Birth Control for Giants

Joel Spencer
Courant Institute
New York, NY, USA
spencer@cims.nyu.edu

Nicholas Wormald*
Department of Combinatorics and Optimization
University of Waterloo
Waterloo ON, Canada
nwormald@uwaterloo.ca

Abstract

The standard Erdős-Renyi model of random graphs begins with n isolated vertices, and at each round a random edge is added. Parametrizing $\frac{n}{2}$ rounds as one time unit, a phase transition occurs at $t = 1$ when a giant component (one of size constant time n) first appears. Under the influence of statistical mechanics, the investigation of related phase transitions has become an important topic in random graph theory.

We define a broad class of graph evolutions in which at each round one chooses one of two random edges $\{v_1, v_2\}, \{v_3, v_4\}$ to add to the graph. The selection is made by examining the sizes of the components of the four vertices. We consider the susceptibility $S(t)$ at time t, being the expected component size of a uniformly chosen vertex. The expected change in $S(t)$ is found which produces in the limit a differential equation for $S(t)$. There is a critical time t_c so that $S(t) \to \infty$ as t approaches t_c from below. We show that the discrete random process asymptotically follows the differential equation for all subcritical $t < t_c$. Employing classic results of Crámer on branching processes we show that the component sizes of the graph in the subcritical regime have an exponential tail. In particular, the largest component is only logarithmic in size. In the supercritical regime $t > t_c$ we show the existence of a giant component, so that $t = t_c$ may be fairly considered a phase transition.

Computer aided solutions to the possible differential equations for susceptibility allow us to establish lower bounds on the extent to which we can either delay or accelerate the birth of the giant component.

1 The Achlioptas Problem and Process

1.1 Introduction

We consider a problem of Dimitris Achlioptas that has received considerable attention. Paul is given n vertices and a graph G on those vertices that will change with time. Initially G has no edges. Each round two edges of K_n, call them $e_1 = \{v_1, v_2\}$ and $e_2 = \{v_3, v_4\}$ are generated independently and uniformly at random. Paul must select one of those edges and add it to G. Paul's object is to avoid creating a giant component, a component of size $\Omega(n)$, for as long as possible. For us, *size* always denotes number of vertices.

It shall be convenient to parametrize the number of rounds m by $m = t\frac{n}{2}$. We shall think of t as the "time" of the process. For any $t < 1$ Paul can succeed by simply always taking the first edge – the graph then selected is the usual Erdős-Rényi random graph which has component sizes

*Research supported by the Australian Research Council and the Canada Research Chairs Program. Research partly carried out while the author was at the Department of Mathematics and Statistics, University of Melbourne.

$O(\ln n)$. (Our statements about Paul's achievements are all with probability tending to 1 as $n \to \infty$.) Here we shall give an algorithm such that at $t = 1.6587$ (and probably at $t = 1.7799$, if some less rigorous computations can be trusted) the component sizes will be $O(\ln n)$. On the other side, by [3, Theorem 1(d)], for $t = 1.9645$ Paul cannot succeed (with probability tending to 1).

While the Achlioptas problem was our original motivation we have become intrigued by what we shall call an Achlioptas process. Fix any algorithm that determines which edge Paul shall select. Let G_m denote the graph after m rounds. Then $G_0, G_1, \ldots$ forms a random graph process, that evolves from the empty graph to a graph with a giant component and, of course, beyond. For a class of algorithms we shall be able to analyze this process. There are interesting analogies to the well-studied Erdős-Rényi evolution – some of which we can prove and others of which remain conjectures.

For convenience of exposition we imagine in each round that the vertices v_1, v_2, v_3 and v_4 are all chosen uniformly and independently from the vertex set. Thus we allow the possibility that $v_1 = v_2$ or $v_3 = v_4$ as well as the possibility that one or both edges is already in the graph. So Paul may create a loop or multiple edge. However, we also show that these effects are asymptotically negligible and that our results hold if the choice of (v_1, v_2, v_3, v_4) is restricted in a given round so as to avoid these possibilities (and is otherwise uniform).

Our method also permits us to analyze algorithms which attempt the opposite of the original question: how much can Paul accelerate the birth of the giant by judiciously selecting edges? We show for instance that if Paul wants the giant to be born before $t = 0.6671$, he can succeed. Flaxman et al. [5] have shown that this is true when $t = 0.985$ (see Bohman and Kravitz [4] for a stronger result) and that it cannot be true when $t = 0.5015$. To approach both versions of the question, we analyze what we shall call *Bounded Size Algorithms.* Let K be a fixed positive integer. We consider algorithms in which Paul's choice of edge depends only on the sizes of the components containing v_1, v_2, v_3, v_4 and where, furthermore, all components of size greater than K are treated the same. Before describing these algorithms formally we give several definitions.

1.2 Notations and Definitions

We set

$$\Omega = \{1, \ldots, K, \omega\} \tag{1}$$

where ω is a special symbol that will be used to represent (informally) "bigger than K." Let G be a graph on n vertices. We let $C_1, \ldots, C_u$ denote the components of G. For $v \in G$ we let $C(v)$ denote the component of G containing v. Then $|C(v)|$, as usual, is the size of that component. We define $c(v) \in \Omega$ by setting

$$c(v) = \begin{cases} |C(v)| & \text{if } |C(v)| \le K, \\ \omega & \text{if } |C(v)| > K. \end{cases} \tag{2}$$

For notational convenience we set

$$\vec{v} = (v_1, v_2, v_3, v_4) \text{ and } c(\vec{v}) = (c(v_1), c(v_2), c(v_3), c(v_4)).$$

For $i \in \Omega$ we define

$$x_i(G) = \frac{1}{n}|\{v : c(v) = i\}|, \tag{3}$$

the proportion of vertices of G in components of "size" i. We define the *susceptibility* $S(G)$ by

$$S(G) = \frac{1}{n}\sum_v |C(v)| = \frac{1}{n}\sum_{i=1}^{u} |C_i|^2. \tag{4}$$

We define the *essential susceptibility*, denoted $S_\omega(G)$, by

$$S_\omega(G) = \frac{1}{n} \sum_{c(v)=\omega} |C(v)| = \frac{1}{n} \sum_{|C_i|>K} |C_i|^2. \tag{5}$$

The contribution to $S(G)$ from components of size at most K is determined by the $x_i(G)$. This gives the relation:

$$S_\omega(G) = S(G) - \sum_{i=1}^{K} i x_i(G). \tag{6}$$

We define

$$\nabla(G) = \frac{1}{n^2} \sum_{i=1}^{u} |C_i|^4. \tag{7}$$

We say that a graph G on n vertices has a K, c *component tail* if

$$\frac{1}{n} |\{v : |C(v)| \geq s\}| \leq K e^{-cs}$$

for all positive integers s. Let c' satisfy $cc' > 1$. We note that the K, c component tail condition implies, for n sufficiently large, that

$$\max_v |C(v)| < c' \ln n. \tag{8}$$

We shall further say that a nonnegative integer valued random variable X has a K, c *tail* if for all nonnegative integers s

$$\Pr[X \geq s] \leq K e^{-cs}.$$

Now we describe the bounded set algorithms formally. For

$$F \subseteq \Omega^4 \tag{9}$$

we define the F-algorithm as follows. (Here F stands for "first", denoting conditions under which the first edge is the one chosen by Paul.) Suppose in a given round the current value of the graph is G and vertices $\vec{v}$ are given. Then the new value of the graph, call it G^+, is given by

$$G^+ = \begin{cases} G \cup \{v_1, v_2\} & \text{if } c(\vec{v}) \in F, \\ G \cup \{v_3, v_4\} & \text{if } c(\vec{v}) \notin F. \end{cases}$$

We call a round *redundant* if the added edge has both vertices lying in the same component of G. This includes the cases when the two vertices are identical and when they are already adjacent in G.

`A Basic Example:` Let $K = 1$ so that $\Omega = \{1, \omega\}$ and set $F = \{(1, 1, \alpha, \beta) : \alpha, \beta \in \Omega\}$. We can describe this process in words: If v_1, v_2 are both isolated vertices then we add the edge $\{v_1, v_2\}$, otherwise we select the edge $\{v_3, v_4\}$. We shall continue this example throughout this work. We note that this algorithm (actually, a modification of it) was studied [2] by T. Bohman and A. Frieze. They showed that their algorithm did avoid a giant component through $\frac{n}{2}(1 + \epsilon)$ rounds for an absolute positive constant (that they did not try to optimize) ϵ.

`The Erdos-Rényi Evolution:` We may regard the standard Erdős-Rényi evolution as the special case $K = 0$, $\Omega = \{\omega\}$ and $F = \{(\omega, \omega, \omega, \omega)\}$, so that the first edge is always taken.

`Asymptotic Notation:` All asymptotics are as n, the number of vertices, approaches infinity. We say an event occurs **a.a.s.** if the probability of it occuring approaches one as n approaches infinity. A statement such as "$S(G_{nt/2}) = S(t) + o(1)$ a.a.s." should be interpreted as meaning that for any $\epsilon > 0$ the event $|S(G_{nt/2}) - S(t)| < \epsilon$ occurs a.a.s.

1.3 The Main Result

In our statement the functions $x_i(t), S(t)$, which depend on k and F, will be the solution to the differential equation (34,35) developed in Section 2.

Theorem 1.1 *There exists $t_c > 0$ and functions $x_i(t), i \in \Omega$, $S(t)$ such that*

1. *(Points in Small Components) ($i \in \Omega$) $x_i(t)$ is defined for all $t \geq 0$. With probability $1 - o(1)$, for all such t we have $x_i(G_{tn/2}) = x_i(t) + o(1)$.*
2. *(Critical Point) $S(t)$ is defined for $t \in [0, t_c)$, and $\lim_{t \to t_c^-} S(t) = \infty$.*
3. *For any fixed $\epsilon > 0$, with probability $1 - o(1)$, for all $t < t_c - \epsilon$ we have $S(G_{tn/2}) = S(t) + o(1)$.*
4. *(Subcritical Behavior) For all $t < t_c$ with there exist positive K, c so that with probability $1 - o(1)$ $G_{tn/2}$ has a K, c component tail In particular, with $cc' > 1$, $G_{tn/2}$ has all component sizes less than $c' \ln n$.*
5. *(Supercritical Behavior) For all $t > t_c$ there exists positive α so that with probability $1 - o(1)$ $G_{tn/2}$ has a component of size at least αn.*

Note that, in this theorem and similar places, $G_{tn/2}$ denotes $G_{\lfloor tn/2 \rfloor}$ when $tn/2$ is not an integer.

We note here the effect of rounds in which at least one of the two edges v_1v_2 and v_3v_4 forms a loop or multiple edge with one of the edges already in the graph. Call such a round *improper* and the others *proper*, and let us call the process in which improper rounds are forbidden (and, conditional upon this, choices of v_1, v_2, v_3, v_4 are uniformly at random) proper and the unrestricted process improper. It is easy to show that for fixed t, the probability that a given round is improper, conditional upon a given graph G, is $O(1/n)$. Hence, the probability that *all tn* rounds are proper is bounded below by a positive constant. Moreover, the probability of a given trajectory $G_0, G_1, \ldots$ occurring in the proper process is at least the probability that it occurs in the improper process. It follows from these observations that any property a.a.s. true for the improper process must also hold a.a.s. for the proper process. In particular, this applies to all the statements in Theorem 1.1.

Bohman and Kravitz [4] have given an analysis of a family of algorithms which includes some of the Bounded Size Algorithms, essentially, ones in which the decision is made based on looking at v_1v_2 only. Their analysis, which is quite different from ours, gives for those algorithms the critical value $t = t_c$ defined above. In the subcritical region $t < t_c$ their bounds on component sizes are of the form n^α for a positive constant α.

1.4 Analogies to Classical Percolation

The susceptibility $S(G)$ can be regarded as the expected size of the component containing a randomly selected vertex v. As all vertices "look the same" in this random process we may think of $S(G_m)$ as the expected size of the component containing a particular vertex v. In classical percolation on Z^d the susceptibility is denoted by χ (we avoid this notation for graphs for obvious reasons!) and $\chi(p)$ denotes the expected value of $|C(\vec{0})|$, the size of the cluster containing the origin (or any particular vertex). There is a critical value p_c such that $\chi(p) \to \infty$ as p approaches p_c from below. Two deep percolation results examine behavior on both sides of p_c.

• Subcritical Behavior: Fix $p < p_c$. Then the distribution of $|C(\vec{0})|$ decays exponentially. That is, there exist positive K, c so that for all $s \geq 1$

$$\Pr[|C(\vec{0})| \geq s] \leq K e^{-cs}.$$

• Supercritical Behavior: Fix $p > p_c$. Then with probability one there is an infinite component.

