The evolution of subcritical Achlioptas processes

Oliver Riordan and Lutz Warnke*

April 23, 2012

Abstract

In Achlioptas processes, starting from an empty graph, in each step two potential edges are chosen uniformly at random, and using some rule one of them is selected and added to the evolving graph. Although the evolution of such 'local' modifications of the Erdős–Rényi random graph process has received considerable attention during the last decade, so far only rather simple rules are well understood. Indeed, the main focus has been on 'bounded-size' rules, where all component sizes larger than some constant B are treated the same way, and for more complex rules very few rigorous results are known.

In this paper we study Achlioptas processes given by (unbounded) size rules such as the sum and product rules. Using a variant of the neighbourhood exploration process and branching process arguments we show that certain key statistics are tightly concentrated at least until the susceptibility (the expected size of the component containing a randomly chosen vertex) diverges. Our convergence result is most likely best possible for certain rules: in the later evolution the number of vertices in small components may not be concentrated. Furthermore, we believe that for a large class of rules the critical time where the susceptibility 'blows up' coincides with the percolation threshold.

1 Introduction

In 2000 Dimitris Achlioptas suggested a class of variants of the classical random graph process, now called *Achlioptas processes*. Such a process defines, for each n, a random sequence $(G_i)_{i\geqslant 0} = (G_i^{\mathcal{R}})_{i\geqslant 0}$ of graphs with vertex set $[n]$ as follows: start with an empty graph G_0 on n vertices. At each step $i \geqslant 1$, two potential edges e_1 and e_2 are chosen independently and uniformly at random from all $\binom{n}{2}$ possible edges (or from those edges not present in G_{i-1}). One of these edges is selected according to a rule $\mathcal{R}$ and added to the graph, so $G_i = G_{i-1} \cup \{e\}$ for $e = e_1$ or e_2. (As usual in combinatorics, we omit the dependence on n in the notation to avoid clutter.) Achlioptas processes are a special case of the more general class of *ℓ-vertex rules*, where in each step

*Mathematical Institute, University of Oxford, 24–29 St Giles', Oxford OX1 3LB, UK. E-mail: {riordan,warnke}@maths.ox.ac.uk.

$\ell \geqslant 2$ vertices $v_1, \ldots, v_\ell$ are chosen independently and uniformly at random and then at least (usually exactly) one edge between these vertices is added. Always adding $e = e_1$ (or $e = v_1v_2$) gives the Erdős–Rényi random graph process, which has been extensively studied for more than 50 years; by now many of its properties are extremely well understood, in particular the evolution of the component structure, see e.g. [8, 9, 11].

During the last decade the evolution of certain 'simple' Achlioptas processes has received considerable attention, mainly for so-called *bounded-size* rules, see e.g. [3, 4, 6, 12, 13, 18, 21]. These make their decisions based only on the sizes of the components containing the endvertices of e_1 and e_2, with the restriction that all sizes larger than some constant B are treated in the same way. For bounded-size rules a number of results have been established concerning concentration of the number of vertices in components of fixed size [6, 21], the size of the second largest component [18], the existence and location of the percolation phase transition where the (unique) linear size 'giant' component first emerges [6, 21], the 'order' of the phase transition [18], and the rescaled size of the largest component [18]. For one particular bounded-size rule (a variant of a process suggested by Bohman and Frieze [4]), some finer details of the percolation phase transition have recently been investigated [3, 12, 13]. The punchline of the work mentioned above is that bounded-size rules seem to show many qualitative similarities with the 'classical' Erdős–Rényi random graph process; in the language of mathematical physics they appear to be in the same 'universality class'.

In contrast, for more involved Achlioptas processes very few rigorous results are known, although these have been widely studied in recent years. This in particular applies to the class of 'unbounded'-size rules, usually simply called *size rules*, whose choices depend only on the sizes of the four components containing the endvertices of the two offered edges. To illustrate our very limited understanding of these, we mention that in one line of research, stimulated by a conjecture of Achlioptas, D'Souza and Spencer published in *Science* [1] (based on 'conclusive numerical evidence'), it was believed that certain size rules (in particular the *product rule*) could give rise to a discontinuous ('first order') phase transition. However, recently it was rigorously shown in [18, 20] that the phase transition is in fact continuous for all Achlioptas processes (even for a larger class of processes). The surprises that (unbounded) size rules have shown so far indicate that our intuition for these processes still needs to be developed. In fact, obtaining non-trivial results for involved Achlioptas processes using e.g. the product rule is well known to be a technical challenge (see e.g. [1, 10]).

Given a graph G, let $N_k(G)$ denote the number of vertices of G in components of size k, and define the *susceptibility* of G as $S(G) = \sum kN_k(G)/n$, so $S(G)$ is the expected size of the component containing a randomly chosen vertex. Let $L_1(G)$ denote the number of vertices in the (a, if there is a tie) largest component of G. We say that $t_c = t_c^{\mathcal{R}}$ is the *percolation threshold* for the process $(G_i^{\mathcal{R}})$ if for $t < t_c$ whp $L_1(G_{tn}) = o(n)$ while for $t > t_c$ whp $L_1(G_{tn}) = \Omega(n)$ (as usual, we henceforth ignore the irrelevant rounding to integers, and say that an event holds *whp*, if it holds with probability tending to 1 as $n \to \infty$). The anal-

ysis for bounded-size rules in [6, 12, 21] uses $S(G_i)$ as well as $N_k(G_i)$ for fixed k as key statistics. In the 'subcritical' regime $t < t_c$, they use Wormald's differential equation method [22, 23] to establish the existence of functions $\rho_k = \rho_k^{\mathcal{R}}$, $s = s^{\mathcal{R}}$ such that $N_k(G_{tn}) \approx \rho_k(t)n$ and $S(G_{tn}) \approx s(t)$ hold whp. Based on this they show that $t_c = t_c^{\mathcal{R}}$ is given by the blow-up point of the susceptibility: $\lim_{t \nearrow t_c} s(t) = \infty$. As indicated by Spencer and Wormald [21], for general size rules the approximation of the key statistics using the differential equation method seems difficult. Another intricacy for such rules is that the dependencies among the selected edges often seem to be more complex in comparison with bounded-size rules, for which a large subset of the added edges can be thought of as chosen uniformly at random (essentially, when all four endvertices are in components of size larger than B); we return to this in Section 3.

1.1 Main result

In this paper we establish the first rigorous convergence result for Achlioptas processes using unbounded size rules such as the product rule: we show that the number of vertices in components of size $k \geqslant 1$ (and the susceptibility) is tightly concentrated until the susceptibility 'blows up', which happens at a critical time t_b. In fact, our result holds for a very large class of Achlioptas-like processes, including essentially all Achlioptas processes studied so far (see Section 2 for the formal definition of ℓ-vertex size rules). Here $S(G_{tn}^{\mathcal{R}}) \xrightarrow{p} \infty$ as $n \to \infty$ means that for any $C > 0$ we have $\mathbb{P}(S(G_{tn}^{\mathcal{R}}) \leqslant C) \to 0$ as $n \to \infty$, and $N_{\geqslant k}(G)$ denotes the number of vertices of G in components of size at least k.

Theorem 1. *Let $\ell \geqslant 2$ and let $\mathcal{R}$ be an ℓ-vertex size rule. There exist $t_b = t_b^{\mathcal{R}} \in [\frac{1}{\ell(\ell-1)}, 1]$ and functions $(\rho_k)_{k \geqslant 1}$ with $\rho_k = \rho_k^{\mathcal{R}} : [0, t_b) \to [0, 1]$ such that the following holds. For every $t \geqslant t_b$ we have*

$$S(G_{tn}^{\mathcal{R}}) \xrightarrow{p} \infty \tag{1}$$

as $n \to \infty$. For every $t < t_b$ we have $\sum_{k \geqslant 1} \rho_k(t) = 1$. Also, for every $t < t_b$ there exist $a, A, C > 0$ (depending only on $\mathcal{R}, \ell, t$) such that for every $t' \in [0, t]$ we have $\rho_k(t') \leqslant Ae^{-ak}$ for all $k \geqslant 1$. In addition, for $n \geqslant n_0(\mathcal{R}, \ell, t)$ the following holds with probability at least $1 - n^{-99}$: for every $0 \leqslant i \leqslant tn$ we have

$$|N_k(G_i^{\mathcal{R}}) - \rho_k(i/n)n| \leqslant (\log n)^C n^{1/2} \quad \text{for all } k \geqslant 1, \tag{2}$$

$$|S(G_i^{\mathcal{R}}) - \textstyle\sum_{k \geqslant 1} k\rho_k(i/n)| \leqslant (\log n)^C n^{-1/2}, \tag{3}$$

and $N_{\geqslant k}(G_i^{\mathcal{R}}) \leqslant Ae^{-ak}n$ for all $k \geqslant 1$.

To interpret this result, we think of the functions $\rho_k(t)$ as describing the 'scaling limit' of the component size distribution at 'time' $t < t_b$, where time is the number of steps divided by n. A key aspect of the result is that this limit does not depend on n; in fact, most of our technical work is devoted to establishing this property – to show only that $N_k(G_{tn}^{\mathcal{R}})$ is concentrated around its expectation, simpler arguments would suffice. The tail bound on ρ_k given

in Theorem 1 states that the idealized component size distribution has an exponential tail for $t < t_{\mathrm{b}}$, as one would expect in a strictly sub-critical random graph. It implies that $s(t) = \sum_{k \geqslant 1} k\rho_k(t) < \infty$ if $t < t_{\mathrm{b}}$, so (3) implies that that for $t < t_{\mathrm{b}}$ we have

$$S(G^{\mathcal{R}}_{tn}) \xrightarrow{\mathrm{p}} s(t) < \infty, \tag{4}$$

where $\xrightarrow{\mathrm{p}}$ denotes convergence in probability. The proof of Theorem 1 will show that $s(t_{\mathrm{b}} - \varepsilon) \geqslant (\ell(\ell-1)\varepsilon)^{-1}$, so the (idealized) susceptibility $s(t)$ blows up at t_{b}. The last statement of the theorem implies that for $t < t_{\mathrm{b}}$ we have $L_1(G^{\mathcal{R}}_{tn}) \leqslant B_t \log n$ whp, for some constant B_t that depends on t. Finally, we shall show in the Appendix that (unless $\mathcal{R}$ directly adds cycles to the graph), for $t < t_{\mathrm{b}}$ whp almost all components are trees, with the rest unicyclic.

Theorem 1 allows us to say something about what happens at time $t = t_{\mathrm{b}}$. Indeed, the definition of an ℓ-vertex rule ensures that in one step, at most ℓ components are destroyed and at most ℓ (in fact at most $\ell/2$) are created, so $|N_k(G^{\mathcal{R}}_{i+1}) - N_k(G^{\mathcal{R}}_i)| \leqslant \ell k$. It follows that each ρ_k is Lipschitz continuous on $[0, t_{\mathrm{b}})$ with constant $k\ell$. Hence we can extend each ρ_k continuously to the point t_{b}, and (2) and the Lipschitz properties of N_k and ρ_k imply that

$$N_k(G^{\mathcal{R}}_{t_{\mathrm{b}}n})/n \xrightarrow{\mathrm{p}} \rho_k(t_{\mathrm{b}}). \tag{5}$$

Together with Theorem 1, the continuity results of [18, 20] imply that $L_1(G_{t_{\mathrm{b}}n})/n \xrightarrow{\mathrm{p}} 0$, and that $\sum_k \rho_k(t_{\mathrm{b}}) = 1$, so the numbers $(\rho_k(t_{\mathrm{b}}))_{k \geqslant 1}$ do capture the asymptotic component size distribution of $G^{\mathcal{R}}_{t_{\mathrm{b}}n}$, although we do not have such tight error bounds as in (2).

The proof of Theorem 1 is based on a variant of the neighbourhood exploration process and relies on branching process (approximation) arguments. This is quite different from previous approaches in this area, which are based on the differential equation method; for certain (restricted) classes of rules these establish *local convergence*, i.e., that there exist functions $\rho_k = \rho^{\mathcal{R}}_k : \mathbb{R}^+ \to [0,1]$ such that, for each fixed $k \geqslant 1$ and $t \geqslant 0$, we have $N_k(G^{\mathcal{R}}_{tn})/n \xrightarrow{\mathrm{p}} \rho_k(t)$ as $n \to \infty$. The limitations of these approaches are that they (i) only apply to certain bounded-size rules [6, 21], or (ii) when applied to size rules need the additional assumption that certain systems of differential equations have unique solutions [19], which is not known to hold for the product rule, for example. So, Theorem 1 establishes for the first time (a strong form of) local convergence for unbounded size rules (such as the product rule) until the susceptibility diverges. We believe that this convergence result is best possible: on the basis of heuristics and simulations presented in [17] we believe that there are certain natural size rules for which beyond $t = t_{\mathrm{b}}$ a giant component emerges whose size is not concentrated. In these rules the numbers of vertices in components of each fixed size k are presumably also not concentrated after this point.

Theorem 1 has some analogies with 'classical' percolation theory on, for example, the infinite lattice $\mathbb{Z}^d$, where there are two *a priori* different critical probabilities p_{H} and p_{T}. Intuitively, these correspond to the thresholds for (i) having (with positive probability) an infinite cluster and (ii) the expected cluster

size being infinite. For essentially all 'natural' lattices of interest it is nowadays known that $p_{\mathrm{H}} = p_{\mathrm{T}}$ (see e.g. [2, 16]), but this fact is not at all obvious! Note that in the finite setting of this paper these two properties correspond to (i) having a linear size component, and (ii) diverging susceptibility. More formally, define $t_{\mathrm{c}} = t_{\mathrm{c}}^{\mathcal{R}}$ as the supremum of the set of $t \geqslant 0$ for which $L_1(G_{tn}^{\mathcal{R}})/n \xrightarrow{\mathrm{p}} 0$ as $n \to \infty$, and $t_{\mathrm{b}} = t_{\mathrm{b}}^{\mathcal{R}}$ as the supremum of the set of $t \geqslant 0$ for which $S(G_{tn}^{\mathcal{R}})$ is bounded in probability. Note that $L_1(G) \leqslant \sqrt{nS(G)}$ implies $t_{\mathrm{b}} \leqslant t_{\mathrm{c}}$. The remark after Theorem 1 entails that for size rules t_{b} is equal to the infimum of the set of $t \geqslant 0$ for which (1) holds, and that $S(G_{t_{\mathrm{b}}n}^{\mathcal{R}}) \xrightarrow{\mathrm{p}} \infty$. In fact, we believe that both thresholds coincide for size rules (analogous to the 'classical' case).

Conjecture 2. *Let $\ell \geqslant 2$ and let $\mathcal{R}$ be an ℓ-vertex size rule. Then $t_{\mathrm{b}}^{\mathcal{R}} = t_{\mathrm{c}}^{\mathcal{R}}$. More precisely, for any $t > t_{\mathrm{b}}^{\mathcal{R}}$ and $\varepsilon > 0$ there exist $\delta, n_0 > 0$ (depending only on $\mathcal{R}, \ell, t, \varepsilon$) such that $\mathbb{P}(L_1(G_{tn}^{\mathcal{R}}) \geqslant \delta n) \geqslant 1 - \varepsilon$ for $n \geqslant n_0$.*

Recall that Achlioptas processes (where the choice is between two edges) are a sub-class of 4-vertex rules. Conjecture 2 was proved for bounded-size Achlioptas processes by Spencer and Wormald [21], and for a subset of these processes by Bohman and Kravitz [6]. We shall show in Section 3.1 that it holds for all bounded-size ℓ-vertex rules, as well as many other size rules, including the 'reverse product rule', for example. However, it does not hold for general ℓ-vertex rules. Indeed, in Section 3.2 we show that modified size rules with one additional feature, namely that they may once switch their behaviour based on the number n of vertices and the number i of steps (or the value of the susceptibility), can delay the appearance of a linear size component for $\Omega(n)$ steps beyond the point where the susceptibility diverges.

