ still remains 1. How does $\mathrm{val}(G^k)$ decay with k if $\mathrm{val}(G) < 1$? Turns out even this simple operation of repeating a game in parallel has a counter-intuitive effect on the value of the game. It is easy to see that $\mathrm{val}(G^k) \geq \mathrm{val}(G)^k$ as provers can use a same strategy as in G to answer each query (x_i, y_i). The first surprise is $\mathrm{val}(G^k)$ is *not* $\mathrm{val}(G)^k$, but sometimes can be *much larger* than $\mathrm{val}(G)^k$. Fortnow [For89] presented a game G for which $\mathrm{val}(G^2) > \mathrm{val}(G)^2$, Feige [Fei91] improved this by giving an example of game G with $\mathrm{val}(G) < 1$ but $\mathrm{val}(G^2) = \mathrm{val}(G)$. Indeed, there are known examples [Raz11] of projection games where $\mathrm{val}(G) = (1 - \varepsilon)$ but $\mathrm{val}(G^k) \geq \left(1 - \varepsilon\sqrt{k}\right)$ for a large range of k.

The first non trivial upper bound on $\mathrm{val}(G^k)$ was proven by Verbitsky [Ver96] who showed that if $\mathrm{val}(G) < 1$ then the value $\mathrm{val}(G^k)$ must go to zero as k goes to infinity. It is indeed true that $\mathrm{val}(G^k)$ decays exponentially with k (if $\mathrm{val}(G) < 1$). This breakthrough was first proved by Raz [Raz98], and has subsequently seen various simplifications and improvements in parameters [Hol09, Rao11, DS14, BG14]. The following statements are due to Holenstein [Hol09], Dinur and Steurer [DS14] respectively.

Theorem 1.1 (Parallel repetition theorem for general games)**.** *Suppose G is a two-prover game such that* $\mathrm{val}(G) \leq 1 - \varepsilon$ *and let* $|\Sigma_X| \, |\Sigma_Y| \leq s$*. Then, for any* $k \geq 0$*,*

$$\mathrm{val}(G^k) \quad \leq \quad \left(1 - \varepsilon^3/2\right)^{\Omega(k/\log s)} .$$

Theorem 1.2 (Parallel repetition theorem for projection games)**.** *Suppose G is a projection game such that* $\mathrm{val}(G) \leq \rho$*. Then, for any* $k \geq 0$*,*

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

Although a lot of these results are substantial simplifications of earlier proofs, they continue to be involved and delicate. Arguably, one might still hesitate to call them *elementary* proofs.

Recently, Moshkovitz [Mos14] came up with an ingenious method to prove a parallel repetition theorem for certain projection games by slightly modifying the underlying game via a process that the author called *fortification*. The method of fortification suggested in [Mos14] was a rather mild change to the underlying game and proving parallel repetition for such *fortified projection games* was sufficient for most applications. The advantage of fortification was that parallel repetition theorem for fortified games had a simple, elementary and elegant proof as seen in [Mos14].

1.1 Fortified games

Fortified games will be described more formally in Section 2, but we give a very rough overview here. Moshkovitz showed that there is an easy way to bound the value of repeated game if we knew that the game was *robust on large rectangles*. We shall first need the notion of *symmetrized projection games*.

Symmetrized Projection games. Given a projection game G on $((X,Y),E)$, the symmetrized game G_{sym} is a game on $((X,X),E')$ such that for every $y \in Y$ with $(x,y),(x',y) \in E$, there is an edge $(x,x') \in E'$ with the constraint $\pi_{(x,y)}(\sigma_x) = \pi_{(x',y)}(\sigma_{x'})$.

For projection games, it would be more convenient to work with the above symmetrized version for reasons that shall be explained shortly. It is not hard to see that $\text{val}(G)$ and $\text{val}(G_{\text{sym}})$ are within a quadratic factor of each other. Thus for projection games, we shall work with the game G_{sym} instead of the original game G.

Definition 1.3 ((δ,ε)-robust games)**.** *Let G be a two-prover game on $((X,Y),E)$. For any pair of sets $S \subseteq X, T \subseteq Y$, let $G_{S\times T}$ be the game where the verifier chooses his random query $(x,y) \in E$ conditioned on the event that $x \in S$ and $y \in T$.*

*G is said to be (δ,ε)-*robust *if for every $S, T \subseteq X$ with $|S| \geq \delta|X|$ and $|T| \geq \delta|Y|$ we have that*

$$\text{val}(G_{S\times T}) \quad \leq \quad \text{val}(G) + \varepsilon.$$

Theorem 1.4 (Parallel repetition for robust projection games [Mos14])**.** *Let G be a projection game on a bi-regular bipartite graph $((X,Y),E)$ with alphabets Σ_X and Σ_Y. For any positive integer k, if $\varepsilon, \delta > 0$ are parameters such that $2\delta|\Sigma_Y|^{k-1} \leq \varepsilon$ and G_{sym} is (δ,ε)-robust, then*[1]

$$\text{val}(G_{sym}^k) \quad \leq \quad \left(\text{val}(G_{sym}) + \varepsilon\right)^k + k\varepsilon.$$

Not all projection games are robust on large rectangles, but Moshkovitz suggested a neat way of slightly modifying a projection game and making it robust. This process was called *fortification*.

On a high level, for any two-prover game, the verifier chooses to verify a constraint corresponding to an edge (x,y) but is instead going to sample several other dummy vertices and give the provers two sets of D vertices $\{x_1,\ldots,x_D\}$ and $\{y_1,\ldots,y_D\}$ such that $x = x_i$ and $y = y_j$ for some i and j. The provers are expected to return labels of all D vertices sent to them but the verifier checks consistency on just the edge (x,y). This is very similar to the "confuse/match" perspective of Feige and Kilian [FK94].

To derandomize this construction, Moshkovitz [Mos14] uses a pseudo-random bipartite graph where given a vertex w, the provers are expected to return labels of all its neighbours (Definition 2.1). The most natural candidate of such a pseudo-random graph is an (δ,ε)-extractor, as we really want to ensure that conditioned on "large enough events" S and T, the underlying distribution on the constraints does not change much. This makes a lot of intuitive sense, since on choosing a random element of S and then a random neighbour, the extractor property guarantee that the induced distribution on vertices in X is ε-close to uniform. Thus, it is natural to expect that conditioning on the events S and T should not change the underlying distribution on the constraints by more than $O(\varepsilon)$. This was the rough argument in [Mos14], which unfortunately turns out to be false. We elaborate on this in Section 3.2 and Appendix A.

[1] The following is the corrected statement from [Mos15].

A recent updated version [Mos15] of [Mos14] provides an different argument for the fortification lemma using a stronger extractor. We discuss this at the end of Section 1.2.

1.2 Our contributions

We present a fix to the approach of [Mos14], by describing a way to transform any given game instance G into a robust instance G^* with the same value following the framework of [Mos14] but using a different graph for concatenation, and a different analysis.

We first describe a concrete counter-example to the original argument of [Mos14] in Section 3.2, that shows concatenating (Definition 2.1) with an arbitrary (δ, ε)-extractor is insufficient. In fact, as we show in Appendix B, concatenating with *any* left-regular graph with left-degree by $o(1/\varepsilon\delta)$ fails to make arbitrary instances (δ, ε)-robust. We instead use bipartite graphs called *fortifiers*, defined below.

Definition 1.5 (Fortifiers)**.** *A bipartite graph $H = ((W, X), E_H)$ is an $(\delta, \varepsilon_1, \varepsilon_2)$-*fortifier *if for any set $S \subseteq W$ such that $|S| \geq \delta|W|$, if π is the probability distribution on X induced by picking a uniformly random element w from S, and a uniformly random neighbor x of w, then*

$$\begin{aligned} |\pi - \mathbf{u}|_1 &\leq \varepsilon_1, \\ \|\pi - \mathbf{u}\|^2 &\leq \frac{\varepsilon_2}{|X|}. \end{aligned}$$

Notice that a fortifier is an extractor, with the additional condition that the ℓ_2-distance of π from the uniform distribution is small. This is what enables us to show that concatenation with a fortifier produces a robust instance.

