CSP Gaps and Reductions in the Lasserre Hierarchy

Madhur Tulsiani[*]

November 22, 2008

Abstract

We study integrality gaps for SDP relaxations of constraint satisfaction problems, in the hierarchy of SDPs defined by Lasserre. Schoenebeck [25] recently showed the first integrality gaps for these problems, showing that for MAX k-XOR, the ratio of the SDP optimum to the integer optimum may be as large as 2 even after $\Omega(n)$ rounds of the Lasserre hierarchy.

We show that for the general MAX k-CSP problem over binary domain, the ratio of SDP optimum to the value achieved by the optimal assignment, can be as large as $2^k/2k - \epsilon$ even after $\Omega(n)$ rounds of the Lasserre hierarchy. For alphabet size q which is a prime, we give a lower bound of $q^k/q(q-1)k - \epsilon$ for $\Omega(n)$ rounds. The method of proof also gives optimal integrality gaps for a predicate chosen at random.

We also explore how to translate gaps for CSP into integrality gaps for other problems using reductions, and establish SDP gaps for Maximum Independent Set, Approximate Graph Coloring, Chromatic Number and Minimum Vertex Cover. For Independent Set and Chromatic Number, we show integrality gaps of $n/2^{O(\sqrt{\log n \log \log n})}$ even after $2^{\Omega(\sqrt{\log n \log \log n})}$ rounds. In case of Approximate Graph Coloring, for every constant l, we construct graphs with chromatic number $\Omega(2^{l/2}/l^2)$, which admit a vector l-coloring for the SDP obtained by $\Omega(n)$ rounds. For Vertex Cover, we show an integrality gap of 1.36 for $\Omega(n^\delta)$ rounds, for a small constant δ.

The results for CSPs provide the first examples of $\Omega(n)$ round integrality gaps matching hardness results known only under the Unique Games Conjecture. This and some additional properties of the integrality gap instance, allow for gaps for in case of Independent Set and Chromatic Number which are stronger than the NP-hardness results known even under the Unique Games Conjecture.

Keywords: semidefinite programming, integrality gaps, constraint satisfaction

1 Introduction

Semidefinite Programming (SDP) hierarchies, such as the ones defined by Lovász and Schrijver [23] and Lasserre [21] provide a nice way to formalize the following question: what is the best approximation that can be achieved for a problem using "stronger and stronger" semidefnite programs? Since SDPs are known to provide some of the best known approximation algorithms for many optimization problems, this is a natural question to consider.

Starting from a basic integer program for the combinatorial problem, in which variables take values only 0 and 1, these hierarchies define a sequence of SDP relaxations (known as "rounds" or "levels" of the hierarchy) for the problem, which are increasingly more constrained. A program obtained by by r rounds can be solved in $n^{O(r)}$ time and hence algorithms obtained by a constant number of rounds run in polynomial time. Also, on a starting integer program with n variables, the SDP obtained by n rounds of the hierarchy gives the true optimum of the combinatorial problem (but may take exponential time to solve).

Most currently known SDP based algorithms can be derived by a constant number of rounds of the Lasserre hierarchy (which is the stronger of the two - see [22] for a comparison). It was also shown by Chlamtac and Singh [9], that for some problems, such as Hypergraph Indpendent Set, the approximation quality improves as one considers more and more rounds of the Lasserre hierarchy.

It is then of interest to quantify the tradeoff between approximation achieved by the SDP and the number of rounds of the hierarchy. The quality of approximation (for say, a maximization problem) achieved by the SDP is measured by the the ratio of SDP optimum to the optimum of the original combinatorial problem, which is known as the integrality gap [1] of the SDP. Lower bounds showing that the integrality gap remains large after many rounds of the hierarchy can be viewed as unconditional lower bounds in a strong computational model, which captures most known SDP algorithms.

Problems studied

Taking the analogy of the computational model a bit further, we study if integrality gaps for one problem can be used to show an integrality gap for another problem using reductions (see proof overview for the structure of reductions for integrality gaps). We remark that reductions between integrality gaps were considered before, for example by [1] and [25] - we merely show how to do somewhat more complicated ones. We use the arguments from known reductions in hardness of approximation literature, which usually start from the hardness of a constraint satisfaction problem and use it to conclude the hardness of another problem. To this end, we study the following problems:

1. MAX k-CSP: This is the problem of finding an assignment to binary variables $x_1, \ldots, x_n$ to satisfy the maximum number of constraints in a given set $C_1, \ldots, C_m$. Each constraint is a boolean function of at most k of the variables and is said to be satisfied when it attains value 1. When the domain of the variables is is allowed to be of size q, we denote the problem by MAX k-CSP$_q$.

   Charikar, Makarychev and Makarychev [7] give an SDP based algorithm for this problem which provably satisfies a ck/q^k fraction of the constraints satisfied by the optimal assignment, for a

---

[1]Unfortunately, there are different conventions regarding the definition of the integrality gap (ratio of SDP optimum to integer optimum and vice-versa). In this paper, we always define integrality gap so that it is a number bigger than 1. Hence, the larger the gap, the farther the SDP is from the true optimum.

constant $c > 0.44$. It was shown by Arora, Alekhnovich and Tourlakis [1] that the integrality gap for the MAX k-SAT problem, which is a special case of MAX k-CSP, remains $2^k/(2^k - 1)$ even after $\Omega(n)$ rounds of the Lovász-Schrijver hierarchy. Schoenebeck [25] recently showed a gap of factor 2 for MAX k-XOR for $\Omega(n)$ and also strengthened the MAX k-SAT results for $\Omega(n)$ rounds of Lasserre.

2. **Maximum Independent Set:** This is the problem of finding the largest set of vertices in a graph, not containing any edges. The best known approximation algorithm by Boppana and Haldórsson [5] achieves an approximation ratio of $O\left(\frac{n}{(\log n)^2}\right)$. Also, Feige [13] showed that the integrality gap for the Lovász ϑ-function, which is an SDP relaxation equivalent to 1 round of Lasserre, is at least $n/2^{c\sqrt{\log n}}$ for some constant c.

3. **Approximate Graph Coloring:** This is the problem of coloring a graph (with different colors for adjacent vertices) with minimum number of colors, when the graph is *known* to be colorable with a small constant number of colors. The best known algorithm, due to Chlamtac [8] colors a 3-colorable graph with at most $n^{0.2072}$ colors.

4. **Chromatic Number:** This is the general problem of finding the minimum number of colors for coloring a graph when no guarantee as the one above is given. For this problem, Feige, Langberg and Schechtman [15] show that a gap can be as large as $n/\mathrm{polylog}(n)$ for an SDP relaxation which is weaker than 1 round of Lasserre.

5. **Minimum Vertex Cover:** In this problem, it is required to find the smallest possible subset of vertices in a graph, which touches every edge. Integrality gap of a factor 7/6 for $\Omega(n)$ rounds of the Lovász-Schrijver SDP hierarchy were shown by Schonebeck, Trevisan and Tulsiani [26], and later strengthened to the Lasserre hierarchy by Schoenebeck. An integrality gap of factor $2 - \epsilon$ was also shown by for $\Omega(\sqrt{\log n/\log \log n})$ rounds of the Lovász-Schrijver hierarchy by Georgiou et. al. [16].

Our Results

We prove strong integrality gaps for MAX k-CSP and develop techniques to carry out PCP based reductions between SDPs, to obtain gaps for the other problems above. We present a summary of the known NP-hardness results and the integrality gap results we obtain, in the table below. UG-hardness denotes the hardness assuming the Unique Games Conjecture. For Approximate Graph Coloring the hardness mentions the tradeoff for a graph known to be l-colorable.

	NP-hardness		UG-hardness		Integrality Gap	No. of Rounds
MAX k-CSP	$\frac{2^k}{2^{\sqrt{2k}}}$	[11]	$\frac{2^k}{k+o(k)}$	[2]	$\frac{2^k}{2k}$	$\Omega(n)$
MAX k-CSP$_q$ (for prime q)			$\frac{q^k}{q(q-1)k}$	[2]	$\frac{q^k}{q(q-1)k}$	$\Omega(n)$
Maximum Independent Set	$\frac{n}{2^{(\log n)^{3/4+\epsilon}}}$	[19]			$\frac{n}{2^{c_1\sqrt{\log n \log \log n}}}$	$2^{c_2\sqrt{\log n \log \log n}}$
Approximate Graph Coloring	l vs. $2^{\frac{1}{25}\log^2 l}$	[18]			l vs. $\frac{2^{l/2}}{4l^2}$	$\Omega(n)$
Chromatic Number	$\frac{n}{2^{(\log n)^{3/4+\epsilon}}}$	[19]			$\frac{n}{2^{c_1\sqrt{\log n \log \log n}}}$	$2^{c_2\sqrt{\log n \log \log n}}$
Minimum Vertex Cover	1.36	[10]	2 - ϵ	[20]	1.36	$\Omega(n^\delta)$

Remark 1.1 It was pointed out by Luca Trevisan that for Maximum Independent Set and Chromatic Number the integrality gap can be shown to be *at most* n/t after t rounds. Hence, the number of rounds in the above results for these problems is optimal up to a polynomial.

Organization of the paper

We present in section 3 an overview of the main ideas in the proofs. Sections 4, 5, 6 and 7 contain the gaps for CSPs, Independent Set, Graph Coloring and Vertex Cover respectively. The integrality gaps for graph problems only use the results from section 4, and may be read only knowing the statement of Corollary 4.8. The gaps for Graph Coloring and Vertex Cover are independent of each other, but depend on the gaps for Independent Set.

2 Preliminaries and notation

2.1 Constraint Satisfaction Problems

For an instance of Φ MAX k-CSP$_q$, we denote the variables by $\{x_1, \ldots, x_n\}$, their domain $\{0, \ldots, q-1\}$ by $[q]$ and the constraints by $C_1, \ldots, C_m$. Each constraint is a function of the form $C_i : [q]^{T_i} \to \{0,1\}$ depending only on the values of the variables in an ordered tuple[2] T_i with $|T_i| \leq k$. We denote the number of constraints satisfied by the best assignment by $\mathsf{OPT}(\Phi)$.

For a given set $S \subseteq [n]$, we denote by $[q]^S$ the set of all mappings from the set S to $[q]$. In context of variables, these mappings can be understood as partial assignments to a given subset of variables. For $\alpha \in [q]^S$, we denote its projection to $S' \subseteq S$ as $\alpha(S')$. Also, for $\alpha_1 \in [q]^{S_1}, \alpha_2 \in [q]^{S_2}$ such that $\alpha_1(S_1 \cap S_2) = \alpha_2(S_1 \cap S_2)$, we denote by $\alpha_1 \circ \alpha_2$ the assignment over $S_1 \cup S_2$ defined by α_1 and α_2. Hence, $(\alpha_1 \circ \alpha_2)(j)$ equals $\alpha_1(j)$ for $j \in S_1$ and $\alpha_2(j)$ for $j \in S_2 \setminus S_1$. We only use the notation $\alpha_1 \circ \alpha_2$ when it is well defined for $\alpha_1, \alpha_2, S_1, S_2$.

We shall prove results for constraint satisfaction problems where every constraint is specified by the same predicate $P : [q]^k \to \{0,1\}$. We denote the set of inputs which the predicate accepts (outputs 1 on) by $P^{-1}(1)$. To generate an instance of the problem each constraint is of the form of P applied to a k-tuple of *literals*. For variable x over domain $[q]$, we can generalize the notion of a literal as $x + a$ (computed modulo q) for $a \in [q]$.

Definition 2.1 *For a given $P : [q]^k \to \{0,1\}$, an instance Φ of MAX k-CSP(P) is a set of constraints $C_1, \ldots, C_m$ where each constraint C_i is over a k-tuple of variables $T_i = \{x_{i_1}, \ldots, x_{i_k}\}$ and is of the form $P(x_{i_1} + a_{i_1}, \ldots, x_{i_k} + a_{i_k})$ for some $a_{i_1}, \ldots, a_{i_k} \in [q]$.*

Given a predicate P, we will consider a random instance Φ of the MAX k-CSP(P) problem. To generate a random instance with m constraints, for every constraint C_i, we randomly select a k-tuple of distinct variables $T_i = \{x_{i_1}, \ldots, x_{i_k}\}$ and $a_{i_1}, \ldots, a_{i_k} \in [q]$, and put $C_i \equiv P(x_{i_1} + a_{i_1}, \ldots, x_{i_k} + a_{i_k})$. It is well known and used in various works on integrality gaps and proof complexity (e.g. [6], [1], [26] and [25]), that random instances of CSPs are highly unsatisfiable and at the same time highly expanding i.e. for every set of constraints which is not too large, most variables occur only in one constraint. We capture the properties we need in the lemma below. A proof is provided in the appendix for the sake of completeness.

Lemma 2.2 *Let $\epsilon, \delta > 0$ and a predicate $P : [q]^k \to \{0,1\}$ be given. Then there exist $\beta = O(q^k \log q/\epsilon^2)$, $\eta = \Omega((1/\beta)^{5/\delta})$ and $N \in \mathbb{N}$, such that if $n \geq N$ and Φ is a random instance of MAX k-CSP(P) with $m = \beta n$ constraints, then with probability $1 - o(1)$*

    1. $\mathsf{OPT}(\Phi) \leq \frac{|P^{-1}(1)|}{q^k}(1 + \epsilon) \cdot m.$

---

[2]We will often ignore the order of the variables in T_i and also refer to T_i as a set of variables.

2. For all $s \leq \eta n$, every set of s constraints involves at least $(k-1-\delta)s$ variables.

The instances we will mostly be concerned with, are described by systems of equations over finite fields. Let q be prime and $A \in (\mathbb{F}_q)^{d \times k}$ be a matrix with $rank(A) = d \leq k$. We define the predicate $P_A : [q]^k \to \{0, 1\}$ such that

$$P_A(x_1, \ldots, x_k) = 1 \quad \Leftrightarrow \quad A \cdot (x_1 \ldots x_k)^\mathsf{T} = 0$$

To generate a constraint C_i in an instance of MAX k-CSP(P_A), we consider $P_A(x_{i_1} + a_{i_1}, \ldots, x_{i_k} + a_{i_k})$ which is 1 iff $A \cdot (x_{i_1} \ldots x_{i_k})^\mathsf{T} = b^{(i)}$ for $b^{(i)} = -A \cdot (a_{i_1} \ldots a_{i_k})^\mathsf{T}$. We define the problem MAX k-CSP(P_A) below which is a special case of the MAX k-CSP(P) problem.

Definition 2.3 *For a given $A \in (\mathbb{F}_q)^{d \times k}$, an instance Φ of MAX k-CSP(P_A) is a set of constraints $C_1, \ldots, C_m$ where each constraint C_i is over a k-tuple $T_i = \{x_{i_1}, \ldots, x_{i_k}\}$ and is of the form $A \cdot (x_{i_1}, \ldots, x_{i_k})^\mathsf{T} = b^{(i)}$ for some $b^{(i)} \in (\mathbb{F}_q)^d$.*

2.2 Linear Codes

Recall that a linear code of distance 3 and length k over $\mathbb{F}_q$ is a subspace of $(\mathbb{F}_q)^k$ such that every non-zero vector in the subspace has at least 3 non-zero coordinates. We shall prove our results for predicates P_A where A is a matrix whose rows form a basis for such a code. Such a matrix is called the generator matrix of the code. To get the optimal bounds, we shall use Hamming codes which have the largest dimension for the fixed distance 3. We refer to the code below as Hamming code of length k.

Fact 2.4 *Let $(q^{r-1} - 1)/(q-1) < k \leq (q^r - 1)/(q-1)$. Then there exists a linear code of length k and distance 3 over $\mathbb{F}_q$, with dimension $k - r$.*

2.3 The Lasserre hierarchy

The Lasserre hierarchy gives a sequence of increasingly tight semidefinite programming relaxations for a quadratic integer program for variables taking values 0 and 1. The semidefinite program after t rounds of the hierarchy introduces a new variable for the product of every t variables in the original program. Thus, if the original program had n variables and polynomial number of constraints, the new program has size $n^{(O(t))}$.

We first consider the program for Maximum Independent Set. Given a graph $G = (V, E)$, the integer program would have a variable X_i for each $i \in V$ with $X_i = 1$ if i is in the independent set and 0 otherwise. To ensure that the solution is an independent set, we would enforce that $X_i X_j = 0$ for all $(i, j) \in E$. To obtain the Lasserre relaxation, we first think of a an integer program which has a variable X_S for each $S \subseteq V, |S| \leq t$ where the intended solution in $X_S = 1$ iff all vertices in S are in the independent set. We can then add the constraint that the product $X_{S_1} \cdot X_{S_2}$ must only depend on $S_1 \cup S_2$. For homogenization, we introduce an extra variable $X_\emptyset$ which is always supposed to be 1. Replacing the integer variables X_S by vectors $\mathbf{U}_S$ gives the semidefinite relaxation as below. Note that the program for t rounds only has vectors for sets of size at most t. It can be shown that for any set S with $|S| \leq t$, the vectors $\mathbf{U}_{S'}, S' \subseteq S$ induce a probability distribution over valid independent sets of the subgraph induced by S.

Similarly, we can write down a relaxation for the MAX k-CSP$_q$ problem. Note that an integer solution to the problem will be given by a single mapping $\alpha_0 \in [q]^{[n]}$, which is an assignment to

---

Lasserre SDP for Maximum Independent Set

$$\text{maximize} \qquad \sum_{i \in V} \left\| \mathbf{U}_{\{i\}} \right\|^2$$

$$\text{subject to} \qquad \left\langle \mathbf{U}_{\{i\}}, \mathbf{U}_{\{j\}} \right\rangle = 0 \qquad\qquad\qquad\qquad \forall \, (i,j) \in E$$

$$\left\langle \mathbf{U}_{S_1}, \mathbf{U}_{S_2} \right\rangle = \left\langle \mathbf{U}_{S_3}, \mathbf{U}_{S_4} \right\rangle \qquad\qquad \forall \, S_1 \cup S_2 = S_3 \cup S_4$$

$$\left\langle \mathbf{U}_{S_1}, \mathbf{U}_{S_2} \right\rangle \in [0,1] \qquad\qquad\qquad\qquad \forall S_1, S_2$$

$$\left\| \mathbf{U}_\emptyset \right\| = 1$$

---

all the variables. Using this, we can define $0/1$ variables $X_{(S,\alpha)}$ for each $S \subseteq [n]$ such that $|S| \leq t$ and $\alpha \in [q]^S$. The intended solution is $X_{(S,\alpha)} = 1$ if $\alpha_0(S) = \alpha$ and 0 otherwise. As before, we introduce $X_{(\emptyset,\emptyset)}$ which is intended to be 1. Replacing them by vectors gives the SDP relaxation. Note that we denote the vectors corresponding to a set of variables and a partial assignment by $\mathbf{V}$ and a vector for a set (as above) by $\mathbf{U}$ to avoid confusion.

Lasserre SDP for MAX k-CSP$_q$

$$\text{maximize} \qquad \sum_{i=1}^{m} \sum_{\alpha \in [q]^{T_i}} C_i(\alpha) \left\| \mathbf{V}_{(T_i, \alpha)} \right\|^2$$

$$\text{subject to} \quad \left\langle \mathbf{V}_{(S_1,\alpha_1)}, \mathbf{V}_{(S_2,\alpha_2)} \right\rangle = 0 \qquad\qquad\qquad \forall \, \alpha_1(S_1 \cap S_2) \neq \alpha_2(S_1 \cap S_2)$$

$$\left\langle \mathbf{V}_{(S_1,\alpha_1)}, \mathbf{V}_{(S_2,\alpha_2)} \right\rangle = \left\langle \mathbf{V}_{(S_3,\alpha_3)}, \mathbf{V}_{(S_4,\alpha_4)} \right\rangle \quad \forall \, S_1 \cup S_2 = S_3 \cup S_4, \; \alpha_1 \circ \alpha_2 = \alpha_3 \circ \alpha_4$$

$$\sum_{j \in [q]} \left\| \mathbf{V}_{(\{i\},j)} \right\|^2 = 1 \qquad\qquad\qquad\qquad \forall i \in [n]$$

$$\left\langle \mathbf{V}_{(S_1,\alpha_1)}, \mathbf{V}_{(S_2,\alpha_2)} \right\rangle \geq 0 \qquad\qquad\qquad\qquad \forall S_1, S_2, \alpha_1, \alpha_2$$

$$\left\| \mathbf{V}_{(\emptyset,\emptyset)} \right\| = 1$$

For any set S with $|S| \leq t$, the vectors $\mathbf{V}_{(S,\alpha)}$ induce a probability distribution over $[q]^S$ such that the assignment $\alpha \in [q]^S$ appears with probability $\left\| \mathbf{V}_{(S,\alpha)} \right\|^2$. The constraints can be understood by thinking of valid solution as coming from a distribution of assignments for all the variable and of $\left\langle \mathbf{V}_{(S_1,\alpha_1)}, \mathbf{V}_{(S_2,\alpha_2)} \right\rangle$ as the probability of the event that variables in S_1 get value according to α_1 and those in S_2 according to α_2. The last two constraints simply state the fact that a probability must be positive and the probability of the event which does not restrict any variable to anything must be 1.

The first constraint says that the probability of a variable simultaneously getting two different values must be 0. The second one says that if we calculate the probability that all variables in $(S_1 \cup S_2)$ $(= S_3 \cup S_4)$ get values according to $\alpha_1 \circ \alpha_2$ $(= \alpha_3 \circ \alpha_4)$ in two different ways, it must still be the same. In the third constraint, $\mathbf{V}_{(\{i\},j)}$ denotes the map which assigns value to j to variable x_i. The constraint can be understood as saying that the variable x_i must take exactly one value. This constraint is sometimes also written as $\sum_{i \in [q]} \mathbf{V}_{(\{i\},j)} = \mathbf{V}_{(\emptyset,\emptyset)}$, which is equivalent as can be seen by noting that $\left\| \sum_{i \in [q]} \mathbf{V}_{(\{i\},j)} - \mathbf{V}_{(\emptyset,\emptyset)} \right\|^2 = \sum_{j \in [q]} \left\| \mathbf{V}_{(\{i\},j)} \right\|^2 - 1$.

