

# Kingman's coalescent and Brownian motion.

**Julien Berestycki and Nathanaël Berestycki**

Julien Berestycki, Université UPMC, Laboratoire de Probabilités et Modèles Aléatoires.  
<http://www.proba.jussieu.fr/~berest/>

*E-mail address:* [julien.berestycki@upmc.fr](mailto:julien.berestycki@upmc.fr)

Nathanaël Berestycki, University of Cambridge, DPMMS, Statistical Laboratory,  
<http://www.statslab.cam.ac.uk/~beresty/>

*E-mail address:* [N.Berestycki@statslab.cam.ac.uk](mailto:N.Berestycki@statslab.cam.ac.uk)

**Abstract.** We describe a simple construction of Kingman's coalescent in terms of a Brownian excursion. This construction is closely related to, and sheds some new light on, earlier work by Aldous (1998) and Warren (1999). Our approach also yields some new results: for instance, we obtain the full multifractal spectrum of Kingman's coalescent. This complements earlier work on Beta-coalescents by Berestycki et al. (2007). Surprisingly, the thick part of the spectrum is *not* obtained by taking the limit as  $\alpha \rightarrow 2$  in the result for Beta-coalescents mentioned above. Other analogies and differences between the case of Beta-coalescents and Kingman's coalescent are discussed.

## 1. Introduction and results

Let  $(\Pi_t, t \geq 0)$  be Kingman's coalescent. This is a Markov process taking its values in the set  $\mathcal{P}$  of partitions of  $\mathbb{N} = \{1, \dots\}$ , such that initially  $\Pi(0)$  is the trivial partition composed exclusively of singletons, and such that each pair of blocks merges at rate 1. Kingman's coalescent was introduced in 1982 by Kingman in his seminal paper Kingman (1982). A fascinating mathematical object in its own right, Kingman's coalescent is also a cornerstone of mathematical population genetics, serving as the basic model for genealogies in a well-mixed, selectively neutral population.

In this paper we provide a new, simple construction of Kingman's coalescent in terms of a Brownian excursion. We apply this construction to the study of some fine properties of this process, paying particular attention to its behaviour near the beginning of time, when microscopic clusters coalesce into a finite number of macroscopic ones. This phenomenon is known as *coming down from infinity*, and we are able to describe the precise multifractal spectrum corresponding to clusters of atypical sizes at small times. Our construction is closely related to (and, in some sense, almost implicit in) earlier work by Aldous (1998), Warren (1999) and others (see section 2 below). However, there are important differences, which will also be discussed in Section 2 (for instance, the aforementioned application could not have

---

*Received by the editors , accepted .*

2000 *Mathematics Subject Classification.* 60J65, 60J80.

*Key words and phrases.* Kingman coalescent, Brownian excursion, Branching processes.

been deduced from these works). This paper complements the picture developed in Berestycki et al. (2007) and Berestycki et al. (2008a) on the relation between coalescents, continuum random trees, and Fleming-Viot type particle systems.

We now describe our construction of Kingman's coalescent. Let  $(B_t, t \geq 0)$  be an excursion of Brownian motion conditioned to hit level 1. That is, let  $\nu$  denote Itô's excursion measure on the space of continuous excursions  $\Omega^* := \bigcup_{\zeta > 0} \Omega_\zeta$ , where

$$\Omega_\zeta := \{f : [0, \zeta] \rightarrow \mathbb{R} \text{ continuous, } f(x) = 0 \iff x \in \{0, \zeta\}\}.$$

Let then  $(B_t, 0 \leq t \leq \zeta)$  be a realization of  $\nu(\cdot | \sup_{s > 0} B_s \geq 1)$ . (With a slight abuse of notation, we may consider  $B$  to be a function defined on  $[0, \infty)$  by declaring  $B(s) = 0$  for all  $s \geq \zeta$ .) Let  $\{L(t, x)\}_{t \geq 0, x \geq 0}$  denote a jointly continuous version of the local-time process of  $B$ , and define

$$Z_x := L(\zeta, x), \quad x \geq 0.$$

Thus  $Z_x$  is the total local time accumulated at level  $x$  by the excursion  $(B_s, s \geq 0)$ . Define a  $\mathcal{P}$ -valued process  $(\Pi_u, 0 \leq u \leq 1)$  as follows. Consider the set  $\{\varepsilon_i\}_{i=1}^\infty$  of excursions of  $B$  above level 1, ordered according to their height: that is,

$$\sup_{s > 0} \varepsilon_1(s) > \sup_{s > 0} \varepsilon_2(s) > \dots$$

Now, fix  $0 < u < 1$ , and consider the set of excursions  $\{e_k\}_{1 \leq k \leq N}$  above level  $u$  reaching level 1, where  $N = N(u)$  is the number of such excursions. Note that for every  $i \geq 1$ ,  $\varepsilon_i$  belongs to exactly one excursion  $e_k$  for some  $1 \leq k \leq N$ , and let  $\phi_u(i) = k \in \{1, \dots, N\}$  be this index. Then define  $\Pi_u$  by declaring that for every  $i, j \geq 1$ ,  $i$  and  $j$  are in the same block of  $\Pi_u$  if and only if  $\phi_u(i) = \phi_u(j)$ , that is, if and only if  $\varepsilon_i$  and  $\varepsilon_j$  fall into the same excursion  $e_k$  for some  $k \leq N$ . Our main result states that  $\Pi_u$  is in fact a time-change of Kingman's coalescent.

**Theorem 1.1.** *The process  $(\Pi_{U(t)}, t \geq 0)$  has the same law as Kingman's coalescent, where for all  $t > 0$ ,*

$$U(t) = \sup \left\{ s > 0 : \int_s^1 \frac{4}{Z_u} du > t \right\}. \quad (1.1)$$

**Remark 1.2.**

- (1) As the reader has surely guessed, the ordering of the excursions  $(\varepsilon_i)_{i=1}^\infty$  by their height is not crucial to this result, but is in the spirit of our use of the Donnelly-Kurtz lookdown approach (see section 3 and Berestycki et al. (2007)).
- (2) The time-change  $\{U(t), t \geq 0\}$  satisfies the following properties:  $U(0) = 1$ ,  $\lim_{t \rightarrow \infty} U(t) = 0$ , and  $U$  is continuous and monotone decreasing.

As promised at the beginning of this introduction, we now give some applications of Theorem 1.1 to the study of the small-time behaviour of Kingman's coalescent. Let  $(\kappa_t, t \geq 0)$  be Kingman's coalescent, and let

$$F(t) = \lim_{n \rightarrow \infty} \frac{1}{n} \#\{1 \leq i \leq n : i \text{ is in the same block as 1 in } \kappa_t\}. \quad (1.2)$$

$F(t)$  is the frequency of the block containing 1 at time  $t$ , and the existence of the almost sure limit in (1.2) (for all times simultaneously) follows from general theory on coalescent processes and exchangeable partitions: see, e.g., Proposition 2.8 in Bertoin (2006). In the same manner, one may define  $F_i(t)$  for all  $i \geq 1$  to be the

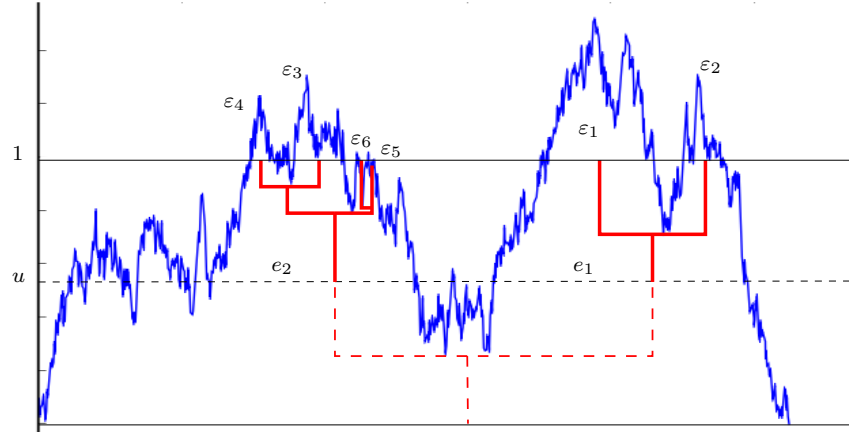


Figure 1.arabic@figure: Construction in Theorem 1. In this picture, we have:  $\Pi_u = (\{1, 2\}, \{3, 4, 5, 6\})$ .

asymptotic frequency of the block containing  $i$  at time  $t$ , thus  $F(t) = F_1(t)$  in (1.2). The first corollary gives us the behaviour of the typical block size near time zero. While this result is well known, our proof is new. Along the way we also provide an alternative path to a result of Aldous concerning the asymptotic of the number of blocks with a given size.

*Corollary 1.3.* As  $t \rightarrow 0$ ,

$$\frac{2F(t)}{t} \xrightarrow{d} E + E'$$

where  $\xrightarrow{d}$  stands for convergence in distribution, and where  $E$  and  $E'$  are two independent exponential variables with parameter 1.

Our second application is, to the best of our knowledge, new. It concerns the existence and the number of blocks (in the sense of Hausdorff dimensions to be specified below) with atypical sizes as  $t \rightarrow 0$ , that is, blocks of size of order  $t^\gamma$  with  $\gamma \neq 1$ . It turns out that for  $\gamma < 1$  (i.e., for anomalously large blocks) we have to look at a more precise, logarithmic, scale as there are almost surely no blocks whose size will be  $t^\gamma$  for small  $t$ . In particular, the sizes of the smallest and largest blocks at small time  $t > 0$  are identified.

