Abstract

In the case of neutral populations of fixed sizes in equilibrium whose genealogies are described by the Kingman \(N\)-coalescent back from time \(t\) consider the associated processes of total tree length as \(t\) increases. We show that the (cadlag) process to which the sequence of compensated tree length processes converges as \(N\) tends to infinity is a process of infinite quadratic variation; therefore this process cannot be a semimartingale. This answers a question posed in [PWW11].

MSC 2000 subject classifications. Primary 60K35; Secondary 92D25

Key words and phrases. Kingman coalescent, tree length process, quadratic variation, look-down graph.

1 Introduction and main result

The Kingman coalescent is a classical model in mathematical population genetics used for describing the genealogies for a wide class of population models (see e.g [Wa08]). The population models in question are neutral, exchangeable and with an offspring distribution of finite variation. One particular example is the Moran model ([Mo58]). This is a stationary continuous-time model for populations of fixed size \(N\) in which the reproduction takes place according to the following rule: starting with a population of size \(N\), after an exponential time of parameter \(\frac{1}{2}\) a pair of individuals is picked uniformly at random from the population, out of which one individual dies and the other one gives birth to one child.

The ancestry of a Moran population of size \(N\) started at time \(-\infty\) is at any time \(t \in \mathbb{R}\) described by the Kingman \(N\)-coalescent. This is a process with values in the set of partitions of \(\{1, \ldots, N\}\) which starts in the partition in singletons and has the following dynamics (backwards in time): given the process is in state \(\pi_k\), it jumps at rate \(\frac{k}{2}\) to a state \(\pi_{k-1}\) which is obtained by merging two randomly chosen elements of \(\pi_k\). The process can be represented graphically as a binary rooted tree which, when traced back from its \(N\) leaves (and correspondingly \(N\) external branches), exhibits a binary merger at rate \(\frac{k}{2}\) while there are \(k\) branches left.

One particular feature of coalescent trees that has been intensively investigated in the literature, due to its relevance in statistical studies of genetic data, is their total length (the sum of the lengths of all the branches of the tree). In the case of the Kingman coalescent tree started with \(N\) leaves

---

*Work partially supported by the DFG Priority Programme SPP 1590 “Probabilistic Structures in Evolution”.
†Institut für Mathematik, Goethe-Universität, 60054 Frankfurt am Main, Germany.
‡Fachrichtung Mathematik, Universität des Saarlandes, Postfach 151150, 66041 Saarbrücken, Germany.
the total length is in expectation equal to twice the harmonic number \( h_{N-1} = \sum_{i=1}^{N-1} \frac{1}{i} \), and when \( N \) tends to infinity (half of) the total length compensated by \( \log N \) converges in law to a Gumbel distributed random variable. In the case of coalescent processes with multiple mergers the total length has been studied in various papers, for instance [BBL12], [BBS07], [BBS08], [DIMR07], [LM07], [Ke12], [Mo10].

As time \( t \) increases the Moran population evolves and its genealogy changes, giving rise to a tree-valued process \( R^N = (R^N_t)_{t \in \mathbb{R}} \), the evolving Kingman \( N \)-coalescent. The associated process of total tree length was investigated in [PWW11]. Following their notation let \( \ell \) denote the mapping that sends a finite tree to its length and call

\[
\mathcal{L}^N = \left( \mathcal{L}^N_t \right)_{t \in \mathbb{R}} := (\ell(R^N_t) - 2 \log N)_{t \in \mathbb{R}}
\]

the compensated tree length process. Pfaffelhuber et al. [PWW11] investigated the asymptotic behaviour of the tree length process as the population size \( N \to \infty \) and showed that there exists a process \( \mathcal{L} = (\mathcal{L}_t)_{t \in \mathbb{R}} \) with sample paths in \( \mathbb{D} \), the space of càdlàg functions equipped with the Skorokhod topology, such that

\[
\mathcal{L}^N \to \mathcal{L} \text{ in law as } N \to \infty.
\]

The process \( \mathcal{L} \) is the Kingman tree length process.

![Tree Length](image)

**Figure 1:** A realisation of the compensated tree length process \( \mathcal{L}^N \) for \( N = 30 \) (courtesy of Peter Pfaffelhuber)

The weak convergence (2) can be lifted to convergence in probability, provided a representation for Moran populations on the same probability space for all population sizes \( N \in \mathbb{N} \) is considered. Such a representation is given by the look-down construction of Donnelly and Kurtz ([DK96], [DK99]) which encodes the evolving coalescent in a path-wise consistent way for increasing \( N \). If \( \mathcal{L}^{ld,N}_t \) denotes the compensated length of the tree at time \( t \) in the look-down representation, then, as shown in [PWW11], there exists a process \( \mathcal{L}^{ld} \), having the same distribution as \( \mathcal{L} \), such that

\[
d_{Sk}(\mathcal{L}^{ld,N}_t, \mathcal{L}^{ld}_t) \longrightarrow 0
\]

holds in probability as the population size \( N \to \infty \), where \( d_{Sk} \) denotes the Skorokhod metric.

