

Short rainbow cycles for families of matchings and triangles

He Guo*

November 20, 2022; revised March 31, 2024

Abstract

A generalization of the famous Caccetta–Haggkvist conjecture, suggested by Aharoni [2], is that any family $\mathcal{F} = (F_1, \dots, F_n)$ of sets of edges in K_n , each of size k , has a rainbow cycle of length at most $\lceil \frac{n}{k} \rceil$. In [3, 1] it was shown that asymptotically this can be improved to $O(\log n)$ if all sets are matchings of size 2, or all are triangles. We show that the same is true in the mixed case, i.e., if each F_i is either a matching of size 2 or a triangle. We also study the case that each F_i is a matching of size 2 or a single edge, or each F_i is a triangle or a single edge, and in each of these cases we determine the threshold proportion between the types, beyond which the rainbow girth goes from linear to logarithmic.

Keywords: rainbow girth, generalized Caccetta–Haggkvist conjecture, short rainbow cycles

1 Introduction

The *directed girth* $dgirth(G)$ of a digraph G is the minimal length of a directed cycle in it (∞ if there is no directed cycle). A famous conjecture of Caccetta and Haggkvist [5] (below — CHC) states that any digraph G on n vertices satisfies $dgirth(G) \leq \lceil \frac{n}{\delta^+(G)} \rceil$, where $\delta^+(G)$ is the minimum out-degree of G . See [6, 8, 9, 11] for progress on this problem. In particular it has been shown that

- (a) The CHC is true if $n \geq 2\delta^+(G)^2 - 3\delta^+(G) + 1$ [12], and
- (b) $dgirth(G) \leq n/\delta^+(G) + 73$ for all G [13].

In [2] a possible generalization of CHC was suggested. Given a family $\mathcal{F} = (F_1, \dots, F_m)$ of sets of edges, a set F of edges is said to be *rainbow* for \mathcal{F} if each of the edges in F is taken from a distinct F_i (if the sets F_i are disjoint, this means that $|F \cap F_i| \leq 1$ for each i). The *rainbow girth* $rgirth(\mathcal{F})$ of \mathcal{F} is the minimal length of a rainbow cycle with respect to \mathcal{F} .

Conjecture 1. *For any family $\mathcal{F} = (F_1, \dots, F_n)$ of subsets of $E(K_n)$ such that $|F_i| = k$ for each $1 \leq i \leq n$, we have $rgirth(\mathcal{F}) \leq \lceil n/k \rceil$.*

We may clearly assume that the sets F_i are disjoint, since otherwise there is a rainbow digon, meaning that the rainbow girth is 2. Given $\mathcal{F} = (F_1, \dots, F_m)$ with $F_i \neq \emptyset$ for each $i \in [m]$ and $\cup_{i=1}^m F_i = E(G)$ for some graph G , we shall refer to \mathcal{F} as an *edge coloring* of G , the indices $i \in [m]$ as *colors*, the sets F_i as *color classes*, and $rgirth(\mathcal{F})$ as the *rainbow girth* of G with respect to the edge coloring \mathcal{F} .

*Faculty of Mathematics, Technion, Haifa 32000, Israel. E-mail: hguo@campus.technion.ac.il.

Devos et. al. [7] proved Conjecture 1 for $k = 2$. In [1] a stronger version of the conjecture was proved, in which the sets F_i are of size 1 or 2. In [10] it was shown that the order of magnitude is right: there exists a constant $C > 0$ such that for any k, n and \mathcal{F} satisfying the assumption of Conjecture 1, we have $rgirth(\mathcal{F}) \leq Cn/k$. An explanation why Conjecture 1 implies the CHC can be found in [3] and [1].

All known extreme examples for Conjecture 1 are obtained from those of the CHC, taking the color classes as stars. This suggests looking at the case when the sets of edges are not stars, and trying to improve the upper bound on the girth. Indeed, in [3], it is proved that if an n -vertex graph is edge-colored by n colors such that each color class is a matching of size 2, then the rainbow girth is $O(\log n)$, asymptotically improving the conclusion of the conjecture.

A set of edges not containing a matching of size 2 is either a star or a triangle, hence the next interesting case is that of families of triangles. In [1], it is proved that a family of n triangles in K_n has rainbow girth $O(\log n)$. Furthermore, it was shown there that $\log n$ is the right order of magnitude: an n -vertex graph is constructed, consisting of n edge-disjoint triangles whose rainbow girth is $\Omega(\log n)$.

