Extremal Structural Results for Feedback Arc Sets and Graph Inversions

Kai Lum [1,2]

[1]MIT PRIMES-USA 2025
[2]Basis Independent Fremont

Abstract

In a digraph, a feedback arc set is a set of edges whose removal eliminates every directed cycle, and the minimum size of such a set is denoted by $\beta(G)$. A digraph is r-free if it contains no directed cycles of length at most r. In this paper, we investigate the minimum feedback arc set in digraphs that are $(r-1)$-free. We prove that $\beta(G) \leq 1$ if $r > \lfloor \frac{2n}{3} \rfloor$, and $\beta(G) \leq 2$ if $r > \frac{n}{2}$ with a forbidden structure. We also present an efficient linear-time algorithm to identify the minimum feedback arc set when $\beta(G) = 1$. For tournaments, we further refine the extremal parameter $\text{inv}_k(n)$. It is the minimum number of inversions required to transform an n-vertex tournament into an acyclic tournament, where each step involves reversing all edges within a subset of at most k vertices. We improve the known upper bound for $\text{inv}_4(n)$ using techniques involving Ramsey numbers with monochromatic subgraph structures, and the bound for $\text{inv}_k(n)$ with two different approaches.

1 Introduction

In this paper, every directed graph (abbreviated as *digraph*) is oriented without loops, parallel, or antiparallel edges. In a digraph G, a *directed cycle* C consists of vertices $V = \{v_1, v_2, \dots, v_{|C|}\}$ such that (v_i, v_{i+1}) are edges for all $i < |C|$, and $(v_{|C|}, v_1)$ is also an edge. A digraph is *acyclic* if it has no directed cycles. A *tournament* is a directed graph obtained by assigning a direction to every edge of a complete graph. In this paper, we study extremal problems involving edge removal or reversal to make a graph acyclic.

Definition 1.1. Given a digraph $G = (V, E)$, a *feedback arc set* is a set of edges $S \subset E$ where $G' = (V, E \setminus S)$ is acyclic. A *minimum feedback arc set* is a feedback arc set of minimum size. Let $\beta(G)$ denote the size of the minimum feedback arc set of G.

Definition 1.2. A digraph G is *r-free* if it does not contain any directed cycle C where $|C| \leq r$.

Directed graphs provide a way to model systems whose elements have one-way dependencies, such as priority orderings or causal relationships. In real-world applications, cycles typically represent inter-dependencies between elements, significantly impacting the underlying structure and making the study of cyclic complexity an important combinatorial problem. A standard measure for this complexity is the size of the minimum feedback arc set, denoted as $\beta(G)$, which indicates how far the digraph is from being acyclic, thus quantifying its cyclic complexity. The minimum feedback arc sets were first studied by Slater [18] on statistically motivated feedback arc minimization for tournaments. Feedback arc sets have applications in areas such as circuit retiming, tournament ranking, precedence-constrained scheduling, and Bayesian-network learning [13, 17, 8].

Definition 1.3. Given a digraph $G = (V, E)$ with vertices $v_1, v_2, \ldots, v_n$, an edge $v_i \to v_j$ is a *forward edge* if $i < j$, otherwise it is a *backward edge.*

For all digraphs G with m edges, $\beta(G) \leq \frac{m}{2}$ can be obtained by splitting the edges into forward and backward edges. Removing the smaller of the two sets creates an acyclic digraph, proving $\beta(G) \leq \frac{m}{2}$. To further improve the bound, Poljak, Rödl, and Spencer [14] show $\beta(T) \leq \frac{m}{2} - cm^{\frac{3}{4}}$ for a constant $c > 0$ on a tournament T. Berger and Shor [2] provide a more general proof for all digraphs G with m edges and Δ being the maximum degree, where $\beta(G) \leq \frac{m}{2} - c\frac{m}{\sqrt{\Delta}}$. Despite progress made for digraphs that exclude short directed cycles [5, 3], optimal bounds for $\beta(G)$ when G is $(r-1)$-free remain open. This problem space includes a more general conjecture by Sullivan, who proposed that all r-free digraphs satisfy

$$\beta(G) \leq \frac{2\gamma(G)}{(r+1)(r-2)} \text{ [19].}$$

More recently, Fox, Himwich, and Mani [6] significantly refined this result for digraphs forbidding certain fixed bipartite subgraphs, demonstrating bounds approaching $\beta(G) = \frac{m}{2} - \Omega(m^r)$ for all rational exponents r between $\frac{3}{4}$ and 1. They also conjectured that $\beta(G) \leq 1$ when a digraph G with n vertices is $(r-1)$-free and $r > \frac{2n}{3}$. We prove this conjecture and show that the bound is tight. We also extend the proof to find the bound for $\beta(G) \leq 2$ with a forbidden structure.

Definition 1.4. For directed cycle C and path $P = (c_1, c_2) \in C$, if there exists a path $c_1 \to v \to c_2$, then $v \notin C$ is a *bypass vertex* of P.

Theorem 1.5. *Let G be an $(r-1)$-free digraph with n vertices. If $r > \lfloor \frac{2n}{3} \rfloor$, then $\beta(G) \leq 1$.*

Theorem 1.6. *Let the structure X be a subgraph with two paths $P_1, P_2 \in C$ sharing a bypass vertex, and $V(P_1) \cap V(P_2) = \emptyset$, then any X-free, $(r-1)$-free digraph G with $r > \frac{n}{2}$ satisfies $\beta(G) \leq 2$.*

Computing a minimum feedback arc set in a digraph $G = (V, E)$ is an NP-hard problem [11]. However, with the restriction of $\beta(G) = 1$ and G being strongly connected, the complexity can be significantly reduced. A straightforward approach involves removing edges one by one and determining whether the resulting subgraph contains a directed cycle. Detecting directed cycles in a digraph is efficiently achieved by applying depth-first search (DFS) from a vertex v. If the search returns to v, then there exists a directed cycle in the graph. Thus, the overall time complexity is $O(|V| \cdot |E|)$. In this work, we present a linear-time algorithm, running in $O(|V| + |E|)$, to identify the minimum feedback arc set when $\beta(G) = 1$.

Definition 1.7. *Inversion* of a tournament T is the operation of reversing the direction of all edges whose endpoints are in a specified subset of vertices.

Definition 1.8. A tournament T is *transitive* if for every three vertices $u, v, w \in T$, whenever there are directed edges $u \to v$ and $v \to w$, there is also a directed edge $u \to w$. Equivalently, a transitive tournament has no directed cycle.