2 Evolution of Achlioptas processes with an initial graph

In this paper we consider the evolution of Achlioptas processes starting with an initial graph F with vertex set $V = [n]$; we restrict our attention to *ℓ-vertex size rules* $\mathcal{R}$, whose decisions depend only on the sizes of the components containing the randomly chosen vertices. More precisely, each such rule $\mathcal{R}$ yields a random sequence $(F_i^{\mathcal{R}})_{i \geqslant 0}$ of graphs on V with $F_0^{\mathcal{R}} = F$. For every $i \geqslant 0$ we draw ℓ vertices $\underline{v}_{i+1} = (v_1, \ldots, v_\ell)$ from V independently and uniformly at random, and then, writing $\underline{c}_{i+1} = (c_1, \ldots, c_\ell)$ for the sizes of the components containing $v_1, \ldots, v_\ell$ in $F_i^{\mathcal{R}}$, we obtain obtain $F_{i+1}^{\mathcal{R}}$ by adding a non-empty set of edges E_{i+1} to $F_i^{\mathcal{R}}$, where $\mathcal{R}$ deterministically selects E_{i+1} as a subset of all pairs between vertices in $\underline{v}_{i+1}$ based only on the list of component sizes $\underline{c}_{i+1}$. Usually exactly one edge is added, but there is no reason to insist on this.

When $F = G_0$ is the empty graph on n vertices we obtain the 'standard' Achlioptas processes using ℓ-vertex size rules $\mathcal{R}$ as defined in [18]. As usual, we can allow for small variations in the above definition; this includes, for example, each time picking an ℓ-tuple of *distinct* vertices, or picking (the ends of) $\ell/2$

randomly selected (distinct) edges not already present in $G_i^{\mathcal{R}}$, see also [18]. For $\ell = 2$ we thus recover the 'classical' Erdős–Rényi random graph process by always adding the pair v_1v_2. In addition, our proofs can be written to allow $\mathcal{R}$ to make randomized decisions(with the probability of adding some set of edges depending only on $\underline{c}_{i+1}$), and, furthermore, to allow $\mathcal{R}$ to know which vertices in $\underline{v}_{i+1}$ are in the same components of $F_i^{\mathcal{R}}$ (for compatibility with [18] we then require $E_{i+1} \neq \emptyset$ whenever all v_j are in distinct components, although nothing in the proof of Theorem 1, except for the bound $t_{\mathrm{b}} \leqslant 1$, needs this).

One difficulty in the proof of Theorem 1 is that there is a complicated dependence between the decisions of $\mathcal{R}$ in each round (and their order is also important). Indeed, changes can 'propagate' throughout the process: if the sizes of a few components are modified (e.g. by altering decisions of $\mathcal{R}$ or tuples $\underline{v} = (v_1, \dots, v_\ell)$ offered), then this might change many future decisions of $\mathcal{R}$, which in turn might alter further decisions, etc. To overcome this our proof proceeds by induction, always establishing concentration only for a small number of steps; this is also the reason why we study the more general evolution starting from an initial graph F. Each time we rely on a two-round exposure argument: in the first round we reveal which tuples are selected, and in the second we then expose their order. For size rules *not* all tuples and components of F 'influence' the size of the component in $F_i^{\mathcal{R}}$ containing v: only those which can be reached from v after adding *all* pairs of each ℓ-tuple to the graph (every rule only selects a subset of these pairs). The key observation is now that given the corresponding 'relevant' tuples and components of F of the first round, the order of these tuples (exposed in the second round) determines the size of the component containing v. It turns out that if we only consider σn rounds for σ sufficiently small, then an exploration process determining these relevant tuples and components in the first round can be closely approximated by a subcritical branching process $\mathfrak{X}_\sigma$ which is defined *without* reference to n. Since the outcome of the second round is a (random) function of the first one, it thus seems plausible that $\mathbb{E}N_k(F_{\sigma n}^{\mathcal{R}})/n$ is independent of n (up to small error terms). In addition, since the first round is subcritical, this means that there typically are not too many tuples and components which influence the size of the component containing v. At least on an intuitive level this makes it plausible that it should be possible to establish concentration of $N_k(F_i^{\mathcal{R}})$ around its expectation by applying McDiarmid's inequality.

The rest of this paper is organized as follows. In the next section we state our main technical result (Theorem 3), and then show in Section 2.2 how it implies Theorem 1. Afterwards, in Section 2.3 we present some branching process preliminaries; these are used in Section 2.4, where we establish Theorem 3. In Section 3 we discuss Conjecture 2, giving examples of classes of size rules for which we can prove the conjecture, and examples of non-size rules for which it does not hold. Finally, in the appendix we consider the cycle structure of Achlioptas processes.

2.1 Main technical result

Our main technical result establishes concentration during the evolution of Achlioptas processes starting with an initial graph F. The special case of an Erdős–Rényi evolution from an initial graph F (which can be seen as an evolving version of a special case of the inhomogeneous random graph model of Bollobás, Janson and Riordan [7]) has been previously studied by Spencer and Wormald [21] and Janson and Spencer [12], the main focus being on the size of the largest component. In this context the susceptibility turns out to be the key parameter, and both papers use in essential ways that the Erdős–Rényi evolution corresponds to the addition of *uniform* random edges (or pairs of vertices). In contrast, when studying the evolution of Achlioptas processes, we need to deal with intricate dependencies between the edges added.

Using susceptibility as a guide, we now briefly motivate the number of steps our result applies to. Suppose that, starting with F satisfying $S(F) = L$, we use the rule $\mathcal{I}$ which in each step joins all ℓ random vertices by edges. Set $s(t) = S(F^{\mathcal{I}}_{tn})$. If the sizes of the joined components are c_i, then (assuming that all components are distinct) the susceptibility changes by $(\sum c_i)^2/n - \sum c_i^2/n = \sum_{i\neq j} c_i c_j/n$. So, since the vertices of each tuple are chosen uniformly at random, it seems plausible that typicality we have $s'(t) \approx n\mathbb{E}(S(F^{\mathcal{I}}_{tn+1}) - S(F^{\mathcal{I}}_{tn})) \approx \ell(\ell-1)s(t)^2$. For $t < [\ell(\ell-1)L]^{-1} = t_{\mathrm{s}}$ this suggests $s(t) \approx [1/L - \ell(\ell-1)t]^{-1}$. Since in each step any rule $\mathcal{R}$ only adds a subset of all $\binom{\ell}{2}$ pairs to the graph, this indicates that the susceptibility does not 'blow up' as long as $t < t_{\mathrm{s}}$. The following result confirms this heuristic argument and shows that, under suitable conditions, for $t < t_{\mathrm{s}}$ the number of vertices in components of size $k \geqslant 1$ is also tightly concentrated (the function ρ intuitively results from an 'infinite' version of the rule $\mathcal{R}$). Here we set $\chi(\varphi) = \sum_{k\geqslant 1} k\varphi(k)$ and $\chi(\rho, t) = \sum_{k\geqslant 1} k\rho(k,t)$, and write $x = a \pm b$ as shorthand for $x \in [a-b, a+b]$.

Theorem 3. *Let $\ell \geqslant 2$ and let $\mathcal{R}$ be an ℓ-vertex size rule. Suppose $\beta > 1$, $B > 0$, $L \geqslant 1$ and $\varphi : \mathbb{N} \to [0,1]$ satisfy*

$$\sum_{k\geqslant 1}\varphi(k) = 1, \tag{6}$$

$$\sum_{k\geqslant 1}\varphi(k)\beta^k \leqslant B, \tag{7}$$

$$\chi(\varphi) \leqslant L. \tag{8}$$

There is a function $\rho : \mathbb{N}\times\mathbb{R}^+ \to [0,1]$ (depending only on $\varphi, \mathcal{R}, \ell$) such that for all $\sigma \geqslant 0$ satisfying

$$\sigma < [\ell(\ell-1)L]^{-1} \tag{9}$$

there exist $\tilde{\beta}, \tilde{B}, \tilde{L} > 1$ (depending only on $\ell, L, \sigma, \beta, B$) such that for every $t \in [0,\sigma]$ equations (6)–(8) hold when $\beta, B, L, \varphi(\cdot)$ are replaced by $\tilde{\beta}, \tilde{B}, \tilde{L}, \rho(\cdot, t)$. If in addition F is a graph on n vertices which for $C \geqslant 0$ satisfies

$$N_k(F) = \varphi(k)n \pm (\log n)^C n^{1/2} \quad \text{for all } k \geqslant 1, \tag{10}$$

$$\sum_{k\in[n]} N_k(F)\beta^j \leqslant Bn, \tag{11}$$

$$S(F) \leqslant L, \tag{12}$$

then, setting $\tilde{C} = C + 9$*, for* $n \geqslant n_0(\ell, L, \sigma, \beta, B, C)$ *the following holds with probability at least* $1 - n^{-200}$*: for every* $0 \leqslant i \leqslant \sigma n$ *we have*

$$S(F_i^{\mathcal{R}}) = \chi(\rho, i/n) \pm (\log n)^{\tilde{C}} n^{-1/2}, \tag{13}$$

and equations (10)–(12) *hold when* $\beta, B, L, C, F, \varphi(\cdot)$ *are replaced by* $\tilde{\beta}, \tilde{B}, \tilde{L}, \tilde{C}, F_i^{\mathcal{R}}, \rho(\cdot, i/n)$*.*

The proof of Theorem 3 is quite involved and is deferred to Section 2.4. It is useful to observe that since $\beta > 1$ holds, (7) and (11) imply the tail bounds $\max\{\sum_{j \geqslant k} \varphi(j), N_{\geqslant k}(F)/n\} \leqslant B\beta^{-k}$ for all $k \geqslant 1$, so $L_1(F) = O(\log n)$. By Theorem 3 analogous estimates also hold for $F_i^{\mathcal{R}}$ with $i \leqslant \sigma n$. In fact, for (11), (12) to hold with β, B, L, F replaced by $\tilde{\beta}, \tilde{B}, \tilde{L}, F_i^{\mathcal{R}}$, a minor modification of our proof shows that it suffices to assume (9), (11) and (12) only; for the special case $\ell = 2$ this was established by Spencer and Wormald [21] under similar conditions. However, the key point of Theorem 3 is (10), i.e., that we obtain concentration of number of vertices in components of size k.

Turning to the susceptibility, by combining (10) with the tail bounds following from (7) and (11), for each fixed j we readily obtain rather precise estimates for $S_j(F_i^{\mathcal{R}}) = \sum_{k \in [n]} k^j N_k(F_i^{\mathcal{R}})/n$ with $i \leqslant \sigma n$, similar to (13). Furthermore, since $L_1(F_{\sigma n}^{\mathcal{R}}) = O(\log n)$ whp, we can easily use the differential equation method to make our heuristic discussion regarding the susceptibility rigorous, which e.g. yields $\chi(\rho, \sigma) \leqslant [1/L - \ell(\ell-1)\sigma]^{-1}$ (for the special case $\ell = 2$ this was noted by Bohman et. al [5]; it is also implicit in [21]). However, this crude bound, which follows from always connecting all ℓ vertices by edges in each step, is generally far from the truth; for this reason it does not suffice in our inductive application of Theorem 3, where we use the 'correct' value given by (13).

2.2 Proof of Theorem 1

This section is devoted to the proof of Theorem 1, which we establish by an inductive application of Theorem 3: each time we show concentration during a small number of steps (and maintain certain technical conditions), where the lengths of these intervals decrease as the susceptibility increases. This is also the main idea of the following rather technical construction: as we shall see in the proof of Lemma 4, for each interval of length Δ_j it determines the scaling limits ρ_k (and certain tail bounds) in a way that does *not* depend on n.

We inductively define a sequence $(\beta_j, B_j, \rho_j, \Delta_j, L_j)_{j \geqslant 0}$ with $\beta_j > 1$, $B_j > 0$, $\Delta_j \geqslant 0$, $L_j \geqslant 1$ and $\rho_j : \mathbb{N} \times \mathbb{R}^+ \to [0,1]$, where the β_j are decreasing ($\beta_{j+1} \leqslant \beta_j$) and the B_j are increasing ($B_{j+1} \geqslant B_j$). In addition, for each $j \geqslant 0$ the sequence satisfies the following invariant: for every $t \in [0, \Delta_j]$ equations (6) and (7) hold with $\beta, B, \varphi(\cdot)$ replaced by $\beta_j, B_j, \rho_j(\cdot, t)$. We start by setting $\beta_0 = B_0 = L_0 = 2$, $\Delta_0 = 0$ and defining $\rho_0 : \mathbb{N} \times \mathbb{R}^+ \to [0,1]$ with $\rho_0(1,0) = 1$ and $\rho_0(k,t) = 0$ otherwise. Given $j \geqslant 1$, recall that $\chi(\rho_{j-1}, t) = \sum_{k \geqslant 1} k\rho_{j-1}(k,t)$ and set

$$L_j = \chi(\rho_{j-1}, \Delta_{j-1}) + 1 \quad \text{and} \quad \Delta_j = [\ell(\ell-1)\,(L_j+1)]^{-1}. \tag{14}$$

Applying the first part of Theorem 3 with $\beta = \beta_{j-1}$, $B = B_{j-1}$, $L = L_j$ and $\varphi(k) = \rho_{j-1}(k, \Delta_{j-1})$, we use the resulting ρ to define $\rho_j = \rho$. Furthermore, by considering $\sigma = \Delta_j$ we obtain $\tilde{\beta}, \tilde{B}$ and set $\beta_j = \min\{\tilde{\beta}, \beta_{j-1}\}$ and $B_j = \max\{\tilde{B}, B_{j-1}\}$; by Theorem 3 these satisfy the required invariant. Furthermore, it is not difficult to see that the entire sequence $(\beta_j, B_j, \rho_j, \Delta_j, L_j)_{j\geqslant 0}$ depends only on $\mathcal{R}, \ell$.

Next we combine the ρ_j (each valid on an interval of length Δ_j) to form $\varphi(k,t)$, which will eventually be $\rho_k(t)$ in Theorem 1; this notation avoids confusion with the ρ_j used. For $t \geqslant 0$ we define r_t as the smallest r such that $t \leqslant \sum_{0\leqslant j\leqslant r} \Delta_j$ and set $r = \infty$ if no such r exists. For all $(k,t) \in \mathbb{N} \times \mathbb{R}^+$ set

$$\varphi(k,t) = \begin{cases} \rho_{r_t}(k, t - \sum_{0\leqslant j<r_t} \Delta_j), & \text{if } r_t < \infty, \\ 0, & \text{otherwise.} \end{cases} \tag{15}$$

Transferring this definition to the invariant of the sequence introduced above, for all $t \geqslant 0$ with $r_t < \infty$ it follows that

$$\sum_{k\geqslant 1} \varphi(k,t) = 1, \tag{16}$$

and that for every $t' \in [0,t]$ we have

$$\sum_{k\geqslant 1} \varphi(k,t')\beta_{r_t}^k \leqslant B_{r_t}. \tag{17}$$

Now we are ready to prove the following concentration result, which also implies that in the previous construction we always have $r_t < \infty$ if (1) fails.

Lemma 4. *Let $\ell \geqslant 2$ and let $\mathcal{R}$ be an ℓ-vertex size rule. For every $t \geqslant 0$ for which (1) fails we have $r_t < \infty$, and there exist $a, A, C > 0$ (depending only on $\mathcal{R}, \ell, t$) such that the following holds for $n \geqslant n_0(\mathcal{R}, \ell, t)$ with probability at least $1 - n^{-99}$: for every $0 \leqslant i \leqslant tn$ we have*

$$N_k(G_i^{\mathcal{R}}) = \varphi(k, i/n)n \pm (\log n)^C n^{1/2} \quad \text{for all } k \geqslant 1,$$
$$S(G_i^{\mathcal{R}}) = \textstyle\sum_{k\geqslant 1} k\varphi(k, i/n) \pm (\log n)^C n^{-1/2},$$

and $N_{\geqslant k}(G_i^{\mathcal{R}}) \leqslant Ae^{-ak}n$ for all $k \geqslant 1$.

