

Semi-inducibility of 4-vertex graphs

Levente Bodnár* and Oleg Pikhurko†

Mathematics Institute and DIMAP

University of Warwick

Coventry CV4 7AL, UK

October 29, 2025

Abstract

For a graph H whose edges are coloured blue or red, the *H-semi-inducibility problem* asks for the maximum, over all graphs G of given order n , of the number of injections from the vertex set of H into the vertex set of G that send red (resp. blue) edges of H to edges (resp. non-edges) of G . We consider all possible 4-vertex non-complete graphs H and essentially resolve all remaining cases except when H is the 3-edge path coloured blue-blue-red in this order (or is equivalent to this case). Some of our proofs are computer-generated, using the flag algebra method of Razborov.

1 Introduction

Basit, Granet, Horsley, Kündgen and Staden [6] introduced the following problem for a given blue-red edge-coloured graph H with κ vertices. For an (uncoloured) graph G , let $\Lambda_H(G)$ be the number of *embeddings* of H into G , that is, injections $V(H) \rightarrow V(G)$ that map red (resp. blue) edges of H to edges (resp. non-edges) of G . If G has $n \geq \kappa$ vertices then we also consider the normalised version

$$\lambda_H(G) := \frac{\Lambda_H(G)}{n^{(\kappa)}},$$

where $n^{(\kappa)} := n(n-1) \dots (n-\kappa+1)$ is the κ -th falling power of n . Thus $\lambda_H(G)$ is the probability that a random injection $V(H) \rightarrow V(G)$ is an embedding of H into G . The *H-semi-inducibility problem* asks for $\Lambda_H(n)$, the maximum of $\Lambda_H(G)$ over all graphs G of order n (i.e. with n vertices) or, equivalently, for its normalised version $\lambda_H(n) := \Lambda_H(n)/n^{(\kappa)}$ for $n \geq \kappa$. It is easy to show (see [6, Proposition 3.1]) that the limit

$$\lambda_H := \lim_{n \rightarrow \infty} \lambda_H(n)$$

exists. We call λ_H the *semi-inducibility constant* of H .

*Email: bodnalev@gmail.com

†Email: O.Pikhurko@warwick.ac.uk

Note that our parameter $\Lambda_H(G)$ is slightly different from the one used in [6] where the authors count the pairs of sets of edges and non-edges in G which, when coloured red and blue respectively, form a copy of H . Of course, each such pair corresponds to the same number of embeddings of H into G (which is the number of colour-preserving automorphisms of H), so it is easy to change between these two parameters. We use $\Lambda_H(G)$ as it admits the natural probabilistic version $\lambda_H(G)$, which is convenient in some arguments.

If H is a colouring of the complete graph on $V(H)$ then the H -semi-inducibility problem amounts to the well-known and actively studied inducibility problem (that was introduced by Pippenger and Golumbic [22] in 1975) for the red subgraph of H . This was the main motivation in [6] for introducing this question.

Various results on the semi-inducibility problem have been proved in [6, 7, 10, 11]; we refer the reader to these papers for further details.

Here we systematically analyse the case when H has 4 vertices (as all cases with $v(H) \leq 3$ are trivial). We exclude the cases when H is complete (as then we obtain the well-studied graph inducibility problem), or H has no edges of some colour (as then trivially $\lambda_H(n) = 1$), or H has an isolated vertex (as this reduces to a 3-vertex case). Since the extremal function $\lambda_H(n)$ does not change when we exchange the two colours in H , it is enough to consider only one graph from each such pair. With the above conventions and exclusions, there are, up to isomorphism, 18 possible 4-vertex 2-coloured graphs H to consider (some of which have been previously resolved) and these all are listed in Table 1.

In brief, our contribution is to resolve all but one of the remaining cases in the table.

This project initially started as our attempt to see what the flag algebra method of Razborov would give for this new extremal problem. The method turned out to be highly successful in proving upper bounds on the semi-inducibility constant, giving the correct value of it for all but one non-complete 4-vertex graphs H . These upper bounds on $\lambda_H(H)$ are recorded in Table 1. The table also gives a construction giving the matching lower bound on λ_{H_i} , where we denote the complete partite graph with m parts of sizes $n_0, \dots, n_{m-1}$ by $T_{n_0, \dots, n_{m-1}}$ (and we omit ceiling and floor signs when these are not essential). Although flag algebras work with a limit version of the problem as $n \rightarrow \infty$ and directly give asymptotic results only, we complemented these calculations with various stability-type arguments to get more precise structural results about (almost) extremal graphs.

In six of the resolved cases H_i , namely for $i = 4, 5, 6, 9, 10, 11$, we proved that, for all sufficiently large n , every extremal graph G is a blow-up of some small m -clique K_m (with $m = m(i)$ being at most 5 in each of these cases). Thus one can consider the problem as combinatorially solved for large n : the determination of the exact value of $\Lambda_{H_i}(n)$ (resp. the set of extremal graphs) amounts to maximising an explicit polynomial over non-negative integers $n_0, \dots, n_{m-1}$ that sum up to n (resp. describing all optimal m -tuples $(n_0, \dots, n_{m-1})$). In fact, we prove in each of these cases that there is a unique m -tuple of non-negative reals $x_0 \geq \dots \geq x_{m-1}$ summing up to 1 such that every almost extremal graph of order $n \rightarrow \infty$ can be made into $T_{x_0 n, \dots, x_{m-1} n}$ by changing at most $o(n^2)$ adjacencies. Thus all almost extremal graphs are close to each other in the edit distance.

The remaining studied cases, namely H_0 , H_1 and H_{15} , although exhibiting some variety of extremal structures, are rather straightforward and we have been able to obtain an essentially full picture in these cases.

Finally, let us discuss the two cases where we could not determine the semi-inducibility constant via the plain flag algebra method.

One of these two cases is H_5 which is the red 3-edge star $T_{3,1}$ with one blue edge added. Here a lower bound

$$\lambda_{H_5} \geq \frac{23637 - 513\sqrt{57}}{10976} \quad (1)$$

follows by taking the complete 5-partite graphs with part sizes $(\alpha n, \alpha n, \alpha n, \alpha n, (1-4\alpha)n)$, where $\alpha := (13 - \sqrt{57})/56$. The plain flag algebra method seems to return the same numerical value as in the right-hand side of (1) but we have not been able to convert it to an exact-arithmetic proof. Instead, we observe (in Lemma 3.22 here) that the symmetrisation method of Brown and Sidorenko [8] can be applied to this case to show that, for every $n \geq 4$, at least one extremal graph for $\lambda_{H_5}(n)$ is complete partite. Working with complete partite graphs greatly reduces the scale of flag algebra calculations and we were able to produce an exact-arithmetic certificate with entries in the field $\mathbb{Q}(\alpha)$, proving that (1) is equality.

The only case in Table 1 with λ_H still unknown is H_3 , which is the 3-edge path coloured blue-blue-red, say visiting vertices 0, 1, 2, 3 in this order. Here we have the following construction G of order $n \rightarrow \infty$: take a disjoint union of an almost γn -regular graph on a vertex set B with $|B| = (\beta + o(1))n$ and a clique on a set A with $|A| = n - |B|$. The number of embeddings of H_3 into G can be approximated as follows: we sum over all maps f of $1, 2 \in V(H)$ to non-adjacent vertices $f(1), f(2) \in V(G)$ of the product of the non-degree of $f(1)$ and the degree of $f(2)$, as these are the number of choices for $f(0)$ and $f(3)$ within factor $1 + o(1)$. Thus $\lambda_H(G)$, within additive $o(1)$, is

$$p(\beta, \gamma) := \beta(1 - \beta) \cdot (1 - \gamma)(1 - \beta) + (1 - \beta)\beta \cdot \gamma\beta + \beta(\beta - \gamma) \cdot (1 - \gamma)\gamma, \quad (2)$$

where the 3 products give the density of embeddings $f : H \rightarrow G$ with the pair $(f(1), f(2))$ belonging to $B \times A$, $A \times B$ and $B \times B$ respectively. Routine optimisation shows that the maximum of p over $0 \leq \gamma \leq \beta \leq 1$ is attained when $\beta = 0.39829918\dots$ and $\gamma = 0.28158008\dots$ satisfy

$$\beta^5 - \frac{10}{3}\beta^4 + \frac{2521}{576}\beta^3 - \frac{407}{144}\beta^2 + \frac{43}{48}\beta - \frac{1}{9} = 0, \quad (3)$$

$$\gamma^4 - \frac{37}{16}\gamma^3 + \frac{57}{32}\gamma^2 - \frac{9}{16}\gamma + \frac{1}{16} = 0, \quad (4)$$

giving 0.1500834073... as the maximal value. Here, the floating-point value returned by flag algebras seems to match this lower bound. Unfortunately, we have not been able to round it, so we leave this as an open problem.

Conjecture 1.1 *For the 4-vertex graph H_3 , which has blue edges $\{0, 1\}$ and $\{1, 2\}$, and a red edge $\{2, 3\}$, it holds that $\lambda_{H_3} = p(\beta, \gamma)$, where p is defined in (2) while β and γ are the appropriate roots of (3) and (4) respectively.*

i	H_i	λ_{H_i}	Construction	Reference
0		$\frac{1}{4}$	Graphs of edge density $\frac{1}{2}$	Section 3.2
1		$\frac{4}{27}$	$\frac{n}{3}$ -regular graphs	Section 3.3
2		$\frac{4}{27}$	$\frac{n}{3}$ -regular graphs	[7, Theorem 6.2], [10, Theorem 1.2]
3		$\leq 0.1500834092\dots$	$\geq 0.1500834073\dots$	Conjecture 1.1
4		$\frac{4}{27}$	$T_{n/3, n/3, n/3}$	Section 3.4
5		$\frac{171\sqrt{57}+7879}{43904}$	$T_{\alpha n, \alpha n, \alpha n, \alpha n, (1-4\alpha)n}$	Section 3.8
6		$\frac{1}{8}$	$T_{n/2, n/2}$	Section 3.7
7		$\frac{27}{256}$	$\frac{1}{4}$ -quasirandom graphs	[6, Theorem 1.9]
8		$\frac{4}{27}$	$T_{n/3, n/3, n/3}$	[6, Theorem 8.1]
9		$\frac{12}{125}$	$T_{n/5, n/5, n/5, n/5, n/5}$	Section 3.6
10		$\frac{1}{8}$	$T_{n/2, n/2}$	Section 3.6
11		$\frac{1}{8}$	$T_{n/2, n/2}$	Section 3.6
12		$\frac{1}{8}$	$T_{n/2, n/2}$	[6, Theorem 1.7]
13		$\frac{1}{8}$	$T_{n/2, n/2}$	[6, Theorem 1.6]
14		$\frac{1}{8}$	$T_{n/2, n/2}$	[6, Section 8.1]
15		$\frac{1}{27}$	Triangle-free $\frac{n}{3}$ -regular graphs	Section 3.5
16		$\frac{1}{8}$	$T_{n/2, n/2}$	[6, Section 8.1]
17		$\frac{1}{8}$	$T_{n/2, n/2}$	[6, Section 8.1]

Table 1: The values of the H -semi-inducibility constant for 4-vertex non-complete graphs H .

This paper is organised as follows. Section 2 contains various definitions and auxiliary results needed in the paper. Then we discuss each individual graph H_i in Section 3 (with Table 1 also containing references to the corresponding sub-sections of Section 3). First, we deal with those cases where we were able to prove the results without using flag algebras, which are the graphs H_0 , H_1 , H_4 and H_{15} . Among the remaining cases, we leave H_5 until the very end since it requires recalling some basic results from the theory of limits of complete partite graphs.

We provide a computer-generated certificate (and the scripts that can be used to generate or verify it) for each upper bound stated in Table 1, including also the cases that have been resolved elsewhere. On the other hand, when dealing with individual cases in Section 3, we

replaced computer-generated claims by human-readable proofs, whenever we could find a proof that had reasonable length and gave a better understanding of the problem.

Independently of us, József Balogh, Bernard Lidický, Dhruv Mubayi, Florian Pfender and Jan Volec (personal communication) also studied the semi-inducibility problem using flag algebras.

2 Preliminaries

In this section, we present some definitions and auxiliary results that will be needed in the paper.

Let $\mathbb{R}$ denote the set of reals. Let $\mathbb{N}$ denote the set of non-negative integers and, for $n \in \mathbb{N}$, we define $[n] := \{0, \dots, n-1\}$. Note that we start indexing from 0, merely to be consistent with the same convention as in our code. If the meaning is clear, we may abbreviate an unordered pair $\{u, w\}$ as uw , including the case when u and w are single-digit numbers. For a set X and an integer $\kappa \geq 0$, the set of all κ -subsets of X is denoted by $\binom{X}{\kappa}$. Recall that $n^{(\kappa)} := \prod_{i=0}^{\kappa-1} (n-i)$ denotes the κ -th falling power of n . We may omit ceiling/floor signs when they are not essential.

For $n, s \in \mathbb{N}$, let $\pi_s(n)$ be the maximum product of s non-negative integers that sum to n . Clearly, this is attained if and only if the s summands are nearly equal. So we can equivalently define

$$\pi_s(n) := \prod_{i=0}^{s-1} \left\lfloor \frac{n+i}{s} \right\rfloor. \quad (5)$$

2.1 Graphs

Let K_n denote the complete graph with n vertices, P_n denote the n -vertex path, and $T_{n_0, \dots, n_{m-1}}$ denote the complete m -partite graph with parts of sizes $n_0, \dots, n_{m-1}$. For integers m and n , the *Turán graph* T_n^m is the complete m -partite graph on $[n]$ with almost equal parts, that is, $T_n^m \cong T_{n_0, \dots, n_{m-1}}$ with $n_i = \lfloor (n+i)/m \rfloor$ for $i \in [m]$. Also, $F + G$ denotes the union of vertex-disjoint copies of graphs F and G .

For a graph $G = (V(G), E(G))$, its *order* is $v(G) := |V(G)|$. The *neighbourhood* of a vertex $u \in V(G)$ is the set

$$\Gamma_G(u) := \{w \in V(G) : \{u, w\} \in E(G)\}.$$

The *degree* of u is $\deg_G(u) := |\Gamma_G(u)|$. Also, the *complement* of a graph G is

$$\overline{G} := \left(V(G), \binom{V(G)}{2} \setminus E(G) \right)$$

and, for $X \subseteq V(G)$, the subgraph *induced* by X is

$$G[X] := (X, \{uw \in E(G) : u, w \in X\}).$$

For distinct vertices $u, w \in V(G)$, let $G \oplus uw$ denote the graph obtained from G by *flipping* (i.e. changing) the adjacency between u and w .