For a graph G, we denote the optimum of the independent set SDP by $\mathsf{FRAC}(G)$, and for the CSP instance Φ we denote it by $\mathsf{FRAC}(\Phi)$. The integrality gap of both the SDPs is then given by (SDP optimum)/(integer optimum). Note that according to our convention, the integrality gap is always at least 1.

3   Overview of proofs

3.1   CSP Gaps

For proving integrality gaps for CSPs, our starting point is the result by Schoenebeck [25]. Even though the result is stated as an integrality gap for MAX-k-XOR, it is useful to view the argument as having the following two parts:

1. Given a system of equations over $\mathbb{F}_2$ with no "small contradictions", it shows how to construct vectors $\mathbf{V}_{(S,\alpha)}$ (satisfying consistency conditions), such that for every set S of variables with $|S| \leq t$, $\left\| \mathbf{V}_{(S,\alpha)} \right\| > 0$ if and only if α satisfies all equations involving variables from S.

2. If one considers a random instance Φ of a CSP, with every constraint being a linear equation on at most k variables (MAX k-XOR), then the system obtained by combining them has no "small contradictions".

It can be shown that the first step implies $\mathsf{FRAC}(\Phi) = m$ (m is the number of constraints) even after t rounds. On the other hand, in a system of equations with each chosen randomly, only about half are satisfiable by any assignment. Hence, one gets an integrality gap of factor 2, for t rounds. Here t depends on the size of "small contradictions" for step 2 and can be chosen to be $\Omega(n)$.

We note that every constraint does not *have* to be a single equation for second step to work. In particular, we consider each constraint to be the form $A \cdot (x_1, \ldots, x_k)^{\mathsf{T}} = b$ for some $A \in (\mathbb{F}_2)^{d \times k}$, which is the same for all constraints and $b \in \mathbb{F}_2^d$. The constraint is said to be satisfied only when all these d equations are satisfied. Now, if A is full rank, then one can show that in a random CSP instance Φ, only about $1/2^d$ fraction of the constraints are satisfiable by any assignment.

Note that all equations obtained by combining all the ones in the constraints are no longer independent (the ones in each constraint are correlated because of a fixed choice of A). However, one can still show that if A satisfies some extra properties, like any linear combination of the d equations given by A involves at least 3 variables (i.e. A is the generator matrix of a distance 3 code over $\mathbb{F}_2$), then the conclusion in the second step can still be made. Step 1 still allows us to conclude that $\mathsf{FRAC}(\Phi) = m$, thereby obtaining an integrality gap of factor 2^d. Optimizing over d gives the claimed results for MAX k-CSP. We also generalize the first step, to work for equations over arbitraty prime fields $\mathbb{F}_q$, to obtain a gap for MAX k-CSP$_q$.

3.2   Reductions

Consider a reduction from a constraint satisfaction problem, to another problem, say Maximum Independent Set for concreteness. Starting from a CSP instance Φ, this reduction creates a graph G_Φ and one needs to argue the following two things:

- *Completeness*: If Φ has an assignment satisfying many constraints, then G_Φ has a large independent set.

- *Soundness*: If Φ has no good assignment, then G_Φ has no large independent sets.

If Φ is an integrality gap instance for an SDP, then Φ has no good assignment but has a good SDP solution. Showing that G_Φ is an integrality gap instance, amounts to making the following two claims simultaneously:

- *Vector Completeness*: Since Φ has a good SDP solution, so does G_Φ.

- *Soundness*: Since Φ has no good assignment, G_Φ has no large independent sets.

Notice that if we are using a known NP-hardness reduction, then the soundness condition is already available. Showing an integrality gap reduces to generalizing "completeness" to "vector completeness". We do this by giving various transformation, that transform an SDP solution for Φ, into one for the problem we are reducing to.

3.2.1   Maximum Independent Set

The transformations for independent set are conceptually the simplest and form a basis for our other results as well. We consider the FGLSS graph G_Φ obtained from Φ, which has vertices of the form (C_i, α), where C_i is a constraint and α_i is a partial assignment to variables in C_i. Vertices corresponding to contradicting assignments are connected.

Since Φ has an SDP solution for t rounds of Lasserre (say for a large t), we have vectors $\mathbf{V}_{(S,\alpha)}$ where S is a set of at most t variables and α is a partial assignment to those variables. We need to produce vectors $\mathbf{U}_{\mathcal{S}}$ where $\mathcal{S}$ is a set of vertices in the FGLSS graph. However, a set of vertices is simply a set of constraints and partial assignments to all the variables involved. Let S' be the set of all variables in all the constraints in $\mathcal{S}$ and let α' be the joint partial assignment defined by all vertices in $\mathcal{S}$ (assuming for now, that no partial assignments in $\mathcal{S}$ contradict). We take $\mathbf{U}_{\mathcal{S}} = \mathbf{V}_{(S', \alpha')}$.

The reduction (by [3]) proceeds by taking products of the graph G_Φ to get G_Φ^r and randomly sampling a certain vertex-induced subgraph. It turns out to be sufficient however, to create an SDP solution for G_Φ^r. Each vertex in G_Φ^r is an r-tuple of vertices in G_Φ with an edge between two r-tuples if vertices in any of the r coordinates are adjacent in G_Φ. A set $\overline{\mathcal{S}}$ of vertices in G_Φ^r is a set of r-tuples and we consider sets $\mathcal{S}_1, \ldots, \mathcal{S}_r$ where $\mathcal{S}_j$ projection of $\overline{\mathcal{S}}$ to the jth coordinate. For G_Φ^r, we simply take $\mathbf{U}_{\overline{\mathcal{S}}} = \mathbf{U}_{\mathcal{S}_1} \otimes \ldots \otimes \mathbf{U}_{\mathcal{S}_r}$. This corresponds to the intuition that every independent set in G_Φ^r corresponds to picking one independent set in each copy of G_Φ.

3.2.2   Approximate Graph Coloring **and** Chromatic Number

To obtain a gap for Approximate Graph Coloring, we modify the FGLSS reduction slightly. The gap for Chromatic Number is derived from this by taking graph products and tensoring vectors as before. The modified reduction below is in the spirit of randomized PCPs of Feige and Killian [14].

Consider an instance Φ when the constraints are known to be of the type $A \cdot (x_1, \ldots, x_k)^\mathsf{T} = b$ and consider G_Φ as before. Supposing that we had an assignment satisfying all constraints, this would give a large independent set. For the the graph to be l colorable, we need l independent sets covering the graph. Let l be the nullity of the matrix A and consider the vectors $w_1, \ldots, w_l$ such that $A \cdot w_l = 0$. If α is a partial assignment to variables $(x_1, \ldots, x_k)$ which satisfies the above constraint, then so is $\alpha + w_j$ for $1 \le j \le l$. [3]

The problem is this does not give us an independent set. If $x_1, \ldots, x_n$ was an assignment, and we looked at the restriction of this assignment to every constraint and added w_j to every restriction, this does not give a consistent assignment to $x_1, \ldots, x_n$. However, this *is* an independent set if we

---

[3]A reader familiar with [14] may recognize this as an attempt to get a zero-knowledge protocol for showing that Φ is satisfiable. However, we do not argue the soundness of this protocol - we simply enforce it in the choice of Φ.

slightly modify the FGLSS graph. Every constraint C_i is on an *ordered tuple* T_i of vertices. For assignment α_1 to T_{i_i} and α_2 to T_{i_2}, we connect them only if α_1, α_2 differ on a variable in the *same coordinate* in T_{i_1}, T_{i_2}. One can then verify that the transformation above then takes us from one independent set to another. Changing the graph may affect the soundness, but it is possible to choose Φ so that the resulting graph still has no large independent sets.

Using this intuition, we now need to produce a vector every vertex and every color (and sets of vertices and sets of colors). For a vertex (C_i, α) in G_Φ, we take the vectors corresponding to the l colors as $\mathbf{V}_{(T_i, \alpha+w_1)}, \ldots, \mathbf{V}_{(T_i, \alpha+w_l)}$.

3.2.3 Minimum Vertex Cover

For Minimum Vertex Cover, we use the reduction by Dinur and Safra [10], which is a little complicated to describe. However, an interesting point comes up in analyzing a (small) step which is not "local", as opposed to all the steps in the previous reductions. The step is a simple Chernoff bound in the completeness part, showing how large independent sets intersect certain other sets of vertices in a graph. Generalizing this to "vector-completeness" seems to require analyzing the "local distributions" defined by the vectors, and combining the bounds globally using properties of the SDP solution. Another component of the proof is a transformation of vectors for long-code based reductions.

4 Gaps for k-CSPs

4.1 Lasserre vectors for linear equations

We first state a generalization of Schoenebeck's result [25] that we shall need. As mentioned before, it is more convenient to view it as constructing vectors for a system of linear equations. We show that if the width-t bounded resolution (defined below) of the system of equations cannot derive a contradiction then there exists vectors for sets of size up to $t/2$ which satisfy all consistency constraints in the Lasserre relaxation. Also these vectors "satisfy" all the equations in the sense that if we think of $\left\| \mathbf{V}_{(S,\alpha)} \right\|^2$ as the probability that the variables in set S get the assignment α, then for any set S, the only assignments having non-zero probability are the ones which satisfy all the equations involving variables from S.

We write a linear equation in variables $\{x_1, \ldots, x_n\}$ as $\omega \cdot x = r$ where ω is a vector of coefficients and x is the vector containing all variables. We also denote by $\overrightarrow{0}$ the coefficient vector with all coefficients equal to 0. For a system Λ of linear equations over $\mathbb{F}_q$, we formally define resolution as below

Definition 4.1 *Let Λ be a system of linear equations over a prime field $\mathbb{F}_q$. Then $\mathsf{Res}(\Lambda, t)$ is defined as the set of all equations involving at most t variables, each of which can be derived by a linear combination of at most two equations from Λ.*

Also, for a set $S \subseteq [n]$, let A_S denote the set of all partial assignments to variables in S, which satisfy all equations in Λ involving only variables from S. We then have the following theorem

Theorem 4.2 *Let q be a prime. Suppose Λ is a system of linear equations in $\mathbb{F}_q$ such that $(\overrightarrow{0} \cdot x = r) \in \Lambda \Leftrightarrow r = 0$ and $\mathsf{Res}(\Lambda, 2t) = \Lambda$. Then there are vectors $\mathbf{V}_{(S,\alpha)}$, for all S with $|S| < t$ and for all $\alpha \in [q]^S$, such that*

8

1. $\left\langle \mathbf{V}_{(S_1,\alpha_1)}, \mathbf{V}_{(S_2,\alpha_2)} \right\rangle \geq 0$ *for all $S_1, S_2, \alpha_1, \alpha_2$.*

2. $\left\langle \mathbf{V}_{(S_1,\alpha_1)}, \mathbf{V}_{(S_2,\alpha_2)} \right\rangle = 0$ *if $\alpha_1(S_1 \cap S_2) \neq \alpha_2(S_1 \cap S_2)$.*

3. *If $\alpha_i \in [q]^{S_i}$, $1 \leq i \leq 4$ are such that $\alpha_1 \circ \alpha_2$ and $\alpha_3 \circ \alpha_4$ are both defined and equal, then*

$$\left\langle \mathbf{V}_{(S_1,\alpha_1)}, \mathbf{V}_{(S_2,\alpha_2)} \right\rangle = \left\langle \mathbf{V}_{(S_3,\alpha_3)}, \mathbf{V}_{(S_4,\alpha_4)} \right\rangle$$

4. $\mathbf{V}_{(S,\alpha)} = 0$ *for $\alpha \notin A_S$ and $\sum_{\alpha \in A_S} \left\| \mathbf{V}_{(S,\alpha)} \right\|^2 = 1$*

Note that the theorem is stated in terms of a system Λ, which is closed under the applications of the operator $\mathsf{Res}(\cdot, 2t)$. Our proof of the theorem essentially follows Schoenebeck's proof except for one modification, which allows the generalization to the q-ary case. We defer the details to the appendix.

4.2 Deriving the gaps for MAX k-CSP$_q$

We will prove an integrality gap for general instances of MAX k-CSP(P_A). Specializing these to Hamming codes will then give the claimed results for MAX k-CSP$_q$. Let $A \in (\mathbb{F}_q)^{d \times k}$ be a matrix with linearly independent rows. Recall that an instance of MAX k-CSP(P_A) is specified by constraints $C_1, \ldots, C_m$ where each constraint C_i is a system of linear equations over $\mathbb{F}_q$ of the form $A \cdot (x_{i_1} \ldots x_{i_k})^{\mathsf{T}} = b^{(i)}$. We first show that if an instance Φ of MAX k-CSP(P_A) has good expansion, then $\mathsf{FRAC}(\Phi) = m$ even after large number of rounds. The integrality gap will then follow as an easy consequence.

Theorem 4.3 *Let $A \in (\mathbb{F}_q)^{d \times k}$ be the generator matrix of a distance 3 code. Let Φ be an instance of MAX k-CSP(P_A) with $m = \beta n$ constraints such that for $s \leq \eta n$, every set of s constraints contains at least $(k - 1 - \delta)s$ variables for $\delta \leq 1/4$. Then, $\mathsf{FRAC}(\Phi) = m$ for the SDP relaxation of MAX k-CSP(P_A) obtained by $(\frac{\eta^3}{8\beta^2})n$ rounds of the Lasserre hierarchy.*

Proof: We take Λ_0 to the set of all linear equations appearing in Φ, and take Λ to be closure of of it under $\mathsf{Res}(\cdot, t)$ for $t = (\frac{\eta^3}{4\beta^2})n$. It will follow easily from Theorem 4.2 that $\mathsf{FRAC}(\Phi) = m$ once we establish that Λ satisfies the necessary conditions to apply the theorem i.e. $(\overrightarrow{0} \cdot x = r) \in \Lambda \Leftrightarrow r = 0$. Expansion based arguments have often been used in previous works in proof complexity. For example, [25] used the argument due to Ben-Sasson and Wigderson [4]. We will show that even when the equations are not independent as in [4] and [25] (equations for every constraint come from A), the property that A has distance 3 shall turn out to be sufficient for (a modification of) the argument.

Let us assume that $(\overrightarrow{0} \cdot x = r_0) \in \Lambda$ for some $r_0 \neq 0$. We shall show that any derivation of this must have an intermediate equation involving too many variables, thus deriving a contradiction. Consider a minimal derivation tree of $(\overrightarrow{0} \cdot x = r_0)$ for $r_0 \neq 0$. By definition of $\mathsf{Res}(\cdot, t)$, this will be a binary tree with the equation $(\overrightarrow{0} \cdot x = r_0)$ at the root and equations in Λ_0 at the leaves.

Let $(\omega \cdot x = r)$ be any intermediate equation in the tree. We denote the node by (ω, r). We denote by $\nu(\omega, r)$ the number of constraints of Φ used in deriving (ω, r). It is immediate that if (ω_3, r_3) can be derived from (ω_1, r_1) and (ω_2, r_3), then $\nu(\omega_3, r_3) \leq \nu(\omega_1, r_1) + \nu(\omega_2, r_2)$. We first observe that $\nu(\overrightarrow{0}, r_0)$ must be large.

Claim 4.4 $\nu(\overrightarrow{0}, r_0) \geq \eta n$

Proof: let $\nu(\overrightarrow{0}, r_0) = s$. Then, the derivation for $(\overrightarrow{0}, r_0)$ involves a linear combination of equations from s constraints. Since every constraint comes from a distance 3 code, every linear combination of equations *within* a constraint must involve at least 3 variables. Hence, the linear combination requied to derive $(\overrightarrow{0}, r_0)$ must include at least 3 variables from each from each of the s constraints. However, to derive $\overrightarrow{0}$ each of these variables must occur an even number of times and hence the s contraints can involve at most $ks - 3s/2$ variables in total. Since every set of up to ηn constraints is highly expanding, this is only possible when $s \geq \eta n$. ∎

In the spirit of [4], we show that $\nu(\omega, r)$ decreases slowly as we proceed from the root to the intermediate nodes. This will allow us to find a node which requires many, but fewer than ηn constraints to derive. Since for less than ηn constraints we can use expansion, we will be able to argue that this node has an equation involving many variables.

Claim 4.5 *For all $j \leq \log_{3/2}(\eta n)$, there is a node (ω_j, r_j) in the tree such that*

$$\left(\frac{1}{3}\right)^j \nu(\overrightarrow{0}, r_0) \ \leq \ \nu(\omega_j, r_j) \ \leq \ \left(\frac{2}{3}\right)^j \nu(\overrightarrow{0}, r_0)$$

Proof: For $j = 0$, we can take $\omega_0 = \overrightarrow{0}$. It suffices to show that given (ω_j, r_j) one can find a node (ω_{j+1}, r_{j+1}) such that $\nu(\omega_j, r_j)/3 \leq \nu((\omega_{j+1}, r_{j+1})) \leq 2\nu(\omega_j, r_j)/3$. Let $(\omega_j^{(1)}, r_j^{(1)}), (\omega_j^{(2)}, r_j^{(2)})$ be the two children of (ω_j, r_j). Since $\nu(\omega_j^{(1)}, r_j^{(1)}) + \nu(\omega_j^{(2)}, r_j^{(2)}) \leq \nu(\omega_j, r_j)$, at least one of them, say $(\omega_j^{(1)}, r_j^{(1)})$, must require more than $\nu(\omega_j, r_j)/3$ constraints to derive it.

If $\nu(\omega_j^{(1)}, r_j^{(1)}) \leq 2\nu(\omega_j, r_j)/3$ then we are done. Else at least one of the children of $(\omega_j^{(1)}, r_j^{(1)})$ must require more than $\nu(\omega_j, r_j)/3$ constraints for its derivation and we can continue the argument on this node. Since we always go down one level in the tree and find a node requiring at least $2\nu(\omega_j, r_j)/3$ constraints, we must stop at some node as the leaves require only one constraint. We take the node we stop at to be (ω_{j+1}, r_{j+1}). ∎

Choosing $j = \lceil \log_{3/2}(s/\eta n) \rceil$, and using $\eta n \leq s \leq \beta n$ gives

$$(\eta^3/\beta^2)n \ \leq \ (\eta n)^3/s^2 \ \leq \ \nu(\omega_j, r_j) \ \leq \ \eta n$$

Consider the number of variables in ω_j. Each one of the constraints used in deriving it, contributes at least 3 variable occurrences. Also, since $\nu(\omega_j, r_j) \leq \eta n$, all the constraints must contain at least $(k-1-\delta)\nu(\omega_j, r_j)$ variables in total, which gives that at most $(1+\delta)\nu(\omega_j, r_j)$ variables occuring in more than one constraint. Out of all the variable occurrences in ω_j, the ones that can cancel out are the ones occurring in more than one constraint. Hence, ω_j must have at least $(3 - 2(1+\delta))\nu(\omega_j, r_j)$ variables. For $\delta \leq 1/4$, this is greater than $(\frac{\eta^3}{4\beta^2})n$ which is a contradiction.

Hence, Λ cannot contain an equation of the form $\overrightarrow{0} \cdot x = r_0$ for $r_0 \neq 0$. Since Λ is closed under $\mathsf{Res}(\cdot, (\frac{\eta^3}{4\beta^2})n)$ by definition, we can apply Theorem 4.2 to get vectors for all sets of size upto $(\frac{\eta^3}{8\beta^2})n$. The vectors also satisfy all the required consistency conditions. Finally, we note that

$$\sum_{i=1}^{m} \sum_{\alpha \in [q]^{T_i}} C_i(\alpha) \left\| \mathbf{V}_{(T_i, \alpha)} \right\|^2 \ = \ \sum_{i=1}^{m} \sum_{\alpha \in A_{T_i}} \left\| \mathbf{V}_{(T_i, \alpha)} \right\|^2 \ = \ \sum_{i=1}^{m} 1 \ = \ m$$

which shows that $\mathsf{FRAC}(\Phi) = m$. ∎

Since random instances are both unsatisfiable and expanding, it is now easy to derive the integrality gap for MAX k-CSP(P_A).

Theorem 4.6 *Let $A \in (\mathbb{F}_q)^{d \times k}$ be the generator matrix of a distance 3 code and let $\zeta > 0$ be given. Then there is a constant $c = c(q, k, \zeta)$ such that for large enough n, the integrality gap for the Lasserre SDP relaxation of MAX k-CSP(P_A) on n variables obtained by cn rounds is at least $q^d - \zeta$.*

Proof: We take $\epsilon = \zeta \cdot \frac{1}{q^d}$, $\delta = 1/4$ and consider a random instance Φ with $m = \beta n$ constraints as in Lemma 2.2, such that the Φ satisfies both the properties in the conclusion of the lemma. Then, by Theorem 4.3 FRAC$(\Phi) = m$ even after $(\frac{\eta^3}{8\beta^2})n$ rounds of the Lasserre hierarchy. On the other hand, by Lemma 2.2, OPT$(\Phi) \leq \frac{|P_A^{-1}(1)|}{q^k}(1 + \epsilon) \cdot m = \frac{1}{q^d}(1 + \epsilon) \cdot m$. Hence, the integrality gap is at least $q^d/(1 + \epsilon) \geq 1/q^d - \zeta$. ∎

We now derive near optimal integrality gaps for $\Omega(n)$ rounds of the Lasserre relaxation of the binary and q-ary MAX k-CSP problems. Note that the integrality gap becomes larger with the dimension of the code. Thus, to optimize the gap, we consider Hamming codes which have the largest dimension for a given length.

Corollary 4.7 *Let $k \geq 3$, q be a prime number and let $\zeta > 0$ be given. Then there exists $c = c(k, q, \zeta) > 0$ such that the for sufficiently large n, the integrality gap for the Lasserre relaxation of the MAX k-CSP problem on n variables with domain size q is at least $\dfrac{q^k}{kq(q - 1) - q(q - 2)} - \zeta$ after cn rounds.*

Proof: Let $(q^{r-1} - 1)/(q - 1) < k \leq (q^r - 1)/(q - 1)$. We take A to be the generator matrix of the Hamming code of length k. Note that the above implies that $q^r - q \leq (k - 1)q(q - 1)$ which gives $q^r \leq kq(q - 1) - q(q - 2)$.