Let T be an n-vertex tournament, where $\mathrm{inv}_k(n)$ is the minimum length of a sequence of inversions, each involving at most k vertices, required to transform T into a transitive tournament. Inversion-number problems for small parameters k $(3 \leq k \leq \frac{n}{2})$ remain poorly understood, especially the exact asymptotic behavior of $\mathrm{inv}_k(n)$ for fixed small k [21]. Yuster [21] provides bounds for $\mathrm{inv}_k(n)$, stating that for all $k \geq 3$,

$$(1 + o(1))\frac{\mathrm{inv}_k(n)}{n^2} \in \left[\frac{1}{2k(k-1)} + \delta_k,\ \frac{1}{2\lfloor \frac{k^2}{2} \rfloor} - \epsilon_k\right],$$

where δ_k and ϵ_k are positive constants. Thus, for $k = 4$, Yuster's upper bound is

$$\frac{1}{2\lfloor \frac{4^2}{2} \rfloor} - \epsilon_4 = \frac{1}{2 \cdot 8} - \epsilon_4 = \frac{1}{16} - \epsilon_4,$$

where $\epsilon_4 > 0$ is a positive constant, indicating these bounds could be improved by finding more precise values for δ_k and ϵ_k. Yuster's arguments for random tournaments utilize results on edge-disjoint subgraph packings, such as those by Frankl and Rödl [7]. Similarly, advancements in bounding monochromatic triangle packings, like those from Gruslys and Letzter [9], are highly relevant to these coloring-based inversion problems. We provide a tighter upper bound for $\mathrm{inv}_4(n)$ using Ramsey numbers [16, 15].

Theorem 1.9. *Let $G = (V, E)$ be an n-vertex tournament. Then*

$$\mathrm{inv}_4(n) \leq \frac{85n^2}{1392} + o(n^2).$$

With a similar method, we can generalize the bound for $\mathrm{inv}_k(n)$. Let k be the maximum number of vertices involved in each inversion operation. The parameter x represents the number of monochromatic cliques K_k identified iteratively through the Ramsey-theoretic construction in Theorem 1.10.

Theorem 1.10. *Let $G = (V, E)$ be an n-vertex tournament. For all positive integers k and x,*

$$\mathrm{inv}_k(n) \ \leq \ \frac{n^2\big(R(k) + (x-1)(k-1)\big)^2 - xn^2(k^2 - k) + 4xn^2 \lfloor \frac{k^2}{4} \rfloor}{(4\big(R(k) + (x-1)(k-1)\big)^2 - 4R(k) - 4(x-1)(k-1)) \lfloor \frac{k^2}{4} \rfloor} \ + \ o\big(n^2\big).$$

To achieve the tightest possible bound, the value of x is chosen carefully based on k and the associated Ramsey number $R(k)$.

Another bound without using Ramsey numbers, and instead only relying on n and k is derived as

Theorem 1.11. *Let $G = (V, E)$ be an n-vertex tournament. The upper bound of*

$$\mathrm{inv}_k(n) \leq \frac{n^2}{4\lfloor \frac{k^2}{4} \rfloor} - 4^{-(1-\frac{\sqrt{2}}{2}k)-o(k)}.$$

The remainder of the paper is organized as follows. Section 2 provides proofs of Theorem 1.5 and Theorem 1.6. Section 3 shows the detailed linear-time algorithm for finding the minimum feedback arc set when $\beta(G) = 1$. The proof of Theorem 1.9, Theorem 1.10, and Theorem 1.11 appears in Section 4.

2 Feedback Arc Set Bounds

Definition 2.1. The *inner vertices* of a path P are the vertices on P, excluding its two endpoints.

Definition 2.2. Given a directed cycle C and a path $P \subset C$, a *deviate path* P' for P is a path that shares the same starting and ending vertex as P, with none of its inner vertices on C. A set of paths is considered a set of *disjoint deviate paths* if none of a path's inner vertices belong to any other deviate path in the set.

Definition 2.3. In a digraph $G = (V, E)$, an edge e can be *safely removed* if $G' = (V, E \setminus \{e\})$ has at least one directed cycle. A vertex v can be *safely removed* if removing all edges incident to v leaves at least one directed cycle in G. A vertex or edge is *safe* if it can be safely removed.

2.1 Proof of Theorem 1.5

In this section, we study how forbidding short directed cycles affects the size of the minimum feedback arc set of graphs. We prove Theorem 1.5, which claims any n-vertex digraph with no directed cycle of length at most $\lfloor \frac{2n}{3} \rfloor$ has an edge whose removal breaks every directed cycle in the graph. We prove the theorem by finding the number of vertices that are not on the smallest directed cycle C required to construct alternative paths that bypass all edges in C. We then show that, given the constraints of the shortest cycles and the total number of vertices, at least one edge in C cannot be bypassed.

Lemma 2.4. *In a digraph G, let C be the shortest directed cycle in G. Suppose path $P \subset C$ has a disjoint deviate path P', and let d denote the number of inner vertices of P'. Then, at most $d+1$ edges on P are safe.*

Proof. Since C is the shortest directed cycle, it follows that $|V(P')| \geq |V(P)|$. Otherwise, replacing P with path P' would yield a directed cycle shorter than C, contradicting the minimality of C. All edges on P are safe since we can remove any of them and still have a directed cycle intact through P'. Since $|V(P')| \geq |V(P)|$, with the d inner vertices of P', there are at most $d+1$ safe edges on P. □

Proof of Theorem 1.5. Let C be one of the shortest cycles in a digraph $G = (V, E)$ where $|V| = n$ vertices. Since G is $(r-1)$-free with $r \geq \lfloor \frac{2n}{3} \rfloor + 1$, the number of vertices in C is $|C| \geq \lfloor \frac{2n}{3} \rfloor + 1$. If there is another directed cycle C', then the number of vertices shared by C and C' is

$$|C| + |C'| - |G| \geq \left(\left\lfloor \frac{2n}{3} \right\rfloor + 1 \right) + \left(\left\lfloor \frac{2n}{3} \right\rfloor + 1 \right) - n > 1.$$

At least two vertices are shared by C and C'. We note that $\beta(G) \leq 1$ is equivalent to the statement that not all edges can be safely removed, as breaking an edge that cannot be safely removed will remove all cycles.

Let $U = G - V(C)$, and define D as the set of all existing deviate paths for paths on C.

Suppose two deviate paths $P', Q' \in D$ intersect at some vertex $v_u \in U$. Given the set of four endpoints in P' and Q', the two starting vertices are adjacent within the set of endpoints, and the two ending vertices are adjacent in the set of endpoints. We show this by contradiction. If the two starting vertices are not adjacent within the set of endpoints, then we can label the endpoints v_1, v_2, v_3, and v_4 in cyclic order, where v_1 and v_3 are starting vertices, and v_2 and v_4 are ending vertices. Let v_u be the vertex that the deviate paths share. There must exist paths from v_1 and v_3 to v_u, and from v_u to v_2 and v_4. This implies the existence of two deviate paths, one from v_1 to v_4, and one from v_3 to v_2. With these two paths, each vertex appears at least once, with v_2 and v_3 appearing twice. Hence, the total length of the paths is at least $|C| + 2$, so at least one of the paths has length at least $\frac{|C|}{2} + 1$. However, this would require at least $\frac{|C|}{2}$ vertices from U, which contradicts the fact that $|U| < \frac{|C|}{2}$. Thus, we can conclude that any two deviate paths must have their starting vertices as well as ending vertices adjacent in the cyclic order of the endpoints.

Since all pairs of deviate paths have the starting points adjacent within the set of endpoints, for any two deviate paths that go through the same vertex $v_u \in U$, let the endpoints be v_1, v_2, v_3, and v_4 in cyclic order such that v_1 and v_2 are starting vertices, and v_3 and v_4 are ending vertices. Let v_u be the vertex that the two deviate paths share. We can always simplify both deviate paths into a single deviate path without losing any safe edges by replacing the initial deviate paths with $v_1 \to v_u \to v_4$. This deviate path encompasses the same safe edges as the original paths. Iteratively applying this simplification to all intersecting pairs yields a set of deviate paths that are pairwise disjoint within U.

