

Edge Version of the Inducibility via the Entropy Method

Yichen Wang^a Xiamiao Zhao^a Mei Lu^a

Submitted: Jul 27, 2025; Accepted: Mar 17, 2026; Published: Apr 24, 2026

© The authors. Released under the CC BY-ND license (International 4.0).

Abstract

The inducibility of a graph H is about the maximum number of induced copies of H in a graph on n vertices. We consider its edge version, that is, the maximum number of induced copies of H in a graph with m edges. Let $c(G, H)$ be the number of induced copies of H in G and $\rho(H, m) = \max\{c(G, H) \mid |E(G)| = m\}$. For any graph H , we prove that $\rho(H, m) = \Theta(m^{\alpha_f(H)})$ where $\alpha_f(H)$ is the fractional independence number of H . Therefore, we now focus on the constant factor in front of $m^{\alpha_f(H)}$. In this paper, we give some results of $\rho(H, m)$ when H is a cycle or path. We conjecture that for any cycle C_k with $k \geq 5$, $\rho(C_k, m) = (1 + o(1)) (m/k)^{k/2}$ and the bound is achieved by the blow up of C_k . For even cycles, we establish an upper bound with an extra constant factor. For odd cycles, we can only establish an upper bound with an extra factor depending on k . We prove that $\rho(P_{2l}, m) \leq \frac{m^l}{2^{(l-1)^{l-1}}}$ and $\rho(P_{2l+1}, m) \leq \frac{m^{l+1}}{4^l}$, where $l \geq 2$. We also conjecture the asymptotic value of $\rho(P_k, m)$. The entropy method is mainly used to prove our results.

Mathematics Subject Classifications: 05C35, 05C38

1 Introduction

Let G and H be two graphs. The **induced density** of H in G , denoted by $i(G, H)$, is the number of induced copies of H in G divided by $\binom{|V(G)|}{|V(H)|}$. Let $i(H, n)$ denote the maximum induced density of H in a graph with n vertices, that is, $i(H, n) = \max\{i(G, H) \mid |V(G)| = n\}$. The **inducibility** of H , denoted by $ind(H)$, is the limit of $i(H, n)$ when n goes to infinity, that is $ind(H) = \lim_{n \rightarrow +\infty} i(H, n)$. This classic topic in extremal combinatorics is introduced by Pippenger and Golumbic [14] in 1975. In the same paper [14], they showed that the inducibility of every k -vertex graph H is at least $k!/(k^k - k)$ and conjectured that this bound is tight for a cycle C_k with $k \geq 5$.

^aDepartment of Mathematical Sciences, Tsinghua University, Beijing 100084, China
(wangyich22@mails.tsinghua.edu.cn); Corresponding author: X. Zhao
(zxm23@mails.tsinghua.edu.cn); (lumei@tsinghua.edu.cn).

Conjecture 1 (Pippenger and Golumbic [14]). The inducibility of a cycle C_k of length $k \geq 5$ is equal to $\frac{k!}{k^k - k}$.

The conjecture for $k = 5$ is proved by Balogh, Hu, Lidický and Pfender [3]. For larger k , Pippenger and Golumbic [14] showed $\text{ind}(C_k) \leq (2e + o(1)) \frac{k!}{k^k}$. The extra factor $2e$ has been improved to $128e/81$ by Hefetz and Tyomkyn [9] in 2018. In 2019, Král', Norin and Volec [12] proved that every n -vertex graph has at most $2n^k/k^k$ induced cycles of length k , that is, $\text{ind}(C_k) \leq 2 \frac{k!}{k^k}$.

Now we fix the number of edges of graphs rather than the number of vertices. Let $c(G, H)$ be the number of induced copies of H in G and $\rho(H, m)$ the maximum number of induced copies of H in a graph with m edges. That is, $\rho(H, m) = \max\{c(G, H) \mid |E(G)| = m\}$. If $H = K_t$, then the problem becomes the famous Kruskal-Katona Theorem [10, 11] which is of central significance in extremal set theory.

Let H be a graph. Its independence number can be defined as the solution of an integer program. Let $\omega : V(H) \rightarrow \{0, 1\}$. We wish to maximize $\sum_{v \in V(H)} \omega(v)$ subject to the constraints that $\omega(u) + \omega(v) \leq 1$ for all edges $uv \in E(H)$. If this integer program is relaxed such that the weights are in the interval $[0, 1]$, then the solution is the **fractional independence number** of H , denoted by $\alpha_f(H)$. That is, maximizing $\sum_{v \in V(H)} \omega(v)$ subject to the constraints that $\omega(u) + \omega(v) \leq 1$ for all edges $uv \in E(H)$ and $0 \leq \omega(v) \leq 1$ for all $v \in V(H)$. Furthermore, a maximum weighting of the fractional independence number can always be obtained when restricting the weights to be in the set $\{0, 1/2, 1\}$ [13, 15]. It means that equivalently we can assume $\omega(v) \in \{0, 1/2, 1\}$ for all $v \in V(H)$ when $\sum_{v \in V(H)} \omega(v)$ achieves maximum. In the non-induced version of the problem, the fractional independence number plays a key role. Alon [1], Friedgut and Kahn [7], and Galvin [8] have studied the maximum number of (not necessarily induced) copies of G in a graph with m edges.

Theorem 2 (Alon [1], Friedgut and Kahn [7], Galvin [8]). *Let H be a graph without isolated vertices. Denote the maximum number of copies of H in a graph with m edges by $\pi(H, m)$. Then*

$$(1 - o(1)) \left(\frac{m}{|E(H)|} \right)^{\alpha_f(H)} \leq \pi(H, m) \leq \frac{(2m)^{\alpha_f(H)}}{|Aut(H)|},$$

where $|E(H)|$ is the number of edges in H , $Aut(H)$ is the automorphism group of H and $|Aut(H)|$ denotes its order.

Note that if H contains isolated vertices, the problem is meaningless since we can add infinitely many isolated vertices to G such that G contains infinitely many induced copies of H .

For the induced version ($\rho(H, m)$), the upper bound in Theorem 2 also works. For the lower bound of $\rho(H, m)$, we can construct a graph by blowing up each vertex of H to an independent set of size $\left(\frac{m}{|E(H)|} \right)^{\omega(v)}$, where $\omega(v)$ is the weight of v in a maximum weighting of the fractional independence number. By calculating the number of induced copies of

H in this blow up graph, we have the following theorem. It is worth mentioning that the lower bound construction is also the lower bound for the non-induced version $(\pi(H, m))$ in Theorem 2.

