Kingman's coalescent as a random metric space

Steven N. Evans *

Department of Statistics, University of California at Berkeley

Abstract

Kingman's coalescent is a Markov process with state–space the collection of partitions of the positive integers. Its initial state is the trivial partition of singletons and it evolves by successive pairwise mergers of blocks. The coalescent induces a metric on the positive integers: the distance between two integers is the time until they both belong to the same block. We investigate the completion of this (random) metric space. We show that almost surely it is a compact metric space with Hausdorff and packing dimension both 1, and it has positive capacities in precisely the same gauges as the unit interval.

1 Introduction

Kingman's coalescent was introduced in [Kin82b, Kin82a] as a model for genealogies in the context of population genetics. This process has since been the subject of a large amount of applied and theoretical work. We refer the reader to [Ald97] for a recent survey and bibliography covering coalescent models in general, and [Tav84, Wat84] for an indication of some of the applications of Kingman's coalescent in genetics.

Here is a quick description of Kingman's coalescent (which we will hereafter simply refer to as the coalescent). Recall that a *partition* of a set S is a collection $\{A_\lambda\}$ of subsets of S (the A_λ are called the *blocks* of the partition) such that $A_\lambda \cap A_\mu = \emptyset$ for $\lambda \neq \mu$ and $\bigcup_\lambda A_\lambda = S$. Let Π denote the collection of partitions of $\mathbb{N} := \{1, 2, \ldots\}$. For $n \in \mathbb{N}$ let Π_n denote the collection of partitions of $\mathbb{N}_n := \{1, 2, \ldots, n\}$. Each partition π in Π (resp. Π_n) corresponds to an *equivalence*

*Research supported in part by NSF grant DMS-9703845

relation $\sim_\pi$ on $\mathbb{N}$ (resp. $\mathbb{N}_n$) by setting $i \sim_\pi j$ if i and j belong to the same block of π. Write ρ_n for the natural restriction map from Π onto Π_n. Kingman [Kin82b] showed that there was a (unique in law) Π–valued Markov process ξ such that for all $n \in \mathbb{N}$ the restricted process $\xi_n := \rho_n \circ \xi$ is a Π_n–valued, time–homogeneous Markov chain with initial state $\xi_n(0)$ the trivial partition $\{\{1\}, \ldots, \{n\}\}$ and the following transition rates: if ξ_n is in a state with k blocks, then

- a jump occurs at rate $\binom{k}{2}$,
- the new state is one of the $\binom{k}{2}$ partitions that can be obtained by merging two blocks of the current state,
- and all such possibilities are equally likely.

There is a natural (random) metric δ on $\mathbb{N}$ defined by

$$\delta(i,j) := \inf\{t \geq 0 : i \sim_{\xi(t)} j\}.$$

In the original interpretation of the coalescent as the random genealogical tree of a countable collection of individuals (with time run backwards from the present), the distance $\delta(i,j)$ is just how long before the present the respective lines of descent of i and j diverged. Note that δ is actually an *ultrametric* on $\mathbb{N}$; that is,

$$\delta(i,j) \leq \delta(i,k) \vee \delta(k,j) \text{ for all } i,j,k \in \mathbb{N}.$$

Let $(\mathbb{S}, \delta)$ denote the *completion* of $(\mathbb{N}, \delta)$. Clearly, the extension of δ to $\mathbb{S}$ is also an ultrametric. Before presenting our main theorem giving some of the properties of $\mathbb{S}$, we will take some time to sketch a description of $\mathbb{S}$ that some readers might find helpful.

Recall that a *rooted tree* is a directed graph with the properties that, with the exception of a a unique vertex (the *root*), every vertex has exactly one directed edge leading to it and the corresponding undirected graph is connected and acyclic. If two vertices v and w of a rooted tree are conncected by a directed edge leading from v to w, then w is said to be a *child* of v.

