Rainbow cycles for families of matchings

Ron Aharoni* and He Guo†

October 27, 2021; Revised on November 10, 2023

Abstract

Given a graph G and a coloring of its edges, a subgraph of G is called *rainbow* if its edges have distinct colors. The *rainbow girth* of an edge coloring of G is the minimum length of a rainbow cycle in G. A generalization of the famous Caccetta-Häggkvist conjecture, proposed by the first author, is that if in an coloring of the edge set of an n-vertex graph by n colors, in which each color class is of size k, the rainbow girth is at most $\lceil \frac{n}{k} \rceil$. In the known examples for sharpness of this conjecture the color classes are stars, suggesting that when the color classes are matchings, the result may be improved. We show that the rainbow girth of n matchings of size at least 2 is $O(\log n)$.

1 Introduction

The *girth* $g(G)$ of a graph G is the minimal length of a cycle in it. Given a graph G and a (not necessarily proper) coloring of its edges, a subgraph of G is called *rainbow* if its edges have distinct colors. The *rainbow girth* $rg(G)$ of G (actually, of its edge-coloring) is the minimum length of a rainbow cycle. All the above definitions apply to both the directed and undirected cases, where in the directed case the cycles are assumed to be directed.

The famous Caccetta-Häggkvist conjecture [5] (below - CHC) is that any digraph G on n vertices satisfies $g(G) \le \lceil \frac{n}{\delta^+(G)} \rceil$, where $\delta^+(G)$ is the minimal out-degree of a vertex. There has been constant progress [6, 8, 9, 10] on the problem. In particular it has been shown that

(a) The CHC is true if $n \ge 2\delta^+(G)^2 - 3\delta^+(G) + 1$ [12], and

(b) $g(G) \le n/\delta^+(G) + 73$ for all G [13].

In [2] a possible generalization of CHC was suggested.

Conjecture 1.1. *Let G be an undirected n-vertex graph. For any edge coloring of $E(G)$ with n colors such that each color class has size at least k, we have $rg(G) \le \lceil n/k \rceil$.*

Devos et. al. [7] proved this conjecture for $r = 2$. In [1] a stronger version of the conjecture was proved when all sets are of size 1 or 2.

For a directed edge $e = uv$ let $n(e)$ be the undirected pair $\{u, v\}$. To see that Conjecture 1.1 is a generalization of CHC, given a directed graph G, for every vertex u let $S(u) = \{n(uv) \mid uv \in E(G)\}$ be the star of edges leaving u, with their direction removed. We claim that an undirected rainbow cycle $v_1 v_2 \dots v_k$ for the sets $S(v)$ gives rise to a directed cycle in G. Otherwise there exists a vertex v_i such that $\{v_i, v_{i+1}\} = n(v_i v_{i+1})$ and $\{v_i, v_{i-1}\} = n(v_i v_{i-1})$. But this contradicts the fact that only one edge is chosen from $S(v_i)$ for participation in the rainbow cycle. Thus the CHC is equivalent to the case of Conjecture 1.1 in which the color classes are n stars, each centered at a different vertex. Hence the sharpness of CHC implies that of Conjecture 1.1.

The standard example showing that the former is the case is the graph on $\{1, 2, \dots, n\}$ with edges $\{i, i+1\}, \{i, i+2\}, \dots, \{i, i+k\}$ for $i = 1, 2, \dots, n$ (indices taken modulo n). But this example is not unique. Bondy [4] noted that if G and H are two graphs witnessing the sharpness of CHC, then blowing each vertex

*Faculty of Mathematics, Technion, Haifa 32000, Israel and MIPT. E-mail: ra@technion.ac.il. Research supported by the Israel Science Foundation (ISF) grant no. 2023464 and the Discount Bank Chair at the Technion, and the European Union's Horizon 2020 research and innovation programme under the Marie Skldowska-Curie grant agreement no. 823748.

