Received: 27 March 2024 | Revised: 14 June 2024 | Accepted: 7 July 2024

DOI: 10.1002/rsa.21251

RESEARCH ARTICLE

WILEY

Seymour's second neighbourhood conjecture: random graphs and reductions

Alberto Espuny Díaz[1] | **António Girão**[2] | **Bertille Granet**[1] | **Gal Kronenberg**[2]

[1]Institut für Informatik, Universität Heidelberg, Heidelberg, Germany
[2]Mathematical Institute, University of Oxford, Oxford, United Kingdom

Correspondence
Bertille Granet, Institut für Informatik, Universität Heidelberg, 69120 Heidelberg, Germany.
Email: granet@informatik.uni-heidelberg.de

Funding information
Deutsche Forschungsgemeinschaft (DFG), Grant/Award Numbers: 447645533, 513704762, 428212407; ERC, Grant/Award Number: 883810; the Alexander von Humboldt Foundation; the Royal Commission for the Exhibition of 1851.

Abstract

A longstanding conjecture of Seymour states that in every oriented graph there is a vertex whose second outneighbourhood is at least as large as its outneighbourhood. In this short note we show that, for any fixed $p \in [0, 1/2)$, a.a.s. every orientation of $G(n, p)$ satisfies Seymour's conjecture (as well as a related conjecture of Sullivan). This improves on a recent result of Botler, Moura and Naia. Moreover, we show that $p = 1/2$ is a natural barrier for this problem, in the following sense: for any fixed $p \in (1/2, 1)$, Seymour's conjecture is actually equivalent to saying that, with probability bounded away from 0, every orientation of $G(n, p)$ satisfies Seymour's conjecture. This provides a first reduction of the problem. For a second reduction, we consider minimum degrees and show that, if Seymour's conjecture is false, then there must exist arbitrarily large strongly-connected counterexamples with bounded minimum outdegree. Contrasting this, we show that vertex-minimal counterexamples must have large minimum outdegree.

KEYWORDS

oriented graphs, random graphs, second neighbourhood conjecture

1 | INTRODUCTION

Given any digraph D and a vertex $v \in V(D)$, let us write $N_1(v) := \{w \in V \setminus \{v\} : vw \in E(D)\}$ for the *outneighbourhood* of v, and let $N_2(v) := \{w \in V \setminus (\{v\} \cup N_1(v)) : E(N_1(v), w) \neq \emptyset\}$ denote the *second outneighbourhood* of v. We say that v is a *Seymour vertex* if $|N_2(v)| \geq |N_1(v)|$. In this note, the word *graph* refers to finite, simple graphs. Seymour conjectured the following (see [8]).

Conjecture 1. *Every oriented graph has a Seymour vertex.*

Seymour's conjecture, which is closely tied to a particular case of the well-known conjecture of Caccetta and Häggkvist [4], has attracted a lot of attention due to its simplicity, and yet it remains widely open. The statement has only been proved for some special classes of oriented graphs. Of particular interest is the case of tournaments, where the conjecture is known to hold: solving a conjecture of Dean [8], this was first proved by Fisher [12], and later reproved by Havet and Thomassé [14]. Other special classes of graphs or variants of the conjecture have been considered in different papers (see, e.g., [1, 5, 6, 11, 13, 16, 17] and the references therein).

1.1 | Orientations of random graphs

In this note, we focus primarily on the case of orientations of random graphs. The first results in this direction are due to Cohn, Godbole, Wright Harkness and Zhang [7], who considered random orientations of the binomial random graph $G(n, p)$ (i.e., a graph on n vertices where each possible edge is included independently with probability p). They showed that, when $C_1(\log n/n)^{1/2} \leq p \leq 1 - C_2(\log n/n)^{1/2}$, a.a.s. (i.e., with probability tending to 1 as $n \to \infty$) a random orientation of $G(n, p)$ contains a Seymour vertex. This implies, in particular, that almost every oriented graph has a Seymour vertex. A more interesting approach was taken by Botler, Moura and Naia [3], who considered *all* orientations of the random graph $G(n, p)$. They showed that a.a.s. *every* orientation of $G(n, p)$ contains a Seymour vertex whenever $p < 1/4$. Here we improve the range of p for which this holds.

Theorem 2. *Let $p < 1/2$. Then, a.a.s. every orientation of $G(n, p)$ contains a Seymour vertex.*

Our methods also apply to a related conjecture of Sullivan [18]; see Appendix A.

Moreover, it turns out that $p = 1/2$ is a natural barrier for this problem. Indeed, we show that, if we could extend Theorem 2 to any value of $p > 1/2$, then we would prove Conjecture 1 in full generality. This means that proving Conjecture 1 in general is as hard as proving that it holds for the random graph $G(n, p)$ for $p \in (1/2, 1)$, providing a first reduction of the problem.

Theorem 3. *Let $p \in (1/2, 1)$. If Conjecture 1 is false, then a.a.s. there is an orientation of $G(n, p)$ with no Seymour vertex.*

1.2 | Counterexamples and minimum degrees

We show that if Conjecture 1 is false, then a vertex-minimal counterexample must have high minimum outdegree.

Proposition 4. *Let D be a vertex-minimal counterexample to Conjecture 1. Then, $\delta^+(D) > \sqrt{|V(D)|}$.*

On the other hand, the following result essentially says that, if there exist any counterexamples to Conjecture 1, then there must exist arbitrarily large strongly-connected counterexamples with bounded minimum outdegree. This gives a second reduction of Seymour's conjecture.

