LARGE-SCALE BEHAVIOUR AND HYDRODYNAMIC LIMIT OF BETA COALESCENTS

BY LUKE MILLER AND HELMUT H. PITTERS*

University of Oxford and University of California, Berkeley

We quantify the behaviour at large scales of the beta coalescent \( \Pi = \{ \Pi(t), t \geq 0 \} \) with parameters \( a, b > 0 \). Specifically, we study the rescaled block size spectrum of \( \Pi(t) \) and of its restriction \( \Pi_n(t) \) to \( \{1, \ldots, n\} \). Our main result is a Law of Large Numbers type of result if \( \Pi \) comes down from infinity. In the case of Kingman’s coalescent the derivation of this so-called hydrodynamic limit has been known since the work of Smoluchowski [30]. We extend Smoluchowski’s result to beta coalescents and show that if \( \Pi \) comes down from infinity both rescaled spectra

\[
\left( c_1 \Pi(t_{\tau_n}), \ldots, c_n \Pi(t_{\tau_n}) \right), \quad \text{and} \quad \left( c_1 \Pi_n(t_{\tau_n}), \ldots, c_n \Pi_n(t_{\tau_n}) \right),
\]

converge to (different) deterministic limits that we compute explicitly in terms of partial Bell polynomials. Here \( c_i \pi \) counts the number of blocks of size \( i \) in a partition \( \pi \), and \( (\tau_n) \) is a sequence such that \( \tau_n \sim n^{-(1-a)} \) as \( n \to \infty \).

Along the way we study the non-trivial limits of the rescaled block counting processes \( \{ n^\alpha \Pi_n(t_{\tau_n}), t \geq 0 \} \), and \( \{ n^\alpha \Pi(t_{\tau_n}), t \geq 0 \} \), where \( \alpha \in [-1, -2/(3-a)] \), and \( \tau_n \sim n^{\alpha(1-a)} \) if \( \Pi \) comes down from infinity.

1. Introduction. Population geneticists are often interested in understanding the genealogy of randomly sampled individuals for a variety of populations. Usually, one starts by describing the evolution of the population forwards in time, classical models for which are the celebrated Wright-Fisher model as well as the Moran model, and numerous variants incorporating more general offspring distributions (the most general framework being the Cannings models), or incorporating phenomena such as mutation, selection, age structure, or spatial structure. Finding the genealogy that corresponds to a large population whose evolution is specified forwards in time involves some technical machinery from stochastic processes and was first carried out rigorously by Kingman in [13, 15, 14] for a host of population models. Their genealogies turn out to be governed by what is now called Kingman’s coalescent. This coalescent, restricted to a sample of \( n \) individuals, starts with their \( n \) lines of descent. As we trace these ancestral lineages back in time, any pair of lineages

*To whom correspondence should be addressed.

MSC 2010 subject classifications: Primary 60F99; secondary 60K35

Keywords and phrases: coalescent processes, Law of Large Numbers, hydrodynamic limit, Bell polynomials

1
merges at rate 1, but no more than two lineages may merge at any given time. For decades after its discovery this stochastic process has been (and still is) utilized by population biologists to model genealogies. However, (a) because of the relevance of non-neutral populations, i.e. populations with some form of natural selection acting on their individuals, and (b) because of an increasing interest in populations with high fecundity, i.e. populations where single individuals may beget a number of offspring that is on the order of the total population size, their genealogies have been studied and found to be no longer adequately modeled by Kingman’s coalescent. As starting points for more information on this topic the interested reader may consult [29] for developments on populations with selection, and the introduction in [7] for developments on populations with high fecundity. The genealogies of samples drawn from these populations turn out to be governed by so-called multiple merger coalescent processes that were introduced independently by Donnelly and Kurtz [8], Pitman [22] and Sagitov [25]. The multiple merger \( n \)-coalescent process \( \Pi_n = \{\Pi_n(t), t \geq 0\} \) starts with the \( n \) lines of descent as does Kingman’s coalescent. However, unlike Kingman’s coalescent \( \Pi_n \) allows for more than two ancestral lines to merge into a single line. In fact, with positive probability all ancestral lines may merge in a single event.

So far we have described the restriction \( \Pi_n \) of a multiple merger coalescent to a sample of \( n \) individuals. Apparently, we could have restricted ourselves to any sample size \( n \geq 2 \), indicating that there might exist an underlying process \( \Pi = \{\Pi(t), t \geq 0\} \) governing the mergers of an infinite number of ancestral lines indexed by the natural numbers \( \mathbb{N} := \{1, 2, \ldots\} \), such that the restriction of \( \Pi \) to \( \{1, \ldots, n\} \) is a process with the same distribution as \( \Pi_n \). It turns out that such a process \( \Pi \) indeed exists, provided the \( (\Pi_n)_{n \geq 2} \) meet some suitable assumption which seems rather natural from the point of view of sampling. To motivate this assumption, picture a geneticist collecting a (random) sample of size \( n + k \) from a specific population. On his way to the lab the geneticist looses \( k \) data items in his sample. Clearly, the genealogy of the remaining data is governed by \( \Pi_{n+k} \) restricted to \( n \) individuals. However, it seems natural to assume that this genealogy should have been the same (in distribution), had the geneticist only collected a sample of size \( n \) in the first place. More formally, we assume that for any integers \( n \geq 2 \) and \( k \geq 1 \) the process \( \Pi_n \) and the restriction of \( \Pi_{n+k} \) to \( \{1, \ldots, n\} \) have the same distribution. Under this consistency assumption the projective limit \( \Pi \) exists, cf. [22]. Pitman [22] characterized this consistency requirement in terms of finite measures \( \Lambda \) on \([0, 1]\), and we will now use this characterization for a formal definition of \( \Pi \).

For any finite measure \( \Lambda \) on the unit interval there exists a (unique in law) Markov process \( \Pi \) with state space \( \mathcal{P}_\mathbb{N} \), the set of all set partitions of the positive integers \( \mathbb{N} \), such that for each \( n \in \mathbb{N} \) the restriction \( \Pi_n \) of \( \Pi \) to \( [n] := \{1, \ldots, n\} \) is a continuous-time Markov chain with the following dynamics: when \( \Pi_n \) has \( m \) blocks, any \( 2 \leq k \leq m \) specific blocks merge into a single block at rate \( \lambda_{m,k} := \int_0^1 x^{k-2}(1-x)^{m-k}\Lambda(dx) \). The process \( \Pi \) is called a multiple merger coalescent process or \( \Lambda \) coalescent. To each \( \Lambda \) coalescent \( \Pi \) one can assign a corresponding tree, the coalescent tree, in an obvious fashion. This is made precise in Section 2.1.1.
2. Background and main results. We focus on the behaviour of the block size spectrum and the block counting process of beta($a, b$) coalescents $\Pi$ at large scales. One idea is that the evolution of $\Pi$ could be approximated by the evolution of $\Pi_n$ as the sample size $n$ grows to infinity. This is reminiscent of what physicists call a hydrodynamic limit, describing the macroscopic evolution of a system comprised of a large number of particles (usually quantified by ordinary or partial differential equations), and deduced from rules dictating the microscopic stochastic interactions between individual particles.

Since normalizing the finite measure $\Lambda$ corresponds to a linear time change of $\Pi$, in what follows we restrict our attention to probability measures, i.e. $\Lambda([0,1]) = 1$. A $\Lambda$ coalescent $\Pi$ is said to come down from infinity if $\mathbb{P}\{\#\Pi(t) < \infty\} = 1$ for all $t > 0$ and $\Pi$ is said to stay infinite if $\mathbb{P}\{\#\Pi(t) = \infty\} = 1$ for all $t \geq 0$. Pitman [22, Proposition 23] showed the dichotomy that if $\Lambda$ does not charge 1, i.e. $\Lambda(\{1\}) = 0$, then the corresponding coalescent either comes down from infinity or stays infinite. Schweinsberg [27] showed that a $\Lambda$ coalescent that does not charge 1 comes down from infinity if and only if

$$\sum_{n \geq 2} (\gamma_n^{(1)})^{-1} < \infty,$$

where $\gamma_n^{(1)} := \sum_{l=2}^{n} \binom{n}{l} \lambda_{n,l}(l-1)$ is the rate at which the number of blocks decreases. We will see that the asymptotic behaviour of

$$\gamma_n^{(k)} := \sum_{l=2}^{n} \binom{n}{l} \lambda_{n,l}(l-1)^k,$$

as $n \to \infty$ for $k = 1, 2, 3$ plays an important rôle in our analysis of the large scale behaviour of $\Pi$. There is another characterization of the coming down from infinity of $\Pi$. Consider the Laplace exponent

$$\Psi^*(q) := \int_0^1 (e^{-qx} - 1 + qx)/x^2 \Lambda(dx) \quad (q \geq 0)$$

of a spectrally positive Lévy process, which is therefore the branching mechanism of a Continuous State Branching Process $X = \{X(t), t \geq 0\}$, say. Bertoin and Le Gall [5] showed that $\Pi$ comes down from infinity if and only if $X$ becomes extinct in finite time almost surely. According to the so-called Grey’s condition, cf. [11] and [6], $X$ becomes extinct in finite time almost surely if and only if

$$\int_{1}^{\infty} dq q \Psi^*(q) < \infty.$$

2.1. Block size spectrum. Before we turn to our main result on the large-scale limit of the block size spectrum of beta coalescents, we discuss some preliminaries.
2.1.1. Preliminaries. A partition of a set $A$ is a set, $\pi$ say, of nonempty pairwise disjoint subsets of $A$ whose union is $A$. The members of $\pi$ are called the blocks of $\pi$. Let $\# A$ denote the cardinality of $A$ and let $\mathcal{P}_A$ denote the set of all partitions of $A$. We use the convention that empty sums equal 0 and empty products equal 1 throughout.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ denote the probability space underlying a $\Lambda$ coalescent $\Pi$. If $\Pi$ comes down from infinity, we will (as is often done implicitly) identify for any $\omega \in \Omega$ the path $t \mapsto \Pi(t)(\omega)$ with a rooted tree whose leaves are labelled by $\mathbb{N}$. More formally, the set of nodes of the tree corresponding to $t \mapsto \Pi(t)(\omega)$ is 

$$T(\omega) := \{(t, B) : t \geq 0, B \in \Pi(t)(\omega)\}.$$ 

If we interpret $T(\omega)$ as a genealogical tree, $(s, B) \in T(\omega)$ means that individual $B$ is alive at time $s$, and if for two points $(s, B), (t, C) \in T(\omega)$ we have that $s \leq t$ and $B \subseteq C$, then $C$ is interpreted as an ancestor of $B$ alive at time $t$. For any two points $(s, B), (t, C) \in T(\omega)$ let 

$$m((s, B), (t, C))(\omega) := \inf\{u > s \vee t : \text{both } B \text{ and } C \text{ are subsets of a common block in } \Pi(u)(\omega)\}$$ 

denote the time back to the most recent common ancestor of $(s, B)$ and $(t, C)$. It can be shown that $T(\omega)$ together with the metric $d(\omega)$ defined by

$$d((s, B), (t, C))(\omega) := (m((s, B), (t, C))(\omega) - s) + (m((s, B), (t, C))(\omega) - t)$$

is an $\mathbb{R}$-tree. This is done formally in Example 3.41 of [10]. For more information on $\mathbb{R}$-trees, the reader is referred to Evans’ lecture notes [10]. Informally, $d(\omega)$ yields the genealogical distance between any two points in $T(\omega)$.

One may wonder whether the (random) tree $T$ can be described more explicitly. One way to study $T$ is by “exploring” it via subtrees, namely, if we consider any $n$ of its leaves labelled $l_1, \ldots, l_n \in \mathbb{N}$, their spanning tree will correspond to a $\Lambda$ $n$-coalescent with leaves labelled $l_1, \ldots, l_n$, as is apparent from the consistency of $\Lambda$ coalescents. As we increase the sample size $n$, we explore larger and larger subtrees of $T$. However, the topology of a subtree spanned by $n$ leaves is rather involved.

