Peter Allen, Julia Böttcher, Jan Hladký, Dian Piguet

A density Corrádi-Hajnal Theorem

Article (Accepted version)
(Refereed)

Original citation:
Allen, Peter, Böttcher, Julia, Hladký, Jan and Piguet, Diana (2014) *A density Corrádi-Hajnal theorem.* Canadian Journal of Mathematics, 67 (4). pp. 721-758. ISSN 0008-414X

DOI: 10.4153/CJM-2014-030-6

A DENSITY CORRÁDI–HAJNAL THEOREM

PETER ALLEN*, JULIA BÖTTCHER*, JAN HLADKÝ†, AND DIANA PIGUET‡

Abstract. We find, for all sufficiently large n and each k, the maximum number of edges in an n-vertex graph which does not contain $k+1$ vertex-disjoint triangles.

This extends a result of Moon [Canad. J. Math. 20 (1968), 96–102] which is in turn an extension of Mantel's Theorem. Our result can also be viewed as a density version of the Corrádi–Hajnal Theorem.

1. Introduction

A classic result of Mantel asserts that each n-vertex graph G with more than $\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil$ edges contains a triangle. What can we say about the number of triangles in a graph with more than $\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil$ edges?

There are three natural interpretations of this question. We can ask how many *vertex-disjoint* triangles are guaranteed, how many *edge-disjoint* triangles are guaranteed, or simply how many triangles are guaranteed in *total*. The answer to each of the first two questions is 1 (which is trivial) and Rademacher proved (see [Erd62a]) that the answer to the last is $\lfloor \frac{n}{2} \rfloor$; in each case the extremal example consists of a complete balanced bipartite graph with one edge added to the larger part. It is then natural to ask the same questions of n-vertex graphs G with at least $\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil + m$ edges, for any $m \geq 1$.

These questions are much harder. Lovász and Simonovits [LS83] gave a conjectured lower bound on the number of triangles present in any n-vertex

---

2010 *Mathematics Subject Classification.* 05C35 (primary).

* Department of Mathematics, London School of Economics, Houghton Street, London, WC2A 2AE, UK. *E-mail*: `p.d.allen|j.boettcher@lse.ac.uk`.

† DIMAP and Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK. The author is an EPSRC Research Fellow. *E-mail*: `honzahladky@gmail.com`.

‡ School of Mathematics, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK *E-mail*: `D.Piguet@bham.ac.uk`.

PA, JH, and DP were supported by DIMAP, EPSRC award EP/D063191/1. PA was partially supported by FAPESP (Proc. 2010/09555-7), and JB by FAPESP (Proc. 2009/17831-7). PA and JB are grateful to NUMEC/USP, Núcleo de Modelagem Estocástica e Complexidade of the University of São Paulo, for supporting this research. The research leading to this result has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. PIEF-GA-2009-253925. An extended abstract of this paper appeared in the proceedings of the Eurocomb 2011 conference. Support by the Institut Mittag-Leffler (Djursholm, Sweden) is gratefully acknowledged.

graph G with at least $\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil + m$ edges, which Erdős [Erd62a] had already proved correct for m small enough compared to n. The conjecture remains open, although a celebrated recent result of Razborov [Raz08], using his method of flag algebras, is that the conjectured lower bound—a complicated continuous but only piecewise differentiable function in m—is asymptotically correct for all m. The number of edge-disjoint triangles was studied by Győri [Győ91], but only for $m \le 2n - 10$ were exact results proved, and for large m it is not clear what the right answer should be.

In this paper we solve (for sufficiently large n) the problem of how many vertex-disjoint triangles are guaranteed to exist in an n-vertex graph G with a given number of edges. It is convenient to rephrase the problem in the following way.

Problem 1. *How many edges can an n-vertex graph G possess if it does not contain $k+1$ vertex-disjoint triangles?*

This problem was first studied by Erdős [Erd62b] and by Moon [Moo68]. The former proved the exact result when $n \ge 400k^2$, and the latter when $n \ge 9k/2 + 4$, giving the following theorem.

Theorem 2 (Moon [Moo68])**.** *Suppose that $n \ge 9k/2 + 4$. Let G be an n-vertex graph which does not contain $k+1$ vertex-disjoint triangles. Then*

$$e(G) \le \binom{k}{2} + k(n-k) + \left\lceil \frac{n-k}{2} \right\rceil \left\lfloor \frac{n-k}{2} \right\rfloor .$$

We give an exact solution to Problem 1 for all values of k when n is greater than an absolute constant n_0. Our main result, Theorem 6, states that the answer is given by four different extremal (families of) graphs in four different regimes of k.

Interestingly, although Moon states that his result "almost certainly remains valid for somewhat smaller values of n also", in fact he almost reaches a natural barrier: the graph which Moon proved to be extremal (the graph $E_1(n,k)$ in Definition 5) is only extremal when $n \ge 9k/2 + 3$.

We remark that our result can also be seen as a variation of two other classical theorems in extremal graph theory. Firstly, Erdős and Gallai [EG59] answered the analogous question for edges instead of triangles.

Theorem 3 (Erdős and Gallai [EG59])**.** *For any n-vertex graph G without $k+1$ vertex-disjoint edges, $e(G) \le \max\{k(n-k) + \binom{k}{2}, \binom{2k+1}{2}\}$.*

In fact, they showed that, depending on k, the extremal graph for this problem either consists of k vertices which are complete to all vertices, or of a $(2k+1)$-clique and a disjoint independent set. In this sense the appearance of various very different extremal structures in our result is not surprising.

Secondly, Corrádi and Hajnal [CH63] considered the variant of Problem 1 where the number of edges is replaced by the minimum degree and proved the following well-known theorem.

Theorem 4 (Corrádi and Hajnal [CH63]). *For any n-vertex graph G which does not contain $k+1$ vertex disjoint triangles, $\delta(G) \leq k + \left\lfloor \frac{n-k}{2} \right\rfloor$.*

In other words, our result is the 'density version' of the Corrádi-Hajnal Theorem.

1.1. **Organisation of the paper.** We state and discuss our main result, Theorem 6, in Section 2. We outline its proof in Section 3. The main combinatorial work of the proof is to be found in Sections 4 and 5. In Section 6 we show how to deduce Theorem 6 from these combinatorial arguments and some maximisation problems. In Section 7 we prove an auxiliary lemma which is one of the key points of the proof of Theorem 6, building on our previous work [ABHP]. In Section 8 we then discuss possibilities of extending our result. Our proof of Theorem 6 requires tedious maximisation arguments, which we state as they are needed but whose derivations are postponed to Appendix A.

The proof relies on a number of elementary but lengthy calculations. After performing these calculations by hand, we verified some using the computer algebra software Maxima. The output pdf file as well as all the data in the wxMaxima format are available as ancillary files on the arXiv.

2. Our result

Given an integer ℓ and a graph H, we write $\ell \times H$ to denote the disjoint union of ℓ copies of H. We say that a graph is *$\ell \times H$-free* if it does not contain ℓ vertex disjoint (not necessarily induced) copies of H. In Theorem 6 we determine the maximal number of edges in a $(k+1)\times K_3$-free graph on n vertices for every $0 \leq k < \frac{n}{3}$. The extremal formula is a somewhat opaque maximum of four different terms, so in preference to presenting it we shall describe four constructions of n-vertex $(k+1)\times K_3$-free graphs corresponding to these four terms. We say that an edge e (or more generally a set of vertices) *meets* a set of vertices X if e and X intersect. The edge e meets X in X' if $X' = X \cap e$.

Definition 5 (extremal graphs). *Let n and k be non-negative integers with $k \leq \frac{n}{3}$. We define the following four graphs (see also Figure 1).*[1]

$E_1(n,k)$: *Let $X \dot\cup Y_1 \dot\cup Y_2$ with $|X| = k$, $|Y_1| = \left\lceil \frac{n-k}{2} \right\rceil$, and $|Y_2| = \left\lfloor \frac{n-k}{2} \right\rfloor$ be the vertices of $E_1(n,k)$. Insert all edges intersecting X, and between Y_1 and Y_2.*

$E_2(n,k)$: *The second class of extremal graphs is defined only for $k < \frac{n-1}{4}$. Let $X \dot\cup Y_1 \dot\cup Y_2$ with $|X| = 2k+1$, $|Y_1| = \lfloor \frac{n}{2} \rfloor$, and $|Y_2| = \lceil \frac{n}{2} \rceil - 2k - 1$ (or $|Y_1| = \lceil \frac{n}{2} \rceil$, and $|Y_2| = \lfloor \frac{n}{2} \rfloor - 2k - 1$) be the vertices of $E_2(n,k)$. Insert all edges within X, and between Y_1 and $X \cup Y_2$. If n is odd, this construction captures two graphs, if n is even just one.*

[1]The constructions for $E_2(n,k)$ and $E_4(n,k)$ do not give unique graphs. We collectively denote all graphs constructed in this way by $E_2(n,k)$ and $E_4(n,k)$, respectively. In the following we only use properties of these graphs that are shared by all of them.

$E_3(n,k)$: *Let* $X\dot{\cup}Y_1$ *with* $|X| = 2k+1$ *and* $|Y_1| = n-2k-1$ *be the vertices of* $E_3(n,k)$. *Insert all edges intersecting* X.

$E_4(n,k)$: *The fourth class of extremal graphs is defined only for* $k \geq \frac{n}{6} - 2$. *When* $k \geq \frac{n-2}{3}$ *take* $E_4(n,k)$ *to be the complete graph* K_n. *Otherwise, the vertex set is formed by five disjoint sets* X, Y_1, Y_2, Y_3, *and* Y_4, *with* $|Y_1| = |Y_3|$, $|Y_2| = |Y_4|$, $|Y_1| + |Y_2| = n-3k-2$, *and* $|X| = 6k-n+4$. *Insert all edges in* X, *between* X *and* $Y_1 \cup Y_2$, *and between* $Y_1 \cup Y_4$ *and* $Y_2 \cup Y_3$. *Thus the choice of* $|Y_1|$ *determines a particular graph in the class* $E_4(n,k)$. *All graphs in* $E_4(n,k)$ *have the same number of edges.*

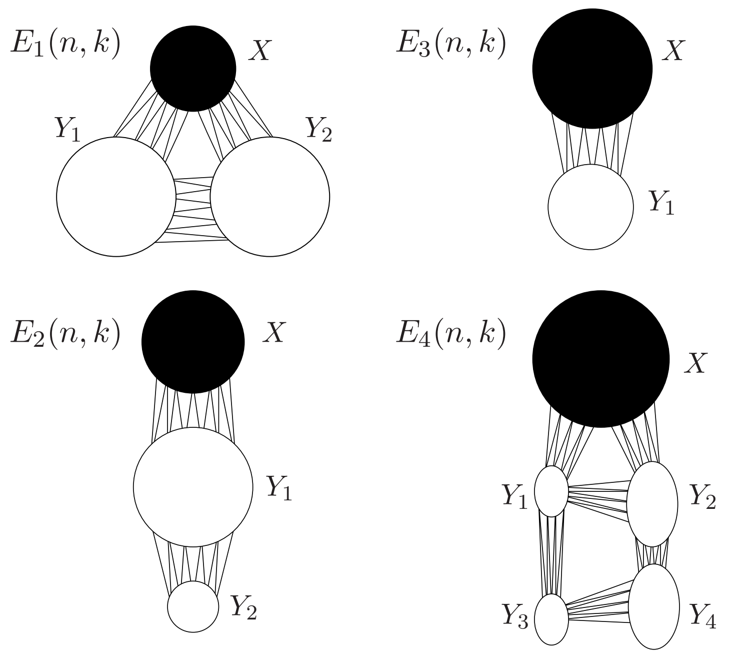

FIGURE 1. The extremal graphs.

Our main result is the following.

Theorem 6. *There exists* n_0 *such that for each* $n > n_0$ *and each* k, $0 \leq k \leq \frac{n}{3}$ *we have the following. Let* G *be a* $(k+1) \times K_3$*-free graph on* n *vertices. Then*

$$e(G) \leq \max_{j \in [4]} e\big(E_j(n,k)\big) \,. \tag{1}$$

Clearly, the graphs $E_i(n,k)$ are edge-maximal subject to not containing $(k+1) \times K_3$. The only exception is $E_4(n,k)$ for $k \lesssim \frac{n}{4}$—in fact, in the range $0 \leq k \lesssim \frac{n}{4}$ the number of edge-disjoint triangles in $E_4(n,k)$ depends on the choice of $|Y_1|$. However $E_4(n,k)$ is in any case not the extremal graph in this range; see the discussion below and Table 1. The graphs $E_i(n,k)$ have the following numbers of edges (after an exact formula we identify the leading terms; to this end we use the symbol $\approx$).

$$
\begin{aligned}
e\big(E_1(n,k)\big) &= \binom{k}{2} + k(n-k) + \left\lceil \frac{n-k}{2} \right\rceil \left\lfloor \frac{n-k}{2} \right\rfloor \approx \tfrac{1}{4}n^2 - \tfrac{1}{4}k^2 + \tfrac{1}{2}kn\ , \\
e\big(E_2(n,k)\big) &= \binom{2k+1}{2} + \left\lceil \frac{n}{2} \right\rceil \left\lfloor \frac{n}{2} \right\rfloor \approx \tfrac{1}{4}n^2 + 2k^2\ , \\
e\big(E_3(n,k)\big) &= \binom{2k+1}{2} + (2k+1)(n-2k-1) \approx 2kn - 2k^2\ , \\
e\big(E_4(n,k)\big) &= \binom{6k-n+4}{2} + (6k-n+4)(n-3k-2) + (n-3k-2)^2 \\
&\approx \frac{n^2}{2} - 3kn + 9k^2\ .
\end{aligned}
\tag{2}
$$

Comparing these edge numbers reveals that, as k grows from 0 to $n/3$, the extremal graphs dominate in the following order (for n sufficiently large). In the beginning $E_1(n,k)$ has the most edges of these four graphs until $k \approx \frac{2n}{9}$, where it is surpassed by $E_2(n,k)$. At $k \approx \frac{n}{4}$ this extremal structure ceases to exist and is replaced by $E_3(n,k)$, until finally at $k \approx (5+\sqrt{3})n/22$ the graph $E_4(n,k)$ takes over. The exact thresholds are listed in Table 1. Further, the edge numbers of the graphs $E_i(n,k)$ are plotted in Figure 2.

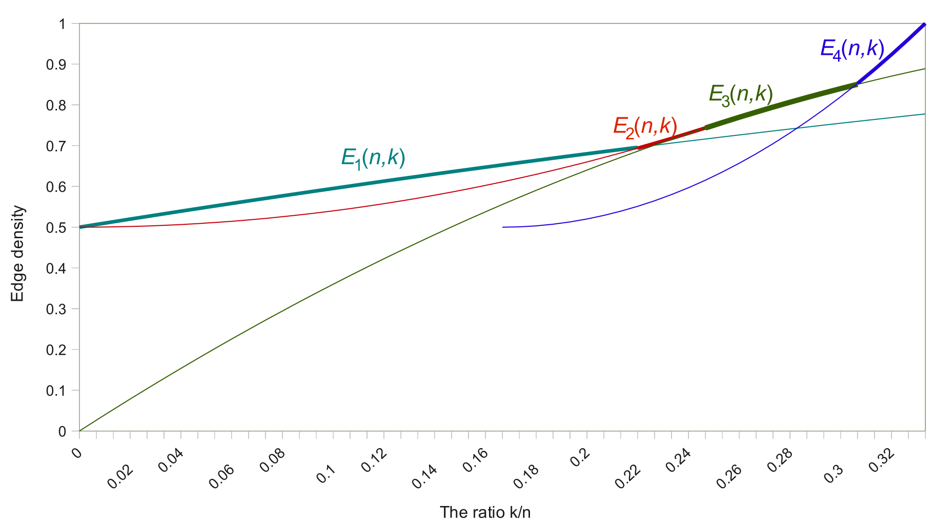

FIGURE 2. Edge densities of the graphs $E_i(n,k)$ where k ranges from 0 to $\frac{n}{3}$.

Observe that for fixed n, as k increases, the transitions of the extremal graphs from $E_1(n,k)$ to $E_2(n,k)$ and from $E_3(n,k)$ to $E_4(n,k)$ are not continuous: $\Theta(n^2)$ edges must be edited to change from the former to the latter structure. The transition from $E_2(n,k)$ to $E_3(n,k)$ however is continuous.

graph	**extremal for**
$E_1(n,k)$	$1 \le k \le \frac{2n-6}{9}$
$E_2(n,k)$	$\frac{2n-6}{9} \le k \le \frac{n-1}{4}$
$E_3(n,k)$	$\frac{n-1}{4} \le k \le \frac{5n-12+\sqrt{3n^2-10n+12}}{22}$
$E_4(n,k)$	$\frac{5n-12+\sqrt{3n^2-10n+12}}{22} \le k \le \frac{n}{3}$

Table 1. Transitions between the extremal graphs.

3. Proof outline and setup

The basic idea of our proof is straightforward: we show that we can partition the vertices of any $(k+1) \times K_3$-free graph into six parts, and establish some upper bounds on the numbers of edges within and between these parts in terms of their sizes only. This defines a function (of six variables) which is an upper bound on the number of edges of a graph with parts of the given sizes. Then maximising this function (subject to n and k being fixed) we obtain an upper bound on the number of edges of a $(k+1) \times K_3$-free graph with n vertices, and observe that this matches the lower bounds provided by the extremal structures given in Definition 5.

We shall now fix the basic setup for our proof, i.e., we will specify the above mentioned six parts, which will be called $\mathcal{T}_1$, $\mathcal{T}_2$, $\mathcal{T}_3$, $\mathcal{T}_4$, $\mathcal{M}$, and $\mathcal{I}$. We need the following definition. Let G be a graph, uv be an edge in G and xyz a triangle in G. We say that uv *sees* vertex x of xyz if uvx is a triangle in G. The edge uv sees xyz if uv sees one of the vertices x, y, or z. Similarly, we say that a vertex u sees (the edge xy of) the triangle xyz if uxy is a triangle in G.

Throughout we will assume the following setup.