Proof. Given $t \geqslant 0$, if (1) fails there exists $\varepsilon > 0$ and an infinite subsequence $\bar{n}$ of $\mathbb{N}$ (depending only on $\mathcal{R}, \ell, t$) satisfying

$$\mathbb{P}(S(G_{t\bar{n}}^{\mathcal{R}}) \leqslant \varepsilon^{-1}) \geqslant \varepsilon. \tag{18}$$

Let $\overline{L} = \varepsilon^{-1} + 3$ and $K = \lceil t\ell(\ell-1)\overline{L} \rceil + 1$. Let $t_0 = 0$, and for $j \geqslant 1$ let

$$t_j = \begin{cases} t_{j-1}, & \text{if } t_{j-1} > t, \\ t_{j-1} + \Delta_j, & \text{otherwise.} \end{cases}$$

For $n \geqslant n_0(\mathcal{R}, \ell, t)$ we inductively show that for every $0 \leqslant j \leqslant K$, setting $C_j = 9j + 2$, with probability at least $1 - jn^{-200}$, for every $0 \leqslant i \leqslant t_j n$ we have

$$N_k(G_i^{\mathcal{R}}) = \varphi(k, i/n)n \pm (\log n)^{C_j} n^{1/2} \quad \text{for all } k \geqslant 1, \tag{19}$$

$$\textstyle\sum_{k \in [n]} N_k(G_i^{\mathcal{R}}) \beta_j^k \leqslant B_j n, \tag{20}$$

$$S(G_i^{\mathcal{R}}) = \chi(\varphi, i/n) \pm (\log n)^{C_j} n^{-1/2}, \tag{21}$$

and for every $0 \leqslant s \leqslant j$ with $t_s \leqslant t$ we have

$$\chi(\varphi, t_s) = \chi(\rho_s, \Delta_s) < \overline{L} - 2. \tag{22}$$

Note that if $t_K < t$, then substituting (22) into (14) yields $\Delta_j > [\ell(\ell-1)\overline{L}]^{-1}$ for all $1 \leqslant j \leqslant K$. From $K > t\ell(\ell-1)\overline{L}$ it thus follows that $t_K > t$, a contradiction. Thus (22) implies $t_K \geqslant t$, i.e., $r_t < \infty$. Recall that $(\beta_j, B_j)_{j \geqslant 1}$ and K depend only on $\mathcal{R}, \ell$ and on $\mathcal{R}, \ell, t$ respectively. Hence the induction hypothesis for $j = K$ implies Lemma 4, where the tail bounds follow from (20) as $\beta_K > 1$.

For the base case $j = 0$ we start with an empty graph on n vertices, and it is easy to see that (19)–(22) hold with $\beta_0 = B_0 = C_0 = 2$ and $\varphi(k, 0) = \rho_0(k, 0)$, as defined above (14).

Given $j \geqslant 1$, for the induction step we may assume that $t_{j-1} \leqslant t$ (otherwise $t_j = t_{j-1}$, and there is nothing to prove). We first assume that $G^{\mathcal{R}}_{t_{j-1}n}$ satisfies the induction hypothesis, i.e., (19)–(22) with j replaced by $j-1$. In particular, (16) and (17) hold for $t = t_{j-1}$ with $r_t = j - 1$, and we have $S(G^{\mathcal{R}}_{t_{j-1}n}) \leqslant \chi(\varphi, t_{j-1}) + 1 = \chi(\rho_{j-1}, \Delta_{j-1}) + 1 = L_j$ for $n \geqslant n_0(C_{j-1})$. Now we condition on $G^{\mathcal{R}}_{t_{j-1}n} = F$ and, analogous as after (14), apply Theorem 3 with $\beta = \beta_{j-1}$, $B = B_{j-1}$, $L = L_j$, $\sigma = \Delta_j$, $C = C_{j-1}$ and $\varphi(k) = \rho_{j-1}(k, \Delta_{j-1}) = \varphi(k, t_{j-1})$, which is possible by the induction hypothesis (and the properties established above). So, for $n \geqslant n_0(\ell, L_j, \Delta_j, \beta_{j-1}, B_{j-1}, C_{j-1})$, with probability at least $1 - n^{-200}$, for every $0 \leqslant i \leqslant \Delta_j n$ the graph $F_i^{\mathcal{R}}$ satisfies (6)–(7), (10)–(11) and (13) when $\beta, B, C, F, \varphi(\cdot)$ are replaced by $\tilde{\beta}, \tilde{B}, \tilde{C}, F_i^{\mathcal{R}}, \rho(\cdot, i/n)$, where $\tilde{C} = C_{j-1}+9$. Note that for size rules $F^{\mathcal{R}}_{\Delta_j n}$ is exactly $G^{\mathcal{R}}_{t_j n}$ conditional on $G^{\mathcal{R}}_{t_{j-1}n} = F$. It is crucial that $\tilde{\beta}, \tilde{B}, \tilde{C}, \rho$ do *not* depend on the initial graph F, and that by construction $\rho = \rho_j$, $\beta_j \leqslant \tilde{\beta}$, $B_j \geqslant \tilde{B}$ and $C_j = \tilde{C}$. So, by appealing to the induction hypothesis and recalling (15), it follows that with probability at least $1 - (j-1)n^{-200} - n^{-200}$ equations (19)–(21) hold. It remains to show that (22) holds. To this end recall that (21) holds with probability at least $1 - jn^{-200} > 1 - \varepsilon$ for all $n \geqslant n_0(C_j, K, \varepsilon)$. So, using that susceptibility is monotone increasing, by (18) it follows that for all $0 \leqslant t' \leqslant \min\{t, t_j\}$ we have $\chi(\varphi, t') < \varepsilon^{-1} + 1 = \overline{L} - 2$, say. Now (22) follows by combining the previous estimate with the observation that for every $s \leqslant j$ with $t_s \leqslant t$ we have $\chi(\varphi, t_s) = \chi(\rho_s, \Delta_s)$. This completes the induction step.

Finally, to see that $n \geqslant n_0(\mathcal{R}, \ell, t)$ suffices note that in each of the K steps we only used $n \geqslant n_0(\ell, L_j, \Delta_j, \beta_{j-1}, B_{j-1}, C_{j-1}, C_j, K, \varepsilon)$, where $C_j = 9j + 2$ and $L_j, \Delta_j, \beta_{j-1}, B_{j-1}$ depend only on $\mathcal{R}, \ell$. This concludes the proof since ε (and thus K) only depends on $\mathcal{R}, \ell, t$. □

Now we define $t_{\mathrm{b}} = t_{\mathrm{b}}^{\mathcal{R}}$ as the infimum of the set of $t \geqslant 0$ for which (1) holds as $n \to \infty$; so (1) fails for $t < t_{\mathrm{b}}$. The remark after the proof of Lemma 4 in [18] implies that for $t > 1$ we whp have $L_1(G_{tn}^{\mathcal{R}}) \geqslant cn$ for $c = c(\ell, t) > 0$, yielding $S(G_{tn}^{\mathcal{R}}) \geqslant [L_1(G_{tn}^{\mathcal{R}})]^2/n \geqslant c^2 n$; so $t_{\mathrm{b}} \leqslant 1$. Furthermore, for $t < [\ell(\ell-1)]^{-1}$ an application of Theorem 3 to the empty graph $F = G_0^{\mathcal{R}}$ on n vertices with $\sigma = t$ and $L = 1$ (similar as in the proof of Lemma 4) readily shows $S(G_{\sigma n}^{\mathcal{R}}) \leqslant \tilde{L}$ whp, so $t_{\mathrm{b}} \geqslant [\ell(\ell-1)]^{-1}$. Now suppose that (1) fails for $t = t_{\mathrm{b}}$. The proof of Lemma 4 then shows that whp (19)–(21) hold for $i = t_{\mathrm{b}} n$, and that $\chi(\varphi, t_{\mathrm{b}}) < \overline{L} - 2$. It follows that we can apply Theorem 3 with $\sigma = [\ell^2 \overline{L}]^{-1}$ and $L = \overline{L}$; this implies $S(G_{(t_{\mathrm{b}}+\sigma)n}^{\mathcal{R}}) \leqslant \tilde{L}$ whp, contradicting the definition of t_{b}. So, since the susceptibility is monotone increasing, it follows that (1) holds for all $t \geqslant t_{\mathrm{b}}$. Combining our findings, Lemma 4, (16) and (17) now yield Theorem 1 with $\rho_k(t) = \varphi(k, t)$.

2.3 Branching processes preliminaries

The following basic results for branching processes will be used in the proof of Theorem 3. They are similar to Theorems 3.2 and 3.3 in [21], where they are attributed to much earlier results of Crámer. Given a non-negative integer valued random variable X, let $F_X(z) = \mathbb{E}z^X$ denote the *(probability) generating function* of X. Note that $F_X(z)$ is convex and monotone increasing for $z \geqslant 0$.

The first lemma essentially states that a two-generation branching process has (uniform) exponential tails provided that the generating function of each offspring distribution has radius of convergence strictly larger than one (and thus also exhibits exponential decay).

Lemma 5. *Let $X, Y \geqslant 0$ be integer valued random variables with $F_X(\alpha) \leqslant A$ and $F_Y(\beta) \leqslant B$, where $\alpha, \beta > 1$. Let Z be the number of grandchildren in the two-generation branching process in which the root node has X children and then each child, independently, has Y children. There exists $a > 0$ (depending only on α, β, A, B) such that $\mathbb{P}(Z \geqslant s) \leqslant Ae^{-as}$ for all $s \geqslant 0$.*

Proof. Pick $C \geqslant \max\{B, 2\}$ such that $x = 1 + (\alpha-1)(\beta-1)/(C-1) \leqslant \beta$. Using $F_Y(1) = 1$ and $F_Y(\beta) \leqslant B \leqslant C$, convexity yields $F_Y(z) \leqslant [(z-1)C + (\beta - z)]/(\beta - 1)$ for all $z \in [1, \beta]$. So, by choice of x we have $F_Y(x) \leqslant \alpha$. Observing that $F_Z(z) = F_X(F_Y(z))$, using monotonicity we obtain $F_Z(x) = F_X(F_Y(x)) \leqslant F_X(\alpha) \leqslant A$. Since $x > 1$ implies $F_Z(x) \geqslant \mathbb{P}(Z \geqslant s)x^s$ for every $s \geqslant 0$, we deduce $\mathbb{P}(Z \geqslant s) \leqslant Ae^{-as}$ for $a = \log x > 0$, completing the proof. □

The second lemma is a standard result for subcritical Galton–Watson branching process: these exhibit (uniform) exponential decay if the offspring distribution itself has (uniform) exponential tails.

Lemma 6. *Let $Z \geqslant 0$ be an integer valued random variable with $\mathbb{E}Z \leqslant \mu < 1$ and $F_Z(\beta) \leqslant B$, where $\beta > 1$. Let T be the total size of the Galton–Watson branching process in which each node, independently, has Z children. There exist $\delta > 1$ and $D > 0$ (depending only on β, B, μ) such that $F_T(\delta) \leqslant D$.*

Proof. Let $f(t) = \mathbb{E}(e^{t(Z-1)})$. Observe that $f(0) = 1$ and $f'(0) = \mathbb{E}(Z-1) \leqslant \mu - 1$. As in the proof of Lemma 5, $F_Z(\beta) \leqslant B$ for $\beta > 1$ yields $\mathbb{P}(Z \geqslant s) \leqslant B\beta^{-s}$, which in turn readily implies that for some $C = C(\beta, B)$ we have $f''(t) = \mathbb{E}((Z-1)^2 e^{t(Z-1)}) \leqslant C$ for all $0 \leqslant t \leqslant (\log\beta)/2$, say. So, using Taylor's theorem, for $0 \leqslant t \leqslant (\log\beta)/2$ we deduce $f(t) \leqslant 1 + (\mu-1)t + Ct^2/2 = h(t)$. Let $x = \min\{(\log\beta)/2, (1-\mu)/C\} > 0$, and observe that $c = \max\{h(x), 1/2\} > 0$ satisfies $f(x) \leqslant c < 1$. Exploring the branching process tree as usual in breadth-first search order, we see that $T > s$ implies $\sum_{i=1}^{s} Z_i \geqslant s$, where the Z_i are independent copies of Z (corresponding to the number of children of the i-th node). Now, using Markov's inequality and independence of the Z_i, for every $s \geqslant 0$ we obtain $\mathbb{P}(T > s) \leqslant \mathbb{E}(e^{x(\sum_{i=1}^{s} Z_i)})e^{-xs} = f(x)^s \leqslant c^s$. Finally, picking $1 < \delta < 1/c$, it follows that $F_T(\delta) \leqslant D = D(\delta, c)$, as claimed. □

2.4 Proof of Theorem 3

The proof of Theorem 3 relies on a two-round exposure argument: we first reveal the random tuples selected, and afterwards expose their order of appearance. It will be convenient to work with a continuous-time random graph model, where the n^ℓ tuples arrive according to independent Poisson processes with rates $1/n^{\ell-1}$. So tuples appear with rate n, and each tuple is chosen uniformly at random and independently of all previous choices. Let E_t denote the set of tuples which arrive in $[0,t]$; so $|E_t| \sim \mathrm{Po}(tn)$. Observe that for each tuple $\underline{u} \in [n]^\ell$ the number $A_{\underline{u}}(t)$ of its arrivals in $[0,t]$ satisfies $A_{\underline{u}}(t) \sim \mathrm{Po}(t/n^{\ell-1})$, and that these random variables are independent for different tuples. Furthermore, writing $x = t/n^{\ell-1}$ and using $e^{-x} \geqslant 1 - x$ twice, note that for $\ell \geqslant 2$ we have

$$\mathbb{P}(A_{\underline{u}}(t) \geqslant 2) = 1 - e^{-x} - xe^{-x} \leqslant x(1 - e^{-x}) \leqslant x^2 \leqslant t^2/n^\ell. \tag{23}$$

Similarly

$$\mathbb{P}(A_{\underline{u}}(t) \geqslant 1) = 1 - e^{-x} \leqslant x = t/n^{\ell-1}. \tag{24}$$

Starting with F, for each tuple $\underline{u} \in E_t$ we join all $\binom{\ell}{2}$ pairs of vertices by edges, and we denote the resulting graph by H_t. We define $H_t^{\mathcal{R}}$ as the graph which we obtain by starting with F, and then presenting the tuples to $\mathcal{R}$ (together with the component sizes of the vertices) in a random order, always updating the graph according to the decisions of $\mathcal{R}$ (adding the pairs selected by $\mathcal{R}$). Since conditioned on $|E_t| = i$ we have i tuples chosen independently and uniformly at random, it follows that

$$\mathbb{E}(N_k(H_t^{\mathcal{R}}) \mid |E_t| = i) = \mathbb{E}N_k(F_i^{\mathcal{R}}). \tag{25}$$

Furthermore, mimicking the proof of *Pittel's inequality* (see e.g. [8]) for $0 < tn < n^\ell$, a short calculation shows that for any graph property $\mathcal{Q}$ we have

$$\mathbb{P}(F_{tn}^{\mathcal{R}} \notin \mathcal{Q}) \leqslant 3\sqrt{tn} \cdot \mathbb{P}(H_t^{\mathcal{R}} \notin \mathcal{Q}). \tag{26}$$

In the following sections we always tacitly assume that the assumptions of Theorem 3 hold and consider $t = t(n)$ satisfying

$$0 \leqslant t \leqslant \sigma \leqslant 1, \tag{27}$$

where $\sigma \leqslant 1$ follows from (9). Furthermore, unless stated otherwise, we will use the continuous-time random graph models H_t and $H_t^{\mathcal{R}}$. For later usage let

$$U = (\log n)^{6/5}. \tag{28}$$

2.4.1 Component exploration process for $\ell = 2$

Our main ingredient for analyzing the first exposure round is a certain exploration process. Given a (random) vertex v, it finds all tuples in E_t and components of F that are 'relevant' in the second exposure round for determining $|C_v(H_t^{\mathcal{R}})|$, where we write $C_v(G)$ for the set of vertices of G that are in the same component as v. As certain details are rather technical for Achlioptas processes, here we first outline some of the basic ideas and techniques for the simpler case of an Erdős–Rényi evolution starting from an initial graph F (in this special case similar ideas were used by Spencer and Wormald [21]). This formally corresponds to the special case $\ell = 2$ and the rule which always adds the offered pair v_1v_2 to the evolving graph; so $H_t = H_t^{\mathcal{R}}$.

One major difference to the Erdős–Rényi case (where we start with an empty graph on n vertices) is that here we have two sources of edges: (i) the initial graph F and (ii) the random pairs in E_t. As edges of type (i) are deterministic and those of type (ii) are random, our exploration process explicitly considers them separately. In the first round we start with a randomly chosen v and mark all $u \in C_v(F)$ as *reached*; all other vertices are *unreached*. In each later round we sequentially go through the vertices w reached in the previous round (the order does not matter here) and determine all its so far unreached neighbours u in E_t (corresponding to pairs $(u_1, u_2) \in E_t$ containing u and w), each time marking all $\tilde{u} \in C_u(F)$ as reached. Note that upon termination $C_v(H_t)$ equals the set of all reached vertices.

