On a Random Graph with Immigrating Vertices: Emergence of the Giant Component

David J. Aldous*
University of California
Department of Statistics
367 Evans Hall # 3860
Berkeley CA 94720-3860
aldous@stat.berkeley.edu

Boris Pittel†
Ohio State University
Department of Mathematics
231 W 18th Avenue
Columbus OH 43210-1174
bgp@math.ohio-state.edu

March 2, 2000

Abstract

A randomly evolving graph, with vertices immigrating at rate n and each possible edge appearing at rate $1/n$, is studied. The detailed picture of emergence of giant components with $O(n^{2/3})$ vertices is shown to be the same as in the Erdős - Rényi graph process with the number of vertices fixed at n at the start. A major difference is that now the transition occurs about a time $t = \pi/2$, rather than $t = 1$. The proof has three ingredients. The size of the largest component in the subcritical phase is bounded by comparison with a certain multitype branching process. With this bound at hand, the growth of the sum-of-squares and sum-of-cubes of component sizes is shown, via martingale methods, to follow closely a solution of the Smoluchowsky-type equations. The approximation allows us to apply results of Aldous (1997) on emergence of giant components in the multiplicative coalescent, i.e. a non-uniform random graph process.

AMS 1991 subject classification: 05C80,60C05,60F17

Key words and phrases. Branching Markov chain, giant component, martingale, multiplicative coalescent, random graph, weak convergence.

*Research supported by N.S.F. Grant DMS96-22859

†Research supported by N.S.F. Grant DMS98-03410, and also by the Microsoft Research in Spring, 1999

1 Introduction

In the Erdős - Rényi [6] random graph $G(n, M)$ there are n vertices and M edges and, setting $N = \binom{n}{2}$, all $\binom{N}{M}$ possible graphs are equally likely. Dynamically, $G(n, M)$ is the M-th stage of the random graph process, which is a Markov chain whose states are graphs on n vertices. $G(n, M)$ is obtained from $G(n, M-1)$ by choosing uniformly at random one of the $N-(M-1)$ pairs of vertices not yet connected and inserting an edge between them. Conceptually close, but different, is a Bernoulli-type random graph $G(n,p)$ (Stepanov [18]); here for each unordered pair $\{v, w\}$ of vertices there is probability p that $\{v, w\}$ is an edge, and all N of those events are independent. The dynamic version of this graph is obtained if we set $p = p(t) = 1-e^{-t}$, effectively assuming that the individual edge birth-times are independent and exponentially distributed. The heuristic principle, rigorously confirmed in lots of cases, is that $G(n, M)$ and $G(n,p)$ behave similarly in the limit $n \to \infty$ if $M = M(n)$ and $p = p(n)$ are such that $M = pN$. Besides, $\{G(n, M)\}_{M \geq 0}$ can be viewed as the sequence of snapshots of $\{G(n, p(t))\}_{t \geq 0}$ at the moments $t_M = \inf\{t : |E(t)| = M\}$, with $E(t)$ standing for the edge set of $G(n, p(t))$.

Many aspects of these random graphs have been studied in great detail. Special attention has been paid to the distribution of (connected) component sizes and the *emergence of the giant component* with size of order $n^{2/3}$, as M (p resp.) increases over the critical range $n/2 \pm O(n^{2/3})$, ($1/n \pm O(n^{-4/3})$ resp.) – see Bollobás [4], Stepanov [19], Janson et al [9], Kolchin [11], Luczak et al [13]. This phenomenon may be considered as a prototype for phase transitions in more general discrete structures (cf. the survey by Luczak [12]), or as a mean-field analog of d-dimensional percolation theory (cf. Grimmett [8]).

For the scaled time t/n, we can view the Erdős - Rényi - Stepanov (ERS) process as the graph-valued process specified by the following two rules.

(i) At time 0 there are n vertices and no edges.

(ii) for each pair of vertices, an edge appears at (exponential) rate $1/n$. With this scaling the critical time interval becomes $1 \pm O(n^{-1/3})$. The subject of this paper is a modification of the process which we name the *random graph with immigrating vertices* (RGIV) process, specified as follows. Rule (i) is replaced by

(iii) At time 0 there are no vertices; then vertices immigrate at rate n and rule (ii) remains in force, applying to each pair of existing vertices (see section 2.1 for precise definition).

So at time t there is a random graph $\mathcal{G}(n,t)$ with about nt vertices, and about $nt^3/6$ edges. (We suggest the reader prove the latter as a warm-up exercise.) It is easy to see heuristically (section 1.3) that the analogous critical time should be $\pi/2$. Our main purpose is to show (Theorem 1) that some more detailed structure of the emergence of the large components of size $n^{2/3}$ is the same in the RGIV process as in the ERS process. (In retrospect this may seem unsurprising, but intuition here can be shaky: in section 1.4 we mention a close model with different behavior.) We leave a study of sub(super)critical behavior of $\mathcal{G}(n,t)$ for the future. Our preliminary results indicate that, analogously to the ERS process, for $t < \pi/2 - \varepsilon$, (with high probability) the components of $\mathcal{G}(n,t)$ are mostly trees, and that there are no components with two or more cycles. For $t > \pi/2 + \varepsilon$, the graph $\mathcal{G}(n,t)$ has a giant component of asymptotic size $c(t)n$, for deterministic $c(t) > 0$. One could view the RGIV process as an attempt to model evolution of a (social) net whose size-growth rate significantly exceeds link-formation rate. Our results establish existence of a very localized time zone when a dominant component first emerges, via coalescence, from a sea of small (tree-type) components.

1.1 Statement of result

Our result will parallel the following description of the emergence of the giant component in the ERS process given by Aldous [1]. Let $\mathbf{C}^{(n)}(s) = (C_1^{(n)}(s), C_2^{(n)}(s), \ldots)$ be the component sizes of $G(n, p(s/n))$, in decreasing order. We may regard $\mathbf{C}^{(n)}(s)$ as defined for all $-\infty < s < \infty$ by interpreting $G(n,p)$ as $G(n,0)$ for $p \leq 0$. And we know that the action takes place for $|s-1| = O(n^{-1/3})$. So, setting $s = 1 + tn^{-1/3}$, we define

$$\bar{\mathbf{C}}^{(n)}(t) = n^{-2/3}\mathbf{C}^{(n)}(1 + tn^{-1/3}), \quad -\infty < t < \infty.$$

Consider $\bar{\mathbf{C}}^{(n)}(t)$ as a random element of the space

$$l^2_\downarrow = \{(x_1, x_2, \ldots) : x_1 \geq x_2 \geq \ldots \geq 0, \ \sum_i x_i^2 < \infty\},$$

and consider $\bar{\mathbf{C}}^{(n)} = (\bar{\mathbf{C}}^{(n)}(t), -\infty < t < \infty)$ as a stochastic process with values in $l^2_\downarrow$. Corollary 2 of Aldous [1] says

$$\bar{\mathbf{C}}^{(n)} \to_d \mathbf{X} \text{ as } n \to \infty \tag{1}$$

for a certain $l^2_\downarrow$-valued limit process $\mathbf{X} = (\mathbf{X}(t), -\infty < t < \infty)$, the *standard multiplicative coalescent.* Here convergence is weak convergence of processes on the time interval $(-\infty, \infty)$, with $l^2_\downarrow$ inheriting the usual topology of l^2.

That is, $\bar{\mathbf{C}}^{(n)}$ and $\mathbf{X}$ are random elements of the space $D((-\infty,\infty),l^2_\downarrow)$ of right-continuous left-limit functions from $(-\infty,\infty)$ to $l^2_\downarrow$ (cf. [7] section 3.5). Asymptotic distributional results for large component sizes in the critical interval, proved in [9], [13] may be reinterpreted as *exact* distributional assertions for the process $\mathbf{X}$. It is probably the nature of the beast that most (not all!) of these results are complicated and involve so called Wright-Stepanov's constants coming from enumerating connected graphs by vertex and edge numbers.

It seems plausible that the way the giant component emerges in the RGIV process is the same as in the ERS process. The following theorem establishes a close connection between the two processes in a near-critical time zone. Interpret $\mathcal{G}(n,t)$ as the empty graph for $t \leq 0$.

Theorem 1 *Let $\mathbf{C}^{(n)}(s) = (C_1^{(n)}(s), C_2^{(n)}(s), \ldots)$ be the component sizes in the RGIV process $\mathcal{G}(n,s)$, in decreasing order. Rescale, to define*

$$\bar{\mathbf{C}}^{(n)}(t) = (2/3)^{1/3} n^{-2/3} \mathbf{C}^{(n)}(\pi/2 + (2/3)^{2/3} t n^{-1/3}), \quad -\infty < t < \infty.$$

Then

$$\bar{\mathbf{C}}^{(n)} \to_d \mathbf{X} \text{ as } n \to \infty$$

where $\mathbf{X}$ is the standard multiplicative coalescent and convergence is weak convergence on $D((-\infty,\infty),l^2_\downarrow)$.

