The Annals of Applied Probability
2012, Vol. 22, No. 5, 2086–2107
DOI: 10.1214/11-AAP827

THE ASYMPTOTIC DISTRIBUTION OF THE LENGTH OF BETA-COALESCENT TREES

BY GÖTZ KERSTING

Goethe Universität

We derive the asymptotic distribution of the total length L_n of a Beta$(2-\alpha, \alpha)$-coalescent tree for $1 < \alpha < 2$, starting from n individuals. There are two regimes: If $\alpha \le \frac{1}{2}(1+\sqrt{5})$, then L_n suitably rescaled has a stable limit distribution of index α. Otherwise L_n just has to be shifted by a constant (depending on n) to get convergence to a nondegenerate limit distribution. As a consequence, we obtain the limit distribution of the number S_n of segregation sites. These are points (mutations), which are placed on the tree's branches according to a Poisson point process with constant rate.

1. Introduction and result. In this paper we investigate the asymptotic distribution of the suitably normalized length L_n of a n-coalescent of the Beta$(2-\alpha, \alpha)$-type with $1 < \alpha < 2$. As a corollary we obtain the asymptotic distribution of the associated number S_n of segregating sites, which is the basis of the Watterson estimator [19] for the rate θ of mutation of the DNA. Here we recall that coalescents with multiple merging such as Beta-coalescents have been considered in the literature as a model for the genealogical relationship within certain maritime species [7, 10].

Beta-coalescents (and more generally Λ-coalescents, as introduced by Pitman [16] and Sagitov [17]) possess a rich underlying partition structure, which is nicely presented in detail in Berestycki [3]. For our purposes it is not necessary to recall all these details, and we refer to the following condensed description of a n-coalescent:

Imagine n particles (blocks in a partition), which coalesce into a single particle within a random number of steps. This happens in the manner of a continuous time Markov chain. Namely, if there are currently $m > 1$ particles, then they merge to l particles at a rate $\rho_{m,l}$ with $1 \le l \le m-1$. Thus

$$\rho_m = \rho_{m,1} + \cdots + \rho_{m,m-1}$$

is the total merging rate, and

$$P_{m,l} = \frac{\rho_{m,l}}{\rho_m}, \qquad 1 \le l \le m-1$$

gives the probability of a jump from m to l.

Received August 2011; revised October 2011.
MSC2010 subject classifications. Primary 60F05; secondary 60K35, 60G50, 60G55.
Key words and phrases. Beta-coalescent, tree, coupling, point process, stable distribution.

In these models the rates $\rho_{m,l}$ have a specific consistency structure arising from the merging mechanism. As follows from Pitman [16], they are, in general, of the form

$$\rho_{m,m-k+1} = \binom{m}{k} \int_0^1 t^{k-2}(1-t)^{m-k} \Lambda(dt), \qquad 2 \le k \le m,$$

where $\Lambda(dt)$ is a finite measure on $[0, 1]$. The choice $\Lambda = \delta_0$ corresponds to the original model due to Kingman [13], then $\rho_{m,l} = 0$ for $l \neq m-1$. In this paper we assume

$$\Lambda(dt) = \frac{1}{\Gamma(2-\alpha)\Gamma(\alpha)} t^{1-\alpha}(1-t)^{\alpha-1}\, dt,$$

thus

$$\rho_{m,m-k+1} = \frac{1}{\Gamma(2-\alpha)\Gamma(\alpha)} \binom{m}{k} B(k-\alpha, m-k+\alpha),$$

where $B(a, b)$ denotes the ordinary Beta-function. Then the underlying coalescent is called the Beta$(2-\alpha, \alpha)$-coalescent. For $\alpha = 1$ it is the Bolthausen–Sznitman coalescent [6] and the case $\alpha \to 2$ can be linked with Kingman's coalescent.

The situation can be described as follows: There are the merging times $0 = T_0 < T_1 < \cdots < T_{\tau_n}$ and there is the embedded time discrete Markov chain $n = X_0 > X_1 > \cdots > X_{\tau_n} = 1$, where X_i is the number of particles (partition blocks) after i merging events, and τ_n is the number of all merging events. This Markov chain has transition probabilities $P_{m,l}$ and, given the event $X_i = m$ with $m > 1$, the waiting time $T_{i+1} - T_i$ to the next jump is exponential with expectation $1/\rho_m$. Since a point process description is convenient later, and we name the point process

$$\mu_n = \sum_{i=0}^{\tau_n - 1} \delta_{X_i} \tag{1}$$

on $\{2, 3, \ldots\}$ the *coalescent's point process* downwards from n, abbreviated CPP(n).

These dynamics can be visualized by a coalescent tree with a root and n leaves. The leaves are located at height $T_0 = 0$ and the root at height T_{τ_n} above. At height T_i there are X_i nodes representing the particles after i coalescing events. The total branch length of this tree is given by

$$L_n = \sum_{i=0}^{\tau_n - 1} X_i (T_{i+1} - T_i). \tag{2}$$

For $1 < \alpha < 2$, the asymptotic magnitude of L_n is obtained by Berestycki et al. in [2]; it is proportional to $n^{2-\alpha}$. The asymptotic distribution of L_n is easily derived for Kingman's coalescent (see [8]); it is Gumbel. The case of a Bolthausen–Sznitman coalescent is treated by Drmota et al. [9], here L_n properly normalized is asymptotically stable. The case $0 < \alpha < 1$ of a Beta-coalescent is contained in more general results of Möhle [14]. Partial results for the Beta-coalescent with $1 < \alpha < 2$ have been obtained by Delmas et al. [8].

In this paper we derive the asymptotic distribution of the length of the Beta-coalescent for $1 < \alpha < 2$. Let ς denote a real-valued stable random variable with index α, which is normalized by the properties

$$\mathbf{E}(\varsigma) = 0, \qquad \mathbf{P}(\varsigma > x) = o(x^{-\alpha}), \qquad \mathbf{P}(\varsigma < -x) \sim x^{-\alpha} \tag{3}$$

for $x \to \infty$. Thus it is maximally skewed among the stable distributions of index α.

Also let

$$c_1 = \frac{\Gamma(\alpha)\alpha(\alpha-1)}{2-\alpha}, \qquad c_2 = \frac{\Gamma(\alpha)\alpha(\alpha-1)^{1+1/\alpha}}{\Gamma(2-\alpha)^{1/\alpha}}.$$

THEOREM 1. *For the Beta-coalescent with* $1 < \alpha < 2$:

(i) *If* $1 < \alpha < \frac{1}{2}(1+\sqrt{5})$ (*thus* $1 + \alpha - \alpha^2 > 0$), *then*

$$\frac{L_n - c_1 n^{2-\alpha}}{n^{1/\alpha+1-\alpha}} \xrightarrow{d} \frac{c_2 \varsigma}{(1+\alpha-\alpha^2)^{1/\alpha}}.$$

(ii) *If* $\alpha = \frac{1}{2}(1+\sqrt{5})$, *then*

$$\frac{L_n - c_1 n^{2-\alpha}}{(\log n)^{1/\alpha}} \xrightarrow{d} c_2 \varsigma.$$

(iii) *If* $\frac{1}{2}(1+\sqrt{5}) < \alpha < 2$, *then*

$$L_n - c_1 n^{2-\alpha} \xrightarrow{d} \eta,$$

where η *is a nondegenerate random variable.*

In fact it is not difficult to see from the proof that η has a density with respect to Lebesgue measure.

This transition at the golden ratio $\frac{1}{2}(1+\sqrt{5})$ is manifested in the results of Delmas et al. [8]. They also show that the number τ_n of collisions, properly rescaled, has an asymptotically stable distribution. This latter result has been independently obtained by Gnedin and Yakubovitch [12].

The region within the coalescent tree, where the random fluctuations of L_n asymptotically arise, are different in the three cases. In case (i) fluctuations come from everywhere between the root and the leaves, whereas in case (iii) they mainly originate at the neighborhood of the root. Then we have to take care of those summands $X_i(T_{i+1} - T_i)$ within L_n, which have an index i close to τ_n. In the intermediate case (ii) the primary contribution stems from summands with index i such that $\tau_n - n^{1-\varepsilon} \le i \le \tau_n - n^{\varepsilon}$ with $0 < \varepsilon < \frac{1}{2}$.

