Directed paths: from Ramsey to Ruzsa and Szemerédi

Po-Shen Loh *

Abstract

Starting from an innocent Ramsey-theoretic question regarding directed paths in tournaments, we discover a series of rich and surprising connections that lead into the theory around a fundamental problem in Combinatorics: the Ruzsa-Szemerédi induced matching problem. Using these relationships, we prove that every coloring of the edges of the transitive n-vertex tournament using three colors contains a directed path of length at least $\sqrt{n} \cdot e^{\log^* n}$ which entirely avoids some color. We also completely resolve the analogous question for ordinary monochromatic directed paths in general tournaments, as well as natural generalizations of the Ruzsa-Szemerédi problem which we encounter through our investigation.

1 Introduction

Ramsey theory is a central area in Combinatorics which concerns structure that must exist in arbitrary (but large) configurations. The classical starting point for Graph Ramsey Theory considers the number of vertices required in order for every 2-edge-coloring of a sufficiently large complete graph to contain a monochromatic copy of a fixed complete graph K_t [8, 22]. Over the century, many other target subgraphs were considered, from the specific, such as paths [13], to trees and cliques [4], to entire classes such as graphs with bounded degree [5, 6, 14] and degeneracy [3].

This paper explores several interesting Ramsey-type problems about paths, and so we briefly outline some of the history. In undirected graphs, an old result of Gerencsér and Gyárfás [13] established that the corresponding 2-color Ramsey number of an n-vertex path is precisely $\lfloor \frac{3n-2}{2} \rfloor$. When the number of colors increases beyond two, however, less is known. The result for three colors was proved asymptotically by Figaj and Łuczak [9], and recently for large n by Gyárfás, Ruszinkó, Sárközy, and Szemerédi [15].

Another flavor of Ramsey-type results, which also inspires the work in this paper, stems from the following celebrated theorem of Erdős and Szekeres [8]: every sequence of n distinct real numbers contains a monotone subsequence of length at least $\sqrt{n}$. This is often proven by the pigeonhole principle. Indeed, assign to each number an ordered pair (x, y), where x tracks the length of the longest increasing subsequence ending at that number, and y tracks the decreasing analogue. Since the numbers are distinct, it is easy to see that the ordered pairs must also be distinct, and therefore some ordered pair has an element at least $\sqrt{n}$. Many extensions have been found for this theorem (see, e.g., any of [11, 17, 21, 25, 26]).

The same pigeonhole argument also proves the following straightforward result. Consider the transitive tournament on the vertex set $\{1, \ldots, n\}$, where the edge between $i < j$ is oriented in the

*Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213. E-mail: ploh@cmu.edu. Research supported in part by NSF grants DMS-1041500 and DMS-1201380, and by a USA-Israel BSF Grant.

direction $\overrightarrow{ij}$. Then, every 2-coloring of the edges of T_n has a monochromatic (directed) path with at least $\sqrt{n}$ vertices (hereafter called the *vertex-length*).

This paper discusses several questions that start from this classical argument, leading to surprising and interesting connections across Combinatorics. This serendipitous tour passes through directed graph Ramsey theory, to k-majority tournaments, and ultimately to the borders of a foundational problem of Ruzsa and Szemerédi on induced matchings. Our first result considers general tournaments (for which the pigeonhole argument completely falls apart), and completely solves the problem there, showing as a corollary that transitive tournaments are the worst case.

Theorem 1.1. *For all positive integers r and n, every r-coloring of the edges of every n-vertex tournament contains a monochromatic path with at least $n^{1/r}$ vertices. This is sharp for every r and n: there are r-edge-colorings of n-vertex tournaments, where all monochromatic paths have at most $\lceil n^{1/r} \rceil$ vertices.*

Another direction appears to lead to a challenging problem. With two colors, the concept of "monochromatic" is unambiguous. With three or more colors, however, there is another natural generalization of the basic concept: one can seek substructures which are *1-color-avoiding*, i.e., in which there is at least one color which does not appear at all. (In the 2-color setting, these two concepts are identical.) For the generalization of Theorem 1.1 to 1-color-avoiding paths in the 3-color setting, even the transitive tournament case is not well-understood.