Loosely speaking, the large components sizes of the RGIV process at time $\pi/2 + (2/3)^{2/3} t n^{-1/3}$ multiplied by $(2/3)^{1/3}$ are close, in distribution, to those of the ERS process at time $1 + tn^{-1/3}$.

1.2 Structure of proof

The usual methods for the ERS process rely on combinatorial enumeration [4] and generating functions [9, 13, 15, 16]. The method in [1] maps graphs to walks and then applies the martingale functional CLT to the walk associated with the ERS process. None of these methods seems to apply very directly to the RGIV process, because one does not have a direct description of the structure of the component containing a specified vertex, or of the dependence between different components. We use a slightly roundabout method, which avoids needing to deal directly with distributions over the critical interval by relying on a known result which goes roughly as follows. The multiplicative coalescent is the $l^2_\downarrow$-valued Markov process which evolves in a way that generalizes the evolution of component sizes in the ERS process. That is, state $\mathbf{x} = (x_i)$ represents a configuration of clusters of masses

x_i; for each pair of clusters of masses $\{x, y\}$ say, the two clusters merge into one cluster of mass $x + y$ at stochastic rate xy. The usual ERS process, suitably rescaled, is the multiplicative coalescent whose initial configuration consists of n clusters of mass $n^{-2/3}$ each. The convergence statement (1) gives the standard multiplicative coalescent as a limit of the multiplicative coalescents started in this particular uniform way, but one also gets the same limit starting from *non-uniform* configurations satisfying certain hypotheses ([1] Proposition 4). It is this nonuniformity which suggests that results on emergence of the giant component may be fairly "robust" under varying details of the random graph model. Now in the RGIV process, it turns out that the vertices which immigrate during the critical time interval $\frac{\pi}{2} \pm O(n^{-1/3})$ make negligible contribution. That is, over the critical interval the RGIV process behaves as the multiplicative coalescent. So to prove Theorem 1 we need only check that the state of the RGIV process *just before* the critical interval satisfies the relevant hypotheses. This is the assertion of Proposition 2. The details of the argument above, deducing Theorem 1 from Proposition 2, are given in section 5.

1.3 The RGIV process just before the critical interval

Recall $\mathbf{C}^{(n)}(t) = (C_1^{(n)}(t), C_2^{(n)}(t), \ldots)$ are the component sizes in the RGIV process $\mathcal{G}(n,t)$, in decreasing order. Define

$$\begin{aligned} I_n(t) &= C_1^{(n)}(t), \\ X_n(t) &= \sum_i (C_i^{(n)}(t))^2, \\ Y_n(t) &= \sum_i (C_i^{(n)}(t))^3, \\ V_n(t) &= \text{number of vertices of } \mathcal{G}(n,t). \end{aligned}$$

Write $\xrightarrow{p}$ for convergence in probability.

Proposition 2 *Let $\alpha \in (1/9, 1/6)$ and define $t_n = \pi/2 - n^{-\alpha}$. Then as $n \to \infty$*

$$\frac{n^2 Y_n(t_n)}{X_n^3(t_n)} \xrightarrow{p} 2/3; \tag{2}$$

$$\frac{n^{4/3}}{X_n(t_n)} + n^{-\alpha+1/3} \xrightarrow{p} 0; \tag{3}$$

$$\frac{n^{2/3} I_n(t_n)}{X_n(t_n)} \xrightarrow{p} 0. \tag{4}$$

Here is a heuristic explanation for why the critical time should be $\pi/2$. *Suppose* there is a deterministic approximation

$$x(t) \approx \tfrac{1}{n} X_n(t).$$

Then $x(t)$ approximately satisfies the differential equation

$$\tfrac{dx}{dt} = x^2 + 1$$

where the "x^2" term captures coalescence of clusters and the "+1" term captures immigration of vertices. Since $x(0) = 0$, the solution is $x(t) = \tan t$, which diverges at $t = \pi/2$. Now in the ERS setting the same argument leads to the differential equation

$$\tfrac{dx}{dt} = x^2, \quad x(0) = 1$$

with solution $x(t) = (1-t)^{-1}$, which diverges at $t = 1$. It is well known ([4], [9]) that in the ERS process, with high probability a cluster of size $n^{2/3}$ is born within an $n^{-1/3}$-window centered at $t = 1$. It seems reasonable to expect some analogous behavior for the IVRG process around $t = \pi/2$, though it is not at all obvious *a priori* how close the analogy will be. This type of differential equation argument is common in the physics literature on related deterministic coagulation models, e.g. Spouge [17]. Aldous [2] studies a certain stochastic coalescence model, a one-parameter family extending the ERS process, invented to make such differential equation analysis tractable.

Our proof of Proposition 2 has two strands. In section 3 we bound the size of the largest component by showing that the component containing a particular vertex is dominated by a certain branching Markov chain. The tail of the distribution of the total progeny of this branching Markov chain is bounded using exponential martingales. In section 4 we use martingale methods to show that realizations of $X_n(t)$ and $Y_n(t)$ stay close to their deterministic approximations before the critical interval. This is similar in spirit, though quite different in implementation, to martingale analyses of stochastic coalescent models in Aldous and Limic ([3] section 3), Aldous [2], and Norris [14]. We had initially hoped that more direct generating function analysis of the subcritical IVRG process would be feasible (as is the case for the ERS process), but were unable to carry through such analysis.

1.4 A related model

As mentioned above, in a deterministic analog of the RGIV process it is known to physicists that some kind of phase transition occurs at time $\pi/2$

[17]. The only related work we know in the mathematical literature is [5] by Capobianco and Frank, who studied a discrete-time graph process which is basically a Bernoulli-type sequence with "new vertex appears" and "edge is inserted" as the possible outcomes in each trial.

As another variation, modify the RGIV process by making each component subject to removal (the vertices and edges are deleted) at stochastic rate λ. Such a process has a stationary distribution, say $\tilde{G}(n, \lambda)$, in which the mean number of vertices equals n/λ. In work-in-progress, it is shown that there is a "phase transition" in the sense that the largest component has size $o(n)$ iff $\lambda \geq 2$. One might hope that in the critical case of $\tilde{G}(n, 2)$, the component sizes would behave similarly to the critical case of the ERS process, but in fact the behavior is considerably different (e.g. the mean size of the component containing a specified vertex is bounded). In our view, determination of the "domain of attraction" of Erdős - Rényi - Stepanov type of critical behavior is a challenging problem.

2 Model and notation

2.1 The RGIV model

We start with a precise definition of the model. Fix $n > 0$. Let vertices appear (immigrate) at the times $0 < a_1^{(n)} < a_2^{(n)} < \ldots$ of a Poisson process of rate n. For each unordered pair $\{k, m\}$ let an undirected edge between the k'th and the m'th arriving vertex appear at time $\max(a_k^{(n)}, a_m^{(n)}) + E_{k,m}^{(n)}$, where the E's are independent for different pairs and have exponential (rate $1/n$) distribution. Then $\mathcal{G}(n, t)$ is the graph consisting of the vertices and edges which have appeared before time t. Recall that we are interested in the component sizes $(C_i^{(n)}(t), i \geq 1)$. These sizes are unchanged if we adopt a slightly different description of the edge-process, as follows. Assume that each *ordered* pair of vertices already present develops edges at times of a Poisson process of rate $1/(2n)$, so that multiple edges and loops are allowed. Clearly, the first edge between a pair of vertices appears with the required rate $1/n$.

By the way, the model makes sense for non-integer n; we use "n" just for comparison with the ERS process. Also our results and proofs would need only trivial changes to cover a more general case when edge-formation rate is c/n.

2.2 Notation

For events $A^{(n)}$ we say "$A^{(n)}$ occurs whp (with high probability)" if $\mathbf{Pr}\{A^{(n)}\} \to 1$ as $n \to \infty$. Unless otherwise stated, all limits are as $n \to \infty$. For non-negative random variables or constants $S^{(n)}, B^{(n)}$, we write

$$S^{(n)} = O_p(B^{(n)})$$

to mean

$$\mathbf{Pr}\{S^{(n)} > w^{(n)} B^{(n)}\} \to 0 \text{ for every } w^{(n)} \to \infty.$$

In other words, $(S^{(n)}/B^{(n)})$ is *tight.* Write

$$S^{(n)} \leq_b B^{(n)}$$

to mean that $S^{(n)}/B^{(n)}$ is bounded by a constant not depending on n. (In the deterministic case this just means $S^{(n)} = O(B^{(n)})$, but the big-O notation is ambiguous for random variables). To ease notation we usually omit the superscript n.