For this result, the framework introduced by Evans (2000) is very convenient. Consider a random metric space  $(S, d)$ , defined as follows. Define a distance  $d$  on  $\mathbb{N}$  by declaring that for every  $i, j \geq 1$ ,  $d(i, j)$  is the time at which the integers  $i$  and  $j$  coalesce, and let  $S$  be the completion of  $\mathbb{N}$  with respect to  $d$ . It can be shown that to every ball of radius  $t > 0$  say, corresponds a unique block of the partition  $\kappa_t$ . The space  $S$  is thus naturally endowed with a measure  $\eta$  such that for every

$x \in S$  and for every  $t > 0$ ,  $\eta(x, t) := \eta(B(x, t))$  is the asymptotic frequency of the block of  $\kappa_t$  associated with  $B(x, t)$ . In this setting, the question mentioned above becomes: are there points  $x \in S$  such that  $\eta(B(x, t))$  is approximately  $t^\gamma$  as  $t \rightarrow 0$ , and if so, what is their Hausdorff dimension? Define for  $\gamma > 1$

$$S_{\text{thin}}(\gamma) = \left\{ x \in S : \limsup_{t \rightarrow 0} \frac{\log(\eta(x, t))}{\log t} = \gamma \right\}. \quad (1.3)$$

This set corresponds to points of  $S$  with atypically small  $\eta(x, t)$ . For thick points, consider for all  $\theta \geq 0$ ,

$$S_{\text{thick}}(\theta) = \left\{ x \in S : \limsup_{t \rightarrow 0} \frac{\eta(x, t)}{t |\log t|} = \theta \right\}. \quad (1.4)$$

**Theorem 1.4.**

1. If  $0 \leq \theta \leq 1$  then

$$\dim S_{\text{thick}}(\theta) = 1 - \theta, \text{ a.s.}$$

When  $\theta = 1$ ,  $S_{\text{thick}}(\theta) \neq \emptyset$  almost surely, but if  $\theta > 1$  then  $S_{\text{thick}}(\theta)$  is empty almost surely.

2. If  $1 < \gamma \leq 2$  then

$$\dim S_{\text{thin}}(\gamma) = \frac{2}{\gamma} - 1, \text{ a.s.}$$

If  $\gamma > 2$  then  $S_{\text{thin}}(\gamma) = \emptyset$  a.s. but  $S_{\text{thin}}(2) \neq \emptyset$  almost surely.

It may be deduced from the above result that the Hausdorff dimension of  $S$  is equal to 1 almost surely, a result which was first proved by Evans (2000).

This result should be compared to Theorem 5 in Berestycki et al. (2007) which describes the multifractal spectrum for Beta( $2 - \alpha, \alpha$ ) coalescent,  $\alpha \in (1, 2)$ . Kingman's coalescent is a limit case for this family and corresponds formally to the case  $\alpha = 2$ , since the Beta( $2 - \alpha, \alpha$ ) distribution converges weakly to a Dirac mass at 0 when  $\alpha \rightarrow 2$ . Interestingly, only the ‘‘thin points’’ side of the spectrum is obtained by taking a limit as  $\alpha \rightarrow 2$  in that result: at the power-law scale, the ‘‘thick points’’ part of the spectrum is empty, although the limit as  $\alpha \rightarrow 2$  exists and is non-degenerate. For instance, here the smallest block turns out to be of order  $t^2$  as  $t \rightarrow 0$ , while the largest block is of order  $t \log(1/t)$ . In the case of Beta-coalescents with parameter  $1 < \alpha < 2$ , these quantities are respectively  $t^{\alpha/(\alpha-1)^2}$  and  $t^{1/\alpha}$ , which does not coincide with  $t \log(1/t)$  when  $\alpha \rightarrow 2^-$ .

We emphasize that the proof of Theorem 1.4 is in some sense purely conceptual: all the work consists of identifying Evans' metric space with a certain Galton-Watson tree (which here turns out to be the Yule tree), equipped with its branching measure. Theorem 1.4 then follows automatically from the results of Mörters and Shieh (2002, 2004) on the multifractal spectrum of these measures. These results themselves rely on elegant percolative methods introduced by Lyons (1990).

## 2. Previous results and discussion

We review some earlier results concerning the relation between Kingman's coalescent and Brownian processes. We start by discussing the ideas contained in Aldous (1999) and Warren (1999), which are most directly related to the representation of Theorem 1.1.

2.1. *Excursions conditioned on their local time.* The notion of continuum random tree (CRT) was developed by D. Aldous in his seminal papers Aldous (1991, 1993), in which a careful treatment of the correspondence between excursions and continuum random trees is given. Particular attention is paid to the normalized Brownian excursion and to the tree that it encodes through this correspondence. Early effort bore on the identification of the law of the tree spanned by a finite number of leaves chosen suitably at random. Given an excursion function  $e \in \Omega^*$  and  $p \in \mathbb{N}$ , let  $t_1, t_2, \dots, t_p \in (0, \zeta)$  be pairwise distinct times, and define a planar labelled tree with edge-lengths  $T = T_p(e, t_1, \dots, t_p)$  as follows.

- $T$  contains a root at height 0 and  $p$  leaves, with the height of the  $k^{\text{th}}$  leaf being  $e(t_k)$ .
- The path from the root to the leaves  $j$  and  $k$  ( $1 \leq j \neq k \leq p$ ) splits at height  $\inf\{e(s), s \in (t_j, t_k)\}$ .
- At each branch point, branches are labelled "left" and "right".

It is further assumed that all branch points are distinct (i.e., branch points are binary). Thus  $T$  has  $p - 1$  branch points. Le Gall (1993) (see his Theorem 3) first identified the distribution of  $T_p(e, t_1, \dots, t_p)$  on the space of trees when  $e$  is a normalized Brownian excursion and the  $t_k$  are independent and uniform random variables on  $[0, \zeta(e)]$  (and further gives the conditional distribution of  $e$  conditionally given the tree  $T$ ). Later on, Aldous (1998) gave a decomposition result for this distribution by identifying the conditional distribution of  $T$  given the local time profile  $(\ell_x, x \geq 0)$  where  $\ell_x$  is the total local time accumulated at  $x$  by  $e$ . This conditional distribution is constructed from a certain inhomogeneous (i.e., time-dependent) coalescent process whose main ingredient is Kingman's coalescent. More precisely, this process is defined as follows. First, the heights of the leaves  $(x_1, \dots, x_p)$  are i.i.d. samples from  $\ell_x dx$  (which is a probability distribution because  $\zeta(e) = 1$  almost surely). Then, thinking of time as running from  $x = \infty$  down to  $x = 0$  and the heights  $x_i$  as being the birth times of particles, the law of  $T$  is such that as  $x$  decreases, the particles that are present merge pairwise, each independently at rate  $4/\ell_x$ . In other words, the conditional distribution of  $T$  given  $\ell$  can be thought of as *an inhomogeneous coalescent with immigration rate  $\ell_x$  and coalescing rate  $4/\ell_x$* . This result may be regarded as the continuum random tree counterpart to Perkins' disintegration theorem in the context of superprocesses Perkins (1991). The proof of Aldous is based on a discrete approximation to the continuum random tree, but soon afterwards Warren (1999) gave two alternative, direct proofs of Aldous's result. The tools used in those arguments are closely related to earlier work by Le Gall (1993) and Warren and Yor (1998).

It thus comes as no surprise that one can embed Kingman's coalescent into a Brownian excursion. However, we emphasize that our construction is rather different in that the tree which we consider is spanned by vertices at distance 1 from the root, rather than by leaves selected uniformly at random. Moreover, it seems difficult to use the description of  $T$  given above to deduce results about Kingman's coalescent. On the other hand, as the reader will see, here these conclusions will follow in a straightforward fashion once Theorem 1.1 is proved. Finally, we believe that the computations leading up to Theorem 1.1 are new, and hope that the reader will find them interesting in themselves.

*2.2. Analogies and differences with Beta-coalescents.* The present paper complements earlier results of Berestycki et al. (2008b, 2007), which focuses on small-time properties of Beta-coalescents with parameter  $\alpha \in (1, 2)$ . Beta-coalescents are a family of  $\Lambda$ -coalescent (i.e., coalescents with multiple collisions), where the measure  $\Lambda$  is the density of a Beta( $2 - \alpha, \alpha$ ) random variable. (The interested reader may consult Berestycki et al. (2007) and references therein.) Kingman's coalescent constitutes a formal limit case for the Beta-coalescents as  $\alpha \rightarrow 2$  since the Beta( $2 - \alpha, \alpha$ ) converges weakly to the Dirac mass in zero as  $\alpha \rightarrow 2$ . It was proved in Berestycki et al. (2007) that for all  $1 < \alpha < 2$ , Beta-coalescents can be embedded in continuum stable random trees associated with  $\alpha$ -stable branching processes rather than Brownian motion. (An excellent introduction to continuous random trees can be found in Duquesne and Le Gall (2002)). When  $\alpha = 2$  the  $\alpha$ -stable branching process is the Feller diffusion, which is closely related to Brownian motion. It is therefore natural to suspect a relation between Kingman's coalescent and Brownian motion.

