Emergence of the Giant Component in Special Marcus-Lushnikov Processes

David Aldous*
Department of Statistics
University of California
Berkeley CA 94720
aldous@stat.berkeley.edu
http://www.stat.berkeley.edu/users/aldous

June 17, 1997

Abstract

Component sizes in the usual random graph process are a special case of the Marcus-Lushnikov process discussed in the scientific literature, so it is natural to ask how theory surrounding emergence of the giant component generalizes to the Marcus-Lushnikov process. Essentially no rigorous results are known; we make a start by proving a weak result, but our main purpose is to draw this topic to the attention of random graph theorists.

1 Introduction

1.1 Background

At time zero there are n separate "atoms"; as time increases, these atoms coalesce into clusters according to the rule

> for each pair of clusters, of sizes $\{x, y\}$ say, they coalesce into a single cluster of size $x + y$ at rate $K(x, y)/n$

where $K(x, y) = K(y, x) \geq 0$ is some specified rate kernel. This rule specifies a continuous-time finite-state Markov process which we shall call the

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

Marcus-Lushnikov process. The model was introduced by Marcus [16], and further studied by Lushnikov [15] as a model of gelation. Observe that in the special case $K(x,y) = xy$ the Marcus-Lushnikov process describes the component sizes in the random graph process $G(n, P(\text{edge}) = p(t))$ with $p(t) = 1 - \exp(-t/n) \approx t/n$. Topics surrounding the "emergence of the giant component" in the random graph process have been studied by mathematicians in great detail (see [4, 11] and citations therein), so it seems natural to ask how far the known behavior for $K(x,y) = xy$ extends to more general *gelling kernels* (see (4) below). It turns out that there is a large scientific literature relevant to the Marcus-Lushnikov process, mostly focusing on its deterministic approximation (2). Curiously, this literature has been largely ignored by random graph theorists; a survey aimed at probabilists is given in [2], and we now summarize some relevant aspects.

The state of the Marcus-Lushnikov process at time t may be described in two equivalent ways: as a vector $(N_x(t), x \geq 1)$ where

$$N_x(t) = \text{number of size-}x \text{ clusters}$$

or as a vector $(X_i(t), i \geq 1)$, where

$$X_i(t) = \text{size of } i\text{'th cluster}$$

and the clusters are ordered so that $X_1(t) \geq X_2(t) \geq \ldots$. Heuristically, if we suppose there exists a deterministic limit

$$n^{-1} N_x(t) \ \xrightarrow{p}\ n(x,t) \text{ as } n \to \infty \tag{1}$$

then the limit should satisfy the *Smoluchowski coagulation equations*

$$\tfrac{d}{dt} n(x,t) = \tfrac{1}{2} \sum_{y=1}^{x-1} K(y, x-y) n(y,t) n(x-y,t) - n(x,t) \sum_{y=1}^{\infty} K(x,y) n(y,t) \tag{2}$$

with $n(x,0) = 1_{(x=1)}$. These equations are a classical subject: the 1972 survey by Drake [8] cites 250 papers. In the Marcus-Lushnikov process we have conservation of mass ($\sum_x x N_x(t) = n \ \forall t$) and so one might expect conservation of mass

$$\sum_{x=1}^{\infty} x n(x,t) = 1, \ 0 \leq t < \infty \tag{3}$$

for the solution of the Smoluchowski coagulation equations. When (3) holds, call K a *non-gelling* kernel; it is easy to show rigorously [25] that $K(x,y) \leq$

$k_0(1+x+y)$ is sufficient for K to be non-gelling. If (3) fails, K is a *gelling* kernel with *gelation time* $0 \le T_{\text{gel}} < \infty$ such that

$$\sum_{x=1}^{\infty} xn(x,t) \quad \begin{array}{lcl} = 1 & : & t < T_{\text{gel}} \\ < 1 & : & t > T_{\text{gel}}. \end{array} \tag{4}$$

For the kernel $K(x,y) = xy$ corresponding to the random graph process, $T_{\text{gel}} = 1$. Intuitively speaking, gelation occurs when some non-vanishing proportion of the mass is in clusters whose size is not $O(1)$. Say K has *exponent* γ if

$$K(cx,cy) \sim c^{\gamma}\bar{K}(x,y) \text{ as } c \to \infty.$$

It is widely believed [23, 24] that any "reasonable" kernel K with exponent $1 < \gamma \le 2$ is gelling and has $T_{\text{gel}} > 0$, based on arguments showing that the second moment $\sum_x x^2 n(x,t)$ diverges at some finite time. But there are essentially no rigorous results to uphold this belief, except for variations of xy (e.g. $K(x,y) = axy + b(x+y) + c$: [21]) and degenerate cases like $K(x,y) = x^{\gamma} 1_{(y=x)}$ [5].