2.3 Simple bounds

It is convenient to record two simple bounds which are needed several times later. Write $V_n(\pi/2)$ and $\mathcal{E}_n(\pi/2)$ for the numbers of vertices and edges of $\mathcal{G}(n, \pi/2)$. Since the distribution of $V_n(\pi/2)$ is Poisson with mean $n\pi/2$, routine large deviations bounds show

$$\exists c_1 : \quad \mathbf{Pr}\{V_n(\pi/2) \geq c_1 n\} = O(2^{-n}). \tag{5}$$

Given $V_n(\pi/2) = v$, the distribution of $\mathcal{E}_n(\pi/2)$ is stochastically smaller than the Poisson distribution with mean $v^2/(2n)$, so from (5) and another use of the large deviation bound

$$\exists c_2 : \quad \mathbf{Pr}\{\mathcal{E}_n(\pi/2) \geq c_2 n\} = O(2^{-n}). \tag{6}$$

3 The largest component in the subcritical interval

3.1 Relation with a typical component

Recall that $I_n(t)$ is the size of the largest component of $\mathcal{G}(n, t)$. Our goal in section 3 is to prove the following bound.

Proposition 3 *Let $\alpha \in (0,\infty)$. Then, for $\beta > 1+2\alpha$,*

$$\mathbf{Pr}\{I_n(t) \le \mu(n,t), \quad \forall\, t < \pi/2 - n^{-\alpha}\} \to 1$$

where

$$\mu(n,t) = 4 + 2\left(\frac{\beta \log n}{\log \frac{\pi}{2t}}\right)^2. \tag{7}$$

Let $C_n(t)$ be the size of the component at time t which contains the first-arriving vertex, conditional on this vertex arriving at time 0. Our first lemma will enable us to study $I_n(t)$ in terms of $C_n(t)$.

Lemma 4

$$\mathbf{Pr}\{I_n(t) > \mu(n,t),\ \textit{for some } t < \pi/2\} \le \frac{n\pi}{2}\,\mathbf{Pr}\{C_n(t) > \mu(n,t),\ \textit{for some } t < \pi/2\}.$$

Proof. For each t, write the components of $\mathcal{G}(n,t)$ as $\{C_n^s(t)\}_{s\le t}$, according to the immigration-times s of the first-appearing vertex of the component. Notice that $C_n^s(t) = \emptyset$ if the component of $\mathcal{G}(n,t)$, that contains s, also contains a vertex which immigrated before s. Let $B_n^s(t)$ denote a maximal connected subgraph of $\mathcal{G}(n,t)$, which contains the vertex s, with a property that no vertex of this subgraph immigrated before s. ($B_n^s(t)$ is the component containing a vertex arriving at s, if we change the rules and forbid such a component to merge with a component whose first vertex arrived earlier.) Clearly $B_n^s(t) = C_n^s(t)$ iff $C_n^s(t) \ne \emptyset$. Importantly, the process $\{B_n^s(t+s)\}_{t\ge 0}$ has the same distribution as $\{C_n(t)\}_{t\ge 0}$. Then

$$\begin{aligned}
& \mathbf{Pr}\{I_n(t) > \mu(n,t),\ \text{for some } t < \pi/2\} \\
= \;& \mathbf{Pr}\{\exists s,t : s \le t < \pi/2,\ |C_n^s(t)| > \mu(n,t)\} \\
\le \;& \mathbf{E}|\{s < \pi/2 : \exists t \in [s,\pi/2),\ |C_n^s(t)| > \mu(n,t)\}| \\
\le \;& \mathbf{E}|\{s < \pi/2 : \exists t \in [s,s+\pi/2),\ |C_n^s(t)| > \mu(n,t-s)\}| \\
= \;& n\int_0^{\pi/2} ds\ \mathbf{Pr}\{\exists t < \pi/2 :\ |C_n^s(t+s)| > \mu(n,t)\} \\
\le \;& n\int_0^{\pi/2} ds\ \mathbf{Pr}\{\exists t < \pi/2 :\ |B_n^s(t+s)| > \mu(n,t)\} \\
= \;& n\int_0^{\pi/2} ds\ \mathbf{Pr}\{\exists t < \pi/2 :\ |C_n(t)| > \mu(n,t)\}.
\end{aligned}$$

■

We turn to the study of $C_n(t)$ in section 3.3, where we bound it by comparison with the process studied next.

3.2 The branching Markov chain

Fix $0 < t < \pi/2$. We study the following discrete-time branching Markov chain (in other words, multitype branching process) on type-space $[0, t]$. An individual of type s has a random number of children of types s', according to a non-homogeneous Poisson process of rate

$$K_t(s, s') = \min(s, s'), \quad 0 \le s' \le t. \tag{8}$$

The process starts with one individual in generation 0, of type t. Let $G_j(t)$ be the number of individuals in generation j, and let $G(t) = \sum_0^\infty G_j(t)$ be the total population. The specific result we need is the tail bound for $G(t)$ provided by Proposition 6 below. As a preliminary, define

$$y(s) = \sin \tfrac{\pi s}{2t}, \quad 0 \le s \le t, \tag{9}$$

where we now drop the subscript t. Then

$$\int_0^t K(s, u) y(u) du = (\tfrac{2t}{\pi})^2 y(s). \tag{10}$$

In other words, $y(\cdot)$ is the first eigenfunction of the integral operator associated with K_t, with corresponding eigenvalue $(\frac{2t}{\pi})^2 < 1$. Identity (10) is classical, e.g. because K_t is the Green function for the differential operator $-y''$, with the boundary conditions $y(0) = 0$, $y'(t) = 0$.

Lemma 5 $\mathbf{Pr}\{G_j(t) > 0\} \le \mathbf{E}G_j(t) \le t(2t/\pi)^{2j-1}$.

Proof. The probabilistic interpretation of the j-fold convolution $K^{(j)}(s, u)$ is that, for a generation-0 individual of type s, the mean number of generation-j descendants of type $[u, u + du]$ equals $K^{(j)}(s, u)du$. Thus by conditioning on the first generation,

$$\mathbf{E}G_j(t) = \int_0^t ds K(t, s) \int_0^t K^{(j-1)}(s, u) du.$$

But by iterating (10) and observing that K is symmetric,

$$\int_0^t y(s) K^{(j-1)}(s, u) ds = (2t/\pi)^{2(j-1)} y(u),$$

and so

$$\int_0^t \int_0^t y(s) K^{(j-1)}(s, u) ds du = (2t/\pi)^{2(j-1)} \int_0^t y(u) du = (2t/\pi)^{2j-1}.$$

So

$$\mathbf{E}G_j(t) \le \sup_{0\le s\le t} \frac{K(t,s)}{y(s)} \left(\frac{2t}{\pi}\right)^{2j-1}$$

and the sup is attained at t. $\blacksquare$

While Lemma 5 implies a tail bound on $G(t)$, what we actually need is the following stronger tail bound.

Proposition 6 *There exists a constant $c_0 > 0$ such that, uniformly for all $0 < t < \pi/2$ and $m > 1$,*

$$\mathbf{Pr}\{G(t) \ge m\} = O\left[\left(\tfrac{2t}{\pi}\right)^{c_0 m^{1/2}}\right].$$

Proof. First note that $G(t)$ is stochastically bounded by the total progeny $Z(t)$, say, of a Galton-Watson branching process for Poisson offspring distribution, with mean

$$\mu = \int_0^t K(t,u)\,du = \frac{t^2}{2}.$$

In particular, $Z(t) < \infty$ almost surely for $t < \sqrt{2}$. From the classical explicit formula

$$\mathbf{Pr}\{Z(t) = z\} = \frac{(\mu z)^{z-1}}{z!} e^{-\mu z},\ z = 1, 2, \ldots,$$

it is easy to show that the stated bound holds uniformly in $0 < t \le 1.4$ and $m > 1$. We then see that, by adjusting the constant c_0, it is enough to obtain the bound for t near $\pi/2$, i.e. uniformly in $\pi/2 - \eta \le t < \pi/2$ and $m > 1$, for some $\eta \in (0, \pi/2)$.

Write $(\xi_i^{(j)}, 1 \le i \le G_j(t))$ for the types of the j'th generation individuals, and write

$$Y_j = \sum_i y(\xi_i^{(j)})$$

for $y(\cdot)$ at (9).