Proposition 5. *Suppose Conjecture 1 is false. Then, for every function $d = d(n) = \omega(1)$, there exist infinitely many $n \in \mathbb{N}$ for which there exists an n-vertex strongly-connected oriented graph D with $\delta^+(D) < d$ which contains no Seymour vertex.*

2 | PROOFS

Our notation will be standard. For any positive integer n, we write $[n] := \{1, \dots, n\}$. Throughout, for the sake of readability, we assume that n is sufficiently large when needed and ignore rounding for asymptotic statements. Given a graph $G = (V, E)$, for any disjoint sets $A, B \subseteq V$ we write $e(A, B)$ for the number of edges with one endpoint in A and the other in B. Given an orientation $\vec{G}$ of G, we will write $\vec{e}(A, B)$ to denote the number of (oriented) edges from A to B. Moreover, for any vertex $v \in V$, we write $d^+(v) := |N_1(v)|$ for the outdegree of v in $\vec{G}$. Given a set $A \subseteq V(D)$, we define $N_1(A) := (\bigcup_{v \in A} N_1(v)) \setminus A$. The graphs or oriented graphs to which the notation refers will always be clear from the context.

2.1 | Orientations of random graphs

We will use the following version of Chernoff's bound (see, e.g., [15, Corollary 2.3]).

Lemma 6. *Let $X \sim \mathrm{Bin}(n, p)$ be a binomial random variable. Then, for all $0 < \delta < 1$ we have that $\mathbb{P}[|X - np| \geq \delta np] \leq 2\mathrm{e}^{-\delta^2 np/3}$.*

We begin with the proof of Theorem 2.

Proof of Theorem 2. Consider $G(n, p)$. By the result of Botler, Moura and Naia [3, Theorem 2], we may assume, for example, that $p \geq 1/8$ (our proof works for any constant $p \in (0, 1/2)$). We will need some probabilistic estimates. Let $\varepsilon := 1/2 - p$, so $\varepsilon \in (0, 3/8]$. Let $\delta := \varepsilon/3$. Let $C = C(\varepsilon)$ be sufficiently large, and let $C' := 2C/\varepsilon$. We will use the following properties.

Claim 1. A.a.s. $G \sim G(n, p)$ satisfies the following properties:

1. Every vertex $v \in V(G)$ satisfies that $d(v) \leq (1 - \varepsilon)n/2$.
2. For every pair of disjoint sets $A, B \subseteq V(G)$ with $|A|, |B| \geq C \log n$ we have that

$$(p - \delta)|A||B| \leq e(A, B) \leq (1 - \varepsilon)|A||B|/2.$$

Proof. Property (1) is standard, and follows by a direct application of Lemma 6 and a union bound over all vertices. Property (2) follows similarly, though the union bound is a bit more involved, so we include the details.

It suffices to show that a.a.s. $e(A, B) = (p \pm \delta)|A||B|$ for all the required pairs of sets. Fix any disjoint sets $A, B \subseteq V(G)$ with $|A|, |B| \geq C \log n$. By Lemma 6, we have that

$$\mathbb{P}[e(A, B) \neq (p \pm \delta)|A||B|] = \mathbb{P}[e(A, B) \neq (1 \pm \delta/p)p|A||B|] \leq 2\exp(-\delta^2|A||B|/3p).$$

We may assume without loss of generality that $|A| \leq |B|$. By a union bound over all choices of A and B, the probability P that the statement fails satisfies

$$P \leq 2 \sum_{a=C\log n}^{n/2} \sum_{b=a}^{n} \binom{n}{a}\binom{n}{b} \exp\left(-\frac{\delta^2}{3p}ab\right)$$
$$\leq \sum_{a=C\log n}^{n/2} \sum_{b=a}^{n} n^a n^b \exp\left(-\frac{\delta^2}{3p}ab\right) = \sum_{a=C\log n}^{n/2} \sum_{b=a}^{n} \exp\left((a+b)\log n - \frac{\delta^2}{3p}ab\right).$$

By making the constant C sufficiently large, we may ensure that $a\log n, b\log n \leq \delta^2 ab/12p$, and so

$$P \leq \sum_{a=C\log n}^{n/2} \sum_{b=a}^{n} \exp\left(-\frac{\delta^2}{6p}ab\right) \leq \Theta(n^2)n^{-\Theta(\log n)} = o(1). \qquad \blacksquare$$

Now, let G be any n-vertex graph satisfying the properties of Claim 1, and let $\vec{G}$ be an arbitrary orientation of G. We are going to show that, if n is sufficiently large, then $\vec{G}$ contains a Seymour vertex. Let $x_1, \dots, x_n$ be a labelling of $V(G)$ such that for all $i, j \in [n]$ with $i \leq j$ we have that $d^+(x_i) \geq d^+(x_j)$; that is, the labels are assigned by decreasing order of the outdegrees of the vertices. Let $X_0 := \emptyset$ and, for each $i \in [n]$, let $X_i := \{x_j : j \in [i]\}$.

Claim 2. If there is some $i \in [n]$ such that $|N_1(x_i) \cap X_{i-1}| \geq C' \log n$, then x_i is a Seymour vertex.