Returning to the Marcus-Lushnikov process, one expects the "weak law of large numbers" (1) to hold for $t < T_{\text{gel}}$, and though this is classically true for $K(x,y) = xy$, general kernels have only recently become the object of rigorous study (Jeon [12, 13]). More qualitatively, we expect the significance of T_{gel} in the Marcus-Lushnikov process to be as follows. Write $C_n(t)$ for the size of the cluster containing a prespecified atom: $P(C_n(t) = x) = \frac{x}{n} EN_x(t)$. Then we expect T_{gel} to be the threshold for tightness of $(C_n(t); n \ge 1)$, that is

$$\lim_{x\to\infty} \limsup_n P(C_n(t) > x) \quad = 0, \quad t < T_{\text{gel}} \tag{5}$$

$$\lim_{x\to\infty} \liminf_n P(C_n(t) > x) \quad > 0, \quad t > T_{\text{gel}}. \tag{6}$$

1.2 Statement of result

Our result concerns kernels of a special form. Recall $X_1(t)$ and $X_2(t)$ are the two largest cluster-sizes, and that $C_n(t)$ is the size of the cluster containing a prespecified atom.

Theorem 1 *Fix $1 < \gamma < 2$ and write $f(x) = x^{\gamma}$. Consider the Marcus-Lushnikov process with kernel*

$$K(x,y) = \frac{2f(x)f(y)}{f(x+y) - f(x) - f(y)}. \tag{7}$$

(a) For fixed $t < 1$

$$\lim_{x\to\infty} \limsup_n P(C_n(t) > x) = 0.$$

(b) For fixed $t < 1$ there exists $h_n = o(n^\varepsilon)\ \forall \varepsilon > 0$ such that

$$P(X_1(t) > h_n) \to 0.$$

(c) There exist random times $U_n \xrightarrow{d} 1$ such that $\inf_{t\geq U_n} X_1(t)/X_2(t) \xrightarrow{d} \infty$.

Discussion. Note that $f(x) = x^2$ would give $K(x, y) = xy$, the random graph process. Theorem 1 provides weak formalizations of the idea that a giant cluster emerges around time 1. What is important for our proof is the form (7) of the kernel and that K has exponent γ, rather than the exact form of $f(x)$, and the proof should extend essentially unchanged to the case where f is regularly varying with exponent γ. The special feature of kernels of form (7) is that, writing $s(t) = \sum_x f(x)n(x,t)$, the Smoluchowski coagulation equations yield

$$\tfrac{ds(t)}{dt} = s^2(t) \tag{8}$$

implying

$$s(t) = (1-t)^{-1},\ 0 \leq t < 1. \tag{9}$$

(We haven't seen this idea stated explicitly, but with hindsight it seems implicit in Ziff [26], appendix). Now (9) implies $T_{\text{gel}} \geq 1$, and strongly suggests $T_{\text{gel}} = 1$, but we are unable to prove this, or to prove its stochastic analog (6). For the random graph process it is classical that $X_1(t) = O(\log n)$ for fixed $t < 1$, so (b) is a comparatively weak assertion. We cannot prove the complementary assertion that $X_1(t) = \Omega(n^{1-o(1)})$ for fixed $t > 1$. Assertion (c) establishes existence at some time near 1 of a "giant cluster" (as measured by relative size) which persists as time increases, but we are unable to show that such a giant cluster does not appear *before* time 1. Finally, note that part (a) could alternatively be deduced from the results of Jeon [12, 13] mentioned above.

While the conclusions of Theorem 1 are much weaker, and the hypothesis on K much more restrictive, than one would like, the point is that Theorem 1 is the first rigorous result dealing with the Marcus-Lushnikov process for a family of gelling kernels with general exponent $1 < \gamma < 2$. Indeed, the only detailed study of this question we have found is van Dongen [22], who gives a non-rigorous treatment of scaling properties near the critical point

T_{gel}. His analysis suggests in particular ([22] eq. (8.6)) that the size of the emerging giant cluster is $\Theta(n^{\frac{2}{1+\gamma}})$, which is of course consistent with the known size $\Theta(n^{2/3})$ for the usual random graph process. Rigorous proof of such refined conjectures presents a major challenge, as does study of general gelling kernels of exponent γ.

Theorem 1 is proved in sections 2 – 4 via fairly routine stochastic calculus techniques. Section 2 develops the stochastic analog of (8,9): see (10) and Proposition 2. Section 3 records two essentially standard exponential tail inequalities for continuous-time martingale-like processes with bounded jumps. In section 4 these tools are used to prove Theorem 1.

Some Monte Carlo simulations are shown in section 5.1. A different approach to the analysis of Marcus-Lushnikov processes for a different special class of kernels is mentioned in section 5.2. Finally, we mention that stochastic calculus techniques are also useful in analysis [1, 3] of the *multiplicative coalescent*, i.e. the $n \to \infty$ limit continuous-space process arising from the random graph process.

2 The process $S(t)$

Recall that the Marcus-Lushnikov process is described equivalently as a vector $(N_x(t), x \geq 1)$ where

$$N_x(t) = \text{number of size-}x\text{ clusters}$$

or as a vector $(X_i(t), i \geq 1)$, where

$$X_i(t) = \text{size of } i\text{'th cluster}$$

and the clusters are ordered so that $X_1(t) \geq X_2(t) \geq \ldots$. Assume K is of the form (7) specified in Theorem 1. Write $\mathcal{F}(t)$ for the natural filtration. The "stochastic calculus" we use is no more than estimates of conditional means and variances; instead of the usual theoretical probabilist's notation (e.g. [20]) we use more intuitive "infinitesimal" notation. That is, $E(dZ(t)|\mathcal{F}(t) = a(t)dt$ means that $Z(t) - A(t)$ is a local martingale for $A(t) = \int_0^t a(s)ds$, and var $(dZ(t)|\mathcal{F}(t)) = v(t)dt$ means that $(Z(t) - A(t))^2 - \int_0^t v(s)ds$ is a local martingale. All asymptotics are as $n \to \infty$; we suppress dependence on n in our notation.

