Combinatorics, Probability and Computing (2019), **28**, pp. 720–732
doi:10.1017/S0963548319000026

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Separation Choosability and Dense Bipartite Induced Subgraphs†

Louis Esperet[1‡], Ross J. Kang[2,*,§], and Stéphan Thomassé[3]

[1]Université Grenoble Alpes, CNRS, G-SCOP, 46 Avenue Félix Viallet, 38000 Grenoble, France, [2]Radboud University Nijmegen, PO box 9010, 6500 GL Nijmegen, Netherlands and [3]Laboratoire d'Informatique du Parallélisme, École Normale Supérieure de Lyon, 46 Allée d'Italie, 69364 Lyon, France

*Corresponding author. Email: r.kang@math.ru.nl

(Received 17 February 2018; revised 1 December 2018; first published online 26 February 2019)

Abstract

We study a restricted form of list colouring, for which every pair of lists that correspond to adjacent vertices may not share more than one colour. The optimal list size such that a proper list colouring is always possible given this restriction, we call separation choosability. We show for bipartite graphs that separation choosability increases with (the logarithm of) the minimum degree. This strengthens results of Molloy and Thron and, partially, of Alon. One attempt to drop the bipartiteness assumption precipitates a natural class of Ramsey-type questions, of independent interest. For example, does every triangle-free graph of minimum degree d contain a bipartite induced subgraph of minimum degree $\Omega(\log d)$ as $d \to \infty$?

2010 MSC Codes: Primary 05C15; Secondary 05C35

1. Introduction

The concept of *list colouring*, where an adversary prescribes which colours may be used per vertex, was introduced independently by Erdős, Rubin and Taylor [14] and Vizing [33]. The starting point for our paper is a restriction to list colouring proposed by Kratochvíl, Tuza and Voigt [24], which one might imagine makes the colouring task much easier. Let $G = (V, E)$ be a graph. For a positive integer k, a mapping $L: V \to \binom{\mathbb{Z}^+}{k}$ is called a *k-list assignment* of G, and a colouring c of V is called an *L-colouring* if $c(v) \in L(v)$ for any $v \in V$. We say a k-list assignment L has *maximum separation* if $|L(u) \cap L(v)| \leqslant 1$ for any edge uv of G, and G is *separation k-choosable* if there is a proper L-colouring of G for any k-list assignment L with maximum separation. The *separation choosability* $\mathrm{ch}_{\mathrm{sep}}(G)$ of G is the least k such that G is separation k-choosable. Kratochvíl, Tuza and Voigt also considered weaker separation, where $|L(u) \cap L(v)| \leqslant q$ for any edge uv, but we only treat the most restrictive non-trivial case; note that the colouring task is trivial if $q = 0$. The original parameter *choosability* $\mathrm{ch}(G)$ of G demands no separation on the lists, and so we have $\mathrm{ch}(G) \geqslant \mathrm{ch}_{\mathrm{sep}}(G)$ always.

† This research was partly supported by a Van Gogh grant, reference 35513NM.

‡This author is partially supported by ANR Projects STINT (ANR-13-BS02-0007) and GATO (ANR-16-CE40-0009-01), and LabEx PERSYVAL-Lab (ANR-11-LABX-0025).

§This author is partially supported by a Vidi grant (639.032.614) of the Netherlands Organisation for Scientific Research (NWO).

© Cambridge University Press 2019. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

https://doi.org/10.1017/S0963548319000026 Published online by Cambridge University Press

Let us first review some well-known results for choosability. It was already known to Erdős, Rubin and Taylor [14] that there are bipartite graphs with arbitrarily large choosability. More specifically, by a connection to the extremal behaviour of the set-theoretic 'Property B', they showed, as $d \to \infty$, that the complete d-regular bipartite graphs $K_{d,d}$ satisfy

$$\mathrm{ch}\,(K_{d,d}) = (1 + o(1)) \log_2 d. \tag{1.1}$$

Somewhat later, Alon demonstrated with a probabilistic argument that just having large minimum degree forces the choosability of a graph to be large. The following considerably generalizes the lower bound in (1.1).

Theorem 1.1. (Alon [3]). *There is a constant $C > 0$ such that* $\mathrm{ch}\,(G) \geqslant C \log d$ *for any graph G with minimum degree d.*

Using containers, Saxton and Thomason [31] recently improved this lower bound to $(1 + o(1)) \log_2 d$ as $d \to \infty$, which is asymptotically optimal by (1.1).

One might wonder to what extent the results above remain valid when we impose maximum separation on the lists. Füredi, Kostochka and Kumbhat [15] already considered the complete d-regular bipartite graphs. By the use of an elegant extremal set-theoretic construction (see Lemma 2.1 below), they strengthened (1.1) by showing that, as $d \to \infty$,

$$\mathrm{ch}_{\mathrm{sep}}\,(K_{d,d}) = (1 + o(1)) \log_2 d. \tag{1.2}$$

In the first part of this paper, we generalize the lower bound in (1.2).

Theorem 1.2. *There is a constant $C > 0$ such that* $\mathrm{ch}_{\mathrm{sep}}\,(G) \geqslant C \log d$ *for any bipartite graph G with minimum degree d.*

This partially extends Theorem 1.1. It also incidentally improves upon a result of Molloy and Thron [29] about adaptable choosability, an implication we describe in Section 2.2. We give the proof of Theorem 1.2 in Section 2.1. The proof combines the construction of [15] with the probabilistic argument of [3], and it requires the assumption that G be bipartite.

Curiously, the question remains whether a form of Theorem 1.2 holds if we drop the bipartiteness assumption.

Conjecture 1.3. *There is a function $x_1(d)$ satisfying $x_1(d) \to \infty$ as $d \to \infty$ such that* $\mathrm{ch}_{\mathrm{sep}}\,(G) \geqslant x_1(d)$ *for any graph G with minimum degree d.*

If we only knew Theorem 1.1 for bipartite G, we could still conclude the same way for general G due to the fact that every graph of average degree at least d contains a bipartite subgraph of minimum degree at least $d/2$. This fact and Theorem 1.2 are not enough to derive the same conclusion for separation choosability of general G: a (bad) list assignment that has maximum separation does not necessarily keep it upon the addition of edges. On the other hand, Kratochvíl, Tuza and Voigt [23] have shown that the complete graph K_{d+1} on $d + 1$ vertices satisfies $\mathrm{ch}_{\mathrm{sep}}\,(K_{d+1}) \sim \sqrt{d}$ as $d \to \infty$. (So separation *does* help in this situation.) A natural tack therefore is, in any graph of given minimum degree, to look for either a large complete subgraph or a dense bipartite *induced* subgraph. More precisely, the following, if true, would imply Conjecture 1.3.

