Linear Level Lasserre Lower Bounds for Certain k-CSPs

Grant Schoenebeck*

Abstract

We show that for $k \geq 3$ even the $\Omega(n)$ level of the Lasserre hierarchy cannot disprove a random k-CSP instance over any predicate type implied by k-XOR constraints, for example k-SAT or k-XOR. (One constant is said to imply another if the latter is true whenever the former is. For example k-XOR constraints imply k-CNF constraints.) As a result the $\Omega(n)$ level Lasserre relaxation fails to approximate such CSPs better than the trivial, random algorithm. As corollaries, we obtain $\Omega(n)$ level integrality gaps for the Lasserre hierarchy of $\frac{7}{6} - \varepsilon$ for VertexCover, $2 - \varepsilon$ for k-UniformHypergraphVertexCover, and any constant for k-UniformHypergraphIndependentSet. This is the first construction of a Lasserre integrality gap.

Our construction is notable for its simplicity. It simplifies, strengthens, and helps to explain several previous results.

1 Introduction

The Lasserre hierarchy [Las01] is a sequence of semidefinite relaxations for certain 0-1 polynomial programs, each one more constrained than the last. The kth level of the Lasserre hierarchy requires that any set of k original vectors be self-consistent in a very strong way. If an integer program has n variables, the nth level of the Lasserre hierarchy is sufficient to obtain a tight relaxation where the only feasible solutions are convex combinations of integral solutions. This is because the nth level requires that the entire set of n vectors are consistent. If one starts from a k-CSP with poly(n) constraints, then it is possible to optimize over the set of solutions defined by the kth level of Lasserre in time $O(n^{O(k)})$, which is sub-exponential for $k = o(n/\log n)$.

The Lasserre hierarchy is similar to the Lovasz-Schrijver hierarchies [LS91], denoted LS and $LS+$ for the linear and semidefinite versions respectively, and the Sherali-Adams [SA90] hierarchy, denoted SA; however, the Lasserre hierarchy is stronger [Lau03]. The region of feasible solutions in ℓth level of the Lasserre hierarchy is always contained in the region of feasible solutions in ℓth level of LS, $LS+$, and SA[1]. A more complete comparison can be found in [Lau03]. While there have been a growing number of integrality gap lower bounds for the LS[ABL02, ABLT06, Tou06, STT07b], the $LS+$[BOGH+03, AAT05, STT07a, GMPT06], and the SA[dlVKM07, CMM07] hierarchies, similar bounds for the Lasserre hierarchy have remained elusive.

The study of these hierarchies is motivated by the success of semidefinite programs in approximation algorithms. In many interesting cases, for small constant ℓ, the ℓth level of the Lasserre hierarchy

*`grant@cs.berkeley.edu`, Computer Science Division, UC Berkeley. This material is based upon work supported under a National Science Foundation Graduate Research Fellowship. Work for this paper was done while author was visiting Princeton University and Microsoft Research-Silicon Valley

[1] In our definition, for ease of presentation, the ℓth level of Lasserre for a k-CSP is only meaningful if $\ell \geq k$, but this can be modified.

provides the best known polynomial-time computable approximation. For example, the first level of the Lasserre hierarchy for the IndependentSet problem implies the Lovasz θ-function and for the MaxCut problem gives the Goemans-Williamson relaxation. The ARV relaxation of the SparsestCut problem is no stronger than the relaxation given in the third level of Lasserre.

In addition, recent work by Eden Chlamtac [Chl07] has shown improved approximation algorithms for coloring and independent set in 3-uniform hypergraphs. In [Chl07] the Lasserre hierarchy was used to find and/or analyze the constraints which led to improved approximations. This work is unlike the aforementioned work, where it was only later realized that the approximation results could be viewed as an application of semidefinite program hierarchies.

Integrality gap results for Lasserre are thus very strong unconditional negative results, as they apply to a "model of computation" that includes the best known algorithms for several problems.

1.1 Previous Lower-Bounds Work

While this is the only work known to us on Lasserre integrality gaps, results are already known about the weaker hierarchical models for several problems, including many problems we study here.

Buresh-Oppenheim, Galesy, Hoory, Magen and Pitassi [BOGH+03], and Alekhnovich, Arora, Tourlakis [AAT05] prove $\Omega(n)$ LS+ round lower bounds for proving the unsatisfiability of random instances of 3-SAT (and, in general, k-SAT with $k \geq 3$) and $\Omega(n)$[2] round lower bounds for achieving approximation factors better than $7/8 - \varepsilon$ for Max 3-SAT, better than $(1-\varepsilon)\ln n$ for Set Cover, and better than $k-1-\varepsilon$ for HypergraphVertexCover in k-uniform hypergraphs. They leave open the question of proving LS+ round lower bounds for approximating the Vertex Cover problem.

Much work has been done on Vertex Cover. Schoenebeck, Tulsiani, and Trevisan[STT07b] show an integrality gap of $2-\varepsilon$ remains after $\Omega(n)$ rounds of LS, which is optimal. This build on the previous work of Arora, Bollobas, Lovasz, and Tourlakis [ABL02, ABLT06, Tou06] who prove that even after $\Omega(\log n)$ rounds the integrality gap of LS is at least $2-\varepsilon$, and that even after $\Omega((\log n)^2)$ rounds the integrality gap of LS is at least $1.5-\varepsilon$.

Somewhat weaker results are known for $LS+$. The best known results are incomparable and were show by shown by Georgiou, Magen, Pitassi, and Tourlakis[GMPT06] and Schoenebeck, Tulsiani, and Trevisan [STT07a]. The former result [GMPT06] builds on the previous ideas of Goemans and Kleinberg [KG98] and Charikar [Cha02], and shows that an integrality gap of $2-\varepsilon$ survives $\Omega(\sqrt{\frac{\log n}{\log \log n}})$ rounds of LS+. The later result shows an integrality gap of $\frac{7}{6}-\varepsilon$ survives $\Omega(n)$ rounds. This result builds on past research which we review here as it is relevant for understanding the results of this paper.

The result of Feige and Ofek [FO06] immediately implies a $17/16-\varepsilon$ integrality gap for one round of LS+, and the way in which they prove their result implies also the stronger $7/6-\varepsilon$ bound. The standard reduction from Max 3-SAT to VertexCover shows that if one is able to approximate VertexCover within a factor better than 17/16 then one can approximate Max 3-SAT within a factor better than 7/8. This fact, and the $7/8-\varepsilon$ integrality gap for Max 3-SAT of [AAT05], however do not suffice to derive an $LS+$ integrality gap result for VertexCover. The reason is that reducing an instance of Max 3SAT to a graph, and then applying a VertexCover relaxation to the graph, defines a semidefinite program that is possibly tighter than the one obtained by a direct relaxation of the Max 3-SAT problem. Feige and Ofek [FO06] are able to analyze the value of the Lovasz θ-function of the graph obtained by taking a random 3-SAT instance and then

[2] In all integrality gap containing an ε, the constant in the Ω depends on ε.

reducing it to an instance of IndependentSet (or, equivalently, of VertexCover).

For the Sherali-Adams hierarchy, Charikar, Makarychev, and Makarychev [CMM07] show that, for some ε, after n^{ε} rounds an integrality gap of $2 - o(1)$ remains.

Other results by Charikar [Cha02] and Hatami, Magen, and Markakis [HMM06] prove a $2 - o(1)$ integrality gap result for semidefinite programming relaxations of Vertex Cover that include additional inequalities. Charikar's relaxation is implied by the relaxation obtained after two rounds of Lasserre. The semidefinite lower bound of Hatami et al is implied after five rounds of Lasserre.

It was compatible with previous results that after a constant number of rounds of Lasserre the integrality gap for Vertex Cover could become $1 + o(1)$.

Our Result

The main result of this paper, is a proof that, for $k \geq 3$, the $\Omega(n)$th level of Lasserre cannot prove that a random k-CSP over any predicate implied by k-XOR is unsatisfiable. From this main results it quickly follows that the $\Omega(n)$th level of Lasserre:

- cannot prove a random k-XOR formula unsatisfiable.
- cannot prove a random k-SAT formula unsatisfiable.
- contains integrality gaps of $1/2 + \varepsilon$ for Max-k-XOR
- contains integrality gaps of $1 - \frac{1}{2^k} + \varepsilon$ for Max k-SAT.
- contains integrality gaps of $\frac{7}{6} - \varepsilon$ for VertexCover.
- contains integrality gaps of any constant for k-UniformHypergraphVertexCover.
- contains integrality gaps of $\Omega(1)$ for k-UniformHypergraphIndependentSet.

In addition to the power of our result, it is also very short and simple. It extends and simplifies results in [STT07a] and [AAT05]. To a large extent it also explains the proofs of [FO06] and [STT07a], and can be seen as being inspired by these results.

Road Map

In Section 2 we will define notation and provide background to our results. In Section 3 we will prove the main result. In Section 4 we will state and prove the remaining results, which are corollaries of the main result.

2 Background and Notation

We denote the set of Boolean variables $[n] = \{1, \ldots, n\}$. Let the range of variables be denoted $\mathbf{x} = \{x_i\}_{i \in [n]} = \{0,1\}^n$. For $I \subseteq \{1, \ldots, n\}$, let $\mathbf{x}_I = \{x_i\}_{i \in I}$ be the projection of $\mathbf{x}$ to the coordinates of I. We will consider programs where each constraint is local, captured by the following definition:

Definition 1 *A k-constraint C_I^f is a function $f : \mathbf{x}_I \rightarrow \{0,1\}$ where $|I| \leq k$.*

Note that given a constraint C_I^f where $f : \mathbf{x}_I \to \{0,1\}$, we can naturally extend it to a constraint C_J^f where $I \subseteq J$ and $f : \mathbf{x}_J \to \{0,1\}$ by first projecting to the variables of I and then applying f. Sometimes we abuse notation and denote by C_I^f the set $\{x_I \in \mathbf{x}_I : f(x_I) = 1\}$.