Theorem 3. *For any graph H without isolated vertices, we have that*

$$(1 - o(1)) \left(\frac{m}{|E(H)|} \right)^{\alpha_f(H)} \leq \rho(H, m) \leq \frac{(2m)^{\alpha_f(H)}}{|Aut(H)|}.$$

Theorem 3 describes the asymptotic order of $\rho(H, m)$. When $H = K_t$, $\rho(K_t, m)$ is maximized in an (almost) complete graph G . At that time, $\binom{|V(G)|}{2} \approx m$, that is, $|V(G)| \approx \sqrt{2m}$, and $c(G, K_t) \approx \binom{|V(G)|}{t} \approx \binom{\sqrt{2m}}{t} \approx 2^{t/2} m^{t/2} / t!$. Note that $\alpha_f(K_t) = t/2$ and $|Aut(K_t)| = t!$. So the upper bound in Theorem 3 is tight for K_t .

Now we focus on the constant factor in front of $m^{\alpha_f(H)}$. We use C_k to denote the cycle on k vertices and P_k the path on k vertices. Then $\alpha_f(C_k) = k/2$ and $\alpha_f(P_k) = \lfloor (k+1)/2 \rfloor$. We have the following theorems about the number of induced copies of cycle and path.

Theorem 4. *We have $(1 + o(1))(2l + 1) \frac{m^l}{(2l+1)^l} \leq \rho(P_{2l}, m) \leq \frac{m^l}{2^{(l-1)^{l-1}}}$, where $l \geq 2$.*

Theorem 5. *We have $(1 + o(1))4 \frac{m^{l+1}}{(2l+2)^{l+1}} \leq \rho(P_{2l+1}, m) \leq \frac{m^{l+1}}{4^l}$, where $l \geq 2$.*

Note that when $l = 1$, $\rho(P_3, m) \leq \binom{m}{2}$ and the bound is achieved by a star.

Theorem 6. *We have the following results for even cycles.*

1. *We have $\rho(C_4, m) = (1 + o(1)) \frac{m^2}{4}$.*
2. *We have $(1 + o(1)) \left(\frac{m}{6}\right)^3 \leq \rho(C_6, m) \leq 4 \left(\frac{m}{6}\right)^3$.*
3. *When $k = 2l \geq 8$, we have*

$$(1 + o(1)) \left(\frac{m}{k}\right)^{k/2} \leq \rho(C_k, m) \leq \left(\frac{m}{2l}\right)^l \left(1 + \frac{1}{l-1}\right)^{l-1} \leq e \left(\frac{m}{k}\right)^{k/2}.$$

Theorem 7. *Let $l \geq 2$. Then, we have*

$$(1 + o(1)) \left(\frac{m}{2l+1}\right)^{(2l+1)/2} \leq \rho(C_{2l+1}, m) \leq \frac{(2l+1)^{l-1/2}}{2^{(l-1)(2l+1)/(2l)}} \left(\frac{m}{2l+1}\right)^{(2l+1)/2}.$$

The lower bound in Theorem 6 and Theorem 7 comes from Theorem 3, except for C_4 . The lower bound of $\rho(C_4, m)$ is achieved by the complete balanced bipartite graph. It is easy to verify that $\alpha_f(C_k) = k/2$. For the odd cycle C_{2l+1} , the best weighting is letting $\omega(v) = 1/2$ for every vertex v . For the even cycle C_{2l} , the best weighting is letting $\omega(v) = x \in [0, 1]$ for an independent set of size l and $\omega(u) = 1 - x$ for the other vertices (see Figure 1 for the corresponding construction when $k = 6$).

We conjecture that the above lower bounds are tight.

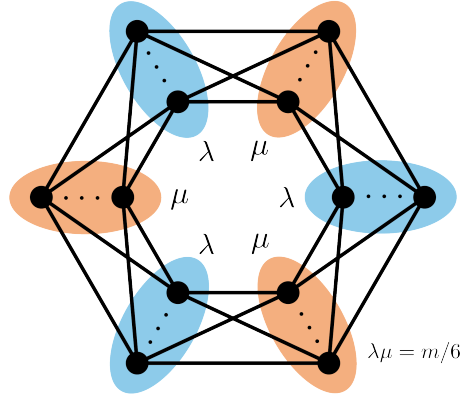


Figure 1: Unbalanced blow up when $k = 6$. The orange (resp. blue) parts have μ (resp. λ) vertices and $\lambda\mu \approx m/6$.

Conjecture 8. For any $k \geq 5$, $\rho(C_k, m) = (1 + o(1))\left(\frac{m}{k}\right)^{k/2}$.

Now we give the lower bounds of $\rho(P_k, m)$ in Theorems 4 and 5. Note that $\rho(P_k, m) \geq (1 + o(1))(k + 1)\left(\frac{m}{k+1}\right)^{\frac{k}{2}}$ where k is an even number and the bound is achieved by the balanced blow up of C_{k+1} . When k is an odd number, we have that $\rho(P_k, m) \geq (1 + o(1))4\left(\frac{m}{k+1}\right)^{(k+1)/2}$ by the construction of the unbalanced blow up of P_k (see Figure 2). Let $P_k[V_1, \dots, V_k]$ be a blow up of P_k and denote $n_i = |V_i|$, $1 \leq i \leq k$. Let $n_i = 1$ for even i . Then $m = n_1 + 2n_3 + 2n_5 + \dots + 2n_{k-2} + n_k$ and the number of induced copies of P_k is $n_1 n_3 \dots n_k \leq 4(m/(k+1))^{(k+1)/2}$ and the equality holds when $n_1 = 2n_3 = 2n_5 = \dots = 2n_{k-2} = n_k$. So we conjecture those bounds to be tight.

Conjecture 9. Let $k \geq 4$ be an even integer, then $\rho(P_k, m) = (1 + o(1))(k + 1)\left(\frac{m}{k+1}\right)^{\frac{k}{2}}$.

Conjecture 10. Let $k \geq 5$ be an odd integer, then $\rho(P_k, m) = (1 + o(1))4\left(\frac{m}{k+1}\right)^{(k+1)/2}$.

Comparing Theorem 4 with Conjecture 9, our theorem has an extra factor $\frac{(2l+1)^{l-1}}{2^{(l-1)^{l-1}}} \approx 2^{l-2}e^{3/2}$ when l goes to infinity. Similarly, comparing Theorem 7 to Conjecture 8, there is an extra factor $\frac{(2l+1)^{l-1/2}}{2^{(l-1)^{(l-1)/(2l)}}} \approx 2^{l-3/2}e^{3/2}$ when l goes to infinity. That is because Theorem 7 is actually a corollary from Theorem 4 using Shearer's Lemma (Lemma 16) which we will mention in Section 2. The extra error factor when estimating induced odd cycle is due to our overestimation about induced odd path.

We use the entropy method to prove our main results. The entropy method is a powerful tool in such problems, for example, counting rainbow triangles when fixing the number of edges in each color class [6]; counting t -cliques when fixing the p -norm of the degree sequence [5].

This paper is organized as follows. In Section 2, we introduce basic properties of the entropy method. In Section 3, we will prove Theorem 5. In Section 4, we will prove Theorems 4 and 7. In Section 5, we will prove Theorem 6.

