Bull. Aust. Math. Soc. **111** (2025), 405–415
doi:10.1017/S0004972724000972

HOW MANY TRIANGLES CAN A GRAPH HAVE?

TEERADEJ KITTIPASSORN ✉ **and KAMIL POPIELARZ**

(Received 17 May 2024; accepted 14 September 2024; first published online 14 November 2024)

Abstract

We investigate the set T_n of the possible number of triangles in a graph on n vertices. The first main result says that every natural number less than $\binom{n}{3} - (\sqrt{2} + o(1))n^{3/2}$ belongs to T_n. On the other hand, we show that there is a number $m = \binom{n}{3} - (\sqrt{2} + o(1))n^{3/2}$ that is not a member of T_n. In addition, we prove that there are two interlacing sequences $\binom{n}{3} - (\sqrt{2} + o(1))n^{3/2} = c_1 \le d_1 \le c_2 \le d_2 \le \cdots \le c_s \le d_s = \binom{n}{3}$ with $|c_t - d_t| = n - 2 - \binom{s-t+1}{2}$ such that $(c_t, d_t) \cap T_n = \emptyset$ for all t. Moreover, we obtain a generalisation of these results for the set of the possible number of copies of a connected graph H in a graph on n vertices.

2020 *Mathematics subject classification*: primary 05C35.

Keywords and phrases: the number of triangles in a graph, realisable numbers, complete graphs in a graph, subgraphs in a graph.

1. Introduction

We ask the following natural question: given a graph H and a natural number n, what are the possible values of m such that there exists a graph on n vertices with *exactly* m copies of H? Surprisingly, very little is known about this problem.

A question of this flavour was considered by Kittipassorn and Mészáros [9] who studied the set F_n of the possible number of *frustrated triangles*, that is, triples of vertices inducing an odd number of edges. They proved that about two thirds of the numbers in $[0, n^{3/2}]$ do not appear in F_n and that every even number between $(1 + o(1))n^{3/2}$ and $\binom{n}{3} - (1 + o(1))n^{3/2}$ is a member of F_n for sufficiently large even n.

Much more attention has been given to the problem of maximising or minimising the number of subgraphs of certain types of graphs with a given number of vertices and edges. For example, Rademacher proved that every graph with $\lfloor n^2/4 \rfloor + 1$ edges contains at least $\lfloor n/2 \rfloor$ triangles. Erdős [5] posed a conjecture, which was later proved by Lovász and Simonovits [11], that a graph of size $\lfloor n^2/4 \rfloor + k$ contains at least $k\lfloor n/2 \rfloor$ triangles if $k < n/2$. On the other hand, Alon [1] investigated the maximum number

The first author was partially supported by Office of the Permanent Secretary, Ministry of Higher Education, Science, Research and Innovation grant RGNS 63-016.

© The Author(s), 2024. Published by Cambridge University Press on behalf of Australian Mathematical Publishing Association Inc.

https://doi.org/10.1017/S0004972724000972 Published online by Cambridge University Press

of subgraphs isomorphic to some given graph, where the maximum is taken over all graphs of a certain size. We refer to [3, 4, 6–8, 10, 12, 13] for related results.

We consider a fixed connected graph H. For a graph G, we define $k_H(G)$ to be the number of copies of H in G. The main object of interest of this paper is

$$S_H^{(n)} = \{k_H(G) : |G| = n\},$$

the set of the possible number of copies of H in a graph on n vertices. Our first result says that almost every number (in the appropriate range) is realisable as a number of copies of H in some graph of order n.

THEOREM 1.1. $[0, (1 - o(1))k_H(K_n)] \subset S_H^{(n)}$ *as* $n \to \infty$.

It is not unreasonable to expect that $o(1)$ above could be replaced by 0. As it happens, this is not the case. Before we state the second result, we introduce some notation that plays a major role in the rest of the paper. For a graph $G = (V, E)$ of order n, let $f_H(G)$ be the number of subgraphs F of $K(V)$ isomorphic to H such that $E(F) \cap E \neq \emptyset$, where $K(V)$ denotes the complete graph on the vertex set V. For example, $f_{K_3}(G)$ is the number of triples of vertices in G inducing at least one edge. Observe that $k_H(G) = k_H(K_n) - f_H(\overline{G})$, and therefore we work instead with the complement of G which is easier to draw when G is dense. We write $a_H^{(n)}(t) = f_H(S_t^{(n)})$ and $b_H^{(n)}(t) = f_H(M_t^{(n)})$, where $S_t^{(n)}$ is a graph on n vertices with t edges forming a star with $t \leq n - 1$ and $M_t^{(n)}$ is a graph on n vertices with t edges forming a matching with $t \leq \lfloor n/2 \rfloor$. We prove the existence of some gaps in $S_H^{(n)}$.

THEOREM 1.2. *For all sufficiently large* n *and all* $t \geq 0$,

$$|(k_H(K_n) - a_H^{(n)}(t + 1), k_H(K_n) - b_H^{(n)}(t)) \cap S_H^{(n)}| = o(n^{h-2}).$$

In particular, there exists a number $m = k_H(K_n) - (c_H + o(1))n^{h-3/2}$ *that is not a member of* $S_H^{(n)}$ *for some constant* c_H.

