AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **66(2)** (2016), Pages 240–255

Density version of the Ramsey problem and the directed Ramsey problem

Zoltán Lóránt Nagy

MTA–ELTE Geometric and Algebraic Combinatorics Research Group
Eötvös Loránd University
H–1117 Budapest, Pázmány P. sétány 1/C
Hungary
nagyzoli@cs.elte.hu

Abstract

We discuss a variant of the Ramsey and the directed Ramsey problem. First, consider a complete graph on n vertices and a two-coloring of the edges such that every edge is colored with at least one color and the number of bicolored edges $|E_{RB}|$ is given. The aim is to find the maximal size f of a monochromatic clique which is guaranteed by such a coloring. Analogously, in the second problem we consider a semicomplete digraph on n vertices such that the number of bi-oriented edges $|E_{bi}|$ is given. The aim is to bound the size F of the maximal transitive subtournament that is guaranteed by such a digraph.

Applying probabilistic and analytic tools and constructive methods, we show that if $|E_{RB}| = |E_{bi}| = p\binom{n}{2}$, $(p \in [0, 1))$, then $f, F < C_p \log(n)$ where C_p only depends on p, while if $m = \binom{n}{2} - |E_{RB}| < n^{3/2}$ then $f = \Theta(\frac{n^2}{m+n})$. The latter case is strongly connected to Turán-type extremal graph theory.

1 Introduction

Ramsey's theorem concerns one of the classic questions in graph theory, aiming to determine accurate bounds on the Ramsey number $R(k)$, the smallest number n such that in any two-coloring of the edges of a complete graph on n vertices, there is guaranteed to be a monochromatic clique of size k. An inverse approach to this problem is the following. Let $c : E(K_n) \to \{\text{red, blue}\}$ be a 2-coloring of the edges of the complete graph K_n. What is the largest number $f(n)$ such that there exists a monochromatic clique of size $f(n)$ in any coloring c of K_n?

Clearly, $f(R(k)) = k$; moreover $t < R(k)$ implies $f(t) < k$.

Due to Erdős, Szekeres and Spencer [11, 14, 30], it is well-known that the following bounds hold for $f(n)$ if $n \geq 2$:

$$\frac{1}{2}\log_2(n) \leq f(n) \leq 2\log_2(n). \tag{1}$$

Although it is widely investigated, even the most significant improvements so far have not had any effect on the main terms, of both bounds [9, 30].

In this paper we propose a variant of this problem. Let us color the edges of K_n in such a way that every edge is either unicolored (red, blue), or bicolored. That is, our modified general two-coloring function is of form

$$c : E(K_n) \to \{\text{red, blue, red and blue}\}.$$

Let $E_{RB}(K_n)$ denote the set of bicolored edges of K_n, and $m = \binom{n}{2} - |E_{RB}(K_n)|$ denote the number of unicolored edges in a two-coloring of K_n. A subgraph H of K_n is called monochromatic in color red, respectively blue, if each edge of H is colored with red, respectively blue color. (That is, this extension allows a subgraph to be monochromatic in both colors, if every edge is bicolored.)

Problem 1.1. *Find the the size $f(n,m)$ of the largest monochromatic clique which is contained in any two-coloring of K_n with m unicolored edges.*

Our goal is to study the order of magnitude of $f(n,m)$ in terms of n and m, where m is considered as a function of n.

Our second problem concerns the directed Ramsey number $\vec{R}(n)$ which was defined by Erdős and Moser, and it had been studied by various authors [3, 12, 25, 27, 28, 31, 32]. Here the aim is to determine the size of the largest transitive subtournament which appears in any tournament on n vertices. A *tournament* T is a digraph in which every pair of vertices is joined by exactly one directed edge. A *subtournament* T' of T is a tournament induced by $V(T') \subseteq V(T)$. A tournament is *transitive* if it is acyclic. Recall that this is equivalent to the property that it has a topological ordering, that is, a linear ordering of its vertices such that for every directed edge $\vec{uv}$, u comes before v in the ordering.

The inverse approach appeared also in this problem, in the following sense. Let $F(n)$ denote the greatest integer F such that all tournaments of order n contain the transitive subtournament of order F. In [31], Stearns showed that $F(2n) \geq F(n)+1$ which provides $F(n) \geq \lfloor \log_2 n \rfloor + 1$. On the other hand, Erdős and Moser proved that $F(n) \leq \lfloor 2\log_2 n \rfloor + 1$, and conjectured that the lower bound of Stearns in fact holds with equality. However this turned out to be false [27]. Later the lower bound was improved by Sanchez-Flores [28], who proved $F(n) \geq \lfloor \log_2 (n/55) \rfloor + 7$ by determining the function $F(n)$ for values $F \leq 55$ using computer techniques, and applying Stearns' recursion. Note that although this result improves the lower bound on $F(n)$, it does not disprove the asymptotic $F(n) = (1+o(1))\log_2(n)$ version of the conjecture.

A *biorientation* of an undirected graph G is a digraph obtained from G by replacing each edge $\{v_i, v_j\}$ of G with either the arc $\overrightarrow{v_i v_j}$ or the arc $\overrightarrow{v_j v_i}$ or both of them. A *semicomplete digraph* is a biorientation of a complete graph K_n. If both directed edges lie in the edge set, then we call the undirected edge (v_i, v_j) bi-oriented. Hence, both tournaments and complete digraphs are examples for this digraph family.

