Electronic Journal of Probability

Vol. 3 (1998) Paper no. 3, pages 1–59.

Journal URL
http://www.math.washington.edu/~ejpecp/
Paper URL
http://www.math.washington.edu/~ejpecp/EjpVol3/paper3.abs.html

The Entrance Boundary of the Multiplicative Coalescent

David Aldous* and Vlada Limic
Department of Statistics
University of California
Berkeley CA 94720
e-mail: aldous@stat.berkeley.edu

Abstract

The multiplicative coalescent $\mathbf{X}(t)$ is a l^2-valued Markov process representing coalescence of clusters of mass, where each pair of clusters merges at rate proportional to product of masses. From random graph asymptotics it is known (Aldous (1997)) that there exists a *standard* version of this process starting with infinitesimally small clusters at time $-\infty$. In this paper, stochastic calculus techniques are used to describe all versions $(\mathbf{X}(t); -\infty < t < \infty)$ of the multiplicative coalescent. Roughly, an extreme version is specified by translation and scale parameters, and a vector $\mathbf{c} \in l^3$ of relative sizes of large clusters at time $-\infty$. Such a version may be characterized in three ways: via its $t \to -\infty$ behavior, via a representation of the marginal distribution $\mathbf{X}(t)$ in terms of excursion-lengths of a Lévy-type process, or via a weak limit of processes derived from the standard multiplicative coalescent using a "coloring" construction.

AMS 1991 subject classifications. 60J50, 60J75
Key words and phrases. Markov process, entrance boundary, excursion,
Lévy process, random graph, stochastic coalescent, weak convergence.
Running Head: Coalescent entrance boundary
Submitted to EJP on September 11, 1997. *Final version accepted* on January 19, 1998.

[a]Research supported by N.S.F. Grant DMS96-22859

1 Introduction

1.1 The multiplicative coalescent

Consider the Markov process whose states are unordered collections $\mathbf{x} = \{x_i\}$ of positive real numbers (visualize $\mathbf{x}$ as a configuration of clusters of matter, with x_i as the mass of the i'th cluster) and whose dynamics are described by

$$\text{for each pair of clusters of masses } (x, y)\text{, the pair merges at rate } xy$$
$$\text{into a single cluster of mass } x + y. \tag{1}$$

For a given initial state $\mathbf{x}(0)$ with a finite number of clusters, (1) specifies a continuous-time finite-state Markov process. This of course would remain true if we replaced the merger rate xy in (1) by a more general rate $K(x, y)$ (see section 1.6), but the case $K(x, y) = xy$ has the following equivalent interpretation. Regard each cluster i of the initial configuration $\mathbf{x}(0)$ as a vertex, and for each pair $\{i, j\}$ let $\xi_{i,j}$ be independent with exponential (rate $x_i(0)x_j(0)$) distribution. At time $t \geq 0$ consider the graph whose edge-set is $\{(i, j) : \xi_{i,j} \leq t\}$ and let $\mathbf{X}(t)$ be the collection of masses of the connected components of that graph. Then $(\mathbf{X}(t), 0 \leq t < \infty)$ is a construction of the process (1). Aldous [1] shows that we can extend the state space from the "finite-length" setting to the "l^2" setting. Precisely, let us represent unordered vectors via their decreasing ordering. Define $(l^2_\searrow, d)$ to be the metric space of infinite sequences $\mathbf{x} = (x_1, x_2, \ldots)$ with $x_1 \geq x_2 \geq \ldots \geq 0$ and $\sum_i x_i^2 < \infty$, where d is the natural metric $d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_i (x_i - y_i)^2}$. Then the "graphical construction" above defines a Markov process (the *multiplicative coalescent*) which is a Feller process ([1] Proposition 5) on $l^2_\searrow$ and which evolves according to (1). The focus in [1] was on the existence and properties of a particular process, the *standard* multiplicative coalescent $(\mathbf{X}^*(t), -\infty < t < \infty)$, which arises as a limit of the classical random graph process near the phase transition (see section 1.3). In particular, the marginal distribution $\mathbf{X}^*(t)$ can be described as follows. Let $(W(s), 0 \leq s < \infty)$ be standard Brownian motion and define

$$W^t(s) = W(s) + ts - \tfrac{1}{2}s^2, \ s \geq 0. \tag{2}$$

So W^t is the inhomogeneous Brownian motion with drift $t - s$ as time s. Now construct the "reflected" analog of W^t via

$$B^t(s) = W^t(s) - \min_{0 \leq s' \leq s} W^t(s'), \ s \geq 0. \tag{3}$$

The reflected process B^t has a set of excursions from zero. Then ([1] Corollary 2) the ordered sequence of excursion lengths of B^t is distributed as $\mathbf{X}^*(t)$. Note in particular that the total mass $\sum_i X_i^*(t)$ is infinite.

1.2 The entrance boundary

The purpose of this paper is to describe (in Theorems 2 – 4 below) the entrance boundary at time $-\infty$ of the multiplicative coalescent. Call a multiplicative coalescent defined for $-\infty < t < \infty$ *eternal*. General Markov process theory (see e.g. [7] section 10 for a concise treatment) says that any eternal multiplicative coalescent is a mixture of *extreme* eternal multiplicative coalescents, and the extreme ones are characterized by the property that the tail σ-field at time $-\infty$ is trivial. We often refer to different multiplicative coalescents as different *versions* of the multiplicative coalescent; this isn't the usual use of the word *version*, but we don't have a better word.

Write $l^3_\searrow$ for the space of infinite sequences $\mathbf{c} = (c_1, c_2, \ldots)$ with $c_1 \geq c_2 \geq \ldots \geq 0$ and $\sum_i c_i^3 < \infty$. Define parameter spaces

$$\bar{\mathcal{I}} = [0, \infty) \times (-\infty, \infty) \times l^3_\searrow$$

$$\mathcal{I}_+ = (0, \infty) \times (-\infty, \infty) \times l^3_\searrow .$$

For $\mathbf{c} \in l^3_\searrow$ let $(\xi_j, j \geq 1)$ be independent with exponential (rate c_j) distributions and consider

$$V^{\mathbf{c}}(s) = \sum_j \left(c_j 1_{(\xi_j \leq s)} - c_j^2 s \right), \ s \geq 0. \tag{4}$$

We may regard (cf. section 2.5) $V^{\mathbf{c}}$ as a "Lévy process without replacement". It is easy to see (section 2.1) that the condition for (4) to yield a well-defined process is precisely the condition $\sum_i c_i^3 < \infty$. Now modify (2,3) by defining, for $(\kappa, \tau, \mathbf{c}) \in \bar{\mathcal{I}}$,

$$\widetilde{W}^{\kappa,\tau}(s) = \kappa^{1/2} W(s) + \tau s - \tfrac{1}{2}\kappa s^2, \ s \geq 0 \tag{5}$$

$$W^{\kappa,\tau,\mathbf{c}}(s) = \widetilde{W}^{\kappa,\tau}(s) + V^{\mathbf{c}}(s), \ s \geq 0 \tag{6}$$

$$B^{\kappa,\tau,\mathbf{c}}(s) = W^{\kappa,\tau,\mathbf{c}}(s) - \min_{0 \leq s' \leq s} W^{\kappa,\tau,\mathbf{c}}(s'), \ s \geq 0. \tag{7}$$

So $B^{\kappa,\tau,\mathbf{c}}(s)$ is a reflected process with some set of excursions from zero. Now define l_0 to be the set of $\mathbf{c} \in l^3_\searrow$ such that, for each $-\infty < \tau < \infty$ and each $\delta > 0$,

$B^{0,\tau,\mathbf{c}}$ has a.s. no infinite excursion and only finitely many

$$\text{excursions with length } \geq \delta \tag{8}$$

(note here $\kappa = 0$). If $\sum_i c_i^2 < t$, the process $W^{0,t,\mathbf{c}}(s)$ has asymptotic ($s \to \infty$) drift rate $t - \sum_i c_i^2 > 0$ and hence $B^{0,\tau,\mathbf{c}}$ ends with an infinite incomplete excursion. So $l_0 \subseteq l^3_\searrow \setminus l^2_\searrow$. In fact we shall prove (section 5.5)

Lemma 1 $l_0 = l^3_\searrow \setminus l^2_\searrow$.

Defining

$$\mathcal{I} = \mathcal{I}_+ \cup (\{0\} \times (-\infty, \infty) \times l_0),$$

we can now state the main results.

Theorem 2 *For each $(\kappa, \tau, \mathbf{c}) \in \mathcal{I}$ there exists an eternal multiplicative coalescent $\mathbf{X}$ such that for each $-\infty < t < \infty$, $\mathbf{X}(t)$ is distributed as the ordered sequence of excursion lengths of $B^{\kappa, t-\tau, \mathbf{c}}$.*

Write $\mu(\kappa, \tau, \mathbf{c})$ for the distribution of the process $\mathbf{X}$ in Theorem 2. Note also that the constant process

$$\mathbf{X}(t) = (y, 0, 0, 0, \ldots), \ -\infty < t < \infty \tag{9}$$

for $y \geq 0$ is an eternal multiplicative coalescent: write $\hat{\mu}(y)$ for its distribution.

Theorem 3 *The set of extreme eternal multiplicative coalescent distributions is $\{\mu(\kappa, \tau, \mathbf{c}) : \ (\kappa, \tau, \mathbf{c}) \in \mathcal{I}\} \cup \{\hat{\mu}(y) : 0 \leq y < \infty\}$.*

Underlying Theorem 3 is an intrinsic characterization of the process with distribution $\mu(\kappa, \tau, \mathbf{c})$. From the definition of $l^2_\searrow$, in $\mathbf{X}(t) = (X_1(t), X_2(t), \ldots)$ the cluster masses are written in decreasing order, so that $X_j(t)$ is the mass of the j'th largest cluster. Write

$$S(t) = \sum_i X_i^2(t)$$

$$S_r(t) = \sum_i X_i^r(t) \ , \ r = 3, 4.$$

Theorem 4 *Let $(\kappa, \tau, \mathbf{c}) \in \mathcal{I}$. An eternal multiplicative coalescent $\mathbf{X}$ has distribution $\mu(\kappa, \tau, \mathbf{c})$ if and only if*

$$|t|^3 S_3(t) \to \kappa + \textstyle\sum_j c_j^3 \quad \textit{a.s. as } t \to -\infty \tag{10}$$

$$t + \frac{1}{S(t)} \to \tau \quad \textit{a.s. as } t \to -\infty \tag{11}$$

$$|t| X_j(t) \to c_j \quad \textit{a.s. as } t \to -\infty, \ \textit{each } j \geq 1. \tag{12}$$

With this parametrization, the standard multiplicative coalescent has distribution $\mu(1,0,\mathbf{0})$. The parameters τ and κ are time-centering and scaling parameters:

$$\text{if } \mathbf{X} \text{ has distribution } \mu(1,0,\mathbf{c})$$
$$\text{then } \widetilde{\mathbf{X}}(t) = \kappa^{-1/3}\mathbf{X}(\kappa^{-2/3}(t-\tau)) \text{ has distribution } \mu(\kappa,\tau,\kappa^{1/3}\mathbf{c}) \ . \quad (13)$$

From (12) we may interpret $\mathbf{c}$ as the relative sizes of distinguished large clusters at time $-\infty$. Further interpretations of $\mathbf{c}$ are addressed in the next two sections, leading to a recipe for constructing the general such process from the standard multiplicative coalescent.

While Theorems 2 – 4 provide concise mathematical characterizations of the processes $\mu(\kappa,\tau,\mathbf{c})$, they are not very intuitively informative about the nature of these processes. Indeed we have no appealing intuitive explanation of why excursions of a stochastic process are relevant, except via the proof technique (section 2.3) which represents masses of clusters in the multiplicative coalescent as lengths of excursions of certain walks. The technical reason for using Theorem 2 (rather than Theorem 4) as the definition of $\mu(\kappa,\tau,\mathbf{c})$ is that we can appeal to familiar weak convergence theory to establish existence of the multiplicative coalescent with $\kappa = 0$, which we do not know how to establish otherwise.

1.3 Relation to random graph processes

The "random graph" interpretation of the standard multiplicative coalescent $\mathbf{X}^*$ is as follows ([1] Corollary 1). In $\mathcal{G}(n, P(\text{edge}) = \frac{1}{n} + \frac{t}{n^{4/3}})$, let $C_1^n(t) \geq C_2^n(t) \geq \ldots$ be the sizes of the connected components. Then as $n \to \infty$, for each fixed t

$$n^{-2/3}(C_1^n(t), C_2^n(t), \ldots) \ \stackrel{d}{\to}\ \mathbf{X}^*(t) \text{ on } l^2_{\searrow}.$$

Consider $\mathbf{c} = (c_1, \ldots, c_k, 0, 0, \ldots)$, and write $v = \sum_i c_i^2$. The eternal multiplicative coalescent with distribution $\mu(1,-v,\mathbf{c})$ arises as the corresponding limit, where in addition to the random edges there are initially k "planted" components of sizes $\lfloor c_i n^{1/3} \rfloor$. (This is a special case of Proposition 7.) From the viewpoint of random graph asymptotics, it is hard to see which infinite vectors $\mathbf{c}$ are appropriate, but in the next section we reformulate the same idea directly in terms of multiplicative coalescents.

The genesis of this paper and [1] was a question posed informally by Joel Spencer (personal communication) in the random graphs context, asking roughly for the existence and essential uniqueness of some process like the standard multiplicative coalescent. The following corollary of Theorems 3 and 4 is perhaps the simplest formalization of "essential uniqueness".

Corollary 5 *The standard multiplicative coalescent has the property*

$$X_1(t)/S(t) \to 0 \text{ a.s. as } t \to -\infty.$$

Any eternal multiplicative coalescent with this property is a mixture of linearly-rescaled standard multiplicative coalescents.

1.4 The coloring construction

The same idea of "initially planted clusters" can be formulated directly in terms of the multiplicative coalescent. Given a configuration $\mathbf{x} \in l^2_{\searrow}$ and a constant $c > 0$, a random configuration $\mathrm{COL}(\mathbf{x}; c)$ can be defined as follows (see section 5.1 for more details of the following). Imagine distinguishing and coloring atoms according to a Poisson process of rate c per unit mass, so that the i'th cluster (which has mass x_i) contains at least one colored atom with chance $1 - e^{-cx_i}$. Then merge all the clusters containing colored atoms into a single cluster. The notation COL is a mnemonic for "color and collapse". This operation commutes with the evolution of the multiplicative coalescent; that is, for a version $(\mathbf{X}(t), t_1 \le t \le t_2)$, the distribution $\mathrm{COL}(\mathbf{X}(t_2); c)$ is the same as the time-t_2 distribution of the version started at time t_1 with distribution $\mathrm{COL}(\mathbf{X}(t_1); c)$. So given an eternal version $\mathbf{X}$ of the multiplicative coalescent, we can define another eternal version $\mathrm{COL}(\mathbf{X}; c)$ whose marginal distribution at time t is the distribution of $\mathrm{COL}(\mathbf{X}(t); c)$. For finite $\mathbf{c} = (c_1, \ldots, c_k)$ we can construct $\mathrm{COL}(\mathbf{X}; \mathbf{c})$ recursively as $\mathrm{COL}(\mathrm{COL}(\mathbf{X}; (c_1, \ldots, c_{k-1})); c_k)$. It turns out that the construction extends to $\mathbf{c} \in l^2_{\searrow}$.

Theorem 6 *(a) Let $\mathbf{X}^*$ be the standard multiplicative coalescent, and let $\mathbf{c} \in l^2_{\searrow}$. Then $\mathrm{COL}(\mathbf{X}^*; \mathbf{c})$ is the eternal multiplicative coalescent with distribution $\mu(1, -\sum_i c_i^2, \mathbf{c})$.*
(b) For $\mathbf{c} \in l^3_{\searrow}$,

$$\mu(1, 0, (c_1, \ldots, c_k, 0, 0, \ldots)) \xrightarrow{d} \mu(1, 0, \mathbf{c}).$$

(c) For $\mathbf{c} \in l_0$,

$$\mu(\kappa, \tau, \mathbf{c}) \xrightarrow{d} \mu(0, \tau, \mathbf{c}) \text{ as } \kappa \downarrow 0.$$

This allows us to give a "constructive" description of the entrance boundary. We remark that if $\mathbf{c}$ is not in $l^2_\searrow$ then $\mathrm{COL}(\mathbf{X}^*;\mathbf{c})$ does not exist, so one might first guess that (up to rescaling) the entrance boundary consisted essentially only of the processes $\mathrm{COL}(\mathbf{X}^*;\mathbf{c})$ for $\mathbf{c} \in l^2_\searrow$. But Theorem 6 says that the process

$$\mathbf{Y}^k(t) = \mathrm{COL}\left(\mathbf{X}^*(t - \sum_{i=1}^k c_i^2); (c_1, \ldots, c_k)\right) \tag{14}$$

has distribution $\mu(1, 0, (c_1, \ldots, c_k))$, which as $k \to \infty$ converges weakly to $\mu(1, 0, \mathbf{c})$ for all $\mathbf{c} \in l^3_\searrow$, not just $l^2_\searrow$. The point is that the increase in l^2-norm caused by the color-and-collapse operation can be compensated by the time-shift. This is loosely analogous to the construction of Lévy processes as limits of linearly-compensated compound Poisson processes. Now by linear rescaling (13) we can define $\mu(\kappa, \tau, \mathbf{c})$ for $\kappa > 0$, and the final surprise (from the viewpoint of the coloring construction) is that for $\mathbf{c} \in l_0$ one can let $\kappa \to 0$ to obtain a process $\mu(0, \tau, \mathbf{c})$.

A related intuitive picture was kindly suggested by a referee. As noted earlier, from (12) we may interpret $\mathbf{c}$ as the relative masses of distinguished large clusters in the $t \to -\infty$ limit. In this limit, these clusters do not interact with each other, but instead serve as nuclei, sweeping up the smaller clusters in such a way that relative masses converge. The asymptotic non-interaction allows a comparison where these large clusters may be successively removed, reducing the process to the standard multiplicative coalescent. In the case where the limit (12) holds with $\mathbf{c} \in l^2_\searrow$ this is the right picture, and is formalized in Proposition 41. But as described above, the case $\mathbf{c} \in l^3_\searrow$ is more complicated.

1.5 Remarks on the proofs

The proofs involve three rather different techniques, developed in turn in sections 2, 3 and 5. Proposition 7, proved in section 2, gives a "domain of attraction" result generalizing the "if" assertion of Theorem 4. As discussed in section 2, the argument follows in outline the argument ("code as random walk, and use standard weak convergence methods") in [1] for a special case, so we shall omit some details. In section 3 we use stochastic calculus to show (Proposition 18) that for a non-constant extreme eternal multiplicative coalescent the limits (10 – 12) must exist. Section 4 brings these results together with the Feller property to prove Theorems 2 – 4. Section 5 develops the coloring construction, where a central idea (cf. Proposition 41) is that replacing $\mathbf{X}$ by $\mathrm{COL}(\mathbf{X};\mathbf{c})$ has the effect of appending the vector $\mathbf{c}$ to the vector of $t \to -\infty$ limits in (12). The proof of Theorem 6 is completed in section 5.4.

For at least two of the intermediate technical results (Propositions 14(b) and 30) our proofs are clumsy and we suspect simpler proofs exist.

1.6 General stochastic coalescents

Replacing the merger rate xy in (1) by a general kernel $K(x,y)$ gives a more general *stochastic coalescent.* Such processes, and their deterministic analogs, are the subject of an extensive scientific literature, surveyed for probabilists in [2]. Rigorous study of general stochastic kernels with infinitely many clusters has only recently begun. Evans and Pitman [9] work in the l_1 setting, where the total mass is normalized to 1, and give general sufficient conditions for the Feller property. This is inadequate for our setting, where $\mathbf{X}^*(t)$ has infinite total mass: Theorem 3 implies that in l_1 the only eternal multiplicative coalescents are the constants. But the l_1 setting does seem appropriate for many kernels. For the case $K(x,y) = 1$ the "standard" coalescent is essentially Kingman's coalescent [11], say $\mathbf{Z} = (Z(t); 0 < t < \infty)$, and it is easy to prove that $\mathbf{Z}$ is the unique version satisfying $\max_i Z_i(t) \to 0$ a.s. as $t \downarrow 0$. The "additive" case $K(x,y) = x + y$ seems harder: the "standard" version of the additive coalescent is discussed in [9, 4] and the entrance boundary is currently under study.

The stochastic calculus techniques based on (68) used in section 3 can partially be extended to certain other kernels and yield information about the "phase transition" analogous to the emergence of the giant component in random graph theory: see [3].

Acknowledgement. We thank a conscientious referee for helpful comments on this paper.

2 The weak convergence argument

2.1 A preliminary calculation

Let (ξ_j) be independent with exponential (rate c_j) distributions. For fixed j, it is elementary that $c_j 1_{(\xi_j \le s)} - c_j^2 s$ is a supermartingale with Doob-Meyer decomposition $M_j(s) - c_j^2(s-\xi_j)^+$, where M_j is a martingale with quadratic variation $\langle M_j \rangle(s) = c_j^3 \min(s, \xi_j)$. Consider, as at (4),

$$V^{\mathbf{c}}(s) = \sum_j \left(c_j 1_{(\xi_j \le s)} - c_j^2 s \right), \ s \ge 0.$$

For $\mathbf{c} \in l^3_\searrow$ we claim that $V^{\mathbf{c}}$ is a supermartingale with Doob-Meyer decomposition

$$V^{\mathbf{c}}(s) = M^{\mathbf{c}}(s) - A^{\mathbf{c}}(s) \tag{15}$$

where

$$A^{\mathbf{c}}(s) = \sum_j c_j^2 (s - \xi_j)^+$$

and where $M^{\mathbf{c}}$ is a martingale with quadratic variation

$$\langle M^{\mathbf{c}} \rangle(s) = \sum_j c_j^3 \min(s, \xi_j). \tag{16}$$

To verify the claim, it is enough to show that $A^{\mathbf{c}}(s)$ and the right side of (16) are finite. The latter is clear, and the former holds because

$$E(s - \xi_j)^+ \le sP(\xi_j \le s) \le s^2 c_j. \tag{17}$$

Thus the processes $W^{\kappa,\tau,\mathbf{c}}$ and $B^{\kappa,\tau,\mathbf{c}}$ featuring in the statement of Theorem 2 are well-defined.

