ON THE STRUCTURE OF ORIENTED GRAPHS AND DIGRAPHS WITH FORBIDDEN TOURNAMENTS OR CYCLES

DANIELA KÜHN, DERYK OSTHUS, TIMOTHY TOWNSEND, YI ZHAO

ABSTRACT. Motivated by his work on the classification of countable homogeneous oriented graphs, Cherlin asked about the typical structure of oriented graphs (i) without a transitive triangle, or (ii) without an oriented triangle. We give an answer to these questions (which is not quite the predicted one). Our approach is based on the recent 'hypergraph containers' method, developed independently by Saxton and Thomason as well as by Balogh, Morris and Samotij. Moreover, our results generalise to forbidden transitive tournaments and forbidden oriented cycles of any order, and also apply to digraphs. Along the way we prove several stability results for extremal digraph problems, which we believe are of independent interest.

1. INTRODUCTION

1.1. H-free graphs. Given a fixed graph H, a graph is called H-*free* if it does not contain H as a (not necessarily induced) subgraph. In 1976 Erdős, Kleitman and Rothschild [17] asymptotically determined the logarithm of the number of K_k-free graphs on n vertices, for every $k \geq 3$. This was strengthened by Kolaitis, Prömel and Rothschild [20], who showed that almost all K_k-free graphs are $(k-1)$-partite, for every $k \geq 3$ (the case $k = 3$ of this was already proved in [17]). This was one of the starting points for a vast body of work concerning the number and structure of H-free graphs on n vertices (see, e.g. [5, 6, 7, 9, 11, 16, 20, 23, 25]). The strongest of these results essentially state that for a large class of graphs $\mathcal{H}$, and any $H \in \mathcal{H}$, almost all H-free graphs have a similar structure to that of the extremal H-free graph. More recently, some related results have been proved for hypergraphs (see, e.g. [10, 24]).

However, the corresponding questions for digraphs and oriented graphs are almost all wide open, and are the subject of this paper. Until now the only results of the above type for oriented graphs were proved by Balogh, Bollobás and Morris [3, 4] who classified the possible 'growth speeds' of oriented graphs with a given property. Moreover Robinson [26, 27], and independently Stanley [30], counted the number of acyclic digraphs. A related problem was considered by Alon and Yuster [2], who determined $\max_G D(G, T)$ over all n-vertex graphs G for sufficiently large n, where T is a fixed tournament and $D(G, T)$ denotes the number of T-free orientations of G (note that $\sum_G D(G, T)$ is the number of T-free oriented graphs on n vertices).

1.2. Oriented graphs and digraphs with forbidden tournaments or cycles. A *digraph* is a pair (V, E) where V is a set of vertices and E is a set of ordered pairs of distinct vertices in V (note that this means that in a digraph we do not allow loops or multiple edges in the same direction). An *oriented graph* is a digraph with at most one edge between two vertices, so may be considered as an orientation of a simple undirected graph. A *tournament* is an orientation of a complete graph. We denote a transitive tournament on k vertices by T_k, and a directed cycle on k vertices by C_k. We only consider labelled graphs and digraphs.

Given a class of graphs $\mathcal{A}$, we let $\mathcal{A}_n$ denote the set of all graphs in $\mathcal{A}$ that have precisely n vertices, and we say that *almost all graphs in $\mathcal{A}$ have property $\mathcal{B}$* if

$$\lim_{n\to\infty} \frac{|\{G \in \mathcal{A}_n : G \text{ has property } \mathcal{B}\}|}{|\mathcal{A}_n|} = 1.$$

Clearly any transitive tournament is C_k-free for any k, and any bipartite digraph is T_3-free. In 1998 Cherlin [13] gave a classification of countable homogeneous oriented graphs. He remarked that 'the striking work of [20] does not appear to go over to the directed case' and made the following conjectures.[1]

Conjecture 1.1 (Cherlin)**.**

  (i) *Almost all T_3-free oriented graphs are tripartite.*
 (ii) *Almost all C_3-free oriented graphs are acyclic, i.e. they are subgraphs of transitive tournaments.*

Our first main result not only verifies part (i) of this conjecture, but shows for all $k \geq 2$ that almost all T_{k+1}-free oriented graphs are k-partite. Note that in particular this shows that in fact almost all T_3-free oriented graphs are actually even bipartite. We also prove the analogous result for digraphs.

Theorem 1.2. *Let $k \in \mathbb{N}$ with $k \geq 2$. Then the following hold.*

  (i) *Almost all T_{k+1}-free oriented graphs are k-partite.*
 (ii) *Almost all T_{k+1}-free digraphs are k-partite.*

Theorem 1.2 can be viewed as a directed version of the theorem of Kolaitis, Prömel and Rothschild [20] mentioned earlier. Note also that (i) means that the typical structure of a T_3-free oriented graph is *not* close to that of the extremal T_3-free oriented graph: it is easy to see that the latter is the blow up of a directed triangle (this fact was probably the motivation for Conjecture 1.1(i)).

Our next main result shows in particular that part (ii) of Conjecture 1.1 is in fact false. We actually show something stronger, namely that for all $k \geq 3$ and for almost all C_k-free oriented graphs on n vertices, the number of edges we must change in order to get an acyclic oriented graph is $\Omega(n/\log n)$. We also prove an analogous version of this result for digraphs. However, Conjecture 1.1(ii) is not too far from being true, as we prove also that almost all C_k-free oriented graphs are close to acyclic, in the sense that we only need to change sub-quadratically many edges in order to obtain an acyclic oriented graph. In the case when k is even we prove an analogous version of this result for digraphs too. We also obtain a (less restrictive) structural result for odd k.

In order to state the theorem precisely we need to introduce a little terminology. Given a labelled digraph or oriented graph G with vertex labels $1,\ldots,n$ and an ordering $\sigma : [n] \to [n]$, a *backwards edge* in G with respect to this ordering is any edge directed from a vertex labelled i to a vertex labelled j, where $\sigma(i) > \sigma(j)$. A *transitive-optimal ordering* of G is any ordering of $V(G)$ that minimises the number of backwards edges in G with respect to the ordering. We say that a directed graph is a *transitive-bipartite blow up* if it can be obtained from a transitive tournament by replacing some of its vertices by complete balanced bipartite digraphs. More formally, a directed graph $G = (V, E)$ is a transitive-bipartite blow up if V admits a partition $A_1,\ldots,A_t$ such that for all $i,j \in [t]$ with $i < j$, the graph induced on G by A_i is either a single vertex or a complete balanced bipartite digraph (with edges in both directions) and the edges in E between A_i and A_j are precisely those edges directed from A_i to A_j (and no others).

---

[1]Note that oriented graphs are referred to as digraphs in [13].

Theorem 1.3. *Let $k, n \in \mathbb{N}$ with $k \geq 3$. There exists $c > 0$ such that for every $\alpha > 0$ the following hold.*

  (i) *Almost all C_k-free oriented graphs on n vertices have between $cn/\log n$ and αn^2 backwards edges in a transitive-optimal ordering.*

 (ii) *Almost all C_k-free digraphs on n vertices have at least $cn/\log n$ backwards edges in a transitive-optimal ordering. Moreover,*

    (a) *if k is even then almost all C_k-free digraphs on n vertices have at most αn^2 backwards edges in a transitive-optimal ordering,*

    (b) *if k is odd then almost all C_k-free digraphs on n vertices can be made into a subgraph of a transitive-bipartite blow up by changing at most αn^2 edges.*

We believe that in fact almost all C_k-free oriented graphs have linearly many backwards edges, and that an analogous result holds for C_k-free digraphs in the case when k is even.

Conjecture 1.4. *Let $k, n \in \mathbb{N}$ with $k \geq 3$. Then the following hold.*

  (i) *Almost all C_k-free oriented graphs on n vertices have $\Theta(n)$ backwards edges in a transitive-optimal ordering.*

 (ii) *If k is even then almost all C_k-free digraphs on n vertices have $\Theta(n)$ backwards edges in a transitive-optimal ordering.*

It is not clear to us what to expect in the case when k is odd.

Question 1.5. *Suppose that k is odd and $\alpha > 0$. Do almost all C_k-free digraphs on n vertices have at most αn^2 backwards edges in a transitive-optimal ordering?*

An undirected version of Theorem 1.3 for forbidden odd cycles was proved by Lamken and Rothschild [21], who showed that for odd k, almost all C_k-free graphs are bipartite. (So the situation for oriented graphs is very different from the undirected one.) For even k the undirected problem is far more difficult. Despite major recent progress by Morris and Saxton [22] the problem of counting the number of C_k-free graphs is still open for even k.

We remark that in Theorem 1.2 we actually get exponential bounds on the proportion of T_{k+1}-free oriented graphs and digraphs that are not k-partite. We also get similar exponential bounds in Theorem 1.3.

1.3. **Sketch of proofs.** A key tool in our proofs is a recent and very powerful result of Saxton and Thomason [28] as well as Balogh, Morris and Samotij [8], which gives an upper bound on the number of independent sets in certain hypergraphs. Briefly, the result states that under suitable conditions on a uniform hypergraph G, there is a small collection $\mathcal{C}$ of small subsets (known as containers) of $V(G)$ such that every independent set of vertices in G is a subset of some element of $\mathcal{C}$. We will use the formulation of Saxton and Thomason [28]. The precise statement of this result (Theorem 3.2) is deferred until Section 3. Roughly speaking, the use of hypergraph containers allows us to reduce an asymptotic counting problem to an extremal problem. It should be noted that the method of hypergraph containers is much more general than Theorem 3.2. Saxton and Thomason [29] also gave a short proof of a somewhat weaker version of Theorem 3.2, which would still have been sufficient for our purposes.

Our approach to proving our main results is as follows. Firstly, in Section 3 we use the main result of [28] to derive a container result which is applicable to digraphs. Then in Section 4 we apply this digraph containers result in a relatively standard way to show that almost all T_{k+1}-free oriented graphs, and almost all T_{k+1}-free digraphs, are close to k-partite (see Lemma 4.5). In Section 5 we combine Lemma 4.5 with an inductive argument to prove the results on the exact structure of typical T_{k+1}-free oriented graphs and digraphs given by Theorem 1.2.

In Section 6 we use the digraph containers result to show that almost all C_k-free oriented graphs are close to acyclic, and that the analogous result for digraphs holds in the case when k is even (see

Lemma 6.8(i), (ii)). For odd k we show that almost all C_k-free digraphs are close to a subgraph of a transitive-bipartite blow up (see Lemma 6.8(iii)). Finally, in Section 7 we complete the proof of Theorem 1.3 by giving a lower bound on the number of backwards edges in C_k-free oriented graphs and digraphs, the upper bounds in Theorem 1.3 being given by Lemma 6.8.

As part of the proofs in Sections 4 and 6 we prove several stability results on digraphs which we believe are of independent interest:

(i) Suppose $k \in \mathbb{N}$ and G is a T_{k+1}-free digraph on n vertices with $e(G) \geq \mathrm{ex}_{di}(n, T_{k+1}) - o(n^2)$. Then G is close to a complete balanced k-partite digraph. (See Lemma 4.3).

(ii) Suppose $k \in \mathbb{N}$ with $k \geq 4$ and k even, and suppose G is a C_k-free digraph on n vertices with $e(G) \geq \mathrm{ex}_{di}(n, C_k) - o(n^2)$. Then G is close to a transitive tournament. (See Lemma 6.5).

(iii) Suppose $k \in \mathbb{N}$ with $k \geq 3$ and k odd, and suppose G is a C_k-free digraph on n vertices with $e(G) \geq \mathrm{ex}_{di}(n, C_k) - o(n^2)$. Then G is close to a transitive-bipartite blow up. (See Lemma 6.6).

Here $\mathrm{ex}_{di}(n, H)$ denotes the maximum number of edges among all H-free digraphs on n vertices. The corresponding Turán type results which determine $\mathrm{ex}_{di}(n, H)$ for $H = T_k$ and $H = C_k$ were proved by Brown and Harary [12] and Häggkvist and Thomassen [19] respectively. These stability results are used in the proofs of Theorems 1.2(ii) and 1.3(ii). We actually prove 'weighted' generalisations of (i) and (ii) which can be used to prove the assertions about oriented graphs in Theorems 1.2(i) and 1.3(i).

Before starting on any of this however, we lay out some notation and set out some useful tools in Section 2, below.

2. Notation and tools

For a set X we let $X^{(r)}$ denote the set of all (unordered) subsets of X of size r. An r-*uniform hypergraph*, or r-*graph*, is a pair (V, E) where V is a set of vertices and $E \subseteq V^{(r)}$. If $G = (V, E)$ is a graph, digraph, oriented graph or r-graph, we let $V(G) := V$, $E(G) := E$, $v(G) := |V(G)|$, and $e(G) := |E(G)|$. For a digraph $G = (V, E)$ define $\Delta^0(G)$ as the maximum of $d^+(v)$ and $d^-(v)$ among all $v \in V$. We write uv for the edge directed from u to v. For a vertex $v \in V$, define the out-neighbourhood of v in G to be $N_G^+(v) := \{u \in V : vu \in E\}$, and similarly define the in-neighbourhood of v in G to be $N_G^-(v) := \{u \in V : uv \in E\}$. Given a set $U \subseteq V$, we sometimes also write $N_U^+(v) := N_G^+(v) \cap U$ and define $N_U^-(v)$ similarly. For disjoint subsets $U, U' \subseteq V$ we let $G[U, U']$ denote the subdigraph of G with vertex set $U \cup U'$ whose edge set consists of all edges between U and U' in G (in both directions). We let $e(U, U') := e(G[U, U'])$. If Q is a k-partition of $[n]$ with partition classes $V_1 \ldots, V_k$, and G is a (di)graph or oriented graph with vertex set $[n]$, we say that Q is a k-*partition of G* if for every $i \in [k]$ we have that $E(G)$ contains no edges uv with $u, v \in V_i$. We assume k-partitions to be unordered unless otherwise stated. For two digraphs G and G' on vertex set $[n]$, we write $G = G' \pm \varepsilon n^2$ if G can be obtained from G' by changing (i.e. adding, deleting, or changing the orientation of) at most εn^2 edges. Given an r-graph $H = (V, E)$ and $\sigma \in V^{(d)}$, where $0 \leq d \leq r - 1$, let $d_H(\sigma) := |\{e \in E : \sigma \subseteq e\}|$ be the *degree* of σ in H. We may simply write $d(\sigma)$ for $d_H(\sigma)$ when it is obvious which r-graph H we are working with. The *average vertex degree* of H is defined to be $(1/|V|) \sum_{v \in V} d_H(\{v\})$. Given an oriented graph H and a digraph H',

- let $f(n, H)$ denote the number of labelled H-free oriented graphs on n vertices,
- let $T(n, k)$ denote the number of labelled k-partite oriented graphs on n vertices,
- let $f^*(n, H')$ denote the number of labelled H'-free digraphs on n vertices,
- let $T^*(n, k)$ denote the number of labelled k-partite digraphs on n vertices.

In some proofs, given $a, b \in \mathbb{R}$ with $0 < a, b < 1$, we will use the notation $a \ll b$ to mean that we can find an increasing function g for which all of the conditions in the proof are satisfied whenever

$a \leq g(b)$. Throughout the paper we write $\log x$ to mean $\log_2 x$, and we assume all graphs, oriented graphs, and digraphs to be labelled unless otherwise stated. We also assume all large numbers to be integers, so that we may sometimes omit floors and ceilings for the sake of clarity.

We define $H(p) := -p \log p - (1-p) \log(1-p)$, the binary entropy function. The following bound will prove useful to us. For $n \geq 1$ and $0 < p < 1/2$,

$$(2.1) \qquad \binom{n}{\leq pn} := \sum_{i=0}^{\lfloor pn \rfloor} \binom{n}{i} \leq 2^{H(p)n}.$$

In a number of our proofs we shall also use the following Chernoff bound.

Theorem 2.1 (Chernoff bound). *Let X have binomial distribution and let $a > 0$. Then*

$$P(X < \mathbb{E}[X] - a) < \exp\left(-\frac{a^2}{2\mathbb{E}[X]}\right).$$

3. Digraph containers

Our main tool is Theorem 3.2 from [28]. Given a hypergraph G satisfying certain degree conditions, it gives a small set of almost independent sets in G (containers) which together contain all independent sets of G. In our applications the vertex set of G will be the edge set of the complete digraph, and the hyperedges will correspond to copies of the forbidden subdigraph. To formulate the degree conditions we need the following definition.

Definition 3.1. *Let G be an r-graph on n vertices with average vertex degree d. Let $\tau > 0$. Given $v \in V(G)$ and $1 \leq j \leq r, n$, let*

$$d^{(j)}(v) := \max\{d(\sigma) : v \in \sigma \subseteq V(G), |\sigma| = j\}.$$

If $d > 0$ we define $\delta_j = \delta_j(\tau)$ by the equation

$$\delta_j \tau^{j-1} nd = \sum_{v \in V(G)} d^{(j)}(v).$$

Then the co-degree function $\delta(G, \tau)$ is defined by

$$\delta(G, \tau) := 2^{\binom{r}{2}-1} \sum_{j=2}^{r} 2^{-\binom{j-1}{2}} \delta_j.$$

If $d = 0$ we define $\delta(G, \tau) := 0$.

