THEORY OF COMPUTING, Volume 18 (5), 2022, pp. 1–28
www.theoryofcomputing.org

SPECIAL ISSUE: CCC 2019

UG-hardness to NP-hardness by Losing Half

Amey Bhangale* Subhash Khot†

Received July 24, 2019; Revised March 4, 2022; Published March 15, 2022

Abstract. The 2-to-2 Games Theorem (Khot *et al.*, STOC'17, Dinur *et al.*, STOC'18 [2 papers], Khot *et al.*, FOCS'18) shows that for all constants $\varepsilon > 0$, it is NP-hard to distinguish between Unique Games instances with some assignment satisfying at least a $(\frac{1}{2} - \varepsilon)$ fraction of the constraints *vs.* no assignment satisfying more than an ε fraction of the constraints. We show that the reduction can be transformed in a non-trivial way to give stronger completeness: For at least a $(\frac{1}{2} - \varepsilon)$ fraction of the vertices on one side, all the constraints associated with them in the Unique Games instance can be satisfied.

We use this guarantee to convert known UG-hardness results to NP-hardness. We show:

1. Tight inapproximability of the maximum size of independent sets in degree-d graphs within a factor of $\Omega\left(\frac{d}{\log^2 d}\right)$, for all sufficiently large constants d.

A preliminary version of this paper appeared in the Proceedings of the 34th Computational Complexity Conference (CCC'19) [12].

*Research supported by Irit Dinur's ERC-CoG grant 772839.

†Research supported by NSF CCF-1813438, Simons Collaboration on Algorithms and Geometry, and Simons Investigator Award.

ACM Classification: F.2

AMS Classification: 68Q17

Key words and phrases: NP-hardness, inapproximability, Unique Games Conjecture

 DOI: 10.4086/toc.2022.v018a005

2. For all constants $\varepsilon > 0$, NP-hardness of approximating the size of the Maximum Acyclic Subgraph within a factor of $\frac{2}{3} + \varepsilon$, improving the previous ratio of $\frac{14}{15} + \varepsilon$ by (Austrin *et al.*, Theory of Computing, 2015).
3. For all constants $\varepsilon > 0$ and for any predicate $P^{-1}(1) \subseteq [q]^k$ supporting a balanced pairwise independent distribution, given a P-CSP instance with value at least $\frac{1}{2} - \varepsilon$, it is NP-hard to satisfy more than a $\frac{|P^{-1}(1)|}{q^k} + \varepsilon$ fraction of constraints.

1 Introduction

The Unique Games Conjecture is a central open problem in the theory of computing. It states that for a certain constraint satisfaction problem over a large alphabet, called *Unique Games* (UG), it is NP-hard to decide whether a given instance has an assignment that satisfies a $(1 - \varepsilon)$ fraction of the constraints or there is no assignment that satisfies even an ε fraction of the constraints for an arbitrarily small constant $\varepsilon > 0$. Since the formulation of the conjecture, it has found interesting connections to tight hardness-of-approximation results for many optimization problems [18, 19, 23, 27, 17, 22, 24, 25]. One of the most notable implications is the result of Raghavendra [27] which informally can be stated as follows: Assuming the NP-hardness of approximating this single CSP (Unique Games) implies tight hardness for approximating every other constraint satisfaction problem, stated in terms of the integrality gap of a certain canonical SDP.

The Unique Games Conjecture was inspired by the NP-hardness of approximating a problem called Label Cover [1]. A Label Cover instance $G = (A, B, E, \Sigma_A, \Sigma_B, \{\pi_e\}_{e \in E})$ consists oftwo sets of variables, A and B, and a bipartite graph between them with the edge set E. The variables from A take values from some alphabet Σ_A and variables from B take values from Σ_B. Every edge e in E has a d-to-1 projection[1] constraint $\pi_e : \Sigma_A \to \Sigma_B$.

For an edge $e(a, b)$, a label α to a and a label β to b satisfy the edge e iff $\pi_e(\alpha) = \beta$. In this language, Unique Games is a Label Cover instance where all the constraints are 1-to-1. We denote an instance of Unique Games by $G = (A, B, E, [L], \{\pi_e\}_{e \in E})$ where $\Sigma_A = \Sigma_B = [L]$.

Given an instance of Unique Games, the goal is to find an assignment to the vertices that satisfies a *good* fraction of the edges. An instance is called ε-satisfiable if there exists an assignment $\sigma : A \cup B \to [L]$, that satisfies at least an ε fraction of the edges in the graph. The Unique Games Conjecture [18] states that for every $\varepsilon > 0$, there exists L such that given a Unique Games instance which is $(1 - \varepsilon)$-satisfiable, it is NP-hard to find an ε-satisfying assignment. Note that there is a polynomial-time algorithm that given a 1-satisfiable instance of Unique Games, finds a 1-satisfying assignment. Recent work [20, 14, 15, 21] has shown that for every constant $\varepsilon > 0$, it is NP-hard to find an ε-satisfying assignment for a given Label Cover instance with 2-to-1 projection constraints, even if the instance is $(1 - \varepsilon)$-satisfiable. This directly implies the following inapproximability for Unique Games.

[1]A constraint $\pi_e : \Sigma_A \to \Sigma_B$ is called a d-to-1 projection constraint, if every $\beta \in \Sigma_B$ has exactly d preimages.

Theorem 1.1 (UG-hardness with completeness 1/2 [20, 14, 15, 21])**.** *For every constant $\varepsilon > 0$, there exists Σ such that for Unique Games instance over Σ, it is* NP*-hard to distinguish between the following two cases:*

- *YES case: The instance is $(\frac{1}{2} - \varepsilon)$-satisfiable.*
- *NO case: No assignment satisfies an ε fraction of the constraints.*

Although we do not improve upon this theorem in terms of the inapproximability gap, we show a stronger guarantee in the YES case. Specifically, we show that in the YES case, there is at least a $\frac{1}{2} - \varepsilon$ fraction of the *vertices* on, say, the left side such that *all* the edges incident on them are satisfied by some assignment and also the instance is left-regular. This clearly implies the above theorem. Formally, the main theorem that we prove is as follows. (See Definition 2.1 for a formal definition of Unique Games.)

Theorem 1.2 (Main)**.** *For every constant $\delta > 0$, there exists $L \in \mathbb{N}$ such that the following holds. Given an instance $G = (A, B, E, [L], \{\pi_e\}_{e \in E})$ of Unique Games, which is regular on the A side, it is* NP*-hard to distinguish between the following two cases:*

- *YES case: There exists a set $A' \subseteq A$ of size $(\frac{1}{2} - \delta)|A|$, and an assignment that satisfies all the edges incident on A'.*
- *NO case: Every assignment satisfies at most a δ fraction of the edges.*

We will denote by $\mathsf{val}(G)$ the maximum fraction, over all assignments, of the edges satisfied and $\mathsf{sval}(G)$ to be the maximum fraction, over all assignments, of the vertices in A such that all the edges incident on them are satisfied. Thus, the above theorem says that for every $\delta > 0$ there exists a label set $[L]$ such that it is NP-hard to distinguish between the cases $\mathsf{sval}(G) \geqslant \frac{1}{2} - \delta$ and $\mathsf{val}(G) \leqslant \delta$.

1.1 $(\frac{1}{2} - \varepsilon)$-satisfiable UG vs. $(1 - \varepsilon)$-satisfiable UG

Let $\varepsilon > 0$ be a very small constant. In the $(1-\varepsilon)$-satisfiable Unique Games instance, by simple averaging argument it follows that for any satisfying assignment $\sigma : A \cup B \to [L]$, there exists $A' \subseteq A$, $|A'| \geqslant (1 - \sqrt{\varepsilon})|A|$ such that for all $v \in A'$, at least a $(1 - \sqrt{\varepsilon})$ fraction of the edges incident on v are satisfied. Having such a large A' is crucial in many UG-reductions. For example, a typical k-query PCP, used in proving UG-hardness of approximation for k-ary CSPs, samples $v \in A$ uniformly at random and k neighbors of $u_1, u_2, \ldots, u_k$ of v uniformly at random. Thus, with probability at least $(1-\sqrt{\varepsilon})(1-k\sqrt{\varepsilon}) \approx 1$ all the edges (u, v_i) are satisfiedby any $(1-\varepsilon)$-satisfying assignment σ.

In contrast to this, if we take a $(1/2)$-satisfiable UG instance then the probability that all the edges (v, u_i) are satisfied is at most $1/2^k$ in the worst case. Therefore, in converting a known UG-hardness result to an NP-hardness result using the NP-hardness of Unique Games with gap $(\frac{1}{2} - \varepsilon, \varepsilon)$, it is not always the case that we lose *'only half'* in completeness.

Another important property of the Unique Games instance which was used in many reductions is that it has stronger completeness; there is a $(1-\delta)$ fraction of the vertices on one side such that *all* the edges incident on them are satisfied, i. e., $\mathsf{sval}(G) \geqslant 1-\delta$ instead of $\mathsf{val}(G) \geqslant 1-\delta$. For example, this property was crucial in the hardness of approximating the maximum size of independent sets in bounded-degree graphs [4] and in many other reductions [8, 11].

As shown in [23], having completeness $\mathsf{val}(G) \geqslant 1-\delta$ for all sufficiently small $\delta > 0$ is equivalent to having completeness $\mathsf{sval}(G) \geqslant 1-\delta'$ for all sufficiently small $\delta' > 0$. It was crucial in the reduction that the $\mathsf{val}(G)$ is arbitrarily close to 1 for the equivalence to hold. We do not know a black-box way of showing the equivalence of $\mathsf{val}(G) = c$ and $\mathsf{sval}(G) = c$ for any $c < 1$. Thus, in order to prove Theorem 1.2 with a stronger completeness guarantee, we crucially exploit the structure of the game given by the known proofs of the 2-to-1 theorems [20, 14, 15, 21] mentioned in the introduction.

1.2 Implications

Using Theorem 1.2, we show the following hardness results by going over the known reductions based on the Unique Games Conjecture.

Independent sets in degree-d graphs The first application is approximating the maximum sized independent set in a degree-d graph, where d is a large constant.

Theorem 1.3. *It is* NP*-hard (under randomized reductions) to approximate independent sets in a degree-d graph within a factor of $O\left(\frac{d}{\log^2 d}\right)$, where d is a constant independent of the number of vertices in the graph.*

This improves the NP-hardness of approximation within a factor $O\left(\frac{d}{\log^4 d}\right)$, as shown in Chan [13] as well as shows the tightness of the randomized polynomial-time approximation algorithm given by Bansal *et al*. [7].

Max-Acyclic Subgraph Given a directed graph $G(V,E)$, the Max-Acyclic Subgraph problem is to determine the maximum fraction of edges $E' \subseteq E$ such that removal of $E \setminus E'$ makes the graph acyclic (removes all the cycles). We can always make a graph acyclic by removing at most half the edges; take any arbitrary ordering of the vertices and remove either all the forward edges or all the backward edges whichever is minimum. This gives a trivial $(1/2)$-approximation algorithm. Guruswami *et al*. [17] showed this is tight by showing that assuming the Unique Games Conjecture, it is NP-hard to approximate Max-Acyclic Subgraph within a factor of $\frac{1}{2}+\varepsilon$ for all $\varepsilon > 0$. In terms of NP-hardness, Austrin *et al.* [5] showed NP-hardness of approximating Max-Acyclic Subgraph within a ratio of $\frac{14}{15}+\varepsilon$, improving upon the previous bound of $\frac{65}{66}+\varepsilon$ by Newman [26]. Our next theorem shows an improved inapproximability of $\frac{2}{3}+\varepsilon$. One interesting feature of our reduction is that it shows that in the worst case, it is NP-hard to perform better than the trivial algorithm described above on instances with value at least $3/4$.

Theorem 1.4. *For every constant $\varepsilon > 0$, it is* NP*-hard to approximate the Max-Acyclic Subgraph problem within a factor of $\frac{2}{3} + \varepsilon$.*

We note that Theorem 1.1 along with the reduction from [17] implies NP-hardness of approximation of the Max-Acyclic Subgraph problem within a factor of $\frac{4}{5} + \varepsilon$. (See Remark 5.5 for a proof sketch.) Therefore, Theorem 1.4 improves upon this bound too.