In Section 4, we see that, when $t < c\sqrt{n}$, where c is some nonnegative constant depending only on H, the length of the interval $a_H^{(n)}(t + 1) - b_H^{(n)}(t)$ is of order n^{h-2}. When H is a triangle, we prove sharp analogues of Theorems 1.1 and 1.2.

THEOREM 1.3. *We have:*

(i) $[0, \binom{n}{3} - (\sqrt{2} + o(1))n^{3/2}] \subset S_{K_3}^{(n)}$ *as* $n \to \infty$*; and*

(ii) $(\binom{n}{3} - a_{K_3}^{(n)}(t + 1), \binom{n}{3} - b_{K_3}^{(n)}(t)) \cap S_{K_3}^{(n)} = \emptyset$ *for all* $n, t \geq 0$.

We see that $a_{K_3}^{(n)}(t) = t(n - 2) - \binom{t}{2}$ and $b_{K_3}^{(n)}(t) = t(n - 2)$, so it is easy to check that the interval in the second part of the theorem is not empty as long as $t \lesssim \sqrt{2n}$. Thus, there exists a number $m = \binom{n}{3} - (\sqrt{2} + o(1))n^{3/2}$ that is not a member of T_n. Therefore, the first part of Theorem 1.3 is sharp.

The same idea that we use in the case of triangles also works for cherries, that is, paths with two edges, P_2.

https://doi.org/10.1017/S0004972724000972 Published online by Cambridge University Press

THEOREM 1.4. *We have:*

(i) $[0, 3\binom{n}{3} - (4 + o(1))n^{3/2}] \subset S^{(n)}_{P_2}$ *as* $n \to \infty$*; and*

(ii) $(3\binom{n}{3} - a^{(n)}_{P_2}(t+1), 3\binom{n}{3} - b^{(n)}_{P_2}(t)) \cap S^{(n)}_{P_2} = \emptyset$ *for all* $n, t \geq 0$.

As in the case of triangles, the first part of Theorem 1.4 is sharp.

The rest of this paper is organised as follows. In Section 2, we prove some preliminary lemmas. We prove the sharp results for triangles (Theorem 1.3) in Section 3. The proofs of Theorems 1.1 and 1.2 are presented in Section 4. We conclude the paper in Section 5 with some open problems.

2. Complete graphs

In this section, we consider the case when $H = K_r$ is a complete graph and prove some lemmas that we use to prove our main theorems.

We start by showing that the first $(\lfloor (n - r^2)/r \rfloor)^r$ natural numbers are realisable: that is, for every $k \leq (\lfloor (n - r^2)/r \rfloor)^r$, there exists an r-partite graph on n vertices with exactly k copies of K_r.

LEMMA 2.1. *For natural numbers* n, r *with* $n \geq r \geq 2$*, and any nonnegative integer* $k \leq (\lfloor (n - r^2)/r \rfloor)^r$*, there is an* r*-partite graph with* k *copies of* K_r.

PROOF. Let $P_{n,r}$ be the set of the possible number of copies of K_r in an r-partite graph on n vertices. Clearly, $P_{n,r} \subseteq S^{(n)}_{K_r}$.

We use induction on r. The base case $r = 2$ is trivial. Suppose that the assertion holds for some $r \geq 2$. Let $G = (V_1 \sqcup V_2 \sqcup V_3, E)$, where $|V_1| = \lceil r(n-r)/(r+1) \rceil$, $|V_2| = \lfloor (n-r)/(r+1) \rfloor$, every vertex in V_1 is joined to every vertex in V_2, and V_3 induces a K_r. By the induction hypothesis, we can replace V_1 by an r-partite graph having k copies of K_r, and therefore we obtain an $(r+1)$-partite graph having $\lfloor (n-r)/(r+1) \rfloor k$ copies of K_{r+1}, for any

$$k \leq \left(\left\lfloor \frac{\frac{r(n-r)}{r+1} - r^2}{r} \right\rfloor \right)^r = \left(\left\lfloor \frac{n-r}{r+1} - r \right\rfloor \right)^r.$$

Therefore, we get an increasing sequence in $P_{n,r+1}$, starting with 0 and ending with

$$\left\lfloor \frac{n-r}{r+1} \right\rfloor \left(\left\lfloor \frac{n-r}{r+1} - r \right\rfloor \right)^r \geq \left(\left\lfloor \frac{n-r}{r+1} - r \right\rfloor \right)^{r+1} = \left(\left\lfloor \frac{n - r(r+2)}{r+1} \right\rfloor \right)^{r+1} \geq \left(\left\lfloor \frac{n-(r+1)^2}{r+1} \right\rfloor \right)^{r+1},$$

such that the difference between consecutive terms is equal to $|V_2|$. To obtain the missing numbers between consecutive terms, notice that it is enough to join the needed number of vertices in V_2 to every vertex in V_3. Clearly, all graphs in the sequence are $(r+1)$-partite. □

Recall that $f_H(G)$ is the number of subgraphs F of $K(V)$ isomorphic to H such that $E(F) \cap E \neq \emptyset$, where $K(V)$ denotes the complete graph on the vertex set V. Moreover, $a^{(n)}_H(t) = f_H(S^{(n)}_t)$ and $b^{(n)}_H(t) = f_H(M^{(n)}_t)$, where $S^{(n)}_t$ is a graph on n vertices with t edges forming a star and $M^{(n)}_t$ is a graph on n vertices with t edges forming a matching.

