Gibbs distributions for random partitions generated by a fragmentation process

Nathanael Berestycki, Jim Pitman

Gibbs distributions for random partitions generated by a fragmentation process

Nathanaël Berestycki and Jim Pitman*

University of British Columbia,
and Department of Statistics, U.C. Berkeley

July 24, 2006

Abstract

In this paper we study random partitions of $\{1, \ldots, n\}$ where every cluster of size j can be in any of w_j possible internal states. The Gibbs (n, k, w) distribution is obtained by sampling uniformly among such partitions with k clusters. We provide conditions on the weight sequence w allowing construction of a partition valued random process where at step k the state has the Gibbs (n, k, w) distribution, so the partition is subject to irreversible fragmentation as time evolves. For a particular one-parameter family of weight sequences w_j, the time-reversed process is the discrete Marcus-Lushnikov coalescent process with affine collision rate $K_{i,j} = a + b(i + j)$ for some real numbers a and b. Under further restrictions on a and b, the fragmentation process can be realized by conditioning a Galton-Watson tree with suitable offspring distribution to have n nodes, and cutting the edges of this tree by random sampling of edges without replacement, to partition the tree into a collection of subtrees. Suitable offspring distributions include the binomial, negative binomial and Poisson distributions.

Keywords Fragmentation processes, Gibbs distributions, Marcus-Lushnikov processes, Gould convolution identities.

*Research supported in part by N.S.F. Grant DMS-0405779

1 Introduction

Gibbs models for random partitions generated by random processes of coagulation and fragmentation have been widely studied ([48], [49], [50], [33]). They typically arise as equilibrium distributions of time-reversible processes of coagulation and fragmentation (see for instance [15], and [5] for general results about exchangeable fragmentation-coalescence processes in equilibrium). There is a much smaller literature in which Gibbs models are derived from an irreversible Markovian coagulation process [28]. This paper presents a Gibbs model for an irreversible Markovian fragmentation process. While Gibbs models for physical processes of fragmentation have been treated before, such models typically allow the possibility of both fragmentation and coagulation at the microscopic level, resulting in a Gibbs equilibrium at any given time. However as time evolves, this equilibrium is moving towards a more fragmented state, so using the language of thermodynamics this equilibrium should be seen as a quasistatic equilibrium of an adiabatic process. The point here is to provide a rigorous Markovian model of irreversible fragmentation at the microscopic level with no possibility of coagulation allowed.

The simplest way to describe the process treated here is to specify its time reversal. This is the Marcus-Lushnikov coalescent process with collision rate kernel $K_{i,j} = a+b(i+j)$ for some constants a and b, where $K_{i,j}$ represents rate of collisions between clusters of i particles and clusters of j particles. This model was solved by Hendriks et al. [28], who showed that the distribution at time t in such a coalescent process started from a monodisperse initial condition is a mixture of microcanonical Gibbs distributions with mixing coefficients depending on t. Here, we derive what turns out to be essentially the same model, modulo time reversal and a formulation in discrete rather than continuous time, but from a different set of assumptions describing the evolution of the process with time running in the direction of fragmentation. The probabilistic link between the two sets of assumptions is a time reversal calculation using Bayes rule. The most interesting feature of this calculation is that starting from a natural recursive assumption for the fragmentation process in terms of Gibbs distributions, there is only one-parameter family of possible solutions to the problem, with the parameter corresponding to the ratio of the two parameters a and b in the collision rate kernel of the reversed time process. Another interesting feature is that for some but not all a and b, the fragmentation process for partitions of a set of size n can be realized by conditioning a Galton-Watson process with suitable offspring distribution to have a family tree of size n, then cutting the edges of this tree by a process of random sampling without replacement. The suitable offspring distributions include the binomial, negative binomial and Poisson distributions.

1.1 Canonical and microcanonical Gibbs distribution

Typically, the state of a coagulation/fragmentation process is represented by a random partition of n, that is a random variable with values in the set $\mathcal{P}_n$ of all partitions of

n. In later sections of this paper the state of the process will be represented rather as a random partition of the set $\{1, 2, \ldots, n\}$, as this device simplifies a number of calculations. But the rest of this introduction follows the more common convention of working with the set $\mathcal{P}_n$ of partitions of the integer n. Let

$$\lambda = 1^{c_1} 2^{c_2} \cdots n^{c_n} \tag{1}$$

denote a typical partition of n. Regarding the state of the system as a partition of n particles into clusters of various sizes, the state λ in (1) indicates that there are c_j clusters of size j for each $1 \leq j \leq n$. Note that $\sum_j j c_j = n$, the total number of particles. The total number of clusters is $k := \sum_j c_j$. The numbers $c_1, c_2, \ldots, c_j \ldots$ may be called numbers of *monomers, dimers, ... j-mers*, or numbers of *singletons, doubletons, ... j-tons*. The Gibbs model most commonly derived from equilibrium considerations is the *canonical Gibbs distribution on partitions of n with weight sequence* (w_j) defined by

$$P(\lambda \mid n; w_1, w_2, \ldots, w_n) = \frac{n!}{Y_n} \prod_{i=1}^{n} \frac{1}{c_i!} \left(\frac{w_i}{i!}\right)^{c_i} \tag{2}$$

where

$$Y_n = Y_n(w_1, w_2, \ldots, w_n) \tag{3}$$

is a normalization constant. This polynomial in the first n weights $w_1, w_2, \ldots, w_n$ is known in the combinatorics literature as the *complete Bell* (or *exponential*) *polynomial* [9]. In the physics literature the Gibbs formula (2) is commonly written in terms of $x_i = w_i/i!$ instead of w_i, and the polynomial

$$Z_n(x_1, x_2, \ldots, w_n) := n! Y_n(1!x_1, 2!x_2, \ldots n!x_n) \tag{4}$$

is called the *canonical partition function*. For textbook treatments of such models, and references to earlier work see [44]. Typically, the canonical Gibbs distribution (2) is derived either from thermodynamic considerations, or from a set of detailed balance equations corresponding to a reversible equilibrium between processes of fragmentation and coagulation. In the latter case the canonical Gibbs distribution is represented as the equilibrium distribution of a time-reversible Markov chain with state space $\mathcal{P}_n$. For related models, see [2], and Vershik [47], who also considers a variation of Gibbs distributions in the context of quantum statistical physics, where he derives asymptotics for the limiting shape of a Gibbs partition associated with the Bose-Einstein statistics by considering a certain variational problem. See also [22] where partitions are subject to a natural additional constraint of consistency corresponding to infinite exchangeability.

Conditioning a canonical Gibbs distribution on the number of clusters k yields a corresponding *microcanonical Gibbs distribution* for each $1 \leq k \leq n$. This distribution assigns to the partition λ displayed in (1) the probability

$$P(\lambda \mid n, k; w_1, w_2, \ldots, w_n) = \frac{n!}{B_{n,k}} \prod_{i=1}^{n} \frac{1}{c_i!} \left(\frac{w_i}{i!}\right)^{c_i} \tag{5}$$

where $B_{n,k} = B_{n,k}(w_1, w_2, \ldots)$ is a *partial Bell (or exponential) polynomial*, and

$$Z_{n,k}(x_1, x_2, \ldots, x_n) := n! B_{n,k}(1!x_1, 2!x_2, \ldots, n!x_n) \tag{6}$$

is known as a *microcanonical partition function*. A great many expressions, representations and recursions for these polynomials $B_{n,k}$ and $Z_{n,k}$ are known [9, 44]. These formulae are useful whenever the weight sequence (w_j) is such that the associated polynomials admit an explicit formula as functions of n and k, or can be suitably approximated (see e.g. [30]). Some of these results are reviewed in [44]. The class of weight sequences (w_j) for which the microcanonical Gibbs model is "solvable", meaning there is an explicit formula for the $B_{n,k}$, is quite large.

In the study of irreversible partition-valued processes, it has been found in several cases (discussed below) that the distribution of the process at time t is a probabilistic mixture over k of microcanonical Gibbs distributions, that is to say a probability distribution of the form

$$P(\lambda) = \sum_{k=1}^{n} q_{n,k} P(\lambda \mid n, k; w_1, w_2, \ldots, w_n) \tag{7}$$

where $q_{n,k}$ represents the probability that λ has k components, so $q_{n,k} \geq 0$ and $\sum_{k=1}^{n} q_{n,k} = 1$, and both $q_{n,k}$ and the weight sequence w_i may be functions of t. We call any distribution of the form (7) a *Gibbs distribution with weights* (w_j), thereby including both canonical and microcanonical Gibbs distributions. For example, Lushnikov [39] showed that the coalescent model with monodisperse initial condition and collision rates $K_{x,y} = xf(y) + yf(x)$ leads to such Gibbs distributions. Hendriks et. al [28] showed that this is also the case for $K_{x,y} = a + b(x + y)$ for constants a and b. See also [44, Section 1.5] for an interpretation of Gibbs distributions with integer weights in terms of composite combinatorial structures.

1.2 Organization and summary of the paper.

The rest of this paper is organized as follows. Section 2 presents some background material, and introduces the formalism of Gibbs distributions over partitions of the set $\{1, 2, \ldots, n\}$. Results for irreversible fragmentation processes are presented in Section 3. Section 4 presents the main result of the paper. This result states that, under an additional set of assumptions (most notably, the *linear selection rule*), it is possible to construct a Gibbs fragmentation process with weight sequence (w_j) if and only if $w_j = \prod_{m=2}^{j}(mc + jb)$ for some constants b and c: in this case it is shown that the time-reversal of the fragmentation process is the discrete Marcus-Lushnikov coalesent with affine coalescent rate: $K_{i,j} = a + b(i + j)$. Section 5 provides some background material on generating functions and branching processes which is needed for the evaluation of a particular Bell polynomial. Section 6 then presents the proofs of the results of Section 4. These results leave out an important case, which is that of the sequence $w_j = (j-1)!$.

In section 7, we approach this problem from the angle of Kingman's coalescent process and the Ewens sampling formula. In particular we construct a continuous analogue of the desired process. However, we show that the existence of this process with discrete time cannot be obtained by taking the discrete-time chain embedded in the continuous process, so that its existence remains an open question.

2 Preliminaries

Let $[n]$ denote the set $\{1, \ldots, n\}$. A *partition of* $[n]$ is an unordered collection of non-empty disjoint subsets of $[n]$ whose union is $[n]$. A generic partition of $[n]$ into k sets (sometimes also referred to as blocks or components of the partition) will be denoted

$$\pi_k = \{A_1, \ldots, A_k\}, \text{ where } 1 \leq k \leq n \tag{8}$$

and where blocks are numbered according e.g. to their least element. Let $\mathcal{P}_{[n]}$ denote the set of all partitions of $[n]$, and let $\mathcal{P}_{[n,k]}$ be the subset of $\mathcal{P}_{[n]}$ comprising all partitions of $[n]$ into k components. Given a sequence of weights $(w_j, j = 1, 2, \cdots)$ define the *microcanonical Gibbs distribution on* $\mathcal{P}_{[n,k]}$ *with weights* $(w_1, w_2, \ldots)$ to be the probability distribution on $\mathcal{P}_{[n,k]}$ which assigns to each partition π_k as in (8) the probability

$$p_{n,k}(\pi_k; w_1, w_2, \ldots w_n) = \frac{1}{B_{n,k}} \prod_{i=1}^{k} w_{\#A_i}, \tag{9}$$

where $\#A_i$ denotes the number of elements of A_i and

$$B_{n,k} := B_{n,k}(w_1, \ldots, w_n) := \sum_{\pi_k \in \mathcal{P}_{[n,k]}} \prod_{i=1}^{k} w_{n_i(\pi_k)}, \tag{10}$$

where for $\pi_k \in \mathcal{P}_{[n,k]}$, the $n_i(\pi_k)$ for $1 \leq i \leq k$ are the sizes of the components of π_k in some arbitrary order. Throughout this paper we will use the notation $p_{n,k}(\,\cdot\,; w_1, \ldots, w_n)$, or simply $p_{n,k}$ if no confusion is possible, for the microcanonical Gibbs distribution (9) determined by the sequence $(w_1, \ldots, w_n)$. Remark that as soon as $k \geq 2$, the microcanonical Gibbs distribution $p_{n,k}$ actually only depends on $(w_1, \ldots, w_{n-1})$. Given a partition π of $[n]$, the *corresponding partition* λ *of* n is $\lambda := 1^{c_1} 2^{c_2} \cdots n^{c_n}$ as in (1) where c_i is the number of components of π of size i. For each vector of non-negative integer counts $(c_1, \ldots, c_n)$ with $\sum_i i c_i = n$ the number of partitions π of $[n]$ corresponding to the partition $1^{c_1} 2^{c_2} \cdots n^{c_n}$ of n is well known to be

$$\frac{n!}{\prod_{i=1}^{n} c_i! (i!)^{c_i}}. \tag{11}$$

The probability distribution on partitions of n induced by the microcanonical Gibbs distribution on $\mathcal{P}_{[n,k]}$ with weights $(w_1, w_2, \ldots)$ therefore identical to the microcanonical

Gibbs distribution on $\mathcal{P}_n$ with the same weights $(w_1, w_2, \ldots)$ as defined in (5), and there is the following standard expression for $B_{n,k}$ [9]:

$$B_{n,k} = n! \sum_{\lambda_k} \prod_{i=1}^{n} \frac{1}{c_i!} \left(\frac{w_i}{i!}\right)^{c_i} \tag{12}$$

where the sum is over all partitions λ_k of n into k components, and $c_i = c_i(\lambda_k)$ is the number of components of λ_k of size i. Thus transferring from Gibbs distributions on partitions of n into k components to Gibbs distributions on partitions of the set $[n]$ into k components is just a matter of keeping track of the universal combinatorial factor (11).