Definition 2 *A k-constraint C_I^f* implies *another k-constraint C_I^g if $C_I^f \subseteq C_I^g$. We say that a predicate is XOR-implied if it is implied by either parity or its negation.*

For notational convenience, we will denote by $C_I^{x_I}$ the constraint C_I^f where $f(\bar{x}_I) = 1$ if $\bar{x}_I = x_I$ and 0 otherwise. Also, we denote by $C_I^{\vec{1}}$ the constraint $C_I^{x_I}$, where x_I is one in each coordinate; by C_I^* the constraint that is always satisfied; and by $C_I^{\emptyset}$ the constraint that is never satisfied.

We will look at relaxations for two types of integer programs. In the first, we have a set of constraints, and would like to know if there is any feasible solution. In the second, we have a set of constraints and would like to maximize some objective function subject to satisfying the constraints. We formalize the notions here:

Definition 3 *A k-*constraint satisfiability problem *$\langle \mathbf{x}, \mathbf{C_I^f} \rangle$ is a set of n Boolean variables $\mathbf{x} = \{0,1\}^n$, and a set of k-constraints $\{C_I^f\}$.*

Definition 4 *A k-*constraint maximization (or minimization) problem *$\langle \mathbf{x}, \mathbf{C_I^f}, M \rangle$ is a set of n Boolean variables $\mathbf{x} = \{0,1\}^n$, a set of k-constraints $\{C_I^f\}$, and a polynomial objective function M of total degree at most k such that $M : \mathbf{x} \to \mathbf{Z}$ is to be maximized (or minimized).*

Lasserre Let $\langle \mathbf{x}, \mathbf{C_I^f}, M \rangle$ be a constraint maximization (or minimization) problem. Ideally, we would like to say that a solution $(y_1, \ldots, y_n)$ in the feasible region of any level of the Lasserre hierarchy must be the convex combination of integer solutions; however, enforcing this directly is difficult. Instead we note that if $(y_1, \ldots, y_n) = \sum_{j=1}^m p_j (z_1^j, \ldots, z_n^j)$ is from a probability distribution of integral solutions, that is $(y_1, \ldots, y_n) = \sum_{j=1}^m p_j (z_1^j, \ldots, z_n^j)$ where $z_i^j \in \{0,1\}$, $z^j = (z_1^j, \ldots, z_n^j)$ are a feasible integral solutions, and $\sum_{j=1}^m p_j = 1$ then, for each possible k-constraint C_I^f we can produce a vector[3]

$$v_{C_I^f}(j) = \begin{cases} \sqrt{p_j} & C_I^f(z^j) = 1 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

If we define scalar variables $x_{C_I^f}$ so that $x_{C_I^f} = ||v_{C_I^f}||^2$, and think of $(y_1, \ldots, y_n)$ as a probability distribution over integer solutions, then $x_{C_I^f}$ is the probability that a randomly drawn solution satisfies the constraint C_I^f. These vectors will satisfy all the constraints of the Lasserre hierarchy at any level. If the reader is unfamiliar with the definition of Lasserre, then it is a straightforward and useful exercise to verify this fact.

Definition 5 *The rth round of Lasserre on the k-constraint maximization problem $\langle \mathbf{x}, \mathbf{C_I^f}, M \rangle$ is the semidefinite program with a variable $v_{C_I^f}$ for every k-constraint C_I^f (not just those in $\mathbf{C_I^f}$). Let $M = \sum_{i=1}^m w_i \prod_{j \in I_i} x_j$ be the objective function expressed as the weighted sum of monomials. We*

[3]That is for *any* function $f : x_I \to \{0,1\}$ where $|I| \leq k$, not simply the constraints that appear in the constraint maximization problem.

will denote by v_0 the vector for the constraint $C^*_\emptyset$, denote by v_{x_I} the vector for the constraint $C_I^{x_I}$, denote by v_i the vector for the constraint $C^{\vec{1}}_{\{i\}}$, and denote by $v^{\vec{1}}_{x_I}$ the vector for the constraint $C^{\vec{1}}_I$.

$$\max \sum_{i=1}^{m} w_i ||v^{\vec{1}}_{I_i}||^2$$

where

$$\forall I \subseteq [n] \qquad ||v_{C_I^*}||^2 = 1 \tag{2}$$

$$\forall f_I \in \mathbf{C_I^f} \qquad ||v_{C_I^f}||^2 = 1 \tag{3}$$

$$\begin{array}{rc} \forall C_I^f, C_J^g, C_{I'}^{f'}, C_{J'}^{g'} \text{ where} & \\ I \cup J = I' \cup J' & \\ (C_I^f \cdot C_J^g) \equiv (C_{I'}^{f'} \cdot C_{J'}^{g'}) & \langle v_{C_I^f}, v_{C_J^g} \rangle = \langle v_{C_{I'}^{f'}}, v_{C_{J'}^{g'}} \rangle \end{array} \tag{4}$$

$$\forall C_I^f \qquad v_{C_I^f} = \sum_{x_I \in C_I^f} v_{x_I} \tag{5}$$

We compare the equivalence of $(C_I^f \cdot C_J^g)$ and $(C_{I'}^{f'} \cdot C_{J'}^{g'})$ as functions on the domain $\mathbf{x}_{I \cup J}$.

The semidefinite program for the rth Lasserre round of a satisfiability problem is the same, but we only check for the existence of feasibility, we do not try to maximize over any objective function. [4]

While the equations are confusing, the intuition is that the vectors define a probability distribution on any set of up to r coordinates (Equations 2, 4, and 5); that the probability distributions always satisfy the constraints (Equation 5); and that the probability distributions properly patch together (Equation 4). It is easy to check the suggested vectors satisfy all these constraints.

Claim 6 *Fix $I \subseteq [n]$ such that $|I| \le r$. Then we can get probability distribution over the elements of $x_I \in \mathbf{x}_I$ by defining the probability of x_I to be $||v_{x_I}||^2$. Actually, these vectors are all orthogonal, and if you sum over them, you get v_0.*

PROOF: If $x_I, x_I' \in \mathbf{x}_I$, then v_{x_I} and $v_{x_I'}$ are orthogonal because $C_I^{x_I} \cdot C_I^{x_I'} = C_I^\emptyset$ and so by Equation 4 $\langle v_{x_I}, v_{x_I'} \rangle = ||v_{C_I^\emptyset}||^2$ and by Equation 5 $||v_{C_I^\emptyset}||^2 = 0$

Thus, by Equation 2 then Equation 5: $1 = ||v_{C_I^*}||^2 = ||\sum_{x_I \in \mathbf{x}_I} v_{x_I}||^2 = \sum_{x_I \in \mathbf{x}_I} ||v_{x_I}||^2$. So indeed we have a probability distribution.

By Equations 2 and 4 $\forall I \subseteq [n]$, $v_{C_I^*} = v_0$ because $\langle v_{C_I^*} - v_0, v_{C_I^*} \rangle = ||v_{C_I^*}||^2 - \langle v_{C_I^*}, v_0 \rangle = 1 - 1 = 0$. □

In applications, it is usually important that we have vectors and not simply local distributions that patch together. The fact that we have vectors gives some global orientation. The Goemans-Williamson MAXCUT algorithm generates a global cut with a hyperplane. It is not clear how to do this with a local distributions alone.

[4] This definition is slightly different, but equivalent to other definitions of the kth level of the Lasserre hierarchy. The way that it is stated, it would require double exponential time to solve the rth level. This is easily remedied by only defining vectors for the constraints $C_I^{\vec{1}}$ and using linear combinations of these vectors to define the remaining vectors. We present it like this for ease of notation.

Claim 7 *If Equations 2, 4 and 5 are satisfied, then Equation 3 is equivalent to requiring that* $||v_{x_I}||^2 = 0$ *for all* x_I *where* $x_I \notin C_I^f$ *for some* $C_I^f \in \mathbf{C_I^f}$.

Proof: We only used Equations 2, 4 and 5 to show Claim 6. So we know that the v_{x_I} are all orthogonal and by Equation 5 additionally know that

$$\begin{aligned} 0 &= 1 - 1 = ||v_0||^2 - 1 = ||C_I^*||^2 - 1 = ||\sum_{x_I \in \mathbf{x}_I} v_{x_I}||^2 - ||\sum_{x_I \in C_I^f} v_{x_I}||^2 \\ &= \sum_{x_I \in \mathbf{x}_I} ||v_{x_I}||^2 - \sum_{x_I \in C_I^f} ||v_{x_I}||^2 = \sum_{x_I \notin C_I^f} ||v_{x_I}||^2 \end{aligned}$$

□

Problems Studied Let $\mathcal{P}$ be a set of boolean predicates on k-variables. In a k-CSP-$\mathcal{P}$ we are given a set of predicates (or clauses) which are each taken from $\mathcal{P}$. Each clause becomes a constraint, and we are asked if all the constraints can be simultaneously satisfied. In Max- k-CSP-$\mathcal{P}$ we want to find the maximum number of clauses that can be satisfied.

We can define a distribution $\mathcal{D}$ over predicates in some $\mathcal{P}$. To sample a random k-CSP-$\mathcal{P}$ formula with Δn clauses, we uniformly and draw Δn clauses from the set of $2^k\binom{n}{k}|\mathcal{P}|$ possible clauses by first uniformly and independently sample each set of k variables and the sign applied to each variable, and then draw a predicate from $\mathcal{D}$.

