Conditional Hardness for Approximate Coloring

Irit Dinur[*]   Elchanan Mossel[†]   Oded Regev[‡]

November 3, 2005

Abstract

We study the APPROXIMATE-COLORING(q, Q) problem: Given a graph G, decide whether $\chi(G) \leq q$ or $\chi(G) \geq Q$ (where $\chi(G)$ is the chromatic number of G). We derive conditional hardness for this problem for any constant $3 \leq q < Q$. For $q \geq 4$, our result is based on Khot's 2-to-1 conjecture [Khot'02]. For $q = 3$, we base our hardness result on a certain '$\bowtie$ shaped' variant of his conjecture.

We also prove that the problem ALMOST-3-COLORING$_\varepsilon$ is hard for any constant $\varepsilon > 0$, assuming Khot's Unique Games conjecture. This is the problem of deciding for a given graph, between the case where one can 3-color all but a ε fraction of the vertices without monochromatic edges, and the case where the graph contains no independent set of relative size at least ε.

Our result is based on bounding various generalized noise-stability quantities using the invariance principle of Mossel et al [MOO'05].

1 Introduction

For a graph $G = (V, E)$ we let $\chi(G)$ be the chromatic number of G, i.e., the smallest number of colors needed to color the vertices of G without monochromatic edges. We study the following problem,

APPROXIMATE-COLORING(q,Q) : Given a graph G, decide between $\chi(G) \le q$ and $\chi(G) \ge Q$.

The problem APPROXIMATE-COLORING$(3, Q)$ is notorious for the wide gap between the value of Q for which an efficient algorithm is known and that for which a hardness result exists. The best known polynomial-time algorithm solves the problem for $Q = \tilde{O}(n^{3/14})$ colors, where n is the number of vertices [3] (In fact, that algorithm solves the *search problem* of finding a Q-coloring given a q-colorable graph. Since we are mainly interested in hardness results, we restrict our attention to the decision version of the problem). In contrast, the strongest known hardness result shows that the problem is NP-hard for $Q = 5$ [12, 8]. Thus, the problem is open for all $5 < Q < \tilde{O}(n^{3/14})$. In this paper we give some evidence that this problem is hard for any constant value of Q. We remark that any hardness result for $q = 3$ immediately carries over for all $q > 3$.

The best algorithm known for larger values of q is due to Halperin et al. [9], improving on a previous result of Karger et al [11]. Their algorithm solves APPROXIMATE-COLORING(q, Q) for $Q = n^{\alpha_q}$ where $0 < \alpha_q < 1$ is some function of q. For example, $\alpha_4 \approx 0.37$. Improving on an earlier result of Fürer [7], Khot has shown [13] that for any large enough constant q and $Q = q^{\frac{\log q}{25}}$, APPROXIMATE-COLORING$(q, Q)$ is NP-hard. A related problem is that of approximating the chromatic number $\chi(\cdot)$ of a given graph. For this problem, an inapproximability result of $n^{1-o(1)}$ is known [6, 13].

Constructions: Our constructions follow the standard composition paradigm initiated in [2, 10], which has yielded numerous inapproximability results by now. In our context, this means that we show reductions from variants of a problem known as label-cover to approximate graph coloring problems. In the label-cover problem, we are given an undirected graph and a number R. Each edge is associated with a binary relation on $\{1, \ldots, R\}$ and we refer to it as a *constraint*. The goal is to label the vertices with values from $\{1, \ldots, R\}$ such that the number of satisfied constraints is maximized, where a constraint is satisfied if its two incident vertices satisfy the relation associated with it.

As is the case with other composition-based reductions, our reductions work by replacing each vertex of the label-cover instance with a block of vertices, known as a *gadget*. In other reductions, the gadget is often the binary hypercube $\{0, 1\}^R$, sometimes known as the long-code. In our case, the gadget is the q-ary hypercube, $\{1, \ldots, q\}^R$. We then connect the gadgets in a way that "encodes" the label-cover constraints. The idea is to ensure that any Q-coloring of the graph (where Q is some constant greater than q), can be "decoded" into a labeling for the underlying label-cover instance that satisfies many label-cover constraints.

We note that the idea of using the q-ary hypercube as a gadget has been around for a number of years. This idea has been studied in [1] and some partial results were obtained. The recent progress of [17] has provided the necessary tool for achieving our result.

Conjectures: Let us now turn our attention to the label-cover problem. None of the known NP-hard variants of the label-cover problem (or even more general PCP systems) seems suitable for composition in our setting. An increasingly popular approach is to rely on the 'Unique-Games' conjecture of Khot [14]. The conjecture states that a very restricted version of label-cover is hard. The strength of this restriction is that in a sense, it reduces the analysis of the entire construction, to the analysis of the gadget alone.

However, this conjecture suffers from inherent imperfect completeness, which seems to prevent it from being used in a reduction to approximate coloring (although it is useful for *almost* approximate coloring). Therefore, we consider some restrictions of label-cover that do have perfect completeness. Our approach is

to search for the least-restricted such label-cover problem that would still yield the desired result. In all, we consider three variants, which result in three different reductions.

- We show that ALMOST-3-COLORING is as hard as Khot's Unique Games problem.

- We show that APPROXIMATE-COLORING$(4, Q)$ is as hard as Khot's 2-to-1 problem for any constant $Q > 4$. This also holds for APPROXIMATE-COLORING(q, Q) for any $q \geq 4$.

- We introduce a new conjecture, which states that label-cover is hard when the constraints are restricted to a form we call $\bowtie$-constraints (read: alpha constraints). We show that for any constant $Q > 3$, APPROXIMATE-COLORING$(3, Q)$ is as hard as solving the $\bowtie$-label-cover problem. We remark that $\bowtie$-constraints have already appeared in [5].

The plausibility of the Unique Games Conjecture, as well as that of other variants, is uncertain. Recently, Trevisan [18] showed that these conjectures are false when the parameters are pushed beyond certain sub-constant values. This has been followed by additional algorithms [**?**, **?**] for various sub-constant settings. Hopefully, these results will trigger more attempts to understand these type of constraint systems from both the algorithmic side, and the inapproximability side.

Techniques: Our main technique is based on the recent progress of Mossel et al [17]. There, they present a powerful technique for bounding the stability of functions under noise operators. For example, one special case of their result says that among all balanced Boolean functions that do not depend too strongly on any one coordinate, the one that is most stable under noise is the majority function. In other words, among all such functions, the majority function is least likely to flip its value if we flip each of its input bits with some small constant probability. In fact, this special case was presented as a conjecture in the work of [15] on MAXCUT and motivated the result of [17].

The technique of [17] is based on what is called an *invariance principle*. This principle allows one to translate questions in the discrete setting (such as the above question on the Boolean hypercube) to corresponding questions in other spaces, and in particular Gaussian space. One then applies known (and powerful) results in Gaussian space.

In this paper we extend their approach is several respects.

- We consider more general noise operators that are given by some arbitrary Markov operator. We then apply this to three operators, one for each of the aforementioned constructions.

- We show that when the inner product under noise of two functions deviates notably from that of two majority functions, there must exist a variable that is influential in *both* functions (see Theorem 3.1). A direct application of [17] only yields a variable that is influential in *one* of the functions. This latter statement was enough for the application to MAXCUT in [15].

- We also present another result tailored for the $\bowtie$ constraints (see Theorem B.2).

We believe that the general framework developed here will find many applications. It also demonstrates the flexibility of the approach of [17].

Future work: Our constructions can be extended in several ways. First, using similar techniques, one can show hardness of APPROXIMATE-COLORING(q, Q) based on the d-to-1 conjecture of Khot for larger values of d (and not only $d = 2$ as we do here). It would be interesting to find out how q depends on d. Second, by strengthening the current conjectures to sub-constant values, one can obtain hardness for Q that depends

on n, the number of vertices in the graph. Again, it is interesting to see how large Q can be. Finally, let us mention that in all our reductions we in fact show in the soundness case that there are no independent sets of relative size larger than ε for arbitrarily small constant ε (note that this is somewhat stronger than showing that there is no Q-coloring). In fact, a more careful analysis can be used to obtain the stronger statement that are no 'almost-independent' sets of relative size larger than ε.

2 Preliminaries

2.1 Approximate coloring problems

We now define the coloring problems that we study in this paper. For any graph G, let $\chi(G)$ be its chromatic number. Namely, $\chi(G)$ is the smallest number of colors needed in order to color the vertices of G without monochromatic edges. For any $3 \le q < Q$, we define

APPROXIMATE-COLORING(q, Q): Given a graph G, decide between $\chi(G) \le q$ or $\chi(G) \ge Q$.

For any $\varepsilon > 0$ we define

ALMOST-3-COLORING$_\varepsilon$: Given a graph $G = (V, E)$, decide between
- There exists a set $V' \subseteq V$, $|V'| \ge (1-\varepsilon)|V|$ such that $\chi(G|_{V'}) = 3$ where $G|_{V'}$ is the graph induced by V'.
- Every independent set $S \subseteq V$ in G has size $|S| \le \varepsilon|V|$.

Observe that these two items are mutually exclusive for $\varepsilon < 1/4$.

2.2 Functions on the q-ary hypercube

Let $[q]$ denote the set $\{0, \ldots, q-1\}$. For an element x of $[q]^n$ write $|x|_a$ for the number of coordinates k of x such that $x_k = a$ and $|x| = \sum_{a \ne 0} |x|_i$ for the number of nonzero coordinates.

In this paper we are interested in functions from $[q]^n$ to $\mathbb{R}$. We define an inner product on this space by $\langle f, g \rangle = \frac{1}{q^n} \sum_x f(x) g(x)$. In our applications, we usually take q to be some constant (say, 3) and n to be large.

Definition 2.1 *Let $f : [q]^n \to \mathbb{R}$ be a function. The influence of the i'th variable on f, denoted $I_i(f)$ is defined by*
$$I_i(f) = \mathbf{E}[\mathbf{V}_{x_i}[f(x)|x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n]]$$
where $x_1, \ldots, x_n$ are uniformly distributed.

Consider a sequence of vectors $\alpha_0 = 1, \alpha_1, \ldots, \alpha_{q-1} \in \mathbb{R}^q$ forming an orthonormal basis of $\mathbb{R}^q$. Equivalently, we can think of these vectors as functions from $[q]$ to $\mathbb{R}$. These vectors can be used to form an orthonormal basis of the space of functions from $[q]^n$ to $\mathbb{R}$, as follows.

Definition 2.2 *Let $\alpha_0 = 1, \alpha_1, \ldots, \alpha_{q-1}$ be an orthonormal basis of $\mathbb{R}^q$. For $x \in [q]^n$, write α_x for*
$$\alpha_{x_1} \otimes \alpha_{x_2} \otimes \cdots \otimes \alpha_{x_n}.$$
Equivalently, we can define α_x as the function mapping $y \in [q]^n$ to $\alpha_{x_1}(y_1)\alpha_{x_2}(y_2)\cdots\alpha_{x_n}(y_n)$.

Clearly, any function in $[q]^n \to \mathbb{R}$ can be written as a linear combination of α_x for $x \in [q]^n$. This leads to the following definition.

Definition 2.3 *For a function $f : [q]^n \to \mathbb{R}$, define $\hat{f}(\alpha_x) = \langle f, \alpha_x \rangle$ and notice that $f = \sum_x \hat{f}(\alpha_x)\alpha_x$.*

Claim 2.4 *For any function $f : [q]^n \to \mathbb{R}$ and any $i \in \{1, \ldots, n\}$,*

$$I_i(f) = \sum_{x : x_i \neq 0} \hat{f}^2(\alpha_x).$$

We include the proof in Appendix F. Notice that this claim holds for any choice of orthonormal basis $\alpha_0, \ldots, \alpha_{q-1}$ as long as $\alpha_0 = 1$.

Definition 2.5 *Let $f : [q]^n \to \mathbb{R}$ be a function, and let $k \leq n$. The low-level influence of the i'th variable on f is defined by*

$$I_i^{\leq k}(f) = \sum_{x : x_i \neq 0, |x| \leq k} \hat{f}^2(\alpha_x).$$

It is easy to see that for any function f, $\sum_i I_i^{\leq k}(f) = \sum_{x : |x| \leq k} \hat{f}^2(\alpha_x) |x| \leq k \sum_x \hat{f}^2(\alpha_x) = k\|f\|_2^2$. In particular, for any function f obtaining values in $[0, 1]$, $\sum_i I_i^{\leq k}(f) \leq k$. Moreover, let us mention that $I_i^{\leq k}$ is in fact independent of the particular choice of basis $\alpha_0, \alpha_1, \ldots, \alpha_{q-1}$ as long as $\alpha_0 = 1$. This can be verified from the above definition.

There is a natural equivalence between $[q]^{2n}$ and $[q^2]^n$. As this equivalence is used often in this paper, we introduce the following notation.