The previous procedure yields an associated 'exploration tree' $\mathcal{T}_v(H_t)$ in a rather natural way: loosely speaking, u is a child of w if u was 'reached' via w. With an eye to the upcoming analysis for size rules, here we already introduce different types of nodes: *vertex nodes*, *component nodes*, and *root nodes*. More precisely, we define $\mathcal{T}_v(H_t)$ inductively as follows: it has a root node v, whose children are vertex nodes $u \in C_v(F)$. Then, given any vertex node w, each of its so far unreached neighbours u in E_t yields a component node as a child, which in turn has vertex nodes $\tilde{u} \in C_u(F)$ as children. It follows that the set of all vertex nodes in $\mathcal{T}_v(H_t)$ equals $C_v(H_t)$. The main point is that, even after ignoring all labels, the *structure* of $\mathcal{T}_v(H_t)$ is enough to determine $|C_v(H_t)|$.

The key idea is now to approximate $\mathcal{T}_v(H_t)$ by an 'idealized' branching process, similar as in the 'classical' Erdős–Rényi case (exploiting, as usual, that by construction every edge is tested at most once). Recall that in $\mathcal{T}_v(H_t)$ already reached vertices are 'ignored'. So, noting that endpoints of random pairs in E_t correspond to random vertices, and that each edge gives rise to two *ordered* tuples, it seems plausible that $\mathcal{T}_v(H_t)$ is dominated by (may be regarded as a subset of) a branching process $\mathfrak{T}_{v,t}$ where (ignoring for simplicity the root and all labels) every vertex node, independently, has $\mathrm{Po}(2t)$ component nodes as

children, each which in turn, independently, has N vertex node descendants, where $N \sim |C_u(F)|$ for a randomly chosen vertex u. Now, using (9) and (12) each vertex node has in expectation $2t \cdot S(F) \leqslant 2\sigma L < 1$ vertex nodes as grandchildren, so we expect that $\mathfrak{T}_{v,t}$ resembles a subcritical branching process which has $O(\log n)$ size with very high probability. From this it follows that $\mathcal{T}_v(H_t)$ and $\mathfrak{T}_{v,t}$ are both small and have similar offspring distribution (as not too many vertices are reached and thus ignored), so it seems plausible that we can couple them so that they agree whp. Note that $\mathfrak{T}_{v,t}$ still depends on n and the initial graph F. Define $\mathbb{P}(R = k) = \varphi(k)$, where φ is given by Theorem 3. The point is now that using (10) it follows that R is very close to N. So, denoting by $\mathfrak{X}_{\varphi,t}$ the 'idealized' version of $\mathfrak{T}_{v,t}$ where we use R instead of N, the former considerations suggest that there is a coupling such that whp $\mathfrak{T}_{v,t} \cong \mathfrak{X}_{\varphi,t}$ holds (ignoring the labels of the vertices). To summarize, we just outlined that using the 'intermediate' process $\mathfrak{T}_{v,t}$ we can couple $\mathcal{T}_v(H_t)$ and $\mathfrak{X}_{\varphi,t}$ so that they typically agree up to isomorphisms. Consequently, the distribution of $|C_v(H_t)|$ can be approximated using $\mathfrak{X}_{\varphi,t}$, which does *not* depend on n or F.

In the above construction and analysis we used in essential ways that in each round only one pair of vertices is chosen and connected by an edge. In contrast, when considering Achlioptas processes several vertices $\underline{v} = (v_1, \dots, v_\ell)$ are chosen in each round, and only a subset of the edges between these vertices is added to the evolving graph. Furthermore, in the second exposure round the order in which the tuples $\underline{v}$ are presented matters (as well as the order of the vertices in each tuple). This motivates the more involved exploration processes used in the next section, whose associated exploration tree captures more detailed structural information (also using more types of nodes).

2.4.2 Component exploration process (the general case)

In this section we consider the first exposure round, where the selected set of tuples E_t is revealed. Note that this defines H_t, which we obtain by starting with F and then joining all ℓ vertices of each tuple in E_t by edges. Using a natural variant of the standard neighbourhood exploration process, for any vertex v we can determine $C_v(H_t)$ as follows. First we determine $C_v(F)$, i.e., find all other vertices which are in the same component of F as v. Then, for each $w \in C_v(F)$ we find all tuples $\underline{u} = (u_1, \dots, u_\ell) \in E_t$ containing w, and repeat the same procedure (recursively) for each $u_j \neq w$, see Figure 1. Observe that for determining $C_v(H_t)$ it suffices to consider only those vertices $u_j \neq w$ which we have not already reached in some previous exploration step.

In the analysis it is easier to start with a *random* vertex v and break down the above exploration process into small steps, constructing an an associated *exploration tree* $\mathcal{T}_{v,t} = \mathcal{T}_{v,t}(F)$. As we shall see, one key property of $\mathcal{T}_{v,t}$ is that we can (typically) reconstruct the vertices and components which have been reached, as well as the tuples which have been 'tested' so far. The vertices of each exploration tree have different types: *vertex nodes*, *component nodes* and *tuple nodes* will represent vertices, components of F and ℓ-tuples, respectively. For technical reasons we also have *root nodes* and *index nodes*. We denote the

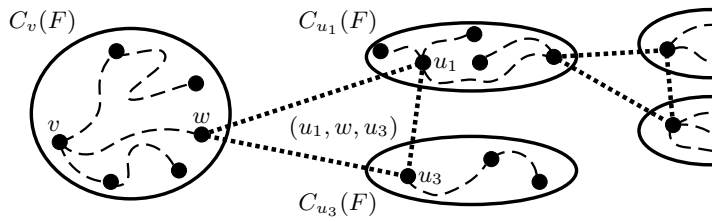

Figure 1: Example of the neighbourhood exploration process for $\ell = 3$. It determines $C_v(H_t)$ by first finding other vertices in the same component of F, then finding tuples containing them; afterwards it repeats this procedure for the new vertices in those tuples, and so on.

vertex nodes of $\mathcal{T}_{v,t}$ by $\mathcal{V}_{v,t}$.

As mentioned above, our exploration starts with a random vertex v, which serves as the root node of $\mathcal{T}_{v,t}$, see Figure 2. Next we (deterministically) find all vertices $w \in C_v(F)$ and then add the vertex nodes w as children of the root. In the following we sequentially traverse each level containing vertex nodes (which essentially corresponds to a breadth first search). Given a vertex node w, we add ℓ index nodes $w_1, \dots, w_\ell$ as children, where w_j is an index node of type j. For each $j = 1, \dots, \ell$ we sequentially test for the presence and multiplicity of all so far untested tuples $\underline{u} = (u_1, \dots, u_\ell)$ with $u_j = w$; we denote the resulting multiset of found tuples by $S_{j,w}$. Now we sequentially traverse the $\underline{u} \in S_{j,w}$. For each such $\underline{u} = (u_1, \dots, u_\ell)$ we add a tuple node $\underline{u}$ and traverse the u_i with $i \neq j$ sequentially. For each $i \neq j$, we add a component node u_i of type $\lambda_j(i)$ as a child of $\underline{u}$, where $\lambda_j(i) = i$ for $i < j$ and $\lambda_j(i) = i - 1$ for $i > j$ (so that the component nodes $\{u_i\}$ with $i \neq j$ have types $1, \dots, \ell - 1$). If u_i is already contained in $\mathcal{T}_{v,t}$ then we 'ignore' this component node. Otherwise we add vertex nodes $w \in C_{u_i}(F)$ as children of u_i, see Figure 2. Note that $C_v(H_t)$ consists exactly of the union of all vertex nodes of $\mathcal{T}_{v,t}$, so

$$C_v(H_t^{\mathcal{R}}) \subseteq C_v(H_t) = \mathcal{V}_{v,t}. \tag{29}$$

The main point is that whenever no component nodes are ignored, then from $\mathcal{T}_{v,t}$ we can reconstruct all explored tuples (in E_t) and components (of F), which for size rules are the only ones relevant for determining the size of $C_v(H_t^{\mathcal{R}})$. In fact, up to relabellings, we can reconstruct these tuples and the relevant component sizes of F *without* looking at the vertex labels (the tree structure, including the node types, is enough). Motivated by this we say that $S_{j,w}$ is *bad* if one of the following conditions hold:

- $S_{j,w}$ contains some tuple $\underline{u} = (u_1, \dots, u_\ell)$ multiple times.
- $S_{j,w}$ contains a tuple $\underline{u} = (u_1, \dots, u_\ell)$ where u_i with $i \neq j$ is already a vertex node of $\mathcal{T}_{v,t}$ constructed so far.

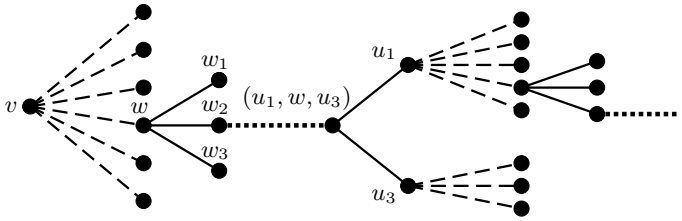

Figure 2: Example of the exploration tree $\mathcal{T}_{v,t}$ for $\ell = 3$. The children of the root vertex v are $w \in C_v(F)$ (vertex nodes), which in turn each have children w_1, w_2, w_3 (index nodes of types $1, 2, 3$). Every w_j has all (so far untested) tuples $\underline{u} = (u_1, u_2, u_3) \in E_t$ with $u_j = w$ as children (tuple nodes), whose descendants are component nodes u_i with $i \neq j$ (of types $1, 2$). If u_i is not already a vertex node of $\mathcal{T}_{v,t}$, then its children are $w \in C_{u_i}(F)$ (vertex nodes), for which we repeat the above construction.

- $S_{j,w}$ contains a tuple $\underline{u} = (u_1, \dots, u_\ell)$ where u_i and u_k with $i \neq k$ are in the same component of F (note that this holds for $u_i \in C_w(F)$ for $i \neq j$).
- $S_{j,w}$ contains tuples $\underline{u} = (u_1, \dots, u_\ell)$ and $\underline{v} = (v_1, \dots, v_\ell)$ for which u_i and v_k with $i, k \neq j$ are in the same component of F.

Otherwise $S_{j,w}$ is *good*; Observe that if $S_{j,w}$ is good, then in $\mathcal{T}_{v,t}$ none of w's component node descendants u_i with $\underline{u} = (u_1, \dots, u_\ell) \in S_{j,w}$ are ignored. For this reason we call $\mathcal{T}_{v,t}$ *good* if every $S_{j,w}$ is good. In the following we estimate the probability that $S_{j,w}$ is *bad*. Clearly, there are at most $n^{\ell-1}$ different tuples with $u_j = w$. Recalling that $\mathcal{V}_{v,t}$ denotes the vertex nodes of $\mathcal{T}_{v,t}$, there are at most $\ell n^{\ell-2}|\mathcal{V}_{v,t}|$ different tuples satisfying the second condition, and at most $\ell^2 n^{\ell-2}|L_1(F)|$ tuples to which the third condition applies. Similarly, there are at most $\ell^2 n^{2(\ell-2)+1}|L_1(F)|$ pairs of tuples which satisfy the last condition. Recall that the random variables $A_{\underline{u}}(t)$, which count the number of times $\underline{u}$ is in E_t, are independent for different tuples $\underline{u}$. So, using (23), (24) and $t \leqslant 1$, whenever $\max\{|\mathcal{V}_{v,t}|, |L_1(F)|\} \leqslant U$ holds we see that the probability of $S_{j,w}$ being bad is at most

$$n^{\ell-1} \cdot t^2/n^\ell + 2\ell^2 n^{\ell-2} U \cdot t/n^{\ell-1} + \ell^2 n^{2\ell-3} U \cdot t^2/n^{2(\ell-1)} \leqslant 4\ell^2 U/n. \qquad (30)$$

To understand the structural properties of $\mathcal{T}_{v,t}$ it will be useful to compare it with a closely related process that is simpler to analyze. Recall that when determining the $S_{j,w}$ we only consider so far untested tuples. Thus each $S_{j,w}$ is dominated (with respect to the subset relation) by $\tilde{S}_{j,w}$, where for each of the $n^{\ell-1}$ tuples $\underline{u} = (u_1, \dots, u_\ell)$ with $u_j = w$, independently, the number of its arrivals is given by a $\mathrm{Po}(t/n^{\ell-1})$ distribution. There is a natural coupling between $S_{j,w}$ and $\tilde{S}_{j,w}$ which only fails if $\tilde{S}_{j,w}$ contains $\underline{u}$ which are forbidden for $S_{j,w}$. Since each of these 'bad' tuples contains at least one vertex from $\mathcal{V}_{v,t}$, there are at most $\ell|\mathcal{V}_{v,t}|n^{\ell-2}$ of them. So, with (24) and $t \leqslant 1$ in mind,

by considering the probability that $\tilde{S}_{j,w}$ selects at least one of them, whenever $|\mathcal{V}_{v,t}| \leqslant U$ holds it follows that

$$\mathrm{d_{TV}}\big(S_{j,w}\,,\tilde{S}_{j,w}\big) \leqslant \ell|\mathcal{V}_{v,t}|n^{\ell-2}\cdot t/n^{\ell-1} \leqslant \ell U/n. \tag{31}$$

We now define $\mathfrak{T}_{v,t} = \mathfrak{T}_{v,t}(F)$ similarly to $\mathcal{T}_{v,t}$: we employ the same construction except that we use (independent copies of) $\tilde{S}_{j,w}$ instead of $S_{j,w}$ and always proceed as if $\tilde{S}_{j,w}$ is good. Since each $S_{j,w}$ is dominated by (may be regarded as a subset of) $\tilde{S}_{j,w}$, it follows that $\mathcal{T}_{v,t}$ is dominated by $\mathfrak{T}_{v,t}$ with respect to the subgraph relation. Denoting the set of vertex nodes of $\mathfrak{T}_{v,t}$ by $\mathfrak{V}_{v,t}$, we see that $\mathcal{V}_{v,t}$ is dominated by $\mathfrak{V}_{v,t}$.

The next lemma states that the number of vertex nodes in $\mathcal{T}_{v,t}$ and $\mathfrak{T}_{v,t}$ have (uniform) exponential decay.

Lemma 7. *Suppose that* (9) *and* (11)–(12) *hold with* $\beta > 1$. *There exist* $a, A > 0$ *(depending only on* $\ell, L, \sigma, \beta, B$*) such that for all* $0 \leqslant t \leqslant \sigma$ *and* $s \geqslant 0$ *we have* $\mathbb{P}(|\mathcal{V}_{v,t}| \geqslant s) \leqslant \mathbb{P}(|\mathfrak{V}_{v,t}| \geqslant s) \leqslant Ae^{-as}$, $\mathbb{E}N_{\geqslant s}(H_t^{\mathcal{R}}) \leqslant Ae^{-as}n$ *and* $\mathbb{P}(L_1(H_t^{\mathcal{R}}) \geqslant s) \leqslant Ae^{-as}n$.

Before giving the proof of this result, which is based on branching processes arguments, we use it to show that $\mathcal{T}_{v,t}$ and $\mathfrak{T}_{v,t}$ can be coupled so that they typically agree. Note that at distance $4i+1, 4i+2, 4i+3, 4i+4$ from the root $\mathcal{T}_{v,t}$ and $\mathfrak{T}_{v,t}$ always have vertex, index, tuple and component nodes.

Lemma 8. *Suppose* $n \geqslant n_0(\ell, L, \sigma, \beta, B)$ *and that the assumptions of Theorem 3 as well as* (27) *hold. There exists a coupling of* $\mathcal{T}_{v,t}$ *and* $\mathfrak{T}_{v,t}$ *so that with probability at least* $1 - (\log n)^4/n$ *we have* $\mathcal{T}_{v,t} = \mathfrak{T}_{v,t}$ *and* $\mathcal{T}_{v,t}$ *is good.*

Proof. We write T^i for the *restriction* of a rooted tree T to all vertices within distance at most i from the root. Let $\mathcal{V}_{v,t}^i$ and $\mathfrak{V}_{v,t}^i$ denote the vertex nodes in $\mathcal{T}_{v,t}^i$ and $\mathfrak{T}_{v,t}^i$, respectively. Recall that $U = (\log n)^{6/5}$. Since $\beta > 1$, note that $L_1(F) \leqslant U$ follows from (11) for $n \geqslant n_0(B, \beta)$.

