Conditioned multi-type Galton-Watson trees

Article *in* ESAIM Probability and Statistics · June 2015

DOI: 10.1051/ps/2016019 · Source: arXiv

Eric Cator
Radboud University

Constructing conditioned multi-type Galton-Watson trees

Eric Cator, Henk Don

Institute for Mathematics, Astrophysics, and Particle Physics

Faculty of Science, Radboud University Nijmegen; e-mail:e.cator@math.ru.nl, henkdon@gmail.com

June 11, 2015

Abstract

We consider multi-type Galton Watson trees, and find the distribution of these trees when conditioning on very general types of recursive events. It turns out that the conditioned tree is again a multi-type Galton Watson tree, possibly with more types and with offspring distributions, depending on the type of the father node and on the height of the father node. These distributions are given explicitly. We give some interesting examples for the kind of conditioning we can handle, showing that our methods have a wide range of applications.

1 Introduction

The asymptotic shape of conditioned Galton-Watson trees has been widely studied. For example, one could condition on the number of nodes of the tree being n, and letting $n \to \infty$. In a lot of cases, the limiting tree is quite well understood, see for example the survey paper by Janson [4]. Some work on finite conditioned trees has been done by Geiger and Kersting [2], who studies the shape of a tree conditioned on having height exactly equal to n. In this context, we also mention the spinal construction of a Galton-Watson tree conditioned to reach generation k, as derived in [3].

In this paper, we will investigate conditioning multi-type Galton-Watson trees on events of a recursive nature (as explained in Section 2), one example being conditioning

on survival to a given level. The main idea is that we consider different classes of trees, where the class of a tree is determined by the types and classes of her children. The offspring distribution of a node depends on its type and on the level of the tree where this node is living. In fact, we show that the conditioned tree again is a multi-type Galton-Watson tree and how this can be used to directly construct such a conditioned tree. Our approach can be seen as a generalization of the well-known decomposition of a supercritical Galton-Watson tree into nodes whose offspring survives forever and nodes whose offspring eventually goes to extinction as discussed in [5].
Section 2.2 discusses a couple of examples that illustrate the applicability of our results. We give an example concerning mutants in a population, we discuss an alternative to Geiger's construction of a tree conditioned on having height exactly k and we show how to condition on the size of the kth generation.

1.1 Notation and preliminaries

We will consider rooted multi-type Galton-Watson trees with arbitrary offspring distribution, that can depend on the current generation. In such a tree each node has a type, which we indicate by a natural number $t \in \Theta := \{1, \dots, \theta\}$. If the root of a tree has type t, we use a bold-face t to denote this root. Define the set of trees of heigth 0 as $\mathcal{T}_0 = \{\mathbf{1}, \mathbf{2}, \dots, \boldsymbol{\theta}\}$. Then we define inductively for $k \geq 1$ the set of trees of height at most k by

$$\mathcal{T}_k = \mathcal{T}_0 \sqcup \bigsqcup_{n=1}^{\infty} \{(t, [T_1, \dots, T_n]) \mid T_i \in \mathcal{T}_{k-1}, t \in \Theta\},$$

and denote the set of all trees by $\mathcal{T} = \bigcup_{k=0}^{\infty} \mathcal{T}_k$. For a tree $T = (t, [T_1, \dots, T_n]) \in \Theta \times (\mathcal{T}_{k-1})^n$, the trees $T_1, \dots, T_n$ will be called the children of T (notation: $T_i \prec T$). The type of T will be just the type of its root and will be denoted by $r(T)$. We now define a function $N : \mathcal{T} \to \mathbb{N}^\theta$ that counts how many children of each type a tree T has:

$$N(T) = (N_1(T), \dots, N_\theta(T)), \quad \text{where} \quad N_t(T) = \# \left\{ \tilde{T} \prec T : r(\tilde{T}) = t \right\}.$$

