Combinatorics, Probability and Computing (2016) 25, 154–171.
doi:10.1017/S0963548315000218

Convergence of Achlioptas Processes via Differential Equations with Unique Solutions

OLIVER RIORDAN[1] and LUTZ WARNKE[2]

[1] Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, UK
(e-mail: riordan@maths.ox.ac.uk)

[2] Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, UK
(e-mail: L.Warnke@dpmms.cam.ac.uk)

Received 27 June 2013; revised 7 April 2015; first published online 14 October 2015

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. The evolution of the rescaled size of the largest component in such variations of the Erdős–Rényi random graph process has recently received considerable attention, in particular for Bollobás's 'product rule'. In this paper we establish the following result for rules such as the product rule: the limit of the rescaled size of the 'giant' component exists and is continuous provided that a certain system of differential equations has a unique solution. In fact, our result applies to a very large class of Achlioptas-like processes.

Our proof relies on a general idea which relates the evolution of stochastic processes to an associated system of differential equations. Provided that the latter has a unique solution, our approach shows that certain discrete quantities converge (after appropriate rescaling) to this solution.

2010 *Mathematics subject classification*: Primary 05C80
Secondary 34F05, 60C05

1. Introduction

More than 50 years ago Erdős and Rényi initiated the systematic study of the *random graph process*, which is the random sequence of graphs obtained by starting with an empty graph on n vertices and then in each step adding a new random edge. Already in their seminal 1960 paper [13] they investigated the size of the largest component in great detail. Suppressing, as usual, the dependence on n, let $L_1(m)$ denote the size of the largest component after m steps. Their results imply, for example, that there is a continuous function $\rho = \rho^{\mathrm{ER}} : [0,\infty) \to [0,1)$ such that for any fixed $t \geqslant 0$ we have $L_1(\lfloor tn \rfloor)/n \xrightarrow{\mathrm{p}} \rho(t)$ as $n \to \infty$, where $\xrightarrow{\mathrm{p}}$ denotes convergence in probability. Nowadays the evolution of the

component structure, in particular the size of the largest component, is one of the most studied properties in the theory of random graphs; see, *e.g.*, the many references in [8, 9].

In order to create processes with potentially different behaviour, in 2000 Dimitris Achlioptas suggested certain variants of the classical random graph process (inspired by the 'power of random choices' paradigm [3]). These also start with an empty graph $G(0)$ on n vertices. At each later step $m \geqslant 1$, 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 $G(m-1)$). One of these edges is selected according to a rule $\mathcal{R}$ and added to the graph, so $G(m) = G(m-1) \cup \{e\}$ for $e = e_1$ or e_2. Processes of this type are now known as *Achlioptas processes*; always adding $e = e_1$ gives the Erdős–Rényi random graph process.

During the past decade the evolution of the largest component in Achlioptas processes has received considerable attention. In one line of research the location (and existence) of the phase transition has been investigated; see, *e.g.*, [6, 7, 28]. This is motivated by Achlioptas' original question, namely, whether the 'freedom of choice' in each step can be used to substantially delay or accelerate the appearance of the linear-size 'giant' component. The above results answer this affirmatively by considering so-called 'bounded-size' rules, whose decisions only depend 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 the same way.

A more recent direction of research concerns finer details of the phase transition in Achlioptas processes (see, *e.g.*, [4, 5, 19, 20]), investigating similarities and differences to the well-understood classical random graph process. In this context in particular the *product rule* (suggested early on by Bollobás as the best rule to delay the phase transition) has received considerable attention: given two potential edges, it picks the one minimizing the product of the sizes of the components of its endvertices. Based on extensive simulations, Achlioptas, D'Souza and Spencer conjectured in *Science* [1] that the rescaled size of the largest component undergoes a discontinuous phase transition for the product rule, that is, there exists a constant $\delta > 0$ so that $L_1(m)/n$ 'jumps' from $o(1)$ to at least δ in $o(n)$ steps. Called *explosive percolation*, this phenomenon has been of great interest to physicists; see, *e.g.*, [10, 11, 12, 14, 22, 31]. However, recently it has been rigorously shown in [23, 25] that the simulations were misleading, and that the phase transition is actually continuous for *all* Achlioptas processes.

The discussion above, and much of the physics literature, takes an important question for granted: Does the scaling limit even exist? More precisely, as in [23, 25] we say that a rule $\mathcal{R}$ is *globally convergent* if there exists an increasing function $\rho = \rho^{\mathcal{R}} : [0, \infty) \to [0, 1]$ such that for any t at which ρ is continuous we have

$$L_1(\lfloor tn \rfloor)/n \xrightarrow{\mathrm{p}} \rho(t) \tag{1.1}$$

as $n \to \infty$. The function $\rho = \rho^{\mathcal{R}}$ is called the *scaling limit* of (the size of the giant component of) $\mathcal{R}$. Writing $N_k(m)$ for the number of vertices of $G(m)$ in components with k vertices, we call a rule $\mathcal{R}$ *locally convergent* if there exist functions $\rho_k = \rho_k^{\mathcal{R}} : [0, \infty) \to [0, 1]$ such that, for each fixed $k \geqslant 1$ and $t \geqslant 0$, we have $N_k(\lfloor tn \rfloor)/n \xrightarrow{\mathrm{p}} \rho_k(t)$ as $n \to \infty$ (such

https://doi.org/10.1017/S0963548315000218 Published online by Cambridge University Press

functions, if they exist, are necessarily continuous). Spencer and Wormald [28] showed that all bounded-size rules are locally convergent, and conjectured that they are globally convergent. In [25] it was shown that global convergence (and continuity of $\rho(t)$) follows from local convergence, settling this conjecture.

For general size rules, the problem of establishing convergence (local and hence global) is still open, although there are partial results: in [26] convergence was established up to the critical time t_b at which the 'susceptibility' (the average size of the component containing a random vertex) diverges. According to Achlioptas, D'Souza and Spencer [1], complex rules such as the product rule seem to be 'beyond the reach of current mathematical techniques', so it is not too surprising that for these no convergence results are known beyond t_b, that is, in the later evolution. Svante Janson [17] also remarks that most likely new methods are needed to understand the detailed behaviour of such rules.

1.1. Main result

In this paper we address the convergence question for Achlioptas processes: we show that rules such as the product rule are globally convergent (for all $t \in [0,\infty)$) provided that a certain associated system of differential equations (defined in Section 2.2) has a unique solution. Our result applies to a very large class of Achlioptas processes, including all bounded-size rules and the product rule. For the definitions of ℓ-vertex rule, merging and well-behaved, see Section 2.

Theorem 1.1. *Let $\ell \geqslant 2$ and let $\mathcal{R}$ be a merging ℓ-vertex rule that is well-behaved. Suppose the associated system of differential equations given by* (2.6)–(2.9) *has a unique solution $(\hat{\rho}_k(t))_{k\geqslant 1}$. Then $\mathcal{R}$ is locally and globally convergent. In particular, for each fixed $k \geqslant 1$ and $t \geqslant 0$, we have*

$$N_k(\lfloor tn \rfloor)/n \xrightarrow{p} \hat{\rho}_k(t) \tag{1.2}$$

as $n \to \infty$. The scaling limit $\rho = \rho^{\mathcal{R}}$ is continuous and satisfies $\rho(t) = 1 - \sum_{k\geqslant 1} \hat{\rho}_k(t)$.

Remark 1. We shall show that under the conditions of Theorem 1.1, the system of differential equations has at least one solution (see Lemma 3.1). The key assumption of the theorem is that it does not have more than one solution.

Remark 2. If we only assume uniqueness on an interval $I = [0, t^*]$ or $I = [0, t^*)$, then the conclusions of Theorem 1.1 hold for any $t \in I$ (see Section 4).

The merging assumption in Theorem 1.1 seems to be necessary (even for size rules): in [24] examples of 'natural' non-merging rules are given, where simulations strongly suggest that convergence fails. (This phenomenon is numerically much more robust than the continuity question addressed in [1].) All rules for which convergence has previously been established are merging and well-behaved, including the classical Erdős–Rényi case [13], all bounded-size rules [28] (such as the Bohman–Frieze rule [19]) as well as the dCDGM rule [11] and the adjacent edge rule [12]. In fact, for all such rules there is a $K \geqslant 1$ such that each ρ_k' in (2.7) can be written as a function of $\rho_1, \dots, \rho_{\max\{k,K\}}$ (such

https://doi.org/10.1017/S0963548315000218 Published online by Cambridge University Press

rules are called *nice* in [25]). In this case the form of the differential equations (2.6)–(2.9) implies by standard results (see, *e.g.*, [16]) that their solution is unique. So Theorem 1.1 generalizes these previous convergence results.

