Carnegie Mellon University
Research Showcase

Computer Science Department | School of Computer Science

7-1-2010

SDP gaps for 2-to-1 and other Label-Cover variants

Venkatesan Guruswami
Carnegie Mellon University, guruswami@cmu.edu

Subash Khot
New York University

Ryan O'Donnell
Carnegie Mellon University, odonnell@cs.cmu.edu

Preyas Popat
New York University

Madhur Tulsiani
Institute for Advanced Study

See next page for additional authors

Follow this and additional works at: http://repository.cmu.edu/compsci

Recommended Citation
Guruswami, Venkatesan; Khot, Subash; O'Donnell, Ryan; Popat, Preyas; Tulsiani, Madhur; and Wu, Yi, "SDP gaps for 2-to-1 and other Label-Cover variants" (2010). *Computer Science Department.* Paper 1154.
http://repository.cmu.edu/compsci/1154

This Conference Proceeding is brought to you for free and open access by the School of Computer Science at Research Showcase. It has been accepted for inclusion in Computer Science Department by an authorized administrator of Research Showcase. For more information, please contact research-showcase@andrew.cmu.edu.

Authors
Venkatesan Guruswami, Subash Khot, Ryan O'Donnell, Preyas Popat, Madhur Tulsiani, and Yi Wu

This conference proceeding is available at Research Showcase: http://repository.cmu.edu/compsci/1154

SDP gaps for 2-to-1 and other Label-Cover variants

Venkatesan Guruswami*[1], Subhash Khot**[2], Ryan O'Donnell***[1], Preyas Popat**[2], Madhur Tulsiani†[3], and Yi Wu***[1]

[1] Computer Science Department
Carnegie Mellon University
[2] Computer Science Department
New York University
[3] School of Mathematics
Institute for Advanced Study

Abstract. In this paper we present semidefinite programming (SDP) gap instances for the following variants of the Label-Cover problem, closely related to the Unique Games Conjecture: (i) 2-to-1 Label-Cover; (ii) 2-to-2 Label-Cover; (iii) α-constraint Label-Cover. All of our gap instances have perfect SDP solutions. For alphabet size K, the integral optimal solutions have value: (i) $O(1/\sqrt{\log K})$; (ii) $O(1/\log K)$; (iii) $O(1/\sqrt{\log K})$.
Prior to this work, there were no known SDP gap instances for any of these problems with perfect SDP value and integral optimum tending to 0.

1 Introduction

1.1 The Unique Games Conjecture and its variants

Since its introduction in 2002, the *Unique Games Conjecture* (UGC) of Khot [8] has proved highly influential and powerful in the study of probabilistically checkable proofs (PCPs) and approximation algorithms. Assuming the UGC yields many strong — and often, optimal — hardness of approximation results that we have been unable to obtain assuming only P $\neq$ NP. Perhaps the acme of this line of research so far is the work of Raghavendra [12], who showed the following result:

Theorem 1. *([12], informally.) Let $\mathcal{C}$ be any bounded-arity constraint satisfaction problem (CSP). Assume the Unique Games Conjecture. Then for a certain*

* Research supported by a Packard Fellowship and US-Israel BSF-2008293.
** Research supported by NSF CAREER grant CCF-0833228, NSF Expeditions grant CCF-0832795, and BSF grant 2008059.
*** Research supported by NSF grants CCF-0747250 and CCF-0915893, BSF grant 2008477, and Sloan and Okawa fellowships.
† Research supported by NSF Grant CCF-0832797.

semidefinite programming (SDP) relaxation of $\mathcal{C}$, the SDP gap for $\mathcal{C}$ is the same as the optimal polynomial-time approximability gap for $\mathcal{C}$, up to an additive constant $\epsilon > 0$ which can be arbitrarily small.

Unfortunately, because of the additive ϵ term, Raghavendra's work is not applicable (even granting the UGC or any related conjecture) for the important case of completely *satisfiable* CSPs; equivalently, PCPs with *perfect completeness*. A good example of this comes from *coloring problems*; e.g., the very well known problem of coloring 3-colorable graphs. The UGC does not help in deducing any hardness result for such problems. Indeed the first strong hardness result for it, due to Dinur, Mossel, and Regev [3], used instead certain variants of UGC which have perfect completeness, namely, the "2-to-1 Conjecture", the "2-to-2 Conjecture", and the "α-Constraint Conjecture". (These conjectures will be described formally in Section 3.) An instance of Label-Cover with α-constraints was also implicit in the result of Dinur and Safra [4] on the hardness of approximating minimum vertex cover.

Recently, several more works have needed to use these alternate conjectures with perfect completeness: e.g., O'Donnell and Wu [11] and Tang [16] on Max-3CSP, Guruswami and Sinop [6] on Max-k-Colorable-Subgraph.

1.2 Statements of the conjectures

Let us briefly give some definitions so that we may state some of the aforementioned conjectures more precisely.

Definition 1. *A* Label-Cover *instance $\mathcal{L}$ is defined by a tuple $((V,E),R,\Psi)$. Here (V,E) is a graph, R is a positive integer and Ψ is a set of constraints (relations), one for each edge: $\Psi = \{\psi_e \subseteq \{1,\ldots,R\}^2 \mid e \in E\}$. A labeling A is a mapping $A : V \to [R]$. We say that an edge $e = (u,v)$ is* satisfied *by A if $(A(u), A(v)) \in \psi_e$. We define:*

$$\mathrm{OPT}(\mathcal{L}) = \max_{A:V\to[R]} \quad \mathbf{Pr}_{e=(u,v)\in E}[(A(u), A(v)) \in \psi_e]$$

Here the probability is over the uniform distribution of edges, i.e. each edge is equally likely to be picked.

Definition 2. *A constraint $\psi \subseteq \{1,\ldots,R\}^2$ is said to be a d-*to-1 projection *if there is a map $\pi : [R] \to [R]$ such that for each element $j \in [R]$ we have $|\pi^{-1}(j)| \le d$, and $(i,j) \in \psi$ if and only if $j = \pi(i)$. A Label-Cover instance is said to be d-to-1 if all its constraints are d-to-1 projections.*

We now state some conjectures on the inapproximability of Unique and 2-to-1 Label-Cover. In Section 3.1 we will also discuss the two other variants of the Label-Cover problem, based on 2-to-2 and "α" constraints, and the associated 2-to-2 and α-Constraint Conjectures on their inapproximability.

Conjecture 1. [7] (**Unique Games Conjecture**) For any $\epsilon, \delta > 0$, it is NP-hard to decide whether a 1-to-1 bipartite Label-Cover instance $\mathcal{L}$ has $\mathrm{OPT}(\mathcal{L}) \ge 1-\epsilon$ or has $\mathrm{OPT}(\mathcal{L}) \le \delta$.

Notice that the above problem is in P when $\epsilon = 0$.

Conjecture 2. [7] (**2-to-1 Conjecture**) For any $\delta > 0$, it is NP-hard to decide whether a 2-to-1 bipartite Label-Cover instance $\mathcal{L}$ has $\mathrm{OPT}(\mathcal{L}) = 1$ or has $\mathrm{OPT}(\mathcal{L}) \leq \delta$.