The set of trees of heigth at most k having a root of type t is denoted by

$$\mathcal{T}_k^t = \{T \in \mathcal{T}_k : r(T) = t\} = \{\mathbf{t}\} \sqcup \bigsqcup_{n=1}^{\infty} \{(t, [T_1, \dots, T_n]) \mid T_i \in \mathcal{T}_{k-1}\}.$$

Let $\mathcal{T}^t = \bigcup_{k=0}^{\infty} \mathcal{T}_k^t$. Denote the offspring distribution of a type t node at height $l \geq 0$ by μ_l^t, for arbitrary probability measures μ_l^t on $\mathbb{N}^\theta = \{0, 1, \dots\}^\theta$. Define independent random variables $W_l^t \sim \mu_l^t$. For a vector $\mathbf{x} \in \mathbb{N}^\theta$, we write the corresponding multinomial coefficient as

$$D(\mathbf{x}) = \binom{|\mathbf{x}|_1}{x_1, \ldots, x_\theta} = \frac{\left(\sum_{i=1}^{\theta} x_i\right)!}{\prod_{i=1}^{\theta} x_i!}.$$

We now introduce the Galton-Watson probability measures on $\mathcal{T}_k^t$. Firstly, let P_0^t be the trivial probability measure on $\mathcal{T}_0^t$, so $P_0^t(\mathbf{t}) = 1$. Now define inductively the probability measure P_{lk}^t for $0 \le l \le k$ and $k \ge 1$ as the following probability measure on $\mathcal{T}_{k-l}^t$: if $l = k$, then $P_{kk}^t = P_0^t$. Otherwise for all $T \in \mathcal{T}_{k-l}^t$

$$P_{lk}^t(T) = \frac{\mathbb{P}(W_l^t = N(T))}{D(N(T))} \prod_{\tilde{T} \prec T} P_{l+1,k}^{r(\tilde{T})}(\tilde{T}),$$

where empty products are taken to be 1. The intuition is that the second sub-index determines the size of the final tree we are considering, whereas the first sub-index determines at which level we are building up the tree (so P_{lk}^t generates trees of type t at level l of size $k-l$). We are interested in P_{0k}^t, which is the Galton-Watson probability measure on $\mathcal{T}_k^t$ (trees cut off at height k with a root of type t).

In the next section we will introduce a class of recursive-type events on which we would like to condition, and discuss several examples of such events. In Section 3 we will introduce the conditional measures corresponding to our events, and in Section 4 we show that these conditional measures indeed coincide with the original Galton-Watson measure, conditioned on our event.

2 Conditioning on recursive events

In this section we introduce a class of recursive-type events on which we would like to condition, such as the event that the tree survives until a specific level.

2.1 Partitioning the set of trees

We will now set up our general framework and show how some examples fit into it. We start by choosing $k_0 \in \mathbb{N}$ and partitioning $\mathcal{T}_{k_0}$ into m classes $A_{k_0}^{(1)}, \ldots, A_{k_0}^{(m)}$. Typically, all trees in such a class have some property that all trees in the other classes do not have. One of the simplest examples would be a partition into two classes, where trees that survive until some level k are in the first class and all other trees in the second.

The partition of $\mathcal{T}_{k_0}$ will be the starting point to recursively define partitions of $\mathcal{T}_l$, $k_0 < l \leq k$ into sets $A_l^{(i)}, i = 1, \ldots, m$, where k is the (maximum) height of the trees that we are considering. Suppose that the partition of $\mathcal{T}_{l-1}$ is already defined. Then we are able to introduce a counting matrix for trees in $\mathcal{T}_l$. Define $C_l : \mathcal{T}_l \to \mathbb{N}^{m \times \theta}$ such that for $T \in \mathcal{T}_l$ the (i,j)th position is given by

$$C_l^{(i,j)}(T) = \# \left\{ \tilde{T} \prec T : \tilde{T} \in A_{l-1}^{(i)} \cap \mathcal{T}^j \right\},$$