Theorem 3.2. [28, Corollary 2.7] *Suppose that $0 < \varepsilon < \frac{1}{2}$ and $\tau \leq \frac{1}{144 r^2! r}$. Let G be an r-graph with vertex set $[n]$ satisfying $\delta(G, \tau) \leq \frac{\varepsilon}{12 r!}$. Then there exists a constant $c = c(r)$ and a collection $\mathcal{C}$ of subsets of $[n]$ with the following properties.*

  (a) *For every independent set I of G there exists $C \in \mathcal{C}$ such that $I \subseteq C$.*
  (b) *$e(G[C]) \leq \varepsilon e(G)$ for all $C \in \mathcal{C}$.*
  (c) *$\log |\mathcal{C}| \leq c \log(\frac{1}{\varepsilon}) n\tau \log(\frac{1}{\tau})$.*

We will apply Theorem 3.2 to prove Theorem 3.3 below, which is a digraph analogue to [28, Theorem 1.3]. To state this we need the following definitions. Given a digraph $G = (V, E)$, let $f_1(G)$ be the number of pairs $u, v \in V$ such that exactly one of uv and vu is an edge of G, and let $f_2(G)$ be the number of pairs $u, v \in V$ such that both uv and vu are edges of G. The following definition of the *weighted size* of G will be crucial in this paper. For $a \in \mathbb{R}$ with $a \geq 1$ we define

$$e_a(G) := a \cdot f_2(G) + f_1(G).$$

This definition allows for a unified approach to extremal problems on oriented graphs and digraphs. We will be mainly interested in the cases $a = 2$ and $a = \log 3$. The former is useful because each

digraph G contains $4^{f_2(G)}2^{f_1(G)} = 2^{e_2(G)}$ (labelled) subdigraphs, and the latter is useful because each digraph G contains $3^{f_2(G)}2^{f_1(G)} = 2^{e_{\log 3}(G)}$ (labelled) oriented subgraphs. Given a digraph H, define the *weighted Turán number* $\mathrm{ex}_a(n, H)$ as the maximum $e_a(G)$ among all H-free digraphs G on n vertices. (So $\mathrm{ex}_2(n, H)$ equals $\mathrm{ex}_{di}(n, H)$ which was defined in Section 1.3.) For $A, B \subseteq V$ we will sometimes write $e_a(A, B)$ to denote $e_a(G[A, B])$.

Given an oriented graph H with $e(H) \geq 2$, we let

$$m(H) = \max_{H' \subseteq H, e(H') > 1} \frac{e(H') - 1}{v(H') - 2}.$$

Theorem 3.3. *Let H be an oriented graph with $h := v(H)$ and $e(H) \geq 2$, and let $a \in \mathbb{R}$ with $a \geq 1$. For every $\varepsilon > 0$, there exists $c > 0$ such that for all sufficiently large N, there exists a collection $\mathcal{C}$ of digraphs on vertex set $[N]$ with the following properties.*

  (a) *For every H-free digraph I on $[N]$ there exists $G \in \mathcal{C}$ such that $I \subseteq G$.*
  (b) *Every digraph $G \in \mathcal{C}$ contains at most εN^h copies of H, and $e_a(G) \leq \mathrm{ex}_a(N, H) + \varepsilon N^2$.*
  (c) $\log |\mathcal{C}| \leq c N^{2 - 1/m(H)} \log N$.

Note that in (a), since I, G are labelled digraphs, $I \subseteq G$ means that I is contained in G in the labelled sense, i.e. the copy of I in G has the same vertex labels as I.

The following corollary is a straightforward consequence of Theorem 3.3. We will not use it elsewhere in the paper, but we include it to illustrate what one can achieve even just with a 'direct' application of Theorem 3.3.

Corollary 3.4. *For every oriented graph H with $e(H) \geq 2$, we have $f(n, H) = 2^{\mathrm{ex}_{\log 3}(n, H) + o(n^2)}$ and $f^*(n, H) = 2^{\mathrm{ex}_2(n, H) + o(n^2)}$.*

Proof.   We only prove the first part here; the proof of the second part is almost identical. Clearly $f(n, H) \geq 2^{\mathrm{ex}_{\log 3}(n, H)}$. By Theorem 3.3, for every $\varepsilon > 0$ there is a collection $\mathcal{C}$ of digraphs on $[n]$ satisfying properties (a)–(c). We know that every digraph $G \in \mathcal{C}$ contains $2^{e_{\log 3}(G)}$ oriented subgraphs. Since each H-free oriented graph is contained in some $G \in \mathcal{C}$, and $|\mathcal{C}| \leq 2^{cn^{2-1/m(H)} \log n}$,

$$f(n, H) \leq \sum_{G \in \mathcal{C}} 2^{e_{\log 3}(G)} \leq 2^{\mathrm{ex}_{\log 3}(n, H) + \varepsilon n^2 + o(n^2)}.$$

We are done by letting $\varepsilon \to 0$.   □

It is not possible to extend Theorem 3.3 or Corollary 3.4 to all digraphs H. For example, let DK_3 denote the double triangle (which consists of three vertices and all possible ordered pairs as edges). Assume that n is even and let $A \cup B$ be a balanced partition of $[n]$. Let $\mathcal{G}$ denote the family of all digraphs G on $[n]$ such that $G[A]$ and $G[B]$ are oriented graphs. It is clear that no $G \in \mathcal{G}$ contains a copy of DK_3, and $|\mathcal{G}| = 2^{n^2/2} 3^{2\binom{n/2}{2}}$ (there are 2 choices for each of the $n^2/2$ ordered pairs between A and B and there are 3 choices for each of the $\binom{n/2}{2}$ unordered pairs on either A or B). Furthermore, it is easy to see that $\mathrm{ex}_2(n, DK_3) = \frac{n^2}{2} + 2\binom{n/2}{2}$ (and the digraphs in $\mathcal{G}$ with the maximum number of edges are extremal digraphs). Thus

$$f^*(n, DK_3) \geq |\mathcal{G}| = 2^{\frac{n^2}{2}} 3^{2\binom{n/2}{2}} \gg 2^{\frac{n^2}{2} + 2\binom{n/2}{2}} = 2^{\mathrm{ex}_2(n, DK_3)}.$$

The proof of Theorem 3.3 is similar to that of [28, Theorem 1.3]. We first define the hypergraph $D(N, H)$, which will play the role of G in Theorem 3.2.

Definition 3.5. *Let H be an oriented graph, let $r := e(H)$ and let $N \in \mathbb{N}$. The r-graph $D(N, H)$ has vertex set $U = ([N] \times [N]) \setminus \{(i, i) : i \in [N]\}$, where $B \in U^{(r)}$ is an edge whenever B, considered as a digraph with vertices in $[N]$, is isomorphic to H.*

We wish to apply Theorem 3.2 to $D(N, H)$. To do this we require an upper bound on $\delta(D(N, H), \tau)$ for some suitable value of τ. We give one in the following lemma, the proof of which is identical to that of [28, Lemma 9.2] and is therefore omitted here.

Lemma 3.6. *Let H be an oriented graph with $r := e(H) \geq 2$, and let $\gamma \leq 1$. For N sufficiently large, $\delta\left(D(N, H), \gamma^{-1} N^{-1/m(H)}\right) \leq r 2^{r^2} v(H)!^2 \gamma$.*

We now state a supersaturation result, which we will use to bound the number of edges in containers. It is the digraph analogue of the well-known supersaturation result of Erdős and Simonovits [18]. Its proof is almost the same, and is omitted here.

Lemma 3.7 (Supersaturation). *Let H be a digraph on h vertices, and let $a \in \mathbb{R}$ with $a \geq 1$. For any $\varepsilon > 0$, there exists $\delta > 0$ such that the following holds for all sufficiently large n. For any digraph G on n vertices, if G contains at most δn^h copies of H, then $e_a(G) \leq \mathrm{ex}_a(n, H) + \varepsilon n^2$.*

We may now apply Theorem 3.2 to $D(N, H)$, using Lemmas 3.6 and 3.7, to obtain Theorem 3.3. The details of this are identical to the proof of Theorem 1.3 in [28] and are omitted here.

4. Rough structure of typical T_{k+1}-free oriented graphs and digraphs

In this section we prove a stability result for T_{k+1}-free digraphs. We apply this (together with Theorem 3.3) at the end of this section to determine the 'rough' structure of typical T_{k+1}-free oriented graphs and digraphs.

The *Turán graph* $\mathrm{Tu}_k(n)$ is the largest complete k-partite graph on n vertices (thus each vertex class has $\lfloor n/k \rfloor$ or $\lceil n/k \rceil$ vertices). Let $t_k(n) := e(\mathrm{Tu}_k(n))$. Let $DT_k(n)$ be the digraph obtained from $\mathrm{Tu}_k(n)$ by replacing each edge of $\mathrm{Tu}_k(n)$ by two edges of opposite directions. Obviously, for all $k \in \mathbb{N}$, $DT_k(n)$ is T_{k+1}-free so $\mathrm{ex}_a(n, T_{k+1}) \geq e_a(DT_k(n)) = a \cdot t_k(n)$. In Lemma 4.1 below we show that $DT_k(n)$ is the unique extremal digraph for T_{k+1}. This result is not needed for any of our proofs, but we believe it is of independent interest, in addition to being useful for illustrating the general proof method of Lemma 4.3. Note that the case $a = 2$ of Lemma 4.1 is already due to Brown and Harary [12].

Lemma 4.1. *Let $a \in \mathbb{R}$ with $3/2 < a \leq 2$ and let $k, n \in \mathbb{N}$. Then $\mathrm{ex}_a(n, T_{k+1}) = a \cdot t_k(n)$, and $DT_k(n)$ is the unique extremal T_{k+1}-free digraph on n vertices.*

Proof. Note that $DT_k(n)$ is the unique k-partite digraph D on n vertices which maximises $e_a(D)$. Moreover, $e_a(DT_k(n)) = a \cdot t_k(n)$. Thus, it suffices to show that for every non-k-partite T_{k+1}-free digraph G on n vertices, there exists a k-partite digraph H on the same vertex set such that $e_a(G) < e_a(H)$.

We prove this by induction on k. This is trivial for the base case $k = 1$, as the only T_2-free digraph is the empty graph. Suppose that $k > 1$ and that the claim holds for $k - 1$. Let $G = (V, E)$ be a non-k-partite T_{k+1}-free digraph on n vertices. Without loss of generality, suppose that $d^+(x) = \Delta^0(G)$ for some vertex $x \in V$. Let $S := N^+(x)$ and $T := V \setminus S$. Since G is T_{k+1}-free we have that $G[S]$ is T_k-free. By induction hypothesis, either (i) there is a $(k-1)$-partite digraph H' on S such that $e_a(G[S]) < e_a(H')$, or (ii) $G[S]$ is $(k-1)$-partite and hence there is trivially a $(k-1)$-partite digraph H' on S such that $e_a(G[S]) = e_a(H')$. Next we want to replace all the edges inside T with edges between S and T as follows. Suppose $y, z \in T$ with $yz \in E$. Then there are $y', z' \in S$ with $yy' \notin E$ and $z'z \notin E$ otherwise $d^+(y) \geq |S| + 1$ or $d^-(z) \geq |S| + 1$, contradicting the assumption $\Delta^0(G) = d^+(x) = |S|$. We now replace yz with yy' and $z'z$. By the definition of $e_a(\cdot)$, and since $a \leq 2$, the gain of adding the edges yy' and $z'z$ is at least $2(a-1)$, while the loss of removing yz is at most one. Thus, since $a > 3/2$, we have that $e_a(G)$ increases by $2a - 3 > 0$. Note that this procedure does not change the in-degree or out-degree of any vertex in T. We repeat this for every edge inside T. At the end we obtain a digraph G' with no edge inside T. We replace $G'[S] = G[S]$ with the $(k-1)$-partite digraph H' obtained before to obtain a k-partite digraph

H. We now consider the two cases (i) and (ii) discussed previously. If $e_a(G[S]) < e_a(H')$ then it is clear that $e_a(G) < e_a(H)$, and we are done. Otherwise, $G[S]$ is $(k-1)$-partite. In this case, since G is not k-partite, there must be an edge inside T, so that there exists y, z as above. Hence $e_a(G) < e_a(G')$, and so clearly $e_a(G) < e_a(H)$ in this case too, as required.   □

We now prove a stability version of Lemma 4.1. The proof idea builds on that of Lemma 4.1. The proof will also make use of the following proposition, which can be proved by a simple but tedious calculation, which we omit here.

Proposition 4.2. *Let* $k, n \in \mathbb{N}$ *with* $n \geq k \geq 2$ *and let* $s > 0$. *Suppose* G *is a* k-partite graph on n vertices in which some vertex class A satisfies $|A - n/k| \geq s$. Then

$$e(G) \leq t_k(n) - s\left(\frac{s}{2} - k\right).$$

Lemma 4.3 (Stability). *Let* $a \in \mathbb{R}$ *with* $3/2 < a \leq 2$ *and let* $k \in \mathbb{N}$. *For any* $\beta > 0$ *there exists* $\gamma > 0$ *such that the following holds for all sufficiently large* n. *If a digraph* G *on* n *vertices is* T_{k+1}-free, and $e_a(G) \geq \mathrm{ex}_a(n, T_{k+1}) - \gamma n^2$, then $G = DT_k(n) \pm \beta n^2$.

Proof.   Choose γ and n_0 such that $1/n_0 \ll \gamma \ll \beta$, and consider any $n \geq n_0$. We follow the proof of Lemma 4.1, in which we fix a vertex $x \in V$ with $d^+(x) = \Delta^0(G)$, and let $S := N^+(x)$ and $T := V \setminus S$, and proceed by induction on k. Again, the base case $k = 1$ is trivial as the only T_2-free digraph is the empty graph. Let m_1 be the number of edges of $G[T]$, and let m_2 be the number of non-edges between T and S (in G).

Let G' be the digraph obtained from G by replacing each edge inside T with two edges between T and S as in the proof of Lemma 4.1. Then

$$(4.1) \qquad\qquad e_a(G) \leq e_a(G') - (2a-3)m_1.$$

Since there are $m_2 - 2m_1$ non-edges between T and S in G', and adding any one of them would increase $e_a(G')$ by at least $a - 1$ (since $a \leq 2$), we have that

$$(4.2) \qquad\qquad e_a(G') \leq |T||S|a + e_a(G[S]) - (m_2 - 2m_1)(a-1).$$

Let

$$(4.3) \qquad\qquad m_3 := \mathrm{ex}_a(|S|, T_k) - e_a(G[S]).$$

Then $m_3 \geq 0$ because $G[S]$ is T_k-free.

Let H be the k-partite digraph obtained from $DT_{k-1}(|S|)$ (on S) by adding the vertex set T together with all the edges (in both directions) between S and T. Then

$$(4.4) \qquad\qquad e_a(H) = |T||S|a + \mathrm{ex}_a(|S|, T_k).$$

Altogether, this gives that

$$e_a(G) \overset{(4.1)}{\leq} e_a(G') - (2a-3)m_1 \overset{(4.2)}{\leq} |T||S|a + e_a(G[S]) - (m_2 - 2m_1)(a-1) - (2a-3)m_1$$

$$\overset{(4.3)}{=} |T||S|a + \mathrm{ex}_a(|S|, T_k) - m_3 - (m_2 - 2m_1)(a-1) - (2a-3)m_1$$

$$\overset{(4.4)}{=} e_a(H) - m_3 - (m_2 - 2m_1)(a-1) - (2a-3)m_1.$$

Let $s := ||T| - \frac{n}{k}|$. By Proposition 4.2 we have that if $s \geq 4k$ then $\mathrm{ex}_a(n, T_{k+1}) \geq a \cdot t_k(n) \geq e_a(H) + as^2/4$. So if $s \geq 4k$ then

$$e_a(G) \leq \mathrm{ex}_a(n, T_{k+1}) - \frac{as^2}{4} - m_3 - (m_2 - 2m_1)(a-1) - (2a-3)m_1,$$

and if $s < 4k$ then

$$e_a(G) \leq \mathrm{ex}_a(n, T_{k+1}) - m_3 - (m_2 - 2m_1)(a-1) - (2a-3)m_1.$$

In either case, since $\mathrm{ex}_a(n, T_{k+1}) - \gamma n^2 \leq e_a(G)$ by assumption, we have that $m_1 \leq \frac{\gamma}{2a-3} n^2$, $m_2 \leq (\frac{\gamma}{a-1} + \frac{2\gamma}{2a-3}) n^2$, $m_3 \leq \gamma n^2$, and $s^2 \leq 4\gamma n^2/a$.

Recall that $e_a(G[S]) = \mathrm{ex}_a(|S|, T_k) - m_3$. Hence we have by induction hypothesis that, since $\gamma \ll \beta$ and $|S| = \Delta^0(G)$ is sufficiently large, $G[S] = DT_{k-1}(|S|) \pm (\beta/2)|S|^2$. Note that we can obtain the digraph $DT_k(n)$ from G by removing m_1 edges inside T, adding m_2 edges between T and S, changing at most $(\beta/2)n^2$ edges inside S, and changing the adjacency of at most s vertices. Thus

$$G = DT_k(n) \pm (m_1 + m_2 + (\beta/2)n^2 + 2sn),$$

and so we have that $G = DT_k(n) \pm \beta n^2$, as required. □

We also need the Digraph Removal Lemma of Alon and Shapira [1].

