SIAM J. Comput.
Vol. 47, No. 3, pp. 1039–1055




# ON THE LOVÁSZ THETA FUNCTION FOR INDEPENDENT SETS IN SPARSE GRAPHS*

NIKHIL BANSAL†, ANUPAM GUPTA‡, AND GURU GURUGANESH‡

**Abstract.** We consider the maximum independent set problem on sparse graphs with maximum degree $d$. We show that the Lovász $\vartheta$-function based semidefinite program (SDP) has an integrality gap of $\widetilde{O}(d/\log^{3/2} d)$, improving on the previous best result of $\widetilde{O}(d/\log d)$. This improvement is based on a new Ramsey-theoretic bound on the independence number of $K_r$-free graphs for large values of $r$. We also show that for stronger SDPs, namely, those obtained using polylog($d$) levels of the $SA^+$ semidefinite hierarchy, the integrality gap reduces to $\widetilde{O}(d/\log^2 d)$. This matches the best unique-games-based hardness result up to lower-order poly($\log\log d$) factors. Finally, we give an algorithmic version of this $SA^+$-based integrality gap result, albeit using $d$ levels of $SA^+$, via a coloring algorithm of Johansson.

**Key words.** algorithms, Ramsey theory, semidefinite programmming

**AMS subject classification.** 68R10

**DOI.** 10.1137/15M1051002

**1. Introduction.** Given a graph $G = (V, E)$, an independent set is a subset of vertices $S$ such that no two vertices in $S$ are adjacent. The maximum independent set problem is one of the most well-studied problems in algorithms and graph theory, and its study has led to various remarkable developments such as the seminal result of Lovász [Lov79] in which he introduced the $\vartheta$-function based on semidefinite programming, as well as several surprising results in Ramsey theory and extremal combinatorics.

In general graphs, the problem is notoriously hard to approximate. Given a graph on $n$ vertices, the best known algorithm is due to Feige [Fei04] and achieves an approximation ratio of $\widetilde{O}(n/\log^3 n)$; here $\widetilde{O}(\cdot)$ suppresses some $\log\log n$ factors. On the hardness side, a result of Håstad [Has96] shows that no $n^{1-\varepsilon}$ approximation exists for any constant $\varepsilon > 0$, assuming NP $\not\subseteq$ ZPP. The hardness has been improved more recently to $n/\exp((\log n)^{3/4+\varepsilon})$ by Khot and Ponnuswami [KP06].

In this paper, we focus on the case of bounded-degree graphs, with maximum degree $d$. Recall that the naïve algorithm (that repeatedly picks an arbitrary vertex $v$ and deletes its neighborhood) produces an independent set of size at least $n/(d+1)$ and hence is a $(d+1)$-approximation. The first $o(d)$-approximation was obtained by Halldórsson and Radhakrishnan [HR94], who gave an $O(d/\log\log d)$ guarantee, based on a Ramsey theoretic result of Ajtai et al. [AEKS81]. Subsequently, an $O(d\,\frac{\log\log d}{\log d})$-approximation was obtained independently by several researchers [AK98,

*Received by the editors December 2, 2015; accepted for publication (in revised form) September 8, 2016; published electronically June 26, 2018. Results from this paper have previously appeared in extended abstracts presented at the 2015 ACM-SIAM Symposium on Discrete Algorithms (SODA) and at the 2015 ACM Symposium on the Theory of Computing (STOC).

http://www.siam.org/journals/sicomp/47-3/M105100.html

**Funding:** The first author was supported by NWO grant 639.022.211 and ERC consolidator grant 617951. The second and third authors were partly supported by NSF awards CCF-1016799 and CCF-1319811.

†Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands, and CWI, 1098 XG Amsterdam, The Netherlands (n.bansal@tue.nl).

‡Computer Science Department, Carnegie Mellon University, Pittsburgh, PA 15213 (anupamg@cs.cmu.edu, ggurugan@cs.cmu.edu).

Hal02, Hal00] using the ideas of Karger, Motwani, and Sudan [KMS98] to round the natural semidefinite program (SDP) for the problem, which was itself based on the Lovász $\vartheta$-function.

On the negative side, Austrin, Khot, and Safra [AKS11] showed an $\Omega(d/\log^2 d)$ hardness of approximation, assuming the Unique Games Conjecture. Assuming P $\neq$ NP, a hardness of $d/\log^4 d$ was recently shown by Chan [Cha13]. We remark that these hardness results seem to hold only when $d$ is a constant or a very mildly increasing function of $n$. In fact for $d = \Omega(n)$, the $\Omega(d/\log^2 d)$ hardness of [AKS11] is inconsistent with the known $O(n/\log^3 n)$ approximation [Fei04]. Hence throughout this paper, it will be convenient to view $d$ as being a sufficiently large but fixed constant.

Roughly speaking, the gap between the $\Omega(d/\log^2 d)$-hardness and the $\widetilde{O}(d/\log d)$-approximation arises for the following fundamental reason. Approaches based on the SDP work extremely well if the $\vartheta$-function has value more than $\widetilde{O}(n/\log d)$, but not below this threshold. In order to to show an $\Omega(d/\log d)$-hardness result, at the very least, one needs an instance with SDP value around $n/\log d$ but optimum integral value about $n/d$. While graphs with the latter property clearly exist (e.g., a graph consisting of $n/(d+1)$ disjoint cliques $K_{d+1}$), the SDP value for such graphs seems to be low. In particular, having a large SDP value imposes various constraints on the graph (for example, they cannot contain many large cliques) which might allow the optimum to be nontrivially larger than $n/d$, for example, due to Ramsey-theoretic reasons.

**1.1. Our results.** Our results resolve some of these questions. Our first result considers the integrality gap of the standard SDP relaxation for an independent set (without applying any lift-and-project steps). We show that it is more powerful than the guarantee given by Alon and Kahale [AK98] and Halperin [Hal02].

THEOREM 1.1. *On graphs with maximum degree $d$, the standard $\vartheta$-function-based SDP formulation for the independent set problem has an integrality gap of $\widetilde{O}(d/\log^{3/2} d)$.*[1]

The proof of Theorem 1.1 is nonconstructive; while it shows that the SDP value is within the claimed factor of the optimal IS size, it does not give an efficient algorithm to find such an approximate solution. Finding such an algorithm remains an open question.

The main technical ingredient behind Theorem 1.1 is the following new Ramsey-type result about the existence of large independent sets in $K_r$-free graphs. This builds on a long line of previous results in Ramsey theory (some of which we discuss in section 2) and is of independent interest. (Recall that $\alpha(G)$ is the maximum independent set size in $G$.)

THEOREM 1.2. *For any $r > 0$, if $G$ is a $K_r$-free graph with maximum degree $d$, then*

$$\alpha(G) = \Omega\left(\frac{n}{d} \cdot \max\left(\frac{\log d}{r \log\log d}, \left(\frac{\log d}{\log r}\right)^{1/2}\right)\right). \tag{1}$$

Previously, the best known bound for $K_r$-free graphs was $\Omega(\frac{n}{d}\frac{\log d}{r\log\log d})$ given by Shearer [She95]. Observe the dependence on $r$: when $r \geq \frac{\log d}{\log\log d}$, i.e., when we are only guaranteed to exclude very large cliques, Shearer's result does not give anything

[1] Here and subsequently, $\widetilde{O}(\cdot)$ suppresses poly$(\log\log d)$ factors.

better than the trivial $n/d$ bound. It is in this range of $r \geq \log d$ that the second term in the maximization in (1) starts to perform better and give a nontrivial improvement. In particular, if $G$ does not contain cliques of size $r = O(\log^{3/2} d)$ (which will be the interesting case for Theorem 1.1), Theorem 1.2 gives a bound of $\widetilde{\Omega}(\frac{n}{d}\,(\log d)^{1/2})$. Even for substantially larger values such as $r = \exp(\log^{1-2\varepsilon} d)$, this gives a nontrivial bound of $\widetilde{O}(\frac{n}{d}\,\log^{\varepsilon} d)$.

Improving on Shearer's bound has been a long-standing open problem in the area, and it is conceivable that the right answer for $K_r$-free graphs of maximum degree $d$ is $\alpha(G) \geq \frac{n}{d}\frac{\log d}{\log r}$. This would be best possible, since in section 3.1 we give a simple construction showing an upper bound of $\alpha(G) = O(\frac{n}{d}\frac{\log d}{\log r})$ for $r \geq \log d$, which to the best of our knowledge is the smallest upper bound currently known. The gap between our lower bound and this upper bound remains an intriguing one to close; in fact it follows from our proof of Theorem 1.1 that such a lower bound would imply an $\widetilde{O}(d/\log^2 d)$ integrality gap for the standard SDP. Alon [Alon96] shows that this bound is achievable under the stronger condition that the neighborhood of each vertex is $(r-1)$-colorable.

