Acyclic sets and colorings in digraphs under restrictions on degrees and cycle lengths

Ararat Harutyunyan* Colin McDiarmid† Gil Puig i Surroca*

August 16, 2024

Abstract

Given a digraph D, we denote by $\vec{\alpha}(D)$ the maximum size of an acyclic set of D (i.e. a set of vertices which induces a subdigraph with no directed cycles), and by $\vec{\chi}(D)$ the minimum number of acyclic sets into which $V(D)$ can be partitioned. In this paper, we study $\vec{\alpha}(D)$ and $\vec{\chi}(D)$ from various perspectives, including restrictions on degrees and cycle lengths. A main result is that, if D is a random r-regular digon-free simple digraph of order n, then $\vec{\alpha}(D) = \Theta(n \log r/r)$ with high probability. This extends a result of Spencer and Subramanian on the Erdős–Rényi random digraph model [29]. Along the way, we derive some related results and propose some conjectures. An example of this is an analog of the theorem of Bondy which bounds the chromatic number of a graph by the circumference of any strong orientation.

1 Introduction

A graph is *simple* if it has no loops or parallel edges. In this paper, our graphs will always be simple. Similarly, a digraph is *simple* if it has no loops or parallel arcs. We are mostly interested in digraphs that are simple. Sometimes, we shall require a digraph to be an *oriented graph*, meaning that directed cycles of length two (called *digons*) are also forbidden. Digraphs are usually denoted by $D = (V, A)$, where V is the set of vertices and A is the set of arcs. A subset S of vertices of a digraph D is called *acyclic* if the induced subdigraph on S contains no directed cycle. We denote by $\vec{\alpha}(D)$ the maximum size of an acyclic set in D. The *dichromatic number* $\vec{\chi}(D)$ of

*LAMSADE, Université Paris Dauphine - PSL, Paris, France
†Department of Statistics, University of Oxford, Oxford, UK

D is the smallest integer k such that $V(D)$ can be partitioned into k sets $V_1, \ldots, V_k$ where each V_i is acyclic. Note that, equivalently, the dichromatic number is the smallest integer k, such that the vertices of D can be colored with k colors so that there is no monochromatic directed cycle. It is easy to see that, for any undirected graph G and its bidirected digraph D obtained from G by replacing each edge by two oppositely oriented arcs, we have $\chi(G) = \vec{\chi}(D)$, where $\chi(G)$ is the *chromatic number* of G—the minimum number of colors needed to color the vertices of G such that no two adjacent vertices have the same color. The dichromatic number was first introduced by Neumann-Lara [26].

In recent years, there has been considerable attention devoted to this topic, and many results have demonstrated that this digraph invariant generalizes many results on the graph chromatic number (see, for example [4, 5, 7, 19, 20]). Some evidence of this surprising relationship includes the generalization of Gallai's classical theorem on list coloring to digraphs in [20], the extension of the important result of Erdős that sparse graphs can have large chromatic number to digraphs in [7], the derivation of an analog of a classical result due to Bollobás in [19], etc.

In the present paper, we study the existence of large acyclic sets in general digraphs as well as bounds on the dichromatic number. We use n for the number of vertices and m for the number of edges or arcs. The following two conjectures motivated our results. The first conjecture is due to Aharoni, Berger and Kfir.

Conjecture 1.1. [1] *If D is an oriented graph then*

$$\vec{\alpha}(D) \geq (1 + o(1)) \frac{n^2}{m} \log_2 \frac{m}{n}.$$

For an n-vertex tournament D we have $m = n(n-1)/2$, so for large tournaments the conjecture says that

$$\vec{\alpha}(D) \geq (2 + o(1)) \log_2 n \quad \text{as} \quad n \to \infty.$$

This is known up to a factor of 2; every n-vertex tournament contains a transitive subtournament of order $\lfloor \log_2 n \rfloor + 1$. Whether the factor 2 can indeed be gained is a major open problem which was already discussed by Erdős and Moser [17].

An equivalent way of stating Conjecture 1.1 is that if D has average outdegree $d^+ = \frac{m}{n}$, then

$$\vec{\alpha}(D) \geq (1 + o(1))\, n \log_2 d^+ / d^+.$$

We note that the meaningful interpretation of the term $o(1)$ is for d^+ large: if, for a digraph D, $\vec{\alpha}(D) \leq (1-\varepsilon)n\log_2 d^+/d^+$ (which is the case, for example, for the Paley tournament on seven vertices [17], or some other specific orders [27]), then the disjoint union of any number of copies of D will also satisfy this inequality.

On the other hand, if digons are permitted, then clearly one cannot hope to find an acyclic set of size greater than $\Theta(n/d^+)$; this directly follows from the fact that in a symmetric digraph an acyclic set corresponds to an independent set of the underlying graph. The extra logarithmic factor can only appear if one considers digraphs of digirth at least three. Interestingly, in the context of undirected graphs, a similar phenomenon occurs with independent sets.

Theorem 1.2. [2] *Every triangle-free graph G of average degree d has $\alpha(G) \geq \frac{n \ln d}{100d}$.*

Shearer [28] improved the constant $1/100$ to $1+o(1)$ (as $d \to \infty$). Later, Johansson [22, 25] addressed the coloring version of the problem.

Theorem 1.3. *There is a constant c such that, for any triangle-free graph G with maximum degree $\Delta \geq 2$, $\chi(G) \leq c\Delta/\ln\Delta$.*

Also in this case, the constant has been brought to $1+o(1)$ (as $\Delta \to \infty$) [24]. For oriented graphs, Erdős and Neumann-Lara [15] conjecture that the analog of Johansson's theorem (and an extension of Conjecture 1.1) should hold: one can color D with $O(\Delta/\log\Delta)$ colors such that each color class is acyclic.

The next conjecture that motivated our paper is the following one:

Conjecture 1.4. *For every tournament H, there is $\varepsilon_H > 0$ such that any H-free tournament T on n vertices satisfies $\vec{\alpha}(T) \geq n^{\varepsilon_H}$.*

(In this paper, by an *H-free digraph D* we will mean a digraph D which has no subdigraph isomorphic to H.) Alon, Pach and Solymosi [3] proved that the conjecture above is equivalent to the Erdős–Hajnal conjecture, one of the central questions in extremal graph theory.

Conjecture 1.5. [16] *For every graph H, there is $\varepsilon_H > 0$ such that any graph G on n vertices which does not have H as an induced subgraph satisfies $\max\{\alpha(G), \omega(G)\} \geq n^{\varepsilon_H}$.*

The structure of the paper is as follows. In Section 2, we prove a theorem on digraph coloring which is reminiscent of a well-known theorem of Bondy.

The results of this section give an upper bound on the dichromatic number when the largest cycle of a digraph has bounded length. In Section 3, we propose a strengthening of the Erdős–Hajnal conjecture, motivating it with several special cases. Lastly, Section 4 concerns random digraphs: we show that, up to a multiplicative constant, random r-regular oriented graphs satisfy the Aharoni–Berger–Kfir conjecture with high probability. Moreover, we also give an upper bound, which essentially differs from the lower bound by a factor of 4.

2 Digraph Coloring and Bondy's Theorem

By the *circumference* of a digraph D, we mean the maximum length of a directed cycle (when we talk about the circumference of D, we assume that D has a directed cycle). Bondy proved the following classical theorem.

