Inducibility of 4-vertex tournaments

Dalton Burke* Bernard Lidický† Florian Pfender‡ Michael Phillips§

March 12, 2021

Abstract

We determine the inducibility of all tournaments with at most 4 vertices together with the extremal constructions. The 4-vertex tournament containing an oriented C_3 and one source vertex has a particularly interesting extremal construction. It is an unbalanced blow-up of an edge, where the sink vertex is replaced by a quasi-random tournament and the source vertex is iteratively replaced by a copy of the construction itself.

1 Introduction

One of the central questions in extremal graph theory is to maximize the number of induced copies of a given graph H in a larger host graph on a fixed number of vertices. For a graph G, we denote the number of vertices by $|H|$. For this, let $I(H,G)$ be the number of vertex subsets of G which induce a graph isomorphic to H, and let

$$I(H,n) = \max_{|G|=n} I(H,G).$$

For $k = |H|$, we normalize the answer and write $i(H,n)\binom{n}{k} = I(H,n)$. Clearly, $0 \leq i(H,n) \leq 1$, so we can think of $i(H,n)$ as a subgraph density. An easy averaging argument shows that $i(H,n)$ is monotone non-increasing and thus converging for $n \to \infty$. Pippenger and Golumbic [21] define the *inducibility of H* as

$$i(H) = \lim_{n\to\infty} i(H,n).$$

Determining inducibilities is notoriously difficult, and the answer is known only for very few explicit graphs H. A major breakthrough for the problem was the introduction of the flag algebra method by Razborov [22] in 2007, and since then the inducibility of a good number of small graphs has been determined with the help of this method [1, 2, 7]. Nevertheless, we do not even

*Department of Mathematical and Statistical Sciences, University of Colorado Denver, E-mail: Dalton.Burke@ucdenver.edu .

†Department of Mathematics, Iowa State University, Ames, IA, E-mail: lidicky@iastate.edu. Research of this author is supported in part by NSF grant DMS-1855653.

‡Department of Mathematical and Statistical Sciences, University of Colorado Denver, E-mail: Florian.Pfender@ucdenver.edu. Research is partially supported by NSF grant DMS-1855622.

§Department of Mathematical and Statistical Sciences, University of Colorado Denver, E-mail: Michael.2.Phillips@ucdenver.edu .

know $i(P_4)$, i.e. the inducibility of the path on four vertices, and we do not even have a conjecture for the answer.

On the other end of the spectrum, Fox, Huang, and Lee [9], and independently Yuster [23], have determined exact values for $i(H,n)$ and thus $i(H)$ for all n and almost all large enough graphs H by studying random graphs. There are numerous other results on inducibility [3, 6, 10–12, 16–19].

All of these questions can be studied for directed graphs as well. Falgas-Ravry and Vaughan [8] studied inducibility of small outstars. Huang [15] extended the result to all outstars. This was further generalized to other stars by Hu, Ma, Norin, and Wu [14]. Short paths with further restrictions were considered in [4] and orientation of a 4-cycle in [13]. In an REU in 2018, Burgher and Burke studied and conjectured extremal constructions for most oriented graphs (directed graphs without 2-cycles) of up to 4 vertices using the flag algebra method. In a similar and independent project, Bożyk, Grzesik and Kielak [24] established more bounds and found more constructions for oriented graphs.

In this paper, we look closer at the tournaments in this list, directed graphs with exactly one arc between any pair of vertices. The number of non-isomorphic tournaments on k vertices is slightly smaller than the number of graphs, and flag algebra computations tend to have similar power. The two projects mentioned in the previous paragraph both found inducibility bounds and closely matching constructions for all tournaments on up to 4 vertices, where the results are easy or trivial for all but three of these 8 small tournaments. These last three tournaments on 4 vertices have very interesting constructions, and in this paper we prove that these constructions are indeed optimal.

When edges are colored, Mubayi and Razborov [20] showed that for every tournament T on $k \geq 4$ vertices whose edges are colored by $\binom{k}{2}$ distinct colors, a structure on $n \geq k$ vertices that maximizes the number of copies of T is a balanced iterated blow-up of T. This implies that $i(T) = \frac{k!}{k^k - k}$ in this rainbow setting.

2 Results

Here we discuss tournaments on at most four vertices. For the tournaments T_1 and T_2 on one and two vertices, respectively, any tournament T has $i(T_k, T) = 1$. Similarly, for all transitive tournaments TT_k on $k \geq 3$ vertices, the transitive tournament TT_n on $n \geq k$ vertices is the unique tournament on n vertices with $i(TT_k, T) = 1$. On the other hand, $i(TT_3, T)$ is minimized exactly if T has all out-degrees in $\{\frac{n-2}{2}, \frac{n-1}{2}, \frac{n}{2}\}$. This easily follows from counting TT_3 by first choosing the source vertex, and then any two out-neighbors. As a consequence, we have for the only other tournament on three vertices C_3:

Proposition 1. *The number of induced copies of C_3 is maximized if and only if every vertex of a tournament has out-degree in $\{\frac{n-2}{2}, \frac{n-1}{2}, \frac{n}{2}\}$.*

This implies $i(C_3) = 1/4$ and leaves us with three 4-vertex tournaments to consider, see Figure 1: the tournaments we get from C_3 by adding a source vertex (C_3^+), a sink vertex (C_3^-), and by adding a vertex of out-degree 1 or 2 (C_4).

We will now define the class $\mathscr{C}_n$ of *carousels* on $n \geq 3$ vertices. A tournament T is in $\mathscr{C}_n$ if its vertices can be labeled $\{v_1, v_2, \ldots, v_n\}$ such that $v_i v_j \in E(T)$ if $0 < j - i < \frac{n}{2}$ or if $-n < j - i < -\frac{n}{2}$.

Figure 1: The four 4-vertex tournaments.