Our analysis centers on the process

$$S(t) = n^{-1} \sum_x f(x) N_x(t) = n^{-1} \sum_i f(X_i(t)).$$

Note $S(0) = 1$ and $S(t)$ is increasing. This section is devoted to the proof of the following stochastic analog of (9).

Proposition 2 *For fixed* $t_0 < 1$

$$\sup_{0 \le t \le t_0} |S(t) - \tfrac{1}{1-t}| \ \xrightarrow{p}\ 0.$$

The proof proceeds via a series of lemmas.

Lemma 3

$$E(dS(t)|\mathcal{F}(t)) = (S^2(t) - n^{-1}Y(t))\ dt \tag{10}$$

where

$$Y(t) = n^{-1} \sum_x f^2(x) N_x(t) = n^{-1} \sum_i f^2(X_i(t)).$$

Proof. From the definition of the Marcus-Lushnikov process,

$$E(dS(t)|\mathcal{F}(t)) =$$

$$\tfrac{1}{2n} \sum_x \sum_y (f(x+y) - f(x) - f(y)) \tfrac{K(x,y)}{n} (N_x(t)N_y(t) - N_x(t)1_{(y=x)})\ dt.$$

Using the special form (7) of K, this becomes

$$\tfrac{1}{n^2} \sum_x \sum_y f(x)f(y)(N_x(t)N_y(t) - N_x(t)1_{(y=x)})\ dt = (S^2(t) - n^{-1}Y(t))\ dt.$$

□

Set $\alpha = (\frac{1}{2} + \frac{1}{\gamma})/2$, so that $\frac{1}{2} < \alpha < \frac{1}{\gamma}$, and note in particular that $\alpha\gamma - 1 < 0$. Define

$$T = \min\{t : X_1(t) \ge n^\alpha\}.$$

Lemma 4 *For fixed* $t_0 < 1$

$$\sup_{0 \le t \le \min(t_0, T)} |S(t) - \tfrac{1}{1-t}| \ \xrightarrow{p}\ 0. \tag{11}$$

Proof. The estimates in this proof hold for $t \le T$. Write $\Delta S(t) = S(t) - S(t-)$ for the jump-sizes of S. Then

$$\sup_{t \le T} \Delta S(t) \le n^{-1}(f(2n^\alpha) - 2f(n^\alpha)) \le 2n^{\alpha\gamma - 1} \tag{12}$$

$$Y(t) \le S(t) f(X_1(t)) \le 4n^{\alpha\gamma} S(t). \tag{13}$$

And

$$\begin{aligned} \text{var } (dS(t)|\mathcal{F}(t)) &\leq 2n^{\alpha\gamma-1}E(dS(t)|\mathcal{F}(t)) \text{ by (12)} \\ &\leq 2n^{\alpha\gamma-1}S^2(t)dt \text{ by (10).} \end{aligned} \quad (14)$$

Consider

$$Q(t) = 1 - \frac{1}{S(t)} - t.$$

Then

$$dQ(t) = -dt + \frac{dS(t)}{S(t)(S(t) + dS(t))}$$

and so

$$0 \geq dQ(t) + dt - \frac{dS(t)}{S^2(t)} \geq -\frac{(dS(t))^2}{S^3(t)}. \quad (15)$$

Combining (10,13,14),

$$0 \geq E(dQ(t)|\mathcal{F}(t)) \geq -6n^{\alpha\gamma-1}/S(t)\ dt \geq -6n^{\alpha\gamma-1}dt$$

$$\text{var } (dQ(t)|\mathcal{F}(t)) \leq \frac{\text{var } (dS(t)|\mathcal{F}(t))}{S^4(t)} \leq \frac{2n^{\alpha\gamma-1}dt}{S^2(t)} \leq 2n^{\alpha\gamma-1}dt.$$

Recall that all these estimates are asserted only for $t \leq T$. By a straightforward application of the L^2 maximal inequality, as $n \to \infty$

$$\sup_{0 \leq t \leq \min(2,T)} |Q(t)| \ \stackrel{p}{\to}\ 0.$$

Because $S(t) = 1/(1 - t - Q(t))$, this implies that T is asymptotically at most 1:

$$(T-1)^+ \ \stackrel{p}{\to}\ 0 \quad (16)$$

and also establishes the Lemma. □

Lemma 5

$$K(x, y) \leq \tfrac{A}{2}(xy^{\gamma-1} + yx^{\gamma-1}) \quad (17)$$

where A depends only on γ.

Proof. By considering the ratios, and scaling to make $x = 1$, we need to verify

$$\sup_y \frac{y^\gamma}{((1+y)^\gamma - 1 - y^\gamma)(y^{\gamma-1} + y)} < \infty.$$

But the ratio has finite limits at 0 and ∞. $\square$

Next consider

$$V(t) = n^{-1}\sum_i X_i^2(t) = n^{-1}\sum_x x^2 N_x(t).$$

Lemma 6 *Write*

$$R(t) = \frac{V^{1/A}(t)}{S(t)}$$

for A as in Lemma 5. Then $E(dR(t)|\mathcal{F}(t)) \leq 6n^{\alpha\gamma-1}R(t)dt$ on $t \leq T$.

