NON-UNIFORM DEGREES AND RAINBOW VERSIONS OF THE CACCETTA-HÄGGKVIST CONJECTURE

RON AHARONI, ELI BERGER, MARIA CHUDNOVSKY, HE GUO, AND SHIRA ZERBIB

ABSTRACT. The Caccetta-Häggkvist conjecture (denoted below CHC) states that the directed girth (the smallest length of a directed cycle) $dgirth(D)$ of a directed graph D on n vertices is at most $\lceil \frac{n}{\delta^+(D)} \rceil$, where $\delta^+(D)$ is the minimum out-degree of D. We consider a version involving all out-degrees, not merely the minimum one, and prove that if D does not contain a sink, then $dgirth(D) \leq 2 \sum_{v \in V(D)} \frac{1}{deg^+(v)+1}$. In the spirit of a generalization of the CHC to rainbow cycles in [1], this suggests the conjecture that given non-empty sets $F_1, \ldots, F_n$ of edges of K_n, there exists a rainbow cycle of length at most $2 \sum_{1 \leq i \leq n} \frac{1}{|F_i|+1}$. We prove a bit stronger result when $1 \leq |F_i| \leq 2$, thereby strengthening a result of DeVos et. al [6]. We prove a logarithmic bound on the rainbow girth in the case that the sets F_i are triangles.

1. INTRODUCTION

The *directed girth* $dgirth(D)$ of a directed graph (digraph) D is the smallest length of a directed cycle in D (∞ if there is no directed cycle). A famous conjecture of Caccetta and Häggkvist [4] is that

$$dgirth(D) \leq \left\lceil \frac{n}{\delta^+(D)} \right\rceil,$$

where $n = |V(D)|$ and $\delta^+(D)$ is the minimum out-degree over all vertices of D. We use the acronym CHC for it. See [14] for a survey of known results on this conjecture up to the year 2006.

The CHC is known to be true asymptotically: in [13] it was proved that

$$(1) \qquad dgirth(D) \leq \left\lceil \frac{n}{\delta^+(D)} \right\rceil + 73.$$

Much of the research on the conjecture has addressed the case $dgirth(D) = 3$. The best result so far is due to Hladký, Král', and Norin [8].

Theorem 1.1. *Every n-vertex digraph with minimum out-degree at least 0.3465n contains a directed triangle.*

A natural question is finding upper bounds on $dgirth(D)$ in terms of all out-degrees of the vertices of D, rather than merely the minimum out-degree. Let

$$\psi(D) := \sum_{v \in V(D)} \frac{1}{\deg^+(v)}.$$

Seymour asked (see [9]) whether CHC could be generalized to

$$(2) \qquad dgirth(D) \leq \lceil \psi(D) \rceil.$$

This was answered in the negative by Hompe [9]. Here we prove "half" of this result, namely:

Theorem 1.2. *For any digraph D, we have*

$$(3) \qquad dgirth(D) \leq 2\psi(D).$$

In fact, we use a slightly different function. Let

$$\varphi(D) := \sum_{v \in V(D)} \frac{1}{deg^+(v) + 1}.$$

Theorem 1.3. *If all out-degrees in D are positive, then $dgirth(D) \leq 2\varphi(D)$.*

This is proved in Section 2.

In Section 3 and 4 we discuss a rainbow, undirected generalization of the CHC, suggested in [1].

Definition 1.4. Let $\mathcal{F} = (F_1, \ldots, F_m)$ be a family of subsets of $E(K_n)$. A *rainbow cycle* for $\mathcal{F}$ is a cycle whose edges are chosen each from a different F_i. The *rainbow girth* $rgirth(\mathcal{F})$ of $\mathcal{F}$ is the smallest length of a rainbow cycle.

Note that an edge belonging to two different sets F_i yields a rainbow digon (that is, a rainbow cycle of length 2). Thus for our purposes we can assume disjointness of the sets F_i. The generalized CHC is:

Conjecture 1.5. *For $\mathcal{F} = (F_1, \ldots, F_n)$ a family of subsets of $E(K_n)$, we have $rgirth(\mathcal{F}) \leq \lceil \frac{n}{\min_{1 \leq i \leq n} |F_i|} \rceil$.*

As explained in Section 3, the CHC is the case in which the sets F_i are stars, with distinct apexes.

Remark 1.6. An advantage of the rainbow version is that it detaches the link between the number of sets and the number n of vertices. The question makes sense for any number of sets. Here are two results on the case $rgirth(\mathcal{F}) = 3$:

Theorem 1.7. [7] $n^2/8 + o(n)$ *triangles on n vertices have a rainbow triangle.*

Theorem 1.8. [1] $\frac{9}{8}n$ *(or more) sets of edges in K_n, each of size $\frac{n}{3}$ or more, have a rainbow triangle.*

In [10] a slight improvement was proved, $\frac{9}{8}n$ being replaced by $1.1077n$.