2.1 Combinatorial Interpretation

The following well-known interpretations provide both motivation and intuition for the study of Gibbs distributions and Bell polynomials. Suppose that n particles labelled by elements of the set $[n]$ are partitioned into *clusters* in such a way that each particle belongs to a unique cluster. Formally, the collection of clusters is represented by a partition of $[n]$. Suppose further that each cluster of size j can be in any one of w_j different *internal states* for some sequence of non-negative integers (w_j). Let the *configuration* of the system of n particles be the partition of the set of n particles into clusters, together with the assignment of an internal state to each cluster. For each partition π of $[n]$ with k components of sizes $n_1, \ldots, n_k$, there are $\prod_{i=1}^{k} w_{n_i}$ different configurations with that partition π. So $B_{n,k}(w_1, w_2, \ldots)$ defined by (10) gives *the number of configurations with k clusters*; the Gibbs distribution (9) with weight sequence (w_j) is *the distribution of the random partition of $[n]$ if all configurations with k clusters are equally likely*, and formula (5) describes the corresponding Gibbs distribution on partitions of n induced by the same hypothesis.

Many particular choices of (w_j) have natural interpretations, both combinatorial and physical. In particular, the following four examples have been extensively studied. Many more combinatorial examples are known where Gibbs distributions arise naturally from an assumption of equally likely outcomes on a suitable configuration space. Related problems of enumeration and asymptotic distributions have been extensively studied [30, 37, 21, 24, 45].

2.2 Some important examples

We recall here some natural examples of Gibbs distributions for particular sequences of weights (w_j), and their combinatorial interpretations, which motivate our work in following sections.

Example 1. *Uniform distribution on partitions of $[n]$.* Take $w_j = 1$ for all j. Then a configuration is just a partition of $[n]$, so that $B_{n,k}(1, 1, \ldots)$ is the number of partitions of $[n]$ into k components, known as a *Stirling number of the second kind*. The

microcanonical Gibbs model $p_{n,k}$ corresponds to assuming that all partitions of $[n]$ into k components are equally likely.

Example 2. *Uniform distribution on permutations.* Suppose that the internal state of a cluster C of size j is one of the $(j-1)!$ cyclic permutations of C. Then $w_j = (j-1)!$, and each configuration corresponds to a permutation of $[n]$. Therefore $B_{n,k}(0!, 1!, 2! \ldots)$ is the number of permutations of $[n]$ with k cycles, known as an *unsigned Stirling number of the first kind*. The microcanonical Gibbs distribution $p_{n,k}$ is the distribution on $\mathcal{P}_n$ induced by a permutation uniformly chosen among all permutations with k cycles.

Example 3. *Cutting a rooted random segment* [44]. Suppose that the internal state of a cluster C of size j is one of $j!$ linear orderings of the set C. Identify each cluster as a directed graph in which there is a directed edge from a to b if and only if a is the immediate predecessor of b in the linear ordering. Call such a graph a *rooted segment*. Then $B_{n,k}(1!, 2!, 3! \ldots)$ is the number of directed graphs whose vertices are labelled by

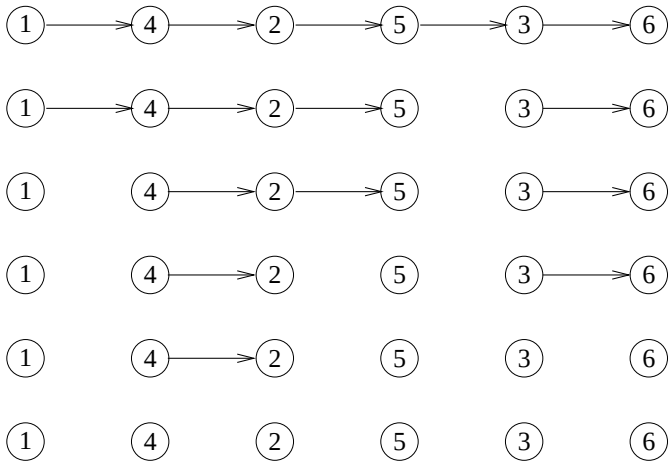

Figure 1: Cutting a rooted random segment.

$[n]$ with k such rooted segments as its components. In the previous two examples, explicit formulae for the $B_{n,k}$ are fairly complicated. But this time there is a simple formula:

$$B_{n,k}(1!, 2!, 3! \ldots) = \binom{n-1}{k-1} \frac{n!}{k!} \tag{13}$$

is known as a *Lah number* [9, p. 135]. The Gibbs model in this instance is a variation of Flory's model for a linear polymerization process [20]. Another interpretation is provided by Kingman's coalescent [1, 34]. It is easily shown in this case that a sequence of random partitions $(\Pi_k, 1 \leq k \leq n)$ such that Π_k has the microcanonical Gibbs distribution with k blocks, can be obtained as follows. Let G_1 be a uniformly distributed random rooted segment labelled by $[n]$, and let G_k be derived from G_1 by deletion of a

set of $k-1$ edges picked uniformly at random from the set of $n-1$ edges of G_1, and let Π_k be the partition induced by the components of G_k. If the $n-1$ edges of G_1 are deleted sequentially, one by one, the random sequence $(\Pi_1, \Pi_2, \ldots, \Pi_n)$ is a refining sequence of random partitions such that Π_k has the Gibbs microcanonical distribution (9). This is illustrated in figure 1. The time-reversed sequence $(\Pi_n, \Pi_{n-1}, \ldots, \Pi_1)$ is then governed by the rules of *Kingman's coalescent*: conditionally given Π_k with k components, Π_{k-1} is equally likely to be any one of the $\binom{k}{2}$ different partitions of $[n]$ obtained by merging two of the components of Π_k. Equivalently, the sequence $(\Pi_1, \Pi_2, \ldots, \Pi_n)$ has uniform distribution over the set $\mathcal{R}_n$ of all refining sequences of partitions of $[n]$ such that the kth term of the sequence has k components. The consequent enumeration $\#\mathcal{R}_n = n!(n-1)!/2^{n-1}$ was obtained by Erdös et al [18]. The fact that Π_k determined by this model has the microcanonical Gibbs distribution with k blocks and weight sequence $w_j = j!$ was obtained by Bayewitz et. al. [4] and Kingman [34].

Example 4. *Cutting a rooted random tree* [44]. Suppose the internal state of a cluster C of size j is one of the j^{j-1} rooted trees labelled by C. Then $B_{n,k}(1^{1-1}, 2^{2-1}, 3^{3-1}, \ldots)$

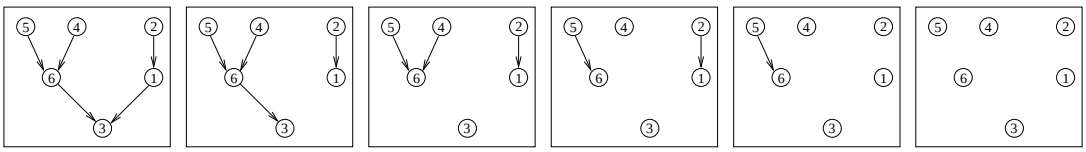

Figure 2: Cutting a rooted random tree with 5 edges

is the number of forests of k rooted trees labelled $[n]$. This time again there is a simple formula for $B_{n,k}$. As a consequence of Cayley's enumeration of forests [42, 43]

$$B_{n,k}(1^{1-1}, 2^{2-1}, 3^{3-1}, \ldots) = \binom{n-1}{k-1} n^{n-k} \tag{14}$$

The Gibbs model in this instance corresponds to assuming that all forests of k rooted trees are equally likely. This model turns up naturally in the theory of random graphs and has been studied and applied in several other contexts. The coalescent obtained by reversing the process of deleting the edges at random is the *additive coalescent* as discussed in [43].

3 Fragmentation Processes

Recall that $\mathcal{P}_{[n]}$ is the set of partitions of $[n] := \{1, \ldots, n\}$. Call a $\mathcal{P}_{[n]}$-valued random process $(\Pi_t, t \in I)$, with index set I a subset of real numbers, a *fragmentation process* if with probability one both

(i) for every pair of times s and t in I with $s < t$ the partition Π_t is a refinement of Π_s, and

(ii) for each $1 \leq k \leq n$ there is some $t \in I$ such that Π_t has k components.

When a confusion is possible we will use the notation $\Pi(t)$ if I is a continuous interval and Π_t in the case where I is a discrete subset of the real numbers. We emphasize that throughout the paper, due to condition (ii), the fragmentation processes we consider are binary. In other words, whenever a split occurs, the split is a binary split in which one and only one block of the partition splits in two, thereby incrementing the number of components by 1. This condition also forces $\Pi(t)$ to be the partition of $[n]$ with one component of size n for all sufficiently small $t \in I$, and to be the partition of $[n]$ into n singletons for all sufficiently large $t \in I$.

Given a sequence of numbers $(w_1, w_2, \ldots, w_{n-1})$, call $(\Pi_t, t \in I)$ a *Gibbs fragmentation process with weights* $(w_1, \ldots, w_{n-1})$ if for every $t \in I$ and $1 \leq k \leq n$, the conditional distribution of Π_t given that Π_t has k components is the microcanonical Gibbs distribution $p_{n,k}$ on $\mathcal{P}_{[n,k]}$ as defined by (9). Note that if $(\Pi_k, k \in [n])$ is a Gibbs fragmentation process, then the unconditional distribution of Π_k is also $p_{n,k}$, because condition (ii) implies that Π_k has k components with probability 1. Finally, the time-reversal $(\Pi_n, \Pi_{n-1}, \ldots, \Pi_1)$ of any fragmentation chain $(\Pi_1, \ldots, \Pi_n)$, is called a *coalescent*.

A basic problem, only partially solved in this paper, is the following:

Problem 1. *For which weight sequences $(w_1, \ldots, w_{n-1})$ does there exist a $\mathcal{P}_{[n]}$-valued Gibbs fragmentation process with these weights?*

The above definitions were made in terms of $\mathcal{P}_{[n]}$-valued processes, as this formalism seems most convenient for computations with Gibbs distributions. Parallel definitions can be made in terms of $\mathcal{P}_n$-valued processes, using the partial ordering of refinement on $\mathcal{P}_n$ defined as follows: for partitions λ and μ of n, λ is a refinement of μ if and only if there exist corresponding partitions λ' and μ' of $[n]$ such that λ' is a refinement of μ'. Less formally, some parts of λ can be coagulated to form the parts of μ. The notions of Gibbs distributions and refining sequences transfer between $\mathcal{P}_n$ and $\mathcal{P}_{[n]}$ in such a way that the following results can be formulated with either state space. The many-to-one correspondence between partitions of the set $[n]$ and partitions of the integer n, quantified by (11), provides a many-to-one correspondence between $\mathcal{P}_n$-valued and $\mathcal{P}_{[n]}$-valued processes, in such a way that the partial ordering of refinement is preserved. Thus for a $\mathcal{P}_{[n]}$-valued Gibbs fragmentation process there is a corresponding $\mathcal{P}_n$-valued fragmentation process, and vice-versa.

To see that Problem 1 is of some interest, note that Example 3 (cutting a random rooted segment) provides a $\mathcal{P}_{[n]}$-valued Gibbs fragmentation process with weights $w_j = j!$ for each n. Example 4 (cutting a random rooted tree) does the same thing for the weights $w_j = j^{j-1}$. What about for the sequence $w_j = 1$ of Example 1 (uniform random partitions) or the sequence $w_j = (j-1)!$ of Example 2 (uniform random permutations)?

In these examples it is not obvious how to construct a Gibbs fragmentation process for $n \geq 4$. (Note that for $n \leq 3$ there exists a $\mathcal{P}_{[n]}$-valued Gibbs fragmentation process for arbitrary positive weights w_1 and w_2, for trivial reasons.) The question for $w_j = 1$ is largely settled by the following proposition, which can be traced as far back as [27] (see also [31]). See Section 7 regarding $w_j = (j-1)!$.