Definition 2.6 *For any $x \in [q]^{2n}$ we denote by $\overline{x}$ the element of $[q^2]^n$ given by $\overline{x} = ((x_1, x_2), \ldots, (x_{2n-1}, x_{2n}))$. For any $y \in [q^2]^n$ we denote by $\underline{y}$ the element of $[q]^{2n}$ given by $\underline{y} = (y_{1,1}, y_{1,2}, y_{2,1}, y_{2,2}, \ldots, y_{n,1}, y_{n,2})$. For a function f on $[q]^{2n}$ we denote by $\overline{f}$ the function on $[q^2]^n$ defined by $\overline{f}(y) = f(\underline{y})$. Similarly, for a function f on $[q^2]^n$ we denote by $\underline{f}$ the function on $[q]^{2n}$ defined by $\underline{f}(x) = f(\overline{x})$.*

Claim 2.7 *For any function $f : [q]^{2n} \to \mathbb{R}$, any $i \in \{1, \ldots, n\}$, and any $k \geq 1$,*

$$I_i^{\leq k}(\overline{f}) \leq I_{2i-1}^{\leq 2k}(f) + I_{2i}^{\leq 2k}(f).$$

Proof: Fix some basis α_x of $[q]^{2n}$ as above and let $\alpha_{\overline{x}}$ be the basis of $[q^2]^n$ defined by $\alpha_{\overline{x}}(\overline{y}) = \alpha_x(y)$. Then, it is easy to see that $\hat{\overline{f}}(\alpha_{\overline{x}}) = \hat{f}(\alpha_x)$. Hence,

$$I_i^{\leq k}(\overline{f}) = \sum_{\overline{x} : \overline{x}_i \neq (0,0), |\overline{x}| \leq k} \hat{\overline{f}}^2(\alpha_{\overline{x}}) \leq \sum_{x : x_{2i-1} \neq 0, |x| \leq 2k} \hat{f}^2(\alpha_x) + \sum_{x : x_{2i} \neq 0, |x| \leq 2k} \hat{f}^2(\alpha_x) = I_{2i-1}^{\leq 2k}(f) + I_{2i}^{\leq 2k}(f)$$

where we used that $|x| \leq 2|\overline{x}|$. ✄

For the following definition, recall that we say that a Markov operator T is *symmetric* if it is reversible with respect to the uniform distribution, i.e., if the transition matrix representing T is symmetric.

Definition 2.8 *Let T be a symmetric Markov operator on $[q]$. Let $1 = \lambda_0 \geq \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{q-1}$ be the eigenvalues of T. We define $r(T)$, the spectral radius of T, by $r(T) = \max\{|\lambda_1|, |\lambda_{q-1}|\}$.*

For T as above, we may define a Markov operator $T^{\otimes n}$ on $[q]^n$ in the standard way. Note that if T is symmetric then $T^{\otimes n}$ is also symmetric and $r(T^{\otimes n}) = r(T)$. If we choose $\alpha_0, \ldots, \alpha_{q-1}$ to be an orthonormal set of eigenvectors for T with corresponding eigenvalues $\lambda_0, \ldots, \lambda_{q-1}$ (so $\alpha_0 = 1$), we see that

$$T^{\otimes n}\alpha_x = \left(\prod_{a \neq 0} \lambda_a^{|x|_a}\right) \alpha_x.$$

4

and hence

$$T^{\otimes n} f = \sum_x \left(\prod_{a \neq 0} \lambda_a^{|x|_a} \right) \hat{f}(\alpha_x) \alpha_x.$$

holds for any function $f : [q]^n \to \mathbb{R}$.

We now describe two operators that we use in this paper. The first is the Beckner operator, T_ρ. For any $\rho \in [-\frac{1}{q-1}, 1]$, it is defined by $T_\rho(x \to x) = \frac{1}{q} + (1 - \frac{1}{q})\rho$ and $T_\rho(x \to y) = \frac{1}{q}(1 - \rho)$ for any $x \neq y$ in $[q]$. It can be seen that T_ρ is a Markov operator as in Definition 2.8 with $\lambda_1 = \ldots = \lambda_{q-1} = \rho$ and hence its spectral radius is $|\rho|$.

Another useful operator is the averaging operator, A_S. For a subset $S \subseteq \{1, \ldots, n\}$, it acts on functions on $[q]^n$ by averaging over coordinates in S, namely, $A_S(f) = \mathbf{E}_{x_S}[f]$. Notice that the function $A_S(f)$ is independent of the coordinates in S.

2.3 Functions in Gaussian space

We let γ denote the standard Gaussian measure on $\mathbb{R}^n$. We denote by $\mathbf{E}_\gamma$ the expected value with respect to γ and by $\langle \cdot, \cdot \rangle_\gamma$ the inner product on $L^2(\mathbb{R}^n, \gamma)$. Notice that $\mathbf{E}_\gamma[f] = \langle f, \mathbf{1} \rangle_\gamma$ where $\mathbf{1}$ is the constant 1 function. For $\rho \in [-1, 1]$, we denote by U_ρ the *Ornstein-Uhlenbeck operator*, which acts on $L^2(\mathbb{R}, \gamma)$ by

$$U_\rho f(x) = \mathbf{E}_{y \sim \gamma}[f(\rho x + \sqrt{1 - \rho^2} y)].$$

Finally, for $0 < \mu < 1$, let $F_\mu : \mathbb{R} \to \{0, 1\}$ denote the function $F_\mu(x) = 1_{x < t}$ where t is chosen in such a way that $\mathbf{E}_\gamma[F_\mu] = \mu$. One useful value that will appear later is $\langle F_\eta, U_\rho F_\nu \rangle_\gamma$. For our purposes it is useful to know that for any $\nu, \eta > 0$, and any $\rho \in [-1, 1]$, it holds that

$$\langle F_\tau, U_\rho F_\tau \rangle_\gamma \leq \langle F_\eta, U_\rho F_\nu \rangle_\gamma \leq \tau,$$

where $\tau = \min(\eta, \nu)$. Moreover, for all $\tau > 0$ and $\rho > -1$ it holds that

$$\langle F_\tau, U_\rho F_\tau \rangle_\gamma > 0.$$

3 An Inequality for Noise Operators

The main result of this section, Theorem 3.1, is a generalization of the result of [17]. It shows that if the inner product of two functions f and g under some noise operator deviates from a certain range then there must exist an index i such that the low-level influence of the ith variable is large in both f and g. This range depends on the expected value of f and g, and the spectral radius of the operator T.

Theorem 3.1 *Let q be a fixed integer and let T be a symmetric Markov operator on $[q]$ such that $\rho = r(T) < 1$. Then for any $\varepsilon > 0$ there exist $\delta > 0$ and $k \in \mathbb{N}$ such that if $f, g : [q]^n \to [0, 1]$ are two functions satisfying $\mathbf{E}[f] = \mu, \mathbf{E}[g] = \nu$ and*

$$\min \left(I_i^{\leq k}(f), I_i^{\leq k}(g) \right) < \delta$$

for all i, then it holds that

$$\langle f, T^{\otimes n} g \rangle \geq \langle F_\mu, U_\rho(1 - F_{1-\nu}) \rangle_\gamma - \varepsilon \tag{1}$$

and

$$\langle f, T^{\otimes n} g \rangle \leq \langle F_\mu, U_\rho F_\nu \rangle_\gamma + \varepsilon. \tag{2}$$

Note that (1) follows from (2). Indeed, apply (2) to $1-g$ to obtain $\langle f, T^{\otimes n}(1-g)\rangle \leq \langle F_\mu, U_\rho F_{1-\nu}\rangle_\gamma + \varepsilon$ and then use

$$\langle f, T^{\otimes n}(1-g)\rangle = \langle f, 1\rangle - \langle f, T^{\otimes n}g\rangle = \mu - \langle f, T^{\otimes n}g\rangle = \langle F_\mu, U_\rho 1\rangle_\gamma - \langle f, T^{\otimes n}g\rangle.$$

From now on we focus on proving (2).

Following the approach of [17], the proof consists of two powerful techniques. The first is an inequality by Christer Borell [4] on continuous Gaussian space. The second is an invariance principle shown in [17] that allows us to translate our discrete question to the continuous Gaussian space.

Definition 3.2 (Gaussian analogue of an operator) *Let T be an operator as in Definition 2.8. We define its Gaussian analogue as the operator $\tilde{T}$ on $L^2(\mathbb{R}^{q-1}, \gamma)$ given by $\tilde{T} = U_{\lambda_1} \otimes U_{\lambda_2} \otimes \ldots \otimes U_{\lambda_{q-1}}$.*

For example, the Gaussian analogue of T_ρ is $U_\rho^{\otimes(q-1)}$. We need the following powerful theorem by Borell [4]. It says that the functions that maximize the inner product under the operator U_ρ are the indicator functions of half-spaces.

Theorem 3.3 (Borell [4]) *Let $f, g : \mathbb{R}^n \to [0, 1]$ be two functions and let $\mu = \mathbf{E}_\gamma[f], \nu = \mathbf{E}_\gamma[g]$. Then $\langle f, U_\rho^{\otimes n}g\rangle_\gamma \leq \langle F_\mu, U_\rho F_\nu\rangle_\gamma$.*

The above theorem only applies to the Ornstein-Uhlenbeck operator. In the following corollary we derive a similar statement for more general operators. The proof follows by writing a general operator as a product of the Ornstein-Uhlenbeck operator and some other operator.

Corollary 3.4 *Let $f, g : \mathbb{R}^{(q-1)n} \to [0, 1]$ be two functions satisfying $\mathbf{E}_\gamma[f] = \mu, \mathbf{E}_\gamma[g] = \nu$. Let T be an operator as in Definition 2.8 and let $\rho = r(T)$. Then $\langle f, \tilde{T}^{\otimes n}g\rangle_\gamma \leq \langle F_\mu, U_\rho F_\nu\rangle_\gamma$.*

Proof: For $1 \leq i \leq q-1$, let $\delta_i = \lambda_i/\rho$. Note that $|\delta_i| \leq 1$ for all i. Let S be the operator defined by $S = U_{\delta_1} \otimes U_{\delta_2} \otimes \ldots \otimes U_{\delta_{q-1}}$. Then,

$$U_\rho^{\otimes(q-1)}S = U_\rho U_{\delta_1} \otimes \ldots \otimes U_\rho U_{\delta_{q-1}} = U_{\rho\delta_1} \otimes \ldots \otimes U_{\rho\delta_{q-1}} = \tilde{T}$$

(this is often called the *semi-group property*). It follows that $\tilde{T}^{\otimes n} = U_\rho^{\otimes(q-1)n}S^{\otimes n}$. The function $S^{\otimes n}g$ obtains values in $[0, 1]$ and satisfies $\mathbf{E}_\gamma[S^{\otimes n}g] = \mathbf{E}_\gamma[g]$. Thus the claim follows by applying Theorem 3.3 to the functions f and $S^{\otimes n}g$. ✎

Definition 3.5 (Real analogue of a function) *Let $f : [q]^n \to \mathbb{R}$ be a function with decomposition $f = \sum \hat{f}(\alpha_x)\alpha_x$. Consider the $(q-1)n$ variables $z_1^1, \ldots, z_{q-1}^1, \ldots, z_1^n, \ldots, z_{q-1}^n$ and let $\Gamma_x = \prod_{i=1, x_i \neq 0}^n z_{x_i}^i$. We define the real analogue of f to be the function $\tilde{f} : \mathbb{R}^{n(q-1)} \to \mathbb{R}$ given by $\tilde{f} = \sum \hat{f}(\alpha_x)\Gamma_x$.*

Claim 3.6 *For any two functions $f, g : [q]^n \to \mathbb{R}$ and operator T on $[q]^n$, $\langle f, g\rangle = \langle \tilde{f}, \tilde{g}\rangle_\gamma$ and $\langle f, T^{\otimes n}g\rangle = \langle \tilde{f}, \tilde{T}^{\otimes n}\tilde{g}\rangle_\gamma$ where $\tilde{f}, \tilde{g}$ denote the real analogues of f, g respectively and $\tilde{T}$ denotes the Gaussian analogue of T.*

Proof: Both α_x and Γ_x form an orthonormal set of functions hence both sides of the first equality are $\sum_x \hat{f}(\alpha_x)\hat{g}(\alpha_x)$. For the second claim, notice that for every x, α_x is an eigenvector of $T^{\otimes n}$ and Γ_x is an eigenvector of $\tilde{T}^{\otimes n}$ and both correspond to the eigenvalue $\prod_{a \neq 0}\lambda_a^{|x|_a}$. Hence, both sides of the second equality are

$$\sum_x \left(\prod_{a \neq 0}\lambda_a^{|x|_a}\right) \hat{f}(\alpha_x)\hat{g}(\alpha_x).$$
✎

Definition 3.7 *For any function f with range $\mathbb{R}$, define the function $\mathrm{chop}(f)$ as*

$$\mathrm{chop}(f)(x) = \begin{cases} f(x) & \text{if } f(x) \in [0,1] \\ 0 & \text{if } f(x) < 0 \\ 1 & \text{if } f(x) > 1 \end{cases}$$

The following theorem is proven in [17]. It shows that under certain conditions, if a function f obtains values in $[0,1]$ then $\tilde{f}$ and $\mathrm{chop}(\tilde{f})$ are close. Its proof is non-trivial and builds on the main technical result of [17], a result that is known as an invariance principal. In essence, it shows that the distribution of values obtained by f and that of values obtained by $\tilde{f}$ are close. In particular, since f never deviates from $[0,1]$, it implies that $\tilde{f}$ rarely deviates from $[0,1]$ and hence $\tilde{f}$ and $\mathrm{chop}(\tilde{f})$ are close. See [17] for more details.

Theorem 3.8 ([17, Theorem 3.18]) *For any $\eta < 1$ and $\varepsilon > 0$ there exists a $\delta > 0$ such that the following holds. For any function $f : [q]^n \rightarrow [0,1]$ such that $\forall x \ |\hat{f}(\alpha_x)| \leq \eta^{|x|}$ and $\forall i \ I_i(f) < \delta$, then $\|\tilde{f} - \mathrm{chop}(\tilde{f})\| \leq \varepsilon$.*

We are now ready to prove the first step in the proof of Theorem 3.1. It is here that we use the invariance principle and Borell's inequality.