Conjecture 1.4. *There are functions $x_2(d)$ and $x_3(d)$ satisfying $x_2(d) \to \infty$ and $x_3(d) \to \infty$ as $d \to \infty$ such that any graph with minimum degree at least d contains a complete subgraph on $x_2(d)$ vertices or a bipartite induced subgraph with minimum degree at least $x_3(d)$.*

https://doi.org/10.1017/S0963548319000026 Published online by Cambridge University Press

In the second part of this paper, we consider some basic aspects of this conjecture. Irrespective of separation choosability, this basic Ramsey-type problem is intriguing.

In fact, there is a range of (quantitative) variations on Conjecture 1.4 which are also natural and interesting. Here follow two such variations.

Conjecture 1.5. *There is a constant $C > 0$ such that any triangle-free graph with minimum degree at least d contains a bipartite induced subgraph of minimum degree at least $C \log d$.*

If true, this would be sharp up to the choice of C (Theorem 3.7).

Conjecture 1.6. *There exist d_0 and g_0 such that any graph of girth at least g_0 with minimum degree at least d_0 contains a bipartite induced subgraph of minimum degree at least* 3.

If 3 is replaced by 2 in Conjecture 1.6, then the conclusion is equivalent to containing an even hole. This weaker statement is true with $g_0 = 4$ and $d_0 = 3$ by a result of Radovanović and Vušković [30, Theorem 1.6]. Note that even holes can be detected efficiently [11, 12], but detection of bipartite induced subgraphs of minimum degree at least 3 is an NP-complete problem [18].

In Section 3, we offer partial progress, towards Conjecture 1.5 especially. To that end we have found it especially useful that, to find a dense bipartite induced subgraph in a graph of given minimum degree, it suffices to find a good proper colouring.

Theorem 1.7. *Any graph with chromatic number at most k and minimum degree d has a bipartite induced subgraph of minimum degree at least $d/2k$.*

We show this as a corollary to a more general result below, Theorem 3.1. In Section 3.1, we illustrate how, together with classic results about colourings and stable sets in triangle-free graphs, Theorem 1.7 yields special cases of Conjecture 1.5: in particular, if the triangle-free graph is nearly regular (Theorem 3.3) or if it has sufficiently large minimum degree with respect to the number of vertices (Theorem 3.4). By a related but different method, we prove a weaker yet still sharp form of Conjecture 1.5 where edges are permitted for one of the two parts of the sought bipartite subgraph (Theorem 3.5).

2. Separation choosability of bipartite graphs

In this section we are primarily concerned with proving Theorem 1.2, but we also discuss some of the implications for another list colouring notion called adaptable choosability, which we discuss in more detail further on.

Although it is not needed for the proof, it may reveal a broader context to notice how Theorems 1.1 and 1.2 relate choosability and separation choosability to graph density. Recall the *degeneracy* of a graph is the maximum over all subgraphs of the minimum degree; a graph is *d-degenerate* if its degeneracy is at most d. A simple greedy argument yields $\mathrm{ch}(G) \leqslant d + 1$ for any d-degenerate graph G. Theorem 1.1 implies that choosability grows with degeneracy: there exists $C > 0$ such that

$$C \log d \leqslant \mathrm{ch}(G) \leqslant d + 1 \tag{2.1}$$

for any graph G with degeneracy d. Similarly, Theorem 1.2 allows us conclude for $\mathrm{ch}_{\mathrm{sep}}(G)$ as in (2.1) except only for *bipartite* G of degeneracy d.

We note that by an adaptation of a construction of Kostochka and Zhu [22], for every integer d there is a d-degenerate graph that is not separation d-choosable. Moreover, Alon, Kostochka, Reiniger, West and Zhu [6] have recently shown, with the help of an extremely elegant graph

https://doi.org/10.1017/S0963548319000026 Published online by Cambridge University Press

construction, that for every g and k there is a bipartite graph G of girth at least g that is not separation k-choosable, every proper subgraph of which has average degree at most $2(k-1)$.

2.1 Proof of Theorem 1.2

For the proof of Theorem 1.2, we require the construction [15, Corollary 1] used in the proof of (1.2). Two hypergraphs H_1 and H_2 on the same vertex set are *nearly disjoint* if every edge of H_1 meets every edge of H_2 in at most one vertex. The stability number $\alpha(H)$ of a hypergraph H is the size of a largest subset of vertices of H that does not contain any edge of H.

Lemma 2.1. (Füredi, Kostochka and Kumbhat [15]). *Fix $r \geqslant 2$. There exist two nearly disjoint r-uniform hypergraphs H_1 and H_2 each on vertex set $[4r^4]$ and having $16r^4 2^r$ edges such that $\alpha(H_1), \alpha(H_2) < 2r^4$.*

We essentially substitute this construction into the proof of Theorem 1.1.

Proof of Theorem 1.2. Let $G = (V, E)$ be a bipartite graph of minimum degree at least d, where without loss of generality $d \geqslant d_0$ for some sufficiently large fixed d_0. Let $V = V_1 \cup V_2$ be the bipartition and suppose that $|V_1| \leqslant |V_2|$. We assume for a contradiction that $\mathrm{ch}\,(G) \leqslant k$, where $k = \lceil C \log d \rceil - 1$ for some fixed $C > 0$ chosen strictly smaller than $(1/2 - \varepsilon)/\log 2$ with $\varepsilon > 0$. Let H_1 and H_2 be the hypergraphs certified by Lemma 2.1 applied with $r = k$. Write $W = [4k^4]$ for their common vertex set, and F_1 and F_2 for their respective edge sets.

As already mentioned, the strategy is the same as in [3], where we incorporate the use of F_1 and F_2. There are two stages of randomness. In the first stage we choose a small random vertex subset A of V_1 and assign lists from F_1 to the vertices of A randomly. After showing that with positive probability there is a subset $A \subseteq V_1$ for which there are many *good* vertices in V_2 (to be defined below), we fix such a subset A and an assignment of lists. In the second stage we assign lists uniformly to the good vertices, and only from F_2 to ensure separation of the list assignment. Goodness helps to guarantee in the second stage that with positive probability the remaining vertices cannot be coloured from their (random) lists.