The subcritical behavior of the finite process G_i mirrors that of classical (infinite) percolation. Theorem 1.1 states that $|C(v)|$, with v chosen uniformly, has a distribution which decays exponentially. As initially all vertices are identical this implies that for any particular vertex v the distribution of $|C(v)|$ decays exponentially. In subcritical classical percolation the components, while finite, can be arbitrarily large. We feel that "finite" in classical percolation corresponds to $O(\ln n)$ in finite random structures.

The subcritical behavior of the finite process G_i also mirrors that of classical (infinite) percolation. Of course, the finite process cannot contain an infinite component. We feel that "infinite" in classical percolation corresponds to $\Omega(n)$ (i.e., a positive proportion of the vertices) in finite random structures.

2 Expected Change in One Round

Fix the current value G of the graph. Let G^+ be the value after one more round. Here G^+ has a distribution. Our probability space will be the uniform generation of $\vec{v}$. Note that the bounded size algorithms are Markovian so that G^+ is determined by G and $\vec{v}$. We examine the expected change in the graph functions x_i, $i \in \Omega$, and S. For notational convenience we shall set, in this section,

$$x_i := x_i(G); \quad S := S(G); \quad S_\omega := S_\omega(G) \tag{10}$$

and

$$x_i^+ := x_i(G^+); \quad S^+ := S(G^+). \tag{11}$$

Note that the x_i^+, S^+ are now random variables. We shall see that the expected changes $E[x_i^+ - x_i]$ and $E[S^+ - S]$ can *almost* be described as a function of the x_i, $i \in \Omega$, S and S_ω, independent of n. A difficulty will arise with redundant rounds. In those cases $x_i^+ = x_i$ and $S^+ = S$ as the component structure does not change. A key to the analysis will be to separate out the effect of redundant rounds, which we shall think of as an "error" term, and to eventually show that it does not have an asymptotic effect in the subcritical phase. On an intuitive level this is not surprising: when there is no giant component the probability that the edge added joins two vertices already in the same component is $o(1)$. That said, the arguments will not be easy.

For further notational convenience we shall set, in this section,

$$\vec{j} = (j_1, j_2, j_3, j_4). \tag{12}$$

Sums over $\vec{j}$ shall be over all $\vec{j} \in \Omega^4$. We shall naturally split the random choice of $\vec{v}$ into choice of $\vec{v}$ with $c(\vec{v}) = \vec{j}$ and then sum over all $\vec{j}$. As the v_i are chosen independently and uniformly

$$\Pr[c(\vec{v}) = \vec{j}] = x_{j_1} x_{j_2} x_{j_3} x_{j_4}. \tag{13}$$

2.1 Points in Small Components

Consider $i \in \Omega$ and $\vec{j} \in \Omega^4$. We define $\Delta(\vec{j}; i)$ to be one half (a factor that will be convenient later) the change in the number of vertices in components of size i from G to G^+ when $\vec{v}$ has $c(\vec{v}) = \vec{j}$ and the round is not redundant, i.e. the new edge does not join two vertices already in the same component. Using symmetry based on $\vec{j} \in F$ versus $\vec{j} \notin F$ there are four basic cases.

`Case I:` $\vec{j} \in F$, $j_1, j_2 \neq \omega$. Then components of size j_1, j_2 disappear. We have $\Delta(\vec{j}, j_1) = -\frac{1}{2} j_1$ and $\Delta(\vec{j}, j_2) = -\frac{1}{2} j_2$, except that if $j_1 = j_2$ we have $\Delta(\vec{j}, j_1) = -j_1$. A component of size $j_1 + j_2$ is created. When $j_1 + j_2 > K$ we have $\Delta(\vec{j}, \omega) = \frac{1}{2}(j_1 + j_2)$, otherwise $\Delta(\vec{j}, j_1 + j_2) = \frac{1}{2}(j_1 + j_2)$.

`Case II:` $\vec{j} \in F$, $j_1 \neq \omega$, $j_2 = \omega$. Then a large component absorbs a component of size j_1 so that $\Delta(\vec{j}, j_1) = -\frac{1}{2} j_1$ and $\Delta(\vec{j}, \omega) = +\frac{1}{2} j_1$.

`Case III:` $\vec{j} \in F$, $j_1 = \omega$, $j_2 \neq \omega$. Then a large component absorbs a component of size j_2 so that $\Delta(\vec{j}, j_2) = -\frac{1}{2} j_2$ and $\Delta(\vec{j}, \omega) = +\frac{1}{2} j_2$.

`Case IV:` $\vec{j} \in F$, $j_1, j_2 = \omega$. Two large components merge to form an even larger component but this is not reflected in these functions and all $\Delta(\vec{j}, i) = 0$.

`Cases V-VIII:` $\vec{j} \notin F$. The cases in which $\vec{j} \notin F$ are identical with (j_1, j_2) and (j_3, j_4) exchanging roles.

All other values of $\Delta(\vec{j}, i)$ shall be zero, including all i in the case of redundant rounds. Note that in all cases 2Δ is an integer and

$$|\Delta| \leq K, \tag{14}$$

the extreme case being the merger of two components of size K.

We now define random variables

$$x_i^* = x_i + \frac{2}{n}\Delta_i \tag{15}$$

$$e_i = x_i^+ - x_i^* \tag{16}$$

where $\Delta_i = \Delta(\vec{j}, i)$ with $\vec{j}$ random. Here e_i represents the "error" in the calculation which occurs when the round is redundant and $\Delta_i \neq 0$. From (15,13)

$$\frac{E[x_i^* - x_i]}{2/n} = E[\Delta_i] = \sum_{\vec{j}} \Delta(\vec{j}; i) x_{j_1} x_{j_2} x_{j_3} x_{j_4}.$$

Here we have deliberately divided by the change in "time" $2/n$ so as to give the discrete analogue of a derivative.

Now we bound e_i. If $e_i \neq 0$ then the round must be redundant or the added edge has two vertices in the same component. Further, the component containing the selected edge must have size at most K. (Recall Case IV: when an edge is placed between two vertices both already in large components all of the Δ_i are 0.) This occurs with probability at most $2Kn^{-1}$. Further, in this case $x_i^+ = x_i$ so $e_i = -\frac{2}{n}\Delta_i$ and so, from (14), $|e_i| \leq \frac{2K}{n}$. Thus

$$\frac{E[e_i]}{2/n} \leq 2K^2 n^{-1}. \tag{17}$$

In summary, from (16–17)

$$\frac{E[x_i^+ - x_i]}{2/n} = \sum_{\vec{j}} \Delta(\vec{j}; i) x_{j_1} x_{j_2} x_{j_3} x_{j_4} + O(K^2 n^{-1}). \tag{18}$$

We also shall use the maximal change $|x_i(G^+) - x_i(G)|$. When the edge selected joins two vertices already in the same component this is zero, otherwise (from (14)) it is at most $2Kn^{-1}$ so that

$$|x_i(G^+) - x_i(G)| \leq 2Kn^{-1} \tag{19}$$

always.

`Countinuing the Basic Example:` Here we have $\Delta(1, 1, \alpha, \beta; 1) = -1$ as two vertices are no longer isolated. With $(\alpha, \beta) \neq (1, 1)$ we have $\Delta(\alpha, \beta, \gamma, \delta; 1)$ is $-\frac{1}{2}$ times the number of γ, δ which equal one. Combining terms

$$\frac{E[x_1^* - x_1]}{2/n} = -x_1^2(G) - (1 - x_1^2(G)) x_1(G). \tag{20}$$

We can see this by noting that with probability $x_1^2(G)$ we select the first edge and $\Delta_1 = -1$ while with probability $1 - x_1^2(G)$ we select the second edge which has two uniformly chosen vertices and so the expected number of isolated vertices deleted is $2x_1(G)$.

2.2 Change in Susceptibility

Now we examine the change in the susceptibility, $S^+ - S$. We define

$$S^*(\vec{v}) = \begin{cases} S(G) + \frac{1}{n}2|C(v_1)| \cdot |C(v_2)| & \text{if } c(\vec{v}) \in F, \\ S(G) + \frac{1}{n}2|C(v_3)| \cdot |C(v_4)| & \text{if } c(\vec{v}) \notin F \end{cases} \tag{21}$$

$$e_S = S^+ - S^*. \tag{22}$$

When $\{v_1, v_2\}$ is selected in a nonredundant round, components of size $|C(v_1)|$, $|C(v_2)|$ merge to form a component of size $|C(v_1)| + |C(v_2)|$ in G^+ so that $S^+ = S + \frac{1}{n}2|C(v_1)| \cdot |C(v_2)| = S^*$ and $e_S = 0$. When $\{v_3, v_4\}$ is selected the situation is identical.

We first bound $E[e_S]$. For each component C_i there is probability at most $2|C_i|^2 n^{-2}$ that either both $v_1, v_2 \in C_i$ or both $v_3, v_4 \in C_i$. When this occurs $0 \geq e_S \geq -\frac{1}{n}2|C_i|^2$. Thus

$$0 \geq E[e_S] \geq -\frac{4}{n^3}\sum_{i=1}^{u}|C_i|^4 = -\frac{2}{n}(4\nabla(G)) \tag{23}$$

(see (7)). We comment that this error shall prove more troublesome than that for the change in x_i.

We turn now to the expected change $E[S^* - S]$. We divide by $2/n$ (to give the discrete derivative) and split according to values $\vec{j} = c(\vec{v})$.

$$\frac{E[S^* - S]}{2/n} = \sum_{\vec{j}} I(\vec{j}). \tag{24}$$

Here we set $I(\vec{j})$ equal to n^{-4} times $\frac{n}{2}$ times the sum of the values $S^* - S$ over all choices of $\vec{v}$ with $c(\vec{v}) = \vec{j}$.

`Case I:` $\vec{j} \in F$, $j_1, j_2 \neq \omega$. Here $S^* - S = \frac{2}{n}j_1 j_2$. There are $x_{j_1}x_{j_2}x_{j_3}x_{j_4}n^4$ such terms and so

$$I(\vec{j}) = j_1 j_2 x_{j_1}x_{j_2}x_{j_3}x_{j_4}. \tag{25}$$

`Case II:` $\vec{j} \in F$, $j_1 \neq \omega, j_2 = \omega$. Here $S^* - S = \frac{2}{n}j_1|C(v_2)|$. There are $x_{j_1}x_{j_3}x_{j_4}n^3$ choices for v_1, v_3, v_4. For each the sum over v_2 with $c(v_2) = \omega$ of $|C(v_2)|$ is the sum of the squares of the sizes of the components of size greater than K. This is precisely nS_ω and so

$$I(\vec{j}) = j_1 x_{j_1}x_{j_3}x_{j_4}S_\omega. \tag{26}$$

`Case III:` $\vec{j} \in F$, $j_1 = \omega, j_2 \neq \omega$. Reversing the roles of j_1, j_2

$$I(\vec{j}) = j_2 x_{j_2}x_{j_3}x_{j_4}S_\omega. \tag{27}$$

`Case IV:` $\vec{j} \in F$, $j_1 = j_2 = \omega$. This is the case that drives the process to a "percolation point", as the large components get still larger. Here $S^* - S = \frac{2}{n}|C(v_1)||C(v_2)|$. There are $x_{j_3}x_{j_4}n^2$ choices for v_3, v_4. For each consider the sum over all v_1, v_2 with $c(v_1) = c(v_2) = \omega$ of $C(v_1) \cdot C(v_2)$. This is precisely the square of the sum over all v_1 with $c(v_1) = \omega$ of $C(v_1)$. This sum over v_1 is precisely, as done in Case II, nS_ω. Hence the sum over v_1, v_2 is $n^2 S_\omega^2(G)$. Hence

$$I(\vec{j}) = x_{j_3}x_{j_4}S_\omega^2. \tag{28}$$

`Cases V-VIII` The cases in which $\vec{j} \notin F$ are identical with (j_1, j_2) and (j_3, j_4) exchanging roles.