Theorem 1.6 (Fortifiers imply robustness)**.** *Suppose G is a general two-prover game on a bi-regular graph $((X, Y), E)$. Then, for any $\varepsilon, \delta > 0$, if $H_1 = ((W, X), E_1)$ and $H_2 = ((Z, Y), E_2)$ are $(\delta, \varepsilon, \varepsilon)$-fortifiers, then the concatenated game $G^* = H_1 \circ G \circ H_2$ is $(\delta, O(\varepsilon))$-robust.*

In particular, bipartite spectral expanders are good fortifiers, as Lemma 2.8 shows. This gives us our main result which follows from Lemma 2.8 and Theorem 1.6:

Corollary 1.7. *Let G be a general two-prover game on a bi-regular graph $((X, Y), E)$. For any $\varepsilon, \delta > 0$, if $H_1 = ((W, X), E_1)$ and $H_2 = ((Z, Y), E_2)$ are two λ-expanders (Definition 2.3) with $\lambda < \varepsilon\sqrt{\delta}$ then concatenated game $G^* = H_1 \circ G \circ H_2$ is $(\delta, 4\varepsilon)$-robust.*

As one would expect, the condition on the fortifier can be relaxed if the underlying graph of the original label cover instance is a spectral-expander. We prove the following theorem. Theorem 1.6 follows from this theorem by setting $\lambda_0 = 1$.

Theorem 1.8. *Let G be a two-prover game on bi-regular graph $((X, Y), E)$ where G is an λ_0-expander. Then for any $\varepsilon, \delta > 0$, if $H_1 = ((W, X), E_1)$ and $H_2 = ((Z, Y), E_2)$ are $(\delta, \varepsilon, (\varepsilon/\lambda_0))$-fortifiers, then the concatenated game $G^* = H_1 \circ G \circ H_2$ is $(\delta, O(\varepsilon))$-robust.*

One could ask if the definition of a fortifier is too strong, or if a weaker object would suffice. We argue in Section 3.1 that if we proceed through concatenation, fortifiers are indeed necessary to make a game robust.

Bipartite Ramanujan graphs of degree $\Theta(1/\varepsilon^2\delta)$ have $\lambda < \varepsilon\sqrt{\delta}$ and are therefore good fortifiers. In Appendix B, we show that this is almost optimal by proving a lower bound of $\Omega(1/\varepsilon\delta)$ on the left-degree of any graph that can achieve (δ, ε)-robustness. This shows that our construction of using expanders to achieve robustness is almost optimal, in terms of the degree of the fortifier graph. Note that the degree of the fortifier is important as the alphabet size of the concatenated game is the alphabet size of the original game raised to the degree. There are known explicit constructions of bi-regular (δ, ε)-extractors with left-degree $\mathrm{poly}(1/\varepsilon)\mathrm{poly}\log(1/\delta)$. But the lower bound in Section 3.1 shows that (δ, ε)-extractors are not fortifiers if $\delta \ll \varepsilon$, which is usually the relevant setting (see Theorem 1.4 and Lemma 1.9).

Though all the above results are stated for bi-regular games, any two-prover game can be easily converted to one on a bi-regular graph or roughly the same value via standard tricks. We outline such a construction (similar to the construction in [DH13] for projection games) in Appendix D.

Independently, the author of [Mos14] came up with a different argument to obtain robustness of projection games by using a $(\delta, \varepsilon\delta)$-extractor. This is described in an updated version [Mos15] present on the author's homepage.

It is also seen from Theorem 1.8 that bi-regular $(\delta, \varepsilon\delta)$-extractors are indeed $(\delta, \varepsilon, \varepsilon)$-fortifiers as well. Using an expander instead is arguably simpler, and is almost optimal.

Remark. *Although this fix provides a proof of a Parallel Repetition Theorem for projection games following the framework of [Mos14], the degree of the fortifier is too large to get the required PCP for proving optimal hardness of the* SET-COVER *problem that Dinur and Steurer [DS14] obtained. See [Mos15] for a discussion on this.*

Remark about parallel repetition for general games

A fairly straightforward generalization Theorem 1.4 to robust general games on bi-regular graphs is the following.

Lemma 1.9 (Parallel repetition for general robust games)**.** *Let G be a general two-prover game on a bi-regular graph $((X, Y), E)$ with alphabets Σ_X and Σ_Y. For any positive integer k, if $\varepsilon, \delta > 0$ are parameters such that $2\delta|\Sigma_X \times \Sigma_Y|^{k-1} \leq \varepsilon$ and G is (δ, ε)-robust, then*

$$\mathrm{val}(G^k) \quad \leq \quad (\mathrm{val}(G) + \varepsilon)^k + k\varepsilon.$$

But it is to be noted that the fortification procedure via concatenating a fortifier makes $|\Sigma_X| = \exp(1/\delta)$ and in such scenarios $\delta|\Sigma_X| \gg 1$ making it infeasible to ensure $2\delta|\Sigma_X \times \Sigma_Y|^{k-1} \leq \varepsilon$. Hence, though Lemma 1.9 may be useful in cases where we know that the game G is robust via other means, the technique of fortification via concatenation increases the alphabet size too much for Lemma 1.9 to be applicable.

For the case of projection games, this is not an issue as Theorem 1.4 only requires $2\delta|\Sigma_Y|^{k-1} < \varepsilon$ and concatenating G_{sym} by a fortifier only increases $|\Sigma_X|$ and keeps Σ_Y unchanged. Thus, one can indeed choose ε and δ small enough to give a parallel repetition theorem for a robust version of an arbitrary projection game.

2 Preliminaries

Notation

- For any vector $\mathbf{a}$, let $|\mathbf{a}|_1 := \sum_i |\mathbf{a}_i|$, and $\|\mathbf{a}\| := \sqrt{\sum_i \mathbf{a}_i^2}$ be the ℓ_1 and ℓ_2-norms respectively.

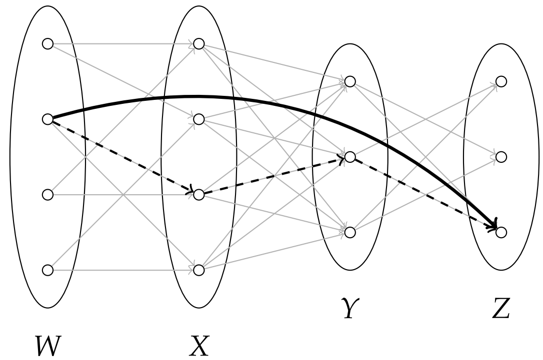

Figure 1: Concatenated Games

- We shall use $\mathbf{u}_S$ to refer to the uniform distribution on a set S. Normally, the set S would be clear from context and in such case we shall drop the subscript S.
- For any vector $\mathbf{a}$, we shall use $\mathbf{a}^{\parallel}$ to refer to the component along the direction of $\mathbf{u}$, and $\mathbf{a}^{\perp}$ to refer to the component orthogonal to $\mathbf{u}$.
- We shall assume that the underlying graph for the games is bi-regular. This is more or less without loss of generality via standard sampling tricks (see Appendix D).

We define the *concatenation* operation of a two-prover games with a bipartite graph that was alluded to in Section 1.1.

Definition 2.1 (Concatenation). *Given a two-prover game on a graph $G = ((X,Y), E)$ with a set of constraints ψ, a pair of alphabets Σ_X and Σ_Y, bipartite graphs $H_1 = ((W,X), E_1)$ with left degree D_1, and $H_2 = ((Z,Y), E_2)$ with left-degree D_2, the* concatenated game *is a game on the (multi) graph $H_1 \circ G \circ H_2 = ((W,Z), E_{H_1 \circ G \circ H_2})$ with $\Sigma_W = \Sigma_X^{D_1}$ and $\Sigma_Z = \Sigma_Y^{D_2}$. Label of a vertex $w \in W$ ($z \in Z$) can be thought of labels to its neighbors in $H_1(H_2)$ in a fixed order. For any edge $(w,z) \in E_{H_1 \circ G \circ H_2}$, there exists $(x,y) \in E$ such that $(w,x) \in E_1$, and $(z,y) \in E_2$. The constraint for this edge first obtains the label of x from w, and similarly obtains the label for y from the label of z, and checks the constraint $\psi_{(x,y)}$ according to the game G.*

Remark. *As mentioned earlier for projection games, as in [Mos14] we shall work with symmetrized version G_{sym}. In G_{sym} which is played on $((X,X), E_{sym})$, concatenating both sides with the same H_1 ensures that the resulting game G^* is still a symmetrized projection game, and that the concatenation operation only changes Σ_X and leaves Σ_Y unchanged for the underlying projection game.*

We state the results in a general setting as the focus here would be mainly on the study of distributions of edges of sub-graphs of concatenated graphs.