Let $p = 1/\sqrt{d}$. Note that $p < 1/8$ for d sufficiently large. Let A be a random subset of V_1 with each vertex of V_1 chosen to be in A independently at random with probability p. Since $\mathbb{E}(|A|) = p|V_1|$, by Markov's inequality we have $\mathbb{P}(|A| > 2p|V_1|) \leqslant 1/2$. We define a random list assignment L_A of $G[A]$ as follows. For each $v \in A$ independently, let $L_A(v)$ be a uniformly random element of F_1. We call a vertex $v \in V_2$ *good* if for any $f \in F_1$ there is at least one neighbour v' of v such that $v' \in A$ and $L_A(v') = f$.

The probability that a vertex $v \in V_2$ is not good is the probability that for some $f \in F_1$ there is no neighbour v' of v such that $v' \in A$ and $L_A(v') = f$. For each fixed $f \in F_1$, since there are at least d neighbours of v, the probability that there is no neighbour v' of v such that $v' \in A$ and $L_A(v') = f$ must by the choice of C be at most

$$(1 - p/|F_1|)^d \leqslant \exp\left(-dp/|F_1|\right) \leqslant \exp\left(-\sqrt{d}/(16k^4 2^k)\right) \leqslant \exp\left(-d^{\varepsilon}\right)$$

for d sufficiently large. Therefore the probability that there is some set $f \in F_1$ that certifies that v is not good must be at most $16k^4 2^k \exp(-d^{\varepsilon}) < 1/4$ for d sufficiently large. It then follows by Markov's inequality that the number of vertices that are not good is at most $|V_2|/2$ with probability strictly greater than $1/2$. So there exists a set $A \subseteq V_1$ and a list assignment L_A (with lists from the set F_1) such that $|A| \leqslant 2p|V_1|$ and the number of good vertices is at least $|V_2|/2$. Fix such a set A and list assignment L_A for the remainder of the proof. Let $A^* \subseteq V_2$ be the set of good vertices.

There are at most $k^{|A|}$ possibilities for an arbitrary (proper) L_A-colouring of $G[A]$. We fix one such colouring c_A and show that with high probability there is a k-list assignment of maximum separation extending L_A, such that c_A cannot be extended to a proper list colouring of $G[A \cup A^*]$.

https://doi.org/10.1017/S0963548319000026 Published online by Cambridge University Press

We define a random list assignment L_{A^*} of $G[A^*]$ by letting, for each $v \in A^*$ independently, $L_{A^*}(v)$ be a uniformly random element of F_2. It is important to notice that, since G is bipartite and H_1 and H_2 are nearly disjoint, the resulting k-list assignment is guaranteed to have maximum separation.

Consider a vertex $v \in A^*$. We define C_v to be the set of colours c such that there is some neighbour v' of v such that $v' \in A$ and $c_A(v') = c$. Recall that $W = [4k^4]$ denotes the common vertex set of H_1 and H_2. We first observe that, since v is good, C_v meets every member of F_1. Thus $W \setminus C_v$ is a stable set of H_1, from which we conclude that $|W \setminus C_v| < 2k^4$. Moreover, there must be some $f_v \in F_2$ such that $f_v \subseteq C_v$; otherwise, C_v is a stable set of H_2 and so $|C_v| < 2k^4$, which in combination with the previous inequality is a contradiction to $|W| = 4k^4$. By the definition of C_v, if the random assignment results in $L_{A^*}(v) = f_v$, then v cannot be properly coloured from its list. Since the choice of $L_{A^*}(v)$ is independent, the probability that all vertices of A^* can be coloured from their lists is at most

$$(1-1/|F_2|)^{|V_2|/2} \leqslant \exp(-|V_2|/(2|F_2|)).$$

Note that, due to the choices of k, C and p and the fact that $|V_2| \geqslant |V_1| \geqslant d$,

$$k^{|A|} \exp(-|V_2|/(2|F_2|)) \leqslant k^{2p|V_1|} \exp(-|V_2|/(2|F_2|)) \leqslant k^{2p|V_2|} \exp(-|V_2|/d^{1/2-\varepsilon}) < 1$$

for d sufficiently large. It then follows that with positive probability there is a k-list assignment L of $G[A \cup A^*]$ with maximum separation such that there is no proper L-colouring of $G[A \cup A^*]$. □

2.2 Adaptable choosability

Given a graph $G = (V, E)$ and a labelling $\ell: E \to [k]$ of the edges, a (not-necessarily-proper) vertex colouring $c: V \to [k]$ is *adapted* to ℓ if, for every edge $e = uv \in E$, not all of $c(u)$, $c(v)$ and $\ell(e)$ are the same value. We say that G is *adaptable k-choosable* if, for any k-list assignment L and any labelling ℓ of the edges of G, there is an L-colouring of G that is adapted to ℓ. The *adaptable choosability* $\mathrm{ch}_a(G)$ of G is the least k such that G is adaptable k-choosable. Every proper colouring is adapted to any labelling ℓ, so $\mathrm{ch}(G) \geqslant \mathrm{ch}_a(G)$ always. This parameter was proposed by Kostochka and Zhu [22]. Molloy and Thron [29] proved that the adaptable choosability grows with choosability. Recall (2.1). More precisely, they proved the following.

Theorem 2.2. (Molloy and Thron [29]). *There is a constant $C > 0$ such that $\mathrm{ch}_a(G) \geqslant C \log^{1/5} d$ for any graph G with minimum degree d.*

We observe separation choosability is at most adaptable choosability.

Proposition 2.3. *For any graph G, $\mathrm{ch}_a(G) \geqslant \mathrm{ch}_{\mathrm{sep}}(G)$.*

Proof. Fix $G = (V, E)$ and let $k = \mathrm{ch}_a(G)$. Let L be a k-list assignment of maximum separation. Let ℓ be a labelling defined for each $e = uv \in E$ by taking $\ell(e)$ as the unique element of $L(u) \cap L(v)$ if it is non-empty, and arbitrary otherwise. By the choice of k, there is guaranteed to be an L-colouring c that is adapted to ℓ. Due to the maximum separation property of L and the definition of ℓ, the colouring c must be proper. □

By this observation and monotonicity of ch_a with respect to subgraph inclusion, Theorem 1.2 implies the following improvement upon Theorem 2.2.