https://doi.org/10.1017/S0004972724000972 Published online by Cambridge University Press

We fix $n, r \geq 2$ and, for brevity, write $f(G) = f_{K_r}(G)$, $a_t = a^{(n)}_{K_r}(t)$, $b_t = b^{(n)}_{K_r}(t)$. Recall that the number of copies of K_r in G is equal to $\binom{n}{r} - f(\overline{G})$, where $f(G)$ is the number of r-sets of vertices in G that induce at least one edge. Therefore, we work instead with the complement of G, which is easier to deal with when G is dense. Observe that $a_t = \sum_{i=1}^{t} \binom{n-1-i}{r-2}$ and hence $a_{t+1} - a_t = \binom{n-1-(t+1)}{r-2}$.

LEMMA 2.2. *For a graph G on n vertices and $e \leq \frac{1}{2}(n-1)$ edges, $f(G) \in [a_e, b_e]$.*

PROOF. We show this by induction on the number of edges e. In the base case when $e \leq 1$, there is nothing to show. For $e > 1$, assume that $f(G') \in [a_{e-1}, b_{e-1}]$ for any graph G' on n vertices and $e-1$ edges.

First, we show that $f(G) \leq b_e$. Take an edge $xy \in G$ such that $d(y) > 1$ and an isolated vertex $w \in G$. Let G' be a graph obtained by removing xy from G and replacing it by xw. Note that $f(G') \geq f(G)$. By repeating this process for any nonindependent edge, we eventually obtain a matching, since $e \leq \frac{1}{2}(n-1)$, without decreasing the value of f.

To show that $f(G) \geq a_e$, pick an edge $xy \in G$ and let G' be a graph obtained by removing xy from G. We show that $f(G) - f(G') \geq \binom{n-1-e}{r-2} = a_e - a_{e-1}$, which completes the proof, as then $f(G) \geq f(G') + a_e - a_{e-1} \geq a_e$, by the induction hypothesis applied to G'. Let $A = V(G) \backslash (N_G(x) \cup N_G(y))$ (observe that $x, y \neq A$). Write M for a largest independent set contained in A and write e_A for the number of edges induced by A. By deleting an endpoint of each edge in A, we see that $|M| \geq |A| - e_A$. Therefore,

$$f(G) - f(G') \geq \binom{|M|}{r-2} \geq \binom{n - |N(x) \cup N(y)| - e_A}{r-2} \geq \binom{n-e-1}{r-2}. \qquad \square$$

We remark that Lemma 2.2 does not imply that $f(G) \notin (b_t, a_{t+1})$ for any t. However, the result follows immediately from the monotonicity of f.

LEMMA 2.3. *For any graph G on n vertices and any $t \geq 0$, $f(G) \notin (b_t, a_{t+1})$.*

PROOF. Write $t_{\max} = \max\{t : b_{t-1} + 1 < a_t\}$ for the last t where there is a gap between the intervals $[a_{t-1}, b_{t-1}]$ and $[a_t, b_t]$. It is enough to show that $f(G) \in [a_t, b_t]$ for some $t \leq t_{\max}$ or $f(G) \geq a_{t_{\max}}$. This follows from Lemma 2.2 if $e(G) \leq t_{\max}$. So we can assume that $e(G) > t_{\max}$. Let G' be a graph obtained from G by deleting some edges until there are exactly $t_{\max}$ edges left. By the monotonicity of f and Lemma 2.2, $f(G) \geq f(G') \geq a_{t_{\max}}$. $\square$

We remark that $t_{\max} = \Theta(\sqrt{n})$.

We now turn our attention to the problem of determining which numbers can be expressed as a *restricted* sum of binomials. More precisely, we are interested in the set, A_k, of integers that can be expressed as $\sum_{i=1}^{j} \binom{d_i}{2}$, where $\sum_{i=1}^{j} d_i = k$ and each d_i is a positive integer. Let s_k be the largest integer such that any integer between 0 and s_k belongs to A_k. To the best of our knowledge, this sequence was first studied by Dogan (see [2, Lemma 20]) who showed that $s_k \geq \frac{1}{2}(k - 2\sqrt{k})^2$. Here, we present a

https://doi.org/10.1017/S0004972724000972 Published online by Cambridge University Press

simpler proof of an asymptotically stronger result of this kind. We start with a simple observation.

OBSERVATION 2.4. Let z_k be an increasing sequence such that $s_k \geq z_k > 0$ for every k. For any positive integer k, we have $s_k \geq \binom{k^*}{2}$, where k^* is the smallest nonnegative integer such that $z_{k-k^*} < k^*$.

PROOF. Take any $\ell \leq \binom{k^*}{2}$. We show that ℓ can be written as a *restricted* sum of binomials. When $\ell = \binom{k^*}{2}$, there is nothing to show, so we can assume that $\ell < \binom{k^*}{2}$. Let k' be the largest integer such that $\binom{k'}{2} \leq \ell$. Write ℓ' for $\ell - \binom{k'}{2}$ and observe that $\ell' = \ell - \binom{k'}{2} < \binom{k'+1}{2} - \binom{k'}{2} = k' < k^*$. By the definition of k^*, $s_{k-k'} \geq z_{k-k'} \geq k' > \ell'$, and the result follows from the definition of $s_{k-k'}$. □

We obtain a lower bound for s_k by applying Observation 2.4 several times.