so this is the number of children of T having type j and being an element of the ith partition class. Now for $k_0 < l \leq k$ we partition $\mathbb{N}^{m \times \theta}$ into subsets $B_{l,1}, \ldots, B_{l,m}$. This partition is the key for the recursive definition of $A_l^{(i)}$. The set $A_l^{(i)}$ will contain exactly those trees for which the counting matrix $C_l(T)$ is in $B_{l,i}$:

$$A_l^{(i)} = \{ T \in \mathcal{T}_l : C_l(T) \in B_{l,i} \} . \tag{2.1}$$

2.2 Examples

Before going into the details of the construction of conditioned trees, we will discuss some examples of recursive events that can be handled by our approach.

2.2.1 Genetic mutations

Suppose we have a population in which sometimes an individual (mutant) is born having a particular mutation in its genetic material. This mutation can be inherited by subsequent generations. Suppose we know the probability that the root is mutated. Such a population can be described as a two-type Galton-Watson process in which the offspring distribution is type- and possibly level-dependent. We take $\Theta = \{1, 2\}$ to be the set of types, where mutants have type 1.

Suppose we would like to condition on the event "there is at least one mutant in the kth generation". Choose $k_0 = 0$ and partition $\mathcal{T}_0 = \{\mathbf{1}, \mathbf{2}\}$ into the classes $A_0^{(1)} = \{\mathbf{1}\}$ and $A_0^{(2)} = \{\mathbf{2}\}$. For $0 < l \leq k$, we want to define $A_l^{(1)}$ and $A_l^{(2)}$ by

$$A_l^{(1)} = \{ T \in \mathcal{T}_l : \text{there is a mutant at level } l \} , \qquad A_l^{(2)} = \mathcal{T}_l \backslash A_l^{(1)}.$$

These events satisfy a recursive relation: $A_l^{(1)}$ contains exactly those trees that have at least one child in $A_{l-1}^{(1)}$. For $T \in \mathcal{T}_l$, the first row of the 2×2-counting matrix $C_l(T)$

counts the children of T that are in $A_{l-1}^{(1)}$. Therefore, for all $0 < l \leq k$ we let

$$B_{l,1} = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{N}^{2\times 2} : a + b \geq 1 \right\}, B_{l,2} = \mathbb{N}^{2\times 2} \backslash B_{l,1},$$

and now (2.1) gives the desired partition of T_l. For the sake of illustration, we note that with a minor change, we can condition on "there is at least one mutant in the kth generation inheriting its mutation from the root". To achieve this, it suffices to merely redefine $B_{l,1}$ and $B_{l,2}$ for all $0 < l \leq k$ as follows

$$B_{l,1} = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{N}^{2\times 2} : a \geq 1 \right\}, B_{l,2} = \mathbb{N}^{2\times 2} \backslash B_{l,1}.$$

In these two examples, we defined one partition class $A_l^{(1)} \subseteq \mathcal{T}_l$ by the event on which conditioning is required. The only other partition class was just the complement of the first one. Finding a suitable partition of the set of trees is not always that obvious, as is demonstrated in the next example. We will show how to condition on the slightly more complicated event "All mutants in the tree inherit their mutation from the root and at least one mutant is present in generation k". As before, define one partition class $A_l^{(1)}$ as the set of trees satisfying the condition. Here it is not sufficient to define only one other partition class. One obstacle is that some trees (namely those with a "spontaneous mutation") in the complement $\mathcal{T}_l \setminus A_l^{(1)}$ are forbidden as a child of trees in $A_{l+1}^{(1)}$ and others are not.
Nevertheless, with a slightly more elaborate partition, we can still handle this case. We distinguish four classes and partition $\mathcal{T}_0$ into

$$A_0^{(1)} = \{\mathbf{1}\}, \quad A_0^{(2)} = \{\mathbf{2}\} \quad \text{and} \quad A_0^{(3)} = A_0^{(4)} = \emptyset.$$

