COUNTING RESTRICTED ORIENTATIONS OF RANDOM GRAPHS

MAURÍCIO COLLARES, YOSHIHARU KOHAYAKAWA, ROBERT MORRIS,
AND GUILHERME OLIVEIRA MOTA

ABSTRACT. We count orientations of $G(n,p)$ avoiding certain classes of oriented graphs. In particular, we study $T_r(n,p)$, the number of orientations of the binomial random graph $G(n,p)$ in which every copy of K_r is transitive, and $S_r(n,p)$, the number of orientations of $G(n,p)$ containing no strongly connected copy of K_r. We give the correct order of growth of $\log T_r(n,p)$ and $\log S_r(n,p)$ up to polylogarithmic factors; for orientations with no cyclic triangle, this significantly improves a result of Allen, Kohayakawa, Mota and Parente. We also discuss the problem for a single forbidden oriented graph, and state a number of open problems and conjectures.

1. INTRODUCTION

An orientation $\vec{H}$ of a graph H is an oriented graph obtained by assigning an orientation to every edge of H. Over 40 years ago, Erdős [7] initiated the study of $D(G, \vec{H})$, the number of $\vec{H}$-free orientations of a graph G, and in particular posed the problem of determining $D(n, \vec{H}) := \max \left\{ D(G, \vec{H}) : |V(G)| = n \right\}$. For tournaments, this problem was resolved by Alon and Yuster [4], who proved that $D(n, T_k) = 2^{\text{ex}(n, K_k)}$ holds for any tournament T_k, and all sufficiently large $n \in \mathbb{N}$.

Recently, Allen, Kohayakawa, Mota and Parente [2] introduced a related problem in the context of random graphs: that of determining the typical number of $C_r^{\circlearrowleft}$-free orientations of the Erdős–Rényi random graph $G(n,p)$, where $C_r^{\circlearrowleft}$ is the directed cycle of length r. The main result of [2] is as follows.

Theorem 1.1. *Let $r \geqslant 3$. Then, with high probability as $n \to \infty$,*

$$
\log_2 D\big(G(n,p), C_r^{\circlearrowleft}\big) = \begin{cases} \big(1 + o(1)\big)p\binom{n}{2} & \text{if } n^{-2} \ll p \ll n^{-(r-2)/(r-1)}, \\ o\big(pn^2\big) & \text{if } p \gg n^{-(r-2)/(r-1)}. \end{cases}
$$

Our first theorem provides the following improvement in the case $r = 3$; the $\widetilde{\Theta}(\cdot)$-notation indicates upper and lower bounds that differ by a polylogarithmic factor.

Theorem 1.2. *The following bounds hold with high probability as $n \to \infty$.*

$$\log_2 D\big(G(n,p), C_3^{\circlearrowleft}\big) = \begin{cases} (1 + o(1))p\binom{n}{2} & \text{if } n^{-2} \ll p \ll n^{-1/2}, \\[2mm] \widetilde{\Theta}(n/p) & \text{if } p \gg n^{-1/2}. \end{cases} \tag{1}$$

We will in fact prove a similar upper bound for $D(G(n,p), C_r^{\circlearrowleft})$ for all $r \geqslant 3$ (see Theorem 6.1). However, we do not believe that our bound is sharp when $r \geqslant 4$; instead, we believe that the natural generalization of the lower bound construction in Theorem 1.2 (see Section 2) is sharp up to polylogarithmic factors.

Conjecture 1.3. *Let $r \geqslant 4$. If $p \gg n^{-(r-2)/(r-1)}$, then*

$$\log D\big(G(n,p), C_r^{\circlearrowleft}\big) = \widetilde{\Theta}\left(\frac{n}{p^{1/(r-2)}}\right)$$

with high probability as $n \to \infty$.

Note that $D(G, C_3^{\circlearrowleft})$ is the number of orientations of G in which every triangle is transitive, or equivalently, in which no triangle is strongly connected. This suggests two natural generalizations of Theorem 1.2, which we discuss below.

Avoiding non-transitive tournaments. Our first generalization deals with the number of orientations of $G(n,p)$ avoiding non-transitive tournaments of a given size.

Definition 1.4. Let $T_r(n,p)$ denote the random variable which counts the number of orientations of the random graph $G(n,p)$ in which every copy of K_r is transitively oriented (we will simply say that "every K_r is transitive").

Our next main result generalizes Theorem 1.2 by determining the typical value of $\log T_r(n,p)$ up to polylogarithmic factors for every $r \geqslant 3$.

Theorem 1.5. *Let $r \geqslant 3$. The following bounds hold with high probability as $n \to \infty$.*

$$\log_2 T_r(n,p) = \begin{cases} (1 + o(1))p\binom{n}{2} & \text{if } n^{-2} \ll p \ll n^{-2/(r+1)}, \\[2mm] \widetilde{\Theta}\big(p^{2-\binom{r}{2}}n^{4-r}\big) & \text{if } n^{-2/(r+1)} \ll p \ll n^{-2/(r+2)}, \tag{2} \\[2mm] \widetilde{\Theta}(n/p) & \text{if } p \gg n^{-2/(r+2)}. \tag{3} \end{cases}$$

Note that the functions in (2) and (3) coincide when $r = 3$. Figure 1 illustrates the typical behaviour of $T_r(n,p)$ given by Theorem 1.5 when $r \geqslant 4$. We remark that, despite the more complicated behaviour of $T_r(n,p)$ when $r \geqslant 4$, the proof of Theorem 1.5 is not significantly more difficult than that of Theorem 1.2.
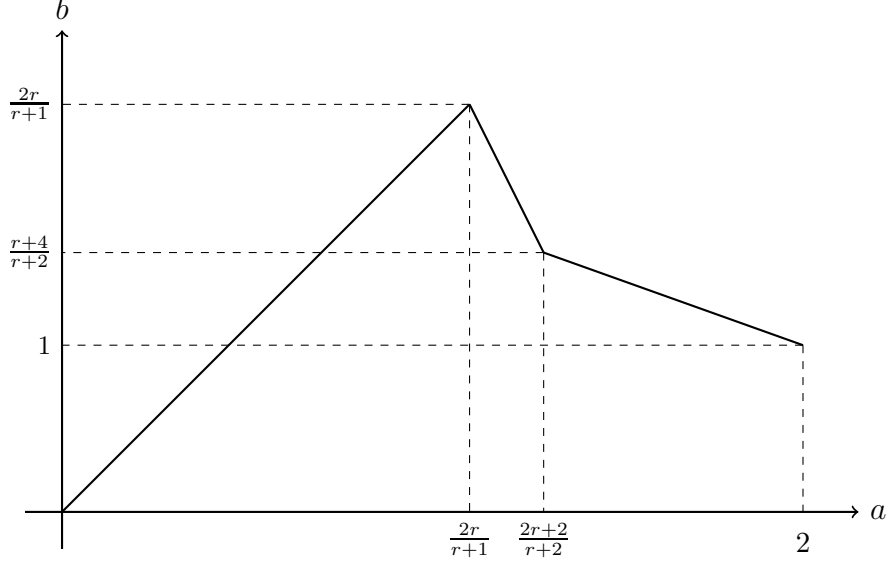
FIGURE 1. The graph of $b = b(a)$, where $T_r(n, p) = \exp\left(n^{b+o(1)}\right)$ and $p\binom{n}{2} = n^a$.

Avoiding strongly connected tournaments. The construction used to prove the lower bounds in Theorems 1.2 and 1.5 and Conjecture 1.3 can also be used to prove the lower bound in the following conjecture.

Conjecture 1.6. *Let $\vec{H}$ be a strongly connected tournament on $r \geqslant 3$ vertices, and suppose that $p \gg n^{-2/(r+1)}$. Then*

$$\log D\big(G(n, p), \vec{H}\big) = \widetilde{\Theta}\left(\frac{n}{p^{(r-1)/2}}\right)$$

with high probability as $n \to \infty$.

Theorem 1.2 proves the conjecture in the case $r = 3$. Moreover, we are able to prove that the upper bound holds if we instead forbid *all* strongly connected tournaments on r vertices.

Definition 1.7. Let $S_r(n, p)$ denote the random variable which counts the number of orientations of the random graph $G(n, p)$ in which no copy of K_r is strongly connected.

Note that $T_r(n, p) \leqslant S_r(n, p)$, and that $D(G(n, p), C_r^{\circlearrowleft}) \leqslant S_r(n, p)$, since every strongly connected orientation of K_r contains a Hamiltonian cycle. The following theorem determines the typical value of $\log S_r(n, p)$ up to polylogarithmic factors for every $r \geqslant 3$.