Perhaps the main contribution of this paper is a new approach for proving convergence. Previous results in this area apply Wormald's 'differential equation method' [29, 30], which is now widely used in probabilistic combinatorics. This shows that under certain conditions, suitable sequences of random variables converge to the solution of a system of differential equations. The key point is that these conditions imply that the differential equations have a unique solution, but are not implied by this. By establishing a more direct connection between the random process and the differential equations, we need only assume that the system of differential equations has a unique solution. Thus, our method is potentially applicable to a much larger class of Achlioptas processes. The general proof idea outlined in Section 3 might also be useful to establish convergence in other stochastic processes.

In fact, our approach establishes more than convergence: for each 'typical' outcome, it shows that the evolution of suitable random variables follows *some* solution of the associated system of differential equations. Hence our method allows us to transfer properties common to all solutions back to the random process. We demonstrate the usefulness of this feature in Section 4, where for Achlioptas processes we narrow the bounds on the interval in which the giant component emerges.

Theorem 1.1 may be seen as a first step towards resolving the convergence question in Achlioptas processes. In particular, further investigation of the system of differential equations (2.6)–(2.9) associated with the product rule (and other complicated rules) seems to be needed: Does it have a unique solution? When the equations do have a unique solution many questions remain, for example: Which conditions are needed to establish asymptotic normality as in [27]?

In the next section we define the processes under consideration and state the system of differential equations associated with them. In Section 3 we first outline a general idea for proving convergence in stochastic processes, and then use this approach to establish our main result. Finally, in Section 4 we investigate the emergence of the giant component via properties of the differential equations.

2. Preliminaries and notation

Our core argument will involve considering sequences of points ω_n in different probability spaces. For this reason we indicate the dependence on n explicitly in the notation. We now recall the relevant definitions from [25]. Fix $\ell \geqslant 2$. An *ℓ-vertex rule* $\mathcal{R}$ yields for each n a random sequence $(G_{n,m})_{m\geqslant 0}$ of graphs with vertex set $[n] = \{1, \dots, n\}$, where $G_{n,0}$ is the empty graph. For each $m \geqslant 0$ we draw ℓ vertices $\underline{v}_{n,m+1} = (v_1, \dots, v_\ell)$ from $[n]$ independently and uniformly at random, and then obtain $G_{n,m+1}$ by adding a (possibly empty) set of edges $E_{n,m+1}$ to $G_{n,m}$, where $\mathcal{R}$ selects $E_{n,m+1}$ as a subset of all pairs between vertices in $\underline{v}_{n,m+1}$. To avoid 'trivial' rules (which never add edges) we require that $E_{n,m+1} \neq \emptyset$ if all ℓ vertices in $\underline{v}_{n,m+1}$ are in distinct components of $G_{n,m}$ (it would also suffice that the conditional probability of this event is bounded away from 0). Formally, we assume the existence of a sample space Ω_n and a filtration $\mathcal{F}_{n,0} \subseteq \mathcal{F}_{n,1} \subseteq \cdots$ such that $\underline{v}_{n,m+1}$ is

https://doi.org/10.1017/S0963548315000218 Published online by Cambridge University Press

$\mathcal{F}_{n,m+1}$-measurable and independent of $\mathcal{F}_{n,m}$, and require $E_{n,m+1}$ (and hence $G_{n,m+1}$) to be $\mathcal{F}_{n,m+1}$-measurable. For later use we let $\underline{c}_{n,m+1} = (c_1, \dots, c_\ell)$ denote the sizes of the components containing the chosen vertices $\underline{v}_{n,m+1} = (v_1, \dots, v_\ell)$ in $G_{n,m}$.

For the purposes of this paper the above definitions are robust with respect to small changes, since our arguments have $o(1)$ elbow room in each step of the process. So we may weaken the conditions on $\underline{v}_{n,m+1}$: it suffices if, for $m = O(n)$, say, the conditional distribution of $\underline{v}_{n,m+1}$ given $\mathcal{F}_{n,m}$ is close to (at total variation distance $\alpha_n = o(1)$ from) the one defined above. This includes variations such as picking an ℓ-tuple of distinct vertices, or picking (the ends of) $\ell/2$ randomly selected (distinct) edges not already present in $G_{n,m}$. Hence, as in [25], we may treat the original examples of Achlioptas as 4-vertex rules where $\mathcal{R}$ always selects one of the pairs $\{v_1, v_2\}, \{v_3, v_4\}$; below we call such $\mathcal{R}$ *Achlioptas rules*.

We say that an ℓ-vertex rule is *merging* if, whenever C, C' are distinct components with $|C|, |C'| \geqslant \varepsilon n$, then in the next step we have probability at least ε^ℓ of joining C to C' (this can be slightly weakened; see [25]). In particular all Achlioptas rules are merging, since with probability at least ε^4 both potential pairs join C to C'.

Turning to notation, we write $N_{n,k,m}$ for the number of vertices of $G_{n,m}$ in components of size k, and let $N_{n,\leqslant k,m} = \sum_{1 \leqslant j \leqslant k} N_{n,j,m}$. We define $N_{n,\geqslant k,m}$ and $L_{1,n,m}$ in an analogous way. To avoid clutter, in Sections 1 and 4 we sometimes suppress n (which is clear from the context), writing, for example,

$$G(m) = G_{n,m}, \tag{2.1}$$

$$N_k(m) = N_{k,n,m}, \tag{2.2}$$

$$L_1(m) = L_{1,n,m}. \tag{2.3}$$

Finally, throughout we write $x = a \pm b$ as shorthand for $x \in [a - b, a + b]$.

2.1. Well-behaved rules

We say that an ℓ-vertex rule $\mathcal{R}$ is *well-behaved* (*at infinity*) if there are functions

$$d_k : (\mathbb{N} \cup \{\infty\})^\ell \to \mathbb{R} \quad \text{and} \quad g : \mathbb{N} \to \mathbb{N} \quad \text{with } g(s) \geqslant s$$

such that the following conditions hold.

(i) Whenever all vertices v_j are in different components we have

$$\mathbb{E}(N_{n,k,m+1} - N_{n,k,m} \mid \mathcal{F}_{n,m}, \underline{v}_{n,m+1}) = d_k(c_1, \dots, c_\ell), \tag{2.4}$$

where $\underline{c}_{n,m+1} = (c_1, \dots, c_\ell)$ lists the sizes of the components containing the selected vertices.

(ii) Suppose there are $I \subseteq [\ell]$ and $S \geqslant k$ such that all v_j with $j \in I$ are in the same component of size $c_j > g(S)$, whereas all other vertices are in different components with sizes $c_j \leqslant S$. Whenever this holds we have

$$\mathbb{E}(N_{n,k,m+1} - N_{n,k,m} \mid \mathcal{F}_{n,m}, \underline{v}_{n,m+1}) = d_k(\tilde{c}_1, \dots, \tilde{c}_\ell), \tag{2.5}$$

where $\tilde{c}_j = \infty$ for $j \in I$ and $\tilde{c}_j = c_j$ otherwise.

https://doi.org/10.1017/S0963548315000218 Published online by Cambridge University Press

In fact, taking $I = \emptyset$ in (2.5) gives (2.4), but we note (2.4) separately for clarity. As we shall discuss below, these conditions are very mild and hold for essentially all Achlioptas processes previously studied, including 'unbounded rules' such as the product and sum rules (the latter is a variant of the product rule which minimizes the sum as opposed to the product). All rules which have been considered so far are *size rules*, which only use $\underline{c}_{n,m+1}$ to decide which edge(s) to add. For these the change of $N_{n,k,m}$ in (2.4) is deterministic given $\underline{c}_{n,m+1}$, but considering the conditional expected change is slightly more general (we can also allow for small deviations in (2.4) and (2.5), but leave this to the interested reader). Intuitively, the second condition ensures that whenever one component is significantly larger than all others, then we can decide which relevant pairs are joined *without* knowing its exact size (this fails, for example, if the change depends on the parity of $\lfloor \log(\max_{j\in[\ell]} c_j) \rfloor$). This mild assumption holds for a large class of rules; for example, $g(s) = s^2$, $g(s) = 2s$ and $g(s) = \max\{s, K\}$ suffice for the product rule, the sum rule, and all nice rules as defined in [25] (which includes bounded-size rules, where $K = B$). Note that since $N_{n,k,m}$ always changes by at most ℓk per step, we have $|d_k(\cdot)| \leqslant \ell k$.

Remark 3. One might also consider rules $\mathcal{R}$ that depend on n in a limited way, such as 'truncated' versions of size rules with a cutoff $B = B(n)$ that grows with n (these follow the original rule when all vertices v_j are in components of size at most B). Our results extend to this case with modified (weaker) conditions (2.4) and (2.5). Given $B(n)$ satisfying $B(n) \to \infty$ and $B(n)/n \to 0$, for (2.5) we modify the size condition to $c_j > \max\{g(S), B(n)\}$ for all $j \in I$ and $c_j \leqslant \min\{S, B(n)\}$ for all $j \in [\ell] \setminus I$; the special case (2.4) is modified similarly. (Note that the functions d_k cannot depend on n.)