Lemma 3.9 *Let q be a fixed integer and let T be a symmetric Markov operator on $[q]$ such that $\rho = r(T) < 1$. Then for any $\varepsilon > 0$, $\eta < 1$, there exists a $\delta > 0$ such that for any functions $f, g : [q]^n \rightarrow [0,1]$ satisfying $\mathbf{E}[f] = \mu, \mathbf{E}[g] = \nu$, $\forall i \ \max(I_i(f), I_i(g)) < \delta$ and $\forall x \ |\hat{f}(\alpha_x)| \leq \eta^{|x|}, \ |\hat{g}(\alpha_x)| \leq \eta^{|x|}$, it holds that*

$$\langle f, T^{\otimes n} g \rangle \leq \langle F_\mu, U_\rho F_\nu \rangle_\gamma + \varepsilon.$$

Proof: Let $\mu' = \mathbf{E}_\gamma[\mathrm{chop}(\tilde{f})]$ and $\nu' = \mathbf{E}_\gamma[\mathrm{chop}(\tilde{g})]$. We note that $\langle F_\mu, U_\rho F_\nu \rangle_\gamma$ is a uniformly continuous function of μ and ν. Let ε_1 be chosen such that if $|\mu - \mu'| \leq \varepsilon_1$ and $|\nu - \nu'| \leq \varepsilon_1$ then it holds that

$$|\langle F_\mu, U_\rho F_\nu \rangle_\gamma - \langle F_{\mu'}, U_\rho F_{\nu'} \rangle_\gamma| \leq \varepsilon/2.$$

Let $\varepsilon_2 = \min(\varepsilon/4, \varepsilon_1)$ and let $\delta = \delta(\eta, \varepsilon_2)$ be the value given by Theorem 3.8 with ε taken to be ε_2. Then, using the Cauchy-Schwartz inequality,

$$|\mu' - \mu| = |\mathbf{E}_\gamma[\mathrm{chop}(\tilde{f}) - \tilde{f}]| = |\langle \mathrm{chop}(\tilde{f}) - \tilde{f}, \mathbf{1} \rangle_\gamma| \leq \|\mathrm{chop}(\tilde{f}) - \tilde{f}\| \leq \varepsilon_2 \leq \varepsilon_1.$$

Similarly, we have $|\nu' - \nu| \leq \varepsilon_1$. Now,

$$
\begin{aligned}
\langle f, T^{\otimes n} g \rangle &= \langle \tilde{f}, \tilde{T}^{\otimes n} \tilde{g} \rangle_\gamma && \text{(Claim 3.6)} \\
&= \langle \mathrm{chop}(\tilde{f}), \tilde{T}^{\otimes n} \mathrm{chop}(\tilde{g}) \rangle_\gamma + \\
&\quad \langle \mathrm{chop}(\tilde{f}), \tilde{T}^{\otimes n}(\tilde{g} - \mathrm{chop}(\tilde{g})) \rangle_\gamma + \langle \tilde{f} - \mathrm{chop}(\tilde{f}), \tilde{T}^{\otimes n} \tilde{g} \rangle_\gamma \\
&\leq \langle \mathrm{chop}(\tilde{f}), \tilde{T}^{\otimes n} \mathrm{chop}(\tilde{g}) \rangle_\gamma + 2\varepsilon_2 \\
&\leq \langle F_{\mu'}, U_\rho F_{\nu'} \rangle_\gamma + 2\varepsilon_2 && \text{(Corollary 3.4)} \\
&\leq \langle F_\mu, U_\rho F_\nu \rangle_\gamma + \varepsilon/2 + 2\varepsilon_2 \leq \langle F_\mu, U_\rho F_\nu \rangle_\gamma + \varepsilon
\end{aligned}
$$

where the first inequality follows from the Cauchy-Schwartz inequality together with the fact that $\mathrm{chop}(\tilde{f})$ and $\tilde{g}$ have L_2 norm at most 1 and that $\tilde{T}^{\otimes n}$ is a contraction on L_2. ✎

Due to lack of space, the rest of the proof is deferred to Appendix B. There, we complete the proof of the theorem by essentially showing how to replace the $\max(I_i(f), I_i(g))$ by $\min(I_i(f), I_i(g))$, as in the statement of the theorem. This is based on the idea of using two influence 'thresholds'. We also present there a theorem tailored to the $\bowtie$ constraint, which is somewhat more involved technically.

4 Approximate Coloring

In this section we describe and prove reductions to the three problems described in Section 2, based on three conjectures on the hardness of label-cover. These conjectures, along with some definitions, are described in Section 4.1. The three reductions are very similar, each combining a conjecture with an appropriately constructed noise operator. In Section 4.2 we describe the three noise operators, and in Section 4.3 we spell out the constructions. Due to space limitations, Section 4.3 contains only one of the three reductions (along with proofs of completeness and soundness). The remaining two reductions are deferred to Appendix E.

4.1 Label-cover problems

Definition 4.1 *A label-cover instance is a triple $G = ((V, E), R, \Psi)$ where (V, E) is a graph, R is an integer, and $\Psi = \left\{ \psi_e \subseteq \{1, \ldots, R\}^2 \mid e \in E \right\}$ is a set of constraints (relations), one for each edge. For a given labeling $L : V \to \{1, \ldots, R\}$, let*

$$\mathsf{sat}_L(G) = \Pr_{e=(u,v) \in E}[(L(u), L(v)) \in \psi_e], \qquad \mathsf{sat}(G) = \max_L(\mathsf{sat}_L(G)).$$

For $t, R \in \mathbb{N}$ let $\binom{R}{\leq t}$ denote the collection of all subsets of $\{1, \ldots, R\}$ whose size is at most t.

Definition 4.2 *A t-labeling is a function $L : V \to \binom{R}{\leq t}$ that labels each vertex $v \in V$ with a subset of values $L(v) \subseteq \{1, \ldots, R\}$ such that $|L(v)| \leq t$ for all $v \in V$. A t-labeling L is said to* satisfy *a constraint $\psi \subseteq \{1, \ldots, R\}^2$ over variables u and v iff there are $a \in L(u)$, $b \in L(v)$ such that $(a, b) \in \psi$. In other words, iff $(L(u) \times L(v)) \cap \psi \neq \emptyset$.*

For the special case of $t = 1$, a 1-labeling is like a labeling $L : V \to \{1, \ldots, R\}$ (except that some vertices get no label). In this case, a constraint ψ over u, v is satisfied by L iff $(L(u), L(v)) \in \psi$.

Similar to the definition of $\mathsf{sat}(G)$, we also define $\mathsf{isat}(G)$ ("induced-sat") to be the relative size of the largest set of vertices for which there is a labeling that satisfies *all* of the induced edges.

$$\mathsf{isat}(G) = \max_S \left\{ \frac{|S|}{|V|} \;\middle|\; \exists L : S \to \{1, \ldots, R\} \text{ that satisfies all the constraints induced by } S \subseteq V \right\}.$$

Let $\mathsf{isat}_t(G)$ denote the relative size of the largest set of vertices $S \subseteq V$ for which there is a t-labeling that satisfies *all* the constraints induced by S.

$$\mathsf{isat}_t(G) = \max_S \left\{ \frac{|S|}{|V|} \;\middle|\; \exists L : S \to \binom{R}{\leq t} \text{ that satisfies all the constraints induced by } S \subseteq V \right\}.$$

We now consider three conjectures on which our reductions are based. The main difference between the three conjectures is in the type of constraints that are allowed. The three types are illustrated in Figure A.1. Due to lack of space, we only describe the first conjecture and defer the remaining two to Appendix C.

Definition 4.3 (1↔1-**constraint**) *A 1↔1 constraint is a relation $\{(i, \pi(i))\}_{i=1}^R$, where $\pi : \{1, \ldots, R\} \to \{1, \ldots, R\}$ is any arbitrary permutation. The constraint is satisfied by (a, b) iff $b = \pi(a)$.*

Conjecture 4.4 (1↔1 **Conjecture**) *For any $\varepsilon, \zeta > 0$ and $t \in \mathbb{N}$ there exists some $R \in \mathbb{N}$ such that given a label-cover instance $G = \langle (V, E), R, \Psi \rangle$ where all constraints are 1↔1-constraints, it is NP-hard to decide between the case $\mathsf{isat}(G) \geq 1 - \zeta$ and the case $\mathsf{isat}_t(G) < \varepsilon$.*

It is easy to see that the above problem is in P when $\zeta = 0$.

4.2 Noise operators

We now define the noise operators corresponding to the $1\leftrightarrow1$-constraints, $\bowtie$-constraints, and $2\leftrightarrow2$-constraints. The noise operator that corresponds to the $1\leftrightarrow1$-constraints is the simplest, and acts on $\{0,1,2\}$. For the other two cases, since the constraints involve pairs of coordinates, we obtain an operator on $\{0,1,2\}^2$ and an operator on $\{0,1,2,3\}^2$. See Figure A.2 for an illustration.

Lemma 4.5 *There exists a symmetric Markov operator T on $\{0,1,2\}$ such that $r(T) < 1$ and such that if $T(x \leftrightarrow y) > 0$ then $x \neq y$.*

Proof: Take the operator given by

$$T = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}.$$

The two remaining noise operators are described in Appendix D.

4.3 The Reductions

The basic idea in all three reductions is to take a label-cover instance and to replace each vertex with a block of q^R vertices, corresponding to the q-ary hypercube $[q]^R$. The intended way to q-color this block is by coloring $x \in [q]^R$ according to x_i where i is the label given to this block. One can think of this coloring as an encoding of the label i. We will essentially prove that any other coloring of this block that uses relatively few colors, can be "list-decoded" into at most t labels from $\{1, \ldots, R\}$. By properly defining edges connecting these blocks, we can guarantee that the lists decoded from two blocks can be used as t-labelings for the label-cover instance.

In this section we will describe only the reduction for ALMOST-3-COLORING. We defer the two (similar) remaining reductions (for APPROXIMATE-COLORING$(4, Q)$ and for APPROXIMATE-COLORING$(3, Q)$) to Appendix E.

In the rest of this section, we use the following notation. For a vector $x = (x_1, \ldots, x_n)$ and a permutation π on $\{1, \ldots, n\}$, we define $x^\pi = (x_{\pi(1)}, \ldots, x_{\pi(n)})$.

ALMOST-3-COLORING

Let $G = ((V, E), R, \Psi)$ be a label-cover instance as in Conjecture 4.4. For $v \in V$ write $[v]$ for a collection of vertices, one per point in $\{0,1,2\}^R$. Let $e = (v, w) \in E$, and let ψ be the $1\leftrightarrow1$-constraint associated with e. By Definition 4.3 there is a permutation π such that $(a, b) \in \psi$ iff $b = \pi(a)$. We now write $[v, w]$ for the following collection of edges. We put an edge (x, y) for $x = (x_1, \ldots, x_R) \in [v]$ and $y = (y_1, \ldots, y_R) \in [w]$ iff

$$\forall i \in \{1, \ldots, R\}, \quad T\left(x_i \leftrightarrow y_{\pi(i)}\right) \neq 0$$

where T is the noise operator from Lemma 4.5. In other words, x is adjacent to y whenever

$$T^{\otimes R}\left(x \leftrightarrow y^\pi\right) = \prod_{i=1}^{R} T\left(x_i \leftrightarrow y_{\pi(i)}\right) \neq 0.$$

The reduction outputs the graph $[G] = ([V], [E])$ where $[V]$ is the disjoint union of all blocks $[v]$ and $[E]$ is the disjoint union of all collections of edges $[v, w]$.

9

Completeness. If $\mathrm{isat}(G) \geq 1 - \varepsilon$, then there is some $S \subseteq V$ of size $(1 - \varepsilon)|V|$ and a labeling $\ell : S \to R$ that satisfies all of the constraints induced by S. We 3-color all of the vertices in $\cup_{v \in S}[v]$ as follows. Let $c : \cup_{v \in S}[v] \to \{0, 1, 2\}$ be defined as follows. For every $v \in S$, the color of $x = (x_1, \ldots, x_R) \in \{0, 1, 2\}^R = [v]$ is defined to be $c(x) := x_i$, where $i = \ell(v) \in \{1, \ldots, R\}$.

To see that c is a legal coloring on $\cup_{v \in S}[v]$, observe that if $x \in [v]$ and $y \in [w]$ share the same color, then $x_i = y_j$ for $i = \ell(v)$ and $j = \ell(w)$. Since ℓ satisfies every constraint induced by S, it follows that if (v, w) is a constraint with an associated permutation π, then $j = \pi(i)$. Since $T(z \leftrightarrow z) = 0$ for all $z \in \{0, 1, 2\}$, there is no edge between x and y.

Soundness. Before presenting the soundness proof, we need the following corollary. It is simply a special case of Theorem 3.1 stated in the contrapositive, with ε playing the role of ν and μ. Here we use the fact that $\langle F_\varepsilon, U_\rho(1 - F_{1-\varepsilon}) \rangle_\gamma > 0$ whenever $\varepsilon > 0$.