In k-XOR we are given a set of clauses which are each of the form $\bigoplus_{i \in I} x_i = 0/1$ where $|I| \leq k$. We will denote the clause $\bigoplus_{i \in I} x_i = b$ by $C_I^{\oplus I = b}$. Each clause becomes a constraint, and we are asked if all the constraints can be simultaneously satisfied. To sample a random k-XOR formula with Δn clauses, we uniformly and independently draw Δn clauses from the set of $2\binom{n}{k}$ possible clauses.

In k-SAT we are given a set of clauses which are each of the form $\vee_{i \in I} x_i$ where $|I| \leq k$. Each clause becomes a constraint, and we are asked if all the constraints can be simultaneously satisfied. In Max k-SAT we want to find the maximum number of clauses that can be satisfied. To sample a random k-SAT formula with Δn clauses, we uniformly and independently draw Δn clauses from the set of $2^k\binom{n}{k}$ possible clauses.

Definition 8 *Give a distribution* $\mathcal{D}$ *over predicates in some* $\mathcal{P}$ *we define* $r(\mathcal{P})$ *to be the probability that a random assignment satisfies a predicate drawn from* $\mathcal{D}$.

For example, in k-XOR, $\mathcal{D}$ is uniformly distributed between k-parity and its negation, and $r(k\text{-XOR}) = 1/2$.

In VertexCover we are given a graph $G = (V, E)$. There is a Boolean variable x_i for each vertex $i \in V$. For each edge $(i, j) \in E$ we have a constraint which says that both x_i and x_j cannot be zero. We are asked to minimize $\sum_{i \in V} x_i$.

In k-UniformHypergraphIndependentSet we are given a k-uniform hypergraph $G = (V, E)$. There is a variable x_i for each vertex $v \in V$. For each edge $(i_1, \ldots, i_k) \in E$ we have a constraint which says that not all $x_{i_1}, \ldots, x_{i_k}$ can be one. We are asked to maximize $\sum_{i \in V} x_i$.

k-UniformHypergraphVertexCover is the same as k-UniformHypergraphIndependentSet except that for each edge $(i_1, \ldots, i_k) \in E$ we have a constraint which says that at least one of $x_{i_1}, \ldots, x_{i_k}$ must be one. We are asked to minimize. $\sum_{i \in V} x_i$.

Background Results Sufficiently dense random k-CSP formulae are far from being satisfiable as the next proposition states.

Proposition 9 *For any $\delta > 0$, with probability $1 - o(1)$, if φ is a random k-CSP chosen from the distribution $\mathcal{D}$ with Δn clauses where $\Delta \geq \frac{\ln 2}{2\delta^2} + 1$, at most a $r(\mathcal{D}) + \delta$ fraction of the clauses of φ can be simultaneously satisfied.*

Proposition 9 is well known in the literature, we provide a proof in the appendix for completion.

Definition 10 Width-w resolution *on an XOR formula φ, successively builds up new clauses by deriving a new clause $\bigoplus_{i \in I \Delta J} x_i = b \oplus b'$ whenever the symmetric difference $|I \Delta J| \leq w$ and the clauses $\bigoplus_{i \in I} x_i = b$ and $\bigoplus_{i \in I} x_i = b'$ had either already been derived or belong to φ.*

Width-w resolution proves a formula φ unsatisfiable if it derives the clause $0 = 1$. The following theorem shows that for random 3-XOR formula, even for quite large w, width-w resolution fails to produce a contradiction.

Theorem 11 *For $k \geq 3$, $d > 0$, $\gamma > 0$, and $0 \leq \varepsilon < k/2 - 1$, if φ is a random k-XOR formula with density dn^ε, then with probability $1 - o(1)$ φ cannot be disproved by width $\alpha n^{1 - \frac{\varepsilon}{k/2 - \gamma - 1}}$ resolution nor can any variable be resolved to true or false. Furthermore, this is true even if the parity sign (whether the predicate is parity or its negation) of each clause is adversatively chosen.*

Wigderson and Ben-Sasson [BSW01] show that a variant of Theorem 11 holds for k-SAT formula. The proof of [BSW01] extends to show Theorem 11 using standard techniques. We include a proof in the appendix for completeness.

3 k-CSPs over XOR-Implied Predicates

We now present the main theorem of the paper.

Theorem 12 *Let $\mathcal{D}$ be a distribution over a set of XOR-implied k-CSP predicates $\mathcal{P}$. Then for every $\delta, \gamma, d > 0$ and $0 \leq \varepsilon < k/2 - 1$ (such that if $\varepsilon = 0$, then $d \geq \frac{\ln 2}{2\delta^2} + 1$) there exists some constants $\alpha \geq 0$, such that with probability $1 - o(1)$, if φ is a random k-CSP formula drawn according to $\mathcal{D}$ with Δn clauses where $\Delta = dn^\varepsilon$ both the following are true:*

1. *at most a $r(D) + \delta$ fraction of the clauses of φ can be simultaneously satisfied.*
2. *The $\alpha n^{1 - \frac{\varepsilon}{k/2 - 1 - \gamma}}$ level of the Lasserre hierarchy permits a feasible solution.*

This theorem implies integrality gaps for *XOR*-implied k-CSPs because the Lasserre relaxation cannot refute that all clauses can be simultaneously satisfied, but, in fact, at most $r(D) + \delta$ clauses can be simultaneously satisfied. Notice that an algorithm that simply guesses a random assignment would expect to satisfy an $r(D)$ fraction of clauses in expectation. In particular this theorem shows that with high probability a random k-XOR formula cannot be refuted by $\Omega(n)$ rounds of Lasserre which gives an integrality gap of $1/2+\varepsilon$ for $\Omega(n)$ rounds of Lasserre for Max k-XOR. Also, this theorem shows that with high probability a random 3-CNF formula cannot be refuted by $\Omega(n)$ rounds of Lasserre which gives an integrality gap of $7/8+\varepsilon$ for $\Omega(n)$ rounds of Lasserre for Max k-SAT.

Theorem 12 follows almost immediately from Theorem 11, Proposition 9, and the following Lemma.

Lemma 13 (Main Lemma) *If a k-XOR formula φ cannot be disproved by width-w resolution, then the $\frac{w}{4}$th round of the Lasserre hierarchy permits a feasible solution.*

Proof:[of Theorem 12] Fix $\delta, \gamma, d, \varepsilon, \mathcal{P}, \mathcal{D}$ as allowed in theorem statement, and let φ be a random k-CSP $\mathcal{P}$ formula with Δn clauses where $\Delta = dn^{\varepsilon}$. By Proposition 9, 1) holds with probability $1-o(1)$ because for sufficiently large n, $\Delta = dn^{\varepsilon} > \frac{\ln 2}{2\delta^2}+1$.

We can write φ as a k-XOR formula φ_{XOR} so that $\varphi_{XOR} \Rightarrow \varphi$. Now the Lasserre relaxation for φ_{XOR} is strictly tighter than that for φ. Let α' be as guaranteed in Theorem 11 using k, d, γ, and ε as inputs so that by Theorem 11 we know that with probability $1-o(1)$ it is the case that φ_{XOR} cannot be disproved by width-$\alpha' n^{1-\frac{\varepsilon}{k/2-\gamma-1}}$ resolution. Let $\alpha = \frac{\alpha'}{4}$. By Lemma 13, φ_{XOR} cannot be proven unsatisfiable by $\frac{\alpha'}{4} n^{1-\frac{\varepsilon}{k/2-\gamma-1}} = \alpha n^{1-\frac{\varepsilon}{k/2-\gamma-1}}$ rounds of Lasserre. Because the Lasserre relaxation for φ_{XOR} is tighter than that for φ_{XOR} it must be the case that φ cannot be proven unsatisfiable by Lasserre either. $\square$

Lemma 13 is the main original technical contribution of this work. In the rest of this section we first provide some intuition for the proof of Lemma 13 and then provide its proof.

For a first attempt to prove the lemma we can observe that for any particular set I of at most $w/4$ variables, we can construct vectors for all C_I^f as follows: 1) Run bounded width resolution to derive a set of constraints that any satisfying assignment must satisfy. 2) Consider the set SAT_I where $SAT_I = \{x_I \in \mathbf{x}_I : x_I$ satisfies all the constaints derived by the resolution whose support is contained in $I\}$. Randomize over SAT_I and construct the vectors as we saw in Equation 1. These vectors will satisfy the Lasserre Equations 2, 3, and 5; however, these vectors will fail miserably to satisfy Equation 4 of the Lasserre constraints. We have set up valid local distributions; however, these distribution do not patch together consistently. The problem is that when the take the dot product of v_{x_I} and v_{x_J}, the values in each coordinate mean something completely different.

To remedy this misalignment we design a space of equivalence classes of XORs of at most $w/2$ variables which we will use to index the coordinates of each vector. We will say that $\bigoplus_{i\in I} x_i \sim \bigoplus_{j\in J} x_j$ if for all assignments that satisfy the derived resolution clauses, $\bigoplus_{i\in I} x_i$ determines $\bigoplus_{i\in J} x_i$ and vice versa. For example, if φ contained the clause $x_1 \oplus x_2 \oplus x_3 = 0$ then $x_1 \oplus x_2 \sim x_3$ because whatever $x_1 \oplus x_2$ is, x_3 must be the opposite. With some $\sim$ equivalent clauses, fixing one clause automatically fixes the $\sim$ equivalent clause to the opposite value (as above). With other $\sim$ equivalent clauses, fixing one clause automatically fixes the $\sim$ equivalent clause to the same value. Using this fact, we can split each equivalence class of $\sim$ equivalent clauses into two parts, so that the $\sim$ equivalent clauses in each part always fix each other to the same value, and $\sim$ equivalent clauses in opposite parts always fix each other to the opposite value. We can arbitrarily label one part $+$ and the other $-$. There is a bijection between the set SAT_I and the equivalence classes of $\bigoplus_{i\in J} x_i$ where $J \subseteq I$, because, intuitively, each time resolution derives a new relation, the dimension of each of these sets is reduced by 1.

