

Rainbow paths and large rainbow matchings

Ron Aharoni*

Department of Mathematics
Technion and MIPT
Haifa, Israel
ra@technion.ac.il

Eli Berger

Department of Mathematics
University of Haifa
Haifa, Israel
berger@math.haifa.ac.il

Maria Chudnovsky[†]

Department of Mathematics
Princeton University
Princeton, NJ, U.S.A.
mchudnov@math.princeton.edu

Shira Zerbib[‡]

Department of Mathematics
Iowa State University
Ames, IA, U.S.A.
zerbib@iastate.edu

Submitted: Jan 16, 2021; Accepted: Oct 7, 2021; Published: Jan 28, 2022
© The authors. Released under the CC BY-ND license (International 4.0).

Abstract

A conjecture of the first two authors is that n matchings of size n in any graph have a rainbow matching of size $n - 1$. We prove a lower bound of $\frac{2}{3}n - 1$, improving on the trivial $\frac{1}{2}n$, and an analogous result for hypergraphs. For $\{C_3, C_5\}$ -free graphs and for disjoint matchings we obtain a lower bound of $\frac{3n}{4} - O(1)$. We also discuss a conjecture on rainbow alternating paths, that if true would yield a lower bound of $n - \sqrt{2n}$. We prove the non-alternating (ordinary paths) version of this conjecture.

Mathematics Subject Classifications: 05C70, 05D15

*We acknowledge the financial support from the Ministry of Educational and Science of the Russian Federation in the framework of MegaGrant no. 075-15-2019-1926 when the first author worked on Section 3 of the paper. The research of R. Aharoni was supported in part by the Israel Science Foundation (ISF) grant no. 2023464 and the Discount Bank Chair at the Technion. This paper is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 823748.

[†]Supported by NSF Grant DMS-1763817.

[‡]Supported by NSF grant DMS-1953929. The authors were supported by US-Israel Binational Science Foundation (BSF) grant no. 2016077.

1 Introduction

Using a common terminology, a *family* is a multiset, namely elements may repeat. While the notation for sets uses curly brackets, families are denoted using ordinary parentheses, so in the notation $(a_i \mid i \in I)$ equality may hold between some a_i s.

Given a family $\mathcal{S} = (S_1, \dots, S_n)$ of sets, we write $\|\mathcal{S}\|$ for $\sum_{i \leq n} |S_i|$.

For such \mathcal{S} , an \mathcal{S} -rainbow set (or just “rainbow set” if the identity of \mathcal{S} is clear from the context) is the range of a partial choice function. That is, it is a set $\{x_{i_j} \mid 1 \leq i_1 < i_2 < \dots < i_m \leq n\}$, where $x_{i_j} \in S_{i_j}$ ($1 \leq j \leq m$).

For integers a, b, c, r we write $(a, b) \rightarrow_r c$ if every family of a matchings in an r -uniform hypergraph, each of size b , has a rainbow matching of size c . If we demand this condition only for r -partite hypergraphs, we write $(a, b) \rightarrow_r^P c$. If $r = 2$ we omit its mention and write just $(a, b) \rightarrow c$ and $(a, b) \rightarrow^P c$, respectively.

A famous conjecture of Ryser-Bruualdi-Stein [8, 14] is that in an $n \times n$ Latin square there exists a transversal (sub-permutation submatrix with distinct symbols) of size $n - 1$, and that for n odd there exists a transversal of size n . In [1] the first part of this conjecture was strengthened to:

Conjecture 1 (Aharoni-Berger [1]). $(n, n) \rightarrow n - 1$.

In [1] this was conjectured only for bipartite graphs, but there is no counterexample known also for general graphs. Here are some known facts:

(F1) $(n, \frac{3}{2}n) \rightarrow^P n$ [5].

(F2) $(n, \frac{3}{2}n + o(n)) \rightarrow n$ [9].

(F3) $(2n - 1, n) \rightarrow^P n$ (The Latin rectangle case was proved in [11], the general case in [1]).

(F4) $2n - 1$ matchings in a bipartite graph, of respective sizes $1, 2, 3, \dots, n - 1, n, n, \dots, n$ (n is repeated n times) have a rainbow matching of size n . [6].

(F5) Let $m \geq n$. Any $2n - 1$ matchings of size m have a matching of size m representing n of them [1].

(F6) $(n, n) \rightarrow^P n - \sqrt{n}$ [15].

(F7) $(\lfloor \frac{k+2}{k+1}n \rfloor - (k + 1), n) \rightarrow^P n - k$ [7].

(F8) $(3n - 3, n) \rightarrow n$ [3].

(F9) $2n$ matchings of size n have a rainbow set with a fractional matching of size n [4].

(F10) $(n, n + o(n)) \rightarrow^P n$ [13].

(F11) n disjoint matchings of size $n + o(n)$ in any graph have a rainbow matching of size n . [12].

It was conjectured in [7] and in [1] that $(2n, n) \rightarrow n$. This would follow, in particular, from Conjecture 1.

In r -uniform hypergraphs a greedy argument yields $(n, n) \rightarrow_r \lceil \frac{n}{r} \rceil$. We shall improve this to $(n, n) \rightarrow_r \frac{n}{r-\frac{1}{2}} + O(\frac{1}{r})$. In the case $r = 2$, the explicit calculation yields $(n, n) \rightarrow \frac{2}{3}n - \frac{4}{3}$. This improves upon the result $(n, n) \rightarrow \frac{2}{3}n - o(n)$, that follows from (F2). When the graph is $\{C_3, C_5\}$ -free or the given matchings are disjoint, the bound can be improved to $\frac{3}{4}n - O(1)$:

Theorem 2.