Lemma 2.2 (Concatenation preserves value). *[Mos14] Given any two-prover game on a bi-regular graph G, if H_1 and H_2 are bi-regular graphs, then we have:*

$$\mathrm{val}(H_1 \circ G \circ H_2) = \mathrm{val}(G).$$

Expanders, extractors and fortifiers

Definition 2.3 (Expanders)**.** *For any bi-regular bipartite graph $H = ((X,Y),E)$ with $|X| = |Y|$ and (left) degree D, we shall use $\lambda(H)$ to denote*

$$\lambda(H) \stackrel{def}{=} \max_{\mathbf{v}\perp\mathbf{u}} \frac{\|H\mathbf{v}\|}{\|\mathbf{v}\|}$$

*where the matrix H is an $|Y| \times |X|$ matrix (rows indexed by vertices in Y, and columns by vertices in X) defined by $H(y,x) = 1/D$ if $(x,y) \in E$ and it is 0 otherwise. For any $\lambda > 0$, a bi-regular bipartite graph H is an λ-*expander *if $\lambda(H) \leq \lambda$.*

More generally[2]*, if $|X| \neq |Y|$, we define $\lambda(H)$ as follows:*

$$\lambda(H) \stackrel{def}{=} \max_{\mathbf{v}\perp\mathbf{u}} \frac{\|H\mathbf{v}\|}{\|\mathbf{v}\|} \cdot \left(\frac{\|\mathbf{u}_X\|}{\|H\mathbf{u}_X\|} \right).$$

Informally, $\lambda(H)$ measures "how much *more* does the matrix H shrink $\mathbf{v} \perp \mathbf{u}_X$ compared to $\mathbf{u}_X$"?

Lemma 2.4 (Explicit expanders [BL06])**.** *For every $D > 0$, there exists a fully explicit family of bipartite graphs $\{G_i\}$, such that G_i is D-regular on both sides and $\lambda(G_i) \leq D^{-1/2}(\log D)^{3/2}$.*

Definition 2.5 (Extractors)**.** *A bipartite graph $H = ((X,Y),E)$ is an (δ,ε)-extractor if for every subset $S \subseteq X$ such that $|S| \geq \delta|X|$, if π is the induced probability distribution on Y by taking a random element of S and a random neighbour, then*

$$|\pi - \mathbf{u}|_1 \quad \leq \quad \varepsilon.$$

Lemma 2.6 (Explicit Extractors [RVW00])**.** *There exists explicit (δ,ε)-extractors $G = (X,Y,E)$ such that $|X| = O(|Y|/\delta)$ and each vertex of X has degree $D = O(\exp(\text{poly}(\log\log(1/\delta))) \cdot (1/\varepsilon^2))$.*

Our earlier definition of a fortifier (Definition 1.5) has properties of both an expander and an extractor. Indeed, we can build fortifiers by just taking a product an expander and an extractor.

Lemma 2.7. *If $H_1 = ((V,W),E_1)$ is a bi-regular (δ,ε)-extractor, and if $H_2 = ((W,X),E_2)$ is a bi-regular λ-expander, then the product graph $H_1 \cdot H_2$ is an $(\delta,\varepsilon,\lambda^2\varepsilon/\delta)$-fortifier.*

Proof. Let H_2 be the normalized adjacency matrix of graph H_2 and let π_S denote the probability distribution on W obtained by picking an element of $S \subseteq V$ uniformly and then choosing a random neighbour in H_1. Thus, $H_2\pi_S$ is the probability distribution on X induced by the uniform distribution on S and a random neighbour in $H_1 \cdot H_2$. We want to show for all S such that $|S| \geq \delta|V|$,

$$|H_2\pi_S - \mathbf{u}|_1 \leq \varepsilon \ \text{ and } \ \|H_2\pi_S - \mathbf{u}\|^2 \leq \frac{\lambda^2\varepsilon/\delta}{|X|}.$$

The first inequality is obtained as $|H_2\pi_S - \mathbf{u}|_1 = |H_2(\pi_S - \mathbf{u})|_1 \leq |\pi_S - \mathbf{u}|_1 \leq \varepsilon$, where we use the fact that $|H_2 v|_1 \leq |v|_1$ for any v and any normalized adjacency matrix, and $|\pi_S - \mathbf{u}|_1 \leq \varepsilon$ follows form the extractor property of H_1.

[2] We are not sure if this definition is standard, but is a natural generalization and precisely what we need in our proof.

As for the second inequality, observe that

$$\|\pi_S - \mathbf{u}\|^2 \le \max_{w\in W}(\pi_S(w)) \cdot |\pi_S - \mathbf{u}|_1 \le \varepsilon \cdot \max_{w\in W}(\pi_S(w)).$$

For a bi-regular extractor[3] H_1 of left-degree D, the degree of any $w \in W$ is $(|V| \cdot D/|W|)$ and the number of edges out of S is least $\delta|V| \cdot D$. Hence, $\max_w \pi_S(w) \le 1/(\delta|W|)$, which is achieved if all neighbours of w are in S. Therefore,

$$\begin{aligned} \|\pi_S - \mathbf{u}\|^2 &\le \frac{(\varepsilon/\delta)}{|W|} \\ \implies \|H_2(\pi_S - \mathbf{u})\|^2 &\le \lambda^2 \frac{|W|}{|X|} \|\pi_S - \mathbf{u}\|^2 \le \frac{|W|}{|X|} \cdot \frac{\lambda^2 \cdot (\varepsilon/\delta)}{|W|} = \frac{\lambda^2 \cdot (\varepsilon/\delta)}{|X|}. \end{aligned}$$

□

In particular, any bi-regular (δ, ε)-extractor is a $(\delta, \varepsilon, \varepsilon/\delta)$-fortifier. Hence, if the underlying graph G of the two-prover game is a $\sqrt{\delta}$-expander, then Theorem 1.8 states that merely using an (δ, ε)-extractor as suggested in [Mos14] would be sufficient to make it $(\delta, O(\varepsilon))$-robust.

Also, since any graph is trivially a 1-expander, a bi-regular $(\delta, \varepsilon\delta)$-extractor is also an $(\delta, \varepsilon, \varepsilon)$-fortifier. The following lemma also shows that expanders are also fortifiers with reasonable parameters as well.

Lemma 2.8. *Let $H = ((W, X), E_H)$ be any λ-expander. Then, for every $\delta > 0$, H is also a $(\delta, \sqrt{\lambda^2/\delta}, \lambda^2/\delta)$-fortifier.*

In particular, if $\lambda \le \varepsilon\sqrt{\delta}$, then H is an $(\delta, \varepsilon, \varepsilon)$-fortifier.

Proof. Let H be the normalized adjacency matrix of H. Let $S \subseteq W$ such that $|S| \ge \delta|W|$. We have,

$$\|\mathbf{u}_S^\perp\|^2 \le \frac{1}{\delta|W|}.$$

Hence, by the expansion property of H,

$$\|H\mathbf{u}_S - \mathbf{u}\|^2 := \|H\mathbf{u}_S^\perp\|^2 \le \lambda^2 \cdot \frac{|W|}{|X|} \cdot \|\mathbf{u}_S^\perp\|^2 \le \frac{\lambda^2/\delta}{|X|}.$$

$|H\mathbf{u}_S - \mathbf{u}|_1 \le \sqrt{\lambda^2/\delta}$ follows from above and Cauchy-Schwarz inequality. □

Although Lemma 2.8 shows that expanders are also fortifiers for reasonable parameters, the construction in Lemma 2.7 is more useful when the underlying graph for the two-prover game is already a good expander. For example, if the underlying graph G was a δ-expander, then Theorem 1.8 suggests that we only require a $(\delta, \varepsilon, \varepsilon/\delta)$-fortifier. Lemma 2.7 implies that an (δ, ε)-extractor is already a $(\delta, \varepsilon, \varepsilon/\delta)$-fortifier and hence is sufficient to make the game robust. The main advantage of this is the degree of δ-expanders must be $\Omega(1/\delta^2)$ whereas we have explicit (δ, ε)-extractors of degree $(1/\varepsilon^2)\exp(\mathrm{poly}\log\log(1/\delta))$ which has a much better dependence in δ. This dependence on δ is crucial for certain applications.