Finally, for each $x_I \in SAT_I$ we construct a vector where each coordinate is indexed by the $\sim$ equivalence classes we just produced. This vector is $\pm\frac{1}{|SAT_I|}$ in each equivalence class that contains an set $J \subseteq I$, and is 0 elsewhere. The sign is $+$ if x_I agrees with the $+$ side of the equivalence class and $-$ otherwise. If we project onto only the non-zero coordinates, then the mapping of our previously constructed vectors (that failed to satisfy Equation 4) to these new vectors is simply a rotation. This implies that all the Lasserre equations that were previously satisfied will still be satisfied. This rotation is similar to taking a Fourier transform. If any two distinct vectors v_{x_I} and v_{x_J} disagree in any non-zero location, they disagree in exactly 1/2 of the places and are orthogonal. Otherwise the signs agree in every non-zero coordinate. These facts allow us to show that these vectors do satisfy Equation 4 of the Lasserre constraints.

We will now prove Lemma 13.

Proof:[Lemma 13]

Construction of Vectors

We first define a set $\mathcal{C}$ which later will be used to index the coordinates of the vectors.

Let φ be a k-XOR formula that has no width-w resolution. Let res-$\mathcal{C}$ be the collection of clauses generated by width-w resolution running on φ. By the hypothesis of the Lemma we are guaranteed that we cannot derive a contradiction.

Let $\frac{w}{2}$-$\mathcal{C}$ be the collection of all possible XOR clauses over at most $\frac{w}{2}$ variables. Now consider the set $\frac{w}{2}$-$\mathcal{C}$/res-$\mathcal{C}$. That is we partition $\frac{w}{2}$-$\mathcal{C}$ into equivalence classes where $C_I^{\oplus I=b} \sim_{\text{res-}\mathcal{C}} C_J^{\oplus J=b'} \Leftrightarrow C_{I\Delta J}^{\oplus(I\Delta J)=(b\oplus b')} \in$ res-$\mathcal{C}$.

Let $S^{|w/2|} = \{\bigoplus_{i\in I} x_i : I \subseteq [n] : |I| \leq w/2\}$. Let $F \subseteq S^{|w/2|}$ be the XOR functions fixed by res-$\mathcal{C}$, that is $I \in F \Leftrightarrow \exists b \in \{0,1\}$ where $C_I^{\oplus I=b} \in$ res-$\mathcal{C}$. Consider the set $\mathcal{C} = S^{|w/2|}/F$–that is $I \sim_F J \Leftrightarrow I\Delta J \in F$.

For each equivalence class $[I] \in \mathcal{C}$, we arbitrarily choose some $I_0 \in [I]$ (for notational convenience, we always choose $\emptyset \in [\emptyset]$). We define a function $\pi : S^{|w/2|} \to \{+1, -1\}$ such that

$$\pi(I) = \begin{cases} +1 & C_I^{\oplus I=0} \sim_{\text{res-}\mathcal{C}} C_{I_0}^{\oplus I_0=0} \\ -1 & C_I^{\oplus I=0} \sim_{\text{res-}\mathcal{C}} C_{I_0}^{\oplus I_0=1} \end{cases}$$

Claim 14 *$\sim_F$ and $\sim_{\text{res-}\mathcal{C}}$ are equivalence relations and π is well defined.*

Fix some $V \subseteq [n]$ where $|V| \leq w/4$. Let $S_V^{|w/2|} = \{\bigoplus_{i\in I} x_i : I \subseteq V\}$ denote the collection of all possible XOR functions over V. Similarly define F_V and (res-$\mathcal{C}$)$_V$. We will denote $S_V^{|w/2|}$ simply S_V in the future.

If we interpret S_V as a group under the action of symmetric difference, then F_V is a subgroup of S_V. We can then consider the group S_V/F–that is $I \sim_F J \Leftrightarrow I\Delta J \in F$. Note also that $S_v/F \cong S_v/F_v$.

Let $SAT_V \subseteq \mathbf{x}_V$ where $x_V \in SAT_V$ if it satisfies all the clauses in (res-$\mathcal{C}$)$_V$. $|SAT_\emptyset| = 1$ by convention.

Claim 15 $|SAT_V| = |S_V/F_V|$

Proof: Consider the natural bijection between $\mathbf{x}_V$ and S_V where $x_V \in \mathbf{x}_V$ is considered as a function that XORs the non-zero bits of x_V. Then we can move freely between those two spaces, so it makes sense to write $\mathbf{x}_V/F_V \cong S_V/F_V$. Let $[\emptyset]$ be the equivalence class of $\emptyset$ in S_V/F_V, and let

$\bar{x}_V \in \mathbf{x}_V$ be a satisfying assignment to all the clauses in (res-$\mathcal{C})_V$. Then for $x_V \in \mathbf{x}_V$, $x_V \in SAT_V$ if and only if $x_V \in \bar{x}_V + [\emptyset]$. It follows that $|SAT_V| = |\mathbf{x}_V/F_V| = |S_V/F_V|$. $\square$

We now define the vectors. Let $|I| \leq w/4$. If $x_I \notin SAT_I$, $v_{x_I} = 0$. Otherwise:

$$v_{x_I}([J]) = \begin{cases} \frac{(-1)^{\oplus_{k\in K} x_k}\pi(K)}{|SAT_I|} & K \in [J] \text{ s.t. } K \subseteq I \\ 0 & \forall K \in [J] \;\; K \not\subseteq I \end{cases}$$

where the range of the coordinates is $[J] \in \mathcal{C}$

Recall that v_{x_I} is the constraint in $\mathbf{x}_I$ that is only satisfied on input x_I.

Intuitively, each vector v_{x_I} is non-zero exactly in the coordinates corresponding to $[J] \in S_I/F_I$. The sign of each non-zero coordinate $[J] \in S_I/F_I$ corresponds to the evaluation of $\oplus_{j\in J} x_j$ (and the $\pi(K)$ terms makes the vectors well defined).

We obtain the vectors for other constraints by taking linear combinations of these vectors.

$$v_{C_I^f} = \sum_{x_I : C_I^f(x_I)=1} v_{x_I}$$

Remark 1 *If the width-bounded resolution not only does not refute φ, but also does not fix any variable x_i to either true or false, then for every $i \in [n]$, $|SAT_{\{i\}}| = 2$ and so $||v_i||^2 = 1/2$.*

Proof that constructed vectors satisfy Lasserre constraints

To show that Equations 2 and 3 are satisfied, we use the following claim.

Claim 16

$$\langle v_{x_I}, v_{x_J}\rangle = \begin{cases} \frac{1}{|SAT_{I\cup J}|} & \forall [K] \in (S_I/F \cap S_J/F), sign(v_{x_I}([K])) = sign(v_{x_J}([K])) \\ 0 & otherwise \end{cases}$$

For any $x_I \in \mathbf{x}_I$ and $x_J \in \mathbf{x}_J$, only coordinates where both v_{x_I} and v_{x_J} are non-zero will contribute to the value of $\langle v_{x_I}, v_{x_J}\rangle$. Claim 16 states that either there is no cancelation amongst non-zero coordinates (all the coordinates where both v_{x_I} and v_{x_J} are non-zero have the same sign), or there is complete cancelation ($\langle v_{x_I}, v_{x_J}\rangle = 0$).

Proof: Note that v_{x_I} and v_{x_J} are both non-zero only in the coordinates $[K] \in S_I/F \cap S_J/F$. Let's say that for some $[K] \in (S_I/F \cap S_J/F)$, $\text{sign}(v_{x_I}([K])) \neq \text{sign}(v_{x_J}([K]))$. We can use this $[K]$ to show that the signs of v_{x_I} and v_{x_J} differ on exactly half of the coordinates of $S_I/F \cap S_J/F$ and thus $\langle v_{x_I}, v_{x_J}\rangle = 0$.

Let $[K_1], [K_2], \ldots, [K_\ell]$ be a complete list of the elements in $(S_I/F \cap S_J/F)$. Now consider the list $[K_1 \circ K], [K_2 \circ K], \ldots, [K_\ell \circ K]$ which is a permutation of the first list. Because $sign([K_i \circ K]) = sign([K_i]) \cdot sign([K])$, if $\text{sign}(v_{x_I}([K_i])) = \text{sign}(v_{x_J}([K_i]))$ then $\text{sign}(v_{x_I}([K_i \circ K])) \neq \text{sign}(v_{x_J}([K_i \circ K]))$, and if $\text{sign}(v_{x_I}([K_i])) \neq \text{sign}(v_{x_J}([K_i]))$ then $\text{sign}(v_{x_I}([K_i \circ K])) = \text{sign}(v_{x_J}([K_i \circ K]))$. But because we simply permuted the coordinates, we know that the number of agreements (and respectively disagreements) in the first list are the same as the number of agreements (and respectively disagreements) in the second list. And so, the number of agreeing and disagreeing coordinates are equal.

