Received: 13 June 2018 | Revised: 13 February 2019 | Accepted: 25 February 2019

DOI: 10.1002/jgt.22457

ARTICLE

WILEY

Rainbow triangles and the Caccetta-Häggkvist conjecture

Ron Aharoni[1] | **Matthew DeVos**[2] | **Ron Holzman**[1]

[1]Department of Mathematics, Technion, Haifa, Israel

[2]Department of Mathematics, Simon Fraser University, Burnaby, British Columbia, Canada

Correspondence
Ron Aharoni, Department of Mathematics, Technion, Haifa 32000, Israel.
Email: raharoni@gmail.com

Funding information
RISE, Grant/Award Number: 823748; United States-Israel Binational Science Foundation, Grant/Award Number: 2006099; Israel Science Foundation, Grant/Award Numbers: 2023464, 409/16; ISF, Grant/Award Number: 409/16

Abstract

A famous conjecture of Caccetta and Häggkvist is that in a digraph on n vertices and minimum outdegree at least n/r there is a directed cycle of length r or less. We consider the following generalization: in an undirected graph on n vertices, any collection of n disjoint sets of edges, each of size at least n/r, has a rainbow cycle of length r or less. We focus on the case $r = 3$ and prove the existence of a rainbow triangle under somewhat stronger conditions than in the conjecture. In our main result, whenever n is larger than a suitable polynomial in k, we determine the maximum number of edges in an n-vertex edge-colored graph where all color classes have size at most k and there is no rainbow triangle. Moreover, we characterize the extremal graphs for this problem.

KEYWORDS

Caccetta-Häggkvist, edge-colored graph, rainbow triangles

1 | INTRODUCTION

The following conjecture is one of the best known in graph theory:

Conjecture 1.1 (Caccetta and Häggkvist [3]). *For every positive integer r, every digraph on n vertices with minimum outdegree at least n/r has a directed cycle of length at most r.*

This conjecture is trivial for $r \leq 2$ and has been proved for $r \geq \sqrt{2n}$ by Shen [22]. The special case $r = 3$ has received considerable attention [2,9,19,21]; the best known bound in this case is

[Correction added after online publication on 03th March, 2019: Missing Equation inserted on Page 7, lines 34 and 35; Page 8, lines 14, 40, and 41; Page 9, lines 30 and 31; also on Page 4, line 35, the square indicating the end of Proof has been removed and on Page 10, lines 19 to 24, italicized font has been changed to roman font. On Page 4, lines 10 and 11; Page 12, lines 26 and 29; Page 13, lines 11 and 12 the prime should be attached to the subscript and not to the line-level symbol.]

 © 2019 Wiley Periodicals, Inc.

given by a theorem of Hladký, Král', and Norin [10] asserting that every simple digraph on n vertices with minimum outdegree at least $0.3465n$ contains a directed triangle.

The following conjecture of the first author (unpublished) is a generalization of this conjecture, which concerns undirected, simple graphs. It replaces the notion of "directed cycle" by that of "rainbow cycle." Define an *edge-colored graph* to be a graph $G = (V, E)$ equipped with a distinguished partition $\{C_1, C_2, ..., C_p\}$ of E, called a *color partition*. The sets C_i are called *color classes*, the indices $i \in \{1, ..., p\}$ are called *colors*, and we say that an edge e has color i if $e \in C_i$. A subgraph H of G is called *rainbow* if distinct edges of H have distinct colors.

Conjecture 1.2. *Let r be a positive integer and let G be an edge-colored graph on n vertices. If there are n color classes, all of size at least n/r, then G contains a rainbow cycle of length at most r.*

To see that Conjecture 1.2 implies the original, let r be a positive integer and let D be a directed simple graph on a set V of n vertices with minimum outdegree at least n/r. Let G be the undirected underlying graph of D (which has the same vertex and edge set). Extend G to an edge-colored graph with the color partition $\{A_D^+(v) | v \in V\}$ (here $A_D^+(v)$ denotes the set of edges of D incident with v and directed away from it). Conjecture 1.2 implies that G has a rainbow cycle of length at most r, and it is easy to see that it corresponds to a directed cycle of length at most r in the original digraph D.

Since the Caccetta-Häggkvist conjecture is sharp, if true, so is this conjecture. The standard example showing that the requirement on the size of the color classes cannot be relaxed is obtained by taking a cycle of length $n = kr + 1$ with vertices $v_1, ..., v_n$, and defining C_i $(1 \le i \le n)$ to be the set of all edges $v_i v_{i+j}$, $1 \le j \le k$ (indices taken cyclically). Conjecture 1.2 is sharp also in another sense, namely, the requirement on the number of colors cannot be relaxed. An example showing that $n - 1$ colors do not suffice is obtained by adding a new vertex z, which is not part of the cycle, to the previous example (so now $n = kr + 2$), and adding to each C_i $(1 \le i \le n - 1)$ the edge $v_i z$. (Below, in Example 3.5, the case $r = 3$ of this example will serve yet another purpose.)

In this paper, we focus on the case $r = 3$ of Conjecture 1.2. This leads us to seek conditions on an edge-colored graph that guarantee the existence of a rainbow triangle. Such conditions have been studied in the literature [4,5,7,8,11–13]. The feature special to the present paper is that the size of the color classes plays the key role.

2 | TWO OBSERVATIONS

Some straightforward arguments give reasonably good approximations to Conjecture 1.2 for the special case $r = 3$. Henceforth we use the term *triangle* for a subgraph isomorphic to K_3. The case $r = 3$ of Conjecture 1.2 says that every edge-colored graph on n vertices with n color classes, each of size at least $n/3$, contains a rainbow triangle. The following theorem gives two approximations to this.