Lemma 4.4 (Removal Lemma). *For any fixed digraph H on h vertices, and any $\gamma > 0$ there exists $\varepsilon' > 0$ such that the following holds for all sufficiently large n. If a digraph G on n vertices contains at most $\varepsilon' n^h$ copies of H, then G can be made H-free by deleting at most γn^2 edges.*

We are now ready to combine Theorem 3.3 with Lemma 4.3 to show that almost all T_{k+1}-free oriented graphs and almost all T_{k+1}-free digraphs are almost k-partite.

Lemma 4.5. *For every $k \in \mathbb{N}$ with $k \geq 2$ and any $\alpha > 0$ there exists $\varepsilon > 0$ such that the following holds for all sufficiently large n.*

(i) *All but at most $f(n, T_{k+1})2^{-\varepsilon n^2}$ T_{k+1}-free oriented graphs on n vertices can be made k-partite by changing at most αn^2 edges.*

(ii) *All but at most $f^*(n, T_{k+1})2^{-\varepsilon n^2}$ T_{k+1}-free digraphs on n vertices can be made k-partite by changing at most αn^2 edges.*

Proof. We only prove (i) here; the proof of (ii) is almost identical. Let $a := \log 3$. Choose $n_0 \in \mathbb{N}$ and $\varepsilon, \gamma, \beta > 0$ such that $1/n_0 \ll \varepsilon \ll \gamma \ll \beta \ll \alpha, 1/k$. Let $\varepsilon' := 2\varepsilon$ and $n \geq n_0$. By Theorem 3.3 (with T_{k+1}, n and ε' taking the roles of H, N and ε respectively) there is a collection $\mathcal{C}$ of digraphs on vertex set $[n]$ satisfying properties (a)–(c). In particular, by (a), every T_{k+1}-free oriented graph on vertex set $[n]$ is contained in some digraph $G \in \mathcal{C}$. Let $\mathcal{C}_1$ be the family of all those $G \in \mathcal{C}$ for which $e_{\log 3}(G) \geq \mathrm{ex}_{\log 3}(n, T_{k+1}) - \varepsilon' n^2$. Then the number of (labelled) T_{k+1}-free oriented graphs not contained in some $G \in \mathcal{C}_1$ is at most

$$|\mathcal{C}|\, 2^{\mathrm{ex}_{\log 3}(n, T_{k+1}) - \varepsilon' n^2} \leq 2^{-\varepsilon n^2} f(n, T_{k+1}),$$

because $|\mathcal{C}| \leq 2^{n^{2-\varepsilon'}}$, by (c), and $f(n, T_{k+1}) \geq 2^{\mathrm{ex}_{\log 3}(n, T_{k+1})}$. Thus it suffices to show that every digraph $G \in \mathcal{C}_1$ satisfies $G = DT_k(n) \pm \alpha n^2$. By (b), each $G \in \mathcal{C}_1$ contains at most $\varepsilon' n^{k+1}$ copies of T_{k+1}. Thus by Lemma 4.4 we obtain a T_{k+1}-free digraph G' after deleting at most γn^2 edges from G. Then $e_{\log 3}(G') \geq \mathrm{ex}_{\log 3}(n, T_{k+1}) - (\varepsilon' + \gamma)n^2$. We next apply Lemma 4.3 to G' and derive that $G' = DT_k(n) \pm \beta n^2$. As a result, the original digraph G satisfies $G = DT_k(n) \pm (\beta + \gamma)n^2$, and hence $G = DT_k(n) \pm \alpha n^2$ as required. □

5. Exact structure of typical T_{k+1}-free oriented graphs and digraphs

From Section 4 we know that a typical T_{k+1}-free oriented graph is almost k-partite (and similarly for digraphs). In this section we use this information to show inductively that we can omit the 'almost' in this statement (see Lemma 5.6 and the proof of Theorem 1.2 at the end of this section). The use of induction to obtain such exact results from approximate ones has been a useful tool in the past (already used in [17]), but involves obstacles which are specific to the problem at hand. Lemma 5.6 relies on several simple observations about the typical structure of almost k-partite oriented graphs and digraphs (see Lemmas 5.2, 5.4 and 5.5).

Recall that $t_k(n)$ denotes the maximum number of edges in a k-partite (undirected) graph on n vertices, i.e. the number of edges in the k-partite Turán graph on n vertices. We say that a k-partition of vertices is *balanced* if the sizes of any two partition classes differ by at most one. Given a k-partition Q of $[n]$ with partition classes $V_1, \ldots, V_k$, and a graph, oriented graph or digraph $G = (V, E)$ on vertex set $[n]$, and an edge $e = uv \in E$ with $u \in V_i$ and $v \in V_j$, we call e a *crossing edge* if $i \neq j$. In Lemma 5.1 below we give upper and lower bounds on $T(n, k)$ and $T^*(n, k)$, in terms of $t_k(n)$ (recall that $T(n, k)$ and $T^*(n, k)$ were defined in Section 2). Lemma 5.1 is used in the proof of Theorem 1.2.

Lemma 5.1. *Let $k \geq 2$. For sufficiently large n we have the following:*

(i) $\frac{k^n 3^{t_k(n)}}{2k!n^{k-1}} \leq \frac{1}{2k!}\left(\begin{smallmatrix} n \\ \lfloor \frac{n}{k} \rfloor, \ldots, \lfloor \frac{n+k-1}{k} \rfloor \end{smallmatrix}\right) 3^{t_k(n)} < T(n, k) < k^n 3^{t_k(n)}.$

(ii) $\frac{k^n 4^{t_k(n)}}{2k!n^{k-1}} < T^*(n, k) < k^n 4^{t_k(n)}.$

Proof. We only prove (i) here; the proof of (ii) is similar. For the upper bound note that k^n counts the number of ordered k-partitions of $[n]$, and that for each such k-partition Q the number of oriented graphs for which every edge is a crossing edge with respect to Q is at most $3^{t_k(n)}$.

For the lower bound we will count the number of (unordered) balanced k-partitions. Each such k-partition gives rise to $3^{t_k(n)}$ k-partite oriented graphs. Since the vertex classes of a balanced k-partition of $[n]$ have sizes $\lfloor \frac{n}{k} \rfloor, \ldots, \lfloor \frac{n+k-1}{k} \rfloor$, the number of such k-partitions is

$$\frac{1}{k!}\left(\begin{matrix} n \\ \lfloor \frac{n}{k} \rfloor, \ldots, \lfloor \frac{n+k-1}{k} \rfloor \end{matrix}\right).$$

We now show that for any given balanced k-partition Q, almost all k-partite oriented graphs for which Q is a k-partition have no other possible k-partitions. Given a balanced k-partition Q of $[n]$ with partition classes $A_1, \ldots, A_k$, consider a random oriented graph where for each potential crossing edge we choose the edge to be either directed one way, directed the other way, or not present, each with probability $1/3$, independently. So each k-partite oriented graph for which Q is a k-partition is equally likely to be generated. Given a set of vertices A in a digraph G, we define their common out-neighbourhood $N^+(A) := \bigcap_{v \in A} N_G^+(v)$. By Theorem 2.1 we have that almost all graphs in the probability space satisfy the following:

(α) whenever $\ell \leq k$ and $i \in [k]$ and $v_1, \ldots, v_\ell \in V(G) \setminus A_i$, we have that

$$|N^+(\{v_1, \ldots, v_\ell\}) \cap A_i| \geq (n/k)(1/3)^{\ell+1}.$$

We now claim that if a k-partite oriented graph G has k-partition Q and satisfies (α) then Q is the unique k-partition of G. Indeed, suppose that Q' is a k-partition of G with vertex classes $A'_1, \ldots, A'_k$. We will show that $Q' = Q$. Consider any k vertices $v_1, \ldots, v_k$ that are such that $G[\{v_1, \ldots, v_k\}]$ is a transitive tournament. Such a set of k vertices exists by (α). Clearly no two of these vertices can be in the same vertex class of Q or Q'. Without loss of generality let us assume that $v_i \in A_i$ and $v_i \in A'_i$ for every $i \in [k]$. Define $N_i := N^+(\{v_1, \ldots, v_k\} \setminus \{v_i\})$. Since N_i is the common out-neighbourhood of $\{v_1, \ldots, v_k\} \setminus \{v_i\}$ it must be that N_i is a subset of A_i and a subset of A'_i. Note that Q and Q' agree on all vertices so far assigned to a partition class of Q'. Now consider any vertex w not yet assigned to a partition class of Q', and suppose $w \in A_j$ for some $j \in [k]$. For every $i \in [k]$ with $i \neq j$ we have by (α) that

$$|N^+(w) \cap N_i| = |N^+(\{w, v_1, \ldots, v_k\} \setminus \{v_i\}) \cap A_i| \geq (n/k)(1/3)^{k+1} \geq 1.$$

This together with the previous observation that $N_i \subseteq A'_i$ implies that $w \notin A'_i$. So $w \in A'_j$. Since w was an arbitrary unassigned vertex we have that $A_i = A'_i$ for every $i \in [k]$, and so $Q = Q'$, which implies the claim. This completes the proof of the middle inequality in Lemma 5.1.

To prove the first inequality note that if $a_1 + \cdots + a_k = n$ then $\binom{n}{a_1,\ldots,a_k}$ is maximised by taking $a_j := \left\lfloor \frac{n+j-1}{k} \right\rfloor$ for every $j \in [k]$. This implies that

$$k^n = \sum_{a_1 + \cdots + a_k = n} \binom{n}{a_1,\ldots,a_k} \leq n^{k-1} \binom{n}{\lfloor \frac{n}{k} \rfloor, \ldots, \lfloor \frac{n+k-1}{k} \rfloor},$$

which in turn implies the first inequality in Lemma 5.1, and hence completes the proof. $\square$

For a given oriented graph or digraph G on vertex set $[n]$ we call a k-partition Q of $[n]$ *optimal* if the number of non-crossing edges in G with respect to Q is at most the number of non-crossing edges in G with respect to Q' for every other k-partition Q' of $[n]$.

Given $k \geq 2$ and $\eta > 0$ we define $F(n, T_{k+1}, \eta)$ to be the set of all labelled T_{k+1}-free oriented graphs on n vertices that have at most ηn^2 non-crossing edges in an optimal k-partition. We define $F_Q(n, T_{k+1}, \eta) \subseteq F(n, T_{k+1}, \eta)$ to be the set of all such oriented graphs for which Q is an optimal k-partition. Similarly, we define $F^*(n, T_{k+1}, \eta)$ to be the set of all labelled T_{k+1}-free digraphs on n vertices that have at most ηn^2 non-crossing edges in an optimal k-partition, and we define $F_Q^*(n, T_{k+1}, \eta) \subseteq F^*(n, T_{k+1}, \eta)$ to be the set of all such digraphs for which Q is an optimal k-partition. Define

$$f(n, T_{k+1}, \eta) := |F(n, T_{k+1}, \eta)| \quad \text{and} \quad f_Q(n, T_{k+1}, \eta) := |F_Q(n, T_{k+1}, \eta)|,$$

and similarly

$$f^*(n, T_{k+1}, \eta) := |F^*(n, T_{k+1}, \eta)| \quad \text{and} \quad f_Q^*(n, T_{k+1}, \eta) := |F_Q^*(n, T_{k+1}, \eta)|.$$

Then Lemma 4.5 implies that for every $\eta > 0$ there exists $\varepsilon' > 0$ such that

$$(5.1) \quad f(n, T_{k+1}) \leq f(n, T_{k+1}, \eta)(1 + 2^{-\varepsilon' n^2}) \quad \text{and} \quad f^*(n, T_{k+1}) \leq f^*(n, T_{k+1}, \eta)(1 + 2^{-\varepsilon' n^2})$$

for all sufficiently large n. (So $\varepsilon' = 2\varepsilon$, where ε is as given by Lemma 4.5.)

Given an oriented graph or digraph G on vertex set V and disjoint subsets $U, U' \subseteq V$ we let $\overrightarrow{e_G}(U, U')$ denote the number of edges in $E(G)$ directed from vertices in U to vertices in U'. For convenience we will sometimes write $\overrightarrow{e}(U, U')$ for $\overrightarrow{e_G}(U, U')$ if this creates no ambiguity. Given $k \in \mathbb{N}$, $\eta, \mu > 0$, and a k-partition Q of $[n]$ with vertex classes $A_1, \ldots, A_k$, we define $F_Q(n, \eta, \mu)$ (respectively $F_Q^*(n, \eta, \mu)$) to be the set of all labelled oriented (respectively directed) graphs on n vertices for which Q is an optimal k-partition and that satisfy the following:

(F1) the number of non-crossing edges with respect to Q is at most ηn^2,
(F2) if $U_i \subseteq A_i$ and $U_j \subseteq A_j$ with $|U_i|, |U_j| \geq \mu n$ for distinct $i, j \in [k]$, then $\overrightarrow{e}(U_i, U_j), \overrightarrow{e}(U_j, U_i) \geq |U_i||U_j|/6$,
(F3) $||A_i| - n/k| \leq \mu n$ for every $i \in [k]$.

Note that property (F2) is similar to the property that the bipartite graph on vertex classes A_i, A_j whose edges are directed from A_i to A_j is μ-regular of density at least $1/6$ (and similarly for edges directed from A_j to A_i) and that the 'reduced graph' R that has vertex set $\{A_1, \ldots, A_k\}$ and edges between pairs that are μ-regular of density at least $1/6$ is a complete digraph.

Define $F_Q(n, T_{k+1}, \eta, \mu)$ to be the set of all oriented graphs in $F_Q(n, \eta, \mu)$ that are T_{k+1}-free. Similarly define $F_Q^*(n, T_{k+1}, \eta, \mu)$ to be the set of all digraphs in $F_Q^*(n, \eta, \mu)$ that are T_{k+1}-free. Note that $F_Q(n, T_{k+1}, \eta, \mu) \subseteq F_Q^*(n, T_{k+1}, \eta, \mu)$. Define $f_Q(n, T_{k+1}, \eta, \mu) := |F_Q(n, T_{k+1}, \eta, \mu)|$ and $f_Q^*(n, T_{k+1}, \eta, \mu) := |F_Q^*(n, T_{k+1}, \eta, \mu)|$.

The next lemma shows that $f_Q(n, T_{k+1}, \eta)$ and $f_Q(n, T_{k+1}, \eta, \mu)$ are asymptotically equal for any k-partition Q and suitable parameter values, (and similarly for $f_Q^*(n, T_{k+1}, \eta)$ and $f_Q^*(n, T_{k+1}, \eta, \mu)$).

Lemma 5.2. *Let $k \geq 2$ and let $0 < \eta, \mu < 1$ be such that $\mu^2 \geq 24H(\eta)$. There exists an integer $n_0 = n_0(\mu, k)$ such that the following hold for all $n \geq n_0$ and for every k-partition Q of $[n]$:*

(i) $f_Q(n, T_{k+1}, \eta) - f_Q(n, T_{k+1}, \eta, \mu) \leq 3^{t_k(n) - \frac{\mu^2 n^2}{100}}$.

(ii) $f_Q^*(n, T_{k+1}, \eta) - f_Q^*(n, T_{k+1}, \eta, \mu) \leq 4^{t_k(n) - \frac{\mu^2 n^2}{100}}$.

Proof. We only prove (i) here; the proof of (ii) is similar. We choose n_0 such that $1/n_0 \ll \mu, 1/k$. We wish to count the number of $G \in F_Q(n, T_{k+1}, \eta) \setminus F_Q(n, T_{k+1}, \eta, \mu)$. Let Q have vertex classes $A_1, \ldots, A_k$. The number of ways that at most ηn^2 non-crossing edges can be placed is at most

$$\binom{n^2}{\leq \eta n^2} \overset{(2.1)}{\leq} 2^{H(\eta) n^2}.$$

If $||A_i| - n/k| > \mu n$ for some $i \in [k]$ then by Proposition 4.2 the number of possible crossing edges is at most

$$t_k(n) - \mu n \left(\frac{\mu n}{2} - k \right) \leq t_k(n) - \frac{\mu^2 n^2}{3}.$$

We can conclude that the number of $G \in F_Q(n, T_{k+1}, \eta) \setminus F_Q(n, T_{k+1}, \eta, \mu)$ that fail to satisfy (F3) is at most

$$2^{H(\eta) n^2} 3^{t_k(n) - \mu^2 n^2 / 3}.$$

Every $G \in F_Q(n, T_{k+1}, \eta) \setminus F_Q(n, T_{k+1}, \eta, \mu)$ that satisfies property (F3) must fail to satisfy property (F2). For a given choice of at most ηn^2 non-crossing edges, consider the random oriented graph H where for each possible crossing edge with respect to Q we choose the edge to be either directed in one direction, directed in the other direction, or not present, each with probability $1/3$, independently. Note that the total number of ways to choose the crossing edges is at most $3^{t_k(n)}$, and each possible configuration of crossing edges is equally likely. So an upper bound on the number of $G \in F_Q(n, T_{k+1}, \eta) \setminus F_Q(n, T_{k+1}, \eta, \mu)$ that fail to satisfy property (F2) is

$$2^{H(\eta) n^2} 3^{t_k(n)} \mathbb{P}(H \text{ on } A_1, \ldots, A_k \text{ fails to satisfy (F2)}).$$

Note that the number of choices for $U_i \subseteq A_i$ and $U_j \subseteq A_j$ as in property (F2) is at most $(2^n)^2$ and that $\mathbb{E}(\overrightarrow{e_H}(U_i, U_j)) = |U_i||U_j|/3 \geq \mu^2 n^2 / 3$. Hence by Theorem 2.1 we get that

$$\mathbb{P}(H \text{ on } A_1, \ldots, A_k \text{ fails to satisfy (F2)}) \leq (2^n)^2 \exp \left(-\frac{\mathbb{E}(\overrightarrow{e_H}(U_i, U_j))}{8} \right) \leq 2^{2n} \exp \left(-\frac{\mu^2 n^2}{24} \right).$$

So summing these upper bounds gives us that

$$f_Q(n, T_{k+1}, \eta) - f_Q(n, T_{k+1}, \eta, \mu) \leq 2^{H(\eta) n^2} 3^{t_k(n)} \left(3^{-\mu^2 n^2 / 3} + 2^{2n} e^{-\frac{\mu^2 n^2}{24}} \right)$$

$$\leq 3^{t_k(n)} 3^{-\frac{\mu^2 n^2}{24} (\log_3 e - \log_3 2)} 2^{2n+1} \leq 3^{t_k(n) - \frac{\mu^2 n^2}{100}},$$

where we use that $\mu^2 \geq 24 H(\eta)$ and that $1/n_0 \ll \mu$. $\qquad \square$

The following proposition allows us to find many disjoint copies of T_k in any graph in $F_Q^*(n, \eta, \mu)$. It will be useful in proving Lemmas 5.4 and 5.6. We omit the proof, since it amounts to embedding a small oriented subgraph into the μ-regular blow-up of a complete digraph which can be done greedily (see e.g. [15, Lemma 7.5.2] for the 'undirected' argument).