By Lemma 2.4, the ratio of the number of the inner vertices on P' to the safe edges on P is $\frac{d}{d+1}$. For all $d \geq 1$, we have $\frac{d}{d+1} \geq \frac{1}{2}$. Thus, each vertex on a disjoint deviate path in U allows us to safely remove at most two edges on C. Since $|C| \geq \lfloor \frac{2n}{3} \rfloor + 1$,

$$|U| = |G| - |C| \leq n - \left(\left\lfloor \frac{2n}{3} \right\rfloor + 1 \right) = \left\lceil \frac{n}{3} \right\rceil - 1.$$

Thus there are at most $2 \cdot |U| = 2 \cdot \left(\lceil \frac{n}{3} \rceil - 1 \right)$ safe edges on C. However, if two paths on C overlap with shared edges, their deviate paths will both identify the shared edges as safe edges, causing double-counting. This reduces the number of distinct safe edges.

Let $\{S_1, S_2, S_3, \ldots, S_z\}$ be the set of safe edges identified by all the disjoint deviate paths in D, where $z = |D|$. The number of safe edges for G without double-counting is

$$S_{all} = |S_1 \cup S_2 \cup S_3 \cup \ldots \cup S_z| \leq |S_1| + |S_2| + |S_3| + \ldots + |S_z| \leq 2 \cdot \left\lceil \frac{n}{3} \right\rceil - 2.$$

For all $n > 3$, since the directed cycle size

$$|C| \geq \left\lfloor \frac{2n}{3} \right\rfloor + 1 > 2 \cdot \left\lceil \frac{n}{3} \right\rceil - 2 \geq S_{all},$$

there exists at least one edge on C that cannot be safely removed, indicating $\beta(G) \leq 1$. □

Remark 2.5. Theorem 1.5 is tight as realized by an $(r-1)$-free digraph $G = (V, E)$ with $n = |V| = 6x, x \in \mathbb{Z}_N$. Let $4x$ vertices form a directed cycle with $v_1, v_2, \ldots, v_{4x}$, and each of the remaining $2x$ vertices connects v_{2i-1} and v_{2i+1} where $1 \leq i \leq 2x$. Therefore, every edge is covered by a deviate path and removing any single edge from C still leaves at least one directed cycle intact, implying $\beta(G) \geq 2$ while $r = 4x = \frac{2n}{3}$. We conclude that the bound $r \leq \lfloor \frac{2n}{3} \rfloor$ is tight.

2.2 Proof of Theorem 1.6

Lemma 2.6. *Let G be an X-free $(r-1)$-free digraph with n vertices. There exists a vertex which cannot be safely removed if $r > \frac{n}{2}$ and $n > 3$.*

Proof. Let C be the shortest cycle in G, with vertices $V(C) = \{v_1, v_2, \ldots, v_m\}$, and edges from v_i to v_{i+1} for any $i < m$, and from v_m to v_1. Since G is $(r-1)$-free with $r > \frac{n}{2}$, the number of vertices in C is $|C| > \frac{n}{2}$. There are n vertices in G and every cycle has more than $\frac{n}{2}$ vertices. If there is another cycle C', the number of vertices shared by C and C' is

$$|C| + |C'| - n > 0.$$

Therefore, C' shares at least one vertex with C.

Suppose two deviate paths $P', Q' \in D$ intersect at vertex $v_u \in U$. Given the set of four endpoints in P' and Q', the two starting vertices are adjacent within the set of endpoints, and the two ending vertices are adjacent in the set of endpoints. We show this by contradiction.

If the two starting vertices are not adjacent within the set of endpoints, we label the endpoints v_1, v_2, v_3, and v_4 in cyclic order, where v_1 and v_3 are starting vertices, and v_2 and v_4 are ending vertices. The path from v_1 to v_2 along C and the path from v_3 to v_4 along C share a bypass vertex, creating the structure X. This forms a contradiction. Thus, it is impossible for two intersecting deviate paths to have endpoints v_1, v_2, v_3, v_4 in cyclic order where v_1 and v_3 are starting vertices and v_2 and v_4 are ending vertices, so any two deviate paths must have their starting vertices as well as ending vertices adjacent in the cyclic order of the endpoints.

Since any two intersecting deviate paths have the starting vertices adjacent within the set of endpoints, for any two deviate paths that go through the same vertex $v_u \in U$, let endpoints v_1,

v_2, v_3, and v_4 be vertices in cyclic order such that v_1 and v_2 are starting vertices, and v_3 and v_4 are ending vertices. Let v_u be the bypass vertex that the two deviate paths share. v_u is a bypass vertex for the path from $v_1 \to v_4$, which encompasses both original paths, meaning we can replace the two paths by a single path $v_1 \to v_4$ without removing any safe vertices. We can repeat these replacements for intersecting deviate paths until all deviate paths are pairwise disjoint.

For each deviate path in D, if a deviate path has v_{D_i} vertices, the path has length $v_{D_i} + 1$, so the corresponding path along the smallest cycle with the same starting and ending vertices will have length at most $v_{D_i} + 1$. The vertices on this path other than the starting and ending vertex can be safely removed, meaning the deviate path corresponds with at most v_{D_i} vertices that can be safely removed. Each vertex on a disjoint deviate path in D can allow us to safely remove at most 1 vertex on C. Since $|C| > \frac{n}{2}$,

$$|U| = |G| - |C| < \frac{n}{2}.$$

With less than $\frac{n}{2}$ vertices in U, we can safely remove at most $|U|$ vertices on C where $|U| < \frac{n}{2} < |C|$. However, if two paths on C overlap with shared edges, their deviate paths will both identify the shared edges as safe edges causing double-counting. This reduces the number of distinct safe edges.

Let $\{S_1, S_2, S_3, \ldots S_z\}$ be the set of safe vertices identified by all the disjoint deviate paths in D, where $z = |D|$. The number of safe vertices for G without double-counting is

$$|S_1 \cup S_2 \cup S_3 \cup \ldots \cup S_z| \leq |S_1| + |S_2| + |S_3| + \ldots + |S_z| < \frac{n}{2}.$$

Since the cycle length $|C|$ is strictly greater than $\frac{n}{2}$, there exists at least one vertex on C that cannot be safely removed. □

Proof of Theorem 1.6. We can prove by induction. Let C be the shortest cycle in G.

Base case ($n \leq 6$):

By Lemma 2.6, there exists a vertex $v \in G$ which cannot be safely removed. When $n \leq 6$, v has at most 5 edges, which yields at most 2 incoming edges or at most 2 outgoing edges. Removing either all the incoming or all the outgoing edges from v can eliminate every cycle. Therefore, $\beta(G) \leq 2$.

Induction Hypothesis:

Suppose for all $k < n$, any $(r-1)$-free digraph G with k vertices satisfies $\beta(G) \leq 2$, where $r > \frac{k}{2}$.

Inductive step ($n \geq 7$):

Let $N^+(v)$ be the out-neighbors of v, $N^-(v)$ be the in-neighbors of v, and

$$N^{++}(v) = \{u \mid \exists w : (v, w), (w, u) \in E(G)\}.$$

If $|N^+(v)| \leq 2$, removing all the outgoing edges can make G acyclic so $\beta(G) \leq 2$.