The question that we address in this paper is one formulated in [PWW11], namely whether \( \mathcal{L} \) is a semimartingale. In [PWW11] it was proved that \( \mathcal{L} \) has “infinite infinitesimal variance”, more
precisely, \( \frac{1}{2 \log t} \mathbb{E}[(\mathcal{L}_t - \mathcal{L}_0)^2] \to 4 \) as \( t \to 0 \). However, the question whether \( \mathcal{L} \) still might have finite quadratic variation remained open in that paper. We will show that \( \mathcal{L} \) has a.s. infinite quadratic variation and hence cannot be a semimartingale. This will be achieved by investigating the jumps of the process \( \mathcal{L}^{ld} \).

Let us now give a brief description of the look-down construction and explain heuristically our approach. A formal description of the look-down graph will then be given in the next section. The main idea behind the look-down representation is to label the individuals in the population according to the persistence (or longevity) of their offspring: label 1 for the individual with the most persistent progeny, 2 for the second and so on.

We consider a system of countably many particles describing the sample genealogies ordered by persistence. At any time, each level 1, 2, ... is occupied by precisely one particle, and the system evolves as follows: for every pair \( i < k \) at rate 1 the particle currently at level \( i \) shoots an arrow towards level \( k \), independently of everything else. At this time the particle at level \( i \) gives birth to a new particle which is placed at level \( k \), while for each \( j \geq k \) the particle located on level \( j \) changes its level from \( j \) to \( j + 1 \). To each birth event we associate a line which records the time evolution of the levels occupied by the new-born particle (see the graphical representation in Figure 2). This line is pushed up to the next level each time a birth event happens on a level to the left of the current level of the line. We say that the line ends (dies) at the time it reaches level \( \infty \). The countable system of all the lines (including the immortal line that sits at level 1) makes up the look-down graph with infinitely many levels (or infinite look-down graph for short); the corresponding representation for a particle system of finite size \( N \) is obtained by projecting the infinite look-down graph onto the first \( N \) levels. When considering the system with \( N \) particles only, we say that a line dies when it is pushed out of level \( N \). Like in the case of the Moran model, the realisations of the sample genealogy can be read off from the look-down graph.

![Figure 2: Detail of one realisation of the infinite look-down graph.](image)

For a line \( G \) in the infinite look-down graph we denote by \( T^G \) its life-length, i.e. the time span between the birth and the death time of \( G \). If we restrict the graph to its first \( N \) levels, then \( T^{G,N} \), the life-length of the line up to \( N \) will denote the time span the line needs until it exits level \( N \).
In terms of trees, the life-length of a line that dies at some time \( t \) in the look-down graph with \( N \) levels corresponds to the length of the external branch that falls off the genealogical tree at time \( t \). Therefore, the jumps of the (compensated) tree length process \( \mathcal{L}^{ld,N} \) happen at the times lines exist level \( N \) in the \( N \)-look-down graph and they have sizes equal to the life-lengths up to \( N \) of these lines. Hence for \( s < t \) we can write

\[
\mathcal{L}_t^{ld,N} - \mathcal{L}_s^{ld,N} = N(t-s) - \sum_G T^{G,N},
\]

where the sum is taken over all lines \( G \) that exit level \( N \) in the time interval \( (s,t] \). It was proved in [PWW11] (see Proposition 3.1 and the proof of Proposition 6.1 therein) that for any fixed times \( s < t \)

\[
\mathcal{L}_t^{ld,N} - \mathcal{L}_s^{ld,N} \rightarrow \mathcal{L}_t^{ld} - \mathcal{L}_s^{ld}
\]

holds in \( L^2 \), and therefore almost surely along a subsequence \((N_k)_{k \in \mathbb{N}}\).

Let us now consider the lines in the infinite look-down graph that die in the time interval \( (s,t] \). For every such line there exists an \( N \) such that for all \( N' \geq N \) this line exits from level \( N' \) in the time interval \( (s,t] \). Conversely, for any line that exits at level \( \infty \) in the complement of the time interval \( (s,t] \) there exists an \( N \) such for all \( N' \geq N \) that this line does not exit from level \( N' \) in the time interval \( (s,t] \). Therefore, with probability one, it is the life-lengths up to \( N_k \) of precisely those lines that reach level \( \infty \) in \( (s,t] \), which appear as summands on the right-hand side of (3) for large enough \( k \), and thus contribute to the limit \( \mathcal{L}_t^{ld,N_k} - \mathcal{L}_s^{ld,N_k} \) as \( N_k \rightarrow \infty \).

Therefore, in order to understand the jumps of the limiting process \( \mathcal{L}^{ld} \) that occur in \( (s,t] \) one key issue is to understand the behaviour of the life-lengths of the lines that die in the infinite look-down graph in this time interval. The following theorem on the squared life-lengths of these lines is the central ingredient for proving our main result, which is stated in Theorem 2 below.

**Theorem 1.** For any \( s < t \) the sum of the squared life-lengths of the lines that die in the time interval \( (s,t] \) in the infinite look-down graph is almost surely infinite.