[3]The bound on the right-degree guaranteed by bi-regularity is crucial for this claim. Without this, extractors are not sufficient for fortification (Section 3.2).

3 Sub-games on large rectangles

Consider a concatenated general game $G^* = H_1 \circ G \circ H_2$ on $((W, Z), E_{H_1 \circ G \circ H_2})$ and $S \subseteq W$ and $T \subseteq Z$. Let μ_S (or μ_T) denote the induced distributions on X(or Y) obtained by picking a uniformly random element of S (or T) and taking a uniformly random neighbour in H_1 (or H_2). That is, the degree of any $x \in X$ (or $y \in Y$) within the set S (or T) is proportional to $\mu_S(x)$ (or $\mu_T(y)$) (See Figure 2).

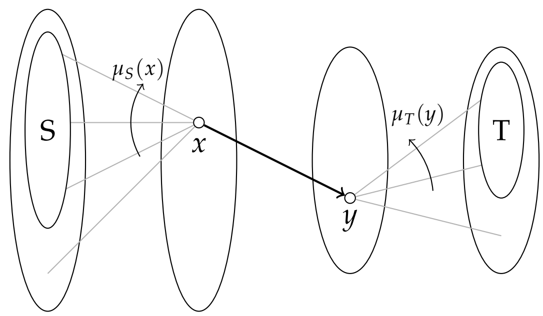

Figure 2: Sub-games on large rectangles

In a subgame $(G^*)_{S\times T}$, the distribution on verifier checking the underlying constraint on (x, y) is given by the following expression:

$$\pi_{x,y} \quad = \quad \frac{\mu_S(x)\mu_T(y)}{\sum\limits_{(x,y)\in G} \mu_S(x)\mu_T(y)}. \tag{3.1}$$

One way to show that the concatenated game G^* is $(\delta, O(\varepsilon))$-robust would be to show that the above distribution $\pi_{x,y}$ is $O(\varepsilon)$-close to uniform whenever $|S|, |T|$ have density at least δ because then the distribution on constraints that the verifier is going to check in $G^*_{S\times T}$ is $O(\varepsilon)$ close to the distribution on constraints in G. Hence, up to additive factor of $O(\varepsilon)$ the quantity $\mathrm{val}(G^*)_{S\times T}$ is same as $\mathrm{val}(G)$. The main question here what properties should H_1 and H_2 satisfy so that the above distribution is close to uniform?

3.1 Fortifiers are necessary

To prove that fortifiers are necessary, we shall restrict ourselves to games on graphs $G = ((X, X), E)$. In such a setting, we can choose to concatenate with the same graph H both sides. We show that if a bipartite graph $H = ((W, X), E_H)$, makes a game on a particular graph G, $(\delta, O(\varepsilon))$-robust, then H is a good fortifier.

As mentioned earlier, if the graph G had some expansion properties, then the requirements on the graph H to concatenate with can be relaxed. Thus, naturally, the worst case graph G is one that expands the least — a matching.

Lemma 3.1 (Fortifiers are necessary). *Let $\varepsilon, \delta > 0$ be small constants. Let $H = ((W, X), E_H)$ be a bi-regular graph, and let $G = ((X, X), E)$ be a matching. Suppose that for every subset $S, T \subseteq W$ with*

$|S|, |T| \geq \delta|W|$, the distribution (defined in Equation (3.1)) induced by the game $(H \circ G \circ H)_{S\times T}$ on the edges of G is ε-close to uniform. Then, for every $S \subseteq W$ with $|S| \geq \delta|W|$,

$$|\mu_S - \mathbf{u}|_1 = \varepsilon, \tag{3.2}$$

$$\|\mu_S - \mathbf{u}\|^2 = \frac{O(\varepsilon)}{|X|}. \tag{3.3}$$

Proof. It is clear that (3.2) is necessary as the distribution on constraints in the sub-game $(H \circ G \circ H)_{S\times W}$ (as defined in (3.1)) is essentially μ_S (as μ_T in this case is uniform).

As for (3.3), let us assume that

$$\|\mu_S - \mathbf{u}\|^2 = \frac{c}{|X|}.$$

Taking $T = S$, we obtain that the distribution (defined in Equation (3.1)) induced by the game $(H \circ G \circ H)_{S\times S}$ on the edges of G is given by

$$\pi_{x,x} = \frac{\mu_S(x)^2}{\sum_x \mu_S(x)^2} = \left(\frac{|X|}{1+c}\right) \cdot \mu_S(x)^2,$$

where the last equality used the fact that $\|\mu_S\|^2 = \left\|\mu_S^{\perp}\right\|^2 + \|\mathbf{u}\|^2$.

$$\begin{aligned}
\sum_{x\in X} \left| \left(\frac{|X|}{c+1}\right) \cdot \mu_S(x)^2 - \frac{1}{|X|} \right| &= \left(\frac{|X|}{1+c}\right) \cdot \sum_{x\in X} \left| \mu_S(x)^2 - \frac{c+1}{|X|^2} \right| \\
&= \left(\frac{|X|}{1+c}\right) \cdot \sum_{x\in X} \left| \mu_S(x) - \frac{\sqrt{c+1}}{|X|} \right| \cdot \left(\mu_S(x) + \frac{\sqrt{c+1}}{|X|} \right) \\
&\geq \left(\frac{1}{\sqrt{1+c}}\right) \cdot \sum_{x\in X} \left| \mu_S(x) - \frac{\sqrt{c+1}}{|X|} \right| \\
&\geq \left(\frac{1}{\sqrt{1+c}}\right) \cdot \left(\left(\sqrt{1+c} - 1\right) - \sum_{x\in X} \left| \mu_S(x) - \frac{1}{|X|} \right| \right) \\
&\geq \left(\frac{1}{\sqrt{1+c}}\right) \cdot \left(\left(\sqrt{1+c} - 1\right) - \varepsilon \right).
\end{aligned}$$

Thus, if the distribution on constraints is ε-close to uniform, then the above lower bound forces $c = O(\varepsilon)$. □

3.2 General (non-regular) extractors are insufficient

Suppose $H = ((W, X), E_H)$ is an arbitrary $(\delta, O(\varepsilon))$-extractor. Consider a possible scenario where there is a subset $S \subseteq W$ with $|S| \geq \delta|W|$ such that μ_S is of the form

$$\mu_S = \left(\varepsilon, \frac{1-\varepsilon}{|X|-1}, \ldots, \frac{1-\varepsilon}{|X|-1}\right).$$

Notice that this is a legitimate distribution that may be obtained from a large subset S as $|\mu_S - \mathbf{u}|_1$ is easily seen to be at most 2ε. However, if $G = ((X, X), E)$ was d-regular with $d = o(|X|)$, then using (3.1), the probability mass on the edge $(1,1)$ on the sub-game over $S \times S$ is

$$\pi_{1,1} = \left(\frac{\varepsilon^2}{\varepsilon^2 + O\left(\frac{\varepsilon d}{|X|}\right)} \right) \approx 1.$$

In other words, if such a distribution μ_S can be induced by the extractor, then the provers can achieve value close to 1 in the game $(H \circ G \circ H)_{S \times S}$ by just labelling the edge $(1,1)$ correctly. Thus, $(H \circ G \circ H)$ is not even $(\delta, 0.9)$-robust.

In Appendix A we show that we can adversarially construct a $(\delta, O(\varepsilon))$-extractor, although non-regular, that induces such a skew distribution. In Appendix B we also show that left-regular graphs of left-degree $o(1/\delta\varepsilon)$ are not fortifiers.

4 Robustness from fortifiers

In this section, we show that concatenating any two-prover game by fortifier(s) yields a robust game as claimed by Theorem 1.8.

Lemma 4.1 (Distributions from large rectangles are close to uniform)**.** *Let μ_S and μ_T be two probability distributions such that*

$$\left|\mu_S^{\perp}\right|_1 \leq \varepsilon_1 \quad \textit{and} \quad \left|\mu_T^{\perp}\right|_1 \leq \varepsilon_1, \tag{4.1}$$

$$\left\|\mu_S^{\perp}\right\|^2 \leq \left(\frac{\varepsilon_2}{|X|}\right) \quad \textit{and} \quad \left\|\mu_T^{\perp}\right\|^2 \leq \left(\frac{\varepsilon_2}{|Y|}\right). \tag{4.2}$$

Then for any bi-regular graph $G = ((X,Y), E)$ that is a λ_0-expander, the distribution on edge (x,y) (where $x \in X$ and $y \in Y$) given by (3.1) is $(2\varepsilon_1 + \varepsilon_1^2 + 2\lambda_0 \cdot \varepsilon_2)$-close to uniform.

