AN 83/84 APPROXIMATION FOR THREE-OCCURRENCE HORN FORMULAS ABOVE THE UNIQUE GAMES THRESHOLDS

FRANK VEGA

hexagon:2611.00827v2

Abstract. We study an unweighted MAX-HORN-2SAT class, which is also a class of MAX 2-SAT instances, in which every Boolean variable has exactly three literal occurrences. For a simple cubic graph $H = (V, E)$ the formula T_H satisfies $\mathrm{OPT}(T_H) = 4|V| + |E| + \alpha(H)$, and every assignment of value A is repaired in polynomial time to an independent set of size at least $A - (4|V| + |E|)$. This gives a polynomial-time 83/84 approximation from a cubic independent-set approximation and a near-bipartite cut algorithm, an L-reduction from cubic MIS, and APX-completeness. For simple d-regular graphs and every $d \geq 1$, the optimum is $(d + 1)|V| + |E| + \alpha(H)$ and the same constructive repair holds: for $d \in \{1, 2\}$ (and more generally for graphs of maximum degree at most 2), $\mathrm{OPT}(T_H)$ is computed exactly in polynomial time (factor $1 > 83/84$), whereas for $d \geq 3$ the problem is APX-complete and the factors are $(3d + 2)/(3d + 3)$ from the all-true assignment, which is at least 83/84 for $d \geq 27$; 83/84 from the Berman–Fujito guarantee for $d \geq 22$ and, with a semidefinite cut rounding, for $d = 21$; $1 - \frac{d-2+\varepsilon}{(3d+3)(d+3+\varepsilon)}$ for every fixed $\varepsilon > 0$ and $4 \leq d \leq 20$ (exceeding $(3d+2)/(3d+3)$); and $(21d+20)/(21d+21)$ from any 6/7-approximate independent set. Consequently, for every $d \geq 1$ the d-regular class has a polynomial-time ratio at least 0.978 (the smallest limiting factor is $185/189 \approx 0.97884$, at $d = 6$), which exceeds the Unique Games thresholds $\alpha^* \approx 0.94616$ and $\beta^\star \approx 0.94017$. For arbitrary simple graphs the optimum is $n + 3m + \alpha(H)$. The central result is that for every $d \geq 1$, the conflict-graph formula $T_{H(G)}$ obtained from a simple MAX-CUT instance G of maximum degree at most d with at least one edge admits a polynomial-time approximation of ratio at least $0.9932 > 83/84$ via the Goemans–Williamson semidefinite relaxation; the guarantee holds for every graph G, and the conflict graph $H(G)$ has maximum degree at most $d + 1$. Since $83/84 \approx 0.9881$ exceeds the Unique Games thresholds $\alpha^* \approx 0.94616$ for weighted MAX-HORN-2SAT and $\beta^\star \approx 0.94017$ for MAX 2-SAT, the UGC barrier is broken on the bounded-occurrence class $\{T_{H(G)} : G \text{ has maximum degree at most } d\}$. While exceeding these thresholds on *all* instances would imply P = NP under the Unique Games Conjecture, our result is restricted to bounded-occurrence instances and is therefore fully consistent with the conjecture.

1. Introduction

MAX-HORN-2SAT maximizes the number of satisfied unary and binary Horn clauses. A Horn clause has at most one positive literal. The formula T_H below is written with the opposite polarity: each of its clauses has at most one negative literal. Renaming every variable to its negation turns every clause of T_H into a Horn clause, and an assignment satisfies a clause exactly when the complemented assignment satisfies the complemented clause, so all objective values are unchanged. The formula T_H is therefore a MAX-HORN-2SAT instance up to this renaming, and the Lean development proves both statements. Every clause has at most two literals, so each formula is also an instance of MAX 2-SAT, and every guarantee proved below for the Horn class is a guarantee for this class of MAX 2-SAT instances, for cubic graphs and d-regular graphs alike. The construction below starts from cubic Maximum Independent Set (MIS), uses four clones for every graph vertex, and connects them by a directed cycle of implications. One clone participates in a unit clause and the other three in the three incident edge clauses. Each clone additionally occurs in its incoming and outgoing implication. Consequently every formula variable occurs exactly three times. The reduction from cubic MIS to MAX-HORN-2SAT composes unit, edge, and consistency clauses; the quantitative contribution is the exact optimum identity and the explicit cardinality bound of the repair, together with their Lean formalization. The contribution is quantitative rather than merely satisfiability preserving. We prove an explicit repair of every inconsistent clone assignment, obtain an exact affine optimum identity, and transfer the Berman–Fujito cubic-MIS guarantee [2] to the Horn class. Cubic MIS

Date: November 23, 2026.

2020 *Mathematics Subject Classification.* 68Q25, 68W25, 90C27.

Key words and phrases. approximate computing; MAX-SAT; independent set problem; Lean 4; NP-hard; satisfiability; Unique Games Conjecture.

is APX-complete [1], and the same repair proves APX-completeness of the image class. The central result is Theorem 18: for every $d \geq 1$, the conflict-graph formula $T_{H(G)}$ obtained from a simple MAX-CUT instance G of maximum degree at most d with at least one edge admits a polynomial-time approximation of ratio at least $0.9932 > 83/84$. Since $83/84$ exceeds the Unique Games thresholds $\alpha^* \approx 0.94616$ and $\beta^\star \approx 0.94017$, the UGC barrier is broken on the bounded-occurrence class $\{T_{H(G)} : G \text{ has maximum degree at most } d\}$.

2. The cubic three-occurrence construction

Let $H = (V, E)$ be a finite simple cubic graph, with $n = |V|$ and $m = |E|$. Thus $m = 3n/2$. At each vertex v, label its incident edges with the three ports $1, 2, 3$. Introduce four Boolean variables

$$z_{v,0}, z_{v,1}, z_{v,2}, z_{v,3}.$$

The formula T_H contains the following clauses.

(1) For every vertex v, one unit clause $(\neg z_{v,0})$.
(2) For every edge $e = uv$, whose endpoint ports are p at u and q at v, one edge clause $(z_{u,p} \vee z_{v,q})$.
(3) For every vertex v and $i \in \mathbb{Z}/4\mathbb{Z}$, one consistency implication $(\neg z_{v,i} \vee z_{v,i+1})$.

Every clause has at most one negative literal. Every clone occurs in exactly one source clause, one outgoing implication, and one incoming implication; hence it occurs exactly three times. There are $4n$ variables and $n + m + 4n = 5n + m$ clauses. For an assignment z, let

$$\begin{aligned} U(z) &= |\{v : z_{v,0} = 0\}|, \\ F(z) &= |\{uv \in E : z_{u,p} = z_{v,q} = 0\}|, \\ T(z) &= |\{(v, i) : z_{v,i} = 1,\ z_{v,i+1} = 0\}|. \end{aligned}$$

These are respectively the satisfied unit clauses, falsified edge clauses, and falsified consistency implications. Therefore

$$\mathrm{val}_{T_H}(z) = U(z) + (m - F(z)) + (4n - T(z)). \tag{1}$$

3. Constructive repair and exact optimum

Call a vertex selected when $z_{v,0} = 0$, and clean when it is selected and its four-bit cycle has no 1-to-0 transition.

Lemma 1 (Cycle lemma)**.** *If $z_{v,0} = 0$ and the cycle at v has no 1-to-0 transition, then $z_{v,i} = 0$ for all i.*

Proof. A nonconstant cyclic binary word has both a 0-to-1 and a 1-to-0 transition. Since the word begins with zero, absence of a 1-to-0 transition forces it to be constantly zero. □

Let $C(z)$ be the clean selected vertices. Delete the left endpoint of every edge induced by $C(z)$, ignoring repetitions, and call the remaining set $R(z)$.

Lemma 2 (Quantitative repair)**.** *$R(z)$ is independent and*

$$|R(z)| \geq \mathrm{val}_{T_H}(z) - (4n + m).$$

The set is computable in polynomial time.