Theorem 1.8. *Let $r \geqslant 3$. The following bounds hold with high probability as $n \to \infty$.*

$$\log_2 S_r(n, p) = \begin{cases} \big(1 + o(1)\big)p\dbinom{n}{2} & \text{if } n^{-2} \ll p \ll n^{-2/(r+1)}, \\[2mm] \widetilde{\Theta}\left(\dfrac{n}{p^{(r-1)/2}}\right) & \text{if } p \gg n^{-2/(r+1)}. \end{cases}$$

3

We consider this result to be reasonably strong evidence in favour of Conjecture 1.6. In Section 7 we will discuss other forbidden oriented graphs; we do not know of any oriented graph $\vec{H}$ containing a cycle for which our lower bound construction fails to give sharp bounds on $D\bigl(G(n,p),\vec{H}\bigr)$.

The rest of the paper is organised as follows: in Section 2 we prove the various lower bounds; in Sections 3, 4 and 5 we prove the upper bounds in Theorems 1.2, 1.5 and 1.8 respectively; in Section 6 we prove an upper bound for arbitrary cycles, and in Section 7 we discuss some further open problems and conjectures.

2. Lower bounds

Each of the bounds proved in this section (and the general lower bound given in Section 7.2) follows from the same simple construction. Roughly speaking, we fix a linear order on the vertex set (let us identify it with the set $[n] = \{1, \ldots, n\}$), choose a "critical length" a, and orient all edges of length at least a in the same direction ("forward"). The edges shorter than a may be oriented in either direction, as long as they are not at risk of creating a forbidden substructure; by choosing a carefully, we can guarantee that (with high probability) there are many such "free" edges.

To illustrate this construction with a simple example, let us begin by proving the lower bound in equation (1) in the case $p = o(1)$. (The case $p = \Theta(1)$ follows from Proposition 2.2, below.)

Proposition 2.1. *If $\omega \gg 1$ and $n^{-1/2} \leqslant p = o(1)$, then*

$$\log D\bigl(G(n,p), C_3^{\circlearrowleft}\bigr) \geqslant \frac{n}{\omega \cdot p} \tag{4}$$

with high probability as $n \to \infty$.

Proof. We may assume that $\omega \to \infty$ sufficiently slowly. We will show that with high probability there exists a set of at least $n/\omega p$ edges which can be oriented freely without creating a non-transitive triangle. To do so, set $a = 2 \cdot p^{-2}/\omega$, and note that $a = o(n)$, since $p \geqslant n^{-1/2}$ and $\omega \to \infty$, and that $a \gg 1$, since $p = o(1)$ and $\omega \to \infty$ sufficiently slowly. Let us say that an edge uv is a-short if its length $|u - v|$ is less than a, and a-long otherwise. Note that if we orient all a-long edges forward, then any non-transitive triangle must contain at least two a-short edges (a "dangerous triangle"), at least one of which must be oriented backward.

Now, observe (e.g., by Chernoff's inequality) that with high probability the number of a-short edges in $G(n,p)$ is $\bigl(1 + o(1)\bigr)pan$, and that the expected number of triangles in $G(n,p)$ containing at least two a-short edges is $O\bigl(p^3 a^2 n\bigr) \ll pan$. By Markov's inequality, it follows that with high probability there exists a set of at least $pan/2 = n/\omega p$ edges that are not in any such triangle and therefore can be oriented freely, as required. $\qquad \square$

We conjecture (see Conjecture 7.5) that the lower bound on $D\bigl(G(n,p), C_3^{\circlearrowleft}\bigr)$ given by Proposition 2.1 is sharp up to a constant factor in the exponent when p is not too large.

On the other hand, when $p \geqslant 1/\log n$ we can obtain a stronger lower bound by varying the linear order used in the construction. The following proposition also provides a suitable lower bound for all non-acyclic $\vec{H}$ when $p = \Theta(1)$.

Proposition 2.2. *If $\vec{H}$ contains a cycle and $p \gg (\log n)/n$, then*

$$D\big(G(n,p), \vec{H}\big) \geqslant p^n \cdot n! \, / \, e^{o(n)}$$

with high probability as $n \to \infty$.

Proposition 2.2 is a straightforward consequence of a result of Goddard, Kenyon, King and Schulman [9, Theorem 2.5] which gives a lower bound on the number of acyclic orientations of an arbitrary graph G in terms of its degree sequence (see also [10]). However, since the proof requires some tedious calculation using Stirling's formula, we provide the details of the proof in the appendix.

It is also straightforward to replace the factor of $1/\omega$ in (4) by a small fixed constant using the second moment method. To illustrate this, we will prove the following bound on $T_r(n, p)$.

Proposition 2.3. *Let $r \geqslant 3$. There exists $c > 0$ such that if $n^{-2/(r+1)} \ll p \ll n^{-2/(r+2)}$, then*

$$\log T_r(n,p) \geqslant c \cdot \frac{n^{4-r}}{p^{\binom{r}{2}-2}}$$

with high probability as $n \to \infty$.

Observe that Propositions 2.1 and 2.2 imply the lower bound in equation (3), and Proposition 2.3 implies the lower bound in (2). Indeed, the lower bound in (3) follows simply by noticing that in a $C_3^{\circlearrowleft}$-free oriented graph, all copies of K_r are transitive.

Since the proof of Proposition 2.3 requires some slightly tedious calculations, we will give here just a sketch of the proof; for the full details, see the appendix.

Sketch proof of Proposition 2.3. We repeat the proof strategy of Proposition 2.1, using the second moment method instead of Markov's inequality. Choose $a = c \cdot n^{3-r} p^{1-\binom{r}{2}}$, and observe that the lower bound on p implies that $a = o(n)$, and the upper bound implies that $a = \Omega(p^{-2})$. As before, say that an edge is a-short if its length is less than a, and a-long otherwise. Note that any non-transitive copy of K_r contains a cyclic triangle, so if we orient all a-long edges forward, then any non-transitive K_r must contain at least two a-short edges sharing a vertex, at least one of which must be oriented backward.

Let X denote the number of copies of K_r in $G(n,p)$ that are "dangerous", in the sense that they contain two a-short edges incident to a single vertex. A straightforward calculation gives

$$\mathbb{E}[X] \; = \; \Theta\big(n^{r-2}a^2 p^{\binom{r}{2}}\big) \; \leqslant \; pan/2r^2,$$

if $c > 0$ is sufficiently small, and

$$\mathrm{Var}(X) \; = \; O\Big(n^{2r-5}a^3 p^{2\binom{r}{2}-1} + n^{2r-5}a^2 p^{2\binom{r}{2}-3} + n^{r-1}a^2 p^{\binom{r}{2}+r-1}\Big) \; \ll \; (pan)^2,$$

since $pan \gg 1$ and $a = \Omega(p^{-2})$. It follows, by Chebyshev's inequality, that $X \leqslant pan/r^2$ with high probability, and hence there exists a set of at least $pan/2$ edges that can be oriented freely, as required. $\qquad\square$

It is straightforward to generalize this idea to prove the lower bounds in Theorem 1.8 and Conjectures 1.3 and 1.6, so we will be somewhat brief with the details. To avoid tedious calculations, we will prove the slightly weaker bounds given by Markov's inequality, rather than the slightly stronger bounds given by the second moment method.

Proposition 2.4. *Let $r \geqslant 3$ and $\omega \gg 1$. If $p \geqslant n^{-(r-2)/(r-1)}$, then*

$$\log D\big(G(n,p), C_r^{\circlearrowleft}\big) \geqslant \frac{n}{\omega \cdot p^{1/(r-2)}}$$

with high probability as $n \to \infty$.

Proof. The lower bound in Proposition 2.2 is stronger when $p = \Omega(1)$, so we may assume $p = o(1)$. Set $a = p^{-(r-1)/(r-2)}/\omega = o(n)$, and consider any orientation $\vec{G}$ of $G(n,p)$ in which all a-long edges are oriented forward. We claim that if a copy $C = (c_1, \ldots, c_r)$ of C_r in $G(n,p)$ is oriented cyclically in $\vec{G}$, then $|c_i - c_j| \leqslant (r-1)(a-1)$ for every $i, j \in [r]$. To see this, simply note that every backwards edge has length at most $a - 1$, and at most $r - 1$ of the edges can be directed backwards.

In particular, if we orient only the a-long edges of $G(n,p)$ forward, the number of choices for $V(C)$ of size r such that C could form a $C_r^{\circlearrowleft}$ for some orientation of the a-short edges is at most $n(ra)^{r-1}$. Therefore, the expected number of such copies of C_r in $G(n,p)$ is $O\big(p^r a^{r-1} n\big) \ll pan$, so the claimed bound follows as in the proof of Proposition 2.1. $\qquad\square$

The next proposition implies the lower bounds in Conjecture 1.6 and Theorem 1.8.

