Randomized graph products, chromatic numbers, and the Lovász ϑ-function

Uriel Feige*

September 24, 1996

Abstract

For a graph G, let $\alpha(G)$ denote the size of the largest independent set in G, and let $\vartheta(G)$ denote the Lovász ϑ-function on G. We prove that for some $c > 0$, there exists an infinite family of graphs such that $\vartheta(G) > \alpha(G)n/2^{c\sqrt{\log n}}$, where n denotes the number of vertices in a graph. This disproves a known conjecture regarding the ϑ function.

As part of our proof, we analyse the behavior of the chromatic number in graphs under a randomized version of graph products. This analysis extends earlier work of Linial and Vazirani, and of Berman and Schnitger, and may be of independent interest.

1 Introduction

Lovász [21] introduced the ϑ function in order to study the so called "Shannon Capacity" of graphs. For every graph G, the ϑ function enjoys the following *sandwich* property:

$$\alpha(G) \leq \vartheta(G) \leq \bar{\chi}(G)$$

where $\alpha(G)$ is the size of the largest independent set in G, and $\bar{\chi}(G)$ is the clique cover number of G ($\bar{\chi}(G) = \chi(\bar{G})$, the chromatic number of the complement of G). This sandwich property shows that $\vartheta(G) = \alpha(G)$ for *perfect graphs*, and suggests that perhaps the ϑ function provides a good estimation of $\alpha(G)$ for general graphs. This would be desirable, since the ϑ function can be approximated with arbitrary precision in polynomial time [14],

*Department of Applied Math and Computer Science, The Weizmann Institute, Rehovot 76100, Israel. feige@wisdom.weizmann.ac.il. Incumbent of the Joseph and Celia Reskin Career Development Chair. Yigal Alon Fellow.

whereas computing $\alpha(G)$ is NP-hard. However, Lovász showed that for random graphs on n vertices, the gap between $\alpha(G)$ and $\vartheta(G)$ is large – in the order of $\sqrt{n}/\log n$.

The following conjecture is attributed to Lovász (see [18]):

Conjecture 1.1 *There exists some constant c, such that for any graph G, $\vartheta(G) < c\alpha(G)\sqrt{n}$.*

Observe that if the conjecture were true, there would be a polynomial time algorithm that estimates $\alpha(G)$ within a factor of $O(\sqrt{n})$. Though this ratio of approximation seems very poor, it is still considerably better than the best ratio of approximation known to be achievable in polynomial time, which currently is $O(n/\log^2 n)$ [8]. Moreover, in a sequence of papers [9, 3, 2] it was established that there is some $\epsilon > 0$ such that it is NP-hard to approximate $\alpha(G)$ within a ratio of n^ϵ. Subsequent works established hardness results for approximating $\alpha(G)$ with ratios close to $n^{1/3}$ [4]. The question of whether there is some ϵ such that a polynomial time algorithm approximates $\alpha(G)$ within $n^{1-\epsilon}$ is open [1]. The conjecture suggested an affirmative answer to this open question.

Until recently, there were virtually no techniques for establishing upper bounds on the ratio $\vartheta(G)/\alpha(G)$ for general graphs, or on the ratio $\bar{\chi}(G)/\vartheta(G)$. The work of Goemans and Williamson [13] on MAX CUT and MAX 2SAT changed this situation, by introducing the vector rounding technique for semidefinite programming (computing the ϑ function is a semidefinite programming task). Karger, Motwani and Sudan [17] extended this technique, and considered the relation between $\chi(G)$ and a solution to a semidefinite program that is equivalent to $\vartheta(\bar{G})$. They showed how to color 3-colorable graphs in polynomial time with $O(n^{1/4})$ colors, which is an improvement over the previous best results [25, 7]. More generally, their results imply that $\bar{\chi}(G) \leq n^{1-c/\vartheta(G)}$, for some constant $c > 0$. They also obtained negative results concerning the ϑ function, constructing examples that show that for some $\epsilon > 0$, there are graphs with $\bar{\chi}(G) = n^\epsilon$ and $\vartheta(G) \leq 3$. Alon and Kahale [1] extended the work of [17], and showed how to use the ϑ function in order to obtain a quantitative improvement over [8]'s algorithm for finding large independent sets in graphs that have linear size independent sets.

