Aspects of exchangeable coalescent processes

Helmut H. Pitters

St Peter's College, Oxford
Trinity Term, 2015

A thesis submitted for the degree of
Doctor of Philosophy of the University of Oxford.

Statement of authorship

I declare that the material in this thesis is my own work, except where otherwise stated.

Abstract

Aspects of exchangeable coalescent processes

Helmut H. Pitters
St Peter's College, Oxford

D.Phil. Thesis
Trinity Term, 2015

In mathematical population genetics a multiple merger n-coalescent process, or Λ n-coalescent process, $\{\Pi^n(t), t \geq 0\}$, models the genealogical tree of a sample of size n (e.g. of DNA sequences) drawn from a large population of haploid individuals. We study various properties of Λ coalescents. Novel in our approach is that we introduce the partition lattice as well as cumulants into the study of functionals of coalescent processes. We illustrate the success of this approach on several examples.

Cumulants allow us to reveal the relation between the tree height, T_n, respectively the total branch length, L_n, of the genealogical tree of Kingman's n-coalescent, arguably the most celebrated coalescent process, and the Riemann zeta function. Drawing on results from lattice theory, we give a spectral decomposition for the generator of both the Kingman and the Bolthausen-Sznitman n-coalescent, the latter of which emerges as a genealogy in models of populations undergoing selection [11].

Taking mutations into account, let M_j count the number of mutations that are shared by j individuals in the sample. The random vector $(M_1, \ldots, M_{n-1})$, known as the site frequency spectrum, can be measured from genetical data and is therefore an important statistic from the point of view of applications. Fu [32] worked out the expected value, the variance and the covariance of the marginals of

the site frequency spectrum. Using the partition lattice we derive a formula for the cumulants of arbitrary order of the marginals of the site frequency spectrum.

Following another line of research, we provide a law of large numbers for a family of Λ coalescents. To be more specific, we show that the process $\{\#\Pi^n(t), t \geq 0\}$ recording the number $\#\Pi^n(t)$ of individuals in the coalescent at time t, coverges, after a suitable rescaling, towards a deterministic limit as the sample size n grows without bound. In the statistical physics literature this limit is known as a hydrodynamic limit. Up to date the hydrodynamic limit was known for Kingman's coalescent, see [5], but not for other Λ coalescents. We work out the hydrodynamic limit for beta coalescents that come down from infinity, which is an important subclass of the Λ coalescents.

Contents

List of symbols and abbreviations

$\leq$	partial order on $\mathcal{P}_{[n]}$ by inclusion	Section 2.1
$\lessdot$	inclusion relation on $\mathcal{P}_{[n]}$ allowing for mergers of two blocks only	Section 3.3.1
$\prec$	inclusion relation on $\mathcal{P}_{[n]}$ allowing for one simultaneous merger only	Section 3.3.1
$\#A$	cardinality of set A	Section 2.1
$a \sim_\pi b$	a and b are elements of a common block in π	Section 2.1
$\mathrm{Beta}(a, b)$	Beta distribution	(2.12)
B_n	nth Bell number	Examples 2.3
$B(a, b)$	Beta function	(2.13)
$B_n((v_k), (w_k))$	nth complete Bell polynomial	(2.2)
$B_{n,k}((w_k))$	(n, k)th partial Bell polynomial	(2.3)
CDI	coming down from infinity	Section 2.3.1
$\mathscr{C}_i(X)$	ith cumulant of r.v. X	Section 4.1.1
$\mathscr{C}(X_1, \ldots, X_n)$	joint cumulant of $(X_1, \ldots, X_n)$	Section 4.2
$\mathscr{C}_{i_1,\ldots,i_n}(X_1, \ldots, X_n)$	joint cumulant of $(X_1, \ldots, X_n)$ of order $(i_1, \ldots, i_n)$	Section 4.2
$\mathscr{C}(X\|Y)$	conditional cumulant of X given Y	Section 4.2
$\mathit{\Delta}_A$	$\{\{a\} \colon a \in A\}$	Section 2.1
d_n	number of derangements of $\{1, \ldots, n\}$	(4.13)
(g_{ij})	Green's matrix of N	(3.26)
$(g_{\pi\rho})$	Green's matrix of $\mathit{\Pi}$	(3.36)
$\mathit{\Gamma}(x)$	Gamma function	(2.15)
$\mathrm{Gamma}(a, b)$	gamma distribution with parameters a, b	(4.41)
h_n	$(n-1)$th harmonic number	Section 4.2
$h(i, j)$	hitting probabilities of N	(3.5)
$h(\pi, \rho)$	hitting probabilities of $\mathit{\Pi}$	(3.38)
$J = \{J_k\}_{k=1}^{n-1}$	jump chain of $\mathit{\Pi}^n$	Definition 2.4
$\mathit{\Lambda}$	finite measure on $[0, 1]$	Definition 2.4
$\lambda_{m,k}$	infinitesimal rates of $\mathit{\Pi}$	Definition 2.4

L_n	total branch length of Kingman's n-coalescent tree	(4.26)
$\left\lfloor {n \atop k} \right\rfloor$	(signless) Lah numbers	(2.7)
$\mathrm{Li}_n(u)$	polylogarithm	(4.39)
$m_n(X)$	nth moment of r.v. X	(4.6)
$m'_n(X)$	nth central moment of r.v. X	(4.10)
$m(\pi,\rho)$	number of maximal chains in $[\pi,\rho]$	(3.41)
$(M_1,\ldots,M_{n-1})$	site frequency spectrum	Section 4.2
$[n]$	$\{1,\ldots,n\}$	Section 1.1
$\mathbb{N}$	$\{1,2,\ldots\}$	
$\mathbb{N}_0$	$\mathbb{N}\cup\{0\}$	
$N(t)$	ancestral process	Sections 3.1, 3.2
$N^n(t)$	ancestral process started in state n	Sections 3.1, 3.2
$\mathrm{Nb}(\alpha,p)$	negative binomial distribution	(4.40)
$\varPi$	$\varLambda$ coalescent	Definition 2.4
$\varPi^n$	$\varLambda$ n-coalescent	Definition 2.4
$\mathcal{P}_A$	set of all partitions of set A	Section 2.1
$\mathcal{P}_{A,i}$	set of all partitions of set A with i blocks	Section 2.1
$\pi\vert_B$	restriction of partition π to set B	Section 2.1
$[\pi,\rho]$	interval in $\mathcal{P}_{[n]}$	Section 2.1
$\pi \sqsubset \mathscr{T}$	tree $\mathscr{T}$ contains partition π	Section 3.3.3
$p(\mathscr{T})$	label set of tree $\mathscr{T}$	Section 3.3.3
$P\cong Q$	partially ordered sets P, Q are isomorphic	Section 2.1
$\mathrm{Pn}(\alpha)$	Poisson distribution	Section 4.2
$q_{\pi,\rho}$, q_π	infinitesimal rates of $\varPi$	Section 3.3
$q_{i,j}$, q_i	infinitesimal rates of ancestral process N	Sections 3.1, 3.2
$\mathcal{R}_\pi$	Markov chain with values in the random recursive trees	Section 2.3.1
R_π	jump chain of $\mathcal{R}_\pi$	Section 3.3.3
SFS	site frequency spectrum	Section 4.2
$\left[{n \atop k} \right]$	(signless) Stirling cycle numbers	(2.4)
$s(n,k)$	(signed) Stirling cycle numbers	(2.5)
$\left\{ {n \atop k} \right\}$	Stirling partition numbers	(2.6)
$S_{n,k}^{\alpha,\beta}$	generalized Stirling numbers	(2.8)
$\mathcal{T}$	coalescent tree corresponding to $\varPi$	Section 2.3.2
$\mathcal{T}^n$	coalescent tree corresponding to $\varPi^n$	Section 2.3.2
τ_k	waiting time when $\varPi$ is in a state with k blocks	Section 2.3
T_n	absorption time of $\varPi^n$	(3.6)
$\mathfrak{T}_\pi$	random recursive tree on label set π	Section 2.3.1
$c\mathscr{T}$	tree obtained by cutting an edge in tree $\mathscr{T}$ uniformly	Section 2.3.1

$x^{\overline{n}}$	nth ascending factorial power	Corollary 3.17
$x^{\underline{n}} = x^{\underline{n,1}}$	nth descending factorial power	(2.9)
$x^{\underline{n,\alpha}}$	nth descending factorial power with increment α	(2.9)
$\zeta(x)$	Riemann zeta function	Section 4.1

Chapter 1

Introduction

1.1 Background

The main forces in the evolution of biological populations are believed to be genetic drift (also known as random mating), mutation, natural selection, recombination, spatial structure and size of a population, and competition between different populations. This is a crude abstraction from phenomena observed in nature, but in order to obtain mathematically tractable models, even more abstractions have to be made, i.e. even more of these forces have to be neglected.

The most celebrated model in population genetics for the evolution of a single population starts from just four assumptions, namely that the population

- evolves in discrete, non-overlapping generations,
- is in equilibrium, i.e. has a fixed size at all times,
- is neutral, and
- its individuals are haploid.

Neutral means, that there is no natural selection acting, that is no individual is favoured over another. Haploid means, that each individual only has one parent. The model just mentioned is known as the (neutral) Wright-Fisher model and we do not provide a mathematically precise definition of it here. For an introduction to mathematical models in population genetics the reader is referred to Etheridge's lecture notes [27].

The only way to gather empirical information about a population in nature is to sample from its current state and infer its properties based on this data. We might then be interested in the so-called genealogy of a sample from this population, that is for any two individuals (and their ancestors) we want to know their (common) ancestor (if any) and the times of their births and deaths. For a large population size, N, it turns out that the genealogy of a sample of size n, drawn from this population, can be approximated by Kingman's n-coalescent. Kingman's n-coalescent can be thought of as a stochastic process with state space the partitions of $[n] := \{1, \dots, n\}$, initial state $\{\{1\}, \dots, \{n\}\}$ and absorbing state $\{[n]\}$ (we occasionally drop the outer parantheses for this partition). If the current state of the process consists of $b > 0$ blocks, any two blocks merge at rate 1 to form the next state, otherwise no further transition occurs.

A fascinating fact about Kingman's coalescent is its robustness: we can change a lot of details in the reproduction mechanism of the (forwards in time) population model (e.g. we may allow for a fluctuating population size, subject to some constraints, of course), and still obtain a genealogy that is governed by Kingman's coalescent in the limit of large population size. So, for a lot of discrete population models Kingman's coalescent plays the rôle that Normal distribution plays for i.i.d. sequences of random variables in the Central Limit Theorem, or that Brownian motion plays for sequences of random walks in Donsker's invariance principle. In probability theory such phenomena are termed universality phenomena, and the limiting object, in our case Kingman's coalescent, is called a universal scaling limit.

Natural generalizations of the Wright-Fisher model are models of randomly mating neutral populations of constant size, also known as Cannings models [20], [21]. Möhle and Sagitov [54] showed that if the population size is large, the genealogy of a suitably rescaled Cannings model can be approximated by a certain class of coalescent processes, so-called exchangeable coalescents or Ξ coalescents, that allow for simultaneous multiple mergers. Here multiple means that more than two blocks may merge in one instant (unlike Kingman's coalescent, that only allows for mergers of two blocks), and simultaneous means that the blocks merging do not all have to end up in the same block, but may belong to different groups each of which gives rise to a new block. In recent years a subclass of the Ξ coalescents has attracted the increasing attention of researchers. The coalescent processes in this subclass only allow for multiple (but not simultaneous) mergers, and are also known as Λ coalescents, see Section 2.3 for a precise definition. Multiple merger coalescents have been introduced independently by Donnelly and Kurtz [24], Pitman [60] and Sagitov [64].

In this thesis we will study various properties of multiple merger coalescent processes.

1.2 Motivation

In some sense, the starting point of this thesis is the question of how to characterize the site frequency spectrum (SFS) of Kingman's n-coalescent. The SFS is the random vector $(M_1, \ldots, M_{n-1})$, where M_j denotes the number of mutations that are shared by j individuals in the sample. For a definition of the coalescent with mutations see Definition 2.7. Since the SFS can be measured in genetical data, this quantity is highly important from the point of view of applications, e.g. for statistical inference. Fu [32] has studied distributional properties, i.e. expected value, variance and covariance, of the SFS for Kingman's coalescent, whereas Birkner, Blath and Eldon [16] extended Fu's results to recursions for the expected value, the variance and covariance of the SFS for multiple merger coalescents.

The SFS is a rather complex quantity, and even though we make some steps towards characterizing its distribution, we cannot yet give a complete characterization. Along the way, the author thought of solving some simpler problems instead, that might shed some light on what a good approach to the SFS should be.

One of these simpler problems is concerned with the cumulants of the absorption time T_n of Kingman's n-coalescent $\Pi^n = \{\Pi^n(t), t \geq 0\}$. When Martin Möhle told the author that calculations suggested that the moments $m_j(T_n) = \mathbb{E}T_n^j$ of T_n should be polynomials in π^2, the author got hooked. Why should this be the case? Instead of working out the moments of T_n directly, the author noticed that its cumulants $\mathscr{C}_j(T_n)$ are much more tractable because of their linearity with respect to convolutions (i.e. the independence of two real random variables X, Y implies $\mathscr{C}_j(X+Y) = \mathscr{C}_j(X) + \mathscr{C}_j(Y)$), and since

$$T_n = \sum_{k=2}^{n} \tau_k,$$

where the waiting times (τ_k) of Π^n are independent exponentials. This observation reveals the connection between $\mathscr{C}_j(T_n)$ and the Riemann zeta function, see (4.4), which in turn implies that $\mathscr{C}_j(T_n)$ are polynomials in π^2. The moments $m_j(T_n)$ are then found by the well-known transformation (see [61, Equation (1.30)])

$$m_j(T_n) = \sum_{\pi \in \mathcal{P}_{[n]}} \prod_{B \in \pi} \mathscr{C}_{\#B}(T_n),$$

where $\mathcal{P}_{[n]}$ denotes the set of all partitions of $[n]$.

The fact that the waiting times of Kingman's coalescent are independent (and independent of the jump chain) can be further exploited to characterize functionals of the coalescent in terms of cumulants. Most studies in coalescent theory are concerned with the limit as $n \to \infty$ of the distribution of some quantitative property, or functional, $f(\Pi^n)$, of the n-coalescent Π^n. Though this is an interesting and sensible endeavour, there are good reasons to study functionals $f(\Pi^n)$ when the sample size n is fixed. One reason being that for statistical inference knowledge of the limiting law of $f(\Pi^n)$ as $n \to \infty$ is really only useful if the rate of convergence is known (and sufficiently high). In this case the limiting law may be used as an approximation for the law of $f(\Pi^n)$.

Another such functional that we investigate is the total tree length

$$L_n = \sum_{k=2}^{n} k\tau_k$$

of Kingman's n-coalecsent. A well known result, see Etheridge's lecture notes [27], is the convergence

$$L_n/2 - \log n \to G \tag{1.1}$$

in distribution as $n \to \infty$, where G is a random variable distributed according to a standard Gumbel law. We will see in Theorem 4.5 that it is straightforward to derive (1.1) via cumulants, which then implies that the convergence holds in L^p for any $p \in (0, \infty)$. (In fact, Theorem 4.5 states that (1.1) even holds almost surely.)

These results and the independence of the waiting times of Kingman's coalescent suggest that cumulants seem to be rather well-suited for the characterization of other functionals, as for instance the SFS. The use of cumulants in the study of functionals of coalescent processes seems to be new in the literature. Before we return to the SFS, let us make another observation. Both the absorption time T_n and the total tree length L_n of Kingman's n-coalescent are simple in the sense that they are fully determined by the knowledge of the ancestral process $N = \{N(t), t \geq 0\}$, where $N(t)$ counts the number of blocks in the coalescent at time t. However, if we want to study more complex functionals (for instance, the SFS), we need a better understanding of $\mathcal{P}_{[n]}$, the set of all partitions of $[n]$, which is the state space of the n-coalescent. To the best of the author's knowledge the lattice structure of $\mathcal{P}_{[n]}$ has not been studied and exploited before in the context of coalescent processes.

From this perspective the results on the spectral decomposition of the partition-valued coalescents (Kingman and Bolthausen-Sznitman, to be precise) in Section 3.3 can be viewed as an exercise in the connection between the n-coalescent and its state space $\mathcal{P}_{[n]}$. On the other hand, the spectral decomposition of a stochastic process is a basic problem whose answer has numerous applications (in our case it allows us to derive, among other things, transition probabilities, see Corollary 3.17, and hitting probabilities, respectively Green's functions, see (3.37)), and is therefore of interest in its own right.

Equipped with an understanding of cumulants and the partition lattice $\mathcal{P}_{[n]}$ we finally characterize the marginals M_j of the SFS in terms of cumulants in Theorem 4.14. This result should be considered as a first step towards the (complete) characterization of the law of the SFS. To derive this result, we work out the cumulants of the negative binomial distribution in Lemma 4.15, which are given by the polylogarithm. To the best of the author's knowledge these cumulants have not been derived before in the literature.

Finally, we study in Chapter 5 the hydrodynamic limit of beta coalescents (defined in Examples 2.5) that come down from infinity. This is a law of large numbers type of result quantifying the manner in which a beta coalescent comes down from infinity. So far in coalescent theory the hydrodynamic limit has only been known for Kingman's coalescent, see Aldous [5]. The main reason for this is that Kingman's coalescent (unlike other multiple merger coalescents) fits into the framework of Smoluchowski's coagulation equations. For beta coalescents that come down from infinity the hydrodynamic limit, $\{c(t), t \geq 0\}$, is a deterministic limit as $n \to \infty$ of the suitably rescaled block counting process (or ancestral process) $N^n = \{N^n(t), t \geq 0\}$, where $N^n(t)$ counts the number of blocks of the n-coalescent Π^n at time t. We show in Theorem 5.10 that

$$c(t) = \left(1 + \frac{\Gamma(a+b)}{(2-a)\Gamma(b)}t\right)^{\frac{1}{a-1}} \qquad (t \geq 0),$$

where a, b are parameters characterizing the beta coalescent.

In the second step, we consider the block size spectrum

$$\{\mathfrak{c}_1\Pi^n(t), \ldots, \mathfrak{c}_n\Pi^n(t), t \geq 0\}$$

of $\varPi^n(t)$, where $\mathfrak{c}_i\varPi^n(t)$ counts the number of blocks of size i in $\varPi^n(t)$. Applying the same rescaling as before, we again find a deterministic limit $\{(c_i(t))_{i\geq 1}, t \geq 0\}$ as $n \to \infty$, see Corollary 5.16, where $c_i(t)$ is given by the ith complete Bell polynomial (defined in Definition 2.2) with suitable parameters.

1.3 Summary

This thesis is laid out as follows. In Chapter 2 we introduce basic notions such as the partition lattice, $\mathcal{P}_A$, the multiple merger coalescent, $\varPi$, and the corresponding coalescent tree, $\mathcal{T}$. In Chapter 3 we provide spectral decompositions for the generator of $\varPi^n$ and the corresponding ancestral process, or block counting process, N^n, in the case of the Kingman and the Bolthausen-Sznitman coalescent. For Kingman's n-coalescent we study the height and total tree length of $\mathcal{T}^n$ and the site frequency spectrum in the setting of finite sample size n in Chapter 4. Finally, we turn to the block size spectrum $(\mathfrak{c}_1\varPi^n, \ldots, \mathfrak{c}_n\varPi^n)$, where $\mathfrak{c}_i\pi$ counts the number of blocks of size i in $\pi \in \mathcal{P}_{[n]}$, and its scaling limit, the so-called hydrodynamic limit, in Chapter 5. In the setting of finite sample size we characterize the distribution of the block size spectrum for the Kingman and the Bolthausen-Sznitman n-coalescent. For a beta coalescent (defined in Examples 2.5) that comes down from infinity (see Section 2.3.1), we work out the scaling limit of the block size spectrum that quantifies the manner in which the beta coalescent comes down from infinity.

Chapter 2

Basic notions

In this Chapter we introduce basic notions and results that will be used throughout this thesis. We start with the partition lattice $\mathcal{P}_{[n]}$ in Section 2.1 which has been studied extensively in combinatorics, see e.g. [70], and lattice theory, see e.g. [15], [35]. Throughout our investigations Bell polynomials will appear in various forms (some special cases being different kinds of Stirling numbers) and we therefore define them in Section 2.1. The partition lattice is the state space of (multiple merger) coalescent processes defined in Section 2.3. We give some examples of the most popular coalescent processes and some of their properties. Usually, a coalescent is identified with a random tree, and this identification is made precise in Section 2.3.2.

2.1 Partition lattice

A *partition* of a non-empty set A is a set, π say, of nonempty pairwise disjoint subsets of A whose union is A. The members of π are called the *blocks* of π. Let

$\#A$ denote the cardinality of A and let $\mathcal{P}_A$ denote the *set of all partitions of* A. For $i \in \mathbb{N}$ we denote by $\mathcal{P}_{A,i}$ the *set of all partitions of* A *that consist of* i *blocks.* For any two elements $a, b \in A$ and $\pi \in \mathcal{P}_A$ we write $a \sim_\pi b$ if and only if a and b are elements of a common block in π, i.e. if there exists a $B \in \pi$ such that $a, b \in B$.

Recall that a *partially ordered set* (*poset* for short) $(P, \leq)$ is a set P together with a binary relation $\leq$ satisfying: for all $p, q, r \in P$

i) $p \leq p$ (reflexivity),

ii) if $p \leq q$ and $q \leq p$, then $p = q$ (antisymmetry),

iii) if $p \leq q$ and $q \leq r$, then $p \leq r$ (transitivity).

Often we call P the poset and omit the binary relation $\leq$ in the notation if there is no risk of confusion. The binary relation $\leq$ is said to be a *partial order* on P. A map $\phi\colon P \to Q$ between two posets $(P, \leq_P)$, $(Q, \leq_Q)$ is called *order-preserving*, if $p \leq_P p'$ implies $\phi(p) \leq_Q \phi(p')$ for any $p, p' \in P$. Two posets P, Q are called *isomorphic*, in which case we write $P \cong Q$, if there exists an order-preserving bijection $\phi\colon P \to Q$ whose inverse is order-preserving. The *product* $P \times Q$ *of two posets* P, Q is defined on the set $\{(p, q)\colon p \in P, q \in Q\}$ by letting $(p, q) \leq (p', q')$ in $P \times Q$ if and only if $p \leq p'$ in P and $q \leq q'$ in Q. For $p, q \in P$ an *upper bound* of p and q is an element $r \in P$ such that $p, q \leq r$. A *least upper bound* of p and q is an upper bound $s \in P$ of p and q such that for any upper bound r of p and q one has $s \leq r$. Clearly, if a least upper bound of two elements p and q exists, it is unique. A greatest lower bound is defined in complete analogy. A *lattice* L is a poset with the property that any two of its elements have a least upper bound and a greatest lower bound. It is well-known that $\mathcal{P}_{[n]}$ together with the relation $\leq$, defined by $\pi \leq \rho$ if and only if ρ can be obtained by merging some blocks of π, is a lattice, the so-called *partition lattice*. For $\pi, \rho \in \mathcal{P}_{[n]}$ with $\pi \leq \rho$ we call the set $[\pi, \rho] \coloneqq \{\sigma \in \mathcal{P}_{[n]} \colon \pi \leq \sigma \leq \rho\}$ an *interval*. For any two sets A and $B \subseteq A$ and a partition $\pi \in \mathcal{P}_A$ we call $\pi|_B \coloneqq \{C \cap B \colon C \in \pi, C \cap B \neq \emptyset\} \in \mathcal{P}_B$ the *restriction of* π *to* B. We will make repeated use of the isomorphism

$$[\pi, \rho] \cong \bigtimes_{B \in \rho} \mathcal{P}_{\pi|_B} \cong \bigtimes_{B \in \rho} \mathcal{P}_{\#\pi|_B}, \tag{2.1}$$

where the right hand side in (2.1) denotes the product of lattices $\mathcal{P}_{\pi|_B}$ as defined earlier, cf. Example 3.10.4 in [70]. For more information on posets in general, and the partition lattice in particular, the reader is referred to [70]. Evidently, we have $\mathcal{P}_{[n]} = [\Delta_{[n]}, \{[n]\}]$, where for any set A we let $\Delta_A \coloneqq \{\{a\} \colon a \in A\}$ be the partition of A into singletons. Moreover, we write Δ_n as a shorthand for $\Delta_{[n]}$.

2.2 Bell polynomials and combinatorics

Various Bell polynomials of specific parameters appear repeatedly in this thesis, for example, Bell numbers and Stirling numbers of different kinds. Once we have defined the Bell polynomials, this should not be too surprising, given the exchangeability property of the coalescents that we will see later. In this section we closely follow Pitman's lecture notes [61] and the reader is referred to these notes as well as to Comtet's book [22] for further information on Bell polynomials.

Definition 2.1 (Bell polynomials)**.** For any two sequences $v_\bullet = (v_k)_{k \in \mathbb{N}}$ and

$w_\bullet = (w_k)_{k\in\mathbb{N}}$ denote by

$$B_n(v_\bullet, w_\bullet) \coloneqq \sum_{k=1}^{n} v_k B_{n,k}(w_\bullet) \tag{2.2}$$

the *nth complete Bell polynomial* (associated with $(v_\bullet, w_\bullet)$), where

$$B_{n,k}(w_\bullet) \coloneqq \sum_{\pi\in\mathcal{P}_{[n],k}} \prod_{B\in\pi} w_{\#B} \tag{2.3}$$

denotes the (n,k)*th partial Bell polynomial* (associated with $w_\bullet$).

Remark 2.2 (Bell polynomials and composite structures)**.** For sequences $v_\bullet = (v_k)$, $w_\bullet = (w_k)$ of non-negative integers, the nth Bell polynomial has a straightforward combinatorial interpretation. Let V denote a construction that associates with each finite set F_n of $n = \#F_n$ elements a set $V(F_n)$ of V-structures on F_n with $v_n = \#V(F_n)$ elements. Likewise, W is a species of combinatorial structures that assigns w_n W-structures to any finite set F_n of n elements. Then $(V \circ W)(F_n)$ denotes the composite structure on F_n, which contains as elements all ways to partition F_n into blocks $\{A_1, \dots, A_k\}$, $k \in [n]$, assign this collection of blocks a V-structure, and assign each block A_l a W-structure. Then, for each set F_n of cardinality n, we have

$$\#(V \circ W)(F_n) = B_n(v_\bullet, w_\bullet),$$

and $B_{n,k}(w_\bullet)$ counts the number of ways to partition F_n into k blocks and assign each block a W-structure. For more on species of combinatorial structures the reader is referred to [13].

Examples 2.3. (i) The number of permutations of $n \in \mathbb{N}$ with $k \in \mathbb{N}$ cycles is given by

$$\begin{bmatrix} n \\ k \end{bmatrix} := B_{n,k}((\bullet - 1)!). \tag{2.4}$$

These quantities are known as the *(signless) Stirling numbers of the first kind*, or *Stirling cycle numbers*. We call the quantities

$$s(n,k) := (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix} \tag{2.5}$$

the *signed Stirling numbers of the first kind*. By convention, we let $\left[\begin{smallmatrix} n \\ 0 \end{smallmatrix}\right] = 0$ and $\left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right] = 1$.

(ii) The number $\#\mathcal{P}_{[n],k}$ of partitions π of $[n] \in \mathbb{N}$ with $\#\pi = k \in \mathbb{N}$ blocks is given by

$$\begin{Bmatrix} n \\ k \end{Bmatrix} := B_{n,k}(\mathbf{1}), \tag{2.6}$$

where $\mathbf{1}$ is a shorthand for the sequence $(1, 1, \ldots)$. These quantities are known as the *Stirling numbers of the second kind*, or *Stirling partition numbers*. By convention, we let $\left\{\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right\} = 1$.

(iii) The number of partitions π of $[n]$ with k *linearly ordered blocks*, i.e. in each block of π the elements are ordered (according to their absolute value, for instance), is given by $B_{n,k}(\bullet!)$, where $\bullet!$ is a shorthand for the sequence $(k!)_{k \geq 1}$.

It can be shown that

$$\left\lfloor {n \atop k} \right\rfloor := B_{n,k}(\bullet!) = \binom{n}{k}\frac{(n-1)!}{(k-1)!}. \tag{2.7}$$

The numbers $\left\lfloor {n \atop k} \right\rfloor$ are known as the *(signless) Lah numbers* or *Stirling numbers of the third kind.*

(iv) For $n \in \mathbb{N}$ let B_n denote the number $\#\mathcal{P}_{[n]}$ of partitions of $[n]$. From the definition of the Bell polynomials and the Stirling partition numbers we have

$$B_n = \sum_{k=1}^{n} \left\{ {n \atop k} \right\} = B_n(\mathbf{1}, \mathbf{1}).$$

We refer to B_n as the nth *Bell number.*

(v) More generally, for distinct $\alpha, \beta \in \mathbb{R}$ and non-negative integers $n, k \in \mathbb{N}_0$ we define the *generalized Stirling numbers* $S^{\alpha,\beta}_{n,k}$ implicitly via

$$x^{\underline{n,\alpha}} = \sum_{k=0}^{n} S^{\alpha,\beta}_{n,k} x^{\underline{k,\beta}}, \tag{2.8}$$

where

$$x^{\underline{n,\alpha}} := x(x-\alpha)\cdots(x-(n-1)\alpha) = \prod_{j=0}^{n-1}(x-j\alpha) = \alpha^n (x/\alpha)^{\underline{n}} \tag{2.9}$$

denotes the nth *descending factorial power with increment α.* For the *nth descending factorial power* $x^{\underline{n,1}}$ we shortly write $x^{\underline{n}}$. Notice that $s(n,k) = S^{1,0}_{n,k}$ and $\left\{ {n \atop k} \right\} = S^{0,1}_{n,k}$. This is but one of several different generalizations of Stirling numbers used in the literature. Recall that the generalized Stirling numbers may be written as

$$S_{n,k}^{\alpha,\beta} = \sum_{j=k}^{n} \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix} \alpha^{n-j} \beta^{j-k}, \tag{2.10}$$

or $S_{n,k}^{\alpha,\beta} = B_{n,k}((\beta-\alpha)^{\underline{\bullet-1,\alpha}})$, see [61, pp. 20, 21].

2.3 Multiple merger coalescent processes

The study of coalescent processes goes back at least to Kingman (1982). In [44] he showed that the genealogy of the rescaled Wright-Fisher population model, and other similar models, can be approximated by a Markov process Π (with state space $\mathcal{P}_{\mathbb{N}}$), now known as Kingman's coalescent. In population genetics, the restriction Π^n of Π to $[n]$ customarily models the genealogy of a sample of size n drawn from a large population of haploid individuals. It is the most celebrated model in population genetics besides the Wright-Fisher model. Since its introduction it has been rediscovered in a variety of different contexts from interacting particle systems in statistical physics, e.g. [3], to Smoluchowski's coagulation equations, e.g. [5] to Hastings-Levitov models of planar random growth [38].

However, when the reproduction mechanism of the population model allows for occasional large offspring numbers, large meaning on the order of total population size, the genealogy will no longer be governed by Kingman's coalescent (put in more probabilistic terms: this population model is not in the range of attraction of Kingman's coalescent). Möhle and Sagitov [54] showed that the genealogies of models of randomly mating neutral populations of constant size, also known as Cannings models [20], [21], are coalescent processes that in general exhibit simultaneous multiple mergers. These coalescent processes are known as Ξ coalescents, sometimes called exchangeable coalescents, and were also introduced and studied by Schweinsberg [65], independently of Möhle and Sagitov.

An important subclass of Ξ coalescents, the multiple merger coalescents, or Λ coalescents, has attracted the increasing attention of researchers in recent years. As their name suggests, these processes only allow for multiple (but not simultaneous) mergers. They were introduced independently by Donnelly and Kurtz [24], Pitman [60] and Sagitov [64]. Multiple merger coalescents are the main object of study in this thesis and are now defined formally.

Definition 2.4 (Multiple merger coalescent)**.** For any finite measure Λ on the unit interval there exists a (unique in law) Markov process Π with state space $\mathcal{P}_{\mathbb{N}}$, such that for each $n \in \mathbb{N}$ the restriction Π^n of Π to $[n] := \{1, \ldots, n\}$ is a continuous-time Markov chain with initial state $\Delta_{[n]}$, absorbing state $\{[n]\}$, and the following dynamics: when Π^n has b blocks, any $2 \le k \le b$ specific blocks merge into a single block at rate $\lambda_{b,k} := \int_0^1 x^{k-2}(1-x)^{b-k}\Lambda(dx)$. The process Π is called the *multiple merger coalescent process corresponding to* Λ, or Λ *coalescent*, cf. [60]. We refer to Π^n as the Λ *n-coalescent*.

In population genetics, the interpretation of Π^n is that two individuals $i, j \in \{1, \ldots, n\}$ in a sample of size n drawn from a large population of haploid individuals at time 0 have a common ancestor at time $-t$ if and only if $\{i, j\}$ is contained in a block of $\Pi^n(t)$. For convenience of notation we let $\Pi^1(t) := \{\{1\}\}$ for all $t \ge 0$.

We denote by $0 = S_0 < S_1 < \cdots < S_{\#J}$ the *jump times of* Π^n with $\#J$ the *total number of jumps*, which in general is random. Since the so-called *ancestral process* or *block counting process* $N^n := \{N^n(t), t \ge 0\}$, where $N^n(t) := \#\Pi^n(t)$ counts the total number of blocks in $\Pi^n(t)$, is a death process starting in n, we have $\#J \leq n-1$ almost surely. For convenience of notation let $S_k := S_{\#J}$ whenever $k > \#J$. Let $\tau_k := S_{n-k+1} - S_{n-k}$ denote the *waiting time after the* k*th jump of* Π^n. Moreover, let $J = \{J_k, k = 1, \ldots, n\}$ denote the *jump chain of* Π^n, where $J_k := \Pi^n(S_{n-k})$ denotes the *state of* Π^n *after its* $(n-k)$*th jump*. Let $\lambda_b := \sum_{k=2}^{b} \binom{b}{k} \lambda_{b,k}$ denote the *total rate in any state with* $b = 2, \ldots, n$ *blocks*.

Examples 2.5. 1. Choosing $\Lambda = \delta_0$, the Dirac measure with mass 1 in 0, we recover the rates

$$\lambda_{m,k} = \begin{cases} 1 & \text{if } k = 2, \\ 0 & \text{otherwise} \end{cases} \qquad (2 \leq k \leq m)$$

of the so-called *Kingman coalescent*.

2. For $\Lambda = \delta_1$, the Dirac measure with mass 1 in 1, we obtain the rates

$$\lambda_{m,k} = \begin{cases} 1 & \text{if } k = m, \\ 0 & \text{otherwise} \end{cases} \qquad (2 \leq k \leq m).$$

The corresponding coalescent process Π is called the *star-shaped coalescent*, since there occurs only one jump, in which all individuals merge, and the corresponding genealogical tree has the shape of a star.

3. For Λ the Lebesgue measure on $[0, 1]$, the rates are given by

$$\lambda_{m,k} = \frac{(k-2)!(m-k)!}{(m-1)!} \qquad (2 \leq k \leq m). \tag{2.11}$$

The corresponding coalescent process Π is known as the *Bolthausen-Sznitman* *coalescent* and was discovered by Bolthausen and Sznitman [17] in the context of spin glasses.

4. Take Λ to be the Beta(a, b) *distribution* with density

$$\Lambda(dx) = B(a,b)^{-1} x^{a-1}(1-x)^{b-1} dx \quad (x \in (0,1)), \tag{2.12}$$

where for $a, b > 0$ the *Beta function* with parameters a, b is defined as

$$B(a,b) := \int_0^1 x^{a-1}(1-x)^{b-1} dx. \tag{2.13}$$

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

$$\lambda_{m,k} := \int_0^1 x^{k-2}(1-x)^{m-k} \Lambda(dx) = \frac{B(k-2+a, m-k+b)}{B(a,b)} = \frac{a^{\overline{k-2}} b^{\overline{m-k}}}{(a+b)^{\overline{m-2}}}, \tag{2.14}$$

where for any $x \in \mathbb{R}$ and $n \in \mathbb{N}$ we denote by $x^{\overline{n}} := x(x+1)\cdots(x+n-1)$ the *rising factorial power.* We used here the identity

$$B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b) \quad (a, b > 0),$$

see [6, Theorem 1.1.4], where for a positive real number $a \in (0, \infty)$ the *Gamma function* is defined as

$$\Gamma(a) := \int_0^1 e^{-x} x^{a-1} dx, \tag{2.15}$$

and may be meromorphically continued to the entire complex plane. Since $\mathit{\Gamma}(a + 1) = a\mathit{\Gamma}(a)$ for $a \geq 0$, we have

$$\mathit{\Gamma}(a+n) = a^{\overline{n}}\mathit{\Gamma}(a) \quad (n \in \mathbb{N}).$$

The corresponding coalescent process $\mathit{\Pi}$ is known as the *beta*(a, b) *coalescent*.

2.3.1 Some properties of multiple merger coalescents

We collect some further notions and properties of multiple merger coalescent processes for later use.

Definition 2.6 (Finite exchangeable partition)**.** A random element $\mathit{\Sigma}$ in $\mathcal{P}_{[n]}$ is called an *exchangeable partition* if and only if for any $l \in [n]$ there exists a symmetric function $\mathrm{p}\colon \mathbb{N}^l \to [0, 1]$, such that

$$\mathbb{P}\left\{\mathit{\Sigma} = \{B_1, \ldots, B_l\}\right\} = \mathrm{p}(\#B_1, \ldots, \#B_l) \tag{2.16}$$

for any partition $\{B_1, \ldots, B_l\} \in \mathcal{P}_{[n]}$. Here symmetric means that $\mathrm{p}(n_1, \ldots, n_l) = \mathrm{p}(n_{\sigma 1}, \ldots, p_{\sigma l})$ for any permutation σ of $\{1, \ldots, l\}$, and p is called the *exchangeable partition probability function.* In our notation, we suppress the dependence of p on l.

It should be clear from the definition of $\mathit{\Pi}$, Definition 2.4, that for any $t \geq 0$, the random partition $\mathit{\Pi}(t)$ of $\{1, \ldots, n\}$ is exchangeable.

Coming down from infinity

A Λ coalescent Π is said to *come down from infinity (CDI)* if $\mathbb{P}\{\#\Pi(t) < \infty\} = 1$ for all $t > 0$ and is said to *stay infinite* if $\mathbb{P}\{\#\Pi(t) = \infty\} = 1$ for all $t > 0$. Pitman [60, Proposition 23] showed the dichotomy that if Λ does not charge 1, i.e. $\Lambda(\{1\}) = 0$, then the corresponding coalescent either comes down from infinity or stays infinite almost surely. Schweinsberg [66] showed that a Λ coalescent that does not charge 1 comes down from infinity if and only if

$$\sum_{n\geq 2}(\gamma_n^{(1)})^{-1} < \infty, \tag{2.17}$$

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.

Bolthausen-Sznitman coalescent and random recursive tree

We present a beautiful construction due to Goldschmidt and Martin [33] of the Bolthausen-Sznitman n-coalescent Π^n via a Markovian cutting procedure applied to the random recursive tree. We will later use this construction in our arguments, e.g. in Section 3.2.2 and Section 3.3.3. Let us recall the notion of a random recursive tree, where we closely follow Goldschmidt and Martin [33]. We call a tree on n nodes labelled by $1, 2, \ldots, n$ an *increasing tree* if the root has label 1 and the labels in any path from the root to another node are increasing. Figure 2.1 shows all 6 increasing trees on $\{\{1\}, \{2\}, \{3\}, \{4\}\}$. If we have an increasing tree on $n-1$ nodes, we obtain an increasing tree on n nodes by adding a node labelled n and attaching it by an edge to one of the nodes in the given tree. Starting from the tree that only consists of the root node 1 this gives an explicit construction of all

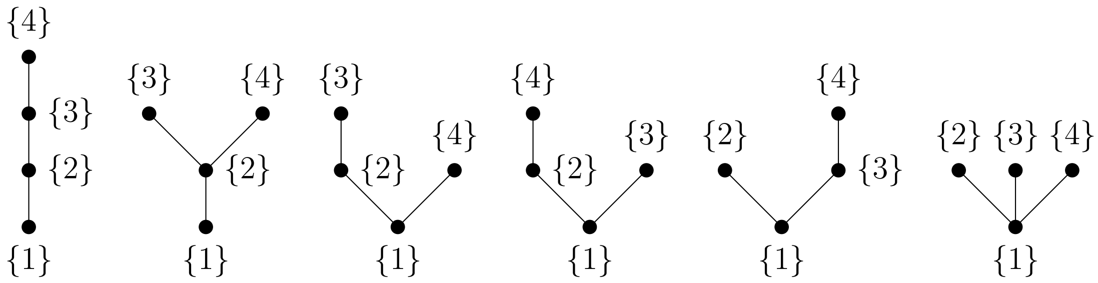

Figure 2.1: All 6 increasing trees on $\{\{1\},\{2\},\{3\},\{4\}\}$. Only the third and the sixth tree contain $\{\{1,2,3\},\{4\}\}$.

increasing trees on n nodes. Consequently, there are $(n-1)!$ increasing trees on n nodes. A *random recursive tree* is a tree chosen uniformly at random from all increasing trees on n nodes. An explicit construction of a random recursive tree on n nodes is the following. Start with the root node labelled 1. If the tree currently has k nodes, choose one of these nodes uniformly at random and attach to it node $k+1$ by an edge. Stop after attaching node n.

For any partition $\pi \in \mathcal{P}_{[n]}$ an *increasing tree on* π is a tree with $\#\pi$ nodes that are labelled by the blocks in π such that the labels in any path from the root to another node are increasing with respect to their least element. We denote a random recursive tree on π by $\mathfrak{T}_\pi$.

Crucial to the construction of the Bolthausen-Sznitman coalescent via random recursive trees is the following cutting procedure. When a tree $\mathscr{T}$ on π is given, a *cut* is performed by picking an edge, removing the subtree above this edge (here we picture trees as they grow in nature: from the root at the bottom to the leaves at the top) and adding the labels of this subtree to the labels of the node below the edge. For a simple tree the cutting procedure is depicted in Figure 2.2. We denote by $c\mathscr{T}$ the tree obtained by cutting $\mathscr{T}$ at an edge chosen uniformly at random. A

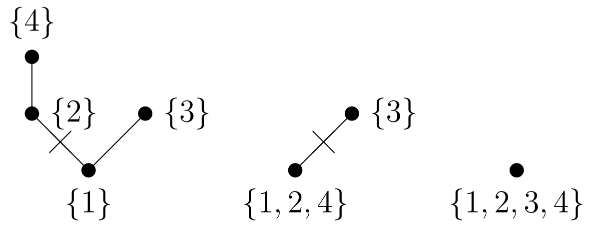

Figure 2.2: An increasing tree on $\{\{1\},\{2\},\{3\},\{4\}\}$ cut down successively to a tree on one node $\{1,2,3,4\}$.

striking result then is Proposition 2.1 in Goldschmidt and Martin [33], which states that after cutting a random recursive tree $\mathfrak{T}_\pi$ on π at an edge chosen uniformly at random, the new tree $c\mathfrak{T}_\pi$ is a random recursive tree on the new label set (which is again a partition of $[n]$).