Problem 1.1. *Determine $f(n)$, the maximum number such that every 3-coloring of the edges of the n-vertex transitive tournament contains a directed path with at least $f(n)$ vertices, which avoids at least one of the colors.*

By merging two of the three color classes, one may obtain a bound $f(n) \geq \sqrt{n}$ using the 2-color Theorem 1.1, and the following standard construction achieves $f(n) \leq n^{2/3}$ when n is a perfect cube: identify the vertex set with the triples $\{1, \ldots, n\}^3$, ordered lexicographically, and color an edge $(x, y, z) \to (x', y', z')$ based upon the leftmost position at which the triples differ. The best result, however, is still unknown.

It turns out that there are interesting connections between this problem and a longstanding open question which traces back to a result of Ruzsa and Szemerédi [24] on induced matchings, which has rich connections to Szemerédi's Regularity Lemma (see, e.g., the survey [18]).

Problem 1.2. *Let $M(n, k)$ be the maximum number of edges in an n-vertex graph $G = (V, E)$ whose edge set is the union of k induced matchings. That is, $E = E_1 \cup \ldots \cup E_k$, where each E_i is precisely the set of edges induced by $G[V_i]$ for some subset $V_i \subset V$, and each E_i is a matching. Determine the asymptotics of $M(n, k)$.*

Leveraging this connection, and using the current best bounds on Problem 1.2 (due to Fox's [10] result on the Triangle Removal Lemma), we obtain a non-trivial lower bound on $f(n)$. Here, $\log^*$ is the iterated logarithm, or the inverse of the tower function $T(n) = 2^{T(n-1)}$, $T(0) = 1$. Also, we write $f(n) = \Omega(g(n))$ to mean that there exists a positive constant c such that $f(n) \geq cg(n)$ for all sufficiently large n.

Theorem 1.2. *Every 3-coloring of the edges of the n-vertex transitive tournament contains a directed path of length at least $\Omega(\sqrt{n} \cdot e^{\log^* n})$.*

Furthermore, we discover that Problem 1.1 is related to an "ordered" variant of the following generalization of the Ruzsa-Szemerédi problem.

Problem 1.3. *Let $M_l(n,k)$ be the maximum number of edges in a n-vertex bipartite graph $G = (V,E)$ whose edge set is the union of k matchings which are **l-separated**. That is, $E = E_1 \cup \ldots \cup E_k$, where each E_i is a matching, and there is no sequence of vertices $a_0, a_1, \ldots, a_t$ with $t \leq l$, such that a_0 and a_t are both incident to distinct edges from the same E_i, and each of a_0a_1, a_1a_2, $\ldots$, $a_{t-1}a_t$ are edges of E. Determine the asymptotics of $M_l(n,k)$.*

Note that when $l = 1$, this corresponds to the induced matchings from Problem 1.2. In our application, we are most interested in the symmetric bipartite case, which corresponds to $M_l(2n,n)$. We completely resolve the unordered version stated above. It turns out that already for $l = 2$, the exponent immediately drops, and remains level for all higher l, so that the unique behavior is special to $l = 1$.

Theorem 1.3. *For all positive integers $l \geq 2$ and n, we have $M_l(2n,n) = (1 + o(1))n^{3/2}$.*

Finally, we show that the same bound on the ordered variant (formulated in Section 3.3) would improve Theorem 1.2 to achieve the asymptotically sharp bound $f(n) \geq \Omega(n^{2/3})$.

This paper is organized as follows. The next section proves Theorem 1.1 on monochromatic directed paths in general tournaments. Section 3 covers the rest of the results, weaving through connections to k-majority tournaments, Ruzsa-Szemerédi induced matchings, and ordered generalizations. The final section summarizes some interesting open problems.

2 Monochromatic

In this section, we prove Theorem 1.1, which concerns monochromatic directed paths in general tournaments. The original approach for transitive tournaments (based upon the pigeonhole-principle) completely falls apart, because there is no natural vertex ordering. Instead, we exhibit a short proof which employs a useful theorem which was independently discovered and rediscovered by Gallai [12], Hasse [16], Roy [23], and Vitaver [27].

Theorem 2.1. *(Gallai-Hasse-Roy-Vitaver.) Let G be an undirected graph with chromatic number k. No matter how its edges are oriented, the resulting directed graph contains a directed path of vertex-length at least k.*

Remark. This finds a directed *path*, in which no vertex is used more than once.

