Connectedness of the degree bounded star process

Catherine Greenhill*
Department of Mathematics and Statistics
University of Melbourne VIC 3010
Australia
csg@ms.unimelb.edu.au

Andrzej Ruciński†
Department of Discrete Mathematics
Adam Mickiewicz University
Poznań, Poland
rucinski@amu.edu.pl

Nicholas C. Wormald‡
Department of Mathematics and Statistics
University of Melbourne VIC 3010
Australia
nick@ms.unimelb.edu.au

Abstract

In this paper we consider a random star d-process which begins with n isolated vertices and in each step chooses randomly a vertex of current minimum degree δ, and connects it with $d - \delta$ random vertices of degree less than d. We show that for $d \geq 3$ the resulting final graph is connected with probability $1 - o(1)$, and moreover that for sufficiently large d it is d-connected with probability $1 - o(1)$.

1 Introduction

The standard models of random graphs (see [2]) have a limited potential for applications. This is because in social and natural sciences several processes and structures demand restrictions on the number of connections between objects. Hence, sometimes more suited is the model of a random d-regular graph, where each vertex has degree d.

Its dynamic counterpart is the random d-process introduced in [6], where to the initially empty graph with n vertices, edges are added randomly one by one, chosen uniformly from all available pairs of vertices whose current degrees are less than d. We say that an event holds *asymptotically almost surely* (briefly a.a.s.) if it holds with probability $1 - o(1)$ as $n \to \infty$. It is proved in [6] that a.a.s. the final graph of a random

*Research supported by an Australian Research Council Postdoctoral Fellowship

†Research supported by the KBN grant 2 P03A 032 16. Part of this research was done during the author's visit to the University of Melbourne

‡Research supported by the Australian Research Council

d-process is d-regular (or with one vertex of degree $d-1$ in the case when dn is odd). Moreover, for $d \geq 3$, such a graph is proved in [7] to be a.a.s. connected.

A quicker way to generate a random d-regular graph is, with some luck, to "saturate" vertices one by one. One way of doing this was proposed in [3]. A random *star d-process* introduced in [3] begins with n isolated vertices and in each step chooses randomly a vertex of current minimum degree δ, and connects it with $d-\delta$ random vertices of degree less than d. It is proved in [5] that a.a.s. the final graph is, indeed, d-regular (provided dn is even). In this paper we prove that a.a.s. it is connected for all $d \geq 3$ and d-connected for sufficiently large d (cf. Theorem 2 in Section 2 and Theorem 3 in Section 3).

The situation for a random star 2-process is quite different. When $d = 2$, the final graph is connected if and only if it is a Hamilton cycle, in which case it is also 2-connected. Robalewska [4, Theorem 3.2] showed that the probability of this event is asymptotically $O(n^{-1/2})$.

It is worthwhile to mention that the 'static', uniform random d-regular graph is a.a.s. d-connected for $d \geq 3$ (see [8]). Although the three models of random regular graphs (static, d-process, star d-process) are not asymptotically equivalent, it is believed that they all are contiguous (see, e.g., [2] for the definition of contiguity).

The paper is organized as follows. In Sections 2 and 3, we prove respectively, Theorems 2 and 3, the former establishing the connectedness and the latter – d-connectedness of a random star d-process. The last section contains a proof of a deterministic result needed for the proof of Theorem 3. In the remainder of this introductory section we give some definitions and facts to be used later.

First, we describe the star d-process formally. Fix $n > d \geq 3$, and let G_0 be the empty graph on n vertices, and G_t be the graph created at step t of the star d-process. To obtain G_{t+1} from G_t, in step $t+1$ we proceed by choosing a vertex v of minimum degree δ_t in G_t uniformly at random, and choosing $d - \delta_t$ *active* vertices (that is those with degree less than d) uniformly at random from G_t to be the neighbours of v in G_{t+1}.

If at some time t_0 there are not more than $d-\delta_{t_0}$ active vertices left, then any further transition is impossible, and we say that the process arrived at its final stage and call the graph constructed so far, that is, G_{t_0}, *the final graph* of the process. (It is possible that some active vertex of the final graph, but not of *minimum* degree, can be made saturated. The rules of the star d-process do not allow that, which, in view of Theorem 1 below, is not essential, at least asymptotically. Clearly, the whole process takes less than n steps, but the actual length may vary. For convenience, we artificially extend its duration beyond the final stage, by defining $G_t = G_{t_0}$ for $t = t_0 + 1, \ldots, n$.

While the minimum degree δ_t of G_t is i, the process is said to be in phase i. Define T_i as the first time t such that $\delta_t > i$; if no such t exists, set $T_i = \infty$. So, the stopping times $T_0, \ldots, T_{d-2}$ divide the entire process into d phases (some of the final ones can be empty). Note that at any time the set of active vertices is independent, and that if $T_{d-2} < \infty$ then the last phase is nonempty and consists of adding (sequentially) a random matching saturating all but at most one active vertices. Thus, in such a case, the final graph has precisely $\lfloor dn/2 \rfloor$ edges, and for dn even it is d-regular.

It is easy to see that $T_0 \geq n/(d+1)$. In our proofs we will rely upon the following result from [5], which roughly says that with probability very close to 1, each phase lasts some constant times n number of steps, and, in particular, the final graph is d-regular.

Theorem 1 *For fixed $d \geq 3$ there exist constants $0 < s_0 < \cdots < s_{d-2}$ and $c > 0$ such that, a.a.s. for $i = 0, \ldots, d-2$, we have $T_i = s_i n + o(n)$, and the number of active vertices at time T_{d-2} is at least cn.*

Thus, although there are trajectories of the star d-process which end up in a final graph with minimum degree $\delta_{t_0} < d - 1$ (even $\delta_{t_0} = 0$ is possible), it may only happen with very small probability.

By Theorem 1, there exists a positive constant c such that a.a.s. there are at least cn active vertices remaining by the end of phase $d - 2$. Fixing one such c, let $T(c)$ be the first time t for which the number of active vertices in G_t is less than cn; if no such t exists, set $T(c) = \infty$. Further, let $T = \min\{T(c), T_{d-2}\}$. By the choice of c,

$$\mathbf{P}(T = T_{d-2}) = 1 - o(1). \tag{1}$$

In our proofs we will often condition on the event $\mathcal{T}_t = \{T > t\}$.