Setup 7. *Let G be an n-vertex graph which is edge-maximal subject to not containing $(k+1) \times K_3$. Let $\mathcal{T}$ be a set of k vertex-disjoint triangles in G, let $\mathcal{M}$ be a maximum matching outside $\mathcal{T}$, and presume $\mathcal{T}$ is chosen to maximise the size of $\mathcal{M}$. The remaining vertices of G, which form an independent set, we call $\mathcal{I}$.*

We now split the set $\mathcal{T}$ into four parts as follows, forming together with $\mathcal{M}$ and $\mathcal{I}$ the six above-mentioned parts of G. Let $\mathcal{T}_1$ be the set of triangles in $\mathcal{T}$ seen by at least two $\mathcal{M}$-edges. Let $\mathcal{T}_2$ be the set of triangles in $\mathcal{T} - \mathcal{T}_1$ seen by either an $\mathcal{M}$-edge and at least one $\mathcal{I}$-vertex or by two $\mathcal{I}$-vertices. Finally, we aim to partition the remaining triangles of $\mathcal{T}$ into a 'sparse part' $\mathcal{T}_3$ and a 'dense part' $\mathcal{T}_4$ by applying the following algorithm. We start with D equal to the set of all triangles in $\mathcal{T} - (\mathcal{T}_1 \cup \mathcal{T}_2)$, and $S = \emptyset$. If there is a triangle in D which sends at most $8(|D| - 1)$ edges to the other triangles in D, we move it to S. We repeat until D contains no more such triangles. We then set $\mathcal{T}_3 := S$, and $\mathcal{T}_4 := D$. Note that every triangle in $\mathcal{T}_4$ sends more than $8(|\mathcal{T}_4| - 1)$ edges to the other triangles in $\mathcal{T}_4$.

We define $m := |\mathcal{M}|$, $i := |\mathcal{I}|$, and $t_j := |\mathcal{T}_j|$ for all $j \in [4]$.

We remark that the outcome of the algorithm for constructing $\mathcal{T}_3$ and $\mathcal{T}_4$ is not uniquely determined. However, any possible pair $\mathcal{T}_3$ and $\mathcal{T}_4$ resulting from the construction we described is suitable for our purposes.

Further, we emphasise that $k = |\mathcal{T}|$ is the number of triangles in $\mathcal{T}$, which cover $3k$ vertices (and similarly $\mathcal{M}$ covers $2m$ vertices). The function $e(\bullet)$ counts the number of edges in G induced by the structure $\bullet$, e.g., $e(\mathcal{T}_3) = e(G[V(\mathcal{T}_3)])$. Similarly, $e(\bullet, \star)$ counts edges in the bipartite graph between the structures $\bullet$ and $\star$.

Before we proceed, let us give some motivation for the above defined partition of G by applying it to our four extremal graphs from Definition 5. First consider the graph $E_1(n,k)$. It is easy to check that for this graph we have $\mathcal{T} = \mathcal{T}_1$, and all vertices (except perhaps one) outside $\mathcal{T}$ are in $\mathcal{M}$. Any pair of triangles of $\mathcal{T}$ has seven edges between them in $E_1(n,k)$, the set $\mathcal{M}$ induces m^2 edges, and $e(\mathcal{M}, \mathcal{T}_1) = 4mt_1$. We shall show in our proof that *in any graph* G, the definition of $\mathcal{T}_1$ forces that any two triangles of $\mathcal{T}_1$ have at most seven edges between them (see Lemma 10(f)), the set $\mathcal{M}$ induces at most m^2 edges (see Lemma 10(c)), and $e(\mathcal{M}, \mathcal{T}_1) \leq 4mt_1$ (see Lemma 10(d)). Together with bounds which we will prove on the number of edges touching $\mathcal{I}$, we conclude that if $\mathcal{T} = \mathcal{T}_1$ then $e(G) \leq e\big(E_1(n,k)\big)$.

Similarly, the definition of $\mathcal{T}_2$ and $\mathcal{T}_4$ is motivated by the fact that in both $E_2(n,k)$ and $E_3(n,k)$ we have $\mathcal{T} = \mathcal{T}_2$, while in $E_4(n,k)$ we have $\mathcal{T} = \mathcal{T}_4$. (The set $\mathcal{T}_3$ is always empty in the extremal graphs.) It turns out that, for $E_2(n,k)$ and $E_3(n,k)$ we will be able to use a similar strategy as lined out for $E_1(n,k)$, i.e., we shall infer from the definition of $\mathcal{T}_2$ that $E_2(n,k)$ and $E_3(n,k)$ have a maximal number of edges in $\mathcal{T}_2$ (see Lemma 10(h)) and then show that $\mathcal{T} = \mathcal{T}_2$ in an extremal graph (for the appropriate range of k). For $E_4(n,k)$ we must work harder: the definition of $\mathcal{T}_4$ permits nine edges to exist between a pair of triangles, yet in $E_4(n,k)$ only some pairs of triangles actually have nine edges between them.

As explained, our main goal in the following will be to establish bounds on the number of edges within and between the six parts of G. One concept that will turn out to be very fruitful in this context is that of a rotation.

Definition 8 (rotation)**.** *Let G' be a graph and let $\mathcal{T}'$ be a triangle factor in G'. An* improving rotation *on a set V' is a set of vertex disjoint triangles $\tilde{\mathcal{T}}$ in V' which witnesses either that $\mathcal{T}'$ is not of maximum size, or that its choice does not maximise the matching number of $G' - V(\mathcal{T}')$: We can replace those triangles of $\mathcal{T}'$ which are contained in V' by the triangles $\tilde{\mathcal{T}}$ and obtain a triangle factor $\mathcal{T}''$ with one of the following two properties. Either $|\mathcal{T}''| > |\mathcal{T}'|$, or $|\mathcal{T}''| = |\mathcal{T}'|$ but the matching number of $G' - V(\mathcal{T}'')$ is bigger than that of $G' - V(\mathcal{T}')$. If, on the other hand, $|\mathcal{T}''| = |\mathcal{T}'|$ and the matching number of $G' - V(\mathcal{T}'')$ equals that of $G' - V(\mathcal{T}')$ then V' is a* non-improving rotation *or simply rotation. In both cases we also say that we can* rotate *from $\mathcal{T}'$ to $\mathcal{T}''$.*

Typically, the rotations that we will consider are local structures. To give an example, let G and $\mathcal{T}$ be as in Setup 7. By definition, there are no improving rotations in G. Suppose, however, that we find outside $\mathcal{T}$ two vertex-disjoint edges uv and $u'v'$, and a triangle xyz of $\mathcal{T}$ with the property that x is a common neighbour of uv, and y of $u'v'$. This structure allows us to rotate by replacing xyz with uvx and $u'v'y$, a contradiction. The non-existence of this structure leads to an upper bound on the number of edges between $\mathcal{M}$ and $\mathcal{T}$.

4. Small rotations

In this section we will describe several rotations involving small numbers (one or two) of triangles, and show that their non-existence gives good bounds on the maximum number of edges within and between $\mathcal{T}_1$, $\mathcal{T}_2$, $\mathcal{T}_3$, $\mathcal{M}$ and $\mathcal{I}$. The bounds obtained on edges involving $\mathcal{T}_4$ are not strong enough for the proof of Theorem 6, but they are strong enough to prove the following lemma, which serves both as an illustration of our technique and as a necessary step in the proof of Theorem 6.

Lemma 9. *Let $k \le \frac{n-8}{5}$ be an integer and let G be a $(k+1)\times K_3$-free graph on n vertices. Then $e(G) \le e\big(E_1(n,k)\big)$.*

Observe that, in contrast to Theorem 6 we do not require any lower bound on n in this lemma. Observe also that since $\frac{n-8}{5} < \frac{2n-8}{9}$, the result follows from Theorem 2: but its proof will exemplify our techniques and put us into position to explain the remaining steps to obtain Theorem 6.

We assume in the following Setup 7. We start with some simple upper bounds.

Lemma 10. *The following bounds hold.*

(*a*) $e(\mathcal{I}) = 0$.
(*b*) $e(\mathcal{I}, \mathcal{M}) \le im$.
(*c*) $e(\mathcal{M}) \le m^2$.
(*d*) $e(\mathcal{M}, \mathcal{T}_1) \le 4mt_1$.
(*e*) $e(\mathcal{I}, \mathcal{T}_1) \le 2it_1$.

(*f*) $e(\mathcal{T}_1) \le 7\binom{t_1}{2} + 3t_1$.
(*g*) $e(\mathcal{I}, \mathcal{T}_2) \le 2it_2$.
(*h*) $e(\mathcal{T}_2) \le 8\binom{t_2}{2} + 3t_2$.
(*i*) $e(\mathcal{T}_3) + e(\mathcal{T}_3, \mathcal{T}_4) \le 8\binom{t_3}{2} + 8t_3t_4 + 3t_3$.

Proof. We leave to the reader the proof of (*a*).

Suppose that a vertex $u \in \mathcal{I}$ sends more than m edges to $\mathcal{M}$. Then there is some edge vw of $\mathcal{M}$ which receives two edges from u. So uvw is a triangle of G, contradicting maximality of $|\mathcal{T}|$. Summing over vertices of $\mathcal{I}$, bound (*b*) follows. Similarly, if a vertex of $\mathcal{M}$ was adjacent to more than m other vertices of $\mathcal{M}$ this would contradict maximality of $\mathcal{T}$. Bound (*c*) follows.

If an edge uv of $\mathcal{M}$ sends more than four edges to any triangle T of $\mathcal{T}_1$, then it must see two vertices of T. Since by definition of $\mathcal{T}_1$ there is another edge $u'v'$ of $\mathcal{M}$ which sees a vertex of T, there are two vertices x, x' of T such that uvx and $u'v'x'$ are triangles of G. This is an improving rotation which contradicts the maximality of $\mathcal{T}$. Therefore, no such edge exists. Bound (*d*) follows by summation. Similarly, if a vertex u of $\mathcal{I}$ were to send three edges to a triangle T of $\mathcal{T}_1$, then (using an edge of $\mathcal{M}$ which sees T) we would have an improving rotation increasing the size of $\mathcal{T}$. Bound (*e*) follows.

Now suppose there were two triangles uvw and $u'v'w'$ of $\mathcal{T}_1$ with more than seven edges between them. By definition of $\mathcal{T}_1$ we can find disjoint edges xy and $x'y'$ of $\mathcal{M}$ such that xy sees u and $x'y'$ sees u'. Because there are at least eight edges between uvw and $u'v'w'$, there must be at least three edges between vw and $v'w'$. In particular, there is a triangle contained in $\{v, w, v', w'\}$. Together with xyu and $x'y'u'$ this is an improving rotation increasing $\mathcal{T}$, contradicting the maximality of $|\mathcal{T}|$. This implies bound (*f*).

Next, suppose there is a vertex u of $\mathcal{I}$ which sends three edges to a triangle xyz of $\mathcal{T}_2$. We utilise the definition of $\mathcal{T}_2$ and infer that one of the two cases must occur. Either there is a second vertex u' of $\mathcal{I}$ which sees two vertices $\{x, y\}$ of that triangle. Hence we can rotate and replace the triangle xyz and the vertices u and u' by the triangle xyu' and the edge uz, a contradiction. The other case when xyz is seen by an edge of $\mathcal{M}$ can be treated similarly. It follows that no vertex of $\mathcal{I}$ sends three edges to any triangle of $\mathcal{T}_2$, hence bound (*g*).

We now turn to proving (*h*). Suppose that there is a pair of triangles xyz and $x'y'z'$ of $\mathcal{T}_2$ forming a copy of K_6. By the definition of $\mathcal{T}_2$ we either have that there are distinct vertices $u, u' \in \mathcal{I}$ which see respectively xy and $x'y'$, or that there is a vertex $u \in I$ which sees xy and an edge $ab \in \mathcal{M}$ disjoint from u which is seen by x'. Suppose the former case. Then we have a similar improving rotation as above: we form xyu, $x'y'u'$, and zz', a contradiction. An analogous improving rotation exists in the other case. This yields our bound (*h*).

Finally, we must show that $e(\mathcal{T}_3) + e(\mathcal{T}_3, \mathcal{T}_4) \le 8\binom{t_3}{2} + 8t_3t_4 + 3t_3$. This bound does not come from a rotation. Instead, recall that $\mathcal{T}_3$ is formed sequentially. We claim that the bound applies to every pair of sets S and D during the construction in Setup 7, that is, that $e(S) + e(S, D) \leq 8\binom{|S|}{2} + 8|S||D| + 3|S|$. This is trivially true at the first stage, when $S = \emptyset$. Now a triangle is moved from D to S when it sends at most $8(|D| - 1|)$ edges to the rest of D. So $|S|$ is increased by one, and $e(S) + e(S, D)$ is increased by at most $3 + 8(|D| - 1)$. Bound (i) follows by induction. □

We next come to two bounds on edges within $\mathcal{T}$.

Lemma 11. *The following bounds hold.*

(j) *When* $t_1 \neq 1$ *and* $j \geq 2$*, then* $e(\mathcal{T}_1, \mathcal{T}_j) \leq 7t_1t_j$.
(k) *When* $t_2 \neq 1$ *and* $j \geq 3$*, then* $e(\mathcal{T}_2, \mathcal{T}_j) \leq 8t_2t_j$.

Proof. We first show (j). Since the case $t_1 = 0$ is trivial, we assume that $t_1 \geq 2$. Let xyz be a triangle in $\mathcal{T}_j$, for some $j \geq 2$, and suppose that there are at least $7t_1 + 1$ edges from $\mathcal{T}_1$ to xyz. Then certainly there is a triangle $uvw \in \mathcal{T}_1$ which sends at least eight edges to xyz. There are two possibilities.

First, suppose uvw sends exactly eight edges to xyz. Then there is another triangle $u'v'w' \in \mathcal{T}_1$ which sends at least seven edges to xyz. By definition of $\mathcal{T}_1$, there are distinct edges ab and $a'b'$ of $\mathcal{M}$ such that ab sees u and $a'b'$ sees u'. Since there are seven edges from $u'v'w'$ to xyz, $v'w'$ must have a common neighbour x; since there are eight edges from xyz to uvw, yz must have two common neighbours in uvw, and in particular one, say v, which is not u. Then replacing uvw, $u'v'w'$ and xyz with abu, $a'b'u'$, $v'w'x$ and yzv is an improving rotation, a contradiction.

Second, suppose uvw sends nine edge to xyz. Then there is another triangle $u'v'w'$ of $\mathcal{T}_1$ which sends at least six edges to xyz. Again we assume $ab \in \mathcal{M}$ sees u, and $a'b' \in \mathcal{M}$ sees u'. Now at least one of v' and w', say v', must have two neighbours in xyz, say x and y. Since xyz sends nine edges to uvw, zvw is a triangle. Then replacing uvw, $u'v'w'$ and xyz with abu, $a'b'u'$, $v'xy$ and zvw is an improving rotation, a contradiction. The bound (j) follows by summation.

We now show (k). Again, we assume $t_2 \geq 2$ and suppose $xyz \in \mathcal{T}_j$ for some $j \geq 3$ sends at least $8t_2 + 1$ edges to $\mathcal{T}_2$. Then there are triangles uvw and $u'v'w'$ of $\mathcal{T}_2$ to which xyz sends respectively nine and at least eight edges. We now use the fact that $uvw, u'v'w' \in \mathcal{T}_2$ to infer the following: either there are distinct vertices a and a' of $\mathcal{I}$ which see respectively uv and $u'v'$, or there is a vertex $a \in \mathcal{I}$ and an edge $bc \in \mathcal{M}$ such that a sees uv and u' sees bc. Let us consider the first case. Now w' is adjacent to at least two vertices of xyz, say x and y, and zw is an edge. Therefore replacing uvw, $u'v'w'$ and xyz by auv, $a'u'v'$ and $w'xy$ maintains the number of triangles of $\mathcal{T}$, but allows us to add zw to $\mathcal{M}$, and is thus an improving rotation, a contradiction. Next we consider the case when there is a vertex $a \in \mathcal{I}$ and an edge $bc \in \mathcal{M}$ such that a sees uv and u' sees bc. There is a vertex of xyz which sees $v'w'$, say x. Then replacing uvw, $u'v'w'$ and xyz by bcu',

$v'w'x$, yzw, and auv is an improving rotation, again a contradiction. The bound (k) follows by summation. □

Our next task is to bound the edges between $\mathcal{M}$ and $\mathcal{T}_j$, $j \geq 2$, and between $\mathcal{I}$ and $\mathcal{T}_j$, $j \geq 3$. We combine these bounds with those given in Lemma 11 because they permit us to handle the cases $t_1 = 1$ and $t_2 = 1$ which were not dealt with in Lemma 11. However, in the proof of Theorem 6 we will find that we require both sets of bounds.

Lemma 12. *The following bounds hold.*

(l)

$$e(\mathcal{T}_1, \mathcal{T}_2) + e(\mathcal{M}, \mathcal{T}_2) \leq \begin{cases} 7t_1t_2 + (2+3m)\,t_2 & \text{if } m \geq 1 \text{ and} \\ 0 & \text{if } m = 0\,. \end{cases}$$

(m) *When* $j = 3, 4$ *we have*

$$e(\mathcal{T}_1, \mathcal{T}_j) + e(\mathcal{M}, \mathcal{T}_j) \leq \begin{cases} 7t_1t_j + (3+3m)\,t_j & \text{if } m \geq 1 \text{ and} \\ 0 & \text{if } m = 0\,. \end{cases}$$

(n) *When* $j = 3, 4$ *we have*

$$e(\mathcal{T}_2, \mathcal{T}_j) + e(\mathcal{I}, \mathcal{T}_j) \leq \begin{cases} 8t_2t_j + (2+i)\,t_j & \text{if } i \geq 1 \text{ and} \\ 0 & \text{if } i = 0\,. \end{cases}$$

Proof. First we prove (l). Observe that if $m = 0$ then by definition of $\mathcal{T}_1$ we have also $t_1 = 0$, and the bound follows. Now by definition of $\mathcal{T}_2$, any triangle $xyz \in \mathcal{T}_2$ is seen by at most one edge ab in $\mathcal{M}$. It follows that all other edges of $\mathcal{M}$ send at most three edges to xyz. Furthermore, if ab sent six edges to xyz, then we would find an improving rotation as follows. Let $c \in \mathcal{I}$ be a vertex which sees (say) the edge xy in xyz, whose existence is guaranteed by definition of $\mathcal{T}_2$. Now cxy and abz are disjoint triangles which can replace xyz to increase the size of $\mathcal{T}$. It follows that xyz sends at most $5 + 3(m-1) = 3m + 2$ edges to $\mathcal{M}$.