Proof. Every edge induced by $C(z)$ loses an endpoint, so $R(z)$ is independent. At most $T(z)$ selected vertices are nonclean, because each such vertex contributes at least one transition. By Lemma 1, both endpoint clones of every edge induced by $C(z)$ are zero; hence every such edge is counted by $F(z)$. The number of distinct removed endpoints is at most the number of edges induced by $C(z)$, which is at most $F(z)$. Thus

$$|R(z)| \geq U(z) - T(z) - F(z).$$

Substituting (1) gives the claimed inequality. □

Theorem 3 (Exact identity)**.** *For every simple cubic graph H,*

$$\mathrm{OPT}(T_H) = 4n + m + \alpha(H).$$

Moreover, an assignment of value A decodes in polynomial time to an independent set of size at least $A - (4n + m)$.

Proof. The upper bound follows from Lemma 2. For the lower bound, let I be an independent set and set all four clones of v to zero exactly when $v \in I$. Every consistency implication and every edge clause is satisfied, while precisely $|I|$ unit clauses are satisfied. The value is $4n+m+|I|$. Taking I maximum proves equality. □

4. The 83/84 approximation

For a simple cubic graph put $B = 4n+m = 11n/2$. Every independent set sends all its incident edges to its complement, so $\alpha \le n/2$. If $|I| \ge 6\alpha/7$, the ratio $(B+|I|)/(B+\alpha)$ is at least $(11+6/7)/12 = 83/84$. We give an endpoint argument that only requires the published cubic MIS factor arbitrarily close to 7/6 [2].

Theorem 4. *On the class $\{T_H : H$ is simple and cubic$\}$ there is a polynomial-time 83/84 approximation.*

Proof. Return the better lift of two independent sets. First use Berman–Fujito with fixed parameters giving solution fraction $r = 6/7-10^{-6}$. If $2\alpha/n \le 9999/10000$, then $(83-84r)\alpha \le (11+84\cdot 10^{-6})(9999/20000)n \le 11n/2 = B$, which gives the required ratio. For the second set, apply the near-satisfiable Max Cut guarantee of Trevisan [6]. For fixed $\epsilon = 10^{-4}$, a graph having a cut of fraction at least $1-\epsilon$ admits a polynomially computable cut of fraction at least 0.95 for all sufficiently large n: the stated bound is $1 - 4\sqrt{\epsilon} + 8\epsilon - o(1) > 0.95$. Fixed small orders may be solved exhaustively, including the empty graph. The independent set of size α and its complement cut 3α edges, a fraction $2\alpha/n$. Thus the cut guarantee applies in the remaining case. From each side delete one endpoint of each internal edge, taking the union of the deletions on that side. The larger remaining independent set has size at least $(n-0.05m)/2 = 37n/80 > 3n/7 \ge 6\alpha/7$. The endpoint calculation above now applies. The algorithm runs both branches without knowing α. □

4.1. **Generalization to all degrees $d \ge 1$.** The same construction and ratio transfer extend to every fixed degree $d \ge 1$. Let $H = (V, E)$ be simple and d-regular, so $m = dn/2$. At each vertex introduce $d+1$ Boolean clones together with one unit clause, d edge clauses (via a port labelling of the incident edges), and a directed cycle of $d+1$ consistency implications. Every clone still occurs in exactly three places. The score identity becomes

$$\mathrm{OPT}(T_H) = (d+1)n + m + \alpha(H).$$

The quantitative repair of Lemma 2 is independent of the cycle length $d+1 \ge 2$, and the same deletion argument yields an independent set of size at least $A - ((d+1)n+m)$. Write

$$B = (d+1)n + m = n\Big(\tfrac{3d}{2}+1\Big), \qquad \beta = \tfrac{3d}{2}+1.$$

For every $d \ge 1$, any independent set cuts $d\alpha$ of the $m = dn/2$ edges, hence $\alpha \le n/2$. Consequently any independent set of size at least $6\alpha/7$ already produces the ratio

$$\frac{B+|I|}{B+\alpha} \ge \frac{\beta+3/7}{\beta+1/2}.$$

The right-hand side is strictly increasing in β for $\beta > 0$ and equals 83/84 when $d = 3$. Therefore the ratio is at least 83/84 for every $d \ge 3$.

For $d \in \{1, 2\}$ (and more generally for any simple graph H of maximum degree at most 2), Maximum Independent Set is solvable *exactly* in linear time, which yields an exact polynomial-time algorithm for T_H.

Proposition 5. *Let $H = (V, E)$ be a finite simple graph of maximum degree at most 2. Then $\mathrm{OPT}(T_H)$ and an optimal assignment of T_H can be computed in $O(|V|)$ time (an approximation ratio of $1 > 83/84$). In particular:*

1. *If H is simple and 1-regular (a disjoint union of $m = n/2$ edges), then $\alpha(H) = n/2$ and $\mathrm{OPT}(T_H) = 3n$.*
2. *If H is simple and 2-regular (a disjoint union of simple cycles $C_{\ell_1} \sqcup \cdots \sqcup C_{\ell_k}$ with $\sum_{j=1}^{k} \ell_j = n$), then $\alpha(H) = \sum_{j=1}^{k} \lfloor \ell_j/2 \rfloor$ and $\mathrm{OPT}(T_H) = 4n + \sum_{j=1}^{k} \lfloor \ell_j/2 \rfloor$.*

Proof. Every connected component of a simple graph of maximum degree at most 2 is a path or a cycle, so a maximum independent set I of H is found in $O(|V|)$ time by greedy traversal along paths and cycles. Lifting I via Lemma 15 (setting $z_{v,i} = 0$ for $v \in I$ and $z_{v,i} = 1$ otherwise) gives an assignment of value $n + 3m + \alpha(H) = \mathrm{OPT}(T_H)$. For 1-regular graphs $m = n/2$ and $\alpha(H) = n/2$, giving $\mathrm{OPT}(T_H) = 2n+m+n/2 = 3n$. For 2-regular graphs $m = n$ and $\alpha(C_\ell) = \lfloor \ell/2 \rfloor$, giving $\mathrm{OPT}(T_H) = 3n+m+\alpha(H) = 4n+\sum_{j=1}^{k} \lfloor \ell_j/2 \rfloor$. □

Proposition 6. *Let H be simple and d-regular, $d \geq 1$, and $B = (d+1)n + m$.*

1. *The assignment with every clone true has value B. It is a $\frac{3d+2}{3d+3}$-approximation on $\{T_H\}$.*
2. *If $d \in \{1, 2\}$, there is an exact polynomial-time algorithm on $\{T_H : H \text{ is simple and } d\text{-regular}\}$ (ratio $1 > \frac{83}{84}$).*
3. *If $d \geq 27$, then $\frac{3d+2}{3d+3} \geq \frac{83}{84}$.*
4. *If $d \geq 22$ and an independent set I satisfies $5\alpha \leq (d+3)|I|$, the Berman–Fujito guarantee [2], its lift has value at least $\frac{83}{84}\,\mathrm{OPT}(T_H)$.*
5. *If an independent set I satisfies $7|I| \geq 6\alpha$, its lift has value at least $\frac{21d+20}{21d+21}\,\mathrm{OPT}(T_H)$.*
6. *If $d = 21$, there is a polynomial-time $83/84$ approximation on $\{T_H : H \text{ is simple and } 21\text{-regular}\}$.*
7. *If an independent set I satisfies $28|I| + d\alpha \geq 27\alpha$, that is $|I| \geq \rho(d)\alpha$ with $\rho(d) = \frac{27-d}{28}$, its lift has value at least $\frac{83}{84}\,\mathrm{OPT}(T_H)$.*
8. *For every fixed $\varepsilon > 0$ and $d \geq 3$ there is a polynomial-time algorithm on $\{T_H : H \text{ is simple and } d\text{-regular}\}$ whose value is at least $\left(1 - \frac{d-2+\varepsilon}{(3d+3)(d+3+\varepsilon)}\right)\mathrm{OPT}(T_H)$. This exceeds $\frac{3d+2}{3d+3}$ for every $d \geq 3$, and the limit $\varepsilon \to 0$ is $1 - \frac{d-2}{3(d+1)(d+3)}$.*