However, we emphasize that there is an essential difference between Theorem 1.1 here and Theorem 1 in Berestycki et al. (2007). The analogue of Theorem 1 in Berestycki et al. (2007) is the following. Let  $(B_s, s \geq 0)$  be a reflecting Brownian motion and let  $\tau_1 = \inf\{t > 0 : L(t, 0) > 1\}$ , where  $L(t, x)$  is the joint local time process of  $B$ . Let  $v > 0$  be such that  $v < \sup_{s \leq \tau_1} B(s)$ , and for all  $0 \leq u \leq v$ , define a partition  $\Pi_u^v$  in exactly the same way as in the construction given above Theorem 1.1, except that where we used the level 1 we now use level  $v$ : thus, here the excursions  $(\varepsilon_i)_{i=1}^\infty$  which we consider are those above level  $v$  (instead of 1) and the  $e_k$  are the excursions above  $u$  that reach  $v$ . Therefore, using this notation, the partition  $\Pi_u$  defined for Theorem 1.1 is simply  $\Pi_u^1$  (with, however, the difference that here  $B$  is not a single excursion).

Now, for all  $t > 0$ , let

$$V(t) := \inf \left\{ s > 0 : \int_0^s 4Z_u^{-1} du > t \right\}$$

where  $Z_x = L(\tau_1, x)$  for all  $x \geq 0$ . (Note that  $V(t) < \sup_{s \leq \tau_1} B(s)$ ). Fix  $T > 0$  and for  $0 \leq t \leq T$ , consider the partition

$$\hat{\Pi}_t := \Pi_{V(T-t)}^{V(T)}.$$

We can prove the following result:

*Proposition 2.1.*  $(\hat{\Pi}_s, 0 \leq s \leq T)$  has the same law as  $(\kappa_s, 0 \leq s \leq T)$ , Kingman's coalescent run for time  $T$ .

Observe that the main difference between Proposition 2.1 and Theorem 1 is that here the tree is spanned by vertices at a random distance  $V(T)$  from the root of the continuum random tree, while these vertices are at a deterministic distance 1 in Theorem 1.1. This was the source of considerable technical difficulties in Berestycki et al. (2007). As the reader will see, the proof of Theorem 1.4 is much more straightforward than that of Theorem 5 in Berestycki et al. (2007).

A direct proof of Proposition 2.1 is provided in Section 3. However, it could also be deduced from Perkins' disintegration theorem together with Le Gall's excursion representation of Dawson-Watanabe processes. For this, one must use the fact that the empirical measure for an exchangeable system of particles fully determines the law of the individual trajectories.

In light of Theorem 1.1 and the above discussion, one can wonder if, conversely, Theorem 1.1 could not be extended to Beta-coalescents. Interestingly, we have rather strong evidence that this cannot be true. Assuming that the construction was valid for a Beta-coalescent, we would conclude that the number of blocks lost at every collision would be an i.i.d. sequence. Indeed, from results in Duquesne and Le Gall (2002), one can see that the tree spanned by vertices at a deterministic level (the “reduced tree”) forms a time-change of a Galton-Watson process. However, in that case the distribution of the number of blocks involved at every collision must agree with the distribution in equation (10) of Berestycki et al. (2008b). Simulations indicate that this is not the case for the final collision. However, we are not aware of a rigorous proof that this cannot be true, although there are also some theoretical arguments in that direction. We note that the essential reason for which the proof of Theorem 1.1 breaks down in the case  $\alpha < 2$  is that Brownian motion possesses an extra symmetry property (reflection about a horizontal line) compared to the height process of other stable continuous-state branching processes.

*2.3. More on duality.* Theorem 1.1 brings to mind another result which relates Kingman's coalescent to the behaviour of Yule processes. (As the reader will see in section 5, Yule processes are indeed embedded in the construction of Theorem 1.1). Consider the jump chain of Kingman's coalescent i.e., the chain  $X(1), \dots, X(n), \dots$  where  $X(n)$  is the element of the  $n^{\text{th}}$  simplex given by the block frequencies of Kingman's coalescent when it has  $n$  blocks. In Bertoin and Goldschmidt (2004), Bertoin and Goldschmidt show that this chain  $X$  can be related to the fragmentation chain obtained by taking a Yule process  $(Y_t, t \geq 0)$  and conditioning on  $\{W = w > 0\}$ , where  $W = \lim_{t \rightarrow \infty} e^{-t} Y_t$  almost surely. This fragmentation process  $(G_t, t \geq 0)$  is defined by considering how the mass  $w$  is being split between the children of an individual when it dies. More precisely, the two children of this individual each give rise to their own independent Yule process  $Y^{(1)}$  and  $Y^{(2)}$ , say. To each one we can associate the corresponding random variable  $W^{(1)}$  and  $W^{(2)}$ . Note that if  $\tau$  is the time of the split, we must have  $W = e^{-\tau} W^{(1)} + e^{-\tau} W^{(2)}$ . Define the fragmentation process  $(G_t, t \geq 0)$  by saying that at time  $\tau$ , the mass  $W$  splits into  $e^{-\tau} W^{(1)}$  and  $e^{-\tau} W^{(2)}$ , and so on as time evolves. If we do not condition on  $\{W = w\}$ , then  $G$  is a fragmentation in the sense that fragments evolve independently with individuals fragments behaving as a rescaled copy of the original process. This is what is usually called a *homogeneous fragmentation process* (see, e.g., Bertoin (2006, Section 3.1)) but note that here the total mass is random. It is then natural to ask what happens when we condition on  $\{W = w\}$  to keep the total mass of this fragmentation deterministic. Theorem 3.1 of Bertoin and Goldschmidt (2004) then states that if  $(N_t, t \geq 0)$  is an independent Poisson process, then conditionally on  $\{W = w\}$ ,

$$\{G_{\log(1+t)}, t \geq 0\} \stackrel{d}{=} \{wX(N_{wt}), t \geq 0\}. \quad (2.1)$$

That is, up to a time-change, the fragmentation  $G$  is the time-reversal of the jump chain of Kingman's coalescent.

In turn, we remind the reader that there is a well-known relation between Yule processes and Brownian excursions: namely, Le Gall (1989) proved that the reduced tree of a Brownian excursion is a time changed Yule tree (this result, Theorem 8, is actually formulated in terms of branching processes, but a tree version can easily

be deduced). The occurrence of the Yule process in the work of Goldschmidt and Bertoin and in the proof of Theorem 1.1 is therefore not surprising.

It is an open problem to decide whether a similar representation exists for the case of Beta-coalescents. We indicate that Goldschmidt (2008) has recently computed explicitly the Martin boundary of the continuous Galton-Watson process  $(Y_t, t \geq 0)$  associated with the reduced tree at level 1 of a stable continuum random tree with index  $\alpha$ . This can be used to describe explicitly the behaviour of  $Y^{(w)}$ , where  $Y^{(w)}$  denotes the process  $Y$  conditioned on  $\{W = w\}$ , (here again,  $W = \lim_{t \rightarrow \infty} e^{-\lambda t} Y_t$  and  $\lambda = 1/(\alpha - 1)$ ).

In particular, from her result it follows that the transition rates of  $Y^{(w)}$  are not independent of the current state of  $Y^{(w)}$ . Since it turns out that  $W = Z_1$  almost surely, ( $Z_1$  being the quantity of local time of the stable CRT at level 1), it follows that conditionally on  $\{Z_1 = z\}$ , the number of children of individuals in the tree is not i.i.d. In particular, the above objection against the extension of Theorem 1.1 does not hold here.

### 3. Brownian construction of Kingman's coalescent.

#### 3.1. Proof of Theorem 1.1.

**Proof.** The proof we give is based on a few simple calculations with excursion theory. For basic definitions and facts, we refer the reader to Chapter XII of Revuz and Yor (1999).

Fix  $0 < t < 1$ , and recall our notation  $Z_x = L(\zeta, x)$  for all  $x \geq 0$ , where  $L(t, x)$  is the joint local time of a Brownian excursion  $(B_s, s \geq 0)$  conditioned to exceed level 1. Define a filtration

$$\{\mathcal{G}_u = \sigma(Z_s, u \leq s \leq 1)\}_{0 \leq u \leq 1}. \quad (3.1)$$

In parallel, define a family of  $\sigma$ -algebras  $\{\mathcal{H}_u\}_{0 \leq u \leq 1}$  by putting

$$\mathcal{H}_u = \sigma(B_{\alpha^u(s)}, s \geq 0) \quad (3.2)$$

where for all  $s > 0$ :

$$\alpha^u(s) := \inf \left\{ t > 0 : \int_0^t \mathbf{1}_{\{B_\tau > u\}} d\tau > s \right\}.$$

In words  $\mathcal{H}_u$  contains all the information about the trajectory above level  $u$ . It is a tedious but easy exercise to see that  $\{\mathcal{H}_{1-t}\}_{0 \leq t \leq 1}$  is a filtration and  $\mathcal{H}_u \supseteq \mathcal{G}_u$  for all  $0 \leq u \leq 1$ . Let  $\delta > 0$ , and let  $u = 1 - t$ . Define the event  $A_{k,j}(\delta)$  as follows. Call  $N$  the number of excursions above  $u$  which reach 1, and let  $(e_1, \dots, e_N)$  be these excursions with an order given by their height (as usual, we view each  $e_i$  as an element of  $\Omega^*$ ). Let  $1 \leq k, j \leq N$  and let

$$A_{k,j}(\delta) := \{\text{excursions } e_k \text{ and } e_j \text{ have coalesced at level } u - \delta\},$$

where we say that  $e_k$  and  $e_j$  have coalesced at level  $u - \delta$  if  $e_k$  and  $e_j$  are part of a single excursion above level  $u - \delta$  which reaches level 1.