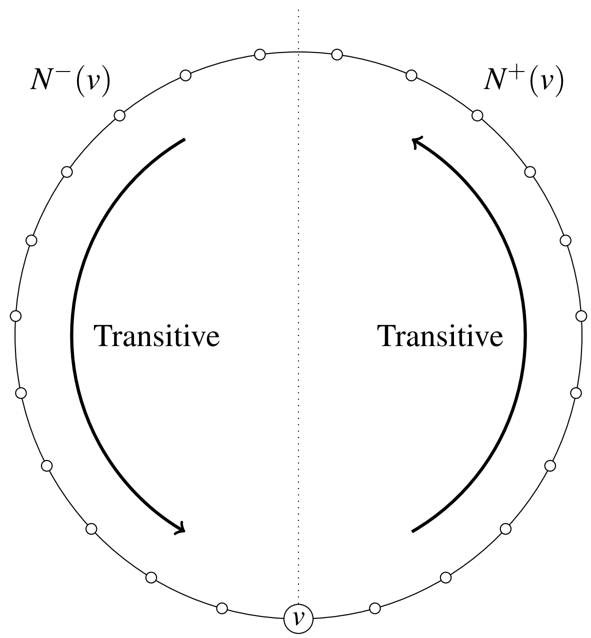


Figure 2: For odd n, the carousel $C \in \mathscr{C}_n$ is unique and vertex transitive. In the even n case, an additional vertex v' would be placed opposite v on the circle with the arc containing v and v' chosen arbitrarily.

Said another way, a tournament T is in $\mathscr{C}_n$ if for every $x \in V(T)$, the in- and out-neighborhoods induce transitive tournaments (T is *locally transitive*) and are as balanced as possible (T is *balanced* when $|V(T)|$ is odd, or *nearly balanced* when $|V(T)|$ is even). See Figure 2 for an illustration.

Observe that for odd n and for $n = 4$, $\mathscr{C}_n$ contains exactly one tournament, and we will call this unique carousel C_n. For even $n \geq 6$, $\mathscr{C}_n$ contains more than one tournament, depending on the directions of the arcs $v_i v_{i+\frac{n}{2}}$. For even n, we denote by $C_n \in \mathscr{C}_n$ the unique tournament we get from deleting one vertex in C_{n+1}. Note that one can alternatively construct C_n from C_{n-1} by duplicating one vertex.

Theorem 2. *For $n \geq 4$, the tournaments maximizing $I(C_4, T)$ are precisely the tournaments in $\mathscr{C}_n$. Consequently, $i(C_4) = \frac{1}{2}$, and for every n, we have*

$$I(C_4, n) = \begin{cases} \frac{n(n^2-1)(n-3)}{48} & \text{if } n \text{ is odd,} \\ \frac{n(n^2-4)(n-3)}{48} & \text{if } n \text{ is even.} \end{cases}$$

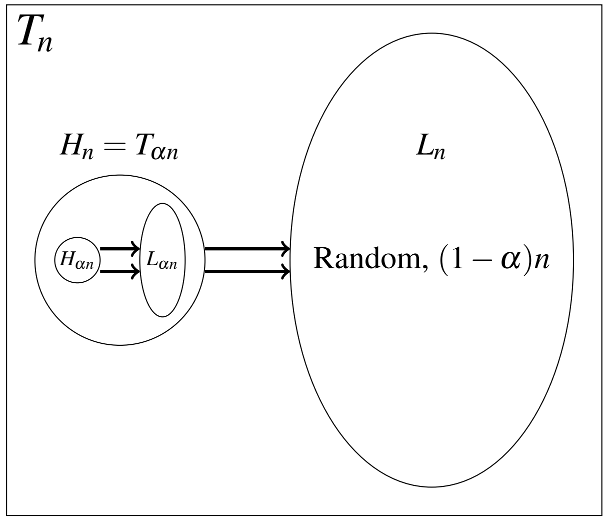


Figure 3: Construction maximizing the number of copies of C_3^+. For $\alpha \in [0,1]$ and n sufficiently large, the extremal construction T_n can be decomposed into subtournaments L_n, of size about $(1-\alpha)n$, and H_n, of size about αn with the properties shown above.

Note that the asymptotic statement that $i(C_4) = \frac{1}{2}$ is also proved in [24], with a proof very similar to the one we provide in the next section. Flag algebra computations indicate that a similar statement is also true for C_5, C_6, C_7 and C_8, and we conjecture it is true for all k. Observe that for $k \geq 5$ and even $n \geq k$, C_n contains more copies of C_k than the other members of $\mathscr{C}_n$, so our conjectured class of extremal tournaments is a bit smaller here.

Conjecture 3. *For all $k \geq 5$ and $n \geq k$, the unique n-vertex tournaments maximizing $I(C_k, T)$ are the tournaments C_n.*

The only tournaments on 4 vertices left are the two tournaments C_3^- and C_3^+. As one gets C_3^- from C_3^+ by reversal of all arcs, the tournaments extremal for C_3^- are precisely the reversals of the tournaments extremal for C_3^+, so it suffices to only study C_3^+. Consider the following construction of a tournament T_n on n vertices. For some fixed $\alpha \in (0,1)$, partition the vertices into two sets H_n (for high out-degree) and L_n (for low out-degree) of size $\lceil \alpha n \rceil$ and $\lfloor (1-\alpha)n \rfloor$, respectively. On the set L_n, direct the edges uniformly at random, i.e. insert a random tournament $R_{\lfloor (1-\alpha)n \rfloor}$. All arcs between the sets are directed from H_n to L_n. On the set H_n, iterate the construction, i.e. insert the tournament $T_{\lceil \alpha n \rceil}$ inductively. See Figure 3 for a sketch of the iterated construction.