Proposition 2.5. *Let $r \geqslant 3$ and $\omega \gg 1$. If $p \geqslant n^{-2/(r+1)}$, then*

$$\log S_r(n,p) \geqslant \frac{n}{\omega \cdot p^{(r-1)/2}}$$

with high probability as $n \to \infty$.

Proof. By Proposition 2.2, we may again assume that $p = o(1)$. Set $a = p^{-(r+1)/2}/\omega = o(n)$. Denoting by v and w the leftmost and rightmost vertices of a strongly connected copy of K_r, as in Proposition 2.4, we see that $|w - v| \leqslant (r-1)(a-1)$. Therefore, all vertices in a "dangerous" copy of K_r are within distance $O(a)$ of each other. The expected number of such copies of K_r in $G(n,p)$ is $O\big(p^{\binom{r}{2}} a^{r-1} n\big) \ll pan$, so the claimed bound again follows as before. $\qquad\square$

Finally, the lower bounds of the form $\big(1 + o(1)\big)p\binom{n}{2}$ in Theorems 1.5 and 1.8 follow easily from the observation that, by Markov's inequality, if $p \ll n^{-2/(r+1)}$ then with high probability $G(n,p)$ contains $o(pn^2)$ copies of K_r. Indeed, if we orient the edges that are contained in a copy of K_r according to a fixed linear order of the vertex set, then we may orient the remaining edges arbitrarily.

3. Upper bound for triangles

In this section we prove the following theorem, which implies the upper bound in Theorem 1.2, and whose proof contains the key idea introduced in this paper.

Theorem 3.1. *Let $0 < p \leqslant 1$. Then, with high probability as $n \to \infty$,*

$$D\big(G(n,p), C_3^{\circlearrowleft}\big) \leqslant \exp\left(\frac{6n(\log n)^2}{p}\right).$$

In order to prove Theorem 3.1, we start with the simple (but key) observation that knowing the orientation of two edges of a triangle in a $C_3^{\circlearrowleft}$-free graph is sometimes enough to uniquely determine the orientation of the third edge. Given any $C_3^{\circlearrowleft}$-free orientation $\vec{G}$ of $G(n,p)$, we will use this fact to find a small set of "non-redundant" edges $S \subset E(\vec{G})$ such that the prescribed orientation $\vec{G}$ is uniquely determined by S. This is formalized in the deterministic lemma below. We remark that the next sections will use similar deterministic statements (in Claim 4.4, Claim 5.4, and Lemma 6.4).

Lemma 3.2. *Let $\vec{G}$ be a $C_3^{\circlearrowleft}$-free orientation of a graph G on n vertices. There exists a set $S \subset E(\vec{G})$ with*

$$|S| \leqslant 2n \cdot \alpha(G)$$

such that $\vec{G}$ is the unique $C_3^{\circlearrowleft}$-free orientation of G containing S.

Proof. We prove the lemma by induction on n. For $n = 1$ it is trivial, so let $n \geqslant 2$ and assume that the statement is true for $n - 1$. Pick a vertex $v \in V(G)$, and let S' be the set given by the lemma applied to $\vec{G}' = \vec{G} - v$. Note that $|S'| \leqslant 2(n-1)\alpha(G)$, since $\alpha(G') \leqslant \alpha(G)$, and that $\vec{G}'$ is the unique $C_3^{\circlearrowleft}$-free orientation of G' containing S'.

Now, let $T \subset E(\vec{G}) \setminus E(\vec{G}')$ be minimal such that $\vec{G}$ is the unique $C_3^{\circlearrowleft}$-free orientation of G containing $S' \cup T$. (Note that such a set exists, since $E(\vec{G}) \setminus E(\vec{G}')$ has this property.) Set $T^+ := \{(v,w) : (v,w) \in T\}$ be the edges of T oriented away from v; we claim that

$$|T^+| \leqslant \alpha(G). \tag{5}$$

Indeed, suppose that there exists an edge $(u,w) \in E(\vec{G}')$ such that $(v,u), (v,w) \in T^+$. Recall that only one orientation of the edge uw can appear together with S' in a $C_3^{\circlearrowleft}$-free graph, and therefore every $C_3^{\circlearrowleft}$-free graph containing $T \setminus \{(v,w)\}$ contains the edge (v,w), as we can deduce the orientation of vw from those of uv and uw. Thus, setting $T' := T \setminus \{(v,w)\}$, it follows that $\vec{G}$ is the unique $C_3^{\circlearrowleft}$-free orientation of G containing $S' \cup T'$, contradicting the minimality of T. Hence the set $\{w : (v,w) \in T^+\}$ must in fact be independent, and therefore we have $|T^+| \leqslant \alpha(G)$, as claimed.

To complete the proof, simply note that the same bound holds for $T^- := T \setminus T^+$, by symmetry, and hence, setting $S := S' \cup T$, we have $|S| \leqslant 2n \cdot \alpha(G)$, as required. $\square$

We remark that Lemma 3.2 can be stated as an extremal result about a deterministic process that resembles graph bootstrap percolation (also known as weak saturation), see e.g. [5, 6].

The following corollary is an immediate consequence of Lemma 3.2.

Corollary 3.3. *A graph G on n vertices admits at most*

$$\sum_{i=0}^{2n\alpha(G)} \binom{e(G)}{i} 2^i$$

$C_3^{\circlearrowleft}$-*free orientations.*

We can now prove Theorem 3.1.

Proof of Theorem 3.1. Note that if $p^2 n \ll (\log n)^2$ then the trivial bound $D(G, C_3^{\circlearrowleft}) \leqslant 2^{e(G)}$ is sufficient, so we may assume this is not the case. It follows that, with high probability, we have $e\big(G(n,p)\big) = \big(1 + o(1)\big)p\binom{n}{2}$ and $\alpha\big(G(n,p)\big) \leqslant \frac{3\log n}{p}$ (by the first moment method, since the expected number of independent sets of this size is $o(1)$). Therefore, by Corollary 3.3,

$$D\big(G(n,p), C_3^{\circlearrowleft}\big) \leqslant \sum_{i=0}^{2n \cdot \alpha(G(n,p))} \binom{e\big(G(n,p)\big)}{i} 2^i$$

$$\leqslant \frac{6n\log n}{p} \cdot \binom{pn^2}{\frac{6n\log n}{p}} \cdot n^{6n/p} \leqslant \exp\left(\frac{6n(\log n)^2}{p}\right)$$

with high probability, as required. $\qquad\square$

4. Avoiding non-transitive tournaments

Recall that $T_r(n,p)$ denotes the number of orientations of $G(n,p)$ in which every copy of K_r is transitive. In this section we prove the following two theorems, which (together with Theorem 3.1) imply the upper bounds in Theorem 1.5.

Theorem 4.1. *Let $r \geqslant 4$. If $p > n^{-2/(r+2)}$, then*

$$T_r(n,p) \leqslant \exp\left(\frac{O\big(n(\log n)^2\big)}{p}\right)$$

with high probability as $n \to \infty$.

Theorem 4.2. *Let $r \geqslant 4$. If $p \leqslant n^{-2/(r+2)}$, then*

$$T_r(n,p) \leqslant \exp\left(\frac{O\big(n^{4-r}(\log n)^2\big)}{p^{\binom{r}{2}-2}}\right)$$

with high probability as $n \to \infty$.

Before proving these theorems, let us first note that a slightly weaker version of Theorem 4.1 follows easily from Theorem 3.1. Indeed, if

$$p \gg n^{-2/(r+2)}(\log n)^{2/(r+2)(r-3)},$$

then with high probability every triangle in $G(n,p)$ is contained in a copy of K_r (see [15]), and hence every orientation of $G(n,p)$ in which every K_r is transitive is also $C_3^{\circlearrowleft}$-free.

In order to remove this polylogarithmic factor, and to prove Theorem 4.2, we will use the following slightly technical lemma, which follows easily from the Janson inequalities (see, e.g., [3, 11]). For completeness, we provide a proof in the appendix.

Lemma 4.3. *For each $r \geqslant 4$, there exists $C > 0$ such that if*

$$p \gg n^{-2/(r+1)} (\log n)^{4/(r+1)(r-2)},$$

then the following holds with high probability as $n \to \infty$. Set

$$t_r(n,p) := \begin{cases} Cp^{2-\binom{r}{2}} n^{3-r} \log n & \text{if } p \leqslant n^{-2/(r+2)} \\ \frac{C \log n}{p} & \text{if } p > n^{-2/(r+2)}. \end{cases} \tag{6}$$

For every $v \in V\big(G(n,p)\big)$ and every $T \subset N(v)$ of size at least $t_r(n,p)$, there exists a copy of K_r in $G(n,p)$ containing v and at least two vertices of T.

To prove Theorems 4.1 and 4.2, we now simply repeat the proof of Theorem 3.1, replacing $\alpha(G)$ by $t_r(n,p)$.

