Homomorphisms into loop-threshold graphs

Jonathan Cutler[*]

Department of Mathematics
Montclair State University
Montclair, NJ 07043 USA

jonathan.cutler@montclair.edu

Nicholas Kass

Department of Mathematics
University of Nebraska Omaha
Omaha, NE 68182 USA

nkass@unomaha.edu

Abstract

Many problems in extremal graph theory correspond to questions involving homomorphisms into a fixed image graph. Recently, there has been interest in maximizing the number of homomorphisms from graphs with a fixed number of vertices and edges into small image graphs. For the image graph H_{ind}, the graph on two adjacent vertices, one of which is looped, each homomorphism from G to H_{ind} corresponds to an independent set in G. It follows from the Kruskal-Katona theorem that the number of homomorphisms to H_{ind} is maximized by the lex graph, whose edges form an initial segment of the lex order.

A *loop-threshold graph* is a graph built recursively from a single vertex, which may be looped or unlooped, by successively adding either a looped dominating vertex or an unlooped isolated vertex at each stage. Thus, the graph H_{ind} is a loop-threshold graph. We survey known results for maximizing the number of homomorphisms into small loop-threshold image graphs. The only extremal homomorphism problem with a loop-threshold image graph on at most three vertices not yet solved is $H_{\text{ind}} \cup E_1$, where extremal graphs are the union of a lex graph and an empty graph. The only question that remains is the size of the lex component of the extremal graph. While we cannot give an exact answer for every number of vertices and edges, we establish the significance of and give bounds on $\ell(m)$, the number of vertices in the lex component of the extremal graph with m edges and at least $m + 1$ vertices.

Mathematics Subject Classifications: 05C35,05C60

1 Introduction

Many problems in classical extremal graph theory can be stated in terms of graph homomorphisms. A *homomorphism* from a graph G to a graph H is a function $\phi : V(G) \to V(H)$ such that $\phi(x)\phi(y) \in E(H)$ whenever $xy \in E(G)$, i.e., ϕ is an edge-preserving map. We let $\mathrm{Hom}(G, H)$ be the set of homomorphisms from G to H and $\hom(G, H) = |\mathrm{Hom}(G, H)|$. If we take the image graph to be H_{ind}, a path on two vertices with one vertex looped (see Figure 1), then elements of $\hom(G, H_{\mathrm{ind}})$ correspond to independent sets in G. This is because vertices mapped to the unlooped vertex of H_{ind} form an independent set and there are no other restrictions on the map. Another

Figure 1: The graph H_{ind}.

natural example arises from taking the image graph to be K_q, in which case elements of $\mathrm{Hom}(G, K_q)$ correspond to proper q-colorings of G.

An interesting class of problems arises by fixing an image graph H and a class of graphs $\mathcal{G}$ and trying to determine which graphs $G \in \mathcal{G}$ maximize $\hom(G, H)$. A lot of recent research has focused on this problem for the image graph H_{ind}, which we saw above corresponds to independent sets, and so we let $i(G) = \hom(G, H_{\mathrm{ind}})$. Kahn [14] gave a bound for $i(G)$ when G is an r-regular bipartite graph on n vertices and Zhao [19] extended this to general regular graphs.

Theorem 1 (Kahn, Zhao). *If G is an r-regular graph on n vertices, then*

$$i(G) \leqslant i(K_{r,r})^{n/2r}.$$

Galvin and Tetali [12] extended the bipartite case of Theorem 1 to general homomorphisms to any image graph. Interestingly, the extension to nonbipartite graphs in the general case is not true. (See [11] for more details.) Also, the problem of maximizing the number of independent sets in graphs with given minimum degree has been well-studied (see, e.g., [1, 2, 7, 8, 9, 10, 13, 20]).

In this paper, we will be interested in maximizing the number of homomorphisms over graphs of fixed order (number of vertices) and size (number of edges). We let $\mathcal{G}(n, m)$ be the set of loopless graphs with n vertices and m edges. The solution to the problem of maximizing the number of independent sets over $\mathcal{G}(n, m)$ follows from the Kruskal-Katona [16, 15] theorem. This is because independent sets in a graph G correspond exactly to those sets that are not in the upper shadow of G, where G is thought of as a set system on the vertex set. Recall that the *lexicographic order* on subsets of $[n]$ is defined by $A < B$ if $\min(A \triangle B) \in A$. Define the *lex graph of order n and size m*, denoted $L(n, m)$, to be the graph with vertex set $[n] = \{1, 2, \ldots, n\}$ and edges consisting of the initial m elements of $\binom{[n]}{2} = \{e \subset [n] : |e| = 2\}$ according to the lexicographic order. We have the following, as was noted in [4].

Theorem 2. *If $G \in \mathcal{G}(n, m)$, then*

$$i(G) \leqslant i(L(n, m)),$$

with equality if and only if $G \cong L(n, m)$.

Theorem 2 was also proved independently by Wood [18]. In fact, since the Kruskal-Katona theorem implies that the lex graph is extremal for independent sets of any fixed size and, as was noted in [5], we have the following, where $i_t(G)$ is the number of independent sets of size t in G.

Theorem 3. *If $G \in \mathcal{G}(n, m)$, then*

$$i_t(G) \leqslant i_t(L(n, m)),$$

with equality if and only if $G \cong L(n, m)$.

This paper will begin with a survey of results about maximizing $\hom(G, H)$ where $G \in \mathcal{G}(n, m)$ and H is a small fixed image graph. Throughout the paper, we assume G is a simple graph, but H may contain loops. In Section 2, we introduce threshold graphs which will turn out to be extremal graphs for a set of image graphs consisting of what we call loop-threshold graphs. In Section 3, we begin our investigation into the one remaining open extremal problem for loop-threshold image graphs on at most three vertices. Finally, in Section 4, we give bounds on a function related to this extremal problem.

2 Threshold and loop-threshold graphs

Threshold graphs play a key role in the investigation of extremal problems related to graph homomorphisms. There are many characterizations of threshold graphs (see, e.g., [17]), but the most useful one for our purposes is as follows.