Proof. Assume there is some $i \in [n]$ satisfying the condition. Let $A := N_1(x_i) \cap X_{i-1}$, so $|A| \geq C' \log n$. Let $m := |N_1(x_i)|$ (and note that $m \geq |A|$). Let $S \subseteq A$ be the set of all vertices $v \in A$ with $\vec{e}(v, N_1(x_i)) \geq (1+\varepsilon)m/2$, and let $T \subseteq A$ be the set of all vertices $v \in A$ with $\vec{e}(v, V \setminus (\{x_i\} \cup N_1(x_i))) \geq (1-\varepsilon)m/2$. Note that, by definition,

$$\vec{e}(T, N_2(x_i)) = \vec{e}(T, V \setminus (\{x_i\} \cup N_1(x_i))) \geq |T|(1-\varepsilon)m/2. \qquad (1)$$

Observe that $|S| < C\log n$ (indeed, if we assume otherwise, then for any set $S' \subseteq S$ of size $|S'| = C\log n$ we have that

$$e(S', N_1(x_i) \setminus S') \geq \sum_{v\in S'} \vec{e}(v, N_1(x_i) \setminus S') \geq |S'|((1+\varepsilon)m/2 - C\log n) \geq |S'|m/2,$$

which contradicts Claim 1 (2)). Since $A \setminus S \subseteq T$, this implies that $|T| \geq C\log n$. It thus follows from Claim 1 (2) that

$$\vec{e}(T, N_2(x_i)) \leq (1-\varepsilon)|T||N_2(x_i)|/2. \qquad (2)$$

Combining (1) and (2), it follows that x_i is a Seymour vertex. ■

By Claim 2, we may assume that $\vec{G}$ satisfies that

$$\text{for all } i \in [n] \text{ we have } |N_1(x_i) \cap X_{i-1}| \leq C' \log n. \qquad (3)$$

Now let $B := X_{2\delta n}$ be partitioned into $B_1 := X_{\delta n}$ and $B_2 := B \setminus B_1$. By Claim 1 (2), we have that $e(B_1, B_2) \geq (p-\delta)|B_1||B_2| = (p-\delta)\delta^2 n^2$. By (3), for each $v \in B_2$ we have $\vec{e}(v, B_1) \leq C' \log n$, and so $\vec{e}(B_1, B_2) \geq (p-2\delta)\delta^2 n^2$. It follows by averaging that there is some vertex $y \in B_1$ such that $\vec{e}(y, B_2) \geq (p-2\delta)\delta n$. Let $Y := N_1(y) \cap B_2$, and let Z be the set of vertices $z \in V \setminus B$ such that $\vec{e}(Y, z) = 0$.

We claim that $|Z| < C \log n$. Indeed, assume that $|Z| \geq C \log n$, let $Z' \subseteq Z$ with $|Z'| = C \log n$, and partition Y arbitrarily into $\Theta(n/\log n)$ sets of size at least $C \log n$. By Claim 1 (2), there are edges of G between Z' and each of these sets, so together $e(Y, Z') = \Omega(n/\log n)$. By averaging, there must be some vertex $z \in Z'$ with $e(Y, z) = \Omega(n/\log^2 n)$. Since $\vec{e}(Y, z) = 0$ by definition, we must have that $\vec{e}(z, Y) = \Omega(n/\log^2 n)$. However, this contradicts (3).

Now, since $V \setminus (B \cup Z) \subseteq N_1(y) \cup N_2(y)$, it follows that $|N_1(y) \cup N_2(y)| \geq (1-3\delta)n$. By Claim 1 (1), we have that $|N_1(y)| \leq (1-\varepsilon)n/2$, and hence $|N_2(y)| \geq (1-3\delta)n - (1-\varepsilon)n/2 \geq |N_1(y)|$, so y is a Seymour vertex. ■

Next, we prove Theorem 3.

Proof of Theorem 3. Let D be an oriented graph for which Conjecture 1 fails, and let H be its underlying graph. Let $h := |V(H)|$. We say that a graph G on $n \geq h$ vertices has a *good ordering* if there is a labelling $x_1, \ldots, x_n$ of its vertices satisfying the following properties:

(P1) $G[\{x_1, \ldots, x_h\}]$ induces a copy of H, and
(P2) for every $i \in \{h+1, \ldots, n\}$ we have $|N(x_i) \cap \{x_1, \ldots, x_{i-1}\}| \geq i/2$.

We claim that, if G has a good ordering $x_1, \ldots, x_n$, then there is an orientation of G with no Seymour vertex. Indeed, consider the orientation $\vec{G}$ of G where $\vec{G}[\{x_1, \ldots, x_h\}]$ induces a copy of D (which is possible by (P1)) and all other edges are oriented towards the vertex with smaller label in the ordering. It is clear that none of $x_1, \ldots, x_h$ can be Seymour vertices in $\vec{G}$: by construction, their outneighbourhoods are defined only by (a subgraph isomorphic to) D, so if they were a Seymour vertex for $\vec{G}$, they would also be a Seymour vertex for D. Moreover, for any $i \in \{h+1, \ldots, n\}$, both $N_1(x_i)$ and $N_2(x_i)$ are contained in $\{x_1, \ldots, x_{i-1}\}$. By (P2), we are guaranteed that $|N_1(x_i)| \geq i/2$, and so $|N_2(x_i)| < |N_1(x_i)|$.

It remains to show that a.a.s. $G(n,p)$ has a good ordering. We will make use of the following standard properties.

Claim 3. Let $\varepsilon := p - 1/2$. A.a.s. $G \sim G(n,p)$ satisfies the following properties.

1. G contains an induced copy of H.
2. Every vertex $v \in V(G)$ satisfies $d(v) \geq (1+\varepsilon/2)n/2$.
3. For every set $X \subseteq V(G)$ with $|X| \leq (1-\varepsilon/4)n$, there exists a vertex $v \in V(G) \setminus X$ such that $e_G(v, X) \geq |X|/2$.

