Maximum density of induced C_5 is achieved by an iterated blow-up of C_5

József Balogh* Ping Hu† Bernard Lidický‡ Florian Pfender§

October 7, 2014

Abstract

Let $C(n)$ denote the maximum number of induced copies of C_5 in graphs on n vertices. For n large enough, we show that $C(n) = a \cdot b \cdot c \cdot d \cdot e + C(a) + C(b) + C(c) + C(d) + C(e)$, where $a + b + c + d + e = n$ and a, b, c, d, e are as equal as possible.

Moreover, if n is a power of 5, we show that the unique graph on n vertices maximizing the number of induced 5-cycles is an iterated blow-up of a 5-cycle.

The proof uses flag algebra computations and stability methods.

1 Introduction

In 1975, Pippinger and Golumbic [23] conjectured that the maximum induced density of a k-cycle is $k!/(k^k - k)$ if $k \geq 5$. In this paper we solve their conjecture for $k = 5$. In addition, we also show that the extremal limit object is unique. The problem of maximizing the induced density of C_5 is also presented on `http://flagmatic.org` as one of the problems where the plain flag algebra method was applied but failed to provide an exact result. It was also mentioned by Razborov [28] during his talk at the Probabilistic and Extremal Combinatorics Workshop, which was part of the IMA Annual Program 2014.

Problems of maximizing the number of induced copies of a fixed small graph H have attracted a lot of attention recently [9, 15, 32]. For a list of other results on this so called inducibility of small graphs of order up to 5, see he work of Even-Zohar and Linial [9].

*Department of Mathematics, University of Illinois, Urbana, IL 61801, USA and Bolyai Institute, University of Szeged, Szeged, Hungary E-mail: `jobal@math.uiuc.edu`. Research is partially supported by Simons Fellowship, NSF CAREER Grant DMS-0745185, Arnold O. Beckman Research Award (UIUC Campus Research Board 13039) and Marie Curie FP7-PEOPLE-2012-IIF 327763.

†Department of Mathematics, University of Illinois, Urbana, IL 61801, USA, E-mail: `pinghu1@math.uiuc.edu`.

‡Department of Mathematics, Iowa State University, Ames, IA, E-mail: `lidicky@iastate.edu`.

§Department of Mathematical and Statistical Sciences, University of Colorado Denver, E-mail: `Florian.Pfender@ucdenver.edu`. Research is partially supported by a collaboration grant from the Simons Foundation.

In this paper, we use a method we originally developed for maximizing the number of rainbow triangles in 3-edge-colored complete graphs [5]. However, the application of the method to the C_5 problem is less technical, and therefore this paper is a more accessible exposition of this new method.

Denote the $(k-1)$-times iterated blow-up of C_5 by $C_5^{k\times}$, see Figure 1. Let $\mathcal{G}_n$ be the set of all graphs on n vertices, and denote by $C(G)$ the number of induced copies of C_5 in a graph G. Define

$$C(n) = \max_{G \in \mathcal{G}_n} C(G).$$

We say a graph $G \in \mathcal{G}_n$ is *extremal* if $C(G) = C(n)$. Notice that, as C_5 is self-complementary, G is extremal if and only if its complement is extremal. If n is a power of 5, we can exactly determine the unique extremal graph and thus $C(n)$.

Theorem 1. *For $k \geq 1$, the unique extremal graph in $\mathcal{G}_{5^k}$ is $C_5^{k\times}$.*

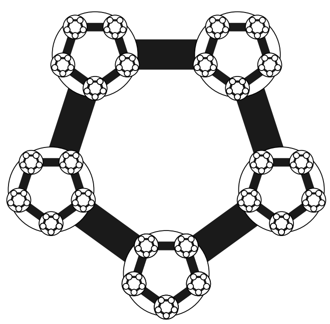

Figure 1: Graph $C_5^{k\times}$ maximizing number of induced C_5s.

To proof Theorem 1, we first proof the following theorem. Note that this theorem is sufficient to determine the unique limit object (the graphon) maximizing the density of induced copies of C_5.

Theorem 2. *There exists n_0 such that for every $n \geq n_0$*

$$C(n) = a \cdot b \cdot c \cdot d \cdot e + C(a) + C(b) + C(c) + C(d) + C(e),$$

where $a+b+c+d+e = n$ and a, b, c, d, e are as equal as possible.

Moreover, if $G \in \mathcal{G}_n$ is an extremal graph, then $V(G)$ can be partitioned into five sets X_1, X_2, X_3, X_4 and X_5 of sizes a, b, c, d and e respectively, such that for $1 \leq i < j \leq 5$ and $x_i \in X_i$, $x_j \in X_j$, we have $x_i x_j \in E(G)$ if and only if $j - i \in \{1, 4\}$.

In the next section, we give a brief overview of our method. In Section 3 we prove Theorem 2, and Theorem 1 in Section 4.

2 Method and Flag Algebras

Our method relies on the theory of flag algebras, a tool developed by Razborov [25]. Flag algebras can be used as a general tool to attack problems from extremal combinatorics. Flag algebras were used for a wide range of problems, for example the Caccetta-Häggkvist conjecture [17, 24], Turán-type problems in graphs [8, 12, 14, 20, 22, 26, 29, 30], 3-graphs [3, 10, 11] and hypercubes [1, 4], extremal problems in a colored environment [2, 5, 7], and also to problems in geometry [19] or extremal theory of permutations [6]. For more details on these applications, see a recent survey of Razborov [27].