Consider a random instance of MAX k-CSP(P_A). By Theorem 4.6, there exists a c and instances on n variables such that the integrality gap after cn rounds for MAX k-CSP(P_A) is at least $q^d - \zeta = q^k/q^r - \zeta$. Finally, using that $q^r \leq kq(q-1) - q(q-2)$ gives the gap is at least $\dfrac{q^k}{kq(q - 1) - q(q - 2)} - \zeta$ as claimed. ∎

Note that in case of $q = 2$, the generator matrix of the binary Hamming code, simply produces the predicate considered by Samorodnitsky and Trevisan [24]. We also state the following more precise version of the gap for binary k-CSPs. The constants are arbitrary, but we shall need the nature of the tradeoffs between k, β and c to get the gaps for Independent Set and Chromatic Number.

Corollary 4.8 *Let a number k and $\epsilon > 0$ be given and let A be the generator matrix for the Hamming code of length k. Then there exist $\beta = O(2^k/\epsilon^2)$ and $c = \Omega((1/\beta)^{80})$ such that if Φ is a random instance of MAX k-CSP(P_A) on $n \gg 1/c$ variables and $m = \beta n$ constraints, then with probability $1 - o(1)$*

  1. OPT$(\Phi) \leq \frac{2k}{2^k}(1 + \epsilon) \cdot m$

  2. For the SDP given by cn rounds of the Lasserre hierarchy, FRAC$(\Phi) = m$.

Proof: Invoking Lemma 2.2 with $\delta = 1/5$ gives $\beta = O(2^k/\epsilon^2)$ and $\eta = O((1/\beta)^{25})$. Theorem 4.3 gives FRAC$(\Phi) = m$ after cn rounds, for $c = \Omega(\eta^3/\beta^2) = \Omega((1/\beta)^{80})$. The dimension d for a Hamming code is $k - 2^{\lceil \log(k+1) \rceil} \geq k - \log(2k)$. Hence OPT$(\Phi) \leq \frac{2k}{2^k}(1 + \epsilon)m$. ∎

4.3 Implications for Random Predicates

We now derive integrality gaps for MAX k-CSP(P) where $P : [q]^k \to \{0, 1\}$ is chosen at random by selecting each input to be in $P^{-1}(1)$ independently with probability p. We denote this distribution over q-ary predicates with k inputs as $Q(p, q, k)$. For any predicate P, a random assignment satisfies $|P^{-1}(1)|/q^k$ fraction of constraints in MAX k-CSP(P) and hence the largest integrality gap one can have is $q^k/|P^{-1}(1)|$. We will show that for almost every random predicate P, the integrality gap for MAX k-CSP(P) is *at least* $q^k/|P^{-1}(1)| - \zeta$ even after $\Omega(n)$ rounds of the Lasserre hierarchy.

The result will follow quite easily using a theorem of Håstad [17] which basically says that a random predicate "contains a copy of" P_A where A is the generator matrix of the Hamming code over $\mathbb{F}_q$. We first define what it means for a predicate to contain a copy of another.

Definition 4.9 *We say that a predicate P_1 contains a predicate equivalent to P_2 if there exists a permutation $\pi : [k] \to [k]$ of the inputs and $b_1, \ldots, b_k \in [q]$, such that*

$$P_2(x_{\pi(1)} + b_1, \ldots, x_{\pi(k)} + b_k) = 1 \implies P_1(x_1, \ldots, x_k) = 1$$

We can now state the theorem of Håstad referred to above. Håstad actually states the result only for random boolean predicates but it is easy to verify that the same proof can be extended to q-ary predicates.

Theorem 4.10 (Håstad [17]) *Let q be a prime and let $(q^{r-1} - 1)/(q - 1) < k \le (q^r - 1)/(q - 1)$. Let A be the generator matrix of the Hamming code over $\mathbb{F}_q$ with length k. Then there is a value c of the form $c = kq^{-r}(1 - o(1))$, such that, with probability $1 - o(1)$, a random predicate chosen according to $Q(p, q, k)$ with $p \ge k^{-c}$ contains a predicate equivalent to P_A.*

Using this theorem, we can now prove optimal integrality gap for almost every predicate in the distribution $Q(p, q, k)$ with the appropriate value of p.

Theorem 4.11 *Let a prime q and $\zeta > 0$ and $k \ge 3$ be given and let $(q^{r-1} - 1)/(q - 1) < k \le (q^r - 1)/(q - 1)$ for some r. Then there exist constants $c = kq^{-r}(1 - o(1))$ and $c' = c'(q, k, \zeta)$ such that if P is a random predicate chosen according to $Q(p, q, k)$ with $p \ge k^{-c}$, then with probability $1 - o(1)$ over the choice of P, the integrality gap for MAX k-CSP(P) after $c'n$ rounds of the Lasserre hierarchy is at least $q^k/|P^{-1}(1)| - \zeta$.*

Proof: Using Theorem 4.10 we know that with probability $1 - o(1)$, a random P contains a predicate equivalent to P_A, where A is the generator matrix of the Hamming code over $\mathbb{F}_q$ with length k. For the rest of the proof, we fix such a P. For this P there exists a permutation π and literals $b_1, \ldots, b_k$ such that $P_A(x_{\pi(1)} + b_1, \ldots, x_{\pi(k)} + b_k) = 1 \implies P(x_1, \ldots, x_k) = 1$.

With every instance Φ of MAX k-CSP(P), we now associate an instance Φ_A of MAX k-CSP(P_A). For every constraint $C_i \equiv P(x_{i_1} + a_{i_1}, \ldots, x_{i_k} + a_{i_k})$ in Φ, we add a constraint C_i' to Φ_A of the form

$$C_i' \equiv P_A(x_{i_{\pi(1)}} + a_{i_{\pi(1)}} + b_1, \ldots, x_{i_{\pi(k)}} + a_{i_{\pi(k)}} + b_k)$$

Thus, if the constraint C_i' is satisfied by an assignment, then so is C_i. Also, if Φ is distributed as a random instance of MAX k-CSP(P), then Φ_A is distributed as a random instance of MAX k-CSP(P_A) with the same number of constraints.

Let $\epsilon = \zeta \cdot |P^{-1}(1)|/q^k$ and $\delta = 1/4$. We consider a random instance Φ of MAX k-CSP(P) with $m = \beta n$ constraints as in Lemma 2.2. By Lemma 2.2 we will have with probability $1 - o(1)$ over Φ that

- $\mathsf{OPT}(\Phi) \leq \dfrac{|P^{-1}(1)|}{q^k}(1+\epsilon) \cdot m$, and

- Every set of $s \leq \eta n$ constraints in Φ_A contains at least $(k - 1 - \delta)s$ variables.

By Theorem 4.3, we have $\mathsf{FRAC}(\Phi_A) = m$ for the SDP obtained by $t = (\frac{\eta^3}{8\beta^2})n$ rounds of the Lasserre hierarchy. Hence, there exist vectors $\mathbf{V}_{(S,\alpha)}$ for all $S \subseteq [n], |S| \leq t$ and $\alpha \in [q]^S$ satisfying all the consistency constraints and such that

$$\sum_{\alpha \in [q]^{T_i}} C_i'(\alpha) \left\| \mathbf{V}_{(T_i,\alpha)} \right\|^2 = 1 \qquad \forall 1 \leq i \leq m$$

However, the same vectors also show that $\mathsf{FRAC}(\Phi) = m$ after t rounds, since

$$\sum_{i=1}^{m} \sum_{\alpha \in [q]^{T_i}} C_i(\alpha) \left\| \mathbf{V}_{(T_i,\alpha)} \right\|^2 \geq \sum_{i=1}^{m} \sum_{\alpha \in [q]^{T_i}} C_i'(\alpha) \left\| \mathbf{V}_{(T_i,\alpha)} \right\|^2 = \sum_{i=1}^{m} 1 = m$$

Hence, the integrality gap for $\mathsf{MAX\ k\text{-}CSP}(P)$ after $(\frac{\eta^3}{8\beta^2})n$ rounds of the Lasserre hierarchy is at least $\mathsf{FRAC}(\Phi)/\mathsf{OPT}(\Phi) \geq \dfrac{q^k}{|P^{-1}(1)|(1+\epsilon)} \geq \dfrac{q^k}{|P^{-1}(1)|} - \zeta.$ ∎

5 Integrality Gap for Maximum Independent Set

To obtain the integrality gaps for Maximum Independent Set we use the reductions by Feige et. al. [12] and by Bellare, Goldreich and Sudan [3]. However, before getting to the proof of the integrality gap, we describe how to transform vectors for a general FGLSS reduction. This transformation shall be useful for our other results as well.

5.1 Vectors for products of the FGLSS graph

Let Φ be an instance of MAX k-CSP with constraints $C_1, \ldots, C_m$ on tuples $T_1, \ldots, T_m$ and the domain of variables as $\{0, 1\}$. Assume that each constraint has exactly l satisfying assignments. We describe below the reduction by [3] from Φ to an independent set problem.

1. Given Φ, create the FGLSS graph $G_\Phi = (V_\Phi, E_\Phi)$ with a vertex for every constraint C_i and every partial assignment to variables in the corresponding tuple T_i which satisfies the constraint C_i. Two vertices $(C_{i_1}, \alpha_1), (C_{i_2}, \alpha_2)$ are connected if α_1 and α_2 assign different values to some variable. Formally

$$\begin{aligned} V_\Phi &= \{(C_i, \alpha) \mid \alpha \in \{0,1\}^{T_i},\ C_i(\alpha) = 1\} \\ E_\Phi &= \{\{(C_{i_1}, \alpha_1), (C_{i_2}, \alpha_2)\} \mid \alpha_1(T_{i_1} \cap T_{i_2}) \neq \alpha_2(T_{i_1} \cap T_{i_2})\} \end{aligned}$$

2. Construct the product graph $G_\Phi^r = (V_\Phi^r, E')$ with vertices of G_Φ^r being r-tuples of vertices in G_Φ. Two vertices $\{(C_{i_1}, \alpha_1), \ldots, (C_{i_r}, \alpha_r)\}$ and $\{(C_{i_1'}, \alpha_1'), \ldots, (C_{i_r'}, \alpha_r')\}$ are connected if for some j, $\{(C_{i_j}, \alpha_j), (C_{i_j'}, \alpha_j')\} \in E_\Phi$.

13

Note that if Φ had m constraints, then G_Φ^r has $l^r \cdot m^r$ vertices, with there being m^r disjoint cliques of l^r vertices, corresponding to every r-tuple of constraints. We denote the clique corresponding to constraints $C_{i_1}, \ldots, C_{i_r}$ as $\mathcal{C}(i_1, \ldots, i_r)$. Formally,

$$\mathcal{C}(i_1, \ldots, i_r) = \left\{ \{(C_{i_1}, \alpha_1), \ldots, (C_{i_r}, \alpha_r)\} \mid \wedge_{j=1}^r C_{i_j}(\alpha_j) = 1 \right\}$$

The largest independent set in G_Φ^r can have at most m^r vertices. We claim that a good SDP solution for Φ can be transformed into a good solution for the independent set SDP on G_Φ^r.

Lemma 5.1 *Let Φ be an instance of* MAX k-CSP *as above with m constraints. If* $\mathsf{FRAC}(\Phi) = m$ *after t rounds of the Lasserre hierarchy, then* $\mathsf{FRAC}(G_\Phi^r) \geq m^r$ *for the independent set SDP obtained after t/k rounds. Moreover, the contribution to the SDP value from vertices in each clique* $\mathcal{C}(i_1, \ldots, i_r)$ *is 1.*

Proof: We first define an independent set solution for t/k rounds on G_Φ and then show how to extend it to G_Φ^r. Consider a set $\mathcal{S}$ of $h \leq t/k$ vertices in G_Φ. It is specified by h constraints and partial assignments $\{(C_{i_1}, \alpha_1), \ldots, (C_{i_h}, \alpha_h)\}$. Define $\mathbf{U}_\mathcal{S}$ as

$$\mathbf{U}_\mathcal{S} = \begin{cases} 0 & \exists j_1, j_2 \leq h \text{ s.t. } \alpha_{j_1}(T_{i_{j_1}} \cap T_{i_{j_2}}) \neq \alpha_{j_2}(T_{i_{j_1}} \cap T_{i_{j_2}}) \\ \mathbf{V}_{(\cup_j T_{i_j}, \alpha_1 \circ \ldots \circ \alpha_h)} & \text{otherwise} \end{cases}$$

We now consider a set $\overline{\mathcal{S}}$ of vertices in G_Φ^r. It is a *set of r-tuples* of vertices in G_Φ. Let $\mathcal{S}_j$ denote the set of vertices of G_Φ which occur in the jth coordinate of the r-tuples in $\overline{\mathcal{S}}$. Define the vector $\mathbf{U}_{\overline{\mathcal{S}}}$ as

$$\mathbf{U}_{\overline{\mathcal{S}}} = \mathbf{U}_{\mathcal{S}_1} \otimes \ldots \otimes \mathbf{U}_{\mathcal{S}_r}$$

Let $\mathbf{U}_{\overline{\emptyset}}$ denote the vector for the empty set of vertices in G_Φ^r. We take $\mathbf{U}_{\overline{\emptyset}} = \mathbf{U}_\emptyset \otimes \ldots \otimes \mathbf{U}_\emptyset$. The vectors $\mathbf{U}_{\overline{\mathcal{S}}}$ are defined for all sets $\overline{\mathcal{S}}$ with at most t/k vertices. We now show that they satisfy all Lasserre constraints.

Claim 5.2 *The vectors $\mathbf{U}_{\overline{\mathcal{S}}}$ satisfy all conditions of the (t/k)-round independent set SDP on G_Φ^r.*

Proof: Since all vectors $\mathbf{U}_{\overline{\mathcal{S}}}$ are tensors of valid Lasserre vectors for the SDP for Φ, all inner products are between 0 and 1. We only need to verify that the vectors corresponding to two vertices connected by an edge are orthogonal, and that $\left\langle \mathbf{U}_{\overline{\mathcal{S}}_1}, \mathbf{U}_{\overline{\mathcal{S}}_2} \right\rangle$ depends only on $\overline{\mathcal{S}}_1 \cup \overline{\mathcal{S}}_2$.

- Consider two vertices $\{(C_{i_1}, \alpha_1), \ldots, (C_{i_r}, \alpha_r)\}$ and $\{(C_{i'_1}, \alpha'_1), \ldots, (C_{i'_r}, \alpha'_r)\}$ connected by an edge. The corresponding vectors are $\mathbf{V}_{(T_{i_1}, \alpha_1)} \otimes \ldots \otimes \mathbf{V}_{(T_{i_r}, \alpha_r)}$ and $\mathbf{V}_{(T_{i'_1}, \alpha'_1)} \otimes \ldots \otimes \mathbf{V}_{(T_{i'_r}, \alpha'_r)}$. The fact that there is an edge between the vertices means that for some $j \leq r$, $\alpha_j(T_{i_j} \cap T_{i'_j}) \neq \alpha'_j(T_{i_j} \cap T_{i'_j})$. Hence $\left\langle \mathbf{V}_{(T_{i_j}, \alpha_j)}, \mathbf{V}_{(T_{i'_j}, \alpha'_j)} \right\rangle = 0$ since the vectors $\mathbf{V}_{(\cdot, \cdot)}$ form a valid Lasserre solution. This gives

$$\left\langle \mathbf{V}_{(T_{i_1}, \alpha_1)} \otimes \ldots \otimes \mathbf{V}_{(T_{i_r}, \alpha_r)}, \mathbf{V}_{(T_{i'_1}, \alpha'_1)} \otimes \ldots \otimes \mathbf{V}_{(T_{i'_r}, \alpha'_r)} \right\rangle = \prod_{j=1}^r \left\langle \mathbf{V}_{(T_{i_j}, \alpha_j)}, \mathbf{V}_{(T_{i'_j}, \alpha'_j)} \right\rangle = 0$$

- Next, consider sets $\overline{\mathcal{S}}_1, \overline{\mathcal{S}}_2, \overline{\mathcal{S}}_3, \overline{\mathcal{S}}_4$ such that $\overline{\mathcal{S}}_1 \cup \overline{\mathcal{S}}_2 = \overline{\mathcal{S}}_3 \cup \overline{\mathcal{S}}_4$. For $1 \leq u \leq 4$, let $\mathcal{S}_j^{(u)}$ denote the union of elements in the jth coordinate of the r-tuples in $\overline{\mathcal{S}}_u$. $\overline{\mathcal{S}}_1 \cup \overline{\mathcal{S}}_2 = \overline{\mathcal{S}}_3 \cup \overline{\mathcal{S}}_4$ means that in particular $\mathcal{S}_j^{(1)} \cup \mathcal{S}_j^{(2)} = \mathcal{S}_j^{(3)} \cup \mathcal{S}_j^{(4)}$ for all $1 \leq j \leq r$. For a fixed j, let $\mathcal{S}_j^{(1)} \cup \mathcal{S}_j^{(2)} = \mathcal{S}_j^{(3)} \cup \mathcal{S}_j^{(4)} = \{(C_{i_1}, \alpha_1), \ldots, (C_{i_h}, \alpha_h)\}$. If the set contains two contradicting partial assignments, then either one of $\mathbf{U}_{\mathcal{S}_1}^{(j)}$ and $\mathbf{U}_{\mathcal{S}_2}^{(j)}$ is 0, or they are equal to Lasserre vectors corresponding to contradicting partial assignments. In either case $\left\langle \mathbf{U}_{\mathcal{S}_1}^{(j)}, \mathbf{U}_{\mathcal{S}_2}^{(j)} \right\rangle = 0$ and similarly $\left\langle \mathbf{U}_{\mathcal{S}_3}^{(j)}, \mathbf{U}_{\mathcal{S}_4}^{(j)} \right\rangle = 0$. If there are no contradicting partial assignments, then the tuples in $\mathcal{S}_1^{(j)} \cup \mathcal{S}_2^{(j)}$ can be extended to a unique partial assignment $\alpha_1 \circ \ldots \circ \alpha_h$ over $\cup_{j=1}^h T_{j_j}$. Since the set of all tuples, and hence the assignment, is same for $\mathcal{S}_3^{(j)} \cup \mathcal{S}_4^{(j)}$, and the corresponding CSP vectors are consistent, we get $\left\langle \mathbf{U}_{\mathcal{S}_1}^{(j)}, \mathbf{U}_{\mathcal{S}_2}^{(j)} \right\rangle = \left\langle \mathbf{U}_{\mathcal{S}_3}^{(j)}, \mathbf{U}_{\mathcal{S}_4}^{(j)} \right\rangle$ for all j, which implies $\left\langle \mathbf{U}_{\overline{\mathcal{S}}_1}, \mathbf{U}_{\overline{\mathcal{S}}_2} \right\rangle = \left\langle \mathbf{U}_{\overline{\mathcal{S}}_3}, \mathbf{U}_{\overline{\mathcal{S}}_4} \right\rangle$. ∎

We show that the value for all the vertices in any clique $\mathcal{C}(i_1, \ldots, i_r)$ is 1. Letting $\alpha_1, \ldots, \alpha_r$ range over all satisfying assignments to $C_{i_1}, \ldots, C_{i_j}$, the contribution of vertices in this clique to the SDP objective value is

$$\sum_{\alpha_1, \ldots, \alpha_r} \prod_{j=1}^r \left\langle \mathbf{V}_{(T_{i_j}, \alpha_j)}, \mathbf{V}_{(\emptyset, \emptyset)} \right\rangle = \prod_{j=1}^r \left\langle \sum_{\alpha_j} \mathbf{V}_{(T_{i_j}, \alpha_j)}, \mathbf{V}_{(\emptyset, \emptyset)} \right\rangle = \prod_{j=1}^r (1) = 1$$

where $\left\langle \sum_{\alpha_j} \mathbf{V}_{(T_{i_j}, \alpha_j)}, \mathbf{V}_{(\emptyset, \emptyset)} \right\rangle = 1$ since the contribution of the constraint C_{i_j} to the SDP for MAX k-CSP is 1. ∎

5.2 Obtaining the Integrality Gap

We can now prove the following integrality gap for Maximum Independent Set.

Theorem 5.3 *There exist constants $c_1, c_2 > 0$ and graphs on N vertices for arbitraily large N, such that the integrality gap for the SDP for independent set obtained by $2^{c_2 \sqrt{\log N \log \log N}}$ rounds of the Lasserre hierarchy, is at least $\dfrac{N}{2^{c_1 \sqrt{\log N \log \log N}}}$.*

Proof: Our integrality gap instance will be a subgraph of G_Φ^r for appropriate choices of Φ and r. We construct the graph $G = (V, E)$ by randomly picking M cliques of the form $\mathcal{C}(i_1, \ldots, i_r)$, and taking G to be the subgraph induced by the vertices in these clique. An easy Chernoff bound shows that if only a small fraction of constraints in Φ were satisfiable, then the size of the largest independent set in G is small.

Claim 5.4 *Let $s = \mathsf{OPT}(\Phi)/m$. Then for $M \geq \frac{100nr}{s^r}$, with probability $1 - o(1)$, all independent sets in G have size at most $2s^r M$.*

Proof: It is easy to see that any independent set in G_Φ can be extended to an assignment to the variables $x_1, \ldots, x_n$ and has size equal to the number of constraints in Φ satisfied by the assignment. Hence, the size of the largest independent set in G_Φ is at most $s \cdot m$. Also, an independent set in G_Φ^r

15

is a set of r-tuples of vertices in G_Φ such that if we consider the set of vertices in the jth coordinate for any j, they form an independent set in G_Φ. Hence, any independent set in G_Φ^r has size at most $(s \cdot m)^r$. Also, note that since an independent set of G_Φ^r can be extended to an assignment to $x_1, \ldots, x_n$ in each of the r coordinatess, there are at most 2^{nr} different independent sets.