Proof. Since a merger of clusters of sizes $\{x,y\}$ causes an increase of $2xy/n$ in V,

$$\begin{aligned} E(dV(t)|\mathcal{F}(t)) &= \tfrac{1}{2}\sum_x\sum_y \frac{2xy}{n}\frac{K(x,y)}{n}(N_x(t)N_y(t) - N_x(t)1_{(y=x)})\ dt \\ &\leq \frac{A}{2n^2}\sum_x\sum_y xy(xy^{\gamma-1} + yx^{\gamma-1})N_x(t)N_y(t)\ dt \qquad \text{by (17)} \\ &= AV(t)S(t)\ dt. \qquad (18) \end{aligned}$$

Expanding $d(1/S(t))$ as at (15),

$$dR(t) \leq \tfrac{1}{A}\frac{V^{\frac{1}{A}-1}(t)}{S(t)}dV(t) - V^{1/A}(t)\frac{dS(t)}{S^2(t)} + V^{1/A}(t)\frac{(dS(t))^2}{S^3(t)}.$$

Evaluate $E(\cdot|\mathcal{F}(t))$ for each term, using (18,10,14):

$$\text{first term } \leq V^{1/A}(t)\ dt$$

$$\text{second term } \leq -V^{1/A}(t)(1 - n^{-1}\tfrac{Y(t)}{S^2(t)})\ dt$$

$$\text{third term } \leq V^{1/A}(t)\frac{2n^{\alpha\gamma-1}}{S(t)}\ dt.$$

Bounding $Y(t)$ by (13), the bound reduces to the bound stated in the lemma. $\square$

Proof of Proposition 2. Set $\tilde{R}(t) = (1-6n^{\alpha\gamma-1}t)^+R(t)$. Lemma 6 implies that $\tilde{R}(\min(t,T))$ is a positive supermartingale, and since $\tilde{R}(0) = 1$ we see that $(\tilde{R}(T))$ is tight (as $n \to \infty$). Then using (16)

$$R(T) \text{ is tight }. \qquad (19)$$

But by definition of T we have $V(T) \geq n^\varepsilon$ for $\varepsilon = 2\alpha - 1 > 0$, and hence $R(T) \geq n^{\frac{\varepsilon}{A}}/S(T)$. So by (19) $S(T) \xrightarrow{p} \infty$. But from (11) this implies $T \xrightarrow{p} 1$ and so the Proposition follows from Lemma 4.

3 Exponential tail bounds

There are well-developed techniques for proving exponential tail bounds for stochastic processes via construction of an "exponential martingale". Results of this type at a high level of generality are presented in section 4.13 of [14] in the continuous-time setting relevant here. (Part of the discrete-time analog is the "method of bounded martingale differences" [17] popularized in the 1980's.) We need the following two results.

Lemma 7 *Let $(M(t))$ be a continuous-time martingale with $M(0) = 0$ and with quadratic variation $Q(t)$, that is $E((dM(t))^2|\mathcal{F}(t)) = dQ(t)$. Suppose $\sup_t |M(t) - M(t-)| \leq 1$. Then*

$$\log P(M(t) \geq a, Q(t) \leq q) \leq \frac{-a^2}{2q} b(a/q) \tag{20}$$

where $b(y) = 2y^{-2}((1+y)\log(1+y) - y)$. In particular, there exists $a_0(q)$ such that

$$\log P(M(t) \geq a, Q(t) \leq q) \leq -\tfrac{1}{2} a \log a, \ \ a \geq a_0(q). \tag{21}$$

Comments. Here (21) follows from (20) by noting $b(y) \sim 2y^{-1}\log y$ as $y \to \infty$. And (20) is stated as a standard fact in [7] equation (2): they write $P(M(t) \geq a) - P(Q(t) > q)$ in place of $P(M(t) \geq a, Q(t) \leq q)$, but our form is equivalent by simply stopping the martingale at $\inf\{t : Q(t) > q\}$. In [7] the result is cited as a reformulation of [14] Theorem 4.13.5.

The second lemma, though not explicitly stated in [14] section 4.13, can be proved using the same set of ideas. We shall just outline the intuitive ideas underlying a proof.

Lemma 8 *Let $(D(t), t \geq 0)$ be a process such that*
(a) $E(dD(t)|\mathcal{F}(t)) \geq bR(t)dt$
(b) $\mathrm{var}\ (dD(t)|\mathcal{F}(t)) \leq aR(t)dt$
(c) $\sup_t |D(t) - D(t-)| \leq 1$
for some process $R(t) \geq 0$ and some constants $0 < a, b < \infty$. Then

$$P(D(t) \leq D(0) - c \text{ for some } t > 0) \leq e^{-\theta c}\ , c > 0$$

where $\theta > 0$ is the solution of $\theta = \frac{a}{b}(e^\theta - 1 - \theta)$.