To get hold of these fluctuations, in proving the theorem, we, loosely speaking, turn around the order of summation in $L_n = \sum_{i=0}^{\tau_n - 1} X_i(T_{i+1} - T_i)$. We shall handle the reversed order by means of two point processes μ and ν on $\{2, 3, \ldots\}$. The first one, which we call the *coalescent's point process downwards from* ∞ [CPP(∞)], gives the asymptotic particle numbers seen from the root of the tree. Here we use Schweinsberg's result [18] implying that the Beta-coalescent comes down from infinity for $1 < \alpha < 2$; see [3], Corollary 3.2. (Therefore our method of proof does not apply to the case of the Bolthausen–Sznitman coalescent.) The second one is a classical stationary renewal point process, which can be reversed without difficulty. Two different couplings establish the links. Thereby the exponential holding times are left aside at first stage. In this respect, our approach to the Beta-coalescent differs from others as in Birkner et al. [5] or Berestyki et al. [1]. Certainly our proof can be extended to a larger class of Λ-coalescents having regular variation with index α between 1 and 2 (compare Definition 4.1 in [3]), which would require some additional technical efforts. It seems less obvious, whether our concept of a coalescent's point process downwards from ∞ can be realized for a much broader family of Λ-coalescents coming down from infinity.

Coalescent trees are used as a model for the genealogical relationship of n individuals backward to their most recent ancestor. Then one imagines that mutations are assigned to positions on the tree's branches in the manner of a Poisson point process with rate θ. Let S_n be the number of these segregation sites; see [3], Section 2.3.4. Given L_n the distribution of S_n is Poisson with mean θL_n. To get the asymptotic distribution one splits S_n into parts.

$$S_n - \theta c_1 n^{2-\alpha} = (S_n - \theta L_n) + \theta(L_n - c_1 n^{2-\alpha}).$$

Since $L_n / c_1 n^{2-\alpha}$ converges to 1 in probability, the first summand is asymptotically normal and also asymptotically independent from the second one. Its normalizing constant is $(\theta L_n)^{-1/2} \sim (\theta c_1)^{-1/2} n^{\alpha/2-1}$. Again there are two regimes. $n^{1-\alpha/2} = o(n^{1/\alpha+1-\alpha})$, if and only if $\alpha < \sqrt{2}$. Partial results are contained in Delmas et al. [8]. We obtain

COROLLARY 2. *Let ζ denote a standard normal random variable, which is independent of ς.*

(i) *If $1 < \alpha < \sqrt{2}$, then*

$$\frac{S_n - \theta c_1 n^{2-\alpha}}{n^{1/\alpha+1-\alpha}} \xrightarrow{d} \frac{\theta c_2 \varsigma}{(1+\alpha-\alpha^2)^{1/\alpha}}.$$

(ii) *If $\alpha = \sqrt{2}$, then*

$$\frac{S_n - \theta c_1 n^{2-\alpha}}{n^{1-\alpha/2}} \xrightarrow{d} \sqrt{\theta c_1}\zeta + \frac{\theta c_2 \varsigma}{(1+\alpha-\alpha^2)^{1/\alpha}}.$$

(iii) *If* $\sqrt{2} < \alpha < 2$, *then*

$$\frac{S_n - \theta c_1 n^{2-\alpha}}{n^{1-\alpha/2}} \xrightarrow{d} \sqrt{\theta c_1}\zeta.$$

This is the organization of the paper: Section 2 contains an elementary coupling of two $\mathbb{N}$-valued random variables. It is used in Section 3, where we introduce and analyze coalescent's point processes, and in Section 4, where we couple these point processes to stationary point processes. Section 5 assembles two auxiliary results on sums of independent random variables. Finally the proof of Theorem 1 is given in Section 6.

2. A coupling. In this section, let the natural number m be fixed. We introduce a coupling of the transition probabilities $P_{m,l}$ and a distribution, which does not depend on m. From the representation of the Beta-function by means of the Γ-function and its functional equation, we have

$$\begin{aligned}\rho_{m,m-k+1} &= \frac{1}{\Gamma(2-\alpha)\Gamma(\alpha)}\frac{m!}{\Gamma(m)}\frac{\Gamma(k-\alpha)}{k!}\frac{\Gamma(m-k+\alpha)}{(m-k)!}\\ &= \frac{1}{\Gamma(2-\alpha)\Gamma(\alpha)}\frac{\Gamma(k-\alpha)}{\Gamma(k+1)}\frac{(m-k+1)\cdots m}{(m-k+\alpha)\cdots(m-1+\alpha)}\frac{\Gamma(m+\alpha)}{\Gamma(m)},\end{aligned}$$

thus

$$P_{m,m-k} = d_{mk}\frac{\Gamma(k+1-\alpha)}{\Gamma(k+2)}, \qquad k \geq 1$$

with

$$d_{mk} = d_m \frac{(m-k)\cdots(m-1)}{(m+\alpha-k-1)\cdots(m+\alpha-2)}$$

and a normalizing constant $d_m > 0$ (also dependent on α). Recall from the Introduction that given $X_0 = m$ the quantities $P_{m,m-k}$ are the weights of the distribution of the downward jump $U = X_0 - X_1$. For a more detailed discussion of this "law of first jump," we refer to Delmas et al. [8].

It is natural to relate this distribution to the distribution of some random variable V with values in $\mathbb{N}$ and distribution given by

$$\mathbf{P}(V = k) = \frac{\alpha}{\Gamma(2-\alpha)}\frac{\Gamma(k+1-\alpha)}{\Gamma(k+2)}, \qquad k \geq 1. \tag{4}$$

This kind of distribution appears for Beta-coalescents already in Bertoin and Le Gall [4] (see their Lemma 4), in Berestycki et al. [1] (in the context of frequency spectra) as well as in Delmas et al. [8]. There the normalizing constant is determined and the following formulas derived:

$$\mathbf{E}(V) = \frac{1}{\alpha - 1} \quad \text{and} \quad \mathbf{P}(V \geq k) = \frac{1}{\Gamma(2-\alpha)}\frac{\Gamma(k+1-\alpha)}{\Gamma(k+1)}. \tag{5}$$

From Stirling's approximation,

$$\mathbf{P}(V=k)\sim\frac{\alpha}{\Gamma(2-\alpha)}k^{-\alpha-1}\quad\text{and}\quad\mathbf{P}(V\geq k)\sim\frac{1}{\Gamma(2-\alpha)}k^{-\alpha}.\tag{6}$$

The sequence d_{mk} is decreasing in k for fixed m, and thus the same is true for $P_{m,m-k}/\mathbf{P}(V=k)$. Therefore V stochastically dominates the jump size U, that is, for all $k\geq 1$,

$$\mathbf{P}(U\geq k\mid X_0=m)\leq\mathbf{P}(V\geq k).\tag{7}$$

We like to investigate a coupling of U and V, where $U\leq V$ a.s. It is fairly obvious that this can be achieved in such a way that

$$\mathbf{P}(U=k\mid V=k)=1\wedge\frac{P_{m,m-k}}{\mathbf{P}(V=k)}=1\wedge\frac{d_{mk}}{d}.\tag{8}$$

[Indeed one may put

$$\mathbf{P}(U=j\mid V=k)=\left(1-\frac{P_{m,m-k}}{\mathbf{P}(V=k)}\right)^{+}\frac{(P_{m,m-j}-\mathbf{P}(V=j))^{+}}{\mathbf{P}(U<k_m)-\mathbf{P}(V<k_m)}$$

for $j\neq k$ with $k_m=\min\{k\geq 1: P_{m,m-k}\leq\mathbf{P}(V=k)\}$. There are other possibilities; later it will only be important that we commit to one of them.]

LEMMA 3. *For a coupling (U,V) fulfilling (8), it holds*

$$\mathbf{P}(U\neq V)\leq\frac{1}{(\alpha-1)m}\quad\textit{and}\quad\mathbf{P}(V\geq k\mid U\neq V)\leq ck^{1-\alpha}$$

for all $k\geq 1$ and some $c<\infty$, which does not depend on m.