In this note we fine-tune the above results, by showing that the rainbow girth is $O(\log n)$ in the mixed case that each F_i is either a matching of size 2 or a triangle. We also study the case that each F_i is a matching of size 2 or a single edge, or each F_i is a triangle or a single edge, and in each of these cases we determine the threshold proportion between the types, beyond which the rainbow girth goes from linear to logarithmic.

2 Graph theoretical and probabilistic tools

As in [10, 3], a key ingredient in the proofs is a result by Bollobás and Szemerédi [4] on the girth of sparse graphs.

Theorem 2. *For $N \geq 4$ and $k \geq 2$, every N -vertex graph with $N + k$ edges has girth at most*

$$\frac{2(N + k)}{3k}(\log_2 k + \log_2 \log_2 k + 4).$$

We shall use two well-known concentration inequalities.

Theorem 3 (Chernoff). *Let X be a binomial random variable $\text{Bin}(n, p)$. For any $t \geq 0$, we have*

$$\mathbb{P}(X \geq \mathbb{E}X + t) \leq \exp\left(-\frac{t^2}{2(\mathbb{E}X + t/3)}\right).$$

Theorem 4 (Chebyshev). *Let X be a random variable. For any $t > 0$, we have*

$$\mathbb{P}(|X - \mathbb{E}X| \geq t) \leq \frac{\text{Var } X}{t^2},$$

where $\text{Var } X$ is the variance of X .

3 Main results

3.1 Matchings and single edges

Theorem 5. *For any $\alpha > 1/2$, there exists C such that for any edge coloring $\mathcal{F} = (F_1, \dots, F_n)$ of an n -vertex graph G with $F_i \neq \emptyset$ for each $i \in [n]$, if at least αn color classes in \mathcal{F} are matchings of size 2, then $\text{rgirth}(\mathcal{F}) \leq C \log n$.*

We shall need a slightly stronger result, which allows the number of size-two matchings to be less than αn and the total number of color classes to be less than n .

Theorem 6. *For any $\alpha > 1/2$, there exist $\xi = \xi(\alpha) > 0$ and $C = C(\alpha)$ such that the following holds. Let G be an n -vertex graph and $\mathcal{F} = (F_1, \dots, F_m)$ be an edge coloring of G . If $\mathcal{F} = \mathcal{F}_M \sqcup \mathcal{F}_E$, where*

1. every $F_i \in \mathcal{F}_M$ is a matching of size 2,
2. every $F_i \in \mathcal{F}_E$ is a single edge,
3. $|\mathcal{F}_M| \geq (\alpha - \xi)n$ and $|\mathcal{F}_E| \geq (1 - \alpha - \xi)n$,

then $\text{rgirth}(\mathcal{F}) \leq C \log n$.

Theorem 6 will follow from Theorem 2, and the following:

Theorem 7. *For any $\alpha > 1/2$, there exist $\beta, c > 0$ such that for any large enough n , given an n -vertex graph G and an edge coloring of G satisfying the assumption in Theorem 6, there exists a subset S of $V(G)$ of size at most βn containing a rainbow edge set of size at least $(\beta + c)n$.*

Once this is proved, Theorem 6 follows by applying Theorem 2 with $N = \beta n$ and $k = cn$.

In Section 3.1.1, we give a simplified proof of Theorem 6. But we still think the proof of Theorem 6 via Theorem 7 involves some interesting techniques.

The idea used to prove Theorem 7 is choosing a random subset S of $V(G)$ and considering the induced subgraph $G[S]$. The crux of the argument is that the expected number of vertices $\mathbb{E}|S|$ is less than the expected number of rainbow edges in $G[S]$, and their difference is linear in n . Furthermore, these two random numbers are concentrated around their expectations, which follows from the concentration inequalities in Section 2.