A typical application of the so called *plain flag algebra method* provides a bound on densities of substructures. In some cases the bound is sharp. This happens most often when the extremal construction is 'clean', for example a blow-up of a small graph. Obtaining an exact result from the sharp bound usually consists of first bounding the densities of some small substructures by $o(1)$, which can be read off from the flag algebra computation. Forbidding these structures can yield a lot of structure of the extremal structure. Finally, stability arguments are used to extract the precise extremal structure.

Blow-ups of small graphs appear very often as extremal graphs, in fact there are large families of graphs whose extremal graphs for the inducibility are of this type, see a recent paper by Hatami, Hirst and Norin [13]. However, there are also many questions where the extremal construction is an iterated blow-up as shown by Pikhurko [21].

For our problem, the conjectured extremal graph has such an iterated structure, for which it is quite rare to obtain the precise density from plain flag algebra computations alone. One such rare example is the problem to determine the inducibility of small out-stars in oriented graphs [10] (note that the problem of inducibility of all out-stars was recently solved by Huang [18] using different techniques). Hladký, Kráľ and Norin [16] announced that they found the inducibility of the oriented path of length 2, which also has an iterated extremal construction, via a flag algebra method. Other than these two examples and [5], we are not aware of any applications of flag algebras which completely determined an iterative structure.

For our problem, a direct application of the plain method gives an upper bound on the limit value and shows that $\lim_{n\to\infty} \frac{C(n)}{\binom{n}{5}} < 0.03846157$, which is slightly more than the density of C_5 in the conjectured extremal construction, which is $\frac{1}{26} \approx 0.03846154$. This difference may appear very small, but the bounds on densities of subgraphs not appearing in the extremal structure are too weak to allow the standard methods to work.

In our method, we instead use flag algebras to find bounds on densities of other subgraphs, which appear with fairly high density in the extremal graph. This enables us to better control the slight lack of performance of the flag algebra bounds as these small errors have a weaker relative effect on larger densities.

3 Proof of Theorem 2

:thmrecurs

In our proofs we consider densities of 7-vertex subgraphs. Guided by their prevalence in the conjectured extremal graph, the following two types of graphs will play an important role. We call a graph $C22111$ if it can be obtained from C_5 by duplicating two vertices. We call a graph $C31111$ if it can be obtained from C_5 by tripling one vertex. The edges between the original vertices and their copies are not specified, and there are two complementary types of $C22111$, depending on the adjacency of the two doubled vertices in C_5. So technically $C22111$ and $C31111$ denote collections of several graphs. Examples of $C22111$ and $C31111$ are depicted in Figure 2. We slightly abuse notation by using $C22111$ and $C31111$ also to denote the densities of these graphs, i.e., the probability that randomly chosen 7 vertices induce the appropriate 7-vertex blow-up of C_5. Moreover, for a set of vertices Z we denote by $C22111(Z)$ and $C31111(Z)$ the densities of $C22111$ and $C31111$ containing Z, i.e., for a graph G on n vertices, $C22111(Z)$ ($C31111(Z)$) is the number of $C22111$($C31111$) containing Z divided by $\binom{n-|Z|}{7-|Z|}$.

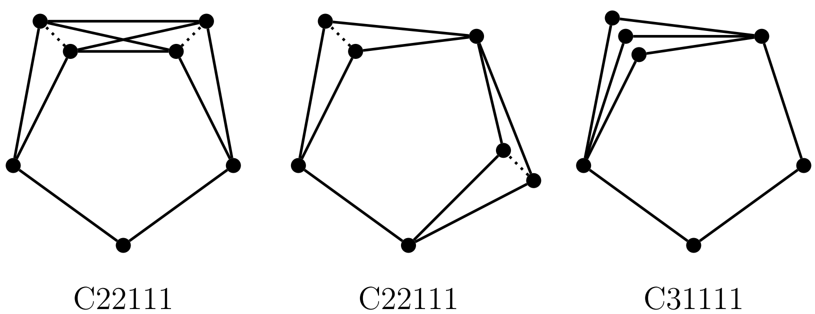

Figure 2: C22111 and C31111. fig-conf

We start with the following statement.

prop:flag **Proposition 3.** *There exists n_0 such that every extremal graph G on at least n_0 vertices satisfies:*

$$C_5 < 0.03846157; \qquad 0.0032241809 < 4 \cdot C22111 - 11.94 \cdot C31111. \tag{1}$$

diff

Proof. This follows from a standard application of the plain flag algebra method. For the second inequality, we minimize the right side with the extra constraint that $C_5 \geq \frac{1}{26}$. For certificates, see `http://math.uiuc.edu/~jobal/cikk/c5/`. □

The bounds in Proposition 3 result from flag algebra computations on 7-vertex graphs. While it is possible to perform flag algebra computations on 8 vertices, the computational effort is very large, and the resulting slightly improved bounds have very little effect on the remainder of the proof. In particular, the bounds are not strong enough for the standard method.