In [11] it was shown that the order of magnitude in the conjecture is correct:

Theorem 1.9. *There exists a constant $0 < C \leq 10^{11}$ such that for any n and any family $\mathcal{F} = (F_1, \ldots, F_n)$ of subsets of $E(K_n)$, we have $rgirth(\mathcal{F}) \leq C \cdot \frac{n}{\min_{1 \leq i \leq n} |F_i|}$.*

A natural challenge is to improve the bound on C.

In [1] a triangles version was proved:

Theorem 1.10. *n sets of edges in K_n, each of size $0.4n$ or more, have a rainbow triangle.*

Compare with the coefficient 0.3465 appearing in Theorem 1.1. In [10], $0.4n$ was replaced by $0.3988n$.

In [6] the following was proved:

Theorem 1.11. *Conjecture 1.5 is true when $|F_i| = 2$ for all $1 \leq i \leq n$.*

The rainbow analogue of Theorem 1.3 is:

Conjecture 1.12. *$rgirth(\mathcal{F}) \leq 2 \sum_{1 \leq i \leq n} \frac{1}{|F_i|+1}$ for any family $\mathcal{F} = (F_1, \ldots, F_n)$ of subsets of $E(K_n)$.*

If true, this would enable taking $C = 2$ in Theorem 1.9. Here we prove:

Theorem 1.13. *Conjecture 1.12 is true when $1 \leq |F_i| \leq 2$ for all $1 \leq i \leq n$.*

We shall prove this (in Section 4) via a result generalizing Theorem 1.11. Let

$$\psi(\mathcal{F}) := \sum_{1 \leq i \leq n} \frac{1}{|F_i|}.$$

We show:

Theorem 1.14. *If $|F_i| \leq 2$ for all $1 \leq i \leq n$, then $rgirth(\mathcal{F}) \leq \lceil \psi(\mathcal{F}) \rceil$.*

Section 3 deals with a special case of Conjecture 1.12, in which all sets F_i are triangles. This case is of particular interest, for the following reason. We know (from the original CHC) that $\min |F_i| \cdot rgirth(\mathcal{F})$ may be close to n, and that this can be exactly n when the sets F_i are stars. In [2] it was proved that if each F_i is a matching of size 2 then $rgirth(\mathcal{F}) = O(\log n)$. Note that a set of graph edges not containing two disjoint edges is a star or a triangle. So, the remaining case, in terms of some uniform assumption on the sets F_i, is that of triangles. We show that this case is close to the case of matchings of size 2:

Theorem 1.15. *For any constant $\alpha > 1/2$ there exists a constant C such that for any n and any family $\mathcal{F} = (F_1, ..., F_{\lceil \alpha n \rceil})$ of subsets of $E(K_n)$ where each F_i is a triangle, there is a rainbow cycle of length at most $C \log n$.*

We also prove, via a random construction, that this result is best possible, in the sense that there are families of n triangles, in which the rainbow girth is $\Omega(\log n)$. (For a stronger version, see Theorem 3.5.)

Theorem 1.16. *There exists a positive constant c such that for any n, there exists a family $\mathcal{F}_n$ of n triangles on n vertices satisfying $rgirth(\mathcal{F}_n) \geq c \log n$.*

2. Non-uniform out-degrees

As mentioned in Section 1, Seymour asked (see [9]) whether the directed girth of a digraph can be bounded from above by an expression involving all out-degrees. A natural such expression is

$$\psi(D) = \sum_{v \in V(D)} \frac{1}{\deg^+(v)}.$$

Hompe [9] showed that $\psi(D)$ is not always an upper bound on the directed girth. His counterexample is obtained from a directed cycle by replacing each vertex of the cycle by a transitive tournament T_k with k vertices, for some k. If the cycle is of length ℓ then the resulting graph D satisfies $dgirth(D) = \ell$ and $\delta^+(D) = k$. Furthermore, $\psi(D) = dgirth(D) \cdot \sum_{i=\delta^+(D)}^{2\delta^+(D)-1} \frac{1}{i}$, and $\varphi(D) = dgirth(D) \cdot \sum_{i=\delta^+(D)}^{2\delta^+(D)-1} \frac{1}{i+1}$, and $\lim_{|V(D)| \to \infty} \frac{dgirth(D)}{\psi(D)} = \lim_{|V(D)| \to \infty} \frac{dgirth(D)}{\varphi(D)} = \log_2 e$.
Possibly, this example is best:

Question 2.1. *Is it true that for any digraph D, $dgirth(D) \le \lceil \log_2 e \cdot \psi(D) \rceil$?*

A *sink* in a digraph is a vertex with out-degree 0. If D contains a sink, then $\psi(D) = \infty$, and thus $dgirth(D) \le \psi(D)$. Thus the interesting case for us is that in which no sink exists. In this case we can prove twice the bound suggested by Seymour, in fact a bit better. Recall that

$$\varphi(D) = \sum_{v \in V(D)} \frac{1}{\deg^+(v) + 1}.$$

Theorem 2.2. *If a digraph D has no sink, then $dgirth(D) \le 2\varphi(D)$. Equality holds if and only if D is a Hamilton cycle (in which $dgirth(D) = 2\varphi(D) = |V(D)|$) or a complete digraph (in which case $dgirth(D) = 2\varphi(D) = 2$).*

In [5] the inequality was proved in the case that all out-degrees are equal.