Proof of Theorem 1.1. Fix an arbitrary r-edge coloring of a tournament. This corresponds to an r-edge coloring of the underlying undirected graph K_n, which partitions of K_n into r edge-disjoint graphs G_1, G_2, $\ldots$, G_r, each corresponding to a color class in the tournament. In particular, if $\chi(G)$ denotes the vertex-chromatic number of G, we have

$$\chi(G_1)\chi(G_2)\cdots\chi(G_r) \geq \chi(K_n) = n,$$

because the vertex-chromatic number is sub-multiplicative. Therefore, one of the color classes of the tournament has underlying undirected graph with $\chi(G_i) \geq n^{1/r}$, and Theorem 2.1 finds a directed path in that color i, completing the proof of the lower bound.

It remains to show that the lower bound is tight. We achieve this by constructing, for every positive integers r and k, an r-coloring of the edges of a k^r-vertex tournament, in which the longest monochromatic path has k vertices. Identify the vertices as r-tuples from $\{1, 2, \ldots, k\}^r$. For every two distinct vertices $(x_1, \ldots, x_r)$ and $(y_1, \ldots, y_r)$, identify the smallest index i for which $x_i \neq y_i$, and place the directed edge from $(x_1, \ldots, x_r) \to (y_1, \ldots, y_r)$ if $x_i < y_i$, introducing the reverse edge otherwise. In either case, color that edge using the i-th color. It is then clear that every monochromatic directed path is of the form where the first t coordinates are fixed, and the $(t+1)$-st coordinate is strictly increasing along the path; the length is then at most k vertices, as claimed. □

3 Avoiding one color

The rest of this paper focuses on directed paths that avoid at least one color. In this setting, the result is not even known for transitive tournaments, and so we focus on that case here. Recall that $f(N)$ is defined as the maximum number n such that every 3-coloring of the edges of the transitive N-vertex tournament contains a directed path of vertex-length at least n, which avoids at least one of the colors.

3.1 Reformulation and k-majority tournaments

We begin by stating an equivalent and independently natural problem. The following statement is more convenient for relating with k-majority tournaments, and this connection will be detailed at the end of this section.

Problem 3.1. *Determine $F(n)$, the maximum number of triples in a sequence $L_1, L_2, \ldots$ such that every coordinate is an integer between 1 and n inclusive, and for every $i < j$, there are at least two coordinates in which L_i is strictly less than L_j.*

Example. The sequence $(3, 2, 7)$, $(7, 3, 2)$, $(1, 8, 8)$, $(2, 9, 9)$, $\ldots$ is valid. Indeed, when comparing the second and fourth triples, note that although the first coordinate decreases from 7 to 2, the second and third coordinates both increase, from 3 to 8 and 2 to 8, respectively.

It turns out that Problem 1.1 (Ramsey for 1-color-avoiding paths) and Problem 3.1 (above) are completely equivalent, as established by the following lemma.

Lemma 3.1. *For any positive integers n and N:*

(i) *if $f(N) \le n$, then $F(n) \ge N$; and*

(ii) *if $f(N) > n$, then $F(n) < N$; and thus*

(iii) *the values of F are precisely the arguments at which f increases: $F(n) = N$ if and only if $f(N) = n$ and $f(N+1) > n$.*

Proof. For part (i), $f(N) \le n$ means that there is a 3-coloring of the edges of the transitive tournament on vertex set $\{1, \ldots N\}$, with all 1-color-avoiding paths having vertex-length at most n. Consider this coloring, and follow the Erdős-Szekeres proof. Associate a triple of numbers to each vertex, with $L_k = (x_k, y_k, z_k)$ corresponding to the k-th vertex, where x_k is the number of vertices in the longest path that ends at the k-th vertex and avoids the first color, and y_k and z_k are the corresponding statistics for avoiding the second and third colors, respectively.

Consider any two triples L_i and L_j with $i < j$. Since there are three colors overall, there are two colors which differ from the color of edge $\overrightarrow{ij}$. Let c be one of them. Then, the longest path which avoids color c and ends at the i-th vertex can be extended by edge $\overrightarrow{ij}$, to end at j. Therefore, the corresponding coordinate (for color c) of L_i must be strictly less than that of L_j. This holds also for the other color which differs from $\overrightarrow{ij}$, and so L_i is strictly less than L_j in at least two coordinates. We have verified the necessary properties produce a valid construction for Problem 3.1 with N triples, all of whose entries are at most n. This implies $F(n) \geq N$.