Definition 4. A graph G is a *threshold graph* if it can be constructed recursively from K_1 by successively adding either a dominating vertex or an isolated vertex.

With this recursive definition in hand, we can observe that threshold graphs on n vertices can by represented by binary sequences of length $n - 1$, which we call the *code* of the threshold graph. We write 1 for a dominating vertex and a 0 for an isolated vertex. As the first vertex is irrelevant in this construction, we omit it from the code. Following convention, we write the code from right to left. Note that for a given code, 1s are adjacent to all vertices to their right in the code and 1s to their left. 0s are adjacent only to 1s to their left. Note that threshold graphs have at most one nontrivial component.

We often use superscripts to denote a string of the same symbol in the code of a threshold graphs so, $0^p 1^q$ is the code with q 1s followed by p 0s (from right to left). We note that any lex graph is a threshold graph. In general, the code of a lex graph is of the form $1^p 0^q 1^a 0^r$, where p, q, and r are all nonnegative integers and a may be either 0 or 1 (where we write x^0 for the empty string).

We say that a graph G is H-*extremal* for a graph H (which may have loops) if

$$\text{hom}(G, H) = \max\left\{\text{hom}(G', H) \ : \ n(G') = n(G) \text{ and } e(G') = e(G)\right\}.$$

It turns out that for many image graphs H, one is able to prove that there is always an H-extremal graph that is threshold. One such class of image graphs are called loop-threshold graphs.

Definition 5. A graph with (perhaps) loops is *loop-threshold* if it can be obtained from K_1, or K_1 with a loop, by successively adding an unlooped isolated vertex, or a looped dominating vertex.

Once again, loop-threshold graphs on n vertices can be associated with a binary code, except now the length of the code is n (rather than $n-1$). This is due to the fact that the code of the first vertex is now relevant since it determines whether that vertex is looped or not. We note that H_{ind} is a loop-threshold graph with code 10. The following was proved in [6].

Theorem 6. *If H is any loop-threshold graph and n and m are any non-negative integers with $0 \leqslant m \leqslant \binom{n}{2}$, then there is a threshold graph $T \in \mathcal{G}(n, m)$ such that T is H-extremal.*

The aim of the remainder of this section is to survey extremal results for $\text{hom}(G, H)$ where H is a loop-threshold graph on at most three vertices. We begin by making a couple of trivial observations. The first involves the situation when the image graph is the empty graph.

Observation 7. *If $q \geqslant 1$ is an integer and H is the loop-threshold graph with code 0^q, i.e., $H = E_q$, then $\text{hom}(G, H) = 0$ for any graph G with at least one edge.*

Another simple result holds for any image loop-threshold graph which is complete with loops on all vertices.

Observation 8. *If $p \geqslant 1$ is an integer and H is a loop-threshold graph with code 1^p and G is a graph with n vertices, then $\text{hom}(G, H) = p^n$.*

There is one more simple case, namely when H is a loop-threshold graph with code of the form $0^q 1^p$. In this case, H is the disjoint union of a fully-looped K_p and E_q.

Proposition 9. *If $p \geqslant 1$ and $q \geqslant 0$ are integers and H is a loop-threshold graph with code $0^q 1^p$ and G is a graph on n vertices with c isolated vertices, then $\text{hom}(G, H) = (p+q)^c p^{n-c}$.*

Proof. Any of the c isolates in G can be mapped to any vertex of H, while the other $n-c$ non-isolates in G can only be mapped to any of the p non-isolates in H. $\qquad\square$

If H is a loop-threshold graph with code $0^q 1^p$, then Proposition 9 implies $\text{hom}(G, H)$ is a simple function of the number of isolates when $G \in \mathcal{G}(n, m)$. To maximize $\text{hom}(G, H)$
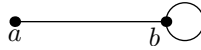
Figure 2: A labeled H_{ind}

over $G \in \mathcal{G}(n, m)$, one simply has to maximize the number of isolates in G. This can be done by a variety of graphs, including the colex[1] graph $C(n, m)$.

The final general case we would like to present in this section is that of image loop-threshold graphs with code $1^p 0^q$. We note that if H is a loop-threshold graph with code $1^p 0^q$ and we consider a homomorphism $\phi \in \text{Hom}(G, H)$ for some graph G, then the vertices of G mapped to the vertices with code 0 in H form an independent set in G. Thus, as one might expect, we can deduce the extremal question for maximizing homomorphisms into H from earlier results about independent sets.

As one does often in statistical physics, we can weight homomorphisms to a graph H with respect to a weighting function $\beta : V(H) \to [0, \infty)$. Then, we "count" weighted homomorphisms from a graph G to H via the partition function:

$$\text{hom}_\beta(G, H) = \sum_{\phi \in \text{Hom}(G,H)} \prod_{v \in G} \beta\left(\phi(v)\right).$$

Note that this reduces to counting homomorphisms if $\beta \equiv 1$, that is, $\text{hom}_\beta(G, H) = \text{hom}(G, H)$. We are interested in the case of weighted independent sets and so, if we label the vertices of H_{ind} as in Figure 2, we define the weight function β^λ to be the weight function on H_{ind} defined by

$$\beta^\lambda(x) = \begin{cases} \lambda & \text{if } x = a \\ 1 & \text{if } x = b. \end{cases}$$

We see that, with this weighting of H_{ind}, an independent set I is assigned weight $\lambda^{|I|}$. But then the weighted homomorphism model introduced above yields the independence polynomial of the graph, i.e., $P_G(\lambda) = \text{hom}_{\beta^\lambda}(G, H)$. Since $P_G(\lambda)$ depends only on the number of independent sets of various sizes and the lex graph maximizes the number of independent sets of all sizes simultaneously, by Theorem 3, we have the following.