In order to work out explicitly the asymptotic behaviour of a subtree when the number $n$ of leaves grows without bound in what follows, we restrict ourselves to beta coalescents, i.e. we take $\Lambda$ to be the beta($a, b$) distribution with density

$$\Lambda(dx) = B(a, b)^{-1}x^{a-1}(1-x)^{b-1}dx \quad (x \in (0, 1)), \quad (2.4)$$ 

where for $a, b > 0$ the beta function with parameters $a, b$ is defined as $B(a, b) := \int_0^1 x^{a-1}(1-x)^{b-1}dx$. Notice that according to Schweinsberg’s characterization of coalescents that come down from infinity, equation (2.1), the beta($a, b$) coalescent comes down from infinity if and only if $a \in (0, 1)$, cf. Example 15 in [27].
We capture the topology of the coalescent tree by recording at each time $t$ the number of blocks of different sizes. We now give a formal definition of this so-called block size spectrum. For any partition $\pi$ of $[n]$ let $\text{cr} := (\ell_1, \ldots, \ell_n)$ denote the so-called type of $\pi$ defined by $\ell_i := \#\{B \in \pi : \#B = i\}$ for any $i \in \mathbb{N}$. Notice that $\sum_i \ell_i = n$. For $d \in \mathbb{N}$ let the rescaled block size spectrum $(C_{n,i})_{i=1}^d := (C_{n,i}(t), t \geq 0)_{i=1}^d$ be defined by
\begin{equation}
C_{n,i}(t) := n^{-1}c_i \Pi_n(t \tau_n), \quad i \in [d],
\end{equation}
for some sequence $(\tau_n)_{n \geq 2}$ of real numbers that we will specify below.

It is well known that Kingman’s coalescent can be seen as the limit of a Marcus-Lushnikov process, cf. [1], thus its hydrodynamic limit fits into the framework of the so-called Smoluchowski coagulation equations. For $\tau_n \sim n^{-1}$ as $n \to \infty$ the work of Smoluchowski [30] implies that the hydrodynamic limit $(c_1(t), \ldots, c_d(t)) = \lim_{n \to \infty} (C_{n,1}(t), \ldots, C_{n,d}(t))$ is deterministic and given by
\begin{equation}
c_i(t) = c(t)^2(1 - c(t))^{i-1}, \quad c(t) = \frac{2}{2 + t},
\end{equation}
for any two sequences $v_\bullet = (v_k)_{k \in \mathbb{N}}$ and $w_\bullet = (w_k)_{k \in \mathbb{N}}$
\begin{equation}
B_i(v_\bullet, w_\bullet) := \sum_{l=1}^i v_l B_{i,l}(w_\bullet) \quad (i \in \mathbb{N}),
\end{equation}
denotes the $i$th complete Bell polynomial (associated with $(v_\bullet, w_\bullet)$), where
\begin{equation}
B_{i,l}(w_\bullet) := \sum_{\pi \in \mathcal{P}[i]} \prod_{B \in \pi} w_{#B} \quad (1 \leq l \leq i),
\end{equation}
denotes the $(i,l)$th partial Bell polynomial (associated with $w_\bullet$) and $\mathcal{P}[i]$ denotes the set of all partitions of $[i]$ that contain $l$ blocks. Moreover, for $x \in \mathbb{R}$, $k \in \mathbb{N}$ let $x^k := x(x-1) \cdots (x-k+1)$ denote the falling factorial power, $x^{-k} := x(x+1) \cdots (x+k-1)$ the rising factorial power and we agree on $x^0 := x^1 := 1$, and for any function $f : \mathbb{N} \to \mathbb{R}$ we write $f(\bullet)$ as shorthand for the sequence $(f(k))_{k \geq 1}$.

**Theorem 1.** Fix $d \in \mathbb{N}$, and $a \in (0,1)$. If $\tau_n \sim n^{-(1-a)}$ as $n \to \infty$ one has convergence
\begin{equation}
(C_{n,1}(t), \ldots, C_{n,d}(t)) \to (c_1(t), \ldots, c_d(t)),
\end{equation}

\begin{flushright}
imsart-aap ver. 2014/10/16 file: hydrodynamic_limit.tex date: June 9, 2021
\end{flushright}
in $D_{[0,1]^d}([0,\infty))$ in the Skorokhod topology. The latter process is deterministic with initial state $(c_1(0),\ldots,c_d(0)) = (1,0,\ldots,0)$. It can be characterised via the generating function

\begin{equation}
\mathcal{G}(t,x) = \sum_{i \geq 1} c_i(t) x^i = c(t) - \left(1 - x\right)^{a-1} + \frac{\Gamma(a+b)}{(2-a)\Gamma(b)} t^{\frac{1}{1-a}} \quad (x \in (-1,1), t \geq 0),
\end{equation}

and one has

\begin{equation}
c_i(t) = \frac{c(t)^{2-a}}{i!} B_i \left(\left(\frac{1}{1-a}\right)^\bullet (1-a)^\bullet, 1-a\right) \quad (i \in [d], t \geq 0),
\end{equation}

where $B_n(v, w)$ denotes the $n$th Bell polynomial as defined below, and

\begin{equation}
c(t) = \left(1 + \frac{\Gamma(a+b)}{(2-a)\Gamma(b)} t\right)^{\frac{1}{1-a}} \quad (t \geq 0)
\end{equation}

is the limit of the suitably rescaled block counting process $\#\Pi_n(t)$ as derived in Theorem 3 below. By $\delta_{ij}$ we denote Kronecker’s delta which equals one if $i = j$ and zero otherwise.

We now turn to the scaling limit of the block size spectrum of $\Pi$. Define the process $(C_{n,1}^*(t),\ldots,C_{n,d}^*(t), t \geq 0)$ via $C_{n,i}^*(t) := n^{-1} c_i(t \tau_n)$ for $i \in [d]$. In particular, the process $(C_{n,1}^*,\ldots,C_{n,d}^*)$ has initial state $(\infty,0,\ldots,0)$, and state space $(\mathbb{R}_+ \cup \{\infty\})^d$.

**Theorem 2.** Fix $d \in \mathbb{N}$, $a \in (0,1)$. If $\tau_n \sim n^{-(1-a)}$ as $n \to \infty$ one has convergence

\begin{equation}
(C_{n,1}^*(t),\ldots,C_{n,d}^*(t), t \geq 0) \to (c_1^*(t),\ldots,c_d^*(t), t \geq 0),
\end{equation}

as $n \to \infty$ in $D_{[0,\infty]^d}([0,\infty))$ in the Skorokhod topology, where the latter process is deterministic with initial state $(\infty,0,\ldots,0)$,

\begin{equation}
c_i^*(t) = \frac{c^*(t)^{2-a}}{i!} B_i \left(\left(\frac{1}{1-a}\right)^\bullet, 1-a\right) \quad (i \in [d], t \geq 0),
\end{equation}

and

\begin{equation}
c^*(t) = \left(\frac{\Gamma(a+b)}{(2-a)\Gamma(b)} t\right)^{\frac{1}{1-a}} \quad (t \geq 0)
\end{equation}

is the limit of the suitably rescaled block counting process $\#\Pi(t)$ as derived in Theorem 4 below.
2.2. Block counting process. The small-time behaviour of multiple merger coalescents has been studied before, cf. [5, 4, 3, 18, 17, 19]. Berestycki, Berestycki and Schweinsberg [4] showed
\[
\lim_{t \to 0^+} t^{\frac{1}{\Delta - 1}} \# \Pi(t) = \left( \frac{\Delta}{\Gamma(2 - \Delta)} \right)^{\frac{1}{\Delta - 1}} \text{ a.s. (2.13)}
\]
for coalescents such that \( \Lambda(dx) = f(x)dx \), where \( f(x) \sim Ax^{1-\Delta} \) as \( x \to 0^+ \) for some \( \Delta \in (1, 2) \). This yields \( A = 1/\left(\Gamma(\Delta)\Gamma(2 - \Delta)\right) \) if \( \Lambda \) is a beta\((2 - \Delta, \beta)\) distribution for some \( \beta > 0 \). The limiting result in (2.13) was obtained in probability for a larger class of multiple merger coalescents in Bertoin and LeGall [5, Lemma 3]. Berestycki, Berestycki and Limic [3] show a Law of Large Numbers type of result, more specifically, they show convergence
\[
\lim_{t \to 0^+} \frac{\# \Pi(t)}{v^*(t)} = 1 \text{ a.s., (2.14)}
\]
where \( v^* \) is uniquely determined by \( \int_{v^*(t)}^\infty dq/\Psi^*(q) = t \), \( t > 0 \), and \( \Psi^* \) is defined in (2.2).

Any function \( v^* \) satisfying (2.14) is called a speed of coming down from infinity for \( \Pi \). We exclude from our study the special case \( a = 1 \) that includes the Bolthausen-Sznitman coalescent. For work on the scaling limit of the block counting process of the Bolthausen-Sznitman coalescent, corresponding to \( \Lambda \) being the uniform distribution on the unit interval, the reader is referred to [28, 2, 20, 16].

As opposed to the results presented above, in some sense our results are stronger as we study process-valued limits of the rescaled block counting process \( C_n := \{C_n(t), t \geq 0\} \) defined by
\[
C_n(t) := n^\alpha \# \Pi_n(t \tau_n), \quad (2.15)
\]
where \( (\tau_n) \) is a sequence of real numbers such that
\[
\tau_n \sim \begin{cases} n^\alpha(1-a) & \text{if } a < 1, \\ 1 & \text{if } a > 1, \end{cases} \quad (2.16)
\]
as \( n \to \infty \). Here \( \alpha \) is a new parameter. We find a nontrivial scaling limit for \( C_n \) whenever \( \alpha \in [-1, -2/(3-a)] \).

For the sake of brevity we will occasionally use \( \phi \equiv \alpha(a-1) \).

At the same time, our results are weaker than (2.13) and (2.14) in the sense that we show weak convergence of processes, but not almost sure convergence of the marginals.

The study of the second-order asymptotics of the block counting process of multiple merger coalescents (that we do not pursue here) was initiated in [18], and also pursued in [17, 19]. Limic and Talarczyk [18] work with the Laplace exponent \( \Psi: [1, \infty) \to \mathbb{R}_+ \) defined by
\[
\Psi(q) := \int_0^1 \frac{(1 - y)^q - 1 + qy}{y^2} \frac{\Lambda(dy)}{y}
\]
instead of $\Psi^*$, and show that $v: \mathbb{R}_+ \to \mathbb{R}_+$ defined by $\int_{v(t)}^\infty dq/\Psi(q) = t$ is a speed of coming down from infinity for $\Pi$, too. They assume that $\Lambda(\{0\}) = \Lambda(\{1\}) = 0$, and that there exists a $y_0 \leq 1$ such that

$$\Lambda(dy) = g(y)dy, \quad y \in [0, y_0] \quad \text{and} \quad \lim_{y \to 0^+} g(y)y^\beta = A$$

for some $\beta \in (0, 1)$ and $A \in (0, \infty)$. Their main result is the convergence in the Skorokhod space $D_{\mathbb{R}}([0, \infty))$ equipped with the $J_1$ topology as $\varepsilon \to 0$

$$\{X_\varepsilon(t), t \geq 0\} \to \{Z(t), t \geq 0\}, \quad (2.17)$$

where $X_\varepsilon(0) := 0$,

$$X_\varepsilon(t) := \varepsilon^{-1/(1+\beta)} \left( \frac{\#\Pi(\varepsilon t)}{v(\varepsilon t)} - 1 \right) \quad (t > 0), \quad (2.18)$$

and $\{Z(t), t \geq 0\}$ is a $(1 + \beta)$ stable process given by $Z(0) = 0$, and

$$Z(t) = -\frac{K}{t} \int_0^t u\mathcal{M}(du) \quad (t > 0),$$

and an independently scattered $(1 + \beta)$ stable random measure $\mathcal{M}$ on $\mathbb{R}$ with skewness intensity 1, cf. [26, Definition 3.3.1], and some suitably defined positive constant $K$. Our next results concerning the small-time behaviour of $\Pi_n$ (corresponding to $\alpha \in [-1, 0]$) seem to be implicit in [18]. However, our proofs are entirely different, and are based on working with the generators both in the case of the block counting process, as well as in the case of the block size spectrum.