Corollary 2.4. *There is a constant $C > 0$ such that $\mathrm{ch}_a(G) \geqslant C \log d$ for any graph G with minimum degree d.*

https://doi.org/10.1017/S0963548319000026 Published online by Cambridge University Press

This is sharp up to the choice of constant C by (1.1). By (2.1), it follows for some $C > 0$ that $\mathrm{ch}_a(G) \geqslant C \log \mathrm{ch}(G)$ for any graph G. However, it has not been ruled out that $\mathrm{ch}_a(G)$ is $\Omega(\sqrt{\mathrm{ch}(G)})$ in general.

3. Dense bipartite induced subgraphs

In this section, we focus mostly on Conjecture 1.5. That is because we find the triangle-free case the most elegant setting for this class of problems. Furthermore, the techniques carry over routinely to larger forbidden cliques; we briefly discuss this at the end of the section.

As every graph of average degree d contains a (bipartite) subgraph of minimum degree at least $d/2$, and since we do not worry ourselves about constant factors, we work with average or minimum degree interchangeably.

A first intuition one might have upon meeting this class of problems is that a good proper colouring in a graph should be helpful to find a dense bipartite induced subgraph. If there are few colour classes in such a colouring, then by the pigeonhole principle one expects some pair of the colour classes to have a relatively large number of edges between them. Although this idea in itself does not quite lead directly to the desired induced subgraph (and thus to Theorem 1.7), this intuition can be formalized and also strengthened, as we now show.

Given a graph $G = (V, E)$, let us say that a probability distribution $\mathcal{S}$ over the stable sets of G satisfies property Q^*_r if $\mathbb{P}(v \in \mathbf{S}) \geqslant r$ for every vertex $v \in V$ and $\mathbf{S}$ a random stable set taken according to $\mathcal{S}$. Recall that the *fractional chromatic number* of G may be defined as the smallest k such that there is a probability distribution over the stable sets of G that satisfies property $\mathrm{Q}^*_{1/k}$.

Theorem 3.1. *Any graph with fractional chromatic number k and average degree d has a bipartite induced subgraph of average degree at least d/k.*

By considering the distribution that takes a colour class of a proper colouring uniformly at random, it is easily seen that the fractional is bounded by the usual chromatic number, and so this immediately implies Theorem 1.7. As we will see in Section 3.3, this lower bound cannot in general be improved by more than a constant factor.

Proof of Theorem 3.1. Let $G = (V, E)$ be a graph and let $\mathcal{S}$ be a probability distribution over the stable sets of G with property $\mathrm{Q}^*_{1/k}$. Without loss of generality, we may assume that $\mathbb{P}(v \in \mathbf{S}) = 1/k$ for every vertex $v \in V$ and $\mathbf{S}$ a random stable set taken according to $\mathcal{S}$. Let $\mathbf{S_1}$ and $\mathbf{S_2}$ be two stable sets taken independently at random according to $\mathcal{S}$. Note that $\mathbb{E}(|\mathbf{S_1}|) = \mathbb{E}(|\mathbf{S_2}|) = n/k$ by the assumption on $\mathcal{S}$. Moreover, for any edge $e = uv \in E$, the probability that e is in the induced subgraph $G[\mathbf{S_1} \cup \mathbf{S_2}]$ is

$$\mathbb{P}(u \in \mathbf{S_1})\,\mathbb{P}(v \in \mathbf{S_2}) + \mathbb{P}(u \in \mathbf{S_2})\,\mathbb{P}(v \in \mathbf{S_1}) = \frac{2}{k^2}.$$

By linearity of expectation, we have that

$$\mathbb{E}\left(|E(G[\mathbf{S_1} \cup \mathbf{S_2}])| - (|\mathbf{S_1}| + |\mathbf{S_2}|)\frac{d}{2k}\right) = \frac{2|E|}{k^2} - \frac{nd}{k^2} = 0.$$

The probabilistic method guarantees stable sets S_1 and S_2 of G such that

$$|E(G[S_1 \cup S_2])| \geqslant (|S_1| + |S_2|)\frac{d}{2k}.$$

Discarding the vertices of $S_1 \cap S_2$ (if any) yields a bipartite induced subgraph of average degree at least d/k. □

https://doi.org/10.1017/S0963548319000026 Published online by Cambridge University Press

3.1 Subcases of Conjecture 1.5

In this subsection, we discuss some special cases of Conjecture 1.5. These observations are mainly consequences of Theorem 3.1 above.

First let us note that, by an induction on the number of vertices, we may assume in Conjecture 1.5 that any proper subgraph of the triangle-free graph has minimum degree less than d. (Note that the base case is implied by Turán's theorem.) So the graph may be assumed to be d-degenerate. By Theorem 3.1, Conjecture 1.5 is thus implied by a recent conjecture of Harris.

Conjecture 3.2. (Harris [17]). *There is some $C > 0$ such that the fractional chromatic number of any d-degenerate triangle-free graph is at most $Cd/\log d$.*

Next we observe that, due to classic results on the chromatic number of triangle-free graphs and on off-diagonal Ramsey numbers, we may also assume in Conjecture 1.5 that the maximum degree of the graph is large with respect to d and that the minimum degree d is not too large with respect to the number of vertices.