Outline of proof. The issue is to show that $\exp(-\theta D(t))$ is a supermartingale, for then the desired inequality follows in the usual way via the optional sampling theorem and Markov's inequality. Writing informally $\Delta = dD(t)$ and $dr = R(t)dt$, the supermartingale requirement is: if

$$|\Delta| \leq 1, \quad E\Delta \geq bdr, \quad \text{var } \Delta \leq adr \tag{22}$$

$$\text{then } E\exp(-\theta\Delta) \leq 1 + o(dr).$$

But consider the problem of maximizing $E\exp(-\theta\Delta)$ subject to the constraints (22). It is easy to see that the maximizing distribution of Δ must be the distribution on $\{-1, x\}$ for the x such that the latter two inequalities in (22) are equalities. This distribution is (up to $o(dr)$ terms)

$$P(\Delta = -1) = adr, \quad P(\Delta = (a+b)dr) = 1 - adr$$

and so satisfies

$$\begin{aligned} E\exp(-\theta\Delta) &= e^{\theta}adr \ + \ e^{-\theta(a+b)dr}(1 - adr) + o(dr) \\ &= 1 + (e^{\theta}a - \theta(a+b) - a)dr + o(dr) \\ &= 1 + o(dr) \text{ by definition of } \theta. \end{aligned}$$

4 Proof of Theorem 1

4.1 Part (a)

Recall $C_n(t)$ denotes the size of the cluster containing a prespecified atom. Writing $\phi(x) = f(x)/x$, by Markov's inequality

$$P(C_n(t) > x | S(t)) \leq \frac{E(\phi(C_n(t)) | S(t))}{\phi(x)} = \frac{S(t)}{\phi(x)}$$

and part (a) of Theorem 1 follows from Proposition 2.

4.2 Part (b)

The idea of the proof is to follow the growth of a particular cluster. Distinguish one atom a, and let $\hat{Z}_a(t)$ be the size of the cluster containing atom a at time t. Let T be the first time that the cluster merges with some strictly larger cluster, and let $Z_a(t)$ be the "stopped" process $Z_a(t) = \hat{Z}_a(\min(t, T-))$.

Lemma 9

$$P(X_1(t) \geq x, S(t) \leq s) \leq nP(Z_a(t) \geq x, S(t) \leq s).$$

Proof. If $X_1(t) \geq x$, distinguish some cluster at time t with at least x atoms, and as time decreases, at each split distinguish the larger of the two clusters (choosing arbitrarily if equal). At time 0 we obtain a distinguished atom. Reversing time, we see that the event $\{X_1(t) \geq x\}$ is the union over atoms a of the events $\{Z_a(t) \geq x\}$. The lemma follows. □

We proceed to analyze $Z(t) = Z_a(t)$ via stochastic calculus.

Lemma 10 *For $t < T$,*

$$\begin{aligned} E(dZ(t)|\mathcal{F}(t)) &\leq AZ(t)S(t)dt \\ \text{var}\,(dZ(t)|\mathcal{F}(t)) &\leq 3AZ^2(t)S(t)dt \end{aligned}$$

where A is the constant in Lemma 5.

Proof. $E(d\hat{Z}(t)|\mathcal{F}(t)) = \hat{b}(\hat{Z}(t))dt$, where

$$\hat{b}(z) = \tfrac{1}{n}\sum_i X_i(t)K(z, X_i(t)) \tag{23}$$

and where the sum is over clusters not containing the distinguished atom. For the stopped process Z, we retain only the summands with $X_i(t) \leq z$, for which (by Lemma 5) $K(z, X_i(t)) \leq AzX_i^{\gamma-1}(t)$. So $E(dZ(t)|\mathcal{F}(t)) = b(Z(t))dt$, where

$$b(z) \leq \tfrac{1}{n}\sum_i X_i(t)\ AzX_i^{\gamma-1}(t) = AzS(t).$$

This establishes the first assertion of the lemma. The argument for the second assertion is similar: in calculating var $(dZ(t)|\mathcal{F}(t))$ the term $X_i(t)$ in (23) is replaced by

$$(z + X_i(t))^2 - z^2 = 2zX_i(t) + X_i^2(t) \leq 3zX_i(t)$$

when $X_i(t) \leq z$, and the argument goes through with this extra factor of $3z$. □

Proof of (b). Write $W(t) = \log Z(t)$. Since $dW(t) \leq \frac{dZ(t)}{Z(t)}$, Lemma 10 implies

$$\begin{aligned} E(dW(t)|\mathcal{F}(t)) &\leq AS(t)dt & (24) \\ \text{var}\,(dW(t)|\mathcal{F}(t)) &\leq 3AS(t)dt & (25) \end{aligned}$$