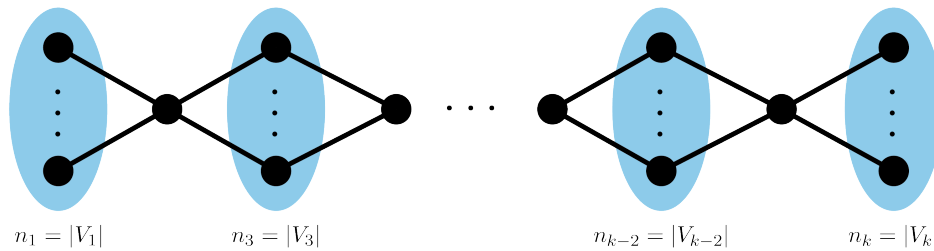


Figure 2: The blow up of P_k when k is odd.

2 Preliminaries

In this section, we introduce basic properties of the entropy method. For a complete discussion and proofs for properties of the entropy method, readers may refer to Section 15.7 in [2].

Definition 11. For any discrete random variable X with finite support S , the **entropy** of X is defined as

$$H(X) = - \sum_{x \in S} p_X(x) \log p_X(x),$$

where $p_X(x)$ denotes $\mathbb{P}(X = x)$.

The entropy $H(X_1, \dots, X_n)$ of several discrete random variables X_1, \dots, X_n is defined as the entropy of the random variable $X = (X_1, \dots, X_n)$.

Definition 12 (Conditional Entropy). Let X, Y be two random variables with finite supports S, T , respectively. Denote by $p_{X,Y}(x, y)$ the probability $\mathbb{P}(X = x, Y = y)$. For each $y \in T$, let $X|Y = y$ be the random variable X conditioned on $Y = y$, and $H(X | Y = y)$ be its entropy. The conditional entropy of X given Y is defined as

$$H(X | Y) = \sum_{y \in T} p_Y(y) H(X | Y = y) = - \sum_{x \in S, y \in T} p_{X,Y}(x, y) \log \left(\frac{p_{X,Y}(x, y)}{p_Y(y)} \right).$$

Proposition 13. For any random variable X with finite support S , we have $H(X) \leq \log |S|$. The equality holds when X is uniformly distributed over S .

Proposition 14 (Chain Rule). For any two random variables X and Y with finite supports, we have

$$H(X, Y) = H(X) + H(Y | X) = H(Y) + H(X | Y).$$

More generally, for any n random variables X_1, \dots, X_n with finite supports, we have

$$H(X_1, \dots, X_n) = H(X_1) + H(X_2 | X_1) + \dots + H(X_n | X_1, \dots, X_{n-1}).$$

Proposition 15 (Drop Condition). *For any two random variables X and Y with finite supports, we have*

$$H(X | Y) \leq H(X).$$

Lemma 16 (Shearer's Lemma). *Let X_1, \dots, X_n be random variables with finite supports. Let \mathcal{A} be a collection of subsets of $[n]$ such that each $i \in [n]$ is contained in at least r sets in \mathcal{A} . Then*

$$H(X_1, X_2, \dots, X_n) \leq \frac{1}{r} \sum_{A \in \mathcal{A}} H((X_i)_{i \in A}).$$

We say an ordered edge tuple (e_1, e_2, \dots, e_k) is an ordered induced copy of C_k if e_i intersects e_{i-1} and e_{i+1} on two different vertices for $i = 1, \dots, k$ (let $e_0 = e_k, e_{k+1} = e_1$) and there is no chord inside. Similarly, we say $(e_1, e_2, \dots, e_{k-1})$ is an ordered induced copy of P_k if e_i intersects e_{i-1} and e_{i+1} on two different vertices for $i = 2, \dots, k-2$ and there are no other edges connecting e_i and e_j for any $1 \leq i \neq j \leq k-1$.

Given a graph G , we say a sequence of edges (e_1, e_2, \dots, e_t) is **well-ordered** if $v_i u_{i+1} \in E(G)$ for all $1 \leq i \leq t-1$ but there is no other edge among $\{u_i u_j, u_i v_j, v_i u_j, v_i v_j\}_{i \neq j}$, where $e_i = u_i v_i, 1 \leq i \leq t$. For example, in an ordered induced copy of $C_k = (e_1, \dots, e_k)$, $(e_1, e_3, \dots, e_{2t-1})$ ($2t-1 < k-1$) is well-ordered. In the following, when there is no ambiguity, we usually ignore specifying the graph G .

We say a well-ordered edge tuple $(e_1, e_3, \dots, e_{2t-1})$ is **embedded** in an ordered induced copy of $P_k = (e'_1, e'_2, \dots, e'_{k-1})$ if $e_{2i-1} = e'_{2i-1}$ for all $1 \leq i \leq t$. For convenience, we write $e_{<2t+1}^o = e_1, e_3, \dots, e_{2t-1}$. Let $e_{<2t+1}^o = e_1, e_3, \dots, e_{2t-1}$ and $\tilde{e}_{<2t+1}^o = \tilde{e}_1, \tilde{e}_3, \dots, \tilde{e}_{2t-1}$. We say $e_{<2t+1}^o = \tilde{e}_{<2t+1}^o$ if $e_i = \tilde{e}_i$ for $i = 1, 3, \dots, 2t-1$.

3 Proof of the upper bound of Theorem 5

Let $k = 2l + 1 \geq 5$ be an odd integer and G the extremal graph maximizing the number of induced copies of P_k . Denote the number of induced copies of P_k in G by Υ . Let $P = (e_1, e_2, \dots, e_{k-1})$ be a random variable uniformly chosen from all ordered induced copies of P_k in G . Hence, we have $H(P) = \log(2\Upsilon)$ by Proposition 13. For a well-ordered edge tuple $(e_1, e_3, \dots, e_{2t-1})$ ($1 \leq t \leq l-2$), let $\beta(e_1, e_3, \dots, e_{2t-1})$ be the number of ordered induced copies of $P_k = (e'_1, e'_2, \dots, e'_k)$ such that $e_{2i-1} = e'_{2i-1}$ for all $1 \leq i \leq t$ and $\alpha(e_1, e_3, \dots, e_{2t-1})$ be the number of edges e such that $(e_1, e_3, \dots, e_{2t-1}, e)$ is well-ordered. For a well-ordered edge tuple $(e_1, e_3, \dots, e_{2l-3})$, let $\mathcal{S}(e_1, e_3, \dots, e_{2l-3})$ be an edge set containing e such that there exists an ordered induced path $P_k = (e'_1, e'_2, \dots, e'_{2l})$ satisfying $e_{2i-1} = e'_{2i-1}$ for all $1 \leq i \leq l-1$ and $e'_{2l} = e$. Write $\gamma_0(e_1, e_3, \dots, e_{2l-3}) = |\mathcal{S}(e_1, e_3, \dots, e_{2l-3})|$. For a well-ordered edge tuple $(e_1, e_3, \dots, e_{2l-3})$ and an edge $e_{2l} \in \mathcal{S}(e_1, e_3, \dots, e_{2l-3})$, let $\gamma_1(e_1, e_3, \dots, e_{2l-3}, e_{2l})$ (resp. $\gamma_2(e_1, e_3, \dots, e_{2l-3}, e_{2l})$) be the number of e such that there exists an ordered induced path $P_k = (e'_1, e'_2, \dots, e'_{2l})$ satisfying $e_{2i-1} = e'_{2i-1}$ for all $1 \leq i \leq l-1$, $e'_{2l} = e_{2l}$ and $e'_{2l-2} = e$ (resp. $e'_{2l-1} = e$).