The next set of results consider using lift-and-project techniques to address the approximability of the problem. We consider the standard LP formulation for the independent set problem strengthened by $\ell$ levels of the Sherali–Adams hierarchy, together with semidefinite constraints at the first level (see section 2 for details). We will refer to this as $\ell$ levels of the mixed hierarchy (this is also referred to as the $SA^+$ hierarchy) and denote this relaxation by $SA^+_{(\ell)}$. Our first result is the following.

THEOREM 1.3. *The value of the $O(\log^4 d)$-level $SA^+$ semidefinite relaxation has an integrality gap of $O(d(\log\log d)^2/\log^2 d)$.*

The main observation behind this result is that as the $SA^+$ relaxation specifies a local distribution on independent sets, and if the relaxation has high objective value then it must be that any polylog$(d)$ size subset of vertices $X$ must contain a large independent subset. We can then use a result of Alon [Alon96], in turn based on the above-mentioned result of Shearer [She95], to show that such graphs have non-trivially large independents sets.

Unfortunately, Alon's argument is non-algorithmic; it shows that the lifted SDP has a small integrality gap, but does not give a corresponding approximation algorithm with running time sub-exponential in $n$. Our next result makes this integrality gap result algorithmic, although at the expense of more levels and a higher running time.

THEOREM 1.4. *There is an $\widetilde{O}(d/\log^2 d)$-approximation algorithm with running time*[2] $\mathrm{poly}(n) \cdot 2^{O(d)}$*, based on rounding a $d$-level $SA^+$ semidefinite relaxation.*

The improvement is simple and is based on bringing the right tool to bear on the problem: instead of using the nonconstructive argument of Alon [Alon96], we use an ingenious and remarkable (and stronger) result of Johansson [Joh96b], who shows that the list-chromatic number of such locally-colorable graphs is $\chi_\ell(G) = O(d\frac{\log k}{\log d})$. His result is based on a very clever application of the Rödl "nibble" method, together with the Lovász Local Lemma to tightly control the various parameters of the process at every vertex in the graph. Applying Johansson's result to our problem gives us the desired algorithm.

[2]While a $d$-level $SA^+$ relaxation has size $n^{O(d)}$ in general, our relaxation only uses variables corresponding to subsets of vertices that lie in the neighborhood of some vertex $v$ and thus has $n \cdot 2^{O(d)}$ variables. We comment on this in section 4.2.

Unfortunately, Johansson's preprint (back from 1996) was never published. (We thank Alan Frieze for sharing a copy with us.) While his main results are summarized in Appendix A, we give the proof in its entirety in the arXiv version of this paper [BGG15] for completeness and to facilitate verification. Our presentation closely follows his but streamlines some arguments based on recent developments such as concentration bounds for low-degree polynomials of random variables, and the algorithmic version of the Lovász Local Lemma. (Johansson's previous preprint [Joh96a] showing the analogous list-coloring result for triangle-free graphs is also unavailable publicly but is presented in the graph coloring book by Molloy and Reed [MR02] and has received considerable attention since, in both the math [AKS99, Vu02, FM13] and computer science communities [GP00, CPS14].)

Finally, the proof of Theorem 1.4 also implies the following new results about the LP-based Sherali–Adams ($SA$) hierarchies, without any SDP constraints.

COROLLARY 1.5. *The LP relaxation with clique constraints on sets of size up to* $\log d$ *(and hence the relaxation* $SA_{(\log d)}$*) has an integrality gap of* $\widetilde{O}(d/\log d)$*. Moreover, the relaxation* $SA_{(d)}$ *can be used to find an independent set achieving an* $\widetilde{O}(d/\log d)$ *approximation in time* $\mathrm{poly}(n) \cdot 2^{O(d)}$*.*

Since LP-based relaxations have traditionally been found to be very weak for the independent set problem, it may be somewhat surprising that a few rounds of the Sherali–Adams hierarchy improves the integrality gap by a nontrivial amount.

Theorem 1.2, our Ramsey-theoretic result, extends to the case when $d$ is the average degree of the graph by first deleting the (at most $n/2$) vertices with degree more than $2\overline{d}$ and then applying the results. In Appendix C, we show that weaker versions of our SDP-based approximation results hold when $d$ is replaced by the average degree instead of the maximum degree. Moreover, we show that the loss in approximation ratio when going from max-degree to average degree is inherent.

## 2. Preliminaries.

Given the input graph $G = (V, E)$, we will denote the vertex set $V$ by $[n] = \{1, \ldots, n\}$. Let $\alpha(G)$ denote the size of a maximum independent set in $G$, and let $d$ denote the maximum degree in $G$. The naïve greedy algorithm implies $\alpha(G) \geq n/(d+1)$ for every $G$. As the greedy guarantee is tight in general (e.g., if the graph is a disjoint union of $n/(d+1)$ copies of the clique $K_{d+1}$), the trivial upper bound of $\alpha(G) \leq n$ cannot give an approximation better than $d+1$ and hence stronger upper bounds are needed. A natural bound is the clique-cover number $\overline{\chi}(G)$, defined as the minimum number of vertex-disjoint cliques needed to cover $V$. As any independent set can contain at most one vertex from any clique, $\alpha(G) \leq \overline{\chi}(G)$.

**Standard LP/ SDP relaxations.** In the standard LP relaxation for the independent set problem, there is variable $x_i$ for each vertex $i$ that is intended to be 1 if $i$ lies in the independent set and 0 otherwise. The LP is the following:

$$\max \sum_i x_i \quad \text{s.t.} \quad x_i + x_j \leq 1 \quad \forall (i,j) \in E \quad \text{and} \quad x_i \in [0,1] \quad \forall i \in [n]. \tag{2}$$

Observe that this linear program is very weak and cannot give an approximation better than $(d+1)/2$: even if the graph consists of $n/(d+1)$ copies of $K_{d+1}$, the solution $x_i = 1/2$ for each $i$ is a feasible one.

In the standard SDP relaxation, there is a special unit vector $v_0$ (intended to indicate 1) and a vector $v_i$ for each vertex $i$. The vector $v_i$ is intended to be $v_0$ if $i$

lies in the independent set and be **0** otherwise. This gives the following relaxation:

(3)
$$\max \sum_i v_i \cdot v_0 \;\text{ s.t. }\; v_0 \cdot v_0 = 1, \;\; v_0 \cdot v_i = v_i \cdot v_i \;\; \forall i \in [n], \;\text{ and }\; v_i \cdot v_j = 0 \;\; \forall (i,j) \in E.$$

Let $X$ denote the $(n+1) \times (n+1)$ Gram matrix with entries $x_{ij} = v_i \cdot v_j$ for $i, j \in \{0, \dots, n\}$. Then we have the equivalent relaxation

(4)
$$\max \sum_i x_{0i} \;\text{ s.t. }\; x_{00} = 1, \;\; x_{0i} = x_{ii} \;\; \forall i \in [n], \;\; x_{ij} = 0 \;\; \forall (i,j) \in E \;\text{ and }\; X \succeq 0.$$

The above SDP, which is equivalent to the well-known $\vartheta$-function of Lovász (see, e.g., [Lau, Lemma 3.4.4]), satisfies $\alpha(G) \le \vartheta(G) \le \overline{\chi}(G)$. The $O(d\frac{\log\log d}{\log d})$ approximations due to [AK98, Hal02, Hal00] are all based on this SDP. Indeed, we use the following important result due to Halperin [Hal02] about the performance of the SDP.

THEOREM 2.1 (Halperin [Hal02, Lemma 5.2]). *Let $\eta \in [0, \frac{1}{2}]$ be a parameter and let $Z$ be the collection of vectors $v_i$ satisfying $\|v_i\|^2 \ge \eta$ in the SDP solution. Then there is an algorithm that returns an independent set of size $\Omega(\frac{d^{2\eta}}{d\sqrt{\ln d}}|Z|)$.*

The statement above differs slightly from the one in [Hal02] since Halperin works with a $\{-1, 1\}$ formulation; a proof of its equivalence appears in Appendix B. Note that if $\eta = c \log\log d / \log d$, then for $c \le 1/4$ Theorem 2.1 does not return any non-trivial independent set. On the other hand, for $c \ge 1/4$ the size of the independent set returned rises exponentially fast with $c$.