To prove (ii), fix an arbitrary sequence of triples $L_1, L_2, \dots, L_N$ with the property that for every $i < j$, there are at least two coordinates in which L_i is strictly less than L_j, and all coordinates are positive integers. We will show that some triple contains an integer strictly greater than n. Indeed, let us construct a 3-edge-coloring of a transitive tournament on vertex set $\{1, \dots, N\}$, where edges are directed in the natural forward increasing ordering. For each pair $i < j$, there is always a choice of one of the three coordinates such that L_i is strictly less than L_j in the other two coordinates. (If L_i is strictly less than L_j in all three coordinates, then there are three choices for this coordinate; disambiguate by selecting the first coordinate.) Use this coordinate as the color of the edge. Since $f(N) > n$ by assumption, there is a 1-color-avoiding directed path with more than n vertices.

At the same time, the following dynamic programming algorithm computes the length of the longest 1-color-avoiding directed path in a transitive tournament. Iteratively construct triples $T_1, \dots, T_N$, where each $T_k = (x_k, y_k, z_k)$ records the vertex-lengths of the longest 1-color-avoiding paths ending at the k-th vertex, avoiding the first, second, and third color, respectively. This can be achieved by starting with $T_1 = (1, 1, 1)$, and at each step, given $T_1, \dots, T_{k-1}$, constructing T_k by setting

$$
\begin{aligned}
x_k &= 1 + \max\{x_j : 1 \leq j < k \text{ and } \overrightarrow{jk} \text{ not in color } 1\} \\
y_k &= 1 + \max\{y_j : 1 \leq j < k \text{ and } \overrightarrow{jk} \text{ not in color } 2\} \\
z_k &= 1 + \max\{z_j : 1 \leq j < k \text{ and } \overrightarrow{jk} \text{ not in color } 3\},
\end{aligned}
$$

with the convention that the maximum is taken to be 0 if it is taken over the empty set. (This is actually consistent with the definition of T_1.)

By construction of the tournament coloring, we inductively have that for each k, the triple T_k is less than or equal to the given triple L_k in every coordinate. Since there is a 1-color-avoiding directed path with more than n vertices, this implies that some entry of some L_k exceeds n, as claimed.

The final statement (iii) hinges on two observations: the sequence $f(1), f(2), f(3), \dots$ is monotonic, and consecutive terms increase by at most one. To see that the latter is true, consider an arbitrary 3-edge-coloring of the transitive N-vertex tournament, and extend it to a 3-edge-coloring of the transitive $(N+1)$-vertex tournament by coloring all edges to the $(N+1)$-st vertex in the first color. This would then contain a 1-color-avoiding path of vertex-length at least $f(N+1)$, which interacts with the $(N+1)$-st vertex in its final vertex, if at all. Therefore, the original 3-edge-coloring of the N-vertex tournament contained a 1-color-avoiding path of vertex-length at least $f(N+1) - 1$, hence $f(N) \geq f(N+1) - 1$, as claimed.

Now suppose that $f(N) = n$ and $f(N+1) > n$. By (i), we have $F(n) \geq N$, and by (ii), we have $F(n) < N + 1$. Since $F(n)$ is an integer, this implies $F(n) = N$. Furthermore, since we have established that $f(1), f(2), \dots$ is monotonic with consecutive terms increasing by at most one, and we also know that $f(1) = 1$ and the sequence is unbounded, this determines $F(n)$ for every positive integer n. Since each $F(n)$ can only take one value, this establishes both directions of the final if-and-only-if statement. $\square$

Problem 3.1 is independently interesting in the above formulation, as it bears some relation to the area of k-majority tournaments.