Corollary 10. *If G is a graph on n vertices and m edges and $\lambda > 0$, then*

$$P_G(\lambda) \leqslant P_{L(n,m)}(\lambda).$$

Finally, we can return to our original question regarding image loop-threshold graphs with codes of the form $1^p 0^q$. We let $S^\circ(p, q)$ be the clique-looped split graph $K_p^\circ \vee E_q$, in which each vertex of K_p is looped.[2] In other words, $S^\circ(p, q)$ is a loop-threshold graph with code $1^p 0^q$. We have the following.

---

[1]The *colexicographic order* on subsets of $[n]$ sets $A < B$ if $\max(A \triangle B) \in B$. The *colex graph with n vertices and m edges*, denoted $C(n, m)$, has vertex set $[n]$ and edge set consisting of the initial m elements of $\binom{[n]}{2}$ in the colexicographic order.

[2]The *join* of graphs G and H, denoted $G \vee H$, has vertex set $V(G) \cup V(H)$ and edge set $E(G) \cup E(H) \cup \{xy : x \in V(G), y \in V(H)\}$.

Corollary 11. *Let $p, q \geqslant 1$ and G be a graph with n vertices and m edges. Then*

$$\hom(G, S^\circ(p,q)) \leqslant \hom(L(n,m), S^\circ(p,q)).$$

Proof. Set $\lambda = p/q$. Then, we have $\hom(G, S^\circ(p,q)) = q^n P_G(\lambda)$ and so the result follows from Corollary 10. □

Armed with the above results, we present in Table 1 all loop-threshold graphs on three or fewer vertices. All cases except for the last two are covered by one of the preceding results in this section. The penultimate image graph F (with code 101) is known as "the fox" (or "the wrench" by Brightwell and Winkler [3]) and the extremal question for this graph was studied by Cutler and Radcliffe [6]. They were able to determine a class of five threshold graphs which formed a minimal class of F-extremal graphs.

The final image graph (with code 010), which we denote J, will be the main focus of this paper. While it is relatively easy to determine the form of all J-extremal graphs, we will determine the size of the non-trivial component asymptotically up to a constant factor.

3 The remaining case

As mentioned in the previous section, the only case remaining in finding a minimal set of H-extremal graphs, where H is a loop-threshold graph on at most three vertices, is that of $H = J$. (See Figure 3.) The code of J is 010.
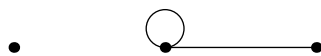

Figure 3: The loop-threshold graph J.

Given the fact that $J = E_1 \cup H_{\mathrm{ind}}$, it is not surprising that the threshold J-extremal graphs are comprised of a lex component with some number of added isolated vertices. Note that, by Theorem 6, for any n and m, there is at least one threshold J-extremal graph. For a graph G, we let $j(G) = \hom(G, J)$.

Theorem 12. *Every J-extremal threshold graph is the union of a lex graph and an empty graph.*

Proof. The result is trivial if $m = 0$. If $m > 0$ and T is a threshold graph on n vertices and m edges then it contains a single nontrivial component, C. Let $q = |V(C)|$. Since each vertex outside of that component can map to any vertex of J under a homomorphism in $\mathrm{Hom}(T, J)$, and every vertex contained in that component must map to the H_{ind} subgraph of J, the number of homomorphisms from T to J is

$$j(T) = 3^{n-q} i(C),$$
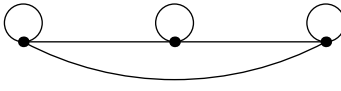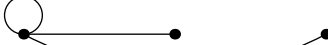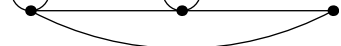
Code of H	H	H-extremal graph
0		Any
1		Any
00		Any
11		Any
01		$C(n,m)$
10		$L(n,m)$
000		Any
111		Any
001		$C(n,m)$
011		$C(n,m)$
100		$L(n,m)$
110		$L(n,m)$
101		See [6]
010		See Section 3

Table 1: H-extremal graphs for all loop-threshold graphs H on at most three vertices

where $i(C)$ is the number of independent sets in C. As Theorem 2 establishes that $i(C) \leqslant i(L(q,m))$, with equality if and only if $C \cong L(q,m)$, we have that

$$
\begin{aligned}
j(T) &= 3^{n-q}\, i(C) \\
&\leqslant 3^{n-q}\, i(L(q,m)) \\
&= j(E_{n-q} \cup L(q,m)),
\end{aligned}
$$

where equality holds in the second step if and only if $T \cong E_{n-q} \cup L(q,m)$. $\qquad\square$

As this construction of lex graphs with added isolated vertices occurs so frequently it is natural to define it as a function of the total number of vertices, the number of vertices in the lex component, and the number of edges.

Definition 13. For $n, q, m \in \mathbb{N}$ where $q \leqslant \min\{n, m+1\}$ and $0 \leqslant m \leqslant \binom{q}{2}$, let $R(n,q;m)$ be the graph on n vertices and m edges defined by the union

$$
R(n,q;m) = L(q,m) \cup E_{n-q}.
$$

The restriction $q \leqslant m+1$ assures that $L(q, m)$ contained no isolated vertices. We also note that

$$j\left(R(n, q ; m)\right)=3^{n-q} i\left(L(q, m)\right) . \tag{1}$$

4 An extremal question

In light of Theorem 12, finding a J-extremal threshold graph on n vertices and m edges is thus reducible to finding q within the bounds of $m \leqslant \binom{q}{2}$ and $q \leqslant m+1$ such that $R(n, q ; m)$ is J-extremal. That is to say that the question of determining J-extremality is solely a matter of finding an appropriate number of vertices on which to form a lex component while leaving the rest as isolated vertices. This is the problem with which the remainder of this paper will be concerned.