Now assume that $\forall [K] \in (S_I/F \cap S_I/F)$, $\text{sign}(v_{x_I}([K])) = \text{sign}(v_{x_J}([K]))$. If we view S_I/F and S_J/F as subgroups of $S_{I\cup J}/F$ we see that $\frac{|S_I/F||S_J/F|}{|S_I/F \cup S_J/F|} = |S_{I\cup J}/F|$. Combining this with the fact that $|S_{I\cup J}/F| = |S_{I\cup J}/F_{I\cup J}| = |SAT_{I\cup J}|$ from Claim 15 we get that

$$\langle v_{x_I}, v_{x_J}\rangle = \frac{|S_I/F \cap S_J/F|}{|SAT_I||SAT_J|} = \frac{|S_I/F \cap S_J/F|}{|S_I/F||S_J/F|} = \frac{1}{|S_{I\cup J}/F|} = \frac{1}{|SAT_{I\cup J}|}$$

□

Equation 2, that $||v_{C_I^*}||^2 = 1$, is satisfied, because

$$||v_{C_I^*}||^2 = ||\sum_{x_I \in \mathbf{x}_I} v_{x_I}||^2 = \sum_{x_I \in \mathbf{x}_I} ||v_{x_I}||^2 = \sum_{x_I \in SAT_I} ||v_{x_I}||^2 = |SAT_I|\frac{1}{|SAT_I|}$$

Also recall, by our construction $v_0 = (1, 0, \ldots, 0)$ where the 1 is in the location indexed by $[\emptyset]$.
Equation 3 states that $\forall f_I \in \mathbf{C_I^f} ||v_{C_I^f}||^2 = 1$. By construction $||v_{x_I}||^2 = \frac{1}{|SAT_I|}$ if $x_I \in SAT_I$ (because each non-zero coordinate of v_{x_I} is $\pm\frac{1}{|SAT_I|}$ and there are $|SAT_I|$ such coordinates) and $||v_{x_I}||^2 = 0$ if $x_I \notin SAT_I$ by construction. Also, by Claim 16 v_{x_I} and $v_{x'_I}$ are orthogonal for $x_I \neq x'_I$.
To show that Equation 3 is satisfied we must show that for any clause $C_I^{\oplus I=b}$ that appears in φ, $||v_{C_I^{\oplus I=b}}||^2 = 1$. Fix some such clause $C_I^{\oplus I=b} \in \varphi$. Then $||v_{C_I^f}||^2 = \sum_{x_I \in SAT_I} ||v_{x_I}||^2 = 1$.
Equation 4 states that for all $C_I^f, C_J^g, C_{I'}^{f'}, C_{J'}^{g'}$ we have that $\langle v_{C_I^f}, v_{C_J^g}\rangle = \langle v_{C_{I'}^{f'}}, v_{C_{J'}^{g'}}\rangle$ whenever $I \cup J = I' \cup J'$ and $(C_I^f \cdot C_J^g) \equiv (C_{I'}^{f'} \cdot C_{J'}^{g'})$. To show that Equation 4 is satisfied we show that

$$\langle v_{C_I^f}, v_{C_J^g}\rangle = \frac{|SAT_{I\cup J} \cap SAT_{C_I^f} \cap SAT_{C_J^g}|}{|SAT_{I\cup J}|}$$

where $SAT_{C_I^f}$ is the subset of $\mathbf{x}_I$ that satisfies the clause C_I^f. It follows that Equation 4 is satisfied because for any $C_I^f, C_J^g, C_{I'}^{f'}, C_{J'}^{g'}$ where $I \cup J = I' \cup J'$ and $(f \cdot g) = (f' \cdot g')$, $SAT_{C_I^f} \cap SAT_{C_J^g} = SAT_{C_{I'}^{f'}} \cap SAT_{C_{J'}^{g'}}$.

$$\langle v_{C_I^f}, v_{C_J^g}\rangle = \langle \sum_{x_I \in C_I^f} v_{x_I}, \sum_{x_J \in C_J^g} v_{x_J}\rangle = \sum_{x_I \in C_I^f} \sum_{x_J \in C_J^g} \langle v_{x_I}, v_{x_J}\rangle = \frac{|SAT_{I\cup J} \cap SAT_{C_I^f} \cap SAT_{C_I^f}|}{|SAT_{I\cup J}|}$$

The last equality follows from Claim 16.
Finally, Equation 5, which states that $\forall C_I^f v_{C_I^f} = \sum_{x_I \in C_I^f} v_{x_I}$ is satisfied by construction.

□

4 Extensions

We now mention the corollaries of Theorem 12 and its proof.

Corollary 17 *For every ε, there exists some constants $\alpha \geq 0$, such that the αn level of Lasserre, an integrality gap of $\frac{7}{6} - \varepsilon$ for* VertexCover *persists.*

The idea of the proof is to rewrite a 3-XOR formula φ as a vertex cover problem on a graph G_φ using the standard FGLSS reduction. We will do it in such a way that any vectors that satisfy the Lasserre relaxation for the 3-XOR instance φ will also satisfy the vertex cover Lasserre relaxation for G_φ.

To prove this corollary, we use the following lemma which states that for a certain type of transformations most of the Lasserre constraints continue to be satisfied:

Lemma 18 *Let $\langle \mathbf{x}, \mathbf{C}^{\mathbf{f}}_{\mathbf{I}}, M\rangle$ and $\langle \bar{\mathbf{x}}, \bar{\mathbf{C}}^{\bar{\mathbf{f}}}_{\bar{\mathbf{I}}}, \bar{M}\rangle$ be two constraint maximization or minimization problems. Let g be a map from constraints on $\mathbf{x} = \{0,1\}^n$ to constraints on $\bar{\mathbf{x}}$ such that*

1) For each $i \in [n]$, $g(i) = \bar{C}^{\bar{f}}_{\bar{I}}$, where $\bar{C}^{\bar{f}}_{\bar{I}}$ is some constraint over $\bar{\mathbf{x}}$.

2) For $x_I \in \mathbf{x}_I$, $g(x_I) = \wedge_{i\in I, x_i=1} g(i) \wedge_{i\in I, x_i=0} \neg g(i)$.

3) For each C^f_I, $g(C^f_I) = \vee_{x_I \in C^f_I} g(x_I)$.

Let $k = \max |\bar{I}| : g(i) = \bar{C}^{\bar{f}}_{\bar{I}}$. Then if a collection of vectors $\{\bar{v}_{\bar{C}^{\bar{f}}_{\bar{I}}}\}_{\bar{C}^{\bar{f}}_{\bar{I}}}$ satisfy the Lasserre constraints after r rounds for $\langle \bar{\mathbf{x}}, \bar{\mathbf{C}}^{\bar{\mathbf{f}}}_{\bar{\mathbf{I}}}, \bar{M}\rangle$, then the collection of vectors $\{v_{C^f_I}\}_{C^f_I}$ where $v_{C^f_I} \equiv \bar{v}_{g(C^f_I)}$ satisfy Equations 2, 4, and 5 for $\lfloor r/k \rfloor$ rounds of Lasserre.

Proof: That we only run for $\lfloor r/k \rfloor$ rounds of Lasserre makes all the vectors well-defined. Each constraint for which we define a vector depends on at most $\lfloor r/k \rfloor$, and so the corresponding vector depends on at most r variables.

We first show Equation 2, that $\forall I \subseteq [n], ||v_{C^*_I}||^2 = 1$ is satisfied. $||v_{C^*_I}||^2 = ||\bar{v}_{g(C^*_I)}||^2 = ||\bar{v}_{\vee_{x_I \in \mathbf{x}_I} g(x_I)}||^2$. Let $\bar{\mathbf{x}}_{\bar{I}}$ be the domain of $g(x_I)$. Then $\vee_{x_I \in \mathbf{x}_I} g(x_I) = \bar{C}^*_{\bar{I}}$ because $g(x_I) = \wedge_{i\in I, x_i=1} g(i) \wedge_{i\in I, x_i=0} \neg g(i)$ and so every $\bar{x} \in \bar{\mathbf{x}}$ satisfies exactly one of these $g(x_I)$. It follows that $||\bar{v}_{\vee_{x_I \in \mathbf{x}_I} g(x_I)}||^2 = ||\bar{v}_{\bar{C}^*_{\bar{I}}}||^2 = 1$ because the vectors $\{\bar{v}_{\bar{C}^{\bar{f}}_{\bar{I}}}\}_{\bar{C}^{\bar{f}}_{\bar{I}}}$ satisfy the Lasserre constraints.

Equation 5 is satisfied by construction.

Finally, Equation 4 is satisfied. We show that $\forall C^f_I, C^g_J, C^{f'}_{I'}, C^{g'}_{J'}$ where $I \cup J = I' \cup J'$ and $(C^f_I \cdot C^g_J) \equiv (C^{f'}_{I'} \cdot C^{g'}_{J'})$ we have that

$$\langle v_{C^f_I}, v_{C^g_J}\rangle = \sum_{x_{I\cup J} \in C^f_I \cdot C^g_J} ||\bar{v}_{g(x_{I\cup J})}||^2$$

This is sufficient because the right hand side will be the same for $\langle v_{C^{f'}_{I'}}, v_{C^{g'}_{J'}}\rangle$.

$$\begin{aligned}
\langle v_{C^f_I}, v_{C^g_J}\rangle &= \sum_{x_I \in C^f_I} \sum_{x_J \in C^g_J} \langle v_{x_I}, v_{x_J}\rangle = \sum_{x_I \in C^f_I} \sum_{x_J \in C^g_J} \langle \bar{v}_{g(x_I)}, \bar{v}_{g(x_J)}\rangle \\
&= \sum_{x_I \in C^f_I} \sum_{x_J \in C^g_J} \langle \bar{v}_{\wedge_{i\in I, x_i=1} g(i) \wedge_{i\in I, x_i=0} \neg g(i)}, \bar{v}_{\wedge_{j\in J, x_j=1} g(j) \wedge_{j\in J, x_j=0} \neg g(j)}\rangle
\end{aligned}$$

Now if x_I and x_J do not agree on the overlap, then the term is 0. For if they are inconsistent, then one vector will require that $g(i)$ holds, and the other that it does not hold, and thus the dot product will be 0. So we can assume that they do agree on the overlap. But then we are simply summing over assignments to $x_{I\cup J}$ that satisfy $C^f_I \cdot C^g_J$. And thus the above can be rewritten as

$$\sum_{x_{I\cup J} \in C^f_I \cdot C^g_J} ||\bar{v}_{\wedge_{i\in I\cup J, x_i=1} g(i) \wedge_{i\in I\cup J, x_i=0} g(i)}||^2 = \sum_{x_{I\cup J} \in C^f_I \cdot C^g_J} ||\bar{v}_{g(x_{I\cup J})}||^2$$

□

We now prove Corollary 17

PROOF: [Corollary 17] Given a 3XOR instance φ with $\Delta n = m$ equation, we define the FGLSS graph G_φ of φ as follows: G_φ has $4m$ vertices, one for each equation of φ and for each assignment to the three variables that satisfies the equation. We think of each vertex i as being labeled by a partial assignment to three variables $L(i)$. Two vertices i and j are connected if and only if $L(i)$ and $L(j)$ are inconsistent. For example, for each equation, the four vertices corresponding to that equation form a clique. It is easy to see that $\mathrm{opt}(\varphi)$ is precisely the size of the largest independent set of G_φ because there is a bijection between maximal independent sets and assignment to the n variables. Note that, in particular, the independent set size of G_φ is at most $N/4$, where $N = 4m$ is the number of vertices. Thus the smallest vertex cover of G_φ is $3N/4$ (because the complement of any independent set is a vertex cover).