1. $(n, n) \rightarrow \frac{2}{3}n - \frac{4}{3}$.
2. $(n, n) \rightarrow_r \frac{n - \frac{1}{2} - \frac{3}{4r-6}}{r - \frac{1}{2}}$.
3. If $\mathcal{F} = (F_1, \dots, F_n)$ is a family of matchings of size n in a $\{C_3, C_5\}$ -free graph, then there exists a rainbow matching of size at least $\frac{3}{4}n - \frac{9}{4}$.
4. If $\mathcal{F} = (F_1, \dots, F_n)$ is a family of disjoint matchings of size n in a graph, then there exists a rainbow matching of size at least $\frac{3}{4}n - \frac{9}{2}$.

Theorem 2 is proved in Section 2.

A main tool used in the study of rainbow matchings is alternating paths. For example, among facts (F1-F10), only for two, (F4) and (F8), there is no alternating paths proof known (the existing proofs use topology). Often the proof goes through results on rainbow directed paths. To state the latter, we need some definitions.

Though the graphs we are considering are all undirected, all paths will be assumed below to be directed. The initial and terminal vertices of a path P are denoted by $in(P)$ and $ter(P)$, respectively.

Definition 3. Given two sets S, T of vertices, a directed path P is called an $S - T$ -path if $in(P) \in S$, $ter(P) \in T$, and $V(P) \cap (S \cup T) = \{in(P), ter(P)\}$.

Definition 4.

1. Let \mathcal{H} be a family of (not necessarily distinct or disjoint) sets of paths. A path is called *strongly rainbow* if each of its edges is chosen from a path belonging to a different $F \in \mathcal{H}$.
2. Let M be a matching, and let \mathcal{H} be a family of sets of M -alternating paths. An M -alternating path is called *strongly rainbow* if each of its non- M edges is chosen from a path belonging to a different $F \in \mathcal{H}$.

This is “double rainbow-ness”: a rainbow set of edges, one from each path in a rainbow set of paths.

- Let $\mathcal{F} = (P_1, \dots, P_m)$ be an ordered family of directed paths. A directed path P is called *rainbow-monotone* (with respect to \mathcal{F}) if its edges, as ordered by P , are (e_1, \dots, e_m) , satisfy $e_i \in E(P_{j_i})$, where $j_1 < j_2 < \dots < j_m$.

Similar definitions apply to M -alternating paths, where M is a matching, and the edges referred to in the definitions are the non- M edges.

Here is a list of known facts about rainbow paths. Part (3) was proved in [2], Part (4) was proved in [5], and the others are taken from [6]. We use the following notation: G is an undirected graph, M is a matching in G , S, T are subsets of $V(G)$, and $Y = V(G) \setminus (S \cup T)$. In each of these facts we add something fact that was in fact included in the original proof, without explicit mention - rainbow monotonicity.

Theorem 5.

- Let F be a sequence of (not necessarily distinct) directed $S - T$ paths. If $S \cap T = \emptyset$ and $|F| > |Y|$ then there exists a directed $S - T$ rainbow-monotone path.
- (Corollary of (1)): Let G be bipartite. Let F be sequence of augmenting M -alternating paths with $|F| > |M|$. Then there exists a rainbow-monotone augmenting M -alternating path.
- nothing In a general graph, if F is a family of augmenting M -alternating paths with $|F| > 2|M|$ then there exists a rainbow augmenting M -alternating path. (No monotonicity claim in this case).
- (Strengthening of (1)): Let \mathcal{H} be a family of sets of disjoint directed $S - T$ paths. If $S \cap T = \emptyset$ and $\|\mathcal{H}\| > |Y|$ then there exists a directed strongly rainbow $S - T$ path. (Recall, $\|\mathcal{H}\| = \sum_{H \in \mathcal{H}} |H|$.)
The proof in [6] yields a monotonicity, stronger version:
- For S, T be as above, any sequence of more than $|Y|$ $S - T$ paths has a rainbow-monotone $S - T$ path.
- (Corollary of (4)): If G is bipartite, M is a matching, \mathcal{H} is a family of sets of sets of disjoint augmenting M -alternating paths and $\|\mathcal{H}\| > |M|$, then there exists a strongly rainbow augmenting alternating path. Again, this is a corollary of a monotonicity version.

All these are sharp. For example, the sharpness of (4) is shown by the family \mathcal{H} consisting of the path $sv_1v_2 \dots v_mt$ repeated m times (here $S = \{s\}, T = \{t\}$).

The proof of fact (F5) uses strongly rainbow alternating paths. Extending that proof to general graphs will require proving the following:

Conjecture 6. Let G be a graph and let M be a matching in G . Let \mathcal{H} be a family of sets of disjoint augmenting M -alternating paths in G . If $\|\mathcal{H}\| > 2|M|$ then there exists a strongly rainbow augmenting M -alternating path.

Let us show how this conjecture implies $(n, n) \rightarrow n - \sqrt{2n}$. Let F_1, \dots, F_n be matchings of size n in a general graph. Let M be a maximal size rainbow matching. We claim that $|M| \geq n - \sqrt{2n}$. Assume for contradiction that $|M| < n - \sqrt{2n}$. For every $i \in I$ we have $|F_i| - |M| > \sqrt{2n}$, implying that $M \Delta F_i$ contains a set H_i of $\sqrt{2n}$ M -augmenting alternating paths. Let $\mathcal{H} = (H_i, i \in I)$. Then $\|\mathcal{H}\| > 2n > 2|M|$, and by the conjecture there exists a strongly rainbow M -augmenting path P . Then $M \Delta P$ is a strongly rainbow matching larger than $|M|$, a contradiction.

Our second main theorem is the ordinary (non-alternating) path version of Conjecture 6 – a possible first step.

Theorem 7. *Let the vertex set of an undirected graph be partitioned into two sets, S, Y . Let \mathcal{H} be a family of sets of paths, each consisting of disjoint $S - S$ paths. If $\|\mathcal{H}\| > 2|Y|$ then there exists a strongly rainbow $S - S$ path.*

This is sharp, as shown by the following construction. Let $S = \{u, v\}$, and let $Y = \{y_1, \dots, y_m\}$. Let \mathcal{H} consist of $2m$ families $\mathcal{F}_1, \dots, \mathcal{F}_{2m}$, each \mathcal{F}_i consisting of a single path F_i , where $F_i = uy_1y_2 \dots, y_mv$ for $i \leq m$, and $F_i = vy_my_{m-1} \dots, y_1v$ for $m < i \leq 2m$.