Ideally, one would be able to find the size of the lex component in the J-extremal graph for any n and m. In general however, this does not appear to be an amenable question. For example, if one fixes n and asks how the size of the lex component changes as m increases, one might hope that this change is at least monotonic even if not easy to precisely quantify. This, unfortunately, is not the case. Figure 4 illustrates this when $n = 50$ which appears to typify the behavior throughout all values of n based on computer testing.
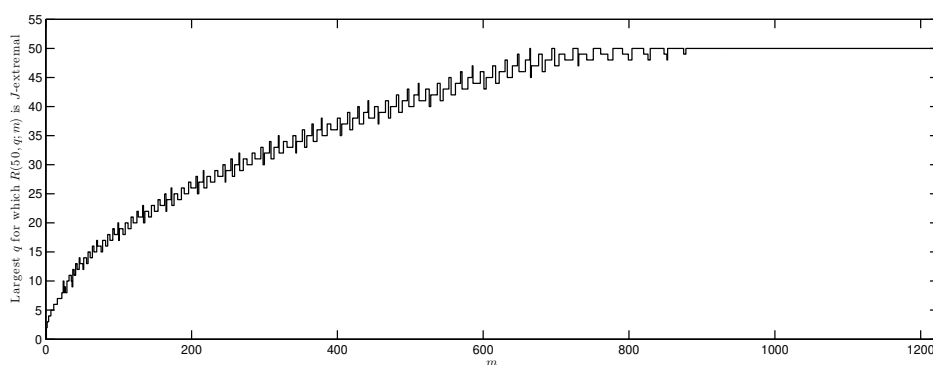
Figure 4: Computer testing showing non-monotone behavior of the lex component size in J-extremal graphs in the case where $n = 50$.

Conversely, if one fixes the number of edges, m, the size of the lex component should remain fixed (at least in the limit) as n increases to accord with equation (1) above since the size of the lex component cannot exceed $m + 1$ vertices. In light of this non-monotonicity in terms of increasing values of m and in order to capitalize on the relative stability of the problem in terms of increasing values of n, we instead consider the size of the lex component when m is fixed. We define a function related to this behavior as follows.

Definition 14. For $m \in \mathbb{N}$, let $\ell(m)$ be the maximum number of vertices in the non-trivial component of a J-extremal threshold graph on $m+1$ vertices and m edges. Equivalently,

$$\ell(m) = \max\{q \ : \ R(m+1, q; m) \text{ is } J\text{-extremal}\}$$

Here any uniqueness issues involving the extremal graphs are eliminated by the use of the maximum in our definition. The choice of defining the composition of the extremal graph on exactly $m+1$ vertices relative to m edges is not arbitrary; it corresponds to the largest possible size of the lex component in an extremal graph. The following theorem proves the stronger claim that there is a single size of the lex component in a J-extremal graph on m edges once n is large enough.

Theorem 15. *The graph $R(n, \ell(m); m)$ on n vertices and m edges is J-extremal for $n \geqslant \ell(m)$.*

Proof. Since we have required $n \geqslant \ell(m)$ we may always add or remove $c = m + 1 - n$ isolated vertices to the graph $R(n, \ell(m); m)$ to produce the graph $R(m+1, \ell(m); m)$, in accordance with the sign of c. A simple counting argument on these isolates yields

$$j\left(R(n, \ell(m); m)\right) = 3^{-c} j\left(R(m+1, \ell(m); m)\right),$$

whereby the J-extremality of $R(n, \ell(m); m)$ is now seen to be a consequence of the J-extremality of $R(m+1, \ell(m); m)$ which holds by the definition of $\ell(m)$. $\qquad\square$

Having established the significance of $\ell(m)$ in the construction of J-extremal graphs, we will concentrate on establishing an upper bound on $\ell(m)$. As the number of homomorphisms to J is so closely related to the number of independent sets in the lex component, the matter of bounding $\ell(m)$ from above can be approached only once reasonable bounds are found for $i(L(n, m))$. Before embarking on these calculations, we pause to investigate some of the properties of the lex graph. The lex graph $L(n, m)$ is related to (and sometimes isomorphic to) the *split graph*. We define the *split graph* $S(n, k)$ is the graph with n vertices defined as the join of K_k and E_{n-k}. Thus, the number of edges in $S(n, k)$ is

$$e(S(n, k)) = \binom{k}{2} + k(n - k).$$

If in the split graph $S(n, k) = K_k \vee E_{n-k}$, we label the vertices of the K_k with $\{1, 2, \ldots, k\}$ and those of the E_{n-k} with $\{k+1, k+2, \ldots, n\}$, then we can use this to describe the structure of the lex graph. If $m = \binom{k}{2} + k(n-k) + w$, we note that the lex graph with n vertices and m edges is

$$L(n, k, w) := L(n, m) = S(n, k) + \{(k+1)x \ : \ k+2 \leqslant x \leqslant k + w + 2\}.$$

Given n and m, we would like to determine k and w so that $L(n, m) = L(n, k, w)$. Throughout this paper, we assume that $0 \leqslant w \leqslant n - k - 2$ so that there is no ambiguity about the value of k and w for a given n and m. Essentially, the lex graph $L(n, k, w)$
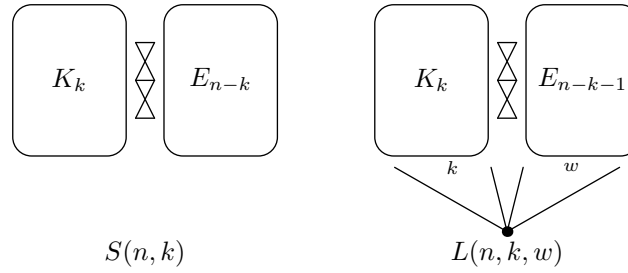
Figure 5: Schematics of the split graph and the lex graph

consists of the split graph $S(n, k)$ along with a star with w edges inside the E_{n-k}. (See Figure 5.) We note that in the threshold code of the lex graph, the center of the star of size w in $L(n, k, w)$ corresponds to the lone 1 inside of the string of initial 0s. From this point forward, we refer to the two types of vertices in the lex graph as those from the *complete part* and those from the *empty part* (even though the empty part may contain a star). Given the lex graph $L(n, k, w)$, it is easy to determine the number of independent sets in it, as we do in the following lemma.