Note that $H(e_2, e_4, \dots, e_{2l-4} | e_{<2l-1}^o) = 0$ by Proposition 13 since $e_2, e_4, \dots, e_{2l-4}$ are uniquely determined by $e_{<2l-1}^o$. Moreover, we have $H(e_{2l-2}, e_{2l-1} | e_{<2l-1}^o, e_{2l}) = H(e_{2l-2} |$

$e_{<2l-1}^o, e_{2l}) = H(e_{2l-1} | e_{<2l-1}^o, e_{2l})$ since e_{2l-2} (resp. e_{2l-1}) is uniquely determined by $e_{<2l-1}^o, e_{2l}$ and e_{2l-1} (resp. e_{2l-2}). Combining Propositions 14 and 13, we have

$$\begin{aligned} H(P) &= \log(2\Upsilon) \\ &\leq H(e_1) + \sum_{i=1}^{l-2} H(e_{2i+1} | e_{<2i+1}^o) + H(e_{2l} | e_{<2l-1}^o) \\ &\quad + \frac{1}{2} (H(e_{2l-2} | e_{<2l-1}^o, e_{2l}) + H(e_{2l-1} | e_{<2l-1}^o, e_{2l})). \end{aligned} \tag{1}$$

Now we analyze each term in (1). Firstly, we have

$$\begin{aligned} H(e_{2i+1} | e_{<2i+1}^o) &\leq \sum_{\tilde{e}_{<2i+1}^o} \mathbb{P}(e_{<2i+1}^o = \tilde{e}_{<2i+1}^o) \log \alpha(\tilde{e}_{<2i+1}^o) \\ &= \sum_{\tilde{e}_{<2i+1}^o} \frac{\beta(\tilde{e}_{<2i+1}^o)}{2\Upsilon} \log \alpha(\tilde{e}_{<2i+1}^o) \\ &= \sum_{\tilde{e}_{<2i+1}^o} \sum_{\text{Ordered induced } P_k: \tilde{P}} \frac{1_{\tilde{e}_{<2i+1}^o \text{ is embedded in } \tilde{P}}}{2\Upsilon} \log \alpha(\tilde{e}_{<2i+1}^o) \\ &= \sum_{\text{Ordered induced } P_k: (\tilde{e}_1, \dots, \tilde{e}_{k-1})} \frac{1}{2\Upsilon} \log \alpha(\tilde{e}_{<2i+1}^o). \end{aligned} \tag{2}$$

Similarly, we have

$$\begin{aligned} H(e_{2l} | e_{<2l-1}^o) &\leq \sum_{\text{Ordered induced } P_k: (\tilde{e}_1, \dots, \tilde{e}_{k-1})} \frac{1}{2\Upsilon} \log \gamma_0(\tilde{e}_{<2l-1}^o), \\ H(e_{2l-2} | e_{<2l-1}^o, e_{2l}) &\leq \sum_{\text{Ordered induced } P_k: (\tilde{e}_1, \dots, \tilde{e}_{k-1})} \frac{1}{2\Upsilon} \log \gamma_1(\tilde{e}_{<2l-1}^o, \tilde{e}_{2l}), \\ H(e_{2l-1} | e_{<2l-1}^o, e_{2l}) &\leq \sum_{\text{Ordered induced } P_k: (\tilde{e}_1, \dots, \tilde{e}_{k-1})} \frac{1}{2\Upsilon} \log \gamma_2(\tilde{e}_{<2l-1}^o, \tilde{e}_{2l}). \end{aligned}$$

When fixing the ordered induced path $P_k = (\tilde{e}_1, \dots, \tilde{e}_{k-1})$, $\{\alpha(\tilde{e}_{<2i+1}^o)\}_{i=1,2,\dots,l-2}$, $\gamma_0(\tilde{e}_{<2l-1}^o)$, $\gamma_1(\tilde{e}_{<2l-1}^o, \tilde{e}_{2l})$ and $\gamma_2(\tilde{e}_{<2l-1}^o, \tilde{e}_{2l})$ are all cardinality of some edge sets. By the definitions, those edge sets are disjoint, since the paths are induced. Hence, we have

$$\sum_{i=1}^{l-2} \alpha(\tilde{e}_{<2i+1}^o) + \gamma_0(\tilde{e}_{<2l-1}^o) + \gamma_1(\tilde{e}_{<2l-1}^o, \tilde{e}_{2l}) + \gamma_2(\tilde{e}_{<2l-1}^o, \tilde{e}_{2l}) \leq m.$$

By AM-GM inequality, we have that

$$\begin{aligned}
& \prod_{i=1}^{l-2} \alpha(\tilde{e}_{<2i+1}^o)^2 \cdot \gamma_0(\tilde{e}_{<2l-1}^o)^2 \cdot \gamma_1(\tilde{e}_{<2l-1}^o, \tilde{e}_{2l}) \cdot \gamma_2(\tilde{e}_{<2l-1}^o, \tilde{e}_{2l}) \\
&= \frac{1}{4} \prod_{i=1}^{l-2} \alpha(\tilde{e}_{<2i+1}^o)^2 \cdot \gamma_0(\tilde{e}_{<2l-1}^o)^2 \cdot (2\gamma_1(\tilde{e}_{<2l-1}^o, \tilde{e}_{2l})) \cdot (2\gamma_2(\tilde{e}_{<2l-1}^o, \tilde{e}_{2l})) \\
&\leq \frac{1}{4} \left(\frac{2 \left(\sum_{i=1}^{l-2} \alpha(\tilde{e}_{<2i+1}^o) + \gamma_0(\tilde{e}_{<2l-1}^o) + \gamma_1(\tilde{e}_{<2l-1}^o, \tilde{e}_{2l}) + \gamma_2(\tilde{e}_{<2l-1}^o, \tilde{e}_{2l}) \right)}{2l} \right)^{2l} \\
&\leq \frac{1}{4} \left(\frac{2m}{2l} \right)^{2l} = \frac{1}{4} \left(\frac{m}{l} \right)^{2l}.
\end{aligned}$$

Substituting the above inequalities into (1), we have

$$\begin{aligned}
H(P) &= \log(2\Upsilon) \\
&\leq \log m + \sum_{\text{Ordered induced } P_k: (\tilde{e}_1, \dots, \tilde{e}_{k-1})} \frac{1}{4\Upsilon} \log \left(\frac{1}{4} \left(\frac{m}{l} \right)^{2l} \right) \\
&\leq \log \frac{m^{l+1}}{2l^l},
\end{aligned}$$

which implies Theorem 5 holds.