1.3 Evidence for and against

Despite significant work, the status of the Unique Games Conjecture — as well as the 2-to-1, 2-to-2, and α-Constraint Conjectures — is unresolved. Towards *disproving* the conjectures, the best algorithms known are due to Charikar, Makarychev, and Makarychev [2]. Using somewhat strong SDP relaxations, those authors gave polynomial-time SDP-rounding algorithms which achieve:

- Value $K^{-\epsilon/(2-\epsilon)}$ (roughly) for Unique Label-Cover instances with SDP value $1 - \epsilon$ over alphabets of size K.
- Value $K^{-3+2\sqrt{2}-\epsilon}$ for 2-to-1 Label-Cover instances with SDP value $1 - \Theta(\epsilon)$ over alphabets of size K.

The best evidence in *favor* of the Unique Games Conjecture is probably the existence of strong SDP gaps. The first such gap was given by Khot and Vishnoi [10]: they constructed a family of Unique Label-Cover instances over alphabet size K with SDP value $1 - \epsilon$ and integral optimal value $K^{-\Theta(\epsilon)}$. In addition to roughly matching the CMM algorithm, the Khot–Vishnoi gaps have the nice property that they even hold with triangle inequality constraints added into the SDP. Even stronger SDP gaps for UGC were obtained recently by Raghavendra and Steurer [13].

Standing in stark contrast to this is the situation for the 2-to-1 Conjecture and related variants with perfect completeness. Prior to this work, there were *no known* SDP gap families for these problems with SDP value 1 and integral optimal value tending to 0 with the alphabet size. Indeed, there was hardly any evidence for these conjectures, beyond the fact that Charikar, Makarychev, and Makarychev failed to disprove them.

1.4 SDP gaps as a reduction tool

In addition to being the only real evidence towards the validity of the UGC, SDP gaps for Unique Games have served another important role: they are the starting points for strong SDP gaps for other important optimization problems. A notable example of this comes in the work of Khot and Vishnoi [10] who used the UG gap instance to construct a super-constant integrality gap for the Sparsest Cut-SDP with triangle inequalities, thereby refuting the Goemans-Linial conjecture that the gap was bounded by $O(1)$. They also used this approach to show that the integrality gap of the Max-Cut SDP remains 0.878 when triangle inequalities are added. Indeed the approach via Unique Games remains the *only known* way to get such strong gaps for Max Cut. Recently, even stronger gaps for Max-Cut

were shown using this framework in [9, 13]. Another example of a basic problem for which a SDP gap construction is only known via the reduction from Unique Games is Maximum Acyclic Subgraph [5].

In view of these results, it is fair to say that SDP gaps for Unique Games are significant unconditionally, regardless of the truth of the UGC. Given the importance of 2-to-1 and related conjectures in reductions to satisfiable CSPs and other problems like coloring where perfect completeness is crucial, SDP gaps for 2-to-1 Label-Cover and variants are worthy of study even beyond the motivation of garnering evidence towards the associated conjectures on their inapproximability.

2 Our results

Label-Cover admits a natural semidefinite programming relaxation (see Figure 1). In this paper, we show the following results on the limitations of the basic semidefinite programming relaxation for Label-Cover instances with 2-to-1, 2-to-2, and α constraints:

- There is an instance of 2-to-2 Label-Cover with alphabet size K and optimum value $O(1/\log K)$ on which the SDP has value 1.
- There are instances of 2-to-1 and α-constraint Label-Cover with alphabet size K and optimum value $O(1/\sqrt{\log K})$ on which the SDP has value 1.

In both cases the instances have size $2^{\Omega(K)}$.

We note that if we only require the SDP value to be $1-\epsilon$ instead of 1, then integrality gaps for all these problems easily follow from gaps from Unique Games, constructed by Khot and Vishnoi [10] (by duplicating labels appropriately to modify the constraints). However, the motivation behind these conjectures is applications where it is important that the completeness is 1. Another difference between the 2-to-1 Label-Cover and the Unique Label-Cover is the fact that for 2-to-1 instances, it is consistent with known algorithmic results of [2] that OPT be as low as K^{-c} for some $c > 0$ independent of ϵ, when the SDP value is $1-\epsilon$. It is an interesting question if OPT can indeed be this low even when the SDP value is 1. Our constructions do not address this question, as we only show $\mathrm{OPT} = O(1/\sqrt{\log K})$.

We also point out that our integrality gaps are for special cases of the Label-Cover problem where the constraints can be expressed as difference equations over $\mathbb{F}_2$-vector spaces. For example, for 2-to-2 Label-Cover, each constraint ϕ_e is of the form $x - y \in \{\alpha, \alpha+\gamma\}$ where $\alpha, \gamma \in \mathbb{F}_2^k$ are constants. For such constraints, the probable of deciding whether an instance is completely satisfiable ($\mathrm{OPT} = 1$) or not ($\mathrm{OPT} < 1$) is in fact in P. To see this, one can treat the coordinates $(x_1, \ldots, x_k)$ and $(y_1, \ldots, y_k)$ as separate boolean variables and introduce an auxiliary boolean variable z_e for each constraint. We can then rewrite the constraint as a conjunction of linear equations over $\mathbb{F}_2$:

$$\bigwedge_{i=1}^{k} \left(x_i - y_i - z_e \cdot \gamma_i = \alpha_i\right).$$

Here $x_i, y_i, \alpha_i, \gamma_i$ denote the ith coordinates of the corresponding vectors. Deciding whether a system of linear equations is completely satisfiable is of course in P. Alternatively, one can note that constraints $x - y \in \{\alpha, \alpha + \gamma\} \bmod \mathbb{F}_2^k$ are *Mal'tsev constraints*, and hence deciding satisfiability of CSPs based on them is in P by the work of Bulatov and Dalmau [1].

Despite this tractability, the SDPs fail badly to decide satisfiability. This situation is similar to the very strong SDP gaps known for problems such as 3-XOR (see [14], [17]) for which deciding complete satisfiability is easy.

3 Preliminaries and Notation

3.1 Label-Cover Problems

In Figure 1, we write down a natural SDP relaxation for the Label-Cover problem. The relaxation is over the vector variables $\mathbf{z}_{(v,i)}$ for every vertex $v \in V$ and label $i \in [R]$.

$$\begin{array}{lll} \text{maximize} & \mathop{\mathbf{E}}_{e=(u,v)\in E}\Big[\sum_{i,j\in\psi_e} \big\langle \mathbf{z}_{(u,i)}, \mathbf{z}_{(v,j)} \big\rangle\Big] & \\ \text{subject to} & \sum_{i\in[R]} \big\|\mathbf{z}_{(v,i)}\big\|^2 = 1 & \forall\, v \in V \\ & \big\langle \mathbf{z}_{(v,i)}, \mathbf{z}_{(v,j)} \big\rangle = 0 & \forall\, i \neq j \in [R], v \in V \end{array}$$

Fig. 1. SDP for Label-Cover

Our goal in this work is to study integrality gaps for the above SDP for various special cases of the Label-Cover problem. We already discussed the Unique Games and 2-to-1 conjectures on the hardness of certain very special cases of Label-Cover. We now discuss two other variants of Label-Cover and their conjectured inapproximability.

Definition 3. *A constraint $\psi \subseteq \{1, \ldots, 2R\}^2$ is said to be a* 2-to-2 constraint *if there are two permutations σ_1, $\sigma_2 : \{1, \ldots, 2R\} \to \{1, \ldots, 2R\}$ such that $(i, j) \in \psi$ if and only if $(\sigma_1(i), \sigma_2(j)) \in T$ where*

$$T := \{(2l-1, 2l-1), (2l-1, 2l), (2l, 2l-1), (2l, 2l)\}_{l=1}^{R}.$$

A Label-Cover instance is said to be 2-to-2 if all its constraints are 2-to-2 constraints.