Theorem 3.3. *There is a constant $C > 0$ such that any triangle-free graph with average degree d and maximum degree Δ contains a bipartite induced subgraph of average degree at least $(d/C\Delta)\log\Delta$.*

Proof. Let G be a triangle-free graph with average degree d and maximum degree Δ. A result of Johansson [20], recently improved significantly by Molloy [28], says that there is a constant $C > 0$ such that any triangle-free graph with maximum degree Δ, and thus G, has chromatic number at most $C\Delta/\log\Delta$. The result now follows from Theorem 3.1. □

Theorem 3.4. *There is a constant C_1 such that any triangle-free graph on n vertices with average degree $d \geqslant Cn^{2/3}\sqrt{\log n}$ contains a bipartite induced subgraph of average degree at least $(CC_1/2)\log n$ for all n large enough.*

Proof. Let G be a triangle-free graph on n vertices with average degree $d \geqslant Cn^{2/3}\sqrt{\log n}$. By a classic result of Ajtai, Komlós and Szemerédi [1], the off-diagonal Ramsey numbers $R(3,\ell)$ satisfy $R(3,\ell) = O(\ell^2/\log\ell)$ as $\ell \to \infty$. Thus, for some $C_1 > 0$, every induced subgraph of G with at least $n^{2/3}$ vertices contains a stable set of size at least $C_1 n^{1/3}\sqrt{\log n}$. We now remove such stable sets (of size at least $C_1 n^{1/3}\sqrt{\log n}$) until fewer than $n^{2/3}$ vertices remain in G. Note that the subgraph H of G induced by the union of these stable sets has chromatic number at most

$$\frac{n}{C_1 n^{1/3}\sqrt{\log n}} \leqslant \frac{n^{2/3}}{C_1\sqrt{\log n}}.$$

For large enough n, the number of edges in H is at least

$$\frac{nd}{2} - n\cdot n^{2/3} \geqslant \frac{C}{4}n^{5/3}\sqrt{\log n},$$

and so H has average degree at least $(C/2)n^{2/3}\sqrt{\log n}$. By Theorem 3.1, G has a bipartite induced subgraph of average degree at least $(CC_1/2)\log n$. □

In summary, Theorem 3.1 and the discussion above imply that in Conjecture 1.5 it suffices to consider, for any $\varepsilon > 0$, a triangle-free graph G on n vertices of average degree d, where $d < \varepsilon n^{2/3}\sqrt{\log n}$, such that G is d-degenerate and contains a vertex of degree more than d/ε. Moreover, it suffices to show in these circumstances that the fractional chromatic number of G is at most $Cd/\log d$ for some $C > 0$.

https://doi.org/10.1017/S0963548319000026 Published online by Cambridge University Press

3.2 Dense semi-bipartite subgraphs

Instead of a dense bipartite *induced subgraph*, we might be happy with a dense bipartite *subgraph* where we only demand (at least) one of the parts induces a stable set. Given a graph $G = (V, E)$, let us call an induced subgraph $G' = (V', E')$ of G *semi-bipartite* if it admits a partition $V' = V_1 \cup V_2$ such that V_1 is a stable set of G, and define the average degree of G' with respect to the semi-bipartition as the average degree of the bipartite subgraph $G[V_1, V_2]$ between V_1 and V_2 (and so we ignore any edges in V_2).

In this subsection, we prove the following.

Theorem 3.5. *Any triangle-free graph of minimum degree $d \geqslant 1$ contains a semi-bipartite induced subgraph of average degree at least $\frac{1}{2}\log d$.*

As we will show in Section 3.3, this is sharp up to a constant factor.

Near the end of the book by Alon and Spencer [5, pp. 321–322], there is a proof of a result of Ajtai, Komlós and Szemerédi [2] that any n-vertex triangle-free graph of maximum degree Δ has a stable set of size at least $n \log \Delta/(8\Delta)$. That proof heavily inspired the proof of the following result. This result directly implies Theorem 3.5.

Lemma 3.6. *Let G be a triangle-free graph of minimum degree $d \geqslant 1$, and $\mathbf{S}$ a random stable set of G chosen uniformly. Then*

$$\mathbb{E}\left(\sum_{v\in\mathbf{S}} d(v)\right) \geqslant \frac{1}{4}\sum_{v\in G}\log d(v).$$

Proof. For each vertex v, let $X_v = d(v)\cdot|\{v\}\cap\mathbf{S}| + |N(v)\cap\mathbf{S}|$. We claim that $\mathbb{E}(X_v) \geqslant \frac{1}{2}\log d(v)$.

To see that this claim holds, consider any stable set T of $G - N[v]$ (where $N[v]$ denotes the closed neighbourhood of v in G). It is enough to prove that for each choice of T as above,

$$\mathbb{E}(X_v \mid \mathbf{S}\cap(V(G) - N[v]) = T) \geqslant \frac{1}{2}\log d(v).$$

Given a stable set T as above, let k be the number of neighbours of v with no neighbour in T. Conditioning on $\mathbf{S}\cap(V(G) - N[v]) = T$, there are precisely $2^k + 1$ equally likely possibilities for $\mathbf{S}$ (T together with v, or T together with one of the 2^k subsets of the set of neighbours of v that have no neighbour in T). It follows that

$$\mathbb{E}(X_v \mid \mathbf{S}\cap(V(G) - N[v]) = T) = \frac{1}{2^k+1}(d(v) + k2^{k-1}).$$

It can be verified that this last quantity is at least $\frac{1}{2}\log d(v)$ for any $d(v) \geqslant 1$ and $k \geqslant 0$; more details are given in the Appendix. This proves the claim.

Note that $\sum_{v\in G} X_v = 2\sum_{v\in S} d(v)$. So by linearity of expectation and the claim above,

$$\mathbb{E}\left(\sum_{v\in S} d(v)\right) = \frac{1}{2}\sum_{v\in G}\mathbb{E}(X_v) \geqslant \frac{1}{4}\sum_{v\in G}\log d(v). \qquad \square$$

Proof of Theorem 3.5. Let G be an n-vertex triangle-free graph of minimum degree $d \geqslant 1$. Let $\mathbf{S}$ be a stable set of G chosen uniformly at random. By Lemma 3.6, the expected number of edges between $\mathbf{S}$ and its complement is at least $\frac{1}{4}n\log d$, and thus the expected average degree of the corresponding semi-bipartite subgraph is at least $\frac{1}{2}\log d$. This proves the existence of the desired semi-bipartite induced subgraph. $\square$

https://doi.org/10.1017/S0963548319000026 Published online by Cambridge University Press

Conjecture 1.5 holds if we can find a distribution $\mathcal{S}$ over the stable sets of the triangle-free graph G that satisfies property $Q^*_{\log d/Cd}$. From the proof of Lemma 3.6, it is possible to find a probability distribution $\mathcal{S}$ over the stable sets of G such that $\mathbb{E}(|N(v) \cap \mathbf{S}|) \geqslant \frac{1}{8} \log d$ or $\mathbb{P}(v \in \mathbf{S}) \geqslant \log d/4d$ for every vertex v and $\mathbf{S}$ a random stable set taken according to $\mathcal{S}$. Unfortunately this property does not seem to be powerful enough to prove Conjecture 1.5.