Proof. With every clone true, the n unit clauses fail and the m edge clauses and all $(d+1)n$ consistency implications hold, so the value is B. Since $\alpha \leq n/2$ for every $d \geq 1$ and $2B = (3d+2)n$, we get $B \geq (3d+2)\alpha$, which gives the first claim. The second claim is Proposition 5. For the third, $3d + 2 \geq 83$. For the Berman–Fujito claim, multiply the required inequality $B + 84|I| \geq 83\alpha$ by $d + 3$ and use $(d+3)|I| \geq 5\alpha$, $(d+3)B \geq (d+3)(3d+2)\alpha$ and $3d^2 - 72d + 177 \geq 0$ for $d \geq 22$. The fifth claim is the endpoint computation above: using $7|I| \geq 6\alpha$, $(21d + 21)(B + |I|) - (21d + 20)(B + \alpha) \geq B - (3d + 2)\alpha \geq 0$. For $d = 21$ put $a = \alpha/n$, so $B = 32.5\,n$ and $m = 10.5\,n$. The algorithm returns the better lift of two independent sets, without knowing a. The first is a Berman–Fujito set with $\varepsilon = 1/50$, which has size at least $50\alpha/241$ since the guarantee for degree 21 is $24/5 + \varepsilon = 241/50$. The second is obtained from a cut: solve the Goemans–Williamson semidefinite relaxation of Max Cut on H to additive error 10^{-3}, round it, and derandomize the rounding by the method of Mahajan and Ramesh [4, 5]; then delete one endpoint of every edge inside each side, and keep the larger side. If $a \leq 99/200$, the first set suffices, because

$$84 \cdot \frac{50}{241}\alpha \geq 83\alpha - B \quad \Longleftrightarrow \quad \frac{15803}{241}\,a \leq 32.5,$$

and $\frac{15803}{241} \cdot \frac{99}{200} = \frac{1564497}{48200} < 32.5$. If $a > 99/200$, the independent set of size α cuts 21α of the m edges, a fraction $2a > 0.99$, so the relaxation has value at least $c = 0.989$. By the analysis of Goemans and Williamson, rounding yields expected cut fraction at least $\arccos(1-2c)/\pi$ whenever $c \geq 0.845$; this holds because $\arccos(1-2x)/\pi$ is convex on $[1/2, 1]$ and its convex minorant coincides with it for $x \geq 0.845$. Here $\arccos(-0.978)/\pi > 0.93$. Deleting one endpoint of each uncut edge leaves an independent set of size at least $(n-0.07m)/2 = (1-0.735)n/2 > 0.12\,n$ on the larger side. Then $B+84|I| \geq 32.5\,n+10.08\,n > 41.5\,n \geq 83\alpha$, using $\alpha \leq n/2$. In both cases $B + 84|I| \geq 83\alpha$, which is the required ratio. For the seventh claim, $B \geq (3d+2)\alpha$ by the argument of the first claim, and the hypothesis gives $84|I| \geq 3(27-d)\alpha$. Adding the two gives $B + 84|I| \geq 83\alpha$. For the eighth claim, the Berman–Fujito set I with $5\alpha \leq (d+3+\varepsilon)|I|$ satisfies $\alpha - |I| \leq \frac{d-2+\varepsilon}{d+3+\varepsilon}\alpha$, and $B + \alpha \geq (3d+3)\alpha$, so $\frac{B+|I|}{B+\alpha} = 1 - \frac{\alpha-|I|}{B+\alpha} \geq 1 - \frac{d-2+\varepsilon}{(3d+3)(d+3+\varepsilon)}$; the loss is below $\frac{1}{3d+3}$ because $d - 2 + \varepsilon < d + 3 + \varepsilon$. □

The threshold $\rho(d) = (27-d)/28$ equals $6/7$ for $d = 3$. Selected values:

d	3	4	5	6	8	10	12	16
$\rho(d)$	6/7	23/28	11/14	3/4	19/28	17/28	15/28	11/28

The value 83/84 for $d = 3$ uses the Berman–Fujito branch and the near-bipartite branch of Theorem 4. The 6/7 hypothesis in the fifth claim is the hypothesis whose cubic instance is supplied by that theorem. For $\varepsilon \to 0$ the ratio of the eighth claim is $1 - \frac{d-2}{3(d+1)(d+3)}$. Selected values:

d	4	5	6	20
$1 - \frac{d-2}{3(d+1)(d+3)}$	$\frac{103}{105}$	$\frac{47}{48}$	$\frac{185}{189}$	$\frac{159}{161}$
$\frac{3d+2}{3d+3}$	$\frac{14}{15}$	$\frac{17}{18}$	$\frac{20}{21}$	$\frac{62}{63}$

The values of the limit $1 - \frac{d-2}{3(d+1)(d+3)}$ of the eighth claim for $4 \le d \le 20$, with the all-true ratio $\frac{3d+2}{3d+3}$:

d	$1 - \frac{d-2}{3(d+1)(d+3)}$	decimal	$\frac{3d+2}{3d+3}$
4	103/105	0.98095	0.93333
5	47/48	0.97917	0.94444
6	185/189	0.97884	0.95238
7	47/48	0.97917	0.95833
8	97/99	0.97980	0.96296
9	353/360	0.98056	0.96667
10	421/429	0.98135	0.96970
11	55/56	0.98214	0.97222
12	115/117	0.98291	0.97436
13	661/672	0.98363	0.97619
14	251/255	0.98431	0.97778
15	851/864	0.98495	0.97917
16	955/969	0.98555	0.98039
17	71/72	0.98611	0.98148
18	1181/1197	0.98663	0.98246
19	1303/1320	0.98712	0.98333
20	159/161	0.98758	0.98413

Every value in this table is greater than $\alpha^* \approx 0.94616$ and greater than $\beta^\star \approx 0.94017$. The smallest value is $\frac{185}{189} \approx 0.97884$, attained at $d = 6$, because $63(d-2) \le 4(d+1)(d+3)$ is equivalent to $(4d-23)(d-6) \ge 0$, which holds for every integer d. Explicitly, the endpoint ratio obtained from a 6/7-approximate independent set is

$$\frac{21d + 20}{21d + 21}.$$

Selected values:

d	1	2	3	4	5	6	8	10	12
ratio	41/42	62/63	83/84	104/105	125/126	146/147	188/189	230/231	272/273
	≈ 0.9762	≈ 0.9841	≈ 0.9881	≈ 0.9905	≈ 0.9921	≈ 0.9932	≈ 0.9947	≈ 0.9957	≈ 0.9963

(For $d \in \{1, 2\}$, exact MIS yields ratio 1 by Proposition 5.) As $d \to \infty$ both $\frac{3d+2}{3d+3}$ and $\frac{21d+20}{21d+21}$ tend to 1.

Theorem 7 (Complete d-regular classification for all $d \ge 1$)**.** *For every integer $d \ge 1$ and every simple d-regular graph $H = (V, E)$ with $n = |V|$ and $m = |E| = dn/2$, the three-occurrence Horn formula T_H satisfies $\mathrm{OPT}(T_H) = (d+1)n+m+\alpha(H)$, and every assignment of value A decodes in polynomial time to an independent set of H of size at least $A-((d+1)n+m)$. Moreover, the class $\{T_H : H$ is simple and d-regular$\}$ admits polynomial-time approximation algorithms with the following guarantees across all $d \ge 1$:*

1. *For $d \in \{1, 2\}$, $\mathrm{OPT}(T_H)$ and an optimal assignment are computed exactly in $O(n)$ time (approximation ratio $1 > \frac{83}{84}$).*
2. *For $d = 3$, there is a polynomial-time $\frac{83}{84}$-approximation.*
3. *For $4 \le d \le 20$ and every fixed $\varepsilon > 0$, there is a polynomial-time $\left(1 - \frac{d-2+\varepsilon}{(3d+3)(d+3+\varepsilon)}\right)$-approximation, strictly exceeding the all-true baseline $\frac{3d+2}{3d+3}$.*
4. *For $d = 21$, there is a polynomial-time $\frac{83}{84}$-approximation.*
5. *For $22 \le d \le 26$, there is a polynomial-time $\frac{83}{84}$-approximation.*
6. *For $d \ge 27$, the all-true assignment achieves a factor of at least $\frac{3d+2}{3d+3} \ge \frac{83}{84}$.*

In addition, for every $d \geq 1$, any independent set of size at least $\frac{6}{7}\alpha(H)$ lifts in polynomial time to a $\frac{21d+20}{21d+21}$-approximation of T_H. Finally, for every $d \geq 1$ the class admits a polynomial-time approximation of ratio at least 0.978 (with $\varepsilon = 1/10$ in item (3)); the smallest limiting factor over all $d \geq 1$ is $\frac{185}{189} \approx 0.97884$, attained at $d = 6$ as $\varepsilon \to 0$. Consequently every class $\{T_H : H \text{ is simple and } d\text{-regular}\}$, $d \geq 1$, has a polynomial-time ratio at least $0.978 > \alpha^* \approx 0.94616 > \beta^\star \approx 0.94017$, so that both Unique Games thresholds are exceeded for every degree $d \geq 1$.