Proposition 5.3. *Let* $n, k \in \mathbb{N}$, *let* $\eta, \mu > 0$, *let* Q *be a* k-*partition of* $[n]$ *with vertex classes* $A_1, \ldots, A_k$, *and suppose* $G \in F_Q^*(n, \eta, \mu)$. *For every* $i \in [k]$ *let* $B_i \subseteq A_i$ *with* $|B_i| \geq 12^{k-2} \mu n$. *Let* σ *be a permutation of* $[k]$. *Then* G *contains a copy of* T_k *on vertices* $v_1, \ldots, v_k$ *where for all distinct* $i, j \in [k]$ *we have that* $v_i \in B_i$ *and that there is an edge from* v_i *to* v_j *if and only if* $\sigma(i) < \sigma(j)$.

We now show that in an optimal partition each vertex is contained in only a small number of non-crossing edges.

Lemma 5.4. *Let $n, k \geq 2$, let $\eta, \mu > 0$, let Q be a k-partition of $[n]$ with vertex classes $A_1, \ldots, A_k$, and suppose $G \in F_Q^*(n, T_{k+1}, \eta, \mu)$. Then for every $i \in [k]$ and every $x \in A_i$ we have that*

$$|N_{A_i}^+(x)| + |N_{A_i}^-(x)| \leq 12^{k-2} 2\mu n.$$

Proof. Suppose not, so that there exists $x \in A_i$, for some $i \in [k]$, such that $|N_{A_i}^+(x)| + |N_{A_i}^-(x)| > 12^{k-2} 2\mu n$. Since Q is an optimal k-partition of G, it must be that

$$|N_{A_j}^+(x)| + |N_{A_j}^-(x)| \geq |N_{A_i}^+(x)| + |N_{A_i}^-(x)| > 12^{k-2} 2\mu n$$

for every $j \in [k]$.

For every $j \in [k]$ define B_j to be $N_{A_j}^+(x)$ if $|N_{A_j}^+(x)| \geq |N_{A_j}^-(x)|$, and $N_{A_j}^-(x)$ otherwise. So $|B_j| \geq 12^{k-2}\mu n$. Let J^+ be the set of all $j \in [k]$ such that $B_j = N_{A_j}^+(x)$, and let $J^- := [k] \setminus J^+$. Fix a permutation σ of $[k]$ with the property that $\sigma(i) < \sigma(j)$ whenever $i \in J^-$ and $j \in J^+$. Now Proposition 5.3 implies that G contains a copy of T_k on vertices $v_1, \ldots, v_k$ where for all distinct $i, j \in [k]$ we have that $v_i \in B_i$ and that the edge between v_i and v_j is directed towards v_j if and only if $\sigma(i) < \sigma(j)$. By the definition of σ, x together with this copy of T_k forms a copy of T_{k+1}. This is a contradiction, since $G \in F_Q^*(n, T_{k+1}, \eta, \mu)$, and so this completes the proof. $\square$

The following result shows that an optimal partition of a graph does not change too much upon the removal of just two vertices from the graph.

Lemma 5.5. *Let $k \geq 2$, and let $0 < \mu < 1/(3k^2)^{12}$ and $0 < \eta < \mu^2/3$. There exists an integer $n_0 = n_0(\mu, k)$ such that the following holds for all $n \geq n_0$. Let Q be a partition of $[n]$ with vertex classes $A_1, \ldots, A_k$ and let x, y be distinct elements of A_1. Then there is a set $\mathcal{P}$ of k-partitions of $[n] \setminus \{x, y\}$, with $|\mathcal{P}| \leq e^{\mu^{2/3}n}$, such that, for every $G \in F_Q^*(n, T_{k+1}, \eta, \mu)$, every optimal k-partition of $G - \{x, y\}$ is an element of $\mathcal{P}$.*

Proof. First note that, for any $G \in F_Q^*(n, T_{k+1}, \eta, \mu)$, we have by definition that the number of non-crossing edges in G with respect to Q is at most ηn^2. So certainly the number of non-crossing edges in $G - \{x, y\}$ with respect to the partition $A_1 \setminus \{x, y\}, A_2 \ldots, A_k$ is at most ηn^2.

Consider an arbitrary k-partition $B_1, \ldots, B_k$ of $[n] \setminus \{x, y\}$. We claim that if there exists $i \in [k]$ and distinct $j, j' \in [k]$ such that $|A_j \cap B_i|, |A_{j'} \cap B_i| \geq \mu n$, then for any $G \in F_Q^*(n, T_{k+1}, \eta, \mu)$ the number of non-crossing edges in $G - \{x, y\}$ with respect to the partition $B_1, \ldots, B_k$ is larger than ηn^2 (and hence $B_1, \ldots, B_k$ cannot be an optimal k-partition of $G - \{x, y\}$). Indeed, if we find such i, j, j' then by (F2) we have that the number of non-crossing edges in $G - \{x, y\}$ with respect to the partition $B_1, \ldots, B_k$ is at least

$$e_G(B_i) \geq e_G(A_j \cap B_i, A_{j'} \cap B_i) \geq 2 \cdot \frac{1}{6}(\mu n)^2 > \eta n^2,$$

which proves the claim.

We let $\mathcal{P}$ be the set of all k-partitions of $[n] \setminus \{x, y\}$ for which no such i, j, j' exist. So by the above claim we have that for every $G \in F_Q^*(n, T_{k+1}, \eta, \mu)$, every optimal k-partition of $G - \{x, y\}$ is an element of $\mathcal{P}$. So it remains to show that $|\mathcal{P}| \leq e^{\mu^{2/3}n}$. Consider an element of $\mathcal{P}$ with partition classes $B_1, \ldots, B_k$. For every $i \in [k]$, let $S_i := \{j : |A_j \cap B_i| \geq \mu n\}$. Note that for every $i \in [k]$ we have that $|S_i| \leq 1$, by definition of $\mathcal{P}$. Note also that $|A_j| \geq n/k - \mu n > k\mu n$ for every $j \in [k]$, and thus for every $i \in [k]$ we have that $|S_i| = 1$. Let $A_1' := A_1 \setminus \{x, y\}$ and let $A_i' := A_i$ for every $i \in \{2, \ldots, k\}$. So every element of $\mathcal{P}$ can be obtained by starting with the k-partition $A_1', \ldots, A_k'$, applying a permutation of $[k]$ to the partition class labels, and then for every ordered pair of partition classes moving at most μn elements from the first partition class to the second.

Hence, since $|A_j| \leq n/k + \mu n \leq 2n/k$, we have that

$$|\mathcal{P}| \leq k! \left(\binom{2n/k}{\leq \mu n}^{k-1} \right)^k \leq k! \left(\mu n \left(\frac{2en/k}{\mu n} \right)^{\mu n} \right)^{k(k-1)} \leq k!(\mu n)^{k^2} \left(\frac{1}{\mu^2} \right)^{\mu k^2 n} \leq e^{\mu^{2/3} n},$$

as required. $\qquad \square$

Define $F'_Q(n, T_{k+1}, \eta)$ to be the set of all oriented graphs in $F_Q(n, T_{k+1}, \eta)$ that have at least one non-crossing edge with respect to Q. Define $f'_Q(n, T_{k+1}, \eta) := |F'_Q(n, T_{k+1}, \eta)|$. Similarly define $F^{*'}_Q(n, T_{k+1}, \eta)$ to be the set of all digraphs in $F^*_Q(n, T_{k+1}, \eta)$ that have at least one non-crossing edge with respect to Q, and define $f^{*'}_Q(n, T_{k+1}, \eta) := |F^{*'}_Q(n, T_{k+1}, \eta)|$. In the following result we use Lemmas 4.5, 5.2, 5.4 and 5.5 to give upper bounds on $f'_Q(n, T_{k+1}, \eta)$ and $f^{*'}_Q(n, T_{k+1}, \eta)$ for any k-partition Q and suitable parameter values.

Lemma 5.6. *For all $k \geq 2$ there exist $\eta > 0$ and $C \in \mathbb{N}$ such that for all $n \in \mathbb{N}$ and all k-partitions Q of $[n]$ the following hold.*
  (i) $f'_Q(n, T_{k+1}, \eta) \leq 3^{t_k(n)} C 2^{-\eta n}$.
  (ii) $f^{*'}_Q(n, T_{k+1}, \eta) \leq 4^{t_k(n)} C 2^{-\eta n}$.

Proof. We only prove (i) here; the proof of (ii) is similar. Choose $C, n_0 \in \mathbb{N}$ and $\varepsilon, \eta, \mu > 0$ such that

$$1/C \ll 1/n_0 \ll \varepsilon \ll \eta \ll \mu \ll 1/k.$$

Define $F_Q(n, T_{k+1})$ to be the set of all T_{k+1}-free oriented graphs on n vertices for which Q is an optimal k-partition, and define $f_Q(n, T_{k+1}) = |F_Q(n, T_{k+1})|$.

The proof proceeds by induction on n. In fact, in addition to (i) we will inductively show that

$$(5.2) \qquad f_Q(n, T_{k+1}) \leq 3^{t_k(n)}(1 + C 2^{-\eta n}).$$

The result holds trivially for $n < n_0$ since $1/C \ll 1/n_0$. So let $n \geq n_0$ and let us assume that for every k-partition Q' of $[n-2]$ we have that

$$(5.3) \qquad f_{Q'}(n-2, T_{k+1}) \leq 3^{t_k(n-2)}(1 + C 2^{-\eta(n-2)}).$$

Let Q have partition classes $A_1, \ldots, A_k$. Define $F'_Q(n, T_{k+1}, \eta, \mu)$ to be the set of all graphs in $F_Q(n, T_{k+1}, \eta, \mu)$ that have at least one non-crossing edge with respect to Q. Define $f'_Q(n, T_{k+1}, \eta, \mu) := |F'_Q(n, T_{k+1}, \eta, \mu)|$. We will first find an upper bound for $f'_Q(n, T_{k+1}, \eta, \mu)$. We will find this bound in four steps. Note that (F3) implies that $f'_Q(n, T_{k+1}, \eta, \mu) = 0$ unless $||A_i| - n/k| \leq \mu n$, so we may assume that this inequality holds.

Step 1: Let B_1 be the number of ways to choose a single non-crossing edge xy with respect to Q. Then $B_1 \leq n^2$. Let A_i be the partition class of Q containing x and y.

Step 2: Let B_2 be the number of ways to choose the edges that do not have an endpoint in $\{x, y\}$. By Lemma 5.5 there is a set $\mathcal{P}$ of k-partitions of $[n] \setminus \{x, y\}$, with $|\mathcal{P}| \leq e^{\mu^{2/3} n}$, such that, for every $G \in F_Q(n, T_{k+1}, \eta, \mu)$, every optimal k-partition of $G - \{x, y\}$ is an element of $\mathcal{P}$. So we have by our inductive hypothesis that

$$B_2 \leq \sum_{Q' \in \mathcal{P}} f_{Q'}(n-2, T_{k+1}) \overset{(5.3)}{\leq} e^{\mu^{2/3} n} 3^{t_k(n-2)}(1 + C 2^{-\eta(n-2)}) \leq 3^{t_k(n-2)} C e^{\mu^{1/2} n}.$$

Step 3: Let B_3 be the number of possible ways to construct the edges between x, y and the vertices outside A_i. Let U be the set of edges chosen in Step 2. One can view U as a subset of the edge set of a graph G in $F'_Q(n, T_{k+1}, \eta, \mu)$. Let U' be the subset of U consisting of all those edges

in U that do not have an endpoint in A_i. So U' can be viewed as the edge set of a subgraph G' of G with $G' \in F_{\tilde{Q}}(n - |A_i|, 5\eta, 3\mu)$, where the set of partition classes of $\tilde{Q}$ is $\{A_1, \ldots, A_k\} \setminus \{A_i\}$. By repeatedly applying Proposition 5.3 to G' we can find at least $n/k - \mu n - 12^{k-3}3\mu n$ vertex-disjoint copies of T_{k-1} in G', each with precisely one vertex in each of the A_j for $j \neq i$. Consider the $2(k-1)$ potential edges between x, y and the vertices of any such T_{k-1}. If we wish for our graph to remain T_{k+1}-free then not all of the possible $3^{2(k-1)}$ sets of such edges are allowed. So since the number of vertices outside A_i not contained in one of these T_{k-1} is at most $(k-1)(2\mu n + 12^{k-3}3\mu n) \leq \mu^{1/2} n \log_3 e/2$, we have that

$$B_3 \leq \left(3^{2(k-1)} - 1\right)^{n/k} 3^{2\left(\mu^{1/2} n \log_3 e/2\right)} < \left(3^{2(k-1)}\left(1 - 3^{-2k}\right)\right)^{n/k} e^{\mu^{1/2} n} \leq 3^{2\frac{k-1}{k}n} e^{-\frac{n}{9^k k}} e^{\mu^{1/2} n}.$$

Step 4: Let B_4 be the number of possible ways to construct the edges between x, y and the other vertices in A_i. Note that by Lemma 5.4, x and y each have at most $12^{k-2}2\mu n$ neighbours inside A_i, and for each of these the edge between them may be oriented in either direction. So since $|A_i| \leq n$, we have that

$$B_4 \leq \binom{n}{\leq 12^{k-2}2\mu n}^2 \left(2^{12^{k-2}2\mu n}\right)^2 \leq \binom{n}{12^{k-2}2\mu n}^2 \left(2^{12^{k-2}2\mu n}\right)^2 \left(2^{12^{k-2}2\mu n}\right)^2$$

$$\leq \left(\frac{2e}{12^{k-2}\mu}\right)^{12^{k-2}4\mu n} \leq e^{\mu^{1/2} n}.$$

In Steps 1–4 we have considered all possible edges, and so $f'_Q(n, T_{k+1}, \eta, \mu) \leq B_1 \cdot B_2 \cdot B_3 \cdot B_4$. Together with the fact that $t_k(n) \geq t_k(n-2) + 2((k-1)/k)(n-2)$ this implies that

$$(5.4) \qquad f'_Q(n, T_{k+1}, \eta, \mu) \leq n^2 3^{t_k(n-2)} C 3^{2((k-1)/k)n} e^{3\mu^{1/2}n} e^{-n/(k9^k)}$$

$$\leq 3^{t_k(n)} C e^{-n/(2k9^k)} \leq 3^{t_k(n)} C 2^{-3\eta n}.$$

Now, note that since $F_Q(n, T_{k+1}, \eta, \mu), F'_Q(n, T_{k+1}, \eta) \subseteq F_Q(n, T_{k+1}, \eta)$ we have that

$$f'_Q(n, T_{k+1}, \eta) = |F'_Q(n, T_{k+1}, \eta) \cap F_Q(n, T_{k+1}, \eta, \mu)| + |F'_Q(n, T_{k+1}, \eta) \setminus F_Q(n, T_{k+1}, \eta, \mu)|$$

$$= |F'_Q(n, T_{k+1}, \eta, \mu)| + |F'_Q(n, T_{k+1}, \eta) \setminus F_Q(n, T_{k+1}, \eta, \mu)|$$

$$\leq |F'_Q(n, T_{k+1}, \eta, \mu)| + |F_Q(n, T_{k+1}, \eta) \setminus F_Q(n, T_{k+1}, \eta, \mu)|$$

$$= f'_Q(n, T_{k+1}, \eta, \mu) + (f_Q(n, T_{k+1}, \eta) - f_Q(n, T_{k+1}, \eta, \mu)).$$

This together with Lemma 5.2(i) gives us that

$$f'_Q(n, T_{k+1}, \eta) \leq f'_Q(n, T_{k+1}, \eta, \mu) + 3^{t_k(n) - \frac{\mu^2 n^2}{100}} \overset{(5.4)}{\leq} 3^{t_k(n)} C 2^{-\eta n},$$

which proves (i). So it remains to prove (5.2).

