THE FEASIBLE REGION OF INDUCED GRAPHS

XIZHI LIU, DHRUV MUBAYI, AND CHRISTIAN REIHER

Abstract. The feasible region $\Omega_{\text{ind}}(F)$ of a graph F is the collection of points (x, y) in the unit square such that there exists a sequence of graphs whose edge densities approach x and whose induced F-densities approach y. A complete description of $\Omega_{\text{ind}}(F)$ is not known for any F with at least four vertices that is not a clique or an independent set. The feasible region provides a lot of combinatorial information about F. For example, the supremum of y over all $(x, y) \in \Omega_{\text{ind}}(F)$ is the inducibility of F and $\Omega_{\text{ind}}(K_r)$ yields the Kruskal-Katona and clique density theorems.

We begin a systematic study of $\Omega_{\text{ind}}(F)$ by proving some general statements about the shape of $\Omega_{\text{ind}}(F)$ and giving results for some specific graphs F. Many of our theorems apply to the more general setting of quantum graphs. For example, we prove a bound for quantum graphs that generalizes an old result of Bollobás for the number of cliques in a graph with given edge density. We also consider the problems of determining $\Omega_{\text{ind}}(F)$ when $F = K_r^-$, F is a star, or F is a complete bipartite graph. In the case of K_r^- our results sharpen those predicted by the edge-statistics conjecture of Alon et. al. while also extending a theorem of Hirst for K_4^- that was proved using computer aided techniques and flag algebras. The case of the 4-cycle seems particularly interesting and we conjecture that $\Omega_{\text{ind}}(C_4)$ is determined by the solution to the triangle density problem, which has been solved by Razborov.

§1. Introduction

1.1. **Feasible regions.** Given a graph G denote by $V(G)$ and $E(G)$ the vertex set and edge set of G respectively. Let $v(G) = |V(G)|$, $e(G) = |E(G)|$, and call $\varrho(G) = e(G)/\binom{v(G)}{2}$ the *edge density* of G. For two graphs F and G denote by $N(F, G)$ the number of induced copies of F in G, and let $\varrho(F, G) = N(F, G)/\binom{v(G)}{v(F)}$ be the *induced F-density* of G.

A *quantum graph* Q is a formal linear combination of finitely many graphs, i.e., an expression of the form

$$Q = \sum_{i=1}^{m} \lambda_i F_i\,,$$

where m is a nonnegative integer, the numbers $\lambda_1, \ldots, \lambda_m$ are real, and $F_1, \ldots, F_m$ are graphs. We call F_i a *constituent* of Q if $\lambda_i \neq 0$. Two quantum graphs Q, Q' are equal if

Key words and phrases. inducibility of graphs, feasible regions.

The first and second author's research is partially supported by NSF awards DMS-1763317 and DMS-1952767.

they have the same constituents and the same (nonzero) coefficients for each constituent. The *complement* of Q is $\overline{Q} = \sum_{i=1}^{m} \lambda_i \overline{F}_i$, where $\overline{F}_i$ denotes the complement of F_i for each $i \in [m]$. A quantum graph Q is *self-complementary* if $Q = \overline{Q}$. Every graph parameter f can be extended linearly to quantum graphs by stipulating $f(Q) = \sum_{i=1}^{m} \lambda_i f(F_i)$. In particular,

$$N(Q,G) = \sum_{i=1}^{m} \lambda_i N(F_i, G) \quad \text{and} \quad \varrho(Q,G) = \sum_{i=1}^{m} \lambda_i \varrho(F_i, G)\,.$$

The main notion investigated in this article is the following.

Definition 1.1 (Feasible region)**.** *Let $Q = \sum_{i=1}^{m} \lambda_i F_i$ be a quantum graph.*

- *A sequence $(G_n)_{n=1}^{\infty}$ of graphs is Q-*good *if $\lim_{n\to\infty} v(G_n) = \infty$, $\lim_{n\to\infty} \varrho(G_n)$ exists, and for every $i \in [m]$ the limit $\lim_{n\to\infty} \varrho(F_i, G_n)$ exists.*
- *A Q-good sequence of graphs $(G_n)_{n=1}^{\infty}$* realizes *a* point *$(x, y) \in [0,1] \times \mathbb{R}$ if*

$$\lim_{n\to\infty} \varrho(G_n) = x \quad \text{and} \quad \lim_{n\to\infty} \varrho(Q, G_n) = y.$$

- *The* feasible region *$\Omega_{\mathrm{ind}}(Q)$ of (induced) Q is the collection of points $(x, y) \in [0,1] \times \mathbb{R}$ realized by some Q-good sequence $(G_n)_{n=1}^{\infty}$.*

We commence a systematic study of the feasible region of quantum graphs Q. As we shall see soon, $\Omega_{\mathrm{ind}}(Q)$ is determined by its boundary, so it suffices to consider for every $x \in [0,1]$ the numbers

$$i(Q,x) = \inf\{y \colon (x,y) \in \Omega_{\mathrm{ind}}(Q)\} \qquad \text{and} \qquad I(Q,x) = \sup\{y \colon (x,y) \in \Omega_{\mathrm{ind}}(Q)\}\,.$$

Determining the values of $i(Q,x)$ and $I(Q,x)$ under some constraints is a central topic in extremal combinatorics. For example, the classical Kruskal-Katona theorem [13, 14] implies

$$I(K_r, x) = x^{r/2} \quad \text{for all } r \geqslant 2 \text{ and } x \in [0,1]\,.$$

Turán's seminal theorem [27] and supersaturation show that for every integer $r \geqslant 3$,

$$i(K_r, x) > 0 \qquad \Longleftrightarrow \qquad x > (r-2)/(r-1)\,.$$

Determining $i(K_r, x)$ for all $x > (r-2)/(r-1)$ is highly nontrivial and was solved for $r = 3$ by Razborov [22], for $r = 4$ by Nikiforov [19], and for all r by the third author [23].

Regarding quantum graphs with at least two constituents, a classical result of Goodman [8] says that $i(K_3 + \overline{K}_3, x) \geqslant 1/4$ and equality holds only for $x = 1/2$. Erdős [5] conjectured that $i(K_r + \overline{K}_r, x) \geqslant 2^{1-\binom{r}{2}}$ for $r \geqslant 4$ with equality for $x = 1/2$. This conjecture was disproved by Thomason [26] for all $r \geqslant 4$, but even for $r = 4$ the minimum value of $i(K_r + \overline{K}_r, x)$ is still unknown.

For a single graph F the function $I(F,x)$ is closely related to the *inducibility*

$$\operatorname{ind}(F) = \lim_{n\to\infty} \max\{\varrho(F,G)\colon v(G) = n\}$$

of F introduced by Pippenger and Golumbic [21]. In fact, $\operatorname{ind}(F) = \max\{I(F,x)\colon x \in [0,1]\}$, where the maximum exists due to the continuity of $I(F,x)$ (see Theorem 1.2 below).

Determining the feasible region $\Omega_{\text{ind}}(F)$ of a single graph F is a special case of the more general problem to determine the *graph profile* $T(\mathcal{F})$ of a given finite family of graphs $\mathcal{F} = \{F_1,\dots,F_k\}$. Here $T(\mathcal{F}) \subseteq [0,1]^k$ is the collection of limit points of $((\varrho(F_1,G_i),\dots,\varrho(F_k,G_i)))_{i=1}^{\infty}$ with $v(G_i) \to \infty$. Besides the clique density theorem, very few results are known about graph profiles (see [4, 9, 10, 12]).

Our results are of two flavors.

- We prove some general results about the shape of $\Omega_{\text{ind}}(Q)$. Our main result here is Theorem 1.2, which states that $I(Q,x)$ and $i(Q,x)$ are continuous and almost everywhere differentiable.
- We study $\Omega_{\text{ind}}(Q)$ for some specific choices of Q for which $\operatorname{ind}(Q)$ has been investigated by many researchers. We focus on quantum graphs whose constituents are complete multipartite graphs and prove a general upper bound for $I(Q,x)$. Prior to this work, $\Omega_{\text{ind}}(F)$ for a single graph F was determined only when F is a clique or an independent set. Here we extend this to the case $F = K_{1,2}$ and also obtain results for complete bipartite graphs. Furthermore we study $\Omega_{\text{ind}}(K_r^-)$, where K_r^- arises from the clique K_r by the deletion of a single edge. As a consequence of our results, we determine the inducibility $\operatorname{ind}(K_r^-)$, which is new for $r \geqslant 5$.

1.2. **General results.** The following result describes the shape of the feasible region of an arbitrary quantum graph.

Theorem 1.2. *For every quantum graph Q we have*

$$\Omega_{\text{ind}}(Q) = \big\{(x,y) \in [0,1]\times\mathbb{R}\colon i(Q,F) \leqslant y \leqslant I(Q,F)\big\}\,.$$

Moreover, the boundary functions $i(Q,x)$ and $I(Q,F)$ are continuous and almost everywhere differentiable.

In contrast to Theorem 1.2 Hatami and Norin [9] gave an example of a finite family $\mathcal{F}$ of graphs such that the intersection of the graph profile $T(\mathcal{F})$ with some hyperplane has a nowhere differentiable boundary.

For every quantum graph Q the feasible regions of Q, $-Q$ and $\overline{Q}$ are closely related. Indeed, using the formulae

$$N(F,G) = N(\overline{F},\overline{G}) \quad \text{and} \quad \varrho(F,G) = \varrho(\overline{F},\overline{G}),$$

which are valid for all graphs F and G, one easily confirms the following observation.

Fact 1.3. *Let Q be a quantum graph.*

(*a*) *The feasible regions of Q and $-Q$ are symmetric to each other about the x-axis. Hence, $I(-Q,x) = -i(Q,x)$ and $i(-Q,x) = -I(Q,x)$ hold for all $x \in [0,1]$.*

(*b*) *The feasible regions of Q and $\overline{Q}$ are symmetric to each other about the line $x = 1/2$. Thus we have $I(Q,x) = I(\overline{Q}, 1-x)$ and $i(Q,x) = i(\overline{Q}, 1-x)$ for every $x \in [0,1]$. In particular, if Q is self-complementary, then $I(Q,x) = I(Q,1-x)$ and $i(Q,x) = i(Q,1-x)$, i.e. the functions $I(Q,x)$ and $i(Q,x)$ are symmetric around $x = 1/2$.* □

The next result shows that for most single graphs F the lower boundary function $i(F,x)$ vanishes identically. The only exceptions occur when F is a clique and $i(F,x)$ is given by the clique density theorem (see Theorem 1.10), or if F is the complement of a clique and $i(F,x)$ is given by the Kruskal-Katona theorem (and Fact 1.3 (*b*)).

Proposition 1.4. *If F denotes a graph which is neither complete nor empty, then $i(F,x) = 0$ for all $x \in [0,1]$.*

We proceed with some estimates based on random graphs. Given a quantum graph $Q = \sum_{i=1}^{m} \lambda_i F_i$ we define

$$\operatorname{rand}(Q,x) = \sum_{i \in [m]} \lambda_i \frac{(v(F_i))!}{|\operatorname{Aut}(F_i)|} x^{e(F_i)} (1-x)^{e(\overline{F}_i)} \qquad \text{for every } x \in [0,1]\,,$$

where $\operatorname{Aut}(F_i)$ is the automorphism group of F_i for $i \in [m]$. Equivalently,

$$\operatorname{rand}(Q,x) = \lim_{n \to \infty} \mathbb{E}\, \varrho(Q, G(n,x))\,,$$

where $G(n,x)$ denotes the standard binomial random graph. It is well known that the random variables $\varrho(G(n,x))$, $\varrho(Q, G(n,x))$ are tightly concentrated around their expectations. This shows the following observation.