We shall prove a monotonicity version:

Theorem 8. *Let S, Y be as above, and let \mathcal{F} be a sequence of $2|Y| + 1$ $S - S$ directed paths. Then there exists a rainbow-monotone $S - S$ directed path.*

Theorem 8 implies Theorem 7. Given a system of sets of disjoint directed paths $\mathcal{H} = (H_1, \dots, H_m)$, order $\bigcup \mathcal{H}$ so that all paths in H_i appear before all paths in H_j whenever $i < j$. A rainbow-monotone path is then clearly strongly rainbow.

In Section 3 we shall give two proofs for Theorem 8. They are quite different, and may point at two possible proof strategies in the conjectured alternating paths case.

Remark 9. Since the submission of this paper, the following beautiful results have been proved by Correia, Pokrovskiy and Sudakov [10]:

- (a) $(n, n) \rightarrow n - o(n)$,
- (b) $(n, n + o(n)) \rightarrow n$.

Part (b) is easily seen to imply part (a). Interestingly, the authors prove (a), and give a general probabilistic construction showing that (a) implies (b).

2 Proof of Theorem 2

2.1 Definitions and lemmas

Definition 10. Two sets of edges, A, B are *orthogonal* if $|a \cap b| = 1$ for every $a \in A, b \in B$.

The following lemma is (up to niceties) a well-known fact, usually expressed as “there do not exist n mutually orthogonal Latin squares of order n ”.

Lemma 11. Let M_0, M_1, \dots, M_t be mutually orthogonal matchings in an r -uniform hypergraph, where $M_0 = \{h\}$ and $|M_i| = r$ for $1 \leq i \leq t$. Then $t \leq r$.

Proof. Let $e \in M_1$ and let $u \in e \setminus h$. Since the edges in M_i meet e at distinct vertices, each M_i , $2 \leq i \leq t$, contains an edge e_i meeting h at u . Orthogonality implies that the edges e_i , $2 \leq i \leq t$, meet h at distinct vertices in $h \setminus e$, and hence their number $t - 1$ is at most $|h \setminus e| = r - 1$. \square

The proofs of all four parts of Theorem 2 start the same way. Let $\mathcal{F} = (F_1, \dots, F_n)$ be a collection of matchings of size n in an r -uniform hypergraph (in the last three parts of the theorem $r = 2$). Let R be a rainbow matching of maximal size, say q . Let \mathcal{G} be the collection of matchings in \mathcal{F} not represented in R .

Given $G \in \mathcal{G}$ and $e \in R$ we say that e is G -wasteful if either (1) e meets at most $r - 1$ edges from G (meaning e is not using its full hitting potential with respect to G), or (2) there exists an edge $g \in G$ meeting both e and another edge $f \in R$ (meaning e is not essential for hitting g). For an edge $e \in R$ let $T(e)$ be the set of matchings $G \in \mathcal{G}$ for which e is non- G -wasteful. By the definition of “wastefulness” the following holds:

(*) $G \in T(e)$ if and only if there exist r edges in G intersecting e and not intersecting any other edge of R .

For a matching $G \in \mathcal{G}$ let T^G be the set of edges $e \in R$ such that $G \in T(e)$. Write $t(G) = |T^G|$.

Lemma 12. $|T(e)| \leq r$ for every $e \in R$.

Proof. Suppose there are $r + 1$ matchings in $T(e)$. For each $G \in T(e)$ let G_e be the set of edges of G meeting e . Then G_e is a matching of size r orthogonal to e . By Lemma 11 there exist $A, B \in T(e)$ such that A_e, B_e are not cross-intersecting, meaning that there are $a \in A_e$ and $b \in B_e$, such that $a \cap b = \emptyset$. Since $A, B \in T(e)$, a and b do not intersect any edge in $R \setminus \{e\}$. Then $R \cup \{a, b\} \setminus \{e\}$ is a rainbow matching, and it is larger than R , contradicting the maximality of R . \square

Lemma 12 implies

$$\sum_{e \in R} |T(e)| \leq rq. \tag{1}$$

2.2 Parts (1) and (2) of Theorem 2

Given $t = t(G)$, the maximal size of G is attained when the edges in $R \setminus T^G$ are arranged in pairs, so that the two edges in each pair meet at most $2r - 1$ distinct edges of G , and the last edge, if it exists (namely if R is odd), meets at most $r - 1$ additional edges of G .

Since $\bigcup R$ is a cover for G , this implies

$$n = |G| \leq rt + (q - t)\left(r - \frac{1}{2}\right) = qr - \frac{1}{2}(q - t).$$

So, $t \geq 2n - 2qr + q$. Summing up over all $G \in \mathcal{G}$ and changing the order of summation, we get

$$\sum_{e \in R} |T(e)| = \sum_{G \in \mathcal{G}} |T^G| \geq (2n - q(2r - 1))|\mathcal{G}| = (2n - q(2r - 1))(n - q).$$

Combining this with (1) yields

$$(2n - q(2r - 1))(n - q) \leq rq. \quad (2)$$

Claim 13. $q > \frac{n - \frac{1}{2} - \frac{3}{4r-6}}{r - \frac{1}{2}}$.