To alleviate notations we also write  $A(\delta)$  when there is no risk of confusion. Theorem 1.1 follows from the following claim:

$$\mathbb{P}(A_{k,j}(\delta) | \mathcal{G}_u; N \geq \max(j, k)) = \frac{4}{Z_u} \delta + o(\delta), \quad a.s. \quad (3.3)$$



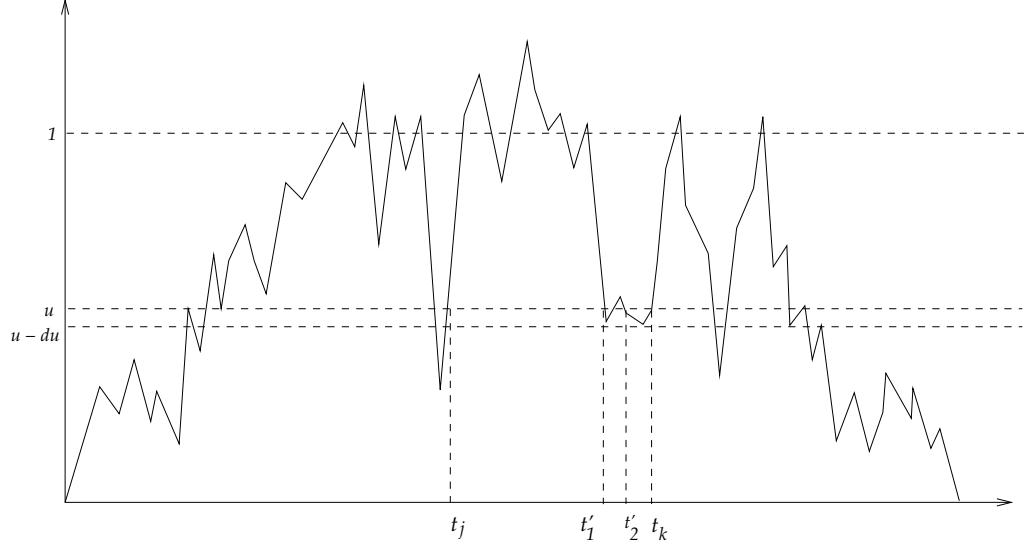


Figure 3.arabic@figure: Between levels  $u$  and  $u - du$ , excursions  $e_j$  and  $e_k$  (which start respectively at times  $t_j$  and  $t_k$ ) coalesce. For this to occur, the excursions below  $u$  starting at time  $t'_1$  and  $t'_2$  must not reach below  $u - du$ .

Note that to every excursion  $e_i$  with  $1 \leq i \leq n$ , we may associate a starting time  $t_i \in [0, \zeta]$  such that the excursion starts at time  $t_i$ , that is:  $B(t + t_i) = u + e_i(t)$ , for all  $t \leq \zeta(e_i)$ . Assume without loss of generality that  $t_j < t_k$ . Among all excursions below level  $u$ , consider the subset of those excursions ( $e'_i, i \geq 1$ ) which separate  $e_k$  and  $e_j$ , that is, whose starting time  $t'_i$  lies strictly in the interval  $(t_j, t_k)$  (see Figure 3.arabic@figure). Observe that for  $e_j$  and  $e_k$  to coalesce by level  $u - \delta$ , it is both necessary and sufficient that:

$$\inf_{s>0} e'_i(s) > -\delta, \quad \text{for all } i \geq 1. \quad (3.4)$$

Since  $L(t, u)$  may only increase at times  $t$  such that  $B_t = u$ , it follows that throughout  $e_j$ , the local time at level  $u$  is constant, let  $\ell_j$  be this quantity (thus,  $\ell_j = L(t_j, u)$ ). Similarly, let  $\ell_k$  be the local time at level  $u$  throughout excursion  $e_k$ , and let  $\ell = |\ell_j - \ell_k|$ . Then we claim that, conditionally on  $\ell_j$  and  $\ell_k$ ,  $A(\delta)$  has probability  $\exp(-\ell/(2\delta) + \ell/(2u))$ . Indeed, by excursion theory and elementary properties of Brownian motion, we have that, conditionally on  $\mathcal{H}_u$ ,

$$\sum_{i \geq 1} \delta_{(\ell'_i, e'_i)}$$

is a Poisson point process on  $[\ell_j, \ell_k] \times \Omega^*$  with intensity  $\mathbf{1}_{\{[\ell_j, \ell_k]\}} d\ell \otimes \nu^u(de)$ , where  $d\ell$  is the one-dimensional Lebesgue measure, and  $\nu^u$  is Itô's excursion measure restricted to  $\Omega_u = \{e \in \Omega^* : \inf_{s>0} e(s) > -u\}$ . Indeed, observe that the set of excursions below  $u$  forms a Poisson point process with intensity  $dt \otimes \nu_-$  (where  $\nu_-$  denotes the restriction of  $\nu$  to negative excursions), and that  $Z_u$  is equal to the first time there is a point in the set  $\Omega^* \setminus \Omega_u$ . Conditionally on  $\mathcal{H}_u$  and  $Z_u = z$ , this point process restricted to  $[0, z]$  is still Poisson but with intensity  $\mathbf{1}_{\{\ell \in [0, z]\}} d\ell \otimes \nu^u(de)$ .

It follows that

$$\sum_{i \geq 1} \delta_{(\ell_i, \inf_{s > 0} e'_i(s))}$$

is a Poisson point process on  $[\ell_i, \ell_j] \times \mathbb{R}_-$  with intensity  $d\ell \mathbf{1}_{\{\ell \in [\ell_i, \ell_j]\}} \otimes \mathbf{1}_{\{-u \leq h \leq 0\}} dh / 2h^2$  (see, e.g., (2.10) in Chapter XII of Revuz and Yor (1999)). Thus,

$$\mathbb{P}(A(\delta) | \mathcal{H}_u) = \exp\left(-|\ell_j - \ell_k| \left(\frac{1}{2\delta} - \frac{1}{2u}\right)\right).$$

Now, observe that, conditionally on  $\mathcal{G}_u$ ,

$$\sum_{i \geq 1} \delta_{(\ell_i, \sup_{s > 0} e_i(s))}$$

is a Poisson point process on  $[0, z] \times \mathbb{R}_+$  with intensity  $d\ell \mathbf{1}_{\{\ell \leq z\}} \otimes \mathbf{1}_{\{h \geq 1-u\}} dh / 2h^2$ , because  $e_i$  are precisely the set of excursions above  $u$  that reach level 1. Therefore, using elementary properties of Poisson processes, conditionally on  $\mathcal{G}_u$ , and conditionally on  $\{N \geq \max(j, k)\}$  (which we must assume if we are to talk about the coalescence of  $e_j$  and  $e_k$ ), then  $\ell_j$  and  $\ell_k$  are two independent random variables uniformly distributed on  $[0, z]$  (this does not depend on the precise intensity of the point process; all that is required is that the time-intensity is the Lebesgue measure and the space-intensity has no atom). In particular, conditionally on  $\mathcal{G}_u$  and  $N \geq \max(j, k)$ ,  $\ell := |\ell_j - \ell_k|$  may be written as:

$$\ell = z|U - V|$$

where  $U, V$  are uniform random variables on  $(0, 1)$  and  $z = Z_u$ . Putting these pieces together, we obtain:

$$\mathbb{P}(A(\delta) | \mathcal{G}_u; N \geq \max(j, k)) = \mathbb{E}\left[\exp\left(\frac{z}{2u}|U - V| - \frac{z}{2\delta}|U - V|\right)\right]$$

Note that the density of  $|U - V|$  is  $2 - 2x$  on  $[0, 1]$ , so making a change of variable  $x = \delta t$  leads to

$$\begin{aligned} \mathbb{P}(A(\delta) | \mathcal{G}_u; N \geq \max(j, k)) &= \int_0^1 (2 - 2x) \exp\left(\frac{zx}{2u} - \frac{zx}{2\delta}\right) dx \\ &= 2\delta \int_0^{1/\delta} \exp\left(\frac{z\delta t}{2u} - \frac{zt}{2}\right) dt - 2\delta^2 \int_0^{1/\delta} t \exp\left(\frac{z\delta t}{2u} - \frac{zt}{2}\right) dt \\ &= \frac{4\delta}{z} + o(\delta) \end{aligned}$$

as  $\delta \rightarrow 0$ . This is precisely (3.3), and so this implies Theorem 1.1.  $\square$

**3.2. Proof of Proposition 2.1.** The proof of Proposition 2.1 involves similar ideas to the proof of Theorem 1.1 above. However, the calculations become, somewhat surprisingly, slightly more complicated. In particular, there is a remarkable cancellation towards the end, which illustrates the following fact. Roughly speaking, we try to compute the rate at which the  $k^{\text{th}}$  highest excursion at a current level  $u > 0$  splits to give the  $k^{\text{th}}$  and  $j^{\text{th}}$  highest excursions, with  $j \geq k$ . We are trying to show that this rate is  $4/Z_u$  and is in particular independent of  $k$  and  $j$ . This may seem hard to believe at first: when splitting, it is easier to create an excursion of small size rather than large. However, the excursion heights accumulate near zero and therefore creating an interval whose size falls exactly between the sizes of the  $n^{\text{th}}$

smallest and  $n + 1$ th smallest intervals also becomes harder when  $n$  tends to infinity. These two effects exactly compensate each other and imply the aforementioned result!