Predicates supporting balanced pairwise independent distributions The next result is approximating Max-k-CSP(P) for a predicate $P : [q]^k \to \{0,1\}$ where $P^{-1}(1)$ supports a balanced pairwise independent distribution, i. e., there exists a distribution on $P^{-1}(1)$ such that 1) the marginal distribution on each coordinate is uniform and 2) the distribution is pairwise independent. Given an instance of Max-k-CSP(P), a random assignment satisfies a $\frac{|P^{-1}(1)|}{q^k}$ fraction of the constraints in expectation. Austrin–Mossel [6] showed that given a $(1-\varepsilon)$-satisfiable instance of Max-k-CSP(P), it is UG-hard to find an assignment that satisfies more than a $\frac{|P^{-1}(1)|}{q^k} + \varepsilon$ fraction of the constraints for any constant $\varepsilon > 0$. This notion is called *approximation resistance* of a predicate P where the best efficient algorithm cannot perform any better than just choosing a random assignment, even if the instance is almost satisfiable. Showing approximation resistance of such predicates *unconditionally* (i. e., assuming only P $\neq$ NP) has received significant attention. For instance, Chan [13] showed that a predicate P is approximation resistant if $P^{-1}(1)$ contains a subgroup that supports a balanced pairwise independent distribution. Also, Barak et al. [9] showed that certain class of algorithms, namely sum-of-squares, cannot be used to perform better than the random assignment on almost satisfiable instances of Max-k-CSP(P) where P supports a balanced pairwise independent distribution.

Our main theorem shows approximation resistance of such predicates but on instances which are $(1/2)$-satisfiable. If we use Theorem 1.2 as a starting point of the reduction in [6], we get the following NP-hardness result.

Theorem 1.5. *If a predicate $P : [q]^k \to \{0,1\}$ supports a balanced pairwise independent distribution, then for every constant $\varepsilon > 0$, it is* NP*-hard to find a solution with value $\frac{|P^{-1}(1)|}{q^k} + \varepsilon$ if a given P-CSP instance is $(\frac{1}{2} - \varepsilon)$-satisfiable.*

Other Results Theorem 1.2 implies many more NP-hardness results in a straightforward way by going over the known UG-reductions but we shall restrict ourselves to proving only the above three theorems. We only state the following important implication which follows from the result of Raghavendra [27] and our main theorem. We refer to [27] for the definition of (c,s) SDP integrality gap of a P-CSP instance.

Theorem 1.6. *(Informal) For every constant $\varepsilon > 0$, if a P-CSP has a (c,s) SDP integrality gap instance, then it is* NP*-hard to distinguish between $(\frac{c}{2} - \varepsilon)$-satisfiable instances from at most $(s+\varepsilon)$-satisfiable instances.*

The reduction actually gives a stronger result; instead of completeness $(\frac{c}{2} - \varepsilon)$ one can get $(\frac{c}{2} + \frac{r}{2} - \varepsilon)$ where $r = \frac{|P^{-1}(1)|}{q^k}$ for a predicate $P : [q]^k \to \{0,1\}$.

1.3 Overview of the proof

In this section, we give an overview of the proof of Theorem 1.2. The main idea that goes into proving Theorem 1.2 is very simple and we elaborate it next.

Let $V = \mathbb{F}_2^n$ and let $\mathrm{Gr}(V,\ell)$ denote the set of all ℓ-dimensional linear subspaces of V. Consider the tables $F : \mathrm{Gr}(V,\ell) \to \mathbb{F}_2^\ell$ and $H : \mathrm{Gr}(V,\ell-1) \to \mathbb{F}_2^{\ell-1}$, where for a subspace L (L'), $F[L]$ ($H[L']$) represents a linear function $F[L] : L \to \mathbb{F}_2$ ($H[L'] : L' \to \mathbb{F}_2$) on the subspace, by fixing an arbitrarily chosen basis of L (L').[2] The intention is that there should be some global linear function $g : V \to \mathbb{F}_2$, such that for every L, $F[L] = g|_L$, and for every L', $H[L'] = g|_{L'}$. We consider the Grassmann 2-to-1 test $\mathcal{T}_1$ from Figure 1.3 as a means to verify that the tables, F and H, are indeed obtained by restricting some global linear function.

Given tables F and H, where F assigns to each ℓ-dimensional subspace $L \in \mathrm{Gr}(V,\ell)$ a linear function $F[L] : L \to \mathbb{F}_2$, and H assigns to each $(\ell-1)$-dimensional subspace $L' \in \mathrm{Gr}(V,\ell-1)$ a linear function $H[L'] : L' \to \mathbb{F}_2$.

- Select an $(\ell-1)$-dimensional subspace $L' \subseteq V$ uniformly at random.
- Select an ℓ-dimensional subspace L containing L' uniformly at random.
- Check if $F[L]|_{L'} = H[L']$.

Figure 1: 2-to-1 Test $\mathcal{T}_1$

We have the following two easy observations about the test.

- The test is 2-to-1 in the following sense. Fix any assignment for L', which is a linear function $\beta : L' \to \mathbb{F}_2$. Consider any linear function $\alpha : L \to \mathbb{F}_2$ such that $\alpha|_{L'} = \beta$. Since $L' \subseteq L$ and the co-dimension of L' in L is exactly 1, there are exactly two possible choices for α such that $\alpha|_{L'} = \beta$. Therefore, for every fixing of $H[L']$, there are exactly two settings of $F[L]$ such that the test accepts the pair $(F[L], H[L'])$.
- The test has *perfect completeness*, i. e., if F and H are indeed the restrictions of a global linear function $g : V \to \mathbb{F}_2$, then the test passes with probability 1.

The soundness of this test is analyzed in [21] with additional contributions from [14, 10]. They show that if the test passes with probability at least $\delta > 0$, then the tables must be non-trivially consistent with some global linear function $g : V \to \mathbb{F}_2$, where $\delta > 0$ is an arbitrarily small constant independent of ℓ and n.

In this paper, we focus on converting the test to a unique test, i. e., a 1-to-1 test, with some additional structure. One way to convert a 2-to-1 test to a unique test is by choosing a random

[2] Any linear function $f : L \to \mathbb{F}_2$ on an ℓ-dimensional space L can be specified by specifying $(f(b_1), f(b_2), \ldots, f(b_\ell)) \in \mathbb{F}_2^\ell$ where $b_1, b_2, \ldots, b_\ell \in L$ form a basis of L.

$i \in \{1,2\}$ for every pair (L,L') such that $L' \subseteq L$ and for every linear function β on L', and adding just one accepting pair (α_i,β) where $\{(\alpha_1,\beta),(\alpha_2,\beta)\}$ are the original accepting assignments from the 2-to-1 test $\mathcal{T}_1$. This does give a unique test and if F and H are restrictions of a global linear function to the subspaces, then the test passes with probability $1/2$. One drawback of this test is as follows. Consider a bipartite graph on $\mathrm{Gr}(V,\ell) \times \mathrm{Gr}(V,\ell-1)$ where two subspaces L,L' are adjacent iff $L' \subseteq L$. Note that the uniform distribution on the edges of this bipartite graph is the same as the test distribution $\mathcal{T}_1$. We say that an edge (L',L) is *satisfied* by the tables F and H, if $F[L]|_{L'} = H[L']$. For any global linear function $g : V \to \mathbb{F}_2$, if F and H are the restrictions of g, then we can only argue that half the edges are satisfied in the sense of the unique test. Hence, the similar guarantee of satisfying around half the edges stays in the final Unique Games instance created from the articles [20, 14, 15, 21] and thus falls short of proving Theorem 1.2.

Now we convert it into a Unique Test $\mathcal{T}_2$ (Figure 1.3) with a guarantee that for around half of the vertices on one side of the bipartite test graph, *all* the edges incident on them are satisfied if the tables F and H are the restrictions of some global linear function.

Given tables F and H, where $F : \mathrm{Gr}(V,\ell) \times (2^{[\ell]} \setminus \{\emptyset\}) \times \{0,1\} \to \mathbb{F}_2^{\ell-1}$ and $H : \mathrm{Gr}(V,\ell-1) \to \mathbb{F}_2^{\ell-1}$,

- Select an $(\ell-1)$-dimensional subspace L' uniformly at random.
- Select an ℓ-dimensional subspace L containing L', $x \in L \setminus L'$ and $b \in \{0,1\}$ uniformly at random.
- Check if $F[L,x,b]|_{L'} = H[L']$.

Here, $F[L,x,b]$ is thought of as a linear function $f : L \to \mathbb{F}_2$ such that $f(x) = b$, and $H[L']$ is thought of as a linear function $L' \to \mathbb{F}_2$, by choosing arbitrary bases of the linear spaces $L_x = \{y \in L \mid y \perp x\}$ and L', respectively.

Figure 2: Unique Test $\mathcal{T}_2$

Towards this goal, we modify the domain of F. We consider two tables, $F : \mathrm{Gr}(V,\ell) \times (2^{[\ell]} \setminus \{\emptyset\}) \times \{0,1\} \to \mathbb{F}_2^{\ell-1}$ and $H : \mathrm{Gr}(V,\ell-1) \to \mathbb{F}_2^{\ell-1}$. We fix an arbitrary one-to-one correspondence between the *non-zero* elements of the ℓ-dimensional subspace L and $2^{[\ell]} \setminus \{\emptyset\}$ for every L. Thus, we can now interpret F as defined on a tuple (L,x,b) where $x \in L \setminus \{0\}$ and $b \in \{0,1\}$. We consider the entries $F[L,x,b]$ and $H[L']$ as linear functions on the spaces L and L', respectively, as follows. As before, we select an arbitrary basis for every $(\ell-1)$-dimensional subspaces $\mathrm{Gr}(V,\ell-1)$. Now $F[L,x,b] \in \mathbb{F}_2^{\ell-1}$ is thought of as a linear function $L \to \mathbb{F}_2$ such that

1. at point x it evaluates to b, and
2. the evaluations of the linear function on the subspace $L_x = \{y \in L \mid y \perp x\}$, which is an $(\ell-1)$-dimensional subspace, is given by $F[L,x,b]$ in terms of the already chosen basis of

L_x).

As before, $H[L']$ is thought of as a linear function $L' \to \mathbb{F}_2$ in terms of the chosen basis for L'.

Consider the following bipartite graph $\big(\mathrm{Gr}(V,\ell) \times (2^{[\ell]} \setminus \{\emptyset\}) \times \{0,1\}, \mathrm{Gr}(V,\ell-1), E\big)$ where (L,x,b) is connected to L' iff $x \notin L'$ and $L' \subseteq L$. The test distribution from the Unique Test $\mathcal{T}_2$ is uniform on the edges of this graph.

We now put permutation constraints on the edges of the graph. For an edge $e \in E$ between (L,x,b) and L', we set the following permutation constraint. Extend the linear function given by $H[L']$ to a linear function $\tilde{H}_{L',x,b}[L]$ on $L = \mathtt{span}\{L',x\}$ by setting $\tilde{H}_{L',x,b}[L](x) = b$. The accepting labels for an edge e are $F[L,x,b]$ and $H[L']$ such that $\tilde{H}_{L',x,b}[L]$ and $F[L]$ are identical when thought of as linear functions on L. Note that the constraint is one-to-one.

Fix any global linear function $g : V \to \mathbb{F}_2$. From this, we define $H[L']$ as the restriction of g on L'. We define the table F partially by setting $F[L,x,g(x)]$ as the restriction of g on L. Thus for every $(L,x,g(x))$, it is clear that all the edges attached to it are *satisfied* by the tables H and F. The set $\{(L,x,g(x))\}$ also constitutes half of one side of the bipartite test graph. This particular structure on the set of satisfying edges from this bipartite graph goes into the final Unique Games instance that we construct. Therefore, in the YES case of the final reduction, we have that there is an assignment such that for at least half the vertices on one side, all the edges attached to them are satisfied. This establishes the completeness of our reduction from the main theorem.

Finally, the soundness of the Unique Test $\mathcal{T}_2$ follows directly from the soundness of the 2-to-1 Test $\mathcal{T}_1$ and we use this to prove the soundness of our reduction.

2 Preliminaries

We start by defining the Unique Games.

Definition 2.1 (Unique Games)**.** An instance $G = (A,B,E,[L],\{\pi_e\}_{e\in E})$ of the Unique Games constraint satisfaction problem consists of a bipartite graph (A,B,E), an alphabet $[L]$ and a permutation map $\pi_e : [L] \to [L]$ for every edge $e \in E$. Given a labeling $\sigma : A \cup B \to [L]$, an edge $e = (u,v)$ is said to be satisfied by σ if $\pi_e(\sigma(v)) = \sigma(u)$.

G is said to be *at most δ-satisfiable* if every labeling satisfies at most a δ fraction of the edges.