The *edit distance* $\delta_{\text{edit}}(G, G')$ between two graphs G and G' of the same order is the minimum value of $|f(E(G)) \Delta E(G')|$ over all bijections $f : V(G) \rightarrow V(G')$; in other words, it is the smallest number of *edits* (changes in adjacency) we have to do in one graph to make it isomorphic to the other. The distance from a graph G to a graph family $\mathcal{G}$ is

$$\delta_{\text{edit}}(G, \mathcal{G}) := \min\{\delta_{\text{edit}}(G, G') : G' \in \mathcal{G}, v(G') = v(G)\}.$$

For $m \in \mathbb{N}$, let $\mathcal{F}_m$ denote the family of graphs of order m consisting of one representative from each isomorphism class. For graphs F and G with κ and n vertices respectively, let $P(F, G)$ be the number of κ -subsets $X \subseteq V(G)$ that induce a subgraph isomorphic to F in G . If $\kappa \leq n$ then we define $p(F, G) := P(F, G)/\binom{n}{\kappa}$, calling it the (*induced*) *density* of F in G .

For a graph B on $[m]$ and pairwise disjoint sets $V_0, \dots, V_{m-1}$ (some possibly empty), the *blowup* $B(V_0, \dots, V_{m-1})$ of B is the graph on $V = \cup_{i=0}^{m-1} V_i$ where $x \in V_i$ and $y \in V_j$ are adjacent if and only if $\{i, j\} \in E(B)$. In particular, each part V_i is an independent set consisting of *clones*, that is, (non-adjacent) vertices having the same sets of neighbours. Let $B()$ denote the family of all blowups of B . We may call B the *base graph*.

2.2 Extremal λ_γ -problem

Given an integer $\kappa \geq 2$ and a function $\gamma : \mathcal{F}_\kappa \rightarrow \mathbb{R}$, we consider the following function on graphs:

$$\lambda_\gamma(G) := \sum_{F \in \mathcal{F}_\kappa} \gamma(F) p(F, G), \quad \text{for a graph } G \text{ with } v(G) \geq \kappa. \quad (6)$$

The corresponding extremal λ_γ -*problem* asks for the value of

$$\lambda_\gamma(n) := \max\{\lambda_H(G) : v(G) = n\}, \quad \text{for } n \geq \kappa. \quad (7)$$

Also, we define $\lambda_\gamma := \lim_{n \rightarrow \infty} \lambda_\gamma(n)$, where the existence of the limit follows from a simple averaging argument, see for example [21, Lemma 2.2].

The $(m-1)$ -dimensional simplex is

$$\mathbb{S}_m := \{(x_0, \dots, x_{m-1}) \in \mathbb{R}^m : x_0 + \dots + x_{m-1} = 1 \text{ and } \forall i \in [m] \ x_i \geq 0\}. \quad (8)$$

For $\mathbf{x} = (x_0, \dots, x_{m-1}) \in \mathbb{S}_m$ and a graph B on $[m]$, let $\lambda_\gamma(B(\mathbf{x}))$ be the limit as $n \rightarrow \infty$ of $\lambda_\gamma(B(V_0, \dots, V_{m-1}))$, where $|V_i| = (x_i + o(1))n$ for $i \in [m]$. This is a continuous function on $\mathbb{S}_m$ (in fact, a polynomial). Let

$$\lambda_\gamma(B()) := \sup\{\lambda_\gamma(B(\mathbf{x})) : \mathbf{x} \in \mathbb{S}_m\}.$$

By the compactness of $\mathbb{S}_m$ the supremum is attained by at least one $\mathbf{x} \in \mathbb{S}_m$; such vectors will be called (λ_γ, B) -*optimal*. The graph B is called λ_γ -*minimal* if for every graph B' obtained from B by removing a vertex, it holds that $\lambda_\gamma(B'()) < \lambda_\gamma(B())$. A standard compactness argument shows that B is λ_γ -minimal if and only if every (λ_γ, B) -optimal vector has all entries positive. Also, B is called λ_γ -*optimal* if $\lambda_\gamma = \lambda_\gamma(B())$, that is, we can attain the asymptotically optimal constant λ_γ by some blowups of B .

Let us pause and see how some of these definitions relate to the semi-inducibility problem for a 2-coloured graph H . Observe that the graph function λ_H can be represented as λ_γ : take $\kappa := v(H)$ and define the value of γ on $F \in \mathcal{F}_\kappa$ to be $\lambda_H(F)$ (which is the probability that a random bijection $V(H) \rightarrow V(F)$ gives an embedding of H into F). Whenever H is given, these will be the default values of κ and γ ; then the functions $\lambda_H(\cdot)$ and $\lambda_\gamma(\cdot)$ will be the same. Let us call a (not necessarily injective) map $f : V(H) \rightarrow V(B)$ a *homomorphism* from H to B and write it as $f : H \rightarrow B$ if the image of every red edge is an edge while every blue edge is sent to a non-edge or to two equal vertices of B ; equivalently, it is a possible assignment of the vertices of H to the parts under an embedding of H into a blowup of B . Using this notation, we have

$$\lambda_H(B(\mathbf{x})) = \sum_{f: H \rightarrow B} \prod_{i \in V(H)} x_{f(i)}, \quad \text{for } \mathbf{x} \in \mathbb{S}_m. \quad (9)$$

We will also need some definitions from [21].

A graph B on $[m]$ is λ_γ -*flip-averse* if there is $\delta > 0$ such that, for every blow-up $G = B(V_0, \dots, V_{m-1})$ with $n \geq 1/\delta$ vertices and $\lambda_\gamma(G) \geq \lambda_\gamma(B) - \delta$, it holds for any distinct vertices $u, w \in V(G)$ (possibly from the same part) that

$$\lambda_\gamma(G) - \lambda_\gamma(G \oplus uw) \geq \delta n^{-2}. \quad (10)$$

Informally, this definition states that if we flip any pair in a large (almost) optimal blowup of B then we decrease the objective function by the maximum possible amount in the order of magnitude. By compactness of $\mathbb{S}_m$, this can be equivalently re-stated in terms of the optimal vectors in $\mathbb{S}_m$. For example, B is λ_H -flip-averse if and only if, for every (λ_H, B) -optimal vector $\mathbf{x} \in \mathbb{S}_m$ and for every (not necessarily distinct) vertices $u, w \in V(B)$, it holds that

$$\sum_{i \in V(H)} \sum_{j \in V(H) \setminus \{i\}} \left(\sum_{\substack{f: H \rightarrow B \\ f(i)=u, f(j)=w}} \prod_{\ell \in V(H) \setminus \{i, j\}} x_{f(\ell)} - \sum_{\substack{f: H \rightarrow B \oplus uw \\ f(i)=u, f(j)=w}} \prod_{\ell \in V(H) \setminus \{i, j\}} x_{f(\ell)} \right) > 0. \quad (11)$$

If we know the set of optimal vectors and it is finite, then (11) can be checked on computer.

Next, we give the definition, also from [21], when an extremal problem is strict. Roughly speaking, it states that if we take a large optimal blowup G of B and attach a new vertex u then the normalised contribution of u to the objective function is equal to $\lambda_\gamma + o(1)$ (the value observed for vertices in G) only if u is almost a clone of an existing vertex. Formally, a graph B on $[m]$ is called λ_γ -*strict* if B is λ_γ -minimal and, for every (λ_γ, B) -optimal $\mathbf{x} \in \mathbb{S}_m$ and every vector $\mathbf{y} \in [0, 1]^m$, the polynomial $p(\mathbf{x}, \mathbf{y})$ defined below in (12) assumes value λ_γ only if there is $i \in [m]$ with each y_j being the indicator function of $ij \in E(B)$. To define the polynomial p , let B' be the blow-up of B with parts $V_i := \{i, i + m\}$ for $i \in [m]$ and let B'' be obtained from B' by adding $2m$ as a new vertex with $\Gamma_{B''}(2m) = [m]$. (In other words, we add a clone of each vertex of B and then add a new vertex adjacent to the original vertices of B .) We define

$$p(\mathbf{x}, \mathbf{y}) := \sum_{F \in \mathcal{F}_\kappa} \gamma(F) \sum_f \prod_{j \in V(F) \setminus f^{-1}(2m)} z_{f(j)}, \quad (12)$$

where

$$\mathbf{z} := (y_0 x_0, \dots, y_{m-1} x_{m-1}, (1 - y_0) x_0, \dots, (1 - y_{m-1}) x_{m-1})$$

and sum is taken over all (including non-injective) maps $f : V(F) \rightarrow V(B'')$ that send edges to edges and non-edges to non-edges while the pre-image of the special vertex $2m$ consists of exactly one vertex.

In the special case when $\lambda_\gamma = \lambda_H$, we have that

$$p(\mathbf{x}, \mathbf{y}) = \sum_{\substack{f: H \rightarrow B'' \\ |f^{-1}(2m)|=1}} \prod_{\substack{j \in V(H) \\ f(j) \neq 2m}} z_{f(j)}.$$

We refer the reader to Claim 3.8 to see an example where the strictness property is verified.

Checking strictness can also be performed by the package, if we already have all (λ_γ, B) -optimal vectors $\mathbf{x}$, as follows. Assume that none of these vectors has a zero entry (as otherwise the graph B is not λ_γ -minimal and thus is not λ_γ -strict by definition). For each (λ_γ, B) -optimal vector $\mathbf{x}$, we have to find all elements $\mathbf{y} \in [0, 1]^m$ for which the polynomial $p(\mathbf{x}, \mathbf{y})$ attains its maximum value λ_γ . We run over all 3^n choices of assigning to each y_i value 0, value 1 or leaving y_i free. For each choice, we use Buchberger's algorithm to check if there are vectors $\mathbf{y}$ as above such that $p(\mathbf{x}, \mathbf{y}) = \lambda_\gamma$, $\partial p(\mathbf{x}, \mathbf{y}) / \partial y_i = 0$ for each free variable y_i , $\mathbf{y} \in [0, 1]^m$ but $\mathbf{y}$ is not the $\{0, 1\}$ -valued vector encoding the B -adjacency of some $i \in [m]$.

For $\mathbf{x} \in \mathbb{S}_m$, we call the λ_γ -problem *Erdős–Simonovits* $(B, \mathbf{x})$ -stable if for every $\varepsilon > 0$ there are $\delta > 0$ and n_0 such that if G is a graph with $n \geq n_0$ vertices and $\lambda_\gamma(G) \geq \lambda_\gamma - \delta$ then G is $\varepsilon \binom{n}{2}$ -close in the edit distance to a blowup $B(V_0, \dots, V_{m-1})$ with $||V_i| - x_i n| < \varepsilon n$. Recall that the last statement means that there is a partition $V(G) = V_0 \cup \dots \cup V_{m-1}$ such that

$$|E(G) \Delta E(B(V_0, \dots, V_{m-1}))| \leq \varepsilon \binom{n}{2}. \quad (13)$$

This property is very useful as the first step towards characterising graphs of sufficiently large order n with $\lambda_\gamma(G) = \lambda_\gamma(n)$. This approach was pioneered by Erdős [13] and Simonovits [26].

Following [21], we call the λ_γ -problem *perfectly B-stable* for a graph B if there is $C > 0$ such that for every graph G of order $n \geq C$ there is a blowup G' of B such that

$$\delta_{\text{edit}}(G, G') \leq C (\lambda_\gamma(n) - \lambda_\gamma(G)) n^2. \quad (14)$$

In particular, it follows that, for all $n \geq C$, every order- n graph G with $\lambda_\gamma(G) = \lambda_\gamma(n)$ is a blowup of B . Perfect stability has been established for various extremal problems, see e.g. [15, 17–19, 21, 25].

Let us remark that the property of being perfect B -stable remains true if we add new vertices to B with arbitrary attachments. In all cases of this paper when we establish perfect stability, the listed graph B is λ_γ -minimal, that is, has no redundant vertices.

Since γ or H will be fixed and clear from the context, we may omit them from our notation. Asymptotic notation, such as $o(1)$, is taken with respect to $n \rightarrow \infty$ (where n is usually the order of the unknown graph G); the constants hidden in it may depend on κ and γ but not on any other parameters. We call a sequence of graphs $(G_n)_{n \in \mathbb{N}}$ with increasing orders *almost extremal* (resp. *almost c-regular*) if $\lambda_\gamma(G_n) = \lambda_\gamma + o(1)$ as $n \rightarrow \infty$ (resp. for every $\varepsilon > 0$ there is n_0 such that, for every $n \geq n_0$, at least $(1 - \varepsilon)v(G)$ vertices u of G satisfy $|\deg_G(u)/v(G) - c| \leq \varepsilon$).

2.3 Flag Algebras

Since the flag algebra approach of Razborov [23] is well established by now (and is described in detail in e.g. [5, 12, 16, 24]), we give only high-level description. Suppose we are given an integer κ and a function $\gamma : \mathcal{F}_\kappa \rightarrow \mathbb{R}$ and would like to upper bound the corresponding extremal function $\lambda_\gamma(n)$ defined in (7). A flag algebra proof of $\lambda_\gamma(n) \leq u + o(1)$ amounts to an identity that holds for every graph G of order $n \rightarrow \infty$ of the form

$$u - \sum_{F \in \mathcal{F}_N} \lambda_\gamma(F) p(F, G) = \text{SOS} + \sum_{F \in \mathcal{F}_N} c_F p(F, G) + o(1), \quad (15)$$

where $N \geq \kappa$ is an integer, each c_F is a non-negative real, called the *slack* of $F \in \mathcal{F}_N$ while SOS is a sum of squares (of certain rooted densities) which can also be equivalently written as $\sum_{F \in \mathcal{F}_N} a_F p(F, G) + o(1)$ for some reals a_F . (See e.g. (28) for an example of an SOS-expression in the flag algebra context.) The corresponding *certificate* lists u , N , the slacks c_F and the positive semi-definite matrices that define the SOS-term. All of these are independent of n , and are stored using exact arithmetic.

All presented certificates were generated by the SageMath package of the first author. This package is still under development. A short guide on how to install it and an overview of its current functionality can be found in the GitHub repository <https://github.com/bodnalev/sage>.

The scripts that we use to generate the certificates, verify the identity in (15) as well as any additional claimed properties can be found in a separate GitHub repository https://github.com/bodnalev/supplementary_files inside the folder `graph_semi_inducibility`, while the certificates themselves can be found in its sub-folder `certificates`. Alternatively, all these files can be found in the ancillary folder of the arXiv version of this paper. Each certificate starts with `semiind` plus the index as in Table 1. Alternatively, a reader can use their own verifier; a description of how the data are arranged in each certificate file can be found in the [readme.md](#) file.

3 Individual cases

Here we investigate individual graphs H_i from Table 1. First, we list those cases (namely H_0 , H_1 , H_4 and H_{15}) where we were able to prove all results without using flag algebras. Our proof for the graph H_5 requires some basics of the limit theory of complete partite graphs, so we postpone it to the end.

3.1 Stars and matchings

Here we provide two easy observations (that in particular apply to the graphs H_0 and H_1 of Table 1). They follow from the facts that $x_0 := \ell/(\ell+m)$ is the unique maximiser of $x^\ell(1-x)^m$ on the interval $[0, 1]$ and that, since we maximise a continuous function on a compact set, all $x \in [0, 1]$ that almost attain the maximum value must be close to x_0 .