Theorem 2.1. [9] *Let G be a graph and D an orientation of G which is strongly connected. Suppose that D has circumference s. Then $\chi(G) \leq s$.*

A *k-list-assignment* L to a digraph D is an assignment $L : V(D) \to \mathcal{P}(\mathbb{Z}^+)$ of sets of positive integers to the vertices of D such that $|L(v)| \geq k$ for every vertex v. D is *L-colorable* if $V(D)$ can be partitioned into acyclic sets $V_1, \ldots, V_s$ so that, for every vertex v, $v \in \cup_{i \in L(v)} V_i$. Let $\vec{\chi}_\ell(D)$ denote the *list dichromatic number* of D, that is, the minimum k such that D is L-colorable for any k-list-assignment to D. List colorings of digraphs were introduced in [20] and were later studied in [5].

Theorem 2.2. *Let D be a simple digraph in which all (directed) cycle lengths belong to a set K with $|K| = k$. Then $\vec{\chi}_\ell(D) \leq k + 1$; and indeed, given any corresponding lists of size $k + 1$, we can find a list acyclic coloring in polynomial time.*

We note that both Theorem 2.1 and Theorem 2.2 are tight; indeed, K_n and the bidirected complete digraph on n vertices, respectively, serve as examples. The following corollary of Theorem 2.2 is in the spirit of Bondy's result (Theorem 2.1).

Corollary 2.3. *Let D be an oriented graph with circumference $s \geq 3$. Then $\vec{\chi}_\ell(D) \leq s - 1$.*

To deduce the corollary, observe that in D there are at most $s-2$ distinct cycle lengths. The theorem follows from the next three easy lemmas. A digraph D is *k-degenerate* if in each subdigraph (including D itself) there is a vertex with indegree or outdegree at most k. A *k-degeneracy order* for D is a listing of the vertices as $v_1, \ldots, v_n$ such that for each $j = 2, \ldots, n$ vertex v_j has indegree or outdegree at most k in the subdigraph of D induced on $\{v_1, \ldots, v_j\}$.

Lemma 2.4. *Let K be a set of k positive integers, and let D be a simple digraph in which all (directed) cycle lengths are in K. Then D is k-degenerate.*

Proof. Let D' be a subdigraph of D. Let $P = (v_1, v_2, v_3, \ldots)$ be a longest (directed) path in D'. If v_1 has an in-neighbor w in D' then w must lie on P, and w can only be a vertex v_j for some $j \in K$. Thus v_1 has indegree at most k, completing the proof. □

Lemma 2.5. *Let the digraph D be k-degenerate. Then we can find a k-degeneracy order for D in polynomial time.*

Proof. Compute the indegree and outdegree of each vertex. Start with an empty list. Repeatedly, choose a vertex v with indegree or outdegree at most k, put v at the start of the current list, and delete v from D and update the remaining indegrees and outdegrees. □

Lemma 2.6. *Let the loop-free digraph D have a given k-degeneracy order $v_1, \ldots, v_n$. Then given any lists of size $k+1$, we can read off a list acyclic coloring f in polynomial time.*

Proof. For each $j = 1, \ldots, n$ proceed as follows. Let A_j be a smaller of the two sets $N^-(v_j) \cap \{v_1, \ldots, v_{j-1}\}$ and $N^+(v_j) \cap \{v_1, \ldots, v_{j-1}\}$, so $|A_j| \leq k$; and set $f(v_j)$ to be any color in $L(v_j) \setminus \{f(w) : w \in A_j\}$. No (directed) cycle C is monochromatic: for if j is the largest index such that v_j is in C, then C must contain a vertex $w \in A_j$ and $f(v_j) \neq f(w)$. □

This completes the proof of Theorem 2.2. We next show that for tournaments we can do much better than Theorem 2.2 and Corollary 2.3. Let us note first a simple but useful lemma.

Lemma 2.7. *For every digraph D, $\vec{\chi}(D)$ (resp. $\vec{\chi}_\ell(D)$) equals the maximum value of $\vec{\chi}(D')$ (resp. $\vec{\chi}_\ell(D')$) over the strongly connected components D' of D.*

Proof. Let t be the maximum value of $\vec{\chi}(D')$ over the strongly connected components D' of D. We may use colors from $[t]$ to properly color each strongly connected component. This gives a proper coloring of D, since each directed cycle is contained within one of the components. The same proof works for list colorings. □

For each $n \in \mathbb{N}$, let $t(n)$ be the maximum value of $\vec{\chi}_\ell(T)$ for T ranging over all n-vertex tournaments. Equivalently, $t(n)$ is the maximum value of $\vec{\chi}_\ell(D)$ for D ranging over all n-vertex oriented graphs. From [17] and [5] we have

$$(1/2 + o(1))\, n/\log_2 n \leq t(n) \leq (1 + o(1))\, n/\log_2 n \;\text{ as } n \to \infty \tag{1}$$

(the lower bound comes from random tournaments). Recall that a strongly connected tournament has a (directed) Hamilton cycle [10]; and thus for a tournament the circumference equals the maximum order of a strongly connected component. Hence, by the last lemma, if T is a tournament of circumference at most s then $\vec{\chi}_\ell(T) \leq t(s)$. By (1) we now have:

Theorem 2.8. *If T is a tournament of circumference at most s then*

$$\vec{\chi}_\ell(T) \;\leq (1 + o(1))\, s/\log_2 s \;\; \text{as } s \to \infty\,.$$

We now show that Theorem 2.8 is essentially best possible (up to a constant factor). Indeed, one may just take a random tournament on $n = s$ vertices. It is known that this tournament has no acyclic set of size at least $2\log_2 n + 2$ with positive probability (in fact, with probability tending to 1 as $n \to \infty$ [17]). The next theorem shows that we can also choose a tournament of arbitrary order.

Theorem 2.9. *For all positive integers s and n with $s \leq n$, there exists an n-vertex tournament T with circumference at most s such that $\vec{\alpha}(T) \leq \frac{4n}{s}\log_2(2s)$, and so $\vec{\chi}(T) \geq \frac{s}{4\log_2(2s)}$.*

Proof. Let T be the following tournament. Partition the vertices into $\lceil \frac{n}{s} \rceil$ sets, $A_1, \ldots, A_{\lceil n/s \rceil}$, each of size at most s. On the vertices of each A_i, orient the edges so that the largest acyclic set in A_i has size less than $2\log_2 s + 2$ (this is always possible since the random tournament on s vertices achieves this bound with positive probability). For any two vertices $v_i \in A_i$ and $v_j \in A_j$, where $i < j$, put an arc from v_i to v_j. Note that each cycle C in T is contained within one of the sets A_i, so C has length at most s. Moreover, since $\vec{\alpha}(A_i) < 2\log_2 s + 2$, it follows that

$$\vec{\alpha}(T) < \lceil n/s \rceil (2\log_2 s + 2) < \tfrac{4n}{s}\, \log_2(2s).$$

Finally, $\vec{\chi}(T) \geq n/\vec{\alpha}(T) > \frac{s}{4\log_2(2s)}$. □

We now consider the problem for general oriented graphs. We shall see in Corollary 2.13 that we can gain a factor of 2 over the non-list version of Corollary 2.3 above (since the digirth is at least 3).

Theorem 2.10. *Let D be a simple digraph with circumference s and digirth g. Then $\vec{\chi}(D) \le \lceil \frac{s}{g-1} \rceil$.*

We note that a slightly weaker version of this theorem is known: it was proved in [13] that $\vec{\chi}(D) \le \lceil \frac{s-1}{g-1} \rceil + 1$. Our bound is clearly at most this value and for many values of s and g it is an improvement of one. Additionally, our proof is somewhat shorter and less technical. To prove Theorem 2.10, we will use a result of Bessy and Thomassé.