Proof. Property (1) is well known (see, e.g., [9, Proposition 11.3.1]). Property (2) is standard and follows by an application of Lemma 6 to the degree of a vertex and a union bound over all vertices. Here we only show the details for the (also simple) proof of property (3). Let

$$a := \frac{3\log 16}{p\left(1-\frac{1}{2p}\right)^2}$$

(note that this is independent of n).

Consider first all sets $X \subseteq V(G)$ of size $|X| < a$. We first show that a.a.s. all vertices in each such set have a common neighbour in G. Indeed, fix a set X with $|X| = k < a$. The probability that any given vertex $v \in V(G) \setminus X$ is a neighbour of all vertices in X is p^k, which is a constant. Thus, the probability that no vertex is a common neighbour of X is $(1-p^k)^{n-|X|} \leq \mathrm{e}^{-p^k n/2} = o(n^{-2k})$. By a union bound over all sets $X \subseteq V(G)$ with $|X| = k$ and over the possible values of k, we reach the desired conclusion.

Consider now any set $X \subseteq V(G)$ with $a \leq |X| \leq n/2$. For any vertex $v \in V(G) \setminus X$, we have by Lemma 6 that $\mathbb{P}[e(v, X) < |X|/2] \leq 1/8$. Thus, the probability that all vertices $v \in V(G) \setminus X$ have fewer than $|X|/2$ neighbours in X is at most $8^{-(n-|X|)} \leq 2^{-3n/2}$. The conclusion follows by a union bound over the at most 2^n choices for X.

Lastly, fix any set $X \subseteq V(G)$ with $n/2 \leq |X| \leq (1-\varepsilon/4)n$. Similarly as above, for any vertex $v \in V(G) \setminus X$ we have by Lemma 6 that $\mathbb{P}[e(v, X) < |X|/2] \leq \mathrm{e}^{-\Theta(n)}$, so the probability that all vertices $v \in V(G) \setminus X$ have fewer than $|X|/2$ neighbours in X is at most $\mathrm{e}^{-\Theta(n^2)}$. One last union bound completes the proof of the claim. ■

Condition now on the event that $G \sim G(n, p)$ satisfies the properties of Claim 3, which holds a.a.s. We construct a good ordering for G as follows. First, choose an arbitrary set of h vertices which induces a copy of H (which exists by Claim 3 (1)) and label the vertices $x_1, \ldots, x_h$ arbitrarily. Thus, (P1) is satisfied. Next, for each $i \in \{h+1, \ldots, (1-\varepsilon/4)n\}$ in turn, by Claim 3 (3), there exists some vertex $v \in V(G) \setminus \{x_1, \ldots, x_{i-1}\}$ satisfying that $|N(v) \cap \{x_1, \ldots, x_{i-1}\}| \geq i/2$; let x_i be an arbitrary vertex satisfying this property. Finally, by Claim 3 (2), all vertices $v \in V(G) \setminus \{x_1, \ldots, x_{(1-\varepsilon/4)n}\}$ satisfy that $d(v) \geq (1+\varepsilon/2)n/2$, and thus, $|N(v) \cap \{x_1, \ldots, x_{(1-\varepsilon/4)n}\}| \geq n/2$. Thus, the remaining vertices can be labelled arbitrarily while guaranteeing that (P2) holds. ■

2.2 | Counterexamples and minimum degrees

We begin with the simple proof of Proposition 5.

Proof of Proposition 5. Let D_0 be a counterexample to Conjecture 1. We define $n_0 := |V(D_0)|$ and $d_0 := \delta^+(D_0)$. Let $N_0 \in \mathbb{N}$ be such that $d(n) > d_0 + n_0$ for all $n \geq N_0 n_0$. For any $N \geq N_0$, we construct a strongly-connected counterexample D to Conjecture 1 with Nn_0 vertices and $\delta^+(D) < d(Nn_0)$.

The construction is a "blow-up" of a consistently-oriented cycle where each vertex is replaced by a copy of D_0. Formally, let $D_0, D_1, \ldots, D_{N-1}$ be vertex-disjoint copies of D_0 and, for each $i \in [N]$, add a complete bipartite graph between the vertices of D_i and those of D_{i+1}, with all edges oriented towards D_{i+1} (where we take indices modulo N). Note that the resulting oriented graph D is strongly connected (simply by considering the complete bipartite oriented graphs between the different copies of D_0) and that $\delta^+(D) = d_0 + n_0 < d(Nn_0)$, as desired. By the symmetry of the construction, we now only need to verify that no vertex in D_1 is a Seymour vertex for D. Let us write $N'(\cdot)$ to denote neighbourhoods in D_1, while we use $N(\cdot)$ for neighbourhoods in D. Note that, for every $v \in V(D_1)$, we have that $N_1(v) = N'_1(v) \cup V(D_2)$ and $N_2(v) = N'_2(v) \cup V(D_3)$, and so $|N_2(v)| \geq |N_1(v)|$ if and only if $|N'_2(v)| \geq |N'_1(v)|$. But D_1 contains no Seymour vertices, so neither does D. ■

Now note that any vertex-minimal counterexample D to Conjecture 1 must be strongly connected. Indeed, otherwise one may consider the auxiliary directed acyclic graph obtained from the strongly-connected components of D and note that any strong component which has no outneighbours in this auxiliary graph would itself be a smaller counterexample to the conjecture.