PROOF. Because of $\alpha<2$,

$$\frac{(m-k)\cdots(m-1)}{(m+\alpha-k-1)\cdots(m+\alpha-2)}\geq\frac{(m-k)\cdots(m-1)}{(m-k+1)\cdots m}=\frac{m-k}{m},$$

and because of $\alpha>1$

$$\frac{(m+\alpha-k-1)\cdots(m+\alpha-2)}{(m-k)\cdots(m-1)}\geq\left(\frac{m+\alpha-1}{m}\right)^k\geq 1+k\frac{\alpha-1}{m},$$

consequently

$$1-\frac{k}{m}\leq\frac{d_{mk}}{d_m}\leq\frac{1}{1+(\alpha-1)k/m}.$$

It follows

$$\left(1-\frac{k}{m}\right)\mathbf{P}(V=k)\leq\frac{d}{d_m}P_{m,m-k}\leq\mathbf{P}(V=k)$$

for all $k \geq 1$ with $d = \alpha/\Gamma(2-\alpha)$. Summing over k yields

$$1 - \frac{1}{m}\mathbf{E}(V) \leq \frac{d}{d_m} \leq 1 \quad \text{or} \quad 1 \leq \frac{d_m}{d} \leq \frac{1}{(1 - 1/((\alpha-1)m))^+}.$$

Combining the estimates, we end up with

$$1 - \frac{k}{m} \leq \frac{d_{mk}}{d} \leq \frac{1}{(1 + (\alpha-1)k/m)(1 - 1/((\alpha-1)m))^+} \tag{9}$$

for all $k \geq 1$.

Now from (8), (9)

$$\begin{aligned}\mathbf{P}(U \neq V) &= \sum_{k \geq 1} \big(\mathbf{P}(V = k) - P_{m,m-k}\big)^+ \\ &= \sum_{k \geq 1} \mathbf{P}(V = k)\left(1 - \frac{d_{mk}}{d}\right)^+ \leq \sum_{k \geq 1} \mathbf{P}(V = k)\frac{k}{m},\end{aligned}$$

and thus from (5),

$$\mathbf{P}(U \neq V) \leq \frac{1}{(\alpha-1)m}$$

which is our first claim.

Also, letting $m \geq 2/(\alpha-1)$ and $k' = 2\lceil(\alpha-1)^{-2} + (\alpha-1)^{-1}\rceil$, then

$$\begin{aligned}\left(1 + (\alpha-1)\frac{k'}{m}\right)\left(1 - \frac{1}{(\alpha-1)m}\right)^+ &= 1 + (\alpha-1)\frac{k'}{m} - \frac{1}{(\alpha-1)m} - \frac{k'}{m^2} \\ &\geq 1 + \frac{\alpha-1}{2}\frac{k'}{m} - \frac{1}{(\alpha-1)m} \geq 1 + \frac{1}{m}.\end{aligned}$$

From (9),

$$1 - \frac{d_{mk'}}{d} \geq 1 - \frac{1}{1 + 1/m} \geq \frac{1}{2m},$$

and from (8),

$$\mathbf{P}(U \neq V) \geq \mathbf{P}(U \neq k', V = k') = \left(1 - \frac{d_{mk'}}{d}\right)^+ \mathbf{P}(V = k') \geq \frac{1}{2m}\mathbf{P}(V = k')$$

for $m \geq 2/(\alpha-1)$. It follows that there is a $\eta > 0$ such that for all $m \geq 1$

$$\mathbf{P}(U \neq V) \geq \frac{1}{\eta m}.$$

Now from (8) and (9),

$$\begin{aligned}\mathbf{P}(V = k \mid U \neq V) &= \frac{\mathbf{P}(U \neq k, V = k)}{\mathbf{P}(U \neq V)} \\ &= \frac{(1 - d_{mk}/d)^+}{\mathbf{P}(U \neq V)}\mathbf{P}(V = k) \leq \eta k \mathbf{P}(V = k),\end{aligned}$$

and the second claim follows from (6). $\square$

3. The coalescent's point process. Let μ denote a point process on $\{2, 3, \dots\}$. For any interval I, let μ_I be the point process on $\{2, 3, \dots\}$ given by

$$\mu_I(B) = \mu(B \cap I), \qquad B \subset \{2, 3, \dots\}.$$

We call μ a *coalescent's point process* downwards from ∞, shortly a CPP(∞), if the following properties hold:

- $\mu(\{2, 3, \dots\}) = \infty$ and $\mu(\{n\}) = 0$ or 1 for any $n \geq 2$ a.s.
- For $n \geq 2$ we have that, given the event $\mu(\{n\}) = 1$, and given $\mu_{[n+1,\infty)}$ the point process $\mu_{[2,n]}$ is a CPP(n) a.s.

Recall that a point process is called a CPP(n), if it can be represented as in (1).

THEOREM 4. *Let $1 < \alpha < 2$. Then the* CPP(∞) *exists and is unique in distribution.*

We prepare the proof by two lemmas.

LEMMA 5. *Let μ be a* CPP(n) *with $1 < n \leq \infty$. Then for any $\varepsilon > 0$ there is a natural number r such that for any interval $I = [a, b]$ with $2 \leq a < b < n$ and $b - a \geq r$, we have*

$$\mathbf{P}\big(\mu(I) = 0\big) \leq \varepsilon.$$

PROOF. For $I = [a, b]$,

$$\{\mu(I) = 0\} = \bigcup_{m=b+1}^{n} \big\{\mu(\{m\}) = 1, \mu([a, m-1]) = 0\big\} \qquad \text{a.s.},$$

since $\mu(\{n\}) = 1$ for $n < \infty$ and $\mu(\{2, 3, \dots\}) = \infty$ a.s. for $n = \infty$. Thus from (1),

$$\mathbf{P}\big(\mu(I) = 0\big) \leq \sum_{m=b+1}^{n} \mathbf{P}(X_1 < a \mid X_0 = m).$$

Applying (7) to $U = X_0 - X_1$ it follows that

$$\mathbf{P}\big(\mu(I) = 0\big) \leq \sum_{m=b+1}^{n} \mathbf{P}(V > m - a) \leq \sum_{k=1}^{\infty} \mathbf{P}(V > b - a + k). \tag{10}$$

Since $\mathbf{E}(V) < \infty$, this series is convergent and the claim follows. □

The next lemma prepares a coupling of CPPs.

LEMMA 6. *Let μ, μ' be two independent CPPs coming down from $n, n' \leq \infty$. Then for any $\varepsilon > 0$ there is a natural number s such that for any b sufficiently large and $n, n' > b$, we have*

$$\mathbf{P}\big(\mu(\{j\}) = \mu'(\{j\}) = 1 \textit{ for some } j = b - s, \dots, b\big) \geq 1 - \varepsilon.$$

PROOF. First let $n < \infty$. We construct a coupling of a CPP(n) μ to an i.i.d. random sequence. Consider random variables $U_1, V_1, U_2, V_2, \dots$ and $n = X_0, X_1, \dots$ with $X_i = n - U_1 - \cdots - U_i$, which are constructed inductively as follows: If $U_1, V_1, \dots, U_i, V_i$ are already gotten, then given the values of these random variables let V_{i+1} be a copy of the random variable V from Section 2 and couple U_{i+1} to V_{i+1} as in Section 2, with $m = X_i$. For definiteness, put $U_{i+1} = 0$ if $X_i = 1$. Then $V_1, V_2, \dots$ are i.i.d. random variables with distribution (4), and $X_0 > X_1 > \cdots > X_{\tau_n - 1}$ are the points of a CPP(n) μ down from n, where τ_n is the natural number i such that $X_i = 1$ for the first time.

Now let k be a natural number. Then $X_{i-1} \geq n - U_1 - \cdots - U_k$ for $i \leq k$. Thus for any $\eta > 0$ and $n \geq 6k\eta^{-1}\mathbf{E}(V) + 2$, from Lemma 3,

$$\mathbf{P}\big(U_i \neq V_i \text{ for some } i \leq k, U_1 + \cdots + U_k \leq 6k\eta^{-1}\mathbf{E}(V)\big)$$

$$\leq \sum_{i=1}^{k} \mathbf{P}\big(U_i \neq V_i, X_{i-1} \geq n - 6k\eta^{-1}\mathbf{E}(V)\big) \leq \frac{k}{(\alpha-1)(n - 6k\eta^{-1}\mathbf{E}(V))},$$

thus

$$\mathbf{P}\big(U_i \neq V_i \text{ for some } i \leq k, U_1 + \cdots + U_k \leq 6k\eta^{-1}\mathbf{E}(V)\big) \leq \frac{\eta}{6}$$

if n is large enough. Also $\mathbf{E}(U_i) \leq \mathbf{E}(V)$ because of (7). Thus from Markov's inequality,

$$\mathbf{P}\big(U_1 + \cdots + U_k > 6k\eta^{-1}\mathbf{E}(V)\big) \leq \frac{\eta}{6}, \tag{11}$$

and consequently

$$\mathbf{P}(U_i \neq V_i \text{ for some } i \leq k) \leq \frac{\eta}{3}$$

if n is sufficiently large (depending on η and k).