Proof. The exact identity and quantitative repair follow from Lemma 15 since $\sum_{v\in V}(\deg_H(v)+1)+m = (d+1)n+m$. Item (1) is Proposition 5; item (2) is Theorem 4; item (3) is Proposition 6(8); item (4) is Proposition 6(6); item (5) is Proposition 6(4); item (6) is Proposition 6(1) and (3); and the final statement is Proposition 6(5). For the last statement, the ratios of items (1), (2), (4), (5) and (6) are at least $\frac{83}{84} > 0.978$. For $4 \leq d \leq 20$ the loss of item (3) is increasing in ε, so it suffices to take $\varepsilon = \frac{1}{10}$, and the loss $\frac{d-2+1/10}{(3d+3)(d+3+1/10)}$ is at most 0.022 because $0.022\,(3d+3)(d+3.1)-(d-1.9) = 0.066d^2-0.7294d+2.1046 > 0$ (its discriminant is negative). The limiting factors $1-\frac{d-2}{3(d+1)(d+3)}$ are at least $\frac{185}{189}$ because $(4d-23)(d-6) \geq 0$ for every integer d, with equality at $d = 6$. Finally $0.978 > 0.94616 > 0.94017$. □

5. APX-completeness

Theorem 8. *The cubic three-occurrence Horn class is APX-complete under L-reductions. The reduction from cubic MIS has constants 23 and 1.*

Proof. A greedy maximal independent set in a graph of maximum degree three has size at least $n/4$, so $n \leq 4\alpha(H)$. Since $m = 3n/2$,

$$\mathrm{OPT}(T_H) = \frac{11n}{2} + \alpha(H) \leq 22\alpha(H) + \alpha(H) = 23\alpha(H).$$

For an assignment of value A, Lemma 2 returns R with $|R| \geq A - (4n+m)$. Hence

$$\alpha(H) - |R| \leq (4n + m + \alpha(H)) - A = \mathrm{OPT}(T_H) - A.$$

These are the two L-reduction inequalities. Cubic MIS is APX-complete and Theorem 4 places the image class in APX. □

5.1. **Hardness for larger degree.** Whereas $\{T_H : H \text{ is simple and } d\text{-regular}\}$ lies in P for $d \in \{1, 2\}$ (Proposition 5), the same L-reduction argument yields APX-completeness for every fixed $d \geq 3$. Maximum Independent Set remains APX-complete on simple d-regular graphs (and on graphs of maximum degree d) for each fixed $d \geq 3$ [1, 2]. The construction of T_H from a d-regular graph is an L-reduction whose constants depend only on d. A greedy maximal independent set in a degree-d graph has size at least $n/(d+1)$, so $n \leq (d+1)\alpha$. With $m = dn/2$ one obtains

$$\mathrm{OPT}(T_H) = (d+1)n + m + \alpha \leq (d+1)^2\Big(\tfrac{3d}{2}+1\Big)\alpha + \alpha.$$

Thus the first L-reduction constant is $O(d^2)$. The second constant remains exactly 1 by the quantitative repair: every assignment of value A produces an independent set R satisfying $\alpha - |R| \leq \mathrm{OPT}(T_H) - A$. Consequently the image class

$$\{T_H : H \text{ is simple and } d\text{-regular}\}$$

is APX-complete for every fixed $d \geq 3$. By Proposition 6(1) the image class is approximable to the factor $1 - \frac{1}{3d+3}$ for every $d \geq 1$, so it lies in APX and is APX-complete for all $d \geq 3$.

6. MAX-CUT as MAX 2-SAT

Proposition 9. *Let $G = (V, E)$ be a graph with n vertices and m edges. Let F_G be the formula with one variable x_v for every vertex and, for every edge uv, the two clauses $(x_u \vee x_v)$ and $(\neg x_u \vee \neg x_v)$. Then:*

1. *an assignment z satisfies exactly $m + \mathrm{cut}(S_z)$ clauses, where $S_z = \{v : z(x_v) = \mathit{true}\}$ and $\mathrm{cut}(S)$ is the number of edges with exactly one endpoint in S; hence $\mathrm{OPT}(F_G) = m + \mathrm{maxcut}(G)$;*
2. *the variable x_v occurs in exactly $2\deg(v)$ clauses;*

(3) *if an assignment satisfies at least* $\rho\,\mathrm{OPT}(F_G)$ *clauses, then the cut* S_z *has at least* $\rho\,(m + \mathrm{maxcut}(G)) - m$ *edges.*

Proof. For an edge uv whose endpoints receive different values both clauses hold. If both endpoints are true only $(x_u \vee x_v)$ holds, and if both are false only $(\neg x_u \vee \neg x_v)$ holds. Summing over the edges gives the first claim, and the second and third follow. □

The clause $(\neg x_u \vee \neg x_v)$ has two negative literals, so F_G is an instance of MAX 2-SAT that does not belong to the Horn class; no guarantee of this paper is claimed for it.

Proposition 10. *Let* G *be a simple graph with* n *vertices and* m *edges. Let* H *be the conflict graph of* F_G*: its* $N = 4m$ *vertices are the literal occurrences of* F_G*, and two vertices are adjacent if they lie in the same clause or are complementary literals. Then* H *is simple, and*

$$\alpha(H) = \mathrm{OPT}(F_G) = m + \mathrm{maxcut}(G).$$

Consequently, with $M = |E(H)|$ *and* $B = N + 3M$*, the formula* T_H *satisfies* $\mathrm{OPT}(T_H) = B + m + \mathrm{maxcut}(G)$*, and an assignment of* T_H *of value* A *yields in polynomial time a cut of* G *with at least* $A - B - m$ *edges.*

Proof. Every literal occurrence has exactly one neighbour in its clause. A positive occurrence of x_u is adjacent to the $\deg(u)$ negative occurrences of x_u, and symmetrically; an in-clause neighbour is a literal of another variable, so no edge is repeated. An independent set contains at most one literal of each clause and no two complementary literals, so it is a consistent partial assignment satisfying as many clauses as the set has vertices; conversely, an assignment and one satisfied literal per satisfied clause give an independent set. Thus $\alpha(H) = \mathrm{OPT}(F_G)$, which equals $m + \mathrm{maxcut}(G)$ by Proposition 9. The identity for T_H is the exact identity for arbitrary graphs. By the quantitative repair, an assignment of value A gives an independent set of size at least $A - B$, hence an assignment of F_G satisfying at least $A - B$ clauses, hence a cut with at least $A - B - m$ edges by Proposition 9. □

Corollary 11. *Let* $\mathrm{maxcut}(G) = \gamma m$ *and* $0 \le \rho \le 1$*.*

(1) *If an assignment of* F_G *satisfies at least* $\rho\,\mathrm{OPT}(F_G)$ *clauses, its cut has at least* $\big(\rho - \frac{1-\rho}{\gamma}\big)\,\mathrm{maxcut}(G)$ *edges.*

(2) *If an assignment of* T_H *has value at least* $\rho\,\mathrm{OPT}(T_H)$*, the cut obtained from it has at least* $\big(\rho - \frac{(1-\rho)(B/m+1)}{\gamma}\big)\,\mathrm{maxcut}(G)$ *edges.*

Proof. By Proposition 9, the cut has at least $\rho(m + \gamma m) - m$ edges; dividing by γm gives the first claim. By Proposition 10 it has at least $\rho(B + m + \gamma m) - B - m$ edges, which gives the second. □

Theorem 12. *For every* $d \ge 1$ *there is a polynomial-time algorithm that, given a simple* d*-regular graph* G*, returns an assignment of* $T_{H(G)}$ *of value at least*

$$\Big(1 - \frac{1 - \alpha_\varepsilon}{6d + 12}\Big)\,\mathrm{OPT}(T_{H(G)}) \;\ge\; \frac{83}{84}\,\mathrm{OPT}(T_{H(G)}),$$

where $H(G)$ *is the conflict graph of Proposition 10 and* $\alpha_\varepsilon = \alpha_{\mathrm{GW}} - 10^{-5} \ge 0.8785$*, with* $\alpha_{\mathrm{GW}} \approx 0.87856$ *the Goemans–Williamson constant.*