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

of G by a copy of H yields another such example (the blow-up is called also the *lexicographic product* of G and H). In [11] more examples are given (accompanied by a conjecture that these exhaust all possible cases of equality in the conjecture).

By the above argument, every example witnessing the sharpness of CHC gives rise to an example witnessing the sharpness of Conjecture 1.1. In fact, all known extreme examples for this conjecture are obtained this way, in particular they have stars as the sets of edges. This suggests that in the antipodal case to that of stars, when the sets of edges are matchings, the conjecture can be strengthened. ("Antipodal" is with respect to the covering number, which is 1 in a star, and the number of edges in a matching.) Indeed, a simple observation is that for the first open case of CHC, that of $\delta^+ = \frac{n}{3}$ (in which a directed triangle is conjectured to exist) the rainbow undirected version is trivial when the sets of edges are matchings. In this case, it is enough that the arithmetic mean of the sizes of the sets is larger than $\frac{n}{4}$, because then by Mantel's theorem there exists a triangle contained in the union of the sets, and if the sets are matchings then a triangle is necessarily rainbow.

2 Rainbow cycles for matchings

Our main result is a corroboration of the intuition that sets of matchings have small rainbow girth.

Theorem 2.1. *There exists a constant C such that for any n-vertex graph G and edge coloring of G with n colors, if each color class is a matching of size 2, then the rainbow girth of G is at most $C \log n$.*

Remark 2.2. *The assumption that G is a graph and not a multigraph breeds no loss of generality, since a double edge is a rainbow digon, meaning that the rainbow girth is 2.*

A key ingredient in the proof is a result by Bollobás and Szemerédi [3] on the girth of sparse graphs.

Theorem 2.3. *For $n \geq 4$ and $k \geq 2$, every n-vertex graph with $n + k$ edges has girth at most*

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

The logarithms are to the base 2. Theorem 2.1 will follow from this result, and the following:

Theorem 2.4. *There exist universal $c, \delta > 0$, such that for any large enough n, given an n-vertex graph G and an edge coloring of G with n colors such that each color class is a matching of size 2, there exists a subset S of $V(G)$ of size at most cn containing the edges of a rainbow set of edges of size at least $(c + \delta)n$.*

Note that the last condition entails $c + \delta \leq 1$. Once this is proved, Theorem 2.1 follows by applying Theorem 2.3 with $k = \delta n$.

The idea of proving Theorem 2.4 is that we take a random subset S of $V(G)$ and consider the induced subgraph $G[S]$. The crux of the argument is that the expected number of vertices $\mathbb{E}|S|$ is polynomially in n less than the expected number of colors of the edges in $G[S]$. Furthermore, these two random numbers are concentrated around their expectations, which follows from two well-known concentration inequalities.

Theorem 2.5 (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 2.6 (Chebyshev's inequality)**.** *Let X be a random variable. For any $\epsilon > 0$, we have*

$$\mathbb{P}(|X - \mathbb{E}X| \geq \epsilon\mathbb{E}X) \leq \operatorname{Var} X/(\epsilon\mathbb{E}X)^2,$$

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

Proof of Theorem 2.4. Denote the i-th color class (which, by our assumption, consists of two disjoint edges) by M_i. Our assumption that G is a graph and not a multigraph implies that the matchings M_i are disjoint.

A vertex v of G is called *heavy* if there are at least $\epsilon^2 n/10^6$ rainbow edges incident to it, where $\epsilon > 0$ is a small constant to be determined later. Let D be the set of heavy vertices of G. Then we have

$$|D| \leq \frac{2 \cdot 2 \cdot n}{(\epsilon^2 n/10^6)} \leq \frac{10^7}{\epsilon^2}. \tag{1}$$

Let $S = D \cup Z$ be a random vertex subset of $V(G)$, where each vertex of $V(G) \setminus D$ is included in Z independently with probability p, for some constant p to be determined later. Then

$$\mathbb{E}|Z| = (n - |D|) \cdot p \quad \text{and} \quad \mathbb{E}|S| = \mathbb{E}|Z| + |D| \sim np. \tag{2}$$