First, we need some notation. Let D be a strong simple digraph on vertex set V. An enumeration $E = v_1, \ldots, v_n$ of V is *elementary equivalent* to another enumeration E' if one of the following holds: $E' = v_n, v_1, \ldots, v_{n-1}$, or $E' = v_2, v_1, v_3, \ldots, v_n$ and neither v_1v_2 nor v_2v_1 is an arc in D. Two enumerations E, E' of V are said to be *equivalent* if there is a sequence $E = E_1, \ldots, E_k = E'$ such that E_i and E_{i+1} are elementary equivalent for each i. The equivalence classes of this equivalence relation are called the *cyclic orders* of D. Given an enumeration $E = v_1, \ldots, v_n$ we say that an arc v_iv_j is a *forward arc* (with respect to E) if $i < j$, and a *backward arc* otherwise. A directed path in D is called a *forward path* if all its arcs are forward arcs. The *index* of directed cycle C, $i_E(C)$, is the number of backward arcs of C. Importantly, the index of a cycle is invariant in a given cyclic order $\mathcal{C}$. A cycle is *simple* if it has index one. A cyclic order $\mathcal{C}$ is *coherent* if for every enumeration E of $\mathcal{C}$ and every backward arc v_jv_i in E, there is a forward path from v_i to v_j. The next lemma is similar to Lemma 1 of [6].

Lemma 2.11. *Let D be a strong simple digraph and let C be a directed cycle of D. Then D has a coherent cyclic order such that C is simple.*

Proof. Amongst all cyclic orders of D such that C is simple, pick a cyclic order $\mathcal{C}$ which minimizes the sum of all cycle indices. Then the proof of Lemma 1 of [6] shows that $\mathcal{C}$ is coherent. □

Lemma 2.12. [6, Section 4] *Let $D = (V, A)$ be a strong simple digraph and C be a longest cycle in D, of length k. Suppose $\mathcal{C}$ is a coherent cyclic order of D such that C is simple in $\mathcal{C}$. Then there is an enumeration $E = v_1, \ldots, v_{i_1}, v_{i_1+1}, \ldots, v_{i_2}, v_{i_2+1}, \ldots, v_{i_k}$ of $\mathcal{C}$ such that $\{v_{i_j+1}, \ldots, v_{i_{j+1}}\}$ is a stable set for all $j = 0, \ldots, k-1$ (with $i_0 := 0$).*

Proof of Theorem 2.10. Note that it is sufficient to prove the result for strong digraphs. Indeed, if D has r strongly connected components $D_1, \ldots, D_r$, we color each D_i with at most $\lceil \frac{s}{g-1} \rceil$ colors. Since every directed cycle of D is contained entirely in a single D_i, this gives a proper coloring of D.

Thus, we may assume that D is strongly connected. Let C be a cycle of length s in D. By Lemma 2.11, D has a coherent cyclic order $\mathcal{C}$ such that C is simple. Now by Lemma 2.12, there is an enumeration $E = v_1, \dots, v_{i_1}, v_{i_1+1}, \dots, v_{i_2}, v_{i_2+1}, \dots, v_{i_s}$ of $\mathcal{C}$ such that $I_j := \{v_{i_{j-1}+1}, \dots, v_{i_j}\}$ is a stable set for each $j = 1, \dots, s$. We claim that $I_j \cup I_{j+1} \cup \ldots \cup I_{j'}$ is an acyclic set for each $1 \le j \le j' \le \min\{j+g-2, s\}$. Indeed, since I_j, $I_{j+1}, \dots, I_{j'}$ are all stable sets, any cycle on $I_j \cup \ldots \cup I_{j'}$ must use a backward arc $v_q v_p$, with $v_q \in I_{j+r}$ and $v_p \in I_{j+k}$, where $r > k \ge 0$. Now, since D is assumed to be of digirth g, there is no forward path from v_p to v_q. This contradicts the fact that $\mathcal{C}$ is coherent. Thus, we can color the digraph induced by $I_j \cup \ldots \cup I_{j'}$ with one color. Coloring consecutive $(g-1)$-tuples $I_j, \dots, I_{j+g-2}$ (and perhaps one shorter tuple) with a single color gives a coloring with at most $\left\lceil \frac{s}{g-1} \right\rceil$ colors. □

Corollary 2.13. *Let D be an oriented graph with circumference s. Then $\vec{\chi}(D) \le \lceil \frac{s}{2} \rceil$.*

The corollary is clearly tight for $s = 3$ or $s = 4$, but we think that it is not optimal for large s (see Theorem 2.8).

Conjecture 2.14. *If D is an oriented graph with no directed cycle of length greater than s, then $\vec{\chi}(D) = O(s/\log s)$ as $s \to \infty$.*

It was proved by Neumann-Lara [26] that if D is an oriented graph only containing odd cycles or only containing even cycles then $\vec{\chi}(D) \le 2$. On the other hand, Chen, Ma and Zang showed the following.

Theorem 2.15. [11] *Let k and r be integers with $k \ge 2$ and $0 \le r \le k-1$. If a simple digraph D contains no directed cycle of length r modulo k, then $\vec{\chi}(D) \le k$.*

We show that this theorem cannot be strengthened to a list coloring version in the case 0 modulo k.

Proposition 2.16. *For all positive integers $k \ge 3$ and t, there is an oriented graph D such that all cycles of D have length 0 modulo k, and $\vec{\chi}_\ell(D) > t$.*

Proof. Let $k \ge 3$ and $t \ge 1$. Set $c = k(t-1)+1$. Consider the following digraph D. Take the directed cycle $\vec{C_k}$ with vertices $v_1, \dots, v_k$ and blow-up each vertex v_i into a set B_i of independent vertices of size $\binom{c}{t}$. Now, put a complete bipartite graph between every pair (B_i, B_{i+1}) with edges going from B_i to B_{i+1} (here, B_{k+1} is B_1). Denote this digraph by D. Clearly, all cycle lengths of D are multiples of k. Let L be an assignment of lists of size t for D such that on each set B_i we see each t-subset of $\{1, \ldots, c\}$.

Now suppose that D is L-colorable. In any such coloring, at most $t-1$ of the c colors are absent on any given B_i. Indeed, suppose that on some B_i we do not see at least t colors, say, the colors $\{1, \ldots, t\}$. This is clearly not possible since then the vertex in B_i with the list $\{1, \ldots, t\}$ was not colored. Thus, in total, there are at most $k(t-1)$ colors missing from all the B_i. Thus, there is some color j that appears on all the B_i. But this is a contradiction since we obtain a directed cycle of color j. Thus, D is not L-colorable and $\vec{\chi}_\ell(D) > t$. □

3 Strong Erdős–Hajnal conjecture

In this section, we propose the following strengthening of the Erdős–Hajnal conjecture (see Conjectures 1.4 and 1.5).

Conjecture 3.1. *For every tournament T there exists some $\varepsilon_T > 0$ such that for any n-vertex T-free simple digraph D, $\vec{\alpha}(D) > n^{\varepsilon_T}$.*

We say that a tournament T has the *Strong Erdős–Hajnal property* if the above conjecture is satisfied for T. In this case, we let ζ_T be the limit, when $k \to \infty$, of the supremum of all ε_T such that $\vec{\alpha}(D) \geq n^{\varepsilon_T}$ for each T-free simple digraph D of order $n \geq k$.