4 Proofs of the upper bound of Theorems 4 and 7

We first prove Theorem 7 when assuming Theorem 4 holds.

Proof of Theorem 7:

Let G be the extremal graph with m edges maximizing the number of induced copies of C_{2l+1} , where $l \geq 2$. We say an ordered vertex tuple $(v_1, v_2, \dots, v_{2l+1})$ is an induced copy of C_{2l+1} if $v_i v_{i+1} \in E(G)$ for all $1 \leq i \leq 2l+1$ (let $v_{2l+2} = v_1$), and there are no other edges among these vertices. In other words, $v_1 v_2 \dots v_{2l+1} v_1$ is an induced cycle in G . Similarly, we say an ordered vertex tuple $(v_1, v_2, \dots, v_{2l})$ is an induced copy of P_{2l} if $v_i v_{i+1} \in E(G)$ for all $1 \leq i \leq 2l-1$, and there are no other edges among these vertices. Let $C = (v_1, v_2, \dots, v_{2l+1})$ be a random vertex tuple uniformly chosen from all ordered induced copies of C_{2l+1} in G . Then the entropy of C is $H(C) = \log(2(2l+1)\Gamma)$, where Γ denotes the number of induced copies of C_{2l+1} in G .

By Shearer's Lemma (Lemma 16), we have

$$H(C) \leq \frac{1}{2l} \sum_{i=1}^{2l+1} H(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_{2l+1}).$$

Note that $(v_{i+1}, \dots, v_{2l+1}, v_1, \dots, v_{i-1})$ is always an ordered induced copy of P_{2l} in G . Let Υ denote the number of induced copies of P_{2l} in G . Then the number of ordered P_{2l}

is 2Υ and we have that $H(v_{i+1}, \dots, v_{2l+1}, v_1, \dots, v_{i-1}) \leq \log(2\Upsilon)$ for any $1 \leq i \leq 2l + 1$. Consequently, we have

$$H(C) = \log(2(2l + 1)\Gamma) \leq \frac{2l + 1}{2l} \log(2\Upsilon).$$

It is worth mentioning that the above inequality is not tight in our conjectured extremal graph, blow up of C_{2l+1} . Combining Theorem 4, we have that

$$\begin{aligned} \Gamma &\leq \frac{1}{2(2l + 1)} (2\Upsilon)^{\frac{2l+1}{2l}} \\ &\leq \frac{(2l + 1)^{l-1/2}}{2(l - 1)^{(l-1)(2l+1)/(2l)}} \left(\frac{m}{2l + 1} \right)^{(2l+1)/2}, \end{aligned}$$

which proves Theorem 7. □

In the following, we give the proof of Theorem 4.

Proof of Theorem 4: Let G be the extremal graph with m edges maximizing the number of induced copies of P_{2l} . Let Υ denote the number of induced copies of P_{2l} in G . For a well-ordered edge tuple $(e_1, e_3, \dots, e_{2t-1})$ ($1 \leq t \leq l - 2$), let $\beta(e_1, e_3, \dots, e_{2t-1})$ be the number of ordered induced copies of $P_{2l} = (e'_1, e'_2, \dots, e'_{2l})$ such that $e_{2i-1} = e'_{2i-1}$ for all $1 \leq i \leq t$ and $\alpha(e'_1, e'_3, \dots, e'_{2t-1})$ be the number of edges e such that $(e'_1, e'_3, \dots, e'_{2t-1}, e)$ is well-ordered.

Let $P = (e_1, e_2, \dots, e_{2l-1})$ be a random variable uniformly chosen from all ordered induced copies of P_{2l} in G . Then the entropy of P is $H(P) = \log(2\Upsilon)$, and from Proposition 14 we have

$$H(P) = H(e_1) + \sum_{i=1}^{l-1} H(e_{2i+1} \mid e_{<2i+1}^o) + H(e_2, e_4, \dots, e_{2l-2} \mid e_1, e_3, \dots, e_{2l-1}).$$

Firstly, note that $H(e_2, \dots, e_{2l-2} \mid e_1, e_3, \dots, e_{2l-1}) = 0$ by Proposition 13 since e_2, \dots, e_{2l-2} are uniquely determined by $e_1, e_3, \dots, e_{2l-1}$. Moreover, similar to Eq (2), we have

$$H(e_{2i+1} \mid e_{<2i+1}^o) \leq \sum_{\text{Ordered induced } P_{2l}: (\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_{2l-1})} \frac{1}{2\Upsilon} (\log \alpha(\tilde{e}_1, \tilde{e}_3, \dots, \tilde{e}_{2i-1})).$$

For convenience, write $\alpha_i = \alpha(\tilde{e}_1, \tilde{e}_3, \dots, \tilde{e}_{2i-1})$, and then we have

$$\begin{aligned}
 H(P) &= H(e_1) + \sum_{i=1}^{l-1} H(e_{2i+1} \mid e_{<2i+1}^o) \\
 &\leq \log m + \sum_{\text{Ordered induced } P_{2l}: (\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_{2l-1})} \sum_{i=1}^{l-1} \frac{1}{2\Upsilon} (\log \alpha_i) \\
 &\leq \log m + \sum_{\text{Ordered induced } P_{2l}: (\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_{2l-1})} \frac{1}{2\Upsilon} \left(\log \prod_{i=1}^{l-1} \alpha_i \right) \\
 &\leq \log m + \sum_{\text{Ordered induced } P_{2l}: (\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_{2l-1})} \frac{1}{2\Upsilon} \left(\log \left(\frac{\sum_{i=1}^{l-1} \alpha_i}{l-1} \right)^{l-1} \right).
 \end{aligned}$$

Note that, for a fixed induced copy of P_{2l} , each α_i is a cardinality of some edge set, and those edge sets are disjoint from each other by their definitions. Therefore, we have that $\sum_{i=1}^{l-1} \alpha_i \leq m$. Thus, we obtain

$$\begin{aligned}
 H(P) = \log(2\Upsilon) &\leq \log m + \sum_{\text{Ordered induced } P_{2l}: (\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_{2l-1})} \frac{1}{2\Upsilon} \log \left(\frac{m}{l-1} \right)^{l-1} \\
 &= \log m + \log \left(\frac{m}{l-1} \right)^{l-1} \\
 &= \log \frac{m^l}{(l-1)^{l-1}},
 \end{aligned}$$

which proves our theorem. □

5 Proof of the upper bound of Theorem 6

Assume $k = 2l$, we say an ordered edge tuple (e_1, e_2, \dots, e_l) **characterizes** an induced copy of C_k if we write $e_i = u_i v_i$ and then $v_i u_{i+1}$ is an edge for $i = 1, \dots, l$ (let $u_{l+1} = u_0$) and there is no other edge among $\{u_i u_j, u_i v_j, v_i u_j, v_i v_j\}_{i \neq j}$. For example, if $(e_1, e_2, \dots, e_{2l})$ is an ordered induced copy of C_{2l} , then $(e_1, e_3, e_5, \dots, e_{2l-1})$ characterizes an ordered induced C_{2l} .