We will define the following two quantities related to the satisfiability of the Unique Games instance.

$$\mathsf{val}(G) := \max_{\sigma:A\cup B\to[L]} \{\text{fraction of edges in } G \text{ satisfied by } \sigma\}.$$

$$\mathsf{sval}(G) := \max_{\sigma:A\cup B\to[L]} \left\{ \frac{|A'|}{|A|} \,\middle|\, \forall e(u,v) \text{ s.t. } u \in A', e \text{ is satisfied by } \sigma \right\}.$$

The following is a conjecture by Khot [18] which has been used to prove many *tight* inapproximability results.

Conjecture 2.2 (Unique Games Conjecture[18])**.** *For every sufficiently small $\delta > 0$ there exists $L \in \mathbb{N}$ such that given an instance $\mathcal{G} = (A,B,E,\{\pi_e\}_{e\in E},[L])$ of Unique Games it is* NP*-hard to distinguish between the following two cases:*

- *YES case:* $\mathsf{val}(G) \geqslant 1-\delta$.
- *NO case:* $\mathsf{val}(G) \leqslant \delta$.

For a subset U of a vector space V, $\mathtt{span}(U)$ denotes its span. The notation $\mathtt{span}(U_1, U_2)$ stands for $\mathtt{span}(U_1 \cup U_2)$ and for $x \in V$, $\mathtt{span}(x, U)$ stands for $\mathtt{span}(\{x\}, U)$. Note that for subspaces L_1, L_2 we have $\mathtt{span}(L_1, L_2) = L_1 + L_2 := \{x_1 + x_2 \mid x_i \in L_i\}$ (the sumset). If $L_1 \cap L_2 = \{0\}$ then $L_1 + L_2$ is the *direct sum*, also denoted $L_1 \oplus L_2$.

For $0 < \ell < n$, let $\mathrm{Gr}(\mathbb{F}_2^n, \ell)$ be the set of all ℓ-dimensional subspaces of $\mathbb{F}_2^n$. Similarly, for a subspace L of $\mathbb{F}_2^n$ such that $\mathtt{dim}(L) > \ell$, let $\mathrm{Gr}(L, \ell)$ be the set of all ℓ-dimensional subspaces of $\mathbb{F}_2^n$ *contained* in L.

3 The Reduction

In this section, we go over the reduction in [15] from a gap 3LIN instance to a 2-to-1 Label Cover instance and then show how to reduce it to a Unique Games instance in Section 3.4. We retain most of the notations from [15].

3.1 Outer Game

The starting point of the reduction is the following problem:

Definition 3.1 (REG-3LIN). The instance (X, Eq) of REG-3LIN consists of a set X of n variables, $X = \{x_1, x_2, \ldots, x_n\}$, taking values in $\mathbb{F}_2$, and a collection Eq of m $\mathbb{F}_2$-linear constraints where each constraint in Eq is a linear constraint on 3 variables. The instance is regular in the following ways: every equation consists of 3 distinct variables, every variable x_i appears in exactly 5 constraints, and every two distinct constraints share at most one variable.

An instance (X, Eq) is said to be t-satisfiable if there exists an assignment to X which satisfies at least a t fraction of the constraints. We have the following theorem implied by the PCP theorem of [2, 3, 16].

Theorem 3.2. *There exists an absolute constant $s < 1$ such that for every constant $\varepsilon > 0$ it is* NP*-hard to distinguish between the cases when a given* REG-3LIN *instance is at least $(1-\varepsilon)$-satisfiable vs. at most s-satisfiable.*

We now define an outer 2-prover 1-round game, parameterized by $k, q \in \mathbb{Z}^+$ and $\beta \in (0,1)$, which will be the starting point of our reduction. The verifier selects k constraints $e_1, e_2, \ldots, e_k$ from the instance (X, Eq) uniformly at random with repetition. If e_i and e_j share a variable for some $i \neq j$ then accept. Otherwise, let $x_{i,1}, x_{i,2}, x_{i,3}$ be the variables in constraint e_i. Let $X_1 = \bigcup_{i=1}^{k} \{x_{i,1}, x_{i,2}, x_{i,3}\}$. The verifier then selects a subset X_2 of X_1 as follows: for each $i \in [k]$, with probability $(1-\beta)$ add $x_{i,1}, x_{i,2}, x_{i,3}$ to X_2 and with probability β, select a variable from $\{x_{i,1}, x_{i,2}, x_{i,3}\}$ uniformly at random and add it to X_2.

On top of this, the verifier selects q pairs of *advice* strings (s_j, s_j^*) where $s_j \in \{0,1\}^{X_2}$, and $s_j^* \in \{0,1\}^{X_1}$ for $1 \leq j \leq q$ as follows : For each $j \in [q]$, select $s_j \in \{0,1\}^{X_2}$ uniformly at random.

The string s_j can be thought of as assigning $\mathbb{F}_2$ values to each of the variables from X_2. The string $s_j^* \in \{0,1\}^{3k}$ is deterministically selected such that its projection on X_2 is the same as s_j and the rest of the coordinates are assigned 0.

The verifier sends $(X_1, s_1^*, s_2^*, \ldots, s_q^*)$ to Prover 1 and $(X_2, s_1, s_2, \ldots, s_q)$ to Prover 2. The verifier expects an assignment to variables in X_i from Prover i. Finally, the verifier accepts if and only if the assignment to X_1 given by Prover 1 satisfies all the equations $e_1, e_2, \ldots, e_k$ and the assignment X_2 given by Prover 2 is consistent with the answer of Prover 1.

Completeness: It is easy to see the completeness. If the instance (X, Eq) is $(1-\varepsilon)$-satisfiable then there is a provers' strategy which makes the verifier accept with probability at least $(1-k\varepsilon)$. The strategy is to use a fixed $(1-\varepsilon)$-satisfying assignment and answer according to it. In this case, with probability at least $(1-k\varepsilon)$, the verifier chooses k constraints which are all satisfied by the fixed assignment and hence the verifier will accept provers' answers.

Soundness: Consider the case when the instance (X, Eq) is at most s-satisfiable for $s < 1$ from Theorem 3.2. If the provers were given only X_1 and X_2 without the advice strings, then the Parallel Repetition Theorem of Raz [28] directly implies that for any provers' strategy, they can make the verifier accept with probability at most $2^{-\Omega(\beta k)}$. This follows because in expectation, there are βk constraints out of k where Prover 2 receives one variable from the constraint. It turns out that a few advice strings will not give provers any significant advantage.

To see this, for each of these βk constraints, with probability 2^{-q}, all the advice strings get assigned value 000 to the variables in the constraints and therefore does not *leak* any information, about which variable from the constraint was being sent to Prover 2, to Prover 1. Thus in expectation, there are $\frac{\beta k}{2^q}$ constraints vs. variable questions where Prover 1 knows nothing about which variable was being sent to Prover 2. One can then argue, by using Raz's Parallel Repetition Theorem, that any provers' strategy can make the verifier accept with probability at most $2^{-\Omega(\beta k/2^q)}$.

The soundness is formally proved in [20].

Theorem 3.3 (Section 3 in [20])**.** *If the* REG-3LIN *instance (X, Eq) is at most s-satisfiable ($s < 1$ from Theorem 3.2) then there is no strategy with which the provers can make the verifier accept with probability greater than $2^{-\Omega(\beta k/2^q)}$.*

Remark 3.4. The importance of the advice strings will come later in the proof of soundness. Specifically, the proof of Theorem 3.14 (from [15] which we use as a black-box) crucially uses the advice strings given to the provers.

To prove our main theorem, the reduction is carried out in three steps:

$$
\begin{array}{ll}
\text{Outer Game} & \\
\quad\downarrow [15] & \\
G_{\texttt{unfolded}}(A,B,E,\Pi,\Sigma_A,\Sigma_B) & \text{(unfolded 2-to-1 Game)} \\
\quad\downarrow [15] & \\
G_{\texttt{folded}}(\tilde{A},B,\tilde{E},\tilde{\Pi},\Sigma_A,\Sigma_B) & \text{(folded 2-to-1 Game)} \\
\quad\downarrow \text{(This work)} & \\
\mathsf{UG}_{\texttt{folded}}(\widehat{A},B,\widehat{E},\widehat{\Pi},\Sigma) & \text{(Unique Game)}
\end{array}
$$

The first two steps are explained in the next two subsections. These follow from [15]. The main contribution of our work is the last step which is given in Section 3.4.

3.2 Unfolded 2-to-1 Game

In this section we reduce a Reg-3Lin instance (X, Eq) to an instance of 2-to-1 Label Cover $G_{\texttt{unfolded}} = (A,B,E,\Pi,\Sigma_A,\Sigma_B)$.

For an equation $e \in Eq$, let $\mathtt{supp}(e) = \{i_1,i_2,i_3\}$ if e is a linear constraint on $x_{i_1}, x_{i_2}, x_{i_3}$. A set of k equations $(e_1,e_2,\dots,e_k)$ is *legitimate* if the supports of the equations are pairwise disjoint and for every two different equations e_i and e_j and for any $x \in e_i$ and $y \in e_j$, the pair $\{x,y\}$ does not appear in any equation in Eq. Define $\mathcal{U}$ to be the following set family.

$$
\mathcal{U} = \left\{ S \in \binom{[n]}{3k} \;\middle|\; S = \bigcup_{i=1}^{k} \mathtt{supp}(e_i) \text{ and } (e_1,e_2,\dots,e_k) \text{ is legitimate} \right\}.
$$

Note that by definition, there is a one-to-one correspondence between the set of legitimate k tuples of equations and $\mathcal{U}$. For $U \in \mathcal{U}$, let $X_U \subseteq \mathbb{F}_2^n$ be the linear subspace whose elements have support in U. For an equation $e \in Eq$ on $x_{i_1}, x_{i_2}, x_{i_3}$, let x_e be the vector in X_U where $x_{i_1} = x_{i_2} = x_{i_3} = 1$ and rest of the coordinates are 0. Denote by $b_e \in \mathbb{F}_2$ the RHS of the equation e. Let H_U be the span of $\{x_e \mid x_e \in X_U\}$. Finally, let $\mathcal{V}$ be the collection of all sets of variables up to size $3k$ (thought of as subsets of $[n]$). Similar to X_U, for $V \in \mathcal{V}$, let $X_V \subseteq \mathbb{F}_2^n$ be the linear subspace whose elements have support in V.

Vertices (A,B): Let $\ell \ll k$ which we will set later. The vertex set of the game $G_{\texttt{unfolded}}$ is defined as follows:

$$
A = \{(U,L) \mid U \in \mathcal{U}, L \in \mathrm{Gr}(X_U,\ell), L \cap H_U = \{0\}\}.
$$

$$
B = \{(V,L') \mid V \in \mathcal{V}, L' \in \mathrm{Gr}(X_V,\ell-1)\}.
$$

Edges E: The distribution on the edges in $G_{\texttt{unfolded}}$ is defined by the following process: Choose X_1 and X_2 as per the distribution given in the outer verifier conditioned on $U \in \mathcal{U}$, where $U = \bigcup_{x_i \in X_1}\{i\}$. Let $V = \bigcup_{x_i \in X_2}\{i\}$. Choose a random subspace $L' \in \mathrm{Gr}(X_V, \ell - 1)$ and a random $L \in \mathrm{Gr}(X_U, \ell)$ such that $L' \subseteq L$. Output $((U,L),(V,L')) \in (A,B)$.

Labels (Σ_A, Σ_B): The label set $\Sigma_A = \mathbb{F}_2^{\ell}$ and the label set $\Sigma_B = \mathbb{F}_2^{\ell-1}$. A label $\sigma \in \Sigma_A$ to (U,L) can be thought of as a linear function $\sigma : L \to \mathbb{F}_2$. Similarly the label $\sigma' \in \Sigma_B$ to a vertex (V,L') is though of as a linear function $\sigma' : L' \to \mathbb{F}_2$. This can be done by fixing arbitrary basis of the respective spaces.

3.3 Folded 2-to-1 Game

For every assignment to the 3LIN instance, there are many vertices in the graph $G_{\texttt{unfolded}}$ which get the same label in accordance with the strategy of labeling the vertices in $G_{\texttt{unfolded}}$ with respect to a fixed assignment to X. So we might as well enforce this constraint on the variables in $G_{\texttt{unfolded}}$. This is achieved by *folding*. In this section, we convert $G_{\texttt{unfolded}}$ to the following Game $G_{\texttt{folded}} = (\tilde{A}, B, \tilde{E}, \tilde{\Pi}, \Sigma_A, \Sigma_B)$.