Proof. Rearranging (2), we get

$$(2r - 1)q^2 - (n + 2rn + r)q + 2n^2 \leq 0.$$

Solving for q , we have $q \geq q_1$, where

$$q_1 = \frac{n + 2rn + r - \sqrt{(n + 2rn + r)^2 - 8(2r - 1)n^2}}{2(2r - 1)}. \quad (3)$$

Let

$$\begin{aligned} \Delta &= (n + 2rn + r)^2 - 8(2r - 1)n^2 \\ &= n^2(4r^2 - 12r + 9) + n(4r^2 + 2r) + r^2 \end{aligned}$$

be the discriminant in (3). Completing Δ to a square entails

$$\Delta < n^2(2r - 3)^2 + 2n(2r^2 + r) + \left(\frac{2r^2 + r}{2r - 3}\right)^2,$$

implying that

$$\sqrt{\Delta} < n(2r - 3) + \frac{2r^2 + r}{2r - 3} = n(2r - 3) + r + 2 + \frac{6}{2r - 3}.$$

Thus

$$\begin{aligned} q_1 &= \frac{n + 2rn + r - \sqrt{\Delta}}{2(2r - 1)} \\ &> \frac{n + 2rn + r - n(2r - 3) - r - 2 - \frac{6}{2r-3}}{2(2r - 1)} \\ &= \frac{n - \frac{1}{2} - \frac{3}{4r-6}}{r - \frac{1}{2}}. \end{aligned}$$

This shows

$$q > \frac{n - \frac{1}{2} - \frac{3}{4r-6}}{r - \frac{1}{2}}. \quad \square$$

The claim entails Part (2) of Theorem 2. Part (1) follows upon plugging in $r = 2$.

2.3 More definitions

For the proofs of parts (3) and (4) of Theorem 2 we need additional definitions. For a matching $G \in \mathcal{G}$, we say that a pair of edges $\{e, f\}$ in R is *half- G -wasteful* if there are three edges $g_e, g_f, g_{ef} \in G$ satisfying the following:

- g_e intersects e and no other edge in R ,
- g_f intersects f and no other edge in R , and
- g_{ef} intersects both e and f .

We also say then that each of e and f are *half- G -wasteful*. Denote by $HW(e)$ the set of all $G \in \mathcal{G}$ for which the edge $e \in R$ is half- G -wasteful. Let W^G be the set of edges $e \in R$ such that $G \in HW(e)$, and write $w = |W^G|$.

Let B be the bipartite graph with respective sides \mathcal{G} and R , in which $G \in \mathcal{G}$ is adjacent to $e \in R$ if and only if $e \in HW(G)$. Then

$$\sum_{e \in R} |HW(e)| = \sum_{e \in R} \deg_B(e) = |E(B)|. \quad (4)$$

Let $N_B(e)$ be the neighborhood of $e \in R$ in the graph B .

The maximal size of $G \in \mathcal{G}$ is attained when the edges in $R \setminus (T^G \cup W^G)$ are arranged in triples, so that the three edges in each triple meet at most four distinct edges of G . Since $\bigcup R$ is a cover of every $G \in \mathcal{G}$, we have

$$n = |G| \leq 2t + \frac{3}{2}w + \frac{4}{3}(q - t - w),$$

implying $3n - 4q \leq 2t + \frac{w}{2}$. Summing over all $G \in \mathcal{G}$ we obtain

$$\begin{aligned} 2 \sum_{e \in R} |T(e)| + \frac{1}{2} \sum_{e \in R} |HW(e)| &= \sum_{G \in \mathcal{G}} (2|T^G| + \frac{1}{2}|W^G|) \\ &\geq (3n - 4q)|\mathcal{G}| \geq (3n - 4q)(n - q). \end{aligned} \quad (5)$$

2.4 Part (3) of Theorem 2

Assume the conditions of Part (3) of Theorem 2 hold, namely that there is no C_3 or C_5 in $\bigcup_{i=1}^n F_i$.

Lemma 14. $|T(e)| = 1$, hence $\sum_{e \in R} |T(e)| \leq q$.

Proof. Assume $G, H \in T(e)$. Let $g, g' \in G$ and $h, h' \in H$ be edges intersecting e and not any other edge in R , such that g, h meet at the same vertex of e . Since the graph contains no triangles, the set $(R \setminus e) \cup \{g, h'\}$ is a rainbow matching of size $q + 1$, contradicting the maximality of R . \square

Lemma 15. Let $f \in R$ and $g_e, g_f, g_{ef} \in G$ be edges witnessing the fact that $G \in HW(e)$. Suppose $G' \in \mathcal{G}$, $G \neq G'$, and there exists $g'_e \in G'$ intersecting e in one vertex and not intersecting any other edge in R . Then $g_e \cap g'_e \cap e \neq \emptyset$.

Proof. We have to show that g_e and g'_e meet e at the same vertex. If not, then since the graph contains no triangles, the set $(R \setminus e) \cup \{g_e, g'_e\}$ is a rainbow matching of size $q + 1$, contradicting the maximality of R . \square

Lemma 16. $\sum_{e \in R} |HW(e)| \leq 2q$.

Proof. By (4) it is enough to show $|E(B)| \leq 2q$.

Claim 17. Let $G_1, G_2 \in \mathcal{G}$ and $e, f \in R$. If (e, f) is a half-wasteful pair for both G_1, G_2 then $\deg_B(e) = 2$.

Proof of the claim. Clearly, $\deg_B(e) \geq 2$ since $G_1, G_2 \in N_B(e)$. Assume to the contrary that $N_B(e)$ contains $G_3 \in \mathcal{G}$ so that $G_3 \neq G_1, G_2$. Then by Lemma 15 the edges in G_1, G_2, G_3 that meet only e intersect at the same vertex of e . Since the graph does not contain C_3 , this implies that the edges in G_1, G_2 that meet both e and f coincide.

Let $a \in G_1$ be an edge meeting only f in R , let $b \in G_2$ be an edge meeting e, f in R , and let $c \in G_3$ be an edge meeting only e in R . Note that a, c do not intersect since the graph contains no C_5 . Replacing e, f in R with a, b, c we obtain a rainbow matching of size $q + 1$, a contradiction. \diamond

Let $D = \{e \in R \mid \deg_B(e) \geq 3\}$.