**Theorem 2.** The Kingman tree length process \( \mathcal{L} \) has a.s. infinite quadratic variation.

The quadratic variation process of a semimartingale is a process with càdlàg paths ([Pr04] Chapter 3, Theorem 22). This implies that for every time \( t \), the quadratic variation of a semimartingale at time \( t \) is finite. It hence follows from Theorem 2 that \( \mathcal{L} \) cannot be a semimartingale.

A key ingredient in the proof of Theorem 1 is the proposition stated below. This result is also of interest in its own right since it sheds light on the overall structure of the look-down graph and the large amount of independence which is built into it. From the brief description of the look-down graph given above (and from the formal definition provided in the next section) it is immediate that the birth times of lines on some level \( k \geq 2 \) in the look-down graph form a Poisson process with rate \( k-1 \). It turns out that the death times of these lines are also points of a Poisson process with the same rate. For the particular case \( k = 2 \) two different proofs of this result were given in [DK06] and [PW06].

**Proposition 1.** In the infinite look-down graph, for every \( k \in \mathbb{N}, k \geq 2 \) consider the process \( \eta_k \) of time points at which the lines that were born on level \( k \) reach level \( \infty \). The processes \( \eta_k \) are mutually independent Poisson with rate \( k-1 \).

For each \( k = 2, 3, \ldots \) the process \( \mathcal{L}^{ld} \) has a jump in each of the points of \( \eta_k \). The size of this jump is equal to the life-length \( T^G \) of the line \( G \) that dies at this time point (see the proof of Theorem 2).
The expectation of $T^G$ is $2/k$ (see (4) below) and for large $k$ the distribution of $T^G$ is concentrated around $2/k$ (see the proof of Theorem 1, which uses a result of [DPS13]). Since the points of $\eta_k$ come at rate $k - 1$, the Lévy measure of a Lévy process which would mimick the jump behaviour of $\mathcal{L}^{ld}$ should (for large $k$) have mass $k - 1$ concentrated around $2/k$. Such a measure would violate the condition $\int_0^1 h^2 \nu(dh) < \infty$ and would therefore not be appropriate as a Lévy measure. As a matter of fact, $\mathcal{L}^{ld}$ is not a Lévy process since its jump sizes are dependent (while its jump times are independent). On the other hand, part of the strategy of the proof of our main result reflects in the simple fact that $\sum (k - 1)(2^k)^2 = \infty$.

2 The look-down process

The look-down construction of Donnelly and Kurtz ([DK96], [DK99]) is an alternative way of representing the evolution of Moran (and more general exchangeable) populations, which proves to be a very powerful instrument in investigating population dynamics. As already mentioned in the introduction, this representation of populations of sizes $N$ is done on one and the same probability space for all $N \in \mathbb{N}$ in such a way that the path-wise consistency of the genealogies is ensured as $N \to \infty$.

The main idea of the look-down representation is the labelling of the individuals according to the persistence of their offspring in the population. In the first paper [DK96] the persistence of the offspring is taken to hold in probability, whereas in the "modified" look-down construction introduced in [DK99], this holds almost surely. We will use this second version of the model which we describe below following [PW06].

We consider a population of infinite size and denote by $\mathcal{V}$ the set $\mathbb{R} \times \mathbb{N}$. An element $(s, i)$ in $\mathcal{V}$ denotes the individual that occupies level $i$ at time $s$. The levels represent indices given to the individuals in the population according to the persistence of their offspring in the following way: the offspring of the individual that lives at time $s$ at level $i$ almost surely outlives the offspring of any other individual alive at time $s$ on a level $k > i$. The process evolves as follows: to every pair of levels $i, k \in \mathbb{N}$ with $i < k$, we attach a (rate one) Poisson point process on $\mathbb{R}$ which we denote by $C_{ik}$. All these Poisson point processes are independent. Each time the clock $C_{ik}$ rings, "level $k$ looks down to level $i$", that is, the current individual at level $i$ reproduces and its offspring is placed at level $k$. For $k \geq 2$ and $s_0 \in \bigcup_{i<k} C_{ik}$ we associate with the individual born at time $s_0$ at level $k$ the set of points

$$G = \bigcup_{j \in \mathbb{N}_0} [s_j, s_{j+1}] \times \{k + j\},$$

where $s_j$, for $j = 1, 2, \ldots$ is $s_j := \inf \left\{ s > s_{j-1} : s \in \bigcup_{l<m<k+j} C_{lm} \right\}$. We call $G$ the line born at time $s_0$ on level $j$ and say that at time $s_j$ the line is pushed from level $k + j - 1$ to level $k + j$. Note that a line is pushed one level upwards every time one of the Poisson point processes associated with levels smaller than or equal to the current level of the line experiences an event. Lines are born on a level $k > 1$ at the times of a Poisson point process with rate $(k - 1)$ and a line at level $k$ is pushed up with rate $(k^2)$ because there are $(k^2)$ independent (rate one) Poisson point processes which trigger the look-down between the levels that are smaller than or equal to $k$.