Proposition 1. *There is an $n_0 < \infty$ such that for all $n \geq n_0$ there does not exist a $\mathcal{P}_{[n]}$-valued Gibbs fragmentation process $(\Pi_k, k \in [n])$ with equal weights $w_1 = w_2 = \cdots = w_{n-1}$.*

Proof. Let $\Pi_{[n,k]}$ denote a random partition with the Gibbs distribution on $\mathcal{P}_{[n,k]}$ with equal weights $w_1 = w_2 = \cdots = w_{n-1}$, meaning that $\Pi_{[n,k]}$ has uniform distribution on $\mathcal{P}_{[n,k]}$. Let

$$X_{(n,k,1)} \geq X_{(n,k,2)} \geq \cdots \geq X_{(n,k,k)} \tag{15}$$

denote the sizes of components of $\Pi_{[n,k]}$ arranged in decreasing order. Then for each fixed i and k with $1 \leq i \leq k$ the ith largest component of $\Pi_{[n,k]}$ has relative size $X_{(n,k,i)}/n$ which converges in probability to $1/k$ as $n \to \infty$. This follows easily from the law of large numbers, and the elementary fact that $\Pi_{[n,k]}$ has the same distribution as $\Pi^*_{n,k}$ given that $\Pi^*_{n,k}$ has k components, where $\Pi^*_{n,k}$ is the random partition of $[n]$ generated by n independent random variables $U_1, \ldots, U_n$ each with uniform distribution on $[k]$. (So i and j are in the same component of $\Pi^*_{n,k}$ if and only if $U_i = U_j$.) In particular, there is an $n_0 < \infty$ such that for all $n \geq n_0$ both

$$\mathbb{P}(X_{(n,2,2)} > (5/12)n) > 1/2 \tag{16}$$

and also

$$\mathbb{P}(X_{(n,3,1)} > (5/12)n) < 1/2 \tag{17}$$

But if $(\Pi_{[n,k]}, 1 \leq k \leq n)$ were a fragmentation process, then $\Pi_{[n,3]}$ would be derived from $\Pi_{[n,2]}$ by splitting one of the two components of $\Pi_{[n,2]}$, and hence $X_{(n,2,2)} \leq X_{(n,3,1)}$ with probability one. Thus for a fragmentation process, (16) implies the reverse of the inequality (17), and this contradiction yields the result. □

The above argument proves the non-existence for large n of a Gibbs fragmentation process for any weight sequence (w_j) such that for $k = 2$ or $k = 3$ the components in the Gibbs partition of $[n]$ into k components are approximately equal in size with high probability. For the weight sequence $w_j = j!$ of Example 3, what happens instead is that the sequence of ranked sizes (15), normalized by n, has a non-degenerate limit distribution for each k. As observed in [34, §5], this limit distribution on $[0,1]^k$ is the distribution of the ranked lengths of k subintervals of $[0,1]$ obtained by cutting $[0,1]$ at $k-1$ points picked independently and uniformly at random from $[0,1]$. This asymptotic distribution has been extensively studied [29]. For the weight sequence $w_j = j^{j-1}$ of Example 4, the behavior is different again. What happens is that for each fixed k

the sequence of ranked sizes (15), when normalized by n, converges in probability to $(1, 0, \ldots 0)$. That is to say, for any fixed k, for sufficiently large n, after k steps in the fragmentation process, there is with high probability one big component of relative mass nearly 1, and $k-1$ small components with combined relative mass nearly zero. To be more precise, it is easily shown that the $k-1$ small components, when kept in the order they are broken off the big component, have unnormalized sizes $X_{n,1}, \ldots, X_{n,k-1}$ that are approximately independent for large n with asymptotic distribution

$$\lim_{n\to\infty} \mathbb{P}(X_{n,i} = j) = \frac{j^{j-1}}{j!} e^{-j} \quad (i, j = 1, 2, \ldots) \tag{18}$$

which is the *Borel distribution* of the total progeny of a critical Poisson-Galton-Watson process with Poisson(1) offspring distribution started with one individual, which can be read from (47) and (49). See [43, §4.1] for proofs and various generalizations. As a consequence of (18) and the asymptotic independence of the $X_{n,1}, \ldots, X_{n,k-1}$, the asymptotic distribution of the combined size $X_{n,1} + \ldots + X_{n,k-1}$ of all but the largest component of the partition of $[n]$ into k components is the distribution of the total progeny of the Poisson-Galton-Watson process with Poisson(1) offspring distribution started with k individuals, which is the *Borel-Tanner* distribution [10]

$$\lim_{n\to\infty} \mathbb{P}(X_{n,1} + \ldots + X_{n,k-1} = m) = \frac{k-1}{m} \frac{m^{m-k+1} e^{-m}}{(m-k+1)!} \tag{19}$$

which can also be read from (47) and (49). According to the classification of Barbour and Granovsky [3], the examples corresponding to $w_j = j!$, $w_j = j^{j-1}$ and $w_j = (j-1)!$ belong respectively to the expansive, convergent and logarithmic structures. These structures exhibit quite a different asymptotic behavior, which may account for the differences observed here between these three examples.

As a contrast to Proposition 1, it is known [50, 15] that for any strictly positive sequence of weights (w_j), there is a reversible coagulation-fragmentation process on $\mathcal{P}_{[n]}$ with the canonical Gibbs distribution (2) as its equilibrium distribution.

4 Existence of Gibbs fragmentation processes

The problem of the existence of a Gibbs fragmentation $(\Pi_1, \ldots, \Pi_n)$ for a given integer n and weight sequence $(w_1, \ldots, w_{n-1})$ is one of existence of an increasing process on a partially ordered set with contraints on the marginal distributions of the process. In principle, this is solved by the work of Strassen on measures with given marginals. See for instance [32, Theorem 1 and Proposition 4]. According to this result, for the existence of a Gibbs fragmentation process it is both necessary and sufficient that for all $A \subset \mathcal{P}_{[n,k]}$,

$$\sum_{\pi\in A} p_{n,k}(\pi) \le \sum_{\pi'\in A'} p_{n,k+1}(\pi') \tag{20}$$

where A' is the set of partitions π' that can be obtained by splitting a single block of some partition $\pi \in A$, and $p_{n,k}$ is the Gibbs measure on partitions of $[n]$ into k blocks. A variation of this condition can also be given in terms of integer partitions rather than set partitions. Unfortunately, it seems hard to use this general criterion to prove any existence result. However (20) can be used as an algorithm to check the existence of a Gibbs fragmentation for a finite n and a given sequence $(w_1, \ldots, w_{n-1})$. For instance, when $w_j = 1$ it is possible to check that the first n for which the existence of a Gibbs fragmentation fails is $n = 20$, as mentioned in [27].

If there exists any $\mathcal{P}_{[n]}$-valued Gibbs fragmentation process governed by $(w_1, \ldots, w_{n-1})$, then there exists one that is a Markov chain. For given a non-Markovian process, one can always create a Markov chain with the same one-step transition probabilities and the same marginal distributions. So Problem 1 reduces to:

Problem 2. *For which weight sequences $(w_1, \ldots, w_{n-1})$ does there exist a transition matrix $\{P(\pi, \nu)\}$ indexed by $\mathcal{P}_{[n]}$ such that $P(\pi, \nu) > 0$ only if ν is a refinement of π, and if $\nu \in \mathcal{P}_{[n,k]}$*

$$\sum_{\pi \in \mathcal{P}_{[n]}} p_{n,k-1}(\pi) P(\pi, \nu) = p_{n,k}(\nu) \quad (2 \leq k \leq n) \tag{21}$$

where $p_{n,k}(\nu)$ is given by the microcanonical Gibbs formula (9) ?

Such a transition matrix $P(\pi, \nu)$ corresponds to a *splitting rule* which describes for each $1 \leq k \leq n-1$ and each partition π of $[n]$ into $k-1$ components, the probability that π splits into a partition ν of $[n]$ into k components. Given that $\Pi_{k-1} = \pi_{k-1}$ with $\pi_{k-1} = \{A'_1, \ldots, A'_{k-1}\}$ say, the only possible values π_k of Π_k are those $\pi_k = \{A_1, \ldots, A_k\}$ such that two of the A_j form a partition of one of the A'_i, and the remaining A_j are identical to the remaining A'_i. The initial splitting rule starting with $\pi_1 = \{[n]\}$ is assumed to be specified by the Gibbs distribution $p_{n,2}$ determined by the weight sequence $(w_1, \ldots, w_{n-1})$ for n_1 and n_2 with $n_1 + n_2 = n$. That is from (9):

$$\mathbb{P}(\Pi_2 = \{A_1, A_2\}) = \frac{w_{\#A_1} w_{\#A_2}}{B_{n,2}(w_1, \ldots, w_{n-1})} \tag{22}$$

for $B_{n,2}$ as in (10). The simplest way to continue is to use the following:

Recursive Gibbs Rule: for each $1 \leq k \leq n-1$, given that $\Pi_{k-1} = \{A'_1, \ldots, A'_{k-1}\}$ and that some particular block $A \in \{A'_1, \ldots, A'_{k-1}\}$ is split with $\#A = m \geq 2$, Π_k is obtained by splitting this block is split into $\{A_1, A_2\}$ with probability given by the right side of (22) with m instead of n.

To complete the description of a splitting rule, it is also necessary to specify for each partition $\pi_{k-1} = \{A'_1, \ldots, A'_{k-1}\}$ the probability that the next component to be split is A'_i, for each $1 \leq i \leq k-1$. The simplest possible assumption seems to be the following:

Linear Selection Rule: Given $\pi_{k-1} = \{A'_1, \ldots, A'_{k-1}\}$, split A'_i with probability proportional to $\#A'_i - 1$, that is with probability $(\#A'_i - 1)/(n-k+1)$.

While this selection rule is somewhat arbitrary, it is natural to investigate its implications for the following reasons. Firstly, blocks of size 1 cannot be split, so the probability of picking a block to split must depend on size. This probability must be 0 for a block of size 1, and 1 for a block of size $n-k+1$. The simplest way to achieve this is by linear interpolation. Secondly, both the segment splitting model and the tree splitting model described in Examples 3 and 4 follow this rule. In each of these examples a block of size m is a graph component with $m-1$ edges, so the linear selection rule corresponds to picking an edge uniformly at random from the set of all edges in the random graph whose components define Π_{k-1}. Given two natural combinatorial examples with the same selection rule, it is natural to ask what other models might follow the same rule.

More complex splitting rules are also of interest. Consider for instance a continuous time Markov fragmentation chain which fragments blocks of size j according to infinitesimal rates dictated by $\lambda_j p_{j,2}$ for some $\lambda_j > 0$. The embedded discrete time chain is then a recursive Gibbs fragmentation chain with the property that the probability to select a particular block of size n_i for the next fragmentation is proportional to λ_{n_i}. But we do not know any nice description of the law of Π_k in this case beyond saying that it is the solution of some Kolmogorov forward equations. See also [26] for theory of discrete and continuous time Markov fragmentation chains which are exchangeable and consistent as n varies, meaning that they can be associated with fragmentations of a mass continuum.

4.1 Main results

This section presents the main results, whose proofs are provided in the next two sections. Recall the definition of the discrete Marcus-Lushnikov coalescent process on $\mathcal{P}_{[n]}$ with affine kernel: this is the unique Markov chain on $\mathcal{P}_{[n]}$ such that π_1 is the partition consisting of singletons and π_k is obtained from π_{k-1} by merging each pair of blocks of sizes i and j with probability proportional to $K_{i,j} = a + b(i+j)$ for some constants a and b. In the case $a = 1$ and $b = 0$ this is Kingman's n-coalescent, as described in Example 3 (blocks coalesce at rate 1), while if $a = 0$ and $b = 1$ this is the additive coalescent mentioned in Example 4.

Theorem 2. *Fix $n \geq 4$, and let $(w_j, 1 \leq j \leq n-1)$ be a sequence of positive weights with $w_1 = 1$. The following two statements are equivalent:*

(i) *The $\mathcal{P}_{[n]}$-valued fragmentation process $(\Pi_k, 1 \leq k \leq n)$ defined by the recursive Gibbs splitting rule derived from these weights, with the linear selection rule, is such that for each $1 \leq k < n$ the random partition Π_k has the microcanonical Gibbs distribution $p_{n,k}$ with the same weights.*

(ii) *The weight sequence* w_j *is of the form*

$$w_j = w_j^{b,c} := \prod_{i=2}^{j}(ic + jb) \ , \ (j = 2, \ldots, n-1) \tag{23}$$

for some real b *and* c *such that*

$$b + c > 0 \text{ and either } b \geq 0 \text{ or } b < 0 \text{ and } c > -(n-1)b/2. \tag{24}$$

(iii) *The time reversal of* $(\Pi_k, 2 \leq k \leq n)$ *is a discrete Marcus-Lushnikov coalescent with affine kernel* $K_{i,j} = a + b(i + j)$. *Moreover, in this case the* b *is the same as in* (ii) *and* $a = 2c$.

Note that c and b appearing in (ii) and (iii) are unique only up to a constant common factor.

Given a continuous time $\mathcal{P}_{[n]}$-valued coalescent or fragmentation process $(\Pi(t), t \in I)$, define the *discrete skeleton* of $(\Pi(t), t \in I)$ to be the $\mathcal{P}_{[n]}$-valued process $(\Pi_k^*, 1 \leq k \leq n)$ where Π_k^* is the common value of $\Pi(t)$ for all $t \in I$ such that $\#\Pi(t) = k$. Provided either $b \geq 0$ or $b < 0$ and $c > -nb/2$ the time-reversed process in part (iii) of the above theorem is the discrete time skeleton of the continuous time affine coalescent with collision rate kernel $K_{i,j} := 2c + b(i + j)$. As observed by Hendriks et al. [28], this kernel has the special property that the process $(\#\Pi(t), t \geq 0)$ is independent of the discrete skeleton of $(\Pi(t), t \geq 0)$ (in the special case of Kingman's coalescent this had been proved earlier in [34]). Thus Theorem 2 implies the result of Hendriks et al. [28] that for an affine coalescent in continuous time the distribution of Π_t is a Gibbs distribution with weights $w_j^{b,c}$ as in (23), that is a mixture over k, with mixing weights depending on t, of the microcanonical Gibbs distributions $p_{n,k}^{b,c}$ featured in Theorem 2. The fact that $(\Pi(t), t \geq 0)$ is a Gibbs coalescent with a particular sequence of weights w_j is related in this instance to the fact that its discrete skeleton is a Gibbs coalescent with the same weights. But this equivalence relies on the independence of the process $(\#\Pi(t), t \geq 0)$ and its discrete skeleton. It is not always true that the discrete skeleton of a continuous time Gibbs coalescent is a discrete time Gibbs coalescent, as illustrated by example in Section 7.