If $|N^+(v)| > 2$, let

$$G' = G - \left(\{v\} \cup N^+(v)\right).$$

To preserve cycles corresponding to the original cycles through v, add an edge from every vertex in $N^-(v)$ to every vertex in $N^{++}(v)$. The shortest cycle in G' then has the length $|C| - 2$, so G' is $(r'-1)$-free where

$$r' = r - 2 > \frac{n}{2} - 2 = \frac{n-4}{2}.$$

Since we remove the vertex v and all the vertices in $N^+(v)$, and $|N^+(v)| > 2$, there are at most $n-4$ vertices in G'. By induction hypothesis, with $n-4$ vertices, $r' > \frac{n-4}{2}$, and G' is $(r'-1)$-free, $\beta(G') \leq 2$.

Therefore, with the base case for $n \leq 6$ and the inductive results, we conclude that for any X-free digraph G with n vertices that is $(r-1)$-free for $r > \frac{n}{2}$, it follows that $\beta(G) \leq 2$.

□

Remark 2.7. To show the bound in Theorem 1.6 is tight, consider the boundary condition $r \leq \frac{n}{2}$. That is, there exist graphs with the shortest cycle length $|C| = \lfloor \frac{n}{2} \rfloor$ and $\beta(G) \geq 3$. One such graph is a blow-up graph with $\frac{n}{2}$ groups of 2 vertices each, assuming n is even and $n \geq 6$. The graph is constructed by a $\frac{n}{2}$-cycle, and then each vertex is replaced by a pair of vertices. Let the pairs of vertices be $p_1, p_2, \ldots p_{\frac{n}{2}}$. All vertices in p_i are adjacent to all vertices in p_{i-1} and p_{i+1} where $1 < i < \frac{n}{2}$, and all vertices in $p_{\frac{n}{2}}$ are adjacent to all vertices in p_1. To make the graph acyclic, we can remove all edges that connect two adjacent pairs. Since each pair contains two vertices, removing all edges between two adjacent pairs requires removing $2 \cdot 2 = 4$ edges. Hence, $\beta(G) = 4$ and we conclude that the bound $r > \frac{n}{2}$ is tight.

3 Linear Time Algorithm for Finding the Minimum Feedback Arc Set

In this section, we present an algorithm to find the minimum feedback arc set in linear time when $\beta(G) = 1$.

Definition 3.1. A *strictly positive flow* is an assignment of positive real values (flows) to every edge in a digraph $G = (V, E)$, where every edge must carry a flow $f(e) > 0, e \in E(G)$.

Definition 3.2. A flow is *conserved* if the total incoming flow equals the total outgoing flow for every vertex. The step of *restoring conservation* refers to adjusting flow values after modifications to guarantee the graph is conserved.

Let $G = (V, E)$ be a strongly connected digraph with $\beta(G) = 1$. Since $\beta(G) = 1$, all the directed cycles in G share at least one edge, where the shared edges are the minimum feedback arc set of G. Suppose there exists a strictly positive conserved flow f on the edges of G, satisfying

$$\sum_{e \in N^-(v)} f(e) = \sum_{e \in N^+(v)} f(e)$$

for every vertex $v \in V$, where $f(e)$ is the flow on each edge $e \in E$. Let S_c be the set of directed cycles in G. By the Flow Decomposition Theorem [1], any conserved flow f can be represented as a sum of positive flows along the directed cycles in S_c. Let F_{max} be the total flow through all directed cycles. The edge in the minimum feedback arc set appears in every directed cycle, so it carries flow F_{max}. That is, the edges with the maximum flow values are where all the flows converge, hence they are in the minimum feedback arc set traversed by every directed cycle in the digraph.

Algorithm 1 FINDMINFAS($G = (V, E)$)

// *Step 1: Simplify the graph*
1: Identify the SCC using Tarjan's algorithm. ▷ $O(|V| + |E|)$
// *Step 2: Construct a strictly positive conserved flow f*
2: Calculate in-degrees $\deg_{\text{in}}(v)$ and out-degrees $\deg_{\text{out}}(v)$ for all $v \in V$. ▷ $O(|V|)$
3: $d_v \leftarrow -\epsilon(\deg_{\text{out}}(v) - \deg_{\text{in}}(v)) \quad \forall v \in V$ ▷ *Define demands for correction flow g with a positive offset* ϵ, $O(|V|)$
4: $g \leftarrow$ FEASIBLEFLOW(G, d) ▷ *Compute the flow* $g \geq 0$, $O(|V| + |E|)$
5: $f_e \leftarrow \epsilon + g_e \quad \forall e \in E$ ▷ *Construct the final flow* $f = \epsilon + g$, $O(|E|)$
// *Step 3: Identify edges with maximum flow value*
6: Determine $S_{FAS} \leftarrow \{e \in E \mid f(e) = \max_{e' \in E} f(e')\}$ ▷ *Identify max flow edges,* $O(|E|)$
7: **return** S_{FAS}

Algorithm 1 computes the minimum feedback arc set S_{FAS} in three key steps: graph simplification, flow construction, and identification of maximum flow edges. The algorithm achieves linear time complexity by using network flow techniques on the digraph structure.

The first step simplifies the graph by identifying strongly connected components (SCCs) using Tarjan's algorithm, which runs in $O(|V|+|E|)$ time. In the case where $\beta(G) = 1$, exactly one SCC remains, which contains the minimum feedback arc set.

The second step calculates the in-degrees and out-degrees for all the vertices.

In steps $3 - 5$, the function $FeasibleFlow(G, d)$ calculates the flow value for each edge using a feasible flow algorithm in $O(|V| + |E|)$ time [4, 10]. Since the flow could contain edges with zero flow, the algorithm ensures strictly positive flows by applying a lower bound adjusting technique [1]. A target minimum flow $\epsilon > 0$ is set on all edges and a non-negative correction flow g is computed to restore conservation. The resulting flow $f = \epsilon + g$ is both strictly positive and conserved.

The final step identifies edges with maximum flow in $O(|E|)$ time, yielding the minimum feedback arc set. The overall time complexity for Algorithm 1 is $O(|V| + |E|)$.

4 Improved Bound for Inversions

For all integers $s, t \geq 1$, Ramsey's theorem states that there exists a Ramsey number $R(s,t) \leq \binom{s+t-2}{s-1}$ such that any two-coloring of the edges of a tournament with at least $R(s,t)$ vertices contains either a monochromatic complete subgraph (clique) of s vertices with the first color or a monochromatic clique of t vertices with the second color.

To prove Theorem 1.9 and Theorem 1.10, we interpret the two colors as the forward and backward edges in a tournament with $R(s,t)$ vertices. A monochromatic clique represents a complete subgraph with edges being all forward or all backward, hence no directed cycle in the clique.

4.1 Proof of Bounds for $\text{inv}_4(n)$

Inversion of a tournament T is the operation of reversing the direction of all edges whose endpoints are in a specified subset of vertices. Let $\text{inv}_k(T)$ denote the minimum length of a sequence of inversions, each involving at most k vertices, required to transform T into a transitive tournament.