*A constraint $\psi \subseteq \{1, \ldots, 2R\}^2$ is said to be an α-*constraint *if there are two permutations σ_1, $\sigma_2 : \{1, \ldots, 2R\} \to \{1, \ldots, 2R\}$ such that $(i, j) \in \psi$ if and only if $(\sigma_1(i), \sigma_2(i)) \in T'$ where*

$$T' := \{(2l-1, 2l-1), (2l-1, 2l), (2l, 2l-1)\}_{l=1}^{R}.$$

A Label-Cover instance is said to be α if all its constraints are α constraints.

Conjecture 3. [3] (**2-to-2 Conjecture**) For any $\delta > 0$, it is NP-hard to decide whether a 2-to-2 Label-Cover instance $\mathcal{L}$ has $\mathrm{OPT}(\mathcal{L}) = 1$ or has $\mathrm{OPT}(\mathcal{L}) \leq \delta$.

It was shown in [3] that the 2-to-2 Conjecture is no stronger than the 2-to-1 Conjecture.

Conjecture 4. [3] (α **Conjecture**) For any $\delta > 0$, it is NP-hard to decide whether a α Label-Cover instance $\mathcal{L}$ has $\mathrm{OPT}(\mathcal{L}) = 1$ or has $\mathrm{OPT}(\mathcal{L}) \leq \delta$.

3.2 Fourier Analysis

Let $\mathcal{V} := \{f : \mathbb{F}_2^k \to \mathcal{R}\}$ denote the vector space of all real functions on $\mathbb{F}_2^k$, where addition is defined as point-wise addition. We always think of $\mathbb{F}_2^k$ as a probability space under the uniform distribution, and therefore use notation such as $\|f\|_p := \mathbf{E}_{x \in \mathbb{F}_2^k}[|f(x)|^p]$. For f, $g \in \mathcal{F}$, we also define the inner product $\langle f, g \rangle := \mathbf{E}[f(x)g(x)]$.

For any $\alpha \in \mathbb{F}_2^k$ the Fourier character $\chi_\alpha \in \mathcal{F}$ is defined by $\chi_\alpha(x) := (-1)^{\alpha \cdot x}$. The Fourier characters form an orthonormal basis for $\mathcal{V}$ with respect to the above inner product, hence every function $f \in \mathcal{V}$ has a unique representation as $f = \sum_{\alpha \in \mathbb{F}_2^k} \widehat{f}(\alpha)\chi_\alpha$, where the Fourier coefficient $\widehat{f}(\alpha) := \langle f, \chi_\alpha \rangle$.

We also sometimes identify each α with the set $S_\alpha = \{i \mid \alpha_i = 1\}$ and denote the Fourier coefficients as $\widehat{f}(S)$. We use the notation $|\alpha|$ for $|S_\alpha|$, the number of coordinates where α is 1.

The following well-known fact states that the norm of a function on $\mathbb{F}_2^k$ is unchanged when expressing it in the basis of the characters.

Proposition 1. *(***Parseval's identity***) For any $f : \mathbb{F}_2^k \to \mathbb{R}$, $\sum_{\alpha \in \mathbb{F}_2^k} \widehat{f}(\alpha)^2 = \|f\|_2^2 = \mathbf{E}[f(x)^2]$.*

We shall also need the following result due to Talagrand ("Proposition 2.3" in [15]), proven using hypercontractivity methods:

Theorem 2. *Suppose $F : \mathbb{F}_2^k \to \mathbb{R}$ has $\mathbf{E}[F] = 0$. Then*

$$\sum_{\alpha \in \mathbb{F}_2^k \setminus \{0\}} \widehat{F}(\alpha)^2/|\alpha| = O\left(\frac{\|F\|_2^2}{\ln(\|F\|_2/(e\|F\|_1))}\right).$$

More precisely, we will need the following easy corollary:

Corollary 1. *If $F : \mathbb{F}_2^k \to \{0, 1\}$ has mean $1/K$, then*

$$\widehat{F}(0)^2 + \sum_{\alpha \in \mathbb{F}_2^k \setminus \{0\}} \widehat{F}(\alpha)^2/|\alpha| = O\left(1/(K \log K)\right)$$

Proof. We have $\widehat{F}(0)^2 = \mathbf{E}[F]^2 = 1/K^2 \leq O(1/(K \log K))$, so we can disregard this term. As for the sum, we apply Theorem 2 to the function $F' = F - 1/K$, which has mean 0 as required for the theorem. It is easy to calculate that $\|F'\|_2 = \Theta(1/\sqrt{K})$ and $\|F'\|_1 = \Theta(1/K)$, and so the result follows.

4 Integrality Gap for 2-to-2 Games

We first give an integrality gap for label cover with 2-to-2 constraints. The instance for 2-to-1 label cover will be an extension of the one below. In fact, our analysis of OPT in the 2-to-1 case will follow simply by reducing it to the analysis of OPT for the 2-to-2 instance below.

The vertex set V in our instance is same as the vertex set of the Unique Games integrality gap instance constructed in [10]. Let $\mathcal{F} := \{f : \mathbb{F}_2^k \mapsto \{-1, 1\}\}$ denote the family of all boolean functions on $\mathbb{F}_2^k$. For $f, g \in \mathcal{F}$, define the product fg as $(fg)(x) := f(x)g(x)$. Consider the equivalence relation $\sim$ on $\mathcal{F}$ defined as $f \sim g \Leftrightarrow \exists \alpha \in \mathbb{F}_2^k \; s.t. \; f \equiv g\chi_\alpha$. This relation partitions $\mathcal{F}$ into equivalence classes $\mathcal{P}_1, \ldots, \mathcal{P}_n$, with $n := 2^K/K$. The vertex set V consists of the equivalence classes $\{\mathcal{P}_i\}_{i\in[n]}$. We denote by $[\mathcal{P}_i]$ the lexicographically smallest function in the class $\mathcal{P}_i$ and by $\mathcal{P}_f$, the class containing f.

We take the label set to be of size K and identify $[K]$ with $\mathbb{F}_2^k$ in the obvious way. For each tuple of the form (γ, f, g) where $\gamma \in \mathbb{F}_2^k \setminus \{0\}$ and $f, g \in \mathcal{F}$ are such that $(1 + \chi_\gamma)f \equiv (1 + \chi_\gamma)g$, we add a constraint $\psi_{(\gamma,f,g)}$ between the vertices $\mathcal{P}_f$ and $\mathcal{P}_g$. Note that the condition on f and g is equivalent to saying that $\chi_\gamma(x) = 1 \implies f(x) = g(x)$. If $f = [\mathcal{P}_f]\chi_\alpha$ and $g = [\mathcal{P}_g]\chi_\beta$ and if $A : [n] \to \mathbb{F}_2^k$ denotes the labeling, the relation $\psi_{(\gamma,f,g)}$ is defined as

$$(A(\mathcal{P}_f), A(\mathcal{P}_g)) \in \psi_{(\gamma,f,g)} \quad \Leftrightarrow \quad (A(\mathcal{P}_f) + \alpha) - (A(\mathcal{P}_g) + \beta) \;\in\; \{0, \gamma\}.$$

Note that for any $\omega \in \mathbb{F}_2^k$, the constraint maps the labels $\{\omega, \omega + \gamma\}$ for $\mathcal{P}_f$ to the labels $\{\omega+\alpha-\beta, \omega+\alpha-\beta+\gamma\}$ for $\mathcal{P}_g$ in a 2-to-2 fashion. We denote the set of all constraints by Ψ. We remark that, as in [10], our integrality gap instances contain multiple constraints on each pair of vertices.