When we consider unordered induced copies of C_k , we simply write $C_k = e_1 e_2 \dots e_k$ when (e_1, e_2, \dots, e_k) is an ordered induced copy of C_k . To distinguish the ordered and unordered versions, we use the notation $C_k = (e_1, e_2, \dots, e_k)$ and $C_k = e_1 e_2 \dots e_k$ for the ordered and unordered versions, respectively. For convenience, all cycles we mention in this section are induced cycles.

For a well-ordered edge tuple $(e_1, e_3, \dots, e_{2t-1})$ ($1 \leq t \leq l-1$), let $\beta(e_1, \dots, e_{2t-1})$ be the number of ordered induced copies of $C_k = (e'_1, e'_2, \dots, e'_k)$ such that $e_{2i-1} = e'_{2i-1}$ for

all $1 \leq i \leq t$. When $t \leq l - 2$, let $\alpha(e_1, e_3, \dots, e_{2t-1})$ be the number of edges e such that $(e_1, e_3, \dots, e_{2t-1}, e)$ is well-ordered. When $t = l - 1$, let $\alpha(e_1, e_3, \dots, e_{2l-3})$ be the number of edges e such that $(e_1, e_3, \dots, e_{2l-3}, e)$ characterizes an induced copy of C_k .

5.1 Case 1: $k = 4$

Given a graph G , for choosing an ordered induced copy of $C_4 = (e_1, e_2, e_3, e_4)$ in G , there are at most m choices for the first edge e_1 and also at most m choices for e_3 . Each induced C_4 is counted four times in this way. So there are at most $m^2/4$ induced copies of C_4 in G .

5.2 Case 2: $k \geq 8$

In this subsection, assume $k = 2l \geq 8$. Let G be the extremal graph maximizing the number of induced copies of C_k . Let $C = (e_1, \dots, e_{2l})$ be a random variable uniformly chosen from all ordered induced copies of C_k in G . Then the entropy of C is $H(C) = \log(2k\Gamma) = \log(4l\Gamma)$, where Γ denotes the number of induced copies of C_{2l} in G .

By Proposition 14, we have

$$H(e_1, \dots, e_{2l}) \leq H(e_1) + H(e_3 \mid e_1) + \sum_{i=2}^{l-1} H(e_{2i+1} \mid e_{<2i+1}^o) + H(e_2, \dots, e_{2l} \mid e_{<2l+1}^o).$$

Note that once we determine $e_1, e_3, \dots, e_{2l-1}$, the remaining edges e_2, e_4, \dots, e_{2l} are uniquely determined since C is an induced cycle. Hence, by Proposition 13, we have $H(e_2, \dots, e_{2l} \mid e_{<2l+1}^o) = 0$. Since (e_1, e_3) is well-ordered in C , we have

$$\begin{aligned} H(e_3 \mid e_1) &\leq \sum_{\tilde{e}_1} \mathbb{P}(e_1 = \tilde{e}_1) \log \alpha(\tilde{e}_1) \\ &= \sum_{\tilde{e}_1} \frac{\beta(\tilde{e}_1)}{4l\Gamma} \log \alpha(\tilde{e}_1) \\ &= \sum_{\tilde{e}_1} \sum_{\text{unordered induced } C_{2l}: \tilde{C}} \frac{1_{\tilde{e}_1 \text{ is in } \tilde{C}}}{2l\Gamma} \log \alpha(\tilde{e}_1) \\ &= \sum_{C_{2l} = \tilde{e}_1 \tilde{e}_2 \tilde{e}_3 \dots \tilde{e}_{2l}} \frac{1}{2l\Gamma} \sum_{i=1}^{2l} \log \alpha(\tilde{e}_i). \end{aligned}$$

Here the second equality is due to for each induced cycle \tilde{C} containing \tilde{e}_1 , there are two ordered \tilde{C} such that \tilde{e}_1 is the first coordinate. Similarly, note that $e_{<2i+1}^o$ is well-ordered in C for $i = 2, \dots, l - 1$. For convenience, let the indices of \tilde{e}_j be in mod $2l$ in the remainder

of this section. That is, $\tilde{e}_{j+2l} = \tilde{e}_j$ for $j = 1, \dots, 2l$. Then we have

$$\begin{aligned}
 & H(e_{2i+1} \mid e_{<2i+1}^o) \\
 & \leq \sum_{\tilde{e}_{<2i+1}^o} \mathbb{P}(e_{<2i+1}^o = \tilde{e}_{<2i+1}^o) \log \alpha(\tilde{e}_{<2i+1}^o) \\
 & = \sum_{\tilde{e}_{<2i+1}^o} \frac{\beta(\tilde{e}_{<2i+1}^o)}{4l\Gamma} \log \alpha(\tilde{e}_{<2i+1}^o) \\
 & = \sum_{C_{2l}=(\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_{2l})} \frac{1}{4l\Gamma} \log \alpha(\tilde{e}_{<2i+1}^o) \\
 & = \sum_{C_{2l}=\tilde{e}_1\tilde{e}_2\cdots\tilde{e}_{2l}} \frac{1}{4l\Gamma} \left(\sum_{j=1}^{2l} \log \alpha(\tilde{e}_j, \tilde{e}_{j+2}, \dots, \tilde{e}_{j+2(i-1)}) + \sum_{j=1}^{2l} \log \alpha(\tilde{e}_j, \tilde{e}_{j-2}, \dots, \tilde{e}_{j-2(i-1)}) \right).
 \end{aligned}$$

For convenience, let $\alpha_{j,i}^+ = \alpha(\tilde{e}_j, \tilde{e}_{j+2}, \dots, \tilde{e}_{j+2(i-1)})$ and $\alpha_{j,i}^- = \alpha(\tilde{e}_j, \tilde{e}_{j-2}, \dots, \tilde{e}_{j-2(i-1)})$ for $2 \leq i \leq l-1$ and $1 \leq j \leq 2l$. Here both $\alpha_{j,i}^+$ and $\alpha_{j,i}^-$ depend on some induced cycle copy of $C_{2l} = \tilde{e}_1\tilde{e}_2 \cdots \tilde{e}_{2l}\tilde{e}_1$ which is not reflected in the notation. Hence in the summation below, the value of $\alpha_{j,i}^+$ and $\alpha_{j,i}^-$ may change under different induced cycle copies of C_{2l} . Combining those inequalities above, we have