For more details on SDPs, and the Lovász $\vartheta$-function, we refer the reader to [GLS88, GM12].

**Lower bounds on the independence number.** As SDPs can handle cliques, looking at $\vartheta(G)$ naturally leads to Ramsey theoretic considerations. In particular, if $\vartheta(G)$ is small, then the trivial $n/(d+1)$ solution already gives a good approximation. Otherwise, if $\vartheta(G)$ is large, then this essentially means that there are no large cliques and one must argue that a large independent set exists (and can be found efficiently).

For bounded degree graphs, a well-known result of this type is that $\alpha(G) = \Omega(n\frac{\log d}{d})$ for triangle-free graphs [AKS80, She83] (i.e., if there are no cliques of size 3). A particularly elegant proof (based on an idea due to Shearer [She95]) is in [AS92]. Moreover this bound is tight, and simple probabilistic constructions show that this bound cannot be improved even for graphs with large girth.

For the case of $K_r$-free graphs with $r \ge 4$, the situation is less clear. Ajtai et al. [AEKS81] showed that $K_r$-free graphs have $\alpha(G) = \Omega(n(\log(\log d/r))/d)$, which implies that $\alpha(G) = \Omega(n \log\log d/d)$ for $r \ll \log d$. This result was the basis of the $O(d/\log\log d)$ approximation due to [HR94]. Shearer [She95] improved this result substantially and showed that

$$\alpha(G) = \Omega\left(\frac{1}{r}\frac{n}{d}\frac{\log d}{\log\log d}\right)$$

for $K_r$-free graphs. This result is based on an elegant entropy-based approach that has subsequently found many applications. However, it is not known how to make this method algorithmic. This bound still seems far from optimum. In particular, it is possible that the dependence on $r$ could be $\log r$. Note that the above bound is trivial when $r \geq \frac{\log d}{\log \log d}$. For constant $r$, in particular $r = 4$, it is also an important open question whether the $\log \log d$ factor above can be removed.

Interestingly, Shearer's bound also implies another (nonalgorithmic) proof that the SDP has integrality gap $O((d \log \log d)/\log d)$. To see this, suppose the SDP objective is $n/r$. This essentially implies that the graph is $K_r$-free as roughly each vertex contributes about $x_i = 1/r$ (formally, one can delete the vertices with $x_i \leq 1/(2r)$ and consider the residual graph). Then the integrality gap is $(n/r)/\alpha(G)$, which by Shearer's bounds is at least $(d \log \log d)/\log d$. It is interesting to note that both Halperin's approach and Shearer's bound seem to get stuck at the same point.

Alon [Alon96] generalized the triangle-free result in a different direction, also using the entropy method. He considered locally $k$-colorable graphs, where the neighborhood of every vertex is $k$-colorable, and showed that $\alpha(G) = \Omega(\frac{n}{d}\frac{\log d}{\log k+1})$. Note that triangle-free graphs are locally 1-colorable. This result also holds under weaker conditions and plays a key role in bounding the integrality gap of $SA^+$ relaxations.

**Bounds on the chromatic number.** Many of the above results also generalize to the much more demanding setting of list coloring. All of them are based on the "nibble" method but require increasingly sophisticated ideas. In particular these results give a bound of $\tilde{\Omega}_r(d/\log d)$ on the list chromatic number of $K_r$-free or locally $r$-colorable graphs. The intuition for why $O(d/\log d)$ arises can be seen via a coupon-collector argument: if each vertex in the neighborhood $N(v)$ chooses a color from $s$ colors independently and uniformly at random, they will use up all $s$ colors unless $d \leq O(s \log s)$, or $s \geq \Omega(d/\log d)$. (Of course, the colors at the neighbors are not chosen uniformly or independently which substantially complicates the arguments.) Kim showed that $\chi_\ell(G) = O(d/\log d)$ for graphs with girth at least 5 [Kim95]. His idea was that for any $v$, and $u, w \in N(v)$, $N(u) \cap N(w) = \{v\}$ because of the girth, and hence the available colors at $u, w$ evolve essentially independently and hence conform to the intuition.

These ideas fail for triangle-free graphs (of girth 4): we could have a vertex $v$, with $u, w \in N(v)$, and $N(u) = N(w)$ (i.e., all their neighbors are common). In this case the lists of available colors at $u$ and $w$ are far from independent: they would be *completely identical*. Johansson [Joh96a] had the crucial insight that this positive correlation is not a problem, since there is no edge between $u$ and $w$ (because of triangle-freeness!). His clever proof introduced the crucial notions of entropy and energy to capture and control the positive correlation along edges in such $K_3$-free graphs.

If there are triangles, say, if the graphs are only locally $k$-colorable, then using these ideas naïvely fails. Another key idea, also introduced by Johansson [Joh96b], is to actually modify the standard nibble process by introducing a *probability reshuffling* step at each vertex depending on its local graph structure, which makes it more complicated. In [BGG15], we give his result for locally-colorable and for $K_r$-free graphs in its entirety.

**Lift-and-project hierarchies.** An excellent introduction to hierarchies and their algorithmic uses can be found in [CT12, Lau03]. Here, we only describe the most basic facts needed for this paper.

The Sherali–Adams hierarchy defines a hierarchy of linear programs with increasingly tighter relaxations. At level $t$, there is a variable $Y_S$ for each subset $S \subseteq [n]$ with $|S| \leq t+1$. Intuitively, one views $X_S$ as the probability that all the variables in $S$ are set to 1. Such a solution can be viewed as specifying a local distribution over valid $\{0,1\}$-solutions for each set $S$ of size at most $t+1$. A formal description of the

$t$-round Sherali–Adams LP $SA_{(t)}$ for the independent set problem can be found in [CT12, Lemma 1].

More formally, for the independent set problem we have the following theorem from [CT12].

THEOREM 2.2 (see [CT12, Lemma 1]). *Consider a family of distributions $\{\mathcal{D}(S)\}_{S\subseteq[n]:|S|\le t+2}$, where each $\mathcal{D}(S)$ is defined over $\{0,1\}^S$. If the distributions satisfy*

1. *for all $(i,j)\in E$ and $S\supseteq\{i,j\}$, it holds that $\Pr_{\mathcal{D}(S)}[(x_i=1)\cap(x_j=1)]=0$, and*
2. *for all $S'\subseteq S\subset[n]$ with $|S|\le t+1$, the distribution $\mathcal{D}(S'),\mathcal{D}(S)$ agree on $S'$,*

*then $X_S=\mathrm{P}_{\mathcal{D}(S)}[\wedge_{i\in S}(x_i=1)]$ is a feasible solution for the level-$t$ Sherali–Adams relaxation.*

*Conversely, for any feasible solution $\{X'_S\}$ for the level-$(t+1)$ Sherali–Adams relaxation, there exists a family of distributions satisfying the above properties, as well as $X_{S'}=\mathrm{P}_{\mathcal{D}(S)}[\wedge_{i\in S'}(x_i=1)]=X'_{S'}$ for all $S'\subseteq S\subset[n]$ such that $|S|\le t+1$.*

Here, condition 1 implies that for a subset of vertices $S$ with $|S|\le t+1$, the local-distribution $\mathcal{D}(S)$ has support on the valid independent sets in the graph induced on $S$, and condition 2 guarantees that different local distributions induce a consistent distribution on the common elements.

For our purposes, we will also impose the positive semidefinite constraint on the variables $x_{ij}$ at the first level (i.e., we add the constraints in (4) on $x_{ij}$ variables). We will call this the $t$-level $SA^+$ formulation and denote it by $SA^+_{(t)}$. Such a solution specifies values $x_S$ for multisets $S$ with $|S|\le\ell+1$. To keep the notation consistent with the LP (2), we will use $x_i$ to denote the marginals $x_{ii}$ on singleton vertices.

## 3. Integrality gap of the standard SDP.

In this section, we show Theorem 1.1, that the integrality gap of the standard Lovász $\vartheta$-function based SDP relaxation is

$$O\left(d\left(\frac{\log\log d}{\log d}\right)^{3/2}\right)=\widetilde{O}\big(d/\log^{3/2}d\big).$$

To show this we prove the following result (which is Theorem 1.2, restated).