In summary, from (22–24),

$$\frac{E[S^+ - S]}{2/n} = \sum_{\vec{j}} I(\vec{j}) + O(\nabla(G)). \tag{29}$$

`The Erdos-Rényi Evolution:` Here $x_\omega(G) = 1$ tautologically and Case IV gives that

$$\frac{E[S^* - S(G)]}{2/n} = S_\omega^2.$$

Suppose we (audaciously) assume that the term e_S is negligible and that $S(G_i)$ evolves according to its expectation. Set $g(t) = S(G_{tn/2})$, further replacing a random variable by a single value. Moving from G_i to G_{i+1} is increasing the "time" t by $2/n$. Thus we would have

$$\frac{g(t + \frac{2}{n}) - g(t)}{2/n} = g^2(t). \tag{30}$$

The left hand side is like a derivative which suggests the differential equation $g'(t) = g^2(t)$. With the initial condition $g(0) = 1$ (as this is the value of S on the empty graph) we have the solution $g'(t) = (1-t)^{-1}$. This function approaches infinity as $t \to 1^-$. This matches the known results about the Erdős-Rényi Evolution: when $t < 1$ the random graph on $t\frac{n}{2}$ edges has $S \sim (1-t)^{-1}$ and at $t = 1$ there is the famous "double jump", or percolation point, and a "giant component" quickly appears.

`Continuing the Basic Example:` In our basic example

$$\frac{E[S^* - S]}{2/n} = x_1^2(G)(1) + (1 - x_1^2(G))S^2(G). \tag{31}$$

With probability $x_1^2(G)$ an edge is added between two isolated vertices, raising S by $(2^2 - 1^2 - 1^2)/n = 2/n$. Otherwise, with probability $1 - x_1^2(G)$ a random edge is added and this raises S by an expected value $\frac{2}{n}S^2$.

2.3 The System of Differential Equations

We define a system of differential equations on functions $x_i(t)$, $i \in \Omega$ and $S(t)$. The proof of the connection with Bounded Size Algorithms will be established in Section 4. The function $x_\omega(t)$ shall be technically redundant as we shall have $\sum_{i\in\Omega} x_i(t) = 1$. For convenience we define $S_\omega(t)$ (following (6)) by

$$S_\omega(t) = S(t) - \sum_{i=1}^{K} i x_i(G). \tag{32}$$

This function is also technically redundant but shall allow us a cleaner statement of the equations.

The initial values of the system are at $t = 0$ with

$$x_1(0) = 1; \quad x_i(0) = 0 \text{ for all } i \neq 1; \quad S(0) = 1. \tag{33}$$

For $i \in \Omega$ we have the equation (motivated by (18))

$$x_i'(t) = \sum_{\vec{j}\in\Omega} \Delta(\vec{j}; i) x_{j_1}(t) x_{j_2}(t) x_{j_3}(t) x_{j_4}(t). \tag{34}$$

Finally, and critically, we have (motivated by (29))

$$S'(t) = \sum_{\vec{j}} I(\vec{j}, t) \tag{35}$$

where the values of I are given by (25,26,27,28) and the similar cases V-VIII, in which one replaces the x_j with $x_j(t)$ and the S_ω with $S_\omega(t)$.

`Continuing the Basic Example:` In our basic example from (20) we have the equation

$$x_1'(t) = -x_1^2(t) - (1 - x_1^2(t))x_1(t).$$

With initial condition $x_1(0) = 1$ this has a unique solution, a function $x(t)$ defined over the nonnegative reals. One can show that $x(t)$ is a strictly decreasing function with $\lim_{t\to\infty} x(t) = 0$. From (31) we have the equation

$$S'(t) = x_1^2(t) + (1 - x_1^2(t))S^2(t).$$

As $0 \le x_1(t) \le 1$ the function $S(t)$ will be strictly increasing. As $1 - x_1(t)$ will be uniformly bounded away from zero for, say, $t \ge 0.1$ we can show that the $S^2(t)$ term forces the function $S(t)$ to "explode" in finite time (c.f. Theorem 2.2). That is, there exists t_c such that S is defined for $0 \le t < t_c$ and $S(t)$ approaches infinity as t approaches t_c from below. Our general results show that this t_c is the percolation point for this process. For any $t < t_c$ all components have size $O(\ln n)$ while for any $t > t_c$ there will be a giant component of size $\Omega(n)$.

2.4 The Differential Equations: Technical Analysis

Theorem 2.1 *Let $x_i(t)$, $i \in \Omega$ be the solution to the system (34) with initial conditions (33). Then*

1. *$x_i(t)$ is defined for all $t \ge 0$.*
2. *$\sum_i x_i(t) = 1$ for all $t \ge 0$.*
3. *$x_i(t) > 0$ for all $t > 0$.*
4. *$x_\omega'(t) > 0$ for all $t > 0$.*

`Proof:` We note that these results are not surprising as they reflect the properties of analogous functions for the discrete process G_i. As (34) is of the form $\vec{x}\,' = f(\vec{x})$ with f a C^∞ (in fact, polynomial) function of $\vec{x}$ the system has a solution in some neighborhood of $t = 0$ (see [6, Chapter 2, Theorem 11] for example).

For all $\vec{j}$ we have $\sum_{i\in\Omega} \Delta(\vec{j}, i) = 0$. Summing (34) over $i \in \Omega$ gives $\sum_{i\in\Omega} x_i' = 0$ so $\sum_{i\in\Omega} x_i(t) = 1$ for all t for which it is defined.

For $i \in \Omega$, $s \ge 0$, set $x_i^{(s)}$ equal the value of the s-th derivative of x_i at $t = 0$. By the 0-th derivative we mean the function itself. So $x_1^{(0)} = 1$ and $x_i^{(0)} = 0$ for $i \ne 1$. Differentiating the equation $\vec{x}\,' = f(\vec{x})$ repeatedly we can find $x_i^{(s)}$ as polynomials in $x_j^{(r)}$ with $r < s$ and hence they are determined. We claim that for each $i \in \Omega$ not all $x_i^{(s)}$ are zero and that the first nonzero value is positive.

For $i = 1$ we have simply $x_1^{(0)} = 1$. For $i \ne 1$ the polynomial for x_i' may have both positive and negative coefficients but the negative coefficients (reflecting a component of size i being absorbed in a larger component – these do not appear for $i = \omega$) are for quartics of the form $x_i x_j x_k x_l$. In expanding out the s-th derivative of this term each addend will have some $x_i^{(r)}$ with $r < s$. Thus these terms cannot contribute to the first nonzero value. For each $i \ne 1$ the polynomial for x_i' will have an addend

$ix_{i-1}x_1x_{i-1}x_1$ with K positive (we interpret $\omega-1$ as K here) reflecting the fact that if the four vertices are in components of sizes $i-1,1,i-1,1$ respectively then any algorithm will create a component of size i. By induction on i there is some s' with $x_{i-1}^{(s)}\neq 0$. Then $x_i^{(2s'+1)}$ will have an addend $ix_1x_{i-1}^{(s')}x_1x_{i-1}^{(s')}$ which will be nonzero. Now consider the first s for which the expansion of $x_i^{(s)}$ has a nonzero addend. That addend will be a positive constant times the product of four terms of the form $x_j^{(r)}$. Furthermore for all four such terms we must have $x_j^{(r-r')}=0$ for $r'>0$ as, if not, there would be a positive addend in the evaluation of $x_i^{(s-r')}$. By induction on s we have that when $x_i^{(s')}$ has its first nonzero addend at $s'=s$ that all such addends will be positive. This completes the claim.

As its first nonzero derivative is positive we deduce that $x_i(t)>0$ in some interval $(0,\epsilon)$. Now we show that $x_i(t)$ is always positive for all i. If not there would be a minimal t for which some $x_i(t)=0$. Fix $t'\in(0,t)$ (say, $t'=\frac{t}{2}$) so that $x_i(t')>0$. In the interval $[t',t)$ all x_j are positive and hence all $x_j\in[0,1]$. The negative addends for x_i' are all of the form $-Kx_ix_jx_kx_l$ which is at least $-Kx_i$. Adding these over the finite number of addends we find $x_i'>-K'x_i$. Thus the function x_i in $[t',t]$ is bounded from below by an exponential function and this contradicts $x_i(t)=0$.

As all of the $x_i(t)$ are nonnegative and they sum to one, all $x_i(t)\in[0,1]$ when defined. Note that f is continuous and the equation $\vec{x}\,'=f(\vec{x})$ takes place in a compact space so $\vec{x}(t)$ is defined for all $t\geq 0$.

Finally, the polynomial for $x_\omega'(t)$ has only positive coefficients times $x_ix_jx_kx_l$. (It has at least one such term, $(K+1)x_Kx_1x_Kx_1$.) At any $t>0$ we have $x_ix_jx_kx_l>0$ and the sum of positive terms must be positive. This completes the verification of the four parts of the theorem. ∎

Theorem 2.2 *Let $x_i(t),S(t)$ be the solution to the system (34,35) with initial conditions (33). Then there is a $t_c>0$ such that $S(t)$ is defined for all $t\in[0,t_c)$ and*

$$\lim_{t\to t_c^-}S(t)=+\infty. \tag{36}$$

Furthermore $S(t)$ is a strictly increasing function on $[0,t_c)$

`Proof:` All of the terms in the expansion (35) of $S'(t)$ are nonnegative. Taking, say, $\vec{j}=(1,1,1,1)$ there is an addend $x_1^4(t)$ from (25). As $x_1(t)$ is strictly positive $S'(t)$ is therefore strictly positive and so $S(t)$ is a strictly increasing function.

To show (36) we actually examine S_ω, defined in (32). Equation (35) becomes

$$S_\omega'(t)=\sum_{\vec{j}}I_\omega(\vec{j},t). \tag{37}$$

Define S_ω^* analogously to S^*, to refer only to components bigger than K only: if $c(\vec{v})\in F$ then

$$S_\omega^*(\vec{v})=\begin{cases} S(G)+\frac{1}{n}2|C(v_1)|\cdot|C(v_2)| & \text{if } |C(v_1)|>K \text{ and } |C(v_1)|>K,\\ S(G)+\frac{1}{n}(|C(v_1)|^2+2|C(v_1)|\cdot|C(v_2)|) & \text{if } |C(v_1)|\leq K \text{ and } |C(v_1)|>K,\\ S(G)+\frac{1}{n}(|C(v_1)|+|C(v_2)|)^2 & \text{if } |C(v_1)|\leq K \text{ and } |C(v_1)|\leq K,\end{cases}$$

and analgously for the fourth case, and symmetrically if $c(\vec{v})\notin F$. Computing I_ω as I was done in Section 2.2, we find there are only a few changes, as follows.

Case I requires the extra condition $j_1+j_2>K$ (otherwise $I_\omega=0$) and gives $I_\omega=\frac{1}{2}(j_1+j_2)^2x_{j_1}x_{j_2}x_{j_3}x_{j_4}$,

Case II becomes $I=(\frac{1}{2}j_1^2x_\omega+j_1S_\omega)x_{j_1}x_{j_3}x_{j_4}$,

Case III is symmetric to II,

Case IV is unaltered.

As all of the $x_i(t)$ are nonnegative, all of the addends in (37) are nonnegative. We look at the particular $\vec{j} = (\omega, \omega, \omega, \omega)$. (This corresponds to all four vertices being in large components so that one must join two vertices in large components and greatly increase S.) Regardless of whether or not $\vec{j} \in F$ we have, from (28) $I(\vec{j}) = x_\omega^2 S_\omega^2$. This gives the lower bound

$$S'_\omega(t) \geq x_\omega^2(t) S_\omega^2. \tag{38}$$

Suppose S_ω is defined at some $t' > 0$ and set $a = S_\omega(t')$. Set $c = x_\omega^2(t')$. As x_ω is an increasing function we have $S'_\omega(t) \geq c S_\omega^2(t)$ for all $t \geq t'$. Thus $S_\omega(t) \geq f(t)$ for $t \geq t'$ where $a = f(t')$ and $f'(t) = cf^2(t)$. But this equation has the explicit solution $f(t) = (a^{-1} - c(t - t'))^{-1}$ so $f(t) \to \infty$ as $t \to t' + (ac)^{-1}$. We don't have an explicit solution for S_ω but this implies there will be some $t_c \in [t', t' + (ac)^{-1}]$ so that S_ω is defined on $[0, t_c)$ and approaches infinity as t approaches t_c from below. As $S_\omega \leq S \leq S_\omega + K$ (36) holds with the same t_c.

3 Evolution from an Initial Graph

3.1 Statement of Results

Consider the following evolution. Begin with a graph G on n vertices which is suitably sparse. Add to it a random graph H with edge probability $p = \frac{t}{n}$ on the same vertex set. When G is empty we have, of course, the standard Erdős-Rényi evolution with the phase transition (referred to in the original papers as the "double jump") at $t = 1$. We show in this section that, with a suitable side condition, there will be a critical value t_c before which all components are of size $O(\ln n)$ and after which a giant component of size $\Omega(n)$ has emerged. We will apply these results to our analysis of Achlioptas processes in later sections. Since, however, these results are of natural interest we note that the arguments given here are independent of the remainder of this work.