We now turn to the rescaled block counting process $C_n$ of $\Pi_n$. If $\Pi$ does not come down from infinity, the limit of $C_n$ as $n \to \infty$ is neither deterministic nor a diffusion limit. If $\Pi$ comes down from infinity, $C_n$ approaches a deterministic limit, however, only for small times, i.e. only for $\alpha \in [-1, -2/(3 - a))$.

**Theorem 3.** Fix $a < 1$, $\alpha \in [-1, -2/(3 - a))$ and let $\tau_n$ be a sequence such that $\tau_n \sim n^{(1-a)\alpha}$ as $n \to \infty$. Then as $n \to \infty$ we have convergence

$$\{C_n(t), t \geq 0\} \to \{c(t), t \geq 0\}, \quad (2.19)$$

in $D_{[0,1]}([0, \infty))$ in the Skorokhod topology, where $c(t)$ solves the ordinary differential equation of Bernoulli type

$$\frac{d}{dt}c(t) = -\frac{\Gamma(a + b)}{(1-a)(2-a)\Gamma(b)} c(t)^{2-a} \quad (t \geq 0), \quad (2.20)$$
with boundary condition

\begin{equation}
(2.21) \quad c(0) = \begin{cases} 
1 & \text{if } \alpha = -1, \\
\infty & \text{if } \alpha \in (-1, -\frac{2}{3-a}).
\end{cases}
\end{equation}

The solution of (2.20) is given by

\begin{equation}
(2.22) \quad c(t) = \left( \delta_{\alpha,-1} + \frac{\Gamma(a + b)}{(2 - a)\Gamma(b)} t \right)^{\frac{1}{a-1}}
\end{equation}

If \( a > 1 \) then \( C_n \) does not admit a diffusion limit as \( n \to \infty \).

**Remark 1.** We omit the case \( \alpha < -1 \) as it corresponds to the initial condition \( c(0) = 0 \) and trivial solution \( c(t) = 0, t \geq 0 \).

**Remark 2.** (1) The special case \( a = 1/2, \alpha = -1 \) is interesting as it contains the arcsine coalescent which was recently studied in [24]. Using Legendre’s duplication formula \( \Gamma(2z) = 2^{2z-1}\Gamma(z)\Gamma(z + \frac{1}{2})/\sqrt{\pi} \), we find

\begin{equation}
(2.23) \quad c(t) = \left( \frac{3\Gamma(b)^2}{3\Gamma(b)^2 + 4^{1-b}\sqrt{\pi}\Gamma(2b)t} \right)^2 \quad (t \geq 0).
\end{equation}

Consequently, for the arcsine coalescent, that is the beta coalescent with parameters \( a = b = \frac{1}{2} \), we obtain

\[ c(t) = \left( \frac{3\sqrt{\pi}}{3\sqrt{\pi} + 2t} \right)^2 \quad (t \geq 0). \]

(2) In the limiting case \( a \to 0 \) we recover for the rescaled block counting process the well-known hydrodynamic limit

\[ c(t) = \frac{2}{2 + t} \quad (t \geq 0), \]

of Kingman’s coalescent, cf. [32, Equation (2.15)].

Interestingly, provided \( \Pi \) comes down from infinity, i.e. if \( a < 1 \), for \( \alpha \in (-1, -2/(3-a)) \) the limit of \( C_n \) agrees with the limit of

\begin{equation}
(2.24) \quad \{ n^\alpha \#(t\tau_n), t \geq 0 \}
\end{equation}

for \( \tau_n \sim n^{\alpha(1-a)} \) as \( n \to \infty \). For \( \alpha = -1 \) we still obtain a non-trivial scaling limit for (2.15), however, this limit does not agree with the one for (2.24).

We now rescale the process \( \Pi \), namely for each \( n \in \mathbb{N} \) let \( C_n^* = \{ C_n^*(t), t \geq 0 \} \) be defined by \( C_n^*(t) := n^\alpha \#(t\tau_n) \). In particular, notice that the initial state of this process is \( C_n^*(0) = \infty \), irrespective of \( \alpha \).
Theorem 4. As $n \to \infty$ we have for $a < 1$, $\alpha \in [-1, -2/(3-a))$ convergence

\begin{equation}
\{C_n^*(t), t \geq 0\} \to \{c^*(t), t \geq 0\}
\end{equation}

in $D_{[0,\infty]}([0,\infty))$ in the Skorokhod topology, where

\begin{equation}
c^*(t) := \left( \frac{\Gamma(a+b)}{(2-a)\Gamma(b)} \right)^{\frac{1}{1-a}}.
\end{equation}

Notice that Theorem 4 suggests the following scaling invariance of $c^*$, which is easily verified. For any real numbers $\alpha, m$ we have $m^\alpha c^*(tm^\alpha(1-a)) = c^*(t)$.

We now turn to the proofs of our results.

3. Proofs. We first prove the results on the block size spectrum in subsection 3.1. The results on the block counting process are proved in subsection 3.2. Before we turn to the proofs, let us collect some facts and fix some further notation.

Recall that $N_n(t) = \#\Pi_n(t)$ counts the number of blocks in $\Pi_n(t)$. Moreover, recall that the Gamma function is defined as $\Gamma(a) := \int_0^1 e^{-x}x^{a-1}dx$ for any positive real number $a \in (0,\infty)$. We will repeatedly use the identity $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ for $a,b > 0$.

Recall that the process $C_n = \{C_n(t), t \geq 0\}$ is defined by $C_n(t) = n^{\alpha}N_n(t_n)$. It is a continuous-time Markov chain with state space $E_n := n^\alpha[1]$, initial state $C_n(0) = n^{1+\alpha}$, absorbing state $n^\alpha$ and it evolves according to the following dynamics:

\begin{equation}
a transition \quad c \mapsto c - n^\alpha(l-1) \quad occurs \ at \ rate \ \left( \frac{n^\alpha c}{l} \right)\lambda_{n^\alpha,c,l} \quad (2 \leq l \leq n^{-\alpha}c).
\end{equation}

For $2 \leq k \leq m$ if there are currently $m$ blocks in $\Pi$, we will see any $k$ specific blocks merge at rate

\begin{equation}
\lambda_{m,k} := \int_0^1 x^{k-2}(1-x)^{m-k}A(dx) = \frac{B(k-2+a,m-k+b)}{B(a,b)} = \frac{a^{k-2}b^{m-k}}{(a+b)^{m-k}}.
\end{equation}

When there are $n$ blocks and $a \notin \{1,2\}$, the total rate to see a merger is given by

\begin{equation}
\gamma_n = \frac{1}{(1-a)(2-a)} \left( (a+b-1)(a+b-2) + b((1-a)n-b+1)(1+b)^{n-2} \right),
\end{equation}

as we will see in Lemma 4.

Remark 3. Equation (3.2) yields the recursive formula

\begin{equation}
\lambda_{m,k+1} = \frac{a+k-2}{b+m-k-1}\lambda_{m,k} \quad (2 \leq k \leq m-1).
\end{equation}
This should be compared to the recursive formula
\[ \lambda_{m,k} = \lambda_{m+1,k} + \lambda_{m+1,k+1} \quad (2 \leq k \leq m) \]
for arbitrary \( \Lambda \) given by Pitman in [22], Lemma 18. Combined, these formulae yield
\[ \lambda_{m+1,k} = \frac{b + m - k}{a + b + m - 2} \lambda_{m,k} \quad (2 \leq k \leq m), \]
and can be used to efficiently compute the rates of the beta\((a, b)\) coalescent by an algorithm.

We collect some Lemmas that we need for our proofs.

**Lemma 1.** For \( a, b > 0 \), a natural number \( n \in \mathbb{N} \) and an integer \( z \in \mathbb{Z} \) we have
\begin{equation}
\frac{a^n}{b^{n+z}} \sim \frac{\Gamma(b)}{\Gamma(a)} n^{a-b-z},
\end{equation}
as \( n \to \infty \).

**Proof.** We calculate
\[ \frac{a^n}{b^{n+z}} = \frac{\Gamma(a+n)}{\Gamma(a)} \frac{\Gamma(b)}{\Gamma(b+n+z)} \sim \frac{\Gamma(b)}{\Gamma(a)} n^{a-b-z}, \]
as \( n \to \infty \). \( \square \)

We now need some notation. For \( d \in \mathbb{N} \), \( x \in \mathbb{R}^d \) and \( k \in \mathbb{N}_0^d \) let \( x^k := \prod_{i=1}^d x_i^{k_i} \) and \( |x| = \sum_{i=1}^d |x_i| \).

**Lemma 2.** Fix \( d \in \mathbb{N} \) and \( k, n \in \mathbb{N}_0^d \). Then for \( a, b \in \mathbb{R} \)
\begin{equation}
\sum_{l \in \mathbb{N}_0^d} \binom{n_k}{l} a^{n_l} b^{n-|l|} \prod_{i=1}^d \binom{n_i}{k_i} = a^{n_k} \sum_{|k|} n_k (a + b + |k|)^{|n|-|k|}.
\end{equation}

**Proof.** We first prove the statement for \( k = 0 \) by an induction on \( d \). Hence, for \( d = 1 \) the statement reads
\[ \sum_{l=0}^n \binom{n}{l} a^l b^{n-l} = (a + b)^n, \]
and this is true, since the sequence \((a^k)_{k \geq 1}\) of rising factorial powers is a sequence of polynomials of binomial type, as is well known. Suppose now that (3.6) holds for some
d \in \mathbb{N}. Then
\begin{align*}
\sum_{l \in \mathbb{N}_{0}^{d+1}} a^{l} b^{n_{l}-|l|} \prod_{i=1}^{d+1} \binom{n_{i}}{l_{i}} &= \sum_{l \in \mathbb{N}_{0}^{d}} a^{l} b^{n_{l}} \prod_{i=1}^{d} \binom{n_{i}}{l_{i}} \sum_{l_{d+1}=0}^{d+1} \binom{n_{d+1}}{l_{d+1}} (a + |l|)^{l_{d+1}} (b + n_{1} + \ldots + n_{d} - |l|)^{l_{d+1}} \\
&= (a + b)^{n_{1} + \ldots + n_{d}} (a + b + n_{1} + \ldots + n_{d})^{\sum_{l_{d+1}=0}^{d+1}} = (a + b)^{|n|},
\end{align*}
where we used the induction hypothesis in the second equality. Now suppose |k| > 0. Performing the index shift \( m = l - k \) in the first step and applying the statement for \( k = 0 \) in the last step, we obtain
\begin{align*}
\sum_{l \in \mathbb{N}_{0}^{d}} l^{k} a^{l} b^{n_{l}-|l|} \prod_{i=1}^{d} \binom{n_{i}}{l_{i}} &= \sum_{m \in \mathbb{N}_{0}^{d}} (m + k)^{\sum_{i=1}^{d} \binom{n_{i}}{m_{i} + k_{i}}} \prod_{i=1}^{d} \binom{n_{i}}{m_{i} + k_{i}} \\
&= a^{\|k\|} \sum_{m \in \mathbb{N}_{0}^{d}} (a + |k|) \sum_{m \in \mathbb{N}_{0}^{d}} (a + |k|) \prod_{i=1}^{d} \binom{n_{i} - k_{i}}{m_{i}} \\
&= a^{\|k\|} n^{k} (a + b + |k|)^{|n| - |k|},
\end{align*}
which completes the proof. \(\square\)