Note that the number of graphs in $F_Q(n, T_{k+1}, \eta, \mu)$ for which every edge is a crossing edge with respect to Q is at most $3^{t_k(n)}$. Since $f_Q(n, T_{k+1}, \eta, \mu) - f'_Q(n, T_{k+1}, \eta, \mu)$ is precisely the number of such graphs, we have that

$$f_Q(n, T_{k+1}, \eta, \mu) - f'_Q(n, T_{k+1}, \eta, \mu) \leq 3^{t_k(n)}.$$

This together with (5.4) implies that

$$f_Q(n, T_{k+1}, \eta, \mu) \leq 3^{t_k(n)}\left(1 + C2^{-3\eta n}\right).$$

Together with Lemma 5.2(i) this implies that

$$(5.5) \qquad f_Q(n, T_{k+1}, \eta) \leq f_Q(n, T_{k+1}, \eta, \mu) + 3^{t_k(n) - \frac{\mu^2 n^2}{100}} \leq 3^{t_k(n)}\left(1 + C2^{-2\eta n}\right).$$

On the other hand, Lemma 4.5(i) implies that

$$f(n, T_{k+1}) - f(n, T_{k+1}, \eta) \;\overset{}{\leq}\; f(n, T_{k+1})2^{-\varepsilon n^2} \overset{(5.1)}{\leq} 2f(n, T_{k+1}, \eta)2^{-\varepsilon n^2}$$
$$\overset{(5.5)}{\leq} 2k^n 3^{t_k(n)} \left(1 + C2^{-2\eta n}\right) 2^{-\varepsilon n^2} \leq 3^{t_k(n)} C2^{-2\eta n}.$$

Now this together with (5.5) implies that

$$f_Q(n, T_{k+1}) \leq f_Q(n, T_{k+1}, \eta) + (f(n, T_{k+1}) - f(n, T_{k+1}, \eta)) \leq 3^{t_k(n)} \left(1 + C2^{-\eta n}\right).$$

This completes the proof of (5.2). □

We can now finally prove Theorem 1.2 using Lemmas 5.1 and 5.6 together with the bounds in (5.1).

Proof of Theorem 1.2. We only prove (i) here; the proof of (ii) is similar. Let η, C be given by Lemma 5.6 and choose n_0 and ε such that $1/n_0 \ll \varepsilon \ll \eta, 1/k$. Consider any $n \geq n_0$ and let $\mathcal{Q}$ be the set of all k-partitions of $[n]$. Note that

$$f'(n, T_{k+1}, \eta) \leq \sum_{Q \in \mathcal{Q}} f'_Q(n, T_{k+1}, \eta).$$

So by Lemma 5.6(i) and the fact that $|\mathcal{Q}| \leq k^n$ we have that

$$f'(n, T_{k+1}, \eta) \leq k^n 3^{t_k(n)} C2^{-\eta n}.$$

Recall from (5.1) that

$$f(n, T_{k+1}) \leq f(n, T_{k+1}, \eta)(1 + 2^{-\varepsilon n^2}).$$

Together with the fact that $f(n, T_{k+1}, \eta) = f'(n, T_{k+1}, \eta) + T(n, k)$ and the upper bound in Lemma 5.1(i), this implies that

$$f(n, T_{k+1}) - T(n, k) \leq f'(n, T_{k+1}, \eta) + f(n, T_{k+1}, \eta)2^{-\varepsilon n^2}$$
$$= f'(n, T_{k+1}, \eta) + (f'(n, T_{k+1}, \eta) + T(n, k))2^{-\varepsilon n^2}$$
$$\leq k^n 3^{t_k(n)} C2^{-\eta n} + k^n 3^{t_k(n)}(1 + C2^{-\eta n})2^{-\varepsilon n^2}.$$

Now the lower bound in Lemma 5.1(i) gives us that $f(n, T_{k+1}) - T(n, k) = o(T(n, k))$. So $f(n, T_{k+1}) = T(n, k)(1 + o(1))$, as required. □

6. ROUGH STRUCTURE OF TYPICAL C_k-FREE ORIENTED GRAPHS AND DIGRAPHS

In this section we prove several stability results for C_k-free digraphs (Lemmas 6.4, 6.5 and 6.6). These are used (together with Theorem 3.3) at the end of the section to determine the 'rough' structure of typical C_k-free oriented graphs and digraphs.

We will make use of the following definitions. For disjoint sets of vertices A, A', we define $\overrightarrow{K}(A, A')$ to be the oriented graph on vertex set $A \cup A'$ with edge set consisting of all the $|A||A'|$ edges that are directed from A to A'. Given a digraph G, $A \subseteq V(G)$ and $x \in V(G) \setminus A$, we say that $G[A, \{x\}]$ is an *in-star* if $G[A, \{x\}] = \overrightarrow{K}(A, \{x\})$, and we say that $G[A, \{x\}]$ is an *out-star* if $G[A, \{x\}] = \overrightarrow{K}(\{x\}, A)$. The following proposition will prove useful to us many times in this section.

Proposition 6.1. *Let $a \in \mathbb{R}$ with $1 \leq a \leq 2$, let $k \in \mathbb{N}$ with $k \geq 2$ and let G be a C_{k+1}-free digraph. Suppose G contains a copy C of C_k with vertex set $A \subseteq V(G)$, and let $x \in V(G) \setminus A$. Then the following hold:*

(i) $e_a(A, \{x\}) \leq k$,
(ii) *if $G[A, \{x\}]$ is not an in-star or an out-star, then $e_a(A, \{x\}) \leq k - 2 + a$,*

(iii) *if $G[A, \{x\}]$ is not an in-star or an out-star, and contains no double edges, then $e_a(A, \{x\}) \leq k - 1$.*

Suppose moreover that for some $\ell \in \{k - 1, k\}$, G contains a copy C' of C_ℓ with vertex set $A' \subseteq V(G)$, where $A \cap A' = \emptyset$. Then the following hold:

(iv) $e_a(A, A') \leq k\ell$,

 (v) *if $G[A, A'] \notin \{\overrightarrow{K}(A, A'), \overrightarrow{K}(A', A)\}$, then $e_a(A, A') \leq k\ell - 2 + a$,*

(vi) *if $G[A, A'] \notin \{\overrightarrow{K}(A, A'), \overrightarrow{K}(A', A)\}$, and moreover $G[A, A']$ contains no double edges, then $e_a(A, A') \leq k\ell - 1$.*

Proof. Write $C = v_1 v_2 \ldots v_k$. For $i \in [k]$ let $Q_i := \{v_i x, x v_{i+1}\}$, where $v_{k+1} := v_1$. Since G is C_{k+1}-free we have that $|E(G) \cap Q_i| \leq 1$ for every $i \in [k]$. Hence $e(A, \{x\}) \leq k$. We can now prove (i)–(vi).

 (i) This follows since $e_a(A, \{x\}) \leq e(A, \{x\}) \leq k$.

 (ii) Suppose that $G[A, \{x\}]$ is not an in-star or an out-star. Note that if for some $j \in [k]$ we have that $E(G) \cap Q_j = \emptyset$ then, since $|E(G) \cap Q_i| \leq 1$ for all $i \in [k]$, $e_a(A, \{x\}) \leq e(A, \{x\}) \leq k - 1 \leq k - 2 + a$ as required. So we may assume that $|E(G) \cap Q_i| = 1$ for every $i \in [k]$. Since $G[A, \{x\}]$ is not an in-star or an out-star, there exists some $j \in [k]$ such that $E(G) \cap Q_j = \{x v_{j+1}\}$ and $E(G) \cap Q_{j+1} = \{v_{j+1} x\}$; that is, there exists a double edge in $G[A, \{x\}]$. So since $e(G[A, \{x\}]) \leq k$ we have that $e_a(A, \{x\}) \leq k - 2 + a$, as required.

(iii) Suppose that $G[A, \{x\}]$ is not an in-star or an out-star, and contains no double edges. Just as in the proof of (ii), we have that if $|E(G) \cap Q_i| = 1$ for every $i \in [k]$ then there exists a double edge in $G[A, \{x\}]$. So we may assume that for some $j \in [k]$ we have that $E(G) \cap Q_j = \emptyset$. This implies that $e_a(A, \{x\}) \leq e(A, \{x\}) \leq k - 1$, as required.

(iv) This is immediate from (i).

 (v) Suppose that $G[A, A'] \notin \{\overrightarrow{K}(A, A'), \overrightarrow{K}(A', A)\}$. We claim that there exists $x \in A'$ such that $G[A, \{x\}]$ is not an in-star or an out-star. Indeed, suppose not. Then since $G[A, A'] \notin \{\overrightarrow{K}(A, A'), \overrightarrow{K}(A', A)\}$, there must exist distinct vertices $y', z' \in A'$ such that $G[A, \{y'\}]$ is an in-star and $G[A, \{z'\}]$ is an out-star. Let P' be the subpath of C' from y' to z'. Thus P' has length s for some $1 \leq s \leq k - 1$. Let $y, z \in A$ be not necessarily distinct vertices such that the subpath P of C from y to z has length $k - s - 1$. Then $y P z y' P' z' y$ is a copy of C_{k+1} in G, which contradicts the assumption that G is C_{k+1}-free. Hence there does exist $x \in A$ such that $G[A, \{x\}]$ is not an in-star or an out-star. So by (i) and (ii) we have that $e_a(A, A') \leq k\ell - 2 + a$ as required.

(vi) This proof is almost identical to that of (v), just using (iii) instead of (ii), and so is omitted. $\qquad\square$

For $k \in \mathbb{N}$ define $T_{n,k}^+$ (up to isomorphism) to be the digraph on vertex set $[n]$ with all edges ij where $i < j$ and all edges ji where $i < j$ and $\lfloor (i-1)/k \rfloor = \lfloor (j-1)/k \rfloor$. So if $n = sk$ for some $s \in \mathbb{N}$ then $T_{n,k}^+$ is obtained from T_s by blowing up each vertex to a copy of the complete digraph DK_k. Note that $T_{n,k}^+$ is C_{k+1}-free, and for all $a \in \mathbb{R}$ with $1 \leq a \leq 2$ we have that

$$(6.1) \qquad e_a(T_{n,k}^+) = \binom{n}{2} + \left\lfloor \frac{n}{k} \right\rfloor \binom{k}{2}(a-1) + \binom{n - k\left\lfloor \frac{n}{k} \right\rfloor}{2}(a-1).$$

We will first show that $T_{n,k}^+$ is an extremal digraph for C_{k+1}. The resulting formula for $\mathrm{ex}_a(n, C_{k+1})$ will be used in the proofs of Lemmas 6.4 and 6.5, but we will not refer to $T_{n,k}^+$ itself again. Note that the case $a = 2$ of Lemma 6.2 corresponds to finding the digraph Turán number of C_{k+1}, and is already due to Häggkvist and Thomassen [19].

Lemma 6.2. *Let $a \in \mathbb{R}$ with $1 \leq a \leq 2$ and let $k \in \mathbb{N}$. Then*

$$\mathrm{ex}_a(n, C_{k+1}) = e_a(T_{n,k}^+).$$

Proof.   The proof proceeds by induction on k. The case $k = 1$ is trivial. So suppose G is a C_{k+1}-free digraph on n vertices for some $k > 1$. The proof now proceeds by induction on n. The cases $n = 1, \ldots, k$ are trivial. So suppose $n > k$. Note that if G is also C_k-free then by our inductive hypothesis on k,

$$e_a(G) \leq e_a(T_{n,k-1}^+) \leq e_a(T_{n,k}^+).$$

Otherwise, G contains a copy of C_k, say on vertex set $A \subseteq V(G)$. So by Proposition 6.1(i) we have that $e_a(A, \{x\}) \leq k$ for every $x \in V(G) \setminus A$. Hence $e_a(A, V(G) \setminus A) \leq k(n - k)$. Note that by our inductive hypothesis on n, $e_a(G[V(G) \setminus A]) \leq e_a(T_{n-k,k}^+)$. Hence,

$$
\begin{aligned}
e_a(G) &= e_a(G[V(G) \setminus A]) + e_a(A, V(G) \setminus A) + e_a(G[A]) \\
&\overset{(6.1)}{\leq} \binom{n-k}{2} + \left(\left\lfloor \frac{n-k}{k} \right\rfloor \binom{k}{2} + \binom{(n-k) - k \left\lfloor \frac{n-k}{k} \right\rfloor}{2} \right)(a-1) + k(n-k) + a\binom{k}{2} \\
&= \binom{n}{2} + \left\lfloor \frac{n}{k} \right\rfloor \binom{k}{2}(a-1) + \binom{n - k \left\lfloor \frac{n}{k} \right\rfloor}{2}(a-1) = e_a(T_{n,k}^+).
\end{aligned}
$$

So indeed $\mathrm{ex}_a(n, C_{k+1}) = e_a(T_{n,k}^+)$, as required. $\square$

We will next prove three stability results for C_{k+1}-free digraphs. The first will cover the case $1 \leq a < 2$ (where a, as usual, is the parameter in the definition of the weighted size of a digraph) and will be used to prove a structural result on C_{k+1}-free oriented graphs. The second covers the case $a = 2$ and k odd, and will be used to prove an analogous structural result on C_{k+1}-free digraphs for odd k. The third covers the case $a = 2$ and k even, and will be used to prove a (less restrictive) structural result on C_{k+1}-free digraphs for even k. The proofs of the first two of these stability results will make use of a result of Chudnovsky, Seymour and Sullivan [14], which we state below. To do so we need to introduce the following notation. Let $\beta(G)$ denote the number of backwards edges in G with respect to a transitive-optimal ordering of G. Let $\gamma(G)$ denote the number of unordered non-adjacent pairs of vertices in G; that is unordered pairs u, v of vertices such that $uv \notin E(G)$ and $vu \notin E(G)$.

Lemma 6.3. [14] *Let G be a $\{C_2, C_3\}$-free digraph. Then $\beta(G) \leq \gamma(G)$.*

It is conjectured in [14] that in fact $\beta(G) \leq \gamma(G)/2$ for all $\{C_2, C_3\}$-free digraphs G. If true, this would be best possible.

Lemma 6.4 (Stability when $a < 2$)**.** *Let $a \in \mathbb{R}$ with $1 \leq a < 2$ and let $k \in \mathbb{N}$ with $k \geq 2$. Then for all $\varepsilon > 0$ there exist $\delta > 0$ and $n_0 \in \mathbb{N}$ such that every C_{k+1}-free digraph G on $n \geq n_0$ vertices with*

$$e_a(G) \geq \binom{n}{2} - \delta n^2$$

satisfies $G = T_n \pm \varepsilon n^2$.

Proof.   We prove the lemma via the following claim.

Claim: *Let $k \in \mathbb{N}$ with $k \geq 2$ and let $\varepsilon > 0$. Suppose that there exist $\delta' > 0$ and $n_0' \in \mathbb{N}$ such that every $\{C_k, C_{k+1}\}$-free digraph G on $n' \geq n_0'$ vertices with*

$$e_a(G) \geq \binom{n'}{2} - \delta' n'^2$$

satisfies $G = T_{n'} \pm \varepsilon n'^2/(2k^2)$. *Then there exist* $\delta > 0$ *and* $n_0 \in \mathbb{N}$ *such that every* C_{k+1}-*free digraph* G *on* $n \geq n_0$ *vertices with*

$$e_a(G) \geq \binom{n}{2} - \delta n^2$$

satisfies $G = T_n \pm \varepsilon n^2$.

In order to check that the claim implies the lemma, we proceed by induction on k. For the base case $k = 2$ the assumption of the claim is satisfied, since if $\delta' := \varepsilon/(4k^2)$ and if G is a $\{C_2, C_3\}$-free digraph on n' vertices with $e_a(G) \geq \binom{n'}{2} - \delta' n'^2$ then $\gamma(G) \leq \delta' n'^2$, and so applying Lemma 6.3 yields the assumption of the claim. So the conclusion of the claim holds, which is precisely the statement of the lemma for $k = 2$. For $k > 2$ the assumption of the claim is satisfied by the inductive hypothesis (since any $\{C_k, C_{k+1}\}$-free digraph is certainly a C_k-free digraph) and so the conclusion of the claim holds, which is precisely the statement of the lemma for k. So by induction the lemma holds and we are done.

Thus it remains to prove the claim. (Note that, apart from in the base case $k = 2$, in the above argument it would suffice for the assumption in the statement of the claim to refer to C_k-free digraphs, rather than $\{C_k, C_{k+1}\}$-free digraphs. As such, this claim is stronger than strictly necessary for our purposes, since the assumption is weaker. However, this approach allows us to prove the base case at the same time as the inductive step, and so yields a shorter proof.)

Proof of claim: Choose δ and n_0 such that $1/n_0 \ll \delta \ll 1/k, 2 - a, \delta'$ and $1/n_0 \ll 1/n_0', \varepsilon$. Let G be a C_{k+1}-free digraph on $n \geq n_0$ vertices with

$$(6.2) \qquad\qquad\qquad e_a(G) \geq \binom{n}{2} - \delta n^2.$$

Let $t \geq 0$ denote the maximum number of vertex-disjoint copies of C_k in G. Let $\mathcal{C} = \{C^1, \ldots, C^t\}$ be a set of t vertex-disjoint copies of C_k in G. Let $V_1 := V(C^1) \cup \cdots \cup V(C^t)$ and $V_2 := V(G) \setminus V_1$. Let $n_1 := |V_1|$ and $n_2 := |V_2|$. Note that $G[V_2]$ is C_k-free.