Fix  $k, j \geq 1$  and assume that  $k \leq j$ . Recall that here  $(B_s, s \leq \tau_1)$  is no longer a single Brownian excursion but a collection of excursions which accumulate one unit of local time at level 0. We shall still denote by  $Z_t = L(\tau_1, t)$  the total local time accumulated at level  $t$  by  $(B_s, 0 \leq s \leq \tau_1)$ . Recall that, by the Ray-Knight theorem (Theorem (2.3) in Chapter XI of Revuz and Yor (1999)),  $(Z_t, t \geq 0)$  is the Feller diffusion, defined by:

$$\begin{cases} dZ_t = \sqrt{Z_t} dW_t \\ Z_0 = 1 \end{cases}$$

where  $(W_t, t \geq 0)$  is a standard Brownian motion. Let

$$\mathcal{K}_u = \sigma(B_{\gamma^u(s)}, s \geq 0) \quad (3.5)$$

where for all  $s > 0$ :

$$\gamma^u(s) := \inf \left\{ t > 0 : \int_0^t \mathbf{1}_{\{B_\tau \leq u\}} d\tau > s \right\}.$$

In words  $\mathcal{K}_u$  contains all the information about the trajectory of  $B$  below level  $u$ . Just as in the proof of Theorem 1.1,  $\{\mathcal{K}_t\}_{t \geq 0}$  is a filtration. We wish to compute the rate at which the  $j^{\text{th}}$  excursion “looks down” on the  $k^{\text{th}}$  excursion to adopt its label. We claim that, independently of  $k$  and  $j$ , and independently of  $\mathcal{H}_t$  (defined in (3.2)) the infinitesimal rate at which this happens is:

$$\text{rate}(j \text{ looks down on } k) = \frac{4}{Z_t}. \quad (3.6)$$

Here and in what follows, the event that  $j$  looks down on  $k$  means the following thing. Define a process  $\{\xi_i(t), t \geq 0\}_{i \geq 1}$  as follows. Initially,  $\{\xi_i(0)\}_{i \geq 1}$  are i.i.d uniform  $(0, 1)$  random variables. We think of  $\xi_i(0)$  as the label of the  $i^{\text{th}}$  highest excursion of  $(B_s, 0 \leq s \leq \tau_1)$  above 0. As time evolves,  $\xi_i(t)$  keeps this interpretation, with  $\xi_i(t)$  being the label of the  $i^{\text{th}}$  highest excursion above level  $t$  of  $(B_s, 0 \leq s \leq \tau_1)$ . The rule of evolution is that when the level  $t$  increases, an excursion may split into two parts. The highest of these two parts is necessarily still ranked  $i^{\text{th}}$ , while the second part has a larger rank  $j > i$ . In that case, we say that  $j$  looks down on  $i$ . When this occurs,  $\xi_j(t^+)$  becomes  $\xi_i(t^-)$ , while for  $k \geq j$ , what was previously the  $k^{\text{th}}$  excursion now becomes the  $(k + 1)^{\text{th}}$  excursion: thus in that case, we define  $\xi_{k+1}(t) = \xi_k(t^-)$ . Visually,  $j$  adopts the label of  $i$  (it looks down on it) and all labels corresponding to excursions with a higher rank are pushed up by 1. Thus (3.6) tells us that, in the terminology of Donnelly and Kurtz,  $\{\xi_i(t)\}_{i \geq 1}$  is a lookdown process with infinitesimal rate  $4/Z_t$  at time  $t$ , conditionally on  $(Z_t, t \geq 0)$ . We do not need the full definition of these processes, which is somewhat complicated, and can be found, for instance, in Chapter V of Etheridge (2000). Thus (3.6) is exactly the Brownian analogue of Theorem 14 in Berestycki et al. (2007), where further details on this approach can be found. In particular, Proposition 2.1 follows directly from (3.6) and Lemma 5.6 in Etheridge (2000).

We give a direct proof of (3.6) based on calculations with Itô's excursion measure. Fix a small  $\delta > 0$ . For all  $t \geq 0$ , let  $(e_{t,i})_{i \geq 1}$  be the excursions of  $B$  above level  $t$

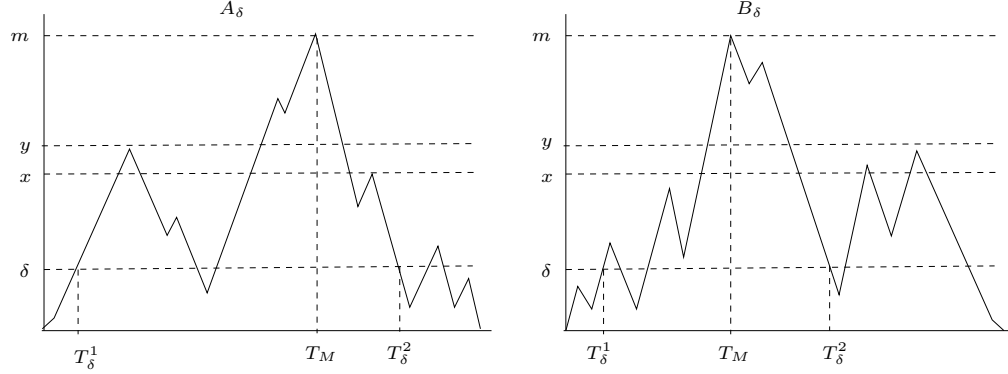


Figure 3.arabic@figure: Events  $A_{x,y}$  on the left and  $B_{x,y}$  on the right

ordered by their height (recall that  $e_{t,i} \in \Omega^*$  for all  $t \geq 0, i \geq 1$ ). For  $k \geq 1$ , let

$$M_k(t) = \sup_{s>0} e_{t,k}(s).$$

A simple analysis of the construction shows that in time  $\delta$  (that is, between times  $s$  and  $t := s + \delta$ ), the  $j^{\text{th}}$  excursion looks down on the  $k^{\text{th}}$  excursion if there is a local minimum located within the  $k^{\text{th}}$  highest excursion at level  $s \leq m \leq t = s + \delta$  such that of the two branches going out of this local minimum, one has a height equal to  $M_k(t^-) = M_k(t)$  and the other has a height  $h_{\text{new}}$  such that

$$M_j(t^-) < h_{\text{new}} < M_{j+1}(t^-). \quad (3.7)$$

Let  $(e_k(x), 0 \leq k \leq \zeta)$  be this excursion. Given  $\{M_k(t) = m\}$ ,  $e_k$  has the law of an Itô excursion conditioned on the event:  $\{\sup_{x \in (0, \zeta)} e_k(x) = m\}$ . For  $m > 0$ , let  $Q^{(m)}$  denote the probability measure on  $C([0, \infty))$  defined by the conditioning of  $\nu$  given that  $\{\sup_{s>0} e(s) = m\}$ . Define:

$$\begin{aligned} T_\delta^1 &= \inf\{x > 0 : e_k(x) > \delta\}, \\ T_M &= \text{the unique } x \geq 0 \text{ such that } e_k(x) = M_k(t) = m, \\ T_\delta^2 &= \inf\{x > T_M : e_k(x) < \delta\}. \end{aligned}$$

We note that (3.7) occurs if and only if the event  $A_{M_{j+1}, M_j} \cup B_{M_{j+1}, M_j}$  occurs, where

$$A_{x,y} := \left\{ \exists T_\delta^1 < s < T_M : x < \sup_{r \in [T_\delta^1, s]} e_k(r) < y \text{ and } e_k(s) < \delta \right\}$$

and

$$B_{x,y} := \left\{ x < \sup_{r \in [T_\delta^2, \zeta]} e_k(r) < y \right\}.$$

For future reference, we also let  $A = A_{M_{j+1}, M_j}$  and  $B = B_{M_{j+1}, M_j}$ .

We claim that

$$\mathbb{P}(A|\mathcal{K}_t) = \mathbb{P}(B|\mathcal{K}_t) = \frac{2\delta}{Z_t} + o(\delta), \quad a.s. \quad (3.8)$$

and that  $\mathbb{P}(A \cap B | \mathcal{K}_t) = O(\delta^2)$ , *a.s.*, so that, up to terms of order  $o(\delta)$ ,  $\mathbb{P}(A \cup B | \mathcal{K}_t) = 4\delta/Z_t$ , *a.s.*, as required in (3.6). By time-reversal symmetry of the Brownian excursion, it is clear that  $\mathbb{P}(A | \mathcal{K}_t) = \mathbb{P}(B | \mathcal{K}_t)$  so we only do the calculations in the case  $B$ .

The key idea to compute  $\mathbb{P}(B)$  is to use the Markov property at time  $T_\delta^2$ . (In fact an additional step is required, since under this conditioning  $e_k$  is not exactly Markov). Let  $m > 0$  and let  $0 < x < y < m$  be two fixed levels (we think of  $y = M_j$  and  $x = M_{j+1}$ ). We claim that  $\mathbb{P}(B)$  can be computed as follows:

$$Q^{(m)}(B_{x,y}) = \frac{\delta}{x} \left(1 - \frac{x}{y}\right) + o(\delta) \quad (3.9)$$

Indeed, consider the unconditional problem first. Recall that the part of the trajectory of a Brownian excursion that comes after  $T_\delta^1$  is simply a Brownian motion (started from  $\delta$ ) run until it hits zero. Let  $\eta > 0$  and let  $M$  be the event that, starting from  $\delta$ , it reaches level  $m$  and from there returns to zero without reaching further than  $m + \eta$ . Starting from  $\delta$ , the probability of reaching  $m$  before hitting 0 is  $\phi_\delta(m) = \delta/m$  by the standard gambler's ruin probability. Once there, the probability of returning to 0 before hitting  $m + \eta$  is  $\eta/(m + \eta)$ . Thus

$$\mathbb{P}^\delta(M) = \phi_\delta(m)\eta/m + o(\eta) \quad (3.10)$$

On the other hand, the event that, started from  $\delta$ , both  $M$  and  $B_{x,y}$  occur can be expressed, by using the strong Markov property at time  $T_\delta^2$ , as:

$$\mathbb{P}^\delta(M \cap B_{x,y}) = \phi_\delta(m) \frac{\eta}{m - \delta} \frac{\delta}{x} \left(1 - \frac{x}{y}\right) + o(\eta) \quad (3.11)$$