Proof of Theorem 2.2. Let us first prove the inequality. We call a digraph K not containing a sink φ-*critical* if for every vertex $v \in V(K)$ either $\varphi(K - v) > \varphi(K)$ or $K - v$ contains a sink.

Claim 2.2.1. *A φ-critical graph is vertex-disjoint union of directed cycles.*

Proof of Theorem 2.2 based on Claim 2.2.1. We remove vertices one by one from D, while keeping the graph sink-less and not increasing φ, until we reach a φ-critical graph K that is vertex-disjoint union of directed cycles. Since K is union of cycles, we have $dgirth(K) \le |V(K)| = 2\varphi(K)$. Since K is a subgraph of D, we have $dgirth(D) \le dgirth(K)$. Since we keep φ not increasing during the removal, we have $\varphi(K) \le \varphi(D)$. Combining these, we have

$$dgirth(D) \le dgrith(K) \le 2\varphi(K) \le 2\varphi(D),$$

which completes the proof. $\square$

To prove Claim 2.2.1 we observe:

Claim 2.2.2. *In any digraph D, there exists a vertex v for which*

$$(4) \qquad \frac{1}{deg^+(v) + 1} \ge \sum_{u \in N^-(v)} \frac{1}{deg^+(u)} \frac{1}{deg^+(u) + 1}$$

Proof. The claim will follow if we show that the sums, over all vertices of D, of the two sides, are the same. On the left-hand side the sum is, by definition, $\varphi(D)$. On the right-hand side, the number of times every vertex u appears is $deg^+(u)$, and hence we get $\sum_{u \in V(D)} \frac{1}{deg^+(u)+1}$, which is again $\varphi(D)$. □

Proof of Claim 2.2.1. Let D be a φ-critical graph and A be the set of vertices v satisfying (4). Note that for any $v \in A$,

$$\varphi(D) - \varphi(D-v) = \frac{1}{deg^+(v)+1} - \sum_{u \in N^-(v)} \left(\frac{1}{deg^+(u)} - \frac{1}{deg^+(u)+1} \right) \geq 0.$$

As D is φ-critical, for every $v \in A$, $D - v$ has a sink, which means there exists a vertex w such that $N^+(w) = v$. Then the w-term in the right-hand side of (4) is $\frac{1}{2}$, while the left-hand side is at most $\frac{1}{2}$ as D is sink-less, and thus $N^-(v) = \{w\}$ and $deg^+(v) = 1$. Namely, both in-degree and out-degree of v are 1. It follows that for every $v \in A$ equality holds in (4), and since the sums over all vertices v of the right-hand sides and the left-hand sides in (4) are equal, it implies that $A = V(D)$. Therefore every vertex of D has both in-degree and out-degree equal to 1, which means D is vertex-disjoint union of directed cycles. □

This concludes the proof of the inequality in Theorem 2.2.

For the second part of the theorem, assume that $dgirth(D) = 2\varphi(D)$. Tracking the proof of the inequality, for $0 \leq i \leq t$ let $D_i = D - \{v_j \mid 1 \leq j \leq i\}$ (so $D_0 = D$), where $v_1, \ldots, v_t$ are the removed vertices from D (if any), in this order. Then

$$dgirth(D_{i-1}) \leq dgirth(D_i) \leq 2\varphi(D_i) \leq 2\varphi(D_{i-1}),$$

where the second inequality is by the first part of this theorem. By the assumption that $dgirth(D) = 2\varphi(D)$, equalities hold throughout, namely $dgirth(D_i) = dgirth(D_{i-1}) = 2\varphi(D_i)$ and $\varphi(D_i) = \varphi(D_{i-1})$. Let $K = D_t$. By the construction and Claim 2.2.1, the φ-critical graph K is the vertex-disjoint union of directed cycles, and since $dgirth(K) = 2\varphi(K)$ it is a single cycle, namely it is a Hamilton cycle.

If $V(K) = V(D)$, then D itself is a Hamilton cycle, proving the desired result. So, we can assume that $V(K) \subsetneq V(D)$.

Claim 2.2.3. *If $V(K) \subsetneq V(D)$, then K is a directed 2-cycle, i.e., a directed digon.*

To show this, let $p = |N^+_{D_{t-1}}(v_t) \cap V(K)|$ and $q = |N^-_{D_{t-1}}(v_t) \cap V(K)|$. Then the fact that $\varphi(D_{t-1}) = \varphi(K)$ implies that

$$\frac{1}{p+1} = q \left(\frac{1}{2} - \frac{1}{3} \right) = \frac{q}{6}.$$

Therefore we have $(p, q) = (5, 1), (2, 2),$ or $(1, 3)$. Since $dgirth(K) = dgirth(D_{t-1})$ we have $(p, q) = (2, 2)$ and K is a digon, otherwise D_{t-1} has a shorter directed cycle than K. This proves the claim, and implies that D_{t-1} is a complete directed graph on three vertices.