Lemma 7 *If $\varepsilon < \min\,(1, (\frac{\pi}{2t})^2 - 1)$ then $(\exp(\varepsilon Y_j),\ j \ge 0)$ is a supermartingale.*

Proof. Fix $s \in [0,t]$ and let (ξ_i) be the types of the children of a type-s parent. It is enough to show

$$\mathbf{E}\exp\left(\varepsilon \sum_i y(\xi_i)\right) \le \exp(\varepsilon y(s)).$$

But the usual Laplace functional formula for Poisson processes (e.g. [10] Lemma 10.2) shows the left side equals

$$\exp\left(\int_0^t K(s,u)(e^{\varepsilon y(u)} - 1)\, du\right). \tag{11}$$

Since $\varepsilon \leq 1$ and $y(u) \leq 1$,

$$e^{\varepsilon y(u)} - 1 \leq \varepsilon y(u) + (\varepsilon y(u))^2 \leq \varepsilon(1+\varepsilon)y(u).$$

Writing $\gamma = (2t/\pi)^2$ and using (10), the quantity (11) is at most

$$\exp(\varepsilon(1+\varepsilon)\gamma y(s)).$$

So we need only check $(1+\varepsilon)\gamma \leq 1$, and this holds by the hypothesis on ε. ■

Because $\sin x \geq \frac{x}{\pi/2}$ on $0 \leq x \leq \pi/2$ we have

$$y(u) = \sin(\pi u/(2t)) \geq u/t$$

and so

$$Y_j = \sum_i y(\xi_i^{(j)}) \geq t^{-1} \sum_i \xi_i^{(j)}.$$

Now, the conditional distribution of $G_{j+1}(t)$ (the total size of the $(j+1)$-th generation) given the first j generations is Poisson with parameter

$$\Lambda_{j+1} = \sum_i \int_0^t \min(\xi_i^{(j)}, u)\, du \leq t \sum_i \xi_i^{(j)}.$$

From the previous inequality we see $\Lambda_{j+1} \leq t^2 Y_j$. Applying the optional stopping theorem for the supermartingale $\{\exp(\varepsilon Y_j)\}$ we find, for arbitrary $\lambda > 0$,

$$\begin{aligned}
\mathbf{Pr}\{\sup_{j\geq 0} \Lambda_{j+1} > \lambda\} &\leq \mathbf{Pr}\left\{\sup_{j\geq 0} \exp(\varepsilon Y_j) > \exp(\varepsilon\lambda t^{-2})\right\} \\
&\leq \frac{\mathbf{E}\exp(\varepsilon Y_0)}{\exp(\varepsilon\lambda t^{-2})} \\
&= \exp(\varepsilon(1-\lambda t^{-2}))
\end{aligned} \tag{12}$$

provided ε satisfies the hypothesis of Lemma 7. Now consider integers $k, m \geq 1$. Then by Lemma 7 and (12)

$$\mathbf{Pr}\{G(t) \geq m+1\} \tag{13}$$

$$\leq \mathbf{Pr}\{G_{k+1}(t) > 0\} + \mathbf{Pr}\{\sum_{j=1}^{k} G_j(t) \geq m\} \tag{14}$$

$$\leq \mathbf{Pr}\{G_{k+1}(t) > 0\} + \mathbf{Pr}\{\sup_{j\geq 0} \Lambda_{j+1} > \lambda\} + \mathbf{Pr}\{\sum_{j=1}^{k} \mathcal{P}_j(\lambda) \geq m\}$$

where the $\mathcal{P}_j(\lambda)$ are independent with Poisson (mean λ) law

$$\leq t(2t/\pi)^{2k+1} + \exp(\varepsilon(1-\lambda t^{-2})) + \mathbf{Pr}\{\mathcal{P}(k\lambda) \geq m\} \tag{15}$$

where $\mathcal{P}(k\lambda)$ has Poisson (mean $k\lambda$) law. To complete the proof of Proposition 6 it is enough to show that, choosing

$$k = \lfloor m^{1/2} \rfloor, \quad \lambda = \delta k,$$

for suitably small δ, each of the three terms in (15) is $O((2t/\pi)^{c_0 m^{1/2}})$ uniformly in $m \geq 1$ and $t \in (\pi/2 - \eta, \pi/2)$, for suitable c_0 and η. For the first term this is clear. For the second term, the issue is to get ε such that

$$\exp(-\varepsilon) \leq (2t/\pi)^c, \quad \pi/2 - \eta \leq t < \pi/2,$$

and this holds (for sufficiently small c and η) by the hypothesis on ε in Lemma 7. The third term is at most $\mathbf{Pr}\{\mathcal{P}(\delta m^{1/2}) \geq m\}$ and by routine large deviation bounds this is at most $(1/2)^m$ when δ is sufficiently small. ■

3.3 The comparison argument

Return to the setting of section 3.1, where we considered the RGIV process conditioned on the first arrival (vertex v_*, say) being at time 0. Fix t. The component at time t containing v_* is not necessarily a tree, but we can consider the "breadth-first spanning tree" $\mathcal{T}(n,t)$ in which we first include edges from v_* to each neighbor $v_1, v_2, \ldots$, then include edges from v_1 to its neighbors not already present, then include edges from v_2 to its neighbors not already present, and so on. So $C_n(t)$ is the number of vertices in $\mathcal{T}(n,t)$. Say that a vertex which arrived at time s has *type* $t-s$.

Lemma 8 *Fix $0 < t < \pi/2$. We can construct jointly $\mathcal{T}(n,t)$ and the branching Markov chain ($\mathcal{T}^*(t)$, say) such that $\mathcal{T}(n,t)$ is a subtree of $\mathcal{T}^*(t)$. In particular, by Proposition 6*

$$\mathbf{Pr}\{C_n(t) \geq m\} = O\left[\left(\tfrac{2t}{\pi}\right)^{c_0 m^{1/2}}\right]$$

uniformly in $0 < t < \pi/2$ and $m > 1$.

Proof. In $\mathcal{T}(n,t)$ the types of the children of the root vertex v_* form a Poisson process of rate $\rho(s)$, where

$$\rho(s)ds = n\ ds\ \cdot\ (1-\exp(-s/n)),\ 0 \le s \le t.$$

The first term is the arrival rate of all vertices, and the second term is the chance that a type-s vertex (arriving at time $t-s$) forms at least one edge to v_* before time t. Notice that $\rho(s) \le s = K(t,s)$. So we can obtain this Poisson $(\rho(s))$ process from the Poisson $(K(t,s))$ process of children of the root of $\mathcal{T}^*(t)$ by retaining or deleting a type-s child of the latter process with probability $\rho(s)/K(t,s)$ or $1-\rho(s)/K(t,s)$, independently of other children. If we implement recursively this "thinning-out" procedure for surviving members of a current generation of the tree $\mathcal{T}^*(t)$, the resulting subtree will have the same distribution as $\mathcal{T}(n,t)$. ■

3.4 Proof of Proposition 3

Pick $\beta > 0$, and set

$$m(n,t) = 2 + \lfloor \left(\beta \frac{\log n}{c_0 \log \frac{\pi}{2t}}\right)^2 \rfloor.$$

Then Lemma 8 implies

$$\mathbf{Pr}\{C_n(t) \ge m(n,t)\} = O(n^{-\beta}). \tag{16}$$

We shall use (16) to show

$$\mathbf{Pr}\{C_n(t) > 2m(n,t),\ \text{for some}\ t < \pi/2 - n^{-\alpha}\} = O(n^{-(\beta-1-2\alpha)}\log^2 n), \tag{17}$$

and then Proposition 3 follows from Lemma 4 if we select $\beta > 1+2\alpha$.

To argue (17), first fix $0 < s < t < \pi/2$. Let $q(s,t)dsdt$ be the chance that

(i) At time t there is some component (of size $C_n^s(t)$, say) whose chronologically first vertex immigrated during $[s, s+ds]$;
(ii) $\max(C_n(t), C_n^s(t)) > m(n,t)$;
(iii) $C_n(t) \le 2m(n,t)$;
(iv) the components of sizes $C_n(t)$ and $C_n^s(t)$ merge during $[t, t+dt]$;
(v) $V_n(\pi/2) \le n\pi$.

We claim

$$q(s,t) \le n \times 2\mathbf{Pr}\{C_n(t) > m(n,t)\} \times n^{-1} \cdot 2m(n,t) \cdot n\pi. \tag{18}$$

Here the first term is immigration rate at time s. The second term is a bound on the chance of (ii), since (as in the proof of Lemma 4) $C_n^s(t)$ is stochastically smaller than $C_n(t-s)$. The third term bounds the merger rate (iv); conditionally, the rate is $n^{-1}C_n(t)C_n^s(t)$, and the second and third factors are then bounded using (iii) and (v).