Theorem 2.1. *Let G be an edge-colored graph on n vertices. If either of the following conditions is satisfied, then G has a rainbow triangle.*

(1) *There are at least $(9/8)n$ color classes, each of size at least $n/3$.*
(2) *There are at least n color classes, each of size at least $2n/5$.*

To prove these statements, for every graph G let $t(G)$ denote the number of triangles in G.

Lemma 2.2. *Let G be an edge-colored graph with m edges without rainbow triangles. If every color class has size at most k, then $t(G) \leq (1/2)m(k-1)$.*

Proof. Let $\{C_1, ..., C_p\}$ be the color partition of G. For a triangle, there is a color class containing the majority of its edges. Every color i can be the majority for at most $\binom{|C_i|}{2}$ triangles. Thus

$$t(G) \leq \sum_{i=1}^{p} \binom{|C_i|}{2} = \frac{1}{2}\sum_{i=1}^{p} |C_i|^2 - \frac{1}{2}m \leq \frac{1}{2}\left(\frac{m}{k}k^2\right) - \frac{1}{2}m = \frac{1}{2}m(k-1).$$

□

The lemma bounds from above the number of triangles which can appear in an edge-colored graph without rainbow triangles. Contrarily, there is a rather straightforward lower bound on the number of triangles in a graph with n vertices and m edges, which follows from a nice application of the Cauchy-Schwarz inequality.

Theorem 2.3 (Goodman [6]; see also Nordhaus and Stewart [17]). *If G is a graph with n vertices and m edges, then*

$$t(G) \geq \frac{4m}{3n}\left(m - \frac{n^2}{4}\right).$$

Note that this implies Mantel's theorem [15]—the special case of Turán's theorem which asserts that an n-vertex graph with more than $n^2/4$ edges contains a triangle. Goodman's theorem was improved by Bollobás [1], and more recently Razborov [18] proved a theorem establishing the precise dependence between the edge density and triangle density. (This result was then extended by Nikiforov [16] to K_4 subgraphs and by Reiher [20] for cliques of arbitrary size.) However, for the purpose of our application, using Razborov's theorem instead of Goodman's provides no improvement.

Proof of Theorem 2.1. Assume (for a contradiction) that G is an edge-minimal graph which violates one of the conditions. Then the size of each color class is k, where $k = \lceil n/3 \rceil$ in the first case and $k = \lceil 2n/5 \rceil$ in the second. Now, combining Lemma 2.2 and Theorem 2.3, we obtain the inequality $3n(k-1) \geq 8m - 2n^2$. As $m \geq (9/8)nk$ in the first case and $m \geq nk$ in the second, we get a contradiction in both cases. □

3 | BOUNDED-SIZE COLOR CLASSES

For positive integers n, k define $g(n, k)$ to be the largest integer m so that there exists an edge-colored graph with n vertices and m edges, so that all color classes have size at most k and there is no rainbow triangle. If we knew that $g(n, \lceil n/3 \rceil) < n\lceil n/3 \rceil$, then the $r = 3$ case of Conjecture

1.2 would follow. Thus, determining $g(n, k)$ for all n and k appears at least as difficult as the $r = 3$ case of the Cacceta-Häggkvist conjecture. But it is possible to calculate $g(n, k)$ for n suitably larger than k. This is the main content of this paper.

We start by introducing a family of edge-colored graphs that, as we shall later prove, achieve the optimal bound for $g(n, k)$ whenever n is sufficiently large. To define this family, let k, a, b be positive integers. Consider a complete bipartite graph $K_{a,b}$ with bipartition (A, B) where $|A| = a$ and $|B| = b$. Form a new graph $H^k_{a,b}$ from this graph by adding $\lfloor b/k \rfloor$ disjoint copies of K_k on the set B, and then if $r = b - k\lfloor b/k \rfloor > 0$ we add another clique of size r to B, disjoint from the rest. Next we equip this graph with an edge coloring. Let $B_1, \ldots, B_{\lceil b/k \rceil}$ be the cliques which were added to B, and for every $v \in A$ and $1 \le i \le \lceil b/k \rceil$ declare the set of edges between v and B_i to be a color class. Then, for each i, order the vertices of B_i as $u_1, \ldots, u_{k'}$ (where $k' = |B_i|$), and for every $1 \le j \le k' - 1$ declare the set of edges between u_j and $\{u_{j+1}, \ldots, u_{k'}\}$ to be a color class. It is straightforward to verify that all color classes have size at most k, and there is no rainbow triangle.

Remark: The color classes which contain edges with both ends in B have sizes ranging from 1 up to $k - 1$. If one wishes to obtain a coloring where all color classes have size equal to k, this can easily be arranged in the case when $2k$ divides b by merging color classes from this construction.

When $k = 1$ all subgraphs are rainbow, so the Turán graph $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ (which is also isomorphic to $H^1_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$) has the maximum number of edges without a rainbow triangle. Thus

$$g(n, 1) = \left\lfloor \frac{n}{2} \right\rfloor \left\lceil \frac{n}{2} \right\rceil.$$

Next consider the case $k = 2$.

Theorem 3.1. *The values of $g(n, 2)$ are given by*

(1) $g(5, 2) = 8$,
(2) $g(n, 2) = \lfloor n/2 \rfloor \lceil n/2 \rceil + \lfloor (1/2) \lceil n/2 \rceil \rfloor$ *for* $n \neq 5$.