As we recall in Section 2, almost surely the random partition $\xi(t)$ has finitely many blocks for all $t > 0$, ξ evolves by blocks coalescing in pairs, and $\xi(t)$ consists of the single block $\mathbb{N}$ for all t sufficiently large. Let $\mathcal{A}$ denote the collection of subsets A of the integers such that A is a block of $\xi(t)$ for some $t > 0$. We can think of the elements of $\mathcal{A}$ as the vertices of a tree rooted at $\mathbb{N}$: a block $A \in \mathcal{A}$ is the child of a block $C \in \mathcal{A}$ if $A \in \xi(s)$ and $C \in \xi(u)$ for some pair of times $s < u$ and there is another (unique) block $B \in \mathcal{A}$ such that A coalesces with B at some time $s < t \leq u$ to form C. Almost surely, each vertex in this rooted tree has two children.

In the usual terminology, an *end* of this rooted tree is an infinite directed path starting at the root, that is, an infinite sequence $\mathbb{N} = A_1, A_2, \ldots$ of blocks such that A_{n+1} is a child of A_n for all n. It is not hard to show that there is a one–to–one relationship between $\mathbb{S}$ and the set of ends.

The correspondence between coalescing partitions, tree structures and ultrametrics is a familiar idea, particularly in the physics literature (see, for example, [MPV87]). Some properties of the space $(\mathbb{N}, \delta)$ are considered explicitly in Section 4 of [Ald93].

It is our aim in this paper to investigate some of the dimension and *capacity* properties of $(\mathbb{S}, \delta)$. We remark that there is a large literature on such “fractal” properties of random trees constructed in various ways from Galton–Watson branching processes; for example, [Haw81] computed the Hausdorff dimension of the boundary of a Galton–Watson tree equipped with a natural metric (see also [Lyo90, LS98]). We refer the reader to [LP96] for an account, including an extensive bibliography, of this and other facets of probability on trees.

We remind the reader of the definitions and basic properties of Hausdorff dimension, packing dimension, energy and capacity. For more detail see [Mat95], [RT61] and [TT85]

Let (T, ρ) be a metric space. Consider a non-decreasing, continuous function $g : \mathbb{R}_+ \to \mathbb{R}_+$ with $g(0) = 0$. The *Hausdorff outer measure* of a set $A \subseteq T$ with respect to the *measure function* g is defined by

$$m_g(A) := \lim_{\varepsilon \downarrow 0} m_g^\varepsilon(A),$$

where

$$m_g^\varepsilon(A) := \inf \left\{ \sum_i g(\mathrm{diam}(B_i)) \right\},$$

with the infimum taken over all countable collections of open balls $B_1, B_2, \ldots$ such that $A \subseteq \bigcup_i B_i$ (that is, $B_1, B_2, \ldots$ is a *cover* of A) and $\sup_i \mathrm{diam}(B_i) \leq \varepsilon$. The *Hausdorff dimension* of such a set A is given by

$$\inf\{\alpha > 0 : m_{g_\alpha}(A) = 0\} = \sup\{\alpha > 0 : m_{g_\alpha}(A) = \infty\},$$

where $g_\alpha(s) := s^\alpha$, $s \geq 0$.

Let (T, ρ), g and A be as above. The *packing premeasure* of A with respect to the measure function g is defined by

$$P_g(A) := \lim_{\varepsilon \downarrow 0} P_g^\varepsilon(A),$$

where

$$P_g^\varepsilon(A) := \sup \left\{ \sum_i g(\mathrm{diam}(B_i)) \right\},$$

with the supremum taken over all countable collections of pairwise disjoint open balls $B_1, B_2, \ldots$ with centres in A (that is, $B_1, B_2, \ldots$ is a *packing* of A) such that $\sup_i \mathrm{diam}(B_i) \leq \varepsilon$. (We remark that it (T, ρ) is an ultrametric space, then any point of a ball B is a centre, and so in this case the requirement that the balls $B_1, B_2, \ldots$ are centred in A is equivalent to the requirement that they intersect A.) The *packing outer measure* of A with respect to the measure function g is then defined to be

$$p_g(A) := \inf \left\{ \sum_i P_g(A_i) \right\},$$

where the infimum is over all countable collections of Borel sets $A_1, A_2, \ldots$ such that $A \subseteq \bigcup_i A_i$. The *packing dimension* of A is given by

$$\inf\{\alpha > 0 : p_{g_\alpha}(A) = 0\} = \sup\{\alpha > 0 : p_{g_\alpha}(A) = \infty\},$$

with g_α as above.