We utilize Ramsey numbers, specifically $R(4,4) = 18$, which guarantees the existence of a monochromatic clique K_4 in every two-edge-coloring of an 18-vertex tournament. The two colors can represent the forward and backward edges in a digraph. This structure allows us to find a tighter bound for $\text{inv}_4(n)$.

Lemma 4.1. *In any tournament, every set of* 30 *vertices contains at least five edge-disjoint transitive subgraphs of size* 4.

Proof. From Ramsey theory, any set of 18 vertices contains at least one monochromatic K_4. Given 30 vertices, we can successively extract monochromatic K_4 subgraphs with the following process.

First, we select a monochromatic K_4, leaving at least 26 vertices unused. From these 26 vertices, select another monochromatic K_4, leaving 22 vertices unused. Continuing similarly, select a third monochromatic K_4, leaving 18 vertices unused. From these 18 vertices, select a fourth monochromatic K_4, leaving exactly 14 vertices plus one vertex from each previously selected group (4 vertices), forming another set of exactly 18 vertices with no overlapping edges. Ramsey theory guarantees a fifth monochromatic K_4, resulting in at least five disjoint monochromatic K_4 subgraphs. □

Lemma 4.2. *For a two-color tournament with size of sufficiently large n, there exist at least $\frac{n^2}{174} - o(n^2)$ edge-disjoint monochromatic K_4 subgraphs.*

Proof. Let c be a constant indicating the highest number where every tournament with n vertices has at least $c \cdot n^2 - o(n^2)$ edge-disjoint K_4 graphs.

Partition the n vertices into 30 equal-sized groups $G_1, G_2, \ldots, G_{30}$. Choose the largest prime $p \leq \frac{n}{30}$, noting p approaches $\frac{n}{30}$ as n increases by the Prime Number Theorem. Construct subsets of size 30 by selecting vertices indexed as $i + j \cdot k \mod p$ within each group G_k, for $0 \leq i, j < p$. No two subsets share more than one vertex from any group, as the congruences $i_0 + j_0 \cdot x \equiv i_1 + j_1 \cdot x \mod p, \quad i_0 + j_0 \cdot y \equiv i_1 + j_1 \cdot y \mod p$ imply identical subsets if more than one vertex is shared. Thus, we generate approximately $p^2 \approx \frac{n^2}{900}$ unique subsets, each containing five monochromatic K_4 subgraphs (Lemma 4.1), yielding at least $5 \cdot \frac{n^2}{900} - o(n^2) = \frac{n^2}{180} - o(n^2)$ total monochromatic K_4 subgraphs. We used no edges between two vertices within the same group, so we can repeat the same process to find K_4 graphs within each group, leaving a total of an additional $30 \cdot c \cdot (\frac{n}{30})^2 - o(n^2) = \frac{c}{30} \cdot n^2 - o(n^2)$ K_4 graphs. We can now find $\frac{n^2}{180} - o(n^2) + \frac{c}{30} \cdot n^2 - o(n^2) \leq c \cdot n^2 \Rightarrow \frac{n^2}{180} \leq \frac{29 \cdot c}{30} \cdot n^2 \Rightarrow c \geq \frac{1}{174}$. We conclude for sufficiently large n, there exist at least $\frac{n^2}{174} - o(n^2)$ edge-disjoint monochromatic K_4 subgraphs. □

Proof of Theorem 1.9. To remove all cycles in a tournament, we can orient the vertices as a permutation, consider the edges as either forward or backward in the permutation, and invert the edges so that they are all forward or all backward. Consider the vertices being arranged randomly. The expected number of forward edges is $\frac{n^2}{4} - o(n^2)$. By Lemma 4.2, there are at least $\frac{n^2}{174}$ edge-disjoint monochromatic K_4 subgraphs, giving an expected count of forward-directed, edge-disjoint K_4 subgraphs of at least $\frac{n^2}{348} - o(n^2)$.

Given four disjoint sets S_1, S_2, S_3, and S_4 with two vertices in each set, inverting $\{S_1, S_2\}$, $\{S_2, S_3\}$, $\{S_3, S_4\}$ and $\{S_1, S_4\}$ will result in inverting the 4 edges between each pair. These edges are a subset of the edges in the bipartite graph between $\{S_1, S_3\}$ and $\{S_2, S_4\}$. By the Kővári-Sós-Turán theorem, the edges can always be removed until there are $o(n^2)$ edges left. These four inversions can result in removing $4 \cdot 4 = 16$ edges when applied to a $K(4, 4)$ bipartite graph.

For our final bound calculation, initially, inversions within monochromatic K_4 subgraphs remove $6 \cdot \frac{n^2}{348} = \frac{n^2}{58}$ edges, leaving $\frac{n^2}{4} - \frac{n^2}{58} = \frac{27n^2}{116}$ edges. Removing these remaining edges via $K_{4,4}$ inversions requires $\frac{4}{16} \cdot \frac{27n^2}{116} = \frac{27n^2}{464}$ additional inversions. Including initial inversions for monochromatic K_4 subgraphs ($\frac{n^2}{348}$ inversions), the total inversions required is

$$\frac{27n^2}{464} + \frac{n^2}{348} = \frac{85n^2}{1392} + o(n^2).$$

Thus, the established upper bound is

$$\text{inv}_4(n) \leq \frac{85n^2}{1392} + o(n^2).$$

□

4.2 Proof of Bounds for $\mathrm{inv}_k(n)$

We continue to utilize Ramsey numbers, specifically the existence of a monochromatic clique K_k in every 2-coloring of an $R(k)$-vertex tournament. We start with proving lemmas to find bounds on the number of K_k subgraphs, and use these results to construct the main proof for Theorem 1.10.

Lemma 4.3. *For all positive integers x and k, a two-color tournament of size $R(k)+(x-1)(k-1)$ has at least x K_k monochromatic tournaments.*

Proof. By Ramsey theory, every two-color tournament with $R(k)$ vertices contains at least one monochromatic clique K_k. Given a two-color tournament with $R(k) + (x-1)(k-1)$ vertices, we can construct monochromatic x K_k subgraphs iteratively.

First, select any monochromatic clique K_k. Remove $k-1$ vertices of this clique from future consideration, while keeping exactly one vertex for possible reuse in later steps. This ensures that subsequent monochromatic cliques remain edge-disjoint, as two distinct cliques formed from this process will share at most one vertex and hence share no edges. Each iteration removes $k-1$ vertices from future consideration. Repeating this process $x-1$ times results in $x-1$ edge-disjoint monochromatic K_k cliques. After these steps, exactly $R(k)$ vertices remain, which contain at least one additional monochromatic K_k by Ramsey's theorem.

Thus, we conclude that a total of x monochromatic K_k cliques are guaranteed to exist within the given two-color tournament. □

Lemma 4.4. *In a two-color tournament of size n, for sufficiently large n, and any positive integer x, there exist at least*

$$\frac{x \cdot n^2}{(R(k) + (x-1)(k-1))^2 - R(k) - (x-1)(k-1)} - o(n^2)$$

edge-disjoint monochromatic K_k subgraphs.