Proof of Proposition 4. Suppose for a contradiction that the statement is false, that is, that there is a vertex-minimal counterexample D with n vertices and $d := \delta^+(D) \le \sqrt{n}$. Pick a vertex $x \in V(D)$ of smallest outdegree, that is, $d^+(x) = d$, and let $A_1 := \{x\} \cup N_1(x)$, $X_1 := N_2(x)$, and $B_1 := V(D) \setminus (A_1 \cup X_1)$. Note that, by assumption, x is not a Seymour vertex, so $|X_1| < d$. Moreover, $E(A_1, B_1) = \emptyset$.

Now we go through the following iterative process: while X_i contains some non-empty subset X_i' with $|N_1(X_i') \cap B_i| < |X_i'|$, define $A_{i+1} := A_i \cup X_i'$, $X_{i+1} := (X_i \setminus X_i') \cup (N_1(X_i') \cap B_i)$ and $B_{i+1} := B_i \setminus N_1(X_i')$, and proceed to the next iteration. Observe that, throughout the process, we maintain the property that $E(A_i, B_i) = \emptyset$. In particular, since D is strongly connected, this guarantees that, if $B_i \neq \emptyset$, then $E(X_i, B_i) \neq \emptyset$. In the i-th iteration of the above procedure we decrease the size of X_i by at least 1, and decrease the size of B_i by at most $|X_i|$. Hence, after at most $d-1$ steps, the procedure must stop, say, at step $t < d$ with $X_t \neq \emptyset$ and $|B_t| \ge n - \binom{d}{2} \ge 1$ by the choice of d.

By the minimality of D, $D[A_t]$ must contain a Seymour vertex, say z. We claim that z is also a Seymour vertex for D. Indeed, let us write $N'(\cdot)$ to denote neighbourhoods in $D[A_t]$. Then, using that $E(A_t, B_t) = \emptyset$, we have that

$$|N_2(z)| - |N_1(N_1(z) \cap X_t) \cap B_t| \ge |N_2'(z)| \ge |N_1'(z)| = |N_1(z)| - |N_1(z) \cap X_t|.$$

Since the procedure above has ended, we have that $|N_1(N_1(z) \cap X_t) \cap B_t| \ge |N_1(z) \cap X_t|$, and so it follows that $|N_2(z)| \ge |N_1(z)|$. This contradicts the fact that D has no Seymour vertex. ■

3 | CONCLUDING REMARKS

In Theorem 2 we showed that, for fixed $p < 1/2$, a.a.s. $G(n,p)$ satisfies Seymour's conjecture, and in Theorem 3 we showed that, for fixed $p \in (1/2, 1)$, Conjecture 1 is equivalent to knowing that, with probability bounded away from 0, all orientations of the random graph $G(n,p)$ contain a Seymour vertex. It would be interesting to see whether the latter provides a useful avenue for tackling Conjecture 1. We note that Botler, Moura and Naia [3] showed that, for fixed $p \in [1/2, 2/3)$, a.a.s. every orientation of $G(n,p)$ with minimum outdegree $\omega(\sqrt{n})$ has a Seymour vertex, so for p in this range it now suffices to study orientations with small minimum outdegree.

Neither of our results applies when $p = (1 \pm o(1))/2$, which is arguably one of the more interesting values for p. In particular, it would be interesting to understand whether every orientation of almost all graphs (which corresponds precisely to the case $p = 1/2$) contains a Seymour vertex. In the spirit of applying a similar method as in our proof of Theorem 3, we ask the following question.

Question 7. *Does $G(n, 1/2)$ a.a.s. have an ordering $x_1, \dots, x_n$ of its vertices such that $|N(x_i) \cap \{x_1, \dots x_{i-1}\}| \ge i/2$ for all $i \in [n]$?*

Theorem 3 also leaves a gap for the cases when $p = 1 - o(1)$. It would thus be interesting to consider orientations of $G(n,p)$ when p tends to 1 (with the case $p = 1$, corresponding to tournaments, being solved already [12, 14]).

Lastly, Propositions 4 and 5 provide information about the minimum degree of (assumed) counterexamples to Conjecture 1. Proposition 5 shows that, if Conjecture 1 is false, then we must be able to construct arbitrarily large strongly-connected counterexamples with bounded minimum outdegree. Proposition 4, on the other hand, shows that the search for (vertex-minimal) counterexamples may be limited to oriented graphs with large minimum outdegree.

ACKNOWLEDGMENTS

We are grateful to the anonymous referees for helpful comments on an earlier version of this paper. Alberto Espuny Díaz was partially funded by the Deutsche Forschungsgemeinschaft (DFG) through projects nos. 447645533 and 513704762. António Girão is supported by ERC advanced grant 883810. Bertille Granet was partially supported by the DFG project 428212407 and by the Alexander von Humboldt Foundation. Gal Kronenberg is supported by the Royal Commission for the Exhibition of 1851. Open Access funding enabled and organized by Projekt DEAL.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Bertille Granet https://orcid.org/0009-0002-6453-2295