Theorem 2.8 of [3] states that for any $\gamma > 3\sqrt{6}/8 \approx 0.9186$ and n -vertex graph with an edge coloring with at least γn colors, if each color class is a matching of size 2, then the rainbow girth is $O(\log n)$. For completeness, we add an alternative proof of this statement in the Appendix A, which was observed by Michael Krivelevich.¹ Note that in Theorem 6 we can take $\xi(\alpha) > 0$ arbitrarily small. In particular, if $\alpha > 3\sqrt{6}/8$, we can guarantee that $\alpha - \xi > 3\sqrt{6}/8$. Therefore here we may assume that $\alpha \leq 3\sqrt{6}/8$. Again since ξ can be chosen arbitrarily small, we may assume that $\min\{\alpha - \xi, 1 - \alpha - \xi\} \geq 1/40$ so that $\min\{|\mathcal{F}_M|, |\mathcal{F}_E|\} \geq n/40$. And without loss of generality, we may assume that $\max\{|\mathcal{F}_M|, |\mathcal{F}_E|\} \leq n$.

First we claim that for $p \in (0, 1)$ close enough to 1, we have

$$\alpha(2p^2 - p^4) + (1 - \alpha)p^2 > p. \tag{1}$$

¹The bound on γ is a bit weaker, but enough for our later use.

In fact, when $\alpha = 1/2 + \delta$ for some $\delta > 0$, the above is equivalent to

$$(1/2 + \delta)(2p - p^3) + (1/2 - \delta)p > 1.$$

Writing $p = 1 - \tau$, we have

$$\begin{aligned} (1/2 + \delta)(2p - p^3) + (1/2 - \delta)p &= \frac{3}{2}p + \delta p - \left(\frac{1}{2} + \delta\right)p^3 \\ &= \frac{3}{2}(1 - \tau) + \delta(1 - \tau) - \left(\frac{1}{2} + \delta\right)(1 - \tau)^3 \\ &= 1 + 2\delta\tau + \tau^2(-3/2 - 3\delta + \tau/2 + \delta\tau), \end{aligned}$$

which is greater than 1 for $\tau = \tau(\delta) > 0$ small enough.

For any $p \in (0, 1)$ close enough to 1 satisfying (1), there exist constants $\epsilon(\alpha, p), \xi(\alpha, p) > 0$ small enough so that

$$(1 - \epsilon)(\alpha - \xi)(2p^2 - p^4) + (1 - \epsilon)(1 - \alpha - \xi)p^2 > p. \quad (2)$$

Fix $1/2 \leq p < 1$ and $\epsilon, \xi > 0$ that satisfy (2).

A vertex v is called a *heavy* vertex if there are at least $(\epsilon^2/10^6)n$ many rainbow edges, i.e., edges in $(\cup \mathcal{F}_M) \cup (\cup \mathcal{F}_E)$, incident to it. Let D be the set of all heavy vertices. Then

$$|D| \leq (2 \cdot 2|\mathcal{F}_M| + 2|\mathcal{F}_E|)/((\epsilon^2/10^6)n) \leq 6n/((\epsilon^2/10^6)n) \leq 10^7/\epsilon^2. \quad (3)$$

We construct a random vertex set

$$S := D \cup (V \setminus D)_p$$

i.e., S contains the set D of heavy vertices and includes each vertex of $V \setminus D$ independently with probability p .

Lemma 8. *With high probability², $|S| \leq np + n^{2/3}$.*

Proof. Note that by construction, $|S|$ has the same probability distribution as $|D| + \text{Bin}(n - |D|, p)$. Set $Z = \text{Bin}(n - |D|, p)$. Applying Chernoff's bound (Theorem 3) and (3), we have

$$\mathbb{P}(Z \geq np + n^{2/3} - |D|) \leq \mathbb{P}(Z \geq \mathbb{E}Z + n^{2/3} - |D|) \leq \exp(-n^{\Omega(1)}).$$

Therefore with high probability, $|S| \leq |D| + Z \leq np + n^{2/3}$. □

Lemma 9. *With probability at least 0.9, the number of color classes in \mathcal{F}_M that have at least one edge contained in S is at least $(1 - \epsilon)|\mathcal{F}_M| \cdot (2p^2 - p^4)$.*

Proof. For $F_i \in \mathcal{F}_M$, let

$$X_i := \mathbb{1}_{\{\text{at least one edge in } F_i \text{ is contained in } S\}}$$

be the indicator random variable that some edge e in F_i is contained in S . Since each vertex is included in S independently with probability at least p and X_i is an increasing random variable with respect to the probability that a vertex is included in S , by inclusion-exclusion we have

$$\mathbb{E}X_i \geq 2p^2 - p^4.$$

²An event holds with high probability if the probability of that event tends to 1 as n tends to infinity.