Consider the martingale part of $W(t)$, that is the martingale M with $M(0) = W(0) = 0$ and $dM(t) = dW(t) - E(dW(t)|\mathcal{F}(t))$. Note that by integrating (24),

$$W(t) \leq M(t) + tAS(t) \tag{26}$$

and by integrating (25) the quadratic variation process $Q(t)$ of $M(t)$ satisfies

$$Q(t) \leq 3tAS(t).$$

By construction of the stopped process Z we have $Z(u) \leq 2Z(u-)$ and hence $W(u) - W(u-) \leq \log 2 < 1$, implying $|M(u) - M(u-)| \leq 1$. Fixing $t < 1$ and applying the general martingale tail bound (21),

$$\log P(M(t) \geq x, S(t) \leq \tfrac{2}{1-t}) \leq -\tfrac{1}{2}x \ \log x$$

for sufficiently large x, not depending on n. Take x_n such that

$$x_n = o(\log n), \ \ x_n \log x_n \geq 3 \log n.$$

Then

$$nP(M(t) \geq x_n, S(t) \leq \tfrac{2}{1-t}) \to 0.$$

Applying (26), we see that there exist $h_n = O(\exp(x_n)) = o(n^\varepsilon)$ for all $\varepsilon > 0$ such that

$$nP(Z(t) \geq h_n, S(t) \leq \tfrac{2}{1-t}) \to 0.$$

Lemma 9 now implies

$$P(X_1(t) \geq h_n, S(t) \leq \tfrac{2}{1-t}) \to 0$$

and the proof of part (b) is completed by appealing to Proposition 2.

4.3 Part (c)

Fix $0 < \eta < 1$ and define

$$U = \min\{t : f(X_1(t)) \geq (1-\eta)nS(t)\}.$$

For $t < U$ we have

$$Y(t) \leq f(X_1(t))S(t) \leq (1-\eta)nS^2(t)$$

and so by (10)

$$E(dS(t)|\mathcal{F}(t)) \geq \eta S^2(t)dt, \ \ t \leq U. \tag{27}$$

Define

$$\hat{U}_j = \min\{t : S(t) \geq 2^j\}$$

$$U_j = \min(\hat{U}_j, U).$$

From (27) and the optional sampling theorem

$$E(S(U_{j+1}) - S(U_j)) \geq \eta 2^{2j} E(U_{j+1} - U_j).$$

It is easy to see that the jumps $\Delta S(t) = S(t) - S(t-)$ satisfy $\Delta S(t) \leq S(t-)$, so $S(U_{j+1}) - S(U_j) \leq 4 \cdot 2^j$, and so

$$E(U_{j+1} - U_j) \leq 4\eta^{-1} 2^{-j}.$$

Summing over $j \geq k$,

$$E(U - U_k)^+ \leq 8\eta^{-1} 2^{-k}.$$

But for fixed k, Proposition 2 implies $(U_k - 1)^+ \stackrel{p}{\to} 0$ as $n \to \infty$, and hence

$$(U - 1)^+ \stackrel{p}{\to} 0. \tag{28}$$

From the definition of U we have $f(X_1(U)) \geq (1-\eta)(f(X_1(U)) + f(X_2(U)))$, and so

$$f(X_1(U))/f(X_2(U)) \geq \tfrac{1-\eta}{\eta}. \tag{29}$$

The definition of U also gives the final inequality in

$$n - X_1(U) = \sum_{i \geq 2} X_i(U) \leq \sum_{i \geq 2} f(X_i(U)) = nS(U) - f(X_1(U)) \leq \eta n S(U).$$

If $S(U) \leq \frac{1}{2\eta}$ then $X_1(U) \geq n/2$ implying $S(U) \geq f(X_1(U)) \geq (n/2)^\gamma$, a contradiction for large n. So $S(U)$ is asymptotically at least $\frac{1}{2\eta}$, implying by Proposition 2 that U is asymptotically at least $1 - 2\eta$:

$$(1 - 2\eta - U)^+ \stackrel{p}{\to} 0. \tag{30}$$

Assume we know

Lemma 11

$$P\left(\inf_{t \geq U} \frac{f(X_1(t))}{f(X_2(t))} \leq r\right) \leq \left(\frac{r\eta}{1-\eta}\right)^\theta, \ r \geq 4$$

where $\theta > 0$ depends only on γ.

Then (28,29,30) give $n \to \infty$ asymptotics for $U = U_n(\eta)$ for each fixed η, and therefore hold for $\eta_n \to 0$ sufficiently slowly, in which setting $U \stackrel{p}{\to} 1$ and then Lemma 11 establishes part (c) of Theorem 1.

Proof of Lemma 11. Recall $S(t) = n^{-1} \sum_{i \geq 1} f(X_i(t))$. We separate the contribution of the largest cluster from the remainder, by writing

$$L(t) = n^{-1} f(X_1(t)), \quad R(t) = n^{-1} \sum_{i \geq 2} f(X_i(t)).$$

The definition of U may be rephrased as

$$U = \inf\{t : \tfrac{L(t)}{R(t)} \geq \tfrac{1-\eta}{\eta}\}.$$

Note that $L(t)$ is an increasing process. Although $R(t)$ is not increasing, we can define an increasing process $(R^*(t), t \geq U)$ by censoring the negative jumps:

$$R^*(t) = R(U) + \sum_{U < s \leq t} (R(s) - R(s-))^+ \geq R(t).$$

Take $\eta < 1/5$, so that $L(U)/R(U) > 4$, and define

$$V = \inf\{t > U : L(t)/R(t) < 4\}.$$

Our goal is to obtain estimates for the process $D(t) = \log(L(t)/R^*(t))$, $U \leq t \leq V$. By copying the proof of Lemma 3

$$E(dR^*(t)|\mathcal{F}(t)) \leq R^2(t)dt \tag{31}$$

and so

$$E(d \log R^*(t)|\mathcal{F}(t)) \leq \frac{R^2(t)dt}{R^*(t)} \leq R(t)dt. \tag{32}$$