Throughout we assume that dn is even, though our methods would work if dn were odd, except that in Theorem 3 we could only show that the final graph is a.a.s. $(d-1)$-connected (the obvious reason being the presence of a vertex of degree $d-1$). Since dn is assumed to be even, the number of active vertices left at the end of phase $d - 2$ is also even, and the last phase makes up a matching saturating all of them. It is easy to verify that this matching occurs uniformly at random.

For a subgraph H of G_t, denote by $A(H)$ the number of active vertices of G_t contained in H, and call it the *active size* of H. For convenience a connected component will be viewed sometimes as a subgraph and sometimes as a set of (its) vertices. In either case the above definition will apply, that is, a connected component C of G_t has *active size* $A(C)$ equal to the number of active vertices it contains. A *singleton* component is one consisting of an isolated vertex.

2 Connectedness of the star d-process

In this section we prove the following theorem.

Theorem 2 *For fixed $d \geq 3$, the final graph of the random star d-process is a.a.s. connected.*

First we will show that a.a.s., the graph $G_{T_{d-2}}$ created by the process contains no connected component of active size less than d. Then we will show that adding a randomly chosen perfect matching to the remaining active vertices a.a.s. connects the graph.

Say that a connected component C of G_t is *small* if $A(C) \leq d^2$, and say that the addition of a star in step $t + 1$ is a *rare step* if some three vertices of the star belong to the same small component of G_t.

Lemma 1 *With probability $1 - o(1)$, there are no rare steps before the time T.*

Proof. Let $\mathcal{R}_t$ be the event that a rare step occurs at time t, for $t = 1, \ldots, n$. Our goal is to show that

$$\mathbf{P}\left(\bigcup_{t=0}^{n}(\mathcal{R}_t \cap \mathcal{T}_t)\right) = o(1).$$

By the law of total probability,

$$\mathbf{P}(\mathcal{R}_t \cap \mathcal{T}_t) = \sum_{G} \mathbf{P}(\mathcal{R}_t \cap \mathcal{T}_t | G_t = G)\mathbf{P}(G_t = G),$$

where the sum runs over all graphs G with $\mathbf{P}(G_t = G) \neq 0$. Note that the conditional probability vanishes if $A(G) < cn$ and that otherwise it equals $\mathbf{P}(\mathcal{R}_t | G_t = G)$.

Fix a graph G with $A(G) \geq cn$. A rare step occurs if either some three neighbours of the star centre, or the star centre and two of its neighbours, are chosen from the same small component of G_t. Regard the chosen vertices as being selected sequentially, so in the first case there are at most $\binom{d}{3}$ choices for when these three neighbours are selected. The first of these three fixes a small component C, and the probability of choosing the other two vertices also from C is $O(n^{-2})$, since $A(C) \leq d^2$ and there are at least $cn - d$ active vertices. The second case is similar, but with the star centre fixing the small component. Thus $\mathbf{P}(\mathcal{R}_t | G_t = G) = O(n^{-2})$ and consequently $\mathbf{P}(\mathcal{R}_t \cap \mathcal{T}_t) = O(n^{-2})$. Summing over $t = 1, \ldots, n$ yields the bound $O(1/n) = o(1)$ on the probability that some rare step occurs at all before the time T. $\square$

For trajectories of the process with no rare steps and such that $T = T_{d-2}$, a deterministic proof by induction shows that each connected component of $G_{T_{d-2}}$ has at least d active vertices.

Lemma 2 *With probability $1 - o(1)$, all connected components of $G_{T_{d-2}}$ have active size at least d.*

Proof. For $t \geq 0$, let $\mathcal{A}_t$ be the event that $A(C) \geq d$ for all non-singleton connected components C in G_t. Fix a trajectory $(G_0, G_1, \ldots)$ of the star d-process with no rare steps, and for which $T = T_{d-2}$. We will show that it has property $\mathcal{A}_t$ for all $t \leq T$. In particular, it follows that A_T holds. By Lemma 1, equation (1), and the fact that there are no singleton components in $G_{T_{d-2}}$, this will imply Lemma 2.

For technical reasons we have to treat phase 0 somewhat differently. Therefore, we define

$$g(C) = \sum_{w \in C} (d - \deg_{G_t}(w)) \tag{2}$$

and $\mathcal{G}_t$ as the property that $g(C) \geq (d-1)^2 + 1$ for all non-singleton components C in G_t. Note that $\mathcal{G}_t$ implies $\mathcal{A}_t$ and that $\mathcal{G}_0$ holds vacuously. We will first prove by induction on t that $\mathcal{G}_t$ holds for all $0 \leq t \leq T_0$. In particular, A_{T_0} will be satisfied. Then we will complete the proof of Lemma 2 by showing, again by induction on t, that $\mathcal{A}_t$ holds for all $T_0 \le t \le T$.

Phase 0. Fix $0 \le t < T_0$ and assume that $\mathcal{G}_t$ holds. Recall that in phase 0, step $t+1$ consists of choosing an isolated vertex v, and choosing d active vertices of G_t to become the neighbours of v in G_{t+1}. Denoting by C the component of G_{t+1} containing v, if two of these neighbours lie in distinct non-singleton components C_1, C_2 of G, then, since $\mathcal{G}_t$ holds,

$$g(C) \ge g(C_1) + g(C_2) - d \ge (d-1)^2 + 1,$$

and $\mathcal{G}_{t+1}$ holds too. If i of the neighbours lie in a non-singleton component C_1 and $d-i$ are isolated vertices in G then the component C formed in G_{t+1} satisfies

$$g(C) = g(C_1) - i + (d-1)(d-i) = g(C_1) + d(d-1-i).$$

Hence $g(C) \ge g(C_1)$ as long as $0 \le i \le d-1$. Thus the only way in which $\mathcal{G}_{t+1}$ can fail is if all d neighbours of v are chosen from C_1. Still, if $g(C_1) \ge (d-1)^2 + 1 + d$, then $\mathcal{G}_{t+1}$ must hold. So we may assume that, say, $g(C_1) \le d^2$, and hence $A(C_1) \le d^2$ by (2). But, since $d \ge 3$, this would constitute a rare step, contradicting our assumption.