2.2. An associated system of differential equations

Suppose that $\mathcal{R}$ is a well-behaved ℓ-vertex rule. In the following equations, each $\rho_k(t)$ is a (differentiable) function on $[0, \infty)$ satisfying

$$0 \leqslant \rho_k(t) \leqslant 1 \quad \text{and} \quad 0 \leqslant \sum_{k\geqslant 1} \rho_k(t) \leqslant 1. \tag{2.6}$$

The system of differential equations associated with $\mathcal{R}$ is given by

$$\rho_k'(t) = \sum_{c_1,\ldots,c_\ell \in \mathbb{N}\cup\{\infty\}} d_k(c_1, \ldots, c_\ell) \prod_{j\in[\ell]} \rho_{c_j}(t) \tag{2.7}$$

for all $k \geqslant 1$, where

$$\rho(t) = \rho_\infty(t) = 1 - \sum_{k\geqslant 1} \rho_k(t), \tag{2.8}$$

together with the initial conditions

$$\rho_k(0) = \begin{cases} 1 & \text{if } k = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{2.9}$$

For $t = 0$, the derivative in (2.7) is taken to be the right-derivative. Note that for all $t \geqslant 0$ we have $|\rho_k'(t)| \leqslant \max_{\underline{c}} |d_k(\underline{c})| \leqslant \ell k$.

https://doi.org/10.1017/S0963548315000218 Published online by Cambridge University Press

As a basic example, consider the Erdős–Rényi random graph process, for which we have $d_k(c_1,c_2) \in \{-2k,-k,0,k\}$. It is not difficult to see that in this case (2.7) simplifies to

$$\rho_k'(t) = -2k\rho_k(t) + k \sum_{c_1+c_2=k} \rho_{c_1}(t)\rho_{c_2}(t), \tag{2.10}$$

which is a special case of *Smoluchowski's coagulation equations* in a form where sol-gel interaction is considered; see, *e.g.*, [2, 21] and the references therein. Here uniqueness follows easily from standard results (see, *e.g.*, [16]), since ρ_k' depends only on $\rho_1,\dots,\rho_k$.

3. Proof of the main result

We start by outlining a rather general idea for proving convergence to the unique solution of a system of differential equations, which we shall later use to establish Theorem 1.1. We consider a discrete stochastic process with sample space Ω_n and filtration $\mathcal{F}_{n,0} \subseteq \mathcal{F}_{n,1} \subseteq \cdots$. For each (discrete) step m we introduce (continuous) time $t = m/s_n$, where the scaling satisfies $s_n \to \infty$ as $n \to \infty$. Suppose our objective is to find a collection of random variables $X_{n,k,m}$ and (continuous) functions $x_k(t)$ together with (deterministic) scaling parameters $S_{n,k}$, such that for each fixed $k \geqslant 1$ and $t \geqslant 0$, we have

$$X_{n,k,ts_n}/S_{n,k} \xrightarrow{\mathrm{p}} x_k(t)$$

as $n \to \infty$, where we ignore the rounding to integers. The two main steps of our approach are as follows.

(1) Defining the one-step change as $\Delta X_{n,k,m+1} = X_{n,k,m+1} - X_{n,k,m}$, we use martingale techniques (the Azuma–Hoeffding inequality along with an absolute bound on $|\Delta X_{n,k,m+1}|$) to show that, with probability tending to 1 as $n \to \infty$, the following holds: for each fixed k and all $m_1, m_2 \geqslant 0$ with $m_2 - m_1 = O(s_n)$ we have

$$X_{n,k,m_2} - X_{n,k,m_1} = \Big(\sum_{m_1 \leqslant m < m_2} \mathbb{E}(\Delta X_{n,k,m+1} \mid \mathcal{F}_{n,m}) \Big) + o(S_{n,k}). \tag{3.1}$$

(2) Suppose we are given a sequence of sample points $\omega_n \in \Omega_n$, defined for some infinite set of $n \in \mathbb{N}$, for which (3.1) and some additional typical properties (technical conditions) hold. Proceeding as in the proof of Helly's selection theorem (see, *e.g.*, Theorem 5.8.1 in [15]), we pick a subsequence $(\omega_{\tilde{n}})$ such that for each $t \geqslant 0$ and $k \geqslant 1$, for some limiting value $x_k(t)$ we have

$$X_{\tilde{n},k,ts_{\tilde{n}}}(\omega_{\tilde{n}})/S_{\tilde{n},k} \to x_k(t) \tag{3.2}$$

as $\tilde{n} \to \infty$ (here we exploit that each $X_{\tilde{n},k,ts_{\tilde{n}}}(\omega_{\tilde{n}})/S_{\tilde{n},k}$ satisfies a Lipschitz condition as a function of t). For this subsequence, we show that for all $t \geqslant 0$, $k \geqslant 1$ and $\varepsilon > 0$ there exists $\delta > 0$ such that for $\tilde{n}$ large enough the following holds: for each $m \geqslant 0$ satisfying $|m - ts_{\tilde{n}}| \leqslant \delta s_{\tilde{n}}$ we have

$$\mathbb{E}(\Delta X_{\tilde{n},k,m+1} \mid \mathcal{F}_{\tilde{n},m})(\omega_{\tilde{n}}) = (f_k(t) \pm \varepsilon/3) S_{\tilde{n},k}/s_{\tilde{n}}, \tag{3.3}$$

where $f_k(t) = f_k(t, x_1(t), x_2(t), \dots)$ is a function of the scaling limits of the selected subsequence. To establish (3.3) we combine coupling arguments with 'typical' properties of the underlying stochastic process.

https://doi.org/10.1017/S0963548315000218 Published online by Cambridge University Press

Now, using (3.1)–(3.3) it is straightforward to show that for all $t \geqslant 0$, $k \geqslant 1$ and $\varepsilon > 0$ there exists $\delta > 0$ such that for all h with $0 < |h| \leqslant \delta$ and $t + h \geqslant 0$ we have

$$\left| \frac{x_k(t+h) - x_k(t)}{h} - f_k(t) \right| < \varepsilon,$$

that is, the $x_k(t)$ satisfy the differential equation

$$x_k'(t) = f_k(t, x_1(t), x_2(t), \ldots).$$

If the associated system of differential equations has a unique solution, then this implies that the limiting functions $x_k(t)$ in (3.2) do *not* depend on the selected subsequence, which establishes the desired convergence (by the well-known subsubsequence principle; see, *e.g.*, Section 1.2 in [18]). Finally, let us remark that by comparison with the underlying process we can (typically) derive additional properties of the x_k; it suffices to establish uniqueness of the solution to the system of differential equations augmented by these extra restrictions.

In the rest of this section we use the above approach to establish Theorem 1.1. Aiming at $N_{n,k,tn}/n \xrightarrow{\mathrm{p}} \rho_k(t)$, we closely follow steps (1) and (2) in Sections 3.1 and 3.2, respectively, with $X_{n,k,m} = N_{n,k,m}$, $x_k(t) = \rho_k(t)$, $S_{n,k} = n$ and $s_n = n$.

3.1. Proof of Theorem 1.1

Our proof of Theorem 1.1 relies on a technical lemma which requires some preparation. To this end we shall introduce several events needed to implement step (1) from our proof outline, which capture various typical properties of $G_{n,m}$. For concreteness, set

$$\eta(n) = (\log\log\log n)^{-1}, \tag{3.4}$$

say (the particular form does not matter, we shall later only use that $\eta(n) \to 0$ as $n \to \infty$). Let $\mathcal{U}_n$ denote the event that at every step m there is at most one component of size at least $\eta(n)n$. Since $\mathcal{R}$ is merging, by the discussion following Theorem 2 in [25] we know that $\mathbb{P}(\mathcal{U}_n) \to 1$ as $n \to \infty$. By Theorem 2 of [25], for any constant $\gamma > 0$ there is a constant $K(\gamma)$ such that

$$\mathbb{P}\big(\forall m : N_{n,\geqslant K(\gamma),m} < L_{1,n,m} + \gamma n\big) \to 1$$

as $n \to \infty$, where $L_{1,n,m}$ is the number of vertices in the largest component of $G_{n,m}$. By a standard argument (considering, say, $\gamma = 2^{-i}$ for each $i \in \mathbb{N}$), we may allow γ to tend to zero at some rate. More precisely, there exist functions $K(\gamma)$ and $\xi(n)$ with $\xi(n) \to 0$ as $n \to \infty$ such that, defining $\mathcal{K}_n$ as the event that for all $m \geqslant 0$ we have

$$\forall \gamma \geqslant \xi(n) : \; N_{n,\geqslant K(\gamma),m} < L_{1,n,m} + \gamma n, \tag{3.5}$$

we have $\mathbb{P}(\mathcal{K}_n) \to 1$ as $n \to \infty$ (note that $K(\gamma)$ does *not* depend on n).