Lemma 16. *For the lex graph $L(n, k, w)$, we have*

$$i(L(n, k, w)) = 2^{n-k-1} + 2^{n-k-w-1} + k.$$

Proof. Note first that if any of the vertices in the complete part of the lex graph $L(n, k, w)$ are included in an independent set, then no other vertices can be in the independent set since these vertices are dominating. There are k such independent sets. If, on the other hand, none of the vertices in the complete side are in an independent set I and the center of the star of size w in the empty part is also not in I, then there are 2^{n-k-1} such independent sets (including the empty set) since the rest of the vertices form an independent set. If the center of the star of size w in the empty set is in I, then none of the other vertices of the star can be in I, and so there are $2^{n-k-w-1}$ such I. These account for all independent sets in $L(n, k, w)$, proving the lemma. $\square$

By extension we may also determine $j(R(n, n'; m))$ in a general form. Noting that isolates can be mapped to any vertex of J by an element of $\text{Hom}(R(n, n'; m))$, we have by Lemma 16,

$$j(R(n, n'; m)) = 3^{n-n'} \left[k + 2^{n'-k-1} + 2^{n'-k-w-1} \right]$$

where $k = k(n', m)$ and $w = w(n', m)$.

Now, we are finally ready to state and prove the main theorem of the paper, which gives asymptotically sharp bounds on $\ell(m)$.

Theorem 17. *For every positive integer $m \geqslant \binom{10}{2}$, there are explicit constants c and d such that*

$$\beta \sqrt{m} - d \sqrt[4]{m} \leqslant \ell(m) \leqslant \beta \sqrt{m} + c \sqrt[4]{m},$$

where $\beta = \frac{1}{\sqrt{\alpha + 1/2}}(\alpha + 1)$ and $\alpha = 1/\log_2(3/2)$, so that $\beta \approx 1.8228$.

Proof. We first prove the upper bound on $\ell(m)$. The idea is compare $j(L(n, m))$ to $j(R(n, n - r; m)) = 3^r j(L(n - r, m))$. When the latter of these two is larger, we can remove r vertices from the lex component of a potentially extremal graph (i.e., the disjoint union of a lex graph and an empty graph) and increase the number of homomorphisms into J. Since we need to compare the number of homomorphisms into J, we let k and w be the integers so that $L(n, m) = L(n, k, w)$. Likewise, we suppose that k_r and w_r are the integers so that $L(n - r, m) = L(n - r, k_r, w_r)$. Note that $k_r \geqslant k$ as the removal of vertices from a lex graph forces the new lex component to be denser. We will show that the number of homomorphisms into J goes up provided that k_r is not too much larger than k. Let d_r be defined by

$$d_r = r \log_2 \left(\frac{3}{2} \right).$$

Now, suppose that $k_r \leqslant k + d_r - 1$. (We will determine when this happens in just a bit.) In this case, we have

$$
\begin{aligned}
i(L(n, k, w)) &= 2^{n-k-1} + 2^{n-k-w-1} + k \\
&\leqslant 2^{n-k} + k \\
&= 2^{n-k} \cdot 3^r \cdot 2^{-(r+d_r-1)-1} + k \\
&< 3^r (2^{n-r-(k+d_r-1)-1} + k_r) \\
&\leqslant 3^r (2^{n-r-(k+d_r-1)-1} + 2^{(n-r)-(k+d_r-1)-w_r+1} + k_r) \\
&\leqslant 3^r i(L(n - r, k_r, w_r)).
\end{aligned}
\tag{2}
$$

The strict inequality holds provided that $r \geqslant 1$. The question remains as to when $k_r \leqslant k + d_r - 1$. Suppose, to the contrary, that $k_r \geqslant k + d_r$. Then we know that, since $w \leqslant n - k - 2$,

$$
\begin{aligned}
e(L(n, k, n - k - 2)) &\geqslant e(L(n, k, w)) \\
&= e(L(n - r, k_r, w_r)) \\
&\geqslant e(L(n - r, k_r, 0)) \\
&\geqslant e(L(n - r, k + d_r, 0)).
\end{aligned}
$$

Thus, writing out the first and last expressions in the above, we have

$$\binom{k}{2} + k(n - k) + n - k - 2 \geqslant \binom{k + d_r}{2} + (k + d_r)(n - r - k - d_r).$$

This reduces to

$$n - k - 2 \geqslant \frac{d_r(2k + d_r - 1)}{2} + k(-r - d_r) + d_r(n - r - k - d_r).$$

Rearranging, we get

$$n(1 - d_r) \geqslant k(1 - d_r - r) - \binom{d_r + 1}{2} - rd_r + 2.$$

Note that if $d_r > 1$, which happens provided that $r \geqslant 2$, then $1 - d_r < 0$. So, if $k_r \geqslant k + d_r$ and $r \geqslant 2$, we know that

$$n \leqslant \left(\frac{1 - d_r - r}{1 - d_r}\right)k + \frac{1}{1 - d_r}\left(2 - \binom{d_r + 1}{2} - rd_r\right)$$
$$= \left(\frac{r + d_r - 1}{d_r - 1}\right)k + \frac{1}{1 - d_r}\left(2 - \binom{d_r + 1}{2} - rd_r\right).$$

Thus, provided $r \geqslant 2$, we know that if

$$n > \left(\frac{r + d_r - 1}{d_r - 1}\right)k + \frac{1}{1 - d_r}\left(2 - \binom{d_r + 1}{2} - rd_r\right), \tag{3}$$

then $k_r \leqslant k + d_r - 1$ and so (2) holds.

So, we have a lower bound on n in terms of r and k (which depends on n and m) that implies that (2) holds. We would like to show that we can find an r so that we can remove r vertices from $L(n, m)$ and increase the number of homomorphisms into J. We will show that such an r exists provided $n > (1 + \alpha)k + c\sqrt{k}$ for some explicit constant c. In fact, we claim that if $n > (1 + \alpha)k + c\sqrt{k}$, then

$$r := \left\lceil \frac{n - (1 + \alpha)k + \sqrt{(n - (1 + \alpha)k)^2 - 96\alpha^2(3 + \frac{1}{\alpha})k}}{8(3 + \frac{1}{\alpha})} \right\rceil \geqslant 2,$$

and we can remove r vertices from $L(n, m)$ and increase the number of homomorphisms into J.