The following corollary was suggested by comparison of Theorem 2 with the branching process interpretation of Bell polynomials provided in Section 5.2. To simplify the argument we introduce a regularity condition on the offspring distribution (25), but this assumption may not be strictly necessary for the result to stay valid.

Corollary 3. *Fix* $n \geq 4$. *Let* T *denote a Galton-Watson tree with offspring distribution* (p_j) *such that*

$$p_j > 0 \text{ if and only if } 0 \leq j < j_1 \text{ for some } 1 \leq j_1 \leq \infty. \tag{25}$$

Let F_1 *be* T *conditioned to have* n *nodes, regarded as a random plane tree (a tree with ordered branches), and for* $2 \leq k \leq n$ *let* F_k *be the plane forest of* k *trees obtained by first*

cutting $k-1$ edges of F_1 picked by a process of random sampling without replacement, and then putting these k trees in random order, with all $k!$ orders equally likely, where the cutting and ordering processes are independent of each other and of F_1. Then the following conditions are equivalent:

(i) *The offspring distribution is such that*

$$\frac{p_j}{p_0} = \frac{1}{j!}\prod_{i=1}^{j}(b-(i-2)c) \tag{26}$$

for some real parameters b and c with $b+c>0$ and such that the product is non-negative for all $1 \le j \le n-1$.

(ii) *The forest of two trees F_2 is distributed like two independent copies of T conditioned to have a total of n nodes.*

(iii) *For every $1 \le k \le n$ the forest of k trees F_k is distributed like k independent copies of T conditioned to have a total of n nodes.*

For such a sequence of forests $(F_k, 1 \le k \le n)$ let $(\Pi_k, 1 \le k \le n)$ be the refining sequence of partitions of $[n]$ defined by labelling the n nodes of tree F_1 by a random permutation independent of F_1, and letting the blocks of Π_k be the tree components of F_k. Then

(iv) *the sequence of partitions $(\Pi_1, \ldots, \Pi_n)$ develops by recursive Gibbs fragmentation with linear selection, for the weight sequence $(w_j^{b,c})$ as in (23), and $(\Pi_n, \ldots, \Pi_1)$ is a Marcus-Lushnikov coalescent with the affine kernel $K_{i,j} = 2c + b(i+j)$.*

The implications (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) of this Corollary were provided in [43] for the case $c = 0$, when the offspring distribution can be Poisson with mean b for any $b > 0$. Note that (i) only specifies the conditional offspring distribution given at most $n-1$ children, as is necessary for the converse for a fixed n. The conditions on b and c imposed in (i), which are necessary for construction of the forest-valued fragmentation (F_k), imply but are not implied by the conditions (24) which are necessary for construction of the partition-valued fragmentation (Π_k). To illustrate for $n = 4$, the conditions (24) are that $b+c > 0$ and $3b+2c > 0$, whereas those in (i) above are $b+c > 0$ and either $b-c \ge 0$ or $b = 0$. In either case, $b \ge 0$, hence $3b + 2c > 0$, but not conversely. The b and c such that the conditions (24) hold for all n are those with $b \ge 0$ and $b+c > 0$. Whereas there is the forest-valued representation for all n if and only if one of the following further conditions holds, as discussed later in Section 5.2.

- $c = 0$: the offspring distribution is then Poisson(b);
- $c > 0$ and $b = (a-1)c$ for a positive integer a: the offspring distribution is then binomial(a, p) for $p = c/(c+1)$;

- $-1 < c < 0$ and $b = (a-1)c$ with $-a = r > 0$: the offspring distribution is then negative binomial(r, p) for $p = c + 1$.

The cases when a is an integer admit further combinatorial interpretations, which we will discuss in more detail elsewhere. For instance, when $a = -1$ we obtain a representation of the affine coalescent with collision kernel $K_{i,j} = i + j - 1$ by time reversal of a process of coalescent plane forests $(F_k, 1 \le k \le n)$, where the forest with k trees has the uniform distribution on the set of

$$\frac{k}{n}\binom{2n-k-1}{n-k}$$

plane forests with k trees. And when a is a positive integer, there is an interpretation of the $(a-1)(i+j)+2$ coalescent in terms of trees where each node has either 0 or a children.

5 Preliminaries

5.1 Generating functions

Let $w(z) := \sum_{n=1}^{\infty} w_n z^n/n!$ be the exponential generating function associated with the sequence of weights (w_n). It follows easily from (10) that

$$B_{n,k}(w_1, w_2, \ldots) = \frac{n!}{k!}[z^n]w(z)^k \tag{27}$$

where $[z^n]w(z)^k$ denotes the coefficient of z^n in the expansion of $w(z)^k$ in powers of z. In particular,

$$B_{n,2}(w_1, \ldots, w_{n-1}) = \frac{1}{2}\sum_{l=1}^{n-1}\binom{n}{l} w_l w_{n-l}. \tag{28}$$

Assuming the weights are such that $w(\xi) < \infty$ for some $\xi > 0$, the formula

$$\mathbb{P}(Y = n) = \frac{w_n \xi^n}{n! w(\xi)}$$

defines the distribution of a non-negative random variable Y whose probability generating function is

$$E(z^Y) = w(\xi z)/w(\xi). \tag{29}$$

If $Y_1, Y_2, \ldots$ is a sequence of independent random variables with the same distribution as Y, then

$$\mathbb{P}(Y_1 + \cdots Y_k = n) = [z^n]\left(\frac{w(z\xi)}{w(\xi}\right)^k = \frac{k! B_{n,k}\xi^n}{n! w(\xi)^k}, \tag{30}$$

which appears for instance in (1.3.1) of [36] and [17, Lemma 3.1]. This implies the *Kolchin representation of block sizes in a Gibbs partition* [44, Theorem 1.2]: for a random

partition of $[n]$ with the microcanonical Gibbs distribution $p_{n,k}$ derived from (w_j), when the k blocks are put in a random order, with each of $k!$ possible orders equally likely, independently of the sizes of the blocks, the sequence of block sizes is distributed as

$$(Y_1, \ldots, Y_k) \text{ given } Y_1 + \cdots Y_k = n. \tag{31}$$

Easily from (27) there is the *exponential formula*

$$e^{xw(z)} = \sum_{n=0}^{\infty} C_n(x) z^n \tag{32}$$

where $C_0(x) = 1$ and $C_n(x)$ for $n = 1, 2, \ldots$ is the polynomial

$$C_n(x) = (n!)^{-1} \sum_{k=1}^{n} B_{n,k}(w_1, w_2, \ldots) x^k.$$

The polynomials $C_n(x)$ are then of *convolution type*, meaning that for $n \geq 1$,

$$C_n(x+y) = \sum_{k=0}^{n} C_k(x) C_{n-k}(y). \tag{33}$$

Assuming now that $w_1 = 1$, let $w^{(-1)}$ denote the compositional inverse of w defined by $w^{(-1)}(w(z)) = z$. According to the Lagrange inversion formula [46, Theorem 5.4.2]

$$[z^n] w(z)^k = \frac{k}{n} [z^{n-k}] \left(\frac{z}{w^{(-1)}(z)} \right)^n = \frac{k}{n} \widehat{C}_{n-k}(n) \tag{34}$$

where $\widehat{C}_n(x)$ is the sequence of polynomials of convolution type defined by

$$\left(\frac{z}{w^{(-1)}(z)} \right)^x = \sum_{n=0}^{\infty} \widehat{C}_n(x) z^n. \tag{35}$$

Combining (27) and (34) we obtain the following lemma. See also Knuth [38] for a similar discussion.

Lemma 4. *Each sequence of real weights $(w_1, w_2, \ldots)$ with $w_1 = 1$ admits the representation*

$$w_n = (n-1)! \, \widehat{C}_{n-1}(n) \tag{36}$$

for a unique sequence of polynomials $\widehat{C}_n(x)$ of convolution type, namely that determined by (35)*, in which case for $n \geq 1$,*

$$B_{n,k}(1, w_2, w_3, \ldots) = \frac{(n-1)!}{(k-1)!} \widehat{C}_{n-k}(n). \tag{37}$$

Many sequences of polynomials of convolution type are known [13, Examples 2.2.16], each providing a sequence of weights (w_n) for which the Bell polynomials can be explicitly evaluated using (37). As a general rule, weight sequences with manageable formulas for the $B_{n,k}$ are those with a simple formula for $z/w^{(-1)}(z)$ rather than for $w(z)$. A rich source of such examples is provided by the theory of Galton-Watson branching processes.

5.2 Branching Processes

Given a weight sequence (w_j) with $w_1 = 1$ and exponential generating function $w(z) = \sum_{n\geq 1} w_n z^n/n!$, let

$$G(z) := \frac{z}{w^{(-1)}(z)} = \sum_{n=0}^{\infty} \widehat{C}_n(1) z^n. \tag{38}$$

where $\widehat{C}_0(1) = 1$. Then provided

$$\widehat{C}_n(1) \geq 0 \text{ for all } n \geq 1 \text{ and } G(\eta) < \infty \text{ for some } \eta > 0 \tag{39}$$

the formula

$$g(z) := \frac{G(z\eta)}{G(\eta)} = \sum_{n=0}^{\infty} \frac{\widehat{C}_n(1)\eta^n}{G(\eta)}\, z^n \tag{40}$$

defines the probability generating function of a non-negative integer valued random variable X with distribution

$$\mathbb{P}(X = n) = \frac{\widehat{C}_n(1)\eta^n}{G(\eta)} \qquad (n = 0, 1, 2, \ldots) \tag{41}$$

Conversely, for each distribution of X with $\mathbb{P}(X = 0) > 0$ and each $\eta > 0$ it is easily seen that there is a unique sequence of convolution polynomials $\widehat{C}_n$ such that (41) holds. This is a particular case of [13, Theorem 2.1.14]. Let Y be the total progeny in a Galton-Watson branching process with generic offspring variable X. It is well known that the probability generating function

$$h(z) := \sum_{n=1}^{\infty} \mathbb{P}(Y = n) z^n \tag{42}$$

can be characterized as the unique solution of the functional equation

$$h(z) = zg(h(z)) \tag{43}$$

which is obtained by conditioning on the number of offspring of the root individual [44, Section 6.1]. Note in particular that given the generating function h of the total progeny, the offspring probability generating is determined by

$$g(v) = \frac{v}{h^{(-1)}(v)} \tag{44}$$

Also,

$$h(1) = 1, \text{ meaning } \mathbb{P}(Y < \infty) = 1,$$

if and only if mean of the offspring distribution is at most 1, that is by (40)

$$g'(1) = \eta G'(\eta)/G(\eta) \leq 1. \tag{45}$$

Note that the assumed form (41) of the offspring distribution forces $\mathbb{P}(X = 0) > 0$ (since $\hat{C}_0(1) = 1$), and so forbids the degenerate case with $\mathbb{P}(X = 1) = 1$. Combining this discussion with Lemma 4 we obtain:

Proposition 5. *Let $w_1 = 1, w_2, \ldots$ be a sequence of non-negative weights with exponential generating function $w(z) := \sum_{n=1}^{\infty} w_n z^n/n!$ and let $w^{(-1)}$ be the compositional inverse of w defined by $w^{-1}(w(z)) = z$. The following two conditions are equivalent:*

(i) *there exists $\xi > 0$ such that $w(\xi) < \infty$ and the random variables Y_i in Kolchin's representation* (31) *of Gibbs partitions, with generating function $w(z\xi)/w(\xi)$, are distributed like the total progeny of some Galton-Watson branching process started with one individual.*

(ii) *The power series*

$$G(z) := \frac{z}{w^{(-1)}(z)} = \sum_{n=0}^{\infty} \widehat{C}_n(1) z^n$$

has non-negative coefficients $\widehat{C}_n(1)$ and $G(\eta) < \infty$ for some $\eta > 0$.

When these conditions hold, the offspring distribution is as displayed in (41)*, with generating function $g(z) = G(\eta z)/G(\eta)$ for $\eta = w(\xi)$, and g must satisfy $g'(1) \leq 1$. The associated evaluation of Bell polynomials is then*

$$B_{n,k}(1, w_2, w_3, \ldots) = \frac{(n-1)!}{(k-1)!} \widehat{C}_{n-k}(n) = \frac{n! w(\xi)^k}{k! \xi^n} \mathbb{P}(Y_1 + \cdots + Y_k = n) \qquad (46)$$

where $\widehat{C}_n(x) := [z^n] G(z)^x$ and $Y_1 + \cdots + Y_k$ represents the total progeny of the branching process started with k individuals.