4.1 SDP Solution

We give below a set of feasible vectors $\mathbf{z}_{(\mathcal{P}_i,\alpha)} \in \mathbb{R}^K$ for every equivalence class $\mathcal{P}_i$ and every label α, achieving SDP value 1. Identifying each coordinate with an $x \in \mathbb{F}_2^k$, we define the vectors as

$$\mathbf{z}_{(\mathcal{P}_i,\alpha)}(x) \;:=\; \frac{1}{K}([\mathcal{P}_i]\chi_\alpha)(x).$$

It is easy to check that $\left\|\mathbf{z}_{(\mathcal{P}_i,\alpha)}\right\|^2 = 1/K$ for each of the vectors, which satisfies the first constraint. Also, $\mathbf{z}_{(\mathcal{P}_i,\alpha)}$ and $\mathbf{z}_{(\mathcal{P}_i,\beta)}$ are orthogonal for $\alpha \neq \beta$ since

$$\left\langle \mathbf{z}_{(\mathcal{P}_i,\alpha)}, \mathbf{z}_{(\mathcal{P}_i,\beta)} \right\rangle \;=\; \frac{1}{K^2}\left\langle [\mathcal{P}_i]\chi_\alpha, [\mathcal{P}_i]\chi_\beta \right\rangle \;=\; \frac{1}{K^2}\left\langle \chi_\alpha, \chi_\beta \right\rangle \;=\; 0$$

using the fact that $[\mathcal{P}_i]^2 = 1$. The following claim proves that the solution achieves SDP value 1.

Claim. For any edge e indexed by a tuple (γ, f, g) with $f(1 + \chi_\gamma) \equiv g(1 + \chi_\gamma)$, we have

$$\sum_{\omega_1,\omega_2 \in \psi_{(\gamma,f,g)}} \left\langle \mathbf{z}_{(\mathcal{P}_f,\omega_1)}, \mathbf{z}_{(\mathcal{P}_g,\omega_2)} \right\rangle \;=\; 1$$

Proof. Let $f \equiv [\mathcal{P}_f]\chi_\alpha$ and $g \equiv [\mathcal{P}_g]\chi_\beta$. Then, $(\omega_1, \omega_2) \in \psi_e$ iff $(\omega_1 + \alpha) - (\omega_2 + \beta) \in \{0, \gamma\}$. Therefore, the above quantity equals (divided by 2 to account for double counting of ω)

$$\begin{aligned} & \frac{1}{2} \cdot \sum_{\omega} \Big(\big\langle \mathbf{z}_{(\mathcal{P}_f,\omega+\alpha)}, \mathbf{z}_{(\mathcal{P}_g,\omega+\beta)} \big\rangle + \big\langle \mathbf{z}_{(\mathcal{P}_f,\omega+\alpha+\gamma)}, \mathbf{z}_{(\mathcal{P}_g,\omega+\beta)} \big\rangle \\ & \qquad + \big\langle \mathbf{z}_{(\mathcal{P}_f,\omega+\alpha)}, \mathbf{z}_{(\mathcal{P}_g,\omega+\beta+\gamma)} \big\rangle + \big\langle \mathbf{z}_{(\mathcal{P}_f,\omega+\alpha+\gamma)}, \mathbf{z}_{(\mathcal{P}_g,\omega+\beta+\gamma)} \big\rangle \Big) \\ = \; & \frac{1}{2} \sum_{\omega} \big\langle \mathbf{z}_{(\mathcal{P}_f,\omega+\alpha)} + \mathbf{z}_{(\mathcal{P}_f,\omega+\alpha+\gamma)}, \mathbf{z}_{(\mathcal{P}_f,\omega+\beta)} + \mathbf{z}_{(\mathcal{P}_f,\omega+\beta+\gamma)} \big\rangle \qquad (1) \end{aligned}$$

However, for each ω, we have $\mathbf{z}_{(\mathcal{P}_f,\omega+\alpha)} + \mathbf{z}_{(\mathcal{P}_f,\omega+\alpha+\gamma)} = \mathbf{z}_{(\mathcal{P}_f,\omega+\beta)} + \mathbf{z}_{(\mathcal{P}_f,\omega+\beta+\gamma)}$, since for all coordinates x,

$$\begin{aligned} \mathbf{z}_{(\mathcal{P}_f,\omega+\alpha)}(x) \; + \; \mathbf{z}_{(\mathcal{P}_f,\omega+\alpha+\gamma)}(x) \; &= \; \frac{1}{K}([\mathcal{P}_f]\chi_{\omega+\alpha}(x) \; + \; [\mathcal{P}_f]\chi_{\omega+\alpha+\gamma}(x)) \\ &= \frac{1}{K}(f(x) + f\chi_\gamma)\chi_\omega(x) = \frac{1}{K}(g(x) + g\chi_\gamma)\chi_\omega(x) \\ = \; \frac{1}{K}([\mathcal{P}_g]\chi_{\omega+\beta}(x) \; + \; [\mathcal{P}_g]\chi_{\omega+\beta+\gamma}(x)) \; &= \; \mathbf{z}_{(\mathcal{P}_f,\omega+\beta)}(x) \; + \; \mathbf{z}_{(\mathcal{P}_f,\omega+\beta+\gamma)}(x). \end{aligned}$$

This completes the proof as the value of (1) then becomes

$$\frac{1}{2} \sum_{\omega} \left\| \mathbf{z}_{(\mathcal{P}_f,\omega+\alpha)} + \mathbf{z}_{(\mathcal{P}_f,\omega+\alpha+\gamma)} \right\|^2 = \frac{1}{2} \sum_{\omega} \left(\left\| \mathbf{z}_{(\mathcal{P}_f,\omega+\alpha)} \right\|^2 + \left\| \mathbf{z}_{(\mathcal{P}_f,\omega+\alpha+\gamma)} \right\|^2 \right) = 1.$$

4.2 Soundness

We now prove that any labeling of the instance described above, satisfies at most $O(1/\log K)$ fraction of the constraints. Let $A : V \to \mathbb{F}_2^k$ be a labeling of the vertices. We extend it to a labeling of all the functions in $\mathcal{F}$ by defining $A([\mathcal{P}_i]\chi_\alpha) := A(\mathcal{P}_i) + \alpha$.

For each $\alpha \in \mathbb{F}_2^k$, define $A_\alpha : \mathcal{F} \to \{0, 1\}$ to be the indicator that A's value is α. By definition, the fraction of constraints satisfied by the labeling A is

$$\begin{aligned} \mathrm{val}(A) &= \mathop{\mathbf{E}}_{(\gamma,f,g)\in\Psi} \left[\sum_{\alpha\in\mathbb{F}_2^k} A_\alpha(f)(A_\alpha(g) + A_{\alpha+\gamma}(g)) \right] \\ &= \mathop{\mathbf{E}}_{(\gamma,f,g)\in\Psi} \left[\sum_{\alpha\in\mathbb{F}_2^k} A_\alpha(f)(A_\alpha(g) + A_\alpha(g\chi_\gamma)) \right] \\ &= 2 \cdot \mathop{\mathbf{E}}_{(\gamma,f,g)\in\Psi} \left[\sum_{\alpha\in\mathbb{F}_2^k} A_\alpha(f)(A_\alpha(g) \right] \qquad (2) \end{aligned}$$

where the last equality used the fact that for every tuple $(\gamma, f, g) \in \Psi$, we also have $(\gamma, f, g\chi_\gamma) \in \Psi$.