Theorem 3.2. *If T is a tournament with $t \geq 3$ vertices which has the Strong Erdős–Hajnal property, then $\zeta_T \leq 2/t$.*

The proof if probabilistic. We say that events $A_1, A_2, \ldots$ hold *with high probability (whp)* if $\mathbb{P}(A_n) \to 1$ as $n \to \infty$. We denote by $D_{n,p}$ the random digraph on n vertices where an edge is placed between two vertices independently with probability $2p$, and each present edge is oriented independently in one of the two possible ways with probability $1/2$.

Proof. Consider the random digraph $D_{n,p}$ with $p = \frac{1}{2}n^{-2/t}$. Subramanian and Spencer [29] showed that whp $\vec{\alpha}(D_{n,p}) \leq \frac{2\ln(np)}{\ln(1-p)^{-1}}(1 + o(1))$. Thus, $\vec{\alpha}(D_{n,p}) \leq 4n^{2/t} \ln n$ whp. Let X_T count the number of copies of T in $D_{n,p}$. Clearly,

$$\mathbb{E}[X_T] \leq \binom{n}{t} t!\, p^{\binom{t}{2}} \leq n^t n^{-t+1}/2^{t(t-1)/2} \leq n/8.$$

Now, by Markov's inequality, $\mathbb{P}(X_T \geq n/4) \leq 1/2$. Therefore, there is a digraph D with number of copies of T at most $n/4$ and with $\vec{\alpha}(D) \leq$

$4n^{2/t}\ln n$. We may delete a vertex from each copy of T in D to obtain a T-free digraph D' with at least $3n/4$ vertices. But $\vec{\alpha}(D') \le \vec{\alpha}(D) \le 4n^{2/t}\ln n$, and the theorem follows. □

The following special case is worthy of attention.

Conjecture 3.3. *The directed triangle has the Strong Erdős–Hajnal property.*

We do not have any means of approaching the above special case, though an inductive argument would show that every n-vertex simple digraph without directed triangles has an acyclic set of size $\omega(\log n)$. Interestingly, when one forbids the transitive tournament on three vertices, the proof is straightforward.

Proposition 3.4. *Let T be the transitive tournament on three vertices. Then, T has the Strong Erdős–Hajnal property. Moreover, for any n-vertex T-free simple digraph D with maximum (total) degree Δ, $\vec{\chi}(D) = O(\Delta/\log\Delta)$.*

Proof. Let D be a T-free digraph. Take an ordering of vertices, and form a graph G_f from the digraph D by only keeping the forward edges. We remark that since D does not contain T, G_f is triangle-free. Note that an independent set in G_f is an acyclic set in D. Now, since G_f is triangle-free, by Ramsey Theory (the fact that $R(3,t) \sim t^2/\log t$, see [23]), it follows that $\alpha(G_f) = \Omega(\sqrt{n\log n})$. Thus, there is an acyclic set of size $\Omega(\sqrt{n\log n})$ in D. The second part of the claim follows from Johansson's theorem on colorings [22, 25] applied to the triangle-free graph G_f. □

4 Random regular digraphs

A (multi)digraph is *r-regular* if every vertex has exactly r in-arcs and r out-arcs. For random r-regular n-vertex oriented graphs, we prove that the size of the largest acyclic set is $\Theta(n\ln r/r)$ whp; see Theorems 4.3 and 4.5. This matches the behavior of the binomial random oriented graph $D_{n,p}$ with $p = r/n$, where vertices have expected indegree and expected outdegree $r(1-1/n)$. Indeed, by a result of Spencer and Subramanian (see Corollary 1.1 in [29]), as $r \to \infty$,

$$\vec{\alpha}(D_{n,r/n}) = (1+o(1))\,\tfrac{2n\ln r}{r} \quad \text{whp.}$$

Random regular graphs can be constructed by means of the configuration model [8, Section 2.4]. Let n, and r be positive integers such that nr

is even. For each vertex $i \in [n]$ we create an r-set $G[i]$, where the sets $G[1], \dots, G[n]$ are pairwise disjoint. We then put a uniformly random pairing (a *configuration*) between all the elements of $\cup_{i=1}^{n} G[i]$. Let $G^*(n, r)$, or simply G^*, be the r-regular multigraph obtained on vertex set $[n]$, where there is an edge ij for each element of $G[i]$ that is paired with an element of $G[j]$. The probability that G^* is simple is bounded away from 0 by a constant, and every r-regular n-vertex (simple) graph has the same probability of appearing as G^* [8]. Let us denote by $\mathscr{G}^*(n, r)$ the set of all r-regular n-vertex multigraphs with vertex set $[n]$, by $\mathscr{G}(n, r)$ the subset of (simple) r-regular n-vertex graphs, and by $P_{n,r}$ the probability measure on $\mathscr{G}^*(n, r)$ associated with G^*.

In the directed setting, we consider the following analog of the configuration model. For each vertex $i \in [n]$ we create two r-sets: $D^+[i]$ and $D^-[i]$, where $D^+[1], D^-[1], \dots, D^+[n], D^-[n]$ are pairwise disjoint. We denote by D^+ and D^- the unions $\cup_{i=1}^{n} D^+[i]$ and $\cup_{i=1}^{n} D^-[i]$, respectively. Next, we put a pairing between the elements of D^+ and D^- (a *directed configuration*), uniformly at random. Let $D^*(n, r)$, or simply D^*, be the r-regular multidigraph obtained on vertex set $[n]$, where there is an arc ij for each element $u \in D^+[i]$ which is paired with some $v \in D^-[j]$. Here, the different elements of a fixed r-set $D^+[i]$ or $D^-[i]$ play an undistinguishable role, so permuting them does not affect the resulting multidigraph. More precisely, these permutations generate a group that acts on the set of all directed configurations, and each orbit corresponds to (an r-regular n-vertex) multidigraph D. Thus, D arises from exactly $r!^{2n} \Pi_{a \in A(D)} \frac{1}{\mathrm{mult}(a)!}$ directed configurations, where $A(D)$ is the set of arcs of D and $\mathrm{mult}(a)$ is the multiplicity of the arc a. In particular, every simple digraph has the same probability of appearing as D^*.

If we forget the orientations of the arcs of D^*, we then obtain a $2r$-regular n-vertex multigraph that we call $\operatorname{forg} D^*$. We denote by $Q_{n,r}$ the probability measure on $\mathscr{G}^*(n, 2r)$ associated with $\operatorname{forg} D^*$.

Remark 4.1. We can establish a link between $\operatorname{forg} D^*(n, r)$ and $G^*(n, 2r)$ through the enumeration of Eulerian orientations. When considering orientations of multigraphs, we have to clarify whether the edges are labelled or not. Unless specified, we will make no distinction between multiple copies of the same edge. An *Eulerian orientation* of a multigraph G is an orientation D of G such that $\mathrm{indeg}_D(v) = \mathrm{outdeg}_D(v)$ for every vertex v. Let $E^*_{n,r}(G)$ be the number of labelled (i.e. edges are labelled) Eulerian orientations of $G \in \mathscr{G}^*(n, r)$, with the convention that loops can be oriented in two ways.

Then

$$\frac{E^*_{n,2r}}{\mathbb{E}[E^*_{n,2r}]} = \frac{Q_{n,r}}{P_{n,2r}}, \tag{2}$$

where the expectation is taken on $G^*(n, 2r)$.