Observation 3.1 *If H is the matching of size $\ell + m$ with ℓ red and m blue edges then*

$$\lambda_H = \frac{\ell^\ell m^m}{(\ell + m)^{\ell+m}},$$

and a large graph is almost extremal if and only if its edge density is $\ell/(\ell + m) + o(1)$. ■

Observation 3.2 *If H is the star $T_{\ell+m,1}$ with ℓ red and m blue edges then*

$$\lambda_H = \frac{\ell^\ell m^m}{(\ell + m)^{\ell+m}},$$

and a large graph is almost extremal if and only if it is almost $\ell/(\ell + m)$ -regular. ■

3.2 Exact result for H_0

Observation 3.1 gives that $\lambda_H(H_0) = 1/4$ and describes all almost extremal graphs. In this section, we determine the extremal function exactly. For integers n and $0 \leq e \leq \binom{n}{2}$, let the *quasi-clique* $QC(n, e)$ be obtained from the clique K_{k+1} by removing $\binom{k+1}{2} - e$ edges forming a star and then adding $n - k - 1$ isolated vertices, where k be the maximum integer with $\binom{k}{2} < e$. Clearly, $QC(n, e)$ has n vertices and e edges.

Theorem 3.3 *Let H denote H_0 , which is the graph on $[4]$ with a red edge 02 and a blue edge 13. Take any $n \geq 4$. Define $k_0 := \lfloor x_0 \rfloor$, where*

$$x_0 := \frac{3 + \sqrt{2n^2 - 10n + 13}}{2}. \quad (16)$$

Then $\Lambda_H(n) = \Lambda_H(QC(n, \binom{k_0+1}{2}))$. Moreover, if $2n^2 - 10n + 13$ is not a square then $QC(n, \binom{k_0+1}{2})$ and its complement are the only extremal graphs; otherwise there are at least $2(k_0 + 1)$ non-isomorphic extremal graphs.

Proof. The easy case $n = 4$ (when $\Lambda_H(n) = 12$ and there are exactly two extremal graphs) satisfies the theorem, so assume that $n \geq 5$.

First, let us prove the stated upper bound on $\Lambda_H(n)$. Take any extremal graph G with n vertices. Let e denote the number of edges in G . We have

$$\Lambda_H(G) = 4e \left(\binom{n}{2} - e \right) - 4 \sum_{u \in V(G)} \deg(u)(n - 1 - \deg(u)), \quad (17)$$

since every pair consisting of an edge uv and a non-edge wx gives 4 embeddings of H , except we have to exclude intersecting pairs (i.e. those with $\{u, v\} \cap \{w, x\} \neq \emptyset$). Thus, if we fix the number of edges of G then our objective is to maximise the sum of the squares of degrees. Ahlswede and Katona [2, Theorem 2] proved that, for every (n, e) , the maximum value is attained by $QC(n, e)$ or by the complement of $QC(n, \binom{n}{2} - e)$. Since $\Lambda_H(G) = \Lambda_H(\overline{G})$, let us assume that $G = QC(n, e)$. Let k is the maximum integer with $\binom{k}{2} < e$ and let $\ell := e - \binom{k}{2}$. Thus, apart

from isolated vertices, G can be viewed as obtained from the clique on a k -set A by adding ℓ edges, $1 \leq \ell \leq k$, all incident to the same vertex u outside of A . If we fix n and k then $\Lambda_H(G)$ is a linear function of ℓ with the coefficient at ℓ being

$$c_{n,k} := \binom{n-k-1}{2} + (n-k-1)(k-1) - \binom{k-1}{2} = \frac{n^2 - 5n}{2} - k^2 + 3k + 1. \quad (18)$$

We can assume that $\ell = k$, for otherwise $c_{n,k} = 0$ (since we can both increase and decrease ℓ) and thus $\Lambda_H(G)$ does not depend on ℓ . Thus $G = K_{k+1} + \overline{K}_{n-k-1}$. We have that

$$\Lambda_H(G) - \Lambda_H(K_k + \overline{K}_{n-k}) = k c_{n,k}.$$

The function $c_{n,k}$, for a real variable k , has a root at $k = x_0$, is negative for $k > x_0$ and positive for $0 < k < x_0$ (since $n \geq 5$). Thus $\Lambda_H(n) = \Lambda_H(G) \leq \Lambda_H(QC(n, \binom{k_0+1}{2}))$, proving the first claim of the theorem.

Now, let us investigate the structure of extremal graphs. Take any extremal graph G of order n . By (17), it also maximises $S(G) := \sum_{u \in V(G)} \deg^2(u)$ given its size $e = e(G)$. It is easy to see that G has either an isolated or a universal vertex. Indeed, if u is a vertex of the minimal degree and w is a vertex of the maximal degree and, say, $uw \in E(G)$ then $\deg(w) = n - 1$: if there existed a non-neighbour v of w then we could strictly increase $S(G)$ by replacing the edge uw by vw . Using a version of this argument, it can be derived via a suitable induction statement that G is a *threshold graph*, that is, it can be obtained from a single vertex by adding new isolated and universal vertices in some order. While the result of Ahlswede and Katona [2] shows that at least one graph that maximises $S(G)$ for given (n, e) has very simple structure, the task of describing all extremal graphs for this problem turns out to be rather complicated, even though it is enough to consider threshold graphs only. This task is addressed in [1, 9, 20]. In fact, there are six families that between them contain every extremal graph.

Let us return back to our problem. By taking complements, we can assume that G has an isolated vertex. Note that $c_{n,k}$ as a function of real $k \geq 0$ is zero only if $k = x_0$, where x_0 and $c_{n,k}$ and x_0 are defined in (16) and (18) respectively.

Suppose first that $2n^2 - 10n + 13$ is not a square. Then $c_{n,k}$ is non-zero for every integer $k \geq 0$. Thus the parameter ℓ the proof of the upper bound is necessarily k , that is, $e = \binom{k_0+1}{2}$. For such a special value of e (plus our assumption that G has an isolated vertex), Theorem 2.4 in [1] implies that G is isomorphic to $QC(n, \binom{k_0+1}{2})$, as desired.

Finally, suppose that $2n^2 - 10n + 13$ is a square. Then $x_0 = k_0$ is an integer and it holds that $k_0 \leq n - 2$ (by $n \geq 6$) and $c_{n,k_0} = 0$. Thus $QC(n, \binom{k_0}{2} + \ell)$ for $0 \leq \ell \leq k_0$ are all pairwise non-isomorphic extremal graphs (which, by containing an isolated vertex, are non-isomorphic to any of their complements). ■

Let us remark that there are infinitely many n such that $2n^2 - 10n + 13$ is a square of some integer d . The corresponding Diophantine equation can be re-written as $2d^2 - 1 = m^2$, using substitution $m = 2n - 5$. This is a special case of the well-studied negative Pell equation and an explicit description of all (infinitely many) pairs of integers (d, m) with $2d^2 - 1 = m^2$ can be found in e.g. [4, Page 144]. For such n (when we are free to choose any $0 \leq \ell \leq k_0$), there are

choices of ℓ when some other extremal graphs appear; we refer the reader to Cases 2.2 and 2.3 of [1, Theorem 2.4] for details.

3.3 Exact result for H_1

For H_1 , the exact function can be easily determined. Recall that $\pi_k(n)$, defined in (5), is the maximum product of k non-negative integers summing up to n .

Lemma 3.4 *Let $H := H_1$, which is the graph on [4] that consists of a red edge 01 and two blue edges 02 and 03. Take any integer $n \geq 4$.*

If n is congruent to 1, 2 or 4 modulo 6 then it holds that $\Lambda_{H_1}(n) = n\pi_3(2n-3)/2$ and a graph G of order n attains equality if and only if G is $\lfloor n/3 \rfloor$ -regular.

If $n = 6k$ for an integer k then $\Lambda_{H_1}(n) = n\pi_3(2n-3)/2 - 6k(4k-1)$ and a graph G of order n attains equality if and only if every vertex has degree $2k-1$ or $2k$.

If $n = 6k+3$ for an integer k then $\Lambda_{H_1}(n) = n\pi_3(2n-3)/2 - (6k+3)(4k+1)$ and a graph G of order n attains equality if and only if every vertex has degree $2k$ or $2k+1$.

Finally, if $n = 6k+5$ for an integer k then $\Lambda_{H_1}(n) = n\pi_3(2n-3)/2 - 8k - 4$ and a graph G of order n attains equality if and only if it has $n-1$ vertices of degree $2k+1$ and 1 vertex of degree $2k+2$.

Proof. Given an integer $n \geq 1$, let $r := \lfloor n/3 \rfloor$ and define

$$D(d) := \pi_3(2n-2) - 2d(n-1-d)(n-2-d), \quad \text{for } d \in \mathbb{N}.$$

Clearly, it holds that $D(d) \geq 0$ for any integer $d \in [n]$. Thus, for an n -vertex graph G , we have that

$$\Lambda_{H_1}(G) = \frac{1}{2} \sum_{u_0 \in V(G)} (\pi_3(2n-3) - D(\deg(u_0))) \leq \frac{n\pi_3(2n-3)}{2}. \quad (19)$$

If n is 0, 2 or 4 modulo 6 then, as it is routine to see, G attains equality in (19) if and only if it is r -regular (and at least one such graph exists since nr is even in each of these cases); thus the lemma holds.

Next, suppose that $n = 6k$. Here $2n-3 = 12k-3$. We have that $\pi_3(12k-3) = (4k-1)^3$. Furthermore, the product of three non-negative integers summing up to $12k-3$ is at most $(4k-2)(4k-1)4k = \pi_3(12k-3) - 4k+1$, with equality if and only if these numbers are $4k-2$, $4k-1$ and $4k$ in some order. Thus, for $d \in [n+1]$, we have that $D(d) \geq 4k-1$ with equality if and only if $2d$ makes a triple of consecutive integers with $6k-2-d$ and $6k-1-d$, that is, $d = 2k-1$ or $d = 2k$. Thus $\sum_{u_0 \in V(G)} D(\deg(u_0)) \geq n \cdot (4k-1)$, giving the stated upper bound on $\Lambda_{H_1}(G)$, with equality if and only if every degree is $2k-1$ or $2k$. Furthermore, the set of such graphs is non-empty (e.g. by the Erdős–Gallai theorem [14]).

The case $n = 6k+3$ is completely analogous to the previous one, so we move the remaining case $n = 6k+5$ with $k \in \mathbb{N}$. Here $2n-3 = 12k+7$ and $\pi_3(12k+7) = (4k+2)^2(4k+3)$. For

$d \in [n]$, the product $2d(6k+3-d)(6k+4-d)$ is equal to $\pi_3(12k+7)$ if and only if $2d = 6k+3-d$, that is, $d = 2k+1$. However, $(2k+1)n$ is odd so at least one vertex of G has degree different from $2k+1$. As it is routine to see, the minimum of $D(d)$ for $d \in [n] \setminus \{2k+1\}$ is $8k+4$ and it is attained if and only if $d = 2k+2$. All claims of the lemma in this case routinely follow from the above observations. ■

3.4 Graph H_4

Here we resolve the case of H_4 , with perfect stability.

Theorem 3.5 *Let H be H_4 , that is, the graph on $[4]$ with the red edge set $\{02, 03, 13\}$ and the blue edge set $\{01\}$. Then all of the following statements hold.*

1. $\lambda_H = 4/27$.
2. The semi-inducibility problem for H is Erdős-Simonovits $(K_3, (1/3, 1/3, 1/3))$ -stable.
3. The semi-inducibility problem for H is perfectly K_3 -stable.

Proof. The 3-partite Turán graphs T_n^3 with $n \rightarrow \infty$ show that $\lambda_H \geq 4/27$. To prove the matching upper bound, take an arbitrary graph G with $n \geq 4$ vertices. By considering where $0 \in V(H)$ is mapped, we derive that

$$\Lambda_H(G) \leq \sum_{u_0 \in V(G)} (\deg(u_0))^2 \cdot (n-1-\deg(u_0)) \leq \sum_{u_0 \in V(G)} \frac{\pi_3(2n-2)}{2} = \frac{n \pi_3(2n-2)}{2}. \quad (20)$$

This is $(4/27 + o(1))n^4$ as $n \rightarrow \infty$, proving the desired upper bound on λ_H .

As a side remark, observe that if we let G be the Turán graph T_n^3 then all inequalities in (20) are equalities if n is divisible by 3 (and we have a gap of only $O(n^2)$ for general n).

Let us turn to proving the Erdős-Simonovits stability. Let $n \rightarrow \infty$ and G be any graph with $\lambda_H(G) = (4/27 + o(1))n^4$. Recall that $\overline{P}_3 = ([3], \{01\})$ is the graph with 3 vertices and exactly one edge.

Claim 3.6 $t(\overline{P}_3, G) = o(1)$.

Proof of Claim. Suppose that the claim is false. Then there are $\Omega(n)$ vertices u_0 for which the number of copies of $\overline{P}_3$ in G with 0 sent to u_0 is $\Omega(n^3)$. Each such vertex u_0 has degree $\Omega(n)$ and thus contributes a gap of $\Omega(n^3)$ to (20). Indeed, when counting embeddings f of H into G with $f(0) = u_0$, we upper bound the number of choices of the pair $(f(1), f(3))$ by $\deg(u_0)(n-1-\deg(u_0))$; however $\Omega(n^2)$ of such pairs (those that create a copy of $\overline{P}_3$ with u_0) are non-adjacent and do not extend to an embedding of H_4 . Thus the total gap in (20) is $\Omega(n^4)$, a contradiction. ■

By the Induced Removal Lemma [3], we can change $o(n^2)$ adjacencies in G so that no copy of $\overline{P}_3$ occurs. This means that G is complete partite.

Claim 3.7 *After removing $o(n)$ vertices, G has three parts, each of size $(1/3 + o(1))n$.*

Proof of Claim. Let the parts of G be $V_0, \dots, V_{m-1}$ of sizes $x_0 n \geq \dots \geq x_{m-1} n$. Let ℓ be the minimal index such that $x_\ell < 2/7$ (with $\ell := m$ if $x_{m-1} \geq 2/7$). Clearly, $0 \leq \ell \leq \lfloor 7/2 \rfloor = 3$. The argument leading to the upper bound (20) gives that

$$\lambda_H(G) \leq \sum_{i=0}^{m-1} x_i \cdot x_i (1 - x_i)^2 + o(1) \leq \frac{4}{27} \sum_{i=0}^{\ell-1} x_i + \frac{2}{7} \left(1 - \frac{2}{7}\right)^2 \cdot \sum_{i=\ell}^{m-1} x_i + o(1).$$

By $\lambda_H(G) = 4/27 + o(1)$ and $\sum_{i=0}^{m-1} x_i = 1$, this implies that $\sum_{i=\ell}^{m-1} x_i = o(1)$ and that, for each $0 \leq i < \ell$, we have $x_i = 1/3 + o(1)$. It follows that $\ell = 3$. Thus we can make the graph G 3-partite with each part of size $(1/3 + o(1))n$, by removing all vertices in $\cup_{i=\ell}^{m-1} V_i$, as desired. ■

Thus we can make G into T_n^3 by editing $o(n^2)$ adjacencies, which is exactly the claimed Erdős-Simonovits stability.