Let

$$X := \sum_{F_i \in \mathcal{F}_M} X_i. \quad (4)$$

We have

$$\mathbb{E}X \geq |\mathcal{F}_M| \cdot (2p^2 - p^4). \quad (5)$$

To prove the lemma, we shall apply Chebyshev's inequality (Theorem 4). For this purpose we have to estimate $\text{Var } X$. With a look at (4), we have

$$\text{Var } X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \sum_{F_i, F_j \in \mathcal{F}_M} (\mathbb{E}X_i X_j - \mathbb{E}X_i \mathbb{E}X_j). \quad (6)$$

Note that if the matchings of the colors i, j are vertex-disjoint, then X_i and X_j are independent and $\mathbb{E}X_i X_j - \mathbb{E}X_i \mathbb{E}X_j = 0$.

Since the matchings in \mathcal{F}_M are disjoint, for every $F_i \in \mathcal{F}_M$, at most $6 = \binom{4}{2}$ matchings $F_j \in \mathcal{F}_M$ can have an edge contained in $\bigcup F_i$. This means that there exist at most $2 \cdot 6n$ pairs (F_i, F_j) such that an edge from F_j is contained in $\bigcup F_i$, or vice versa. Thus the contribution of such pairs to $\text{Var } X$ is at most $12n$.

Apart from vertex-disjointness, there are two more possible forms of $F_i \cup F_j$:

- I. three connected components: one 2-path and two disjoint edges, or
- II. two vertex-disjoint 2-paths.

Examine Case I. Let $F_i = \{a, b\}$, where $a = \{x, y\}$, $b = \{u, v\}$, and let $F_j = \{c, d\}$, where $c = \{x, z\}$ and $d = \{s, t\}$.

If x is a heavy vertex, then $\mathbb{E}X_i X_j - \mathbb{E}X_i \mathbb{E}X_j = 0$. Indeed, since the events $\{a \subseteq S\} = \{y \in S\}$, $\{b \subseteq S\}$, $\{c \subseteq S\} = \{z \in S\}$, and $\{d \subseteq S\}$ are mutually independent, we have

$$\begin{aligned} & \mathbb{E}X_i \mathbb{E}X_j \\ &= \left(\mathbb{P}(a \subseteq S \text{ and } b \not\subseteq S) + \mathbb{P}(a \not\subseteq S \text{ and } b \subseteq S) + \mathbb{P}(a \subseteq S \text{ and } b \subseteq S) \right) \\ & \quad \cdot \left(\mathbb{P}(c \subseteq S \text{ and } d \not\subseteq S) + \mathbb{P}(c \not\subseteq S \text{ and } d \subseteq S) + \mathbb{P}(c \subseteq S \text{ and } d \subseteq S) \right) \\ &= \left(\mathbb{P}(a \subseteq S) \mathbb{P}(b \not\subseteq S) + \mathbb{P}(a \not\subseteq S) \mathbb{P}(b \subseteq S) + \mathbb{P}(a \subseteq S) \mathbb{P}(b \subseteq S) \right) \\ & \quad \cdot \left(\mathbb{P}(c \subseteq S) \mathbb{P}(d \not\subseteq S) + \mathbb{P}(c \not\subseteq S) \mathbb{P}(d \subseteq S) + \mathbb{P}(c \subseteq S) \mathbb{P}(d \subseteq S) \right) \\ &= \mathbb{P}(a, c \subseteq S \text{ and } b, d \not\subseteq S) + \mathbb{P}(a, d \subseteq S \text{ and } b, c \not\subseteq S) + \mathbb{P}(b, c \subseteq S \text{ and } a, d \not\subseteq S) \\ & \quad + \mathbb{P}(b, d \subseteq S \text{ and } a, c \not\subseteq S) + \mathbb{P}(a, b, c \subseteq S \text{ and } d \not\subseteq S) + \mathbb{P}(a, b, d \subseteq S \text{ and } c \not\subseteq S) \\ & \quad + \mathbb{P}(a, c, d \subseteq S \text{ and } b \not\subseteq S) + \mathbb{P}(b, c, d \subseteq S \text{ and } a \not\subseteq S) + \mathbb{P}(a, b, c, d \subseteq S) \\ &= \mathbb{E}X_i X_j. \end{aligned}$$

If x is not a heavy vertex, the contribution to $\text{Var } X$ for such (F_i, F_j) is at most $4 \cdot n \cdot \frac{\epsilon^2 n}{10^6}$, since there are at most n ways to choose F_i , four ways to choose x , and at most $\frac{\epsilon^2 n}{10^6}$ ways to choose F_j .