After a bit of work and using the fact that $d_r = r/\alpha$, we have that (3) is equivalent to

$$n - (1 + \alpha)k > \frac{2\alpha^2 k + (2 + \frac{1}{\alpha})r^2 + r}{2(r - \alpha)}.$$

We will find a stronger inequality that implies this (and is cleaner) provided $r \geqslant 2$. In what follows, we will use the fact that if $r \geqslant 2$, then $r - \alpha \geqslant r/8$ several times. Thus,

$$\frac{2\alpha^2 k + (2 + \frac{1}{\alpha})r^2 + r}{2(r - \alpha)} < \frac{\alpha^2 k}{r - \alpha} + \frac{(3 + 1/\alpha)r^2}{2(r - \alpha)}$$
$$< \frac{8\alpha^2 k}{r} + \frac{4(3 + 1/\alpha)r^2}{r}$$
$$= \frac{8\alpha^2 k}{r} + 4(3 + 1/\alpha).$$

And so, we know that (3) is implied by $r \geqslant 2$ and

$$n - (1 + \alpha)k > 4\left(3 + \frac{1}{\alpha}\right)r + \frac{8\alpha^2 k}{r}.$$

Multiplying through by r yields a quadratic inequality in r and so the above is equivalent to $r \geqslant 2$ being between its two roots

$$\frac{n - (1+\alpha)k \pm \sqrt{(n-(1+\alpha)k)^2 - 96\alpha^2(3 + \frac{1}{\alpha})k}}{8(3 + \frac{1}{\alpha})}. \tag{4}$$

We are left to show that there is an $r \geqslant 2$ between these two roots. We show that this occurs when $n > (1+\alpha)k + c\sqrt{k}$ where $c = \sqrt{384(3 + 1/\alpha)\alpha^2}$. If $n > (1+\alpha)k + c\sqrt{k}$, then $(n - (1+\alpha)k)^2 > c^2 k$ which in turn implies that

$$96(3 + 1/\alpha)\alpha^2 k < \frac{1}{4}(n - (1+\alpha)k)^2. \tag{5}$$

Also, note that the smaller root in (4) is positive. Thus, if we can show that the gap between the two roots is at least two, then there is an $r \geqslant 2$ that does the job for us, i.e.,

$$\frac{\sqrt{(n - (1+\alpha)k)^2 - 96(3 + 1/\alpha)\alpha^2 k}}{4(3 + 1/\alpha)} > 2.$$

Using (5), the above is implied if

$$\frac{\frac{\sqrt{3}}{2}[n - (1+\alpha)k]}{4(3 + \frac{1}{\alpha})} > 2, \quad \text{or} \quad n - (1+\alpha)k > \frac{16(3 + \frac{1}{\alpha})}{\sqrt{3}}.$$

But we know that $n - (1+\alpha)k > c\sqrt{k} \geqslant c$ and one can check that $c > \frac{16(3+1/\alpha)}{\sqrt{3}}$. So, there is an $r \geqslant 2$ between the two roots in (4) and thus, we have shown that if $n > (1+\alpha)k + c\sqrt{k}$, we can remove r (as defined above) vertices from $L(n,m)$ and increase the number of homomorphisms into J.

We are left to show what bound on $\ell(m)$ this yields. Let $n = (1+\alpha)k + c\sqrt{k}$ so that we may not be able to remove r vertices and so $\ell(m) \leqslant n$. Then

$$m = \binom{k}{2} + k(n-k) + w \quad \Longleftrightarrow \quad f(k) := \left(\frac{1}{2} + \alpha\right)k^2 + ck^{3/2} - \frac{k}{2} - m + w = 0.$$

It is easy to check that f is continuous and increasing on $[0, \infty)$ and that $f(0) < 0$. Further, if $x = \sqrt{\frac{m}{\frac{1}{2}+\alpha}}$, then $f(x) > 0$, so, since k is the root of f, we know $k < x$. Thus, we have

$$\ell(m) \leqslant n$$
$$= (1+\alpha)k + c\sqrt{k}$$
$$\leqslant (1+\alpha)\sqrt{\frac{m}{\frac{1}{2} + \alpha}} + c\left(\frac{m}{\frac{1}{2} + \alpha}\right)^{1/4},$$

where the last step follows from the fact that $(1+\alpha)x + c\sqrt{x}$ is increasing for all $x > 0$. Thus, we have proved the upper bound.

The proof of the lower bound, while similar to that of the upper bound, is a bit different and so we include it for completeness. In this case, we add add r vertices to $L(n, m)$. Let $L(n, m) = L(n, r, w)$ and define k'_r and w'_r to be the integers for which $L(n + r, m) = L(n + r, k'_r, w'_r)$. In this case, note that $k \geqslant k'_r$. Again, let $d_r = r \log_2(3/2)$. We compare $j(L(n, m) \cup E_r)$ to $j(L(n + r, m))$.

$$
\begin{aligned}
3^r i(L(n, k, w)) &= 3^r (2^{n-k-1} + 2^{n-k-w-1} + k) \\
&\leqslant 3^r (2^{n-k} + k) \\
&= 3^r \cdot 2^{n-k} \cdot 3^{-r} 2^r 2^{r \log_2(3/2)} + 3^r \cdot k \\
&= 2^{n+r-(k-d_r-2)-2} + 3^r \cdot k \\
&\leqslant 2^{n+r-k'_r-2} + 3^r \cdot k \\
&< 2^{n+r-k'_r-2} + 2^{n+r-k'_r-2} \\
&\leqslant 2^{n+r-k'_r-1} + 2^{n+r-k'_r-w'_r-1} + k'_r \\
&= i(L(n + r, k'_r, w'_r)).
\end{aligned}
\tag{6}
$$

Note that the fifth step holds if $k'_r \geqslant k - d_r - 2$ and the sixth step holds provided $3^r \cdot k < 2^{n+r-k'_r-2}$. We first discuss the former as it is similar to the upper bound and then deal with the latter.