$$\begin{aligned}
 & \log(4l\Gamma) = H(e_1, \dots, e_{2l}) \\
 & \leq \log m + \sum_{C_{2l}=\tilde{e}_1\tilde{e}_2\cdots\tilde{e}_{2l}} \frac{1}{2l\Gamma} \sum_{j=1}^{2l} \log \alpha(\tilde{e}_j) + \sum_{i=2}^{l-1} \sum_{C_{2l}=\tilde{e}_1\tilde{e}_2\cdots\tilde{e}_{2l}} \frac{1}{4l\Gamma} \left(\sum_{j=1}^{2l} \log \alpha_{j,i}^+ + \sum_{j=1}^{2l} \log \alpha_{j,i}^- \right) \\
 & \leq \log m + \sum_{C_{2l}=\tilde{e}_1\tilde{e}_2\cdots\tilde{e}_{2l}} \frac{1}{4l\Gamma} \left(\sum_{j=1}^{2l} \left(\log \alpha(\tilde{e}_j) + \sum_{i=2}^{l-1} \log \alpha_{j,i}^+ \right) + \left(\log \alpha(\tilde{e}_j) + \sum_{i=2}^{l-1} \log \alpha_{j,i}^- \right) \right) \\
 & \leq \log m + \sum_{C_{2l}=\tilde{e}_1\tilde{e}_2\cdots\tilde{e}_{2l}} \frac{1}{4l\Gamma} \left(\log \left(\prod_{j=1}^{2l} \left(2^{\frac{\alpha(\tilde{e}_j)}{2}} \prod_{i=2}^{l-1} \alpha_{j,i}^+ \right) \right) + \log \left(\prod_{j=1}^{2l} \left(2^{\frac{\alpha(\tilde{e}_j)}{2}} \prod_{i=2}^{l-1} \alpha_{j,i}^- \right) \right) \right).
 \end{aligned}$$

Using AM-GM inequality, we have

$$\log \left(\prod_{j=1}^{2l} \left(2^{\frac{\alpha(\tilde{e}_j)}{2}} \prod_{i=2}^{l-1} Q_{j,i} \right) \right) \leq \log \left(2^{\frac{1}{l-1}} \frac{\sum_{j=1}^{2l} \left(\frac{\alpha(\tilde{e}_j)}{2} + \sum_{i=2}^{l-1} Q_{j,i} \right)}{2l(l-1)} \right)^{2l(l-1)},$$

where $Q_{j,i} \in \{\alpha_{j,i}^+, \alpha_{j,i}^-\}$. Write $S_j^+ = \frac{\alpha(\tilde{e}_j)}{2} + \sum_{i=2}^{l-1} \alpha_{j,i}^+$ and $S_j^- = \frac{\alpha(\tilde{e}_j)}{2} + \sum_{i=2}^{l-1} \alpha_{j,i}^-$ for $j = 1, \dots, 2l$. Then we have the following claim.

Claim 17. For any (unordered) induced cycle $C_{2l} = \tilde{e}_1\tilde{e}_2 \cdots \tilde{e}_{2l}\tilde{e}_1$, we have $\sum_{j=1}^{2l} S_j^+ \leq ml$ and $\sum_{j=1}^{2l} S_j^- \leq ml$.

Let us first assume Claim 17 holds. Then, we have

$$\begin{aligned} \log(4l\Gamma) &= H(e_1, \dots, e_{2l}) \\ &\leq \log m + \sum_{C_{2l}=\tilde{e}_1\tilde{e}_2\cdots\tilde{e}_{2l}} \frac{1}{4l\Gamma} 2 \log \left(2^{\frac{1}{l-1}} \frac{ml}{2l(l-1)} \right)^{2l(l-1)} \\ &\leq \log \left(2m^l \left(\frac{1}{2(l-1)} \right)^{l-1} \right). \end{aligned}$$

As a conclusion, we have that $\rho(C_k, m) = \Gamma \leq \left(\frac{m}{2l}\right)^l \left(1 + \frac{1}{l-1}\right)^{l-1} \leq e \left(\frac{m}{2l}\right)^l = e \left(\frac{m}{k}\right)^{k/2}$ and we are done. In the following, we give the proof of Claim 17.

Proof of Claim 17:

Assume the induced cycle C_{2l} has vertices v_1, v_2, \dots, v_{2l} and $\tilde{e}_i = v_i v_{i+1}, i = 1, \dots, 2l$ where $v_{2l+1} = v_1$. We only prove $\sum_{j=1}^{2l} S_j^+ \leq ml$ since the proof for $\sum_{j=1}^{2l} S_j^- \leq ml$ is similar. Note that $\alpha(\tilde{e}_j)$ and $\alpha_{j,i}^+$ ($i = 2, \dots, l-1$) are defined by the cardinality of some edge set. To estimate S_j^+ , we only need to count for each edge how much it contributes to S_j^+ . In fact, we can easily verify that for a fixed j , each edge contributes at most $3/2$ (when counted in both $\frac{\alpha(\tilde{e}_j)}{2}$ and $\alpha_{j,2}^+$) to S_j^+ . So a natural bound is $\sum_{j=1}^k S_j^+ \leq 3ml$.

By a refined analysis, we can derive the claim. We say an edge $e = uv$ is adjacent to w if either $uw \in E(G)$ or $vw \in E(G)$. Given e , let $J_e = \{j \mid v_j \text{ is adjacent to } e\}$. For convenience, let the indices of v_j be in mod $2l$ in the remainder of this section. That is, $v_{j+2l} = v_j$ for $j = 1, \dots, 2l$. Here, e may intersect the cycle but this does not affect our proof.

We now analyze the contribution of e to S_j^+ :

1. When $j \in J_e$ and $j + 1 \in J_e$, e contributes 0 to S_j^+ .
2. When $j \in J_e$ and $j + 1 \notin J_e$,
 - (a) if $j + 2, j + 3, \dots, j + 2l - 4 \notin J_e$ and $j - 3 \in J_e$, e contributes at most $3/2$ to S_j^+ ,
 - (b) otherwise, e contributes at most $1/2$ to S_j^+ .
3. When $j \notin J_e$ and $j + 1 \in J_e$, e contributes at most $1/2$ to S_j^+ .
4. When $j, j + 1 \notin J_e$, let j' be the next element in J_e after j .
 - (a) If $|j' - j|$ is odd, e contributes at most 1 to S_j^+ .
 - (b) If $|j' - j|$ is even, then e contributes 0 to S_j^+ ,

Now consider the contribution of e to $\sum_{j=1}^{2l} S_j^+$. If e contributes $3/2$ to S_j^+ for some j , then $\{j, j - 3\} \subseteq J_e \subseteq \{j, j - 1, j - 2, j - 3\}$ (see Figure 3). It is worth mentioning that if $k = 6$, Claim 17 fails on this occasion when e contributes $3/2$ to S_j^+ . Enumerate all possible J_e and we have e contributes at most l to $\sum_{j=1}^{2l} S_j^+$.