LEMMA 2.5. $s_k \geq \frac{1}{2}(k - \sqrt{2k})^2 + O(k)$ *as* $k \to \infty$.

PROOF. First, $s_k \geq \lfloor k/2 \rfloor$, as we can write any $\ell \leq \lfloor k/2 \rfloor$ as $\sum_{i=1}^{\ell} \binom{2}{2} + \sum_{i=1}^{k-2\ell} \binom{1}{2}$. For every $k' \leq \lfloor k/3 \rfloor$, we have $\lfloor (k-k')/2 \rfloor \geq k'$; therefore, $s_k \geq \binom{\lfloor k/3 \rfloor}{2}$ by Observation 2.4. Applying Observation 2.4 again yields $s_k \geq \binom{k-3\sqrt{2k}}{2}$, and by applying it one more time, we arrive at the statement of the lemma. □

3. Triangles

We now consider the case when $H = K_3$, that is, when H is a triangle. For brevity, we write $T_n = S^{(n)}_{K_3}$, $f(G) = f_{K_3}(G)$, $a_t = a^{(n)}_{K_3}(t)$ and $b_t = b^{(n)}_{K_3}(t)$ in this section.

Before we improve Lemma 2.1 to Theorem 1.3(i), we consider Theorem 1.3(ii) and look for nonmembers of T_n. Recall that the number of triangles in G is equal to $\binom{n}{3} - f(\overline{G})$, where $f(G)$ is the number of triples of vertices in G that induce at least one edge. Notice that we have a simple formula, $f(G) = e(G)(n-2) - n_c + n_t$, where n_c is the number of cherries (that is, paths with two edges, P_2) and n_t is the number of triangles in G. This comes from the fact that each edge is contained in exactly $n-2$ triples, but we double or triple count the triples which contain more than one edge. Using this formula, it is easy to see that $a_t = t(n-2) - \binom{t}{2}$ and $b_t = t(n-2)$.

We showed in Lemma 2.3 that $f(G) \notin (b_t, a_{t+1})$ for all $t \geq 0$. On the other hand, we now prove that every number bigger than $(\sqrt{2} + o(1))n^{3/2}$ is realisable.

Observe that if G is triangle-free, then

$$f(G) = e(G)(n-2) - n_c = e(G)(n-2) - \sum_{x \in G} \binom{d(x)}{2}.$$

LEMMA 3.1. *Let k be such that $s_k \geq n-2$ and $2k \leq n$. For every m in the interval $[k(n-2), \binom{n-2k}{3} + k(n-2)]$, there is a graph G on n vertices such that $f(G) = m$.*

https://doi.org/10.1017/S0004972724000972 Published online by Cambridge University Press

PROOF. Take any $m \in [k(n-2), \binom{n-2k}{3} + k(n-2)]$. We construct a graph G on the vertex set $U \sqcup V$, where $|U| = n-2k$, $|V| = 2k$, such that $f(G) = m$. Observe, first, that since adding an edge to a graph increases the value of f by at most $n-2$, we can find a configuration of edges in U that contributes m' to f, where $0 \le m' + k(n-2) - m < n-2$. We now find a configuration of edges in V that contributes exactly $m' - m$ to f. By the definition of s_k, we can find a sequence of positive integers $d_1, d_2, \dots, d_j$ which sum up to k and such that $\sum_{i=1}^{j} \binom{d_i}{2} = m' + k(n-2) - m$. Let $G[V]$ be the union of stars $S_{d_1}, S_{d_2}, \dots, S_{d_j}$ and an independent set. Then $f(G) = m' + k(n-2) - (m' + k(n-2) - m) = m$. □

COROLLARY 3.2. *For any sufficiently large n, if $m \in [\sqrt{2}n^{3/2} + 4n^{5/4}, (1-o(1))\binom{n}{3}]$, then there is a graph G on n vertices such that $f(G) = m$.*

PROOF. It follows from Lemma 2.5 that $s_k \ge \binom{k-\sqrt{2k}}{2}$, and hence, if $k \ge \sqrt{2n} + 4n^{1/4}$, then $s_k \ge n-2$ for sufficiently large n. The rest follows from Lemma 3.1. □

We are now ready to deduce Theorem 1.3 from Lemmas 2.1, 2.3 and Corollary 3.2.

PROOF OF THEOREM 1.3. (i) Recall that the number of triangles in G is $\binom{n}{3} - f(\overline{G})$. Therefore, Corollary 3.2 implies that $[o(\binom{n}{3}), \binom{n}{3} - (\sqrt{2} + o(1))n^{3/2}] \subset T_n$. Together with Lemma 2.1, which, for $r = 3$, says that $[0, ((n-9)/3)^3] \subset T_n$, we conclude that $[0, \binom{n}{3} - (\sqrt{2} + o(1))n^{3/2}] \subseteq T_n$ for sufficiently large n.