We would like to figure out when $k'_r \geqslant k - d_r - 2$. So we assume the contrary, so that $k'_r \leqslant k - d_r - 3$. Then, as above, we have

$$
\begin{aligned}
e(L(n, k, 0)) &\leqslant e(L(n, k, w)) \\
&= e(L(n + r, k'_r, w'_r)) \\
&\leqslant e(L(n + r, k - d_r - 3, w'_r)) \\
&\leqslant e(L(n + r, k - d_r - 3, n + r - k + d_r + 2)).
\end{aligned}
$$

This yields that

$$
\binom{k}{2} + k(n - k) \leqslant \binom{k - d_r - 3}{2} + (k - d_r - 3)(n + r - k + d_r + 3) + n + r - k + d_r + 2,
$$

which can be rearranged to give

$$
(d_r + 2)n \leqslant (r + d_r + 2)k - \frac{d_r^2}{2} - \frac{3d_r}{2} - 1,
$$

and so

$$
\begin{aligned}
n &\leqslant \left(\frac{r + d_r + 2}{d_r + 2} \right) k - \frac{d_r^2 + 3d_r + 2}{2(d_r + 2)} \\
&= \left(\frac{r + d_r + 2}{d_r + 2} \right) k - \frac{d_r + 1}{2}.
\end{aligned}
$$

Thus, we know that $k'_r \geqslant k - d_r - 2$, and so we can add r vertices, provided

$$n > \left(\frac{r + d_r + 2}{d_r + 2}\right) k - \frac{d_r + 1}{2} = (1 + \alpha)k - \left(\frac{2\alpha^2}{r + 2\alpha}\right) k - \frac{r + \alpha}{2\alpha},$$

or

$$(1 + \alpha)k - n > \left(\frac{2\alpha^2}{r + 2\alpha}\right) k + \frac{r + \alpha}{2\alpha}.$$

If $r \geqslant 2$, then we have that $r + \alpha \leqslant 2r$, and so the above is implied by

$$(1 + \alpha)k - n > \left(\frac{2\alpha^2}{r + 2\alpha}\right) k + \frac{r}{\alpha} \quad \Longleftrightarrow \quad (r + 2\alpha)((1 + \alpha)k - n) > 2\alpha^2 k + \frac{r}{\alpha}(r + 2\alpha).$$

Since $r + 2\alpha \leqslant 3r$ if $r \geqslant 2$, the above is implied by (provided $r \geqslant 2$)

$$r[(1 + \alpha)k - n] > 2\alpha^2 k + \frac{3}{\alpha}r^2.$$

Thus, we can add r vertices provided there is an integer $r \geqslant 2$ between the two roots

$$\frac{(1 + \alpha)k - n \pm \sqrt{((1 + \alpha)k - n)^2 - 24\alpha k}}{6/\alpha}. \tag{7}$$

We claim this is true provided that $n < (1 + \alpha)k - d\sqrt{k}$ where $d = \sqrt{92\alpha}$. As in the proof of the upper bound, we first note that the smaller root is positive and show that the gap between the two roots is greater than two, i.e.,

$$\frac{\sqrt{((1 + \alpha)k - n)^2 - 24\alpha k}}{3/\alpha} > 2.$$

Using the fact that $n < (1 + \alpha)k - d\sqrt{k}$, we know that $24\alpha k < \frac{1}{4}((1 + \alpha)k - n)^2$, and so the above is implied by

$$(1 + \alpha)k - n > \frac{12}{\alpha\sqrt{3}}.$$

But $(1 + \alpha)k - n > d\sqrt{k} \geqslant d$ and one can check that $d > 12/(\alpha\sqrt{3})$. Thus, an $r \geqslant 2$ between the roots in (7) exists provided $n < (1 + \alpha)k - d\sqrt{k}$.

Now we must return to our earlier question of whether $3^r \cdot k < 2^{n+r-k'_r-2}$. This inequality is equivalent to $r \log_2(3) + \log_2(k) + 2 < n + r - k'_r$ which, in turn, is equivalent to $r \log_2(3) - r < n - k'_r - \log_2(k) - 2$. Thus, the desired inequality holds provided

$$r < \frac{n - k'_r - \log_2(k) - 2}{\log_2(3) - 1}.$$

Noting that $k'_r \leqslant k$ and $\log_2(k) \leqslant k$, it is sufficient to prove that

$$r < \frac{n - 2k - 2}{\log_2(3) - 1}.$$

Thus, we are once again faced with proving the existence of a r obeying an inequality. In this case, we once again show that the gap between this upper bound on r and the smaller root in (7) is greater than 2. Since the smaller root in (7) is positive, this guarantees the existence of an $r \geqslant 2$ satisfying all of the requisite inequalities. So, we want to show that

$$\frac{n - 2k - 2}{\log_2(3) - 1} - \frac{(1+\alpha)k - n - \sqrt{((1+\alpha)k - n)^2 - 24\alpha k}}{6/\alpha} > 2.$$

Since $24\alpha k \leqslant \frac{1}{4}((1+\alpha)k - n)^2$, this is implied by

$$\frac{n - 2k - 2}{\log_2(3) - 1} - \frac{(2 - \sqrt{3})((1+\alpha)k - n)}{2(3/\alpha)} > 2.$$

Rearranging, we see that this is equivalent to

$$\left(\frac{1}{\log_2(3) - 1} + \frac{\alpha(2 - \sqrt{3})}{6}\right) n > \left(\frac{\alpha(2 - \sqrt{3})(1 + \alpha)}{6} + \frac{2}{\log_2(3)}\right) k + \frac{2}{\log_2(3)} + 2.$$

Roughly speaking, this is $1.786n > 1.469k + 3.262$ or $n > .822k + 1.828$, which is certainly true for $n \geqslant 10$ since $n > k$. The smallest m for which we can guarantee that $n \geqslant 10$ is $m = \binom{10}{2}$. Thus, the existence of a suitable r is guaranteed provided $n \leqslant (1+\alpha)k - d\sqrt{k}$ and $m \geqslant \binom{10}{2}$.