Note that the extended labeling $A : \mathcal{F} \to \mathbb{F}_2^k$ takes on each value in $\mathbb{F}_2^k$ an equal number of times. Hence

$$\mathop{\mathbf{E}}_{f}[A_\alpha(f)] = \mathop{\mathbf{Pr}}_{f}[A(f) = \alpha] = 1/K \quad \text{for each } \alpha \in \mathbb{F}_2^k. \tag{3}$$

For our preliminary analysis, we will use only this fact to show that for any $\alpha \in \mathbb{F}_2^k$ it holds that

$$\mathop{\mathbf{E}}_{(\gamma,f,g)\in\Psi}[A_\alpha(f)A_\alpha(g)] \le O(1/(K \log K)). \tag{4}$$

It will then follow that the soundness (2) is at most $O(1/\log K)$. Although this tends to 0, it does so only at a rate proportional to the logarithm of the alphabet size, which is $K = 2^k$.

Beginning with the left-hand side of (4), let's write $F = A_\alpha$ for simplicity. We think of the functions f and g being chosen as follows. We first choose a function $h : \gamma^\perp \to \{-1, 1\}$. Note that $\gamma^\perp \subseteq \mathbb{F}_2^k$ is the set of inputs where $\chi_\gamma = 1$ and hence $f = g$, and we let $f(x) = g(x) = h(x)$ for $x \in \gamma_\perp$. The values of f and g on the remaining inputs are chosen independently at random. Then

$$\begin{aligned}\mathop{\mathbf{E}}_{(\gamma,f,g)\in\Psi}[F(f)F(g)] &= \mathop{\mathbf{E}}_{\gamma}\mathop{\mathbf{E}}_{h:\gamma^\perp\to\{-1,1\}}\left[\mathop{\mathbf{E}}_{f,g|h}[F(f)F(g)]\right] \\ &= \mathop{\mathbf{E}}_{\gamma}\mathop{\mathbf{E}}_{h:\gamma^\perp\to\{-1,1\}}\left[\mathop{\mathbf{E}}_{f|h}[F(f)]\mathop{\mathbf{E}}_{g|h}[F(g)]\right]. \end{aligned} \tag{5}$$

Let us write $P_\gamma F(h)$ for $\mathbf{E}_{f|h}\, F(f)$, which is also equal to $\mathbf{E}_{g|h}\, F(g)$. We now use the Fourier expansion of F. Note that the domain here is $\{-1, 1\}^K$ instead of $\mathbb{F}_2^k$. To avoid confusion with characters and Fourier coefficients for functions on $\mathbb{F}_2^k$, we will index the Fourier coefficients below by sets $S \subseteq \mathbb{F}_2^k$. Given an $f \in V$, we'll write f^S for $\prod_{x\in S} f(x)$ (which is a Fourier character for the domain $\{-1, 1\}^K$). Now for fixed γ and h,

$$P_\gamma F(h) = \mathop{\mathbf{E}}_{f|h}[F(f)] = \mathop{\mathbf{E}}_{f|h}\left[\sum_{S\subseteq\mathbb{F}_2^k}\widehat{F}(S)f^S\right] = \sum_{S\subseteq\mathbb{F}_2^k}\widehat{F}(S)\cdot\mathop{\mathbf{E}}_{f|h}[f^S].$$

The quantity $\mathbf{E}_{f|h}[f^S]$ is equal to h^S if $S \subseteq \gamma^\perp$ as is 0 otherwise. Thus, using the Parseval identity, we deduce that (5) equals

$$\mathop{\mathbf{E}}_{\gamma}\mathop{\mathbf{E}}_{h:\gamma^\perp\to\{-1,1\}}\left[(P_\gamma F(h))^2\right] = \mathop{\mathbf{E}}_{\gamma}\left[\sum_{S\subseteq\gamma^\perp}\left(\widehat{F}(S)\right)^2\right] = \sum_{S\subseteq\mathbb{F}_2^k}\mathop{\mathbf{Pr}}_{\gamma}[S \subseteq \gamma^\perp]\cdot\left(\widehat{F}(S)\right)^2.$$

Recalling that $\gamma \in \mathbb{F}_2^k \setminus \{0\}$ is chosen uniformly, we have that

$$\sum_{S\subseteq\mathbb{F}_2^k}\mathop{\mathbf{Pr}}_{\gamma}[S \subseteq \gamma^\perp]\cdot\left(\widehat{F}(S)\right)^2 = \sum_{S\subseteq\mathbb{F}_2^k}2^{-\dim(S)}\cdot\left(\widehat{F}(S)\right)^2,$$

where we are writing $\dim(S) = \dim(\text{span } S)$ for shortness (and defining $\dim(\emptyset) = 0$). For $|S| \geq 1$ we have $\dim(S) \geq \log_2 |S|$ and hence $2^{-\dim(S)} \geq 1/|S|$. Thus

$$\sum_{S \subseteq \mathbb{F}_2^k} 2^{-\dim(S)} \cdot \widehat{F}(S)^2 \leq \widehat{F}(\emptyset)^2 + \sum_{\emptyset \neq S \subseteq \mathbb{F}_2^k} \widehat{F}(S)^2/|S|.$$

Corollary 1 shows that this is at most $O(1/(K \log K))$. This completes the proof:

$$\text{val}(A) \;=\; 2\cdot \sum_{\alpha \in \mathbb{F}_2^k} \mathop{\mathbf{E}}_{(\gamma,f,g)\in\Psi} [A_\alpha(f) A_\alpha(g)] \;\leq\; 2\cdot \sum_{\alpha \in \mathbb{F}_2^k} 2^{-\dim(S)} \widehat{A_\alpha}(S)^2 \;=\; O(1/\log K).$$

5 Integrality gap for 2-to-1 label cover

The instances for 2-to-1 label cover are bipartite. We denote such instances as (U, V, E, R_1, R_2, Π) where $R_2 = 2R_1$ denote the alphabet sizes on the two sides. For a bipartite instance, the label cover SDP can be written in the following form involving vectors $\mathbf{y}_{(u,i)}$ for each $u \in U, i \in [R_1]$ and vectors $\mathbf{z}_{(v,j)}$ for each $v \in v, j \in [R_2]$.

$$\begin{array}{lll}
\text{maximize} & \displaystyle\mathop{\mathbf{E}}_{e=(u,v)\in E}\Big[\sum_{i\in[R_2]} \big\langle \mathbf{y}_{(u,\pi_e(i))}, \mathbf{z}_{(v,j)} \big\rangle\Big] & \\
\text{subject to} & \displaystyle\sum_{i\in[R_1]} \big\|\mathbf{y}_{(u,i)}\big\|^2 = 1 & \forall\, u \in U \\
& \displaystyle\sum_{i\in[R_2]} \big\|\mathbf{z}_{(v,i)}\big\|^2 = 1 & \forall\, v \in V \\
& \big\langle \mathbf{y}_{(u,i)}, \mathbf{y}_{(u,j)} \big\rangle = 0 & \forall\, i \neq j \in [R_1], u \in U \\
& \big\langle \mathbf{z}_{(v,i)}, \mathbf{z}_{(v,j)} \big\rangle = 0 & \forall\, i \neq j \in [R_2], v \in V
\end{array}$$

Fig. 2. SDP for 2-to-1 games

5.1 Gap Instance

As in the case of 2-to-2 games, the set V consists of equivalence classes $\mathcal{P}_1, \ldots, \mathcal{P}_n$, which partition the set of functions $\mathcal{F} = \{f : \mathbb{F}_2^k \to \{-1, 1\}\}$, according to the equivalence relation $\sim$ defined as $f \sim g \;\Leftrightarrow\; \exists \alpha \in \mathbb{F}_2^k \; s.t. f \equiv g\chi_\alpha$. The label set $[R_2]$ is again identified with $\mathbb{F}_2^k$ and is of size $K = 2^k$.