Claim 18. For every $e \in D$ there exists a subset $S(e) \subseteq R$ with the following properties:

1. $|S(e)| = \deg_B(e)$.
2. For every $f \in S(e)$, $\deg_B(f) = 1$.
3. $S(e) \cap S(f) = \emptyset$ whenever $e, f \in D$ and $e \neq f$.

Proof of the claim. Let $e \in D$ and write $d = \deg_B(e)$. Let $N_B(e) = \{G_1, \dots, G_d\}$, and let $f_1, \dots, f_d \in R$ be edges such that (e, f_i) is a half- G_i -wasteful pair. Let $S(e) = \{f_1, \dots, f_d\}$.

To prove (1), we have to show that $f_i \neq f_j$ if $i \neq j$. Indeed, if $f := f_i = f_j$ then (e, f) is a half-wasteful pair for both G_i, G_j , implying by Claim 17 that $\deg_B(e) = 2$, contradicting the fact that $e \in D$.

To prove (2), we show that $\deg_B(f_i) = 1$ for every i . First, G_i is adjacent to f_i in B , showing $\deg_B(f_i) \geq 1$. Assume for contradiction that there exists $G \in \mathcal{G}$, $G \neq G_i$, so that f_i is also a half- G -wasteful edge. Let $j \in [d]$ so that $G_j \neq G_i, G$ (such j exists since $d \geq 3$). Let $a \in G_i, b \in G_j$ be edges meeting only e in R . Let $c \in G_i$ be an edge meeting both e, f_i . Let $d \in G_i$ and $g \in G$ be edges meeting only f in R . Then by Lemma 15, a, b meet e at the same vertex and d, g meet f at the same vertex. Since the graph contains no C_3 , this implies that $b \cap c = c \cap g = \emptyset$. Since the graph contains no C_5 , we have also

$b \cap g = \emptyset$, implying that $(R \setminus \{e, f\}) \cup \{b, c, g\}$ is a rainbow matching of size $q + 1$, a contradiction.

To prove (3), let $e \neq f \in D$ and suppose $h \in S(e) \cap S(f)$. Let $G_1 \in HW(e, h)$ and $G_2 \in HW(f, h)$. Since G_1 is a matching $G_1 \neq G_2$. Let $G_3 \in N_B(e)$ so that $G_3 \neq G_1, G_2$ (it exists because $deg_B(e) \geq 3$). By Lemma 15, the edges in G_1, G_2 meeting only h intersect at the same vertex of h . Let $b \in G_2$ be an edge meeting only h in R , let $a \in G_1$ be an edge meeting e, h , and let $c \in G_3$ be an edge meeting only e in R . Again, $a \cap b = a \cap c = \emptyset$ for otherwise we have a C_3 , and $b \cap c = \emptyset$ for otherwise we have a C_5 . Then like before, replacing e, h in R by a, b, c results with a larger rainbow matching, a contradiction. \diamond

For $i = 1, 2$ let $U_i = \{e \in R \mid deg_B(e) = i\}$, and let $U_3 = D = R \setminus (U_1 \cup U_2)$. For $i \in [3]$ let E_i be the set of edges in B adjacent to a vertex in U_i .

By Claim 18 $|E_3| \leq |U_1| = |E_1|$. Therefore,

$$|E(B)| = \sum_{i=1}^3 |E_i| \leq 2|E_1| + |E_2| \leq 2 \sum_{i=1}^3 |U_i| \leq 2q.$$

This completes the proof of the lemma. \square

Combining (5) with Lemmas 14 and 16 we get

$$(3n - 4q)(n - q) \leq 3q,$$

entailing (using similar calculations to those in Claim 13) $q > \frac{3n}{4} - \frac{9}{4}$. This completes the proof of Part (3) of the theorem.

2.5 Part (4) of Theorem 2

Assume the conditions of Part (4) of Theorem 2 hold. That is, the matchings F_1, \dots, F_n are pairwise disjoint.

For $e \in R$, let E_e be the set of edges $g \in \bigcup HW(e)$ intersecting e and no other edge in R .

Lemma 19. *Suppose $e \in R$ has $deg_B(e) \geq 3$. Then all the edges in E_e intersect e at the same vertex (i.e., $\bigcap E_e \cap e \neq \emptyset$).*

Proof. Write $e = uv$. Let G_1, G_2, G_3 be three distinct matchings in $HW(e)$, and let $g_i \in G_i$ be an edge intersecting e and no other edges in R . It is enough to show that g_1, g_2, g_3 meet e at the same vertex. Assume g_1, g_2 meet e at v , and g_3 meets e at u . Since G_1, G_2 are disjoint, $g_1 \neq g_2$, and thus g_3 cannot meet both g_1 and g_2 . Say $g_1 \cap g_3 = \emptyset$. Then the set $(R \setminus e) \cup \{g_1, g_3\}$ is a rainbow matching of size $q + 1$, contradicting the maximality of R . \square

Let $D = \{e \in R \mid deg_B(e) \geq 5\}$.

Lemma 20. *For every $e \in D$ there exists a subset $S(e) \subseteq R$ with the following properties:*

1. $|S(e)| \geq \frac{\deg_B(e)}{2}$.
2. For every $f \in S(e)$, $\deg_B(f) \leq 2$.
3. If $e, f \in D$ and $e \neq f$, then $S(e) \cap S(f) = \emptyset$.

Proof. Let $e \in D$ and write $d = \deg_B(e)$. Let $N_B(e) = \{G_1, \dots, G_d\}$, and let $f_1, \dots, f_d \in R$ be edges so that (e, f_i) is a half- G_i -wasteful pair. Let $S(e) = \{f_1, \dots, f_d\}$.