THEOREM 3.1. *Let $G$ be a $K_r$-free graph with maximum degree $d$. Then*

$$\alpha(G)=\Omega\left(\frac{n}{d}\max\left(\frac{\log d}{r\log\log d},\left(\frac{\log d}{\log r}\right)^{1/2}\right)\right).$$

*In particular, for $r=\log^c d$ with $c\ge1$, we get $\alpha(G)=\Omega\big(\frac{n}{d}\big(\frac{\log d}{c\log\log d}\big)^{1/2}\big)$.*

We need the following basic facts. The first follows from a simple counting argument (see [Alon96, Lemma 2.2] for a proof).

LEMMA 3.2. *Let $F$ be a family of $2^{\varepsilon x}$ distinct subsets of an $x$-element set $X$. Then the average size of a member of $F$ is at least $\varepsilon x/(10\log(1+1/\varepsilon))$.*

FACT 3.3. *Let $G$ be a $K_r$-free graph on $x$ vertices; then*

$$\alpha(G)\ge\max\left(\frac{x^{1/r}}{2},\frac{\log x}{\log(2r)}\right).$$

*Note that the latter bound is stronger when $r$ is large, i.e., roughly when $r\ge\log x/\log\log x$.*

*Proof.* Let $R(s,t)$ denote the off-diagonal $(s,t)$-Ramsey number, defined as the smallest number $n$ such that any graph on $n$ vertices contains either an independent set of size $s$ or a clique of size $t$.

It is well known that $R(s,t) \le \binom{s+t-2}{s-1}$ [ES35]. Approximating the binomial gives us the bounds $R(s,t) \le (2s)^t$ and $R(s,t) \le (2t)^s$; the former is useful for $t \le s$ and the latter for $s \le t$. If we set $R(s,t) = x$ and $t = r$, the first bound gives $s \ge (1/2)x^{1/r}$ and the second bound gives $s \ge \log x / \log(2r)$. ◻

We will be interested in lower bounding the number of independent sets $\mathcal{I}$ in a $K_r$-free graph. Clearly, $\mathcal{I} \ge 2^{\alpha(G)}$ (consider every subset of a maximum independent set). However the following improved estimate will play a key role in Theorem 3.1. Roughly speaking it says that if $\alpha(G)$ is small, in particular of size logarithmic in $x$, then the independent sets are spread all over $G$, and hence their number is close to $x^{\Omega(\alpha(G))}$.

THEOREM 3.4. *Let $G$ be a $K_r$-free graph on $x$ vertices, and let $\mathcal{I}$ denote the number of independent sets in $G$. Then we have*

$$\log \mathcal{I} \ge \max\left(\frac{x^{1/r}}{2}, \frac{\log^2 x}{6 \log 2r}\right).$$

*Proof.* The first bound follows trivially from Fact 3.3, and hence we focus on the second bound. Also, assume $r \ge 3$ and $x \ge 64$, else the second bound is trivial.

Define $s := \log x / \log(2r)$. Let $G'$ be the graph obtained by sampling each vertex of $G$ independently with probability $p := 2/x^{1/2}$. The expected number of vertices in $G'$ is $px = 2x^{1/2}$. Let $\mathcal{G}$ denote the good event that $G'$ has at least $x^{1/2}$ vertices. Clearly, $Pr[\mathcal{G}] \ge 1/2$ (in fact it is exponentially close to 1). Since the graph $G'$ is also $K_r$-free, conditioned on the event $\mathcal{G}$, by Fact 3.3 it has an independent set of size at least $\log(x^{1/2})/\log(2r) = s/2$. Thus the expected number of independent sets of size $s/2$ in $G'$ is at least $1/2$.

Now consider some independent set $Y$ of size $s/2$ in $G$. The probability that $Y$ survives in $G'$ is exactly $p^{s/2}$. As the expected number of independent sets of size $s/2$ in $G'$ is at least $1/2$, it follows that $G$ must contain at least $(1/2)(1/p^{s/2})$ independent sets of $s/2$. This gives us that

$$\log \mathcal{I} \ge \frac{s}{2}\log\left(\frac{1}{p}\right) - 1 \ge \frac{s}{2}\log x^{1/2} - \frac{s}{2} - 1 \ge \frac{s}{6}\log x,$$

where the last inequality assumes that $x$ is large enough. ◻

We are now ready to prove Theorem 3.1.

*Proof.* We can assume that $d \ge 16$, else the claim is trivial. Our arguments follow the probabilistic approach of [She95, Alon96]. Let $W$ be a random independent set of vertices in $G$, chosen uniformly among all independent sets in $G$. For each vertex $v$, let $X_v$ be a random variable defined as $X_v = d|\{v\} \cap W| + |N(v) \cap W|$.

Observe that $|W|$ can be written as $\sum_v |v \cap W|$; moreover, it satisfies $|W| \ge (1/d)\sum_v |N(v) \cap W|$, since a vertex in $W$ can be in at most $d$ sets $N(v)$. Hence we have that

$$|W| \ge \frac{1}{2d}\sum_v X_v.$$

Let $\gamma = \max\left(\frac{\log d}{r\log\log d}, \left(\frac{\log d}{\log r}\right)^{1/2}\right)$ denote the improvement factor in Theorem 3.1 over the trivial bound of $n/d$. Thus to show that $\alpha(G)$ is large, it suffices to show that

$$\mathbb{E}[X_v] \geq c\gamma \tag{5}$$

for each vertex $v$ and some fixed constant $c$.

In fact, we show that (5) holds for every conditioning of the choice of the independent set in $V - (N(v) \cup \{v\})$. In particular, let $H$ denote the subgraph of $G$ induced on $V - (N(v) \cup \{v\})$. For each possible independent set $S$ in $H$, we will show that

$$\mathbb{E}[X_v \mid W \cap V(H) = S] \geq c\gamma.$$

Fix a choice of $S$. Let $X$ denote the nonneighbors of $S$ in $N(v)$, and let $x = |X|$. Let $\varepsilon$ be such that $2^{\varepsilon x}$ denotes the number of independent sets in the induced subgraph $G[X]$. Now, conditioning on the intersection $W \cap V(H) = S$, there are precisely $2^{\varepsilon x} + 1$ possibilities for W: one in which $W = S \cup \{v\}$ and $2^{\varepsilon x}$ possibilities in which $v \notin W$ and $W$ is the union of $S$ with an independent set in $G[X]$.

By Lemma 3.2, the average size of an independent set in $X$ is at least $\frac{\varepsilon x}{10\log(1/\varepsilon+1)}$ and thus we have that

$$\mathbb{E}[X_v \mid W \cap V(H) = S] \geq d\frac{1}{2^{\varepsilon x}+1} + \frac{\varepsilon x}{10\log(1/\varepsilon+1)}\frac{2^{\varepsilon x}}{2^{\varepsilon x}+1}. \tag{6}$$

Now, if $2^{\varepsilon x} + 1 \leq \sqrt{d}$, then the first term is at least $\sqrt{d}$, and we've shown (5) with room to spare. So we can assume that $\varepsilon x \geq (1/2)\log d$. Moreover, by Theorem 3.4,

$$\varepsilon x \geq \max\left(\frac{x^{1/r}}{2}, \frac{\log^2 x}{6\log(2r)}\right)$$

and hence the right-hand side in (6) is at least

$$\begin{aligned} &\frac{1}{20\log(1/\varepsilon+1)}\max\left(\frac{\log d}{2}, \frac{x^{1/r}}{2}, \frac{\log^2 x}{6\log 2r}\right) \\ &\geq \frac{1}{20\log(x+1)}\max\left(\frac{\log d}{2}, \frac{x^{1/r}}{2}, \frac{\log^2 x}{6\log 2r}\right), \end{aligned} \tag{7}$$

where the inequality uses $\varepsilon \geq 1/x$ (since $\varepsilon x \geq (1/2)\log d \geq 1$).

First, let's consider the first two expressions in (7). If $x \geq \log^r d$, then as $x^{1/r}/\log(x+1)$ is increasing in $x$, it follows that the right-hand side of (7) is at least

$$\frac{x^{1/r}}{40\log(x+1)} = \Omega\left(\frac{\log d}{r\log\log d}\right).$$

On the other hand if $x \leq \log^r d$, then we have that the right-hand side is again at least

$$\frac{1}{20\log(x+1)}\frac{\log d}{2} = \Omega\left(\frac{\log d}{r\log\log d}\right).$$

Now, consider the first and third expressions in (7). Using the fact that $\max(a,b) \geq \sqrt{ab}$ with $a = (\log d)/2$ and $b = (\log^2 x)/(6\log 2r)$, we get that (7) is at least $\Omega(\frac{\log d}{\log r})^{1/2}$. Hence, for every value of $x$ we get that (7) is at least $\Omega(\gamma)$ as desired in (5); this completes the proof of Theorem 3.1. ∎

We can now show the main result of this section.