The expressions from Proposition 3 compare to the following limiting values in the iterated blow-up $C_5^{k\times}$, where $k \to \infty$:

$$C_5 = \frac{1}{26} \approx 0.03846154; \quad 4 \cdot C22111 - 11.94 \cdot C31111 = 4 \cdot \frac{5}{31} - 11.94 \cdot \frac{5}{93} \approx 0.0032258.$$

Notice that in the iterated blow-up of C_5, in the limit $4 \cdot C22111 - 12 \cdot C31111 = 0$. For our method to work, we need a lower bound greater than zero. On the other hand, experiments convinced us that the method works best if the bound is only slightly above zero, where a suitable factor is again determined by experiments.

Let G be an extremal graph on n vertices, where n is sufficiently large to apply Proposition 3. Denote the set of all induced C_5s in G by $\mathcal{Z}$. We assume that $a \in \mathbb{R}$ and $Z = z_1z_2z_3z_4z_5$ is an induced C_5 maximizing $C22111(Z) - a \cdot C31111(Z)$. Then

$$\begin{aligned}(C22111(Z) - a \cdot C31111(Z))\binom{n-5}{2} &\geq \frac{1}{|\mathcal{Z}|}\sum_{Y \in \mathcal{Z}}(C22111(Y) - a \cdot C31111(Y))\binom{n-5}{2} \\ \geq \frac{(4 \cdot C22111 - 3a \cdot C31111)\binom{n}{7}}{C_5\binom{n}{5}} &= \frac{\frac{4}{21}C22111 - \frac{a}{7}C31111}{C_5}\binom{n-5}{2}.\end{aligned}$$

As mentioned above, experiments indicate that we get the most useful bounds if $C22111(Z) - a \cdot C31111(Z)$ is close but not too close to 0. Using (1) and $a = 3.98$, we get

$$C22111(Z) - 3.98 \cdot C31111(Z) > 0.00399184. \tag{2}$$

main2

For $1 \leq i \leq 5$, we define sets of vertices Z_i which look like z_i to the other vertices of Z. Formally,

$$Z_i := \{v \in V(G) : G[(Z \setminus z_i) \cup v] \cong C_5\} \text{ for } 1 \leq i \leq 5.$$

Note that $Z_i \cap Z_j = \emptyset$ for $i \neq j$. We call a pair v_iv_j *funky*, if v_iv_j is an edge while z_iz_j is not an edge or vice versa, where $v_i \in Z_i$, $v_j \in Z_j$, $1 \leq i < j \leq 5$. In other words, $G[Z \cup \{v_i, v_j\}] \ncong C22111$, i.e., every funky pair destroys a potential copy of $C22111(Z)$. Denote by E_f the set of funky pairs. With this notation, (2) implies that

$$\sum_{1 \leq i < j \leq 5} |Z_i||Z_j| - |E_f| - 3.98 \sum_{i \in [5]} |Z_i|^2/2 > 0.00399184\binom{n-5}{2}.$$

For any choice of sets $X_i \subseteq Z_i$, where $i \in [5]$, let $X_0 := V(G) \setminus \bigcup X_i$. Let f be the number of funky pairs not incident to vertices in X_0, divided by n^2 for normalization, and denote $x_i = \frac{1}{n}|X_i|$ for $i \in \{0, \ldots, 5\}$. Choose the X_i (possibly $X_i = Z_i$) such that the left hand side in

$$2\sum_{1 \leq i < j \leq 5} x_ix_j - 2f - 3.98\sum_{i \in [5]} x_i^2 > 0.00399184 \tag{3}$$

main3

is maximized. In order to simplify notation, we use $X_{i+5} = X_i$ and $x_{i+5} = x_i$ for all $i \geq 1$.

Claim 4. *The following equations are satisfied:*

$$0.198855061 < x_i < 0.201144939 \quad \text{for} \quad i \in [5]; \tag{4} \qquad \boxed{\texttt{xbound}}$$

$$x_0 < 0.00102055255; \tag{5} \qquad \boxed{\texttt{Xzeromax}}$$

$$f < 0.00000408012713. \tag{6} \qquad \boxed{\texttt{funky}}$$

Proof. We solve a program (P) where the objective function is one of the desired bounds and constraints are (3) and $\sum_{i=0}^{5} x_i = 1$. We first solve (4). By symmetry, bounds for x_1 will work also for x_2, x_3, x_4 and x_5. Hence it suffices to bound only x_1. We claim that if (P) has a feasible solution S, then there exists a feasible solution S' where

$$S'(x_1) = S(x_1), \quad S'(f) = 0, \quad S'(x_0) = S(x_0)$$
$$S'(x_2) = S'(x_3) = S'(x_4) = S'(x_5) = \frac{1}{4}\big(1 - S(x_1) - S(x_0)\big).$$

Since x_2, x_3, x_4 and x_5 appear only in constraints, we only need to check if (3) is satisfied. The left hand side of (3) can be rewritten as

$$\begin{aligned} &2x_1 \sum_{2\le i<j\le 5} x_i + 2 \sum_{2\le i<j\le 5} x_i x_j - 3.98 \sum_{1\le i<j\le 5} x_i^2 - 2f \\ =\ &2x_1 \sum_{2\le i<j\le 5} x_i - \sum_{2\le i<j\le 5} (x_i - x_j)^2 - 0.98 \sum_{2\le i<j\le 5} x_i^2 - 3.98 x_1^2 - 2f \end{aligned}$$