We need to show what bound on $\ell(m)$ this gives and so let $n = (1+\alpha)k - d\sqrt{k}$ so that we may not be able to add r vertices. Thus, $\ell(m) \geqslant n$. So

$$m = \binom{k}{2} + k(n - k) + w \quad \Longleftrightarrow \quad g(k) := \left(\frac{1}{2} + \alpha\right)k^2 - ck^{3/2} - \frac{k}{2} - m + w = 0.$$

In this case, $g(0) < 0$ and there is one critical point $x_c > 0$ with $g(x)$ decreasing before x_c and increasing after it. So, g has a unique root k. If $x = \sqrt{\frac{m}{1+\alpha}}$, then $g(x) < 0$ so $x < k$. Thus,

$$\begin{aligned}
\ell(m) &\geqslant n \\
&= (1+\alpha)k - d\sqrt{k} \\
&\geqslant (1+\alpha)\sqrt{\frac{m}{\frac{1}{2} + \alpha}} - d\left(\frac{m}{1+\alpha}\right)^{1/4},
\end{aligned}$$

where the last step follows from the fact the $(1+\alpha)x - d\sqrt{x}$ is a increasing function of x for $x \geqslant 1/(2d(1+\alpha))^2$. Thus, we have proved the lower bound. $\qquad \square$

5 Future directions

While we have partially solved the question of maximizing the number of homomorphisms into J from $\mathcal{G}(n, m)$, one might hope for a more detailed answer that is able to account

Code of H	H-extremal graph
0000 or 1111	Any
0001, 0011, or 0111	Colex
0010	Lex with isolates
0101	One of the extremal graphs from [6] with isolates
0110	Lex with isolates
1000, 1100, or 1110	Lex
1001	??
1010	??
1011	??
1101	??

Table 2: H-extremal graphs for all loop-threshold graphs H on four vertices

for the non-monotonicity of the lex component size. In addition to this, many of the questions regarding maximization of homomorphisms into loop-threshold graphs H of order four remain open as well. While some are trivial consequences of the observations in Section 2, others are similar to J and others seem to be completely open. The current status of these problems is given in Table 2.

Acknowledgements

References

[1] James Alexander, Jonathan Cutler, and Tim Mink, *Independent sets in graphs with given minimum degree*, Electron. J. Combin. **19** (2012), no. 3, #P37.

[2] James Alexander and Tim Mink, *A new method for enumerating independent sets of a fixed size in general graphs*, J. Graph Theory **81** (2016), no. 1, 57–72.

[3] Graham R. Brightwell and Peter Winkler, *Graph homomorphisms and phase transitions*, J. Combin. Theory Ser. B **77** (1999), no. 2, 221–262.

[4] Jonathan Cutler and A. J. Radcliffe, *Extremal graphs for homomorphisms*, J. Graph Theory **67** (2011), 261–284.

[5] Jonathan Cutler and A. J. Radcliffe, *Extremal problems for independent set enumeration*, Electron. J. Combin. **18** (2011), no. 1, #R169.

[6] Jonathan Cutler and A. J. Radcliffe, *Extremal graphs for homomorphisms II*, J. Graph Theory **76** (2014), no. 1, 42–59.

[7] Jonathan Cutler and A. J. Radcliffe, *The maximum number of complete subgraphs in a graph with given maximum degree*, J. Combin. Theory Ser. B **104** (2014), 60–71.

[8] Jonathan Cutler and A. J. Radcliffe, *The maximum number of complete subgraphs of fixed size in a graph with given maximum degree*, J. Graph Theory **84** (2017), no. 2, 134–145.

[9] John Engbers and David Galvin, *Counting independent sets of a fixed size in graphs with a given minimum degree*, J. Graph Theory **76** (2014), no. 2, 149–168.

[10] David Galvin, *Two problems on independent sets in graphs*, Discrete Math. **311** (2011), 2105–2112.

[11] David Galvin, *Maximizing H-colorings of a regular graph*, J. Graph Theory **73** (2013), no. 1, 66–84.

[12] David Galvin and Prasad Tetali, *On weighted graph homomorphisms*, Graphs, morphisms and statistical physics, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., vol. 63, Amer. Math. Soc., Providence, RI, 2004, pp. 97–104.

[13] Wenying Gan, Po-Shen Loh, and Benny Sudakov, *Maximizing the number of independent sets of a fixed size*, Combin. Probab. Comput. **24** (2015), no. 3, 521–527.

[14] Jeff Kahn, *An entropy approach to the hard-core model on bipartite graphs*, Combin. Probab. Comput. **10** (2001), no. 3, 219–237.

[15] G. Katona, *A theorem of finite sets*, Theory of graphs (Proc. Colloq., Tihany, 1966), Academic Press, New York, 1968, pp. 187–207.

[16] Joseph B. Kruskal, *The number of simplices in a complex*, Mathematical optimization techniques, Univ. of California Press, Berkeley, Calif., 1963, pp. 251–278.

[17] N. V. R. Mahadev and U. N. Peled, *Threshold graphs and related topics*, Annals of Discrete Mathematics, vol. 56, North-Holland Publishing Co., Amsterdam, 1995.

[18] David R. Wood, *On the maximum number of cliques in a graph*, Graphs Combin. **23** (2007), no. 3, 337–352.

[19] Yufei Zhao, *The number of independent sets in a regular graph*, Combin. Probab. Comput. **19** (2010), no. 2, 315–320.

[20] Yufei Zhao, *Extremal regular graphs: Independent sets and graph homomorphisms*, The American Mathematical Monthly **124** (2017), no. 9, 827–843.