This was the first step in the inductive proof of the following claim:

Claim 2.2.4. *If $V(K) \subsetneq V(D)$, then D_i is the complete digraph on $2+t-i$ vertices for all $0 \leq i \leq t$.*

We prove this by induction on $|V(D)| - i$. Assuming that D_i is complete digraph on $2 + t - i$ vertices, let $p = |N^+_{D_{i-1}}(v_i) \cap V(D_i)|$ and $q = |N^-_{D_{i-1}}(v_i) \cap V(D_i)|$. Since $\varphi(D_i) = \varphi(D_{i-1})$, we have

$$\frac{1}{p+1} = q\Big(\frac{1}{|V(D_i)|} - \frac{1}{|V(D_i)| + 1}\Big) = \frac{q}{|V(D_i)|(|V(D_i)| + 1)}.$$

Since $0 \le p, q \le |V(D_i)|$, we have $p = q = |V(D_i)|$, so D_{i-1} is a complete digraph on $|V(D_i)| + 1 = 2 + t - (i-1)$ vertices. This completes the proof of the claim.

Putting $i = 0$ proves the statement in the theorem. □

3. The rainbow version of CHC for triangles

In this section and the next we consider the rainbow, undirected generalization of the CHC.

Here is an explanation why Conjecture 1.5 is a generalization of CHC. For a directed edge $e = (u, v)$ let $n(e)$ be the undirected pair $\{u, v\}$. Given a digraph D, for every vertex $u \in V(D)$ let $F_u = \{n(uv) \mid (u, v) \in E(D)\}$ be the star of edges leaving u, with their direction removed. Let $G(D)$ be an undirected graph with vertex set $V(D)$ and edge set $\cup_{u \in V(D)} F_u$. Note that sets F_u are stars with distinct apexes in G. It is easy to verify that a sequence of vertices $v_1 v_2 \ldots v_k$ forms a rainbow cycle in G if and only if they form a directed cycle in D.

The CHC holds asymptotically: it is known that $dgirth(D) \le \lceil \frac{n}{\delta^+(D)} \rceil + 73$ (see [13]). In the undirected rainbow version the gap between the conjecture and the known bounds is much larger.

In [2] it was proved that there exists a constant C for which every set of n matchings of size 2 in K_n has a rainbow cycle of length at most $C \log n$. If $\mathcal{F} = (F_1, \ldots, F_n)$ are n stars with distinct apexes then directing all edges in F_i away from the apex yields, by Theorem 2.2, we have that $rgirth(\mathcal{F}) \le 2\psi(\mathcal{F})$. We cannot prove the same if the apexes are allowed to coincide:

Problem 3.1. *Prove (or disprove) $rgirth(\mathcal{F}) \le 2\psi(\mathcal{F})$ for any set of n stars in K_n.*

Since a set of edges not containing a matching of size 2 is either a star or a triangle, the remaining case (assuming all sets F_i are of size at least 2) is that of triangles. Like in the case of sets of edges containing each a pair of disjoint edges, a better than linear bound can be proved in this case:

Theorem 3.2. *For any constant $\alpha > 1/2$ there exists a constant C such that for any n and any family $\mathcal{F} = (F_1, ..., F_{\lceil \alpha n \rceil})$ of subsets of $E(K_n)$ where each F_i is a triangle, there is a rainbow cycle of length at most $C \log n$.*

The proof uses the following result of Bollobás and Szemerédi [3].

Theorem 3.3. *For $n \ge 4$ and $k \ge 2$, every n-vertex graph with $n + k$ edges has girth at most*

$$\frac{2(n+k)}{3k}(\log k + \log \log k + 4).$$

Proof of Theorem 3.2. As noted above, we may assume that the sets F_i are edge-disjoint, or else $rgirth(\mathcal{F}) = 2$. Choosing any two edges from each F_i, we obtain an n-vertex graph with at least $(1 + \delta)n$ edges, where $\delta = 2\alpha - 1 > 0$. Then Theorem 3.3 implies that there is a cycle of length at most $C \log n$ for some positive $C(\alpha)$. If such a cycle is not rainbow, we can replace two edges in the same edge

set F_i by the other edge in the triangle F_i to get a shorter cycle. Do it repeatedly until we obtain a rainbow cycle, which is of length at most $C \log n$.   □

The next example, the crown-like graph, shows that the condition $\alpha > \frac{1}{2}$ is necessary, namely for $\alpha = \frac{1}{2}$ the rainbow girth can be linear in n, not logarithmic.

Example 3.4. Let $m = \lfloor \frac{1}{2}n \rfloor$. Let K be a cycle on m vertices with edges $e_1, \ldots, e_m$. Let $v_1, \ldots, v_m$ be distinct vertices not on K, and let F_i be the triangle with vertex set $e_i \cup \{v_i\}$. The rainbow girth is m.