Theorem 3.1 *Let L, K, c be positive real numbers. Let G be a graph on n vertices with a K, c component tail. Let H be a random graph with edge probability $p = \frac{t}{n}$ on the same vertex set, where t is fixed. Set $G^+ = G \cup H$.*

1. *(subcritical) Assume $S(G) \leq L$ for all n. Let $tL < 1$. Then there exist K^+, c^+ (dependent on K, c, L, t but not on n nor G) such that a.a.s. G^+ has a K^+, c^+ component tail. In particular, all components have size $O(\ln n)$.*

2. *(supercritical) Assume $S(G) > L$. Let $tL > 1$. Then a.a.s. G^+ has a giant component. More precisely, there exists $\gamma > 0$ (dependent on K, c, L, t but not on n nor G) such that G^+ has a component of size at least γn.*

3.2 Three Branching Process Results

Here we give three results on branching processes. The first two shall be used to show the subcritical case, and the third the supercritical case, of Theorem 3.1. Theorem 3.2 below was essentially shown by Crámer in 1920. The proofs in all three cases are complicated by the requirement of getting "explicit" constants.

Theorem 3.2 *Let K, c be positive reals. Let Z be a nonnegative integer valued random variable with a K, c tail and with $E[Z] = \mu < 1$. Let T be the size of the Galton-Watson branching process in which each node, independently, has Z children. Then there exist positive K^+, c^+, dependent only on μ, K, c, such that T has a K^+, c^+ tail.*

`Proof:` Fix any positive $\lambda < c$, say $\lambda = \frac{c}{2}$ for definiteness. The Laplace Transform $E[e^{tZ}]$ is then defined for all $0 \le t \le \lambda$. For such t we also have $E[(Z-1)^2 e^{tZ}] \le E[(Z-1)^2 e^{\lambda Z}]$ which is bounded by a convergent sum. Let M be an explicit upper bound on $E[(Z-1)^2 e^{tZ}]$. Now, using the standard association of branching processes with random walks, we have

$$\Pr[T \ge s+1] \le \Pr[Z_1 + \ldots + Z_s \ge s] \tag{39}$$

where the Z_i (number of children of the i-th node) are independent, each with distribution Z. For any $0 \le t \le \lambda$ we have the Chernoff bound

$$\Pr[Z_1 + \ldots + Z_s \ge s] \le E[e^{t(Z_1+\ldots+Z_s)}]e^{-ts} = E[e^{tZ}e^{-t}]^s. \tag{40}$$

Let us set

$$\phi(t) := e^{-t}E[e^{tZ}] = E[e^{t(Z-1)}]. \tag{41}$$

We have $\phi(0) = 1$ and

$$\phi'(0) = E[Z-1] = \mu - 1 \tag{42}$$

which, critically, is negative. Also,

$$\phi''(t) = E[(Z-1)^2 e^{t(Z-1)}] \le M \tag{43}$$

for all $t \in [0, \lambda]$. Hence

$$\phi(t) \le 1 + (\mu-1)t + (1+M)\frac{t^2}{2} \tag{44}$$

for all $t \in [0, \lambda]$. We set $t = \frac{1-\mu}{1+M}$ if this value is less than λ, otherwise we set $t = \lambda$. Either way we get an explicit $U < 1$ and the bound $\phi(t) \le U$. This gives Theorem 3.2 with $K^+ = U^{-1}$ and $c^+ = -\ln U$.

Theorem 3.3 *Let K_1, c_1, K_2, c_2 be positive reals. Let X, Y be nonnegative integer values random variables with X having a K_1, c_1 tail and Y having a K_2, c_2 tail. Consider the two generation branching process in which the root node has X children and then each child independently has Y children. Let Z be the number of grandchildren. Then there exist K, c, dependent only on K_1, c_1, K_2, c_2, such that Z has a K, c tail.*

`Proof:` Let $F_X(x) = \sum \Pr[X = i]x^i$, $F_Y(x) = \sum \Pr[Y = i]x^i$, $F_Z(x) = \sum \Pr[Z = i]x^i$ be the standard generating functions for X, Y, Z. These are related by $F_Z(x) = F_X(F_Y(x))$. Fix $\lambda_1 < c_1$, say $\frac{1}{2}c_1$ for definiteness. Then $F_X(e^{\lambda_1})$ has an explicit bound $L := \sum_s K_1 e^{(\lambda_1 - c_1)s}$. By monotonicity $F_X(t) \le L$ for all $1 \le x \le e^{\lambda_1}$. Fix $\lambda_2 < c_2$, say $\frac{1}{2}c_2$ for definiteness. Then for $1 \le x \le e^{\lambda_2}$

$$F_Y'(x) = \sum i x^{i-1}\Pr[Y = i] \le \sum i x^{i-1} K_2 e^{-c_2 i} \le K_2 \sum_i i e^{(\lambda_2 - c_2)i}.$$

Let M denote this bound on $F_Y'(x)$. Then

$$F_Y(x) \le 1 + M(x-1) \tag{45}$$

for $1 \le x \le e^{\lambda_2}$. Set

$$t = \min[\frac{1}{M}(e^{\lambda_1} - 1), e^{\lambda_1}].$$

Then

$$F_Z(t) \le F_X(1 + M(t-1)) \le F_X(e^{\lambda_1}) \le L. \tag{46}$$

For any s we have

$$L \ge F_Z(t) \ge \Pr[Z \ge s]t^s \tag{47}$$

which gives Theorem 3.3 with $K = L$ and $c = \ln t$.

Theorem 3.4 *Let Z be a distribution on the nonnegative integers with $E[Z] = \mu > 1$ and $Var[Z] = \sigma^2$. Let $Z_1, Z_2, \ldots$ be independent, each with distribution Z. Then*

$$\Pr[\sum_{i=1}^{s} Z_i \leq s-1] \leq \beta^s. \tag{48}$$

Also, consider the Galton-Watson tree in which each node independently has Z children. That tree is finite with probability at most y. Here $\beta < 1$, $y < 1$ and both β, y depend only on μ, σ.

`Proof:` Set $\phi(t) = E[e^{-t(Z-1)}]$. Then $\phi(0) = 1$, $\phi'(0) = E[Z-1] = \mu - 1$ and, for $0 \leq t \leq 1$,

$$\phi''(t) = E[(Z-1)^2 e^{-t(Z-1)}] \leq \Pr[Z=0]e^t + E[Z^2] \leq e + \mu^2 + \sigma^2$$

so that $\phi(t) \leq 1 + (\mu-1)t + K\frac{t^2}{2}$ with $K = e + \mu^2 + \sigma^2$. We select t to minimize this quadradic, or $t = 1$ if the minimum is not in $[0,1]$, to give an explicit $\phi(t) < 1$.

$$\Pr[\sum_{i=1}^{s} Z_i \leq s-1] \leq \Pr[\sum_{i=1}^{s}(Z_i - 1) \leq 0] \leq E[e^{-t\sum(Z_i-1)}] = \phi(t)^s$$

giving (48). Let w be the probability the Galton-Watson tree is finite. Then w is the minimal nonnegative solution to the equation $w = F_Z(w)$ where $F_Z(w) = \sum_s \Pr[Z=s]w^s$. Then $F_Z(1) = 1$, $F_Z'(1) = E[Z] = \mu$ and for $t \in [0,1]$ $F_Z''(t) \leq F_Z''(1) = E[Z(Z-1)] < K$ with $K = \sigma^2 + \mu^2$. Thus, for $t \in [0,1]$, $F_Z(1-t) < 1 - t\mu + K\frac{t^2}{2}$. At $t = 2(\mu-1)/K$, $F_Z(1-t) < 1-t$. Hence $w \leq 1 - 2(\mu-1)/K$.

3.3 The Subcritical Case

`Notation:` $C(v)$ shall refer to the component containing v in the graph G. $C^+(v)$ shall refer to the component containing v in the graph $G^+ = G \cup H$ as given by Theorem 3.1. We call w a child of v if there exists z so that $\{v,z\} \in H$ and $w \in C(z)$. (We include the case $w = z$.) We call w a descendant of v if there is a sequence $v = v_0, \ldots, v_r = w$ with v_{i+1} a child of v_i for $0 \leq i < r$. (This includes $w = v$ with the trivial sequence $v = v_0$.) We let $C^-(v)$ denote the set of descendants of v. Note that $C^-(v) \subseteq C^+(v)$, the exact relation being given by (50).

Theorem 3.5 *In the notation of Theorem 3.1 let $Z(v)$ denote the number of children of v, where v is an arbitrary vertex. Then $Z(v)$ has a K_1, c_1 tail, where K_1, c_1 depend only on the L, K, c, t of Theorem 3.1. Further $E[Z(v)] \leq tS(G)$.*

`Proof:` Define Z^+ as follows. Take a random subset S of the vertices where $\Pr[z \in S] = p = \frac{t}{n}$ and set $Z^+ = \sum_{z \in S} |C(z)|$, with the sum interpreted as zero when $S = \emptyset$. For any fixed v the (random) neighborhood, call it $S(v)$, of v in H has the distribution of S except that v cannot be a neighbor of itself. Further,

$$Z(v) = |\cup_{z \in S(v)} C(z)| \leq \sum_{z \in S(v)} |C(v)| \tag{49}$$

so that $Z(v)$ is dominated by Z^+. It suffices, therefore, to show that Z^+ has a K_1, c_1 tail.

Z^+ *almost* fits the assumptions of Theorem 3.3. In the notation of that theorem we set $X = |S|$ which has Binomial Distribution $B[n,p]$. Conditioning on $X = a$ the elements $v_1, \ldots, v_a$ form a random a-element set. Let Y have the distribution of $|C(w)|$ with w a uniformly chosen vertex. The problem is that the $|C(v_i)|$ are not quite independent as we must have $v_i \neq v_j$. We take care of this by first setting $\alpha = \ln(1-p)$. As $p = \frac{t}{n} = o(1)$, $\alpha \sim p$. Consider a two generation tree where X has Poisson distribution with mean $n\alpha$. For each of the X children we uniformly and independently select a vertex v from the graph (allowing repetition) and give that child $Y = |C(v)|$ children. Let Z^* denote the number of grandchildren in this tree. We apply Theorem 3.3 so that $Z^* = \sum_{i=1}^{X} |C(v_i)|$ has a K_1, c_1 tail. For each v the number of times that v appears amongst the X children is a Poisson distribution with mean α and so the probability that v appears is p. Thus if we take the sum for Z^* but only add $|C(v)|$ one time when v appears more than once we have the distribution for Z^+. So Z^* dominates Z^+ which dominates $Z(v)$ and hence $Z(v)$ has a K_1, c_1 tail. Furthermore $E[Z(v)] \leq E[Z^*] = E[X]E[Y] = tS(G)$.

We consider the probability space consisting of the generation of the random graph H and the uniform generation of v.

Theorem 3.6 *Fix L, K, c, t in Theorem 3.1. Then $|C^-(v)|$ has a K_2, c_2 tail.*

Let Z^* be the distribution defined in the proof of Theorem 3.5. Let T^* denote the total size of a Galton-Watson process in which each node independently has Z^* children. From Theorem 3.5 Z^* has a K_1, c_1 tail and $E[Z^*] < 1$ hence from Theorem 3.2 T^* has a K_2, c_2 tail.

We now generate the vertices of $C^-(v)$ by, say, breadth first search. When probing for the children of v', however, we look only for w not already in the tree, then adding $C(w)$ whenever $\{v', w\} \in H$. When this occurs call all $w' \in C(w)$ "new children" of v'. This gives $C^-(v)$ a tree structure with no vertex appearing more than once. When probing for new children of v' we are checking H for edges $\{v', w\}$ and these have not been probed before. Hence the conditional distribution of the number of new children of v', conditioning on previous history, is dominated by Z^*. Hence the distribution of the size of the tree, $|C^-(v)|$, is dominated by T^* and it therefore has a K_2, c_2 tail.