Fact 1.5. *If Q denotes a quantum graph and $x \in [0,1]$, then*

$$I(Q,x) \geqslant \operatorname{rand}(Q,x) \geqslant i(Q,x)\,.$$

In particular, for a single graph F the inequality $I(F,x) > 0$ holds for all $x \in (0,1)$. □

Let $P_{4,1}$ be the 5-vertex graph that is the disjoint union of a path on 4 vertices and an isolated vertex. It was asked in [6] whether the inducibility of some graph is achieved by a random graph and, in particular, whether the inducibility $\operatorname{ind}(P_{4,1})$ is achieved by the Erdős-Rényi random graph $G(n, 3/10)$. Here we pose an easier question of a similar flavor.

Problem 1.6. *Do there exist a graph F and some $x \in (0,1)$ such that $I(F,x) = \operatorname{rand}(F,x)$?*

1.3. **Complete multipartite graphs.** We now present our results on $I(Q,x)$ for specific quantum graphs Q. Our focus is on quantum graphs whose constituents are complete multipartite graphs (a graph whose edge set is empty is viewed as complete multipartite with only one part). A case of particular interest is $Q = K_r + \overline{K}_r$ for $r \geqslant 3$. Goodman [8] proved that for every graph G on n vertices $\varrho(K_3 + \overline{K}_3, G) \geqslant 1/4 + o(1)$ and the random graph $G(n,1/2)$ shows that this bound is tight. Therefore, $i(K_3 + \overline{K}_3, x) \geqslant 1/4$ and equality holds when $x = 1/2$. Combining Goodman's result [8] with a theorem of Olpp [20] one can determine $\Omega_{\mathrm{ind}}(K_3 + \overline{K}_3)$ completely.

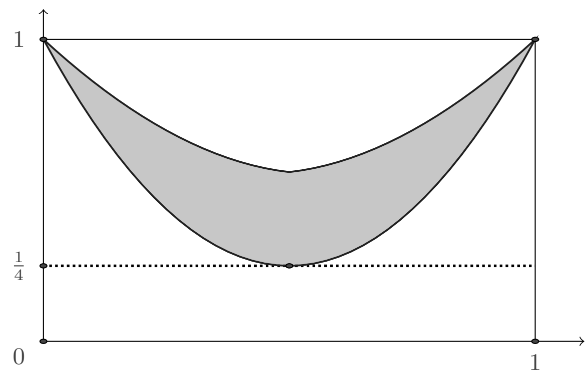

FIGURE 1.1. $\Omega_{\mathrm{ind}}(K_3 + \overline{K}_3)$ is the shaded area above.

Theorem 1.7 (Goodman [8], Olpp [20]). *For every $x \in [0,1]$ we have*

$$\begin{aligned} i(K_3 + \overline{K}_3, x) &= 1 - 3x + 3x^2 \quad \text{and} \\ I(K_3 + \overline{K}_3, x) &= 1 - 3\min\left\{x - x^{3/2}, (1-x) - (1-x)^{3/2}\right\}. \end{aligned}$$ □

For $r \geqslant 4$ determining $\Omega_{\mathrm{ind}}(K_r + \overline{K}_r)$ seems beyond current methods.

Problem 1.8. *Determine $\Omega_{\mathrm{ind}}(K_r + \overline{K}_r)$ for $r \geqslant 4$.*

Another well-studied problem concerns the determination of $\Omega_{\mathrm{ind}}(K_r)$ for $r \geqslant 3$. We already mentioned that $I(K_r, x) = x^{r/2}$ follows from the Kruskal-Katona theorem [13, 14]. For the lower bound $i(K_r, x)$ we consider (independently of r) the following complete multipartite graphs.

Construction 1.9. *For integers $n \geqslant k \geqslant 2$ and real $x \in \left(\frac{k-2}{k-1}, \frac{k-1}{k}\right]$ let $H^\star(n,x)$ be the complete k-partite graph on n vertices with parts $V_1, \dots, V_k$ of sizes $|V_1| = \dots = |V_{k-1}| = \lfloor \alpha_k n \rfloor$ and $|V_k| = n - (k-1)\lfloor \alpha_k n \rfloor$, where*

$$\alpha_k = \frac{1}{k}\left(1 + \sqrt{1 - \frac{k}{k-1}x}\right).$$

Moreover, $H^\star(n,0)$ and $H^\star(n,1)$ denote the empty and the complete graph on n vertices.

One checks immediately that $\lim_{n\to\infty} \varrho(H^\star(n,x)) = x$ holds for every $x \in [0,1]$. Consequently, for every $r \geqslant 2$ the function $g_r(x) = \lim_{n\to\infty} \varrho(K_r, H^\star(n,x))$ is an upper bound on $i(K_r, x)$.

A more explicit description of g_r is as follows. Clearly $g_r(x) = 0$ holds for every $x \leqslant \frac{r-2}{r-1}$ and $g(1) = 1$. If $x \in \big(\frac{r-2}{r-1}, 1\big)$ there exists a unique integer $k \geqslant r$ such that $x \in \big(\frac{k-2}{k-1}, \frac{k-1}{k}\big]$ and a short calculation reveals

$$g_r(x) = \frac{(k)_r}{k^r}\left(1+\sqrt{1-\frac{k}{k-1}x}\right)^{r-1}\left(1-(r-1)\sqrt{1-\frac{k}{k-1}x}\right),$$

where $(k)_r = k(k-1)\cdots(k-r+1)$. Lovász and Simonovits conjectured in the seventies that this function coincides with $i(K_r, x)$ and the third author proved that this is indeed the case.

Theorem 1.10 (Clique density theorem, Reiher [23])**.** *For all integers $r \geqslant 3$ and real $x \in [0,1]$ we have $i(K_r, x) = g_r(x)$.* □

The non-asymptotic problem to determine for given natural numbers n and m the exact minimum number of r-cliques an n-vertex graph with m edges needs to contain is still wide open in general. But for triangles there has recently been spectacular progress by Liu, Pikhurko, and Staden [16].

Easy calculations show that the function $g_r(x)$ is non-differentiable at the critical values $x = 1 - 1/q$, where $q \geqslant r-1$ denotes an integer. Moreover, $g_r(x)$ is piecewise concave between any two consecutive critical values. An old result of Bollobás [2] (proved long before the clique density theorem) asserts that the piece-wise linear function interpolating between the critical values of $g_r(x)$ is a lower bound on $i(K_r, x)$. Here we extend this result to quantum graphs whose constituents are complete multipartite graphs.

To state this generalization we need the following concepts. For every positive integer $r \geqslant 2$ and every quantum graph Q we define the *complete r-partite feasible region* $\Omega_{\mathrm{ind-r}}(Q)$ to be the collection of all points in $[0, (r-1)/r] \times \mathbb{R}$ that can be realized by a Q-good sequence $(G_n)_{n=1}^\infty$ of complete r-partite graphs (isolated vertices are not allowed). For $x \in [0, (r-1)/r]$, let

$$i_r(Q,x) = \inf\{y\colon (x,y) \in \Omega_{\mathrm{ind-r}}(Q)\} \quad \text{and} \quad I_r(Q,x) = \sup\{y\colon (x,y) \in \Omega_{\mathrm{ind-r}}(Q)\}\,.$$

Optimizing over r we put

$$m(Q,x) = \inf\left\{i_r(Q,x)\colon r \geqslant \left\lceil \frac{1}{1-x}\right\rceil\right\} \quad \text{and} \quad M(Q,x) = \sup\left\{I_r(Q,x)\colon r \geqslant \left\lceil \frac{1}{1-x}\right\rceil\right\}$$

for every quantum graph Q and every real $x \in [0,1)$ as well as

$$m(Q,1) = M(Q,1) = \lim_{n\to\infty} \varrho(Q, K_n)\,.$$

Clearly, we have

$$i(Q,x) \leqslant m(Q,x) \leqslant M(Q,x) \leqslant I(Q,x)\,.$$

Next we observe that for every bounded function $f\colon [0,1] \longrightarrow \mathbb{R}$ there exist a pointwise minimum concave function $\mathrm{cap}(f) \geqslant f$ and, similarly, a maximum convex function $\mathrm{cup}(f) \leqslant f$. In fact, $\mathrm{cap}(f)$ is given by

$$\mathrm{cap}(f)(x) = \sup\Big\{\lambda_1 f(x_1) + \dots + \lambda_n f(x_n)\colon n \geqslant 1, (\lambda_1, \dots, \lambda_n) \in \Delta_{n-1}, \text{ and } \sum_{i=1}^{n} \lambda_i x_i = x\Big\}$$

for all $x \in [0,1]$, where

$$\Delta_{n-1} = \big\{(\lambda_1, \dots, \lambda_n) \in [0,1]^n\colon \lambda_1 + \dots + \lambda_n = 1\big\}$$

denotes the $(n-1)$-dimensional standard simplex. Moreover, replacing the supremum by an infimum one obtains a formula for $\mathrm{cup}(f)(x)$.

Theorem 1.11. *Let $Q = \sum_{i=1}^{m} \lambda_i F_i$ be a quantum graph all of whose constituents are complete multipartite graphs.*

(a) If every F_i with $\lambda_i > 0$ is complete, then

$$i(Q,x) \geqslant \mathrm{cup}\,(m(Q,x)) \qquad \textit{for all } x \in [0,1]\,.$$

(b) If every F_i with $\lambda_i < 0$ is complete, then

$$I(Q,x) \leqslant \mathrm{cap}\,(M(Q,x)) \qquad \textit{for all } x \in [0,1]\,.$$

The aforementioned result of Bollobás is the case $Q = K_r$ of Theorem 1.11 *(a)*.

1.4. **Almost complete graphs.** For every integer $t \geqslant 3$ we let K_t^- denote the graph obtained from a clique K_t by deleting one edge. As these graphs are neither complete nor empty, Proposition 1.4 tells us that the feasible regions $\Omega_{\mathrm{ind}}(K_t^-)$ are completely determined by the functions $I(K_t^-, x)$. For $t = 3$ we have the following exact result showing that the graphs $H^\star(n,x)$ minimizing the triangle density also maximize the induced K_3^--density.

Theorem 1.12. *The equality $I(K_3^-, x) = \frac{3}{2}\,(x - g_3(x))$ holds for all $x \in [0,1]$.*

For $t \geqslant 4$ we show a piecewise linear upper bound on $I(K_t^-, x)$ that yields the correct value of the inducibility $\mathrm{ind}(K_t^-)$. In the statement that follows, we set

$$k(t) = \begin{cases} \lceil (t+1)(3t-8)/6 \rceil & \text{if } t \neq 5, 8, 11, 14, 17, 20 \\ (t-2)(3t+1)/6 & \text{if } t = 5, 8, 11, 14, 17, 20. \end{cases}$$

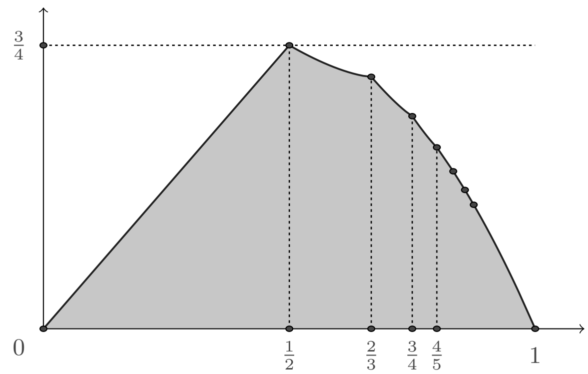

FIGURE 1.2. $\Omega_{\mathrm{ind}}(K_3^-)$.