Now let τ be the first moment t such that $C_n(t) > 2m(n,t)$. At τ the component which contains the vertex which arrived at time 0 merges with another component, whose first vertex immigrated at a time $s < \tau$, and these components must satisfy conditions (i)-(iii) above. It follows that

$$\mathbf{Pr}\{\tau \in [t, t+dt], V_n(\pi/2) \le n\pi\} \le dt \int_0^t q(s,t)\, ds.$$

And so

$$\begin{aligned}\mathbf{Pr}\{C_n(t) > 2m(n,t), \text{ for some } t < \pi/2 - n^{-\alpha}\} &= \mathbf{Pr}\{\tau < \pi/2 - n^{-\alpha}\} \\ &\le \int\int_{0\le s\le t\le \pi/2-n^{-\alpha}} q(s,t)\, dsdt \quad + \mathbf{Pr}\{V_n(\pi/2) > n\pi\}.\end{aligned}$$

The second term is exponentially small. The integrand in the first term is $O(n^{-\beta+1}m(n,t)) = O(n^{-(\beta-1-2\alpha)}\log^2 n)$ by (18) and (16). ■

4 Martingale analysis of the RGIV process

Introduce the discrete time moments τ_m, $m \ge 1$, when something happens (either a vertex immigrates or an edge appears) in the RGIV process. Let G_m be the random multigraph immediately after the moment τ_m. An expression "at time m" will be used instead of "at time τ_m". Obviously, m equals the total number of vertices and arcs in the multigraph G_m. Write $V_m = V(G_m)$ for the total number of vertices, and write X_m, Y_m for the sum of squared and the sum of cubed component sizes in the graph G_m. So $X_m = X_n(\tau_m), Y_m = Y_n(\tau_m)$, and $I_m = I_n(\tau_m)$. Our analysis rests upon the following formulas.

Lemma 9 *Denoting* $\mathbf{E}(\{\}|(G_m,\tau_m))$ *by* $\mathbf{E}(\{\}|\bullet)$, *we have*
(i) $\mathbf{E}(V_{m+1} - V_m|\bullet) = \frac{1}{1+(V_m/n)^2/2}$
(ii) $\mathbf{E}(X_{m+1} - X_m|\bullet) = \frac{1+(X_m/n)^2 - n^{-1}(Z_m/n)}{1+(V_m/n)^2/2}$
(iii) $\mathbf{E}(Y_{m+1} - Y_m|\bullet) = \frac{1+3(X_m/n)(Y_m/n) - 3n^{-1}(U_m/n)}{1+(V_m/n)^2/2}$
where Z_m and U_m are the sums of the fourth powers and the fifth powers of the component sizes for the graph G_m.

Proof. At time m the total current rate of arc formation is $V_m^2/(2n)$ and the rate of vertex immigration is n.

So the probability that the next event is "a vertex immigrates" equals

$$\frac{n}{n+V_m^2/(2n)} = \frac{1}{1+(V_m/n)^2/2}.$$

This is (i). Further, let $\ell = \ell(G_m)$ denote the total number of components of G_m, and let $i_1, \ldots, i_\ell$ denote their sizes. Then

$$\begin{aligned}
\mathbf{E}(X_{m+1} - X_m|\bullet) &= \frac{n}{n+V_m^2/(2n)} \cdot 1^2 + \frac{1/(2n)}{n+V_m^2/(2n)} \sum_{1\le\alpha\neq\beta\le\ell} i_\alpha i_\beta (2i_\alpha i_\beta) \\
&= \frac{1}{1+(V_m/n)^2/2} + \frac{1/(2n^2)}{1+(V_m/n)^2/2}\left(2X_m^2 - 2\sum_{\alpha=1}^{\ell} i_\alpha^4\right),
\end{aligned}$$

which leads to (ii). The proof of (iii) follows along similar lines. ■

Motivated by Lemma 9, we introduce the initial value problem: solve

$$\frac{dv}{dt} = \frac{1}{1+v^2/2}, \quad \frac{dx}{dt} = \frac{1+x^2}{1+v^2/2}, \quad \frac{dy}{dt} = \frac{1+3xy}{1+v^2/2} \tag{19}$$

subject to $v(0) = x(0) = y(0) = 0$. Its (parametric) solution is given by

$$x = \tan v, \quad y = \sec^3 v \left(\frac{1}{12}\sin 3v + \frac{3}{4}\sin v\right). \tag{20}$$

This is obtained by integrating

$$\frac{dx}{dv} = 1 + x^2, \quad x(0) = 0, \tag{21}$$

and then integrating

$$\frac{dy}{dv} = 1 + 3xy = 1 + 3y\tan v, \quad y(0) = 0. \tag{22}$$

The system (19) has two integrals. From (21) $\int \frac{dx}{1+x^2} = \int dv$, that is

$$J(x,v) := \arctan x - v \equiv \text{ const.} \tag{23}$$

From the last two equations in (19)

$$\frac{dy}{dx} = \frac{1+3xy}{1+x^2},$$

which has a general solution

$$y = x + \tfrac{2}{3}x^3 + \text{ const } (1+x^2)^{3/2}.$$

That is,

$$H(x,y) := \frac{y - x - \frac{2}{3}x^3}{(1+x^2)^{3/2}} \equiv \text{ const.} \tag{24}$$

Since $J(x,v)$, $H(x,y)$ are kept constant along every trajectory of (19), the gradients **grad** J and **grad** H are orthogonal to the vector formed by the right hand expressions in (19). The proof of Proposition 2 consists of showing that $J(X_m/n, V_m/n)$ and $H(X_m/n, Y_m/n)$ remain nearly constant along most realizations of the process $\{G_m\}$, at least as long as $\tau_m \le \frac{\pi}{2} - n^{-\alpha}$, $\alpha < 1/6$. The proof occupies the remainder of section 4.

4.1 Analysis of J

Fix

$$1/9 < \alpha < 1/6, \quad 1/3 < \delta < \beta < 1/2. \tag{25}$$

Recall from Proposition 3 the definition of $\mu(n,\tau)$.

Lemma 10 *With high probability, for all m such that $\tau_m \le \frac{\pi}{2} - n^{-\alpha}$,*
(i) $\left|\frac{V_m}{n} - \tau_m\right| \le n^{-\beta}$
(ii) $I_m \le \mu(n, \tau_m)$
(iii) $|J(X_m/n, V_m/n)| \le n^{-\delta}$.

Proof. Proposition 3 verifies (ii). The functional CLT for the rate-n Poisson process $V(\cdot)$ counting number of vertices of $\mathcal{G}(n,t)$ implies

$$\sup_{s \le \pi/2} \left| \frac{V(s)}{n} - s \right| = O_p(n^{-1/2})$$

establishing (i).

Let T^0 be the stopping time defined by

$$T^0 := \min\{m : \tau_m > \tfrac{\pi}{2} - n^{-\alpha} \text{ or (i) or (ii) or (iii) fails at } m\}.$$

To complete the proof by verifying (iii), we need to show

$$\mathbf{Pr}\{|J_{T^0}| > n^{-\delta}\} \to 0. \tag{26}$$

Introduce the random sequence $J_m = J(X_m/n, V_m/n)$, where $J(x,v) = \arctan x - v$ as at (23). Let us estimate $\mathbf{E}(J_{m+1}^2|\bullet)$, assuming that $m < T^0$. First we write

$$J_{m+1}^2 - J_m^2 = (\mathbf{grad}\ J_m^2)^* \Delta_m + O(R_m) \tag{27}$$

where

$$\begin{aligned} \Delta_m &= \left(\frac{X_{m+1} - X_m}{n}, \frac{V_{m+1} - V_m}{n} \right), \\ R_m &= |J_{xx}^2(\tilde{x}, \tilde{v})| n^{-2} (X_{m+1} - X_m)^2 \\ &\quad + |J_{xv}^2(\tilde{x}, \tilde{v})| n^{-2} (X_{m+1} - X_m) \cdot (V_{m+1} - V_m) \\ &\quad + |J_{vv}^2(\tilde{x}, \tilde{v})| n^{-2} (V_{m+1} - V_m)^2, \end{aligned} \tag{28}$$

for some $(\tilde{x}, \tilde{v})$ on the line segment joining $(X_m/n, V_m/n)$ and $(X_{m+1}/n, V_{m+1}/n)$. Here **grad** J_m^2 means "**grad** J^2 evaluated at $(X_m/n, V_m/n)$". Now **grad** $J = (\frac{1}{1+x^2}, -1)$ and so

$$\mathbf{grad}\ J^2 = \tfrac{2J}{1+x^2}(1, -(1+x^2)).$$

So, consulting equations (i) and (ii) of Lemma 9, we see that

$$(\mathbf{grad}\ J_m^2)^* \mathbf{E}(\Delta_m|\bullet) = \frac{2J_m}{1 + (X_m/n)^2}\ \frac{1}{n}\ \frac{(-n^{-1}) Z_m(n)/n}{1 + (V_m/n)^2/2},$$

where Z_m is the sum of component sizes raised to the fourth power. We need to get explicit bounds for the last expression and for $\mathbf{E}(R_m|\bullet)$. Clearly $Z_m \le I_m^2 X_m$ and $|J_m| \le n^{-\delta}$ by condition (iii), since we work on $\{m < T^0\}$. So

$$\left| (\mathbf{grad}\ J_m^2)^* \mathbf{E}(\Delta_m|\bullet) \right| \le_b n^{-(2+\delta)} \frac{I_m^2 X_m/n}{1 + (X_m/n)^2}. \tag{29}$$