Fix $0 < \lambda < 1/4$, say $\lambda = 1/8$ for concreteness. For each $m \geqslant 0$ set

$$\Delta N_{n,k,m+1} = N_{n,k,m+1} - N_{n,k,m}$$

and

$$Y_{n,k,m+1} = \Delta N_{n,k,m+1} - \mathbb{E}(\Delta N_{n,k,m+1} \mid \mathcal{F}_{n,m}).$$

https://doi.org/10.1017/S0963548315000218 Published online by Cambridge University Press

Set

$$Z_{n,k,j} = \sum_{0 \leqslant m < j} Y_{n,k,m+1}.$$

Let $\mathcal{D}_n$ denote the event that for all $1 \leqslant k \leqslant n^{\lambda}$ and $1 \leqslant m_1 \leqslant m_2 \leqslant n^2$ with $m_2 - m_1 \leqslant n^{1+\lambda}$ we have $|Z_{n,k,m_2} - Z_{n,k,m_1}| < n^{1/2+2\lambda}$. Note that by rearranging terms, for all such k, m_1, m_2 the event $\mathcal{D}_n$ implies

$$N_{n,k,m_2} - N_{n,k,m_1} = \Big(\sum_{m_1 \leqslant m < m_2} \mathbb{E}(\Delta N_{n,k,m+1} \mid \mathcal{F}_{n,m}) \Big) \pm n^{1/2+2\lambda}, \tag{3.6}$$

which is the rigorous analogue of (3.1) from step (1) of our proof outline. Since the number of vertices in components of size k changes by at most ℓk per step, we have $|\Delta N_{n,k,m+1}| \leqslant \ell k$ and thus $|Z_{n,k,m+1} - Z_{n,k,m}| = |Y_{n,k,m+1}| \leqslant 2\ell k$. Furthermore $\mathbb{E}(Y_{n,k,m+1} \mid \mathcal{F}_{n,m}) = 0$, so $(Z_{n,k,j})_{j \geqslant m_1}$ is a martingale. Thus, for fixed k, m_1, m_2 satisfying the conditions above, by the Azuma–Hoeffding inequality we have, say,

$$\mathbb{P}(|Z_{n,k,m_2} - Z_{n,k,m_1}| \geqslant n^{1/2+2\lambda}) \leqslant 2e^{-n^{3\lambda}/(8\ell^2 k^2)} \leqslant 2e^{-n^{\lambda}/(8\ell^2)} \leqslant n^{-9}$$

for n large enough. Taking a union bound (to account for all choices of k, m_1, m_2) yields $\mathbb{P}(\mathcal{D}_n) \to 1$ as $n \to \infty$, with room to spare.

Finally, define the 'good' event $\mathcal{G}_n = \mathcal{D}_n \cap \mathcal{K}_n \cap \mathcal{U}_n$; we have shown that $\mathbb{P}(\mathcal{G}_n) \to 1$ as $n \to \infty$. We are now ready to state the main technical lemma. As usual, we ignore the irrelevant rounding to integers.

Lemma 3.1. *Let $\ell \geqslant 2$ and let $\mathcal{R}$ be a merging ℓ-vertex rule that is well-behaved. Let (ω_n) with $\omega_n \in \mathcal{G}_n \subseteq \Omega_n$ be defined for an infinite set of $n \in \mathbb{N}$. Then there exists a subsequence $(\omega_{\tilde{n}})$ of (ω_n) such that for each $t \geqslant 0$ and $k \geqslant 1$ we have $N_{\tilde{n},k,t\tilde{n}}(\omega_{\tilde{n}})/\tilde{n} \to \rho_k(t)$, where the $(\rho_k(t))_{k \geqslant 1}$ are functions on $\mathbb{R}^+$ satisfying the system of differential equations (2.6)–(2.9) associated to $\mathcal{R}$.*

Note that Lemma 3.1 implies that the system of differential equations (2.6)–(2.9) has at least one solution. By comparison with the underlying process we can establish additional properties of the $\rho_k(t)$, for example that $\rho_{\leqslant k}(t) = \sum_{1 \leqslant j \leqslant k} \rho_j(t)$ is monotone decreasing in t. Before giving the proof of Lemma 3.1, we first show how it implies Theorem 1.1 (as we shall see, by Theorem 3 in [25] it suffices to establish (1.2), *i.e.*, local convergence). Aiming at a contradiction, suppose that (1.2) fails, that is, there exists $t_0 \geqslant 0$, $k_0 \geqslant 1$, $\varepsilon > 0$ and an infinite set of $\bar{n} \in \mathbb{N}$ such that $|N_{\bar{n},k_0,t_0\bar{n}}/\bar{n} - \hat{\rho}_{k_0}(t_0)| > \varepsilon$ holds with probability at least ε, where $\hat{\rho}_{k_0}(t)$ is given by the (by assumption) unique solution to (2.6)–(2.9). Since $\mathbb{P}(\mathcal{G}_n) \to 1$ as $n \to \infty$, this implies (by discarding a finite number of elements in the beginning) that there exists an infinite sequence of sample points $(\omega_{\bar{n}})$ with $\omega_{\bar{n}} \in \mathcal{G}_{\bar{n}} \subseteq \Omega_{\bar{n}}$ that satisfy

$$|N_{\bar{n},k_0,t_0\bar{n}}(\omega_{\bar{n}})/\bar{n} - \hat{\rho}_{k_0}(t_0)| > \varepsilon. \tag{3.7}$$

Now Lemma 3.1 gives a subsequence $(\omega_{\tilde{n}})$ satisfying $N_{\tilde{n},k,t\tilde{n}}(\omega_{\tilde{n}})/\tilde{n} \to \rho_k(t)$ for each $t \geqslant 0$ and $k \geqslant 1$, where the $(\rho_k(t))_{k \geqslant 1}$ solve (2.6)–(2.9) on $\mathbb{R}^+$. But by (3.7) we also have

$$|\rho_{k_0}(t_0) - \hat{\rho}_{k_0}(t_0)| > \varepsilon/2,$$

https://doi.org/10.1017/S0963548315000218 Published online by Cambridge University Press

contradicting uniqueness of the solution to (2.6)–(2.9). It follows that (1.2) holds for each fixed $k \geqslant 1$ and $t \geqslant 0$, that is, $\mathcal{R}$ is locally convergent. Now Theorem 3 in [25] implies that $\mathcal{R}$ is also globally convergent, with continuous scaling limit $\rho^{\mathcal{R}}(t) = 1 - \sum_{k \geqslant 1} \hat{\rho}_k(t)$, completing the proof of Theorem 1.1.

3.2. Proof of Lemma 3.1

Following step (2) of our proof outline, we start by selecting a 'nice' subsequence of (ω_n), proceeding as in the proof of Helly's selection theorem (see, *e.g.*, Theorem 5.8.1 in [15]). Define $F_n(k,t) = N_{n,k,tn}(\omega_n)/n$ if $1 \leqslant k \leqslant n$; otherwise set $F_n(k,t) = 0$. Clearly, $F_n(k,t) \in [0,1]$. Furthermore, $F_n(1,0) = 1$ and $F_n(k,0) = 0$ for $k \geqslant 2$. Let $(q_r)_{r \geqslant 1}$ be an enumeration of $\mathbb{Q}^+$. A standard diagonal argument yields a subsequence $(\omega_{\tilde{n}})$ such that for all $(k, q_r) \in \mathbb{N} \times \mathbb{Q}^+$ the value of $F_{\tilde{n}}(k, q_r)$ converges to some limit s_{k,q_r}. For each $k \in \mathbb{N}$ we now define $\rho_k(q_r) = s_{k,q_r}$ for all $q_r \in \mathbb{Q}^+$. Since $N_{n,k,m}$ changes by at most ℓk per step, as a function of t each $F_n(k,t)$ is Lipschitz on $\mathbb{R}^+$ with constant ℓk, so ρ_k has this property on $\mathbb{Q}^+$. For each $k \in \mathbb{N}$ we can thus extend ρ_k to a Lipschitz-continuous function on $\mathbb{R}^+$. Henceforth we always work with the subsequence selected above, but write n instead of $\tilde{n}$ for ease of notation. For each $t \in \mathbb{R}^+$ and $k \in \mathbb{N}$ we then have

$$N_{n,k,tn}(\omega_n)/n \to \rho_k(t), \tag{3.8}$$

which is the rigorous analogue of (3.2) from our proof outline. Turning to some basic properties of the $\rho_k(t)$, by counting vertices we see that $0 \leqslant \rho_k(t) \leqslant 1$ and

$$0 \leqslant \sum_{k \geqslant 1} \rho_k(t) \leqslant 1. \tag{3.9}$$

Furthermore, the initial conditions $\rho_k(0) = \mathbb{1}_{\{k=1\}}$ hold for $k \geqslant 1$. To summarize, so far we have established (2.6) and (2.9).