Let G be a semicomplete digraph on n vertices. We call G transitive, if it contains a transitive subtournament T such that $V(T) = V(G)$. Let $E_{bi}(G)$ be the set of bi-oriented edges of the semicomplete digraph G, and $m = \binom{n}{2} - |E_{bi}(G)|$ denotes the number of edges of one orientation in G.

Following the concept of the general two-coloring and Problem 1.1, we study the following.

Problem 1.2. *Find the size $F(n,m)$ of the largest transitive tournament which is contained in any semicomplete digraph G of m edges of one orientation, on n vertices.*

The aim of the present paper is to study the order of magnitude of the considered functions $f(n,m)$ and $F(n,m)$ in terms of the cardinality of the bi-oriented or bi-colored edges which increases from $O(\log(n))$ to n while $|E_{RB}(K_n)|$, respectively $|E_{bi}(G)|$ increase until $\binom{n}{2}$.

This problem has a connection to the problem of Erdős and Rado too [13], studied also by Larson and Mitchell [24]. The question they discuss is the following. Given $N > 2$ and $k > 1$, what is the smallest n so that every digraph on a set of n vertices either has an independent set of size k or contains a transitive tournament TT_N on N vertices. We also mention the concept of Ramsey-Turán theory due to T. Sós and M. Simonovits, in which the function $RT(n, L, \mu)$, the maximum number of edges of an L-free graph on n vertices with independence number less than μ plays a key role (see e.g. [2, 29]). In the center of interest concerning this function, it is the asymptotic behavior of $RT(n, L, \phi(n))$ and its so called 'phase transition', that is, when and how the asymptotic behavior of $RT(n, L, \phi)$ changes sharply when we replace ϕ by a slightly smaller function ϕ'.

Our first result concerning the relation of the two functions in view is the following.

Proposition 1.3. $f(n,m) \leq F(n,m)$.

The following observation is straightforward.

Observation 1.4. *$f(n, m = \binom{n}{2}) = f(n)$, $f(n, m = 0) = n$,*
$F(n, m = \binom{n}{2}) = F(n)$, $F(n, m = 0) = n$,
$f(n,m) \leq f(n,m')$ and $F(n,m) \leq F(n,m')$ if $m' \leq m$ (monotonicity).

The main result of this paper describes some sort of phase transition as well. In the case $|E_{RB}(K_n)|$, respectively $|E_{bi}(G)|$ are small, and $m = \Omega(n^2)$, $f(n,m)$ and $F(n,m)$ are of order $\log n$, while $m = o(n^{3/2})$ implies that $f(n,m)$ is of order $\frac{n^2}{m+n}$. Also, seemingly random like extremal structures give their place to well-organized, Turán-type structures. Formally this can be expressed as below.

Notion 1.5. Let $m = \Theta(n^2)$; then introduce the parameter p as the density of the 2-colored or bi-oriented edges,

$$p := \frac{|E_{RB}(K_n)|}{|E(K_n)|} < 1 \text{ in Problem 1.1 and similarly } p := \frac{|E_{bi}(G)|}{|E(K_n)|} < 1 \text{ in Problem 1.2.}$$

We call a 2-coloring, respectively a semicomplete digraph G *p-dense*, if p is the ratio of the bicolored edges of K_n, respectively the bi-oriented edges of G. Suppose that $p \in [0,1)$ is fixed. Then $f_p(n)$ denotes the largest monochromatic clique that is guaranteed in a p-dense coloring, that is, $f_p(n) := f(n,(1-p)\binom{n}{2})$. $F_p(n)$ denotes the largest transitive tournament that is guaranteed in a p-dense digraph, that is, $F_p(n) := F(n,(1-p)\binom{n}{2})$.

Theorem 1.6.

$$F_p(n) \leq 2\left\lceil \frac{1}{1-p} \right\rceil \log_2(n(1-p))$$

and

$$\frac{1}{5}\frac{1}{1-p}\frac{1}{\log_2 \frac{2}{1-p}} \log_2 n \leq f_p(n), \quad \text{if} \quad p \geq 2/3.$$

Note that $\frac{1}{2}\log_2 n \leq f_p(n)$ holds for every p.

In the other case, let $m = o(n^2)$.

Theorem 1.7.

(i) If $m \leq n$, then

$$f(n,m) = n - \left\lfloor \frac{m}{2} \right\rfloor \quad \text{and} \quad F(n,m) = n - \left\lfloor \frac{m}{3} \right\rfloor.$$

(ii) If $m \geq n$, then

$$\text{(A)} \quad f(n,m) \geq \frac{n^2}{m+n} \quad \text{and} \quad \text{(B)} \quad F(n,m) \geq \frac{2n^2}{2m+n}.$$

(iii) If $m \leq n^{3/2}$, then

$$f(n,m) \leq \frac{n}{\lfloor \frac{m}{n} \rfloor + 1}.$$

Our paper is built up as follows. In Section 2, we begin with the proof of Proposition 1.3. Next we prove Theorem 1.6 applying constructive and probabilistic methods, and a quantitative variant of the Erdős–Szemerédi theorem [15].

In Section 3, we examine the case $m = o(n^2)$ and prove Theorem 1.7. First we discuss the case $m \leq n$ when exact results can be obtained. Then we prove the remaining statements of Theorem 1.7 (ii) and (iii). The lower bounds are derived from the Caro-Wei bound [4, 34] and the results of Alon, Kahn and Seymour on k-degenerate graphs [1], while the upper bounds are related to equitable colorings, graph packings, and the directed feedback vertex set problem. Finally, a slight improvement is presented on the upper bound of $F(n)$ using probabilistic techniques in Section 4.