Next let l be a natural number and $n' = n + l$. Let $U'_1, V'_1, U'_2, V'_2, \dots$ and $n' = X'_0, X'_1, \dots$ an analog construction with random variables, which are independent of $U_1, V_1, U_2, V_2, \dots$. Then also

$$\mathbf{P}(U'_i \neq V'_i \text{ for some } i \leq k) \leq \frac{\eta}{3}.$$

Moreover because V has finite expectation and because of independence from classical results on recurrent random walks,

$$\mathbf{P}\left(\sum_{i=1}^{j} V_i \neq \sum_{i=1}^{j} V'_i - l \text{ for all } j \leq k\right) \leq \frac{\eta}{6},$$

if only k is sufficiently large (depending on l). Combining the estimates we obtain

$$\mathbf{P}\left(\sum_{i=1}^{j} U_i \neq \sum_{i=1}^{j} U'_i - l \text{ for all } j \leq k\right) \leq \frac{5\eta}{6}.$$

For the corresponding independent CPPs μ and μ' coming down from n and $n' = n + l$, this implies, together with (11),

$$\mathbf{P}\big(\mu(\{j\}) = \mu'(\{j\}) = 1 \text{ for some } j \in [n - 6k\eta^{-1}\mathbf{E}(V), n]\big) \geq 1 - \eta.$$

Leaving aside the coupling procedure we have proved the following: Let $\eta > 0$, let l be a natural number and let μ and μ' denote independent CPPs coming down from $n < \infty$ and $n' = n + l$. Then there is a natural number r' such that

$$\mathbf{P}\big(\mu(\{j\}) = \mu'(\{j\}) = 1 \text{ for some } j = n - r', \dots, n\big) \geq 1 - \eta, \tag{12}$$

if only n is large enough.

With this preparation we come to the proof of the lemma. Let $\varepsilon > 0$, $b \geq 2$ and let $n, n' > b$. Denote

$$M = \max\{k \leq b : \mu(\{k\}) = 1\}, \qquad M' = \max\{k \leq b : \mu'(\{k\}) = 1\}$$

(with the convention $M = 1$, if $\mu([2, b]) = 0$). From Lemma 5,

$$\mathbf{P}(M, M' \in [b - r, b]) \geq 1 - \frac{\varepsilon}{2}$$

for some r and $b > r + 2$. Then

$$\begin{aligned}
&\mathbf{P}\big(\mu(\{j\}) = \mu'(\{j\}) = 1 \text{ for no } j \in [b - r' - r, b]\big) \\
&\quad \leq \frac{\varepsilon}{2} + \mathbf{P}\big(\mu(\{j\}) = \mu'(\{j\}) = 1 \\
&\qquad\qquad \text{for no } j \in [b - r' - r, b]; b - r \leq M, M' \leq b\big) \\
&\quad \leq \frac{\varepsilon}{2} + 2 \sum_{b - r \leq m < m' \leq b} \mathbf{P}\big(\mu(\{j\}) = \mu'(\{j\}) = 1 \\
&\qquad\qquad\qquad \text{for no } j = m - r', \dots, m \mid X_0 = m, X_0' = m'\big).
\end{aligned}$$

From (12) it follows that the right-hand probabilities are bounded by $\eta = \varepsilon/4r^2$, if b is only sufficiently large. Then

$$\mathbf{P}\big(\mu(\{j\}) = \mu'(\{j\}) = 1 \text{ for no } j \in [b - r' - r, b]\big) \leq \varepsilon,$$

which is our claim with $s = r + r'$. $\square$

As a corollary, we note:

LEMMA 7. *Let μ and μ' be two independent* CPP(∞). *Then a.s.* $\mu(\{j\}) = \mu'(\{j\}) = 1$ *for infinitely many* $j \in \mathbb{N}$.

PROOF. From the preceding lemma there are numbers $b_1 < b_2 < \cdots$ such that

$$\mathbf{P}\big(\mu(\{j\}) = \mu'(\{j\}) = 1 \text{ for no } j = b_k, \dots, b_{k+1}\big) \leq 2^{-k}.$$

Now an application of the Borel–Cantelli lemma gives the claim. $\square$

PROOF OF THEOREM 4. The existence follows from the fact that for $\alpha > 1$ the corresponding Beta-coalescent $(\Pi_t)_{t\geq 0}$ comes down from infinity [18], which means that the number of blocks in Π_t is a finite number N_t for each $t > 0$. Put $\mu(\{k\}) = 1$, if and only if $N_t = k$ for some $t > 0$.

Uniqueness follows from the last lemma and a standard coupling argument. $\square$

4. A bigger coupling. Now let ν be a stationary renewal point process on $\{2, 3, \ldots\}$; that is, if we denote the points of ν by $2 \leq R_1 < R_2 < \cdots$, then the increments $R_{i+1} - R_i$ are independent for $i \geq 0$ (with $R_0 = 1$) and $R_{i+1} - R_i$ has for $i \geq 1$ the distribution (4). A stationary version of the process exists, since $\mathbf{E}(V) < \infty$, such that the distribution of R_1 may be adjusted in the usual way to obtain stationarity, that is,

$$\mathbf{P}(R_1 = r) = \frac{\mathbf{P}(V \geq r - 1)}{\mathbf{E}(V)}, \qquad r = 2, 3, \ldots. \tag{13}$$

Stationarity is of advantage for us. Then ν may be considered as restriction of a stationary point process on $\mathbb{Z}$. Such a process is invariant in distribution under the transformation $z \mapsto z_0 - z$, $z \in \mathbb{Z}$ with $z_0 \in \mathbb{Z}$. Therefore ν, restricted to $\{2, \ldots, n\}$ looks the same, when considered upwards or downwards.

In this section we introduce a coupling between ν and the CPP(∞) μ, which allows us later to replace μ by ν. Given $b \geq 2$ let, as above,

$$M = \max\bigl\{k \leq b : \mu(\{k\}) = 1\bigr\}, \qquad M' = \max\bigl\{k \leq b : \nu(\{k\}) = 1\bigr\}.$$

Again, if there is no $k \leq b$ such that $\mu(\{k\}) = 1$, we put $M = 1$, and similary for M'. Let λ_b and λ'_b denote the distributions of M and M' (both dependent on b).

Now for $r \in \mathbb{N}$ we consider the following construction of μ and ν, restricted to $[2^{r-1} + 1, 2^r]$. Take any coupling (M, M') of λ_{2^r} and λ'_{2^r}. Given (M, M') construct random variables $U_1, V_1, U_2, V_2, \ldots$ inductively as in the proof of Lemma 6, using the coupling of Section 2. Here we start with $X_0 = M$. Also let $Y_0 = M'$,

$$X_i = M - U_1 - \cdots - U_i, \qquad Y_i = M' - V_1 - \cdots - V_i, \qquad i \geq 1, \tag{14}$$

and

$$N = \min\{i \geq 0 : X_i \leq 2^{r-1}\}, \qquad N' = \min\{i \geq 0 : Y_i \leq 2^{r-1}\}. \tag{15}$$

The whole construction is interrupted at the moment $N \vee N'$. Maybe $M, M' \leq 2^{r-1}$, then no step of the construction is required. Clearly the following statements are true:

- The point process $\sum_{i=0}^{N-1} \delta_{X_i}$ is equal in distribution to μ, restricted to $[2^{r-1} + 1, 2^r]$.

- The point process $\sum_{i=0}^{N'-1} \delta_{Y_i}$ is equal in distribution to ν, restricted to $[2^{r-1}+1, 2^r]$.
- X_N and $Y_{N'} \vee 1$ have the distributions $\lambda_{2^{r-1}}$ and $\lambda'_{2^{r-1}}$.