To describe the set U, we further partition the vertices in V according to other equivalence relations. For each $\gamma \in \mathbb{F}_2^k, \gamma \neq 0$, we define an equivalence relation $\cong_\gamma$ on the set $\mathcal{P}_1, \ldots, \mathcal{P}_n$ as

$$\mathcal{P}_i \cong_\gamma \mathcal{P}_j \quad \Leftrightarrow \quad \exists f \in \mathcal{P}_i, g \in \mathcal{P}_j \;\; s.t. \;\; f(1 + \chi_\gamma) \equiv g(1 + \chi_\gamma)$$

This is equivalent to saying:

$$\mathcal{P}_i \cong_\gamma \mathcal{P}_j \qquad \Leftrightarrow \qquad \exists f \in \mathcal{P}_i, g \in \mathcal{P}_j \text{ s.t. } fg(x) = -1 \Rightarrow \chi_\gamma(x) = -1 \ \forall x \in \mathbb{F}_2^k.$$

This partitions $\mathcal{P}_1, \ldots, \mathcal{P}_n$ (and hence also the set $\mathcal{F}$) into equivalence classes $\mathcal{Q}_1^\gamma, \ldots, \mathcal{Q}_m^\gamma$. Here $m = 2^{K/2+1}/K$ (this is immediate from the second definition and the fact that $n = 2^K/K$) and the partition is different for each γ. The set U has one vertex for each class of the form $\mathcal{Q}_i^\gamma$ for all $i \in [m]$ and $\gamma \in \mathbb{F}_2^k \setminus \{0\}$. As before, we denote by $[\mathcal{Q}_i^\gamma]$ the lexicographically smallest function in the class $\mathcal{Q}_i^\gamma$, and by $\mathcal{Q}_f^\gamma$ the class under $\cong_\gamma$ containing f. Note that if $f \in \mathcal{Q}_i^\gamma$, then there exists a $\beta \in \mathbb{F}_2^k$ such that $f(1+\chi_\gamma) \equiv [\mathcal{Q}_i^\gamma]\chi_\beta(1+\chi_\gamma)$.

The label set R_1 has size $K/2$. For each vertex $\mathcal{Q}_i^\gamma \in U$, we think of the labels as pairs of the form $\{\alpha, \alpha+\gamma\}$ for $\alpha \in \mathbb{F}_2^k$. More formally, we identify it with the space $\mathbb{F}_2^k/\langle\gamma\rangle$. We impose one constraint for every pair of the form (γ, f) between the vertices $\mathcal{P}_f$ and $\mathcal{Q}_f^\gamma$. If $f \equiv [\mathcal{P}_f]\chi_\alpha$ and $f(1+\chi_\gamma) \equiv [\mathcal{Q}_i^\gamma]\chi_\beta(1+\chi_\gamma)$, then the corresponding relation $\psi_{(\gamma,f)}$ is defined by requiring that for any labelings $A : V \to [R_2]$ and $B : U \to [R_1]$,

$$(B(\mathcal{Q}_f^\gamma), A(\mathcal{P}_f)) \in \psi_{(\gamma,f)} \quad \Leftrightarrow \quad A(\mathcal{P}_f) + \alpha \ \in \ B(\mathcal{Q}_f^\gamma) + \beta.$$

Here, if $B(\mathcal{Q}_f^\gamma)$ is a pair of the form $\{\omega, \omega+\gamma\}$, then $B(\mathcal{Q}_f^\gamma)+\beta$ denotes the pair $\{\omega+\beta, \omega+\gamma+\beta\}$.

5.2 SDP Value

As before, we give a set of vectors $\mathbf{y}_{(\mathcal{Q}_i^\gamma, \{\alpha,\alpha+\gamma\})}$ and $\mathbf{z}_{(\mathcal{P}_i,\alpha)}$ in $\mathbb{R}^K$, identifying each coordinate with an $x \in \mathbb{F}_2^k$. We define the vectors as

$$\begin{aligned} \mathbf{y}_{(\mathcal{Q}_i^\gamma, \{\alpha,\alpha+\gamma\})}(x) &:= \frac{1}{K}\left([\mathcal{Q}_i^\gamma]\chi_\alpha(1+\chi_\gamma)\right)(x), \\ \mathbf{z}_{(\mathcal{P}_i,\alpha)}(x) &:= \frac{1}{K}\left([\mathcal{P}_f]\chi_\alpha\right)(x). \end{aligned}$$

We have already shown that $\left\langle \mathbf{z}_{(\mathcal{P}_i,\alpha)}, \mathbf{z}_{(\mathcal{P}_i,\beta)} \right\rangle = 0$ for $\alpha \neq \beta$ and $\left\|\mathbf{z}_{(\mathcal{P}_i,\alpha)}\right\|^2 = 1/K$. It again follows by the orthogonality of characters that for disjoint pairs $\{\alpha, \alpha+\gamma\}$ and $\{\beta, \beta+\gamma\}$, the vectors $\mathbf{y}_{(\mathcal{Q}_i^\gamma, \{\alpha,\alpha+\gamma\})}$ and $\mathbf{y}_{(\mathcal{Q}_i^\gamma, \{\beta,\beta+\gamma\})}$ are orthogonal. It is also easy to verify that $\left\|\mathbf{y}_{(\mathcal{Q}_i^\gamma, \{\alpha,\alpha+\gamma\})}\right\|^2 = 2/K$. Hence, the vectors form a feasible solution.

To show that the SDP value is equal to 1, we consider an arbitrary constraint indexed by the pair (γ, f). Let $f \equiv [\mathcal{P}_f]\chi_\alpha$ and $f(1+\chi_\gamma) \equiv [\mathcal{Q}_i^\gamma]\chi_\beta(1+\chi_\gamma)$. Then for any $\omega \in \mathbb{F}_2^k$, this constraint maps the label $\omega+\alpha$ for $\mathcal{P}_f$ to the pair $\{\omega+\beta, \omega+\gamma+\beta\}$ for $\mathcal{Q}_f^\gamma$. Hence, the value of the SDP solution on this constraint is given by

$$\sum_{\omega \in \mathbb{F}_2^k} \left\langle \mathbf{y}_{(\mathcal{Q}_i^\gamma, \{\omega+\beta,\omega+\beta+\gamma\})}, \mathbf{z}_{(\mathcal{P}_i,\alpha+\omega)} \right\rangle$$

We will show that for every ω, $\mathbf{y}_{(\mathcal{Q}_i^\gamma,\{\omega+\beta,\omega+\beta+\gamma\})} = \mathbf{z}_{(\mathcal{P}_i,\alpha+\omega)} + \mathbf{z}_{(\mathcal{P}_i,\alpha+\omega+\gamma)}$. This will complete the proof as the above expression then becomes

$$\sum_{\omega\in\mathbb{F}_2^k} \left\langle \mathbf{z}_{(\mathcal{P}_i,\alpha+\omega)} + \mathbf{z}_{(\mathcal{P}_i,\alpha+\omega+\gamma)}, \mathbf{z}_{(\mathcal{P}_i,\alpha+\omega)} \right\rangle \;=\; \sum_{\omega\in\mathbb{F}_2^k} \left\|\mathbf{z}_{(\mathcal{P}_i,\alpha+\omega)}\right\|^2 \;=\; 1.$$