In case II, let $F_i = \{a, b\}$, where $a = \{x, y\}$, $b = \{u, v\}$, and let $F_j = \{c, d\}$, where $c = \{x, z\}$ and $d = \{u, t\}$.

If both x and u are heavy vertices, then the events $\{a \subseteq S\} = \{y \in S\}$, $\{b \subseteq S\} = \{v \in S\}$, $\{c \subseteq S\} = \{z \in S\}$, and $\{d \subseteq S\} = \{t \in S\}$ are mutually independent. Similarly to case I, we have $\mathbb{E}X_i X_j - \mathbb{E}X_i \mathbb{E}X_j = 0$. Otherwise the contribution to $\text{Var } X$ for such (F_i, F_j) is at most $4 \cdot n \cdot \frac{\epsilon^2 n}{10^6}$.

Summing, for n large enough (so that $12 \leq \frac{2\epsilon^2}{10^6}n$) we have

$$\text{Var } X \leq 12n + 2 \cdot 4 \cdot n \cdot \frac{\epsilon^2 n}{10^6} \leq \frac{\epsilon^2 n^2}{10^5}.$$

Since $1/2 \leq p < 1$, we have $2p^2 - p^4 \geq 7/16$. Applying Chebyshev's inequality and using the assumption $|\mathcal{F}_M| \geq n/40$, we have

$$\begin{aligned} \mathbb{P}(X \leq (1 - \epsilon)|\mathcal{F}_M| \cdot (2p^2 - p^4)) &\leq \mathbb{P}(X \leq \mathbb{E}X - \epsilon|\mathcal{F}_M| \cdot (2p^2 - p^4)) \\ &\leq \frac{\text{Var } X}{(\epsilon|\mathcal{F}_M| \cdot (2p^2 - p^4))^2} \leq \frac{\epsilon^2 n^2}{10^5 \epsilon^2 (n/40)^2 (7/16)^2} \leq 1/10, \end{aligned}$$

which completes the proof. \square

Lemma 10. *With probability at least 0.9, the number of color classes in \mathcal{F}_E that are contained in S is at least $(1 - \epsilon)|\mathcal{F}_E| \cdot p^2$.*

Proof. For $F_i \in \mathcal{F}_E$, let

$$Y_i := \mathbb{1}_{\{F_i \subseteq S\}}$$

be the indicator random variable that the edge of color i is contained in S . Then $\mathbb{E}Y_i \geq p^2$. Let

$$Y := \sum_{F_i \in \mathcal{F}_E} Y_i.$$

We have

$$\mathbb{E}Y \geq |\mathcal{F}_E| \cdot p^2. \quad (7)$$

To prove the lemma, we shall apply Chebyshev's inequality (Theorem 4). For this purpose we have to estimate $\text{Var } Y$. We have

$$\text{Var } Y = \mathbb{E}Y^2 - (\mathbb{E}Y)^2 = \sum_{F_i, F_j \in \mathcal{F}_E} (\mathbb{E}Y_i Y_j - \mathbb{E}Y_i \mathbb{E}Y_j). \quad (8)$$

Note that if the edges of colors i, j are vertex-disjoint, then Y_i and Y_j are independent and $\mathbb{E}Y_i Y_j - \mathbb{E}Y_i \mathbb{E}Y_j = 0$.

Furthermore, if the edges of two distinct color classes $F_i, F_j \in \mathcal{F}_E$, say uv and uw , intersect at a heavy vertex u that is in D , then $\mathbb{E}Y_i Y_j - \mathbb{E}Y_i \mathbb{E}Y_j = \mathbb{P}(v, w \in S) - \mathbb{P}(v \in S)\mathbb{P}(w \in S) = 0$.