Proof. Let $c = \mathrm{maxcut}(G)$. Solving the Goemans–Williamson relaxation to additive error 10^{-5} and rounding with derandomization [4, 5] returns in polynomial time a set S with $\mathrm{cut}(S) \ge \alpha_\varepsilon\, c$ (the relaxation error 10^{-5} is absorbed in α_ε). Set x_v true exactly for $v \in S$ and choose one true literal in every satisfied clause of F_G. These $m + \mathrm{cut}(S)$ literal occurrences are pairwise non-adjacent in $H(G)$, because they lie in distinct clauses and no two are complementary. So they form an independent set I with $|I| = m + \mathrm{cut}(S)$, while $\alpha(H(G)) = m + c$ by Proposition 10. Since $c \le m$,

$$|I| \ge m + \alpha_\varepsilon\, c \ge \frac{1 + \alpha_\varepsilon}{2}\,(m + c) = \frac{1 + \alpha_\varepsilon}{2}\,\alpha(H(G)),$$

because $(1-\alpha_\varepsilon)(m-c) \geq 0$. Let $D = d+1$, so $H(G)$ is D-regular and $B = (D+1)N+M \geq (3D+2)\alpha(H(G))$ by $\alpha(H(G)) \leq N/2$. The lift of I has value $B + |I|$, hence

$$\frac{B+|I|}{B+\alpha(H(G))} \geq 1 - \frac{\frac{1-\alpha_\varepsilon}{2}\,\alpha(H(G))}{B+\alpha(H(G))} \geq 1 - \frac{1-\alpha_\varepsilon}{6d+12}.$$

Finally $\frac{1-\alpha_\varepsilon}{6d+12} \leq \frac{1-\alpha_\varepsilon}{18} < \frac{1}{84}$ for $d \geq 1$, since $1-\alpha_\varepsilon \leq 0.1215 < 18/84 \approx 0.214$. □

Proposition 13. *For every integer $d \geq 1$ and every simple graph G with at least one edge and maximum degree at most d, the polynomial-time algorithm of Theorem 16 returns an assignment of the conflict-graph formula $T_{H(G)}$ of value at least $0.9932\,\mathrm{OPT}(T_{H(G)})$, which is strictly greater than $\frac{83}{84}\,\mathrm{OPT}(T_{H(G)})$.*

Proof. Let G be a simple graph with $m \geq 1$ edges and maximum degree at most d. By Theorem 16, the derandomized Goemans–Williamson algorithm yields a cut that lifts to an independent set I in $H(G)$ satisfying $|I| \geq \frac{1+\alpha_\varepsilon}{2}\alpha(H(G)) \geq 0.93925\alpha(H(G))$. The number of vertices of $H(G)$ is $N = 4m$, and its number of edges is $M = 2m+\sum_v \deg(v)^2 \geq 4m$, because $\deg(v)^2 \geq \deg(v)$ for every integer $\deg(v) \geq 0$ and $\sum_v \deg(v) = 2m$. Thus $B(H(G)) = N+3M \geq 16m$. On the other hand, $\alpha(H(G)) = m+\mathrm{maxcut}(G) \leq 2m$, so $B(H(G)) \geq 8\alpha(H(G))$. The approximation ratio of the lifted assignment is therefore at least

$$1 - \frac{1-\alpha_\varepsilon}{2}\,\frac{\alpha(H(G))}{B(H(G))+\alpha(H(G))} \geq 1 - \frac{1-\alpha_\varepsilon}{18} > 0.9932.$$

Since $0.9932 > 83/84$, the condition is strictly satisfied. □

Remark 14. *The bound on the maximum degree of G is not used in the ratio 0.9932: it holds for every graph with at least one edge. It controls the degrees of the conflict graph, because every vertex of $H(G)$ has degree $\deg_G(v)+1 \leq d+1$ (proof of Theorem 16), so that $T_{H(G)}$ has at most $d+2$ clones for every vertex of $H(G)$. By Proposition 13, the ratio $0.9932 > \frac{83}{84}$ is therefore achieved on the conflict-graph class $\{T_{H(G)} : G \text{ has maximum degree at most } d\}$ for every $d \geq 1$; for d-regular G Theorem 12 gives the factor $1 - \frac{1-\alpha_\varepsilon}{6d+12}$.*

6.1. **Generalization to arbitrary simple graphs.** The three-occurrence construction extends without change of method to any finite simple graph $H = (V, E)$ with $n = |V|$ and $m = |E|$, even when H is not regular. At each vertex $v \in V$ of degree $d_v = \deg_H(v)$, label its incident edges with the ports $1, \ldots, d_v$ and introduce $d_v + 1$ Boolean variables

$$z_{v,0}, z_{v,1}, \ldots, z_{v,d_v}.$$

The formula T_H contains:

(1) for every vertex $v \in V$, one unit clause $(\neg z_{v,0})$;
(2) for every edge $e = uv \in E$ with endpoint ports p at u and q at v, one edge clause $(z_{u,p} \vee z_{v,q})$;
(3) for every vertex $v \in V$ and $i \in \mathbb{Z}/(d_v+1)\mathbb{Z}$, one consistency implication $(\neg z_{v,i} \vee z_{v,i+1})$.

Every clause has at most one negative literal. Each clone $z_{v,i}$ occurs once in a source clause (the unit clause at v for $i = 0$, or the edge clause at port i for $1 \leq i \leq d_v$), once negatively in $(\neg z_{v,i} \vee z_{v,i+1})$, and once positively in $(\neg z_{v,i-1} \vee z_{v,i})$. Thus every variable of T_H still has exactly three literal occurrences. Using $\sum_{v\in V} d_v = 2m$, the total number of variables is

$$\sum_{v\in V}(d_v+1) = n + 2m,$$

and the total number of clauses is $n + m + (n + 2m) = 2n + 3m$. Put

$$B(H) = \sum_{v\in V}(d_v+1) + m = n + 3m.$$

Lemma 15 (General quantitative repair and exact identity)**.** *For every finite simple graph $H = (V, E)$ with $n = |V|$ and $m = |E|$,*

$$\mathrm{OPT}(T_H) = n + 3m + \alpha(H).$$

Moreover, every assignment z of value A decodes in polynomial time to an independent set $R(z) \subseteq V$ of size at least $A - (n + 3m)$.

Proof. For an assignment z, let $U(z) = |\{v \in V : z_{v,0} = 0\}|$, let $F(z)$ be the number of falsified edge clauses, and let $T(z) = \sum_{v\in V} |\{i \in \mathbb{Z}/(d_v+1)\mathbb{Z} : z_{v,i} = 1,\ z_{v,i+1} = 0\}|$. Since there are $n+2m$ consistency implications in total,

$$\mathrm{val}_{T_H}(z) = U(z) + (m - F(z)) + (n + 2m - T(z)) = n + 3m + U(z) - F(z) - T(z).$$

Call a vertex $v \in V$ selected when $z_{v,0} = 0$, and clean when it is selected and its (d_v+1)-bit cycle has no 1-to-0 transition. By Lemma 1 (which holds for any cycle length $d_v + 1 \geq 1$), every clone $z_{v,i}$ of a clean selected vertex v is zero. Let $C(z)$ be the set of clean selected vertices, and form $R(z)$ by deleting one endpoint of each edge induced by $C(z)$, ignoring repeated deletions. Then $R(z)$ is independent in H. Each nonclean selected vertex contributes at least one 1-to-0 transition, so at most $T(z)$ selected vertices are nonclean. Every edge induced by $C(z)$ has both endpoint clones equal to zero and is therefore counted by $F(z)$, so at most $F(z)$ vertices are removed from $C(z)$. Consequently,

$$|R(z)| \geq |C(z)| - F(z) \geq U(z) - T(z) - F(z) = \mathrm{val}_{T_H}(z) - (n + 3m).$$