For the proof of Lemma 3.1 it remains to show that the $\big(\rho_k(t)\big)_{k \geqslant 1}$ satisfy the differential equations (2.7). Here our main ingredient will be the *deterministic* result Lemma 3.2 below, which is the rigorous analogue of (3.3) from our proof outline. Recall that $\Delta N_{n,k,m+1} = N_{n,k,m+1} - N_{n,k,m}$. For brevity, we write $f_k(t) = f_k(t, \rho_1(t), \rho_2(t), \ldots)$ for the right-hand side of (2.7).

Lemma 3.2. *Let $\omega_n \in \mathcal{G}_n \subseteq \Omega_n$ be defined for an infinite set of $n \in \mathbb{N}$, and suppose that* (3.8) *holds. Then for all $t \geqslant 0$, $k \geqslant 1$ and $\varepsilon > 0$ there exists $0 < \delta \leqslant 1$ such that for n large enough the following holds: for all $m \geqslant 0$ satisfying $|m - tn| \leqslant \delta n$ we have*

$$\mathbb{E}(\Delta N_{n,k,m+1} \mid \mathcal{F}_{n,m})(\omega_n) = f_k(t) \pm \varepsilon/3. \tag{3.10}$$

Proof. Recall that $\omega_n \in \mathcal{K}_n \cap \mathcal{U}_n$ satisfies (3.8). Given $t \geqslant 0$, $k \geqslant 1$ and $\varepsilon > 0$, pick $0 < \gamma \leqslant \varepsilon/(90\ell^2 k)$. Recall that by definition $\rho(t) = 1 - \sum_{k \geqslant 1} \rho_k(t) \in [0,1]$; see (2.8) and (3.9). Let $\rho_{<s}(t) = \sum_{1 \leqslant k < s} \rho_k(t)$, which is increasing in s with limit $1 - \rho(t)$. Choose an integer $S \geqslant k$ such that $S \geqslant K(\gamma)$ and $\rho_{<S}(t) \geqslant 1 - \rho(t) - \gamma$, and let $0 < \delta \leqslant \min\{\gamma/(9\ell S^2), 1\}$. By

https://doi.org/10.1017/S0963548315000218 Published online by Cambridge University Press

choice of S we have

$$\sum_{k\geqslant S}\rho_k(t)=\sum_{k\geqslant 1}\rho_k(t)-\rho_{<S}(t)=1-\rho(t)-\rho_{<S}(t)\leqslant\gamma. \tag{3.11}$$

Consider $m\geqslant 0$ satisfying $|m-tn|\leqslant\delta n$. Since $\eta(n)\leqslant\gamma$ for n large enough (see (3.4)) using $\omega_n\in\mathcal{U}_n$ we see that

$$N_{n,\geqslant\gamma n,m}(\omega_n)>0 \quad \text{implies} \quad N_{n,\geqslant\gamma n,m}(\omega_n)=L_{1,n,m}(\omega_n). \tag{3.12}$$

Furthermore, since $\omega_n\in\mathcal{K}_n$ and $S\geqslant K(\gamma)$, by (3.5) we have $N_{n,\geqslant S,m}(\omega_n)\leqslant L_{1,n,m}(\omega_n)+\gamma n$ for n large enough. So, by distinguishing whether $L_{1,n,m}(\omega_n)$ is larger or smaller than γn, we infer

$$N_{n,\geqslant S,m}(\omega_n)-N_{n,\geqslant\gamma n,m}(\omega_n)\leqslant 2\gamma n. \tag{3.13}$$

We shall now evaluate $\mathbb{E}(\Delta N_{n,k,m+1}\mid\mathcal{F}_{n,m})(\omega_n)$. For this we regard the graph $G_{n,m}(\omega_n)$ as fixed, and the vertices $\underline{v}_{n,m+1}=(v_1,\dots,v_\ell)$ as random. So, in the following all probabilities $\mathbb{P}(*)$ are shorthand for $\mathbb{P}(*\mid\mathcal{F}_{n,m})(\omega_n)$. Recall the definitions of $\underline{v}_{n,m+1}=(v_1,\dots,v_\ell)$ and $\underline{c}_{n,m+1}=(c_1,\dots,c_\ell)$: the vertices $v_1,\dots,v_\ell$ are chosen independently and uniformly at random from $[n]$, and c_j denotes the size of the component in $G_{n,m}(\omega_n)$ containing v_j. So, for each $s\in[n]$ we have

$$\mathbb{P}(c_j=s)=N_{n,s,m}(\omega_n)/n.$$

We define $\mathcal{T}$ as the event that (i) all vertices v_j with $c_j\leqslant S$ are in different components, and (ii) there are no vertices v_j with $S<c_j<\gamma n$. Let $g(\cdot)$ be the function appearing in the definition (2.5) of well-behaved, which satisfies $g(S)\geqslant S$. Clearly, $g(S)<\gamma n$ for n large enough. Note that whenever $\mathcal{T}$ holds, by (3.12) all v_j in components of size at least $\gamma n>g(S)$ are in the same component (the largest), so (2.4) or (2.5) applies, giving

$$\mathbb{E}(\Delta N_{n,k,m+1}\mid\mathcal{F}_{n,m},\underline{v}_{n,m+1})=d_k(\tilde{c}_1,\dots,\tilde{c}_\ell), \tag{3.14}$$

where $\tilde{c}_j=\infty$ if $c_j\geqslant\gamma n$, and $\tilde{c}_j=c_j$ otherwise. Whether or not $\mathcal{T}$ holds, the two sides of (3.14) are bounded by ℓk. Using (3.13) we see that $\mathbb{P}(\neg\mathcal{T})\leqslant\ell^2S/n+2\ell\gamma$ (by separately estimating the probability that (i) or (ii) fails), and so by choice of γ we have, say,

$$4\ell k\cdot\mathbb{P}(\neg\mathcal{T})\leqslant 4\ell k\cdot(\ell^2S/n+2\ell\gamma)\leqslant 10\ell^2k\gamma\leqslant\varepsilon/9$$

for n large enough. Setting $\mathfrak{S}=[S]\cup\{s\in[n]:s\geqslant\gamma n\}$, by taking expectations of both sides of (3.14), it follows that

$$\mathbb{E}(\Delta N_{n,k,m+1}\mid\mathcal{F}_{n,m})(\omega_n)=\sum_{s_1,\dots,s_\ell\in\mathfrak{S}}d_k(\mathfrak{s}_1,\dots,\mathfrak{s}_\ell)\prod_{j\in[\ell]}\mathbb{P}\big(c_j=s_j\big)\pm\varepsilon/9, \tag{3.15}$$

where $\mathfrak{s}_j=\infty$ if $s_j\geqslant\gamma n$, and $\mathfrak{s}_j=s_j$ otherwise.

For the reader interested in the robustness remark of Section 2, we point out that in the estimates above we had plenty of elbow room. In particular, if the conditional distribution of $\underline{v}_{n,m+1}$ is at total variation distance $\alpha_n=o(1)$ from the one used above, then (recalling $|\Delta N_{n,k,m+1}|\leqslant\ell k$ and $|d_k(\cdot)|\leqslant\ell k$) a simple coupling argument shows that this only adds an additive error of at most, say, $4\ell k\alpha_n$, which is negligible (say at most $\varepsilon/99$) for n large enough. With this in mind, (3.15) is easily seen to still hold in such slight variations.

https://doi.org/10.1017/S0963548315000218 Published online by Cambridge University Press

After this short interlude, we define a probability distribution Y as follows for $s \in \mathbb{N} \cup \{\infty\}$:

$$\mathbb{P}\big(Y = s\big) \;=\; \begin{cases} N_{n,\geqslant \gamma n,m}(\omega_n)/n & \text{if } s = \infty, \\ N_{n,s,m}(\omega_n)/n & \text{if } s < \gamma n, \\ 0 & \text{otherwise.} \end{cases} \tag{3.16}$$

Let $Y_1, \ldots, Y_\ell$ be i.i.d. with distribution Y and observe that $\mathbb{P}(Y_j = s) = \mathbb{P}(c_j = s)$ for $s \leqslant S < \gamma n$. Since $\tilde{s}_j = \infty$ for $s_j \geqslant \gamma n$, using $\sum_{s \geqslant \gamma n} \mathbb{P}\big(c_j = s\big) = \mathbb{P}\big(Y = \infty\big)$ it follows that (3.15) gives

$$\mathbb{E}(\Delta N_{n,k,m+1} \mid \mathcal{F}_{n,m})(\omega_n) = \sum_{s_1,\ldots,s_\ell \in [S] \cup \{\infty\}} d_k(s_1, \ldots, s_\ell) \prod_{j \in [\ell]} \mathbb{P}\big(Y_j = s_j\big) \pm \varepsilon/9. \tag{3.17}$$