Corollary 4.6 *Let q be a fixed integer and let T be a reversible Markov operator on $[q]$ such that $r(T) < 1$. Then for any $\varepsilon > 0$ there exist $\delta > 0$ and $k \in \mathbb{N}$ such that the following holds. For any $f, g : [q]^n \to [0, 1]$, if $E[f] \geq \varepsilon$, $E[g] \geq \varepsilon$, and $\langle f, T^{\otimes n} g \rangle = 0$, then*

$$\exists i \in \{1, \ldots, n\}, \quad I_i^{\leq k}(f) \geq \delta \quad \text{and} \quad I_i^{\leq k}(g) \geq \delta.$$

We will show that if $[G]$ has an independent set $S \subseteq [V]$ of relative size $\geq 2\varepsilon$, then $\mathrm{isat}_t(G) \geq \varepsilon$ for a fixed constant $t > 0$ that depends only on ε. More explicitly, we will find a set $J \subseteq V$, and a t-labeling $L : J \to \binom{R}{\leq t}$ such that $|J| \geq \varepsilon|V|$ and L satisfies all the constraints of G induced by J. In other words, for every constraint ψ over an edge $(u, v) \in E \cap J^2$, there are values $a \in L(u)$ and $b \in L(v)$ such that $(a, b) \in \psi$.

Let J be the set of all vertices $v \in V$ such that the fraction of vertices belonging to S in $[v]$ is at least ε. Then, since $|S| \geq 2\varepsilon|[V]|$, Markov's inequality implies $|J| \geq \varepsilon|V|$.

For each $v \in J$ let $f_v : \{0, 1, 2\}^R \to \{0, 1\}$ be the characteristic function of S restricted to $[v]$, so $\mathbf{E}[f_v] \geq \varepsilon$. Select δ, k according to Corollary 4.6 with ε and the operator T of Lemma 4.5, and set

$$L(v) = \left\{ i \in \{1, \ldots, R\} \,\middle|\, I_i^{\leq k}(f_v) \geq \delta \right\}.$$

Clearly, $|L(v)| \leq k/\delta$ because $\sum_{i=1}^R I_i^{\leq k}(f) \leq k$. Thus, L is a t-labeling for $t = k/\delta$. The main point to prove is that for every edge $e = (v_1, v_2) \in E \cap J^2$ induced on J, there is some $a \in L(v_1)$ and $b \in L(v_2)$ such that $(a, b) \in \psi_e$. This would imply that $\mathrm{isat}_t(G) \geq |J|/|V| \geq \varepsilon$.

Fix $(v_1, v_2) \in E \cap J^2$, and let π be the permutation associated with the $1 \leftrightarrow 1$ constraint on this edge. (It may be easier to first think of $\pi = id$.) Recall that the edges in $[v_1, v_2]$ were defined based on π, and on the noise operator T defined in Lemma 4.5. Let $f = f_{v_1}$, and define g by $g(x^\pi) = f_{v_2}(x)$. Since S is an independent set, $f(x) = f_{v_1}(x) = 1$ and $g(y^\pi) = f_{v_2}(y) = 1$ implies that x, y are not adjacent, so by construction $T^{\otimes R}(x \leftrightarrow y^\pi) = 0$. Therefore,

$$\langle f, T^{\otimes R} g \rangle = 3^{-R} \sum_x f(x) T^{\otimes R} g(x) = 3^{-R} \sum_x f(x) \sum_{y^\pi} T^{\otimes R}(x \leftrightarrow y^\pi) g(y^\pi) = \sum_{x, y^\pi} 0 = 0.$$

Also, by assumption, $E[g] \geq \varepsilon$ and $E[f] \geq \varepsilon$. Corollary 4.6 implies that there is some index $i \in \{1, \ldots, R\}$ for which both $I_i^{\leq k}(f) \geq \delta$ and $I_i^{\leq k}(g) \geq \delta$. By definition of L, $i \in L(v_1)$. Since the i-th variable in g is the $\pi(i)$-th variable in f_{v_2}, $\pi(i) \in L(v_2)$. It follows that there are values $i \in L(v_1)$ and $\pi(i) \in L(v_2)$ such that $(i, \pi(i))$ satisfies the constraint on (v_1, v_2). This means that $\mathrm{isat}_t(G) \geq |J|/|V| \geq \varepsilon$.

5 Omitted proofs and discussions

We have omitted from this extended abstract a comparison between the conjectures we have used and Khot's original conjectures. A detailed comparison is given in Appendix G. The rest of the main Fourier theorem is given in Appendix B. The reductions for APPROXIMATE-COLORING are given in Appendix E.

Acknowledgements

References

[1] N. Alon, I. Dinur, E. Friedgut, and B. Sudakov. Graph products, fourier analysis and spectral techniques. *GAFA*, 14(5):913–940, 2004.

[2] M. Bellare, O. Goldreich, and M. Sudan. Free bits, PCPs, and nonapproximability—towards tight results. *SIAM J. Comput.*, 27(3):804–915, 1998.

[3] A. Blum and D. Karger. An $\tilde{O}(n^{3/14})$-coloring algorithm for 3-colorable graphs. *Inform. Process. Lett.*, 61(1):49–53, 1997.

[4] C. Borell. Geometric bounds on the Ornstein-Uhlenbeck velocity process. *Z. Wahrsch. Verw. Gebiete*, 70(1):1–13, 1985.

[5] I. Dinur and S. Safra. On the hardness of approximating minimum vertex cover. *Annals of Mathematics*, 2005. To appear.

[6] U. Feige and J. Kilian. Zero knowledge and the chromatic number. *J. Comput. System Sci.*, 57(2):187–199, 1998.

[7] M. Fürer. Improved hardness results for approximating the chromatic number. In *Proc. 36th IEEE Symp. on Foundations of Computer Science*, pages 414–421, 1995.

[8] V. Guruswami and S. Khanna. On the hardness of 4-coloring a 3-colorable graph. In *15th Annual IEEE Conference on Computational Complexity (Florence, 2000)*, pages 188–197. IEEE Computer Soc., Los Alamitos, CA, 2000.

[9] E. Halperin, R. Nathaniel, and U. Zwick. Coloring k-colorable graphs using relatively small palettes. *J. Algorithms*, 45(1):72–90, 2002.

[10] J. Håstad. Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001.

[11] D. Karger, R. Motwani, and M. Sudan. Approximate graph coloring by semidefinite programming. *Journal of the ACM*, 45(2):246–265, 1998.

[12] S. Khanna, N. Linial, and S. Safra. On the hardness of approximating the chromatic number. *Combinatorica*, 20(3):393–415, 2000.

[13] S. Khot. Improved inapproximability results for MaxClique, chromatic number and approximate graph coloring. In *Proc. 42nd IEEE Symp. on Foundations of Computer Science*, pages 600–609, 2001.

[14] S. Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, pages 767–775. ACM Press, 2002.

[15] S. Khot, G. Kindler, E. Mossel, and R. ODonnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? In *Proc. 45th IEEE Symp. on Foundations of Computer Science*, pages 146–154, 2004.

[16] S. Khot and O. Regev. Vertex cover might be hard to approximate to within $2 - \varepsilon$. In *Proc. of 18th IEEE Annual Conference on Computational Complexity (CCC)*, pages 379–386, 2003.

[17] E. Mossel, R. O'Donnell, and K. Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. Preprint, 2005.

[18] L. Trevisan. Approximation algorithms for unique games. In *ECCCTR: Electronic Colloquium on Computational Complexity, technical reports*, 2005.
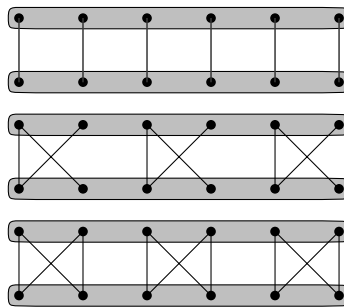
A   Figures

Figure A.1: Three types of constraints (top to bottom): $1 \leftrightarrow 1, \bowtie, 2 \leftrightarrow 2$
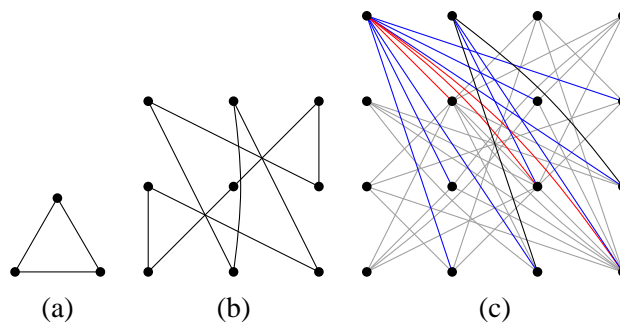

Figure A.2: Three noise operators (edge weights not shown) corresponding to: (a) $1 \leftrightarrow 1$, (b) $\bowtie$, and (c) $2 \leftrightarrow 2$.

B   Rest of Proof of Theorem 3.1

We complete the proof of Theorem 3.1 by proving:

Lemma B.1 *Let q be a fixed integer and let T be a symmetric Markov operator on $[q]$ such that $\rho = r(T) < 1$. Then for any $\varepsilon > 0$, there exists a $\delta > 0$ and an integer k such that for any functions $f, g : [q]^n \to [0, 1]$ satisfying $\mathbf{E}[f] = \mu, \mathbf{E}[g] = \nu$ and*

$$\forall i \quad \min\left(I_i^{\leq k}(f), I_i^{\leq k}(g)\right) < \delta \tag{3}$$

then

$$\langle f, T^{\otimes n} g \rangle \leq \langle F_\mu, U_\rho F_\nu \rangle_\gamma + \varepsilon. \tag{4}$$

Proof: Let $f_1 = T_\eta^{\otimes n} f$ and $g_1 = T_\eta^{\otimes n} g$ where $\eta < 1$ is chosen so that $\rho^j(1 - \eta^{2j}) < \varepsilon/4$ for all j. Then

$$\left|\langle f_1, T^{\otimes n} g_1 \rangle - \langle f, T^{\otimes n} g \rangle\right| = \left|\sum_x \hat{f}(\alpha_x)\hat{g}(\alpha_x) \prod_{a \neq 0} \lambda_a^{|x|_a}(1 - \eta^{2|x|})\right|$$

$$\leq \sum_x \rho^{|x|}(1 - \eta^{2|x|})\left|\hat{f}(\alpha_x)\hat{g}(\alpha_x)\right| \leq \varepsilon/4$$

where the last inequality follows from the Cauchy-Schwartz inequality. Thus, in order to prove (4) it suffices to prove

$$\langle f_1, T^{\otimes n} g_1 \rangle \leq \langle F_\mu, U_\rho F_\nu \rangle_\gamma + 3\varepsilon/4. \tag{5}$$

Let $\delta(\varepsilon/4, \eta)$ be the value given by Lemma 3.9 plugging in $\varepsilon/4$ for ε. Let $\delta' = \delta(\varepsilon/4, \eta)/2$. Let k be chosen so that $\eta^{2k} < \min(\delta', \varepsilon/4)$. Define $C = k/\delta'$ and $\delta = (\varepsilon/8C)^2 < \delta'$. Let

$$B_f = \{i : I_i^{\leq k}(f) \geq \delta'\}, \quad B_g = \{i : I_i^{\leq k}(g) \geq \delta'\}.$$

We note that B_f and B_g are of size at most $C = k/\delta'$. By (3), we have that whenever $i \in B_f$, $I_i^{\leq k}(g) < \delta$. Similarly, for every $i \in B_g$ we have $I_i^{\leq k}(f) < \delta$. In particular, B_f and B_g are disjoint.

Recall the averaging operator A. We now let

$$f_2(x) = A_{B_f}(f_1) = \sum_{x:x_{B_f}=0} \hat{f}(\alpha_x)\alpha_x \eta^{|x|},$$

$$g_2(x) = A_{B_g}(g_1) = \sum_{x:x_{B_g}=0} \hat{g}(\alpha_x)\alpha_x \eta^{|x|}.$$

Clearly, $\mathbf{E}[f_2] = \mathbf{E}[f]$ and $\mathbf{E}[g_2] = \mathbf{E}[g]$, and for all x $f_2(x), g_2(x) \in [0, 1]$. It is easy to see that $I_i(f_2) = 0$ if $i \in B_f$ and $I_i(f_2) \leq I_i^{\leq k}(f) + \eta^{2k} < 2\delta'$ otherwise and similarly for g_2. Thus, for any $i, \max(I_i(f_2), I_i(g_2)) < 2\delta'$. We also see that for any $x, |\hat{f}_2(\alpha_x)| \leq \eta^{|x|}$ and the same for g_2. Thus, we can apply Lemma 3.9 to obtain that

$$\langle f_2, T^{\otimes n} g_2 \rangle \leq \langle F_\mu, U_\rho F_\nu \rangle_\gamma + \varepsilon/4.$$

In order to show (5) and complete the proof, we show that

$$\left|\langle f_1, T^{\otimes n} g_1 \rangle - \langle f_2, T^{\otimes n} g_2 \rangle\right| \leq \varepsilon/2.$$

This follows by

$$
\begin{aligned}
|\langle f_1, T^{\otimes n} g_1 \rangle - \langle f_2, T^{\otimes n} g_2 \rangle| &= \Big| \sum_{x : x_{B_f \cup B_g} \neq 0} \hat{f}(\alpha_x) \hat{g}(\alpha_x) \prod_{a \neq 0} \lambda_a^{|x|_a} \eta^{2|x|} \Big| \\
&\leq \eta^{2k} \sum_{x : |x| \geq k} \left| \hat{f}(\alpha_x) \hat{g}(\alpha_x) \right| + \sum \left\{ \left| \hat{f}(\alpha_x) \hat{g}(\alpha_x) \right| : x_{B_f \cup B_g} \neq 0, |x| \leq k \right\} \\
&\leq \varepsilon/4 + \sum_{i \in B_f \cup B_g} \sum \left\{ \left| \hat{f}(\alpha_x) \hat{g}(\alpha_x) \right| : x_i \neq 0, |x| \leq k \right\} \\
&\leq \varepsilon/4 + \sum_{i \in B_f \cup B_g} \sqrt{I_i^{\leq k}(f)} \sqrt{I_i^{\leq k}(g)} \\
&\leq \varepsilon/4 + \sqrt{\delta}(|B_f| + |B_g|) \\
&\leq \varepsilon/4 + 2C\sqrt{\delta} = \varepsilon/2,
\end{aligned}
$$

where the next-to-last inequality holds because for each $i \in B_f \cup B_g$ one of $I_i^{\leq k}(f), I_i^{\leq k}(g)$ is at most δ and the other is at most 1. ✎

The final theorem of this section is needed only for the APPROXIMATE-COLORING$(3, Q)$ result. Here, the operator T acts on $[q^2]$ and is assumed to have an additional property. Before proceeding, it is helpful to recall Definition 2.6.