If $t_1 \neq 1$, then summing over $\mathcal{T}_2$ together with the bound (j) of Lemma 11 gives the desired bound (l). If $t_1 = 1$, then we must work a little harder. Either $xyz \in \mathcal{T}_2$ sends at most seven edges to the triangle $uvw \in \mathcal{T}_1$, in which case xyz sends in total at most $3m + 7 + 2$ edges to $\mathcal{T}_1 \cup \mathcal{M}$, or xyz sends more than seven edges to uvw. In this case, we claim that no edge of $\mathcal{M}$ sees xyz, or we would have an improving rotation exactly as in the proof of bound (f) of Lemma 11. It follows that xyz sends at most $3m$ edges to $\mathcal{M}$, and so in total again at most $3m + 9$ edges to $\mathcal{T}_1 \cup \mathcal{M}$. Now summation yields the desired bound (l).

We next prove (m). Suppose $j \in \{3, 4\}$. Again the $m = 0$ case is trivial. Again by definition of $\mathcal{T}_j$, at most one edge in $\mathcal{M}$ sees the triangle $xyz \in \mathcal{T}_j$, and thus we have that xyz sends at most $3m + 3$ edges to $\mathcal{M}$. Again, if

$t_1 \neq 1$ then summation combined with the bound (j) of Lemma 11 yields the desired bound (m). Again, if $t_1 = 1$ then we either have that xyz sends at most seven edges to the triangle $uvw \in \mathcal{T}_1$, and so in total $3m + 10$ edges to $\mathcal{T}_1 \cup \mathcal{M}$, or it sends more than seven edges to uvw but is not seen by any edge of $\mathcal{M}$ (or this would create an improving rotation), and so sends at most $3m + 9$ edges to $\mathcal{T}_1 \cup \mathcal{M}$. Again the desired bound (m) follows by summation.

Finally we prove the bound (n). Suppose $j \in \{3, 4\}$. Observe that if $i = 0$ then we have by definition of $\mathcal{T}_2$ that $t_2 = 0$ and hence the bound follows. Now by definition of $\mathcal{T}_j$, at most one vertex of $\mathcal{I}$ sees the triangle $xyz \in \mathcal{T}_j$, and all other vertices of $\mathcal{I}$ therefore send at most one edge to xyz. We conclude that xyz sends at most $3 + (i - 1) = i + 2$ edges to $\mathcal{I}$. If $t_2 \neq 1$, then summation and the bound (k) of Lemma 11 yield the desired bound (n). If $t_2 = 1$, then there are two possibilities. First, xyz sends at most eight edges to the triangle $abc \in \mathcal{T}_2$, in which case it sends in total at most $10 + i$ edges to $\mathcal{T}_2 \cup \mathcal{I}$. Second, xyz sends nine edges to abc, in which case there can exist no vertex of $\mathcal{I}$ which sees xyz or we would have an improving rotation exactly as in the proof of bound (h) of Lemma 11. Then xyz sends in total at most $i + 9$ edges to $\mathcal{T}_2 \cup \mathcal{I}$. The desired bound (n) follows by summation. □

Observe that, at this stage, we provided bounds for all (bipartite or internal) edge sets but $e(\mathcal{T}_4)$. These bounds, with the exception of the bounds on edges in $\mathcal{T}_4 \cup \mathcal{M} \cup \mathcal{I}$, will turn out to be strong enough for all parts of the proof of Theorem 6. It is convenient to summarise them in one function. First, let

$$\begin{aligned} f'(\tau_1, \tau_2, \tau_3, \tau_4, \mu, \iota) :=& 4\mu\tau_1 + 2\iota\tau_1 + 7\binom{\tau_1}{2} + 3\tau_1 + 2\iota\tau_2 \\ & + 8\binom{\tau_2}{2} + 3\tau_2 + 8\binom{\tau_3}{2} + 8\tau_3\tau_4 + 3\tau_3 \\ & + 7\tau_1\tau_2 + (2 + 3\mu)\tau_2 + 7\tau_1(\tau_3 + \tau_4) \\ & + (3 + 3\mu)\tau_3 + 8\tau_2(\tau_3 + \tau_4) + (2 + \iota)\tau_3 \,. \end{aligned} \tag{3}$$

We now define $f(\tau_1, \tau_2, \tau_3, \tau_4, \mu, \iota)$ by

$$f := \begin{cases} f'(\tau_1, \tau_2, \tau_3, \tau_4, \mu, \iota) & \text{when } \mu \geq 1 \text{ and } \iota \geq 1 \\ f'(\tau_1, \tau_2, \tau_3, \tau_4, \mu, \iota) - (2\tau_2 + 3\tau_3) & \text{when } \mu = 0 \text{ and } \iota \geq 1 \\ f'(\tau_1, \tau_2, \tau_3, \tau_4, \mu, \iota) - 2\tau_3 & \text{when } \mu \geq 1 \text{ and } \iota = 0 \\ f'(\tau_1, \tau_2, \tau_3, \tau_4, \mu, \iota) - (2\tau_2 + 5\tau_3) & \text{when } \mu = 0 \text{ and } \iota = 0 \end{cases} \tag{4}$$

The purpose of the functions f and f' is the following. When $t_1, t_2 \neq 1$, we have by summing the bounds in parts (d)–(i) of Lemma 10, the $j = 4$ cases of parts (j) and (k) of Lemma 11, part (l) of Lemma 12 and the $j = 3$ cases of parts (m) and (n) of Lemma 12 that

$$e(G) - e(\mathcal{T}_4 \cup \mathcal{M} \cup \mathcal{I}) \leq f(t_1, t_2, t_3, t_4, m, i) \,.$$

We observe that the reason that f and f' differ is that Lemma 12 yields different bounds depending on whether m or i is zero, i.e., we have

$$e(G) - e(\mathcal{T}_4 \cup \mathcal{M} \cup \mathcal{I}) \leq f'(t_1, t_2, t_3, t_4, m, i) .$$

We further observe that although $e(G) - e(\mathcal{T}_4 \cup \mathcal{M} \cup \mathcal{I}) \leq f(t_1, t_2, t_3, t_4, m, i)$ is valid in general only when $t_1, t_2 \neq 0$, by parts (m) and (n) of Lemma 12 the following is always valid.

$$\begin{aligned} &e(G) - e(\mathcal{T}_4 \cup \mathcal{M} \cup \mathcal{I}) + e(\mathcal{T}_4, \mathcal{M} \cup \mathcal{I}) \\ &\quad \leq e(G) - e(\mathcal{T}_4 \cup \mathcal{M} \cup \mathcal{I}) + e(\mathcal{T}_4, \mathcal{M} \cup \mathcal{I}) + e(\mathcal{T}_1 \cup \mathcal{T}_2, \mathcal{T}_4) \\ &\quad \leq f(t_1, t_2, t_3, t_4, m, i) + (3 + 3m)t_4 + (2 + i)t_4 . \end{aligned} \tag{5}$$

As previously mentioned, our proof has a combinatorial part and an arithmetic part: we need to know the maxima of several functions, of which f is the first. We state the required lemma here, but defer the proof to Appendix A. Let

$$\begin{aligned} F(n, k) := \big\{ &(\tau_1, \tau_2, \tau_3, \tau_4, \mu, \iota) \in \mathbb{N}_0^6 : \\ &\tau_1 + \tau_2 + \tau_3 + \tau_4 = k \, , 2\mu + \iota = n - 3k \big\} . \end{aligned} \tag{6}$$

Lemma 13. *When $n \geq 3k + 2$ we have*

$$\max_{(\tau_1, \tau_2, 0, 0, \mu, \iota) \in F(n,k)} \big(f(\tau_1, \tau_2, 0, 0, \mu, \iota) + \iota\mu + \mu^2 \big) = \max_{j \in [3]} e\big(E_j(n, k)\big) .$$

A trivial upper bound for $e(\mathcal{T}_4)$ is given by

$$e(\mathcal{T}_4) \leq \binom{3t_4}{2}. \tag{7}$$

It turns out that this trivial bound suffices to prove Lemma 9 (but not Theorem 6). We define $h(\tau_1, \tau_2, \tau_3, \tau_4, \mu, \iota)$ by

$$h := f(\tau_1, \tau_2, \tau_3, \tau_4, \mu, \iota) + \iota\mu + \mu^2 + (3 + 3\mu)\tau_4 + (2 + \iota)\tau_4 + \binom{3\tau_4}{2} . \tag{8}$$

Proof of Lemma 9. Let $k \leq \frac{n-8}{5}$. Let G and its decomposition be as in Setup 7. In particular, we obtain numbers $t_1, \ldots, t_4, m, i$. By (5), (a)–(c) of Lemma 10, and (7) we have $e(G) \leq h(t_1, t_2, t_3, t_4, m, i)$ for the function h defined in (8). From (3), (4), and (8) one can check that

$$h(t_1, t_2, 0, t_3 + t_4, m, i) \geq h(t_1, t_2, t_3, t_4, m, i) .$$

Also from (8) we have the following.

$$\begin{aligned} &h(t_1 + t_3 + t_4, t_2, 0, 0, m, i) - h(t_1, t_2, 0, t_3 + t_4, m, i) \\ &= (t_3 + t_4)(m + i - t_2 - t_3 - t_4 - 4) \\ &\geq (t_3 + t_4)\frac{n - 5k - 8}{2} , \end{aligned} \tag{9}$$

where the inequality comes from $t_2 + t_3 + t_4 \leq k$ and $2m + i = n - 3k$. Since $n - 5k - 8 \geq 0$, we have

$$h(t_1 + t_3 + t_4, t_2, 0, 0, m, i) \geq h(t_1, t_2, t_3, t_4, m, i) .$$

Now $h(t_1+t_3+t_4,t_2,0,0,m,i)=f(t_1+t_3+t_4,t_2,0,0,m,i)+im+m^2$, so by Lemma 13 we have

$$e(G)\le h(t_1,t_2,t_3,t_4,m,i)\le \max_{j\in[3]} e\big(E_j(n,k)\big)\,.$$

Finally, according to Table 1, this maximum is given by $e\big(E_1(n,k)\big)$, completing the proof. □

5. Large rotations

In order to prove Theorem 6 we need to improve the bounds given in the previous section on the number of edges touching $\mathcal{T}_4$; in particular, we need stronger bounds than the trivial $e(\mathcal{T}_4)\le\binom{3|\mathcal{T}_4|}{2}$. We will obtain these stronger bounds by describing rotations using many more—up to 29—triangles. In constructing these rotations, we will need to assume that $\mathcal{T}_4$ does not contain too *few* edges, which will lead to a case distinction in the proof of Theorem 6.

Recall that by definition of $\mathcal{T}_4$, every triangle in $\mathcal{T}_4$ sends more than $8(t_4-1)$ edges to the other triangles of $\mathcal{T}_4$, which should be seen as something like a 'minimum degree' condition. Imposing the further condition $e(\mathcal{T}_4)\ge 8\binom{t_4}{2}+10t_4-27$ has the consequence that there must exist some pairs of triangles in $\mathcal{T}_4$ which are connected by nine edges; the combination of the two features makes $\mathcal{T}_4$ an exceptionally good place for construction of complex rotations. Our aim is to take advantage of this in order to provide a good bound on $e(\mathcal{T}_4\cup\mathcal{M}\cup\mathcal{I})$.

Unfortunately, this will mean that we can no longer use Lemma 12 to provide us with our upper bounds on $e(\mathcal{T}_1\cup\mathcal{M},\mathcal{T}_4)$ and $e(\mathcal{T}_2\cup\mathcal{I},\mathcal{T}_4)$, and we will be forced to use instead Lemma 11. This lemma only gives bounds on $e(\mathcal{T}_1,\mathcal{T}_4)$ when $t_1\ne 1$, and on $e(\mathcal{T}_2,\mathcal{T}_4)$ when $t_2\ne 1$, which causes a problem that we must now deal with. Consequently, if either $t_1=1$ and the triangle in $\mathcal{T}_1$ sends more than $7t_4+18$ edges to $\mathcal{T}_4$, or $t_2=1$ and the triangle in $\mathcal{T}_2$ sends more than $8t_4$ edges to $\mathcal{T}_4$, or both, we will have to handle these one or two exceptional triangles along with $\mathcal{T}_4$. Fortunately, this adds only a slight complication.

Let $\mathcal{T}_5$ contain all triangles of $\mathcal{T}_4$, together with $\mathcal{T}_1$ if $t_1=1$ and $e(\mathcal{T}_1,\mathcal{T}_4)>7t_4+18$, and with $\mathcal{T}_2$ if $t_2=1$ and $e(\mathcal{T}_2,\mathcal{T}_4)>8t_4$. Let $t_5=|\mathcal{T}_5|$. That is, we have $t_4\le t_5\le t_4+2$.

First, the fact that every triangle in $\mathcal{T}_4$ sends more than $8(t_4-1)$ edges to the other triangles of $\mathcal{T}_4$ makes $\mathcal{T}_4$ well connected. The following definition makes this precise.

Definition 14 (connect, favour)**.** *Given two triangles T and T', we say that a third triangle T''* connects *T to T', or that there is a* connection *from T to T' via T'', if one of the following two conditions holds.*

(i) There are at least 8 edges from T'' to both T and T', or
(ii) There are 9 edges from T'' to T, and at least 7 from T'' to T'.

To emphasise that the definition is not symmetric in T and T' we say that the connection favours T *and also write* $T \rightsquigarrow T'' \rightsquigarrow T'$.

We show that two triangles in $\mathcal{T}_4$ can be connected in many different ways.

Lemma 15. *For any pair of distinct triangles T and T' of $\mathcal{T}_4$, there are at least $\frac{1}{12}(t_4 - 2)$ triangles $T'' \in \mathcal{T}_4$ with $T \rightsquigarrow T'' \rightsquigarrow T'$.*

Proof. Suppose first that there are at least $\frac{7}{12}(t_4-2)$ triangles in $\mathcal{T}_4 \setminus \{T, T'\}$ which send 8 or more edges to T. By the definition of $\mathcal{T}_4$ we have $e(T', \mathcal{T}_4 \setminus \{T'\}) > 8(t_4 - 1)$, and so in particular there are at most $\frac{1}{2}(t_4 - 2)$ triangles of $\mathcal{T}_4 \setminus \{T, T'\}$ which send seven or less edges to T'. Hence at least $\frac{1}{12}(t_4-2)$ triangles of $\mathcal{T}_4$ must send at least eight edges to both T and T', as required.

If on the other hand there are less than $\frac{7}{12}(t_4-2)$ triangles in $\mathcal{T}_4 \setminus \{T, T'\}$ sending eight or more edges to T, then there are more than $\frac{5}{12}(t_4 - 2)$ triangles of $\mathcal{T}_4 \setminus \{T, T'\}$ which send at most seven edges to T. Hence, since $e(T, \mathcal{T}_4 \setminus \{T, T'\}) \geq 8(t_4-2)$, there must also be more than $\frac{5}{12}(t_4-2)$ triangles in $\mathcal{T}_4$ which send nine edges to T. Again by definition of $\mathcal{T}_4$, of these, at least $\frac{1}{12}(t_4 - 2)$ must also send seven or more edges to T', as required. □

Our next Lemma now uses this observation to obtain structural information about $\mathcal{T}_5 \cup \mathcal{M} \cup \mathcal{I}$. Here we need that t_4 is sufficiently large.

Lemma 16. *Provided that $e(\mathcal{T}_4) \geq 8\binom{t_4}{2} + 10t_4 - 27$ and $t_4 \geq 176$, there is no set of vertex-disjoint triangles induced by $V(\mathcal{T}_5 \cup \mathcal{M} \cup \mathcal{I})$ which covers three or more vertices of $\mathcal{M} \cup \mathcal{I}$.*

Proof. Suppose the statement is false, that is, there exists a set of vertex-disjoint triangles in $\mathcal{T}_5 \cup \mathcal{M} \cup \mathcal{I}$ which covers three or more vertices of $\mathcal{M} \cup \mathcal{I}$.

Then we have one of the following three Situations.

(i) There are three triangles which each consist of a vertex of $\mathcal{M} \cup \mathcal{I}$ and an edge in $\mathcal{T}_5$.

(ii) There is one such triangle and one triangle consisting of an edge in $\mathcal{M} \cup \mathcal{I}$ and a vertex of $\mathcal{T}_5$.

(iii) There are two triangles of the latter type.

We denote the set of these two or three vertex disjoint triangles by $\mathcal{S}$ and call them *extra triangles*. We denote the set of vertices in these triangles that are in $\mathcal{T}_5$ by Z. Observe that $|Z| \leq 6$ and therefore Z meets at most six triangles in $\mathcal{T}_5$ which we denote by $\mathcal{Z}_5 \subseteq \mathcal{T}_5$.

The idea now is as follows. If we are in Case (i) and $\mathcal{Z}_5$ contained only two triangles we immediately arrived at a contradiction since we could replace $\mathcal{Z}_5$ by $\mathcal{S}$ and obtain a triangle factor with one triangle more than $\mathcal{T}$. Similarly, if we are in Case (ii) or (iii) we cannot have $|\mathcal{Z}_5| = 1$. These two observations together mean that we cannot have $|\mathcal{Z}_5| < |\mathcal{S}|$. We will show in the following that, by way of a sequence of rotations, we can turn any configuration of $\mathcal{Z}_5$ into a configuration resembling such a situation and hence arrive at a contradiction.

More precisely, we shall proceed as follows. Let V_5 be the set of vertices covered by $\mathcal{Z}_5 \cup \mathcal{S}$. Throughout our process we shall keep track of a set of *new triangles* $\mathcal{N}'$ and a set of *deleted triangles* $\mathcal{D}'$ such that

$$\mathcal{N}' \cap \mathcal{T}_5 = \emptyset \qquad \text{and} \qquad \mathcal{D}' \subseteq \mathcal{T}_5 \qquad \text{and} \qquad |\mathcal{D}'| = |\mathcal{N}'| \leq 29\,. \tag{10}$$

In the beginning we set $\mathcal{D}' = \mathcal{N}' = \emptyset$. In each step, we will consider the set of triangles

$$\mathcal{T}_5' := (\mathcal{T}_5 \setminus \mathcal{D}') \cup \mathcal{N}' \cup \mathcal{S}\,.$$

It will not be true in general throughout the process that $\mathcal{T}_5'$ is a triangle factor (observe that this for example fails initially). On the other hand we will always have that

$$\text{each vertex of } V_5 \text{ is covered either by one or by two triangles of } \mathcal{T}_5'\,. \tag{11}$$

We will denote the set of vertices covered by two triangles by Z' and call them the *marked vertices*. We let $\mathcal{Z}_5'$ be the set of those triangles of $\mathcal{T}_5$ which contain a marked vertex and we call these triangles the *marked triangles*. Note that in the beginning we have $Z' = Z$ and $\mathcal{Z}_5' = \mathcal{Z}_5$. Further, in each step we will have that

$$\text{every marked vertex is contained in a triangle of } \mathcal{T}_5 \setminus \mathcal{D}'\,, \tag{12}$$

which implies that in each step $\mathcal{T}_5' \setminus \mathcal{Z}_5'$ is a triangle factor of size $|\mathcal{T}_5| - |\mathcal{Z}_5'| + |\mathcal{S}|$ by (10).