3.3 Upper bounds

In this subsection, we prove that Conjecture 1.5, if true, would be sharp up to a constant factor.

Theorem 3.7. *There are constants $C, C', C'' > 0$ such that for every large enough n there is a triangle-free graph on n vertices with all degrees between $C'n^{1/3}$ and $C''n^{1/3}$ that contains no semi-bipartite induced subgraph of average degree at least $C \log n$.*

This is a relatively routine probabilistic construction using the binomial random graph, but we include the details for completeness. Theorem 3.7 also certifies sharpness of the bounds in Theorems 1.7, 3.1, 3.3 and 3.5.

Before continuing, let us make the convenient observation that, if we do not mind constant factors, it suffices to consider only semi-bipartite subgraphs with both parts of equal size.

Proposition 3.8. *Suppose $A, B \subseteq G$ are disjoint with $|A| \geqslant |B|$ and satisfy that the average degree of $G[A, B]$ is d. Then there exists $A' \subseteq A$ with $|A'| = |B|$ such that the average degree of $G[A', B]$ is at least $d/2$.*

Proof. Take A' to be the vertices in A with the $|B|$ largest degrees. The number of edges in $G[A', B]$ is at least

$$\frac{|B|}{|A|}|E(G[A, B])| \geqslant \frac{|B|}{|A|} \cdot \frac{d}{2}(|A| + |B|).$$

Thus the average degree of $G[A', B]$ is at least $d(|A| + |B|)/2|A| \geqslant d/2$. □

In the following result we determine up to a constant factor the largest average degree over all semi-bipartite induced subgraphs in the binomial random graph. For Theorem 3.7, we only need the upper bound.

Proposition 3.9. *In the binomial random graph on $[n]$ with edge probability p, where $np \to \infty$ as $n \to \infty$ and $p < 0.99$, the largest average degree of a semi-bipartite induced subgraph is $\Theta(\log np)$ with high probability.*

Proof. Fix $\varepsilon > 0$. With high probability, the stability number is at most $(2 + \varepsilon) \log_b np$, where $b = 1/(1 - p)$ (*cf.* [16]). By Proposition 3.8, we can just consider disjoint vertex subsets A, B with $|A| = |B| = k$, where $k \leqslant (2 + \varepsilon) \log_b np$. For a given k, the expected number of such A, B where $G[A, B]$ has average degree at least $C \log np$ is as $n \to \infty$ at most

$$\binom{n}{2k}\binom{2k}{k} \mathbb{P}(\text{Bin}(k^2, p) \geqslant Ck \log np) \leqslant \left(\frac{en}{k}\right)^{2k} \exp(-Ck \log np) < \exp\left(-\frac{C}{2}k \log np\right),$$

where we used a Chernoff bound (*cf.* [19, Corollary 2.4]) and took a sufficiently large fixed choice of $C > 0$. Crudely summing this estimate in the range $C \log np \leqslant k \leqslant (2 + \varepsilon) \log_b np$, Markov's inequality implies the chance there is a semi-bipartite induced subgraph with average degree at least $C \log np$ is, as $n \to \infty$, at most

https://doi.org/10.1017/S0963548319000026 Published online by Cambridge University Press

$$(2+\varepsilon)(\log_b np)\exp\left(-\frac{C^2}{2}\log^2 np\right)\to 0$$

for C chosen large enough.

For the lower bound, first note by a Chernoff bound (*cf.* [19, Corollary 2.3]) that the number of edges is $(1+o(1))n^2p/2$, and so the minimum degree is at least $(1+o(1))np/2$, with high probability. By classic results on the chromatic number of the random graph [8, 26, 27], with high probability the chromatic number is $(1+o(1))n/(2\log_b np)$. It therefore follows from Theorem 3.1 that there is a bipartite induced subgraph with average degree at least $(1+o(1))p\log_b np$ with high probability, as required. □

Proof of Theorem 3.7. Let $p = Dn^{-2/3}$ for some fixed $0 < D < 2^{-1/4}$. Consider the binomial random graph on $[n]$ with edge probability p. The expected degree of a vertex is $p(n-1) \sim Dn^{1/3}$. By a Chernoff (*cf.* [19, Corollary 2.3]) and a union bound, the minimum degree is less than $Dn^{1/3}/2$ or the maximum degree is more than $3Dn^{1/3}/2$ with probability at most $2n\exp(-Dn^{1/3}/12)\to 0$ as $n\to\infty$. The expected number of triangles is at most $n^3p^3/6 = D^3n/6$. By Markov's inequality, the probability that there are at least $D^3n/3$ triangles is at most $1/2$. By Proposition 3.9 there exists $C>0$ such that, for all n large enough, with positive probability there is a graph G that has minimum degree at least $Dn^{1/3}/2$, maximum degree at most $3Dn^{1/3}/2$, fewer than $D^3n/3$ triangles, and no semi-bipartite induced subgraph of average degree at least $C\log np$. Assuming n is large enough, fix such a graph G and remove an arbitrary vertex from each triangle. This leaves a triangle-free graph G^* with at least $(1-D^3/3)n>0$ vertices and more than $Dn^{4/3}/4 - (D^3n/3)(3Dn^{1/3}/2) = D(1-2D^4)n^{4/3}/4 > 0$ edges, so all degrees between $D(1-2D^4)n^{1/3}/4$ and $3Dn^{1/3}/2$. Moreover, by monotonicity G^* contains no semi-bipartite induced subgraph of average degree at least $C\log n \geqslant C\log np$. □

3.4 K_r-free graphs

In this section we explain how to extend the results obtained previously for triangle-free graphs to the case of K_r-free graphs, with $r\geqslant 4$.

The following result has the same proof as that of Theorem 3.3, but uses instead the fact that K_r-free graphs of maximum degree Δ have chromatic number at most $200r\Delta\log\log\Delta/\log\Delta$. This was recently proved by Molloy [28], improving an earlier result of Johansson [21].

Theorem 3.10. *There is some constant $C>0$ such that, for every $r\geqslant 4$, any K_r-free graph with average degree d and maximum degree Δ contains a bipartite induced subgraph of average degree at least*

$$\frac{d\log\Delta}{Cr\Delta\log\log\Delta}.$$

□

The next result is shown in the same way as Theorem 3.4, but using a slightly more general bound [1] for the off-diagonal Ramsey numbers: $R(r,\ell) = O(\ell^{r-1}/(\log\ell)^{r-2})$ as $\ell\to\infty$.