Definition 3.1. *Given a set of k linear orders on a common ground set X (with k odd), the corresponding **k-majority tournament** has vertex set X, with an edge directed from x to y if and only if a majority of the k orders rank $x < y$.*

These types of tournaments have been actively studied in social science [19], as well as in Combinatorics, where previous researchers studied extremal problems regarding the sizes of dominating sets [1] and transitive subtournaments [20]. Problem 3.1 is connected to bounding transitive subtournaments in a natural class of constructions of k-majority tournaments. Indeed, a convenient ground set is the collection of k-tuples of integers between 1 and n inclusive, i.e., $X = \{1, \ldots, n\}^k$. Every permutation σ of $\{1, \ldots, k\}$ gives rise to a linear order on X, where any two k-tuples are compared lexicographically, with coordinate $\sigma(1)$ inspected first, breaking ties with coordinate $\sigma(2)$, and so on. Therefore, a natural construction for a k-majority tournament consists of selecting k such linear orderings which take turns primarily emphasizing each of the k coordinates. Specifically, this is a collection of permutations $\sigma_1, \ldots, \sigma_k$, where each $\sigma_i(1) = i$.

Strictly speaking, to estimate the size of the largest transitive subtournament in this class of construction, one must consider all elements of each permutation, because ties need to be broken. However, when n is large compared to k, the main asymptotic may be dictated by transitive subtournaments which are entirely resolved by the leading elements of every permutation (with no ties). This type of transitive subtournament is then a sequence of k-tuples in which every pair in the sequence strictly increases in a majority of the coordinates. When $k = 3$, this is precisely Problem 3.1.

3.2 Induced matchings and Ruzsa-Szemerédi

In light of the previous section, it suffices to focus on Problem 3.1, which asks to upper-bound the lengths of sequences of triples $L_1, L_2, \ldots$ from $\{1, 2, \ldots, n\}^3$, in which every pair of triples increases in at least two coordinates. The following trivial observation implies that the sequence length is at most n^2.

Observation 3.1. *For each fixed x^*, y^*, the sequence has at most one triple of the form (x^*, y^*, z).*

To improve upon this trivial bound $F(n) \leq n^2$, we uncover a nice connection to the celebrated Ruzsa-Szemerédi induced matching problem by using the sequence of triples to construct the following bipartite graph G. Let the vertex set be $\{a_1, \ldots, a_n, b_1, \ldots, b_n\}$. For each triple (x, y, z), place an edge between a_x and b_y, and label it with z.

Lemma 3.2. *For each label $z \in \{1, \ldots, n\}$, the set E_z of edges labeled z forms an **induced matching**: the subgraph of G (which includes all edges of all labels) induced by all vertices spanned by those edges of E_z is a matching.*

Proof. Fix z. By Observation 3.1, for each x, there is at most one triple of the form $(x, \star, z)$, and so each vertex a_x is incident to at most one edge labeled z. Similarly, each vertex b_y is incident to at most one edge labeled z. Therefore, in the subgraph of G induced by all vertices spanned by those edges of E_z, the edges of E_z form a matching.

It remains to show that no other edges (from other labels) are present in this induced subgraph. Indeed, suppose for contradiction that vertices a_x and b_y are both incident to edges labeled z, but the edge $a_x b_y$ itself is labeled $z' \neq z$. This means that the triple sequence contains the following triples, in some order (not necessarily the order presented):

$$(x, y', z)$$
$$(x', y, z)$$
$$(x, y, z')$$

Without loss of generality, suppose that the first two triples come in the relative order presented. Since both share the same third coordinate z, but must strictly increase in at least two coordinates, we have

$$x < x' \quad \text{and} \quad y' < y. \tag{1}$$

Similarly, since the first and third triples both agree in the first coordinate x, there must be a strict increase in both of the other two coordinates from one to the other. Since we already have $y' < y$ from (1), we must therefore also have $z < z'$.

Finally, since the second and third triples both agree in the second coordinate y, there must be a strict increase in both of the other two coordinates from one to the other. However, we already have $x' > x$ and $z < z'$, which are in opposite directions, which is a contradiction. □

We are now ready to prove our theorem.