2.2 The weak convergence result

The rest of section 2 is devoted to the proof of the following "domain of attraction" result. Given $\mathbf{x} \in l^2_\searrow$ define

$$\sigma_r(\mathbf{x}) = \sum_i x_i^r, \ r = 1, 2, 3, 4, 5.$$

Proposition 7 *For each $n \geq 1$ let $(\mathbf{X}^{(n)}(t); t \geq 0)$ be the multiplicative coalescent with initial state $\mathbf{x}^{(n)}$, a finite-length vector. Suppose that, as $n \to \infty$,*

$$\frac{\sigma_3(\mathbf{x}^{(n)})}{(\sigma_2(\mathbf{x}^{(n)}))^3} \to \kappa + \sum_j c_j^3 \tag{18}$$

$$\frac{x_j^{(n)}}{\sigma_2(\mathbf{x}^{(n)})} \to c_j, \ j \geq 1 \tag{19}$$

$$\sigma_2(\mathbf{x}^{(n)}) \to 0 \tag{20}$$

where $0 \leq \kappa < \infty$ and $\mathbf{c} \in l^3_\searrow$. If $(\kappa, 0, \mathbf{c}) \in \mathcal{I}$ then for each fixed t,

$$\mathbf{X}^{(n)}\left(\frac{1}{\sigma_2(\mathbf{x}^{(n)})} + t\right) \xrightarrow{d} \mathbf{Z} \tag{21}$$

where $\mathbf{Z}$ is distributed as the ordered sequence of excursion lengths of $B^{\kappa,t,\mathbf{c}}$. If $\kappa = 0$ and $\mathbf{c} \notin l_0$ then the left side of (21) is not convergent.

The special case where $\mathbf{c} = 0$ and $\kappa = 1$ is Proposition 4 of [1]. The proof of the general case is similar in outline, so we shall omit some details.

It will be important later that Proposition 7 is never vacuous, in the following sense.

Lemma 8 *For any $(\kappa, 0, \mathbf{c}) \in \mathcal{I}$ we can choose $(\mathbf{x}^{(n)})$ to satisfy (18 - 20).*

Proof. In the case $\kappa > 0$ we may (cf. the random graph setting, section 1.3) take $\mathbf{x}^{(n)}$ to consist of n entries of size $\kappa^{-1/3}n^{-2/3}$, preceded by entries $(c_1\kappa^{-2/3}n^{-1/3}, \ldots, c_{l(n)}\kappa^{-2/3}n^{-1/3})$, where $l(n) \to \infty$ sufficiently slowly. In the case $\kappa = 0$ and $\mathbf{c} \in l_0$, take $\mathbf{x}^{(n)}$ to consist of entries $(c_1 n^{-1/3}, \ldots, c_{l(n)}n^{-1/3})$, where $l(n) \to \infty$ fast enough so that $\sum_{i=1}^{l(n)} c_i^2 \sim n^{1/3}$.

2.3 Breadth-first walk

Fix a finite-length initial configuration $\mathbf{x} \in l^2_\searrow$, and fix $0 < q < \infty$. In section 1.1 we described the graphical construction of $\mathbf{X}(q)$, the state at q of the multiplicative coalescent started at state $\mathbf{x}$. Here we copy from [1] section 3.1 an augmentation of the graphical construction, in which there is another notion of "time" and each vertex has a "birth-time" and most vertices have a "parent" vertex. Simultaneously we construct the *breadth-first walk* associated with $\mathbf{X}(q)$.

For each ordered pair (i,j), let $U_{i,j}$ have exponential(qx_j) distribution, independent over pairs. The construction itself will not involve $U_{i,i}$, but they will be useful in some later considerations. Note that with the above choice of rates

$$P(\text{ edge } i \leftrightarrow j \text{ appears before time } q) = 1 - \exp(-x_i x_j q) = P(U_{i,j} \le x_i)\,.$$

Choose $v(1)$ by size-biased sampling, i.e. vertex v is chosen with probability proportional to x_v. Define $\{v : U_{v(1),v} \le x_{v(1)}\}$ to be the set of children of $v(1)$, and order these children as $v(2), v(3), \ldots$ so that $U_{v(1),v(i)}$ is increasing. The children of $v(1)$ can be thought of as those vertices of the multiplicative coalescent connected to $v(1)$ via an edge by time q. Start the walk $z(\cdot)$ with $z(0) = 0$ and let

$$z(u) = -u + \sum_v x_v 1_{(U_{v(1),v} \le u)},\ 0 \le u \le x_{v(1)}\,.$$

So
$$z(x_{v(1)}) = -x_{v(1)} + \sum_{v \text{ child of } v(1)} x_v.$$

Inductively, write $\tau_{i-1} = \sum_{j \le i-1} x_{v(j)}$. If $v(i)$ is in the same component as $v(1)$, then the set

$$\{v \notin \{v(1), \ldots, v(i-1)\} : v \text{ is a child of one of } \{v(1), \ldots, v(i-1)\}\}$$

consists of $v(i), \ldots, v(l(i))$ for some $l(i) \ge i$. Let the children of $v(i)$ be $\{v \notin \{v(1), \ldots, v(l(i))\} : U_{v(i),v} \le x_{v(i)}\}$, and order them as $v(l(i)+1), v(l(i)+2), \ldots$ such that $U_{v(i),v}$ is increasing. Set

$$z(\tau_{i-1} + u) = z(\tau_{i-1}) - u + \sum_{v \text{ child of } v(i)} x_v\ 1_{(U_{v(i),v} \le u)},\ 0 \le u \le x_{v(i)}. \quad (22)$$

After exhausting the component containing $v(1)$, choose the next vertex by size-biased sampling, i.e. each available vertex v is chosen with probability proportional to x_v. Continue. After exhausting all vertices, the above construction produces a forest. Each tree in this forest is also a connected component of the corresponding multiplicative coalescent. After adding extra edges (i,j) for each pair such that $i < j \le l(i)$ and $U_{v(i),v(j)} \le x_{v(i)}$, the forest becomes the graphical construction of the multiplicative coalescent. By construction, both the vertices and the components appear in the size-biased random order.

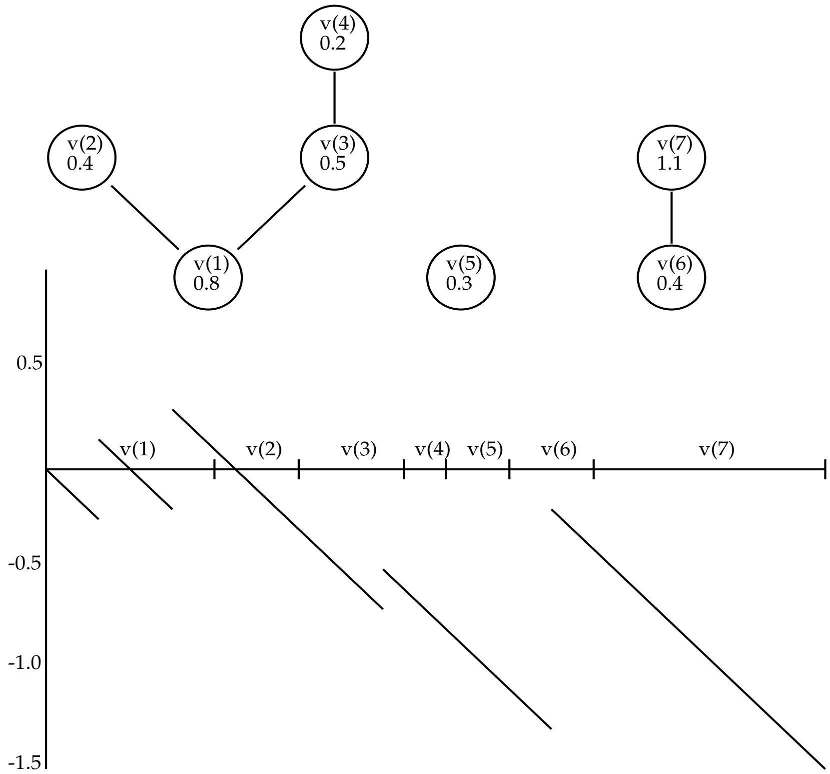

Figure 1

Figure 1 illustrates a helpful way to think about the construction, picturing the successive vertices $v(i)$ occupying successive intervals of the "time" axis, the length of the interval for v being the weight x_v. During this time interval we "search for" children of $v(i)$, and any such child $v(j)$ causes a jump in $z(\cdot)$ of size $x_{v(j)}$. The time of this jump is the *birth time* $\beta(j)$ of $v(j)$, which in this case (i.e. provided $v(j)$ is not the first vertex of its component) is $\beta(j) = \tau_{i-1} + U_{v(i),v(j)}$. These jumps are superimposed on a constant drift of rate -1. If $v(j)$ is the first vertex of its component, its birth time is the start of its time interval: $\beta(j) = \tau_{j-1}$. If a component consists of vertices

$\{v(i), v(i+1), \ldots, v(j)\}$, then the walk $z(\cdot)$ satisfies

$$z(\tau_j) = z(\tau_{i-1}) - x_{v(i)},$$

$$z(u) \geq z(\tau_j) \text{ on } \tau_{i-1} < u < \tau_j.$$

The interval $[\tau_{j-1}, \tau_i]$ corresponding to a component of the graph has length equal to the mass of the component (i.e. of a cluster in the multiplicative coalescent), and this interval is essentially an "excursion above past minima" of the breadth-first walk. This connection is the purpose of the breadth-first walk, and will asymptotically lead to the Theorem 2 description of eternal multiplicative coalescents in terms of excursions of $W^{\kappa,\tau,\mathbf{c}}$.

2.4 Weak convergence of breadth-first walks

Fix $(\mathbf{x}^{(n)}, n \geq 1)$ and t satisfying the hypotheses of Proposition 7. Let $(Z_n(s), 0 \leq s \leq \sigma_1)$ be the breadth-first walk associated with the state at time $q = \frac{1}{\sigma_2(\mathbf{x}^{(n)})} + t$ of the multiplicative coalescent started in state $\mathbf{x}^{(n)}$. Our first goal is to prove (Proposition 9) weak convergence of $\frac{1}{\sigma_2(\mathbf{x}^{(n)})} Z_n(s)$. The argument is an extension of the proof of [1] Proposition 10, in which the terms $R_n(s)$ did not arise.

By hypotheses (18 - 20) we may choose $m(n)$ to be an integer sequence which increases to infinity sufficiently slowly that

$$\left| \sum_{i=1}^{m(n)} \frac{(x_i^{(n)})^2}{(\sigma_2(\mathbf{x}^{(n)}))^2} - \sum_{i=1}^{m(n)} c_i^2 \right| \to 0 \ , \quad \left| \sum_{i=1}^{m(n)} \left(\frac{x_i^{(n)}}{\sigma_2(\mathbf{x}^{(n)})} - c_i \right)^3 \right| \to 0 \ , \tag{23}$$

$$\text{and} \quad \sigma_2(\mathbf{x}^{(n)}) \sum_{i=1}^{m(n)} c_i^2 \to 0 \ . \tag{24}$$

In the sequel, we will sometimes omit n from the notation, and in particular we write σ_2 for $\sigma_2(\mathbf{x}^{(n)})$ and write $\mathbf{x}$ for $\mathbf{x}^{(n)}$. Consider the decomposition

$$Z_n(s) = Y_n(s) + R_n(s) \ , \text{ where } \ R_n(s) = \sum_{i=1}^{m(n)} \left(\hat{x}_i 1_{\{\xi_i^n \leq s\}} - \frac{x_i^2}{\sigma_2} s \right) \ ,$$

with $\xi_i^n = \beta(j)$ when $v(j) = i$, and

$$\begin{aligned} \hat{x}_i &= x_i \ , \ \text{ if } i \text{ is not the first vertex in its component} \\ &= 0 \ , \ \text{ else } . \end{aligned}$$

Define

$$\bar{Z}_n(s) = \frac{1}{\sigma_2} Z_n(s) = \frac{1}{\sigma_2} Y_n(s) + \frac{1}{\sigma_2} R_n(s) = \bar{Y}_n(s) + \bar{R}_n(s), \text{ say.}$$

Proposition 9 *As $n \to \infty$ $(\bar{Y}_n, \bar{R}_n) \ \stackrel{d}{\to}\ (\widetilde{W}^{\kappa,t}, V^{\mathbf{c}})$, where $V^{\mathbf{c}}$ and $\widetilde{W}^{\kappa,t}$ are independent, and therefore $\bar{Z}_n \ \stackrel{d}{\to}\ W^{\kappa,t,\mathbf{c}}$.*

First we deal with $\bar{Y}_n(s) = \bar{Z}_n(s) - \bar{R}_n(s)$. We can write $Y_n = M_n + A_n$, $M_n^2 = Q_n + B_n$ where M_n, Q_n are martingales and A_n, B_n are continuous, bounded variation processes. If we show that, for any fixed s_0,

$$\sup_{s \le s_0} \left| \frac{A_n(s)}{\sigma_2} + \frac{\kappa s^2}{2} - ts \right| \ \stackrel{p}{\to}\ 0 \tag{25}$$

$$\frac{1}{\sigma_2^2}\ B_n(s_0)\ \stackrel{p}{\to}\ \kappa\, s_0 \tag{26}$$

$$\frac{1}{\sigma_2^2}\ E \sup_{s \le s_0} |M_n(s) - M_n(s-)|^2 \to 0\ , \tag{27}$$

then by the standard functional CLT for continuous-time martingales (e.g [8] Theorem 7.1.4(b)) we deduce convergence of $\bar{Y}_n$ to $\widetilde{W}^{\kappa,t}$. Note that (27) is an immediate consequence of (19, 23) and the fact that the largest possible jump of M_n has size $x_{m(n)+1}$. Here and below write m for $m(n)$. Define

$$\bar{\sigma}_r = \sigma_r - \sum_{i=1}^{m} x_i^r\ ,\ r = 2, 3, 4, 5\ .$$

It is easy to check that hypotheses (18, 19) and conditions (23, 24) imply

$$\bar{\sigma}_2 \sim \sigma_2\ ,\ \frac{\bar{\sigma}_3}{\bar{\sigma}_2^3} \to \kappa \text{ as } n \to \infty. \tag{28}$$

Lemma 11 of [1] extends easily to the present setting, as follows.

Lemma 10

$$dA_n(s) = (-1 + \sum_{i=1}^{m} x_i^2/\sigma_2)\ ds\ \ + (t + \tfrac{1}{\sigma_2})(\bar{\sigma}_2 - Q_2(s) - \tilde{Q}_2(s))\ ds$$

$$dB_n(s) = (t + \tfrac{1}{\sigma_2})(\bar{\sigma}_3 - Q_3(s) - \tilde{Q}_3(s))\ ds$$

where, for $\tau_{i-1} \le s < \tau_i$*,*

$$Q_2(s) = \sum_{j\le i} x^2_{v(j)}, \quad Q_3(s) = \sum_{j\le i} x^3_{v(j)}$$

$$\tilde{Q}_2(s) = \sum_{j>i,\beta(j)<s} x^2_{v(j)}, \quad \tilde{Q}_3(s) = \sum_{j>i,\beta(j)<s} x^3_{v(j)},$$

and all sums above are over vertices j *with* $v(j) \notin \{1, 2, \ldots, m\}$*.*

Because $\frac{1}{2}\kappa s^2 - ts = \int_0^s (\kappa u - t)\, du$, showing (25) reduces, by Lemma 10, to showing

$$\sup_{u\le s_0} |d(u)| \stackrel{p}{\to} 0$$

where

$$d(u) = \frac{-1 + (t + \frac{1}{\sigma_2})(\sigma_2 - Q_2(u) - \tilde{Q}_2(u)) - t\sum_{i=1}^m x_i^2}{\sigma_2} + (\kappa u - t).$$

Using (23, 24) and hypotheses (18) and (20), this in turn reduces to proving Lemmas 11 and 12 below. Similarly, (26) reduces to showing

$$\frac{Q_3(s_0) + \tilde{Q}_3(s_0)}{\sigma_2^3} \stackrel{p}{\to} 0.$$

Since $Q_3(s_0) \le x_{m+1}Q_2(s_0)$ and $\tilde{Q}_3(s_0) \le x_{m+1}\tilde{Q}_2(s_0)$, this also follows from Lemma 11 and 12, using (28) and hypothesis (19).

Lemma 11 $\sup_{u\le s_0} \tilde{Q}_2(u)/\sigma_2^2 \stackrel{p}{\to} 0$*.*

Lemma 12

$$\sup_{u\le s_0} \left| \frac{1}{\sigma_2^2} Q_2(u) - \frac{\bar{\sigma}_3}{\sigma_2^3} u \right| \stackrel{p}{\to} 0.$$

Proof of Lemma 11. $\tilde{Q}_2(s) \le x_{m+1}\tilde{Q}_1(s)$, where for $\tau_{i-1} \le s < \tau_i$

$$\tilde{Q}_1(s) = \sum_{j>i,\beta(j)<s} x_{v(j)} ,$$

and the index j of summation above is not additionally constrained. By (23) and hypothesis (19) we have $x_{m+1}/\sigma_2 \to 0$, so it is enough to prove

$$\frac{1}{\sigma_2} \sup_{s\le s_0} \tilde{Q}_1(s) \text{ is stochastically bounded as } n \to \infty.$$

This was proved in [1] under the hypothesis $x_1/\sigma_2 \to 0$, but examining the argument reveals that our weaker hypothesis (19) is sufficient for the conclusion.

Proof of Lemma 12. This argument, too, is only a slight variation on the one given in [1]. We exploit the fact that the $(v(i))$ are in size-biased random order. Introduce an artificial time parameter θ, let (T_i) be independent with exponential(x_i) distribution, and consider

$$\begin{aligned} D_1(\theta) &= \sum_{j\geq 1} x_j 1_{(T_j\leq\theta)} - \sigma_2\theta \\ D_2(\theta) &= \sum_{j\geq m+1} x_j^2 1_{(T_j\leq\theta)} - \bar{\sigma}_3\theta \\ D_0(\theta) &= \frac{1}{\sigma_2^2} D_2(\theta) - \frac{\bar{\sigma}_3}{\sigma_2^3} D_1(\theta). \end{aligned}$$

Ordering vertices i according to the (increasing) values of T_i gives the size-biased ordering. So the process

$$\left(\frac{1}{\sigma_2^2} Q_2(\tau_i) - \frac{\bar{\sigma}_3}{\sigma_2^3}\tau_i, i \geq 0\right)$$

is distributed as the process $(D_0(\theta_i), i \geq 0)$, where

$$\theta_i = \min\{\theta : T_j \leq \theta \text{ for exactly } i \text{ different } j\text{'s from } \{m+1, m+2, \ldots\}\ \}.$$

In order to prove Lemma 12 it is enough to show that

$$D(s_0) = \sup\{|D_0(\theta)| \ : D_1(\theta) + \sigma_2\theta \leq s_0\} \xrightarrow{p} 0\ . \tag{29}$$

For $u = 1, 2$ the process $D_u(\theta)$ is a supermartingale, and so by a maximal inequality ([12] Lemma 2.54.5), for $\varepsilon > 0$

$$\tfrac{1}{3}\,\varepsilon P(\sup_{\theta'\leq\theta} |D_u(\theta')| > 3\varepsilon) \leq E|D_u(\theta)| \leq \left(|ED_u(\theta)| + \sqrt{\text{var } D_u(\theta)}\right).$$

Now

$$\begin{aligned} |ED_2(\theta)| &= -ED_2(\theta) \\ &= \sum_{j\geq m+1} x_j^2(x_j\theta + \exp(-x_j\theta) - 1) \\ &\leq \sum_{j\geq m+1} x_j^2\ (x_j\theta)^2/2 \\ &= \theta^2\bar{\sigma}_4/2 \end{aligned}$$

$$\begin{aligned}\text{var } D_2(\theta) &= \sum_{j \geq m+1} x_j^4 P(T_j \leq \theta) P(T_j > \theta) \\ &\leq \sum_{j \geq m+1} x_j^4 \ (x_j \theta) \\ &= \theta \bar{\sigma}_5 .\end{aligned}$$

Similarly

$$|ED_1(\theta)| \leq \theta^2 \sigma_3 / 2; \text{ var } D_1(\theta) \leq \theta \sigma_3. \tag{30}$$

Combining these bounds,

$$\tfrac{1}{3} \varepsilon P(\sup_{\theta' \leq \theta} |D_0(\theta')| > 6\varepsilon) \leq \frac{1}{\sigma_2^2} \left(\frac{\theta^2 \bar{\sigma}_4}{2} + \theta^{1/2} \bar{\sigma}_5^{1/2} \right) + \frac{\sigma_3}{\sigma_2^3} \left(\frac{\theta^2 \sigma_3}{2} + \theta^{1/2} \sigma_3^{1/2} \right).$$

Setting $\theta = 2s_0/\sigma_2$ and using the bounds $\bar{\sigma}_4 \leq x_{m+1}\sigma_3$, $\bar{\sigma}_5 \leq x_{m+1}^2 \sigma_3$, the bound becomes

$$O\left(\frac{x_{m+1}\sigma_3}{\sigma_2^4} + \frac{\sigma_3^{1/2} x_{m+1}}{\sigma_2^{5/2}} + \frac{\sigma_3^2}{\sigma_2^5} + \frac{\sigma_3^{3/2}}{\sigma_2^{7/2}} \right)$$

and this $\to 0$ using (18,19,20). A simple Chebyshev inequality argument shows that for $\theta = 2s_0/\sigma_2$,

$$P(D_1(\theta) + \sigma_2 \theta \leq s_0) \to 0 \ ,$$

which together with (29) verifies Lemma 12. $\square$

We have now shown that $\bar{Y}_n \xrightarrow{d} \widetilde{W}^{\kappa,t}$. In order to complete the proof of Proposition 9 we need to show $\bar{R}_n \xrightarrow{d} V^{\mathbf{c}}$, and moreover that $(\bar{Y}_n, \bar{R}_n) \xrightarrow{d} (\widetilde{W}^{\kappa,t}, V^{\mathbf{c}})$, where $V^{\mathbf{c}}$ and $\widetilde{W}^{\kappa,t}$ are independent.