Theorem 3.7 *There exist K', c' so that the random variable $|C^+(v)|$ has a K', c' tail.*

We first note that

$$C^+(v) = \cup_{w \in C(v)} C^-(v). \tag{50}$$

Say $C(v) = \{v_1, \ldots, v_X\}$. Analogously to the proof of Theorem 3.6 we generate the descendents of v_i sequentially. We find $C^-(v_i)^*$, the set of vertices of $C^j(v_i)$ not already in $\cup_{j<i} C^-(v_j)$. Set $R = \cup_{j<i} C^-(v_j)$. When $v_i \in R$ we simply set $C^-(v_j) = \emptyset$. Otherwise we find $C^-(v_j)^*$ as in Theorem 3.6 except that we do not probe for any $w \in R$. As before, conditioning on the previous history does not change the adjacency probabilities of H as these pairs have not been probed before. Hence the distribution of $C^-(v_j)^*$ is dominated by T^* which has a K_2, c_2 tail. Then $|C^+(v)|$ is dominated by a two generation process with distributions X, T^* respectively. By Theorem 3.3 it has a K', c' tail.

We now prove the subcritical case of Theorem 3.1 for any K^+, c^+ with $K^+ > K'$ and $c^+ < c'$. Our probability space is now the choice of the random H. For each vertex v let A_v be the event $|C^+(v)| \geq s$ and I_v the indicator random variable of A_v. Set $Y_s = \sum_v I_v$, the number of vertices in components of size at least s. From Theorem 3.7 and linearity of expectation $E[Y_s] \leq nK'e^{-c's}$. From Markov's Inequality

$$\Pr[Y_s \geq nK^+e^{-c^+s}] \leq \frac{nK^+e^{-c^+s}}{nK'e^{-c's}} = O(e^{(c'-c^+)s}). \tag{51}$$

For s "small" (for example $s = O(1)$) this inequality is not good enough and instead we bound the variance

$$Var[Y_s] = \sum_{v,w} E[I_v I_w] - E[I_v]E[I_w]. \tag{52}$$

We claim that for any v, w

$$\Pr[A_v \wedge A_w] \leq \Pr[A_v]\Pr[A_w] + \Pr[C^+(v) = C^+(w)]. \tag{53}$$

We use here a theorem of Reimer [7]. In general form, let $R \subset \Omega$ be a random subset of a universe Ω with the events $r \in R$ mutually independent over all $r \in \Omega$. Let A_1, A_2 be monotonic events in the sense that if they hold when $R = S_1$ then they hold when $R = S_2$ for any $S_1 \subseteq S_2$. Let $A_1 * A_2$ be the event that the set R contains two *disjoint* sets S_1, S_2 such that A_1 holds when $R = S_1$ and A_2 holds when $R = S_2$. Reimer's Theorem then says that $\Pr[A_1 * A_2] \leq \Pr[A_1]\Pr[A_2]$. In our case Ω is the set of pairs $\{v, w\}$ of vertices and $R = G \cup H$. That is, $\Pr[\{v,w\} \in R] = p$ if $\{v,w\} \notin G$, $\Pr[\{v,w\} \in R] = 1$ if $\{v,w\} \in G$. A_1, A_2 are A_v, A_w respectively. If A_v and A_w and $C^+(v) \neq C^+(w)$ then we let S_1 be the set of edges of $C^+(v)$ and S_2 be the set of edges of $C^+(w)$. Thus

$$\Pr[A_v \wedge A_w \wedge C^+(v) \neq C^+(w)] \leq \Pr[A_v * A_w] \leq \Pr[A_v]\Pr[A_w]$$

from whice (53) follows.

Now summing (52) over all v, w and applying (53) we have

$$Var[Y_s] \leq \sum_{v,w} \Pr[C^+(v) = C^+(w)].$$

For each fixed v this sum over w is simply $E[|C^+(v)|]$ and hence

$$Var[Y_s] \leq \sum_v E[|C^+(v)|].$$

But the right hand side is precisely $nE[|C^+(v)|]$ with v chosen uniformly. From Theorem 3.7 this variable has a K', c' tail and therefore has a bounded expectation. That is

$$Var[Y_s] = O(n)$$

where the implicit constant does not depend on s.

From Chebyschev's Inequality

$$\Pr[Y_s \geq nK^+e^{-c^+s}] \leq Var[Y_s](n(K^+e^{-c^+s} - K'e^{-c's}))^{-2} = O(n^{-1}e^{2c's}). \tag{54}$$

We use Markov's Inequality (51) when, say, $s \geq \ln\ln n$ and Chebyschev's Inequality (54) when $s < \ln\ln n$ so that

$$\sum_{s=0}^{\infty} \Pr[Y_s \geq nK^+e^{-c^+s}] = o(1)$$

and thus the subcritical case of Theorem 3.1 holds.

3.4 The Supercritical Case

Here we show the supercritical case of Theorem 3.1. We first reduce to the case when the component sizes of G are bounded. Given L, K, c, t satisfying the conditions of Theorem 3.1 let $\delta > 0$ with $(L - \delta)t > 1$ and (to avoid trivialities) $\delta < 1$. Pick M sufficiently large so that $\sum_{s>M} Ke^{-cs}s < \delta$. Let W be the set of vertices in components of size at most M and let G^-, H^- be the restrictions of G, H to W. Then G^- has m vertices with $1 - \frac{m}{n} < \sum_{s>M} Ke^{-cs} < \delta$ and $S(G^-) \geq \frac{n}{m}(S(G) - \sum_{s>M} Ke^{-cs}s) > \frac{n}{m}(L - \delta)$. On W the random H^- has probability $p = \frac{t'}{m}$ with $t' = \frac{m}{n}t$. Then $t'S(G^-) > 1$. If Theorem 3.1 holds for G^- then $G^- \cup H^-$ has a component of size $\Omega(m)$. But $m \geq n(1 - \delta)$ so this would be a component of size $\Omega(n)$ inside of $G \cup H$.

Hence we may, and shall, assume that G has all components of size at most M. Set $x_i = x_i(G)$, for $1 \leq i \leq M$. Fix t with $tL > 1$ where $S(G) = L + o(1)$. Select a small positive γ so that

$t(L - 2\gamma M^2) > 1$. We shall actually show that some $|C^-(v)| \geq \gamma n$. (Recall $C^-(v)$ from the start of the previous subsection.)

We apply the following well known procedure. Take a random vertex v_1 and generate $C^-(v_1)$. If $|C^-(v_1)| \geq \gamma n$ we call the procedure a success and terminate. Set $\omega(n) = \ln \ln n$, though any sufficiently slow growing function of n would suffice. If $\omega(n) \leq |C^-(v_1)| < \gamma n$ we call the procedure a failure and terminate. Otherwise, we remove the vertices of $C^-(v_1)$, select a random v_2 from the remaining vertices and generate $C^-(v_2)$ in that subgraph. We iterate this procedure. If at some stage more than γn vertices have been removed (and the procedure has not yet been terminated) we terminate the procedure as a failure. We show that this procedure terminates as a success with probability $1 - o(1)$.

Consider any iteration of this procedure. Let R be the set of remaining vertices. We select $v \in R$ uniformly and generate $C^-(v)$ by breadth first search, terminating if and when the size reaches γn. When searching for the children of some w, at most $2\gamma n$ of the original vertices cannot be used. Set $y_i = \max\{x_i - 2\gamma, 0\}$. Then

$$\sum i y_i \geq \sum i x_i - \sum_{i=1}^{M} 2\gamma i \geq L - 2\gamma M^2$$

where we have not attempted to optimize. There are at least $y_i n$ vertices in components of size i, $1 \leq i \leq M$, that are available. Let Z be the distribution obtained by taking $\lfloor \frac{1}{i} y_i n \rfloor$ components of size i, selecting each vertex independently with probability t/n, and counting the number of vertices in components of vertices selected. The actual distribution of the number of children of w dominates this Z. Here $E[Z] \sim t \sum i y_i \geq t(L - 2\gamma M^2) > 1$. Also $Var[Z] \sim E[Z]$ as the selecting of a point in a component is independent over the different components. Now we employ the natural connection between tree size and random walk. If $C^-(v)$ has size s then the s vertices generated precisely $s - 1$ new children. The probability that they generate at most $s - 1$ new children is bounded above by the probability that $Z_1 + \ldots + Z_s \leq s - 1$ where the Z_i independently have distribution Z and this is bounded, using Theorem 3.4, by β^s for a β bounded below one. Thus the probability that the iteration of the procedure we are considering terminates in failure is $O(\beta^{\omega(n)}) = o(1)$. The probability that the iteration terminates in success is bounded from below by the probability that the Galton-Watson tree with distribution Z has size at least γn. This is bounded from below by the probability that the tree is infinite. From Theorem 3.4 the probability the iteration terminates in success is at least y, a postive constant. Let A be an arbitrary constant. Consider the process through the first A iterations or until it terminates, whichever comes first. The probability of it terminating in failure is at most A times $o(1)$ which is $o(1)$. If at a stage there has been neither success nor failure then at most $A\omega(n) < \gamma n$ vertices have been used so the next iteration is performed. The probability that none of the A stages end in success is then at most $(1 - y)^A$. As A was arbitrary the probability of success is of the form $1 - o(1)$.

4 The Subcritical Phase

In this section we fix $t < t_c - \epsilon$ where t_c is given by (36). We shall show that the functions $x_i(G_{tn/2})$, $i \in \Omega$ and $S(G_{tn/2})$ are concentrated around the values $x_i(t)$ and $S(t)$ respectively. The functions x_i will not be so difficult. The proof for S is more subtle and, in proving it, we shall actually show further properties of $G_{tn/2}$ that are interesting in their own right and that shall be useful in studying the supercritical phase.

4.1 Points in Small Components

From Theorem 2.1 the functions $x_i(t)$ are defined for all $t \geq 0$. We claim that for any fixed positive t the functions $x_j(G_{tn/2})$ are concentrated around the values $x_j(t)$. This follows using far more general results as developed by the second author. The differential equations (34) are of the form

$$\vec{x}'(t) = F(\vec{x})$$

where $F : R^{K+1} \to R^{K+1}$ is a C^∞ function. Indeed, each coordinate function is a polynomial of degree 4 with real coefficients over the variables $x_1, \ldots, x_K, x_\omega$. We define a discrete vector valued sequence of random variables $\vec{X}_i = (x_1(G_i), \ldots, x_K(G_i), x_\omega(G_i))$. What (18) tells us is that

$$\left| \frac{E[\vec{X}_{i+1} - \vec{X}_i | G_0, \ldots, G_i]}{2/n} - F(\vec{X}_i) \right|_\infty = O(K^2 n^{-1}). \tag{55}$$

Indeed, the difference in the coordinate corresponding to $x_j(G)$ is precisely the "error" e_j as bounded by (17).

Theorem 4.1 *Fix $\tau > 0$. With probability $1 - O(\exp(-n^{1/5}))$,*

$$\left| \vec{X}_i - \vec{x}(2i/n) \right|_\infty = O(n^{-1/4})$$

uniformly for $0 \leq i \leq \tau n/2$.

`Proof:` To establish the concentration of $x_i(G_{tn/2})$, we will apply [8, Theorem 5.1]. (The reader may wish to consult [9, Theorem 3], which is rather simpler to apply, but will not suffice for the second application, in the next section.) That theorem is expressed in terms of the natural unscaled variables, which appear in the present context as $Y_j = nx_j$, $j = 1, \ldots, K$. There are three hypotheses that need to be verified. The boundedness hypothesis is implied by (19). (Referring to [8, Theorem 5.1], β is a constant and $\gamma = 0$.) The trend hypothesis is implied by (55). To be specific, for any fixed positive τ we choose the compact domain $\mathcal{D} = [0, \tau] \times [0, 1]^{K+1}$, and let D be a bounded open set containing $\mathcal{D}$. The stopping time T_D is defined in the general situation to be the minimum i such that $(2i/n, \vec{X}_i) \notin D$; in this case necessarily $T_D = \lfloor \tau n/2 \rfloor + 1$. Note that (55) holds for $i < T_D$. (Referring to [8, Theorem 5.1], $\lambda_1 = O(K^2 n^{-1})$. There is also a factor of 2 difference in the scaling of time: here we have $t = 2i/n$ whilst in the theorem being applied $t = i/n$, but this merely needs to be remembered when interpreting the conclusion. The Lipschitz hypothesis requires appropriate behaviour of $F(\vec{X}_i)$, regarded as a function of $(2i/n, \vec{X}_i)$, on D. It suffices to note that the derivative of F is continuous on the closure of D.