Proof. To see that $g(5, 2) \ge 8$ let G be the 5-wheel with center z and cycle vertices v_1, v_2, v_3, v_4, and let the color classes be $C_i = \{v_i v_{i+1}, v_i z\}$ $(1 \le i \le 4)$, where $v_5 := v_1$. To see that $g(n, 2) \ge \lfloor n/2 \rfloor \lceil n/2 \rceil + \lfloor (1/2) \lceil n/2 \rceil \rfloor$, notice that the graph $H^2_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ has n vertices and $\lfloor n/2 \rfloor \lceil n/2 \rceil + \lfloor (1/2) \lceil n/2 \rceil \rfloor$ edges, setting this as a lower bound on $g(n, 2)$.

For an upper bound, we shall use the following stability version of Turán's theorem.

Theorem 3.2 (Lovász and Simonovits [14]). *If G is a graph with n vertices and $\lfloor n/2 \rfloor \lceil n/2 \rceil + s$ edges, where $s < \lfloor n/2 \rfloor$, then $t(G) \ge s \lfloor n/2 \rfloor$.*

The upper bound can easily be checked for $n \le 5$. We henceforth assume that G is a graph with $n \ge 6$ vertices and m edges, with edges colored so that the color classes are of size at most 2, and G has no rainbow triangles. Let $s = m - \lfloor n/2 \rfloor \lceil n/2 \rceil$. We have to show that $s \le (1/2) \lceil n/2 \rceil$.

Assume first that $s < \lfloor n/2 \rfloor$. In this case, combining Lemma 2.2 (for $k = 2$) with Theorem 3.2 gives

$$s\left\lfloor \frac{n}{2} \right\rfloor \leq \frac{m}{2}.$$

Replacing m by $\lfloor n/2 \rfloor \lceil n/2 \rceil + s$ and multiplying by $2/\lfloor n/2 \rfloor$ yields

$$2s \leq \left\lceil \frac{n}{2} \right\rceil + \frac{s}{\lfloor n/2 \rfloor}.$$

Since $s/\lfloor n/2 \rfloor$ is less than 1, it may be removed from the right-hand side, and we obtain $s \leq (1/2)\lceil n/2 \rceil$ as desired.

If $s \geq \lfloor n/2 \rfloor$, we can remove edges from the graph until the number of edges is $\lfloor n/2 \rfloor \lceil n/2 \rceil + \lfloor n/2 \rfloor - 1$. By the above argument, the number of edges in the new graph does not exceed $\lfloor n/2 \rfloor \lceil n/2 \rceil + (1/2)\lceil n/2 \rceil$. For all values of $n \geq 6$ except $n = 7$, this is a contradiction because $\lfloor n/2 \rfloor - 1 > (1/2)\lceil n/2 \rceil$. It remains to rule out the case when G has 7 vertices and 15 edges. In this case, the average degree in G is less than 5. By removing a vertex of degree at most 4, we are left with a graph on 6 vertices having at least 11 edges, contradicting the upper bound for $n = 6$. □

For $k > 2$ we do not know the value of $g(n, k)$ for general n. But the following theorem gives the answer for large enough n.

Theorem 3.3. *Let $n \geq 36k^{10}(k+1)^4$. Then there exists a so that $H^k_{a,n-a}$ has the maximum number of edges over all n-vertex edge-colored graphs for which all color classes have size at most k and there is no rainbow triangle. Moreover, any graph attaining this maximum is isomorphic to $H^k_{a,n-a}$ for some a.*

This theorem is proved in the next section. Roughly speaking, it asserts that

$$g(n, k) \approx \frac{n^2}{4} + \frac{(k-1)n}{4},$$

when n is large. To determine the exact value of $g(n, k)$ for all large enough n, we need to find the optimal choice of a within the family of graphs $H^k_{a,n-a}$. This is a routine maximization problem, whose solution is stated without proof in the following proposition. It turns out that the answer depends on the remainder n leaves upon division by $2k$.

Proposition 3.4. *Let $n = t(2k) + r$, where $1 \leq r \leq 2k$.*

(1) *If $1 \leq r \leq k$ then the maximum of $|E(H^k_{a,b})|$ under the constraint $a + b = n$ is attained (solely) in the following two cases: $a = tk$, $b = tk + r$, and $a = tk + 1$, $b = tk + r - 1$. The value of the maximum is*

$$\frac{n^2}{4} + \frac{(k-1)n}{4} - \frac{r(k+1-r)}{4}.$$

(2) *If $k + 1 \leq r \leq 2k$ then the maximum of $|E(H^k_{a,b})|$ under the constraint $a + b = n$ is attained (solely) when $a = (t-1)k + r$, $b = (t+1)k$. The value of the maximum is*

$$\frac{n^2}{4} + \frac{(k-1)n}{4} + \frac{(r-k-1)(2k-r)}{4}.$$

Theorem 3.3 shows that if n is large enough with respect to k, then $g(n, k)$ is attained at some graph $H^k_{a,b}$. This is not true for all values of n. We have already noticed it in the case $n = 5,\ k = 2$ (see Theorem 3.1), where an example was given demonstrating $g(5, 2) \geq 8$, while the optimal value of $|E(H^2_{a,b})|$ is 7. This example can be generalized, as follows:

Example 3.5. Let $k \geq 2$ and $n = 3k - 1$. Take a cycle $v_1 v_2 ... v_{3k-2}$, add a special vertex z, and connect every v_i to all vertices v_j at distance at most $k - 1$ from it on the cycle, and to z. The $3k - 2$ color classes are the sets $C_i = \{v_i v_{i+j} | 1 \leq j \leq k - 1\} \cup \{v_i z\}$ for all $1 \leq i \leq 3k - 2$, where counting is modulo $3k - 2$.

It is easy to check that there are no rainbow triangles. The number of edges is $(3k - 2)k = 3k^2 - 2k$, while by Proposition 3.4 (applied with $t = 1,\ r = k - 1$) the optimal value of $|E(H^k_{a,b})|$ is $3k^2 - 3k + 1$.