Now the constraints for the Lasserre hierarchy for VertexCover on this graph G_φ, are defined over the vertices of this graph. Formally, let $\mathbf{x} = \{0,1\}^{V(G_\varphi)}$. Let $\mathbf{C^f_I}$ contain the constraint $C^{x_i \vee x_j}_{\{i,j\}}$ for each edge $(i,j) \in E(G_\varphi)$, so that the constraint is satisfied if and only if at least one of the vertices incident to the edge is in the cover. Let $M = \sum_{i \in V(G_\varphi)} x_i$. Then VertexCover is the 2-constraint minimization problem $\langle \mathbf{x}, \mathbf{C^f_I}, M \rangle$. We can convert it into a Lasserre instance using Definition 5.

However, each vertex is really just an assignment to 3 variables of the 3XOR instance. Intuitively, there is a natural map from constraints on the instance of G_φ to constraints on φ. Using this intuition, we will solve the instance of φ using the vectors of Theorem 12, define each vector in the instance of G_φ to be identical to the vector assigned to the negation of the analogous constraint of φ, and then show the all the Lasserre constraints are satisfied. For vertex i define $g(i) = \neg L(i)$ to be the negation of the constraint implied by the label of vertex i. Now extend g as in Lemma 18. By Lemma 18 Equations 2, 4, and 5 for $\Omega(n)$ rounds of Lasserre.

We still must show that Equation 3 is satisfied. We must show that for each edge $(i,j) \in E(G_\varphi)$ that $||v_{i,j}||^2 + ||v_{i,\neg j}||^2 + ||v_{\neg i,j}||^2 = 1$. By Claim 7 we can simply show that $||v_{\neg i,\neg j}||^2 = 0$. Let $(i,j) \in E(G_\varphi)$, then

$$||v_{\neg i,\neg j}||^2 = ||\bar{v}_{\neg g(i),\neg g(j)}||^2 = ||\bar{v}_{L(i),L(j)}||^2 = 0$$

The first equality follows from Equation 4. The last equality is true because $L(i)$ and $L(j)$ contradict each other. We know this because i and j are joined by an edge.

Knowing that the Lasserre constraints are satisfied, we compute the objective function $\sum_{i \in V(G_\varphi)} ||v_i||^2$. Four distinct vertices were created for each of the N clauses. We show that the sum of the $||v_i||^2$ over the four vertices in any clause is always 3. Let $C^f_I \in \varphi$ be such a clause, let $i_j : 1 \leq j \leq 4$ be the four vertices corresponding to C^f_I, and let $L(i_j)$ be the label corresponding to vertex i_j. Then $\sum_{j=1}^4 \bar{v}_{L(i)} = v_0$ by Claim 6 and the fact that the vector corresponding to an unsatisfying assignment is $\vec{0}$. And so

$$\sum_{j=1}^{4} ||v_{i_j}||^2 = \sum_{j=1}^{4} ||\bar{v}_{\neg L(i_j)}||^2 = 3 \sum_{j=1}^{4} ||\bar{v}_{L(i_j)}||^2 = 3$$

However, at most $(1/2 + \varepsilon)n$ of the clauses of φ can be satisfied, and so G_φ has an independent set of at most $(\frac{1}{8} + \varepsilon)N$, and by taking the complement a vertex cover of size at most $\frac{7}{8} - \varepsilon$. We get the integrality gap of $(\frac{7}{8} + \varepsilon)N/(3N/4) = \frac{7}{6} - \varepsilon$ □

Corollary 19 *There exists some constants $\alpha \geq 0$, such at the αn level of Lasserre, an integrality gap of any constant persists for* 3-UniformHypergraphIndependentSet.

We will use the following well known proposition which is proved in the appendix for completeness:

Proposition 20 *For every $k \geq 3, \varepsilon > 0$, there exists $\delta > 0$, such that if H is a random k-uniform hypergraph with Δn edges, where $\Delta \geq \delta$, then with probability $1 - o(1)$, H has no independent set of size εn, and, equivalently, H has no vertex cover of size $(1 - \varepsilon)n$.*

Proof: Let $\varepsilon = \frac{1}{2c}$ and let Δ be such that $(1-\varepsilon)^{\Delta} \leq \frac{\varepsilon}{e}$. Let H be a random uniform hypergraph with Δn edges. By proposition 20 we know that with high probability H has no independent set of size εn. We now must show that there exists a good solution to the Lasserre relaxation.

We note that the CSP instance is $\langle \mathbf{x}, \mathbf{C}^{\mathbf{f}}_{\mathbf{I}}, M \rangle$ where $\mathbf{x} = \{0,1\}^{V(H)}$, $M = \sum_{i \in V} x_i$, and for each edge $(v_1, \ldots, v_k) \in E(H)$ we add the constraint $\vee_{i=1}^{k} \neg x_i$ to $\mathbf{C}^{\mathbf{f}}_{\mathbf{I}}$ which we can transform into a Lasserre relaxation according to Definition 5. Note that any constraint of the form $\vee_{i=1}^{k} \neg x_i$ is implied by either $\oplus_{i=1}^{k} x_i = 1$ if k is even or $\oplus_{i=1}^{k} x_i = 0$ if k is odd. Consider then the k-XOR formula φ_H with Δn clauses which implies $\mathbf{C}^{\mathbf{f}}_{\mathbf{I}}$. We see that by Theorem 11 that φ cannot be disproved by width $\Omega(n)$ resolution and no single variable can be fixed. Even though the k-XOR clauses are not truly random because the constants are all the same, the theorem still applies. By Theorem 12 φ cannot be disproved by $\Omega(n)$ levels of Lasserre. Moreover by Remark 1 we have that $||v_i||^2 = 1/2$ for all i. Thus $M = \sum ||v_i||^2 = n/2$.

So the ratio of the Lasserre optimum to the actual optimum is $\frac{n/2}{\varepsilon n} = c$. □

Corollary 21 *There exists some constants $\alpha \geq 0$, such at the αn level of Lasserre, an integrality gap of 2 persists for* 3-UniformHypergraphVertexCover.

The proof of Corollary 19 is very similar to that of Corollary 21

Proof: Let $\varepsilon = \frac{1}{2c}$ and let Δ be such that $(1-\varepsilon)^{\Delta} \leq \frac{\varepsilon}{e}$. Let H be a random uniform hypergraph with Δn edges. By proposition 20 we know that with high probability H has no vertex cover of size $(1-\varepsilon)n$. We now must show that there exists a good solution to the Lasserre relaxation.

We note that the CSP instance is $\langle \mathbf{x}, \mathbf{C}^{\mathbf{f}}_{\mathbf{I}}, M \rangle$ where $\mathbf{x} = \{0,1\}^{V(H)}$, $M = \sum_{i \in V} x_i$, and for each edge $(v_1, \ldots, v_k) \in E(H)$ we add the constraint $\vee_{i=1}^{k} x_i$ to $\mathbf{C}^{\mathbf{f}}_{\mathbf{I}}$ which we can transform into a Lasserre relaxation according to Definition 5. Note that any constraint of the form $\vee_{i=1}^{k} x_i$ is implied by $\oplus_{i=1}^{k} x_i = 1$. Consider then the k-XOR formula φ_H with Δn clauses which implies $\mathbf{C}^{\mathbf{f}}_{\mathbf{I}}$. We see that by Theorem 11 that φ cannot be disproved by width $\Omega(n)$ resolution and no single variable can be fixed. Even though the k-XOR clauses are not truly random because the constants are all the same, the theorem still applies. By Theorem 12 φ cannot be disproved by $\Omega(n)$ levels of Lasserre. Moreover by Remark 1 we have that $||v_i||^2 = 1/2$ for all i. Thus $M = \sum ||v_i||^2 = n/2$.

So the ratio of the actual optimum to the Lasserre optimum is $\frac{(1-\varepsilon)n}{n/2} = \frac{1-\varepsilon}{2}$. □

5 Acknowledgements

I would like to both acknowledge and thank Alexandra Kolla and Madhur Tulsiani for many fruitful discussions without which this work would not have been completed. Also, I would like to thank Moses Charikar, Eden Chlamtac, Uriel Feige, and Avi Wigderson for helpful thoughts and pointers. Finally, I would like to thank the anonymous FOCS reviewers for helpful comments.

6 Conclusion

We have shown the first known integrality gaps for Lasserre. On the one hand you can see the main theorem (Theorem 12) as showing gaps for problems that are already known or thought to be **NP**-hard. We say that a predicate A is *approximation resistant* if, given a constraint satisfaction problem over A predicates, it is **NP**-hard to approximate the fraction of such predicates which can be simultaneously satisfied better than the trivial algorithm which randomly guesses an assignment and returns the fraction of predicates it satisfies. In [Zwi98], Zwick shows that the only 3-CPSs which are approximation resistant are exactly those which are implied by parity or its negation. So, for $k = 3$, the main theorem applies exactly to those problems which we already know are **NP**-hard.