We say that a line dies when it reaches level infinity and denote the death time of line $G$ by

$$d^G := \lim_{j \to \infty} s_j.$$
Since the rate at which a line born at a level bigger than or equal to 2 is pushed up is quadratical, we conclude that the time it takes for a line to die is finite almost surely. Level 1 is never hit by arrows and therefore the offspring of the individuals living on this level persist forever in the population. We call the line $\mathbb{R} \times \{1\}$ the immortal line.

The set of all the lines is countable and it forms a partition of $\mathcal{V}$. The random graph obtained in this way is called the \textit{look-down graph} (with infinitely many levels). This graph records the evolution of a population of infinite size. Embedded in the look-down process are all the $N$-particle look-down processes corresponding to populations of sizes $N \in \mathbb{N}$. The $N$-particle look-down process is constructed in a similar way, but the graph has only $N$ levels and we say that a line dies when it exits level $N$. Any $N$-particle look-down process can be recovered as the projection of the infinite look-down process on the first $N$ levels.

The ordering by persistence, which is an essential feature of the look-down process, makes it no longer exchangeable with respect to the indices of the individuals, as it happens in the Moran model. Nevertheless, the ancestral process of the population has the same distribution in the two models (DK99). In order to see that the ancestral process induced by the look-down process is the Kingman coalescent, consider two lines $G$ and $G'$. For $(s, l) \in G$ and $(t, i) \in G'$ with $s \leq t$ we say that $(s, l)$ is the \textit{ancestor} of $(t, i)$ and we write

$$A_s(t, i) = l,$$

if either the two lines are the same, or there are some lines $G_1, \ldots, G_m$ such that $G_1$ descends from $G$, $G_k$ descends from $G_{k-1}$, for $k = 2, \ldots, m$ and $G'$ descends from $G_m$. Two individuals $(t, i)$ and $(t, j)$ living at time $t$ have the same ancestor at time $s$ if $A_s(t, i) = A_s(t, j)$ and we write $i \sim j$ with $u = t - s$. The random equivalence relation $\sim$ defines the ancestral process of the population alive at time $t$, $\mathcal{R}_{t}^{ld} = (\mathcal{R}_{t}^{ld}(u))_{u \in \mathbb{R}}$. It is not difficult to check that for each $t$ the restriction of $\mathcal{R}_{t}^{ld}$ to $\{1, \ldots, N\}$ has the distribution of $\mathcal{R}^N$, the $N$-Kingman coalescent. The consistency property then implies that the genealogy $\mathcal{R}_{t}^{ld}$ of the infinite population has the distribution of the Kingman coalescent $\mathcal{R}$.

In the look-down graph with infinitely many levels let us consider a line $G$ born on level $l^G \geq 2$. The time this line needs in order to reach level infinity is

$$T^G = \sum_{j=l^G}^{\infty} X^G_j,$$

where the time $X^G_j$ spent by the line on level $j$ is an exponentially distributed random variable with parameter $(j^2)$ and the $X_j$’s are independent from one another for different $j$’s. We call $T^G$ the \textit{life-length} of the line $G$. In terms of trees, the life-length of a line that dies at some time $t$ represents the length of the external branch that falls off the genealogical tree at time $t$. When restricting to the first $N$ levels in the graph, we define

$$T^{G,N} := \sum_{j=l^G}^{N} X^G_j,$$

(5)

to be the \textit{life-length up to level $N$} of the line $G$. 

6
3 Proof of Theorem 1

We start by proving Proposition 1.

Proof of Proposition 1

For every \( n \in \mathbb{N}, n \geq 2 \) and \( 2 \leq k \leq n \) let us write \( \eta^n_k \) for the process of arrival times at level \( n \) of lines born on level \( k \). For \( k = n \), the process \( \eta^n_n \) equals the process of time points were new lines are born on level \( n \). Since (independently of everything else) new lines on level \( n \) are born from level \( m \) at rate 1, \( m = 1, \ldots, n - 1 \), it is clear that for every \( n \geq 2 \), \( \eta^n_n \) is a Poisson process with rate \( n - 1 \) that is independent of \( (\eta^n_2, \ldots, \eta^n_{n-1}) \).

It is thus sufficient to prove the following claim:

(*) for every \( n \in \mathbb{N}, n \geq 2 \), the processes \( \eta^n_k, k = 2, \ldots, n - 1 \), are Poisson processes of rate \( k - 1 \) and they are independent from one another for \( n \) fixed and different values of \( k \).

Assuming this claim holds, remember that for a line \( G = \bigcup_{j \in \mathbb{N}_0} [s^G_j, s^G_{j+1}) \times \{k + j\} \) born on level \( k \), the time point \( s^G_j \) is the time the line reaches level \( k + j \) and that its death time

\[
\begin{align*}
a^G & = \lim_{j \to \infty} s^G_j
\end{align*}
\]

is finite almost surely. Now, denoting by \( G_k \) the set of all the lines in the look-down graph which are born on level \( k \), it follows that the time points \( \{s^G_j\}_{G \in G_k} \) are the points of the process \( \eta^{k+j}_k \), whereas the points \( \{d^G\}_{G \in G_k} \) are the points of the process \( \eta_k \). Thus, the assertion of the Proposition follows from the claim.