By arguments similar to those in Lemma 5.11 of [TT85] it is possible to show that the inequality $m_g(A) \leq p_g(A)$ always holds and so, in particular, the Hausdorff dimension of a set is at most its packing dimension.

We now recall the definitions of energy and capacity. Again let (T, ρ) be a metric space. Write $M_1(T)$ for the collection of (Borel) probability measures on T. A *gauge* is a function $f : [0, \infty[\to [0, \infty]$, such that:

- f is continuous and non-increasing,
- $f(0) = \infty$,
- $f(r) < \infty$ for $r > 0$,
- $\lim_{r\to\infty} f(r) = 0$.

Given $\mu \in M_1(T)$ and a gauge f, the *energy of* μ *in the gauge* f is the quantity

$$\mathcal{E}_f(\mu) := \int \mu(dx) \int \mu(dy)\, f(\rho(x, y)).$$

The *capacity of* $A \subseteq T$ *in the gauge* f is the quantity

$$\mathrm{Cap}_f(A) := (\inf\{\mathcal{E}_f(\mu)\})^{-1},$$

where the infimum is over probability measures $\mu \in M_1(T)$ with closed support contained in A (note by our assumptions on f that we need only consider diffuse $\mu \in M_1(T)$ in the infimum). The *capacity dimension* of a set $A \subseteq T$ is given by

$$\inf\{\alpha > 0 : \mathrm{Cap}_{f_\alpha}(A) = 0\} = \sup\{\alpha > 0 : \mathrm{Cap}_{f_\alpha}(A) = \infty\},$$

where $f_\alpha(s) := s^{-\alpha}$, $s > 0$.

The capacity dimension of a set equals its Hausdorff dimension (see Ch. 8 of [Mat95]), and hence the capacity dimension is also dominated by the packing dimension.

Our main result is the following, which, *inter alia*, asserts in the terminology of [PP95] (see, also, [BP92, PPS96, Per96]) that $\mathbb{S}$ is a.s. *capacity–equivalent* to the unit interval $[0, 1]$.

Theorem 1.1 *Almost surely, the metric space* $(\mathbb{S}, \delta)$ *is compact, and the Hausdorff and packing dimensions of* $\mathbb{S}$ *are both* 1. *There exist random variables* C^*, C^{**} *such that almost surely* $0 < C^* \leq C^{**} < \infty$ *and for every gauge* f

$$C^*\mathrm{Cap}_f([0, 1]) \leq \mathrm{Cap}_f(\mathbb{S}) \leq C^{**}\mathrm{Cap}_f([0, 1]). \tag{1.1}$$

Let us say a little about the interpretation of Theorem 1.1. Capacities, Hausdorff measures and packing measures are all ways of capturing how large a set is. By definition, knowing that $\mathbb{S}$ has positive capacity in some gauge indicates that $\mathbb{S}$ is large enough to allow mass to be spread "smoothly" on it. It is possible to establish analogues for $\mathbb{S}$ of density results for Euclidean space Hausdorff and packing measures (see Theorems 2.1 and 5.4 of [TT85] for statements of the Euclidean results). These results show that knowing $\mathbb{S}$ has positive Hausdorff or packing measure for some measure function is again equivalent to knowing that $\mathbb{S}$ supports a measure that is "smooth" in an appropriate sense.

The compactness claim of Theorem 1.1 and the fact that the Hausdorff and packing dimensions are at most 1 are established in Section 3. The capacity–equivalence (1.1) is proved in Section 6, and, by the general relationships between Hausdorff, packing and capacity dimensions, this also establishes the required lower bound on the Hausdorff and packing dimensions.

The results of this paper suggest a number of problems for future study. A process of coalescing partitions of $\mathbb{N}$ is constructed in [DEF$^+$98] using coalescing Brownian motions on the circle. The techniques of the present paper are used there to show that the corresponding metric space has Hausdorff and packing dimensions both equal to $\frac{1}{2}$ and that this space is capacity–equivalent to the middle–$\frac{1}{2}$ Cantor set (and hence, by the results of [PPS96], to the Brownian zero set). It is, of course, natural to investigate the existence of exact Hausdorff and packing measure functions for this Brownian model and the model considered here. In this regard, the random measure Γ constructed in Section 5 and its analogue in the setting of [DEF$^+$98] are the natural candidates for applying the abovementioned analogues of the density theorems of [RT61] and [TT85], provided one can obtain the requisite upper and lower densities. Lastly, a number of other coalescing partition–valued processes arising from models in chemistry, cosmology and physics are considered in [EP98, Pit97, BS97] and it would be interesting to investigate the "fractal" properties of the corresponding metric spaces, which are typically not compact.