In each step we will now perform a rotation by adding some vertex disjoint triangles in $G[V_5]$ to the set of new triangles $\mathcal{N}'$, and deleting as many triangles from $\mathcal{T}_5' \cap \mathcal{T}_5$, i.e., we will add these triangles to the set of deleted triangles $\mathcal{D}'$. We will have three preparation steps (Preparation 1–3) and three main rotation types (Type 1–3). No step will change the size of Z' and

$$\text{each rotation of Type 1, 2 or 3 will decrease the size of } \mathcal{Z}_5'\,. \tag{13}$$

We will stop when $|\mathcal{Z}_5'| < |\mathcal{S}|$, since then $\mathcal{T}_5' \setminus \mathcal{Z}_5'$ is a triangle factor with more triangles than $\mathcal{T}_5$, a contradiction.

It remains to construct $\mathcal{Z}_5'$ with these properties. We will first carry out three preparatory steps: roughly, these consist of locating two disjoint copies of K_6 in $\mathcal{T}_4$ (Preparation 1) and showing that we can 'move' Z' to $\mathcal{T}_4$ (which is useful because Lemma 15 then applies), in Preparations 2 and 3. After this we have either two, three, or six marked vertices in $\mathcal{T}_4$. Our next aim is to 'move around' these vertices within $\mathcal{T}_4$ such that they are contained in one, one or two (respectively) triangles of $\mathcal{T}_4$. Achieving this immediately gives us $|\mathcal{Z}_5'| < |\mathcal{S}|$, which is what we want. To do this we make use of our main rotation Types 1, 2 and 3. We will now give details of the Preparation steps and the main rotation Types.

Preparation 1. There are two disjoint pairs (T_1, T_1') and (T_2, T_2') of triangles in $\mathcal{T}_4$ which do not meet Z and are such that $V(T_1) \cup V(T_1')$ and $V(T_2) \cup V(T_2')$

each induce a copy of K_6 in G. We set $\mathcal{K}_6 := \{T_1, T_1', T_2, T_2'\}$ and call (T_1, T_1') and (T_2, T_2') the K_6-copies of $\mathcal{K}_6$.

To see this, let $H = (\mathcal{T}_4, E_H)$ be the auxiliary graph with edges exactly between those triangles $T, T' \in \mathcal{T}_4$ which are connected by nine edges. Since $e(\mathcal{T}_4) \geq 8\binom{t_4}{2} + 10t_4 - 27$ by assumption and

$$e(\mathcal{T}_4) \leq 9e(H) + 8\big(\tbinom{t_4}{2} - e(H)\big) + 3t_4 = e(H) + 8\tbinom{t_4}{2} + 3t_4,$$

we conclude that $e(H) \geq 7t_4 - 27$. Since $\max\big(7(t_4 - 7 + \binom{7}{2}), \binom{17}{2}\big) = 7t_4 - 28 < e(H)$, we can apply Theorem 3 to H and infer that there are at least eight independent edges in H, and hence at least two independent edges in H which do not meet $\mathcal{Z}_5$. These two edges give us the pairs (T_1, T_1') and (T_2, T_2').

Preparation 2. Suppose that $\{uvw\} = \mathcal{T}_1 \subseteq \mathcal{T}_5$. We distinguish four cases.

Case 1: In the case when $Z \cap \mathcal{T}_1 = \emptyset$ we do not do anything.

Case 2: If $Z \cap \mathcal{T}_1 = \{u\}$, then we consider the edges between uvw and $\mathcal{T}_4$. Because there are in total at least $7t_4 + 19$ such edges (recall that this was the condition for inclusion of $\mathcal{T}_1$ in $\mathcal{T}_5$), in particular there must be at least ten triangles of $\mathcal{T}_4$ to which uvw sends more than seven edges. Now at most 9 of these triangles are in $\mathcal{Z}_5 \cup \mathcal{K}_6$, as no triangle of $\mathcal{T}_4$ covers $u \in Z$. Therefore there is a triangle $xyz \in \mathcal{T}_4 \setminus (\mathcal{Z}_5 \cup \mathcal{K}_6)$ to which uvw sends at least eight edges. Thus vw has a common neighbour, say x, in xyz. We add xvw to the set of new triangles $\mathcal{N}'$, and uvw to the set of deleted triangles $\mathcal{D}'$. The upshot is that u is no longer marked, but x, which lies in a triangle of $\mathcal{T}_4 \setminus (\mathcal{Z}_5 \cup \mathcal{K}_6)$, is.

Case 3: If $Z \cap \mathcal{T}_1 = \{u, v\}$ then we work similarly: again, there is a triangle $xyz \in \mathcal{T}_4 \setminus (\mathcal{Z}_5 \cup \mathcal{K}_6)$ to which uvw sends at least eight edges, and we may assume w is adjacent to both x and y. We add xyw to $\mathcal{N}'$, and uvw to $\mathcal{D}'$. The result is that u and v are no longer marked, but x and y are.

Case 4: If $Z \cap \mathcal{T}_1 = \mathcal{T}_1$, we may again simply ignore $\mathcal{T}_1$. The only possibility is that we are in situation (i) or in situation (ii). We rule out situation (ii) as follows. If auv and bcw are the two triangles from situation (ii) then replacing $uvw \in \mathcal{T}_1$ by auv and bcw is an improving rotation, a contradiction.

Preparation 3. Suppose that $\{u'v'w'\} = \mathcal{T}_2 \subseteq \mathcal{T}_5$. We behave exactly as above, with the exception that since $e(\mathcal{T}_2, \mathcal{T}_4) > 8t_4$ there must be at least $t_4/2$ triangles of $\mathcal{T}_4$ to which $u'v'w'$ sends at least eight edges. We require $t_4/2 \geq 10$, but this is guaranteed by our assumption $t_4 \geq 176$.

Before describing the main rotation Types, let us briefly recap the current situation. We have a set Z' of marked vertices, which contains either six, three or two vertices (in Situation (i), (ii) or (iii) respectively). If $|\mathcal{T}_1| = |\mathcal{T}_2| = 1$ and there are six marked vertices are in $\mathcal{T}_1 \cup \mathcal{T}_2$ then removing the two triangles $\mathcal{T}_1 \cup \mathcal{T}_2$ from $\mathcal{T}$ and adding the three triangles $\mathcal{S}$ is an improving rotation, which is a contradiction. It follows that either all the marked vertices are in $\mathcal{T}_4$, or we have six marked vertices, of which either three are in the unique triangle of $\mathcal{T}_1$ or three are in the unique triangle of $\mathcal{T}_2$, and the remaining three are in $\mathcal{T}_4$. We have a set of at most two deleted triangles $\mathcal{D}'$ (at most one from each of Preparation 2 and 3) none of which are in $\mathcal{T}_4$. Finally, we have a set $\mathcal{K}_6$ consisting of four triangles of $\mathcal{T}_4$ which span two disjoint copies of K_6, none of whose vertices are marked.

We now describe the main rotation Types.

Type 1. Suppose that $\big|\mathcal{T}_4 \cap (\mathcal{D}' \cup \mathcal{Z}'_5)\big| \leq 11$, and that there are two triangles uvw and $u'v'w'$ of $\mathcal{Z}'_5$, such that $Z' \cap \{u,v,w,u',v',w'\} = \{u,u'\}$. We can add two triangles to $\mathcal{D}'$, neither in $\mathcal{K}_6$, and two triangles to $\mathcal{N}'$, and obtain $\big|Z' \cap \{u,v,w\}\big| = 2$.

This type of rotation can be constructed for the following reason. By Lemma 15 there are $\frac{1}{12}(t_4-2) > 14$ triangles xyz in $\mathcal{T}_4$ such that $uvw \rightsquigarrow xyz \rightsquigarrow u'v'w'$. Of these, at most 4 are in $\mathcal{K}_6$, and, because xyz is neither uvw nor $u'v'w'$, at most 9 are in $\mathcal{T}_4 \cap (\mathcal{D}' \cup \mathcal{Z}'_5)$. It follows that we may choose xyz in $\mathcal{T}_4 \setminus (\mathcal{D}' \cup \mathcal{K}_6 \cup \mathcal{Z}'_5)$ such that $uvw \rightsquigarrow xyz \rightsquigarrow u'v'w'$. Because of this connection, at least one vertex of xyz, say x, is adjacent to both v' and w'. In addition, because the connection favours uvw, at least two vertices of uvw are adjacent to y and z. In particular, one vertex of uvw different from u, say v, forms a triangle with y and z. Now we can rotate by adding the triangles $u'v'w'$ and xyz to the set $\mathcal{D}'$ of deleted triangles and the triangles $v'w'x$ and vyz to the set $\mathcal{N}'$ of new triangles. Observe that this rotation satisfies (10) and (11). Further, it removes u' from Z' and $u'v'w'$ from $\mathcal{Z}'_5$ and adds v to Z' and no new triangle to $\mathcal{Z}'_5$. Hence (12) and (13) are also satisfied.

Type 2. Suppose that $\big|\mathcal{T}_4 \cap (\mathcal{D}' \cup \mathcal{K}_6 \cup \mathcal{Z}'_5)\big| \leq 15$, that at least one of the copies of K_6 in $\mathcal{K}_6$ does not meet $\mathcal{D}'$, and that there are two triangles uvw and $u'v'w'$ of $\mathcal{Z}'_5$, such that $Z' \cap \{u,v,w,u',v',w'\} = \{u,v,u'\}$. We can add five triangles to $\mathcal{D}'$, exactly two of which are in $\mathcal{K}_6$, and five triangles to $\mathcal{N}'$, and obtain $Z' \cap \{u,v,w\} = \{u,v,w\}$.

Let the copy of K_6 in $\mathcal{K}_6$ not meeting $\mathcal{D}'$ be on the triangles abc, def of $\mathcal{T}_4$. By Lemma 15 there are at least $\frac{1}{12}(t_4-2)$ triangles xyz in $\mathcal{T}_4$ with $uvw \rightsquigarrow xyz \rightsquigarrow abc$. Since xyz is neither uvw nor abc, by assumption there are at least two choices of $xyz \notin \mathcal{D}' \cup \mathcal{K}_6 \cup \mathcal{Z}'_5$. We fix one. Similarly, by Lemma 15 there is a choice of triangle $x'y'z'$ in $\mathcal{T}_4 \setminus (\mathcal{D}' \cup \mathcal{K}_6 \cup \mathcal{Z}'_5 \cup \{uvw\})$ with $u'v'w' \rightsquigarrow x'y'z' \rightsquigarrow def$ (see also Figure 3). Because of the second connection, at least one vertex of $x'y'z'$, say x', is a common neighbour of $v'w'$, and at least one vertex of def, say e, is a common neighbour of $y'z'$. We conclude that $x'v'w'$ and $ey'z'$ are triangles.

Now we distinguish two possibilities concerning the connection between uvw and abc. First, there are at least eight edges from xyz to both uvw and abc. In this case, we are guaranteed that at least two vertices, say x and y, of xyz are adjacent to w, and at least two vertices, say a and b, of abc are adjacent to z. Hence xyw, abz and cdf are triangles in G. Second, since the

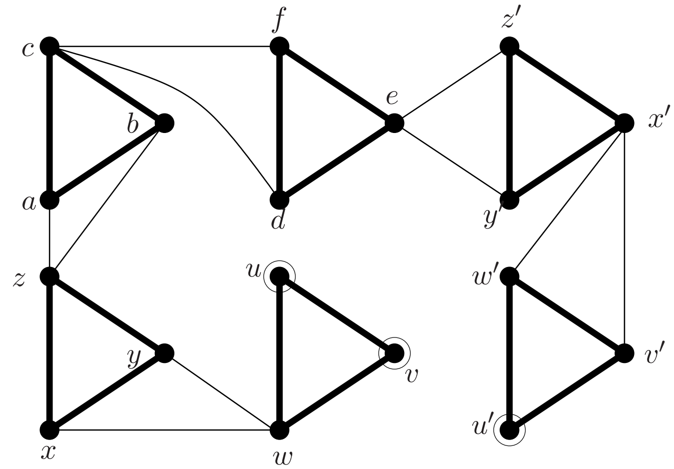

FIGURE 3. The second rotation type.

connection favours uvw, from xyz there are nine edges to uvw and seven to abc. Some vertex of xyz, say z, is adjacent to both, a and b. Therefore, again, abz, xyw, and cdf are triangles.

Accordingly we can rotate by adding $x'v'w'$, $ey'z'$, cdf, abz, and xyw to $\mathcal{N}'$, and $u'v'w'$, $x'y'z'$, def, abc, and xyz to $\mathcal{D}'$. This deletes u' from Z' and hence $u'v'w'$ from $\mathcal{Z}'_5$, it adds w to Z' and no triangle to $\mathcal{Z}'_5$. Hence, as can easily be checked, this rotation satisfies (10), (11), (12), and (13).

Type 3. Suppose that $\left|\mathcal{T}_4 \cap (\mathcal{D}' \cup \mathcal{K}_6 \cup \mathcal{Z}'_5)\right| \leq 12$ and $\mathcal{K}_6 \cap \mathcal{D}' = \emptyset$, and that there are three triangles uvw, $u'v'w'$, and $u''v''w''$ of $\mathcal{Z}'_4 \cap \mathcal{T}_4$ such that $Z' = \{u, v, u', v', u'', v''\}$. Then we can add at most ten triangles to $\mathcal{D}'$ and ten triangles to $\mathcal{N}'$, and obtain $Z' = \{u, v, w, u', v', w'\}$.

Let the two copies of K_6 in $\mathcal{K}_6$ be (abc, def) and $(a'b'c', d'e'f')$. We apply Lemma 15 five times to obtain the following connections which avoid each other and whose connecting triangles are from $\mathcal{T}_4 \setminus (\mathcal{D}' \cup \mathcal{K}_6 \cup \mathcal{Z}'_5)$: $uvw \rightsquigarrow xyz \rightsquigarrow abc$, $u'v'w' \rightsquigarrow x'y'z' \rightsquigarrow def$, $u''v''w'' \rightsquigarrow x''y''z'' \rightsquigarrow a'b'c'$, $abc \rightsquigarrow a''b''c'' \rightsquigarrow def$, and $d'e'f' \rightsquigarrow d''e''f'' \rightsquigarrow a''b''c''$. Observe that this is possible since at each application Lemma 15 guarantees at least 15 connecting triangles in $\mathcal{T}_4$, while at each application there are by assumption at most $12 - 2 = 10$ triangles of $\mathcal{D}' \cup \mathcal{K}_6 \cup \mathcal{Z}'_5$ to avoid (since two triangles from this set are being connected and are thus automatically avoided), together with the at most four previously determined connecting triangles which must also be avoided.

Arguing similarly as before, these connections guarantee, without loss of generality, the triangles wxy, $w'x'y'$, $w''x''y''$, zab, $z'de$, and $z''a'b'$. Since cf belongs to a K_6, it is an edge, and because of the connection $abc \rightsquigarrow a''b''c'' \rightsquigarrow def$, there is a vertex, say a'', of $a''b''c''$ which is adjacent to both c and f, and so cfa'' is a triangle of G. Finally, using the connection $d'e'f' \rightsquigarrow d''e''f'' \rightsquigarrow a''b''c''$, we can find a common neighbour, say d'', of $b''c''$

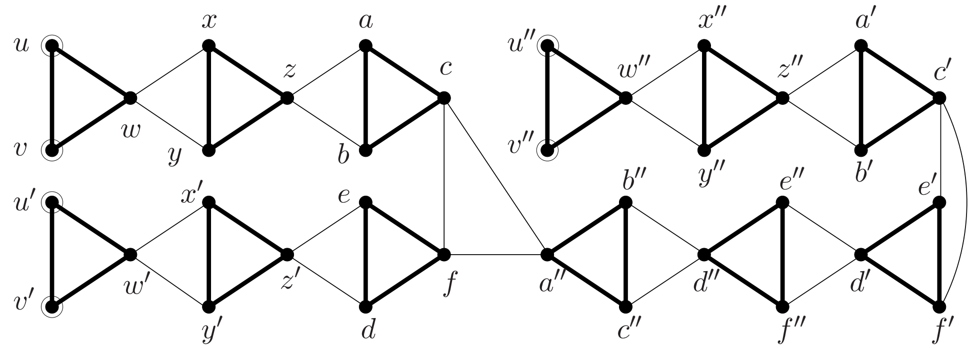

FIGURE 4. The third rotation type.

in $d''e''f''$, and and a common neighbour, say d', of $e''f''$ in $d'e'f'$. Hence, $b''c''d''$, $d'e''f''$, and $c'e'f'$ are triangles of G. See Figure 4.

We conclude that we can rotate by adding the ten triangles wxy, $w'x'y'$, $w''x''y''$, zab, $z'de$, $z''a'b'$, cfa'', $b''c''d''$, $d'e''f''$, $c'e'f'$ to $\mathcal{N}'$ and adding the ten triangles $u''v''w''$, xyz, $x'y'z'$, $x''y''z''$, abc, def, $a'b'c'$, $d'e'f'$, $a''b''c''$, $d''e''f''$ to $\mathcal{D}'$. This removes u'' and v'' from Z' and hence $u''v''w''$ from $\mathcal{Z}_5'$, and adds w and w' to Z' and no new triangle to $\mathcal{Z}_5'$. Again, it is easy to check that this rotation satisfies (10), (11), (12), and (13).

We now explain how we apply these rotation Types. If we started in Situation (iii), then Z' consists of two vertices in $\mathcal{T}_4$. These two vertices are in distinct triangles of $\mathcal{T}_4$ (since $|\mathcal{Z}_5'| \geq |\mathcal{S}| = 2$). We apply rotation Type 1 to the two triangles of $\mathcal{Z}_5'$, which we can do since $\left|\mathcal{T}_4 \cap (\mathcal{D}' \cup \mathcal{Z}_5')\right| = 2$. This adds two triangles to each of $\mathcal{D}'$ and $\mathcal{N}'$, and reduces $\mathcal{Z}_5'$ to one triangle. So we have $|\mathcal{Z}_5'| < |\mathcal{S}|$ and we are done.