The complete coupling is

$$\begin{aligned} \Phi^r(M, M') &= \left(\sum_{i=0}^{N-1} \delta_{X_i}, \sum_{i=0}^{N'-1} \delta_{Y_i}, X_N, Y_{N'} \vee 1\right) \\ &= (\phi_1^r, \phi_2^r, \phi_3^r, \phi_4^r) \qquad \text{(say)}. \end{aligned} \tag{16}$$

Its distribution is uniquely determined by the distribution of the coupling (U, V) from Section 2. The following continuity property is obvious:

- If we have a sequence (M_n, M_n') of couplings of λ_{2^r} and λ'_{2^r} such that $(M_n, M_n') \xrightarrow{d} (M, M')$, then (M, M') is also a coupling of λ_{2^r} and λ'_{2^r} and

$$\Phi^r(M_n, M_n') \xrightarrow{d} \Phi^r(M, M').$$

Another obvious fact is that this construction can be iterated: Given $\Phi^r(M, M')$ we construct $\Phi^{r-1}(\phi_3^r, \phi_4^r)$ and so forth. Thus starting with the independent coupling (M, M') (i.e., M and M' are independent), we obtain the tupel

$$\Psi^r = (\Phi^{1,r}(M_{1,r}, M'_{1,r}), \Phi^{2,r}(M_{2,r}, M'_{2,r}), \dots, \Phi^{r,r}(M_{r,r}, M'_{r,r})),$$

where $(M_{r,r}, M'_{r,r}) = (M, M')$ and $(M_{s,r}, M'_{s,r}) = (\phi_3^{s+1,r}, \phi_4^{s+1,r})$ for $s < r$. Since $M_{s,r}$ and $M'_{s,r}$ are no longer independent in general, the tupels Ψ^r are initially not consistent for different r. To enforce consistency note that for fixed s the distributions of $(M_{s,r}, M'_{s,r})$ are tight for $r \geq s$, since they take values in the finite set $\{1, \dots, 2^s\} \times \{1, \dots, 2^s\}$. Thus by a diagonalization argument, we may obtain a sequence $1 \leq r_1 < r_2 < \cdots$ such that

$$(M_{s,r_n}, M_{s,r_n}) \xrightarrow{d} (M_{s,\infty}, M'_{s,\infty})$$

for certain couplings $(M_{s,\infty}, M'_{s,\infty})$ of λ_{2^s} and λ'_{2^s}.

If we make use instead of the independent coupling (M, M'), now $(M_{r,\infty}, M'_{r,\infty})$ as starting configuration in the construction of Ψ^r, then we gain consistency in the sense that

$$\Psi^{r-1} \stackrel{d}{=} (\Phi^{1,r}(M_{1,r}, M'_{1,r}), \Phi^{2,r}(M_{2,r}, M'_{2,r}), \dots, \Phi^{r-1,r}(M_{r-1,r}, M'_{r-1,r})).$$

Proceeding to the projective limit, we obtain the "big coupling,"

$$\Psi^\infty = (\Phi^{1,\infty}(M_{1,\infty}, M'_{1,\infty}), \Phi^{2,\infty}(M_{2,\infty}, M'_{2,\infty}), \dots). \tag{17}$$

It has the property that

$$\mu = \sum_{r=1}^\infty \phi_1^{r,\infty} \quad \text{and} \quad \nu = \sum_{r=1}^\infty \phi_2^{r,\infty} \tag{18}$$

are coupled copies of our CPP(∞) and stationary point process.

In order to estimate the difference between both point processes, we go back to (14), (15) and estimate the tail of the distribution of

$$D_r = \max_{i \le N \wedge N'} |X_i - Y_i|. \tag{19}$$

LEMMA 8. *There is a constant $c > 0$ such that for all $r \ge 1$ and all $t > 0$,*

$$\mathbf{P}(D_r > t) \le ct^{1-\alpha}.$$

PROOF. For $i \le N \wedge N'$, we have

$$\begin{aligned} |X_i - Y_i| &\le \sum_{j \le N \wedge N'} |U_j - V_j| + |X_0 - Y_0| \\ &\le \sum_{j \le N \wedge N'} |U_j - V_j| + (2^r - M) + (2^r - M'). \end{aligned} \tag{20}$$

From (6), (10),

$$\mathbf{P}(2^r - M > t) \le \sum_{k \ge t} \mathbf{P}(V \ge k) \le ct^{1-\alpha} \tag{21}$$

for a suitable $c > 0$.

Because of stationarity $2^r - M'$ and $(R_1 - 2) \wedge (2^r - 1)$ are equal in distribution, therefore because of (6), (13)

$$\mathbf{P}(2^r - M' > t) \le \mathbf{P}(R_1 > t) \le ct^{1-\alpha} \tag{22}$$

for a suitable $c > 0$.

Finally from Lemma 3 $U_j \ne V_j$ occurs for $j \le N$ at most with probability $p = 2^{1-r}/(\alpha - 1)$ and then $|U_j - V_j| \le V_j$ a.s. Also because of Lemma 3 these V_j can be stochastically dominated by random variables $a + b\zeta_j$ with constants $a, b > 0$ and positive i.i.d. random variables ζ_j, which possess a stable distribution of index $\alpha - 1$ and Laplace transform $\exp(-\lambda^{\alpha-1})$. Also $N \wedge N' \le 2^{r-1} = w$ (say). Thus $\sum_{j \le N \wedge N'} |U_j - V_j|$ is stochastically dominated by the random variable

$$W = \sum_{j=0}^{w} (a + b\zeta_j) I_j,$$

where I_j are i.i.d. Bernoulli with success probability p. Let $\varphi(\lambda) = \exp(-a\lambda - (b\lambda)^{\alpha-1})$ be the Laplace transform of $a + b\zeta_j$. Then W has the Laplace transform

$$\sigma(\lambda) = \bigl(1 - p\bigl(1 - \varphi(\lambda)\bigr)\bigr)^w.$$

It follows $1-\sigma(\lambda) \le wp(1-\varphi(\lambda)) \le (1-\varphi(\lambda))/(\alpha-1)$. From the well-known identity $\lambda \int_0^\infty e^{-\lambda x} \mathbf{P}(W > x)\,dx = 1-\sigma(\lambda)$, it follows that

$$\begin{aligned} e^{-1}\mathbf{P}(W > t) &\le t^{-1}\int_0^\infty e^{-x/t}\mathbf{P}(W > x)\,dx \\ &= 1-\sigma(1/t) \le \frac{1}{\alpha-1}\big(1-\exp\big(-at^{-1}-(bt)^{1-\alpha}\big)\big). \end{aligned}$$

Thus

$$\mathbf{P}\biggl(\sum_{j\le N\wedge N'} |U_j - V_j| > t\biggr) \le \mathbf{P}(W > t) \le ct^{1-\alpha} \tag{23}$$

for a suitable $c > 0$. Using estimates (21) to (23) in (20) yields our claim. □

Additionally, we note that

$$|N - N'| \le D_r. \tag{24}$$

Indeed, if $N < N'$, then $X_N \le 2^{r-1}$, thus $Y_N \le 2^{r-1} + D_r$. Further $Y_{N'-1} > 2^{r-1}$, which implies $N' - 1 - N \le Y_N - Y_{N'-1} \le D_r - 1$. The case $N' < N$ is treated in the same way.

5. On sums of independent random variables.

The following lemma can be deduced from well-known results (see, e.g., Petrov [15]), but a direct proof seems more convenient. Let

$$\gamma = \frac{1}{\alpha-1}.$$

LEMMA 9. *Let $V_1, V_2, \ldots$ be i.i.d. copies of the random variable* (4). *Then for any $\beta \in \mathbb{R}$ and any $\varepsilon > 0$ a.s.,*

$$\sum_{k=1}^n k^{-\beta}(V_k - \gamma) = \eta_n + o(n^{1/\alpha-\beta+\varepsilon}),$$

where η_n is a.s. convergent.

PROOF. Let $\varepsilon > 0$. A short calculation gives that $\mathbf{E}(V_k^2; V_k \le k^{1/\alpha+\varepsilon})$ is of order $k^{2/\alpha-1+(2-\alpha)\varepsilon}$; thus

$$\sum_{k=1}^\infty k^{-1/\alpha-\varepsilon}\big(V_k 1_{V_k \le k^{1/\alpha+\varepsilon}} - \mathbf{E}(V_k; V_k \le k^{1/\alpha+\varepsilon})\big)$$

is a.s. convergent. Also $\mathbf{E}(V_k; V_k > k^{1/\alpha+\varepsilon})$ is of order less than $k^{1/\alpha-1}$ and $\mathbf{P}(V_k > k^{1/\alpha+\varepsilon})$ is of order $k^{-1-\alpha\varepsilon}$ such that $V_k > k^{1/\alpha+\varepsilon}$ occurs only finitely often a.s. Thus

$$\sum_{k=1}^\infty k^{-1/\alpha-\varepsilon}(V_k - \gamma)$$

is a.s. convergent for all $\varepsilon > 0$.