As described in Section 3, if H_1 and H_2 are $(\delta, \varepsilon_1, \varepsilon_2)$-fortifiers, then for any set S and T of density at least δ, the distribution on the constraints of $(H_1 \circ G \circ H_2)_{S \times T}$ is given by (3.1). From the above lemma, it follows that the value of the game on any large rectangle can change only by the above bound on the statistical distance. By setting the parameters, Theorem 1.8 follows immediately from Lemma 4.1. Further, Corollary 1.7 also follows from Lemma 4.1 and Lemma 2.8 as any graph is trivially a 1-expander.

The rest of this section would be devoted to the proof of Lemma 4.1. For brevity, let us assume that $|X| = n$, $|Y| = m$ and let d be the left-degree of G. We shall prove Lemma 4.1 by proving the following two claims.

Claim 4.2.

$$\sum_{(x,y)\in G} \left| \frac{\mu_S(x)\mu_T(y)}{\sum\limits_{(x,y)\in G} \mu_S(x)\mu_T(y)} - \frac{\mu_S(x)\mu_T(y)}{d/m} \right| \quad \leq \quad \lambda_0 \cdot \varepsilon_2$$

Claim 4.3.

$$\sum_{(x,y)\in G} \left| \frac{\mu_S(x)\mu_T(y)}{d/m} - \frac{1}{n \cdot d} \right| \quad \leq \quad 2\varepsilon_1 + \varepsilon_1^2 + \lambda_0 \cdot \varepsilon_2$$

Clearly, Lemma 4.1 follows from Claim 4.2 and Claim 4.3.

Proof of Claim 4.2. If G denotes the normalized adjacency matrix of the graph G (that is, normalized so that $G\mathbf{u}_X = \mathbf{u}_Y$), then observe that $\sum_{(x,y)\in G} \mu_S(x)\mu_T(y) = d \cdot \langle G\mu_S, \mu_T\rangle$. If we resolve μ_S and μ_T in the direction of the uniform distribution and the orthogonal component, we have

$$\begin{aligned}
\langle G\mu_S, \mu_T\rangle &= \langle \mathbf{u}_Y, \mathbf{u}_Y\rangle + \left\langle G\mu_S^{\perp}, \mu_T^{\perp}\right\rangle = \frac{1}{m} + \left\langle G\mu_S^{\perp}, \mu_T^{\perp}\right\rangle \\
\implies \quad \left|\langle G\mu_S, \mu_T\rangle - \frac{1}{m}\right| &\leq \lambda_0 \cdot \left\|\mu_S^{\perp}\right\| \cdot \left\|\mu_T^{\perp}\right\| \cdot \sqrt{\frac{n}{m}} \\
&\leq \left(\frac{\lambda_0 \cdot \varepsilon_2}{m}\right). \qquad \text{(using (4.2))}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\sum_{(x,y)\in G} \left|\frac{\mu_S(x)\mu_T(y)}{d\,\langle G\mu_S, \mu_T\rangle} - \frac{\mu_S(x)\mu_T(y)}{d/m}\right| &\leq \sum_{(x,y)\in G} \left(\frac{\mu_S(x)\mu_T(y)}{d\,\langle G\mu_S, \mu_T\rangle}\right) |1 - m\,\langle G\mu_S, \mu_T\rangle| \\
&\leq \lambda_0 \cdot \varepsilon_2.
\end{aligned}$$

□

Proof of Claim 4.3.

$$\sum_{(x,y)\in G} \left|\frac{\mu_S(x)\mu_T(y)}{d/m} - \frac{1}{n\cdot d}\right| = \left(\frac{m}{d}\right) \sum_{(x,y)\in G} \left|\mu_S(x)\mu_T(y) - \frac{1}{n \cdot m}\right|.$$

Since $\mu_S(x) = \frac{1}{n} + \mu_S^{\perp}(x)$ and $\mu_T(y) = \frac{1}{m} + \mu_T^{\perp}(y)$,

$$\begin{aligned}
\left(\frac{m}{d}\right) \sum_{(x,y)\in G} \left|\mu_S(x)\mu_T(y) - \frac{1}{n\cdot m}\right| &= \left(\frac{m}{d}\right) \sum_{(x,y)\in G} \left|\frac{\mu_S^{\perp}(x)}{m} + \frac{\mu_T^{\perp}(y)}{n} + \mu_S^{\perp}(x)\mu_T^{\perp}(y)\right| \\
\text{(Using triangle inequality)} \quad &\leq \frac{1}{d} \sum_{(x,y)\in G} \left|\mu_S^{\perp}(x)\right| + \frac{m}{nd} \sum_{(x,y)\in G} \left|\mu_T^{\perp}(y)\right| \\
&\quad + \left(\frac{m}{d}\right) \sum_{(x,y)\in G} \left|\mu_S^{\perp}(x)\mu_T^{\perp}(y)\right| \\
&= \left|\mu_S^{\perp}\right|_1 + \left|\mu_T^{\perp}\right|_1 + \left(\frac{m}{d}\right) \sum_{(x,y)\in G} \left|\mu_S^{\perp}(x)\mu_T^{\perp}(y)\right|,
\end{aligned}$$

where the last equality uses the fact that G is a bi-regular graph. Define $f_S(x) \equiv |\mu_S^{\perp}(x)|$ is a vector with the entrywise absolute values of $\mu_S^{\perp}$, and similarly f_T. Then, the RHS above equation reduces to

$$\begin{aligned}
\left|\mu_S^{\perp}\right|_1 + \left|\mu_T^{\perp}\right|_1 + \left(\frac{m}{d}\right) \sum_{(x,y)\in G} \left|\mu_S^{\perp}(x)\mu_T^{\perp}(y)\right| &= \left|\mu_S^{\perp}\right|_1 + \left|\mu_T^{\perp}\right|_1 \\
&\quad + \left(\frac{m}{d}\right) \cdot \sum_{(x,y)\in G} f_S(x) f_T(y) \\
&= \left|\mu_S^{\perp}\right|_1 + \left|\mu_T^{\perp}\right|_1 + m\,\langle G f_S, f_T\rangle \\
\text{(Using (4.1))} \quad &\leq 2\varepsilon_1 + m \cdot \langle G f_S, f_T\rangle. \qquad (4.3)
\end{aligned}$$

A simple bound for $m \cdot \langle G f_S, f_T \rangle$ would $m \left\| G\mu_S^{\perp} \right\| \left\| \mu_T^{\perp} \right\|$ by Cauchy-Schwarz inequality. We can use the expansion of G again to estimate this better. Consider the decomposition $f_S = \alpha_1 \cdot \mathbf{u}_X + f_S^{\perp}$ and $f_T = \alpha_2 \cdot \mathbf{u}_Y + f_T^{\perp}$. It follows that $\alpha_1 = |f_S|_1$ and $\alpha_2 = |f_T|_1$, and hence $\alpha_1, \alpha_2 \leq \varepsilon_1$ by (4.1). Hence,

$$\begin{aligned} m \cdot \langle G f_S, f_T \rangle &= \alpha_1 \cdot \alpha_2 \, + \, m \cdot \left\langle G f_S^{\perp}, f_T^{\perp} \right\rangle \\ &\leq \varepsilon_1^2 \, + \, m \cdot \lambda_0 \cdot \left\| f_S^{\perp} \right\| \cdot \left\| f_T^{\perp} \right\| \cdot \sqrt{\frac{n}{m}} \\ &\leq \varepsilon_1^2 \, + \, m \cdot \lambda_0 \cdot \left\| \mu_S^{\perp} \right\| \cdot \left\| \mu_T^{\perp} \right\| \cdot \sqrt{\frac{n}{m}} \\ \text{(Using (4.2))} \qquad &\leq \varepsilon_1^2 \, + \, \lambda_0 \varepsilon_2. \end{aligned}$$

Combining this with (4.3), we get

$$\sum_{(x,y)\in G} \left| \frac{\mu_S(x)\mu_T(y)}{d/m} - \frac{1}{n \cdot d} \right| \quad \leq \quad 2\varepsilon_1 + \varepsilon_1^2 + \lambda_0 \varepsilon_2. \qquad \square$$

Acknowledgements

We would like to thank Dana Moshkovitz for several discussions and clarifications regarding the initial counter-example. We would also like to thank Mohammad Bavarian for pointing out that our proof might generalize for general two-prover games, and would like to thank Anup Rao for pointing out subtleties involving parallel repetition for general games. We also would like to thank Prahladh Harsha, Irit Dinur and Amir Shpilka for many fruitful conversations and comments on the write-up.