If we started in Situation (ii), then Z' consists of three vertices in $\mathcal{T}_4$. These may either lie in two or three triangles of $\mathcal{T}_4$ (since $|\mathcal{Z}_5'| \geq |\mathcal{S}| = 2$). In the former case we apply rotation Type 1 to two of the triangles of $\mathcal{Z}_5'$ (which we may do for the same reason as above), which reduces $\mathcal{Z}_5'$ to two triangles. Now since $\mathcal{Z}_5'$ has two triangles, so one contains two vertices of Z' and the other contains one. We apply rotation Type 2 to the two triangles of $\mathcal{Z}_5'$, which we may do since $\left|\mathcal{T}_4 \cap (\mathcal{D}' \cup \mathcal{K}_6 \cup \mathcal{Z}_5')\right| \leq 2 + 4 + 2 = 8$, and obtain $|\mathcal{Z}_5'| = 1 < |\mathcal{S}|$: we are done.

Finally, suppose we started in Situation (i). Now Z' contains six vertices and $|\mathcal{S}| = 3$. These cannot all lie in two triangles of $\mathcal{T}_5$, since otherwise deleting these two triangles and adding $\mathcal{S}$ to $\mathcal{T}$ is an improving rotation. If three vertices of Z' lie in one triangle of $\mathcal{T}_5$, and the remaining three lie in either two or three triangles (which must be in $\mathcal{T}_4$) then we apply the identical rotation strategy as in Situation (ii). We may have $\mathcal{Z}_5'$ larger by one than there, but nevertheless the rotations exist. The remaining possibility is that all six vertices of Z' lie in $\mathcal{T}_4$, and no three are contained in any one triangle of $\mathcal{T}_4$. We separate several possibilities.

First, if the six vertices lie in three triangles of $\mathcal{T}_4$, then we apply rotation Type 3, which we may do since $\left|\mathcal{T}_4 \cap (\mathcal{D}' \cup \mathcal{K}_6 \cup \mathcal{Z}'_5)\right| = 0 + 4 + 3 = 7$, and obtain $|\mathcal{Z}'_5| = 2 < |\mathcal{S}|$: we are done.

If the six vertices lie in five or six triangles of $\mathcal{T}_4$, then we apply rotation Type 1 either once or twice. In the first application we have $\left|\mathcal{T}_4 \cap (\mathcal{D}' \cup \mathcal{Z}'_5)\right| \leq 0 + 6 = 6$, while in the second application (if we apply it twice) $\left|\mathcal{T}_4 \cap (\mathcal{D}' \cup \mathcal{Z}'_5)\right| \leq 2 + 5 = 7$, so we are permitted to do this. We add either two or four triangles to each of $\mathcal{D}'$ and $\mathcal{N}'$, and reduce $\mathcal{Z}'_5$ to four triangles, which is our final case.

The final case we have to handle is that $\mathcal{Z}'_5$ contains four triangles, of which two contain two vertices of Z' each and two contain one each. We apply rotation Type 2 twice. In the first application we have $\left|\mathcal{T}_4 \cap (\mathcal{D}' \cup \mathcal{K}_6 \cup \mathcal{Z}'_5)\right| \leq 4+4+4 = 12$, while in the second we have $\left|\mathcal{T}_4 \cap (\mathcal{D}' \cup \mathcal{K}_6 \cup \mathcal{Z}'_5)\right| \leq 9+2+3 = 14$, since the first application adds five triangles to $\mathcal{D}'$, two of which are in $\mathcal{K}_6$, and therefore we can indeed construct these rotations. After the second application of rotation Type 2 we have $|\mathcal{Z}'_5| = 2 < |\mathcal{S}|$ and we are done. □

We are able to convert the structural information provided by Lemma 16 into an upper bound on $e(\mathcal{T}_5 \cup \mathcal{M} \cup \mathcal{I})$. We need to define the following function.

$$p(h,a) := \begin{cases} a(h-a) + \binom{h-2a}{2} + 6h & 2a \leq h < 9a\,, \\ (a-2)(h-a+2) + \binom{h-2a+4}{2} & 9a \leq h\,. \end{cases} \tag{14}$$

The connection between this function and $e(\mathcal{T}_5 \cup \mathcal{M} \cup \mathcal{I})$ is provided by the following lemma, whose proof we defer to Section 7.

Lemma 17. *There exists κ_0 such that the following holds. Let H be a graph of order $h \geq \kappa_0$. Suppose that A is a subset of $V(H)$, with $3 \leq |A| \leq h/2$ and the property that there is no set of vertex-disjoint triangles in H which covers three or more vertices of A. Then $e(H) \leq p(h, |A|)$.*

Putting this lemma together with Lemma 16 allows us to strengthen the bound (7). This is the missing ingredient for the proof of Theorem 6.

Lemma 18. *There exists κ_0 such that the following holds. Provided that $e(\mathcal{T}_4) \geq 8\binom{t_4}{2} + 10t_4 - 27$ and $t_4 \geq \max\left(176, \kappa_0, \frac{2m+i}{3}\right)$ we have*

$$e(\mathcal{T}_5 \cup \mathcal{M} \cup \mathcal{I}) \leq p(3t_5 + 2m + i, 2m + i)\,.$$

Proof. Suppose that $e(\mathcal{T}_4) \geq 8\binom{t_4}{2} + 10t_4 - 27$ and $t_4 \geq \max(176, \kappa_0)$. By Lemma 16 there is no set of vertex-disjoint triangles induced by $V(\mathcal{T}_5 \cup \mathcal{M} \cup \mathcal{I})$ which covers three or more vertices of $\mathcal{M} \cup \mathcal{I}$. We then apply Lemma 17 to $G[\mathcal{T}_5 \cup \mathcal{M} \cup \mathcal{I}]$, with the partition into $\mathcal{T}_5$ and $\mathcal{M} \cup \mathcal{I}$. We conclude that the number of edges in this graph is at most $p(3t_5 + 2m + i, 2m + i)$ as desired. □

6. Proof of Theorem 6

We are now in a position to prove Theorem 6. The basic idea is the same as for the proof of Lemma 9. We assume Setup 7, and put together our various upper bounds on edges between parts to obtain a function of six variables (the sizes of the six parts) which upper bounds the number of edges in G. We then show that this function is maximised, subject to the constraints $t_1+t_2+t_3+t_4=k$ and $2m+i=n-3k$, by $e\big(E_i(n,k)\big)$ for some $i\in[4]$.

A small problem with this strategy is that Lemma 18, which we would like to use to provide one of our upper bounds, only applies if $e(\mathcal{T}_4)\geq 8\binom{t_4}{2}+10t_4-27$. We therefore have to handle the case that $e(\mathcal{T}_4)\leq 8\binom{t_4}{2}+10t_4-28$ separately. We need to define a function, which we obtain as follows. Summing the bounds in Lemmas 10 and 12, together with the assumption $e(\mathcal{T}_4)\leq 8\binom{t_4}{2}+10t_4-28$, we see that the following function bounds above $e(G)$.

$$\begin{aligned}g_s(t_1,t_2,t_3,t_4,m,i) :=& f(t_1,t_2,t_3,t_4,m,i)+im+m^2\\ &+(3+3m)t_4+(2+i)t_4+8\binom{t_4}{2}+10t_4-28\,.\end{aligned} \tag{15}$$

The maximisation of $g_s(t_1,t_2,t_3,t_4,m,i)$ subject to $t_1+t_2+t_3+t_4=k$ and $2m+i=n-3k$ is a matter of calculation which we defer to Appendix A.

Lemma 19. *If $n\geq 8406$ and $(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota)\in F(n,k)$ then*

$$g_s(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota)\leq\max_{j\in[4]}e\big(E_j(n,k)\big)\,.$$

The final function, g_ℓ, that we need to define, which we will show bounds above $e(G)$ provided that $e(\mathcal{T}_4)\geq 8\binom{t_4}{2}+10t_4-27$, is a little more complicated. Its definition is as follows.

If $t_4<\max\big(176,\kappa_0,\frac{2m+i}{3}\big)$, then $g_\ell(t_1,t_2,t_3,t_4,m,i)$ is defined by

$$g_\ell := f(t_1,t_2,t_3,t_4,m,i)+im+m^2+(3+3m)t_4+(2+i)t_4+\binom{3t_4}{2}\,. \tag{16}$$

If $t_4\geq\max\big(176,\kappa_0,\frac{2m+i}{3}\big)$ and $t_1\neq 1$ then we set

$$g_\ell(t_1,t_2,t_3,t_4,m,i) := f(t_1,t_2,t_3,t_4,m,i)+p(3t_4+2m+i,2m+i)\,. \tag{17}$$

If $t_4\geq\max\big(176,\kappa_0,\frac{2m+i}{3}\big)$ and $t_1=1$ then we set

$$g_\ell(t_1,t_2,t_3,t_4,m,i) := f(t_1,t_2,t_3,t_4,m,i)+p(3t_4+2m+i,2m+i)+20\,. \tag{18}$$

The following lemma, whose proof we defer to Appendix A, states that g_ℓ is upper bounded as desired.

Lemma 20. *Let $n\geq\max(4\cdot 10^4,900\kappa_0)$ and $k\in\mathbb{N}$ be given. If $n\leq 5k+8$, we have*

$$\max_{(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota)\in F(n,k)} g_\ell(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota)\leq\max_{j\in[4]}e\big(E_j(n,k)\big)\,.$$

The proof of Theorem 6 now amounts to verification that the functions g_s and g_ℓ indeed upper bound $e(G)$ as required.

Proof of Theorem 6. Given n and k, let G be an n-vertex graph which does not contain $(k+1)\times K_3$. Further, assume that $n \geq \max(4\cdot 10^4, 900\kappa_0)$. We assume G is decomposed as in Setup 7.

If $n > 5k+8$, then by Lemma 9 we have

$$e(G) \leq e\big(E_1(n,k)\big)\,,$$

so we may now assume that $n \leq 5k+8$.

If $e(\mathcal{T}_4) \leq 8\binom{t_4}{2} + 10t_4 - 28$, then our situation is exactly as in (15), i.e., by Lemma 19 we have

$$e(G) \leq g_s(t_1,t_2,t_3,t_4,m,i) \leq \max_{j\in[4]} e\big(E_j(n,k)\big)\ ,$$

which completes the proof in this case.

If on the other hand $e(\mathcal{T}_4) \geq 8\binom{t_4}{2} + 10t_4 - 27$, we have the following fact.

Claim 6.1. If $e(\mathcal{T}_4) \geq 8\binom{t_4}{2} + 10t_4 - 27$ then there exist $c_1, c_2 \in \{0,1\}$ such that we have

$$e(G) \leq g_\ell(t_1-c_1, t_2-c_2, t_3, t_4+c_1+c_2, m, i)\,.$$

Furthermore, we have $t_1 - c_1 \geq 0$ and $t_2 - c_2 \geq 0$.

Proof of Claim 6.1. We distinguish five cases.

Case 1: $t_4 < \max\big(176, \kappa_0, \frac{2m+i}{3}\big)$.

We take $c_1 = c_2 := 0$, and sum the bounds (5) and (a)–(c) of Lemma 10 together with the trivial bound $e(\mathcal{T}_4) \leq \binom{3t_4}{2}$. We obtain

$$e(G) \leq f(t_1,t_2,t_3,t_4,m,i) + im + m^2 + (3+3m)t_4 + (2+i)t_4 + \binom{3t_4}{2}$$

and so by (16) we have $e(G) \leq g_\ell(t_1,t_2,t_3,t_4,m,i)$.

Case 2: $t_4 \geq \max\big(176, \kappa_0, \frac{2m+i}{3}\big)$, $e(\mathcal{T}_1,\mathcal{T}_4) \leq 7t_1t_4 + 18$ and $e(\mathcal{T}_2,\mathcal{T}_4) \leq 8t_2t_4$.

We take again $c_1 = c_2 := 0$. By definition we have $\mathcal{T}_5 = \mathcal{T}_4$. We sum the bounds (d)–(i) of Lemma 10, the bounds (l) and the $j=3$ cases of (m) and (n) of Lemma 12, the bound $e(\mathcal{T}_4\cup\mathcal{M}\cup\mathcal{I}) \leq p(3t_4+2m+i, 2m+i)$ from Lemma 18 and the assumed $e(\mathcal{T}_2,\mathcal{T}_4) \leq 8t_2t_4$. These bounds cover all the edges of G except $e(\mathcal{T}_1,\mathcal{T}_4)$, and we have

$$e(G) - e(\mathcal{T}_1,\mathcal{T}_4) \leq f(t_1,t_2,t_3,t_4,m,i) + p(3t_4+2m+i, 2m+i) - 7t_1t_4\,.$$

If $t_1 \neq 1$, then Lemma 11 part (j) gives us that $e(\mathcal{T}_1,\mathcal{T}_4) \leq 7t_1t_4$, and we obtain

$$e(G) \leq f(t_1,t_2,t_3,t_4,m,i) + p(3t_4+2m+i, 2m+i)\ ,$$

which is in correspondence with (17). If $t_1 = 1$, then the assumed $e(\mathcal{T}_1,\mathcal{T}_4) \leq 7t_1t_4 + 18$ gives us

$$e(G) \leq f(t_1,t_2,t_3,t_4,m,i) + p(3t_4+2m+i, 2m+i) + 18\ ,$$

and by (18) we have $e(G) \leq g_\ell(t_1, t_2, t_3, t_4, m, i)$.

Case 3: $t_4 \geq \max\left(176, \kappa_0, \frac{2m+i}{3}\right)$, $e(\mathcal{T}_1, \mathcal{T}_4) \leq 7t_1t_4 + 18$ and $e(\mathcal{T}_2, \mathcal{T}_4) > 8t_2t_4$.

We take $c_1 := 0$ and $c_2 := 1$. Observe that by Lemma 11 part (k) $e(\mathcal{T}_2, \mathcal{T}_4) > 8t_2t_4$ implies that $t_2 = 1$. By definition of $\mathcal{T}_5$ we have $\mathcal{T}_5 = \mathcal{T}_4 \cup \mathcal{T}_2$, and by Lemma 18 we have $e(\mathcal{T}_5 \cup \mathcal{M} \cup \mathcal{I}) \leq p(3t_4 + 2m + i + 3, 2m + i)$.

We use the bounds in parts (d)–(f) and (i) of Lemma 10 and the $j = 3$ cases of parts (m) and (n) of Lemma 12. Together with the above bound on $e(\mathcal{T}_5 \cup \mathcal{M} \cup \mathcal{I})$, these bounds cover all the edges of G except $e(\mathcal{T}_1, \mathcal{T}_2 \cup \mathcal{T}_4)$, and we have

$$\begin{aligned} e(G) - e(\mathcal{T}_1, \mathcal{T}_2 \cup \mathcal{T}_4) \leq f(t_1, 0, t_3, t_4, m, i) \\ + p(3t_4 + 2m + i + 3, 2m + i) - 7t_1t_4 + 8t_2t_3 \,. \end{aligned} \tag{19}$$

If $t_1 \neq 1$, then the $j = 2$ and $j = 4$ cases of Lemma 11 part (j) yield $e(\mathcal{T}_1, \mathcal{T}_2 \cup \mathcal{T}_4) \leq 7t_1(t_2 + t_4) = 7t_1t_5$, and we obtain from (19) that

$$e(G) \leq f(t_1, 0, t_3, t_5, m, i) + p(3t_4 + 2m + i + 3, 2m + i)\,.$$

By (17) we have $e(G) \leq g_\ell(t_1, 0, t_3, t_4 + 1, m, i)$. If $t_1 = 1$, then we use instead the trivial $e(\mathcal{T}_1, \mathcal{T}_2) \leq 9$ and the assumed $e(\mathcal{T}_1, \mathcal{T}_4) \leq 7t_1t_4 + 18$ to obtain $e(\mathcal{T}_1, \mathcal{T}_2 \cup \mathcal{T}_4) \leq 7t_1t_5 + 20$, and hence

$$e(G) \leq f(t_1, 0, t_3, t_5, m, i) + p(3t_4 + 2m + i + 3, 2m + i) + 20\,,$$

and by (18) we have $e(G) \leq g_\ell(t_1, 0, t_3, t_4 + 1, m, i)$.

Case 4: $t_4 \geq \max\left(176, \kappa_0, \frac{2m+i}{3}\right)$, $e(\mathcal{T}_1, \mathcal{T}_4) > 7t_1t_4 + 18$ and $e(\mathcal{T}_2, \mathcal{T}_4) \leq 8t_2t_4$.

We take $c_1 := 1$ and $c_2 := 0$. By Lemma 11 part (j) we have $t_1 = 1$. Thus we have $\mathcal{T}_5 = \mathcal{T}_4 \cup \mathcal{T}_1$. Summing the bounds in parts (g)–(i) of Lemma 10 and those in part (l) and the $j = 3$ cases of parts (m) and (n) of Lemma 12, together with the assumed $e(\mathcal{T}_2, \mathcal{T}_4) \leq 8t_2t_4$ and the bound $e(\mathcal{T}_5 \cup \mathcal{M} \cup \mathcal{I}) \leq p(3t_4 + 2m + i + 3, 2m + i)$ from Lemma 18, we obtain

$$e(G) \leq f(0, t_2, t_3, t_5, m, i) + p(3t_4 + 2m + i + 3, 2m + i) + 20\,.$$

By (18) we have $e(G) \leq g_\ell(0, t_2, t_3, t_4 + 1, m, i)$.

Case 5: $t_4 \geq \max\left(176, \frac{2m+i}{3}\right)$, $e(\mathcal{T}_1, \mathcal{T}_4) > 7t_1t_4 + 18$ and $e(\mathcal{T}_2, \mathcal{T}_4) > 8t_2t_4$.