2 The case $m = \Theta(n^2)$

In this section, we explain the connection between our two problems in consideration, then give upper bounds on the functions f and F when $m = \Theta(n^2)$.

Proof of Proposition 1.3. It is enough to prove that if a bound $F(n, m) < k$ holds for a $k \in \mathbb{Z}^+$, then $f(n, m) < k$ also holds. Consider a semicomplete digraph G having exactly m one-oriented edges, and assume that no transitive subtournament exists on k vertices in G. Label its vertices by $\{1, 2, \ldots, n\}$ arbitrary, and assign red color to ascending edges $\overrightarrow{ij}$ (where $i < j$), blue color to descending edges $\overrightarrow{ij}$ (where $i > j$). The coloring we thus obtain cannot contain a monochromatic clique of size k. Indeed, that would provide a transitive subtournament in G. $\square$

2.1 Probabilistic argument, upper bound on $F_p(n)$

Our aim is to generalize the theorem of Erdős and Moser [12] to obtain upper bounds on $F_p(n)$, to confirm Theorem 1.6 and prove another upper bound. To this end, we prove a lemma first. $\mathbb{E}(X)$ will denote the expected value of a random variable X.

Lemma 2.1. *Let $p \in [0, 1]$ and let Y be a random variable of distribution $Y \sim$ Binom(n, p) and Z be a random variable of distribution $Z \sim$ Hypergeom(N, pN, n). Then for every $c \geq 1$,*

$$\mathbb{E}(c^Z) \leq \mathbb{E}(c^Y).$$

Proof.

$$\mathbb{E}(c^Z) = \mathbb{E}\left((1 + c - 1)^Z\right) = \sum_{k=0}^{n} \mathbb{E}\binom{Z}{k}(c-1)^k,$$

and similar proposition holds for Y. It is enough to show that

$$\mathbb{E}\binom{Z}{k} \leq \mathbb{E}\binom{Y}{k},$$

since this inequality implies the lemma. For the binomial distribution, we have

$$\mathbb{E}\binom{Y}{k} = \sum_{t=0}^{n} \binom{t}{k} P(Y = t) = \sum_{t=k}^{n} \binom{t}{k} p^t (1-p)^{n-t} \binom{n}{t} = \binom{n}{k} p^k$$

in view of the binomial theorem and the identity

$$\binom{n}{t}\binom{t}{k} = \binom{n}{k}\binom{n-k}{t-k}.$$

On the other hand,

$$\mathbb{E}\binom{Z}{k} = \sum_{t=0}^{n} \binom{t}{k} P(Z = t) = \sum_{t=k}^{n} \binom{t}{k} \frac{\binom{pN}{t}\binom{(1-p)N}{n-t}}{\binom{N}{n}} = \frac{\binom{n}{k}\binom{pN}{k}}{\binom{N}{k}}.$$

Indeed, if we multiply the term $\frac{\binom{n}{k}\binom{pN}{k}}{\binom{N}{k}}$ by

$$1 = \sum_{t=k}^{pN} \frac{\binom{pN-k}{t-k}\binom{(1-p)N}{n-t}}{\binom{N-k}{n-k}},$$

the identity corresponding to the sum of the probabilities of a hypergeometric distribution $(N-k, pN-k, n-k)$, we end up at the equation

$$\sum_{t=k}^{n} \binom{t}{k} \frac{\binom{pN}{t}\binom{(1-p)N}{n-t}}{\binom{N}{n}} = \frac{\binom{n}{k}\binom{pN}{k}}{\binom{N}{k}} \sum_{t=k}^{pN} \frac{\binom{pN-k}{t-k}\binom{(1-p)N}{n-t}}{\binom{N-k}{n-k}}.$$

By expanding the binomial coefficients, we get that the expressions on the two sides are identical. Hence, the claim $\mathbb{E}(c^Z) \leq \mathbb{E}(c^Y)$ is equivalent to the straightforward inequality

$$\frac{\binom{pN}{k}}{\binom{N}{k}} \leq p^k \quad \text{for } p \in [0,1].$$

□

Theorem 2.2. *If $p < 1$, then*

$$F_p(n) < \frac{2}{\log_2 \frac{2}{1+p}} \log_2 n + 1 \quad \textit{holds.}$$

Proof. Let G_n be a random p-dense semicomplete digraph, that is, bi-oriented edges of cardinality $p\binom{n}{2}$ are chosen randomly and the remaining edges are oriented at random, independently from the others. Let A_i be the event that a given k-subset $X_i^{(k)}$ of $V(G_n)$ induces a transitive semicomplete digraph, and A_i^* be the variable that counts distinct transitive subtournaments in $X_i^{(k)}$. Note that $P(\bigcap_i \overline{A_i}) > 0$ implies $F_p(n) < k$.

By definition, the event $\overline{A_i}$ is equivalent to the event $\{A_i^* = 0\}$, so by Markov's inequality, we get that $\mathbb{E}(\sum_i A_i^*) < 1$ implies $F_p(n) < k$.