We inductively couple $\mathcal{T}_{v,t}^{4i+1}$ and $\mathfrak{T}_{v,t}^{4i+1}$ for $0 \leqslant i \leqslant U$ so that with probability at least $1 - i \cdot 5\ell^3U^2/n$ we have either $\max\{|\mathcal{V}_{v,t}^{4i+1}|, |\mathfrak{V}_{v,t}^{4i+1}|\} \geqslant U$, or $\mathcal{T}_{v,t}^{4i+1} = \mathfrak{T}_{v,t}^{4i+1}$ with all $S_{j,w}$ of $\mathcal{T}_{v,t}^{4i+1}$ being good. The base case $i = 0$ is straightforward, as both use the same procedure for generating the root and its children. Now suppose that we have constructed $\mathcal{T}_{v,t}^{4i+1}$ and $\mathfrak{T}_{v,t}^{4i+1}$ coupled as above. In the following we sequentially consider vertex nodes w at distance $4i+1$ from the root and extend the coupling to their descendants with distance up to $4(i+1)+1$; here we clearly may assume $|\mathcal{V}_{v,t}^{4i+1}| = |\mathfrak{V}_{v,t}^{4i+1}| < U$. For each vertex node w we create ℓ index nodes $w_1, \dots w_\ell$ (of types $1, \dots, \ell$). We abandon our coupling whenever we have found more than U vertex nodes (in which case we are done), so (31) holds. Thus we can couple $S_{j,w}$ and $\tilde{S}_{j,w}$ so that they agree with probability at least $1 - \ell U/n$. Now we also abandon our coupling whenever $S_{j,w}$ is bad, which happens with probability at most $4\ell^2U/n$ by (30). The point is that given good $S_{j,w} = \tilde{S}_{j,w}$, in both cases the same deterministic construction is used for generating the descendants of w_j with distance up to $4(i+1)+1$ from the

root. So, by repeating this for $w_1, \dots, w_\ell$, with probability at least $1 - 5\ell^3 U/n$ we can couple the descendants of w with distance up to $4(i+1)+1$ from the root. Since we follow this argument for each of the at most U vertex nodes at distance $4i$ from the root, we see that we can extend our coupling to $\mathcal{T}_{v,t}^{4(i+1)+1}$ and $\mathfrak{T}_{v,t}^{4(i+1)+1}$ with probability at least $1 - 5\ell^3 U^2/n$, establishing the claim.

Finally, by Lemma 7 we know that $V = \max\{|\mathcal{V}_{v,t}|, |\mathfrak{V}_{v,t}|\} < U/10$ holds with probability at least, say, $1 - n^{-9}$ for $n \geqslant n_0(a, A)$. This together with the above coupling completes the proof (as there are no vertex nodes with distance larger than $4V + 1$ from the root). □

We now introduce an idealized 'infinite' version $\mathfrak{X}_{\varphi,t}$ of the exploration tree that is defined *without* reference to n or F, and in which 'bad' things (such as 'ignored' component nodes) cannot happen by definition. Let R be the random variable with $\mathbb{P}(R = k) = \varphi(k)$ for each $k \geqslant 1$, where φ is given by Theorem 3. We start $\mathfrak{X}_{\varphi,t}$ with a root node and add R vertex nodes as children. Then, given any vertex node, we deterministically create ℓ children (index nodes of types $1, \dots, \ell$). Each of these, independently, has $Z \sim \mathrm{Po}(t)$ children (tuple nodes). For each of these grandchildren we assign again (deterministically) $\ell - 1$ children (component nodes of types $1, \dots, \ell - 1$). All of these, independently, give birth to R many descendants (vertex nodes).

For our subsequent analysis it will be key to observe that if we are only interested in equality up to isomorphisms, then we can generate $\mathfrak{T}_{v,t}$ in a more convenient way, similarly to $\mathfrak{X}_{\varphi,t}$. Indeed, using standard properties of Poisson processes and noting that selecting a uniform tuple $\underline{u} = (u_1, \dots, u_\ell)$ with $u_j = w$ is equivalent to picking $\ell - 1$ random vertices, we can generate the descendants of w_j constructed by $\tilde{S}_{j,w}$ using the following three-generation tree process: the root has $Z \sim \mathrm{Po}(t)$ children (tuple nodes); then for each of the resulting children we construct (deterministically) $\ell - 1$ grandchildren (component nodes of types $1, \dots, \ell - 1$), which each in turn give birth to N descendants (vertex nodes), where $N \sim |C_u(F)|$ for a uniformly and independently chosen vertex u. Comparing the resulting construction with $\mathfrak{X}_{\varphi,t}$, it follows that we can generate $\mathfrak{T}_{v,t}$ up to relabellings in the same way as $\mathfrak{X}_{\varphi,t}$, with the only difference that we use N instead of R.

Proof of Lemma 7. Since $\mathcal{V}_{v,t}$ is dominated by (may be regarded as a subset of) $\mathfrak{V}_{v,t}$, we have $\mathbb{P}(|\mathcal{V}_{v,t}| \geqslant s) \leqslant \mathbb{P}(|\mathfrak{V}_{v,t}| \geqslant s)$. Using this inequality, we claim that it is enough to prove existence of $a, A > 0$ (depending only on $\ell, L, \sigma, \beta, B$) satisfying

$$\mathbb{P}(|\mathfrak{V}_{v,t}| \geqslant s) \leqslant Ae^{-as} \quad \text{for all } s \geqslant 0. \tag{32}$$

Indeed, recall that v is chosen uniformly at random, so that $\mathbb{P}(|C_v(H_t^{\mathcal{R}})| \geqslant s \mid H_t^{\mathcal{R}} = G) = N_{\geqslant s}(G)/n$. Taking expectations, we see that $\mathbb{E}N_{\geqslant s}(H_t^{\mathcal{R}}) = n\mathbb{P}(|C_v(H_t^{\mathcal{R}})| \geqslant s)$. Using (29) we have $|C_v(H_t^{\mathcal{R}})| \leqslant |\mathcal{V}_{v,t}|$, so $\mathbb{P}(|C_v(H_t^{\mathcal{R}})| \geqslant s) \leqslant Ae^{-as}$ by (32). Now Markov's inequality gives $\mathbb{P}(L_1(H_t^{\mathcal{R}}) \geqslant s) \leqslant Ae^{-as}n$.

In the remainder we establish (32) using Lemmas 5 and 6. Let Z_j be independent copies of $Z \sim \mathrm{Po}(t)$, and let $v_{j,r,k}$ be uniformly and independently

chosen random vertices. We henceforth construct $\mathfrak{T}_{v,t}$ up to relabellings, as described in the paragraph proceeding this proof. Given a vertex node w with distance $4i+1$ from the root, in this tree construction it has

$$W = \sum_{1 \leqslant j \leqslant \ell} \sum_{1 \leqslant r \leqslant Z_j} \sum_{1 \leqslant k \leqslant \ell - 1} |C_{v_{j,r,k}}(F)|$$

vertex node descendants at distance $4(i+1)+1$ from the root, where $\mathbb{E}(W) = \ell t(\ell-1)S(F) \leqslant \ell\sigma(\ell-1)L < 1$ due to $t \leqslant \sigma$ and (9). Note that $F_W(z) = [F_Z([F_N(z)]^{\ell-1})]^{\ell}$, where $N \sim |C_u(F)|$ for a uniformly chosen vertex u. By (11) we have $[F_N(\beta)]^{\ell-1} \leqslant B^{\ell-1}$. Now, since $Z \sim \mathrm{Po}(t)$ and $0 \leqslant t \leqslant \sigma$, it easily follows that $F_Z(z) = e^{t(z-1)} \leqslant e^{\sigma z}$ for $z \geqslant 0$, so $F_W(\beta) \leqslant \tilde{B} = \tilde{B}(\ell, \sigma, B)$. Let W^+ be the size of the Galton–Watson branching process in which each node, independently, has W children. Lemma 6 yields $F_{W^+}(\delta) \leqslant D$, where $\delta > 1$ and $D > 0$ depend only on $\ell, L, \sigma, \beta, \tilde{B}$. Since the distribution of W does not depend on the w or i considered above, it in particular follows that each vertex node with distance 1 from the root has W^+ vertex node descendants in $\mathfrak{T}_{v,t}$.

Finally, note that $\mathfrak{T}_{v,t}$ starts with a root vertex which gives birth to N vertex node children, each of whose vertex nodes descendants is given by independent copies of W^+. With this in mind $|\mathfrak{V}_{v,t}| \sim T$, where T is a two-generation branching process where the root has N children, and then each of these, independently, has W^+ children. Recall that $F_N(\beta) \leqslant \tilde{B}$ and $F_{W^+}(\delta) \leqslant D$ for $\beta, \delta > 1$ and $\tilde{B}, D > 0$. So, Lemma 5 yields (32) for $A = \tilde{B}$ and $a > 0$ depending only on $\beta, \delta, \tilde{B}, D$. As explained, this completes the proof. □

Recall that $\mathfrak{X}_{\varphi,t}$ uses the same construction as $\mathfrak{T}_{v,t}$, with the difference that it employs R instead of N. When establishing the exponential decay in the proof of Lemma 7, note that the only properties of N used are $\mathbb{E}N = S(F) \leqslant L$ and $F_N(\beta) \leqslant B$. Since $\mathbb{E}R = \chi(\varphi) \leqslant L$ and $F_R(\beta) \leqslant B$ by (6)–(8), the same argument thus carries over word-by-word when applied to the vertex nodes of $\mathfrak{X}_{\varphi,t}$, which we denote by $\mathfrak{V}_{\varphi,t}$.

Lemma 9. *Suppose that* (6)–(8) *and* (9) *hold with* $\beta > 1$. *There exist* $a, A > 0$ *(depending only on* $\ell, L, \sigma, \beta, B$*) such that for all* $0 \leqslant t \leqslant \sigma$ *and* $s \geqslant 0$ *we have* $\mathbb{P}(|\mathfrak{V}_{\varphi,t}| \geqslant s) \leqslant Ae^{-as}$, *where* a, A *are defined in the same way as in Lemma 7.* □

After these preparations, we are now ready to show that we can couple $\mathcal{T}_{v,t}$ and $\mathfrak{X}_{\varphi,t}$ so that they typically agree up to isomorphisms (by using $\mathfrak{T}_{v,t}$ as an 'intermediate' process).

Lemma 10. *Suppose* $n \geqslant n_0(\ell, L, \sigma, \beta, B)$ *and that the assumptions of Theorem 3 as well as* (27) *hold. There exists a coupling of* $\mathcal{T}_{v,t}$ *and* $\mathfrak{X}_{\varphi,t}$ *so that with probability at least* $1 - (\log n)^{C+5} n^{-1/2}$ *we have* $\mathcal{T}_{v,t} \cong \mathfrak{X}_{\varphi,t}$ *and* $\mathcal{T}_{v,t}$ *is good.*

Proof. Recall that $U = (\log n)^{6/5}$. By Lemma 8 it suffices to couple $\mathfrak{T}_{v,t}$ and $\mathfrak{X}_{\varphi,t}$ so that with probability at least $1 - 4\ell^2 U^4 (\log n)^C n^{-1/2}$ we have $\mathfrak{T}_{v,t} \cong \mathfrak{X}_{\varphi,t}$. To this end we use a similar but simpler argument as in the proof of Lemma 8,

inductively extending our coupling from distance $4i+1$ to $4(i+1)+1$ from the root. As before, using Lemma 7 and 9 we can safely abandon our coupling whenever we have seen at least U vertex nodes, or when we reach distance U from the root. In the inductive step, the only difference between $\mathfrak{X}_{\varphi,t}$ and $\mathfrak{T}_{v,t}$ is that $\mathfrak{X}_{\varphi,t}$ uses R whereas $\mathfrak{T}_{v,t}$ uses N. Recall that $\mathbb{P}(R=k)=\varphi(k)$ and $\mathbb{P}(N=k)=N_k(F)/n$. It is not difficult to see that (7) and (11) imply $\mathbb{P}(R \geqslant U) \leqslant n^{-2}$ and $\mathbb{P}(N \geqslant U) = 0$ for $n \geqslant n_0(\beta, B)$. Using these tail estimates together with (10), by distinguishing values smaller and larger than U we obtain

$$\mathrm{d_{TV}}\big(R\,,N\big) \leqslant U \cdot (\log n)^C n^{-1/2} + n^{-2} \leqslant 2U(\log n)^C n^{-1/2}. \tag{33}$$

We furthermore may safely abandon our coupling whenever some index node has $Z \geqslant U$ children, since (using $t \leqslant \sigma$) this occurs with probability at most n^{-9} for $n \geqslant n_0(\sigma)$. The point is that this ensures that we only need to couple R and N at most $\ell^2 U^2$ times when going from distance $4i+1$ to $4(i+1)+1$. So, each time we can extend the coupling inductively with probability at least, say, $1 - 3\ell^2 U^3 (\log n)^C n^{-1/2}$. Arguing as in the proof of Lemma 8, this completes the coupling argument. □

2.4.3 Expected component sizes

After analyzing the tuple and component structure induced by E_t, we now consider the second exposure round, where the selected tuples are presented in random order to $\mathcal{R}$. Intuitively, the coupling given by Lemma 10 allows us to estimate $\mathbb{E}N_k(H_t^{\mathcal{R}})$ using $\mathfrak{X}_{\varphi,t}$. As we shall see, this also carries over to $\mathbb{E}N_k(F_{tn}^{\mathcal{R}})$.

Recall that if the exploration tree $\mathcal{T}_{v,t} \cong T$ is good, then during its construction no component nodes are ignored. As mentioned in Section 2.4.2, the key point is that if no nodes are ignored (i.e., all component nodes have at least one child), then from the structure of T (which includes the vertex types) we can reconstruct all tuples in E_t and component sizes of F (up to relabellings) which are relevant for determining $|C_v(H_t^{\mathcal{R}})|$. We denote the corresponding set of tuples and component sizes by $\mathcal{T}_T$ and $\mathcal{C}_T$, respectively. As the above 'reconstruction' procedure only uses the tree-structure of T, it in fact can be applied to any exploration tree in which each component node has at least one child; so, in particular, to $\mathfrak{X}_{\varphi,t} \cong T$. In the following we define $|C^{\mathcal{R}}(T)|$ for any exploration tree T, where we formally set $|C^{\mathcal{R}}(T)| = 0$ if T contains a component node with 0 descendants. Otherwise, we traverse in (uniform) random order the tuples in $\mathcal{T}_T$; for each tuple we present the component sizes of its vertices to $\mathcal{R}$ and update the list of components (and their sizes) according to the decisions of $\mathcal{R}$ (by adding the pairs selected by $\mathcal{R}$). Finally, we define $|C^{\mathcal{R}}(T)|$ as the size of the resulting component which contains the root vertex of T. Since the second exposure round of $H_t^{\mathcal{R}}$ presents the tuples in E_t to $\mathcal{R}$ in random order, a moment's thought reveals that conditional on $\mathcal{T}_{v,t} \cong T$ being good, both $|C_v(H_t^{\mathcal{R}})|$ and $|C^{\mathcal{R}}(\mathcal{T}_{v,t})|$ have exactly the same distribution for size rules. So, for all $k \geqslant 1$

we have

$$\mathbb{P}(|C_v(H_t^{\mathcal{R}})| = k \mid \mathcal{T}_{v,t} \cong T \text{ is good}) = \mathbb{P}(|C^{\mathcal{R}}(\mathcal{T}_{v,t})| = k \mid \mathcal{T}_{v,t} \cong T \text{ is good}). \tag{34}$$

Before using this observation to estimate $\mathbb{E}N_k(H_t^{\mathcal{R}})$, we first collect some basic properties of the function ρ, where we set

$$\rho(k,t) = \mathbb{P}(|C^{\mathcal{R}}(\mathfrak{X}_{\varphi,t})| = k) \quad \text{for all } (k,t) \in \mathbb{N} \times \mathbb{R}^+. \tag{35}$$