References

[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. *Proof verification and the hardness of approximation problems*. J. ACM, 45(3):501–555, May 1998. doi:10.1145/278298.278306.

[AS92] Noga Alon and Joel Spencer. *The Probabilistic Method*. John Wiley, 1992.

[AS98] Sanjeev Arora and Shmuel Safra. *Probabilistic checking of proofs: A new characterization of NP*. J. ACM, 45(1):70–122, January 1998. doi:10.1145/273865.273901.

[BG14] Mark Braverman and Ankit Garg. *Small value parallel repetition for general games*. Electronic Colloquium on Computational Complexity (ECCC), 21:95, 2014. To appear in STOC 2015. eccc:TR14/095.

[BL06] Yonatan Bilu and Nathan Linial. *Lifts, discrepancy and nearly optimal spectral gap**. Combinatorica, 26(5):495–519, 2006. doi:10.1007/s00493-006-0029-7.

[DH13] Irit Dinur and Prahladh Harsha. *Composition of Low-Error 2-Query PCPs Using Decodable PCPs*. SIAM J. Comput., 42(6):2452–2486, 2013. doi:10.1137/100788161.

[DS14] Irit Dinur and David Steurer. *Analytical approach to parallel repetition*. In *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 624–633. 2014. doi:10.1145/2591796.2591884.

[Fei91] Uriel Feige. *On the success probability of the two provers in one-round proof systems*. In *Structure in Complexity Theory Conference, 1991., Proceedings of the Sixth Annual*, pages 116–123. Jun 1991. doi:10.1109/SCT.1991.160251.

[FK94] Uriel Feige and Joe Kilian. *Two prover protocols: low error at affordable rates*. In *Proceedings of the Twenty-Sixth Annual ACM Symposium on Theory of Computing, 23-25 May 1994, Montréal, Québec, Canada*, pages 172–183. 1994. doi:10.1145/195058.195128.

[For89] Lance Jeremy Fortnow. *Complexity-theoretic aspects of interactive proof systems*. Ph.D. thesis, Massachusetts Institute of Technology, 1989.

[Hol09] Thomas Holenstein. *Parallel repetition: Simplification and the no-signaling case*. Theory of Computing, 5(1):141–172, 2009. doi:10.4086/toc.2009.v005a008.

[Mos14] Dana Moshkovitz. *Parallel repetition from fortification*. In *55th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2014, Philadelphia, PA, USA, October 18-21, 2014*, pages 414–423. 2014. doi:10.1109/FOCS.2014.51.

[Mos15] ———. *Parallel repetition from fortification*, 2015. http://people.csail.mit.edu/dmoshkov/papers/par-rep/final3.pdf. eccc:TR14/054/revision/2.

[Rao11] Anup Rao. *Parallel repetition in projection games and a concentration bound*. SIAM J. Comput., 40(6):1871–1891, 2011. doi:10.1137/080734042.

[Raz98] Ran Raz. *A parallel repetition theorem*. SIAM J. Comput., 27(3):763–803, 1998. doi:10.1137/S0097539795280895.

[Raz11] ———. *A counterexample to strong parallel repetition*. SIAM J. Comput., 40(3):771–777, June 2011. doi:10.1137/090747270.

[RT00] Jaikumar Radhakrishnan and Amnon Ta-Shma. *Bounds for dispersers, extractors, and depth-two superconcentrators*. SIAM J. Discrete Math., 13(1):2–24, 2000. doi:10.1137/S0895480197329508.

[RVW00] Omer Reingold, Salil Vadhan, and Avi Wigderson. *Entropy waves, the zig-zag graph product, and new constant-degree expanders and extractors*. In *Foundations of Computer Science, 2000. Proceedings. 41st Annual Symposium on*, pages 3–13. 2000. doi:10.1109/SFCS.2000.892006.

[Ver96] Oleg Verbitsky. *Towards the parallel repetition conjecture*. Theor. Comput. Sci., 157(2):277–282, May 1996. doi:10.1016/0304-3975(95)00165-4.

A An explicit extractor that does not provide robustness

Let $H = ((W, X), E_H)$ be any (δ, ε)-extractor. Let us assume that the extractor is left-regular with left-degree D, and let $m = |W|$ and $n = |X|$. For any $x \in X$ and $S \subseteq W$, let $d_S(x)$ denote the degree of x in S. Let us fix one $S \subset W$ such that $|S| = \delta|W|$.

We will transform the graph H so that the distribution induced by the set S looks like the counter-example described in Section 3.2 in the following two steps by altering the edges in the subgraph $S \times X$:

1. First change the degree into X from S to be exactly uniform.
2. Next further change the degrees into X from S to be like the counterexample

Both these operations can be achieved in a monotone fashion: for every $x \in X$, the neighborhood of every vertex is either a superset, or a subset of its neighborhood before each operation.

We will show that moving the edges this way does not perturb the indegree distribution from other large sets by too much, and the resulting graph is a $(\delta, O(\varepsilon))$ extractor as long as the number of edges we relocate is at most $O(\varepsilon\delta \cdot mD)$. This process will preserve the left-regularity of H but would *not* preserve bi-regularity.

First let us move edges (monotonically) from S into X create the uniform distribution on X. When doing this, the degree of each vertex changes by $\Delta_S(x) := |d_S(x) - \frac{\delta m D}{n}|$, where $d_S(x)$ was the old degree. From the extractor property, we know that:

$$\sum_{x \in X} \Delta_S(x) \;=\; \sum_{x \in X} (\delta m D) \left| \frac{d_S(x)}{\sum d_S(x)} - \left(\frac{1}{n}\right) \right| \;\leq\; \varepsilon\delta \cdot mD. \tag{A.1}$$

Every vertex $x \in X$ now has degree d^S_{avg}. Fix some vertex $x_1 \in X$, and relocate from every other $x \neq x_1$ any set of $\varepsilon \cdot d^S_{\text{avg}}$ edges to be incident on x_1. Thus, if $d'_S(x)$ refers to the new degrees, we have $d'_S(x_1)$ is $(1 + \varepsilon n) d^S_{\text{avg}}$ where as $d'_S(x)$ is $(1 - \varepsilon) d^S_{\text{avg}}$ for every other $x \neq x_1$.

The further change in degrees incurred on any $x \in X$ is $\Delta'_S(x) := \left| d'_S(x) - \frac{\delta m D}{n} \right|$. Since we this process only relocates $O(\varepsilon \cdot d^S_{\text{avg}} |X|)$ edges, we have

$$\sum_{x \in X} \Delta'_S(x) \;=\; \sum_{x \in X} \left| d'_S(x) - d^S_{\text{avg}} \right| \;\leq\; O(n \cdot \varepsilon \cdot d^S_{\text{avg}}) \;=\; O(\varepsilon\delta \cdot mD). \tag{A.2}$$

Thus, the neighbourhood of any vertex x has changed additively by at most $\Delta_S(x) + \Delta'_S(x)$. Therefore, for any subset $T \subseteq W$ of size at least $\delta|W|$,

$$\begin{aligned}
\sum_{x \in X} \left| d'_T(x) - d^T_{\text{avg}} \right| &\leq \sum_{x \in X} \left| d_T(x) - d^T_{\text{avg}} \right| + \sum_{x \in X} \left| d'_T(x) - d_T(x) \right| \\
&\leq \varepsilon |T| D + \sum_{x \in X} \left(\Delta_S(x) + \Delta'_S(x) \right) \\
&\leq \varepsilon |T| D + O(\varepsilon\delta \cdot mD) \quad \text{(using (A.1) and (A.2))} \\
&\leq O(\varepsilon \cdot |T| D).
\end{aligned}$$

Thus, the new graph after relocating edges is still an $(\delta, O(\varepsilon))$-extractor. This extractor, induces a distribution similar to the one described in Section 3.2 and hence cannot provide robustness.