REFERENCES

1. J. Ai, S. Gerke, G. Gutin, S. Wang, A. Yeo, and Y. Zhou. On Seymour's and Sullivan's second neighbourhood conjectures. arXiv e-prints arXiv: 2306.03493. 2023.
2. N. Alon and J. H. Spencer, *The probabilistic method*. Wiley series in discrete mathematics and optimization, 4th ed., John Wiley & Sons, Inc, Hoboken, NJ, 2016.
3. F. Botler, P. F. S. Moura, and T. Naia, *Seymour's second neighborhood conjecture for orientations of (pseudo)random graphs*, Discret. Math. **346** (2023), no. 12, 113583. https://doi.org/10.1016/j.disc.2023.113583.
4. L. Caccetta and R. Häggkvist. On minimal digraphs with given girth. Proceedings of the ninth Southeastern conference on combinatorics, graph theory, and computing (Florida Atlantic Univ., Boca Raton, Fla., 1978), congress. Numer., vol. XXI, Utilitas Math., Winnipeg, MB. 1978 pp. 181–187.
5. B. Chen and A. Chang, *A note on Seymour's second neighborhood conjecture*, Discret. Appl. Math. **337** (2023), 272–277. https://doi.org/10.1016/j.dam.2023.05.012.
6. G. Chen, J. Shen, and R. Yuster, *Second neighborhood via first neighborhood in digraphs*, Ann. Comb. **7** (2003), no. 1, 15–20. https://doi.org/10.1007/s000260300001.
7. Z. Cohn, A. Godbole, E. Wright Harkness, and Y. Zhang, *The number of Seymour vertices in random tournaments and digraphs*, Graphs and Combinatorics **32** (2016), no. 5, 1805–1816. https://doi.org/10.1007/s00373-015-1672-9.
8. N. Dean and B. J. Latka, *Squaring the tournament—An open problem*, Congr. Numerantium **109** (1995), 73–80.
9. R. Diestel, *Graph theory*, *graduate texts in mathematics*, Vol **173**, Springer, Berlin, 2017.
10. P. Erdős and A. Rényi, *On random graphs. I*, Publ. Math. Debrecen **6** (1959), 290–297.
11. D. Fidler and R. Yuster, *Remarks on the second neighborhood problem*, J. Graph Theory **55** (2007), no. 3, 208–220. https://doi.org/10.1002/jgt.20229.
12. D. C. Fisher, *Squaring a tournament: A proof of Dean's conjecture*, J. Graph Theory **23** (1996), no. 1, 43–48. https://doi.org/10.1002/(SICI)1097-0118(199609)23:1<43::AID-JGT4>3.0.CO;2-K.
13. S. Ghazal, *Seymour's second neighborhood conjecture for tournaments missing a generalized star*, J. Graph Theory **71** (2012), no. 1, 89–94. https://doi.org/10.1002/jgt.20634.
14. F. Havet and S. Thomassé, *Median orders of tournaments: A tool for the second neighborhood problem and Sumner's conjecture*, J. Graph Theory **35** (2000), no. 4, 244–256. https://doi.org/10.1002/1097-0118(200012)35:4<244::AID-JGT2>3.0.CO;2-H.
15. S. Janson, T. Łuczak, and A. Ruciński, Random graphs, Wiley-Interscience series in discrete mathematics and optimization, Wiley-Interscience, New York, 2000. https://doi.org/10.1002/9781118032718.
16. A. Lladó, *On the second neighborhood conjecture of Seymour for regular digraphs with almost optimal connectivity*, Eur. J. Comb. **34** (2013), no. 8, 1406–1410. https://doi.org/10.1016/j.ejc.2013.05.023.

17. T. Seacrest, *The arc-weighted version of the second neighborhood conjecture*, J. Graph Theory **78** (2015), no. 3, 219–228. https://doi.org/10.1002/jgt.21800.
18. B. D. Sullivan. A summary of problems and results related to the Caccetta-Häggkvist conjecture. arXiv Mathematics e-prints arXiv: math/0605646. 2006.

How to cite this article: A. Espuny Díaz, A. Girão, B. Granet, and G. Kronenberg, *Seymour's second neighbourhood conjecture: random graphs and reductions*, Random Struct. Alg. (2024), 1–12. https://doi.org/10.1002/rsa.21251

APPENDIX A: ON SULLIVAN'S SECOND NEIGHBOURHOOD CONJECTURE FOR RANDOM GRAPHS

Sullivan [18] proposed multiple variants and strengthenings of Conjecture 1. Here we consider one of them. Given any digraph D and a vertex $v \in V(D)$, let us now write $N^+(v) := N_1(v)$ for the outneighbourhood of v, $N^-(v) := \{w \in V \setminus \{v\} : wv \in E(D)\}$ for the inneighbourhood of v, and $N_2^+(v) := N_2(v)$ for the second outneighbourhood of v. Given a set $A \subseteq V(D)$, we define $N^+(A) := (\bigcup_{v \in A} N^+(v)) \setminus A$. We say that v is a *Sullivan vertex* if $|N_2^+(v)| \geq |N^-(v)|$. (Notice that, contrary to the definition of a Seymour vertex, $N_2^+(v)$ and $N^-(v)$ may intersect.)

Conjecture 8 ([18, Conjecture 6.6]). *Every oriented graph has a Sullivan vertex.*

Conjecture 8 has received far less attention than Conjecture 1. Only recently, Ai, Gerke, Gutin, Wang, Yeo and Zhou [1] showed that it holds for certain classes of graphs, including tournaments, planar oriented graphs and some families of oriented split graphs. Moreover, they also showed that almost all oriented graphs satisfy Conjecture 8, as a.a.s. a random orientation of $G(n, p)$ (for $p \in (0, 1)$ independent of n) satisfies the conjecture. Here we use similar ideas to those presented in the paper to prove a result analogous to Theorem 2 for Sullivan's conjecture.