Any independent set in the sampled graph G also extends to an assignment to $x_1, \ldots, x_n$ in each coordinate and can be thought of as the intersection of an independent set I of G_Φ^r with the sampled blocks. Fix any independent set I of G_Φ^r. We sample M out of the m^r blocks in G_Φ and each block has at most one vertex belonging to I (because each block is a clique). Hence, by Chernoff bounds, the probability that more than $2s^r \cdot M$ vertices in G belong to I is at most $\exp(-s^r \cdot M/50)$. Taking a union bound over all choices of I, we get that with probability at least $1 - \exp(-s^r \cdot M/50 + nr)$, all independent set of G have size at most $2s^r \cdot M$. Choosing $M \geq 100nr/s^r$ ensures that the probability is $1 - o(1)$. ∎

We now make the choices for all the parameters involved. For a large n, let $k = \delta \log n$ for some small constant δ, and let $r = \log n/(\log \log n)$. Consider an instance Φ of MAX k-CSP as given by Corollary 4.8. By chosing $\epsilon = 1/2$, we can get that $k/2^k \leq s \leq 3k/2^k$. Also, since the constraints are based on the Hamming code, the number of satisfying assignments to each constraint is at most $l \leq 2k$. We pick $M = 100nr \cdot (2^{kr}/k^r)$.

By the previous claim, the size of the maximum independent set in G is at most $2Ms^r$ (w.h.p. over the choice of G). We take the SDP solution to be the same as constructed for G_Φ^r. By Lemma 5.1, the contribution of the vectors in each clique to the SDP value is 1. Hence, the value of the SDP solution for G is M, which gives an integrality gap of $(1/2s^r) \geq (1/2) \cdot (2^k/3k)^r$. On the other hand, the number of vertices in G is

$$N = M \cdot l^r \leq (100nr \cdot 2^{kr}/k^r) \cdot (2k)^r = O(nr \cdot 2^{(k+1)r})$$

With our choice of parameters, the integrality gap is at least $\dfrac{N}{2^{c_1\sqrt{\log N \log \log N}}}$ for some constant c_1.

To verify the number of rounds, note that Corollary 4.8 gives $\beta = O(2^k) = O(n^\delta)$ and $c = \Omega((1/n^\delta)^{80})$. Hence, we have SDP solutions for $cn = \Omega(n^{1-80\delta})$ rounds for Φ and consequently for $\Omega(n^{1-80\delta}/k)$ rounds for the independent set SDP on G. For $\delta < 1/80$, this is at least $2^{c_2\sqrt{\log N \log \log N}}$ for some constant c_2. ∎

6 Gaps for Graph Coloring

In this section, we show that SDPs in the Lasserre hierarchy fail to approximate the chromatic number of a graph. Gaps for chromatic number are syntactically different from the usual integrality gaps for SDPs because the value of the chromatic number is not a linear function of the inner products of vectors in an SDP. Instead for any l, one can write down an SDP for which a *feasible solution* gives a *vector l-coloring* of the graph. We show graphs for which an l-coloring remains feasible even after many rounds of the Lasserre hierarchy, even though the actual chromatic number of the graph is much larger than l.

We show that for any constant l, there are graphs with chromatic number at least $\frac{2^{l/2}}{4l^2}$ which admit a vector l-coloring even after $\Omega(n)$ rounds of the Lasserre hierarchy. For Chromatic Number, we show that the ratio of the chormatic number of the graph and the number of colors in the vector coloring obtained by $2^{\Omega(\sqrt{\log n \log \log n})}$ rounds of the Lasserre hierarchy can be as high as $\frac{n}{2^{O(\sqrt{\log n \log \log n})}}$. We write the Lasserre SDP for l-coloring as a constraint satisfaction problem, with

16

the additional restriction that *all* the constraints which say colors of two adjacent vertices must be different, are satisfied. This formulation is equivalent to the one considered by Chlamtac [8]. [4] To avoid confusion with the SDP for MAX k-CSP we denote the sets of vertices here by $\mathcal{S}$, partial assignments by γ and vectors for coloring by $\overline{\mathbf{V}}_{(\mathcal{S},\gamma)}$.

---

t-round Lasserre SDP for l-coloring of graph $G = (V, E)$

Minimize l s.t. there exist vectors $\overline{\mathbf{V}}_{(\mathcal{S},\gamma)}$ for all $|\mathcal{S}| \leq t, \gamma \in [l]^{\mathcal{S}}$ satisfying

$$\left\langle \overline{\mathbf{V}}_{(\{u_1\},\gamma)}, \overline{\mathbf{V}}_{(\{u_2\},\gamma)} \right\rangle = 0 \qquad\qquad\qquad \forall (u_1, u_2) \in E, \gamma \in [l]$$

$$\left\langle \overline{\mathbf{V}}_{(\mathcal{S}_1,\gamma_1)}, \overline{\mathbf{V}}_{(\mathcal{S}_2,\gamma_2)} \right\rangle = 0 \qquad\qquad\qquad \forall\ \gamma_1(\mathcal{S}_1 \cap \mathcal{S}_2) \neq \gamma_2(\mathcal{S}_1 \cap \mathcal{S}_2)$$

$$\left\langle \overline{\mathbf{V}}_{(\mathcal{S}_1,\gamma_1)}, \overline{\mathbf{V}}_{(\mathcal{S}_2,\gamma_2)} \right\rangle = \left\langle \overline{\mathbf{V}}_{(\mathcal{S}_3,\gamma_3)}, \overline{\mathbf{V}}_{(\mathcal{S}_4,\gamma_4)} \right\rangle \qquad \forall\ \mathcal{S}_1 \cup \mathcal{S}_2 = \mathcal{S}_3 \cup \mathcal{S}_4,\ \gamma_1 \circ \gamma_2 = \gamma_3 \circ \gamma_4$$

$$\sum_{j \in [l]} \left\| \overline{\mathbf{V}}_{(\{u\},j)} \right\|^2 = 1 \qquad\qquad\qquad \forall u \in V$$

$$\left\langle \overline{\mathbf{V}}_{(\mathcal{S}_1,\gamma_1)}, \overline{\mathbf{V}}_{(\mathcal{S}_2,\gamma_2)} \right\rangle \geq 0 \qquad\qquad\qquad \forall \mathcal{S}_1, \mathcal{S}_2, \gamma_1, \gamma_2$$

$$\left\| \overline{\mathbf{V}}_{(\emptyset,\emptyset)} \right\| = 1$$

The reduction we describe in this section is a slightly modified version of the reduction for independent set and is specifically designed to work for problems of the type MAX k-CSP(P_A), where the constraints being linear equations in $\mathbb{F}_2$. It is inspired by what could be a "zero-knowledge protocol" for such predicates, in the sense of Feige and Killian [14]. Here, we describe the reduction without going through the protocol, at the cost of defining the following additional (and somewhat unnatural) solution concept for a CSP instance.

Definition 6.1 *Let Φ be an instance of* MAX k-CSP *with constraints $C_1, \ldots, C_m$ on ordered k-tuples $T_1, \ldots, T_m$ of variables. For a constraint C_i on variables $(x_{i_1}, \ldots, x_{i_k})$, we say that the constraint* k-satisfied *by assignments $\Pi_1, \ldots, \Pi_k$ if $C_i(\Pi_1(x_{i_1}), \ldots, \Pi_k(x_{i_k})) = 1$. We denote by $\mathsf{OPT}_k(\Phi)$, the maximum number of constraints in Φ that are k-satisfied by any assignments $\Pi_1, \ldots, \Pi_k$.*

Note that the above definition crucially uses the fact that the constraint is defined on an *ordered tuple* of variables as we read the value of the first variable from Π_1, the second from Π_2 and so on. By slightly strengthening the notion of unsatisfiability for a random CSP instance in Lemma 2.2, we can strengthen Corollary 4.8 as below.

Corollary 6.2 *Let a number k and $\epsilon > 0$ be given and let A be the generator matrix for the Hamming code of length k. Then there exist $\beta = O(k2^k/\epsilon^2)$ and $c = \Omega((1/\beta)^{80})$ such that if Φ is a random instance of* MAX k-CSP*(P_A) on $n \gg 1/c$ variables and $m = \beta n$ constraints, then with probability $1 - o(1)$*

1. $\mathsf{OPT}_k(\Phi) \leq \frac{2k}{2^k}(1 + \epsilon) \cdot m$

2. *For the SDP given by cn rounds of the Lasserre hierarchy,* $\mathsf{FRAC}(\Phi) = m$.

---

[4]In fact, the SDPs in Lasserre hierarchy are fairly independent of the representation used. It is easy to switch between different SDPs for a problem, by losing at most a constant factor in the number of rounds.

6.1 Gaps for Approximate Graph Coloring

We reduce from a CSP instance Φ as in Corollary 6.2. For an instance Φ of MAX k-CSP(P_A), consider the vectors $w \in \{0,1\}^k$ such that $A \cdot w^{\mathsf{T}} = 0$ over $\mathbb{F}_2$. If A is the generator matrix of the Hamming code of length k, there are $2^{\lceil \log(k+1) \rceil}$ such vectors. We shall show that the graph produced by our reduction has a vector coloring with $l = 2^{\lceil \log(k+1) \rceil}$ colors, where we shall identify the domain $[l]$ of the coloring CSP with the vectors $w_1, \ldots, w_l$ satisfying $A \cdot w_j^{\mathsf{T}} = 0$.

We now give the reduction from Φ as above, to Approximate Graph Coloring. Similar to the case of independent set, we create the FGLSS graph with a vertex for every constraint and every satisfying partial assignment to the variables in that constraint. However, we have fewer edges: we connect two vertices (C_{i_1}, α_1) and (C_{i_2}, α_2) iff α_1 and α_2 disagree on a variable that occurs *at the same position* in the ordered tuples T_{i_1} and T_{i_2}. Formally, we create the graph $G_\Phi = (V_\Phi, E_\Phi)$ such that

$$V_\Phi = \{(C_i, \alpha) \mid \alpha \in \{0,1\}^{T_i},\ C_i(\alpha) = 1\}$$
$$E_\Phi = \{\{(C_{i_1}, \alpha_1), (C_{i_2}, \alpha_2)\} \mid \exists 1 \le j \le k.\ [T_{i_1,j} = T_{i_2,j}] \wedge [\alpha_1(T_{i_1,j}) \ne \alpha_2(T_{i_2,j})]\}$$

where $T_{i,j}$ is used to denote the variable in the jth position in the ordered tuple corresponding to the ith constraint. To show that G_Φ has large chromatic number, we claim that all independent sets in G_Φ are small.

Claim 6.3 *The size of the maximum independent set in G_Φ is $\mathsf{OPT}_k(\Phi)$.*

Proof: Let I be an independent set in Φ. Hence I is a set of pairs of the form (C_i, α), where C_i is a constraint and α is a partial assignment giving values for variables in T_i. Since all vertices corresponding to a single constraint are connected, I can include at most one vertex corresponding to one constraint.

Consider the values given to all the variables $x_1, \ldots, x_k$ by all the partial assignments in I, when the variable is present in the jth position in the tuple. Since all the partial assignments to constraints must be consistent in the values at the jth position, these values can be extended to a unique assignment, say $\Pi_j = (a_1, \ldots, a_n)$ to the variables $x_1, \ldots, x_n$. Similarly, we can define assignments $\Pi_1, \ldots, \Pi_k$ for each of the k positions.

Hence, the independent set corresponds to picking at most one of the satisfying assignments for every constraint, with the jth variable in the tuple set according to Π_j. This gives that the size of the largest independent set is at most $\mathsf{OPT}_k(\Phi)$. ∎

Lemma 6.4 *Let Φ be an instance of MAX k-CSP(P_A), with m constraints such that each constraint has exactly l satisfying assignments. If $\mathsf{FRAC}(\Phi) = m$ after t rounds of the Lasserre hierarchy, then there is a feasible solution to the SDP for l-coloring of G_Φ obtained by $t/2k$ rounds of the Lasserre hierarchy.*

Proof: We now define the vectors $\overline{\mathbf{V}}_{(\mathcal{S}, \gamma)}$ for a set $\mathcal{S} \subseteq V_\Phi, |\mathcal{S}| \le t/k$ and $\gamma \in [l]^{\mathcal{S}}$. Let $(\mathcal{S}, \gamma) = (\{(C_{i_1}, \alpha_1), \ldots, (C_{i_h}, \alpha_h)\}, \gamma)$ Recall that the domain $[l]$ is identified with the vectors $w_1, \ldots, w_l \in \{0,1\}^k$ which satisfy $A \cdot w_j^{\mathsf{T}} = 0$ for $1 \le j \le l$. Hence the partial assignment γ assigns a *vector* in $\mathbb{F}_2^k$ to each vertex (C_{i_j}, α_j). We use the vectors given by γ to modify the assignments to each C_{i_j}. This can be viewed as the zero-knowledge step of randomizing over all the satisfying assignments to each constraint. Formally, we change α_j to $\alpha_j + \gamma((C_{i_j}, \alpha_j))$ where $\gamma((C_{i_j}, \alpha_j))$ is the vector in $\mathbb{F}_2^k$ (the "color") assigned by γ to the vertex (C_{i_j}, α_j) and the '+' is over $\mathbb{F}_2$. Let $[\alpha_j, \gamma]$ denote this assignment to T_{i_j} which is shifted by γ.

18

With this interpretation, we define the vectors as 0 if these shifted partial assignments contradict, and otherwise as the Lasserre vectors corresponding to the assignment defined collectively by all the shifted assignments. For all $|\mathcal{S}| \leq t/k$ and $\gamma \in [l]^{\mathcal{S}}$, we define

$$\overline{\mathbf{V}}_{(\mathcal{S},\gamma)} = \begin{cases} 0 & \exists j_1, j_2.\ [\alpha_{j_1}, \gamma](T_{i_{j_1}} \cap T_{i_{j_2}}) \neq [\alpha_{j_2}, \gamma](T_{i_{j_1}} \cap T_{i_{j_2}}) \\ \mathbf{V}_{(\cup T_{i_j}, [\alpha_1, \gamma] \circ \ldots \circ [\alpha_h, \gamma])} & \text{otherwise} \end{cases}$$

We now need to verify that the vectors satisfy all the SDP conditions.

- For an edge $\{(C_{i_1}, \alpha_1), (C_{i_2}, \alpha_2)\}$, we must have that $\left\langle \overline{\mathbf{V}}_{(\{(C_{i_1}, \alpha_1)\}, \gamma)}, \overline{\mathbf{V}}_{(\{(C_{i_2}, \alpha_2)\}, \gamma)} \right\rangle = 0$. Note that if (C_{i_1}, α_1) and (C_{i_2}, α_2) have an edge, then for some j, T_{i_1} and T_{i_2} have the same variable in the jth position and α_1, α_2 disagree on that variable. Then $[\alpha_1, \gamma]$ and $[\alpha_2, \gamma]$, which are equal to $\alpha_1 + w$ and $\alpha_2 + w$ for some w in the null space of A, would also disagree on that variable. Hence, by validity of the Lasserre solution for the CSP

$$\left\langle \overline{\mathbf{V}}_{(\{(C_{i_1}, \alpha_1)\}, \gamma)}, \overline{\mathbf{V}}_{(\{(C_{i_2}, \alpha_2)\}, \gamma)} \right\rangle = \left\langle \mathbf{V}_{((T_{i_1}, [\alpha_1, \gamma])}, \mathbf{V}_{(T_{i_2}, [\alpha_2, \gamma])} \right\rangle = 0$$

- We next verify that $\left\langle \overline{\mathbf{V}}_{(\mathcal{S}_1, \gamma_1)}, \overline{\mathbf{V}}_{(\mathcal{S}_2, \gamma_2)} \right\rangle = 0$ whenever γ_1, γ_2 disagree on $\mathcal{S}_1 \cap \mathcal{S}_2$. The disagreement means that there is some vertex $(C_{i_j}, \alpha_j) \in \mathcal{S}_1 \cap \mathcal{S}_2$ such that $\gamma_1((C_{i_j}, \alpha_j)) \neq \gamma_2((C_{i_j}, \alpha_j))$. If T_{i_j} is the tuple of variables corresponding to C_{i_j}, then $[\alpha_j, \gamma_1](T_{i_j}) \neq [\alpha_j, \gamma_2](T_{i_j})$. Assuming neither of $\overline{\mathbf{V}}_{(\mathcal{S}_1, \gamma_1)}$ and $\overline{\mathbf{V}}_{(\mathcal{S}_2, \gamma_2)}$ is zero, we must have that $\overline{\mathbf{V}}_{(\mathcal{S}_1, \gamma_1)} = \mathbf{V}_{(S_1', \alpha_1')}$ and $\overline{\mathbf{V}}_{(\mathcal{S}_2, \gamma_2)} = \mathbf{V}_{(S_2', \alpha_2')}$ for some $S_1', S_2' \subseteq [n]$ and partial assignments α_1', α_2'. Also, we have that $T_{i_j} \subseteq \mathcal{S}_1 \cap \mathcal{S}_2$ and $\alpha_1'(T_{i_j}) = [\alpha_j, \gamma_1](T_{i_j}) \neq [\alpha_j, \gamma_2](T_{i_j}) = \alpha_2'(T_{i_j})$. This gives $\left\langle \mathbf{V}_{(S_1', \alpha_1')}, \mathbf{V}_{(S_2', \alpha_2')} \right\rangle = 0$.

- We also need to show that $\left\langle \overline{\mathbf{V}}_{(\mathcal{S}_1, \gamma_1)}, \overline{\mathbf{V}}_{(\mathcal{S}_2, \gamma_2)} \right\rangle = \left\langle \overline{\mathbf{V}}_{(\mathcal{S}_3, \gamma_3)}, \overline{\mathbf{V}}_{(\mathcal{S}_4, \gamma_4)} \right\rangle$ whenever $\mathcal{S}_1 \cup \mathcal{S}_2 = \mathcal{S}_3 \cup \mathcal{S}_4$ and $\gamma_1 \circ \gamma_2 = \gamma_3 \circ \gamma_4$. For convenience, we only show this when all sets have size at most $t/2k$ by showing that $\left\langle \overline{\mathbf{V}}_{(\mathcal{S}_1, \gamma_1)}, \overline{\mathbf{V}}_{(\mathcal{S}_2, \gamma_2)} \right\rangle = \left\langle \overline{\mathbf{V}}_{(\mathcal{S}_1 \cup \mathcal{S}_2, \gamma_1 \circ \gamma_2)}, \overline{\mathbf{V}}_{(\emptyset, \emptyset)} \right\rangle$. Again, assuming neither of these vectors are zero, let $\overline{\mathbf{V}}_{(\mathcal{S}_1, \gamma_1)} = \mathbf{V}_{(S_1', \alpha_1')}$ and $\overline{\mathbf{V}}_{(\mathcal{S}_2, \gamma_2)} = \mathbf{V}_{(S_2', \alpha_2')}$.

  If α_1' and α_2' contradict (note that this may happen even when $\gamma_1 \circ \gamma_2$ is defined) then there must be some vertices $(C_{i_1}, \alpha_1) \in \mathcal{S}_1$ and $(C_{i_2}, \alpha_2) \in \mathcal{S}_2$ such that C_{i_1} and C_{i_2} involve a common variable, on which the shifted assignments $[\alpha_1, \gamma_1]$ and $[\alpha_2, \gamma_2]$ disagree. But then both these vertices will also be present in $\mathcal{S}_1 \cup \mathcal{S}_2$ and the assignments shifted according to $\gamma_1 \circ \gamma_2$ will also disagree. Hence, $\overline{\mathbf{V}}_{(\mathcal{S}_1 \cup \mathcal{S}_2, \gamma_1 \circ \gamma_2)} = 0$ which satisfies the condition in this case. If not, we must have that $\overline{\mathbf{V}}_{(\mathcal{S}_1 \cup \mathcal{S}_2, \gamma_1 \circ \gamma_2)} = \mathbf{V}_{(S_3', \alpha_3')}$. Since the vectors $\mathbf{V}_{(\cdot, \cdot)}$ for a valid CSP solution, it will be sufficient to show that $S_3' = S_1' \cup S_2'$ and $\alpha_3' = \alpha_1' \circ \alpha_2'$. Since S_3' contains all the variables involved in constraints present either in $\mathcal{S}_1$ or $\mathcal{S}_2$, it must include all variables in $S_1' \cup S_2'$. Finally, for any $(C_{i_j}, \alpha_j) \in \mathcal{S}_1 \cup \mathcal{S}_2$, $\alpha_3'(T_{i_j}) = [\alpha_j, \gamma_1 \circ \gamma_2](T_{i_j}) = (\alpha_1' \circ \alpha_2')(T_{i_j})$, which proves the required condition.

- Finally, we need to verify that for every vertex (C_i, α) of G_Φ, $\sum_{j \in [l]} \left\| \overline{\mathbf{V}}_{(\{(C_i, \alpha)\}, j)} \right\|^2 = 1$. Note that $\overline{\mathbf{V}}_{(\{(C_i, \alpha)\}, j)} = \mathbf{V}_{(T_i, \alpha + w_j)}$ where w_j is a vector on $\mathbb{F}_2$ such that $A \cdot w_j^\mathsf{T} = 0$. If the constraint C_i is of the form $A \cdot x = b_i$ and if α is a satisfying assignment, then as j ranges from 1 to l, $(\alpha + w_j)$ ranges over all the satisfying assignments to the constraint C_i. Hence, we have that

$$\sum_{j \in [l]} \left\| \overline{\mathbf{V}}_{(\{(C_i, \alpha)\}, j)} \right\|^2 = \sum_{j \in [l]} \left\| \mathbf{V}_{(T_i, \alpha + w_j)} \right\|^2 = \sum_{\alpha \in \{0,1\}^{T_i}} C_i(\alpha) \left\| \mathbf{V}_{(T_i, \alpha)} \right\|^2 = 1$$

19

where the last equality used the fact that $\mathsf{FRAC}(\Phi) = m$ and hence the contribution to the SDP value, from assignments of each constraint, is 1.

∎

This now gives the claimed gap for Approximate Graph Coloring.