2 Some observations on the coalescent

We begin by recalling some results about the coalescent from [Kin82b]. Let $N(t)$ denote the number of blocks of the partition $\xi(t)$. Almost surely, $N(t) < \infty$ for all $t > 0$ and the process N is a pure–death Markov chain that jumps from k to $k-1$ at rate $\binom{k}{2}$ for $k > 1$ (the state 1 is a trap). For $k \in \mathbb{N}$, put $\sigma_k := \inf\{t \geq 0 : N(t) = k\}$. The process ξ is constant on each of the intervals $[\sigma_k, \sigma_{k-1}[$, $k > 1$. Write $I_1(t) < \cdots < I_{N(t)}(t)$ for an ordered listing of the least elements of the various blocks of $\xi(t)$. Almost surely, for all $t > 0$ the asymptotic block frequencies

$$F_i(t) := \lim_{n\to\infty} n^{-1} \left|\{j \in \mathbb{N}_n : j \sim_{\xi(t)} I_i(t)\}\right|, \ 1 \leq i \leq N(t),$$

exist (where we use $|A|$ to denote the cardinality of a set A) and

$$F_1(t) + \cdots + F_{N(t)}(t) = 1.$$

It follows from the arguments that lead to Equation (35) in [Ald97] that

$$\lim_{t\downarrow 0} tN(t) = 2, \ a.s. \tag{2.1}$$

Finally, we claim that

$$\lim_{t\downarrow 0} t^{-1} \sum_{i=1}^{N(t)} F_i(t)^2 = 1, \ a.s. \tag{2.2}$$

To see this, set $X_{n,i} := F_i(\sigma_n)$ for $n \in \mathbb{N}$ and $1 \leq i \leq n$, and observe from (2.1) that it suffices to establish

$$\lim_{n\to\infty} n \sum_{i=1}^{n} X_{n,i}^2 = 2, \ a.s. \tag{2.3}$$

By the "paintbox" construction in Section 5 of [Kin82b] (see also Section 4.2 of [Ald97] for an exposition with some helpful pictures) the random variable $\sum_{i=1}^n X_{n,i}^2$ has the same law as $U_{(1)}^2 + (U_{(2)} - U_{(1)})^2 + \cdots + (U_{(n-1)} - U_{(n-2)})^2 + (1 - U_{(n-1)})^2$, where $U_{(1)} \leq \ldots \leq U_{(n-1)}$ are the order statistics corresponding to i.i.d. random variables $U_1, \ldots, U_{n-1}$ that are uniformly distributed on $[0,1]$. By a classical result on the spacings between order statistics of i.i.d. uniform random variables (see, for example, Section III.3.(e) of [Fel71]), the law of $\sum_{i=1}^n X_{n,i}^2$ is thus the same as that of $(\sum_{i=1}^n T_i^2)/(\sum_{i=1}^n T_i)^2$, where $T_1, \ldots, T_n$ are i.i.d. mean one exponential random variables.

Now for any $0 < \varepsilon < 1$ we have, recalling $\mathbb{P}[T_i^2] = 2$,

$$\begin{aligned} &\mathbb{P}\left\{\left(\sum_{i=1}^n T_i^2\right) \Big/ \left(\sum_{i=1}^n T_i\right)^2 > (1+\varepsilon)(1-\varepsilon)^{-2} 2n^{-1}\right\} \\ &\quad \leq \mathbb{P}\left\{\sum_{i=1}^n (T_i^2 - \mathbb{P}[T_i^2]) > 2\varepsilon n\right\} + \mathbb{P}\left\{\sum_{i=1}^n (T_i - \mathbb{P}[T_i]) < -\varepsilon n\right\}. \end{aligned}$$