A somewhat different direction was taken by Szegedy [24] who showed that the following two conjectures are equivalent:

1. For some $\epsilon > 0$ and every graph G, $\vartheta(G) \leq \alpha(G)n^{1-\epsilon}$.
2. For some $\epsilon > 0$ and every graph G, $\bar{\chi}(G) \leq \vartheta(G)n^{1-\epsilon}$.

[1] See recent developments in Section 4.

In this paper we disprove the above two conjectures, and as a corollary we also disprove Conjecture 1.1.

Theorem 1.2 *For some constant $c > 0$ and an infinite family of graphs, $\bar{\chi}(G) > n^{1-c/\log \vartheta(G)}$. In particular, there exists an infinite family of graphs for which $\vartheta(G) < 2^{\sqrt{\log n}}$ and $\bar{\chi}(G) > n/2^{c\sqrt{\log n}}$.*

Likewise, for some constant $c > 0$ and an infinite family of graphs, $\vartheta(G) > n^{1-c/\log \alpha(G)}$. In particular, there exists an infinite family of graphs for which $\alpha(G) < 2^{\sqrt{\log n}}$ and $\vartheta(G) > n/2^{c\sqrt{\log n}}$.

We prove our theorem by giving a non-explicit (randomized) construction of the bad family of graphs. Our construction was inspired by the "randomized graph products" technique of Berman and Schnitger [5], and in particular, by Blum's use of randomized graph products in order to show that if $\alpha(G)$ can be approximated in polynomial time within a factor of $n^{1-\epsilon}$, then this leads to a randomized polynomial time algorithm that colors three colorable graphs with $O(\log n)$ colors [6].

In Section 2 we present the technique of randomized graph products, and show how it affects the chromatic number of graphs. In Section 3 we define the ϑ function, present some of its known properties, and prove Theorem 1.2. In Section 4 we discuss other potential applications of randomized graph products. These include gap amplification for NP-hardness results for approximating chromatic numbers, and explicit construction of Ramsey graphs.

2 Randomized graph products and the chromatic number

In this section, we develop a general theory for the relation between randomized graph products and the chromatic number. Graph products come in many variants, and we concentrate here on the variant that seems most appropriate in the context of the chromatic number.

2.1 Background and definitions

Recall that for any graph G, its chromatic number $\chi(G)$ is the minimum number of colors that suffice in order to color the vertices of G, where adjacent vertices must receive different colors. Linial and Vazirani [19] analysed the behavior of the chromatic number under graph products.

Definition 2.1 *For a graph G, $V(G)$ denotes its vertex set and $E(G)$ denotes its edge set. The* inclusive graph product *$G \times H$ is the graph with vertex set $V(G \times H) = V(G) \times V(H)$, and edge set $([x, y], [x', y']) \in E(G \times H)$ iff $(x, x') \in E(G)$ or $(y, y') \in E(H)$. The k-fold inclusive graph product of a graph G with itself is denoted by G^k.*

Proposition 2.2 *For any graphs G and H, $\alpha(G \times H) = \alpha(G) \cdot \alpha(H)$.*

Proof: Consider any independent sets in G and H. Then their direct product is an independent set in $G \times H$. Conversely, consider any independent set C in $G \times H$. Consider the set of vertices $V_{G|C} \subset V(G)$ such that $v \in V_{G|C}$ iff $\exists u \in H$ such that $(v, u) \in C$. Then by the fact that C is an independent set in $G \times H$ and the definition of graph products, it follows that $V_{G|C}$ is an independent set in G. A similar argument applies to $V_{H|C}$.