Theorem 9. *Let* $p = p(n) \in [0, 1]$ *be such that* $\limsup_{n\to\infty} p < 1/2$. *Then, a.a.s. every orientation of* $G(n, p)$ *contains a Sullivan vertex.*

As a difference with respect to Theorem 2, we note that here we must consider the cases when p tends to 0; for Seymour's conjecture, these cases were covered by the earlier results of Botler, Moura and Naia [3].

We note first that, if a graph G has an isolated vertex, then so do all orientations of G, and isolated vertices are Sullivan vertices. Thus, it follows from the well-known threshold for connectivity [10] that, if $p \leq (1 - \varepsilon) \log n/n$, then a.a.s. every orientation of $G(n, p)$ has a Sullivan vertex. Thus, in order to prove Theorem 9 we only need to consider larger values of p. In particular, Theorem 9 follows directly from the following two results.

Proposition 10. *Let* $3 \log n/4n \leq p = p(n) = o(1)$. *Then, a.a.s. every orientation of* $G(n, p)$ *contains a Sullivan vertex.*

Proposition 11. *Let* $p = p(n) \in [0, 1]$ *with* $p = \omega(\log^{-1/3} n)$ *and* $\limsup_{n\to\infty} p < 1/2$. *Then, a.a.s. every orientation of* $G(n, p)$ *contains a Sullivan vertex.*

We will need to bound large deviations for the upper tail of some binomial random variables. For this, we will use the following Chernoff bound (see, e.g., [2, Corollary A.1.10]).

Lemma 12. *Let* $X \sim \mathrm{Bin}(n,p)$ *be a binomial random variable. Then, for all* $\delta > 0$ *we have that* $\mathbb{P}[X \geq (1+\delta)np] \leq (\mathrm{e}^{\delta}/(1+\delta)^{1+\delta})^{np}$.

Proof of Proposition 10. We begin our proof by observing that a.a.s. $G(n,p)$ satisfies certain properties.

Claim 4. A.a.s. $G \sim G(n,p)$ satisfies the following properties:

1. For each pair of distinct vertices $u, v \in V(G)$, there are fewer than $np/2400$ paths of length 2 joining u and v.
2. Every vertex $v \in V(G)$ satisfies $d(v) \leq 4np$.
3. For every set $A \subseteq V(G)$ of $n/3$ vertices we have $e(A) \geq n^2p/25$.

Proof. All claims follow from direct applications of Lemmas 6 and 12 (with Lemma 12 being needed for (1) and (2)) and a union bound. ∎

Now let G be an arbitrary n-vertex graph satisfying the properties from Claim 4. Note that, if in an orientation $\vec{G}$ of G there is some vertex $v \in V(G)$ such that $|N_2^+(v)| > 4np$, then trivially v must be a Sullivan vertex, since by Claim 4 (2) all vertices have indegree at most $4np$. We are going to show that, in fact, every orientation $\vec{G}$ of G contains at least one such vertex, thus deriving the desired result.

We argue by contradiction. Suppose that there is an orientation $\vec{G}$ of G such that every vertex $v \in V(G)$ satisfies $|N_2^+(v)| \leq 4np$. We are going to count consistently oriented paths of length 2 in two different ways to reach a contradiction. Let P denote the number of such paths. First, we may count P by adding over all vertices the number of paths starting at that vertex. For each vertex, the number of such paths is less than $n^2p^2/300$: indeed, any such path must have its endpoint in $N^+(v) \cup N_2^+(v)$, which is a set of size at most $8np$ by Claim 4 (2) and the assumption on the orientation, and by Claim 4 (1) there are fewer than $np/2400$ paths joining any given pair of vertices. Thus,

$$P < n^3p^2/300. \tag{A1}$$

Next, we may obtain P by adding over all vertices $v \in V(G)$ the number of paths of length 2 whose middle vertex is v. Observe that, for each fixed v, this number is $d^-(v)d^+(v)$. We claim that at least $2n/3$ vertices have outdegree at least $np/10$. Indeed, if we assume otherwise, there is a set $A \subseteq V(G)$ of $n/3$ vertices of outdegree at most $np/10$, so by Claim 4 (3) we have that

$$\frac{n^2p}{25} \leq e(A) \leq \sum_{v\in A} d^+(v) \leq \frac{n}{3}\frac{np}{10} = \frac{n^2p}{30},$$

a contradiction. Similarly, at least $2n/3$ vertices must have indegree at least $np/10$. But this means that at least $n/3$ vertices have both in- and outdegree at least $np/10$, which immediately implies that

$$P \geq n^3p^2/300.$$

This results in a contradiction with (A1). ∎

Proof of Proposition 11. Consider $G(n,p)$. As $\limsup_{n\to\infty} p < 1/2$, there exist $n_0 \in \mathbb{N}$ and a constant $\varepsilon > 0$ such that $p \leq 1/2 - \varepsilon$ for all $n \geq n_0$. Let $C = C(\varepsilon)$ be sufficiently large and let $\delta := \varepsilon/5$. We will use the following properties, which follow from standard applications of Lemma 6.

Claim 5. A.a.s. $G \sim G(n,p)$ satisfies the following properties:

1. Every vertex $v \in V(G)$ satisfies that $d(v) = (1 \pm \delta)np$.
2. For every set $A \subseteq V(G)$ with $|A| \geq \delta n$ we have that $e_G(A) = (1 \pm \delta)\binom{|A|}{2}p$.
3. For every pair of disjoint sets $A, B \subseteq V(G)$ with $|A| \geq C \log n$ and $|B| \geq np^2$ we have that $e(A,B) = (1 \pm \delta)|A||B|p$.
4. For every pair of disjoint sets $A, B \subseteq V(G)$ with $|A|, |B| \geq n^{3/4}$ we have that $e(A,B) = (1 \pm \delta)|A||B|p$.