A fourth moment computation and Markov's inequality show that both terms on the right–hand side are bounded above by $c(\varepsilon) n^{-2}$ for a suitable constant $c(\varepsilon)$. A similar bound holds for

$$\mathbb{P}\left\{\left(\sum_{i=1}^n T_i^2\right) \Big/ \left(\sum_{i=1}^n T_i\right)^2 < (1-\varepsilon)(1+\varepsilon)^{-2} 2n^{-1}\right\}.$$

The claim (2.3) and hence (2.2) now follows by an application of the Borel–Cantelli Lemma. As the referee remarked, the tail estimates needed for the Borel–Cantelli Lemma can also be obtained from Markov's inequality and known moment formulae for Dirichlet distributions.

3 Compactness and upper bounds on dimensions

Given $B \subseteq \mathbb{S}$, write $\mathrm{cl}B$ for the closure of B. Each of the sets

$$\begin{aligned} U_i(t) &= \mathrm{cl}\{j \in \mathbb{N} : j \sim_{\xi(t)} I_i(t)\} \\ &= \mathrm{cl}\{j \in \mathbb{N} : \delta(j, I_i(t)) \leq t\} \\ &= \{y \in \mathbb{S} : \delta(y, I_i(t)) \leq t\} \end{aligned}$$

is a closed ball with diameter at most t (in an ultrametric space, the diameter and radius of a ball are equal). The closed balls of $\mathbb{S}$ are also the open balls and every ball is of the form $U_i(t)$ for some $t > 0$ (see, for example, Proposition 18.4 of [Sch84]) and, in fact, every ball is of the form $U_i(\sigma_k)$ for some $k \in \mathbb{N}$ and $1 \leq i \leq k$. In particular, the collection of balls is countable. Any ball of diameter at most t is contained in a unique one of the $U_i(t)$, and any ball of diameter at least t contains one or more of the $U_i(t)$ (see, for example, Proposition 18.5 of [Sch84]). Moreover, $k-1$ of the balls $U_i(\sigma_k)$, $1 \leq i \leq k$, are of the form $U_j(\sigma_{k+1})$ for some $1 \leq j \leq k+1$; and the remaining ball is of the form $U_h(\sigma_{k+1}) \cup U_\ell(\sigma_{k+1})$ for some pair $1 \leq h, \ell \leq k+1$.

Because $\mathbb{S}$ is complete by definition, in order to show that $\mathbb{S}$ is a.s. compact it suffices by Ascoli's theorem to show that $\mathbb{S}$ is a.s. totally bounded. However, for any $t > 0$ we have a.s. that $\mathbb{S}$ is covered by $N(t) < \infty$ closed balls of diameter at most t.

In order to establish that both the Hausdorff and packing dimensions of $\mathbb{S}$ are a.s. at most 1 it suffices to consider the packing dimension, because packing dimension always dominates Hausdorff dimension.

Recall that a packing of $\mathbb{S}$ is a pairwise disjoint collection of balls in $\mathbb{S}$. By definition of packing dimension, in order to establish that the packing dimension is at most 1 a.s. it suffices to show for each $\alpha > 1$ that there is a random variable C such that $C < \infty$ a.s. and, for any packing $B_1, B_2, \ldots$ of $\mathbb{S}$ with balls of diameter at most 1, we have $\sum_k \operatorname{diam}(B_k)^\alpha \le C$. As we observed above, if $2^{-p} \le \operatorname{diam}(B_k) < 2^{-(p-1)}$ for some $p = 0, 1, 2, \ldots$, then B_k contains one or more of the balls $U_i(2^{-p})$. Thus

$$|\{k \in \mathbb{N} : 2^{-p} \le \operatorname{diam}(B_k) < 2^{-(p-1)}\}| \le N(2^{-p})$$

and

$$\sum_k \operatorname{diam}(B_k)^\alpha \le \sum_{p=0}^{\infty} N(2^{-p}) 2^{-(p-1)\alpha} < \infty,$$

by (2.1), as required.

4 An alternative expression for energy

We need to adapt to our setting the alternative expression for energy obtained by summation–by–parts in Section 2 of [PP95].