It follows that every maximal independent set in $G \times H$ is the direct product of maximal independent sets in G and H, and that any maximum independent set in $G \times H$ is the direct product of maximum independent sets in G and H. $\square$

Linial and Vazirani [19] show:

Theorem 2.3 *For any two graphs on n vertices*

$$\frac{\chi(G) \cdot \chi(H)}{\ln n} \leq \chi(G \times H) \leq \chi(G) \cdot \chi(H)$$

The upper bound on $\chi(G \times H)$ in the above theorem is trivial (the direct product of the minimum coloring of G and the minimum coloring of H gives a coloring of $G \times H$). Linial and Vazirani presented a combinatorial proof for the lower bound. However, it is more convenient to study graph products within the framework of the *fractional chromatic number*, as is done implicitly in [20]. We survey this approach.

Observe that the chromatic number of a graph G is the minimum number of independent sets that cover each vertex of G. Hence it is the solution of an integer program with exponentially many variables. Each maximal independent set S_i of G is associated with a variable y_i.

Minimize $\sum y_i$ subject to:

$y_i \in \{0, 1\}$, for each y_i.

$\sum_{\{i|v \in S_i\}} y_i \geq 1$, for each vertex $v \in V(G)$.

We can relax the above to a linear program as follows:

Minimize $\sum y_i$ subject to:

$y_i \geq 0$, for each y_i.

$\sum_{\{i|v \in S_i\}} y_i \geq 1$, for each vertex $v \in V(G)$.

Hence each maximal independent set is assigned a nonnegative weight such that each vertex is "covered" by weights that sum up to at least 1.

Definition 2.4 *The* fractional chromatic number *of graph G, denoted by $\chi_f(G)$, is the minimum value of the above (relaxed) linear program.*

Lovász [20] has shown:

Lemma 2.5 *For any graph G on n vertices, $\chi(G)/(1+\ln(\alpha(G)) < \chi_f(G) \leq \chi(G)$.*

To understand the behavior of the fractional chromatic number under graph products, it is convenient to consider the dual to the linear program that defines the fractional chromatic number. We associate a variable x_i with each vertex v_i.

Maximize $\sum_i x_i$, subject to:

$x_i \geq 0$, for each x_i.

$\sum_{\{i|v_i \in S_j\}} x_i \leq 1$, for each maximal independent set S_j.

Definition 2.6 *The* fractional matching number *of graph G, denoted by $\tau_f(G)$, is the maximum value of the above linear program.*

Proposition 2.7 *For any graph G, $\chi_f(G) = \tau_f(G)$.*

Proof: By duality theory for linear programming. $\square$

Lemma 2.8 *For any graphs G and H, $\chi_f(G \times H) = \chi_f(G) \cdot \chi_f(H)$.*

Proof: By taking direct products of the optimum solutions to the defining linear programs, we have that $\chi_f(G \times H) \leq \chi_f(G) \cdot \chi_f(H)$, and $\tau_f(G \times H) \geq \tau_f(G) \cdot \tau_f(H)$. This follows from the fact that every maximal independent set in $G \times H$ is the direct product of maximal independent sets in G and H (see proof of Proposition 2.2). $\square$

Lemma 2.8 together with Lemma 2.5 imply a lower bound similar (though not identical) to the one in Theorem 2.3. We have the following corollary.

Corollary 2.9 *For any graph G, $(\chi(G))^k/(\ln(\alpha(G))+1)^k \leq \chi(G^k) \leq (\chi(G))^k$.*

2.2 The sampling problem

A randomized graph product is obtained by taking a random subset of the vertices of a product graph, and considering the vertex induced subgraph that results. We are interested in the following problem.

Problem 2.10 *Given a graph G on n vertices and an integer k, construct a graph $\hat{G}_k$ with the following properties:*

1. *$\hat{G}_k$ is a vertex induced subgraph of G^k.*
2. *$\chi_f(G^k)/2 \leq |V(\hat{G}_k)| \leq 2\chi_f(G^k)$.*
3. *$\alpha(\hat{G}_k) \leq \frac{kn}{\ln(kn)}$. Hence $\chi_f(\hat{G}_k) \geq |V(\hat{G}_k)| \frac{\ln(kn)}{kn}$.*

Theorem 2.11 *For large enough n, for every graph G on n vertices and for any positive integer k, graphs $\hat{G}_k$ as described in Problem 2.10 exist.*