For $\beta > \frac{1}{\alpha}$, it follows that the sum $\sum_{k=1}^{n} k^{-\beta}(V_k - \gamma)$ is a.s. convergent, which is our claim [then the term $o(n^{1/\alpha-\beta+\varepsilon})$ is superfluous]. In the case $\beta \le \frac{1}{\alpha}$ by Kronecker's lemma a.s.,

$$\sum_{k=1}^{n} k^{-\beta}(V_k - \gamma) = o(n^{1/\alpha-\beta+\varepsilon}),$$

which again is our claim (now η_n is superfluous). $\square$

Next recall that ς denotes a random variable with maximally skewed stable distribution of index α as in (3). The following result can be deduced from a general statement on triangular arrays of independent random variables; see [11], Chapter XVII, Section 7; however, a direct proof seems easier.

LEMMA 10. *Let $V_1, V_2, \ldots$ be independent copies of the random variable* (4). *Then the following holds true*:

(i) *Let $1 < \alpha < \frac{1}{2}(1+\sqrt{5})$. Then*

$$n^{\alpha-1-1/\alpha} \sum_{k=1}^{n} k^{1-\alpha}(V_k - \gamma) \xrightarrow{d} -c\varsigma,$$

where

$$c = \big((1+\alpha-\alpha^2)\Gamma(2-\alpha)\big)^{-1/\alpha}.$$

(ii) *For $\alpha = \frac{1}{2}(1+\sqrt{5})$*

$$(\log n)^{-1/\alpha} \sum_{k=1}^{n} k^{1-\alpha}(V_k - \gamma) \xrightarrow{d} \frac{-\varsigma}{\Gamma(2-\alpha)^{1/\alpha}}.$$

PROOF. (i): From (5), (6) and the theory of stable laws, it follows that

$$n^{-1/\alpha}(V_1 + \cdots + V_n - \gamma n) \xrightarrow{d} \frac{-\varsigma}{\Gamma(2-\alpha)^{1/\alpha}}.$$

We express this relation by means of the characteristic functions $\varphi(u)$ and $e^{\psi(u)}$ of $V - \gamma$ and $-\varsigma/\Gamma(2-\alpha)^{1/\alpha} : \varphi(n^{-1/\alpha}u)^n \to e^{\psi(u)}$ for all $u \in \mathbb{R}$, or slightly more generally,

$$\varphi(v_n n^{-1/\alpha}u)^n \to e^{\psi(u)},$$

if $v_n \to 1$. Since $\varphi_n(u) = \varphi(v_n n^{-1/\alpha}u)$ is again a characteristic function, it follows from Feller [11], Chapter XVII.1, Theorem 1, that for $n \to \infty$

$$n\big(\varphi(v_n n^{-1/\alpha}u) - 1\big) \to \psi(u)$$

or

$$\varphi(su) - 1 \sim s^\alpha \psi(u) \qquad \text{as } s \to 0$$

for all real u. Since $\alpha - \alpha^2 > -1$ for $\alpha < \frac{1}{2}(1+\sqrt{5})$, it follows that with $\zeta = (1+\alpha-\alpha^2)^{1/\alpha}$

$$\sum_{k=1}^{n}\left(\varphi\left(\frac{\zeta k^{1-\alpha}}{n^{1-\alpha+1/\alpha}}u\right) - 1\right) \sim \psi(u)\sum_{k=1}^{n}\left(\frac{\zeta k^{1-\alpha}}{n^{1-\alpha+1/\alpha}}\right)^{\alpha} \to \psi(u).$$

Similarly,

$$\sum_{k=1}^{n}\left|\varphi\left(\frac{\zeta k^{1-\alpha}}{n^{1-\alpha+1/\alpha}}u\right) - 1\right| \to |\psi(u)|,$$

and consequently,

$$\begin{aligned}&\sum_{k=1}^{n}\left|\varphi\left(\frac{\zeta k^{1-\alpha}}{n^{1-\alpha+1/\alpha}}u\right) - 1\right|^2 \\ &\qquad \le \max_{k=1,\dots,n}\left|\varphi\left(\frac{\zeta k^{1-\alpha}}{n^{1-\alpha+1/\alpha}}u\right) - 1\right|\sum_{k=1}^{n}\left|\varphi\left(\frac{\zeta k^{1-\alpha}}{n^{1-\alpha+1/\alpha}}u\right) - 1\right| \to 0\end{aligned}$$

for $n \to \infty$.

In order to transfer these limit results to characteristic functions we use that for all complex numbers z with $|z| \le 1$,

$$|z - e^{z-1}| \le c|z-1|^2$$

for some $c > 0$. Therefore, if $|z_1|, \dots, |z_n| \le 1$,

$$\left|z_1 \cdots z_n - e^{(z_1-1)+\cdots+(z_n-1)}\right| \le \sum_{k=1}^{n}|z_k - e^{z_k-1}| \le c\sum_{k=1}^{n}|z_k - 1|^2.$$

We put $z_k = z_{kn}(u) = \varphi(\frac{\zeta k^{1-\alpha}}{n^{1-\alpha+1/\alpha}}u)$. Then the right-hand side goes to zero, and we obtain

$$z_{1n}(u)\cdots z_{nn}(u) \to e^{\psi(u)}.$$

Since the product on the left-hand side is the characteristic function of $\zeta n^{\alpha-1-1/\alpha} \times \sum_{k=1}^{n} k^{1-\alpha}(V_k - \frac{1}{\alpha-1})$ the claim follows.

(ii): This proof goes along the same lines using

$$\sum_{k=1}^{n}\left(\varphi\left(\frac{k^{1-\alpha}}{(\log n)^{1/\alpha}}u\right) - 1\right) \sim \psi(u)\sum_{k=1}^{n}\left(\frac{k^{1-\alpha}}{(\log n)^{1/\alpha}}\right)^{\alpha}.$$

Now $\alpha - \alpha^2 = -1$, thus

$$\sum_{k=1}^{n}\left(\varphi\left(\frac{k^{1-\alpha}}{(\log n)^{1/\alpha}}u\right) - 1\right) \sim \psi(u)\frac{1}{\log n}\sum_{k=1}^{n}\frac{1}{k} \sim \psi(u),$$

and the claim follows. $\square$

6. Proof of Theorem 1. Again let $2 \le R_1 < R_2 < \cdots$ be the points of the stationary point process ν, and denote

$$V_j = R_{j+1} - R_j, \qquad j \ge 1.$$

The random variables $V_1, V_2, \dots$ are i.i.d. with distribution (4).

LEMMA 11. *We have*

$$\int_{[2,n]} x^{1-\alpha} \nu(dx) = \frac{n^{2-\alpha}}{\gamma(2-\alpha)} - \gamma^{-\alpha} \sum_{k \le n/\gamma} k^{1-\alpha}(V_k - \gamma) + \delta_n$$

with

$$\delta_n = \eta_n + o_P(n^{1/\alpha^2+1-\alpha+\varepsilon})$$

for any $\varepsilon > 0$, *where the random variables* η_n *are convergent in probability.*

PROOF. Our starting point is

$$\int_{[2,n]} x^{1-\alpha} \nu(dx) = \sum_{i=1}^{r_n} R_i^{1-\alpha},$$

where r_n is such that $R_{r_n} \le n < R_{r_n+1}$. From Lemma 9 we have $R_n - \gamma n = o(n^{1/\alpha+\varepsilon})$ a.s., which implies $r_n - \frac{n}{\gamma} = o(n^{1/\alpha+\varepsilon})$ a.s.