As a preliminary, note that $\sigma_2^2 \leq \sigma_1 \sigma_3$ by the Cauchy-Schwarz inequality, and so by (18,20)

$$\sigma_1 \geq \sigma_2^2 / \sigma_3 \to \infty. \tag{31}$$

Define $\tilde{\xi}_i^n$ to be the first time of the form $\tau_{k-1} + U_{v(k),v(i)}$ for some k with $U_{v(k),v(i)} \leq x_{v(k)}$ (this is where we need $U_{v(i),v(i)}$). In case $U_{v(k),v(i)} > x_{v(k)}$ for all k, let $\tilde{\xi}_i^n = \sigma_1 + \xi_i^*$, with $\xi_i^* \stackrel{d}{=}$ exponential$((t + \frac{1}{\sigma_2})x_i)$, independent of the walk. By elementary properties of exponentials, $\tilde{\xi}_i^n$ has exponential$((t + \frac{1}{\sigma_2})x_i)$ distribution, and $\tilde{\xi}_i^n$, $\tilde{\xi}_j^n$ are independent for $i \neq j$, $i, j \leq n$. Obviously $\xi_i^n \leq \tilde{\xi}_i^n$ and $\tilde{\xi}_i^n = \xi_i^n$ on the event

$$\{ \text{ vertex } i \text{ is not the first in its component} \}. \tag{32}$$

By (19) we conclude that $\tilde{\xi}_i^n \overset{d}{\to} \xi_i \overset{d}{=}$ exponential(c_i), implying that for any fixed integer M

$$\left(\sum_{i=1}^{M}\left(\frac{x_i}{\sigma_2}1_{\{\tilde{\xi}_i^n \le u\}} - \frac{x_i^2}{\sigma_2^2}u\right), 0 \le u \le s\right) \overset{d}{\to} \left(\sum_{i=1}^{M}\left(c_i 1_{\{\xi_i \le u\}} - c_i^2 u\right), 0 \le u \le s\right), \quad (33)$$

as $n \to \infty$, where the limit $(\xi_i, i \ge 1)$ are independent with $\xi_i \overset{d}{=}$ exponential (rate c_i), $i \ge 1$. We now need a uniform tail bound.

Lemma 13 *For each* $\varepsilon > 0$

$$\lim_{M\to\infty} \limsup_{n\to\infty} P\left(\sup_{u \le s}\left|\sum_{i=M}^{m(n)}\left(\frac{x_i}{\sigma_2}1_{\{\tilde{\xi}_i^n \le u\}} - \frac{x_i^2}{\sigma_2^2}u\right)\right| > \varepsilon\right) = 0\,. \qquad (34)$$

Proof. First split the sum in (34) into

$$\sum_{i=M}^{m}[(t+\tfrac{1}{\sigma_2})x_i 1_{\{\tilde{\xi}_i^n \le u\}} - (t+\tfrac{1}{\sigma_2})^2 x_i^2 u] + \sum_{i=M}^{m}[-t x_i 1_{\{\tilde{\xi}_i^n \le u\}} + t x_i^2 u^2 + 2tu\tfrac{x_i^2}{\sigma_2}]\,,$$

where we recognize the leading term as the supermartingale $V^{\tilde{\mathbf{c}}}(u)$ with

$$\tilde{\mathbf{c}} \equiv (t+\tfrac{1}{\sigma_2})(x_M, x_{M+1}, \ldots, x_m, 0, \ldots)\ .$$

Each remaining term gets asymptotically (uniformly in u) small, as $n \to \infty$, uniformly in M. For example, for the first one we calculate

$$E\left(\sup_{u\le s} t\sum_{i=M}^{m} x_i 1_{\{\tilde{\xi}_i^n \le u\}}\right) \le t\,s\sum_{i=M}^{m} x_i^2 + s\,\frac{\sum_{i=M}^{m} x_i^2}{\sigma_2} \to 0 \text{ by } (20,\, 24)\ ,$$

and the other two terms are even easier to bound. So it is enough to show

$$\lim_{M\to\infty} \limsup_{n\to\infty} P\left(\sup_{u\le s}\left|\sum_{i=M}^{m}(t+\tfrac{1}{\sigma_2})x_i 1_{\{\tilde{\xi}_i^n \le u\}} - (t+\tfrac{1}{\sigma_2})^2 x_i^2 u\right| > \varepsilon\right) = 0\ . \qquad (35)$$

For $(M^{\tilde{\mathbf{c}}}, A^{\tilde{\mathbf{c}}})$ defined in (15), estimates (16, 17) give

$$E(M^{\tilde{\mathbf{c}}})^2(s) \le \sum_{i=M}^{m}(t+\tfrac{1}{\sigma_2})^3 x_i^3\, s\ ,\ \ EA^{\tilde{\mathbf{c}}}(s) \le \sum_{i=M}^{m}(t+\tfrac{1}{\sigma_2})^3 x_i^3\, s^2.$$

Hypotheses (18 – 20) and conditions (23, 24) imply

$$\lim_{n\to\infty}\sum_{i=M}^{m}(t+\tfrac{1}{\sigma_2})^3 x_i^3 = \sum_{i=M}^{\infty} c_i^3,$$

and now a standard maximal inequality application yields (35). □

Lemma 13 and (33) imply, by a standard weak convergence result ([6] Theorem 4.2),

$$\sum_{i=1}^{m}\left(\frac{x_i}{\sigma_2}1_{\{\tilde{\xi}_i^n\le s\}}-\frac{x_i^2}{\sigma_2^2}s\right)\ \xrightarrow{d}\ V^{\mathrm{c}}(s)\ . \tag{36}$$

Because $(\tilde{\xi}_i^n,\ i<m)$ is independent of $Y_n(s)$, we have joint convergence

$$\left(\bar{Y}_n(s),\sum_{i=1}^{m}\left(\frac{x_i}{\sigma_2}1_{\{\tilde{\xi}_i^n\le s\}}-\frac{x_i^2}{\sigma_2^2}s\right)\right)\ \xrightarrow{d}\ (\widetilde{W}^{\kappa,t},V^{\mathrm{c}}(s))\ , \tag{37}$$

with independent $\widetilde{W}^{\kappa,t}$ and $V^{\mathrm{c}}(s)$. To complete the proof of Proposition 9 it is enough to show that

$$\begin{aligned}
\sup_{u\le s}\left|\sum_{i=1}^{m}\left(\frac{x_i}{\sigma_2}1_{\{\tilde{\xi}_i^n\le u\}}-\frac{\hat{x}_i}{\sigma_2}1_{\{\xi_i^n\le u\}}\right)\right| &= \sup_{u\le s}\left|\sum_{i=1}^{m}\frac{x_i}{\sigma_2}1_{\{\xi_i^n<\tilde{\xi}_i^n\le u\}}\right| \\
&= \sum_{i=1}^{m}\frac{x_i}{\sigma_2}1_{\{\xi_i^n<\tilde{\xi}_i^n\le s\}} \qquad (38)\\
&\xrightarrow{p} 0. \qquad (39)
\end{aligned}$$

For the event $\{\xi_i^n<\tilde{\xi}_i^n\le s\}$ to occur at some random time $U\le s$ the walk must exhaust some finite number of components ending with a vertex $v(J)$, and then $v(J+1)$ must be i. Since $v(J+1)$ is chosen by size-biased sampling,

$$P(v(J+1)=i|U,J)=\frac{x_i}{\sigma_1-U}\le\frac{x_i}{\sigma_1-s}\quad\text{on }\{U\le s\}\text{ a.s.}.$$

In other words $P(\xi_i^n<\tilde{\xi}_i^n\le s)\le\frac{x_i}{\sigma_1-s}$ on $\{U\le s\}$. So the expectation of (38) equals

$$\sum_{i=1}^{m}\frac{x_i}{\sigma_2}P(\xi_i^n<\tilde{\xi}_i^n\le s)\le\sum_{i=1}^{m}\frac{x_i}{\sigma_2}\cdot\frac{x_i}{\sigma_1-s}\le\frac{1}{\sigma_1-s}\to 0\ ,$$

by (31).

2.5 Properties of the Lévy-type limit process

Having proved Proposition 9, to prove Proposition 7 we need to verify that the excursions of the reflected version of the normalized walk $\bar{Z}_n$ converge in $(l^2_\searrow, d)$ to the excursions of reflected $W^{\kappa,t,\mathbf{c}}$, which is defined to be $B^{\kappa,t,\mathbf{c}}$. This will be done in section 2.6. As a preliminary, we need the following properties of the limit process, which were routine, and hence not explicitly displayed, in the "purely Brownian" $\mathbf{c} = \mathbf{0}$ setting of [1]. In principle one should be able to prove Lemma 1 also directly from the definition, but we are unable to do so.

Proposition 14 *Let* $(\kappa, t, \mathbf{c}) \in \mathcal{I}$ *and write* $W(s) = W^{\kappa,t,\mathbf{c}}(s)$ *and* $B(s) = B^{\kappa,t,\mathbf{c}}(s)$. *Then*
(a) $W(s) \xrightarrow{p} -\infty$ *as* $s \to \infty$.
(b) $P(B(s) = 0) = 0, s > 0$.
(c) $\max\{y_2 - y_1 : \ y_2 > y_1 \geq s_0, (y_1, y_2)$ *is an excursion of* $B(\cdot)\} \xrightarrow{p} 0$ *as* $s_0 \to \infty$.
(d) With probability 1, *the set* $\{s : B(s) = 0\}$ *contains no isolated points.*

The proof occupies the rest of section 2.5. As at (15) write

$$W(s) = \kappa^{1/2} W^*(s) + ts - \frac{1}{2}\kappa s^2 + M^{\mathbf{c}}(s) - A^{\mathbf{c}}(s)\,, \qquad (40)$$

where W^* is a standard Brownian motion. It follows easily from (16) that

$$s^{-1} M^{\mathbf{c}}(s) \to 0 \ \text{ a.s. as } \ s \to \infty. \qquad (41)$$

Moreover $(1 - \frac{\xi_i}{s})^+ c_i^2 \uparrow c_i^2$ for all i, so by the monotone convergence theorem

$$\frac{A^{\mathbf{c}}(s)}{s} = \sum_i \left(1 - \frac{\xi_i}{s}\right)^+ c_i^2 \to \sum c_i^2 \leq \infty \ \text{ a.s.}.$$

Recall that $\sum_i c_i^2 = \infty$ if $\kappa = 0$. Since of course $s^{-1} W^*(s) \to 0$ a.s., representation (40) implies $s^{-1} W(s) \to -\infty$ a.s., which gives assertion (a).

Assertion (c) is true by definition of l_0 in the case $\kappa = 0$. If $\kappa > 0$ we may assume $\kappa = 1$ by rescaling. Restate (c) as follows: for each $\varepsilon > 0$

$$\text{number of (excursions of } B \text{ with length} > 2\varepsilon) \ < \infty \ \text{ a.s.} \qquad (42)$$

Fix $\varepsilon > 0$ and define events $C_n = \{\sup_{s \in [(n-1)\varepsilon, n\varepsilon]} (W((n+1)\varepsilon) - W(s)) > 0\}$. For $t_1, t_2 \in R$, write $t_1 \sim t_2$ if both t_1 and t_2 are straddled by the same excursion of B. Note that $\{(n-1)\varepsilon \not\sim n\varepsilon \sim (n+1)\varepsilon\} \subset C_n$, so it suffices to show $P(C_n \ i.o.) = 0$. In fact, by (41) it is enough to show

$$\sum_{n \geq s_0/\varepsilon} P(C_n \cap C^{s_0}) < \infty\,, \ \text{for all large } s_0\,, \qquad (43)$$

where $C^{s_0} = \{\sup_{s \geq s_0} |M(s)/s| \leq \frac{\varepsilon^2}{4}\}$. From (40) we get

$$C_n \subseteq \{ \sup_{s \in [(n-1)\varepsilon, n\varepsilon]} W^*(\varepsilon(n+1)) - W^*(s) > (-2t\varepsilon) \wedge (-\varepsilon t)$$

$$+\frac{\varepsilon^2}{2}(2n+1) + A((n+1)\varepsilon) - A(n\varepsilon) - \sup_{s \in [(n-1)\varepsilon, n\varepsilon]} M((n+1)\varepsilon) - M(s)\} .$$

Consider n large enough so that $n-1 \geq s_0/\varepsilon$ and $2t\varepsilon < (2n+1)\varepsilon^2/8$. Then, while on C^{s_0},

$$\sup_{s \in [(n-1)\varepsilon, (n+1)\varepsilon]} \left| \frac{M(s)}{(2n+1)\varepsilon} \right| < \frac{\varepsilon^2}{8} ,$$

and so

$$C_n \cap C^{s_0} \subseteq \left\{ \sup_{s \in [(n-1)\varepsilon, n\varepsilon]} W^*(\varepsilon(n+1)) - W^*(s) \geq \frac{\varepsilon^2}{8}(2n+1) \right\} .$$

Since the increment distribution of W^* doesn't change by shifting time, nor by reversing time and sign,

$$\begin{aligned} P(C_n \cap C^{s_0}) &\leq P(\sup_{s \in [\varepsilon, 2\varepsilon]} W^*(s) > \frac{\varepsilon^2}{8}(2n+1)) \\ &\leq P(W^*(\varepsilon) > \frac{\varepsilon^2}{16}(2n+1)) + P(\sup_{s \in [0,\varepsilon]} W^*(s) > \frac{\varepsilon^2}{16}(2n+1)) \\ &\leq \frac{512}{\varepsilon^3 (2n+1)^2} , \text{ by a maximal inequality.} \end{aligned}$$

This establishes (43) and hence assertion (c).

The proofs of (b) and (d) involve comparisons with Lévy processes, as we now discuss. Given $(\kappa, t, \mathbf{c}) \in \mathcal{I}$, one can define the Lévy process

$$L(s) = \kappa^{1/2} W^*(s) + ts + \sum_i (c_i N_i(s) - c_i^2 s)$$

where W^* is standard Brownian motion and $(N_i(\cdot), i \geq 1)$ are independent Poisson counting processes of rates c_i. Clearly

$$W(s) \leq L(s), \ s \geq 0. \tag{44}$$

By a result on Lévy processes (Bertoin [5] Theorem VII.1)

$$T_{(0,-\infty)} := \inf\{s > 0 : L(s) < 0\} = 0 \text{ a.s.} \tag{45}$$

This is already enough to prove (d), as follows (cf. [5] Proposition VI.4). For a stopping time S for $W(\cdot)$, the incremental process $(W(S+s)-W(S), s \geq 0)$ conditioned on the pre-S σ-field is distributed as $W^{\kappa,t',\mathbf{c}'}(s)$ for some random $t', \mathbf{c}'$. Applying this observation to $S_r = \inf\{s \geq r : W(s) = \inf_{0 \leq u \leq s} W(u)\}$, and then applying (44,45) to $W^{\kappa,t',\mathbf{c}'}(s)$ and the corresponding Lévy process, we see that S_r is a.s. not an isolated zero. This fact, for all rationals r, implies (d).

It remains to prove assertion (b). We first do the (easy) case $\kappa > 0$. Fix s, look at the process just before time s, and Brownian-scale to define

$$W_\varepsilon(u) = \varepsilon^{-1/2}(W(s-u\varepsilon) - W(s)),\ 0 \leq u \leq 1.$$

We claim that

$$W_\varepsilon(\cdot) \stackrel{d}{\to} \kappa^{1/2} W^*(\cdot) \text{ as } \varepsilon \to 0. \tag{46}$$

Clearly $\kappa^{1/2}W^*$ arises as the limit of the $\kappa^{1/2}W^*(u) + tu - \frac{1}{2}\kappa u^2$ terms of W, so it is enough to show

$$(\varepsilon^{-1/2}(V^{\mathbf{c}}(s-u\varepsilon) - V^{\mathbf{c}}(s)),\ 0 \leq u \leq 1) \stackrel{d}{\to} 0 \text{ as } \varepsilon \to 0. \tag{47}$$

Now the contribution to the left side of (47) from the $(c_j 1_{(\xi_j \leq u)} - c_j^2 u)$ term of $V^{\mathbf{c}}$ is asymptotically zero for each fixed j; and as in section 2.1 a routine variance calculation enables us to bound $\sup_{0 \leq u \leq 1} \varepsilon^{-1/2}(V^{\mathbf{c}}(s-u\varepsilon) - V^{\mathbf{c}}(s))$ in terms of $\sum_j c_j^3$. This establishes (47) and thence (46). Combining (46) with the fact that $\inf_{0 \leq u \leq 1} \kappa^{1/2}W^*(u) < 0$ a.s. implies $P(\inf_{s-\varepsilon \leq u \leq s} W(u) < W(s)) \to 1$ as $\varepsilon \to 0$. Thus $P(W(s) = \inf_{u \leq s} W(u)) = 0$, which is (b).

It remains to prove assertion (b) in the case $\kappa = 0$. Recall that in this case $\sum_i c_i^2 = \infty$ and $\sum_i c_i^3 < \infty$. By ([5] Theorem VII.2 and page 158) the analog of (b) holds for $L(\cdot)$: for fixed s_0,

$$P(L(s_0) = \inf_{0 \leq u \leq s_0} L(u)) = 0. \tag{48}$$

To prove (b) we will need an inequality in the direction opposite to (44): this will be given at (50).

Fix s_0. Define a mixture of Lévy processes by

$$Q_m(s) = \left(t - \sum_{i=1}^{m} c_i^2\right) s + \sum_{i \geq m+1} \left(c_i 1_{A_i} M_i(s) - c_i^2 s\right)$$

where (A_i) are independent events with $P(A_i) = (1 - 2c_i s_0)^+$ and where $(M_i(\cdot))$ are independent Poisson processes of rates $\bar{c}_i = c_i e^{-c_i s_0}$. The sum converges by comparing with the sum defining $L(s)$, because $E(N_i(s) - 1_{A_i} M_i(s)) = O(c_i^2)$ as $i \to \infty$. Applying to (Q_m) the Lévy process result (48),

$$P\left(Q_m(s_0) = \inf_{0 \le u \le s_0} Q_m(u)\right) = 0. \tag{49}$$

We shall show that the processes $(Q_{2m}(u), 0 \le u \le s_0)$ and $(W(u), 0 \le u \le s_0)$ can be coupled so that

$$P(Q_{2m}(s_0) - Q_{2m}(u) \le W(s_0) - W(u) \text{ for all } 0 \le u \le s_0) \to 1 \text{ as } m \to \infty. \tag{50}$$

Then (49,50) imply

$$P\left(W(s_0) = \inf_{0 \le u \le s_0} W(u)\right) = 0$$

which is assertion (b).

We shall need the following "thinning" lemma.

Lemma 15 *Given $s_0, \lambda, \lambda_i, i \ge 1$ such that $\lambda \le \lambda_i e^{-\lambda_i s_0}$, let $(\xi_i, i \ge 1)$ be independent, with ξ_i having exponential(λ_i) distribution. Then we can construct a rate-λ Poisson point process on $[0, s_0]$ whose points are a subset of $(\xi_i, 1 \le i \le V)$ where $V - 1$ has Poisson(λs_0) distribution.*

Proof. If $\xi_1 > s_0$, set $V = 1$. If $\xi_1 = s \le s_0$, toss a coin independently, with $\frac{\lambda e^{-\lambda s}}{\lambda_1 e^{-\lambda_1 s}}$ probability of landing heads. If tails, delete the point ξ_1 and set $V = 1$. If heads, the point ξ_1 becomes the first arrival of the Poisson process. Next consider the interval $I_1 = [\xi_1, s_0] = [s, s_0]$ and the point ξ_2. If $\xi_2 \notin I_1$, set $V = 2$. Else, the point $\xi_2 = s + t$ becomes the second arrival of the Poisson process with probability $\frac{\lambda e^{-\lambda t}}{\lambda_2 e^{-\lambda_2 (s+t)}}$. Continue in the same manner. $\square$

Recall that, in the present $\kappa = 0$ setting,

$$W(s) = ts + \sum_i (c_i 1_{(\xi_i \le s)} - c_i^2 s)$$

where ξ_i has exponential(c_i) distribution. Write $(\xi_{i,j}, j \in J_i)$ for the set of points of $M_i(\cdot)$ in $[0, s_0]$ if A_i occurs, but to be the empty set if A_i does not occur. We seek to couple the points $(\xi_{i,j}, 2m < i < \infty, j \in J_i)$ and the points $(\xi_i, 1 \le i < \infty)$ in such a way that $\xi_{i,j} = \xi_{h(i,j)}$ for some random $h(i,j) \le i$ such that the values $\{h(i,j) : i > 2m, j \in J_i\}$ are distinct. Say the coupling is *successful* if we can do this. Clearly a successful coupling induces a coupling of $(Q_{2m}(u), 0 \leq u \leq s_0)$ and $(W(u), 0 \leq u \leq s_0)$ such that the inequality in (50) holds. So it will suffice to show that the probability of a successful coupling tends to 1 as $m \to \infty$. The construction following involves ξ_i for $i \geq m$. Since we are proving the $m \to \infty$ limit assertion (51), we may suppose that $c_i e^{-c_i s_0}$ is non-increasing in $i \geq m$ and that c_i is sufficiently small to satisfy several constraints imposed later.