Vertices $(\tilde{A}, B)$: Consider the following relation on the vertices in A:

$$(U_1, L_1) \sim_R (U_2, L_2) \text{ iff } L_1 + H_{U_1} + H_{U_2} = L_2 + H_{U_1} + H_{U_2}.$$

The following Lemma 3.5 says that $\sim_R$ is indeed an equivalence relation, i. e., $A = \mathcal{C}_1 \sqcup \mathcal{C}_2 \sqcup \ldots$, where each $\mathcal{C}_i$ is one of the equivalence classes. In other words, if $(U_0, L_0) \in \mathcal{C}_i$, then

$$\mathcal{C}_i = \{(U,L) \in A \mid L + H_U + H_{U_0} = L_0 + H_U + H_{U_0}\}.$$

The proof of the lemma crucially uses the facts that U corresponds to a *legitimate* set of equations and that the REG-3LIN instance is *regular*, namely, every equation consists of three distinct variables and every two distinct constraints share at most one variable.

Lemma 3.5 (Lemma 3.2 in [15]). *The relation $\sim_R$ defined above is an equivalence relation: for every i, there exists an ℓ-dimensional subspace $R_{\mathcal{C}_i}$ such that for all $(U,L) \in \mathcal{C}_i$,*

$$H_U + L = R_{\mathcal{C}_i} + H_U.$$

We define the vertex set $\tilde{A}$ as follows:

$$\tilde{A} = \{\mathcal{C}(U,L) \mid (U,L) \in A\},$$

where $\mathcal{C}(U,L)$ is the equivalence class of (U,L). In other words, in $\tilde{A}$, there is a vertex for every equivalence class.

Edges $\tilde{E}$: The distribution on the edges in $G_{\texttt{folded}}$ is defined by the following process: sample $((U,L),(V,L'))$ with respect to E and output $(\mathcal{C}(U,L),(V,L'))$.

Labels (Σ_A, Σ_B)**:** The label set $\Sigma_A = \mathbb{F}_2^\ell$, a label σ to $\mathcal{C} \in \tilde{A}$ can be thought of as a linear function $\sigma : R_\mathcal{C} \to \mathbb{F}_2$. As before, the label $\sigma' \in \Sigma_B$ to a vertex (V, L') is though of as a linear function $\sigma' : L' \to \mathbb{F}_2$.

In order to define the constraints on the edges, we need the following definitions:

Definition 3.6. For a space $H_U + L$ such that $L \cap H_U = \{0\}$ and a linear function $\sigma : L \to \mathbb{F}_2$, the extension of σ, respecting the side conditions, to the whole space $H_U + L$ is a linear function $\beta : H_U + L \to \mathbb{F}_2$ such that for all $x_e \in X_U$, $\beta(x_e) = b_e$ and $\beta|_L = \sigma$.

Note that there is a one-to-one mapping between linear functions on L and their extensions as all the equations in U are disjoint and hence $\{x_e \mid x_e \in X_U\}$ form a basis for the space H_U.

Definition 3.7. Consider a label σ to a vertex $\mathcal{C}$ which is a linear function on $R_\mathcal{C}$. The unfolding of it to the elements of the $\mathcal{C}$ is given as follows: For $(U, L) \in \mathcal{C}$, define a linear function $\tilde{\sigma}_U : H_U + L \to \mathbb{F}_2$ such that it is equal to the extension of σ to $H_U + R_\mathcal{C}$ respecting the side conditions.

The spaces $H_U + L$ and $H_U + R_\mathcal{C}$ are the same and hence the above definition makes sense. We are now ready to define the constraints.

Constraints $\tilde{\Pi}$**:** Consider linear functions $\sigma : R_\mathcal{C} \to \mathbb{F}_2$ and $\sigma' : L' \to \mathbb{F}_2$. A pair (σ, σ') satisfies the edge $(\mathcal{C}, (V, L')) \in \tilde{E}$, if for every $(U, L) \in \mathcal{C}$ such that $((U, L), (V, L')) \in E$, the unfolding $\tilde{\sigma}_U$ when restricted to the subspace L' is σ', i. e., $\tilde{\sigma}_U|_{L'} = \sigma'$.

We have the following completeness and soundness guarantee of the reduction from [15].

Lemma 3.8 (Completeness)**.** *(Lemma 4.1 in [15]) For every constant $\varepsilon > 0$, if the* Reg-3Lin *instance (X, Eq) is $(1-\varepsilon)$-satisfiable then there exists $\tilde{A}' \subseteq \tilde{A}$, $|\tilde{A}'| \geqslant (1 - k\varepsilon)|\tilde{A}|$, and a labeling to the 2-to-1 Label Cover instance $G_{\texttt{folded}}$ such that* all *the edges incident on $\tilde{A}'$ are satisfied.*

Lemma 3.9 (Soundness)**.** *(Lemma 4.2 in [15], and [21]) For all $\delta > 0$, there exist $q, k \geq 1$ and $\beta \in (0, 1)$, such that if the* Reg-3Lin *instance (X, Eq) is at most s-satisfiable (where s is from Theorem 3.2) then every labeling to $G_{\texttt{folded}}$ satisfies at most a δ fraction of the edges.*

3.4 Reduction to Unique Games

In this section, we convert the Label Cover instance $G_{\texttt{folded}}$ to a Unique Games instance with the stronger completeness guarantee that we are after. We will reduce an instance $G_{\texttt{folded}} = (\tilde{A}, B, \tilde{E}, \tilde{\Pi}, \Sigma_A, \Sigma_B)$ to an instance of Unique Game $\mathsf{UG}_{\texttt{folded}} = (\widehat{A}, B, \widehat{E}, \widehat{\Pi}, \Sigma)$.

Vertices $(\widehat{A}, B)$**:** We will split each vertex $\mathcal{C} \in \tilde{A}$ into many copies. Fix an ℓ-dimensional subspace $R_\mathcal{C}$ given by Lemma 3.5. For every $x \in R_\mathcal{C} \setminus \{0\}$ and $b \in \{0, 1\}$ we add a copy $\mathcal{C}_{x,b}$ to $\widehat{A}$.

$$\widehat{A} = \bigcup_{\mathcal{C} \in \tilde{A}} \{\mathcal{C}_{x,b} \mid x \in R_\mathcal{C} \setminus \{0\}, b \in \{0, 1\}\}.$$

Edges $\widehat{E}$: The distribution on the edge set $\widehat{E}$ is as follows: We first pick $((U,L),(V,L'))$ from the distribution E. Let $(U,L) \in \mathcal{C}$. We then select $y \in (H_U + L) \setminus (H_U + L')$ and $b \in \{0,1\}$ independently and uniformly at random. Note that $\mathtt{dim}(\mathtt{span}\{y, H_U\} \cap R_{\mathcal{C}}) = 1$ since $y \notin H_U$ and $y \in H_U + L = H_U + R_{\mathcal{C}}$. Let $x \in \mathtt{span}\{y, H_U\} \cap R_{\mathcal{C}}$ be the non-zero vector. Output $(\mathcal{C}_{x,b}, (V,L'))$.

Claim 3.10. *x is distributed uniformly in $R_{\mathcal{C}} \setminus (H_U + L')$ conditioned on (U,V,L,L').*

Proof. We first claim that $x \in R_{\mathcal{C}} \setminus (H_U + L')$ by showing $x \notin H_U + L'$. Suppose not, then we can write $x = h + x'$ where $h \in H_U$ and $x' \in L'$. We also know that $x \in \mathtt{span}\{y, H_U\} \cap R_{\mathcal{C}}$ and $R_{\mathcal{C}} \cap H_U = \{0\}$. Thus, x can be written as $x = \tilde{h} + y$ where $\tilde{h} \in H_U$. This implies that $h + x' = \tilde{h} + y$. In other words, $y = h + \tilde{h} + x' \in H_U + L'$, a contradiction.

Since each $y \in (H_U + L) \setminus (H_U + L')$ gives a unique non-zero $x \in \mathtt{span}\{y, H_U\} \cap R_{\mathcal{C}}$, we will show that the number of $y \in (H_U + L) \setminus (H_U + L')$ that give a fixed x is same for all $x \in R_{\mathcal{C}} \setminus (H_U + L')$ and this will prove the claim.

Fix any $\tilde{x} \in R_{\mathcal{C}} \setminus (H_U + L')$, we now claim that the set of all $y \in (H_U + L) \setminus (H_U + L')$ that give $\tilde{x}$ is $\mathtt{span}\{\tilde{x}, H_U\} \setminus H_U$. Clearly, for any $y \notin \mathtt{span}\{\tilde{x}, H_U\} \setminus H_U$, $\tilde{x} \notin \mathtt{span}\{y, H_U\}$ and also for every $y \in \mathtt{span}\{\tilde{x}, H_U\} \setminus H_U$, $\tilde{x} \in \mathtt{span}\{y, H_U\}$. Thus, it remains to show that $\mathtt{span}\{x, H_U\} \setminus H_U \subseteq (H_U + L) \setminus (H_U + L')$ for all $x \in R_{\mathcal{C}} \setminus (H_U + L')$.

To prove the inclusion, suppose for contradiction $(\mathtt{span}\{x, H_U\} \setminus H_U) \cap (H_U + L') \neq \emptyset$. This means $x + h = \tilde{h} + v'$ for some $h, \tilde{h} \in H_U$ and $v' \in L'$. This implies $x = h + \tilde{h} + v' \in H_U + L'$ contradicting $x \in R_{\mathcal{C}} \setminus (H_U + L')$. $\square$

Labels Σ: The label set is $\Sigma = \mathbb{F}_2^{\ell-1}$. A label σ to $\mathcal{C}_{x,b}$ can be thought of as a linear function $\sigma : R_{\mathcal{C}} \to \mathbb{F}_2$ such that $\sigma(x) = b$. This is done by fixing an arbitrary $\ell - 1$ basis elements from the subspace $\{y \in R_{\mathcal{C}} \mid y \perp x\}$.

It is easy to see that there is a one-to-one correspondence between the labels Σ and the linear functions σ on $R_{\mathcal{C}}$ such that $\sigma(x) = b$. Similar to the previous case of $G_{\mathtt{folded}}$, a label from $\Sigma(= \Sigma_B)$ to a vertex (V,L') in B is interpreted as a linear function $\sigma' : L' \to \mathbb{F}_2$.

We define an analogous unfolding of the labels, to the vertices in $\widehat{A}$, to the elements of the corresponding equivalence class. Since the label set here is different from $G_{\mathtt{folded}}$, for a label σ to $\mathcal{C}_{x,b}$ (thought of as a linear function on $R_{\mathcal{C}}$ respecting $\sigma(x) = b$) we use the notation $\widehat{\sigma}_U$ to denote its unfolding to $(U,L) \in \mathcal{C}_{x,b}$.

1**-to-**1 **Constraints $\widehat{\Pi}$:** Finally the constraint $\pi_e : \Sigma \to \Sigma$ between the endpoints of an edge $e = (\mathcal{C}_{x,b}, (V,L'))$ is given as follows: Consider linear functions $\sigma : R_{\mathcal{C}} \to \mathbb{F}_2$ respecting $\sigma(x) = b$ and $\sigma' : L' \to \mathbb{F}_2$. A pair $(\sigma, \sigma') \in \pi_e$ if for every $(U,L) \in \mathcal{C}$ such that $((U,L),(V,L')) \in E$ and $\mathtt{span}\{x, H_U\} \cap L' = \{0\}$, the unfolding $\tilde{\sigma}_U$ satisfies $\tilde{\sigma}_U|_{L'} = \sigma'$.

To see that every σ' has a unique preimage in π_e, for any linear function $\sigma' : L' \to \mathbb{F}_2$, there is a unique linear function $\sigma : R_{\mathcal{C}} \to \mathbb{F}_2$ such that $\sigma(x) = b$ satisfying the above conditions. This is because of the following claim.

Claim 3.11. *Any basis for L' along with x and $\{x_e \mid x_e \in X_U\}$ forms a basis for $H_U + R_{\mathcal{C}}$ for every $(U,L) \in \mathcal{C}$.*

Proof. Let us unwrap the conditions for putting an edge between (V,L') and $\mathcal{C}_{x,b}$. One necessary condition is that $(\mathcal{C},(V,L'))$ should be an edge in $\tilde{E}$. By the definition of $\tilde{E}$, there exists $(U,L) \in \mathcal{C}$ such that $L' \subseteq L$. Recall, x is such that there exists a $y \in (H_U+L)\setminus(H_U+L')$ such that $\mathtt{dim}(\mathtt{span}\{y,H_U\} \cap R_{\mathcal{C}}) = 1$ and $x \in \mathtt{span}(y,H_U) \cap R_{\mathcal{C}}$. Therefore $x \in (H_U+L)\setminus(H_U+L')$ and hence $\mathtt{dim}(\mathtt{span}\{x,H_U+L'\}) = k+\ell$ (as $H_U \cap L' = \{0\}$). This implies that any basis of L', the basis $\{x_e \mid x_e \in X_U\}$ of H_U, and x span H_U+L. Since by Lemma 3.5 the subspace H_U+L is same as the subspace $H_U+R_{\mathcal{C}}$, the claim follows. □

We now show the completeness and the soundness of the overall reduction to the Unique Games instance.