This gives the constructive repair and the upper bound $\mathrm{OPT}(T_H) \leq n + 3m + \alpha(H)$. Conversely, for any independent set $I \subseteq V$, setting $z_{v,i} = 0$ for all $v \in I$ and $0 \leq i \leq d_v$, and $z_{v,i} = 1$ otherwise, satisfies all $n + 2m$ consistency implications, all m edge clauses, and $|I|$ unit clauses, realizing value $n + 3m + |I|$. Taking I maximum yields equality. □

Theorem 16 (Non-regular conflict-graph family)**.** *Let $G = (V_G, E_G)$ be any finite simple graph with $m = |E_G| \geq 1$ edges, and let $H(G)$ be the conflict graph of F_G from Proposition 10. Then every variable of $T_{H(G)}$ has exactly three literal occurrences, if G has maximum degree at most d then $H(G)$ has maximum degree at most $d + 1$, and there is a polynomial-time algorithm returning an assignment of $T_{H(G)}$ of value at least*

$$\Big(1 - \frac{1 - \alpha_\varepsilon}{18}\Big)\,\mathrm{OPT}(T_{H(G)}) > 0.9932\ \mathrm{OPT}(T_{H(G)}).$$

Proof. Each vertex $v \in V_G$ of degree $d_v = \deg_G(v)$ contributes d_v positive literal occurrences and d_v negative literal occurrences to F_G (none when $d_v = 0$). Thus $H(G)$ has $N = 2\sum_{v\in V_G} d_v = 4m$ vertices. Every literal occurrence of the variable x_v has 1 in-clause neighbor and d_v complementary neighbors in $H(G)$ (with no overlap, as in Proposition 10), so those $2d_v$ vertices each have degree $d_v + 1$ in $H(G)$; in particular the maximum degree of $H(G)$ is $1 + \max_v d_v$. Hence the number of edges of $H(G)$ is

$$M = |E(H(G))| = \frac{1}{2}\sum_{v\in V_G} 2d_v(d_v + 1) = \sum_{v\in V_G}(d_v^2 + d_v) = 2m + \sum_{v\in V_G} d_v^2.$$

Since $d_v^2 \geq d_v$ for every integer $d_v \geq 0$, $\sum_{v\in V_G} d_v^2 \geq \sum_{v\in V_G} d_v = 2m$ and therefore $M \geq 4m$. By Lemma 15,

$$\mathrm{OPT}(T_{H(G)}) = B(H(G)) + \alpha(H(G)), \qquad B(H(G)) = N + 3M \geq 4m + 12m = 16m.$$

On the other hand, as in Proposition 10, an independent set in $H(G)$ selects at most one literal per clause of F_G and no complementary pair, so $\alpha(H(G)) = \mathrm{OPT}(F_G) = m + \mathrm{maxcut}(G) \leq 2m$. Consequently, $B(H(G)) \geq 16m \geq 8\,\alpha(H(G))$. Now run the derandomized Goemans–Williamson algorithm [4, 5] on G to obtain in polynomial time a cut S with $\mathrm{cut}(S) \geq \alpha_\varepsilon\ \mathrm{maxcut}(G)$. Selecting one true literal per satisfied clause of F_G under S yields an independent set I in $H(G)$ of size $|I| = m + \mathrm{cut}(S)$. Because $\mathrm{maxcut}(G) \leq m$, we have $(1 - \alpha_\varepsilon)(m - \mathrm{maxcut}(G)) \geq 0$, and hence

$$|I| \geq m + \alpha_\varepsilon\ \mathrm{maxcut}(G) \geq \frac{1 + \alpha_\varepsilon}{2}(m + \mathrm{maxcut}(G)) = \frac{1 + \alpha_\varepsilon}{2}\,\alpha(H(G)).$$

Lifting I to $T_{H(G)}$ gives an assignment of value $B(H(G)) + |I|$, whose approximation ratio is

$$\begin{aligned}\frac{B(H(G)) + |I|}{B(H(G)) + \alpha(H(G))} &= 1 - \frac{\alpha(H(G)) - |I|}{B(H(G)) + \alpha(H(G))} \\ &\geq 1 - \frac{\frac{1-\alpha_\varepsilon}{2}\,\alpha(H(G))}{9\,\alpha(H(G))} \\ &= 1 - \frac{1-\alpha_\varepsilon}{18} \\ &> 0.9932.\end{aligned}$$

□

7. Unique Games thresholds and the 83/84 barrier

Let α^* denote the constant of Brakensiek, Huang and Zwick for weighted MAX $\{1,2\}$-HORN-SAT, approximately 0.94615981. Their Theorem 5.12 gives UG-hardness at $\alpha^* + \varepsilon$ for every $\varepsilon > 0$ [3]. The constant $\beta^\star \approx 0.94016567$ of Lewin, Livnat and Zwick is the UG-hard threshold for MAX 2-SAT [3]. Both UG-hardness statements quantify over *all* (unrestricted) instances.

Proposition 17 (UGC threshold and P = NP)**.** *Under the Unique Games Conjecture, the existence of a polynomial-time algorithm that achieves an approximation ratio strictly above α^* for weighted MAX $\{1,2\}$-HORN-SAT or strictly above $\beta^\star$ for MAX 2-SAT on* all *instances would imply* P = NP*. However, the guarantees established in this paper apply exclusively to restricted classes where every variable has exactly three literal occurrences. Exceeding the UGC threshold on such bounded-occurrence classes does not contradict the conjecture, because the hardness statements quantify over all instances.*

Theorem 18 (Ratio above the Unique Games thresholds)**.** *For every $d \geq 1$, there is a polynomial-time algorithm that, given a simple graph G with at least one edge and maximum degree at most d, returns an assignment of the three-occurrence Horn formula $T_{H(G)}$ of value at least $0.9932\,\mathrm{OPT}(T_{H(G)}) > \frac{83}{84}\,\mathrm{OPT}(T_{H(G)})$, and*

$$\frac{83}{84} > \alpha^* \approx 0.94616 > \beta^\star \approx 0.94017.$$

Under the Unique Games Conjecture, unless P = NP*, no polynomial-time algorithm has a factor above $\alpha^* + \varepsilon$ on all instances of weighted MAX $\{1,2\}$-HORN-SAT, or above $\beta^\star + \varepsilon$ on all instances of MAX 2-SAT, for any $\varepsilon > 0$. The class $\{T_{H(G)} : G$ has maximum degree at most $d\}$ is a class of Horn formulas, hence of MAX 2-SAT instances, in which every variable has exactly three literal occurrences, and it is consistent with the conjecture.*

Proof. The algorithm and the factor $0.9932 > 83/84$ are Proposition 13. The numerical inequalities and the hardness statements are those quoted above and [3]. □

Theorem 19 (General conflict-graph theorem)**.**

(1) *Let $H = (V, E)$ be any finite simple graph with $n = |V|$ and $m = |E|$. The formula T_H constructed above for an arbitrary simple graph has exactly three literal occurrences for every variable, at most two literals and at most one negative literal in every clause, and*

$$\mathrm{OPT}(T_H) = n + 3m + \alpha(H).$$

Every assignment of value A decodes in polynomial time to an independent set of H of size at least $A - (n + 3m)$.

(2) *Let G be a finite simple graph with $m \geq 1$ edges, let $H(G)$ be its conflict graph, with $N = 4m$ vertices and $M = 2m + \sum_v \deg_G(v)^2$ edges, and put $B = N + 3M$. Then*

$$\mathrm{OPT}(T_{H(G)}) = B + m + \mathrm{maxcut}(G),$$

and from an assignment of $T_{H(G)}$ of value A one computes in polynomial time a cut of G with at least $A - B - m$ edges.