Proof of Theorem 1.2. The induced matching property places us squarely in the context of Problem 1.2, which studies $M(n, k)$, the maximum number of edges in an n-vertex graph whose edge set is the union of k induced matchings. Indeed, we immediately have $F(n) \leq M(2n, n)$, because our bipartite graph G has exactly $F(n)$ edges. The original result of Ruzsa and Szemerédi [24] already breaks below the trivial quadratic bound, and the current best Triangle Removal Lemma bound of Fox [10] gives $M(2n, n) \leq O(n^2/e^{\log^* n})$, which, in light of Lemma 3.1, proves our result. □

3.3 Ruzsa-Szemerédi generalizations

It turns out that the bipartite graphs constructed from our triple sequences have richer properties beyond the induced matching phenomena. Indeed, the set of edges labeled with each fixed value z forms an "ordered" induced matching, as illustrated in Figure 1.

Observation 3.2. *For each label $z \in \{1, \ldots, n\}$, the edges labeled z are of the form (a_{i_1}, b_{j_1}), (a_{i_2}, b_{j_2}), ..., with $a_{i_1} < a_{i_2} < \cdots$ and $b_{j_1} < b_{j_2} < \cdots$.*

This additional ordering property alone is insufficient to significantly improve the result, because Ruzsa and Szemerédi's lower bound construction [24], based upon Behrend's [2, 7] three-term arithmetic progression free sets, actually already produces ordered induced matchings. Their graph has about $n^2/e^{\sqrt{\log n}}$ edges, so even if the additional order condition has an effect, it cannot improve the exponent on n. However, our collection of ordered induced matchings still satisfies an additional condition, which we call Σ-*free* (illustrated in Figure 2).

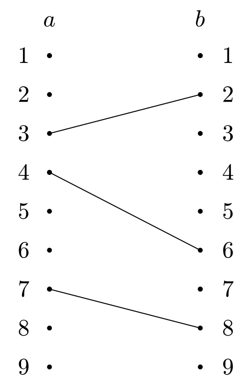

Figure 1: Ordered matching. Note that the matching edges do not cross each other.

Lemma 3.3. *There are no vertices b_h, a_i, b_j, a_k, b_l such that $i < k$ and $h \leq j \leq l$, edges $a_i b_h$ and $a_k b_l$ are both labeled z, and edges $a_i b_j$ and $a_k b_j$ are both in the graph (but may carry labels different from z).*

Remark. By choosing $j = l$ or $h = j$, we recover the ordered induced matching condition, so this is stronger than the traditional context of Ruzsa-Szemerédi.

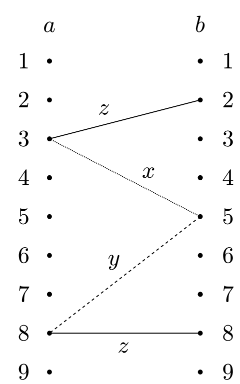

Figure 2: Prohibited structure in Σ-free matchings. The two outer edges both share the same label, but the connecting edges are differently labeled.

Proof of Lemma 3.3. Lemma 3.2 already established that the edge labels provide a partition into induced matchings, so we may assume that h, j, and l are distinct. Let x be the label of the edge $a_i b_j$. We then have two triples (i, h, z) and (i, j, x) with $h < j$, so since one triple must increase in at least two coordinates, this forces $z < x$ as well. Let y be the label of the edge $a_k b_j$. A similar argument considering (i, j, x) and (k, j, y) implies that $x < y$. However, the final edge $a_k b_l$ gives rise to (k, l, z), and when we compare that triple to (k, j, y), we find that they match in one coordinate, we already have deduced that $z < x < y$ and $j < l$, so neither can exceed the other in two coordinates. This contradiction establishes the claim. □

Lemma 3.3 now motivates Problem 1.3 from the introduction, which is a natural direction of generalization to the Ruzsa-Szemerédi induced matching problem: rather than having two edges of the same matching linked by a single edge of another matching, disallow edges of the same matching to be linked by a path of length up to l (possibly using edges of different labels). Recall that we defined $M_l(n,k)$ to be the maximum number of edges in a n-vertex bipartite graph whose edge set is the union of k matchings which are l-separated.

Lemma 3.3 actually provides an "ordered" condition which is more restrictive than the forbidden structures in Problem 1.3, so upper bounds on $M_l(2n,n)$ do not necessarily transfer over to our original question. However, due to its proximity to the well-studied problem of Ruzsa and Szemerédi, Problem 1.3 is of independent interest. We now proceed to prove the sharp upper bound on $M_l(2n,n)$ for all $l \geq 2$.