Lemma 11. *Suppose that* (6)–(8) *and* (9) *hold with* $\beta > 1$. *The function* $\rho : \mathbb{N} \times \mathbb{R}^+ \to [0,1]$ *defined in* (35) *depends only on* $\varphi, \mathcal{R}, \ell$ *and satisfies* $\sum_{k \geqslant 1} \rho(k,t) = 1$ *for all* $0 \leqslant t \leqslant \sigma$. *Furthermore, there exist* $a, A > 0$ *(depending only on* $\ell, L, \sigma, \beta, B$*) such that for all* $0 \leqslant t \leqslant \sigma$ *and* $s \geqslant 0$ *we have* $\rho(s,t) \leqslant Ae^{-as}$, *where* a, A *are given by Lemma 9.*

Proof. The definitions of $C^{\mathcal{R}}(\cdot)$ and of $\mathfrak{X}_{\varphi,t}$ depend only on $\mathcal{R}, \ell$ and on $\varphi, \mathcal{R}, \ell, t$ respectively. So, from (35) we see that $\rho : \mathbb{N} \times \mathbb{R}^+ \to [0,1]$ depends only on $\varphi, \mathcal{R}, \ell$. Since the component containing the root vertex of $\mathfrak{X}_{\varphi,t}$ can only contain vertex nodes of $\mathfrak{X}_{\varphi,t}$, we see that $1 \leqslant |C^{\mathcal{R}}(\mathfrak{X}_{\varphi,t})| \leqslant |\mathfrak{V}_{\varphi,t}|$ holds, from which $\rho(0,t) = 0$ follows. Furthermore, Lemma 9 implies $\rho(s,t) \leqslant \mathbb{P}(|\mathfrak{V}_{\varphi,t}| \geqslant s) \leqslant Ae^{-as}$ for all $s \geqslant 1$, where $a, A > 0$ depend only on $\ell, L, \sigma, \beta, B$. Similarly, for all $s \geqslant 0$ we have $\mathbb{P}(\mathfrak{X}_{\varphi,t} \text{ is infinite}) \leqslant \mathbb{P}(|\mathfrak{V}_{\varphi,t}| \geqslant s) \leqslant Ae^{-as}$. But $Ae^{-as} \to 0$ as $s \to \infty$, so $\mathbb{P}(\mathfrak{X}_{\varphi,t} \text{ is infinite}) = 0$, which in turn yields $\sum_{k \geqslant 1} \rho(k,t) = 1$. □

Lemma 12. *Suppose* $n \geqslant n_0(\ell, L, \sigma, \beta, B)$ *and that the assumptions of Theorem 3 as well as* (27) *hold. We have*

$$\mathbb{E}N_k(H_t^{\mathcal{R}}) = \rho(k,t)n \pm (\log n)^{C+6} n^{1/2} \quad \textit{for all } k \geqslant 1. \tag{36}$$

Proof. Similar as in the proof of Lemma 7, since v is chosen uniformly at random we have $\mathbb{E}N_k(H_t^{\mathcal{R}}) = n\mathbb{P}(|C_v(H_t^{\mathcal{R}})| = k)$. To prove the claim it thus suffices to relate $\mathbb{P}(|C_v(H_t^{\mathcal{R}})| = k)$ and $\rho(k,t) = \mathbb{P}(|C^{\mathcal{R}}(\mathfrak{X}_{\varphi,t})| = k)$. The coupling of Lemma 10 implies that $\mathfrak{X}_{\varphi,t} \cong \mathcal{T}_{v,t}$ holds with probability at least $1 - (\log n)^{C+5} n^{-1/2}$ for $n \geqslant n_0(\ell, L, \sigma, \beta, B)$. Hence

$$\mathbb{P}(|C^{\mathcal{R}}(\mathcal{T}_{v,t})| = k) = \mathbb{P}(|C^{\mathcal{R}}(\mathfrak{X}_{\varphi,t})| = k) \pm 2(\log n)^{C+5} n^{-1/2}.$$

Since this coupling also implies that $\mathcal{T}_{v,t}$ is good, using (34) it follows that

$$\mathbb{P}(|C_v(H_t^{\mathcal{R}})| = k) = \mathbb{P}(|C^{\mathcal{R}}(\mathcal{T}_{v,t})| = k) \pm 2(\log n)^{C+5} n^{-1/2}.$$

Finally, combining our findings and recalling (35), we readily obtain (36). □

Now we relate $H_t^{\mathcal{R}}$ with $F_{tn}^{\mathcal{R}}$ by establishing that $\mathbb{E}N_k(H_t^{\mathcal{R}}) \approx \mathbb{E}N_k(F_{tn}^{\mathcal{R}})$.

Lemma 13. *Suppose that* $0 \leqslant t \leqslant 1$. *Then for* $n \geqslant n_0(\ell)$ *we have*

$$\mathbb{E}N_k(F_{tn}^{\mathcal{R}}) = \mathbb{E}N_k(H_t^{\mathcal{R}}) \pm k(\log n) n^{1/2} \quad \textit{for all } k \geqslant 1. \tag{37}$$

Proof. Observe that N_k changes by at most ℓk per step. So, for $r \leqslant s$ we have $\mathbb{E}(N_k(F_s^{\mathcal{R}}) \mid F_r^{\mathcal{R}} = G) = N_k(G) \pm (s-r)\ell k$. Taking expectations and restricting our attention to $r \in \{tn - i, tn\}$ shows that for each $i \geqslant 0$ we have

$$\mathbb{E}N_k(F_{tn\pm i}^{\mathcal{R}}) = \mathbb{E}N_k(F_{tn}^{\mathcal{R}}) \pm \ell k i. \tag{38}$$

Set $s = 3\sqrt{n \log n}$. Using $t \leqslant 1$, standard Chernoff bounds yield that $|E_t| = tn \pm s$ with probability at least $1 - n^{-2}$ for $n \geqslant n_0$. Combining this with (25) and (38), we readily obtain

$$\mathbb{E}N_k(H_t^{\mathcal{R}}) = \mathbb{E}N_k(F_{tn}^{\mathcal{R}}) \pm \ell k s \pm n \cdot n^{-2},$$

which implies (37) for $n \geqslant n_0(\ell)$, with room to spare. □

2.4.4 Concentration of component sizes

In this section we establish concentration of $N_k(F_i^{\mathcal{R}})$ around its expected value. The main technical difficulty here is that few changes of the offered tuples might alter many decisions of size rules (as the component sizes observed in later rounds can change); as we shall see, the bounds for $L_1(\cdot)$ implied by Lemma 7 will be a crucial ingredient for showing that this is typically not the case.

Lemma 14. *Suppose $n \geqslant n_0(\ell, L, \sigma, \beta, B)$ and that the assumptions of Theorem 3 hold. With probability at least $1 - n^{-250}$, for every $0 \leqslant i \leqslant \sigma n$ we have*

$$N_k(F_i^{\mathcal{R}}) = \mathbb{E}N_k(F_i^{\mathcal{R}}) \pm (\log n)^2 n^{1/2} \quad \textit{for all } 1 \leqslant k \leqslant (\log n)^2. \tag{39}$$

Proof. We sequentially draw σn random tuples and consider two associated graph sequences $F_i^{\mathcal{R}}$ and $F_i^{\mathcal{I}}$, where the 'influence' rule $\mathcal{I}$ in each step simply joins all ℓ randomly chosen vertices by edges. Note that $F_i^{\mathcal{R}} \subseteq F_i^{\mathcal{I}}$ always holds. Let $\mathcal{L}$ denote the event that $L_1(F_{\sigma n}^{\mathcal{I}}) < U = (\log n)^{6/5}$, which by monotonicity implies $L_1(F_i^{\mathcal{I}}) < U$ for all $0 \leqslant i \leqslant \sigma n$. Combining Lemma 7 with (26), for $n \geqslant n_0(a, A, \sigma)$ we have, say,

$$\mathbb{P}(\neg\mathcal{L}) \leqslant 3\sqrt{\sigma n} \cdot \mathbb{P}(L_1(H_\sigma^{\mathcal{I}}) \geqslant U) \leqslant n^{-300}. \tag{40}$$

For every $1 \leqslant i \leqslant \sigma n$ let $X_{k,i}$ denote the number of vertices which satisfy $|C_v(F_i^{\mathcal{R}})| = k$ and $|C_v(F_i^{\mathcal{I}})| < U$. When $\mathcal{L}$ holds no vertices are 'ignored' due to $|C_v(F_i^{\mathcal{I}})| \geqslant U$, so we have $X_{k,i} = N_k(F_i^{\mathcal{R}})$. Together with (40) this readily gives, say, $\mathbb{E}X_{k,i} = \mathbb{E}N_k(F_i^{\mathcal{R}}) \pm n^{-1}$. So, for $\Delta = U^{3/2} n^{1/2}$ it follows that

$$\mathbb{P}(\{|N_k(F_i^{\mathcal{R}}) - \mathbb{E}N_k(F_i^{\mathcal{R}})| \geqslant 2\Delta\} \cap \mathcal{L}) \leqslant \mathbb{P}(|X_{k,i} - \mathbb{E}X_{k,i}| \geqslant \Delta). \tag{41}$$

Note that for every size rule $\mathcal{R}$ the random variable $X_{k,i}$ can be written as $X_{k,i} = f(\underline{v}_1, \ldots, \underline{v}_i)$, where the $\underline{v}_j$ denote the ℓ-tuples generated by the ℓ-vertex process in each step (uniformly and independently). We claim that the function f satisfies $|f(\omega) - f(\tilde{\omega})| \leqslant 4\ell U$ whenever ω and $\tilde{\omega}$ differ in one coordinate, i.e., in one tuple. Assuming that $(\underline{v}_1, \ldots, \underline{v}_i)$ yield $F_i^{\mathcal{R}}$ and $F_i^{\mathcal{I}}$, respectively, let $\tilde{F}_i^{\mathcal{R}}$ and

$\tilde{F}_i^{\mathcal{I}}$ denote the graphs which result by changing $\underline{v}_j$ to $\underline{\tilde{v}}_j$. Since $F_i^{\mathcal{I}}$ and $\tilde{F}_i^{\mathcal{I}}$ only differ in the edges induced by $\underline{v}_j$ and $\underline{\tilde{v}}_j$, there is a set of vertices W containing at most 2ℓ components in each of $F_i^{\mathcal{I}}$ and $\tilde{F}_i^{\mathcal{I}}$ so that outside of W the component structure of both graphs is the same (to see this note that the order is irrelevant for $\mathcal{I}$, so we may assume $i = j$; then defining W as the union of the components containing the vertices of $\underline{v}_j$ and $\underline{\tilde{v}}_j$ in $F_{i-1}^{\mathcal{I}} = \tilde{F}_{i-1}^{\mathcal{I}}$ suffices). The key point is now that for size rules the decisions of $\mathcal{R}$ in $F_i^{\mathcal{R}}$ and $\tilde{F}_i^{\mathcal{R}}$ are the same for all tuples which contain no vertices from W (indeed, if a decision of $\mathcal{R}$ is modified then any changes of the resulting component sizes can only 'propagate' inside the components of $F_i^{\mathcal{I}}$ and $\tilde{F}_i^{\mathcal{I}}$; so only tuples containing vertices from W can be affected). It follows that the component structure outside of W is also the same in $F_i^{\mathcal{R}}$ and $\tilde{F}_i^{\mathcal{R}}$. Recall that W contains at most 2ℓ components in each of $F_i^{\mathcal{I}}$ and $\tilde{F}_i^{\mathcal{I}}$. So, since $X_{k,i}$ only counts those vertices v with $|C_v(F_i^{\mathcal{I}})| < U$, we see that a change of one tuple can alter f by at most $2 \cdot 2\ell \cdot U$, as claimed. So, recalling that $1 \leqslant i \leqslant \sigma n$, for $n \geqslant n_0(\ell, \sigma)$ McDiarmid's inequality [15] implies

$$\mathbb{P}(|X_{k,i} - \mathbb{E}X_{k,i}| \geqslant \Delta) \leqslant \exp\left(-\frac{2\Delta^2}{i(4\ell U)^2}\right) \leqslant n^{-300} \quad . \tag{42}$$

Finally, after combining (40)–(42), taking a union bound to account for all choices of $1 \leqslant i \leqslant \sigma n$ and $1 \leqslant k \leqslant (\log n)^2$ completes the proof (noting that the claim is trivial for $i = 0$). □

Using the main idea of the above proof we can directly show that $\mathbb{E}N_k(F_i^{\mathcal{R}})$ is essentially independent of the initial graph $F_0^{\mathcal{R}} = F$ for $i \leqslant \sigma n$: for any two graphs $F, \tilde{F}$ satisfying the assumptions of Theorem 3 their expected values can differ by at most, say, $(\log n)^{C+3} n^{1/2}$. The key point is that for such graphs we can construct a bijection Ψ between their vertex sets which, up to an exceptional set W of at most, say, $4U(\log n)^C n^{1/2}$ vertices, preserves the component structure of F and $\tilde{F}$, respectively. Now, using Ψ we couple $F_i^{\mathcal{R}}, F_i^{\mathcal{I}}$ and $\tilde{F}_i^{\mathcal{R}}, \tilde{F}_i^{\mathcal{I}}$ in a measure preserving way. Since changes can only propagate inside the components of the 'influence' graphs, only those vertices whose components in $F_i^{\mathcal{I}}$ or $\tilde{F}_i^{\mathcal{I}}$ contain vertices of W or $\Psi(W)$ can be 'spoiled'. Intuitively, since the components usually have size at most U, under this coupling N_k thus typically differs by at most $2|W| \cdot U$ for both graphs. Taking the error probability of $\max\{L_1(F_{\sigma n}^{\mathcal{I}}), L_1(\tilde{F}_{\sigma n}^{\mathcal{I}})\} < U$ into account, the claim now follows without much work.

2.4.5 Putting things together

In this section we combine our findings to prove Theorem 3. Lemma 11 easily implies the first part, i.e., existence of $\rho : \mathbb{N} \times \mathbb{R}^+ \to [0, 1]$ with the desired properties. Indeed, it ensures that for every $\sigma \geqslant 0$ satisfying (9) there exist $a, A > 0$ (depending only on $\ell, L, \sigma, \beta, B$) such that for every $t \in [0, \sigma]$ we have $\sum_{k \geqslant 1} \rho(k, t) = 1$ and

$$\rho(s, t) \leqslant Ae^{-as} \quad \text{for all } s \geqslant 0. \tag{43}$$

For later usage we remark that Lemma 7 holds for the same a, A. Let $D = 300/a > 0$ and $\tilde{\beta} = \min\{e^{a/2}, e^{1/(4D)}\} > 1$. Now, using (43) we see that

$$\sum_{k\geqslant 1} \rho(k,t)\tilde{\beta}^k \leqslant A\sum_{k\geqslant 1} e^{-ak/2} = \tilde{B} - 1, \tag{44}$$

with $1 < \tilde{B} < \infty$ depending only on a, A. Similarly, we obtain

$$\chi(\rho,t) = \sum_{k\geqslant 1} k\rho(k,t) \leqslant A\sum_{k\geqslant 1} ke^{-ak} = \tilde{L} - 1, \tag{45}$$

with $1 < \tilde{L} < \infty$ depending only on a, A. Summarizing, equations (6)–(8) hold when $\beta, B, L, \varphi(\cdot)$ are replaced by $\tilde{\beta}, \tilde{B}, \tilde{L}, \rho(\cdot, t)$, with room to spare.