Proof. Condition (i) is that $w(z\xi)/w(\xi) = h(z)$ where h is derived from some probability generating function g via (43). Let h^{-1} denote the compositional inverse of h, defined by $h(h^{-1}(z)) = z$. The equation $h(z) = v$ is solved by $z = w^{(-1)}(vw(\xi))/\xi$, so using (44) g is recovered as

$$g(v) = \frac{v}{h^{-1}(v)} = \frac{v\xi}{w^{(-1)}(vw(\xi))} = \frac{\xi}{w(\xi)} \frac{vw(\xi)}{w^{(-1)}(vw(\xi))} = \frac{G(\eta v)}{G(\eta)}$$

where $\eta = w(\xi)$ so that $G(\eta) = G(w(\xi)) = \frac{w(\xi)}{w^{(-1)}(w(\xi))} = \frac{w(\xi)}{\xi}$. The rest is read from Lemma 4. □

The conditions of the previous proposition force the branching process to be critical or subcritical. For arbitrary η with $G(\eta) < \infty$, and a branching process with offspring generating function $g(z) := G(z\eta)/G(\eta)$, the Lagrange inversion formula shows that the distribution of the total progeny of the branching process started with k individuals is given by the formula

$$\mathbb{P}(Y_1 + \cdots + Y_k = n) = [z]^n h(z)^k = \frac{k}{n} [z^{n-k}] g(z)^n = \frac{\eta^{n-k}}{G(\eta)^n} \frac{k}{n} \widehat{C}_{n-k}(n) \qquad (47)$$

where the Y_i are independent and identically distributed according to this formula for $k = 1$. Formula (47) can be rewritten using Lemma 4 as

$$\mathbb{P}(Y_1 + \cdots Y_k = n) = \frac{\xi^n}{\eta^k} \frac{k!}{n!} B_{n,k}(1, w_2, w_3, \ldots) \tag{48}$$

which is also consistent with (30). Here $\xi := \eta/G(\eta)$, and necessarily $w(\xi) \leq \eta$, with $w(\xi) = \eta$ and $\mathbb{P}(Y_1 + \cdots + Y_k < \infty) = 1$ only in the critical or subcritical case $g'(1) \leq 1$.

To illustrate these results, consider first the generating function $G(z) = e^{bz}$ so that

$$G(z)^x = e^{bzx} = \sum_{n=0}^{\infty} \frac{b^n x^n}{n!} z^n$$

The associated sequence of convolution polynomials is

$$\widehat{C}_n(x) = b^n x^n / n!.$$

The convolution identity (33) is the binomial theorem. The corresponding weight sequence is

$$w_n = (n-1)! \widehat{C}_{n-1}(n) = b^{n-1} n^{n-1}$$

and the Bell polynomial evaluation is

$$B_{n,k} = \frac{(n-1)!}{(k-1)!} \widehat{C}_{n-k}(n) = \binom{n-1}{k-1} b^{n-k} n^{n-k} \tag{49}$$

as indicated earlier in (14). The branching process interpretation is that for Poisson offspring distribution with mean b, the distribution of the total progeny of the branching process started with k individuals is given by formula (47) with the above substitutions for $\eta = 1$ and $\xi = 1/G(1) = e^{-b}$.

Consider next the generating function $G(z) = (1 + cz)^a$ for some pair of real parameters a and c, so that

$$G(z)^x = (1 + cz)^{ax} = \sum_{n=0}^{\infty} \binom{ax}{n} c^n z^n.$$

The associated sequence of convolution polynomials is

$$\widehat{C}_n(x) = \binom{ax}{n} c^n.$$

In this case, the convolution identity (33) is called the Chu-Vandermonde identity (see, e.g., [25]). The corresponding weight sequence is

$$w_n = (n-1)! \widehat{C}_{n-1}(n) = (n-1)! \binom{an}{n-1} c^{n-1} \tag{50}$$

and the Bell polynomial evaluation is

$$B_{n,k} = \frac{(n-1)!}{(k-1)!}\widehat{C}_{n-k}(n) = \frac{(n-1)!}{(k-1)!}\binom{an}{n-k}c^{n-k}. \tag{51}$$

Two cases of this formula have well known probabilistic interpretations [11, 10, 12], as indicated in the next two paragraphs.

If a is a positive integer and $c > 0$, then $\widehat{C}_n(1) \geq 0$ for all n. For $\eta = 1$ the probability generating function (40) is

$$g(z) = \frac{G(z)}{G(1)} = \left(\frac{1+cz}{1+c}\right)^a = (q+pz)^a$$

for $p := c/(1+c)$ and $q := 1-p$. This represents the binomial distribution with parameters a and p. For $a = 1$ the evaluation (51) reduces to the previous evaluation (13) of the Lah numbers. The branching process in this case is a rather trivial one, with each individual having either 0 or 1 offspring. So the random family tree is just a line of vertices whose length is geometrically distributed. Cutting the edges in such a segment of random length by an independent process of Bernoulli trials yields a geometrically distributed number of components, which given their number have independent and identically distributed lengths with another geometric distribution. According to Corollary 3 a similar interpretation of the microcanonical Gibbs distributions with weights (50) can be provided in terms of random cutting of edges of a Galton-Watson tree both in the case of binomial(a, p) offspring distribution for $a = 1, 2, 3, \ldots$, and in the following case of negative binomial offspring distribution.

If $a = -r$ and $c = -q$ for $r > 0$ and $0 < q < 1$, again $\widehat{C}_n(1) \geq 0$ for all n. For $\eta = 1$ the probability generating function (40) is

$$g(z) = \frac{G(z)}{G(1)} = \left(\frac{1+cz}{1+c}\right)^{-r} = \left(\frac{1-q}{1-qz}\right)^r$$

which is the generating function of the negative binomial distribution with parameters $r > 0$ and $p = 1 - q = 1 + c \in (0, 1)$.

It is easily seen that the coefficients $\binom{a}{n}c^n$ are non-negative for all n only in the two cases just discussed. So only in these cases does the Bell polynomial (51) admit the interpretation of Lemma 5 in terms of the total progeny of a branching process for all n. Still, the weights w_n in (50) are non-negative for other choices of real a and c, for instance $a > 1$ and $c > 0$. These weights still define a Gibbs distribution on partitions, and there is the Kolchin representation (31) for the sizes of blocks of such a partition. A natural probabilistic construction of such random partitions is provided by Theorem 2. The interesting intermediate case, when the coefficients $\binom{a}{j}c^j$ are non-negative only for $j < j_1$ for some $j_1 < \infty$, corresponds to Corollary 3. Then $\widehat{C}_n(1)$ can be set equal to 0 for $j \geq j_1$, and the previous branching process formulas remain valid provided n is restricted to $n \leq j_1$.

5.3 Evaluation of a Bell polynomial

The results of the last two Bell polynomial evaluations (49) and (51), which are implicit in the standard theory of branching processes, are unified algebraically by the following lemma. The evaluation (54) is also implicit in [28, (19)-(21)], and plays a key role in our treatment of Gibbs models for fragmentation processes.

Lemma 6. *For each pair of real parameters b and c, the polynomials*

$$\widehat{C}_n^{b,c}(x) := \frac{1}{n!}\prod_{j=0}^{n-1}(bx + cx - cj) \tag{52}$$

are of convolution type. For the corresponding weight sequence

$$w_n^{b,c} := (n-1)!\widehat{C}_{n-1}^{b,c}(n) = \prod_{i=2}^{n}(ic + nb) \tag{53}$$

there is the Bell polynomial evaluation

$$B_{n,k}(1, w_2^{b,c}, w_3^{b,c}, \ldots) = (n-k)!\widehat{C}_{n-k}^{b,c}(n) = \binom{n-1}{k-1}\prod_{i=k+1}^{n}(ic + nb). \tag{54}$$

Proof. This is read from the previous example with generating function $G(z) = (1+cz)^a$ for $a = b/c + 1$. The limiting case $c = 0$ corresponds to $G(z) = e^{zb}$. □

It is convenient to record here as well an immediate consequence of (54):

Lemma 7. *The sequence of weights $w_n = w_n^{b,c}$ is the unique solution of the recursion $w_1 = 1$, $w_2 = 2b + 2c$, and*

$$w_n = \frac{2c + nb}{(n-1)}B_{n,2}(w_1, \ldots, w_{n-1}) \qquad (n = 2, 3, \ldots) \tag{55}$$

for $B_{n,2}$ as in (28).

6 Proofs

6.1 Proof of the main result

The proof of Theorem 2 is based on the next two lemmas.

Lemma 8. *Fix $n \geq 4$ and $3 \leq k \leq n-1$, and let (Π_{k-1}, Π_k) be a pair of random partitions of $[n]$ such that Π_{k-1} is distributed according to the microcanonical Gibbs distribution $p_{n,k-1}$ with weights $w_1 = 1, w_2, \ldots, w_n$, and Π_k is derived from Π_{k-1} by the recursive Gibbs splitting rule with these weights, and the linear selection rule. The following two conditions are equivalent:*

(i) Π_k *has the microcanonical Gibbs distribution* $p_{n,k}$ *with the same weights.*

(ii) *The function*

$$f(m) := \frac{(m-1)w_m}{B_{m,2}(w_1, \ldots, w_{n-1})} \qquad (2 \le m \le n-1) \tag{56}$$

satisfies

$$\sum_{1 \le i < j \le k} f(n_i + n_j) = g(n,k) \tag{57}$$

for all sequences of k *positive integers* $(n_1, \ldots, n_k)$ *with* $\sum_{i=1}^k n_i = n$ *and some function* $g(n,k)$.

When these conditions hold,

$$g(n,k) = \frac{(n-k+1)B_{n,k-1}}{B_{n,k}} \tag{58}$$

and the reverse transition from Π_k *to* Π_{k-1} *is governed by the Marcus-Lushnikov coagulation mechanism with kernel* $K_{i,j} = f(i+j)$. *In the case* $k = 3$ *these conditions are equivalent to*

$$f(m) = 2c + mb \text{ for all } 2 \le m \le n-1 \tag{59}$$

and hence to $w_j = w_j^{b,c}$ *as in* (23), *for some real* b *and* c.

Proof. Let π_k denote any particular partition of $[n]$ into k blocks, say $\{A_1, \ldots, A_k\}$ with $\#A_i = n_i, 1 \le i \le k$. For $1 \le i < j \le n$ let $\pi_{k-1}^{i,j}$ be the partition of $[n]$ into $k-1$ blocks derived from $\{A_1, \ldots, A_k\}$ by merging of A_i and A_j. The hypothesis of the lemma implies that

$$\mathbb{P}(\Pi_{k-1} = \pi_{k-1}^{i,j}, \Pi_k = \pi_k) = \mathbb{P}(\Pi_{k-1} = \pi_{k-1}^{i,j}) \frac{(n_i + n_j - 1)}{(n-k+1)} \frac{w_{n_i} w_{n_j}}{B_{n_i+n_j,2}} \tag{60}$$

and that

$$\mathbb{P}(\Pi_{k-1} = \pi_{k-1}^{i,j}) = \frac{w_{n_i+n_j} \prod_{l=1}^k w_{n_l}}{B_{n,k-1}\, w_{n_i} w_{n_j}} \tag{61}$$

Substituting (61) into (60) gives

$$\mathbb{P}(\Pi_{k-1} = \pi_{k-1}^{i,j}, \Pi_k = \pi_k) = \frac{f(n_i + n_j) \prod_{l=1}^k w_{n_l}}{(n-k+1)B_{n,k-1}} \tag{62}$$

for f derived from the weights as in (56). Summing this probability over all possible choices of (i,j) with $1 \le i < j \le k$ yields $\mathbb{P}(\Pi_k = \pi_k)$, so the equivalence of conditions (i) and (ii) is clear. Assuming these conditions hold, (58) follows at once: dividing (62) by the Gibbs formula for $\mathbb{P}(\Pi_k = \pi_k)$ gives

$$\mathbb{P}(\Pi_{k-1} = \pi_{k-1}^{i,j} | \Pi_k = \pi_k) = f(n_i + n_j)/g(n,k).$$

as claimed. In the case $k = 3$, we deduce (59) from the following lemma, and the weights are then determined by Lemma 7. □

Lemma 9. *Fix $n \geq 3$ and let $(f(m), 2 \leq m \leq n-1)$ be a sequence such that for every triple of positive integers (n_1, n_2, n_3) with $n_1 + n_2 + n_3 = n$*

$$f(n_1 + n_2) + f(n_2 + n_3) + f(n_1 + n_3) = C \tag{63}$$

for some constant C. Then there exist constants b and c such that $f(m) = 2c + mb$ for every $2 \leq m \leq n-1$, and $C = 2(3c + nb)$.

Proof. For $n = 3$ or $n = 4$ the conclusion is trivial, so assume $n \geq 5$. Since $f(m)$ is defined only for $2 \leq m \leq n-1$, it is enough to show that

$$f(l) - f(l-1) = f(l-1) - f(l-2) \text{ for all } 4 \leq l \leq n-1 \tag{64}$$

Let i be the integer part of $l/2$ and $j = l - i$. Then $i \geq 2$ and either $j = i$ or $j = i + 1$, so $j \geq 2$ too. Write $EQ(n_1, n_2, n_3)$ for the equation (63) determined by a particular choice of (n_1, n_2, n_3). Keeping in mind that $l = i + j$, we have

$$EQ(i-1, j-1, n-l+2): \quad f(l-2) + f(n-i+1) + f(n-j+1) = C \tag{65}$$

$$EQ(i-1, j, n-l+1): \quad f(l-1) + f(n-i+1) + f(n-j) = C \tag{66}$$

$$EQ(i, j-1, n-l+1): \quad f(l-1) + f(n-i) + f(n-j+1) = C \tag{67}$$

$$EQ(i, j, n-l): \quad f(l) + f(n-i) + f(n-j) = C \tag{68}$$

Subtract (65) from (66) to obtain

$$f(l-1) - f(l-2) = f(n-j+1) - f(n-j) \tag{69}$$

and subtract (67) from (68) to obtain

$$f(l) - f(l-1) = f(n-j+1) - f(n-j) \tag{70}$$

and Lemma 9 follows. □

We can now finish the proof of Theorem 2. Fix $n \geq 4$, let $(w_j, 1 \leq j \leq n-1)$ be a sequence of positive weights with $w_1 = 1$.