Theorem 6.5 *For every constant l there is a $c = c(l)$ and an infinite family of graphs $G = (V, E)$ with chromatic number $\Omega\left(\frac{2^{l/2}}{l^2}\right)$, and such that G has a vector coloring with l colors for the SDP obtained by $c \cdot |V|$ rounds of the Lasserre hierarchy.*

Proof: For any l, there is a k such that $l/2 \leq 2^{\lceil \log(k+1)\rceil} \leq l$. For this k, consider an instance Φ of MAX k-CSP with n variables and $m = \beta n$ constraints as given by Corollary 6.2, choosing $\epsilon = 1/2$. We take our graph G to be G_Φ as defined above. Claim 6.3 shows that the largest independent set has size at most $\mathsf{OPT}_k(\Phi)$, which is at most $(3l/2^{k+1}) \cdot m$ by Corollary 6.2. Since the number of vertices in G (say N) is at least $k \cdot m$, its chromatic number is $\Omega(2^k/l^2) = \Omega(2^{l/2}/l^2)$.

On the other hand, we have SDP solutions for Φ for $c'n$ rounds (with $c' = c'(k)$) with $\mathsf{FRAC}(\Phi) = m$. By Lemma 6.4 G has a vector coloring $2^{\lceil \log(k+1)\rceil}$ colors for the SDP obtained by $c'n/k = cN$ rounds of Lasserre, where c depends only on k (which depends only on l). ∎

6.2 Gaps for Chromatic Number

We now modify the graph and the SDP solution constructed in the previous section to get strong gaps for Chromatic Number. As in the case of independent sets, we define the product graph $G_\Phi^r = (V_\Phi^r, E')$ for G_Φ defined above. Two vertices $\{(C_{i_1}, \alpha_1), \ldots, (C_{i_r}, \alpha_r)\}$ and $\{(C_{i'_1}, \alpha'_1), \ldots, (C_{i'_r}, \alpha'_r)\}$ in V_Φ^r are connected if for some j, $\{(C_{i_j}, \alpha_j), (C_{i'_j}, \alpha'_j)\} \in E_\Phi$. Note that the edge set E_Φ is slightly different than it was in the case of independent set. $\mathcal{C}(i_1, \ldots, i_r)$ is defined as before

$$\mathcal{C}(i_1, \ldots, i_r) = \{\{(C_{i_1}, \alpha_1), \ldots, (C_{i_r}, \alpha_r)\} \mid \wedge_{j=1}^r C_{i_j}(\alpha_j) = 1\}$$

We argue that if G_Φ has a vector coloring with l colors, then G_Φ has a vector coloring with l^r colors. We think of the l^r colors as r-tuples of values in $[l]$. Hence, a partial assignment assigns to each vertex a tuple in $[l]^r$.

Claim 6.6 *If there is a feasible solution for the l-coloring SDP for G_Φ obtained by t rounds of the Lasserre hierarchy, then there is also a feasible solution for the SDP for l^r-coloring the graph G_Φ^r obtained by t rounds.*

Proof: We define the vector $\overline{\mathbf{V}}_{(\overline{\mathcal{S}}, \overline{\gamma})}$ for all $\overline{\mathcal{S}} \subseteq V_\Phi^r$, $|\overline{\mathcal{S}}| \leq t$ and $\overline{\gamma} \in ([l]^r)^{\overline{\mathcal{S}}}$. Each vertex $v \in \overline{\mathcal{S}}$ is of the form $\{((C_{i_1}, \alpha_1), \ldots, (C_{i_r}, \alpha_r))\}$. For a such vertex v, let $[v]_j$ denote the element in jth coordinate of v i.e. (C_{i_j}, α_j). Also, $\overline{\gamma}(v)$ is an r-tuple $(l_1, \ldots, l_r)$ and we denote the jth coordinate l_j by $[\overline{\gamma}(v)]_j$. Given a pair $(\overline{\mathcal{S}}, \overline{\gamma})$, we break it into different projection sets P_j for each $1 \leq j \leq r$

$$P_j = \{([v]_j, [\overline{\gamma}(v)]_j) \mid v \in \overline{\mathcal{S}}\}$$

Each element in P_j corresponds to a vertex in G_Φ (given by $[v]_j$) and a color in $[l]$ for the vertex (given by $[\overline{\gamma}(v)]_j$). Note that there can be two different elements $((C_i, \alpha), l_j)$ and $((C_i, \alpha), l'_j)$ which

assign different colors to the same vertex. If this is the case for any set P_j, we take $\overline{\mathbf{V}}_{(\overline{S},\overline{\gamma})} = 0$. Otherwise, for each set P_j, we can define the set $\mathcal{S}_j$ of vertices of G_Φ that are contained in P_j and also a partial assignment $\gamma_j \in [l]^{\mathcal{S}_j}$, since every vertex of $\mathcal{S}_j$ gets a unique color by assumption. In this case, we define $\overline{\mathbf{V}}_{(\overline{S},\overline{\gamma})}$ by tensoring assignment vectors in each coordinate. Formally,

$$\overline{\mathbf{V}}_{(\overline{S},\overline{\gamma})} = \begin{cases} 0 & \exists 1 \leq j \leq r \ \& \ ((C_i,\alpha),l_j),((C_i,\alpha),l'_j) \in P_j \ s.t. \ l_j \neq l'_j \\ \overline{\mathbf{V}}_{(\mathcal{S}_1,\gamma_1)} \otimes \ldots \otimes \overline{\mathbf{V}}_{(\mathcal{S}_r,\gamma_r)} & \text{otherwise} \end{cases}$$

It is easy to verify that the vectors satisfy all the required SDP conditions.

- Let $u_1 = \{(C_{i_1},\alpha_1),\ldots,(C_{i_r},\alpha_r)\}$ and $u_2 = \{(C_{i_1},\alpha_1),\ldots,(C_{i_r},\alpha_r)\}$ be two adjacent vertices, and let $\overline{\gamma} \in [l]^r$ be any color $(l_1,\ldots,l_r)$. Then, by adjacency, we must have that for some $j \leq r$, $\{(C_{i_j},\alpha_j),(C_{i'_j},\alpha'_j)\} \in E_\Phi$. Hence,

$$\left\langle \overline{\mathbf{V}}_{(\{u_1\},\overline{\gamma})}, \overline{\mathbf{V}}_{(\{u_2\},\overline{\gamma})} \right\rangle = \prod_{j=1}^{r} \left\langle \overline{\mathbf{V}}_{(\{(C_{i_j},\alpha_j)\},l_j)}, \overline{\mathbf{V}}_{(\{(C_{i'_j},\alpha'_j)\},l_j)} \right\rangle = 0$$

- Similarly, if $(\overline{S}_1,\overline{\gamma}_1)$ and $(\overline{S}_2,\overline{\gamma}_2)$ have a contradiction, or $\overline{S}_1 \cup \overline{S}_2 = \overline{S}_3 \cup \overline{S}_4$ and $\overline{\gamma}_1 \circ \overline{\gamma}_2 = \overline{\gamma}_3 \circ \overline{\gamma}_4$, then these conditions will also hold in each of the coordinatewise projections. Hence, the SDP conditions will be satisfied in these cases.

- To verify that for each $u \in V$, $\sum_{j \in [l]^r} \left\| \overline{\mathbf{V}}_{(\{u\},j)} \right\|^2 = 1$ we again note that

$$\sum_{l_1,\ldots,l_r} \left\| \overline{\mathbf{V}}_{(\{(C_{i_1},\alpha_1),\ldots,(C_{i_r},\alpha_r)\},(l_1,\ldots,l_r))} \right\|^2 = \sum_{l_1,\ldots,l_r} \prod_{j=1}^{r} \left\langle \overline{\mathbf{V}}_{(\{(C_{i_j},\alpha_j)\},l_j)}, \overline{\mathbf{V}}_{(\{(C_{i_j},\alpha_j)\},l_j)} \right\rangle$$

$$= \prod_{j=1}^{r} \sum_{l_j} \left\| \overline{\mathbf{V}}_{(\{(C_{i_j},\alpha_j)\},l_j)} \right\|^2 = 1$$

■

We now prove the integrality gap for Chromatic Number by similar arguments as in Theorem 5.3.

Theorem 6.7 *There exist constants $c_1, c_2, c_3 > 0$ and graphs G on N vertices, for arbitrarily large N such that*

1. *The chromatic number of G is $\Omega\left(\frac{N}{2^{c_1\sqrt{\log N \log\log N}}}\right)$.*

2. *The SDP for coloring G obtained by $\Omega\left(2^{c_2\sqrt{\log N \log\log N}}\right)$ rounds of the Lasserre hierarchy admits a vector coloring with $O\left(2^{c_3\sqrt{\log N \log\log N}}\right)$ colors.*

Proof: We construct the graph G by sampling M cliques of the form $\mathcal{C}(i_1,\ldots,i_r)$ from G_Φ^r, and considering the subgraph induces by their vertices. The size of the independent sets is small w.h.p. over the choice of G.

Claim 6.8 *Let $s = \mathsf{OPT}_k(\Phi)/m$. Then for $M \geq \frac{100nr}{s^r}$, with probability $1 - o(1)$, all independent sets in G have size at most $2s^r M$.*

Proof: By Chernoff bound arguments identical to those in Claim 5.4. ∎

We again choose $k = \delta \log n$ for some small constant δ, and let $r = \log n / (\log \log n)$ for a large n. Applying Corollary 6.2 with $\epsilon = 1/2$ gives an instance Φ of MAX k-CSP(P_A) with $k/2^k \leq s \leq 3k/2^k$. (Note that here $s = \mathsf{OPT}_k(\Phi)/m$). The number of assignments to each constraint is exactly $l = 2^{\lceil \log(k+1) \rceil} \leq 2k$. We again pick $M = 100nr \cdot (2^{kr}/k^r)$.

With high probability over the choice of G, the size of the maximum independent set in G is at most $2Ms^r$. The number of vertices in G is

$$N \;=\; M \cdot l^r \;\leq\; (100nr \cdot 2^{kr}/k^r) \cdot (2k)^r \;=\; O(nr \cdot 2^{(k+1)r})$$

and hence the chromatic number of G is at least $(l^r/2s^r)$, which is $\dfrac{N}{2^{c_1 \sqrt{\log N \log \log N}}}$ for some constant c_1, with our parameters.

We can again take the Lasserre vectors corresponding to sets of vectors in G, to be the same as the vectors for the corresponding sets in G_Φ^r. By Claim 6.6 the number of colors in the vector coloring is l^r, which is at most $2^{c_2 \sqrt{\log N \log \log N}}$ for some constant c_2. Also, the number of rounds is $\Omega\left(\frac{n}{k \cdot \beta^{80}}\right)$, which is $\Omega\left(2^{c_3 \sqrt{\log N \log \log N}}\right)$ for $c_3 > 0$, if δ (in choosing $k = \delta \log n$) is small enough. ∎

7 Integrality Gaps for Vertex Cover

In this section, we prove an intergrality gap of 1.36 for Minimum Vertex Cover using the reduction by Dinur and Safra [10]. One can start with an integer program for Minimum Vertex Cover and obtain the Lasserre SDP for Vertex Cover by introducing vector variables for every set of t vertices in the graph. Equivalently (and this is the form we will be using), one can simply work with the t-round SDP for Maximum Independent Set and modify the objective.

We collect the relation between SDP solutions for Maximum Independent Set and Minimum Vertex Cover, together with some other (known) characterizations of the independent set solution which we shall need, in the lemma below. Proof of the lemma is deferred to the appendix.

Lemma 7.1 *Let the vectors $\mathbf{U}_S$ for $|S| \leq t$ form a solution to the t-round Lasserre SDP for* Maximum Independent Set *on a weighted graph $G = (V, E)$, with SDP value $\mathsf{FRAC}(G)$. Then there exist vectors $\mathbf{V}_{(S,\alpha)}$ for all $|S| \leq t/2, \alpha \in \{0,1\}^S$, determined by the vectors $\mathbf{U}_S$, such that*

1. $\mathbf{U}_\emptyset = \mathbf{V}_{(\emptyset,\emptyset)}$ *and* $\mathbf{U}_S = \mathbf{V}_{(S,1_S)} \forall S$, *where 1_S is the partial assignment which assigns 1 to all elements in S.*

2. *The vectors $\mathbf{V}_{(S,\alpha)}$ satisfy all conditions of the SDP for constraint satisfaction problems.*

3. *For any S, the vectors $\{\mathbf{V}_{(S,\alpha)} \mid \alpha \in \{0,1\}^S\}$ induce a probability distribution over $\{0,1\}^S$. The events measurable in this probability space correspond to all $\alpha' \in \{0,1\}^{S'}$ for all $S' \subseteq S$, and $\mathbb{P}[\alpha'] = \left\|\mathbf{V}_{(S',\alpha')}\right\|^2$.*

4. *The vectors $\mathbf{V}_{(S,0_S)}$ form a solution to the t-round SDP for Minimum Vertex Cover with objective value $\sum_{v \in V} w(v) - \mathsf{FRAC}(G)$, where $w(v)$ denotes the weight of vertex v and 0_S denotes the all-zero assignment.*

Through the remaining part of this section, we shall only consider the independent set SDP on all the graphs in the reduction. We will show that the value of the fractional independent set is large for all intermediate graphs obtained in the reduction. On the other hand, it willll be possible to conclude directly from the correctness of the Dinur-Safra proof that the size of the actual independent set in these graphs is small. Comparing the corresponding values for vertex cover will give us the required integrality gap.

7.1 The starting graphs for the Dinur-Safra reduction

We first describe the graphs required for the reduction by Dinur and Safra [10]. They require few key properties of the graph which they use to argue the soundness of the reduction i.e. the graphs produces by the reduction have no large independent set. In fact, graphs of the form G_Φ^r as defined in section 5 turn out to satisfy all the required conditions. Also, we already have vector solutions for the independent set SDP on these graphs. We only need to argue that these vectors can be transformed appropriately through the steps of the reduction. Dinur and Safra define the notion of "co-partite" graphs as below.

Definition 7.2 *We say that a graph $G = (M \times L, E)$ is (m, l) co-partite, if it is composed of $m = |M|$ disjoint cliques, each of size $l = |L|$. The edges that go between the cliques may be arbitrary. Formally, for all $i \in M$ and $j_1 \neq j_2 \in L$, we require that $\{(i, j_1), (i, j_2)\} \in E$.*

Let Φ be any CSP instance with m constraints and each constraint having exactly l satisfying assignments. Then it is easy to see that the FGLSS graph G_Φ is (m, l) co-partite. Also, the graph G_Φ^r is (m^r, l^r) co-partite. The reduction in [10] also requires an (m, l) co-partite graphs such that for some fixed constants $\epsilon_0, h > 0$ every subset of vertices $I \subseteq M \times L$ with $|I| \geq \epsilon_0 m$, contains a clique of size h. It also follows from their argument (proof of Theorem 2.1 in [10]) [5] that if $\mathsf{OPT}(\Phi) \leq s \cdot m$ for some $s < 1$, then G_Φ^r satisfies this property for an appropriate r.

Theorem 7.3 ([10]) *Let Φ be a CSP instance with m constraints, each having l satisfying assignments, and such that any assignment satisfies at most $s < 1$ fraction of the constraints. Also, let $\epsilon_0, h > 0$ be given. Then there exists an $r = O(\log(h/\epsilon))$ such that any set of vertices in G_Φ^r, which does not contain an h-clique, has size at most $\epsilon_0 \cdot m^r$.*

7.2 The graphs with block-assignments

The next step of the reduction, which is crucial for the soundness, transforms a graph $G = (V, E)$ which is (m_0, l_0) co-partite into a new graph $G_{\mathcal{B}}$ which is (m_1, l_1) co-partite and has some additional properties required for the soundness.

We consider the set of *blocks* of d vertices in V, i.e. the set

$$\mathcal{B} = \binom{V}{d} = \{B \subseteq V \mid |B| = d\}$$

Also, for each block B, let L_B denote all "large" partial assignments to vertices in B. Formally, $L_B = \{\alpha \in \{0,1\}^B \mid |\alpha| \geq d_{\mathsf{T}}\}$, where $d_{\mathsf{T}} = d/2l_0$ and $|\alpha|$ is the number of 1s in the image of α.

---

[5]The result in [10] is actually stated not for the graph G_Φ^r, but for a graph G' obtained by converting the CSP Φ to a two-player game and considering the graph obtained by parallel repetition. However, the graph G' defined in their paper is a spanning subgraph of G_Φ^r, and hence if a subset of vertices contains an h-clique in G', then it also contains one in G_Φ^r.

The vertex set of the graph $G_\mathcal{B}$ is taken to be set of all pairs of the form (B, α), where $\alpha \in L_B$. To define the edges, we consider a pair of blocks whose symmetric difference is just a pair of vertices (v_1, v_2) such that $(v_1, v_2) \in E$. We connect two partial assignments corresponding to such a pair of blocks, if they form an "obvious contradiction" i.e. they disagree on the intersection or they assign 1 to both the vertices in the symmetric difference. It is important for the soundness analysis in [10] that for any such pair of blocks (B_1, B_2), any $\alpha_1 \in L_{B_1}$ is *not connected* to at most two partial assignments in B_2 (and vice-versa). We also add edges between all partial assignments within a block. Thus, we define

$$
\begin{aligned}
V_\mathcal{B} &= \{(B, \alpha) \mid B \in \mathcal{B}, \alpha \in L_B\} \\
E_\mathcal{B}^{(1)} &= \bigcup_{\substack{\hat{B} \in \binom{V}{d-1} \\ (v_1, v_2) \in E}} \left\{ (\hat{B} \cup \{v_1\}, \alpha_1), (\hat{B} \cup \{v_2\}, \alpha_2) \mid \alpha_1(\hat{B}) \neq \alpha_2(\hat{B}) \text{ or } \alpha_1(v_1) = \alpha_2(v_2) = 1 \right\} \\
E_\mathcal{B} &= E_\mathcal{B}^{(1)} \cup \left(\bigcup_B \{(B, \alpha_1), (B, \alpha_2) \mid \alpha_1 \neq \alpha_2, \ \alpha_1, \alpha_2 \in L_B\} \right)
\end{aligned}
$$

Note that $G_\mathcal{B}$ is also (m_1, l_1) co-partite with $m_1 = |\mathcal{B}|$ and $l_1 = |L_B|$ (which is the same for all B). We now show that if G has a good SDP solution, then so does $G_\mathcal{B}$.

Lemma 7.4 *Let $G = (V, E)$ be an (m_0, l_0) co-partite graph such that for the independent set SDP obtained by t rounds of the Lasserre hierarchy, $\mathsf{FRAC}(G) = m_0$. For given $\epsilon > 0$ and $d_T > 2/\epsilon_1$, let $G_\mathcal{B}$ be constructed as above. Then, for the independent set SDP obtained by t/d rounds, $\mathsf{FRAC}(G_\mathcal{B}) \geq (1 - \epsilon_1)|\mathcal{B}|$.*

Proof: Since G admits a solution for t rounds of the Lasserre SDP, we also have vectors $\mathbf{V}_{(S,\alpha)}$ for all $|S| \leq t, \alpha \in \{0, 1\}^S$ as described in Lemma 7.1. We now use these vectors to define the vector solution for the SDP on $G_\mathcal{B}$. Each vertex of $G_\mathcal{B}$ is of the form (B, α) where $B \in \mathcal{B}$ and $\alpha \in L_B$. Consider a set $\mathcal{S}$ of $i \leq t/d$ vertices, $\mathcal{S} = \{(B_1, \alpha_1), \ldots, (B_i, \alpha_i)\}$. As in the Section 5, we let the vector corresponding to this set be 0 if any two assignments in the set contradict and equal to the vector for the partial assignment jointly defined by $\alpha_1, \ldots, \alpha_i$ otherwise.

$$
\mathbf{U}_\mathcal{S} = \begin{cases} 0 & \exists j_1, j_2 \text{ s.t. } \alpha_{j_1}(B_{j_1} \cap B_{j_2}) \neq \alpha_{j_2}(B_{j_1} \cap B_{j_2}) \\ \mathbf{V}_{(\cup_j B_j, \alpha_1 \circ \ldots \circ \alpha_i)} & \text{otherwise} \end{cases}
$$

If (B_1, α_1) and (B_2, α_2) have an edge between them because α_1 and α_2 contradict, then we must have $\langle \mathbf{U}_{\{(B_1, \alpha_1)\}}, \mathbf{U}_{\{(B_2, \alpha_2)\}} \rangle = \langle \mathbf{V}_{(B_1, \alpha_1)}, \mathbf{V}_{(B_2, \alpha_2)} \rangle = 0$ since CSP vectors corresponding to contradicting partial assignments must be orthogonal. For an edge between $(\hat{B} \cup \{v_1\}, \alpha_1)$ and $(\hat{B} \cup \{v_2\}, \alpha_2)$ where $(v_1, v_2) \in E$, $\alpha_1(v_1) = \alpha_2(v_2) = 1$ and $\alpha_1(\hat{B}) = \alpha_2(\hat{B}) = \alpha$ (say), we have

$$
\left\langle \mathbf{U}_{\{(\hat{B} \cup \{v_1\}, \alpha_1)\}}, \mathbf{U}_{\{(\hat{B} \cup \{v_2\}, \alpha_2)\}} \right\rangle = \left\langle \mathbf{V}_{(\hat{B}, \alpha)}, \mathbf{V}_{(\{v_1, v_2\}, (1,1))} \right\rangle = \left\langle \mathbf{V}_{(\hat{B}, \alpha)}, 0 \right\rangle = 0
$$

because in the independent solution on graph G

$$
\left\| \mathbf{V}_{(\{v_1, v_2\}, (1,1))} \right\|^2 = \left\langle \mathbf{V}_{(\{v_1\}, 1)}, \mathbf{V}_{(\{v_2\}, 1)} \right\rangle = 0
$$

The proof that the vectors defined above satisfy the other SDP conditions is identical to that in Claim 5.2 and we omit the details.

The interesting part of the argument will be to show that the value of the independent set SDP for $G_\mathcal{B}$ will be large. In the completeness part of the Dinur-Safra reduction, one needs to say that if G

24

has a *large independent set* then so does $G_\mathcal{B}$, and it follows very easily using a Chernoff bound on how different blocks intersect the large independent set of G. We need to make the same conclusion about the SDP value and the argument is no longer applicable. However, it is possible to get the conclusion by looking at the "local distributions" defined by the vector $\mathbf{V}_{(B,\alpha)}$ as mentioned in Lemma 7.1. We then combine the bounds obtained in each block globally using properties of the vector solution.