3.4.1 Completeness

Lemma 3.12 (Completeness). *For every constant $\varepsilon > 0$, if there exists $\tilde{A}' \subseteq \tilde{A}$, $|\tilde{A}'| \geqslant (1-k\varepsilon)|\tilde{A}|$, and a labeling to the* 2-to-1 *Label Cover instance $G_{\mathtt{folded}}$ such that* all *the edges incident on $\tilde{A}'$ are satisfied, then there exists $\widehat{A}' \subseteq \widehat{A}$, $|\widehat{A}'| \geqslant (\frac{1-k\varepsilon}{2})|\widehat{A}|$, and a labeling to Unique Games instance $\mathsf{UG}_{\mathtt{folded}}$ such that* all *the edges incident on $\widehat{A}'$ are satisfied.*

Proof. Fix a labeling $(\tilde{\mathscr{A}},\tilde{\mathscr{B}})$ to $G_{\mathtt{folded}}$ where $\tilde{\mathscr{A}} : \tilde{A} \to \Sigma_A$ and $\tilde{\mathscr{B}} : B \to \Sigma_B$ that satisfies all the edges incident on a $(1-k\varepsilon)$ fraction of the vertices in $\tilde{A}$. We will construct a labeling $(\widehat{\mathscr{A}},\widehat{\mathscr{B}})$ to the instance $\mathsf{UG}_{\mathtt{folded}}$, where $\widehat{\mathscr{A}} : \widehat{A} \to \Sigma$ and $\widehat{\mathscr{B}} : B \to \Sigma$ that will satisfy all the edges adjacent to at least a $\frac{(1-k\varepsilon)}{2}$ fraction of vertices $\widehat{A}$ in $\mathsf{UG}_{\mathtt{folded}}$.

We will set $\widehat{\mathscr{B}} = \tilde{\mathscr{B}}$. Now to assign a label to $\mathcal{C}_{x,b} \in \widehat{A}$, we look at the label $\sigma := \tilde{\mathscr{A}}(\mathcal{C}) \in \mathbb{F}_2^{\ell}$ as a linear function $\sigma : R_{\mathcal{C}} \to \mathbb{F}_2$. If $\sigma(x) = b$, we set $\widehat{\mathscr{A}}(\mathcal{C}_{x,b})$ to be the same linear function $\sigma : R_{\mathcal{C}} \to \mathbb{F}_2$ respecting $\sigma(x) = b$. Otherwise, we set $\widehat{\mathscr{A}}(\mathcal{C}_{x,b}) = \perp$. It is obvious that exactly half the vertices in $\widehat{A}$ got assigned a label in Σ.

Claim 3.13. *If the label $\tilde{\mathscr{A}}(\mathcal{C})$ to $\mathcal{C}$ satisfies all the edges incident on it, then for all $x \in R_{\mathcal{C}} \setminus \{0\}$, there exists a unique $b \in \{0,1\}$ such that the label $\widehat{\mathscr{A}}(\mathcal{C}_{x,b})$ satisfies all the edges incident on $\mathcal{C}_{x,b}$, unless $\widehat{\mathscr{A}}(\mathcal{C}_{x,b}) = \perp$.*

Proof. For convenience let $\sigma = \tilde{\mathscr{A}}(\mathcal{C})$. If we let $\Gamma(\mathcal{C}) \subseteq B$ to be the neighbors of $\mathcal{C}$ in $G_{\mathtt{folded}}$, then the set of neighbors of $\mathcal{C}_{x,b}$ is a subset of $\Gamma(\mathcal{C})$. Furthermore if (V,L') is connected to $\mathcal{C}_{x,b}$ in $\mathsf{UG}_{\mathtt{folded}}$ then $x \notin L'$ and $x \in R_{\mathcal{C}}$. The condition that the edge $(\mathcal{C},(V,L'))$ is satisfied by $\tilde{\mathscr{A}}$ means that for all $(U,L) \in \mathcal{C}$ such that $L' \subseteq L$, the unfolding of σ satisfies $\tilde{\sigma}_U|_{L'} = \tilde{B}((V,L'))$. Since the unfolding of the label $\widehat{\mathscr{A}}(\mathcal{C}_{x,b})$ to $\mathcal{C}_{x,b}$ gives the same linear function $\tilde{\sigma}$, it follows that $\tilde{\sigma}_U|_{L'} = \widehat{B}((V,L'))$ for every $(U,L) \in \mathcal{C}$ and every $(V,L') \in \Gamma(\mathcal{C})$ such that $L' \subseteq L$. Therefore $\widehat{\mathscr{A}}$ satisfies all the edges incident on $\mathcal{C}_{x,b}$ if $\widehat{\mathscr{A}}(\mathcal{C}_{x,b}) \neq \perp$. □

Let $\tilde{A}' \subseteq \tilde{A}$ be the set of vertices such that all the edges incident on them are satisfied by the labeling $(\tilde{\mathscr{A}},\tilde{\mathscr{B}})$. By assumption $|\tilde{A}'| \geqslant (1-k\varepsilon)|\tilde{A}|$. Consider the subset $\widehat{A}' \subseteq \widehat{A}$

$$\widehat{A}' = \{\mathcal{C}_{x,b} \mid \widehat{\mathscr{A}}(\mathcal{C}_{x,b}) \neq \perp, \mathcal{C} \in \tilde{A}'\}.$$

Now, $|\widehat{A}'| \geqslant \frac{1-k\varepsilon}{2}|\widehat{A}|$ and from the above claim, all the edges incident on $\widehat{A}'$ are satisfied by the labeling $(\widehat{\mathscr{A}},\widehat{\mathscr{B}})$. □

3.4.2 Soundness

Let $F_U : \{L + H_U \mid L \in \mathrm{Gr}(X_U,\ell), L\cap H_U = \{0\}\} \to \mathbb{F}_2^{\ell}$. $F_U[L+H_U]$ can be thought of as a linear function $L+H_U \to \mathbb{F}_2$ respecting the side conditions. This is again by fixing an arbitrary basis of L. Define $\texttt{agreement}(F_U)$ as the probability of the following event:

- Select an $(\ell-1)$-dimensional subspace $L' \in X_U$ such that $L'\cap H_U = \{0\}$ uniformly at random.
  - Select ℓ-dimensional subspaces L_1 and L_2 containing L' such that $L_1 \cap H_U = L_2 \cap H_U = \{0\}$ uniformly at random.
  - Check if $F_U[L_1+H_U]|_{L'} = F_U[L_1+H_U]|_{L'}$.

The main technical theorem which was conjectured in [15] and proved in [21] is that if $\texttt{agreement}(F_U)$ is a positive constant, then there is a global linear function $g : X_U \to \mathbb{F}_2$ respecting the side conditions and a special (not too small) subset S of $\{L+H_U \mid L \in \mathrm{Gr}(X_U,\ell), L\cap H_U = \{0\}\}$ such that for a constant fraction of elements in S, F_U agrees with g. We will not need the details of this theorem. Instead, we state the main soundness lemma from [15] which crucially used the aforementioned structural theorem and also the advice strings as mentioned in Remark 3.4.

Theorem 3.14 (Implied by Lemma 4.1 in [15])**.** *For every constant $\delta > 0$, there exist large enough $\ell \ll k$, $q \in \mathbb{Z}^+$ and $\beta \in (0,1)$ such that if there is an unfolded assignment $\mathscr{A} : A \to \Sigma_A$ to $G_{\texttt{unfolded}}$ such that for at least δ fraction of U, $\texttt{agreement}(F_U) \geqslant \delta$, $F_U[L+H_U] = \mathscr{A}(U,L)$ for every $L \in \mathrm{Gr}(X_U,\ell)$, then there exists a provers' strategy which makes the outer verifier accept with probability at least p_δ, where p_δ is independent of k.*

Armed with this theorem, we are ready to prove the soundness of the reduction to the Unique Games instance $\mathsf{UG}_{\texttt{folded}}$.

Lemma 3.15 (Soundness)**.** *Let $\delta > 0$ and fix $q \in \mathbb{Z}^+$, $\beta \in (0,1)$ and $\ell \ll k$ as in Theorem 3.14. If $\mathsf{UG}_{\texttt{folded}}$ is δ-satisfiable then there exists a provers' strategy which makes the outer verifier accept with probability at least $p_{\frac{\delta^4}{2^{16}}}$.*

Proof. Fix any δ-satisfying assignment $(\widehat{\mathscr{A}},\widehat{\mathscr{B}})$, $\widehat{\mathscr{A}} : \widehat{A} \to \Sigma$, $\widehat{\mathscr{B}} : \widehat{B} \to \Sigma$ to the Unique Games instance $\mathsf{UG}_{\texttt{folded}}$. We first get a randomized labeling $(\tilde{\mathscr{A}},\tilde{\mathscr{B}})$ to $G_{\texttt{folded}}$ where $\tilde{\mathscr{A}} : \tilde{A} \to \Sigma_A$ and $\tilde{\mathscr{B}} : B \to \Sigma_B$ as follows: We will keep $\tilde{\mathscr{B}} = \widehat{\mathscr{B}}$. For every $\mathcal{C} \in \tilde{A}$, we pick a random $x \in R_{\mathcal{C}}$ and $b \in \{0,1\}$ and set $\tilde{\mathscr{A}}(\mathcal{C}) = \widehat{\mathscr{A}}(\mathcal{C}_{x,b})$. We now unfold the assignment $\tilde{\mathscr{A}}$ to $\mathscr{A}$. Define $F_U[L+H_U] = \mathscr{A}(U,L)$ for every $L \in \mathrm{Gr}(X_U,\ell)$. Note that F_U is a random assignment.

Let $p(U)$ denote the probability that an edge in $\mathsf{UG}_{\texttt{folded}}$ is satisfied conditioned on U. Consider U such that $p(U) \geqslant \frac{\delta}{2}$. By an averaging argument, there are at least a $\delta/2$ fraction of U such that $p(U) \geqslant \frac{\delta}{2}$.

Claim 3.16. $\mathbf{E}_{F_U}[\texttt{agreement}(F_U)] \geqslant \frac{p(U)^4}{2^{11}} - o_k(1)$.

Proof. Fix a particular U. Define a randomized assignment $F'_U[L']$ as follows: Select a random $V \subseteq U$ conditioned on the event that $L' \subseteq X_V$. Set $F'_U[L'] = \widehat{\mathscr{B}}(V, L')$. Consider the following two distributions:
Distribution $\mathcal{D}_U$:

- Select $\mathbf{V}$ uniformly at random from $\{V \mid ((U,L),(V,L')) \in E \text{ for some } L, L'\}$
- Select $\mathbf{L}'$ uniformly at random from $\mathrm{Gr}(X_V, \ell-1)$
- Select $\mathbf{L}$ uniformly at random from $\{L \mid L \in \mathrm{Gr}(X_U, \ell) \text{ and } L' \subseteq L\}$
- Let $\mathcal{C}$ be the equivalence class such that $(U,L) \in \mathcal{C}$, select $\mathbf{x} \sim R_{\mathcal{C}}$ as in the edge distribution $\widehat{E}$.
- Select $\mathbf{b} \in \{0,1\}$ uniformly at random.

Distribution $\mathcal{D}'_U$:

- Select $\mathbf{L}'$ uniformly at random from $\mathrm{Gr}(X_U, \ell-1)$
- Select $\mathbf{V}$ uniformly at random from $\{V \mid ((U,L),(V,L')) \in E, L' \in \mathrm{Gr}(X_V, \ell-1) \text{ and } L \in \mathrm{Gr}(X_U, \ell)\}$
- Select $\mathbf{L}$ uniformly at random from $\{L \mid L \in \mathrm{Gr}(X_U, \ell) \text{ and } L' \subseteq L\}$
- Let $\mathcal{C}$ be the equivalence class such that $(U,L) \in \mathcal{C}$, select $\mathbf{x} \sim R_{\mathcal{C}}$ as in the edge distribution $\widehat{E}$.
- Select $\mathbf{b} \in \{0,1\}$ uniformly at random.