Turning to properties of $F_i^{\mathcal{R}}$, from Lemmas 12–14 it follows that with probability at least $1 - n^{-250}$, for every $0 \leqslant i \leqslant \sigma n$ (by considering $t = i/n \in [0, \sigma]$) we have

$$N_k(F_i^{\mathcal{R}}) = \rho(k, i/n)n \pm 3(\log n)^{C+6}n^{1/2} \quad \text{for all } 1 \leqslant k \leqslant (\log n)^2 \tag{46}$$

for $n \geqslant n_0(\ell, L, \sigma, \beta, B)$. Recall that Lemma 7 holds with the a, A chosen above. By definition of D it follows that, with probability at least $1 - n^{-250}$, we have

$$L_1(F_{\sigma n}^{\mathcal{R}}) \leqslant D\log n \tag{47}$$

for $n \geqslant n_0(A)$. In the remainder we assume that (46)–(47) hold. Recalling (43) and the definition of D, note that for all $k \geqslant D\log n$ and $0 \leqslant i \leqslant \sigma n$ we have, say, $|\rho(k, i/n)| \leqslant n^{-9}$ for $n \geqslant n_0(A)$. Using (47) it follows that

$$N_k(F_i^{\mathcal{R}}) = \rho(k, i/n)n \pm (\log n)^C n^{1/2} \quad \text{for all } k \geqslant D\log n \text{ and } 0 \leqslant i \leqslant \sigma n.$$

Together with (46), for every $0 \leqslant i \leqslant \sigma n$ this establishes (10) with $C, F, \varphi(\cdot)$ replaced by $\tilde{C}, F_i^{\mathcal{R}}, \rho(\cdot, i/n)$ for $n \geqslant n_0(D)$, where $\tilde{C} = C + 9$. Now, using (44) and (46)–(47) we see that for every $0 \leqslant i \leqslant \sigma n$ we have (as $i/n \in [0, \sigma]$)

$$\begin{aligned}\sum_{k\in[n]} N_k(F_i^{\mathcal{R}})\tilde{\beta}^k &\leqslant n \sum_{1\leqslant k\leqslant D\log n} \rho(k, i/n)\tilde{\beta}^s + 3D(\log n)^{C+6}n^{1/2} \sum_{1\leqslant k\leqslant D\log n} \tilde{\beta}^k \\ &\leqslant n\sum_{k\geqslant 1} \rho(k, i/n)\tilde{\beta}^k + 3D^2(\log n)^{C+7}n^{3/4} \leqslant \tilde{B}n\end{aligned}$$

for $n \geqslant n_0(C, D)$, which establishes (11) with β, B, F replaced by $\tilde{\beta}, \tilde{B}, F_i^{\mathcal{R}}$. It remains to show that (12)–(13) hold. Recall that $S(G) = \sum_{k\in[n]} kN_k(G)/n$. Now, assuming $n \geqslant n_0(a, A, D)$ and using (45)–(47) similarly as above, for every $0 \leqslant i \leqslant \sigma n$ we have

$$\begin{aligned}S(F_i^{\mathcal{R}}) &= \sum_{k\geqslant 1} k\rho(k, i/n) \pm A\sum_{k\geqslant D\log n} ke^{-ak} \pm 3D^2(\log n)^{C+8}n^{-1/2} \\ &= \chi(\rho, i/n) \pm 4D^2(\log n)^{C+8}n^{-1/2},\end{aligned}$$

which establishes (13) for $n \geqslant n_0(D)$. Finally, recalling (45), it follows that (12) holds with L, F replaced by $\tilde{L}, F_i^{\mathcal{R}}$ for $n \geqslant n_0(C, D)$, which completes the proof of Theorem 3.

3 When does $t_b = t_c$?

In this section we discuss Conjecture 2, first showing that it does hold for many size rules and then, in Section 3.2, showing that it cannot be extended to general ℓ-vertex rules, i.e., that the critical point where the susceptibility blows up need not always coincide with the percolation threshold.

3.1 Rules with uniform random edges

It is well known, and not hard to check, that under suitable assumptions the graph $F_{\theta n}$ given by adding θn independent and uniformly random edges to a given n-vertex initial graph F can be viewed as an instance of the inhomogeneous random graph model of Bollobás, Janson and Riordan [7]. To make this precise, consider instead the (multi-)graph $\widetilde{F}_\theta$ obtained from F by adding a Poisson number $\mathrm{Po}(2\theta/n)$ of copies of each of the $\binom{n}{2}$ possible edges, with these numbers independent; we may then ignore multiple edges, as we are only interested in the component structure. Since $\mathrm{Po}(\theta(n-1))$ edges are added in total, and there will be few multiple edges, $\widetilde{F}_\theta$ and $F_{\theta n}$ are essentially interchangeable (one may use domination arguments comparing them for different θ to make this precise). Given two components C_1 and C_2 of F, the number of edges between them in $\widetilde{F}_\theta$ is Poisson with mean $|C_1||C_2|2\theta/n$. Making (for a change) the n dependence explicit, let H_n be the random graph whose vertices are the components of F, with an edge between two vertices if these components are joined by an edge of $\widetilde{F}_\theta$. We say that a vertex of H_n has *type* k if the corresponding component of F has k vertices. Then the probability of an edge between a given type-i vertex and a given type-j vertex of H_n is $1-e^{-2\theta ij/n}$, which is around $2\theta ij/n$ if i and j are not too big, and the events that different edges are present are independent.

More precisely, let $\kappa(i,j) = 2\theta ij$ for all positive integers i and j. Suppose that μ is a finite measure on $\mathbb{Z}^+$, i.e., that $\mu_k = \mu(\{k\}) \geqslant 0$ for all k and $0 < \sum_{k\geqslant 1} \mu_k < \infty$. Let $F = F_n$ be a random n-vertex starting graph. Suppose that, for each fixed $k \geqslant 1$,

$$\frac{N_k(F_n)}{kn} \xrightarrow{\mathrm{p}} \mu_k \tag{48}$$

as $n \to \infty$, i.e., that H_n has asymptotically $\mu_k n$ vertices of type k, and that

$$\sum_{k\geqslant 1} k\mu_k = 1. \tag{49}$$

Then one can use (49), the fact that F_n has n vertices and (48) to show that whenever $K(n) \to \infty$ we have

$$N_{\geqslant K(n)}(F_n)/n \xrightarrow{\mathrm{p}} 0, \tag{50}$$

and it follows that for any $A \subset \mathbb{Z}^+$ we have

$$\sum_{k\in A} \frac{N_k(F_n)}{kn} \xrightarrow{\mathrm{p}} \sum_{k\in A} \mu_k. \tag{51}$$

In the terminology of [7], this means that the (random) sets of vertices of the graphs H_n, together with their types, form a *generalized vertex space* on the *generalized ground space* $(\mathbb{Z}^+, \mu)$. Taking $A = \mathbb{Z}^+$ in (51), we have in particular that $|H_n|/n \xrightarrow{\mathrm{p}} \mu(\mathbb{Z}^+) \in (0, \infty)$. By (49) the function κ forms an integrable kernel on the ground space $(\mathbb{Z}^+, \mu)$, with integral 2θ. Finally, the technical 'graphicality' condition of [7] is met since $\widetilde{F}_\theta$ has asymptotically θn edges. It follows that under these assumptions, the results of [7] apply to H_n (see Remark 2.4 there). The most important of these results is [7, Theorem 3.1], which tells us that H_n will whp contain a giant component (one with $\Theta(n)$ vertices) if and only if $||T_\kappa|| > 1$, where T_κ is a certain integral operator associated to κ. In particular, if $||T_\kappa|| > 1$ then there is some constant $\alpha = \alpha(\kappa, \mu) > 0$ (anything smaller than the quantity $\rho(\kappa)$ in [7]) such that whp H_n has a component with at least αn vertices. For the particular κ considered here, which is 'rank 1', we have $||T_\kappa|| = \sum_k 2\theta k^2 \mu_k$; see (16.8) in [7]. Note that if H_n contains a component with at least αn vertices, then so does $\widetilde{F}_\theta$ – the union of the components of F corresponding to these vertices of H_n. So, in short, if (48) and (49) hold, then $\widetilde{F}_\theta$ will have a giant component (whp) if (and, one can check, only if) $\sum_k 2\theta k^2 \mu_k > 1$. Moreover, it is not hard to check that these conclusions remain true if we delete some subset of the components of F_n, and adjust μ, as long as (48) holds for the new graph and $\Theta(n)$ components remain; this is because (51) still holds, and the kernel is still graphical.

We shall apply the observations above with initial graph $F = F_n = G^{\mathcal{R}}_{t_\mathrm{b} n}$, where $\mathcal{R}$ is some ℓ-vertex size rule. By (5), the condition (48) holds with $\mu_k = \rho_k(t_\mathrm{b})/k$. Furthermore, as noted after (5), we have $\sum_k \rho_k(t_\mathrm{b}) = 1$, which gives (49). Finally, note that

$$||T_\kappa|| = \sum_k 2\theta k \rho_k(t_\mathrm{b}) = 2\theta s(t_\mathrm{b}) = \infty, \tag{52}$$

since $s(t) = \sum_k \rho_k(t)$ diverges at $t = t_\mathrm{b}$. So far this tells us only that if we run any size rule up to time $t = t^{\mathcal{R}}_\mathrm{b}$ and then switch to adding uniformly random edges, after any constant times n further edges a giant component will emerge. The key point is that variants of this argument can be used to study the further evolution of $G^{\mathcal{R}}_i$ for suitable rules $\mathcal{R}$. A related approach was taken in [21] and [12].

Theorem 15. *Let $\mathcal{R}$ be a bounded-size ℓ-vertex rule. Then the conclusion of Conjecture 2 holds for $\mathcal{R}$; in particular, $t^{\mathcal{R}}_\mathrm{c} = t^{\mathcal{R}}_\mathrm{b}$, and moreover for any $\varepsilon > 0$ there is an $\alpha > 0$ such that whp $L_1(G^{\mathcal{R}}_{(t_\mathrm{b}+\varepsilon)n}) \geqslant \alpha n$.*

Note that this result was proved for some bounded-size 4-vertex rules (ones in which either v_1v_2 or v_3v_4 is added) already by Spencer and Wormald [21].

Proof. By definition of bounded-size rules, there is a constant B such that $\mathcal{R}$ treats all components of size greater than B in the same way. Consider the graph $G^{\mathcal{R}}_{t_\mathrm{b} n}$ generated by the rule after $t_\mathrm{b} n$ steps. Let W be the set of vertices of this graph in components of size greater than B, and let $F = F_n$ be the subgraph

of $G^{\mathcal{R}}_{t_\mathrm{b}n}$ induced by W. Noting that $s(t_\mathrm{b}) = \sum_k k\rho_k(t_\mathrm{b}) = \infty > B$, we have $\rho_k(t_\mathrm{b}) > 0$ for some $k > B$, and it follows that for some constant $\beta > 0$, we have $|W| \geqslant \beta n$ whp. From now on we assume that this is the case. In all subsequent steps of our original process $G^{\mathcal{R}}_i$, every vertex of W is in a component of size greater than B. Fix $\varepsilon > 0$. Let us call a step *good* if in this step all ℓ selected vertices are in W. Then each step is good with probability at least β^ℓ, and it follows that whp at least θn of the next εn steps are good, where $\theta = \varepsilon\beta^\ell/2$ is a positive constant. Again using the definition of a bounded-size rule, in each good step at least one edge is added and by symmetry it is chosen uniformly at random from all possible edges with ends in W. It follows that we may couple $G^{\mathcal{R}}_{(t_\mathrm{b}+\varepsilon)n}$ and $\widetilde{F}_\theta$ so that whp the former contains the latter as a subgraph. But F satisfies the assumptions above with $\mu_k = \rho_k(t_\mathrm{b})/k$ for $k > B$ and $\mu_k = 0$ for $k \leqslant B$. Since the sum in (52) remains infinite after removing the first B terms, Theorem 3.1 of [7] and the discussion above imply that for some positive α, whp $\widetilde{F}_\theta$ contains a component with at least αn vertices. □

Our next result concerns a different generalization of the Bohman–Frieze process [4]. Let us call an Achlioptas rule $\mathcal{R}$ *take-it-or-leave-it* if, when presented with a choice of two edges e_1 and e_2, the rule decides which to select depending only on the current graph and on e_1. In other words, the rule first sees e_1 and must decide whether to take this edge of not; if not, it selects the uniformly random edge e_2. Bounded-size rules of this type were studied, for example, by Bohman and Kravitz [6]; here we do not assume that the rule is bounded-size.

Theorem 16. *Let $\mathcal{R}$ be a take-it-or-leave-it size rule. Then the conclusions of Conjecture 2 and Theorem 15 hold for $\mathcal{R}$.*

Proof. Consider the process $(G^{\mathcal{R}}_{t_\mathrm{b}n+i})_{i\geqslant 0}$, i.e., our Achlioptas process started at step $t_\mathrm{b}n$. As above, set $F = F_n = G^{\mathcal{R}}_{t_\mathrm{b}n}$. Since $\mathcal{R}$ is a take-it-or-leave-it rule, the further evolution may be described as follows. Let $\mathcal{L}_1$ and $\mathcal{L}_2$ be independent lists of independent (potential) edges each chosen uniformly at random from all $\binom{n}{2}$ possibilities. In step i of our process (step $t_\mathrm{b}n + i$ of the original), take for e_1 the ith element of $\mathcal{L}_1$. The rule now decides whether to add this edge to the current graph. If not, take for e_2 the *next* edge from $\mathcal{L}_2$, and add that. Thus, the jth time that the rule declines the first edge, we take the jth edge from $\mathcal{L}_2$.

Since the edges in $\mathcal{L}_2$ are uniformly random, the discussion before Theorem 15 shows that for any constant $\delta > 0$, whp the first δn edges from $\mathcal{L}_2$ will, when added to $F = G^{\mathcal{R}}_{t_\mathrm{b}n}$, be enough to form a giant component. Fix $\varepsilon > 0$, and define as above a graph H_n whose vertices are the components of F, with edges corresponding to the first εn edges from $\mathcal{L}_1$. As noted before, this graph H_n may be viewed as an instance of the model studied in [7], and there is some $\alpha > 0$ such that whp H_n has a component with at least αn vertices. Furthermore, by the stability result [7, Theorem 3.9], there is some $\delta > 0$ such that whp H_n has the property that deleting any δn edges still leaves a component with at least $\alpha n/2$ vertices of H_n. Hence, whp $\mathcal{L}_1$ has the property that if we add any subset of at least $(\varepsilon - \delta)n$ of the first εn edges to F, we will create a component of size at least $\alpha n/2$, and whp $\mathcal{L}_2$ has the property that adding its

first δn edges to F creates a component of size at least some constant times n. But when both properties hold, then whatever the rule does, $G^{\mathcal{R}}_{(t_b+\varepsilon)n}$ will have a giant component. □

For our final result, let us call an Achlioptas rule *large-biased* if there exists some constant B such that if both endvertices of e_1 are isolated vertices (components of size one) and both endvertices of e_2 are in components of size greater than B, then the rule will select e_2. Perhaps the most interesting examples of such rules are the *reverse product rule*, where we select the (a if there is a tie) edge maximizing the product of the sizes of the components containing its endvertices, or the *reverse sum rule*, defined similarly but with product replaced by sum. Perhaps surprisingly (given the difficulty of analyzing the usual product rule), we can prove Conjecture 2 for such rules.

Theorem 17. *Let $\mathcal{R}$ be a large-biased size rule. Then the conclusions of Conjecture 2 and Theorem 15 hold for $\mathcal{R}$.*

Proof. The proof is very similar to that of Theorem 15. Indeed, as usual we start from $F = G^{\mathcal{R}}_{t_b n}$. As before, let W be the set of vertices of F in components of size greater than B. Call a subsequent step *good* if e_1 joins two vertices in components of size one and e_2 joins two vertices in W. Since there are whp at least some constant times n isolated vertices in $G^{\mathcal{R}}_{(t_b+1)n}$, and (as before), W whp has size at least a constant times n, off an event of small probability the conditional probability (given the history) that the next step is good is always at least some positive constant. Furthermore, when a step is good, the added edge is uniformly random among all possible edges inside W. The remainder of the argument is as for Theorem 15; we omit the details. □

The results above all illustrate the idea that if we can find a reasonable number of uniformly random edges among the edges selected by our process, then the process will be 'well behaved' (will have $t_c = t_b$). This approach can be used to prove Conjecture 2 for other special classes of size rules, but it seems that additional ideas are needed for the general case.

3.2 Examples of delayed percolation

Having given several partial results supporting our belief in Conjecture 2, in this section we show that the conjecture cannot be extended to arbitrary ℓ-vertex rules. More concretely, we give examples of simple rules that can delay the appearance of linear size components for $\Omega(n)$ steps beyond the point where the susceptibility diverges. The rules we use behave like size rules almost all the time.

We start by introducing the *r-sum rule* $\mathcal{S}_r$, which is a $2r$-vertex size rule. Given vertices $(v_1, \ldots, v_\ell)$ and the corresponding list of component sizes $(c_1, \ldots, c_\ell)$, the r-sum rule adds the pair $v_{2j-1}v_{2j}$ with the (smallest, if there are ties) $j \in [r]$ that minimizes the sum $c_{2j-1} + c_{2j}$ of the component sizes. Recall the definition of $F^{\mathcal{R}}_i$ given in Section 2: informally it denotes the graph that we obtain

by starting with the initial graph $F_0^{\mathcal{R}} = F$ and then following i steps of an Achlioptas process using the rule $\mathcal{R}$ (to decide which edges to add in each step). Intuitively, the next lemma states that the r-sum rule does not substantially change (uniform) polynomial tails for $N_{\geqslant k}$ during some δn steps (here we use $\mathcal{S}_r$ for concreteness; other size rules exhibit similar behaviour).