Proof of (2). Note that in the configuration model with parameters n, $2r$ there are $c_{n,2r} := \frac{(2nr)!}{(nr)!2^{nr}}$ possible pairings, and in the directed version of the configuration model with parameters n, r there are $d_{n,r} := (nr)!$ possible pairings. Thus, we see that, for every $2r$-regular n-vertex multigraph G,

$$P_{n,2r}(G) = \frac{(2r)!^n}{2^{\ell(G)}c_{n,2r}} \prod_{e\in E(G)} \frac{1}{\text{mult}(e)!},$$

where $\ell(G)$ is the number of loops of G and $E(G)$ is its set of edges. Similarly,

$$Q_{n,r}(G) = \frac{r!^{2n}}{d_{n,r}} \sum_{D\in \text{EO}(G)} \prod_{a\in A(D)} \frac{1}{\text{mult}(a)!},$$

where $\text{EO}(G)$ is the set of Eulerian orientations of G and $A(D)$ is the set of arcs of D. On the other hand,

$$E^*_{n,2r}(G) = 2^{\ell(G)} \sum_{D\in \text{EO}(G)} \prod_{e\in E'(G)} \binom{\text{mult}_G(e)}{\text{mult}_D(e^+)},$$

where $E'(G)$ is the set of non-loop edges of G and, for each $e \in E'(G)$, e^+ is a fixed orientation of e (notice that the previous expression is independent of this choice). The claim follows from the fact that

$$\mathbb{E}[E^*_{n,2r}] = \frac{2^{nr}\binom{2r}{r}^n}{\binom{2nr}{nr}}$$

(see the proof of Theorem 3.47 in [14]: there, $E^*_{n,2r}$ is defined in an alternative way). □

Parallel to the undirected case, D^* is an oriented graph with probability at least a positive constant. We could not find this result in the literature, so we prove it here for completeness. (In contrast, the probability that D^* is simple has been studied; see for instance [12, Proposition 4.2].)

Lemma 4.2. *For every positive integer i, let $\mu_i = \frac{(2r-1)^i+1}{2i}$. Then,*

$$\lim_{n\to\infty} \mathbb{P}(D^*(n, r) \text{ is an oriented (simple) graph}) = e^{-\mu_1-\mu_2}.$$

Proof. Given integers k, j, let us denote $k(k-1)\dots(k-j+1)$ by $(k)_j$. For every positive integer i, let $X_{i,n}$ be the random variable counting the number of cycles of length i in $G^*(n,2r)$. It is shown in [14, Lemma 3.51] that

$$\lim_{n\to\infty}\frac{\mathbb{E}[E^*_{n,2r}(X_{1,n})_{j_1}\dots(X_{k,n})_{j_k}]}{\mathbb{E}[E^*_{n,2r}]}=\prod_{i=1}^{k}\mu_i^{j_i}$$

for any set of non-negative integers $j_1,\dots,j_k$, where the expectation is taken on $G^*(n,2r)$. By Remark 4.1,

$$\frac{\mathbb{E}[E^*_{n,2r}(X_{1,n})_{j_1}\dots(X_{k,n})_{j_k}]}{\mathbb{E}[E^*_{n,2r}]}=\mathbb{E}\left[\tfrac{Q_{n,r}}{P_{n,2r}}(X_{1,n})_{j_1}\dots(X_{k,n})_{j_k}\right]=$$

$$\sum_{G\in\mathscr{G}^*(n,2r)}\left(\tfrac{Q_{n,r}}{P_{n,2r}}(X_{1,n})_{j_1}\dots(X_{k,n})_{j_k}\right)(G)P_{n,2r}(G)=\mathbb{E}_{Q_{n,r}}[(X_{1,n})_{j_1}\dots(X_{k,n})_{j_k}],$$

where $\mathbb{E}_{Q_{n,r}}$ is the expectation taken on $\operatorname{forg} D^*(n,r)$. Therefore, by the method of moments (see [21, Theorem 6.10]), under the measures $Q_{n,r}$, $X_{i,n}\xrightarrow{d}\tilde{X}_i$ as $n\to\infty$, jointly for all i, where $\tilde{X}_i\in\mathrm{Po}(\mu_i)$ are independent Poisson random variables (see also [21, Lemma 9.17]). Hence,

$$\begin{aligned}&\lim_{n\to\infty}\mathbb{P}(D^*(n,r)\text{ is an oriented graph})\\&=\lim_{n\to\infty}\mathbb{P}(\operatorname{forg} D^*(n,r)\in\mathscr{G}(n,2r))=\lim_{n\to\infty}Q_{n,r}(\mathscr{G}(n,2r))\\&=\lim_{n\to\infty}Q_{n,r}(X_{1,n}=X_{2,n}=0)=e^{-\mu_1-\mu_2}.\end{aligned}$$

□

We say that events $A_1,A_2,\dots$ hold *with very high probability (wvhp)* if $\mathbb{P}(A_n)=1-e^{-\Omega(n)}$ as $n\to\infty$.

Theorem 4.3. *Let r be a positive integer and let D be a random digraph, chosen uniformly among all r-regular n-vertex oriented graphs with labelled vertices. Then, $\vec{\alpha}(D)\le\frac{2\ln r+4}{r}n$ wvhp.*

Proof. Let D^* be the random r-regular n-vertex multidigraph obtained with the directed version of the configuration model. By Lemma 4.2, D^* is an oriented graph (i.e., has no parallel arcs, loops, or digons) with probability bounded away from 0. Moreover, every r-regular n-vertex oriented graph has the same probability of appearing as D^*. Thus, it suffices to prove that $\mathbb{P}(\vec{\alpha}(D^*)\ge\frac{2\ln r+4}{r}n)=e^{-\Omega(n)}$ as $n\to\infty$.

Let k be a positive integer and $0 < \beta < 1$ a real number, for now both of them unspecified. Let ℓ be the integer divisible by k in the interval $[\beta n, \beta n + k)$. Suppose that some $A \subseteq V(D^*)$ of size $|A| = \ell$ is acyclic. Then, there is an ordering of A $\sigma : \{1, \ldots, \ell\} \to A$ such that each arc in $D^*[A]$ is of the form $\sigma(i) \to \sigma(j)$ for some $i < j$. This implies that A can be partitioned into k subsets $A_1, \ldots, A_k$ in a way that

(a) $|A_i| = \frac{\ell}{k}$;

(b) for each pair $1 \le i < j \le k$, there is no arc from any element of A_j to any element of A_i.

If $\vec{\alpha}(D^*) \ge \ell$, then for one of the ℓ-subsets A of $V(D^*)$ the above condition must hold. The number of ways to choose A is $\binom{n}{\ell}$ and the number of ways to partition such a set A into k parts $A_1, \ldots, A_k$ is easily at most k^ℓ. Thus, in total there are at most $\binom{n}{\ell} k^\ell \le (\frac{ekn}{\ell})^\ell$ choices.