For $1 \le i \le n$ let X_i be the indicator random variable that an edge of color i is contained in S, i.e., $X_i := \mathbb{1}_{\{\text{an edge of color } i \text{ is contained in } S\}}$. Since each vertex is included with probability at least p and X_i is an increasing event with respect to the probability that a vertex is included in S, by inclusion-exclusion we have

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

Let

$$X := \sum_{i=1}^{n} X_i. \tag{3}$$

We have

$$\mathbb{E}X \ge n(2p^2 - p^4). \tag{4}$$

By Theorem 2.5, for fixed $0 < p < 1$ and $\epsilon > 0$, when n is large enough we have

$$\mathbb{P}\Big(|S| \ge (1+\epsilon)np\Big) \le \mathbb{P}\Big(|Z| \ge (1+\epsilon/2)\mathbb{E}|Z|\Big) \le \exp(-\Omega(n)). \tag{5}$$

So, with probability tending to 1 as n tends to infinity,

$$|S| \le (1+\epsilon)np. \tag{6}$$

Writing $p - (2p^2 - p^4) = p(p-1)(p^2 + p - 1)$, we see that for $\frac{-1+\sqrt{5}}{2} < p < 1$, we have $p < 2p^2 - p^4$, yielding the separation between $\mathbb{E}X$ and $\mathbb{E}|S|$, needed for the application of Theorem 2.3.

Claim 2.7. *There exist constants $p \in (\frac{-1+\sqrt{5}}{2}, 1)$ and $\epsilon > 0$ such that*

$$(1-\epsilon)(2p^2 - p^4) - (1+\epsilon)p \ge [(2p^2 - p^4) - p]/3 > 0, \tag{7}$$

and with probability at least 0.9 for all large n,

$$X \ge (1-\epsilon)n(2p^2 - p^4). \tag{8}$$

We fix $0.618 \approx \frac{-1+\sqrt{5}}{2} < p < 1$ and $\epsilon(p) > 0$ satisfying (7).

To prove (8), we shall apply Chebyshev's inequality. For this purpose we have to estimate $\operatorname{Var} X$. With a look at (3), we have

$$\operatorname{Var} X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \sum_{i,j} (\mathbb{E}X_iX_j - \mathbb{E}X_i\mathbb{E}X_j). \tag{9}$$

Note that if the edges in the color classes i, j are vertex-disjoint, then X_i and X_j are independent and $\mathbb{E}X_iX_j - \mathbb{E}X_i\mathbb{E}X_j = 0$.

Since the matchings M_j are disjoint, for every $i \in [n]$ at most $6 = \binom{4}{2}$ matchings M_j can have an edge contained in $\bigcup M_i$. This means that there exist at most $2 \cdot 6n$ pairs (M_i, M_j) such that an edge from M_j is contained in $\bigcup M_i$, or vice versa. Thus the contribution of such pairs to $\operatorname{Var} X$ is at most $O(n)$.

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

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

II. two vertex-disjoint 2-paths.

Examine Case I. Let $M_i = \{a, b\}$, where $a = xy$, $b = uv$, and let $M_j = \{c, d\}$, where $c = xz$ and $d = st$.

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

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

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

In case II, let $M_i = \{a, b\}$, where $a = xy,\ b = uv$, and let $M_j = \{c, d\}$, where $c = xz$ and $d = ut$.

If both x and u are heavy vertices, then similarly to case I, we have $\mathbb{E}X_iX_j - \mathbb{E}X_i\mathbb{E}X_j = 0$. Otherwise the contribution to $\operatorname{Var} X$ for such (M_i, M_j) is at most $4 \cdot n \cdot \frac{\epsilon^2 n}{10^6}$.