Note that the term $\sum_{2\le i<j\le 5}(x_i - x_j)^2$ is minimized if $x_i = x_j$ for all $i, j \in \{2, 3, 4, 5\}$. The term $x_2^2 + x_3^2 + x_4^2 + x_5^2$, subject to $x_2 + x_3 + x_4 + x_5$ being a constant, is also minimized if $x_i = x_j$ for all $i, j \in \{2, 3, 4, 5\}$. Since $f \ge 0$, the term $2f$ is minimized when $f = 0$. Hence (3) is satisfied by S' and we can add the constraints $x_2 = x_3 = x_4 = x_5$ and $f = 0$ to bound x_1. The resulting program (P') is

$$(P') \begin{cases} \text{minimize} & x_1 \\ \text{subject to} & x_0 + x_1 + 4y = 1, \\ & 8x_1 y - 0.98 \cdot 4y^2 - 3.98 x_1^2 \ge 0.00399184, \\ & x_0, x_1, y \ge 0. \end{cases}$$

We solve (P') using Lagrange multipliers. We delegate the work to Sage [31] and we provide the Sage script at `http://math.uiuc.edu/~jobal/cikk/c5/`. Finding an upper bound on x_1 is done by changing the objective to maximization.

Similarly, we can set $x_1 = x_2 = x_3 = x_4 = x_5 = 1/5$ to get an upper bound on f. We can set $f = 0$ and $x_1 = x_2 = x_3 = x_4 = x_5 = (1 - x_0)/5$ to get an upper bound on x_0. We omit the details. Sage scripts for solving the resulting programs are provided at `http://math.uiuc.edu/~jobal/cikk/c5/`. □

Furthermore, for any vertex $v \in X_i, i \in [5]$ we use $d_f(v)$ to denote the number of funky pairs from v to $(X_1 \cup X_2 \cup X_3 \cup X_4 \cup X_5) \setminus X_i$ after normalizing by n. If we move v from X_1 to X_0, then the left hand of (3) will decrease by

$$\frac{1}{n}\left(2(x_2+x_3+x_4+x_5) - 2d_f(v) - 2\cdot 3.98\cdot x_1 + o(1)\right).$$

If this quantity was negative, then the left hand of (3) could be increased by moving v to X_0, contradicting our choice of X_i. This together with (4) implies that

$$d_f(v) \le x_2+x_3+x_4+x_5 - 3.98\cdot x_1 + o(1) \le 1 - 4.98\cdot x_1 + o(1) \le 0.009701795. \quad (7)$$ maxfunky

Symmetric statements hold also for every vertex $v \in X_2 \cup X_3 \cup X_4 \cup X_5$.

nofunky **Claim 5.** *There are no funky pairs.*

Proof. Assume there is a funky pair uv. It is enough to consider two cases, either $u \in X_1, v \in X_2$ or $u \in X_1, v \in X_3$. Here we check the case where $u \in X_1, v \in X_2$, so uv is not an edge. The other case follows from considering the complement of G.

Let G' be a graph obtained from G by adding the edge uv, i.e., changing uv to be not funky. We compare the number of induced C_5s containing $\{u,v\}$ in G and in G'. In G', there are at least

$$[x_3x_4x_5 - (d_f(u)+d_f(v))\max\{x_3x_4, x_3x_5, x_4x_5\} - f\cdot\max\{x_3,x_4,x_5\}]\, n^3$$

induced C_5s containing uv, since we can pick one vertex from each of X_3, X_4, X_5 to form an induced C_5 as long as none of the resulting nine pairs is funky.

Now we count the number of induced C_5s in G containing $\{u,v\}$. The number of such C_5s which contain vertices from X_0 is upper bounded by $x_0n^3/2$. Next we count the number of such C_5s avoiding X_0. Observe that there are no C_5s avoiding X_0 in which uv is the only funky pair.

The number of C_5s containing another funky pair $u'v'$ with $\{u,v\} \cap \{u',v'\} = \emptyset$ can be upper bounded by fn^3. We are left to count C_5s where the other funky pairs contain u or v. The number of C_5s containing at least two vertices other than u and v which are in funky pairs can be upper bounded by $(d_f(u)^2/2 + d_f(v)^2/2 + d_f(u)d_f(v))n^3$.

It remains to count only C_5s containing exactly one vertex w where uw and vw are the options for funky pairs. The number of choices of w is at most $(d_f(u)+d_f(v))n$. As $\{u,v,w\}$ is a part of an induced C_5, the set $\{u,v,w\}$ induces a path in either G or the complement of G. Let the middle vertex of that path be in X_i. If $G[\{u,v,w\}]$ is a path, then the remaining two vertices of a C_5 cannot be in $X_{i+1} \cup X_{i+4}$. If $G[\{u,v,w\}]$ is the complement of a path, then the remaining two vertices cannot be in $X_{i+2} \cup X_{i+3}$. Hence the remaining two vertices of a C_5 containing $\{u,v,w\}$ can be chosen from at most $3\max\{x_i\}n$ vertices. So we get an upper bound of $(d_f(u)+d_f(v))n\binom{3\max\{x_i\}n}{2}$ such C_5s.