We want to establish bound (31) below. If $I_m \le C \log^2 n$ for large C then a better bound than (31) follows from (29) and boundedness of the function $x \to \frac{x}{1+x^2}$. So suppose $I_m > C \log^2 n$. By condition (ii) and the definition of $\mu(n,t)$ in Proposition 3, on $\{m < T^0\}$ we have

$$I_m \le \mu(n, \tau_m) \le_b \max(1, (\tfrac{\pi}{2} - \tau_m)^{-2} \log^2 n).$$

The only way it's possible to have $I_m > C \log^2 n$ is if the max is attained by the second argument, in which case the bound (29) becomes

$$\left| (\mathbf{grad}\ J_m^2)^* \mathbf{E}(\Delta_m|\bullet) \right| \le_b n^{-(2+\delta)} \frac{1}{X_m/n} (\tfrac{\pi}{2} - \tau_m)^{-4} \log^4 n. \tag{30}$$

In addition,

$$\begin{aligned} X_m/n &= \tan(V_m/n + O(n^{-\delta})) \text{ by condition (iii)} \\ &= \tan(\tau_m + O(n^{-\delta} + n^{-\beta})) \text{ by condition (i)} \\ &= \tan(\tau_m + O(n^{-\delta})) \quad (\delta < \beta) \\ &= (1 + o(1)) \tan \tau_m = (1 + o(1))(\tfrac{\pi}{2} - \tau_m)^{-1} \end{aligned}$$

the penultimate equality holding because $\delta > \alpha$ and, on $\{m < T^0\}$, $\frac{\pi}{2} - \tau_m > n^{-\alpha}$. Substituting into (30) gives

$$|(\mathbf{grad}\ J^2_m)^* \mathbf{E}(\Delta_m | \bullet)| \leq_b n^{-(2+\delta-3\alpha)} \log^4 n. \tag{31}$$

Next consider R_m at (28). First note

$$\frac{X_{m+1}}{n} - \frac{X_m}{n} \leq \frac{2I^2_n}{n}, \quad \frac{V_{m+1}}{n} - \frac{V_m}{n} \leq \frac{1}{n}. \tag{32}$$

From $J^2_x = 2J/(1+x^2)$ we calculate

$$J^2_{xx}(x,v) = \frac{2 - 4xJ}{(1+x^2)^2}, \ J^2_{xv}(x,v) = \frac{-2}{1+x^2}, \ J^2_{vv}(x,v) = 2. \tag{33}$$

For $x \in [X_m/n, X_{m+1}/n]$ and $v \in [V_m/n, V_{m+1}/n]$,

$$\begin{aligned} |J(x,v) - J(X_m/n, V_m/n)| &\leq n^{-1} + (X_{m+1}/n - X_m/n) J_x(X_m/n, V_m/n) \\ &\leq n^{-1} + \frac{2I^2_m/n}{1 + (X_m/n)^2}, \quad \text{(by (32))} \end{aligned}$$

and using $|J(X_m/n, V_m/n)| \leq n^{-\delta}$,

$$|J(x,v)| \leq_b n^{-\delta} + \frac{I^2_m/n}{1+(X_m/n)^2}. \tag{34}$$

Combining (28,32,33),

$$R_m \leq_b \frac{|J(\tilde{x},\tilde{v})| X_m/n \ + 1}{(1 + X^2_m/n^2)^2} \frac{I^4_m}{n^2} + \frac{1}{1 + X^2_m/n^2} \frac{I^2_m}{n^2} + \frac{1}{n^2} \tag{35}$$

We want to show

$$R_m \leq_b n^{-2} \log^8 n. \tag{36}$$

If $I_m \leq C \log^2 n$ for large C then (36) follows from (35) and boundedness of J. If $I_m \geq C \log^2 n$, we know already that $I_m = O((\frac{\pi}{2} - \tau_m)^{-1} \log n)^2$ and that $X_m/n = (1+o(1))(\frac{\pi}{2}-\tau_m)^{-1}$. Plugging these estimates and (34) into (35), we get

$$\begin{aligned} R_m \quad \leq_b \quad & \left[(\tfrac{\pi}{2}-\tau_m)^3\left(n^{-\delta}+\frac{\log^4 n}{n(\frac{\pi}{2}-\tau_m)^2}\right)+(\tfrac{\pi}{2}-\tau_m)^4\right] \cdot \frac{\log^8 n}{n^2(\frac{\pi}{2}-\tau_m)^4} \\ & +\frac{\log^4 n}{n^2}+\frac{1}{n^2}. \end{aligned}$$

The dominant term in the big square brackets is $(\frac{\pi}{2}-\tau_m)^4$, because $\frac{\pi}{2}-\tau_m \geq n^{-\alpha}$ with $\alpha < 1/3 < \delta$. Thus the bound (36) holds also for $I_m \geq C\log^2 n$.

Combining (27,31,36) we see there is a constant c such that

$$\mathbf{E}(J_{m+1}^2 - J_m^2|\bullet) \leq cn^{-(1+\omega_1)}\log^8 n,\ \ 0 \leq m < T^0$$

where

$$\omega_1 := \min\{1+\delta-3\alpha, 1\} > 0.$$

Thus, the sequence $\{J_m^2 - cmn^{-(1+\omega_1)}\log^8 n,\ 0 \leq m \leq T^0\}$ is a supermartingale, whose value is zero at $m=0$. By (5,6)

$$\mathbf{Pr}\{T^0 > (c_1+c_2)n\} = O(2^{-n}).$$

Applying the optional stopping theorem for $\min(T^0,(c_1+c_2)n)$,

$$\mathbf{E}(J^2_{\min(T^0,(c_1+c_2)n)}) \leq cn^{-(1+\omega_1)}\log^8 n \cdot \mathbf{E}\min(T^0,(c_1+c_2)n) \leq_b n^{-\omega_1}\log^8 n.$$

Consequently,

$$\begin{aligned} \mathbf{Pr}(|J_{T^0}| > n^{-\delta}) \quad &\leq \quad n^{-2\delta}\mathbf{E}(J_{T^0}^2) + O(2^{-n}) \leq_b n^{-\lambda_1}\log^8 n, \\ \lambda_1 \quad &= \quad \min\{1-\delta-3\alpha, 1-2\delta\} > 0. \end{aligned}$$

Because $\alpha < 1/6$, this inequality establishes (26) and completes the proof of Lemma 10. ■

Proof of assertions (3, 4) of Proposition 2. Recall $t_n = \frac{\pi}{2} - n^{-\alpha}$. Define T^* by:

$$\tau_{T^*-1} < t_n \leq \tau_{T^*}.$$

It is easy to check that

$$\tau_{T^*} - t_n = O_p(1/n), \quad t_n - \tau_{T^*-1} = O_p(1/n).$$

By Lemma 10, whp

$$\begin{aligned}
X(t_n)/n &= X_{T^*-1}/n \\
&= \tan(V_{T^*-1}/n + O(n^{-\delta})) \text{ by (iii)} \\
&= \tan(\tau_{T^*-1} + O(n^{-\delta} + n^{-\beta})) \text{ by (i)} \\
&= \tan(\tfrac{\pi}{2} - n^{-\alpha} + O(n^{-\delta})). \qquad (37)
\end{aligned}$$

So whp

$$\begin{aligned}
\frac{n}{X(t_n)} &= \tan(n^{-\alpha} + O(n^{-\delta})) \\
&= n^{-\alpha} + O(n^{-\delta}) + O(n^{-3\alpha}) \\
&= n^{-\alpha} + o(n^{-1/3}) \qquad (38)
\end{aligned}$$

establishing (3). Moreover by (ii), whp

$$I(t_n) \leq \mu(n, t_n) = O(n^{2\alpha} \log^2 n)$$

and so using (38), whp

$$\begin{aligned}
\frac{I(t_n)}{X(t_n)} &= O\left(\frac{n^{2\alpha}\log^2 n}{n^{1+\alpha}}\right) \\
&= O(n^{\alpha-1}\log^2 n) \\
&= o(n^{-2/3})
\end{aligned}$$

establishing (4). ■

4.2 Analysis of H

We complete the proof of Proposition 2 by showing that $H(X_m/n, Y_m/n)$ also remains nearly constant along most realizations of the process $\{G_m\}$. Recall $1/9 < \alpha < 1/6$ and fix $0 < \sigma < 1/2$.