Proof of Theorems 4.1 and 4.2. We begin with a deterministic claim corresponding to Lemma 3.2. Given $n \in \mathbb{N}$ and $p \in (0,1)$, let G be a graph on n vertices that satisfies the conclusion of Lemma 4.3, that is, for every $v \in V(G)$ and every $T \subset N(v)$ of size at least $t_r(n,p)$, there exists a copy of K_r in G containing v and at least two vertices of T.

Claim 4.4. *Let $\vec{G}$ be an orientation of G in which every K_r is transitive. There exists a set $S \subset E(\vec{G})$ with*

$$|S| \leqslant 2n \cdot t_r(n,p)$$

such that $\vec{G}$ is the unique orientation of G containing S in which every K_r is transitive.

Proof of Claim 4.4. Observe first that, since every copy of K_r in G is transitive in $\vec{G}$, every triangle in G that is contained in a copy of K_r is also transitive. Fix an ordering $v_1, \ldots, v_n$ of $V(G)$, let $k \in [n]$, and suppose that we have already found a set $S_k \subset E\big(\vec{G}[\{v_1, \ldots, v_k\}]\big)$ such that

$$|S_k| \leqslant 2k \cdot t_r(n,p),$$

and $\vec{G}[\{v_1, \ldots, v_k\}]$ is the unique orientation of $G[\{v_1, \ldots, v_k\}]$ containing S_k in which every triangle that is contained in a copy of K_r in G is transitive.

Now, let $S_{k+1} \subset E\big(\vec{G}[\{v_1, \ldots, v_{k+1}\}]\big)$ be minimal such that $S_{k+1} \supset S_k$, and such that $\vec{G}[\{v_1, \ldots, v_{k+1}\}]$ is the unique orientation of $G[\{v_1, \ldots, v_{k+1}\}]$ containing S_{k+1} in which every triangle that is contained in a copy of K_r in G is transitive. Setting $T^+ := \{w : (v_{k+1}, w) \in S_{k+1}\}$, we claim that

$$|T^+| \leqslant t_r(n,p). \tag{7}$$

Indeed, if $|T^+| > t_r(n,p)$ then there exists a copy of K_r in G containing v_{k+1} and at least two vertices $u, w \in T^+$, and hence the triangle uwv_{k+1} must be transitive in $\vec{G}$. But this means that (as in the proof of Lemma 3.2) we can deduce the orientation of either

9

uv_{k+1} or wv_{k+1} from that of the other, together with that of uw, and hence S_{k+1} is not minimal. This contradiction proves (7). Similarly, defining $T^- := \{w : (w, v_{k+1}) \in S_{k+1}\}$, an analogous argument shows that $|T^-| \leqslant t_r(n, p)$. Proceeding inductively, we obtain a set $S = S_n$ as claimed. $\qquad \square$

We now prove Theorem 4.1. Suppose $p > n^{-2/(r+2)}$. It follows from Claim 4.4 (and Lemma 4.3) that, with high probability,

$$T_r(n, p) \leqslant \sum_{i=0}^{2n \cdot t_r(n,p)} \binom{pn^2}{i} 2^i \leqslant n^{3Cn/p} \cdot \left(\frac{pn^2}{\frac{2Cn \log n}{p}} \right) \leqslant \exp\left(\frac{2Cn(\log n)^2}{p} \right),$$

as required.

To prove Theorem 4.2, we suppose from now on that $p \leqslant n^{-2/(r+2)}$. Note that if $p^{\binom{r}{2}-1} n^{r-2} \leqslant (\log n)^2$ then the trivial bound $2^{e(G)}$ is sufficient, so we may assume this is not the case. It follows that, with high probability, we have $e(G(n,p)) = (1 + o(1))p\binom{n}{2}$ and $G(n, p)$ satisfies the conclusion of Lemma 4.3. Hence, by Claim 4.4,

$$T_r(n, p) \leqslant \sum_{i=0}^{2n \cdot t_r(n,p)} \binom{pn^2}{i} 2^i \leqslant \exp\left(O\big(t_r(n, p) \cdot n \log n\big) \right)$$

with high probability, as required. $\qquad \square$

5. Avoiding strongly connected tournaments

Recall that $S_r(n, p)$ denotes the number of orientations of $G(n, p)$ in which no copy of K_r is strongly connected. In this section we prove the following theorems, which (together with Theorem 3.1) imply the upper bound in Theorem 1.8.

Theorem 5.1. *Let $r \geqslant 4$. If $n^{-2/(r+1)} \ll p \leqslant (\log n)^{-2/(r-2)}$, then*

$$S_r(n, p) \leqslant \exp\left(\frac{O\big(n(\log n)^{1+1/(r-2)}\big)}{p^{(r-1)/2}} \right) \tag{8}$$

with high probability as $n \to \infty$.

Theorem 5.2. *If $p > (\log n)^{-2/(r-2)}$, then*

$$S_r(n, p) \leqslant \exp\left(\frac{O\big(n(\log n)^2\big)}{p} \right) \tag{9}$$

with high probability as $n \to \infty$.

The proofs of Theorems 5.1 and 5.2 are similar to the proofs of Theorems 4.1 and 4.2. Instead of Lemma 4.3, we will use the following straightforward fact, which also follows easily from the Janson inequalities (see the appendix).

Lemma 5.3. *For every $r \geqslant 3$, there exists $C > 0$ such that the following holds with high probability as $n \to \infty$. Set*

$$s_r(n, p) = \begin{cases} \frac{C(\log n)^{1/(r-1)}}{p^{r/2}} & \text{if } p \leqslant (\log n)^{-2/(r-1)} \\ \frac{C \log n}{p} & \text{if } p > (\log n)^{-2/(r-1)}. \end{cases}$$

Then every set $S \subset V\big(G(n, p)\big)$ with $|S| \geqslant s_r(n, p)$ contains a copy of K_r.

We can now easily deduce Theorem 5.1, using the method of the previous two sections.

Proof of Theorems 5.1 and 5.2. Once again, we begin with a deterministic claim (cf. Lemma 3.2 and Claim 4.4). Given $n \in \mathbb{N}$ and $p \in (0, 1)$, let G be a graph on n vertices that satisfies the conclusion of Lemma 5.3 for $r - 1$, that is, every set of vertices of G of size at least $s_{r-1}(n, p)$ contains a copy of K_{r-1}.

Claim 5.4. *Let $\vec{G}$ be an orientation of G in which no copy of K_r is strongly connected. There exists a set $S \subset E(\vec{G})$ with*

$$|S| \leqslant 2n \cdot s_{r-1}(n, p)$$

such that $\vec{G}$ is the unique orientation of G containing S with no strongly connected K_r.

Proof of Claim 5.4. We will use the simple observation that every orientation of K_{r-1} contains a Hamiltonian path. Now, we can simply choose the set S greedily, vertex by vertex, as before. To be precise, fix an ordering $v_1, \ldots, v_n$ of $V(G)$, let $k \in [n]$, and suppose that we have already found a set $S_k \subset E\big(\vec{G}[\{v_1, \ldots, v_k\}]\big)$ such that

$$|S_k| \leqslant 2k \cdot s_{r-1}(n, p),$$

and $\vec{G}[\{v_1, \ldots, v_k\}]$ is the unique orientation of $G[\{v_1, \ldots, v_k\}]$ containing S_k in which no copy of K_r is strongly connected.

Now, let $S_{k+1} \subset E\big(\vec{G}[\{v_1, \ldots, v_{k+1}\}]\big)$ be minimal such that $S_{k+1} \supset S_k$, and such that $\vec{G}[\{v_1, \ldots, v_{k+1}\}]$ is the unique orientation of $G[\{v_1, \ldots, v_{k+1}\}]$ containing S_{k+1} in which no copy of K_r is strongly connected. Setting $T^+ := \{w : (v_{k+1}, w) \in S_{k+1}\}$, we claim that

$$|T^+| \leqslant s_{r-1}(n, p). \tag{10}$$

Indeed, if T^+ were larger than this, then there would exist a copy of K_{r-1} in $G[T^+]$, and this copy of K_{r-1} contains a Hamiltonian path in $\vec{G}$, from a to b, say. Moreover, the orientations of the edges in this path are determined by S_k, and we can therefore deduce the orientation of the edge bv_{k+1} from $S_{k+1} \setminus \{(v_{k+1}, b)\}$. This contradicts the minimality of S_{k+1}, and hence proves (10). Defining $T^- := \{w : (w, v_{k+1}) \in S_{k+1}\}$, a similar argument shows that $|T^-| \leqslant s_{r-1}(n, p)$. Taking $S = S_n$ finishes the proof. $\square$

Suppose now that $n^{-2/(r+1)} \ll p \leqslant (\log n)^{-2/(r-2)}$. It follows from Claim 5.4 (and the first case of Lemma 5.3) that, with high probability,

$$S_r(n, p) \leqslant \sum_{i=0}^{2n \cdot s_{r-1}(n,p)} \binom{pn^2}{i} 2^i \leqslant \exp\left(\frac{O\big(n(\log n)^{1+1/(r-2)}\big)}{p^{(r-1)/2}}\right),$$

11

and Theorem 5.1 is proved. To prove Theorem 5.2, note that if $p > (\log n)^{-2/(r-2)}$, then a similar calculation using Lemma 5.3 shows that

$$S_r(n,p) \leqslant \sum_{i=0}^{2n \cdot s_{r-1}(n,p)} \binom{pn^2}{i} 2^i \leqslant \exp\left(\frac{O\big(n(\log n)^2\big)}{p}\right)$$

with high probability, as claimed. $\qquad\square$

6. Avoiding longer cycles

In this section, we show an upper bound for the number of orientations avoiding oriented cycles of length r (denoted by $C_r^{\circlearrowleft}$). As stated in Conjecture 1.3, we believe substantially better upper bounds are possible for $r \geqslant 4$.