We now proof the claim (*) by an induction argument.

For the basic step of the induction let \( n = 3 \). On level 2, lines are born at the times of the Poisson process \( \eta^2_2 \) and every time a line is born, the line that occupied the level 2 is pushed up to level 3. Therefore, a line born on level 2 arrives at level 3 at the next point of \( \eta^2_2 \) after the line’s birth time. It follows that the set of points of \( \eta^2_3 \) is equal to the set of points of \( \eta^2_2 \) and hence \( \eta^3_2 \) is a Poisson process with rate 1.

We assume now that the claim holds for \( n \) and prove it for \( n + 1 \). From the induction assumption and the last sentence in the first paragraph of this proof it follows that the processes \( \eta^n_k, k = 2, \ldots, n \), are independent Poisson processes of rate \( k - 1 \). A fortiori, the process \( \eta^n_{n+1} \) of arrival times at level \( n \), obtained by superposing the independent processes \( \eta^n_2, \ldots, \eta^n_n \) is Poisson with rate \( \binom{n}{2} \). A line currently on level \( n \) is pushed to level \( n + 1 \) at the next point of \( \eta^n_{n+1} \) after the line’s arrival on level \( n \). Therefore, there is a bijective function \( \phi \) from the collection of points of the process \( \eta^n_{n+1} \) into itself which maps the time a line arrives at level \( n \) onto the time it is pushed up (and arrives at level \( n + 1 \)) (see Figure 3).
To each point of \( \eta_{\cdot\cdot,n} \) we associate a label which records the level at which the line arriving at this point was born. By the induction assumption these labels are independent and take value \( k \) with probability \((k-1)/{n}\). The birth level of a line arriving at time \( t \) at level \( n+1 \) is the birth level of the line arriving at time \( \phi^{-1}(t) \) at level \( n \), and hence it is the label of the point \( \phi^{-1}(t) \). The induction step is now completed by the following elementary observation: Consider an independent \( M \), then the new labelling has the same distribution as the old one. \( \square \)

**Proof of Theorem 1**

Let \( s < t \in \mathbb{R} \) be fixed. For every \( k \in \mathbb{N} \) we consider the sequence of lines born on level \( k \) that die after time \( s \), indexed by their death times. Let \( (T_{ik})_{i \geq 1} \) be the sequence of their life-lengths and let \( M_k \) denote the number of these lines which die by time \( t \). According to Proposition 1 the numbers \( M_k \) are Poisson distributed with parameter \((k-1)(t-s)\) and independent from one another for different \( k \)'s. We show that

\[
S_{s,t} := \sum_{k=2}^{\infty} \sum_{i=1}^{M_k} T_{ik}^2
\]

is infinite almost surely.

For each \( 1 \leq i \leq M_k \) the following law of large numbers holds (see [Al99] and also [PWW11], Lemma 4.6 and [DPS13]):

\[
kT_{ik} \xrightarrow{k \to \infty} 2 \quad \text{almost surely.}
\]

For each \( k \geq 2 \) it holds that

\[
P\left(M_k \notin \left[\frac{1}{2}(k-1)(t-s), 2(k-1)(t-s)\right]\right) \quad \text{(6)}
\]

\[
\text{or } \{ M_k \in \left[\frac{1}{2}(k-1)(t-s), 2(k-1)(t-s)\right] \text{ and } T_{ik} \notin \left[\frac{1}{k'}, \frac{3}{k'}\right] \text{ for some } i = 1, \ldots, M_k \}
\]

\[
\leq P\left(M_k \notin \left[\frac{1}{2}(k-1)(t-s), 2(k-1)(t-s)\right]\right) + \left[2(k-1)(t-s)\right]P\left(T_{ik} \notin \left[\frac{1}{k'}, \frac{3}{k'}\right]\right).
\]

Cramér’s theorem guarantees that \( P\left(M_k \notin \left[\frac{1}{2}(k-1)(t-s), 2(k-1)(t-s)\right]\right) \) decays exponentially in \( k \) and hence the first term on the right-hand side is summable. For the second term we use Theorem 1 of [DPS13] which says that the distribution of \( kT_{ik} \) satisfies a large deviation principle with scale \( k \) and a good rate function. Since

\[
P\left(T_{ik} \notin \left[\frac{1}{k'}, \frac{3}{k'}\right]\right) = P\left(|T_{ik} - \frac{2}{k}| > \frac{1}{k}\right)
\]

it follows that the second term on the right-hand side of (6) is also summable. By the Borel-Cantelli lemma we obtain that there exists an \( \mathbb{N} \)-valued random variable \( K_1 \geq 2 \) such that for all \( k \geq K_1 \)

\[
M_k \in \left[\frac{1}{2}(k-1)(t-s), 2(k-1)(t-s)\right] \quad \text{and} \quad T_{ik} \in \left[\frac{1}{k'}, \frac{3}{k'}\right] \quad \text{for all } i = 1, \ldots, M_k \quad \text{almost surely}
\]

and in particular

\[
M_k \geq \frac{1}{2}(k-1)(t-s) \quad \text{and} \quad T_{ik} \geq \frac{1}{k} \quad \text{for all } i = 1, \ldots, M_k \quad \text{almost surely.}
\]
Therefore, it holds that almost surely
\[
\sum_{k=K_1}^{\infty} \sum_{i=1}^{M_k} T_{ik}^2 \geq \sum_{k=K_1}^{\infty} \left[ \frac{1}{2} (k-1)(t-s) \right] \cdot \frac{1}{k^2}.
\]