(ii) Since the number of triangles in G is equal to $\binom{n}{3} - f(\overline{G})$, we obtain, using Lemma 2.3, that $(\binom{n}{3} - a_{t+1}, \binom{n}{3} - b_t) \cap T_n = \emptyset$ for all $n, t \ge 0$. □

4. General H

Now, we consider the case when H is an arbitrary connected graph on $h \ge 3$ vertices. We start by showing that the first $(1-o(1))k_H(K_n)$ numbers are realisable when n goes to infinity.

Our strategy is to recursively partition the vertex set into two subsets and modify the edges between vertices in each of the classes, but without adding edges between the classes. Let $g_H = g_H^{(n)}$ be the maximum number of new copies of H obtained by adding an edge to a graph, over all graphs on n vertices. We claim that there is a constant $c_H > 0$ such that $g_H \le c_H n^{h-2}$. Indeed, a new copy must contain both of the end vertices of the newly added edge, there are $\binom{n-2}{h-2}$ h-sets of vertices in G containing two fixed vertices and each h-set may contain at most $h!$ copies of H. Therefore, $g_H \le h!\binom{n-2}{h-2} \le c_H n^{h-2}$.

The next two lemmas are needed in our construction.

LEMMA 4.1. *If $[0, cn^\alpha] \subset S_H^{(n)}$ for all sufficiently large n, where $\alpha \le h-2$, then, for all sufficiently large n and some new constant $c_1 > 0$, we have $[0, c_1 n^{\alpha h/(h-2)}] \subset S_H^{(n)}$.*

https://doi.org/10.1017/S0004972724000972 Published online by Cambridge University Press

PROOF. Consider an empty graph G with vertex set $V = V_1 \sqcup V_2$, where $n_1 = |V_1| = c'n^{\alpha/(h-2)}$ and $n_2 = |V_2| = n - |V_1|$, and where $c' > 0$ will be chosen later. Let $G_0 = G$ and let G_{i+1} be a graph obtained by adding an edge between vertices of V_1 in G_i. Since H is connected,

$$k_H(G_{i+1}) - k_H(G_i) \le g_H^{(n_1)} \le c_H n_1^{h-2} = c_H c'^{h-2} n^{\alpha}.$$

Therefore, we obtain an increasing sequence in $S_H^{(n)}$, starting with 0 and ending with $k_H(G_{\binom{n_1}{2}})$, such that the differences between consecutive terms are at most $c_H c'^{h-2} n^{\alpha}$. We modify $G_i[V_2]$ to obtain the missing numbers between consecutive terms. By the hypothesis, we can modify $G_i[V_2]$ to obtain $G_i'[V_2]$ containing any number k of copies of H, where $k \in [0, cn_2^{\alpha}]$. Hence, it suffices to find $c' > 0$ such that $c_H c'^{h-2} n^{\alpha} < cn_2^{\alpha}$. We consider two cases depending on α.

(1) If $\alpha = h - 2$, then $cn_2^{\alpha} = c(n - c'n)^{\alpha} = c(1 - c')^{\alpha} n^{\alpha}$. Therefore, it suffices to choose $c' > 0$ such that $c_H c'^{h-2} < c(1 - c')$. This gives $g_H^{(n_1)} < cn_2^{\alpha}$.
(2) If $\alpha < h - 2$, then $cn_2^{\alpha} = c(n - o(n))^{\alpha} \sim cn^{\alpha}$. If we choose $c' > 0$ such that $c' < (c/c_H)^{1/(h-2)}$, then, for sufficiently large n, we again have $g_H^{(n_1)} < cn_2^{\alpha}$.

Therefore, every number less than $k_H(G_{\binom{n_1}{2}}) = k_H(K_{n_1}) > c''n_1^h = c_1 n^{\alpha h/(h-2)}$ is realisable. □

If we have the same hypothesis, but with $\alpha > h - 2$ instead, from the next lemma we see that, for sufficiently large n, we can construct a graph on n vertices with k copies of H for any $k \le (1 - o(1))k_H(K_n)$.

LEMMA 4.2. *If $[0, cn^{\alpha}] \subset S_H^{(n)}$ for sufficiently large n, where $\alpha > h - 2$, then, for sufficiently large n, $[0, (1 - o(1))k_H(K_n)] \subset S_H^{(n)}$.*

PROOF. Proceeding as in Lemma 4.1, choose $\beta \in ((h-2)/\alpha, 1)$ and let $n_2 = |V_2| = n^{\beta}$ and $n_1 = |V_1| = n - |V_2|$. Note that $g_H^{(n_1)} = O(n^{h-2})$. By the hypothesis, we can modify $G[V_2]$ to obtain any number of copies of H up to cn_2^{α}, where $cn_2^{\alpha} = \omega(n^{h-2})$. Therefore, every number in the interval $[0, k_H(K_{n_1})]$ is realisable. However, $n_1 = (1 - o(1))n$, and hence $k_H(K_{n_1}) = (1 - o(1))k_H(K_n)$. □

PROOF OF THEOREM 1.1. We start by showing that, trivially, $[0, \lfloor n/h \rfloor] \subset S_H^{(n)}$. To achieve that, note that, for any $k \le \lfloor n/h \rfloor$, we can simply construct a graph on n vertices consisting of k disjoint copies of H.