From (3.13) and the definition of Y we have $\mathbb{P}(S < Y < \infty) = \mathbb{P}(S < Y < \gamma n) \leqslant 2\gamma$. Since $|d_k(\cdot)| \leqslant \ell k$, we can extend the sum to all $s_1, \ldots, s_\ell \in \mathbb{N} \cup \{\infty\}$ at the price of an additive error of $4\gamma \ell^2 k$, say. Since $4\gamma \ell^2 k \leqslant \varepsilon/18$ by choice of γ, this gives

$$\mathbb{E}(\Delta N_{n,k,m+1} \mid \mathcal{F}_{n,m})(\omega_n) = \mathbb{E}(d_k(Y_1, \ldots, Y_\ell)) \pm \varepsilon/6. \tag{3.18}$$

For $s \leqslant S$ note that $N_{n,s,m}$ changes by at most $\ell s \leqslant \ell S$ in each step, so $|m - tn| \leqslant \delta n$ implies $|N_{n,s,m}(\omega_n) - N_{n,s,tn}(\omega_n)| \leqslant \ell S \delta n$. Hence, using the definition of δ and (3.8), for $s \leqslant S$ and n large enough we have

$$|N_{n,s,m}(\omega_n)/n - \rho_s(t)| \leqslant \ell S \delta + \gamma/(2S) \leqslant \gamma/S. \tag{3.19}$$

Using this observation we shall now show that the right-hand side of (3.18) is essentially determined by the $(\rho_k(t))_{k \geqslant 1}$; this is key for our approach. To this end consider the probability distribution Z, which is defined as follows for $s \in \mathbb{N} \cup \{\infty\}$:

$$\mathbb{P}\big(Z = s\big) \;=\; \begin{cases} \rho(t) = 1 - \sum_{k \geqslant 1} \rho_k(t) & \text{if } s = \infty, \\ \rho_s(t) & \text{otherwise.} \end{cases} \tag{3.20}$$

Claim 3.3. *For n large enough we have*

$$\mathrm{d_{TV}}\big(Y, Z\big) \leqslant 4\gamma.$$

Proof. Recall that the total variation distance is given by

$$\mathrm{d_{TV}}\big(Y, Z\big) = \frac{1}{2} \sum_{s \in \mathbb{N} \cup \{\infty\}} |\mathbb{P}(Y = s) - \mathbb{P}(Z = s)|. \tag{3.21}$$

For $s \leqslant S$, by (3.19) we have

$$\sum_{s \in [S]} |\mathbb{P}(Y = s) - \mathbb{P}(Z = s)| \leqslant \gamma$$

for n large enough. Next, we consider the summands where $s \in \mathbb{N} \setminus [S]$. Recalling (3.16) and (3.13), we have $\mathbb{P}\big(Y \in \mathbb{N} \setminus [S]\big) \leqslant 2\gamma$. Similarly, from (3.20) and (3.11) we have

https://doi.org/10.1017/S0963548315000218 Published online by Cambridge University Press

$\mathbb{P}\big(Z \in \mathbb{N} \setminus [S]\big) \leqslant \gamma$. Thus

$$\sum_{s \in \mathbb{N} \setminus [S]} |\mathbb{P}(Y = s) - \mathbb{P}(Z = s)| \leqslant 3\gamma.$$

Finally, since Y and Z are probability distributions, they differ on $s = \infty$ by no more than the sum of the differences of the other values, that is, by at most 4γ. In view of (3.21), the claim follows. □

Taking $Z_1, \dots, Z_\ell$ i.i.d. with distribution Z, using Claim 3.3 the distributions of $(Y_1, \dots, Y_\ell)$ and $(Z_1, \dots, Z_\ell)$ can be coupled such that they agree with probability at least $1 - 4\ell\gamma$. So, since $|d_k(\cdot)| \leqslant \ell k$, in (3.18) we may replace all occurrences of Y_j by Z_j at the price of an additive error of $8\gamma\ell^2 k$, say. Since $8\gamma\ell^2 k \leqslant \varepsilon/6$ by choice of γ, it follows that

$$\mathbb{E}(\Delta N_{n,k,m+1} \mid \mathcal{F}_{n,m})(\omega_n) = \mathbb{E}(d_k(Z_1, \dots, Z_\ell)) \pm \varepsilon/3.$$

Recall that $f_k(t) = f_k(t, \rho_1(t), \rho_2(t), \dots)$ is shorthand for the right-hand side of (2.7). By definition of the Z_j (see (3.20)), it follows that $\mathbb{E}(d_k(Z_1, \dots, Z_\ell)) = f_k(t)$, completing the proof of (3.10), that is, of Lemma 3.2. □

Finally, we are now ready to complete the proof of Lemma 3.1. As discussed previously in Lemma 3.2, passing to a subsequence we may assume (3.8), and it remains to establish (2.7). Aiming at a contradiction, suppose that (2.7) fails for some $t \geqslant 0$ and $k \geqslant 1$. By definition of $\rho_k'(t)$ and $f_k(t) = f_k(t, \rho_1(t), \rho_2(t), \dots)$ this implies that there exists $\varepsilon > 0$ such that for all $\delta > 0$ there is an h with $0 < |h| < \delta$ such that

$$\left| \frac{\rho_k(t+h) - \rho_k(t)}{h} - f_k(t) \right| \geqslant \varepsilon, \tag{3.22}$$

where if $t = 0$ we may also assume $t + h > 0$ (since we only consider the right-derivative for $t = 0$.) For these $t \geqslant 0$, $k \geqslant 1$ and $\varepsilon > 0$, we now pick $0 < \delta \leqslant 1$ as given by Lemma 3.2. Decreasing δ if necessary, in the case of $t > 0$ we may also assume that $\delta \leqslant t/2$ holds, say. For each h with $0 < |h| \leqslant \delta$ and $t + h \geqslant 0$ we now write m_1, m_2 for the minimum and maximum of $\{(t+h)n, tn\}$, which satisfy $m_1 \geqslant 0$ and $0 < m_2 - m_1 < n^{1+\lambda}$. Recall that $\omega_n \in \mathcal{D}_n$, and note that $k \leqslant n^\lambda$ for n large enough. Now, using (3.6) and (3.10) we see that for n large enough

$$\begin{aligned} &N_{n,k,(t+h)n}(\omega_n) - N_{n,k,tn}(\omega_n) \\ &\quad = \operatorname{sgn}(h) \cdot \Big(\sum_{m_1 \leqslant m < m_2} \mathbb{E}(\Delta N_{n,k,m+1} \mid \mathcal{F}_{n,m})(\omega_n) \Big) \pm n^{1/2+2\lambda} \\ &\quad = hn \cdot (f_k(t) \pm \varepsilon/3) \pm n^{1/2+2\lambda}. \end{aligned}$$

Rearranging terms, using (3.8) and $\lambda < 1/4$ we deduce that for n large enough we have, say,

$$\left| \frac{\rho_k(t+h) - \rho_k(t)}{h} - f_k(t) \right| \leqslant \varepsilon/2 + n^{-1/2+2\lambda}/|h| < \varepsilon, \tag{3.23}$$

https://doi.org/10.1017/S0963548315000218 Published online by Cambridge University Press

contradicting (3.22) above. This establishes (2.7) for all $t \geqslant 0$ and $k \geqslant 1$, completing the proof of Lemma 3.1.

4. Emergence of the giant component

In this section we demonstrate that our approach may still yield useful information in the presence of multiple solutions to the associated system of differential equations: using the emergence of the giant component as an example, we show that properties common to all solutions of the differential equations often transfer to the discrete random process.

We start by briefly recalling the strategy used in the proof of Theorem 1.1. Namely, we first defined events $\mathcal{G}_n$ with $\mathbb{P}(\mathcal{G}_n) \to 1$ as $n \to \infty$, and then showed that any sequence (ω_n) of 'runs' of an Achlioptas process with $\omega_n \in \mathcal{G}_n$ has a subsequence $(\omega_{\tilde{n}})$ where $(N_{\tilde{n},k,t\tilde{n}}(\omega_{\tilde{n}})/\tilde{n})_{k\geqslant 1}$ converges to a solution $(\rho_k(t))_{k\geqslant 1}$ of the associated system of differential equations, also with continuous $\rho(t) = 1 - \sum_{k\geqslant 1} \rho_k(t)$ by Theorem 3 in [25]. With this in mind, Remark 2 follows since the proof of Theorem 3 in [25] carries over *mutatis mutandis* to any interval of form $I = [0, t^*]$ or $I = [0, t^*)$. In other words, for any such interval I we obtain convergence to the (by assumption) unique solution $(\hat{\rho}_k(t))_{k\geqslant 1}$, with continuous $\rho(t) = 1 - \sum_{k\geqslant 1} \rho_k(t)$. This is important since it may well be that uniqueness for the system of differential equations can be established only up to some point; in particular, uniqueness after 'gelation' (see, *e.g.*, [2]), that is, when $\sum_{k\geqslant 1} \rho_k(t) < 1$, seems to be much harder to establish.