Fix $M > 2m$. We work by backwards induction on $k = M, M-1, M-2, \ldots, 2m+1$. Suppose we have defined a joint distribution of $(\xi_{i,j}, M \geq i \geq k+1, j \in J_i)$ and $(\xi_i, M \geq i \geq k+1-D(k+1))$, for some random $D(k+1) \geq 0$. For the inductive step, if A_k does not occur then the set J_k is empty, so the induction goes through for $D(k) = (D(k+1)-1)^+$. If A_k does occur, we appeal to Lemma 15 to construct the points $(\xi_{k,j}, j \in J_k)$ as a subset of the points $(\xi_i, k - D(k+1) \geq i \geq k - D(k))$, where $D(k) - D(k+1)$ has Poisson$(\bar{c}_k s_0)$ distribution and is independent of $D(k+1)$. We continue this construction until

$$T_M = \max\{k < M : D(k) \geq k - m\},$$

after which point the "$\lambda \leq \lambda_i$" condition in the Lemma 15 might not hold. Provided $T_M < 2m$ we get a successful coupling. Thus it is enough to prove

$$\lim_{m\to\infty} \limsup_{M\to\infty} P(T_M \geq 2m) = 0. \tag{51}$$

By construction, $(D(k) : M \geq k \geq T_M - m)$ is the non-homogeneous Markov chain specified by $D(M+1) = 0$ and

$$\begin{aligned} D(k) = (D(k+1) - 1)^+ &\text{ on an event of probability } \min(1, 2c_k s_0); \\ &\text{otherwise } D(k) - D(k+1) \text{ has Poisson}(\bar{c}_k s_0) \text{ distribution .} \end{aligned} \tag{52}$$

We analyze this chain by standard exponential martingale techniques.

Lemma 16 *There exist $\theta > 1$ and $\alpha > 0$ such that, provided c_k is sufficiently small,*

$$E(\theta^{D(k)} | D(k+1) = d) \leq \theta^d \exp(-\alpha c_k), \ d \geq 1.$$

Proof. We may take $2c_k s_0 < 1$, and then the quantity in the lemma equals

$$2c_k s_0 \theta^{d-1} \ + \ (1 - 2c_k s_0)\theta^d \exp((\theta - 1)\bar{c}_k s_0).$$

Since $\bar{c}_k \leq c_k$, we obtain a bound $\theta^d f(c_k)$ where

$$f(c) = 2cs_0\theta^{-1} \ + \ (1 - 2cs_0)\exp((\theta - 1)cs_0).$$

So $f(0) = 1$ and

$$\begin{aligned} f'(0) &= 2s_0(\theta^{-1} - 1) + (\theta - 1)s_0 \\ &= -\tfrac{2\,s_0}{15}, \text{ choosing } \theta = 6/5. \end{aligned}$$

So $\alpha = s_0/8$ will serve. $\square$

Now fix i and consider $\zeta_i = \max\{j \leq i : D(j) = 0\}$. Lemma 16 implies that the process

$$\Lambda(k) = \theta^{D(k)} \exp(\alpha(c_{i-1} + \ldots + c_k)),\ i \geq k \geq \zeta_i$$

is a supermartingale. On the event $\{T_M \geq \max(2m, \zeta_i)\}$ we have

$$\Lambda(T_M) \geq \theta^{T_M - m} \exp(\alpha(c_{i-1} + \ldots + c_{T_M})) \geq \theta^m \exp(\alpha(c_{i-1} + \ldots + c_{2m}))$$

the second inequality because we may assume c_{2m} is sufficiently small that $\exp(\alpha c_{2m}) < \theta$. Since $\Lambda(i) = \theta^{D(i)}$, the optional sampling theorem implies

$$P(i \geq T_M \geq \max(2m, \zeta_i) | D(i)) \leq \theta^{D(i) - m} e^{-\alpha\,(c_{i-1} + \ldots + c_{2m})} \text{ on } \{D(i) \geq 1\}$$

and the conditional probability is zero on $\{D(i) = 0\}$. From the transition probability (52) for the step from $i+1$ to i,

$$\begin{aligned} E(\theta^{D(i)} 1_{(D(i) \geq 1)} | D(i+1) = 0) &= \exp((\theta - 1)\bar{c}_i s_0) - \exp(-\bar{c}_i s_0) \\ &\leq \theta \bar{c}_i s_0 \leq \theta c_i s_0. \end{aligned}$$

Combining with the previous inequality,

$$P(i \geq T_M \geq \max(2m, \zeta_i) | D(i+1) = 0) \leq \theta^{-m} e^{-\alpha(c_{i-1} + \ldots + c_{2m})}\ \theta c_i s_0\,. \quad (53)$$

By considering the smallest $i \geq T_M$ such that $D(i+1) = 0$,

$$\begin{aligned} P(T_M \geq 2m) &= \sum_{i=2m}^{M} P(D(i+1) = 0, i \geq T_M \geq \max(2m, \zeta_i)) \\ &\leq \sum_{i=2m}^{M} P(i \geq T_M \geq \max(2m, \zeta_i) | D(i+1) = 0). \end{aligned}$$

Substituting into (53), to prove (51) it suffices to prove

$$\theta^{-m} \sum_{i=2m}^{\infty} c_i \exp(-\alpha(c_{i-1} + \ldots + c_{2m})) \to 0 \text{ as } m \to \infty. \quad (54)$$

In fact the sum is bounded in m, as we now show. For integer $q \geq 0$ write $A(q) = \{i : q \leq c_{i-1} + \ldots + c_{2m} < q + 1\}$. Then (since we may take each $c_i < 1$) we have $\sum_{i \in A(q)} c_i \leq 2$. So

$$\sum_{i \in A(q)} c_i \exp(-\alpha(c_{i-1} + \ldots + c_m)) \leq 2 \exp(-\alpha q)$$

and the sum over q is finite, establishing (54).

2.6 Weak convergence in l^2

The remainder of the proof of Proposition 7 follows the logical structure of the proof of Proposition 4 in [1]. We are able to rely on the theory of size-biased orderings for random sequences in $l^2_\searrow$ and convergence, developed in [1], section 3.3, which we now repeat.

For a countable index set Γ write $l^2_+(\Gamma)$ for the set of sequences $\mathbf{x} = (x_\gamma; \gamma \in \Gamma)$ such that each $x_\gamma \geq 0$ and $\sum_\gamma x_\gamma^2 < \infty$. Write ord $: l^2_+(\Gamma) \to l^2_\searrow$ for the "decreasing ordering" map.

Given $\mathbf{Y} = \{Y_\gamma : \gamma \in \Gamma\}$ with each $Y_\gamma > 0$, construct r.v.'s (ξ_γ) such that, conditional on $\mathbf{Y}$, the (ξ_γ) are independent and ξ_γ has exponential(Y_γ) distribution. These define a random linear ordering on Γ, i.e. $\gamma_1 \leq \gamma_2$ iff $\xi_{\gamma_1} \leq \xi_{\gamma_2}$. For $0 \leq a < \infty$ define

$$S(a) = \sum\{Y_\gamma : \xi_\gamma < a\}. \tag{55}$$

Note that

$$E(S(a)|\mathbf{Y}) = \sum_\gamma Y_\gamma(1 - \exp(-aY_\gamma)) \leq a \sum_\gamma Y_\gamma^2.$$

So if $\mathbf{Y} \in l^2_+(\Gamma)$ then we have $S(a) < \infty$ a.s.. Next we can define $S_\gamma = S(\xi_\gamma) < \infty$ and finally define the *size-biased point process* (SBPP) associated with $\mathbf{Y}$ to be the set $\Xi = \{(S_\gamma, Y_\gamma) : \gamma \in \Gamma\}$. So Ξ is a random element of $\mathcal{M}$, the space of configurations of points on $[0, \infty) \times (0, \infty)$ with only finitely many points in each compact rectangle $[0, s_0] \times [\delta, 1/\delta]$. Note that Ξ depends only on the ordering, rather than the actual values, of the ξ's. Writing π for the "project onto the y-axis" map

$$\pi(\{(s_\gamma, y_\gamma)\}) = \{y_\gamma\} \tag{56}$$

we can recover ord $\mathbf{Y}$ from Ξ via ord $\mathbf{Y} =$ ord $\pi(\Xi)$.

Convergence in (57) below is the natural notion of vague convergence of counting measures on $[0, \infty) \times (0, \infty)$: see e.g. [10].

Proposition 17 ([1] Proposition 15) *Let $\mathbf{Y}^{(n)} \in l^2_+(\Gamma^n)$ for each $1 < n \le \infty$, and let $\Xi^{(n)}$ be the associated SBPP. Suppose*

$$\Xi^{(n)} \xrightarrow{d} \Xi^{(\infty)}, \tag{57}$$

where $\Xi^{(\infty)}$ is a point process satisfying

$$\sup\{s : (s,y) \in \Xi^{(\infty)} \text{ for some } y\} = \infty \quad a.s. \tag{58}$$

$$\text{if } (s,y) \in \Xi^{(\infty)} \text{ then } \sum\{y' : (s',y') \in \Xi^{(\infty)}, s' < s\} = s \quad a.s. \tag{59}$$

$$\max\{y : (s,y) \in \Xi^{(\infty)} \text{ for some } s > s_0\} \xrightarrow{p} 0 \text{ as } s_0 \to \infty. \tag{60}$$

Then $\mathbf{Y}^{(\infty)} = \mathrm{ord}\ \pi(\Xi^{(\infty)})$ is in $l^2_\searrow$, and $\mathrm{ord}\ \mathbf{Y}^{(n)} \xrightarrow{d} \mathrm{ord}\ \mathbf{Y}^{(\infty)}$.

Let $\Xi^{(\infty)}$ be the point process with points

$$\{(l(\gamma), |\gamma|), \gamma \text{ an excursion of } B^{\kappa,t,\mathbf{c}}\ \},$$

where $l(\gamma)$ and $|\gamma|$ are the leftmost point and the length of an excursion γ. In the setting of Proposition 7, let $\mathbf{Y}^{(n)}$ be the set of component sizes of the multiplicative coalescent at time $t + 1/\sigma_2$. If the k'th component of the breadth-first walk consists of vertices $\{v(i), v(i+1), \ldots, v(j)\}$, let $l(n,k) = \tau_{i-1}$ and $C(n,k) = \tau_j - \tau_{i-1}$ be the leftmost point and the size of the corresponding excursion. Let $\Xi^{(n)}$ be the point process with points $\{(l(n,i), C(n,i)) : i \ge 1\}$. Since the components of the breadth-first walk are in size-biased order, $\Xi^{(n)}$ is distributed exactly as the SBPP associated with $\mathbf{Y}^{(n)}$. So the proof of convergence (21) in Proposition 7 will be completed when we check the hypotheses of Proposition 17. But (58,59,60) are direct consequences of Proposition 14(a,b,c). Moreover, the weak convergence $\bar{Z}_n \xrightarrow{d} W$ given by Proposition 9, combined with the property of Proposition 14(d), implies by routine arguments (cf. [1] Lemma 7) the weak convergence (57) of starting-times and durations of excursions: we omit the details.

To establish the "non-convergence" assertion of Proposition 7, suppose $\kappa = 0$ and $\mathbf{c} \in l^3_\searrow \setminus l_0$. So for some t and δ,

$$P(B^{0,t,\mathbf{c}} \text{ has infinitely many excursions of length } > \delta) > \delta \tag{61}$$

(infinite excursions are impossible by Proposition 14(a)). Choose $\mathbf{x}^{(n)}$ satisfying the hypotheses of Proposition 7. Proposition 9 and the argument in the paragraph above show that for some $\omega(n) \to \infty$

$$\liminf_n P\left(\mathbf{X}^{(n)}\left(\frac{1}{\sigma_2(\mathbf{x}^{(n)})}+t\right) \text{ contains } \geq \omega(n) \text{ clusters of size } \geq \delta\right) > \delta \tag{62}$$

implying non-convergence in $l^2_\searrow$.

3 Analysis of eternal multiplicative coalescents

This section is devoted to the proof of

Proposition 18 *Let* $\mathbf{X}$ *be an extreme eternal multiplicative coalescent. Then either* $\mathbf{X}$ *is a constant process (9) or else*

$$|t|^3 S_3(t) \quad \to a \qquad a.s. \text{ as } t \to -\infty \tag{63}$$

$$t + \frac{1}{S(t)} \quad \to \tau \qquad a.s. \text{ as } t \to -\infty \tag{64}$$

$$|t| X_j(t) \quad \to c_j \qquad a.s. \text{ as } t \to -\infty, \text{ each } j \geq 1 \tag{65}$$

where $\mathbf{c} \in l^3_\searrow$, $-\infty < \tau < \infty$, $a > 0$ *and* $\sum_j c_j^3 \leq a < \infty$.

3.1 Preliminaries

As observed in [1] section 4, when $\mathbf{X}(0) = \mathbf{x}$ is finite-length, the dynamics (1) of the multiplicative coalescent can be expressed in martingale form as follows. Let $\mathbf{x}^{(i+j)}$ be the configuration obtained from $\mathbf{x}$ by merging the i'th and j'th clusters, i.e. $\mathbf{x}^{(i+j)} = (x_1, \ldots, x_{u-1}, x_i + x_j, x_u, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{j-1}, x_{j+1}, \ldots)$ for some u. Write $\mathcal{F}(t) = \sigma\{\mathbf{X}(u); u \leq t\}$. Then

$$E(\Delta g(\mathbf{X}(t))|\mathcal{F}(t)) = \sum_i \sum_{j>i} X_i(t) X_j(t) \left(g(\mathbf{X}^{(i+j)}(t)) - g(\mathbf{X}(t))\right) dt \tag{66}$$

for all $g : l^2_\searrow \to R$ (for *all* g because there are only finitely many possible states). Of course, our "infinitesimal" notation $E(\Delta Y(t)|\mathcal{F}(t)) = A(t)dt$ is just an intuitive way of expressing the rigorous assertion that $M(t) = Y(t) - \int_0^t A(s)ds$ is a local martingale; similarly the notation var $(\Delta Y(t)|\mathcal{F}(t)) = B(t)dt$ means that $M^2(t) - \int_0^t B(s)ds$ is a local martingale. Throughout section 3 we apply (66) and the strong Markov property to derive inequalities for general ($l^2_\searrow$-valued rather than just finite-length) versions of the multiplicative coalescent. These can be justified by passage to the limit from the finite-length setting, using the Feller property ([1] Proposition 5). In this way we finesse the issue of proving the strong Markov property in the $l^2_\searrow$-valued setting, or discussing exactly which functions g satisfy (66) in the $l^2_\searrow$-valued setting.

Because a merge of clusters of sizes x_i and x_j causes an increase in S of size $(x_i + x_j)^2 - x_i^2 - x_j^2 = 2x_ix_j$, (66) specializes to

$$E(\Delta f(S(t))|\mathcal{F}(t)) = \sum_i \sum_{j>i} X_i(t)X_j(t)\left(f(S(t) + 2X_i(t)X_j(t)) - f(S(t))\right) dt \tag{67}$$

which further specializes to

$$E(\Delta S(t)|\mathcal{F}_t) = 2\sum_i \sum_{j>i} X_i^2(t)X_j^2(t)dt = \left(S^2(t) - \sum_i X_i^4(t)\right)\ dt. \tag{68}$$

3.2 Martingale estimates

In this section we give bounds which apply to the multiplicative coalescent $(\mathbf{X}(t), t \geq 0)$ with arbitrary initial distribution.

Lemma 19 *Let $T = \min\{t \geq 0 : S(t) \geq 2S(0)\}$. Then*

$$E\int_0^T (S(t) - X_1^2(t))dt \leq 5.$$

Proof. Assume $\mathbf{X}(0)$ is deterministic, and write $s(0) = S(0)$. Since $\sum_i X_i^4(t) \leq X_1^2(t)S(t)$, by (68)

$$E(\Delta S(t)|\mathcal{F}_t) \geq S(t)(S(t) - X_1^2(t))\ dt \geq s(0)(S(t) - X_1^2(t))\ dt.$$

By the optional sampling theorem,

$$ES(T) - s(0) \geq s(0)E\int_0^T (S(t) - X_1^2(t))\ dt.$$

But $S(T-) \leq 2s(0)$ and $S(T) - S(T-) \leq 2X_1^2(T-) \leq 2S(T-)$, so $S(T) \leq 6s(0)$, establishing the inequality asserted in the Lemma. □

Lemma 20 ([1] Lemma 19) *Write $Z(t) = t + \frac{1}{S(t)}$. Then*

$$0 \leq E(\Delta Z(t)|\mathcal{F}(t)) \leq \left(\frac{S_4(t)}{S^2(t)} + \frac{2(S_3(t))^2}{S^3(t)}\right)\ dt \tag{69}$$

and

$$\text{var }(\Delta Z(t)|\mathcal{F}(t)) \leq \frac{2(S_3(t))^2}{S^4(t)}\,dt. \tag{70}$$

Lemma 21 *Define* $Y(t) = \log S_3(t) - 3\log S(t)$. *Then*
(i) $|E(\Delta Y(t)|\mathcal{F}(t))| \leq 15X_1^2(t)dt$
(ii) $\text{var }(\Delta Y(t)|\mathcal{F}(t)) \leq 36X_1^2(t)dt$.

Proof. Write $S_2(t)$ for $S(t)$. By (66) $E(\Delta Y(t)|\mathcal{F}(t))$ equals dt times the following expression, where we have dropped the "t":

$$\sum_i\sum_{j>i} X_iX_j[\log(S_3 + 3X_i^2X_j + 3X_iX_j^2) - \log S_3]$$

$$-3\sum_i\sum_{j>i} X_iX_j[\log(S_2 + 2X_iX_j) - \log S_2].$$

Using the inequality $|\log(a+b) - \log a - \frac{b}{a}| \leq \frac{b^2}{2a^2}$ for $a, b > 0$, we see that $|E(\Delta Y(t)|\mathcal{F}(t))|$ is bounded by dt times

$$\left|\frac{1}{S_3}\sum_i\sum_{j>i} X_iX_j(3X_i^2X_j + 3X_iX_j^2) \;-\; \frac{3}{S_2}\sum_i\sum_{j>i} 2X_i^2X_j^2\right| \tag{71}$$

$$+\sum_i\sum_{j>i} X_iX_j\frac{(3X_i^2X_j + 3X_iX_j^2)^2}{2S_3^2} \;+\; 3\sum_i\sum_{j>i} X_iX_j\frac{(2X_iX_j)^2}{2S_2^2}. \tag{72}$$

The quantity (72) is at most

$$\frac{9}{2}\left(\frac{S_5}{S_3} + \frac{S_4^2}{S_3^2}\right) + \frac{3S_3^2}{S_2^2} \tag{73}$$

which is bounded by $12X_1^2$, using the fact $S_{r+1}/S_r \leq X_1$. The quantity in (71) equals $\frac{3|Z|}{S_2S_3}$, where

$$\begin{aligned} Z &= \sum_i\sum_{j\neq i}\sum_k X_i^3X_j^2X_k^2 - \sum_i\sum_j\sum_{k\neq j} X_i^3X_j^2X_k^2 \\ &= -S_5S_2 + S_3S_4, \end{aligned}$$

and so the quantity in (71) equals $3\left|\frac{S_4}{S_2} - \frac{S_5}{S_3}\right| \leq 3X_1^2$. Combining these bounds gives (i). For (ii), note that var $(\Delta Y(t)|\mathcal{F}(t))$ equals dt times an expression bounded from above by

$$\sum_i\sum_{j>i} X_iX_j[\log(S_3 + 3X_i^2X_j + 3X_iX_j^2) - \log S_3]^2$$

$$+9\sum_i\sum_{j>i} X_iX_j[\log(S_2+2X_iX_j)-\log S_2]^2.$$

Since $(\log(a+b)-\log a)^2 \leq (b/a)^2$ for $a,b>0$, we repeat the argument via (72,73), with slightly different constants, to get the bound

$$9\left(\frac{S_5}{S_3}+\frac{S_4^2}{S_3^2}\right)+\frac{18S_3^2}{S_2^2}\leq 36X_1^2. \quad \square$$

Now imagine distinguishing some cluster at time 0, and following that distinguished cluster as it merges with other clusters. Write $X_*(t)$ for its size at time t. It is straightforward to obtain the following estimates.

Lemma 22
$$E(\Delta X_*(t)|\mathcal{F}(t)) = (X_*(t)S(t)-X_*^3(t))\,dt$$
$$\text{var }(\Delta X_*(t)|\mathcal{F}(t)) \leq X_*(t)S_3(t)\,dt.$$

Our final estimate relies on the graphical construction of the multiplicative coalescent, rather than on martingale analysis.

Lemma 23
$$P(S(t)\leq t\,X_1(0),\ X_2(t)\geq \delta|\mathbf{X}(0)) \leq \delta^{-2}tX_1(0)\exp(-\delta tX_1(0)).$$

Proof. We may assume $\mathbf{X}(0)$ is a non-random configuration $\mathbf{x}(0)$. We may construct $\mathbf{X}(t)$ in two steps. First, let $\mathbf{Y}(t)=(Y_j(t),j\geq 1)$ be the state at time t of the multiplicative coalescent with initial state $(x_2(0),x_3(0),\ldots)$. Second, for each $j\geq 1$ merge the cluster $Y_j(t)$ with the cluster of size $x_1(0)$ with probability $1-\exp(-tx_1(0)Y_j(t))$, independently as j varies. Write N for the number of j such that $Y_j(t)\geq\delta$, and let $M\leq N$ be the number of these j which do not get merged with the cluster of size $x_1(0)$. Since $S(t)\geq N\delta^2$, the probability we seek to bound is at most

$$\begin{aligned}
P(N\leq\delta^{-2}tx_1(0),M\geq 1) &\leq P(M\geq 1|N\leq\delta^{-2}tx_1(0))\\
&\leq E(M|N\leq\delta^{-2}tx_1(0))\\
&\leq \delta^{-2}tx_1(0)\exp(-tx_1(0)\delta).
\end{aligned}$$

3.3 Integrability for eternal processes

Proposition 24 *Let* $\mathbf{X}$ *be an extreme eternal multiplicative coalescent. Then either* $\mathbf{X}$ *is a constant process (9) or else*

$$\limsup_{t\to-\infty}|t|X_1(t)<\infty \ \ a.s. \tag{74}$$

$$\int_{-\infty}^{0} X_1^2(t)dt < \infty \ a.s. \tag{75}$$

and

$$\int_{-\infty}^{0} S^2(t)dt < \infty \ a.s. \ . \tag{76}$$

Proof. By extremality

$$\limsup_{t\to-\infty} |t|X_1(t) = C \text{ a.s.} \tag{77}$$

for some constant $0 \le C \le \infty$. Suppose $C = \infty$. Fix a large constant M and define

$$T_n = \inf\{t \ge -n : |t|X_1(t) \ge M\}.$$

Applying Lemma 23 to $(\mathbf{X}(T_n + u), u \ge 0)$ and $t = -T_n$, on the event $\{T_n < 0\}$, gives

$$P(T_n < 0, S(0) \le |T_n|X_1(T_n), X_2(0) \ge \delta)$$

$$\le \delta^{-2} E|T_n|X_1(T_n)\exp(-\delta|T_n|X_1(T_n)).$$

The supposition $C = \infty$ implies that $P(T_n < 0, |T_n|X_1(T_n) \ge M) \to 1$ as $n \to \infty$, and so

$$P(S(0) \le M, X_2(0) \ge \delta) \le \delta^{-2} \ \sup\{y\exp(-\delta y) : y \ge M\}. \tag{78}$$

Letting $M \to \infty$ we see $P(X_2(0) \ge \delta) = 0$. But δ is arbitrary, and so $\mathbf{X}(0)$ consists of just a single cluster. The same argument applies to $\mathbf{X}(t)$ for each t, and hence $\mathbf{X}$ is a constant process. So now consider the case $C < \infty$, i.e. where (74) holds. Then (75) follows immediately. So we can define $U > -\infty$ by

$$\int_{-\infty}^{U} X_1^2(t)dt = 1.$$

Note that a.s. $\lim_{t\to-\infty} S(t) = 0$. Otherwise we would have $S(t) \to s^*$, $s^* > 0$ a constant by extremality, and Lemma 19 would give $X_1(t) \to \sqrt{s^*}$, contradicting (74). Now define

$$\widetilde{T}_m = \inf\{t : S(t) \ge 2^{-m}\}, \ m = 1, 2, \ldots$$

$$T_m = \min(\widetilde{T}_m, U)\,,$$

so that $P(T_m > -\infty) = 1$ for all m and $T_m \downarrow -\infty$. Using Lemma 19,

$$E \int_{T_{m+1}}^{T_m} (S(t) - X_1^2(t))dt \leq 5.$$

Because $S(t) \leq 2^{-m}$ on $T_{m+1} \leq t < T_m$,

$$E \int_{T_{m+1}}^{T_m} S^2(t)dt \leq 2^{-m} E \int_{T_{m+1}}^{T_m} S(t)dt \leq 2^{-m} \left(5 + E \int_{T_{m+1}}^{T_m} X_1^2(t)dt\right).$$

Summing over m,

$$E \int_{-\infty}^{T_1} S^2(t)dt \leq 5 + E \int_{-\infty}^{T_1} X_1^2(t)dt \leq 6$$

establishing (76).