Now, let us assume that we have fixed a set $A \subseteq V(D^*)$ with $|A| = \ell$, and a partition $A_1, \ldots, A_k$ of A with $|A_i| = \frac{\ell}{k}$. Without loss of generality, we may assume that $A = \{1, \ldots, \ell\}$ and that $A_1 = \{1, \ldots, \frac{\ell}{k}\}, \ldots, A_k = \{(k-1)\frac{\ell}{k} + 1, \ldots, \ell\}$. We would like to compute the probability that there is no *backward* arc, i.e., no arc from A_j to A_i for any $j > i$. Let E_1 be the event that there is no arc from A_k to any of the A_i, for all $i < k$. Clearly,

$$\mathbb{P}(E_1) = \prod_{j=0}^{r\frac{\ell}{k}-1} \frac{rn - r(k-1)\frac{\ell}{k} - j}{rn - j} \le \left(1 - \frac{(k-1)\ell}{kn}\right)^{\frac{r\ell}{k}}.$$

In general, let E_i be the event that no vertex of A_{k-i+1} has a backward arc. Then

$$\begin{aligned} \mathbb{P}(E_i | E_1, \ldots, E_{i-1}) &= \prod_{j=0}^{r\frac{\ell}{k}-1} \frac{rn - r(i-1)\frac{\ell}{k} - r(k-i)\frac{\ell}{k} - j}{rn - r(i-1)\frac{\ell}{k} - j} \\ &\le \left(1 - \frac{(k-i)\frac{\ell}{k}}{n - (i-1)\frac{\ell}{k}}\right)^{\frac{r\ell}{k}} \le \left(1 - \frac{(k-i)\ell}{kn}\right)^{\frac{r\ell}{k}} \\ &\le \exp\left(-\frac{r(k-i)\ell^2}{k^2 n}\right). \end{aligned}$$

Thus the probability that A with the partition $A_1, \ldots, A_k$ satisfies (b) is at most

$$\exp\left(-\sum_{i=1}^{k} \frac{r(k-i)\ell^2}{k^2 n}\right) = \exp\left(-\frac{r(1-\frac{1}{k})\ell^2}{2n}\right).$$

Hence, the probability that there is an acyclic set of size ℓ is at most

$$\left(\frac{ekn}{\ell}\right)^{\ell} \exp\left(-\frac{r(1-\frac{1}{k})\ell^2}{2n}\right) \leq \exp\left\{\ell\left(1+\ln k - \ln\beta - \frac{\beta r}{2}\left(1-\tfrac{1}{k}\right)\right)\right\},$$

where we used the facts that $\frac{ekn}{\ell} \leq \frac{ek}{\beta}$ and $r(1-\frac{1}{k})\ell^2 \geq r(1-\frac{1}{k})\beta n\ell$.

Now, we fix $\beta = \frac{2\ln(3r/4)+4}{r}$ and $k = \lceil \beta r/2 \rceil$. Clearly, we can assume that $r \geq 2$. This implies that $\beta > \frac{4}{r}$, so we have the bound $k < \frac{\beta r}{2} + 1 < \frac{3\beta r}{4}$. Denote by $c_r := 1 + \ln k - \ln\beta - \frac{\beta r}{2}\left(1-\frac{1}{k}\right)$ and note that $c_r < 1 + \ln\frac{3r}{4} - \frac{\beta r}{2} + 1 = 0$. Moreover, note that c_r is independent of n. Thus, for n large enough,

$$\mathbb{P}(\vec{\alpha}(D^*) \geq \tfrac{2\ln r+4}{r}n) \leq \mathbb{P}(\vec{\alpha}(D^*) \geq \beta n + k) \leq \mathbb{P}(\vec{\alpha}(D^*) \geq \ell) \leq e^{c_r\ell} \leq e^{c_r\beta n},$$

which completes the proof. □

Remark 4.4. Unfortunately, the bound of Theorem 4.3 is meaningless for small r. It makes sense to push the analysis above to try to find a constant $c < 1$ such that wvhp $\vec{\alpha}(D) \leq cn$. Note that we cannot expect that to work for $r = 1$. Indeed, it is well-known that the number of cycles of the uniform random permutation $\pi \in S_n$ is concentrated around its mean $1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}$ [18, Example III.4]. It follows that, when $r = 1$, $\vec{\alpha}(D^*) = n - (1 + o(1))\ln n$ with probability tending to 1 as $n \to \infty$. Below we show that, for $r \geq 2$, one can take $c = 99/100$. In any case, the bounds from Theorems 4.3 and 4.5 are still far from each other, and bringing them closer remains an open problem.

In the proof of Theorem 4.3, we now bound the probability of the event E_i as follows:

$$\mathbb{P}(E_i) = \prod_{j=0}^{r\frac{\ell}{k}-1} \frac{rn + r\frac{\ell}{k} - r\ell - j}{rn - r(i-1)\frac{\ell}{k} - j} \leq \left(\frac{n + \frac{\ell}{k} - \ell}{n - (i-1)\frac{\ell}{k}}\right)^{\frac{r\ell}{k}} \leq \left(\frac{\frac{k}{\beta} + 1 - k}{\frac{k}{\beta} + 1 - i}\right)^{\frac{r\ell}{k}},$$

and so the product $\prod_{i=1}^{k} \mathbb{P}(E_i)$ is upper-bounded by

$$\left(\left(\frac{k}{\beta} + 1 - k\right)^{k} \prod_{i=1}^{k} \frac{1}{\frac{k}{\beta} + 1 - i}\right)^{\frac{r\ell}{k}}.$$

We have that

$$\sum_{i=1}^{k} \ln\left(\tfrac{k}{\beta} + 1 - i\right) \geq \int_{1}^{k} \ln\left(\tfrac{k}{\beta} + 1 - x\right) dx$$

$$= \left[-x - \left(\tfrac{k}{\beta} + 1 - x \right) \ln \left(\tfrac{k}{\beta} + 1 - x \right) \right]_1^k$$
$$= 1 - k - \left(\tfrac{k}{\beta} + 1 - k \right) \ln \left(\tfrac{k}{\beta} + 1 - k \right) + \tfrac{k}{\beta} \ln \tfrac{k}{\beta},$$

so $\mathbb{P}(\vec{\alpha}(D^*) \geq \ell)$ is at most

$$\exp \left\{ \ell \left(1 + r \left(1 - \tfrac{1}{k}\right) + \left(1 - \tfrac{r}{\beta}\right) \ln \tfrac{k}{\beta} + r \left(\tfrac{1}{\beta} + \tfrac{1}{k} \right) \ln \left(\tfrac{k}{\beta} + 1 - k \right) \right) \right\}.$$

Therefore, it is enough to ask that

$$1 + r \left(1 - \tfrac{1}{k}\right) + \left(1 - \tfrac{r}{\beta}\right) \ln \tfrac{k}{\beta} + r \left(\tfrac{1}{\beta} + \tfrac{1}{k} \right) \ln \left(\tfrac{k}{\beta} + 1 - k \right) < 0,$$

which, for $r \geq 2$, is satisfied by $k = 100$ and $\beta = 99/100$.

There is a natural (fast) greedy algorithm which yields an acyclic set in a loop-free digraph D. (The U is for 'unused'.)

input: a digraph D and a total order $\preceq$ on $V(D)$
set $A = U = \emptyset$ and $W = V(D)$
while $W \neq \emptyset$
  let w be the first (smallest) vertex in W and reveal $N^+(w)$
  move w into A and move $N^+(w) \cap (W \backslash \{w\})$ into U
output: A

Observe that the output set A is acyclic if we ignore any loops, since all arcs point 'backwards' to vertices added earlier. We call A the *greedy acyclic set* of D with respect to $\preceq$, and denote its size by $\vec{\alpha}(D, \preceq)$. (If we wanted to find a large acyclic set in a general loop-free digraph, not necessarily random, we would pick $\preceq$ uniformly at random.)