By a Taylor expansion,

$$\begin{aligned} R_i^{1-\alpha} &= (\gamma i)^{1-\alpha} + (1-\alpha)(\gamma i)^{-\alpha}(R_i - \gamma i) + \delta_i' \\ &= (\gamma i)^{1-\alpha} + (1-\alpha)(\gamma i)^{-\alpha} \sum_{j=1}^{i-1}(V_j - \gamma) + \delta_i'', \end{aligned} \tag{25}$$

where the remainder is a.s. of the order

$$\delta_i'' = O\big(i^{-\alpha-1}(R_i - \gamma i)^2\big) + O(i^{-\alpha}) = o(i^{2/\alpha-\alpha-1+\varepsilon}).$$

We consider now the sums of the different terms in (25).

$$\sum_{i=1}^{r_n} (\gamma i)^{1-\alpha} = \frac{\gamma^{1-\alpha}}{2-\alpha} r_n^{2-\alpha} + \eta_n', \tag{26}$$

where η_n' is a.s. convergent. Further, putting $a_n = (\alpha - 1)\sum_{i>n} i^{-\alpha}$,

$$\begin{aligned} (1-\alpha)\sum_{i=1}^{r_n} i^{-\alpha} \sum_{j=1}^{i-1}(V_j - \gamma) &= (1-\alpha) \sum_{j=1}^{r_n-1}(V_j - \gamma) \sum_{i=j+1}^{r_n} i^{-\alpha} \\ &= a_{r_n}(R_{r_n+1} - R_1 - \gamma r_n) - \sum_{j=1}^{r_n} a_j(V_j - \gamma). \end{aligned}$$

The distribution of $R_{r_n+1} - n$ does not depend on n because of stationarity, thus $a_{r_n}(R_{r_n+1} - R_1 - n) = O_P(n^{1-\alpha})$. Also $\sum_{j=1}^{n}(a_j - j^{1-\alpha})(V_j - \gamma)$ is a.s. convergent for $\alpha > 1$, since $a_n - n^{1-\alpha} = O(n^{-\alpha})$ and since V has finite expectation. It follows

$$(1-\alpha)\sum_{i=1}^{r_n} i^{-\alpha}\sum_{j=1}^{i-1}(V_j - \gamma) = r_n^{1-\alpha}(n - \gamma r_n) - \sum_{j=1}^{r_n} j^{1-\alpha}(V_j - \gamma) + \eta_n'' + O_P(n^{1-\alpha}), \tag{27}$$

where η_n'' is a.s. convergent. Next

$$\sum_{i=1}^{r_n}\delta_i'' = \eta_n''' + o(n^{2/\alpha-\alpha+\varepsilon}) \qquad \text{a.s.} \tag{28}$$

for all $\varepsilon > 0$, where η_n''' is a.s. convergent. Note that this formula covers two cases: If $\frac{2}{\alpha} < \alpha$, then the sum is a.s. convergent and the right-hand term is superfluous. Otherwise the term η_n''' can be neglected.

Furthermore another Taylor expansion gives

$$\frac{n^{2-\alpha}}{2-\alpha} = \frac{(\gamma r_n)^{2-\alpha}}{2-\alpha} + (\gamma r_n)^{1-\alpha}(n - \gamma r_n) + o(n^{2/\alpha-\alpha+\varepsilon}) \qquad \text{a.s.} \tag{29}$$

Combining (25) to (29) gives

$$\sum_{i=1}^{r_n} R_i^{1-\alpha} = \frac{n^{2-\alpha}}{\gamma(2-\alpha)} - \gamma^{-\alpha}\sum_{j=1}^{r_n} j^{1-\alpha}(V_j - \gamma) + \eta_n + o(n^{2/\alpha-\alpha+\varepsilon}) \qquad \text{a.s.,} \tag{30}$$

where η_n is convergent in probability.

Finally we consider the (loosely notated) difference

$$\sum_{j=r_n+1}^{n/\gamma} j^{1-\alpha}(V_j - \gamma) = \sum_{j \le n/\gamma} j^{1-\alpha}(V_j - \gamma) - \sum_{j=1}^{r_n} j^{1-\alpha}(V_j - \gamma).$$

For any random sequence of natural numbers s_n such that $s_n = o(n^{1/\alpha+\varepsilon})$ a.s. for all $\varepsilon > 0$,

$$\sum_{i \le s_n}(V_i - \gamma) = R_{s_n+1} - R_1 - \gamma s_n = o(s_n^{1/\alpha+\varepsilon}) = o(n^{1/\alpha^2+2\varepsilon+\varepsilon^2}) \qquad \text{a.s.}$$

Since $r_n - n/\gamma = o(n^{1/\alpha+\varepsilon})$ a.s. for any $\varepsilon > 0$, this implies for any $\varepsilon > 0$ in probability,

$$\sum_{j=r_n+1}^{n/\gamma}(V_j - \gamma) = o_P(n^{1/\alpha^2+\varepsilon}).$$

This implies $\sum_{j=r_n+1}^{n/\gamma}(V_j+\gamma)=o_P(n^{1/\alpha+\varepsilon})$. Therefore

$$\begin{aligned}
&\left|\sum_{j=r_n+1}^{n/\gamma} j^{1-\alpha}(V_j-\gamma)\right| \\
&\qquad \le r_n^{1-\alpha}\left|\sum_{j=r_n+1}^{n/\gamma}(V_j-\gamma)\right| + \left|\left(\frac{n}{\gamma}\right)^{1-\alpha} - r_n^{1-\alpha}\right| \sum_{j=r_n+1}^{n/\gamma}(V_j+\gamma) \\
&\qquad = o_P(n^{1/\alpha^2+1-\alpha+\varepsilon}) + O\big(n^{-\alpha}(n-\gamma r_n)\big)o_P(n^{1/\alpha+\varepsilon}) \\
&\qquad = o_P(n^{1/\alpha^2+1-\alpha+\varepsilon}) + o_P(n^{2/\alpha-\alpha+2\varepsilon}).
\end{aligned}$$

Since $\frac{1}{\alpha^2}+1\ge\frac{2}{\alpha}$, we end up with

$$\sum_{j=r_n+1}^{n/\gamma} j^{1-\alpha}(V_j-\gamma)=o_P(n^{1/\alpha^2+1-\alpha+\varepsilon}).$$

Combining this estimate with (30) gives the claim. $\square$

PROOF OF THEOREM 1. The total length (2) of the n-coalescent can be rewritten as

$$L_n=\sum_{i=0}^{\tau_n-1}\frac{X_i}{\rho_{X_i}}E_i,$$

where $E_0, E_1, \ldots$ denote exponential random variables with expectation 1, independent among themselves and from the X_i.

From Lemma 2.2 in Delmas et al. [8], we have for $m\to\infty$

$$\rho_m=\frac{1}{\alpha\Gamma(\alpha)}m^\alpha+O(m^{\alpha-1}). \tag{31}$$

In the first step we replace the points $n=X_0>X_1>\cdots$ of a CPP(n) by points of a CPP(∞): If we take independent versions of both then for given $\varepsilon>0$ by Lemma 6, there is a natural number $s\ge 1$ such that with probability at least $1-\varepsilon$ they meet before $n-s$. From this moment both CPPs can be coupled. Thus, letting $n\ge X_0'>X_1'>\cdots$ be the points of the coupled CPP(∞) within $[2,n]$, independent of $E_0, E_1, \ldots,$ and

$$L_n'=\sum_{i=0}^{\tau_n'-1}\frac{X_i'}{\rho_{X_i'}}E_i,$$

then due to the the coupling and (31) for n sufficiently big

$$\mathbf{P}\big(|L_n-L_n'|>3\alpha\Gamma(\alpha)n^{1-\alpha}(E_0+\cdots+E_s)\big)\le\varepsilon.$$

Since $\alpha > 1$, $L_n - L_n' = o_P(1)$, thus we may replace L_n by L_n' in our asymptotic considerations.

Thus we work now with a CPP(∞) μ, which we couple to a stationary point process ν according to (17) and (18). Also let $E_0, E_1, \dots$ be independent of the whole coupling. We use the formula

$$L_n' = \int_{[2,n]} \frac{xE_x}{\rho_x} \mu(dx), \tag{32}$$

in which the exponential random variables now are ordered differently. Since $\sum_{x\geq 1} x^{-\alpha} E_x < \infty$ a.s., it follows from (31) that

$$L_n' = \alpha\Gamma(\alpha) \int_{[2,n]} \frac{E_x}{x^{\alpha-1}} \mu(dx) + \eta_{1,n},$$

where $\eta_{1,n}$ is a.s. convergent.