In this construction, all copies of C_3^+ lie completely in H_n, completely in L_n, or have exactly one vertex in H_n and three vertices forming a C_3 in L_n. Notice that $i(C_3, R_{\lfloor (1-\alpha)n \rfloor}) = 1/4 + o(1)$ and $i(C_3^+, R_{\lfloor (1-\alpha)n \rfloor}) = 1/8 + o(1)$. As $i(C_3^+, T_{\lceil \alpha n \rceil}) = i(C_3^+, T_n) + o(1)$, we have

$$\begin{aligned} i(C_3^+, T_n) &= \alpha^4 i(C_3^+, T_n) + 4\alpha(1-\alpha)^3 i(C_3, R_{\lfloor (1-\alpha)n \rfloor}) + (1-\alpha)^4 i(C_3^+, R_{\lfloor (1-\alpha)n \rfloor}) + o(1) \\ &= \alpha^4 i(C_3^+, T_n) + 4\alpha(1-\alpha)^3 \tfrac{1}{4} + (1-\alpha)^4 \tfrac{1}{8} + o(1), \end{aligned}$$

so

$$i(C_3^+, T_n) = \frac{\alpha(1-\alpha)^3 + \frac{1}{8}(1-\alpha)^4}{1-\alpha^4} + o(1).$$

Maximizing this quantity gives us $\alpha = \frac{1}{5}(2\sqrt[3]{9} - 2 - \sqrt[3]{3}) \approx 0.1435836$, and

$$i(C_3^+, T_n) = \frac{1}{8}\left(8 - 9\sqrt[3]{3} + 3\sqrt[3]{9}\right) + o(1) \approx 0.1575006670 + o(1).$$

We show that all large extremal tournaments for C_3^+ essentially look this way, and that the limit object is unique.

Theorem 4. *Let $(T_n)_{n=1}^\infty$ be a sequence of tournaments on n vertices with $I(C_3^+, T_n) = I(C_3^+, n)$. Let $\alpha = \frac{1}{5}(2\sqrt[3]{9} - 2 - \sqrt[3]{3})$. For sufficiently large n, the vertex set of T_n can be partitioned into sets L_n and H_n so that $|H_n| = \alpha n + o(1)$, all arcs between these sets are from H_n to L_n, the sequence of tournaments $(T_n[L_n])_{n=1}^\infty$ is quasi-random, and $I(C_3^+, T_n[H_n]) = I(C_3^+, |H_n|)$. Hence*

$$i(C_3^+) = \frac{1}{8}\left(8 - 9\sqrt[3]{3} + 3\sqrt[3]{9}\right) \approx 0.1575006670.$$

This construction is also mentioned in [24], together with an almost matching upper bound obtained by the flag algebra method. While C_3^+ may not be the most interesting tournament to consider, we find this extremal construction combining quasi-random parts with iterated blow-ups fascinating.

3 Proof of Theorem 2

Proof of Theorem 2. We begin by observing the following identity for all tournaments T on at least 4 vertices:

$$i(C_3, T) = \tfrac{1}{2} i(C_4, T) + \tfrac{1}{4} i(C_3^+, T) + \tfrac{1}{4} i(C_3^-, T). \qquad (1)$$

This follows from the fact that the probability to find a C_3 when picking three vertices at random is equal to the probability to first find C_4, C_3^+, or C_3^- when picking four vertices, times the appropriate probability that removing one of these vertices leaves a C_3.

Multiplying both sides by $\binom{n}{4}$, we can express this relationship in terms of a direct count of induced C_4 for any tournament T:

$$I(C_3, T) \cdot \tfrac{n-3}{4} = \tfrac{1}{2} I(C_4, T) + \tfrac{1}{4}(I(C_3^+, T) + I(C_3^-, T)),$$

implying that

$$I(C_4, T) = I(C_3, T) \cdot \tfrac{n-3}{2} - \tfrac{1}{2}(I(C_3^+, T) + I(C_3^-, T)).$$

Let $T \in \mathscr{C}_n$. Then every vertex in T has out-degree in $\{\frac{n-2}{2}, \frac{n-1}{2}, \frac{n}{2}\}$, so by Proposition 1, $I(C_3, T)$ is maximized. On the other hand, the out-neighborhoods and in-neighborhoods of all vertices in T induce transitive tournaments, so $I(C_3^+, T) = I(C_3^-, T) = 0$. This shows that T maximizes $I(C_4, T)$.

It remains to show that no other tournament shares this property. For this, let T be any $\{C_3^+, C_3^-\}$-free, (near) regular tournament, and let $v_1 \in V(T)$ with $d^+(v_1) = k \in \{\frac{n-2}{2}, \frac{n-1}{2}, \frac{n}{2}\}$. As T is C_3^+-free, the out-neighborhood of v_1 is C_3-free and therefore transitive, and we may relabel the out-neighbors in this induced order as $\{v_2, v_3, \dots, v_{k+1}\}$. Similarly, the in-neighborhood is transitive, and we may relabel it in the induced order as $\{v_{k+2}, \dots, v_n\}$.

Now suppose, for the sake of contradiction, that $T \notin \mathscr{C}_n$, and thus there exists an arc $v_i v_j$ with $0 < i - j < \frac{n}{2}$ or if $-n < i - j < -\frac{n}{2}$. Let us first assume that $0 < i - j < \frac{n}{2}$. As $\{v_2, v_3, \dots, v_{k+1}\}$ and

$\{v_{k+2},\ldots,v_n\}$ are transitively ordered, we have $j \le k$ and $i \ge k+1$. As v_j has out-degree at least $\frac{n-1}{2}$, v_j has an out-neighbor $v_{i'}$ with $i' > i$, implying that $v_i v_{i'} \in E(T)$. But now $T[v_1, v_j, v_i, v_{i'}] \simeq C_3^+$, a contradiction.