Let $k_{\max}$ be the largest integer k such that $(h/(h-2))^k \le h$ (note that $(h/(h-2))^{k_{\max}} \in (h-2, h]$). We claim that $[0, c_k n^{(h/(h-2))^k}] \subset S_H^{(n)}$ for every positive integer $k \le k_{\max}$ and sufficiently large n. We prove the claim by induction on k. For $k = 0$, we already have $[0, c_0 n] \subset S_H^{(n)}$. Suppose that $[0, c_k n^{(h/(h-2))^k}] \in S_H^{(n)}$ and $k < k_{\max}$. Observe, first, that $(h/(h-2))^k \le h - 2$, as otherwise $(h/(h-2))^{k_{\max}}$ would be greater than h. Hence, we can apply Lemma 4.1 and conclude that $[0, c_{k+1} n^{(h/(h-2))^{k+1}}] \in S_H^{(n)}$ for large enough n. Note that we apply Lemma 4.1 only finitely many times and hence n remains finite.

https://doi.org/10.1017/S0004972724000972 Published online by Cambridge University Press

Therefore, for n large enough, we have $[0, cn^{\alpha}] \subset S_H^{(n)}$ with $\alpha = (h/(h-2))^{k_{\max}} \in (h-2, h]$, and hence we can apply Lemma 4.2 and conclude that, for n sufficiently large, $[0, (1-o(1))k_H(K_n)] \subset S_H^{(n)}$. □

We recall the major notation. For a graph $G = (V, E)$ of order n, let $f_H(G)$ be the number of subgraphs F of $K(V)$ isomorphic to H such that $E(F) \cap E \neq \emptyset$, where $K(V)$ denotes the complete graph on the vertex set V. Then the number of copies of H in G is equal to $k_H(K_n) - f_H(\overline{G})$. In the next lemma, we describe a formula for f_H.

LEMMA 4.3. *For any graphs H on h vertices and G on n vertices,*

$$f_H(G) = c_H e(G)\binom{n-2}{h-2} - \sum_{k=2}^{e(H)} \sum_{\substack{e(F)=k \\ \delta(F)>0}} (-1)^k c_H(F) k_F(G) \binom{n-|F|}{h-|F|}.$$

PROOF. Let $E(G) = \{e_1, \dots, e_m\}$, where $m = e(G)$. Denote the set of subgraphs of the complete graph on $V(G)$ isomorphic to H containing the edge e_i by A_i, that is, $A_i = \{F \subset K(V) : e_i \in E(F), F \cong H\}$. Note that $f_H(G) = |\bigcup_{i=1}^m A_i|$. Therefore, by the inclusion–exclusion principle,

$$f_H(G) = \sum_{k=1}^{m} \sum_{i_1 < \cdots < i_k} (-1)^{k+1} |A_{i_1} \cap \cdots \cap A_{i_k}|.$$

For a graph F on at most h vertices, let $c_H(F)$ be the number of copies of H in the complete graph K_h containing a fixed subgraph of the complete graph K_h, isomorphic to F. Let $F = G[e_{i_1}, \dots, e_{i_k}]$ be the graph induced by the edges $e_{i_1}, \dots, e_{i_k}$. Observe that $|A_{i_1} \cap \cdots \cap A_{i_k}| = c_H(F)\binom{n-|F|}{h-|F|}$. Therefore,

$$\begin{aligned} f_H(G) &= \sum_{k=1}^{m} \sum_{i_1 < \cdots < i_k} (-1)^{k+1} |A_{i_1} \cap \cdots \cap A_{i_k}| \\ &= \sum_{k=1}^{m} \sum_{\substack{F \subseteq H \\ e(F)=k \\ \delta(F)>0}} (-1)^{k+1} k_F(G) c_H(F) \binom{n-|F|}{h-|F|} \\ &= c_H m \binom{n-2}{h-2} - \sum_{k=2}^{e(H)} \sum_{\substack{e(F)=k \\ \delta F>0}} (-1)^k c_H(F) k_F(G) \binom{n-|F|}{h-|F|}. \end{aligned}$$

□

The following easy lemma gives us an upper bound for $k_F(G)$.

LEMMA 4.4. *If F is a graph on f vertices with no isolated vertices, then, for every graph G on e edges, the number of copies of F in G is at most e^{f-1}.*

PROOF. We proceed by induction. The base case $f = 2$ is trivial. Assume that $f > 2$ and consider two cases. If F is a matching on $f = 2l$ vertices, then the result follows easily: the number of copies of F in G is at most $\binom{e}{l} \leq e^l \leq e^{2l-1} = e^{f-1}$. In the other

https://doi.org/10.1017/S0004972724000972 Published online by Cambridge University Press

case, when F is not a matching, there exists a vertex $v \in F$ such that $F' = F - v$ has no isolated vertices. By the induction hypothesis, there are at most e^{f-2} copies of F' in G. Each copy of F' in G can be extended to at most e copies of F in G, since, by assumption, v must be adjacent to some vertex of F'. Therefore, there are at most e^{f-1} copies of F in G. □