Proof of Theorem 1.3. We start by proving that $M_2(2n,n) \leq n^{3/2}$. Suppose that we have a $(2n)$-vertex bipartite graph which is a disjoint union of 2-separated matchings $E_1, \ldots, E_n$. Associate a set S_v of indices i to each vertex v, which records which E_i contain edges incident to v. Now, if there is a vertex v such that two distinct neighbors x and y are both incident to edges of the same E_i, then the 2-separation condition is violated. (Here we use the assumption that the graph is bipartite: they cannot both be incident to the same edge.) Thus, for each vertex w, its neighbors have disjoint S_v.

Since each E_i is a matching, each vertex is incident to at most one edge of each E_i. Therefore, for every vertex, the sum of the degrees of its neighbors is at most the number of matchings n; summing over all $2n$ vertices, the number of two-edge walks in the graph is at most $(2n)(n) = 2n^2$. Since this is also precisely the sum of the squares of the degrees, convexity implies

$$2n^2 = \sum_v d_v^2 \geq (2n)(\overline{d})^2,$$

where $\overline{d}$ is the average degree. Therefore, the average degree is at most $\sqrt{n}$, which implies that there are at most $\frac{1}{2}(2n)\sqrt{n}$ edges, and $M_2(2n,n) \leq n^{3/2}$ as claimed. Since $M_l(2n,n)$ is clearly monotonic in l, this then immediately implies that $M_l(2n,n) \leq n^{3/2}$ for all $l \geq 2$ as well.

For the lower bound, we will prove that this bound is exactly sharp for every perfect square n through an explicit construction. Start with the disjoint union of $\sqrt{n}$ copies of the complete bipartite graph $K_{\sqrt{n},\sqrt{n}}$. Each component has exactly n edges; assign one to each E_i. Now, every pair of distinct edges from the same E_i actually comes from two different connected components, and are therefore at infinite distance in the graph. Hence, for every perfect square n and every $l \geq 2$, we have $M_l(2n,n) = n^{3/2}$, which completes the proof. □

Remark. If the bound in Theorem 1.3 could be established in the weaker Σ-free setting, we would immediately achieve $F(n) \leq n^{3/2}$, which by Lemma 3.1 would produce the sharp bound $f(n) \geq n^{2/3}$ on our original Ramsey question on 1-color-avoiding paths.

4 Concluding remarks

This study has exposed many interesting problems, ranging from Ramsey theory, to digraphs, and on to Regularity. The original open problem is to improve the Ramsey bounds for 1-color-avoiding paths in 3-edge-colorings of transitive tournaments (Problem 1.1). It is also interesting to consider general tournaments. Our proof of Theorem 1.1 handily solved the two-color case using a completely different conceptual argument, and so it is not inconceivable that general tournaments could be handled by a direct approach. There may also be interesting phenomena when more than three colors are involved. Indeed, there is a whole spectrum of Ramsey-type problems where the notion of monochromatic slides across the scale from 1-color to 1-color-avoiding.

There are also interesting new questions which border on the area of Szemerédi's Regularity Lemma. Our Theorem 1.3 handles (unordered) l-separated matchings. It would be interesting to prove the analogue of the $M_2(2n, n)$ upper bound for Σ-free matchings (which are ordered). The Σ-free condition naturally generalizes to ordered notions of l-separation between matchings, and indeed, our graphs satisfy all properties along this successively richer hierarchy. Any improvement anywhere along this hierarchy would directly translate into improvements for our problem.

Acknowledgments

The author would like to thank Jacob Fox for interesting discussions on the topic of k-majority tournaments, which were the actual inspiration for this investigation. Also, several student staff members from the USA Math Olympiad Summer Program were involved in early work regarding the formulation in Problem 3.1, including Linus Hamilton, Jenny Iglesias, and Matt Superdock.

References

[1] N. Alon, G. Brightwell, H. Kierstead, A. Kostochka, and P. Winkler, Dominating sets in k-majority tournaments, *J. Combin. Theory Ser. B* **96** (2006), 374–387.

[2] F. Behrend, On sets of integers which contain no three terms in arithmetic progression, *Proc. Nat. Acad. Sci.* **32** (1946), 331–332.