Theorem B.2 *Let q be a fixed integer and let T be a symmetric Markov operator on $[q^2]$ such that $\rho = r(T) < 1$. Suppose moreover, that T has the following property. Given (x_1, x_2) chosen uniformly at random and (y_1, y_2) chosen according to T applied to (x_1, x_2) we have that (x_2, y_2) is distributed uniformly at random. Then for any $\varepsilon > 0$, there exists a $\delta > 0$ and an integer k such that for any functions $f, g : [q]^{2n} \to [0, 1]$ satisfying $\mathbf{E}[f] = \mu, \mathbf{E}[g] = \nu$, and for $i = 1, \ldots, n$*

$$
\min \left(I_{2i-1}^{\leq k}(f), I_{2i-1}^{\leq k}(g) \right) < \delta, \qquad \min \left(I_{2i-1}^{\leq k}(f), I_{2i}^{\leq k}(g) \right) < \delta, \text{ and} \qquad \min \left(I_{2i}^{\leq k}(f), I_{2i-1}^{\leq k}(g) \right) < \delta
$$

it holds that

$$
\langle \overline{f}, T^{\otimes n} \overline{g} \rangle \geq \langle F_\mu, U_\rho(1 - F_{1-\nu}) \rangle_\gamma - \varepsilon \tag{6}
$$

and

$$
\langle \overline{f}, T^{\otimes n} \overline{g} \rangle \leq \langle F_\mu, U_\rho F_\nu \rangle_\gamma + \varepsilon. \tag{7}
$$

Proof: As in Theorem 3.1, (6) follows from (7) so it is enough to prove (7). Assume first that in addition to the three conditions above we also have that for all $i = 1, \ldots, n$,

$$
\min \left(I_{2i}^{\leq k}(f), I_{2i}^{\leq k}(g) \right) < \delta. \tag{8}
$$

Then it follows that for all i, either both $I_{2i-1}^{\leq k}(f)$ and $I_{2i}^{\leq k}(f)$ are smaller than δ or both $I_{2i-1}^{\leq k}(g)$ and $I_{2i}^{\leq k}(g)$ are smaller than δ. Hence, by Claim 2.7, we know that for all i we have

$$
\min \left(I_i^{\leq k/2}(\overline{f}), I_i^{\leq k/2}(\overline{g}) \right) < 2\delta
$$

and the result then follows from Lemma B.1. However, we do not have this extra condition and hence we have to deal with 'bad' coordinates i for which $\min(I_{2i}^{\leq k}(f), I_{2i}^{\leq k}(g)) \geq \delta$. Notice for such i it must be the case that both $I_{2i-1}^{\leq k}(f)$ and $I_{2i-1}^{\leq k}(g)$ are smaller than δ. Informally, the proof proceeds as follows. We first

14

define functions f_1, g_1 that are obtained from f, g by adding a small amount of noise. We then obtain f_2, g_2 from f_1, g_1 by averaging the coordinates $2i - 1$ for bad i. Finally, we obtain f_3, g_3 from f_2, g_2 by averaging the coordinate $2i$ for bad i. The point here is to maintain $\langle \overline{f}, T^{\otimes n}\overline{g} \rangle \approx \langle \overline{f_1}, T^{\otimes n}\overline{g_1} \rangle \approx \langle \overline{f_2}, T^{\otimes n}\overline{g_2} \rangle \approx \langle \overline{f_3}, T^{\otimes n}\overline{g_3} \rangle$. The condition in Equation 8 now applies to f_3, g_3 and we can apply Lemma B.1, as described above. We now describe the proof in more detail.

We first define $\overline{f_1} = T_\eta^{\otimes n}\overline{f}$ and $\overline{g_1} = T_\eta^{\otimes n}\overline{g}$ where $\eta < 1$ is chosen so that $\rho^j(1 - \eta^{2j}) < \varepsilon/4$ for all j. As in the previous lemma it is easy to see that

$$|\langle \overline{f_1}, T^{\otimes n}\overline{g_1} \rangle - \langle \overline{f}, T^{\otimes n}\overline{g} \rangle| < \varepsilon/4$$

and thus it suffices to prove that

$$\langle \overline{f_1}, T^{\otimes n}\overline{g_1} \rangle \leq \langle F_\mu, U_\rho F_\nu \rangle_\gamma + 3\varepsilon/4.$$

Let $\delta(\varepsilon/2, \eta), k(\varepsilon/2, \eta)$ be the values given by Lemma B.1 with ε taken to be $\varepsilon/2$. Let $\delta' = \delta(\varepsilon/2, \eta)/2$. Choose a large enough k so that $128k\eta^k < \varepsilon^2\delta'$ and $k/2 > k(\varepsilon/2, \eta)$. We let $C = k/\delta'$ and $\delta = \varepsilon^2/128C$. Notice that $\delta < \delta'$ and $\eta^k < \delta$. Finally, let

$$B = \left\{ i \ \middle| \ I_{2i}^{\leq k}(f) \geq \delta', I_{2i}^{\leq k}(g) \geq \delta' \right\}.$$

We note that B is of size at most C. We also note that if $i \in B$ then we have $I_{2i-1}^{\leq k}(f) < \delta$ and $I_{2i-1}^{\leq k}(g) < \delta$. We claim that this implies that $I_{2i-1}(f_1) \leq \delta + \eta^k < 2\delta$ and similarly for g. To see that, take any orthonormal basis $\beta_0 = 1, \beta_1, \ldots, \beta_{q-1}$ of $\mathbb{R}^q$ and notice that we can write

$$f_1 = \sum_{x \in [q]^{2n}} \hat{f}(\beta_x)\eta^{|x|}\beta_x.$$

Hence,

$$I_{2i-1}(f_1) = \sum_{\substack{x \in [q]^{2n} \\ x_{2i-1} \neq 0}} \hat{f}(\beta_x)^2\eta^{2|x|} < \delta + \eta^k \sum_{\substack{x \in [q]^{2n} \\ |x| > k}} \hat{f}(\beta_x)^2 \leq \delta + \eta^k$$

where we used that the number of nonzero elements in $\overline{x}$ is at least half of that in x.

Next, we define $f_2 = A_{2B-1}(f_1)$ and $g_2 = A_{2B-1}(g_1)$ where A is the averaging operator and $2B - 1$ denotes the set $\{2i - 1 \mid i \in B\}$. Note that

$$\|\overline{f_2} - \overline{f_1}\|_2^2 = \|f_2 - f_1\|_2^2 \leq \sum_{i \in B} I_{2i-1}(f_1) \leq 2C\delta.$$

and similarly,

$$\|\overline{g_2} - \overline{g_1}\|_2^2 = \|g_2 - g_1\|_2^2 \leq 2C\delta.$$

Thus

$$|\langle \overline{f_1}, T^{\otimes n}\overline{g_1} \rangle - \langle \overline{f_2}, T^{\otimes n}\overline{g_2} \rangle| \leq |\langle \overline{f_1}, T^{\otimes n}\overline{g_1} \rangle - \langle \overline{f_1}, T^{\otimes n}\overline{g_2} \rangle| + |\langle \overline{f_1}, T^{\otimes n}\overline{g_2} \rangle - \langle \overline{f_2}, T^{\otimes n}\overline{g_2} \rangle|$$
$$\leq 2\sqrt{2C\delta} = \varepsilon/4$$

where the last inequality follows from the Cauchy-Schwartz inequality together with the fact that $\|\overline{f_1}\|_2 \leq 1$ and also $\|T^{\otimes n}\overline{g_2}\|_2 \leq 1$. Hence, it suffices to prove

$$\langle \overline{f_2}, T^{\otimes n}\overline{g_2} \rangle \leq \langle F_\mu, U_\rho F_\nu \rangle_\gamma + \varepsilon/2.$$

We now define $f_3 = A_{2B}(f_2)$ and $g_3 = A_{2B}(g_2)$. Equivalently, we have $\overline{f_3} = A_B(\overline{f_1})$ and $\overline{g_3} = A_B(\overline{g_1})$. We show that $\langle \overline{f_2}, T^{\otimes n}\overline{g_2} \rangle = \langle \overline{f_3}, T^{\otimes n}\overline{g_3} \rangle$. Let α_x, $x \in [q^2]^n$, be an orthonormal basis of eigenvectors of $T^{\otimes n}$. Then

$$\langle \overline{f_3}, T^{\otimes n}\overline{g_3} \rangle = \sum_{x,y \in [q^2]^n, x_B = y_B = 0} \hat{\overline{f_1}}(\alpha_x)\hat{\overline{g_1}}(\alpha_y)\langle \alpha_x, T^{\otimes n}\alpha_y \rangle.$$

Moreover, since A is a linear operator and f_1 can be written as $\sum_{x \in [q^2]^n} \hat{\overline{f_1}}(\alpha_x)\underline{\alpha_x}$ and similarly for g_1, we have

$$\langle \overline{f_2}, T^{\otimes n}\overline{g_2} \rangle = \sum_{x,y \in [q^2]^n} \hat{\overline{f_1}}(\alpha_x)\hat{\overline{g_1}}(\alpha_y)\langle \overline{A_{2B-1}(\underline{\alpha_x})}, T^{\otimes n}\overline{A_{2B-1}(\underline{\alpha_y})} \rangle.$$

First, notice that when $x_B = 0$, $\overline{A_{2B-1}(\underline{\alpha_x})} = \alpha_x$ since α_x does not depend on coordinates in B. Hence, in order to show that the two expressions above are equal, it suffices to show that

$$\langle \overline{A_{2B-1}(\underline{\alpha_x})}, T^{\otimes n}\overline{A_{2B-1}(\underline{\alpha_y})} \rangle = 0$$

unless $x_B = y_B = 0$. So assume without loss of generality that $i \in B$ is such that $x_i \neq 0$. The above inner product can be equivalently written as

$$\mathbf{E}_{z,z' \in [q^2]^n}[\overline{A_{2B-1}(\underline{\alpha_x})}(z) \cdot \overline{A_{2B-1}(\underline{\alpha_y})}(z')]$$

where z is chosen uniformly at random and z' is chosen according to $T^{\otimes n}$ applied to z. Fix some arbitrary values to $z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_n$ and $z'_1, \ldots, z'_{i-1}, z'_{i+1}, \ldots, z'_n$ and let us show that

$$\mathbf{E}_{z_i, z'_i \in [q^2]}[\overline{A_{2B-1}(\underline{\alpha_x})}(z) \cdot \overline{A_{2B-1}(\underline{\alpha_y})}(z')] = 0.$$

Since $i \in B$, the two expressions inside the expectation do not depend on $z_{i,1}$ and $z'_{i,1}$ (where by $z_{i,1}$ we mean the first coordinate of $\underline{z_i}$). Moreover, by our assumption on T, $z_{i,2}$ and $z'_{i,2}$ are independent. Hence, the above expectation is equal to

$$\mathbf{E}_{z_i \in [q^2]}[\overline{A_{2B-1}(\underline{\alpha_x})}(z)] \cdot \mathbf{E}_{z'_i \in [q^2]}[\overline{A_{2B-1}(\underline{\alpha_y})}(z')].$$

Since $x_i \neq 0$, the first expectation is zero. This establishes that $\langle \overline{f_2}, T^{\otimes n}\overline{g_2} \rangle = \langle \overline{f_3}, T^{\otimes n}\overline{g_3} \rangle$.

The functions f_3, g_3 satisfy the property that for every $i = 1, \ldots, n$, either both $I_{2i-1}^{\leq k}(f_3)$ and $I_{2i}^{\leq k}(f_3)$ are smaller than δ' or both $I_{2i-1}^{\leq k}(g_3)$ and $I_{2i}^{\leq k}(g_3)$ are smaller than δ'. By Claim 2.7, we get that for $i = 1, \ldots, n$, either $I_i^{\leq k/2}(\overline{f_3})$ or $I_i^{\leq k/2}(\overline{g_3})$ is smaller $2\delta'$. We can now apply Lemma B.1 to obtain

$$\langle \overline{f_3}, T^{\otimes n}\overline{g_3} \rangle \leq \langle F_\mu, U_\rho F_\nu \rangle_\gamma + \varepsilon/2.$$

□

C The Two Remaining Conjectures

Definition C.1 (2↔2**-constraint**) *A 2↔2 constraint is defined by a pair of permutations $\pi_1, \pi_2 : \{1, \ldots, 2R\} \to \{1, \ldots, 2R\}$ and the relation*

$$2\leftrightarrow 2 = \{(2i, 2i), (2i, 2i-1), (2i-1, 2i), (2i-1, 2i-1)\}_{i=1}^R.$$

The constraint is satisfied by (a, b) iff $(\pi_1^{-1}(a), \pi_2^{-1}(b)) \in 2\leftrightarrow 2$.