3.4 Proof of Proposition 16

We quote a version of the L^2 maximal inequality and the L^2 convergence theorem for (reversed) martingales.

Lemma 25 *Let $(Y(t); -\infty < t \leq 0)$ be a process adapted to $(\mathcal{F}(t))$ and satisfying*

$$|E(\Delta Y(t)|\mathcal{F}(t))| \leq \alpha(t)\ dt, \quad \mathrm{var}\ (\Delta Y(t)|\mathcal{F}(t)) \leq \beta(t)\ dt.$$

(a)For $T_0 < T_1$ bounded $\mathcal{F}(t)$-stopping times

$$P\left(\sup_{T_0 \leq t \leq T_1} |Y(t) - Y(T_0)| \geq \int_{T_0}^{T_1} \alpha(t)dt\ + y\right) \leq P\left(\int_{T_0}^{T_1} \beta(t)dt \geq b\right) + b/y^2.$$

(b) If

$$\int_{-\infty}^{0} \alpha(t)\ dt < \infty \text{ a.s. and } \int_{-\infty}^{0} \beta(t)\ dt < \infty \text{ a.s.}$$

then $\lim_{t \to -\infty} Y(t)$ exists and is finite a.s.

To prove Proposition 18 we consider an extreme eternal multiplicative coalescent $\mathbf{X}$ which is not constant, so that by Proposition 24 we have the integrability results

$$\int_{-\infty}^{0} X_1^2(t)dt < \infty \text{ a.s.} \quad \int_{-\infty}^{0} S^2(t)dt < \infty \text{ a.s.} \quad \int_{-\infty}^{0} X_1(t)S(t)dt < \infty \text{ a.s.} \tag{79}$$

the final inequality via Cauchy-Schwarz. First, apply Lemma 25(b) to the process $Y(t)$ defined in Lemma 21: we deduce that as $t \to -\infty$

$$\frac{S_3(t)}{S^3(t)} \to a \text{ a.s.} \tag{80}$$

where $0 < a < \infty$ is a constant, by extremality. Next we want to apply Lemma 25(b) to the process $Z(t)$ defined in Lemma 20. Because $S_4(t) \leq X_1(t)S_3(t)$ and because $S_3(t) = O(S^3(t))$ by (80), the bounds in Lemma 20 are $O(X_1(t)S(t) + S^3(t) + S^2(t))$, which are integrable by (79). So Lemma 25(b) is applicable to $Z(t) = \frac{1}{t} + S(t)$, and we conclude

$$t + \frac{1}{S(t)} \to \tau \text{ a.s.} \tag{81}$$

where $-\infty < \tau < \infty$ is also a constant. Note in particular the consequence

$$\lim_{t \to -\infty} |t|S(t) = 1 \text{ a.s..}$$

So (80,81) establish (63,64), and it remains only to establish (65).

Recall that Lemma 22 deals with the notion of the size $X_*(t)$ at time $t \geq 0$ of a cluster which was distinguished at time 0. Consider the (rather imprecise: see below) corresponding notion of a distinguished cluster $(X_*(t), t > -\infty)$ in the context of the eternal multiplicative coalescent $\mathbf{X}$. Given such a cluster, for $t < 0$ consider $Y(t) = |t|X_*(t)$. Using Lemma 22,

$$\begin{aligned} E(\Delta Y(t)|\mathcal{F}(t)) &= \Big((|t|S(t) - 1)X_*(t) - |t|X_*^3(t) \Big)\, dt \qquad (82) \\ \text{var } (\Delta Y(t)|\mathcal{F}(t)) &\leq t^2 X_*(t) S_3(t)\, dt. \qquad (83) \end{aligned}$$

To verify that the bounds are integrable, note that (81) implies $(|t|S(t)-1) = O(S(t))$ and Proposition 24 implies $|t|X_*(t) = O(1)$, so the first bound is $O(X_*(t)S(t) + X_*^2(t))$ which is integrable. And $|t|S(t) \to 1$, so using (80) $t^2 S_3(t) = O(S(t))$, hence the second bound is $O(X_*(t)S(t))$. Thus Lemma 25(b) is applicable, and we deduce that as $t \to -\infty$

$$|t|X_*(t) \to c_* \text{ a.s.}$$

for some constant $c_* \geq 0$. This argument is rather imprecise, since the notion of "a distinguished cluster starting at time $-\infty$" presupposes some way of specifying the cluster at time $-\infty$. We shall give a precise argument for

$$|t|X_1(t) \to c_1 \text{ a.s. for some constant } c_1 \geq 0\,, \tag{84}$$

and the general case can be done by induction: we omit details. For $\eta > 0$ define

$$m(\eta) = \limsup_{t\to-\infty} \max\{j : |t|X_j(t) > \eta\}\,,$$

with the convention max{empty set} = 0. Then $m(\eta) < \infty$ by (80) and $m(\eta)$ is constant by extremality. In fact, $m(\eta)$ is a decreasing, right-continuous step function. Define $c_1 = \sup\{\eta : m(\eta) = 1\}$. By definition of $m(\eta)$ we have

$$\limsup_{t\to-\infty} |t|X_1(t) = c_1\,.$$

Fix $\eta_1 < c_1$, a continuity point of $m(\cdot)$, with $m(\eta_1) = 1$. Define

$$T_n = \min\{t \geq -n : |t|X_1(t) \geq \eta_1\} \wedge -1$$

so that $T_n \downarrow -\infty$ by definition. At time T_n consider the largest cluster, and its subsequent growth as a distinguished cluster, its size denoted by $X_*(t)$. As before let $Y(t) = |t|X_*(t)$, $t < 0$. Define events

$$C_1(t') = \{|t|S(t) \leq 2\,, \forall t < t'\}\,,\quad C_2(t') = \{|t + \frac{1}{S(t)}| \leq 1\,, \forall t < t'\}\,,$$

$$C_3(t') = \{|t|^3 S_3(t) \leq a+1\,, \forall t < t'\}\,,\quad C_4(t') = \{|t|X_1(t) \leq c_1 + 1\,, \forall t < t'\}\,,$$

and $C(t') = \bigcap_{i=1}^4 C_i(t')$, so that $P(C(t')) \to 1$ as $t' \to -\infty$. For $t < t'$ and while on $C(t')$, the quantities (82, 83) are bounded in absolute value by

$$\alpha(t) = \frac{2(c_1+1)+a+1}{|t|^2}\,,\qquad \beta(t) = \frac{(c_1+1)(a+1)}{|t|^2}\,.$$

Thus, if $T_n < t'$, then for all $k \geq 1$

$$\int_{T_{n+k}}^{T_n} \alpha(t)dt = O\left(\frac{1}{|T_n|}\right) = O\left(\frac{1}{|t'|}\right) \quad\text{and}\quad \int_{T_{n+k}}^{T_n} \beta(t)dt = O\left(\frac{1}{|t'|}\right)\,.$$

Now Lemma 25(a) gives

$$P\left(\sup_{T_{n+k}\leq t\leq T_n} |Y(t) - Y(T_{n+k})| \geq \varepsilon\right) \leq 1 - P(C(t')) + P(T_n > t') + O\left(\frac{1}{|t'|}\right)\,.$$

Taking limits as $k \to \infty$, $n \to \infty$ and $t' \to \infty$, in this order, yields

$$P(\liminf_{t\to-\infty} |t|X_1(t) \leq c_1 - \varepsilon) = 0,\ \text{for all } \varepsilon > 0\,,$$

and (84) follows.

4 Proof of Theorems 2 - 4

Proposition 5 of [1] established the Feller property of the multiplicative coalescent as a $l^2_\searrow$-valued Markov process. Roughly speaking, Theorems 2 - 4 are easy consequences of Propositions 7 and 18 and the Feller property, though we shall see that a subtle complication arises.

We first record the following immediate consequence of the Feller property and the Kolmogorov extension theorem.

Lemma 26 *For $n = 1, 2, \ldots$ let $(\mathbf{X}^{(n)}(t), t \geq t_n)$ be versions of the multiplicative coalescent. Suppose $t_n \to -\infty$ and $\mathbf{X}^{(n)}(t) \ \stackrel{d}{\to}\ \mathbf{X}^{(\infty)}(t)$, say, for each fixed $-\infty < t < \infty$. Then there exists an eternal multiplicative coalescent $\mathbf{X}$ such that $\mathbf{X}(t) \ \stackrel{d}{=}\ \mathbf{X}^{(\infty)}(t)$ for each t.*

Proof of Theorem 2. Fix $(\kappa, 0, \mathbf{c}) \in \mathcal{I}$ and use Lemma 8 to choose $(\mathbf{x}^{(n)})$ satisfying $(18-20)$. Time-shift and regard Proposition 7 as specifying versions $(\mathbf{X}^{(n)}(t); t \geq -\frac{1}{\sigma_2(\mathbf{x}^{(n)})})$ of the multiplicative coalescent with initial states $\mathbf{X}^{(n)}(-\frac{1}{\sigma_2(\mathbf{x}^{(n)})}) = \mathbf{x}^{(n)}$. Then Proposition 7 asserts the existence, for fixed t, of the limit

$$\mathbf{X}^{(n)}(t) \ \stackrel{d}{\to}\ \mathbf{Z}(t).$$

By Lemma 26 there exists an eternal multiplicative coalescent $\mathbf{Z}$ with these marginal distributions. Define $\mu(\kappa, 0, \mathbf{c})$ as the distribution of this $\mathbf{Z}$, and for $-\infty < \tau < \infty$ define $\mu(\kappa, \tau, \mathbf{c})$ as the distribution of $(\mathbf{Z}(t-\tau), -\infty < t < \infty)$. $\square$

Before continuing to the proofs of Theorems 3 - 4, we record a few consequences of the Feller property.

Lemma 27 *Suppose $(\mathbf{X}^{(n)})$ are versions of the multiplicative coalescent such that $\mathbf{X}^{(n)}(t) \ \stackrel{d}{\to}\ \mathbf{X}(t)$ for each t. If $t_n \to t$ then $\mathbf{X}^{(n)}(t_n) \ \stackrel{d}{\to}\ \mathbf{X}(t)$.*

Remark. Conceptually, this holds because by general theory the Feller property implies weak convergence of processes in the Skorokhod topology. Rather than relying on such general theory (which is usually [12] developed in the locally compact setting: of course $l^2_\searrow$ isn't locally compact) we give a concrete argument.

Proof of Lemma 27. Write $||\mathbf{x}||$ for the l^2 norm. For any version of the multiplicative coalescent, $t \to ||\mathbf{X}(t)||$ is increasing and (by [1] Lemma 20) continuous in probability. Then convergence $||\mathbf{X}^{(n)}(t)|| \ \stackrel{d}{\to}\ ||\mathbf{X}(t)||$ easily implies $||\mathbf{X}^{(n)}(t_n)|| - ||\mathbf{X}^{(n)}(t)|| \ \stackrel{d}{\to}\ 0$. But then by [1] Lemma 17 (restated as Lemma 36(i) below) $||\mathbf{X}^{(n)}(t_n) - \mathbf{X}^{(n)}(t)|| \xrightarrow{d} 0$, establishing the lemma. $\square$

Recall that Proposition 7 was stated for finite-length initial states $\mathbf{x}^{(n)}$. The last step needed for the generalization to all $\mathbf{x}^{(n)} \in l^2_{\searrow}$ is the following

Lemma 28 *Suppose $\mathbf{x}^{(n)} \in l^2_{\searrow}$, $n \geq 1$ satisfies (18-20). Define $\mathbf{x}^{(n,k)}$ to be the truncated vector $(x_1^{(n)}, \ldots, x_k^{(n)})$ and $\mathbf{X}^{(n,k)}$ the corresponding multiplicative coalescent. Take $k = k(n) \to \infty$ sufficiently fast so that $(\mathbf{x}^{(n,k)}, n \geq 1)$ satisfies (18 - 20) with the same limits as does $(\mathbf{x}^{(n)}, n \geq 1)$, and so that Then there exists a coupling $(\mathbf{X}^{(n,k)}, \mathbf{X}^{(n)})$ such that*

$$\left\| \mathbf{X}^{(n)} \left(t + \frac{1}{\sigma_2(\mathbf{x}^{(n)})} \right) - \mathbf{X}^{(n,k)} \left(t + \frac{1}{\sigma_2(\mathbf{x}^{(n)})} \right) \right\| \xrightarrow{p} 0 \text{ as } n \to \infty .$$

The proof uses estimates derived later, and is given after the proof of Lemma 35.

Lemma 29 *Proposition 7 remains valid for any sequence $\mathbf{x}^{(n)} \in l^2_{\searrow}$ satisfying (18-20).*

Proof. Combining the conclusion of Proposition 7 for $(\mathbf{x}^{(n,k)}, n \geq 1)$ with Lemmas 27 and 28 gives the conclusion of Proposition 7 for $(\mathbf{x}^{(n)}, n \geq 1)$. $\square$

Proof of Theorems 3 and 4. The "if" part of Theorem 4 follows from Proposition 7 (in the extended setting of Lemma 29) by taking the initial state $\mathbf{x}^{(n)}$ in Proposition 7 to be $\mathbf{X}(-n)$.

Now consider an arbitrary extreme eternal non-constant multiplicative coalescent $\mathbf{X}$. Proposition 18 shows the existence as $t \to -\infty$ limits of constants $(\kappa, \tau, \mathbf{c})$, where $\kappa := a - \sum_i c_i^3 \geq 0$. Applying Proposition 7 with initial state $\mathbf{X}(-n)$, we see that $(\kappa, \tau, \mathbf{c}) \in \mathcal{I}$ and that $\mathbf{X}$ has distribution $\mu(\kappa, \tau, \mathbf{c})$. (Note that here we use the non-convergence part of Proposition 7.) It follows that the extreme points of the set of eternal multiplicative coalescents are $\{\mu(\kappa, \tau, \mathbf{c}) : (\kappa, \tau, \mathbf{c}) \in \mathcal{J}\} \cup \{\hat{\mu}(y) : 0 \leq y < \infty\}$ for some $\mathcal{J} \subseteq \mathcal{I}$. Now consider $\mathbf{X}$ with distribution $\mu(\kappa, \tau, \mathbf{c})$. As above, Proposition 18 implies the existence of some (maybe random) limits for the left sides of (10 – 12), but – and this is the subtle point – do not directly imply these limits are the specific constants asserted on the right side of (10 – 12). All we can deduce is that the distribution of $\mathbf{X}$ is a mixture:

$$\mu(\kappa, \tau, \mathbf{c}) = \int_{\mathcal{I}} \mu(\cdot) d\nu(\cdot), \ \nu \text{ a distribution on } \mathcal{I} \tag{85}$$

where ν is supported on $\mathcal{J} \subset \mathcal{I}$. We need to show the representation (85) holds only for the measure ν degenerate at the point $(\kappa, \tau, \mathbf{c})$. This will imply $\mathcal{J} = \mathcal{I}$ and establish the "only if" part of Theorem 4, completing the proof of Theorems 3 and 4.

Proposition 30 *Given $(\kappa, \tau, \mathbf{c}) \in \mathcal{I}$, the representation (85) holds only for the measure ν degenerate at the point $(\kappa, \tau, \mathbf{c})$.*

In principle it should be possible to prove Proposition 30 from Theorem 2 by showing that the parameters $\kappa, \tau, \mathbf{c}$ can be recovered as $t \to \pm\infty$ limits of functionals of the excursion lengths of $B^{\kappa, t-\tau, \mathbf{c}}$ (cf. Lemma 33 below). But it seems hard to recover τ in that manner. Instead we modify the proof of Proposition 18 to obtain Lemma 31 below.

Proof of Proposition 30. By scaling, we may assume $\tau = 0$ and $\kappa = 1$ or $\kappa = 0$. As at the beginning of this section, for initial states $(\mathbf{x}^{(n)})$ given by Lemma 8,

$$(\mathbf{X}^{(n)}(t); t \geq -\tfrac{1}{\sigma_2(\mathbf{x}^{(n)})}) \stackrel{d}{\to} (\mathbf{X}(t); -\infty < t < \infty)$$

where $\mathbf{X}$ has distribution $\mu(\kappa, \tau, \mathbf{c})$.

Lemma 31 *As $t \to -\infty$*

(i) $\quad \dfrac{S_3(t)}{S_2^3(t)} \to \kappa + \sum_i c_i^3 \ , \ a.s.$

(ii) $\quad t + \dfrac{1}{S_2(t)} \to \tau \ , \ a.s.$

(iii) $\quad$ *For each $j \geq 1$ with $c_j > 0$*

$$|t| X_k(t) \to c_j \quad a.s., \ . \tag{86}$$

for some finite integer-valued random variable $k \geq 1$.