To prove (1) it is enough to show that there do not exist three indices $j, k, \ell \in [d]$ such that $f_j = f_k = f_\ell$. Assume for contradiction that such j, k, ℓ exist, and let $f := f_j = f_k = f_\ell$. For $i \in \{j, k, \ell\}$, let $g_e^i, g_f^i, g_{ef}^i \in G_i$ be edges witnessing the fact that (e, f) is a half- G_i -wasteful pair. By Lemma 19, $g_e^j \cap g_e^k \cap g_e^\ell \cap e \neq \emptyset$, implying $g_{ef}^j \cap g_{ef}^k \cap g_{ef}^\ell \cap e \neq \emptyset$. Since the edges $g_{ef}^j, g_{ef}^k, g_{ef}^\ell$ meet also f , at least two of them must be equal, contradicting the fact that the matchings are pairwise disjoint.

To prove (2), let $f \in S(e)$ and assume $\deg_B(f) \geq 3$. Let $H \in HW(e, f)$, with edges $h_e, h_f, h_{ef} \in H$ witnessing this fact. Let $G, G' \in HW(f)$ be two matchings different than H . By Lemma 19, all the edges in E_f meet $f \cap h_f$, and all the edges in E_e meet $e \cap h_e$. In particular, no edge in $E_f \cup E_e$ intersects h_{ef} .

Let $g \in E_f \cap G$ and $g' \in E_f \cap G'$. Then g, g' are distinct. Since $\deg_B(e) \geq 4$, $E_e \setminus (H \cup G' \cup G'')$ contains at least one edge. Thus there exist edges $a \in \{g, g'\}$ and $b \in E_e \setminus (H \cup G' \cup G'')$ such that a, b do not intersect. Replacing e, f in R with a, b and h_{ef} , we get a larger rainbow matching, a contradiction.

To prove (3), let $e \neq f \in D$ and suppose $h \in S(e) \cap S(f)$. Let $A \in HW(e, h)$ and $B \in HW(f, h)$. Since A, B are matchings, $A \neq B$. Let $a_e, a_h, a_{eh} \in A$ and $b_f, b_h, b_{fh} \in B$ be edges witnessing $A \in HW(e, h)$ and $B \in HW(f, h)$, respectively. By Lemma 19, all the edges in E_e meet e at the same vertex, and all the edges in E_f meet f at the same vertex.

Split into two cases.

Case 1: The edges a_h and b_h intersect h at the same vertex.

Since $|E_e \setminus \{a_e, b_e\}| \geq 2$ there exists an edge in $c \in E_e \setminus \{a_e, b_e\}$ not intersecting b_h . Thus $(R \cup \{b_h, a_{eh}, c\}) \setminus \{e, h\}$ is a larger rainbow matching, a contradiction.

Case 2: The edges a_h and b_h intersect h at two different vertices.

In this case a_{eh}, b_{fh} do not intersect, and clearly, no edge in $E_e \cup E_f$ intersect either a_{eh} or b_{fh} . Since $\deg_B(e), \deg_B(f) \geq 5$, there exist edges $c \in E_e \setminus \{a_e, b_e\}$ and $d \in E_f \setminus \{a_f, b_f\}$ so that $\{c, d, a_{eh}, b_{fh}\}$ is a rainbow matching in \mathcal{G} . Thus $(R \cup \{c, d, a_{eh}, b_{fh}\}) \setminus \{e, f, h\}$ is a rainbow matching of size $q + 1$, a contradiction. \square

Lemma 21. $\sum_{e \in R} |HW(e)| \leq 4q$

Proof. For $i \in [4]$ let U_i be the set of edges in R of degree i in B , and let $U_5 = D$. For $i \in [5]$, let E_i be the set of edges in B adjacent to a vertex in U_i . By Lemma 20 we have,

$$|E_5| \leq 2(|U_1| + |U_2|).$$

Therefore,

$$|E(B)| = \sum_{i=1}^5 |E_i| \leq \sum_{i=1}^4 i|U_i| + 2(|U_1| + |U_2|) \leq 4 \sum_{i=1}^5 |U_i| \leq 4q.$$

This completes the proof of the lemma. \square

Combining (1), (5) and Lemma 21 we get

$$(3n - 4q)(n - q) \leq 6q,$$

implying $q > \frac{3n}{4} - \frac{9}{2}$ and proving Part (4). This completes the proof of Theorem 2.

Question 22. Does there exist a function $f(r)$ such that $(n, n) \rightarrow_r n - f(r)$?

3 Two proofs of Theorem 8

For a rooted tree G and a vertex v on it let Gv be the path on G from the root to v . The tree G is called *rainbow-monotone* if the directed path Gv is rainbow-monotone for every $v \in V(G)$. For a forest \mathcal{F} whose components are rooted trees and a vertex $v \in V(\mathcal{F})$, we write $\mathcal{F}v$ for the path Gv , where G is the component of \mathcal{F} containing v . Similarly, vG denotes the sub-tree of G rooted at v . If K, L are trees sharing a vertex v , we denote by KvL the edge-wise union of the path Kv and the tree vL . This definition does not imply that KvL is a tree, but in our applications it will be.

First proof of Theorem 8. Inducting on i , we grow trees $T_i(s)$ rooted at s for every $s \in S$ and $i \leq m$. At each step we shall denote by Y_i^k the set of vertices $y \in Y$ that are reached by precisely k trees $T_i(s)$ ($k \geq 0$).

The inductive construction will maintain two conditions:

1. $Y_i^k = \emptyset$ for all $k > 2$, and
2. $T_i(s)$ is rainbow-monotone for every $s \in S$.

For the base, $i = 0$, let $T_0(s) = (s)$, a single-vertex tree, for every $s \in S$. Condition (2) is obvious, and (1) is true since $Y_0^0 = Y$ and $Y_0^k = \emptyset$ for every $k > 0$.

Suppose $T_i(s)$ have been defined for all $s \in S$, and that they satisfy (1) and (2). Let $in(P_{i+1}) = p$ and $ter(P_{i+1}) = q$. Let $X = V(T_i(p)) \cup Y_i^2$. Then $p \in X$. Since $q \in S$ we may assume that $q \notin V(T_i(s))$ for any $s \neq p$, since otherwise the path $T_i(s)q$ is the desired path.

So, there is an edge $xy \in E(P_{i+1})$ such that $x \in X$, $y \notin X$.