For each block B, let D_B denote the distribution over $\{0,1\}^B$ defined by the vectors $\mathbf{V}_{(B,\alpha)}$ for all $\alpha \in \{0,1\}^B$. For each block, we define a random variable Z_B determined by the event $\alpha \in \{0,1\}^B$ with value $Z_B = |\alpha|$. One can then convert the statement about the SDP value to a statement about the local distributions, by noting that

$$\mathsf{FRAC}(G_\mathcal{B}) = \sum_{B \in \mathcal{B}} \sum_{\alpha \in L_B} \left\| \mathbf{V}_{(B,\alpha)} \right\|^2 = \sum_{B \in \mathcal{B}} \left(1 - \sum_{\alpha \notin L_B} \left\| \mathbf{V}_{(B,\alpha)} \right\|^2 \right) = |\mathcal{B}| - \sum_{B \in \mathcal{B}} \mathbb{P}_{D_B} [Z_B < d_\mathsf{T}]$$

The problem thus reduces to showing that $\sum_{B \in \mathcal{B}} \mathbb{P}_{D_B}[Z_B < d_\mathsf{T}] \leq \epsilon_1 |\mathcal{B}|$. By a second moment analysis on every local distribution, we have

$$\mathbb{P}_{D_B}[Z_B < d_\mathsf{T}] \ \leq \ \mathbb{P}_{D_B}\left[|Z_B - 2d_\mathsf{T}| > d_\mathsf{T}\right] \ \leq \ \frac{\mathbb{E}_{D_B}\left[(Z_B - 2d_\mathsf{T})^2\right]}{d_\mathsf{T}^2} \tag{1}$$

The following claim provides the necessary estimates to bound the sum of probabilities.

Claim 7.5

$$\sum_{B \in \mathcal{B}} \mathbb{E}_{D_B}[Z_B] = 2d_\mathsf{T} \cdot |\mathcal{B}| \quad and \quad \sum_{B \in \mathcal{B}} \mathbb{E}_{D_B}[Z_B^2] \leq (4d_\mathsf{T}^2 + 2d_\mathsf{T}) \cdot |\mathcal{B}|$$

Proof: For all blocks B and $v \in B$, define the random variable $X_{v,B}$ which is 1 if a random α chosen according to D_B assigns 1 to the vertex v and 0 otherwise. By using the fact that the distribution is defined by the vectors $\mathbf{V}_{(B,\alpha)}$, we get that for $v \in B$

$$\mathbb{E}_{D_B}[X_{v,B}] \ = \ \mathbb{P}_{D_B}[\alpha(v) = 1] \ = \ \left\| \mathbf{V}_{(\{v\},1)} \right\|^2 \ = \ \left\| \mathbf{U}_{\{v\}} \right\|^2$$

where $\mathbf{U}_{\{v\}}$ denotes the vector for vertex v in the solution for graph G. Similarly,

$$\mathbb{E}_{D_B}[X_{v_1,B} X_{v_2,B}] \ = \ \mathbb{P}_{D_B}[\alpha(\{v_1, v_2\}) = (1,1)] \ = \ \left\| \mathbf{V}_{(\{v_1,v_2\},(1,1))} \right\|^2 \ = \ \left\langle \mathbf{U}_{\{v_1\}}, \mathbf{U}_{\{v_2\}} \right\rangle$$

With the above relations, and using the facts that each vertex appears in exactly $d/m_0 l_0$ fraction of the blocks and $\mathsf{FRAC}(G) = m_0$, we can compute the sum of expectations of Z_B as

$$\sum_{B \in \mathcal{B}} \mathbb{E}_{D_B}[Z_B] \ = \ \sum_{B \in \mathcal{B}} \mathbb{E}_{D_B}\left[\sum_{v \in B} X_{v,B}\right] \ = \ \sum_{B \in \mathcal{B}} \sum_{v \in B} \left\| \mathbf{U}_{\{v\}} \right\|^2 \ = \ \frac{d}{m_0 l_0} |\mathcal{B}| \left(\sum_{v \in V} \left\| \mathbf{U}_{\{v\}} \right\|^2 \right)$$

$$= \ \frac{d}{m_0 l_0} |\mathcal{B}| \cdot m_0 \ = \ 2d_\mathsf{T} \cdot |\mathcal{B}|$$

Similarly, for the expectations of the squares, we get

$$\sum_{B \in \mathcal{B}} \mathbb{E}_{D_B}[Z_B^2] \ = \ \sum_{B \in \mathcal{B}} \sum_{v_1,v_2 \in B} \mathbb{E}_{D_B}[X_{v_1,B} X_{v_2,B}] \ = \ \sum_{B \in \mathcal{B}} \sum_{v_1,v_2 \in B} \left\langle \mathbf{U}_{\{v_1\}}, \mathbf{U}_{\{v_2\}} \right\rangle$$

25

Again, each pair (v_1, v_2) such that $v_1 \neq v_2$ appears in less than $(d^2/m_0^2 l_0^2)$ fraction of the blocks, and a pair such that $v_1 = v_2$ appears in $d/m_0 l_0$ fraction. Hence,

$$
\begin{aligned}
\sum_{B \in \mathcal{B}} \mathbb{E}_{D_B}[Z_B^2] & \leq |\mathcal{B}| \left(\frac{d^2}{m_0^2 l_0^2} \right) \sum_{v_1, v_2} \langle \mathbf{U}_{\{v_1\}}, \mathbf{U}_{\{v_2\}} \rangle + |\mathcal{B}| \left(\frac{d}{m_0 l_0} \right) \sum_v \|\mathbf{U}_{\{v\}}\|^2 \\
& = |\mathcal{B}| \left(\frac{d^2}{m_0^2 l_0^2} \right) \left\| \sum_v \mathbf{U}_{\{v\}} \right\|^2 + |\mathcal{B}| \cdot 2 d_\mathsf{T}
\end{aligned}
$$

Finally, to calculate the first term, we shall need the fact that G is (m_0, l_0) co-partite. Let $V = M_0 \times L_0$. We write each v as (i, j) for $i \in M_0, j \in L_0$. Using the fact that all vectors within a single clique are orthogonal, we get

$$
\left\| \sum_{(i,j) \in M_0 \times L_0} \mathbf{U}_{\{(i,j)\}} \right\|^2 \leq m_0 \cdot \sum_{i \in M_0} \left\| \sum_{j \in L_0} \mathbf{U}_{\{(i,j)\}} \right\|^2 = m_0 \cdot \sum_{i \in M_0} \sum_{j \in L_0} \|\mathbf{U}_{\{(i,j)\}}\|^2 = m_0^2.
$$

Using this bound we get that $\sum_{B \in \mathcal{B}} \mathbb{E}_{D_B}[Z_B^2] \leq (4 d_\mathsf{T}^2 + 2 d_\mathsf{T})|\mathcal{B}|$, which proves the claim. ∎

Using equation (1) and the previous claim, we get that

$$
\sum_{B \in \mathcal{B}} \mathbb{P}_{D_B}[Z_B < d_\mathsf{T}] \leq \frac{1}{d_\mathsf{T}^2} \sum_{B \in \mathcal{B}} (\mathbb{E}_{D_B}[Z_B^2] - 2 d_\mathsf{T} \mathbb{E}_{D_B}[Z_B] + 4 d_\mathsf{T}^2) \leq \frac{2}{d_\mathsf{T}} |\mathcal{B}|
$$

Hence, for $d > 2/\epsilon_1$, the SDP value is at least $(1 - \epsilon_1)|\mathcal{B}|$. ∎

7.3 The long-code step

The next step of the reduction defines a weighted graph starting from an (m_1, l_1) co-partite graph. Let $G = (V, E)$ be the given graph and $V = M_1 \times L_1$. We then define the graph $G_{\mathcal{LC}}$ which has a vertex for every $i \in M_1$ and every $J \subseteq L_1$. Also, the graph is weighted with each vertex (i, J) having a weight $w(i, J)$ depending on $|J|$.

$$
\begin{aligned}
V_{\mathcal{LC}} & = \{(i, J) \mid i \in M_1, J \subseteq L_1\} \\
E_{\mathcal{LC}} & = \{\{(i_1, J_1), (i_2, J_2)\} \mid \forall j_1 \in J_1, j_2 \in J_2. \; \{(i_1, j_1), (i_2, j_2)\} \in E\} \\
w(i, J) & = \frac{1}{m_1} p^{|J|} (1 - p)^{|L_1 \setminus J|}
\end{aligned}
$$

Lemma 7.6 *Let $G = (V, E)$ be an (m_1, l_1) co-partite graph such that independent set SDP for t-rounds on G has value at least $\mathsf{FRAC}(G)$. Let $G_{\mathcal{LC}}$ be defined as above. Then there is a solution to the $t/2$-round SDP for independent set on $G_{\mathcal{LC}}$ with value at least $p \cdot \frac{\mathsf{FRAC}(G)}{m_1}$.*

Proof: Let us denote the vectors in G by $\mathbf{U}_S$ and those in $G_{\mathcal{LC}}$ by $\overline{\mathbf{U}}_S$. We define the vectors $\mathbf{U}_S$ for each $\mathcal{S} = \{(i_1, J_1), \ldots, (i_r, J_r)\}$ for $r \leq t$ as

$$
\overline{\mathbf{U}}_S = \sum_{j_1 \in J_1, \ldots, j_r \in J_r} \mathbf{U}_{\{(i_1, j_1), \ldots, (i_r, j_r)\}}
$$

We now need to verify that they satisfy all the SDP conditions and the SDP value is as claimed.

26

- We first check that vectors for adjacent vertices are orthogonal. Let $\{(i_1, J_1), (i_2, J_2)\}$ be an edge in $G_{\mathcal{LC}}$. Then

$$\left\langle \overline{\mathbf{U}}_{\{(i_1,J_1)\}}, \overline{\mathbf{U}}_{\{(i_2,J_2)\}} \right\rangle \;=\; \sum_{j_1 \in J_1, j_2 \in J_2} \left\langle \mathbf{U}_{\{(i_1,j_1)\}}, \mathbf{U}_{\{(i_2,j_2)\}} \right\rangle \;=\; 0$$

since all pairs $\{(i_1, j_1), (i_2, j_2)\}$ form edges in G.

- For convenience, we shall only verify $\left\langle \overline{\mathbf{U}}_{\mathcal{S}_1}, \overline{\mathbf{U}}_{\mathcal{S}_2} \right\rangle = \left\langle \overline{\mathbf{U}}_{\mathcal{S}_1 \cup \mathcal{S}_2}, \overline{\mathbf{U}}_{\emptyset} \right\rangle$ for all $\mathcal{S}_1, \mathcal{S}_2$ with at most $t/2$ elements. This will prove that vectors $\overline{\mathbf{U}}_{\mathcal{S}}$ with $|\mathcal{S}| \leq t/2$ satisfy the SDP conditions. First, we observe that using the orthogonality of vectors in G corresponding to (i, j_1) and (i, j_2) for $j_1 \neq j_2$,

$$\overline{\mathbf{U}}_{\{(i,J_1),(i,J_2)\}} = \sum_{j_1 \in J_1, j_2 \in J_2} \mathbf{U}_{\{(i,j_1),(i,j_2)\}} = \sum_{j \in J_1 \cap J_2} \mathbf{U}_{\{(i,j)\}} = \overline{\mathbf{U}}_{\{(i,J_1 \cap J_2)\}} \tag{2}$$

By inducting this argument, it is always possible to assume without loss of generality that for a set $\mathcal{S} = \{(i_1, J_1), \ldots, (i_r, J_r)\}$, the elements $i_1, \ldots, i_r$ are all distinct. We need the above fact and a little more notation to verify the required condition. For $\mathcal{S} = \{(i_1, J_1), \ldots, (i_r, J_r)\}$, let

$$\mathcal{F}(\mathcal{S}) = \{(i_1, j_1), \ldots, (i_r, j_r) \mid j_1 \in J_1, \ldots, j_r \in J_r\}$$

with this notation,

$$\left\langle \overline{\mathbf{U}}_{\mathcal{S}_1}, \overline{\mathbf{U}}_{\mathcal{S}_2} \right\rangle = \sum_{\substack{T_1 \in \mathcal{F}(\mathcal{S}_1) \\ T_2 \in \mathcal{F}(\mathcal{S}_2)}} \left\langle \mathbf{U}_{T_1}, \mathbf{U}_{T_2} \right\rangle$$

Also, $\left\langle \mathbf{U}_{T_1}, \mathbf{U}_{T_2} \right\rangle \neq 0$ if and only if $\forall (i, j_1) \in T_1, (i, j_2) \in T_2$, we have $j_1 = j_2$ (again using orthogonality of vectors for vertices in a single clique in G). However, this means that $T_1 \cup T_2 \in \mathcal{F}(\mathcal{S}_1 \cup \mathcal{S}_2)$. Also, since all is in $\mathcal{S}_1, \mathcal{S}_2$ are distinct as observed using (2), every element $T \in \mathcal{F}(\mathcal{S}_1 \cup \mathcal{S}_2)$ corresponds to a unique pair $T_1 \in \mathcal{F}(\mathcal{S}_1), T_2 \in \mathcal{F}(\mathcal{S}_2)$. This gives

$$\left\langle \overline{\mathbf{U}}_{\mathcal{S}_1}, \overline{\mathbf{U}}_{\mathcal{S}_2} \right\rangle = \sum_{\substack{T_1 \in \mathcal{F}(\mathcal{S}_1) \\ T_2 \in \mathcal{F}(\mathcal{S}_2)}} \left\langle \mathbf{U}_{T_1}, \mathbf{U}_{T_2} \right\rangle = \sum_{T \in \mathcal{F}(\mathcal{S}_1 \cup \mathcal{S}_2)} \left\langle \mathbf{U}_T, \mathbf{U}_{\emptyset} \right\rangle = \left\langle \overline{\mathbf{U}}_{\mathcal{S}_1 \cup \mathcal{S}_2}, \overline{\mathbf{U}}_{\emptyset} \right\rangle$$

- From the above condition, it follows that for all $\mathcal{S}$, $\left\| \overline{\mathbf{U}}_{\mathcal{S}} \right\|^2 = \left\langle \overline{\mathbf{U}}_{\mathcal{S}}, \overline{\mathbf{U}}_{\emptyset} \right\rangle \leq \left\| \overline{\mathbf{U}}_{\mathcal{S}} \right\|$, using $\left\| \overline{\mathbf{U}}_{\emptyset} \right\| = 1$. Since the length of all vectors is at most 1, and all inner products are positive in G, all inner products for the vector solution above are between 0 and 1.

To verify the SDP value, we simply need to use the fact that all vectors within a single clique in G are orthogonal. Hence the SDP value is equal to

$$\frac{1}{m_1} \sum_{i \in M_1, J \subseteq L_1} p^{|J|}(1-p)^{|L_1 \setminus J|} \left\| \overline{\mathbf{U}}_{\{(i,J)\}} \right\|^2 \;=\; \frac{1}{m_1} \sum_{i \in M_1, J \subseteq L_1} p^{|J|}(1-p)^{|L_1 \setminus J|} \sum_{j \in J} \left\| \mathbf{U}_{\{(i,j)\}} \right\|^2$$

$$=\; \frac{1}{m_1} \sum_{i \in M_1, j \in L_1} p \cdot \left\| \mathbf{U}_{\{(i,j)\}} \right\|^2 \;=\; \frac{p}{m_1} \cdot \mathsf{FRAC}(G)$$

$\blacksquare$

7.4 Putting things together

For an (m_0, l_0) co-partite graph G, let $DS(G, \epsilon_1, d_\mathsf{T})$ denote the graph obtained by staring from G and performing the block assignments step and long-code step, where the reduction has parameters d_T and ϵ_1 for the block assignments step. Let $p_{max} = \frac{3-\sqrt{5}}{2}$. The soundness analysis of the Dinur-Safra reduction can be summarized in the following theorem (stated in a way adapted to our application).

Theorem 7.7 ([10]) *For given $\epsilon_1 > 0, p \in (1/3, p_{max})$, there exist constants $\epsilon_0, h, d'_\mathsf{T}$ such that if G is an m_0, l_0 co-partite graph such that every set of $\epsilon_0 \cdot m_0$ vertices in G contains an h-clique, then the weight of the maximum independent set in $DS(G, \epsilon_1, d_\mathsf{T})$ for any $d_\mathsf{T} \geq d'_\mathsf{T}$ is at most $4p^3 - 3p^4 + \epsilon_1$.*

Using the above theorem and the previous discussion, we can now prove an integrality gap for Minimum Vertex Cover.

Theorem 7.8 *For any given $\epsilon > 0$, there exists $\delta = \delta(\epsilon) > 0$ and an infinite family of graphs such that for graphs in the family, with N vertices, the integrality gap for the SDP relaxation for* Minimum Vertex Cover *obtained by $\Omega(N^\delta)$ rounds of the Lasserre hierarchy, remains at least $1.3606 - \epsilon$.*

Proof: Let $p \in (1/3, p_{max})$ be such that $\frac{1 - 4p^3 + 3p^4}{1-p} = 1.3606$ and $\epsilon_1 = \epsilon/10$. Let ϵ_0, h be as given by Theorem 7.7 and let $d_\mathsf{T} = \max(d'_\mathsf{T}, 2/\epsilon_1)$. For large enough n, let Φ be an instance of a constraint satisfaction problem as given by Corollary 4.8 (by using $\epsilon = 1/2$ to invoke the corollary) for $k = 3$. This is simply an instance of MAX 3-XOR on n variables with $m = O(n)$ constraints in which $s < 2/3$ fraction of the constraints are satisfiable and $\mathsf{FRAC}(\Phi) = m$ even after $\Omega(n)$ rounds.

By Theorem 7.3, there exists an $r = O(\log(h/\epsilon_0))$, such that G_Φ^r, which is an (m_0, l_0) co-partite graph for $m_0 = m^r$ and $l_0 = 4^r$, has no h-clique-free subset with $\epsilon_0 m_0$ vertices. Then the weight of the maximum independent set in $DS(G_\Phi^r, \epsilon_1, d_\mathsf{T})$ is at most $4p^3 - 3p^4 + \epsilon_1$, and hence the weight of the minimum vertex cover is at least $1 - 4p^3 + 3p^4 - \epsilon_1$.

On the other hand, by lemmata 5.1, 7.4 and 7.6, $\mathsf{FRAC}(DS(G_\Phi^r, \epsilon_1, d_\mathsf{T})) \geq p(1 - \epsilon_1)$ for the independent set SDP obtained by $\Omega(n/l_0 d_\mathsf{T})$ rounds. Hence the gap for Minimum Vertex Cover is at least $\frac{1 - 4p^3 + 3p^4 - \epsilon_1}{p - \epsilon_1} \geq 1.3606 - \epsilon$. It remains to express the number of rounds in terms of the number of vertices in $DS(G_\Phi^r, \epsilon_1, d_\mathsf{T})$. However, note that at all parameters in the reduction are constants and the size of the graph grows by a polynomial factor at each step of the reduction. Hence, the number of rounds equals $\Omega(n) = \Omega(N^\delta)$ for constant δ depending on p and ϵ, where N denotes $|DS(G_\Phi^r, \epsilon_1, d_\mathsf{T})|$. ∎

Acknowledgements

References

[1] Michael Alekhnovich, Sanjeev Arora, and Iannis Tourlakis. Towards strong nonapproximability results in the Lovasz-Schrijver hierarchy. In *Proceedings of the 37th ACM Symposium on Theory of Computing*, pages 294–303, 2005.

[2] Per Austrin and Elchanan Mossel. Approximation resistant predicates from pairwise independence. In *Proceedings of the 23rd IEEE Conference on Computational Complexity*, pages 249–258, Los Alamitos, CA, USA, 2008. IEEE Computer Society.

[3] M. Bellare, O. Goldreich, and M. Sudan. Free bits, PCP's and non-approximability – towards tight results. *SIAM Journal on Computing*, 27(3):804–915, 1998. Preliminary version in *Proc. of FOCS'95*.

[4] Eli Ben-Sasson and Avi Wigderson. Short proofs are narrow: Resolution made simple. *Journal of the ACM*, 48(2), 2001.

[5] R. Boppana and M.M. Halldórsson. Approximating maximum independent set by excluding subgraphs. *Bit*, 32:180–196, 1992.

[6] Josh Buresh-Oppenheim, Nicola Galesi, Shlomo Hoory, Avner Magen, and Toniann Pitassi. Rank bounds and integrality gaps for cutting planes procedures. In *Proceedings of the 44th IEEE Symposium on Foundations of Computer Science*, pages 318–327, 2003.

[7] Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Near-optimal algorithms for maximum constraint satisfaction problems. In *Proceedings of the 18th ACM-SIAM Symposium on Discrete Algorithms*, pages 62–68, 2007.

[8] Eden Chlamtac. Approximation algorithms using hierarchies of semidefinite programming relaxations. In *FOCS*, pages 691–701, 2007.

[9] Eden Chlamtac and Gyanit Singh. Improved approximation guarantees through higher levels of sdp hierarchies. In *APPROX-RANDOM*, pages 49–62, 2008.

[10] Irit Dinur and Shmuel Safra. On the hardness of approximating minimum vertex-cover. *Annals of Mathematics*, 162(1):439–486, 2005.

[11] Lars Engebretsen and Jonas Holmerin. More efficient queries in PCPs for NP and improved approximation hardness of maximum CSP. In *Proceedings of the 22th Symposium on Theoretical Aspects of Computer Science*, pages 194–205, 2005.

[12] U. Feige, S. Goldwasser, L. Lovász, S. Safra, and M. Szegedy. Approximating clique is almost NP-complete. In *Proceedings of the 32nd IEEE Symposium on Foundations of Computer Science*, pages 2–12, 1991.

[13] Uriel Feige. Randomized graph products, chromatic numbers, and the lovász vartheta-funktion. *Combinatorica*, 17(1):79–90, 1997.

[14] Uriel Feige and Joe Kilian. Zero knowledge and the chromatic number. *Journal of Computer and System Sciences*, 57(2):187–199, 1998.

[15] Uriel Feige, Michael Langberg, and Gideon Schechtman. Graphs with tiny vector chromatic numbers and huge chromatic numbers. *SIAM J. Comput.*, 33(6):1338–1368, 2004.