The following theorem implies that the $\log n$ bound in Theorem 3.2 is the right order of magnitude. The following is a fine-tuned version of Theorem 1.16 from the introduction:

Theorem 3.5. *For any $\alpha > 0$, there exists a constant $c > 0$ such that for any integer n, there exists an n-vertex graph G formed by at least αn edge-disjoint triangles such that any rainbow cycle in G has length at least $c \log n$.*

We use two probabilistic tools, the inequalities of Chernoff and Markov.

Theorem 3.6 (Chernoff). *Let X be a binomial random variable $\mathrm{Bin}(n, p)$. For any $0 < \epsilon < 1$, we have*

$$\mathbb{P}(X \geq (1 + \epsilon)\mathbb{E}X) \leq \exp(-\epsilon^2 \mathbb{E}X/3).$$

Theorem 3.7 (Markov). *Let X be a non-negative random variable. For any $t > 0$, we have*

$$\mathbb{P}(X \geq t) \leq \mathbb{E}X/t.$$

Proof of Theorem 3.5. Let $p := \frac{25\alpha}{n^2}$. Denote by $G^{(3)}(n, p) =: H$ the system of triples in which each element of $\binom{[n]}{3}$ is included independently with probability p. The example proving the theorem will be the set of triangles induced by the triples in $G^{(3)}(n, p)$, with some triples removed.

Here are the details. We have

$$\mathbb{E}|H| = \binom{n}{3}p \geq 4\alpha n.$$

Chernoff's inequality yields

$$\mathbb{P}(|H| \leq 3\alpha n) \leq \mathbb{P}(|H| \leq 0.9 \cdot \mathbb{E}|H|) = o(1).$$

Let $\mathcal{A}$ be the event $\{H : |H| \geq 3\alpha n\}$. Then

(5) $$\mathbb{P}(\mathcal{A}) = 1 - o(1)$$

Let

$$Y := |\{(A_1, A_2) : A_i \in H \text{ for } i = 1, 2, A_1 \neq A_2 \text{ and } |A_1 \cap A_2| = 2\}|$$

be the number of pairs of distinct triples in H that intersect at two vertices.

Then

$$\mathbb{E}Y \leq \binom{n}{3} \cdot 3 \cdot np^2 = o(n),$$

as there are at most $\binom{n}{3}$ ways to choose $A_1 \in \binom{[n]}{3}$, 3 ways to choose a pair Π of vertices in the intersection, and then at most n ways to complete Π to the

triple $A_2 \in \binom{[n]}{3}$, and the probability that both A_1, A_2 are in H is p^2. Then by Markov's inequality, we have

$$\mathbb{P}(Y \geq \alpha n) = o(1).$$

Let $\mathcal{B}$ be the event that $Y \leq \alpha n$. By the above

(6)
$$\mathbb{P}(\mathcal{B}) = 1 - o(1).$$

Given a 3-graph F on $[n]$, let $E^{(2)}(F) := \{e \in \binom{[n]}{2} : e \subseteq A \text{ for some } A \in F\}$. Note that for a fixed $e \in \binom{[n]}{2}$,

$$\mathbb{P}(e \in E^{(2)}(H)) = 1 - \mathbb{P}(e \not\subseteq A \text{ for any } A \in H) = 1 - (1-p)^{n-2} =: q.$$

By Bernoulli's inequality, we have $(1-p)^{n-2} \geq 1 - (n-2)p$. Therefore

(7)
$$q \leq (n-2)p \leq \frac{25\alpha}{n}.$$

Let C be a k-cycle in K_n with edges $e_1, \dots, e_k$. We say that C is *distinguishable* if each $e_i \subseteq A_i$ for some $A_i \in H$ and $e_j \not\subseteq A_i$ if $i \neq j$.

We consider the probability that $C \subseteq E^{(2)}(H)$ and C is distinguishable. We shall bound this probability from above by q^k. Indeed, for $e_1, \dots, e_k$ we define the event $\mathcal{S}_{e_i} = \mathcal{S}_{e_i}(e_1, \dots, e_{i-1})$ as

$$\mathcal{S}_{e_i} := \{\text{there exists } A_i \in \binom{[n]}{3} \text{ with } e_i \subseteq A_i \text{ and } e_j \not\subseteq A_i \text{ for } j < i \text{ such that } A_i \in H\}.$$

Then we have

$$\mathbb{P}(C \subseteq E^{(2)}(H) \text{ is distinguishable})$$

$$\leq \mathbb{P}(\cap_{i=1}^{k} \mathcal{S}_{e_i}) = \prod_{i=1}^{k} \mathbb{P}(\mathcal{S}_{e_i} \mid \cap_{j=1}^{i-1} \mathcal{S}_{e_j}) \leq (1 - (1-p)^{n-2})^k = q^k,$$

where the second to last inequality is because there are at most $n-2$ many $A \in \binom{[n]}{3}$ satisfying that $e_i \subseteq A$ and $e_j \not\subseteq A$ for all $j < i$, thus $\mathbb{P}(\neg \mathcal{S}_{e_i} \mid \cap_{j=1}^{i-1} \mathcal{S}_{e_j}) \geq (1-p)^{n-2}$.