4 | PROOF OF THEOREM 3.3

Here is a pleasing fact about complete edge-colored graphs:

Lemma 4.1. *Every edge-colored K_n without rainbow triangles has a color class of size at least $n - 1$.*

The lemma is an easy corollary of the following pretty theorem of Gallai.

Theorem 4.2 (Gallai [5]; Gyárfás and Simonyi [8]). *If G is an edge-colored complete graph on at least two vertices without a rainbow triangle, there is a nontrivial partition $\mathscr{P}$ of $V(G)$ satisfying:*

(1) *If $P, Q \in \mathscr{P}$ satisfy $P \neq Q$, then all edges with one end in P and the other in Q have the same color.*
(2) *The set of edges with ends in distinct blocks of $\mathscr{P}$ has at most two colors.*

Proof of Lemma 4.1. We may assume $n \geq 2$ and apply the above theorem to choose a partition $\mathscr{P}$. If this is a bipartition, then the edge-cut it defines is monochromatic and has size at least $n - 1$. If $|\mathscr{P}| = 3$, then of the three cuts it defines two must be of the same color, meaning that the graph has a nontrivial monochromatic edge-cut (which must have size $\geq n - 1$). Otherwise $|\mathscr{P}| \geq 4$ and the number of edges between blocks of $\mathscr{P}$ must be at least $2(n - 1)$ which again forces a color class to have size $\geq n - 1$ (to see this last claim, note that the number of edges between blocks is one-half of the sum over all blocks $P \in \mathscr{P}$ of the number of edges with exactly one end in P). □

For a graph G, let $\delta(G)$ denote the minimum degree in G. If X is a set of vertices of G, let $G[X]$ denote the graph induced on X, let $E(X)$ denote the set of edges of G with both ends in X,

and let $e(X) = |E(X)|$. Also, for $x \in X$ let $d_X(x)$ denote the degree of x in $G[X]$. If Y is a set of vertices disjoint from X, then let $E(X, Y)$ denote the set of edges with one end in X and the other in Y, and let $e(X, Y) = |E(X, Y)|$. Also, $\bar{e}(X, Y) = |X|\cdot|Y| - e(X, Y)$.

The main ingredient in the proof of Theorem 3.3 is as follows:

Lemma 4.3. *Let $G = (V, E)$ be an edge-colored graph with the property that every color class has size at most k and there is no rainbow triangle. If $n = |V|$ satisfies $n \geq n_1(k) := 6k^5(k+1)^2$ and $m = |E|$ satisfies*

$$m \geq \frac{n^2}{4} + \frac{(k-1)n}{4} - \frac{k(k-1)}{2}, \tag{1}$$

and

$$\delta(G) \geq \left\lfloor \frac{n}{2} \right\rfloor, \tag{2}$$

then there exists a partition $\{X, Y\}$ of V so that X is an independent set and every component of $G[Y]$ has size at most k.

Proof. For $k = 1$ the lemma follows from Turán's theorem, so we may assume that $k \geq 2$. Let $\mathscr{T}$ be the set of triangles of G. Now Lemma 2.2 together with a straightforward application of Cauchy-Schwarz yields

$$\begin{aligned}\sum_{uv \in E} |N(u) \cup N(v)| &= \sum_{uv \in E} (d(u) + d(v) - |\{T \in \mathscr{T} | uv \in E(T)\}|) \\ &= \sum_{v \in V} (d(v))^2 - 3|\mathscr{T}| \\ &\geq \frac{4m^2}{n} - \frac{3}{2}m(k-1).\end{aligned}$$

Thus there exists an edge x_0y_0 such that $|N(x_0) \cup N(y_0)| \geq (4m/n) - (3/2)(k-1) \geq n - (2/3)(k-1)$, where for the second inequality we have used (1) and $n \geq n_1(k)$. Choose disjoint sets $X_0 \subseteq N(x_0)$ and $Y_0 \subseteq N(y_0)$ so that $X_0 \cup Y_0 = N(x_0) \cup N(y_0)$. Next we establish a claim which gives upper bounds on $e(X_0)$ and $e(Y_0)$, that we will use to get a lower bound on $e(X_0, Y_0)$. This, in turn, will give a lower bound on the size of a maximum cut in G, which is the sole purpose of this part of the proof; the sets X_0, Y_0 themselves will play no further role.

Claim. If $W \subseteq N(w)$ for some $w \in V$, then $e(W) \leq (k-1)|W|$.

Proof. Let I be the set of colors appearing on the edges $E(\{w\}, W)$ and for every $i \in I$ let d_i be the number of edges in $E(\{w\}, W)$ of color i. Let F be the set of edges $uv \in E(W)$ for which wu and wv have the same color. Then $|F| \leq \sum_{i \in I} \binom{d_i}{2}$. Since there are no rainbow

triangles, every edge $uv \in E(W)\backslash F$ is colored by some $i \in I$. Since for every $i \in I$ the number of edges in $E(W)$ colored i is at most $k - d_i$, we have

$$\begin{aligned} |E(W)| &\leq \sum_{i\in I} \left(\binom{d_i}{2} + (k - d_i)\right) \\ &= \sum_{i\in I} \left((k-1)d_i - (d_i - 1)\left(k - \tfrac{1}{2}d_i\right)\right) \\ &\leq (k-1)\sum_{i\in I} d_i \\ &= (k-1)|W|, \end{aligned}$$

establishing the claim.