Theorem 6.1. *Let $r \geqslant 3$. Then*

$$\log D\big(G(n,p), C_r^{\circlearrowleft}\big) = \widetilde{O}\left(\frac{n}{p}\right)$$

with high probability as $n \to \infty$.

The proof of Theorem 6.1 follows directly from Lemma 6.4 below, which is a generalisation of Lemma 3.2 to longer cycles. To prove Lemma 6.4, we start with the following simple observation.

Lemma 6.2. *Let $\vec{G}$ be a $C_r^{\circlearrowleft}$-free graph, v a vertex of $\vec{G}$ and $P = (w_1, \ldots, w_k)$ a directed path in $\vec{G}[N(v)]$ such that $w_i \in N^+(v)$ for $1 \leqslant i \leqslant r-2$. Then $P \subset N^+(v)$.*

Proof. If the conclusion were not true, there would exist a minimal i such that $w_i \in N^-(v)$. By hypothesis, we would have $i > r-2$. But then $(v, w_{i-(r-2)}, \ldots, w_i, v)$ would be a directed cycle of length r, a contradiction. $\qquad\square$

By reversing the orientation of all edges of G, we can deduce from Lemma 6.2 that if the last $r-2$ vertices of a directed path contained in $N(v)$ are in $N^-(v)$, then the whole path is in $N^-(v)$.

The main additional ingredient in the proof of Lemma 6.4 is the Gallai–Milgram theorem [8]; the proof of Theorem 6.1 was inspired by a similar application in [1].

Theorem 6.3 (Gallai–Milgram [8])**.** *The vertex set of every directed graph $\vec{G}$ can be partitioned into at most $\alpha(G)$ vertex-disjoint directed paths.*

We are now ready to prove the main lemma of this section.

Lemma 6.4. *Let $\vec{G}$ be a $C_r^{\circlearrowleft}$-free orientation of a graph G on n vertices. There exists a set $S \subset E(\vec{G})$ with*

$$|S| \leqslant 2n \cdot (r-2)\alpha(G)$$

such that $\vec{G}$ is the unique $C_r^{\circlearrowleft}$-free orientation of G containing S.

Proof. We proceed by induction on n. Let $v \in V(G)$ be any vertex of G, and let $G' = G \setminus v$. By induction, there exists a set $S' \subset E(\vec{G}')$ of size $2(n-1) \cdot (r-2)\alpha(G)$ such that $\vec{G}' = \vec{G} \setminus v$ is the unique $C_r^{\circlearrowright}$-free orientation of G' containing S'. Our aim is to find a set of edges T of size $2(r-2)\alpha(G)$ such that $\vec{G}$ is the unique $C_r^{\circlearrowright}$-free orientation of G containing $S' \cup T$.

We start by applying Theorem 6.3 to partition the graph $\vec{G}[N^+(v)]$ into a collection $\mathcal{P}^+$ of at most $\alpha(G)$ oriented paths, and define T^+ to be the set of edges given by

$$T^+ = \big\{ (v, w) : w \text{ is one of the first } r-2 \text{ vertices in some } P \in \mathcal{P}^+ \big\}.$$

We define T^- similarly by decomposing $\vec{G}[N^-(v)]$ into at most $\alpha(G)$ oriented paths and taking the last $r-2$ vertices of each path. We claim that $T = T^+ \cup T^-$ has the desired property.

To check the claim, we must show that any $C_r^{\circlearrowright}$-free orientation $\vec{H}$ of G containing $S' \cup T$ equals $\vec{G}$. By the induction hypothesis, $\vec{G}[V(G')] = \vec{H}[V(G')]$, so it suffices to show that $N_G^+(v) \subset N_H^+(v)$ and $N_G^-(v) \subset N_H^-(v)$. Since $\vec{H}$ contains $T^+ \subset T$, we know $N_H^+(v)$ contains the first $r-2$ vertices in each path of $\mathcal{P}^+$. By Lemma 6.2, $N_H^+(v)$ contains all paths in $\mathcal{P}^+$, and since $\mathcal{P}^+$ was a partition of $N_G^+(v)$, we conclude that $N_G^+(v) \subset N_H^+(v)$. A similar argument works for T^-, and therefore we have checked the claim. Taking $S = S' \cup T$ finishes the proof. $\square$

Proof of Theorem 6.1. This proof closely mirrors that of Theorem 3.1, with Lemma 6.4 replacing Lemma 3.2. We obtain

$$D\big(G(n,p), C_r^{\circlearrowright}\big) \leqslant \sum_{i=0}^{2n(r-2)\cdot\alpha(G(n,p))} \binom{e\big(G(n,p)\big)}{i} 2^i \leqslant \exp\left(\frac{6n(r-2)(\log n)^2}{p}\right)$$

with high probability, as required. $\square$

7. Open problems

In this section we will mention some further open problems and possible directions for future research; in particular, we will discuss the problem of removing the polylogarithmic factors that separate the upper and lower bounds in Theorems 1.2, 1.5 and 1.8, and the problem of determining the behaviour of $D\big(G(n,p), \vec{H}\big)$, the number of $\vec{H}$-free orientations of $G(n,p)$, for an arbitrary oriented graph $\vec{H}$.

First, we remark that if $\vec{H}$ is contained in a transitive tournament then the situation is different; more precisely, the following theorem is an easy consequence of a well-known theorem of Rödl and Ruciński [14].

Theorem 7.1. *Let $r \geqslant 3$ and $p \gg n^{-2/(r+1)}$. With high probability, every orientation of $G(n,p)$ contains a transitive copy of K_r.*

To deduce Theorem 7.1 from the main theorem of [14], simply fix a linear order of the vertices, and define an edge-colouring from an orientation by colouring forward-pointing edges blue and backwards edges red. In this setting, a monochromatic copy of K_r corresponds to a transitively-oriented K_r (but not vice-versa). If $p \gg n^{-2/(r+1)}$, then [14, Theorem 1] ensures (with high probability) the existence of a monochromatic copy of K_r in any 2-colouring of $G(n, p)$, and therefore for this range of p it is impossible to avoid a transitively-oriented copy of K_r.

We remark that Theorem 7.1 does *not* give the correct threshold for the event "every orientation of $G(n, p)$ contains a transitive triangle", since every orientation of K_4 contains a transitive triangle, and the event $\{K_4 \subset G(n, p)\}$ has a threshold at $\Theta(n^{-2/3})$. Nevertheless, we suspect that $n^{-2/(r+1)}$ is the correct threshold for the event "every orientation of $G(n, p)$ contains a transitive copy of K_r" for every $r \geqslant 4$.

Therefore, we will assume throughout this section that $\vec{H}$ contains a cycle. For this case, we state (somewhat imprecisely) the central question that is suggested by the work in this paper.

Question 7.2. *Is the lower bound construction described in Section 2 always sharp?*

The results proved in this paper provide some evidence in favour of a positive answer to this question (at least in a weak sense). It is moreover plausible that it is true in a much stronger sense: that (with high probability) almost all $\vec{H}$-free orientations of $G(n, p)$ are "close" to one of the orientations given by the construction described in Section 2.[1]

Problem 7.3. *Determine the typical structure of an $\vec{H}$-free orientation of $G(n, p)$.*

For example, in the case $\vec{H} = C_3^{\circlearrowleft}$ one might hope to prove that if $p \gg n^{-1/2}$, then the following holds with high probability: for almost all $C_3^{\circlearrowleft}$-free orientations of $G(n, p)$, there exists an ordering of the vertices such that $\Theta(n/p)$ edges are oriented backwards, and all but $o(n/p)$ of those edges have length $O(1/p^2)$.