Let us now assume that $-n < i-j < -\frac{n}{2}$, and so $i \le k$ and $j > k$. Similarly as before, there now exists a j' with $k < j' < j$ and $v_{j'}v_i \in E$, which again implies that $T[v_1, v_i, v_{j'}, v_j] \simeq C_3^+$, a contradiction proving the theorem. □

4 Proof of Theorem 4

Proof of Theorem 4. We start with an upper bound for the inducibility of C_3^+ using standard flag algebra methods. Notice that the upper bound is not sharp, which is common for extremal constructions involving iterations. We will always assume that n is large enough that we are allowed to suppress lower order terms in our computations.

Claim 4.1. $i(C_3^+, n) \in (0.157500667, 0.157500672)$.

Proof. We know that $i(C_3^+, n) > 0.157500667$ by our construction. Using standard plain flag algebra techniques, we find that

$$i(C_3^+) \le \frac{22050094023517191892318610820045970412528560}{140000000000000000000000000000000000000000000} < 0.157500672.$$

Certificates are too large to be presented here, and do not add much insight. They can be found at http://lidicky.name/pub/tournaments. □

Next, a standard symmetrization argument gives that every vertex is in roughly the same number of C_3^+. Recall that $(T_n)_{n=1}^{\infty}$ is a sequence of tournaments on n vertices with $I(C_3^+, T_n) = I(C_3^+, n)$.

Claim 4.2. *Every vertex is in* $i(C_3^+, T_n)\binom{n-1}{3} + O(n^2)$ *many copies of* C_3^+.

Proof. By definition, the average number of copies a vertex is in is $i(C_3^+, T_n)\binom{n-1}{3}$. Let v be a vertex which is in the fewest copies $C_3^+(v)$, and let w be a vertex which is in the most copies $C_3^+(w)$. Let $C_3^+(vw)$ be the number of copies containing both v and w. If we delete v, and add a copy of w, we gain

$$C_3^+(w) - C_3^+(v) - C_3^+(vw)$$

copies of C_3^+. As T_n is extremal, this quantity must be non-positive. Observing that $C_3^+(vw) = O(n^2)$ shows the claim. □

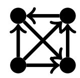
The traditional way to extract structure from flag algebra computations giving sharp bounds is to look for subgraphs, for which the computations tell you that they have zero density in every extremal construction. If the computations do not give sharp bounds like here, another approach is to do the opposite, and to compute bounds on subgraphs which occur with high density. Neither of these approaches has much promise in this problem. As a large part of the conjectured extremal tournament is quasi-random, all subgraphs appear with a fairly moderate frequency.

Inspired by the conjectured extremal tournament, we are looking for other features. A first observation is that the degree distribution is concentrated around a few values. All vertices in L_n have about the same fairly small out-degree, and all vertices in H_n have very large out-degree. A second observation is that all arcs between L_n and H_n are directed from H_n to L_n. We use flag algebra computations to prove that these two observations are true in every extremal tournament, and from this we are able to prove the theorem.

Let $deg(x)$ be the normalized out-degree distribution function for an extremal tournament T_n:

$$deg(x) = \frac{1}{n}\left|\{v \in V(T) : d^+(v) = xn\}\right|.$$

For the remainder of the proof, "normalized" will be suppressed for simplicity. To make our computations more intuitive to follow, we will often denote the quantity $i(H,T_n)$ by a picture of the graph H, so we might write

 $= i(C_3^+, T_n).$

We now show that T_n has a degree distribution similar to the conjectured example.

Claim 4.3. *For all $x \in [0, 0.416] \cup [0.44057, 0.8849]$, $deg(x) = 0$.*

Proof. We prove this claim by showing three bounds. First we investigate vertices v with $d^+(v) \leq 0.85n$ and obtain lower and upper bounds on $d^+(v)$, namely that $d^+(v) \in (0.416, 0.44057)$. For the third bound, we switch to vertices v with $d^+(v) \geq 0.85n$ and show that actually $d^+(v) > 0.8849n$.

We begin with the lower bound of the support of $deg(x)$. Fix some vertex $v \in V(T_n)$ and color all vertices in $N^+(v)$ black and color $N^-(v)$ white. We will use flag algebras to bound the proportion of black vertices in $T_n - v$, and to this end we begin setting up a program that can be bounded by the plain flag algebra method. Since $i(C_3^+, T_n - v) > 0.157500667$, we know the sum of the densities of all 2-colorings of C_3^+ is at least 0.157500667. Since all vertices in $V(T_n)\backslash v$ are colored, we have $\bullet + \circ = 1$. We reduce our search space with the constraint that $\bullet \leq 0.85$, interpreted as v having normalized out-degree at most 0.85.

Ignoring lower order terms, we also know that every vertex is in the same number of C_3^+ (see Claim 4.2), so we can add an additional constraint to reflect this fact. If v plays the role of the source vertex in the C_3^+, then the remaining three vertices are all in $N^+(v)$ and induce a C_3. Otherwise, v plays the role of one of the vertices in the C_3, and the other three vertices induce a transitive triangle where the source and sink are in $N^-(v)$ and the last vertex is in $N^+(x)$. Our coloring scheme thus allows us to include the final bound in the following program:

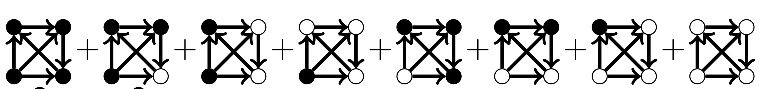
Objective:

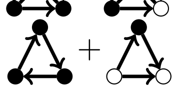
minimize $\bullet$

Constraints:

$\bullet \leq 0.85$

$\bullet + \circ = 1$

$0.157500667 \leq$ 

$0.157500667 \leq$ 

From this program, we find that $\bullet > 0.416$. More precisely,

$$\bullet \geq \frac{374513205377685397063529268488234341918782335394533100}{900000000000000000000000000000000000000000000000000000} > 0.416.$$