On the other hand, the main theorem applies to results that are known to be in **P**. Deciding if a k-XOR formula is satisfiable is equivalent to solving a set of linear equations over $\mathbb{F}_2$, which can be done with Gaussian elimination.

References

[AAT05] 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. 1, 2, 3

[ABL02] Sanjeev Arora, Béla Bollobás, and László Lovász. Proving integrality gaps without knowing the linear program. In *Proceedings of the 43rd IEEE Symposium on Foundations of Computer Science*, pages 313–322, 2002. 1, 2

[ABLT06] Sanjeev Arora, Béla Bollobás, László Lovász, and Iannis Tourlakis. Proving integrality gaps without knowing the linear program. *Theory of Computing*, 2(2):19–51, 2006. 1, 2

[BOGH+03] 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. 1, 2

[BSW01] Eli Ben-Sasson and Avi Wigderson. Short proofs are narrow-resolution made simple. *J. ACM*, 48(2):149–169, 2001. 7

[Cha02] Moses Charikar. On semidefinite programming relaxations for graph coloring and Vertex Cover. In *Proceedings of the 13th ACM-SIAM Symposium on Discrete Algorithms*, pages 616–620, 2002. 2, 3

[Chl07] Eden Chlamtac. Approximation algorithms using hierarchies of semidefinite programming relaxations. In *FOCS '07: Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, pages 691–701, Washington, DC, USA, 2007. IEEE Computer Society. 2

[CMM07] Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Integrality gaps for Sherali-Adams relaxations, 2007. manuscript. 1, 2 [dlVKM07] Wenceslas Fernandez de la Vega and Claire Kenyon-Mathieu. Linear programming relaxations of maxcut. In *SODA '07: Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 53–61, Philadelphia, PA, USA, 2007. Society for Industrial and Applied Mathematics. 1

[FO06] Uriel Feige and Eran Ofek. Random 3CNF formulas elude the Lovász theta function. Manuscript, 2006. 2, 3

[GMPT06] Konstantinos Georgiou, Avner Magen, Toniann Pitassi, and Iannis Tourlakis. Tight integrality gaps for Vertex Cover SDPs in the Lovasz-Schrijver hierarchy. Technical Report TR06-152, Electronic Colloquium on Computational Complexity, 2006. 1, 2

[HMM06] Hamed Hatami, Avner Magen, and Vangelis Markakis. Integrality gaps of semidefinite programs for Vertex Cover and relations to ℓ_1 embeddability of negative type metrics. Manuscript, 2006. 3

[KG98] Jon M. Kleinberg and Michel X. Goemans. The Lovász Theta function and a semidefinite programming relaxation of Vertex Cover. *SIAM Journal on Discrete Mathematics*, 11:196–204, 1998. 2

[Las01] Jean B. Lasserre. An explicit exact sdp relaxation for nonlinear 0-1 programs. In *Proceedings of the 8th International IPCO Conference on Integer Programming and Combinatorial Optimization*, pages 293–303, London, UK, 2001. Springer-Verlag. 1

[Lau03] Monique Laurent. A comparison of the sherali-adams, lovász-schrijver, and lasserre relaxations for 0–1 programming. *Math. Oper. Res.*, 28(3):470–496, 2003. 1

[LS91] L. Lovasz and A. Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM J. on Optimization*, 1(12):166–190, 1991. 1

[SA90] Hanif D. Sherali and Warren P. Adams. A hierarchy of relaxation between the continuous and convex hull representations. *SIAM J. Discret. Math.*, 3(3):411–430, 1990. 1

[STT07a] Grant Schoenebeck, Luca Trevisan, and Madhur Tulsiani. A linear round lower bound for Lovasz-Schrijver SDP relaxations of Vertex Cover. In *Proceedings of the 39th ACM Symposium on Theory of Computing*, 2007. Earlier version appeared as Technical Report TR06-098 on Electronic Colloquium on Computational Complexity. 1, 2, 3

[STT07b] Grant Schoenebeck, Luca Trevisan, and Madhur Tulsiani. Tight integrality gaps for Lovasz-Schrijver LP relaxations of Vertex Cover and Max Cut. In *Proceedings of the 39th ACM Symposium on Theory of Computing*, 2007. Earlier version appeared as Technical Report TR06-132 on Electronic Colloquium on Computational Complexity. 1, 2

[Tou06] Iannis Tourlakis. New lower bounds for Vertex Cover in the Lovasz-Schrijver hierarchy. In *Proceedings of the 21st IEEE Conference on Computational Complexity*, 2006. 1, 2

[Zwi98] Uri Zwick. Approximation algorithms for constraint satisfaction problems involving at most three variables per constraint. In *SODA '98: Proceedings of the ninth annual ACM-SIAM symposium on Discrete algorithms*, pages 201–210, Philadelphia, PA, USA, 1998. Society for Industrial and Applied Mathematics. 15

7 Appendix

Proposition 22 *For any $\delta > 0$, with probability $1 - o(1)$, if φ is a random k-CSP chosen from the distribution $\mathcal{D}$ with Δn clauses where $\Delta \geq \frac{\ln 2}{2\delta^2} + 1$, at most a $r(\mathcal{D}) + \delta$ fraction of the clauses of φ can be simultaneously satisfied.*

Proof: Fix an assignment to n variables. Now if we choose, $m = \Delta n$ clauses at random, the probability that more than a $r(\mathcal{D}) + \delta$ fraction of them are satisfied is at most $\exp(-2\delta^2 m) = \exp(-2\delta^2 \Delta n)$. To get this, we use the Chernoff Bound that says

$$\Pr[X \geq \mathbb{E}[X] + \lambda] \leq \exp(-2\lambda^2/m)$$

where X is the number of satisfied clauses, $\mathbb{E}[X] = r(\mathcal{D})m$, $\lambda = \delta m$. Picking a random formula and random assignment, the probability that more than a $r(\mathcal{D}) + \delta$ fraction of the clauses are satisfied is $\exp(-2\delta^2 \Delta n)$. Taking a union bound over all assignments, we get

$$\begin{aligned}\Pr[\text{any assignment satisfies} \geq (1/2 + \delta)m \text{ clauses}] &\leq \exp(-2\delta^2 \Delta n) \cdot 2^n \\ &= \exp(n(\ln 2 - 2\delta^2 \Delta) = \exp(-2\delta^2 n)\end{aligned}$$

because $\Delta \geq \frac{\ln 2}{2\delta^2} + 1$. $\Box$

Theorem 23 *For $k \geq 3$, $d > 0$, $\gamma > 0$, and $0 \leq \varepsilon < k/2 - 1$, if φ is a random k-XOR formula with density dn^ε, then with probability $1 - o(1)$ φ cannot be disproved by width $\alpha n^{1 - \frac{\varepsilon}{k/2 - \gamma - 1}}$ resolution nor can any variable be resolved to true or false. Furthermore, this is true even if the parity sign (whether the predicate is parity or its negation) of each clause is adversatively chosen.*

We use the following Proposition:

Proposition 24 *For any $k \geq 3$, $d > 0$, $\gamma > 0$, and $0 \leq \varepsilon < k/2 - 1$, there exists $\beta > 0$ such that if φ is a random k-XOR formula with density dn^ε then with probability $1 - o(1)$:*

1. *Every subformula $\varphi' \subseteq \varphi$ where $|\varphi'| \leq \beta n^{1 - \frac{\varepsilon}{k/2 - 1}}$ is satisfiable even after fixing one variable.*
2. *For every subformula $\varphi' \subseteq \varphi$ where $|\varphi'| \in [\frac{1}{3}\beta n^{1 - \frac{\varepsilon}{k/2 - \gamma - 1}}, \frac{2}{3}\beta n^{1 - \frac{\varepsilon}{k/2 - \gamma - 1}}]$, we have that $2V(\varphi') - k|\varphi'| \geq 2\gamma|\varphi'|$ where $V(\varphi')$ is the number of variables in φ'.*

Proof:[Theorem 23] Let φ be a random XOR formula as in the theorem statement and let C be any clause over the variables of φ. We define $\mu(C)$ to be the smallest size of a subformula $\varphi' \subseteq \varphi$ such that we can start from φ' and imply C using resolution. We note that in any resolution tree, if C_1 and C_2 together imply C_3, then $\mu(C_1) + \mu(C_2) \geq \mu(C_3)$.

From the first part of Proposition 24 we know that with high probability $\mu(0 = 1) \geq \beta n^{1 - \frac{\varepsilon}{k/2 - 1}}$.

Now consider a resolution tree that derives $0 = 1$, that is, a contradiction. We will show that this tree must contain a clause C with many variables. By the subadditivity of μ as we move up the resolution tree, this tree must contain some clause C such that $\mu(C) \in [\frac{1}{3}\beta n^{1 - \frac{\varepsilon}{k/2 - \gamma - 1}}, \frac{2}{3}\beta n^{1 - \frac{\varepsilon}{k/2 - \gamma - 1}}]$.

We will now show that with high probability C contains $\frac{\gamma\beta}{6} n^{1 - \frac{\varepsilon}{k/2 - \gamma - 1}}$ variables and thus that the width of the resolution is at least as large. Let φ' be a subformula of size $\mu(C)$ which implies C. By the second part of Proposition 24 we know that $2V(\varphi') - k|\varphi'| \geq \gamma|\varphi'|$. Each variable of φ' must appear either in two of the clauses of φ' or in C itself. If a variable appears in one clause, but not in C; then no matter what the value of the other variables of that clause, the clause could still be satisfied by flipping this one variable. Therefore this clause can always be satisfied independently of the rest of φ' and is not required to imply C. This violates minimality of φ'. So

$$|C| + \frac{k}{2}|\varphi'| \geq V(\varphi') \Rightarrow |C| \geq \frac{1}{2}(2V(\varphi') - k|\varphi'|) \geq \gamma|\varphi'| \geq \frac{\gamma\beta}{3} n^{1-\frac{\varepsilon}{k/2-\gamma-1}}$$

so let $\alpha = \frac{\gamma\beta}{3}$.