Proof: By a probabilistic argument. From the definition of $\tau_f(G)$, the dual of $\chi_f(G)$, it follows that there are non-negative values x_i that can be assigned to the vertices v_i of G^k, such that $\sum x_i = \chi_f(G^k)$, and on any maximal independent set S_j of G^k, $\sum_{\{i|v_i \in S_j\}} x_i \leq 1$. Construct $\hat{G}_k$ by performing the following random sampling procedure. For each vertex $v_i \in G^k$, select it with probability x_i, independently of other vertices. Let $\hat{G}$ denote the vertex induced subgraph that is obtained by such a process. ($\hat{G}$ is a random variable.) We show that with high probability, $\hat{G}$ has all the properties specified in Problem 2.10.

The first property holds, as by construction, $\hat{G}$ is a vertex induced subgraph of G^k. To see that the second property holds, observe that the expected number of vertices in $\hat{G}$ is exactly $\chi_f(G^k)$. The probability that the number of vertices in $\hat{G}$ deviates from its expectation by a factor of more than two is small, and can be ignored. (Observe: this last probability is highest when $\chi_f(G^k)$ is small. But even if G is an independent set and $\chi_f(G^k) = 1$, still there is constant probability that $1/2 \leq \hat{G} \leq 2$, regardless of the value of n.) To see that the third property holds, consider any maximal independent set $S_j \in G^k$. We bound the probability that more than t of its vertices end up in $\hat{G}$. By the constraint $\sum_{\{i|v_i \in S_j\}} x_i \leq 1$, we obtain that the expected number of vertices of S_j that end up in $\hat{G}$ is at most 1. The proof of the following lemma follows well known techniques (see for example [16]). We present it for completeness.

Lemma 2.12 *Let X_1, ..., X_n be independent random Boolean variables, with $\sum_{i=1}^{n} E[X_i] \leq 1$, where $E[.]$ denotes expectation. Then $Pr[\sum X_i \geq t] \leq e^{-t\ln t + t - 1}$.*

Proof: For any $h > 0$, using Markov's inequality, and the independence of the variables:

$$Pr[\sum X_i \geq t] = Pr[e^{h(-t+\sum X_i)} \geq 1]$$

$$\leq E[e^{h(-t+\sum X_i)}] = e^{-ht} \prod E[e^{hX_i}]$$

If we denote $p_i = Pr[X_i = 1]$ we obtain that $E[e^{hX_i}] = 1 + p_i(e^h - 1)$, and $\ln E[e^{hX_i}] \leq p_i(e^h - 1)$. Since $\sum p_i \leq 1$, we have that $\sum \ln E[e^{hX_i}] \leq e^h - 1$ and $\prod E[e^{hX_i}] \leq e^{e^h - 1}$. Select $h = \ln t$. $\square$

It follows that the probability that more than $kn/\ln kn$ vertices of S_j get selected is at most $e^{-kn(1-o(1))}$, where $o(1) \to 0$ as $kn \to \infty$.

Lemma 2.13 *The number of maximal independent sets in G^k is less than $3^{kn/3}$.*

Proof: The number of maximal independent sets in G is at most $3^{n/3}$ [23]. Any maximal independent set in G^k is the direct product of k maximal independent sets in G. $\square$

We conclude that the probability that the third property of Problem 2.10 does not hold is at most $3^{kn/3}e^{-kn(1-o(1))}$. But since $3^{1/3}/e < 1$, this probability tends to 0 as n grows. $\square$

We proved Theorem 2.11 by using a probabilistic argument. We sample vertices of G^k at random, with a certain well defined probability distribution. However, this probability distribution is hard to compute, unless $P = NP$. One may ask whether using a uniform distribution over the vertices of G^k would suffice. Unfortunately, in general, the answer is negative. This can easily be seen by considering a graph G which has a single clique of size $\chi(G)$, and in which all other vertices are isolated.

However, in an important special case, we can select vertices of G^k uniformly at random. Recall that $\alpha(G)$ denotes the size of the maximum independent set, and that $\alpha(G)\cdot\chi(G) \geq n$.