[16] Konstantinos Georgiou, Avner Magen, Toniann Pitassi, and Iannis Tourlakis. Integrality gaps of 2 - o(1) for vertex cover SDPs in the Lovász-Schrijver hierarchy. In *Proceedings of the 48th IEEE Symposium on Foundations of Computer Science*, pages 702–712, 2007.

[17] Johan Håstad. On the approximation resistance of a random predicate. In *APPROX-RANDOM*, pages 149–163, 2007.

[18] Subhash Khot. Improved inaproximability results for maxclique, chromatic number and approximate graph coloring. In *FOCS*, pages 600–609, 2001.

[19] Subhash Khot and Ashok Kumar Ponnuswami. Better inapproximability results for maxclique, chromatic number and min-3lin-deletion. In *ICALP (1)*, pages 226–237, 2006.

[20] Subhash Khot and Oded Regev. Vertex cover might be hard to approximate to within $2 - \epsilon$. In *Proceedings of the 18th IEEE Conference on Computational Complexity*, 2003.

[21] Jean B. Lasserre. An explicit exact sdp relaxation for nonlinear 0-1 programs. In *ipco01*, pages 293–303, London, UK, 2001. Springer-Verlag.

[22] M. Laurent, Mathematisch Centrum (smc, The Dutch Foundation, and Monique Laurent. A comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre relaxations for 0-1 programming. *Mathematics of Operations Research*, 28:470–496, 2003.

[23] L. Lovász and A. Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM J. on Optimization*, 1(12):166–190, 1991.

[24] Alex Samorodnitsky and Luca Trevisan. Gowers uniformity, influence of variables, and PCPs. In *Proceedings of the 38th ACM Symposium on Theory of Computing*, pages 11–20, 2006.

[25] Grant Schoenebeck. Linear level Lasserre lower bounds for certain k-csps. In *Proceedings of the 49th IEEE Symposium on Foundations of Computer Science*, 2008.

[26] Grant Schoenebeck, Luca Trevisan, and Madhur Tulsiani. A linear round lower bound for Lovász-Schrijver SDP relaxations of vertex cover. In *Proceedings of the 22nd IEEE Conference on Computational Complexity*, 2007.

A Proofs from Section 2

Lemma A.1 *Let $\epsilon, \delta > 0$ and a predicate $P : [q]^k \to \{0, 1\}$ be given. Then there exist $\beta = O(q^k/\epsilon^2)$, $\eta = \Omega((1/\beta)^{5/\delta})$ and $N \in \mathbb{N}$, such that if $n \geq N$ and Φ is a random instance of* MAX k-CSP(P) *with $m = \beta n$ constraints, then with probability $1 - o(1)$*

1. OPT$(\Phi) \leq \frac{|P^{-1}(1)|}{q^k}(1 + \epsilon) \cdot m$.

2. For all $s \leq \eta n$, every set of s constraints involves at least $(k - 1 - \delta)s$ variables.

Proof: Let $\alpha \in [q]^n$ be any fixed assignment. For a fixed α, the events that a constraint C_i is satisfied are independent and happen with probability $|P^{-1}(1)|/q^k$ each. Hence, the probability over the choice of Φ that α satisfies more than $\frac{|P^{-1}(1)|}{q^k}(1 + \epsilon) \cdot \beta n$ constraints is at most

$\exp(-\epsilon^2\beta n|P^{-1}(1)|/3q^k)$ by Chernoff bounds. By a union bound, the probability that *any* assignment satisfies more than $\frac{|P^{-1}(1)|}{q^k}(1+\epsilon)\cdot\beta n$ constraints is at most $q^n\cdot\exp\left(-\frac{\epsilon^2\beta n|P^{-1}(1)|}{3q^k}\right)=\exp\left(n\ln q-\frac{\epsilon^2\beta n|P^{-1}(1)|}{3q^k}\right)$ which is $o(1)$ for $\beta=\frac{6q^k\ln q}{\epsilon^2}$.

For showing the next property, we consider the probability that a set of s constraints contains at most cs variables, where $c=k-1-\delta$. This is upper bounded by

$$\binom{n}{cs}\cdot\binom{\binom{cs}{k}}{s}\cdot s!\binom{\beta n}{s}\cdot\binom{n}{k}^{-s}$$

Here $\binom{n}{cs}$ is the number of ways of choosing the cs variables involved, $\binom{\binom{cs}{k}}{s}$ is the number of ways of picking s tuples out of all possible k-tuples on cs variables and $s!\binom{\beta n}{s}$ is the number of ways of selecting the s constraints. The number $\binom{n}{k}^s$ is simply the number of ways of picking s of these k-tuples in an unconstrained way. Using $(\frac{a}{b})^b\le\binom{a}{b}\le(\frac{a\cdot e}{b})^b$, $s!\le s^s$ and collecting terms, we can bound this expression by

$$\left(\frac{s}{n}\right)^{\delta s}\left(e^{2k+1-\delta}k^{1+\delta}\beta\right)^s\le\left(\frac{s}{n}\right)^{\delta s}(\beta^5)^s=\left(\frac{s\beta^{5/\delta}}{n}\right)^{\delta s}$$

We need to show that the probability that a set of s constraints contains less than cs variables for *any* $s\le\eta n$ is $o(1)$. Thus, we sum this probability over all $s\le\eta n$ to get

$$\sum_{s=1}^{\eta n}\left(\frac{s\beta^{5/\delta}}{n}\right)^{\delta s}=\sum_{s=1}^{\ln^2 n}\left(\frac{s\beta^{5/\delta}}{n}\right)^{\delta s}+\sum_{s=\ln^2 n+1}^{\eta n}\left(\frac{s\beta^{5/\delta}}{n}\right)^{\delta s}$$
$$\le O\left(\frac{\beta^5}{n^\delta}\ln^2 n\right)+O\left(\left(\eta\cdot\beta^{5/\delta}\right)^{\delta\ln^2 n}\right)$$

The first term is $o(1)$ and is small for large n. The second term is also $o(1)$ for $\eta=1/(100\beta^{5/\delta})$. ∎

Fact A.2 *Let* $(q^{r-1}-1)/(q-1)<k\le(q^r-1)/(q-1)$*. Then there exists a linear code of length* k *and distance* 3 *over* $\mathbb{F}_q$*, with dimension* $k-r$*.*

Proof: Let $l=(q^r-1)/(q-1)$. We can first construct a code of length equal to l and dimension $l-r$ by specifying the $r\times l$ check matrix (a matrix whose rows span the space orthogonal to the code). The requirement that the code must have distance 3 means that no two columns of the check matrix should be linearly dependent. We can choose (for example) all non-zero vectors in $(\mathbb{F}_q)^r$ having their first nonzero element as 1 to get a matrix with r rows having this property. It is easy to check that there are $l=(q^r-1)/(q-1)$ such columns. To reduce the code length, we delete the last $l-k$ columns of the matrix to get the check matrix of a code with length k, distance 3 and dimension $k-r$. ∎

B   Proof of Theorem 4.2

Recall that for a system Λ of linear equations, $\mathsf{Res}(\Lambda,t)$ was defined as the system of equations, each having at most t variables and obtained by combining at most 2 equations in Λ. We shall also require some additional notation for the proof. For a linear equation $\omega\cdot x=r$, we denote by

$Supp(\omega)$ the set of non-zero coordinates in ω and by Ω_S the set of all coefficient vectors ω such that $Supp(\omega) \subseteq S$. Λ_S denotes all equations $(\omega \cdot x = r) \in \Lambda$ such that $\omega \in \Omega_S$.

Suppose $\exists r_1 \neq r_2$ such that $(\omega \cdot x = r_1) \in \Lambda$ and $(\omega \cdot x = r_2) \in \Lambda$ with $|Supp(w)| < t$. Then, $(\overrightarrow{0} \cdot x = r_1 - r_2) \in \mathsf{Res}(\Lambda, 2t)$. Conversely, if we know that $(\overrightarrow{0} \cdot x = r) \notin \mathsf{Res}(\Lambda, 2t)$ for any $r \neq 0$, then we can assume that for any $\omega \in \Omega$, there is at most one value of r such that $(\omega, r) \in \Lambda$. We will abuse notation to write $\omega \in \Lambda$ when some such r exists and is guaranteed to be unique. Also, for ω such that $(\omega \cdot x = r) \in \Lambda$ we define the function $\lambda(\omega)$, which specifies what the value of $\omega \cdot x$ should be for any satisfying assignment (or partial assignment) according to Λ. We take $\lambda(\omega) = r$ if there exists a unique r such that $(\omega \cdot x = r) \in \Lambda$ and undefined otherwise.

Theorem B.1 *Let q be a prime. Suppose Λ is a system of linear equations in $\mathbb{F}_q$ such that $(\overrightarrow{0} \cdot x = r) \in \Lambda \Leftrightarrow r = 0$ and $\mathsf{Res}(\Lambda, 2t) = \Lambda$. Then there are vectors $\mathbf{V}_{(S,\alpha)}$, for all S with $|S| < t$ and for all $\alpha \in [q]^S$, such that*

1. *$\left\langle \mathbf{V}_{(S_1,\alpha_1)}, \mathbf{V}_{(S_2,\alpha_2)} \right\rangle \geq 0$ for all $S_1, S_2, \alpha_1, \alpha_2$.*

2. *$\left\langle \mathbf{V}_{(S_1,\alpha_1)}, \mathbf{V}_{(S_2,\alpha_2)} \right\rangle = 0$ if $\alpha_1(S_1 \cap S_2) \neq \alpha_2(S_1 \cap S_2)$.*

3. *If $\alpha_i \in [q]^{S_i}$, $1 \leq i \leq 4$ are such that $\alpha_1 \circ \alpha_2$ and $\alpha_3 \circ \alpha_4$ are both defined and equal, then*

$$\left\langle \mathbf{V}_{(S_1,\alpha_1)}, \mathbf{V}_{(S_2,\alpha_2)} \right\rangle = \left\langle \mathbf{V}_{(S_3,\alpha_3)}, \mathbf{V}_{(S_4,\alpha_4)} \right\rangle$$

4. *$\mathbf{V}_{(S,\alpha)} = 0$ for $\alpha \notin A_S$ and $\sum_{\alpha \in A_S} \left\| \mathbf{V}_{(S,\alpha)} \right\|^2 = 1$*

Proof Idea: The idea of the proof is to "encode" the partial assignments in the vectors in such a way so that it is easy to enforce consistency according to the given system of constraints. Since the constraints are in the form of linear equations, it is easiest to specify the value of all the linear forms $\omega \cdot x$.

In the binary case, one can think that in the vector $\mathbf{V}_{(S,\alpha)}$ we have a coordinate for each ω in which we specify $\mathbb{E}(-1)^{\omega \cdot x}$, where the expectation is over all x which agree with the assignment α. This specifies all the Fourier coefficients of the function which is 1 iff x is consistent with α, and hence "encodes" α. When $Supp(\omega) \not\subseteq S$, this expectation vanishes and hence one only needs to bother about ω in the set $\Omega = \cup_{|S|<t}\Omega_S$. Furthermore, because of the nature of dependencies by linear constraints, the value of $(\omega_1 - \omega_2) \cdot x$ is either completely determined by Λ (when $(\omega_1 - \omega_2) \in \Lambda$) or is completely undetermined and hence $\omega_1 \cdot x$ and $\omega_2 \cdot x$ are independent of each other. We thus partition Ω into various equivalence classes based on the set of equations Λ such that all linear forms within a class are completely determined by fixing the value of any one of them. The vectors we construct will have one coordinate corresponding to each of these classes which will enforce the dependencies due to Λ automatically.

Finally, to generalize this to q-ary equations, the natural analogue would be to consider the powers of the roots of unity i.e. expressions of the form $\exp(\frac{2\pi i(\omega \cdot x)}{q})$. However, this is not an option, since the coordinates of the vectors are not allowed to be complex. It is easy to check though that the proof in the binary case requires only one property of the coordinates that if two vectors disagree on the value in one coordinate (i.e. the product is -1), then the disagreements and agreements are balanced over all the coordinates and hence the inner product of the vectors is zero.

In the q-ary case, it can be proved that if the difference in the value of $\omega \cdot x$ in some coordinate according to two vectors is $\Delta \in [q], \Delta \neq 0$, then over all the coordinates, all values of Δ (including

0) occur equally often. We then choose each "coordinate" to be a small $q - 1$ dimensional vector such that the product of the vectors at a coordinate is $-1/(q-1)$ if $\Delta \neq 0$ and 1 if $\Delta = 0$. This, combined with the balance property, still gives the orthogonality of the vectors which correspond to inconsistent assignments, and suffices for our purposes. The details are given in the proof below.

Proof of Theorem B.1: Let $\Omega = \cup_{|S|<t}\Omega_S$. For $\omega_1, \omega_2 \in \Omega$, we say $\omega_1 \sim \omega_2$ if $(\omega_1 - \omega_2) \in \Lambda$. Since this is an equivalence relation, this partitions Ω into equivalence classes $\mathcal{C}_1, \ldots, \mathcal{C}_N$. We write $\mathcal{C}(\omega)$ to denote the class containing ω. Next, we choose a representative (say the lexicographically first element) for each class. We use $[\mathcal{C}]$ to denote the representative for the class $\mathcal{C}$.

For constructing the vector $\mathbf{V}_{(S,\alpha)}$, we assign it a $q - 1$ dimensional "coordinate" corresponding to each equivalence class of the above relation i.e. for each class we choose a $q - 1$ dimensional vector and the final vector $\mathbf{V}_{(S,\alpha)}$ is the direct sum of all these vectors. Let $e_0, e_1, \ldots, e_{q-1}$ denote the q maximally separated unit vectors in $q - 1$ dimensions such that $\langle e_i, e_j \rangle = -\frac{1}{q-1}$ if $i \neq j$ and 1 if $i = j$. Using $\mathbf{V}_{(S,\alpha)}(\mathcal{C})$ to denote the coordinate corresponsing to $\mathcal{C}$, we define $\mathbf{V}_{(S,\alpha)}$ as:

$$\mathbf{V}_{(S,\alpha)}(\mathcal{C}) = \begin{cases} 0 & \text{if } \alpha \text{ disagrees with some equation in } \Lambda_S \\ 0 & \text{if } \mathcal{C} \cap \Omega_S = \emptyset \\ \frac{1}{|A_S|} \cdot e_{\omega \cdot \alpha + \lambda([\mathcal{C}(\omega)]-\omega)} & \text{for any } \omega \in \mathcal{C} \cap \Omega_S \end{cases}$$

Here $\omega \cdot \alpha$ is defined as $\sum_{i \in S} \omega_i \alpha(i)$ is the inner product of ω with the partial assignment α, which can be computed since $Supp(\omega) \subseteq S$. The expression $\omega \cdot \alpha + \lambda([\mathcal{C}(\omega)] - \omega)$ is computed modulo q. To show that the vector is well defined, we first need to argue that the coordinate $\mathbf{V}_{(S,\alpha)}(\mathcal{C})$ does not depend on which ω we choose from $\mathcal{C} \cap \Omega_S$.

Claim B.2 *If $\alpha \in [q]^S$ satisfies all equations in Λ_S, then for any class $\mathcal{C}$ and $\omega_1, \omega_2 \in \mathcal{C} \cap \Omega_S$*

$$\omega_1 \cdot \alpha + \lambda([\mathcal{C}] - \omega_1) = \omega_2 \cdot \alpha + \lambda([\mathcal{C}] - \omega_2)$$

Proof: Since $([\mathcal{C}] - \omega_1), ([\mathcal{C}] - \omega_2), (\omega_1 - \omega_2) \in \Lambda$ and each have support size at most $2t$, it must be the case that $\lambda([\mathcal{C}] - \omega_2) - \lambda([\mathcal{C}] - \omega_1) = \lambda(\omega_1 - \omega_2)$ otherwise we can derive $\overrightarrow{0} \cdot x = r$ for $r \neq 0$. Also, since α is consistent with Λ_S, it must satisfy $(\omega_1 - \omega_2) \cdot \alpha = \lambda(\omega_1 - \omega_2)$. The claim follows. ∎

The next claim shows that the only way two vectors $\mathbf{V}_{(S_1,\alpha_1)}$ and $\mathbf{V}_{(S_2,\alpha_2)}$ can have non-zero inner product is by having $\langle \mathbf{V}_{(S_1,\alpha_1)}(\mathcal{C}), \mathbf{V}_{(S_2,\alpha_2)}(\mathcal{C}) \rangle = \frac{1}{|A_{S_1}||A_{S_2}|}$ for each coordinate $\mathcal{C}$ in which it is non-zero. The theorem essentially follows from this claim using a simple counting argument.

Claim B.3 *Let $\alpha_1 \in [q]^{S_1}$ and $\alpha_2 \in [q]^{S_2}$ be two partial assignments. If $\langle \mathbf{V}_{(S_1,\alpha_1)}(\mathcal{C}), \mathbf{V}_{(S_2,\alpha_2)}(\mathcal{C}) \rangle < 0$ for some class $\mathcal{C}$, then $\langle \mathbf{V}_{(S_1,\alpha_1)}, \mathbf{V}_{(S_2,\alpha_2)} \rangle = 0$.*

Proof: For any class $\mathcal{C}'$ with $\langle \mathbf{V}_{(S_1,\alpha_1)}(\mathcal{C}'), \mathbf{V}_{(S_2,\alpha_2)}(\mathcal{C}') \rangle \neq 0$, we have some $\omega_1' \in \mathcal{C}' \cap \Omega_{S_1}$ and $\omega_2' \in \mathcal{C}' \cap \Omega_{S_2}$. We will work with the quantity $\lambda([\mathcal{C}(\omega_1')] - \omega_1') + \omega_1' \cdot \alpha_1 - \lambda([\mathcal{C}(\omega_2')] - \omega_2') - \omega_2' \cdot \alpha_2$ denoted by $\Delta(\mathcal{C}')$., which is the difference of the indices of $\mathbf{V}_{(S_1,\alpha_1)}(\mathcal{C}')$ and $\mathbf{V}_{(S_2,\alpha_2)}(\mathcal{C}')$. It is easy to see that

$$\langle \mathbf{V}_{(S_1,\alpha_1)}(\mathcal{C}'), \mathbf{V}_{(S_2,\alpha_2)}(\mathcal{C}') \rangle = \begin{cases} \frac{1}{|A_{S_1}||A_{S_2}|} & \text{if } \Delta(\mathcal{C}') = 0 \\ -\frac{1}{(q-1)|A_{S_1}||A_{S_2}|} & \text{otherwise} \end{cases}$$

33

For any r_1, r_2 we will give an injective map which maps a class $\mathcal{C}_{i_1}$ having $\Delta(\mathcal{C}_{i_1}) = r_1$ to a class $\mathcal{C}_{i_2}$ having $\Delta(\mathcal{C}_{i_2}) = r_2$. Hence over all the classes, all values of $\Delta(\mathcal{C}')$ must occur equally often. This would imply the claim, since if $\langle \mathbf{V}_{(S_1, \alpha_1)}(\mathcal{C}'), \mathbf{V}_{(S_2, \alpha_2)}(\mathcal{C}') \rangle \neq 0$ for N_0 classes $\mathcal{C}'$, then

$$\langle \mathbf{V}_{(S_1, \alpha_1)}, \mathbf{V}_{(S_2, \alpha_2)} \rangle = \frac{N_0}{q} \cdot \frac{1}{|A_{S_1}||A_{S_2}|} + \frac{N_0(q-1)}{q} \cdot \left(-\frac{1}{(q-1)|A_{S_1}||A_{S_2}|} \right) = 0$$

We now construct the above map using the class $\mathcal{C}$. Let $\omega_1 \in \mathcal{C} \cap \Omega_{S_1}$ and $\omega_2 \in \mathcal{C} \cap \Omega_{S_2}$. If $\langle \mathbf{V}_{(S_1, \alpha_1)}(\mathcal{C}), \mathbf{V}_{(S_2, \alpha_2)}(\mathcal{C}) \rangle < 0$, then we have that

$$\lambda (\omega_1 - \omega_2) + \omega_1 \cdot \alpha_1 - \omega_2 \cdot \alpha_2 = \Delta(\mathcal{C}) \neq 0$$

Here we used the fact that

$$\lambda ([\mathcal{C}] - \omega_1) - \lambda ([\mathcal{C}] - \omega_2) = \lambda(\omega_2 - \omega_1)$$

Let $\mathcal{C}'$ be any class such that $\langle \mathbf{V}_{(S_1, \alpha_1)}(\mathcal{C}'), \mathbf{V}_{(S_2, \alpha_2)}(\mathcal{C}') \rangle \neq 0$ Let $\omega_1' \in \mathcal{C} \cap \Omega_{S_1}$ and $\omega_2' \in \mathcal{C}' \cap \Omega_{S_2}$. Then

$$\lambda (\omega_1' - \omega_2') + \omega_1' \cdot \alpha_1 - \omega_2' \cdot \alpha_2 = \Delta(\mathcal{C}')$$

where $\Delta(\mathcal{C}')$ may now also be 0. Thus, for all $\mu \in [q]$, we get that

$$\mu \lambda (\omega_2 - \omega_1) + \lambda (\omega_2' - \omega_1') + (\mu \omega_1 + \omega_1') \cdot a_1 - (\mu \omega_2 + \omega_2') \cdot a_2 \quad = \quad \mu \Delta(\mathcal{C}) + \Delta(\mathcal{C}') \qquad (3)$$

Since $(\omega_2 - \omega_1) \in \Lambda$ and $(\omega_2' - \omega_1') \in \Lambda$ and each involves at most t variables, we also have $(\mu \omega_1 + \omega_1') - (\mu \omega_2 + \omega_2') \in \Lambda$. Hence, $(\mu \omega_1 + \omega_1')$ and $(\mu \omega_2 + \omega_2')$ must be in the same class, say $\mathcal{C}''$, and we can write

$$\mu \lambda (\omega_2 - \omega_1) + \lambda (\omega_2' - \omega_1') \quad = \quad \lambda ((\mu \omega_2 + \omega_2') - (\mu \omega_1 + \omega_1'))$$
$$= \quad \lambda ([\mathcal{C}''] - (\mu \omega_1 + \omega_1')) - \lambda ([\mathcal{C}''] - (\mu \omega_2 + \omega_2'))$$

Combining this with equation 3, we get that

$$\Delta(\mathcal{C}'') = \mu \Delta(\mathcal{C}) + \Delta(\mathcal{C}')$$

Since $\Delta(\mathcal{C}) \neq 0$, for any r_1, r_2, choosing $\mu = (r_2 - r_1)/\Delta(\mathcal{C})$ gives a mapping in which the image of a class $\mathcal{C}'$ with $\Delta(\mathcal{C}') = r_1$ is the class $\mathcal{C}(\mathcal{C}'')$ with $\Delta(\mathcal{C}'') = r_2$. It is also easy to check that this mapping is injective and hence the claim follows. ∎

From the above claim, we get property (1) in the theorem, since if two vectors have non-zero inner product, it must be positive *in every coordinate*. From the above claim it also follows that the inner product is only non-zero for vectors corresponding to partial assignments which are "mutually consistent" in the sense described below. The following also proves property (2) as stated in the theorem.