Define $Z = V\backslash(X_0 \cup Y_0)$. Applying the claim to each of $E(X_0)$ and $E(Y_0)$, we have

$$\begin{aligned} m &= |E(X_0)| + |E(Y_0)| + |E(X_0, Y_0)| + |E(Z)| + |E(Z, X_0 \cup Y_0)| \\ &\leq (k-1)(|X_0| + |Y_0|) + e(X_0, Y_0) + \sum_{z\in Z} d(z). \end{aligned}$$

Since $|X_0| + |Y_0| \leq n$ and $|Z| \leq (2/3)(k-1)$, this implies

$$m \leq e(X_0, Y_0) + \frac{5}{3}(k-1)n. \tag{3}$$

Let $E(X, Y)$ be a maximum cut, meaning that $\{X, Y\}$ is a partition of V for which $e(X, Y)$ is maximum. Let a_X be the average degree of $G[X]$, let a_Y be the average degree of $G[Y]$ and without loss of generality assume that $a_X \leq a_Y$. We proceed to prove a series of properties of X and Y, eventually showing that they satisfy the conclusions of the lemma.

(1) $e(X, Y) \geq (n^2/4) - (3/2)(k-1)n$ and $\bar{e}(X, Y) \leq (3/2)(k-1)n$. The first part follows from Inequality (3), the fact that $e(X, Y) \geq e(X_0, Y_0)$ (since (X, Y) is a maximum cut), Inequality (1), and the assumption that $n \geq n_1(k)$. The second part follows from the first since

$$\bar{e}(X, Y) = |X|\cdot|Y| - e(X, Y) \leq \frac{n^2}{4} - e(X, Y) \leq \frac{3}{2}(k-1)n.$$

(2) Every $x \in X$ satisfies $e(x, Y) \geq (1/2)\lfloor n/2 \rfloor$ and every $y \in Y$ satisfies $e(y, X) \geq (1/2)\lfloor n/2 \rfloor$. This follows from the assumption that the minimum degree is at least $\lfloor n/2 \rfloor$, and the fact that a maximum cut is "unfriendly," namely, every vertex is adjacent to at least as many vertices on the other side as on its own side.

(3) Both graphs $G[X]$ and $G[Y]$ have maximum degree less than $7k$. Suppose for a contradiction that this is false, and assume (without loss of generality) that $x \in X$ has $d_X(x) \geq 7k$. Choose a set $X' \subseteq N(x) \cap X$ with $|X'| = 7k$ and let I be the set of colors appearing on the edges $E(x, X')$. Let $Y' = N(x) \cap Y$ and note that by (2) we must

have $|Y'| \geq (1/2)\lfloor n/2 \rfloor$. Fix $y' \in Y'$. For each $x' \in N(y') \cap X'$, we have a triangle on x, x', y'. The majority color in that triangle can be the color of xy' for at most $k-1$ choices of x'. For all other choices of x', the color of $x'y'$ must be in I. Hence $|N(y') \cap X'| \leq k - 1 + |\{e \in E(y', X') | e \text{ has a color in } I\}|$. It follows that

$$\bar{e}(y', X') \geq 6k + 1 - |\{e \in E(y', X') | e \text{ has a color in } I\}|.$$

The total number of edges in G with a color in I is at most $7k^2$. Thus summing the above inequalities over all $y' \in Y'$ yields $\bar{e}(Y', X') \geq (6k+1)(1/2)\lfloor n/2 \rfloor - 7k^2$ which exceeds $(3/2)(k-1)n$ for $n \geq n_1(k)$, contradicting (1).

(4) $\lfloor n/2 \rfloor - 7k < |X|, |Y| < \lceil n/2 \rceil + 7k$. To show, for example, that $|X| > \lfloor n/2 \rfloor - 7k$, take any vertex $y \in Y$. By assumption (2), its degree in G is at least $\lfloor n/2 \rfloor$, and by the previous fact its degree in $G[Y]$ is less than $7k$. Hence $e(y, X) > \lfloor n/2 \rfloor - 7k$, implying in particular $|X| > \lfloor n/2 \rfloor - 7k$. Combining the last inequalities with (2) yields:

(5) Every $x \in X$ satisfies $\bar{e}(x, Y) < 14k$ and every $y \in Y$ satisfies $\bar{e}(y, X) < 14k$.

(6) Every component of $G[X]$ and $G[Y]$ has size at most k; in particular, $G[X]$ and $G[Y]$ have maximum degree less than k. Suppose (for a contradiction) that there is a component of $G[X]$ with size greater than k and choose a tree T in $G[X]$ with exactly $k+1$ vertices. It follows from (5) that the number of common neighbors of $V(T)$ in the set Y is at least $|Y| - 14k(k+1) \geq (n/2) - 14(k+1)^2 > k^2$, using $n \geq n_1(k)$. Since there are at most k^2 edges which have the same color as an edge of T, it follows that there is at least one vertex $y \in Y$ so that y is adjacent to every vertex in $V(T)$ with the added property that no edge in $E(y, V(T))$ shares a color with an edge in $E(T)$. However, this gives a contradiction: it is not possible for all edges in $E(y, V(T))$ to have the same color (since this set has size $k+1$), but otherwise there is a rainbow triangle containing y and an edge of T.