THEOREM 3.5. *The standard SDP for an independent set has an integrality gap of*

$$O\left(d\left(\frac{\log\log d}{\log d}\right)^{3/2}\right).$$

*Proof.* Given a graph $G$ on $n$ vertices, let $\beta \in [0,1]$ be such that the SDP on $G$ has objective value $\beta n$. If $\beta \le 2/\log^{3/2} d$, the naïve greedy algorithm already implies a $d/\log^{3/2} d$ approximation. Thus, we will assume that $\beta \ge 2/\log^{3/2} d$. Recall we use $x_i$ to denote the marginals $x_{ii}$ on singleton vertices in the SDP.

Let us delete all the vertices that contribute $x_i \le \beta/2$ to the objective. The residual graph has objective value at least $\beta n - (\beta/2)n = \beta n/2$.

Let $\eta = 2\log\log d/\log d$. If there are more than $n/\log^2 d$ vertices with $x_i \ge \eta$, applying Theorem 2.1 to the collection of these vertices already gives an independent set of size at least

$$\Omega\left(\frac{d^{2\eta}}{d\sqrt{\ln d}}\cdot\frac{n}{\log^2 d}\right) = \Omega\left(\frac{n\log^{3/2} d}{d}\right)$$

and hence a $O(d/\log^{3/2} d)$ approximation.

Thus we can assume that fewer than $n/\log^2 d$ vertices have $x_i \ge \eta$. As each vertex can contribute at most 1 to the objective, the SDP objective on the residual graph obtained by deleting the vertices with $x_i \ge \eta$ is at least $\beta n/2 - n/(\log^2 d)$ which is at least $\beta n/3$, since $\beta \ge 2/\log^{3/2} d$.

So we have a feasible SDP solution on a subgraph $G'$ of $G$, where the objective is at least $\beta n/3$ (here $n$ is the number of vertices in $G$ and not $G'$) and each surviving vertex $i$ has value $x_i$ in the range $[\beta/2, \eta]$.

As $x_i \le \eta$ for each $i$, and the SDP objective is at least $\beta n/3$, the number of vertices $n'$ in $G'$ satisfies $n' \ge (\beta n/3)/\eta = \Omega(n\beta/\eta)$. Moreover, as $x_i \ge \beta/2$ for each vertex $i \in G'$, and the SDP does not put more than one unit of probability mass on any clique, it follows that $G'$ is $K_r$-free for $r = 2/\beta = \log^{3/2} d$. Applying Theorem 3.1 to $G'$ with parameter $r = \log^{3/2} d$, we obtain that $G'$ has an independent set of size

$$\Omega\left(\frac{n'}{d}\sqrt{\frac{\log d}{\log r}}\right) = \Omega\left(\frac{n'}{d}\sqrt{\frac{\log d}{\log\log d}}\right) = \Omega\left(\frac{n\,\beta}{d\,\eta}\sqrt{1/\eta}\right) = \Omega\left(\frac{\beta n}{d}\cdot\eta^{-3/2}\right).$$

The SDP objective for $G$ was $\beta n$, so the integrality gap is $O(d\eta^{3/2}) = O(d(\frac{\log\log d}{\log d})^{3/2})$. ∎

### 3.1. An upper bound.

LEMMA 3.6. *There exists $K_r$-free graphs $G$ with maximum degree $d$ such that $\alpha(G) \le O\big(\frac{n}{d}\frac{\log d}{\log r}\big)$ whenever $r \ge \log d$.*

*Proof.* We use the standard lower bound $R(s,t) = \Omega\big(\big(\frac{t}{\log t}\big)^{s/2}\big)$ for off-diagonal Ramsey numbers for $t \ge s$. While stronger lower bounds exist (see Theorem 1.2 in [BK10]), this one suffices for the lemma. Setting $t = r$ and $s = O(\log d/(\log r - \log\log r))$, it follows that there exist $K_r$-free graphs $H$ on $d$ vertices such that $\alpha(H) = O(\log d/(\log r - \log\log r))$ whenever $t \ge s$. We now apply two simplifications. First, as $\log d \ge s$, we can simplify the condition $t \ge s$ to be $r \ge \log d$. Second, as $\log r \ge 2\log\log r$ for all $r > 1$, we can say $\alpha(H) = O\big(\frac{\log d}{\log r}\big)$. Setting $G$ to be $n/d$ disjoint copies of $H$ completes the lemma. ∎

**4. An algorithm using lift-and-project.** In this section, we give our results bounding the integrality gap of the $SA^+$ mixed hierarchy. We first show that the $O(\log^4 d)$-level $SA^+$ relaxation has an integrality gap of $\tilde{O}(d/\log^2 d)$. This result, however, does not give an effective procedure to find such a large independent set. Then we show how to round a vector solution to the $d$-level $SA^+$ relaxation to get an independent set of size at least an $\tilde{O}(d/\log^2 d)$ factor of the optimal independent set.

Consider the $SA^+_{(t)}$ relaxation on $G$; for the subsequent results we choose the value of $t$ to be $O(\log^4 d)$ and $d$, respectively. Let $\mathsf{sdp}_t(G)$ denote its value. We can assume that

$$\mathsf{sdp}_t(G) \geq n/\log^2 d; \tag{8}$$

otherwise the naïve greedy algorithm already gives an $O(d/\log^2 d)$ approximation.

Let $\eta = 3 \log\log d/\log d$, and let $Z$ denote the set of vertices $i$ with $x_i \geq \eta$. We can assume that $|Z| \leq n/(4\log^2 d)$; otherwise applying Theorem 2.1 gives an independent set of size $\Omega(|Z|\cdot d^{2\eta}/(d\sqrt{\log d})) = \Omega(n \log^2 d/d)$. Note that we can apply Theorem 2.1, since our solution belongs to $SA^+$ and hence is a valid SDP solution. Hence,

$$\mathsf{sdp}_t(G) \leq |Z| \cdot 1 + (n - |Z|) \cdot \eta \leq \left(n/\left(4\log^2 d\right)\right) \cdot 1 + n \cdot \eta \leq 2\eta n.$$

Let $V'$ denote the set of vertices $i$ with $x_i \in [1/(4\log^2 d), \eta]$, and let $G' = G[V']$ be the graph induced on these vertices.

Claim 4.1. $|V'| \geq \mathsf{sdp}_t(G)/(2\eta)$.

*Proof.* The total contribution to $\mathsf{sdp}_t(G)$ of vertices $i$ with $x_i \leq 1/(4\log^2 d)$ can be at most $n/(4\log^2 d)$, which by (8) is at most $\mathsf{sdp}_t(G)/4$. Similarly, the contribution of vertices in $Z$ is at most $|Z|$, which is again at most $\mathsf{sdp}_t(G)/4$. Together this gives $\mathsf{sdp}_t(G') \geq \mathsf{sdp}_t(G)/2$. As each vertex in $V'$ has $x_i \leq \eta$, the claim follows. ∎

Lemma 4.2. *Let $t \geq \log^4 d$. For the graph $G' = G[V']$, let $v \in V'$ and let $T \subseteq N_{G'}(v)$ be a subset of neighbors of $v$ in $G'$ having size at most $t$.*

(a) *$T$ contains an independent set of size at least $|T|/\log^2 d$.*

(b) *The induced subgraph $G'[T]$ is $O(\log^3 d)$-colorable.*

*Proof.* Consider the solution $SA^+_{(t)}$ restricted to $G'$. Since $|T| \leq t$ and $x_i \geq 1/(4\log^2 d)$ for all $i \in N_{G'}(v)$, we use Theorem 2.2 to deduce that the $SA^+_{(t)}$ solution defines a "local distribution" $\{X_S\}_{S\subseteq T}$ over subsets of $T$ with the following properties:

(i) $X_S \geq 0$ and $\sum_{S\subseteq T} X_S = 1$,

(ii) $X_S > 0$ only if $S$ is independent in the induced subgraph $G'[T]$ (and hence in $G$), and

(iii) for each vertex $i \in T$, it holds that

$$x_i = \sum_{S\subseteq T: i\in S} x_S \geq 1/(4\log^2 d).$$

Now scaling up the solution $\{X_S\}$ by $4\log^2 d$ gives a valid fractional coloring of $T$ using $4\log^2 d$ colors. This means at least one of the color classes must have size at least $|T|/(4\log^2 d)$. This proves (a).