Problem 2.14 *Given a graph G and an integer k, construct a graph $\hat{G}_k$ with the following properties:*

1. *$\hat{G}_k$ is a vertex induced subgraph of G^k.*
2. *$(n/\alpha(G))^k/2 \leq |V(\hat{G}_k)| \leq 2(n/\alpha(G))^k$.*
3. *$\alpha(\hat{G}_k) \leq k\alpha(G)\ln n/\ln \alpha(G)$. Hence $\chi(\hat{G}_k) \geq |V(\hat{G}_k)| \ln \alpha(G)/k\alpha(G)\ln n$.*

Corollary 2.15 *For every graph G and positive integer k, selecting independently and uniformly at random vertices of G^k, each with probability $\alpha(G)^{-k}$, gives induced subgraphs of G^k that satisfy the requirements in Problem 2.14 with high probability.*

Proof: Similar to the proof of Theorem 2.11, using $x_i = \alpha(G)^{-k}$, and observing that the number of maximal independent sets in G is at most $n^{\alpha(G)}$. $\square$

3 The ϑ function

3.1 Definitions and preliminaries

There are many equivalent definitions of the ϑ function. We shall follow a definition that was suggested in [17]. Our notion of *vector clique cover* is derived from their notion of *vector coloring*, and we deal with the complements of the graphs that they consider.

Definition 3.1 *Given a graph $G = (V, E)$ on n vertices, a* vector k-clique cover *of G is an assignment of unit vectors u_i from the space $\mathcal{R}^n$ to each vertex $i \in V$, such that for any two nonadjacent vertices i and j, the inner product of their vectors satisfies the equality*

$$u_i u_j = -\frac{1}{k-1}$$

Definition 3.2 *For any graph G, $\vartheta(G)$ is the least k for which G admits a vector k-clique cover.*

Remark: For the equivalence between the above definition and other definitions used in [21], see [17].

Observe that if we place k unit vectors such as to maximize the angle between any two of them, their endpoints form the vertices of a simplex in $\mathcal{R}^{k-1}$, and the inner product between any two such vectors is $-1/(k-1)$. It follows that $\vartheta(G) \leq \bar{\chi}(G)$, since the vertices of each clique of the minimal clique cover can be identified with a particular unit vector to give a vector $\bar{\chi}(G)$-clique cover. Moreover, $\alpha(G) \leq \vartheta(G)$, since the maximum angle between two vectors associated with vertices of the maximum independent set can be at most $\cos^{-1}(-1/(\alpha(G)-1))$.

We introduced the notion of inclusive graph products, denoted by $G \times H$, in Definition 2.1. For the purpose of the current section, it is more natural to use a "complement" notion of graph products. We define the strong graph product $G \cdot H$ as a graph with vertex set $V(G \cdot H) = V(G) \times V(H)$, and edge set $([x, y], [x', y']) \in E(G \cdot H)$ iff $\{(x, x') \in E(G)$ or $x = x'\}$ *and* $\{(y, y') \in E(H)$ or $y = y'\}$. This definition is extended in the obvious way to a k-fold graph product $G^{\cdot k}$ of a graph with itself. (The notation $G^{\cdot k}$ is used in order to distinguish between the current version of graph products and the one of Definition 2.1.)

The following property is crucial for using the ϑ function in the context of the Shannon Capacity. (See proof in [21]).

Lemma 3.3 *For every two graphs G and H,*

$$\vartheta(G \cdot H) = \vartheta(G) \cdot \vartheta(H)$$

Let $\omega(G)$ denote the size of the maximum clique in G. Similar to Proposition 2.2 we have:

Proposition 3.4 *For every two graphs G and H,*

$$\omega(G \cdot H) = \omega(G) \cdot \omega(H)$$

Observe that in general, $\alpha(G \cdot H)$ may be strictly larger than $\alpha(G) \cdot \alpha(H)$. In fact, the ϑ function was introduced in order to upper bound $(\alpha(G^{\cdot k}))^{1/k}$. Denote the complement of a graph G by $\bar{G}$. We have that $\alpha(G \cdot \bar{G}) \geq n$, since the vertices (i, i) for $1 \leq i \leq n$ form an independent set in $G \cdot \bar{G}$.