In general, we do not know if there is a unique 'critical' *gelation point* t_c (with $\sum_{k\geqslant 1} \rho_k(t) = 1$ for $t < t_c$ and $\sum_{k\geqslant 1} \rho_k(t) < 1$ for $t > t_c$); different solutions might give different gelation points in some range. However, with Theorems 4.1 and 4.2 we shall show that the giant component emerges at some point within this range (without assuming any uniqueness).

Our first result gives conditions sufficient to guarantee that all components have sublinear size. In view of (2.8) our assumption is natural: that every solution to the differential equations satisfies $\sum_{k\geqslant 1} \rho_k(t) = 1$ (*i.e.*, $\rho(t) = 0$).

Theorem 4.1. *Let $\ell \geqslant 2$ and let $\mathcal{R}$ be a merging ℓ-vertex rule that is well-behaved. Assume that for some $t^* \in [0, \infty)$ every solution $(\tilde{\rho}_k(t))_{k\geqslant 1}$ to the associated system of differential equations given by (2.6)–(2.9) satisfies $\sum_{k\geqslant 1} \tilde{\rho}_k(t^*) = 1$. Then for any $0 \leqslant t \leqslant t^*$ we have $L_1(tn)/n \xrightarrow{p} 0$.*

Proof. By monotonicity it suffices to show that $L_1(t^*n)/n \xrightarrow{p} 0$. Recall that in the proofs we indicate the dependence on n explicitly, writing, for example, L_{1,n,t^*n} for $L_1(t^*n)$; see (2.1)–(2.3). Proceeding along the lines of the proof of Theorem 1.1, suppose there exists $\delta > 0$ and an infinite set of $\bar{n} \in \mathbb{N}$ with $\mathbb{P}(L_{1,\bar{n},t^*\bar{n}}/\bar{n} \geqslant \delta) \geqslant \delta$. Then, since $\mathbb{P}(\mathcal{G}_n) \to 1$ as $n \to \infty$, there exists an infinite sequence of sample points $(\omega_{\bar{n}})$ with $\omega_{\bar{n}} \in \mathcal{G}_{\bar{n}} \subseteq \Omega_{\bar{n}}$ and

$$L_{1,\bar{n},t^*\bar{n}}(\omega_{\bar{n}})/\bar{n} \geqslant \delta. \tag{4.1}$$

Now Lemma 3.1 gives a subsequence $(\omega_{\tilde{n}})$ with $N_{\tilde{n},k,t\tilde{n}}(\omega_{\tilde{n}})/\tilde{n} \to \rho_k(t)$ for each $t \geqslant 0$ and $k \geqslant 1$, where the $(\rho_k(t))_{k\geqslant 1}$ solve (2.6)–(2.9). Hence, by assumption we have

https://doi.org/10.1017/S0963548315000218 Published online by Cambridge University Press

$\sum_{k\geqslant 1}\rho_k(t^*)=1$, and so for some K we have $\sum_{1\leqslant k\leqslant K}\rho_k(t^*)\geqslant 1-\delta/4$. Since

$$L_{1,\tilde{n},m}\leqslant \max\{N_{\tilde{n},\geqslant K+1,m},K\}\leqslant \tilde{n}-N_{\tilde{n},\leqslant K,m}+K,$$

for $\tilde{n}$ sufficiently large we infer

$$\begin{aligned} L_{1,\tilde{n},t^*\tilde{n}}(\omega_{\tilde{n}})/\tilde{n} &\leqslant 1-N_{\tilde{n},\leqslant K,t^*\tilde{n}}(\omega_{\tilde{n}})/\tilde{n}+K/\tilde{n}\\ &\leqslant 1-\sum_{1\leqslant k\leqslant K}\rho_k(t^*)+\delta/4\leqslant \delta/2, \end{aligned} \tag{4.2}$$

contradicting (4.1). □

Our next result gives conditions sufficient to guarantee the emergence of a linear-size component. The main assumption will be that every solution to the differential equations has $\rho(t^*)>0$ (*i.e.*, $\sum_{k\geqslant 1}\rho_k(t^*)<1$). This can be restated as the *non-existence* of a solution with $\rho(t^*)=0$; when effectively (as here) imposing the condition $\rho(t^*)=0$, we may simplify the equations, replacing (2.7) by

$$\rho_k'(t)=\sum_{c_1,\dots,c_\ell\in\mathbb{N}} d_k(c_1,\dots,c_\ell)\prod_{j\in[\ell]}\rho_{c_j}(t). \tag{4.3}$$

This generalizes the Smoluchowksi coagulation equations (see, *e.g.*, [2, 21]) in a form without sol-gel interaction. The advantage is that it allows us to drop condition (2.5) in the following theorem.

Theorem 4.2. *Let $\ell\geqslant 2$ and let $\mathcal{R}$ be a merging ℓ-vertex rule that satisfies assumption* (2.4). *Assume that for some $t^*\in[0,\infty)$ every solution $(\tilde{\rho}_k(t))_{k\geqslant 1}$ on $[0,t^*]$ to the associated system of differential equations given by* (2.6), (2.9) *and* (4.3) *satisfies $\sum_{k\geqslant 1}\tilde{\rho}_k(t^*)<1$. Then for any $t^*\leqslant t<\infty$ and $\varepsilon>0$ there exist $\alpha,n_0>0$ such that $\mathbb{P}(L_1(tn)\geqslant \alpha n)\geqslant 1-\varepsilon$ for all $n\geqslant n_0$.*

Proof. By monotonicity it suffices to establish the claim for $t=t^*$. As in earlier proofs, we make the dependence on n explicit; see Theorem 4.1 and (2.1)–(2.3). Aiming at a contradiction, suppose there exists $\varepsilon>0$ such that for all $\alpha,n_0>0$ we have $\mathbb{P}(L_{1,n,t^*n}\leqslant \alpha n)\geqslant\varepsilon$ for some $n\geqslant n_0$. It follows as usual that there exists $\hat{\alpha}(n)\to 0$ as $n\to\infty$ and an infinite set of $\bar{n}\in\mathbb{N}$ such that $\mathbb{P}(L_{1,\bar{n},t^*\bar{n}}\leqslant\hat{\alpha}(\bar{n})\bar{n})\geqslant\varepsilon$. Define $\mathcal{L}_n$ as the event that $L_{1,n,t^*n}\leqslant\hat{\alpha}(n)n$. Since $\mathbb{P}(\mathcal{G}_n)\to 1$ as $n\to\infty$ there exists an infinite sequence of sample points $(\omega_{\bar{n}})$ with $\omega_{\bar{n}}\in\mathcal{G}_{\bar{n}}\cap\mathcal{L}_{\bar{n}}\subseteq\Omega_{\bar{n}}$, for which we now prove the following variant of Lemma 3.1.

Claim 4.3. *There is a subsequence $(\omega_{\bar{n}})$ of $(\omega_{\bar{n}})$ such that for each $0\leqslant t\leqslant t^*$ and $k\geqslant 1$ we have*

$$N_{\tilde{n},k,t\tilde{n}}(\omega_{\tilde{n}})/\tilde{n}\to\rho_k(t), \tag{4.4}$$

where the $(\rho_k(t))_{k\geqslant 1}$ are functions satisfying the system of differential equations (2.6), (2.9), (4.3) *on* $[0,t^*]$.

https://doi.org/10.1017/S0963548315000218 Published online by Cambridge University Press

Proof. Defining $d_k(c_1, \ldots, c_\ell) = 0$ if any argument is infinite, note that (4.3) equals (2.7). So, in view of Section 3.2, it suffices to prove Lemma 3.2 for $0 \leqslant t \leqslant t^*$. We closely follow the original argument, only changing some minor details (we also write n instead of $\tilde{n}$ for ease of notation). When selecting the parameters γ, S, δ we use $S \geqslant K(\gamma/9)$ instead of $S \geqslant K(\gamma)$. Observe that $N_{n,\geqslant S,m}$ increases by at most ℓS in each step, so by choice of δ we have

$$N_{n,\geqslant S,(t^*+\delta)n} \leqslant N_{n,\geqslant S,t^*n} + \ell\delta Sn \leqslant N_{n,\geqslant S,t^*n} + \gamma n/9.$$

Since $\omega_n \in \mathcal{K}_n \cap \mathcal{L}_n$ and $S \geqslant K(\gamma/9)$, we also have

$$N_{n,\geqslant S,t^*n}(\omega_n) \leqslant L_{1,n,t^*n}(\omega_n) + \gamma n/9 \leqslant (\hat{\alpha}(n) + \gamma/9)n \leqslant \gamma n/3, \tag{4.5}$$

for n large enough. For $t \leqslant t^*$, by combining these estimates with monotonicity, we deduce that for n sufficiently large we have

$$L_{1,n,m}(\omega_n) \leqslant L_{1,n,(t^*+\delta)n}(\omega_n) \leqslant \max\{N_{n,\geqslant S,(t^*+\delta)n}(\omega_n), S\} \leqslant \gamma n/2 \tag{4.6}$$

for every $m \geqslant 0$ with $|m - tn| \leqslant \delta$.