For $\pi \in \mathcal{P}_{[n]}$ define a time-homogeneous continuous-time Markov chain $\mathcal{R}_\pi :=\{\mathcal{R}_\pi(t), t \geq 0\}$ with values in the set of random recursive trees as follows. The initial state $\mathcal{R}_\pi(0)$ is the random recursive tree $\mathfrak{T}_\pi$ on π. The process $\mathcal{R}_\pi$ evolves according to the following dynamics. Suppose $\mathcal{R}_\pi$ is in state $\mathscr{T}$. If $\mathscr{T}$ has only one vertex, do nothing. Otherwise, attach to each edge in $\mathscr{T}$ an exponential 1 clock, all clocks being independent. When the first clock rings, cut $\mathscr{T}$ at the associated edge to obtain the next state of $\mathcal{R}_\pi$. Proposition 2.2 of Goldschmidt and Martin [33] establishes that the Bolthausen-Sznitman n-coalescent $\mathit{\Pi}^n$ is equal in distribution to

$$\{p(\mathcal{R}_{\Delta_n}(t)), t \geq 0\},$$

where $p(\mathscr{T}) := \pi$ for any tree $\mathscr{T}$ on the label set π.

2.3.2 Coalescent tree

Often, if the Λ coalescent Π comes down from infinity, it is implicitly identified with a random tree. These seems so natural an identification, that most of the time it is not even mentioned in the literature, and we will also make it continuously in this thesis. However, this identification can be made precise via the notion of $\mathbb{R}$-trees. In order to do this, we closely follow the lecture notes of Evans [30] to which we refer the reader for further information on random $\mathbb{R}$-trees.

Real tree induced by coalescing partitions

Consider a non-decreasing (with respect to the partial order $\leq$ on $\mathcal{P}_{\mathbb{N}}$), right-continuous function $\rho\colon [0,\infty) \to \mathcal{P}_{\mathbb{N}}$ with $\rho(0) = \Delta_{\mathbb{N}}$ and $\rho(t) = \{\mathbb{N}\}$ for t sufficiently large, and recall that $\mathcal{P}_{\mathbb{N}}$ is endowed with the discrete topology. Then ρ encodes the evolution of coalescing partitions, just as any path $t \mapsto \Pi(t)(\omega)$, $\omega \in \Omega$, of a coalescent Π that comes down from infinity, where $(\Omega, \mathcal{F}, \mathbb{P})$ is the probability space underlying Π. The tree corresponding to ρ is encoded by the set of nodes

$$\mathcal{T} := \{(t,B)\colon t \in [0,\infty), B \in \Pi(t)\}$$

together with the metric d that we now define. For any two points $(t,A),(s,B) \in \mathcal{T}$ let

$$\begin{aligned} m((t,A),(s,B)) := \inf\{u > s \vee t\colon & \text{ both } A \text{ and } B \text{ are} \\ & \text{subsets of a common block in } \Pi(u)\} \end{aligned}$$

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

$$d((t, A), (s, B)) := (m((t, A), (s, B)) - s) + (m((t, A), (s, B)) - t)$$

is an $\mathbb{R}$-tree. This is done formally in Example 3.41 of [30]. Informally, d yields the genealogical distance between any two points in $\mathcal{T}$.

If we interpret $\mathcal{T}$ as a genealogical tree, $(s, B) \in \mathcal{T}$ means that individual B is alive at time s, and if for two points $(t, A), (s, B) \in \mathcal{T}$ we have that $s \leq t$ and $B \subsetneq A$, then A is interpreted as an ancestor of B alive at time t.

This yields a pathwise construction of the random tree $\mathcal{T}$ corresponding to Π, provided that Π comes down from infinity. We refer to $\mathcal{T}$ as the *coalescent tree*. In general, Π^n can be identified in complete analogy with an $\mathbb{R}$-tree $\mathcal{T}^n$ on n leaves labelled $1, \dots, n$. If Π comes down from infinity, $\mathcal{T}^n$ corresponds to the subtree in $\mathcal{T}$ spanned by the leaves $1, \dots, n$.

We will also be interested in modeling mutations that may affect individuals in the sample, and that we encode in the coalescent process. The corresponding process is called a Λ coalescent with mutation. It is best introduced via the corresponding coalescent tree.

Definition 2.7 (Λ coalescent with mutation)**.** For $\omega \in \Omega$ consider the coalescent tree $\mathcal{T}(\omega)$ corresponding to $\Pi(\omega) = \{\Pi(t)(\omega), t \geq 0\}$. Run a homogeneous Poisson process N with parameter $r > 0$, independently of $\Pi(\omega)$ (respectively $\mathcal{T}(\omega)$), along the branches of $\mathcal{T}(\omega)$, and at each jump of N place a mark at the corresponding position of the branch. Each mark is interpreted as a mutation that affects all individuals (nodes) supported by the corresponding branch. We call this process the Λ *coalescent with mutation.* We call r the *mutation rate.*

In this chapter we have introduced the Bell polynomials and have seen that some of their special cases agree with various kinds of Stirling numbers. We defined the multiple merger coalescents (respectively n-coalescents) and introduced some notation for the description of their state space $\mathcal{P}_{\mathbb{N}}$ (respectively $\mathcal{P}_{[n]}$), the partition lattice. We defined what it means for a coalescent to come down from infinity and made precise the identification of a CDI coalescent with the corresponding coalescent tree.

Chapter 3

Spectral decompositions

The spectral decomposition of the generator of a Markov process is of interest in its own right, but also provides a route to calculating further properties of the process under consideration. In this Chapter we work out for both the Kingman and the Bolthausen-Sznitman coalescent spectral decompositions for the generator of Π^n, respectively the generator of the corresponding ancestral process, or block counting process.

As applications we obtain explicit formulae for the transition probabilities, Corollaries 3.3 and 3.17, respectively Green's functions and hitting probabilities, Corollaries 3.5 and 3.18. In the case of the Bolthausen-Sznitman coalescent we obtain closed-form expressions for the factorial moments, see (3.7), the cumulative distribution function, see (3.8), and the Laplace transform, see (3.9), of the absorption time T_n of Π^n.

The transition probabilities for the Bolthausen-Sznitman coalescent and the hitting probabilities for the Kingman n-coalescent have been provided before by Bolthausen and Sznitman, respectively Kingman. All other results are new (unless stated otherwise).

3.1 Ancestral process of the Kingman coalescent

A number of authors have studied N^n in Kingman's case extensively, and we mention this here for the sake of completeness. An overview of the properties of N^n is given in Tavarés lecture notes [72]. Somewhat more generally, we give a spectral decomposition of the generator of a pure death process starting from a finite number of individuals, which obviously includes Kingman's n-coalescent. The following Lemma 3.1 has been derived in [51, Hilfssatz 6.1.6].

Lemma 3.1 (Spectral decomposition of a pure death process)**.** *Let $n \in \mathbb{N}$ and let $X := \{X(t), t \geq 0\}$ be a pure death process with state space $\{1, \ldots, n\}$ and pairwise distinct death rates $d_1, \ldots, d_n$. Then, the transition probabilities $p_{ij}(t) := \mathbb{P}\{X(t) = j \,|\, X(0) = i\}$ are given by*

$$p_{ij}(t) \;=\; \sum_{k=j}^{i} e^{-d_k t} r_{ik} l_{kj}, \qquad i, j \in \{1, \ldots, n\}, \tag{3.1}$$

where the $n \times n$ matrices $R = (r_{ij})$ and $L = (l_{ij})$ are lower left triangular and defined via $r_{ij} := l_{ij} := 0$ for $i < j$ and

$$r_{ij} \;:=\; \prod_{l=j+1}^{i} \frac{d_l}{d_l - d_j} \qquad \text{and} \qquad l_{ij} \;:=\; \prod_{l=j}^{i-1} \frac{d_{l+1}}{d_l - d_i} \qquad \text{for } i \geq j.$$

Here and elsewhere we agree on the convention that empty products equal 1 and empty sums equal 0.

Proof. With some effort it can be checked that $RL = I$, where $I = (\delta_{ij})_{1 \leq i,j \leq n}$

denotes the $n \times n$ unit matrix. Let $Q = (q_{ij})_{1 \le i,j \le n}$ denote the generator matrix of X, i.e. $q_{ii} := -d_i$, $q_{i,i-1} := d_i$ and $q_{ij} = 0$ otherwise. Furthermore, let D denote the diagonal matrix with entries $d_{ij} := -d_i$ for $i = j$ and $d_{ij} := 0$ otherwise. It is readily checked that $RD = QR$, and, hence, $RDL = Q$. The transition probabilities $p_{ij}(t)$ of the process X are now obtained from the spectral decomposition $P(t) := e^{tQ} = e^{tRDL} = R(e^{tD})L$ of the transition matrix $P(t)$ as $p_{ij}(t) = \sum_{k=1}^{n} e^{-d_k t} r_{ik} l_{kj} = \sum_{k=j}^{i} e^{-d_k t} r_{ik} l_{kj}$. $\square$

Plugging in the transition rates $q_{ij} = i(i-1)/2$ for Kingman's coalescent, we find $d_{ij} = -i(i-1)/2$ for $i = j$ and $d_{ij} = 0$ otherwise, and for $i \ge j$

$$
\begin{aligned}
r_{ij} &= \prod_{k=j+1}^{i} \frac{q_k}{q_k - q_j} = \prod_{k=j+1}^{i} \frac{k(k-1)}{k(k-1) - j(j-1)} \\
&= \prod_{k=j+1}^{i} \frac{k(k-1)}{(k-j)(k+j-1)} = \frac{i!(i-1)!(2j-1)!}{j!(j-1)!(i-j)!(i+j-1)!},
\end{aligned}
$$

and

$$
\begin{aligned}
l_{ij} &= \prod_{k=j}^{i-1} \frac{q_{k+1}}{q_k - q_i} \\
&= \prod_{k=j}^{i-1} \frac{k(k+1)}{(k-i)(k+i-1)} = (-1)^{i-j} \frac{(i-1)!i!(i+j-2)!}{(j-1)!j!(i-j)!(2i-2)!}.
\end{aligned}
$$

The formula for r_{ij} asks for a combinatorial interpretation. In fact, we will see in Corollary 3.21 that the spectral decomposition of the generator of Π^n reveals the structure of both R and L more clearly.

3.2 Ancestral process of the Bolthausen-Sznitman coalescent

This subsection is a modified version of the publication [56] which was written in collaboration with Martin Möhle from the University of Tübingen.

A spectral decomposition for the generator and the transition probabilities of the block counting process N of the Bolthausen-Sznitman coalescent Π is derived. This decomposition is closely related to the generalized Stirling numbers, and in particular to the Stirling numbers of the first and second kind. The proof is based on generating functions and exploits a certain factorization property of the Bolthausen-Sznitman coalescent. As an application we derive a formula for the hitting probability $h(i, j)$ that N ever visits state j when started from state $i \geq j$. Moreover, explicit formulae are derived for the moments and the distribution function of the absorption time T_n of Π^n. We provide an elementary proof for the well known convergence of $T_n - \log\log n$ in distribution to the standard Gumbel distribution. It is shown that the speed of this convergence is of order $1/\log n$.

3.2.1 Introduction and results

Recall that the Bolthausen-Sznitman coalescent is the particular Λ coalescent Π with Λ being the uniform distribution on the unit interval $[0, 1]$ and that $N(t) = \#\Pi(t)$ denotes the number of blocks of $\Pi(t)$. It is well known that $N = \{N(t), t \geq 0\}$ is a Markovian process. Let $Q = (q_{ij})_{i,j\in\mathbb{N}}$ denote the generator of N. The ancestral process N makes a transition from i to j at rate $\binom{i}{i-j+1}\lambda_{i,i-j+1}$ if $i > j$ and at rate 0 otherwise. Consequently, Q is a lower left triangular matrix and, using (2.11), we obtain for its entries

$$\begin{aligned} q_{ij} = \binom{i}{i-j+1}\lambda_{i,i-j+1} &= \frac{i!}{(i-j+1)!(j-1)!}\frac{(j-1)!(i-j-1)!}{(i-1)!} \\ &= \frac{i}{(i-j)(i-j+1)}, \qquad \text{for } i > j, \end{aligned} \tag{3.2}$$

$q_{ii} = -\sum_{j=1}^{i-1} q_{ij} = 1-i$ and $q_{ij} = 0$ for $i < j$. Let $q_i := -q_{ii} = i-1$, $i \in \mathbb{N}$.

The main result of this chapter, Theorem 3.1, provides a spectral decomposition for Q, which is closely related to the classical Stirling numbers. For the definition of Stirling numbers see Section 2.1. The proof of Theorem 3.1 is given in Section 3.2.2.

Theorem 3.1. (Spectral decomposition of the generator) *The infinitesimal generator $Q = (q_{ij})_{i,j\in\mathbb{N}}$ of the block counting process of the Bolthausen-Sznitman coalescent has spectral decomposition $Q = RDL$, where $D = (d_{ij})_{i,j\in\mathbb{N}}$ is the diagonal matrix with entries $d_{ij} = 1-i$ for $i = j$ and $d_{ij} = 0$ for $i \neq j$, and $R = (r_{ij})_{i,j\in\mathbb{N}}$ and $L = (l_{ij})_{i,j\in\mathbb{N}}$ are lower left triangular matrices with entries*

$$r_{ij} \;=\; \frac{(j-1)!}{(i-1)!}\begin{bmatrix} i \\ j \end{bmatrix} \qquad \textit{and} \qquad l_{ij} \;=\; (-1)^{i+j}\frac{(j-1)!}{(i-1)!}\begin{Bmatrix} i \\ j \end{Bmatrix}, \qquad i,j \in \mathbb{N}. \tag{3.3}$$

In particular, $r_{i1} = r_{ii} = l_{ii} = 1$ and $l_{i1} = (-1)^{i-1}/(i-1)!$, $i \in \mathbb{N}$.

Remark 3.2. For alternative formulae for r_{ij} and l_{ij} we refer the reader to (3.21), (3.22) and (3.24). The first entries of R and L are provided in (3.25).

The next corollary provides the spectral decomposition of the transition matrix of the block counting process N of the Bolthausen-Sznitman coalescent.

Corollary 3.3. (Spectral decomposition of the transition matrix) *For all $t \in [0,\infty)$ the transition matrix $P(t) := (\mathbb{P}\{N(t) = j \mid N(0) = i\})_{i,j\in\mathbb{N}}$ of the block* *counting process* $\{N(t), t \geq 0\}$ *of the Bolthausen-Sznitman coalescent has spectral decomposition* $P(t) = Re^{tD}L$*, i.e.*

$$p_{ij}(t) \;:=\; \mathbb{P}\{N(t)=j\,|\,N(0)=i\} \;=\; \sum_{k=j}^{i} e^{-(k-1)t} r_{ik} l_{kj} \;=\; e^{t(1-i)} \frac{(j-1)!}{(i-1)!} S_{i,j}^{-e^{-t},1}, \tag{3.4}$$

$i, j \in \mathbb{N}$*, where* D *is the diagonal matrix defined in Theorem 3.1 and* $R = (r_{ij})_{i,j\in\mathbb{N}}$ *and* $L = (l_{ij})_{i,j\in\mathbb{N}}$ *are defined via* (3.3).

Remark 3.4. As a consequence of Corollary 3.3, the block counting process has resolvent (see, for example, Norris [59, p. 146])

$$R_\lambda := \int_0^\infty e^{-\lambda t} P(t)\,\mathrm{d}t = \int_0^\infty e^{-\lambda t} R e^{tD} L\,\mathrm{d}t = R \int_0^\infty e^{-\lambda t} e^{tD}\,\mathrm{d}t L = RD(\lambda)L,$$

$\lambda \in (0, \infty)$, where $D(\lambda)$ is the diagonal matrix with entries $d_{ii}(\lambda) = 1/(\lambda + i - 1)$, $i \in \mathbb{N}$.

As an application we provide a formula for the probability that the block counting process ever visits state $j \in \mathbb{N}$ when started from state $i \in \mathbb{N}$ with $i \geq j$.

Corollary 3.5. (Hitting probabilities) *The probability*

$$h(i,j) := \mathbb{P}\{N(t) = j \text{ for some } t \geq 0\,|\,N(0) = i\} \tag{3.5}$$

that the block counting process hits state $j \in \mathbb{N}$ *when started from state* $i \in \mathbb{N}$ *with* $i \geq j$ *is given by* $h(i, 1) = 1$ *and*

$$h(i,j) \;=\; (-1)^{i+j} \frac{(j-1)(j-1)!}{(i-1)!} \sum_{k=j}^{i} \frac{s(i,k)\left\{{k \atop j}\right\}}{k-1}, \qquad 2 \leq j \leq i.$$

Remark 3.6. Further formulae for $h(i,j)$ are provided in [55, Theorem 2.1 with $\alpha = 1$] and [58, Eq. (11)]. For the asymptotics of $h(i,j)$ as $i \to \infty$ we refer the reader to [37, Corollary 3.4] and [58, Theorem 1.1 and Corollary 1.3].

As a further application we provide in the following formulae for the moments, the distribution function and the Laplace transform of the *absorption time*

$$T_n := \inf\{t \geq 0 \colon \#\Pi^n(t) = 1\} \tag{3.6}$$

of Π^n, that is the time until Π^n reaches its absorbing state. In the biological context T_n is called the *time back to the most recent common ancestor* of a sample of size n.

Corollary 3.7. (Moments of the absorption time) *The absorption time T_n of the Bolthausen-Sznitman n-coalescent has moments*

$$\mathbb{E}T_n^j \;=\; \frac{j!}{(n-1)!}\sum_{k=2}^{n}(-1)^k \frac{\left[{n \atop k}\right]}{(k-1)^j}, \qquad n, j \in \mathbb{N}. \tag{3.7}$$

Remark 3.8. Note that (3.7) is simpler than the formula in [31, Proposition 1.1]. Some concrete values of the first and second moment of T_n are listed in [31, p. 403, Table 1]. For the asymptotics of $\mathbb{E}T_n^j$ as $n \to \infty$ we refer the reader to [31, Theorem 1.1].

Recall that for a positive real number $a \in (0,\infty)$ the Gamma function is defined as $\Gamma(a) := \int_0^1 e^{-x}x^{a-1}dx$.

Corollary 3.9. (Distribution and asymptotics of the absorption time) *The absorption time T_n of the Bolthausen-Sznitman coalescent started in a partition* *with n blocks has distribution function*

$$\mathbb{P}\{T_n \le t\} = \prod_{j=1}^{n-1} \frac{j - e^{-t}}{j} = \binom{n-1-e^{-t}}{n-1} = \frac{\Gamma(n-e^{-t})}{\Gamma(n)\Gamma(1-e^{-t})} \quad (t \in (0,\infty)), \tag{3.8}$$

and Laplace transform

$$\mathbb{E}e^{-\lambda T_n} \;=\; 1 + \frac{(-1)^{n-1}}{(n-1)!} \sum_{k=2}^{n} s(n,k) \frac{\lambda}{k-1+\lambda} \qquad (\lambda \in [0,\infty)). \tag{3.9}$$

In particular, $T_n - \log\log n \to T$ in distribution as $n \to \infty$, where T is standard Gumbel distributed with distribution function $F(t) := e^{-e^{-t}}$, $t \in \mathbb{R}$.

Remark 3.10. 1. The first equation in (3.8) is implicitly contained in [60, Theorem 14, Eq. (17) with $k=1$]. The convergence in distribution of $T_n - \log\log n$ to the standard Gumbel distribution is well known from the literature (see, for example, [33, Proposition 3.4] or [31, Corollary 1.2]). Our alternative proof of this convergence is based on the last expression in (3.8) and is rather elementary as it follows from a straightforward application of Stirling's formula for $\Gamma(x)$ as $x \to \infty$.

2. Let $\Psi(x) := (\mathrm{d}/\mathrm{d}x)(\log\Gamma(x)) = \Gamma'(x)/\Gamma(x)$ denote the digamma function (logarithmic derivative of the Gamma function). Taking the derivative with respect to t in (3.8) it is readily seen that for $n \ge 2$ the absorption time T_n has density

$$\begin{aligned} f_{T_n}(t) &= \frac{e^{-t}\Gamma(n-e^{-t})}{\Gamma(n)\Gamma(1-e^{-t})}\big(\Psi(n-e^{-t}) - \Psi(1-e^{-t})\big) \\ &= e^{-t}\mathbb{P}\{T_n \le t\} \sum_{k=1}^{n-1} \frac{1}{k-e^{-t}}, \qquad n \ge 2, t \in (0,\infty). \end{aligned}$$

Note that, for $n \ge 2$, $\mathbb{P}\{T_n \le t\} \sim t/(n-1)$ as $t \searrow 0$ and, hence, $\lim_{t\searrow 0} f_{T_n}(t) =$

$1/(n-1)$.

3. The fact that the distribution function (3.8) of the absorption time T_n is known explicitly can be further exploited. For example, it is readily checked that for all $x \in \mathbb{R}$, $\mathbb{P}\{T_n - \log\log n \le x\} - F(x) \sim -\gamma e^{-x} F(x)/\log n$ as $n \to \infty$, where $\gamma \approx 0.577216$ denotes Euler's constant. Thus, the speed of the convergence of $T_n - \log\log n$ to the Gumbel distribution is of order $1/\log n$.

For the star-shaped coalescent the spectral decomposition $Q = RDL$ can be readily calculated and is omitted here. Finding the spectral decomposition of the generator Q of the block counting process for other exchangeable coalescents, for example for the $\beta(a,b)$ coalescent with parameters $a, b > 0$, is an open problem. Note that the spectral decomposition $Q = RDL$ exists for any exchangeable coalescent, since the generator Q is lower left triangular. More precisely, R and L are recursively defined via the first equation in (3.12) and (3.13) respectively, where q_{ij} and q_i are the rates and total rates of the exchangeable coalescent.

3.2.2 Proofs

Before Theorem 3.1 will be proven, we mention two recursions for the Stirling numbers. These recursions will be used in the proof of Theorem 3.1.

Lemma 3.11. *The absolute Stirling numbers of the first kind satisfy the recursion*

$$\begin{bmatrix} i \\ j \end{bmatrix} = (i-1)! \sum_{k=j-1}^{i-1} \frac{\begin{bmatrix} k \\ j-1 \end{bmatrix}}{k!}, \qquad i, j \in \mathbb{N}, \tag{3.10}$$

whereas the Stirling numbers of the second kind satisfy the recursion

$$\left\{ {i \atop j} \right\} = \sum_{k=j-1}^{i-1} \binom{i-1}{k} \left\{ {k \atop j-1} \right\}, \qquad i,j \in \mathbb{N}. \tag{3.11}$$

Proof. Let us verify (3.10) by induction on i. Obviously, (3.10) holds for $i = 1$, since $\left[{1 \atop j} \right] = \delta_{j1} = \left[{0 \atop j-1} \right]$ for all $j \in \mathbb{N}$. The induction from i to $i+1$ works as follows. We have, by induction,

$$\begin{aligned} i! \sum_{k=j-1}^{i} \frac{\left[{k \atop j-1} \right]}{k!} &= i! \sum_{k=j-1}^{i-1} \frac{\left[{k \atop j-1} \right]}{k!} + \left[{i \atop j-1} \right] \\ &= i! \frac{\left[{i \atop j} \right]}{(i-1)!} + \left[{i \atop j-1} \right] = i \left[{i \atop j} \right] + \left[{i \atop j-1} \right] = \left[{i+1 \atop j} \right], \end{aligned}$$

by the well known recursion $\left[{i+1 \atop j} \right] = i\left[{i \atop j} \right] + \left[{i \atop j-1} \right]$, see [70, Lemma 1.3.6]. The induction is complete.

We verify (3.11) as follows. There are $\binom{i-1}{k}$ possibilities to choose a subset A of size k from $\{1,\ldots,i-1\}$ and there are $\left\{ {k \atop j-1} \right\}$ possibilities to partition this subset A into $j-1$ nonempty subsets. Together with the nonempty set $(\{1,\ldots,i-1\} \setminus A) \cup \{i\}$ one obtains, after summing over all possible values of k, the number $\left\{ {i \atop j} \right\}$ of partitions of $\{1,\ldots,i\}$ into j nonempty subsets. $\square$

Let us now turn to the proof of the spectral decomposition $Q = RDL$ (see Theorem 3.1). The following proof is based on generating functions and has the advantage that the solution (3.3) for R and L does not need to be known in advance in order to perform the proof. The solution (3.3) pops up naturally during the proof as the coefficients of appropriately chosen generating functions. Crucial for the proof is the factorization formula (3.14), which essentially goes back to the particular form of the rates q_{ij} in (3.2).

Proof. (of Theorem 3.1) Let $D = (d_{ij})_{i,j\in\mathbb{N}}$ be the diagonal matrix with entries $d_{ii} := q_{ii} = -q_i = 1-i$, $i \in \mathbb{N}$, and $d_{ij} := 0$ for $i \neq j$. Furthermore, let $R = (r_{ij})_{i,j\in\mathbb{N}}$ be the lower left triangular matrix with entries recursively defined for each $j \in \mathbb{N}$ via $r_{jj} := 1$ and

$$r_{ij} \;:=\; \frac{1}{q_i - q_j}\sum_{k=j}^{i-1} q_{ik}r_{kj} \;=\; \frac{1}{i-j}\sum_{k=j}^{i-1} q_{ik}r_{kj}, \qquad i \in \{j+1, j+2, \ldots\}. \quad (3.12)$$

Since $q_{ii} = -q_i$, $i \in \mathbb{N}$, we conclude that $r_{ij}q_{jj} = \sum_{k=j}^{i} q_{ik}r_{kj}$. Thus, the entries r_{ij} of R are defined such that $RD = QR$. Define $L := R^{-1}$. Then, the spectral decomposition $Q = RDL$ holds. Moreover, $DL = LQ$ and, hence, $q_{ii}l_{ij} = \sum_{k=j}^{i} l_{ik}q_{kj}$, $i, j \in \mathbb{N}$. Since $q_{ii} = -q_i$, $i \in \mathbb{N}$, we obtain for each $i \in \mathbb{N}$ the backward recursion $l_{ii} = 1$ and

$$l_{ij} \;=\; \frac{1}{q_j - q_i}\sum_{k=j+1}^{i} l_{ik}q_{kj} \;=\; \frac{1}{j-i}\sum_{k=j+1}^{i} l_{ik}q_{kj}, \qquad j = i-1, i-2, \ldots, 2, 1. \quad (3.13)$$

Let us verify by induction on i ($\geq j$) that

$$r_{ij} \;\leq\; \prod_{k=j+1}^{i} \frac{q_k}{q_k - q_j} \;=\; \prod_{k=j+1}^{i} \frac{k-1}{k-j} \;=\; \binom{i-1}{j-1}, \qquad i, j \in \mathbb{N}, i \geq j,$$

with the convention that empty products equal 1. For $i = j$ this inequality obviously holds, since $r_{jj} = 1$. The induction step from $\{j, \ldots, i-1\}$ to i ($> j$) works as follows. By (3.12) and by induction,

$$\begin{aligned} r_{ij} &= \frac{1}{q_i - q_j}\sum_{k=j}^{i-1} q_{ik} r_{kj} \ \le\ \frac{1}{q_i - q_j}\sum_{k=j}^{i-1} q_{ik} \prod_{l=j+1}^{k} \frac{q_l}{q_l - q_j} \\ &\le \frac{1}{q_i - q_j}\sum_{k=j}^{i-1} q_{ik} \prod_{l=j+1}^{i-1} \frac{q_l}{q_l - q_j} \ \le\ \frac{1}{q_i - q_j} q_i \prod_{l=j+1}^{i-1} \frac{q_l}{q_l - q_j} \ =\ \prod_{l=j+1}^{i} \frac{q_l}{q_l - q_j}, \end{aligned}$$

which completes the induction.

In order to solve the recursion (3.12) we exploit generating function techniques similar to those used in [31], [55] and [58]. Let $U := \{z \in \mathbb{C} : |z| < 1\}$ denote the open unit disc. For $j \in \mathbb{N}$ define the generating function $r_j : U \to \mathbb{C}$ via

$$r_j(z) := \sum_{i=j}^{\infty} r_{ij} z^i,$$

$z \in U$. Note that, for $j \in \mathbb{N}$ and $z \in U$,

$$|r_j(z)| \le \sum_{i=j}^{\infty} r_{ij} |z|^i \le \sum_{i=j}^{\infty} \binom{i-1}{j-1} |z|^j = |z|^j/(1-|z|)^j < \infty.$$

Thus, for each $j \in \mathbb{N}$, the function $r_j : U \to \mathbb{C}$ is well defined. Consider the modified function $f_j : U \to \mathbb{C}$ defined via $f_j(z) := \sum_{i=j}^{\infty} (i-j) i^{-1} r_{ij} z^i$, $z \in U$. Clearly, $|f_j(z)| \le \sum_{i=j}^{\infty} r_{ij} |z|^i = r_j(|z|) < \infty$, so $f_j : U \to \mathbb{C}$ is well defined. We have

$$\begin{aligned} f_j(z) = \sum_{i=j}^{\infty} r_{ij} z^i - j \sum_{i=j}^{\infty} \frac{r_{ij}}{i} z^i &= r_j(z) - j \int_0^z \sum_{i=j}^{\infty} r_{ij} s^{i-1}\, \mathrm{d}s \\ &= r_j(z) - j \int_0^z \frac{r_j(s)}{s}\, \mathrm{d}s. \end{aligned}$$

On the other hand, by the recursion (3.12), we obtain the factorization

$$\begin{aligned} f_j(z) &= \sum_{i=j+1}^{\infty}(i-j)r_{ij}\frac{z^i}{i} \;=\; \sum_{i=j+1}^{\infty}\sum_{k=j}^{i-1}q_{ik}r_{kj}\frac{z^i}{i} \;=\; \sum_{k=j}^{\infty}r_{kj}\sum_{i=k+1}^{\infty}\frac{q_{ik}}{i}z^i \\ &= \sum_{k=j}^{\infty}r_{kj}z^k\sum_{i=k+1}^{\infty}\frac{z^{i-k}}{(i-k)(i-k+1)} \\ &= \sum_{k=j}^{\infty}r_{kj}z^k\sum_{n=1}^{\infty}\frac{z^n}{n(n+1)} \;=\; r_j(z)a(z), \end{aligned} \tag{3.14}$$

where the function $a : U \to \mathbb{C}$ is defined via $a(z) := \sum_{n=1}^{\infty} z^n/(n(n+1)) = 1-(1-z)L(z)/z$, $z \in U$, with $L(z) := -\log(1-z)$, $z \in U$. Thus, r_j satisfies the integral equation $j\int_0^z r_j(s)/s\,\mathrm{d}s = (1-a(z))\,r_j(z)$. Differentiating with respect to z leads to $jr_j(z)/z = -a'(z)r_j(z)+(1-a(z))r_j'(z)$ or, equivalently, $(1-a(z))r_j'(z) = (a'(z)+j/z)r_j(z)$. Plugging in $1-a(z) = (1-z)L(z)/z$ and $a'(z) = L(z)/z^2-1/z$ leads to

$$r_j'(z) \;=\; \left(\frac{1}{z(1-z)}+\frac{j-1}{(1-z)L(z)}\right)r_j(z). \tag{3.15}$$

The solution of this homogeneous differential equation with initial conditions $r_j(0) = r_j'(0) = \cdots = r_j^{(j-1)}(0) = 0$ and $r_j^{(j)}(0) = j!$ is

$$r_j(z) \;=\; \frac{z}{1-z}(L(z))^{j-1}, \qquad j \in \mathbb{N}, |z| < 1. \tag{3.16}$$

In the following, for a power series $f(z) = \sum_{n=0}^{\infty} f_n z^n$, $[z^n]f(z) := f_n$ denotes the coefficient of z^n in the series expansion of f. By [1, p. 824], $(L(z))^j/j! = \sum_{k=j}^{\infty} z^k {k \brack j}/k!$. Thus, $[z^k](L(z))^j = j!{k \brack j}/k!$. For the coefficient $r_{ij} = [z^i]r_j(z)$ we therefore obtain

$$
\begin{aligned}
r_{ij} &= [z^i]r_j(z) = [z^i]\Big(\frac{z}{1-z}(L(z))^{j-1}\Big) = \sum_{k=0}^{i-1}[z^{i-k}]\frac{z}{1-z}\,[z^k](L(z))^{j-1}\\
&= \sum_{k=0}^{i-1}[z^k](L(z))^{j-1} = (j-1)!\sum_{k=j-1}^{i-1}\frac{\left[{k \atop j-1}\right]}{k!} = \frac{(j-1)!}{(i-1)!}\left[{i \atop j}\right] \qquad (3.17)
\end{aligned}
$$

by (3.10), which is the first formula in (3.3).

Let us now turn to $L := R^{-1}$. We have $(z, z^2, \ldots)R = (r_1(z), r_2(z), \ldots)$. Multiplying by L and noting that $RL = (\delta_{ij})_{i,j\in\mathbb{N}}$ it follows that $(z, z^2, \ldots) = (r_1(z), r_2(z), \ldots)L$. Thus, $z^j = \sum_{i=j}^{\infty} r_i(z)l_{ij} = z(1-z)^{-1}\sum_{i=j}^{\infty}(-\log(1-z))^{i-1}l_{ij}$, or, equivalently, $(1-z)z^{j-1} = \sum_{i=j}^{\infty}(-\log(1-z))^{i-1}l_{ij}$. Replacing z by $1-e^{-z}$ leads to $e^{-z}(1-e^{-z})^{j-1} = \sum_{i=j}^{\infty} l_{ij}z^{i-1}$. Thus,

$$
l_j(z) := \sum_{i=j}^{\infty} l_{ij}z^i = ze^{-z}(1-e^{-z})^{j-1}, \qquad j \in \mathbb{N}. \qquad (3.18)
$$

Let us extract the coefficient $l_{ij} = [z^i]l_j(z)$. We have

$$
\begin{aligned}
(1-e^{-z})^j &= \sum_{i=0}^{j}\binom{j}{i}(-e^{-z})^i = \sum_{i=0}^{j}\binom{j}{i}(-1)^i\sum_{k=0}^{\infty}\frac{(-zi)^k}{k!}\\
&= \sum_{k=0}^{\infty}\frac{z^k}{k!}(-1)^k\sum_{i=0}^{j}(-1)^i\binom{j}{i}i^k\\
&= \sum_{k=0}^{\infty}\frac{z^k}{k!}(-1)^{j+k}\sum_{i=0}^{j}(-1)^{j-i}\binom{j}{i}i^k = \sum_{k=0}^{\infty}\frac{z^k}{k!}(-1)^{j+k}j!\left\{{k \atop j}\right\},
\end{aligned}
$$

where the last equality follows from the explicit formula

$$
\left\{{k \atop j}\right\} = \frac{1}{j!}\sum_{i=0}^{j}(-1)^{j-i}\binom{j}{i}i^k, \qquad k, j \in \mathbb{N}_0, \qquad (3.19)
$$

for the Stirling numbers of the second kind, see [70, Equation (1.94a)]. Thus, $[z^k](1-e^{-z})^{j-1} = (-1)^{j+k-1}(j-1)!\left\{{k \atop j-1}\right\}/k!$ and, therefore,

$$\begin{aligned}
l_{ij} &= [z^i]l_j(z) \;=\; [z^i](ze^{-z}(1-e^{-z})^{j-1}) \;=\; \sum_{k=0}^{i-1}[z^{i-k}](ze^{-z})\,[z^k](1-e^{-z})^{j-1}\\
&= \sum_{k=j-1}^{i-1}\frac{(-1)^{i-1-k}}{(i-1-k)!}\frac{(-1)^{j+k-1}(j-1)!\left\{{k \atop j-1}\right\}}{k!}\\
&= (-1)^{i+j}(j-1)!\sum_{k=j-1}^{i-1}\frac{\left\{{k \atop j-1}\right\}}{(i-1-k)!k!}\\
&= (-1)^{i+j}\frac{(j-1)!}{(i-1)!}\sum_{k=j-1}^{i-1}\binom{i-1}{k}\left\{{k \atop j-1}\right\} \;=\; (-1)^{i+j}\frac{(j-1)!}{(i-1)!}\left\{{i \atop j}\right\}
\end{aligned}$$

by (3.11), which is the second formula in (4.21). The proof is complete. $\square$

Remark 3.12. 1. Let us provide some additional information on R and L. Plugging in the formula

$$s(i,j) \;=\; (-1)^{i+j}\frac{i!}{j!}\sum_{\substack{k_1,\ldots,k_j\in\mathbb{N}\\ k_1+\cdots+k_j=i}}\frac{1}{k_1\cdots k_j}, \qquad i,j\in\mathbb{N} \tag{3.20}$$

for the Stirling numbers of the first kind into the first formula in (3.3) leads to

$$r_{ij} \;=\; \frac{i}{j}\sum_{\substack{k_1,\ldots,k_j\in\mathbb{N}\\ k_1+\cdots+k_j=i}}\frac{1}{k_1\cdots k_j}, \qquad i,j\in\mathbb{N}. \tag{3.21}$$

In particular, $r_{i2} = (i/2)\sum_{k=1}^{i-1}1/(k(i-k)) = (1/2)\sum_{k=1}^{i-1}(1/k+1/(i-k)) = h_{i-1}$, $i\in\mathbb{N}$, where $h_0 := 0$ and $h_n := \sum_{k=1}^n 1/k$ denotes the nth harmonic number, $n\in\mathbb{N}$. Thus, $r_{i2}\sim\log i$ as $i\to\infty$. More generally, for arbitrary but fixed $j\in\mathbb{N}$ the absolute Stirling numbers of the first kind $\left[{i \atop j}\right]$ satisfy the asymptotics

$\left[{i \atop j}\right] \sim (i-1)!(\log i)^{j-1}/(j-1)!$ as $i \to \infty$, see for example [1, p. 824]. Thus, for each $j \in \mathbb{N}$ we conclude from (3.3) that $r_{ij} \sim (\log i)^{j-1}$ as $i \to \infty$. For each $j \in \mathbb{N}$ it follows from the middle expression in (3.17) that the sequence $(r_{ij})_{i \in \mathbb{N}}$ is monotone increasing in i. Plugging (3.19) into the second formula in (3.3) leads to

$$l_{ij} \quad = \quad (-1)^{i+j} \frac{1}{j(i-1)!} \sum_{k=0}^{j} (-1)^{j-k} \binom{j}{k} k^i, \qquad i, j \in \mathbb{N}. \tag{3.22}$$

Alternatively one may also plug in the formula

$$\left\{ {i \atop j} \right\} \quad = \quad \frac{i!}{j!} \sum_{\substack{k_1, \ldots, k_j \in \mathbb{N} \\ k_1 + \cdots + k_j = i}} \frac{1}{k_1! \cdots k_j!}, \qquad i, j \in \mathbb{N} \tag{3.23}$$

for the Stirling numbers of the second kind into the second formula in (3.3) leading to

$$l_{ij} \;=\; (-1)^{i+j} \frac{i}{j} \sum_{\substack{k_1, \ldots, k_j \in \mathbb{N} \\ k_1 + \cdots + k_j = i}} \frac{1}{k_1! \cdots k_j!}, \qquad i, j \in \mathbb{N}. \tag{3.24}$$

Note that formula (3.24) for l_{ij} has structural similarities with formula (3.21) for r_{ij}. These similarities point to the spectral decomposition of the partition valued Bolthausen–Sznitman n-coalescent, which is provided in Theorem 3.16.

2. Eqs. (3.21), (3.22) and (3.24) are useful to compute the entries of R and L

numerically. One obtains

$$R = \begin{pmatrix} 1 & & & & & \\ 1 & 1 & & & & \\ 1 & \frac{3}{2} & 1 & & & \\ 1 & \frac{11}{6} & 2 & 1 & & \\ 1 & \frac{25}{12} & \frac{35}{12} & \frac{5}{2} & 1 & \\ \vdots & & & & & \ddots \end{pmatrix} \quad \text{and} \quad L = \begin{pmatrix} 1 & & & & & \\ -1 & 1 & & & & \\ \frac{1}{2} & -\frac{3}{2} & 1 & & & \\ -\frac{1}{6} & \frac{7}{6} & -2 & 1 & & \\ \frac{1}{24} & -\frac{5}{8} & \frac{25}{12} & -\frac{5}{2} & 1 & \\ \vdots & & & & & \ddots \end{pmatrix}. \tag{3.25}$$

Proof. (of Corollary 3.3) Fix $t \in [0, \infty)$. By Theorem 3.1, $Q = RDL$, where R and $L := R^{-1}$ have entries (3.3). Therefore, the spectral decomposition of the transition matrix is $P(t) = e^{tQ} = e^{tRDL} = Re^{tD}L$. Thus, $p_{ij}(t) := \mathbb{P}\{N(t) = j \mid N_0 = i\} = \sum_{k=j}^{i} e^{-q_k t} r_{ik} l_{kj}$ for all $i, j \in \mathbb{N}$. The last equality in (3.4) follows from (3.3) and formula (2.10):

$$\begin{aligned} p_{ij}(t) &= \sum_{k=j}^{i} e^{-q_k t} r_{ik} l_{kj} = \frac{(j-1)!}{(i-1)!} \sum_{k=j}^{i} e^{-(k-1)t} \begin{bmatrix} i \\ k \end{bmatrix} \begin{Bmatrix} k \\ j \end{Bmatrix} \\ &= e^{(1-i)t} \frac{(j-1)!}{(i-1)!} \sum_{k=j}^{i} e^{t(i-k)} (-1)^{i-k} \begin{bmatrix} i \\ k \end{bmatrix} \begin{Bmatrix} k \\ j \end{Bmatrix} = e^{(1-i)t} \frac{(j-1)!}{(i-1)!} S_{i,j}^{-e^t,1}. \end{aligned}$$

□

Proof. (of Corollary 3.5) The Green's matrix $G = (g_{ij})_{i,j \in \mathbb{N}}$ of $\{N(t), t \geq 0\}$ is given by

$$G := \int_0^\infty P(t)\,\mathrm{d}t. \tag{3.26}$$

Using the spectral decomposition of $P(t)$ provided in Corollary 3.3 we obtain, for $2 \leq j \leq i$,

$$g_{ij} \;=\; \int_0^\infty p_{ij}(t)\,\mathrm{d}t \;=\; \sum_{k=j}^{i} r_{ik} l_{kj} \int_0^\infty e^{-(k-1)t}\,\mathrm{d}t \;=\; (-1)^{i+j}\frac{(j-1)!}{(i-1)!}\sum_{k=j}^{i}\frac{s(i,k)\left\{{k \atop j}\right\}}{k-1},$$

where the last equality follows from (4.21). We have $g_{ij} = h(i,j)/(q_j(1-f_j))$, where $h(i,j)$ is the probability of hitting j starting from i and f_j is the return probability for j. For $(N(t))_{t\geq 0}$ we evidently have $f_j = 0$ for all $j \geq 2$ and $q_j = j-1$, $j \in \mathbb{N}$. Hence,

$$h(i,j) \;=\; q_j g_{ij} \;=\; (j-1)(-1)^{i+j}\frac{(j-1)!}{(i-1)!}\sum_{k=j}^{i}\frac{s(i,k)\left\{{k \atop j}\right\}}{k-1}, \qquad 2 \leq j \leq i.$$

Clearly, $h(i,1) = 1$, since N^i reaches its absorbing state 1 almost surely. $\square$

Proof. (of Corollary 3.7) By Corollary 3.3, for all $t \in (0,\infty)$ and all $n \in \mathbb{N}$, $\mathbb{P}\{T_n > t\} = 1 - p_{n1}(t) = -\sum_{k=2}^{n} r_{nk} l_{k1} e^{-q_k t}$. Thus, for all $j \in \mathbb{N}$,

$$\mathbb{E}T_n^j \;=\; \int_0^\infty j t^{j-1}\mathbb{P}\{T_n > t\}\,\mathrm{d}t \;=\; -\sum_{k=2}^{n} r_{nk} l_{k1}\int_0^\infty j t^{j-1} e^{-q_k t}\,\mathrm{d}t \;=\; -\sum_{k=2}^{n} r_{nk} l_{k1}\frac{j!}{q_k^j}.$$

Plugging in the formulae $r_{nk} = (k-1)!\left[{n \atop k}\right]/(n-1)!$ and $l_{k1} = (-1)^{k-1}/(k-1)!$ from Theorem 3.1 and noting that $q_k = k-1$ the formula (3.7) for $\mathbb{E}T_n^j$, $n, j \in \mathbb{N}$, follows immediately. $\square$

Proof. (of Corollary 3.9) From Corollary 3.3 it follows that the absorption time T_n of Π^n has distribution function $\mathbb{P}\{T_n \le t\} = p_{n1}(t) = \sum_{k=1}^n r_{nk} l_{k1} e^{-q_k t}$. By (4.21),

$$r_{nk} l_{k1} \;=\; \frac{(k-1)!}{(n-1)!} \begin{bmatrix} n \\ k \end{bmatrix} \frac{(-1)^{k-1}}{(k-1)!} \;=\; \frac{(-1)^{n-1}}{(n-1)!} s(n,k).$$

Thus, T_n has distribution function

$$\begin{aligned} \mathbb{P}\{T_n \le t\} &= \frac{(-1)^{n-1}}{(n-1)!} \sum_{k=1}^n s(n,k)\, e^{-(k-1)t} = \frac{(-1)^{n-1}}{(n-1)!} e^t \sum_{k=1}^n s(n,k)(e^{-t})^k \qquad (3.27) \\ &= \frac{(-1)^{n-1}}{(n-1)!} e^t (e^{-t})_n = \prod_{j=1}^{n-1} \frac{j - e^{-t}}{j} = \binom{n-1-e^{-t}}{n-1} \\ &= \frac{\Gamma(n - e^{-t})}{\Gamma(n)\Gamma(1-e^{-t})}, \end{aligned}$$

$n \in \mathbb{N}$, $t \in (0,\infty)$, where we have used the formula $x^{\underline{n}} = \sum_{k=1}^n s(n,k) x^k$, $n \in \mathbb{N}$, $x \in \mathbb{R}$, see (2.8). Note that (3.27) is in agreement with [60, Proposition 29, Eq. (44)], when taking [60, Eq. (49)] into account.