Now we compare the number of induced C_5s containing uv in G and in G' . We use x_{max} and x_{min} to denote the upper and lower bound respectively from (4), use d_f to denote the upper bound on $d_f(u)$ and $d_f(v)$ from (7), and also use bounds from (5) and (6). The number of C_5s containing uv divided by n^3 is

$$\text{in } G: \leq x_0/2 + f + 2d_f^2 + 9d_f x_{max}^2 \leq 0.0043;$$
$$\text{in } G': \geq (x_{min} - 2d_f)x_{min}^2 - f x_{max} \geq 0.007.$$

This contradicts the extremality of G. □

Next, we want to show that $X_0 = \emptyset$. For this, suppose that there exists $x \in X_0$. We will add x to one of the X_i, $i \in [5]$ such that $d_f(x)$ is minimal. By symmetry, we may assume that x is added to X_1. Note that adding a single vertex to X_1 does not change any of the density bounds we used above by more than $o(1)$.

Claim 6. *For every $x \in X_0$, if x is added to X_1 then $d_f(x) \geq 0.0919109388238$.*

Proof. Let xw be a funky pair, where $w \in X_2$. The case where $w \in X_3$ can be argued the same way by considering the complement of G. Let G' be obtained from G by adding the edge xw. Since G is extremal, we have $C(G') \leq C(G)$. The following analysis is similar to the proof of Claim 5. However, we can say a bit more since every funky pair contains x.

First we count induced C_5s containing xw in G. The number of induced C_5s containing xw and other vertices from X_0 is easily bounded from above by $x_0 n^3/2$.

Let F be an induced C_5 in G containing xw and avoiding $X_0 \setminus \{x\}$. Since all funky pairs contain x, $F - x$ is an induced path $p_0p_1p_2p_3$ without funky pairs. Either $p_j \in X_2$ for all $j \in \{0,1,2,3\}$ or there is $i \in \{1,2,3,4,5\}$ such that $p_j \in X_{i+j}$ for all $j \in \{0,1,2,3\}$. The first case is depicted in Figure 3(a). Consider now the second case. If $i \in \{2,3,4\}$, then $xp_0p_1p_2p_3$ does not satisfy the definition of F. Hence $i \in \{1,5\}$ and the possible C_5s are depicted in Figure 3(b)(c). In all cases, F contains exactly two funky pairs, xw and xy. The location of y entirely determines $F - x$. Hence the number of induced C_5s containing xw is at most $d_f(x)x_{max}^2 n^3$.

In G', there are at least $\big(x_3x_4x_5 - d_f(x)\max\{x_3x_4, x_3x_5, x_4x_5\}\big)n^3$ induced C_5s containing xw. We obtain

$$C(G)/n^3 \leq d_f(x)x_{max}^2 + x_0/2;$$
$$C(G')/n^3 \geq (x_{min} - d_f(x))x_{min}^2.$$

Since $C(G') \leq C(G)$, we have

$$(x_{min} - d_f(x))x_{min}^2 \leq d_f(x)x_{max}^2 + x_0/2,$$

which together with (4) and (5) gives $d_f(x) \geq 0.0919109388238$.

□

Claim 7. *Every vertex of the extremal graph G is in at least $(1/26+o(1))\binom{n}{4} \approx 0.001602564n^4$ induced C_5s.*

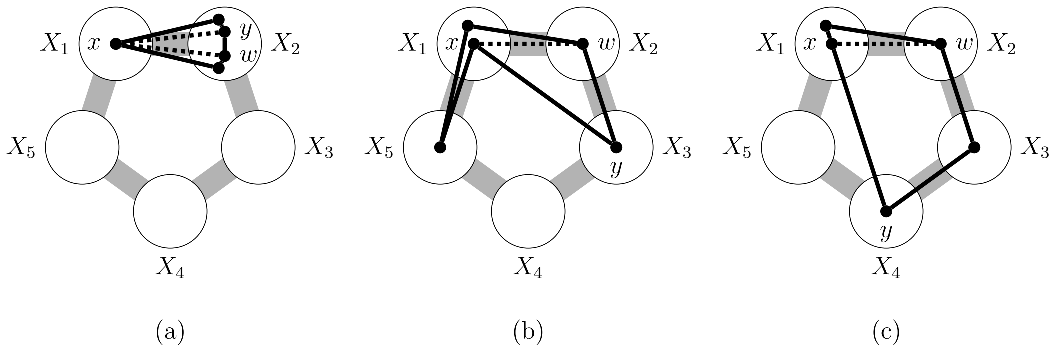

Figure 3: Possible C_5s with funky pair xw. They all have exactly one other funky pair xy. fig-funky

Proof. For every vertex $u \in V(G)$, denote by C_5^u the number of C_5s in G containing u. For any two vertices $u, v \in V(G)$, we show that $C_5^u - C_5^v < n^3$. This implies Claim 7. Denote by C_5^{uv} the number of C_5s in G containing both u and v. A trivial bound is $C_5^{uv} \le \binom{n-2}{3}$.

Let G' be obtained from G by deleting v and duplicating u to u', i.e., for every vertex x we add the edge xu' iff xu is an edge. As G is extremal we have

$$0 \ge C(G') - C(G) \ge C_5^u - C_5^v - C_5^{uv} \ge C_5^u - C_5^v - \binom{n-2}{3}.$$

□

X0empty **Claim 8.** *X_0 is empty.*

Proof. Assume $x \in X_0$, then we count the number of induced C_5s containing x. Our goal is to show that C_5^x is smaller than the value in Claim 7. Let $a_i n$ be the number of neighbors of x in X_i and $b_i n$ be the number of non-neighbors of x in X_i for $i \in \{0, 1, 2, 3, 4, 5\}$.