LEMMA 4.5. *Let G be a graph on n vertices with $e = O(n^{1/2})$ edges. Then*

$$f_H(G) = c_H e(G)\binom{n-2}{h-2} - c_H(P_2)k_{P_2}(G)\binom{n-3}{h-3} + c_H(K_3)k_{K_3}(G)\binom{n-3}{h-3} + o(n^{h-2}).$$

PROOF. We consider the term $c_H(F)k_F(G)\binom{n-|F|}{h-|F|}$, where F is a graph on at least four vertices. From Lemma 4.4, $k_F(G) \le e^{|F|-1}$ and so $k_F(G)\binom{n-|F|}{h-|F|} \le e^{|F|-1}n^{h-|F|}$. Hence, under the assumption that $e = O(n^{1/2})$,

$$k_F(G)\binom{n-|F|}{h-|F|} = O(n^{|F|/2-1/2}n^{h-|F|}) = O(n^{h-1/2-|F|/2}) = O(n^{h-5/2}) = o(n^{h-2}).$$

As there are only three graphs on fewer than four vertices with no isolated vertices, namely, K_2, K_3 and P_2, and the number of terms in the summation in $f(G)$ depends only on H, we can write

$$f_H(G) = c_H e(G)\binom{n-2}{h-2} - c_H(P_2)k_{P_2}(G)\binom{n-3}{h-3} + c_H(K_3)k_{K_3}(G)\binom{n-3}{h-3} + o(n^{h-2}). \quad \square$$

The next lemma, which we use to prove that there are gaps in $S_H^{(n)}$, tells us that, for sufficiently large n, stars and matchings are asymptotically extremal examples of graphs for $f_H(G)$: that is, for a graph G on t edges,

$$a_H^{(n)}(t) - o(n^{h-2}) = f_H(S_t^{(n)}) - o(n^{h-2}) \le f_H(G) \le f_H(M_t^{(n)}) + o(n^{h-2}) = b_H^{(n)}(t) + o(n^{h-2}).$$

LEMMA 4.6. *Let G be a graph on n vertices with $e = O(n^{1/2})$ edges. As $n \to \infty$:*

(i) $f_H(G) \ge a_H^{(n)}(e) - o(n^{h-2}) = c_H e\binom{n-2}{h-2} - c_H(P_2)\binom{e}{2}\binom{n-3}{h-3} + o(n^{h-2})$; *and*

(ii) $f_H(G) \le b_H^{(n)}(e) + o(n^{h-2}) = c_H e\binom{n-2}{h-2} + o(n^{h-2})$.

PROOF. This an immediate corollary of Lemma 4.5. Observe that $k_{K_3}(S_e^{(n)}) = 0$, as stars contain no triangles and $k_{P_2}(S_e^{(n)}) = \binom{e}{2}$. Therefore, by Lemma 4.5,

$$a_H^{(n)}(e) = c_H e\binom{n-2}{h-2} - c_H(P_2)\binom{e}{2}\binom{n-3}{h-3} + o(n^{h-2}).$$

On the other hand, matchings contain no copies of K_3 nor P_2, and hence, again, by Lemma 4.5,

$$b_H^{(n)}(e) = c_H e\binom{n-2}{h-2} + o(n^{h-2}).$$

https://doi.org/10.1017/S0004972724000972 Published online by Cambridge University Press

For any graph G on e edges, we have $k_{P_2}(G) \le \binom{e}{2}$, so it follows from the above estimate on $a_H^{(n)}(e)$ and from Lemma 4.5 that

$$f_H(G) \ge c_H e\binom{n-2}{h-2} - c_H(P_2)\binom{e}{2}\binom{n-3}{h-3} + o(n^{h-2}) = a_H^{(n)}(e) - o(n^{h-2}).$$

On the other hand, it is easy to see that, for any graph G, we have $c_H(P_2)k_{P_2}(G) \ge c_H(K_3)k_{K_3}(G)$. So, by Lemma 4.5,

$$f_H(G) \le c_H e\binom{n-2}{h-2} + o(n^{h-2}) = b_H^{(n)}(e) + o(n^{h-2}). \qquad \square$$

PROOF OF THEOREM 1.2. Let $t_{\max} = \max\{t : b_H^{(n)}(t) < a_H^{(n)}(t+1)\}$. First, we show that $t_{\max} = \Theta(\sqrt{n})$. Indeed,

$$\begin{aligned} a_H^{(n)}(t+1) - b_H^{(n)}(t) &= c_H\binom{n-2}{h-2} - c_H(P_2)\binom{t+1}{2}\binom{n-3}{h-3} + o(n^{h-2}) \\ &= c_1 n^{h-2} - c_2 t^2 n^{h-3} + o(n^{h-2}), \end{aligned}$$

for some constants $c_1, c_2 > 0$ depending only on H. It follows that there is a constant C, depending only on H, such that, for all sufficiently large n, we have $a_H^{(n)}(t+1) - b_H^{(n)}(t) \ge 0$ if $t < C\sqrt{n}$ and $a_H^{(n)}(t+1) - b_H^{(n)}(t) \le 0$ otherwise. From Lemma 4.6, for all sufficiently large n,

$$f_H(G) \in [a_H^{(n)}(e) - o(n^{h-2}), b_H^{(n)}(e) + o(n^{h-2})]$$

for a graph G with $e \le t_{\max}$ edges. For a graph G with more than $t_{\max}$ edges, let G' be a graph obtained from G by deleting some edges until there are exactly $t_{\max}$ edges left. By monotonicity of f, we have $f_H(G) \ge f_H(G') \ge a_H^{(n)}(t_{\max}) - o(n^{h-2})$. Therefore, for any $t < t_{\max}$, the number of integers m in the interval $(b_H^{(n)}(t), a_H^{(n)}(t+1))$ such that there is a graph G on n vertices with $f_H(G) = m$ is at most $o(n^{h-2})$. It follows that $|(k_H(K_n) - a_H^{(n)}(t+1), k_H(K_n) - b_H^{(n)}(t)) \cap S_H^{(n)}| = o(n^{h-2})$ for every t. $\square$