Theorem 3.11. *For each $r\geqslant 3$, any K_r-free graph G on n vertices with average degree $d = \Omega(n^{1-1/r}(\log n)^{1/(r-1)})$ contains a bipartite induced subgraph of average degree $\Omega(\log n)$ as $n\to\infty$.*

Note also that a refined analysis in the proof of Lemma 3.6, using [32, Lemma 1] instead of the simple counting argument, yields that there is an absolute constant $C>0$, such that in a K_r-free graph G, the expected sum of the degrees of the vertices in a random stable set (from the uniform distribution) is at least the sum of

https://doi.org/10.1017/S0963548319000026 Published online by Cambridge University Press

$$C\frac{\log d(v)}{r\log\log d(v)},$$

over all the vertices $v \in G$. As a consequence, we obtain the following K_r-free analogue of Theorem 3.5.

Theorem 3.12. *There is a constant $C > 0$ such that, for any $r \geqslant 4$, any K_r-free graph of minimum degree $d \geqslant 3$ contains a semi-bipartite induced subgraph of average degree at least*

$$C\frac{\log d}{r\log\log d}.$$ □

For completeness, we merely state two upper bounds relevant to Conjectures 1.4 and 1.6. The proofs routinely adapt the proof of Theorem 3.7.

Proposition 3.13. *Fix an integer $g \geqslant 4$. There is a constant $C > 0$ such that for every large enough n there is a graph on n vertices of girth at least g with minimum and maximum degrees both $\Theta(n^{1/(g-1)})$ that contains no semi-bipartite induced subgraph of average degree at least $C\log n$.*

Proposition 3.14. *Fix an integer $r \geqslant 3$. There is a constant $C > 0$ such that for every large enough n there is a K_r-free graph on n vertices with minimum and maximum degrees both $\Theta(n^{1-2/r})$ that contains no semi-bipartite induced subgraph of average degree at least $C\log n$.*

4. Conclusion

In the course of our research, we found alternative derivations of Theorems 3.3–3.5 and of the lower bound in Proposition 3.9. We have omitted these for brevity, but mention here that, like Theorem 3.1, they all rely on some appropriate proper colouring. Since it is easy to construct graphs with very high degree bipartite induced subgraphs and also high (fractional) chromatic number, naturally we wonder if there are other, possibly more 'direct', methods to produce dense bipartite induced subgraphs. In particular, this might help to improve upon the constant factors in our bounds.

Regardless of the eventual status of Conjecture 1.5, Theorems 3.4 and 3.7 hint at the following more refined problem.

Problem 4.1. *Fix $\eta \in (0, 1)$. As $n \to \infty$, determine the asymptotic infimum $f_\eta(n)$ of the largest minimum degree of a bipartite induced subgraph over all triangle-free graphs of minimum degree n^η.*

Theorem 3.7 implies that $f_{1/3}(n)$ is at most logarithmic in n, while a simple modification of Theorem 3.4 implies that $f_{2/3+\varepsilon}(n)$, for fixed $\varepsilon > 0$, is at least polynomial in n. If a guiding 'paradigm' that couples the triangle-free process with the Erdős–Rényi process (*cf. e.g.* [7]) holds as well for our problem, then Proposition 3.9 suggests that perhaps η being around $1/2$ is the transition point between logarithmic and polynomial behaviour for $f_\eta(n)$. Note that the $\eta = 1$ case is related (via Theorem 1.7) to a problem of Erdős and Simonovits: *cf.* [9].

We find it difficult to imagine that Conjecture 1.3 is not true, but keep in mind that we have only tried one particular approach.

Our study in fact began by considering the separation choosability analogue of a conjecture of Alon and Krivelevich [4]. Although our paper is already brimming with conjecture, we still find this worth highlighting. Both this and the original conjecture (for choosability) remain open, and, if true, would considerably generalize the *upper* bounds in (1.2) and (1.1), respectively.

https://doi.org/10.1017/S0963548319000026 Published online by Cambridge University Press

Conjecture 4.2. *There is a constant $C > 0$ such that* $\mathrm{ch}_{\mathrm{sep}}(G) \leqslant C \log \Delta$ *for any bipartite graph G with maximum degree Δ.*

Acknowledgements

We are grateful to François Pirot for permission to include his improvement on our original Theorem 3.1. We are also thankful for the helpful comments from the anonymous referee.

Note added

Since the posting of our manuscript to arXiv, the second half of our paper has precipitated a number of significant further developments.

- Up to a logarithmic factor, Problem 4.1 has been solved, independently, by Cames van Batenburg, de Joannis de Verclos, Pirot and the second author [10] and by Kwan, Letzter, Sudakov and Tran [25].
- In [25], they moreover proved Conjecture 1.4 (and thus confirmed Conjecture 1.3) and have nearly settled Conjecture 1.5 (and thus confirmed Conjecture 1.6), in that they have established $\Omega(\log d/\log\log d)$ bipartite induced minimum degree in K_r-free graphs for every fixed $r \geqslant 3$.
- Work of Davies, de Joannis de Verclos, Pirot and the second author [13] has improved the asymptotic leading constant to 2 in Theorem 3.5.