The number of C_5s where the other four vertices are in $X_1 \cup X_2 \cup X_3 \cup X_4 \cup X_5$ is upper bounded by

$$\left(a_1 b_2 b_3 a_4 + a_2 b_3 b_4 a_5 + a_3 b_4 b_5 a_1 + a_4 b_5 b_1 a_2 + a_5 b_1 b_2 a_3 + \frac{1}{4} \sum_{i=1}^{5} a_i^2 b_i^2 \right) n^4.$$

The variables a_i, b_i satisfy (4) and Claim 6. Moreover, we also need to include the cases that the C_5s can contain vertices from X_0, which we do very generously by increasing all variables by a_0 or b_0.

So we want to solve the following program:

$$(P)\begin{cases}\text{maximize} & \sum_{i=1}^{5}(a_i+a_0)(b_{i+1}+b_0)(b_{i+2}+b_0)(a_{i+3}+a_0)+\frac{1}{4}\sum_{i=1}^{5}a_i^2b_i^2\\ \text{subject to} & \sum_{i=0}^{5}(a_i+b_i)=1,\\ & 0.198855061\le a_i+b_i\le 0.201144939 \text{ for } i\in\{1,2,3,4,5\},\\ & a_0+b_0\le 0.00102055255,\\ & b_2+b_5+a_3+a_4\ge 0.0919109388238,\\ & b_1+b_3+a_4+a_5\ge 0.0919109388238,\\ & b_2+b_4+a_1+a_5\ge 0.0919109388238,\\ & b_3+b_5+a_1+a_2\ge 0.0919109388238,\\ & b_4+b_1+a_2+a_3\ge 0.0919109388238,\\ & a_i,b_i\ge 0 \text{ for } i\in\{0,1,2,3,4,5\}.\end{cases}$$

Instead of solving (P) we solve a slight relaxation (P') with increased upper bounds on a_i+b_i, which allows us to drop a_0 and b_0. Since the objective function is maximizing, we can claim that a_i+b_i is always as large as possible, which decreases the degrees of freedom.

$$(P')\begin{cases}\text{maximize} & f=\sum_{i=1}^{5}a_ib_{i+1}b_{i+2}a_{i+3}+\frac{1}{4}\sum_{i=1}^{5}a_i^2b_i^2\\ \text{subject to} & a_i+b_i=0.201144939+0.00102055255 \text{ for } i\in\{1,2,3,4,5\},\\ & b_2+b_5+a_3+a_4\ge 0.0919109388238,\\ & b_1+b_3+a_4+a_5\ge 0.0919109388238,\\ & b_2+b_4+a_1+a_5\ge 0.0919109388238,\\ & b_3+b_5+a_1+a_2\ge 0.0919109388238,\\ & b_4+b_1+a_2+a_3\ge 0.0919109388238,\\ & a_i,b_i\ge 0 \text{ for } i\in\{1,2,3,4,5\}.\end{cases}$$

Note that the resulting program (P') has only 5 degrees of freedom.

We find an upper bound of the solution of (P') by a brute force method. We discretize the space of possible solutions, and bound the gradient of the target function to control the behavior between the grid points. For this, we fix a constant s which will correspond to the number of steps. For every a_i we check $s+1$ equally spaced values between 0 and $0.201144939+0.00102055255$ that include the boundaries. By this we have a grid of s^5 boxes where every feasible solution of (P'), and hence also of (P), is in one of the boxes.

Next we need to find the partial derivatives of f. Since f is symmetric, we only check the partial derivative with respect to a_1.

$$\frac{\partial f}{\partial a_1}=b_2b_3a_4+a_3b_4b_5+\frac{1}{2}a_1b_1^2$$

We want to find an upper bound on $\frac{\partial f}{\partial a_1}$. We can pick 0.21 as an upper bound on a_i+b_i. Hence we assume $a_1+b_1=a_3+b_3=a_4+b_4=b_2=b_5=0.21$ and we maximize

$$b_3a_4+a_3b_4=(0.21-a_3)a_4+a_3(0.21-a_4)=0.21a_4+0.21a_3-2a_3a_4.$$

This is maximized if $a_3 = 0, a_4 = 0.21$ or $a_3 = 0.21, a_4 = 0$ and gives the value 0.21^2. Hence

$$a_1 b_1^2 = 4a_1 \cdot \frac{b_1}{2} \cdot \frac{b_1}{2} \leq \frac{4(a_1 + b_1)^3}{3^3} = \frac{4 \cdot 0.21^3}{27}.$$

The resulting upper bound is

$$\frac{\partial f}{\partial a_1} \leq 0.21^2 + \frac{2 \cdot 0.21^3}{27} \leq 0.045.$$

Hence in a box with side length t the value of f cannot be bigger than the value at a corner plus $5t/2 \cdot 0.045$. The factor $5t/2$ comes from the fact that the closest corner is in distance at most $t/2$ in each of the 5 coordinates.