Proof. Proposition 18 and representation (85) ensure the existence of the above limits as $t \to -\infty$. So it suffices to show the corresponding assertions where the limits are taken over a deterministic sequence of times $t_m \to -\infty$. Consider assertion (i). To simplify the notation, define $f : l^2_\searrow \to R^+$ by $f(\cdot) = \log \sigma_3(\cdot) - 3 \log \sigma_2(\cdot)$, and let $Y(t) = f(\mathbf{X}(t))$ and $Y^{(n)}(t) = f(\mathbf{X}^{(n)}(t))$. We want to check that $\lim_{t \to -\infty} Y(t) = \log(a)$, where $a = \kappa + \sum_i c_i^3$. We know $\lim_n Y^{(n)}(t) = Y(t)$ and $\lim_n Y^{(n)}(-1/\sigma_2(\mathbf{x}^{(n)})) = \lim_n f(\mathbf{x}^{(n)}) = \log(a)$. The idea of the proof is, of course, $Y(t) \approx Y^{(n)}(t) \approx$

$Y^{(n)}(-1/\sigma_2(\mathbf{x}^{(n)}))$ when t is large negative, and n is large . To make these notions precise, a simple adaptation of the argument in Proposition 24, using $\mathbf{X}^{(n)}(0) \stackrel{d}{\to} \mathbf{X}(0)$ as the only additional ingredient, gives

Corollary 32 *Interpret* $\mathbf{X}^{(m)}(t) = \mathbf{0}$ *for* $t < -\frac{1}{\sigma_2(\mathbf{x}^{(m)})}$. *Then*

$$\limsup_{t\to-\infty} \sup_{m\geq 1} |t| X_1^{(m)}(t) < \infty\,, \quad \lim_{t\to-\infty} \sup_{m\geq 1} \int_{-\infty}^{t} (X_1^{(m)}(u))^2\, du = 0 \quad a.s. \tag{87}$$

and

$$\lim_{t\to-\infty} \sup_{m\geq 1} \int_{-\infty}^{t} (S^{(m)}(u))^2\, du = 0 \quad a.s. \tag{88}$$

Proof. We modify the proof of (74) to show (87). Suppose

$$\limsup_{t\to-\infty} \sup_{m\geq 1} |t| X_1^{(m)}(t) = \infty\,,$$

fix a large constant M and define

$$T_n = \inf\{t \geq -n : \sup_{m\geq 1} |t| X_1^{(m)}(t) \geq M\}\,,$$

and provided $T_n < \infty$, let $k_n = \inf\{m \geq 1 : |T_n| X_1^{(m)}(T_n) \geq M\}$. Applying Lemma 23 to $(\mathbf{X}^{(k_n)}(T_n + u), u \geq 0)$ and $t = -T_n$, on the event $\{T_n < 0\}$, gives

$$\begin{aligned} &P(T_n < 0, S^{(k_n)}(0) \leq |T_n| X_1^{(k_n)}(T_n), X_2^{(k_n)}(0) \geq \delta) \\ &\quad \leq \delta^{-2} E |T_n| X_1^{(k_n)}(T_n) \exp(-\delta |T_n| X_1^{(k_n)}(T_n)). \end{aligned}$$

The supposition $C = \infty$ implies that $P(T_n < 0, |T_n| X_1^{(k_n)}(T_n) \geq M) \to 1$ as $n \to \infty$. Moreover $T_n \to -\infty$ implying $k_n \to \infty$. Let $n \to \infty$ to get (78). □

Now an application of Lemma 25(a), using Lemma 21 and (87), provides a sequence $t_m \downarrow -\infty$ such that

$$\limsup_{n\geq 1} P\left(\sup_{t\in[-1/\sigma_2(\mathbf{x}^{(n)}), t_m]} |Y^{(n)}(t) - Y^{(n)}(\,-1/\sigma_2(\mathbf{x}^{(n)})\,)\,| \geq \varepsilon_m \right) \leq \frac{1}{m^2}\,,$$

where ε_m is some fixed sequence, $\varepsilon_m \to 0$. Then after taking limits as $n \to \infty$, we get

$$P(|Y(t_m) - \log(a)| \geq \varepsilon_m) \leq 1/m^2$$

establishing (i). Now let $g(\cdot,t) = t + \frac{1}{\sigma_2(\cdot)}$, $Z(t) = g(\mathbf{X}(t),t)$, $Z^{(n)}(t) = g(\mathbf{X}^{(n)}(t),t)$. From part (i) we have

$$\limsup_n P\left(\sup_{u\in[-1/\sigma_2(\mathbf{x}^{(n)}),t]} S_3^{(n)}(u)/S_2^{(n)}(u) \geq a+1\right) \to 0 \quad \text{as } t \to -\infty.$$

Let $C_0 = 2(a+1)^2$. Using estimates in Lemma 20, and arguing as in the proof of Proposition 18, we see that

$$|E(\Delta Z^{(n)}(u)|\mathcal{F}(u))| \leq C_0\left(X_1^{(n)}(u)S^{(n)}(u) + (S^{(n)}(u))^3\right)du$$

and

$$\text{var }(\Delta Z^{(n)}(u)|\mathcal{F}(u)) \leq C_0(S^{(n)}(u))^2 du$$

for all $u \in [-\frac{1}{\sigma_2(\mathbf{x}^{(n)})}, t]$ and all large n, with probability converging to 1 as $t \to -\infty$. By Lemma 25(a) and Corollary 32, we find a sequence t_m so that

$$\limsup_{n\geq 1}\ P\left(\sup_{t\in[-1/\sigma_2(\mathbf{x}^{(n)}),t_m]} |Z^{(n)}(t) - Z^{(n)}(\,-1/\sigma_2(\mathbf{x}^{(n)})\,)\,| \geq \varepsilon_m\right) \leq \frac{1}{m^2}\,,$$

implying (ii) and

$$\limsup_{n\geq 1} P\left(\sup_{u\in[-\frac{1}{\sigma_2(\mathbf{x}^{(n)})},t]} \frac{|\,|u|S(u)-1|}{S(u)} \geq \varepsilon_m + |\tau|\right) \to 0\,, t \to -\infty. \tag{89}$$

The proof of (iii) is a similar "extension" of the corresponding argument in Proposition 18. Fix some integer $j \geq 1$, and consider $\mathbf{X}^{(n)}$ for n such that the initial state $\mathbf{x}^{(n)}$ in Lemma 8 has $l(n) \geq j$. For simplicity assume $c_j \neq c_k$, $j \neq k$. Follow the initial cluster of size $c_j/n^{1/3}$ as it merges with other clusters, and denote its size at time t by $X_*^{(n)}(t)$, $t > -\frac{1}{\sigma_2(\mathbf{x}^{(n)})}$. Using (89), we are able to write $E(\Delta(|t|X_*^{(n)}(t))|\mathcal{F}(t))$ and var $(\Delta(|t|X_*^{(n)}(t))|\mathcal{F}(t))$ as $O(X_*^{(n)}(t)S^{(n)}(t)+(X_*^{(n)})^2(t))$, uniformly in n (recall the discussion following (82,83)). Hence

$$\limsup_{n\geq 1}\ P\left(\sup_{t\in[-1/\sigma_2(\mathbf{x}^{(n)}),t_m]} |\,|t|X_*^{(n)}(t) - c_j\,| \geq \varepsilon_m\right) \leq \frac{1}{m^2}\,, \tag{90}$$

for some sequence $t_m \to -\infty$. Let $k_m^{(n)} = \sup\{i \geq 1 : X_i^{(n)}(t_m) \geq X_*^{(n)}(t_m)\}$ be the *rank* of the distinguished cluster. We may assume $\varepsilon_m < c_j/2$ so that $k_m^{(n)}$ is tight since $\mathbf{X}^{(n)}(t_m) \xrightarrow{d} \mathbf{X}(t_m) \in l^2_{\searrow}$. In fact

$$\lim_{M\to\infty} \limsup_n P(\min_{1\leq i\leq M} |\,|t|X_i^{(n)}(t_m) - c_j\,| \geq \varepsilon_m) \leq \frac{1}{m^2}\ , \text{ implying}$$

$$\lim_{M\to\infty} P(\min_{1\leq i\leq M} |\,|t|X_i(t_m) - c_j\,| \geq \varepsilon_m) \leq \frac{1}{m^2}\,.$$

In words, at least one component of $\mathbf{X}(t_m)$ falls into the ε_m neighborhood of c_j with probability at least $1 - 1/m^2$. Denote by k_m the rank of largest such component, and set $k_m = \infty$ if no such component exist. Now we may take $k = \liminf_m k_m$.

In the general case with ties $c_{j-1} > c_j = c_{j+1} = \ldots = c_{j+r-1} > c_{j+r}$ for some finite $r \geq 1$, one needs to follow all r clusters of size $c_j/n^{1/3}$ simultaneously, but since with large probability no pair of these clusters coalesces by time t_m (easy!) the argument does not change essentially. $\square$

Lemma 33 *Let $\psi(t)$ denote the length of the largest excursion of $B^{\kappa,t-\tau,\mathbf{c}}$ with left end-point in $[0,\frac{2}{\kappa}t]$. Then*

$$t^{-1}\psi(t) \xrightarrow{d} 2/\kappa \text{ as } t \to +\infty.$$

Proof. The excursions of $B^{\kappa,t-\tau,\mathbf{c}}$ away from 0 are the excursions of $W = W^{\kappa,t-\tau,\mathbf{c}}$ above past minima. In the sequel, we simply call them *excursions.* Recall the representation (40) from section 2.5. Denote by $\widetilde{W}(s)$ the rescaled process

$$\frac{1}{t^2}W(s\,t) = s - \frac{1}{2}\kappa s^2 + \kappa^{1/2}\frac{1}{t^2}W^*(s\,t) + \frac{1}{t^2}V^{\mathbf{c}}(s\,t) - \frac{\tau}{t}\,.$$

Then the excursion-length process of $\widetilde{W}(\cdot)$ is precisely the excursion-length process of $W(\cdot)$, shrunk by a factor $\frac{1}{t}$. Let $\tilde{\psi}_t$ be the length of the largest $\widetilde{W}$-excursion that started in $[0, 2/\kappa]$. The assertion of the lemma is equivalent to the assertion $\tilde{\psi}_t \xrightarrow{d} \frac{2}{\kappa}$ as $t \to +\infty$. This reduces to showing

$$\sup_{s\in[0,\alpha]} |\widetilde{W}(s) - (s - \tfrac{1}{2}\kappa s^2)| \xrightarrow{d} 0\,,\ t\to\infty \ \text{ for any fixed } \alpha > 0\,.$$

Since $\sup_{s\in[0,\alpha]} W^*(st)/t^2 \xrightarrow{d} 0$ as $t \to +\infty$ it suffices to check

$$\sup_{s\in[0,\alpha]} V^{\mathbf{c}}(st)/t^2 \xrightarrow{d} 0$$

by showing

$$\sup_{s\in[0,\alpha]} M^{\mathbf{c}}(st)/t^2 \xrightarrow{d} 0\,, \quad A^{\mathbf{c}}(\alpha t)/t^2 \xrightarrow{d} 0 \quad \text{as} \quad t \to +\infty\,.$$

But this is easy (cf. Proposition 14). □

Remark. $\psi(t)$ depends on both the lengths and the order in which excursions of $B^{\kappa,t-\tau,\mathbf{c}}$ appear. But since components appear in size-biased order in the breadth-first walk, Proposition 9 implies the excursions of $B^{\kappa,t-\tau,\mathbf{c}}$ appear in size-biased order (in the l^2 sense of section 2.6). Therefore the distribution of $\psi(t)$ is determined by the distribution of the decreasing-ordered vector $\mathbf{X}(t)$ of excursion lengths.

Proof of Proposition 30, continued. Given $\mathbf{x}, \mathbf{x}' \in l^3_{\searrow}$, write $\mathbf{x} \subseteq \mathbf{x}'$ if $\{\mathbf{x}\} \subseteq \{\mathbf{x}'\}$, where $\{\mathbf{x}\}$ denotes the multiset $\{x_1, x_2, \ldots\}$ of $\mathbf{x}$-coordinates, and the relation $\subseteq$ is in terms of mutisets. Let $a = \kappa + \sum_i c_i^3$. Due to Lemma 31, (85) becomes

$$\mu(\kappa,\tau,\mathbf{c}) = \int_{\mathcal{J}_1} \mu(\kappa',\tau',\mathbf{c}')d\nu(\kappa',\tau',\mathbf{c}'),$$

where $\mathcal{J}_1 = \mathcal{J} \cap \{(\kappa',\tau,\mathbf{c}') : \kappa' + \sum_i {c'_i}^3 = a,\ \mathbf{c} \subseteq \mathbf{c}',\ 0 \le \kappa' \le \kappa\}$. If $\kappa = 0$, then all points in $\mathcal{J}_1$ have the first coordinate $\kappa\,(=0)$. The same is true, by Lemma 33, for all positive κ. But $\kappa' = \kappa$ implies $\mathbf{c}' = \mathbf{c}$, so (12) holds with the original $\kappa, \tau, \mathbf{c}$, and the mixing measure ν in (85) is concentrated on the point $\{(\kappa,\tau,\mathbf{c})\}$.

5 The coloring construction

In this section we develop the coloring construction outlined in section 1.4, providing the basis for the proof of Theorem 6.

For $\mathbf{x}, \mathbf{y} \in l^2_{\searrow}$ write $\mathbf{x} \leq \mathbf{y}$ for coordinatewise inequality $x_i \leq y_i \ \forall i$. Write $\mathbf{x} \preceq \mathbf{y}$ if $\mathbf{x}$ is the decreasing ordering of $\{y_{i,j}, i, j \geq 1\}$, where $\sum_j y_{i,j} \leq y_i \ \forall i$. Note that either inequality implies $||\mathbf{x}|| \leq ||\mathbf{y}||$, where $||\cdot||$ is the l^2 norm.

5.1 Graphical construction of $\mathrm{COL}(\mathbf{X}; \mathbf{c})$

We first show that the graphical construction ([1] section 4) of the multiplicative coalescent can be extended to give a construction of the process $\mathrm{COL}(\mathbf{X}; \mathbf{c})$ described informally in section 1.4. Fix nonnegative sequences $\mathbf{x} = (x_i)$ and $\mathbf{c} = (c_k)$, and fix $t \geq 0$.

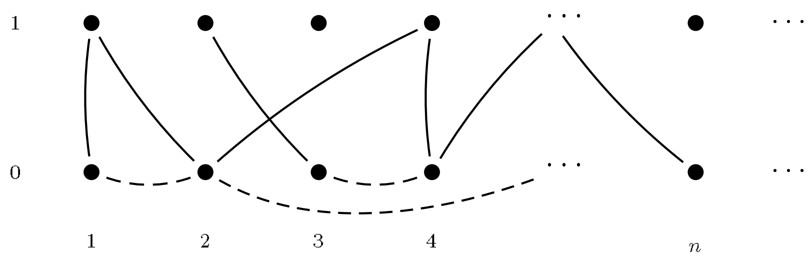

Figure 2

Consider the random graph, illustrated in figure 2, on vertex-set $\{(i, 0), i \geq 1\} \cup \{(k, 1), k \geq 1\}$ with

$$\text{(for } i, j \geq 1\text{) a Poisson}(x_i x_j t) \text{ number of edges } (i, 0) \leftrightarrow (j, 0) \tag{91}$$

$$\text{(for } i, k \geq 1\text{) a Poisson}(x_i c_k) \text{ number of edges } (i, 0) \leftrightarrow (k, 1) \tag{92}$$

where the Poisson numbers are independent for different pairs. All edges connecting a pair of vertices are represented by one single edge in the figure. For each connected component $\mathcal{C}$ of this graph, calculate the "weight" $w(\mathcal{C}) = \sum_{i:(i,0)\in\mathcal{C}} x_i$, and temporarily write $\mathbf{Y}(t)$ for the sequence of component-weights, in decreasing order. At time $t = 0$ only solid line edges (92) exist. Write $\mathrm{COL}(\mathbf{x}; \mathbf{c}) = \mathbf{Y}(0)$. As t varies, link the Poisson random variables in (91) via the natural Poisson processes. New edges (91) are the dashed lines in the figure. Provided we prove $\mathbf{Y}(t)$ is $l^2_{\searrow}$-valued (see

Lemma 34 below), it is clear that $\mathbf{Y}(t)$ evolves as a multiplicative coalescent, because components $\mathcal{C}_a, \mathcal{C}_b$ merge when an edge $(i,0) \leftrightarrow (j,0)$ appears for some $(i,0) \in \mathcal{C}_a$, $(j,0) \in \mathcal{C}_b$, which occurs at rate $\sum_{(i,0)\in\mathcal{C}_a} \sum_{(j,0)\in\mathcal{C}_b} x_i x_j = w(\mathcal{C}_a) w(\mathcal{C}_b)$.

To link this construction to the "coloring" description in section 1.4, first consider $t = 0$. For each k, regard each cluster i as containing a number of "atoms of color k" instead of a number of edges $(i,0) \leftrightarrow (k,1)$. Do the minimal amount of coalescence of clusters required to ensure that all similarly-colored atoms are in the same cluster. Then $\mathrm{COL}(\mathbf{x};\mathbf{c})$ is the vector of cluster-sizes.

Now consider $t > 0$. We could obtain $\mathbf{Y}(t)$ by first drawing the edges (92) and then the edges (91). Because (91) is the original graphical construction of the multiplicative coalescent, this means we are running the multiplicative coalescent from state $\mathrm{COL}(\mathbf{x};\mathbf{c})$ for time t. Alternatively, we could first draw the edges (91) and then the edges (92). This means we run the multiplicative coalescent from initial state $\mathbf{x}$ to obtain $\mathbf{X}(t)$, and then apply the coloring construction $\mathbf{x}' \to \mathrm{COL}(\mathbf{x}';\mathbf{c})$. (In both cases we appeal to the additivity property of the Poisson semigroup). In brief, we say that coloring commutes with the evolution of the multiplicative coalescent, and we may unambiguously write $\mathrm{COL}(\mathbf{X}(t);\mathbf{c})$ for $\mathbf{Y}(t)$.

Lemma 34 *If* $\mathbf{c} \in l^2_\searrow$ *and* $\mathbf{X}(0) \in l^2_\searrow$ *then the process* $(\mathrm{COL}(\mathbf{X}(t);\mathbf{c}), t \geq 0)$ *given by the graphical construction above is distributed as the multiplicative coalescent.*

Proof. By the discussion above, it suffices to show that $\mathrm{COL}(\mathbf{X}(t);\mathbf{c})$ takes values in $l^2_\searrow$, and since $\mathbf{X}(t)$ takes values in $l^2_\searrow$ we only need to show that $\mathrm{COL}(\mathbf{x};\mathbf{c})$ takes values in $l^2_\searrow$ when $\mathbf{x} \in l^2_\searrow$. Consider again the random graph defined by relations (91) and (92), with t set to 1, and add

$$\text{(for } k,l \geq 1) \text{ a Poisson}(c_k c_l) \text{ number of edges } (k,1) \leftrightarrow (l,1).$$

For each connected component $\mathcal{C}$ of the new graph define *full weight* by $w^f(\mathcal{C}) = \sum_{i:(i,0)\in\mathcal{C}} x_i + \sum_{k:(k,1)\in\mathcal{C}} c_k$, and let $\mathbf{Z}(1)$ be the vector of full weights in decreasing order. Then $\mathbf{Z}(1)$ is distributed as the time-1 distribution of the multiplicative coalescent started at time 0 with configuration $\mathbf{x} \bowtie \mathbf{c} \in l^2_\searrow$, where the "join" $\mathbf{b} \bowtie \mathbf{c}$ is defined by

$$\mathbf{b} \bowtie \mathbf{c} \text{ is the decreasing ordering of the multiset } \{b_i, i \geq 1\} \cup \{c_i, i \geq 1\}.$$

In particular, $\mathbf{Z}(1)$ takes values in $l^2_{\searrow}$. Since $\mathrm{COL}(\mathbf{x};\mathbf{c}) \preceq \mathbf{Z}(1)$ a.s. , it follows that $\mathrm{COL}(\mathbf{x};\mathbf{c}) \in l^2_{\searrow}$. $\square$
We will sometimes say "the color-k cluster of $\mathrm{COL}(\mathbf{X}(t);\mathbf{c})$", which is formally the union of clusters $X_i(t)$ for which $(i,0)$ is in the same component as $(k,1)$ in the graphical construction.

It is easy to see from the graphical construction that

$$\mathrm{COL}(\mathrm{COL}(\mathbf{x};\tilde{\mathbf{c}});\hat{\mathbf{c}}) \ \stackrel{d}{=}\ \mathrm{COL}(\mathbf{x};\tilde{\mathbf{c}} \bowtie \hat{\mathbf{c}}) \tag{93}$$

and hence

$$\mathrm{COL}(\mathrm{COL}(\mathbf{X}(t);\tilde{\mathbf{c}});\hat{\mathbf{c}}) \ \stackrel{d}{=}\ \mathrm{COL}(\mathbf{X}(t);\tilde{\mathbf{c}} \bowtie \hat{\mathbf{c}}). \tag{94}$$

For $\mathbf{x},\mathbf{c}$ as above, let $i \sim j$ mean that $(i,0)$ and $(j,0)$ are merged together within a component of $\mathrm{COL}(\mathbf{x};\mathbf{c})$, and let $i \stackrel{c}{\sim} j$ mean that $(i,0)$ and $(j,0)$ share a common color. Clearly $\{i \stackrel{c}{\sim} j\} \subseteq \{i \sim j\}$, for all i, j. We will need the following

Lemma 35 *(i)* $P(i \sim j) \geq P(i \stackrel{c}{\sim} j) = 1 - \prod_k \{1-(1-e^{-x_i c_k})(1-e^{-x_j c_k})\}$
(ii) If $\|\mathbf{c}\|\,\|\mathbf{x}\| < 1$ *, then for* $i \neq j$

$$P(i \sim j) \leq x_i x_j \frac{\|\mathbf{c}\|^2}{1-\|\mathbf{c}\|^2\|\mathbf{x}\|^2}\,.$$

Proof. Assertion (i) is immediate from the construction. For (ii) we have $i \sim j$ iff there exists a finite path in the $\mathrm{COL}(\mathbf{x};\mathbf{c})$-graph which connects vertices $(i,0)$ and $(j,0)$, for instance a path

$$(i,0) \leftrightarrow (k_1,1) \leftrightarrow (i_1,0) \leftrightarrow (k_2,1) \leftrightarrow (i_2,0) \leftrightarrow \ldots (i_{m-1},0) \leftrightarrow (k_m,1) \leftrightarrow (j,0)$$

of length $2m$. The probability of an edge $(i_l,0) \leftrightarrow (k,1)$ is less than $x_{i_l} c_k$, so the chance that the particular path above exists is at most

$$x_i x_j c_{k_1}^2 c_{k_2}^2 \ldots c_{k_m}^2 x_{i_1}^2 \ldots x_{i_{m-1}}^2\,.$$

Summing over all paths of length $2m$ gives

$$P(i \sim j \text{ via a path of length } 2m) \leq x_i x_j \|\mathbf{c}\|^{2m} \|\mathbf{x}\|^{2m-2}.$$

Summing over $m \geq 1$ establishes (ii). $\square$

We digress to give a defered proof.