Theorem 4.5. *Let r be a positive integer and let D be a random digraph, chosen uniformly among all r-regular oriented graphs on $[n]$. Then $\vec{\alpha}(D, \leq) \geq \frac{1}{5}\, n \log_2(r+1)/r$ wvhp. Moreover, for any $\varepsilon > 0$ there exists an r_ε such that, if $r \geq r_\varepsilon$, then $\vec{\alpha}(D, \leq) \geq (\frac{1}{2} - \varepsilon)\, n \ln(r+1)/r$ wvhp.*

Note that our lower bounds on $\vec{\alpha}(D, \leq)$ match the upper bound in Theorem 4.3 on $\vec{\alpha}(D)$ up to a constant factor. From now on, log means $\log_2$. Let $0 < \alpha < 1$ be fixed. To prove Theorem 4.5 we shall use a truncated version of the above greedy algorithm, where we replace the condition 'while $W \neq \emptyset$' by 'while $W \neq \emptyset$ and $|A| \leq \alpha n$'. Later we shall set $\alpha = \frac{1}{5}$.

We shall prove that, when $\alpha = \frac{1}{5}$, for D^* the random n-vertex r-regular multidigraph obtained with the directed version of configuration model (as in the proof of Theorem 4.3), the algorithm yields a set A with

$$|A| \geq \tfrac{1}{5}\, n \log(r+1)/r \quad \text{wvhp.} \tag{3}$$

But by Lemma 4.2 the probability that D^* is an oriented graph is bounded away from 0, and all oriented graphs have the same probability of appearing as D^*, so the theorem will follow. (The proof below shows that, essentially, the bounds of Theorem 4.5 hold also for $\vec{\alpha}(D^*)$, since the expected number of loops in D^* is r.)

We use one preliminary lemma. Let $n \geq 1$ and $0 \leq a, b \leq n$. Let U and V be disjoint n-sets, and let G be the complete bipartite graph with parts U and V. Let M be a random perfect matching in G chosen uniformly from the $n!$ perfect matchings in G. Let $A \subseteq U$ with $|A| = a$ and $B \subseteq V$ with $|B| = b$. Let the random variable $X(n, A, B)$ (or less precisely $X(n, a, b)$) be the number of edges in M between A and B. Observe that $\mathbb{E}[X] = ab/n$.

If X, Y are two random variables, we say that X is *stochastically dominated* by Y if $\mathbb{P}(X \geq t) \leq \mathbb{P}(Y \geq t)$ for every t, and we denote it by $X \leq_s Y$.

Lemma 4.6. *Let $n, n' \geq 1$, let $a, b \leq n$, let $a', b' \leq n'$; and suppose that $n' \leq n$, $a' \geq a$ and $b' \geq b$. Let $Y = X(n, a, b)$ and $Y' = X(n', a', b')$. Then $Y \leq_s Y'$.*

Proof of Lemma 4.6. It suffices to establish the following three simple claims.

$$\text{If } \ a+1 \leq n \ \text{ then } \ X(n, a, b) \leq_s X(n, a+1, b). \tag{4}$$

$$\text{If } \ b+1 \leq n \ \text{ then } \ X(n, a, b) \leq_s X(n, a, b+1). \tag{5}$$

$$X(n+1, a, b) \leq_s X(n, a, b). \tag{6}$$

To prove (4), let $a + 1 \leq n$; let $A \subseteq A' \subseteq U$ with $|A| = a, |A'| = a + 1$; and let $B \subseteq V$ with $|B| = b$. Then always $X(n, A, B) \leq X(n, A', B)$ so (4) holds. We may prove (5) in the same way.

It remains to prove (6). Let U and V be disjoint $(n+1)$-sets, let $A \subseteq U$ with $|A| = a \leq n$ and let $B \subseteq V$ with $|B| = b \leq n$. Let $u \in U \backslash A$, and let v

be the random vertex in V paired with u in the random matching M. Then for each relevant integer i

$$\begin{aligned} & \mathbb{P}(X(n+1,a,b) \geq i) \\ = \ & \mathbb{P}((X(n+1,A,B) \geq i) \wedge (v \in B)) + \mathbb{P}((X(n+1,A,B) \geq i) \wedge (v \notin B)) \\ = \ & \tfrac{b}{n+1}\,\mathbb{P}(X(n,a,b-1) \geq i) + \tfrac{n+1-b}{n+1}\,\mathbb{P}(X(n,a,b) \geq i) \\ \leq \ & \tfrac{b}{n+1}\,\mathbb{P}(X(n,a,b) \geq i) + \tfrac{n+1-b}{n+1}\,\mathbb{P}(X(n,a,b) \geq i) \quad \text{by (5)} \\ = \ & \mathbb{P}(X(n,a,b) \geq i). \end{aligned}$$

Now (6) follows, and the proof if complete. □

Proof of (3), and thus of Theorem 4.5. Consider part way through a run of the algorithm, when we are about to reveal $N^+(w)$. At this time, we know the sets A, U and W of vertices in G; we know all the r arcs out of each vertex in A (that is, we know the edges in the random matching M which are incident to the points corresponding to an out-incidence of a vertex in A), and all these arcs go from A to $A \cup U$. The remaining $r(n-|A|)$ edges in M form a uniformly random perfect matching M' in the bipartite graph over the remaining points. Let the *cost* Y when revealing $N^+(w)$ be the consequent reduction in the size of W. Then

$$Y \leq_s 1 + X(r(n-|A|), r, r|W|) \leq_s 1 + X(r(n-\lfloor \alpha n \rfloor), r, r|W|), \tag{7}$$

where we use Lemma 4.6 for the second inequality $\leq_s$. Note that $1 \leq Y \leq r+1$, and $\mathbb{E}[Y] \leq 1 + \frac{r|W|}{n(1-\alpha)}$.

Let $b \geq 1 + r^{-1}$ be a constant. We assume that n is large enough. Divide the runs of the algorithm into stages $s = 1, 2, \ldots$, where stage s is when $nb^{-s+1} \geq |W| > nb^{-s}$. Consider stage s, where $1 \leq s \leq \log_b(r+1)$. Let $Y_1, Y_2, \ldots$ be the costs of the first, second,... vertices added to A in this stage, where we set $Y_i = 0$ if the algorithm stops before adding the ith vertex or $|W|$ has decreased to at most nb^{-s}. If we add an ith vertex, then at this time $|W| \leq nb^{-s+1}$ and $|A| \leq \lfloor \alpha n \rfloor$, so by (7) (conditional on all history to date)

$$Y_i \leq_s 1 + X(r(n-\lfloor \alpha n \rfloor), r, r|W|) \leq_s Z$$

where $Z \sim 1 + X(r(n-\lfloor \alpha n \rfloor), r, rnb^{-s+1})$. Further, let $Z_1, \ldots, Z_k$ be independent copies of Z: then jointly $(Y_1, \ldots, Y_k) \leq_s (Z_1, \ldots, Z_k)$, and so $\sum_{i=1}^k Y_i \leq_s \sum_{i=1}^k Z_i$. Recall that $1 \leq Z \leq r+1$ and $\mathbb{E}[Z] = 1 + \frac{rnb^{-s+1}}{n-\lfloor \alpha n \rfloor} \leq 1 + \beta r b^{-s+1}$, where $\beta = (1-\alpha)^{-1}$. But $rb^{-s+1} \geq \frac{br}{r+1} \geq 1$, so $\mathbb{E}[Z] \leq (1+\beta) r b^{-s+1}$.