For $0 < l \leq k$, we define the following subsets of $\mathbb{N}^{4\times 2}$:

$$B_{l,1} = \left\{ \begin{bmatrix} a & 0 \\ b & c \\ 0 & 0 \\ d & 0 \end{bmatrix} : a \geq 1 \right\}, \qquad B_{l,2} = \left\{ \begin{bmatrix} 0 & 0 \\ 0 & a \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \right\}$$

$$B_{l,3} = \left\{ \begin{bmatrix} a & b \\ c & d \\ e & f \\ g & h \end{bmatrix} : b + e + f + h \geq 1 \right\}, \quad B_{l,4} = \left\{ \begin{bmatrix} 0 & 0 \\ a & b \\ 0 & 0 \\ c & 0 \end{bmatrix} : a + c \geq 1 \right\}$$

As can be easily checked, these sets are disjoint and $\bigcup_i B_{l,i} = \mathbb{N}^{4\times 2}$, so this indeed is a partition. It follows by induction that the sets $A_l^{(i)}$ partition $\mathcal{T}_l$ in such a way that

- $A_l^{(1)}, 0 < l \leq k$ contains exactly the trees having
  - at least one mutated child of which the mutated progeny reaches level l, and
  - no "spontaneous" mutants in the progeny of their children.
- $A_l^{(2)}, 0 < l \leq k$ contains the trees having only type 2 descendants.
- $A_l^{(3)}, 0 < l \leq k$ contains the trees having a type 2 descendant with a type 1 child ("spontaneous mutation").
- $A_l^{(4)}, 0 < l \leq k$ contains all other trees in $\mathcal{T}_l$.

Note that these classes are defined by properties of the children of a tree and not by the type of the tree itself. For example, a tree in $A_l^{(2)}$ can have a type 1 root, but all its descendants have type 2. The conditional measure we are interested in is now obtained by conditioning P_{0k}^1 on $A_k^{(1)}$.

2.2.2 Conditioning on the size of generation k

As a next example, we show how to condition a single-type Galton-Watson tree on having exactly G individuals in the kth generation. In this case, we partition $\mathcal{T}_0 = \{\mathbf{1}\}$ into $G+2$ classes by defining

$$A_0^{(1)} = \{\mathbf{1}\}, \quad A_0^{(0)} = A_0^{(2)} = A_0^{(3)} = \ldots = A_0^{(G)} = A_0^{(G+1)} = \emptyset.$$

Define $x \in \mathbb{N}^{G+2}$ by $x := [0 \quad 1 \quad 2 \quad \ldots \quad G \quad G+1]^T$. For $0 < l \leq k$, we define

$$B_{l,i} = \left\{y \in \mathbb{N}^{G+2} : x^T y = i\right\}, \quad B_{l,G+1} = \left\{y \in \mathbb{N}^{G+2} : x^T y \geq G+1\right\},$$

where $0 \leq i \leq G$. Partitioning $\mathcal{T}_l$ according to (2.1) gives the following: for $0 \leq i \leq G$, $A_l^{(i)}$ contains the trees of which the lth generation has exactly size i, while $A_l^{(G+1)}$ contains the trees of which the lth generation has at least size $G+1$. Conditioning P_{0k} on $A_k^{(G)}$ gives the result we are looking for.