The conclusion of [9, Theorem 3] has two parts. The first implies that the system of differential equations given by (34) with initial conditions (33) has a unique solution $(t, x_1, \ldots, x_K, x_\omega)$ in $\mathcal{D}$ which extends outside of $\mathcal{D}$ beyond some point at which $t = \sigma_0$ say, on the boundary of $\mathcal{D}$. By Theorem 2.1, $\sigma_0 = \tau$. The second part of the conclusion gives the desired concentration which we may state as given in the theorem (for convenience, taking $\lambda = n^{-1/4}$ in the terminology of [8, Theorem 5.1], and not forgetting the factor of 2 discrepancy in the definition of t). ∎

`The Erdos-Renyi Evolution:` Consider, say, the proportion $x_1(t)$ of isolated vertices in the random graph with $t\frac{n}{2}$ edges or, essentially equivalently, the random graph $G(n, \frac{t}{n})$. It is well known that $x_1(t)$ concentrates around e^{-t}. Note that this function has no special behavior near the value $t = 1$ or, indeed, any other value. That is, there is no percolation with respect to the number of isolated vertices.

4.2 Concentration of Susceptibility

We fix $t_0 < t_c$, in the subcritical phase. From Theorem 2.2 the function $S(t)$ is defined and increasing on $[0, t]$. With foresight we fix a positive integer Q with

$$2S(t_0)\frac{t_0}{Q} < 1$$

and set

$$\epsilon = \frac{t_0}{Q} \quad \text{so} \quad S(t_0)\epsilon < \frac{1}{2}. \tag{56}$$

We split the interval $[0, t_0]$ into Q equal intervals of length ϵ. For $0 \leq i \leq Q$ we define

$$G^i = G_{ni\epsilon/2}\,,$$

the graph at time $i\epsilon$. Here we assume that $ni\epsilon/2$ is an integer; the effect of replacing it with the closest integer will clearly be negligible in the following argument.

Theorem 4.2 *For $0 \leq i \leq Q$ with probability approaching one*
- *G^i has a K, c component tail. Here K, c depend on i.*
- *$S(G^i) = S(i\epsilon) + o(1)$.*

With $i = Q$, as $t_0 < t_c$ was arbitrary, Theorem 4.2 will complete the subcritical behavior sections of our main result, Theorem 1.1. (Recall that the bound on the largest component of $G_{tn/2}$ in the fourth part of Theorem 1.1 follows immediately from (8).)

`Proof:` We use induction on i, $i = 0$ being immediate as G^0 is empty. Assume, by induction, that the hypotheses hold for a fixed $i < Q$. Let H be the graph consisting of *both* edges $\{v_1, v_2\}$ and $\{v_3, v_4\}$ for all rounds j, $ni\epsilon/2 < j \leq n(i+1)\epsilon/2$. Then G^{i+1} is a subgraph of $G \cup H$. Here, H can be taken as a random set of $n\epsilon - O(\log n)$ distinct non-loop edges. The $O(\log n)$ term accounts for loops and repeated edges, which may be ignored (it is easy to show that the number of these is bounded in probability). By standard methods we can treat H like a random graph with edge probability $\frac{2\epsilon}{n}$. To see this, let H' be a random graph with edge probability $\frac{2\epsilon}{n}$. The inductive assumption gives that $S(G^i) = S(i\epsilon) + o(1) < S(t_0)$ and our choice (56) of ϵ was made sufficiently small that the conditions for the subcritical case of Theorem 3.1 apply. Thus $G^i \cup H'$ a.a.s. has a K^+, c^+ component tail. For fixed G^i, the property of H', that $G^i \cup H'$ has a K^+, c^+ component tail, is monotone decreasing and hence convex. Thus, by [10, Fact 1] for example, $G^i \cup H$ a.a.s. has a K^+, c^+ component tail. Hence G^{i+1}, being a subgraph of $G^i \cup H'$, also has a K^+, c^+ component tail.

Now we want to extend the results of Section 4.1 to include the function S. The system (34,35) with initial conditions (33) has a unique solution $\vec{x}(t) = (x_1(t)\ldots, x_K(t), x_{\omega(t)}, S(t))$ which is defined for $t \in [0, t_0]$. The differential equations are of the form

$$\vec{x}'(t) = F(\vec{x}) \tag{57}$$

where $F : R^{n+2} \to R^{n+2}$ is a C^∞ function. The first $n+1$ coordinates (the x_i') have been described before and the final coordinate (the S') is a polynomial function of the x_i involving S and S^2. We define a discrete vector valued sequence of random variables $\vec{X}_j = (x_1(G_j), \ldots, x_K(G_j), x_\omega(G_j), S(G_j))$, where we restrict attention to the range $i\epsilon n/2 \leq j \leq (i+1)\epsilon n/2$. By Theorem 4.1 the initial values $\vec{X}_{i\epsilon n/2}$ and $\vec{x}(i\epsilon)$ are only $o(1)$ apart.

There is now a special difficulty in that we do not have a uniform bound on the change $|S(G_{j+1}) - S(G_j)|$ that we did for the x_i with (19). When the addition of a single edge merges components of sizes α, β the value of S increases by $\frac{2}{n}\alpha\beta$. Fortunately, we have already proven the first part of the induction hypothesis for $i+1$. We know that with probability $1-o(1)$ the graph G^{i+1} has a K, c component tail and so has all components of size $O(\ln n)$.

To take care of the $o(1)$ probability that G^{i+1} has a large component we employ what is sometimes called the coward's sequence. We modify the sequence $\vec{X}_j$ to $\vec{X}_j^*$ as follows. At the initial value $j = i\epsilon n/2$ they are equal. If the sequence G_j (stopping at $j = (i+1)\epsilon n/2$) never has a component of size bigger than $c' \ln n$ then the two sequences $\vec{X}_j, \vec{X}_j^*$ are equal. Otherwise, let j be the first value where G_j has a component of size bigger than $c' \ln n$. Then for $j \le j' < (i+1)\epsilon n/2$ we simply define

$$\vec{X}_{j'+1}^* = \vec{X}_{j'}^* + \frac{2}{n} F(\vec{X}_{j'}^*) \tag{58}$$

where F is given by the differential equation system (57).

The idea behind the coward's sequence can be described in quite general terms. We have a sequence $\vec{X}_j$ which follows the diffential equation system in expectation. Usually throughout the sequence there is never a j and a G_j so that $\vec{X}_{j+1}$ can be substantially different from $\vec{X}_j$. The coward's sequence is, in Mathematics of Finance jargon, risk averse. In a situation in which there might be a large change it "shuts down" and just follows the differential equation. The entire coward's sequence $\vec{X}_j^*$ shall usually equal the entire sequence $\vec{X}_j$ as this risk rarely occurs. The coward's sequence never has large changes and from that we shall deduce that it will rarely stray far from the differential equation. As the sequences are usually completely equal we deduce that the original sequence will rarely stray far from the differential equation.

We need to prove that with probability $1-o(1)$ the final value $\vec{X}_{(i+1)n\epsilon/2}$ is within $o(1)$ of the value of the differential equation $\vec{x}((i+1)\epsilon)$. With probability $1-o(1)$ the final values $\vec{X}_{(i+1)n\epsilon/2}$ and $\vec{X}_{(i+1)n\epsilon/2}^*$ are the same, as this occurs when G^{i+1} does not have a component of size bigger than $c' \ln n$. So it suffices to prove that with probability $1-o(1)$ the final value $\vec{X}_{(i+1)n\epsilon/2}^*$ is within $o(1)$ of the value of the differential equation $\vec{x}((i+1)\epsilon)$.

We claim

$$\left| \frac{E[\vec{X}_{j+1}^* - \vec{X}_j^* | G_j]}{2/n} - F(\vec{X}_j^*) \right|_\infty = O(n^{-1} \ln^3 n). \tag{59}$$

When G_j has a component of size bigger than $c' \ln n$ the left hand side is zero, from our definition (58) of $\vec{X}_{j+1}^*$. Otherwise (the main case) $\vec{X}^* = \vec{X}$. In (55) we have shown that the left hand side vector has all coordinates at most $2K^2 n^{-1}$ except for the S coordinate which we now examine. Let's employ the notation of Section 2.2 with $S = S(G_j), S^+ = S(G_{j+1})$, S^* being what S^+ "would be" ignoring the "error" e_S. Combining (21,22,35) we have that

$$\frac{E[S^+ - S]}{2/n} - F(\vec{X}) = \frac{E[e_S]}{2/n} \tag{60}$$

and from (23)

$$\left| \frac{E[e_S]}{2/n} \right| \le \nabla(G_j). \tag{61}$$

Indeed, we have chosen the notation $\nabla(G)$ in (7) to correspond to differential inequalities in the mathematical study of percolation. Here we are in the situation where the largest component of G_j has size $O(\ln n)$. When all $|C_i| \le u$ a simple convexity argument gives $\sum |C_i|^4 \le nu^3$. Thus

$$|\nabla(G_j)| = O(n^{-1} \ln^3 n)$$

which yields (59).

We further claim that

$$|\vec{X}^*_{j+1} - \vec{X}^*_j|_\infty = O(n^{-1}\ln^2 n) \tag{62}$$

provided that S is bounded above by some constant. When G_j has a component of size bigger than $c'\ln n$ the left hand side is precisely $\frac{2}{n}F(\vec{X}^*_j)_\infty$ which is $O(n^{-1})$ as all the coordinates of $\vec{X}^*_j$ lie in a bounded region. Otherwise (the main case) $\vec{X}^* = \vec{X}$. From (55) we know that the left hand side vector has all coordinates $O(n^{-1})$ except, perhaps, the S coordinate. As G_j has all components of size $O(\ln n)$ a single edge can only change S by $n^{-1}\ln^2 n$, which yields (62).

We may now apply [8, Theorem 5.1] as in Section 4.1. To achieve a suitable upper bound $1/\epsilon$ on S we use the domain $\mathcal{D} = [0,t_0] \times [0,1]^{K+1} \times [0,1/\epsilon]$, and as before let D be a bounded open set containing $\mathcal{D}$. This time, from (59) we have $\lambda_1 = O(\ln^3 n)$ and, from (62), $\beta = O(\ln^2 n)$. It follows from the theorem that the system of differential equations given by (34,35) with initial conditions (33) has a unique solution $(t, x_1, \ldots, x_K, x_\omega, S)$ in $\mathcal{D}$ which extends outside of $\mathcal{D}$ beyond some point at which $t = \sigma_0$ say, on the boundary of $\mathcal{D}$. This we already deduced in Section 2.4, but the point now is that, from (56), $S(t_0) = \epsilon/2$ and so, in view of Theorem 2.2, the boundary reached cannot be $S = \epsilon$. Thus $\sigma_0 = t_0$. We now obtain the desired concentration by taking $\lambda = n^{-1/4}$ in the terminology of [8, Theorem 5.1]. ∎

5 The Supercritical Phase

In this section we fix $\epsilon > 0$. Our object is to show that $G_{(t_c+\epsilon)n/2}$ will have a "giant component".

First select some $t^* \in (0, t_c)$, we can select $t^* = \frac{1}{2}t_c$ for definiteness. Set $\beta = x_\omega(t^*)$. Now, with foresight, select $t^- \in (t^*, t_c)$ such that

$$(S(t^-) - K)\epsilon\beta^4 > 1.$$

Let G^- denote $G_{nt^-/2}$. From our analysis of the subcritical case we know that a.a.s. $S(G^-) = S(t^-) + o(1)$ and that $x_\omega(G^-) > \beta$. Let W denote the set of vertices v with $|C(v)| > K$ and let G denote the restriction of G^- to W. Let $m = |W|$ so that $m > \beta n$. Furthermore

$$S(G) = \frac{n}{m}S_\omega(G^-) \geq S_\omega(G^-) \geq S(G^-) - K.$$

Consider the $\epsilon\frac{n}{2}$ rounds j with $t_c\frac{n}{2} < j \leq (t_c+\epsilon)\frac{n}{2}$. Call a round *good* if all four selected vertices $v_1, v_2, v_3, v_4 \in W$. (Clarification: The set W is set at time t^-, we do not enlarge it when the components of other vertices become large.) Each round is good with independent probability $(m/n)^4 > \beta^4$. There will be more than $\epsilon\beta^4\frac{n}{2} > \epsilon\beta^4\frac{m}{2}$ good rounds. Conditioning on a round being good the v_1, v_2, v_3, v_4 are independent and uniform over W. Suppose $(\omega,\omega,\omega,\omega) \in F$, the other case being identical. Let H be the graph on W consisting of all edges $\{v_1, v_2\}$ from all good rounds j in this region. Then

$$G_{n(t_c+\epsilon)}|_W \supseteq G \cup H$$

where H is a random graph on W with more than $\epsilon\beta^4\frac{m}{2}$ edges. As in the corresponding part of our analysis of the subcritical phase, we can assume using [10, Fact 1] that H is a random graph with edge probability $\epsilon\beta^4/m$. This time, we use the fact that for a fixed graph G, the property of H, that $G \cup H$ contains a component of size γm, is convex.