Corollary 3.5 *For any graph G, $\vartheta(G) \cdot \vartheta(\bar{G}) \geq n$.*

If G is a random graph, than the above corollary gives Lovász's proof that for random graphs, $\vartheta(G)/\alpha(G) = \Omega(\sqrt{n}/\log n)$.

3.2 A large gap

The following was proved in [17].

Lemma 3.6 *For some $\epsilon > 0$ there is an infinite family of graphs with $\vartheta(G) \leq 3$ and $\bar{\chi}(G) \geq n^{\epsilon}$.*

To prove Theorem 1.2, we use a complementary version of Theorem 2.11. Let $\bar{\chi}_f(G)$ denote the fractional clique cover number of graph G (which is the fractional chromatic number of $\bar{G}$). Then by complementing the graphs in Problem 2.10 and in Theorem 2.11, we have:

Proposition 3.7 *For large enough n, given any graph G on n vertices and an integer k, there is a graph $\hat{G}_k$ with the following properties:*

1. *$\hat{G}_k$ is a vertex induced subgraph of $G^{\cdot k}$.*
2. *$\bar{\chi}_f(G^{\cdot k})/2 \leq |V(\hat{G}_k)| \leq 2\bar{\chi}_f(G^{\cdot k})$.*
3. *$\omega(\hat{G}_k) \leq \frac{kn}{\ln(kn)}$. Hence $\bar{\chi}(\hat{G}_k) \geq |V(\hat{G}_k)|\frac{\ln(kn)}{kn}$.*

We are now ready to prove Theorem 1.2.

Proof: Let G be a graph as constructed in Lemma 3.6. Lemma 2.5 implies that $\bar{\chi}_f(\bar{G}) \geq n^{\epsilon}$ (the extra logarithmic term is absorbed in the n^{ϵ}, since we did not specify ϵ). Now apply Proposition 3.7. We obtain a graph $\hat{G}$ on roughly $N = n^{\epsilon k}$ vertices, and $\bar{\chi}(\hat{G}) \geq N/kn \geq N^{1-c/k}$, for some $c > 0$ (that depends on ϵ). By Lemma 3.3, $\vartheta(G^{\cdot k}) \leq 3^k$. Any vector clique cover for $G^{\cdot k}$ is also a vector clique cover for its vertex induced subgraph $\hat{G}$. Hence $\vartheta(\hat{G}) \leq \vartheta(G^{\cdot k})$. Hence for any k we have a graph $\hat{G}$ with $\vartheta(\hat{G}) \leq 3^k$ and $\bar{\chi}(\hat{G}) \geq N^{1-c/k}$, which gives the gap between ϑ and $\bar{\chi}$ claimed in Theorem 1.2. For $k = \Theta(\log n) = \Theta(\sqrt{\log N})$ the gap is largest.

By the results in [24], the above also implies a gap of $n^{1-o(1)}$ between $\vartheta(G)$ and $\alpha(G)$. A more direct proof of this large gap follows from taking the complement of the graph $\hat{G}$. The graph G' obtained in this process has $N \simeq n^{\epsilon k}$ vertices, and $\alpha(G') = \omega(\hat{G}) \leq kn/\ln(kn) \leq N^{c/k}$, for some c that depends on ϵ. By Corollary 3.5, it has $\vartheta(G') \geq N/3^k$. Hence $\theta(G') \geq N^{1-c/\log\alpha(G')}$, for some $c > 0$. $\square$

4 Discussion

Corollary 2.15 is only a special case of Theorem 2.11. However, this seems to be the most useful special case. Often, the lower bounds that we know on $\chi(G)$ are in fact derived from known upper bounds on $\alpha(G)$, using the relation $\alpha(G)\chi(G) \geq n$. This was the case for the graphs of Lemma 3.6. This is also the case for the graphs that are constructed by Lund and Yannakakis [22], when proving that it is NP-hard to approximate $\chi(G)$ within a factor of n^{ϵ}, for some ϵ. Based on the above observations, we suggest other potential applications of Corollary 2.15.