For $t > 0$ write $\mathcal{U}(t)$ for the collection of balls $\{U_1(t), \ldots, U_{N(t)}(t)\}$. Let $\mathcal{U}$ denote the union of these collections over all $t > 0$. As remarked in Section 3, the collection $\mathcal{U}$ is just the countable collection of all balls of $\mathbb{S}$. Given $U \in \mathcal{U}$ with $U \neq \mathbb{S}$, let $U^{\rightarrow}$ denote the unique element of $\mathcal{U}$ such that $U \subsetneq U^{\rightarrow}$ and if $V \in \mathcal{U}$ with $U \subseteq V \subseteq U^{\rightarrow}$ then either $V = U$ or $V = U^{\rightarrow}$. More concretely, such a ball U is in $\mathcal{U}(\sigma_k)$ but not in $\mathcal{U}(\sigma_{k-1})$ for some unique $k > 1$, and $U^{\rightarrow}$ is the unique element of $\mathcal{U}(\sigma_{k-1})$ such that $U \subset U^{\rightarrow}$. Define $\mathbb{S}^{\rightarrow} := \dagger$, where $\dagger$ is an adjoined symbol. Put $\operatorname{diam}(\dagger) = \infty$.

Given a gauge f, write φ_f for the diffuse measure on $[0, \infty[$ such that $\varphi_f([r, \infty[) = \varphi_f(]r, \infty[) = f(r)$, $r \ge 0$. For a diffuse probability measure $\mu \in M_1(\mathbb{S})$ we have, with the convention $f(\infty) = 0$,

$$
\begin{aligned}
\mathcal{E}_f(\mu) &= \int \mu(dx) \int \mu(dy)\, f(\delta(x,y)) \\
&= \int \mu(dx) \int \mu(dy) \sum_{U \in \mathcal{U},\, \{x,y\} \subseteq U} f(\operatorname{diam}(U)) - f(\operatorname{diam}(U^{\rightarrow})) \\
&= \sum_{U \in \mathcal{U}} \left(f(\operatorname{diam}(U)) - f(\operatorname{diam}(U^{\rightarrow})) \right) \int \mu(dx) \int \mu(dy)\, \mathbf{1}\{\{x,y\} \subseteq U\} \\
&= \sum_{U \in \mathcal{U}} \left(f(\operatorname{diam}(U)) - f(\operatorname{diam}(U^{\rightarrow})) \right) \mu(U)^2 \\
&= \sum_{U \in \mathcal{U}} \int_{[0,\infty[} \varphi_f(dt) \mathbf{1}\{U \in \mathcal{U}(t)\} \mu(U)^2 \\
&= \int_{[0,\infty[} \varphi_f(dt) \sum_{U \in \mathcal{U}(t)} \mu(U)^2 .
\end{aligned}
\tag{4.1}
$$

5 Construction of a good measure on $\mathbb{S}$

In order to establish the left–hand side of the capacity–equivalence (1.1), it appears, a priori, that for each gauge f we might need to find a random probability measure Γ depending on f such that $C\mathrm{Cap}_f([0,1]) \leq (\mathcal{E}_f(\Gamma))^{-1}$ for some a.s. non–zero random variable C that does not depend on f. It turns out, however, that we can find a Γ that works simultaneously for all gauges f. We construct Γ as follows.

Let $\mathcal{B}$ denote the algebra of subsets of $\mathbb{S}$ generated by the collection of balls $\mathcal{U}$; so that $\mathcal{B}$ is just the countable collection of finite unions of balls. Of course, the σ–algebra generated by $\mathcal{B}$ is the Borel σ–algebra of $\mathbb{S}$. It is clear that, on an event Ω^* with $\mathbb{P}(\Omega^*) = 1$, the sets in $\mathcal{B}$ are compact, and, moreover, for all $k \in \mathbb{N}$ and indices $1 \leq i \leq k$ if $U_i(\sigma_k) = U_{i_1}(\sigma_{k+1}) \cup U_{i_2}(\sigma_{k+1})$ (that is, if $\{I_{i_1}(\sigma_{k+1}), I_{i_2}(\sigma_{k+1})\} = \{I_\ell(\sigma_{k+1}) : I_\ell(\sigma_{k+1}) \sim_{\xi(\sigma_k)} I_i(\sigma_k)\}$), then $F_i(\sigma_k) = F_{i_1}(\sigma_{k+1}) + F_{i_2}(\sigma_{k+1})$. It is therefore possible on the event Ω^* to define a finitely additive set function Γ on $\mathcal{B}$ such that