To show the vector identity, we simply note that for each coordinate x, we have

$$\begin{aligned}
\mathbf{y}_{(\mathcal{Q}_i^\gamma,\{\omega+\beta,\omega+\beta+\gamma\})}(x) &= \frac{1}{K}\left([\mathcal{Q}_i^\gamma]\chi_\beta(1+\chi_\gamma)\right)(x) \\
&= \frac{1}{K}\left(f(1+\chi_\gamma)\right)(x) \\
&= \frac{1}{K}\left([\mathcal{P}_f]\chi_\alpha + [\mathcal{P}_f]\chi_{\alpha+\gamma}\right)(x) \\
&= \mathbf{z}_{(\mathcal{P}_i,\alpha+\omega)}(x) + \mathbf{z}_{(\mathcal{P}_i,\alpha+\omega+\gamma)}(x).
\end{aligned}$$

5.3 Soundness

We now bound the fraction of constraints satisfied by any pair of labelings $A : V \to [K]$ and $B : U \to [K/2]$. Let $\mathbf{1}_{\{\mathcal{E}\}}$ denote the indicator of the event $\mathcal{E}$, and $N(u)$ denote the neighborhood of a vertex $u \in U$. Then, the fraction of constraints satisfied by any assignments A, B, can be bound by an application of Cauchy-Schwarz as

$$\begin{aligned}
\mathrm{val}(A,B) &= \mathop{\mathbf{E}}_{u\in U}\mathop{\mathbf{E}}_{v\in N(u)}\left[\mathbf{1}_{\{\pi_{uv}(A(v))=B(u)\}}\right] \\
&\le \left(\mathop{\mathbf{E}}_{u\in U}\left(\mathop{\mathbf{E}}_{v\in N(u)}\left[\mathbf{1}_{\{\pi_{uv}(A(v))=B(u)\}}\right]\right)^2\right)^{1/2} \\
&= \left(\mathop{\mathbf{E}}_{u\in U}\mathop{\mathbf{E}}_{v_1,v_2\in N(u)}\left[\mathbf{1}_{\{\pi_{uv_1}(A(v_1))=B(u)=\pi_{uv_2}(A(v_2))\}}\right]\right)^{1/2} \\
&\le \left(\mathop{\mathbf{E}}_{u\in U}\mathop{\mathbf{E}}_{v_1,v_2\in N(u)}\left[\mathbf{1}_{\{\pi_{uv_1}(A(v_1))=\pi_{uv_2}(A(v_2))\}}\right]\right)^{1/2}
\end{aligned}$$

Note that if π_{uv_1} and π_{uv_2} are 2-to-1 projections, then the inner quantity in the last expression denotes the value of a 2-to-2 label cover instance, each of whose constraints is defined by two 2-to-1 constraints in the original instance. For the 2-to-1 instance described above, we will show that the inner quantity in fact denotes the fraction of constraints satisfied by A for the 2-to-2 instance described in Section 4. This will show that the fraction of constraints satisfied by any assignment in the above 2-to-1 instance can be at most $O(1/\sqrt{\log K})$.

To see this, note that a vertex $u \in U$ and a vertex $v_1 \in V$ can be sampled jointly by picking a pair (γ, f) and taking $u = \mathcal{Q}_f^\gamma$ and $v_1 = \mathcal{P}_f$. Sampling $v_2 \in N(u)$ corresponds to choosing a class $\mathcal{P}_i$ such that for some $\beta \in \mathbb{F}_2^k$ $[\mathcal{P}_i]\chi_\beta(1+\chi_\gamma) \equiv f(1+\chi_\gamma)$. Thus, v_2 can be sampled by choosing a random g such that $f(1+\chi_\gamma) \equiv g(1+\chi_\gamma)$ and taking $v_2 = \mathcal{P}_g$.

Also, if $f \equiv [\mathcal{P}_f]\chi_{\alpha_1}$ and $g \equiv [\mathcal{P}_g]\chi_{\alpha_2}$, then the constraint $\pi_{uv_1}(A(v_1)) = \pi_{uv_2}(A(v_2))$ simply requires that for some $\omega \in \mathbb{F}_2^k$, $A(\mathcal{P}_f) + \alpha_1$ and $A(\mathcal{P}_g) + \alpha_2$ both lie in the set $\{\omega, \omega + \gamma\}$ and hence

$$(A(\mathcal{P}_f) + \alpha_1) - (A(\mathcal{P}_g) + \alpha_2) \in \{0, \gamma\}.$$

6 From 2-to-1 constraints to α-constraints

In this section we show that any integrality gap instance for 2-to-1 games, with sufficiently many edges, can be converted to an integrality gap instance for games with α-constraints. The SDP we consider for these games is identical to the ones considered before, except for the objective function.

Theorem 3. *Let $\mathcal{L} = (U, V, E, R, 2R, \Psi)$ be a bipartite instance of 2-to-1 label cover problem with $\mathrm{OPT}(\mathcal{L}) \leq \delta$ and SDP value 1. Also, let $|E| \geq 4(|U| + |V|)\log(R)/\epsilon^2$. Then there exists another instance $\mathcal{L}' = (U, V, E, 2R, \Psi')$ of Label Cover with α-constraints having SDP value 1 and $\mathrm{OPT}(\mathcal{L}') \leq \delta + \epsilon + 1/R$.*

Proof. The proof simply follows by adding R "fake" labels for each vertex $u \in U$, and then randomly augmenting the constraints to make them of the required form. In particular, let the new labels we add for each $u \in U$ be $R+1, \ldots, 2R$. Let $e = (u, v)$ be an edge. Since the constraints in Ψ are 2-to-1 type, there exist permutations $\sigma_{1,e} : [R] \to [R]$ and $\sigma_{2,e} : [2R] \to [2R]$ such that after permuting the labels on each side, the projection π_e maps labels $(2i-1, 2i)$ to i i.e. $\pi_e(\sigma_{2,e}^{-1}(2i-1)) = \pi_e(\sigma_{2,e}^{-1}(2i)) = \sigma_{1,e}^{-1}(i)$.

To incorporate the new labels into the constraint, choose a random bijection $\sigma'_{1,e} : \{R+1, \ldots, 2R\} \to [R]$. We now construct a new permutation $\tilde{\sigma}_{1,e} : [2R] \to [2R]$ as $\tilde{\sigma}_{1,e}(i) = 2\sigma_{1,e}(i) - 1$ if $i \leq R$ and $\tilde{\sigma}_{1,e}(i) = 2\sigma'_{1,e}(i)$ if $i > R$ i.e. the new labels are mapped to the even positions $2, 4, \ldots, 2R$ while the others are mapped to the odd positions.

The original 2-to-1 constraints are satisfied by a labeling A iff the pair $(\tilde{\sigma}_{1,e}(A(u), \sigma_{2,e}(A(v)))$ is of the form $(2i-1, 2i-1)$ or $(2i-1, 2i)$ for some $i \leq R$. We augment the constraint by also allowing $(\tilde{\sigma}_{1,e}(A(u), \sigma_{2,e}(A(v)))$ to be $(2i, 2i-1)$ for some i. Note that if the constraint is satisfied in this way, then u must get one of the new labels. Also, note that the augmentation is random as we choose the map $\sigma'_{1,e}$ independently at random for each edge e.