Let $\gamma < (\beta+1)^{-1}$ and $k = \frac{\gamma n(b-1)}{br}$, and note that $(k+1)\,\mathbb{E}[Z] \leq (\beta + 1)\gamma n(b^{-s+1} - b^{-s}) + (1+\beta)rb^{-s+1}$. Thus $\sum_{i=1}^{\lceil k \rceil} Z_i \leq n(b^{-s+1} - b^{-s}) - (r+1)$ wvhp, by a standard inequality, see for example [25, Section 10.1]. Hence, in this stage wvhp either we add at least $\lceil k \rceil$ vertices to A, or by the end of the stage we have $|A| \geq \alpha n$. This holds for each stage $s = 1, \ldots, \log_b(r+1)$. Hence after these stages, wvhp either $|A| \geq \log_b(r+1)k = \frac{\gamma(b-1)n \log_b(r+1)}{br}$ or $|A| \geq \alpha n$.

Finally set $b = 2$, $\alpha = \frac{1}{5}$ (so $\beta = \frac{5}{4}$), and $\gamma = \frac{2}{5}$. Then

$$\min\{\tfrac{\gamma(b-1)\log_b(r+1)}{br}, \alpha\} = \min\{\tfrac{\log(r+1)}{5r}, \tfrac{1}{5}\} = \tfrac{\log(r+1)}{5r}.$$

Thus after the stages above we have $|A| \geq \frac{\log(r+1)}{5r} n$ wvhp, and we have proved (3) as required. Alternatively, if $b = 1 + r^{-1}$ and α, γ are chosen arbitrarily close to 0 and $\frac{1}{2}$, respectively, and assuming that r is large enough, then

$$\min\{\tfrac{\gamma(b-1)\log_b(r+1)}{br}, \alpha\} = \tfrac{\gamma(b-1)\log_b(r+1)}{br} \geq (\tfrac{1}{2} - \varepsilon)\tfrac{\ln(r+1)}{r}$$

for any given $\varepsilon \in \mathbb{R}^+$. □

We know that for every r-regular simple digraph D, $\vec{\chi}(D) \leq r+1$ and so $\vec{\alpha}(D) \geq n/(r+1)$ (see Lemmas 2.5 and 2.6). The lower bound here on $\vec{\alpha}(D)$ is better than that in Theorem 4.5 for small r.

Finally, note that both Theorem 4.3 and Theorem 4.5 hold also if D is chosen uniformly at random among all r-regular n-vertex simple digraphs (i.e. allowing digons). Indeed, we may use essentially the same proofs: in the first paragraph of the proof of Theorem 4.3 we can just replace 'oriented graph' by 'simple digraph', and we can do the same in the paragraph following (3).

Acknowledgements. The first and third authors are supported by the ANR project 21-CE48-0012 DAGDigDec (DAGs and Digraph Decompositions). The third author is also partially supported by the MICINN project PID2022-137283NB-C22 ACoGe (Algebraic combinatorics and its connections to geometry). Some of this work was carried out during the First Armenian Workshop on Graphs, Combinatorics, Probability and their Applications to Machine Learning (2019) and the authors are grateful to the organizers.

References

[1] R. Aharoni, E. Berger, O. Kfir, Acyclic systems of representatives and acyclic colorings of digraphs, *J. Graph Theory* 59(3): 177–189 (2008).

[2] M. Ajtai, J. Komlós, E. Szemerédi, A Note on Ramsey Numbers, *J. Comb. Theory, Ser. B* 29(3): 354–360 (1980).

[3] N. Alon, J. Pach, J. Solymosi, Ramsey-type Theorems with Forbidden Subgraphs, *Combinatorica* 21(2): 155–170 (2001).

[4] J. Bang-Jensen, T. Bellitto, T. Schweser, M. Stiebitz, Hajos and Ore constructions of digraphs, *Electron. J. Comb.*, 27(1), 2020.

[5] J. Bensmail, A. Harutyunyan, N. K. Le, List coloring digraphs, *J. Graph Theory* 87(4): 492–508 (2018).

[6] S. Bessy, S. Thomassé, Spanning a strong digraph by alpha circuits: A proof of Gallai's conjecture, *Combinatorica* 27(6): 659–667 (2007).

[7] D. Bokal, G. Fijavž, M. Juvan, P. M. Kayll, B. Mohar, The circular chromatic number of a digraph, *J. Graph Theory* 46 (2004), 227–240.

[8] B. Bollobás, *Random Graphs*, Cambridge University Press, 2001.

[9] J. A. Bondy, Diconnected orientations and a conjecture of Las Vergnas, *J. London Math. Soc.* s2-14(2): 277–282 (1976).

[10] P. Camion, Chemins et circuits hamiltoniens des graphes complets, *C. R. Acad. Sci. Paris* 249: 2151–2152 (1959).

[11] Z. Chen, J. Ma and W. Zang, Coloring digraphs with forbidden cycles, *J. Comb. Theory, Ser. B* 115: 210–223 (2015).

[12] N. Chen, M. Olvera-Cravioto, Directed random graphs with given degree distributions, *Stoch. Syst.* 3(1): 147–186 (2013).

[13] N. Cordero-Michel, H. Galeana-Sánchez, New bounds for the dichromatic number of a digraph, *Discrete Math. Theor. Comput. Sci.* 21(1) #7 (2019).

[14] P. Creed, *Counting and Sampling Problems on Eulerian Graphs*, PhD thesis, 2010.

[15] P. Erdős, Problems and results in number theory and graph theory, *Proceedings of Ninth Manitoba Conference on Numerical Math. and Computing*, 3–21, 1979. https://www.renyi.hu/~p_erdos/1980-30.pdf

[16] P. Erdős, A. Hajnal, Ramsey-type theorems, *Discrete Appl. Math.* 25: 37–52 (1989).

[17] P. Erdős, L. Moser, On the representation of directed graphs as unions of orderings, *Publ. Math. Inst. Hung. Acad. Sci., Ser. A* 9: 125–132 (1964).

[18] P. Flajolet, R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, 2009.

[19] A. Harutyunyan, B. Mohar, Two results on the digraph chromatic number, *Discrete Math.* 10(312), 2011.

[20] A. Harutyunyan, B. Mohar, Gallai's Theorem for List Coloring of Digraphs, *SIAM J. Discrete Math.* 25(1): 170–180 (2011).

[21] S. Janson, A. Ruciński, and T. Łuczak, *Random Graphs*, John Wiley & Sons, 2011.

[22] A. Johansson, Asymptotic Choice Number for Triangle Free Graphs, *DIMACS Technical Report* 91-5 (1996).

[23] J. H. Kim, The Ramsey Number $R(3,t)$ Has Order of Magnitude $t^2/\log t$, *Random Struct. Algorithms* 7(3): 173–207 (1995).

[24] M. Molloy, The list chromatic number of graphs with small clique number, *J. Comb. Theory, Ser. B* 134: 264–284 (2019).

[25] M. Molloy and B. Reed, *Graph Colouring and the Probabilistic Method*, Springer, 2002.

[26] V. Neumann-Lara, The dichromatic number of a digraph, *J. Comb. Theory, Ser. B* 33(3): 265–270 (1982).

[27] A. Sanchez-Flores, On Tournaments Free of Large Transitive Subtournaments, *Graphs Comb.* 14: 181–200 (1998).

[28] J. B. Shearer, A note on the independence number of triangle-free graphs, *Discrete Math.* 46(1): 83–87 (1983).

[29] J. Spencer, C. R. Subramanian, On the size of induced acyclic subgraphs in random digraphs, *Discrete Math. Theor. Comput. Sci.* 10(2): 47–54 (2008).