Proof. Let $y = R(k)+(x-1)(k-1)$. Partition the n vertices into y equal-sized groups $G_1, G_2, \ldots, G_y$. Choose the largest prime $p \le \frac{n}{y}$, noting p approaches $\frac{n}{y}$ as n increases by the Prime Number Theorem. We can construct subsets of vertices of size p by selecting vertices indexed as $i + j \cdot k \mod p$ within each group G_k, for $0 \le i, j < p$, as the congruences of the form $i_0 + j_0 \cdot k \equiv i_1 + j_1 \cdot k \mod p$ for distinct pairs (i_0, j_0) and (i_1, j_1) can hold true for at most one value of k, so that any two subsets share at most one vertex.

Thus, we generate approximately $p^2 \approx \frac{n^2}{y^2}$ unique subsets, each containing x monochromatic K_k subgraphs (Lemma 4.3). There are now also y groups of $\frac{n}{y}$ vertices that have no K_k subgraphs with at least two vertices in these groups. For some constant c, the number of subgraphs is of the form $c \cdot n^2$ such that

$$c \cdot n^2 = c \cdot y \cdot \left(\frac{n}{y}\right)^2 + x \cdot \frac{n^2}{y^2}.$$

Since each of the y subgroups have $\frac{n}{y}$ elements, which means there are $y \cdot \left(c \cdot (\frac{n}{y})^2\right)$ subgraphs that exist purely within a group, with $c = \frac{x}{y^2 - y}$, there are

$$cn^2 = \frac{xn^2}{y^2 - y} - o(n^2) = \frac{xn^2}{(R(k) + (x-1)(k-1))^2 - R(k) - (x-1)(k-1)} - o(n^2)$$

total monochromatic K_k subgraphs.

□

Proof of Theorem 1.10. Let $y = R(k) + (x-1)(k-1)$. To remove all cycles in a directed graph, we can orient the vertices as a permutation, consider the edges as either forward or backward in the permutation, and invert the edges so that they are all forward or all backward. Consider the vertices being arranged randomly. The expected number of forward edges is approximately $\frac{n^2}{4} - o(n^2)$. By Lemma 4.4, since on average half of the monochromatic K_k graphs form forward-directed monochromatic K_k subgraphs, and there is an expected value of at least $\frac{x \cdot n^2}{y^2-y}$ K_k graphs, the expected count of forward-directed, edge-disjoint K_k subgraphs is at least

$$\frac{x \cdot n^2}{2y^2 - 2y} - o(n^2).$$

We can partition any $2k$ vertices into four disjoint sets S_1, S_2, S_3, and S_4 with $|S_1| = |S_3| = \lfloor \frac{k}{2} \rfloor$, and $|S_2| = |S_4| = \lceil \frac{k}{2} \rceil$. Inverting $\{S_1, S_2\}$, $\{S_2, S_3\}$, $\{S_3, S_4\}$ and $\{S_1, S_4\}$ will result in inverting precisely all edges connecting vertices in $\{S_1, S_3\}$ and $\{S_2, S_4\}$. These inverted edges are exactly those in the complete bipartite graph $K_{2\lfloor \frac{k}{2} \rfloor, 2\lceil \frac{k}{2} \rceil}$. By the Kővári-Sós-Turán theorem, the edges can always be removed until there are $o(n^2)$ edges left. These four inversions can remove $2\lfloor \frac{k}{2} \rfloor \cdot 2\lceil \frac{k}{2} \rceil = 4\lfloor \frac{k^2}{4} \rfloor$ edges when applied to a $K_{k,k}$ bipartite graph.

For our final bound calculation, initially, inversions within monochromatic K_k subgraphs remove

$$\binom{k}{2} \cdot \frac{x \cdot n^2}{2y^2 - 2y} = \frac{x \cdot n^2 \cdot (k^2 - k)}{4y^2 - 4y}$$

edges, leaving

$$\frac{n^2}{4} - \frac{x \cdot n^2 \cdot (k^2 - k)}{4y^2 - 4y}$$

edges. The number of $K_{k,k}$ inversions required to remove the remaining edges is

$$\begin{aligned} &\frac{4}{4\lfloor \frac{k^2}{4} \rfloor} \cdot \left(\frac{n^2}{4} - \frac{x \cdot n^2 \cdot (k^2 - k)}{4y^2 - 4y} \right) \\ &= \frac{n^2 y^2 - n^2 y - x \cdot n^2 (k^2 - k)}{(4y^2 - 4y)\lfloor \frac{k^2}{4} \rfloor}. \end{aligned}$$

We need to invert the set of monochromatic K_k graphs, the set of complete bipartite graphs, and $o(n^2)$ edges that remain after removing the complete bipartite graphs, as shown by the Kővári-Sós-Turán theorem. As a result, the total inversions required is

$$\begin{aligned} \operatorname{inv}_k(n) &\leq \frac{x \cdot n^2}{2y^2 - 2y} + \frac{n^2 y^2 - n^2 y - x \cdot n^2 (k^2 - k)}{(4y^2 - 4y)\lfloor \frac{k^2}{4} \rfloor} \\ &= n^2 \cdot \frac{2x\lfloor \frac{k^2}{4} \rfloor + y^2 - y - x \cdot k^2 + x \cdot k}{(4y^2 - 4y)\lfloor \frac{k^2}{4} \rfloor} - o(n^2). \end{aligned}$$

Thus, the established upper bound for $\operatorname{inv}_k(n)$ is

$$n^2 \cdot \frac{2x\lfloor \frac{k^2}{4} \rfloor + (R(k) + (x-1)(k-1))^2 - R(k) - (x-1)(k-1) - x \cdot k^2 + x \cdot k}{(4(R(k) + (x-1)(k-1))^2 - 4(R(k) + (x-1)(k-1)))\lfloor \frac{k^2}{4} \rfloor} - o(n^2).$$

□

Remark 4.5. This bound is minimized when choosing an x that maximizes the number of K_k. This corresponds to maximizing the expression

$$\frac{x}{(R(k) + (x-1)(k-1))^2 - R(x) - (x-1)(k-1)} = \frac{x}{y^2 - y},$$

where $y = R(k) + (x-1)(k-1)$, and we find the optimal x by solving

$$\begin{aligned}
\frac{d}{dx}\left[\frac{x}{y^2-y}\right] &= \frac{y^2 - y - x(2y-1)\frac{dy}{dx}}{(y^2-y)^2} \\
&= \frac{(R(k)+(x-1)(k-1))^2 - (R(k)+(x-1)(k-1)) - x(2(R(k)+(x-1)(k-1))-1)(k-1)}{(y^2-y)^2} \\
&= \frac{-k^2x^2 + 2kx^2 + k^2 - 2kR(k) + R(k)^2 - x^2 - k + R(k)}{(y^2-y)^2} \\
&= \frac{-(k-1)^2x^2 + (k-R(k))^2 - k + R(k)}{(y^2-y)^2}.
\end{aligned}$$

For domain $x \in [1, \infty]$, this expression starts being positive, then reaches zero at $x = \frac{\sqrt{(k-R(k))^2-k+R(k)}}{k-1}$, and becomes negative afterwards. Since x is an integer, we can thus optimize $\frac{x \cdot n^2}{(R(k)+(x-1)(k-1))^2} - o(n^2)$ by setting x to either the floor or ceiling of $\frac{\sqrt{(k-R(k))^2-k+R(k)}}{k-1}$.

4.3 An Alternate Bound for $\text{inv}_k(n)$

Ramsey numbers can be used to find bounds for $\text{inv}_k(n)$, however, not all Ramsey numbers are known. In this section, we prove Theorem 1.11, which only takes k and n as parameters for the bound.