Later phases. Fix $T_0 \le t < T_{d-2}$ and assume that $\mathcal{A}_t$ holds. Let $\delta(G_t) = j$, where $1 \le j \le d-2$. In step $t+1$ a vertex v of degree j is chosen, and $d-j$ active vertices are chosen to become the neighbours of v in G_{t+1}. Let C_1 be the component of G containing v. If any neighbours are chosen from a component of G different from C_1, then the component C of G_{t+1} containing v satisfies

$$A(C) \ge 2d - 1 - (d-j) = d + j - 1 \ge d,$$

and $\mathcal{A}_{t+1}$ holds. Thus the only way in which $\mathcal{A}_{t+1}$ can fail is if all $d-j$ neighbours of v are chosen from C_1. Still, if $A(C_1) \ge 2d$ then $\mathcal{A}_{t+1}$ must hold. Otherwise, since $2d < d^2$ and $d - j \ge 2$, we would get a rare step – a contradiction. $\square$

Since $d \ge 3$, Lemma 2 shows that a.a.s. there are no components of active size less than 3 in $G_{T_{d-2}}$. By (1), the set S of active vertices in $G_{T_{d-2}}$ has size at least cn. Recall that the edges added throughout phase $d-1$ form a uniformly chosen perfect matching M on the set of active vertices in $G_{T_{d-2}}$. We will next show that adding this matching a.a.s. connects the graph. The following combinatorial fact will be useful.

Lemma 3 *For positive reals s, r and m, let $s = s_1 + \cdots + s_p$ be an ordered partition of s, with $m \le \min s_i \le r \le s$. Let $\mathcal{I}$ be the set of $I \subseteq [p]$ such that $\sum_{i \in I} s_i = r$. Then*

$$|\mathcal{I}| \le \binom{\lfloor s/m \rfloor}{\lfloor r/m \rfloor}.$$

Proof. By complementation, we can assume that $r \le s/2$. Note that $\mathcal{I}$ is a Sperner system of subsets of $[p]$, that is, no set in $\mathcal{I}$ contains another. Also note that $|I| \le \lfloor r/m \rfloor$ for each $I \in \mathcal{I}$.

First suppose that $\lfloor r/m \rfloor \le \lfloor p/2 \rfloor$. Then by the LYM inequality (c.f. [1])

$$|\mathcal{I}| \le \binom{p}{\lfloor r/m \rfloor} \le \binom{\lfloor s/m \rfloor}{\lfloor r/m \rfloor}$$

where the second inequality comes from the fact that $p \le \lfloor s/m \rfloor$.

On the other hand, if $\lfloor r/m \rfloor > \lfloor p/2 \rfloor$ then by Sperner's theorem (c.f. [1])

$$|\mathcal{I}| \le \binom{p}{\lfloor p/2 \rfloor} \le \binom{\lfloor s/m \rfloor}{\lfloor p/2 \rfloor} \le \binom{\lfloor s/m \rfloor}{\lfloor r/m \rfloor}$$

where the last inequality comes from $\lfloor r/m \rfloor \le \frac{1}{2} \lfloor s/m \rfloor$. $\square$

Finally, we show that adding the randomly chosen perfect matching a.a.s. connects the graph. Here S denotes the set of active vertices in $G_{T_{d-2}}$, which, by (1), a.a.s. contains at least cn vertices (hence $|S| \to \infty$ as $n \to \infty$). Furthermore, p stands for the number of connected components of $G_{T_{d-2}}$, and $S_1, \dots, S_p$ represent their active parts.

Lemma 4 *Let $S_1, \dots, S_p$ be a partition of a set S, with $s = |S|$ even and with $|S_i| \ge 3$ for all i. Let M be a perfect matching on S chosen uniformly at random. Form a graph H_M whose vertices are the sets S_i, with an edge from S_i to S_j whenever an edge of M joins an element of S_i to one of S_j. Then*

$$\lim_{s \to \infty} \mathbf{P}(H_M \text{ is connected}) = 1 - O(s^{-1}).$$

Proof. Let M be a matching on S chosen uniformly at random. Suppose that H_M is disconnected. Then we can partition S as $L \cup R$, where L and R are both unions of sets S_i, such that no edge of M joins an element of L to an element of R. Let $r = |R|$. Then r is even, and $r \ne 2$, since $|S_i| \ge 3$. So, without loss of generality, $4 \le r \le s/2$. Let $m(r)$ denote the number of perfect matchings of r points, r even. Note that from Stirling's formula

$$m(r) = \frac{r!}{2^{r/2}(r/2)!} \sim \sqrt{2} \left(\frac{r}{e}\right)^{r/2}$$

as $r \to \infty$, and hence $m(r) = \Theta((r/e)^{r/2})$ uniformly over all even r (meaning that it is bounded above and below by positive constants times the asymptotic expression). Therefore, setting

$$f(r) = \sqrt{\frac{r^r (s-r)^{s-r}}{s^s}},$$

the probability that no matching edge joins an element of L to an element of R is

$$\frac{m(r)m(s-r)}{m(s)} = O(f(r)).$$

By Lemma 3, there are at most $\binom{\lfloor s/3\rfloor}{\lfloor r/3\rfloor}$ ways that R can be chosen to contain r elements of S. Therefore, denoting $\lfloor s/3\rfloor$ by s_3 and $\lfloor r/3\rfloor$ by r_3 (and ignoring the condition that r is even)

$$\begin{aligned}\mathbf{P}(H_M \text{ is disconnected}) &= O(1)\sum_{r=4}^{s/2}\binom{s_3}{r_3} f(r)\\ &= O(1)\sum_{r=4}^{s/2}\left(\frac{es_3}{r_3}\right)^{r_3} f(r)\\ &= O(1)\sum_{r=4}^{11} s^{r_3-r/2} + O(1)\sum_{r=12}^{s/2} F(r), \qquad (3)\end{aligned}$$

where

$$F(r) = \left(\frac{es}{r}\right)^{r/3} \times f(r) = \left(\frac{es}{r}\right)^{r/3}\sqrt{\frac{r^r(s-r)^{s-r}}{s^s}}.$$