We evaluate the expected value by separating all orderings of $X^{(k)}$ for a possible transitive tournament, as follows.

$$\mathbb{E}(\sum_i A_i^*) = \sum_i \mathbb{E}(A_i^*) = \binom{n}{k} k! \sum_j \frac{2^j}{2^{\binom{k}{2}}} P(j \text{ bi-oriented edges exist in } X_i^{(k)}).$$

Let Z be a random variable, which counts the bi-oriented edges in the edge set of $X_i^{(k)}$. Clearly, Z has distribution $Z \sim \textit{Hypergeom}\left(\binom{n}{2}, p\binom{n}{2}, \binom{k}{2}\right)$. According to Lemma 2.1, $\mathbb{E}(c^Z) \leq \mathbb{E}(c^Y)$, where $Y \sim \textit{Binom}\left(\binom{k}{2}, p\right)$. Applying this in the former equality with $c = 2$, we get

$$\mathbb{E}(\sum_i A_i^*) \leq \binom{n}{k} \frac{k!}{2^{\binom{k}{2}}} \sum_j 2^j p^j (1-p)^{\binom{k}{2}-j} \binom{\binom{k}{2}}{j} = \binom{n}{k} k! \left(\frac{1+p}{2}\right)^{\binom{k}{2}}.$$

Hence if $n^k < \left(\frac{2}{1+p}\right)^{\binom{k}{2}} \Leftrightarrow 2k\log_2 n < k(k-1)\log_2 \frac{2}{1+p}$ holds for k, then $F_p(n) < k$, that is, $F_p(n) < \frac{2}{\log_2 \frac{2}{1+p}} \log_2 n + 1$. □

Theorem 2.2 implies that for any fixed density $p < 1$, both $f_p(n)$ and $F_p(n)$ is of order $\log n$. The following theorem yields the same conclusion based on a construction.

Theorem 2.3. *Let $t \le n$ be a positive integer, and $p := 1 - \frac{1}{t}$. Then $F_p(n) \le tF_0(\lceil \frac{n}{t} \rceil) \le 2t\log_2(\frac{n}{t})$.*

Proof. Partition the vertices into t classes of size $\lceil \frac{n}{t} \rceil$ or $\lfloor \frac{n}{t} \rfloor$. Let each class span an oriented graph containing no transitive tournaments of size more than $F_0(\lceil \frac{n}{t} \rceil)$. Finally, let the edges between the partition classes be bi-oriented. This digraph contains no transitive tournaments of size more than $tF_0(\lceil \frac{n}{t} \rceil)$, while a simple calculation shows that the density of the bi-oriented edges is at least p. The bound thus follows from inequality (1). □

Remark 2.4. In view of Observation 1.4, Theorem 2.3 can be extended to any $p \in (0, 1)$ by applying the theorem for an integer t for which $p < 1 - \frac{1}{t}$ holds.

Observation 2.5. Theorem 2.2 is useful if p is rather small, while the upper bound of Theorem 2.3 is better if $p > 1/2$. Note that Theorem 1.6, part I is a consequence of Theorem2.3.

2.2 The Erdős–Szemerédi argument, lower bound on $f_p(n)$

In their celebrated paper [15], Erdős and Szemerédi presented a variant of the Ramsey problem, where a dependence on the density of the graph is taken into consideration.

Theorem 2.6. *[15] Let G be a graph on n vertices and $\frac{1}{k}\binom{n}{2}$ edges. Then either G or its complement contains a complete graph K_a with $a > C\frac{k}{\log k}\log n$, where C is a constant independent of k and n.*

Following their ideas and the exact result of Čulik, the following quantitative form can be confirmed.

Theorem 2.7. *Let G be a graph on n vertices and $\frac{1}{k}\binom{n}{2}$ edges with $k \ge 6$. Then either G or its complement contains a complete graph K_a with $a > \frac{1}{10}\frac{k}{\log k}\log n$.*

(Note that $\log n$ denotes the natural logarithm.)

Observe that this assertion indeed implies the lower bound of Theorem 1.6, hence

$$\frac{1}{5(1-p)}\frac{1}{\log \frac{2}{1-p}} \log n \le f_p(n)$$

holds if $p \ge 2/3$.

Proof of Theorem 1.6, lower bound. We can assume that the number of edges colored with only red color is at most $\frac{1-p}{2}\binom{n}{2}$. Apply Theorem 2.7 with $k = \frac{2}{1-p}$ where $p > 2/3$ (ensuring $k > 6$). It yields a suitably large monochromatic clique. □

To prove the quantitative variant of the Erdős–Szemerédi theorem, first recall the Zarankiewicz problem about determining the function $\mathrm{z}(\mu, \nu, s, t)$.

Definition 2.8. Given positive integers μ, ν, s, t, let $\mathrm{z}(\mu, \nu, s, t)$ be the maximum number of ones in a $(0, 1)$ matrix of size $\mu \times \nu$ that does not contain an all ones submatrix of size $s \times t$.

Equivalent definition is the following

Definition 2.9. A bipartite graph $G = (A, B; E)$ is $K_{s,t}$-free if it does not contain s nodes in A and t nodes in B that span a subgraph isomorphic to $K_{s,t}$. The maximum number of edges a $K_{s,t}$-free bipartite graph of size (μ, ν) may have is denoted by $\mathrm{z}(\mu, \nu, s, t)$, and is called Zarankiewicz number with the corresponding parameters.

The order of magnitude of the Zarankiewicz numbers in general was obtain by Kővári, Sós and Turán, and reads as follows.

Theorem 2.10. *[23]* $\mathrm{z}(\mu, \nu; s, t) \leq (s-1)^{1/t}(\nu - t + 1)\mu^{1-1/t} + (t-1)\mu$

It was later improved by many authors, see [19, 26].

However, in the special case when one color class is much larger than the other, the value can be determined exactly in view of an early result of Čulík.