In particular, for all $x \in \mathbb{R}$ and n sufficiently large,

$$\begin{aligned} \mathbb{P}\{T_n - \log\log n \le x\} &= \mathbb{P}\{T_n \le x + \log\log n\} = \frac{\Gamma(n - e^{-x-\log\log n})}{\Gamma(n)\Gamma(1-e^{-x-\log\log n})} \\ &= \frac{\Gamma(n - e^{-x}/\log n)}{\Gamma(n)\Gamma(1 - e^{-x}/\log n)} \sim \frac{\Gamma(n + c/\log n)}{\Gamma(n)} \end{aligned}$$

as $n \to \infty$, where the abbreviation $c := -e^{-x}$ is used for convenience. In order to see that the last expression converges to $e^c = e^{-e^{-x}} = F(x)$ as $n \to \infty$, one may apply Stirling's formula $\mathit{\Gamma}(n+1) \sim (n/e)^n\sqrt{2\pi n}$ as $n \to \infty$ to obtain

$$\begin{aligned}\frac{\mathit{\Gamma}(n+c/\log n)}{\mathit{\Gamma}(n)} &\sim \frac{\mathit{\Gamma}(n+c/\log n+1)}{\mathit{\Gamma}(n+1)}\\ &\sim \frac{((n+c/\log n)/e)^{n+c/\log n}\sqrt{2\pi(n+c/\log n)}}{(n/e)^n\sqrt{2\pi n}}\\ &\sim \frac{((n+c/\log n)/e)^{n+c/\log n}}{(n/e)^n}\\ &= e^{-c/\log n}(n+c/\log n)^{c/\log n}(1+c/(n\log n))^n\\ &\sim (n+c/\log n)^{c/\log n}\\ &= (1+c/(n\log n))^{c/\log n}n^{c/\log n} \sim n^{c/\log n} = e^c.\end{aligned}$$

Thus, for all $x \in \mathbb{R}$ it is shown that $\mathbb{P}\{T_n - \log\log n \le x\} \to F(x)$ as $n \to \infty$, so $T_n - \log\log n$ converges in distribution to the standard Gumbel distribution.

It remains to determine the Laplace transform of T_n. Applying the formula

$$\mathbb{E}f(T_n) \;=\; f(0) + \int_0^\infty f'(t)\mathbb{P}\{T_n > t\}\,\mathrm{d}t$$

to the function $f(t) := e^{-\lambda t}$, it follows that T_n has Laplace transform

$$\mathbb{E}e^{-\lambda T_n} \;=\; 1-\lambda\int_0^\infty e^{-\lambda t}(1-\mathbb{P}\{T_n \le t\})\,\mathrm{d}t, \qquad n \in \mathbb{N}, \lambda \in [0,\infty). \tag{3.28}$$

Plugging in the formula (3.27) for the distribution function of T_n it follows that

$$\mathbb{E}e^{-\lambda T_n} \;=\; 1-\lambda\int_0^\infty e^{-\lambda t}\left(1-\frac{(-1)^{n-1}}{(n-1)!}\sum_{k=1}^n s(n,k)e^{-(k-1)t}\right)\mathrm{d}t, \qquad n \in \mathbb{N}, \lambda \in [0,\infty).$$

Since $s(n,1)=(-1)^{n-1}(n-1)!$ we can rewrite this as

$$\begin{aligned}\mathbb{E}e^{-\lambda T_n} &= 1+\lambda\int_0^\infty e^{-\lambda t}\frac{(-1)^{n-1}}{(n-1)!}\sum_{k=2}^n s(n,k)e^{-(k-1)t}\,\mathrm{d}t\\ &= 1+\lambda\frac{(-1)^{n-1}}{(n-1)!}\sum_{k=2}^n s(n,k)\int_0^\infty e^{-(k-1+\lambda)t}\,\mathrm{d}t \;=\; 1+\frac{(-1)^{n-1}}{(n-1)!}\sum_{k=2}^n s(n,k)\frac{\lambda}{k-1+\lambda},\end{aligned}$$

$n\in\mathbb{N}$, $\lambda\in[0,\infty)$. The proof is complete. □

We have seen in (3.27) that the spectral decomposition provides an analytic proof of the formula (3.8) for the distribution function of the absorption time T_n. Alternatively, the connections between the Bolthausen-Sznitman coalescent and the random recursive tree, see 2.3.1, or the Chinese Restaurant process, see Appendix A, may also be used to prove (3.8). In most cases it is enriching to view a problem from different perspectives, and therefore we provide these alternative proofs.

Proof. (Second proof of (3.8)) Alternatively, the product formula (3.8) for the distribution function of T_n follows from the construction of the Bolthausen-Sznitman coalescent via the random recursive tree described in Section 2.3.1. Recall that Π^n is constructed by cutting down the random recursive tree $\mathfrak{T}_{\Delta_{[n]}}$. Hence, the event that Π^n reaches its absorbing state before time t, $\{T_n\le t\}$, occurs if and only if all children of the root in $\mathfrak{T}_{\Delta_{[n]}}$ are cut before time t, i.e. the attached exponential clock has to ring before time t. More formally, the event $\{T_n\le t\}$ coincides with

$$\bigcap_{i=2}^n(\{i\text{ is a child of the root, }E_i\le t\}\cup\{i\text{ is not a child of the root}\}),$$

where $E_2,\ldots,E_n$ are independent standard exponential random variables. These events, indexed by i, are furthermore independent. Formula (3.8) now follows by the straightforward calculation

$$\begin{aligned}
&\mathbb{P}\{T_n \le t\} \\
&= \mathbb{P}\left\{\bigcap_{i=2}^{n}(\{i \text{ is a child of the root, } E_i \le t\} \cup \{i \text{ is not a child of the root}\})\right\} \\
&= \prod_{i=2}^{n}\left(\frac{1}{i-1}(1-e^{-t}) + \left(1-\frac{1}{i-1}\right)\right) = \prod_{j=1}^{n-1}\frac{j-e^{-t}}{j} = \frac{\Gamma(n-e^{-t})}{\Gamma(n)\Gamma(1-e^{-t})},
\end{aligned}$$

$n \in \mathbb{N}$, $t \in (0,\infty)$. □

Proof. (Third proof of (3.8)) The following third approach derives formula (3.8) via the Chinese restaurant process. It is known, see [60, Theorem 14], that $\Pi(t)$ is a $\mathrm{PD}(e^{-t},0)$ random partition. For more information on both the Chinese restaurant process and the $\mathrm{PD}(\alpha,\theta)$ distribution see Appendix A. On the other hand the distribution $\mathrm{PD}(\alpha,\theta)$ is obtained by considering the ranked frequencies of the partition of an (α,θ)-Chinese restaurant process. Thus, the event $\{T_n \le t\}$ coincides with

$$\{\text{the first } n \text{ customers in a } (e^{-t},0)\text{-Chinese restaurant sit at table } 1\}.$$

Since any new customer joins the $m_k > 0$ customers at table k with probability $(m_k - e^{-t})/m$, where $m := \sum_k m_k$ denotes the number of already present customers, we obtain

$$\mathbb{P}\{T_n \le t\} \;=\; \prod_{i=2}^{n}\frac{i-1-e^{-t}}{i-1} \;=\; \frac{\Gamma(n-e^{-t})}{\Gamma(n)\Gamma(1-e^{-t})}$$

as required. □

3.3 Partition-valued process

This section is a modified version of the preprint [46] written in collaboration with Jonas Kukla from the University of Tübingen.

As opposed to Sections 3.1 and 3.2, we now turn to the partition-valued processes. Namely, we consider both the Kingman coalescent $\Pi^K = \{\Pi^K(t), t \geq 0\}$ and the Bolthausen-Sznitman coalescent $\Pi^{BS} = \{\Pi^{BS}(t), t \geq 0\}$. For convenience we drop the superscripts K and BS when there is no risk of ambiguity. Our main results are spectral decompositions of the generator of the Bolthausen-Sznitman n-coalescent, Theorem 3.16, respectively of the generator of Kingman's n-coalescent, Theorem 3.13. As Corollaries we obtain for the Bolthausen-Sznitman coalescent a derivation of the formula for the transition probabilites that goes back to [17], and a formula for its Green's matrix. As a further application we obtain a spectral decomposition of the generator of the block counting process of the Bolthausen-Sznitman, respectively Kingman's coalescent. These derivations are alternative to the ones given in Section 3.1 and Section 3.2.

We will denote the generator for both the coalescent process and the corresponding ancestral process by Q, as it will be clear from the context whether we refer to the partition-valued process, with generator $(q_{\pi\rho})_{\pi,\rho\in\mathcal{P}_n}$, or the integer-valued process, with generator $(q_{ij})_{i,j\in[n]}$.

3.3.1 Kingman's coalescent

Recall that Kingman's n-coalescent $\Pi^{n,K} = \{\Pi^{n,K}(t), t \geq 0\}$ is obtained by choosing Λ to be δ_0, the Dirac measure in 0, that is the corresponding Q-matrix

$Q \coloneqq Q^{n,K} = (q_{\pi\rho})_{\pi,\rho\in\mathcal{P}_n}$ is given by

$$q_{\pi\rho} = \begin{cases} 1 & \text{if } \pi \lessdot \rho, \\ -\binom{\#\pi}{2} & \text{if } \pi = \rho, \\ 0 & \text{otherwise,} \end{cases} \tag{3.29}$$

where $\pi \lessdot \rho$ if and only if $\pi \leq \rho$ and $\#\pi - \#\rho = 1$. In other words, the jump chain of Kingman's n-coalescent is the directed simple random walk on the partition lattice $\mathcal{P}_{[n]}$, where at each step the chain jumps into a coarser partition. In this subsection we only consider Kingman's n-coalescent and therefore write Π^n instead of $\Pi^{n,K}$.

Theorem 3.13 (Spectral decomposition of Kingman's coalescent)**.** *Define the matrices $L = (l_{\pi\rho})_{\pi,\rho\in\mathcal{P}_{[n]}}$ and $R = (r_{\pi\rho})_{\pi,\rho\in\mathcal{P}_{[n]}}$ by*

$$l_{\pi\rho} \coloneqq \begin{cases} (-1)^{\#\pi-\#\rho}\frac{(\#\pi+\#\rho-2)!}{(2\#\pi-2)!}\prod_{B\in\rho}\#\pi|_B! & \text{if } \pi \leq \rho, \\ 0 & \text{otherwise,} \end{cases} \tag{3.30}$$

and

$$r_{\pi\rho} \coloneqq \begin{cases} \frac{(2\#\rho-1)!}{(\#\pi+\#\rho-1)!}\prod_{B\in\rho}\#\pi|_B! & \text{if } \pi \leq \rho, \\ 0 & \text{otherwise.} \end{cases} \tag{3.31}$$

Then a spectral decomposition of Q is given by $Q = RDL$, where $D = (d_{\pi\rho})_{\pi,\rho\in\mathcal{P}_{[n]}}$ is defined by $d_{\pi\pi} \coloneqq -\binom{\#\pi}{2}$ and $d_{\pi\rho} = 0$ if $\pi \neq \rho$. In particular, $l_{\pi\pi} = r_{\pi\pi} = 1$ for any $\pi \in \mathcal{P}_{[n]}$.

Remark 3.14. Although the spectral decomposition of Q, Theorem 3.13, directly yields a spectral decomposition of the matrix of transition probabilites $p_{\pi\rho}(t) := \mathbb{P}\{\Pi(t) = \rho|\Pi(0) = \pi\}$, analoguous to the one given in (3.55) for the Bolthausen-Sznitman n-coalescent, this expression is not particularly convenient. In fact, a better approach to calculating $p_{\pi\rho}(t)$ is the one given by Kingman in [44] Equation (2.5).

Consider the block counting process $\{N^n(t), t \geq 0\}$ of Kingman's n-coalescent defined by $N^n(t) := \#\Pi^n(t)$ and let $Q = (q_{ij})_{i,j\in[n]}$ denote the generator of $N^n(t)$. From Theorem 3.13 we obtain a spectral decomposition of Q.

Corollary 3.15. *Let $L = (l_{ij})_{i,j\in[n]}$, $R = (r_{ij})_{i,j\in[n]}$ be matrices defined by*

$$l_{ij} := (-1)^{i+j}\frac{(i+j-2)!}{(2i-2)!}\left\lfloor \begin{matrix} i \\ j \end{matrix} \right\rfloor, \qquad r_{ij} := \frac{(2j-1)!}{(i+j-1)!}\left\lfloor \begin{matrix} i \\ j \end{matrix} \right\rfloor, \tag{3.32}$$

where $\left\lfloor \begin{smallmatrix} i \\ j \end{smallmatrix} \right\rfloor := \binom{i-1}{j-1}\frac{i!}{j!}$ denotes the unsigned Lah numbers which count the number of partitions into j blocks of a set of i elements, where the elements in each block are ordered. Then $Q = RDL$ is a spectral decomposition of Q, where $D = (d_{ij})_{i,j\in\mathcal{P}_{[n]}}$ is defined by $d_{ii} = -\binom{i}{2}$ and $d_{ij} = 0$ if $i \neq j$.

Notice that using the definition (2.7) for the Lah numbers, the formulae in Corollary 3.15 for the entries of R and L agree with the ones derived after Lemma 3.1.

For a derivation of the hitting probabilities of Kingman's n-coalescent see Remark 3.25.

3.3.2 Bolthausen-Sznitman coalescent

The Bolthausen-Sznitman n-coalescent $\Pi^{n,BS} = \{\Pi^{n,BS}(t), t \geq 0\}$ is obtained by choosing Λ to be the uniform measure on $[0, 1]$, that is the corresponding Q-matrix

$Q := Q^{n,BS} = (q_{\pi\rho})_{\pi,\rho\in\mathcal{P}_{[n]}}$ is given by

$$q_{\pi\rho} = \begin{cases} \frac{(\#\rho-1)!(\#\pi-\#\rho-1)!}{(\#\pi-1)!} & \text{if } \pi \prec \rho, \\ -(\#\pi-1) & \text{if } \pi = \rho, \\ 0 & \text{otherwise,} \end{cases} \tag{3.33}$$

where $\pi \prec \rho$ if and only if ρ is obtained by exactly one merger of blocks of π. In this section we only consider the Bolthausen-Sznitman n-coalescent and therefore write Π^n instead of $\Pi^{n,BS}$.

Theorem 3.16 (Spectral decomposition of the Bolthausen-Sznitman coalescent)**.** *Let $L = (l_{\pi\rho})_{\pi,\rho\in\mathcal{P}_{[n]}}$ and $R = (r_{\pi\rho})_{\pi,\rho\in\mathcal{P}_{[n]}}$ be matrices defined by*

$$l_{\pi\rho} := \begin{cases} (-1)^{\#\pi-\#\rho}\frac{(\#\rho-1)!}{(\#\pi-1)!} & \textit{if } \pi \leq \rho, \\ 0 & \textit{otherwise,} \end{cases} \tag{3.34}$$

and

$$r_{\pi\rho} := \begin{cases} \frac{(\#\rho-1)!}{(\#\pi-1)!}\prod_{B\in\rho}(\#\pi|_B-1)! & \textit{if } \pi \leq \rho, \\ 0 & \textit{otherwise.} \end{cases} \tag{3.35}$$

Recall that $\pi \leq \rho$ if and only if each block of π is contained in a block of ρ. Then a spectral decomposition of Q is given by $Q = RDL$, where $D = (d_{\pi\rho})_{\pi,\rho\in\mathcal{P}_{[n]}}$ is defined by $d_{\pi\pi} = -(\#\pi-1)$ and $d_{\pi\rho} = 0$ if $\pi \neq \rho$. In particular, $l_{\pi\pi} = r_{\pi\pi} = 1$ for any $\pi \in \mathcal{P}_{[n]}$.

The right eigenvectors $r_{\pi\rho}$ may be interpreted as the probability that a random recursive tree on the label set π can be cut down to a tree on the label set ρ, see "Bolthausen-Sznitman coalescent and random recursive tree" in Section 2.3. As an application of this spectral decomposition we derive the well-known formula for the transition probabilities of Π^n given by Bolthausen and Sznitman in [17], Proposition 1.4.

Corollary 3.17 (Transition probabilities of the Bolthausen-Sznitman coalescent)**.** *For any two partitions $\pi, \rho \in \mathcal{P}_{[n]}$ and any time $t \geq 0$ the transition probabilities $p_{\pi\rho}(t) := \mathbb{P}\{\Pi^n(t) = \rho | \Pi^n(0) = \pi\}$ of the Bolthausen-Sznitman n-coalescent are given by*

$$p_{\pi\rho}(t) = (-1)^{\#\rho} e^t \frac{(\#\rho - 1)!}{(\#\pi - 1)!} \prod_{B \in \rho} (-e^{-t})^{\overline{\#\pi|_B}},$$

where for $x \in \mathbb{R}$, $k \in \mathbb{N}$ we denote by $x^{\overline{k}} := x(x+1)\cdots(x+k-1)$ the ascending factorial power with the convention $x^{\overline{0}} := 1$.

Thanks to the spectral decomposition, Theorem 3.16, we obtain a formula for the Green's matrix $G = (g_{\pi\rho})_{\pi,\rho\in\mathcal{P}_{[n]}}$ of the Bolthausen-Sznitman n-coalescent defined by

$$g_{\pi\rho} := \int_0^\infty p_{\pi\rho}(t)dt. \tag{3.36}$$

Recall that $g_{\pi\rho} = \mathbb{E}[\int_0^\infty 1_{\{\Pi^n(t)=\rho\}} dt | \Pi^n(0) = \pi]$ is the expected total time that Π^n spends in ρ starting from π. Recall that we denote by $\left[{i \atop j}\right]$ the Stirling permutation numbers or unsigned Stirling numbers of the first kind, which count the number of permutations of a set of i elements with j cycles.

Corollary 3.18 (Green's matrix of the Bolthausen-Sznitman coalescent)**.** *The Green's matrix* $G = (g_{\pi\rho})_{\pi,\rho\in\mathcal{P}_{[n]}}$ *is given by*

$$g_{\pi\rho} = \begin{cases} (-1)^{\#\rho} \frac{(\#\rho-1)!}{(\#\pi-1)!} \sum_{(k_B)_{B\in\rho}\in\mathbb{N}^{\#\rho}} \frac{(-1)^{|k|}}{|k|-1} \prod_{B\in\rho} \begin{bmatrix} \#\pi|_B \\ k_B \end{bmatrix} & \text{if } \pi \leq \rho \neq \{[n]\}, \\ \infty & \text{if } \pi \leq \rho = \{[n]\}, \\ 0 & \text{otherwise}, \end{cases} \tag{3.37}$$

where $|k| := \sum_{B\in\rho} k_B$.

Remark 3.19. Notice that since the return probability to any state $\pi \in \mathcal{P}_{[n]}$, $\pi \neq \{[n]\}$, equals 0 for any Λ n-coalescent Π^n there is a close connection between the Green's matrix G of Π^n and its hitting probabilities defined by

$$h(\pi,\rho) := \mathbb{P}\left\{\Pi^n \text{ hits } \rho \text{ when started from } \pi\right\}, \tag{3.38}$$

namely via $g_{\pi\rho} = h(\pi,\rho)/q_\rho = h(\pi,\rho)/(1-\#\rho)$, cf. [59], p. 146, where $q_\rho := \sum_{\sigma\in\mathcal{P}_{[n]},\sigma\neq\rho} q_{\rho\sigma}$ is the total rate in ρ.

Recall that $\left\{ {i \atop j} \right\}$ denotes the Stirling partition numbers or Stirling numbers of the second kind, which count the number of partitions into j blocks of a set of i elements.

Remark 3.20. From Corollary 3.18 it follows by a technical but straightforward computation that for any $\pi \in \mathcal{P}_{[n],i}$ and $j \leq i$

$$\sum_{\rho\in\mathcal{P}_{[n],j}} g_{\pi\rho} = (-1)^j \frac{(j-1)!}{(i-1)!} \sum_{k=j}^{i} \frac{(-1)^k}{k-1} \begin{bmatrix} i \\ k \end{bmatrix} \begin{Bmatrix} k \\ j \end{Bmatrix},$$

in agreement with the entries of the Green's matrix of the block counting process of the Bolthausen-Sznitman coalescent provided in the proof of Corollary 1.5 in [56].

As a further application of the spectral decomposition, Theorem 3.16, we present a further derivation of the spectral decomposition of the generator $Q = (q_{ij})_{i,j\in[n]}$ of the block counting process $\{N^n(t), t \geq 0\}$ of the Bolthausen-Sznitman n-coalescent. In Section 3.2 this spectral decomposition of Q was derived by means of generating functions without recourse to the partition-valued process Π^n.

Corollary 3.21. *Let $L = (l_{ij})_{i,j\in[n]}$, $R = (r_{ij})_{i,j\in[n]}$, and $D = (d_{ij})_{i,j\in[n]}$ be matrices given by*

$$l_{ij} := (-1)^{i-j}\frac{(j-1)!}{(i-1)!}\left\{ {i \atop j} \right\}, \quad r_{ij} := \frac{(j-1)!}{(i-1)!}\left[{i \atop j} \right], \quad d_{ij} := (1-i)1_{\{i=j\}}, \tag{3.39}$$

Then a spectral decomposition of Q is given by $Q = RDL$.

3.3.3 Proofs

Recall the notation of Chapter 2 where we introduced the partition lattice $\mathcal{P}_{[n]}$ and the multiple merger coalescent Π. The Λ n-coalescent Π^n is a Markov chain on $\mathcal{P}_{[n]}$ with Q-matrix $Q^n = (q^n_{\pi\rho})_{\pi,\rho\in\mathcal{P}_{[n]}}$ given by

$$q^n_{\pi\rho} := \begin{cases} \lambda_{\#\pi,\#\pi-\#\rho+1} & \text{if } \pi \leq \rho, \\ -\lambda_{\#\pi} & \text{if } \pi = \rho, \\ 0 & \text{otherwise,} \end{cases} \tag{3.40}$$

where $\lambda_b \coloneqq \sum_{k=2}^{b} \binom{b}{k} \lambda_{b,k}$, $b \geq 2$, is the infinitesimal rate. Recall the definition of a partial order, $\leq$, on a set P from Section 2.1. If any two elements $p, q \in P$ of P are comparable, i.e. either $p \leq q$ or $q \leq p$ holds, then $\leq$ is said to be a *linear (or total) order* on P. More precisely, if $(P, \leq)$ is a partially ordered set and the reflexivity condition

i) $p \leq p$ for all $p \in P$

can be replaced by the stronger condition

i) $p, q \in P$ implies $p \leq q$ or $q \leq p$ (totality),

then $(P, \leq)$ is called a *linearly (or totally) ordered set.*

We call a binary relation $\leq_{\text{ex}}$ on P an *extension of* $\leq$ *to a linear order*, if $\leq_{\text{ex}}$ is a linear order on P and for all $p, q \in P$, $p \leq q$ implies $p \leq_{\text{ex}} q$. There are several extensions of $\leq$ to a linear order on $\mathcal{P}_{[n]}$. Let us fix such an extension $\leq_{\text{ex}}$ and notice that the following quantities of Q^n that we are interested in do not depend on the specific extension chosen. The linear order $\leq_{\text{ex}}$ induces a natural bijection ψ from $\mathcal{P}_{[n]}$ to $[B_n]$, where B_n denotes the nth Bell number, which is the number $\#\mathcal{P}_{[n]}$ of partitions of $[n]$, defined inductively by letting $\psi(\mathit{\Delta}_{[n]}) = 1$ and $\psi(\pi) \leq \psi(\rho)$ if and only if $\pi \leq_{\text{ex}} \rho$. Via the bijection ψ, Q^n can be viewed as a matrix with index set $[B_n]$ and is then an upper right triangular matrix. The determinant of Q^n is therefore given by the product of its diagonal entries. Hence, the characteristic polynomial of Q^n is given by

$$\chi_{Q^n}(x) \coloneqq \det(Q^n - x\mathbb{I}_n) = (-1)^{B_n} \prod_{\pi \in \mathcal{P}_n} (\lambda_{\#\pi} + x) = (-1)^{B_n} \prod_{i=1}^{n} (\lambda_i + x)^{\left\{ n \atop i \right\}},$$

where $\mathbb{I}_{\mathcal{P}_{[n]}} = (\delta_{\pi\rho})_{\pi,\rho\in\mathcal{P}_{[n]}}$ is the identity matrix on $\mathcal{P}_n$. Hence, for each $i \in [n]$, $-\lambda_i$ is an eigenvalue of Q^n with algebraic multiplicity $\#\mathcal{P}_{[n],i} = \left\{ {n \atop i} \right\}$. Since Q^n is a Q-matrix, $(1,\dots,1)^\top \in \mathbb{R}^{B_n}$ is an eigenvector corresponding to the eigenvalue 0, and direct inspection yields that $(1,0,\dots,0)^\top \in \mathbb{R}^{B_n}$ is also an eigenvector corresponding to the eigenvalue $-\lambda_n$.

From now on we fix an $n \in \mathbb{N}$, $n \geq 2$, and drop this index in the notation, if there is no risk of confusion.

Kingman's n-coalescent

For any two partitions $\pi, \rho \in \mathcal{P}_{[n]}$ such that $\pi \leq \rho$ one may ask in how many different ways the jump chain of Kingman's coalescent may reach ρ when started in π. Since at each step only one merger of a pair of blocks occurs, there are $\#\pi - \#\rho + 1$ steps to be taken, and so the set of different ways is

$$C(\pi,\rho) := \{(\pi_1,\dots,\pi_m) \colon \pi = \pi_1 \lessdot \cdots \lessdot \pi_m = \rho, m = \#\pi - \#\rho + 1\},$$

where we defined $\lessdot$ in the paragraph preceding Theorem 3.13. We call each element in $C(\pi,\rho)$ a *maximal chain in* $[\pi,\rho]$ and denote by

$$m(\pi,\rho) := \#C(\pi,\rho) \tag{3.41}$$

the *number of maximal chains in* $[\pi,\rho]$. Before we turn to the proof of the spectral decomposition of Q, Theorem 3.13, we count the number of maximal chains $m(\pi,\rho)$ in $[\pi,\rho]$.

Lemma 3.22 (Number of maximal chains)**.** *For $\pi, \rho \in \mathcal{P}_{[n]}$ we have that*

$$m(\pi,\rho) = \begin{cases} 2^{\#\rho-\#\pi}(\#\pi-\#\rho)!\prod_{B\in\rho}\#\pi|_B! & \text{if } \pi \le \rho, \\ 0 & \text{otherwise.} \end{cases} \tag{3.42}$$

Proof. For $\pi \not\le \rho$ the claim is trivial. Suppose then that $\pi \le \rho$. Notice that any maximal chain $(\pi_1, \dots, \pi_n)$ in $[\varDelta_{[n]}, \{[n]\}]$ can be constructed as follows. Let $\pi_1 \coloneqq \varDelta_{[n]}$, and if π_i with $i < n$ is constructed, π_{i+1} is obtained by merging two blocks in π_i, which can be done in $\binom{\#\pi_i}{2}$ ways. When $\pi_n = \{[n]\}$ is reached, the construction is finished. This construction shows that there are $m(\varDelta_{[n]}, \{[n]\}) = \binom{n}{2}\binom{n-1}{2}\cdots\binom{2}{2} = 2^{1-n}n!(n-1)!$ maximal chains in $[\varDelta_{[n]}, \{[n]\}]$. Hence (3.42) holds in the case $(\pi,\rho) = (\varDelta_{[n]}, \{[n]\})$.

For the general case, recall the isomorphism $[\pi,\rho] \cong \bigtimes_{B\in\rho}\mathcal{P}_{\pi|_B}$. As a consequence, any maximal chain in $[\pi,\rho]$ can be built by choosing a maximal chain in each factor $\mathcal{P}_{\pi|_B}$ and then "intertwining" these chains, i.e. ordering their elements (excluding the first element — the partition into singletons — in each chain) in any order subject to the restriction that the order of elements of the same chain is preserved. Consequently, we have

$$\begin{aligned} m(\pi,\rho) &= (\#\pi-\#\rho)!\prod_{B\in\rho}[(\#\pi|_B-1)!]^{-1}\prod_{B\in\rho}m(\varDelta_{\pi|_B}, \{\pi|_B\}) \\ &= (\#\pi-\#\rho)!\prod_{B\in\rho}2^{1-\#\pi|_B}\#\pi|_B! = 2^{\#\rho-\#\pi}(\#\pi-\#\rho)!\prod_{B\in\rho}\#\pi|_B!. \end{aligned}$$

$\square$

Lemma 3.23. *For any $\rho \in \mathcal{P}_{[n]}$ the vector $(r_{\pi\rho})_{\pi \in \mathcal{P}_{[n]}}$ defined by*

$$r_{\pi\rho} := \begin{cases} \frac{2^{\#\pi - \#\rho}(2\#\rho - 1)!}{(\#\pi - \#\rho)!(\#\pi + \#\rho - 1)!} m(\pi, \rho) & \text{if } \pi \leq \rho, \\ 0 & \text{otherwise,} \end{cases} \tag{3.43}$$

$$= \begin{cases} \frac{(2\#\rho - 1)!}{(\#\pi + \#\rho - 1)!} \prod_{B \in \rho} \#\pi|_B! & \text{if } \pi \leq \rho, \\ 0 & \text{otherwise,} \end{cases} \tag{3.44}$$

is a right eigenvector of Q with corresponding eigenvalue $-\binom{\#\rho}{2}$.

Proof. Fix $\pi, \rho \in \mathcal{P}_{[n]}$. If $\pi < \rho$, we have

$$\begin{aligned} &\sum_{\sigma \in \mathcal{P}_n} q_{\pi\sigma} r_{\sigma\rho} \\ &= \sum_{\sigma :\, \pi \lessdot \sigma} r_{\sigma\rho} - \binom{\#\pi}{2} r_{\pi\rho} \\ &= \frac{2^{\#\pi - 1 - \#\rho}(2\#\rho - 1)!}{(\#\pi - 1 - \#\rho)!(\#\pi + \#\rho - 2)!} \sum_{\sigma :\, \pi \lessdot \sigma} m(\sigma, \rho) - \binom{\#\pi}{2} r_{\pi\rho} \\ &= \left(\frac{(\#\pi - \#\rho)(\#\pi + \#\rho - 1)}{2} - \binom{\#\pi}{2} \right) r_{\pi\rho} = -\binom{\#\rho}{2} r_{\pi\rho}, \end{aligned}$$

where we used $\sum_{\sigma :\, \pi \lessdot \sigma} m(\sigma, \rho) = m(\pi, \rho)$. If $\pi = \rho$, we have

$$\sum_{\sigma \in \mathcal{P}_{[n]}} q_{\pi\sigma} r_{\sigma\rho} = q_{\pi\pi} = -\binom{\#\rho}{2} r_{\pi\rho},$$

since $m(\pi, \pi) = 1$, hence $r_{\pi\pi} = 1$. Finally, if $\pi \leq \rho$ does not hold, thus $r_{\pi\rho} = 0$, we cannot have $\pi \leq \sigma \leq \rho$ for any $\sigma \in \mathcal{P}_{[n]}$ and therefore $\sum_{\sigma \in \mathcal{P}_{[n]}} q_{\pi\sigma} r_{\sigma\rho} = 0$. This shows (3.43). Now (3.44) follows from Lemma 3.22 on the number of maximal chains. $\square$

Evidently, the B_n eigenvectors (recall that B_n denotes the nth Bell number) of Q defined by (3.43) are linearly independent. We are now interested in the inverse matrix of $R = (r_{\pi\rho})_{\pi\rho\in\mathcal{P}_{[n]}}$, that is the matrix $L = (l_{\pi\rho})_{\pi,\rho\in\mathcal{P}_{[n]}}$ such that $\delta_{\pi\rho} = \sum_{\sigma\in\mathcal{P}_{[n]}} r_{\pi\sigma} l_{\sigma\rho}$ for all $\pi, \rho \in \mathcal{P}_{[n]}$.

Lemma 3.24. *For any $\pi \in \mathcal{P}_{[n]}$ the vector $(l_{\pi\rho})_{\rho\in\mathcal{P}_{[n]}}$ given by*

$$l_{\pi\rho} := \begin{cases} (-1)^{\#\pi-\#\rho} \frac{2^{\#\pi-\#\rho}(\#\pi+\#\rho-2)!}{(2\#\pi-2)!(\#\pi-\#\rho)!} m(\pi,\rho) & \text{if } \pi \leq \rho, \\ 0 & \text{otherwise,} \end{cases} \tag{3.45}$$

$$= \begin{cases} (-1)^{\#\pi-\#\rho} \frac{(\#\pi+\#\rho-2)!}{(2\#\pi-2)!} \prod_{B\in\rho} \#\pi|_B! & \text{if } \pi \leq \rho, \\ 0 & \text{otherwise,} \end{cases} \tag{3.46}$$

is a left eigenvector of Q with corresponding eigenvalue $-\binom{\#\pi}{2}$.

Proof. Use $\sum_{\sigma:\,\sigma\lessdot\rho} m(\pi,\sigma) = m(\pi,\rho)$ to obtain

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{P}_{[n]}} l_{\pi\sigma} q_{\sigma\rho} \\
&= \sum_{\sigma:\,\sigma\lessdot\rho} l_{\pi\sigma} - \binom{\#\rho}{2} l_{\pi\rho} \\
&= (-1)^{\#\pi-\#\rho-1} \frac{2^{\#\pi-\#\rho-1}(\#\pi+\#\rho-1)!}{(2\#\pi-2)!(\#\pi-\#\rho-1)!} \sum_{\sigma:\,\sigma\lessdot\rho} m(\pi,\sigma) - \binom{\#\rho}{2} l_{\pi\rho} \\
&= \left(-\frac{(\#\pi+\#\rho-1)(\#\pi-\#\rho)}{2} - \binom{\#\rho}{2}\right) l_{\pi\rho} = -\binom{\#\pi}{2} l_{\pi\rho},
\end{aligned}$$

thus (3.45) holds. Equation (3.46) follows from the Lemma on the number of maximal chains, Lemma 3.22. □

Proof. (of Theorem 3.13) The inverse matrix of R, let us call it $U = (u_{\pi\rho})_{\pi,\rho\in\mathcal{P}_{[n]}}$, is uniquely determined, and is a matrix of left eigenvectors of Q, i.e. $UQ = DU$. Moreover, for any $\pi \in \mathcal{P}_{[n]}$ we have by assumption $u_{\pi\pi} = \sum_{\sigma\colon \pi \le \sigma \le \rho} u_{\pi\sigma} r_{\sigma\rho} = \delta_{\pi\pi} = 1$. This uniquely determines a matrix of left eigenvectors of Q, since for any $\pi, \rho \in \mathcal{P}_{[n]}$ with $\pi < \rho$ we have

$$-\binom{\#\rho}{2} u_{\pi\rho} = \sum_{\sigma \in [\pi,\rho]} u_{\pi\sigma} q_{\sigma\rho} = 1_{\{\pi \lessdot \rho\}} + \sum_{\sigma\colon \pi < \sigma \le \rho} u_{\pi\sigma} q_{\sigma\rho},$$

and $u_{\pi\rho} = 0$ for $\pi \nleq \rho$. Since $QR = RD$ by Lemma (3.24), and evidently $l_{\pi\pi} = 1$ for any $\pi \in \mathcal{P}_{[n]}$, we have $U = L$ and the claim follows. □

Remark 3.25. Instead of calculating the hitting probabilites $h(\pi, \rho)$ of Kingman's n-coalescent via the spectral decomposition, Theorem 3.13, we use the observation from Section 3.3.1 that the jump chain of $\mathit{\Pi}$ can be interpreted as the directed simple random walk on $\mathcal{P}_{[n]}$. This implies that the jump chain of $\mathit{\Pi}$ (when started from π) traces out any maximal chain in $[\pi, \{[n]\}]$ with equal probability $m(\pi, \{[n]\})^{-1}$. Clearly, the total number of maximal chains in $[\pi, \{[n]\}]$ that contain ρ is $m(\pi, \rho) m(\rho, \{[n]\})$, and thus

$$h(\pi, \rho) = \frac{m(\pi,\rho) m(\rho, \{[n]\})}{m(\pi, \{[n]\})} = \binom{\#\pi - 1}{\#\rho - 1}^{-1} \frac{\#\rho!}{\#\pi!} \prod_{B \in \rho} \#\pi|_B! = \left\lfloor \begin{matrix} \#\pi \\ \#\rho \end{matrix} \right\rfloor^{-1} \prod_{B \in \rho} \#\pi|_B!,$$

where we used the Lemma on the number of maximal chains, Lemma 3.22, in the second step. In the special case $\pi = \mathit{\Delta}_{[n]}$ this formula was given by Kingman in [44, equation (2.3)].

Proof. (of Corollary 3.15) In complete analogy to the argument in Corollary 3.21, we have $q_{ij} = \sum_{\rho \in \mathcal{P}_{[n],j}} q_{\pi\rho}$, independent of the particular partition $\pi \in \mathcal{P}_{[n],i}$, as the following calculation shows. Using the spectral decomposition of Q, Theorem

3.13, we obtain

$$
\begin{aligned}
q_{ij} &= \sum_{\rho \in \mathcal{P}_{[n],j}} \sum_{\sigma \in [\pi,\rho]} r_{\pi\sigma} d_{\sigma\sigma} l_{\sigma\rho} \\
&= -\sum_{\substack{\sigma\colon \pi \leq \sigma \\ \#\sigma \geq j}} \frac{(2\#\sigma - 1)!}{(i + \#\sigma - 1)!} \left(\prod_{B \in \sigma} \#\pi|_B! \right) \binom{\#\sigma}{2} \\
&\qquad\qquad \times \sum_{\substack{\rho\colon \sigma \leq \rho \\ \#\rho = j}} (-1)^{\#\sigma + j} \frac{(\sigma + j - 2)!}{(2\#\sigma - 2)!} \prod_{B \in \rho} \#\sigma|_B! \\
&= -\sum_{k=j}^{i} \frac{(2k-1)!}{(i+k-1)!} \left\lfloor {i \atop k} \right\rfloor \binom{k}{2} (-1)^{k+j} \frac{(k+j-2)!}{(2k-2)!} \left\lfloor {k \atop j} \right\rfloor = \sum_{k=j}^{i} r_{ik} d_{kk} l_{kj},
\end{aligned}
$$

where in the third step we used the identity $\left\lfloor {i \atop k} \right\rfloor = \sum_{\sigma \in \mathcal{P}_{[i],k}} \prod_{B \in \sigma} \#B!$ twice. This identity for the Lah numbers is obvious from the interpretation of $\left\lfloor {i \atop k} \right\rfloor$ as the number of partitions into k ordered blocks of a set of i elements, where the elements in each block are ordered. $\square$

Bolthausen-Sznitman n-coalescent

In order to prepare the proof of the spectral decomposition of $Q = Q^{n,BS}$, Theorem 3.16, we calculate the left and right eigenvectors of Q. In the sequel we give two proofs for the right eigenvectors of Q that are of rather different flavours. The first proof is completeley self-contained and only makes use of the partition lattice $\mathcal{P}_{[n]}$. Together with the proof of Lemma 3.23 it might serve as a starting point to find a spectral decomposition for more general coalescents, e.g. beta coalescents. A probabilistic interpretation of the right eigenvectors of Q in terms of random recursive trees motivates our second proof.

Lemma 3.26. *For $\rho \in \mathcal{P}_{[n]}$ the vector $(r_{\pi\rho})_{\pi \in \mathcal{P}_{[n]}}$ defined by*

$$r_{\pi\rho} := \begin{cases} \frac{(\#\rho-1)!}{(\#\pi-1)!} \prod_{B\in\rho} (\#\pi|_B - 1)! & \text{if } \pi \leq \rho, \\ 0 & \text{otherwise,} \end{cases} \tag{3.47}$$

is a right eigenvector of Q with corresponding eigenvalue $1 - \#\rho$.