Theorem 1.13. *For all $t \geqslant 4$ and $x \in [0,1]$ we have $I(K_t^-, x) \leqslant h_t(x)$, where h_t denotes the piecewise linear function interpolating between $h_t(0) = 0$ and*

$$h_t(1 - 1/r) = \binom{t}{2} \frac{(r-1)_{t-2}}{r^{t-1}} \qquad \text{for } r \geqslant k(t)\,.$$

Furthermore,

$$\mathrm{ind}(K_t^-) = \binom{t}{2} \frac{(q(t)-1)_{t-2}}{q(t)^{t-1}}, \qquad \text{where } q(t) = \lceil (t-2)(3t+1)/6 \rceil\,. \tag{1.1}$$

For instance, for $t = 4$ we have $q(4) = 5$ and, hence, $\mathrm{ind}(K_4^-) = 72/125$. This was originally proved by Hirst [11], whose computer assisted argument is based on the flag algebra method. Moreover, Theorem 1.13 yields the upper bound $I(K_4^-, x) \leqslant 3x/4$ for $x \in [0, 3/4]$. For small values of x we have the following stronger bound.

Proposition 1.14. *If $x \in [0, 1/2]$, then $I(K_4^-, x) \leqslant 3x^2/2$.*

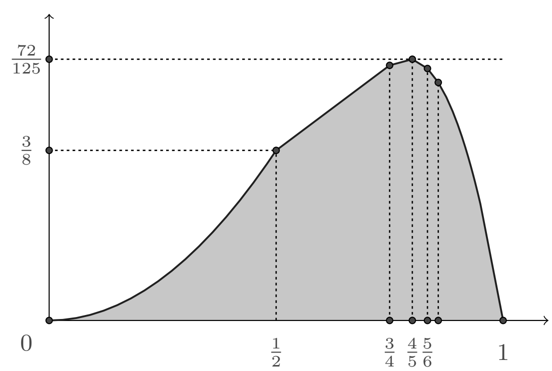

FIGURE 1.3. $\Omega_{\mathrm{ind}}(K_4^-)$ is contained in the shaded area above.

Finally, we remark that our determination of $\mathrm{ind}(K_t^-)$ in (1.1) implies

$$\lim_{t\to\infty} \mathrm{ind}(K_t^-) = 1/e\,. \tag{1.2}$$

This is closely related to the so-called *edge-statistics conjecture* of Alon, Hefetz, Krivelevich, and Tyomkyn [1]. Given positive integers k and $\ell \leqslant \binom{k}{2}$ let the quantum graph $Q_{k,\ell}$ be the sum of all k-vertex graphs with ℓ edges. Alon et al. conjectured $\mathrm{ind}(Q_{k,\ell}) \leqslant 1/e + o_k(1)$ and proved this for some range of ℓ. Following the work of Kwan, Sudakov, and Tran [15], the edges statistics conjecture was resolved by Fox and Sauermann [7] and, independently, by Martinsson, Mousset, Noever, and Trujić [18]. Part of the original motivation for the edges statistics conjecture was the observation that for $\ell = 1$ we have $Q_{k,1} = \overline{K_k^-}$ and $\mathrm{rand}(\overline{K_k^-}, 1/\binom{k}{2}) \to 1/e$ as $k \to \infty$. Thus the asymptotic formula (1.2) follows from the results in [7, 18]. However, the exact values $\mathrm{ind}(K_5^-) = 525/1024$, $\mathrm{ind}(K_6^-) = 178200/13^5$, etc. implied by Theorem 1.13 are new.

1.5. **Stars.** A second case of asymptotic equality in the edge-statistics conjecture occurs for stars. For every positive integer t we denote the star with t edges by S_t. As the case $S_1 = K_2$ is trivial, we may assume $t \geqslant 2$ in the sequel. A quick calculation shows that the induced S_t-density of a complete bipartite graph the sizes of whose vertex classes have roughly the ratio $1 : t$ is $1/e + o_t(1)$.

A precise formula for the inducibility of stars was discovered by Brown and Sidorenko [3] (see Theorem 5.4 below). Here we shall show that for small densities x the values $I(S_t, x)$ of the upper bound function of the feasible region are realized by complete bipartite graphs.

Toward this goal we consider for every real $x \in [0, 1/2]$ a sequence $(B(n,x))_{n=1}^{\infty}$ of complete bipartite graphs with $v(B(n,x)) = n$ for every $n \in \mathbb{N}$ and $\lim_{n\to\infty} \varrho(B(n,x)) = x$. The vertex classes of $B(n,x)$ have the sizes αn and $(1-\alpha)n$ for some $\alpha \in [0, 1/2]$ satisfying $\alpha(1-\alpha) = x/2 + o(1)$. Since $\varrho(S_t, B(n,x)) = (t+1)\big(\alpha(1-\alpha)^t + (1-\alpha)\alpha^t\big) + o_n(1)$ we are lead to the function $s_t \colon [0, 1/2] \longrightarrow \mathbb{R}$ defined by

$$s_t(x) = \lim_{n\to\infty} \varrho(S_t, B(n,x)) = \frac{t+1}{2^t} x \left(\left(1 - \sqrt{1-2x}\right)^{t-1} + \left(1 + \sqrt{1-2x}\right)^{t-1} \right) . \qquad (1.3)$$

As we shall show in Section 5, there is a unique point $x = x^\star(t) \in [0, 1/2]$ at which $s_t(x)$ attains its maximum. Moreover,

$$x^\star(2) = x^\star(3) = \frac{1}{2} \quad \text{and} \quad \frac{2t}{(t+1)^2} < x^\star(t) < \frac{2}{t+1} \text{ holds for } t \geqslant 4 .$$

Using Theorem 1.11 we determine $I(S_t, x)$ for $x \in [0, x^\star(t)]$.

Theorem 1.15. *If $t \geqslant 2$ is an integer and $x \in [0, x^\star(t)]$, then $I(S_t, x) = s_t(x)$.*

Notice that for $t = 2$ this tells us $I(K_3^-, x) = 3x/2$ for $x \in [0, 1/2]$, which follows from Theorem 1.12 as well. It seems hard to determine $I(S_t, x)$ for $t \geqslant 3$ and $x \geqslant x^\star(t)$ (some remarks on this problem are given in Section 7).

For future reference it is convenient to extend the definitions of this subsection to the trivial case $t = 1$ by setting $x^\star(1) = 1/2$ and $s_1(x) = x$ for every $x \in [0, 1/2]$ (which is one half of the values one would obtain by plugging $t = 1$ into (1.3)). It is then still true that we have $I(S_1, x) = s_1(x)$ for every $x \in [0, x^\star(1)]$ and that equality holds for the sequence $(B(n,x))_{n=1}^{\infty}$ of bipartite graphs.

1.6. **Complete bipartite graphs.** For positive integers s and t let $K_{s,t}$ denote the complete bipartite graph whose vertex classes are of size s and t. So $K_{1,t} = S_t$ is a star and it turns out that the calculation of $I(K_{s,t}, x)$ reduces to $I(S_{|s-t|+1}, x)$ for $x \in [0, x^\star(|s-t|+1)]$.

Theorem 1.16. *Let $t \geqslant s \geqslant 2$ be integers. Then for every $x \in [0,1]$ we have*

$$I(K_{s,t}, x) \leqslant \frac{1}{2^{s-1}(t-s+2)}\binom{s+t}{s} x^{s-1} I(S_{t-s+1}, x),$$

and equality holds for $x \leqslant x^\star(t-s+1)$. In particular, for $x \in [0, x^\star(t-s+1)]$,

$$I(K_{s,t}, x) = \begin{cases} \frac{1}{2^t}\binom{2t}{t} x^t & \text{if} \quad t = s, \\ \frac{1}{2^t}\binom{s+t}{s} x^s \left(\left(1-\sqrt{1-2x}\right)^{t-s} + \left(1+\sqrt{1-2x}\right)^{t-s}\right) & \text{if} \quad t > s. \end{cases}$$

The remainder of this subsection focuses on the case $s = t = 2$. Observe that $K_{2,2} = C_4$ is a four-cycle. Theorem 1.16 yields $I(C_4, x) = 3x^2/2$ for every $x \in [0, 1/2]$, where equality is achieved by the sequence $(B(n,x))_{n=1}^{\infty}$ of bipartite graphs. For $x \geqslant 1/2$ we believe that $I(C_4, x)$ is related to the constructions for the clique density theorem (see Construction 1.9).

Conjecture 1.17. *For every real number $x \in [1/2, 1]$ we have*

$$I(C_4, x) = \lim_{n\to\infty} \varrho\left(C_4, H^\star(n,x)\right).$$

This conjecture predicts $I(C_4, 1-1/k) = 3(k-1)/k^3$ for every integer $k \geqslant 2$ and our next result shows that this is indeed the case.

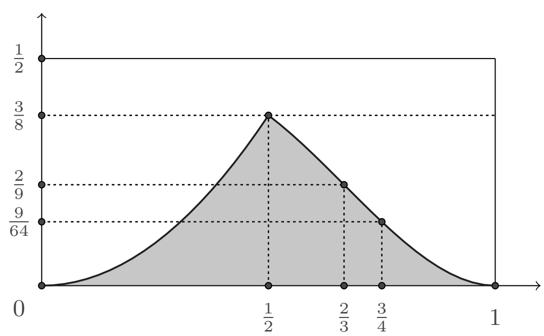

FIGURE 1.4. $\Omega_{\text{ind}}(C_4)$ is contained in the shaded area above.

Theorem 1.18. *If $x \in [1/2, 1]$, then*

$$I(C_4, x) \leqslant 3x(1-x)^2.$$

Moreover, the bound is tight for all $x \in \{(k-1)/k \colon k \in \mathbb{N}$ and $k \geqslant 2\}$.

Organization. For every $x \in \{2,3,4,5,6\}$ the results stated in Subsection 1.x are proved in Section x. Section 7 contains further remarks and open problems.

§2. Proofs of general results

We prove Theorem 1.2 and Proposition 1.4 in this section. The following result can be proved using a similar argument as in [17, Proposition 1.3].

Proposition 2.1. *For every quantum graph Q the set $\Omega_{\mathrm{ind}}(Q)$ is closed.* □

Therefore the definitions of $i(Q,x)$ and $I(Q,x)$ rewrite as

$$i(Q,x) = \min\{y\colon (x,y) \in \Omega_{\mathrm{ind}}(Q)\} \quad \text{and} \quad I(Q,x) = \max\{y\colon (x,y) \in \Omega_{\mathrm{ind}}(Q)\}\,.$$

Next we show that $\Omega_{\mathrm{ind}}(Q)$ is determined by $i(Q,x)$ and $I(Q,x)$.