Let X_k be the number of distinguishable cycles of length k in $E^{(2)}(H)$. With a look at (7), we have

$$\mathbb{E}X_k \leq n^k q^k \leq (25\alpha)^k \leq n^{1/2}$$

for $k \leq c(\alpha) \log n$ and $c > 0$ small enough, therefore

$$\sum_{k=3}^{\lfloor c \log n \rfloor} \mathbb{E}X_k = o(n).$$

Then by Markov's inequality, we have

$$\mathbb{P}(\sum_{k=3}^{\lfloor c \log n \rfloor} X_k \geq \alpha n) = o(1)$$

so that

(8)
$$\mathbb{P}(\mathcal{C}) = 1 - o(1)$$

for the event $\mathcal{C} := \{\sum_{k=3}^{\lfloor c \log n \rfloor} X_k \leq \alpha n\}$.

Combining (5), (6), and (8), we take H when $\mathcal{A} \cap \mathcal{B} \cap \mathcal{C}$ holds, which holds with probability $1 - o(1)$. From H, we remove at most one triple in the pairs counted by Y to get H_1 so that the triples in H_1 intersect with each other in at most

one vertex. In particula, each $e \in E^{(2)}(H_1)$ is contained in exactly one $A \in H_1$. Then $\mathcal{A} \cap \mathcal{B}$ implies that

$$|H_1| \geq 3\alpha n - \alpha n \geq 2\alpha n.$$

View each triple in H_1 as a triangle in $E^{(2)}(H_1)$. The above observation confirms that the triangles are edge-disjoint. If there is a rainbow cycle in $E^{(2)}(H_1)$ with edges $e_1, \ldots, e_k$, then $e_i \subseteq A_i \in H_1$, the rainbow property and the fact that there is exactly one triple in H_1 contains an edge in $E^{(2)}(H_1)$ implies that the cycle is distinguishable. For each rainbow cycle of length at most $c \log n$ in H_1, in order to destroy the rainbow cycle, we choose at most one edge e and remove the triple $A \supseteq e$ from H_1 to get H_2. As $E^{(2)}(H_1) \subseteq E^{(2)}(H)$, the event $\mathcal{C}$ implies that we only need to remove at most αn triples. Therefore

$$|H_2| \geq |H_1| - \alpha n \geq \alpha n.$$

Let $G := E^{(2)}(H_2)$. Then G is a graph formed by at least αn edge-disjoint triangles without rainbow cycles of length less than $c \log n$. This completes the proof. $\square$

4. Rainbow girth when $\max |F_i| = 2$

For $\mathcal{F} = (F_1, \ldots, F_m)$ a family of subsets of $E(K_n)$, recall that

$$\psi(\mathcal{F}) = \sum_{1 \leq i \leq m} \frac{1}{|F_i|}.$$

Theorem 4.1. *Let $\mathcal{F} = (F_1, \ldots, F_n)$ be a family of subsets of $E(K_n)$ such that $1 \leq |F_i| \leq 2$. Then $rgirth(\mathcal{F}) \leq \lceil \psi(\mathcal{F}) \rceil$.*

In [6] Conjecture 1.5 was proved when $|F_i| = 2$ for all i. Theorem 4.1 is a generalization to the case in which some of the sets F_i are singleton sets.

The theorem is easily seen to be equivalent to:

Theorem 4.2. *Let $\mathcal{F} = (F_1, \ldots, F_n)$ be a family of subsets of $E(K_n)$ such that $1 \leq |F_i| \leq 2$. Assume p sets are of size 1, and $n - p$ are of size 2. Then $rgirth(\mathcal{F}) \leq \lceil \frac{n+p}{2} \rceil$.*

We will refer to the edges in F_i as colored by color i.

Proof. We may assume that the sets F_i are disjoint, or else there is a rainbow digon (cycle of length 2). The case where all the sets F_i are of size 2 was proved in [6]. Thus we may assume $|F_1| = 1$. Let $F_1 = \{e\}$.

We construct a subgraph H of G recursively as follows. Let $H_0 = \{e\}$. At each step i, H_i is constructed by adding to H_{i-1} a vertex $x_i \notin V(H_{i-1})$ and two edges $x_i a_i, x_i b_i \notin E(H_{i-1})$ such that $a_i, b_i \in V(H_{i-1})$ and $x_i a_i, x_i b_i$ are colored by the same color i. We stop at step $i = t$ when there are no such two edges to add, and we let $H = H_t$.

For two vertices $u, v \in V(G)$ let $\text{dist}_{r,G}(u, v)$ denote the *rainbow distance* of u, v, that is, the minimum length (number of edges) of a rainbow path in G connecting u, v. For a subgraph G' of G let the *rainbow diameter* of G be defined as $rd(G') := \max_{u,v \in V(G')} \text{dist}_{r,G'}(u, v)$. We omit the subscript G in dist_r and it should be clear according to the context.