Therefore besides the case $F_i = F_j$, the non-zero contribution to $\mathbb{E}Y$ can only come from F_i, F_j that the edges intersect at a non-heavy vertex. Note that by definition, such a non-heavy vertex is incident to at most $(\epsilon^2/10^6)n$ edges in $\cup \mathcal{F}_E$. Therefore there are at most $2 \cdot |\mathcal{F}_E| \cdot 2 \cdot (\epsilon^2/10^6)n$ such pairs (F_i, F_j) .

Therefore using the assumption that $|\mathcal{F}_E| \geq n/40$, for n large enough (so that $1 \leq \frac{4\epsilon^2}{10^6}n$) we have

$$\text{Var } Y \leq |\mathcal{F}_E| + 4|\mathcal{F}_E|(\epsilon^2/10^6)n \leq \frac{8 \cdot 40\epsilon^2}{10^6}|\mathcal{F}_E| \cdot \frac{n}{40} \leq \frac{\epsilon^2}{10^3}|\mathcal{F}_E|^2.$$

Therefore by Chebyshev's inequality and $p \geq 1/2$, we have

$$\mathbb{P}(Y \leq |\mathcal{F}_E| \cdot p^2 - \epsilon|\mathcal{F}_E| \cdot p^2) \leq \mathbb{P}(Y \leq \mathbb{E}Y - \epsilon|\mathcal{F}_E| \cdot p^2) \leq \frac{\text{Var } Y}{(\epsilon|\mathcal{F}_E| \cdot p^2)^2} \leq \frac{\epsilon^2 |\mathcal{F}_E|^2}{10^3 \cdot \epsilon^2 |\mathcal{F}_E|^2 \cdot (1/2)^4} \leq 1/10,$$

which completes the proof. \square

Proof of Theorem 7. Combining Lemma 8–10 and taking a union bound, we know that with positive probability (at least $1/2$), all of the following hold:

$$|S| \leq np + n^{2/3}, \quad X \geq (1 - \epsilon)|\mathcal{F}_M| \cdot (2p^2 - p^4), \quad \text{and} \quad Y \geq (1 - \epsilon)|\mathcal{F}_E| \cdot p^2.$$

Therefore there is some S such that all of the above hold. Then the number of rainbow edges contained in S is at least

$$\begin{aligned} X + Y &\geq (1 - \epsilon)|\mathcal{F}_M| \cdot (2p^2 - p^4) + (1 - \epsilon)|\mathcal{F}_E| \cdot p^2 \\ &\geq (1 - \epsilon)(\alpha - \xi)n(2p^2 - p^4) + (1 - \epsilon)(1 - \alpha - \xi)n \cdot p^2. \end{aligned}$$

With a look at (2) and setting $3c := (1 - \epsilon)(\alpha - \xi)(2p^2 - p^4) + (1 - \epsilon)(1 - \alpha - \xi)p^2 - p > 0$, for n large enough we have that $|S| \leq \beta n$ for $\beta := p + c$ and the number of rainbow edges contained in S is at least $(\beta + c)n = ((p + 3c) - c)n$. Therefore this completes the proof of Theorem 7. \square

3.1.1 A simplified proof of Theorem 6

Indeed, Theorem 6 will follow from the following theorem and Theorem 2 by setting $N = |S| \leq n$ and $k = cn$.

Theorem 11. *For any $\alpha > 1/2$, there exist $c > 0$ such that for any large enough n , given an n -vertex graph G and an edge coloring of G satisfying the assumption in Theorem 6, there exists a subset S of $V(G)$ of such that the size of a rainbow edge set contained in S is at least cn larger than $|S|$.*

Proof. We take $p \in (0, 1)$ and $\xi > 0$ satisfying (2) and set

$$c := (\alpha - \xi)(2p^2 - p^4) + (1 - \alpha - \xi)p^2 - p > 0.$$

And we included each vertex of G into S independently with probability p . Let r_S be the largest size of a rainbow edge set contained in S . By linearity of expectations and following similarly as the arguments in (5) and (7), we have

$$\begin{aligned} &\mathbb{E}(r_S - |S|) \\ &= |\mathcal{F}_M|(2p^2 - p^4) + |\mathcal{F}_E|p^2 - pn \\ &\geq \left((\alpha - \xi)(2p^2 - p^4) + (1 - \alpha - \xi)p^2 - p \right) n \geq cn. \end{aligned}$$

Therefore there exists an instance of S such that the conclusion of the theorem holds. \square

3.1.2 Sharpness of the condition $\alpha > 1/2$

To get the logarithmic in n bound on the rainbow girth, it is necessary to assume $\alpha > 1/2$ in Theorem 5 and Theorem 6. The following n -vertex graph F with $n/2$ matchings of size 2 in \mathcal{F}_M and $n/2$ single edges in \mathcal{F}_E (so that $\alpha = 1/2$) has rainbow girth linear in n .