Note that Proposition 6.1(i) implies that $e_a(V_1, V_2) \leq n_1 n_2$, since G is C_{k+1}-free. Also, for $i = 1, 2$, since $G[V_i]$ is C_{k+1}-free, Lemma 6.2 and (6.1) together imply that

$$e_a(G[V_i]) \leq \mathrm{ex}_a(n_i, C_{k+1}) = \binom{n_i}{2} + \left\lfloor \frac{n_i}{k} \right\rfloor \binom{k}{2}(a-1) + \binom{n_i - k\left\lfloor \frac{n_i}{k} \right\rfloor}{2}(a-1) \leq \binom{n_i}{2} + \delta n^2.$$

(The last inequality holds since $1/n_0 \ll \delta \ll 1/k$.) Together with (6.2) this implies that

$$(6.3) \qquad e_a(G[V_1]) = e_a(G) - e_a(V_1, V_2) - e_a(G[V_2])$$

$$\geq \left(\binom{n}{2} - \delta n^2 \right) - n_1 n_2 - \left(\binom{n_2}{2} + \delta n^2 \right) = \binom{n_1}{2} - 2\delta n^2,$$

and similarly that

$$(6.4) \qquad\qquad e_a(G[V_2]) = e_a(G) - e_a(V_1, V_2) - e_a(G[V_1]) \geq \binom{n_2}{2} - 2\delta n^2,$$

and that

$$(6.5) \qquad\qquad e_a(V_1, V_2) = e_a(G) - e_a(G[V_1]) - e_a(G[V_2]) \geq n_1 n_2 - 3\delta n^2.$$

We now consider the digraph G' defined on vertex set $[t] \cup V_2$ as follows. Firstly, $G'[V_2] := G[V_2]$. For vertices $i, j \in [t]$ we have that $ij \in E(G')$ if and only if $G[V(C^i), V(C^j)] = \overrightarrow{K}(V(C^i), V(C^j))$. For a vertex $x \in V_2$ and an element $i \in [t]$ we have that $ix \in E(G')$ if and only if $G[V(C^i), \{x\}]$ is an in-star and that $xi \in E(G')$ if and only if $G[V(C^i), \{x\}]$ is an out-star.

Note that by Proposition 6.1(iv), $e_a(G[V(C^i), V(C^j)]) \leq k^2$ (for all $i \neq j$). Moreover, Proposition 6.1(v) implies that if $i, j \in [t]$ and $ij, ji \notin E(G'[[t]])$ then $e_a(G[V(C^i), V(C^j)]) \leq k^2 - 2 + a$. Let $s := \binom{t}{2} - e_a(G'[[t]]) = \binom{t}{2} - e(G'[[t]])$. Then

$$(6.6) \qquad \binom{n_1}{2} - 2\delta n^2 \overset{(6.3)}{\leq} e_a(G[V_1]) = \sum_{\substack{i,j \in [t] \\ i < j}} e_a(G[V(C^i), V(C^j)]) + \sum_{i \in [t]} e_a(G[V(C^i)])$$

$$\leq \binom{n_1}{2} - s(2 - a) + t\binom{k}{2}(a - 1).$$

Thus $s(2 - a) \leq 3\delta n^2$, i.e.

$$e_a(G'[[t]]) \geq \binom{t}{2} - \frac{3\delta n^2}{2 - a}.$$

Similarly, Proposition 6.1(ii) implies that if $i \in [t]$ and $x \in V_2$ and $ix, xi \notin E(G'[[t]])$ then $e_a(G[V(C^i), \{x\}]) \leq k - 2 + a$. So we have that

$$e_a(G'[[t], V_2]) \overset{(6.5)}{\geq} tn_2 - \frac{3\delta n^2}{2 - a}.$$

So recalling that $e_a(G'[V_2]) = e_a(G[V_2])$ we have that

$$(6.7) \qquad e_a(G') = e_a(G'[[t]]) + e_a(G'[V_2]) + e_a(G'[[t], V_2]) \overset{(6.4)}{\geq} \binom{t + n_2}{2} - \frac{8\delta n^2}{2 - a}.$$

Since $t + n_2 \geq n/k$ we have that $t + n_2 \geq n_0'$ and that $8\delta n^2/(2 - a) \leq \delta'(t + n_2)^2$, and hence by (6.7) that

$$(6.8) \qquad e_a(G') \geq \binom{t + n_2}{2} - \delta'(t + n_2)^2.$$

We now claim that G' must be $\{C_k, C_{k+1}\}$-free. Indeed, suppose not. If G' contains a copy of C_{k+1} then it is clear that G also contains a copy of C_{k+1}, contradicting our assumption. So we may assume that G' contains a copy of C_k. Since $G'[V_2] = G[V_2]$ is C_k-free by construction, the vertex set of any copy of C_k in G' must contain some $i \in [t]$. But then G would clearly contain a copy of C_{k+1} using two of the vertices in $V(C^i)$, again contradicting our assumption that G is C_{k+1}-free. So G' is $\{C_k, C_{k+1}\}$-free, as claimed.

Thus by (6.8) and the assumption in the statement of the claim we have that $G' = T_{t+n_2} \pm \varepsilon(t + n_2)^2/(2k^2)$. Together with the definition of G' this implies that $G = T_n \pm \varepsilon n^2$, as required. This completes the proof of the claim, and hence completes the proof of the lemma. $\qquad \square$

The rough strategy of the next proof is similar to that of Lemma 6.4.

Lemma 6.5 (Stability when $a = 2$ and k is odd). *Let $k \in \mathbb{N}$ with $k \geq 3$ and k odd. Then for all $\varepsilon > 0$ there exist $\delta > 0$ and $n_0 \in \mathbb{N}$ such that every C_{k+1}-free digraph G on $n \geq n_0$ vertices with*

$$e(G) \geq \binom{n}{2} - \delta n^2$$

satisfies $G = T_n \pm \varepsilon n^2$.

Proof. We prove the lemma via the following claim.

Claim: *Let $k \in \mathbb{N}$ with $k \geq 3$ and k odd, and let $\varepsilon > 0$. Suppose that there exist $\delta' > 0$ and $n_0' \in \mathbb{N}$ such that every $\{C_{k-1}, C_k\}$-free digraph G on $n' \geq n_0'$ vertices with*

$$e(G) \geq \binom{n'}{2} - \delta' n'^2$$

satisfies $G = T_{n'} \pm \varepsilon n'^2/(2k^2)$. *Then there exist* $\delta > 0$ *and* $n_0 \in \mathbb{N}$ *such that every* C_{k+1}*-free digraph* G *on* $n \geq n_0$ *vertices with*

$$e(G) \geq \binom{n}{2} - \delta n^2$$

satisfies $G = T_n \pm \varepsilon n^2$.

In order to check that the claim implies the lemma, we proceed by induction on $\ell := (k+1)/2$. The argument is similar to that in the proof of Lemma 6.4. (As before, Lemma 6.3 implies that in the base case $\ell = 2$ of the induction, the assumption of the claim holds.)

Proof of claim: Choose δ and n_0 such that $1/n_0 \ll \delta \ll 1/k, \delta'$ and $1/n_0 \ll 1/n_0', \varepsilon$. Let G be a C_{k+1}-free digraph on $n \geq n_0$ vertices with

(6.9) $$e(G) \geq \binom{n}{2} - \delta n^2.$$

Let $t \geq 0$ denote the maximum number of vertex-disjoint copies of C_k in G. Let $\mathcal{C} = \{C^1, \ldots, C^t\}$ be a set of t vertex-disjoint copies of C_k in G. Let $V_1 := V(C^1) \cup \cdots \cup V(C^t)$ and $n_1 := |V_1|$. Now let $t^* \geq 0$ denote the maximum number of vertex-disjoint copies of C_{k-1} in $G[V \setminus V_1]$. Let $\mathcal{C}^* = \{C_*^1, \ldots, C_*^{t^*}\}$ be a set of t^* vertex-disjoint copies of C_{k-1} in $G[V \setminus V_1]$. Let $V_2 := V(C_*^1) \cup \cdots \cup V(C_*^{t^*})$ and $n_2 := |V_2|$. Let $V_3 := V(G) \setminus (V_1 \cup V_2)$ and $n_3 := |V_3|$. Note that $G[V_2 \cup V_3]$ is C_k-free and that $G[V_3]$ is $\{C_{k-1}, C_k\}$-free.

Proposition 6.1(i) implies that $e(V_1, V_2) \leq n_1 n_2$ and that $e(V_1, V_3) \leq n_1 n_3$, since G is C_{k+1}-free, and that $e(V_2, V_3) \leq n_2 n_3$, since $G[V_2 \cup V_3]$ is C_k-free. Also, similarly to the proof of Lemma 6.4 we use Lemma 6.2 to get that $e(G[V_i]) \leq \binom{n_i}{2} + \delta n^2$ for $i = 1, 2, 3$. Together with (6.9) this implies that

$$\begin{aligned}
e(G[V_1]) &= e(G) - e(V_1, V_2) - e(V_1, V_3) - e(V_2, V_3) - e(G[V_2]) - e(G[V_3]) \\
&\geq \left(\binom{n}{2} - \delta n^2\right) - n_1 n_2 - n_1 n_3 - n_2 n_3 - \left(\binom{n_2}{2} + \delta n^2\right) - \left(\binom{n_3}{2} + \delta n^2\right) \\
&= \binom{n_1}{2} - 3\delta n^2,
\end{aligned}$$

and similarly that $e(G[V_2]) \geq \binom{n_2}{2} - 3\delta n^2$, and that $e(G[V_3]) \geq \binom{n_3}{2} - 3\delta n^2$, and that $e(V_i, V_j) \geq n_i n_j - 4\delta n^2$ for all $i, j \in \{1, 2, 3\}$, $i \neq j$.

We now consider the digraph G' defined on vertex set $([t] \times \{0\}) \cup ([t^*] \times \{1\}) \cup V_3$ as follows. Firstly, for every vertex $v \in V(G')$ define $f(v)$ to be $\{v\}$ if $v \in V_3$, to be $V(C^i)$ if $v = (i, 0) \in [t] \times \{0\}$, and to be $V(C_*^i)$ if $v = (i, 1) \in [t^*] \times \{1\}$. Now let $G'[V_3] := G[V_3]$ and for vertices $u, v \in V(G')$ with $|\{u, v\} \cap V_3| \leq 1$ define $uv \in E(G')$ if and only if $G[f(u), f(v)] = \overrightarrow{K}(f(u), f(v))$.

Note that by the Kővári-Sós-Turán theorem, G contains at most δn^2 double edges, since G is C_{k+1}-free and $k + 1$ is even, and $1/n_0 \ll \delta \ll 1/k$. Note also that by Proposition 6.1(iv), if $u, v \in [t] \times \{0\}$ then $e(G[f(u), f(v)]) \leq k^2$. If, moreover, $uv, vu \notin E(G'[[t] \times \{0\}])$ and $G[f(u), f(v)]$ contains no double edge, then by Proposition 6.1(vi) we have that $e(G[f(u), f(v)]) \leq k^2 - 1$. Using that $e(G[V_1]) \geq \binom{n_1}{2} - 3\delta n^2$ one can now argue similarly as in (6.6) to see that

$$e(G'[[t] \times \{0\}]) \geq \binom{t}{2} - 5\delta n^2.$$

Also, Proposition 6.1(iv) implies that if $u \in [t] \times \{0\}$ and $v \in [t^*] \times \{1\}$ then $e(G[f(u), f(v)]) \leq k(k-1)$. If, moreover, $uv, vu \notin E(G'[[t] \times \{0\}, [t^*] \times \{1\}])$ and $G[f(u), f(v)]$ contains no double edge, then by Proposition 6.1(vi) we have that $e(G[f(u), f(v)]) \leq k(k-1) - 1$. Using that $e(V_1, V_2) \geq n_1 n_2 - 4\delta n^2$ one can again argue similarly as in (6.6) to see that

$$e(G'[[t] \times \{0\}, [t^*] \times \{1\}]) \geq tt^* - 5\delta n^2.$$

Furthermore, Proposition 6.1(i) implies that if $u \in [t] \times \{0\}$ and $v \in V_3$ then $e(G[f(u), f(v)]) \leq k$. If, moreover, $uv, vu \notin E(G'[[t] \times \{0\}, V_3])$ and $G[f(u), f(v)]$ contains no double edge, then by Proposition 6.1(iii) we have that $e(G[f(u), f(v)]) \leq k - 1$. Using that $e(V_1, V_3) \geq n_1 n_3 - 4\delta n^2$ one can again argue similarly as in (6.6) to see that

$$e(G'[[t] \times \{0\}, V_3]) \geq t n_3 - 5\delta n^2.$$

Using that $G[V_2 \cup V_3]$ is C_k-free, in a similar way we get that $e(G'[[t^*] \times \{1\}]) \geq \binom{t^*}{2} - 5\delta n^2$ and that $e(G'[[t^*] \times \{1\}, V_3]) \geq t^* n_3 - 5\delta n^2$. So recalling that $e(G'[V_3]) = e(G[V_3]) \geq \binom{n_3}{2} - 3\delta n^2$ we have that

$$
\begin{aligned}
(6.10) \qquad e(G') &= e(G'[[t] \times \{0\}]) + e(G'[[t^*] \times \{1\}]) + e(G'[V_3]) \\
&\quad + e(G'[[t] \times \{0\}, [t^*] \times \{1\}]) + e(G'[[t] \times \{0\}, V_3]) + e(G'[[t^*] \times \{1\}], V_3) \\
&\geq \binom{t + t^* + n_3}{2} - 28\delta n^2.
\end{aligned}
$$

Since $t + t^* + n_3 \geq n/k$, we have that $t + t^* + n_3 \geq n_0'$ and that $28\delta n^2 \leq \delta'(t + t^* + n_3)^2$, and hence by (6.10) that

$$(6.11) \qquad\qquad e(G') \geq \binom{t + t^* + n_3}{2} - \delta'(t + t^* + n_3)^2.$$

We now claim that G' must be $\{C_{k-1}, C_k\}$-free. Indeed, suppose not. If G' contains a copy of C_k then since $G'[V_3] = G[V_3]$ is C_k-free by construction, the vertex set of such a copy of C_k in G' must contain some $u \in ([t] \times \{0\}) \cup ([t^*] \times \{1\})$. But then G would clearly contain a copy of C_{k+1} using two of the vertices in $f(u)$, contradicting our assumption that G is C_{k+1}-free. Similarly, if G' contains a copy of C_{k-1} then since $G[V_3]$ is C_{k-1}-free by construction, the vertex set of such a copy of C_{k-1} in G' must contain some $u \in ([t] \times \{0\}) \cup ([t^*] \times \{1\})$. If there exists such a $u \in [t] \times \{0\}$ then G would clearly contain a copy of C_{k+1} using three of the vertices in $f(u)$, since $|f(u)| = k \geq 3$. Otherwise, there exists $u \in [t^*] \times \{1\}$ and a copy of C_{k-1} in G' that uses u but no vertices in $[t] \times \{0\}$. But then $G[V_2 \cup V_3]$ would clearly contain a copy of C_k using two of the vertices in $f(u)$, contradicting our previous observation that $G[V_2 \cup V_3]$ is C_k-free.

So G' is $\{C_{k-1}, C_k\}$-free, as claimed. Thus by (6.11) and the assumption in the statement of the claim we have that $G' = T_{t+t^*+n_3} \pm \varepsilon(t + t^* + n_3)^2/(2k^2)$. Together with the definition of G' this implies that $G = T_n \pm \varepsilon n^2$, as required. This completes the proof of the claim, and hence completes the proof of the lemma. $\qquad\square$

We now prove a digraph stability result for forbidden odd cycles. Here both the Turán and the stability results allow for a richer structure than in the previous two lemmas: For even k, a near extremal graph can be obtained from a transitive tournament by blowing up some of its vertices into complete bipartite graphs of arbitrary size (see Section 1 for the precise definition). This makes the proof more difficult than the previous two.

Lemma 6.6 (Stability when $a = 2$ and k is even). *Let $k \in \mathbb{N}$ be even. Then for all $\varepsilon > 0$ there exist $\delta > 0$ and $n_0 \in \mathbb{N}$ such that every C_{k+1}-free digraph G on $n \geq n_0$ vertices with*

$$(6.12) \qquad\qquad e(G) \geq \binom{n}{2} - \delta n^2$$

can be made into a transitive-bipartite blow up by changing at most εn^2 edges.