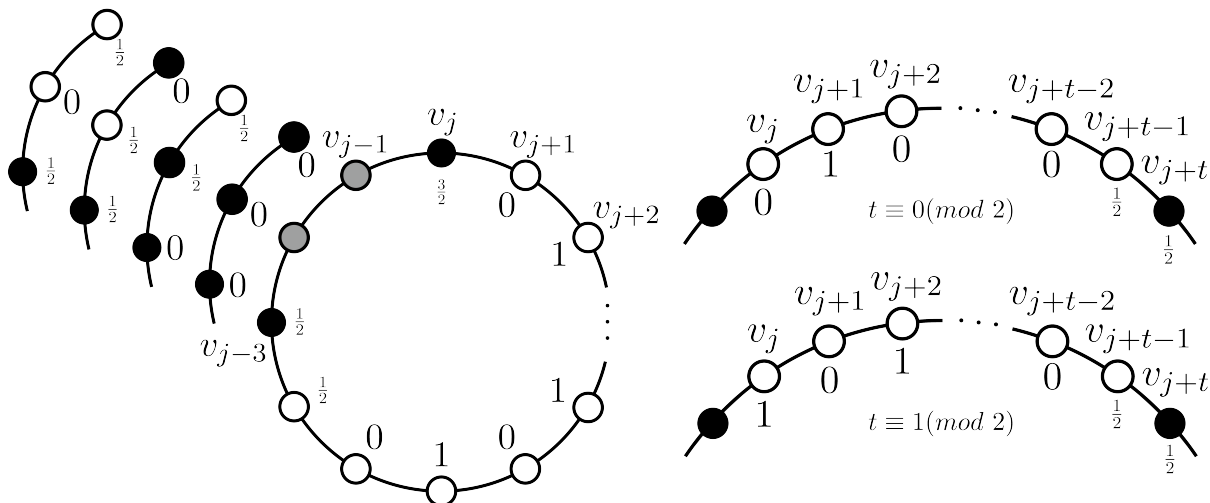


Figure 3: The upper bound of the contribution of e to $\sum_{j=1}^{2l} S_j^+$. The black nodes represent vertices in J_e , the white nodes represent vertices not in J_e and gray nodes represent vertices which are not sure whether in J_e . The number on the side of each v_j is an upper bound of the contribution of e to S_j^+ . The graph on the left shows when e contributes $3/2$ to some S_j^+ . At the upper left, we enumerate all possibilities of the gray nodes. The graph on the right shows the other case, split the cycle into “white node” paths of length t .

Otherwise, e contributes at most 1 to each S_j^+ . Then we can split the cycle into maximal paths such that all vertices in the path are not in J_e . For each such path $v_j, v_{j+1}, \dots, v_{j+t-1}$, the contribution of e to $\sum_{j'=j}^{j+t} S_{j'}^+$ is at most $t+1$ (see Figure 3). Sum them up then we have e contributes at most l to $\sum_{j=1}^{2l} S_j^+$.

As a conclusion, we have $\sum_{j=1}^{2l} S_j^+ \leq ml$. □

5.3 Case 3: $k = 6$

When $k = 2l = 6$, Claim 17 fails. However, we can still get a weaker result that $\sum_{j=1}^k S_j^+ \leq (l+1)m$. It leads to the upper bound $\rho(C_6, m) \leq 4 \left(\frac{m}{6}\right)^3$. The details are omitted here.

Note that in this case, we have a worse constant in the upper bound compared to the case when $k \geq 8$. It is interesting to ask whether we can prove a sharp upper bound for $k = 6$.

Remark

We remark that there is a recent independent work by Chao, Antonir, Li and Yu [4] which also studies the edge inducibility problem. We thank them very much for giving some vital comments, including pointing out that Theorem 3 is the corollary of Theorem 2.

They define the **edge inducibility** of graph G by

$$eind(G) := \limsup_{m \rightarrow \infty} \frac{|Aut(G)|\rho(G, m)}{(2m)^{\alpha_f(G)}} \in [0, 1].$$

They also give the edge inducibility of P_4 and P_6 as the following theorem.

Theorem 18 ([4]).

$$eind(P_4) = \frac{1}{4}, \frac{5}{372} \leqslant eind(P_6) \leqslant \frac{1}{36},$$

which disproves our Conjecture 9 for $k = 4, 6$. Our Conjecture 8 coincides with their Conjecture 1.7.

Acknowledgements

The authors would like to thank the referee for their careful reading and valuable comments which help to improve the presentation of this paper. M. Lu is supported by the National Natural Science Foundation of China (Grant 12571372).

References

- [1] N. Alon. On the number of subgraphs of prescribed type of graphs with a given number of edges. *Israel Journal of Mathematics*, 38(1):116–130, 1981.
- [2] N. Alon and J. H. Spencer. *The probabilistic method*. John Wiley & Sons, 2016.
- [3] J. Balogh, P. Hu, B. Lidický, and F. Pfender. Maximum density of induced 5-cycle is achieved by an iterated blow-up of 5-cycle. *European Journal of Combinatorics*, 52:47–58, 2016.
- [4] T.-W. Chao, A. Cohen Antonir, A. Li, and H.-H. H. Yu. Edge inducibility via local directed graphs. [arXiv:2509.24064](https://arxiv.org/abs/2509.24064) 2025.
- [5] T.-W. Chao, Z. Dong, Z. Shen, and N. Yang. Many cliques with small degree powers. [arXiv:2410.04744](https://arxiv.org/abs/2410.04744), 2024.
- [6] T.-W. Chao and H.-H. H. Yu. Kruskal–Katona-type problems via the entropy method. *Journal of Combinatorial Theory, Series B*, 169:480–506, 2024.
- [7] E. Friedgut and J. Kahn. On the number of copies of one hypergraph in another. *Israel Journal of Mathematics*, 105(1):251–256, 1998.
- [8] D. Galvin. Three tutorial lectures on entropy and counting. [arXiv:1406.7872](https://arxiv.org/abs/1406.7872), 2014.
- [9] D. Hefetz and M. Tyomkyn. On the inducibility of cycles. *Journal of Combinatorial Theory, Series B*, 133:243–258, 2018.
- [10] G. Katona. A theorem of finite sets. *Classic Papers in Combinatorics*, pages 381–401, 1987.
- [11] J. B. Kruskal. The number of simplices in a complex. *Mathematical optimization techniques*, 10:251–278, 1963.

- [12] D. Král', S. Norin, and J. Volec. A bound on the inducibility of cycles. *Journal of Combinatorial Theory, Series A*, 161:359–363, 2019. ISSN 0097-3165.
- [13] G. L. Nemhauser and L. E. Trotter Jr. Vertex packings: structural properties and algorithms. *Mathematical Programming*, 8(1):232–248, 1975.
- [14] N. Pippenger and M. C. Golumbic. The inducibility of graphs. *Journal of Combinatorial Theory, Series B*, 19(3):189–203, 1975.
- [15] W. Willis. Bounds for the independence number of a graph. Virginia Commonwealth University, 2011.