To prove (b), we can use a set-covering argument. The fractional coloring can be viewed as a fractional set cover of $T$, where the sets are all independent sets in $G$.

The (fractional) number of sets used is $4\log^2 d$. Now the integrality gap of the LP relaxation of set cover implies that we can cover $T$ using at most $4\log^2 d \cdot O(\log|T|) = O(\log^3 d)$. ∎

**4.1. The integrality gap result.** We now bound the integrality gap of the $O(\log^4 d)$-round $SA^+$ relaxation. The following Ramsay-theoretic result will be crucial.

THEOREM 4.3 (Alon [Alon96, Theorem 1.1]). *Let $H = (V, E)$ be a graph on $n$ vertices with maximum degree $d \geq 1$ such that for every vertex $v \in V$ the induced subgraph on the set of all neighbors of $v$ is $k$-colorable. Then,*

$$\alpha(H) \geq \Omega\left(\frac{n}{d}\frac{\log d}{\log(k+1)}\right).$$

To reduce the number of rounds of $SA^+$, we use a version of the result above that holds under a considerably weaker condition.

THEOREM 4.4 (Alon [Alon96]). *Let $H = (V, E)$ be a graph on $n$ vertices with maximum degree $d$, and let $k \geq 1$ be an integer. If for every vertex $v$ and every subset $T \subset N(v)$ with $|T| \leq k\log^2 d$, it holds that the subgraph induced on $T$ has an independent set of size at least $|T|/k$, then*

$$\alpha(H) \geq \Omega\left(\frac{n}{d}\frac{\log d}{\log(k+1)}\right).$$

For completeness, a proof of Theorem 4.4 can be found in Appendix B.2.

Now consider the graph $G' = G[V']$. By Lemma 4.2(a) with the parameter $t = \log^2 d$, the graph $G'$ satisfies the requirements in Theorem 4.4 with $k = \log^2 d$; that theorem gives us that

$$\alpha(G') \geq \Omega\left(\frac{|V'|}{d}\frac{\log d}{\log(\log^2 d)}\right).$$

Finally, using Claim 4.1, the integrality gap is

$$\frac{\mathsf{sdp}_{(\log^4 d)}(G)}{\alpha(G)} \leq \frac{\mathsf{sdp}_{(\log^4 d)}(G)}{\alpha(G')} \leq O\left(\frac{d\eta\log(\log^2 d)}{\log d}\right) = \widetilde{O}\left(\frac{d}{\log^2 d}\right),$$

This completes the proof of Theorem 1.3.

**4.2. The algorithmic result.** To get an algorithm, note that Lemma 4.2(b) with $t = d$ implies that the neighborhood of each vertex in $G'$ is $O(\log^3 d)$ colorable. In other words, $G'$ is locally $k$-colorable for $k = O(\log^3 d)$. We now use Johansson's coloring algorithm for locally $k$-colorable graphs (Theorem A.1) to find an independent set of $G'$ with size

$$\mathsf{alg}(G') = \Omega\left(\frac{|V'|}{d}\cdot\frac{\log d}{\log(k+1)}\right).$$

Using $k = O(\log^3 d)$ and Claim 4.1 this implies an algorithm to find independent sets in degree $d$ graphs, with an integrality gap of

$$\frac{\mathsf{sdp}_d(G)}{\mathsf{alg}(G)} \leq \frac{\mathsf{sdp}_d(G)}{\mathsf{alg}(G')} \leq O\left(\frac{d\eta\log(k+1)}{\log d}\right) = \widetilde{O}\left(\frac{d}{\log^2 d}\right).$$

Our algorithm only required a fractional coloring on the neighborhood of vertices. Since there are at most $2^d$ independent sets in each neighborhood, there are at most $n \cdot 2^d$ relevant variables in our SDP. Hence, we can compute the relevant fractional coloring in time $\mathrm{poly}(n) \cdot 2^{O(d)}$. This completes the proof of Theorem 1.4.

**5. LP-based guarantees.** We prove Corollary 1.5, showing the integrality gap of the Sherali–Adams hierarchy (without the SDP constraints).

*Proof.* Consider the standard LP (2) strengthened by the clique inequalities $\sum_{i \in C} x_i \leq 1$ for each clique $C$ with $|C| \leq \log d$. As each clique lies in the neighborhood of some vertex, the number of such cliques is at most $n \cdot \binom{d}{\log d}$. Let $\beta n$ denote the objective value of this LP relaxation. We assume that $\beta \geq 2/\log d$; otherwise the naïve algorithm already gives a $d/\log d$ approximation.

Let $B_0$ denote the set of vertices with $x_i \leq 1/\log d = \beta/2$. For $j = 1, \ldots, k$, where $k = \log\log d$, let $B_j$ denote the set of vertices with $x_i \in (2^{j-1}/\log d, 2^j/\log d]$. Note that $\sum_{j \geq 1} \sum_{i \in B_j} x_i = \beta n - \sum_{i \in B_0} x_i \geq \beta n/2$, and thus there exists some index $j$ such that $\sum_{i \in B_j} x_i \geq \beta n/(2k)$.

Let $\gamma = 2^{j-1}/\log d$; for each $i \in B_j$, $x_i \in (\gamma, 2\gamma]$. Since $x_i > \gamma$ for each $i \in B_j$, the clique constraints ensure that the graph induced on $B_j$ is $K_r$-free for $r = 1/\gamma$. Moreover, since $x_i \leq 2\gamma$ for each $i \in B_j$, $|B_j| \geq \frac{1}{2\gamma} \cdot \frac{\beta n}{2k}$. By Shearer's result for $K_r$-free graphs we obtain

$$\alpha(B_j) = \Omega\left(|B_j| \cdot \frac{\gamma \log d}{d \log\log d}\right) = \Omega\left(\frac{\beta n \log d}{d(\log\log d)^2}\right).$$

This implies the claim about the integrality gap.

A similar argument implies the constructive result. Let $\beta n$ denote the value of the $SA_{(d)}$ relaxation. As before, we assume that $\beta \geq 2/\log d$ and divide the vertices into $1 + \log\log d$ classes. Consider the class $B_j$ with $j \geq 1$ that contributes most to the objective, and use the fact that the graph induced on $B_j$ is locally $k$-colorable for $k = (\log d/2^{j-1} \cdot \log d) = O(\log^2 d)$. As in section 4, we can now use Johansson's coloring algorithm Theorem A.1 to find a large independent set. ∎

**Appendix A. Johansson's algorithm for coloring sparse graphs.** For completeness, we state two results of Johansson [Joh96b] on coloring degree-$d$ graphs: one about graphs where vertex neighborhoods can be colored using few colors ("locally-colorable" graphs), and another about $K_r$-free graphs. Since the original manuscript is not available online, a complete proof is presented in the arXiv version of this paper [BGG15].

THEOREM A.1. *For any $r, \Delta$, there exists a randomized algorithm that, given a graph $G$ with maximum degree $\Delta$ such that the neighborhood of each vertex is $r$-colorable, outputs a proper coloring of $V(G)$ using $O(\frac{\Delta}{\ln \Delta} \ln r)$ colors in expected* $\mathrm{poly}(n2^{\Delta})$ *time.*

THEOREM A.2. *For any $r, \Delta$, there exists a randomized algorithm that, given a graph $G$ with maximum degree $\Delta$ which excludes $K_r$ as a subgraph, outputs a proper coloring of $V(G)$ using $O\big(\frac{\Delta}{\ln \Delta}(r^2 + r \ln\ln \Delta)\big)$ colors in expected* $\mathrm{poly}(n)$ *time.*

**Appendix B. Miscellaneous proofs.**

**B.1. Proof of Theorem 2.1.** Recall the statement of Halperin's theorem: for $\eta \in [0, \frac{1}{2}]$, suppose $Z$ is the collection of vectors $v_i$ satisfying $\|v_i\|^2 \geq \eta$ in the SDP solution. Then we want to find an independent set of size $\Omega(\frac{d^{2\eta}}{d\sqrt{\ln d}}|Z|)$.

Let $a_i = v_i \cdot v_0 = \|v_i\|^2$, and let $w_i = v_i - \langle v_i, v_0\rangle v_0$ denote the projection of $v_i$ to $v_0^{\perp}$, the hyperplane orthogonal to $v_0$. As $\|w_i\|^2 + \langle v_i, v_0\rangle^2 = \|v_i\|^2$, we obtain $\|w_i\|^2 = a_i - a_i^2$. Let $u_i = w_i/|w_i|$.