Case I $x \in V(T_i(p))$. Define then $T_{i+1}(p) = T_i(p) \cup \{xy\}$ (it is a tree since $y \notin T_i(p)$) and $T_{i+1}(s) = T_i(s)$ for all $s \neq p$.

Case II $x \in Y_i^2 \setminus V(T_i(p))$. Since $y \notin Y_i^2$, there exists $r \in S$ (possibly $r = p$) for which $x \in V(T_i(r))$ and $y \notin V(T_i(r))$. Then defining $T_{i+1}(r) = T_i(r) \cup \{xy\}$ and $T_{i+1}(s) = T_i(s)$ for all $s \neq r$ maintains the inductive assumptions.

Since at each of the m steps we are adding a vertex in Y to the trees, and since no vertex in Y appears more than twice, $m \leq 2|Y|$, as desired. \square

The second proof is longer, but less of a hocus-pocus. It uses part (5) of Theorem 5. We give it here since it may contain ideas relevant to Conjecture 6, the alternating paths version of the theorem.

Second Proof of Theorem 8. First - a blueprint, which will then be given a rigorous formulation.

Assume that there is no rainbow-monotone $S - S$ path. We grow inductively rainbow-monotone forests \mathcal{F}_i , each consisting of trees $T_i(s)$ rooted at s , for all $s \in S$. We also keep track of sets $W_i(t)$ of paths for each $t \in S$. These are “temporarily wasted” paths, namely paths not adding an edge to \mathcal{F}_i . A “wasted” path in $W_i(t)$ will be used at the end of the procedure to find a rainbow-monotone path ending at t .

Let \mathcal{F}_0 consist of the single vertex trees (s) , $s \in S$, and $W_0(t) = \emptyset$ for every $t \in S$.

The inductive step: If there exists an edge $xy \in E(P_{i+1})$ such that $x \in V[\mathcal{F}_i]$ and $y \notin V[\mathcal{F}_i]$ choose one such edge and add it to \mathcal{F}_i to form \mathcal{F}_{i+1} . If not, let $\mathcal{F}_{i+1} = \mathcal{F}_i$ and let $W_{i+1}(ter(P_{i+1})) = W_i(ter(P_{i+1})) \cup \{P_{i+1}\}$, and $W_{i+1}(t) = W_i(t)$ for every $t \neq ter(P_{i+1})$.

Each non-wasted path P_i adds a vertex of Y to $V[\mathcal{F}_i]$. Hence the number of such paths is at most $|Y|$. To get the desired inequality $m \leq 2|Y|$, it suffices then to show that the number of wasted paths is at most $|Y|$. This will follow from:

$$|W_m(t)| \leq |V[T_m(t)] \setminus \{t\}| \text{ for every } t \in S.$$

To see this, assume $|W_m(t)| > |V[T_m(t)] \setminus \{t\}|$. Contract $V \setminus V[T_m(t)]$ to a single vertex z . Let Q_i , $i \leq m$ consist of the part of the path P_i contained in $V[T_m(t)]$, and let P'_i be a path in the contracted graph, defined by $E(P'_i) = E(Q_i) \cup \{z \text{ in}(Q_i)\}$ (namely, Q_i with z appended to its initial vertex). Then by part (5) of Theorem 5 there is a rainbow-monotone path (with respect to the paths P'_i) from z to t , which can then be extended (by uncontracting z) to a rainbow-monotone $s - t$ path for some $s \neq t$.

A rigorous argument performs both types of steps together. We use the “wasted” paths as we go along, instead of waiting till the end of the process.

Below we use the symbol \diamond to mark the conclusion of an intermediate step in a proof.

We construct inductively, for $j \leq m$, non-decreasing (containment-wise) disjoint trees $T_j(s)$, $s \in S$, and non-decreasing forests $\mathcal{Z}_j(s)$, $s \in S$.

The inductive construction will maintain the following properties.

- (P1) At each step precisely one $T_j(s)$ or one $\mathcal{Z}_j(s)$ grows, meaning that it is added a vertex from Y , not met as yet.
- (P2) The tree $T_j(s)$ is rooted at $s \in S$, and $V(T_j(s)) \cap V(T_j(s')) = \emptyset$ whenever $s \neq s'$.
- (P3) The union \mathcal{T}_j of the trees $T_j(s)$, $s \in S$, is a forest.
- (P4) For $s \in S$, the forest $\mathcal{Z}_j(s)$ is the union of rooted trees, having a special form: each tree in $\mathcal{Z}_j(s)$ has its root r in $V(T_j(s'))$ for some $s' \in S$, $s' \neq s$, and its other vertices in $V(T_j(s))$. We denote such a tree by $Z_j(s, r)$.

- (P6) At each step a new vertex from Y is added either to \mathcal{T}_j or to $\mathcal{Z}_j(s)$ for some $s \in S$.
- (P7) Edges are added to $Z_i(s, r)$ for $i \geq j$ only after the path $\mathcal{T}_j r$ terminating at r has been constructed.
- (P8) The forests \mathcal{T}_j and $\mathcal{Z}_j(s)$ are rainbow-monotone.
- (P9) If two trees, $U \in \mathcal{Z}_j(s_1)$ and $W \in \mathcal{Z}_j(s_2)$ meet, where $s_1 \neq s_2$, then they share only one vertex, which is the root of at least one of them.

(The secret wish of the trees in $\mathcal{Z}_j(s)$ is to reach s . If fulfilled, the path in $\mathcal{T}_j r \mathcal{Z}_j(r, s)$ going from the the root s' of the tree in \mathcal{T}_j containing r to s , will then be a rainbow $S - S$ path.)