We now apply the supercritical section of Theorem 3.1. $G \cup H$, and hence $G_{n(t_c+\epsilon)}|_W$ and hence $G_{n(t_c+\epsilon)}$ contains a component of size γm. This size is at least $\gamma\beta n$ and hence is the desired giant component of Theorem 1.1.

6 Numerical results

To obtain information on t_c for a given K and F, we solve the equations (34) and (37) numerically. These equations are rewritten below for the purpose of rigorously bounding errors. Various sets (or "rules") F were tried.

Method of computation

Let $\hat{F}$ denote the mulitset obtained from combining the sets of vectors $(j_1, j_2, j_3, j_4) \in F$ and $(j_3, j_4, j_1, j_2) \notin F$ without discarding duplicates. The differential equations (34) can be written as

$$x_i' = f_i = \sum_{\vec{j} \in \hat{F}} \frac{i}{2} \Big(\delta_{i=j_1+j_2} - \delta_{i=j_1} - \delta_{i=j_2} \Big) x_{j_1} x_{j_2} x_{j_3} x_{j_4} \tag{63}$$

$(i \neq \omega)$,

$$x_\omega' = f_\omega = \sum_{\vec{j} \in \hat{F}} \frac{1}{2} \Big(\delta_{\{j_1+j_2>K,\ j_1 \neq \omega,\ j_2 \neq \omega\}} (j_1 + j_2) + \delta_{\{j_1 \neq \omega,\ j_2 = \omega\}} j_1 + \delta_{\{j_1 = \omega,\ j_2 \neq \omega\}} j_2 \Big) x_{j_1} x_{j_2} x_{j_3} x_{j_4} \tag{64}$$

where, for a set of statements H, $\delta_H = 1$ if the statements in H are all true and $\delta_H = 0$ otherwise. Similarly, considering the discussion of I_ω after (37),

$$\begin{aligned} S_\omega' = f_\alpha \quad = \quad & \sum_{\vec{j} \in \hat{F}} \Bigg(\delta_{\{j_1+j_2>K,\ j_1 \neq \omega,\ j_2 \neq \omega\}} \frac{1}{2} (j_1 + j_2)^2 x_{j_1} x_{j_2} + \delta_{j_1 \neq \omega,\ j_2 = \omega} \big(\frac{1}{2} j_1^2 x_{j_1} x_\omega + j_1 x_{j_1} S_\omega \big) \\ & + \delta_{j_1 = \omega,\ j_2 \neq \omega} \big(\frac{1}{2} j_2^2 x_{j_2} x_\omega + j_2 x_{j_2} S_\omega \big) + \delta_{j_1 = j_2 = \omega} S_\omega^2 \Bigg) x_{j_3} x_{j_4}. \end{aligned}$$

Use $\vec{x} = (x_1, \ldots, x_K, x_\omega, x_\alpha)$ where for convenience x_α is defined to equal S_ω. Then an arbitrary variable is x_i where $i \in \Omega^+ = \Omega \cup \{\alpha\}$. To solve the equations by Euler's method, set $x_i^*(0) = x_i(0)$ $(0 \leq i \leq k)$ and then, given $\vec{x}^*(t) = (x_0^*, \ldots, x_k^*)$, compute

$$x_i^*(t + h) = x_i^*(t) + h f_i(\vec{x}^*(t)), \quad i \in \Omega^+. \tag{65}$$

This is iterated for $t = 0, h, 2h, \ldots$. Euler's method is analyzed below for rigorous results.

To speed up computations, we may precompute for each pair (j_1, j_2), the sum of $x_{j_3} x_{j_4}$ over all (j_3, j_4) such that $\vec{j} \in F$ and over all (j_3, j_4) such that $(j_3, j_4, j_1, j_2) \in F$. By running once through all pairs (j_1, j_2), this enables the whole set of derivatives to be computed in time $O(K^2)$.

Approximate results

We solved the differential equations (34) and (37) numerically with initial values given by (33) and $S_\omega(0) = 0$. For large k we used a second order Runge-Kutta method without error analysis but with excellent convergence apparent from running with various step sizes, for various bounded size algorithms. The results are shown in Table 1. This shows the approximate value of t (rounded) at which S_ω reaches 10^4; very shortly after, it will go to infinity by (38). Here "minp" is an algorithm which chooses the pair of vertices such that the product of their component sizes is minimised, that is, it is an F-algorithm in which

$$F(j_1, j_2, j_3, j_4) = \begin{cases} 1 & \text{if } j_1 j_2 < j_3 j_4 \\ 0 & \text{if } j_1 j_2 > j_3 j_4. \end{cases}$$

In the case of a tie $(j_1 j_2 = j_3 j_4)$ the value can clearly be toyed around with but does not seem to make much difference to the results. For the results below, we used 1 iff $j_1 + j_2 \geq j_3 + j_4$. Here and in the next case, if $j_i = \omega$ it is treated as infinite.

The algorithm "minh" minimises the harmonic mean of the component sizes, i.e. maximises the sum of the reciprocals (with the reciprocal of ω treated as 0), and in the case of a tie, minimises the sum of the component sizes. The algorithm "minl", the best performing one we found for getting rigorous bounds for small K, minimises the sum of the third iteration of $1 + \log x$ applied to the component sizes, that is,

$$F(j_1, j_2, j_3, j_4) = \begin{cases} 1 & \text{if } \ln(1+\ln(1+\ln j_1))) + \ln(1+\ln(1+\ln j_2))) \\ & \qquad \leq \ln(1+\ln(1+\ln j_3))) + \ln(1+\ln(1+\ln j_4))) \\ 0 & \text{otherwise} \end{cases}$$

(but here for use with large K, any value ω is evaluated as $K+1$: this gives "better" results, probably because taking ω as infinity penalises the large components too much).

These are the bounded size algorithms we found which delay the birth of the giant the most; minh and minl seem to do this best, and almost equally well. However it is extremely unlikely that there is no better bounded size algorithm for this.

K	minp	minh	minl
50	1.7425	1.7573	1.7576
100	1.7556	1.7679	1.7681
200	1.7639	1.7741	1.7742
400	1.7691	1.7778	1.7777
800	1.7723	1.7799	1.7798
1600	1.774	1.781	1.781

Table 1. Approximate values of t_c (rounded) for various bounded size algorithms.

In the next table we consider algorithms which aim at accelerating the birth of the giant. This time, the best we tried for small K was maxp, which maximises the product of the component sizes.

K	maxp	maxh	maxl
50	0.6546	0.6571	0.6568
100	0.6544	0.6571	0.6566
200	0.6542	0.6571	0.6565
400	0.6540	0.6571	0.6565

Table 2. Approximate values of t_c for algorithms accelerating the birth of the giant.

Rigorous results with error bounds

When computed by machine, we actually have in place of (65)

$$x_i^*(t+h) = x_i^*(t) + hf_i(\vec{x}^*(t)) + \rho_i(t) \tag{66}$$

where ρ_i is the rounding error due to floating point approximation in machine computation.

We first examine the truncation error of the method, i.e. $\tau_i(t) = x_i(t+h) - x_i(t) - hf_i(\vec{x}(t))$. By Taylor's theorem,

$$\tau_i(t) = \frac{h^2}{2}\ddot{x}_i(\xi_i) \tag{67}$$

for some $t \leq \xi_i \leq t+h$, and we have

$$\ddot{x}_i(\xi_i) = \frac{d}{dt} f_i(\vec{x}(t)) \mid_{t=\xi_i} = \sum_{r \in \Omega^+} \frac{\partial f_i}{\partial x_r} f_r \mid_{\vec{x}=\vec{x}(\xi_i)} . \tag{68}$$

(Here we used the fact that the differential equations are autonomous: f_i is independent of t.)

It would be possible to compute fairly accurate bounds on these functions during execution of an algorithm, but the number of terms to be considered would excessively slow down the computation when K is large unless considerable programming effort was employed. Instead, to obtain approximate results, it is enough to use bounds on the absolute value of the f_i and their partial derivatives, and then to replace each summation over $\hat{F}$ by twice the sum of $\vec{j} \in \Omega$. Along the way, we know that all x_i are nonnegative since the only negative terms in f_i have factor x_i. We also know that $\sum_{i\in\Omega} x_i = 1$ since the derivative of this sum is 0 and its initial value is 1.

Doing this to the expression in (63) for $1 \le r \le K$, the contribution from the case $r = j_1$ is

$$\sum_{j_2,j_3,j_4\in\Omega} \Big(\delta_{i=r+j_2} i + \delta_{i=r} i + \delta_{i=j_2} j_2\Big) x_{j_2} x_{j_3} x_{j_4} = \delta_{i>r} i x_{i-r} + \delta_{i=r} i + i x_i.$$

The same bound applies for the case $r = j_2$. For $r = j_3$ it is

$$\sum_{j_1,j_2,j_4\in\Omega} \Big(\delta_{i=j_1+j_2}(j_1+j_2) - \delta_{i=j_1} j_1 - \delta_{i=j_2} j_2\Big) x_{j_1} x_{j_2} x_{j_4} = i \sum_{j<i} x_j x_{i-j} + 2 i x_i,$$

and this bound also applies for the contribution from $r = x_4$. Summing these and writing

$$T(i) = T(i,\vec{x}) = \sum_{j<i} x_j x_{i-j}$$

gives

$$\left|\frac{\partial f_i}{\partial x_r}\right| \le \delta_{i>r} 2 i x_{i-r} + \delta_{i=r} 2i + 6 i x_i + 2 i T(i).$$

This applies for all $i, r \in \Omega \setminus \{\omega\}$. Similar calculations give

$$\left|\frac{\partial f_i}{\partial x_\omega}\right| \le 4 i x_i + 2 i T(i) \quad (1 \le i \le K).$$

For the next cases we will also use

$$P_i^{(r)} = \sum_{j=r}^{K} j^i x_j$$

for $i = 0$, 1 and 2, and

$$P_i = P_i^{(1)}, \qquad Q_i = \sum_{j_1,j_2\in\Omega} \delta_{\{j_1+j_2>K,\ j_1\neq\omega,\ j_2\neq\omega\}} (j_1+j_2)^i x_{j_1} x_{j_2}$$

for $i = 1$ and 2. Then

$$\begin{aligned}
\left|\frac{\partial f_\omega}{\partial x_r}\right| &\le 2P_1^{(K-r+1)} + 2rP_0^{(K-r+1)} + 2r x_\omega + 2Q_1 + 4P_1 x_\omega \quad (1 \le j \le K),\\
\left|\frac{\partial f_\omega}{\partial x_\omega}\right| &\le 2P_1(1+2x_\omega) + 2Q_1\\
\left|\frac{\partial f_\alpha}{\partial x_r}\right| &\le 2P_2^{(K-r+1)} + 4rP_1^{(K-r+1)} + 2r^2P_0^{(K-r+1)} + 2r^2 x_\omega\\
&\quad +4rS_\omega + 2Q_2 + 2P_2 x_\omega + 4P_1 S_\omega + 4S_\omega^2 \quad (1 \le r \le K),\\
\left|\frac{\partial f_\alpha}{\partial x_\omega}\right| &\le 2P_2 + 2Q_2 + 4P_2 x_\omega + 8P_1 S_\omega + 4S_\omega^2,\\
\left|\frac{\partial f_\alpha}{\partial S_\omega}\right| &\le 4P_1 + 4S_\omega.
\end{aligned}$$

In a similar way we obtain

$$\begin{aligned}
|f_r| &\leq rT(r) + 2rx_r \quad (1 \leq r \leq K), \\
|f_\omega| &\leq Q_1 + 2P_1 x_\omega, \\
|f_\alpha| &\leq Q_2 + 2P_2 x_\alpha + 4P_1 S_\omega + 2S_\omega^2.
\end{aligned} \tag{69}$$

To use (68), we need to use the above bounds at $\vec{x}(\xi_i)$ when what is known is only the value of $\vec{x}^*$. It follows from (69) that

$$|f_r| \leq 3r \quad (1 \leq r \leq K) \tag{70}$$

and

$$|f_\alpha| \leq 4K^2 + 4KS_\omega + 2S_\omega^2 = 2K^2 + 2(K + S_\omega)^2. \tag{71}$$

Now write $e_i(t) = x_i^*(t) - x_i(t)$ and suppose that we have already computed bounds $b_i(t) \geq |e_i(t)|$ for all $i \in \Omega^+$. Assuming all x_i^* are nonnegative (which can be checked dynamically), we now have

$$x_r(\xi_i) \leq x_r^* + b_r(t) + 3rh \qquad (1 \leq r \leq K),$$

and from (69)

$$x_\omega(\xi_i) \leq x_\omega^* + b_\omega(t) + 2Kh.$$

One the other hand, from (71), x_α is bounded above by y where $y' = 2K^2 + 2(K + y)^2$ with initial condition $y(t) = x_\alpha^*(t) + b_\alpha(t)$. Hence

$$x_\alpha(\xi_i) \leq K \tan(2Kh + C)) - K \tag{72}$$

where $C = \arctan(1 + (x_\alpha^*(t) + b_\alpha(t))/K)$. Naturally, this requires the argument of tan to be less than $\pi/2$; if not, the error becomes unbounded.