3.1. Block size spectrum. For \( d \in \mathbb{N} \) recall the rescaled block size spectrum \((C_{n,i})_{i=1}^{d+1} := (C_{n,i}(t), t \geq 0)_{i=1}^{d+1}\) defined by
\begin{align*}
C_{n,i}(t) := n^{-1} c_{i} \Pi_{n}(t \tau_{n}), i \in [d], & \quad C_{n,d+1}(t) := n^{-1} \sum_{i=d+1}^{n} c_{i} \Pi_{n}(t \tau_{n}).
\end{align*}
For \( l \in \mathbb{N}_{0}^{d+1} \) with \(|l| > 1\) we say that an \( l \)-merger occurs in \( P_{n} \) if among the merging blocks there are \( l_{1} \) singletons, \( l_{2} \) blocks of size 2, ..., \( l_{d} \) blocks of size \( d \) and \( l_{d+1} \) blocks of size at least \( d + 1 \). The process \((C_{n,i})_{i=1}^{d+1}\) has state space \( E_{n,d+1}^{d} := n^{-1} \{0, \ldots, n\}^{d+1} \setminus \{0\} \), initial state \((1, 0, \ldots, 0)\), absorbing state \((0, 0, \ldots, n^{-1})\) and evolves according to the following dynamics:
\begin{align*}
\text{a transition } c \mapsto c & \quad \text{occurs at rate } \lambda_{n,c,|l|} \prod_{i=1}^{d+1} \binom{nc_{i}}{l_{i}}.
\end{align*}
if $c \in E_n^d$ and $l_i \leq c_i$ for all $i \in \{d+1\}$, where $\|l\| := \sum_{i=1}^{d+1} il_i$ and $e_i = (\delta_{ij})_{j=1}^{d+1}$ denotes the $i$th unit vector in $\mathbb{R}^{d+1}$.

Let $\partial_i = \frac{\partial}{\partial x_i}$ denote the $i$th partial derivative.

**Proposition 1.** Fix $d \in \mathbb{N}$. For a sequence $(\tau_n)$ of order $n^{-(1-a)}$ and $a < 1$ we have convergence

$$(C_{n,1}(t), \ldots, C_{n,d+1}(t)) \rightarrow (c_1(t), \ldots, c_{d+1}(t)),$$

in $D_{[0,1]^{d+1}}([0, \infty))$ in the Skorokhod topology, where the latter process is deterministic with initial state $(c_1(0), \ldots, c_{d+1}(0)) = (1, 0, \ldots, 0)$ and generator

$$Gf(c) := \frac{\Gamma(a+b)}{\Gamma(b)} \sum_{i=1}^{d} \left( -c_i |c|^{1-a} + \sum_{m=2}^{i} \frac{a^{m-2}|c|^{2-a-m}}{2-a} \sum_{l \in \mathbb{N}_0^d \mid \|l\|=m, \|l\|=i} \prod_{k=1}^{d} c_k \right) \partial_i f(c)$$

$$+ \frac{\Gamma(a+b)}{\Gamma(b)} \left( -\frac{|c|^{2-a}}{2-a} + \sum_{r=1}^{d+1} \sum_{m=2}^{r} \frac{a^{m-2}}{2-a} \sum_{l \in \mathbb{N}_0^d \mid \|l\|=m, \|l\|=r} \prod_{k=1}^{d+1} c_k \right) \partial_{d+1} f(c).$$

**Proof.** For a function $f : \mathbb{R}^{d+1} \rightarrow \mathbb{R}$ and a vector $\kappa \in \mathbb{N}_0^{d+1}$ let $D^\kappa := \partial_1^{\kappa_1} \cdots \partial_{d+1}^{\kappa_{d+1}}$ and $f^{(\kappa)}(x) := \partial_1^{\kappa_1} \cdots \partial_{d+1}^{\kappa_{d+1}} f(x)$. Moreover, let $\kappa! := \prod_{i=1}^{d+1} \kappa_i!$ and for any vector $x \in \mathbb{R}^{d+1}$ let $x^\kappa := \prod_{i=1}^{d+1} x_i^{\kappa_i}$. Letting $\lambda_n(c)$ denote the total rate of $(C_{n,i})_{i=1}^{n}$ in state $c \in E_n^d$. Using a Taylor expansion we obtain for the generator $G_n$ of $(C_{n,i})_{i=1}^{d+1}$

$$G_n f(c) := \lambda_n(c) \int (f(c') - f(c)) \mu(c, dc')$$

$$= \tau_n \lambda_n(c) \sum_{l \in \mathbb{N}_0^{d+1}} \frac{(f(c - (l - e_{\|l\|\wedge(d+1)})/n) - f(c)) \lambda_n(c, l)}{\lambda_n(c)} \prod_{i=1}^{d+1} \left( \frac{nc_i}{l_i} \right)$$

$$= \tau_n \sum_{l \in \mathbb{N}_0^{d+1}, |l|>1} \sum_{\kappa \in \mathbb{N}_0^{d+1}, |\kappa|=1} \frac{(l - e_{\|l\|\wedge(d+1)})^\kappa}{n} D^\kappa f(c)$$

$$+ \sum_{\kappa \in \mathbb{N}_0^{d+1}, |\kappa|=2} \frac{(l - e_{\|l\|\wedge(d+1)})^\kappa}{n^2 \kappa!} D^\kappa f \left( c - \partial_n l - e_{\|l\|\wedge(d+1)} \right) \prod_{i=1}^{d+1} \left( \frac{nc_i}{l_i} \right),$$

for any $f \in C^\infty([0, \infty)^{d+1}), c \in E_n^d$ and for some $\partial_n, l \in [0, 1]^{d+1}$. Let $\|f\| := \sup_{x \in [0,1]^{d+1}} |f(x)|$.  

**imsart-aap ver. 2014/10/16 file: hydrodynamic_limit.tex date: June 9, 2021**
Part 1. Let us first consider summands corresponding to $|\kappa| = 2$ by focusing on

$$T^{(2)}(n) := \frac{\tau_n}{n^2} \sum_{l \in \mathbb{N}_0^{d+1}} \sum_{\kappa \in \mathbb{N}_0^{d+1}} \frac{(l - e^{|\kappa|}(d+1))^{\kappa}}{\kappa!} D_{\kappa}^f \left( c - \vartheta_{n,l} \frac{l - e^{|\kappa|}(d+1)}{n} \right)$$

$$\times \lambda_{n,|\kappa|,l} \prod_{i=1}^{d+1} \left( nc_i \right),$$

Evidently, in this case there exist (possibly equal) $i, j \in [d + 1]$ with $\kappa = e_i + e_j$. Notice that

$$|(l - e^{|\kappa|}(d+1))^{\kappa}| \leq l^\kappa \begin{cases} \frac{l_i^2}{1} & \text{if } \kappa = 2e_i \text{ for some } i \in [d + 1], \\ l_i l_j & \text{if } \kappa = e_i + e_j \text{ for some } i, j \in [d + 1], i \neq j. \end{cases}$$

Hence, for fixed $i \in [d + 1]$ the summand in $T^2(n)$ corresponding to $\kappa = 2e_i$ is bounded by

$$\tau_n \sum_{l \in \mathbb{N}_0^{d+1}} \frac{1}{2n^2} (l - e^{|\kappa|}(d+1))^{2e_i} D_{2e_i}^f \left( c - \vartheta_{n,l} \frac{l - e^{|\kappa|}(d+1)}{n} \right) \lambda_{n,|\kappa|,l} \prod_{k=1}^{d+1} \left( nc_k \right)$$

$$\leq \|f^{(2e_i)}\|_\infty \frac{\tau_n}{n^2} \sum_{l \in \mathbb{N}_0^{d+1}} l_i^2 \lambda_{n,|\kappa|,l} \prod_{k=1}^{d+1} \left( nc_k \right)$$

$$= \|f^{(2e_i)}\|_\infty \frac{\tau_n}{(a + b)^{|\kappa|}-2} \frac{a^{|\kappa|}-2}{n^2} \sum_{l \in \mathbb{N}_0^{d+1}} l_i^2 a^{|\kappa|-|l|} \prod_{k=1}^{d+1} \left( nc_k \right),$$

(3.12) $$= \|f^{(2e_i)}\|_\infty \frac{\tau_n}{(1 - a)(2 - a)} \frac{a^{|\kappa|}-2}{n^2} \sum_{l \in \mathbb{N}_0^{d+1}} l_i^2 (a - 2)^{|\kappa|-|l|} \prod_{k=1}^{d+1} \left( nc_k \right).$$

Writing $l_i^2 = l_i^2 + l_i$ and applying Lemma 2 and Lemma 1, we find

$$\frac{\tau_n}{(a + b)^{|\kappa|}-2} \frac{a^{|\kappa|}-2}{n^2} \sum_{l \in \mathbb{N}_0^{d+1}} l_i^2 (a - 2)^{|\kappa|-|l|} \prod_{k=1}^{d+1} \left( nc_k \right)$$

$$= \frac{\tau_n}{(a + b)^{|\kappa|}-2} \frac{(1 - a)(2 - a) nc_i (nc_i - 1)(a + b)^{|\kappa|-2}}{(1 - a)(a - 1) c_i^2 n a - 1} \to 0,$$
as $n \to \infty$, and

$$
\frac{\tau_n}{(a+b)^{|n|c|-2}n^2} \sum_{l \in \mathbb{N}^{d+1}_{0} \setminus \{0\}} l_i (a-2)^{|l|} b^{|n|c|-|l|} \prod_{k=1}^{d+1} \left( \frac{nc_k}{l_k} \right)
$$

(3.13)

$$
= (a-2) \frac{\tau_n}{(a+b)^{|n|c|-2}n^2} \left( nc_i (a+b-1)^{|n|c|-1} - b^{|n|c|-1} nc_i \right)
$$

$$
\sim (a-2) \Gamma(a+b) \left( \frac{c_i n^{a-2}}{\Gamma(a+b-1)} - \frac{|c|^{1-a} n^{-2}}{\Gamma(b)} \right) \to 0,
$$

as $n \to \infty$.

For fixed $i,j \in [d+1]$ such that $i \neq j$, by applying Lemma 1 and Lemma 2 (with $k = e_i + e_j$), we obtain for the summand in $T^2(n)$ corresponding to $\kappa = e_i + e_j$ the bound

$$
\frac{\tau_n}{n^2} \sum_{l \in \mathbb{N}^{d+1}_{0} \setminus \{0\}} l_i l_j D^{e_i+e_j} f \left( c - \vartheta_{n,l} \frac{l - e_i \vartheta_{n,l} (d+1)}{n} \right) \lambda_{n|c|,|l|} \prod_{k=1}^{d+1} \left( \frac{nc_k}{l_k} \right)
$$

$$
\leq \| f^{(e_i+e_j)} \|_{\infty} \frac{\tau_n}{(1-a)(2-a)} \frac{\tau_n}{n^2 (a+b)^{|n|c|-2}} \sum_{l \in \mathbb{N}^{d+1}_{0} \setminus \{0\}} l_i l_j (a-2)^{|l|} b^{|n|c|-|l|} \prod_{k=1}^{d+1} \left( \frac{nc_k}{l_k} \right)
$$

$$
\sim \| f^{(e_i+e_j)} \|_{\infty} (c_i c_j) \frac{\tau_n}{(1-a)(2-a)} \frac{\tau_n}{n^2 (a+b)^{|n|c|-2}} (a-2)(a-1)(c_i c_j n)^2 (a+b)^{|n|c|-2}
$$

$$
= \| f^{(e_i+e_j)} \|_{\infty} (c_i c_j)^2 n^{a-1} \to 0,
$$

as $n \to \infty$. By the triangle inequality it follows that $|T^2(n)|$ vanishes as $n \to \infty$.