Similarly, we obtain $\bullet < 0.44057$, or more precisely that

$$\bullet \leq \frac{396504577626701914630036949179105848845764101947535900}{900000000000000000000000000000000000000000000000000000} < 0.44057,$$

from the following program:

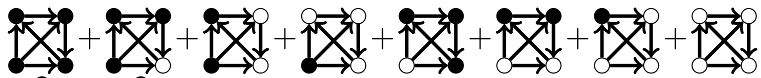
Objective:

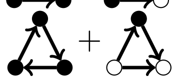
maximize $\bullet$

Constraints:

$\bullet \leq 0.85$

$\bullet + \circ = 1$

$0.157500667 \leq$ 

$0.157500667 \leq$ 

These two results imply that for large enough n, no vertices have normalized out-degree in $[0, 0.416] \cup [0.44057, 0.85]$. We extend this result with the following program restricting the degree of large out-degree vertices:

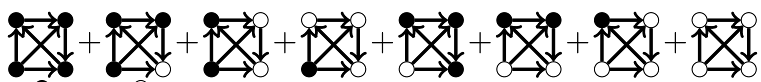
Objective:

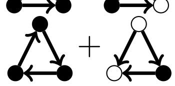
minimize $\bullet$

Constraints:

$\bullet \geq 0.85$

$\bullet + \circ = 1$

$0.157500667 \leq$ 

$0.157500667 \leq$ 

This program outputs the lower bound $\bullet > 0.8849$, completing the proof of this claim. More precisely, it gives

$$\bullet \geq \frac{796434927922041149520351194696218603042290321628262400}{900000000000000000000000000000000000000000000000000000} > 0.88492769769.$$

All certificates can be found at http://lidicky.name/pub/tournaments. □

Let H_n be the set of vertices in T_n with normalized out-degree in $(0.8849, 1]$, and L_n be the set of vertices with normalized out-degree in $(0.416, 0.44057)$. The above claim implies that $H_n \cup L_n = V(T_n)$. We now show that no arcs in T_n are directed from L_n to H_n, once again using a coloring-scheme to acquire localized information in an extremal construction.

Claim 4.4. *For every $x \in L_n$ and $y \in H_n$, $yx \in E(T_n)$.*

Proof. Let $x \in L_n$, $y \in H_n$, so x has normalized out degree in $[0.415, 0.441]$ and y has normalized out-degree at least 0.8849. We color $V(T_n) - \{x, y\}$ with the following scheme, in which the top color represents the relation to x, and the bottom color represents the relation to y (see also Figure 4):

- Assign color black-black ● to $N^+(x) \cap N^+(y)$,
- Assign color black-white ◓ to $N^+(x) \cap N^-(y)$,
- Assign color white-black ◒ to $N^-(x) \cap N^+(y)$,
- Assign color white-white ⊖ to $N^-(x) \cap N^-(y)$.

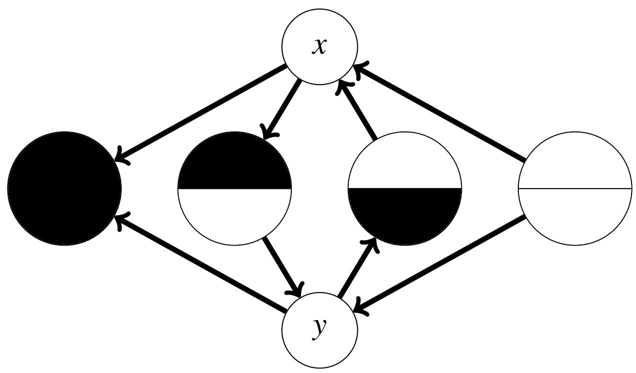


Figure 4: Four-coloring scheme for $T_n - \{x, y\}$
.

In order to model the out-degree assumptions, we will use the following constraints:

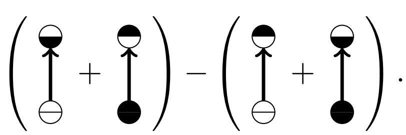
$$0.416 \leq ● + ◓ \leq 0.44057 \quad \text{and} \quad 0.8849 \leq ● + ◒.$$

As in the proof of Claim 4.3, any programs involving this color scheme can include a constraint to ensure that x is in the right number of C_3^+ with vertices in $V(T_n) \backslash y$, and that y is in the right number of C_3^+ with vertices in $V(T_n) \backslash x$.

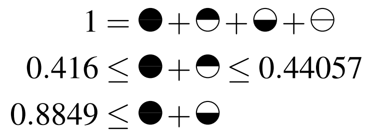
The purpose of this set up is to show that $x \to y$ results in fewer C_3^+ than $y \to x$, so we need to determine how to construct C_3^+ which include both of these vertices. For this, we look again at Figure 4. If $x \to y$, we create a C_3^+ with each arc ⊖→◒ and with each arc ●→◓ . On the other hand, if $y \to x$, we create a C_3^+ with each arc ⊖→◓ and with each arc ●→◒ .

Similarly as above, we can now pose the following program bounding the difference between C_3^+ containing $x \to y$ and containing $y \to x$. Note that there are up to 96 different C_3^+ in $T_n - \{x, y\}$ with 4 colors. Also, when counting the C_3^+ in $T_n - y$ containing x, we have to account for the colors induced by the arcs with y.