Lemma 11 *With high probability, for all m such that $\tau_m \leq \frac{\pi}{2} - n^{-\alpha}$,*
(iv) $|H(X_m/n, Y_m/n)| \leq n^{-\sigma}$.

Proof. The argument follows closely the structure of the proof of Lemma 10. Recall

$$H(x, y) := \frac{y - x - \frac{2}{3}x^3}{(1 + x^2)^{3/2}}.$$

Write $H_m = H(X_m/n, Y_m/n)$. Let T^1 be the stopping time defined as the minimum of T^0 and

$$\min\{m : \tau_m > \tfrac{\pi}{2} - n^{-\alpha} \text{ or (iv) fails at } m\}.$$

We need to show

$$\mathbf{Pr}\{|H_{T^1}| > n^{-\sigma}\} \to 0. \tag{39}$$

As at (27,28),

$$H^2_{m+1} - H^2_m = (\mathbf{grad}\ H^2_m)^* \Delta_m + O(R_m), \tag{40}$$

where

$$\begin{aligned}
\Delta_m &= \left(\frac{X_{m+1} - X_m}{n}, \frac{Y_{m+1} - Y_m}{n} \right), \\
R_m &= |H^2_{xx}(\tilde{x}, \tilde{y})| n^{-2} (X_{m+1} - X_m)^2 \\
&\quad + |H^2_{xy}(\tilde{x}, \tilde{y})| n^{-2} (X_{m+1} - X_m) \cdot (Y_{m+1} - Y_m) \\
&\quad + |H^2_{yy}(\tilde{x}, \tilde{y})| n^{-2} (Y_{m+1} - Y_m)^2,
\end{aligned} \tag{41}$$

for some $(\tilde{x}, \tilde{y})$ on the line segment joining $(X_m/n, Y_m/n)$ and $(X_{m+1}/n, Y_{m+1}/n)$. We calculate

$$\mathbf{grad}\ H^2 = \frac{2H}{(1+x^2)^{5/2}} \cdot (-(1+3xy), 1+x^2).$$

Using Lemma 9(ii,iii)

$$\begin{aligned}
(\mathbf{grad}\ H^2_m)^* \mathbf{E}(\Delta_m | \bullet) &= \frac{2H_m}{(1+(X_m/n)^2)^{5/2}} \frac{n^{-2}}{1+(V_m/n)^2/2} \\
&\quad \cdot \Big((1+3(X_m/n)(Y_m/n)) Z_m/n \ - \ (1+(X_m/n)^2)(3U_m/n) \Big).
\end{aligned}$$

We seek to bound this on $\{m < T^1\}$, where by definition $|H_m| \le n^{-\sigma}$. Since

$$Y_m \le I_m X_m, \quad Z_m \le I^2_m X_m, \quad U_m \le I^3_m X_m$$

we find

$$|(\mathbf{grad}\ H^2_m)^* \mathbf{E}(\Delta_m | \bullet)| \le_b n^{-(2+\sigma)} I^3_m \frac{(X_m/n)^3}{(1+(X_m/n)^2)^{5/2}}. \tag{42}$$

Now repeat the argument that led from (29) to (31), considering separately the cases $I_m \le C \log^2 n$ and $I_m \ge C \log^2 n$, and noting that we may assume that conditions (i,ii,iii) of Lemma 10 hold because $m < T^1 \le T^0$. We conclude

$$|(\mathbf{grad}\ H^2_m)^* \mathbf{E}(\Delta_m|\bullet)| \le_b n^{-(2+\sigma-4\alpha)} \log^6 n. \tag{43}$$

Next we follow the line of argument which led from (32) to (36), via bounding H^2_{xx}, H^2_{xy} and H^2_{yy} at an intermediate point $(\tilde{x}, \tilde{y})$, to obtain

$$R_m \le_b n^{-(2-6\alpha)} \log^{12} n.$$

Plugging this and (43) into (40),

$$\mathbf{E}(H^2_{m+1} - H^2_m|\bullet) \le cn^{-(2-6\alpha)} \log^{12} n.$$

As in the previous section, the optional stopping theorem for the supermartingale $\{H^2_m - cmn^{-(2-6\alpha)} \log^{12} n, 0 \le m \le T^1\}$ implies

$$\mathbf{E}(H^2_{T^1}) \le cn^{-(2-6\alpha)} \log^{12} n \cdot ET^1 \le_b n^{-(1-6\alpha)} \log^{12} n.$$

Since $\alpha > 1/6$ this establishes (39) and completes the proof of Lemma 11.

Proof of assertion (2) of Proposition 2. Lemma 11 implies that whp

$$|H(X(t_n)/n, Y(t_n)/n)| \le n^{-\sigma}.$$

Rearranging the definition of H,

$$\frac{y}{x^3} - \frac{2}{3} = x^{-2} + x^{-3}(1+x^2)^{3/2} H(x,y).$$

By (38) $X(t_n)/n = (1+o(1))n^\alpha$ whp, so whp

$$\begin{aligned} \frac{n^2 Y(t_n)}{X^3(t_n)} - \frac{2}{3} &= \frac{Y(t_n)/n}{(X(t_n)/n)^3} \\ &= O(n^{-2\alpha}) + O(n^{-\sigma}) \\ &= o(1). \end{aligned}$$

This establishes (2) and completes the proof of Proposition 2.

5 Proof of Theorem 1

Recall that Theorem 1 features the rescaled process

$$\bar{\mathbf{C}}^{(n)}(t) = (2/3)^{1/3} n^{-2/3} \mathbf{C}^{(n)}(\pi/2 + (2/3)^{2/3} t n^{-1/3}), \quad (-\infty < t < \infty), \tag{44}$$

which traverses the critical time interval as t increases from $-\infty$ to ∞. The time t_n used in Proposition 2 corresponds to rescaled time $\mathbf{t}_n$, where

$$\tfrac{\pi}{2} - n^{-\alpha} = \tfrac{\pi}{2} + (2/3)^{2/3}\mathbf{t}_n n^{-1/3};$$

in other words

$$\mathbf{t}_n = -(3/2)^{2/3} n^{-\alpha+1/3}.$$

Proposition 2 is equivalent to the following assertions about $\bar{\mathbf{C}}^{(n)}(\mathbf{t}_n)$. The factors $(2/3)$ in the rescaling arise as the rescalings which make the limit in the first assertion be 1.

Proposition 12 *Let $1/9 < \alpha < 1/6$. Then*

$$\frac{\sum_i \left(\bar{C}_i^{(n)}(\mathbf{t}_n)\right)^3}{\left(\sum_i \left(\bar{C}_i^{(n)}(\mathbf{t}_n)\right)^2\right)^3} \xrightarrow{p} 1.$$

$$\frac{1}{\sum_i \left(\bar{C}_i^{(n)}(\mathbf{t}_n)\right)^2} + \mathbf{t}_n \xrightarrow{p} 0.$$

$$\frac{\bar{C}_1^{(n)}(\mathbf{t}_n)}{\sum_i \left(\bar{C}_i^{(n)}(\mathbf{t}_n)\right)^2} \xrightarrow{p} 0.$$

Now let $(\mathcal{G}_-(n,t))_{t\geq 0}$ be the modification of the RGIV process in which vertex creation is stopped at time t_n, but edges continue to form thereafter, with the same rate n^{-1}. Let $\mathbf{C}_-^{(n)}(t)$ be the component sizes of $\mathcal{G}_-(n,t)$, and let $(\bar{\mathbf{C}}_-^{(n)}(t), -\infty < t < \infty)$ be the process derived from $\mathbf{C}_-^{(n)}(t)$ by rescaling as at (44). Then $(\bar{\mathbf{C}}_-^{(n)}(t))_{t\geq \mathbf{t}_n}$ is the multiplicative coalescent started at time $\mathbf{t}_n$ from state $\mathbf{C}_-^{(n)}(\mathbf{t}_n) := \bar{\mathbf{C}}^{(n)}(\mathbf{t}_n)$. The assertions of Proposition 12 are exactly the hypotheses of Proposition 4 of Aldous [1], whose conclusion is

$$\bar{\mathbf{C}}_-^{(n)} \to_d \mathbf{X}, \tag{45}$$

in the sense of weak convergence on $D((-\infty,\infty), l^2_\downarrow)$. For simplicity, let's consider convergence at fixed (rescaled) time 0. Specializing (45) gives

Corollary 13

$$\bar{\mathbf{C}}_-^{(n)}(0) \to_d \mathbf{X}(0),$$

in the sense of weak convergence on $l^2_\downarrow$.