**Part 2.** We now focus on $|\kappa| = 1$, i.e. we consider

$$
T^{(1)}(n) := -\frac{\tau_n}{n^2} \sum_{l \in \mathbb{N}^{d+1}_{0} \setminus \{0\}} \sum_{\kappa \in \mathbb{N}^{d+1}_{0} \setminus \{0\}} (l - e_i \vartheta_{n,l} (d+1))^\kappa D^\kappa f(c) \lambda_{n|c|,|l|} \prod_{k=1}^{d+1} \left( \frac{nc_k}{l_k} \right)
$$

$$
= -\frac{\tau_n}{n(a+b)^{|n|c|-2}} \sum_{i=1}^{d+1} \partial_i f(c) \sum_{l \in \mathbb{N}^{d+1}_{0} \setminus \{0\}} (l_i - 1_{\{i\} \setminus \{|l|\}(d+1)}) a^{|l| - 2 b^{|n|c|-|l|}} \prod_{k=1}^{d+1} \left( \frac{nc_k}{l_k} \right).
$$

Now consider in $T^{(1)}(n)$ the summands corresponding to a fixed $i \in [d+1]$. We partition these summands and analyse their asymptotics separately as follows. Firstly, by Lemma 2
(with $k = e_i$) we have that

$$O_i(n) := \frac{\tau_n}{n(a + b)^{\lvert n \rvert - 2}} \sum_{l \in \mathbb{N}_0^{d+1}} l_i a^{\lvert l \rvert - 2} b^{\lvert n \rvert - \lvert l \rvert} \prod_{k=1}^{d+1} (n c_k)$$

$$= \frac{1}{(a - 1)(a - 2)} \frac{\tau_n}{n(a + b)^{\lvert n \rvert - 2}} \sum_{l \in \mathbb{N}_0^{d+1}} l_i (a - 2)^{\lvert l \rvert} b^{\lvert n \rvert - \lvert l \rvert} \prod_{k=1}^{d+1} (n c_k)$$

\begin{equation}
3.14 \quad = \frac{1}{(a - 1)(a - 2)} \frac{\tau_n}{n(a + b)^{\lvert n \rvert - 2}} \left( (a - 2) n c_i (a + b - 1)^{\lvert n \rvert - 1} - (a - 2) b^{\lvert n \rvert - 1} n c_i \right)
\end{equation}

$$= \frac{c_i}{(a - 1)(a + b)^{\lvert n \rvert - 2}} \left( (a + b - 1)^{\lvert n \rvert - 1} - b^{\lvert n \rvert - 1} - b^{\lvert n \rvert - 1} n c_i \right) \sim \frac{\Gamma(a + b)}{a - 1} \frac{1}{c_i n^{a - 1}}$$

as $n \to \infty$. Secondly, for any fixed $i \in [d + 1]$ set

$$I_i(n) := -\frac{\tau_n}{n(a + b)^{\lvert n \rvert - 2}} \sum_{l \in \mathbb{N}_0^{d+1}} 1_{\{l = [l] \wedge (d+1)\}} a^{\lvert l \rvert - 2} b^{\lvert n \rvert - \lvert l \rvert} \prod_{k=1}^{d+1} (n c_k)$$

For $i \in [d]$ the summand corresponding to the indicator $1_{\{i = [l] \wedge (d+1)\}}$ has asymptotic
behaviour

\[ I_1(n) := -\frac{\tau_n}{n(a + b)^n|c|^{-2}} \sum_{\ell \in \mathbb{N}_0^{d+1}} 1_{|\ell| = |l|} a^{1|l| - 2d} b^{n|c| - |l|} \prod_{k=1}^{d+1} (nc_k) \]

\[ = -\frac{\tau_n}{n(a + b)^n|c|^{-2}} \sum_{\ell \in \mathbb{N}_0^{d+1}} a^{1|l| - 2d} b^{n|c| - |l|} \prod_{k=1}^{d+1} (nc_k) \]

\[ = -\frac{\tau_n}{n(a + b)^n|c|^{-2}} \sum_{m=2}^{n|c| \wedge d+1} a^{m-2} b^{n|c| - m} \sum_{\ell \in \mathbb{N}_0^{d+1}} \prod_{k=1}^{d+1} (nc_k) \]

\[ \sim -\frac{\Gamma(a + b)}{\Gamma(b)} n^{a-2} \sum_{m=2}^{\frac{i}{d+1}} a^{m-2} b^{n|c| - m} \sum_{\ell \in \mathbb{N}_0^{d+1}} \prod_{k=1}^{d+1} \frac{l_k^{c_k}}{k!} \]

as \( n \to \infty \).

However, \( I_{d+1}(n) \) must be treated as a special case. Using the set equality

\[ \{ \ell \in \mathbb{N}_0^{d+1} : |\ell| > 1, |\ell| \geq d + 1 \} = \mathbb{N}_0^{d+1} \setminus \left( \{ \ell \in \mathbb{N}_0^{d+1} : |\ell| \geq 2, |\ell| \in [d] \} \cup \{ \ell \in \mathbb{N}_0^{d+1} : 0 \leq |\ell| \leq 1 \} \right), \]

it follows that

\[ I_{d+1}(n) := \frac{\tau_n}{(1 - a)(2 - a)n(a + b)^n|c|^{-2}} \left( \sum_{\ell \in \mathbb{N}_0^{d+1}} (a - 2)^{1|\ell| - d} b^{n|c| - |\ell|} \prod_{k=1}^{d+1} (nc_k) \right) \]

\[ - \sum_{r=1}^{d} \sum_{m=2}^{r} (a - 2)^{m} b^{n|c| - m} \sum_{\ell \in \mathbb{N}_0^{d+1}} \prod_{k=1}^{d+1} (nc_k) \]

\[ - \sum_{\ell \in \mathbb{N}_0^{d+1}} (a - 2)^{d+1} b^{n|c| - d} \prod_{k=1}^{d+1} (nc_k) - \frac{b^{n|c|}}{1}. \]
We now study each of the summands in $I_{d+1}(n)$. Using Lemma 2 (with $k = 0$) we find that the first summand

$$\frac{\tau_n}{(1 - a)(2 - a)n(a + b)^{|c| - 2}} \sum_{l \in \mathbb{N}_{d+1}} (a - 2)^{|l|} b^{|n| c - l} \prod_{k=1}^{d+1} \binom{nc_k}{l_k}$$

$$= \frac{\tau_n}{(1 - a)(2 - a)n(a + b)^{|c| - 2}} (a + b - 2)^{|n| c}$$

$$\sim \frac{\Gamma(a + b)}{(1 - a)(2 - a)\Gamma(a + b - 2)} n^{-a - 2}$$

vanishes as $n \to \infty$. For the second summand, we find

$$-\frac{\tau_n}{(1 - a)(2 - a)n(a + b)^{|c| - 2}} \sum_{r=1}^{d} \sum_{m=2}^{r} (a - 2)^{|m|} b^{|n| c - m} \sum_{l \in \mathbb{N}_{d+1}} \prod_{|l|=m, |l|=r}^{d+1} \binom{nc_k}{l_k}$$

$$\sim -\frac{\Gamma(a + b)}{\Gamma(b)} |c|^{2-a-m} \sum_{r=1}^{d} \sum_{m=2}^{r} a^{-m-2} \sum_{l \in \mathbb{N}_{d+1}} \prod_{|l|=m, |l|=r}^{d+1} \frac{l_k}{l_k!}$$

as $n \to \infty$.

Using

$$\sum_{l \in \mathbb{N}_{d+1}} (a - 2)^{|l|} b^{|n| c - l} \prod_{|l|=1}^{d+1} \binom{nc_k}{l_k} = \sum_{m=1}^{d+1} (a - 2) b^{|n| c - l} nc_m = (a - 2) b^{|n| c - l} n |c|,$$

as $n \to \infty$ we find for the final summand

$$-\frac{\tau_n}{(1 - a)(2 - a)n(a + b)^{|c| - 2}} \sum_{l \in \mathbb{N}_{d+1}} (a - 2)^{|l|} b^{|n| c - l} \prod_{k=1}^{d+1} \binom{nc_k}{l_k} - b^{|n| c} \sim \frac{\Gamma(a + b)}{(1 - a)\Gamma(b)} |c|^{2-a}$$

as $n \to \infty$.

Since, by definition, $T^{(1)}(n) = -\sum_{i=1}^{d+1} (O_i(n) + I_i(n)) \partial_i f(c)$, we obtain the convergence

$$(3.15) \quad \lim_{n \to \infty} \sup_{c \in E^n_d} |G_n f(c) - G f(c)| = 0,$$

for all $f \in C^*_c([0, \infty)^{d+1})$.

Since $G$ operates on real functions defined on the bounded domain $E^d := [0,1]^{d+1}$ whose boundary is not smooth, a direct analysis of the corresponding semigroup, respectively
process, as done in the one-dimensional case in the proof of Theorem 3, is nontrivial, cf. [31]. Instead, we proceed via the theory of martingale problems. We say that a process is a martingale, if it is a martingale with respect to its natural filtration.

Once we show that the function \( f \) is integrable, and so \( (c_i)_{i=1}^{d+1} \) is a martingale for each \( n \geq 2 \), converges in distribution to \( (c_i)_{i=1}^{d+1} \), then for each \( f \in C^2_c(E^d) \)

\[
\int_0^t \mathcal{G}_n f((C_{n,i}(s))_{i=1}^{d+1}) ds
\]

is a martingale by the continuous mapping theorem (cf. Corollary 1.9 of Chapter 3 in [9]) and Problem 7 of Chapter 7 in [9], since \( (c_i)_{i=1}^{d+1} \) is bounded by 1 and \( M_n(t) \) is uniformly integrable, and so \( (c_i)_{i=1}^{d+1} \) is a solution of the martingale problem for \( \{ (f, \mathcal{G}_f): f \in C^2_c(E) \} \).

Once we show that the function \( b = (b_i)_{i=1}^{d+1} \) from \([0, \infty) \times \mathbb{R}^{d+1} \) to \( \mathbb{R}^{d+1} \), defined for \( t \geq 0, c \in \mathbb{R}^{d+1} \) by

\[
b_i(t, c) := b_i(c) := \frac{\Gamma(a + b)}{\Gamma(b)} \begin{cases}
-c|c| \frac{1}{1-a} + \sum_{m=2}^{i} a^{m-2} |c|^{2-a-m} \sum_{l \in \mathbb{N}^{d}} |l|=m,|l|=i \prod_{k=1}^{d} \frac{c_k}{l_k!} & \text{if } c \in [0, 1]^{d+1}, i \in [d], \\
-c|c|^{2-a} \frac{2}{2-a} + \sum_{r=1}^{d+1} a^{m-2} \sum_{l \in \mathbb{N}^{d}} |l|=m,|l|=r \prod_{k=1}^{d+1} \frac{c_k}{l_k!} & \text{if } c \in [0, 1]^{d+1}, i = d + 1, \\
0 & \text{otherwise}
\end{cases}
\]

satisfies the conditions of Theorem 3.10 of Chapter 5 in [9], then Theorem 2.6 of Chapter 8 implies that the martingale problem for \( \{ (f, \mathcal{G}_f): f \in C^2_c(E) \} \) is well-posed. This is indeed the case, since

\[
cb(c) = \sum_{i=1}^{d+1} c_i b_i(c),
\]

and moreover, using \( c_i \leq |c| \leq 1 \)

\[
\sum_{i=1}^{d} c_i b_i(c) \leq K \sum_{i=1}^{d} c_i |c|^{2-a} K_{d,i} \leq \hat{K}_d |c|^{3-a} \leq \hat{K}_d |c|^2
\]

and

\[
c_{d+1} b_{d+1}(c) \leq |c| \sum_{r=1}^{d+1} \sum_{m=2}^{r} |c|^m K_{d,r,m} \leq K_d |c|^3 \leq K_d |c|^2,
\]
where the $K, K_{d,i}, K_d, \tilde{K}_d, K_{d,r,m}$ denote suitable constants. It is now straightforward to verify the conditions of Corollary 8.16 of Chapter 4 in [9] which implies the convergence in the statement.