7.1. **Removing the polylogarithmic terms.** An important (and probably very challenging) step in the direction of Problem 7.3 would be to remove the polylogarithmic factor between our upper and lower bounds on $\log T_r(n, p)$ and $\log S_r(n, p)$.

Problem 7.4. *Determine the typical values of $\log T_r(n, p)$ and $\log S_r(n, p)$ up to a constant factor for each $r \geqslant 3$ and every function $p \gg n^{-2/(r+1)}$.*

Note that some polylogarithmic factor is necessary, at least when p is large, since

$$T_r(n, 1) = S_r(n, 1) = n!$$

for every $r \geqslant 3$. In the case $r = 3$, we conjecture that a combination of the lower bounds given by Propositions 2.1 and 2.2 is sharp up to the implicit constant factor in the exponent.

---

[1]Of course, one can ask the same questions for a family $\vec{\mathcal{H}}$ of forbidden oriented graphs, such as the family of non-transitive tournaments (or the family of strongly-connected tournaments) of a given size.

Conjecture 7.5. $D\big(G(n,p), C_3^{\circlearrowleft}\big) \leqslant 2^{O(n/p)} \cdot n!$ *for every* $p \geqslant n^{-1/2}$.

7.2. General forbidden structures.

In this section we will discuss the general lower bound on $D\big(G(n,p), \vec{H}\big)$ given by the construction described in Section 2, where $\vec{H}$ is an arbitrary oriented graph that contains a cycle. Let $m_2(\vec{H})$ denote the 2-density of the underlying graph of $\vec{H}$,

$$m_2(\vec{H}) = \max\left\{ \frac{e(\vec{F}) - 1}{v(\vec{F}) - 2} \; : \; \vec{F} \subset \vec{H}, \, v(\vec{F}) \geqslant 3 \right\}.$$

Note that any oriented graph can be decomposed into strongly connected components in a unique way, and let $s(\vec{H})$ denote the number of strongly connected components of $\vec{H}$.

Proposition 7.6. *Suppose that* $\vec{H}$ *contains a cycle. If* $p \gg n^{-1/m_2(\vec{H})}$ *and* $\omega \gg 1$, *then*

$$\log D\big(G(n,p), \vec{H}\big) \geqslant \frac{pn}{\omega} \cdot \max\left\{ \left(p^{e(\vec{F})-1} n^{s(\vec{F})-1}\right)^{\frac{-1}{v(\vec{F})-s(\vec{F})-1}} \; : \; \vec{F} \subset \vec{H}, \, v(\vec{F}) \geqslant s(\vec{F})+2 \right\}$$

with high probability as $n \to \infty$.

Proof. When $p = \Omega(1)$, the claimed lower bound is of order $o(n)$, and therefore follows from Proposition 2.2. When $p = o(1)$, we first need to observe that, since $\vec{H}$ contains a cycle, there is a strongly connected component $\vec{F} \subset \vec{H}$ with $v(\vec{F}) \geqslant 3$, and this implies that there exists $\vec{F} \subset \vec{H}$ with $v(\vec{F}) \geqslant s(\vec{F})+2$ and $p^{e(\vec{F})-1} n^{s(\vec{F})-1} = o(1)$. Given any such $\vec{F}$, consider the construction of Section 2 with

$$a = \frac{1}{\omega} \cdot \left(p^{e(\vec{F})-1} n^{s(\vec{F})-1}\right)^{\frac{-1}{v(\vec{F})-s(\vec{F})-1}},$$

that is, orient all edges of length at least a from left to right. Note that $a \gg 1$, by our choice of $\vec{F}$, and that $p \gg n^{-1/m_2(\vec{H})}$ implies that $a = o(n)$. Now, observe that the expected number of potential copies of $\vec{F}$ in $G(n,p)$ is

$$O\left(n^{s(\vec{F})} a^{v(\vec{F})-s(\vec{F})} p^{e(\vec{F})}\right) \ll pan,$$

since any two vertices in the same strongly connected component of a copy of $\vec{F}$ must lie within distance $O(a)$ of one another, and since $v(\vec{F}) \geqslant s(\vec{F})+2$. By Markov's inequality, it follows that with high probability there exists a set of at least $pan/2$ edges that can be oriented freely without creating a copy of $\vec{F}$, and hence of $\vec{H}$. Since this holds for each $\vec{F} \subset \vec{H}$ with $v(\vec{F}) \geqslant s(\vec{F})+2$ and $p^{e(\vec{F})-1} n^{s(\vec{F})-1} = o(1)$, the claimed bound follows. $\square$

We can now rephrase Question 7.2 more precisely in this setting.

Question 7.7. *Suppose that* $\vec{H}$ *contains a cycle, and let* $p \gg n^{-1/m_2(\vec{H})}$. *Is it true that*

$$\log D\big(G(n,p), \vec{H}\big) = \widetilde{\Theta}\left(\max_{\vec{F} \subset \vec{H}, \, v(\vec{F}) \geqslant s(\vec{F})+2} \left\{ pn \cdot \left(p^{e(\vec{F})-1} n^{s(\vec{F})-1}\right)^{\frac{-1}{v(\vec{F})-s(\vec{F})-1}} \right\} \right)$$

with high probability as $n \to \infty$?

Observe that the oriented subgraph $\vec{F} \subset \vec{H}$ which corresponds to the maximum in Proposition 7.6 depends on p, and in general it can change arbitrarily many times as p increases. A positive answer to Question 7.7 would therefore imply the existence of $\vec{H}$ for which $D\big(G(n,p), \vec{H}\big)$ exhibits arbitrarily many thresholds between $n^{-1/m_2(\vec{H})}$ and 1.

<center>ACKNOWLEDGEMENTS</center>

<center>APPENDIX</center>

In this appendix, we will prove Propositions 2.3 and 2.2 and Lemmas 4.3 and 5.3. We start out by recalling Janson's inequality, which will be used in two of the proofs (see, e.g., [3, 11]).

Lemma A.1 (Janson's inequality). *Suppose that $\{B_i\}_{i \in I}$ is a family of subsets of a finite set X and let $0 \leqslant p \leqslant 1$. Let*

$$\mu = \sum_{i \in I} p^{|B_i|} \qquad \text{and} \qquad \Delta = \sum_{i \sim j} p^{|B_i \cup B_j|}, \tag{A.11}$$

where the sum is over ordered pairs and $i \sim j$ means that $i \neq j$ and $B_i \cap B_j \neq \emptyset$. Then,

$$\mathbb{P}(B_i \not\subset X_p \text{ for all } i \in I) \leqslant \exp\left(-\min\left\{\frac{\mu}{2}, \frac{\mu^2}{2\Delta}\right\}\right).$$

A.1. **Proofs.** We now prove Propositions 2.3 and 2.2 and Lemmas 4.3 and 5.3.

Proof of Proposition 2.3. We will basically repeat the proof strategy of Proposition 2.1; namely, we will count the number of edges which are contained in a "dangerous K_r", that is, a K_r containing two a-short edges incident to a single vertex. This time, we will replace Markov's inequality by an application of the second moment method.

Set $a = n^{3-r} p^{1-\binom{r}{2}}/4r^2$, and observe that the bounds on p imply $n \gg a \geqslant p^{-2}/4r^2$. Note that any non-transitive copy of K_r contains a cyclic triangle, and therefore contains two a-short edges sharing a vertex. Let X be a random variable counting the number of such K_rs in $G(n,p)$. By first choosing the vertex to which the two a-short edges are incident, then choosing the other endpoints of those edges, then choosing the remaining $n-3$ vertices arbitrarily, we obtain

$$\mathbb{E}(X) \leqslant n \binom{2a}{2} \binom{n-3}{r-3} p^{\binom{r}{2}} \leqslant \frac{pan}{2r^2}.$$

The above estimate overcounts by considering three of the vertices as special. Compensating for this, we obtain

$$\mathbb{E}(X) \geqslant n^{r-2} a^2 p^{\binom{r}{2}}/r^r \geqslant \frac{pan}{4r^{r+2}}.$$

<center>16</center>

We are done if we show that $X \leqslant pan/r^2$ with high probability, because forcing every edge from each of the X copies of K_r to be oriented forward leaves at least $pan/2$ free edges.

All that is left is a routine application of the second moment method, the details of which we include for completeness. Note first that $\mathbb{E}(X) = \Omega(pan) = \Omega(n/p) \gg 1$. In order to compute $\mathrm{Var}(X)$, we consider pairs (R_1, R_2) of edge-intersecting copies of K_r in K_n, each of which having two a-short edges sharing a vertex. We split into two cases:

(1) To count pairs with $|R_1 \cap R_2| = 2$, we choose vertices one by one and observe that the restriction on a-short edges implies that three elements of $R_1 \cup R_2$ must be at distance at most a from previous vertices. Therefore, the contribution to the variance from those pairs is $\Theta\left(n^{2r-5}a^3p^{2\binom{r}{2}-1}\right) = \Theta\left((\mathbb{E}(X))^2/pan\right)$.