In the above expression, the second term expresses the fact that the process, having reached distance  $m$ , only has to return to distance  $\delta$  before reaching  $m + \eta$ . The third term corresponds to hitting  $x$  before returning to zero, while the fourth and last term corresponds to the probability that, having reached level  $x$ , the process  $e$  reaches 0 before  $y$ . All these calculations involve the same gambler's ruin probability already used above. Combining (3.10) with (3.11), we obtain as claimed, up to terms of leading order  $\delta$ :

$$Q^{(m)}(B_{x,y}) = \frac{\delta}{x} \left(1 - \frac{x}{y}\right) + o(\delta). \quad (3.12)$$

Here the term in  $o(\delta)$  depends only on  $m = M_k(s)$ . Rewriting (3.12), we find:

$$Q^{(m)}(B_{x,y}) = \delta \left(\frac{1}{x} - \frac{1}{y}\right) + o(\delta). \quad (3.13)$$

Taking the expectation:

$$\begin{aligned} Q^{(m)}(B) &= \delta E \left( \frac{1}{M_j(t)} - \frac{1}{M_{j+1}(t)} \right) + o(\delta) \\ &= \delta [\mathbb{E}(1/M_j(s)) - \mathbb{E}(1/M_{j+1}(s))] + o(\delta). \end{aligned} \quad (3.14)$$

Now, observe that by the Markov property of the underlying continuous random tree, we have above level  $s$  a Poisson point process of excursions with intensity  $\text{It}\delta$ 's excursion measure  $n$ , run for a local time  $z = Z_s$ . Thus, the number  $N_{>a}$  of excursions higher than some fixed level  $a > 0$  is distributed as:

$$N_{>a} = \text{Poisson}(z/2a)$$

since  $n(\sup e > a) = 1/2a$ . Therefore,

$$\begin{aligned}\mathbb{P}(M_j(s) > a) &= \mathbb{P}(N_{>a} \geq j) \\ &= 1 - e^{-z/2a} - \dots - e^{-z/2a} (z/2a)^{j-1} / (j-1)!\end{aligned}$$

from which it follows that  $M_j$  has a density equal to

$$\mathbb{P}(M_j \in (a, a + da)) = \frac{e^{-z/2a} (z/2a)^j}{a(j-1)!}$$

This implies:

$$\begin{aligned}\mathbb{E}(1/M_j) - \mathbb{E}(1/M_{j+1}) &= \int_0^\infty \frac{1}{a} e^{-z/2a} [(z/2a)^j / a(j-1)! - (z/2a)^{j+1} / a(j)!] da \\ &= \frac{1}{j!} \int_0^\infty \frac{1}{a} e^{-z/2a} (z/2a)^j [j/a - z/2a^2] da\end{aligned}$$

using the change of variable  $u = z/2a$ , we obtain:

$$\begin{aligned}\mathbb{E}(1/M_j) - \mathbb{E}(1/M_{j+1}) &= \frac{1}{j!} \left( - \int_0^\infty e^{-u} u^j \frac{2u}{z} \frac{2ju}{z} \frac{z}{2u^2} du + \int_0^\infty e^{-u} u^j \frac{2u}{z} du \right) \\ &= \frac{2}{zj!} \int_0^\infty e^{-u} u^{j+1} du - \frac{2}{z(j-1)!} \int_0^\infty e^{-u} u^j du \\ &= \frac{2}{z} \left( \frac{(j+1)!}{j!} - \frac{j!}{(j-1)!} \right) \\ &= \frac{2}{z}.\end{aligned}$$

The cancellation of terms involving  $j$  is quite remarkable. Since  $z = Z_t$ , we have thus proved the claim (3.8):

$$\mathbb{P}(B) = 2/Z_t.$$

As explained before, this proves the claim (3.6).

#### 4. Proof of Corollary 1.3

Let  $(B_s, 0 \leq s \leq \zeta)$  be a Brownian excursion conditioned to reach 1, and let  $(\kappa_t, t \geq 0)$  be the realization of Kingman's coalescent described in Theorem 1.1. Note that by Theorem 1.1, one can obtain the frequencies of the blocks of  $\kappa_t$  by simply considering the excursions of  $B$  above level  $U(t)$  with a mass proportional to the quantity of local time that they accumulate at level 1. Thus, to prove Corollary 1.3, the first step consists of computing the number of excursions above a given level  $u < 1$  which carry a given amount of local time at 1. For  $0 < u < 1$ , and  $x \geq 0$ , let  $N(u, x)$  be the number of excursions above level  $1 - u$  that carry a local time at 1 greater than  $x$ .

**Lemma 4.1.** *We have the almost sure convergence:*

$$\frac{2u}{Z_1} N(u, 2ux) \longrightarrow e^{-x} \quad (4.1)$$

for all  $x \geq 0$  as  $u \rightarrow 0$ .

**Proof.** To see this, observe that by excursion theory,  $N(u) := N(u, 0)$  is a Poisson random variable with mean  $Z_{1-u}/(2u)$ . Moreover, conditionally on  $N(u, 0) = n$ , the  $n$  excursions that reach level 1 are i.i.d. realizations of the Itô measure conditioned to exceed level  $u$ . Now, it is well-known that such an excursion, upon reaching  $u$ , behaves afterwards as Brownian motion killed when returning to 0, so by standard properties of Brownian motion local time (see Proposition 4.6 chapter VI in Revuz and Yor (1999)) we have that the amount of local time accumulated by such an excursion at level  $u$  is exponentially distributed with mean  $2u$ . Therefore, by Poisson thinning, conditionally on  $Z_{1-u} = z$ ,

$$N(u, x) \stackrel{d}{=} \text{Poisson} \left( \frac{z}{2u} e^{-x/(2u)} \right). \quad (4.2)$$

Considering excursions with local time greater than  $2ux$ , let

$$X_u = \frac{2ue^x}{Z_{1-u}} N(u, 2ux).$$

Note that, by almost sure continuity of  $Z_s$  near  $s = 1$ , it suffices to prove that  $X_u \rightarrow 1$  almost surely as  $u \rightarrow 0$ . Next we recall the following standard Chernoff bound for a Poisson random variable  $Y$  with parameter  $\mu$ : for all  $h \geq 0$ ,

$$\mathbb{P}(Y > \mu(1 + \delta)) \leq e^{-h\mu(1+\delta)} \mathbb{E}(e^{hY}) = \exp(\mu(e^h - 1 - h) - h\mu\delta).$$

Since  $e^h - 1 - h \sim h^2/2$  for  $h \rightarrow 0$ , we see that  $-\lambda = e^h - 1 - h - h\delta < 0$  for sufficiently small  $h$ . Thus, using (4.2), we are led to the estimate:

$$\mathbb{P}(|X_u - 1| > \delta | Z_{1-u} = z) \leq C e^{-\lambda z/2u}$$

for some  $\lambda > 0$  and some  $C > 0$ . Taking the expectation in the above, we get

$$\mathbb{P}(|X_u - 1| > \delta) \leq \mathbb{E}(C e^{-\lambda Z_{1-u}/2u}).$$

For  $k \geq 1$ , let  $u_k = 1/k^2$ . Note that for  $0 \leq s \leq 1$ ,  $Z_s$  dominates stochastically an exponential random variable with mean  $2s$ . Indeed,  $Z_s$  is greater than the local time accumulated by  $B$  at level  $s$  after  $T_1$ , the hitting time of 1 by  $B$ . Since  $(B_{T_1+t}, t \geq 0)$  has the distribution of a Brownian motion started at 1 killed upon hitting zero, we may apply again Proposition 4.6 in Chapter VI of Revuz and Yor (1999), the claim follows. We deduce that

$$\begin{aligned} \sum_{k=2}^{\infty} \mathbb{E}(C e^{-k^2 \lambda Z_{1-u_k}/2}) &\leq C \sum_{k=2}^{\infty} \int_0^{\infty} \frac{e^{-\lambda k^2 x} e^{-x/(2-2/k^2)}}{2-2/k^2} dx \\ &\leq C \sum_{k=2}^{\infty} \frac{1}{\lambda k^2 (2-2/k^2) + 1} < \infty. \end{aligned}$$

Thus by the Borel-Cantelli lemma, we get that

$$X_{u_k} \rightarrow 1 \text{ almost surely as } k \rightarrow \infty. \quad (4.3)$$

Now, to obtain almost sure convergence for other values of  $u$ , let  $k$  be such that  $u_{k+1} < u \leq u_k$ , and consider the process  $t \mapsto N(t, 2ux)$ ,  $t \in [u_{k+1}, u_k]$ . Note that, for a given size  $y \geq 0$  say, the difference  $|N(u, y) - N(u_k, y)|$  is bounded by the total number of excursions that coalesce during the interval  $(u_{k+1}, u_k]$ . To see this observe that  $u \mapsto N(u, y)$  evolves by jumps of size 1 (either two excursions that had masses smaller than  $y$  coalesce to give birth to an excursion of mass at least  $y$  and the jump is positive, or two excursions of masses larger than  $y$  coalesce, in

which case  $N(u, y)$  decreases by 1). Hence for each coalescence event in  $(u_{k+1}, u_k]$  the process  $N(u, y)$  changes by at most 1.