We have the following lemma from [15].

Lemma 3.17 (Lemma 3.6 in [15])**.** *Consider the two marginal distributions on the pair (V, L'), one with respect to $\mathcal{D}_U$ and another with respect to $\mathcal{D}'_U$. If $2^{\ell}\beta \le \frac{1}{8}$, then the statistical distance between the two distributions is at most $\beta\sqrt{k} \cdot 2^{\ell+4}$.*

In the distribution $\mathcal{D}_U$, there is always a constraint between $\mathcal{C}_{x,b}$ and (V, L') in $\mathsf{UG}_{\texttt{folded}}$. Moreover, the distribution of $(\mathcal{C}_{x,b}, (V, L'))$ is same as the edge distribution $\widehat{E}$. Therefore

$$p(U) = \Pr_{\mathcal{D}_U}\left[(\widehat{\mathscr{A}}(\mathcal{C}_{x,b}), \widehat{\mathscr{B}}(V,L')) \text{ satisfy the edge } (\mathcal{C}_{x,b}, (V,L'))\right].$$

Rewriting the above equality,

$$p(U) = \Pr_{\mathcal{D}_U}\left[\widehat{\sigma}_U|_{L'} = \widehat{\mathscr{B}}(V,L') \mid \sigma = \widehat{\mathscr{A}}(\mathcal{C}_{x,b})\right].$$

Using Claim 3.10, the distribution of $F_U[L + H_U]$, conditioned on $x \in R_{\mathcal{C}} \setminus (H_U + L')$, is same as the distribution $\widehat{\mathscr{A}}(\mathcal{C}_{x,b})$ (with appropriate unfolding of it) chosen with respect to $\mathcal{D}_U$. As

$|R_{\mathcal{C}} \setminus (H_U + L')| = |R_{\mathcal{C}}|/2$, for a random $x \in R_{\mathcal{C}}$, the event $x \in R_{\mathcal{C}} \setminus (H_U + L')$ happens with probability $\frac{1}{2}$. Since we pick an uniformly random $x \in R_{\mathcal{C}}$ while defining $\tilde{\mathscr{A}}(\mathcal{C})$, which in turn defines $F_U[L+H_U]$, we have

$$\frac{p(U)}{2} \leqslant \mathop{\mathbf{E}}_{F_U} \Pr_{\mathcal{D}_U}\left[F_U[L+H_U]|_{L'} = \widehat{\mathscr{B}}(V,L')\right],$$

Now,

$$\Pr_{\mathcal{D}_U}\left[F_U[L+H_U]|_{L'} = \widehat{\mathscr{B}}(V,L')\right] \approx \Pr_{\mathcal{D}'_U}\left[F_U[L+H_U]|_{L'} = \widehat{\mathscr{B}}(V,L')\right].$$

follows from the closeness of distributions $\mathcal{D}_U$ and $\mathcal{D}'_U$ on (V,L') given by Lemma 3.17 by setting $\beta \ll \frac{1}{\sqrt{k}}$ (this setting of β is consistent with the setting of β in Theorem 3.14). Conditioned on L', the distribution of (V,L') in $\mathcal{D}'_U$ is same as the distribution we used to assign $F'_U[L']$ and therefore we get

$$\frac{p(U)}{2} - o_k(1) \leqslant \mathop{\mathbf{E}}_{F_U} \Pr_{L' \subseteq L}\left[F_U[L+H_U]|_{L'} = F'_U[L']\right].$$

Let E_1 be the event that $\frac{p(U)}{4} - o_k(1) \leqslant \Pr_{L' \subseteq L}[F_U[L+H_U]|_{L'} = F'_U[L']]$. By an averaging argument, $\Pr[E_1] \geqslant \frac{P(U)}{4} - o_k(1)$. We now fix an F_U for which E_1 occurs. By another averaging argument, there is at least a $p(U)/8$ fraction of $L' \in \mathrm{Gr}(X_U, \ell-1)$ such that $\Pr_{L \supseteq L'}[F_U[L+H_U]|_{L'} = F'_U[L']] \geqslant \frac{p(U)}{8} - o_k(1)$. For each of such L' we have,

$$\begin{aligned}\Pr_{L_1,L_2 \supseteq L'}[F_U[L_1+H_U] = F_U[L_2+H_U]] &= \Pr_{L_1,L_2 \supseteq L'}\left[F_U[L_1+H_U]|_{L'} = F_U[L_2+H_U]|_{L'} = F'_U[L']\right] \\ &\geqslant \frac{p(U)^2}{2^6} - o_k(1).\end{aligned}$$

Thus overall, we get

$$\Pr_{L_1,L_2 \supseteq L'}[F_U[L_1+H_U] = F_U[L_2+H_U] \mid E_1] \geqslant \frac{p(U)^3}{2^9} - o_k(1).$$

Hence,

$$\mathop{\mathbf{E}}_{F_U}[\mathtt{agreement}(F_U)] \geqslant \Pr[E_1] \cdot \Pr_{L_1,L_2 \supseteq L'}[F_U[L_1+H_U] = F_U[L_2+H_U] \mid E_1] \geqslant \frac{p(U)^4}{2^{11}} - o_k(1).$$

□

There is at least a $\delta/2$ fraction of U such that $p(U) \geqslant \delta/2$. This means for at least a $\delta/2$ fraction of U, $\mathbf{E}[\mathtt{agreement}(F_U)] \geqslant \frac{\delta^4}{2^{15}} - o_k(1)$ using the previous claim. Thus, again by an averaging argument, there exists a fixed $\{F_U \mid U \in \mathcal{U}\}$, coming from unfolding of some assignment $\tilde{\mathscr{A}}$, such that for at least a $\delta^4 2^{-16}$ fraction of U, we have $\mathtt{agreement}(F_U) \geqslant \delta^4 2^{-16}$. The lemma now follows from Theorem 3.14. □

We now prove the main theorem.

Proof of Theorem 1.2 Fix $\delta > 0$. We let q, β and $\ell \ll k$ be as given in the setting of Theorem 3.14. Firstly, if we look at the marginal of the edge distribution on $\widehat{A}$ then it is uniform and hence the instance is left-regular.[3] Now, starting with an instance (X, Eq), we have the following two guarantees of the reduction.

1. If the instance (X, Eq) is $(1 - \frac{2\delta}{k})$-satisfiable then by Lemma 3.8 and Lemma 3.12, the Unique Games instance $\mathsf{UG}_{\texttt{folded}}$ has a property that for at least a $(1/2) - \delta$ fraction of the vertices in $\widehat{A}$, all the edges incident on them are satisfied.

2. Consider the other case in which the instance (X, Eq) is at most s-satisfiable where $s < 1$. If the Unique Games instance $\mathsf{UG}_{\texttt{folded}}$ has a δ-satisfying assignment, then by Lemma 3.15 there is a provers' strategy which can make the outer verifier accept with probability at least $p_{\frac{\delta^4}{2^{16}}} \gg 2^{-\Omega(\beta k/2^q)}$ for large enough k. This contradicts Theorem 3.3 and hence in this case, $\mathsf{UG}_{\texttt{folded}}$ has no assignment that satisfies a δ fraction of the edges.

Since by Theorem 3.2, distinguishing between a given instance (X, Eq) being at least $(1 - \frac{2\delta}{k})$-satisfiable or at most s-satisfiable is NP-hard, this proves our main theorem. $\square$

4 Independent set in degree-d graphs

We consider a weighted graph $H = (V, E)$ where the sum of the weights of all the vertices is 1 and also sum of the weights of all the edges is also 1. For $S \subseteq V$, we will denote the total weight of the vertices in S by $w(S)$.

Definition 4.1. A graph H is (δ, ε)-dense if for every $S \subseteq V(H)$ with $w(S) \geqslant \delta$, the total weight of the edges inside S is at least ε.

For $\rho \in [-1, 1]$ and $\beta \in [0, 1]$, the quantity $\Gamma_\rho(\beta)$ is defined as:

$$\Gamma_\rho(\beta) := \Pr[X \leq \phi^{-1}(\beta) \wedge Y \leq \phi^{-1}(\beta)],$$

where X and Y are jointly distributed normal Gaussian random variables with co-variance ρ and ϕ is the cumulative density function of a normal Gaussian random variable.

We will prove the following theorem.

Theorem 4.2. *Fix $\varepsilon > 0, p \in \left(0, \frac{1}{2}\right]$, then for all sufficiently small $\delta > 0$, there exists a polynomial-time reduction from an instance of left-regular Unique Games, $G(A, B, E, [L], \{\pi_e\}_{e \in E})$, to a graph H such that*

1. *If $\mathsf{sval}(G) \geqslant c$, then there is an independent set of weight $c \cdot p$ in H.*

2. *If $\mathsf{val}(G) \leqslant \delta$, then H is $(\beta, \Gamma_\rho(\beta) - \varepsilon)$ dense for every $\beta \in [0, 1]$ and $\rho = -\frac{p}{p-1}$.*

The reduction is exactly the same as the one in [4]. We will only show the completeness (1) from Theorem 4.2 here. The soundness of this theorem is proved in [4]. This theorem will imply Theorem 1.3 using a randomized sparsification technique of [4] to convert the weighted graph into a bounded-degree unweighted graph.

[3] The edges have weights, but it can be made an unweighted left-regular instance by adding multiple edges proportional to its weight with the same constraint.

Let $G(A,B,E,[L],\{\pi_e\}_{e\in E})$ be an instance of Unique Games. The distribution of edges in H is as follows:

- Select $u \in B$ uniformly at random.
- Select two neighbors v_1 and v_2 of u independently and uniformly at random. Let π_1 and π_2 be the constraints between (u,v_1) and (u,v_2) respectively.
- Select $x,y \in \{0,1\}^L$, such that for each $i \in [L]$, the pair (x_i,y_i) is sampled independently from the distribution $\mathcal{D}$.
- Output the edge $(v_1, x\circ\pi_1), (v_2, y\circ\pi_2)$.

Figure 3: Reduction from UG to Independent Set from [4].

4.1 The AKS reduction

Consider the distribution $\mathcal{D}$ on $(a,b) \in \{0,1\}^2$ such that $\Pr[a=b=1]=0$ and each bit is p-biased, i. e., $\Pr[b=1]=\Pr[b=1]=p$. For a string $x \in \{0,1\}^L$ and a permutation $\pi : [L] \to [L]$, define $x\circ\pi \in \{0,1\}^L$ as $(x\circ\pi)_i = x_{\pi(i)}$ for all $i \in [L]$.

Let $G(A,B,E,[L],\{\pi_e\}_{e\in E})$ be an instance of Unique Games which is regular on the A side. We convert it into a weighted graph H. The vertex set of H will be $A \times \{0,1\}^L$. Weight of a vertex (v,x) where $v \in A$ and $x \in \{0,1\}^L$ is $\frac{\mu_p(x)}{|A|}$, where $\mu_p(x) := p^{|x|}(1-p)^{L-|x|}$. The edge distribution is given in Figure 3.

Lemma 4.3 (Completeness). *If $\mathsf{sval}(G) \geqslant c$, then there is an independent set in H of weight $c \cdot p$.*

Proof. Fix an assignment $\ell : A \cup B \to \Sigma$ that gives $\mathsf{sval}(G) \geqslant c$. Let $A' \subseteq A$ be the set of vertices such all the edges incident on A' are satisfied by ℓ. We know that $|A'| \geqslant c \cdot |A|$. Consider the following subset of vertices in H.

$$I = \{(v,x) \mid v \in A', x_{\ell(v)} = 1\}.$$

Firstly, the total weight of set I is $c \cdot p$. We show that I is in fact an independent set in H. Suppose for contradiction, there exists an edge $(v_1,x),(v_2,y)$ in H with both endpoints are in I. Let u be the common neighbor of v_1,v_2 (one such u must exist). If we let π_1 and π_2 be the permutation constraints between (u,v_1) and (u,v_2), then the conditions for being an edge implies that $(x_{\pi_1(\ell(u))}, y_{\pi_2(\ell(u))})$ should have a support in $\mathcal{D}$. Since all the edges incident on A' are satisfied, $\pi_i(\ell(u)) = \ell(v_i)$ for $i \in \{1,2\}$. Therefore, $(x_{\ell(v_1)}, y_{\ell(v_2)})$ is also supported in $\mathcal{D}$ and hence both cannot be 1 which implies that both cannot belong to I. $\square$

Lemma 4.4 (Soundness [4]). *For every constant $\varepsilon > 0$, if H is not $(\beta, \Gamma_\rho(\beta)-\varepsilon)$-dense for some $\beta \in [0,1]$ and $\rho = -\frac{p}{p-1}$, then G is δ-satisfiable for $\delta := \delta(\varepsilon,p) > 0$.*

Lemma 4.3 and Lemma 4.4 prove Theorem 4.2.