(3) *There is a polynomial-time algorithm that returns an assignment of $T_{H(G)}$ of value at least $(1 - \frac{1-\alpha_\varepsilon}{18})\,\mathrm{OPT}(T_{H(G)}) > 0.9932\,\mathrm{OPT}(T_{H(G)})$. If G has maximum degree at most d, then $H(G)$ has maximum degree at most $d + 1$. If G is d-regular, the value is at least $(1 - \frac{1-\alpha_\varepsilon}{6d+12})\,\mathrm{OPT}(T_{H(G)})$.*

(4) $0.9932 > \frac{83}{84} > \alpha^* > \beta^\star$, *so the class* $\{T_{H(G)}\}$ *exceeds both Unique Games thresholds. Since these thresholds concern all instances, the statement is consistent with the Unique Games Conjecture.*

Proof. Part (1) is Lemma 15 together with the occurrence count of the construction in the previous subsection. For part (2), $H(G)$ has $N = 4m$ vertices and $M = 2m + \sum_v \deg_G(v)^2$ edges by the proof of Theorem 16, and $\alpha(H(G)) = m + \mathrm{maxcut}(G)$ by Proposition 10, whose proof of this identity does not use regularity. Lemma 15 applied to $H(G)$ gives $\mathrm{OPT}(T_{H(G)}) = N + 3M + \alpha(H(G))$, which is the stated identity. An assignment of value A decodes to an independent set of $H(G)$ of size at least $A - B$, hence to an assignment of F_G satisfying at least $A - B$ clauses, hence by Proposition 9 to a cut with at least $A - B - m$ edges. Part (3) is Theorem 16 and, for regular G, Theorem 12. Part (4): $0.9932 > 83/84 \approx 0.98810$ and $83/84 > \alpha^* > \beta^\star$ by Theorem 18. □

8. Lean verification

The module `PalomaMax2SAT.CubicThreeOccurrence` represents the cubic graph with three ports at every vertex. It formalizes the transition count, false edge clauses, explicit deletion repair, independence of the repaired set, its cardinality bound, the matching assignment from an independent set, and the exact optimum identity. It also proves the division-free inequality

$$84(4n + m + |I|) \geq 83(4n + m + \alpha)$$

from $2m = 3n$, $2\alpha \leq n$, and $7|I| \geq 6\alpha$, and the rational identity $11 + 6/7 = (83/84) \cdot 12$. The file `Solution.lean` proves directly the general arithmetic transfer

$$84\big((d+1)n + m + |I|\big) \geq 83\big((d+1)n + m + \alpha\big)$$

valid for every $d \geq 3$ under the analogous hypotheses $2m = dn$ and $2\alpha \leq n$, and the monotonicity statement that the endpoint ratio $(\beta + 3/7)/(\beta + 1/2)$ is at least $83/84$ for all $\beta \geq 11/2$. GitHub Actions builds the module and runs the public theorem comparison. The module `PalomaMax2SAT.SemanticHorn` defines T_H explicitly as a family of clauses, each a list of literals over the clone variables, indexed by the vertices (unit clauses), the edges (edge clauses) and the pairs (v, i) (consistency implications). It proves that the number of clauses of T_H satisfied by z equals $U(z) + (m - F(z)) + (4n - T(z))$, that the maximum of this number over all assignments is exactly $4n + m + \alpha$, that every clause has at most one negative literal, that every clause becomes a Horn clause when every variable is renamed to its negation and that this renaming preserves satisfaction of each clause, that T_H has $5n + m$ clauses, and that every variable occurs exactly three times. The Berman–Fujito guarantee, the Trevisan near-bipartite branch, the Goemans–Williamson rounding, the APX-completeness of cubic Maximum Independent Set, the discussion of the Unique Games Conjecture are proved in the text or taken from the literature and are not formalized in Lean. The module `PalomaMax2SAT.DegreeTransfer` formalizes the arithmetic of Proposition 6 for all $d \geq 1$. From $2m = dn$ and $2\alpha \leq n$, the value $(d+1)n + m$ satisfies $(3d+2)(B+\alpha) \leq (3d+3)B$ with $B = (d+1)n + m$. For $d \in \{1, 2\}$ and an exact independent set $|I| = \alpha$, the lift has value $B + \alpha = \mathrm{OPT}(T_H)$ (ratio 1, hence $83(B+\alpha) \leq 84(B+|I|)$). For $d \geq 27$ it satisfies $83(B+\alpha) \leq 84B$. For $d \geq 22$ and $5\alpha \leq (d+3)s$ one gets $83(B+\alpha) \leq 84(B+s)$. From $6\alpha \leq 7s$ one gets $(21d+20)(B+\alpha) \leq (21d+21)(B+s)$ for all $d \geq 1$, and from $27\alpha \leq 28s + d\alpha$ one gets $83(B+\alpha) \leq 84(B+s)$ for all $d \geq 1$. For $d = 21$ it proves the two-branch arithmetic: from $2m = 21n$, $2\alpha \leq n$, $200\alpha \leq 99n \Rightarrow 50\alpha \leq 241s$ and $99n < 200\alpha \Rightarrow 3n \leq 25s$ one gets $83(22n+m+\alpha) \leq 84(22n+m+s)$. It also proves the two branches of the $d = 21$ argument separately, the inequality $\frac{15803}{241} \cdot \frac{99}{200} < \frac{65}{2}$, the fact that $200\alpha > 99n$ gives a cut fraction above 0.99, the step from at most $7m/100$ uncut edges to $3n \leq 25|I|$, and the tabulated values of $\rho(d)$ and of $(21d+20)/(21d+21)$. The independent-set guarantees are hypotheses; Lean does not formalize the semidefinite rounding. The module `PalomaMax2SAT.DegreeFormula` defines T_H for every d as an explicit list of Horn clauses over the $d+1$ clones of each vertex. It proves that the number of satisfied clauses equals $U(z) + (m - F(z)) + ((d+1)n - T(z))$, that the optimum is exactly $(d+1)n + m + \alpha$, that every clause has at most one negative literal and becomes a Horn clause when every variable is renamed to its negation, that T_H has $(d+2)n + m$ clauses and that every variable occurs exactly three times. It derives $2m = dn$ from the port structure and $2\alpha \leq n$ for $d > 0$, proves that the lift of an independent set I has value $(d+1)n + m + |I|$, and combines these with the arithmetic of `DegreeTransfer` into statements about T_H itself: the baseline ratio $\frac{3d+2}{3d+3}$