Now since \( K_1 \) is almost surely finite, it follows that the sum on the right-hand side is infinite almost surely and that
\[
S_{s,t} = \infty \quad \text{almost surely,}
\]
which gives the claim. \( \Box \)

4 Proof of Theorem 2

In order to prove Theorem 2 we make use of the result in Proposition 3.2 of [PWW11]. Since \( d_{sk}(\mathcal{L}^{ld,N}, \mathcal{L}^{ld}) \to 0 \) as \( N \to \infty \) in probability, there is an increasing sequence \((N_k)_{k \in \mathbb{N}}\) in \( \mathbb{N} \) diverging to \( \infty \) such that \( d_{sk}(\mathcal{L}^{ld,N_k}, \mathcal{L}^{ld}) \to 0 \) almost surely as \( k \to \infty \).

The following lemma and proposition are a key ingredients of the proof of Theorem 2.

Lemma 1. Let \((y_k)_{k \geq 1}\), \(y_k: \mathbb{R} \to \mathbb{R}\) be a sequence of càdlàg functions satisfying that there exist two sequences \((\tau_k)_{k \geq 1}\) and \((\gamma_k)_{k \geq 1}\) in \( \mathbb{R} \) such that \(y_k\) has a jump of size \(\gamma_k\) at time \(\tau_k\) for all \(k \geq 1\). Moreover, suppose that the sequence \((y_k)_{k \geq 1}\) converges in the Skorohod topology to a càdlàg function \(y\) and that the sequences \((\tau_k)_{k \geq 1}\) and \((\gamma_k)_{k \geq 1}\) are convergent. Let \(\tau := \lim_{k \to \infty} \tau_k\) and \(\gamma := \lim_{k \to \infty} \gamma_k\) and assume that \(\gamma > 0\). Then the function \(y\) has a jump of size \(\gamma\) at time \(\tau\).

Proof. Let \(\Lambda\) be the set of all strictly increasing and continuous functions \(\lambda: [0, \infty) \to [0, \infty]\). The sequence \((y_k)_{k \geq 1}\) converges in the Skorohod topology to \(y\), which means that
\[
\inf \left\{ \epsilon > 0 \mid \exists \lambda \in \Lambda : \sup_{t \in \mathbb{R}} |t - \lambda(t)| \leq \epsilon, \sup_{t \in \mathbb{R}} |y_k(t) - y(\lambda(t))| \leq \epsilon \right\} \to 0
\]
as \(k \to \infty\).

Let \((\lambda_k)_{k \geq 1}\) be a sequence of functions in \(\Lambda\) such that
\[
\sup_{t} |t - \lambda_k(t)| \to 0 \quad \text{and} \quad \sup_{t} |y_k(t) - y(\lambda_k(t))| \to 0
\]
as \(k \to \infty\). Let us now look at the jump sizes \(\delta_k\) defined by
\[
\delta_k = y \circ \lambda_k(\tau_k) - y \circ \lambda_k(\tau_k-).
\]

Since the size of the jump of the function \(y_k\) at the time \(\tau_k\) is \(\gamma_k = y_k(\tau_k) - y_k(\tau_k-)\), we can infer that
\[
|\gamma_k - \delta_k| = |y_k(\tau_k) - y_k(\tau_k-) - y \circ \lambda_k(\tau_k) + y \circ \lambda_k(\tau_k-)|
\]
\[
= |y_k(\tau_k) - y \circ \lambda_k(\tau_k) + y \circ \lambda_k(\tau_k-) - y_k(\tau_k-)|
\]
\[
\leq |y_k(\tau_k) - y \circ \lambda_k(\tau_k)| + |y \circ \lambda_k(\tau_k-) - y_k(\tau_k-)|. \tag{8}
\]

Observe that the right-hand side of (8) converges to 0 as \(k \to \infty\) and therefore
\[
|\gamma_k - \delta_k| \to 0 \quad \text{as} \quad k \to \infty.
\]
Since \( \gamma_k \to \gamma \) as \( k \to \infty \), it follows that 
\[
\delta_k \to \gamma \quad \text{as} \quad k \to \infty. 
\]

Hence, the function \( y \) has jumps at the time points \( \lambda_k(\tau_k) \) and the sizes \( \delta_k \) of these jumps converge to \( \gamma \).