5 Maximum Acyclic Subgraph

In this section we state the reduction from [17] and analyze the completeness of the reduction. Given a directed graph $H=(V,E)$, we will denote by $\texttt{Val}(H)$ the fraction of edges in the maximum sized acyclic subgraph of H. We need the following definition.

Definition 5.1. A t-ordering of a directed graph $H=(V,E)$ consists of a map $O:V\to[t]$. The value of a t-ordering O is given by

$$\texttt{Val}_t(O)=\Pr_{(a,b)\in E}[O(a)<O(b)]+\frac{1}{2}\cdot\Pr_{(a,b)\in E}[O(a)=O(b)].$$

Define $\texttt{Val}_t(H)$ as:

$$\texttt{Val}_t(H)=\max_{O}\texttt{Val}_t(O).$$

The following lemma [17] will be crucial in the reduction from Unique Games to the Maximum Acyclic Subgraph problem.

Lemma 5.2 ([17]). *Given $\eta>0$ and a positive integer t, for every sufficiently large m, there exists a weighted directed acyclic graph $H(V,E)$ on m vertices along with a distribution $\mathcal{D}$ on the orderings $\{O:V\to[m]\}$ such that:*

1. *For every $u\in V$ and $i\in[m]$, $\Pr_{O\sim\mathcal{D}}[O(u)=i]=\frac{1}{m}$.*
2. *For every directed edge $(a\to b)$, $\Pr_{O\sim\mathcal{D}}[O(a)<O(b)]\geqslant 1-\eta$.*
3. *$\texttt{Val}_t(H)\leqslant\frac{1}{2}+\eta$.*

The reduction is given in Figure 4. For a string $x\in[q]^L$ and a permutation $\pi:[L]\to[L]$, define $x\circ\pi\in[q]^L$ by $(x\circ\pi)_i=x_{\pi(i)}$ for all $i\in[L]$.

Lemma 5.3 (Completeness). *For small enough $\varepsilon,\eta>0$, if the Unique Games instance G has $\mathsf{sval}(G)\geqslant c$ then $\texttt{Val}(\mathcal{G})\geqslant c\cdot(1-2\varepsilon)(1-\eta)+(1-c)\cdot\left(\frac{1}{2}-\frac{1}{2m}\right)$*

Proof. Fix an assignment $\ell:A\cup B\to\Sigma$ that gives $\mathsf{sval}(G)\geqslant c$. Let $A'\subseteq A$ be the set of vertices such that its edges are satisfied by ℓ, we know that $|A'|\geqslant c\cdot|A|$. Consider the following m ordering $\mathcal{O}:B\times[m]^L\to[m]$ of the vertices of $\mathcal{G}$: $\mathcal{O}(v,x)=x_{\ell(v)}$. We will show that $\texttt{Val}_m(\mathcal{O})\geqslant c(1-2\varepsilon)(1-\eta)+(1-c)\cdot\left(\frac{1}{2}-\frac{1}{2m}\right)$. This will prove the lemma.

$$\begin{aligned}\texttt{Val}(\mathcal{G})\geqslant\texttt{Val}_m(\mathcal{O})&\geqslant\Pr[\mathcal{O}((v_1,\tilde{x}\circ\pi_1)<\mathcal{O}(v_2,\tilde{y}\circ\pi_2)]\\&=\Pr[\tilde{x}_{\pi_1(\ell(v_1))}<\tilde{y}_{\pi_2(\ell(v_2))}]\\&\geqslant c\cdot\Pr[\tilde{x}_{\pi_1(\ell(v_1))}<\tilde{y}_{\pi_2(\ell(v_2))}\mid u\in A']\\&\quad+(1-c)\cdot\Pr[\tilde{x}_{\pi_1(\ell(v_1))}<\tilde{y}_{\pi_2(\ell(v_2))}\mid u\notin A'].\end{aligned}\tag{5.1}$$

Let $G(A, B, E, [L], \{\pi_e\}_{e \in E})$ be an instance of Unique Games. Fix a graph $H([m], E_H)$ from Lemma 5.2 with parameters $\eta > 0$ and $t \in \mathbb{Z}^+$, along with the distribution $\mathcal{D}$. Construct a weighted directed graph $\mathcal{G}$ on $B \times [m]^L$ with the following distribution on the edges:

- Select $u \in A$ uniformly at random.
- Select two neighbors v_1 and v_2 of u uniformly at random. Let π_1 and π_2 be the constraints between (u, v_1) and (u, v_2) respectively.
- Pick an edge $e = (a, b) \in E_H$ at random from the graph H.
- Select $x, y \in [m]^L$, such that for each $i \in [L]$, the pair (x_i, y_i) is sampled independently as follows:
  - sample $O \sim \mathcal{D}$, set $x_i = O(a)$ and $y_i = O(b)$.
- Perturb x and y as follows: for each $i \in [L]$, with probability $(1 - \varepsilon)$, set $\tilde{x}_i = x_i$, with probability ε set $\tilde{x}_i$ to be uniformly at random from $[m]$. Do the same thing for y independently to get $\tilde{y}$.
- Output the directed edge $(v_1, \tilde{x} \circ \pi_1) \to (v_2, \tilde{y} \circ \pi_2)$.

Figure 4: Reduction from UG to Max-Acyclic Graph from [17].

Now, if $u \in A'$ then $\pi_1(\ell(v_1)) = \pi_2(\ell(v_2)) = \ell(u)$ and hence,

$$\begin{aligned}\Pr[\tilde{x}_{\pi_1(\ell(v_1))} < \tilde{y}_{\pi_2(\ell(v_2))} \mid u \in A'] &= \Pr[\tilde{x}_{\ell(u)} < \tilde{y}_{\ell(u)}] \\ &\geqslant (1 - 2\varepsilon) \cdot \mathop{\mathbf{E}}_{(a,b) \in E_H} \Pr_{O \sim \mathcal{D}}[O(a) < O(b)] \\ &\geqslant (1 - 2\varepsilon)(1 - \eta). \end{aligned} \tag{5.2}$$

Now, we can show that $\Pr[\tilde{x}_{\pi_1(\ell(v_1))} < \tilde{y}_{\pi_2(\ell(v_2))} \mid u \notin A']$ is at least $(1 - 2\varepsilon)(1 - \eta)$ as above if $\pi_1(\ell(v_1)) = \pi_2(\ell(v_2))$. If $\pi_1(\ell(v_1)) \neq \pi_2(\ell(v_2))$ then $\tilde{x}_{\pi_1(\ell(v_1))}$ and $\tilde{y}_{\pi_1(\ell(v_1))}$ are uncorrelated and are distributed uniformly in $[m]$. Therefore, $\Pr[\tilde{x}_{\pi_1(\ell(v_1))} < \tilde{y}_{\pi_2(\ell(v_2))} \mid u \notin A'] = \frac{\binom{m}{2}}{m^2} = \frac{1}{2} - \frac{1}{2m}$. Thus, for small enough ε and η, we have the following lower-bound.

$$\Pr[\tilde{x}_{\pi_1(\ell(v_1))} < \tilde{y}_{\pi_2(\ell(v_2))} \mid u \notin A'] \geqslant \min\left\{(1 - 2\varepsilon)(1 - \eta), \frac{1}{2} - \frac{1}{2m}\right\} \geqslant \frac{1}{2} - \frac{1}{2m}. \tag{5.3}$$

Plugging equation (5.2) and equation (5.3) into equation (5.1), we get

$$\mathtt{Val}(\mathcal{G}) \geqslant c \cdot (1 - 2\varepsilon)(1 - \eta) + (1 - c) \cdot \left(\frac{1}{2} - \frac{1}{2m}\right).$$

□

The following soundness of the reduction is shown in [17].

Lemma 5.4 (Soundness [17])**.** *If the Unique Games instance G has $\mathsf{val}(G) \leqslant \delta$ then $\mathtt{Val}(\mathcal{G}) \leqslant \frac{1}{2} + \eta + o_t(1) + \delta'$, where $\delta' \to 0$ as $\delta \to 0$.*

Proof of Theorem 1.4 For every $\varepsilon' < 0$, setting $\varepsilon, \eta, \delta > 0$ small enough constants and m large enough, in the completeness of the reduction we have a maximum acyclic subgraph of size at least $\frac{c}{2} + \frac{1}{2} - \varepsilon'$, whereas in the soundness, it is at most $\frac{1}{2} + \varepsilon'$. Since by Theorem 1.2, it is NP-hard to distinguish between $\mathsf{sval}(G) \geqslant \frac{1}{2} - \delta$ and $\mathsf{val}(G) \leqslant \delta$, we get that it is NP-hard to approximate the size of the Maximum Acyclic Subgraph within a factor of $\frac{1/2+\varepsilon'}{1/4+1/2-\varepsilon'-\delta/2} \approx \frac{2}{3}$.

Remark 5.5. Instead of $\mathsf{sval}(G) = \frac{1}{2}$, if we only have $\mathsf{val}(G) = \frac{1}{2}$, then the same construction and the labeling from Lemma 5.3 gives $\mathtt{Val}(\mathcal{G}) \geqslant \frac{5}{8}$. To see this, fix an assignment $\ell : A \cup B \to \Sigma$ which gives $\mathsf{val}(G) \geqslant \frac{1}{2}$. Let α_u denote the fraction of edges incident on u that are satisfied by ℓ. Therefore, we have $\mathsf{val}(G) = \mathbf{E}_{u \in A}[\alpha_u] = \frac{1}{2}$. Using a similar analysis as in the completeness of the reduction, we have $\mathtt{Val}(\mathcal{G}) \geqslant \mathbf{E}_{u \in A}[\alpha_u^2 \cdot (1 - 2\varepsilon)] + \mathbf{E}_{u \in A}[(1 - \alpha_u^2) \cdot \frac{1}{2}] \geqslant (1 - 2\varepsilon)\,\mathbf{E}[\frac{1}{2} + \frac{\alpha_u^2}{2}]$. By the Cauchy–Schwarz inequality $\mathbf{E}[\alpha_u^2] \geqslant (\mathbf{E}[\alpha_u])^2 = \frac{1}{4}$ and hence $\mathtt{Val}(\mathcal{G}) \geqslant (1 - 2\varepsilon) \cdot \frac{5}{8}$. This along with the soundness lemma gives the NP-hardness of $\frac{4}{5}$.

6 Predicates supporting Pairwise Independence

In this section, we prove Theorem 1.5.

6.1 The Austrin-Mossel reduction

Let $\mathcal{D}$ be a distribution on $P^{-1}(1)$ which is balanced and pairwise independent. For a string $x \in [q]^L$ and a permutation $\pi : [L] \to [L]$, define $x \circ \pi \in [q]^L$ such that $(x \circ \pi)_i = x_{\pi(i)}$ for all $i \in [L]$.

Let $G(A, B, E, [L], \{\pi_e\}_{e \in E})$ be an instance of Unique Games. We convert it into a P-CSP instance $\mathcal{I}$ as follows. The variable set of $\mathcal{I}$ is $B \times [q]^L$. The variables are *folded* in the sense that for every assignment $f : B \times [q]^L \to [q]$ to the variables, we enforce that for every $v \in B$, $x \in [q]^L$ and $\alpha \in [q]$,

$$f(v, x + \alpha^L) = f(v, x) + \alpha,$$

where additions are $(\mathtt{mod}\ q)$.

The distribution on the constraints is given in Figure 5:

Lemma 6.1 (Completeness)**.** *If $\mathsf{sval}(G) \geqslant c$, the $\mathcal{I}$ is $(c - \varepsilon)$-satisfiable.*

Proof. Fix an assignment $\ell : A \cup B \to \Sigma$ that gives $\mathsf{sval}(G) \geqslant c$. Let $A' \subseteq A$ be the set of vertices such that all the edges incident on A' are satisfied by ℓ, we know that $|A'| \geqslant c \cdot |A|$. Thus, with probability at least c, $u \in A'$ and all edges attached to it are satisfied by ℓ. Consider the following assignment f to the variables of $\mathcal{I}$: For a variable (v, x), we assign $f(v, x) = x_{\ell(v)}$.