Suppose first as in condition (i) of Theorem 2 that $(\Pi_k, 1 \leq k \leq n)$ is a $\mathcal{P}_{[n]}$-valued fragmentation process defined by the recursive Gibbs splitting rule derived from these weights, with the linear selection rule, and that the distribution of Π_k is $p_{n,k}$ for every k. Then condition (ii) of Lemma 8 holds for all $3 \leq k \leq n-1$, and in particular for $k = 3$. Lemma 9 now forces (59) for some b and c, hence $w_j = w_j^{b,c}$ by Lemma 7.

Conversely, suppose that $(\Pi_k, 1 \le k \le n)$ is a $\mathcal{P}_{[n]}$-valued fragmentation process defined by the recursive Gibbs splitting rule with the weights $w_j = w_j^{b,c}$, and the linear selection rule. Lemma 7 implies that (59) holds, so it is clear that condition (ii) of Lemma 8 holds for $3 \le k \le n-1$ with

$$g(n,k) = (k-1)(kc + nb). \tag{71}$$

Consider the inductive hypothesis that Π_{k-1} has the microcanonical Gibbs distribution $p^{b,c}_{n,k-1}$ with these weights $(w_j^{b,c})$. This is true for $k = 3$ by assumption. Assuming it true for some k, Lemma 8 provides the inductive step from k to $k+1$. Thus the distribution of Π_k is $p^{b,c}_{n,k}$ for every $2 \le k \le n-1$. Thus condition (i) of Theorem 2 is satisfied by the weights $w_j = w_j^{b,c}$.

Condition (iii) of Theorem 2, that the reversed process is an affine coalescent, is now read from the last sentence of Lemma 8.

6.2 Proof of Corollary 3

Recall first that the distribution of an unconditioned Galton-Watson tree, restricted to finite trees, is given by the formula

$$\mathbb{P}(T = t) = \pi(t) := \prod_{v \in V(t)} p_{n(v,t)} \tag{72}$$

where

- t denotes a generic plane tree with a finite number of nodes $\#t$;
- $V(t)$ is the set of nodes of t;
- $n(v,t)$ is the number of children of node v of t;
- p_n is the probability that a node has n children;

The nodes of t are regarded as unlabelled. But the tree has a root node, and the children of each node are assigned a total order, say from left to right. So the nodes of t can be identified or listed by some arbitrary convention, such as depth first search, and any such convention can be used to rigorously identify the set of nodes $V(t)$ as a subset of some ambient countable set. See [42, 44] for background. Fix $n \ge 4$. By definition, F_1 is T conditioned on $\#T = n$, so

$$\mathbb{P}(F_1 = t) = \pi(t)\mathbf{1}_{\{\#t=n\}}/q(n) \tag{73}$$

where $q(n)$ is by definition the probability that T has n nodes:

$$q(n) := \mathbb{P}(\#T = n) = \sum_t \pi(t)\mathbf{1}_{\{\#t=n\}} = p_0^n \frac{w_n}{n!}. \tag{74}$$

In the last formula, read from (48), the weight sequence (w_n) with $w_1 = 1$ is determined as in (44) by its exponential generating function $\sum_n w_n z^n/n!$ which is the compositional inverse of $zp_0/g(z)$ for $g(z) := \sum_{n=0}^{\infty} p_j z^j$ the offspring generating function. The probability (74) is strictly positive for every $n \geq 1$, by the simplifying assumption (25) on the offspring distribution.

Let $\widehat{F}_2$ be the plane forest of two trees obtained by splitting F_1 by deletion of a uniformly chosen random edge of F_1, with subtree containing the root put to the left, and the remaining fringe subtree put to the right. Then the distribution of $\widehat{F}_2$ is given by the following formula: for a generic pair of plane trees (t_1, t_2)

$$\mathbb{P}(\widehat{F}_2 = (t_1, t_2)) = \frac{\pi(t_1)\pi(t_2)\Sigma(t_1)}{q(n)(n-1)} \mathbf{1}_{\{\#t_1 + \#t_2 = n\}} \tag{75}$$

where

$$\Sigma(t) := \sum_{v \in V(t)} r_{n(v,t)} \tag{76}$$

with

$$r_m := (m+1)p_{m+1}/p_m \qquad (0 \leq m \leq n-2)$$

and the particular offspring distribution display in (26) is characterized by the formula

$$r_m = b - (m-1)c \qquad (0 \leq m \leq n-2). \tag{77}$$

Formula (75) is obtained by conditioning on which vertex v of t_1 is the one to which t_2 is attached in t, and given that v has $m+1$ children in t, which of these $m+1$ children is the root of t_1. Tossing a fair coin to decide the order of trees in $\widehat{F}_2$ then yields F_2 with distribution

$$\mathbb{P}(F_2 = (t_1, t_2)) = \frac{\pi(t_1)\pi(t_2)(\Sigma(t_1) + \Sigma(t_2))}{2q(n)(n-1)} \mathbf{1}_{\{\#t_1 + \#t_2 = n\}}. \tag{78}$$

On the other hand, the distribution of F_2^* defined by two independent copies of T conditioned to have a total of n nodes is given by

$$\mathbb{P}(F_2^* = (t_1, t_2)) = \frac{\pi(t_1)\pi(t_2)}{q_2(n)} \mathbf{1}_{\{\#t_1 + \#t_2 = n\}} \tag{79}$$

where

$$q_2(n) = \sum_{m=1}^{n-1} q(m)q(n-m) = p_0^n \frac{2B_{n,2}}{n!}$$

gives the distribution of the total progeny of the branching process started with two individuals, as indicated in (48), with $B_{n,2} = B_{n,2}(1, w_2, \ldots, w_{n-1})$. Condition (ii) of Corollary 3, is the equality in distribution

$$F_2 \stackrel{d}{=} F_2^*. \tag{80}$$

It is clear from (78) and (79) that this equality in distribution is equivalent to the identity

$$\Sigma(t_1) + \Sigma(t_2) = \frac{(n-1)w_n}{B_{n,2}} \tag{81}$$

for all pairs of trees (t_1, t_2) with $\pi(t_1)\pi(t_2)1(\#t_1 + \#t_2 = n) > 0$, where $\Sigma(t) := \sum_{v \in V(t)} r_{n(v,t)}$ for $r_m := (m+1)p_{m+1}/p_m$.
(i) $\Rightarrow$ (ii). If (i) holds then $r_m = b + c - mc$ and hence

$$\Sigma(t_1) + \Sigma(t_2) = nb + nc - (n-2)c = nb + 2c \tag{82}$$

because in every forest of two trees with n nodes the sum of the numbers of children of all nodes is the total number of edges, which is $n-2$, and (81) is now read from (55).
(ii) $\Rightarrow$ (iii) and (iv). Assuming (ii), it follows immediately from (78), (79) and (80) that for $k = 2$ the distribution of Π_k generated by random labelling of tree components of F_k has the Gibbs distribution $p_{n,k}$ with whatever weight sequence $(w_j^{b,c})$ is associated with the distribution of the total progeny of the branching process, and that conditionally given Π_k the k plane trees associated with these components are distributed like independent copies of T conditioned to have the sizes dictated by the block sizes of Π_k. Suppose inductively that this is so for some $k \geq 2$. The process of random edge deletion induces the linear selection rule for components to split, and given that a tree component is split, the inductive hypothesis and the assumption for $k = 2$ implies that the component is split into two independent copies of T conditioned to have the right size. The implication (iii) $\Rightarrow$ (i) of Theorem 2 now provides the inductive step.
(iii) $\Rightarrow$ (ii) is trivial.
(ii) $\Rightarrow$ (i). This follows easily from the identity (81) and the following Lemma:

Lemma 10. *Fix $n \geq 3$. Let $r(m)$ be a real-valued function with domain $S = \{0, 1, \ldots j\}$ for some $1 \leq j \leq n-2$, such that*

$$\sum_{i=1}^{n} r(n_i) = C$$

for some constant C for each choice of $(n_1, \ldots, n_n)$ with $n_i \in S$ for all $1 \leq i \leq n$ and $\sum_{i=1}^{n} n_i = n-2$. Then there exist real a and b such that $r(m) = am + b$ for all $m \in S$.

Proof. Consider for each $1 \leq m \leq j-1$ the sequence $(n_1, \ldots, n_n)$ with the first $n-m-2$ terms equal to 1, the next term equal to m, and the last $m+1$ terms equal to 0. This sequence gives

$$(n-m-2)r(1) + r(m) + (m+1)r(0) = C$$

and the same holds for $m+1$ instead of m. The difference of these two identities gives

$$r(m+1) - r(m) = r(1) - r(0)$$

and the conclusion follows. □

7 Gibbs fragmentations for random permutations in continuous time

Given a symmetric non-negative *collision rate function* $K_{i,j}$ defined for positive integers i and j, call the $\mathcal{P}_{[n]}$-valued continuous time parameter Markovian coalescent process $(\Pi_t, t \geq 0)$, in which each pair of clusters of sizes i and j is merging at rate $K_{i,j}$, the *Marcus-Lushnikov coalescent with collision kernel* $K_{i,j}$. See [1] for background. It is assumed throughout this section, in keeping with the definition of a coalescent process given in the previous section, that such a coalescent process is started with the *monodisperse initial condition*. That is to say Π_0 is the partition of $[n]$ into n singletons. Both Marcus and Lushnikov worked with the corresponding $\mathcal{P}_n$-valued process rather than a $\mathcal{P}_{[n]}$-valued process, but there is no difficulty in translating results from one state-space to the other, by application of the standard criterion for a function of a Markov process to be Markov. Lushnikov [39] found the remarkable result that for a collision kernel of the form $K_{i,j} = if(j) + jf(i)$ for each $t > 0$ the distribution of Π_t is of the form

$$\mathbb{P}(\Pi_t = \pi) = \sum_{k=1}^{n} q_{n,k}(t) p_{n,k}(\pi; w_j(t), j = 1, 2, \ldots) \tag{83}$$

where $p_{n,k}(\pi; w_j, j = 1, 2, \ldots)$ denotes the microcanonical Gibbs distribution on $\mathcal{P}_{[n,k]}$ with weights w_j, and the functions $q_{n,k}(t) = \mathbb{P}(\#\Pi_t = k)$ and the weights $w_j(t)$ are determined by a system of differential equations. As mentioned earlier, Hendriks et al. [28] showed that for $K_{i,j} = a + b(i + j)$ for constants a and b the $w_j(t)$ can be chosen independently of t as $w_j(t) = w_j$ where w_j is determined by a and b via formula (23) for $c = a/2$.

Kingman [34] studied the particular case of the Marcus-Lushnikov coalescent with $a = 1$ and $b = 0$. In this process, at any given time t, given that $\#\Pi_t = k$, each of the $k(k-1)/2$ cluster pairs in existence at time t is merging at rate 1. Call this process with state space $\mathcal{P}_{[n]}$ *Kingman's n-coalescent*, Motivated by applications to genetics, Kingman [34] proposed the following construction. Given a coalescent process $(\Pi_t, t \geq 0)$, suppose that each cluster of Π_t is subject to mutation at rate $\theta/2$ for some $\theta > 0$. Now define a random partition $\tilde{\Pi}_\theta$ of $[n]$ by declaring that i and j are in the same block of $\tilde{\Pi}_\theta$ if and only if no mutation affects the clusters containing i and j in the interval $(0, \tau_{ij})$ where τ_{ij} is the *collision time* of i and j in the coalescent process $(\Pi_t, t \geq 0)$, that is the first time t that i and j are in the same cluster of Π_t. Kingman obtained the following result:

Proposition 11. (Kingman [35]) *Suppose that $(\Pi_t, t \geq 0)$ is Kingman's n-coalescent. Then*

$$\mathbb{P}(\tilde{\Pi}_\theta = \pi) = \frac{\theta^{k-1}}{[\theta + 1]_{n-1}} \prod_{i=1}^{k} (n_i - 1)! \tag{84}$$

for each partition π of $[n]$ into k components of sizes $n_1, \ldots, n_k$

The distribution of $\tilde{\Pi}_\theta$ defined by (84) first appears in [19] and is known as Ewens' Sampling Formula with parameter θ. This distribution has long been recognized as an essential tool in population genetics (see, e.g. [16] recently), and has been applied in a wide variety of contexts in probability. Note that this distribution is a particular mixture over k, with mixing coefficients depending on θ, of the microcanonical Gibbs distributions on $\mathcal{P}_{[n,k]}$ with weights $(j-1)!$, as interpreted in Example 2. Recall that this distribution has been constructed starting from $(\Pi_t, t \geq 0)$, where due to Example 3, for all t, Π_t is a mixture over k, with mixing coefficients depending on t, of the microcanonical Gibbs distributions on $\mathcal{P}_{[n,k]}, 1 \leq k \leq n$, with the different weight sequence $(j!, j \geq 1)$. It does not seem obvious from a combinatorial perspective why there should be such a connection between the Gibbs models with these two weight sequences.