To give an idea of the proof, consider the triangle-free case $k = 2$. In this case, we first consider a maximal collection $\mathcal{A}$ of disjoint double edges in G. It is easy to see that for almost all pairs of (double) edges $u_1 u_2, v_1 v_2 \in \mathcal{A}$, either (i) $G[\{u_1, u_2, v_1, v_2\}]$ is a complete balanced bipartite digraph or (ii) G contains all four possible edges from $\{u_1, u_2\}$ to $\{v_1, v_2\}$ (or vice versa). We consider

the following auxiliary 'semi-oriented graph' G' whose vertex set is $\mathcal{A}$. In case (i), we include an (undirected) red edge between u_1u_2 and v_1v_2 in G'. In case (ii), we include a blue edge directed from u_1u_2 to v_1v_2 in G' (or vice versa). One can now show that the red edges induce a set of disjoint almost complete graphs R in G'. We then contract each such red almost complete graph R into a vertex v_R to obtain an oriented graph J (vertices of G' which are not involved in any of these R are also retained in J). So all edges of J are blue. Crucially, it turns out that J is close to a transitive tournament. Moreover, in G each v_R corresponds to an almost complete bipartite digraph, so altogether this shows that the subgraph of G induced by the edges in $\mathcal{A}$ is close to a transitive-bipartite blow up. One can generalize this argument to incorporate the vertices of G not covered by edges in $\mathcal{A}$ (these will only be incident to blue edges in G' and J and not to any red ones).

To formalize the above argument, we make use of the following definitions. A *semi-oriented graph* is obtained from an undirected graph by first colouring each of the edges either red or blue and then giving an orientation to each of the blue edges. So a semi-oriented graph is a pair $G = (V, E)$, where V is a set of vertices and E is a set of coloured edges, some of which are red and undirected and the rest of which are blue and directed. We define basic notions such as induced subgraphs of G in the obvious way. Given a vertex $v \in V$ we denote the set of all vertices $x \in V$ for which there is a blue directed edge $vx \in E$ by $N_G^+(v)$. We call the vertices in $N_G^+(v)$ *blue out-neighbours of v*. We define the sets $N_G^-(v)$ of *blue in-neighbours of v* and $N_G^{\mathrm{red}}(v)$ of *red neighbours of v* in a similar way. If $x \in N_G^+(v) \cup N_G^-(v)$ we say that *x is a blue neighbour of v*.

We denote the complete bipartite digraph (with edges in both directions) with vertex classes of sizes a and b by $DK_{a,b}$. The following simple proposition will also be used in the proof of Lemma 6.6. We omit its straightforward proof, the details of which can be found in [31].

Proposition 6.7. *Let $k \in \mathbb{N}$ be even and let G be a C_{k+1}-free digraph. Suppose G contains a copy of $DK_{k/2,k/2}$ with vertex classes A, B. Then the following hold.*

(i) *Suppose $x \in V(G) \setminus (A \cup B)$. Then $e(G[\{x\}, A \cup B]) \leq k$, with equality only if $G[\{x\}, A \cup B] = \overrightarrow{K}(\{x\}, A \cup B)$ or $G[\{x\}, A \cup B] = \overrightarrow{K}(A \cup B, \{x\})$ or $G[\{x\} \cup A \cup B] = DK_{k/2+1,k/2}$.*

(ii) *Suppose G contains another copy of $DK_{k/2,k/2}$ with vertex classes C, D such that $(C \cup D) \cap (A \cup B) = \emptyset$. Then $e(G[C \cup D, A \cup B]) \leq k^2$, with equality only if $G[C \cup D, A \cup B] = \overrightarrow{K}(C \cup D, A \cup B)$ or $G[C \cup D, A \cup B] = \overrightarrow{K}(A \cup B, C \cup D)$ or $G[C \cup D \cup A \cup B] = DK_{k,k}$.*

Proof of Lemma 6.6. Choose $n_0, \delta, \varepsilon_1, \varepsilon_2$ such that $1/n_0 \ll \delta \ll \varepsilon_1 \ll \varepsilon_2 \ll 1/k, \varepsilon$. Let G be a C_{k+1}-free digraph on $n \geq n_0$ vertices which satisfies (6.12).

Let $t \geq 0$ denote the maximum number of vertex-disjoint copies of $DK_{k/2,k/2}$ in G. Let $\mathcal{A} = \{A^1, \ldots, A^t\}$ be a collection of vertex sets of t vertex-disjoint copies of $DK_{k/2,k/2}$ in G. Let $V_1 := A^1 \cup \cdots \cup A^t$ and let $V_2 := V(G) \setminus V_1$. Note that $G[V_2]$ is $DK_{k/2,k/2}$-free, and hence by the Kővári-Sós-Turán theorem $G[V_2]$ contains at most δn^2 double edges.

Claim 1: *For each $i \in [t]$ there are at most $6\delta^{1/2}n$ vertices $x \in V_2$ for which $G[A^i \cup \{x\}] = DK_{k/2,k/2+1}$.*

Indeed, suppose that there exists a set X of more than $6\delta^{1/2}n$ such vertices. Proposition 6.7(ii) and the fact that $1/n_0 \ll \delta, 1/k$ together imply that $e(G[V_1]) \leq \binom{|V_1|}{2} + \delta n^2$. Moreover, $e(V_1, V_2) \leq |V_1||V_2|$ by Proposition 6.7(i). Together with our previous observation that $G[V_2]$ contains at most δn^2 double edges and (6.12) this implies that $e(G[X]) \geq \binom{|X|}{2} - 3\delta n^2$. But this means that there are $x, y \in X$ such that $xy \in E(G)$ and such that both x and y are joined with double edges to the same vertex class of $G[A^i] = DK_{k/2,k/2}$, which contradicts the fact that G is C_{k+1}-free.

Let G^* be the digraph obtained from G by deleting the at most δn^2 double edges in $G[V_2]$ and deleting the double edges between A^i and all the vertices $x \in V_2$ for which $G[A^i \cup \{x\}] =$

$DK_{k/2,k/2+1}$ (for each $i \in [t]$). By Claim 1, for each $i \in [t]$, the number of the latter double edges is at most $6\delta^{1/2}n \cdot k/2 = 3k\delta^{1/2}n$. Thus

$$(6.13) \qquad e(G^*) \geq e(G) - 2\left(\delta n^2 + \frac{n}{k} \cdot 3k\delta^{1/2}n\right) \overset{(6.12)}{\geq} \binom{n}{2} - 7\delta^{1/2}n^2.$$

Consider the semi-oriented graph $G' = (V', E')$ where $V' := \mathcal{A} \cup V_2$ and the edge set E' is defined as follows. Firstly, for every vertex $v \in V'$ define $f(v)$ to be v if $v \in \mathcal{A}$, and to be $\{v\}$ if $v \in V_2$. If $u, v \in V'$ then there is a blue edge in E' directed from u to v if $G[f(u), f(v)] = \overrightarrow{K}(f(u), f(v))$. If $u, v \in \mathcal{A}$ then there is a red edge in E' between u and v if $G[f(u) \cup f(v)] = DK_{k,k}$. So $G'[V_2] = G^*[V_2]$.

Note that, since G is C_{k+1}-free, G' cannot contain any copy of C_{k+1} which contains at least one blue edge and in which all the blue edges are oriented consistently (as any such copy of C_{k+1} in G' would correspond to a C_{k+1} in G).

Let V_0' denote the set of all vertices $v \in V'$ for which there are at least $\delta^{1/4}n$ vertices $u \in V'$ such that G' does not contain an edge between v and u. Note that Proposition 6.7 together with (6.13) implies that $|E'| \geq \binom{|V'|}{2} - 7\delta^{1/2}n^2$. Hence $|V_0'| \leq 15\delta^{1/4}n$. Let $G'' := G' - V_0'$, $V'' := V' \setminus V_0'$ and $E'' := E(G'')$.

Claim 2:

(a) *For every vertex $v \in V''$ there are at most $\delta^{1/4}n$ vertices $u \in V''$ such that G'' does not contain an edge between v and u.*

(b) *G'' does not contain a triangle uvw such that both uv and vw are red edges and wu is a blue edge.*

(c) *G'' does not contain a triangle uvw such that uv is a red edge and both vw and wu are (directed) blue edges.*

Indeed, (a) is clear from the definition of G'', while (b) and (c) follow easily from the fact that G is C_{k+1}-free.

Given $q, q' \in \mathbb{N}$, we say that $U \subseteq V''$ is a *red (q, q')-clique* if $|U| \geq q'$ and $|U \setminus N^{\mathrm{red}}_{G''}(u)| \leq q$ for all $u \in U$.

Claim 3: *Suppose that R is a red $(\delta^{1/4}n, \varepsilon_1 n)$-clique. Then the following hold.*

(a) *$G''[R]$ does not contain a blue edge.*

(b) *No vertex $v \in V'' \setminus R$ has both a red and a blue neighbour in R.*

(c) *No vertex $v \in V'' \setminus R$ has both a blue in-neighbour and a blue out-neighbour in R.*

Suppose that R' is another red $(\delta^{1/4}n, \varepsilon_1 n)$-clique such that $R \cap R' = \emptyset$. Then the following hold.

(d) *G'' cannot contain both a red edge and a blue edge between R and R'.*

(e) *G'' cannot contain both a directed blue edge from some vertex in R to some vertex in R' and a directed blue edge from some vertex in R' to some vertex in R.*

First note that (a) follows immediately from Claim 2(b) and the definition of a red $(\delta^{1/4}n, \varepsilon_1 n)$-clique. To prove (b), suppose that some vertex $v \in V'' \setminus R$ has both a red and a blue neighbour in R. Claim 2(a) and the fact that $|R| \geq \varepsilon_1 n$ imply that either v has at least $\varepsilon_1 n/3$ red neighbours in R or at least $\varepsilon_1 n/3$ blue neighbours in R (or both). Suppose that the former holds (the argument for the latter is similar). Let $u \in R$ be a blue neighbour of v. Since R is a red $(\delta^{1/4}n, \varepsilon_1 n)$-clique, all but at most $\delta^{1/4}n < \varepsilon_1 n/3$ vertices of R are red neighbours of u. So there exists a red neighbour $u' \in R$ of u which is also a red neighbour of v. Then the triangle $uu'v$ contradicts Claim 2(b). This proves (b). The argument for (c) is similar. (d) follows from (b) and Claim 2(a), while (e) follows from (b), (c) and Claim 2(a).

Claim 4: *There exists a collection $\mathcal{R}$ of pairwise disjoint red $(\delta^{1/4}n, \varepsilon_1 n)$-cliques such that, writing $V_{\mathcal{R}}$ for the set of all those vertices in V'' covered by these red $(\delta^{1/4}n, \varepsilon_1 n)$-cliques, the following holds:*

    (a) *For every $R \in \mathcal{R}$ and every $v \in R$ all red neighbours of v lie in R.*

    (b) *Every $v \in V'' \setminus V_{\mathcal{R}}$ has less than $\varepsilon_1 n$ red neighbours (and all of these lie in $V'' \setminus V_{\mathcal{R}}$).*

To prove Claim 4, let $\mathcal{R}$ be a collection of pairwise disjoint red $(\delta^{1/4}n, \varepsilon_1 n)$-cliques such that the set $V_{\mathcal{R}}$ of all those vertices in V'' covered by these red $(\delta^{1/4}n, \varepsilon_1 n)$-cliques is maximal and, subject to this condition, such that $|\mathcal{R}|$ is minimal. We will show that $\mathcal{R}$ is as required in Claim 4.

To prove that Claim 4(a) holds, suppose first that there is some vertex $x \in V'' \setminus V_{\mathcal{R}}$ that has a red neighbour in some $R \in \mathcal{R}$. Then Claims 2(a) and 3(b) together imply that $|(R \cup \{x\}) \setminus N^{\mathrm{red}}_{G''}(x)| \leq \delta^{1/4}n$. Moreover, by Claims 2(a) and 3(a),(b) we have that $|(R \cup \{x\}) \setminus N^{\mathrm{red}}_{G''}(v)| \leq \delta^{1/4}n$ for every $v \in R$. So $R \cup \{x\}$ is a red $(\delta^{1/4}n, \varepsilon_1 n)$-clique, contradicting our choice of $\mathcal{R}$.

Suppose next that there are distinct $R, R' \in \mathcal{R}$ such that G'' contains a red edge between R and R'. Then Claims 3(a),(d) imply that $G''[R \cup R']$ does not contain a blue edge. Together with Claim 2(a) this implies that $R \cup R'$ is a red $(\delta^{1/4}n, \varepsilon_1 n)$-clique, again contradicting our choice of $\mathcal{R}$. Altogether this proves Claim 4(a).

To check Claim 4(b), suppose that some $v \in V'' \setminus V_{\mathcal{R}}$ has at least $\varepsilon_1 n$ red neighbours. Claim 2(b) implies that $G''[N^{\mathrm{red}}_{G''}(v)]$ cannot contain a blue edge. Together with Claim 2(a) this implies that $G''[N^{\mathrm{red}}_{G''}(v)]$ is a red $(\delta^{1/4}n, \varepsilon_1 n)$-clique. But Claim 4(a) implies that $N^{\mathrm{red}}_{G''}(v) \subseteq V'' \setminus V_{\mathcal{R}}$, contradicting our choice of $\mathcal{R}$. This completes the proof of Claim 4.

Let G'''' be the semi-oriented graph obtained from G'' by deleting all the red edges which are not covered by some $R \in \mathcal{R}$. Note that by Claim 4(b) at most $\varepsilon_1 n^2$ red edges are deleted. Let J be the oriented graph obtained from G'''' by contracting each $R \in \mathcal{R}$ into a single vertex v_R. So $V(J)$ consists of all these vertices v_R as well as all the vertices in $V'' \setminus V_{\mathcal{R}}$. Let $J_2 := J[V'' \setminus V_{\mathcal{R}}] = G''''[V'' \setminus V_{\mathcal{R}}]$ and let $J_1 := J - V(J_2)$. Claims 3(c),(e) and Claim 4(a) together imply that J is indeed an oriented graph. Moreover, by Claim 2(a) J_1 is a tournament and $J[V(J_1), V(J_2)]$ is a bipartite tournament (i.e. for all $v_R \in V(J_1)$ and $v \in V(J_2)$ either $v_R v$ or $v v_R$ is a directed edge of J).

Claim 5:

    (a) *J does not contain a copy of C_3 having at least one vertex in $V(J_1)$.*

    (b) *J_1 is a transitive tournament.*

    (c) *J can be made into a transitive tournament by changing at most $\varepsilon_2 n^2$ edges in $E(J_2)$.*

Suppose that (a) does not hold and let xyv_R be a copy of C_3 in J. We only consider the case when $x \in V(J_1)$ and $y \in V(J_2)$; the other cases are similar. So let $R' \in \mathcal{R}$ be such that $x = v_{R'}$. Claim 2(a) and the definition of J together imply that R' contains a blue in-neighbour x' of y (in G''). Moreover, Claims 2(a) and 3(c),(e) imply that $|R \setminus N^-_{G''}(x')| \leq \delta^{1/4}n$ and $|R \setminus N^+_{G''}(y)| \leq \delta^{1/4}n$. Together with the fact that R is a red $(\delta^{1/4}n, \varepsilon_1 n)$-clique this implies that R contains a path $P = u \ldots v$ of length $k - 2$ where $u \in N^+_{G''}(y)$ and $v \in N^-_{G''}(x')$. So $Px'y$ is a C_{k+1} in G'' in which all the blue edges are oriented consistently. Using the fact that the edge vx' is blue, it is now easy to see that $Px'y$ corresponds to a C_{k+1} in G, a contradiction. This proves (a). (b) follows from (a) and our previous observation that J_1 is a tournament.

It remains to prove (c). Note that $e(J_2) \geq \binom{|J_2|}{2} - 2\varepsilon_1 n^2$ by Claim 2(a) and the definition of J (and of G''''). Moreover, $J_2 = G''''[V'' \setminus V_{\mathcal{R}}]$ is a C_{k+1}-free oriented graph. So Lemma 6.4 implies that $J_2 = T_{|J_2|} \pm \varepsilon_2 n^2$. Let $\sigma_2 : V(J_2) \to [|J_2|]$ be a transitive-optimal ordering of the vertices of J_2. Let $r := |J_1| = |\mathcal{R}|$ and let $v_{R_1}, \ldots v_{R_r}$ be the unique transitive ordering of the vertices of J_1. We claim that for every vertex $x \in V(J_2)$ there exists an index $i_x \in [r]$ such that all the v_{R_i} with $i \leq i_x$ are in-neighbours of x in J while all the v_{R_i} with $i > i_x$ are out-neighbours of x in J.

(Indeed, suppose not. Since $J[V(J_1), V(J_2)]$ is a bipartite tournament this implies that there are indices $i < j$ such that v_{R_i} is an out-neighbour of x in J and v_{R_j} is an in-neighbour of x in J. But then $xv_{R_i}v_{R_j}$ is a copy of C_3 contradicting (a).) For each $i \in [r]$ let $X_i := \{x \in V(J_2) : i_x = i\}$. Note that there are no indices $i < j$ such that J contains a directed edge from some vertex $x \in X_j$ to some vertex $x' \in X_i$ (otherwise $xx'v_{R_{i+1}}$ would be a copy of C_3 contradicting (a)). Consider the vertex ordering σ obtained from $v_{R_1}, \ldots, v_{R_r}$ by including all the vertices in X_i between v_{R_i} and $v_{R_{i+1}}$ in the ordering induced by σ_2 (for each $i \in [r]$). This vertex ordering shows that (c) holds.

Recall that for each $R \in \mathcal{R}$ the set $\bigcup R$ is a subset of $V(G)$ of size $k|R|$.