Conditioned on $u \in A'$, we will show that $(f(v_1, x^1 \circ \pi_1), f(v_2, x^2 \circ \pi_2), \ldots, f(v_k, x^k \circ \pi_k)) \in P^{-1}(1)$ with probability $(1 - \varepsilon)$ and this will prove the lemma. Now, $(f(v_2, x^2 \circ \pi_2), \ldots, f(v_k, x^k \circ \pi_k)))$

Let $G(A,B,E,[L],\{\pi_e\}_{e\in E})$ be an instance of Unique Games.

- Select $u \in A$ uniformly at random.
- Select k neighbors $\{v_1,v_2,\ldots,v_k\}$ of u uniformly at random. Let π_j be the constraint between (u,v_j) for all $j \in [k]$.
- Select $x^1,x^2,\ldots,x^k \in [q]^L$, such that for each $i \in [L]$ sample $(x_i^1,x_i^2,\ldots,x_i^k)$ independently as follows:
  - with probability $(1-\varepsilon)$, $(x_i^1,x_i^2,\ldots,x_i^k)$ is sampled from the distribution $\mathcal{D}$.
  - with probability ε, $(x_i^1,x_i^2,\ldots,x_i^k)$ is sampled from $[q]^k$ uniformly at random.
- Output the constraint $((v_1,x^1\circ\pi_1),(v_2,x^2\circ\pi_2),\ldots,(v_k,x^k\circ\pi_k))$.

Figure 5: Reduction from UG to a P-CSP instance $\mathcal{I}$ from [6].

is same as $((x^1\circ\pi_1)_{\ell(v_1)},(x^2\circ\pi_2)_{\ell(v_2)},\ldots,(x^k\circ\pi_k)_{\ell(v_k)})$, which in turns equals $(x^1_{\pi_1(\ell(v_1))},x^2_{\pi_2(\ell(v_2))},\ldots,x^k_{\pi_k(\ell(v_k))})$. Since ℓ satisfies all the edges (u,v_i), we have that for all $j\in[k]$, $\pi_j(\ell(v_j))=\ell(u)=: i$ for some $i\in[L]$. Therefore we get $(f(v_1,x^1\circ\pi_1),f(v_2,x^2\circ\pi_2),\ldots,f(v_k,x^k\circ\pi_k))=(x_i^1,x_i^2,\ldots,x_i^k)$, and according to the distribution, it belongs to $P^{-1}(1)$ with probability $(1-\varepsilon)$. □

We have the following soundness of the reduction.

Lemma 6.2 (Soundness [6]). *If the instance $\mathcal{I}$ is $\left(\frac{P^{-1}(1)}{q^k}+\eta\right)$-satisfiable, then G is $\delta := \delta(\eta,\varepsilon,k,q) > 0$ satisfiable, where $\eta \to 0$ as $\delta \to 0$.*

The completeness and soundness of the reduction, along with our main theorem, imply Theorem 1.5.

Acknowledgements. We would like to thank all the anonymous reviewers for their comments which helped in improving the presentation significantly.

References

[1] SANJEEV ARORA, LÁSZLÓ BABAI, JACQUES STERN, AND Z SWEEDYK: The hardness of approximate optima in lattices, codes, and systems of linear equations. *J. Comput. System Sci.*, 54(2):317–331, 1997. Preliminary version in STOC'93. [doi:10.1006/jcss.1997.1472] 2

[2] 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, 1998. Preliminary version in FOCS'92. [doi:10.1145/278298.278306] 9

[3] SANJEEV ARORA AND SHMUEL SAFRA: Probabilistic checking of proofs: A new characterization of NP. *J. ACM*, 45(1):70–122, 1998. Preliminary version in FOCS'92. [doi:10.1145/273865.273901] 9

[4] PER AUSTRIN, SUBHASH A. KHOT, AND MULI SAFRA: Inapproximability of vertex cover and independent set in bounded degree graphs. *Theory of Computing*, 7(3):27–43, 2011. [doi:10.4086/toc.2011.v007a003] 4, 19, 20

[5] PER AUSTRIN, RAJSEKAR MANOKARAN, AND CENNY WENNER: On the NP-hardness of approximating ordering-constraint satisfaction problems. *Theory of Computing*, 11(10):257–283, 2015. [doi:10.4086/toc.2015.v011a010] 4

[6] PER AUSTRIN AND ELCHANAN MOSSEL: Approximation resistant predicates from pairwise independence. *Comput. Complexity*, 18(2):249–271, 2009. [doi:10.1007/s00037-009-0272-6] 5, 24

[7] NIKHIL BANSAL, ANUPAM GUPTA, AND GURU GURUGANESH: On the Lovász theta function for independent sets in sparse graphs. *SIAM J. Comput.*, 47(3):1039–1055, 2018. Preliminary version in STOC'15. [doi:10.1137/15M1051002] 4

[8] NIKHIL BANSAL AND SUBHASH A. KHOT: Inapproximability of hypergraph vertex cover and applications to scheduling problems. In *Proc. 37th Internat. Colloq. on Automata, Languages, and Programming (ICALP'10)*, pp. 250–261. Springer, 2010. [doi:10.1007/978-3-642-14165-2_22] 4

[9] BOAZ BARAK, SIU ON CHAN, AND PRAVESH K. KOTHARI: Sum of squares lower bounds from pairwise independence. In *Proc. 47th STOC*, pp. 97–106. ACM Press, 2015. [doi:10.1145/2746539.2746625] 5

[10] BOAZ BARAK, PRAVESH K. KOTHARI, AND DAVID STEURER: Small-Set Expansion in Shortcode Graph and the 2-to-2 Conjecture. In *Proc. 10th Innovations in Theoret. Comp. Sci. conf. (ITCS'19)*, pp. 9:1–12. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2019. [doi:10.4230/LIPIcs.ITCS.2019.9] 6

[11] AMEY BHANGALE, RAJIV GANDHI, MOHAMMAD TAGHI HAJIAGHAYI, ROHIT KHANDEKAR, AND GUY KORTSARZ: Bi-covering: Covering edges with two small subsets of vertices. *SIAM J. Discr. Math.*, 31(4):2626–2646, 2017. [doi:10.1137/16M1082421] 4

[12] AMEY BHANGALE AND SUBHASH A. KHOT: UG-hardness to NP-hardness by losing half. In *Proc. 34th Comput. Complexity Conf. (CCC'19)*, pp. 3:1–20. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2019. [doi:10.4230/LIPIcs.CCC.2019.3] 1

[13] SIU ON CHAN: Approximation resistance from pairwise-independent subgroups. *J. ACM*, 63(3):27:1–32, 2016. [doi:10.1145/2873054] 4, 5

[14] IRIT DINUR, SUBHASH A. KHOT, GUY KINDLER, DOR MINZER, AND MULI SAFRA: On non-optimally expanding sets in Grassmann graphs. In *Proc. 50th STOC*, pp. 940–951. ACM Press, 2018. [doi:10.1145/3188745.3188806] 2, 3, 4, 6, 7

[15] IRIT DINUR, SUBHASH A. KHOT, GUY KINDLER, DOR MINZER, AND MULI SAFRA: Towards a proof of the 2-to-1 games conjecture? In *Proc. 50th STOC*, pp. 376–389. ACM Press, 2018. [doi:10.1145/3188745.3188804] 2, 3, 4, 7, 9, 10, 11, 12, 13, 16, 17

[16] URIEL FEIGE, SHAFI GOLDWASSER, LÁSZLÓ LOVÁSZ, SHMUEL SAFRA, AND MARIO SZEGEDY: Interactive proofs and the hardness of approximating cliques. *J. Comput. System Sci.*, 43(2):268–292, 1996. [doi:10.1145/226643.226652] 9

[17] VENKATESAN GURUSWAMI, RAJSEKAR MANOKARAN, AND PRASAD RAGHAVENDRA: Beating the random ordering is hard: Inapproximability of maximum acyclic subgraph. In *Proc. 49th FOCS*, pp. 573–582. IEEE Comp. Soc., 2008. [doi:10.1109/FOCS.2008.51] 2, 4, 5, 21, 22, 23

[18] SUBHASH A. KHOT: On the power of unique 2-prover 1-round games. In *Proc. 34th STOC*, pp. 767–775. ACM Press, 2002. [doi:10.1145/509907.510017] 2, 8

[19] SUBHASH A. KHOT, GUY KINDLER, ELCHANAN MOSSEL, AND RYAN O'DONNELL: Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM J. Comput.*, 37(1):319–357, 2007. [doi:10.1137/S0097539705447372] 2

[20] SUBHASH A. KHOT, DOR MINZER, AND MULI SAFRA: On independent sets, 2-to-2 games, and Grassmann graphs. In *Proc. 49th STOC*, pp. 576–589. ACM Press, 2017. [doi:10.1145/3055399.3055432] 2, 3, 4, 7, 10

[21] SUBHASH A. KHOT, DOR MINZER, AND MULI SAFRA: Pseudorandom sets in Grassmann graph have near-perfect expansion. In *Proc. 59th FOCS*, pp. 592–601. IEEE Comp. Soc., 2018. [doi:10.1109/FOCS.2018.00062] 2, 3, 4, 6, 7, 13, 16

[22] SUBHASH A. KHOT AND ASSAF NAOR: Approximate kernel clustering. *Mathematika*, 55(1–2):129–165, 2009. Preliminary version in FOCS'08. [doi:10.1112/S002557930000098X] 2

[23] SUBHASH A. KHOT AND ODED REGEV: Vertex cover might be hard to approximate to within $2-\varepsilon$. *J. Comput. System Sci.*, 74(3):335–349, 2008. [doi:10.1016/j.jcss.2007.06.019] 2, 4

[24] SUBHASH A. KHOT, MADHUR TULSIANI, AND PRATIK WORAH: A characterization of strong approximation resistance. In *Proc. 46th STOC*, pp. 634–643. ACM Press, 2014. [doi:10.1145/2591796.2591817] 2

[25] SUBHASH A. KHOT AND NISHEETH K. VISHNOI: The Unique Games Conjecture, integrality gap for cut problems and embeddability of negative-type metrics into ℓ_1. *J. ACM*, 62(1):8:1–23, 2015. [doi:10.1145/2629614] 2

[26] ALANTHA NEWMAN: The maximum acyclic subgraph problem and degree-3 graphs. In *Proc. 4th Internat. Workshop on Approximation Algorithms for Combinat. Opt. Probl. (APPROX'01)*, pp. 147–158. Springer, 2001. [doi:10.1007/3-540-44666-4_18] 4

[27] PRASAD RAGHAVENDRA: Optimal algorithms and inapproximability results for every CSP? In *Proc. 40th STOC*, pp. 245–254. ACM Press, 2008. [doi:10.1145/1374376.1374414] 2, 5

[28] RAN RAZ: A parallel repetition theorem. *SIAM J. Comput.*, 27(3):763–803, 1998. [doi:10.1137/S0097539795280895] 10

AUTHORS

Amey Bhangale
Assistant professor
University of California, Riverside
California, USA
bhangale@cs.ucr.edu
https://www.cs.ucr.edu/~bhangale/

Subhash Khot
Professor
Courant Institute of Mathematical Sciences
New York University, USA
khot@cs.nyu.edu
https://cs.nyu.edu/~khot/

ABOUT THE AUTHORS

AMEY BHANGALE is an Assistant Professor in the Computer Science Department at the University of California, Riverside.

Amey got his undergraduate degree from VJTI, Mumbai. In the final year of his undergraduate studies, he met Prof. Rajiv C. Gandhi, who was visiting Mumbai in 2010/11 on a Fulbright scholarship and taught a course on approximation algorithms at VJTI. This encounter influenced Amey to pursue research in the theory of computing.

Amey graduated from Rutgers University in 2017 under the supervision of Swastik Kopparty. He spent two wonderful years in Israel where he was a postdoctoral fellow at the Weizmann Institute of Science, hosted by Irit Dinur. He is interested in approximation algorithms, hardness of approximation and the analysis of Boolean functions. He enjoys playing and watching tennis (and hardly ever says no to anyone who invites him to play tennis).

SUBHASH KHOT is a Professor in the Computer Science Department at New York University. He has a Ph. D. from Princeton University, completed in 2003 under the supervision of Sanjeev Arora.

His teacher and mentor at Vyankatrao Highschool, Mr. Vaman G. Gogate played a decisive role in directing his attention to mathematics. If not for Mr. Gogate's guidance and some fortuitous turn of events, the chance of someone from the remote town of Ichalkaranji pursuing mathematical research was strictly nil. Subsequently, Mr. Gogate also mentored Amit Deshpande and Raghav Kulkarni (CS theorists) and Abhijit Gadde (string theorist). He is in touch with all his past students and continues to provide guidance on all aspects of life.