Returning to the original RGIV process, the number of vertices which immigrate during the time interval $[t_n = \pi/2 - n^{-\alpha}, \pi/2]$ has Poisson (mean $n^{1-\alpha}$) law, independent of $\mathbf{C}_-^{(n)}(t_n)$. Let $(\mathcal{G}_+(n,t))_{t\geq 0}$ be the modification of the RGIV process in which these vertices all appear at time t_n and there is no subsequent vertex creation, but edges continue to form thereafter. Let $\mathbf{C}_+^{(n)}(t)$ be the component sizes of $\mathcal{G}_+(n,t)$, and let $(\bar{\mathbf{C}}_+^{(n)}(t), -\infty < t < \infty)$ be the process derived from $\mathbf{C}_+^{(n)}(t)$ by rescaling as at (44).

Lemma 14 $\mathbf{C}^{+(n)}(0) \to_d \mathbf{X}(0)$.

Proof. Just as in the proof of Corollary 13, it is enough to show that the assertions of Proposition 12 remain true when we append to $\bar{\mathbf{C}}^{(n)}(\mathbf{t}_n)$ a Poisson$(n^{1-\alpha})$ number of clusters, each of mass $(2/3)^{1/3}n^{-2/3}$. The first and third assertions are easy to check. For the second assertion, write $S(n) = \sum_i \left(\bar{C}_i^{(n)}(\mathbf{t}_n)\right)^2$ and $D(n)$ for the sum of squares of masses of the extra clusters. We need to show

$$\frac{1}{S(n)+D(n)} - \frac{1}{S(n)} \ \stackrel{p}{\to}\ 0,$$

which reduces to showing

$$D(n)/(S^2(n)) \ \stackrel{p}{\to}\ 0.$$

Now

$$D(n) = O_p(n^{1-\alpha}n^{-4/3}) = O_p(n^{-1/3-\alpha})$$

and by the second assertion of Proposition 12, whp

$$n^{\alpha-1/3} = O(1/\mathbf{t}_n) = O_p(S(n)).$$

So what we need is $-1/3-\alpha < 2(\alpha - 1/3)$, which reduces to the assumption $\alpha > 1/9$ in Proposition 12. ■

We digress to recall ([3] section 5) that there is a natural partial order $\preceq$ on $l^2_\downarrow$. Informally, $\mathbf{x} \preceq \mathbf{y}$ if $\mathbf{y}$ can be obtained from $\mathbf{x}$ by coalescing together clusters of $\mathbf{x}$ and appending new clusters. Lemma 17 of [1] says:

$$\text{if } \mathbf{x} \preceq \mathbf{y} \text{ then } 0 \leq d^2(\mathbf{x},\mathbf{y}) \leq \sum_i y_i^2 \ - \ \sum_i x_i^2 \qquad (46)$$

where d is distance in $l^2_\downarrow$.

Lemma 15 *If random elements* $\mathbf{Y}^{(n)}, \mathbf{Y}^{*(n)}, \mathbf{Y}^{**(n)}$ *of* $l^2_\downarrow$ *satisfy*
(a) $\mathbf{Y}^{(n)} \preceq \mathbf{Y}^{*(n)} \preceq \mathbf{Y}^{**(n)}$
(b) $\mathbf{Y}^{(n)} \stackrel{d}{\to} \mathbf{Y}$ *and* $\mathbf{Y}^{**(n)} \stackrel{d}{\to} \mathbf{Y}$
then $\mathbf{Y}^{*(n)} \stackrel{d}{\to} \mathbf{Y}$.

Proof. By (a) and (46)

$$\begin{aligned} d^2(\mathbf{Y}^{(n)}, \mathbf{Y}^{*(n)}) &\leq \sum_i (Y_i^{*(n)})^2 - \sum_i (Y_i^{(n)})^2 \\ &\leq \sum_i (Y_i^{**(n)})^2 - \sum_i (Y_i^{(n)})^2 \end{aligned} \tag{47}$$

By (b) $\sum_i (Y_i^{(n)})^2 \stackrel{d}{\to} \sum_i Y_i^2$ and $\sum_i (Y_i^{**(n)})^2 \stackrel{d}{\to} \sum_i Y_i^2$. The fact that the right side of (47) is non-negative and is the difference of two terms converging in distribution to the same limit implies (as an easy fact about real-valued random variables) that the right side of (47) tends to 0 in probability. The conclusion of the lemma follows. ■

Returning to the processes defined above Lemma 14, clearly there is a natural coupling of the three processes for which

$$\bar{\mathbf{C}}^{(n)}_-(0) \preceq \bar{\mathbf{C}}^{(n)}(0) \preceq \bar{\mathbf{C}}^{(n)}_+(0). \tag{48}$$

Corollary 13 and Lemma 14 show that the left and right quantities converge in law to the same limit $\mathbf{X}(0)$. Then Lemma 15 implies

$$\bar{\mathbf{C}}^{(n)}(0) \to_d \mathbf{X}(0).$$

Essentially the same argument shows that for any fixed t_0

$$\bar{\mathbf{C}}^{(n)}(t_0) \to_d \mathbf{X}(t_0). \tag{49}$$

Proposition 5 of [1] proved the *Feller property* of the multiplicative coalescent, which asserts that the distribution at time t is a continuous function of the initial distribution (with respect to convergence in distribution, i.e. weak convergence). So (49) implies convergence of the subsequent processes

$$(\bar{\mathbf{C}}^{(n)}(t), t_0 \leq t < \infty) \to_d (\mathbf{X}(t), t_0 \leq t < \infty).$$

That this holds for each $t_0 > -\infty$ implies convergence of the whole processes, which is the conclusion of Theorem 1.

Acknowledgement. We thank the referees for helpful suggestions on improving the exposition.

References

[1] D.J. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.*, 25:812–854, 1997.

[2] D.J. Aldous. Emergence of the giant component in special Marcus-Lushnikov processes. *Random Structures Algorithms*, 12:179–196, 1998.

[3] D.J. Aldous and V. Limic. The entrance boundary of the multiplicative coalescent. *Electron. J. Probab.*, 3:1–59, 1998.

[4] B. Bollobás. *Random Graphs.* Academic Press, London, 1985.

[5] M. Capobianco and O. Frank. Graph evolution by stochastic additions of points and lines. *Discrete Math.*, 46:133-143, 1983.

[6] P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.*, 5:17–61, 1960.

[7] S.N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence.* Wiley, New York, 1986.

[8] G.R. Grimmett. *Percolation.* Springer-Verlag, Berlin, 1989.

[9] S. Janson, D. E. Knuth, T. Łuczak, and B. Pittel. The birth of the giant component. *Random Structures and Algorithms*, 4:233–358, 1993.

[10] O. Kallenberg. *Foundations of Modern Probability.* Springer-Verlag, 1997.

[11] V.F. Kolchin. On the behavior of a random graph near a critical point. *Theory Probab. Appl.*, 31:439-451, 1986.

[12] T. Łuczak. Phase transition phenomena in random discrete structures. *Discrete Math.*, 136:225–242, 1994.

[13] T. Łuczak, B. Pittel, and J.C. Wierman. The structure of a random graph at the point of the phase transition. *Trans. Amer. Math. Soc.*, 341:721–748, 1994.

[14] J.R. Norris. Smoluchowski's coagulation equation: Uniqueness, non-uniqueness and a hydrodynamic limit for the stochastic coalescent. *Ann. Appl. Probab.* 9: 78–109, 1999.

[15] B. Pittel. A random graph with a subcritical number of edges. *Trans. Amer. Math. Soc.*, 309:51–75, 1988.

[16] B. Pittel. On tree census and the giant component in sparse random graphs. *Random Structures Algorithms*, 1:311–342, 1990.

[17] J. L. Spouge. Solutions and critical times for the monodisperse coagulation equation when $a(i,j) = A + B(i+j) + Cij$. *J. Phys. A: Math. Gen.*, 16:767–773, 1983.

[18] V.E. Stepanov. Combinatorial algebra and random graphs. *Theory Probab. Appl.*, 14:373–399, 1969.

[19] V.E. Stepanov. On some features of the structure of a random graph near a critical point. *Theory Probab. Appl.*, 32:573–594, 1988.