If we set $s = 100$, we compute that the maximum over all grid points is less than 0.00133. This can be checked by a computer program `mesh-opt.cpp`. With $t < 0.0021$, we have $5t/2 \cdot 0.045 < 0.00024$. So we conclude that x is in less than $0.00157n^4$ induced C_5s which contradicts Claim 7. □

We have just established the "outside" structure of G. This implies that

$$C(n) = (x_1 \cdot x_2 \cdot x_3 \cdot x_4 \cdot x_5)n^5 + C(x_1 n) + C(x_2 n) + C(x_3 n) + C(x_4 n) + C(x_5 n).$$

Averaging over all subgraph of order $n - 1$, we can easily see that $C(n) \leq C(n-1)$ for all n, so

$$\ell := \lim_{n\to\infty} \frac{C(n)}{\binom{n}{5}}$$

exists. Therefore,

$$\ell + o(1) = 5!\ x_1 \cdot x_2 \cdot x_3 \cdot x_4 \cdot x_5 + \ell(x_1^5 + x_2^5 + x_3^5 + x_4^5 + x_5^5),$$

which implies that $x_i = \frac{1}{5} + o(1)$, and $\ell = \frac{1}{26}$, given the constraints on the x_i.

In order to prove Theorem 2, it remains to show that in fact $|X_i| - |X_j| \leq 1$ for all $i, j \in \{1, \dots, 5\}$.

Claim 9. *For n large enough, we have $|X_i| - |X_j| \leq 1$ for all $i, j \in \{1, \dots, 5\}$.*

Proof. By symmetry, assume for contradiction that $|X_1| - |X_2| \geq 2$. Let v be from X_1 where C_5^v is minimized over the vertices in X_1 and let w be from X_2 where C_5^w is maximized over the vertices in X_2. As G is extremal, $C_5^v + C_5^{vw} - C_5^w \geq 0$; otherwise, we can increase the number of C_5s by replacing v by a copy of w.

Let $y_i := |X_i| = x_i n$. By the monotonicity of $\frac{C(n)}{n^5}$, we have

$$\frac{1}{26} + o(1) \geq \frac{C(y_2)}{\binom{y_2}{5}} \geq \frac{C(y_1)}{\binom{y_1}{5}} \geq \frac{1}{26}.$$

Therefore, using that $y_1 - y_2 \geq 2$,

$$\begin{aligned}
C_5^v + C_5^{vw} - C_5^w &\leq \frac{C(y_1)}{y_1} + y_2y_3y_4y_5 + y_3y_4y_5 - \frac{C(y_2)}{y_2} - y_1y_3y_4y_5 \\
&\leq \frac{y_2C(y_1) - y_1C(y_2)}{y_1y_2} + (y_2 - y_1 + 1)y_3y_4y_5 \\
&\leq \left(\frac{1}{26} + o(1)\right)\frac{1}{y_1y_2}\left(y_2\binom{y_1}{5} - y_1\binom{y_2}{5}\right) + (y_2 - y_1 + 1)y_3y_4y_5 \\
&\leq \left(\frac{1}{26\cdot 5!} + o(1)\right)\left(y_1^4 - y_2^4\right) + (y_2 - y_1 + 1)y_3y_4y_5 \\
&= \left(\frac{1}{26\cdot 5!} + o(1)\right)(y_1 - y_2)\left(y_1^3 + y_1^2y_2 + y_1y_2^2 + y_2^3\right) + (y_2 - y_1 + 1)y_3y_4y_5 \\
&= (y_1 - y_2)\left(\left(\frac{1}{26\cdot 5!} + o(1)\right)\frac{4n^3}{125} - \frac{n^3}{125}\right) + \frac{(1+o(1))n^3}{125} \\
&\leq \left(\frac{2}{26\cdot 5!} + o(1)\right)\frac{4n^3}{125} - \frac{(1+o(1))n^3}{125} < 0,
\end{aligned}$$

a contradiction. □

With this claim, the proof of Theorem 2 is complete.

4 Proof of Theorem 1

Theorem 1 is a consequence of Theorem 2. The main proof idea is to take a minimal counterexample and show that some blow-up of this graph contradicts Theorem 2.

Proof of Theorem 1. Theorem 1 is easily seen to be true for $k = 1$, so suppose for a contradiction that there is a graph G on $n = 5^k$ vertices with $C(G) \geq C(C_5^{k\times})$ that is not isomorphic to $C_5^{k\times}$ for a minimal $k \geq 2$.

If G has the structure described in Theorem 2, then G is isomorphic to $C_5^{k\times}$ by the minimality of k, a contradiction. Therefore, $V(G)$ cannot be partitioned into five sets X_1, X_2, X_3, X_4, X_5 with $|X_i| = 5^{k-1}$ as described in Theorem 2.

Let H be an extremal graph on $5^\ell > n_0$ vertices, where n_0 is taken from the statement of Theorem 2. Blowing up every vertex of $C_5^{k\times}$ by a factor of 5^ℓ, and inserting H in every part, gives an extremal graph G_1 on $5^{k+\ell}$ vertices by ℓ applications of Theorem 2. On the other hand, the graph G_2 obtained by blowing up every vertex of G by a factor of 5^ℓ, and inserting H in every part, contains at least as many C_5s as G_1,

$$C(G_1) = 5^k \cdot C(H) + C(C_5^{k\times}) \cdot (5^\ell)^5, \qquad C(G_2) = 5^k \cdot C(H) + C(G) \cdot (5^\ell)^5,$$

so $C(G_1) \leq C(G_2)$. Hence G_2 must also be extremal. By Theorem 2, $V(G_2)$ can be partitioned into five sets X_1, X_2, X_3, X_4, X_5 with $|X_i| = 5^{k+\ell-1}$ in the described way. In particular, two vertices in G_2 are in the same set X_i if and only if their adjacency pattern agrees on more than half of the remaining vertices. But this implies that for every copy of H inserted into the blow up of G, all vertices are in the same X_i, and thus a partition of $V(G)$ as described in Theorem 2 is induced by this partition of $V(G_2)$, a contradiction. □

Acknowledgement

We would like to thank Jan Volec for fruitful discussions. Thank you for reading all the way to the end. We hope you liked it. Please send us some candy, beer or wine.