$$
\Gamma(U_i(t)) = F_i(t),\ t > 0,\ 1 \leq i \leq N(t), \tag{5.1}
$$

and

$$
\Gamma(\mathbb{S}) = 1. \tag{5.2}
$$

Furthermore, if $A_1 \supseteq A_2 \supseteq \ldots$ is a decreasing sequence of sets in the algebra $\mathcal{B}$ such that $\bigcap_n A_n = \emptyset$, then, by compactness, $A_n = \emptyset$ for all n sufficiently large and it is certainly the case that $\lim_{n\to\infty} \Gamma(A_n) = 0$. A standard extension theorem (see, for example, Theorems 3.1.1 and 3.1.4 of [Dud89]) gives that on the event Ω^* the set function Γ extends to a probability measure (also denoted by Γ) on the Borel σ–algebra of $\mathbb{S}$. Define Γ to be, say, the point mass δ_1 off the event Ω^*.

6 Capacities and lower bounds on dimensions

Establishing the capacity–equivalence (1.1) in the statement of Theorem 1.1 will certainly show that the capacity dimension of $\mathbb{S}$ is 1 a.s. The packing dimension of a set is at least its Hausdorff dimension, which is in turn equal to its capacity dimension. Therefore, (1.1) combined with the results of Section 3 will establish that the packing and Hausdorff dimensions of $\mathbb{S}$ are both 1 a.s.

From (4.1) and (2.2) we see that for some random variable C' (not depending on f) with $0 < C' < \infty$ a.s. we have

$$\begin{aligned}\mathrm{Cap}_f(\mathbb{S}) &\geq (\mathcal{E}_f(\Gamma))^{-1} = \left(\int \varphi_f(dt) \sum_{U \in \mathcal{U}(t)} \Gamma(U)^2\right)^{-1} \\ &= \left(\int \varphi_f(dt) \sum_{i=1}^{N(t)} F_i(t)^2\right)^{-1} \\ &\geq C' \left(\int \varphi_f(dt)(t \wedge 1)\right)^{-1} = C' \left(\int_0^1 f(t)\, dt\right)^{-1}.\end{aligned} \tag{6.1}$$

Note from the Cauchy-Schwarz inequality that for any $\mu \in M_1(\mathbb{S})$

$$1 = \left(\sum_{U \in \mathcal{U}(t)} \mu(U)\right)^2 \leq N(t) \sum_{U \in \mathcal{U}(t)} \mu(U)^2,$$

and so, by (4.1),

$$\mathrm{Cap}_f(\mathbb{S}) \leq \left(\int \varphi_f(dt) N(t)^{-1}\right)^{-1}.$$

This sort of bound appears in Section IV.2 of [Car67]. Applying (2.1), we see that for some random variable C'' (again not depending on f) with $0 < C'' < \infty$ a.s. we have

$$\mathrm{Cap}_f(\mathbb{S}) \leq C'' \left(\int \varphi_f(dt)(t \wedge 1)\right)^{-1} = C'' \left(\int_0^1 f(t)\, dt\right)^{-1} \tag{6.2}$$

The capacity–equivalence (1.1) follows from (6.1) and (6.2) and the fact that there exist constants $0 < c^\# \leq c^{\#\#} < \infty$ such that

$$c^\# \left(\int_0^1 f(t)\, dt\right)^{-1} \leq \mathrm{Cap}_f([0,1]) \leq c^{\#\#} \left(\int_0^1 f(t)\, dt\right)^{-1}$$

(this is described as "classical" in [PPS96] and follows by arguments similar to those used around equation (9) in Section 2 of that paper to prove the analogous inequalities for $[0,1]^2$).

Acknowledgement

Thanks are due to David Aldous and Jim Pitman for helpful discussions about Kingman's coalescent, and to an anonymous referee for several suggestions that improved the presentation.

References

[Ald93] D. Aldous. The continuum random tree III. *Ann. Probab.*, 21:248–289, 1993.