Claim 3.8 *The λ_H -problem is K_3 -flip-averse and K_3 -strict.*

Proof of Claim. By Claim 3.7, the unique optimal vector for K_3 is $(1/3, 1/3, 1/3)$. Thus the flip averseness follows by computing that, for $n \rightarrow \infty$,

$$\lambda_H(T_n^3) - \lambda_H(T_n^3 \oplus uw) + o(1) = \begin{cases} \frac{4}{9} + \frac{4}{9} + \frac{4}{9} - \frac{4}{9}, & \text{if } uw \in E(G), \\ -\frac{4}{9} - 0 - 0 + \frac{8}{9}, & \text{if } uw \in E(\overline{G}), \end{cases} = \frac{4}{9} > 0,$$

where the four terms in each row correspond to the total contribution of maps $V(H) \rightarrow V(T_n^3)$ such that uw is the image of 02, 03, 13 and 01 respectively.

In order to check strictness, we write the polynomial $q(\mathbf{y}) := p((1/3, 1/3, 1/3), \mathbf{y})$ as the sum $\sum_{i=0}^3 q_i(\mathbf{y})$, where p is as in (12) and, for $i \in [4]$, q_i corresponds to those functions f that map i to the special vertex. Thus q_i is the limit as $n \rightarrow \infty$ of $1/n^{(3)}$ times the number of embeddings of H into G' that send $i \in V(H)$ to u , where G' is obtained from the Turán graph T_n^3 by adding a vertex u with $y_i n/3$ neighbours in the i -th part. It is easy to see that

$$\begin{aligned} q_0(\mathbf{y}) &= \frac{y_0 + y_1 + y_2}{3} \cdot \sum_{i \in [4]} \sum_{j \in [4] \setminus \{i\}} \frac{y_i}{3} \cdot \frac{1 - y_j}{3}, \\ q_1(\mathbf{y}) &= \sum_{i \in [4]} \sum_{j \in [4] \setminus \{i\}} \frac{y_i}{3} \cdot \frac{1 - y_j}{3} \cdot \frac{2}{3}, \\ q_2(\mathbf{y}) &= \frac{y_0 + y_1 + y_2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3}, \\ q_3(\mathbf{y}) &= \frac{y_0^2 + y_1^2 + y_2^2}{9} \cdot \frac{2}{3}. \end{aligned}$$

Adding them, we obtain the (symmetric) polynomial $q(\mathbf{y})$. It can be checked to satisfy $q(0, 1, 1) = 4\lambda_H = 16/27$ (as expected) and our task is to show this element and its permutations are the

only elements of $[0, 1]^3$ that attain the maximum value $16/27$. Routine calculations show that, for any $i, j \in [3]$, it holds that

$$\frac{\partial q(\mathbf{y})}{\partial y_i} - \frac{\partial q(\mathbf{y})}{\partial y_j} = \frac{2}{27} \cdot (y_i - y_j) \cdot (4 - y_0 - y_1 - y_2).$$

This is zero for $\mathbf{y} \in [0, 1]^2$ only if $y_i = y_j$. Thus the only points that we need to check are those for which each coordinate is 0, 1 or some common value $y \in (0, 1)$. There are a few cases to consider. For each case, we have to find the values of a univariate polynomial of degree at most 3 on its critical points in $(0, 1)$ and check that we obtain $16/24$ only when the corresponding full vector $\mathbf{y} \in [0, 1]^3$ is a permutation of $(0, 0, 1)$. This is a routine task, so we just do one case for an illustration. For example, one can compute that $q(y, y, y) = (-6y^3 + 4y^2 + 6y)/9$ and its derivative has only one root in $(0, 1)$, namely $y = (2 + \sqrt{31})/9$ with the value of the polynomial being $(356 + 124\sqrt{31})/2187 < 16/27$, so we do not get any counterexample to strictness in this case. We refer the reader to our code that does full verification as described in Section 2 (without using any symmetries of the base graph). ■

Theorem 5.8 in [21] states, modulo minor technical assumptions that hold here, that if, with respect to some base graph B , an extremal λ_γ -problem is Erdős-Simonovits B -stable while the graph B is both λ_γ -flip-averse and λ_γ -strict then the problem is perfectly stable. It applies to the H_4 -semi-inducibility by the above claims, thus proving perfect K_3 -stability. ■

3.5 Graph H_{15}

Recall that H_{15} has red edges 01, 02 and blue edges are 03, 12, 13. Here we determine the exact value of $\Lambda_{H_{15}}(n)$ except when n is 1 or 4 modulo 6 when we will have a gap of order n^2 between our bounds.

As a building block we will need a construction of a triangle-free d -regular n -vertex graph for some special pairs (n, d) . The following two simple construction will be our basic blocks. For integers $m \geq d \geq 0$, let $B_{m,d}$ be the bipartite graph with $2m$ vertices where each part is a copy of the residues modulo m and we connect two residues across if their sum modulo m is in $[d]$. Clearly, the graph $B_{m,d}$ is triangle-free (as it is bipartite) and d -regular. Next, for any positive integers m, k with $3k < m + 2$, let $R_{m,k}$ be the graph whose vertex set consists of the residues modulo $2m + 1$ and two vertices are adjacent if their difference modulo $2m + 1$ belongs to the (symmetric under negation) interval

$$I := \{m + 1 - k, m + 2 - k, \dots, m - 1 + k, m + k\}.$$

In other words, if we view the vertices as arranged on a circle then we connect each vertex to the diametrically opposite interval of size $2k$. Clearly, this graph is $2k$ -regular. In order to check that there are no triangles, it is enough, by the symmetries of the construction, to check that two cyclically shortest hops from 0 jump over the interval I , that is, we have $2(m + 1 - k) > m + k$, which is exactly the assumption that we made.

Theorem 3.9 *Let $H = H_{15}$. For every $n \geq 4$, it holds that $\Lambda_H(n) \leq n \pi_3(n - 1)$.*

If n is 0, 2, or 5 modulo 6 then $\Lambda_H(n) = n \pi_3(n-1)$ and an order- n graph G is extremal if and only if G is $\lceil n/3 \rceil$ -regular and triangle-free.

If $n = 6k + 3$ for $k \in \mathbb{N}$ then $\Lambda_H(n) = n \pi_3(n-1) - 2k$ and an order- n graph G is extremal if and only if G is triangle-free, has exactly $n-1$ vertices of degree $2k+1$, and one vertex of degree $2k$ or $2k+2$.

If $n = 6k + 1$ or $n = 6k + 4$ for integer $k \rightarrow \infty$ then

$$6k^2 + O(k) \leq n \pi_3(n-1) - \Lambda_H(n) \leq \frac{15}{2} k^2 + O(k).$$

In particular, it follows from above that $\lambda_H = 1/27$. Furthermore, a graph G of an arbitrary order $n \rightarrow \infty$ satisfies $\lambda_H(G) = 1/27 + o(1)$ if and only if G has triangle density $o(1)$ and is almost $1/3$ -regular.

Proof. Take any graph G . For a vertex u_i of G , we use the shortcuts $d_i := \deg_G(u_i)$, $\Gamma_i := \Gamma_G(u_i)$ and $\bar{\Gamma}_i := \Gamma_{\bar{G}}(u_i)$. Thus Γ_i and $\bar{\Gamma}_i$ are the sets of respectively neighbours and non-neighbours of u_i in G (which by definition do not contain the vertex u_i itself). Also, for adjacent vertices u_0 and u_1 , we let t denote the number of triangles in G containing the edge $u_0 u_1$.

Let us show that $\Lambda_H(G) \leq n \pi_3(n-1)$ for any n -vertex graph G . Every embedding f of H into G can be obtained by picking adjacent u_0 and u_1 for respectively $f(0)$ and $f(1)$, and then arbitrary $f(2) \in \Gamma_0 \cap \bar{\Gamma}_1$ and $f(3) \in \bar{\Gamma}_0 \cap \Gamma_1$; conversely, every such map f gives an embedding. Since the sets from which we pick $f(2)$ and $f(3)$ partition the $(n-1-d_1)$ -set $\bar{\Gamma}_1$, the number of choices here is at most $\pi_2(n-1-d_1)$. Summing this over all ordered pairs (u_0, u_1) of adjacent vertices, we obtain the desired upper bound

$$\Lambda_H(G) \leq \sum_{u_1 \in V(G)} \sum_{u_0 \in \Gamma_1} \pi_2(n-1-d_1) = \sum_{u_1 \in V(G)} d_1 \pi_2(n-1-d_1) \leq n \pi_3(n-1), \quad (21)$$

where we used the trivial observation that

$$D(m) := \pi_3(n-1) - m \pi_2(n-1-m),$$

is non-negative for every integer $m \in [n]$. Let us call $D(u_1)$ the *defect* of u_1 .

For future reference, we also define the *defect* of adjacent vertices u_0 and u_1 to be

$$\begin{aligned} D(u_0, u_1) &:= \pi_2(n-1-d_1) - |\Gamma_0 \cap \bar{\Gamma}_1| \cdot |\bar{\Gamma}_0 \cap \Gamma_1| \\ &= \pi_2(n-1-d_1) - (d_0-1-t)(n-d_0-d_1+t) \geq 0. \end{aligned}$$

and the (total) *defect* of G to be

$$D(G) := \sum_{u_1 \in V(G)} \left(D(d_1) + \sum_{u_0 \in \Gamma_1} D(u_0, u_1) \right) \geq 0.$$

Then the upper bound proved in (21) is a consequence of the identity

$$\Lambda_{H_{15}}(G) = n \pi_3(n-1) - D(G). \quad (22)$$

Take any n -vertex graph G attaining equality in (21); thus the defect of G is zero, which implies that all vertex and pair defects are zero. Take any adjacent u_0 and u_1 . Since $D(u_0, u_1) = 0$, it holds that

$$-1 \leq (n - d_0 - d_1 + t) - (d_0 - 1 - t) \leq 1.$$

By re-arranging the terms, we obtain

$$n + 2t \leq 2d_0 + d_1 \leq n + 2t + 2. \quad (23)$$

Also, the equality $D(d_1) = 0$ (for any integer d_1) is equivalent to $|2d_1 - (n - 1 - d_1)| \leq 2$, which can be rewritten as

$$n - 3 \leq 3d_1 \leq n + 1. \quad (24)$$

Note that (23) holds for every two adjacent vertices taken in either order, so it remains true if we swap d_0 and d_1 . Thus it follows from (23) and (24) that $n + 2t \leq n + 1$, that is, $t = 0$. Since $u_0 u_1 \in E(G)$ was arbitrary, G is triangle-free.

Let us show that G is $\lceil n/3 \rceil$ -regular. Take any vertex u_1 . By (24), u_1 cannot be an isolated vertex. Take any its neighbour u_0 . If $d_0 = d_1$ then the above equations imply that $n \leq 3d_1 \leq n + 2$ and thus $d_1 = \lceil n/3 \rceil$, as desired. If $d_0 < d_1$ then, by (23) and (24), we have $n \leq 2d_0 + d_1 \leq 3d_1 - 2 \leq n - 1$, a contradiction. If $d_1 > d_0$ then we get a contradiction in the same way except we use (23) with d_0 and d_1 swapped.

Thus if n is 0, 2, or 5 modulo 6 then every n -vertex graph attaining the upper bound is as stated (and conversely). Also, the set of such graphs is non-empty as shown by $B_{3k, 2k}$ for $n = 6k$, $B_{3k+1, 2k+1}$ for $n = 6k + 2$, and $R_{3k-1, k}$ for $n = 6k - 1$ with integer $k \geq 1$.

Next, let $n = 6k + 3$. Take any graph G with $6k + 3$ vertices and $\Lambda_H(G) \geq n\pi_3(n - 1) - 2k$. Note that $D(d_1) = 0$ if and only if d_1 is $2k$ or $2k + 1$; otherwise it is easy to see that

$$D(d_1) \geq \pi_3(6k + 2) - \min \{ (2k + 2)(2k)^2, (2k - 1)(2k + 1)(2k + 2) \} = 2k$$

and equality holds if and only if $d_1 = 2k + 2$. Thus every degree of G is $2k$ or $2k + 1$, except at most one vertex of degree $2k + 2$.

First, suppose that there is a vertex u_1 of degree $2k$. For any $u_0 \in \Gamma_1$, we have that

$$D(u_0, u_1) = \pi_2(4k + 2) - (d_0 - t - 1)(4k + t + 3 - d_0),$$

which is at least 1 if $d_0 = 2k + 1$ and at least 4 if $d_0 = 2k$, with equality in these cases only if $t = 0$. Since at most one vertex has degree $2k + 2$, we have that $\sum_{u_0 \in \Gamma_1} D(u_0, u_1) \geq 2k$ with equality only if each of $2k$ neighbours of u_1 has degree exactly $2k + 1$ and there are no triangles containing u_1 . Thus all other defect values are zero, from which it follows that there is only one vertex of degree $2k$ (and no vertex has degree $2k + 2$); moreover, for any two adjacent vertices u_0 and u_1 of degree $2k + 1$ we have that

$$D(u_0, u_1) = \pi_2(4k + 1) - (2k - t)(2k + t + 1) = t^2 + t,$$

so necessarily $t = 0$. Thus G is triangle-free (and its degrees are as stated in the theorem) and we are done.

Second, suppose that each degree is $2k + 1$ or $2k + 2$. Since n is odd, there is at least one vertex u_1 of degree $2k + 2$. We have

$$D(2k + 2) \geq \pi_3(6k + 2) - (2k + 2)\pi_2(4k) = 2k \cdot (2k + 1)^2 - (2k + 2)(2k)^2 = 2k.$$

Thus all other defects are zero, u_1 is the only vertex of degree $2k + 2$, all other vertices have degree $2k + 1$ and there are no triangles, as desired.

Conversely, the above calculations show that, for each $k \geq 1$, every graph G as stated in the theorem satisfies $\Lambda_H(G) = n\pi_3(n - 1) - 2k$.

For a construction of at least one such graph, we start with $R_{3k+1,k}$. This graph remains triangle-free if we add any number of vertex-disjoint edges, each of the form $\{i, i + 2k\}$ or $\{i, i + 2k + 1\}$ for some residue i . Indeed, a triangle can use at most one of the added edges by their disjointness. On the other hand, each original edge of $R_{3k+1,k}$ has cyclic length at least $2k + 2$ so after two jumps in the graph $R_{3k+1,k}$ (in the same or opposite directions) we are at distance at most $2k - 1$ from where we started but no edges of such length are added. Note that the graph on residues modulo $6k + 3$ consisting only of the edges of length $2k$ is a vertex-disjoint union of 3 cycles C_0 , C_1 and C_2 , each of length $2k + 1$. Let us index them so that $V(C_i) \ni i$ for $i \in [3]$. Let G be obtained from $R_{3k+1,k}$ by adding, to the edge set, the pair $\{0, 2k + 1\}$ and matchings of size k in each of C_i that avoid the vertices $0 \in V(C_0)$ and $2k + 1 \in V(C_1)$. Each vertex in this triangle-free graph has degree $2k + 1$ except the unmatched vertex of C_2 which has degree $2k$, as desired.