Thus we have that for all  $u \in [u_{k+1}, u_k]$ :

$$|N(u, 2ux) - N(u_k, 2ux)| \leq |N(u_{k+1}) - N(u_k)|.$$

Since  $N(u, x)$  is monotone in  $x$ , we obtain that for every  $u \in (u_{k+1}, u_k]$ ,

$$\begin{aligned} |N(u, 2ux) - N(u_k, 2u_kx)| &\leq |N(u_{k+1}) - N(u_k)| + |N(u_k, 2ux) - N(u_k, 2u_kx)| \\ &\leq |N(u_{k+1}) - N(u_k)| + |N(u_k, 2u_{k+1}x) - N(u_k, 2u_kx)| \\ &\leq 2|N(u_{k+1}) - N(u_k)| + |N(u_{k+1}, 2u_{k+1}x) - N(u_k, 2u_kx)| \\ &= \Delta_k, \end{aligned}$$

say, where we have set

$$\Delta_k := 2|N(u_{k+1}) - N(u_k)| + |N(u_{k+1}, 2u_{k+1}x) - N(u_k, 2u_kx)|. \quad (4.4)$$

Multiplying by  $u$  and letting  $X'_u = uN(u, 2ux)$  (so that (4.1) is equivalent to  $X'_u \rightarrow Z_1 e^{-x}/2$  almost surely as  $u \rightarrow 0$  for all  $x \geq 0$ ), we get:

$$\begin{aligned} \sup_{u \in (u_{k+1}, u_k]} |X'_u - X'_{u_k}| &= \sup_{u \in (u_{k+1}, u_k]} |u(N(u, 2ux) - N(u_k, 2u_kx)) \\ &\quad + (u - u_k)N(u_k, 2u_kx)| \\ &\leq \sup_{u \in (u_{k+1}, u_k]} \{u\Delta_k + |u - u_k|N(u_k, 2u_kx)\} \\ &\leq u_k\Delta_k + \frac{|u_{k+1} - u_k|}{u_k} X'_{u_k}. \end{aligned} \quad (4.5)$$

It is plain that the second term on the right-hand side converges almost surely to 0. To see that  $u_k\Delta_k \rightarrow 0$  as well, it suffices to observe that by (4.3) applied respectively with  $x = 0$  and  $x > 0$ , we see that  $N(u_k) \sim C_1 k^2$  while  $N(u_k, 2u_kx) \sim C_2 k^2$  for some random  $C_1, C_2 > 0$ . It follows that both terms in the right-hand side of (4.4) are  $o(k^2)$ , i.e.,  $u_k\Delta_k \rightarrow 0$  almost surely. Thus the left-hand side of (4.5) converges to 0 as well, and this implies  $X'_u \rightarrow Z_1 e^{-x}/2$  almost surely as  $u \rightarrow 0$ . This finishes the proof of the lemma.  $\square$

We trivially obtain from Lemma 4.1 a result first derived by Aldous (1999) (see his equation (35)). Let  $K(t, x)$  be the number of blocks in Kingman's coalescent at time  $t$  that are greater than  $x$ .

**Lemma 4.2.** *For every  $x \geq 0$ , we have the almost sure convergence as  $t \rightarrow 0$ :*

$$\frac{t}{2} K(t, tx/2) \longrightarrow e^{-x}. \quad (4.6)$$

**Proof.** This is a trivial consequence of Theorem 1.1 and Lemma 4.1. Indeed, since every block at time  $t$  corresponds to an excursion above level  $u$  with  $u = U(t)$  by Theorem 1, and since the mass of a block is given by the renormalized amount of local time it accumulates at level 1, one may write:

$$K(t, tx/2) = N(u, y)$$

where  $u = 1 - U(t) \sim tZ_1/4$  as  $t \rightarrow 0$ , and  $y = txZ_1/2$ . Therefore, using the almost sure convergence result in Lemma 4.1 and making the necessary cancellations, we obtain the desired estimate (4.6).  $\square$



Corollary 1.3 follows directly from (4.6), since if  $B(t)$  denotes the mass of a block randomly chosen among the  $K(t, 0)$  blocks present at time  $t$  (uniformly at random), then  $\mathbb{P}(B(t) > tx/2) = \mathbb{E}(K(t, tx/2)/K(t, 0)) \rightarrow e^{-x}$  by the Lebesgue dominated convergence theorem. It follows that:

$$\frac{2B(t)}{t} \xrightarrow{d} E$$

where  $E$  is an exponentially distributed random variable with mean 1. Now, observe that the distribution of the frequency  $F(t)$  of the block containing 1 is nothing but a size-biased version of the law of  $B(t)$ , that is, for every nonnegative Borel function  $f$ , we have (denoting  $B = B(t)$  and  $F = F(t)$ ):

$$E\{f(F)\} = E\{Bf(B)\}.$$

By considering for instance for every  $a > 0$ ,  $\mathbb{P}(F \leq a) = \mathbb{E}(B\mathbf{1}_{\{B \leq a\}})$ , and the Lebesgue convergence theorem, we conclude that as  $t \rightarrow 0$ ,

$$\frac{F}{2t} \xrightarrow{d} \hat{E}$$

where  $\hat{E}$  has a size-biased exponential distribution. That is,  $\hat{E}$  has the distribution  $xe^{-x}dx$  which is a Gamma(2) distribution. Equivalently,  $\hat{E}$  is the sum of two standard exponential random variables  $E + E'$ . This concludes the proof of Corollary 1.3.

## 5. Proof of Theorem 1.4

The main ingredient for the proof of Theorem 1.4 is the definition of *reduced tree* associated with our Brownian excursion  $(B_s, 0 \leq s \leq \zeta)$ , and results about the multifractal spectrum of the Branching measure of Galton Watson trees due to Mörters and Shieh (2002). Formally, the reduced tree  $\mathbb{T}$  can be described by saying that  $(B_s, 0 \leq s \leq \zeta)$  encodes a continuum random tree  $\mathcal{T}$  with a metric  $d$  and a root  $o$  as in Aldous (1991). Each vertex  $z \in \mathcal{T}$  has a unique geodesic  $\gamma_z : [0, 1] \rightarrow \mathcal{T}$  that connects it to  $o$ , such that if  $d(o, z) = \rho$ , then  $d(o, \gamma_z(t)) = \rho t$  for all  $0 \leq t \leq 1$ . We define  $\mathbf{T}$  by

$$\mathbf{T} = \bigcup_{z \in \mathcal{T}: d(z, 0) = 1} \gamma_z([0, 1]).$$

Informally,  $\mathbf{T}$  is a continuum random tree obtained from  $\mathcal{T}$  by taking away (“pruning”) every vertex whose descendance does not reach distance 1 from the root. Equivalently, this is the tree which, at level  $0 \leq u < 1$ , has as many branches as there are excursions above level  $u$  that reach level 1. Thus if we let  $|\mathbf{T}(u)|$  be the number of branches of  $\mathbf{T}$  at level  $0 \leq u < 1$ , then we have by definition  $|\mathbf{T}(u)| = N(u)$ , the number of excursions above  $u$  reaching level 1.

The process  $|\mathbf{T}(u)|$  is a variant of a process already considered by Neveu and Pitman in a seminal paper Neveu and Pitman (1989). The key observation is that:

$$\{|\mathbf{T}(1 - e^{-t})|, t \geq 0\} \stackrel{d}{=} \{Y_t, t \geq 0\}, \quad (5.1)$$

where  $\{Y_t, t \geq 0\}$  is a (rate 1) Yule process (see Le Gall (1989, section 5)). This is a continuous-time Galton-Watson process where individuals die at rate 1 to give birth to exactly two offsprings. In fact, an even stronger property holds. Consider

the random tree  $\mathbb{T}$  obtained from  $\mathbf{T}$  by applying the same exponential time-change as in (5.1). That is, for all  $z \in \mathcal{T}$  with  $d(z, o) = 1$ , define  $\gamma'_z(t) = \gamma_z(1 - e^{-t})$ . Let

$$\mathbb{T} := \bigcup_{z \in \mathcal{T}: d(z, o) = 1} \gamma'_z([0, \infty)). \quad (5.2)$$

Then  $\mathbb{T}$  is a Yule tree. We can thus define the boundary  $\partial\mathbb{T}$  of the tree  $\mathbb{T}$  by taking  $\partial\mathbb{T}$  to be the set of *rays*, i.e., the set of all non-backtracking  $\mathbb{T}$ -valued paths  $(\zeta(t), t \geq 0)$  such that  $\zeta(t)$  is at distance  $t$  from the root. The boundary  $\partial\mathbb{T}$  is naturally equipped with a measure  $\mu$ , called the *branching measure*, which is defined as follows. Thanks to *Kesten-Stigum* theorem:

$$e^{-t}|\mathbb{T}(t)| \xrightarrow{d} W \quad (5.3)$$

where  $W > 0$  almost surely. In the Yule case,  $W$  is well known to be an exponential random variable with mean 1. We next define a metric  $\delta$  on  $\partial\mathbb{T}$  by declaring that for  $\zeta, \zeta' \in \partial\mathbb{T}$ ,  $\delta(\zeta, \zeta') = e^{-t}$  where  $t$  is the time of the “most recent common ancestor”, that is,  $t = \sup\{s \geq 0 : \zeta(s) = \zeta'(s)\}$ . ( $\delta$  is the so-called genealogical metric on  $\partial\mathbb{T}$ ). A ball  $B$  of radius  $t$  for  $\delta$  consists of all rays which pass through a given vertex  $z \in \mathbb{T}$ , at distance  $t$  from the root. The subtree containing  $z$  is itself a Yule tree and hence the Kesten-Stigum theorem applies to it, let  $W(z)$  be the associated Kesten-Stigum random variable as in (5.3). Then define

$$\mu(B) = e^{-t}W(z).$$

It is easy to see that  $\mu$  satisfies the assumptions of Carathéodory’s extension theorem (see, e.g., Durrett (1996), p. 444) and hence defines a finite measure on  $\partial\mathbb{T}$ , with total mass  $\mu(\partial\mathbb{T}) = W(o)$ . It will also be convenient to define the probability measure  $\mu^\sharp(\cdot) = \mu(\cdot)/\mu(\partial\mathbb{T})$ .