Proof. (first proof of Lemma 3.26) In order to carry out the following calculations, for any two partitions $\pi, \rho \in \mathcal{P}_{[n]}$ we need to parameterize the set $\{\sigma \colon \pi \prec \sigma \leq \rho\}$. To construct an arbitrary partition σ such that $\pi \prec \sigma$, we could choose a subset $C \subseteq \pi$ of at least two blocks of π and merge them in order to obtain σ. In this case $\#\sigma = \#\pi - \#C + 1$. If, additionally, we require $\sigma \leq \rho$, we certainly cannot choose any collection C of blocks in π. Instead, all blocks chosen have to be in $\pi|_B$ for some block $B \in \rho$, in which case $\#\sigma|_B = \#\pi|_B - \#C + 1$. To summarize, we have

$$\{\sigma \colon \pi \prec \sigma \leq \rho\} = \left\{ \left\{ \bigcup_{D\in C} D \right\} \cup (\pi \setminus C) : B \in \rho, C \subseteq \pi|_B, \#C \geq 2 \right\}.$$

Using this parametrization, we obtain

$$
\begin{aligned}
&\sum_{\sigma\in\mathcal{P}_{[n]}} q_{\pi\sigma} r_{\sigma\rho} \\
&= \sum_{\sigma:\,\pi\prec\sigma\leq\rho} q_{\pi\sigma} r_{\sigma\rho} - (\#\pi-1) r_{\pi\rho} \\
&= \sum_{B\in\rho} \sum_{\substack{C\subseteq\pi|_B \\ \#C\geq 2}} \frac{(\#\pi-\#C)!(\#C-2)!}{(\#\pi-1)!} \frac{(\#\rho-1)!}{(\#\pi-\#C)!} \\
&\qquad \times (\#\pi|_B-\#C)! \prod_{D\in\rho\setminus\{B\}} (\#\pi|_D-1)! - (\#\pi-1) r_{\pi\rho} \\
&= \sum_{B\in\rho} \sum_{c=2}^{\#\pi|_B} \binom{\#\pi|_B}{c} \frac{(c-2)!}{(\#\pi-1)!} (\#\rho-1)! \frac{(\#\pi|_B-c)!}{(\#\pi|_B-1)!} \prod_{D\in\rho} (\#\pi|_D-1)! - (\#\pi-1) r_{\pi\rho} \\
&= \left(\sum_{B\in\rho} \sum_{c=2}^{\#\pi|_B} \binom{\#\pi|_B}{c} (c-2)! \frac{(\#\pi|_B-c)!}{(\#\pi|_B-1)!} - (\#\pi-1) \right) r_{\pi\rho} \\
&= \left(\sum_{B\in\rho} \#\pi|_B \sum_{c=2}^{\#\pi|_B} \frac{1}{c(c-1)} - (\#\pi-1) \right) r_{\pi\rho} \\
&= \left(\sum_{B\in\rho} (\#\pi|_B-1) - (\#\pi-1) \right) r_{\pi\rho} = (1-\#\rho) r_{\pi\rho},
\end{aligned}
$$

and the claim follows. □

Recall the construction of the Bolthausen-Sznitman coalescent via the random recursive tree in "Bolthausen-Sznitman coalescent and random recursive tree" in Section 2.3. By using this construction we will give a second derivation of Lemma 3.26. Fix two partitions $\pi, \rho \in \mathcal{P}_{[n]}$ such that $\pi \leq \rho$. We say that an increasing *tree* $\mathscr{T}$ *on* π *contains* ρ if and only if one can obtain an increasing tree on ρ by successively cutting $\mathscr{T}$. If $\mathscr{T}$ contains ρ we write $\rho \sqsubset \mathscr{T}$. Let $\{R_\pi(k), k \geq 0\}$ denote the jump chain of $\mathcal{R}_\pi$. From the definition of $\mathcal{R}$ it follows that a process that is equal in distribution to the jump chain R can be constructed recursively by letting

$$R_\pi(0) \coloneqq \mathfrak{T}_\pi, \quad R_\pi(k+1) \coloneqq cR_\pi(k), \quad k \geq 0.$$

A simple question then is: what is the probability $\mathbb{P}\{\rho \sqsubset \mathfrak{T}_\pi\}$ that a random recursive tree on π contains ρ? Notice first, that an increasing tree $\mathscr{T}$ on π contains ρ if for each block $B \in \rho$ we can find a node $v = v(B)$ in $\mathscr{T}$ such that the labels in the subtree above v coincide with the elements in $\pi|_B$. In other words, we can construct all increasing trees on π that contain ρ by first constructing an increasing tree $\mathscr{T}$ on ρ. Then each node $v \in \mathscr{T}$ is labelled by some block $B \in \rho$. Now if $\#\pi|_B > 1$, build an increasing tree on $\pi|_B$ and replace v by this tree. This procedure is carried out for all nodes $v \in \mathscr{T}$ to obtain an increasing tree on π that contains ρ. Therefore, the number of increasing trees on π containing ρ is

$$\#\{\mathscr{T} : \mathscr{T} \text{ increasing tree on } \pi, \rho \sqsubset \mathscr{T}\} = (\#\rho - 1)! \prod_{B \in \rho} (\#\pi|_B - 1)!.$$

On the other hand, the total number of increasing trees on π is $(\#\pi - 1)!$. Since a random recursive tree on π is a tree chosen uniformly at random from all increasing trees on π, we obtain

$$\mathbb{P}\{\rho \sqsubset \mathfrak{T}_\pi\} = \begin{cases} \frac{(\#\rho-1)! \prod_{B\in\rho}(\#\pi|_B-1)!}{(\#\pi-1)!} & \text{if } \pi \leq \rho, \\ 0 & \text{otherwise,} \end{cases} \tag{3.48}$$

which is just $r_{\pi\rho}$. We can now turn to our second proof of Lemma 3.26 in terms of random recursive trees.

Proof. (second proof of Lemma 3.26) We rewrite the statement as

$$\sum_{\sigma:\,\pi\prec\sigma} q_{\pi\sigma} r_{\sigma\rho} = (\#\pi - \#\rho) r_{\pi\rho}. \tag{3.49}$$

Let $J = \{J_k, k = 1, \ldots, n\}$ denote the jump chain of $\Pi^n = \Pi^{BS,n}$. Since

$$\mathbb{P}\{J_{n-1} = \sigma | J_n = \pi\} = q_{\pi\sigma}/q_\pi$$

is the probability that J jumps from π to σ, after dividing (3.49) by $q_\pi = \#\pi - 1$ we obtain for the left hand side

$$\begin{aligned}
\sum_{\sigma:\,\pi\prec\sigma} \frac{q_{\pi\sigma}}{q_\pi} r_{\sigma\rho} &= \sum_{\sigma:\,\pi\prec\sigma} \mathbb{P}\{J_{n-1} = \sigma | J_n = \pi\} \mathbb{P}\{\rho \sqsubset \mathfrak{T}_\sigma\} \\
&= \sum_{\sigma:\,\pi\prec\sigma} \mathbb{P}\{p(c\mathfrak{T}_\pi) = \sigma\} \mathbb{P}\{\rho \sqsubset \mathfrak{T}_\sigma\} \\
&= \mathbb{P}\left\{\rho \sqsubset \mathfrak{T}_{p(c\mathfrak{T}_\pi)}\right\} \\
&= \mathbb{P}\{\rho \sqsubset c\mathfrak{T}_\pi\} \\
&= \mathbb{P}\{\rho \sqsubset c\mathfrak{T}_\pi | \rho \sqsubset \mathfrak{T}_\pi\} \mathbb{P}\{\rho \sqsubset \mathfrak{T}_\pi\},
\end{aligned}$$

where we used that for any random recursive tree $\mathscr{T}$ one has $\mathfrak{T}_{p(\mathscr{T})} =_d \mathscr{T}$. Conditional on $\rho \sqsubset \mathfrak{T}_\pi$, there are precisely $\#\pi - \#\rho$ among the $\#\pi - 1$ edges in $\mathfrak{T}_\pi$ that we may cut in order to obtain a tree $c\mathfrak{T}_\pi$ that contains ρ. Since, by definition, $c\mathfrak{T}_\pi$ is obtained by cutting $\mathfrak{T}_\pi$ at an edge chosen uniformly at random, we have

$$\mathbb{P}\{\rho \sqsubset c\mathfrak{T}_\pi | \rho \sqsubset \mathfrak{T}_\pi\} = \frac{\#\pi - \#\rho}{\#\pi - 1}.$$

The claim follows. $\square$

Lemma 3.27. *For $x \in \mathbb{R} \setminus \{0\}$ we have*

$$\sum_{\sigma \in \mathcal{P}_{[n]}} r_{\pi\sigma} x^{\#\sigma - 1} l_{\sigma\rho} = (-1)^{\#\rho} x^{-1} \frac{(\#\rho - 1)!}{(\#\pi - 1)!} \prod_{B \in \rho} (-x)^{\overline{\#\pi|_B}}. \tag{3.50}$$

Proof. Notice that

$$\begin{aligned}\sum_{\sigma \in \mathcal{P}_{[n]}} r_{\pi\sigma} x^{\#\sigma - 1} l_{\sigma\rho} &= \frac{(\#\rho - 1)!}{(\#\pi - 1)!} \sum_{\sigma \in [\pi,\rho]} (-1)^{\#\sigma - \#\rho} x^{\#\sigma - 1} \prod_{B \in \sigma} (\#\pi|_B - 1)! \\ &= (-1)^{\#\rho} x^{-1} \frac{(\#\rho - 1)!}{(\#\pi - 1)!} \sum_{\sigma \in [\pi,\rho]} \prod_{B \in \sigma} -x(\#\pi|_B - 1)!. \end{aligned} \tag{3.51}$$

From the interpretation of $\left[{n \atop i}\right]$ as counting the number of permutations of $[n]$ with i cycles it is clear that $\left[{n \atop i}\right] = \sum_{\pi \in \mathcal{P}_{[n],i}} \prod_{B \in \pi} (\#B - 1)!$, cf. equation (1.15) in [61]. Using the isomorphism $[\pi, \rho] \cong \bigtimes_{B \in \rho} \mathcal{P}_{\pi|_B}$, we calculate

$$\begin{aligned}\sum_{\sigma \in [\pi,\rho]} \prod_{B \in \sigma} -x(\#\pi|_B - 1)! &= \sum_{\tau' \in \bigtimes_{B \in \rho} \mathcal{P}_{\pi|_B}} \prod_{B \in \rho} \prod_{C \in \tau'_B} -x(\#C - 1)! \\ &= \prod_{B \in \rho} \sum_{\tau \in \mathcal{P}_{\pi|_B}} \prod_{C \in \tau} -x(\#C - 1)! \\ &= \prod_{B \in \rho} \left(\sum_{k=1}^{\#\pi|_B} (-x)^k \sum_{\tau \in \mathcal{P}_{\pi|_B,k}} \prod_{C \in \tau} (\#C - 1)! \right) \\ &= \prod_{B \in \rho} \sum_{k=1}^{\#\pi|_B} (-x)^k \left[{\#\pi|_B \atop k}\right] \qquad (3.52) \\ &= \prod_{B \in \rho} (-x)^{\overline{\#\pi|_B}}, \qquad (3.53) \end{aligned}$$

where in the last step we used $x^{\overline{n}} = \sum_{k=1}^{n} \left[{n \atop k}\right] x^k$, cf. (2.8). $\square$

Lemma 3.28. *The matrix* $L = (l_{\pi\rho})_{\pi,\rho\in\mathcal{P}_{[n]}}$ *defined by*

$$l_{\pi\rho} := \begin{cases} (-1)^{\#\pi-\#\rho}\frac{(\#\rho-1)!}{(\#\pi-1)!} & \text{if } \pi \le \rho, \\ 0 & \text{otherwise,} \end{cases} \tag{3.54}$$

is the inverse matrix of R, *i.e.* $\sum_{\sigma\in\mathcal{P}_{[n]}} r_{\pi\sigma} l_{\sigma\rho} = \delta_{\pi\rho}$.

Proof. Choosing $x = 1$ in Lemma 3.27 we have that $\sum_{\sigma\in\mathcal{P}_{[n]}} r_{\pi\sigma} l_{\sigma\rho} = \delta_{\pi\rho}$. □

Proof. (of Theorem 3.16) The claim follows by Lemmata 3.26 and 3.28. □

Proof. (of Corollary 3.17) Since Π is a continuous-time Markov chain with finite state space, we have for $P(t) = (p_{\pi\rho}(t))_{\pi,\rho\in\mathcal{P}_{[n]}}$ the identity $P(t) = \exp(tRDL) = R\exp(tD)L$. In the last step we made use of the spectral decomposition, Theorem 3.16. In particular, this yields

$$p_{\pi\rho}(t) = \sum_{\sigma\in\mathcal{P}_{[n]}} r_{\pi\sigma} e^{-t(\#\sigma-1)} l_{\sigma\rho} = \sum_{\sigma\in[\pi,\rho]} r_{\pi\sigma} e^{-t(\#\sigma-1)} l_{\sigma\rho}. \tag{3.55}$$

Letting $x = e^{-t}$ in Lemma 3.27 proves the claim. □

Proof. (of Corollary 3.18) From (3.52) we have

$$\begin{aligned} x^{-1} \sum_{\sigma\in[\pi,\rho]} \prod_{B\in\sigma} -x(\#\pi|_B - 1)! &= x^{-1} \prod_{B\in\rho} \sum_{k=1}^{\#\pi|_B} (-x)^k \begin{bmatrix} \#\pi|_B \\ k \end{bmatrix} \\ &= \sum_{(k_B)_{B\in\rho}\in\mathbb{N}^{\#\rho}} (-1)^{|k|} x^{|k|-1} \prod_{B\in\rho} \begin{bmatrix} \#\pi|_B \\ k_B \end{bmatrix}, \end{aligned}$$

where $|k| := \sum_{B\in\rho} k_B$. Letting $x = e^{-t}$ and integrating out with respect to t we see for $\rho \neq \{[n]\}$

$$\int_0^\infty e^t \sum_{\sigma\in[\pi,\rho]} \prod_{B\in\sigma} -e^{-t}(\#\pi|_B - 1)!dt = \sum_{(k_B)_{B\in\rho}\in\mathbb{N}^{\#\rho}} \frac{(-1)^{|k|}}{|k|-1} \prod_{B\in\rho} \begin{bmatrix} \#\pi|_B \\ k_B \end{bmatrix},$$

where for $\rho = \{[n]\}$

$$\int_0^\infty e^t \sum_{\sigma\in[\pi,\rho]} \prod_{B\in\sigma} -e^{-t}(\#\pi|_B - 1)!dt = \sum_{k=1}^{\#\pi} (-1)^k \int_0^\infty e^{-t(k-1)}dt \begin{bmatrix} \#\pi \\ k \end{bmatrix} = -\infty.$$

The claim follows from (3.51) and the definition of $g_{\pi\rho}$. □

Proof. (of Corollary 3.21) Evidently, the rate at which a jump from i to j blocks occurs equals the sum of all rates at which a jump from a partition $\pi \in \mathcal{P}_{[n],i}$ to a partition $\rho \in \mathcal{P}_{[n],j}$ occurs, i.e. $q'_{ij} = \sum_{\rho\in\mathcal{P}_{[n],j}} q_{\pi\rho}$, where $\pi \in \mathcal{P}_{[n],i}$ is fixed arbitrarily. The quantity $\sum_{\rho\in\mathcal{P}_{[n],j}} q_{\pi\rho}$ does not depend on the choice of $\pi \in \mathcal{P}_{[n],i}$, as the following calculation shows. By the spectral decomposition of Q, Theorem 3.16, we obtain

$$\begin{aligned}
q'_{ij} &= \sum_{\rho\in\mathcal{P}_{[n],j}} \sum_{\sigma\in[\pi,\rho]} r_{\pi\sigma} d_{\sigma\sigma} l_{\sigma\rho} \\
&= \sum_{\substack{\sigma:\,\pi\leq\sigma \\ \#\sigma\geq j}} \sum_{\substack{\rho:\,\sigma\leq\rho \\ \#\rho=j}} \frac{(\#\sigma-1)!}{(\#\pi-1)!} \prod_{B\in\sigma} (\#B-1)!(1-\#\sigma)(-1)^{\#\sigma-j} \frac{(j-1)!}{(\#\sigma-1)!} \\
&= \frac{(j-1)!}{(i-1)!}(-1)^j \sum_{\substack{\sigma:\,\pi\leq\sigma \\ \#\sigma\geq j}} (\#\sigma-1)! \prod_{B\in\sigma} (\#B-1)!(1-\#\sigma) \frac{(-1)^{\#\sigma}}{(\#\sigma-1)!} \begin{Bmatrix} \#\sigma \\ j \end{Bmatrix} \\
&= \frac{(j-1)!}{(i-1)!}(-1)^j \sum_{k=j}^{i} (k-1)! \begin{bmatrix} i \\ k \end{bmatrix} (1-k) \frac{(-1)^k}{(k-1)!} \begin{Bmatrix} k \\ j \end{Bmatrix} = \sum_{k=i}^{j} r'_{ik} d'_{kk} l'_{kj},
\end{aligned}$$

hence, $Q' = R'D'L'$. $\square$

For the Kingman coalescent and the Bolthausen-Sznitman coalescent we have derived spectral decompositions for both the generators of Π^n and N^n together with a number of applications.

One may try to extend the generating functions techniques that we used for the block counting process to the setting of beta coalescents. This leads to fractional differential equations whose solutions currently seem to be out of reach (personal communication with Martin Möhle). The author is somewhat more optimistic that the spectral decomposition of Π^n can be “guessed” for beta coalescents, and we leave this for future work.

Chapter 4

Functionals of Kingman's coalescent and cumulants

In recent years the study of functionals $f(\Pi^n)$, respectively $f(\mathcal{T}^n)$, of an n-coalescent process Π^n, respectively of the corresponding coalescent tree $\mathcal{T}^n$, has attracted increasing attention among researchers in probability. Here, f is a measurable function from the set $D_{\mathcal{P}_{[n]}}[0,\infty)$ of $\mathcal{P}_{[n]}$-valued cádlág functions (i.e. right-continuous functions with left limits) on $[0,\infty)$, respectively rooted $\mathbb{R}$-trees on n leaves, endowed with the Skorokhod topology, into $\mathbb{R}^d$ for some $d \in \mathbb{N}$. For instance, f could measure the height or the total length of $\mathcal{T}^n$ or count the number of branches in $\mathcal{T}^n$ that support one leaf. The tree $\mathcal{T}^n$ exhibits a rich combinatorial structure, and—unlike for a Galton-Watson tree—its subtrees are highly interdependent. It is because of this dependency structure that for most functionals $f(\mathcal{T}^n)$ little can be said about their distribution. Often, the strategy, therefore, is to study rescaling limits of $f(\mathcal{T}^n)$ as $n \to \infty$.

In this Chapter, we deliberately focus on the setting of a finite sample size n and characterize the law of some functionals $f(\mathcal{T}^n)$ explicitly. The limiting distribution of $f(\mathcal{T}^n)$ (suitable rescaled) as $n \to \infty$ will only occur as a byproduct, if at all. In order to make some progress, throughout this chapter we restrict ourselves to the Kingman coalescent. We argue that cumulants are a useful tool for the characterization of $f(\mathcal{T}^n)$.

Section 4.1 is a modified version of the article [53] written in collaboration with Martin Möhle from the University of Tübingen. This article was accepted for publication in 2014 by the journal "Markov Processes and Related Fields". Here, f is chosen to be a) the height, T_n, respectively b) the total length, L_n, of $\mathcal{T}^n$.

In Section 4.2 we turn our attention to the site frequency spectrum $(M_1, \ldots, M_{n-1})$ for Kingman's n-coalescent. Recall that M_j counts the number of mutations that affect j individuals in the sample. This quantity can be measured in genetical data and is therefore highly important from the point of view of applications and for statistical inference. In Kingman's case the expected value, variance and covariance have been studied by Fu [32]. A discussion of these results and further information on the SFS can be found in Durrett [26]. For general multiple merger coalescents Birkner, Blath and Eldon [16] have derived recursion formulae for the expected value, the variance and the covariance of the SFS.

Here, we are only concerned with characterizing the law of an arbitrary marginal M_j and the beautiful underlying combinatorial problem. In fact, in Theorem 4.14 we work out the cumulants of arbitrary order of M_j.

4.1 Absorption time and tree length

We provide formulae for the cumulants and the moments of the time T back to the most recent common ancestor of the Kingman coalescent. It is shown that both the jth cumulant and the jth moment of T are linear combinations of the values $\zeta(2m)$, $m \in \{0, \ldots, \lfloor j/2 \rfloor\}$, of the Riemann zeta function ζ with integer coefficients. The proof is based on a solution of a two-dimensional recursion with countably many initial values. A closely related strong convergence result for the tree length L_n of the Kingman coalescent restricted to a sample of size n is derived. The results give reason to revisit the moments and central moments of the classical Gumbel distribution.

4.1.1 Results

Kingman [45] studied the absorption time T of Π. In the biological context T is called the time back to the most recent common ancestor or the age of the most recent common ancestor. It is well known that $T = \sum_{k=2}^{\infty} \tau_k$ is an infinite convolution of independent exponentially distributed random variables τ_k with parameter $\lambda_k = k(k-1)/2$. Using the inversion formula for Fourier transforms it is readily checked that T has a bounded and infinitely differentiable density $g : \mathbb{R} \to [0, \infty)$ with respect to Lebesgue measure on $\mathbb{R}$. Kingman [45, p. 37, Eq. (5.9)] showed that $g(t) = \sum_{k=2}^{\infty} (-1)^k (2k-1) \lambda_k e^{-\lambda_k t}$, $t \in (0, \infty)$. It is furthermore known (Watterson [74, p. 213], Tavaré [71, p. 132]) that T has mean $\mathbb{E}T = 2$ and variance $\mathrm{Var}(T) = 4\pi^2/3 - 12 \approx 1.15947$ and, hence, second moment $\mathbb{E}T^2 = 4\pi^2/3 - 8 \approx 5.15947$. To the best of the authors' knowledge higher moments and cumulants of T have not been derived so far. Proposition 4.1 and Theorem 4.3 below provide full information on the cumulants and moments of T.

Before we present our results, let us recall the notion of a cumulant. Let X be a real random variable whose moment generating function $M_X(s) = \mathbb{E}e^{sX}$ exists for $|s| < \delta$ and some $\delta > 0$. Recall that the *cumulant generating function* of X is defined by $K_X(s) = \log M_X(s)$. The coefficient c_j in the power series expansion $K_X(s) = \sum_{j=1}^{\infty} c_j s^j/j!$ near 0 is called the *jth cumulant* of X, denoted $\mathscr{C}_j(X)$. Notice that $\mathscr{C}_1(X) = \mathbb{E}X$ is the expectation and $\mathscr{C}_2(X) = \mathrm{Var}(X)$ the variance of X. We will repeatedly use the following properties of cumulants. For real random variables X, Y and $j \in \mathbb{N}$, $a, b \in \mathbb{R}$ we have

i) $\mathscr{C}_j(X + b) = \mathscr{C}_j(X)$ if $j \geq 2$,

ii) $\mathscr{C}_j(aX) = a^j \mathscr{C}_j(X)$,

iii) the independence of X, Y implies $\mathscr{C}_j(X + Y) = \mathscr{C}_j(X) + \mathscr{C}_j(Y)$.

Proposition 4.1. *For the Kingman coalescent the absorption time T has cumulants*

$$\mathscr{C}_j(T) \;=\; (j-1)!2^j \sum_{k=2}^{\infty} \frac{1}{k^j(k-1)^j}, \qquad j \in \mathbb{N}, \tag{4.1}$$

and moments

$$\mathbb{E}T^j \;=\; j!2^j \sum_{k=2}^{\infty} \frac{(-1)^k(2k-1)}{k^j(k-1)^j} \;=\; j! \sum_{m=1}^{\infty} \frac{1}{m^j}\left(\frac{4m-1}{(2m-1)^j} - \frac{4m+1}{(2m+1)^j}\right), \qquad j \in \mathbb{N}. \tag{4.2}$$

In particular, $\mathscr{C}_j(T) \sim (j-1)!$ and $\mathbb{E}T^j \sim 3j!$ as $j \to \infty$. Alternatively,

$$\mathbb{E}T^j \;=\; j! \sum_{2 \leq k_1 \leq \cdots \leq k_j} \frac{1}{\lambda_{k_1} \cdots \lambda_{k_j}} \;=\; j!2^j \sum_{2 \leq k_1 \leq \cdots \leq k_j} \prod_{i=1}^{j} \frac{1}{k_i(k_i-1)}, \qquad j \in \mathbb{N}. \tag{4.3}$$

Remark 4.2. For $n \in \mathbb{N}$ let T_n denote the absorption time of the Kingman coalescent restricted to a sample of size n. The proof of Proposition 4.1 provided in Section 4.1.2 shows that $T_n \to T$ almost surely as $n \to \infty$ with convergence of all moments, which implies (see, for example, [42, Proposition 3.12]) the convergence $T_n \to T$ in L^p for any $p \in (0, \infty)$. Moreover, the sequence $(T_n^p)_{n \in \mathbb{N}}$ is uniformly integrable for any $p \in (0, \infty)$.

The following result shows that the cumulants $\mathscr{C}_j(T)$ and the moments $\mathbb{E}T^j$ of T are related to the Riemann zeta function ζ. More precisely, $\mathscr{C}_j(T)$ and $\mathbb{E}T^j$ are both linear combinations of the zeta values $\zeta(2m)$, $m \in \{0, \ldots, \lfloor j/2 \rfloor\}$, with integer coefficients.

Theorem 4.3. *For all* $j \in \mathbb{N}$,

$$\mathscr{C}_j(T) \;=\; (-1)^j 2^{j+1} \sum_{m=0}^{\lfloor j/2 \rfloor} \frac{(2j-2m-1)!}{(j-2m)!} \zeta(2m) \tag{4.4}$$

and

$$\mathbb{E}T^j \;=\; (-1)^j j 2^{j+1} \sum_{m=0}^{\lfloor j/2 \rfloor} (2m-1) \left(1 - \frac{1}{2^{2m-1}}\right) \frac{(2j-2m-2)!}{(j-2m)!} \zeta(2m), \tag{4.5}$$

where ζ denotes the zeta function.

Remarks 4.4. 1. The coefficient in (4.5) in front of $\zeta(2m)$ is integer, since $2^{j+1}(1 - 1/2^{2m-1}) = 2^{j+1} - 2^{j-2m+2} \in \mathbb{Z}$ and $(2j-2m-2)!/(j-2m)! \in \mathbb{Z}$ for all $j \in \mathbb{N}$ and all $m \in \{0, \ldots, \lfloor j/2 \rfloor\}$.

2. Let $B_0, B_1, \ldots$ denote the Bernoulli numbers defined recursively via $B_0 := 1$ and $B_n := -1/(n+1) \sum_{k=0}^{n-1} \binom{n+1}{k} B_k$ for $n \in \mathbb{N}$. For instance, $B_1 = -1/2$, and $B_2 = 1/6$. Since $\zeta(2m) = (-1)^{m-1}(2\pi)^{2m}/(2(2m)!)B_{2m}$ is a rational multiple of

π^{2m}, Theorem 4.3 implies that $\mathscr{C}_j(T)$ and $\mathbb{E}T^j$ are both polynomials in π^2 of degree $\lfloor j/2 \rfloor$ with rational coefficients.

The proofs provided in Section 4.1.2 rely on the fact that the jump chain of the block counting process of the Kingman coalescent is deterministic, which implies that T is an infinite convolution of exponentially distributed random variables. The proof of Theorem 4.3 is based on a solution of a two-dimensional recursion with an infinite number of initial values (see Lemma 4.9).

Our methods do not seem to be directly applicable to (absorption times of) other exchangeable coalescent processes, since the jump chain of the block counting process of a coalescent with multiple collisions is in general not deterministic.

Our methods are partly useful to analyze further functionals of the Kingman n-coalescent (restricted to a sample of size $n \in \mathbb{N}$). As an example we provide detailed information on the tree length L_n (the sum of the lengths of all branches of the n-coalescent tree) of the Kingman n-coalescent. Let G be a standard Gumbel distributed random variable with distribution function $x \mapsto \exp(-\exp(-x))$, $x \in \mathbb{R}$. It is well known (see, for example, [72, p. 22–23], [25, Lemma 7.1] or [27, Lemma 2.21]) that L_n has the same distribution as the maximum of $n-1$ independent and exponentially distributed random variables with parameter $1/2$ and that $L_n/2 - \log n \to G$ in distribution as $n \to \infty$. We verify the following stronger convergence result.

Theorem 4.5 (Strong asymptotics of the tree length)**.** *For the Kingman coalescent, as $n \to \infty$, $G_n := L_n/2 - \log n \to G$ almost surely and in L^p for any $p \in (0, \infty)$, where G is standard Gumbel distributed. Moreover, the sequence $(G_n^p)_{n\in\mathbb{N}}$ is uniformly integrable for any $p \in (0, \infty)$.*

Remarks 4.6. 1. Theorem 4.5 implies the convergence $\mathscr{C}_j(G_n) \to \mathscr{C}_j(G)$ of all cumulants and the convergence $\mathbb{E}G_n^j \to \mathbb{E}G^j$ of all moments, which is one of the starting points of the proof of Theorem 4.5. Formulae for the cumulants and the moments of L_n are provided in (4.27), (4.29) and (4.30).

2. For asymptotic results on the tree length L_n for $beta(a,b)$ coalescents with parameters $a, b \in (0,\infty)$ we refer the reader to [43, Theorem 1] for $0 < a < 1$ and $b := 2-a$ and to [25, Corollary 4.3 and Theorem 5.2] for $a = b = 1$ (Bolthausen–Sznitman coalescent). The asymptotics of L_n for Ξ-coalescents with dust is provided in [52, Theorem 3].

The Gumbel distribution revisited

Theorem 4.5 and its proof smooth the way to establish a new result, Theorem 4.7, on the Gumbel distribution, also called the extreme value distribution of type 1. Recall that a standard Gumbel distributed random variable G has (see, for example, [40, p. 12, Eqs. (22.29) and (22.30)]) cumulants $\mathscr{C}_1 = \mathscr{C}_1(G) = \gamma$ (Euler's constant) and $\mathscr{C}_j = \mathscr{C}_j(G) = (-1)^j \Psi^{(j-1)}(1) = (j-1)!\,\zeta(j)$, $j \geq 2$. Note that $\mathscr{C}_j \sim (j-1)!$ as $j \to \infty$. Moreover, G has moments

$$m_n := \mathbb{E}G^n = \int_0^\infty (-\log u)^n e^{-u}\,\mathrm{d}u = (-1)^n \Gamma^{(n)}(1), \quad n \in \mathbb{N}_0 := \{0,1,2,\ldots\}. \tag{4.6}$$

The integral in (4.6) is sometimes called the nth Euler–Mascheroni integral. In the analytic community its interpretation as the nth moment of the Gumbel distribution often remains unmentioned. It is well known that $m_n \sim n!$ as $n \to \infty$. The moments $m_1, m_2, \ldots$ can be recursively computed via the relation between cumulants and moments

$$m_n = \sum_{k=1}^{n} \binom{n-1}{k-1} \mathscr{C}_k m_{n-k} = \gamma m_{n-1} + (n-1)! \sum_{k=2}^{n} \frac{1}{(n-k)!} \zeta(k) m_{n-k}, \quad n \in \mathbb{N}, \tag{4.7}$$

in agreement with the recursion provided on top of p. 214 in the book of Boros and Moll [18]. A useful and well known formula for the moments is

$$m_n = \sum_{\pi \in \mathcal{P}_{[n]}} \prod_{B \in \pi} \mathscr{C}_{\#B}, \qquad n \in \mathbb{N}. \tag{4.8}$$

Since there exist $n!/(a_1! \cdots a_n! 1!^{a_1} \cdots n!^{a_n})$ partitions $\pi \in \mathcal{P}_{[n]}$ having a_j blocks of size j, $1 \leq j \leq n$, the above formula can be also written as

$$m_n = n! \sum_{a_1,\ldots,a_n} \prod_{i=1}^{n} \frac{1}{a_i!} \left(\frac{\mathscr{C}_i}{i!} \right)^{a_i} = n! \sum_{a_1,\ldots,a_n} \frac{\gamma^{a_1}}{a_1!} \prod_{i=2}^{n} \frac{1}{a_i!} \left(\frac{\zeta(i)}{i} \right)^{a_i}, \qquad n \in \mathbb{N}, \tag{4.9}$$

where the sum $\sum_{a_1,\ldots,a_n}$ extends over all $a_1, \ldots, a_n \in \mathbb{N}_0$ satisfying $\sum_{i=1}^n i a_i = n$. For instance, $m_1 = \gamma$, $m_2 = \gamma^2 + \zeta(2)$, and $m_3 = \gamma^3 + 3\gamma\zeta(2) + 2\zeta(3)$. In particular, m_n is a polynomial of degree n in the variable $x := (x_1, \ldots, x_n) := (\gamma, \zeta(2), \ldots, \zeta(n))$ with non-negative integer coefficients. In the following we focus on the central moments

$$m'_n := \mathbb{E}(G - \gamma)^n \quad (n \in \mathbb{N}_0), \tag{4.10}$$

of the Gumbel distribution. Clearly,

$$m_n = \mathbb{E}(G - \gamma + \gamma)^n = \sum_{j=0}^{n} \binom{n}{j} \gamma^{n-j} m'_j, \qquad n \in \mathbb{N}_0, \tag{4.11}$$

and

$$m'_n \;=\; \sum_{j=0}^{n}\binom{n}{j}(-\gamma)^j m_{n-j} \;=\; n!\sum_{j=0}^{n}\frac{(-\gamma)^j}{j!}\frac{m_{n-j}}{(n-j)!} \;\sim\; n!\sum_{j=0}^{\infty}\frac{(-\gamma)^j}{j!} \;=\; n!e^{-\gamma}$$

as $n\to\infty$. As for the moments we obtain for the central moments the recursion

$$m'_n \;=\; \sum_{k=1}^{n}\binom{n-1}{k-1}\mathscr{C}_k\,(G-\gamma)\,m'_{n-k} \;=\; (n-1)!\sum_{k=2}^{n}\frac{1}{(n-k)!}\zeta(k)m'_{n-k}, \qquad n\in\mathbb{N},$$

with solution

$$m'_n \;=\; \sum_{\pi\in\mathcal{P}_{[n]}}\prod_{B\in\pi}\mathscr{C}_{\#B}\,(G-\gamma) \;=\; n!\sum_{a_2,\ldots,a_n}\prod_{i=2}^{n}\frac{1}{a_i!}\left(\frac{\zeta(i)}{i}\right)^{a_i}, \qquad n\in\mathbb{N}, \tag{4.12}$$

where the last sum $\sum_{a_2,\ldots,a_n}$ extends over all $a_2,\ldots,a_n\in\mathbb{N}_0$ satisfying $\sum_{i=2}^n ia_i=n$. In particular, for all $n\ge 2$, m'_n is a polynomial of degree $\lfloor n/2\rfloor$ in the variable $(\zeta(2),\ldots,\zeta(n))$ with non-negative integer coefficients. Theorem 4.7 below provides an alternative formula for m'_n. In order to state the result we introduce the non-negative integer coefficients

$$d_n \;:=\; n!\sum_{j=0}^{n}\frac{(-1)^j}{j!}, \qquad n\in\mathbb{N}_0. \tag{4.13}$$

Note that d_n is the number of derangements (fixed point free permutations) of n elements. Clearly, $d_n=nd_{n-1}+(-1)^n$, $n\in\mathbb{N}$. For instance, $d_0=1$, $d_1=0$, $d_2=1$, $d_3=2$, $d_4=9$, $d_5=44$, and $d_6=265$. For a set B and indices n_b, $b\in B$, we use in the following the notation $n_B:=\sum_{b\in B}n_b$.

Theorem 4.7 (Alternative formula for the central moments)**.** *A standard Gumbel* *distributed random variable G has central moments $m'_0 = 1$, $m'_1 = 0$ and*

$$m'_n \;=\; n!\sum_{i=1}^{n} \frac{1}{i!} \sum_{\substack{n_1,\ldots,n_i\geq 2\\ n_1+\cdots+n_i=n}} \frac{d_{n_1}\cdots d_{n_i}}{n_1!\cdots n_i!}\, s_i(n_1,\ldots,n_i), \qquad n\geq 2, \tag{4.14}$$

with coefficients d_n defined in (4.13) and $s_i(n_1,\ldots,n_i)$ given via

$$s_i(n_1,\ldots,n_i) := \sum_{\substack{k_1,\ldots,k_i\in\mathbb{N}\\ \text{all distinct}}} \frac{1}{k_1^{n_1}\cdots k_i^{n_i}} \tag{4.15}$$

$$= \sum_{l=1}^{i} (-1)^{i-l} \sum_{\{B_1,\ldots,B_l\}\in\mathcal{P}_{[i]}} (\#B_1-1)!\cdots(\#B_l-1)!\,\zeta(n_{B_1})\cdots\zeta(n_{B_l}) \tag{4.16}$$

$$= \sum_{\pi\in\mathcal{P}_{[i]}} (-1)^{i-\#\pi} \prod_{B\in\pi} (\#B-1)!\,\zeta(n_B), \tag{4.17}$$

where ζ denotes the zeta function.

Remarks 4.8. 1. The proof of Theorem 4.7 is based on the fact that the random variable S_n defined in (4.28) converges to $G-\gamma$ as $n\to\infty$. For more information on this convergence we refer the reader to the proof of Theorem 4.5.

2. Clearly, $s_i(n_1,\ldots,n_i)$ is symmetric with respect to the entries $n_1,\ldots,n_i$. For instance, $s_1(n_1)=\zeta(n_1)$, $s_2(n_1,n_2)=\zeta(n_1)\zeta(n_2)-\zeta(n_1+n_2)$ and $s_3(n_1,n_2,n_3) = \zeta(n_1)\zeta(n_2)\zeta(n_3)-\zeta(n_1)\zeta(n_2+n_3)-\zeta(n_2)\zeta(n_1+n_3)-\zeta(n_3)\zeta(n_1+n_2)+2\zeta(n_1+n_2+n_3)$. A recursion for $s_i(n_1,\ldots,n_i)$ is provided in (4.31). Eqs. (4.16) and (4.17) are reminiscent of sieve formulae, but we have not been able to rigorously relate these equations with some known sieve formula.

3. The values of the central moments m'_n for $0\leq n\leq 10$ are listed in Appendix B. The nth raw moment m_n of the Gumbel distribution is either obtained via (4.8)

or (4.9), or from the central moments m'_j, $0 \le j \le n$, via (4.11). The book of Srivastava and Choi [69, pp. 370–371] contains the values of $(-1)^n m_n = \Gamma^{(n)}(1)$ for $1 \le n \le 10$.

4.1.2 Proofs

Proof. (of Proposition 4.1) For $n \in \mathbb{N}$ let T_n denote the absorption time of the Kingman coalescent restricted to a sample of size n. Clearly (see, for example, Kingman [45, Eq. (5.5)]), $T_n \stackrel{d}{=} \sum_{k=2}^n \tau_k$, where $\tau_2, \tau_3, \dots$ are independent random variables and τ_k is exponentially distributed with parameter $\lambda_k = k(k-1)/2$, $k \ge 2$. Thus (e.g., Ross [63, p. 309]), T_n has a hypoexponential distribution with density $g_n(t) := \sum_{k=2}^n a_{nk}\lambda_k e^{-\lambda_k t}$, $t > 0$, where

$$a_{nk} \;:=\; \prod_{\substack{j=2\\ j\neq k}}^{n} \frac{\lambda_j}{\lambda_j - \lambda_k} \;=\; (-1)^k(2k-1)\frac{n!(n-1)!}{(n-k)!(n+k-1)!}, \qquad 2 \le k \le n.$$

Alternatively, the density g_n of T_n is obtained as follows. Let $D = (D_t)_{t\ge 0}$ denote the block counting process of the Kingman coalescent restricted to a sample of size n. From the Lemma on the spectral decomposition of a pure death process, Lemma 3.1, it follows that T_n has distribution function

$$\begin{aligned}
\mathbb{P}\{T_n \le t\} &= \mathbb{P}\{D_t = 1\} \;=\; \sum_{k=1}^{n} e^{-\lambda_k t} r_{nk} l_{k1}\\
&= \sum_{k=1}^{n} e^{-\lambda_k t}\Big(\prod_{j=k+1}^{n} \frac{\lambda_j}{\lambda_j - \lambda_k}\Big)\Big(\prod_{j=1}^{k-1} \frac{\lambda_{j+1}}{\lambda_j - \lambda_k}\Big)\\
&= 1 - \sum_{k=2}^{n} e^{-\lambda_k t} \prod_{\substack{j=2\\ j\neq k}}^{n} \frac{\lambda_j}{\lambda_j - \lambda_k} \;=\; 1 - \sum_{k=2}^{n} a_{nk} e^{-\lambda_k t}, \qquad t \in [0,\infty),
\end{aligned}$$

where the second last equality holds since $\lambda_1 = 0$. Taking the derivative with respect to t it follows that T_n has density g_n. In particular, T_n has moments

$$\mathbb{E}T_n^j \;=\; \int_0^\infty t^j g_n(t)\,\mathrm{d}t \;=\; \sum_{k=2}^n a_{nk}\int_0^\infty t^j \lambda_k e^{-\lambda_k t}\,\mathrm{d}t \;=\; j!\sum_{k=2}^n \frac{a_{nk}}{\lambda_k^j}, \qquad j\in\mathbb{N}. \tag{4.18}$$

In the following it is shown, essentially by letting $n\to\infty$ in (4.18), that

$$\mathbb{E}T^j \;=\; j!\sum_{k=2}^\infty \frac{(-1)^k(2k-1)}{\lambda_k^j}, \qquad j\in\mathbb{N}. \tag{4.19}$$

Eq. (4.19) holds for $j = 1$, since $\mathbb{E}T = 2$. Assume now that $j\geq 2$. From $T_n \nearrow T \stackrel{d}{=} \sum_{k=2}^\infty \tau_k$ almost surely as $n\to\infty$ it follows by monotone convergence that the left hand side in (4.18) converges to the left hand side in (4.19) as $n\to\infty$. In order to see that the right hand side in (4.18) converges to the right hand side in (4.19) as $n\to\infty$ fix $\varepsilon>0$. Since $j\geq 2$ and $\lambda_k = k(k-1)/2$, the series $\sum_{k=2}^\infty (2k-1)/\lambda_k^j$ is absolutely convergent. Thus, there exists a constant $n_0 = n_0(\varepsilon)$ such that $\sum_{k=n_0+1}^\infty (2k-1)/\lambda_k^j < \varepsilon$. Noting that $a_{nk} = (-1)^k(2k-1)b_{nk}$ with

$$0 \;\leq\; b_{nk} \;:=\; \frac{n!(n-1)!}{(n-k)!(n+k-1)!} \;=\; \frac{n(n-1)\cdots(n-k+1)}{(n+k-1)(n+k-2)\cdots n} \;\leq\; 1$$

it follows for all $n>n_0$ that

$$\begin{aligned}
\Big|\sum_{k=2}^n \frac{a_{nk}}{\lambda_k^j} - \sum_{k=2}^\infty \frac{(-1)^k(2k-1)}{\lambda_k^j}\Big| \;&\leq\; \sum_{k=2}^n \frac{2k-1}{\lambda_k^j}\underbrace{|b_{nk}-1|}_{\leq 1} + \sum_{k=n+1}^\infty \frac{2k-1}{\lambda_k^j}\\
&\leq\; \sum_{k=2}^{n_0} \frac{2k-1}{\lambda_k^j}|b_{nk}-1| + \sum_{k=n_0+1}^\infty \frac{2k-1}{\lambda_k^j}\\
&\leq\; \sum_{k=2}^{n_0} \frac{2k-1}{\lambda_k^j}|b_{nk}-1| + \varepsilon \;\to\; \varepsilon
\end{aligned}$$

as $n \to \infty$, since $b_{nk} \to 1$ as $n \to \infty$ for each fixed $k \in \{2, 3, \ldots\}$. Since $\varepsilon > 0$ can be chosen arbitrarily small, it follows that the right hand side in (4.18) converges to the right hand side in (4.19). Thus, (4.19) is established. Distinguishing in (4.19) even $k = 2m$ and odd $k = 2m + 1$ and summing over all $m \in \mathbb{N}$ it follows that

$$\mathbb{E}T^j \;=\; j! \sum_{m=1}^{\infty} \frac{1}{m^j} \bigg(\frac{4m-1}{(2m-1)^j} - \frac{4m+1}{(2m+1)^j} \bigg), \quad j \in \mathbb{N},$$

and (4.2) is established. Let us now turn to the cumulants of T. Note that the exponential distribution with parameter $\lambda \in (0, \infty)$ has jth cumulant $(j-1)!/\lambda^j$. From $T_n \stackrel{d}{=} \sum_{k=2}^n \tau_k$ it follows that T_n has cumulants

$$\mathscr{C}_j(T_n) \;=\; \sum_{k=2}^n \mathscr{C}_j(\tau_k) \;=\; \sum_{k=2}^n \frac{(j-1)!}{\lambda_k^j} \;=\; (j-1)! 2^j \sum_{k=2}^n \frac{1}{k^j (k-1)^j}, \qquad j \in \mathbb{N}. \tag{4.20}$$

We have verified above that the moments of T_n converge to those of T, which implies that the cumulants of T_n converge to those of T. Letting $n \to \infty$ in (4.20) yields (4.1).