Proposition 2.2. *Let Q be a quantum graph, $x \in [0,1]$ and $y_1 < y_2$. If $(x,y_1) \in \Omega_{\mathrm{ind}}(Q)$ and $(x,y_2) \in \Omega_{\mathrm{ind}}(Q)$, then $(x,y) \in \Omega_{\mathrm{ind}}(Q)$ holds for all $y \in [y_1,y_2]$.*

Proof of Proposition 2.2. Fix $y \in [y_1,y_2]$. Let $(G'_n)_{n=1}^{\infty}$ be a Q-good sequence of graphs that realizes (x,y_1), and let $(G''_n)_{n=1}^{\infty}$ be a Q-good sequence of graphs that realizes (x,y_2). Without loss of generality we may assume that $V(G'_n) = V(G''_n) = [n]$ for $n \geqslant 1$. We shall construct a sequence of graphs $(G_n)_{n=1}^{\infty}$ with $V(G_n) = [n]$ for every $n \geqslant 1$ that realizes (x,y).

For fixed $n \geqslant 1$ we consider a finite sequence of graphs $G_n^1, \dots, G_n^{m(n)}$ with common vertex set $[n]$ which interpolates between $G_n^1 = G'_n$ and $G^{m(n)} = G''_n$ in the sense that

- for $1 \leqslant m < m(n)$ the graph G_n^{m+1} arises from G_n^m by adding or deleting a single edge,
- and $\min\{\varrho(G'_n), \varrho(G''_n)\} \leqslant \varrho(G_n^m) \leqslant \max\{\varrho(G'_n), \varrho(G''_n)\}$ for every $m \in [m(n)]$.

Due to the first bullet we have $\varrho(Q, G_n^{m+1}) = \varrho(Q, G_n^m) + o(1)$ for every $m \in [m(n)-1]$. Combined with $\varrho(Q, G_n^1) = y_1 + o(1)$ and $\varrho(Q, G_n^{m(n)}) = y_2 + o(1)$ this proves that there exists some $k(n) \in [m(n)]$ such that the graph $G_n = G_n^{k(n)}$ satisfies $\varrho(Q, G_n) = y + o(1)$. Owing to the second bullet we also have $\varrho(G_n) = x + o(1)$. □

Towards the continuity of $I(Q,x)$ we now establish the following lemma.

Lemma 2.3. *For every quantum graph Q there exist constants $\ell \geqslant 1$ and $C \geqslant 0$ such that for all x, x' with $0 < x \leqslant x' \leqslant 1$ we have*

$$\frac{I(Q,x')}{(x')^{\ell}} \leqslant \frac{I(Q,x)}{x^{\ell}} + C \cdot \left(\left(\frac{1}{x}\right)^{\ell} - \left(\frac{1}{x'}\right)^{\ell}\right)\,. \tag{2.1}$$

Proof of Lemma 2.3. Fix $0 < x \leqslant x' \leqslant 1$, set $\alpha = (x'/x)^{1/2} - 1$, and consider a Q-good sequence $(G'_n)_{n=1}^{\infty}$ that realizes $(x', I(Q, x'))$. Without loss of generality we may assume $v(G'_n) = n$ for every $n \geqslant 1$. Let G_n be the graph which is the union of G'_n and a set of $\lfloor \alpha n \rfloor$ isolated vertices. Since

$$\varrho\left(G_n\right) = \frac{\varrho\left(G'_n\right)\binom{n}{2}}{\binom{n+\lfloor \alpha n \rfloor}{2}} \to \frac{x'}{(1+\alpha)^2} = x \quad \text{as} \quad n \to \infty\,,$$

we have

$$I(Q, x) \geqslant \limsup_{n\to\infty} \varrho(Q, G_n)\,. \tag{2.2}$$

To estimate the right side we write $Q = \sum_{i\in P} \lambda_i F_i + \sum_{j\in N} \lambda_j F_j$ with $\lambda_i > 0$ for $i \in P$ and $\lambda_j < 0$ for $j \in N$. Set $\ell_i = v(F_i)$ for every $i \in P \cup N$ and $\ell = \max\{\ell_i/2 \colon i \in P \cup N\}$. For every $i \in P$ the fact that G'_n is a subgraph of G_n yields

$$\varrho\left(F_i, G_n\right) \geqslant \frac{\varrho\left(F_i, G'_n\right)\binom{n}{\ell_i}}{\binom{n+\lfloor \alpha n \rfloor}{\ell_i}} \geqslant \frac{\varrho\left(F_i, G'_n\right)}{(1+\alpha)^{\ell_i}} = \frac{\varrho\left(F_i, G'_n\right)}{(x'/x)^{\ell_i/2}} \geqslant \frac{\varrho\left(F_i, G'_n\right)}{(x'/x)^{\ell}}. \tag{2.3}$$

For $j \in N$ we use that every induced copy of F_j in G_n is either already contained in G'_n or involves one of the new isolated vertices, which implies

$$\varrho\left(F_j, G_n\right) \leqslant \frac{\varrho\left(F_j, G'_n\right)\binom{n}{\ell_j} + \alpha n \cdot \binom{n+\lfloor \alpha n \rfloor}{\ell_j - 1}}{\binom{n+\lfloor \alpha n \rfloor}{\ell_j}} \leqslant \varrho(F_j, G'_n) + \frac{\ell_j \cdot \alpha}{1+\alpha} + o_n(1)\,.$$

Taking into account that

$$\varrho(F_j, G'_n) \leqslant \frac{\varrho\left(F_j, G'_n\right)}{(x'/x)^{\ell}} + \left(1 - \left(\frac{x}{x'}\right)^{\ell}\right)$$

and

$$\frac{\alpha}{1+\alpha} = 1 - \left(\frac{x}{x'}\right)^{1/2} \leqslant 1 - \left(\frac{x}{x'}\right)^{\ell}$$

we obtain

$$\varrho\left(F_j, G_n\right) \leqslant \frac{\varrho(F_j, G'_n)}{(x'/x)^{\ell}} + (\ell_j + 1)\left(1 - \left(\frac{x}{x'}\right)^{\ell}\right) + o_n(1)\,.$$

Combined with (2.3) this entails

$$\begin{aligned}
\varrho\left(Q, G_n\right) &= \sum_{i\in P} \lambda_i \varrho\left(F_i, G_n\right) + \sum_{j\in N} \lambda_j \varrho\left(F_j, G_n\right) \\
&\geqslant \sum_{i\in P} \lambda_i \frac{\varrho\left(F_i, G'_n\right)}{(x'/x)^{\ell}} + \sum_{j\in N} \lambda_j \left(\frac{\varrho\left(F_j, G'_n\right)}{(x'/x)^{\ell}} + (\ell_j + 1)\left(1 - \left(\frac{x}{x'}\right)^{\ell}\right)\right) - o_n(1) \\
&\geqslant \frac{\varrho(Q, G'_n)}{(x'/x)^{\ell/2}} - C \cdot \left(1 - \left(\frac{x}{x'}\right)^{\ell}\right) - o_n(1)\,,
\end{aligned}$$

where $C = \sum_{j\in N}(-\lambda_j)(\ell_j + 1) \geqslant 0$. Now (2.2) reveals

$$I(Q, x) \geqslant \frac{I(Q, x')}{(x'/x)^{\ell}} - C \cdot \left(1 - \left(\frac{x}{x'}\right)^{\ell}\right)$$

and upon multiplying both sides by $x^{-\ell}$ the claim follows. $\square$

For later use we record the following consequence.

Corollary 2.4. *Given a quantum graph Q and $x' \in [0,1]$, $\varepsilon > 0$, there exists some $\delta > 0$ such that $I(Q,x) > I(Q,x') - \varepsilon$ holds for all $x \in [0,x')$ with $|x - x'| < \delta$.* $\square$

Now we are ready to prove the main result of Subsection 1.2.

Proof of Theorem 1.2. Given a quantum graph Q the formula

$$\Omega_{\mathrm{ind}}(Q) = \big\{(x,y) \in [0,1] \times \mathbb{R} \colon i(Q,F) \leqslant y \leqslant I(Q,F)\big\}$$

follows immediately from Proposition 2.2. Now, due to Fact 1.3 *(a)* it suffices to show that $I(Q,x)$ is continuous and almost everywhere differentiable.

Let $\ell \geqslant 1$, $C \geqslant 0$ be the constants provided by Lemma 2.3. Owing to (2.1) the function $F\colon (0,1] \longrightarrow \mathbb{R}$ defined by $F(x) = (I(Q,x) + C)/x^{\ell}$ is decreasing. It follows that F is almost everywhere differentiable and that for every $x \in (0,1]$ the left-sided limit $\lim_{x \to x_0^-} F(x)$ exists. Consequently, the function $I(Q,x)$ has the same properties.

Let us show next that $I(Q,x)$ is left-continuous. Given an arbitrary $x_0 \in (0,1]$ we already know that the limit $y_0 = \lim_{x\to x_0^-} I(Q,x)$ exists. Proposition 2.1 yields $(x_0,y_0) \in \Omega_{\mathrm{ind}}(Q)$, whence $I(Q,x_0) \geqslant y_0$. But $I(Q,x_0) > y_0$ would contradict Corollary 2.4 and thus we have indeed $I(Q,x_0) = y_0$. By Fact 1.3 *(b)* the function $I(Q,x) = I(\overline{Q}, 1-x)$ is right-continuous as well. This concludes the proof. $\square$

Proof of Proposition 1.4. For every $n \in \mathbb{N}$ and $x \in [0,1]$ we let $H'(n,x)$ denote the n-vertex graph consisting of a clique of order $\lfloor x^{1/2}n \rfloor$ and $n - \lfloor x^{1/2}n \rfloor$ isolated vertices. Moreover, we set $H''(n,x) = \overline{H'(n,1-x)}$. Notice that $\lim_{n\to\infty} \varrho(H'(n,x)) = \lim_{n\to\infty} \varrho(H''(n,x)) = x$ holds for every $x \in [0,1]$.

Now suppose that F is a graph which is neither complete nor empty. If F has no isolated vertex, then $\varrho(F, H'(n,x)) = 0$ holds for all $n \in \mathbb{N}$ and $x \in [0,1]$, which leads to $i(F,x) = 0$. If F has an isolated vertex we get the same conclusion from $\varrho(F, H''(n,x)) = 0$. $\square$

§3. Proof for complete multipartite graphs

We prove Theorem 1.11 in this section. The following result of Schelp and Thomason [25] will be useful in our argument.

Theorem 3.1 (Schelp-Thomason [25]). *Let $Q = \sum_{i\in[m]} \lambda_i F_i$ be a quantum graph whose constituents are complete multipartite graphs and let $n \in \mathbb{N}$. If every F_i with $\lambda_i < 0$ is complete, then among all n-vertex graphs G maximizing $\varrho(Q,G)$ there is a complete multipartite one.*

Definition 3.2. *Suppose that $H\colon [0,1] \to \mathbb{R}$ is a concave function and $L\colon [0,1] \to \mathbb{R}$ is a linear function. We say L is a tangent line of H at $x_0 \in [0,1]$ if $L(x) \geqslant H(x)$ holds for $x \in [0,1]$ with equality for $x = x_0$.*

It is easy to see that for every concave function $F\colon [0,1] \to \mathbb{R}$ and every $x_0 \in (0,1)$ there always exists a (not necessarily unique) tangent line of F at x_0.