Note that the terms in the first summation in (3) are $O(s^{-1})$ (This can be verified separately for the first two, and then note that for $r \geq 6$, $r_3 - r/2 \leq r/3 - r/2 \leq -1$.) So the first summation is $O(s^{-1})$. We now seek a bound on $\sum_{r=12}^{s/2} F(r)$. The second derivative of $\log F$ is

$$\frac{1}{6r} + \frac{1}{2(s-r)},$$

which is positive for all r. So, F is log-convex, and it is enough to consider only the extremes of the range of r. We have as above that $F(12) = O(s^{12/3-12/2}) = O(s^{-2})$, and

$$F(s/2) = \frac{(2e)^{s/6}}{2^{s/2}} = \left(\frac{e}{4}\right)^{s/6} = O(s^{-2}).$$

Thus $F(r) = O(s^{-2})$ uniformly for all $12 \leq r \leq s/2$ by the log-convexity of F. Therefore $\sum_{r=12}^{s/2} F(r) = O(s^{-1})$ and the lemma follows from (3). ∎

3 Higher connectivity for large d

The aim of this section is to prove the following sharpening of Theorem 2 for sufficiently large degree d.

Theorem 3 *For fixed d large enough, the final graph of the random star d-process is a.a.s. d-connected.*

We do not make any attempt to determine the minimum d for which this argument works, as there are several places where the argument cannot work for d less than 5 say, without major changes. On the other hand, we believe the result is true for all $d \geq 3$.

Let $\delta(H)$ denote the minimum degree of the graph H. Note that in every graph H there is a set S of vertices of size $\delta(H)$ such that $H \setminus S$ contains a singleton component. Say that a graph H is *monosingular* if for all sets $S \subseteq [n]$ with $|S| = \delta(H)$, at most one component in $H \setminus S$ is a singleton. Given a subgraph H of G, for each $S \subseteq [n]$ let $C_1(S), \ldots, C_{j_S}(S)$ be the non-singleton, connected components of $H \setminus S$. We say that H is *linked* in G if for every S of size $\delta(H)$, and every partition $[j_S] = I \cup J$, where I and J are nonempty, there is an edge of G which joins $\bigcup_{i \in I} C_i(S)$ with $\bigcup_{i \in J} C_i(S)$.

The proof of Theorem 3 is based on the following simple observation.

Lemma 5 *If a d-regular graph G contains a monosingular spanning subgraph H, with $\delta(H) = d - 1$, which is linked in G, then G is d-connected.*

Proof. Let S be a $(d-1)$-subset of $[n]$ and $[n] \setminus S = A \cup B$. We will argue that there must be an edge from A to B in G. We are done if there is such an edge already in H. Otherwise, since H is linked in G, there is an $A - B$ edge in G, unless A or B is a union of singleton components of $H \setminus S$. But then, as H is monosingular, either A or B consists of a single vertex. Say, $A = \{v\}$. However, v has d neighbours in G and, since $|S| = d - 1$, at least one of them must belong to B. □

Recall that T_{d-2} is the end of phase $d-2$ as defined in the Introduction, and that, by (1), a.a.s. there are at least cn active vertices remaining in $G_{T_{d-2}}$, for some fixed positive constant c.

Let γ be a fixed positive constant for which the ratio γ/c is sufficiently small (in a sense to be made precise at a point in the argument below). We say that a set $S \subseteq [n]$ is *good* if all non-singleton components of $G_{T_{d-2}} \setminus S$ have each size at least γn and contain each at least $\log^2 n$ active vertices of $G_{T_{d-2}}$. We show later that a.a.s. $G_{T_{d-2}}$ is monosingular (see Lemma 7), and all $(d-1)$-vertex sets of $G_{T_{d-2}}$ are good.

Recall that phase $d-1$ consists of adding a randomly chosen perfect matching M of the vertices of degree $d-1$ in $G_{T_{d-2}}$. We next show that if all $(d-1)$-sets of $G_{T_{d-2}}$ are good, then M a.a.s. makes $G_{T_{d-2}}$ linked in the final graph, which then, due to Lemma 5, becomes d-connected.

Lemma 6 *Suppose that all vertex sets S in $G_{T_{d-2}}$ of size $d-1$ are good. Then a.a.s. $G_{T_{d-2}}$ is linked in the final graph.*

Proof. Since we condition here on the event that all vertex sets S in $G_{T_{d-2}}$ of size $d-1$ are good, we may argue for a fixed graph $H = G_{T_{d-2}}$. For each (good) $(d-1)$-set S of H let $C_1(S), \ldots, C_{j_S}(S)$ be the non-singleton, connected components of $H \setminus S$. Note that $j_S < 1/\gamma$. We will prove that, given S and a partition $[j_S] = I \cup J$, the probability that no edge of M joins $A = \bigcup_{i \in I} C_i(S)$ with $B = \bigcup_{i \in J} C_i(S)$ is $o(n^{-d+1})$. As there are less than n^{d-1} sets S and less than $2^{1/\gamma} = O(1)$ choices of I and J, this will complete the proof.

Suppose that H contains s active vertices. Let a be the number of active vertices in A, and b the number in B. Assume $a \le b$. Because S is good, we have $a \ge \log^2 n$, and note that also $a + b \le s$. We extend the definition of $m(i)$ to be $m(i-1)$ in the case that i is odd. Similar to the proof of Lemma 4, the probability that no matching edge joins an active vertex in A to an active vertex in B is at most

$$n^{d-1}\frac{m(a)m(b)}{m(s)} = O(n^{d-1})\sqrt{\frac{a^a b^b}{s^s}} = O(n^{d-1})(a/s)^{a/2} = O(n^{d-1})2^{-\log^2 n/2} = o(n^{-d+1}),$$

where the factor n^{d-1} accounts for choosing the matching edges to S. □

It remains to show that a.a.s. all sets S in $G_{T_{d-2}}$ of size $d-1$ are good and $G_{T_{d-2}}$ is monosingular. We first prove that a.a.s. there are no $(d-1)$-sets in $G_{T_{d-2}}$ which isolate a relatively small subgraph other than a singleton.