Note that $\mathscr{C}_j(T)/(j-1)! = \sum_{k=2}^\infty 1/\lambda_k^j \to 1$ and that $\mathbb{E}T^j/j! = \sum_{k=2}^n (-1)^k (2k-1)/\lambda_k^j \to 3$ as $j \to \infty$, since $\lambda_2 = 1$ and $\lambda_k \geq \lambda_3 = 3$ for all $k \geq 3$.

It remains to verify the alternative formula (4.3) for the moments of T. For all $j \in \mathbb{N}$ we have

$$\begin{aligned} \mathbb{E}T^j &= \mathbb{E}\bigg(\sum_{k=2}^\infty \tau_k \bigg)^j \;=\; \sum_{k_1, \ldots, k_j = 2}^\infty \mathbb{E}\tau_{k_1} \cdots \tau_{k_j} \\ &= \sum_{k_1, \ldots, k_j = 2}^\infty \prod_{m=2}^\infty \mathbb{E}\tau_m^{a_m} \;=\; \sum_{2 \leq k_1 \leq \cdots \leq k_j} \frac{j!}{\prod_{m=2}^\infty a_m!} \prod_{m=2}^\infty \mathbb{E}\tau_m^{a_m}, \end{aligned}$$

where, for $m \geq 2$, a_m denotes the number of indices $k_1, \ldots, k_j$ being equal to m.

Note that $\sum_{m=2}^{\infty} a_m = j$. Since $\mathbb{E}\tau_m^{a_m} = a_m!/\lambda_m^{a_m}$, the above expression simplifies to

$$\mathbb{E}T^j \;=\; j! \sum_{2\le k_1\le\cdots\le k_j} \frac{1}{\lambda_{k_1}\cdots\lambda_{k_j}} \;=\; j!2^j \sum_{2\le k_1\le\cdots\le k_j} \prod_{i=1}^{j} \frac{1}{k_i(k_i-1)},$$

which is (4.3). □

The proof of Theorem 4.3 is based on the following basic but fundamental lemma, which provides a solution for a certain two dimensional recursion with a countable number of initial values.

Lemma 4.9. *Let $a_1, b_1, a_2, b_2, \ldots \in \mathbb{R}$. For $i, j \in \mathbb{N}$ define s_{ij} recursively via $s_{ij} := s_{i-1,j} - s_{i,j-1}$, with initial values $s_{0k} := a_k$ and $s_{k0} := b_k$, $k \in \mathbb{N}$. Then*

$$s_{ij} \;=\; \sum_{k=1}^{j} (-1)^{j-k} \binom{i+j-k-1}{i-1} a_k + (-1)^j \sum_{k=1}^{i} \binom{i+j-k-1}{j-1} b_k, \qquad i,j \in \mathbb{N}. \tag{4.21}$$

Proof. (of Lemma 4.9) Induction on $i+j$. Clearly, (4.21) holds for $i+j=2$, since $s_{11} = s_{01} - s_{10} = a_1 - b_1$. For the induction step from $i+j-1$ to $i+j$ (>2) three cases are distinguished.

Case 1: If $j=1$, then $i>1$ and, hence, by the recursion and by induction, $s_{i1} = s_{i-1,1} - s_{i0} = (a_1 - \sum_{k=1}^{i-1} b_k) - b_i = a_1 - \sum_{k=1}^{i} b_k$, which is (4.21) for $j=1$.

Case 2: If $i=1$, then $j>1$ and, hence, by the recursion and by induction, $s_{1j} = s_{0j} - s_{1,j-1} = a_j - (\sum_{k=1}^{j-1} (-1)^{j-1-k} a_k + (-1)^{j-1} b_1) = \sum_{k=1}^{j} (-1)^{j-k} a_k + (-1)^j b_1$, which is (4.21) for $i=1$.

Case 3: If $i, j > 1$ then

$$\begin{aligned}
s_{ij} &= s_{i-1,j} - s_{i,j-1} \\
&= \sum_{k=1}^{j} (-1)^{j-k} \binom{i+j-k-2}{i-2} a_k + (-1)^j \sum_{k=1}^{i-1} \binom{i+j-k-2}{j-1} b_k \\
&\qquad - \sum_{k=1}^{j-1} (-1)^{j-1-k} \binom{i+j-k-2}{i-1} a_k - (-1)^{j-1} \sum_{k=1}^{i} \binom{i+j-k-2}{j-2} b_k \\
&= a_j + \sum_{k=1}^{j-1} (-1)^{j-k} \left(\binom{i+j-k-2}{i-2} + \binom{i+j-k-2}{i-1} \right) a_k \\
&\qquad + (-1)^j b_i + (-1)^j \sum_{k=1}^{i-1} \left(\binom{i+j-k-2}{j-2} + \binom{i+j-k-2}{j-1} \right) b_k \\
&= \sum_{k=1}^{j} (-1)^{j-k} \binom{i+j-k-1}{i-1} a_k + (-1)^j \sum_{k=1}^{i} \binom{i+j-k-1}{j-1} b_k,
\end{aligned}$$

which completes the induction. $\square$

Before we come to the proof of Theorem 4.3 we provide a typical application of Lemma 4.9 showing that this lemma can be used to determine the value of certain series.

Example 4.10. For $j \in \mathbb{N}$ we would like to determine the series $\sum_{k=2}^{\infty} 1/(k^j(k-1)^j)$. We proceed as follows. For $i, j \in \mathbb{N}_0 := \{0, 1, 2, \ldots\}$ with $i + j \geq 2$ define $s_{ij} := \sum_{k=2}^{\infty} 1/(k^i(k-1)^j)$. Note that $s_{11} = 1$, $s_{i0} = \zeta(i) - 1$, $i \geq 2$, and that $s_{0j} = \zeta(j)$, $j \geq 2$. For all $i, j \in \mathbb{N}$ with $i + j \geq 3$,

$$s_{ij} = \sum_{k=2}^{\infty} \frac{1}{k^{i-1}(k-1)^{j-1}} \frac{1}{k(k-1)} = \sum_{k=2}^{\infty} \frac{1}{k^{i-1}(k-1)^{j-1}} \left(\frac{1}{k-1} - \frac{1}{k} \right) = s_{i-1,j} - s_{i,j-1}.$$

If we additionally define $s_{01} := 1$ and $s_{10} := 0$, then this recursion holds also for

$i = j = 1$, so for all $i, j \in \mathbb{N}$. By Lemma 4.9, for all $j \in \mathbb{N}$,

$$\begin{aligned} s_{jj} &= \sum_{k=1}^{j} (-1)^{j-k} \binom{2j-k-1}{j-1} s_{0k} + (-1)^j \sum_{k=1}^{j} \binom{2j-k-1}{j-1} s_{k0} \\ &= (-1)^j \sum_{\substack{k=2 \\ k \text{ even}}}^{j} \binom{2j-k-1}{j-1} (s_{0k} + s_{k0}) + (-1)^j \sum_{\substack{k=1 \\ k \text{ odd}}}^{j} \binom{2j-k-1}{j-1} (s_{k0} - s_{0k}). \end{aligned}$$

Plugging in $s_{0k} + s_{k0} = 2\zeta(k) - 1$ for even k and $s_{k0} - s_{0k} = -1$ for odd k we obtain the solution

$$\begin{aligned} \sum_{k=2}^{\infty} \frac{1}{k^j (k-1)^j} &= s_{jj} = (-1)^{j+1} \sum_{k=1}^{j} \binom{2j-k-1}{j-1} + 2(-1)^j \sum_{\substack{k=2 \\ k \text{ even}}}^{j} \binom{2j-k-1}{j-1} \zeta(k) \\ &= (-1)^{j+1} \binom{2j-1}{j} + 2(-1)^j \sum_{m=1}^{\lfloor j/2 \rfloor} \binom{2j-2m-1}{j-1} \zeta(2m) \\ &= 2(-1)^j \sum_{m=0}^{\lfloor j/2 \rfloor} \binom{2j-2m-1}{j-1} \zeta(2m), \qquad j \in \mathbb{N}, \end{aligned} \tag{4.22}$$

since $\zeta(0) = -1/2$. For instance $s_{22} = 2\zeta(2) - 3 = \pi^2/3 - 3 \approx 0.28987$ and $s_{33} = 10 - 6\zeta(2) = 10 - \pi^2 \approx 0.13040$.

The following proof of Theorem 4.3 has much in common with the previous example, but a modified double sequence $(s_{ij})_{i,j\in\mathbb{N}}$ is used having in particular more involved initial values.

Proof. (of Theorem 4.3) The formula (4.4) for the cumulants of T follows directly from (4.1) by multiplying (4.22) with $(j-1)!2^j$. In order to verify the formula (4.5) for the moments of T define for $i, j \in \mathbb{N}_0$ with $i + j \geq 2$

$$s_{ij} := \sum_{k=2}^{\infty} \frac{(-1)^k (2k-1)}{k^i (k-1)^j}.$$

By Proposition 4.1, $\mathbb{E}T^j = j!2^j s_{jj}$. Thus it remains to verify that

$$s_{jj} \;=\; 2(-1)^j \sum_{m=0}^{\lfloor j/2 \rfloor} (2m-1)\left(1-\frac{1}{2^{2m-1}}\right)\frac{(2j-2m-2)!}{(j-1)!(j-2m)!}\zeta(2m), \qquad j \in \mathbb{N}. \tag{4.23}$$

It is straightforward to check that $s_{11} = 1$, $s_{02} = 2\log 2 + \zeta(2)/2$, $s_{20} = 1 - 2\log 2 + \zeta(2)/2$,

$$s_{i0} \;=\; 1 - 2\left(1-\frac{1}{2^{i-2}}\right)\zeta(i-1) + \left(1-\frac{1}{2^{i-1}}\right)\zeta(i), \qquad i \geq 3,$$

and

$$s_{0j} \;=\; 2\left(1-\frac{1}{2^{j-2}}\right)\zeta(j-1) + \left(1-\frac{1}{2^{j-1}}\right)\zeta(j), \qquad j \geq 3.$$

For all $i, j \in \mathbb{N}$ with $i + j \geq 3$ we have

$$s_{ij} \;=\; \sum_{k=2}^{\infty} \frac{(-1)^k(2k-1)}{k^{i-1}(k-1)^{j-1}}\frac{1}{k(k-1)} \;=\; \sum_{k=2}^{\infty} \frac{(-1)^k(2k-1)}{k^{i-1}(k-1)^{j-1}}\left(\frac{1}{k-1}-\frac{1}{k}\right) \;=\; s_{i-1,j}-s_{i,j-1}. \tag{4.24}$$

If we additionally define $s_{01} := 1$ and $s_{10} := 0$, then the recursion (4.24) holds also for $i = j = 1$, so for all $i, j \in \mathbb{N}$. By Lemma 4.9,

$$s_{jj} \;=\; \sum_{k=1}^{j} (-1)^{j-k}\binom{2j-k-1}{j-1} s_{0k} + (-1)^j \sum_{k=1}^{j} \binom{2j-k-1}{j-1} s_{k0}, \qquad j \in \mathbb{N}.$$

Ordering with respect to even and odd k yields

$$(-1)^j s_{jj} \;=\; \sum_{\substack{k=2 \\ k \text{ even}}}^{j} \binom{2j-k-1}{j-1}(s_{0k}+s_{k0}) + \sum_{\substack{k=1 \\ k \text{ odd}}}^{j} \binom{2j-k-1}{j-1}(s_{k0}-s_{0k}).$$

Plugging in $s_{0k} + s_{k0} \;=\; 1 + 2(1 - 1/2^{k-1})\zeta(k)$ for even k, $s_{10} - s_{01} \;=\; -1$, and

$s_{k0} - s_{0k} = 1 - 4(1 - 1/2^{k-2}))\zeta(k-1)$ for odd $k \geq 3$, it follows that

$$
\begin{aligned}
(-1)^j s_{jj} &= \sum_{\substack{k=2 \\ k \text{ even}}}^{j} \binom{2j-k-1}{j-1}\Big(1 + 2\Big(1 - \frac{1}{2^{k-1}}\Big)\zeta(k)\Big) \\
&\quad + \binom{2j-2}{j-1}(-1) + \sum_{\substack{k=3 \\ k \text{ odd}}}^{j} \binom{2j-k-1}{j-1}\Big(1 - 4\Big(1 - \frac{1}{2^{k-2}}\Big)\zeta(k-1)\Big) \\
&= \sum_{k=2}^{j} \binom{2j-k-1}{j-1} - \binom{2j-2}{j-1} + 2\sum_{\substack{k=2 \\ k \text{ even}}}^{j} \binom{2j-k-1}{j-1}\Big(1 - \frac{1}{2^{k-1}}\Big)\zeta(k) \\
&\quad - 4\sum_{\substack{k=3 \\ k \text{ odd}}}^{j} \binom{2j-k-1}{j-1}\Big(1 - \frac{1}{2^{k-2}}\Big)\zeta(k-1) \\
&= \binom{2j-2}{j} - \binom{2j-2}{j-1} + 2\sum_{\substack{k=2 \\ k \text{ even}}}^{j} \binom{2j-k-1}{j-1}\Big(1 - \frac{1}{2^{k-1}}\Big)\zeta(k) \\
&\quad - 4\sum_{\substack{k=2 \\ k \text{ even}}}^{j-1} \binom{2j-k-2}{j-1}\Big(1 - \frac{1}{2^{k-1}}\Big)\zeta(k).
\end{aligned}
$$

Thus,

$$
(-1)^j s_{jj} = c_j + \sum_{\substack{k=2 \\ k \text{ even}}}^{j} c_{jk}\zeta(k), \tag{4.25}
$$

where $c_j := \binom{2j-2}{j} - \binom{2j-2}{j-1} = -(2j-2)!/(j!(j-1)!)$ and

$$
\begin{aligned}
c_{jk} &:= 2\Big(1 - \frac{1}{2^{k-1}}\Big)\Big(\binom{2j-k-1}{j-1} - 2\binom{2j-k-2}{j-1}\Big) \\
&= 2\Big(1 - \frac{1}{2^{k-1}}\Big)\Big(\frac{(2j-k-1)!}{(j-1)!(j-k)!} - 2\frac{(2j-k-2)!}{(j-1)!(j-k-1)!}\Big) \\
&= 2\Big(1 - \frac{1}{2^{k-1}}\Big)\frac{(2j-k-2)!}{(j-1)!(j-k)!}\big((2j-k-1) - 2(j-k)\big) \\
&= 2(k-1)\Big(1 - \frac{1}{2^{k-1}}\Big)\frac{(2j-k-2)!}{(j-1)!(j-k)!}.
\end{aligned}
$$

Substituting $k = 2m$ in (4.25) and noting that $\zeta(0) = -1/2$ yields (4.23). The proof is complete. $\square$

In order to prepare the proof of Theorem 4.5 recall that a standard Gumbel distributed random variable G has (see, for example, [40, p. 12, Eqs. (22.29) and (22.30)]) cumulants $\mathscr{C}_1(G) = \gamma$ (Euler's constant) and $\mathscr{C}_j(G) = (-1)^j \Psi^{(j-1)}(1) = (j-1)!\,\zeta(j)$, $j \geq 2$.

Proof. (of Theorem 4.5) We start in a similar way to the proof of Proposition 4.1. Clearly,

$$L_n := \sum_{k=2}^{n} k\tau_k = \sum_{k=2}^{n} X_k, \tag{4.26}$$

where the random variables $X_k := k\tau_k$, $k \geq 2$, are independent and X_k is exponentially distributed with parameter $\mu_k := \lambda_k/k = (k-1)/2$, $k \geq 2$. Thus, L_n has cumulants

$$\mathscr{C}_j(L_n) = \sum_{k=2}^{n} \mathscr{C}_j(X_k) = \sum_{k=2}^{n} \frac{(j-1)!}{\mu_k^j} = (j-1)!2^j \sum_{k=2}^{n} \frac{1}{(k-1)^j}, \qquad j \in \mathbb{N}. \tag{4.27}$$

For $j = 1$ we have $\mathscr{C}_1(G_n) = \mathscr{C}_1(L_n/2 - \log n) = \mathscr{C}_1(L_n)/2 - \log n = \sum_{k=2}^{n} 1/(k-1) - \log n \to \gamma = \mathscr{C}_1(G)$ as $n \to \infty$. For $j \geq 2$ we have $\mathscr{C}_j(G_n) = \mathscr{C}_j(L_n)/2^j = (j-1)! \sum_{k=2}^{n} 1/(k-1)^j \to (j-1)!\zeta(j) = \mathscr{C}_j(G)$ as $n \to \infty$. Thus, we have convergence $\mathscr{C}_j(G_n) \to \mathscr{C}_j(G)$ as $n \to \infty$ of all cumulants, which implies the convergence $\mathbb{E}G_n^j \to \mathbb{E}G^j$ as $n \to \infty$ of all moments.

For the convergence $G_n := L_n/2 - \log n \to G$ in distribution as $n \to \infty$ we refer the reader to [25, Lemma 7.1] and the references in the remark thereafter. In order to verify that the convergence $G_n \to G$ holds even almost surely define

$$S_n \;:=\; \frac{L_n - \mathbb{E}L_n}{2} \;=\; \sum_{k=2}^{n} \frac{X_k - \mathbb{E}X_k}{2} \;=\; \sum_{k=2}^{n} Y_k, \qquad n \in \mathbb{N}, \tag{4.28}$$

where $Y_k := (X_k - \mathbb{E}X_k)/2$, $k \geq 2$. Note that $Y_2, Y_3, \ldots$ are independent. From $\mathbb{E}L_n = 2\sum_{k=2}^{n} 1/(k-1) = 2\log n + 2\gamma + O(1/n)$ we conclude that $S_n \to G - \gamma$ in distribution as $n \to \infty$. It is well known (see, for example [14, Theorem 22.7]) that a sum $S_n = \sum_{k=2}^{n} Y_k$ of independent random variables converges in distribution if and only if it converges almost surely. Thus we even have $S_n \to G - \gamma$ almost surely. Thus, $G_n \to G$ almost surely as $n \to \infty$. By [42, Proposition 3.12] it follows that $G_n \to G$ in L^p for any $p \in (0, \infty)$, and the sequence $(G_n^p)_{n\in\mathbb{N}}$ is uniformly integrable for any $p \in (0, \infty)$. Note that $\|S_n - (G - \gamma)\|_p \leq \|\gamma + S_n - G_n\|_p + \|G_n - G\|_p = \|\gamma + \log n - \sum_{k=2}^{n} 1/(k-1)\|_p + \|G_n - G\|_p \to 0$ as $n \to \infty$, so we also have $S_n \to G - \gamma$ in L^p for any $p \in (0, \infty)$ and, hence, convergence $\mathbb{E}S_n^j \to \mathbb{E}(G-\gamma)^j$ as $n \to \infty$ of all moments. Furthermore, $(S_n)_{n\in\mathbb{N}}$ is a martingale, but we did not use this property in the proof.

□

Remark 4.11. For the sake of completeness, we provide in this remark formulae for the moments of the total tree length L_n. It is known (see, for example, [25, Lemma 7.1] and the remark thereafter) that L_n has the same distribution as the maximum of $n-1$ independent and exponentially distributed random variables with parameter 1/2. In particular, L_n has distribution function $\mathbb{P}\{L_n \leq t\} = (1 - e^{-t/2})^{n-1}$, $t > 0$, and, hence, moments

$$\begin{aligned}
\mathbb{E}L_n^j &= \int_0^\infty jt^{j-1}\mathbb{P}\left\{L_n > t\right\} \mathrm{d}t \\
&= \int_0^\infty jt^{j-1}\big(1-(1-e^{-t/2})^{n-1}\big) \,\mathrm{d}t \\
&= \int_0^\infty -jt^{j-1}\sum_{k=1}^{n-1}\binom{n-1}{k}(-e^{-t/2})^k \,\mathrm{d}t \\
&= \sum_{k=1}^{n-1}(-1)^{k+1}\binom{n-1}{k}\int_0^\infty jt^{j-1}e^{-kt/2}\,\mathrm{d}t \\
&= \sum_{k=1}^{n-1}(-1)^{k+1}\binom{n-1}{k}\frac{j!}{(k/2)^j} = j!2^j\sum_{k=1}^{n-1}\frac{(-1)^{k+1}}{k^j}\binom{n-1}{k}, \qquad j\in\mathbb{N}.
\end{aligned} \tag{4.29}$$

Alternatively,

$$\begin{aligned}
\mathbb{E}L_n^j &= \mathbb{E}\Bigg(\sum_{k=2}^n X_k\Bigg)^j = \sum_{k_1,\dots,k_j=2}^n \mathbb{E}X_{k_1}\cdots X_{k_j} \\
&= \sum_{k_1,\dots,k_j=2}^n \mathbb{E}X_2^{a_2}\cdots\mathbb{E}X_n^{a_n} = \sum_{2\le k_1\le\cdots\le k_j\le n}\frac{j!}{a_2!\cdots a_n!}\,\mathbb{E}X_2^{a_2}\cdots\mathbb{E}X_n^{a_n},
\end{aligned}$$

where, for $m\in\{2,\dots,n\}$, a_m denotes the number of indices $k_1,\dots,k_j$ being equal to m. Since $\mathbb{E}X_m^{a_m} = a_m!/\mu_k^{a_m}$, the above expression simplifies to

$$\mathbb{E}L_n^j = j!\sum_{2\le k_1\le\cdots\le k_j\le n}\frac{1}{\mu_{k_1}\cdots\mu_{k_j}} = j!2^j\sum_{1\le k_1\le\cdots\le k_j\le n-1}\frac{1}{k_1\cdots k_j}, \quad j\in\mathbb{N}. \tag{4.30}$$

Comparing (4.29) with (4.30) leads to the combinatorial identity

$$\sum_{k=1}^{n-1}\frac{(-1)^{k+1}}{k^j}\binom{n-1}{k} = \sum_{1\le k_1\le\cdots\le k_j\le n-1}\frac{1}{k_1\cdots k_j}, \qquad j\in\mathbb{N}, n\ge 2.$$

Proof. (of Theorem 4.7) Note first that an exponentially distributed random variable X with parameter $\alpha \in (0,\infty)$ has moments $\mathbb{E}X^n = n!/\alpha^n$ and central moments $\mathbb{E}(X-\mathbb{E}X)^n = d_n/\alpha^n$ with d_n defined in (4.13), $n \in \mathbb{N}_0$. Consider the random variable S_N defined via (4.28). From the proof of Theorem 4.5 it is already known that $S_N \to G-\gamma$ almost surely as $N \to \infty$ with convergence of all moments. Moreover, for all $n \in \mathbb{N}_0$,

$$\begin{aligned}
\mathbb{E}S_N^n &= \mathbb{E}(Y_2+\cdots+Y_N)^n \\
&= \sum_{i=1}^n \sum_{2\le k_1<\cdots<k_i\le N} \sum_{\substack{n_1,\ldots,n_i\ge 1\\ n_1+\cdots+n_i=n}} \frac{n!}{n_1!\cdots n_i!}\mathbb{E}Y_{k_1}^{n_1}\cdots\mathbb{E}Y_{k_i}^{n_i} \\
&= \sum_{i=1}^n \frac{1}{i!} \sum_{\substack{k_1,\ldots,k_i=2\\ \text{all distinct}}}^{N} \sum_{\substack{n_1,\ldots,n_i\ge 1\\ n_1+\cdots+n_i=n}} \frac{n!}{n_1!\cdots n_i!}\mathbb{E}Y_{k_1}^{n_1}\cdots\mathbb{E}Y_{k_i}^{n_i} \\
&= n!\sum_{i=1}^n \frac{1}{i!} \sum_{\substack{n_1,\ldots,n_i\ge 1\\ n_1+\cdots+n_i=n}} \frac{1}{n_1!\cdots n_i!} \sum_{\substack{k_1,\ldots,k_i=2\\ \text{all distinct}}}^{N} \frac{d_{n_1}}{(k_1-1)^{n_1}}\cdots\frac{d_{n_i}}{(k_i-1)^{n_i}} \\
&= n!\sum_{i=1}^n \frac{1}{i!} \sum_{\substack{n_1,\ldots,n_i\ge 2\\ n_1+\cdots+n_i=n}} \frac{d_{n_1}\cdots d_{n_i}}{n_1!\cdots n_i!} \sum_{\substack{k_1,\ldots,k_i=1\\ \text{all distinct}}}^{N-1} \frac{1}{k_1^{n_1}\cdots k_i^{n_i}},
\end{aligned}$$

since $d_1=0$. Letting $N\to\infty$ it follows that G has central moments (4.14) with $s_i(n_1,\ldots,n_i)$ defined via (4.15). Obviously, the expressions in (4.16) and (4.17) coincide. Thus, it remains to verify that $s_i(n_1,\ldots,n_i)$ can be expressed in terms of the zeta function via (4.16). We show this by induction on $i\in\mathbb{N}$. Clearly, (4.16) holds for $i=1$, since $s_1(n_1)=\zeta(n_1)$ for all $n_1\ge 2$. Concerning the induction step from $1,\dots,i$ to $i+1$ (≥ 2) note first that, for all $i\in\mathbb{N}$ and all $n_1,\dots,n_i\geq 2$,

$$
\begin{aligned}
&s_{i+1}(n_1,\dots,n_{i+1})\\
&=\sum_{\substack{k_1,\dots,k_i\in\mathbb{N}\\ \text{all distinct}}}\frac{1}{k_1^{n_1}\cdots k_i^{n_i}}\sum_{k_{i+1}\in\mathbb{N}\setminus\{k_1,\dots,k_i\}}\frac{1}{k_{i+1}^{n_{i+1}}}\\
&=\sum_{\substack{k_1,\dots,k_i\in\mathbb{N}\\ \text{all distinct}}}\frac{1}{k_1^{n_1}\cdots k_i^{n_i}}\Bigg(\sum_{k_{i+1}\in\mathbb{N}}\frac{1}{k_{i+1}^{n_{i+1}}}-\sum_{r=1}^{i}\frac{1}{k_r^{n_{i+1}}}\Bigg)\\
&=s_i(n_1,\dots,n_i)\,s_1(n_{i+1})-\sum_{r=1}^{i}s_i(n_1,\dots,n_{r-1},n_r+n_{i+1},n_{r+1},\dots,n_i).
\end{aligned}
\tag{4.31}
$$

By induction we conclude that

$$
\begin{aligned}
&s_i(n_1,\dots,n_i)\,s_1(n_{i+1})\\
&=\sum_{l=1}^{i}(-1)^{i-l}\sum_{\{B_1,\dots,B_l\}\in\mathcal{P}_i}(\#B_1-1)!\cdots(\#B_l-1)!\,\zeta(n_{B_1})\cdots\zeta(n_{B_l})\zeta(n_{i+1})\\
&=\sum_{l=1}^{i}(-1)^{i-l}\sum_{\substack{\{B_1,\dots,B_{l+1}\}\in\mathcal{P}_{i+1}\\ B_{l+1}=\{i+1\}}}(\#B_1-1)!\cdots(\#B_{l+1}-1)!\,\zeta(n_{B_1})\cdots\zeta(n_{B_{l+1}})\\
&=\sum_{l=2}^{i+1}(-1)^{i+1-l}\sum_{\substack{\{B_1,\dots,B_l\}\in\mathcal{P}_{i+1}\\ B_l=\{i+1\}}}(\#B_1-1)!\cdots(\#B_l-1)!\,\zeta(n_{B_1})\cdots\zeta(n_{B_l}).
\end{aligned}
\tag{4.32}
$$

Also by induction it is seen that

$$
\begin{aligned}
&\sum_{r=1}^{i}s_i(n_1,\dots,n_{r-1},n_r+n_{i+1},n_{r+1},\dots,n_i)\\
&=\sum_{r=1}^{i}\sum_{l=1}^{i}(-1)^{i-l}\sum_{\substack{\{A_1,\dots,A_l\}\in\mathcal{P}_i\\ r\in A_l}}(\#A_1-1)!\cdots(\#A_l-1)!\,\zeta(n_{A_1})\cdots\zeta(n_{A_{l-1}})\zeta(n_{A_l\cup\{i+1\}}),
\end{aligned}
$$

where, in the last sum, we ordered without loss of generality the blocks $A_1,\dots,A_l$

of the partition such that the element r belongs to the last block A_l. Reordering the sums on the right hand side yields

$$\begin{aligned}
&\sum_{r=1}^{i} s_i(n_1,\ldots,n_{r-1},n_r+n_{i+1},n_{r+1},\ldots,n_i)\\
&\quad = \sum_{l=1}^{i}(-1)^{i-l}\sum_{\{A_1,\ldots,A_l\}\in\mathcal{P}_i}\sum_{\substack{r=1\\ r\in A_l}}^{i}(\#A_1-1)!\cdots(\#A_l-1)!\,\zeta(n_{A_1})\cdots\zeta(n_{A_{l-1}})\zeta(n_{A_l\cup\{i+1\}}).
\end{aligned}$$

Rewriting this expression in terms of the blocks $B_1 := A_1,\ldots,B_{l-1} := A_{l-1}$ and $B_l := A_l \cup \{i+1\}$ it follows that

$$\begin{aligned}
&\sum_{r=1}^{i} s_i(n_1,\ldots,n_{r-1},n_r+n_{i+1},n_{r+1},\ldots,n_i)\\
&\quad = \sum_{l=1}^{i}(-1)^{i-l}\sum_{\{B_1,\ldots,B_l\}\in\mathcal{P}_{i+1}}\sum_{\substack{r=1\\ r,i+1\in B_l}}^{i}(\#B_1-1)!\cdots(\#B_{l-1}-1)!\,(\#B_l-2)!\,\zeta(n_{B_1})\cdots\zeta(n_{B_l}).
\end{aligned}$$

The last sum (over r) consists of $\#B_l-1$ summands and these summands do not depend on r, which gives rise to a factor $\#B_l-1$ leading to

$$\begin{aligned}
&\sum_{r=1}^{i} s_i(n_1,\ldots,n_{r-1},n_r+n_{i+1},n_{r+1},\ldots,n_i)\\
&=\sum_{l=1}^{i}(-1)^{i-l}\sum_{\substack{\{B_1,\ldots,B_l\}\in\mathcal{P}_{i+1}\\ B_k\neq\{i+1\}\text{ for all }k\in\{1,\ldots,l\}}}(\#B_1-1)!\cdots(\#B_l-1)!\,\zeta(n_{B_1})\cdots\zeta(n_{B_l}).
\end{aligned}\tag{4.33}$$

Subtracting (4.33) from (4.32) and recalling (4.31) it follows that

$$
\begin{aligned}
&s_{i+1}(n_1, \ldots, n_{i+1}) \\
&= \sum_{l=1}^{i+1} (-1)^{i+1-l} \sum_{\{B_1,\ldots,B_l\} \in \mathcal{P}_{i+1}} (\#B_1 - 1)! \cdots (\#B_l - 1)!\, \zeta(n_{B_1}) \cdots \zeta(n_{B_l}),
\end{aligned}
$$

which completes the induction. Thus, (4.16) is established.

□

4.2 Site frequency spectrum

Consider a sample of size $n \geq 2$ drawn from a large population of haploid individuals undergoing mutation. In most settings in population genetics the genealogy of such a sample is modled by Kingman's n-coalescent with mutations. We are interested in the distribution of the site frequency spectrum $(M_1, \ldots, M_{n-1})$, where M_j is the number of mutations that are shared by precisely j individuals in the sample. We derive the higher order cumulants of an arbitrary marginal M_j in terms of the jump chain $J = \{J_k, k = 1, \ldots, n\}$ of Π^n, where J_k denotes the state of Π^n after its $(n-k)$th jump.

For studies of the asymptotic behaviour of the SFS under Kingman's coalescent see Dahmer and Kersting [23], and under general multiple merger coalescents see Berestycki, Berestycki and Limic [8].

Before we turn to our result, let us recall the results of Fu [32] on the expected value, variance and covariance of the SFS without proof. Notice that the mutation parameter θ in Durrett's notation corresponds to $2r$ in our notation.

Theorem 4.12 (Theorem 2.1 in [26]).

$$\mathbb{E}M_j = 2r/j \quad (j \in [n-1]). \tag{4.34}$$

In order to present Fu's results on the variance and covariance, we use the following notation. Let $h_n := \sum_{k=1}^{n-1} 1/k$, $n \in \mathbb{N}$, denote the $(n-1)$th harmonic number, and let

$$\beta_n(i) := \frac{2n}{(n-i+1)(n-i)}(h_{n+1} - h_i) - \frac{2}{n-i} \quad (n, i \in \mathbb{N}). \tag{4.35}$$

Theorem 4.13 (Theorem 2.2 in [26]). *We have for $i, j \in [n-1]$*

$$\operatorname{Var} M_j = 2r/j + 4r^2\sigma_{jj}, \quad \textit{and for } i \neq j \quad \operatorname{Cov}(M_i, M_j) = 4r^2\sigma_{ij}, \tag{4.36}$$

where

$$\sigma_{jj} := \begin{cases} \beta_n(j+1) & \textit{if } j < n/2, \\ 2\frac{h_n - h_j}{n-j} - \frac{1}{j^2} & \textit{if } j = n/2, \\ \beta_n(j) - \frac{1}{j^2} & \textit{if } j > n/2 \end{cases} \tag{4.37}$$

and for $i > j$

$$\sigma_{ij} := \begin{cases} \frac{\beta_n(i+1) - \beta_n(i)}{2} & \textit{if } i + j < n, \\ \frac{h_n - h_i}{n-i} + \frac{h_n - h_i}{n-j} - \frac{\beta_n(i) + \beta_n(j+1)}{2} - \frac{1}{ij} & \textit{if } i + j = n, \\ \frac{\beta_n(j) - \beta_n(j+1)}{2} - \frac{1}{ij} & \textit{if } i + j > n. \end{cases} \tag{4.38}$$

Recall that for non-negative integers $n, k \in \mathbb{N}$ with $k \leq n$ we call $\left\lfloor {n \atop k} \right\rfloor = B_{n,k}(\bullet !) = \binom{n}{k}\frac{(n-1)!}{(k-1)!}$ the Lah numbers. Then $\left\lfloor {n \atop k} \right\rfloor$ counts the number of partitions π of $[n]$ with k linearly ordered blocks, i.e. in each block of π the elements are ordered.

Theorem 4.14. *(Cumulants of marginals of site frequency spectrum) The ith cumulant $\mathscr{C}_i\left(M_j\right)$ of M_j is given by*

$$\sum_{s=1}^{i} \sum_{l_1,\ldots,l_s=2}^{n} \sum_{\sigma \in \mathcal{P}_{[i],s}} \mathscr{C}\left(\left(\mathfrak{c}_j J_{l_k}\right)^{\#\sigma_k} : k \in [\#\sigma]\right) \prod_{m=1}^{\#\sigma} \mathrm{Li}_{1-\#\sigma_m}\left(r\left[r+\binom{l_m}{2}\right]^{-1}\right),$$

where σ_k denotes the kth block of the partition σ, the ordering of the blocks is fixed arbitrarily, and

$$\mathrm{Li}_n\left(u\right) := \sum_{l\geq 1} u^l/l^n \tag{4.39}$$

denotes the polylogarithm. In particular, for the expectation of M_j we obtain $\mathscr{C}_1\left(M_j\right) = \mathbb{E}M_j = 2r/j$ in agreement with Theorem 4.12.

We have not yet been able to show that our result in Theorem 4.14 agrees with Fu's result on the variance M_j in Theorem 4.13. Before we turn to the proof of Theorem 4.14, we calculate the cumulants of a negative binomial distribution.

We define the negative binomial distribution, $\mathrm{Nb}(\alpha, p)$, with parameters $\alpha > 0$, $p \in [0,1]$ to be supported on the non-negative intergers $\mathbb{N}_0$ with probability mass function

$$k \mapsto \binom{\alpha+k-1}{k}(1-p)^\alpha p^k \qquad (k \in \mathbb{N}_0). \tag{4.40}$$

If $\alpha \in \mathbb{N}$, $\mathrm{Nb}(\alpha, p)$ can be interpreted as the distribution of the number of failures in a sequence of independent Bernoulli trials with success probability $1-p$ each, before the αth success occurs. Recall that the special case $\alpha = 1$ yields the geometric distribution with success probability $1-p$.

In the following the notation $a[i \text{ times}]$ is a shorthand for i consecutive symbols "a" seperated by commas.

Lemma 4.15. *(Cumulants of negative binomial distribution) Let $X \sim \mathrm{Nb}(\alpha, p)$. Then the ith cumulant of X is given by*

$$\mathscr{C}_i\left(X\right) = \alpha \mathrm{Li}_{1-i}\left(p\right).$$

Martin Möhle pointed out the following elementary proof of Lemma 4.15.

Proof. (first proof of Lemma 4.15) It follows immediately with the binomial series that the moment generating function of $X \sim \mathrm{Nb}(\alpha, p)$ is

$$\mathbb{E}e^{tX} = \left(\frac{1-p}{1-pe^t}\right)^\alpha \quad (t < -\log p).$$

Hence, the cumulant generating function of X is given by

$$\log \mathbb{E}e^{tX} = \alpha \log \frac{1-p}{1-pe^t} = \alpha(\log(1-p) - \log(1-pe^t)).$$

Now, for $t < -\log p$, using the series representation $\log \frac{1}{1-x} = \sum_{k \geq 1} x^k/k$ for

$-1 \leq x < 1$, see [34, Equation 1.513.4], of the logarithm

$$\begin{aligned}-\log(1-pe^t) &= \sum_{k\geq 1} \frac{(pe^t)^k}{k} = \sum_{k\geq 1} \frac{p^k}{k} \sum_{l\geq 0} \frac{(tk)^l}{l!} = \sum_{l\geq 0} \frac{t^l}{l!} \sum_{k\geq 1} p^k k^{l-1} \\ &= \sum_{k\geq 1} \frac{p^k}{k} + \sum_{l\geq 1} \frac{t^l}{l!} \mathrm{Li}_{1-l}\left(p\right) = -\log(1-p) + \sum_{l\geq 1} \frac{t^l}{l!} \mathrm{Li}_{1-l}\left(p\right),\end{aligned}$$

and therefore

$$\log \mathbb{E} e^{tX} = \alpha \sum_{l\geq 1} \frac{t^l}{l!} \mathrm{Li}_{1-l}\left(p\right),$$

and the claim follows. □

Although this is a short and elementary proof, we consider it instructive to give an alternative proof of Lemma 4.15 via conditional cumulants, which is also a warmup for the proof of Theorem 4.14. However, this proof requires the notion of joint and conditional cumulants, as well as the notion of a mixture distribution, that we now recall.

For the following statements on joint and conditional cumulants we closely follow Speed [67] and Brillinger [19].

For $s \in \mathbb{R}^n$ let $M_{(X_1,\ldots,X_n)}(s) = \mathbb{E}[\exp\{\sum_{l=1}^n s_l X_l\}]$ denote the moment generating function of the real random vector $(X_1, \ldots, X_n)$ and suppose it has radius of convergence $\delta > 0$. Then the cumulant generating function of $(X_1, \ldots, X_n)$ is given by $K_{(X_1,\ldots,X_n)}(s) = \log M_{(X_1,\ldots,X_n)}(s)$ and the joint cumulant $\mathscr{C}_{i_1,\ldots,i_n}\left(X_1, \ldots, X_n\right)$ of order $(i_1, \ldots, i_n) \in \mathbb{N}^d$ is defined as the coefficient $c_{i_1,\ldots,i_n}$ in the series expansion $K_{(X_1,\ldots,X_n)}(s) = \sum_{i_1,\ldots,i_n\geq 1} c_{i_1,\ldots,i_n} \prod_{j=1}^n (s_1^{i_j}/i_j!)$ in s near 0. Often in the literature $\mathscr{C}_{1,\ldots,1}\left(X_1, \ldots, X_n\right)$ is called the joint cumulant of $(X_1, \ldots, X_n)$, and we denote it by $\mathscr{C}(X_1,\dots,X_n)$. A remarkable property of the joint cumulant $\mathscr{C}(X_1,\dots,X_n)$ is captured in the following proposition.

Proposition 4.16. *(*[67, Proposition 4.1]*) If there is a subset $S \subseteq [n]$ such that the random variables $(X_i)_{i\in S}$ and $(X_i)_{i\in[n]\setminus S}$ are independent, then $\mathscr{C}(X_1,\dots,X_n) = 0$.*

Another important property of the joint cumulant is its multilinearity, as stated in the next proposition.

Proposition 4.17. *(Multilinearity of joint cumulant,* [67, Proposition 4.2]*) Let $(X_{ij})_{(i,j)\in[n]\times[m]}$, respectively $(a_{ij})_{(i,j)\in[n]\times[m]}$ be an array of real random variables, respectively real numbers. Then*

$$\mathscr{C}\left(\sum_{j_1} a_{1j_1}X_{1j_1},\dots,\sum_{j_n} a_{nj_n}X_{nj_n}\right) = \sum_{j_1}\cdots\sum_{j_n} a_{1j_1}\cdots a_{nj_n}\mathscr{C}(X_{1j_1},\dots,X_{nj_n}).$$

Finally, we recall a formula for the joint cumulants in terms of conditional cumulants that was first given by Brillinger [19]. Suppose that X, Y are (possibly multivariate) random variables defined on some common probability space. We write $\mathscr{C}(X|Y)$ for the joint cumulant of X given Y. In complete analogy we denote conditional cumulants of higher order.

Proposition 4.18. *(*[67, Proposition 4.4]*)*

$$\mathscr{C}(X_1,\dots,X_n) = \sum_{\pi\in\mathcal{P}_{[n]}} \mathscr{C}\left(\mathscr{C}(X_i \colon i\in B|Y) : B\in\pi\right).$$

In what follows we use the notion of a mixture distribution as defined in Chapter 8 of Johnson et al. [39]. There is a number of contexts in which mixture distributions arise, but here we are only interested in the context of (randomly) varying parameters: suppose the distribution $\mathbb{P}_X = \mathbb{P}_X(\theta_1, \ldots, \theta_m)$, $m \in \mathbb{N}$, of a random variable X valued in $S \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, depends on parameters $\theta_1, \ldots, \theta_m$ that vary randomly according to the joint distribution $\mathbb{P}_{(\Theta_1,\ldots,\Theta_m)}$, that is, conditional on $\{\Theta_1 = \theta_1, \ldots, \Theta_m = \theta_m\}$ the cumulative distribution function of X is given by

$$F_{X;\theta_1,\ldots,\theta_m}(x) \coloneqq \mathbb{P}\left\{X \leq x | \Theta_1 = \theta_1, \ldots, \Theta_m = \theta_m\right\}, \quad (x \in S).$$

Consequently, the (unconditional) cumulative distribution function of X is given by

$$F_X(x) \coloneqq \mathbb{P}\left\{X \leq x\right\} = \int F_{X;\theta_1,\ldots,\theta_m}(x) d\mathbb{P}_{(\Theta_1,\ldots,\Theta_m)}.$$

The law $\mathbb{P}_X$ is called a mixture distribution and $\mathbb{P}_{(\Theta_1,\ldots,\Theta_m)}$ the corresponding mixing distribution. In analogy to the notation of Johnson et al. [39] we write this relation in the symbolic form

$$\mathbb{P}_X =_d \mathbb{P}_X(\theta_1, \ldots, \theta_m) \bigwedge_{(\theta_1,\ldots,\theta_m)} \mathbb{P}_{(\Theta_1,\ldots,\Theta_m)}.$$

For $\alpha, \beta > 0$ we define the Gamma distribution, Gamma(α, β), with parameters (α, β) to be supported on $(0, \infty)$ with density

$$t \mapsto \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-t\beta} \qquad (t \in (0, \infty)). \tag{4.41}$$

The special case $\alpha = 1$ yields the exponential distribution with parameter β. Recall that the negative binomial distribution can be constructed as a mixture distribution

$$\mathrm{Nb}(\alpha, \beta/(1+\beta)) = \mathrm{Pn}(\theta) \bigwedge_{\theta} \mathrm{Gamma}(\alpha, \beta), \tag{4.42}$$

where $\mathrm{Pn}(\theta)$ denotes a Poisson distribution with parameter $\theta > 0$, see e.g. [39].