Proof of Theorem 1.11. By Fact 1.3 *(a)* it suffices to show part *(b)*. Let $Q = \sum_{i\in[m]} \lambda_i F_i$ be a quantum graph whose constituents are complete multipartite graphs such that every F_i with and $\lambda_i < 0$ is complete. For brevity we set $H(x) = \mathrm{cap}\,(M(Q,x))$ for every $x \in [0,1]$. Clearly

$$H(0) = M(Q,0) = \lim_{n\to\infty} \varrho(Q, \overline{K_n}) = I(Q,0)$$

and a similar argument shows $H(1) = I(Q,1)$. So it remains to prove $H(x_0) \geqslant I(Q,x_0)$ for every $x_0 \in (0,1)$. To this end we choose a tangent line $L(x) = kx + p$ of H at x_0, so that

$$H(x) \leqslant kx + p \quad \text{for all } x \in [0,1] \quad \text{and} \quad H(x_0) = kx_0 + p\,. \tag{3.1}$$

Now let $(G_n)_{n=1}^{\infty}$ be a sequence of graphs that realizes $(x_0, I(Q,x_0))$. By Theorem 3.1 applied to the quantum graph $Q^\star = Q - kK_2$ there exists for every $n \geqslant 1$ a multipartite n-vertex graph G'_n such that $v(G'_n) = v(G_n)$ and

$$\varrho(Q,G_n) - k\varrho(G_n) = \varrho(Q^\star, G_n) \leqslant \varrho(Q^\star, G'_n) = \varrho(Q, G'_n) - k\varrho(G'_n)\,. \tag{3.2}$$

By passing to a subsequence of $(G'_n)_{n=1}^{\infty}$ we may assume that the limits $x_1 = \lim_{n\to\infty} \varrho(G'_n)$ and $y_1 = \lim_{n\to\infty} \varrho(Q, G'_n)$ exist. Due to the definition of $M(Q,x_1)$ and (3.1) we have

$$y_1 \leqslant M(Q,x_1) \leqslant H(x_1) \leqslant kx_1 + p$$

and taking the limit $n \to \infty$ in (3.2) it follows that

$$I(Q,x_0) - kx_0 \leqslant y_1 - kx_1 \leqslant p\,.$$

Together with (3.1) this leads to the desired estimate $I(Q,x_0) \leqslant kx_0 + p = H(x_0)$. □

§4. Proofs for almost complete graphs

In this section we prove Theorems 1.12 and 1.13 as well as Proposition 1.14.

4.1. **Cherries.** We begin with the proof of Theorem 1.12. Consider a graph $G = (V,E)$ with $|V| = n$ vertices. Counting the number of pairs $(\{x,y\}, z) \in E \times V$ with $z \neq x, y$ in two different ways, we obtain

$$(n-2)|E| = N(\overline{K_3^-}, G) + 2N(K_3^-, G) + 3N(K_3, G)\,.$$

Dividing by $2\binom{n}{3}$ and rearranging we deduce

$$\varrho(K_3^-, G) = \frac{3}{2}\big(\varrho(K_2, G) - \varrho(K_3, G)\big) - \frac{1}{2}\varrho(\overline{K_3^-}, G)\,.$$

Therefore the clique density theorem yields for every $x \in [0,1]$ the upper bound $I(K_3^-, x) \leqslant \frac{3}{2}(x - g_3(x))$. Moreover, for every $x \in [0,1]$ the sequence of multipartite graphs $(H^\star(n,x))_{n=1}^\infty$ is $\overline{K_3^-}$-free and establishes the lower bound $I(K_3^-, x) \geqslant \frac{3}{2}(x - g_3(x))$.

4.2. **Piecewise linear upper bounds.** Roughly speaking we show in this subsection that a concave piecewise linear function is an upper bound on $I(K_t^-, x)$ if it respects the constraints coming from Turán graphs.

Lemma 4.1. *Suppose that an integer $s \geqslant 1$ and real numbers λ, μ have the property that*

$$\frac{1}{r^{s+1}}\binom{r-1}{s} \leqslant \lambda\frac{r-1}{2r} + \mu \tag{4.1}$$

holds for every positive integer r. If $m \geqslant 1$ and $(\alpha_1, \ldots, \alpha_m) \in \Delta_{m-1}$, then

$$\sum_{i=1}^{m} \sum_{W \in \binom{[m]\setminus\{i\}}{s}} \alpha_i^2 \prod_{j\in W} \alpha_j \leqslant \lambda \sum_{\{i,j\}\in\binom{[m]}{2}} \alpha_i\alpha_j + \mu\,.$$

Proof. Assume for the sake of contradiction that this fails and let m denote the least positive integer for which there exists a counterexample. Appealing to a theorem of Weierstraß, we pick a point $(\alpha_1^\star, \ldots, \alpha_m^\star) \in \Delta_{m-1}$ such that the difference

$$\Phi = \sum_{i=1}^{m} \sum_{W \in \binom{[m]\setminus\{i\}}{s}} (\alpha_i^\star)^2 \prod_{j\in W} \alpha_j^\star - \lambda \sum_{\{i,j\}\in\binom{[m]}{2}} \alpha_i^\star\alpha_j^\star$$

is maximal. Due to our indirect assumption we know $\Phi > \mu$. The case $r = m$ of (4.1) reveals that $\alpha_1^\star = \cdots = \alpha_m^\star = 1/m$ is false. Therefore, we have $m \geqslant 2$ and and for reasons of symmetry we may assume that $\alpha_1^\star < \alpha_2^\star$.

Given two real numbers $\alpha_1, \alpha_2 \geqslant 0$ satisfying

$$\alpha_1 + \alpha_2 = \alpha_1^\star + \alpha_2^\star$$

we write $\Phi(\alpha_1, \alpha_2)$ for the result of replacing $\alpha_1^\star$, $\alpha_2^\star$ in the above formula for Φ by α_1, α_2. So $\Phi(\alpha_1^\star, \alpha_2^\star) = \Phi$ and there are constants $c_1, \ldots, c_5$ depending only on $\alpha_3^\star, \ldots, \alpha_m^\star$, and λ such that

$$\Phi(\alpha_1, \alpha_2) = c_1 + c_2(\alpha_1 + \alpha_2) + c_3(\alpha_1^2 + \alpha_2^2) + c_4\alpha_1\alpha_2 + c_5(\alpha_1 + \alpha_2)\alpha_1\alpha_2\,.$$

Since $\alpha_1 + \alpha_2$ is constant and $\alpha_1^2 + \alpha_2^2$, $2\alpha_1\alpha_2$ add up to the constant $(\alpha_1^\star + \alpha_2^\star)^2$, it follows that there are constants c_6, c_7 such that

$$\Phi(\alpha_1, \alpha_2) = c_6\alpha_1\alpha_2 + c_7\,.$$

If $c_6 \neq 0$ we can find a real number $\xi \neq 0$ such that $|\xi|$ is very small and $\Phi(\alpha_1^\star+\xi, \alpha_2^\star-\xi) > \Phi$ contradicts the maximality of Φ. So $c_6 = 0$ and $\Phi(\alpha_1, \alpha_2) = c_7 = \Phi$ is constant. But now $\Phi(\alpha_1^\star + \alpha_2^\star, 0) = \Phi$ contradicts the minimality of m. This completes the proof. □

Lemma 4.2. *Suppose that $t \geqslant 3$ and that $f\colon [0,1] \longrightarrow \mathbb{R}$ is a piecewise linear concave function. If for every positive integer r we have*

$$f(1-1/r) \geqslant \binom{t}{2}\frac{(r-1)\cdots(r-(t-2))}{r^{t-1}}\,, \tag{4.2}$$

then $I(K_t^-, x) \leqslant f(x)$ holds for every $x \in [0,1]$.

Proof. Since f is the pointwise minimum of a family of linear functions, it suffices to deal with the case that $f(x) = \lambda x + \mu$ is itself linear. By Theorem 1.11 *(b)* it is enough to show $M(K_t^-, x) \leqslant \lambda x + \mu$ for every $x \in [0,1]$. We shall establish the more precise estimate that every complete multipartite graph G on n vertices satisfies

$$N(K_t^-, G) \leqslant (2\lambda|E(G)| + \mu n^2)n^{t-2}/t!\,. \tag{4.3}$$

Let $a_1, \dots, a_m$ be the sizes of the vertex classes of G and set $\alpha_i = a_i/n$ for every $i \in [m]$. Now $\sum_{i=1}^m \alpha_i = 1$ and

$$N(K_t^-, G) = \sum_{i=1}^m \binom{a_i}{2} \sum_{W\in\binom{[m]\smallsetminus\{i\}}{t-2}} \prod_{j\in W} a_j \leqslant \frac{n^t}{2}\sum_{i=1}^m \alpha_i^2 \sum_{W\in\binom{[m]\smallsetminus\{i\}}{t-2}} \prod_{j\in W} \alpha_j$$

and, therefore, instead of (4.3) it suffices to show

$$\sum_{i=1}^m \alpha_i^2 \sum_{W\in\binom{[m]\smallsetminus\{i\}}{t-2}} \prod_{j\in W} \alpha_j \leqslant \frac{4\lambda}{t!} \sum_{\{i,j\}\in\binom{[m]}{2}} \alpha_i\alpha_j + \frac{2\mu}{t!}\,.$$

By Lemma 4.1 applied to $t-2$, $4\lambda/t!$, $2\mu/t!$ here in place of s, λ, μ there this inequality follows from the fact that

$$\frac{1}{r^{t-1}}\binom{r-1}{t-2} \leqslant \frac{4\lambda}{t!}\cdot\frac{r-1}{2r} + \frac{2\mu}{t!} = \frac{2f(1-1/r)}{t!}$$

holds for every $r \geqslant 1$, which is in turn equivalent to the hypothesis (4.2). □

4.3. **Precise calculations.** Fix an integer $t \geqslant 4$. Our next goal is to show that the function h_t introduced in Theorem 1.13 satisfies the assumptions of Lemma 4.2. To this end we set $A_r = \binom{t}{2}\frac{(r-2)_{t-3}}{r^{t-2}}$ for every integer $r \geqslant 2$.

Lemma 4.3. *Let $t \geqslant 4$ and $r \geqslant t-1$ be integers.*

(*a*) *If $r \leqslant (3t^2-5t-4)/6$, then $A_{r-1} < A_r$.*
(*b*) *If $r = (3t^2-5t-2)/6$, then $A_{r-1} < A_r$ or $A_{r-1} > A_r$ holds depending on whether $t \leqslant 20$ or $t > 20$.*
(*c*) *If $r \geqslant (3t^2-5t)/6$, then $A_{r-1} > A_r$.*

In particular, there exists a unique integer $k \geqslant t-2$ *satisfying* $A_k = \max\{A_r \colon r \geqslant t-2\}$*, namely* $k = k(t)$*.*

Proof. One confirms easily that

$$A_{r-1} < A_r \iff 1 - \frac{t-1}{r} < \left(1 - \frac{1}{r}\right)^{t-2}\left(1 - \frac{2}{r}\right).$$

Due to the approximations

$$\sum_{i=0}^{3} \frac{(-1)^i}{r^i}\binom{t-2}{i} \leqslant \left(1 - \frac{1}{r}\right)^{t-2} \leqslant \sum_{i=0}^{4} \frac{(-1)^i}{r^i}\binom{t-2}{i}$$

we obtain the implications

$$\frac{(t+2)(t-2)(t-3)}{6} < r\left(\frac{(t+1)(t-2)}{2} - r\right) \implies A_{r-1} < A_r$$

and

$$\frac{(t+2)(t-2)(t-3)}{6} > r\left(\frac{(t+1)(t-2)}{2} - r\right) + \frac{t+3}{4r}\binom{t-2}{3} \implies A_{r-1} > A_r$$

(see also the proof of Lemma 4.5).