Theorem 2.11 (Čulík [10])**.** *If* $1 \leq s \leq \mu$ *and* $(t-1)\binom{\mu}{s} \leq \nu$*, then*

$$\mathrm{z}(\mu, \nu, s, t) = (s-1)\nu + (t-1)\binom{\mu}{s}.$$

Proof of Theorem 2.7. Let G has n vertices, and average degree $\overline{d} = \frac{n}{k}$. Then a subgraph $G' \subseteq G$ exists such that $\forall v \in V(G') : d_{G'}(v) \leq \frac{2n}{k}$ and $|V(G')| = n/2 =: n'$.

Consider a largest maximal independent set I in G'. We may assume that $\mu = |I| \leq \frac{1}{10}\frac{k}{\log k}\log n$, otherwise we are done. Let $W := V(G') \setminus I$ and let H denote the complement of the bipartite graph induced by I and W w.r.t. the complete bipartite graph on the same clusters. Applying Theorem 2.11 of Čulík, we obtain a large complete subgraph $K_{s,t}$ in H.

Indeed, $E(H) > \mu(n' - \mu - \frac{4n'}{k})$ where $\mu \leq \frac{1}{10}\frac{k}{\log k}\log n$, while setting $s = (1-5/k)\mu \leq (1-\ 5/k)\ C\frac{k}{10\log k}\log n$ implies

$$\mathrm{z}(\mu, n' - \mu, s, t) = (s-1)(n' - \mu) + (t-1)\binom{\mu}{s}$$

if the conditions of the theorem are verified. To this end, set $t = \sqrt{n}$, and consider the bound

$$\binom{\mu}{s} = \binom{\mu}{\mu - s} \leq \left(\frac{\mu \cdot e}{\mu - (1-5/k)\mu}\right)^{\mu-(1-5/k)\mu} \leq \left(\frac{k \cdot e}{5}\right)^{\frac{5}{10}\frac{\log n}{\log k}} < \sqrt{n'/4},$$

thus $(t-1)\binom{\mu}{s} < \sqrt{n}\sqrt{n'/4} \leq n' - \mu$.

Easy computation shows that $\mathrm{z}(\mu, n'-\mu, s, \sqrt{n}) < s(n'-\mu) < \mu(n'-\mu-\frac{4n'}{k}) < E(H)$ holds, as

$$\frac{s}{\mu} = (1-5/k) < 1 - \frac{4n'}{k(n'-\mu)}.$$

Hence in graph H, there exists a set T of $\sqrt{n}$ points in $W \subset V(H)$ which are connected to a set S of s points of I. $\omega(H)$ cannot exceed $\mu - s$, as I was a maximal independent set in G'. Consequently, bounds on the Ramsey numbers gives lower bound on the independence number of T and the clique number of G' as well. Namely, using the bound of Erdős and Szekeres [14],

$$R(\mu - s, \mu) < \binom{2\mu - s}{\mu - s} < \left(\frac{(2\mu - s)\cdot e}{\mu - s}\right)^{\mu - s} \leq \left(\frac{(k+5)\cdot e}{5}\right)^{\frac{5}{10}\frac{\log n}{\log k}} < k^{\frac{1}{2}\frac{\log n}{\log k}} = \sqrt{n},$$

if $k \geq 6$. This bound holds for every $\mu \leq \frac{1}{10}\frac{k}{\log k}\log n$, thus $H\mid_T$ must contain an independent set of size at least $\frac{1}{10}\frac{k}{\log k}\log n$, completing the proof. □

3 The case $m = o(n^2)$

The cardinality of the bioriented or bicolored edges is assumed to be $(1-o(1))\binom{n}{2}$ from now on.

This section is built up as follows. First we recall some Turán-type lemmas, which will be key ingredients in the proofs of Theorem 1.7 (i) and (ii). We continue with the proofs of the upper and lower bounds of Theorem 1.7, distinguishing between the cases $m \leq n$ and $m \geq n$.

3.1 Preliminary lemmas

Theorem 3.1 (Turán theorem, Caro-Wei bound)**.** *[4, 33, 34] Let H be a simple graph with degree sequence $0 < d_1 \leq \ldots\ldots \leq d_n$. Then there exists an induced subgraph H' of H on $n' \geq \sum_{i=1}^{n} \frac{1}{d_i+1}$ nodes, containing no edges. That is,*

$$\alpha(H) \geq \sum_{i=1}^{n} \frac{1}{d_i + 1}$$

holds for the independence number of H.

Corollary 3.2. *[20] Let H be a simple graph on n vertices and of m edges. Then*

$$\alpha(H) \geq \frac{n}{2m/n+1}.$$

The Caro-Wei bound has been generalized in many ways. Caro, Hansberg and Tuza studied [5, 6] d-independent subsets $S \subseteq V$ of the vertex set $V = V(G)$ of a graph G, which are sets of vertices such that the maximum degree in the graph induced by S is at most d. Alon, Kahn and Seymour called a graph H d-degenerate if every subgraph of it contains a vertex of degree smaller than d and bound the maximum number $\alpha_d(G)$ of vertices of an induced d-degenerate subgraph of G in [1]. Observing that $\alpha_1(G) = \alpha(G)$ and 2-degenerate graph are forests, we can formulate their result - similarly to Theorem 3.1 - as follows.

Lemma 3.3. *[1] Let H be a simple graph with degree sequence $0 < d_1 \leq \ldots\ldots \leq d_n$. Then there exists an induced subgraph H' of H on*

$$n' \geq \sum_{i=1}^{n} \frac{2}{d_i+1}$$

nodes, containing no cycles.

For the sake of completeness we give a short proof.