Definition 4.6. Let $R(n_1, e_1, n_2, e_2)$ denote the minimum number of vertices such that a two-color graph with $R(n_1, e_1, n_2, e_2)$ where each edge is either red or blue has either a subgraph of size n_1 with at least e_1 red edges or a subgraph of size n_2 with at least e_2 blue edges.

Lemma 4.7. *If* $\frac{e_1}{\binom{n_1}{2}} + \frac{e_2}{\binom{n_2}{2}} \leq 1$*, then* $R(n_1, e_1, n_2, e_2) \leq \max(n_1, n_2)$.

Proof. We prove by contradiction. Given a graph of size $\max(n_1, n_2)$, and let p be the probability of a uniformly random edge being red.

Choose n_1 vertices uniformly at random. The expected number of red edges they span is $p\binom{n_1}{2}$, so there exists an n_1-vertex subgraph with at least $p\binom{n_1}{2}$ red edges. Similarly, there exists an n_2-vertex subgraph with at least $(1-p)\binom{n_2}{2}$ blue edges. For contradiction, assume that there is no n_1-vertex subgraph with $\geq e_1$ red edges and no n_2-vertex subgraph with $\geq e_2$ blue edges. Then

$$p\binom{n_1}{2} < e_1 \quad \text{and} \quad (1-p)\binom{n_2}{2} < e_2,$$

thus

$$p < \frac{e_1}{\binom{n_1}{2}} \quad \text{and} \quad 1 - p < \frac{e_2}{\binom{n_2}{2}}.$$

Adding them gives

$$1 < \frac{e_1}{\binom{n_1}{2}} + \frac{e_2}{\binom{n_2}{2}},$$

which contradicts the hypothesis. Hence $R(n_1, e_1, n_2, e_2) \leq \max(n_1, n_2)$. □

Lemma 4.8. *Every two-color subgraph of* $4^{(1-\frac{\sqrt{2}}{2})n+o(n)}$ *vertices contains a subgraph of size* n *such that there are at least* $\frac{3n(n-1)}{8}$ *edges with the same color.*

Proof. Let v be a vertex on a two-color graph with red/blue edges, where there exists a subgraph of size n_1 with e_1 red edges or a subgraph of size n_2 with e_2 blue edges. Note that v cannot have $R(n_1-1, e_1-n_1+1, n_2, e_2)$ red edges or more, or these edges will connect to either a blue subgraph of size n_2 with e_2 blue edges, or a red subgraph of size n_1-1 with e_1-n_1+1 red edges, where if this subgraph is combined with v, a subgraph of size n_1 with e_1 edges will be formed. Similarly, v cannot have $R(n_1, e_1, n_2-1, e_2-n_2+1)$ blue edges. Thus,

$$R(n_1, e_1, n_2, e_2) < R(n_1, e_1, n_2-1, e_2-n_2+1) + R(n_1-1, e_1-n_1+1, n_2, e_2),$$

indicating

$$R(a,b) < R(a+1,b) + R(a,b+1).$$

Let $f(a,b) = R\left(n-a, e-\frac{a(2n-a-1)}{2}, k-b, e-\frac{b(2n-b-1)}{2}\right)$.

Thus, $f(0,0) < \sum_{k=0}^{\lceil n(2-\sqrt{2})\rceil} \binom{\lceil n(2-\sqrt{2})\rceil}{k} f(k, \lceil n(2-\sqrt{2})\rceil - k)$. For a value k, let

$$b = \lceil n(2-\sqrt{2})\rceil - k,$$

such that

$$f\left(k, \lceil n(2-\sqrt{2})\rceil - k\right) < R(n-k, \frac{3n(n-1)}{8} - \frac{(2n-1-k)k}{2}, n-b, \frac{3n(n-1)}{8} - \frac{(2n-1-b)b}{2}\Big).$$

To prove this satisfies the condition for Lemma 4.7, we want to prove

$$\frac{\frac{3n(n-1)}{8} - \frac{(2n-1-k)k}{2}}{\frac{(n-k)(n-k-1)}{2}} + \frac{\frac{3n(n-1)}{8} - \frac{(2n-1-b)b}{2}}{\frac{(n-b)(n-b-1)}{2}} \leq 1.$$

The derivative of the left hand side with respect to k is $\frac{(n-1)n(2k-2n+1)}{4(n-k)^2(n-k-1)^2}$. Through symmetry, the derivative of the left hand side with respect to b is

$$\frac{d}{db}\frac{\frac{3n(n-1)}{8} - \frac{(2n-1-b)b}{2}}{\frac{(n-b)(n-b-1)}{2}} = \frac{(n-1)n(2b-2n+1)}{4(n-b)^2(n-b-1)^2}.$$

Since $k+b$ is constant, the overall derivative with respect to k after taking into account b is

$$\begin{aligned}\frac{d}{dk}\left(\frac{\frac{3n(n-1)}{8} - \frac{(2n-1-k)k}{2}}{\frac{(n-k)(n-k-1)}{2}} + \frac{\frac{3n(n-1)}{8} - \frac{(2n-1-b)b}{2}}{\frac{(n-b)(n-b-1)}{2}}\right) &= \frac{(n-1)n(2k-2n+1)}{4\,(n-k)^2\,(n-k-1)^2} - \frac{(n-1)n(2b-2n+1)}{4\,(n-b)^2\,(n-b-1)^2} \\ &\propto \frac{2k-2n+1}{4\,(n-k)^2\,(n-k-1)^2} - \frac{2b-2n+1}{4\,(n-b)^2\,(n-b-1)^2}.\end{aligned}$$

This is increasing over k since $\frac{(2k-2n+1)}{4(n-k)^2(n-k-1)^2}$ increases as k increases, as

$$\frac{d}{dk}\frac{(2k-2n+1)}{4(n-k)^2(n-k-1)^2} = \frac{3(x+\frac{1-2n}{2})}{2(x-n)^3(x-n+1)^3} \leq 0,$$

while $\frac{(2b-2n+1)}{4(x-b)^2(x-b-1)^2}$ decreases as k increases through symmetry. When $k=b$, through symmetry,

$$\frac{(2k-2n+1)}{4(n-k)^2(n-k-1)^2} - \frac{(2b-2n+1)}{4(n-b)^2(n-b-1)^2} = 0.$$

Thus, the expression $\frac{\frac{3n(n-1)}{8}-\frac{(2n-1-k)k}{2}}{\frac{(n-k)(n-k-1)}{2}}+\frac{\frac{3n(n-1)}{8}-\frac{(2n-1-b)b}{2}}{\frac{(n-b)(n-b-1)}{2}}$ has derivative starting negative, equal zero at $k=b$, then becomes positive, showing it minimizes at $k=b$. The function equals 1 when $k=b$. Thus, Lemma 4.7 is satisfied, and

$$f(k,\lceil(2-\sqrt{2})\rceil-k)\leq\max(n_1,n_2)\leq 2n.$$

Thus,

$$\begin{aligned}f(0,0)&\leq\sum_{k=0}^{\lceil n(2-\sqrt{2})\rceil}\binom{\lceil n(2-\sqrt{2})\rceil}{k}f\Big(k,\ \lceil n(2-\sqrt{2})\rceil-k\Big)\\&\leq n\cdot\sum_{k=0}^{\lceil n(2-\sqrt{2})\rceil}\binom{\lceil n(2-\sqrt{2})\rceil}{k}2n\\&=4^{\lceil(1-\frac{\sqrt{2}}{2})n\rceil+o(n)}.\end{aligned}$$