Let $n = 6k + 1$ with integer $k \rightarrow \infty$. Take any order- n graph G . First, let us show that the defect of G is at least $6k^2 + O(k)$. Suppose that, say, $D(G) \leq 6k^2$ as otherwise we are done.

Note that $D(d_1) = (2k)^3 - d_1\pi_2(6k - d_1)$. It is 0 if $d_1 = 2k$, it is $2k$ if $d_1 = 2k \pm 1$ and, for $d_1 = 2k + s$ with $|s| \geq 2$, it is at least

$$(2k)^3 - (2k + s) \left(\frac{4k - s}{2} \right)^2 = \frac{3}{2}ks^2 - \frac{1}{4}s^3 \geq \max \left\{ 6k - 2, \frac{1}{2}ks^2 \right\} \quad (25)$$

since $|s| \leq 4k$. Let A (resp. B) be the set of vertices of degree $2k$ (resp. $2k + 1$). Let $a := |A|$ and $b := |B|$. By above, we have that $(n - a) \cdot 2k \leq D(G) \leq 6k^2$. Thus $a \geq n/2$.

Note that for any two adjacent vertices $u_0, u_1 \in A$, it holds that

$$D(u_0, u_1) = \pi_2(4k) - (2k - 1 - t)(2k + 1 + t) = (t + 1)^2 \geq 1. \quad (26)$$

Thus every edge inside A contributes 2 to $D(G)$, since it is counted for its both orderings. By $D(G) \leq 6k^2$ and (25), we can assume that the maximum degree is at most $2k + \sqrt{12k}$. Thus the number of edges between A and its complement is at most $b(2k + 1) + (n - a - b)(2k + \sqrt{12k})$ and 2 times the number of edges inside A can be lower bounded by $a \cdot 2k$ minus this quantity. By (25) and (26), we have

$$D(G) \geq \left(a \cdot 2k - b(2k + 1) - (n - a - b)(2k + \sqrt{12k}) \right) + b \cdot 2k + (n - a - b) \cdot (6k - 2). \quad (27)$$

This is a linear function of b with the coefficient at b being $-4k + \sqrt{12k} + 1 < 0$. Thus it is at least its value when $b = n - a$, giving $D(G) \geq 2ak + O(k) \geq 6k^2 + O(k)$, as desired.

The stated lower bound on $\Lambda_H(6k+1)$ comes from the following construction. Let $a := \lfloor k/2 \rfloor$. We start with $B_{3k,k}$, denoting its parts as $U := \{u_0, \dots, u_{3k-1}\}$ and $W := \{w_0, \dots, w_{3k-1}\}$ where index arithmetic will be taken modulo $3k$. It will be notationally convenient to rotate one set of the residues, so we connect v_i to u_j in $B_{3k,k}$ if and only if $i + j$ belongs to the interval $\{2k+1, \dots, 4k\}$ modulo $3k$. Add new a edges $\{u_i, w_{i+k-a}\}$ for $i \in \{0, \dots, a-1\}$. Add a new vertex u connected to $2k$ elements $w_k, \dots, w_{3k-1}$. The obtained graph G is clearly bipartite. It has $2k + 2a$ vertices of degree $2k + 1$ (namely, $u_0, \dots, u_{a-1}$ and $w_{k-a}, \dots, w_{3k-1}$) with the remaining $4k + 1 - 2a$ vertices (namely, $u_a, \dots, u_{3k-1}$, $w_0, \dots, w_{k-1-a}$ and u) having degree $2k$, so $\sum_{v \in V(G)} D(v) = (2k + 2a) \cdot 2k$. It is routine to see that, since there are no triangles, if we take two adjacent vertices, one of degree $2k + 1$ and the other of degree $2k + 1$ or $2k$ then their defect (in either ordering) is zero. Also, as we observed earlier (see (26)) the defect of any ordered adjacent pair of vertices each of degree $2k$ is 1. Each such pair contains one of $w_0, \dots, w_{k-a-1}$, while each of these vertices w_i is adjacent to all degree- $(2k + 1)$ vertices in U and thus has exactly $2k - a$ neighbours of degree $2k$. Also, all $2k$ neighbours of the added vertex u have degree $2k + 1$. Thus the total defect is $(2k + 2a) \cdot 2k + 2 \cdot (k - a)(2k - a) = 15k^2/2 + O(k)$, as desired.

Finally, let $n = 6k + 4$. Here the proof of the upper bound is very similar to the case $n = 6k + 1$, so we will be rather brief. Take any extremal graph G . For $u_1 \in V(G)$, it holds that $D(d_1) = 0$ only if $d_1 = 2k + 1$. Furthermore, $D(d_1) = 2k + 1$ if $d_1 = 2k$ or $2k + 2$ and then $D(d_1)$ grows quadratically as d_1 moves away from $2k + 1$. In particular, the maximum degree of G is at most $2k + O(\sqrt{k})$. Also, one can show that every adjacent pair of vertices, each of degree $2k + 1$, has defect at least 1. It follows from the above observations by an easy adaptation of the argument leading to (27) that $D(G) \geq 6k^2 + O(k)$.

For the lower bound, observe that, if the graph is triangle-free, then the defect of any adjacent pair of vertices (in either order), one of degree $2k + 2$ and the other of degree $2k + 1$ or $2k + 2$, is zero. Our construction we will have only vertices of degree $2k + 1$ and $2k + 2$, about $n/2$ of each degree; then we have to minimise the number of edges spanned by a set of vertices of degree $2k + 1$. Let $a := \lceil (3k + 2)/2 \rceil$. Take the union of complete bipartite graphs $T_{a, 3k+2-a}$ with parts $U_0 \cup U_1$ and $T_{3k+2-a, a}$ with parts $W_0 \cup W_1$. Add a copy of $B_{a, a-k}$ with parts U_0 and W_1 and a copy of $B_{3k+2-a, 2k+1-a}$ with parts U_1 and W_0 . All vertices in $U_0 \cup W_1$ (resp. $U_1 \cup W_0$) have degree $2k + 2$ (resp. $2k + 1$). Also, each vertex of degree $2k + 1$ has $k/2 + O(1)$ neighbours of degree $2k + 1$. Thus the total defect is

$$|U_0 \cup W_1| \cdot k/2 + |U_1 \cup W_0| \cdot (2k + 1) + O(k) = \frac{15}{2} k^2 + O(k),$$

as desired.

Finally, it is routine to see from (22) that a graph G of an arbitrary order $n \rightarrow \infty$ is almost extremal (that is, $D(G) = o(n^4)$) if and only if it is as stated in the theorem. ■

The upper bound of Theorem 3.9 for $n = 6k + 1$ and $n = 6k + 4$ can be improved with extra work to $\Lambda_H(n) \leq n \pi_3(n - 1) - (6 + c + o(1))k^2$ for some (small) $c > 0$. However, it seems that pushing $6 + c$ all the way up to $15/2$ (if this is the correct constant at k^2) would be rather messy. So we limit ourselves to the stated bounds which are relatively easy to show.

3.6 Computer-generated results for H_9 , H_{10} and H_{11}

Here we present the cases (namely the graphs H_9 , H_{10} and H_{11}) where the flag algebras not only give the semi-inducibility constant but also prove perfect stability by applying the criterion given by the second author, Sliacan and Tyros [21, Theorem 7.1] that can be automatically checked by computer.

Since the statement of [21, Theorem 7.1] is rather long, we just give an informal description of it. In addition to a flag algebra certificate proving $\lambda_\gamma \leq u$ as in (15) (with the matching lower bound coming from blowups of some base graph B), we have to provide some graph τ with at most $N - 2$ vertices that satisfies certain properties. Some of the required properties of τ are: (a) if we forbid τ as induced subgraph then the maximum value of λ_γ on graphs of order $n \rightarrow \infty$ drops by $\Omega(1)$, (b) τ admits the unique (up to an automorphism of B) homomorphism $f : \tau \rightarrow B$, and (c) the each vertex of B is uniquely identified by its adjacencies to $f(\tau)$. The gist of the proof of [21, Theorem 7.1] is that every almost extremal G of order $n \rightarrow \infty$ has τ -density $\Omega(1)$ by (a), every copy of τ in G gives pairwise disjoint $V_0, \dots, V_{m-1} \subseteq V(G)$ naturally defined by looking at adjacencies of a vertex of G to that copy by (b)–(c) so that any single discrepancy between $E(G)$ and $E(B(V_0, \dots, V_{m-1}))$ gives $\Omega(n^{N-v(\tau)-2})$ N -vertex subgraphs F of G that are ‘wrong’ (i.e. do not admit a homomorphism to B). If, furthermore, every ‘wrong’ F has positive slack $c_F > 0$ then G must be close to $B(V_0, \dots, V_{m-1})$ for a ‘typical’ copy of τ . The proof also needs that at least one of certain Conditions (i) and (ii) of [21, Theorem 7.1] holds, and if Condition (i) holds then the proof gives that the optimal B -vector is unique. In brief, Condition (i) states that the SOS-term in (15) contains squares of τ -rooted subgraph densities such that the positive semi-definite matrix that encodes them has the smallest possible co-rank 1. When we use [21, Theorem 7.1], we just list N and τ , and state which of Conditions (i) and (ii) was verified by our code.

Theorem 3.10 *Let H_9 denote the graph on [4] with the red edge set $\{03\}$ and the blue edge set $\{01, 02, 12, 13\}$. Then it holds that $\lambda_{H_9} = 12/125$, the problem is perfectly B -stable for $B = K_5$ by [21, Theorem 7.1(i)] with $N = 7$ and $\tau = ([5], \{01\})$. Furthermore, the unique maximizer of $\lambda_{H_9}(B(\mathbf{a}))$ is $\mathbf{a} = (1/5, 1/5, 1/5, 1/5, 1/5)$. ■*

Theorem 3.11 *Let H_{10} denote the graph on [4] with the red edge set $\{01, 13\}$ and the blue edge set $\{02, 03\}$. Then it holds that $\lambda_{H_{10}} = 1/8$, the problem is perfectly B -stable for $B = K_2$ by [21, Theorem 7.1(ii)] with $N = 5$ and $\tau = K_1$. Furthermore, the unique maximizer of $\lambda_{H_{10}}(B(\mathbf{a}))$ is $\mathbf{a} = (1/2, 1/2)$.*

Proof. Since Condition (i) of [21, Theorem 7.1] does not apply here, we have to check the uniqueness of the maximiser $\mathbf{a}$. Note that $\Lambda_H(T_{xn, (1-x)n}) = p(x)n^{(4)} + o(n^4)$, where $p(x) := x(1-x)^3 + x^3(1-x)$ since the only part assignment that gives a copy of H is when the vertices 0, 2, 3 are in one part and the vertex 1 in the other. As it is easy to see, the maximum of $p(x)$ on $[0, 1]$ is $1/8$ and this is attained if and only if $x = 1/2$, as desired. ■

Theorem 3.12 *Let H_{11} denote the graph on $[4]$ with the red edge set $\{01, 03\}$ and the blue edge set $\{02, 13\}$. Then it holds that $\lambda_{H_{11}} = 1/8$, the problem is perfectly B -stable for $B = K_2$ by [21, Theorem 7.1(i)] with $N = 4$ and $\tau = K_2$. Furthermore, the unique maximizer of $\lambda_{H_{11}}(B(\mathbf{a}))$ is $\mathbf{a} = (1/2, 1/2)$. ■*

Since the flag algebra proof of $\lambda_{H_{11}} \leq 1/8$ uses only $N = 4$, the identity in (15) can be explicitly written. Namely, after some cleaning up when we try to make as many coefficients inside the certificate as possible to be 0, the proof amounts to checking that, for any order- n graph G , if we expand

$$\frac{n^4}{8} - \Lambda_{H_{11}}(G) - \frac{3}{8} \sum_{uw \in E(G)} (\deg_G(u) - \deg_G(w))^2 - \frac{1}{8} \sum_{uw \in E(\bar{G})} (|\Gamma_G(u) \cap \Gamma_G(w)| - |\Gamma_{\bar{G}}(u) \cap \Gamma_{\bar{G}}(w)|)^2 \quad (28)$$

as a linear combination of 4-vertex subgraph counts up to error term $o(n^4)$ then each coefficient will be non-negative. While this proves that $\lambda_{H_{11}} \leq 1/8$, it does not give much insight. (Unfortunately, computer calculations indicate that there is no proof of $\lambda_{H_{11}} \leq 1/8$ with $N = 4$ as in (28) that has only one sum, either over edges or over non-edges of G .)

3.7 Graph H_6

For the graph H_6 , flag algebras give the sharp upper bound on the semi-inducibility constant. However, perfect stability does not hold. So we prove the Erdős-Stone stability by extracting some further information from the certificate, in addition to just the upper bound $\lambda_H(H_6) \leq 1/8$. Then, by using the standard stability approach, we get exact result for all $n \geq n_0$.

Theorem 3.13 *Let H_6 be the graph on $[4]$ with the red edge set $\{02\}$ and the blue set $\{01, 03, 13\}$. Then $\lambda_{H_6} = 1/8$, where the upper bound is proved via flag algebras with $N = 5$. The lower bound comes from blowups of K_2 with the part ratios given by $\mathbf{a} = (1/2, 1/2)$. Moreover, if a 5-vertex graph F has density $\Omega(1)$ in an almost extremal graph then F is complete bipartite or is obtained from a 4-vertex complete bipartite graph by adding one isolated vertex.*

Proof. The result is obtained by a standard application of flag algebras. Also, our script checks that every 5-vertex graph F with $c_F = 0$ is complete bipartite plus at most one isolated vertex. ■

First, let us observe that the semi-inducibility problem for H_6 does not satisfy perfect stability. Consider the graph $G := T_{x_0 n, x_1 n} + \bar{K}_{(1-x_0-x_1)n}$ (that is, G is obtained from the complete bipartite graph with part sizes $x_0 n$ and $x_1 n$ by adding $(1 - x_0 - x_1)n$ isolated vertices). We have that $\lambda_{H_6}(G) = q(x_0, x_1) + o(1)$, where

$$q(x_0, x_1) := x_0 x_1 \cdot (1 - 2x_0 x_1 + (1 - x_0 - x_1)^2). \quad (29)$$

Indeed, we first choose an unordered pair uw of adjacent vertices in G (to be the image $\{f(0), f(2)\}$); then an ordered pair $(f(1), f(3))$ of two non-adjacent vertices of G gives rise to exactly two possible maps of 02 to uw if both u and w are isolated vertices of G and exactly

one such map otherwise. Now, if we take small $\varepsilon > 0$ and let each of x_0 and x_1 be $1/2 - \varepsilon$ then the density of H_6 is $1/8 - \varepsilon^2 + O(\varepsilon^3)$ while G can be shown to be at the edit distance at least $\Omega(\varepsilon n^2)$ from a blowup of K_2 . Thus there is no perfect stability.