Now, let G be any n-vertex graph satisfying the properties of Claim 5, and let $\vec{G}$ be an arbitrary orientation of G. We are going to show that, if n is sufficiently large, then $\vec{G}$ contains a Sullivan vertex. Let $x_1, \dots, x_n$ be a labelling of $V(G)$ such that for all $i,j \in [n]$ with $i \leq j$ we have that $d^+(x_i) \geq d^+(x_j)$; that is, the labels are assigned by decreasing order of the outdegrees of the vertices. Let $X_0 := \emptyset$ and, for each $i \in [n]$, let $X_i := \{x_j : j \in [i]\}$.

Claim 6. If there is some $i \in [n]$ such that $d^+(x_i) \geq (1+\varepsilon)np^2$ and $|N^+(x_i) \cap X_{i-1}| \geq C \log n$, then x_i is a Sullivan vertex.

Proof. Assume there is some $i \in [n]$ satisfying the conditions. Let $A \subseteq N^+(x_i) \cap X_{i-1}$ be such that $|A| = C \log n$, and let $B := N^+(A)$. By using the ordering of the vertices and Claim 5 (3), it follows that

$$(1 - o(1))(1+\varepsilon)np^2|A| \leq |A|(d^+(x_i) - |A|) \leq \vec{e}(A,B) \leq (1+\delta)|A||B|p,$$

where the first inequality holds by the lower bound on p. It follows that $|B| \geq (1+\delta)np$. As $B \setminus N^+(x_i) \subseteq N_2^+(x_i)$ and $|N^-(x_i)| \leq (1+\delta)np - |N^+(x_i)|$ by Claim 5 (1), it follows that x_i is a Sullivan vertex. ∎

Next we show that there exist many vertices whose outdegree must be sufficiently large for Claim 6 to apply to them.

Claim 7. For every $\alpha \in [\delta, 1-\delta]$ and every $\beta \in \left[0, \frac{1+\delta}{1+3\delta}\frac{\alpha}{2} - \frac{\delta}{(1+3\delta)\alpha}\right]$, there is no set $A \subseteq V(G)$ of size $|A| \geq \alpha n$ such that $d^+(v) < (1+3\delta)\beta np$ for all $v \in A$.

Proof. Let us assume for a contradiction that there exists a set $A \subseteq V(G)$ of size αn such that $d^+(v) < (1+3\delta)\beta np$ for all $v \in A$. By using the properties of Claim 5, we note that

$$(1-\delta)n^2p/2 \leq \sum_{v\in V(G)} d^+(v) = \sum_{v\in A} d^+(v) + \sum_{v\in V(G)\setminus A} d^+(v)$$
$$< (1+3\delta)\alpha\beta n^2 p + (1+\delta)\binom{(1-\alpha)n}{2}p + (1+\delta)\alpha(1-\alpha)n^2p.$$

By reordering, one can readily verify that this implies that

$$\beta > \frac{1+\delta}{1+3\delta}\frac{\alpha}{2} - \frac{\delta}{(1+3\delta)\alpha},$$

a contradiction. ∎

Set $\alpha := 1-2\delta$ and $\beta := (1+\varepsilon)p/(1+3\delta)$. A standard algebraic manipulation shows that $\beta \leq \frac{1+\delta}{1+3\delta}\frac{\alpha}{2} - \frac{\delta}{(1+3\delta)\alpha}$, so we may apply Claim 7 with these parameters to conclude that

$$\text{for all } i \in [2\delta n] \text{ we have } d^+(x_i) \geq (1+\varepsilon)np^2. \tag{A2}$$

By Claim 6, we may thus assume that $\vec{G}$ satisfies that

$$\text{for all } i \in [2\delta n] \text{ we have } |N^+(x_i) \cap X_{i-1}| \leq C \log n. \tag{A3}$$

Now let $B := X_{2\delta n}$ be partitioned into $B_1 := X_{\delta n}$ and $B_2 := B \setminus B_1$. By Claim 5 (4), we have that $e(B_1, B_2) \geq (1-\delta)|B_1||B_2|p = (1-\delta)\delta^2 n^2 p$. By (A3), for each $v \in B_2$ we have $\vec{e}(v, B_1) \leq C \log n$, and so $\vec{e}(B_1, B_2) \geq (1-2\delta)\delta^2 n^2 p$. It follows by averaging that there is some vertex $y \in B_1$ such that $\vec{e}(y, B_2) \geq (1-2\delta)\delta np$. Let $Y \subseteq N^+(y) \cap B_2$ be a set with $|Y| = (1-2\delta)\delta np$, and let $Z := N^+(Y)$. Observe that $|Z| \geq np^2$ by (A2) and (A3), and so, by Claim 5 (4), $\vec{e}(Y, Z) \leq (1+\delta)|Y||Z|p$. On the other hand, using (A2) and Claim 5 (2), we have that $\vec{e}(Y, Z) \geq (1+\varepsilon-\delta+2\delta^2)np^2|Y|$. Thus, we conclude that $|Z| \geq (1+\delta)np$. As $Z \setminus N^+(y) \subseteq N_2^+(y)$ and $|N^-(y)| \leq (1+\delta)np - |N^+(y)|$ by Claim 5 (1), it follows that y is a Sullivan vertex. ∎