Before we turn to the second proof of Lemma 4.15, we need an alternative formula for the polylogarithm $\mathrm{Li}_{-n}(u)$ when $n \in \mathbb{N}_0$.

Lemma 4.19. *For the polylogarithm of negative order we have*

$$\mathrm{Li}_{-n}(u) = \sum_{k=0}^{n} k! \left\{ {n+1 \atop k+1} \right\} \left(\frac{u}{1-u} \right)^{k+1} \qquad (n \in \mathbb{N}). \tag{4.43}$$

Proof. We prove the statement by indunction on n.

For $n = 0$, according to the definition (4.39) of the polylogarithm we have

$$\mathrm{Li}_0(u) = \sum_{l \geq 1} u^l = \frac{1}{1-u} - 1 = \frac{u}{1-u}$$

in agreement with the right hand side in (4.43).

To conclude the induction step, suppose that the statement holds for some $n \in \mathbb{N}$. Taking the derivative of $\mathrm{Li}_{-n}(u)$ with respect to u yields the recursion

$$\frac{d}{du}\mathrm{Li}_{-n}(u) = \frac{d}{du} \sum_{l \geq 1} u^l l^n = \sum_{l \geq 1} u^{l-1} l^{n+1} = u^{-1} \sum_{l \geq 1} u^l l^{n+1} = u^{-1} \mathrm{Li}_{-(n+1)}(u).$$

This recursion together with the initial value $\mathrm{Li}_0(u)$ completely determines the polylogarithm of negative order, $\mathrm{Li}_{-n}(u)$, $n \in \mathbb{N}$. If we can show that the right hand side

$$M_{-n}(u) := \sum_{k=0}^{n} k! \left\{ {n+1 \atop k+1} \right\} \left(\frac{u}{1-u} \right)^{k+1} \qquad (n \in \mathbb{N})$$

in (4.43) satisfies the same recursion, the claim follows.

Notice that for $m \in \mathbb{N}$

$$\begin{aligned} \frac{d}{du} u^m (1-u)^{-m} &= m u^{m-1} (1-u)^{-m} + m u^m (1-u)^{-m-1} \\ &= u^{-1} \left(m (\frac{u}{1-u})^m + m (\frac{u}{1-u})^{m+1} \right). \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{d}{du} M_{-n}(u) &= \frac{d}{du} \sum_{k=0}^{n} k! \left\{ {n+1 \atop k+1} \right\} \left(\frac{u}{1-u} \right)^{k+1} \\ &= u^{-1} \sum_{k=0}^{n} k! \left\{ {n+1 \atop k+1} \right\} \left((k+1)(\frac{u}{1-u})^{k+1} + (k+1)(\frac{u}{1-u})^{k+2} \right) \\ &= u^{-1} \sum_{k=0}^{n} (k+1)! \left\{ {n+1 \atop k+1} \right\} (\frac{u}{1-u})^{k+1} (1 + \frac{u}{1-u}) \\ &= \frac{1}{u(1-u)} \sum_{k=0}^{n} (k+1)! \left\{ {n+1 \atop k+1} \right\} (\frac{u}{1-u})^{k+1} \\ &= u^{-1} M_{-(n+1)}(u), \end{aligned}$$

where the last equation is seen as follows. Recall for $k > 0$ the recurrence relation for the Stirling numbers of the second kind

$$\left\{ {n+1 \atop k} \right\} = \left\{ {n \atop k-1} \right\} + \left\{ {n \atop k} \right\} k.$$

This implies

$$\begin{aligned}
M_{-(n+1)}(u) &= \sum_{k=0}^{n+1} k! \left\{ {n+2 \atop k+1} \right\} (\frac{u}{1-u})^{k+1} \\
&= \sum_{k=0}^{n+1} k! \left\{ {n+1 \atop k} \right\} (\frac{u}{1-u})^{k+1} + \sum_{k=0}^{n+1} k!(k+1) \left\{ {n+1 \atop k+1} \right\} (\frac{u}{1-u})^{k+1} \\
&= \sum_{k=1}^{n+1} k! \left\{ {n+1 \atop k} \right\} (\frac{u}{1-u})^{k+1} + \sum_{k=0}^{n} (k+1)! \left\{ {n+1 \atop k+1} \right\} (\frac{u}{1-u})^{k+1},
\end{aligned}$$

since $\left\{ {n+1 \atop 0} \right\} = \left\{ {n+1 \atop n+2} \right\} = 0$.

Moreover, we have

$$\begin{aligned}
\sum_{k=1}^{n+1} k! \left\{ {n+1 \atop k} \right\} (\frac{u}{1-u})^{k+1} &= \sum_{k=0}^{n} (k+1)! \left\{ {n+1 \atop k+1} \right\} (\frac{u}{1-u})^{k+2} \\
&= \frac{u}{1-u} \sum_{k=0}^{n} (k+1)! \left\{ {n+1 \atop k+1} \right\} (\frac{u}{1-u})^{k+1},
\end{aligned}$$

and consequently

$$M_{-(n+1)}(u) = \frac{1}{1-u} \sum_{k=0}^{n} (k+1)! \left\{ {n+1 \atop k+1} \right\} (\frac{u}{1-u})^{k+1}.$$

□

Proof. (second proof of Lemma 4.15) Recall that the jth cumulant of a random variate with law $\mathrm{Pn}(\theta)$ is θ for all $j \in \mathbb{N}$, and the jth cumulant of a random variate that obeys a $\mathrm{Gamma}(\alpha, \beta)$ law is $\alpha(j-1)!/\beta^j$, $j \in \mathbb{N}$. Let $X \sim \mathrm{Nb}(\alpha, \beta/(1+\beta))$ have a negative binomial distribution. According to the construction (4.42) we compute

$$
\begin{aligned}
\mathscr{C}_i(X) = \mathscr{C}(X[i \text{ times}]) &= \sum_{\sigma \in \mathcal{P}_{[i]}} \mathscr{C}\left(\mathscr{C}_{\#B}(\mathrm{Pn}(\tau)|\tau) : B \in \sigma\right) \\
&= \sum_{\sigma \in \mathcal{P}_{[i]}} \mathscr{C}(\tau \colon B \in \sigma) = \sum_{\sigma \in \mathcal{P}_{[i]}} \mathscr{C}_{\#\sigma}(\tau) \\
&= \alpha \sum_{b=1}^{i} \frac{(b-1)!}{\beta^b} \left\{ {i \atop b} \right\} = \alpha \mathrm{Li}_{1-i}\left(1/(1+\beta)\right),
\end{aligned}
$$

where we used the formula

$$
\mathrm{Li}_{-n}(u) = \sum_{k=0}^{n} k! \left\{ {n+1 \atop k+1} \right\} \left(\frac{u}{1-u} \right)^{k+1}
$$

for the polylogarithm of negative order, see Lemma 4.19. Now let $\beta \coloneqq p/(1-p)$. Then $X \sim \mathrm{Nb}(\alpha, p)$ and, consequently,

$$
\mathscr{C}_i(X) = \alpha \mathrm{Li}_{1-i}(p),
$$

and the claim follows. □

Proof. (Proof of Theorem 4.14) *Part I.* Conditioning on the jump chain J, we have by Proposition 4.18

$$
\mathscr{C}_i(M_j) = \mathscr{C}(M_j[i \text{ times}]) = \sum_{\sigma \in \mathcal{P}_{[i]}} \mathscr{C}\left(\mathscr{C}\left(M_j[\#B \text{ times}] \mid J\right) : B \in \sigma\right),
$$

and

$$\begin{aligned}\mathscr{C}\left(M_j[\#B \text{ times}]|J\right) &= \mathscr{C}\left(\sum_{b=2}^{n} M_j(\#J_b)[\#B \text{ times}]\middle| J\right)\\ &= \sum_{b_1=2}^{n}\cdots\sum_{b_{\#B}=2}^{n}\mathscr{C}\left(M_j(\#J_{b_1}),\dots M_j(\#J_{b_{\#B}})|J\right),\end{aligned}$$

where $M_j(k)$ is defined as the number of mutations that affect j individuals while Π^n has k blocks. Letting $u_l \coloneqq \#\{i\colon b_i = l\}$ we obtain

$$\mathscr{C}\left(M_j(\#J_{b_1}),\dots M_j(\#J_{b_{\#B}})|J\right) = \mathscr{C}\left(M_j(\#J_2)[u_2 \text{ times}],\dots,M_j(\#J_n)[u_n \text{ times}]|J\right)$$

and thus

$$\mathscr{C}\left(M_j[\#B \text{ times}]|J\right) = \sum_{\substack{u_2,\dots,u_n=0\\|u|=\#B}}^{\#B}\mathscr{C}\left(M_j(\#J_2)[u_2 \text{ times}],\dots,M_j(\#J_n)[u_n \text{ times}]|J\right)$$

While in state J_i, Π^n has $\mathfrak{c}_j J_i$ blocks of size j. The time until the next merger is $r\tau_i$, where (τ_i) is a sequence of independent exponential random variables, τ_i having parameter $\binom{i}{2}$, and the number of mutations before the next jump that may affect any of these ancestors independently is Poisson distributed with random parameter $r\tau_i$. This shows that

$$M_j(\#J_i) =_d \bigstar_{\mathfrak{c}_j J_i}\operatorname{Pn}(\theta)\bigwedge_{\theta}\operatorname{Exp}\left(\binom{i}{2}/r\right), \tag{4.44}$$

where $\bigstar_l$ denotes the l-fold convolution of measures. Let $(P_i(r\tau_i))_{i=2}^{n}$ denote a sequence of random variables with law $\bigotimes_{\mathfrak{c}_j J_i} \mathrm{Pn}(\theta) \bigwedge_\theta \mathrm{Exp}\left(\binom{i}{2}/r\right)$, where $\bigotimes_l$ denotes the l-fold product of measures. By Proposition 4.16 and Lemma 4.15

$$\begin{aligned}
&\mathscr{C}\left(P_1(r\tau_2)[u_2 \text{ times}], \ldots, P_n(r\tau_n)[u_n \text{ times}] | J\right) \\
&= \begin{cases} \mathscr{C}\left(P_k(r\tau_k)[\#B \text{ times}] | J\right) & \text{if } u_k = \#B \text{ for some } 2 \le k \le n, \\ 0 & \text{otherwise,} \end{cases} \\
&= \begin{cases} \mathrm{Li}_{1-\#B}\left(r\left[r + \binom{k}{2}\right]^{-1}\right) & \text{if } u_k = \#B \text{ for some } 2 \le k \le n, \\ 0 & \text{otherwise,} \end{cases}
\end{aligned}$$

where we applied Lemma 4.15 in the second step. The multilinearity of the joint cumulant, Proposition 4.17, together with (4.44) yields

$$\begin{aligned}
&\mathscr{C}\left(M_j(\#J_2)[u_2 \text{ times}], \ldots, M_j(\#J_n)[u_n \text{ times}] | J\right) \\
&= \left(\prod_{l=2}^{n} (\mathfrak{c}_j J_l)^{u_l}\right) \mathscr{C}\left(P_1(r\tau_2)[u_2 \text{ times}], \ldots, P_n(r\tau_n)[u_n \text{ times}] | J\right)
\end{aligned}$$

Hence, a further application of the multilinearity of the joint cumulant, Proposition 4.17, yields

$$\begin{aligned}
\mathscr{C}_i\left(M_j\right) &= \sum_{\sigma \in \mathcal{P}_{[i]}} \mathscr{C}\left(\sum_{l=2}^{n} (\mathfrak{c}_j J_l)^{\#B} \mathrm{Li}_{1-\#B}\left(r\left[r + \binom{l}{2}\right]^{-1}\right) : B \in \sigma\right) \\
&= \sum_{\sigma \in \mathcal{P}_{[i]}} \sum_{(l_B)_{B\in\sigma} \in \{2,\ldots,n\}^{\#\sigma}} \mathscr{C}\left((\mathfrak{c}_j J_{l_B})^{\#B} \mathrm{Li}_{1-\#B}\left(r\left[r + \binom{l_B}{2}\right]^{-1}\right) : B \in \sigma\right).
\end{aligned}$$

Notice that

$$\begin{aligned}
&\mathscr{C}\left(\left(\mathfrak{c}_j J_{l_B}\right)^{\#B}\mathrm{Li}_{1-\#B}\left(r\left[r+\binom{l_B}{2}\right]^{-1}\right): B\in\sigma\right)\\
&=\left(\prod_{B\in\sigma}\mathrm{Li}_{1-\#B}\left(r\left[r+\binom{l_B}{2}\right]^{-1}\right)\right)\mathscr{C}\left(\left(\mathfrak{c}_j J_{l_B}\right)^{\#B}\colon B\in\sigma\right),
\end{aligned}$$

and

$$\mathscr{C}\left(\left(\mathfrak{c}_j J_{l_B}\right)^{\#B}\colon B\in\sigma\right)=\mathscr{C}\left(\left(\mathfrak{c}_j J_{l_k}\right)^{\#\sigma_k}\colon k\in[\#\sigma]\right).$$

Putting everything together, we obtain for $\mathscr{C}_i\left(M_j\right)$

$$\sum_{s=1}^{i}\sum_{l_1,\dots,l_s=2}^{n}\sum_{\sigma\in\mathcal{P}_{[i],s}}\mathscr{C}\left(\left(\mathfrak{c}_j J_{l_k}\right)^{\#\sigma_k}\colon k\in[\#\sigma]\right)\prod_{m=1}^{\#\sigma}\mathrm{Li}_{1-\#\sigma_m}\left(r\left[r+\binom{l_m}{2}\right]^{-1}\right).$$

Part II. We are left to show $\mathscr{C}_1\left(M_j\right)=2r/j$. From the above formula we obtain for $i=1$

$$\begin{aligned}
\mathscr{C}_1\left(M_j\right)&=\sum_{l=2}^{n}\mathscr{C}\left(\mathfrak{c}_j J_l\right)\mathrm{Li}_0\left(r\left[r+\binom{l}{2}\right]^{-1}\right)\\
&=r\sum_{l=2}^{n}\binom{l}{2}^{-1}\mathscr{C}\left(\mathfrak{c}_j J_l\right),
\end{aligned}$$

since by the definition of the polylogarithm

$$\mathrm{Li}_0\left(r\left[r+\binom{l}{2}\right]^{-1}\right)=\sum_{k\geq 1}\left(\frac{r}{r+\binom{l}{2}}\right)^k=\frac{r+\binom{l}{2}}{\binom{l}{2}}-1=\frac{r}{\binom{l}{2}}.$$

Recalling the formula for the hitting probabilities $h(\pi,\rho)$ of Kingman's coalescent from Remark 3.25, we obtain

$$\begin{aligned}\mathscr{C}\left(\mathfrak{c}_j J_l\right) = \mathbb{E}\mathfrak{c}_j J_l &= \sum_{\pi\in\mathcal{P}_{[n],l}} \mathbb{E}[\mathfrak{c}_j J_l | J_l = \pi] h(\Delta_n,\pi)\\ &= \begin{bmatrix} n \\ l \end{bmatrix}^{-1} \sum_{\pi\in\mathcal{P}_{[n],l}} \mathfrak{c}_j\pi \prod_{B\in\pi} \#B! = \begin{bmatrix} n \\ l \end{bmatrix}^{-1} \frac{n!}{(n-j)!} \begin{bmatrix} n-j \\ l-1 \end{bmatrix},\end{aligned}$$

where the last step follows from

$$\begin{aligned}\sum_{\pi\in\mathcal{P}_{[n],l}} \mathfrak{c}_j\pi \prod_{B\in\pi} \#B! &= \sum_{\substack{m_1,\dots,m_n\geq 0\\ |m|=l, \|m\|=n}} \frac{n!}{\prod_{l=1}^n l!^{m_l} m_l!} m_j \prod_{l=1}^n l!^{m_l}\\ &= \frac{n!}{(n-j)!} \sum_{\substack{m_1,\dots,m_n\geq 0\\ |m|=l-1, \|m\|=n-j}} \frac{n!}{\prod_{l=1}^n l!^{m_l} m_l!} \prod_{l=1}^n l!^{m_l}\\ &= \frac{n!}{(n-j)!} \begin{bmatrix} n-j \\ l-1 \end{bmatrix}.\end{aligned}$$

We thus have

$$\mathscr{C}_1\left(M_j\right) = r\frac{n!}{(n-j)!} \sum_{l=2}^n \binom{l}{2}^{-1} \begin{bmatrix} n \\ l \end{bmatrix}^{-1} \begin{bmatrix} n-j \\ l-1 \end{bmatrix}.$$

Notice that

$$\begin{aligned}\begin{bmatrix} n \\ l \end{bmatrix}^{-1} \begin{bmatrix} n-j \\ l-1 \end{bmatrix} &= \frac{\binom{n-j-1}{l-2}\frac{(n-j)!}{(l-1)!}}{\binom{n}{l}\frac{(n-1)!}{(l-1)!}}\\ &= \frac{\binom{n-j-1}{l-2}(n-j)!}{\binom{n}{l}(n-1)!}\\ &= \frac{(n-j)!}{(n-1)!}\frac{(n-j-1)!l!(n-l)!}{(l-2)!(n-j-l+1)!n!}\\ &= 2\frac{(n-j)!}{(n-1)!}\binom{l}{2}\frac{(n-j-1)!}{n!}\frac{(n-l)!}{(n-j-l+1)!},\end{aligned}$$

and therefore

$$\begin{aligned}
\mathscr{C}_1\left(M_j\right) &= 2r\frac{(n-j-1)!}{(n-1)!}\sum_{l=2}^{n}\frac{(n-l)!}{(n-j-l+1)!}\\
&= 2r\frac{(j-1)!(n-j-1)!}{(n-1)!}\sum_{l=2}^{n}\binom{n-l}{j-1}\\
&= 2r\frac{\binom{n-1}{j}}{\binom{n-1}{j}j} = 2r/j,
\end{aligned}$$

where we used $\sum_{l=2}^{n}\binom{n-l}{j-1} = \binom{n-1}{j}$ in the last step. □

We restricted ourselves to Kingman's n-coalescent and studied the height and total length of the corresponding genealogical tree, as well as the marginals of the site frequency spectrum. Our calculations suggest that cumulants are rather well-suited for this purpose.

Three directions for future research seem rather obvious, namely a) to work out the joint cumulants of the jump chain J of Kingman's n-coalescent in order to simplify the result in Theorem 4.14, b) to extend this result to the joint cumulants of the site frequency spectrum and c) to study functionals of coalescents other than Kingman's coalescent via cumulants. To this end, the reader should recall that for any n-coalescent Π^n, conditional on its jump chain J the waiting times are independent exponentials (which, of course, is true for any Markov chain), and this fact may be exploited via conditional cumulants. Moreover, cumulants have been used extensively for studying asymptotics, e.g. central limit theorems, as well as for statistical inference, see, for instance, [68] and [49]. Another direction for future research would be to apply this theory to the functionals of Λ coalescents.

Chapter 5

Hydrodynamic limit

Recall that for any multiple merger coalescent Π the partition $\Pi(t)$ is exchangeable for any $t > 0$, where the notion of an exchangeable random partition of $\mathbb{N}$ is explained in Appendix A. In other words, the law of $\Pi(t)$ only depends on the sizes of its blocks, but not on the specific elements contained in a block. Because of this exchangeability the block size spectrum

$$(\mathfrak{c}_1 \Pi^n(t), \dots, \mathfrak{c}_n \Pi^n(t)) \quad (n \geq 2, t \geq 0)$$

contains almost all information on the random partition $\Pi^n(t)$, namely up to the labelling of the elements in the blocks. Recall that $\mathfrak{c}_i \pi$ counts the number of blocks of size i in $\pi \in \mathcal{P}_{[n]}$. In this chapter we study various aspects of the block size spectrum for different coalescent processes. To the best of the authors' knowledge the block size spectrum of Λ coalescents has not been studied before in the literature, except from the hydrodynamic limit for Kingman's coalescent that we will encounter shortly.

In the setting of finite sample size n we characterize in Section 5.1 the block size spectrum for the Kingman and the Bolthausen-Sznitman coalescent via its factorial moments. Crucial for our derivation are the transition probabilities of Π, which so far have only been derived for the Kingman coalescent [44, Equation (2.4)] and the Bolthausen-Sznitman coalescent [17, Proposition 1.4], see also Corollary 3.17. In the case of Bolthausen-Sznitman similar calculations lead us to a beautiful formula for a somewhat mysterious quantity: the logarithmic generating function of the sequence $(u \mapsto \mathbb{E}e^{-uN^k(t)})_{k\geq 1}$ of Laplace transforms of $N^k(t)$. However, the consideration of this quantity was inspired by a similar one that was introduced and studied by Pollaczek [62] (see, for instance, Equation (3.29) in [62]) in queuing theory. As an application of this result we derive the Laplace transform of $N^n(t)$. This can be viewed as complementing the spectral decomposition of the transition matrix of $N^n(t)$ and the formula

$$\mathbb{E}N^n(t) = \frac{\Gamma(n+e^{-t})}{\Gamma(n)\Gamma(1+e^{-t})} \quad (t \geq 0)$$

in [57, Equation (1.2)] for the mean of $N^n(t)$.

In Section 5.2 we turn to a scaling limit of $\{\mathfrak{c}\Pi^n(t), t \geq 0\}$ as $n \to \infty$ for beta coalescents that come down from infinity. If we rescale space by n^{-1} and suitably rescale time, this limit exists and is deterministic. In the statistical physics literature (e.g. systems modeled by Smoluchowski's coagulation equations) this kind of limit is referred to as the hydrodynamic limit. In the case of Kingman's coalescent this hydrodynamic limit is well-known, cf. the survey paper by Aldous [5].

5.1 Block size spectrum

In this section we study the block size spectrum in the setting of finite sample size n for two coalescents: the Kingman and the Bolthausen-Sznitman coalescent.

5.1.1 Kingman

For Kingman's coalescent the transition probabilities of $\varPi^n$ have been derived by Kingman, namely [44, Equations (2.3) and (2.5)] yield

$$\mathbb{P}\left\{\varPi^n(t) = \pi\right\} = \mathbb{P}\left\{N^n(t) = \#\pi\right\} h(\varDelta_n, \pi), \tag{5.1}$$

where $h(\varDelta_n, \pi)$ denotes the probability that $\varPi^n$ hits π when started from $\varDelta_n$, as defined in Remark 3.25. This allows us to derive the following formula for the descending factorials of $\mathfrak{c}\varPi^n(t)$.

Proposition 5.1 (Descending factorials of block size spectrum; Kingman coalescent)**.** *For $k \in \mathbb{N}_0^n$ we have*

$$\mathbb{E}(\mathfrak{c}_1\varPi^n(t))^{\underline{k_1}} \cdots (\mathfrak{c}_n\varPi^n(t))^{\underline{k_n}} = \frac{\mathbb{E}N^n(t)^{\underline{|k|}}(N^n(t)-1)^{\underline{|k|}}(n-N^n(t))^{\underline{\|k\|-|k|}}}{(n-1)^{\underline{\|k\|}}}. \tag{5.2}$$

Proof. First consider the moments given that $\varPi^n(t)$ has m blocks. Using (5.1) we have

$$
\begin{aligned}
&\mathbb{E}(\mathfrak{c}_1 \mathit{\Pi}^n(t))^{\underline{k_1}} \cdots (\mathfrak{c}_n \mathit{\Pi}^n(t))^{\underline{k_n}} 1_{\{\mathit{\Pi}^n(t)=m\}} \qquad (5.3)\\
&= \sum_{\pi \in \mathcal{P}_{[n],m}} (\mathfrak{c}_1 \pi)^{\underline{k_1}} \cdots (\mathfrak{c}_n \pi)^{\underline{k_n}} \mathbb{P}\{\mathit{\Pi}^n(t) = \pi\}\\
&= \mathbb{P}\{N^n(t) = m\} \begin{bmatrix} n \\ m \end{bmatrix}^{-1} \sum_{\pi \in \mathcal{P}_{[n],m}} (\mathfrak{c}_1 \pi)^{\underline{k_1}} \cdots (\mathfrak{c}_n \pi)^{\underline{k_n}} \prod_{B \in \pi} \# B!\\
&= \mathbb{P}\{N^n(t) = m\} \begin{bmatrix} n \\ m \end{bmatrix}^{-1} \frac{n!}{(n - \|k\|)!} \begin{bmatrix} n - \|k\| \\ m - |k| \end{bmatrix}\\
&= \mathbb{P}\{N^n(t) = m\}\, m^{\underline{|k|}} (m-1)^{\underline{|k|}} (n-m)^{\underline{\|k\| - |k|}} / (n-1)^{\underline{\|k\|}},
\end{aligned}
$$

since

$$
\begin{aligned}
&\sum_{\pi \in \mathcal{P}_{[n],m}} (\mathfrak{c}_1 \pi)^{\underline{k_1}} \cdots (\mathfrak{c}_n \pi)^{\underline{k_n}} \prod_{B \in \pi} \# B!\\
&= \sum_{\substack{u_1, \ldots, u_n \geq 0 \\ \|u\| = n, |u| = m}} \frac{n!}{\prod_{i=1}^n (i!)^{u_i} u_i!} \prod_{i=1}^n u_i^{\underline{k_i}} (i!)^{u_i}\\
&= \sum_{\substack{\tilde{u}_1, \ldots, \tilde{u}_n \geq 0 \\ \|\tilde{u}\| = n - \|k\|, |\tilde{u}| = m - |k|}} \frac{n!}{\prod_{i=1}^n (i!)^{\tilde{u}_i} \tilde{u}_i!} \prod_{i=1}^n (i!)^{\tilde{u}_i}\\
&= \frac{n!}{(n - \|k\|)!} \sum_{\substack{\tilde{u}_1, \ldots, \tilde{u}_n \geq 0 \\ \|\tilde{u}\| = n - \|k\|, |\tilde{u}| = m - |k|}} \frac{(n - \|k\|)!}{\prod_{i=1}^n (i!)^{\tilde{u}_i} \tilde{u}_i!} \prod_{i=1}^n (i!)^{\tilde{u}_i}\\
&= \frac{n!}{(n - \|k\|)!} \begin{bmatrix} n - \|k\| \\ m - |k| \end{bmatrix}
\end{aligned}
$$

Summing both sides of equation (5.3) over m we obtain the claim. □

It is somewhat surprising that a formula for the moments of $N^n(t)$ is rather complicated. Griffiths [36, p. 406 with mutation parameter $\theta = 0$] showed that such a formula is given by

$$\mathbb{E}N^n(t)^{\underline{k}} = \sum_{l=k}^{n} e^{-l(l-1)/2}(2l-1)\binom{l-1}{k-1} l^{\overline{k-1}} \frac{n^{\underline{l}}}{n^{\overline{l}}}. \tag{5.4}$$

One can compute the right hand side in (5.2) solely in terms of descending factorial moments of $N^n(t)$, but since this formula is not very illuminating, we omit it here. However, we do give a formula for the mean of the entries in the block size spectrum without proof:

$$\mathbb{E}\mathfrak{c}_j \mathit{\Pi}^n(t) = \sum_{l=0}^{j-1} (-1)^l \binom{j-1}{l} \frac{\mathbb{E}N^n(t)^{\overline{l+2}}}{(n-1)^{\overline{l+1}}}. \tag{5.5}$$

This formula can be proven using the identity

$$x^{\underline{a}}(n-x)^{\underline{b}} = \frac{\sum_{k=0}^{b}(-1)^k \binom{b}{k} x^{\underline{a+k}}(n-(a+k))!}{(n-(a+b))!} \qquad (0 < x \leq n, a, b \in \mathbb{N}),$$

which in turn can be shown by an argument similar to the one given in the proof of [28, Lemma A.2, Appendix A].

5.1.2 Bolthausen-Sznitman

We now study the block size spectrum for the Bolthausen-Sznitman n-coalescent. Recall that for a partition $\pi = \{A_1, \dots, A_k\} \in \mathcal{P}_{[n]}$ such that $n_i = \# A_i$ the transition probabilities of the Bolthausen-Sznitman n-coalescent are given by

$$\mathbb{P}\left\{\mathit{\Pi}^n(t) = \pi\right\} = p_t(n_1, \dots, n_k) = \frac{(k-1)!}{(n-1)!} e^{-(k-1)t} \prod_{i=1}^{k} (1-e^{-t})^{\overline{n_i - 1}}, \tag{5.6}$$

cf. Corollary 3.17. The following lemma allows us to rewrite the right hand side of (5.6) in a form that will be more convenient for our subsequent calculations.

Lemma 5.2. *For $i \in [n]$ let $m_i \coloneqq \#\{j : n_j = i\}$. Then*

$$p_t(n_1, \dots, n_k) = e^t (-1)^k \frac{(k-1)!}{(n-1)!} \prod_{i=1}^{n} \left((-e^{-t})^{\overline{i}} \right)^{m_i}, \tag{5.7}$$

Proof. Note that $(-1)^l x^{\underline{l}} = (-x)^{\overline{l}}$, so $x(1-x)^{\overline{l-1}} = -(-x)^{\overline{l}}$ and therefore

$$\prod_{i=1}^{k} (1-x)^{\overline{n_i - 1}} = x^{-k} \prod_{i=1}^{k} -(-x)^{\overline{n_i}} = x^{-k} \prod_{i=1}^{n} \left(-(-x)^{\overline{i}} \right)^{m_i}.$$

Applying this to (5.6) yields the statement. □

Theorem 5.3 (Logarithmic generating function of Laplace transforms of block counting process)**.** *For $n \in \mathbb{N}$, $u, t, \xi \geq 0$, such that $-1 < e^{-u}[(1-\xi)^{e^{-t}} - 1] \leq 1$ the logarithmic generating function of the Laplace transforms of $N^n(t)$, $n \in \mathbb{N}$, is given by*

$$\Phi(\xi, u, t) \coloneqq \sum_{k \geq 1} \mathbb{E}\left[\exp\left\{ -u N^k(t) \right\} \right] \frac{\xi^k}{k} = -e^t \log\left(1 + e^{-u}[(1-\xi)^{e^{-t}} - 1] \right). \tag{5.8}$$

Proof. The convergence of the series is clear for $-1 \leq \xi < 1$, since $|\mathbb{E}[\exp\{-uN^k(t)\}]| \leq 1$ for all $t, u \geq 0$, $k \in \mathbb{N}$, as $N^k(t) \geq 1$ a.s. for all $u \geq 0$ and $k \in \mathbb{N}$. Using (5.7) and [61, Equation (54)], we have

$$\begin{aligned}
&\mathbb{P}\{N^n(t)=k\} &(5.9)\\
&=(-1)^k e^t \frac{(k-1)!}{(n-1)!} \sum_{\substack{m_1,\dots,m_n\geq 0\\ |m|=k,\|m\|=n}} \frac{n!}{\prod_{i=1}^n (i!)^{m_i} m_i!} \prod_{i=1}^n \left((-e^{-t})^{\overline{i}}\right)^{m_i}\\
&=(-1)^k e^t \frac{(k-1)!}{(n-1)!} B_{n,k}(w_\bullet),
\end{aligned}$$

where the sequence $w_\bullet := (w_k)_{k\in\mathbb{N}}$ is given by $w_k := (-e^{-t})^{\overline{k}}$. Note that the exponential generating function of $(w_k)_{k\geq 1}$ is given by

$$w(\xi) := \sum_{k\geq 1} w_k \frac{\xi^k}{k!} = \sum_{k\geq 1} (-e^{-t})^{\overline{k}} \frac{\xi^k}{k!} = (1-\xi)^{e^{-t}} - 1. \tag{5.10}$$

For the Laplace transform of the number of blocks in $\Pi^n(t)$ we therefore obtain

$$\mathbb{E}\exp\{-uN^n(t)\} = \frac{e^t}{(n-1)!} \sum_{k=1}^n (-e^{-u})^k (k-1)! B_{n,k}(w_\bullet) = e^t \frac{B_n(v_\bullet, w_\bullet)}{(n-1)!}, \tag{5.11}$$

where $v_\bullet = (v_k)_{k\in\mathbb{N}}$ is given by $v_k := (k-1)!(-e^{-u})^k$. The sequence $(v_k)_{k\geq 1}$ has exponential generating function

$$v(\xi) := \sum_{k\geq 1} v_k \frac{\xi^k}{k!} = \sum_{k\geq 1} \frac{(-e^{-u}\xi)^k}{k} = -\log(1+e^{-u}\xi), \tag{5.12}$$

whenever $-1 < e^{-u}\xi \leq 1$. Using [61, Equation (57)] we find

$$\sum_{n\geq 1} \mathbb{E}[\exp\{-uN^n(t)\}] \frac{\xi^n}{n} = e^t v(w(\xi)) = -e^t \log(1+e^{-u}[(1-\xi)^{e^{-t}} - 1]),$$

whenever $-1 < e^{-u}[(1-\xi)^{e^{-t}} - 1] \leq 1$, which is the claim. $\square$

In (5.11) we already saw a formula for the Laplace transform of $N^n(t)$ in terms of complete Bell polynomials. Corollary 5.4 yields a more explicit formula for this Laplace transform in terms of generalized Stirling numbers.

Corollary 5.4 (Laplace transform of restricted block counting process)**.** *For $n \in \mathbb{N}$ the Laplace transform of $N^n(t)$ is given by*

$$\mathbb{E}\exp\{-uN^n(t)\} = \frac{\partial_\xi^n \varPhi(\xi, u, t)|_{\xi=0}}{(n-1)!} = \frac{e^{t-u}}{(n-1)!}\sum_{b=1}^{n}(b-1)!e^{-bt}S_{n,b}^{1,e^{-t}}, \tag{5.13}$$

for $u \geq 0$.

Proof. From the formula for the logarithmic generating function of the Laplace transform of the block counting process, Theorem 5.3, we have

$$\mathbb{E}\exp\{-uN^n(t)\} = (n-1)!^{-1}\partial_\xi^n \varPhi(\xi, u, t)|_{\xi=0}.$$

In order to compute the nth partial derivative of $\varPhi$ with respect to ξ, we first write $\varPhi$ as a functional composition in a convenient way, and then apply Faà di Bruno's formula. Letting

$$g(x) := (1-x)^{e^{-t}}, \tag{5.14}$$

and

$$f(x) := -e^t \log(1 + e^{-u}\{x-1\}),$$

we have

$$\varPhi(\xi, u, t) = -e^{-t} \log(1 + e^{-u}[(1-\xi)^{e^{-t}} - 1]) = (f \circ g)(\xi).$$

Recall that Faà di Bruno's formula [41] states that the ith derivative of the concatenation $f \circ g$ of two functions f, g is given by

$$\frac{d^i}{dx^i}(f \circ g)(x) = \sum_{\pi \in \mathcal{P}_{[i]}} f^{(\#\pi)}(g(x)) \prod_{B \in \pi} g^{(\#B)}(x), \tag{5.15}$$

where f, g are real functions that are at least i times differentiable, and $f^{(j)}$ denotes the jth derivative of f.

In our case, we have $g(0) = 1$,

$$g^{(j)}(x) = (-1)^j (e^{-t})^{\underline{j}} (1-x)^{e^{-t}-j},$$

and, hence, $g^{(j)}(0) = (-1)^j (e^{-t})^{\underline{j}}$. Moreover,

$$\begin{aligned} f^{(j)}(x) &= -e^t (\log \circ (1 + e^{-u}\{\cdot - 1\}))^{(j)}(x) \\ &= -e^{t-u}(-1)^{j-1}(j-1)!(1 + e^{-u}\{x-1\})^{-j}. \end{aligned}$$

This holds true, since, by another application of Faà di Bruno's formula

$$\begin{aligned} \frac{d^l}{dx^l}(\log \circ h)(x) &= \sum_{\pi \in \mathcal{P}_{[l]}} \log^{(\#\pi)}(h(x)) \prod_{B \in \pi} h^{(\#B)}(x) \\ &= \sum_{\pi \in \mathcal{P}_{[l]}} (-1)^{\#\pi - 1}(\#\pi - 1)! h(x)^{-\#\pi} \prod_{B \in \pi} h^{(\#B)}(x), \end{aligned}$$

and, letting $h(x) \coloneqq 1 + e^{-u}\{x-1\}$, we have

$$\frac{d^j}{dx^j}h(x) = \begin{cases} e^{-u} & \text{if j=1,} \\ 0 & \text{otherwise.} \end{cases}$$

Consequently,

$$\begin{aligned}
&\frac{d^j}{dx^j}(\log \circ (1 + e^{-u}\{x-1\}))(x) \\
&= \sum_{\pi \in \mathcal{P}_{[j]}} (-1)^{\#\pi - 1}(\#\pi - 1)!(1 + e^{-u}\{x-1\})^{-\#\pi} \prod_{B \in \pi} h^{(\#B)}(x) \\
&= e^{-u}(-1)^{j-1}(j-1)!(1 + e^{-u}\{x-1\})^{-j}.
\end{aligned}$$

Since $f^{(j)}(g(\xi))|_{\xi=0} = f^{(j)}(1) = -(-1)^{j-1}(j-1)!e^{t-u}$, we compute

$$\begin{aligned}
\partial_\xi^n \Phi(\xi, u, t)|_{\xi=0} &= \sum_{\pi \in \mathcal{P}_{[n]}} f^{(\#\pi)}(g(0)) \prod_{B \in \pi} g^{(\#B)}(0) \\
&= -e^{t-u} \sum_{\pi \in \mathcal{P}_{[n]}} (-1)^{\#\pi - 1}(\#\pi - 1)! \prod_{B \in \pi} (-1)^{\#B}(e^{-t})^{\underline{\#B}} \\
&= e^{t-u} \sum_{b=1}^{n} (b-1)! \sum_{\pi \in \mathcal{P}_{[n]}, b} \prod_{B \in \pi} (e^{-t})^{\underline{\#B}} \\
&= e^{t-u} \sum_{b=1}^{n} (b-1)! e^{-tb} \sum_{\pi \in \mathcal{P}_{[n]}, b} \prod_{B \in \pi} (e^{-t} - 1)^{\underline{\#B - 1}} \\
&= e^{t-u} \sum_{b=1}^{n} (b-1)! e^{-bt} S_{n,b}^{1, e^{-t}}.
\end{aligned}$$

The claim follows. $\square$

We basically derived the Laplace transform of $N^n(t)$ in the Bolthausen-Sznitman case for the sake of completeness, and in order to show that in the two cases of

Kingman and Bolthausen-Sznitman some progress can be made even in the setting of finite sample size n, since the transition probabilities are known in these cases. To the best of the author's knowledge the transition probabilities are not known for other coalescent processes. However, in order to study the scaling limit of the block size spectrum as $n \to \infty$, one does not need the transition probabilities, as the next section shows.

5.2 Hydrodynamic limit

This section is a modified version of the preprint [50] written in collaboration with Luke Miller (University of Oxford). We quantify the manner in which the beta coalescent $\Pi = \{\Pi(t), t \geq 0\}$, with parameters $a \in (0,1)$, $b > 0$, comes down from infinity. Approximating Π by its restriction Π^n to $[n]$, the suitably rescaled block counting process $n^{-1}\#\Pi^n(tn^{a-1})$ has a deterministic limit, $c(t)$, as $n \to \infty$. An explicit formula for $c(t)$ is provided in Theorem 5.10.

The block size spectrum $(\mathfrak{c}_1\Pi^n(t), \ldots, \mathfrak{c}_n\Pi^n(t))$ captures more refined information about the coalescent tree corresponding to Π. Using the corresponding scaling, the block size spectrum also converges to a deterministic limit as $n \to \infty$. This limit is characterized by a system of ordinary differential equations whose ith solution is a complete Bell polynomial, depending only on $c(t)$ and a, that we work out explicitly, see Corollary 5.16.

For other studies of asymptotic properties of beta coalescents we refer the reader to Berestycki, Berestycki and Schweinsberg [9] and [10]. For work on the first- and second-order approximation of the ancestral process $N(t)$ as $t \to 0$ the reader may want to consult Berestycki, Berestycki and Limic [7] and Limic, Talarczyk [47], [48].

5.2.1 Introduction and summary of results

One may wonder whether the (random) coalescent tree $\mathcal{T}$ corresponding to Π (see Section 2.3.2)) can be described more explicitly. One way to study $\mathcal{T}$ 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 Λ n-coalescent with leaves labelled $l_1, \ldots, l_n$, as is apparent from the consistency of Λ coalescents. As we increase the sample size n, we explore larger and larger subtrees of $\mathcal{T}$. However, the topology of a subtree spanned by n leaves is rather involved.

In order to work out explicitly the asymptotic shape of a subtree when the number n of leaves grows without bound in what follows, we restrict ourselves to beta coalescents that come down from infinity. Recall the definition of a beta coalescent from Examples 2.5. Notice that according to Schweinsberg's characterization of coalescents that come down from infinity, equation (2.17), the beta(a, b) coalescent comes down from infinity if and only if $a \in (0, 1)$, cf. Example 15 in [66]. Without further mention, we assume hereafter that Λ is the beta(a, b) distribution for some $a \in (0, 1)$, $b > 0$.

In the first part of our note we study the behaviour of the frequency of the total number of blocks of Π for small times. We show in Theorem 5.10 that with a suitable scaling of time the frequency of the total number of blocks of Π has a non-trivial scaling limit, more precisely, as $n \to \infty$ we have convergence to a deterministic limit

$$\{n^{-1}\#\Pi^n(tn^{a-1}), t \geq 0\} \to \left\{ \left(1 + \frac{\Gamma(a+b)}{(2-a)\Gamma(b)} t\right)^{\frac{1}{a-1}}, t \geq 0 \right\}$$

in $D_{[0,1]}[0, \infty)$, the space of right-continuous functions with left limits from $[0, \infty)$

to $[0,1]$, in the Skorokhod topology. We obtain the limit as the solution $c(t)$ of the ordinary differential equation

$$\frac{d}{dt}c(t) = -\frac{\Gamma(a+b)}{(1-a)(2-a)\Gamma(b)}c(t)^{2-a} \quad (t \geq 0), \quad c(0) = 1, \tag{5.16}$$

of Bernoulli type.