Now for any pair of vertices $i, j$, we have that $w_i \cdot w_j = v_i \cdot v_j - \langle v_i, v_0\rangle\langle v_j, v_0\rangle$. As $v_i \cdot v_j = 0$ if $(i, j) \in E$, we have that

$$w_i \cdot w_j = v_i \cdot v_j - \langle v_i, v_0\rangle\langle v_j, v_0\rangle = -a_i a_j$$

and hence

$$u_i \cdot u_j = -\frac{\sqrt{a_i a_j}}{\sqrt{(1-a_i)(1-a_j)}} \leq -\frac{\eta}{(1-\eta)}.$$

The last step follows as $a_i, a_j \geq \eta$ and as $x/1-x$ is increasing for $x \in [0, 1]$.

Thus the unit vectors $u_i$ can be viewed as a feasible solution to a vector $k$-coloring (in the sense of [KMS98]), where $k$ is such that $1/(k-1) = \eta/(1-\eta)$. This gives $k = 1/\eta$, and now we can use the result of [KMS98, Lemma 7.1] that such graphs have independent sets of size $\Omega(|Z|/d^{1-2/k}\sqrt{\ln d}) = \Omega(|Z|d^{2\eta}/(d\sqrt{\ln d}))$.

**B.2. Proof of Theorem 4.4.** The proof of Theorem 4.4 is similar to that of Theorem 3.1, and we give only the differences. We set $\gamma = \frac{\log d}{\log k}$. As in that proof, we wish to show that $\mathbb{E}[X_v] \geq c\gamma$ for each vertex $v$ and some constant $c > 0$. The next few steps of the proof are identical, culminating in (6), which shows that

$$\mathbb{E}[X_v \mid W \cap V(H') = S] \geq d\frac{1}{2^{\epsilon x}+1} + \frac{\epsilon x}{10\log(1/\epsilon+1)}\frac{2^{\epsilon x}}{2^{\epsilon x}+1}.$$

Again, if $2^{\epsilon x} + 1 \leq \sqrt{d}$, then the first term is at least $\sqrt{d}$ and we are done. Otherwise, it must be that $\epsilon x \geq (1/2)\log d$ and hence the right-hand side in (6) is at least

$$\frac{\log d}{20\log(1/\epsilon+1)}. \tag{9}$$

We now consider two cases depending on the value of $x$. Recall the assumptions on the graph $H$: namely, for any vertex $v$ and any subset $T$ lying in the neighborhood of $v$ of size at most $k\log^2 d$, there is an independent set of size at least $|T|/k$.

- If $x \leq k\log^2 d$, then by our assumption, $X$ contains an independent set of size at least $|X|/k$. Every subset of this is also an independent set, and hence the number of independent sets in $X$ is $2^{\epsilon x} \geq 2^{x/k}$. Hence $\epsilon \geq 1/k$ and so gives that (9) is at least $\frac{\log d}{40\log(k+1)}$.
- If $x \geq k\log^2 d$, then again by our assumption, $X$ contains at least $2^{\log^2 d}$ independent sets, and hence $\epsilon x \geq \log^2 d$. As $x \leq d$, it follows that $\epsilon \geq \log^2 d/d \geq 1/d$ and hence $\log(1/\epsilon + 1) \leq 2\log d$. Thus the right-hand side of (6) is at least
$$\frac{\epsilon x}{20\log(1/\epsilon+1)} \geq \frac{\log^2 d}{40\log d}.$$

In all cases we have an independent set of size $\Omega(\frac{n}{d}\frac{\log d}{\log(k+1)})$, which completes the proof of Theorem 4.4.

**Appendix C. The average-degree case.** In this section, we show that any algorithm for graphs with *maximum* degree $d$ based on (lifts of) the standard SDP can be translated into an algorithm for graphs with *average* degree $\delta$, albeit with a slight loss in performance. For example, an integrality gap of $O(d/\log^2 d)$ translates to one of $O(\delta/\log^{1.5}\delta)$. Moreover, we show that it is unlikely that we can do better.

LEMMA C.1. *Let $\epsilon \le 1$ and $\ell \ge 1$. Suppose the integrality gap of the $\ell$-level $SA^+$ semidefinite relaxation on graphs with* maximum *degree $d$ is*

$$\widetilde{O}\left(\frac{d}{(\log d)^{1+\epsilon}}\right);$$

*then its integrality gap on graphs with* average *degree $\delta$ is at most*

$$\widetilde{O}\left(\frac{\delta}{(\log \delta)^{1+\epsilon/2}}\right).$$

*Proof.* Let $\mathsf{sdp}$ denote the value of the $\ell$-level $SA^+$ semidefinite relaxation on the graph $G$, and define $\beta := 1/(\log^{1+\epsilon/2}\delta)$. We can assume that $\mathsf{sdp} \ge 3\beta n$: indeed, the greedy algorithm finds an independent set of size at least $n/(\delta+1) \ge \mathsf{sdp}/(3\beta(\delta+1))$ and thus bounds the integrality gap by $\tilde{O}(\delta/\log^{1+\epsilon/2}\delta)$.

Define $\eta = \frac{c\log\log\delta}{\log\delta}$ and partition the vertices into three sets as follows:

$$\begin{aligned} A &= \{v \mid x_v \ge \eta\},\\ B &= \{v \mid \beta \le x_v < \eta\},\\ C &= \{v \mid x_v < \beta\}. \end{aligned}$$

Let $x(S) = \sum_{v\in S} x_v$. Since $\mathsf{sdp} \ge 3\beta n$, the SDP value of vertices in $C$ is $x(C) < |C|\beta \le \beta n \le \mathsf{sdp}/3$. Hence, at least one of $x(A)$ or $x(B)$ is greater than $\mathsf{sdp}/3$. In each case, we will exhibit an independent set of size $\Omega((\log^{1+\epsilon/2}\delta)/\delta)\cdot\mathsf{sdp}$, which in turn will bound the integrality gap.

*Case* I. Suppose $x(A) \ge \mathsf{sdp}/3$. This implies that $|A| \ge x(A) \ge \mathsf{sdp}/3 \ge \beta n$. We define a vertex $v$ to be "$A$-high" if $deg(v) \ge a := \frac{2}{\beta}\delta$. By Markov's inequality, there are at most $\beta n/2$ $A$-high vertices in $G$. If we drop $A$-high vertices from $A$, there are at least $|A| - \beta n/2 \ge |A|/2$ vertices in the remaining set $A'$. Furthermore, the graph $G[A']$ has maximum degree $a$. Applying Theorem 2.1 to the set of vectors in the solution induced on vertices from $A'$ gives an independent set of size

$$\Omega\left(|A'|\cdot\frac{a^{2\eta}}{a\sqrt{\log a}}\right) \ge \Omega\left(|A|\cdot\frac{(\log\delta)^{2c}}{\delta(\log\delta)^2}\right) \ge \Omega\left(\frac{(\log\delta)^{2c}}{\delta(\log\delta)^2}\right)\mathsf{sdp}.$$

Setting $c \ge 3/2 + \epsilon/4$ completes this case.

*Case* II. Suppose $x(B) \ge \mathsf{sdp}/3$. This implies that $|B|\cdot\eta \ge x(B) \ge \mathsf{sdp}/3 \ge \beta n$. Hence, we can say that $|B| \ge \frac{\beta}{\eta}n$. We define a vertex $v$ to be "$B$-high" if $deg(v) \ge b := 2\frac{\eta}{\beta}\delta$. By Markov's inequality, there are at most $\frac{\beta}{\eta}n/2$ $B$-high vertices in $G$. If we drop $B$-high vertices from $B$, there are at least $|B|/2$ vertices in the remaining set $B'$. The graph $G[B']$ now has maximum degree $b$. Moreover, $x(B') \ge x(B) - (\frac{\beta}{\eta}n/2)\cdot\eta \ge x(B)/2$, so the optimal value of the $SA^+$ relaxation on the graph $G[B']$ is at least as high. Now we can apply the assumption on the integrality gap of the convex program to $G[B']$ to infer the existence of an independent set in $G[B']$ (and hence in $G$) of size

$$\widetilde{\Omega}\left(\frac{b}{\log^{1+\epsilon}b}x(B')\right) = \widetilde{\Omega}\left(\frac{\delta}{\log^{1+\epsilon/2}\delta}x(B)\right) \ge \widetilde{\Omega}\left(\frac{\delta}{\log^{1+\epsilon/2}\delta}\mathsf{sdp}\right). \qquad \square$$

To show this transformation cannot be improved substantially, consider a graph $G$ showing the integrality gap of the SDP in terms of the maximum degree $d$ is

$\Omega(d/(\log d)^{1+\epsilon})$. Specifically, let $G$ be a graph on $n$ vertices and maximum degree $d$ such that $\alpha(G) = O(n\beta/d)$, but the value of the semidefinite program is $\mathsf{sdp}(G) = \Omega(n\beta/(\log d)^{1+\epsilon})$ for some $\beta \in [1, (\log d)^{\epsilon}]$.[3] From this we construct an instance $H$ with an integrality gap of $\tilde{\Omega}(\delta/(\log \delta)^{1+\epsilon/2})$, where $\delta$ is the average degree.