Proof. Take an arbitrary arrangement of the vertices of H. Label those vertices which have at most one preceding neighbour in the permutation. It immediately follows that labeled vertices cannot span a cycle. If the arrangement is chosen uniformly at random, then every vertex i gets a label with probability $\frac{2}{d_i+1}$, thus the expected value of the labeled vertices is exactly $\sum_{i=1}^{n} \frac{2}{d_i+1}$. This implies the existence of a suitable subgraph H'. □

If $n \leq m$, Lemma 3.3 implies

Corollary 3.4. *[1] Let H be a simple graph on n vertices and of $n \leq m$ edges. Then*

$$\alpha_2(H) \geq \frac{2n}{2m/n+1}.$$

3.2 Subcase $m \leq n$

We start with the case $m \leq n$ to demonstrate the proof techniques of Theorem 1.7, underline the difference between the functions $f(n, m)$ and $F(n, m)$ in consideration, furthermore to determine explicitly its values via Theorem 1.7 (i).

Proof of Theorem 1.7 (i). First we prove $f(n, m) = n - \left\lfloor \frac{m}{2} \right\rfloor$. Consider those edges which are colored with precisely one color. Assume that the color red is assigned to at least as many edges as the color blue. Thus there are at most $\left\lfloor \frac{m}{2} \right\rfloor$ blue edges.

Delete a minimal covering (vertex) set of the blue edges. The remaining vertex set induces a graph where the color red is assigned to every edge. On the other hand, if both the blue and the red edge set is independent, equality holds.

Next, we prove inequality $F(n,m) \leq n - \left\lfloor \frac{m}{3} \right\rfloor$ by a construction. Take $\left\lfloor \frac{m}{3} \right\rfloor$ disjoint triangles, and orient round the edges in each triangle. Let all the edges between different triangles be bi-oriented edges. Clearly, at most is sharp. Indeed, let H be the simple graph constructed from G by deleting the bi-oriented edges, and omitting the orientations of the remaining oriented edges and the isolated vertices. Let $0 < d_1 \leq \ldots.. \leq d_s$ be the degree sequence of H. We can obtain a subgraph H' on at least $\sum_{i=1}^{s} \frac{2}{d_i+1}$ nodes which do not span cycles, according to Lemma 3.3. Consider the directed subgraph of G restricted to the vertex set corresponding to the vertex set of H'. Since H' was a forest, the graph obtained from G by deleting the bi-oriented edges and restricted to the vertex set of H' has a topological ordering. This provides a transitive subtournament of G on $\left\lceil \sum_{i=1}^{s} \frac{2}{d_i+1} \right\rceil$ nodes. Hence Lemma 3.5 completes the proof. □

Lemma 3.5. *Let G be a simple graph with degree sequence $0 < d_1 \leq \ldots \leq d_s$. If G has m edges and $s \leq n$, then*

$$\sum_{i=1}^{s} \frac{2}{d_i+1} + n - s \geq n - \frac{m}{3}.$$

Proof. Observe that the following inequality holds for every positive integer x:

$$\frac{2}{x+1} - 1 \geq -\frac{x}{6}.$$

Since $\sum_{i=1}^{s} d_i = 2m$, summing it for every degree of G confirms the statement. □

3.3 Subcase $n \leq m$

Next we show that $n \leq m$ implies that (A) $f(n,m) \geq \frac{n}{m/n+1}$ and (B) $F(n,m) \geq \frac{2n}{2m/n+1}$ hold.

Proof of Theorem 1.7(ii). Part (A). Consider the edges which have exactly one color, and suppose that the color red is assigned to at least as many edges as the color blue. Thus the cardinality of the blue edges is at most $m/2$. Let us take a maximal independent set on the graph of the blue edges. This provides a monochromatic (red) subgraph in the original K_n. The cardinality of the maximal independent set is at least $\frac{n}{m/n+1}$ due to Corollary 3.2.

Part (B). Apply Corollary 3.4 to the simple graph H obtained from our digraph G by deleting all bi-oriented edges and omitting the orientation for the rest of the edges. Consider the obtained forest subgraph H' of H, and the original orientations of the edges of H'. These edges in G obviously provides an acyclic orientation, with a suitable topological ordering, hence the vertex set of H' guarantees a transitive subtournament on at least $\frac{2n}{2m/n+1}$ vertices. □

We highlight the connection between the feedback vertex set problems (deadlock recovery) and Theorem 1.7 (ii)(B). A feedback vertex set of a graph (or digraph) is a set of vertices whose removal leaves a graph without (directed) cycles. In other words, each feedback vertex set contains at least one vertex of any cycle in the graph. Since the feedback vertex sets play a prominent role in the study of deadlock recovery in operating systems, it has been studied extensively [7, 8, 17, 18, 22]. To find the minimal feedback vertex set is NP-complete in the undirected and directed case as well [22].

Essentially, $F(n,m)$ is equal to the size of the maximal size of an acyclic subgraph which is guaranteed to be in a digraph of m directed edges on n vertices, that is, $n - F(n,m)$ is the minimal size of a feedback vertex set, such that the removal of an appropriate set of $n - F(n,m)$ vertices makes any digraph of m edges acyclic. Instead, in the proof of Theorem 1.7 (ii)(B) we bound the size of the minimal size of a feedback vertex sets of undirected graphs of m edges, which generally provides a fairly weaker bound, even if they both lead to exact results when m is small via Theorem 1.7 (i).

Before stating upper bounds for $f(n,m)$ and prove Theorem 1.7 (iii), recall that $\vec{R}(n)$ denoted the directed Ramsey number, and $F(\vec{R}(k+1)-1) = k$.