So for the proof of part *(a)* it suffices to observe that $(t-1)/3 \leqslant r \leqslant (3t^2 - 5t - 4)/6$ implies

$$r\left(\frac{(t+1)(t-2)}{2} - r\right) \geqslant \frac{3t^2 - 5t - 4}{6} \cdot \frac{t-1}{3} \geqslant \frac{t(t-1)(t-2)}{6} > \frac{(t+2)(t-2)(t-3)}{6}.$$

Similarly, if $r \geqslant (3t^2 - 5t)/6 > (t-2)(t+3)/6$ we have

$$\begin{aligned} r\left(\frac{(t+1)(t-2)}{2} - r\right) + \frac{t+3}{4r}\binom{t-2}{3} &< \frac{3t^2 - 5t}{6} \cdot \frac{t-3}{3} + \frac{(t-3)(t-4)}{4} \\ &< \frac{(t+2)(t-2)(t-3)}{6}, \end{aligned}$$

which proves part *(c)*.

We proceed with the case $r = (3t^2 - 5t - 2)/6$, which requires $t \equiv 2 \pmod 3$. Direct calculations show $A_{r-1} < A_r$ for $t \in \{5, 8, 11, 14, 17, 20\}$ and $A_{r-1} > A_r$ for $t = 23$. As soon as $t \geqslant 26$ we have $8(t-8)r > 3(t+3)(t-3)(t-4)$ and hence

$$\begin{aligned} r\left(\frac{(t+1)(t-2)}{2} - r\right) + \frac{t+3}{4r}\binom{t-2}{3} &< \frac{3t^2 - 5t - 2}{6} \cdot \frac{t-2}{3} + \frac{(t-2)(t-8)}{9} \\ &= \frac{(t+2)(t-2)(t-3)}{6}, \end{aligned}$$

which concludes the discussion of *(b)*. Finally, *(a)*–*(c)* together imply

$$A_{t-2} < A_{t-1} < \cdots < A_{k(t)} \quad \text{and} \quad A_{k(t)} > A_{k(t)+1} > \ldots,$$

whence $A_{k(t)} = \max\{A_r \colon r \geqslant t-2\}$. □

Lemma 4.4. *We have $I(K_t^-, x) \leqslant h_t(x)$ for every $x \in [0,1]$.*

Proof. For later use we observe that the number $k = k(t)$ satisfies

$$k \geqslant \frac{t(t-2)}{2}\,. \tag{4.4}$$

Indeed, if $t \neq 5, 8, 11, 14, 17, 20$, then $k - t(t-2)/2 = \lceil (t-8)/6 \rceil \geqslant \lceil -2/3 \rceil = 0$ and in the remaining cases we have $k - t(t-2)/2 = (t-2)/6 \geqslant 0$.

Next we show

$$h_t(1-1/r) \geqslant \binom{t}{2}\frac{(r-1)\cdots(r-(t-2))}{r^{t-1}}$$

for every positive integer r. The cases $r \leqslant t-2$ and $r \geqslant k$ are clear. Now suppose that $t-1 \leqslant r < k$. Since $1-1/r \leqslant 1-1/k$ and $h_t(x) = A_k \cdot x$ for all $x \in [0, 1-1/k]$ we have

$$h_t(1-1/r) = A_k \cdot (1-1/r) \geqslant A_r \cdot (1-1/r) = \binom{t}{2}\frac{(r-1)\cdots(r-(t-2))}{r^{t-1}}\,,$$

as desired.

According to Lemma 4.2 it only remains to show that h_t is concave. Now $A_k > A_{k+1}$ rewrites as

$$\frac{h_t(1-1/k)}{1-1/k} > \frac{h_t(1-1/(k+1))}{1-1/(k+1)}$$

and, therefore, h_t is concave in some sufficiently small neighbourhood around $x = 1-1/k$. Define $F\colon \left[0, \frac{1}{k}\right] \longrightarrow \mathbb{R}$ by $F(x) = x\prod_{i=1}^{t-2}(1-ix)$. Since $h_t(1-1/r) = \binom{t}{2}F(1/r)$ holds for every $r \geqslant k$, it suffices to show that F is concave. If $x \in [0, 1/k]$, then

$$\sum_{i=1}^{t-2}\frac{i}{1-ix} \leqslant \frac{1+\cdots+(t-2)}{1-(t-2)x} \overset{(4.4)}{\leqslant} \frac{(t-2)(t-1)}{2(1-2/t)} = \frac{(t-1)t}{2} < \frac{2}{x}$$

and thus

$$\frac{F''(x)}{F(x)} = \sum_{1\leqslant i<j\leqslant t-2}\frac{ij}{(1-ix)(1-jx)} - \frac{1}{x}\sum_{i=1}^{t-2}\frac{i}{1-ix} < \frac{1}{2}\Big(\sum_{i=1}^{t-2}\frac{i}{1-ix}\Big)^2 - \frac{1}{x}\sum_{i=1}^{t-2}\frac{i}{1-ix} \leqslant 0\,,$$

which proves that F is indeed concave. □

The only part of Theorem 1.13 still lacking verification is (1.1). Setting $B_r = \binom{t}{2}\frac{(r-1)_{t-2}}{r^{t-1}}$ for every $r \geqslant t-2$ and $f = \lceil (t-2)(3t+1)/6 \rceil$ we are to show $B_f = \max\{B_r \colon r \geqslant t-2\}$. It turns out that this holds in the following slightly stronger form.

Lemma 4.5. *We have $0 = B_{t-2} < B_{t-1} < \cdots < B_f$ and $B_f > B_{f+1} > \ldots$.*

Proof. First we show $B_{r-1} < B_r$ for every integer $r \in [t-1, f]$. The fact that $(t-2)(3t+1)$ is even yields $f \leqslant (t-2)(3t+1)/6 + 2/3 = (t-1)(3t-2)/6$, whence

$$\frac{t-1}{3} \leqslant r \leqslant \binom{t}{2} - \frac{t-1}{3}\,.$$

For this reason we have

$$r\left(\binom{t}{2}-r\right) \geqslant \frac{t-1}{3}\left(\binom{t}{2}-\frac{t-1}{3}\right) > \binom{t}{3},$$

which rewrites as

$$1-\frac{t-1}{r} < 1-\frac{t}{r}+\binom{t}{2}\frac{1}{r^2}-\binom{t}{3}\frac{1}{r^3}.$$

As the right side is at most $(1-1/r)^t$, this proves

$$1 < \frac{(r-1)^t}{r^{t-1}(r-(t-1))} = \frac{B_r}{B_{r-1}},$$

as desired.

Next we show $B_{r-1} > B_r$ for every $r \geqslant f+1$. Due to $r \geqslant (3t^2-5t+4)/6 > \frac{1}{2}\binom{t}{2}$ we have

$$r\left(\binom{t}{2}-r\right) < \frac{3t^2-5t+4}{6}\cdot\frac{t-2}{3} = \binom{t}{3}-\frac{(t-2)^2}{9}.$$

Moreover, $r \geqslant t(t-2)/2$ implies

$$\binom{t}{4}\cdot\frac{1}{r} < \frac{(t-1)(t-3)}{12} < \frac{(t-2)^2}{9}.$$

Adding the previous two estimates we obtain

$$r\left(\binom{t}{2}-r\right)+\binom{t}{4}\cdot\frac{1}{r} < \binom{t}{3},$$

which rewrites as

$$1-\frac{t-1}{r} > 1-\frac{t}{r}+\binom{t}{2}\frac{1}{r^2}-\binom{t}{3}\frac{1}{r^3}+\binom{t}{4}\frac{1}{r^4}.$$

As the right side is an upper bound on $(1-1/r)^t$ we can conclude

$$1 > \frac{(r-1)^t}{r^{t-1}(r-(t-1))} = \frac{B_r}{B_{r-1}}. \qquad \square$$

4.4. **More on K_4^-.** Our last result on $\Omega_{\mathrm{ind}}(K_4^-)$, Proposition 1.14, is an immediate consequence of the following result.

Lemma 4.6. *Every graph G satisfies $N(K_4^-, G) \leqslant \frac{1}{2}\binom{|E(G)|}{2}$.*

Proof. Notice that an abstract K_4^- has two perfect matchings. Now with every induced copy of K_4^- in G we associate its two perfect matchings, viewed as members of $\binom{E(G)}{2}$. We are thereby considering $2N(K_4^-, G)$ pairs of edges of G. Since every pair $\{e,f\} \in \binom{E(G)}{2}$ can be associated to at most one copy of K_4^- in G (namely the copy induced by $e \cup f$, if it exists), this proves the claim. $\square$

§5. Proofs for stars

In this section we prove Theorem 1.15. Recall from Section 1.5 that for every integer $t \geqslant 3$ and every real $x \in [0, 1/2]$ we defined

$$s_t(x) = \frac{t+1}{2^t} x \left(\left(1 - \sqrt{1-2x}\right)^{t-1} + \left(1 + \sqrt{1-2x}\right)^{t-1} \right).$$

We commence by showing that there is a unique $x^\star(t) \in [0, 1/2]$, where the function s_t attains its maximum. For $t = 3$ we have $s_3(x) = 2x(1-x)$ and, hence, $x^\star(3) = 1/2$ is as desired. The case $t \geqslant 4$ is addressed by the next lemma.

Lemma 5.1. *For $t \geqslant 4$ there exists a unique real $x^\star(t) \in \left(\frac{2t}{(t+1)^2}, \frac{2}{t+1}\right)$ such that the function s_t is strictly increasing on $[0, x^\star(t)]$ and strictly decreasing on $[x^\star(t), 1/2]$.*

Proof. Define the auxiliary function $h\colon [0,1] \longrightarrow \mathbb{R}$ by $h(y) = 1 - ty + ty^{t-1} - y^t$. Due to $h''(y) = t(t-1)y^{t-3}(t-2-y) > 0$ for $y \in (0,1]$ this function is strictly convex. Together with $h(0) = 1$, $h(1) = 0$, and $h'(1) = t(t-3) > 0$ this shows that there exists a unique $y^\star = (0,1)$ such that $h(y^\star) = 0$, $h(y) > 0$ for $y \in [0, y^\star)$, and $h(y) < 0$ for $y \in (y^\star, 1)$.

Due to

$$\frac{\mathrm{d}}{\mathrm{d}y} \frac{y + y^t}{(1+y)^{t+1}} = \frac{h(y)}{(1+y)^{t+2}}$$

it follows that $\frac{y+y^t}{(1+y)^{t+1}}$ is strictly increasing on $[0, y^\star)$ and strictly decreasing on $(y^\star, 1]$. As $\frac{2y}{(1+y)^2}$ is strictly increasing on $[0,1]$ and

$$s_t\left(\frac{2y}{(1+y)^2}\right) = \frac{(t+1)(y+y^t)}{(1+y)^{t+1}},$$

it follows that s_t has the desired monotonicity properties for $x^\star(t) = \frac{2y^\star}{(1+y^\star)^2}$.