Definition C.2 ($\bowtie$-constraint) *An $\bowtie$ constraint is defined by a pair of permutations $\pi_1, \pi_2 : \{1, \ldots, 2R\} \to \{1, \ldots, 2R\}$ and the relation*

$$\bowtie = \{(2i-1, 2i-1), (2i, 2i-1), (2i-1, 2i)\}_{i=1}^{R} \ .$$

The constraint is satisfied by (a, b) iff $(\pi_1^{-1}(a), \pi_2^{-1}(b)) \in \bowtie$.

Conjecture C.3 (2↔2 Conjecture) *For any $\varepsilon > 0$ and $t \in \mathbb{N}$ there exists some $R \in \mathbb{N}$ such that given a label-cover instance $G = \langle (V, E), 2R, \Psi \rangle$ where all constraints are 2↔2-constraints, it is NP-hard to decide between the case $\mathsf{sat}(G) = 1$ and the case $\mathsf{isat}_t(G) < \varepsilon$.*

The 1↔1 conjecture and the above conjecture are no stronger than the corresponding conjectures of Khot. Namely, our 1↔1 conjecture is not stronger than Khot's (bipartite) unique games conjecture, and our 2↔2 conjecture is not stronger than Khot's (bipartite) 2→1 conjecture. The former claim was already proven by Khot and Regev in [16]. The latter claim is proven in a similar way. For completeness, we include both proofs in Appendix G. We also make a third conjecture that is used in our reduction to APPROXIMATE-COLORING$(3, Q)$. This conjecture seems stronger than Khot's conjectures.

Conjecture C.4 ($\bowtie$ Conjecture) *For any $\varepsilon > 0$ and $t \in \mathbb{N}$ there exists some $R \in \mathbb{N}$ such that given a label-cover instance $G = \langle (V, E), 2R, \Psi \rangle$ where all constraints are $\bowtie$-constraints, it is NP-hard to decide between the case $\mathsf{sat}(G) = 1$ and the case $\mathsf{isat}_t(G) < \varepsilon$.*

Remark: The (strange-looking) $\bowtie$-shaped constraints have already appeared before, in [5]. There, it is (implicitly) proven that for all $\varepsilon, \zeta > 0$ given a label-cover instance G where all constraints are $\bowtie$-constraints, it is NP-hard to distinguish between the case $\mathsf{isat}(G) > 1 - \zeta$ and the case $\mathsf{isat}_{t=1}(G) < \varepsilon$. The main difference between their case and our conjecture is that in our conjecture we consider any constant t, while in their case t is 1.

D The Two Remaining Operators

Lemma D.1 *There exists a symmetric Markov operator T on $\{0, 1, 2, 3\}^2$ such that $r(T) < 1$ and such that if $T((x_1, x_2) \leftrightarrow (y_1, y_2)) > 0$ then $\{x_1, x_2\} \cap \{y_1, y_2\} = \emptyset$.*

Proof: Our operator has three types of transitions, with transitions probabilities β_1, β_2, and β_3.

- With probability β_1 we have $(x, x) \leftrightarrow (y, y)$ where $x \neq y$.

- With probability β_2 we have $(x, x) \leftrightarrow (y, z)$ where x, y, z are all different.

- With probability β_3 we have $(x, y) \leftrightarrow (z, w)$ where x, y, z, w are all different.

These transitions are illustrated in Figure A.2(c) with red indicating β_1 transitions, blue indicating β_2 transitions, and black indicating β_3 transitions. For T to be symmetric Markov operator, we need that β_1, β_2 and β_3 are non-negative and

$$3\beta_1 + 6\beta_2 = 1, \quad 2\beta_2 + 2\beta_3 = 1.$$

It is easy to see that the two equations above have solutions bounded away from 0 and that the corresponding operator has $r(T) < 1$. For example, choose $\beta_1 = 1/12$, $\beta_2 = 1/8$, and $\beta_3 = 3/8$. ✎

Lemma D.2 *There exists a symmetric Markov operator T on $\{0, 1, 2\}^2$ such that $r(T) < 1$ and such that if $T((x_1, x_2) \leftrightarrow (y_1, y_2)) > 0$ then $x_1 \notin \{y_1, y_2\}$ and $y_1 \notin \{x_1, x_2\}$. Moreover, the noise operator T satisfies the following property. Let (x_1, x_2) be chosen according to the uniform distribution and (y_1, y_2) be chosen according T applied to (x_1, x_2). Then the distribution of (x_2, y_2) is uniform.*

Proof: The proof resembles the previous proof. Again there are 3 types of transitions.

- With probability β_1 we have $(x, x) \leftrightarrow (y, y)$ where $x \neq y$.

- With probability β_2 we have $(x, x) \leftrightarrow (y, z)$ where x, y, z are all different.

- With probability β_3 we have $(x, y) \leftrightarrow (z, y)$ where x, y, z are all different.

For T to be a symmetric Markov operator we require β_1, β_2 and β_3 to be non-negative and

$$2\beta_1 + 2\beta_2 = 1, \quad \beta_2 + \beta_3 = 1.$$

Moreover, the last requirement of uniformity of (x_2, y_2) amounts to the equation

$$\beta_1/3 + 2\beta_2/3 = 2\beta_3/3.$$

It is easy to see that $\beta_2 = \beta_3 = 0.5$ and $\beta_1 = 0$ is the solution of all equations and that the corresponding operator has $r(T) < 1$. This operator is illustrated in Figure A.2(b). ∎

E The Two Remaining Reductions

The basic idea in all three reductions is to take a label-cover instance and to replace each vertex with a block of q^R vertices, corresponding to the q-ary hypercube $[q]^R$. The intended way to q-color this block is by coloring $x \in [q]^R$ according to x_i where i is the label given to this block. One can think of this coloring as an encoding of the label i. We will essentially prove that any other coloring of this block that uses relatively few colors, can be "list-decoded" into at most t labels from $\{1, \ldots, R\}$. By properly defining edges connecting these blocks, we can guarantee that the lists decoded from two blocks can be used as t-labelings for the label-cover instance.

In the rest of this section, we use the following notation. For a vector $x = (x_1, \ldots, x_n)$ and a permutation π on $\{1, \ldots, n\}$, we define $x^\pi = (x_{\pi(1)}, \ldots, x_{\pi(n)})$.

APPROXIMATE-COLORING$(4, Q)$**:** This reduction is nearly identical to the one above, with the following changes:

- The starting point of the reduction is an instance $G = ((V, E), 2R, \Psi)$ as in Conjecture C.3.

- Each vertex v is replaced by a copy of $\{0, 1, 2, 3\}^{2R}$ (which we still denote $[v]$).

- For every $(v, w) \in E$, let ψ be the $2\leftrightarrow2$-constraint associated with e. By Definition C.1 there are two permutations π_1, π_2 such that $(a, b) \in \psi$ iff $(\pi_1^{-1}(a), \pi_2^{-1}(b)) \in 2\leftrightarrow2$. We now write $[v, w]$ for the following collection of edges. We put an edge (x, y) for $x = (x_1, \ldots, x_{2R}) \in [v]$ and $y = (y_1, \ldots, y_{2R}) \in [w]$ if

$$\forall i \in \{1, \ldots, R\}, \quad T((x_{\pi_1(2i-1)}, x_{\pi_1(2i)}) \leftrightarrow (y_{\pi_2(2i-1)}, y_{\pi_2(2i)})) \neq 0$$

where T is the noise operator from Lemma D.1. Equivalently, we put an edge if $T^{\otimes R}(\overline{x^{\pi_1}} \leftrightarrow \overline{y^{\pi_2}}) \neq 0$.

As before, the reduction outputs the graph $[G] = ([V], [E])$ where $[V]$ is the union of all blocks $[v]$ and $[E]$ is the union of collection of the edges $[v, w]$.

APPROXIMATE-COLORING$(3, Q)$**:** Here again the reduction is nearly identical to the above, with the following changes:

- The starting point of the reduction is an instance of label-cover, as in Conjecture C.4.

- Each vertex v is replaced by a copy of $\{0, 1, 2\}^{2R}$ (which we again denote $[v]$).

- For every $(v, w) \in E$, let π_1, π_2 be the permutations associated with the constraint, as in Definition C.2. Define a collection $[v, w]$ of edges, by including the edge $(x, y) \in [v] \times [w]$ iff

$$\forall i \in \{1, \ldots, R\}, \quad T((x_{\pi_1(2i-1)}, x_{\pi_1(2i)}) \leftrightarrow (y_{\pi_2(2i-1)}, y_{\pi_2(2i)})) \neq 0$$

where T is the noise operator from Lemma D.2. As before, this condition can be written as $T^{\otimes R}(\overline{x^{\pi_1}} \leftrightarrow \overline{y^{\pi_2}}) \neq 0$.

As before, we look at the coloring problem of the graph $[G] = ([V], [E])$ where $[V]$ is the union of all blocks $[v]$ and $[E]$ is the union of collection of the edges $[v, w]$.

E.1 Completeness

APPROXIMATE-COLORING$(4, Q)$**:** Let $\ell : V \rightarrow \{1, \ldots, 2R\}$ be a labeling that satisfies all the constraints in G. We define a legal 4-coloring $c : [V] \rightarrow \{0, 1, 2, 3\}$ as follows. For a vertex $x = (x_1, \ldots, x_{2R}) \in \{0, 1, 2, 3\}^{2R} = [v]$ set $c(x) := x_i$, where $i = \ell(v) \in \{1, \ldots, 2R\}$.

To see that c is a legal coloring, fix any $2\leftrightarrow2$ constraint $(v, w) \in E$ and let π_1, π_2 be the permutations associated with it. Let $i = \ell(v)$ and $j = \ell(w)$, so by assumption $(\pi_1^{-1}(i), \pi_2^{-1}(j)) \in 2\leftrightarrow2$. In other words there is some $k \in \{1, \ldots, R\}$ such that $i \in \{\pi_1(2k - 1), \pi_1(2k)\}$ and $j \in \{\pi_2(2k - 1), \pi_2(2k)\}$. If $x \in [v]$ and $y \in [w]$ share the same color, then $x_i = c(x) = c(y) = y_j$. Since

$$x_i \in \left\{ x_{2k-1}^{\pi_1}, x_{2k}^{\pi_1} \right\} \quad \text{and} \quad y_j \in \left\{ y_{2k-1}^{\pi_2}, y_{2k}^{\pi_2} \right\}$$

we have that the above sets intersect. This, by Lemma D.1, implies that $T^{\otimes R}(\overline{x^{\pi_1}} \leftrightarrow \overline{y^{\pi_2}}) = 0$. So the vertices x, y cannot be adjacent, hence the coloring is legal.

APPROXIMATE-COLORING$(3, Q)$**:** Here the argument is nearly identical to the above. Let $\ell : V \rightarrow \{1, \ldots, 2R\}$ be a labeling that satisfies all of the constraints in G. We define a legal 3-coloring $c : [V] \rightarrow \{0, 1, 2\}$ like before: $c(x) := x_i$, where $i = \ell(v) \in \{1, \ldots, 2R\}$. To see that c is a legal coloring, fix any edge $(v, w) \in E$ and let π_1, π_2 be the permutations associated with the $\bowtie$-constraint. Let $i = \ell(v)$ and $j = \ell(w)$, so by assumption $(\pi_1^{-1}(i), \pi_2^{-1}(j)) \in \bowtie$. In other words there is some $k \in \{1, \ldots, R\}$ such that $i \in \{\pi_1(2k - 1), \pi_1(2k)\}$ and $j \in \{\pi_2(2k - 1), \pi_2(2k)\}$ and not both $i = \pi_1(2k)$ and $j = \pi_2(2k)$. Assume, without loss of generality, that $i = \pi_1(2k - 1)$, so $x_i = x_{2k-1}^{\pi_1}$ and $y_j \in \left\{ y_{2k-1}^{\pi_2}, y_{2k}^{\pi_2} \right\}$.

If $x \in [v]$ and $y \in [w]$ share the same color, then $x_i = c(x) = c(y) = y_j$, so

$$x_{2k-1}^{\pi_1} = x_i = y_j \in \left\{ y_{2k-1}^{\pi_2}, y_{2k}^{\pi_2} \right\} .$$

By Lemma D.2 this implies $T((x_{2k-1}^{\pi_1}, x_{2k}^{\pi_1}) \leftrightarrow (y_{2k-1}^{\pi_2}, y_{2k}^{\pi_2})) = 0$, which means there is no edge between x and y.

E.2 Soundness

APPROXIMATE-COLORING$(4, Q)$**:** We outline the argument and emphasize only the modifications. Assume that $[G]$ contains an independent set $S \subseteq [V]$ whose relative size is at least $1/Q$ and set $\varepsilon = 1/2Q$.

- Let $f_v : \{0, 1, 2, 3\}^{2R} \to \{0, 1\}$ be the characteristic function of S in $[v]$. Define the set $J \subseteq V$ as before and for all $v \in J$, define

$$L(v) = \left\{ i \in \{1, \ldots, 2R\} \;\middle|\; I_i^{\leq 2k}(f_v) \geq \frac{\delta}{2} \right\}$$

  where k, δ are the values given by Corollary 4.6 with ε and the operator T of Lemma D.1. As before, $|J| \geq \varepsilon |V|$ and $\mathbf{E}[f_v] \geq \varepsilon$ for $v \in J$. Now L is a t-labeling with $t = 4k/\delta$. Fix an edge $(v, w) \in E \cap J^2$ and let π_1, π_2 be the associated permutations. Define f, g by $f(x^{\pi_1}) := f_{v_1}(x)$ and $g(y^{\pi_2}) := f_{v_2}(y)$.