Combined, these properties yield:

Claim 23. *The path $\mathcal{T}_j r$ does not meet $V(Z_j(s, r)) \setminus \{r\}$, hence $\mathcal{T}_j r Z_j(s, r)$ is a tree.*

Having introduced the protagonists of the proof and their intended inductively preserved properties, let us describe the inductive construction. The preservation of the properties (P1)-(P8) will be mostly easy to validate - we comment only on those that are not obvious. Let $\mathcal{G}_{(0,0)}$ consist of the singleton trees (s) , $s \in S$, and let $\mathcal{Z}_{(0,0)}(s) = \emptyset$ for every $s \in S$. Suppose \mathcal{G}_j and $\mathcal{Z}_j(s)$, $s \in S$, have been defined; let $j^+ = (k, i)$ be the index following j in the lexicographic order, and let $P = P_k^i$. We shall use P to extend either \mathcal{T}_j or $\mathcal{Z}_j(s)$ for some $s \in S$, maintaining monotonicity. Below we assume, for contradiction, that such an extension is not possible.

Claim 24. $V(P) \subseteq V(\mathcal{T}_j)$.

If not, the first edge on P leaving $V(\mathcal{T}_j)$ could be added to \mathcal{T}_j . ◇

Claim 25. *The forests \mathcal{T}_j , as well as the trees $\mathcal{T}_j r Z_j(s, r)$, are rainbow-monotone.*

Proof of the claim. In the construction we go over the paths P_i^k one by one. When an edge e from P_i^k is added, it is at the top of a tree. When we get to P_j^k , $j > i$, we cannot put an edge from it right after e , because P_i^k and P_j^k are vertex disjoint. So e will be put on top of a path in a tree, that consists of edges from P_z^t for $t < k$. This explains the monotonicity of both $\mathcal{T}_j(s)$ and $Z_j(s, r)$. By (3) it follows that their concatenation $\mathcal{T}_j r Z_j(s, r)$ is also rainbow-monotone. ◇

Claim 26. *If $s \in S$ then $s \notin V(\mathcal{Z}_j(s))$.*

Indeed, if $s \in V(Z_j(s, r))$ then the path $\mathcal{T}_j r Z_j(s, r) s$ terminating at s is a rainbow-monotone $S - S$ path. ◇

Let $t = \text{ter}(P)$. By Claim 26, $t \notin V(\mathcal{Z}_j(t))$. Let v be the first vertex on P that belongs to $V(G_j(t)) \setminus V(\mathcal{Z}_j(t))$, and let u be the vertex preceding it on P . We now add the edge uv to the forest $\mathcal{Z}_j(t)$. There are two possible cases.

Case I. $u \in V(G_j(t))$. Then the edge uv is added to the existing tree in $\mathcal{Z}_j(t)$ containing u .

Case II. $u \notin V(G_j(t))$. It is possible that u has been used already as a root of a tree $Z_j(t, u)$ in $\mathcal{Z}_j(t)$. In this case we add the edge uv to $Z_j(t, u)$. Otherwise, a new tree $Z_{j+}(t, u)$ is created in $\mathcal{Z}_{j+}(t)$, consisting of the single edge uv .

Note that in any of these cases $\mathcal{Z}_j(t)$ is enlarged by the addition of the vertex v .

Since each step uses one path P_i and since at each step either a rainbow $S - S$ path emerges, or one of the forests \mathcal{T}_j or \mathcal{Z}_j annexes a vertex of Y , it follows that $m \leq 2|Y|$, as desired. \square

References

- [1] R. Aharoni and E. Berger, Rainbow matchings in r -partite r -graphs, *Electron. J. Combin* 16 (2009), #R119.
- [2] R. Aharoni, E. Berger, M. Chudnovsky, D. Howard, P. Seymour, Large rainbow matchings in general graphs, *European J. Combin.* 79 (2019), 222–227
- [3] R. Aharoni, J. Briggs, J. Kim and M. Kim, Rainbow paths and rainbow matchings in graphs, [arXiv:2004.0759](https://arxiv.org/abs/2004.0759).
- [4] R. Aharoni, R. Holzman and Z. Jiang, Rainbow fractional matchings, *Combinatorica* 39 (2019), 1191–1202.
- [5] R. Aharoni, D. Kotlar and R. Ziv, Representation of large matchings in bipartite graphs, *SIAM Journal of Discrete Mathematics* 31 (2017) 1726–1731.
- [6] R. Aharoni, D. Kotlar and R. Ziv, Uniqueness of the extreme cases in theorems of Drisko and Erdős-Ginzburg-Ziv, *European J. Combin.* 67 (2018), 222–229.
- [7] J. Barát, A. Gyárfás and G. Sárközy, Rainbow matchings in bipartite multigraphs, *Period. Math. Hung.* 74 (2017) 108–111.
- [8] R. A. Brualdi and H. J. Ryser, *How Many Colors Guarantee a Rainbow Matching? Combinatorial Matrix Theory*, Cambridge University Press, Cambridge, UK, 1991.
- [9] D. Clemens and J. Ehrenmüller, An improved bound on the sizes of matchings guaranteeing a rainbow matching, *Electron. J. Combin.* 23 (2016), #P2.11.
- [10] D.M. Correia, A. Pokrovskiy and B. Sudakov, Short proofs of rainbow matching results, *Combin. Probab. Comput.* 30 (2021), 762–780.
- [11] A. A. Drisko, Transversals in row-latin rectangles, *J. Combin. Theory, Ser. A* 84 (1998) 181–195.
- [12] P. Gao, R. Ramadurai, I. Wanless and N. Wormald, Full rainbow matchings in graphs and hypergraphs, to appear in *Combinatorics, Probability and Computing*, <https://doi.org/10.1017/S0963548320000620>.
- [13] A. Pokrovskiy, An approximate version of a conjecture of Aharoni and Berger. *Adv. Math.* 333 (2018), 1197–1241.
- [14] S. K. Stein, Transversals of Latin squares and their generalizations, *Pacific J. Math.* 59 (1975), 567–575.
- [15] D.E. Woolbright, An $n \times n$ Latin square has a transversal with at least $n - \sqrt{n}$ distinct elements, *J. Combin. Theory, Ser. A* 24 (1978), 235–237.