For simplicity, we may assume that $n \geq 8$ is divisible by 4. The vertices of F are $v_{i,j}$ for $i = 1, \dots, n/4$ and $j = 1, \dots, 4$. For each $1 \leq i \leq n/4$, the four vertices $v_{i,j}$ for $1 \leq j \leq 4$ form a 4-cycle: $\{v_{i,1}v_{i,2}, v_{i,3}v_{i,4}\}$ is a matching of size 2 from \mathcal{F}_M and $v_{i,2}v_{i,3}, v_{i,4}v_{i,1}$ are two edges in two colors from \mathcal{F}_E . And $\{v_{i,3}v_{i+1,2}, v_{i,4}v_{i+1,1}\}$ (with the first subscripts module $n/4$) for each $1 \leq i \leq n/4$ is a matching of size 2 from \mathcal{F}_M . It can be verified that the n -vertex graph F satisfies the assumption but has rainbow girth $n/2$.

3.2 Triangles and single edges

Theorem 12. *For any constant $\alpha > 0$, if an n -vertex graph G and edge coloring of G satisfying that at least αn color classes consisting of a triangle and at least $(1 - \alpha)n$ color classes consisting of a single edge, then the rainbow girth is at most $C \log n$ for some constant $C(\alpha) > 0$.*

We can prove a slightly stronger result, which allows the total number of color classes to be less than n .

Theorem 13. *For any constants $\alpha > 0$ and $\xi < \alpha/3$, if an n -vertex graph G and edge coloring of G satisfying that at least $(\alpha - \xi)n$ color classes consisting of a triangle and at least $(1 - \alpha - \xi)n$ color classes consisting of a single edge, then the rainbow girth is at most $C \log n$ for some constant $C(\alpha, \xi) > 0$.*

Proof. We take each of the single edge and two edges from each triangle. Then in the n -vertex graph, we have at least $2 \cdot (\alpha - \xi)n + (1 - \alpha - \xi)n = (1 + \alpha - 3\xi)n$ edges. Since $1 + \alpha - 3\xi > 1$, Theorem 2 implies that there is a cycle of length at most $C \log n$ for some constant $C(\alpha, \xi) > 0$. If this cycle is not rainbow, we can replace two edges of the same color, which must come from the same triangle, by the other edge in the triangle to get a shorter cycle. Do it repeatedly until we obtain a rainbow cycle, which is of length at most $C \log n$. This completes the proof. \square

3.2.1 Sharpness of the condition $\alpha > 0$

To get a logarithmic in n bound on the rainbow girth, we need $\alpha > 0$ in Theorem 12 and Theorem 13. Otherwise for $\alpha = 0$, an n -cycle with n edges in distinct colors has rainbow girth n .

3.3 Matchings and triangles

Theorem 14. *There exists a constant $C > 0$ such that for any n -vertex graph G and edge coloring of G with n colors, if each color class is either a matching of size 2 or a triangle, then the rainbow girth is at most $C \log n$.*

Proof. This result follows immediately from Theorem 5 and Theorem 12. If the number of triangles is at least αn for some $\alpha > 0$, then Theorem 12 implies the result. Otherwise, the number of matchings of size 2 is at least $(1 - o(1))n$, where Theorem 5 implies the result. \square

Acknowledgements.

We thank Ron Aharoni for helpful suggestions during the preparation of this note. We thank Michael Krivelevich for pointing out an alternative proof of a theorem that we cite. We thank Lutz Warnke for meaningful discussion inspiring the simplified proof of Theorem 6. And thanks for the comments from the referees.

References

- [1] R. Aharoni, E. Berger, M. Chudnovsky, H. Guo, and S. Zerbib. Non-uniform degrees and rainbow versions of the Caccetta-Häggkvist conjecture. *SIAM J. Discrete Math.* **37** (2023), 1704–1714.
- [2] R. Aharoni, M. DeVos, and R. Holzman. Rainbow triangles and the Caccetta-Häggkvist conjecture. *J. Graph Theory* **92** (2019), 347–360.