Note that, when $t < \frac{1}{2}t_{\max}$, the gap between $a_H^{(n)}(t+1)$ and $b_H^{(n)}(t)$ is of order n^{h-2}. Hence, as n goes to infinity,

$$\frac{|(k_H(K_n) - a_H^{(n)}(t+1), k_H(K_n) - b_H^{(n)}(t)) \cap S_H^{(n)}|}{a_H^{(n)}(t+1) - b_H^{(n)}(t)} = o(1).$$

5. Open problems

We conclude our paper with some open problems that we feel would merit further study. Let $\phi_H^{(n)} = \min\{m \ge 0 : m \notin S_H^{(n)}\}$ be the smallest nonmember of $S_H^{(n)}$. Therefore, we have proved that $\phi_{K_3}^{(n)} = \binom{n}{3} - (\sqrt{2} + o(1))n^{3/2}$ and $\phi_{P_2}^{(n)} = 3\binom{n}{3} - (4 + o(1))n^{3/2}$.

The ultimate goal is to determine $S_H^{(n)}$. In particular, we ask the following question.

https://doi.org/10.1017/S0004972724000972 Published online by Cambridge University Press

PROBLEM 5.1. What is the asymptotic behaviour of $\binom{n}{r} - \phi_{K_r}^{(n)}$?

We showed, in Lemma 2.3, that if $e(G) = t \le t_{\max}$, then $f_H(G) \in [a_H^{(n)}(t), b_H^{(n)}(t)]$. Simple constructions readily show that, in general, not every number in the interval $[a_H^{(n)}(t), b_H^{(n)}(t)]$ is attained by $f_H(G)$. Therefore, we pose the following problem.

PROBLEM 5.2. Determine the set $[a_H^{(n)}(t), b_H^{(n)}(t)] - \{f_H(G) : e(G) = t\}$ for $t \le t_{\max}$.

Acknowledgement

We thank Tomas Juškevičius for useful discussions.

References

[1] N. Alon, 'On the number of subgraphs of prescribed type of graphs with a given number of edges', *Israel J. Math.* **38**(1–2) (1981), 116–130.
[2] P. Balister and A. Dogan, 'On the edge spectrum of saturated graphs for paths and stars', *J. Graph Theory* **89**(4) (2018), 364–385.
[3] J. Cutler and A. J. Radcliffe, 'The maximum number of complete subgraphs in a graph with given maximum degree', *J. Combin. Theory Ser. B* **104** (2014), 60–71.
[4] P. Erdős, 'On the number of complete subgraphs contained in certain graphs', *Publ. Math. Inst. Hung. Acad. Sci., Ser. A* **7** (1962), 459–464.
[5] P. Erdős, 'On a theorem of Rademacher–Turán', *Illinois J. Math.* **6**(1) (1962), 122–127.
[6] P. Erdős, 'On the number of triangles contained in certain graphs', *Canad. Math. Bull.* **7** (1964), 53–56.
[7] P. Erdős, 'On the number of complete subgraphs and circuits contained in graphs', *Cas. Pestováni Mat.* **94** (1969), 290–296.
[8] Z. Füredi, 'Graphs with maximum number of star-forests', *Studia Sci. Math. Hungar.* **27** (1992), 403–407.
[9] T. Kittipassorn and G. Mészáros, 'Frustrated triangles', *Discrete Math.* **338**(12) (2015), 2363–2373.
[10] D. G. Larman, 'On the number of complete subgraphs and circuits in a graph', *Proc. Roy. Soc. Edinburgh Sect. A* **308**(1494) (1969), 327–342.
[11] L. Lovász and M. Simonovits, 'On the number of complete subgraphs of a graph II', in: *Studies in Pure Mathematics* (eds. P. Erdős, L. Alpár, G. Halász and A. Sárközy) (Springer, Basel, 1983), 459–495; to the memory of Paul Turán.
[12] V. Nikiforov, 'The number of cliques in graphs of given order and size', *Trans. Amer. Math. Soc.* **363** (2011), 1599–1618.
[13] E. A. Nordhaus and B. M Stewart, 'Triangles in an ordinary graph', *Canad. J. Math.* **15** (1963), 33–41.

TEERADEJ KITTIPASSORN, Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand
e-mail: teeradej.k@chula.ac.th

KAMIL POPIELARZ, Zurich, Switzerland
e-mail: kamil.popielarz@gmail.com

https://doi.org/10.1017/S0004972724000972 Published online by Cambridge University Press