- Since S is an independent set, $f(x^{\pi_1}) = f_{v_1}(x) = 1$ and $g(y^{\pi_2}) = f_{v_2}(y) = 1$ implies that x, y are not adjacent, so by construction $T(x^{\pi_1} \leftrightarrow y^{\pi_2}) = 0$. Therefore, $\langle f, Tg \rangle = 0$.

- Now, recalling Definition 2.6, consider the functions $\overline{f}, \overline{g} : (\{0, 1, 2, 3\}^2)^R \to \{0, 1\}$. Applying Corollary 4.6 on $\overline{f}, \overline{g}$ we may deduce the existence of an index $i \in \{1, \ldots, R\}$ for which both $I_i^{\leq k}(\overline{f}) \geq \delta$ and $I_i^{\leq k}(\overline{g}) \geq \delta$. By Claim 2.7, $\delta \leq I_i^{\leq k}(\overline{f}) \leq I_{2i-1}^{\leq 2k}(f) + I_{2i}^{\leq 2k}(f)$, so either $I_{2i-1}^{\leq 2k}(f) \geq \delta/2$ or $I_{2i}^{\leq 2k}(f) \geq \delta/2$. Since the j-th variable in f is the $\pi_1(j)$-th variable in f_{v_1}, this puts either $\pi_1(2i)$ or $\pi_1(2i - 1)$ in $L(v_1)$. Similarly, at least one of $\pi_2(2i), \pi_2(2i - 1)$ is in $L(v_2)$. Thus, there are $a \in L(v_1)$ and $b \in L(v_2)$ such that $(\pi_1^{-1}(a), \pi_2^{-1}(b)) \in 2 \leftrightarrow 2$ so L satisfies the constraint on (v_1, v_2).

We have shown that L satisfies every constraint induced by J, so $\mathrm{isat}_t(G) \geq \varepsilon$.

APPROXIMATE-COLORING$(3, Q)$**:** The argument here is similar to the previous one. The main difference is in the third step, where we replace Corollary 4.6 by the following corollary of Theorem B.2. The corollary follows by letting ε play the role of μ and ν, and using the fact that $\langle F_\varepsilon, U_\rho(1 - F_{1-\varepsilon}) \rangle_\gamma > 0$ whenever $\varepsilon > 0$.

Corollary E.1 *Let T be the operator on $\{0, 1, 2\}^2$ defined in Lemma D.2. For any $\varepsilon > 0$, there exists $\delta > 0, k \in \mathbb{N}$, such that for any functions $f, g : \{0, 1, 2\}^{2R} \to [0, 1]$ satisfying $\mathbf{E}[f] \geq \varepsilon, \mathbf{E}[g] \geq \varepsilon$, there exists some $i \in \{1, \ldots, R\}$ such that either*

$$\min\left(I_{2i-1}^{\leq k}(f), I_{2i-1}^{\leq k}(g)\right) \geq \delta \quad or \quad \min\left(I_{2i-1}^{\leq k}(f), I_{2i}^{\leq k}(g)\right) \geq \delta \quad or \quad \min\left(I_{2i}^{\leq k}(f), I_{2i-1}^{\leq k}(g)\right) \geq \delta.$$

Now we have functions $f_v : \{0, 1, 2\}^{2R} \to \{0, 1\}$, and J is defined as before. Define a labeling

$$L(v) = \left\{ i \in \{1, \ldots, 2R\} \;\middle|\; I_i^{\leq k}(f_v) \geq \delta \right\}$$

where k, δ are the values given by Corollary E.1 with ε. Then L is a t-labeling with $t = k/\delta$.

Let us now show that L is a satisfying t-labeling. Let (v_1, v_2) be a $\bowtie$-constraint with associated permutations π_1, π_2. Define $f(x^{\pi_1}) = f_{v_1}(x), g(x^{\pi_2}) = f_{v_2}(x)$. We apply Corollary E.1 on f, g, and obtain an index $i \in \{1, \ldots, R\}$. Since the j-th variable in f is the $\pi_1(j)$-th variable in f_{v_1}, this puts either $\pi_1(2i)$ or $\pi_1(2i - 1)$ in $L(v_1)$. Similarly, at least one of $\pi_2(2i), \pi_2(2i - 1)$ is in $L(v_2)$. Moreover, we are guaranteed that either $\pi_1(2i - 1) \in L(v_1)$ or $\pi_2(2i - 1) \in L(v_2)$. Thus, there are $a \in L(v_1)$ and $b \in L(v_2)$ such that $(\pi_1^{-1}(a), \pi_2^{-1}(b)) \in \bowtie$ so L satisfies the constraint on (v_1, v_2).

F   Omitted Proofs

F.1   Proof of Claim 2.4

Let us first fix the values of $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$. Then

$$\mathbf{V}_{x_i}[f] = \mathbf{V}_{x_i}\Big[\sum_y \hat{f}(\alpha_y)\alpha_y \Big] = \mathbf{V}_{x_i}\Big[\sum_{y:y_i \neq 0} \hat{f}(\alpha_y)\alpha_y \Big],$$

where the last equality follows from the fact that if $y_i = 0$ then α_y is a constant function of x_i. If $y_i \neq 0$, then the expected value of α_y with respect to x_i is zero. Therefore,

$$\mathbf{V}_{x_i}\Big[\sum_{y:y_i \neq 0} \hat{f}(\alpha_y)\alpha_y \Big] = \mathbf{E}_{x_i}\left[\Big(\sum_{y:y_i \neq 0} \hat{f}(\alpha_y)\alpha_y \Big)^2 \right] = \mathbf{E}_{x_i}\left[\sum_{y,z:y_i \neq 0, z_i \neq 0} \hat{f}(\alpha_y)\hat{f}(\alpha_z)\alpha_y\alpha_z \right].$$

Thus,

$$I_i(f) = \mathbf{E}_x\left[\sum_{y,z:y_i \neq 0, z_i \neq 0} \hat{f}(\alpha_y)\hat{f}(\alpha_z)\alpha_y\alpha_z \right] = \sum_{y,z:y_i \neq 0, z_i \neq 0} \hat{f}(\alpha_y)\hat{f}(\alpha_z)\mathbf{E}_x[\alpha_y\alpha_z] = \sum_{y:y_i \neq 0} \hat{f}^2(\alpha_y),$$

as needed.   ∎

G   Comparison with Khot's Conjectures

Let us first state Khot's original conjectures. For $d \geq 1$, an instance of the weighted bipartite d-to-1 label cover problem is given by a tuple $\Phi = (X, Y, \Psi, W)$. We often refer to vertices in X as *left* vertices and to vertices in Y as *right* vertices. The set Ψ consists of one d-to-1 relation ψ_{xy} for each $x \in X$ and $y \in Y$. More precisely, $\psi_{xy} \subseteq \{1, \ldots, R\} \times \{1, \ldots, R/d\}$ is such that for any $b \in \{1, \ldots, R/d\}$ there are precisely d elements $a \in \{1, \ldots, R\}$ such that $(a, b) \in \psi_{xy}$. The set W includes a non-negative weight $w_{xy} \geq 0$ for each $x \in X$, $y \in Y$. We denote by $w(\Phi, x)$ the sum $\sum_{y \in Y} w_{xy}$ and by $w(\Phi)$ the sum $\sum_{x \in X, y \in Y} w_{xy}$. A *labeling* is a function L mapping X to $\{1, \ldots, R\}$ and Y to $\{1, \ldots, R/d\}$. A constraint ψ_{xy} is *satisfied* by a labeling L if $(L(x), L(y)) \in \psi_{xy}$. Also, for a labeling L, the weight of satisfied constraints, denoted by $w_L(\Phi)$, is $\sum w_{xy}$ where the sum is taken over all $x \in X$ and $y \in Y$ such that ψ_{xy} is satisfied by L. Similarly, we define $w_L(\Phi, x)$ as $\sum w_{xy}$ where the sum is now taken over all $y \in Y$ such that ψ_{xy} is satisfied by L. The following conjectures were presented in [14].

Conjecture G.1 (Bipartite 1-to-1 Conjecture) *For any $\zeta, \gamma > 0$ there exists a constant R such that the following is NP-hard. Given a 1-to-1 label cover instance Φ with label set $\{1, \ldots, R\}$ and $w(\Phi) = 1$ distinguish between the case where there exists a labeling L such that $w_L(\Phi) \geq 1 - \zeta$ and the case where for any labeling L, $w_L(\Phi) \leq \gamma$.*

In the following conjecture, d is any fixed integer greater than 1.

Conjecture G.2 (Bipartite d-to-1 Conjecture) *For any $\gamma > 0$ there exists a constant R such that the following is NP-hard. Given a bipartite d-to-1 label cover instance Φ with label sets $\{1, \ldots, R\}, \{1, \ldots, R/d\}$ and $w(\Phi) = 1$ distinguish between the case where there exists a labeling L such that $w_L(\Phi) = 1$ and the case where for any labeling L, $w_L(\Phi) \leq \gamma$.*

The theorem we prove in this section is the following.

Theorem G.3 *Conjecture 4.4 follows from Conjecture G.1 and Conjecture C.3 follows from Conjecture G.2 for $d = 2$.*[1]

The proof follows by combining Lemmas G.4, G.5, G.7, and G.9. Each lemma presents an elementary transformation between variants of the label cover problem. The first transformation modifies a bipartite label cover instance so that all X variables have the same weight. When we say below that Φ' has the same type of constraints as Φ we mean that the transformation only duplicates existing constraints and hence if Φ consists of d-to-1 constraints for some $d \geq 1$, then so does Φ'.

Lemma G.4 *There exists an efficient procedure that given a weighted bipartite label cover instance $\Phi = (X, Y, \Psi, W)$ with $w(\Phi) = 1$ and a constant ℓ, outputs a weighted bipartite label cover instance $\Phi' = (X', Y, \Psi', W')$ on the same label sets and with the same type of constraints with the following properties:*

- *For all $x \in X'$, $w(\Phi', x) = 1$.*

- *For any $\zeta \geq 0$, if there exists a labeling L to Φ such that $w_L(\Phi) \geq 1 - \zeta$ then there exists a labeling L' to Φ' in which $1 - \sqrt{(1 + \frac{1}{\ell - 1})\zeta}$ of the variables x in X' satisfy that $w_{L'}(\Phi', x) \geq 1 - \sqrt{(1 + \frac{1}{\ell - 1})\zeta}$. In particular, if there exists a labeling L such that $w_L(\Phi) = 1$ then there exists a labeling L' in which all variables satisfy $w_{L'}(\Phi', x) = 1$.*

- *For any $\beta_2, \gamma > 0$, if there exists a labeling L' to Φ' in which β_2 of the variables x in X' satisfy $w_{L'}(\Phi', x) \geq \gamma$, then there exists a labeling L to Φ such that $w_L(\Phi) \geq (1 - \frac{1}{\ell})\beta_2\gamma$.*

Proof: Given Φ as above, we define $\Phi' = (X', Y, \Psi', W')$ as follows. The set X' includes $k(x)$ copies of each $x \in X$, $x^{(1)}, \ldots, x^{(k(x))}$ where $k(x)$ is defined as $\lfloor \ell \cdot |X| \cdot w(\Phi, x) \rfloor$. For every $x \in X$, $y \in Y$ and $i \in \{1, \ldots, k(x)\}$ we define $\psi'_{x^{(i)}y}$ as ψ_{xy} and the weight $w'_{x^{(i)}y}$ as $w_{xy}/w(\Phi, x)$. Notice that $w(\Phi', x) = 1$ for all $x \in X'$ and that $(\ell - 1)|X| \leq |X'| \leq \ell|X|$. Moreover, for any $x \in X$, $y \in Y$, the total weight of constraints created from ψ_{xy} is $k(x)w_{xy}/w(\Phi, x) \leq \ell|X|w_{xy}$.

We now prove the second property. Given a labeling L to Φ that satisfies constraints of weight at least $1 - \zeta$, consider the labeling L' defined by $L'(x^{(i)}) = L(x)$ and $L'(y) = L(y)$. By the property mentioned above, the total weight of unsatisfied constraints in Φ' is at most $\ell|X|\zeta$. Since the total weight in Φ' is at least $(\ell - 1)|X|$, we obtain that the fraction of unsatisfied constraints is at most $(1 + \frac{1}{\ell - 1})\zeta$. Hence, by a Markov argument, we obtain that for at least $1 - \sqrt{(1 + \frac{1}{\ell - 1})\zeta}$ of the X' variables $w_{L'}(\Phi', x) \geq 1 - \sqrt{(1 + \frac{1}{\ell - 1})\zeta}$.

We now prove the third property. Assume we are given a labeling L' to Φ' for which β_2 of the variables have $w_{L'}(\Phi', x) \geq \gamma$. Without loss of generality we can assume that for every $x \in X$, the labeling $L'(x^{(i)})$ is the same for all i. This holds since the constraints between $x^{(i)}$ and the Y variables are the same for all $i \in \{1, \ldots, k(x)\}$. We define the labeling L as $L(x) = L'(x^{(1)})$. The weight of constraints satisfied by L is:

$$\sum_{x \in X} w_L(\Phi, x) \geq \frac{1}{\ell|X|} \sum_{x \in X} k(x) \cdot w_L(\Phi, x)/w(\Phi, x)$$

$$= \frac{1}{\ell|X|} \sum_{x \in X'} w_{L'}(\Phi', x)$$

$$\geq \frac{1}{\ell|X|}\beta_2|X'|\gamma \geq \left(1 - \frac{1}{\ell}\right)\beta_2\gamma$$

---

[1] We in fact show that for any $d \geq 2$, the natural extension of Conjecture C.3 to d-to-d constraints follows from Conjecture G.2 with the same value of d.

where the first inequality follows from the definition of $k(x)$. ✂

The second transformation creates an *unweighted* label cover instance. Such an instance is given by a tuple $\Phi = (X, Y, \Psi, E)$. The multiset E includes pairs $(x, y) \in X \times Y$ and we can think of (X, Y, E) as a bipartite graph (possibly with parallel edges). For each $e \in E$, Ψ includes a constraint, as before. The instances created by this transformation are left-regular, in the sense that the number of constraints $(x, y) \in E$ incident to each $x \in X$ is the same.