Proof of Lemma 28. Write $\sigma_2^{(n)}$ for $\sigma_2(\mathbf{x}^{(n)})$, let $\mathbf{y}^{(n,k)} = (x_{k+1}^{(n)}, x_{k+2}^{(n)}, \ldots)$ and let $\mathbf{Y}^{(n,k)}$ be a multiplicative coalescent with initial state $\mathbf{y}^{(n,k)}$, independent of $\mathbf{X}^{(n,k)}$. From the assumptions of the lemma we have $\|\mathbf{y}^{(n,k)}\|/\sigma_2^{(n)} \to 0$ as $n \to \infty$. Let $\{y_i^{(n,k)} \sim y_j^{(n,k)}\}$ be the event that the i'th and the j'th cluster of $\mathbf{y}^{(n,k)}$ become merged within the same component of $\mathbf{Y}^{(n,k)}(t + 1/\sigma_2^{(n)})$. Since $\{y_i^{(n,k)} \sim y_j^{(n,k)}\}$ happens iff a finite path in the multiplicative coalescent graph connects i and j, and the probability of a link $l \leftrightarrow m$ occurring is bounded by $(t + 1/\sigma_2^{(n)}) y_l^{(n,k)} y_m^{(n,k)}$, we obtain a bound similar to the one in Lemma 35(ii)

$$P(y_i^{(n,k)} \sim y_j^{(n,k)}) \leq \frac{y_i^{(n,k)} y_j^{(n,k)} (t + 1/\sigma_2^{(n)})}{1 - (t + 1/\sigma_2^{(n)})^2 \, \|\mathbf{y}^{(n,k)}\|^2},$$

for all $j > i \geq 1$. Therefore

$$(t + 1/\sigma_2^{(n)}) \, \|\mathbf{Y}^{(n,k)}(t + 1/\sigma_2^{(n)})\| \to 0 \text{ in } L^2 \text{ as } n \to \infty. \tag{95}$$

Write $\mathbf{X}^{(n,k)}$ and $\mathbf{Y}^{(n,k)}$ for $\mathbf{X}^{(n,k)}(t + 1/\sigma_2^{(n)})$ and $\mathbf{Y}^{(n,k)}(t + 1/\sigma_2^{(n)})$, and write $\mathcal{F}^{(n,k)}$ for the σ-field generated by $\mathbf{X}^{(n,k)}, \mathbf{Y}^{(n,k)}$. By the random graph construction, a realization of $\mathbf{X}^{(n)}(t + 1/\sigma_2^{(n)})$ can be obtained from $\mathbf{X}^{(n,k)}$ and $\mathbf{Y}^{(n,k)}$ by appending independent edges connecting components $X_i^{(n,k)}$ and $Y_j^{(n,k)}$ at rate $(t + 1/\sigma_2^{(n)}) X_i^{(n,k)} Y_j^{(n,k)}$ for $i, j \geq 1$, and merging the connected components. The conditional expected second moment mass increase

$$E(S_2^{(n)}(t + 1/\sigma_2) - S_2^{(n,k)}(t + 1/\sigma_2) \mid \mathcal{F}^{(n,k)})$$

is bounded by

$$\begin{aligned} \|\mathbf{Y}^{(n,k)}\|^2 \; &+ \; 2 \sum_{i<j} Y_i^{(n,k)} Y_j^{(n,k)} P(Y_i^{(n,k)} \sim Y_j^{(n,k)} \mid \mathcal{F}^{(n,k)}) \\ &+ \; 2 \sum_{i<j} X_i^{(n,k)} X_j^{(n,k)} P(X_i^{(n,k)} \sim X_j^{(n,k)} \mid \mathcal{F}^{(n,k)}) \\ &+ \; 2 \sum_{i,j} X_i^{(n,k)} Y_j^{(n,k)} P(X_i^{(n,k)} \sim Y_j^{(n,k)} \mid \mathcal{F}^{(n,k)}). \end{aligned}$$

Since $\mathbf{X}^{(n,k)}$ is tight, by (95) we have $(t + 1/\sigma_2) \, \|\mathbf{X}^{(n,k)}\| \, \|\mathbf{Y}^{(n,k)}\| \xrightarrow{p} 0$, and we can apply again the reasoning of Lemma 35(ii) to bound the conditional probabilities above. For example, if we consider $\{Y_i^{(n,k)} \sim Y_j^{(n,k)}\}$ then

$(t+1/\sigma^{(n)})\mathbf{X}^{(n,k)}$ plays the role of the color weights $\mathbf{c}$ in Lemma 35(ii). We conclude (omitting the tedious details) that the conditional probabilities above are respectively $O(Y_i^{(n,k)}Y_j^{(n,k)}\ (t+1/\sigma_2)^2\ \|\mathbf{X}^{(n,k)}\|^2)$, $O(X_i^{(n,k)}X_j^{(n,k)}$ $(t+1/\sigma_2)^2\ \|\mathbf{Y}^{(n,k)}\|^2)$ and $O(X_i^{(n,k)}Y_j^{(n,k)}(t+1/\sigma_2))$, uniformly over all pairs i,j. Together with (95), this establishes the lemma.

5.2 Technical lemmas

For subsequent arguments we need two technical facts. The first lemma gives some topological properties of $(l^2_\searrow, ||\cdot||)$.

Lemma 36 *(i) If* $\mathbf{x} \preceq \mathbf{y}$ *then* $||\mathbf{y}-\mathbf{x}||^2 \leq ||\mathbf{y}||^2 - ||\mathbf{x}||^2$.

(ii) For each $\mathbf{y} \in l^2_\searrow$ *the set* $\{\mathbf{x} : \mathbf{x} \preceq \mathbf{y}\}$ *is pre-compact.*

(iii) Given a sequence $(\mathbf{x}^n, n \geq 1)$ *define* $\mathbf{x}^{[m]}$ *by:* $\mathbf{x}_i^{[m]} = \sup_{n\geq m} x_i^n$. *If* $\mathbf{y}^k \to \mathbf{x}$ *then there exists a subsequence* $\mathbf{x}^n = \mathbf{y}^{k_n}$ *such that* $\mathbf{x}^{[m]} \to \mathbf{x}$.

Proof. Assertion (i) is the (easy) Lemma 17 of [1]. For (ii), if $y_{i,j} \geq 0$ and $\sum_j y_{i,j} \leq y_i$ then $\sum_j y_{i,j}^2 1_{(y_{i,j}\leq\varepsilon)} \leq \varepsilon y_i$. Since also $\sum_j y_{i,j}^2 \leq y_i^2$, we have

$$\sum_j y_{i,j}^2 1_{(y_{i,j}\leq\varepsilon)} \leq \min(\varepsilon y_i, y_i^2).$$

So for fixed $\mathbf{y}$,

$$\sup_{\mathbf{x}\preceq\mathbf{y}} \sum_j x_j^2\ 1_{(x_j\leq\varepsilon)} \leq \varepsilon \sum_{i\leq k} y_i\ +\ \sum_{i>k} y_i^2$$

for any k. Letting $\varepsilon \to 0$ and then $k \to \infty$

$$\lim_{\varepsilon\to 0} \sup_{\mathbf{x}\preceq\mathbf{y}} \sum_j x_j^2\ 1_{(x_j\leq\varepsilon)} = 0.$$

By monotonicity, $\mathbf{x} \preceq \mathbf{y}$ implies $x_k \leq \varepsilon$ for $k \geq \varepsilon^{-2}||\mathbf{y}||^2$, and so

$$\lim_{k\to\infty} \sup_{\mathbf{x}\preceq\mathbf{y}} \sum_{j\geq k} x_j^2\ = 0.$$

This and norm-boundedness imply pre-compactness. For (iii), by replacing $\mathbf{y}^k$ by $\max(\mathbf{y}^k, \mathbf{x})$ we may assume $\mathbf{y}^k \geq \mathbf{x}$, and then by subtracting $\mathbf{x}$ we may assume $\mathbf{x} = 0$. But in that case $\mathbf{x}^{[m]} \leq \sum_{n\geq m} \mathbf{x}^n$, so we need only choose the subsequence $\mathbf{x}^n$ so that $\sum_n ||\mathbf{x}^n|| < \infty$. $\square$

Lemma 36(i) and monotone convergence easily imply that for $\mathbf{x}, \mathbf{c} \in l^2_\searrow$

$$\mathrm{COL}(\mathbf{x}; (c_1, \ldots, c_n, 0, \ldots))\ \xrightarrow{d}\ \mathrm{COL}(\mathbf{x}; \mathbf{c}).$$

The next lemma provides a more general result on joint continuity.

Lemma 37 *Let* $\mathbf{X}^n$, $\mathbf{C}^n$ *be two sequences of* $l^2_\searrow$*-valued random variables such that* $\mathbf{X}^n \ \stackrel{d}{\to}\ \mathbf{X}$ *and* $\mathbf{C}^n \to \mathbf{C}$ *a.s.. Then*

$$\mathrm{COL}(\mathbf{X}^n;\mathbf{C}^n)\ \stackrel{d}{\to}\ \mathrm{COL}(\mathbf{X};\mathbf{C}).$$

Proof. First observe the following. If $\mathbf{x}' \leq \mathbf{x}''$ and $\mathbf{c}' \leq \mathbf{c}''$, then the graphical constructions (91,92) for $(\mathbf{x}',\mathbf{c}')$ and $(\mathbf{x}'',\mathbf{c}'')$ may be coupled in such a way that the components $\mathcal{C}'$ are a refinement of the components $\mathcal{C}''$, and therefore $\mathrm{COL}(\mathbf{x}';\mathbf{c}') \preceq \mathrm{COL}(\mathbf{x}'';\mathbf{c}'')$.

To prove the lemma, it suffices to consider deterministic $\mathbf{x}^n \to \mathbf{x}$ and $\mathbf{c}^n \to \mathbf{c}$. By the subsequence method and Lemma 36(iii) we may assume $\mathbf{x}^{[m]} \to \mathbf{x}$ and $\mathbf{c}^{[m]} \to \mathbf{c}$ in the notation of that lemma, where $\mathbf{x}^{[m]} \geq \mathbf{x}$ and $\mathbf{c}^{[m]} \geq \mathbf{c}$. Make the coupling mentioned above, and apply Lemma 36(i) to obtain

$$||\mathrm{COL}(\mathbf{x}^{[m]};\mathbf{c}^{[m]}) - \mathrm{COL}(\mathbf{x};\mathbf{c})||^2 \leq ||\mathrm{COL}(\mathbf{x}^{[m]};\mathbf{c}^{[m]})||^2 - ||\mathrm{COL}(\mathbf{x};\mathbf{c})||^2.$$

The same argument gives another coupling for which

$$||\mathrm{COL}(\mathbf{x}^{[m]};\mathbf{c}^{[m]}) - \mathrm{COL}(\mathbf{x}^m;\mathbf{c}^m)||^2 \leq ||\mathrm{COL}(\mathbf{x}^{[m]};\mathbf{c}^{[m]})||^2 - ||\mathrm{COL}(\mathbf{x}^m;\mathbf{c}^m)||^2.$$

From the hypothesis of convergence $\mathbf{x}^m \to \mathbf{x}, \mathbf{c}^m \to \mathbf{c}$, these couplings have the property that any pair of clusters (x_i, x_j) of $\mathbf{x}$ which are joined in $\mathrm{COL}(\mathbf{x};\mathbf{c})$ will be joined in $\mathrm{COL}(\mathbf{x}^m;\mathbf{c}^m)$ for all sufficiently large m. It follows via Fatou's lemma that

$$\liminf_m ||\mathrm{COL}(\mathbf{x}^m;\mathbf{c}^m)|| \geq ||\mathrm{COL}(\mathbf{x};\mathbf{c})||.$$

Combining these inequalities, we see that the lemma reduces to proving

Lemma 38

$$||\mathrm{COL}(\mathbf{x}^{[m]};\mathbf{c}^{[m]})|| \to ||\mathrm{COL}(\mathbf{x};\mathbf{c})||. \tag{96}$$

We start by proving the finite-length case.

Lemma 39 *Let* $\mathbf{x}^{[m]}, \mathbf{x}$ *be as above, and let* $\mathbf{c}^{[m]} = (c_1^{[m]},\ldots,c_n^{[m]},0,\ldots)$, $\mathbf{c} = (c_1,\ldots,c_n,0,\ldots)$. *Then*

$$||\mathrm{COL}(\mathbf{x}^{[m]};\mathbf{c}^{[m]})|| \to ||\mathrm{COL}(\mathbf{x};\mathbf{c})|| \text{ and}$$

$$\mathrm{COL}(\mathbf{x}^{[m]};\mathbf{c}^{[m]})\ \stackrel{d}{\to}\ \mathrm{COL}(\mathbf{x};\mathbf{c})\,.$$

Proof. First let $\mathbf{c}^{[m]} = (c^{[m]}, 0, \ldots)$, and think in terms of the (coupled) simple coloring construction of section 1.4.

$$\begin{aligned} E\|\mathrm{COL}(\mathbf{x}^{[m]};\mathbf{c}^{[m]})\|^2 &= \sum_i (x_i^{[m]})^2 + \sum_{i\neq j} x_i^{[m]} x_j^{[m]} (1-e^{-c^{[m]}x_i^{[m]}})(1-e^{-c^{[m]}x_j^{[m]}}) \\ &\to \sum_i x_i^2 + \sum_{i\neq j} x_i x_j (1-e^{-cx_i})(1-e^{-cx_j}) \\ &= E\|\mathrm{COL}(\mathbf{x};\mathbf{c})\|^2 , \end{aligned}$$

by Lemma 36(ii). Since $\|\mathrm{COL}(\mathbf{x}^{[m]};\mathbf{c}^{[m]})\| \geq \|\mathrm{COL}(\mathbf{x};\mathbf{c})\|$ always, the first assertion of Lemma 39 is true in this case, and therefore

$$\mathrm{COL}(\mathbf{x}^{[m]};(c^{[m]},0,\ldots)) \to \mathrm{COL}(\mathbf{x};(c,0,\ldots)) \text{ a.s. },$$

as discussed above. The rest follows by repeated application of (93) and induction. $\Box$

Proof of Lemma 38. To shorten the notation let

$$\mathbf{c}^{(n)} = (c_1,\ldots,c_n,0,\ldots)\ ,\ \mathbf{c}^{[m](n)} = (c_1^{[m]},\ldots,c_n^{[m]},0,\ldots)$$

and let

$$\mathbf{X}^{(n)} \overset{d}{=} \mathrm{COL}(\mathbf{x};\mathbf{c}^{(n)})\ ,\ \mathbf{X} = \mathrm{COL}(\mathbf{X}^{(n)};(c_{n+1},c_{n+2},\ldots))\ ,$$

$$\mathbf{X}^{[m](n)} \overset{d}{=} \mathrm{COL}(\mathbf{x}^{[m]};\mathbf{c}^{[m](n)})\,,\mathbf{X}^{[m]} = \mathrm{COL}(\mathbf{X}^{[m](n)};(c_{n+1}^{[m]},c_{n+2}^{[m]},\ldots))$$

be coupled in such a way that

$$\mathbf{X}^{(n)} \preceq \mathbf{X}^{[m_2](n)} \preceq \mathbf{X}^{[m_1](n)} \quad \text{for all } 1 \leq m_1 \leq m_2, 1 \leq n. \tag{97}$$

From Lemma 39 and (97) we deduce

$$\liminf_{m\to\infty} \|\mathbf{X}^{[m]}\| \geq \|\mathbf{X}\|\,.$$

At the same time, by the nesting (97)

$$B_n := \bigcap_m \{\, \|\mathbf{X}^{[m](n)}\| \sum_{k=n+1}^{\infty} (c_k^{[m]})^2 < 1/2\} \supseteq \{\, \|\mathbf{X}^{[1]}\| \sum_{k=n+1}^{\infty} (c_k^{[1]})^2 < 1/2\}\,,$$

so that $P(B_n) \to 1$. While on B_n, we are able to use estimate Lemma 35(ii) to bound the distance between $\mathbf{X}^{[m]}$ and $\mathbf{X}^{[m](n)}$. Precisely,

$$P(i \sim j \text{ in } \mathbf{X}^{[m]} \mid \mathbf{X}^{[m](n)}) \leq 2\, X_i^{[m](n)}\, X_j^{[m](n)} \sum_{k=n+1}^{\infty} (c_k^{[1]})^2\,,$$

therefore $E(\|\mathbf{X}^{[m]}\|^2 - \|\mathbf{X}^{[m](n)}\|^2 \mid \mathbf{X}^{[m](n)})$

$$\begin{aligned}
&\leq 2\sum_i \sum_{j\neq i} (X_i^{[m](n)})^2 (X_j^{[m](n)})^2 \sum_{k=n+1}^{\infty} (c_k^{[1]})^2 \\
&\leq 2\|\mathbf{X}^{[1]}\|^4 \sum_{k=n+1}^{\infty} (c_k^{[1]})^2 ,
\end{aligned}$$

implying specially, for any $\varepsilon > 0$,

$$\begin{aligned}
\limsup_{m\to\infty} P(\|\mathbf{X}^{[m]}\|^2 - \|\mathbf{X}^{[m](n)}\|^2 > \varepsilon) &\leq P(B_n^c) \\
&+ \varepsilon^{-1} E\left(1 \wedge 2\|\mathbf{X}^{[1]}\|^4 \sum_{k=n+1}^{\infty} (c_k^{[1]})^2\right) .
\end{aligned}$$

The assertion of the lemma follows from Lemma 39 by letting $n \to \infty$. □

5.3 Coloring and $t \to -\infty$ limits

From an eternal multiplicative coalescent $(\mathbf{X}(t), -\infty < t < \infty)$ and $\mathbf{c} \in l^2_{\searrow}$ we may construct another eternal multiplicative coalescent $\mathrm{COL}(\mathbf{X};\mathbf{c}) = (\mathrm{COL}(\mathbf{X}(t);\mathbf{c}), -\infty < t < \infty)$, using Lemma 34 and the Kolmogorov extension. We will assume without explicit comment that $\mathbf{X}(t) \to 0$ as $t \to -\infty$, so that by Proposition 18 the $t \to -\infty$ limits $(a, \tau, \mathbf{c})$ in (63 – 65) exist for $\mathbf{X}$. In the non-extreme case these limits may be random, but even in that case (64) implies

$$|t| S(t) \xrightarrow{p} 1 \text{ as } t \to -\infty. \tag{98}$$

We now study the connection between the $\mathbf{c}$ in the coloring construction and the $\mathbf{c}$ appearing as the $t \to -\infty$ limit (65).

Lemma 40 *(a) Let $\mathbf{Y} = \mathrm{COL}(\mathbf{X}; (c_1, 0, 0, \ldots))$. Let $Y_*(t)$ be the size of the color-1 cluster. Then as $t \to -\infty$*

$$\begin{aligned}
|t|\ Y_*(t) &\xrightarrow{p} c_1 &(99)\\
|t|^2 \left(\sum_i Y_i^2(t) - \sum_i X_i^2(t)\right) &\xrightarrow{p} c_1^2 &(100)\\
|t|^3 \left(\sum_i Y_i^3(t) - \sum_i X_i^3(t)\right) &\xrightarrow{p} c_1^3. &(101)
\end{aligned}$$

(b) Let $\mathbf{Y} = \mathrm{COL}(\mathbf{X};\mathbf{c})$*. Then*

$$|t|\left(Y_{[j]}(t) - X_j(t)\right) \stackrel{p}{\to} 0 \textit{ as } t \to -\infty$$

where $Y_{[j]}(t)$ *is the size of the cluster of* $\mathbf{Y}(t)$ *containing* $X_j(t)$*.*

Proof. (a) As $t \to -\infty$,

$$E(|t|Y_*(t)|\mathbf{X}(t)) = |t|\sum_i X_i(t)(1-\exp(-c_1X_i(t))) \stackrel{p}{\sim} |t|c_1S(t) \stackrel{p}{\to} c_1 \text{ by } (98)$$

$$\mathrm{var}\ (|t|Y_*(t)|\mathbf{X}(t)) \le |t|^2\sum_i X_i^2(t)(1-\exp(-c_1X_i(t))) = O(|t|^2S_3(t)) \stackrel{p}{\to} 0$$

by (63). So (99) follows from Chebyshev's inequality. The left side of (100) equals $|t|^2(Y_*^2(t) - A_2(t))$, where $A_2(t) = \sum X_i^2(t)$ summed over the colored clusters. So it suffices to prove $|t|^2A_2(t) \stackrel{p}{\to} 0$. But arguing as above,

$$\begin{aligned} |t|^2E(A_2(t)|\mathbf{X}(t)) &\le |t|^2\sum_i X_i^2(t)\cdot c_1X_i(t) \\ &= c_1|t|^2S_3(t) \\ &\stackrel{p}{\to} 0 \text{ by } (63). \end{aligned}$$

Similarly, the left side of (101) equals $|t|^3(Y_*^3(t) - A_3(t))$, with $A_3(t) = \sum X_i^3(t)$ summed over the colored clusters. To prove $|t|^3A(t) \stackrel{p}{\to} 0$, we bound as above,

$$\begin{aligned} |t|^3E(A_3(t)|\mathbf{X}(t)) &\le |t|^3\sum_i X_i^3(t)\cdot c_1X_i(t) \\ &= c_1|t|^3S_4(t) \\ &\le c_1X_1(t)\cdot|t|^3S_3(t) \\ &= O(X_1(t)) \stackrel{p}{\to} 0 \text{ by } (63). \end{aligned}$$

For (b), Lemma 35(ii) implies that (provided the denominator is positive)

$$|t|E(Y_{[j]}(t) - X_j(t)|\mathbf{X}(t)) \le |t|X_j(t)\frac{||c||^2}{1-||c||^2||\mathbf{X}(t)||^2}\ \sum_i X_i^2(t)$$

and the right side $\stackrel{p}{\to}$ 0 by (98). □

Here is the central result of section 5.

Proposition 41 *Fix* $\mathbf{c} \in l^2_\searrow$*. An eternal multiplicative coalescent* $\mathbf{X}$ *satisfies (65) with limit* $\mathbf{c}$ *if and only if* $\mathbf{X} \stackrel{d}{=} \mathrm{COL}(\widetilde{\mathbf{X}}; \mathbf{c})$*, where* $\widetilde{\mathbf{X}}$ *is some eternal multiplicative coalescent satisfying (65) with limit* $\widetilde{\mathbf{c}} = 0$*.*

Proof. Let $\mathbf{X}$ be an eternal multiplicative coalescent satisfying (65) with limit $\mathbf{c}$. For a fixed time t and a sequence of negative times $t_n < t$, let $\mathbf{X}^n(t)$ be the time-t distribution of the multiplicative coalescent started at time t_n with distribution $(X_{n+1}(t_n), X_{n+2}(t_n), \ldots)$, and let

$$C_i := X_i(t_n)(t - t_n)\,.$$

Assumption (65), together with the fact $\lim_{t\to-\infty} |t|S_2(t) = 1$, allows us to choose $t_n \to -\infty$ fast enough in order to have

$$(C_1, C_2, \ldots, C_n, 0, \ldots) \quad \to \quad \mathbf{c} \text{ a.s. in } l^2_\searrow \tag{102}$$

$$n \cdot \sum_{i=1}^{n} X_i(t_n) \quad \to \quad 0 \text{ a.s. } , \; n \to \infty. \tag{103}$$

Lemma 42 *For fixed* (n,t) *we may construct a random vector* $\bar{\mathbf{X}}^n(t)$ *jointly with* $\mathrm{COL}(\mathbf{X}^n(t); (C_1, \ldots, C_n, 0, 0, \ldots))$*, so that* $\bar{\mathbf{X}}^n(t) \stackrel{d}{=} \mathbf{X}(t)$ *and*

$$d\left(\bar{\mathbf{X}}^n(t), \mathrm{COL}(\mathbf{X}^n(t); (C_1, \ldots, C_n, 0, 0, \ldots))\right) \stackrel{p}{\to} 0 \textit{ as } n \to \infty.$$

Proof. It is again helpful to think in terms of the corresponding random graphs on the vertices $\{(i,0), i \geq 1\} \cup \{(k,1), k \geq 1\}$. First put

$$\text{(for } i, j \geq 1\text{) a Poisson}((t - t_n)X_{i+n}(t_n)X_{j+n}(t_n)) \text{ edges } (i,0) \leftrightarrow (j,0) \tag{104}$$

$$\text{(for } i \geq 1,\; k = 1, \ldots, n\text{) a Poisson}(X_i(t_n)C_k) \text{ edges } (i,0) \leftrightarrow (k,1). \tag{105}$$

As in the graphical construction, let $\mathbf{Y}^n = \mathrm{COL}(\mathbf{X}^n(t); (C_1, \ldots, C_n, 0, 0, \ldots))$ be the vector of ordered component-weights $w(\mathcal{C})$. Let $\mathbf{Z}^n$ be the vector of ordered full weights $w^f(\mathcal{C})$, where $(k,1)$ is given weight $X_k(t_n)$.