The random partition $\tilde{\Pi}_\theta$ (which is sometimes referred to as the random allelic partition), and more generally the fragmentation process $(\tilde{\Pi}_\theta, \theta \geq 0)$ discussed below, can be defined starting from any coalescent (Π_t), but there seems to be a manageable formula for the distribution of $\tilde{\Pi}_\theta$ only for Kingman's coalescent. See however the recent work of Möhle [41] where an explicit recursion is given for the random allelic partition obtained from a Λ-coalescent (i.e., coalescent with multiple collisions). See also the related work of [14] as well as [6, 7] which has some explicit asymptotic formulae in the particular case of a beta-coalescent.

As a development of Kingman's result, there is the following proposition. See also [23] (or [44, Exercise 5.2.1]) for an alternative construction.

Proposition 12. *There exists a Gibbs fragmentation process $(\tilde{\Pi}_\theta, \theta \geq 0)$ with weight sequence $((j-1)!, j \geq 1)$ such that for each $\theta > 0$ the distribution of $\tilde{\Pi}_\theta$ is the Gibbs distribution on $\mathcal{P}_{[n]}$ with these weights as displayed in* (84).

Proof. Given the path of a Kingman coalescent process $(\Pi_t, t \geq 0)$, construct a random tree $\mathcal{T}$ as follows. Let the vertices of the tree $\mathcal{T}$ be labelled by the random collection $\mathbf{V}$ of all subsets of $[n]$ which appear as clusters in the coalescent at some time in its evolution. Because the coalescent develops via binary mergers, starting with n singletons and terminating $\Pi_t = [n]$ for all sufficiently large t, the set $\mathbf{V}$ comprises the collection of all n singleton subsets of $[n]$, which are the *leaves* of the tree, the whole set $[n]$ which is the *root* of the tree, and $n-2$ further subsets of $[n]$, whose identities depend on how the coalescent evolves, which are the *internal vertices* of the tree. The tree $\mathcal{T}$ has $n + 1 + (n-2) = 2n-1$ vertices all together. Associate with each subset v of $[n]$ that is a vertex of the tree the time $t(v)$ at which the coalescent forms the cluster v. Thus $t(v) = 0$ if and only if v is one of the n singleton leaf vertices, $t([n]) = \inf\{t : \#\Pi_t = 1\}$, and the collection of times $t(v)$ as v ranges over the $n-1$ non-leaf vertices of the tree is the set of times t at which the process $(\#\Pi_t, t \geq 0)$ experiences a downward jump. For each non-leaf vertex v in $\mathcal{T}$, let there be exactly two edges of $\mathcal{T}$ directed from v to v_1 and v_2, where v_1 and v_2 are the two clusters which merged to form v. Let each vertex v of $\mathcal{T}$ be placed at height $t(v)$ equal to the time of its formation, and for $i = 1, 2$ regard the directed edge from v to v_i as a segment of length $t(v) - t(v_i)$. Now Kingman's

construction of $\tilde{\Pi}_\theta$ amounts to supposing that there is a Poisson process of cut points on the edges of this tree, with rate $\theta/2$ per unit length, and identifying the blocks of $\tilde{\Pi}_\theta$ with the restrictions to the set of n leaves of $\mathcal{T}$ (identified with $[n]$) of the components of the random forest obtained by cutting segments of $\mathcal{T}$ at the Poisson cut points. Now conditionally given the tree $\mathcal{T}$, construct the Poisson cut points simultaneously for all $\theta > 0$ so that for each edge of the tree of of length ℓ the moments of cuts of that edge form a homogeneous Poisson process of rate $\ell\theta/2$, and these processes are independent for different edges. Then $\tilde{\Pi}_\theta$ has been constructed simultaneously for each $\theta > 0$ in such a way that $\tilde{\Pi}_\theta$ is obviously a refinement of $\tilde{\Pi}_\phi$ for $\theta > \phi$. Since with probability one there are no ties between the times of cuts on different segments, it is clear that the process $(\tilde{\Pi}_\theta, \theta \geq 0)$ develops by binary splits. Thus $(\tilde{\Pi}_\theta, \theta \geq 0)$ is a Gibbs fragmentation process. □

While the one-dimensional distributions of this process $(\tilde{\Pi}_\theta, \theta \geq 0)$ are given by Ewens' sampling formula (84), the two and higher dimensional distributions seem difficult to describe explicitly. In particular, a calculation of the simplest transition rate associated with the process $(\tilde{\Pi}_\theta, \theta \geq 0)$, provided below, shows that this rate depends on θ. It seems quite difficult to give a full account of all transition rates of $(\tilde{\Pi}_\theta, \theta \geq 0)$, though their general form can be described and a method for their computation for small n will be indicated. For $n \geq 2$ the process $(\tilde{\Pi}_\theta, \theta \geq 0)$ turns out to be non-Markovian, so its distribution is not determined by its transition rates.

In connection with Proposition 12 and such calculations, the following problem arises:

Problem 3. *Does there exist for each n a $\mathcal{P}_{[n]}$-valued Gibbs fragmentation process $(\Pi_k, 1 \leq k \leq n)$ with weight sequence $((j-1)!, j \geq 1)$?*

7.1 Calculations with the tree derived from Kingman's coalescent.

The following calculations (Proposition 13) show that for $n \geq 4$, the discrete-time chain embedded in $(\tilde{\Pi}_\theta, \theta \geq 0)$ (that is, the sequence of successive states of $(\tilde{\Pi}_\theta, \theta > 0)$, or its discrete skeleton) does not provide a solution to Problem 3.

Let $\mathcal{T}_n$ denote the random tree derived as in the proof of Proposition 12 from Kingman's n-coalescent $(\Pi_t, t \geq 0)$, and recall the definition of the fragmentation process $(\tilde{\Pi}_\theta, \theta > 0)$. Let Θ be the time of the first cut in this process, and let $\tilde{\Pi}_\Theta$ be the state of the fragmentation at this random time. Thus almost surely $\tilde{\Pi}_\Theta$ is a partition with two blocks.

Proposition 13. *The law of Θ is determined by*

$$\mathbb{P}(\Theta \in d\theta)/d\theta = \frac{(n-1)!}{[\theta+1]_{n-1}} \sum_{i=1}^{n-1} \frac{1}{i+\theta} \tag{85}$$

For π with two components of sizes n_1 and n_2

$$\mathbb{P}(\tilde{\Pi}_\Theta = \pi \mid \Theta = \theta) = \frac{\sum_{i=1}^{n-1}(i+\theta)^{-1}\binom{n-1}{i}^{-1}\left[\binom{n_2-1}{i-1}+\binom{n_1-1}{i-1}\right]}{\binom{n}{n_1}\sum_{j=1}^{n-1}(j+\theta)^{-1}} \tag{86}$$

and

$$\mathbb{P}(\tilde{\Pi}_\Theta = \pi) = \frac{(n-1)!}{\binom{n}{n_1}}\int_0^\infty \frac{d\theta}{[\theta+1]_{n-1}}\sum_{i=1}^{n-1}\frac{\left[\binom{n_1-1}{i-1}+\binom{n_2-1}{i-1}\right]}{\binom{n-1}{i}(i+\theta)} \tag{87}$$

Proof. We only provide a sketch of the calculations leading to this result as they are somewhat tedious. For $1 \le k \le n$ let $T_k = \inf\{t : \#\Pi_t = k\}$. Let $S_i = (i+1)(T_i - T_{i+1})$, which we call the i^{th} stratum of the tree. It is clear that the total length of all segments in the tree $\mathcal{T}$ is

$$L_n := \sum_{i=1}^{n-1}(i+1)(T_i - T_{i+1}) \tag{88}$$

From the definition of the underlying coalescent process $(\Pi_t, t \ge 0)$, the random variable $T_{i+1} - T_i$ has exponential distribution with rate $i(i+1)/2$, and these random variables are independent for $1 \le i \le n-1$. It follows that

$$E\exp(-\frac{\theta}{2}L_n) = \prod_{i=1}^{n-1}\frac{i(i+1)/2}{\theta(i+1)/2 + i(i+1)/2} = \frac{(n-1)!}{[\theta+1]_{n-1}} \tag{89}$$

where $[\theta+1]_{n-1} = \prod_{i=1}^{n-1}(\theta+i)$. On the other hand, given L_n, the Poisson process with rate $\theta/2$ per unit segment length in the tree has no points with probability $\exp(-(\theta/2)L_n)$. So the expectation calculated in (89) is just the probability that $\tilde{\Pi}_\theta$ is the partition of $[n]$ with one component, or in other words that $\Theta > \theta$. Thus (85) follows by differentiation.

Now, let I denote the index of the stratum in which the first cut point falls at time Θ. Then it follows from the representation of L_n as the sum of independent exponential variables $L_n = \sum_{i=1}^{n-1} S_i$ that the sum over i in (85) corresponds to summing over the possible values i of I. That is, for $1 \le i \le n-1$,

$$\mathbb{P}(\Theta \in \theta, I = i)/d\theta = \frac{(n-1)!}{[\theta+1]_{n-1}}\frac{1}{i+\theta} \tag{90}$$

and hence

$$\mathbb{P}(I = i \mid \Theta = \theta) = \frac{(i+\theta)^{-1}}{\sum_{j=1}^{n-1}(j+\theta)^{-1}}. \tag{91}$$

Observe now that given $\Theta = \theta$ and $I = i$, the partition $\tilde{\Pi}_\Theta$ consists of two components, obtained as the restriction to $[n]$, identified as the set of leaves of the tree $\mathcal{T}$, of the two components of $\mathcal{T}$ separated by the cut at time Θ in stratum i of $\mathcal{T}$. To be precise,

$\tilde{\Pi}_\Theta = \{C, [n] - C\}$ where C is the cluster of Π_t in existence during the time interval (T_{i+1}, T_i) corresponding to the segment of $\mathcal{T}$ which is cut at time Θ. This C is one of the clusters of Π^*_{i+1}, where $(\Pi^*_k, 1 \le k \le n)$ is the discrete skeleton of $(\Pi_t, t \ge 0)$. Now by construction of the Poisson cutting process, and the fact that the discrete skeleton $(\Pi^*_k, 1 \le k \le n)$ of $(\Pi_t, t \ge 0)$ is independent of $(\#\Pi_t, t \ge 0)$, it is clear that the conditional distribution of Π^*_{i+1} given $\Theta = \theta$ and $I = i$ is identical to its unconditional distribution, that is the Gibbs distribution on $\mathcal{P}_{[n,i+1]}$ with weights $(j!, j \ge 1)$, and that C is one of the $i+1$ components of Π^*_{i+1} picked by a mechanism independent of the sizes of these components. Therefore,

$$\mathbb{P}(\#C = n_1 \mid \Theta = \theta, I = i) = \mathbb{P}(\#C_{i+1} = n_1) \tag{92}$$

where C_{i+1} is a random component of Π^*_{i+1}. After some combinatorics, it follows easily that the conditional distribution of $\tilde{\Pi}_\Theta$ given $\Theta = \theta$ and $I = i$ is given by

$$\mathbb{P}(\tilde{\Pi}_\Theta = \pi \mid \Theta = \theta, I = i) = \binom{n}{n_1}^{-1} \binom{n-1}{i}^{-1} \left[\binom{n_2 - 1}{i-1} + \binom{n_1 - 1}{i-1} \right]. \tag{93}$$

Combining this expression with (91) shows that the conditional distribution of $\tilde{\Pi}_\Theta$ given $\Theta = \theta$ is given by (86). We now easily obtain (87) from (86) and (85) by integration. □

We now briefly explain why it can be deduced from the explicit formula (87) that the discrete chain embedded in $(\tilde{\Pi}_\theta, \theta > 0)$ is not a Gibbs fragmentation. We must simply show that (87) does not coincide with the Gibbs microcanonical distribution $p_{n,2}$ associated with the weight sequence $w_j = (j-1)!$. Let J_Θ denote the size of a component of $\tilde{\Pi}_\Theta$ picked by the toss of a fair coin independent of $\tilde{\Pi}_\Theta$. Then using the above, if $n_1 = 1$ and $n_2 = n - 1$,

$$\mathbb{P}(J_\Theta = 1) = \frac{1}{2}(n-2)! \int_0^\infty \frac{d\theta}{[\theta+1]_{n-1}} \left(\sum_{i=1}^{n-1} \frac{i}{i+\theta} + \frac{1}{1+\theta} \right)$$

After a few lines of algebra, using some integration by parts and some partial fractions, we can conclude that

$$\mathbb{P}(J_\Theta = 1) = \frac{1}{2}(n-1)! \sum_{i=1}^{n-1} a_{n,i} \log i + \frac{1}{2}. \tag{94}$$

where $a_i = (-1)^{i-1} \binom{n}{i-1} / n!$.