Lemma G.5 *There exists an efficient procedure that given a weighted bipartite label cover instance $\Phi = (X, Y, \Psi, W)$ with $w(\Phi, x) = 1$ for all $x \in X$ and a constant ℓ, outputs an unweighted bipartite label cover instance $\Phi' = (X, Y, \Psi', E')$ on the same label sets and with the same type of constraints with the following properties:*

- *All left degrees are equal to $\alpha = \ell|Y|$.*

- *For any $\beta, \zeta > 0$, if there exists a labeling L to Φ such that $w_L(\Phi, x) \geq 1 - \zeta$ for at least $1 - \beta$ of the variables in X, then there exists a labeling L' to Φ' in which for at least $1 - \beta$ of the variables in X, at least $1 - \zeta - 1/\ell$ of their incident constraints are satisfied. Moreover, if there exists a labeling L such that $w_L(\Phi, x) = 1$ for all x then there exists a labeling L' to Φ' that satisfies all constraints.*

- *For any $\beta, \gamma > 0$, if there exists a labeling L' to Φ' in which β of the variables in X have γ of their incident constraints satisfied, then there exists a labeling L to Φ such that for β of the variables in X, $w_L(\Phi, x) > \gamma - 1/\ell$.*

Proof: We define the instance $\Phi' = (X, Y, \Psi', E')$ as follows. For each $x \in X$, choose some $y_0(x) \in Y$ such that $w_{xy_0(x)} > 0$. For every $x \in X$, $y \neq y_0(x)$, E' contains $\lfloor \alpha w_{xy} \rfloor$ edges from x to y associated with the constraint ψ_{xy}. Moreover, for every $x \in X$, E' contains $\alpha - \sum_{y \in Y \setminus \{y_0(x)\}} \lfloor \alpha w_{xy} \rfloor$ edges from x to $y_0(x)$ associated with the constraints $\psi_{xy_0(x)}$. Notice that all left degrees are equal to α. Moreover, for any $x, y \neq y_0(x)$, we have that the number of edges between x and y is at most αw_{xy} and the number of edges from x to $y_0(x)$ is at most $\alpha w_{xy_0(x)} + |Y| = \alpha(w_{xy_0(x)} + 1/\ell)$.

Consider a labeling L to Φ and let $x \in X$ be such that $w_L(\Phi, x) > 1 - \zeta$. Then, in Φ', the same labeling satisfies that the number of incident constraints to x that are satisfied is at least $(1 - \zeta - 1/\ell)\alpha$. Moreover, if $w_L(\Phi, x) = 1$ then all its incident constraints in Φ' are satisfied (this uses that $w_{xy_0(x)} > 0$). Finally, consider a labeling L' to Φ' and let $x \in X$ have γ of their incident constraints satisfied. Then, $w_{L'}(\Phi, x) > \gamma - \frac{1}{\ell}$. ✂

In the third lemma we modify a left-regular unweighted label cover instance so that it has the following property: if there exists a labeling to the original instance that for many variables satisfies many of their incident constraints, then the resulting instance has a labeling that for many variables satisfies *all* their incident constraints. But first, we prove a combinatorial claim.

Claim G.6 *For any integers ℓ, d, R and real $0 < \gamma < \frac{1}{\ell^2 d}$, let $\mathcal{F} \subseteq P(\{1, \ldots, R\})$ be a multiset containing subsets of $\{1, \ldots, R\}$ each of size at most d with the property that no element $i \in \{1, \ldots, R\}$ is contained in more than γ fraction of the sets in $\mathcal{F}$. Then, the probability that a sequence of sets $F_1, F_2, \ldots, F_\ell$ chosen uniformly from $\mathcal{F}$ (with repetitions) is pairwise disjoint is at least $1 - \ell^2 d\gamma$.*

Proof: Note that by the union bound it suffices to prove that $\Pr[F_1 \cap F_2 \neq \emptyset] \leq d\gamma$. This follows by fixing F_1 and using the union bound again:

$$\Pr[F_1 \cap F_2 \neq \emptyset] \leq \sum_{x \in F_1} \Pr[x \in F_2] \leq d\gamma.$$

23

Lemma G.7 *There exists an efficient procedure that given an unweighted bipartite d-to-1 label cover instance $\Phi = (X, Y, \Psi, E)$ with all left-degrees equal to some α, and a constant ℓ, outputs an unweighted bipartite d-to-1 label cover instance $\Phi' = (X', Y, \Psi', E')$ on the same label sets with the following properties:*

- *All left degrees are equal to ℓ.*

- *For any $\beta, \zeta \geq 0$, if there exists a labeling L to Φ such that for at least $1 - \beta$ of the variables in X $1 - \zeta$ of their incident constraints are satisfied, then there exists a labeling L' to Φ' in which $(1 - \zeta)^\ell(1 - \beta)$ of the X' variables have all their ℓ constraints satisfied. In particular, if there exists a labeling L to Φ that satisfies all constraints then there exists a labeling L' to Φ' that satisfies all constraints.*

- *For any $\beta > 0$, $0 < \gamma < \frac{1}{\ell^2 d}$, if in any labeling L to Φ at most β of the variables have γ of their incident constraints satisfied, then in any labeling L' to Φ', the fraction of satisfied constraints is at most $\beta + \frac{1}{\ell} + (1 - \beta)\ell^2 d\gamma$.*

Proof: We define $\Phi' = (X', Y, \Psi', E')$ as follows. For each $x \in X$, consider its neighbors $(y_1, \ldots, y_\alpha)$ listed with multiplicities. For each sequence $(y_{i_1}, \ldots, y_{i_\ell})$ where $i_1, \ldots, i_\ell \in \{1, \ldots, \alpha\}$ we create a variable in X'. This variable is connected to $y_{i_1}, \ldots, y_{i_\ell}$ with the same constraints as x, namely $\psi_{xy_{i_1}}, \ldots, \psi_{xy_{i_\ell}}$. Notice that the total number of variables created from each $x \in X$ is α^ℓ. Hence, $|X'| = \alpha^\ell |X|$.

We now prove the second property. Assume that L is a labeling to Φ such that for at least $1 - \beta$ of the variables in X, $1 - \zeta$ of their incident constraints are satisfied. Let L' be the labeling to Φ' assigning to each of the variables created from $x \in X$ the value $L(x)$ and for each $y \in Y$ the value $L(y)$. Consider a variable $x \in X$ that has $1 - \zeta$ of its incident constraints satisfied and let Y_x denote the set of variables $y \in Y$ such that ψ_{xy} is satisfied. Then among the variables in X' created from x, the number of variables that are connected only to variables in Y_x is at least $\alpha^\ell(1 - \zeta)^\ell$. Therefore, the total number of variables all of whose constraints are satisfied by L' is at least

$$\alpha^\ell(1 - \zeta)^\ell(1 - \beta)|X| = (1 - \zeta)^\ell(1 - \beta)|X'|.$$

We now prove the third property. Assume that in any labeling L to Φ at most β of the X variables have γ of their incident constraints satisfied. Let L' be an arbitrary labeling to Φ'. For each $x \in X$ define $\mathcal{F}_x \subseteq P(\{1, \ldots, R\})$ as the multiset that contains for each constraint incident to x the set of labels to x that, together with the labeling to the Y variables given by L', satisfy this constraint. So $\mathcal{F}_x$ contains α sets, each of size d. Moreover, our assumption above implies that for at least $1 - \beta$ of the variables $x \in X$, no element $i \in \{1, \ldots, R\}$ is contained in more than γ fraction of the sets in $\mathcal{F}_x$. By Claim G.6, for such x, at least $1 - \ell^2 d\gamma$ fraction of the variables in X' created from x have the property that it is impossible to satisfy more than one of their incident constraints simultaneously. Hence, the number of constraints in Φ' satisfied by L' is at most

$$
\begin{aligned}
&\alpha^\ell \cdot \beta \cdot |X| \cdot \ell + \alpha^\ell(1 - \beta)|X|\Big((1 - \ell^2 d\gamma) + (\ell^2 d\gamma) \cdot \ell\Big) \\
&= |X'|\big(\beta\ell + (1 - \beta)(1 - \ell^2 d\gamma) + (1 - \beta)(\ell^2 d\gamma)\ell\big) \\
&\leq |E'|\left(\beta + \frac{1}{\ell} + (1 - \beta)\ell^2 d\gamma\right).
\end{aligned}
$$

The last lemma transforms a bipartite label cover into a non-bipartite label cover. This transformation no longer preserves the constraint type: d-to-1 constraints become d-to-d constraints. We first prove a simple combinatorial claim.

Claim G.8 *Let $A_1, \ldots, A_N$ be a sequence of pairwise intersecting sets of size at most T. Then there exists an element contained in at least N/T of the sets.*

Proof: All sets intersect A_1 in at least one element. Since $|A_1| \leq T$, there exists an element of A_1 contained in at least N/T of the sets. ∎

For the following lemma, recall from Definition 4.2 that a t-labeling labels each vertex with a set of at most t labels. Recall also that a constraint on x, y is satisfied by a t-labeling L if there is a label $a \in L(x)$ and $b \in L(y)$ such that (a, b) satisfies the constraint.

Lemma G.9 *There exists an efficient procedure that given an unweighted bipartite d-to-1 label cover instance $\Phi = (X, Y, \Psi, E)$ on label sets $\{1, \ldots, R\}, \{1, \ldots, R/d\}$, with all left-degrees equal to some ℓ, outputs an unweighted d-to-d label cover instance $\Phi' = (X, \Psi', E')$ on label set $\{1, \ldots, R\}$ with the following properties:*

- *For any $\beta \geq 0$, if there exists a labeling L to Φ in which $1 - \beta$ of the X variables have all their ℓ incident constraints satisfied, then there exists a labeling to Φ' and a set of $1 - \beta$ of the variables of X such that all the constraints between them are satisfied. In particular, if there exists a labeling L to Φ that satisfies all constraints then there exists a labeling L' to Φ' that satisfies all constraints.*

- *For any $\beta > 0$ and integer t, if there exists a t-labeling L' to Φ' and a set of β variables of X such that all the constraints between them are satisfied, then there exists a labeling L to Φ that satisfies at least β/t^2 of the constraints.*

Proof: For each pair of constraints $(x_1, y), (x_2, y) \in E$ that share a Y variable we add one constraint $(x_1, x_2) \in E'$. This constraint is satisfied when there exists a labeling to y that agrees with the labeling to x_1 and x_2. More precisely,

$$\psi'_{x_1 x_2} = \Big\{ (a_1, a_2) \in \{1, \ldots, R\} \times \{1, \ldots, R\} \,\Big|\, \exists b \in \{1, \ldots, R/d\} \, (a_1, b) \in \psi_{x_1 y} \wedge (a_2, b) \in \psi_{x_2 y} \Big\}.$$

Notice that if the constraints in Ψ are d-to-1 then the constraints in Ψ' are d-to-d.

We now prove the second property. Let L be a labeling to Φ and let $C \subseteq X$ be of size $|C| \geq (1 - \beta)|X|$ such that all constraints incident to variables in C are satisfied by L. Consider the labeling L' to Φ' given by $L'(x) = L(x)$. Then, we claim that L' satisfies all the constraints in Φ' between variables of C. Indeed, take any two variables $x_1, x_2 \in C$ with a constraint between them. Assume the constraint is created as a result of some $y \in Y$. Then, since $(L(x_1), L(y)) \in \psi_{x_1 y}$ and $(L(x_2), L(y)) \in \psi_{x_2 y}$, we also have $(L(x_1), L(x_2)) \in \psi'_{x_1 x_2}$.

It remains to prove the third property. Let L' be a t-labeling to Φ' and let $C \subseteq X$ be a set of variables of size $|C| \geq \beta|X|$ with the property that any constraint between variables of C is satisfied by L'. We first define a t-labeling L'' to Φ as follows. For each $x \in X$, we define $L''(x) = L(x)$. For each $y \in Y$, we define $L''(y) \in \{1, \ldots, R/d\}$ as the label that maximizes the number of satisfied constraints between C and y. We claim that for each $y \in Y$, L'' satisfies at least $1/t$ of the constraints between C and y. Indeed, for each constraint between C and y consider the set of labels to y that satisfy it. These sets are pairwise intersecting since all constraints in Φ' between variables of C are satisfied by L'. Moreover, since Φ is a

25

d-to-1 label cover, these sets are of size at most t. Claim G.8 asserts the existence of a labeling to y that satisfies at least $1/t$ of the constraints between C and y. Since at least β of the constraints in Φ are incident to C, we obtain that L'' satisfies at least β/t of the constraints in Φ.

To complete the proof, we define a labeling L to Φ by $L(y) = L''(y)$ and $L(x)$ chosen uniformly from $L''(x)$. Since $|L''(x)| \le t$ for all x, the expected number of satisfied constraints is at least β/t^2, as required.
∎