Claim 4.2.1. $rd(H_i) \leq \frac{i}{2} + 1$, *and if i is even, except at most one pair of vertices $u_i, v_i \in V(H_i)$, for any pair of vertices $u, v \in V(H_i)$, we have $\text{dist}_r(u, v) \leq \frac{i}{2}$.*

Proof. If $i \in \{0, 1\}$ the claim is trivial. We proceed by induction on i.

Suppose first that $i + 1$ is odd. By the induction hypothesis, there exists at most one pair of vertices $u_i, v_i \in V(H_i)$ such that $\text{dist}_r(u_i, v_i) = \frac{i}{2} + 1$ and for any other pair of vertices $u, v \in V(H_i)$, $\text{dist}_r(u, v) \leq \frac{i}{2}$. We have to show that for every $y, z \in V(H_{i+1})$, $\text{dist}_r(y, z) \leq \lfloor \frac{i+1}{2} + 1 \rfloor = \frac{i}{2} + 1$. If $y, z \in V(H_i)$ we are done. Suppose $z = x_{i+1}$. If $y \notin \{u_i, v_i\}$ there is a rainbow path from a_{i+1} to y of length at most $\frac{i}{2}$ and thus there is a rainbow path from x_{i+1} to y of length at most $\frac{i}{2} + 1$. If $y \in \{u_i, v_i\}$, say $y = u_i$, then either $a_{i+1} \neq v_i$ or $b_{i+1} \neq v_i$. In both cases there exists a rainbow path from x_{i+1} to y, through a_{i+1} or b_{i+1} respectively, of length at most $\frac{i}{2} + 1$.

Assume now that $i + 2$ is even. By the induction hypothesis, there exists at most one pair $u_i, v_i \in V(H_i)$ such that $\text{dist}_r(u_i, v_i) = \frac{i}{2} + 1$ and any other pair of vertices in $V(H_i)$ is of rainbow distance at most $\frac{i}{2}$. We have to show that there is at most one pair $u_{i+2}, v_{i+2} \in V(H_{i+2})$ such that $\text{dist}_r(u_{i+2}, v_{i+2}) = \frac{i}{2} + 2$ and any other pair of vertices in $V(H_{i+2})$ is of rainbow distance at most $\frac{i}{2} + 1$.

We split into two cases.

Case 1. $x_{i+1} \notin \{a_{i+2}, b_{i+2}\}$.

Choose $u_{i+2} = x_{i+1}, v_{i+2} = x_{i+2}$. We claim that $\text{dist}_r(u_{i+2}, v_{i+2}) \leq \frac{i}{2} + 2$. Indeed, by the induction hypothesis we can choose a vertex $u \in \{a_{i+1}, b_{i+1}\}$ and $v \in \{a_{i+2}, b_{i+2}\}$ such that $\text{dist}_r(u, v) \leq \frac{i}{2}$ by the fact that $a_{i+1}, b_{i+1}, a_{i+2}, b_{i+2} \in V(H_i)$ and $\{u, v\} \neq \{u_i, v_i\}$, and then adding the edges $x_{i+1}u, x_{i+2}v$ we get a rainbow path between x_{i+1}, x_{i+2} of length at most $\frac{i}{2} + 2$. Let $u, v \in V(H_{i+2})$ such that $\{u, v\} \neq \{x_{i+1}, x_{i+2}\}$. Since $a_{i+1}, b_{i+1}, a_{i+2}, b_{i+2} \in V(H_i)$, we have $\text{dist}_r(u, v) \leq \frac{i}{2} + 1$ like in the odd case.

Case 2. $x_{i+1} = a_{i+2}$.

In this case, either $b_{i+2} \neq u_i$ or $b_{i+2} \neq v_i$. Assume WLOG $b_{i+2} \neq v_i$. Choose $u_{i+2} = x_{i+2}, v_{i+2} = v_i$. Then $\text{dist}_r(u_{i+2}, v_{i+2}) \leq \lfloor \frac{i+1}{2} + 1 \rfloor + 1 = \frac{i}{2} + 2$, and like before, $\text{dist}_r(u, v) \leq \frac{i}{2} + 1$ for any other pair of vertices u, v. $\square$

Returning to the proof of the theorem, we proceed by induction on n. Contract H into a single vertex h to obtain a new graph G' (G' may have loops). Note that $n' := |V(G')| = n - t - 1$, the number of colors is $n - t - 1 = n'$ and the number of colors of size 1 is $p' = p - 1$. By induction there exists a rainbow cycle C in G' of size at most $\lceil \frac{n'+p'}{2} \rceil = \lceil \frac{n-t+p-2}{2} \rceil = \lceil \frac{n-t+p}{2} \rceil - 1$.