Given a vector solution $\{\mathbf{y}_{(u,i)}\}_{u \in U, i \in [R]}$ and $\{\mathbf{z}_{(v,j)}\}_{v \in V, j \in [2R]}$ for Ψ, we leave the vectors $\mathbf{z}_{(v,j)}$ unchanged and for each $u \in U$, take $\mathbf{z}_{(u,i)} = \mathbf{y}_{(v,i)}$ if $i \leq R$ and 0 otherwise. It is immediate that the solution is feasible. Also, the value of the objective is the same as the value of the 2-to-1 SDP, as all the additional terms in the objective involve some vector $\mathbf{z}_{(u,i)}$ for some $i > R$ and are hence 0. Thus, the SDP value for the new instance is 1.

To bound the optimal value of any labeling $A : U \cup V \to [2R]$, we split it as

$$\begin{aligned}\mathop{\mathbf{E}}_{e=(u,v)\in E}\left[\mathbf{1}_{\{(A(u),A(v)) \text{ satisfy } e\}}\right] &= \mathop{\mathbf{E}}_{e=(u,v)\in E}\left[\mathbf{1}_{\{A(u)\leq R\}} \cdot \mathbf{1}_{\{(A(u),A(v)) \text{ satisfy } e\}}\right] \\ &+ \mathop{\mathbf{E}}_{e=(u,v)\in E}\left[\mathbf{1}_{\{A(u)> R\}} \cdot \mathbf{1}_{\{(A(u),A(v)) \text{ satisfy } e\}}\right]\end{aligned}$$

Note that the first term is simply the number of 2-to-1 constraints satisfied by A and it at most δ by assumption.

Also, for any fixed labeling A, the probability over the choice of the random maps $\{\sigma'_{1,e}\}_{e\in E}$, that $(A(u), A(v))$ satisfy e given that $A(u) > R$, is at most $1/R$. By a Chernoff bound, the fraction of edges (u, v) satisfied with $A(u) > R$ is at most $1/R+\epsilon$ with probability $\exp(-\epsilon^2|E|/3)$ over the choice of the random maps. By a union bound and the condition on ϵ, the second term is at most $1/R + \epsilon$ for all labelings A, with high probability over the choice of $\{\sigma'_{1,e}\}_{e\in E}$. Picking an instance with appropriate choice of maps $\sigma'_{1,e}$ gives the required instance $\mathcal{L}'$.

7 Discussion

The instances we construct have SDP value 1 only for the most basic semidefinite programming relaxation. It would be desirable to get gaps for stronger SDPs, beginning with the most modest extensions of this basic SDP. For example, in the SDP for 2-to-1 Label Cover from Figure 2, we can add valid nonnegativity constraints for the dot product between every pair of vectors in the set

$$\{\mathbf{y}_{(u,i)} \mid u \in U, i \in [R_1]\} \bigcup \{\mathbf{z}_{(v,j)} \mid v \in V, j \in [R_2]\} \ ,$$

since in the integral solution all these vectors are $\{0, 1\}$-valued. The vectors we construct do *not* obey such a nonnegativity requirement. For the case of Unique Games, Khot and Vishnoi [10] were able to ensure nonnegativity of all dot products by simply taking tensor products of the vectors with themselves and defining new vectors $\mathbf{y}'_{(u,i)} = \mathbf{y}^{\otimes 2}_{(u,i)} = \mathbf{y}_{(u,i)} \otimes \mathbf{y}_{(u,i)}$ and $\mathbf{z}'_{(v,j)} = \mathbf{z}^{\otimes 2}_{(v,j)} = \mathbf{z}_{(v,j)} \otimes \mathbf{z}_{(v,j)}$. Since $\langle \mathbf{a}^{\otimes 2}, \mathbf{b}^{\otimes 2}\rangle = \langle \mathbf{a}, \mathbf{b}\rangle^2$, the desired nonnegativity of dot products is ensured.

We cannot apply this tensoring idea in our construction as it does not preserve the SDP value at 1. For example, for 2-to-1 Label Cover, if we have $\mathbf{y}_{(u,i)} = \mathbf{z}_{(v,j_1)} + \mathbf{z}_{(v,j_2)}$ (so that these vectors contribute 1 to the objective value to the SDP of Figure 2), then upon tensoring we no longer necessarily have $\mathbf{y}^{\otimes 2}_{(u,i)} = \mathbf{z}^{\otimes 2}_{(v,j_1)} + \mathbf{z}^{\otimes 2}_{(v,j_2)}$. Extending our gap instances to obey the nonnegative dot product constraints is therefore a natural question that we leave open. While this seems already quite challenging, one can of course be more ambitious and ask for gap instances for stronger SDPs that correspond to certain number of rounds of some hierarchy, such as the Sherali-Adams hierarchy together with consistency of vector dot products with pairwise marginals. For Unique Games, gap instances for several rounds of such a hierarchy were constructed in [13].

Acknowledgments

Madhur would like to thank Per Austrin for helpful discussions in the early stages of this work. We also thank the anonymous reviewers for their suggestions.

References

1. Bulatov, A., Dalmau, V.: A simple algorithm for Mal'tsev constraints. SICOMP 36(1), 16–27 (2006)
2. Charikar, M., Makarychev, K., Makarychev, Y.: Near-optimal algorithms for unique games. In: Proc. 38th ACM STOC. pp. 205–214 (2006)
3. Dinur, I., Mossel, E., Regev, O.: Conditional hardness for approximate coloring. SICOMP 39(3), 843–873 (2009)
4. Dinur, I., Safra, S.: On the hardness of approximating minimum vertex cover. Ann. Math. 162(1), 439–486 (2005)
5. Guruswami, V., Manokaran, R., Raghavendra, P.: Beating the random ordering is hard: Inapproximability of maximum acyclic subgraph. In: Proc. 49th IEEE FOCS. pp. 573–582 (2008)
6. Guruswami, V., Sinop, A.K.: Improved inapproximability results for maximum k-colorable subgraph. In: Proc. 12th APPROX. pp. 163–176 (2009)
7. Khot, S.: Hardness results for coloring 3-colorable 3-uniform hypergraphs. In: Proc. 43rd IEEE FOCS. pp. 23–32 (2002)
8. Khot, S.: On the power of unique 2-prover 1-round games. In: Proc. 34th ACM STOC. pp. 767–775 (2002)
9. Khot, S., Saket, R.: SDP integrality gaps with local ℓ_1-embeddability. In: Proc. 50th IEEE FOCS. pp. 565–574 (2009)
10. Khot, S., Vishnoi, N.: The Unique Games Conjecture, integrality gap for cut problems and embeddability of negative type metrics into ℓ_1. In: Proc. 46th IEEE FOCS. pp. 53–62 (2005)
11. O'Donnell, R., Wu, Y.: Conditional hardness for satisfiable CSPs. In: Proc. 41st ACM STOC. pp. 493–502 (2009)
12. Raghavendra, P.: Optimal algorithms and inapproximability results for every CSP? In: Proc. 40th ACM STOC. pp. 245–254 (2008)
13. Raghavendra, P., Steurer, D.: Integrality gaps for strong SDP relaxations of unique games. In: Proc. 50th IEEE FOCS. pp. 575–585 (2009)
14. Schoenebeck, G.: Linear level Lasserre lower bounds for certain k-CSPs. In: Proc. 49th IEEE FOCS. pp. 593–602 (2008)
15. Talagrand, M.: On Russo's approximate zero-one law. Ann. Prob. 22(3), 1576–1587 (1994)
16. Tang, L.: Conditional hardness of approximating satisfiable Max 3CSP-q. In: Proc. 20th ISAAC. pp. 923–932 (2009)
17. Tulsiani, M.: CSP gaps and reductions in the Lasserre hierarchy. In: Proc. 41st ACM STOC. pp. 303–312 (2009)