Proposition 3.6. *Let $|E_{bi}(G)| = \binom{n}{2} - m$, where $m = n \cdot (\frac{\vec{R}(k+1)}{2} - 1)$. Suppose that $\vec{R}(k+1) - 1 \mid n$. Then*

$$F(n,m) \leq \frac{k}{\vec{R}(k+1)-1} \cdot n.$$

Proof. There exists a tournament T on $\vec{R}(k+1)-1$ vertices which does not contain a transitive tournament of size $k+1$. Take $\frac{n}{\vec{R}(k+1)-1}$ disjoint copy of it, and join every pair of vertices with bi-oriented edges which are not in the same copy of T. Clearly, the size of the maximal transitive tournament of this graph is at most $k \cdot \frac{n}{\vec{R}(k+1)-1}$. □

Proposition 3.7. *Let $|E_{RB}(G)| = \binom{n}{2} - m$, and let C denote $\lfloor m/n \rfloor$, and suppose $C + 1 \leq n/(C+1) \in \mathbb{Z}^+$. Then*

$$f(n,m) \leq \frac{n}{C+1}.$$

Proof. Let r be the remainder in the division $n : (C+1)$. Take the disjoint union of $\lfloor n/(C+1) \rfloor$ cliques of size $C+1$, and a clique of size r and call it K. Observe first that the number of edges is at most $m/2$ in K. Consider now K' and K'', two copies of K. Since $C+1 \leq n/(C+1)$, it is easy to see that one can pack them into a complete graph K_n, i.e. there exists an edge-disjoint placement of them onto the same set of n vertices. In fact, this is a very special case of the famous Hajnal-Szemerédi theorem [21], which states that if H is an n-vertex graph with $\Delta(G) \leq q$, then H packs with the graph H' whose components are complete graphs of size $\lfloor n/(q+1)\rfloor$ or $\lceil n/(q+1)\rceil$. Color the edges of K' blue and the edges of K'' red. Hence, if all the other edges of K_n are colored with both colors, then at most m will be the number of unicolored edges, while a monochromatic clique has at most $\lfloor n/(C+1)\rfloor + 1$ vertices due to the pigeon-hall principle. □

As a consequence Theorem 1.7 (iii) follows.

Theorem 1.7 (ii) and 3.7 implies together that equality can be attained in the two-colored case, if some divisibility conditions hold and m is not too large.

Corollary 3.8. *If $m/n = C \in \mathbb{Z}^+$, and $C+1 \leq n/(C+1) \in \mathbb{Z}^+$, then*

$$f(n,m) = \frac{n}{m/n+1}.$$

This method thus provides exact result till $m \leq n^{3/2}$ holds.

If the order of magnitude is bigger than $n^{3/2}$, the graph cannot contain a disjoint union of packings of monochromatic cliques of size $C+1$, from both red and blue color, hence the lower bound cannot be attained.

4 Improving the upper bound for $F(n)$

In the following section, we consider simple oriented graphs.

Lemma 4.1 (Lovász Local Lemma)**.** *[16] Let $A_1, A_2, \ldots, A_m$ be a series of events such that each event occurs with probability at most P and such that each event is independent of all the other events except for at most d of them. If $Pe(d+1) < 1$, then there is a nonzero probability that none of the events occurs.*

Theorem 4.2. $F(n) < 2\log_2 n - 1 + o(1)$.

Proof. Fix $n \geq 55$, and let $\sqrt{n} > k > 4$. Let T_n be a random tournament such that every edge is oriented in one direction uniformly at random, independently from the orientations of all the other edges. Let A_i be the event that a given k-subset $X_i^{(k)}$ of $V(T_n)$ is a transitive tournament. Clearly, $P(A_i) = \frac{k!}{2^{\binom{k}{2}}}$.

A_i is independent of all events A_j for which $|X_i^{(k)} \cap X_j^{(k)}| \leq 1$. This yields

$$d \leq \binom{n}{k} - \binom{n-k}{k} - k\binom{n-k}{k-1} = \sum_{j=2}^{k} \binom{k}{j}\binom{n-k}{k-j}.$$

Increasing the right hand side, we obtain

$$d < \binom{k}{2}\binom{n-k}{k-2}\left(1 + \left[\frac{k-2}{3}\frac{k-2}{(n-2k+3)}\right] + \left[\frac{k-2}{3}\frac{k-2}{(n-2k+3)}\right]^2\right.$$

$$+\left[\frac{k-2}{3}\frac{k-2}{(n-2k+3)}\right]^3+\dots\Bigg).$$

Hence

$$d<\binom{k}{2}\binom{n-k}{k-2}\left(\frac{1}{1-\frac{(k-2)^2}{3(n-2k+3)}}\right).$$

Applying $\sqrt{n}>k>4$, we get that the sum of the above geometric progression is less than 1.2.

In view of the Lovász Local Lemma, if $e(d+1)P(A_i)<1$, then P(no transitive k-subtournaments in T_n) >0, thus $f(n)<k$.

Therefore if

$$e\cdot\frac{k^2}{2}\frac{n^{k-2}}{(k-2)!}1.2\cdot\frac{k^2(k-2)!}{2^{\binom{k}{2}}}<1,$$

then $F(n)<k$. This implies

$$(k-2)\log_2 n<\binom{k}{2}-4\log_2 k-0.7.$$

Consequently,

$$2\log_2 n-1<k+\frac{2-8\log_2 k-1.4}{(k-2)},$$

which completes the proof. □

Acknowledgments

I am very grateful to Ágnes Tóth, András Gyárfás and Yair Caro for their valuable suggestions, and to the anonymous referees for their very constructive comments which helped me improve the manuscript significantly.