On the other hand, we can show that this problem satisfies the Erdős-Simonovits stability.

Lemma 3.14 *For every $\varepsilon > 0$ there are $\delta > 0$ and n_0 such that every graph G with $n \geq n_0$ that satisfies $\lambda_{H_6}(G) \geq 1/8 - \delta$ is within edit distance at most εn^2 from $T_{n/2, n/2}$.*

Proof. Suppose that the lemma is false for some $\varepsilon > 0$. Thus we can find an almost extremal graph of order $n \rightarrow \infty$ that is εn^2 far from $T_{n/2, n/2}$. By the second part of Theorem 3.13 and the Induced Removal Lemma [3], we can change $o(n^2)$ adjacencies in G such that every 5-vertex subset spans a complete bipartite graph apart possibly one isolated vertex. In particular, the new graph G contains neither a triangle nor an induced path on 4 vertices. It follows that G is bipartite, for a shortest odd cycle in G , which is necessarily induced, has to contain one of these two configurations. Furthermore, if two vertices u, w are connected by a path of length 3 then u, w must be adjacent. It follows that each component of G is a complete bipartite graph. Since G has no induced matching with 2 edges, it is a complete bipartite graph T_{n_0, n_1} plus $n - n_0 - n_1$ isolated vertices. Let $x_0 := n_0/n$ and $x_1 := n_1/n$. Thus $\lambda_{H_6}(G) = q(x_0, x_1)$, where q is the degree-4 polynomial defined by (29).

It follows from Theorem 3.13 that $q(x, y) \leq 1/8$ on

$$S := \{(x, y) \in \mathbb{R}^2 : x \geq 0, y \geq 0, x + y \leq 1\}.$$

We need to show that equality holds only if $x = y = 1/2$. For this task, we first re-prove the upper bound $1/8$ on the values of q on S using basic calculus. Here we maximise a polynomial on a compact set, so the maximum is attained at a critical point or on the boundary. First, let us find all critical points (x, y) of q in the interior of S . If $y = x$ then

$$q(x, x) = \frac{1}{8} + \frac{1}{8}(2x - 1)^2(4x^2 - 4x - 1)$$

is less than $1/8$ for each $x \in (0, 1/2)$. So assume that $x \neq y$. The following identity is routine to check:

$$\frac{\partial q(x, y)}{\partial y} - \frac{\partial q(x, y)}{\partial x} = (y - x) \left((x - y)^2 + 2(1 - x - y) \right).$$

Its right-hand side is non-zero so (x, y) is not a critical point. Thus it remains to consider the boundary of S . For every real x , we have $q(x, 0) = q(0, x) = 0$ while $q(x, 1 - x) = 1/8 - (1 - 2x)^2/8$ is at most $1/8$ with equality if and only if $x = 1/2$. Thus $(1/2, 1/2)$ is the only point of S where q attains value $1/8$.

By the continuity of the involved functions, it follows that each of x_0 and x_1 is $1/2 + o(1)$. However, then G is within edit distance $o(n^2)$ from $T_{n/2, n/2}$, contradicting our assumption. ■

Although perfect stability does not hold for H_6 , we will show in Theorem 3.16 that every extremal graph G of sufficiently large order is exactly complete bipartite. Our proof will need the following lemma which is routine to verify; in fact, one can explicitly write all optimal a if needed, see e.g. [7, Eq. (16)].

Lemma 3.15 *As $n \rightarrow \infty$, an integer a with $0 \leq a \leq n$ that maximises the polynomial*

$$P_n(x) := x^{(3)}(n-x) + x(n-x)^{(3)}. \quad (30)$$

satisfies $a = n/2 + \sqrt{3n}/2 + O(1)$ and the maximum value is $n^{(4)}/8 + 3n^2/4 + O(n)$. ■

Theorem 3.16 *Let H be the graph H_6 (that has one red edge 02 and blue edges 01, 03, 13). Then every graph G of sufficiently large order n with $\Lambda_H(G) = \Lambda_H(n)$ is a blowup of K_2 .*

Proof. Let $G = (V, E)$ be any extremal graph of sufficiently large order n . For notational convenience, we view n as going to the infinity and use the asymptotic notation accordingly. Let $V_0 \cup V_1 = V$ be a max-cut partition of G . By Lemma 3.14, we have that the number of edges between V_0 and V_1 is at least $(1/4 + o(1))n^2$. It follows that $|V_i| = n/2 + o(n)$ for each $i = 0, 1$ and almost every pair between V_0 and V_1 is an edge of G . Also, it follows that the number of edges inside V_0 or V_1 is $o(n^2)$ as otherwise G will be $\Omega(n^2)$ -far from being bipartite (e.g. by having $\Omega(n^3)$ triangles). Let W consist of all *wrong* pairs, that is, non-edges across the parts and edges inside a part. As we just argued, it holds that $|W| = o(n^2)$.

Take any vertex $u \in V$. For $i = 0, 1$, define $y_i := |\Gamma(u) \cap V_i|/|V_i|$ to be the fraction of vertices in V_i that are neighbours of u .

Let us show that $\Lambda_H(G, u)$, the number of embeddings of H into G that use the vertex u , is equal to $p(y_0, y_1)n^3 + o(n^3)$, where we define

$$p(x, y) := \frac{x^3 + x^2y + xy^2 + y^3}{8} - \frac{xy}{2} - \frac{x}{8} - \frac{y}{8} + \frac{1}{2}. \quad (31)$$

This calculation is straightforward but rather lengthy, so we include some details to make it easier to verify for the reader. Here, we can replace $G - u$ by the blowup of K_2 with the parts $V_0 \setminus \{u\}$ and $V_1 \setminus \{u\}$, since this changes the number of embeddings via u by at most $4!|W|n = o(n^3)$. Clearly, $\Lambda_H(G, u)$ is the sum over all i in $V(H) = [4]$ of the number of embeddings that map i to u . Also, the i -th summand can be written as $p_i(y_0, y_1)n^3 + o(n^3)$, where

$$\begin{aligned} p_0(x, y) &:= \frac{x+y}{2} \cdot \frac{(1-x)^2 + (1-y)^2}{4}, \\ p_1(x, y), p_3(x, y) &:= \frac{(1-x)^2}{8} + \frac{(1-y)^2}{8}, \\ p_2(x, y) &:= \frac{x+y}{8}. \end{aligned}$$

For example, the formula for p_1 follows by observing that 0 and 3 have to be mapped inside the same part V_i and be non-neighbours of u , while 2 has to go to a vertex in the other part V_{1-i} (and, conversely, every such injection into $V \setminus \{u\}$ gives an embedding of H). Finally, it is routine to check that $p = \sum_{i=0}^3 p_i$.

The pairs $(x, y) = (0, 1)$ and $(1, 0)$ correspond to u being adjacent to every vertex in one part and to none in the other. As expected, the value of p on them is $1/2 = 4\lambda_H$. Also, it holds that $p(x, y) \leq 1/2$ for every $(x, y) \in [0, 1]^2$ (as otherwise we would contradict Theorem 3.13

by εn vertices to G with appropriate attachment for some small constant $\varepsilon > 0$). Recall that the strictness property would state here that all optimal (x, y) are of this form. However, this property does not hold for H -semi-inducibility, since p attains the maximum value $1/2$ also on $(0, 0)$. As a step in our exactness result, we need to describe all optimal (x, y) . For this, we have to calculate the maximum of the explicit polynomial p from scratch, using basic calculus.

Claim 3.17 *The maximum of the polynomial $p(x, y)$ (defined in (31)) over $(x, y) \in [0, 1]^2$ is $1/2$ and this is attained if and only if (x, y) is one of $(0, 1)$, $(1, 0)$ and $(0, 0)$.*

Proof of Claim 3.17. We maximise a differentiable function over a compact set so its maximum is attained at a critical point in the interior or at some point on the boundary. The identity

$$\frac{\partial p(x, y)}{\partial x} - \frac{\partial p(x, y)}{\partial y} = \frac{(x - y)(x + y + 2)}{4}$$

shows that all interior critical points are limited to the line $x = y$. We have that $p(x, x) - 1/2 = x(2x^2 - 2x - 1)/4$ and this is negative for all $0 < x < 1$. So it remains to consider the boundary. For $0 \leq x \leq 1$, it holds that $p(x, 0) - 1/2 = x^3/8 - x/8 \leq 0$ with equality if and only if $x = 0$ or $x = 1$, and that $p(x, 1) - 1/2 = x(x^2 + x - 4)/8 \leq 0$ with equality if and only if $x = 0$. This takes care also of the remaining pieces of the boundary by the symmetry between x and y , finishing the proof. ■

The average number of embeddings that use a random vertex of the extremal graph G is $\Lambda_H(n) \cdot 4/n$. Since we cannot increase $\Lambda_H(G)$ by replacing a vertex by a clone of another, every vertex u is in at least $4\Lambda_H(n)n^3 + O(n^2) = (4\Lambda_H + o(1))n^3$ embeddings. This implies by Claim 3.17 that the vector (y_0, y_1) (where $y_i := |\Gamma(u) \cap V_i|/|V_i|$) is $o(1)$ -close to $(0, 1)$, $(1, 0)$, or $(0, 0)$. In the first two cases, the vertex u belongs necessarily to respectively V_0 and V_1 by the max-cut property.

Define V_2 to be the set of those vertices $u \in V$ for which the last alternative holds (which is equivalent to $\deg(u) = o(n)$). The edit distance from G to a complete partite graph is $\Omega(|V_2|n)$ (e.g. by considering the degrees of the vertices in V_2). Hence, by stability (that is, by Lemma 3.14), we have that $|V_2| = o(n)$.

Let $V'_i := V_i \setminus V_2$ for $i = 0, 1$. Define W' to consist of the non-edges between V'_0 and V'_1 , the edges inside V'_0 or V'_1 (so far these the same definitions same as for W) and the edges between $V'_0 \cup V'_1$ and V_2 . By above, we have that the maximum degree $\Delta(W') = o(n)$. As the next step, we can show that W' is in fact empty.

Claim 3.18 *It holds that $W' = \emptyset$.*

Proof of Claim 3.18. Suppose on the contrary that some pair uw is in W' . Let $G' := G \oplus uw$ be obtained from G by changing the adjacency between u and w . We would like to show that $\Lambda_H(G') - \Lambda_H(G) \geq \varepsilon n^2$ for some small constant $\varepsilon > 0$, which will be the desired contradiction. Note that only those injections that use both u and w can contribute non-zero amount to $\Lambda_H(G) - \Lambda_H(G')$. Thus if, in each of G and G' , we remove all vertices of $V_2 \setminus \{u, w\}$ and correct

every pair of W' (or, equivalently, W) except uw , then the difference $\Lambda_H(G') - \Lambda_H(G)$ changes by at most $4! (|V_2|n + |W'| + \Delta(W')n) = o(n^2)$, which can be taken care of by just slightly decreasing ε . Thus we can calculate $\Lambda_H(G') - \Lambda_H(G)$ with respect to the simplified graphs.

First, suppose that uw is an edge inside, say, V'_0 . In this case, G' is obtained from G by making u and w non-adjacent. Each map f that is an embedding of H into G but not into G' has to map 02 to uw in some order (2 choices). Ignoring V_2 , every extension has to map both 1 and 3 to V'_0 , so we have in total at most $2(n/2)^2 + o(n^2)$ such embeddings. On the other hand, we can lower bound the number of the embeddings created by flipping uw as follows. First, we can map $\{1, 3\}$ to uw (2 choices) then 0 to $V'_0 \setminus \{u, w\}$ and 2 to V'_1 $((n/2)^2 + o(n^2))$ choices). Second, we can map 0 into uw and then one of 1, 3 to the remaining vertex in uw (4 choices), then we can map 2 to V'_1 and the remaining vertex of $\{1, 3\}$ to $V'_0 \setminus \{u, w\}$, having here $(n/2)^2 + o(n^2)$ choices. Thus

$$0 \geq \Lambda_H(G') - \Lambda_H(G) \geq 2 \cdot n^2/4 + 4 \cdot n^2/4 - 2 \cdot n^2/2 + o(n^2) = n^2/2 + o(n^2) > 0,$$

a contradiction showing that V'_0 (and similarly V'_1) spans no edges.

Second, suppose that $(u, w) \in V'_0 \times V'_1$ is a non-edge. Here G' is obtained from G by changing uw to an edge. There are no maps that embed H into G but not into G' as each has to map a blue edge to uw leaving no choice for the third vertex of the blue triangle of H . On the other hand, the number of maps which are embeddings into G' but not G is at least $2 \cdot (n/2)^2 + o(n^2)$: map $\{0, 2\}$ to uw in some order and then map 1 and 3 into the same part as 0. This contradiction shows that all pairs between V'_0 and V'_1 are edges.

Third, suppose that $u \in V_2$ and, say, $w \in V'_0$ (and these are adjacent in G). Maps that are embeddings into G but not G' have to map $\{0, 2\}$ to uw . If 0 is mapped to w then 1, 3 have both to be mapped to V'_0 giving at most $(n/2)^2 + o(n^2)$ choices; if 0 is mapped to u then 1 and 3 have to be mapped into a same part (any) so we have at most $n^2/2 + o(n^2)$ choices here. On the other hand, all the following injections give embeddings into G' but not into G . We can map 1, 3 to uw (2 choices), send 0 to $V'_0 \setminus \{w\}$ and send 2 to V'_1 $((n/2)^2 + o(n^2))$ choices). Or we can map 3 to u , 0 to w , 1 to $V'_0 \setminus \{w\}$ and 2 to V'_1 $(n^2/4 + o(n^2))$ choices); we obtain the same bound when we map 1 to u and 0 to w . Thus

$$0 \geq \Lambda_H(G') - \Lambda_H(G) \geq 2(n^2/4) + 2(n^2/4) - (3/4)n^2 + o(n^2) = n^2/4 + o(n^2) > 0,$$

a contradiction. This proves the claim. ■

Denote $n_i := |V'_i|$ for $i = 0, 1$ and $n_2 := |V_2|$. As our next intermediate step, let us show that $n_2 \leq 30$. We do not optimise the bound since we will later show that $n_2 = 0$ (by using that $n_2 = O(1)$).