References

[1] R. Baber. Turán densities of hypercubes. Submitted.

[2] R. Baber and J. Talbot. A solution to the 2/3 conjecture. Submitted.

[3] R. Baber and J. Talbot. Hypergraphs do jump. *Combin. Probab. Comput.*, 20(2):161–171, 2011.

[4] J. Balogh, P. Hu, B. Lidický, and H. Liu. Upper bounds on the size of 4- and 6-cycle-free subgraphs of the hypercube. *European Journal of Combinatorics*, 35:75–85, 2014.

[5] J. Balogh, P. Hu, B. Lidický, F. Pfender, J. Volec, and M. Young. Rainbow triangles in three-colored graphs. Submitted.

[6] J. Balogh, P. Hu, B. Lidický, O. Pikhurko, B. Udvari, and J. Volec. Minimum number of monotone subsequences of length 4 in permutations. *To appear in Combinatorics, Probability and Computing.*

[7] J. Cummings, D. Kráľ, F. Pfender, K. Sperfeld, A. Treglown, and M. Young. Monochromatic triangles in three-coloured graphs. *J. Combin. Theory Ser. B*, 103(4):489–503, 2013.

[8] S. Das, H. Huang, J. Ma, H. Naves, and B. Sudakov. A problem of Erdős on the minimum number of k-cliques. *J. Combin. Theory Ser. B*, 103(3):344–373, 2013.

[9] C. Even-Zohar and N. Linial. A note on the inducibility of 4-vertex graphs, 2013.

[10] V. Falgas-Ravry and E. R. Vaughan. Turán h-densities for 3-graphs. *Electron. J. Combin.*, 19:R40, 2012.

[11] R. Glebov, D. Kráľ, and J. Volec. A problem of Erdős and Sós on 3-graphs. Submitted.

[12] A. Grzesik. On the maximum number of five-cycles in a triangle-free graph. *J. Combin. Theory Ser. B*, 102(5):1061–1066, 2012.

[13] H. Hatami, J. Hirst, and S. Norine. The inducibility of blow-up graphs. *Journal of Combinatorial Theory, Series B*, pages –, 2014.

[14] H. Hatami, J. Hladký, D. Kráľ, S. Norine, and A. Razborov. On the number of pentagons in triangle-free graphs. *J. Combin. Theory Ser. A*, 120(3):722–732, 2013.

[15] J. Hirst. The inducibility of graphs on four vertices. *J. Graph Theory*, 75(3):231–243, 2014.

[16] J. Hladký, D. Kráľ, and S. Norin. Private communication.

[17] J. Hladký, D. Kráľ, and S. Norine. Counting flags in triangle-free digraphs. In *European Conference on Combinatorics, Graph Theory and Applications (EuroComb 2009)*, volume 34 of *Electron. Notes Discrete Math.*, pages 621–625. Elsevier Sci. B. V., Amsterdam, 2009.

[18] H. Huang. On the maximum induced density of directed stars and related problems. *SIAM J. Discrete Math.*, 28(1):92–98, 2014.

[19] D. Kráľ, L. Mach, and J.-S. Sereni. A new lower bound based on Gromov's method of selecting heavily covered points. *Discrete Comput. Geom.*, 48(2):487–498, 2012.

[20] V. Nikiforov. The number of cliques in graphs of given order and size. *Trans. Amer. Math. Soc.*, 363(3):1599–1618, 2011.

[21] O. Pikhurko. On Possible Turán Densities, 2014. to appear in Israel J Math.

[22] O. Pikhurko and E. R. Vaughan. Minimum number of k-cliques in graphs with bounded independence number. *To appear in Combinatorics, Probability and Computing.*

[23] N. Pippenger and M. C. Golumbic. The inducibility of graphs. *J. Combinatorial Theory Ser. B*, 19(3):189–203, 1975.

[24] A. Razborov. On the Caccetta-Häggkvist conjecture with forbidden subgraphs. Submitted.

[25] A. A. Razborov. Flag algebras. *Journal of Symbolic Logic*, 72(4):1239–1282, 2007.

[26] A. A. Razborov. On the minimal density of triangles in graphs. *Combin. Probab. Comput.*, 17(4):603–618, 2008.

[27] A. A. Razborov. Flag Algebras: An Interim Report. In *The Mathematics of Paul Erdős II*, pages 207–232. Springer, 2013.

[28] A. A. Razborov. Flag algebras: an interim report, 2014. Talk at IMA Annual Program `http://www.ima.umn.edu/videos/index.php?id=2756`.

[29] C. Reiher. The clique density theorem, 2012. Submitted.

erfeld:2011 [30] K. Sperfeld. The inducibility of small oriented graphs. Submitted.

sage [31] W. Stein et al. *Sage Mathematics Software (Version 6.3)*. The Sage Development Team, 2014. http://www.sagemath.org.

flagmatic [32] E. Vaughan. Flagmatic 2.0, 2012. http://flagmatic.org.