- [3] R. Aharoni and H. Guo. Rainbow cycles for families of matchings. *Israel J. Math.* **256** (2023), 1–8.
- [4] B. Bollobás and E. Szemerédi. Girth of sparse graphs. *J. Graph Theory* **39** (2002), 194–200.
- [5] L. Caccetta and R. Häggkvist. On minimal digraphs with given girth. *Congress. Numer.* **21** (1978), 181–187.
- [6] V. Chvátal and E. Szemerédi. Short cycles in directed graphs. *J. Combin. Theory Ser. B* **35** (1983), 323–327.
- [7] M. DeVos, M. Drescher, D. Funk, S. González Hermosillo de la Maza, K. Guo, T. Huynh, B. Mohar, and A. Montejano. Short rainbow cycles in graphs and matroids. *J. Graph Theory* **96** (2021), 192–202.
- [8] Y.O. Hamidoune. A note on minimal directed graphs with given girth. *J. Combin. Theory Ser. B* **43** (1987), 343–348.
- [9] C.T. Hoàng and B. Reed. A note on short cycles in diagraphs. *Discrete Math.* **66** (1987), 103–107.
- [10] P. Hompe and S. Spirk. Further approximations for Aharoni’s rainbow generalization of the Caccetta-Häggkvist conjecture. *Electron. J. Combin.* **29** (2022), #P1.55.
- [11] T. Nishimura. Short cycles in digraphs. *Discrete Math.* **38** (1988), 295–298.
- [12] J. Shen. On the girth of digraphs. *Discrete Math.* **211** (2000), 167–181.
- [13] J. Shen. On the Caccetta-Häggkvist conjecture. *Graphs Combin.* **18** (2002), 645–654.

A An alternative proof of rainbow girth for matchings

Michael Krivelevich pointed out the following result to the author in a personal communication.

Theorem 15. *For any $\gamma > 0$ satisfying $1 - (\frac{1}{2})^{4\gamma} < \gamma$, the following holds: For an n -vertex graph G with an edge coloring with at least γn colors, if each color class is a matching of size 2, then the rainbow girth of G is at most $C \log n$ for some $C(\gamma) > 0$.*

Proof. Let \mathcal{F}_M be the collection of γn many size-2 matchings. Let H be the graph with vertex set $V(G) = [n]$ and edge set $\cup \mathcal{F}_M$. Independently we take one edge from each matching $F \in \mathcal{F}_M$ uniformly at random, and we set the chosen edge set as T . Then

$$|T| = \gamma n.$$

Let Z be the vertex subset $\{z \in V(G) : z \notin e \text{ for any } e \in T\}$. If we can show $Z \geq (1 - \gamma')n$ for some constant $\gamma' < \gamma$. Then the edge set T is contained in a vertex set of size at most $\gamma'n$. Then Theorem 2 implies there is a rainbow cycle in T of length $O(\log n)$, which completes the proof.

Note that for each vertex $v \in V(G)$,

$$\mathbb{P}(v \in Z) = \left(\frac{1}{2}\right)^{d_H(v)},$$

i.e., none of edges in H that are incident to v is chosen into T . Therefore

$$\mathbb{E}|Z| = \sum_{v \in V(G)} \left(\frac{1}{2}\right)^{d_H(v)} \geq n \left(\frac{1}{2}\right)^{\sum_{v \in V(G)} d_H(v)/n} = n \left(\frac{1}{2}\right)^{4\gamma},$$

where the inequality is by the convexity of the function $f(x) = (1/2)^x$. Therefore there exists a choice of T such that

$$|Z| \geq n \left(\frac{1}{2}\right)^{4\gamma}.$$

Hence we can set $\gamma' = 1 - (\frac{1}{2})^{4\gamma}$. By assumption we have $\gamma' < \gamma$, and this completes the proof. \square

Though not necessary for proving the theorem, applying the bounded difference inequality, one can also show that Z in the proof is concentrated around $\mathbb{E}Z$, since each choice of edge influences the result by at most 4.

Remark 16. *Numerical solution says for $\gamma \geq 0.922523266904828$, we have $1 - (\frac{1}{2})^{4\gamma} < \gamma$ so that the conclusion of Theorem 15 holds.*