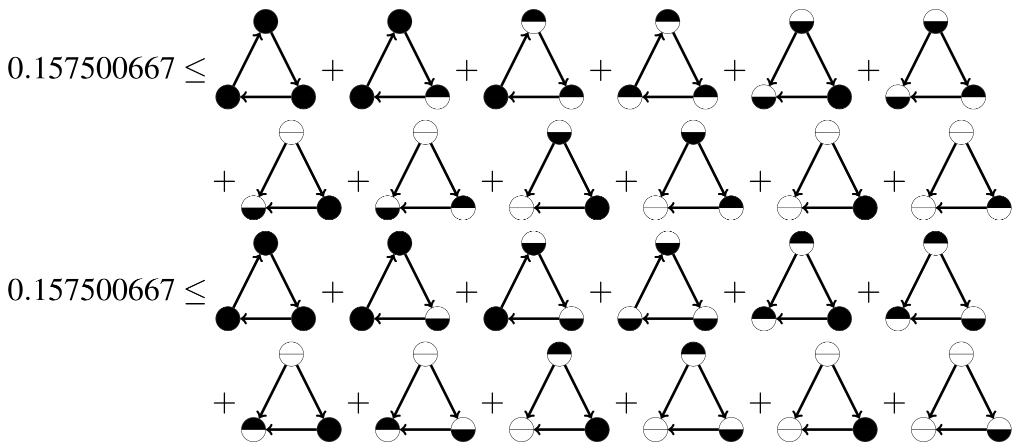
Objective: maximize



Constraints:



$0.157500667 \leq$ sum of all 4-colorings of C_3^+



We find that the solution to this program is bounded above by -0.0768:

$$\frac{-120102985125347247530839107704244200941438 17}{15625000000000000000000000000000000000000000} \approx -0.076866,$$

implying that $(y \to x)$ results in at least $0.0768 \cdot \binom{n}{2}$ more copies of C_3^+ than $(x \to y)$ in T_n for sufficiently large n, proving our claim. Certificates can be found at http://lidicky.name/pub/tournaments. □

Having determined the behavior of the relationship between H_n and L_n, we now focus on the internal behavior of H_n. The following claim implies that, for large enough n, the overall structure of T_n iterates into H_n.

Claim 4.5. $I(C_3^+, T_n[H_n]) = I(C_3^+, |H_n|)$.

Proof. The only copies of C_3^+ in T_n are those chosen completely in H_n, completely in L_n, or with precisely 1 vertex chosen from H_n. The arcs in $T_n[H]$ impact neither the second nor third type of C_3^+. Therefore, $T_n[H_n]$ is extremal and the claim follows. □

We next focus on showing that the sizes of H_n and L_n are correct.

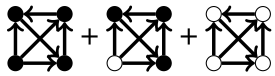
Claim 4.6. $|L_n| < \frac{6}{7}n \approx 0.85714n$.

Proof. First, since there are no arcs from L_n to H_n, the average out-degree of vertices in L_n is $\frac{|L_n|-1}{2}$, so

$$\tfrac{1}{n}|L_n| \in (0.832, 0.88114).$$

We would like a tighter upper bound, so we pose the following program wherein we color the vertices in L_n black and the vertices in H_n white. In this program, we assume that $|L_n| \geq \frac{6}{7}n$ and show that the density of C_3^+ is then too small, implying the claim. We note as well that $\bullet\!\!\longrightarrow\!\circ$ is a forbidden subgraph by Claim 4.4, so we include this as a constraint in the program as well.

Objective:

maximize 

Constraints:

$6/7 \leq ● \leq 0.88114$
$● + ○ = 1$

Forbidden Subgraph:

●⟶○

This program is bounded above by

$$\frac{141750254952413854395355025351649452032106400000000}{900000000000000000000000000000000000000000000000000} < 0.15750003 < i(C_3^+, T_n),$$

thus cannot be extremal, and so implies that T_n must satisfy $|L_n| < \frac{6}{7}n$. Certificates can be found at http://lidicky.name/pub/tournaments. □

We next aim to prove that the sequence $(T_n[L_n])_{n=1}^{\infty}$ is quasi-random. To do so we prove Claim 4.7, a consequence of the characterization of quasi-random tournaments in (Chung and Graham [5]). They list 11 different properties characterizing quasi-random tournaments, but we will only use the first two.

P_1 : Every tournament appears asymptotically with the same density as in the random tournament.

$P_2 : \lim_{n\to\infty} i(C_4, G_n) = \frac{3}{8}$.

Claim 4.7. *A sequence of tournaments $(G_n)_{n=1}^{\infty}$ with $|G| = n$ is quasi-random if and only if*

$$\lim_{n\to\infty} i(C_3, G_n) = \frac{1}{4} \text{ and } \lim_{n\to\infty} i(C_3^+, G_n) = \frac{1}{8}.$$

Proof. The "only if" statement follows immediately from property P_1, so we concern ourselves with proving the "if" statement. Let $(G_n)_{n=1}^{\infty}$ be a sequence of tournaments so that $|G_n| = n$, and recall Proposition 1 that (near) regular tournaments are precisely those tournaments which maximize the number of induced copies of C_3.

So, assume that

$$\lim_{n\to\infty} i(C_3, G_n) = \frac{1}{4} \text{ and } \lim_{n\to\infty} i(C_3^+, G_n) = \frac{1}{8},$$

and observe that the tournaments are tending towards regularity, i.e. there are $n - o(n)$ vertices with out-degree $\frac{n}{2}+o(n)$. Now observe that

$$\begin{aligned}
\frac{1}{8}\binom{n}{4}+o(n^4) &= I(C_3^+,G_n)\\
&= \sum_{v\in V(G_n)} I(C_3,G_n[N^+(v)])\\
&= \sum_{v\in V(G_n)} i(C_3,G_n[N^+(v)])\binom{n/2}{3}+o(n^3), \text{ by regularity}\\
&\le \sum_{v\in V(G_n)} \frac{1}{4}\binom{n/2}{3}+o(n^3)\\
&= \frac{1}{8}\binom{n}{4}+o(n^4).
\end{aligned}$$