4.1 Hardness of approximating the chromatic number

Various kinds of graph products are used in order to strengthen hardness of approximation results for certain graph properties. In fact, this was the intended application of Linial and Vazirani [19]. The notion of randomized graph products (for a type of graph products that is different from the one studied in the current paper) was introduced by Berman and Schnitger [5], and applied in proving computational hardness of approximation for $\alpha(G)$ and $\omega(G)$. Our Corollary 2.15 may have applications in sharpening or simplifying known hardness results for approximating $\chi(G)$. The following theorem shows a possible application of this corollary.

Theorem 4.1 *Assume that it is NP-hard (under randomized reductions) to distinguish between graphs with $\alpha(G) \leq n^{1-c}$ (and hence $\chi(G) \geq n^c$), and graphs with $\chi(G) \leq n^{c(1-\delta)}$, for some $0 < \delta < 1$ and $0 < c < 1$. Then for any $\epsilon > 0$, it is NP-hard (under randomized reductions) to approximate $\chi(G)$ within a factor of $n^{\delta-\epsilon}$.*

Proof: Given an input graph G, it is NP hard to distinguish between $\alpha(G) \leq n^{1-c}$ and $\chi(G) \leq n^{c(1-\delta)}$. Apply Corollary 2.15 on G, with a postulated $\alpha(G) = n^{1-c}$. We obtain a graph $\hat{G}_k$ on roughly n^{ck} vertices, which either has chromatic number above n^{ck-1}, or has chromatic number below $n^{c(1-\delta)k}$. The proof follows, for large enough k. $\square$

In essence, Theorem 4.1 amplifies the "hard gap" from $n^{c\delta}$, where $c < 1$, to $n^{\delta-\epsilon}$, where $\epsilon > 0$ can be made arbitrarily small. The "gap location" moves from n^c to $n^{1-\epsilon}$.

The question of distinguishing between graphs of the above types (one with no large independent set, the other with a small chromatic number) is known to be NP-hard under randomized reductions for $\delta \simeq 1/5$ [4, 12].

4.2 Explicit constructions of Ramsey graphs

Consider the graphs constructed in [17], with $\vartheta(G) \leq 3$ and $\omega(G) \leq n^\delta$, for some $\delta < 1$. By selecting $k \simeq \log n$ in Corollary 2.15, we obtain a Ramsey graph with $n^{O(\log n)}$ vertices and no clique or independent set of n vertices. The best explicit construction known for such graphs has size only $n^{O(\log n/\log\log n)}$. (See Frankl and Wilson [11]. In fact, the constructions in [11] give the proof to Lemma 3.6.)

Open question: Is there an efficient deterministic construction of the graphs $\hat{G}_k$ of Problem 2.14? More specifically, for $0 < \delta < 1$, we need to construct a set $S \subset \{1, \ldots, n\}^{\log n}$ such that for some $0 < \epsilon < 1$ and $c > 0$:

1. $|S| \geq n^{\epsilon \log n}$.

2. For any generalized cube C of size $(n^{\delta})^{\log n}$ we have that $|S \bigcap C| \leq n^c$.

3. S is efficiently constructable. That is, there is an $O(n^{c \log n})$ algorithm (or better still, a $O(\log^2 n)$ space algorithm, or $O(\log^c n)$ time algorithm) that decides for any $x \in \{1, \ldots, n\}^{\log n}$ whether $x \in S$.

A positive answer to this question would improve upon known explicit constructions of Ramsey graphs.

Recent developments: We summarize here recent developments that shed new light on the results of the current paper. Hastad showed [15] that for any $\epsilon > 0$, there is no polynomial time algorithm that can approximate $\alpha(G)$ within a ratio of $n^{1-\epsilon}$, unless NP has randomized polynomial time algorithms. This suggests that there may not be any polynomial time computable function (the ϑ function being only one such example) that approximates the size of the independent set within a ratio of $n^{1-\epsilon}$. Feige and Kilian [10] showed a similar $n^{1-\epsilon}$ hardness result for approximating $\chi(G)$. The proof makes essential use of randomized graph products as in Corollary 2.15.