Lemma 7 *With probability $1 - o(1)$, no $(d-1)$-element set $S \subseteq [n]$ isolates in $G_{T_{d-2}}$ a subgraph containing strictly between 1 and γn vertices. In particular, a.a.s $G_{T_{d-2}}$ is monosingular.*

Proof. In the bulk of this proof, we argue about the process for $t < T$ (defined in Section 1), which permits us essentially to assume that such G_t has at least cn active vertices. To justify this formally requires a little attention, and eventually uses (1).

Fix a $(d-1)$-element set S and an ℓ-element set C, $2 \le \ell < \gamma n$. We will show that the probability that C is a union of connected components of $G_{T_{d-2}} \setminus S$ is extremely small, conditional upon $T = T_{d-2}$. (Note that we do not demand that C forms a connected component.) Then later we will sum over all S and C.

For C to exist as above, all edges incident with vertices in C join to vertices in $C \cup S$. To explain the analysis, we redefine the star process to include fractional times. In between G_t and G_{t+1}, introduce a sequence of graphs defined as follows. For $j = d - \delta(G_t) + 1$, define $G_{t+1/j}$ to be the same as G_t but with a distinguished vertex v of degree $d - j + 1$ randomly chosen. Then form $G_{t+2/j}$ by adding an edge from v to a randomly chosen active vertex. Then repeat for $G_{t+3/j}$, and so on until G_{t+1} is formed as required, with the random star of $j-1$ arms, centred at v. During these non-integer times, we refer to v as the *current* star centre. At integer times, the star is complete and there is no current centre. All previously defined stopping times (T_i, T, etc.) remain with respect to integer times only, but the event $\mathcal{T}_t = \{T > t\}$ is still meaningful for fractional t. Note, however, that now $\mathcal{T}_t$ implies only that $A(G_t) \ge cn - d + 1$.

We first demonstrate our argument in a simplified version and obtain a bound weaker than we need. For a fractional time t, and $k \ge 1$, define $\mathcal{A}_k^t$ to be the event that t is precisely the kth fractional time at which an edge is added with the current star centre in C. Also define $\mathcal{B}_k^t$ to be the event that the edge added at fractional time t is the kth edge with current star centre in C, and moreover joins to a vertex in $C \cup S$. Then

$$\mathbf{P}(\mathcal{B}_k^t \mid \mathcal{A}_k^t \cap \mathcal{T}_t) \le \frac{\ell + d - 2}{cn - d} = p \tag{4}$$

since there are at least $cn - d$ active vertices in G_t (other than v) when $t < T$, and at most $l + d - 2$ in $S \cup C \setminus \{v\}$.

Define $\mathcal{B}_k = \mathcal{B}_k(S, C) = \bigcup_{t<T} \mathcal{B}_k^t$, that is, the event that the kth edge with current star centre in C has its other end in $C \cup S$ and occurs before time T. Since $\mathcal{B}_k \subseteq \bigcup_t (\mathcal{A}_k^t \cap \mathcal{T}_t)$, we have, with p as in (4),

$$\mathbf{P}(\mathcal{B}_k) = \sum_t \mathbf{P}(\mathcal{B}_k^t \mid \mathcal{A}_k^t \cup \mathcal{T}_t)\mathbf{P}(\mathcal{A}_k^t \cap \mathcal{T}_t) \le p \sum_t \mathbf{P}(\mathcal{A}_k^t) \le p.$$

We next show that for any $m \ge 1$,

$$\mathbf{P}\left(\bigcap_{k=1}^{m} \mathcal{B}_k\right) \le p^m. \tag{5}$$

As in (4),

$$\mathbf{P}(\mathcal{B}_2^t \mid \mathcal{B}_1 \cap \mathcal{A}_2^t \cap \mathcal{T}_t) \le p$$

and so

$$\mathbf{P}(\mathcal{B}_1 \cap \mathcal{B}_2) = \sum_t \mathbf{P}(\mathcal{B}_2^t \mid \mathcal{B}_1 \cap \mathcal{A}_2^t \cap \mathcal{T}_t)\mathbf{P}(\mathcal{B}_1 \cap \mathcal{A}_2^t \cap \mathcal{T}_t) \le p\mathbf{P}(\mathcal{B}_1) \le p^2.$$

Iterating this inequality gives (5).

Unfortunately, there may be an insufficient number of star centres chosen in C during the process to yield a large enough value of m for (5) to be useful. So we now strengthen this argument to include the fractional times when edges join a current star centre in S to a vertex in C. Note that there can be at most $d(d-1)$ such times. If at any such time, the next edge added joins the current star centre to a vertex in C, we call this a *rare* event. We will now focus on those times when the star centre in S produces a rare event.

Given $R \subseteq \{1, \ldots, d(d-1)\}$, we can make a slight modification of the argument above, redefining $\mathcal{A}_k^t$ to be the event that t is precisely the kth fractional time at which the current star centre is either in C, or in S for the ith time, for some $i \in R$. Also $\mathcal{B}_k^t$ becomes the sub-event of $\mathcal{A}_k^t$ in which the edge added to G_t has both ends in $C \cup S$ and *at least one* end in C. The result is again (5), but this is for each possible set R. For every deterministic process in which there are m edges with both ends in $C \cup S$ and at least one end in C, there exists a set R corresponding to just those edges coming from a star centre in S. Summing over all possible $D = 2^{d(d-1)}$ sets R, we deduce that the probability that there are precisely m edges occurring before time T, each with an end in C, and such that none of these have an end outside $S \cup C$, is at most Dp^m.

If, as happens a.a.s. $T = T_{d-2}$, then all vertices in C have degree at least $d-1$ in G_T, and so at least $\ell(d-1)/2$ edges have at least one end in C. Hence, the probability that S disconnects C conditional upon $T = T_{d-2}$ can be bounded above by

$$\mathbf{P}\left(\bigcap_{k=1}^{m} \mathcal{B}_k(S, C) \mid T = T_{d-2}\right) \le (D + o(1))p^{\ell(d-1)/2}.$$

Consequently, for each $\ell = 2, \dots, \gamma n$, the expected number of sets of size $d-1$ of G_T disconnecting a component of size ℓ, conditional upon $T = T_{d-2}$, is bounded by summing this over all choices of S and C, resulting in

$$\sum_S \sum_C \mathbf{P}\left(\bigcap_{k=1}^{m} \mathcal{B}_k(S,C) \mid T = T_{d-2}\right) \leq (D+o(1))\binom{n}{d-1}\binom{n}{\ell}\left(\frac{\ell+d}{cn-d}\right)^{\ell(d-1)/2}$$
$$\leq Dn^{d-1}\left(\left(\frac{en}{\ell}\right)\left(\frac{\ell+d}{cn-d}\right)^{(d-1)/2}\right)^{\ell} = Dn^{d-1}\left(\frac{K\ell}{cn}\right)^{\ell(d-3)/2},$$

where K is a suitable constant which depends on c and d only. For example, by bounding $\ell + d \leq 5\ell d/6$ and $cn - d \geq 5cn/6$, one can take $K = (ed)^{\frac{d-1}{d-3}} c^{-\frac{2}{d-3}}$.