[Ald97] D.J. Aldous. Deterministic and stochastic models for coalescence (aggregation, coagulation): a review of the mean–field theory for probabilists. To appear in *Bernoulli*, available via http://www.stat.berkeley.edu/users/aldous, 1997.

[BP92] I. Benjamini and Y. Peres. Random walks on a tree and capacity in the interval. *Ann. Inst. H. Poincaré Probab. Statist.*, 28:557–592, 1992.

[BS97] E. Bolthausen and A.-S. Sznitman. On Ruelle's probability cascades and an abstract cavity method. Preprint, 1997.

[Car67] L. Carleson. *Selected Problems on Exceptional Sets.* van Nostrand, Princeton, 1967.

[DEF+98] P. Donnelly, S.N. Evans, K. Fleischmann, T.G. Kurtz, and X. Zhou. Continuum–sites stepping–stone models, coalescing exchangeable partitions, and random trees. Preprint, 1998.

[Dud89] R.M. Dudley. *Real Analysis and Probability.* Wadsworth, Belmont CA, 1989.

[EP98] S.N. Evans and J. Pitman. Construction of Markovian coalescents. *Ann. Inst. H. Poincaré Probab. Statist.*, 34:339–383, 1998.

[Fel71] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume II. Wiley, New York, 2nd edition, 1971.

[Haw81] J. Hawkes. Trees generated by a simple branching process. *J. London Math. Soc. (2)*, 24:374–384, 1981.

[Kin82a] J.F.C. Kingman. On the genealogy of large populations. In J. Gani and E.J. Hannan, editors, *Essays in Statistical Science*, pages 27–43. Applied Probability Trust, 1982. Special vol. 19A of *J. Appl. Probab.*

[Kin82b] J.F.C. Kingman. The coalescent. *Stochastic Process. Appl.*, 13:235–248, 1982.

[LP96] R.D. Lyons and Y. Peres. Probability on trees and networks. Book in preparation for Cambridge University Press, available via http://php.indiana.edu/˜rdlyons/, 1996.

[LS98] S.P. Lalley and T. Sellke. An extension of Hawke's theorem on the Hausdorff dimension of a Galton–Watson tree. Preprint, 1998.

[Lyo90] R.D. Lyons. Random walks and percolation on trees. *Ann. Probab.*, 18:931–958, 1990.

[Mat95] P. Mattila. *Geometry of Sets and Measures in Euclidean Spaces: Fractals and Rectifiability*, volume 44 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge – New York, 1995.

[MPV87] M. Mezard, G. Parisi, and M.A. Virasoro. *Spin Glass Theory and Beyond*, volume 9 of *World Scientific Lecture Notes in Physics*. World Scientific, Singapore, 1987.

[Per96] Y. Peres. Remarks on intersection–equivalence and capacity–equivalence. *Ann. Inst. H. Poincaré Phys. Théor.*, 64:339–347, 1996.

[Pit97] J. Pitman. Coalescents with multiple collisions. Preprint available via http://www.stat.berkeley.edu/users/pitman, 1997.

[PP95] R. Pemantle and Y. Peres. Galton–Watson trees with the same mean have the same polar sets. *Ann. Probab.*, 23:1102–1124, 1995.

[PPS96] R. Pemantle, Y. Peres, and J.W. Shapiro. The trace of spatial Brownian motion is capacity–equivalent to the unit square. *Probab. Theory Related Fields*, 106:379–399, 1996.

[RT61] C.A. Rogers and S.J. Taylor. Functions continuous and singular with respect to a Hausdorff measure. *Mathematika*, 8:1–31, 1961.

[Sch84] W. H. Schikhof. *Ultrametric Calculus: an Introduction to p-adic Analysis*, volume 4 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge – New York, 1984.

[Tav84] S. Tavaré. Line–of–descent and genealogical processes, and their applications in population genetics. *Theoret. Population Biol.*, 26:119–164, 1984.

[TT85] C. Tricot and S.J. Taylor. Packing measure, and its evaluation for a Brownian path. *Trans. Amer. Math. Soc.*, 288:679–699, 1985.

[Wat84] G. A. Watterson. Lines of descent and the coalescent. *Theoret. Population Biol.*, 26:77–92, 1984.