On $\{U \leq t < V\}$ the largest cluster cannot be overtaken by coalescence of two smaller clusters, so

$$\begin{aligned} &E(dL(t)|\mathcal{F}(t)) \\ &= n^{-1}dt \sum_{i \geq 2} (f(X_1(t)) + f(X_i(t)))n^{-1} \frac{2f(X_1(t))f(X_i(t))}{f(X_1(t) + X_i(t)) - f(X_1(t)) - f(X_i(t))} \\ &\geq dt\ n^{-2} 2f(X_1(t)) \sum_{i \geq 2} f(X_i(t)) \\ &= 2L(t)R(t)dt. \end{aligned} \tag{33}$$

For $U \le t < V$ we have $L(t) \ge 4R(t)$ and so the jumps $\Delta L(t)$ satisfy

$$\Delta L(t) \le ((\tfrac{5}{4})^\gamma - 1)L(t-) \le \tfrac{9}{16}L(t-) \tag{34}$$

and so

$$\Delta \log L(t) \ge \frac{\log \frac{25}{16} - 1}{9/16}\ \frac{\Delta L(t)}{L(t-)} \ge \tfrac{3}{4}\ \frac{\Delta L(t)}{L(t-)}.$$

So from (33)

$$E(d\log L(t)|\mathcal{F}(t)) \ge \tfrac{3}{2}R(t)dt$$

and now (32) implies

$$E(d\log D(t)|\mathcal{F}(t)) \ge \tfrac{1}{2}R(t)dt. \tag{35}$$

For $U \le t < V$ we have $L(t) \ge 4R(t)$ and so

$$f(X_1(t) + X_i(t)) - f(X_1(t)) \le B^{1/2}X_1^{\gamma-1}(t)X_i(t),\ i \ge 2$$

where B depends only on γ. So

$$\begin{aligned}
\text{var }(dL(t)|\mathcal{F}(t)) &\le n^{-2}dt\sum_{i\ge 2}(B^{1/2}X_i(t)X_1^{\gamma-1}(t))^2\ n^{-1}K(X_1(t), X_i(t)) \\
&\le An^{-3}dt\sum_{i\ge 2}BX_i^2(t)X_1^{2\gamma-2}(t)X_1(t)X_i^{\gamma-1}(t) \text{ using Lemma 5} \\
&= ABn^{-3}dt\sum_{i\ge 2}X_1^{2\gamma-1}(t)X_i^{\gamma+1}(t) \\
&\le ABn^{-3}dt\sum_{i\ge 2}X_1^{2\gamma}(t)X_i^{\gamma}(t) \\
&= ABL^2(t)R(t)dt
\end{aligned}$$

and so

$$\text{var }(d\log L(t)|\mathcal{F}(t)) \le ABR(t)dt. \tag{36}$$

And

$$\frac{\Delta R^*(t)}{R(t-)} \le \sup_{x,y>0}\frac{(x+y)^\gamma - x^\gamma - y^\gamma}{x^\gamma + y^\gamma} = 2^{\gamma-1} - 1 \le 1 \tag{37}$$

and so

$$\begin{aligned}
\text{var }(dR^*(t)|\mathcal{F}(t)) &\le R(t)E(dR^*(t)|\mathcal{F}(t)) \\
&\le R^3(t)dt \text{ by (31)}
\end{aligned}$$

and so

$$\mathrm{var}\ (d \log R^*(t) | \mathcal{F}(t)) \leq \tfrac{R^3(t)dt}{(R^*(t))^2} \leq R(t)dt.$$

Combining with (36),

$$\mathrm{var}\ (d \log D(t) | \mathcal{F}(t)) \leq 2(1 + AB) R(t) dt. \tag{38}$$

Now (35,38) verify hypotheses (a,b) of Lemma 8 for $(D(t), U \leq t \leq V)$, and hypothesis (c) follows easily from (34,37). The conclusion of the lemma is

$$P\left(\inf_{0 \leq t \leq V-U} \frac{L(U+t)}{R^*(U+t)} \leq \frac{L(U)}{R(U)} e^{-c} \right) \leq e^{-\theta c},\ c > 0$$

where θ depends only on γ. Since $R^* \geq R$ and $L(U)/R(U) \geq (1-\eta)/\eta$,

$$P\left(\inf_{0 \leq t \leq V-U} \frac{L(U+t)}{R(U+t)} \leq \frac{1-\eta}{\eta} e^{-c} \right) \leq e^{-\theta c},\ c > 0.$$

Provided $\frac{1-\eta}{\eta} e^{-c} > 4$, we may (from definition of V) replace $\inf_{t \leq V-U}$ by $\inf_{t < \infty}$, and so

$$P\left(\inf_{0 \leq t < \infty} \frac{L(U+t)}{R(U+t)} \leq r \right) \leq \left(\frac{r\eta}{1-\eta} \right)^{\theta},\ r > 4.$$

This establishes Lemma 11.

5 Final remarks

5.1 Monte Carlo simulations

Put figures 1, 2, 3 near here.