B Lower bounds on degree of fortifiers

In this section, we will show that an attempt to make a game (δ, ε)-robust by concatenating any left-regular graph with left degree D fails if $D \leq o(1/\varepsilon\delta)$.

Lemma B.1. *Let $H = ((W, X), E_H)$ be a left-regular bipartite graph with left-degree $D = 1/(c \cdot \varepsilon\delta)$ for some $c > 0$, and small enough constants ε, δ. Then, there exists a subset $S \subseteq W$ with $|S| \geq \delta|W|$ such that if p was the distribution on X induced by the uniform distribution on S then*

$$\|p - \mathbf{u}\|^2 \;\geq\; \frac{\Omega(c\varepsilon)}{|X|}.$$

Proof. Let $d_{\text{avg}} = |W|D/|X|$. Note that at most $|X|/2$ vertices x satisfy $\deg(x) \geq 2d_{\text{avg}}$. Further, if there is a set S of $|X|/4$ vertices x that $\deg(x) < (0.5)d_{\text{avg}}$, then if p is the distribution on X induced by the uniform distribution on W, then $|p - \mathbf{u}|_1 > 1/4$ which implies that $\|p - \mathbf{u}\|_2^2 \geq \frac{1}{4|X|}$ by Cauchy-Schwarz.

Otherwise, there exists $X' \subset X$ such that $|X'| = c\,\varepsilon\delta^2|X|$ and for each $x \in X'$ we have $(0.5)d_{\text{avg}} < \deg(x) < 2d_{\text{avg}}$. Consider the set S_0 of all neighbours of X'. If $D < 1/(c\varepsilon\delta)$, we have $|S_0| \leq 2c\,\delta^2\varepsilon \cdot |W|D = 2\delta|W|$ which is a very small fraction of $|W|$ when δ is small enough. Consider an arbitrary set $S_1 \subseteq W$ such that $|S_1| = \delta m$, with $S_1 \cap S_0 = \emptyset$. Let $S_2 = S_0 \cup S_1$. Let π_1, π_2 be the probability distribution on X induced by S_1, S_2 respectively. Note that $|S_2| \leq 3\delta|W|$.

For every $x \in X'$, we know that $\pi_1(x) = 0$ and $\pi_2(x) = \Omega\left(\frac{1}{\delta|X|}\right)$. Therefore,

$$\|\pi_1 - \pi_2\|^2 \geq \Omega\left(\frac{c\delta^2\varepsilon|X|}{\delta^2|X|^2}\right) \;=\; \frac{\Omega(c\varepsilon)}{|X|}.$$

Since $\|\pi_1 - \pi_2\| \leq \|\pi_1 - \mathbf{u}\| + \|\pi_2 - \mathbf{u}\|$, we have that one of the sets S_1 or S_2 shows the validity of the lemma ☐

We thus immediately infer the following:

Corollary B.2. *For all small enough $\delta, \varepsilon > 0$, no left-regular graph $H = ((W, X), E_H)$ with left-degree $D = o(1/\varepsilon\delta)$ is an $(\delta, *, \varepsilon)$-fortifier.*

Note that any $(\delta, \varepsilon, \varepsilon)$-fortifier is in particular an (δ, ε)-extractor, and hence we also have that $D = \Omega((1/\varepsilon^2)\log(1/\delta))$ [RT00]. We also point out that the construction of Lemma 2.8 has left-degree $D = \tilde{O}(1/\varepsilon^2\delta)$. The above essentially shows this construction is almost optimal.

C Parallel repetition from fortification

We present a mild generalization of Theorem 1.4 to general bi-regular games, following essentially the same strategy as in [Mos14].

Lemma C.1. *Let $G = ((X, Y), E)$ be a (δ, ε)-robust general game that is bi-regular with $2\delta\left(|\Sigma_X||\Sigma_Y|\right)^{k-1} < \varepsilon$. Then,*

$$\mathrm{val}(G^k) \quad\leq\quad \mathrm{val}(G^{k-1}) \cdot (\mathrm{val}(G) + \varepsilon) \;+\; \varepsilon.$$

Proof. Consider any deterministic strategy for the provers. These are merely functions

$$f_1 : X^k \to \Sigma_X^k \quad \text{and} \quad f_2 : Y^k \to \Sigma_Y^k$$

that assign labels to the k queries asked by the verifier. For every $(k-1)$-tuple of queries $\bar{v} = (v_1, \ldots, v_{k-1})$ with each $v_i := (x_i, y_i) \in E$, and an arbitrary tuple of $(k-1)$ pairs of labels $\bar{\sigma} := ((\sigma_1, \sigma_1'), \ldots, (\sigma_{k-1}, \sigma_{k-1}')) \in (\Sigma_X \times \Sigma_Y)^{k-1}$, define the rectangle $\mathcal{R}_{\bar{v},\bar{\sigma}} := S_{\bar{v},\bar{\sigma}} \times T_{\bar{v},\bar{\sigma}}$ where

$$S_{\bar{v},\bar{\sigma}} = \{x_k \;:\; f_1(x_1, x_2, \ldots, x_k) \text{ assigns label } \sigma_i \text{ to } x_i \text{ for all } i \leq k-1\},$$
$$T_{\bar{v},\bar{\sigma}} = \{y_k \;:\; f_2(y_1, y_2, \ldots, y_k) \text{ assigns label } \sigma_i' \text{ to } y_i \text{ for all } i \leq k-1\}.$$

Also we shall call a rectangle $\mathcal{R}_{\bar{v},\bar{\sigma}}$ *accepting* if every coordinate (σ_i, σ_i') of $\bar{\sigma}$ satisfies the constraint on $v_i = (x_i, y_i)$ for all $1 \leq i \leq k-1$. In words, an accepting rectangle $\mathcal{R}_{\bar{v},\bar{\sigma}}$ is the set of all possible queries v_k for the last round such that the provers win on the first $(k-1)$ rounds with $x_1, \ldots, x_{k-1}$ and $y_1, \ldots, y_{k-1}$ getting labels $\sigma_1, \ldots, \sigma_{k-1}$ and $\sigma_1', \ldots, \sigma_{k-1}'$ respectively. We shall call a rectangle $\mathcal{R}_{\bar{v},\bar{\sigma}}$ "large" if $S_{\bar{v},\bar{\sigma}}$ and $T_{\bar{v},\bar{\sigma}}$ have density at least δ, and "small" otherwise. We shall partition the space of all possible queries $(v_1, \ldots, v_k)$ into the following sets. Note that v_k belongs to a unique rectangle $\mathcal{R}_{\bar{v},\bar{\sigma}}$.

- $\mathcal{A}_0 = \{(v_1, \ldots, v_k) \;:\; \mathcal{R}_{\bar{v},\bar{\sigma}} \text{ is not accepting}\}$
- $\mathcal{A}_1 = \{(v_1, \ldots, v_k) \;:\; \mathcal{R}_{\bar{v},\bar{\sigma}} \text{ is accepting and "large"}\}$
- $\mathcal{A}_2 = \{(v_1, \ldots, v_k) \;:\; \mathcal{R}_{\bar{v},\bar{\sigma}} \text{ is accepting and "small"}\}$

Observe that $|\mathcal{A}_1| + |\mathcal{A}_2| \leq \mathrm{val}(G^{k-1}) \cdot |E|^k$ because $\mathcal{A}_1 \cup \mathcal{A}_2$ is the set of queries on which the provers succeed on the first $(k-1)$ rounds.

Also, the projection of elements in set $\mathcal{A}_1$ to the kth coordinate, is essentially a union of large rectangles. By the (δ, ε)-robustness of G, any strategy of the provers can succeed on each large rectangle with probability at most $\mathrm{val}(G) + \varepsilon$. Hence, the provers succeed on at most a $(\mathrm{val}(G) + \varepsilon)$-fraction of points in $\mathcal{A}_1$.

Furthermore, since G is regular, we get $|\mathcal{A}_2|$ is at most $|E|^{k-1} \cdot 2\delta|E| \cdot |\Sigma_X \times \Sigma_Y|^{k-1} \leq \varepsilon\,|E|^k$ by the choice of δ and ε.[4]

Hence, the total number of queries on which the provers can succeed is upper bounded by $(\mathrm{val}(G) + \varepsilon)\,|\mathcal{A}_1| + |\mathcal{A}_2|$. It therefore follows that they succeed on at most a $\mathrm{val}(G^{k-1})(\mathrm{val}(G) + \varepsilon) + \varepsilon$ fraction of queries. □

Unfolding the recursion from the above lemma, we get the following generalization of Theorem 1.4.