When establishing (3.14) the assumption (2.4) thus always applies (whenever the event $\mathcal{T}$ holds all vertices are in different components and satisfy $c_j \leqslant S$ by (4.6)). Consequently (3.15) holds, since $d_k(c_1, \ldots, c_\ell) = 0$ if $\infty \in \{c_1, \ldots, c_\ell\}$. Now the remainder of the argument leading to Lemma 3.2 is unchanged, which, as discussed, completes the proof of Claim 4.3. □

Now consider a subsequence $(\omega_{\tilde{n}})$ with the properties guaranteed by Claim 4.3. From (4.5), for n large we have

$$N_{\tilde{n},\leqslant S,t^*\tilde{n}}(\omega_{\tilde{n}}) \geqslant n - N_{\tilde{n},\geqslant S,t^*\tilde{n}}(\omega_{\tilde{n}}) \geqslant (1 - \gamma/3)n,$$

so from (4.4) it follows that

$$\sum_{k\geqslant 1} \rho_k(t^*) \geqslant \sum_{1\leqslant k\leqslant S} \rho_k(t^*) \geqslant 1 - \gamma.$$

Since we could choose the constant γ to be arbitrarily small, we have $\sum_{k\geqslant 1} \rho_k(t^*) \geqslant 1$. Since $(\rho_k(t))_{k\geqslant 1}$ is a solution to (2.6), (2.9) and (4.3) on $[0, t^*]$, this contradicts the assumptions of the theorem. □

Remark 4. Of course, in the assumptions of Theorem 4.2 we may replace (4.3) by (2.7). Indeed, using (3.9), the $(\rho_k(t))_{k\geqslant 1}$ constructed above satisfy $\sum_{k\geqslant 1} \rho_k(t) = 1$ for $0 \leqslant t \leqslant t^*$. Thus they also solve (2.7), since $\rho(t) = 0$ for $0 \leqslant t \leqslant t^*$.

Remark 5. Theorem 4.2 also holds without the merging assumption; we outline the minor modifications needed for the proof. Using Remark 9 in [25], we replace 'at most one component' by 'at most $\ell - 1$ components' in the definition of $\mathcal{U}_n$, and replace $L_{1,n,m}$ by $L_{n,m}$ in the definition of $\mathcal{K}_n$, where $L_{n,m}$ denotes the sum of the sizes of the $\ell - 1$ largest components. Now, thinking of all 'infinity' terms as the probability of being in one of the

https://doi.org/10.1017/S0963548315000218 Published online by Cambridge University Press

$\ell - 1$ largest components (of size $\geqslant \gamma n$), using $L_{n,m} \leqslant \ell \cdot L_{1,n,m}$ it is not difficult to push the argument through; we omit the details.

It might be surprising that the sol-gel interaction and condition (2.5) are used in Theorem 4.1 but not Theorem 4.2 (rather than the other way round). The explanation is that our proofs proceed by contradiction, showing the existence of a gelating solution in the case of Theorem 4.1, and a non-gelating solution in the case of Theorem 4.2. Nevertheless, since condition (2.5) essentially ensures that the giant component, once it emerges, evolves in a regular way, it may well not be needed in Theorem 4.1.

Acknowledgement

We are grateful to the referee for a very careful reading of the paper, and for helpful suggestions concerning the presentation.

References

[1] Achlioptas, D., D'Souza, R. M. and Spencer, J. (2009) Explosive percolation in random networks. *Science* **323** 1453–1455.
[2] Aldous, D. (1999) Deterministic and stochastic models for coalescence (aggregation and coagulation): A review of the mean-field theory for probabilists. *Bernoulli* **5** 3–48.
[3] Azar, Y., Broder, A. Z., Karlin, A. R. and Upfal, E. (1999) Balanced allocations. *SIAM J. Comput.* **29** 180–200.
[4] Bhamidi, S., Budhiraja, A. and Wang, X. (2015) Aggregation models with limited choice and the multiplicative coalescent. *Random Struct. Alg.* **46** 55–116.
[5] Bhamidi, S., Budhiraja, A. and Wang, X. (2014) Bounded-size rules: The barely subcritical regime. *Combin. Probab. Comput.* **23** 505–538.
[6] Bohman, T. and Frieze, A. (2001) Avoiding a giant component. *Random Struct. Alg.* **19** 75–85.
[7] Bohman, T. and Kravitz, D. (2006) Creating a giant component. *Combin. Probab. Comput.* **15** 489–511.
[8] Bollobás, B., Janson, S. and Riordan, O. (2007) The phase transition in inhomogeneous random graphs. *Random Struct. Alg.* **31** 3–122.
[9] Bollobás, B. and Riordan, O. (2009) Random graphs and branching processes. In *Handbook of Large-Scale Random Networks*, Vol. 18 of *Bolyai Society Mathematical Studies*, pp. 15–115.
[10] Cho, Y. S., Kim, S. W., Noh, J. D., Kahng, B. and Kim, D. (2010) Finite-size scaling theory for explosive percolation transitions. *Phys. Rev. E* **82** 042102.
[11] da Costa, R. A., Dorogovtsev, S. N., Goltsev, A. V. and Mendes, J. F. F. (2010) Explosive percolation transition is actually continuous. *Phys. Rev. Lett.* **105** 255701.
[12] D'Souza, R. M. and Mitzenmacher, M. (2010) Local cluster aggregation models of explosive percolation. *Phys. Rev. Lett.* **104** 195702.
[13] Erdős, P. and Rényi, A. (1960) On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl* **5** 17–61.
[14] Friedman, E. J. and Landsberg, A. S. (2009) Construction and analysis of random networks with explosive percolation. *Phys. Rev. Lett.* **103** 255701.
[15] Gut, A. (2005) *Probability: A Graduate Course*, Springer.
[16] Hurewicz, W. (1958) *Lectures on Ordinary Differential Equations*, MIT Press.
[17] Janson, S. (2011) Networking: Smoothly does it. *Science* **333** 298–299.
[18] Janson, S., Łuczak, T. and Ruciński, A. (2000) *Random Graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience.

[19] Janson, S. and Spencer, J. (2012) Phase transitions for modified Erdős–Rényi processes. *Ark. Math.* **50** 305–329.
[20] Kang, M., Perkins, W. and Spencer, J. (2013) The Bohman–Frieze process near criticality. *Random Struct. Alg.* **43** 221–250.
[21] Norris, J. (1999) Smoluchowski's coagulation equation: Uniqueness, nonuniqueness and a hydrodynamic limit for the stochastic coalescent. *Ann. Appl. Probab.* **9** 78–109.
[22] Radicchi, F. and Fortunato, S. (2010) Explosive percolation: A numerical analysis. *Phys. Rev. E* **81** 036110.
[23] Riordan, O. and Warnke, L. (2011) Explosive percolation is continuous. *Science* **333** 322–324.
[24] Riordan, O. and Warnke, L. (2012) Achlioptas processes are not always self-averaging. *Phys. Rev. E* **86** 011129.
[25] Riordan, O. and Warnke, L. (2012) Achlioptas process phase transitions are continuous. *Ann. Appl. Probab.* **22** 1450–1464.
[26] Riordan, O. and Warnke, L. (2015) The evolution of subcritical Achlioptas processes. *Random Struct. Alg.* **47** 174–203.
[27] Seierstad, T. G. (2009) A central limit theorem via differential equations. *Ann. Appl. Probab.* **19** 661–675.
[28] Spencer, J. and Wormald, N. C. (2007) Birth control for giants. *Combinatorica* **27** 587–628.
[29] Wormald, N. C. (1995) Differential equations for random processes and random graphs. *Ann. Appl. Probab.* **5** 1217–1235.
[30] Wormald, N. C. (1999) The differential equation method for random graph processes and greedy algorithms. In *Lectures On Approximation and Randomized Algorithms*, PWN, pp. 73–155.
[31] Ziff, R. M. (2009) Explosive growth in biased dynamic percolation on two-dimensional regular lattice networks. *Phys. Rev. Lett.* **103** 045701.