Observe that by Lemma 3.15 we have

$$\Lambda_H(n) \geq \max_{a \in [n+1]} \Lambda_H(T_{a, n-a}) = \frac{1}{8}n^{(4)} + \frac{3}{4}n^2 + O(n),$$

and the average number of embeddings per vertex in extremal G is at least $(4/n)\Lambda_H(n) = (n-1)^{(3)}/2 + O(n)$. Since we cannot increase G by removing a vertex and adding a clone of

another vertex, it holds for every vertex u that the number of embeddings that use u is at least

$$\Lambda_H(G, u) \geq \frac{(n-1)^{(3)}}{2} - 4! \binom{n}{2} + O(n) = \frac{n^3}{2} - 15n^2 + O(n). \quad (32)$$

Let us estimate $\Lambda_H(G, u)$ for any given $u \in V_2$. First, the number of the embeddings f where the 3 other vertices of H are all mapped into $V'_0 \cup V'_1$ is at most $2(n_0^{(2)}n_1 + n_0n_1^{(2)})$. Indeed, every such f has to map 1 or 3 to u and if, say, $f(1) = u$ then $f(0)$ and $f(3)$ have to be in the same part V'_i while 2 is in the other part V'_{1-i} for some $i \in \{0, 1\}$. Next, let show that the number of embeddings that use exactly one other vertex of V_2 is at most $|V_2|(n^2 + o(n^2))$. We have to sum over all choices of the other vertex $w \in V_2 \setminus \{u\}$. If uw is an edge then 0, 2 have to be mapped to u, w , and then 1, 3 are mapped to the same part V'_i (for $i = 0, 1$); thus we have at most $2n^2/2 + o(n^2) = n^2 + o(n^2)$ embeddings for such w . If uw is not an edge, then 1, 3 have to be mapped to u, w , and then 0, 2 are mapped to different parts among V'_0 and V'_1 ; again we get at most $n^2 + o(n^2)$ embeddings and the claim follows. Finally, the number of embeddings that use at least two other vertices from V_2 is at most $O(n_2^2n) = o(n_2n^2)$. Thus by (32) we have

$$\begin{aligned} 0 &\geq \left(\frac{n^3}{2} - 15n^2 \right) - \left(2 \left(n_0^{(2)}n_1 + n_0n_1^{(2)} \right) + n_2n^2 \right) + o(n_2n^2) \\ &= \frac{(n_0 - n_1)^2(n_0 + n_1)}{2} + 4n_0n_1 + \frac{(n_0 + n_1)^2(n_2 - 30)}{2} + o(n_2n^2), \end{aligned}$$

where the last identity is obtained by substituting $n = n_0 + n_1 + n_2$. This cannot be satisfied for any $n_2 \geq 31$, proving the claimed upper bound on n_2 .

Finally, we are ready to derive that $n_2 = 0$ by estimating the total number of embeddings. We group them depending on how many points of V_2 are used. As we observed before, there are $n_0n_1^{(3)} + n_0^{(3)}n_1$ embeddings that do not use V_2 at all. Next consider the case that exactly one vertex i of H is mapped into V_2 . Then $i \in \{1, 3\}$ (because of the red edge 02). If, say $i = 1$ then 0, 3 have to go to the same part V'_i and 2 has to go to the other part V'_{1-i} . Thus, by the symmetry between $1, 3 \in V(H)$, there are at most $2n_2 \left(n_0n_1^{(2)} + n_0^{(2)}n_1 \right)$ embeddings that use exactly one vertex of V_2 . Also, there are at most $\binom{n_2}{2} \cdot (n - n_2)^2 + o(n^2)$ embeddings with exactly two points in V_2 . Indeed, once we chose an unordered pair $uw \in \binom{V_2}{2}$, its adjacency uniquely determines which pair in $V(H)$ is mapped to it (namely, it is $\{0, 2\}$ if $uw \in E(G)$ and $\{1, 3\}$ otherwise) and then there are at most $(1/2 + o(1))\binom{n-n_2}{2}$ choices of a suitable unordered pair outside of V_2 , and then there are 4 ways to order these two unordered pairs. Finally, the number of embeddings with at least three vertices from V_2 is at most $O(n_2^3n) = O(n)$.

Thus, if we denote $c := (n/2 - n_0)/\sqrt{n}$ (and substitute $n_0 = n/2 + c\sqrt{n}$ and $n_1 = n/2 - c\sqrt{n} - n_2$) and use that $c = o(\sqrt{n})$ and $n_2 = O(1)$, we get that

$$\begin{aligned} 0 &\leq n_0n_1^{(3)} + n_0^{(3)}n_1 + 2n_2 \left(n_0n_1^{(2)} + n_0^{(2)}n_1 \right) + \binom{n_2}{2} \cdot (n - n_2)^2 - \frac{1}{8}n^{(4)} - \frac{3}{4}n^2 + o(n^2) \\ &= p(c, n_2)n^2 + (-4c^3n_2 - 2cn_2^2)n^{3/2} + o(n^2), \end{aligned} \quad (33)$$

where we denote $p(c, n_2) := -2c^4 + 3c^2 - 2c^2n_2 - \frac{n_2^2}{4} + \frac{3n_2}{4} - \frac{9}{8}$. It follows that $c = O(1)$; thus the second term in (33) (that is, the one at $n^{3/2}$) can be absorbed into the error term. We

conclude that $0 \leq p(c, n_2) + o(1)$. If we fix n_2 and vary c , then the derivative of the (even degree-4) polynomial of $p(c, n_2)$ has three roots: $c = 0$ and $\pm\sqrt{3 - 2n_2}$. Thus, if $n_2 \geq 2$ then the only real root is $c = 0$ and $p(c, n_2) \leq p(0, n_2) = -n_2^2/4 + 3n_2/4 - 9/8$, which is at most $p(0, 3/2) = -9/16$, a contradiction. Finally, if $n_2 = 1$ then $p(c, 1) = -2c^4 + c^2 - 5/8$ is at most $p(1/2, 1) = -1/2$, a contradiction again. Thus $n_2 = 0$, proving Theorem 3.16. ■

3.8 Graph H_5

Recall that H_5 is the red 3-edge star $T_{3,1}$ with one blue edge inserted. As we mentioned in the introduction, we have not been able to round the floating-point matrices returned by the plain flag algebra method for the λ_{H_5} -problem. However, we will show here (in Lemma 3.22) that in order to solve the semi-inducibility problem for H_5 it is enough to work with complete partite graphs only. Specifically, in Section 3.8.1 we present a sufficient criterion from [17] for a version of perfect stability to hold. Then, in Section 3.8.2, we show that it applies to the semi-inducibility problem for H_5 . The computational part (when we solve H_5 -semi-inducibility problem in the theory of complete partite graphs) is discussed in Section 3.8.3.

3.8.1 Symmetrizable problems

First, we need to give some further definitions. Given $\kappa \geq 2$ and $\gamma : \mathcal{F}_\kappa \rightarrow \mathbb{R}$, let the parameters $\lambda_\gamma(G)$, $\lambda_\gamma(n)$ and λ_γ be as defined in Section 2.2.

For a graph G and distinct vertices $u, w \in V(G)$, let the graph G_{uw} be obtained from G by making w a clone of u (thus the new neighbourhood of w is $\Gamma_G(u) \setminus \{u\}$). We call the λ_γ -problem *strongly symmetrizable* if, for every graph G and every non-edge $uw \in E(\overline{G})$, it holds that

$$2\lambda_\gamma(G) \leq \lambda_\gamma(G_{wu}) + \lambda_\gamma(G_{uw}). \quad (34)$$

It is easy to show that, if the problem is strongly symmetrizable, then it is symmetrizable, as defined in [17, Definition 1]. Since the latter definition is rather long (in order to apply for a wider class of problem), we do not state it here. Instead, we verify, for example, the first part of [17, Definition 1], namely that one can transform an arbitrary graph G_0 on $[n]$ into a complete partite graph in steps so that each step changes only $o(n^2)$ pairs (in fact, it will be at most $n - 1$ pairs here) and does not decrease the objective function. Inductively on i , we construct a graph G_i whose restriction to $[i - 1]$ is complete partite. Suppose that $i \geq 1$ and have already have G_{i-1} . If i is connected to every vertex of $[i]$ in G_{i-1} then we take $G_i := G_{i-1}$ (with $\{i\}$ forming a new separate part), so suppose otherwise. Take any $j \in [i]$ non-adjacent to i in G_{i-1} . If $\lambda_\gamma((G_{i-1})_{ji}) \geq \lambda_\gamma(G_{i-1})$ then we let $G_i := (G_{i-1})_{ji}$ (where we replace i by a clone of j in G_{i-1}). So suppose that $\lambda_\gamma((G_{i-1})_{ji}) < \lambda_\gamma(G_{i-1})$. By (34) we have that $\lambda_\gamma((G_{i-1})_{ij}) > \lambda_\gamma(G_{i-1})$. Let $J = \{j_0, \dots, j_{m-1}\}$ denote the set of G_{i-1} -clones of j that lie inside $[i]$. We get G_i in m steps where in the t -th step we symmetrise j_t into i in the current graph. By (34) and induction on t , the t -th step increases the objective function (because the opposite symmetrisation decreases it), finishing the proof. A direct consequence

is that, for every integer $n \geq \kappa$, at least one extremal graph attaining the value of $\lambda_\gamma(n)$ is complete partite.

Let $\mathcal{P}$ denote the set of all complete partite graphs (up to isomorphism) and let $\overline{\mathcal{P}}$ be the set of sequences $(x_i)_{i=1}^\infty$ such that $x_1 \geq x_2 \geq \dots$, each x_i is non-negative and $\sum_{i=1}^\infty x_i \leq 1$. There is a natural map that sends a complete m -partite graph $T_{n_1, \dots, n_m}$ with $n_1 \geq \dots \geq n_m$ to the sequence $(n_1/n, \dots, n_m/n, 0, \dots) \in \overline{\mathcal{P}}$, where $n := n_1 + \dots + n_m$. The set $\overline{\mathcal{P}}$ with the pointwise convergence can be viewed as the limit space for complete partite graphs and its elements can be identified with graphons of a very special form (namely, those for which the induced density of the 3-vertex 1-edge graph $\overline{P}_3$ is zero). We refer the reader to [17] for further details. Observe that $x_0 := 1 - \sum_{i=1}^\infty x_i$ may be positive in the limit; combinatorially, it refers for the total proportion of vertices in parts whose relative size tends to 0. Also, the following definition gives, for every element $\mathbf{x} \in \overline{\mathcal{P}}$ a sequence of complete partite graphs that converges to $\mathbf{x}$.

Definition 3.19 *Given $n \in \mathbb{N}$ and $\mathbf{x} = (x_1, x_2, \dots)$ with $x_1 \geq x_2 \geq \dots \geq 0$ and $x_0 := 1 - \sum_{i \geq 1} x_i \geq 0$, define a complete partite graph $G_{n, \mathbf{x}}$ with vertex set $[n]$, parts $V_1, \dots, V_m$ for some m and a set V_0 of universal vertices, i.e., $|V_0|$ singleton parts, as follows. If $x_0 = 0$, take a partition $[n] = V_1 \cup \dots \cup V_m$ with $||V_i| - x_i n| < 1$ and let $V_0 = \emptyset$. Otherwise, for all $i \geq 1$ with $x_i n \geq 2$, let $|V_i| = \lfloor x_i n \rfloor$ and let V_0 consist of the remaining vertices in $[n]$.*

Next, we define a version of perfect stability from [17].

Definition 3.20 (Perfect $\mathcal{P}$ -Stability) *The λ_γ -problem is perfectly $\mathcal{P}$ -stable if there exists $C > 0$ such that for every graph G of order $n \geq C$ there is a complete partite graph G' of the same order n such that*

$$\delta_{\text{edit}}(G, G') \leq C(\lambda_\gamma(n) - \lambda_\gamma(G))n^2.$$

Note that this property, although very closely related to our definition of perfect stability, is slightly different since the number of parts of G' is not specified.

Before we can state the sufficient condition for $\mathcal{P}$ -stability from [17], we need to give some further definitions. Define

$$\lambda_\gamma(\mathbf{x}) := \lim_{n \rightarrow \infty} \lambda_\gamma(G_{n, \mathbf{x}}) \quad (35)$$

and let $\text{OPT}(\lambda_\gamma)$ be the set of $\mathbf{x} \in \overline{\mathcal{P}}$ that maximise this function. One can show that the limit in (35) exists and write an explicit formula for $\lambda_\gamma(\mathbf{x})$ (which is a polynomial of $\mathbf{x}$ if $\mathbf{x}$ has finite support and an absolutely convergent power series otherwise). Also, by the compactness of $\overline{\mathcal{P}}$ and the continuity of the function $\lambda_\gamma : \overline{\mathcal{P}} \rightarrow \mathbb{R}$, the set $\text{OPT}(\lambda_\gamma)$ is non-empty.

In the notation of Definition 3.19, if G' is a graph obtained by adding a new vertex z to $G = G_{n, \mathbf{x}}$, we say that z is a $\mathcal{P}$ -clone of $u \in V(G)$ if $u \in V_0$ and $\Gamma_{G'}(z) = V(G)$, or if $u \notin V_0$ and $\Gamma_{G'}(z) = \Gamma_G(u)$.

Theorem 3.21 ([17, Theorem 1.1]) *Let λ_γ be a symmetrisable function such that the set $\text{OPT}(\lambda_\gamma)$ has finitely many elements. Suppose also that there exists $c > 0$ such that the following hold for every large n and every maximiser $\mathbf{x} \in \text{OPT}(\lambda_\gamma)$, where $G := G_{n, \mathbf{x}}$.*

- (i) For all distinct $u, w \in V(G)$ we have $\lambda_\gamma(G) - \lambda_\gamma(G \oplus uw) \geq cn^{-2}$.
- (ii) If G_u is obtained from G by adding a new vertex u which is complete or empty to each part of G (where each V_i , $i \in [m]$ is a part and we have $|V_0|$ singleton parts) then the minimum number of edits at u needed to make u a $\mathcal{P}$ -clone of some existing vertex of G is at most $n(\lambda_\gamma(G) - \lambda_\gamma(G_u, u))/c$.

Then the λ_γ -problem is perfectly $\mathcal{P}$ -stable.

Note that if every $\mathbf{x} \in \text{OPT}(\lambda_\gamma)$ has $m < \infty$ non-zero terms and satisfies $\sum_{i=1}^m x_i = 1$ then Property (i) of Theorem 3.21 is the same as our λ_γ -flip-averseness of K_m while Property (ii) can be shown to be implied by the λ_γ -strictness of K_m . (Note that when we attach a new vertex u in Property (ii), we consider only those attachments where, for each part V_i , the pairs between u and V_i are all edges or all non-edges; Property (ii) relates the contribution of u to the number of edits so that it handles correctly the general case in [17] when $\mathbf{x}$ may have infinitely many positive entries.)

3.8.2 Sufficient condition for a semi-inducibility problem to be symmetrisable

The following lemma gives a sufficient condition for the H -semi-inducibility problem to be symmetrisable. It generalises the analogous inducibility result of Brown and Sidorenko [8] which is a special case when H is a 2-colouring of a complete graph. It is proved by an easy adaptation of the argument from [8].