This implies that $i(C_3,G_n[N^+(v)]) = \frac{1}{4}+o(1)$ for all but at most $o(n)$ vertices $v\in V(G_n)$. This equality also implies that $i(TT_4,G_n)=\frac{3}{8}+o(1)$. Now

$$\begin{aligned}
\frac{1}{4}+o(1) = i(C_3,G_n) &= \frac{1}{2}i(C_4,G_n)+\frac{1}{4}i(C_3^+,G_n)+\frac{1}{4}i(C_3^-,G_n),\\
\frac{1}{4}+o(1) = \frac{1}{3}i(TT_3,G_n) &= \frac{1}{6}i(C_4,G_n)+\frac{1}{4}i(C_3^+,G_n)+\frac{1}{4}i(C_3^-,G_n)+\frac{1}{3}i(TT_4,G_n),
\end{aligned}$$

so

$$o(1) = i(C_3,G_n)-\frac{1}{3}i(TT_3,G_n) = \frac{1}{3}i(C_4,G_n)-\frac{1}{8}+o(1),$$

and thus $i(C_4,G_n)=\frac{3}{8}+o(1)$. This last statement is property P_2 in [5] which is equivalent to (G_n) being quasi-random. □

Claim 4.8. *In any tournament T on n vertices, $i(C_3^+,T)\le \frac{1}{8}+\frac{2}{3}(\frac{1}{4}-i(C_3,T))+o(1)$.*

Proof. Using flag algebras, we show

$$3i(C_3^+,T)+2i(C_3,T)\le \frac{7}{8}+o(1).$$

The claim follows by rearranging the inequality. Certificates can be found at http://lidicky.name/pub/tournaments. □

Claim 4.9. *The sequence $(T_n[L_n])$ is quasi-random.*

Proof. Let $L=\frac{1}{n}|L_n|$. By Claim 4.5, $i(C_3^+,T_n[H_n]) = i(C_3^+,T_n)+o(1)$. Looking at the density of the C_3^+ which are not completely contained in H_n, we have, ignoring $o(1)$ terms,

$$\begin{aligned}
L^4\tfrac{1}{8}+4(1-L)L^3\tfrac{1}{4} &\le (1-(1-L)^4)i(C_3^+,T_n)\\
&= L^4\cdot i(C_3^+,T_n[L_n])+4(1-L)L^3\cdot i(C_3,T_n[L_n])\\
&\le L^4\cdot\left(\tfrac{1}{8}+\tfrac{2}{3}(\tfrac{1}{4}-i(C_3,T_n[L_n]))\right)+4(1-L)L^3\cdot i(C_3,T_n[L_n])\\
&= \tfrac{7}{24}L^4+L^3 i(C_3,T_n[L_n])(4-\tfrac{14}{3}L)\\
&\le \tfrac{7}{24}L^4+L^3\tfrac{1}{4}(4-\tfrac{14}{3}L) = L^4\tfrac{1}{8}+4(1-L)L^3\tfrac{1}{4}.
\end{aligned}$$

The first inequality is true as the left side is the value we would expect if we replace $T_n[L_n]$ by a random tournament. The second inequality follows from Claim 4.8. For the last inequality, note that $0 < L < \frac{6}{7}$, and so the left side is maximized if and only if C_3 is maximized at $\frac{1}{4}$. As the first and the last term in this chain of inequalities are equal, we have equality throughout. Thus $i(C_3, T_n[L_n]) = \frac{1}{4}$ and $i(C_3^+, T_n[L_n]) = \frac{1}{8}$, proving the claim using Claim 4.7. □

Claim 4.10. *The normalized size of L_n is $L = \frac{1}{5}\left(7 + \sqrt[3]{3} - 2\sqrt[3]{9}\right) + o(1)$, and our conjectured structure is the limit object for the inducibility of C_3^+.*

Proof. We know that $L < 6/7$, that $T_n[L_n]$ is quasi-random, that all arcs between H_n and L_n point towards the vertex in L_n, and that $i(C_3^+, T_n[H_n]) = i(C_3^+, T_n) + o(1)$. Thus,

$$i(C_3^+, T_n) = L^4 \cdot \frac{1}{8} + \binom{4}{1} L^3(1-L) \cdot \frac{1}{4} + (1-L)^4 (i(C_3, T_n) + o(1)).$$

This is maximized when $i(C_3^+, T_n) = \frac{1}{8}\left(8 - 9\sqrt[3]{3} + \sqrt[3]{3^5}\right) + o(1)$ and $L = 1 - \alpha + o(1)$. □

We have thus shown that every extremal tournament matches our construction, completing the proof of this theorem. □

5 Discussion

In this section, we discuss some of the peculiarities of this problem and its solutions, including the novel strategies introduced in this paper. First and foremost, we know of no other inducibility problem for which all extremal constructions include a quasi-random component as in the case of C_3^+ and C_3^- and ask the following question:

Problem 1. *For what classes of graphs (undirected or directed) do the extremal constructions for the corresponding inducibility problem involve non-trivial quasi-random components?*

For C_3^+, the extremal construction was conjectured by noting that our tournament can be decomposed into a source vertex and a C_3; described another way, we begin with an arc and blow up the head into a C_3. Essentially, we ask the following: for a digraph $G = (V, E)$ with cut $C = (S, T)$ and cut-set of size $|S| \cdot |T|$, for what structures $G[S]$ and $G[T]$ does the resulting inducibility problem have as extremal solutions constructions for which $\alpha \cdot 100\%$ of the vertices induce a "typical random graph structure" for some $\alpha \in (0, 1)$? Natural candidates for consideration would include $G[T] \cong C_3$ and $G[S]$ isomorphic to any 2-vertex digraph or 3-vertex tournament.

Historically, flag algebra techniques have been leveraged to determine bounds on global graph densities. The models developed in Claims 4.3 and 4.4, however, resulted in bounds on localized information. In the case of Claim 4.3, we were able to determine something very powerful regarding the distribution of out-degrees in extremal constructions, namely that all vertices have normalized out-degrees in a very specific set. In the case of Claim 4.4, we were able to determine the direction of an arc between any pair of vertices which satisfy basic constraints related to their out-degrees.