On the other hand, let Π have the Gibbs $(n, 2, w)$ distribution with $w_j = (j-1)!$, and J is the size of a randomly picked component. Remark that by decomposing on the size of the cycle containing 1,

$$B_{n,2} = \sum_{j=1}^{n-1} \binom{n-1}{j-1} (j-1)!(n-1-j)! = (n-1)! H_{n-1}$$

where $H_{n-1} := \sum_{j=1}^{n-1} 1/j$. Since the number of permutations with exactly two cycles one of which has size 1 is $n(n-2)!$, we conclude that $\mathbb{P}(J=1) = \frac{1}{2}\frac{n}{(n-1)H_{n-1}}$. This is incompatible with (94). Indeed if this was to be equal to right-hand side in (94) one would get $\log(\prod_{i=1}^{n-1} i^{r_i}) = q$ for some rational number q and $r_i = (-1)^i\binom{n}{i-1} \in \mathbf{Z}$, and thus $e^q = q'$ for some (other) rational numbers q and q'. This contradicts the transcendence of e.

Thus the distribution of the partition of $[n]$ into two parts obtained at the time of the first split is not the common distribution of $\tilde{\Pi}_\theta$ given $\#\tilde{\Pi}_\theta = 2$ for all $\theta > 0$. In particular, for $n = 4$ the formulas above give:

$$\mathbb{P}(J_\Theta = 1) = \mathbb{P}(J_\Theta = 3) = 3(-\log 2 + \frac{1}{2}\log 3) + \frac{1}{2}$$

and

$$\mathbb{P}(J_\Theta = 2) = 6\log 2 - 3\log 3$$

whereas

$$\mathbb{P}(J=1) = \mathbb{P}(J=3) = 4/11$$
$$\mathbb{P}(J=2) = 3/11$$

We conclude that the discrete skeleton of $(\tilde{\Pi}_\theta)_{\theta\geq 0}$, i.e., the discrete fragmentation chain embedded in it, does not give a discrete Gibbs fragmentation associated with $w_j = (j-1)!$.

7.2 A reformulation with walks on the symmetric group

In view of the combinatorial interpretation of Example 2, Problem 3 can be restated as:

Problem 4. *Does there exist for each n a sequence of random permutations $(\sigma_k, 1 \leq k \leq n)$ such that σ_k has uniform distribution on the set of permutations of $[n]$ with k cycles, and for $k \leq \ell$ the partition generated by the cycles of σ_ℓ is a refinement of σ_k?*

This problem may be partially reformulated in terms of random walks on the symmetric group. Suppose we consider the Cayley graph G_n of the symmetric group induced by the set of generators $S = \{\text{all transpositions}\}$, that is, we put an edge between two permutations σ and π if and only if σ may be written as $\sigma = \tau \cdot \pi$ for some transposition τ. It is well-known that multiplying a permutation by a transposition can only result in a coagulation or a fragmentation in the cycle structure. More precisely, suppose $C = (x_1, \ldots, x_k)$ is a cycle of the permutation π. If we multiply by the transposition $\tau = (x_i, x_j)$ then the resulting cycle structure in σ is the same as that of π except that C breaks into $(x_1, \ldots, x_{i-1}, x_j, x_{j+1}, \ldots, x_k)$ on the one hand and $(x_i, \ldots, x_{j-1})$ on the other hand. Conversely, suppose $C = (x_1, \ldots, x_k)$ and $C' = (y_1, \ldots, y_l)$ are two cycles of π and we multiply π by the transposition (x_i, y_j). In the resulting permutation, C and C' will be replaced by a unique cycle $C'' = (x_1, \ldots, x_{i-1}, y_j, y_{j+1}, \ldots, y_{j-1}, x_i, \ldots, x_k)$. In particular, any (random) walk on G_n may be viewed as a coagulation and fragmentation process on the cycle structure of the permutation. Moreover, a well-known result

due to Cayley states that if σ is a permutation then the *graph distance* between σ and the identity permutation I (i.e., the minimum number of edges one must cross to go from I to σ on G_n) is simply $n - \#\text{cycles of } \sigma$. As a consequence, by considering a time-reversal of the process, to solve Problem 4, it is enough to construct a random process $(\sigma_k)_{0 \leq k \leq n-1}$ on G_n which has the following two properties:

1. The sequence $(\sigma_k, 0 \leq k \leq n-1)$ is a random walk on G_n, in the sense that if $\sigma_k = \sigma$ at the next stage σ can only jump to one the neighbors of σ.

2. The permutation σ_k has the following marginal distribution: at stage k the distribution of σ_k is uniform on the sphere of radius k about the identity, that is the set of all permutations whose distance to the identity is k.

Property 1 ensures that the cycles of σ_k perform a coagulation-fragmentation process. In conjunction with property 2, since σ_k must be at distance k, it must be the case that all jumps of σ are produced by some fragmentation. Moreover, by Cayley's result for the distance of a permutation to the identity, if σ_k is uniform on the sphere of radius k then its cycle structure is a realization of the Gibbs distribution (9) with weight sequence $w_j = (j-1)!$ (Note however this is not strictly equivalent to Problem 4 since not all fragmentations at the level of partitions can be represented by moving along some edge of G_n. For instance, it is impossible to get from the permutation (1 2 3 4) to the permutation (1 3)(2 4) in one step).

In this context, a very natural process to consider is the simple random walk on G_n, conditioned to never backtrack. In other words, starting from the identity, at each step choose uniformly among all edges that lead from distance k to distance $k+1$. Although this may seem a very natural candidate for properties 1 and 2, this is far from being the case. Much is known about this process, and in particular it has been shown in [8] that the distribution of this process at time $k = \lfloor an \rfloor$ for any $0 < a < 1$ is asymptotically singular with respect to the uniform distribution on the sphere of radius k.

Acknowledgements

We thank Jomy Alappattu for carefully reading a draft of this paper and for pointing out some mistakes, as well as some helpful discussions. We thank the referees for some useful suggestions.

References

[1] D.J. Aldous. Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. *Bernoulli*, 5: 3–48, 1999.

[2] R. Arratia, A. Barbour and S. Tavaré. *Logarithmic combinatorial structures: a probabilistic approach.* European Math. Society Monographs, 1, 2003.

[3] A. Barbour and B. Granovsky. Random combinatorial structures: the convergent case. *J. Comb. Theory, Ser. A* 109(2): 203-220, 2005.

[4] M.H. Bayewitz, J. Yerushalmi, S. Katz, and R. Shinnar. The extent of correlations in a stochastic coalescence process. *J. Atmos. Sci.*, 31:1604–1614, 1974.

[5] J. Berestycki. Exchangeable fragmentation-coalescence processes and their equilibrium distribution. *Electr. J. Prob.*, 9:770-824, 2004.

[6] J. Berestycki, N. Berestycki and J. Schweinsberg. Small-time behavior of Beta-coalescents. Preprint, `math.PR/0601032`, 2006.

[7] J. Berestycki, N. Berestycki and J. Schweinsberg. Beta-coalescents and continuous stable random trees. Preprint, `math.PR/0602113`, 2006.

[8] N. Berestycki. The hyperbolic geometry of random transpositions. *Ann. Probab.*, 34(2), 429–467, 2006.

[9] L. Comtet. *Advanced Combinatorics.* D. Reidel Pub. Co., Boston, 1974. (translated from French).

[10] P. C. Consul and F. Famoye. *Lagrangian probability distributions.* Birkhäuser Boston Inc., Boston, MA, 2006.

[11] P. C. Consul and L. R. Shenton. Use of Lagrange expansion for generating discrete generalized probability distributions. *SIAM J. Appl. Math.*, 23:239–248, 1972.

[12] L. Devroye. The branching process method in the Lagrange random variate generation, *Communications in Statistics—Simulation*, 21, 1–14, 1992.

[13] A. Di Bucchianico. *Probabilistic and analytical aspects of the umbral calculus*, volume 119 of *CWI Tract.* Stichting Mathematisch Centrum Centrum voor Wiskunde en Informatica, Amsterdam, 1997.

[14] R. Dong, A. Gnedin, and J. Pitman. Exchangeable partitions derived from Markovian coalescents. Preprint, `math.PR/0603745`.

[15] R. Durrett, B. L. Granovsky, and S. Gueron. The equilibrium behavior of reversible coagulation-fragmentation processes. *J. Theoret. Probab.*, 12(2):447-474, 1999.

[16] R. Durrett and J. Schweinsberg. Power laws for family sizes in a duplication model *Ann. Probab.* 33,6: 2094–2126, 2005.

[17] M. Erlihson and B. Granovsky. Reversible coagulation-fragmentation processes and random combinatorial structures: asymptotics for the number of groups. *Rand. Struct. Algor.*, 25, 227–245, 2004.

[18] P. Erdős, R.K. Guy, and J.W. Moon. On refining partitions. *J. London Math. Soc.*, II. Ser 9:565–570, 1975.

[19] W.J. Ewens. The sampling theory of selectively neutral alleles. *Theor. Popul. Biol.*, 3:87–112, 1972.

[20] P. J. Flory. *Principles of polymer chemistry.* Ithaca, London: Cornell University Press 1953

[21] B. Fristedt. The structure of partitions of large integers. *Trans. Amer. Math. Soc.*, 337:703-735, 1993.

[22] A. Gnedin and J. Pitman. Exchangeable Gibbs partitions and Stirling triangles. *Zapiski St. Petersb. Dept. Math. Inst.* 325: 82-103, 2005.

[23] A. Gnedin and J. Pitman. Poisson representation of a Ewens fragmentation process. Preprint, `arXiv:math.PR/0608307`.

[24] W. M. Y. Goh and E. Schmutz. Random set partitions. *SIAM J. Discrete Math.*, 7:419-436, 1994.

[25] H.W. Gould. A series of transformations leading to convolution identities. *Duke Math. J.*, 28:193-202, 1961.

[26] B. Haas, G. Miermont, J. Pitman and M. Winkel. Continuum tree asymptotics of discrete fragmentations and applications to phylogenetic models. Preprint, `math.PR/0604350`.

[27] L.H. Harper. The morphology of partially ordered sets. *J. Combinatorial Theory Ser. A*, 17:44–58, 1974.

[28] E.M. Hendriks, J.L. Spouge, M. Eibl, and M. Shreckenberg. Exact solutions for random coagulation processes. *Z. Phys. B - Condensed Matter*, 58:219–227, 1985.

[29] L. Holst. On the lengths of the pieces of a stick broken at random. *J. Appl. Probab.*, 17:623 – 634, 1980.

[30] L. Holst. On numbers related to objects of unlike partitions and occupancy problems. *European J. Combin.* 2(3):231-237, 1981.

[31] S. Janson. Conditioned Galton-Watson trees do not grow. Preprint, `math.PR/0604141`.

[32] T. Kamae, U. Krengel and G.L. O'Brien, Stochastic inequalities on partially ordered spaces. *Ann. Probab.* 5: 899-912, 1977.

[33] F. Kelly. *Reversibility in stochastic networks*, Wiley, 1979

[34] J. F. C. Kingman. The coalescent. *Stochastic Processes and their Applications*, 13:235–248, 1982.

[35] J. F. C. Kingman. On the genealogy of large populations. In *Essays in Stastical Science* (eds J. Gani and E. J. Hannan, Applied Probability Trust, Sheffield), *J. Appl. Prob.* Spec. Vol 19A, 27–43, 1982.

[36] V. F. Kolchin. *Random graphs.* Cambridge University Press 1999.

[37] V. F. Kolchin. *Random mappings.* Translation series in Mathematics and Engineering. Optimization Software Inc. Publications Division, New York, 1986. MR88:a60022.

[38] D. Knuth. Convolution polynomials. *Mathematica journal*, 2, 4:67–78, 1992. Reprinted as Chapter 41 of *Selected Papers of Discrete Mathematics* (Stanford, California, Center for the Study of Language and Information). Available from the arXiv server as `math.CA/9207221`.

[39] A.A. Lushnikov. Coagulation in finite systems. *J. Colloid and Interface Science*, 65:276–285, 1978.

[40] A.H. Marcus. Stochastic coalescence. *Technometrics*, 10:133–143, 1968.

[41] M. Möhle. On sampling distributions for coalescent processes with simultaneous multiple collisions. *Bernoulli*, 12, 1:35-53, 2006.

[42] J. Pitman. Enumerations of trees and forests related to branching processes and random walks. *Microsurveys in Discrete Probability*, D. Aldous and J. Propp editors. DIMACS Ser. Discrete Math. Theoret. Comp. Sci no. 41 163-180. Amer. Math. Soc. Providence RI, 1998.

[43] J. Pitman. Coalescent random forests. *J. Combin. Theory A.*, 85:165-193, 1999.

[44] J. Pitman. Combinatorial stochastic processes. *Lecture notes in mathematics, Ecole d'Eté de probabilités de Saint-Flour* XXXII-2002. Vol. 1875, Springer, 2006.

[45] A. Rényi. Probabilistic methods in combinatorial mathematics. In R.C. Bose and T.A. Dowlilng, eds. *Combinatorial mathematics and its applications*, page 1-13. Univ. of North Carolina Press, Chapel Hill 1969.

[46] R. P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 1999.

[47] A. M. Vershik. Statistical mechanics of combinatorial partitions and their limit shapes. *Funct. Anal. Appl.*, 30:90–105, 1996 (translation from Russian).

[48] P. Whittle. The equilibrium statistics of a clustering process in uncondensed phase. *Proc. Roy. Lond. Soc. A,* 285:501–519, 1965.

[49] P. Whittle. Statistical processes of aggregation and polymerisation. *Proc. Camb. Phil. Soc.*, 61:475–495, 1965.

[50] P. Whittle. *Systems in Stochastic Equilibrium*. Wiley, 1986.