(7) $a_X + a_Y \geq k - 1 - (3k^2/n)$. Note that $m \leq (1/2)a_X|X| + (1/2)a_Y(n - |X|) + |X|(n - |X|)$. The right-hand side of this inequality is a quadratic function of $|X|$ for which the maximum is attained at $|X| = (n/2) + ((a_X - a_Y)/4)$. This maximum value is equal to

$$\frac{n^2}{4} + \frac{(a_X + a_Y)n}{4} + \frac{(a_X - a_Y)^2}{16}.$$

It follows from (6) that $|a_X - a_Y| \leq k - 1$ and this gives us the following inequality which completes the argument:

$$\frac{n^2}{4} + \frac{(k-1)n}{4} - \frac{k(k-1)}{2} \leq m \leq \frac{n^2}{4} + \frac{(a_X + a_Y)n}{4} + \frac{(k-1)^2}{16}.$$

(8) $a_X \leq 1/(k+1)$. (Remember that we assumed $a_X \leq a_Y$ and our aim is to show that $a_X = 0$.) Suppose (for a contradiction) that $a_X > 1/(k+1)$. Define $\tilde{a}_X = \lceil a_X - (1/(k+1)) \rceil$ and $\tilde{a}_Y = \lceil a_Y - (1/(k+1)) \rceil$ and note that our assumptions imply $\tilde{a}_X, \tilde{a}_Y \geq 1$. Let $X^+ = \{x \in X | d_X(x) \geq \tilde{a}_X\}$ and $X^- = X \backslash X^+$. Then we have

$$a_X|X| = \sum_{x\in X} d_X(x) \le \left(a_X - \frac{1}{k+1}\right)|X^-| + k|X^+| = \left(a_X - \frac{1}{k+1}\right)|X| + \left(k - a_X + \frac{1}{k+1}\right)|X^+|.$$

It follows that $|X^+| \ge |X|/(k(k+1))$. Similarly, setting $Y^+ = \{y \in Y | d_Y(y) \ge \tilde{a}_Y\}$ we find that $|Y^+| \ge |Y|/(k(k+1))$. It follows from (7) that

$$\tilde{a}_X + \tilde{a}_Y \ge a_X + a_Y - \frac{2}{k+1} \ge k - 1 - \frac{3k^2}{n} - \frac{2}{k+1} > k - 2,$$

again using $n \ge n_1(k)$. So, $\tilde{a}_X$ and $\tilde{a}_Y$ are positive integers which sum to at least $k-1$. For every component of $G[X]$ of size at least $\tilde{a}_X + 1$ choose a spanning tree, and let $\mathscr{T}_X$ be the set of these spanning trees. Similarly let $\mathscr{T}_Y$ be a collection consisting of one spanning tree from each component of $G[Y]$ of size at least $\tilde{a}_Y + 1$. We claim that for every $T \in \mathscr{T}_X$ and $U \in \mathscr{T}_Y$ one of the following holds:

(a) $\bar{e}(V(T), V(U)) \ge 1$.

(b) $E(V(T), V(U))$ contains an edge with the same color as an edge of T or U.

Were neither condition above to be satisfied, all edges between $V(T)$ and $V(U)$ would need to have the same color to prevent a rainbow triangle, but then this color would appear too many times. Indeed, $e(V(T), V(U)) \ge (\tilde{a}_X + 1)(\tilde{a}_Y + 1)$ and this product exceeds k, because its two factors are greater than 1 and sum to at least $k+1$. For every $T \in \mathscr{T}_X \cup \mathscr{T}_Y$, the total number of edges in G with the same color as some color in $E(T)$ is less than k^2. Thus, by considering all pairs of trees from $\mathscr{T}_X \times \mathscr{T}_Y$ we get the following bound:

$$\bar{e}(X, Y) \ge |\mathscr{T}_X|\cdot|\mathscr{T}_Y| - k^2|\mathscr{T}_X| - k^2|\mathscr{T}_Y|.$$

Every vertex $x \in X^+$ has $d_X(x) \ge \tilde{a}_X$ and must therefore be in a component of $G[X]$ of size at least $\tilde{a}_X + 1$. It follows that $|\mathscr{T}_X| \ge (|X^+|/k) \ge |X|/(k^2(k+1))$. A similar argument for Y shows that $|\mathscr{T}_Y| \ge |Y|/(k^2(k+1))$. Since (using $n \ge n_1(k)$) both of these quantities are larger than k^2 we can plug these bounds for $|\mathscr{T}_X|$ and $|\mathscr{T}_Y|$ in the previous equation. Doing so gives the following (note that we have used (4) to get the bound $|X|\cdot|Y| \ge (n^2/4) - 49k^2$ and, again, $n \ge n_1(k)$):

$$\begin{aligned}
\bar{e}(X, Y) &\ge \frac{|X|\cdot|Y|}{k^4(k+1)^2} - \frac{|X|+|Y|}{k+1} \\
&\ge \frac{n^2}{4k^4(k+1)^2} - \frac{49}{k^2(k+1)^2} - \frac{n}{k+1} \\
&\ge \left(\frac{3}{2}k - \frac{49}{k^2(k+1)^2 n} - \frac{1}{k+1}\right)n \\
&> \frac{3}{2}(k-1)n.
\end{aligned}$$

This contradicts (1), and completes the proof of (8).

With this last property in hand, we have all we need to complete the proof. It follows from (8) and (7) and $n \ge n_1(k)$ that

$$a_Y \geq k - 1 - \frac{3k^2}{n} - \frac{1}{k+1} \geq k - 1 - \frac{1}{k}.$$

Note that the graph obtained from K_k by removing one edge has average degree $k - 1 - (2/k)$; more generally any graph on at most k vertices which is not isomorphic to K_k has average degree at most $k - 1 - (2/k)$. Setting t to be the number of components of $G[Y]$ isomorphic to K_k we then have

$$\begin{aligned}\left(k - 1 - \tfrac{1}{k}\right)|Y| &\leq a_Y|Y| \\ &\leq tk(k-1) + (|Y| - tk)\left(k - 1 - \tfrac{2}{k}\right) \\ &= |Y|(k-1) - \tfrac{2}{k}(|Y| - tk).\end{aligned}$$