Lemma 18. *Let F be a graph on n vertices. Suppose there are $x, C > 0$ and $K = K(n) \geqslant 1$ such that for all $1 \leqslant k \leqslant K$ we have*

$$N_{\geqslant k}(F) \leqslant Ck^{-x}n. \tag{53}$$

Given $r \geqslant 1 + 1/x$ there exists $\delta = \delta(x, C, r) > 0$ such if n is large enough then, with probability at least $1 - n^{-99}$, for all $1 \leqslant k \leqslant K' = \min\{K, n^{1/[2(1+x)]}\}$ we have

$$N_{\geqslant k}(F_{\delta n}^{\mathcal{S}_r}) \leqslant 2Ck^{-x}n. \tag{54}$$

Proof. Set $\delta = 2^{-[(2+x)r+3]}C^{-(r-1)}$. Let $\mathcal{E}_{i',k'}$ denote the event that for all $0 \leqslant i \leqslant i'$ and $1 \leqslant k \leqslant k'$ we have

$$N_{\geqslant k}(F_i^{\mathcal{S}_r}) \leqslant 2Ck^{-x}n. \tag{55}$$

Observe that it suffices to show that $\mathcal{E}_{\delta n,K'}$ fails with probability at most n^{-99}. For $k \leqslant K'$ let $X_{k,i}$ denote the indicator function of the event $N_{\geqslant k}(F_i^{\mathcal{S}_r}) \neq N_{\geqslant k}(F_{i-1}^{\mathcal{S}_r})$. Set $X_k = \sum_{1 \leqslant i \leqslant \delta n} X_{k,i}$ and $Y_k = \sum_{1 \leqslant i \leqslant \delta n} Y_{k,i}$, where

$$Y_{k,i} \;=\; \begin{cases} X_{k,i}, & \text{if } \mathcal{E}_{i-1,k-1} \text{ holds,} \\ 0, & \text{otherwise.} \end{cases}$$

Note that in each step a new component of size at least k is only created by $\mathcal{S}_r$ if in each pair $v_{2j-1}v_{2j}$ at least one vertex is in a component of size at least $\lceil k/2 \rceil$. So, whenever $\mathcal{E}_{i-1,k-1}$ holds, using (55) and $r \geqslant 1 + 1/x$, we see that the probability that $X_{k,i} = 1$ is at most

$$\left(\frac{2N_{\geqslant \lceil k/2 \rceil}(F_{i-1}^{\mathcal{S}_r})}{n}\right)^r \leqslant \left(\frac{4C}{(k/2)^x}\right)^r = \frac{\left(2^{2+x}C\right)^r}{k^{rx}} \leqslant \frac{\left(2^{2+x}C\right)^r}{k^{1+x}} = \xi_k.$$

Since $Y_{k,i} = 0$ whenever $\mathcal{E}_{i-1,k-1}$ fails, it follows that Y_k is stochastically dominated by a binomial random variable with δn trials and success probability ξ_k. Note that (53) implies $C \geqslant 1$. Now, using $k \leqslant K'$ we have $\delta n \xi_k \geqslant C/8 \cdot n^{1/2} \geqslant 600 \log n$ for $n \geqslant n_0$, so standard Chernoff bounds yield

$$\mathbb{P}(Y_k \geqslant 2\delta n \xi_k) \leqslant e^{-\delta n \xi_k/3} \leqslant n^{-200}. \tag{56}$$

Next we claim that $\mathcal{E}_{\delta n,k-1}$ and $Y_k < 2\delta n \xi_k$ together imply $\mathcal{E}_{\delta n,k}$, so that $\mathbb{P}(\neg\mathcal{E}_{\delta n,k}) \leqslant \mathbb{P}(\neg\mathcal{E}_{\delta n,k-1}) + \mathbb{P}(Y_k \geqslant 2\delta n \xi_k)$. Indeed, by monotonicity $\mathcal{E}_{\delta n,k-1}$ implies $\mathcal{E}_{i,k-1}$ for every $0 \leqslant i \leqslant \delta n$, so $Y_k = X_k$. Now, since $N_{\geqslant k}$ increases by at most $2k$ per step, by choice of δ it follows that

$$N_{\geqslant k}(F_{\delta n}^{\mathcal{S}_r}) - N_{\geqslant k}(F_0^{\mathcal{S}_r}) \leqslant 2kY_k \leqslant 4\delta \left(2^{2+x}C\right)^r k^{-x}n \leqslant Ck^{-x}n,$$

which together with $N_{\geqslant k}(F_0^{\mathcal{S}_r}) = N_{\geqslant k}(F) \leqslant Ck^{-x}n$ implies $N_{\geqslant k}(F_{\delta n}^{\mathcal{S}_r}) \leqslant 2Ck^{-x}n$, as claimed. Iterating the above argument for $k \leqslant K'$ and noting that $\mathcal{E}_{\delta n,1}$ always holds due to $C \geqslant 1$, using (56) we obtain

$$\mathbb{P}(\neg\mathcal{E}_{\delta n,K'}) \leqslant \sum_{2\leqslant k\leqslant K'} \mathbb{P}(Y_k \geqslant 2\delta n\xi_k) \leqslant n^{-99},$$

and the proof is complete. □

Let $\mathcal{D}_r$ denote the rule which always adds the pair v_1v_2 during the first $n/2$ steps (corresponding to an Erdős–Rényi evolution with $\ell = 2$); afterwards it 'switches' and uses the r-sum rule $\mathcal{S}_r$. The point is that many properties of the 'critical' Erdős–Rényi random graph $G_{n,n/2}$ are well known: there exist constants $C, \alpha > 0$ and a function $K = K(n)$ with $K \to \infty$ as $n \to \infty$ such that whp $S(G_{n,n/2}) \geqslant n^\alpha$ and $N_{\geqslant k}(G_{n,n/2}) \leqslant Ck^{-1/2}n$ for all $1 \leqslant k \leqslant K$. So, by conditioning on these properties and then using Lemma 18, we immediately deduce the main result of this section. Indeed, using the rule $\mathcal{D}_r$ for $r \geqslant 3$ we whp have diverging susceptibility after $n/2$ steps, but in δn subsequent steps whp no linear size components appear (in fact, in this case $t_{\mathrm{b}} = 1/2 < t_{\mathrm{c}}$ holds).

Corollary 19. *For every $r \geqslant 3$ there exists $\delta = \delta(r) > 0$ such that we have whp $S(G_{n/2}^{\mathcal{D}_r}) = \omega(1)$ and $L_1(G_{n/2+\delta n}^{\mathcal{D}_r}) = o(n)$.*

Alternatively, using essentially the same line of reasoning, we obtain a similar result by switching after the first step where the susceptibility is at least $L = L(n) = \omega(1)$, for L not too large. Furthermore, we can replace $\mathcal{S}_r$ by other suitable size rules. For example, the rule $\mathcal{M}_\ell$, which always connects two vertices with the two smallest component sizes c_j, satisfies an analogue of Lemma 18 for $\ell \geqslant 2 + 1/x$. So the rule $\mathcal{C}_\ell$, which switches from an Erdős–Rényi evolution (always adding v_1v_2) to $\mathcal{M}_\ell$ after $n/2$ steps, yields another example with $t_{\mathrm{b}} < t_{\mathrm{c}}$.

Corollary 20. *For every $\ell \geqslant 4$ there exists $\delta = \delta(\ell) > 0$ such that we have whp $S(G_{n/2}^{\mathcal{C}_\ell}) = \omega(1)$ and $L_1(G_{n/2+\delta n}^{\mathcal{C}_\ell}) = o(n)$.*

Note that the examples given in Corollary 19 and 20 always behave like size rules except that once between two steps they change the rule used (by only querying natural parameters such as the number of vertices and steps, or the susceptibility). So, one can argue that Conjecture 2 already fails for a rather restricted superset of size rules.

References

[1] D. Achlioptas, R.M. D'Souza, and J. Spencer. Explosive percolation in random networks. *Science* **323** (2009), 1453–1455.

[2] M. Aizenman and D.J. Barsky Sharpness of the phase transition in percolation models. *Comm. Math. Phys.* **108** (1987), 489–526.

[3] S. Bhamidi, A. Budhiraja, and X. Wang. Bohman–Frieze processes at criticality and emergence of the giant component. Preprint, 2011. arXiv:1106.1022.

[4] T. Bohman and A. Frieze. Avoiding a giant component. *Random Struct. Alg.* **19** (2001), 75–85.

[5] T. Bohman, A. Frieze, M. Krivelevich, P.-S. Loh, and B. Sudakov. Ramsey games with giants *Random Struct. Alg.* **38** (2011), 1–32.

[6] T. Bohman and D. Kravitz. Creating a giant component. *Combin. Probab. Comput.* **15** (2006), 489–511.

[7] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Struct. Alg.* **31** (2007), 3–122.

[8] B. Bollobás, *Random Graphs*, 2nd ed., Cambridge University Press (2001).

[9] B. Bollobás and O. Riordan. Random graphs and branching processes. In *Handbook of large-scale random networks*, Bolyai Soc. Math. Stud **18** (2009), pp. 15–115.

[10] S. Janson. Networking – Smoothly does it. *Science* **333** (2011), 298–299.

[11] S. Janson, D.E. Knuth, T. Łuczak and B. Pittel. The birth of the giant component. *Random Struct. Alg.* **4** (1993), 231–358.

[12] S. Janson and J. Spencer. Phase transitions for modified Erdős-Rényi processes. *Ark. Math.*, to appear. arXiv:1005.4494.

[13] M. Kang, W. Perkins, and J. Spencer. The Bohman–Frieze process near criticality. *Random Struct. Alg.*, to appear. arXiv:1106.0484.

[14] R.M. Karp. The transitive closure of a random digraph. *Random Struct. Alg.* **1** (1991), 73–93.

[15] C. McDiarmid. On the method of bounded differences. In *Surveys in Combinatorics* (Norwich, 1989), London Math. Soc. Lecture Note Ser., vol. 141, pp. 148–188. Cambridge Univ. Press, Cambridge, 1989.

[16] M.V. Men'shikov, Coincidence of critical points in percolation problems. *Dokl. Akad. Nauk SSSR* **288** (1986), 1308–1311.

[17] O. Riordan and L. Warnke. Achlioptas processes can be nonconvergent. Preprint, 2011. arXiv:1111.6179.

[18] O. Riordan and L. Warnke. Achlioptas process phase transitions are continuous. *Ann. Appl. Probab.*, to appear. arXiv:1102.5306.

[19] O. Riordan and L. Warnke. Convergence of Achlioptas processes via differential equations with unique solutions. Preprint, 2011. arXiv:1111.6179.

[20] O. Riordan and L. Warnke. Explosive percolation is continuous. *Science* **333** (2011), 322–324.

[21] J. Spencer and N.C. Wormald. Birth control for giants. *Combinatorica* **27** (2007), 587–628.

[22] N.C. Wormald. Differential equations for random processes and random graphs. *Ann. Appl. Probab.* **5** (1995), 1217–1235.

[23] N.C. Wormald. The differential equation method for random graph processes and greedy algorithms. In *Lectures on approximation and randomized algorithms*, pages 73–155. PWN, Warsaw, 1999.

A Appendix

In this appendix we show that, as long as an Achlioptas process contains only small components, it will have a very simple cycle structure: most components will be trees, some will be unicyclic, and there will (whp) be no 'complex' components, i.e., ones containing more than one cycle. In fact, we prove the result for the more general class of ℓ-vertex rules. However, here we need an additional assumption: in each round the set of added edges is a forest. Allowing slightly greater generality, we call a rule *acyclic* if the edges between $v_1, \dots, v_\ell$ added in a single step correspond to a forest on $1, \dots, \ell$. This in particular holds if in each step at most two edges are added. Note that such an assumption is in fact necessary for $\ell \geqslant 3$, since always connecting all vertices in each step quickly creates many cycles and complex components.

Lemma 21. *Let $\ell \geqslant 2$ and let $\mathcal{R}$ be an acyclic ℓ-vertex rule. Given $0 < \delta < 1/4$ and $U = U(n)$, suppose that for $n \geqslant n_0(\ell, \delta)$ we have $1 \leqslant U \leqslant n^{1/4-\delta}$. For $n \geqslant n_0(\ell, \delta)$, with probability at least $1 - n^{-\delta/2}$ the following holds for every $0 \leqslant i \leqslant n^{1+\delta}$: in $G_i^{\mathcal{R}}$ there are no components of size at most U which contain at least two cycles, and the number of vertices in components of size at most U with exactly one cycle is at most $U^2 n^{2\delta}$.*

Proof. Set $m = n^{1+\delta}$. Let $\mathcal{B}_{1,i}$ denote the event that in step i one of the following happens: (a) at least three of the ℓ randomly chosen vertices are contained in the same component of size at most U, or (b) there are randomly chosen vertices w_1, w_2, w_3, w_4 and two components C_1, C_2 of size at most U such that $w_1, w_2 \in C_1$ and $w_3, w_4 \in C_2$. It is easy to see that $\mathcal{B}_{1,i}$ holds with probability at most $\ell^3 U^2/n^2 + \ell^4 U^2/n^2$. So, denoting by $\mathcal{B}_1$ the event that $\mathcal{B}_{1,i}$ holds for some $i \leqslant m$, we see that

$$\mathbb{P}(\neg \mathcal{B}_1) \leqslant m \cdot 2\ell^4 U^2/n^2 \leqslant 2\ell^4 n^{-1/2-\delta}.$$

Let $S_{i,U}$ denote the set of vertices of $G_i^{\mathcal{R}}$ which are in components of size at most U containing exactly one cycle, and let $\mathcal{B}_2$ be the event that $S_{i,U}$ contains at least $2\ell^3 U m/n$ components for some $i \leqslant m$. Then $\neg\mathcal{B}_2$ implies

$|S_{i,U}| \leqslant 2\ell^3U^2m/n = R$ for every $i \leqslant m$, where $R \leqslant U^2n^{2\delta}$ for $n \geqslant n_0(\ell,\delta)$. Since in each step the number of components in $S_{i,U}$ changes by at most ℓ, when $\mathcal{B}_2$ holds there are at least $2\ell^2Um/n$ steps in which the number of components in $S_{i,U}$ increases. This can only happen if at least two randomly chosen vertices are in the same component of size at most U. Since in each step this has probability at most ℓ^2U/n, the expected number of such steps is bounded by $\lambda = \ell^2Um/n$. Using standard Chernoff bounds (and stochastic domination) it follows that

$$\mathbb{P}(\mathcal{B}_2) \leqslant e^{-\lambda/3} \leqslant e^{-n^\delta}.$$

Let $\mathcal{B}_3$ denote the event that in some $G_i^{\mathcal{R}}$ with $i \leqslant m$ there exists a component of size at most U which contains at least two cycles. In each step where $\mathcal{B}_{1,i}$ fails, note that a new complex component of size at most U can only be created if (a) at least two randomly chosen vertices are in $S_{i-1,U}$ or (b) one randomly chosen vertex lies in $S_{i-1,U}$, and two randomly chosen vertices are in the same tree component of size at most U. So, by considering the probability that this happens for some $i \leqslant m$ (assuming $|S_{i-1,U}| \leqslant R$), we see that

$$\mathbb{P}(\neg\mathcal{B}_1 \cap \neg\mathcal{B}_2 \cap \mathcal{B}_3) \leqslant m \cdot (\ell^2R^2/n^2 + \ell^3RU/n^2) \leqslant 5\ell^8U^4m^3/n^4 \leqslant 5\ell^8n^{-\delta},$$

completing the proof for $n \geqslant n_0(\delta,\ell)$. □

Theorem 1 tells us that for size rules, for any fixed $t < t_{\mathrm{b}}$, the largest component of $G_{tn}^{\mathcal{R}}$ whp has size at most $O(\log n)$. Taking $U = (\log n)^2$, say, we see that if $\mathcal{R}$ is acyclic, then whp there are no complex components, and at most $n^{o(1)}$ vertices on cyclic components – in other words, almost all components are trees. (We have not tried to optimize the bound here – the method actually gives some power of $\log n$.)