Lemma 3.22 *Let H be a blue-red edge-coloured graph such that the red graph is the complete partite with parts $R_0 \cup \dots \cup R_{m-1}$ and, for every $i \in [m]$ and every two vertices $u, w \in R_i$, the blue neighbourhoods of u and w in $R_i \setminus \{u, w\}$ are nested (that is, one is a subset of the other, allowing equality). Then the H -semi-inducibility problem is strongly symmetrisable.*

Proof. For a graph G , let $\text{Hom}(H, G)$ be the set of embeddings of H into G . Recall that these are injections $V(H) \rightarrow V(G)$ that map red (resp. blue) edges to edges (resp. non-edges) of G . Thus $|\text{Hom}(H, G)| = \Lambda_H(G)$.

Take any graph G and $xy \in E(\overline{G})$. In order to prove (34) it is enough to prove that, for every distinct $u, w \in V(H)$ and any injection $g : V(H) \setminus \{u, w\} \rightarrow V(G) \setminus \{x, y\}$, we have

$$2 |\text{Hom}(H, G) \cap \text{Ext}(g)| \leq |\text{Hom}(H, G_{xy}) \cap \text{Ext}(g)| + |\text{Hom}(H, G_{yx}) \cap \text{Ext}(g)|, \quad (36)$$

where $\text{Ext}(g)$ consists of all *extensions* f of g , that is, injective maps $V(H) \rightarrow V(G)$ such that $f(v) = g(v)$ for every $v \in V(H) \setminus \{u, w\}$. Indeed, if we sum (36) over all unordered pairs $uw \in \binom{V(H)}{2}$ and all injections $g : V(H) \setminus \{u, w\} \rightarrow V(G) \setminus \{x, y\}$ then every embedding f from H into any one of G , G_{xy} or G_{yx} is counted exactly $\binom{v(H)}{2}$ times.

Take any extension f of g . Let ϕ be the identity map on $V(G)$, except it swaps x and y . Note that ϕ is an involution, that is, $\phi \circ \phi$ is the identity map.

If the image of f is disjoint from $\{x, y\}$ and f is an embedding of H into G then it is trivially an embedding of H into each of G_{xy} and G_{yx} , so it contributes 2 to both sides of (36), as desired.

Next, let us consider those extensions $f \in \text{Ext}(g)$ such that the image of f contains exactly one of x and y . If f is an embedding of H into G and, say, $f(u) = x$ (and thus $f(w) \neq y$) then both f and the composition $\phi \circ f$ are two different embeddings into G_{xy} (since y is a clone of x in G_{xy}), contributing 2 to the right-hand side of (36). We do not overcount in the right-hand side: if $\phi \circ f$ also happens to be an embedding into G , then the contribution of $\phi \circ f$ is also counted for G_{yx} (and the pair $\{f, \phi \circ f\}$ contributes 4 to both sides of (36)).

Finally, it remain to look at extensions f whose image contains both x and y . Since ϕ is an involution, all such maps are partitioned into pairs $\{f, \phi \circ f\}$. Consider the contribution of any such pair. There is nothing to prove if none of f and $\phi \circ f$ is an embedding of H into G . Suppose that both f and $\phi \circ f$ are embeddings of H into G . By $xy \in E(\overline{G})$, the pair uw is not a red edge of H . Thus u and w are in the same part R_i and have the same red neighbourhoods in H . Also, for every vertex $v \in V(H) \setminus \{u, w\}$, which is a blue neighbour of at least one of u and w , we have that $f(v) \in \Gamma_{\overline{G}}(x) \cap \Gamma_{\overline{G}}(y)$. Thus both f and $\phi \circ f$ are embeddings into both G_{xy} and G_{yx} , and pair $\{f, \phi \circ f\}$ contributes 4 to both sides of (36). Finally, suppose that f is an embedding of H into G but $\phi \circ f$ is not. By the former property and by $xy \in E(\overline{G})$, there is $i \in [m]$ such that $u, w \in R_i$. Thus u and w have the same red neighbourhoods in H and nested blue neighbourhoods in $R_i \setminus \{u, w\}$, say that of u being a subset of that of w . Thus the constraints on the possible image of $f(u)$ (for $f \in \text{Ext}(g)$) are weaker than those for $f(w)$. We conclude that both f and $\phi \circ f$ are embeddings into $G_{f(w)f(u)}$. Thus the contribution of $\{f, \phi \circ f\}$ to the right-hand size of (36) is at least 2, as desired. This finishes the proof of the lemma. ■

3.8.3 Calculations for H_5

Clearly, Lemma 3.22 applies when $H = H_5$ is the red 3-star with one blue edge inserted. Thus, in order to determine the semi-inducibility constant of H , it is enough to solve the problem for complete partite graphs. While one should in principle be able to determine the optimal number of parts (5 in this case) and the optimal limiting ratios (the vector $(\alpha, \dots, \alpha, (1 - 4\alpha))$ with $\alpha = (13 - \sqrt{57})/56$) by hand, this task seems still quite messy. Instead, we recourse to flag algebras.

Theorem 3.23 *Let $H = H_5$ be the graph on $[4]$ with the red edge set $\{01, 02, 03\}$ and the blue set $\{13\}$. Then the maximum value of $\lambda_H(G)$ over complete partite graphs G of order $n \rightarrow \infty$ is $(23637 - 513\sqrt{57})/10976 + o(1)$. Moreover, there is C such that for every complete partite graph G of order $n \geq 1/C$ there is a complete 5-partite graph G' with*

$$\delta_{\text{edit}}(G, G') \leq C(\lambda_\gamma(n) - \lambda_\gamma(G))n^2. \quad (37)$$

Proof. The first part of the theorem is proved by the standard flag algebra approach with

$N = 8$ within the theory of complete partite graphs. This is a hereditary family of graphs (characterised by forbidding $\overline{P}_3$ as an induced subgraph) so flag algebras can be applied to it.

Let us show how we derive the second part of the theorem from the obtained certificate.

Our certificate satisfies that the slack c_F is positive for every 8-vertex complete partite graph F with more than 5 parts. Hence up to removing $o(n)$ vertices, every almost extremal complete partite graph is 5 partite. Also, it satisfies [21, Theorem 7.1], including its Condition (i), with $\tau = K_4$. (The statement of [21, Theorem 7.1] allows for any hereditary family of graphs which is closed under taking blowups, in particular it applies to $\mathcal{P}$.) Thus [21, Theorem 7.1] implies in particular that there is the unique (up to an automorphism of K_5 , that is, any permutation of indices) vector $\mathbf{x} \in \mathbb{S}_5$ that maximises $\lambda_H(K_5(\mathbf{x}))$. Thus $\mathbf{a} := (\alpha, \alpha, \alpha, \alpha, 1 - 4\alpha)$ is the unique, up to a permutation, (λ_{H_5}, K_5) -optimal vector.

The claim about the existence of C is another conclusion of [21, Theorem 7.1], namely that the perfect K_5 -stability property (defined in the obvious way) holds within the theory of complete partite graphs. ■

Combining Lemma 3.22 and Theorem 3.23, we can obtain the upper bound (and thus the value of λ_{H_5}) stated in Table 1:

Theorem 3.24 *It holds that $\lambda_{H_5} = (23637 - 513\sqrt{57})/10976$ and the semi-inducibility problem for H_5 is perfectly K_5 -stable.*

Proof. This extremal problem is symmetrisable by Lemma 3.22. Furthermore, Theorem 3.23 gives that the complete partite version of the problem has unique maximiser in the limit space $\overline{\mathcal{P}}$, namely the sequence made from $\mathbf{a} := (\alpha, \alpha, \alpha, \alpha, 1 - 4\alpha)$ by attaching zeros. This corresponds taking blowups of $B := K_5$.

Next, we would like to check the flip averseness and Condition (ii) of Theorem 3.21 for λ_{H_5} . While this task is straightforward, the number of cases is considerable: up to symmetries we have to check 4 possible flip pairs (since there are 2 different part ratios) and 10 possible attachments (namely, the new vertex can be fully attached to any number $i \in [5]$ of parts of ratio α times the two options of being full or empty to the remaining part of ratio $1 - 4\alpha$). So we delegate this task to computer and provide scripts that verify both flip averseness and the strictness using exact arithmetic. Recall that strictness implies Condition (ii) of Theorem 3.21. Thus Theorem 3.21 applies here and gives that the semi-inducibility problem for H_5 is perfectly $\mathcal{P}$ -stable. This, when combined with (37), gives that the problem is in fact perfectly K_5 -stable. ■

4 Concluding remarks

Let us remark that, although we proved that the semi-inducibility problem for H_4 is perfectly K_3 -stable in Section 3.4, our flag algebra certificate with $N = 4$ does not satisfy the sufficient condition of [21, Theorem 7.1]. (Namely the 4-vertex clique K_4 does not admit a homomorphism into the base graph K_3 but has zero slack.) However, the information contained in the certificate

is enough to give alternative proofs of Claims 3.6 and 3.7. For example, every 4-vertex graph F which is not complete partite has positive slack c_F in (15), which implies that the density of F in every almost extremal graph G is $o(1)$. Since (every 4-vertex extension of) $\overline{P}_3$ is not complete partite, Claim 3.6 (which is equivalent to the induced density of $\overline{P}_3$ being $o(1)$) follows.

Flag algebras worked remarkably well for the semi-inducibility of non-complete 4-vertex graphs. Recall that if we are interested only in the value of the semi-inducibility constant λ_{H_i} , $0 \leq i < 18$, then the sharp upper bounds can be determined by this method for every i except perhaps $i = 3$ and $i = 5$. (In fact, it is possible that the method also works for these two cases but we were just unable to round the floating-point matrices returned by the solver.) One informal explanation for such efficiency is that, since $\{2, 3\}$ is not an edge of H_i , we can compute $\Lambda_{H_i}(G)$ up to error $O(n^3)$ by summing over all suitable $f(0), f(1) \in V(G)$ of the product of the number of choices for $f(2)$ and $f(3)$. This can be viewed as some scalar product of the vectors of 3-vertex subgraph densities in G rooted on $(f(0), f(1))$. Since positive semi-definite matrices appear naturally for problems involving scalar products (e.g. as Gram matrices), this gives some intuition why the flag algebra method is so successful here.

Using flag algebras, the first author has obtained a number of exact results on the semi-inducibility constant for 5-vertex graphs. These findings will appear in a subsequent paper.

We should also mention that the computer-generated certificates obtained early in the project were quite helpful in finding matching lower bounds as well as in proving in a computer-free way some of the results (once the authors were confident of their validity). This approach may be useful for other problems. The first author made publicly available not only the verifier code but also the code that was used for generating the certificates. It should not be hard to adapt it to apply flag algebra calculations to other suitable extremal problems. We hope that this package will be generally useful to the research community.

Acknowledgements

Both authors were supported by ERC Advanced Grant 101020255.

References

- [1] B. M. Ábrego, S. Fernández-Merchant, M. G. Neubauer, and W. Watkins, *Sum of squares of degrees in a graph*, JIPAM. J. Inequal. Pure Appl. Math. **10** (2009), Article 64, 34.
- [2] R. Ahlswede and G. O. H. Katona, *Graphs with maximal number of adjacent pairs of edges*, Acta Math. Acad. Sci. Hungar. **32** (1978), 97–120.
- [3] N. Alon, E. Fischer, M. Krivelevich, and M. Szegedy, *Efficient testing of large graphs*, Combinatorica **20** (2000), 451–476.
- [4] T. Andreescu, D. Andrica, and I. Cucurezeanu, *An introduction to Diophantine equations*, Birkhäuser Verlag, New York, 2010. A problem-based approach.
- [5] R. Baber and J. Talbot, *Hypergraphs do jump*, Combin. Probab. Computing **20** (2011), 161–171.
- [6] A. Basit, B. Granet, D. Horsley, A. Kündgen, and K. Staden, *The semi-inducibility problem*, 2025. E-print arxiv:2501.09842.

- [7] L. Bodnár and O. Pikhurko, *Some exact inducibility-type results for graphs via flag algebras*, 2025. E-print arxiv:2507.01596.
- [8] J. I. Brown and A. Sidorenko, *The inducibility of complete bipartite graphs*, J. Graph Theory **18** (1994), 629–645.
- [9] O. D. Byer, *Two path extremal graphs and an application to a Ramsey-type problem*, Discrete Math. **196** (1999), no. 1-3, 51–64.
- [10] H. Chen, F. C. Clemen, and J. A. Noel, *Maximizing alternating paths via entropy*, 2025. E-print arxiv:2505.03903.
- [11] H. Chen and J. A. Noel, *On alternating 6-cycles in edge-coloured graphs*, 2025. E-print arxiv:2505.09809.
- [12] M. K. de Carli Silva, F. M. de Oliveira Filho, and C. M. Sato, *Flag algebras: a first glance*, Nieuw Arch. Wiskd. (5) **17** (2016), 193–199.
- [13] P. Erdős, *On bipartite subgraphs of graphs*, Mat. Lapok **18** (1967), 283–288. In Hungarian.
- [14] P. Erdős and T. Gallai, *Graphs with prescribed degree of vertices*, Mat. Lapok **11** (1960), 264–274.
- [15] Z. Füredi, *A proof of the stability of extremal graphs, Simonovits’ stability from Szemerédi’s regularity*, J. Combin. Theory (B) **115** (2015), 66–71.
- [16] S. Gilboa, R. Glebov, D. Hefetz, N. Linial, and A. Morgenstern, *On the local structure of oriented graphs—a case study in flag algebras*, Electron. J. Combin. **29** (2022), Paper No. 3.39, 53.
- [17] H. Liu, O. Pikhurko, M. Sharifzadeh, and K. Staden, *Stability from graph symmetrisation arguments with applications to inducibility*, J. Lond. Math. Soc. (2) **108** (2023), 1121–1162.
- [18] S. Norin and L. Yepremyan, *Turán number of generalized triangles*, J. Combin. Theory (A) **146** (2017), 312–343.
- [19] S. Norin and L. Yepremyan, *Turán numbers of extensions*, J. Combin. Theory (A) **155** (2018), 476–492.
- [20] U. N. Peled, R. Petreschi, and A. Sterbini, *(n, e) -graphs with maximum sum of squares of degrees*, J. Graph Theory **31** (1999), 283–295.
- [21] O. Pikhurko, J. Šliachan, and K. Tyros, *Strong forms of stability from flag algebra calculations*, J. Combin. Theory (B) **135** (2019), 129–178.
- [22] N. Pippenger and M. C. Golumbic, *The inducibility of graphs*, J. Combin. Theory (B) **19** (1975), 189–203.
- [23] A. Razborov, *Flag algebras*, J. Symb. Logic **72** (2007), 1239–1282.
- [24] A. Razborov, *On 3-hypergraphs with forbidden 4-vertex configurations*, SIAM J. Discr. Math. **24** (2010), 946–963.
- [25] A. Roberts and A. Scott, *Stability results for graphs with a critical edge*, Europ. J. Combin. **74** (2018), 27–38.
- [26] M. Simonovits, *A method for solving extremal problems in graph theory, stability problems*, Theory of Graphs (Proc. Colloq., Tihany, 1966), 1968, pp. 279–319.