Choose $\gamma = c/(2K)$. Summing this over $\ell \geq 3$ and taking d sufficiently large (say $d > 7$) gives the following bound on the probability (conditional upon $T = T_{d-2}$) of existence of a set of size $d-1$ which disconnects in G_T a component of size between 3 and γn, inclusive:

$$n^{d-1}\sum_{\ell=3}^{\gamma n}\left(\frac{K\ell}{cn}\right)^{\ell(d-3)/2} \leq n^{d-1}\sum_{\ell=3}^{\log^2 n}\left(\frac{K\log^2 n}{cn}\right)^{3(d-3)/2} + n^{d-1}\sum_{\ell=\log^2 n}^{\gamma n} 2^{-\frac{d-3}{2}\log^2 n}$$
$$\leq (\log n)^{3(d-3)+2} n^{d-1-3(d-3)/2} + n^d 2^{-\log^2 n} = o(1).$$

Finally, if $\ell = 2$ then the lower bound on m used above can be increased to $m \geq 2d-3$, yielding the estimate

$$Dn^{d-1}\left(\frac{K\ell}{cn}\right)^{2d-3} = O(n^{-d+2}) = o(1)$$

on the expected number of sets S disconnecting a component C of size 2. The lemma follows, since $T = T_{d-2}$ a.a.s. ∎

It remains to prove that, with probability $1 - o(1)$, for all $(d-1)$-sets S, all large components of $G_{T_{d-2}} \setminus S$ have many active vertices. Here "large" means with size at least γn and "many" means at least $\log^2 n$. The remainder of the section is devoted to proving this.

Fix a subset $S \subset [n]$ with $d-1$ vertices. For each time step $t \geq 0$, up until the end of phase $d-2$, we consider the connected components of the graph $G_t \setminus S$. As before, we assume $t < T$, so that there are at least cn active vertices remaining at time t. We will obtain a lower bound on the number of active vertices in each component, relative to the size of the component at that time, and a couple of other parameters including the number of edges from the component to S. Eventually, we will show that, unless something extremely unusual happens, the process does not form any large components with only few active vertices.

To make this precise, we define a component of $G_t \setminus S$ to be *small* if it has active size less than $\log^3 n + d + 1$. For the star added to G_t, define $\mathcal{C}$ to be the set of small components of $G_t \setminus S$ containing vertices of the star, and call the addition of the star to G_t a *rare step with respect to* S if:

(i) the star centre is in a component in $\mathcal{C}$, and

(ii) for some $w > 0$, the number of edges of the star with both ends in components in $\mathcal{C}$ is exactly $w + |\mathcal{C}|$.

Note that the notions of a small component and of a rare step differ from those defined in Section 2.

In the event of such a step, we say that the star centre acquires *a weight* w with respect to S. For a component C of $G_t \setminus S$, we can identify the rare steps that have occurred from time 0 to time t and involved star centres which now belong to C. The *weight* $w(C)$ of C is defined to be the sum of the weights of these vertices.

We next show that rare steps really are rare.

Lemma 8 *With probability* $1 - o(1)$*, for all* $(d-1)$*-element sets* $S \subseteq [n]$ *and all components* C *of* $G_{T_{d-2}} \setminus S$ *we have* $w(C) < d$*.*

Proof. Fix a subset $S \subset [n]$ with $d-1$ vertices and an integer $w > 0$. In view of equation (1), it suffices to show that the probability that some rare steps with respect to S of total weight w have occurred before time T is $O((\log^6 n/n)^w)$. Lemma 8 will follow by taking $w = d$ and summing over the $O(n^{d-1})$ choices for S.

As in the proof of Lemma 7, we compute various probabilities conditioning on $\mathcal{T}_t$. Let $\mathcal{R}_w^t$ be the event that a rare step of weight w occurs at time t. We will bound $\mathbf{P}(\mathcal{R}_w^t \cap \mathcal{T}_t)$ by a bound on $\mathbf{P}(\mathcal{R}_w^t \mid \mathcal{T}_t)$.

To estimate $\mathbf{P}(\mathcal{R}_w^t \mid \mathcal{T}_t)$, suppose that $|\mathcal{C}| = r$. Then the corresponding star has some $r + w$ edges involving r small components. We bound the probability of this event as follows. First select $r - 1$ edges, each terminating in a distinct small component not containing the star centre. These edges, together with the star centre, determine the r small components used. Then select $w + 1$ edges, each terminating in an active vertex contained in one of these pre-selected components. Since there are at least cn active vertices available in the entire graph G_t, the probability that this rare step occurs at a given time t before T is

$$\mathbf{P}(\mathcal{R}_w^t \cap \mathcal{T}_t) \leq \mathbf{P}(\mathcal{R}_w^t \mid \mathcal{T}_t) = O(\log^{3(w+1)} n/n^{w+1}).$$

For a sequence of rare steps, we can use conditioning as in the proof of Lemma 7 just after (5). So the probability that i rare steps of weights $w_1, \ldots, w_i$ occur at times $t_1 < \ldots < t_i < T$ is, with $w = \sum_\ell w_\ell$,

$$\mathbf{P}\left(\bigcap_{\ell=1}^{i} \mathcal{R}_{w_\ell}^{t_\ell} \mid \mathcal{T}_{t_i}\right) = O\left(\prod_{\ell=1}^{i} \left(\frac{\log^3 n}{n}\right)^{w_\ell+1}\right) = O\left(\left(\frac{\log^3 n}{n}\right)^{w+i}\right).$$