Claim 6: *Each red $(\delta^{1/4}n, \varepsilon_1 n)$-clique $R \in \mathcal{R}$ satisfies $G[\bigcup R] = DK_{|R|k/2, |R|k/2} \pm \delta^{1/5}n^2$.*

To prove Claim 6, pick $v \in R$ and write $N_{G''}^{\text{red}}(v) \cap R = \{v_1, \ldots, v_s\}$. Recall that v corresponds to a copy of $DK_{k/2,k/2}$ in G, and let A and B denote the vertex classes of this copy. Similarly, each v_i corresponds to a copy of $DK_{k/2,k/2}$ in G. Let A_i and B_i denote its vertex classes. Recall from the definition of G'' that $G[A \cup A_i \cup B \cup B_i] = DK_{k,k}$. By swapping A_i and B_i if necessary, we may assume that the vertex classes of this copy of $DK_{k,k}$ are $A \cup A_i$ and $B \cup B_i$. Since G is C_{k+1}-free, neither $G[A_1 \cup \cdots \cup A_s]$ nor $G[B_1 \cup \cdots \cup B_s]$ contains an edge. Thus whenever $v_i v_j$ is a red edge in G'' then $G[A_i \cup A_j \cup B_i \cup B_j]$ is a copy of $DK_{k,k}$ with vertex classes $A_i \cup A_j$ and $B_i \cup B_j$. But since R is a red $(\delta^{1/4}n, \varepsilon_1 n)$-clique, for each $i \in [s]$ all but at most $\delta^{1/4}n$ vertices in $\{v_1, \ldots, v_s\}$ are red neighbours of v_i and $|R \setminus \{v_1, \ldots, v_s\}| \le \delta^{1/4}n$. Thus $G[\bigcup R] = DK_{|R|k/2, |R|k/2} \pm \delta^{1/5}n^2$, as required.

Using Claims 5(c) and 6 it is now straightforward to check that G can be made into a transitive-bipartite blow up by changing at most εn^2 edges. □

We now have all the tools we need to show that almost all C_k-free oriented graphs are close to acyclic, and that for all even k almost all C_k-free digraphs are close to acyclic, and that for all odd k almost all C_k-free digraphs are close to a transitive-bipartite blow up. The proof of Lemma 6.8 is almost identical to that of Lemma 4.5, using Lemmas 6.4, 6.5 and 6.6 instead of Lemma 4.3, and so is omitted here.

Lemma 6.8. *For every $k \in \mathbb{N}$ with $k \ge 3$ and any $\alpha > 0$ there exists $\varepsilon > 0$ such that the following holds for all sufficiently large n.*

(i) *All but at most $f(n, C_k)2^{-\varepsilon n^2}$ C_k-free oriented graphs on n vertices can be made into subgraphs of T_n by changing at most αn^2 edges.*

(ii) *If k is even then all but at most $f^*(n, C_k)2^{-\varepsilon n^2}$ C_k-free digraphs on n vertices can be made into subgraphs of T_n by changing at most αn^2 edges.*

(iii) *If k is odd then all but at most $f^*(n, C_k)2^{-\varepsilon n^2}$ C_k-free digraphs on n vertices can be made into a subgraph of a transitive-bipartite blow up by changing at most αn^2 edges.*

7. TYPICAL C_k-FREE ORIENTED GRAPHS AND DIGRAPHS ARE NOT ACYCLIC

Let $\mathcal{O}_{n,k}$ be the set of all labelled C_k-free oriented graphs on n vertices and let $\mathcal{O}_{n,k}^*$ be the set of all labelled C_k-free digraphs on n vertices. We show that almost all graphs in $\mathcal{O}_{n,k}$ and almost all graphs in $\mathcal{O}_{n,k}^*$ have at least $cn/\log n$ backwards edges in a transitive-optimal ordering, for some constant $c > 0$. Let $\mathcal{O}_{n,k,r}$ be the set of all labelled C_k-free oriented graphs on n vertices with exactly r backwards edges in a transitive-optimal ordering. Let $\mathcal{O}_{n,k,\le r} := \bigcup_{i \in \{0,1,\ldots,\lfloor r \rfloor\}} \mathcal{O}_{n,k,i}$, and define the digraph analogues $\mathcal{O}_{n,k,r}^*$ and $\mathcal{O}_{n,k,\le r}^*$ in a similar way.

Lemma 7.1. *Let $k \ge 3$ and let $n \in \mathbb{N}$ be sufficiently large. Then*

(i) $|\mathcal{O}_{n,k,n/2^{13}}| \ge 2^{n/2^{14}} |\mathcal{O}_{n,k,\le n/(2^{14}\log n)}|$,

(ii) $|\mathcal{O}_{n,k,n/2^{13}}^*| \ge 2^{n/2^{14}} |\mathcal{O}_{n,k,\le n/(2^{14}\log n)}^*|$.

Note that Lemma 7.1 together with Lemma 6.8 immediately yields Theorem 1.3.

Proof of Lemma 7.1. We only prove the case $k = 3$ of (i) here; the proofs for (ii) and the case $k > 3$ are very similar. Let $m_2 := \lfloor n/2^{13} \rfloor$. Fix $m_1 \in \mathbb{Z}$ with $0 \leq m_1 \leq m_2/(2 \log n)$. For every oriented graph G fix some transitive-optimal ordering $\sigma_G : V(G) \to [n]$.

Consider an auxiliary bipartite graph H with vertex classes $\mathcal{O}_{n,3,m_1}$ and $\mathcal{O}_{n,3,0}$ whose edge set is defined as follows. Let there be an edge in H between $A \in \mathcal{O}_{n,3,m_1}$ and $B \in \mathcal{O}_{n,3,0}$ if the graph B can be obtained from the graph A by deleting the m_1 backwards edges with respect to σ_A. Note that every graph generated in this way from a graph $A \in \mathcal{O}_{n,3,m_1}$ belongs to $\mathcal{O}_{n,3,0}$, so A certainly has at least one neighbour in $\mathcal{O}_{n,3,0}$.

We claim that, in H, a graph $B \in \mathcal{O}_{n,3,0}$ has at most $\binom{n^2/2}{m_1} 2^{m_1}$ neighbours in $\mathcal{O}_{n,3,m_1}$. Indeed, any graph in $\mathcal{O}_{n,3,m_1}$ that can generate B in the described way can be obtained from B by choosing exactly m_1 of the at most $n^2/2$ pairs of vertices that have no edge between them in B, and then adding edges between them with some orientations (for which there are 2^{m_1} possibilities).

Together with our previous observation that, in H, every graph $A \in \mathcal{O}_{n,3,m_1}$ has at least one neighbour in $\mathcal{O}_{n,3,0}$, this implies that

$$(7.1) \qquad |\mathcal{O}_{n,3,m_1}| \leq \sum_{A \in \mathcal{O}_{n,3,m_1}} d_{\mathcal{O}_{n,3,0}}(A) = \sum_{B \in \mathcal{O}_{n,3,0}} d_{\mathcal{O}_{n,3,m_1}}(B) \leq |\mathcal{O}_{n,3,0}| \binom{n^2/2}{m_1} 2^{m_1}.$$

For a graph $G \in \mathcal{O}_{n,3,0}$ we define a *flippable* 4-*set* in G to be any set of 4 vertices, with labels w, x, y, z say, satisfying the following:

- the vertices w, x, y, z are consecutive in the ordering σ_G; that is $\sigma_G(w) + 3 = \sigma_G(x) + 2 = \sigma_G(y) + 1 = \sigma_G(z)$,
- $\sigma_G(w) - 1$ is divisible by 4.

Note that every graph in $\mathcal{O}_{n,3,0}$ has $\lfloor n/4 \rfloor$ flippable 4-sets.

Now consider an auxiliary bipartite graph H' with vertex classes $\mathcal{O}_{n,3,0}$ and $\mathcal{O}_{n,3,m_2}$ whose edge set is defined as follows. Let there be an edge in H' between $B \in \mathcal{O}_{n,3,0}$ and $C \in \mathcal{O}_{n,3,m_2}$ if the graph C can be obtained from the graph B by choosing exactly m_2 flippable 4-sets in B with respect to σ_B and, for each flippable 4-set w, x, y, z chosen, deleting all edges between the vertices w, x, y, z and then adding the edges of a 4-cycle $wxyz$. Note that every graph generated in this way from a graph $B \in \mathcal{O}_{n,3,0}$ belongs to $\mathcal{O}_{n,3,m_2}$.

We claim that, in H', a graph $B \in \mathcal{O}_{n,3,0}$ has exactly $\binom{\lfloor n/4 \rfloor}{m_2}$ neighbours in $\mathcal{O}_{n,3,m_2}$. Indeed, the neighbours of B are precisely those graphs generated by choosing exactly m_2 of the exactly $\lfloor n/4 \rfloor$ flippable 4-sets in B with respect to σ_B, and then changing the edges between pairs of vertices in these flippable 4-sets in the described way. Each choice of m_2 flippable 4-sets generates a different graph. So the claim holds.

We claim also that, in H', a graph $C \in \mathcal{O}_{n,3,m_2}$ has at most 2^{8m_2} neighbours in $\mathcal{O}_{n,3,0}$. Indeed, first note that any graph in $\mathcal{O}_{n,3,m_2}$ with at least one neighbour in $\mathcal{O}_{n,3,0}$ contains exactly m_2 induced 4-cycles. Any graph in $\mathcal{O}_{n,3,0}$ that can generate C in the described way can be obtained from C by choosing for each of the m_2 induced 4-cycles an ordering of the 4 vertices respecting the order of the 4-cycle (of which there are 4), and then changing the edges between pairs of vertices in these 4-cycles to some transitive configuration with respect to the chosen ordering (for which there are 2^6 possibilities). So indeed the claim holds.

So using these degree bounds gives us that

$$(7.2) \qquad |\mathcal{O}_{n,3,0}| \binom{\lfloor n/4 \rfloor}{m_2} = \sum_{B \in \mathcal{O}_{n,3,0}} d_{\mathcal{O}_{n,3,m_2}}(B) = \sum_{C \in \mathcal{O}_{n,3,m_2}} d_{\mathcal{O}_{n,3,0}}(C) \leq |\mathcal{O}_{n,3,m_2}| 2^{8m_2}.$$

Now (7.1) and (7.2) together imply that

$$(7.3) \qquad \frac{|\mathcal{O}_{n,3,m_2}|}{|\mathcal{O}_{n,3,m_1}|} \geq \frac{\binom{\lfloor n/4 \rfloor}{m_2}}{\binom{n^2/2}{m_1} 2^{m_1} 2^{8m_2}}.$$

Since n is sufficiently large we have that $\binom{\lfloor n/4 \rfloor}{m_2} \geq \left(\frac{n}{8m_2}\right)^{m_2}$ and $\binom{n^2/2}{m_1} 2^{m_1} \leq n^{2m_1}$. Hence the right hand side of (7.3) is at least

$$\left(\frac{n}{2^{11}m_2}\right)^{m_2} n^{-2m_1} \geq \left(\frac{n}{2^{11}m_2}\right)^{m_2} n^{-\frac{m_2}{\log n}} = 2^{m_2 \log\left(\frac{n}{2^{11}m_2}\right) - \frac{m_2}{\log n} \log n} \geq 2^{m_2}.$$

So this together with (7.3) gives us that $|\mathcal{O}_{n,3,m_2}| \geq 2^{m_2}|\mathcal{O}_{n,3,m_1}|$ for any integer $0 \leq m_1 \leq m_2/(2\log n)$. So since n is sufficiently large,

$$|\mathcal{O}_{n,3,m_2}| \geq \frac{2^{m_2}}{n}|\mathcal{O}_{n,3,\leq m_2/(2\log n)}| \geq 2^{n/2^{14}}|\mathcal{O}_{n,3,\leq m_2/(2\log n)}|,$$

as required. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \square$

8. Acknowledgment

References

[1] N. Alon, A. Shapira, Testing subgraphs in directed graphs, *J. Comput. System Sci.* **69** (2004), 354–382.

[2] N. Alon, R. Yuster, The number of orientations having no fixed tournament, *Combinatorica* **26** (2006), 1–16.

[3] J. Balogh, B. Bollobás, R. Morris, Hereditary properties of combinatorial structures: posets and oriented graphs, *J. Graph Theory* **56** (2007), 311–332.

[4] J. Balogh, B. Bollobás, R. Morris, Hereditary properties of tournaments, *Electronic J. Combin.* **14** (2007).

[5] J. Balogh, B. Bollobás, M. Simonovits, The number of graphs without forbidden subgraphs, *J. Combin. Theory Series B* **91** (2004), 1–24.

[6] J. Balogh, B. Bollobás, M. Simonovits, The typical structure of graphs without given excluded subgraphs, *Random Structures and Algorithms* **34** (2009) 305–318.

[7] J. Balogh, B. Bollobás, M. Simonovits, The fine structure of octahedron-free graphs, *J. Combin. Theory Series B* **101** (2011), 67–84.

[8] J. Balogh, R. Morris, W. Samotij, Independent sets in hypergraphs, *J. Amer. Math. Soc.* **28** (2015), 669–709.

[9] J. Balogh, R. Morris, W. Samotij, L. Warnke, The typical structure of sparse K_{r+1}-free graphs, *Trans. Amer. Math. Soc.*, to appear.

[10] J. Balogh, D. Mubayi, Almost all triangle-free triple systems are tripartite, *Combinatorica* **32** (2012), 143–169.

[11] J. Balogh, W. Samotij, The number of $K_{s,t}$-free graphs, *J. London Math. Soc.* **83** (2011), 368–388.

[12] W.G. Brown, F. Harary, Extremal digraphs, in *Combinatorial theory and its applications, Coll. Math. Soc. J. Bolyai* **4** (1970), 135–198.

[13] G. Cherlin, The classification of countable homogeneous directed graphs and countable homogeneous n-tournaments, *AMS Memoir* **131** (1998), AMS, Rhode Island.

[14] M. Chudnovsky, P. Seymour, B. Sullivan, Cycles in dense digraphs, *Combinatorica* **28** (2008), 1–18.

[15] R. Diestel, Graph Theory, *Graduate Texts in Mathematics* **173** 4th ed. (2010), Springer-Verlag, Heidelberg.

[16] P. Erdős, P. Frankl, V. Rödl, The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent, *Graphs and Combin.* **2** (1986), 113–121.

[17] P. Erdős, D.J. Kleitman, B.L. Rothschild, Asymptotic enumeration of K_n-free graphs, in *Colloquio Internazionale sulle Teorie Combinatorie* (Rome, 1973), Vol. II, 19–27. *Atti dei Convegni Lincei* **17**, Accad. Naz. Lincei, Rome, 1976.

[18] P. Erdős, M. Simonovits, Supersaturated graphs and hypergraphs, *Combinatorica* **3** (1983), 181–192.

[19] R. Häggkvist, C. Thomassen, On pancyclic digraphs, *J. Combin. Theory Series B* **20** (1976), 20–40.

[20] Ph.G. Kolaitis, H.J. Prömel, B.L. Rothschild, $K_{\ell+1}$-free graphs: asymptotic structure and a $0-1$ law, *Trans. Amer. Math. Soc.* **303** (1987), 637–671.

[21] E. Lamken, B.L. Rothschild, The number of odd-cycle-free graphs, in *Finite and Infinite Sets (A. Hajnal, L. Lovász, and V.T. Sós, eds.)* (1984), 547–553. *Coll. Math. Soc. J. Bolyai* **37**.

[22] R. Morris, D. Saxton, The number of $C_{2\ell}$-free graphs, *Advances in Math.*, to appear.

[23] D. Osthus, H.J. Prömel, A. Taraz, For which densities are random triangle-free graphs almost surely bipartite?, *Combinatorica* **23** (2003), 105–150.

[24] Y. Person, M. Schacht, Almost all hypergraphs without Fano planes are bipartite, in *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms* (SODA, 2009), 217–226. ACM Press.

[25] H.J. Prömel, A. Steger, The asymptotic number of graphs not containing a fixed color-critical subgraph, *Combinatorica* **12** (1992), 463–473.

[26] R.W. Robinson, Enumeration of acyclic digraphs, in *R. C. Bose et al. (Eds.), Proc. Second Chapel Hill Conf. on Combinatorial Mathematics and its Applications* (Univ. North Carolina, Chapel Hill, N.C., 1970), 391–399.

[27] R.W. Robinson, Counting labeled acyclic digraphs, in *F. Harary (Ed.), New Directions in the Theory of Graphs* (Academic Press, NY, 1973), 239–273.

[28] D. Saxton, A. Thomason, Hypergraph containers, *Inventiones Math.* **201** (2015), 925–992.

[29] D. Saxton, A. Thomason, Simple containers for simple hypergraphs, *Combin. Probab. Comput.*, to appear.

[30] R.P. Stanley, Acyclic orientations of graphs, *Discrete Math.* **5** (1973), 171–178.

[31] T. Townsend, PhD Thesis, *Extremal problems on graphs, directed graphs and hypergraphs*, PhD thesis, University of Birmingham (2015).

Daniela Kühn, Deryk Osthus, Timothy Townsend
School of Mathematics
University of Birmingham
Edgbaston
Birmingham
B15 2TT
UK
E-mail addresses:
{d.kuhn, d.osthus, txt238}@bham.ac.uk

Yi Zhao
Department of Mathematics & Statistics
Georgia State University
30 Pryor Street
Atlanta
GA 30303
USA
E-mail address:
yzhao6@gsu.edu