References

[1] N. Alon, J. Kahn and P. D. Seymour, Large Induced Degenerate Subgraphs, *Graphs Combin.* **3** (1987), 203–211.

[2] J. Balogh, P. Hu and M. Simonovits, Phase transitions in Ramsey-Turán theory, *J. Combin. Theory Ser. B* **114** (2015), 148–169.

[3] J. Bang-Jensen and G. Gutin, *Digraphs: Theory, Algorithms and Applications*, Springer (2009).

[4] Y. Caro, New results on the independence number, *Tech. Report*, Tel-Aviv University, (1979).

[5] Y. Caro and A. Hansberg, New approach to the k-independence number of a graph, *Elect. J. Comb.* **20**(1) (2013), P33.

[6] Y. Caro and Z. Tuza, Improved lower bounds on k-independence, *J. Graph Theory* **15** (1) (1991), 99–107.

[7] J. Chen, F. V. Fomin, Y. Liu, S. Lu and Y. Villanger, Improved algorithms for feedback vertex set problems, *J. Comp. Sys. Sciences* **74** (7) (2008), 1188–1198.

[8] J. Chen, Y. Liu, S. Lu, B. O'Sullivan and I. Razgon, A fixed-parameter algorithm for the directed feedback vertex set problem, *J. ACM* **55** (5) Art. 21, (2008).

[9] D. Conlon, A new upper bound for diagonal Ramsey numbers, *Ann. Math.* **170** (2) (2009), 941–960.

[10] K. Čulík, Teilweise Lösung eines verallgemeinerten Problem von K. Zarankiewicz, *Ann. Soc. Polon. Math.* **3** (1956), 165–168.

[11] P. Erdős, Some remarks on the theory of graphs, *Bull. Amer. Math. Soc.* **53** (4) (1947), 292–294.

[12] P. Erdős and L. Moser, On the representation of directed graphs as unions of orderings, *Publ. Math. Inst. Hungar. Acad. Sci.* **9** (1964), 125–132.

[13] P. Erdős and R. Rado, Partition relations and transitivity domains of binary relations, *J. London Math. Soc.* **1** (1) (1967), 624–633.

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

[15] P. Erdős and E. Szemerédi, On a Ramsey type problem, *Periodica Math. Hung.* **2** (1-4) (1972), 295–299.

[16] P. Erdős and L. Lovász, Problems and results on 3-chromatic hypergraphs and some related questions, *Colloq. Math. Soc. J. Bolyai* 11, North Holland, Amsterdam, (1975), 609–627.

[17] P. Erdős and L. Pósa, On independent circuits contained in a graph, *Canadian J. Math.* **17** (1965), 347–352.

[18] P. Festa, P. M. Pardalos and M. G. C. Resende, Feedback set problems, In: *Handbook of Combinatorial Optimization, Supplement A*, Kluwer Academic, Dordrecht (1999), pp. 209–258.

[19] Z. Füredi, An upper bound on Zarankiewicz's problem, *Comb. Probab. Comput.* **5** (1996), 29–33.

[20] J. R. Griggs, Lower bounds on the independence number in terms of the degrees, *J. Combin. Theory Ser. B* **34** (1983), 22–39.

[21] A. Hajnal and E. Szemerédi, Proof of a conjecture of Erdős, *Combinatorial Theory and its Applications* Vol. II, North-Holland, New York-London (1970), 601–603.

[22] R. M. Karp, Reducibility among combinatorial problems, In: *Complexity of Computer Computations*, Plenum, New York (1972), pp. 85–103.

[23] T. Kővári, V. T. Sós and P. Turán, On a problem of K. Zarankiewicz, *Colloquium Math.* **3** (1954), 50–57.

[24] J. A. Larson and W. J. Mitchell, On a problem of Erdős and Rado, *Ann. Combin.* **1** (1) (1997), 245–252.

[25] V. Neumann-Lara, A Short Proof of a Theorem of Reid and Parker on Tournaments, *Graphs Combin.* **10** (1994), 363–366.

[26] V. Nikiforov, A contribution to the Zarankiewicz problem, *Lin. Alg. Appl.* **432** (6) (2010), 1405–1411.

[27] K. B. Reid and E. T. Parker, Disproof of a conjecture of Erdős and Moser on tournaments, *J. Combin. Theory* **9** (1970), 225–238.

[28] A. Sanchez-Flores, On tournaments free of large transitive subtournaments, *Graphs Combin.* **14** (1998), 181–200.

[29] M. Simonovits and V. T. Sós, Ramsey-Turán theory, *Discrete Math.* **229** (1) (2001), 293–340.

[30] J. Spencer, Ramsey's theorem—a new lower bound, *J. Combin. Theory Ser. A* **18** (1975), 108–115.

[31] R. Stearns, The voting problem, *Amer. Math. Monthly* **66** (1959), 761–763.

[32] C. Tabib, About the inequalities of Erdős and Moser on the largest transitive subtournament of a tournament, *Combinatoire énumérative*, Lec. Notes in Math. 1234, Berlin Heidelberg New York: Springer (1986), 308–320.

[33] P. Turán, An extremal problem in graph theory (in Hungarian), *Mat. Fiz. Lapok* **48** (1941), 436–452.

[34] V. K. Wei, A lower bound on the stability number of a simple graph, *Tech. Memorandum* No. 81-11217-9, Bell Laboratories, 1981.

(Received 28 Jan 2016; revised 7 Sep 2016)