Proposition 1 implies that for each $i \in \mathbb{N}$ $c_i(t)$ is a solution of the ODE

$$c'_i(s) - \frac{\Gamma(a + b)}{(a - 1)\Gamma(b)} c_i(s)c(s)^{1-a} = \frac{\Gamma(a + b)}{\Gamma(b)} \sum_{m=2}^{i} \frac{a^{m-2} c(s)^{2-a-m}}{\prod_{l \in \mathbb{N}_0^d \mid |l| = m, ||l|| = i} l_k!},$$

or

$$c'_i(s) = \frac{G}{i!} \sum_{m=1}^{i} (a - 2)^m c(t)^{2-a-m} \sum_{l \in \mathbb{N}_0^d \mid |l| = m, ||l|| = i} \frac{i!}{\prod_{k=1}^{i} (k!)^{l_k}} \prod_{k=1}^{i} (k!c_k(t))^{l_k}$$

$$= \frac{G}{i!} \sum_{m=1}^{i} (a - 2)^m c(t)^{2-a-m} B_{i,m}(v_\bullet, w_\bullet)$$

$$= \frac{G}{i!} c(t)^{2-a} B_i(v_\bullet, w_\bullet),$$

where

$$G := \frac{\Gamma(a + b)}{(1-a)(2-a)\Gamma(b)}$$

and $v_\bullet = (v_k)$, $w_\bullet = (w_k)$ are sequences defined by $v_k := (a - 2)^k c(t)^{-k}$ and $w_k := k!c_k(t)$.

Consider now the generating function

$$\mathcal{G}(t, x) := \sum_{i \geq 1} c_i(t)x^i \quad (x \in [-1, 1], t \geq 0).$$

We write $\partial_t$ for the partial derivative $\partial / \partial t$ with respect to $t$.

**Lemma 3.** The generating function $\mathcal{G}$ solves the partial differential equation

$$\partial_t \mathcal{G}(t, x) = -\frac{\Gamma(a + b)}{(1-a)(2-a)\Gamma(b)} ((c(t) - \mathcal{G}(t, x))^{2-a} - c(t)^{2-a}) \quad (x \in (-1, 1), t \geq 0)$$

with boundary condition $\mathcal{G}(0, x) = x$ for $x \in [-1, 1]$. 
PROOF. We use the well known fact, cf. [23, Equation (1.11)], that for any two sequences \((v_k), (w_k)\), the exponential generating function of the associated complete Bell polynomials \((B_k(v_\bullet, w_\bullet))\) is given by

\[
\sum_{k \geq 1} B_k(v_\bullet, w_\bullet) \frac{x^k}{k!} = v(w(x)),
\]

where these quantities are defined, and \(v\), respectively \(w\), denotes the exponential generating function of \((v_k)\), respectively \((w_k)\), i.e. \(v(\theta) := \sum_{k \geq 1} v_k \theta^k / k!\), \(w(x) := \sum_{k \geq 1} w_k x^k / k!\). For our particular choice of \((v_k)\) and \((w_k)\), we find

\[
v(\theta) := \sum_{j \geq 1} v_j \frac{\theta^j}{j!} = \sum_{j \geq 1} (a - 2)^j \frac{\theta^j}{j! c(t)^j} = (1 - \frac{\theta}{c(t)})^{2-a} - 1,
\]

\[
w(x) := \sum_{k \geq 1} w_k \frac{x^k}{k!} = \sum_{k \geq 1} c_k(t) x^k = \mathcal{G}(t, x),
\]

for \(|\theta| < c(t), x \in [-1, 1]\). Noticing that \(|\mathcal{G}(t, x)| < c(t)\) for \(|x| < 1\) we obtain

\[
\partial_t \mathcal{G}(t, x) = \sum_{i \geq 1} c_i'(t) x^i = Gc(t)^{2-a} \sum_{j \geq 1} B_j((a - 2)^j c(t)^{a-\bullet}) \frac{x^j}{j! c(t)^j} = Gc(t)^{2-a} v(\mathcal{G}(t, x)) = G((c(t) - \mathcal{G}(t, x))^{2-a} - c(t)^{2-a}).
\]

\[
\square
\]

PROOF. (of (2.8) in Theorem 1) First, for fixed \(x \in (0, 1)\) consider the transformation

\[
g(t, x) := c(t) - \mathcal{G}(t, x) \quad (t \geq 0).
\]

It is straightforward to verify that \(g\) solves the Bernoulli differential equation

\[
\partial_t g(t, x) = -\frac{\Gamma(a + b)}{(1 - a)(2 - a)\Gamma(b)} g(t, x)^{2-a},
\]

with boundary condition \(g(0, x) = 1 - x\). Notice the remarkable similarity between this partial differential equation and the ordinary differential equation in (2.20) for the total number of blocks. We interpret this as a form of self-similarity in terms of generating functions. It is straightforward to solve (3.21) and obtain

\[
g(t, x) = \left( (1 - x)^{a-1} + \frac{\Gamma(a + b)}{(2 - a)\Gamma(b)} t \right)^\frac{1}{a-1}.
\]

\[
\square
Proof. (of (2.9) in Theorem 1) From the definition (3.19) of $G$ it is clear that we can compute its coefficient $c_i(t)$ for instance by evaluating its $i$th partial derivative with respect to $x$ at $x = 0$. To this end, it will prove useful to write $G$ as a composition, namely

$$G(t, x) = c(t) - (f \circ g)(x),$$

where

$$f(x) := \left(x + \frac{\Gamma(a + b)}{(2 - a)\Gamma(b)} t\right)^{1 \over a-1}$$

and

$$g(x) := (1 - x)^{a-1}.$$

We can now find a formula for the $i$th partial derivative of $G$ by an application of Faà di Bruno’s formula, cf. [12], which states that

$$\frac{d^i}{dx^i} (f \circ g)(x) = \sum_{\pi \in \mathcal{P}[i]} f^{(#\pi)}(g(x)) \prod_{B \in \pi} g^{(#B)}(x),$$

for any two real functions $f, g$ that are at least $i$ times differentiable, where $f^{(j)}$ denotes the $j$th derivative of $f$. In our case, for $j \in \mathbb{N}$ the $j$th derivatives are

$$f^{(j)}(x) = \left(\frac{1}{a-1}\right)^j \left(x + \frac{\Gamma(a + b)}{(2 - a)\Gamma(b)} t\right)^{1 \over a-1 - j},$$

and

$$g^{(j)}(x) = (-1)^j (a-1)^j (1 - x)^{a-1-j}.$$

Plugging this into (3.23) yields

$$c_i(t) = -\frac{1}{i!} \left. \frac{\partial^i G}{\partial x^i} \right|_{x=0} \cdot G(t, x)$$

$$= -\frac{1}{i!} \sum_{\pi \in \mathcal{P}[i]} \left(\frac{1}{a-1}\right)^{#\pi} \left(1 + \frac{\Gamma(a + b)}{(2 - a)\Gamma(b)} t\right)^{1 \over a-1 - #\pi} \prod_{B \in \pi} (-1)^{#B} (a-1)^{#B}$$

$$= -\frac{c(t)}{i!} \sum_{\pi \in \mathcal{P}[i]} \left(\frac{1}{a-1}\right)^{#\pi} c(t)^{#\pi (1-a)} \prod_{B \in \pi} (1 - a)^{#B}$$

$$= \frac{c(t)^{2-a}}{i!} \sum_{m=1}^i \left(\frac{1}{1-a}\right)^{m} (-c(t)^{1-a})^{m-1} B_{i,m}((1-a)^\bullet)$$

$$= \frac{c(t)^{2-a}}{i!} B_i \left(\left(\frac{1}{1-a}\right)^\bullet (-c(t)^{1-a})^{\bullet-1}, (1-a)^\bullet\right).$$

$\square$
Corollary 1. In the limiting case $a \to 0$ we find for the limiting frequencies of blocks of size $i$,

$$c_i(t) = \frac{2}{2+t}c(t)^2(1-c(t))^{i-1} \left(\frac{t}{2+t}\right)^{i-1} \quad (t \geq 0).$$

These are the limiting frequencies in the Kingman coalescent in agreement with Smoluchowski’s result, cf. (2.6).

Proof. As $a \to 0$, (2.9) in Theorem 1 yields

$$c_i(t) = \frac{2}{2+a}B_i\left(\frac{1}{1-a}\right)(-c(t))^{i-1}, \quad (1-a)^{i-1} \quad (t \geq 0).$$

Moreover, recall that $B_i(k!) = \binom{i-1}{k}$ is the $(i,k)$th (unsigned) Lah number, cf. [23, Equation (1.55)], counting the number of partitions into $k$ linearly ordered subsets of a set containing $i$ elements. Thus

$$B_i((-c(t))^{i-1}, k!) = \sum_{k=1}^{i} k!(-c(t))^{k-1}B_i(k!) = \sum_{k=1}^{i} \binom{i-1}{k-1}(-c(t))^{k-1} = i!(1-c(t))^{i-1},$$

and the claim follows. \(\square\)

In complete analogy to our discussion of the block counting process of $\Pi_n$, define the process $(C^*,C^*)_{d+1}(t), \ldots, C^*_{n,d+1}(t), t \geq 0)$ via $C^*_i(t) := n^{-1}c_i\Pi(t\tau_n)$ for $i \in [d]$ and $C^*_{n,d+1}(t) = n^{-1}\sum_{i \geq d+1} c_i\Pi(t\tau_n)$. In particular, the process $(C^*_1, \ldots, C^*_{n,d+1})$ has initial state $(\infty, 0, \ldots, 0)$.

We now show a slightly more general result than Theorem 2. Namely, we show that for fixed $d \in \mathbb{N}$ and any sequence $(\tau_n)$ such that $\tau_n \sim n^{a-1}$ as $n \to \infty$ and $a < 1$ we have convergence

$$(C^*_1(t), \ldots, C^*_{n,d+1}(t), t \geq 0) \to (c_1^*(t), \ldots, c_{d+1}^*(t), t \geq 0),$$

as $n \to \infty$ in $D_{[0,\infty)}^{d+1}([0,\infty))$ in the Skorokhod topology, where the latter process is deterministic with initial state $(\infty, 0, \ldots, 0)$ and given by

$$c_i^*(t) = \frac{c_i(t)^{2-a}}{i!}B_i(1-1-a, (1-a)^i) \quad (i \in [d], t \geq 0).$$

Proof. (of Theorem 2) The process $(C^*_1(t), \ldots, C^*_{n,d+1}(t), t \geq 0)$ has state space $E^*_{n,d} = (n^{-1}\mathbb{N} \cup \{\infty\})^{d+1}$ and initial state $(\infty, 0, \ldots, 0)$. The limiting process $(c_1^*(t), \ldots, c_{d+1}^*(t), t \geq$
Theorem 4, the function \(c_d\text{ differential equations (3.16)}\) but with initial conditions \(c_0)\) solves the system of ordinary differential equations (3.16) with initial conditions \(c(3.16)\) but with initial conditions \(c_0) = M\) and \(c_{M,i}(0) = 0\) for \(i \geq 2\). The corresponding generating function

\[
\mathcal{G}_M(t, x) := \sum_{i \geq 1} c_{M,i}(t) x^i
\]

solves the partial differential equation (3.20) with boundary condition \(\mathcal{G}_M(0, x) = Mx\).