Claim B.4 *If $\langle \mathbf{V}_{(S_1, \alpha_1)}, \mathbf{V}_{(S_2, \alpha_2)} \rangle \neq 0$, then α_1 and α_2 agree on all the variables in $S_1 \cap S_2$. Moreover, the assignment over $S_1 \cup S_2$ defined by α_1 and α_2 satisfies all the equations in $\Lambda_{S_1 \cup S_2}$.*

Proof: By Claim B.3, it suffices to show that when α_1 and α_2 disagree on some variable in $S_1 \cap S_2$ or when $\alpha_1 \circ \alpha_2$ violates some equation in $\Lambda_{S_1 \cup S_2}$, then there exists a class $\mathcal{C}$ such that that $\langle \mathbf{V}_{(S_1,\alpha_1)}(\mathcal{C}), \mathbf{V}_{(S_2,\alpha_2)}(\mathcal{C}) \rangle < 0$. Consider the case when for $i \in S_1 \cap S_2$, $\alpha_1(i) \neq \alpha_2(i)$. Let ω be the vector which has coefficient 1 corresponding to x_i and all has other coefficients as zero. Then $\lambda([\mathcal{C}(\omega)] - \omega) + \omega \cdot \alpha_1 - \lambda([\mathcal{C}(\omega)] - \omega) - \omega \cdot \alpha_2 = \alpha_1(i) - \alpha_2(i) \neq 0$, which implies that $\langle \mathbf{V}_{(S_1,\alpha_1)}(\mathcal{C}(\omega)), \mathbf{V}_{(S_2,\alpha_2)}(\mathcal{C}(\omega)) \rangle = \frac{1}{|A_{S_1}||A_{S_2}|} \langle e_{\omega \cdot \alpha_1 + \lambda([\mathcal{C}(\omega)] - \omega)}, e_{\omega \cdot \alpha_2 + \lambda([\mathcal{C}(\omega)] - \omega)} \rangle < 0$.

Next, suppose that α_1 and α_2 agree on $S_1 \cap S_2$, but $\alpha_1 \circ \alpha_2$ violates some equation $(\omega \cdot x = r) \in \Lambda_{S_1 \cup S_2}$ i.e. $(\alpha_1 \circ \alpha_2) \cdot \omega \neq r$. Let ω_1 be the vector which is the same as ω for all coordinates in S_1 and is zero otherwise. It is clear that $\omega_1 \in \Omega_{S_1}$ and $(\omega_1 - \omega) \in \Omega_{S_2}$. Also $\omega_1 \sim (\omega_1 - \omega)$ as their difference is ω which is in Λ. Let both of them be in the class $\mathcal{C}$ and consider $\lambda([\mathcal{C}] - \omega_1) + \alpha_1 \cdot \omega_1 - \lambda([\mathcal{C}] - (\omega_1 - \omega)) - \alpha_2 \cdot (\omega_1 - \omega)$. This is equal to $(\alpha_1 \circ \alpha_2) \cdot \omega - \lambda(\omega)$ which is non-zero by assumption. Hence, $\langle \mathbf{V}_{(S_1,\alpha_1)}(\mathcal{C}), \mathbf{V}_{(S_2,\alpha_2)}(\mathcal{C}) \rangle < 0$ and the claim follows. ■

Properties (2) and (3) will now follow from the claim below.

Claim B.5 *If $\alpha_1(S_1 \cap S_2) = \alpha_2(S_1 \cap S_2)$ and $\alpha_1 \circ \alpha_2$ satisfies all equation in $\Lambda_{S_1 \cup S_2}$, then*

$$\langle \mathbf{V}_{(S_1,\alpha_1)}, \mathbf{V}_{(S,\alpha_2)} \rangle = \frac{1}{|A_{S_1 \cup S_2}|}$$

Proof: Let $\mathcal{C}$ be any class such that $\langle \mathbf{V}_{(S_1,\alpha_1)}(\mathcal{C}), \mathbf{V}_{(S,\alpha_2)}(\mathcal{C}) \rangle \neq 0$. Then there exist some elements $\omega_1 \in \mathcal{C} \cap \Omega_{S_1}$ and $\omega_2 \in \mathcal{C} \cap \Omega_2$. Since $\alpha_1 \circ \alpha_2$ satisfies all equations in $\Lambda_{S_1 \cup S_2}$, we must have that $\lambda(\omega_1 - \omega_2) = (\alpha_1 \circ \alpha_2) \cdot (\omega_1 - \omega_2) = \alpha_1 \cdot \omega_1 - \alpha_2 \cdot \omega_2$. This gives that $\lambda([\mathcal{C}] - \omega_1) - \alpha_1 \cdot \omega_1 = \lambda([\mathcal{C}] - \omega_2) - \alpha_2 \cdot \omega_2$ and hence $\langle \mathbf{V}_{(S_1,\alpha_1)}(\mathcal{C}), \mathbf{V}_{(S,\alpha_2)}(\mathcal{C}) \rangle = \frac{1}{|A_{S_1}||A_{S_2}|}$. So to compute the inner product, we only need to know how many classes intersect both Ω_{S_1} and Ω_{S_2}.

Note that Λ_S (as a set of linear combinations without the RHS) is a subgroup of Ω_S under addition. For calculating the inner product $\langle \mathbf{V}_{(S_1,\alpha_1)}, \mathbf{V}_{(S_2,\alpha_2)} \rangle$, we consider the quotient group $\Omega_{S_1 \cup S_2}/\Lambda_{S_1 \cup S_2}$ with $\Omega_{S_1}/\Lambda_{S_1 \cup S_2}$ and $\Omega_{S_2}/\Lambda_{S_1 \cup S_2}$ being subgroups of it. $\Omega_{S_1}/\Lambda_{S_1}$ (which is the same as $\Omega_{S_1}/\Lambda_{S_1 \cup S_2}$) has one element for each class which has non-empty intersection with Ω_{S_1} (similarly for S_2). Hence, the number of classes intersecting both Ω_{S_1} and Ω_{S_2} is equal to $|\Omega_{S_1}/\Lambda_{S_1} \cap \Omega_{S_2}/\Lambda_{S_2}|$.

For A and B which are subgroups of a group G, we know that $|A + B| = |A||B|/|A \cap B|$. Using this gives

$$|\Omega_{S_1 \cup S_2}/\Lambda_{S_1 \cup S_2}| = |\Omega_{S_1}/\Lambda_{S_1 \cup S_2} + \Omega_{S_2}/\Lambda_{S_1 \cup S_2}| = \frac{|\Omega_{S_1}/\Lambda_{S_1 \cup S_2}||\Omega_{S_2}/\Lambda_{S_1 \cup S_2}|}{|\Omega_{S_1}/\Lambda_{S_1 \cup S_2} \cap \Omega_{S_2}/\Lambda_{S_1 \cup S_2}|}$$

Finally, noting that $\Omega_{S_1}/\Lambda_{S_1 \cup S_2}$ is the same as $\Omega_{S_1}/\Lambda_{S_1}$ and $|\Omega_{S_1}/\Lambda_{S_1}| = |A_{S_1}|$ (similarly for S_2), we get that

$$|\Omega_{S_1}/\Lambda_{S_1} \cap \Omega_{S_2}/\Lambda_{S_2}| = \frac{|A_{S_1}||A_{S_2}|}{|A_{S_1 \cup S_2}|}$$

Thus, we have

$$\langle \mathbf{V}_{(S_1,\alpha_1)}, \mathbf{V}_{(S_2,\alpha_2)} \rangle = |\Omega_{S_1}/\Lambda_{S_1} \cap \Omega_{S_2}/\Lambda_{S_2}| \cdot \frac{1}{|A_{S_1}||A_{S_2}|} = \frac{1}{|A_{S_1 \cup S_2}|}$$

which proves the claim. ■

To check property (3) we note that

$$\left\langle \mathbf{V}_{(S_1,\alpha_1)}, \mathbf{V}_{(S_2,\alpha_2)} \right\rangle = \left\langle \mathbf{V}_{(S_3,\alpha_3)}, \mathbf{V}_{(S_4,\alpha_4)} \right\rangle = 0$$

by Claim B.4 when $\alpha_1 \circ \alpha_2$ violates any constraint in $\Lambda_{S_1 \cup S_2}$. When it does not violate any constraints, Claim B.5 applies and we have

$$\left\langle \mathbf{V}_{(S_1,\alpha_1)}, \mathbf{V}_{(S_2,\alpha_2)} \right\rangle = \left\langle \mathbf{V}_{(S_3,\alpha_3)}, \mathbf{V}_{(S_4,\alpha_4)} \right\rangle = \frac{1}{|A_{S_1 \cup S_2}|}$$

From Claims B.4 and B.5, it is also immediate that $\left\| \mathbf{V}_{(S,\alpha)} \right\|^2 = 1/|A_S|$ if $\alpha \in A_S$ and is 0 otherwise. Hence, $\sum_{\alpha \in A_S} \left\| \mathbf{V}_{(S,\alpha)} \right\|^2 = 1$ which proves property (4). ∎

C Proof of Lemma 7.1

Lemma C.1 *Let the vectors $\mathbf{U}_S$ for $|S| \leq t$ form a solution to the t-round Lasserre SDP for* Maximum Independent Set *on a weighted graph $G = (V, E)$, with SDP value $\mathsf{FRAC}(G)$. Then there exist vectors $\mathbf{V}_{(S,\alpha)}$ for all $|S| \leq t/2, \alpha \in \{0,1\}^S$, determined by the vectors $\mathbf{U}_S$, such that*

1. *$\mathbf{U}_\emptyset = \mathbf{V}_{(\emptyset,\emptyset)}$ and $\mathbf{U}_S = \mathbf{V}_{(S,1_S)} \forall S$, where 1_S is the partial assignment which assigns 1 to all elements in S.*

2. *The vectors $\mathbf{V}_{(S,\alpha)}$ satisfy all conditions of the SDP for constraint satisfaction problems.*

3. *For any S, the vectors $\{\mathbf{V}_{(S,\alpha)} \mid \alpha \in \{0,1\}^S\}$ induce a probability distribution over $\{0,1\}^S$. The events measurable in this probability space correspond to all $\alpha' \in \{0,1\}^{S'}$ for all $S' \subseteq S$, and $\mathbb{P}[\alpha'] = \left\| \mathbf{V}_{(S',\alpha')} \right\|^2$.*

4. *The vectors $\mathbf{V}_{(S,0_S)}$ form a solution to the t-round SDP for* Minimum Vertex Cover *with objective value $\sum_{v \in V} w(v) - \mathsf{FRAC}(G)$, where $w(v)$ denotes the weight of vertex v and 0_S denotes the all-zero assignment.*

Proof: We shall define the vectors for all sets of size upto t. However, for convenience, we shall only prove the SDP conditions for vectors corresponding to sets of size at most $t/2$. For a pair (S, α) where $\alpha \in \{0,1\}^S$, we denote by $\alpha^{-1}(0)$ the set $\{i \in S \mid \alpha(i) = 0\}$ and by $\alpha^{-1}(1)$, the set $S \setminus \alpha^{-1}(0)$. We define the vectors $\mathbf{V}_{(S,\alpha)}$ using inclusion-exclusion, as

$$\mathbf{V}_{(S,\alpha)} = \sum_{T \subseteq \alpha^{-1}(0)} (-1)^{|T|} \mathbf{U}_{T \cup \alpha^{-1}(1)}$$

Also, we define $\mathbf{V}_{(\emptyset,\emptyset)} = \mathbf{U}_\emptyset$.

- Property (1) is then immediate from the definition of the vectors. We also note that for any $i \in \alpha^{-1}(0)$, we can write

$$\mathbf{V}_{(S,\alpha)} = \mathbf{V}_{(S \setminus \{i\}, \alpha(S \setminus \{i\}))} - \mathbf{V}_{(S,\alpha')}$$

where $\alpha' \in \{0,1\}^S$ is such that $\alpha'(j) = \alpha(j) \; \forall j \neq i$ and $\alpha'(i) = 1$.

36

- We now show that the vectors satisfy the SDP conditions. For all S_1, S_2 such that $|S_1 \cup S_2| \leq t$, we will show that

$$\langle \mathbf{V}_{(S_1,\alpha_1)}, \mathbf{V}_{(S_2,\alpha_2)} \rangle = \begin{cases} 0 & \text{when } \alpha_1(S_1 \cap S_2) \neq \alpha_2(S_1 \cap S_2) \\ \langle \mathbf{V}_{(S_1 \cup S_2, \alpha_1 \circ \alpha_2)}, \mathbf{V}_{(\emptyset,\emptyset)} \rangle & \text{otherwise} \end{cases}$$

It is easy to check that this will show that all SDP conditions are satisfied for sets of size at most $t/2$. We will proceed by induction on the total number of "zeroes" in the assignments α_1 and α_2 i.e. on $|\alpha_1^{-1}(0)| + |\alpha_2^{-1}(0)|$. The base case is when $\alpha_1 = 1_{S_1}$ and $\alpha_2 = 1_{S_2}$. Then, the product on the left is simply equal to $\langle \mathbf{U}_{S_1}, \mathbf{U}_{S_2} \rangle$, which is equal to $\|\mathbf{U}_{S_1 \cup S_2}\|^2$ since the vectors $\mathbf{U}_S$ form a valid solution to the independent set SDP.

For the induction step, first consider the case when α_1 and α_2 disagree on some $i \in S_1 \cap S_2$. Say $\alpha_1(i) = 0$ and $\alpha_2(i) = 1$. Then we can rewrite the inner product as

$$\langle \mathbf{V}_{(S_1,\alpha_1)}, \mathbf{V}_{(S_2,\alpha_2)} \rangle = \left\langle \mathbf{V}_{(S_1 \setminus \{i\}, \alpha_1(S_1 \setminus \{i\}))} - \mathbf{V}_{(S_1,\alpha_1')}, \mathbf{V}_{(S_2,\alpha_2)} \right\rangle$$
$$= \left\langle \mathbf{V}_{(S_1 \setminus \{i\}, \alpha_1(S_1 \setminus \{i\}))}, \mathbf{V}_{(S_2,\alpha_2)} \right\rangle - \left\langle \mathbf{V}_{(S_1,\alpha_1')}, \mathbf{V}_{(S_2,\alpha_2)} \right\rangle$$

where, as before, α_1' is equal to α_1 for all $j \neq i$ and has $\alpha_1'(i) = 1$. By the induction hypothesis, the terms on the right are either both equal to 0 or both equal $\left\langle \mathbf{V}_{(S_1 \cup S_2, \alpha_1' \circ \alpha_2)}, \mathbf{V}_{(\emptyset,\emptyset)} \right\rangle$ depending on whether α_1' and α_2 disagree or not. In either case, their difference is 0.

Next, we consider $\langle \mathbf{V}_{(S_1,\alpha_1)}, \mathbf{V}_{(S_2,\alpha_2)} \rangle$ when $\alpha_1 \circ \alpha_2$ is well defined. For all $i \in S_1 \cap S_2$ such that $\alpha_1(i) = \alpha_2(i) = 0$, we can always write $\mathbf{V}_{(S_2,\alpha_2)} = \mathbf{V}_{(S_2 \setminus \{i\}, \alpha_2(S_2 \setminus \{i\}))} - \mathbf{V}_{(S_2,\alpha_2')}$ and note that $\mathbf{V}_{(S_2,\alpha_2')}$ will be orthogonal to $\mathbf{V}_{(S_1,\alpha_1)}$ by the previous case, since it contradicts α_1 on i. Hence, we can always reduce to the case when there is an i in only one of the sets, say S_1, such that $\alpha_1(i) = 0$. Now we again decompose $\mathbf{V}_{(S_1,\alpha_1)}$, and note that

$$\langle \mathbf{V}_{(S_1,\alpha_1)}, \mathbf{V}_{(S_2,\alpha_2)} \rangle = \left\langle \mathbf{V}_{(S_1 \setminus \{i\}, \alpha_1(S_1 \setminus \{i\}))}, \mathbf{V}_{(S_2,\alpha_2)} \right\rangle - \left\langle \mathbf{V}_{(S_1,\alpha_1')}, \mathbf{V}_{(S_2,\alpha_2)} \right\rangle$$
$$= \left\langle \mathbf{V}_{(S_1 \cup S_2 \setminus \{i\}, \alpha_1(S_1 \setminus \{i\}) \circ \alpha_2)}, \mathbf{V}_{(\emptyset,\emptyset)} \right\rangle - \left\langle \mathbf{V}_{(S_1 \cup S_2, \alpha_1' \circ \alpha_2)}, \mathbf{V}_{(\emptyset,\emptyset)} \right\rangle$$
$$= \left\langle \mathbf{V}_{(S_1 \cup S_2, \alpha_1 \circ \alpha_2)}, \mathbf{V}_{(\emptyset,\emptyset)} \right\rangle$$

where we used the induction hypothesis in the second step.

- We claim that to show property (3), it is sufficient to prove that for all S, $\sum_{\alpha \in \{0,1\}^S} \|\mathbf{V}_{(S,\alpha)}\|^2 = 1$ (It is clearly necessary if we intend to interpret these values as probabilities). Before proving it, we note that it is equivalent to the condition that $\sum_{\alpha \in \{0,1\}^S} \mathbf{V}_{(S,\alpha)} = \mathbf{V}_{(\emptyset,\emptyset)}$ because it implies

$$\left\| \sum_{\alpha \in \{0,1\}^S} \mathbf{V}_{(S,\alpha)} - \mathbf{V}_{(\emptyset,\emptyset)} \right\|^2 = \sum_{\alpha \in \{0,1\}^S} \|\mathbf{V}_{(S,\alpha)}\|^2 - 1 = 0 \tag{4}$$

Using the above equivalence, we can then conclude the consistency of the probability distributions as defined in the lemma. Consider a set S and the distribution over $\alpha \in \{0,1\}^S$ given by $\mathbb{P}[\alpha] = \|\mathbf{V}_{(S,\alpha)}\|^2$. In this distribution, if $\alpha' \in \{0,1\}^{S'}$ is an event, then the probability of α' can be calculated by summing the probabilities of all the events $\alpha \in \{0,1\}^S$ such that

37

$\alpha(S') = \alpha'$. This must also equal $\left\|\mathbf{V}_{(S',\alpha')}\right\|^2$, which would be the probability of α' if we just considered the distribution over S'. This follows from (4) because,

$$\sum_{\alpha(S')=\alpha} \left\|\mathbf{V}_{(S,\alpha)}\right\|^2 = \sum_{\alpha(S')=\alpha'} \left\langle\mathbf{V}_{(S,\alpha)}, \mathbf{V}_{(\emptyset,\emptyset)}\right\rangle = \sum_{\alpha_1 \in \{0,1\}^{S\setminus S'}} \left\langle\mathbf{V}_{(S\setminus S',\alpha_1)}, \mathbf{V}_{(S',\alpha')}\right\rangle$$

$$= \left\langle\mathbf{V}_{(\emptyset,\emptyset)}, \mathbf{V}_{(S',\alpha')}\right\rangle = \left\|\mathbf{V}_{(S',\alpha')}\right\|^2$$

To prove that $\sum_{\alpha \in \{0,1\}^S} \left\|\mathbf{V}_{(S,\alpha)}\right\|^2 = 1$ for all S with $|S| \leq t/2$, we proceed by induction on $|S|$. The base case (empty set) is trivial. To do the induction step, note that for any $i \in S$

$$\sum_{\alpha \in \{0,1\}^S} \left\|\mathbf{V}_{(S,\alpha)}\right\|^2 = \sum_{\alpha_1 \in \{0,1\}^{S\setminus\{i\}}} \left\langle\mathbf{V}_{(S\setminus\{i\},\alpha_1)}, \mathbf{V}_{(\{i\},0)}\right\rangle + \sum_{\alpha_1 \in \{0,1\}^{S\setminus\{i\}}} \left\langle\mathbf{V}_{(S\setminus\{i\},\alpha_1)}, \mathbf{V}_{(\{i\},1)}\right\rangle$$

$$= \left\langle\mathbf{V}_{(\emptyset,\emptyset)}, \mathbf{V}_{(\{i\},1)}\right\rangle + \left\langle\mathbf{V}_{(\emptyset,\emptyset)}, \mathbf{V}_{(\{i\},1)}\right\rangle = \left\langle\mathbf{V}_{(\emptyset,\emptyset)}, \mathbf{V}_{(\emptyset,\emptyset)}\right\rangle = 1$$

- Finally, to show that the vectors $\mathbf{V}_{(S,0_S)}$ form a valid solution to the vertex cover SDP, we note that they satisfy all the consistency conditions by the previous arguments. The only extra condition that the SDP would impose is that for any edge (i,j), $\left\langle\mathbf{U} - \mathbf{V}_{(\{i\},0)}, \mathbf{U}_\emptyset - \mathbf{V}_{(\{j\},0)}\right\rangle = 0$. But this is immediate because $\mathbf{U}_\emptyset - \mathbf{V}_{(\{i\},0)} = \mathbf{U}_{\{i\}}$ (similarly for j) and $\left\langle\mathbf{U}_{\{i\}}, \mathbf{U}_{\{j\}}\right\rangle = 0$. ∎