From (7) we know that \(|\tau_k - \lambda_k(\tau_k)| \to 0 \) as \( k \to \infty \) and because \( \tau_k \to \tau \) as \( k \to \infty \), we conclude that 
\[
\lambda_k(\tau_k) \to \tau \quad \text{as} \quad k \to \infty. 
\] (9)

We now claim that there exists a \( k_0 \) such that \( \lambda_k(\tau_k) \) is equal to \( \tau \) for all \( k \geq k_0 \). Indeed, this holds true because otherwise it would mean that there is an infinite number of different values in the sequence \((\lambda_k(\tau_k))_{k \geq 1}\) at which the function \( y \) has jumps of sizes bigger than a threshold, say \( \frac{\gamma}{2} \). But for a càdlàg function there can be at most finitely many points in any finite time interval at which the jump size exceeds a given positive number (see [Bi99 Chapter 3, Section 12, Lemma 1]). Thus, by (9), there must exist a \( k_0 \) such that \( \lambda_k(\tau_k) \) is equal to \( \tau \) for all \( k \geq k_0 \). For \( k \geq k_0 \) we have that \( \delta_k \) is equal to \( y(\tau) - y(\tau-) \) which is precisely the size of the jump of \( y \) at time \( \tau \). Since we already know that the sequence \( \delta_k \) converges to \( \gamma \), this implies that the jump of \( y \) at time \( \tau \) is equal to \( \gamma \), which proves the lemma. \( \square \)

**Proposition 2.** The sum of the squared jump sizes of the process \( \mathcal{L}^{ld} \) occurring in any interval of positive length is infinite almost surely.

**Proof.** Consider the look-down graph and recall that for every \( N \in \mathbb{N} \) the \( N \)-look-down graph can be recovered as the projection of the infinite graph onto its first \( N \) levels.

Let \( \mathcal{G} \) denote the set of all the lines in the infinite look-down graph and for a line \( G = \bigcup_{j \in \mathbb{N}_0} [s^G_j, s^G_{j+1}) \times \{l^G + j\} \) born on level \( l^G \) let us set 
\[
d^{G,N} := \begin{cases} 
s^G_{N-l^G+1} & \text{if} \ N \geq l^G \\
-\infty & \text{otherwise}, \end{cases} 
\]
the exit time from level \( N \) of the line \( G \).

We are interested in the times and the sizes of the jumps of the processes \( \mathcal{L}^{ld,N} \). Jumps occur at the times \( \{d^{G,N}\}_{G \in \mathcal{G}} \) when lines die in the \( N \)-look-down process (i.e. they exit level \( N \)). Since for a fixed \( G \in \mathcal{G} \) the sequences \( \{d^{G,N}\}_{N \in \mathbb{N}} \) and \( \{s^G_{N-l^G+1}\}_{N \in \mathbb{N}} \) are identical for \( N \) large enough, it follows that 
\[
\lim_{N \to \infty} d^{G,N} = d^G, 
\] (10)

where \( d^G \) is the death time of line \( G \). The jump size of the process \( \mathcal{L}^{ld,N} \) at time \( d^{G,N} \) has size equal to the life-length \( T^{G,N} \) of the line \( G \) up to level \( N \) defined in (5). Note that the exponential times \( X^G_j \) do not depend on \( N \). Therefore, we have that 
\[
\lim_{N \to \infty} T^{G,N} = \lim_{N \to \infty} \sum_{j=1}^{N} X^G_j = \sum_{j=1}^{\infty} X^G_j = T^G 
\] (11)

almost surely, where \( T^G \) defined in (4) is the life length of line \( G \) in the infinite look-down graph.

In view of (10) and (11) we now apply for every \( G \in \mathcal{G} \) Lemma 1 to the paths of \( (\mathcal{L}^{ld,N_k})_{k \geq 1} \), the sequence of times \( (d^{G,N_k})_{k \geq 1} \) and the sequence of jump sizes \( (T^{G,N_k})_{k \geq 1} \). Consequently, for each \( G \in \mathcal{G} \), the limiting process \( \mathcal{L}^{ld} \) has a jump of size \( T^G \) at time \( d^G \).
Thus, for the sum of the squared jump sizes of \( L_{ld} \) occurring in an interval \([0, t]\), \( t > 0\),

\[
\sum_{0 \leq s \leq t} (\Delta L_{ld}^s)^2 \geq \sum_{G \in \mathcal{G}} \left( \sum_{d \in [0,t]} (T^G)^2 \right)
\]

holds and since, according to Theorem 1, the right-hand side is infinite almost surely, the Proposition is proved. □

**Proof of Theorem 2**

It remains to note that for any \( t > 0 \), any càdlàg path \( X \) which obeys

\[
\sum_{0 \leq s \leq t} (\Delta X_s)^2 = \infty
\]

has the property

\[
\lim_{n \to \infty} \sum_{j=1}^{l(n)} \left( X_{(n)}^{(j)} - X_{p_{j+1}}^{(n)} \right)^2 = \infty
\]

for any sequence of partitions \((\mathcal{P}_n)_{n \in \mathbb{N}} = \left( (\rho_j^{(n)})_{j=1,...,l(n)} \right)_{n \in \mathbb{N}}\) of \([0, t]\) with mesh size tending to zero for \( n \to \infty \). To this purpose we order the jump sizes of \( X \) that occur in \((0, t)\) according to their sizes and denote by \((t_i)_{i \geq 1}\) the corresponding jump times, i.e. \(|\Delta X_{t_1}| \geq |\Delta X_{t_2}| \geq \ldots\) holds. Then, for every (fixed but arbitrary) \( k \in \mathbb{R} \) there exists an \( m(k) \) such that

\[
\sum_{i=1}^{m(k)} (\Delta X_{t_i})^2 \geq k.
\]