We take $c_1 = c_2 = 1$. By Lemma 11 parts (j) and (k) we have $t_1 = t_2 = 1$, and thus we have $\mathcal{T}_5 = \mathcal{T}_1 \cup \mathcal{T}_2 \cup \mathcal{T}_4$. Summing the bounds in part (i) of Lemma 10 and the $j = 3$ cases of parts (m) and (n) of Lemma 12, together with the bound $e(\mathcal{T}_5 \cup \mathcal{M} \cup \mathcal{I}) \leq p(3t_4 + 2m + i + 6, 2m + i)$ from Lemma 18, we obtain

$$e(G) \leq f(0, 0, t_3, t_5, m, i) + p(3t_4 + 2m + i + 6, 2m + i)\,.$$

By (18) we have $e(G) \leq g_\ell(0, 0, t_3, t_4 + 2, m, i)$. □

Observe that for any n-vertex, $(k + 1) \times K_3$-free graph G decomposed as in Setup 7 we have $(t_1, t_2, t_3, t_4, m, i) \in F(n, k)$. By Claim 6.1 there are $c_1, c_2 \in \{0, 1\}$ such that $(t_1 - c_1, t_2 - c_2, t_3, t_4 + c_1 + c_2, m, i) \in F(n, k)$ and such that $e(G) \le g_\ell(t_1 - c_1, t_2 - c_2, t_3, t_4 + c_1 + c_2, m, i)$. By Lemma 20 we thus obtain

$$e(G) \le \max_{j\in[4]} e\big(E_j(n,k)\big)\,,$$

as desired. □

7. Graphs with few triangles touching a given set

In this section, we prove Lemma 17. The extremal problem of that lemma is not a very natural one. Also, we remark that Lemma 17 is sharp only when $h \ge 9a$. This is the regime in which we need the exact answer.

However the closely related extremal problem of bounding the number of edges in a graph H on h vertices with no triangle touching a given set $A \subseteq V(H)$ of size a is quite natural. We already studied it in two previous papers [ABHP13, ABHP], where we (respectively) determined the extremal function and proved uniqueness and stability for the problem. We need a special case of the extremal result of [ABHP13].

Theorem 21. *Let H be a graph on h vertices, and A be a subset of $V(H)$ of size $a \le \frac{h}{2}$ such that no triangle of H intersects A. Then we have*

$$e(H) \le \binom{h-2a}{2} + a(h-a)\,.$$

We also need a stability version of this theorem, proved in [ABHP]. To this end, we consider the following family $\mathcal{H}_A$ of graphs on the vertex set $[h]$ and with a distinguished set $A \subseteq [h]$, $|A| = a$ which show the optimality of Theorem 21. To construct one graph in $\mathcal{H}_A$, we take any set $B \subseteq [h]$ of size a disjoint from A, put a complete balanced bipartite graph on $A \cup B$ (where the parts of the bipartite graph may be any partition of $A \cup B$) and make all the vertices of $[h] \setminus (A \cup B)$ adjacent to each other and to all the vertices of B.

Theorem 22 ([ABHP])**.** *For every $\varepsilon > 0$ there exist $\gamma > 0$ and h_0 such that the following holds. Let H be a graph of order $h \ge h_0$ and A be a subset of $V(H)$ of size $a \le h/2$ such that no triangle of H intersects A. Suppose furthermore that $e(H) \ge \binom{h-2a}{2} + a(h-a) - \gamma h^2$. Then by editing at most εh^2 pairs in $\binom{V(H)}{2}$ we can obtain a graph in $\mathcal{H}_A$ (without changing the vertices of A).*

We will now show how this implies Lemma 17.

Proof of Lemma 17. We set $\varepsilon = 1/400$ and let $\gamma > 0$ and h_0 be given by Theorem 22. We set

$$\kappa_0 = \max\left(10\gamma^{-1}, h_0, 8000\right). \tag{20}$$

Suppose that $h \ge \kappa_0$ and H is an h-vertex graph and A is a set of a vertices such that no set of vertex disjoint triangles of H covers more than two vertices of A. This implies that we can identify a set of at most two vertex-disjoint triangles covering a maximum number of vertices of A. Taking the vertices of these triangles, and adding further arbitrary vertices if necessary, we obtain a set U of six vertices, with $|A \cap U| = 2$, such that $H - U$ has no triangle intersecting $A \setminus U$. Removing all the edges of H with one or two endpoints in U therefore yields a graph H' in which no triangle intersects A. By Theorem 21, H' has at most $\binom{h-2a}{2} + a(h-a)$ edges.

There are two cases to deal with, corresponding to the two possibilities in the definition of $p(h,a)$ in (14). The easier case is $2a \le h < 9a$, where we do not attempt to prove a sharp extremal result. Since at most $6h$ edges were removed from H to obtain H', we have $e(H) \le \binom{h-2a}{2} + a(h-a) + 6h = p(h,a)$, which completes the proof in this case.

We now turn to the case $3 \le a \le \frac{h}{9}$. Again, if $e(H') < p(h,a) - 6h$ then $e(H) < p(h,a)$ and we are done. So since $p(h,a) \ge a(h-a) - 2(h+2) + \binom{h-2a}{2}$ we may assume

$$e(H') > p(h,a) - 6h \overset{(20)}{\ge} \binom{h-2a}{2} + a(h-a) - \gamma h^2 \,.$$

By Theorem 22 we can edit at most εh^2 pairs in $\binom{V(H')}{2}$ we obtain an extremal graph $G \in \mathcal{H}_A$ on $V(H)$ with no triangle intersecting A. Recall that G consists of a complete balanced bipartite graph on a set of $2a$ vertices $A \cup B$ (where A and B are not necessarily the partition classes). The remaining vertices form a clique, and all the edges between them and B are present. It is easy to check that since $|A| = |B| = a \le \frac{h}{9}$, any set of $\frac{2h}{9}$ vertices of G induces at least

$$\binom{\frac{h}{9}}{2} \overset{(20)}{\ge} \frac{h^2}{200}$$

edges of G. Since G was obtained from H' by editing at most εh^2 pairs in $\binom{V(H')}{2}$, and H' was obtained from H by deleting edges, it follows that any set of $\frac{2h}{9}$ vertices of H induce at least $\left(\frac{1}{200} - \varepsilon\right)h^2 = \frac{1}{400}h^2$ edges of H.

We claim that this implies that any set C of $\frac{2h}{9}$ vertices of H contains a matching with at least seven edges. Indeed, we can find such a matching greedily, and after removing from C at most 6 matching edges and all edges incident to these matching edges, we removed at most $12 \cdot \frac{2h}{9} + 6 < \frac{1}{400}h^2$ edges from C.

Let $A \cap U = \{v_1, v_2\}$ and recall that $e(H') = e(H' - U)$. Theorem 21 applied to the graph $H' - U$ on $h - 6$ vertices and the set $A \setminus U$ with $a - 2$ vertices (it is indeed possible to apply Theorem 21 because $a - 2 \le \frac{1}{2}(h-6)$ by (20) and by $a \le \frac{h}{9}$) gives

$$e(H') \le \binom{h-6-2(a-2)}{2} + (a-2)(h-6-(a-2)) \,.$$

Observe that if $\deg_H(v_1) + \deg_H(v_2) \leq 2h - 6a - 9$ then we have

$$\begin{aligned} e(H) &\leq e(H') + \deg_H(v_1) + \deg_H(v_2) + 4h \\ &\leq \binom{h-6-2(a-2)}{2} + (a-2)\big(h-6-(a-2)\big) + 2h - 6a - 9 + 4h \\ &= p(h,a)\,, \end{aligned} \tag{21}$$

and we are done. We may therefore assume $\deg_H(v_1) + \deg_H(v_2) > 2h - 6a - 9$, and since $\deg_H(v_i) \leq h-1$ for $i \in [2]$ it follows that $\deg_H(v_i) \geq h - 6a - 8 \geq \frac{2h}{9}$, where the final inequality follows from $a \leq \frac{h}{9}$ and (20).

Since any set of $\frac{2h}{9}$ vertices of H contains a matching with at least seven edges, we conclude that $N_H(v_i)$ contains such a matching M_i for $i \in [2]$. Observe that if there were a triangle xyz in H with $x \in A \setminus \{v_1, v_2\}$ then we could use M_1 and M_2 to find greedily a collection of vertex-disjoint triangles in H covering $\{x, v_1, v_2\}$. This is a contradiction to the assumption on H that no such collection exists, and we conclude that there is no triangle of H which intersects $A' := A \setminus \{v_1, v_2\}$.

To complete the proof, we will now show that this final condition that no triangle of H intersects A' implies that $e(H) \leq p(h,a)$. Let $\prec$ be a linear order of the vertices of H. We apply the following 'vertex duplication' operation successively. If there are non-adjacent vertices u_1, u_2 in A' such that either $\deg_H(u_1) < \deg_H(u_2)$, or $\deg_H(u_1) = \deg_H(u_2)$ and $u_1 \prec u_2$, then we change H by resetting the neighbourhood of u_1 to $N_H(u_2)$. Let H'' be the graph obtained by repeatedly applying this operation until every pair of non-adjacent vertices of A' has identical neighbourhoods.

By construction, we have $e(H'') \geq e(H)$, and no triangle in H'' intersects A'. Now $H''[A']$ is a complete partite graph, and since $H''[A']$ contains no triangles it is a complete bipartite graph. Let its parts be Y_3 and Y_4 (the latter of which may have size zero). Moreover, all vertices $y \in Y_3$ have the identical neighbourhood $N_{H''}(y) \setminus A' =: Y_1$. Likewise, all vertices $y \in Y_4$ have the identical neighbourhood $N_{H''}(y) \setminus A' =: Y_2$. If $Y_4 = \emptyset$ then we set $Y_2 = \emptyset$. Since no triangle of H'' intersects A' the sets Y_1 and Y_2 are disjoint independent sets in H''. Finally, let X be the remaining vertices of H''. We have

$$\begin{aligned} e(H) \leq e(H'') &\leq \binom{|X|}{2} + \big(|Y_1| + |Y_2|\big)\big(|X| + |Y_3| + |Y_4|\big) \\ &= \binom{h-a+2-s}{2} + s(h-s)\,, \end{aligned}$$

where $s := |Y_1| + |Y_2|$. This function is maximised by $s = a - \frac{3}{2}$, and the maximum with s an integer occurs at $s = a-1, a-2$, where the function is precisely equal to $p(h,a)$. We conclude that $e(H) \leq p(h,a)$ as desired. □

8. Concluding remarks

Small values of n. We did not try to optimise our arguments in order to reduce n_0. Indeed, the value we obtain depends on the relation between ε and γ provided in Theorem 22, and the proof of that result in [ABHP] makes use of the Stability Theorem of Erdős and Simonovits [Erd68, Sim68] for triangles. But there is no 'heavy machinery' involved which would cause n_0 to become very large. It seems very likely that tracing exact values through these results would lead to a value of n_0 here smaller than 10^{10}. Perhaps Theorem 6 even holds with $n_0 = 1$, but we did not spend much effort on trying to find counterexamples for small values of n. Certainly our proof will not give such a result even with optimisation.

Tilings with larger cliques. It would be natural to ask for an extension of Theorem 6 to $(k+1)\times K_r$-free graphs G rather than $(k+1)\times K_3$-free graphs, thus obtaining a density version of the Hajnal–Szemerédi Theorem [HS70] rather than the Corrádi–Hajnal Theorem. The same basic approach as in our proof of Theorem 6 seems to be a reasonable strategy for proving such a result: We call a family $(\mathcal{K}_r, \mathcal{K}_{r-1}, \dots, \mathcal{K}_1)$ an *r-tiling family* if $\mathcal{K}_i$ is a collection of disjoint copies of the clique K_i inside G, and the sets $\mathcal{K}_r, \mathcal{K}_{r+1}, \dots, \mathcal{K}_1$ partition the vertices of G. We then consider an r-tiling family which maximises the vector $(|\mathcal{K}_r|, |\mathcal{K}_{r+1}|, \dots, |\mathcal{K}_1|)$ in lexicographic order, and try to work out bounds on the edge counts inside the sets $\mathcal{K}_i$ and between $\mathcal{K}_i$ and $\mathcal{K}_j$, relying again on rotation techniques. Some parts of such an argument can be made to work, but there are some additional difficulties for $r \geq 4$ that do not appear for $r = 3$. We are not even sure what the complete family of extremal graphs should be.

Tilings with more general graphs. An extension of Theorem 6 which seems within the reach of existing techniques is to get asymptotically tight bounds on the size of a maximal H-tiling (as a function of the density of the host graph) for any three-colourable graph H. The bipartite counterpart for this is the extension of Theorem 3 by Grosu and Hladký [GH12]. These problems can also be seen as density versions of Komlós's extension [Kom00] of the Hajnal–Szemerédi Theorem to general graphs. It seems likely that the technique developed by Komlós, and adapted to this setting by Grosu and Hladký, is flexible enough to allow such a generalisation for H-tiling with any fixed 3-colourable graph H, and that the extremal graphs for the problem of H-tiling in a graph of a given density will resemble the graphs $E_1, \dots, E_4$ from Definition 5, though the part sizes will not be the same as in that definition.

9. Acknowledgement

The paper was finalised during the participation at the program Graphs, Hypergraphs, and Computing at Institut Mittag–Leffler. We would like to thank the organisers and the staff of the institute for creating a very productive atmosphere. We had to take our little children with us. We would like to acknowledge the support of the London Mathematical Society, and Mathematics Institute at the University of Warwick (JH) for contributing to childcare expenses that were incurred during this trip, and Emili Simonovits for helping us with babysitting.

References

[ABHP] P. Allen, J. Böttcher, J. Hladký, and D. Piguet. An extension of Turán's Theorem, uniqueness, and stability. Preprint available via arXiv.

[ABHP13] P. Allen, J. Böttcher, J. Hladký, and D. Piguet. Turánnical hypergraphs. *Random Structures Algorithms*, 42(1):29–58, 2013.

[CH63] K. Corrádi and A. Hajnal. On the maximal number of independent circuits in a graph. *Acta Math. Acad. Sci. Hungar.*, 14:423–439, 1963.

[EG59] P. Erdős and T. Gallai. On maximal paths and circuits of graphs. *Acta Math. Acad. Sci. Hungar*, 10:337–356 (unbound insert), 1959.

[Erd62a] P. Erdős. On a theorem of Rademacher-Turán. *Illinois Journal of Mathematics*, 6:122–127, 1962.

[Erd62b] P. Erdős. Über ein Extremalproblem in der Graphentheorie. *Arch. Math.*, 13:222–227, 1962.

[Erd68] P. Erdős. On some new inequalities concerning extremal properties of graphs. In *Theory of Graphs (Proc. Colloq., Tihany, 1966)*, pages 77–81. Academic Press, New York, 1968.

[GH12] C. Grosu and J. Hladký. The extremal function for partial bipartite tilings. *European J. Combin.*, 33(5):807–815, 2012.

[Győ91] E. Győri. On the number of edge disjoint cliques in graphs of given size. *Combinatorica*, 11(3):231–243, 1991.

[HS70] A. Hajnal and E. Szemerédi. Proof of a conjecture of P. Erdős. In *Combinatorial theory and its applications, II (Proc. Colloq., Balatonfüred, 1969)*, pages 601–623. North-Holland, Amsterdam, 1970.

[Kom00] J. Komlós. Tiling Turán theorems. *Combinatorica*, 20(2):203–218, 2000.

[LS83] L. Lovász and M. Simonovits. On the number of complete subgraphs of a graph. II. In *Studies in pure mathematics*, pages 459–495. Birkhäuser, Basel, 1983.

[Moo68] J. W. Moon. On independent complete subgraphs in a graph. *Canad. J. Math.*, 20:95–102, 1968.

[Raz08] A. A. Razborov. On the minimal density of triangles in graphs. *Combin. Probab. Comput.*, 17(4):603–618, 2008.

[Sim68] M. Simonovits. A method for solving extremal problems in graph theory, stability problems. In *Theory of Graphs (Proc. Colloq., Tihany, 1966)*, pages 279–319. Academic Press, New York, 1968.

Appendix A. Maximisations

In this section we provide proofs of Lemmas 13, 19, and 20. These lemmas concern maximisations of certain functions. We build our arguments on tedious elementary algebraic manipulations. While some of the statements we need could be obtained by a more routine technique of Lagrange multipliers, this method seems to lead to even lengthier calculations in our setting. This is caused in particular by a high degree of discontinuity, caused by various case distinctions and appearance of the floor/ceiling function, of the functions we want to maximise.

We first collect some useful statements relating to f, all of which are obtained by simple calculation using equations (3) and (4). The three relations (22)–(24) below hold for any $\tau_1, \dots, \tau_4, \mu, \iota \geq 0$.

$$\begin{aligned} f(\tau_1 + x, &\tau_2 - x, \tau_3, \tau_4, \mu, \iota) - f(\tau_1, \tau_2, \tau_3, \tau_4, \mu, \iota) \\ &= \frac{x^2}{2} + \big(\mu - \tau_2 - \tau_3 - \tau_4 + \tfrac{1}{2}\big)x - \begin{cases} 0 & \mu = 0 \\ 2x & \mu > 0 \end{cases} \end{aligned} \tag{22}$$

$$f(\tau_1, \tau_2 + \tau_3, 0, \tau_4, \mu, \iota) - f(\tau_1, \tau_2, \tau_3, \tau_4, \mu, \iota) \geq (\iota - 3)\tau_3 \tag{23}$$

$$\begin{aligned} f(\tau_1, \tau_2, \tau_3, \tau_4, \mu, \iota) \leq 8\binom{\tau_1 + \tau_2 + \tau_3 + \tau_4}{2} &- 8\binom{\tau_4}{2} \\ &+ (4\mu + 2\iota + 6)(\tau_1 + \tau_2 + \tau_3) - \tau_1\tau_4 \end{aligned} \tag{24}$$

Provided that $\min\{\mu, \mu - x, \iota + 2x\} \geq 1$, $\iota \geq 0$, and $x \geq 0$ we have

$$f(\tau_1, \tau_2, \tau_3, \tau_4, \mu - x, \iota + 2x) - f(\tau_1, \tau_2, \tau_3, \tau_4, \mu, \iota) \geq x(\tau_2 - \tau_3)\,. \tag{25}$$

If $\mu \geq 5$, combining (23) and (25) we have

$$\begin{aligned} f(\tau_1, \tau_2 + \tau_3, 0, \tau_4, \mu - 4, \iota + 8) &- f(\tau_1, \tau_2, \tau_3, \tau_4, \mu, \iota) \\ &\geq 4(\tau_2 - \tau_3) + (\iota + 5)\tau_3 \geq 0\,. \end{aligned} \tag{26}$$

We will use the following lemma in our later maximisation results. Observe that Lemma 13 is part (*iii*) of this lemma.

Lemma 23. *Given non-negative integers $\tau_1, \tau_2, \tau_3, \tau_4, \mu, \iota$ the following are true.*

(*i*) *We have*

$$f(\tau_1, \tau_2, \tau_3, \tau_4, \mu, \iota) \leq \max\big(f(\tau_1 + \tau_2, 0, \tau_3, \tau_4, \mu, \iota), f(0, \tau_1 + \tau_2, \tau_3, \tau_4, \mu, \iota)\big)$$

with equality only if either $\tau_1 = 0$ or $\tau_2 = 0$.