Figure 3 is an illustration: the components connected via solid edges (104, 105) generate $\mathbf{Z}^n$. Next include (dashed) extra edges

$$\text{(for } 1 \leq k < l \leq n\text{) a Poisson}(X_k(t_n)C_l) \text{ number of edges } (k,1) \leftrightarrow (l,1)\,, \tag{106}$$

and write $\bar{\mathbf{X}}^n(t)$ for the sequence of (new) full weights $w^f(\mathcal{C})$. Then $\bar{\mathbf{X}}^n(t) \stackrel{d}{=} \mathbf{X}(t)$ (in fact, this construction was previously used in Lemma 23).

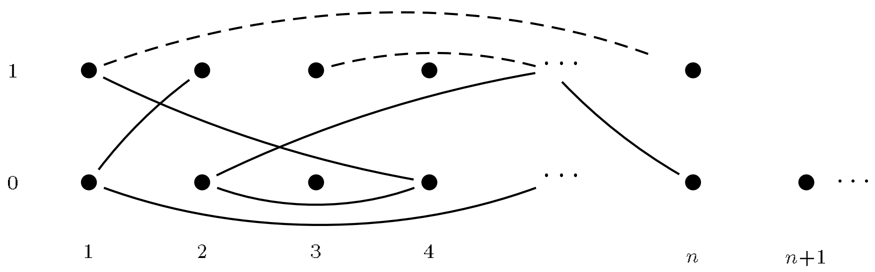

Figure 3

For $1 \leq k \leq n$ and $i \geq 1$ write $k \sim i$ if, after formation of $\mathbf{Y}^n$, vertex $(k,1)$ gets connected via (solid) edge to the i'th component of $\mathbf{Y}^n$. From the construction,

$$\begin{aligned} d^2(\mathbf{Z}^n, \mathbf{Y}^n) &\leq (\sum_{k=1}^n X_k(t_n))^2 + 2\sum_i (Y_i^n \sum_{k=1}^n X_k(t_n) 1_{\{k\sim i\}}) \\ &\leq (\sum_{k=1}^n X_k(t_n))^2 + 2(\sum_{i=1}^\infty (Y_i^n)^2\ )^{1/2} (\sum_{i=1}^\infty (\sum_{k=1}^n X_k(t_n) 1_{\{k\sim i\}})^2\ )^{1/2} \end{aligned}$$

by the Cauchy-Schwarz inequality, and since for each k, there is only one i such that $k \sim i$,

$$d^2(\mathbf{Z}^n, \mathbf{Y}^n) \leq (\sum_{k=1}^n X_k(t_n))^2 + 2\|\mathbf{Y}^n\|^{1/2} \sum_{k=1}^n X_k(t_n) \ \stackrel{p}{\to}\ 0\,,$$

by (102, 103) and tightness of $\|\mathbf{Y}^n\|$ since $\mathbf{Y}^n \preceq \bar{\mathbf{X}}^n(t) \ \stackrel{d}{=}\ \mathbf{X}(t) \in l^2_{\searrow}$.

Since $\bar{\mathbf{X}}(t)$ equals $\mathbf{Y}^n$, on the event $\{$ no dashed edges $\}$, it suffices to show they appear rarely. Precisely, conditional on $\mathbf{X}(t_n)$, the chance that their number is non-zero is at most

$$\begin{aligned} \sum\sum_{1\leq k<l\leq n} X_k(t_n) C_l \quad &= \quad \sum\sum_{1\leq k<l\leq n} (t-t_n) X_k(t_n) X_l(t_n) \\ &\leq \quad (t-t_n) X_1(t_n) \ \times\ n \sum_{k=1}^n X_k(t_n) \\ &\stackrel{p}{\to} \quad 0 \text{ by } (102, 103). \ \Box \end{aligned}$$

Returning to the proof of Proposition 41, note that in the construction above,

$$\mathbf{X}^n(t) \preceq \bar{\mathbf{X}}^n(t) \ \stackrel{d}{=}\ \mathbf{X}(t)$$

and so by Lemma 36(ii) the sequence $(\mathbf{X}^n(t), n \geq 1)$ is tight. By passing to a subsequence we may assume $\mathbf{X}^n(t)$ converges in distribution, and appealing to the Feller property we may assume $\mathbf{X}^n(t) \xrightarrow{d} \widetilde{\mathbf{X}}(t)$ for some eternal multiplicative coalescent $\widetilde{\mathbf{X}}(t)$. By (102) and Lemma 37

$$\mathrm{COL}(\mathbf{X}^n(t); (C_1, \ldots, C_n, 0, 0, \ldots)) \xrightarrow{d} \mathrm{COL}(\widetilde{\mathbf{X}}(t); \mathbf{c})$$

and then Lemma 42 implies $\mathbf{X}(t) \stackrel{d}{=} \mathrm{COL}(\widetilde{\mathbf{X}}(t); \mathbf{c})$. We need to prove $\tilde{\mathbf{c}} = 0$. If not, then $\lim_{t \to -\infty} |t| \widetilde{X}_1(t) = \tilde{c}_1 > 0$, and for simplicity we may suppose $\tilde{c}_1$ is constant and is distinct from the entries of $\mathbf{c}$ (the general case involves only modifying notation). But then Lemma 40(b) would imply that $\tilde{c}_1$ does indeed appear as a limit $\lim_{t \to -\infty} |t| X_{[1]}(t)$ for some cluster of $\mathbf{X}(t)$, a contradiction. This establishes the "only if" part of Proposition 41. Conversely, suppose $\mathbf{X}$ is distributed as $\mathrm{COL}(\widetilde{\mathbf{X}}; \mathbf{c})$, where $\widetilde{\mathbf{X}}$ is some eternal multiplicative coalescent satisfying (65) with limit $\tilde{\mathbf{c}} = 0$. Write $\mathbf{Y}^{(k)} = \mathrm{COL}(\widetilde{\mathbf{X}}; (c_1, \ldots, c_k, 0, 0, \ldots))$. Applying Lemma 40 inductively on k it is easy to see that the limit (65) for $\mathbf{Y}^{(k)}$ is $(c_1, \ldots, c_k, 0, 0, \ldots)$. Write $\mathbf{c}^{(k)} = (c_{k+1}, c_{k+2}, \ldots)$, so that by (94) we may use the representation

$$\mathbf{X} = \mathrm{COL}(\mathbf{Y}^{(k)}; \mathbf{c}^{(k)}).$$

Applying Lemma 40(b) to this representation, we see that the limits (65) for $\mathbf{X}$ include each entry of $(c_1, \ldots, c_k)$ and hence each entry of $\mathbf{c}$. Define $M_k(t)$ to be the maximum size of a cluster of $\mathbf{X}(t)$ which does not contain any of the clusters $Y_1^{(k)}(t), \ldots, Y_k^{(k)}(t)$. We will show

$$\limsup_{t \to -\infty} |t| M_k(t) \xrightarrow{p} 0 \text{ as } k \to \infty \tag{107}$$

and it follows that $\mathbf{X}$ can have no limit (65) except for the entries of $\mathbf{c}$, establishing the "if" part of Proposition 41.

Fix k, and to ease notation write $\mathbf{Y}$ for $\mathbf{Y}^{(k)}$. Write $\bar{S}_2(t) = \sum_i Y_i^2(t)$ and define events

$$\begin{aligned} B_1(t) &= \{\bar{S}_2(t) ||\mathbf{c}^{(k)}||^2 < 1/2\} \\ B_2(t) &= \{|t| Y_k(t) \leq 2c_k\} \\ B_3(t) &= \{|t| \bar{S}_2(t) \leq 2\} \\ B(t) &= \bigcap_{i=1}^{3} B_i(t) \end{aligned}$$

so that $P(B(t)) \to 1$ as $t \to -\infty$. For $i > k$ write $W_i(t)$ for the size of the cluster of $\mathbf{X}(t)$ containing $Y_i(t)$. The following inequalities hold on $B(t)$, for $i > k$.

$$\begin{aligned} E(W_i(t) - Y_i(t)|\mathbf{Y}(t)) &\leq 2\sum_{j>k} Y_j(t) \cdot Y_i(t)Y_j(t)||\mathbf{c}^{(k)}||^2 \\ &\leq 2Y_i(t)\bar{S}_2(t)||\mathbf{c}^{(k)}||^2 , \end{aligned}$$

the first inequality by Lemma 35 and definition of $B_1(t)$. Next, $|t|Y_i(t) \leq 2c_k$ by definition of $B_2(t)$, and so by Markov's inequality

$$\begin{aligned} P(|t|W_i(t) \geq 2c_k + \varepsilon|\mathbf{Y}(t)) &\leq \varepsilon^{-1}|t| \cdot 2Y_i(t)\bar{S}_2(t)||\mathbf{c}^{(k)}||^2 \\ &\leq 4\varepsilon^{-1}Y_i(t)||\mathbf{c}^{(k)}||^2 \text{ by definition of } B_3(t). \end{aligned}$$

If $|t|M_k(t) \geq 2c_k + \varepsilon$ then $\sum_{i>k} |t|Y_i(t)1_{(|t|W_i(t)\geq 2c_k+\varepsilon)} \geq 2c_k + \varepsilon \geq \varepsilon$ and so by Markov's inequality

$$\begin{aligned} P(|t|M_k(t) \geq 2c_k + \varepsilon|\mathbf{Y}(t)) &\leq \varepsilon^{-1}\sum_{i>k} |t|Y_i(t)\ P(|t|W_i(t) \geq 2c_k + \varepsilon|\mathbf{Y}(t)) \\ &\leq 4\varepsilon^{-2}|t|\bar{S}_2(t)||\mathbf{c}^{(k)}||^2 \\ &\leq 8\varepsilon^{-2}||\mathbf{c}^{(k)}||^2. \end{aligned}$$

Letting $t \to -\infty$ and then $k \to \infty$ establishes (107). □

5.4 Proof of Theorem 6

Proposition 7 shows that the standard multiplicative coalescent $\mathbf{X}^* \stackrel{d}{=} \mu(1, 0, \mathbf{0})$ is the limit of $(\mathbf{X}^{(n)}(t); t \geq -\frac{1}{\sigma_2(\mathbf{x}^{(n)})} = -n^{1/3})$ started with configuration $\mathbf{x}^{(n)}$ consisting of n clusters of mass $n^{-2/3}$ each. Write $\tilde{\mathbf{x}}^{(n)} = \mathbf{x}^{(n)} \bowtie (cn^{-1/3}, 0, 0, \ldots)$. Since $\frac{1}{\sigma_2(\mathbf{x}^{(n)})} - \frac{1}{\sigma_2(\tilde{\mathbf{x}}^{(n)})} \sim c^2$, Proposition 7 also shows that $\mu(1, -c^2, (c, 0, \ldots))$ is the limit of $(\widetilde{\mathbf{X}}^{(n)}(t); t \geq -n^{1/3})$ started with configuration $\tilde{\mathbf{x}}^{(n)}$.

For $n \geq 1$ and $t > -n^{-1/3}$ let $c_n(t) = \frac{c}{n^{1/3}}(t + n^{-1/3})$. Arguing as in Lemma 42, it is easy to obtain a version of $\widetilde{\mathbf{X}}^{(n)}(t)$, denoted by $\mathbf{Y}^{(n)}(t)$, made from $\mathrm{COL}(\mathbf{X}^{(n)}(t); (c_n(t), 0, \ldots))$ by appending (non-randomly) a cluster of mass $c/n^{1/3}$ to the color-c component. Since $c_n(t) \to c$, $\mathbf{X}^{(n)}(t) \stackrel{d}{\to} \mathbf{X}^*(t)$ and $\mathbf{Y}^{(n)}(t) \stackrel{d}{\sim} \mathrm{COL}(\mathbf{X}^{(n)}(t); (c_n(t), 0, \ldots))$ as $n \to \infty$, Lemma 37 gives

$$\mu(1, -c^2, (c, 0, \ldots)) \stackrel{d}{=} \mathrm{COL}(\mathbf{X}^*; (c, 0, \ldots)) .$$

Similarly, $\mu(1, -\sum_i c_i^2, \mathbf{c}) \stackrel{d}{=} \mathrm{COL}(\mathbf{X}^*; \mathbf{c})$ for any finite-length $\mathbf{c}$. When $\mathbf{c} \in l^2_\searrow$, let $\mathbf{c}^k = (c_1, \ldots, c_k, 0, 0, \ldots)$. Then for each k

$$\mu(1, -\sum_{i=1}^{k} c_i^2, \mathbf{c}^k) \stackrel{d}{=} \mathrm{COL}(\mathbf{X}^*; \mathbf{c}^k)\,. \tag{108}$$

Another application of Lemma 37 ensures convergence of the right-hand-side to $\mathrm{COL}(\mathbf{X}^*; \mathbf{c})$ as $k \to \infty$. The time-t marginals of $\mu(1, -\sum_{i=1}^{k} c_i^2, \mathbf{c}^k)$ and $\mu(1, -\sum_{i=1}^{\infty} c_i^2, \mathbf{c})$ are the distributions of the ordered excursion lengths of $B^{1,t-\sum_{i=1}^{k} c_i^2, \mathbf{c}^k}$ and $B^{1,t-\sum_{i=1}^{\infty} c_i^2, \mathbf{c}}$, and clearly

$$B^{1,t-\sum_{i=1}^{k} c_i^2, \mathbf{c}^k} \stackrel{d}{\to} B^{1,t-\sum_{i=1}^{\infty} c_i^2, \mathbf{c}}\,.$$

Using Propositions 14 and 17 as in section 2.6 we deduce that the left-hand-side in (108) converges to $\mu(1, -\sum_i c_i^2, \mathbf{c})$. This gives assertion (a) of Theorem 6. Similarly, for (b) we have $B^{1,0,\mathbf{c}^k} \stackrel{d}{\to} B^{1,0,\mathbf{c}}$, implying $\mu(1, 0, \mathbf{c}^k) \stackrel{d}{\to} \mu(1, 0, \mathbf{c})$. And for (c) we have $B^{\kappa,\tau,\mathbf{c}} \stackrel{d}{\to} B^{0,\tau,\mathbf{c}}$, implying $\mu(\kappa, \tau, \mathbf{c}) \stackrel{d}{\to} \mu(0, \tau, \mathbf{c})$.

5.5 Proof of Lemma 1

Suppose l_0 is a strict subset of $l^3_\searrow \setminus l^2_\searrow$ and take $\mathbf{c} \in (l^3_\searrow \setminus l^2_\searrow) \setminus l_0$. So for some fixed t and $\delta > 0$,

$$P(B^{0,t,\mathbf{c}} \text{ has infinitely many excursions of length } > \delta) > \delta. \tag{109}$$

Let $(\mathbf{x}^{(n)}, n \geq 1)$ satisfy the hypotheses of Proposition 7 with limits $\kappa = 0$, $\tau = 0$ and the given $\mathbf{c}$. Let $\mathbf{X}^{(n)}(\cdot)$ be the multiplicative coalescent started at time 0 from configuration $\mathbf{x}^{(n)}$. Let $(Z^{(n)}(s), s \geq 0)$ be the reflected rescaled breadth-first walk associated with $\mathbf{X}^{(n)}(\frac{1}{\sigma_2(\mathbf{x}^{(n)})} + t)$, so that as argued below (61)

$$Z^{(n)} \stackrel{d}{\to} B^{0,t,\mathbf{c}} \text{ on } D[0, \infty). \tag{110}$$

Define $\tilde{\mathbf{x}}^{(n)} = \mathbf{x}^{(n)} \bowtie (\sigma_2(\mathbf{x}^{(n)}), 0, 0, \ldots)$. The sequence of configurations $(\tilde{\mathbf{x}}^{(n)}, n \geq 1)$ satisfies the hypotheses in Proposition 7 with limits $\tilde{\kappa} = 0$, $\tilde{\tau} = -1$, $\tilde{\mathbf{c}} = \mathbf{c} \bowtie (1, 0, 0, \ldots)$. Let $\widetilde{\mathbf{X}}^{(n)}(\cdot)$ be the multiplicative coalescent started at time 0 from configuration $\tilde{\mathbf{x}}^{(n)}$. Let $(\widetilde{Z}^{(n)}(s), s \geq 0)$ be the reflected rescaled breadth-first walk associated with $\widetilde{\mathbf{X}}^{(n)}(\frac{1}{\sigma_2(\mathbf{x}^{(n)})} + t)$, so that as above

$$\widetilde{Z}^{(n)} \stackrel{d}{\to} B^{0,t+1,\tilde{\mathbf{c}}} \text{ on } D[0, \infty).$$

We shall describe a coupling of $\mathbf{X}^{(n)}(\frac{1}{\sigma_2(\mathbf{x}^{(n)})}+t)$ and $\widetilde{\mathbf{X}}^{(n)}(\frac{1}{\sigma_2(\mathbf{x}^{(n)})}+t)$, and an induced coupling of $Z^{(n)}$ and $\widetilde{Z}^{(n)}$, with the property that, with probability bounded away from zero as $n \to \infty$,

$$\text{the first excursion of } \widetilde{Z}^{(n)} \text{ of length} > \delta \text{ has length} > \theta(n) \text{ and starts} \\ \text{no later than the start of the first excursion of } Z^{(n)} \text{ of length} > \delta \quad (111)$$

where $\theta(n) \to \infty$. This implies that with non-zero probability the limit $B^{0,t+1,\tilde{\mathbf{c}}}$ has an infinite excursion, contradicting Proposition 14(a), and the contradiction establishes the lemma.

The coupling uses the coloring construction. Write $c_n = 1 + t\sigma_2(\mathbf{x}^{(n)})$. First construct $\mathrm{COL}(\mathbf{X}^{(n)}(\frac{1}{\sigma_2(\mathbf{x}^{(n)})}+t); c_n)$ and then to the colored component append a cluster of mass $\sigma_2(\mathbf{x}^{(n)})$, and write $\widetilde{\mathbf{X}}^{(n)}(\frac{1}{\sigma_2(\mathbf{x}^{(n)})}+t)$ for the resulting random vector. This has the correct distribution, by the property that coloring commutes with the evolution of the multiplicative coalescent. Moreover we may regard $\mathbf{X}^{(n)}(\frac{1}{\sigma_2(\mathbf{x}^{(n)})}+t)$ and $\widetilde{\mathbf{X}}^{(n)}(\frac{1}{\sigma_2(\mathbf{x}^{(n)})}+t)$ as component sizes of graphs on vertex sets of masses $\mathbf{x}^{(n)} \subseteq \tilde{\mathbf{x}}^{(n)}$, so that the size-biased ordering of $\tilde{\mathbf{x}}^{(n)}$ induces coupled versions of the walks $Z^{(n)}$ and $\widetilde{Z}^{(n)}$. Using (109, 110), with probability bounded away from zero as $n \to \infty$,

(a) at least $\theta(n)$ clusters corresponding to excursions of $Z^{(n)}$ of length $> \delta$ are colored (for some $\theta(n) \to \infty$)

(b) the cluster corresponding to the first excursion of $Z^{(n)}$ of length $> \delta$ is colored.

But if the events (a) and (b) hold then (111) automatically holds.

6 Final remarks

Give the set of distributions of eternal multiplicative coalescents the topology of weak convergence of processes on $D((-\infty,\infty), l^2_{\searrow})$. It is natural to believe (but not obvious) that the subset of extreme distributions is closed. The Theorem 3 characterization induces a topology on the parameter space $\mathcal{I} \cup [0,\infty)$. This is not quite the "natural" topology obtained by considering $\mathcal{I}$ as a subset of $R^2 \times l^3_{\searrow}$, for the following reason. Given arbitrary $\mathbf{c}_m \in l^3_{\searrow}$ and $\kappa_m > 0$, we have

$$\mu(\kappa_m, \tau_m, \mathbf{c}_m) \to \hat{\mu}(0) \text{ if } \tau_m \to -\infty \text{ sufficiently fast}$$

$$\mu(\kappa_m, \tau_m, \mathbf{c}_m) \not\to \hat{\mu}(0) \text{ if } \tau_m \to \infty \text{ sufficiently fast}$$

where $\hat{\mu}(0)$ is the distribution of the zero process. To specify the topology on extreme distributions requires (at least) specifying which sequences τ_m satisfy $\mu(\kappa_m, \tau_m, \mathbf{c}_m) \to \hat{\mu}(0)$, and this is not obvious.

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. To appear in *Bernoulli*. Available via homepage http://www.stat.berkeley.edu/users/aldous, 1997.

[3] D.J. Aldous. Emergence of the giant component in special Marcus-Lushnikov processes. To appear in Random Structures and Algorithms, 1998.

[4] D.J. Aldous and J. Pitman. The standard additive coalescent. Technical Report 489, Dept. Statistics, U.C. Berkeley, 1997. Available via http://www.stat.berkeley.edu/users/pitman.

[5] J. Bertoin. *Lévy Processes*. Cambridge University Press, 1996. Cambridge Tracts in Math. 126.

[6] P. Billingsley. *Convergence of Probability Measures*. Wiley, 1968.

[7] E.B. Dynkin. Sufficient statistics and extreme points. *Ann. Probab.*, 6:705–730, 1978.

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

[9] S.N. Evans and J. Pitman. Construction of Markovian coalescents. Technical Report 465, Dept. Statistics, U.C. Berkeley, 1996. Revised May 1997. To appear in *Ann. Inst. Henri Poincaré*.

[10] O. Kallenberg. *Random Measures*. Akademie-Verlag, Berlin, 1983.

[11] J.F.C. Kingman. The coalescent. *Stochastic Process. Appl.*, 13:235–248, 1982.

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