Figures 1 and 2 present data from a simulation with $\gamma = 1.5$ and $n = 100,000$. Figure 1 indicates slow convergence in Proposition 2. The are several ways to quantify the notion of "size of the emerging giant component". One way is via M_n, the maximum (over time t) of the size of the second-largest cluster at time t. The heuristic analysis mentioned in section 1.2 suggests the conjecture $n^{-\frac{2}{1+\gamma}} M_n \ \stackrel{d}{\to}\ M$ with a non-degenerate limit M. This is known in the "random graph" case $\gamma = 2$ [11, 1]. Figure 3 shows the value of M_n in 10 simulations for varying values of n, and the results are consistent with the conjecture.

5.2 Another approach

Implicit in Lushnikov [15] (see [10, 6] for clearer expositions) is the following result, which gives an exact relationship between the Marcus-Lushnikov process and a finite analog of the Smoluchowski coagulation equations, for certain kernels.

Lemma 12 *Consider the Marcus-Lushnikov process with*

$$K(x,y) = xf(y) + yf(x) \tag{39}$$

for some f. Then

$$P(N_x(t) = n_x, x \geq 1) = n! \prod_x \frac{(b_x(t))^{n_x}}{n_x!}$$

where $(b_x(t))$ are the solutions of the differential equations

$$\frac{d}{dt} b_x(t) = \sum_{i=1}^{x-1} i f(x-i) b_i(t) b_{x-i}(t) - (n-x) f(x) b_x(t)$$

with $b_x(0) = 1_{(x=1)}$.

This result seems close in spirit to mathematical work on random graphs [9, 18, 19] featuring first-order approximations to various stochastic processes by differential equations. It seems plausible that Lemma 12 could be used as a starting point for investigating emergence of the giant cluster for kernels of the form (39) with $f(x) = x^{\gamma-1}$.

Acknowledgements. I thank Michael Ostland for carrying out the simulations in section 5.1, and Vlada Limic for careful reading of a draft version.

References

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

[2] D.J. Aldous. Deterministic and stochastic models for coalescence: a review of the mean-field theory for probabilists. Unpublished. Available via homepage http://www.stat.berkeley.edu/users/aldous, 1997.

[3] D.J. Aldous and V. Limic. The entrance boundary of the multiplicative coalescent. In Preparation, 1997.

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

[5] E. Buffet and J.V. Pulé. Gelation: the diagonal case revisited. *Nonlinearity*, 2:373–381, 1989.

[6] E. Buffet and J.V. Pulé. On Lushnikov's model of gelation. *J. Statist. Phys.*, 58:1041–1058, 1990.

[7] A. Dembo. Moderate deviations for martingales with bounded jumps. *Elect. Comm. in Probab.*, 1:11–17, 1996.

[8] R.L. Drake. A general mathematical survey of the coagulation equation. In G.M. Hidy and J.R. Brock, editors, *Topics in Current Aerosol Research (Part 2)*, volume 3 of *International Reviews in Aerosol Physics and Chemistry*, pages 201–376. Pergammon, 1972.

[9] M. E. Dyer, A. M. Frieze, and B. Pittel. On the average performance of the greedy algorithm for finding a matching in a graph. *Ann. Appl. Probab.*, 3:526–552, 1993.

[10] E.M. Hendriks, J.L. Spouge, M. Eibl, and M. Shreckenberg. Exact solutions for random coagulation processes. *Z. Phys. B - Condensed Matter*, 58:219–227, 1985.

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

[12] I. Jeon. *Gelation Phenomena*. PhD thesis, Ohio State, 1996.

[13] I. Jeon. Gelation phenomena. Paper, in preparation, 1997.

[14] R. Sh. Liptser and A.N. Shiryayev. *Theory of Martingales.* Kluwer, Dordrecht, 1989.

[15] A.A. Lushnikov. Coagulation in finite systems. *J. Colloid and Interface Science*, 65:276–285, 1978.

[16] A.H. Marcus. Stochastic coalescence. *Technometrics*, 10:133–143, 1968.

[17] C. McDiarmid. On the method of bounded differences. In *Surveys in Combinatorics 1989*, pages 148–188. Cambridge Univ. Press, 1989. London Math. Soc. Lecture Notes 141.

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

[19] B. Pittel, J. Spencer, and N. Wormald. Sudden emergence of a giant k-core in a random graph. *J. Combin. Theory Ser. B*, 67:111–151, 1996.

[20] L.C.G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales: Foundations*, volume 1. Wiley, second edition, 1994.

[21] 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.

[22] P.G.J. van Dongen. Fluctuations in coagulating systems II. *J. Statist. Phys.*, 49:927–975, 1987.

[23] P.G.J. van Dongen and M.H. Ernst. On the occurrence of a gelation transition in Smoluchowski's coagulation equation. *J. Statist. Phys.*, 44:785–792, 1986.

[24] P.G.J. van Dongen and M.H. Ernst. Scaling solutions of Smoluchowski's coagulation equation. *J. Statist. Phys.*, 50:295–329, 1988.

[25] W.H. White. A global existence theorem for Smoluchowski's coagulation equation. *Proc. Amer. Math. Soc.*, 80:273–276, 1980.

[26] R. M. Ziff. Kinetics of polymerization. *J. Stat. Physics*, 23:241–263, 1980.