(*ii*) *When $\iota \geq 4$ we have*

$$f(\tau_1, \tau_2, \tau_3, \tau_4, \mu, \iota) \leq f(\tau_1, \tau_2 + \tau_3, 0, \tau_4, \mu, \iota)$$

with equality only if $\tau_3 = 0$.

(*iii*) *When* $n \geq 3k+2$ *we have*

$$\max_{(\tau_1,\tau_2,0,0,\mu,\iota)\in F(n,k)} \big(f(\tau_1,\tau_2,0,0,\mu,\iota)+\iota\mu+\mu^2\big) = \max_{j\in[3]} e\big(E_j(n,k)\big)\,.$$

(*iv*) *When* $n \geq 3k+21$ *we have*

$$\max_{(\tau_1,\tau_2,\tau_3,0,\mu,\iota)\in F(n,k)} f(\tau_1,\tau_2,\tau_3,0,\mu,\iota)+\iota\mu+\mu^2 = \max_{j\in[3]} e\big(E_j(n,k)\big)\,.$$

Proof of Lemma 23. Proof of part (*i*)*:* By (22),

$$f(\tau_1+x,\tau_2-x,\tau_3,\tau_4,\mu,\iota)-f(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota)$$

is a quadratic in x with positive x^2-coefficient. It follows that for any $a \leq b$, the maximum of $f(\tau_1+x,\tau_2-x,\tau_3,\tau_4,\mu,\iota)-f(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota)$ over $a \leq x \leq b$ occurs when either $x=a$ or $x=b$. In particular, we have for all non-negative $\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota$ that

$$f(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota) \leq \max\big(f(\tau_1+\tau_2,0,\tau_3,\tau_4,\mu,\iota), f(0,\tau_1+\tau_2,\tau_3,\tau_4,\mu,\iota)\big)\,,$$

with equality only when $\tau_1=0$ or $\tau_2=0$.

Proof of part (*ii*)*:* By (23), when $\iota \geq 4$, we have

$$f(\tau_1,\tau_2+\tau_3,0,\tau_4,\mu,\iota) \geq f(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota)\,,$$

with equality only when $\tau_3=0$.

Proof of part (*iii*)*:* By part (*i*) the maximum on the left-hand side is attained either when $\tau_1=k,\tau_2=0$, or when $\tau_1=0,\tau_2=k$.

By (3) and (4) we have

$$\begin{aligned}
f(k,0,0,0,\mu,\iota)+\iota\mu+\mu^2 &= 4\mu k+2\iota k+7\binom{k}{2}+3k+\iota\mu+\mu^2\\
&= 7\binom{k}{2}+3k+2(n-3k)k+\mu(n-3k-\mu)\\
&\leq 7\binom{k}{2}+3k+2(n-3k)k+\left\lfloor\frac{n-3k}{2}\right\rfloor\left\lceil\frac{n-3k}{2}\right\rceil\\
&= \binom{k}{2}+k(n-k)+\left\lfloor\frac{n-k}{2}\right\rfloor\left\lceil\frac{n-k}{2}\right\rceil\\
&= e\big(E_1(n,k)\big)\,,
\end{aligned} \tag{27}$$

where the last term on the second line achieves its maximum, $\lfloor\frac{n-3k}{2}\rfloor\lceil\frac{n-3k}{2}\rceil$, exactly when $\mu=n-3k-\mu$ and $\iota=0$, if $n-3k$ is even, or when $\mu = n-3k-\mu-1$ and $\iota=1$, if not (observe that we cannot have $\mu=n-3k-\mu+1$, since then we would have $\iota=-1$).

To deal with the term $\tau_1=0$, $\tau_2=k$ we have to distinguish between the cases $\mu=0$ and $\mu>0$.

When $\mu = 0$ we first observe that the case $k = 0$ trivially satisfies the statement. Thus we assume that $k > 0$. We have

$$\begin{aligned} f(0,k,0,0,0,n-3k) + 0 + 0 \\ &= 2(n-3k)k + 8\binom{k}{2} + 3k \\ &< 8\binom{k}{2} + 2(n-3k-2)k + 3k + 5k \\ &= f(0,k,0,0,1,n-3k-2) \\ &\le f(0,k,0,0,1,n-3k-2) + (n-3k-2)\times 1 + 1^2 . \end{aligned} \tag{28}$$

It follows that $f(0,k,0,0,\mu,\iota) + \iota\mu + \mu^2$ is not maximised on $F(n,k)$ when $\mu = 0$.

When $\mu \ge 1$, again from (3) and (4), we have

$$\begin{aligned} f(0,k,0,0,\mu,\iota) + \iota\mu + \mu^2 &= 2\iota k + 8\binom{k}{2} + 3k + (2+3\mu)k + \iota\mu + \mu^2 \\ &= 8\binom{k}{2} + 5k + 2(n-3k)k + \mu(n-\mu-4k) , \end{aligned} \tag{29}$$

which is maximised on $F(n,k)$ both when

$$\mu = \max\left(1, \left\lfloor \frac{n-4k}{2} \right\rfloor\right) \text{ and } \mu = \max\left(1, \left\lceil \frac{n-4k}{2} \right\rceil\right) .$$

It is straightforward from (3) and (4) to check that for the numbers $\mu_1 := \lfloor \frac{n-4k}{2} \rfloor, \iota_1 := n-3k-2\mu_1$, and $\mu_2 := \lceil \frac{n-4k}{2} \rceil, \iota_2 := n-3k-2\mu_2$ we have

$$\begin{aligned} f(0,k,0,0,\mu_1,\iota_1) + \iota_1\mu_1 + \mu_1^2 &= f(0,k,0,0,\mu_2,\iota_2) + \iota_2\mu_2 + \mu_2^2 \\ &= e\big(E_2(n,k)\big) . \end{aligned}$$

Further

$$f(0,k,0,0,1,n-3k-2) + (n-3k-2)\times 1 + 1^2 = e\big(E_3(n,k)\big) . \tag{30}$$

This completes the proof.

Proof of part (iv): Let $k \ge 0$ and $n \ge 3k+21$ be fixed. By part (*iii*) it is enough to show that the function $f(\tau_1,\tau_2,\tau_3,0,\mu,\iota) + \iota\mu + \mu^2$ is maximised on the set $F(n,k)$ only when $\tau_3 = 0$.

Let $(\tau_1,\tau_2,\tau_3,0,\mu,\iota) \in F(n,k)$. From part (*ii*) we have that if $\iota \ge 4$ then

$$\begin{aligned} f(\tau_1,\tau_2,\tau_3,0,\mu,\iota) \\ &\le \max\big(f(\tau_1+\tau_2+\tau_3,0,0,0,\mu,\iota), f(0,\tau_1+\tau_2+\tau_3,0,0,\mu,\iota)\big) . \end{aligned}$$

with equality only when $\tau_3 = 0$, as desired. In the rest of the proof we assume that $\iota \le 3$. Since $2\mu + \iota = n - 3k \ge 21$, we have $\mu \ge 9$.

We separate two cases. First, suppose that $\tau_2+\tau_3 \geq 13$. Using (25), since $\mu - 6 \geq 3 \geq 1$, we have

$$\begin{aligned} f(\tau_1, \tau_2+\tau_3, 0, 0, \mu-6, \iota+12) + (\iota+12)(\mu-6) + (\mu-6)^2 \\ - \big(f(\tau_1, \tau_2+\tau_3, 0, 0, \mu, \iota) + \iota\mu + \mu^2\big) \\ \geq 6(\tau_2+\tau_3) - 6\iota - 36 > (3-\iota)\tau_3 \,, \end{aligned}$$

since we have $\tau_2+\tau_3 \geq 13$. By (23) we obtain

$$\begin{aligned} f(\tau_1, \tau_2+\tau_3, 0, 0, \mu-6, \iota+12) + (\iota+12)(\mu-6) + (\mu-6)^2 \\ > f(\tau_1, \tau_2, \tau_3, 0, \mu, \iota) + \iota\mu + \mu^2 \,, \end{aligned}$$

as desired.

Second, suppose that $\tau_2+\tau_3 \leq 12$. We have

$$\begin{aligned} f(\tau_1+\tau_2+\tau_3, 0, 0, 0, \mu, \iota) - f(\tau_1, \tau_2+\tau_3, 0, 0, \mu, \iota) \\ \geq (\mu-2)(\tau_2+\tau_3) - \binom{\tau_2+\tau_3}{2} \\ > (\mu+\iota-11)(\tau_2+\tau_3) + (3-\iota)\tau_3 \\ \geq (3-\iota)\tau_3 \,, \end{aligned}$$

where the last inequality comes from $2\mu + \iota = n - 3k \geq 21$. Together with (23) we then have

$$f(\tau_1+\tau_2+\tau_3, 0, 0, 0, \mu, \iota) - f(\tau_1, \tau_2, \tau_3, 0, \mu, \iota) > 0 \,,$$

and hence that

$$f(\tau_1+\tau_2+\tau_3, 0, 0, 0, \mu, \iota) + \iota\mu + \mu^2 > f(\tau_1, \tau_2, \tau_3, 0, \mu, \iota) + \iota\mu + \mu^2 \,,$$

as desired. □

Proof of Lemma 19. Let $\bar{\tau}_1 := \tau_2+\tau_3+\tau_4$. We now provide some preliminary observations and then distinguish four cases to prove the lemma.

From (15) we have

$$g_s(\tau_1, \tau_2, \tau_3, \tau_4, \mu, \iota) - g_s(\tau_1, \bar{\tau}_1, 0, 0, \mu, \iota) \leq 9\tau_4 - (\tau_3+\tau_4)(\iota-3) \,. \tag{31}$$

Moreover, if $k \leq 43n/140$ then we have $2\mu + \iota = n - 3k \geq \frac{11n}{140}$. Hence

$$\mu \geq \frac{11n}{280} - 6 \qquad \text{if } k \leq 43n/140 \text{ and } \iota \leq 12 \,. \tag{32}$$

Case 1: $k > 43n/140$. We shall show that this implies

$$g_s(\tau_1, \tau_2, \tau_3, \tau_4, \mu, \iota) < e\big(E_4(n,k)\big) \,.$$

From (24) we have

$$\begin{aligned} &g_s(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota) \\ \le& 8\binom{\tau_1+\tau_2+\tau_3+\tau_4}{2} + (4\mu+2\iota+15)(\tau_1+\tau_2+\tau_3+\tau_4) - 28 + \iota\mu + \mu^2 \\ \le& 8\binom{k}{2} + 2k(n-3k) + 20k + \Big(\frac{n-3k}{2}\Big)^2, \end{aligned} \tag{33}$$

where the second inequality comes from the fact that $(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota) \in F(n,k)$.

Now solving the quadratic inequality (in variable k)

$$\begin{aligned} 8\binom{k}{2} + 20k + 2(n-3k)k + \Big(\frac{n-3k}{2}\Big)^2 \\ < \binom{6k-n+4}{2} + (6k-n+4)(n-3k-2) + (n-3k-2)^2 \\ = e\big(E_4(n,k)\big) \end{aligned}$$

shows that we have $g_s(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota) < e\big(E_4(n,k)\big)$ if k satisfies

$$k > \frac{n+2}{5} + \frac{\sqrt{14n^2+406n-84}}{35} = \frac{7+\sqrt{14}}{35}n + \frac{14+15\sqrt{14}}{35}. \tag{34}$$

Indeed, (34) is satisfied as $n \ge 8406$ and as $k > \frac{43n}{140}$. Hence we are done in this case.

Case 2: $\iota \ge 12$. Note that the right hand side of (31) is not positive for $\iota \ge 12$. Thus (31) implies $g_s(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota) \le g_s(\tau_1,\bar\tau_1,0,0,\mu,\iota) = f(\tau_1,\bar\tau_1,0,0,\mu,\iota) + \iota\mu + \mu^2 - 28$. Moreover, by Lemma 23 part (*iii*) the function $f(\tau_1,\bar\tau_1,0,0,\mu,\iota) + \iota\mu + \mu^2$, subject to the constraints $\tau_1 + \bar\tau_1 = k$ and $2\mu + \iota = n - 3k$, is bounded from above by $\max_{j\in[3]} e\big(E_j(n,k)\big)$. Hence $g_s(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota) \le \max_{j\in[3]} e\big(E_j(n,k)\big)$ as desired.

Case 3: $k \le 43n/140$, $\iota < 12$, $\bar\tau_1 > 80$. By (32) we have $\mu - 25 > 0$, so from (15) we obtain

$$\begin{aligned} &g_s(\tau_1,\bar\tau_1,0,0,\mu-25,\iota+50) - g_s(\tau_1,\bar\tau_1,0,0,\mu,\iota) \ge 25\bar\tau_1 - 25\iota - 25^2 \\ &\overset{\iota<12}{>} 25\bar\tau_1 - 25\cdot 37 \overset{\bar\tau_1>80}{>} 12\bar\tau_1 + 13\cdot 80 - 25\cdot 37 > 12\bar\tau_1 \ge 12(\tau_3+\tau_4). \end{aligned} \tag{35}$$

Since the right hand side of (31) is at most $12(\tau_3+\tau_4)$, this implies

$$\begin{aligned} g_s(\tau_1,\bar\tau_1,0,0,\mu-25,\iota+50) &> g_s(\tau_1,\bar\tau_1,0,0,\mu,\iota) + 12(\tau_3+\tau_4) \\ &\ge g_s(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota), \end{aligned}$$

and so we conclude from Lemma 23 part (*iii*) that

$$\begin{aligned} g_s(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota) &< g_s(\tau_1,\bar\tau_1,0,0,\mu-25,\iota+50) \\ &< f(\tau_1,\bar\tau_1,0,0,\mu-25,\iota+50) \\ &\quad + (\iota+50)(\mu-25)+(\mu-25)^2 \\ &\le \max_{j\in[3]} e\big(E_j(n,k)\big)\,. \end{aligned}$$

Case 4: $k \le 43n/140$, $\iota < 12$, $\bar\tau_1 \le 80$. Again from (15) we have

$$g_s(\tau_1+\bar\tau_1,0,0,0,\mu,\iota) - g_s(\tau_1,\bar\tau_1,0,0,\mu,\iota) \ge \bar\tau_1\mu - \binom{\bar\tau_1}{2} - 2\bar\tau_1\,. \qquad (36)$$

In addition, by (32) and since $n \ge 8406$ we have $\mu \ge \frac{11n}{280} - 6 > 320$. This implies

$$\bar\tau_1\mu - 2\bar\tau_1 - \binom{\bar\tau_1}{2} = \bar\tau_1\Big(\mu - \frac{\bar\tau_1 - 1}{2} - 2\Big) > \bar\tau_1\cdot 12 \ge 12(\tau_3+\tau_4)\,,$$

and hence we obtain using Lemma 23 part (*iii*) that

$$\begin{aligned} \max_{j\in[3]} e\big(E_j(n,k)\big) &\ge f(\tau_1+\bar\tau_1,0,0,0,\mu,\iota) + \iota\mu + \mu^2 \\ &> g_s(\tau_1+\bar\tau_1,0,0,0,\mu,\iota) \\ &> g_s(\tau_1,\bar\tau_1,0,0,\mu,\iota) + 12(\tau_3+\tau_4) \\ &\ge g_s(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota)\,, \end{aligned}$$

where, again, the last inequality follows from (31). □

Proof of Lemma 20. Our aim is to show that $g_\ell(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota)$ with all variables required to be non-negative integers and with $\tau_1+\tau_2+\tau_3+\tau_4 = k$ and $2\mu+\iota = n-3k$, is bounded above by

$$e(n,k) := \max_{i\in[4]} \big\{ e\big(E_1(n,k)\big),\ldots,e\big(E_4(n,k)\big)\big\}\,.$$

The main difficulty is to show that indeed if g_ℓ is maximised then $\tau_3 = 0$ and at most one of the variables τ_1,τ_2,τ_4 is non-zero, and as mentioned the reason why this seems not to be easy to automate is that g_ℓ is quite discontinuous. There are two regimes in which g_ℓ behaves quite differently. Furthermore, it is occasionally convenient to assume that $2\mu+\iota$ is reasonably large, leading to a third case.

The easier of the two regimes is when $3\tau_4 < 2\mu+\iota$. In this case, g_ℓ is defined by (16). This function is still not quite continuous: when μ or ι are changed from 0 to 1 or vice versa, there is discontinuity in the definition of the function f, but this turns out to be easy to handle.

Case 1: $3\tau_4 < n-3k$ and $n-3k \ge 30$.