Next, due to $h(1/t) = t^{2-t} - t^{-t} > 0$ we have $y^\star > \frac{1}{t}$ and, consequently, $x^\star(t) > \frac{2t}{(t+1)^2}$. Similarly,

$$h\left(\frac{1}{t-1}\right) < -\frac{1}{t-1} + \frac{t}{(t-1)^{t-1}} \leqslant \frac{t-(t-1)^2}{(t-1)^3} < 0$$

yields $y^\star < \frac{1}{t-1}$, whence

$$x^\star(t) < \frac{2(t-1)}{t^2} < \frac{2}{t+1}. \qquad \square$$

Lemma 5.2. *For every integer $t \geqslant 3$ the function s_t is increasing and concave on $[0, x^\star(t)]$.*

Proof. Our choice of $x^\star(t)$ guarantees that s_t is indeed increasing. So it suffices to show that s_t is concave on the interval $I_t = \left[0, \frac{2}{t+1}\right]$. Since

$$s_t(x) = \frac{t+1}{2^{t-1}} \sum_{0 \leqslant n \leqslant (t-1)/2} \binom{t-1}{2n} x(1-2x)^n$$

it suffices to show for every positive integer $n \leqslant (t-1)/2$ that $x(1-2x)^n$ is concave on I_t. This follows immediately from

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} x(1-2x)^n = 4n(1-2x)^{n-2}[(n+1)x-1]\,. \qquad \square$$

Our next step is to show $M(S_t, x) = I_2(S_t, x) = s_t(x)$ for $x \in [0, x^\star(t)]$. To this end we use the following result due to Brown and Sidorenko, which is implicit in the proof of [3, Proposition 2].

Proposition 5.3 (Brown-Sidorenko [3])**.** *Let r, s, t, n be positive integers with $r \geqslant 3$. For every complete r-partite graph G on n vertices there exists a complete $(r-1)$-partite graph G' on the same vertex set such that $e(G') \leqslant e(G)$ and $N(K_{s,t}, G') \geqslant N(K_{s,t}, G)$.*

The proof proceeds by "merging" two smallest vertex classes of G, i.e., if $V_1, \dots, V_r$ with $|V_1| \leqslant \dots \leqslant |V_r|$ are the vertex classes of G, then one constructs G' so as to have the vertex classes $V_1 \cup V_2, V_3, \dots, V_r$. Clearly, $r-2$ iterations of this process lead to a complete bipartite graph G'' such that $V(G'') = V(G)$, $e(G'') \leqslant e(G)$, and $N(K_{s,t}, G') \geqslant N(K_{s,t}, G)$. This shows that for the determination of the inducibility of $K_{s,t}$ only complete bipartite host graphs are relevant. This establishes the following result on stars.

Theorem 5.4 (Brown-Sidorenko [3])**.** *For every integer $t \geqslant 2$ the inducibility of S_t is given by* $\mathrm{ind}(S_t) = I_2(S_t, x^\star(t))$.

We proceed with another simple consequence of Proposition 5.3.

Lemma 5.5. *If $r, t \geqslant 2$ are integers and $x \in [0, x^\star(t)]$, then $I_2(S_t, x) \geqslant I_r(S_t, x)$.*

Proof of Lemma 5.5. Let $y_2 = I_2(S_t, x)$, $y_r = I_r(S_t, x)$ and consider an S_t-good sequence of complete r-partite graphs $(G_n)_{n=1}^\infty$ that realizes (x, y_r). In view of Proposition 5.3 there exists a sequence $(G'_n)_{n=1}^\infty$ of complete bipartite graphs such that

$$V(G'_n) = V(G_n), \quad e(G'_n) \leqslant e(G_n), \quad \text{and} \quad N(K_{s,t}, G'_n) \geqslant N(K_{s,t}, G_n) \tag{5.1}$$

hold for every positive integer n. By passing to a subsequence we may assume that the limits $x' = \lim_{n\to\infty} \varrho(G'_n)$ and $y'_2 = \lim_{n\to\infty} \varrho(S_t, G'_n)$ exist. Now (5.1) implies

$$x' \leqslant x \quad \text{and} \quad y'_2 \geqslant y_r\,, \tag{5.2}$$

and as $(G'_n)_{n=1}^\infty$ is an S_t-good sequence of complete bipartite graphs that realizes (x', y'_2) we have $y'_2 \leqslant I_2(S_t, x')$. Since $I_2(S_t, \cdot) = s_t(\cdot)$ is increasing on $[0, x^\star(t)]$, the first estimate in (5.2) entails $I_2(S_t, x') \leqslant I_2(S_t, x)$. So altogether we obtain

$$y_r \leqslant y'_2 \leqslant I_2(S_t, x') \leqslant I_2(S_t, x)\,,$$

which concludes the proof. $\square$

Now we are ready to prove Theorem 1.15.

Proof of Theorem 1.15. The case $t = 2$ already being understood in Theorem 1.12 we may assume that $t \geqslant 3$. It is clear that $I(S_t, x) \geqslant I_2(S_t, x) = s_t(x)$ holds for $x \in [0, 1/2]$ and thus we just need to show $I(S_t, x) \leqslant s_t(x)$ for $x \in [0, x^\star(t)]$. Define $f\colon [0,1] \longrightarrow \mathbb{R}$ by

$$f(x) = \begin{cases} s_t(x) & \text{for} \quad x \in [0, x^\star(t)] \\ s_t\left(x^\star(t)\right) & \text{for} \quad x \in [x^\star(t), 1]. \end{cases}$$

Lemma 5.2 informs us that f is concave. Moreover, we have $f(x) \geqslant M(S_t, x)$ for all $x \in [0,1]$. Indeed, if $x \in [0, x^\star(t)]$ this follows from Lemma 5.5 and for $x \in [x^\star(t), 1]$ we can appeal to Theorem 5.4 instead. Summarizing, $f(x)$ is a concave upper bound on $M(S_t, x)$. Owing to Theorem 1.11 this proves $I(S_t, x) \leqslant f(x) = s_t(x)$ for every $x \in [0, x^\star(t)]$. □

§6. Proofs for complete bipartite graphs

In this section we prove Theorems 1.16 and 1.18. The upper bound on $I(K_{s,t}, x)$ stated in Theorem 1.16 is an immediate consequence of the following result.

Proposition 6.1. *If $t \geqslant s \geqslant 2$ are positive integers, then for every graph G we have*

$$N(K_{s,t}, G) \leqslant \frac{(t-s+1)!}{s!t!} N(S_{t-s+1}, G) \cdot \left(e(G)\right)^{s-1} .$$

Proof of Proposition 6.1. Notice that for an abstract $K_{s,t}$ the number of ordered partitions $V(K_{s,t}) = U_1 \cup \ldots \cup U_s$ such that U_1 induces a star S_{t-s+1} and each of $U_2, \ldots, U_s$ induces an edge is $\binom{t}{t-s+1}(s-1)!s!$. This is because there are $s\binom{t}{t-s+1}$ possibilities for U_1; moreover, if $i \in [2, s]$ and $U_1, \ldots, U_{i-1}$ are already fixed, then there are $(s-i+1)^2$ possibilities for U_i.

By double counting it follows that $\binom{t}{t-s+1}(s-1)!s!N(K_{s,t}, G)$ is at most the number of s-tuples $(U_1, \ldots, U_s)$ of subsets of G such that $G[U_1] \cong S_{t-s+1}$ and $G[U_i] \cong K_2$ for all $i \in [2, s]$, whence

$$\binom{t}{t-s+1}(s-1)!s!N(K_{s,t}, G) \leqslant N(S_{t-s+1}, G) \cdot \left(e(G)\right)^{s-1} .$$

Now it remains to observe $\binom{t}{t-s+1}(s-1)!s! = \frac{s!t!}{(t-s+1)!}$. □

We remark that this argument is asymptotically optimal if G is a complete bipartite graph. More precisely, for $x \leqslant x^\star(t-s+1)$ the sequence $(B(n,x))_{n=1}^{\infty}$ establishes the equality case in Theorem 1.16. This observation concludes the proof of Theorem 1.16.

In the remainder of this section we show the following explicit version of Theorem 1.18.

Theorem 6.2. *Every graph G on n vertices with $xn^2/2$ edges satisfies*

$$N(C_4, G) \leqslant \frac{x(1-x)^2}{8} n^4 + 2n^3 .$$

For the proof we need the following well-known result due to Goodman [8], whose short proof we include for the sake of completeness.

Proposition 6.3 (Goodman [8])**.** *For every real number $x \in [0,1]$, every positive integer n, and every graph G on n vertices with $xn^2/2$ edges we have*

$$\sum_{v \in V(G)} e(v) \geqslant \sum_{v \in V(G)} d(v)^2 - xn^3/2\,,$$

where $e(v) = e(G[N(v)])$ denotes the number of triangles containing the vertex v.

Proof of Proposition 6.3. Counting the number of pairs $(u, \{v, w\}) \in V(G) \times E(G)$ with $v, w \in N(u)$ in two different ways, we obtain

$$\sum_{u \in V(G)} e(u) \geqslant \sum_{vw \in G} \big(d(v) + d(w) - n\big) = \sum_{v \in V(G)} d(v)^2 - e(G) \cdot n\,. \qquad \square$$

Goodman's formula has the following consequence, which will assist us in the inductive proof of Theorem 6.2.

Corollary 6.4. *Every graph G with n vertices and $xn^2/2$ edges possesses a vertex v satisfying*

$$e(v) \geqslant \frac{d(v)^2}{2} + \frac{(1-4x+3x^2)n^2}{4} - \frac{(1-x)^3 n^3}{4(n-d(v))}\,.$$

Proof. The Cauchy-Schwarz inequality implies $\sum_{v \in V(G)} d(v)^2 \geqslant x^2 n^3$ and because of

$$\sum_{v \in V(G)} (n - d(v)) = (1-x)n^2$$

we also have

$$\sum_{v \in V(G)} \frac{1}{n - d(v)} \geqslant \frac{1}{1-x}\,.$$

Consequently,

$$\begin{aligned}\sum_{v \in V(G)} \Big(\frac{d(v)^2}{2} + \frac{(1-4x+3x^2)n^2}{4} - \frac{(1-x)^3 n^3}{4(n-d(x))}\Big) &\leqslant \sum_{v \in V(G)} \frac{d(v)^2}{2} + \frac{(x^2-x)n^3}{2} \\ &\leqslant \sum_{v \in V(G)} d(v)^2 - xn^3/2\,.\end{aligned}$$

Due to Proposition 6.3 the result now follows by averaging. $\square$

Proof of Theorem 6.2. We argue by induction on n. The base case $n \leqslant 3$ is clear, for there are no 4-cycles in graphs with less than four vertices. Now suppose $n \geqslant 4$ and that our claim holds for every graph on $n-1$ vertices.

Given a graph G on n vertices with $xn^2/2$ edges we invoke Corollary 6.4 and get a vertex $v \in V(G)$ such that

$$e \geqslant \frac{d^2}{2} + \frac{(1-4x+3x^2)n^2}{4} - \frac{(1-x)^3 n^3}{4(n-d)}\,, \tag{6.1}$$

where $d = d(v)$ and $e = e(v)$. We contend that

$$N(C_4, G) \leqslant N(C_4, G-v) + (d^2/2 - e)(n-d)\,, \tag{6.2}$$

or, in other words, that there are at most $(d^2/2 - e)(n-d)$ induced copies of $K_{2,2}$ in G which contain the vertex v. The reason for this is that each such copy contains a pair of non-adjacent members of $N(v)$ and a fourth vertex belonging to $V(G) \smallsetminus N(v)$. Clearly there are at most $d^2/2 - e$ possibilities for such a non-adjacent pair and at most $n - d$ possibilities for the fourth vertex.