This means $R(n_1,e_1,n_2,e_2)\leq 4^{(1-\frac{\sqrt{2}}{2})n+o(n)}$, so every two-color subgraph of $4^{(1-\frac{\sqrt{2}}{2})n}$ vertices contains a subgraph of size n such that there are at least $\frac{3n(n-1)}{8}$ edges with the same color. □

Proof of Theorem 1.11. To remove all cycles in a directed graph, we can orient the vertices as a permutation, consider the edges as either forward or backward in the permutation, and invert the edges so that they are all forward or all backwards. Consider the vertices being arranged randomly. The expected number of forward edges is $\frac{n^2}{4}-o(n^2)$. With Lemma 4.8, we can form edge-disjoint tournaments of size k with greater than $\frac{3k^2}{8}$ forward edges or backward edges until there are no tournaments of size $4^{(1-\frac{\sqrt{2}}{2})k+o(k)}$ without edges chosen in one of the edge disjoint tournaments of size k. Through Turán's Theorem, this would result in inverting $\frac{n^2}{2\cdot 4^{(1-\frac{\sqrt{2}}{2})k+o(k)}-2}=\frac{n^2}{4^{(1-\frac{\sqrt{2}}{2})k+o(k)}}$ edges. Thus, on average, when inverting each of these tournaments with greater than $\frac{3k^2}{8}$ forward edges of each of these tournaments with greater than $\frac{3k^2}{8}$ backward edges, $\frac{n^2}{4^{(1-\frac{\sqrt{2}}{2})k+o(k)}}$ are inverted, with each inversion decreasing either the number of forward edges total or backward edges total by over $\frac{k^2}{4}$.

Consider four disjoint sets S_1, S_2, S_3, and S_4, where S_1 and S_3 contains $\lfloor\frac{k}{2}\rfloor$ elements and S_2 and S_4 contains $\lceil\frac{k}{2}\rceil$ elements. Inverting $\{S_1,S_2\}$, $\{S_2,S_3\}$, $\{S_3,S_4\}$ and $\{S_1,S_4\}$ will result in inverting the $\lfloor\frac{k^2}{4}\rfloor$ edges between each pair. These edges are a subset of the edges in the bipartite graph between $\{S_1,S_3\}$ and $\{S_2,S_4\}$. By the Kővári-Sós-Turán theorem [12], we can apply these inversions to eliminate edges until there are $o(n^2)$ edges left. These four inversions can result in removing $4\cdot 4=16$ edges when applied to a $K_{4,4}$ bipartite graph. There are, on average, $\frac{n^2}{4}-o(n^2)$ forward edges, so through this method of finding sets to invert, we can use $\frac{n^2}{4\lfloor\frac{k^2}{4}\rfloor}+o(n^2)$ inversions to invert all edges. We can now repeat this method after excluding the initial $\frac{n^2}{2}4^{-(1-\frac{\sqrt{2}}{2})k-o(k)}$ edges we invert at the start with greater than $\frac{k^2}{4}$ edges being inverted on average at a time in order to find

$$\mathrm{inv}_k(n)\leq\frac{n^2\cdot 4^{-(1-\frac{\sqrt{2}}{2})k-o(k)}}{4(\lfloor\frac{k^2}{4}\rfloor+1)}+\frac{n^2(1-4^{-(1-\frac{\sqrt{2}}{2})k-o(k)})}{4\lfloor\frac{k^2}{4}\rfloor}=\frac{n^2}{4\lfloor\frac{k^2}{4}\rfloor}-n^2\cdot 4^{-(1-\frac{\sqrt{2}}{2})k-o(k)}.$$

□

5 Acknowledgements

I would like to thank my mentor, Dr. Nitya Mani, for her guidance and support throughout my research. I am also grateful to the MIT PRIMES-USA program for providing this unique and invaluable opportunity.

References

[1] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications.* Prentice Hall, Englewood Cliffs, NJ, 1993.

[2] B. Berger and P. Shor. Approximation algorithms for the maximum acyclic subgraph problem, *Proceedings of the first annual ACM-SIAM Symposium on Discrete Algorithms*, 1990.

[3] M. Chudnovsky, P. Seymour, and B. Sullivan, Cycles in Dense Digraphs, *Combinatorica*, 28, 1–18, 2008.

[4] T. Erlebach and T. Hagerup, Routing flow through a strongly connected graph, *Algorithmica*, 32(3), 467–473, 2002.

[5] J. Fox, P. Keevash, and B. Sudakov, Directed graphs without short cycles, *Combinatorics, Probability and Computing*, 19(2): 285–301, 2010.

[6] J. Fox, Z. Himwich, and N. Mani, Extremal results on feedback arc sets in digraphs, *Random Structures & Algorithms*, 64, 287–308, 2024.

[7] P. Frankl and V. Rödl, Near perfect coverings in graphs and hypergraphs, *European Journal of Combinatorics*, 6(4), 317–326, 1985.

[8] P. Gillot and P. Parviainen, Scalable Bayesian network structure learning via maximum acyclic subgraph, *Proceedings of the 10th International Conference on Probabilistic Graphical Models*, 138, 209–220, 2020.

[9] V. Gruslys and S. Letzter, Monochromatic triangle packings in red-blue graphs, arXiv:2008.05311, 2020.

[10] B. Haeupler and R. E. Tarjan, Finding a feasible flow in a strongly connected network, *Operations Research Letters*, 36(4), 397–398, 2008.

[11] R. M. Karp, *Reducibility among combinatorial problems*, in R. E. Miller and J. W. Thatcher (eds.), Complexity of Computer Computations, Plenum Press, New York, 85–103, 1972.

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

[13] C. E. Leiserson and J. B. Saxe, Retiming synchronous circuitry, *Algorithmica*, 6(1), 5–35, 1991.

[14] S. Poljak, V. Rödl, and J. Spencer, Tournament ranking with expected profit in polynomial time, *SIAM Journal on Discrete Mathematics*, 1(3), 372–376, 1988.

[15] S. P. Radziszowski, Small Ramsey Numbers, *The Electronic Journal of Combinatorics*, Dynamic Survey DS1, revision #17, pp. 4, https://doi.org/10.37236/21, 2024.

[16] F. P. Ramsey, On a Problem of Formal Logic, *Proceedings of the London Mathematical Society*, 30, 264–286, 1930.

[17] S. S. Skiena, *The Algorithm Design Manual*, Vol. 2. Springer, London, 2008.

[18] P. Slater, Inconsistencies in a schedule of paired comparisons, *Biometrika*, 48, 303–312, 1961.

[19] B. Sullivan, *Extremal Problems in Digraphs*, Ph.D. thesis, Princeton University, May 2008.

[20] R. E. Tarjan, Depth-first search and linear graph algorithms, *SIAM Journal on Computing*, 1(2), 146–160, 1972.

[21] R. Yuster, On tournament inversion, *Journal of Graph Theory*, 110, 82–89, 2025.