Then our key claim is that we can identify  $(\partial\mathbb{T}, \mu^\sharp)$  with Evans’ metric space  $(S, \eta)$ :

$$(\partial\mathbb{T}, \mu^\sharp) \cong (S, \eta), \quad (5.4)$$

in the sense that we can find a continuous one-to-one map  $\Phi : S \rightarrow \partial\mathbb{T}$  such that  $x \in S$  and  $\zeta = \Phi(x)$ ,

$$\eta\{B(x, t)\} = \mu^\sharp\{B(\zeta, 1 - U(t))\}, \quad (5.5)$$

where  $U(t)$  is the time-change appearing in Theorem 1.1. To do this, we first define  $\Phi$  on the integers. For  $i \in \mathbb{N}$ , simply define  $\Phi(i)$  to be the ray associated with excursion  $\varepsilon_i$ , where  $\varepsilon_1, \varepsilon_2, \dots$  are the excursions of  $B$  above level 1, ordered by their respective heights. The definition of  $\Phi$  is easily extended to  $x \in S$  by taking a suitable sequence  $x_n \in \mathbb{N}$  with  $x_n \rightarrow x$  (see the proof of Theorem 5 in Berestycki et al. (2007) for further details).

Having made the identification (5.5), it turns out that Theorem 1.4 is now an easy application of Theorem 1.2 in Mörters and Shieh (2002) and Theorem 1 in Mörters and Shieh (2004). These results are derived for a discrete-time Galton-Watson tree  $T$ , and state the following. Assume that  $T$  is a discrete-time Galton-Watson tree and that the number of offsprings  $L$  of an individual is such that for some  $r > 0$ ,  $\mathbb{E}(\exp(rL)) = \infty$  but  $\mathbb{E}(\exp(tL)) < \infty$  if  $t < r$ , and  $L$  is unbounded. Then we have, by Theorem 1.2 in Mörters and Shieh (2002) (corresponding to the thick part of the spectrum), for all  $0 \leq \theta \leq a/r$ :

$$\dim \left\{ \xi \in \partial\mathbb{T} : \limsup_{n \rightarrow \infty} \frac{\mu(B(\xi, e^{-n}))}{m^{-n}} = \theta \right\} = a - r\theta, \quad (5.6)$$

where  $a = \log \mathbb{E}(L)$  and  $m = \mathbb{E}(L)$ . Moreover when  $\theta = a/r$  the above set is non-empty (see Lemma 3.3 (ii) in Mörters and Shieh (2002)). Assuming further that  $\mathbb{P}(L = 1) > 0$  (the *Schröder case*) and letting  $\tau = -\log \mathbb{P}(L = 1)/a$  we have, by Theorem 1 in Mörters and Shieh (2004), (corresponding to the thin part of the spectrum), for all  $a \leq \theta \leq a(1 + 1/\tau)$ :

$$\dim \left\{ \xi \in \partial \mathbb{T} : \limsup_{n \rightarrow \infty} \frac{-\log \mu(B(\xi, e^{-n}))}{n} = \theta \right\} = a \left( \frac{a}{\theta} (1 + \tau) - \tau \right). \quad (5.7)$$

Moreover when  $\theta = a(1 + 1/\tau)$ , the above set is non-empty almost surely. To use these results in our case, let  $T$  be the discrete-time tree obtained by sampling  $\mathbb{T}$  at discrete times  $1, 2, \dots$ , so  $T$  is a discrete-time Galton-Watson process which belongs to the Schröder class. The random variable  $W = W(o)$  is of course unchanged so it is an exponential random variable with mean 1 (note that  $W$  is only equal to half the local time  $Z_1$ , corresponding to a “one-sided” Lévy approximation of  $Z_1$  from below: see, e.g., (1.11) and (1.19) in Chapter VI of Revuz and Yor (1999)). The distribution of  $L$  is not particularly nice to write down but it is unbounded and we may nonetheless identify the parameters  $a$ ,  $r$  and  $\tau$  as follows. Note first that since  $W$  is exponential with parameter 1,  $r = 1$ . Moreover, since

$$\{e^{-t}Y_t, t \geq 0\} \text{ is a martingale}$$

we obtain  $m = \mathbb{E}(Y_1) = \mathbb{E}(L) = e$ , so  $a = \log \mathbb{E}(L) = 1$ . We obtain  $\tau$  by computing  $\mathbb{P}(L = 1) = e^{-1}$  since every individual branches at rate 1, and hence  $\tau = 1$  as well. Using (5.5) together with (5.6) and (5.7), it is now straightforward to deduce Theorem 1.4. The details are left to the reader and are similar (in fact, much easier) than the proof of Lemma 26 in Berestycki et al. (2007). Note in particular that, since the construction of Kingman's coalescent in Theorem 1 holds at a fixed deterministic level 1, we avoid the use of Lemma 24 in Berestycki et al. (2007).

### Acknowledgements

We thank Christina Goldschmidt for sharing with us her thoughts about the Martin boundary of Galton-Watson trees, and we thank Ed Perkins for useful discussions. This work started when J.B. was at Université de Provence in Marseille, and N.B. was a postdoc at University of British Columbia. J.B. thanks the Math department of UBC for their invitation, during which this work started.

### References

- D. Aldous. The continuum random tree. I. *Ann. Probab.* **19** (1), 1–28 (1991).
- D. Aldous. The continuum random tree. III. *Ann. Probab.* **21** (1), 248–289 (1993).
- D. Aldous. Brownian excursion conditioned on its local time. *Electron. Comm. Probab.* **3**, 79–90 (electronic) (1998).
- D. Aldous. Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. *Bernoulli* **5** (1), 3–48 (1999).
- J. Berestycki, N. Berestycki and V. Limic. Interpreting  $\lambda$ -coalescent speed of coming down from infinity via particle representation of super-processes. (2008a). In preparation.
- J. Berestycki, N. Berestycki and J. Schweinsberg. Beta-coalescents and continuous stable random trees. *Ann. Probab.* **35** (5), 1835–1887 (2007).

- J. Berestycki, N. Berestycki and J. Schweinsberg. Small-time behavior of beta coalescents. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** (2), 214–238 (2008b).
- J. Bertoin. *Random fragmentation and coagulation processes*, volume 102 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge (2006).
- J. Bertoin and C. Goldschmidt. Dual random fragmentation and coagulation and an application to the genealogy of Yule processes. In *Mathematics and computer science. III*, Trends Math., pages 295–308. Birkhäuser, Basel (2004).
- T. Duquesne and J.-F. Le Gall. Random trees, Lévy processes and spatial branching processes. *Astérisque* (281), vi+147 (2002).
- R. Durrett. *Probability: theory and examples*. Duxbury Press, Belmont, CA, second edition (1996).
- A. Etheridge. *An introduction to superprocesses*, volume 20 of *University Lecture Series*. American Mathematical Society, Providence, RI (2000).
- S. Evans. Kingman’s coalescent as a random metric space. In *Stochastic models (Ottawa, ON, 1998)*, volume 26 of *CMS Conf. Proc.*, pages 105–114. Amer. Math. Soc., Providence, RI (2000).
- C. Goldschmidt. private communication (2008).
- J. F. C. Kingman. The coalescent. *Stochastic Process. Appl.* **13** (3), 235–248 (1982).
- J.-F. Le Gall. The uniform random tree in a Brownian excursion. *Probab. Theory Related Fields* **96** (3), 369–383 (1993).
- Jean-François Le Gall. Marches aléatoires, mouvement brownien et processus de branchement. In *Séminaire de Probabilités, XXIII*, volume 1372 of *Lecture Notes in Math.*, pages 258–274. Springer, Berlin (1989).
- R. Lyons. Random walks and percolation on trees. *Ann. Probab.* **18** (3), 931–958 (1990).
- P. Mörters and N.-R. Shieh. Thin and thick points for branching measure on a Galton-Watson tree. *Statist. Probab. Lett.* **58** (1), 13–22 (2002).
- P. Mörters and N.-R. Shieh. On the multifractal spectrum of the branching measure on a Galton-Watson tree. *J. Appl. Probab.* **41** (4), 1223–1229 (2004).
- J. Neveu and J. W. Pitman. The branching process in a Brownian excursion. In *Séminaire de Probabilités, XXIII*, volume 1372 of *Lecture Notes in Math.*, pages 248–257. Springer, Berlin (1989).
- E. Perkins. On the continuity of measure-valued processes. In *Seminar on Stochastic Processes, 1990 (Vancouver, BC, 1990)*, volume 24 of *Progr. Probab.*, pages 261–268. Birkhäuser Boston, Boston, MA (1991).
- D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition (1999).
- J. Warren. On a result of David Aldous concerning the trees in a conditioned excursion. *Electron. Comm. Probab.* **4**, 25–29 (electronic) (1999).
- J. Warren and M. Yor. The Brownian burglar: conditioning Brownian motion by its local time process. In *Séminaire de Probabilités, XXXII*, volume 1686 of *Lecture Notes in Math.*, pages 328–342. Springer, Berlin (1998).