Letting \(c_M(t) := \sum_{i \geq 1} c_{M,i}(t)\), hence \(c_M(t) = (M^{a-1} + \frac{\Gamma(a + b)}{(2-a)\Gamma(b)} t)\frac{1}{\alpha-1}\) as in the proof of Theorem 4, the function

\[
g_M(t, x) := c_M(t) - \mathcal{G}_M(t, x)
\]

solves the partial differential equation (3.21) with initial condition \(g_M(0, x) = M(1 - x)\) and therefore

\[
g_M(t, x) = \left((M(1 - x))^{a-1} + \frac{\Gamma(a + b)}{(2-a)\Gamma(b)} t\right)\frac{1}{\alpha-1}.
\]

In complete analogy to the proof of (2.9) in Theorem 1 we let

\[
f(x) := \left(x + \frac{\Gamma(a + b)}{(2-a)\Gamma(b)} t\right)\frac{1}{\alpha-1} \quad \text{and} \quad g_M(x) := (M(1 - x))^{a-1},
\]

so \(g_M(t, x) = (f \circ g_M)(x)\). Since \(g_M^{(j)}(x) = M^{a-1}(-1)^j(a - 1)j!(1 - x)^{a-1-j}\), we obtain for \(i \in [d]\)

\[
c_{M,i}(t) = -\frac{1}{i!} \frac{\partial^i}{\partial x^i} \bigg|_{x=0} \mathcal{G}_M(t, x)
\]

\[
= -\frac{1}{i!} \sum_{\pi \in \mathcal{P}\{i\}} \left(\frac{1}{a-1}\right)^{\#\pi} (M^{a-1} + \frac{\Gamma(a + b)}{(2-a)\Gamma(b)} t)\frac{1}{\alpha-1} - \#\pi \prod_{B \in \pi} (-1)^{\#B} (a - 1)^{\#B} M^{a-1}
\]

\[
= -\frac{c_M(t)}{i!} \sum_{\pi \in \mathcal{P}\{i\}} \left(\frac{1}{a-1}\right)^{\#\pi} c_M(t)^{\#\pi(1-a)} \prod_{B \in \pi} (1 - a)^{\#B} M^{a-1}
\]

\[
= -\frac{c_M(t)}{i!} \sum_{m=1}^i (-1)^m \left(\frac{1}{1-a}\right)^m c_M(t)^m(1-a) B_{i,m}(1-a^x) M^{m(a-1)}
\]

\[
\to \frac{c_{\infty}(t)^{2-a}}{i!} B_i \left(\left(\frac{1}{1-a}\right)^x, (1-a^x)\right)
\]
as $M \to \infty$, since
\[(c_M(t)/M)^{m(1-a)} = \left( 1 + \frac{\Gamma(a + b)t}{(2 - a)\Gamma(b)} M^{1-a} \right)^{-m} \to 1\]
as $M \to \infty$.

3.2. Block counting process. We need some Lemmas in order to establish our proofs.

**Lemma 4.** For $a \notin \{1, 2\}$ one has
\[(3.31) \quad \gamma_n = \frac{1}{(1 - a)(2 - a)} \left( a + b - 1 \right) \frac{(a + b - 2) + b((1 - a)n - b + 1)}{(a + b)^{n-2}} \left( 1 + \Gamma(a + b) \frac{(a + b - 1)(a + b - 2)}{(a - 1)(a - 2)} \right) \]
as $n \to \infty$. As an immediate consequence, we obtain that for $a < 1$ the sequence $(\gamma_n, n \geq 2)$ is strictly increasing, that is for any $n \geq 2$ we have $\gamma_n < \gamma_{n+1}$.

**Remark 4.** When $a < 1$ the asymptotic behaviour of $\gamma_n$ as given in Lemma 4 agrees with Lemma 4 in [5].

**Proof.** We have
\[\gamma_n = \sum_{l=2}^{n} \frac{n}{l} \lambda_{n,l} = \sum_{l=2}^{n} \frac{n}{l} a^{l-2} b^{n-l} \]
\[= \frac{1}{(1 - a)(2 - a)(a + b)^{n-2}} \sum_{l=2}^{n} \frac{n}{l} (a - 2)^{l-1} b^{n-l} \]
\[= \frac{1}{(1 - a)(2 - a)} \left( \frac{(a + b - 2)\pi}{(a + b)^{n-2}} + \frac{(2 - a)n^{n-2} - b\pi}{(a + b)^{n-2}} \right) \]
\[= \frac{1}{(1 - a)(2 - a)} \left( (a + b - 1)(a + b - 2) + b((1 - a)n - b + 1) \right) \frac{(1 + b)^{n-2}}{(a + b)^{n-2}} \]
\[\sim \begin{cases} 
\frac{\Gamma(a + b)}{(2 - a)\Gamma(b)} \frac{n^{2-a}}{(a + b - 1)(a + b - 2)} & \text{if } a < 2 \\
\frac{\Gamma(a + b)}{(a - 1)(a - 2)} & \text{if } a > 2,
\end{cases}
\]
where we applied Lemmas 1 and 2 and distinguished the cases $a + b = 2$ and $a + b \neq 2$. 

\[\square\]
Lemma 5. For $a \notin \{1, 2\}$ as $n \to \infty$ we have that

\[
\gamma_n^{(1)} \sim \begin{cases} 
\frac{\Gamma(a+b)}{(1-a)(2-a)\Gamma(b)} n^{2-a} & \text{if } a < 1 \\
\frac{a+b-1}{a-1} n & \text{if } a > 1.
\end{cases}
\] (3.32)

Proof. We have

\[
\gamma_n^{(1)} = \sum_{l=2}^{n} \binom{n}{l} (l-1) \frac{a^{l-2}b^{n-l}}{(a+b)^{n-2}} = \gamma_n^{(1)} - \gamma_n.
\] (3.33)

Lemma 2 with $d = k = 1$ and some algebra yields

\[
\gamma_n^{(1)} = -\frac{n}{1-a} \left( \frac{(a+b-1)^{n-1}}{(a+b)^{n-2}} - \frac{b^{n-1}}{(a+b)^{n-2}} \right) = -\frac{n}{1-a} \left( a + b - 1 - \frac{b^{n-1}}{(a+b)^{n-2}} \right).
\]

Recall from Lemma 4,

\[
\gamma_n = -\frac{1}{(1-a)(2-a)} \left( (a+b-1)(a+b-2) + (2-a)n \frac{\Gamma(a+b)\Gamma(b+n-1)}{\Gamma(b)\Gamma(a+b+n-2)} - \frac{\Gamma(b+n)(a+b)}{\Gamma(b)\Gamma(a+b+n-2)} \right).
\]

Thus we obtain,

\[
\frac{\gamma_n^{(1)}}{\gamma_n} = \gamma_n^{(1)} - \gamma_n
\] (3.34)

\[
\frac{\gamma_n^{(1)}}{\gamma_n} = -\frac{n}{1-a} \frac{a+b-1}{(a+b)(2-a)} - \frac{a+b-1}{1-a}(a+b-2) + \frac{\Gamma(a+b)\Gamma(b+n)}{(1-a)(2-a)\Gamma(b)\Gamma(a+b+n-2)}.
\] (3.35)

and the claim follows. \qed

We need three more Lemmas in order to prove Theorem 3.

Lemma 6. For $a \notin \{1, 2\}$ we have $\gamma_n^{(2)}, \gamma_n^{(2)} \sim n^2$ as $n \to \infty$.

Proof. We can write $\gamma_n^{(2)} = \gamma_n^{(2)} - \gamma_n^{(1)}$, since $(l-1)^2 = l^2 - (l-1)$. Applying Lemma 2 with $d = 1$ and $k = 2$ we find that

\[
\gamma_n^{(2)} = \sum_{l=2}^{n} \binom{n}{l} \frac{a^{l-2}b^{n-l}}{(a+b)^{n-2}} l^2 = \frac{1}{(a-2)^2(a+b)^{n-2}} \sum_{l=2}^{n} \binom{n}{l} (a-2)^{l-2}b^{n-l}l^2 = n(n-1).
\] (3.36)

From this and Lemma 5 we conclude $\gamma_n^{(2)} = \gamma_n^{(2)} - \gamma_n^{(1)} \sim n^2$ as $n \to \infty$. \qed
Lemmas 7. For $a \notin \{1, 2\}$ we have $\gamma_n^{(3)} \sim \frac{a}{a+b} n^3$.

Proof. Notice that $(l-1)^3 = l^3 + (l-1)^2 - (l-1)^2$, hence $\gamma_n^{(3)} = \gamma_n^{(3)} + \gamma_n^{(2)} - \gamma_n^{(2)}$. From Lemma 6 we have $\gamma_n^{(2)} - \gamma_n^{(2)} \sim 0$ as $n \to \infty$. Now,

$$\gamma_n^{(3)} = \frac{1}{(a-2)(a-1)(a+b)^{-2}} \sum_{l=2}^{n} \binom{n}{l} (a-2) b^{n-l} l^3$$

$$= \frac{1}{(a-2)(a-1)(a+b)^{-2}} \sum_{l=2}^{n} \binom{n}{l} (a-2) b^{n-l} l^3$$

$$= \frac{1}{(a-2)(a-1)(a+b)^{-2}} (a-2) b^{n-3} n^3 (a+b)^{n-3}$$

as $n \to \infty$. \hfill \Box

Lemma 8. The sequence

$$\sup_{d=1}^{n} \left| d^{2-a} - \frac{\Gamma(b+d)}{\Gamma(a+b+d-2)} \right|, n \in \mathbb{N},$$

is bounded.

Proof. Gautschi’s inequality states that for $x > 0$ and $s \in (0, 1)$,

$$x^{1-s} < \frac{\Gamma(x+1)}{\Gamma(x+s)} < (1+x)^{1-s}.$$  \hspace{1cm} (3.38)

Applying this inequality to our case, we obtain for $a < 1$

$$(b+d-2)^{1-a} < \frac{\Gamma(b+d-2+1)}{\Gamma(b+d-2+a)} < (b+d-1)^{1-a}.$$  \hspace{1cm}

Thus the function

$$f: [1, \infty] \to \mathbb{R}, x \mapsto x^{2-a} - \frac{(b+x-1)\Gamma(b+x-1)}{\Gamma(b+x-2+a)}$$

is bounded below by

$$g: [1, \infty] \to \mathbb{R}, x \mapsto x^{2-a} - (b+x-1)^{2-a},$$

and bounded above by

$$h: [1, \infty] \to \mathbb{R}, x \mapsto x^{2-a} - (b+x-1)(b+x-2)^{1-a}.$$
The derivative of $g$ is given by $g'(x) = (2-a)(x^{1-a} - (b+x-1)^{1-a})$. For $b \geq 1$, we have $g' \leq 0$, and the maximum of $g$ is $g(1) = 1 - b^{2-a} \leq 0$. For $b < 1$ we have $g' > 0$, hence on any closed interval $[1,y]$ in the real numbers $g$ obtains its maximum at $y$. Since $x \mapsto x^{p}$ is Lipschitz continuous, there exists a constant $M > 0$ such that

$$|g(y)| = |y^{2-a} - (b+y-1)^{2-a}| \leq M|1-b|.$$  

Consequently, $f$ is bounded below on its domain. A similar argument for $h$ instead of $g$ shows that $f$ is bounded above and below on its domain, and the claim follows.

For a metric space $(E,r)$ we denote by $D_{E}([0,\infty))$ the space of right-continuous functions from $[0,\infty)$ into $E$ having left limits. Moreover, by $C(E)$, respectively $C^{\infty}(E)$, we denote the continuous, respectively smooth functions (that is functions that have derivatives of arbitrary order) from $E$ to $\mathbb{R}$.