With these bounds on x_r, the values of P_1 and P_2 can be bounded above and used in the bounds on derivatives and partial derivatives above, and this produces a bound via (68) and (67) on the error $|\tau_i(t)|$.

Furthermore, by (66),

$$\begin{aligned}
x_i^*(t+h) &= x_i(t) + e_i(t) + hf_i(\vec{x}^*(t)) + \rho_i(t) \\
&= x_i(t) + hf_i(\vec{x}(t)) + h\Delta_i(t) + \rho_i(t) + e_i(t) \\
&= x_i(t+h) - \tau_i(t) + h\Delta_i(t) + \rho_i(t) + e_i(t)
\end{aligned}$$

where $\Delta_i(t) = f_i(\vec{x}^*(t)) - f_i(\vec{x}(t))$. Hence

$$|e_{i+1}(t+h)| \leq |\tau_i(t)| + h|\Delta_i(t)| + |\rho_i(t)| + |e_i(t)|.$$

It is a straightforward, but tedious, matter to compute upper bounds on $\Delta_i(t)$, given $\vec{x}^*(t)$ and upper bounds $b_i(t)$ on $|e_i(t)|$ for each $i \in \Omega^+$. Including numerically computed upper bounds on $\rho_i(t)$ based on the machine accuracy permits recursive computation of $b_i(t)$ along with the computation of $x_i^*(t)$ for $t = 0, h, 2h, \ldots$. A lower bound on t_c is given by the first t examined such that, in (72), $2Kh + C \geq \pi/2$. On the other hand, the analysis following (38) shows that an upper bound is $t + (S_\omega(t) - b_\alpha(t))^{-1}(x_\omega(t) - b_\omega(t))^{-2}$ for each t.

The approximate size of the errors apart from ρ_i tends to be of the form $c_1 e^{c_2 k} h$ for constants c_1 and c_2. The best value of h is determined by a tug-of-war between such a function and the accumulation of the floating point errors of the machine which appear as ρ_i.

Rigorous results with error bounds

The rigorous results obtained by the method described above are shown in Table 3, with lower bounds rounded down and upper bounds rounded up. The assumptions involved in this are that the machine used correctly computes double precision floating point operations to the accuracy that it should, and that there is no error in our programming. Note that for small values of K, some of the algorithms will behave identically. For algorithms which aim to retard the birth of the giant, better lower bounds were obtained for quite large values of K. For maxp, the best found was for $K = 5$, and for maxh, at $K = 4$.

K	algorithm	lower bound	upper bound
1	minp, minh, minl	1.3933	1.3936
2	minp, minh, minl	1.5110	1.5115
	minl	1.5171	1.5176
4	minp	1.5974	1.6035
	minh	1.6125	1.6183
	minl	1.6140	1.6191
8	minp	1.6295	1.7838
	minh	1.6563	1.7674
	minl	1.6587	1.7624
1	maxp,maxh,maxl	0.6887	0.6902
4	maxh	0.6622	0.6689
5	maxp	0.6596	0.6671

Table 3. Rigorous bounds on t_c for various algorithms.

7 Conjectures, Speculations and Comments

Let $L_1(G)$ denote the size of the largest component of G. In their classic work Erdős and Rényi showed that a.a.s. $L_1(G(n, \frac{t}{n}) = n(F(t) + o(1))$. The function $F(t)$ was described completely. For $t < 1$ (the subcritical case) $F(t)$ was zero. The phase transition, in modern language, was not first order. That is, $\lim_{t\to 1^+} F(t) = 0$. We conjecture similar behavior for any bounded size algorithm. That is, we conjecture the existence of a function F such that a.a.s. $L_1(G_{nt/2}) = n(F(t) + o(1))$. From our main result, Theorem 1.1, we would necessarily have $F(t) = 0$ in the subcritical case $t < t_c$ and $F(t) > 0$ in the supercritical case $t > t_c$. We conjecture furthermore that the phase transition is not first order, that $\lim_{t\to t_c^+} F(t) = 0$. We may rephase this conjecture so that it does not presuppose the existence of a limiting function F:

`Conjecture:` For all $\alpha > 0$ there exists $t > t_c$ so that a.a.s $L_1(G_{nt/2}) \leq n\alpha$.

Let $L_2(G)$ denote the size of the second largest component of G. Erdős and Rényi showed that in the supercritical case $t > 1$, $L_2(G(n, \frac{t}{n})) = O(\ln n)$ a.a.s. We conjecture, analogously but perhaps rashly, that the same holds for any bounded size algorithm:

`Conjecture:` For all $t > t_c$ there exists K so that a.a.s. $L_2(G_{nt/2}) \leq K \ln n$.

We understand that the critical window for the Erdős-Rényi evolution is given by the parametrization $p = n^{-1} + \lambda n^{-4/3}$. In the precritical region, $\lambda \to -\infty$, the contribution to the susceptibility from any single component is negligible. In the postcritical region, $\lambda \to +\infty$, the contribution to the susceptibility is dominated by the largest, often call dominant, component. Inside the critical window the largest components are all of size $\Theta(n^{2/3})$ and they all contribute substantially to the susceptibility. We may naturally ask look for similar results for bounded size algorithms. In particular, will there be a similar scaling $t_c n^{-1} + \lambda n^{-4/3}$ to describe their behavior.

The restriction to *bounded* size algorithms may not be fully necessary. We can naturally define a size algorithm as any algorithm where the choice of edge selected depends only on the sizes of the components of the four vertices. One natural algorithm is the "product rule" suggested by Dimitris Achlioptas. With this rule one selects that edge which minimizes the product of the component sizes of its components. This has the nice property that, at least locally, it minimizes the susceptibility.

Is a size algorithm in some sense the limit of bounded size algorithms? More explicitly, given a size algorithm A let A_K be any bounded size algorithm in which we follow the size algorithm when all four vertices are in components of size at most K. Let t_K be the critical point of the algorithm A_K. We conjecture that for any choices of the A_K the t_K approach a limit t_c and that t_c is the critical point for the algorithm A. Our numerical results suggest approximate values of such limits but do not seem to show anything concrete about them.

One may also ask about the critical window for size algorithms. We strongly suspect that the behavior will not (at least, not in all cases) be similar to that of the Erdős-Rényi evolution. For example, we suspect that in the product rule evolution will lead to a much narrower critical window in which there will be two components of size substantially bigger than $n^{2/3}$. Achlioptas has suggested that the percolation may not be first order: There may be a positive constant α so that for at time $t = t_c + \epsilon$, with ϵ an arbitrarily small constant, there already exists a giant component of size αn. Computer simulation lends credence to this possibility, which remains most speculative.

Finally, we consider the difficulty in extending our argument to cover arbitrary algorithms, not bounded size nor even size algorithms. Differential equations cannot be applied to the subcritical phase unless the algorithm behaves smoothly enough. For the supercritical phase, one might ask if the supercritical part of Theorem 3.1 still applies. The answer is no. To make this clear, we give the following precise result.

Lemma 7.1 *For any fixed $L > 0$ and $1/2 > \epsilon > 0$ there is a G_0 with $S(G_0) > L$, and an algorithm, for which the component size in the Achlioptas process after ϵn steps is a.a.s. $O(\log n)$.*

`Proof (sketch):` Take G_0 such that all components have size 1 or M, and the average squared component size is L, where to understand the situation we may imagine $1 \ll L \ll M$, but of course L and M are constants whilst the number n of vertices of G_0 tends to ∞. Then the number of M-components is approximately $LM^{-2}n$.

Consider the following algorithm: when a pair of edges is presented, give preference to the edge which, *if added to* G_0 would create a smaller component. Thus, in choosing beteen the two, ignore all edges apart from those in G_0. Break ties randomly.

Now consider the graph of chosen edges after ϵn steps. Noting that the number of vertices in M-components is approximately $LM^{-1}n$ which is a very small fraction of n, we can say that the number of edges between isolates of G_0 is approximately ϵn, the number between isolates and M-components is approximately $\epsilon(2L/M)^2 n$, and the number with both ends in M-components is approximately $\epsilon(L/M)^4 n$. Furthermore, conditioning on the number of edges within any one of these classes, the edges occur uniformly at random.

Thus, we can model the graph at time ϵn by taking each component of G_0 as a vertex of a graph F, edges occurring with probability $p_0 \approx 2\epsilon/n$ between isolates of G, probability $p_1 \approx \epsilon(2L/M)^2 n/(LM^{-2}n^2) = 4\epsilon L/n$ from an isolate to a vertex of F standing for an M-component, and $p_2 \approx \epsilon(L/M)^4 n/2(LM^{-2}n)^2 = \epsilon L^2/2n$ between two vertices of F standing for M-components.

Now the expected degree of an isolate of G_0 in F is $d_1 + d_2$ where $d_1 \approx np_0 \approx 2\epsilon$ is for edges to other isolates and $d_2 \approx (L/M^2)np_1 \approx 4\epsilon L^2/M^2$ is for edges to M-components. As $\epsilon < 1/2$ (and by taking large M) this is less than some constant less than 1. The expected number of edges in F from a given M-component of G_0 to other M-components is $d_3 \approx p_2 LM^{-2}n \approx \epsilon L^3/2M^2$. By taking M sufficiently large, this can be made much smaller than any desired positive constant. Thus the expected degree of the isolate is bounded above by a constant less than 1.

The only problem is the expected number d_4 of edges from a given M-component to isolated vertices of G_0; $d_4 \approx np_2 \approx \epsilon L^2/2$. However, since $d_2 d_4 \approx 2\epsilon^2 L^4/M^2$ we can choose M large enough for this to be arbitrarily small. Then we may consider the component of F containing a given M-component by taking breadth-first search and moving to any M-components reachable by a path consisting of 1-components in one step. The number of children in each step has expected value less than 1 and has a K, c component tail. By Theorem 3.2 the components of F will also have such a tail, and the result follows on expanding each vertex of F to a component of G_0 (remembering M is fixed). ∎

Acknowledgements: Svante Janson made a very significant input at the start of our researches by suggesting to focus on the sum of squares of component sizes and noting the differential equation for susceptibility in the Erdős-Rényi case. Yuval Peres noted the analogies to classical percolation given in Section 1.4 and it was these analogies that led us to the fruitful examination of graphs with a K, c component tail.

References

[1] N. Alon and J. Spencer, *The Probabilistic Method*, 2nd ed., John Wiley, 2000.

[2] T. Bohman, A. Frieze, Avoiding a giant component, *Random Structures & Algorithms* **19** (2001), 75-85

[3] T. Bohman, A. Frieze and N.C. Wormald, Avoidance of a giant component in half the edge set of a random graph, manuscript.

[4] T. Bohman and D. Kravitz, Creating a giant component, manuscript.

[5] A. Flaxman, D. Garmanik and G.B. Sorkin, Embracing the giant component, pp. 69–79 in *Latin 2004: Theoretical Informatics*, M. Farach-Coltin (ed). Lecture Notes in Computer Science 2976, Springer, 2004.

[6] W. Hurewicz, *Lectures on Ordinary Differential Equations*, M.I.T. Press, Cambridge Massachusetts (1958).

[7] D. Reimer, Proof of the van der Berg-Kesten conjecture, *Combinatorics Probability & Computing* **9** (2000), 27-32.

[8] N.C. Wormald, The differential equation method for random graph processes and greedy algorithms. In *Lectures on Approximation and Randomized Algorithms*, M. Karoński and H.J. Prömel (eds), pp. 73–155. PWN, Warsaw, 1999.

[9] N.C. Wormald, Analysis of greedy algorithms on graphs with bounded degrees, *Discrete Mathematics*, **273** (2003), 235–260.

[10] N.C. Wormald, Random graphs and asymptotics. Section 8.2 in *Handbook of Graph Theory*, J.L. Gross and J. Yellen (eds), pp. 817–836. CRC, Boca Raton, 2004.