References

[1] Ajtai, M., Komlós, J. and Szemerédi, E. (1980) A note on Ramsey numbers. *J. Combin. Theory Ser. A* **29** 354–360.
[2] Ajtai, M., Komlós, J. and Szemerédi, E. (1981) A dense infinite Sidon sequence. *European J. Combin.* **2** 1–11.
[3] Alon, N. (2000) Degrees and choice numbers. *Random Struct. Alg.* **16** 364–368.
[4] Alon, N. and Krivelevich, M. (1998) The choice number of random bipartite graphs. *Ann. Combin.* **2** 291–297.
[5] Alon, N. and Spencer, J. H. (2008) *The Probabilistic Method*, third edition, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley.
[6] Alon, N., Kostochka, A., Reiniger, B., West, D. B. and Zhu, X. (2016) Coloring, sparseness and girth. *Israel J. Math.* **214** 315–331.
[7] Bohman, T. and Keevash, P. (2010) The early evolution of the H-free process. *Invent. Math.* **181** 291–336.
[8] Bollobás, B. (1988) The chromatic number of random graphs. *Combinatorica* **8** 49–55.
[9] Brandt, S. and Thomassé, S. (2011) Dense triangle-free graphs are four-colorable: A solution to the Erdős–Simonovits problem. Manuscript.
[10] Cames van Batenburg, W., de Joannis de Verclos, R., Kang, R. J. and Pirot, F. (2018) Bipartite induced density in triangle-free graphs. arXiv:1808.02512
[11] Chudnovsky, M., Kawarabayashi, K.-i. and Seymour, P. (2005) Detecting even holes. *J. Graph Theory* **48** 85–111.
[12] Conforti, M., Cornuéjols, G., Kapoor, A. and Vušković, K. (2002), Even-hole-free graphs, part II: Recognition algorithm. *J. Graph Theory* **40** 238–266.
[13] Davies, E., de Joannis de Verclos, R., Kang, R. J. and Pirot, F. (2018) Colouring triangle-free graphs with local list sizes. arXiv:1812.11152
[14] Erdős, P., Rubin, A. L. and Taylor, H. (1980) Choosability in graphs. In *Proceedings of the West Coast Conference on Combinatorics, Graph Theory and Computing*, Utilitas Mathematica, Winnipeg, pp. 125–157.
[15] Füredi, Z., Kostochka, A. and Kumbhat, M. (2014) Choosability with separation of complete multipartite graphs and hypergraphs. *J. Graph Theory* **76** 129–137.
[16] Grimmett, G. R. and McDiarmid, C. J. H. (1975) On colouring random graphs. *Math. Proc. Cambridge Philos. Soc.* **77** 313–324.
[17] Harris, D. G. (2016) Some results on chromatic number as a function of triangle count. arXiv:1604.00438
[18] Havet, F. (2017) Finding more than an even hole. Manuscript.
[19] Janson, S., Łuczak, T. and Ruciński, A. (2000) *Random Graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience.

https://doi.org/10.1017/S0963548319000026 Published online by Cambridge University Press

[20] Johansson, A. (1996) Asymptotic choice number for triangle-free graphs. Technical Report 91-5, DIMACS.
[21] Johansson, A. (1996) The choice number of sparse graphs. Manuscript.
[22] Kostochka, A. V. and Zhu, X. (2008) Adapted list coloring of graphs and hypergraphs. *SIAM J. Discrete Math.* **22** 398–408.
[23] Kratochvíl, J., Tuza, Z. and Voigt, M. (1998) Brooks-type theorems for choosability with separation. *J. Graph Theory* **27** 43–49.
[24] Kratochvíl, J., Tuza, Z. and Voigt, M. (1998) Complexity of choosing subsets from color sets. *Discrete Math.* **191** 139–148.
[25] Kwan, M., Letzter, S., Sudakov, B. and Tran, T. (2018) Dense induced bipartite subgraphs in triangle-free graphs. arXiv:1810.12144
[26] Łuczak, T. (1991) The chromatic number of random graphs. *Combinatorica* **11** 45–54.
[27] Matula, D. W. and Kučera, L. (1990) An expose-and-merge algorithm and the chromatic number of a random graph. In *Random Graphs '87* (M. Karoński, J. Jaworski and A. Ruciński, eds), Wiley, pp. 175–187.
[28] Molloy, M. (2019) The list chromatic number of graphs with small clique number. *J. Combin. Theory Ser. B* **134** 264–284.
[29] Molloy, M. and Thron, G. (2011) The adaptable choosability number grows with the choosability number. *Discrete Math.* **311** 2268–2271.
[30] Radovanović, M. and Vušković, K. (2013) A class of three-colorable triangle-free graphs. *J. Graph Theory* **72** 430–439.
[31] Saxton, D. and Thomason, A. (2015) Hypergraph containers. *Invent. Math.* **201** 925–992.
[32] Shearer, J. B. (1995) On the independence number of sparse graphs. *Random Struct. Alg.* **7** 269–271.
[33] Vizing, V. G. (1976) Coloring the vertices of a graph in prescribed colors. *Diskret. Analiz* **29** 10.

Appendix: Technical details in the proof of Lemma 3.6

In this section, we provide further details of the proof that, for any $d \geqslant 1$ and $0 \leqslant k \leqslant d$, it holds that

$$\frac{d + k2^{k-1}}{2^k + 1} \geqslant \frac{1}{2} \log d. \tag{A.1}$$

Inequality (A.1) is trivially satisfied if $d = 1$, so we can assume that $d \geqslant 2$. The inequality certainly holds if $d/(2^k + 1) \geqslant \frac{1}{2} \log d$, so we can also assume that $2^k + 1 > 2d/\log d$, which is equivalent to $k > \log_2 (2d/\log d - 1)$.

Inequality (A.1) also holds provided we can show $k2^{k-1}/(2^k + 1) \geqslant \frac{1}{2} \log d$. Note that the function $x \mapsto x2^{x-1}/(2^x + 1)$ is increasing for $x \geqslant 0$. It follows from our previous assumption that

$$\frac{k2^{k-1}}{2^k + 1} = \frac{1}{2} k(1 - (2^k + 1)^{-1}) > \frac{1}{2}\left(1 - \frac{\log d}{2d}\right) \log_2\left(\frac{2d}{\log d} - 1\right).$$

Thus for (A.1) it suffices to establish for $d \geqslant 2$ that

$$\left(1 - \frac{\log d}{2d}\right) \log_2\left(\frac{2d}{\log d} - 1\right) \geqslant \log d.$$

It is routinely checked that the function

$$x \mapsto \left(1 - \frac{\log x}{2x}\right) \log_2\left(\frac{2x}{\log x} - 1\right) - \log x \quad \text{for } x > 1$$

has a unique minimum of about 0.30 at some $x_0 \approx 9.74$.

We remark that as $d \to \infty$, the right-hand side of inequality (A.1) can be improved to $(1 - o(1)) \log_2 d$.

Cite this article: Esperet L, Kang RJ, and Thomassé S (2019). Separation Choosability and Dense Bipartite Induced Subgraphs. *Combinatorics, Probability and Computing* **28**, 720–732. https://doi.org/10.1017/S0963548319000026

https://doi.org/10.1017/S0963548319000026 Published online by Cambridge University Press