For every jump time \( t_i \) and every partition \( \mathcal{P}_n \) let \( \sigma_{i,n} \) be the largest point in the partition smaller than \( t_i \) and \( \tau_{i,n} \) be the smallest point in the partition larger than or equal to \( t_i \). Then, for \( n \) large enough, there is at most one of the \( t_1, \ldots, t_{m(k)} \) between any two points of the partition \( \mathcal{P}_n \) and thus

\[
\sum_{j=1}^{l(n)} \left( X_{(n)}^{(j)} - X_{p_{j+1}}^{(n)} \right)^2 \geq \sum_{i=1}^{m(k)} \left( X_{\tau_{i,n}} - X_{\sigma_{i,n}} \right)^2
\]

holds for \( n \) large enough. Using the càdlàg property of \( X \) we obtain that

\[
\lim_{n \to \infty} \sum_{i=1}^{m(k)} \left( X_{\tau_{i,n}} - X_{\sigma_{i,n}} \right)^2 \geq \sum_{i=1}^{m(k)} (\Delta X_{t_i})^2
\]

Since \( k \) was arbitrary, (13) follows from the last three inequalities.

From this together with Proposition 2 and Lemma 1, Theorem 2 is immediate. □

**Acknowledgement.** We are grateful to Stephan Gufler and Etienne Pardoux for stimulating discussions.

**References**

[Al99] D.J. Aldous (1999). Deterministic and stochastic models for coalescence (aggregation, coagulation): a review of the mean-field theory for probabilists. *Bernoulli*, 5 3-48.
[BBL12] J. Berestycki, N. Berestycki and V. Limic (2012). Asymptotic sampling formulae for Lambda-coalescents. arXiv:1201.6512.

[BBS07] J. Berestycki, N. Berestycki and J. Schweinsberg (2007). Beta-coalescents and continuous stable random trees. *Ann. Probab.* 35 1835-1887.

[BBS08] J. Berestycki, N. Berestycki and J. Schweinsberg (2008). Small time properties of Beta-coalescents. *Ann. Inst. H. Poincaré* 44 214-238.

[Bi99] P. Billingsley (1999) *Convergence of Probability Measures*, John Wiley and Sons.

[DDSJ09] J-F. Delmas, J.S. Dhersin and A. Siri-Jegousse (2010) On the two oldest families for the Wright-Fisher process, *Electron. J. Probab.* 15 776-800.

[DPS13] A. Depperschmidt, P. Pfaffelhuber and A. Scheuringer (2013) Some large deviations in Kingman’s coalescent, arXiv: 1311.0649

[DK96] P. Donnelly and T.G. Kurtz, (1996). A countable representation for the Fleming-Viot measure-valued diffusion, *Ann. Probab.*, 24(2) 698-742.

[DK99] P. Donnelly and T.G. Kurtz (1999) Particle representations for measure-valued population models, *Ann. Probab.*, 27(1) 166-205.

[DK06] P. Donnelly and T.G. Kurtz (2006) The Eve process, Manuscript, personal communication.

[DIMR07] M. Drmota, A. Iksanov, M. Möhle and U. Rösler (2007). Asymptotic results about the total branch length of the Bolthausen-Sznitman coalescent. *Stoch. Proc. Appl.* 117 1404-1421.

[IM07] A. Iksanov and M. Möhle (2007). A probabilistic proof of a weak limit law for the number of cuts needed to isolate the root of a random recursive tree. *Electron. Comm. Probab.* 12 28-35.

[Ke12] G. Kersting (2012). The asymptotic distribution of the length of Beta-coalescent trees. *Ann. Appl. Probab.* 22 2086-2107.

[Mo10] M. Möhle (2010). Asymptotic results for coalescent processes without proper frequencies and applications to the two-parameter Poisson-Dirichlet coalescent. *Stoch. Process. Appl.* 120 2159-2173.

[Mo58] P.A.P. Moran (1958). Random processes in genetics. *Proc. Cambridge Phil. Soc.* 54 60-71.

[PW06] P. Pfaffelhuber and A. Wakolbinger, (2006). The process of most recent common ancestors in an evolving coalescent, *Stoch. Proc. Appl.*, 116 1836-1859.

[PWW11] P. Pfaffelhuber, A. Wakolbinger and H. Weisshaupt (2011) The tree length of an evolving coalescent, *Prob. Theo. Rel. Fields*, 151 529-557.

[Pr04] P. Protter (2004) *Stochastic Integration and Differential Equations*, Springer.

[Wa08] J. Wakeley (2008). *Coalescent Theory: An Introduction* Roberts & Company.