We define the following auxiliary function.

$$h(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota) := 4\mu\tau_1 + 2\iota\tau_1 + 7\binom{\tau_1}{2} + 3\tau_1 + 2\iota\tau_2 + 8\binom{\tau_2}{2} + 3\tau_2 + 8\binom{\tau_3}{2} + 8\tau_3\tau_4 + 3\tau_3 + 7\tau_1\tau_2 + (2+3\mu)\tau_2 + 7\tau_1(\tau_3+\tau_4) + (3+3\mu)\tau_3 + 8\tau_2(\tau_3+\tau_4) + (2+\iota)\tau_3 + \iota\mu + \mu^2 + (3+3\mu)\tau_4 + (2+\iota)\tau_4 + \binom{3\tau_4}{2}. \tag{37}$$

Observe that this function is almost the same as g_ℓ: indeed, if $\mu,\iota \ge 1$ then they are equal, while otherwise g_ℓ is smaller and the difference is one of $2\tau_2 + 3\tau_3$, $2\tau_3$ and $2\tau_2 + 5\tau_3$ according to (4). We have the following equations (where we write h for $h(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota)$).

$$h(\tau_1+x,\tau_2,\tau_3,\tau_4-x,\mu,\iota) - h = x^2 + \mu x + \iota x - 4x - \tau_3 x - \tau_2 x - 2\tau_4 x \tag{38}$$

$$h(\tau_1,\tau_2+x,\tau_3,\tau_4-x,\mu,\iota) - h = \tfrac{x^2}{2} + \iota x - \tfrac{5}{2}x - \tau_4 x \tag{39}$$

$$h(\tau_1,\tau_2,\tau_3+x,\tau_4-x,\mu,\iota) - h = \tfrac{x^2}{2} + \tfrac{x}{2} - t_4 x \tag{40}$$

These are all quadratic in x with positive x^2 coefficients, and by the above observation the same statement is true (though the linear terms are different) when h is replaced with g_ℓ throughout. It follows that if g_ℓ is maximised, then either $\tau_4 = 0$ or $\tau_4 = \frac{n-3k}{3}$. We have the following equations:

$$h(\tau_1+x,\tau_2,\tau_3-x,\tau_4,\mu,\iota) - h = \tfrac{x^2}{2} + (\mu + \iota - \tau_2 - \tau_3 - \tau_4 - \tfrac{9}{2})x \tag{41}$$

$$h(\tau_1+x,\tau_2-x,\tau_3,\tau_4,\mu,\iota) - h = \tfrac{x^2}{2} + (\mu - \tau_2 - \tau_3 - \tau_4 - \tfrac{3}{2})x \tag{42}$$

$$h(\tau_1,\tau_2+x,\tau_3-x,\tau_4,\mu,\iota) - h = (\iota - 3)x \tag{43}$$

First we will consider the case $\tau_4 = \frac{n-3k}{3} > 0$. The equations (41) and (42) are positive quadratics in x, and the same is true replacing h with g_ℓ throughout. It follows that if g_ℓ is maximised and $\tau_1 > 0$ then $\tau_1 = 2k - \frac{n}{3}$ and $\tau_2 = \tau_3 = 0$. In this case we have $h = g_\ell$, and also

$$\begin{aligned} g_\ell(k,0,0,0,\mu,\iota) - g_\ell(2k-\tfrac{n}{3},0,0,\tfrac{n-3k}{3},\mu,\iota) &= \tfrac{n-3k}{3}\big(-\tfrac{n-3k}{3} + \mu + \iota - 4\big) \\ &= \tfrac{n-3k}{3}\big(\tfrac{n-3k}{6} + \tfrac{\iota}{2} - 4\big)\,, \end{aligned} \tag{44}$$

which, since $n - 3k \ge 30$, is positive. This is a contradiction to g_ℓ being maximised.

It remains to check the case $\tau_1 = 0$. By (43), we have $h \le h(\tau_1,\tau_2+\tau_3,0,\tau_4,\mu,\iota) + 3\tau_3$, and so the total difference between $h(\tau_1,\tau_2+\tau_3,\tau_4,\mu,\iota)$ and $g_\ell(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota)$ is at most $8k \le 3n$. Since we assume $\tau_1 = 0$ and $\tau_4 = \frac{n-3k}{3}$, and since $n \ge 10^4$, it is enough to show that $h(0,2k-\frac{n}{3},0,\frac{n-3k}{3},\mu,\iota)$ is always smaller than $e(n,k)$ by at least $\frac{1}{1000}n^2$. To simplify the analysis, we write $\approx$ to mean we discard all terms only linear in n. Together with the difference between h and g_ℓ, the linear error terms never amount to more than $6n < \frac{1}{1000}n^2$.

We have

$$h(0, 2k - \tfrac{n}{3}, 0, \tfrac{n-3k}{3}, \mu, \iota) = 3\iota k - \tfrac{\iota n}{3} + \tfrac{n^2}{18} - \tfrac{3k}{2} + \tfrac{5n}{6} + 3k\mu - \tfrac{kn}{3} + \tfrac{9k^2}{2} + \iota\mu + \mu^2 \,. \tag{45}$$

Discarding the linear terms and substituting $\iota = n - 3k - 2m$ we get

$$h(0, 2k - \tfrac{n}{3}, 0, \tfrac{n-3k}{3}, \mu, \iota) \approx \tfrac{11kn}{3} - \tfrac{9k^2}{2} - 6k\mu - \tfrac{5n^2}{18} + \tfrac{5\mu n}{3} - \mu^2 \,. \tag{46}$$

This function is a negative quadratic in μ with maximum at $\mu = \frac{5n}{6} - 3k$. Since we are only interested in the case that $\mu \geq 0$ and $\iota \geq 0$, we need to separate some subcases.

Subcase 1: $0.31n \leq k \leq n/3$ and $\mu = 0$. We have

$$h(0, 2k - \tfrac{n}{3}, 0, \tfrac{n-3k}{3}, 0, n - 3k) \approx \tfrac{11kn}{3} - \tfrac{9k^2}{2} - \tfrac{5n^2}{18} \text{ and} \tag{47}$$

$$h(0, 2k - \tfrac{n}{3}, 0, \tfrac{n-3k}{3}, 0, n - 3k) - e(E_4(n,k)) \approx -\tfrac{7n^2}{9} + \tfrac{20kn}{3} - \tfrac{27k^2}{2} \,. \tag{48}$$

This function is maximised at $k = \frac{20n}{81}$, and since $0.31 > \frac{20}{81}$ its maximum in the range $0.31n \leq k \leq n$ is at $k = 0.31n$, where the value attained is less than $-0.007n^2$.

Subcase 2: $\frac{5n}{18} \leq k < 0.31n$ and $\mu = 0$. We have

$$h(0, 2k - \tfrac{n}{3}, 0, \tfrac{n-3k}{3}, 0, n - 3k) - e(E_3(n,k)) \approx \tfrac{5kn}{3} - \tfrac{5k^2}{2} - \tfrac{5n^2}{18} \,, \tag{49}$$

which function is maximised at $k = n/3$ and hence in the range of k always smaller than the value at $k = 0.31n$, which is less than $-0.001n^2$.

Subcase 3: $\frac{2n}{9} \leq k \leq \frac{5n}{18}$, and $\mu = \frac{5n}{6} - 3k$. We have

$$h(0, 2k - \tfrac{n}{3}, 0, \tfrac{n-3k}{3}, \tfrac{5n}{6} - 3k, 3k - \tfrac{2n}{3}) \approx -\tfrac{4kn}{3} + \tfrac{9k^2}{2} + \tfrac{15n^2}{36}, \quad \text{and} \tag{50}$$

$$h(0, 2k - \tfrac{n}{3}, 0, \tfrac{n-3k}{3}, \tfrac{5n}{6} - 3k, 3k - \tfrac{2n}{3}) - e(E_3(n,k)) \approx -\tfrac{10kn}{3} + \tfrac{13k^2}{2} + \tfrac{15n^2}{36} \,, \tag{51}$$

which is a positive quadratic in k. At $k = \frac{5n}{18}$ the value of the LHS of (51) is $\frac{-1205n^2}{648}$, and at $k = \frac{2n}{9}$ we get $\frac{-481n^2}{324}$, the latter of which is the maximum in this range of k.

Subcase 4: $\frac{n}{5} < k < \frac{2n}{9}$ and $\iota = 0$. We get

$$h(0, 2k - \tfrac{n}{3}, 0, \tfrac{n-3k}{3}, \tfrac{n-3k}{2}, 0) \approx \tfrac{-kn}{3} + \tfrac{11n^2}{36} + \tfrac{9k^2}{4}, \quad \text{and} \tag{52}$$

$$h(0, 2k - \tfrac{n}{3}, 0, \tfrac{n-3k}{3}, \tfrac{n-3k}{2}, 0) - e(E_1(n,k)) \approx \tfrac{5k^2}{2} - \tfrac{5kn}{6} + \tfrac{n^2}{18} \,, \tag{53}$$

which is a positive quadratic in k. At $k = \frac{n}{5}$ we get $\frac{-n^2}{90}$, and at $k = \frac{2n}{9}$ the value is $\frac{-n^2}{162}$, the latter of which is the maximum in this range of k.

Since these subcases exhaust the range of k we are considering, we conclude that indeed if $\tau_1 = 0$ and $\tau_4 > 0$ then g_ℓ is not maximised. It follows that the only maxima of g_ℓ with $3\tau_4 < 2\mu + \iota$ are with $\tau_4 = 0$. By Lemma 23 part (iv), it follows that the maximum of $g_\ell(\tau_1, \tau_2, \tau_3, \tau_4, \mu, \iota)$ subject to the conditions $\tau_1 + \tau_2 + \tau_3 + \tau_4 = k$, $2\mu + \iota = n - 3k$, and $3\tau_4 < 2\mu + \iota$, is at most $e(n,k)$ as desired.

Case 2: $3\tau_4 \geq \max(528, 3\kappa_0, n-3k)$.

In this range g_ℓ is defined by either (17) or (18). As in the previous case, the function is not continuous but the discontinuities are small. Again we define an auxiliary function:

$$h(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota) := 4\mu\tau_1 + 2\iota\tau_1 + 7\binom{\tau_1}{2} + 3\tau_1 + 2\iota\tau_2 + 8\binom{\tau_2}{2} + 3\tau_2 + 8\binom{\tau_3}{2} + 8\tau_3\tau_4 + 3\tau_3 + 7\tau_1\tau_2 + (2+3\mu)\tau_2 + 7\tau_1(\tau_3+\tau_4) + (3+3\mu)\tau_3 + 8\tau_2(\tau_3+\tau_4) + (2+\iota)\tau_3 + (2\mu+\iota-2)(3\tau_4+2) + \binom{3\tau_4-2\mu-\iota+4}{2}. \tag{54}$$

The difference between h and g_ℓ is at most $2\tau_2 + 5\tau_3 + 12n + 20$. As in the previous case, we compute some differences of this auxiliary function, where we write h as a shorthand for $h(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota)$.

$$h(\tau_1+x,\tau_2-x,\tau_3,\tau_4,\mu,\iota) - h = \tfrac{x^2}{2} + (\mu-\tau_2-\tau_3-\tau_4-\tfrac{3}{2})x \tag{55}$$
$$h(\tau_1+x,\tau_2,\tau_3-x,\tau_4,\mu,\iota) - h = \tfrac{x^2}{2} + (\mu+\iota-\tau_2-\tau_3-\tau_4-\tfrac{9}{2})x \tag{56}$$
$$h(\tau_1,\tau_2+x,\tau_3-x,\tau_4,\mu,\iota) - h = (\iota-3)x \tag{57}$$
$$h(\tau_1+x,\tau_2,\tau_3,\tau_4-x,\mu,\iota) - h = x^2 + (4\mu+2\iota-\tau_2-\tau_3-2\tau_4-5)x \tag{58}$$
$$h(\tau_1,\tau_2+x,\tau_3,\tau_4-x,\mu,\iota) - h = \tfrac{x^2}{2} + (2\iota+3\mu-\tau_4-\tfrac{7}{2})x \tag{59}$$
$$h(\tau_1,\tau_2,\tau_3+x,\tau_4-x,\mu,\iota) - h = \tfrac{x^2}{2} + (3\mu-\tau_4+\tfrac{1}{2})x \tag{60}$$

We will first consider the case $\tau_4 = \frac{n-3k}{3}$. We will show that in this case $g_\ell < e(n,k)$. By (57), h is larger by at most $3t_3$ than $h(\tau_1,\tau_2+\tau_3,0,\tau_4,\mu,\iota)$, and so $g_\ell(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota)$ is larger than $h(\tau_1,\tau_2+\tau_3,0,\tau_4,\mu,\iota)$ by at most $2\tau_2 + 8\tau_3 + 12n + 20 < 16n$. Since $n \geq 4\cdot 10^4$ it suffices to show that $h(\tau_1,\tau_2,0,\frac{n-3k}{3},\mu,\iota)$ is smaller than $e(n,k)$ by at least $\frac{1}{2000}n^2$.

We may assume that $h(\tau_1,\tau_2,0,\frac{n-3k}{3},\mu,\iota)$ is maximised, and since (55) is a positive quadratic in x this implies that either $\tau_1 = 2k-\frac{n}{3}$ and $\tau_2 = 0$ or vice versa; we separate subcases. As in the previous case, we will discard linear terms, which will never exceed $19n < \frac{1}{2000}n^2$.

Subcase 1: $\tau_1 = 2k-\frac{n}{3}$ and $\tau_2 = 0$; $k \leq \frac{n}{4}$.

We have

$$h(2k-\tfrac{n}{3},0,0,\tfrac{n-3k}{3},\mu,\iota) = \tfrac{7kn}{3} - \tfrac{n^2}{18} - 3k^2 - k + \tfrac{n}{6} + 2 \tag{61}$$
$$\approx \tfrac{7kn}{3} - \tfrac{n^2}{18} - 3k^2\,, \text{and}$$
$$h(2k-\tfrac{n}{3},0,0,\tfrac{n-3k}{3},\mu,\iota) - e(E_2(n,k)) \approx \tfrac{7kn}{3} - \tfrac{11n^2}{36} - 5k^2\,, \tag{62}$$

which is maximised at $k = \frac{7n}{30}$ where we obtain $\frac{-n^2}{45}$.

Subcase 2: $\tau_1 = 2k-\frac{n}{3}$ and $\tau_2 = 0$; $k > \frac{n}{4}$. We have

$$h(2k-\tfrac{n}{3},0,0,\tfrac{n-3k}{3},\mu,\iota) - e(E_3(n,k)) \approx \tfrac{kn}{3} - \tfrac{n^2}{18} - k^2\,, \tag{63}$$

which function is maximised at $k=\frac{n}{6}$ where we obtain $\frac{-n^2}{36}$.

Subcase 3: $\tau_1=0$ and $\tau_2=2k-\frac{n}{3}$; $\frac{n}{5}\le k\le\frac{n}{4}$. Since $2\mu=n-3k-\iota$ we have

$$h(0,2k-\tfrac{n}{3},0,\tfrac{n-3k}{3},\mu,\iota)=\tfrac{7kn}{6}+\tfrac{n^2}{18}+\iota(k-\tfrac{n}{6})+2k-\tfrac{n}{3}+2\,, \tag{64}$$

which is maximised over ι when $\iota=n-3k$. We can therefore assume $\iota=n-3k$ and obtain

$$h(0,2k-\tfrac{n}{3},0,\tfrac{n-3k}{3},\mu,\iota)-e(E_2(n,k))\approx\tfrac{8kn}{3}-\tfrac{13n^2}{36}-5k^2\,, \tag{65}$$

which function is maximised at $k=\frac{4n}{15}$ where the value is $\frac{-n^2}{180}$.

Subcase 4: $\tau_1=0$ and $\tau_2=2k-\frac{n}{3}$; $\frac{n}{4}<k\le\frac{3n}{10}$. As in (64), the function $h(0,2k-\frac{n}{3},0,\frac{n-3k}{3},\mu,\iota)$ is maximised over ι when $\iota=n-3k$ and therefore we assume this is the case. We have

$$h(0,2k-\tfrac{n}{3},0,\tfrac{n-3k}{3},\mu,\iota)-e(E_3(n,k))\approx\tfrac{2kn}{3}-\tfrac{n^2}{9}-k^2\,, \tag{66}$$

which function is maximised at $k=n/3$, so within this range of k the maximum is at $k=\frac{3n}{10}$ where the value is $\frac{-n^2}{900}$.

Subcase 5: $\tau_1=0$ and $\tau_2=2k-\frac{n}{3}$; $\frac{3n}{10}<k\le\frac{n}{3}$. As in (64), the function $h(0,2k-\frac{n}{3},0,\frac{n-3k}{3},\mu,\iota)$ is maximised over ι when $\iota=n-3k$ and therefore we assume this is the case. We have

$$h(0,2k-\tfrac{n}{3},0,\tfrac{n-3k}{3},\mu,\iota)-e(E_4(n,k))\approx\tfrac{17kn}{3}-\tfrac{11n^2}{18}-12k^2\,, \tag{67}$$

which function is maximised at $k=\frac{17n}{72}$ where the value is $\frac{-41n^2}{540}$.

These subcases are exhaustive, and it follows that if $\tau_4=\frac{n-3k}{3}$ then $g_\ell<e(n,k)$.

It remains to consider the possibility $\tau_4>\frac{n-3k}{3}$. Observe that (58), (59) and (60) are quadratics in x with positive x^2 coefficient. In particular, if h is maximised and $3\tau_4>n-3k$ then $\tau_1=\tau_2=\tau_3=0$. The same statement is almost true replacing h with g_ℓ: the only problem is that when τ_1 is varied, the function is discontinuous, being greater by 20 than it 'should be', at $\tau_1=1$, that being where (18) is used rather than (17). Nevertheless, we can conclude that if g_ℓ is maximised then $\tau_2=\tau_3=0$ and $\tau_1\in\{0,1\}$. We have

$$g_\ell(\tau_1,0,0,k-\tau_1,\mu,\iota)\le(2n+3+k-7\tau_1)\tau_1+p(n-3\tau_1,n-3k)+20\,.$$

This is very close to $e(E_4(n,k))$, and if $n/5\le k\le 3n/10$ then it is smaller than $e(E_3(n,k))$ by at least $\frac{n^2}{100}-30n>0$. If on the other hand $k>\frac{3n}{10}$, then $n-3\tau_1>9(n-3k)$ and by (14) we have

$$g_\ell(\tau_1,0,0,k-\tau_1,\mu,\iota)=$$
$$=(2n-8k-\tfrac{3}{2}-\tfrac{5\tau_1}{2})\tau_1-\tfrac{3n}{2}+9k^2+9k-3kn+\tfrac{n^2}{2}+\begin{cases}2 & \tau_1=0\\ 22 & \tau_1=1\end{cases}. \tag{68}$$

Since $2n-8k-\frac{3}{2}-\frac{5\tau_1}{2}<\frac{-2n}{5}<-20$ we conclude that g_ℓ is maximised with $\tau_1=0$. Now we have

$$g_\ell(0,0,0,k,\mu,\iota)=p(n,n-3k)=e(E_4(n,k))\,, \tag{69}$$

as desired.

Case 3: $3\tau_4<\max(528,3\kappa_0)$ and $n-3k<\max(528,3\kappa_0)$.

Observe that, taking the largest terms of each of (16), (17) and (18), we have

$$\begin{aligned} &g_\ell(\tau_1,\tau_2,\tau_3,\tau_4,\mu,\iota)\\ &\qquad\le 8\binom{k-\tau_4}{2}+(8\tau_4+3)(k-\tau_4)+\binom{3\tau_4}{2}+(n-3k)n+20\\ &\qquad\le \tfrac{4n^2}{9}+2\max(528,3\kappa_0)n+\binom{\max(528,3\kappa_0)}{2}+20\\ &\qquad\le e(E_4(n,k))-\tfrac{n^2}{18}+8\max(528,3\kappa_0)n+\max(528,3\kappa_0)^2\\ &\qquad< e(E_4(n,k))\,, \end{aligned}$$

where the final inequality uses $n\ge 300\max(528,3\kappa_0)$.

These cases are exhaustive, completing the proof. □