for all $d \geq 1$, exactness (ratio 1) for $d \in \{1, 2\}$ given an exact independent set, the ratio 83/84 for $d \geq 27$, and the transfers of the Berman–Fujito, 6/7, threshold and $d = 21$ hypotheses. The module `PalomaMax2SAT.GeneralDegreeFormula` defines T_H for a graph with a port labelling and a degree d_v at every vertex, as an explicit list of clauses over $d_v + 1$ clones per vertex; the port structure is a bijection between the $2m$ edge ends and the ports. It proves $\sum_v d_v = 2m$, hence $\sum_v (d_v + 1) + m = n + 3m$; that the number of satisfied clauses equals $U(z) + (m - F(z)) + (\sum_v (d_v + 1) - T(z))$; the cycle lemma and the repair inequality $\mathrm{val}(z) \leq B + |R(z)|$ with $R(z)$ independent; the value $B + |I|$ of the lift of an independent set; and hence the identity $\mathrm{OPT} = B + \alpha$ with $B = \sum_v (d_v + 1) + m$ (Lemma 15). The Horn property of the clauses and the three-occurrence count are formalized only for regular graphs, in `DegreeFormula`. The module `PalomaMax2SAT.CutTransfer` proves Proposition 9(1), that an assignment of F_G satisfies exactly $m + \mathrm{cut}(S)$ clauses, and the arithmetic of the two ratio transfers of Corollary 11. The conflict-graph identity of Proposition 10 is proved in `ConflictGraph`. The module `PalomaMax2SAT.CutLift` proves that an independent set I of H with $18785\,\alpha(H) \leq 20000\,|I|$ lifts to the ratio 83/84 on T_H for every degree $D \geq 1$, which is the last step of Theorem 12; the construction of I from a cut is in `ConflictGraph` and the cut guarantee is taken from the literature. The module `PalomaMax2SAT.CitedResults` states the published guarantees used by the $d = 21$ argument as explicit hypotheses, namely the existence of a cut with at most $7m/100$ uncut edges [4, 5] and of a Berman–Fujito independent set [2]. It proves that such a cut yields an independent set I with $3n \leq 25|I|$ and that, under the two hypotheses, some independent set lifts to a 83/84-approximation of T_H for $d = 21$. It also states the Goemans–Williamson cut guarantee as a numerical hypothesis, from which it proves the independent-set step of Theorem 12. The module `PalomaMax2SAT.ConflictGraph` proves $\alpha(H(G)) = m + \mathrm{maxcut}(G)$ for every graph given by its edge list, and `PalomaMax2SAT.ConflictDegree` proves that the conflict graph is $(d+1)$-regular when G is d-regular, that x_v occurs $2\deg(v)$ times and that $H(G)$ has $4m$ vertices. The module `PalomaMax2SAT.LReduction` proves that a graph of maximum degree Δ has an independent set I with $n \leq (\Delta+1)|I|$, the first L-reduction constant $\mathrm{OPT}(T_H) \leq 23\,\alpha(H)$ for cubic graphs and $2\,\mathrm{OPT}(T_H) \leq ((3d+2)(d+1)+2)\alpha(H)$ for d-regular graphs, and the second constant 1 (Theorem 8); the APX-completeness of cubic Maximum Independent Set is cited. The module `PalomaMax2SAT.PaperArithmetic` proves the comparison $\frac{83}{84} > 0.94615981 > 0.94016567$ of Theorem 18, the two-branch arithmetic of Theorem 4 with the two branch guarantees as hypotheses, and the numerical constants of Theorem 12. The repository is available at [7]. The theorem `ugc_bounded` is in the branch `ugc-bounded-degree` of the repository and is published on the Prove2.me platform [8]. The module `PalomaMax2SAT.UgcRegular` proves the theorem `all_degd`, the formalized part of Theorem 7: for every $d \geq 1$ and every d-regular port graph, the optimum $(d+1)n+m+\alpha$, $2m = dn$ and $2\alpha \leq n$; the repair of every assignment to an independent set $R(z)$ with $\mathrm{score}(z) \leq (d+1)n+m+|R(z)|$; the value $(d+1)n+m+|I|$ of the lift of an independent set I; the ratio $\frac{3d+2}{3d+3}$ of the all-true assignment for every $d \geq 1$ and $\frac{83}{84}$ for $d \geq 27$; the ratio $\frac{21d+20}{21d+21}$ for an independent set with $6\alpha \leq 7|I|$ and $\frac{83}{84}$ for $27\alpha \leq 28|I|+d\alpha$, both for every $d \geq 1$; the Berman–Fujito transfer for $d \geq 22$, its division-free form with $\varepsilon = 1/c$ for every degree $e + 2$, and the two-branch transfer for $d = 21$; and, in the module `PalomaMax2SAT.RegularLemmas`, $\alpha(H) = m = n/2$ with $\mathrm{OPT}(T_H) = 3n$ for $d = 1$, and $m = n$ with $\mathrm{OPT}(T_H) = 4n + \alpha(H)$ for $d = 2$. The linear-time computation of a maximum independent set on paths and cycles and the polynomial running times of items (1) to (6) of Theorem 7 are not formalized. The theorem `all_degd` is in the branch `ugc-regular-degree` of the repository and is published on the Prove2.me platform [9]. The module `PalomaMax2SAT.UgcBroken` proves the theorem `ugc_broken`: the threshold comparison $\beta^\star < \alpha^* < 83/84$, the inequality $19(3d+3) < 20(3d+2)$ for $d \geq 6$, and, for every degree $D \geq 3$, the lift of an independent set I with $6\alpha \leq 7|I|$ to a ratio of at least 83/84 on T_H; the lift is Proposition 6(5) for every degree $D \geq 3$, together with the threshold comparison. The module `PalomaMax2SAT.UgcClass` proves the theorem `ugc_class`, a conjunction of the formalized steps of Theorem 12: the threshold comparison; the fact that for every graph G given by an edge list and every vertex set S the occurrences selected by S form an independent set of $H(G)$ of size $m + \mathrm{cut}(S)$; the identity $\alpha(H(G)) = m + \mathrm{maxcut}(G)$; the $(d+1)$-regularity of $H(G)$ for d-regular G and its $4m$ vertices; the implication $8785\,c \leq 10000\,\mathrm{cut} \Rightarrow 18785\,(m + c) \leq 20000\,(m + \mathrm{cut})$; and the lift of an independent set I with $18785\,\alpha \leq 20000\,|I|$ to a ratio of at least 83/84 on T_H for every degree $D \geq 1$. The module `PalomaMax2SAT.UgcBounded` proves the theorem `ugc_bounded`: the comparison $\beta^\star < \alpha^* < 83/84 < 0.9932$; the same selection and independence-number statements for every graph given

by an edge list; that a vertex of $H(G)$ has degree at most $d+1$ when G has maximum degree at most d; that the degrees of $H(G)$ add up to $4m + 2\sum_v \deg_G(v)^2$, so that $H(G)$ has at least $4m$ edges; the division-free step that $4m \le M$, $\alpha \le 2m$ and $18785\,\alpha \le 20000\,|I|$ give $9932(B+\alpha) \le 10000(B+|I|)$ for $B = 4m+3M$, used in Theorem 16; and, for the formula T_H of a port-labelled graph with degrees d_v, the identities of `GeneralDegreeFormula` together with the ratio $9932/10000 > 83/84$ for the lift of an independent set I with $8\alpha(H) \le B$ and $18785\,\alpha(H) \le 20000\,|I|$. The identification of $T_{H(G)}$ with the formula of a port-labelled graph is not formalized. The following statements of the paper are not formalized in Lean: Theorem 19 as a single statement (its parts are the items described above); the Horn property and the three-occurrence count of T_H for non-regular graphs; the APX-completeness statement of Theorem 8 (only the L-reduction constants are proved); the simplicity of $H(G)$ as a separate lemma; $\mathrm{OPT}(F_G) = m + \mathrm{maxcut}(G)$ and $\mathrm{OPT}(T_{H(G)}) = B + m + \mathrm{maxcut}(G)$ as single statements (their ingredients are proved) and Corollary 11 with real γ and ρ (it is proved division-free); the Trevisan bound, the Goemans–Williamson bound and the exhaustive solution of small orders in Theorem 4 and Proposition 6(6); the polynomial running time of every constructive statement; the semidefinite rounding, the Berman–Fujito and Trevisan guarantees, and the Unique Games Conjecture with its hardness thresholds. The notes in `doc/formalization.md` list each cited result with its reference.

Acknowledgments

The author would like to thank Royce Ortiz for his support, which made it possible for the author to have access to the AI tool used in this work. The author used Claude Sonnet 5.5 to assist in deriving the proofs, in writing the Lean 4 formalization, and in preparing this manuscript. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

References

[1] Paola Alimonti and Viggo Kann. Some APX-completeness results for cubic graphs. *Theoretical Computer Science*, 237:123–134, 2000. doi:10.1016/S0304-3975(98)00158-3.

[2] Piotr Berman and Toshihiro Fujito. On approximation properties of the independent set problem for low degree graphs. *Theory of Computing Systems*, 32:115–132, 1999. doi:10.1007/s002240000113.

[3] Joshua Brakensiek, Neng Huang, and Uri Zwick. Tight approximability of MAX 2-SAT and relatives, under UGC. In *Proceedings of SODA 2024*, pages 1328–1344, 2024. doi:10.1137/1.9781611977912.53.

[4] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42:1115–1145, 1995.

[5] Sanjeev Mahajan and Hariharan Ramesh. Derandomizing approximation algorithms based on semidefinite programming. *SIAM Journal on Computing*, 28:1641–1663, 1999.

[6] Luca Trevisan. Max cut and the smallest eigenvalue. In *Proceedings of STOC 2009*, 2009. https://theory.stanford.edu/~trevisan/pubs/maxcutstoc.pdf.

[7] Frank Vega. Paloma: MAX-HORN-2SAT and MAX-CUT formalization. GitHub, https://github.com/frankvegadelgado/paloma, 2026.

[8] Frank Vega. PalomaMax2SAT.UgcBounded.ugc_bounded: Lean 4 theorem and verified proof. Prove2.me, https://prove2.me/theorems/3ca93908-150c-4b65-9e9e-94d9884fe07a, 2026. Lean 4 source in the `ugc-bounded-degree` branch of the repository https://github.com/frankvegadelgado/paloma.

[9] Frank Vega. PalomaMax2SAT.UgcRegular.all_degd: Lean 4 theorem and verified proof. Prove2.me, https://prove2.me/theorems/0e9bbdb8-656d-40f9-ade9-3d9064d4bb23, 2026. Lean 4 source in the `ugc-regular-degree` branch of the repository https://github.com/frankvegadelgado/paloma.

Information Physics Institute, 840 W 67th St, Hialeah, FL 33012, USA
Email address: vega.frank@gmail.com