**Proof.** (of Theorem 3) The jump chain $(J_{k}^{n})_{k \geq 0}$ of $C_{n}(t) = n^{\alpha} N_{n}(t \tau_{n})$ has transition probabilities

$$
\mu_{n}(c, c - n^{\alpha}(l-1)) := \mathbb{P} \{ J_{k}^{n} = c - n^{\alpha}(l-1) | J_{0}^{n} = c \} = \begin{cases} \binom{n^{\alpha}c}{l} \frac{\lambda_{n^{-\alpha}c}}{\lambda_{n^{-\alpha}c}} & \text{if } c > n^{\alpha}, 2 \leq l \leq n^{-\alpha}c, \\ 1 & \text{if } c = n^{\alpha}, l = 1, \\ 0 & \text{otherwise}. \end{cases}
$$

Denoting by $\lambda_{n}(c)$ the total rate of $C_{n}$ in state $c \in E_{n}$ for any $f \in C^{\infty}([0,1])$ the generator of $C_{n}$ is given by

$$
\mathcal{G}_{n}f(c) = \lambda_{n}(c) \int_{E_{n}} (f(c') - f(c)) \mu_{n}(c, dc')
= \tau_{n} \lambda_{n^{-\alpha}c} \sum_{l=2}^{n^{\alpha}c} (f(c) - f(n^{\alpha}(l-1)) - f(c)) \binom{n^{\alpha}c}{l} \frac{\lambda_{n^{-\alpha}c}}{\lambda_{n^{-\alpha}c}}
= \tau_{n} \sum_{l=2}^{n^{\alpha}c} (-n^{\alpha}(l-1)f'(c) + R_{2}(\vartheta_{n,l}) \binom{n^{\alpha}c}{l} \lambda_{n^{-\alpha}c,l},
$$

where we used Taylor’s approximation in the third equality. Taylor’s approximation ensures the existence of a value $\vartheta_{n,l} \in (c - (l-1)/n, c)$ such that the remainder term $R_{2}(\vartheta_{n,l})$ is given, for instance, by its Lagrange form

$$
R_{2}(\vartheta_{n,l}) = \frac{1}{2} (n^{\alpha}(l-1))^{2} f''(\vartheta_{n,l}).
$$

First notice that by Lemma 5

$$
-n^{\alpha} \tau_{n} \sum_{l=2}^{n^{\alpha}c} \binom{n^{\alpha}c}{l} (l-1) \lambda_{n^{-\alpha}c,l} = -n^{\alpha} \tau_{n} \gamma_{n^{-\alpha}c}^{(1)} \rightarrow -\frac{\Gamma(a+b)}{(1-a)(2-a)\Gamma(b)} \alpha^{2-a}
$$
as \( n \to \infty \), since \( \tau_n \) is chosen to be of order \( n^{(1-a)\alpha} \), and \( a < 1 \) by assumption. Since \( f \) has derivatives of arbitrarily high order on \([0,1]\), \( f'' \) attains its supremum \( \|f''\|_\infty := \sup_{x \in [0,1]} |f''(x)| \). Consequently, applying Lemma 6 in the third step we obtain

\[
\tau_n \sum_{l=2}^{n-\alpha} \binom{n-\alpha}{l} \lambda_{n-\alpha,c,l} R_2(x_{n,l}) \leq n^{2\alpha} \frac{\tau_n}{2} \|f''\|_\infty \sum_{l=2}^{n-\alpha} \binom{n-\alpha}{l} \lambda_{n-\alpha,c,l} (l-1)^2 \\
= n^{2\alpha} \frac{\tau_n}{2} \|f''\|_\infty \gamma_{n-\alpha,c}^{(2)} \sim \frac{1}{2} c^2 \tau_n \|f''\|_\infty
\]

as \( n \to \infty \). For the term on the left hand side to vanish we need \( \phi = (1-a)\alpha < 0 \), which explains the restriction \( \alpha < 0 \) for \( a < 1 \). For \( a > 1 \) the only non-trivial solution of \( m(t) \) is obtained for \( \phi = 0 \). However, using \( \gamma_{n-\alpha,c}^{(3)} \sim \frac{a}{a+b} n^3 \), Lemma 7, one can see that in this case the remainder term of third order in Taylor’s approximation has Lagrange form \( R_3(\vartheta_{n,l}) \) on the order of \( (n^{\alpha}(l-1))^3 f'''(\theta_{n,l}) \) for some \( \theta_{n,l} \in (c-(l-1)/n,c) \), and one can show that

\[
\tau_n \sum_{l=2}^{n-\alpha} \binom{n-\alpha}{l} \lambda_{n-\alpha,c,l} R_3(x_{n,l})
\]

is bounded by a term asymptotically equivalent to

\[
\begin{cases} 
0 & \text{if } \phi < 0, \\
1 & \text{if } \phi = 0, \\
\frac{1}{2} c^2 \|f''\|_\infty & \text{if } \phi > 0,
\end{cases}
\]

as \( n \to \infty \). That is, for \( a > 1 \) and for \( a < 1, \phi = 0 \) we do not obtain a diffusion limit as \( n \to \infty \).

We are now ready to show that as \( n \) grows without bounds,

\[
\sup_{c \in E_n} |G_n f(c) - G f(c)|
\]

converges to zero, where the operator \( G \) is defined by

\[
G f(c) := -\frac{\Gamma(a+b)}{(1-a)(2-a)\Gamma(b)} c^{2-a} f'(c).
\]

For the sake of simplifying the proof, let us henceforth assume that \( \tau_n = n^{(1-a)\alpha} \). First, for any \( c \in E_n \) we have

\[
|G_n f(c) - G f(c)| \leq |n^{(2-a)\alpha} \gamma_{n-\alpha,c}^{(1)} f'(c) - G f(c)| + n^{2\alpha} \frac{\tau_n}{2} \|f''\|_\infty \gamma_{n-\alpha,c}^{(2)}.
\]
We first focus on the last summand in (3.42). Since \( \gamma_n^{(2)} = \gamma_n^{(2)} - \gamma_n^{(1)} \), we find from equations (3.34) and (3.36) that \( \gamma_n^{(2)} \leq Cn^2 \) for some constant \( C > 0 \). Thus as \( n \to \infty \)

\[
 n^{2\alpha + n/2} \| f'' \|_\infty \sup_{c \in E_n} \gamma_n^{(2)} \leq Cn^{(3-a)\alpha} n^2 \to 0,
\]

if \( \alpha < -2/(3-a) \). Now focus on the first summand on the right hand side of (3.42) to show that

\[
 \lim_{n \to \infty} \sup_{c \in E_n} \left| -n^{(2-a)\alpha} \gamma_n^{(1)} f'(c) - G f(c) \right| = 0.
\]

From the proof in Lemma 5 recall that

(3.43)

\[
 \gamma_n^{(1)} = -n^{a + b - 1} (a + b - 1) (a + b - 2) \frac{\Gamma(a + b) + \Gamma(b + n) (1 - a) (2 - a)}{(1 - a) (2 - a) \Gamma(b) (a + b - 2 + n)}.
\]

We focus on the last summand to see that

\[
 \lim_{n \to \infty} \sup_{c \in E_n} \left| -n^{(2-a)\alpha} \frac{\Gamma(a + b)}{(1 - a) (2 - a) \Gamma(b) (a + b + n^{-\alpha} c - 2)} f'(c) - G f(c) \right| = 0.
\]

To this end, observe that

\[
 \sup_{c \in E_n} \left| -n^{(2-a)\alpha} \frac{\Gamma(a + b)}{(1 - a) (2 - a) \Gamma(b) (a + b + n^{-\alpha} c - 2)} f'(c) - G f(c) \right| \]

\[
 = \frac{\Gamma(a + b)}{(1 - a) (2 - a) \Gamma(b) f'(c) \sup_{c \in E_n} \left| c^{(2-a)} - n^{(2-a)\alpha} \frac{\Gamma(b + n^{-\alpha} c)}{(a + b + n^{-\alpha} c - 2)} \right|} \]

\[
 = \frac{n^{(2-a)\alpha} \Gamma(a + b)}{(1 - a) (2 - a) \Gamma(b) f'(c) \sup_{d = 1} \left| d^{(2-a)} - \frac{\Gamma(b + d)}{(a + b + d - 2)} \right|} \to 0
\]

as \( n \to \infty \), since \( \sup_{d = 1} \left| d^{(2-a)} - \frac{\Gamma(b + d)}{(a + b + d - 2)} \right| \) as a sequence in \( n \) is bounded, cf. Lemma 8.

The contribution of the second summand in (3.43) to (3.42) clearly vanishes as \( n \to \infty \).

The contribution of the first summand in (3.43) to (3.42) is of the order \( n^{(3-a)\alpha} \) which also vanishes as \( n \to \infty \) under our requirement that \( \alpha < -2(3-a) \).

Overall, this shows the convergence

(3.44)

\[
 \lim_{n \to \infty} \sup_{c \in E_n} |G_n f(c) - G f(c)| \to 0.
\]

Since \( [0, \infty] \ni c \mapsto -c^{2-a}(a+b)/(1-a)(2-a)\Gamma(b) \) is Lipschitz continuous, Theorem 2.1 in Chapter 8 of [9] yields that the set \( C^\infty([0,\infty]) \) is a core for \( G \), and the closure of \( \{ (f, G f) : f \in C^\infty([0,\infty]) \} \) is single-valued and generates a Feller semigroup \( \{ T(t) \} \) on
By Theorem 2.7 in Chapter 4 of [9] there exists a process \( c \) corresponding to \( \{ T(t) \} \).

To prove that \( C_n \) converges in \( D_{[0,1]}([0,\infty)) \) in the Skorokhod topology to \( c \) as \( n \to \infty \), it suffices by Corollary 8.7 of Chapter 4 to show that (3.44) holds for all \( f \) in a core for the generator \( G \), which we have just done.

Instead of the restriction \( \Pi_n \) of the beta coalescent \( \Pi \), we now rescale the latter process, namely for each \( n \in \mathbb{N} \) let \( C_n^* = \{ C_n^*(t), t \geq 0 \} \) be defined by \( C_n^*(t) := n^a \# \Pi(t \tau_n) \). In particular, notice that the initial state of this process is \( C_n^*(0) = \infty \), irrespective of \( a \).

Proof. (of Theorem 4) For the most part the calculations are identical to the ones in the proof of Theorem 3. Notice that the process \( C_n^* \) has state space \( E_n^* := n^a \mathbb{N} \cup \{ \infty \} \) and initial state \( C_n(0) = \infty \). Because of the consistency of the \( \Lambda \) coalescents, i.e. \( \Pi_n \) is equal in distribution to the restriction of \( \Pi \) to \([n] \), the generators of \( C_n \) and \( C_n^* \) are of precisely the same form, except that the generator of \( C_n \) operates on functions \( f \) mapping \( n^a \mathbb{N} \) to \( \mathbb{R} \), whereas \( C_n^* \) operates on functions \( f \) mapping \( n^a \mathbb{N} \cup \{ \infty \} \) to \( \mathbb{R} \). For this reason the generator calculations for \( C_n^* \) are identical to the ones given in the proof of Theorem 3. In particular, the limit \( c^*(t) \) of \( C_n^*(t) \) as \( n \to \infty \) satisfies the ordinary differential equation (2.20) with boundary condition \( c^*(0) = \infty \). However, we already solved this ODE in equation (2.22).

Remark 5. (1) Applying Legendre’s duplication formula as in Remark 2, we find for \( a = \frac{1}{2} \)

\[
  c^*(t) = \frac{9 \Gamma(b)^4}{16^{1-b} \pi \Gamma(2b)^2} \frac{1}{t^2} \quad (t \geq 0),
\]

which boils down to

\[
  c^*(t) = \frac{9 \pi}{4 t^2} \quad (t \geq 0)
\]

for the arcsine coalescent.

(2) In the limiting case \( a \to 0 \) we obtain

\[
  c^*(t) = \frac{2}{t} \quad (t \geq 0),
\]

which agrees with the result for Kingman’s coalescent.

Acknowledgements. Most of the results in this work are taken from the authors’ PhD theses. They would like to thank their supervisor Alison Etheridge for her advice and helpful discussions. The authors are grateful to an anonymous referee for a thorough
and thoughtful reading and helpful suggestions. H. H. P. also thanks Julien Berestycki and Matthias Birkner as well as Martin Möhle and Elmar Teufl for stimulating discussions, and thankfully acknowledges financial support by the foundation “Private Stiftung Ewald Marquardt für Wissenschaft und Technik, Kunst und Kultur”.

References.