To show that you cannot fix one variable to true or false the proof is almost exactly the same. Instead of showing that $\mu(0 = 1)$ is large, we show that for any x_i, $\mu(x_i = 0)$ and $\mu(x_i = 1)$ are large. This also follows from the first part of Proposition 24.

We note that we never used the parity of individual clauses in the proof, only the variables contained in each clause. Therefore the theorem still applies even if the parity of each clause is adversarially chosen. □

Proof:[Proposition 24] First we bound the probability that for a random formula φ, there exists a set of ℓ clauses containing a total of fewer than $c\ell$ variables by $(O(1)\frac{\ell^{k-c-1}}{n^{k-c-1-\varepsilon}})^{\ell}$;

We can upper bound the probability that there is a set of ℓ clauses containing a total of fewer than $c\ell$ variables by

$$\binom{n}{c\ell} \cdot \binom{\binom{c\ell}{k}}{\ell} \cdot l! \cdot \binom{m}{\ell} \cdot \binom{n}{k}^{-\ell}$$

where $\binom{n}{c\ell}$ is the choice of the variables, $\binom{\binom{c\ell}{k}}{\ell}$ is the choice of the ℓ clauses constructed out of such variables, $\ell! \cdot \binom{m}{\ell}$ is a choice of where to put such clauses in our ordered sequence of m clauses, and $\binom{n}{k}^{-\ell}$ is the probability that such clauses were generated as prescribed.

Using $\binom{N}{K} < (eN/K)^K$, $k! < k^k$, and $m = n^{1+\varepsilon}$ we simplify to obtain the upper bound $\left(O\left(\frac{\ell^{k-c-1}}{n^{k-c-1-\varepsilon}}\right)\right)^{\ell}$.

We first show that the first part of the proposition is true if we do not fix any variables. If $\varphi' \subseteq \varphi$ is a minimal unsatisfiable subformula of φ, then each variable that appears in φ' must occur twice in φ'. Otherwise the clause in which that variable appears is always satisfiable and φ' is not a minimal unsatisfiable subformula. Thus it is sufficient to show that no set of ℓ clauses contains fewer than $\frac{k}{2}\ell$ variables. We will show that if we set $c = k/2$ in the above formula, the sum over ℓ from 1 to $\beta n^{1-\frac{\varepsilon}{\frac{k}{2}-1}}$, can be made $o(1)$ with a sufficiently small β.

$$\sum_{\ell=1}^{\beta n^{1-\frac{\varepsilon}{\frac{k}{2}-1}}} \left(O\left(\frac{\ell^{\frac{k}{2}-1}}{n^{\frac{k}{2}-1-\varepsilon}}\right)\right)^{\ell}$$

Let δ be a sufficiently small constant, and let $\omega(n)$ be some function that grows in an unbounded fashion. We break up the above sum into:

$$\sum_{\ell=1}^{\delta n^{1-\frac{\varepsilon}{\frac{k}{2}-1}}\omega(n)^{-1}} \left(O\left(\frac{\ell^{\frac{k}{2}-1}}{n^{\frac{k}{2}-1-\varepsilon}}\right)\right)^{\ell} + \sum_{\ell=\delta n^{1-\frac{\varepsilon}{\frac{k}{2}-1}}\omega(n)^{-1}+1}^{\beta n^{1-\frac{\varepsilon}{\frac{k}{2}-1}}} \left(O\left(\frac{\ell^{\frac{k}{2}-1}}{n^{\frac{k}{2}-1-\varepsilon}}\right)\right)^{\ell}$$

We then bound each of these terms:

$$\sum_{\ell=1}^{\delta n^{(1-\frac{\varepsilon}{\frac{k}{2}-1})}\omega(n)^{-1}} \left(O\left(\frac{\ell^{\frac{k}{2}-1}}{n^{\frac{k}{2}-1-\varepsilon}}\right)\right)^{\ell} \leq \sum_{\ell=1}^{\infty} \left(O(1)(\delta\omega(n)^{-1})^{k-c-1}\right)^{\ell} = o(1)$$

for sufficiently small δ and sufficiently large n.

$$\begin{aligned}
\sum_{\ell=\delta n^{1-\frac{\varepsilon}{\frac{k}{2}-1}}\omega(n)^{-1}+1}^{\beta n^{1-\frac{\varepsilon}{\frac{k}{2}-1}}} \left(O\left(\frac{\ell^{\frac{k}{2}-1}}{n^{\frac{k}{2}-1-\varepsilon}}\right)\right)^{\ell} &\leq \sum_{\ell=\delta n^{1-\frac{\varepsilon}{\frac{k}{2}-1}}\omega(n)^{-1}+1}^{\infty} \left(O(1)\beta^{k-c-1}\right)^{\ell} \\
&\leq \beta^{\delta n^{1-\frac{\varepsilon}{\frac{k}{2}-1}}\omega(n)^{-1}} \sum_{\ell=1}^{\infty} \left(O(1)\beta^{k-c-1}\right)^{\ell} = o(1)
\end{aligned}$$

for sufficiently small β and sufficiently slowly growing $\omega(n)$.

Now we note that small subformulas are satisfiable even if we fix one variable. We can use all the above machinery, but now require that every set of ℓ clauses contains $\frac{k}{2}+1$ variables. However, this change is absorbed into the O constant in $\left(O\left(\frac{\ell^{k-c-1}}{n^{k-c-1-\varepsilon}}\right)\right)^{\ell}$ because in the above analysis when changing to $\binom{n}{c\ell-1}\cdot\binom{\binom{c\ell-1}{k}}{\ell}\cdot l!\cdot\binom{m}{\ell}\cdot\binom{n}{k}^{-\ell}$ we only get an addition factor of $\frac{c\ell-1}{ne}\left(\frac{c\ell}{c\ell-1}\right)^{k}$ the first factor helps and the second is bounded by 2^k which is a constant.

Now we show the second part of the Proposition.

We saw above that we can bound the probability that there exists a subformula of size ℓ that fails to satisfy $2V(\varphi')-k|\varphi'|\geq 2\gamma|\varphi'|$ by $\left(O\left(\frac{\ell^{\frac{k}{2}-\gamma-1}}{n^{\frac{k}{2}-\gamma-1-\varepsilon}}\right)\right)^{\ell}$. We will fix β later, and now use a union bound to upper bound the probability that there exists a clause φ' such that $|\varphi'|\in[\frac{1}{3}\beta n^{1-\frac{\varepsilon}{k/2-\gamma-1}},\frac{2}{3}\beta n^{1-\frac{\varepsilon}{k/2-\gamma-1}}]$ and $|V(\varphi')|\leq(\frac{k}{2}+\gamma)|\varphi'|$.

$$\begin{aligned}
&\sum_{\ell=\frac{1}{3}\beta n^{1-\frac{\varepsilon}{k/2-\gamma-1}}}^{\frac{1}{3}\beta n^{1-\frac{\varepsilon}{k/2-\gamma-1}}} \left(O\left(\frac{\ell^{\frac{k}{2}-\gamma-1}}{n^{\frac{k}{2}-\gamma-1-\varepsilon}}\right)\right)^{\ell} \\
\leq\ & \left(\frac{1}{3}\beta n^{1-\frac{\varepsilon}{k/2-\gamma-1}}\right)\left(O\left(\frac{\left(\frac{2}{3}\beta n^{1-\frac{\varepsilon}{k/2-\gamma-1}}\right)^{\frac{k}{2}-\gamma-1}}{n^{\frac{k}{2}-\gamma-1-\varepsilon}}\right)\right)^{\left(\frac{1}{3}\beta n^{1-\frac{\varepsilon}{k/2-\gamma-1}}\right)} \\
\leq\ & \left(\frac{1}{3}\beta n^{1-\frac{\varepsilon}{k/2-\gamma-1}}\right)\left(O(\frac{2}{3}\beta)^{k/2-\gamma-1}\right)^{\left(\frac{1}{3}\beta n^{1-\frac{\varepsilon}{k/2-\gamma-1}}\right)} \\
\leq\ & \left(\frac{1}{3}\beta n^{1-\frac{\varepsilon}{k/2-\gamma-1}}\right)(\frac{1}{2})^{\left(\frac{1}{3}\beta n^{1-\frac{\varepsilon}{k/2-\gamma-1}}\right)} = o(1)
\end{aligned}$$

for a sufficiently small choice of β

□

Proposition 25 *For every $k \geq 3, \varepsilon > 0$, there exists $\delta > 0$, such that if H is a random k-uniform hypergraph with Δn edges, where $\Delta \geq \delta$, then with probability $1 - o(1)$, H has no independent set of size εn, and, equivalently, H has no vertex cover of size $(1 - \varepsilon)n$.*

PROOF:

Let δ be such that $(1-\varepsilon)^{\delta} < \frac{\varepsilon}{e}$. Then the probability that H has an independent set of size εn (or has a vertex cover of size $(1-\varepsilon)n$) is bounded by the probability that there is a set of size εn such that no edge contains only vertices from this set:

$$\binom{n}{\varepsilon n}(1-\varepsilon)^{\Delta n} \leq \left(\frac{e}{\varepsilon}\right)^{\varepsilon n}(1-\varepsilon)^{\delta n} \leq \left(\frac{e}{\varepsilon}\right)^{\varepsilon n}\left(\frac{\varepsilon}{e}\right)^{n} = \left(\frac{\varepsilon}{e}\right)^{(1-\varepsilon)n} = o(1)$$

□