Hence, summing over $i \le w$, over the $O(n^i)$ choices of $t_1, \dots, t_i$, and the bounded number of partitions $w_1 + \ldots + w_i$ of w, the probability that some rare steps with respect to S of total weight w have occurred before time T is

$$\sum_{i=1}^{w} \sum_{t_1,\dots,t_i} \sum_{w_1+\cdots+w_i=w} \mathbf{P}\left(\bigcap_{\ell=1}^{i} \mathcal{R}_{w_\ell}^{t_\ell} \cap \mathcal{I}_{t_i}\right) = O\left(n^i \left(\frac{\log^3 n}{n}\right)^{w+i}\right) = O\left(\left(\frac{\log^6 n}{n}\right)^{w}\right).$$

□

We will show in Section 4 the following deterministic lower bound on the number of active vertices contained in a component. Let $s(C)$ be the number of edges from C to S, let $k(C)$ be the number of vertices in C, and recall that $A(C)$ is the number of active vertices contained in C.

Lemma 9 *For sufficiently large d, consider any deterministic trajectory of the star d-process, and any $t < T$. Then for all $(d-1)$-element sets of vertices S, and all components C of $G_t \setminus S$, we have*

$$A(C) \ge f(C) - 4w(C) - 3s(C),$$

where

$$f(C) = \min\{\log^3 n, \sqrt{k(C) + s(C)}\}.$$

Armed with Lemma 9 we may now quickly finish the proof of Theorem 3.

Proof of Theorem 3. By Lemma 7 and Lemma 8, a.a.s. for every subset $S \subset [n]$ of size $d-1$ and every non-singleton component C of $G_{T_{d-2}} \setminus S$ we have

$$k(C) \ge \gamma n, \quad w(C) < d, \quad \text{and} \quad s(C) \le d^2$$

(the last fact is always true). Hence, by Lemma 9,

$$A(C) \ge f(C) - 4d - 3d^2 \ge \min\{\log^3 n, \sqrt{\gamma n}\} - 4d - 3d^2 \ge \log^2 n$$

which proves that with probability $1 - o(1)$, all $(d-1)$-sets in $G_{T_{d-2}}$ are good.

Combining this with Lemma 6 shows that $G_{T_{d-2}}$ is a.a.s. linked in the final graph of the random star d-process. By Lemma 7, $G_{T_{d-2}}$ is also a.a.s. monosingular. Thus, Lemma 5 with $H = G_{T_{d-2}}$ yields that the final graph is a.a.s. d-connected, when d is large enough, thereby proving Theorem 3. □

4 A deterministic result: proof of Lemma 9

To complete the proof of Theorem 3 it only remains to establish Lemma 9. We will use the following simple inequalities, first of which is easily proved by induction on r and the other is elementary (the proofs are omitted).

Lemma 10 *(a) If $r \geq 1$ and $a_1, \ldots, a_r \geq d$ where d is sufficiently large, then*

$$\sqrt{\sum_{\ell=1}^{r} a_\ell} \leq \sum_{\ell=1}^{r} \sqrt{a_\ell} - 4(r-1).$$

(b) If $a \geq 4$ and $0 \leq a_0 \leq a$ then

$$\sqrt{a - a_0} \geq \sqrt{a} - \frac{a_0}{2}.$$

Proof of Lemma 9. Fix a $(d-1)$-element set S. We proceed by induction on t. The inequality in Lemma 9 holds at $t = 0$ when all components are isolated vertices, since they all have active size 1 and $s = w = 0$.

Next take any $t \geq 0$ and assume that the inequality in Lemma 9 holds at time t. Consider the star added to G_t. Let us denote it by X, and its centre by x. Suppose firstly that $x \in S$. Let C' be a component of $G_t \setminus S$, and let $j \geq 0$ be the number of vertices of the star X contained in C'. Then C' is also a component of $G_{t+1} \setminus S$, but in this step the functions A, s and f may change. So when viewed as a component of $G_{t+1} \setminus S$, we denote it by C. Then

$$A(C) \geq A(C') - j, \qquad s(C) = s(C') + j, \qquad f(C) \leq f(C') + j, \qquad w(C) = w(C'),$$

where of course functions such as $A(C')$ are computed before adding X and those like $A(C)$ are computed after. Therefore

$$f(C) - 4w(C) - 3s(C) \leq f(C') - 4w(C') - 3s(C') - 2j \leq A(C') - 2j \leq A(C),$$

as required.

In the remainder of the proof we consider the case that $x \notin S$. Let C denote the component of $G_{t+1} \setminus S$ containing x.

Suppose that at least one of the components of $G_t \setminus S$ contained in C contains at least $\log^3 n + d + 1$ active vertices (that is is not small). Since X contains at most $d+1$ vertices, at least $\log^3 n$ vertices must remain active in C, and thus $A(C) \geq \log^3 n \geq f(C) - 4w(C) - 3s(C)$, as required. So for the remainder of the proof, we may assume that all components of $G_t \setminus S$ contained in C are small.

Note that for every such component C' there are two possibilities. If C' received a star centre at some time in the past, then it has a vertex of degree d and so $k(C') + s(C') \geq k(C') \geq d + 1$. Otherwise, all edges incident with vertices in C' must join to S, from which it follows that $k(C') = 1$. We then call C' a 1-component. If the only vertex in C' becomes inactive upon the addition of X, and is not x, then it must have been incident with all $d - 1$ vertices of S, and so $k(C') + s(C') = d$. All the other components, that is, 1-components either containing x, or remaining active in G_{t+1}, will be called *spare*.

Suppose that X joins m spare components and r other components $C_1, \ldots, C_r$ of $G_t \setminus S$. It follows that

$$k(C_\ell) + s(C_\ell) \geq d, \tag{6}$$

for each $\ell = 1, \ldots, r$.