(2) If $|R_1 \cap R_2| = r'$ for $3 \leqslant r' \leqslant r - 1$, the contribution to the variance, which we will denote by $V_{r'}$, is $\Theta\left(n^{2r-r'-2}a^2p^{2\binom{r}{2}-\binom{r'}{2}}\right)$. Since $V_{r'+1}/V_{r'} = \Theta\left(n^{-1}p^{-r'}\right)$, we can conclude that $V_{r'}$ is unimodal for this range of p. It thus suffices to bound the cases $r' = 3$ or $r' = r - 1$. For those values, $(\mathbb{E}(X))^2/V_3 = \Theta(na^2p^3) \gg pan \gg 1$ and $(\mathbb{E}(X))^2/V_{r-1} = \Theta(ap^{2-r}) \gg 1$.

Therefore, the variance is a sum of $r - 2$ terms, each of which being $o\left((\mathbb{E}(X))^2\right)$. By Chebyshev's inequality, $X = \Theta\left(\mathbb{E}(X)\right)$, as desired. $\qquad\square$

To prove Proposition 2.2 we will need the following lemma of Goddard, Kenyon, King and Schulman [9].

Lemma A.2 ([9], Theorem 2.5). *Let G be a graph. Then the number of acyclic orientations of G is at least*

$$\prod_{v \in V(G)} \left((d(v)+1)!\right)^{1/(d(v)+1)}.$$

Since $d(v) = (1 + o(1))pn$ with high probability whenever $p \gg (\log n)/n$, Proposition 2.2 follows from Lemma A.2 and a standard calculation using Stirling's approximation.

Proof of Proposition 2.2. We will use the following weak version of Stirling's inequality, valid for every $n \geqslant 1$.

$$(2\pi n)^{1/2} \left(\frac{n}{e}\right)^n \leqslant n! \leqslant e\sqrt{n} \left(\frac{n}{e}\right)^n.$$

Let $x_v = d(v) + 1$. Lemma A.2, together with the fact that $x^{1/x}$ is monotone decreasing for $x \geqslant 3$, implies that there are at least

$$\prod_{v \in V} (x_v!)^{1/x_v} \geqslant \prod_{v \in V} \frac{x_v(2\pi x_v)^{1/2x_v}}{e} \geqslant \prod_{v \in V} \frac{x_v \cdot (2n)^{1/2n}\pi^{1/2x_v}}{e}$$

acyclic orientations of a graph G. In the $G(n, p)$ case, for $p \gg (\log n)/n$ it holds that $d(v) = (1 + o(1))pn$ for every v with high probability. Therefore, the number of acyclic

orientations of $G(n, p)$ is at least

$$\frac{(1 + o(1))^n p^n n^n \cdot \sqrt{2n}\pi^{1/4p}}{e^n} = p^n n! \, / \, e^{o(n)},$$

as desired. □

We remark that a simple argument of Manber and Tompa [12] shows that the number of acyclic orientations of any graph G is at most $\prod_{v \in V(G)}(d(v) + 1)$. For random graphs, Reidys [13] showed stronger upper and lower bounds that hold with high probability, as well as concentration of the logarithm of the number of acyclic orientations around its expectation.

We now proceed to the proofs of Lemmas 4.3 and 5.3.

Proof of Lemma 4.3. Recall that

$$t_r(n, p) := \begin{cases} Cp^{2-\binom{r}{2}}n^{3-r}\log n & \text{if } p \leqslant n^{-2/(r+2)}, \\ C(\log n)/p & \text{if } p > n^{-2/(r+2)}. \end{cases}$$

Fix $v \in [n]$ and $T \subset N(v)$ with $|T| = t$. Let $\mathcal{S}$ be the family of possible vertex sets for copies of K_r satisfying the conclusion of the lemma, that is,

$$\mathcal{S} = \left\{ S \in \binom{[n]}{k} : v \in S \text{ and } |S \cap T| = 2 \right\}.$$

Since $T \subset N(v)$, we may assume $G(n, p)$ contains all edges between v and vertices in T. Therefore, for a given $S \in \mathcal{S}$, the edges that must be in $G(n, p)$ for it to form a copy of K_r are

$$F(S) = \binom{S}{2} \setminus (\{v\} \times T).$$

We will use Janson's inequality to show that

$$\mathbb{P}(F(S) \not\subset E(G(n, p)) \text{ for every } S \in \mathcal{S}) \leqslant \exp(-2t \log n). \tag{A.12}$$

This requires bounding the parameters μ and Δ in (A.11). We start with bounding μ to obtain

$$\mu = \sum_{S \in \mathcal{S}} p^{|F(S)|} = \binom{n-1-t}{r-3}\binom{t}{2}p^{\binom{r}{2}-2} \geqslant \frac{Ct \log n}{r!}, \tag{A.13}$$

where in the last inequality we used the definition of t. For bounding Δ, we will need to consider pairs of copies of K_r which intersect. To analyze the possible intersections, let

$$\mathcal{K}(b, c) = \left\{(S_1, S_2) \in \mathcal{S}^2 : |(S_1 \cup S_2) \cap T| = b, |S_1 \cap S_2| = c \text{ and } F(S_1) \sim F(S_2)\right\}$$

be the family of pairs of copies of K_r which contribute to Δ, and note that $\mathcal{K}(b, c)$ is always empty unless $2 \leqslant b \leqslant 4$, due to the definition of $\mathcal{S}$, and unless $2 \leqslant c \leqslant r - 1$, due to the definition of the relation '$\sim$'. We can thus write

$$\Delta(b, c) = \sum_{(S_1, S_2) \in \mathcal{K}(b, c)} p^{|F(S_1) \cup F(S_2)|} \tag{A.14}$$

so that $\Delta = \sum_{b=2}^{4} \sum_{c=2}^{r-1} \Delta(b,c)$. Having defined μ and Δ, our main technical goal will be to bound $\min(\mu/2, \mu^2/2\Delta)$, which will take several steps. First, to get a more explicit expression for $\Delta(b,c)$, we will use the following simple fact.

Claim A.3. $|\mathcal{K}(b,c)| \leqslant \binom{t}{b} \binom{n-1-t}{2r-b-c-1} \cdot 3^{2r}$.

Proof. To estimate $|\mathcal{K}(b,c)|$, we will count the possible ways of choosing $S_1 \cup S_2$ then choose whether each element goes in S_1, S_2, or both. To form $S_1 \cup S_2$, we include v, then b elements from T, then the remaining elements from outside T. A moment's thought reveals that $|(S_1 \cup S_2) \setminus (T \cup \{v\})| = (2r - c) - (b + 1)$, implying the desired bound. $\square$

Note that some simple counting shows that $|F(S_1) \cup F(S_2)| = 2\binom{r}{2} - \binom{c}{2} - b$ for every $(S_1, S_2) \in \mathcal{K}(b,c)$. Applying this and Claim A.3 to (A.14), we obtain

$$\Delta(b,c) \leqslant \binom{t}{b} \binom{n-1-t}{2r-c-b-1} 3^{2r} p^{2\binom{r}{2} - \binom{c}{2} - b} =: D(b,c) \qquad (A.15)$$

To understand the behavior of $\Delta(b,c)$, we will make a series of claims.

Claim A.4. *The sets $\mathcal{K}(2,2)$ and $\mathcal{K}(3,2)$ are empty.*

Proof. Let $(S_1, S_2) \in \mathcal{K}(b,c)$. If $b = 2$, then S_1 and S_2 contain the same two elements x, y of T. Therefore, $\{v, x, y\} \subset S_1 \cap S_2$, and so $c \geqslant 3$. If $b = 3$, however, then S_1 and S_2 share a single element x of T; if we also assume $c = 2$, then the copies of K_r over S_1 and S_2 share a single edge, which must be $\{v, x\}$. But then $F(S_1) \cap F(S_2) = \emptyset$, because edges from v to T were excluded from consideration. $\square$

We now state two observations which will simplify future calculations.

Claim A.5. $t \ll pn/(\log n)$.

Proof. This follows from the hypothesis $p \gg n^{-2/(r+1)} (\log n)^{4/(r+1)(r-2)}$. $\square$

Claim A.6. *For every b and c, it holds that $D(b+1, c) \leqslant D(b,c)$.*

Proof. $D(b+1,c)/D(b,c) = \Theta(t/pn) \ll 1$, by Claim A.5. $\square$

We can now state and prove our main technical claim for proving Lemma 4.3. In the right-hand side of the inequality below, the denominator is sub-optimal for simplicity of proof. It is shown only to make the dependence on C explicit.