It follows from this that $t \geq |Y|/(2k)$. Let s be the number of nontrivial components of $G[X]$ (ie, components with at least two vertices). By Lemma 4.1 there is at least one missing edge between every K_k subgraph of $G[Y]$ and every nontrivial component of $G[X]$. Since each such component contains at most $k(k-1)/2$ edges, we have the bound

$$m \leq |Y|(n - |Y|) + \frac{k-1}{2}|Y| - s\left(t - \frac{k(k-1)}{2}\right).$$

Now $|Y|(n - |Y|) + ((k-1)/2)|Y|$ is a quadratic function of $|Y|$ maximized at $|Y| = (n/2) + ((k-1)/4)$ with maximum value $(n^2/4) + ((k-1)n/4) + ((k-1)^2/16)$, so

$$m \leq \frac{n^2}{4} + \frac{(k-1)n}{4} + \frac{(k-1)^2}{16} - s\left(t - \frac{k(k-1)}{2}\right).$$

Using our initial lower bound on m, and the facts that $t \geq |Y|/(2k)$, $|Y| > \lfloor n/2 \rfloor - 7k$, and $n \geq n_1(k)$, shows that $s > 0$ is impossible. Therefore, $s = 0$ and we have that X is an independent set. We established that all components of $G[Y]$ have size at most k in (6) so the proof is now complete. □

An easy calculation yields that removing a vertex of degree smaller than $\lfloor n/2 \rfloor$ retains Inequality (1), namely:

Lemma 4.4. *Let $G = (V, E)$ be a graph with n vertices and m edges, where*

$$m \geq \frac{n^2}{4} + \frac{(k-1)n}{4} - \frac{k(k-1)}{2}.$$

Let v be a vertex with degree smaller than $\lfloor n/2 \rfloor$, let $m' = |E(G - v)|$ and let $n' = |V(G - v)| = n - 1$. Then

$$m' \geq \frac{n'^2}{4} + \frac{(k-1)n'}{4} - \frac{k(k-1)}{2}.$$

The next lemma shows that given Inequality (1), the process of removing vertices of degree smaller than half the current number of vertices cannot continue for too long.

Lemma 4.5. *Let $G = (V, E)$ be a graph with n vertices and m edges, where $n \geq 2k$ and*

$$m \geq \frac{n^2}{4} + \frac{(k-1)n}{4} - \frac{k(k-1)}{2}. \tag{4}$$

Then there exists a subset U of $V(G)$ of size $n' \geq \sqrt{n}$, such that the graph $G' = G[U]$ has minimal degree at least $\lfloor n'/2 \rfloor$, and its number of edges m' satisfies

$$m' \geq \frac{n'^2}{4} + \frac{(k-1)n'}{4} - \frac{k(k-1)}{2}. \tag{5}$$

Moreover, U can be obtained by removing sequentially vertices of degree smaller than half the current number of vertices.

Proof. If there is no vertex of degree smaller than $\lfloor n/2 \rfloor$, there is nothing to prove. Suppose that G has a vertex, denote it x_n, with degree smaller than $\lfloor n/2 \rfloor$. Let $G_{n-1} = G - v$. By the previous lemma, G_{n-1} satisfies (5). If all vertices in G_{n-1} have degree at least $\lfloor (n-1)/2 \rfloor$, we are done. If not, remove a vertex x_{n-1} of degree smaller than $\lfloor (n-1)/2 \rfloor$. By the previous lemma the graph $G - x_n - x_{n-1}$ satisfies (5). We continue this way, until, having deleted vertices $x_{n'+1}, x_{n'+2}, \ldots, x_n$, we obtain a graph $G_{n'}$ all of whose vertices have degrees at least $\lfloor n'/2 \rfloor$. The proof of the lemma will be complete if we show that $n' \geq \sqrt{n}$.

Let $U = \{x_1, \ldots, x_{n'}\}$ be the remaining set of vertices. There are at most $\binom{n'}{2}$ edges of G contained in U, and each vertex x_i, $i > n'$, is adjacent in G to at most $(i-1)/2$ vertices x_j, $j < i$. Hence

$$m = |E(G)| \leq \binom{n'}{2} + \sum_{i=n'+1}^{n} \frac{i-1}{2} = \binom{n'}{2} + \frac{1}{2}\left(\binom{n}{2} - \binom{n'}{2}\right) = \frac{1}{2}\left(\binom{n}{2} + \binom{n'}{2}\right).$$

This, combined with (4) and the fact that $n \geq 2k$, yields $n' \geq \sqrt{n}$, completing the proof. □

Proof of Theorem 3.3. Let $k \geq 2$ and let $n \geq 36k^{10}(k+1)^4$. Let G be an edge-colored graph on n vertices, with color classes of size at most k and no rainbow triangle, so that its number of edges m is the largest among all such graphs. We shall show that there exists a graph $H^k_{a,b}$ with $a + b = n$ and $|E(H^k_{a,b})| \geq m$.

Clearly

$$m = |E(G)| \geq |E(H^k_{\lfloor n/2 \rfloor, \lceil n/2 \rceil})| \geq \frac{n^2}{4} + \frac{(k-1)n}{4} - \frac{k(k-1)}{2}.$$

Hence, by Lemma 4.5, there exists a subset U of $V = V(G)$ of size $n' \geq \sqrt{n} \geq 6k^5(k+1)^2$, obtained by deleting sequentially vertices of degree less than half the current number of vertices, such that all degrees in $G[U]$ are at least $\lfloor n'/2 \rfloor$ and (5) is satisfied. By Lemma 4.3 applied to $G[U]$, there exists a partition $\{X', Y'\}$ of U so that X' is an independent set, and all components of $G[Y']$ have size at most k.