We next examine how many vertices can become inactive when X is added. Let w_0 denote the weight of x after this step. By definition,

$$w_0 = \max\{|E(X)| - r - m, 0\} ,$$

so

$$|E(X)| + 1 = |V(X)| \le r + m + w_0 + 1. \tag{7}$$

But in all cases $m - \xi$ of these vertices are 1-components of $G_t \setminus S$ which remain active, where

$$\xi = \begin{cases} 1 & x \text{ is in a 1-component,} \\ 0 & \text{otherwise.} \end{cases}$$

Hence the number of vertices which become inactive when X is added is at most $r + w_0 + 1 + \xi$. Since there are initially $A(C_\ell)$ active vertices in C_ℓ, and one in each of the m spare 1-components, we have

$$A(C) \ge \left(\sum_{\ell=1}^{r} A(C_\ell) \right) + m - r - w_0 - 1 - \xi. \tag{8}$$

We aim to show that the component C of $G_{t+1} \setminus S$ satisfies the inequality in Lemma 9. Define $k_\ell = k(C_\ell)$, $s_\ell = s(C_\ell)$ and $w_\ell = w(C_\ell)$ for $1 \le \ell \le r$, and

$$s_0 = s(C) - \sum_{\ell=1}^{r} s_\ell,$$

and note that

$$m = k(C) - \sum_{\ell=1}^{r} k_\ell$$

and

$$w_0 = w(C) - \sum_{\ell=1}^{r} w_\ell.$$

We examine three cases for the rest of the proof.

Case 1. $r = 0$
In this case we have $k(C) = m$, $w_0 = 0$ and $\xi = 1$, and, by (8), we will be done if

$$m + 3s_0 \ge \sqrt{m + s_0} + 2. \tag{9}$$

This easily holds for $s_0 \ge 1$. Otherwise, we have $m = d + 1$ and (9) is valid for $d \ge 3$.

Case 2. $\sqrt{k_\ell + s_\ell} < \log^3 n$ for all $1 \le \ell \le r$.

Consequently, $f(C_\ell) = \sqrt{k_\ell + s_\ell}$ for $1 \le \ell \le r$. By the induction hypothesis,

$$A(C_\ell) \ge f(C_\ell) - 4w_\ell - 3s_\ell = \sqrt{k_\ell + s_\ell} - 4w_\ell - 3s_\ell$$

for $1 \le \ell \le r$. By (6), Lemma 10 applies, and together with (8) implies

$$\begin{aligned} A(C) \ge \left(\sum_{\ell=1}^{r} \sqrt{k_\ell + s_\ell}\right) &- 4(w(C) - w_0) - 3(s(C) - s_0) + m - r - w_0 - 1 - \xi \\ \ge \sqrt{k(C) + s(C) - m - s_0} &- 4w(C) - 3s(C) + 3r + 3s_0 + 3w_0 + m - 5 - \xi \\ \ge \sqrt{k(C) + s(C)} &- 4w(C) - 3s(C) + 3r + \frac{5s_0}{2} + 3w_0 + \frac{m}{2} - 5 - \xi \\ \ge f(C) &- 4w(C) - 3s(C) + 3r + \frac{5s_0}{2} + 3w_0 + \frac{m}{2} - 5 - \xi. \end{aligned}$$

Hence the required inequality in Lemma 9 follows provided

$$6r + 5s_0 + 6w_0 + m \ge 10 + 2\xi. \tag{10}$$

Note that s_0 is *at least* the number of edges in G_t from the m spare 1-components to S. But for these components, *all* incident edges go to S. So, if $m \ge 1$, we have

$$s_0 \ge \delta(G_t) = d_{G_t}(x) \ge d - (r + m + w_0) \tag{11}$$

by (7), because $|E(X)| = d - d_{G_t}(x)$. Hence

$$6r + 5s_0 + 6w_0 + m \ge 5r + d + 4s_0 + 5w_0 \ge 5 + d,$$

which implies (10) when $d \ge 7$. So we may assume $m = 0$, which gives $2 \le |E(X)| \le r + w_0$ by (7), and by the fact that $\delta(G_t) \le d - 2$. This, again, implies (10).

Case 3. $\sqrt{k_1 + s_1} \ge \log^3 n$.

Here $f(C_1) = \log^3 n$, and for $2 \le \ell \le r$ we have by the induction hypothesis and (6) that $A(C_\ell) \ge f(C_\ell) - 4w(C_\ell) - 3s_\ell \ge \sqrt{d} - 4w(C_\ell) - 3s_\ell$, and so (8) implies

$$A(C) \ge \log^3 n - 3(s(C) - s_0) - 4(w(C) - w_0) + (r - 1)\sqrt{d} + m - r - w_0 - 1 - \xi.$$

So, we are done if

$$(r - 1)\sqrt{d} - r + 3s_0 + 3w_0 + m \ge 1 + \xi. \tag{12}$$

Since the right hand side of (12) is at most 2, it is immediately satisfied if $r \ge 2$ and $d \ge 16$. For $r = 1$, inequality (12) follows unless $s_0 = w_0 = 0$. Thus, we are left with having to show $m \ge 2 + \xi$ under the conditions $r = 1$ and $s_0 = w_0 = 0$. If $m \ge 3$, this is trivially true. If $m = 1$ or $m = 2$, we note that (11) is still valid which yields a contradiction for $d \ge 4$. If $m = 0$, then a component of $G_t \setminus S$ contains the entire star X, which contradicts the assumption $w_0 = 0$, as X has at least two edges. $\square$

References

[1] B. Bollobás, *Combinatorics*. Cambridge University Press, Cambridge, 1986.

[2] S. Janson, T. Łuczak, A. Ruciński, *Random Graphs*. Wiley, New York, 2000.

[3] H. Robalewska, Stochastic processes and random graphs, PhD. Thesis, University of Melbourne, 1996.

[4] H. Robalewska, Random graphs generated by the star 2-process, *Random Structures and Algorithms* **11** (1997), pp. 125–149.

[5] H. Robalewska and N. C. Wormald, Random star d-processes, *Combinatorics, Probability and Computing* **9** (2000), pp. 33–43.

[6] A. Ruciński and N. C. Wormald, Random graph processes with degree restrictions, *Combinatorics, Probability and Computing* **1** (1992), pp. 169–180.

[7] A. Ruciński and N. C. Wormald, The probability that a random d-process results in a connected graph, *Journal of the Australian Mathematical Society* **72** (2002), pp. 67–85.

[8] N. C. Wormald, The asymptotic connectivity of labelled regular graphs, *Journal of Combinatorial Theory, Series B* **31** (1981) pp. 156–167.