Corollary C.2. *Let $G = ((X, Y), E)$ be a (δ, ε)-robust general game with $2\delta\,(|\Sigma_X||\Sigma_Y|)^{k-1} < \varepsilon$. Then,*

$$\mathrm{val}(G^k) \leq (\mathrm{val}(G) + \varepsilon)^k + k \cdot \varepsilon.$$

As mentioned earlier, this is not useful when say $|\Sigma_X| = \exp(O(1/\delta))$, which is unfortunately the case when an arbitrary game is made robust by concatenating with a fortifier.

[4] In the case of projection games, the set of $\bar{\sigma}$ that are accepting pairs for $\bar{v}$ can be indexed with Σ_Y^{k-1} instead of $(\Sigma_X \times \Sigma_Y)^{k-1}$, and that gets the better parameters for projection games as in Theorem 1.4.

D Making the graph bi-regular

In this section, we shall show that a general game on a graph can be converted to a slightly larger game on a bi-regular graph with almost the same value.

Lemma D.1. *Given a two-prover game G any graph $((X,Y),E)$. For every $\varepsilon > 0$, there is a polynomial time algorithm to construct a game G' with* $\mathrm{size}(G') = \mathrm{size}(G) \cdot \tilde{O}\left(\left(|\Sigma_X| + |\Sigma_Y|\right)/\varepsilon\right)^5$ *such that G' is on a bi-regular graph and* $\mathrm{val}(G') \leq \mathrm{val}(G) + \varepsilon$.

The rest of this section would be a proof of this. Suppose we have a graph $G = ((X,Y),E)$ that is possibly non-regular. We shall make some transformations on the graph to make it bi-regular such that it does not affect the value of the game by much. This is along the same lines as the technique used by Dinur and Harsha [DH13]. We shall need the following well-known *Expander Mixing Lemma*.

Lemma D.2 (Expander Mixing Lemma)**.** *Let $H = ((P,Q),E)$ be a λ-expander with $|P| = |Q|$. Then, for every subsets $A \subseteq P$ and $B \subseteq Q$,*

$$\left| \frac{|E(A,B)|}{|E|} \; - \; \frac{|A|}{|P|} \cdot \frac{|B|}{|Q|} \right| \; \leq \; \lambda$$

A proof of the above lemma may be found in any text that studies expanders graphs (for example, [AS92, Chapter 5]).

We shall make the graph bi-regular in two steps. We shall first make a transformation that makes it regular on the right side, and then repeat the same process on the left. But first, we would need to ensure that the degree on the Y side is large enough for the transformation to work. This is just done by creating d copies of every edge with the same constraint. The graph therefore becomes a multi-graph but the value remains the same.[5]

Thus, from now on, we assume that we are given a game $G = ((X,Y),E)$, with the minimum degree being "large enough", that we want to make biregular. The transformation of G to make it regular on right side is as follows (Figure 3):

For every vertex $y \in X$ with degree d_y, we shall have a set C_y of d_y vertices. Between the vertices C_y and the neighbourhood of y (in G), we shall add a λ-expander of degree d. The constraint on any edge between $x \in N(y)$ and a vertex in C_y would be the same as $\psi_{(x,y)}$. Let us denote this game by G_λ.

Lemma D.3. $\mathrm{val}(G_\lambda) \leq \mathrm{val}(G) + \lambda\,|\Sigma_Y|$.

Proof. Consider any labelling L_λ of G_λ. From this, let L be the natural randomized labelling for G such that $L(x) = L_\lambda(x)$ for every $x \in X$, and $L(y) = L_\lambda(y_i)$ be where y_i is a random element of C_y. For every $y \in Y$, let δ_y be the expected fraction of edges incident on y that are satisfied by this assignment.

$$\delta_y = \sum_{\sigma \in \Sigma_Y} \Pr[L(y) = \sigma] \cdot \Pr_{x \sim y}[(L(x), \sigma) \text{ satisfies } \psi_{(x,y)}]$$

By the definition of $\mathrm{val}(G)$, we know that $\sum_{y \in Y} d_y \delta_y \leq \mathrm{val}(G) \cdot |E|$.

[5] One could also do this by replicating every vertex d times and adding the edges between them.

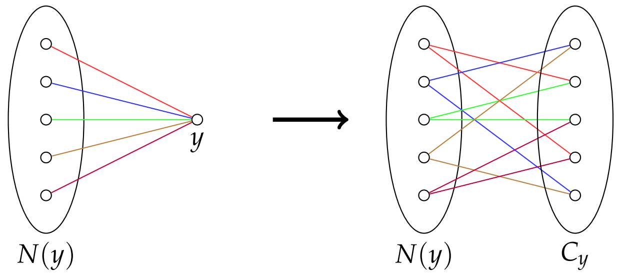

Figure 3: Enforcing bi-regularity

Subclaim D.4. *For every $y \in Y$, the fraction of edges between $N(y)$ and C_y that are satisfied by L_λ is at most $(\delta_y + \lambda|\Sigma_y|)$*

Before we prove this, let us see why this is sufficient to complete the proof of the lemma. The number of edges between C_y and $N(y)$ is exactly $d \cdot d_y$ where d is the degree of the expander. Therefore, the number of edges in G_λ that are satisfied is

$$\begin{aligned}\sum_{y \in Y} d \cdot d_y \cdot (\delta_y + \lambda|\Sigma_y|) &\leq d \cdot \sum_{y \in Y} d_y \delta_y \;+\; O(d\lambda\,|\Sigma_Y|) \cdot \sum_{y \in Y} d_y \\ &\leq (\mathrm{val}(G) + \lambda\,|\Sigma_Y|) \cdot |E_\lambda|\end{aligned}$$

as claimed by the lemma. Thus, it suffices to prove Subclaim D.4.

Proof of Subclaim D.4. The number of edges between C_y and $N(y)$ is $d \cdot d_y$. Partition the vertices of C_y into sets $\{C_{y,\sigma} \,:\, \sigma \in \Sigma_Y\}$ based on the label assigned by L_λ. For every $\sigma \in \Sigma_Y$, let A_σ denote the set of vertices $x \in N(y)$ such that $(L_\lambda(x), \sigma)$ satisfies $\psi_{(x,y)}$. Hence, the set of edges that are satisfied by L_λ is precisely $\bigcup_\sigma E(A_\sigma, C_{y,\sigma})$. By Lemma D.2,

$$\begin{aligned}\left|E(A_\sigma, C_{y,\sigma})\right| &\leq |A_\sigma| \cdot \left|C_{y,\sigma}\right| \cdot \frac{d}{d_y} \;+\; \lambda \cdot d \cdot d_y \\ \Longrightarrow \sum_{\sigma \in \Sigma_Y} \left|E(A_\sigma, C_{y,\sigma})\right| &\leq \sum_{\sigma \in \Sigma_Y} |A_\sigma| \cdot \left|C_{y,\sigma}\right| \cdot \frac{d}{d_y} \;+\; \lambda \cdot |\Sigma_Y| \cdot d \cdot d_y \\ &= (d \cdot d_y) \sum_{\sigma \in \Sigma_Y} \Pr[L(y) = \sigma] \cdot \Pr_{x \sim y}[(L(x), \sigma) \text{ satisfies } \psi_{(x,y)}] \\ &\quad + \;\lambda\,|\Sigma_Y| \cdot d \cdot d_y \\ &= \left(\delta_y \;+\; \lambda \cdot |\Sigma_Y|\right) \cdot (d \cdot d_y)\end{aligned}$$

as claimed, since the number of edges is $d \cdot d_y$. □ (Subclaim D.4)

That hence finishes the proof of the Lemma. □

This operation ensures that the right-degree of the game G_λ is d and the value changes by at most $\varepsilon/2$ if $\lambda < (\varepsilon/2\,|\Sigma_Y|)$. By Lemma 2.4, we can choose explicit constructions of expanders with $d = \tilde{O}(1/\lambda^2) = \tilde{O}((|\Sigma_Y|\,/\varepsilon)^2)$. The graph is now right-regular with degree d, and the degree of every $x \in X$ has increased by a factor of d. Repeating the same process for the other side makes both sides regular and the value changes by at most ε. □(Lemma D.1)