If C does not use the vertex h, we are done. Otherwise, uncontracting h, C either remains a cycle (possibly containing a vertex in h) - in this case we are done; or it may become a path in G, with end vertices $u \neq v \in V(H)$. By Claim 4.2.1, there is a rainbow path P in H connecting u and v of size at most $\frac{t}{2} + 1$. Note that P uses colors not appearing in C. Thus $P + C$ is a rainbow cycle in G of size at most $\lfloor \lceil \frac{n-t+p}{2} \rceil - 1 + \frac{t}{2} + 1 \rfloor = \lceil \frac{n+p}{2} \rceil$. This completes the proof of the theorem. $\square$

Corollary 4.3. *Let D be an n-vertex sink-less digraph. Assume p vertices have out-degree 1. Then $girth(D) \leq \lceil \frac{n+p}{2} \rceil$.*

Remark 4.4. In [12] (Theorem 1) a slightly weaker result was proved: $dgirth(D) \leq \lceil \frac{n+p+1}{2} \rceil$.

Proof of Corollary 4.3. For each vertex v of D that has out-degree more than 2, we remove some arbitrary edges to make v have out-degree exactly 2. Then in

the resulting digraph D', there are p vertices of out-degree 1 and $n - p$ vertices of out-degree 2, and we have $girth(D) \leq girth(D')$. Therefore it is enough to prove that $girth(D') \leq \lceil \frac{n+p}{2} \rceil$. While by the construction to explain why Conjecture 1.5 generalizes CHC in Section 3, we reduce the problem into rainbow undirected version with p stars of size 1 and $n - p$ stars of size 2. Therefore we complete the proof by applying Theorem 4.2. $\qquad\square$

Corollary 4.5. *For a family $\mathcal{F} = (F_1, \ldots, F_n)$ of subsets of $E(K_n)$ satisfying $1 \leq |F_i| \leq 2$ for all $1 \leq i \leq n$, we have $rgirth(\mathcal{F}) \leq 2 \sum_{1 \leq i \leq n} \frac{1}{|F_i|+1}$.*

Proof. Applying Theorem 4.2, we have $rgirth(\mathcal{F}) \leq \lceil \psi(\mathcal{F}) \rceil = \lceil \frac{n+p}{2} \rceil$, where p is the number of sets in $\mathcal{F}$ of size 1. Note that $\lceil \frac{n+p}{2} \rceil \leq \frac{n+p}{2} + \frac{1}{2}$, which is at most $2(\frac{p}{2} + \frac{n-p}{3}) = 2 \sum_{1 \leq i \leq n} \frac{1}{|F_i|+1}$ when $p \leq n - 3$. Furthermore, for $p = n - 2$ or n, $\lceil \psi(\mathcal{F}) \rceil = \lceil \frac{n+p}{2} \rceil = \psi(\mathcal{F}) \leq 2 \sum_{1 \leq i \leq n} \frac{1}{|F_i|+1}$. And in the remaining case $p = n - 1$, we have $rgirth(\mathcal{F}) \leq n - 1 \leq \psi(\mathcal{F}) \leq 2 \sum_{1 \leq i \leq n} \frac{1}{|F_i|+1}$. $\qquad\square$

References

[1] R. Aharoni, M. DeVos, and R. Holzman. Rainbow triangles and the Caccetta-Häggkvist conjecture. *J. Graph Theory* **92** (2019), 347–360.

[2] R. Aharoni and H. Guo. Rainbow cycles for families of matchings. *Israel Journal of Mathematics*, to appear. `arXiv:2110.14332`.

[3] B. Bollobás and E. Szemerédi. Girth of sparse graphs. *J. Graph Theory* **39** (2002), 194–200.

[4] L. Caccetta and R. Häggkvist. On minimal digraphs with given girth. *Congress. Numer.* **21** (1978), 181–187.

[5] V. Chvátal and E. Szemerédi. Short cycles in directed graphs. *J. Combin. Theory Ser. B* **35** (1983), 323–327.

[6] 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.

[7] I. Goorevitch and R. Holzman. Rainbow triangles in families of triangles. In preparation.

[8] J. Hladký, D. Král', and S. Norin. Counting flags in triangle-free digraphs. *Electron. Notes Discrete Math.* **34** (2009), 621–625.

[9] P. Hompe. The girth of digraphs and concatenating bipartite graphs. *Senior Theses.*

[10] P. Hompe, Z. Qu, and S. Spirkl. Improved bounds for the triangle case of Aharoni's rainbow generalization of the Caccetta-Häggkvist conjecture. Preprint (2022). `arXiv:2206.10733`.

[11] P. Hompe and S. Spirkl. Further approximations for Aharoni's rainbow generalization of the Caccetta-Häggkvist conjecture. *Electron. J. Combin.* **29** (2022), #P1.55.

[12] J. Shen. On the girth of digraphs. *Discrete Math.* **211** (2000), 167–181.

[13] J. Shen. On the Caccetta-Häggkvist conjecture. *Graphs Comb.* **18** (2002), 645–654.

[14] B. Sullivan. A summary of results and problems related to the Caccetta-Häggkvist conjecture. Preprint (2006). `arXiv:math/0605646` .