Acknowledgements

I thank Oded Schramm and Noga Alon for many valuable discussions.

References

[1] N. Alon, N. Kahale. "Approximating the independence number via the θ-function". *Manuscript, November 1994.*

[2] S. Arora, C. Lund, R. Motwani, M. Sudan, M. Szegedy. "Proof verification and hardness of approximation problems". *Proc. of 33rd IEEE Symp. on Foundations of Computer Science, 14–23, 1992.*

[3] S. Arora, S. Safra. "Probabilistic checking of proofs: a new characterization of NP". *Proc. of 33rd IEEE Symp. on Foundations of Computer Science, 2–13, 1992.*

[4] M. Bellare, O. Goldreich, M. Sudan. "Free bits, PCPs and nonapproximability – towards tight results". *Proc. of 36th IEEE Symp. on Foundations of Computer Science, 422–431, 1995.*

[5] P. Berman, G. Schnitger. "On the complexity of approximating the independent set problem". *Information and Computation 96 (1992), 77–94.*

[6] A. Blum. "Algorithms for approximate graph coloring". Phd dissertation, MIT, 1991.

[7] A. Blum. "New approximation algorithms for graph coloring." *Journal of the ACM, 41:470–516, 1994.*

[8] R. Boppana, M. Haldorsson. "Approximating maximum independent sets by excluding subgraphs". *Proc. of 2nd SWAT, Springer, LNCS 447 (1990), 13–25.*

[9] U. Feige, S. Goldwasser, L. Lovász, S. Safra, S. Szegedy. "Interactive proofs and the hardness of approximating cliques". *Journal of the ACM, 43(2), 268–292, 1996.*

[10] U. Feige, J. Kilian. "Zero knowledge and the chromatic number". *Proc. of Eleventh Annual IEEE Conference on Computational Complexity, 278–287, 1996.*

[11] P. Frankl, R. Wilson. "Intersection theorems with geometric consequences". *Combinatorica 1, 357–368, 1981.*

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

[13] M. Goemans, D. Williamson. ".878-Approximation algorithms for MAX CUT and MAX 2SAT". *Proc. of 26th ACM Symposium on the Theory of Computing, 422–431, 1994.*

[14] M. Grotschel, L. Lovász, A. Schrijver. "Geometric algorithms and combinatorial optimization". *Springer-Verlag, Berlin, 1987.*

[15] J. Hastad. "Clique is hard to approximate within $n^{1-\epsilon}$". *Proc. of 37th IEEE Symposium of Foundations of Computer Science, 1996*, to appear.

[16] W. Hoeffding. "Probability inequalities for sums of bounded random variables". *Journal of the American Statistical Association, 58 (1963) 13–30.*

[17] D. Karger, R. Motwani, M. Sudan. "Approximate Graph Coloring by Semidefinite Programming". *Proc. of 35th IEEE Symp. on Foundations of Computer Science, 2–13, 1994.*

[18] D. Knuth. "The sandwich theorem". *The electronic journal of combinatorics, 1:1–48, 1994.*

[19] N. Linial, U. Vazirani. "Graph products and chromatic numbers". *Proc. of 30th IEEE Symp. on Foundations of Computer Science, 124–128, 1989.*

[20] L. Lovász. “On the ratio of the optimal integral and fractional covers”. *Discrete Mathematics 13 (1975) 383–390.*

[21] L. Lovász. “On the Shannon Capacity of a Graph”. *IEEE Trans. on Information Theory, Vol. IT-25, No. 1, pp. 1–7, 1979.*

[22] C. Lund, M. Yannakakis. “On the hardness of approximating minimization problems”. *Journal of the ACM, 41(5), 960–981, 1994.*

[23] J. Moon, L. Moser. “On cliques in graphs”. *Israel J. Math. 3 (1965) 23–28.*

[24] M. Szegedy. “A note on the θ number of Lovász and the generalized Delsarte bound”. *Proc. of 35th IEEE Symp. on Foundations of Computer Science, 36–39, 1994.*

[25] A. Wigderson. “Improving the performance guarantee for approximate graph coloring”. *Journal of the ACM, 30(4), 729–735, 1983.*