Finally, we want to make an observation about Conjecture 3. Let $k \geq 5$ be odd, and let $n > k$. Let $X \subset V(C_n)$ be a set of k vertices such that $C_n[X] \cong C_k$. Observe that for every vertex $v \in$

$V(C_n) \setminus X$, we have $C_n[X \cup \{x\}] \cong C_{k+1}$. If we now express $i(C_k, T)$ in a tournament T in terms of densities of $(k+1)$-vertex graphs similarly to (1), we can easily conclude that Conjecture 3 is true for $k+1$ if it is true for k, so it suffices to prove it for all odd k. Standard flag algebra computations show that C_n is $o(n^2)$ arc flips away from every extremal tournament for C_5 and C_7 (and thus for C_6 and C_8 by this observation), but we have not seriously tried to show the full conjecture for these cases.

References

[1] József Balogh, Ping Hu, Bernard Lidický, and Florian Pfender. Maximum density of induced 5-cycle is achieved by an iterated blow-up of 5-cycle. *European J. Combin.*, 52(part A):47–58, 2016. doi:10.1016/j.ejc.2015.08.006.

[2] Adam Blumenthal and Michael Phillips. Inducibility of the net graph, 2021. arXiv:2103.06350.

[3] Jason I. Brown and Alexander Sidorenko. The inducibility of complete bipartite graphs. *J. Graph Theory*, 18(6):629–645, 1994. doi:10.1002/jgt.3190180610.

[4] Ilkyoo Choi, Bernard Lidický, and Florian Pfender. Inducibility of directed paths. *Discrete Math.*, 343(10):112015, 11, 2020. doi:10.1016/j.disc.2020.112015.

[5] Fan R. K. Chung and Ron L. Graham. Quasi-random tournaments. *J. Graph Theory*, 15(2):173–198, 1991. doi:10.1002/jgt.3190150206.

[6] Éva Czabarka, Audace A. V. Dossou-Olory, László A. Székely, and Stephan Wagner. Inducibility of d-ary trees. *Discrete Math.*, 343(2):111671, 15, 2020. doi:10.1016/j.disc.2019.111671.

[7] Chaim Even-Zohar and Nati Linial. A note on the inducibility of 4-vertex graphs. *Graphs Combin.*, 31(5):1367–1380, 2015. doi:10.1007/s00373-014-1475-4.

[8] Victor Falgas-Ravry and Emil R. Vaughan. Turán H-densities for 3-graphs. *Electron. J. Combin.*, 19(3):Paper 40, 26, 2012. doi:10.37236/2733.

[9] Jacob Fox, Hao Huang, and Choongbum Lee. Packing problems. Methods and Challenges in Extremal and Probabilistic Combinatorics workshop, Banff International Research Station (2015). Video available at http://www.birs.ca/events/2015/5-day-workshops/15w5008/videos/watch/201508240900-Fox.html, 2015.

[10] Jacob Fox, Lisa Sauermann, and Fan Wei. On the inducibility problem for random cayley graphs of abelian groups with a few deleted vertices, 2019. arXiv:1904.07682.

[11] Hamed Hatami, James Hirst, and Serguei Norine. The inducibility of blow-up graphs. *J. Combin. Theory Ser. B*, 109:196–212, 2014. doi:10.1016/j.jctb.2014.06.005.

[12] Dan Hefetz and Mykhaylo Tyomkyn. On the inducibility of cycles. *J. Combin. Theory Ser. B*, 133:243–258, 2018. doi:10.1016/j.jctb.2018.04.008.

[13] Ping Hu, Bernard Lidický, Florian Pfender, and Jan Volec. Inducibility of orientations of C_4. Manuscript.

[14] Ping Hu, Jie Ma, Sergey Norin, and Hehui Wu. The inducibility of oriented stars, 2020. arXiv:2008.05430.

[15] Hao Huang. On the maximum induced density of directed stars and related problems. *SIAM J. Discrete Math.*, 28(1):92–98, 2014. doi:10.1137/130912931.

[16] Daniel Král', Sergey Norin, and Jan Volec. A bound on the inducibility of cycles. *J. Combin. Theory Ser. A*, 161:359–363, 2019. doi:10.1016/j.jcta.2018.08.003.

[17] Bernard Lidický, Connor Mattes, and Florian Pfender. C_5 is almost a fractalizer, 2021. arXiv:2102.06773.

[18] Hong Liu, Oleg Pikhurko, Maryam Sharifzadeh, and Katherine Staden. Stability from graph symmetrisation arguments with applications to inducibility, 2020. arXiv:2012.10731.

[19] Natasha Morrison and Alex Scott. Maximising the number of induced cycles in a graph. *J. Combin. Theory Ser. B*, 126:24–61, 2017. doi:10.1016/j.jctb.2017.03.007.

[20] Dhruv Mubayi and Alexander A. Razborov. Polynomial to exponential transition in Ramsey theory. *Proc. Lond. Math. Soc.*, 122:69–92, 2021. doi:10.1112/plms.12320.

[21] Nicholas Pippenger and Martin Charles Golumbic. The inducibility of graphs. *J. Combin. Theory Ser. B*, 19(3):189–203, 1975. doi:10.1016/0095-8956(75)90084-2.

[22] Alexander A. Razborov. Flag algebras. *J. Symbolic Logic*, 72(4):1239–1282, 2007. doi:10.2178/jsl/1203350785.

[23] Raphael Yuster. On the exact maximum induced density of almost all graphs and their inducibility. *J. Combin. Theory Ser. B*, 136:81–109, 2019. doi:10.1016/j.jctb.2018.09.005.

[24] Łukasz Bożyk, Andrzej Grzesik, and Bartłomiej Kielak. On the inducibility of oriented graphs on four vertices, 2020. arXiv:2010.11664.