Claim A.7. *The following relation holds between μ and Δ.*

$$\min\left\{\frac{\mu}{2}, \frac{\mu^2}{2\Delta}\right\} > \frac{C}{6r \cdot 3^{2r} \cdot (r!)^2} \cdot t \log n.$$

Proof of Claim A.7. By (A.13), it suffices to bound μ^2/Δ. We will show that

$$\mu^2 \geqslant \frac{C}{3^{2r} \cdot (r!)^2} \cdot t \log n \cdot \Delta(b,c)$$

19

for every b and c such that $\mathcal{K}(b,c) \neq \emptyset$. We start with $b = 2$ (thus $3 \leqslant c \leqslant r - 1$, by Claim A.4). Using inequalities (A.13) and (A.15) and the fact that $t \geqslant C(\log n)/p$, we can bound

$$\frac{\mu^2}{\Delta(2,c)} \geqslant \frac{1}{3^{2r} \cdot (r!)^2} \cdot t^2 n^{c-3} p^{\binom{c}{2}-2} \geqslant \frac{C}{3^{2r} \cdot (r!)^2} \cdot t \log n,$$

where the last inequality used that $n^{c-3} p^{\binom{c}{2}-3} \geqslant 1$ for this range of p and c (for $c \geqslant 4$, this is equivalent to $p \geqslant n^{-2/(c+2)}$, which is true since $p \gg n^{-2/(r+1)}$ and $c \leqslant r - 1$). By Claim A.6, it only remains to check the case $(b,c) = (4,2)$, which also satisfies the desired inequality since

$$\frac{\mu^2}{\Delta(4,2)} = \Theta(pn) \gg t \log n,$$

by Claim A.5. Summing over valid values of b and c (at most $3r$ of them) and rearranging concludes the proof of Case 2, thereby proving the claim. $\qquad\square$

By choosing C large enough, Claim A.7 and Janson's inequality allow us to deduce (A.12). Since there are at most n choices for v and at most n^t choices for T, the conclusion of the lemma holds with high probability for all v and T by the union bound. $\qquad\square$

Now we prove Lemma 5.3. Since this is a simpler application of Janson's inequality than the previous one, we will be somewhat brief.

Proof of Lemma 5.3. Fix $r \geqslant 3$ and let C be sufficiently large. Recall that

$$s_r(n,p) = \begin{cases} \frac{C(\log n)^{1/(r-1)}}{p^{r/2}} & \text{if } p \leqslant (\log n)^{-2/(r-1)} \\ \frac{C \log n}{p} & \text{if } p > (\log n)^{-2/(r-1)}. \end{cases}$$

Denoting $s := s_r(n,p)$, by the union bound over all subsets of $[n]$ of size s it is enough to show that

$$\mathbb{P}(G(s,p) \not\supset K_r) \ll \binom{n}{s}^{-1}. \tag{A.16}$$

We define the random variable X to be the number of K_rs in $G(s,p)$ and compute the parameters μ and Δ in (A.11). We have

$$\mu = \binom{s}{r} p^{\binom{r}{2}} \geqslant \frac{C^{r-1}}{2r!} \cdot s \log n.$$

Now we need to consider interactions between two edge-intersecting copies of K_r. Partitioning according to the number of vertices in their intersection, we conclude that

$$\Delta = \sum_{a=2}^{r-1} \Delta(a) = \sum_{a=2}^{r-1} \binom{s}{2r-a} p^{2\binom{r}{2}-\binom{a}{2}}.$$

The following claim will allow us to use Janson's inequality.

Claim A.8. *We have*

$$\min\left\{\frac{\mu}{2}, \frac{\mu^2}{2\Delta}\right\} > \frac{1}{2r} \cdot \min\left\{\frac{C^{r-1}}{2r!}, C\right\} \cdot s \log n,$$

20

Proof. We split into two cases corresponding to the definition of s.

Case 1: $p \leqslant (\log n)^{-2/(r-1)}$. We will show that $\mu \geqslant \Delta/r$. We can bound

$$\frac{\mu}{\Delta(a)} \geqslant \frac{s^r \cdot p^{\binom{r}{2}}}{s^{2r-a} \cdot p^{2\binom{r}{2}-\binom{a}{2}}} = \left(s \cdot p^{(a+r-1)/2}\right)^{a-r}$$

$$\geqslant \left((\log n)^{1/(r-1)} p^{(a-1)/2}\right)^{a-r} \geqslant (\log n)^{(2-a)(a-r)/(r-1)} \geqslant 1,$$

where, in the second line, the first inequality follows from the definition of s, the second from the bound on p and the last from $2 \leqslant a \leqslant r-1$. We obtain $\mu \geqslant \Delta(a)$, and summing over a shows that $\mu \geqslant \Delta/r$, concluding this case.

Case 2: $p > (\log n)^{-2/(r-1)}$. In this case, we need to bound μ^2/Δ as well. We will do so by bounding $\mu^2/\Delta(a)$ for every $2 \leqslant a \leqslant r-1$. Replacing $s = C(\log n)/p$ and using the lower bound on p shows that

$$\frac{\mu^2}{\Delta(a)} = s\left(s \cdot p^{a/2}\right)^{a-1} \geqslant C^{a-1} \cdot s \cdot (\log n)^{(a-1)(r+1-a)/(r-1)} \geqslant C^{a-1} \cdot s \cdot \log n.$$

Therefore, $\mu^2 \geqslant C \cdot s \log n \cdot \Delta(a)$. We can again sum over a and rearrange to obtain $\mu^2/\Delta \geqslant (C/r) \cdot s \log n$, which proves this case and therefore the claim. $\qquad\square$

Choosing C large enough, we can apply Janson's inequality and conclude (A.16), which finishes the proof of Lemma 5.3. $\qquad\square$

References

[1] M. Aigner, E. Triesch, and Z. Tuza, *Searching for acyclic orientations of graphs*, Discrete Math. **144** (1995), no. 1-3, 3–10. ↑12

[2] P. Allen, Y. Kohayakawa, G. O. Mota, and R. F. Parente, *On the number of orientations of random graphs with no directed cycles of a given length*, Electron. J. Combin. **21** (2014), no. 1, Paper 1.52, 13. ↑1

[3] N. Alon and J. H. Spencer, *The probabilistic method*, Fourth edition, Wiley Series in Discrete Mathematics and Optimization, John Wiley & Sons, Inc., Hoboken, NJ, 2016. ↑9, 16

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

[5] J. Balogh, B. Bollobás, and R. Morris, *Graph bootstrap percolation*, Random Structures & Algorithms **41** (2012), no. 4, 413–440. ↑7

[6] B. Bollobás, *Weakly k-saturated graphs*, Beiträge zur Graphentheorie (Kolloquium, Manebach, 1967), 1968, pp. 25–31. ↑7

[7] P. Erdős, *Some new applications of probability methods to combinatorial analysis and graph theory*, Proceedings of the Fifth Southeastern Conference on Combinatorics, Graph Theory and Computing (Florida Atlantic Univ., Boca Raton, Fl, 1974), 1974, pp. 39–51. Congressus Numerantium, No. X. ↑1

[8] T. Gallai and A. N. Milgram, *Verallgemeinerung eines graphentheoretischen Satzes von Rédei*, Acta Sci. Math. (Szeged) **21** (1960), 181–186. ↑12

[9] W. Goddard, C. Kenyon, V. King, and L. J. Schulman, *Optimal randomized algorithms for local sorting and set-maxima*, SIAM J. Comput. **22** (1993), no. 2, 272–283. ↑5, 17

[10] B. Iriarte Giraldo, *Combinatorics of acyclic orientations of graphs: algebra, geometry and probability*, Ph.D. Thesis, 2015. ↑5

[11] S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000. ↑9, 16

[12] U. Manber and M. Tompa, *The effect of number of hamiltonian paths on the complexity of a vertex-coloring problem*, SIAM Journal on Computing **13** (1984), no. 1, 109–115. ↑18

[13] C. Reidys, *Acyclic orientations of random graphs*, Advances in Applied Mathematics **21** (1998), no. 2, 181–192. ↑18

[14] V. Rödl and A. Ruciński, *Threshold functions for Ramsey properties*, J. Amer. Math. Soc. **8** (1995), no. 4, 917–942. ↑13, 14

[15] J. Spencer, *Threshold functions for extension statements*, J. Combin. Theory Ser. A **53** (1990), no. 2, 286–305. ↑8

Departamento de Matemática, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil

E-mail address: mauricio@collares.org

Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, SP, Brazil

E-mail address: yoshi@ime.usp.br

IMPA, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, RJ, Brazil

E-mail address: rob@impa.br

Centro de Matemática, Computação e Cognição, Universidade Federal do ABC, Santo André, SP, Brazil

E-mail address: g.mota@ufabc.edu.br