Summing, we have

$$\operatorname{Var} X = O(n) + 2 \cdot 4 \cdot n \cdot \frac{\epsilon^2 n}{10^6} \leq \frac{\epsilon^2 n^2}{10^5}.$$

Since $1/2 \leq p < 1$, we have $2p^2 - p^4 \geq 7/16$. Applying Chebyshev's inequality, we have

$$\begin{aligned}
\mathbb{P}(X \leq (1-\epsilon)n \cdot (2p^2 - p^4)) &\leq \mathbb{P}(X \leq \mathbb{E}X - \epsilon n \cdot (2p^2 - p^4)) \\
&\leq \frac{\operatorname{Var} X}{(\epsilon n \cdot (2p^2 - p^4))^2} \leq \frac{\epsilon^2 n^2}{10^5 \epsilon^2 n^2 (7/16)^2} \leq 1/10.
\end{aligned}$$

This proves Claim 2.7.

Combining (8) with (5), we have that with probability at least 1/2 for all large n,

$$|S| \leq (1+\epsilon)np < (1-\epsilon)n(2p^2 - p^4) \leq X.$$

Theorem 2.4 now follows from (7), upon taking $c := (1+\epsilon)p$ and $\delta := [(2p^2 - p^4) - p]/3$. □

2.1 Fewer than n sets

In the original CHC, each set of edges is a star associated with a vertex, hence it was natural that there are n sets. In the rainbow undirected case there is no natural choice of the number of sets. Indeed, the main theorem is valid also with fewer than n sets.

Theorem 2.8. *For any constant $\alpha > \frac{3\sqrt{6}}{8}$, there exists a constant C such that for any n-vertex graph G and edge coloring of G with αn colors, if each color class is a matching of size 2, then the rainbow girth of G is at most $C \log n$.*

Proof. For the argument in the proof of Theorem 2.4 to work with αn colors, we need to have $p < \alpha(2p^2 - p^4)$ for some p (which would imply separation between $\mathbb{E}|S| \sim pn$ and $\mathbb{E}X \geq \alpha n(2p^2 - p^4)$).

Thus we need to find a minimal $0 < \alpha_0 < 1$ such that $p = \alpha_0(2p^2 - p^4)$ for some $p \in (0, 1)$. This will happen when the two curves $y(p) = p$ and $y(p) = \alpha_0(2p^2 - p^4)$ are tangent, namely $(\alpha_0(2p^2 - p^4))' = p' = 1$. The above two constraints and $\alpha_0 p \neq 0$ imply that $\alpha_0 = \frac{3\sqrt{6}}{8}$ and the only feasible p is $\frac{\sqrt{6}}{3}$. □

It would be interesting to find the optimal value of α in Theorem 2.8. For example, it should be at least 1/2: for even n and $G = C_n$, assume the edges of the cycle in order are $e_1, e_2, \ldots, e_n$. If we color the edges e_i and $e_{i+n/2}$ by color i, then there is no rainbow cycle.
Acknowledgements. We thank the referees for their helpful suggestions. We also thank Xiaozheng Chen, Yuhui Cheng, and Ruonan Li for identifying omissions in earlier versions of this work.

References

[1] R. Aharoni, E. Berger, M. Chudnovski, H. Guo, and S. Zerbib. Non-uniform degrees and rainbow versions of the Cacceta-Häggkivst 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] B. Bollobás and E. Szemerédi. Girth of sparse graphs. *J. Graph Theory* **39** (2002), 194–200.

[4] J.A. Bondy. Counting subgraphs: A new approach to the Caccetta-Häggkvist conjecture. *Discrete Math.* **165/166** (1997), 71–80.

[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] T. Nishimura. Short Cycles in Digraphs. *Discrete Math.* **38** (1988), 295–298.

[11] A.A. Razborov. On the Caccetta-Häggkvist Conjecture with Forbidden Subgraphs. *J. Graph Theory* **74** (2013), 236–248.

[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.