Next $\sum_{x\geq 2} x^{-1/2-\varepsilon}(E_x - 1)$ is a.s. convergent for any $\varepsilon > 0$. It follows that $\sum_{x\geq 2} x^{1-\alpha}(E_x - 1)$ is a.s. convergent for $\alpha > \frac{3}{2}$ and else a.s. of order $O(n^{3/2-\alpha+\varepsilon})$. Given μ, the same holds true for $\int_{[2,n]} \frac{E_x - 1}{x^{\alpha-1}} \mu(dx)$, thus

$$L_n' = \alpha\Gamma(\alpha) \int_{[2,n]} x^{1-\alpha} \mu(dx) + \eta_{2,n} + o(n^{3/2-\alpha+\varepsilon}) \qquad \text{a.s.},$$

where again $\eta_{2,n}$ is a.s. convergent.

Next from (18) with $2^s < n \leq 2^{s+1}$

$$\begin{aligned}\int_{[2,n]} x^{1-\alpha} \mu(dx) &= \int_{[2,n]} x^{1-\alpha} \nu(dx) \\ &\quad + \sum_{r=1}^{s} \int_{[2^{r-1}+1,2^r]} x^{1-\alpha} (\phi_1^{r,\infty}(dx) - \phi_2^{r,\infty}(dx)) \\ &\quad + \int_{[2^s+1,n]} x^{1-\alpha} (\phi_1^{s,\infty}(dx) - \phi_2^{s,\infty}(dx)).\end{aligned}$$

From (19) and (24) we see that

$$\begin{aligned}&\left| \int_{[2^{r-1}+1,2^r]} x^{1-\alpha} (\phi_1^{r,\infty}(dx) - \phi_2^{r,\infty}(dx)) \right| \\ &\qquad \leq 2^{r-1}(\alpha - 1)(2^{r-1})^{-\alpha} D_r + 2(2^{r-1})^{1-\alpha} D_r,\end{aligned}$$

and the same estimate holds for the last term above. In view of Lemma 8 and the Borel–Cantelli lemma, we conclude that

$$\int_{[2,n]} x^{1-\alpha} \mu(dx) = \int_{[2,n]} x^{1-\alpha} \nu(dx) + \eta_{3,n}$$

with $\eta_{3,n}$ a.s. convergent. Altogether

$$L_n' = \alpha\Gamma(\alpha) \int_{[2,n]} x^{1-\alpha} \nu(dx) + \eta_{4,n} + o(n^{3/2-\alpha+\varepsilon}),$$

where $\eta_{4,n}$ is a.s. convergent. Finally Lemma 11 gives a.s.

$$
\begin{aligned}
L_n' = {} & \frac{\Gamma(\alpha)\alpha(\alpha-1)}{(2-\alpha)} n^{2-\alpha} - \Gamma(\alpha)\alpha(\alpha-1)^{\alpha} \sum_{k \le n/\gamma} k^{1-\alpha}(V_k - \gamma) \\
& + \eta_n + o_P(n^{1/\alpha^2+1-\alpha+\varepsilon}) + o(n^{3/2-\alpha+\varepsilon})
\end{aligned} \tag{33}
$$

for all $\varepsilon > 0$, where η_n now is convergent in probability.

We are ready to treat the different cases of Theorem 1:

If $1 < \alpha < (1+\sqrt{5})/2$, then we use that $1/\alpha > 1/\alpha^2$ and $1/\alpha > 1/2$. Therefore the three remainder terms in (33) are all of order $o_P(n^{1/\alpha+1-\alpha})$ and thus may be neglected. The result follows from an application of Lemma 10. The case $\alpha = (1+\sqrt{5})/2$ is treated in the same way.

If $\alpha > (1+\sqrt{5})/2$, then $\frac{1}{\alpha^2} + 1 - \alpha < 0$ and $3/2 - \alpha < 0$. Also from Lemma 9 it follows that $\sum_{k \le n/\gamma} k^{1-\alpha}(V_k - \gamma)$ is a.s. convergent. Thus it follows from (33) that $L_n' - \frac{\Gamma(\alpha)\alpha(\alpha-1)}{(2-\alpha)} n^{2-\alpha}$ is convergent in probability. To see that the limit of L_n' (and thus L_n) is nondegenerate, we go back to (32), respectively,

$$
\begin{aligned}
& L_n' - \frac{\Gamma(\alpha)\alpha(\alpha-1)}{(2-\alpha)} n^{2-\alpha} \\
& \quad = \frac{2\alpha\Gamma(\alpha)}{\rho_2}\mu(\{2\})E_2 + \left(\alpha\Gamma(\alpha)\int_{[3,n]} \frac{xE_x}{\rho_x}\mu(dx) - \frac{\Gamma(\alpha)\alpha(\alpha-1)}{(2-\alpha)} n^{2-\alpha}\right).
\end{aligned}
$$

As shown the term in brackets in convergent in probability. Also $\mu(\{2\}) = 1$ with positive probability. Since the exponential variable E_2 is independent from the rest on the right-hand side, the whole limit has to be nondegenerate. This finishes the proof. □

REFERENCES

[1] BERESTYCKI, J., BERESTYCKI, N. and SCHWEINSBERG, J. (2007). Beta-coalescents and continuous stable random trees. *Ann. Probab.* **35** 1835–1887. MR2349577

[2] BERESTYCKI, J., BERESTYCKI, N. and SCHWEINSBERG, J. (2008). Small-time behavior of beta coalescents. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 214–238. MR2446321

[3] BERESTYCKI, N. (2009). *Recent Progress in Coalescent Theory*. *Ensaios Matemáticos* **16**. Sociedade Brasileira de Matemática, Rio de Janeiro. MR2574323

[4] BERTOIN, J. and LE GALL, J.-F. (2006). Stochastic flows associated to coalescent processes. III. Limit theorems. *Illinois J. Math.* **50** 147–181 (electronic). MR2247827

[5] BIRKNER, M., BLATH, J., CAPALDO, M., ETHERIDGE, A., MÖHLE, M., SCHWEINSBERG, J. and WAKOLBINGER, A. (2005). Alpha-stable branching and beta-coalescents. *Electron. J. Probab.* **10** 303–325 (electronic). MR2120246

[6] BOLTHAUSEN, E. and SZNITMAN, A. S. (1998). On Ruelle's probability cascades and an abstract cavity method. *Comm. Math. Phys.* **197** 247–276. MR1652734

[7] BOOM, E. G., BOULDING, J. D. G. and BECKENBACH, A. T. (1994). Mitochondrial dna variation in introduced populations of Pacific oyster, crassostrea gigas, in British Columbia. *Can. J. Fish. Aquat. Sci.* **51** 1608–1614.

[8] DELMAS, J.-F., DHERSIN, J.-S. and SIRI-JEGOUSSE, A. (2008). Asymptotic results on the length of coalescent trees. *Ann. Appl. Probab.* **18** 997–1025. MR2418236

[9] DRMOTA, M., IKSANOV, A., MOEHLE, M. and ROESLER, U. (2007). Asymptotic results concerning the total branch length of the Bolthausen–Sznitman coalescent. *Stochastic Process. Appl.* **117** 1404–1421. MR2353033

[10] ELDON, B. and WAKELEY, J. (2006). Coalescent processes when the distribution of offspring number among individuals is highly skewed. *Genetics* **172** 2621–2633.

[11] FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications, Vol. II*. Wiley, New York. MR0270403

[12] GNEDIN, A. and YAKUBOVICH, Y. (2007). On the number of collisions in Λ-coalescents. *Electron. J. Probab.* **12** 1547–1567 (electronic). MR2365877

[13] KINGMAN, J. F. C. (1982). The coalescent. *Stochastic Process. Appl.* **13** 235–248. MR0671034

[14] MÖHLE, M. (2006). On the number of segregating sites for populations with large family sizes. *Adv. in Appl. Probab.* **38** 750–767. MR2256876

[15] PETROV, V. V. (1975). *Sums of Independent Random Variables*. Springer, New York. MR0388499

[16] PITMAN, J. (1999). Coalescents with multiple collisions. *Ann. Probab.* **27** 1870–1902. MR1742892

[17] SAGITOV, S. (1999). The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.* **36** 1116–1125. MR1742154

[18] SCHWEINSBERG, J. (2000). A necessary and sufficient condition for the Λ-coalescent to come down from infinity. *Electron. Commun. Probab.* **5** 1–11 (electronic). MR1736720

[19] WATTERSON, G. A. (1975). On the number of segregating sites in genetical models without recombination. *Theoret. Population Biology* **7** 256–276. MR0366430

GOETHE UNIVERSITÄT FRANKFURT
BOX 187
D-60054 FRANKFURT/MAIN
GERMANY
E-MAIL: kersting@math.uni-frankfurt.de