[3] S. Burr and P. Erdős, On the magnitude of generalized Ramsey numbers for graphs, in Infinite and Finite Sets, Vol. 1 (Keszthely, 1973), 214–240, Colloq. Math. Soc. János Bolyai, Vol. 10, North-Holland, Amsterdam, 1975.

[4] V. Chvátal, Tree-complete graph Ramsey numbers, *J. Graph Theory* **1** (1977), 93.

[5] V. Chvátal, V. Rödl, E. Szemerédi, and W. Trotter, The Ramsey number of a graph with bounded maximum degree, *J. Comb. Theory Ser. B* **34** (1983), 239–243.

[6] D. Conlon, J. Fox, and B. Sudakov, *Combinatorica* **32** (2012), 513–535.

[7] M. Elkin, An improved construction of progression-free sets, *Israel Journal of Mathematics* **184** (2011), 93–128.

[8] P. Erdős and G. Szekeres, A combinatorial problem in geometry, *Compositio Math.* **2** (1935), 463–470.

[9] A. Figaj and T. Łuczak, The Ramsey number for a triple of long even cycles, *J. Combin. Theory Ser. B* **97** (2007), 584–596.

[10] J. Fox, A new proof of the graph removal lemma, *Ann. of Math* **174** (2011), 561–579.

[11] J. Fox, J. Pach, B. Sudakov, and A. Suk, Erdős–Szekeres-type theorems for monotone paths and convex bodies, *Proceedings of the London Mathematical Society*, **105** (2012), 953–982.

[12] T. Gallai, On directed graphs and circuits, in: **Theory of graphs (Proceedings of the Colloquium Tihany 1966**, Academic Press, New York (1968), 115–118.

[13] L. Gerencsér and Gyárfás, On Ramsey-type problems, *Annales Universitatis Scientiarum Budapestinensis, Eötvös Sect. Math.* **10** (1967), 167–170.

[14] R. Graham, V. Rödl, and A. Ruciński, On graphs with linear Ramsey numbers, *J. Graph Theory* **35** (2000), 176–192.

[15] A. Gyárfás, M. Ruszinkó, Sárközy, and E. Szemerédi, Three color Ramsey numbers for paths, *Combinatorica* **27** (2007), 35–69.

[16] M. Hasse, Zur algebraischen Begründung der Graphentheorie. I, *Mathematische Nachrichten* **28** (1965), 275–290.

[17] K. Kalmanson, On a theorem of Erdős and Szekeres, *Journal of Combinatorial Theory, Series A*, **15** (1973), 343–346.

[18] J. Komlós, A. Shokoufandeh, M. Simonovits, and E. Szemerédi, The regularity lemma and its applications in graph theory, in: Theoretical aspects of computer science (Tehran, 2000), Lecture Notes in Comput. Sci., 2292, Springer, Berlin, 2002, 84–112.

[19] D. McGarvey, A theorem on the construction of voting paradoxes, *Econometrica* **21** (1953), 608–610.

[20] K. Milans, D. Schreiber, and D. West, Acyclic sets in k-majority tournaments, *Electron. J. Combin.* **18** (2011), P122.

[21] G. Moshkovitz and A. Shapira, Ramsey theory, integer partitions and a new proof of the Erdős–Szekeres Theorem, *Advances in Math.* **262** (2014), 1107–1129.

[22] F. Ramsey, On a problem in formal logic, *Proc. London Math Soc.* **30** (1930), 264–286.

[23] B. Roy, Nombre chromatique et plus longs chemins d'un graphe, *Rev. Française Informat. Recherche Opérationnelle* **1** (1967), 129–132.

[24] I. Z. Ruzsa and E. Szemerédi, Triple systems with no six points carrying three triangles, *Proc. Colloq. Math. Soc. János Bolyai* **18** (1978), 939–945.

[25] M. Steele, Variations on the monotone subsequence theme of Erdős and Szekeres, In *Discrete probability and algorithms*, Springer (1995), 111–131.

[26] T. Szabó and G. Tardos, A multidimensional generalization of the Erdős–Szekeres lemma on monotone subsequences, *Combinatorics, Probability and Computing* **10** (2001), 557–565.

[27] L. Vitaver, Determination of minimal coloring of vertices of a graph by means of Boolean powers of the incidence matrix, *Doklady Akademii Nauk SSSR* **147** (1962), 758–759.