Claim 6.5. *We have*

$$8N(C_4, G-v) \leqslant x(1-x)^2(n^4 - 4n^3) + 2(xn-d)(1-4x+3x^2)n^2 + 16n^3\,.$$

Proof. The induction hypothesis yields

$$8N(C_4, G-v) \leqslant x'(1-x')^2(n-1)^4 + 16(n-1)^3\,, \tag{6.3}$$

where x' is defined by

$$x' = \frac{2|E(G-v)|}{(n-1)^2} = \frac{xn^2 - 2d}{(n-1)^2}\,.$$

The function $h(x) = x(1-x)^2$ has derivatives $h'(x) = 1 - 4x + 3x^2$ and $h''(x) = -4 + 6x$. Therefore we have $\|h'\|_{[0,1]} = 1$ and $\|h''\|_{[0,1]} = 4$, where $\|\cdot\|_{[0,1]}$ denotes the supremum norm with respect to the unit interval. So Taylor's formula and (6.3) imply

$$\begin{aligned}8N(C_4, G-v) \leqslant x(1-x)^2(n-1)^4 &+ (1-4x+3x^2)(x'-x)(n-1)^4 \\ &+ 2(x'-x)^2(n-1)^4 + 16(n-1)^3\,.\end{aligned}$$

Here

$$x(1-x)^2(n-1)^4 \leqslant x(1-x)^2(n^4 - 4n^3 + 6n^2) \leqslant x(1-x)^2(n^4-4n^3) + n^2$$

and due to

$$x' - x = \frac{(2n-1)x - 2d}{(n-1)^2} \tag{6.4}$$

we have $2(x'-x)^2(n-1)^4 = 2\big|(2n-1)x - 2d\big|^2 \leqslant 8n^2$. For these reasons it suffices to establish

$$(1-4x+3x^2)(x'-x)(n-1)^4 \leqslant 2(xn-d)(1-4x+3x^2)n^2 + 7n^2\,. \tag{6.5}$$

Now the triangle inequality yields

$$\begin{aligned}
&\big|(x'-x)(n-1)^4 - 2(xn-d)n^2\big| \\
&\qquad\qquad \leqslant \big|(x'-x)(n-1)^2 - 2(xn-d)\big|(n-1)^2 + 2|xn-d|\big(n^2-(n-1)^2\big) \\
&\qquad\qquad \overset{(6.4)}{\leqslant} x(n-1)^2 + 4n^2 \leqslant 5n^2
\end{aligned}$$

and together with $\|h'\|_{[0,1]} = 1$ this proves (6.5). Thereby Claim 6.5 is proved. □

Now combining (6.1), (6.2), and Claim 6.5 we obtain

$$\begin{aligned}
8N(C_4, G) &\leqslant x(1-x)^2(n^4-4n^3) + 2(xn-d)(1-4x+3x^2)n^2 + 16n^3 \\
&\qquad\qquad - 2(1-4x+3x^2)n^2(n-d) + 2(1-x)^3n^3 \\
&= x(1-x)^2n^4 + 16n^3\,,
\end{aligned}$$

as desired. □

§7. Concluding remarks

7.1. **General questions.** As the example $Q = K_3 + \overline{K}_3$ shows, for a quantum graph Q the function $I(Q, x)$ can have at least two global maxima. We do not know whether this is possible for single graphs F as well.

Problem 7.1. *Does there exist a graph F such that the function $I(F, x)$ has at least two global maxima?*

Two questions of a similar flavor are as follows.

Problem 7.2. *Does there exist a graph F such that for some nontrivial interval J we have $I(F, x) = \mathrm{ind}(F)$ for all $x \in J$?*

Problem 7.3. *Does there exist a graph F such that the function $I(F, x)$ has a nontrivial local maximum (minimum)?*

Recall that for a self-complementary graph F the function $I(F, x)$ is symmetric around $x = 1/2$. One may thus wonder whether some appropriate self-complementary graph F yields an affirmative solution to Problem 7.1. This approach leads to the following question.

Problem 7.4. *Let F be a self-complementary graph. Is it true that $I(F, x) = \mathrm{ind}(F)$ holds if and only if $x = 1/2$?*

7.2. **Problems for specific graphs.** The smallest problem left open by our results on stars in Section 5 is to determine $I(S_3, x)$ for $x \in [1/2, 1]$. In an interesting contrast to the case $S_2 = K_3^-$ one can show that the clique density construction (see Construction 1.9) is not extremal for this problem. For $x \in [4\sqrt{2} - 5, 1]$ the best construction we are aware of is the complement of a clique of order $\lfloor (1-x)^{1/2} n \rfloor$, which leads to the bound

$$I(S_3, x) \geqslant 4(1 - (1-x)^{1/2})(1-x)^{3/2} . \tag{7.1}$$

For $x \in [0.91, 0.93]$ we have a complicated argument based on the results in [24] which shows that equality holds in (7.1). In the range $x \in [1/2, 4\sqrt{2} - 5)$ the complement of two disjoint cliques of order $\lfloor ((1-x)/2)^{1/2} n \rfloor$ shows that $I(S_3, x)$ is strictly larger than the right side of (7.1). We hope to return to this problem in the near future.

Finally, we would like to emphasize Conjecture 1.17 again: Is it true that for $x \in [1/2, 1]$ the graphs in Construction 1.9 minimizing the triangle density maximize the induced C_4 density?

References

[1] N. Alon, D. Hefetz, M. Krivelevich, and M. Tyomkyn, *Edge-statistics on large graphs*, Combin. Probab. Comput. **29** (2020), no. 2, 163–189, DOI 10.1017/s0963548319000294. MR4079631 ↑1.4

[2] B. Bollobás, *On complete subgraphs of different orders*, Math. Proc. Cambridge Philos. Soc. **79** (1976), no. 1, 19–24, DOI 10.1017/S0305004100052063. MR424614 ↑1.3

[3] J. I. Brown and A. Sidorenko, *The inducibility of complete bipartite graphs*, J. Graph Theory **18** (1994), no. 6, 629–645, DOI 10.1002/jgt.3190180610. MR1292981 ↑1.5, 5, 5.3, 5.4

[4] S. Bubeck and N. Linial, *On the local profiles of trees*, J. Graph Theory **81** (2016), no. 2, 109–119, DOI 10.1002/jgt.21865. MR3433633 ↑1.1

[5] P. Erdős, *On the number of complete subgraphs contained in certain graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **7** (1962), 459–464 (English, with Russian summary). MR151956 ↑1.1

[6] C. Even-Zohar and N. Linial, *A note on the inducibility of 4-vertex graphs*, Graphs Combin. **31** (2015), no. 5, 1367–1380, DOI 10.1007/s00373-014-1475-4. MR3386015 ↑1.2

[7] J. Fox and L. Sauermann, *A completion of the proof of the edge-statistics conjecture*, Adv. Comb., posted on 2020, Paper No. 4, 52, DOI 10.19086/aic.12047. MR4125345 ↑1.4

[8] A. W. Goodman, *On sets of acquaintances and strangers at any party*, Amer. Math. Monthly **66** (1959), 778–783, DOI 10.2307/2310464. MR107610 ↑1.1, 1.3, 1.7, 6, 6.3

[9] H. Hatami and S. Norin, *On the boundary of the region defined by homomorphism densities*, J. Comb. **10** (2019), no. 2, 203–219, DOI 10.4310/JOC.2019.v10.n2.a1. MR3912211 ↑1.1, 1.2

[10] H. Hatami and S. Norine, *Undecidability of linear inequalities in graph homomorphism densities*, J. Amer. Math. Soc. **24** (2011), no. 2, 547–565, DOI 10.1090/S0894-0347-2010-00687-X. MR2748400 ↑1.1

[11] J. Hirst, *The inducibility of graphs on four vertices*, J. Graph Theory **75** (2014), no. 3, 231–243, DOI 10.1002/jgt.21733. MR3153118 ↑1.4

[12] H. Huang, N. Linial, H. Naves, Y. Peled, and B. Sudakov, *On the 3-local profiles of graphs*, J. Graph Theory **76** (2014), no. 3, 236–248, DOI 10.1002/jgt.21762. MR3200287 ↑1.1

[13] G. Katona, *A theorem of finite sets*, Theory of graphs (Proc. Colloq., Tihany, 1966), Academic Press, New York, 1968, pp. 187–207. MR0290982 ↑1.1, 1.3

[14] J. B. Kruskal, *The number of simplices in a complex*, Mathematical optimization techniques, Univ. of California Press, Berkeley, Calif., 1963, pp. 251–278. MR0154827 ↑1.1, 1.3

[15] M. Kwan, B. Sudakov, and T. Tran, *Anticoncentration for subgraph statistics*, J. Lond. Math. Soc. (2) **99** (2019), no. 3, 757–777, DOI 10.1112/jlms.12192. MR3977889 ↑1.4

[16] H. Liu, O. Pikhurko, and K. Staden, *The exact minimum number of triangles in graphs with given order and size*, Forum Math. Pi **8** (2020), e8, 144, DOI 10.1017/fmp.2020.7. MR4089395 ↑1.3

[17] X. Liu and D. Mubayi, *The feasible region of hypergraphs*, J. Combin. Theory Ser. B **148** (2021), 23–59, DOI 10.1016/j.jctb.2020.12.004. MR4193065 ↑2

[18] A. Martinsson, F. Mousset, A. Noever, and M. Trujić, *The edge-statistics conjecture for $\ell \ll k^{6/5}$*, Israel J. Math. **234** (2019), no. 2, 677–690, DOI 10.1007/s11856-019-1929-8. MR4040841 ↑1.4

[19] V. Nikiforov, *The number of cliques in graphs of given order and size*, Trans. Amer. Math. Soc. **363** (2011), no. 3, 1599–1618, DOI 10.1090/S0002-9947-2010-05189-X. MR2737279 ↑1.1

[20] D. Olpp, *A conjecture of Goodman and the multiplicities of graphs*, Australas. J. Combin. **14** (1996), 267–282. MR1424340 ↑1.3, 1.7

[21] N. Pippenger and M. C. Golumbic, *The inducibility of graphs*, J. Combinatorial Theory Ser. B **19** (1975), no. 3, 189–203, DOI 10.1016/0095-8956(75)90084-2. MR401552 ↑1.1

[22] A. A. Razborov, *On the minimal density of triangles in graphs*, Combin. Probab. Comput. **17** (2008), no. 4, 603–618, DOI 10.1017/S0963548308009085. MR2433944 ↑1.1

[23] Chr. Reiher, *The clique density theorem*, Ann. of Math. (2) **184** (2016), no. 3, 683–707, DOI 10.4007/annals.2016.184.3.1. MR3549620 ↑1.1, 1.10

[24] Chr. Reiher and S. Wagner, *Maximum star densities*, Studia Sci. Math. Hungar. **55** (2018), no. 2, 238–259, DOI 10.1556/012.2018.55.2.1395. MR3813354 ↑7.2

[25] R. H. Schelp and A. Thomason, *A remark on the number of complete and empty subgraphs*, Combin. Probab. Comput. **7** (1998), no. 2, 217–219, DOI 10.1017/S0963548397003234. MR1617934 ↑3, 3.1

[26] A. Thomason, *A disproof of a conjecture of Erdős in Ramsey theory*, J. London Math. Soc. (2) **39** (1989), no. 2, 246–255, DOI 10.1112/jlms/s2-39.2.246. MR991659 ↑1.1

[27] P. Turán, *Eine Extremalaufgabe aus der Graphentheorie*, Mat. Fiz. Lapok **48** (1941), 436–452 (Hungarian, with German summary). MR18405 ↑1.1

Department of Mathematics, Statistics, and Computer Science, University of Illinois, Chicago, IL 60607 USA

E-mail address: `xliu246@uic.edu`

E-mail address: `mubayi@uic.edu`

Fachbereich Mathematik, Universität Hamburg, Hamburg, Germany

E-mail address: `Christian.Reiher@uni-hamburg.de`