In the second part we work out in Proposition 5.12 a scaling limit for the so-called block size spectrum

$$\{\mathfrak{c}\Pi^n(t), t \geq 0\} \tag{5.17}$$

of Π^n, where for $\pi \in \mathcal{P}_{[n]}$ $\mathfrak{c}\pi := (\mathfrak{c}_1\pi, \ldots, \mathfrak{c}_n\pi)$ denotes the so-called type of π defined by $\mathfrak{c}_i\pi := \#\{B \in \pi\colon \#B = i\}$ for any $i \in \mathbb{N}$. The block size spectrum is a summary statistic that encodes "almost all" information on a subtree of the coalescent tree spanned by n leaves. The tree spanned by the leaves $l_1, \ldots, l_n$, also called the spanning tree, is the smallest subtree in the coalescent tree with leaves $l_1, \ldots, l_n$. To be more specific, given $\{\mathfrak{c}\Pi^n(t), t \geq 0\}$ one can recover the corresponding subtree in the coalescent tree up to the choice of branches that merge at each branch point and up to the labelling of the leaves. In fact, in order to reconstruct this subtree on n leaves, due to the exchangeability of Π all one needs to do is choose at each branch point k branches among the existing branches uniformly at random (if k branches are to merge), and randomly label its leaves by $1, \ldots, n$ (i.e. according to a permutation of $\{1, \ldots, n\}$ picked uniformly at random). We focus on the behaviour of the average number of blocks of a given size, and therefore rescale the state space of the block size spectrum by n^{-1}.

Fix $d \in \mathbb{N}$ arbitrarily. We show that the evolution of the frequency of blocks of size $\leq d$ in Π^n with the same time scaling as before converges to a deterministic limit, namely

$$\{n^{-1}(\mathfrak{c}_1 \Pi^n(tn^{a-1}), \ldots, \mathfrak{c}_d \Pi^n(tn^{a-1})), t \geq 0\} \to \{(c_1(t), \ldots, c_d(t)), t \geq 0\},$$

as $n \to \infty$ in $D_{[0,1]^d}[0, \infty)$ in the Skorokhod topology, where by Corollary 5.16

$$c_i(t) = \frac{c(t)^{2-a}}{i!} B_i \left(\left(\frac{1}{1-a} \right)^{\overline{\bullet}} (-c(t)^{1-a})^{\bullet - 1}, (1-a)^{\overline{\bullet}} \right) \quad (i \in \mathbb{N}), \tag{5.18}$$

and for any function $f\colon \mathbb{N} \to \mathbb{R}$ we write $f(\bullet)$ as a shorthand for the sequence $(f(k))_{k \geq 1}$. We find the functions $c_i(t)$ by computing their generating function

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

Theorem 5.15 states that $\mathscr{G}$ is given by

$$\mathscr{G}(t, x) = c(t) - \left((1-x)^{a-1} + \frac{\Gamma(a+b)}{(2-a)\Gamma(b)} t \right)^{\frac{1}{a-1}} \quad (x \in (-1, 1), t \geq 0). \tag{5.19}$$

It is remarkable to see the similarity between the term subtracted from $c(t)$ in formula (5.19) for $\mathscr{G}$, namely

$$g(t, x) := c(t) - \mathscr{G}(t, x) = \sum_{i \geq 1} c_i(t)(1 - x^i) = \left((1-x)^{a-1} + \frac{\Gamma(a+b)}{(2-a)\Gamma(b)} t \right)^{\frac{1}{a-1}}$$

and $c(t)$. In fact, $g(t,x)$ solves the partial differential equation

$$\partial_t g(t,x) = -\frac{\Gamma(a+b)}{(1-a)(2-a)\Gamma(b)} g(t,x)^{2-a} \quad (x \in (0,1), t \geq 0) \tag{5.20}$$

with boundary condition $g(0,x) = 1-x$, and $g(t,0) = c(t)$. This partial differential equation should be compared to the ordinary differential equation for $c(t)$, equation (5.16). We interpret this as a form of self-similarity of the limiting block frequency spectrum of the beta coalescents expressed in terms of generating functions.

5.2.2 Block counting process

Before we turn to the scaling limit of the block size spectrum we study the simpler block counting process $N^n = \{N^n(t), t \geq 0\}$. Recall that $N^n(t) = \#\Pi^n(t)$ counts the number of blocks in $\Pi^n(t)$. For $a, b > 0$ we will repeatedly use the identities

$$B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b),$$

see [6, Theorem 1.1.4], and, since $\Gamma(a+1) = a\Gamma(a)$,

$$\Gamma(a+n) = a^{\overline{n}}\Gamma(a) \quad (n \in \mathbb{N}).$$

In order to find the right time scaling for N^n, notice that the rate at which the number of blocks decreases has asymptotics

$$\sum_{l=2}^{n} \binom{n}{l} \lambda_{n,l}(l-1) \sim \frac{\Gamma(a+b)}{(1-a)(2-a)\Gamma(b)} n^{2-a},$$

as $n \to \infty$, which is proved in Lemma 5.8. Consequently, in order to see a nontrivial limit of N^n when its state space is rescaled by n^{-1}, we should rescale time by a factor on the order of n^{a-1}. Let therefore $C^n := \{C^n(t), t \geq 0\}$ be defined by

$$C^n(t) := n^{-1} N^n(t\upsilon_n) \qquad (n \geq 2, t \geq 0), \tag{5.21}$$

where $\upsilon_n \sim n^{a-1}$ as $n \to \infty$.

The process C^n is a continuous-time Markov chain with state space $E_n := n^{-1}[n]$, initial state $C^n(0) = 1$, absorbing state n^{-1} and evolves according to the following dynamics: for $c \in E_n$ and $2 \leq l \leq nc$

$$\text{a transition} \qquad c \mapsto c - \frac{l-1}{n} \qquad \text{occurs at rate} \quad \binom{nc}{l} \lambda_{nc,l}. \tag{5.22}$$

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

$$\lambda_{m,k} := \int_0^1 x^{k-2}(1-x)^{m-k} \Lambda(dx) = \frac{B(k-2+a, m-k+b)}{B(a,b)} = \frac{a^{\overline{k-2}} b^{\overline{m-k}}}{(a+b)^{\overline{m-2}}}. \tag{5.23}$$

Remark 5.5. Equation (5.23) yields the recursive formula

$$\lambda_{m,k+1} = \frac{a+k-2}{b+m-k-1} \lambda_{m,k} \qquad (2 \leq k \leq m-1). \tag{5.24}$$

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 arbitray Λ given in [60, Lemma 18]. Combined, these formulae can be used to efficiently compute the rates of the beta(a, b) coalescent iteratively.

We prepare the proof of the diffusion approximation for C^n by establishing some Lemmas. For $k \in \mathbb{N}$ define

$$\gamma_n^{(k)} := \sum_{l=2}^{n} \binom{n}{l} \lambda_{n,l} (l-1)^k, \quad \text{and} \quad \gamma_n^{(\underline{k})} := \sum_{l=2}^{n} \binom{n}{l} \lambda_{n,l} l^{\underline{k}}. \tag{5.25}$$

The asymptotic behaviour of $\gamma_n^{(1)}, \gamma_n^{(2)}$ as $n \to \infty$ will play a crucial rôle in the proof of our first result, Theorem 5.10.

Lemma 5.6. *For $a, b > 0$, a natural number $n \in \mathbb{N}$ and an integer $z \in \mathbb{Z}$ we have*

$$\frac{a^{\overline{n}}}{b^{\overline{n+z}}} \sim \frac{\Gamma(b)}{\Gamma(a)} n^{a-b-z}, \tag{5.26}$$

as $n \to \infty$.

Proof. Since $\Gamma(a+n)/\Gamma(n) \sim n^a$ as $n \to \infty$, we have

$$\frac{a^{\overline{n}}}{b^{\overline{n+z}}} = \frac{\Gamma(a+n)}{\Gamma(a)} \frac{\Gamma(b)}{\Gamma(b+n+z)} \sim \frac{\Gamma(b)}{\Gamma(a)} n^{a-b-z},$$

as $n \to \infty$. $\square$

Moreover, for $d \in \mathbb{N}$, $x \in \mathbb{R}^d$ and $k \in \mathbb{N}_0^d$ let $x^{\overline{k}} := \prod_{i=1}^d x_i^{\overline{k_i}}$ and $|x| = \sum_{i=1}^d |x_i|$.

Lemma 5.7. *Fix $d \in \mathbb{N}$ and $k, n \in \mathbb{N}_0^d$. Then for $a, b \in \mathbb{R}$*

$$\sum_{l \in \mathbb{N}_0^d} l^{\underline{k}} a^{\overline{|l|}} b^{\overline{n-|l|}} \prod_{i=1}^{d} \binom{n_i}{l_i} = a^{\overline{|k|}} n^{\underline{|k|}} (a+b+|k|)^{\overline{|n|-|k|}}. \tag{5.27}$$

Proof. We first prove the statement for $k = 0 \in \mathbb{R}^d$ by an induction on d. Hence, for $d = 1$ the statement reads

$$\sum_{l=0}^{n} \binom{n}{l} a^{\overline{l}} b^{\overline{n-l}} = (a+b)^{\overline{n}},$$

and this is true, since the sequence $(a^{\overline{k}})_{k\geq 1}$ of rising factorial powers is a sequence of polynomials of binomial type, as is well known. Suppose now that (5.27) holds for some $d \in \mathbb{N}$. Then

$$\begin{aligned}
&\sum_{l\in\mathbb{N}_0^{d+1}} a^{\overline{|l|}} b^{\overline{|n|-|l|}} \prod_{i=1}^{d+1} \binom{n_i}{l_i} \\
&= \sum_{l\in\mathbb{N}_0^{d}} a^{\overline{|l|}} b^{\overline{n_1+\ldots+n_d-|l|}} \prod_{i=1}^{d} \binom{n_i}{l_i} \\
&\quad \times \sum_{l_{d+1}=0}^{n_{d+1}} \binom{n_{d+1}}{l_{d+1}} (a+|l|)^{\overline{l_{d+1}}} (b+n_1+\cdots+n_d-|l|)^{\overline{n_{d+1}-l_{d+1}}} \\
&= (a+b)^{\overline{n_1+\cdots+n_d}} (a+b+n_1+\ldots+n_d)^{\overline{n_{d+1}}} = (a+b)^{\overline{|n|}},
\end{aligned}$$

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

$$
\begin{aligned}
&\sum_{l \in \mathbb{N}_0^d} l^{\underline{k}} a^{\overline{|l|}} b^{\overline{|n|-|l|}} \prod_{i=1}^{d} \binom{n_i}{l_i} \\
&= \sum_{m \in \mathbb{N}_0^d} (m+k)^{\underline{k}} a^{\overline{|m|+|k|}} b^{\overline{|n|-|k|-|m|}} \prod_{i=1}^{d} \binom{n_i}{m_i + k_i} \\
&= a^{\overline{|k|}} \sum_{m \in \mathbb{N}_0^d} (a+|k|)^{\overline{|m|}} b^{\overline{|n|-|k|-|m|}} \prod_{i=1}^{d} \left[(m_i + k_i)^{\underline{k_i}} \binom{n_i}{m_i + k_i} \right] \\
&= a^{\overline{|k|}} n^{\underline{k}} \sum_{m \in \mathbb{N}_0^d} (a+|k|)^{\overline{|m|}} b^{\overline{|n|-|k|-|m|}} \prod_{i=1}^{d} \binom{n_i - k_i}{m_i} \\
&= a^{\overline{|k|}} n^{\underline{k}} (a+b+|k|)^{\overline{|n|-|k|}},
\end{aligned}
$$

which completes the proof. $\square$

Lemma 5.8. *For $a \in (0,1)$ as $n \to \infty$ we have that*

$$\gamma_n^{(1)} \sim \frac{\Gamma(a+b)}{(1-a)(2-a)\Gamma(b)} n^{2-a}. \tag{5.28}$$

Proof. Firstly, note the relation $\gamma_n^{(1)} = \gamma_n^{(\underline{1})} - \gamma_n^{(0)}$. Using Lemma 5.7 with $d = k = 1$, we find

$$\gamma_n^{(\underline{1})} = -\frac{n}{1-a} \left(\frac{(a+b-1)^{\overline{n-1}}}{(a+b)^{\overline{n-2}}} - \frac{b^{\overline{n-1}}}{(a+b)^{\overline{\overline{n-2}}}} \right).$$

The first summand in above paranthesis vanishes, if $a + b = 1$. If $a + b \neq 1$ we have

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

as $n \to \infty$ by Lemma 5.6. On the other hand, Lemma 5.6 yields

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

as $n \to \infty$, and, since $a < 1$, we obtain

$$\gamma_n^{(\underline{1})} \sim \frac{\Gamma(a+b)}{(1-a)\Gamma(b)} n^{2-a}, \tag{5.29}$$

as $n \to \infty$.

Now consider $\gamma_n^{(0)}$:

$$\begin{aligned}
\gamma_n^{(0)} = \sum_{l=2}^{n} \binom{n}{l} \lambda_{n,l} &= \frac{1}{(a+b)^{\overline{n-2}}} \sum_{l=2}^{n} \binom{n}{l} a^{\overline{l-2}} b^{\overline{n-l}} \\
&= \frac{1}{(1-a)(2-a)(a+b)^{\overline{n-2}}} \sum_{l=2}^{n} \binom{n}{l} (a-2)^{\overline{l}} b^{\overline{n-l}} \\
&= \frac{1}{(1-a)(2-a)} \left(\frac{(a+b-2)^{\overline{n}}}{(a+b)^{\overline{n-2}}} + \frac{(2-a)nb^{\overline{n-1}} - b^{\overline{n}}}{(a+b)^{\overline{n-2}}} \right) \\
&\sim \frac{\Gamma(a+b)}{(2-a)\Gamma(b)} n^{2-a},
\end{aligned}$$

where we applied Lemmata 5.6 and 5.7 once more and distinguished the cases $a + b = 2$ and $a + b \neq 2$. Putting everything together, the claim follows.

$\square$

Lemma 5.9. *As $n \to \infty$ we have that $\gamma_n^{(2)} \sim n^2$.*

Proof. We can write $\gamma_n^{(2)} = \gamma_n^{(\underline{2})} - \gamma_n^{(1)}$, since $(l-1)^2 = l^{\underline{2}} - (l-1)$. Applying Lemma

5.7 with $d = 1$ and $k = 2$ we find that

$$\begin{aligned}
\gamma_n^{(\underline{2})} &= \sum_{l=2}^{n} \binom{n}{l} \frac{a^{\overline{l-2}} b^{\overline{n-l}}}{(a+b)^{\overline{n-2}}} l^{\underline{2}} \\
&= \frac{1}{(a-2)^{\overline{2}} (a+b)^{\overline{n-2}}} \sum_{l=2}^{n} \binom{n}{l} (a-2)^{\overline{l}} b^{\overline{n-l}} l^{\underline{2}} = n(n-1).
\end{aligned}$$

From this and Lemma 5.8 we conclude $\gamma_n^{(2)} = \gamma_n^{(\underline{2})} - \gamma_n^{(1)} \sim n^2$ as $n \to \infty$. □

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 $\mathbb{R}$.

Theorem 5.10. *Let υ_n be of order n^{a-1}. Then as $n \to \infty$ we have convergence*

$$\{C^n(t), t \geq 0\} \to \{c(t), t \geq 0\} \tag{5.30}$$

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

$$\frac{d}{dt} c(t) = -\frac{\Gamma(a+b)}{(1-a)(2-a)\Gamma(b)} c(t)^{2-a} \quad (t \geq 0), \tag{5.31}$$

with boundary condition $c(0) = 1$. The solution to (5.38) is given by

$$c(t) = \left(1 + \frac{\Gamma(a+b)}{(2-a)\Gamma(b)} t\right)^{\frac{1}{a-1}} = \left(\frac{(2-a)\Gamma(b)}{(2-a)\Gamma(b) + \Gamma(a+b)t}\right)^{\frac{1}{1-a}}. \tag{5.32}$$

Proof. The jump chain $(J^n_k)_{k\geq 0}$ of $C^n(t) = n^{-1}N^n(t\upsilon_n)$ has transition probabilities

$$\mu_n(c, c - \frac{l-1}{n}) := \mathbb{P}\left\{ J^n_1 = c - \frac{l-1}{n} \middle| J^n_0 = c \right\}$$
$$= \begin{cases} \binom{nc}{l}\frac{\lambda_{nc,l}}{\lambda_{nc}} & \text{if } c > n^{-1}, 2 \leq l \leq nc, \\ 1 & \text{if } c = n^{-1}, l = 1, \\ 0 & \text{otherwise.} \end{cases}$$

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

$$\begin{aligned} \mathcal{G}_n f(c) &= \lambda_n(c) \int_{E_n} (f(c') - f(c))\mu_n(c, dc') \\ &= \upsilon_n \lambda_{nc} \sum_{l=2}^{nc} \left(f(c - \frac{l-1}{n}) - f(c) \right) \binom{nc}{l} \frac{\lambda_{nc,l}}{\lambda_{nc}} \\ &= \upsilon_n \sum_{l=2}^{nc} \left(-\frac{l-1}{n} f'(c) + R_2(\vartheta_{n,l}) \right) \binom{nc}{l} \lambda_{nc,l}, \end{aligned} \tag{5.33}$$

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

$$R_2(\vartheta_{n,l}) = \frac{1}{2}\left(\frac{l-1}{n}\right)^2 f''(\vartheta_{n,l}).$$

First notice that by Lemma 5.8

$$-\frac{\upsilon_n}{n} \sum_{l=2}^{nc} \binom{nc}{l} (l-1)\lambda_{nc,l} = -\frac{\upsilon_n}{n}\gamma^{(1)}_{nc} \to -\frac{\Gamma(a+b)}{(1-a)(2-a)\Gamma(b)} c^{2-a}$$

as $n \to \infty$, since υ_n is chosen to be of order n^{a-1}, and $a < 1$ by assumption. Since f is smooth (i.e. f has derivatives of arbitrarily high order) on $[0,1]$, f'' attains its supremum $\|f''\|_\infty \coloneqq \sup_{x\in[0,1]}|f''(x)|$. Consequently, as $n \to \infty$ we obtain

$$\begin{aligned}\upsilon_n \sum_{l=2}^{nc} \binom{nc}{l} \lambda_{nc,l} R_2(x_{n,l}) &\le \frac{\upsilon_n}{2n^2}\|f''\|_\infty \sum_{l=2}^{nc} \binom{nc}{l} \lambda_{nc,l}(l-1)^2 \\ &= \frac{\upsilon_n}{2n^2}\|f''\|_\infty \gamma_{nc}^{(2)} \sim \frac{1}{2}c^2\upsilon_n\|f''\|_\infty \to 0,\end{aligned}$$

where we used Lemma 5.9 in the third step. This shows the convergence

$$\lim_{n\to\infty} \sup_{c\in E_n} |\mathcal{G}_n f(c) - \mathcal{G} f(c)| \to 0, \tag{5.34}$$

where the operator $\mathcal{G}$ is defined by

$$\mathcal{G} f(c) \coloneqq -\frac{\Gamma(a+b)}{(1-a)(2-a)\Gamma(b)} c^{2-a} f'(c). \tag{5.35}$$

Since $[0,1] \ni c \mapsto -c^{2-a}\Gamma(a+b)/(1-a)(2-a)\Gamma(b)$ is Lipschitz continuous, [29, Theorem 2.1 in Chapter 8] yields that the set $C^\infty([0,1])$ is a core for $\mathcal{G}$, and the closure of $\{(f, \mathcal{G}f) \colon f \in C^\infty([0,1])\}$ is single-valued and generates a Feller semigroup $\{T(t)\}$ on $C([0,1])$. By [29, Theorem 2.7 in Chapter 4] there exists a diffusion 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 [29, Corollary 8.7 of Chapter 4] to show that (5.34) holds for all f in a core for the generator $\mathcal{G}$, which we have just done.

□

Remark 5.11. 1. According to Theorem 1.1 in Berestycki, Berestycki and Schweinsberg [10] for beta$(2-\alpha,\alpha)$ coalescents we have

$$\lim_{t\to 0} t^{1/(\alpha-1)} \#\Pi(t) = (\alpha\Gamma(\alpha))^{1/(\alpha-1)} \quad \text{a.s.}$$

Letting $t' := tn^{a-1}$, where $a := 2-\alpha$, this result yields

$$\begin{aligned} n^{-1}\#\Pi(tn^{a-1}) &= \frac{1}{n^{1+\frac{a-1}{\alpha-1}} t^{\frac{1}{\alpha-1}}} (tn^{a-1})^{\frac{1}{\alpha-1}} N(tn^{a-1}) \\ &\to t^{\frac{1}{1-\alpha}} (\alpha\Gamma(\alpha))^{\frac{1}{\alpha-1}} = \left(\frac{t}{(2-a)\Gamma(2-a)}\right)^{\frac{1}{a-1}} \end{aligned}$$

a.s. as $n\to\infty$, and suggests that one can extend Theorem 5.10 for beta(a,b) coalescents to the convergence

$$\{n^{-1}\#\Pi(tn^{a-1}), t\geq 0\} \to \left(\frac{\Gamma(a+b)}{(2-a)\Gamma(b)} t\right)^{\frac{1}{a-1}} \tag{5.36}$$

in $D_{[0,1]}([0,\infty))$ in the Skorokhod topology. In order to see this, notice that the boundary value for the limiting ODE changes, but not the generator calculations in the proof of Theorem 5.10 if we consider the process $\{n^{-1}\#\Pi(tn^{a-1}), t\geq 0\}$. Thus, the limiting ODE satisfies

$$\frac{d}{dt}c(t) = -\frac{\Gamma(a+b)}{(1-a)(2-a)\Gamma(b)} c(t)^{2-a} \quad (t\geq 0), \tag{5.37}$$

with boundary condition $c(0)=\infty$. In order to solve this ODE, notice that the solution to

$$\frac{d}{dt}c(t) = -\frac{\Gamma(a+b)}{(1-a)(2-a)\Gamma(b)} c(t)^{2-a} \quad (t\geq 0), \quad c(0)=C \tag{5.38}$$

is given by

$$c(t) = \left(C^{a-1} + \frac{\Gamma(a+b)t}{(2-a)\Gamma(b)} \right)^{\frac{1}{a-1}},$$

and converges to

$$\left(\frac{\Gamma(a+b)t}{(2-a)\Gamma(b)} \right)^{\frac{1}{a-1}}$$

as $C \to \infty$. It would be interesting to see whether Theorem 1.1 in [10] can be extended to the result in (5.36), and we leave this investigation for future work.

2. Kingman's coalescent can be thought of as the beta(a,b) coalescent obtained in the limit as $a \to 0$. In this case we recover for the rescaled block counting process the hydrodynamic limit

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

which is well known, cf. [75], equation (2.15).

5.2.3 Block size spectrum

For $d \in \mathbb{N}$ let the rescaled block size spectrum $(C_i^n)_{i=1}^{d+1} := (C_i^n(t), t \geq 0)_{i=1}^{d+1}$ be defined by

$$C_i^n(t) := n^{-1}\mathfrak{c}_i \Pi^n(tv_n), i \in [d], \qquad C_{d+1}^n(t) := n^{-1} \sum_{i=d+1}^{n} \mathfrak{c}_i \Pi^n(tv_n). \tag{5.39}$$

For $l \in \mathbb{N}_0^{d+1}$ with $|l| > 1$ we say that an l-merger occurs in Π^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 $(C_i^n)_{i=1}^{d+1}$ has state space

$E_n^d := n^{-1}\{0,\dots,n\}^{d+1} \setminus \{0\}$, initial state $(1,0,\dots,0) \in E_n^d$, absorbing state $(0,0,\dots,n^{-1})$ and evolves according to the following dynamics:

$$\text{a transition} \qquad c \mapsto c - \frac{l - \mathrm{e}_{\|l\| \wedge (d+1)}}{n} \qquad \text{occurs at rate} \quad \lambda_{n|c|,|l|} \prod_{k=1}^{d+1} \binom{nc_k}{l_k}, \tag{5.40}$$

if $c \in E_n^d$ and $l_k \leq c_k$ for all $i \in [d+1]$, where $\|l\| := \sum_{k=1}^{d+1} kl_k$ and $\mathrm{e}_i = (\delta_{ij})_{j=1}^{d+1}$ denotes the ith unit vector in $\mathbb{R}^{d+1}$.

Let $\partial_i = \frac{\partial}{\partial x_i}$ denote the partial derivative with respect to the ith coordinate.

Proposition 5.12. *Fix $d \in \mathbb{N}$. For any sequence (υ_n) such that $\upsilon_n \sim n^{a-1}$ as $n \to \infty$ and $a < 1$ we have convergence*

$$(C_1^n(t), \dots, C_{d+1}^n(t)) \to (c_1(t), \dots, c_{d+1}(t)),$$

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

$$\begin{aligned}\mathcal{G}f(c) := \frac{\Gamma(a+b)}{\Gamma(b)} \sum_{i=1}^{d} \left(-\frac{c_i |c|^{1-a}}{1-a} + \sum_{m=2}^{i} a^{\overline{m-2}} |c|^{2-a-m} \sum_{\substack{l \in \mathbb{N}_0^d \\ |l|=m, \|l\|=i}} \prod_{k=1}^{d} \frac{c_k^{l_k}}{l_k!} \right) \partial_i f(c) \\ + \frac{\Gamma(a+b)}{\Gamma(b)} \left(-\frac{|c|^{2-a}}{2-a} + \sum_{r=1}^{d+1} \sum_{m=2}^{r} a^{\overline{m-2}} \sum_{\substack{l \in \mathbb{N}_0^{d+1} \\ |l|=m, \|l\|=r}} \prod_{k=1}^{d+1} \frac{c_k^{l_k}}{l_k!} \right) \partial_{d+1} f(c),\end{aligned} \tag{5.41}$$

for any $f \in C^\infty([0,1]^{d+1})$.

Proof. For a function $f : \mathbb{R}^{d+1} \to \mathbb{R}$ 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 $x^\kappa := \prod_{i=1}^{d+1} x_i^{\kappa_i}$. Let $\lambda_n(c)$ denote the total rate of $(C_i^n)_{i=1}^{d+1}$ in state $c \in E_n^d$. Using a Taylor expansion we obtain for the generator $\mathcal{G}_n$ of $(C_i^n)_{i=1}^{d+1}$

$$
\begin{aligned}
\mathcal{G}_n f(c) &:= \lambda_n(c) \int (f(c') - f(c)) \mu(c, dc') \\
&= \upsilon_n \lambda_{n|c|} \sum_{\substack{l \in \mathbb{N}_0^{d+1} \\ |l|>1, l \le nc}} \big(f(c - (l - \mathrm{e}_{\|l\| \wedge (d+1)})/n) - f(c)\big) \frac{\lambda_{n|c|,|l|}}{\lambda_{n|c|}} \prod_{k=1}^{d+1} \binom{nc_k}{l_k} \\
&= \upsilon_n \sum_{l \in \mathbb{N}_0^{d+1}, |l|>1} \Bigg(- \sum_{\kappa \in \mathbb{N}_0^{d+1}, |\kappa|=1} \frac{(l - \mathrm{e}_{\|l\| \wedge (d+1)})^\kappa}{n} D^\kappa f(c) \qquad (5.42) \\
&\quad + \sum_{\kappa \in \mathbb{N}_0^{d+1}, |\kappa|=2} \frac{(l - \mathrm{e}_{\|l\| \wedge (d+1)})^\kappa}{n^2 \kappa!} D^\kappa f\left(c - \vartheta_{n,l} \frac{l - \mathrm{e}_{\|l\| \wedge (d+1)}}{n}\right)\Bigg) \qquad (5.43) \\
&\qquad \times \lambda_{n|c|,|l|} \prod_{i=1}^{d+1} \binom{nc_i}{l_i},
\end{aligned}
$$

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

Part 1. Let us first consider summands corresponding to $|\kappa| = 2$ by focusing on

$$
\begin{aligned}
T^{(2)}(n) := \frac{\upsilon_n}{n^2} \sum_{\substack{l \in \mathbb{N}_0^{d+1} \\ |l|>1}} \sum_{\substack{\kappa \in \mathbb{N}_0^{d+1} \\ |\kappa|=2}} \frac{(l - \mathrm{e}_{\|l\| \wedge (d+1)})^\kappa}{\kappa!} D^\kappa f\left(c - \vartheta_{n,l} \frac{l - \mathrm{e}_{\|l\| \wedge (d+1)}}{n}\right) \qquad (5.44) \\
\times \lambda_{n|c|,|l|} \prod_{k=1}^{d+1} \binom{nc_k}{l_k},
\end{aligned}
$$

Evidently, in this case there exist (possibly equal) $i, j \in [d+1]$ with $\kappa = \mathrm{e}_i + \mathrm{e}_j$.

Notice that

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

Hence, for fixed $i \in [d+1]$ and $\kappa = 2\mathrm{e}_i$

$$\begin{aligned} T^2(n) = \upsilon_n \sum_{\substack{l \in \mathbb{N}_0^{d+1} \\ |l|>1}} & \frac{1}{2n^2} (l - \mathrm{e}_{\|l\| \wedge (d+1)})^{2\mathrm{e}_i} D^{2\mathrm{e}_i} f \left(c - \vartheta_{n,l} \frac{l - \mathrm{e}_{\|l\| \wedge (d+1)}}{n} \right) \lambda_{n|c|,|l|} \\ & \times \prod_{k=1}^{d+1} \binom{nc_k}{l_k} \\ & \leq \|f^{(2\mathrm{e}_i)}\|_\infty \frac{\upsilon_n}{n^2} \sum_{\substack{l \in \mathbb{N}_0^{d+1} \\ |l|>1}} l_i^2 \lambda_{n|c|,|l|} \prod_{k=1}^{d+1} \binom{nc_k}{l_k} \\ & = \|f^{(2\mathrm{e}_i)}\|_\infty \frac{\upsilon_n}{(a+b)^{\overline{n|c|-2}} n^2} \sum_{\substack{l \in \mathbb{N}_0^{d+1} \\ |l|>1}} l_i^2 a^{\overline{|l|-2}} b^{\overline{n|c|-|l|}} \prod_{k=1}^{d+1} \binom{nc_k}{l_k} \\ & = \frac{\|f^{(2\mathrm{e}_i)}\|_\infty}{(1-a)(2-a)} \frac{\upsilon_n}{(a+b)^{\overline{n|c|-2}} n^2} \sum_{\substack{l \in \mathbb{N}_0^{d+1} \\ |l|>1}} l_i^2 (a-2)^{\overline{|l|}} b^{\overline{n|c|-|l|}} \prod_{k=1}^{d+1} \binom{nc_k}{l_k}. \end{aligned}$$

Writing $l_i^2 = l_i^{\underline{2}} + l_i$, we find

$$\begin{aligned} & \frac{\upsilon_n}{(a+b)^{\overline{n|c|-2}} n^2} \sum_{\substack{l \in \mathbb{N}_0^{d+1} \\ |l|>1}} l_i^{\underline{2}} (a-2)^{\overline{|l|}} b^{\overline{n|c|-|l|}} \prod_{k=1}^{d+1} \binom{nc_k}{l_k} \\ & = \frac{\upsilon_n}{(a+b)^{\overline{n|c|-2}} n^2} (1-a)(2-a) n|c|(n|c|-1)(a+b)^{\overline{n|c|-2}} \\ & \sim (2-a)(1-a)|c|^2 n^{a-1} \to 0, \end{aligned}$$

as $n \to \infty$, and

$$
\begin{aligned}
&\frac{\upsilon_n}{(a+b)^{\overline{n|c|-2}}n^2} \sum_{\substack{l\in\mathbb{N}_0^{d+1}\\ |l|>1}} l_i (a-2)^{\overline{|l|}} b^{\overline{n|c|-|l|}} \prod_{k=1}^{d+1} \binom{nc_k}{l_k} \\
&= (a-2)\frac{\upsilon_n}{(a+b)^{\overline{n|c|-2}}n^2}\left(n|c|(a+b-1)^{\overline{n|c|-1}} - b^{\overline{n|c|-1}} nc_i\right) \qquad (5.45)\\
&\sim (a-2)\Gamma(a+b)\left(\frac{|c|n^{a-2}}{\Gamma(a+b-1)} - \frac{c_i|c|^{1-a}n^{-1}}{\Gamma(b)}\right) \to 0,
\end{aligned}
$$

as $n \to \infty$, where we applied Lemma 5.7 (with $k = 2\mathrm{e}_i$ in the first case and $k = \mathrm{e}_i$ in the second) and Lemma 5.6.

For fixed $i, j \in [d+1]$ such that $i \neq j$ and $\kappa = \mathrm{e}_i + \mathrm{e}_j$ we obtain

$$
\begin{aligned}
T_n^{(2)} &\le \frac{\upsilon_n}{n^2} \sum_{l\in\mathbb{N}_0^{d+1}, |l|>1} l_i l_j D^{\mathrm{e}_i+\mathrm{e}_j} f\left(c - \vartheta_{n,l}\frac{l-\mathrm{e}_{\|l\|\wedge(d+1)}}{n}\right)\lambda_{n|c|,|l|} \prod_{k=1}^{d+1}\binom{nc_k}{l_k} \\
&\le \frac{\|f^{(\mathrm{e}_i+\mathrm{e}_j)}\|_\infty}{(1-a)(2-a)} \frac{\upsilon_n}{n^2(a+b)^{\overline{n|c|-2}}} \sum_{\substack{l\in\mathbb{N}_0^{d+1}\\ |l|>1}} l_i l_j (a-2)^{\overline{|l|}} b^{\overline{n|c|-|l|}} \prod_{k=1}^{d+1}\binom{nc_k}{l_k} \\
&= \frac{\|f^{(\mathrm{e}_i+\mathrm{e}_j)}\|_\infty}{(1-a)(2-a)} \frac{\upsilon_n}{n^2(a+b)^{\overline{n|c|-2}}} (a-2)(a-1)(n|c|)^{\underline{2}}(a+b)^{\overline{n|c|-2}} \\
&\sim \|f^{(\mathrm{e}_i+\mathrm{e}_j)}\|_\infty |c|^2 n^{a-1} \to 0,
\end{aligned}
$$

as $n \to \infty$, where we applied Lemmata 5.6 and 5.7 (with $k = \mathrm{e}_i + \mathrm{e}_j$). Summarizing, we showed that $T^2(n)$ vanishes as $n \to \infty$.

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

$$
\begin{aligned}
T^{(1)}(n) &\coloneqq -\frac{\upsilon_n}{n} \sum_{\substack{l\in\mathbb{N}_0^{d+1}\\|l|>1}} \sum_{\substack{\kappa\in\mathbb{N}_0^{d+1}\\|\kappa|=1}} (l - \mathrm{e}_{\|l\|\wedge(d+1)})^{\kappa} D^{\kappa} f(c) \lambda_{n|c|,|l|} \prod_{k=1}^{d+1} \binom{nc_k}{l_k} \\
&= -\frac{\upsilon_n}{n(a+b)^{\overline{n|c|-2}}} \sum_{i=1}^{d+1} \partial_i f(c) \\
&\qquad \times \sum_{\substack{l\in\mathbb{N}_0^{d+1}\\|l|>1}} (l_i - 1_{\{i=\|l\|\wedge(d+1)\}}) a^{\overline{|l|-2}} b^{\overline{n|c|-|l|}} \prod_{k=1}^{d+1} \binom{nc_k}{l_k}.
\end{aligned}
$$

Now consider in $T^{(1)}(n)$ the summands corresponding to a fixed $i \in [d+1]$. We partition these summands and analyse their asymptotics seperately as follows. Firstly, by Lemma 5.7 (with $k = \mathrm{e}_i$) we have that

$$
\begin{aligned}
O_i(n) &\coloneqq \frac{\upsilon_n}{n(a+b)^{\overline{n|c|-2}}} \sum_{\substack{l\in\mathbb{N}_0^{d+1}\\|l|>1}} l_i a^{\overline{|l|-2}} b^{\overline{n|c|-|l|}} \prod_{k=1}^{d+1} \binom{nc_k}{l_k} \\
&= \frac{1}{(1-a)(2-a)} \frac{\upsilon_n}{n(a+b)^{\overline{n|c|-2}}} \sum_{\substack{l\in\mathbb{N}_0^{d+1}\\|l|>1}} l_i (a-2)^{\overline{|l|}} b^{\overline{n|c|-|l|}} \prod_{k=1}^{d+1} \binom{nc_k}{l_k} \\
&= \frac{1}{(1-a)(2-a)} \frac{\upsilon_n}{n(a+b)^{\overline{n|c|-2}}} \\
&\qquad \times \left((a-2)nc_i(a+b-1)^{\overline{n|c|-1}} - (a-2)b^{\overline{n|c|-1}} nc_i \right) \\
&= \frac{c_i}{a-1} \frac{\upsilon_n}{(a+b)^{\overline{n|c|-2}}} \left((a+b-1)^{\overline{n|c|-1}} - b^{\overline{n|c|-1}} \right) \\
&\sim \frac{\Gamma(a+b)}{a-1} c_i n^{a-1} \left(\frac{1}{\Gamma(a+b-1)} - \frac{(|c|n)^{1-a}}{\Gamma(b)} \right) \sim -\frac{\Gamma(a+b)}{(a-1)\Gamma(b)} c_i |c|^{1-a},
\end{aligned}
$$

as $n \to \infty$. Secondly, for any fixed $i \in [d]$ the summand corresponding to the indicator $1_{\{i=\|l\|\wedge(d+1)\}}$ has asymptotic behaviour

$$\begin{aligned}
I_i(n) &\coloneqq -\frac{\upsilon_n}{n(a+b)^{\overline{n|c|-2}}} \sum_{\substack{l\in\mathbb{N}_0^{d+1}\\ |l|>1}} 1_{\{i=\|l\|\wedge(d+1)\}} a^{\overline{|l|-2}} b^{\overline{n|c|-|l|}} \prod_{k=1}^{d+1} \binom{nc_k}{l_k} \\
&= -\frac{\upsilon_n}{n(a+b)^{\overline{n|c|-2}}} \sum_{\substack{l\in\mathbb{N}_0^{d+1}\\ |l|>1,\|l\|=i}} a^{\overline{|l|-2}} b^{\overline{n|c|-|l|}} \prod_{k=1}^{d+1} \binom{nc_k}{l_k} \\
&= -\frac{\upsilon_n}{n(a+b)^{\overline{n|c|-2}}} \sum_{m=2}^{n|c|\wedge i} a^{\overline{m-2}} b^{\overline{n|c|-m}} \sum_{\substack{l\in\mathbb{N}_0^{d+1}\\ |l|=m,\|l\|=i}} \prod_{k=1}^{d+1} \binom{nc_k}{l_k} \\
&\sim -\frac{\Gamma(a+b)}{\Gamma(b)} n^{a-2} \sum_{m=2}^{i} a^{\overline{m-2}} |c|^{2-a-m} n^{-a-(m-2)} n^m \sum_{\substack{l\in\mathbb{N}_0^{d+1}\\ |l|=m,\|l\|=i}} \prod_{k=1}^{d+1} \frac{c_k^{l_k}}{l_k!} \\
&= -\frac{\Gamma(a+b)}{\Gamma(b)} \sum_{m=2}^{i} a^{\overline{m-2}} |c|^{2-a-m} \sum_{\substack{l\in\mathbb{N}_0^{d+1}\\ |l|=m,\|l\|=i}} \prod_{k=1}^{d+1} \frac{c_k^{l_k}}{l_k!},
\end{aligned}$$

as $n \to \infty$.

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

$$\begin{aligned}
&\{l \in \mathbb{N}_0^{d+1} \colon |l| > 1, \|l\| \geq d+1\} \\
&= \mathbb{N}_0^{d+1} \setminus \left(\{l \in \mathbb{N}_0^{d+1} \colon |l| \geq 2, \|l\| \in [d]\} \cup \{l \in \mathbb{N}_0^{d+1} \colon |l| \leq 1\} \right),
\end{aligned}$$

it follows that

$$I_{d+1}(n) \coloneqq \frac{\upsilon_n}{(1-a)(2-a)n(a+b)^{\overline{n|c|-2}}} \Bigg(\sum_{l\in\mathbb{N}_0^{d+1}} (a-2)^{\overline{|l|}} b^{\overline{n|c|-l}} \prod_{k=1}^{d+1} \binom{nc_k}{l_k}$$
$$- \sum_{r=1}^{d} \sum_{m=2}^{r} (a-2)^{\overline{m}} b^{\overline{n|c|-m}} \sum_{\substack{l\in\mathbb{N}_0^{d+1} \\ |l|=m, \|l\|=r}} \prod_{k=1}^{d+1} \binom{nc_k}{l_k}$$
$$- \sum_{\substack{l\in\mathbb{N}_0^{d+1} \\ |l|=1}} (a-2)^{\overline{|l|}} b^{\overline{n|c|-|l|}} \prod_{k=1}^{d+1} \binom{nc_k}{l_k} - b^{\overline{n|c|}} \Bigg).$$

We now study each of the summands in $I_{d+1}(n)$. Using Lemma 5.7 (with $k=0$) we find that the first summand

$$\frac{\upsilon_n}{(1-a)(2-a)n(a+b)^{\overline{n|c|-2}}} \sum_{l\in\mathbb{N}_0^{d+1}} (a-2)^{\overline{|l|}} b^{\overline{n|c|-l}} \prod_{k=1}^{d+1} \binom{nc_k}{l_k}$$
$$= \frac{\upsilon_n}{(1-a)(2-a)n(a+b)^{\overline{n|c|-2}}} (a+b-2)^{\overline{n|c|}}$$
$$\sim \frac{\Gamma(a+b)}{(1-a)(2-a)\Gamma(a+b-2)} n^{a-2}$$

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

$$- \frac{\upsilon_n}{(1-a)(2-a)n(a+b)^{\overline{n|c|-2}}} \sum_{r=1}^{d} \sum_{m=2}^{r} (a-2)^{\overline{m}} b^{\overline{n|c|-m}} \sum_{\substack{l\in\mathbb{N}_0^{d+1} \\ |l|=m, \|l\|=r}} \prod_{k=1}^{d+1} \binom{nc_k}{l_k}$$
$$\sim -\frac{\Gamma(a+b)}{\Gamma(b)} \sum_{r=1}^{d} \sum_{m=2}^{r} a^{\overline{m-2}} \sum_{\substack{l\in\mathbb{N}_0^{d+1} \\ |l|=m, \|l\|=r}} \prod_{k=1}^{d+1} \frac{c_k^{l_k}}{l_k!},$$

as $n\to\infty$.

Using

$$\sum_{\substack{l\in\mathbb{N}_0^{d+1}\\|l|=1}}(a-2)^{\overline{|l|}}b^{\overline{n|c|-|l|}}\prod_{k=1}^{d+1}\binom{nc_k}{l_k}=(a-2)b^{\overline{n|c|-1}}n\sum_{m=1}^{d+1}c_m=(a-2)b^{\overline{n|c|-1}}n|c|,$$

as $n\to\infty$ we find for the last summand in $I_{d+1}(n)$

$$\begin{aligned}&-\frac{v_n}{(1-a)(2-a)n(a+b)^{\overline{n|c|-2}}}\left(\sum_{\substack{l\in\mathbb{N}_0^{d+1}\\|l|=1}}(a-2)^{\overline{|l|}}b^{\overline{n|c|-|l|}}\prod_{k=1}^{d+1}\binom{nc_k}{l_k}+b^{\overline{n|c|}}\right)\\&\quad\sim\frac{\Gamma(a+b)}{(2-a)\Gamma(b)}|c|^{2-a}\end{aligned}$$

as $n\to\infty$.

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

$$\lim_{n\to\infty}\sup_{c\in E_n^d}|\mathcal{G}_nf(c)-\mathcal{G}f(c)|=0, \tag{5.46}$$

for all $f\in C_c^\infty([0,1]^{d+1})$.

Since $\mathcal{G}$ 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 5.10, is non-trivial, cf. [73]. 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.