Define $n' = n\,(\log d)^{\epsilon/2}$ and $\delta' = d/(\log d)^{\epsilon/2}$, so that $nd = n'\delta'$. We construct $H$ by taking the union of $G$ with $n'/\delta'$ disjoint copies of $K_{\delta'}$, the complete graph on $\delta'$ vertices. The number of edges in $H$ is at most $nd + n'\delta' = 2nd$ and the number of vertices is $n' + n \in [n', 2n']$; the average degree of $H$ is $\delta := O(\delta')$. Furthermore, $\mathsf{sdp}(H) \geq \mathsf{sdp}(G) = \Omega(n\beta/(\log d)^{1+\epsilon}) = \Omega(n'\beta/(\log \delta)^{1+\epsilon/2})$ and $\alpha(H) = n\beta/d + n'/\delta' \leq 2n'/\delta' = O(n'/\delta)$. Therefore, the integrality gap of the SDP on the instance $H$ is at least $\widetilde{\Omega}(\delta/(\log \delta)^{1+\epsilon/2})$, where $\delta$ is its average degree.

**Acknowledgments.** We thank Alan Frieze for sharing his copy of Johansson's preprint. We thank Noga Alon, Tom Bohman, Alan Frieze, Venkatesan Guruswami, Lap Chi Lau, Konstantin Makarychev, and Prasad Raghavendra for enlightening discussions. Finally, we thank the referees; their thorough reviews helped us fix errors and improve the presentation.

## REFERENCES

[AEKS81] M. Ajtai, P. Erdős, J. Komlós, and E. Szemerédi, *On Turán's theorem for sparse graphs*, Combinatorica, 1 (1981), pp. 313–317.

[AK98] N. Alon and N. Kahale, *Approximating the independence number via the $\vartheta$-function*, Math. Program., 80 (1998), pp. 253–264, http://dx.doi.org/10.1007/BF01581168.

[AKS80] M. Ajtai, J. Komlós, and E. Szemerédi, *A note on Ramsey numbers*, J. Combin. Theory Ser. A, 29 (1980), pp. 354–360.

[AKS99] N. Alon, M. Krivelevich, and B. Sudakov, *Coloring graphs with sparse neighborhoods*, J. Combin. Theory Ser. B, 77 (1999), pp. 73–82.

[AKS11] P. Austrin, S. Khot, and M. Safra, *Inapproximability of vertex cover and independent set in bounded degree graphs*, Theory Comput., 7 (2011), pp. 27–43.

[Alon96] N. Alon, *Independence numbers of locally sparse graphs and a Ramsey type problem*, Random Struct. Algorithms, 9 (1996), pp. 271–278.

[AS92] N. Alon and J. Spencer, *The Probabilistic Method*, Wiley Interscience, New York, 1992.

[BGG15] N. Bansal, A. Gupta, and G. Guruganesh, *On the Lovász Theta Function for Independent Sets in Sparse Graphs*, CoRR, arXiv:1504.04767, 2015.

[BK10] T. Bohman and P. Keevash, *The early evolution of the $H$-free process*, Invent. Math., 181 (2010), pp. 291–336.

[Cha13] S. O. Chan, *Approximation resistance from pairwise independent subgroups*, in Proceedings of STOC, 2013, pp. 447–456, http://doi.acm.org/10.1145/2488608.2488665.

[CPS14] K.-M. Chung, S. Pettie, and H.-H. Su, *Distributed algorithms for the Lovász local lemma and graph coloring*, in Proceedings of PODC, 2014, pp. 134–143, http://doi.acm.org/10.1145/2611462.2611465.

[CT12] E. Chlamtac and M. Tulsiani, *Convex relaxations and integrality gaps*, in Handbook on Semidefinite, Conic and Polynomial Optimization, M. F. Anjos and J. B. Lasserre, eds., Springer, New York, 2012.

[ES35] P. Erdős and G. Szekeres, *A combinatorial problem in geometry*, Compos. Math., 2 (1935), pp. 463–470.

[Fei04] U. Feige, *Approximating maximum clique by removing subgraphs*, SIAM J. Discrete Math., 18 (2004), pp. 219–225.

[FM13] A. Frieze and D. Mubayi, *Coloring simple hypergraphs*, J. Combin. Theory Ser. B, 103 (2013), pp. 767–794.

[GLS88] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, Springer-Verlag, Berlin, 1988.

[3] We may assume $\beta \geq 1$ as $\alpha(G) \geq n/d$ for any $d$-regular graph. If $\mathsf{sdp}(G) \geq \Omega(n/\log d)$, then Theorem 2.1 gives us a better integrality gap. Therefore, we can assume $\beta \in [1, (\log d)^{\epsilon}]$.

[GM12] B. Gärtner and J. Matoušek, *Approximation Algorithms and Semidefinite Programming*, Springer, Heidelberg, 2012, http://dx.doi.org/10.1007/978-3-642-22015-9.

[GP00] D. A. Grable and A. Panconesi, *Fast distributed algorithms for Brooks-Vizing colorings*, J. Algorithms, 37 (2000), pp. 85–120, http://dx.doi.org/10.1006/jagm.2000.1097.

[Hal00] M. M. Halldórsson, *Approximations of weighted independent set and hereditary subset problems*, J. Graph Algorithms Appl., 4 (2000), pp. 1–16, http://dx.doi.org/10.7155/jgaa.00020.

[Hal02] E. Halperin, *Improved approximation algorithms for the vertex cover problem in graphs and hypergraphs*, SIAM J. Comput., 31 (2002), pp. 1608–1623.

[Has96] J. Håstad, *Clique is hard to approximate within* $n^{1-\epsilon}$, in Proceedings of FOCS, 1996, pp. 627–636.

[HR94] M. M. Halldórsson and J. Radhakrishnan, *Improved approximations of independent sets in bounded-degree graphs via subgraph removal*, Nord. J. Comput., 1 (1994), pp. 475–492.

[Joh96a] A. Johansson, *Asymptotic Choice Number for Triangle-Free Graphs*, 1996.

[Joh96b] A. Johansson, *The Choice Number of Sparse Graphs*. preprint, 1996.

[Kim95] J. H. Kim, *On Brooks' theorem for sparse graphs*, Combin. Probab. Comput., 4 (1995), pp. 97–132.

[KMS98] D. R. Karger, R. Motwani, and M. Sudan, *Approximate graph coloring by semidefinite programming*, J. ACM, 45 (1998), pp. 246–265.

[KP06] S. Khot and A. K. Ponnuswami, *Better inapproximability results for MaxClique, chromatic number and Min-3Lin-Deletion*, in Proceedings of ICALP, vol. 1, 2006, pp. 226–237.

[Lau] M. Laurent, *Networks and Semidefinite Programming*, http://homepages.cwi.nl/~monique/lnmb12/Lecture_Notes_NSP_21_01_13.pdf (21 January 2013).

[Lau03] M. Laurent, *A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for* 0-1 *programming*, Math. Oper. Res., 28 (2003), pp. 470–496.

[Lov79] L. Lovász, *On the Shannon capacity of a graph*, IEEE Trans. Inform. Theory, 25 (1979), pp. 1–7, http://dx.doi.org/10.1109/TIT.1979.1055985.

[MR02] M. Molloy and B. Reed, *Graph Colouring and the Probabilistic Method*, Algorithms Combin. 23, Springer-Verlag, Berlin, 2002, http://dx.doi.org/10.1007/978-3-642-04016-0.

[She83] J. B. Shearer, *A note on the independence number of triangle-free graphs*, Discrete Math., 46 (1983), pp. 83–87.

[She95] J. B. Shearer, *On the independence number of sparse graphs*, Random Struct. Algorithms, 7 (1995), pp. 269–272.

[Vu02] V. H. Vu, *A general upper bound on the list chromatic number of locally sparse graphs*, Combin. Probab. Comput., 11 (2002), pp. 103–111.