Starting from $G[U]$ and the partition $\{X', Y'\}$, we show how to construct a graph isomorphic to $H^k_{a,b}$, $a + b = n$, having at least m edges. First, we put vertices $z_{n'+1}, z_{n'+2}, \ldots, z_n$ back into the graph, in this order, instead of the vertices $x_{n'+1}, x_{n'+2}, \ldots, x_n$ that were deleted. At each step we add z_i to the smaller side of the partition that we have at hand at that step (or to any side, if the two sides are equal), and we connect it to all vertices on the other side. By the property of x_i, the addition of z_i adds more edges than the number of edges deleted when x_i was removed. At the end of this process, we get a graph $\hat{G}$ with n vertices and at least m edges, together with a partition $\{X, Y\}$ of $V(\hat{G})$ satisfying the conclusion of Lemma 4.3. Note that at this stage we do not care about edge-colors, we just aim to show that the number of edges of $\hat{G}$ is at most that of $H^k_{|X|,|Y|}$.

If there is a component H of $\hat{G}[Y]$ which is not a complete graph, modify $\hat{G}$ by adding all edges with both ends in $V(H)$ that are missing. If there are two components of $\hat{G}[Y]$, say H_1 and H_2, both of size less than k, say $k > |V(H_1)| \geq |V(H_2)|$, increase $|E(\hat{G})|$ by deleting the edges $E(v, V(H_2) \setminus \{v\})$ for some $v \in V(H_2)$, and then adding the edges $E(v, V(H_1))$. Finally, add to $\hat{G}$ any missing edge between X and Y. Our modified $\hat{G}$ is now isomorphic to $H^k_{|X|,|Y|}$, and has at least m edges. This proves the first part of the theorem. For the "moreover" part, observe that if $n' < n$ or if $\hat{G}$ is modified at all, we end up with strictly more than m edges, contradicting the maximality of m. □

ACKNOWLEDGMENTS

The research of the first author was supported by ISF grant no. 2023464, BSF grant no. 2006099, and by the Discount Bank Chair at the Technion. The research of the third author was supported by ISF grant no. 409/16.

ORCID

Ron Aharoni http://orcid.org/0000-0002-1768-1901

REFERENCES

1. B. Bollobás, Relations between sets of complete subgraphs, Proc. Fifth British Combin. Conf. Aberdeen, 1975, pp. 79–84.
2. J. A. Bondy, *Counting subgraphs: a new approach to the Caccetta-Häggkvist conjecture*, Discrete Math. **165/166** (1997), 71–80.
3. L. Caccetta, R. Häggkvist, *On minimal digraphs with given girth*, Proc. Ninth Southeastern Conf. on Combinatorics, Graph Theory, and Computing, Boca Raton, FLA, 1978, pp. 181–187.
4. P. Erdős, M. Simonovits, and V. T. Sós, *Anti-Ramsey theorems*, Coll. Math. Soc. J. Bolyai **10** (1973), 633–643.

5. T. Gallai, *Transitiv orientierbare graphen*, Acta Math. Hungar. **18** (1967), 25–66.
6. A. W. Goodman, *On sets of acquaintances and strangers at any party*, Am. Math. Mon. **66** (1959), 778–783.
7. A. Gyárfás and G. Sárközy, *Gallai of colorings of non-complete graphs*, Discrete Math. **310** (2010), 977–980.
8. A. Gyárfás and G. Simonyi, *Edge colorings of complete graphs without tricolored triangles*, J. Graph Theory **46** (2004), 211–216.
9. P. Hamburger, P. Haxell, and A. Kostochka, *On directed triangles in digraphs*, Electron. J. Combin. **14** (2007), N19.
10. J. Hladký, D. Král', and S. Norin, *Counting flags in triangle-free digraphs*, Combinatorica **37** (2017), 49–76.
11. B. Li et al., *Rainbow triangles in edge-colored graphs*, Eur. J. Combin. **36** (2014), 453–459.
12. H. Li, *Rainbow C_3's and C_4's in edge-colored graphs*, Discrete Math. **313** (2013), 1893–1896.
13. H. Li and G. Wang, *Color degree and heterochromatic cycles in edge-colored graphs*, Eur. J. Combin. **33** (2012), 1958–1964.
14. L. Lovász and M. Simonovits, *On the number of complete subgraphs of a graph II*, Stud. Pure Math (1983), 459–495.
15. W. Mantel, *Problem 28: solution by H. Gouwentak, W. Mantel, J. Teixeira de Mattes, F. Schuh and W. A. Wythoff*, Wiskundige Opgaven **10** (1907), 60–61.
16. V. Nikiforov, *The number of cliques in graphs of given order and size*, Trans. Amer. Math. Soc. **363** (2011), 1599–1618.
17. E. A. Nordhaus and B. M. Stewart, *Triangles in an ordinary graph*, Can. J. Math. **15** (1963), 33–41.
18. A. A. Razborov, *On the minimal density of triangles in graphs*, Combin. Probab. Comput. **17** (2008), 603–618.
19. A. A. Razborov, *On the Caccetta-Häggkvist conjecture with forbidden subgraphs*, J. Graph Theory **74** (2013), 236–248.
20. C. Reiher, *The clique density theorem*, Ann. Math. **184** (2016), 683–707.
21. J. Shen, *Directed triangles in digraphs*, J. Combin. Theory Ser. B **74** (1998), 405–407.
22. J. Shen, *On the girth of digraphs*, Discrete Math. **211** (2000), 167–181.

How to cite this article: Aharoni R, DeVos M, Holzman R. Rainbow triangles and the Caccetta-Häggkvist conjecture. *J Graph Theory*. 2019;92:347–360.
https://doi.org/10.1002/jgt.22457