Since $\mathcal{G}_n$ is the generator of a Markov jump process,

$$M_n(t) \coloneqq f((C_i^n(t))_{i=1}^{d+1}) - \int_0^t \mathcal{G}_n f((C_i^n(s))_{i=1}^{d+1}) ds$$

is a martingale for each $f \in B(E^d)$ with compact support. Hence, if some subsequence of $\{(C_i^n)_{i=1}^{d+1}, n \geq 2\}$ converges in distribution to $(c_i)_{i=1}^{d+1}$, then for each $f \in C_c^2(E^d)$

$$f((c_i(t))_{i=1}^{d+1}) - \int_0^t \mathcal{G}_n f((c_i(s))_{i=1}^{d+1}) ds$$

is a martingale by the continuous mapping theorem, cf. [29, Corollary 1.9 of Chapter 3]) and [29, Problem 7 of Chapter 7], since $(c_i^n)_{i=1}^{d+1}$ 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 $\{(f, \mathcal{G}f) \colon f \in C_c^2(E^d)\}$. 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

$$b_i(t,c) \coloneqq b_i(c) \coloneqq$$
$$\frac{\Gamma(a+b)}{\Gamma(b)} \begin{cases} -\frac{c_i|c|}{1-a} + \sum_{m=2}^{i} a^{\overline{m-2}} |c|^{2-a-m} \sum_{\substack{l \in \mathbb{N}^d \\ |l|=m, \|l\|=i}} \prod_{k=1}^{d} \frac{c_k^{l_k}}{l_k!} & \text{if } c \in [0,1]^{d+1}, i \in [d], \\ -\frac{|c|^{2-a}}{2-a} + \sum_{r=1}^{d+1} \sum_{m=2}^{r} a^{\overline{m-2}} \sum_{\substack{l \in \mathbb{N}_0^{d+1} \\ |l|=m, \|l\|=r}} \prod_{k=1}^{d+1} \frac{c_k^{l_k}}{l_k!} & \text{if } c \in [0,1]^{d+1}, i = d+1, \\ 0 & \text{otherwise,} \end{cases}$$

satisfies the conditions of [29, Theorem 3.10 of Chapter 5], then [29, Theorem 2.6 of Chapter 8] implies that the martingale problem for $\{(f, \mathcal{G}f) \colon f \in C_c^2(E)\}$ is well-posed. This is indeed the case, since

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

and moreover, using $c_i \leq |c| \leq 1$

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

and

$$c_{d+1} b_{d+1}(c) \leq |c| \sum_{r=1}^{d+1} \sum_{m=2}^{r} |c|^m K \leq K|c|^3 \leq K|c|^2,$$

where the K, $K_{d,i}$ denote suitable constants that may vary from step to step. It is now straightforward to verify the conditions of Corollary 8.16 of Chapter 4 in [29] which implies the convergence in the statement. □

Let $G := \Gamma(a+b)/(1-a)(2-a)\Gamma(b)$.

Lemma 5.13. *We have*

$$\frac{d}{dt} c_i(t) = \frac{G}{i!} c(t)^{2-a} B_i((a-2)^{\overline{\bullet}} c(t)^{-\bullet}, \bullet! c_\bullet(t)) \quad (i \in \mathbb{N}, t \geq 0).$$

Proof. Proposition 5.12 implies that for each $i \in \mathbb{N}$ $c_i(t)$ is a solution of the ODE

$$c_i'(s) - \frac{\Gamma(a+b)}{(a-1)\Gamma(b)} c_i(s) c(s)^{1-a} = \frac{\Gamma(a+b)}{\Gamma(b)} \sum_{m=2}^{i} a^{\overline{m-2}} c(s)^{2-a-m} \sum_{\substack{l \in \mathbb{N}_0^d \\ |l|=m, \|l\|=i}} \prod_{k=1}^{d} \frac{c_k^{l_k}(s)}{l_k!}, \tag{5.47}$$

or

$$
\begin{aligned}
c_i'(s) &= \frac{\Gamma(a+b)}{(a-1)(a-2)\Gamma(b)} \sum_{l\in\mathbb{N}_0^i, \|l\|=i} (a-2)^{\overline{|l|}} c(s)^{2-a-|l|} \prod_{k=1}^{i} \frac{c_k(s)^{l_k}}{l_k!} \\
&= \frac{G}{i!} \sum_{m=1}^{i} (a-2)^{\overline{m}} c(t)^{2-a-m} \sum_{\substack{l\in\mathbb{N}_0^d \\ |l|=m, \|l\|=i}} \frac{i!}{\prod_{k=1}^{i} l_k!(k!)^{l_k}} \prod_{k=1}^{i} (k! c_k(t))^{l_k} \\
&= \frac{G}{i!} \sum_{m=1}^{i} (a-2)^{\overline{m}} c(t)^{2-a-m} B_{i,m}(w_\bullet) \\
&= \frac{G}{i!} c(t)^{2-a} B_i(v_\bullet, w_\bullet),
\end{aligned}
$$

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

Consider now the generating function

$$
\mathscr{G}(t,x) \coloneqq \sum_{i\geq 1} c_i(t) x^i \quad (x \in [-1,1], t \geq 0). \tag{5.48}
$$

We write ∂_t for the partial derivative $\frac{\partial}{\partial t}$ with respect to t.

Lemma 5.14. *The generating function $\mathscr{G}$ solves the partial differential equation*

$$
\partial_t \mathscr{G}(t,x) = -\frac{\Gamma(a+b)}{(1-a)(2-a)\Gamma(b)} \left((c(t) - \mathscr{G}(t,x))^{2-a} - c(t)^{2-a} \right) \tag{5.49}
$$

for $x \in (-1,1), t \geq 0$ with boundary condition $\mathscr{G}(0,x) = x$ for $x \in [-1,1]$.

Proof. We use the well known fact, cf. [61, Equation (1.11)], that for any two sequences $(v_k), (w_k)$, the exponential generating function of the sequence 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 $v_\bullet = (v_k)$, respectively $w_\bullet = (w_k)$, i.e. $v(\theta) \coloneqq \sum_{k\geq 1} v_k \frac{\theta^k}{k!}$, $w(x) \coloneqq \sum_{k\geq 1} w_k \frac{x^k}{k!}$. For our particular choice of (v_k) and (w_k), we find

$$\begin{aligned} v(\theta) &\coloneqq \sum_{j\geq 1} v_j \frac{\theta^j}{j!} = \sum_{j\geq 1}(a-2)^{\bar{j}}\frac{\theta^j}{j!c(t)^j} = (1-\frac{\theta}{c(t)})^{2-a} - 1, \\ w(x) &\coloneqq \sum_{k\geq 1} w_k \frac{x^k}{k!} = \sum_{k\geq 1} c_k(t)x^k = \mathscr{G}(t,x), \end{aligned}$$

for $|\theta| < c(t)$, $x \in [-1,1]$. Noticing that $|\mathscr{G}(t,x)| < c(t)$ for $|x| < 1$, we use Lemma 5.13 to obtain

$$\begin{aligned} \partial_t \mathscr{G}(t,x) &= \sum_{i\geq 1} c_i'(t)x^i \\ &= Gc(t)^{2-a}\sum_{i\geq 1} B_i((a-2)^{\bar{\bullet}}c(t)^{-\bullet}, \bullet! c_\bullet(t))\frac{x^i}{i!} \\ &= Gc(t)^{2-a}v(\mathscr{G}(t,x)) = G((c(t)-\mathscr{G}(t,x))^{2-a} - c(t)^{2-a}). \end{aligned}$$

$\square$

Theorem 5.15. *The generating function $\mathscr{G}$ is given by*

$$\mathscr{G}(t,x) = c(t) - \left((1-x)^{a-1} + \frac{\Gamma(a+b)}{(2-a)\Gamma(b)}t\right)^{\frac{1}{a-1}} \quad (x \in (0,1), t \geq 0). \tag{5.50}$$

Proof. First, for fixed $x \in (0,1)$ consider the transformation

$$g(t,x) \coloneqq c(t) - \mathscr{G}(t,x) \quad (t \geq 0).$$

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

$$\partial_t g(t,x) = -\frac{\Gamma(a+b)}{(1-a)(2-a)\Gamma(b)} g(t)^{2-a}, \tag{5.51}$$

with boundary condition $g(0,x) = 1-x$. Notice the remarkable similarity between this partial differential equation and the ordinary differnetial equation in (5.38) 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 (5.51) and obtain

$$g(t,x) = \left((1-x)^{a-1} + \frac{\Gamma(a+b)}{(2-a)\Gamma(b)} t\right)^{\frac{1}{a-1}}. \tag{5.52}$$

$\square$

Corollary 5.16. *For the deterministic limit $\{(c_1(t), \ldots, c_{d+1}(t)), t \geq 0\}$ derived in Proposition 5.12 we have*

$$c_i(t) = \frac{c(t)^{2-a}}{i!} B_i\left(\left(\frac{1}{1-a}\right)^{\overline{\bullet}} (-c(t)^{1-a})^{\bullet - 1}, (1-a)^{\overline{\bullet}}\right) \quad (i \in [d], t \geq 0). \tag{5.53}$$

Proof. From the definition (5.48) of $\mathscr{G}$ it is clear that we can compute its coefficient $c_i(t)$ for instance by evaluating its ith partial derivative with respect to x at $x = 0$. To this end, it will prove useful to write $\mathscr{G}$ as a composition, namely

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

where

$$f(x) := \left(x + \frac{\Gamma(a+b)}{(2-a)\Gamma(b)}t\right)^{\frac{1}{a-1}} \qquad \text{and} \qquad g(x) := (1-x)^{a-1}.$$

We can now find a formula for the ith partial derivative of $\mathscr{G}$ by an application of Faà di Bruno's formula, see [41], which states that

$$\frac{d^i}{dx^i}(f \circ g)(x) = \sum_{\pi \in \mathcal{P}_{[i]}} f^{(\#\pi)}(g(x)) \prod_{B \in \pi} g^{(\#B)}(x), \tag{5.54}$$

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

$$f^{(j)}(x) = \left(\frac{1}{a-1}\right)^{\underline{j}} \left(x + \frac{\Gamma(a+b)}{(2-a)\Gamma(b)}t\right)^{\frac{1}{a-1}-j}, \tag{5.55}$$

and

$$g^{(j)}(x) = (-1)^j (a-1)^{\underline{j}} (1-x)^{a-1-j}. \tag{5.56}$$

Plugging this into (5.54) and using (5.32) yields

$$
\begin{aligned}
c_i(t) &= \frac{1}{i!}\partial_x^i\bigg|_{x=0} \mathscr{G}(t,x) \\
&= -\frac{1}{i!}\sum_{\pi\in\mathcal{P}_{[i]}}\left(\frac{1}{a-1}\right)^{\#\pi}\left(1+\frac{\Gamma(a+b)}{(2-a)\Gamma(b)}t\right)^{\frac{1}{a-1}-\#\pi}\prod_{B\in\pi}(-1)^{\#B}(a-1)^{\underline{\#B}} \\
&= -\frac{c(t)}{i!}\sum_{\pi\in\mathcal{P}_{[i]}}(-1)^{\#\pi}\left(\frac{1}{1-a}\right)^{\overline{\#\pi}} c(t)^{\#\pi(1-a)}\prod_{B\in\pi}(1-a)^{\overline{\#B}} \\
&= \frac{c(t)^{2-a}}{i!}\sum_{m=1}^{i}\left(\frac{1}{1-a}\right)^{\overline{m}}(-c(t)^{1-a})^{m-1}B_{i,m}((1-a)^{\overline{\bullet}}) \\
&= \frac{c(t)^{2-a}}{i!}B_i\left(\left(\frac{1}{1-a}\right)^{\overline{\bullet}}(-c(t)^{1-a})^{\bullet-1},(1-a)^{\overline{\bullet}}\right).
\end{aligned}
$$

□

Remark 5.17. Notice that for $t \geq 0$ we have $c(t) = \sum_{i\geq 1} c_i(t)$, since by monotone convergence $\mathscr{G}(t,x) = \sum_{i\geq 1} c_i(t)x^i \to \sum_{i\geq 1} c_i(t) = \mathscr{G}(t,1)$ as $x \to 1$, whereas Theorem 5.15 implies $\mathscr{G}(t,x) \to c(t)$ as $x \to 1$, since

$$
\left((1-x)^{a-1}+\frac{\Gamma(a+b)}{(2-a)\Gamma(b)}t\right)^{\frac{1}{a-1}} \sim 1-x
$$

as $x \to 1$.

In this chapter we have studied the block size spectrum $\{\mathfrak{c}\Pi^n(t), t \geq 0\}$ and its scaling limit as $n \to \infty$ for various coalescent processes. In Section 5.1 we used the transition probabilities for the Kingman and the Bolthausen-Sznitman coalescent to characterize the block size spectrum for finite sample size n. Currently, we cannot extend this method to other coalescents, as their transition probabilities are not (yet) known. Finding the transition probabilities for coalescents other than the Kingman and the Bolthausen-Sznitman coalescent seems to be a challenging problem. One route to this goal might be to try to find the spectral decomposition of the corresponding generator. However, in Section 5.2 we turn to the so-called hydrodynamic limit of the block size spectrum for beta coalescents that come down from infinity. The explicit formula for this limit quantifies how the corresponding coalescent comes down from infinity.

Our results on the hydrodynamic limit raise a number of new questions that one could investigate next.

(i) The hydrodynamic limit for beta coalescents, given in Theorem 5.10 and Proposition 5.12, can be viewed as a law of large numbers type of result. A natural question to ask is about the nature of the fluctuations of the relative block sizes around their means c_i, $i \in \mathbb{N}$.

(ii) An obvious task is to understand the limiting behaviour in the regime $a \geq 1$. Of particular interest is the borderline case $a = 1$ that comprises the Bolthausen-Sznitman coalescent.

(iii) The reason that the hydrodynamic limit has already been studied for Kingman's coalescent is that this coalescent fits into the framework of Smoluchowski's coagulation equations, corresponding to the constant coagulation kernel $K(x, y) = 1$, see [5]. Clearly, multiple merger coalescents do not fit into the framework of Smoluchowski's coagulation equations, since a) they allow for mergers of more than two blocks and b) their rates do depend on the total number of blocks present in the system.

Naively, one would expect that Smoluchowski's coagulation equations could be generalized to comprise both the classical kernels as well as kernels corresponding to multiple merger coalescents. If this is the case, the more general coagulation kernel should have the form $K_b(x_1, \ldots, x_k)$ with $x_1 + \ldots + x_k \leq b$, where b denotes the total mass of current particles, and where k varies between 1 and the number of current particles. A particularly interesting problem would then be the following. In Aldous [5] the author reviews in Section 4 finite-volume stochastic models that respectively correspond to the constant kernel $K(x, y) = 1$, the additive kernel $K(x, y) = x + y$ and the multiplicative kernel $K(x, y) = xy$. For all of these kernels the hydrodynamic limit is known explicitly, see [75]. One could hope for closed-form expressions for the hydrodynamic limit for a generalization a) $K_b(x_1 + \ldots + x_k) = x_1 + \ldots + x_k$ of the additive kernel, and a generalization b) $K_b(x_1, \ldots, x_k) = x_1 \cdots x_k$ of the multiplicative kernel. Moreover, it would be interesting to identify and study corresponding finite-volume stochastic models.

Chapter 6

Conclusion

We briefly recall the lines of investigation in this thesis and outline possible directions for future research.

i) Spectral decompositions. We have derived the spectral decompositions of Π^n and N^n and applications thereof in Chapter 3 for the Kingman and the Bolthausen-Sznitman n-coalescent. Finding spectral decompositions for the beta coalescents seems to be a challenging open problem. However, this spectral decomposition is of interest, as it should pave the way to work out the transition probabilities for beta n-coalescents.

ii) In Chapter 4 we studied the height, the total branch lengths and the SFS of Kingman's n-coalescent tree $\mathcal{T}^n$. Our approach suggests that one should investigate further functionals via cumulants and the partition lattice, e.g. the ordered branch lenghts $(L_n^1, \ldots, L_{n-1}^n)$, where L_n^j measures the total length of branches in $\mathcal{T}^n$ subtending j leaves. In an ongoing project we study the multivariate cumulants of the jump chain J of Kingman's n-coalescent in order to simplify the formula for the cumulants of M_j in Theorem 4.14. The final goal in this direction is to find the joint cumulants of the SFS.

More distant goals would be to use cumulants to investigate functionals of other than Kingman's coalescent, and moreover, to apply the statistical theory developed for cumulants to coalescent processes.

iii) We studied the hydrodynamic limit in Chapter 5. In future work one may want to extend this result to a broader class of coalescents. Obvious candidates would be the regularly varying coalescents, see [12, Definition 4.1].

Two other questions that arise naturally are: a) What is the distribution of the fluctuations around the hydrodynamic limit $c(t)$, respectively $(c_i(t))_{i\geq 1}$, in other words, a central limit type of result. b) What can be said about coalescents that do not come down from infinity, e.g. beta(a, b) coalescents in the regime $a \geq 1$? Does there still exist a rescaling that yields a hydrodynamic limit, possibly in special cases only?

Yet another direction of research, which the author considers quite exciting, is the following. Smoluchowski's coagulation equations model systems of particles (of different masses) that merge as time passes, and are parameterized by a so-called *coagulation kernel* K. The value $K(x, y)$ is interpreted as the rate at which two clusters of particles with masses x and y coagulate, or merge, to form a new cluster of mass $x + y$. It is well known that Kingman's coalescent corresponds to the constant kernel $K(x, y) = 1$. The additive kernel, $K(x, y) = x + y$, corresponds to the additive coalescent, which is important in the context of continuum random trees. The multiplicative kernel, $K(x, y) = xy$, corresponds to the multiplicative coalescent, and is intimately related to the Erdős-Rényi random graph. For more information on these models of random coagulation see, for instance, Aldous' survey [5], [2] and [4].

It seems natural to ask for an extension of Smoluchowski's coagulation equations to allow for multiple mergers, so that they would incorporate multiple merger coalescents. Because of their symmetry, the more general kernels

a) $K(x_1, \dots, x_k) \coloneqq 1$,

b) $K(x_1, \dots, x_k) \coloneqq x_1 + \cdots + x_k$,

c) $K(x_1, \dots, x_k) \coloneqq x_1 \cdots x_k$,

would then lend themselves as starting points for further investigations. A particularly exciting problem would be to find the corresponding finite-volume stochastic models.

Appendix A

Poisson-Dirichlet random partitions and Chinese restaurant process

We briefly recall the notion of a Poisson-Dirichlet random partition and its construction via Kingman's paintbox, where we closely follow Berestycki's lecture notes [12]. For further information on this topic the reader is also referred to Pitman's lecture notes [61].

For any partition $\pi \in \mathcal{P}_{[n]}$ we write $i \sim_\pi j$ if and only if i, j are elements of a common block in π. For any permutation σ of $\mathbb{N}$ define

$$\sigma\pi := \{\{\sigma(b) \colon b \in B\} \colon B \in \pi\},$$

in other words, let $\sigma(i) \sim_{\sigma\pi} \sigma(j)$ if and only if $i \sim_\pi j$. Recall that an exchangeable random partition of $\mathbb{N}$ is a random element $\varSigma$ in $\mathcal{P}_\mathbb{N}$ whose law is invariant under any permutation of finitely many elements, more formally,

$$\Sigma =_d \sigma\Sigma$$

for all permutations σ of $\mathbb{N}$ with finite support. Kingman's correspondence, see, for instance, [12, Corollary 1.1] asserts that the laws of exchangeable random partitions on $\mathbb{N}$ and the laws on

$$\mathcal{S}_0 := \left\{ s = (s_1, s_1, \ldots) \colon s_1 \geq s_2 \geq \cdots, \sum_{i \in \mathbb{N}_0}, s_i = 1 \right\},$$

the space of tilings of the unit interval, are in one-to-one correspondence.

We recall how a random exchangeable partition is constructed from a given tiling $s = (s_i)_{i \geq 0} \in \mathcal{S}_0$. Picture the tilings $J_0 = (0, s_0)$, $J_1 = (s_0, s_0 + s_1)$, etc. as paintboxes containing paint of different colours (we say that J_i contains paint of colour i), the tiling J_0 playing an exceptional rôle as we shall see. We now "colour" the natural numbers $1, 2, \ldots$ consecutively as follows: Let $(U_i)_{i \in \mathbb{N}}$ be a sequence of i.i.d. uniform $[0, 1]$ random variables. If $U_i \in J_k$ falls into the kth paintbox for some $k \geq 1$, we paint the number i with colour k. If $U_i \in J_0$ falls into the 0th paintbox, we paint i with a unique, new colour. The random exchangeable partition Σ of $\mathbb{N}$ is defined by $i \sim_\Sigma j$ if and only if i and j are of the same colour. Instead of a deterministic tiling $s \in \mathcal{S}_0$ one could have a random tiling of $[0, 1]$. This construction of a random exchangeable partition from a tiling of the unit interval is referred to as *Kingman's paintbox construction.*

A family of random tilings that is ubiquitous in the field of exchangeable random partitions is known as the Poisson-Dirichlet distribution. This is a two-parameter family usually denoted $\mathrm{PD}(\alpha, \theta)$. However, for our purposes we restrict ourselves to the case $0 < \alpha < 1$, $\theta = 0$. Let $(W_i)_{i \geq 1}$ be a sequence of independet random variables with $W_i \sim \mathrm{Beta}(1-\alpha, i\alpha)$ having a Beta distribution, see (2.12). Construct a sequence $(\tilde{P}_i)_{i \geq 1}$ of random variables by letting

$$\begin{aligned}
\tilde{P}_1 &\coloneqq W_1, \\
&\vdots \\
\tilde{P}_{n+1} &\coloneqq (1 - \tilde{P}_1 - \cdots - \tilde{P}_n) W_n.
\end{aligned}$$

The non-increasing reordering $(P_i)_{i \geq 1}$ of $(\tilde{P}_i)_{i \in \mathbb{N}}$ is said to have a *Poisson-Dirichlet distribution* with parameters $(\alpha, 0)$, denoted $\mathrm{PD}(\alpha, 0)$. This construction of $\mathrm{PD}(\alpha, 0)$ is sometimes called the stick-breaking construction (for obvious reasons). We say that a random exchangeable partition Σ of $\mathbb{N}$ is a $\mathrm{PD}(\alpha, 0)$-*partition* if it can be constructed from the random tiling $(P_i)_{i \geq 1}$ of $[0, 1]$ via Kingman's paintbox construction.

We now turn to an alternative construction of a $\mathrm{PD}(\alpha, 0)$-partition Σ via the so-called *Chinese restaurant process*. Picture a restaurant with an unlimited number of tables. Customers enter one by one and take seats according to the following rule: the first customer sits at the first table. When customer $n+1$ enters and there are k tables occupied, he joins a table occupied by m customers with probability $(m - \alpha)/n$ each, and takes a seat at a new table with probability $k\alpha/n$. The partition Σ is constructed by letting $i \sim_\Sigma j$ if and only if customers i and j sit at the same table.

Appendix B

Central moments of the Gumbel distribution

For the sake of completeness we record the central moment m'_n of the Gumbel distribution for $0 \le n \le 10$. Based on (4.14) we provide them in terms of the coefficients $s_i(n_1, \ldots, n_i)$, in terms of the zeta function, and numerically. The first central moments of the Gumbel distribution are $m'_0 = 1$, $m'_1 = 0$, $m'_2 = s_1(2) = \zeta(2) = \pi^2/6 \approx 1.64493$, $m'_3 = 2s_1(3) = 2\zeta(3) \approx 2.40411$, $m'_4 = 9s_1(4) + 3s_2(2,2) = 6\zeta(4) + 3\zeta^2(2) = 3\pi^4/20 \approx 14.61136$, $m'_5 = 44s_1(5) + 20s_2(2,3) = 24\zeta(5) + 20\zeta(2)\zeta(3) = 24\zeta(5) + (10/3)\pi^2\zeta(3) \approx 64.43235$,

$$\begin{aligned}
m'_6 &= 265s_1(6) + 135s_2(2,4) + 40s_2(3,3) + 15s_3(2,2,2) \\
&= 120\zeta(6) + 90\zeta(2)\zeta(4) + 40\zeta^2(3) + 15\zeta^3(2) \\
&= \frac{61}{168}\pi^6 + 40\zeta^2(3) \;\approx\; 406.87347,
\end{aligned}$$

$$
\begin{aligned}
m_7' &= 1854 s_1(7) + 924 s_2(2,5) + 630 s_2(3,4) + 210 s_3(2,2,3) \\
&= 720\zeta(7) + 504\zeta(2)\zeta(5) + 420\zeta(3)\zeta(4) + 210\zeta^2(2)\zeta(3) \\
&= 720\zeta(7) + 84\pi^2\zeta(5) + \frac{21}{2}\pi^4\zeta(3) \;\approx\; 2\,815.13142,
\end{aligned}
$$

$$
\begin{aligned}
m_8' &= 14833 s_1(8) + 7420 s_2(2,6) + 4928 s_2(3,5) + 2835 s_2(4,4) \\
&\qquad + 1890 s_3(2,2,4) + 1120 s_3(2,3,3) + 105 s_4(2,2,2,2) \\
&= 5040\zeta(8) + 3360\zeta(2)\zeta(6) + 2688\zeta(3)\zeta(5) + 1260\zeta^2(4) \\
&\qquad + 1260\zeta^2(2)\zeta(4) + 1120\zeta(2)\zeta^2(3) + 105\zeta^4(2) \\
&= \frac{1261}{720}\pi^8 + 2688\zeta(3)\zeta(5) + \frac{560}{3}\pi^2\zeta^2(3) \;\approx\; 22630.60731,
\end{aligned}
$$

$$
\begin{aligned}
m_9' &= 133496 s_1(9) + 66744 s_2(2,7) + 44520 s_2(3,6) + 49896 s_2(4,5) \\
&\qquad + 16632 s_3(2,2,5) + 22680 s_3(2,3,4) + 2240 s_3(3,3,3) + 2520 s_4(2,2,2,3) \\
&= 40320\zeta(9) + 25920\zeta(2)\zeta(7) + 20160\zeta(3)\zeta(6) + 18144\zeta(4)\zeta(5) \\
&\qquad + 9072\zeta^2(2)\zeta(5) + 15120\zeta(2)\zeta(3)\zeta(4) + 2240\zeta^3(3) + 2520\zeta^3(2)\zeta(3) \\
&= 40320\zeta(9) + 4320\pi^2\zeta(7) + \frac{2268}{5}\pi^4\zeta(5) + 2240\zeta^3(3) + 61\pi^6\zeta(3) \\
&\approx 203595.03670, \quad \text{and}
\end{aligned}
$$

$$\begin{aligned}
m'_{10} &= 1334961 s_1(10) + 667485 s_2(2,8) + 444960 s_2(3,7) + 500850 s_2(4,6) \\
&\quad + 243936 s_2(5,5) + 166950 s_3(2,2,6) + 221760 s_3(2,3,5) + 127575 s_3(2,4,4) \\
&\quad + 75600 s_3(3,3,4) + 28350 s_4(2,2,2,4) + 25200 s_4(2,2,3,3) + 945 s_5(2,2,2,2,2) \\
&= 362880\zeta(10) + 226800\zeta(2)\zeta(8) + 172800\zeta(3)\zeta(7) + 151200\zeta(4)\zeta(6) \\
&\quad + 72576\zeta^2(5) + 75600\zeta^2(2)\zeta(6) + 120960\zeta(2)\zeta(3)\zeta(5) + 56700\zeta(2)\zeta^2(4) \\
&\quad + 50400\zeta^2(3)\zeta(4) + 18900\zeta^3(2)\zeta(4) + 25200\zeta^2(2)\zeta^2(3) + 945\zeta^5(2) \\
&= \frac{4977}{352}\pi^{10} + 172800\zeta(3)\zeta(7) + 72576\zeta^2(5) + 20160\pi^2\zeta(3)\zeta(5) + 1260\pi^4\zeta^2(3) \\
&\approx 2036946.09776.
\end{aligned}$$

Bibliography

[1] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, 9th printing, Dover, 1972.

[2] David Aldous, *Brownian excursions, critical random graphs and the multiplicative coalescent*, Ann. Probab. **25** (1997), no. 2, 812–854. MR1434128 (98d:60019)

[3] ———, *Interacting particle systems as stochastic social dynamics*, Bernoulli **19** (2013), no. 4, 1122–1149. MR3102546

[4] David Aldous and Jim Pitman, *The standard additive coalescent*, Ann. Probab. **26** (1998), no. 4, 1703–1726. MR1675063 (2000d:60121)

[5] David J. Aldous, *Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists*, Bernoulli **5** (1999), no. 1, 3–48. MR1673235 (2001c:60153)

[6] George E. Andrews, Richard Askey, and Ranjan Roy, *Special functions*, Encyclopedia of Mathematics and its Applications, vol. 71, Cambridge University Press, Cambridge, 1999. MR1688958 (2000g:33001)

[7] Julien Berestycki, Nathanaël Berestycki, and Vlada Limic, *The Λ-coalescent speed of coming down from infinity*, Ann. Probab. **38** (2010), no. 1, 207–233. MR2599198 (2011d:60224)

[8] ———, *Asymptotic sampling formulae for Λ-coalescents*, Ann. Inst. Henri Poincaré Probab. Stat. **50** (2014), no. 3, 715–731. MR3224287

[9] Julien Berestycki, Nathanaël Berestycki, and Jason Schweinsberg, *Beta-coalescents and continuous stable random trees*, Ann. Probab. **35** (2007), no. 5, 1835–1887. MR2349577 (2009d:60244)

[10] ______, *Small-time behavior of beta coalescents*, Ann. Inst. Henri Poincaré Probab. Stat. **44** (2008), no. 2, 214–238. MR2446321 (2010a:60253)

[11] ______, *The genealogy of branching Brownian motion with absorption*, Ann. Probab. **41** (2013), no. 2, 527–618. MR3077519

[12] Nathanaël Berestycki, *Recent progress in coalescent theory*, Ensaios Matemáticos [Mathematical Surveys], vol. 16, Sociedade Brasileira de Matemática, Rio de Janeiro, 2009. MR2574323 (2011d:60225)

[13] F. Bergeron, G. Labelle, and P. Leroux, *Combinatorial species and tree-like structures*, Encyclopedia of Mathematics and its Applications, vol. 67, Cambridge University Press, Cambridge, 1998. Translated from the 1994 French original by Margaret Readdy, With a foreword by Gian-Carlo Rota. MR1629341 (2000a:05008)

[14] Patrick Billingsley, *Probability and measure*, Third, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, 1995. A Wiley-Interscience Publication. MR1324786 (95k:60001)

[15] Garrett Birkhoff, *Lattice theory*, third, American Mathematical Society Colloquium Publications, vol. 25, American Mathematical Society, Providence, R.I., 1979. MR598630 (82a:06001)

[16] M. Birkner, J. Blath, and B. Eldon, *Statistical properties of the site-frequency spectrum associated with Lambda-coalescents*, ArXiv e-prints (May 2013), available at `1305.6043`.

[17] E. Bolthausen and A.-S. Sznitman, *On Ruelle's probability cascades and an abstract cavity method*, Comm. Math. Phys. **197** (1998), no. 2, 247–276.

[18] George Boros and Victor Moll, *Irresistible integrals*, Cambridge University Press, Cambridge, 2004. Symbolics, analysis and experiments in the evaluation of integrals. MR2070237 (2005b:00001)

[19] David Brillinger, *The calculation of cumulants via conditioning*, Annals of the Institute of Statistical Mathematics **21** (1969), no. 1, 215–218.

[20] C. Cannings, *The latent roots of certain Markov chains arising in genetics: a new approach. I. Haploid models*, Advances in Appl. Probability **6** (1974), 260–290. MR0343949 (49 #8689)

[21] ———, *The latent roots of certain Markov chains arising in genetics: a new approach. II. Further haploid models*, Advances in Appl. Probability **7** (1975), 264–282. MR0371430 (51 #7648)

[22] Louis Comtet, *Advanced combinatorics*, enlarged, D. Reidel Publishing Co., Dordrecht, 1974. The art of finite and infinite expansions. MR0460128 (57 #124)

[23] Iulia Dahmer and Götz Kersting, *The internal branch lengths of the Kingman coalescent*, Ann. Appl. Probab. **25** (2015), no. 3, 1325–1348. MR3325275

[24] P. Donnelly and T. G. Kurtz, *Particle representations for measure-valued population models*, The Annals of Probability **27** (199901), no. 1, 166–205.

[25] Michael Drmota, Alex Iksanov, Martin Moehle, and Uwe Roesler, *Asymptotic results concerning the total branch length of the Bolthausen-Sznitman coalescent*, Stochastic Process. Appl. **117** (2007), no. 10, 1404–1421. MR2353033 (2008j:60178)

[26] Richard Durrett, *Probability models for DNA sequence evolution*, Second, Probability and its Applications (New York), Springer, New York, 2008. MR2439767 (2009k:60001)

[27] Alison Etheridge, *Some mathematical models from population genetics*, Lecture Notes in Mathematics, vol. 2012, Springer, Heidelberg, 2011. Lectures from the 39th Probability Summer School held in Saint-Flour, 2009. MR2759587 (2011j:92043)

[28] Alison M. Etheridge, Robert C. Griffiths, and Jesse E. Taylor, *A coalescent dual process in a moran model with genic selection, and the lambda coalescent limit*, Theoretical Population Biology (2010).

[29] Stewart N. Ethier and Thomas G. Kurtz, *Markov processes – characterization and convergence*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, 1986. MR838085 (88a:60130)

[30] Steven N. Evans, *Probability and real trees*, Lecture Notes in Mathematics, vol. 1920, Springer, Berlin, 2008. Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6–23, 2005. MR2351587 (2009d:60014)

[31] F. Freund and M. Möhle, *On the time back to the most recent common ancestor and the external branch length of the Bolthausen-Sznitman coalescent*, Markov Process. Related Fields **15** (2009), no. 3, 387–416. MR2554368 (2011b:60296)

[32] Y. X. Fu, *Statistical Properties of Segregating Sites*, Theoretical Population Biology **48** (October 1995), no. 2, 172–197.

[33] C. Goldschmidt and J. Martin, *Random recursive trees and the Bolthausen-Sznitman coalesent*, Electron. J. Probab. **10** (2005), no. 21, 718–745.

[34] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, Seventh, Elsevier/Academic Press, Amsterdam, 2007. Translated from the Russian, Translation edited and with a preface by Alan Jeffrey and Daniel Zwillinger, With one CD-ROM (Windows, Macintosh and UNIX). MR2360010 (2008g:00005)

[35] George Grätzer, *Lattice theory: foundation*, Birkhäuser/Springer Basel AG, Basel, 2011. MR2768581 (2012f:06001)

[36] Robert C. Griffiths, *Coalescent lineage distributions*, Adv. in Appl. Probab. **38** (2006), no. 2, 405–429. MR2264950 (2007k:60150)

[37] O. Henard, *The fixation line* (2013). Preprint available at http://www.maths.qmul.ac.uk/ henard/fixation.pdf.

[38] Fredrik Johansson Viklund, Alan Sola, and Amanda Turner, *Small-particle limits in a regularized laplacian random growth model*, Communications in Mathematical Physics **334** (2015), no. 1, 331–366 (English).

[39] Norman L. Johnson, Adrienne W. Kemp, and Samuel Kotz, *Univariate discrete distributions*, Third, Wiley Series in Probability and Statistics, Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, 2005. MR2163227 (2006d:62007)

[40] Norman L. Johnson, Samuel Kotz, and N. Balakrishnan, *Continuous univariate distributions. Vol. 2*, Second, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, John Wiley & Sons Inc., New York, 1995. A Wiley-Interscience Publication. MR1326603 (96j:62029)

[41] Warren P. Johnson, *The curious history of Faà di Bruno's formula*, Amer. Math. Monthly **109** (2002), no. 3, 217–234. MR1903577 (2003d:01019)

[42] Olav Kallenberg, *Foundations of modern probability*, Second, Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR1876169 (2002m:60002)

[43] Götz Kersting, *The asymptotic distribution of the length of beta-coalescent trees*, Ann. Appl. Probab. **22** (2012), no. 5, 2086–2107. MR3025690

[44] J. F. C. Kingman, *The coalescent*, Stochastic Process. Appl. **13** (1982), no. 3, 235–248. MR671034 (84a:60079)

[45] ———, *On the genealogy of large populations*, J. Appl. Probab. **Special Vol. 19A** (1982), 27–43. Essays in statistical science. MR633178 (83d:92043)

[46] J. Kukla and H. H. Pitters, *A spectral decomposition for the Bolthausen-Sznitman coalescent and the Kingman coalescent* (2015). Preprint available at https://sites.google.com/site/pittersh/spectraldecomposition.pdf.

[47] V. Limic and A. Talarczyk-Noble, (2013). Preprint available at http://www.math.u-psud.fr/ limic/izmars/lamfluct.pdf.

[48] ———, *Diffusion limits at small times for coalescents with a Kingman component* (2015). Preprint available at http://www.math.u-psud.fr/ limic/izmars/kingfluct.pdf.

[49] Peter McCullagh, *Tensor methods in statistics*, Monographs on statistics and applied probability, Chapman and Hall, London [u.a.], 1987.

[50] L. Miller and H. H. Pitters, *The hydrodynamic limit of beta coalescents that come down from infinity* (2015). Preprint available at https://sites.google.com/site/pittersh/hydrodynamiclimit.pdf.

[51] M. Möhle, *Der Coalescent*, 1988. Diploma thesis, Johannes Gutenberg University of Mainz.

[52] M. Möhle, *Asymptotic results for coalescent processes without proper frequencies and applications to the two-parameter Poisson-Dirichlet coalescent*, Stochastic Process. Appl. **120** (2010), no. 11, 2159–2173. MR2684740 (2011j:60235)

[53] M. Möhle and H. Pitters, *Absorption time and tree length of the Kingman coalescent and the Gumbel distribution* (2013). Preprint available at http://www.math.uni-tuebingen.de/user/moehle/kingman.pdf.

[54] M. Möhle and S. Sagitov, *A classification of coalescent processes for haploid exchangeable population models*, Ann. Probab. **29** (2001), no. 4, 1547–1562. MR1880231 (2003b:60134)

[55] Martin Möhle, *On hitting probabilities of beta coalescents and absorption times of coalescents that come down from infinity*, ALEA Lat. Am. J. Probab. Math. Stat. **11** (2014), no. 1, 141–159. MR3225970

[56] Martin Möhle and Helmut Pitters, *A spectral decomposition for the block counting process of the Bolthausen-Sznitman coalescent*, Electron. Commun. Probab. **19** (2014), no. 47, 11. MR3246966

[57] M. Möhle, *The Mittag-Leffler process and a scaling limit for the block counting process of the Bolthausen-Sznitman coalescent*, ArXiv e-prints (2014), available at `1410.7354`.

[58] Martin Möhle, *Asymptotic hitting probabilities for the Bolthausen-Sznitman coalescent*, J. Appl. Probab. **51A** (201412), 87–97.

[59] J. R. Norris, *Markov chains*, Cambridge series in statistical and probabilistic mathematics, Cambridge University Press, 1998.

[60] J. Pitman, *Coalescents with multiple collisions*, Ann. Probab. **27** (1999), no. 4, 1870–1902. MR1742892 (2001h:60016)

[61] ———, *Combinatorial stochastic processes*, Springer-Verlag, Berlin, 2006.

[62] Félix Pollaczek, *Problèmes stochastiques posés par le phénomène de formation d'une queue d'attente à un guichet et par des phénomènes apparentés*, Mémor. Sci. Math., no. 136, Gauthier-Villars, Paris, 1957. MR0091555 (19,987c)

[63] Sheldon M. Ross, *Introduction to probability models, ninth edition*, Academic Press, Inc., Orlando, FL, USA, 2006.

[64] S. Sagitov, *The general coalescent with asynchronous mergers of ancestral lines*, J. Appl. Probab. **36** (1999), no. 4, 1116–1125. MR1742154 (2001f:92019)

[65] Jason Schweinsberg, *Coalescents with simultaneous multiple collisions*, Electron. J. Probab. **5** (2000), Paper no. 12, 50 pp. (electronic). MR1781024 (2002g:60113)

[66] Jason Schweinsberg, *A necessary and sufficient condition for the Lambda-coalescent to come down from infinity.*, Electronic Communications in Probability [electronic only] **5** (2000), 1–11 (eng).

[67] T. P. Speed, *Cumulants and partition lattices*, Australian Journal of Statistics **25** (1983), no. 2, 378–388.

[68] ———, *Cumulants and partition lattices. II. Generalised k-statistics*, J. Austral. Math. Soc. Ser. A **40** (1986), no. 1, 34–53. MR809724 (87h:62005)

[69] H. M. Srivastava and Junesang Choi, *Zeta and q-zeta functions and associated series and integrals.*, Amsterdam: Elsevier, 2012 (English).

[70] R. P. Stanley, *Enumerative combinatorics: Volume 1*, 2nd ed., Cambridge University Press, New York, NY, USA, 2011.

[71] Simon Tavaré, *Line-of-descent and genealogical processes, and their applications in population genetics models*, Theoret. Population Biol. **26** (1984), no. 2, 119–164. MR770050 (86f:92017)

[72] ———, *Ancestral inference in population genetics*, Lectures on probability theory and statistics, 2004, pp. 1–188. MR2071630 (2005i:62155)

[73] Marlene Gabriele Ulmet, *Properties of semigroups generated by first order differential operators*, Results Math. **22** (1992), no. 3-4, 821–832. MR1189768 (93i:47073)

[74] G. A. Watterson, *The sampling theory of selectively neutral alleles*, Advances in Appl. Probability **6** (1974), 463–488. MR0351504 (50 #3992)

[75] Jonathan A.D. Wattis, *An introduction to mathematical models of coagulation–fragmentation processes: A discrete deterministic mean-field approach*, Physica D: Nonlinear Phenomena **222** (October 2006), no. 1-2, 1–20 (en).

Acknowledgement

After the first phone call with Alison Etheridge I knew I had to look no further for a supervisor for my PhD. What an enthusiasm in her voice, what a determination to offer me a position, even though we had not met before! I am highly indebted for her precious advise and her boundless support, which she offered in so many respects. I am thankful for the freedom she gave me to make my own mistakes and discoveries, for her encouragement to pursue the daunting adventures and to find my own path in the jungle of research.

Thanks to Jonas Kukla, Luke Miller and Martin Möhle for some very enjoyable research collaborations.

Thanks to Anita Winter for supporting me to undertake a D.Phil. in Oxford. I'd like to thank her and Matthias Birkner for their kind invitations to short visits.

For sharing his knowledge and expertise I'd like to thank Bob Griffiths, and for enjoyable discussions that made me think I thank Nick Barton, Jochen Blath, Julien Berestycki, Ryan Christ, Steve Evans, Stefano Favaro, Yuval Filmus, Fabian Freund, Robert Gaunt, Christina Goldschmidt, Alex Hening, Jerome Kelleher, Tom Kurtz, Cyril Labbé, Charline Smadi Lasserre, Elmar Teufl, Amanda Turner and Anton Wakolbinger.

I had the privilege to regularly visit the probability group in the Mathematics Institute in Tübingen. The financial support of the "Private Stiftung Ewald Marquardt für Wissenschaft und Technik, Kunst und Kultur," and a Teaching Assistantship from the Department of Statistics in Oxford are thankfully acknowledged.

Finally, I owe a great deal to my family and friends, particularly to Annette, Benedikt and Bernd. This thesis is dedicated to Hermine and Hermann Pitters.