2.2.3 The tree has heigth exactly k

As a final illustration, we explain how to condition a Galton-Watson tree on having height exactly k, thus producing an alternative for the construction of Geiger and Kersting [2]. We consider trees in $\mathcal{T}_{k+1}$ that are conditioned to reach level k, but not level $k+1$. We start by choosing $k_0 = 2$, and partitioning $\mathcal{T}_2$ into three sets, namely correct trees, short trees and long trees:

$$\begin{aligned} A_2^{(1)} &= \{T \in \mathcal{T}_2 \mid T \text{ reaches level 1, but not level 2}\}, \\ A_2^{(2)} &= \{T \in \mathcal{T}_2 \mid T \text{ does not reach level 1}\} = \{\mathbf{0}\}, \\ A_2^{(3)} &= \{T \in \mathcal{T}_2 \mid T \text{ reaches level 2}\}. \end{aligned}$$

Define for each $2 < l \leq k+1$

$$\begin{aligned} B_{l,1} &= \{n \in \mathbb{N}^3 \mid n_1 \geq 1, n_3 = 0\}, \\ B_{l,2} &= \{n \in \mathbb{N}^3 \mid n_1 = 0, n_3 = 0\}, \\ B_{l,3} &= \{n \in \mathbb{N}^3 \mid n_3 \geq 1\}, \end{aligned}$$

and let $\mathcal{T}_l$ be partitioned as in (2.1). This construction guarantees that if a tree $T \in \mathcal{T}_{k+1}$ is an element of $A_{k+1}^{(1)}$, then it has at least one child that reaches level k, and no children that reach level $k+1$. If $T \in A_{k+1}^{(2)}$, all its children do not reach level k, and if $T \in A_{k+1}^{(3)}$, then at least one child reaches level $k+1$. Conditioning on being in $A_{k+1}^{(1)}$ therefore gives the desired result.

2.3 Remarks following the examples

As it turns out from the examples in the previous section, the setup allows to condition on quite a variety of events. A fundamental requirement on these events is that they are determined only by the *number* of children of a tree having particular properties. So we can (for instance) not distinguish between trees having the same children in a different order.

An additional example is discussed in detail in [1]. As an application of the theory developed in the present paper, the cost of searching a tree to a given level is determined. The proposed model takes into account costs for having a lot of children, but also for walking into dead ends. So both a high expected offspring and a low expected offspring would give high search costs. This gives rise to an optimization problem: which offspring distribution gives minimal costs? For this model the conditional probability measures are explicitly constructed, leading to recursions that enable us to calculate the costs and solve the optimization problem for Poisson offspring.

Conditioning on recursive events as in the examples allows us to compute (conditional) probabilities that are defined in terms of such events. As an illustration: in the example on genetic mutations we can easily compute the probability that the root is mutated, given that there is at least one mutant in generation k. What makes the results even more useful is that they show how to directly construct a tree conditioned on some event. This means that trees conditioned on (rare) events can be studied by just simulating them.

3 Conditional measures

In this section we construct an alternative measure $\tilde{P}^t_{lk}$ on $\mathcal{T}^t_{k-l}$, that depends on the event we want to condition on. As soon as we have this measure, conditioning on the desired event is a triviality. In the next section, we will show that in fact the two measures P^t_{lk} and $\tilde{P}^t_{lk}$ are the same.

Define ${}^t p^{(i)}_{lk}$ for $0 \le l \le k - k_0$ by

$$ {}^t p^{(i)}_{lk} = P^t_{lk}(A^{(i)}_{k-l} \cap \mathcal{T}^t). $$

We can calculate this probability in a recursive way. Denote, for $q \in [0,1]^m$ with $\sum q_i = 1$, by $\mathrm{Multi}(n, q) \in \mathbb{N}^m$ the multinomial distribution where we distribute n elements over m classes, according to the probabilities q_i. We also choose independent random vectors $W^t_l \sim \mu^t_l$ according to the offspring distribution of a type t node at level l and denote the jth coordinate by $W^t_{l,j}$. Then, for $l < k - k_0$

$$ {}^t p^{(i)}_{lk} = \mathbb{P}\left(\bigotimes_{j=1}^{\theta} \mathrm{Multi}\left(W^t_{l,j}, ({}^j p^{(1)}_{l+1,k}, \ldots, {}^j p^{(m)}_{l+1,k}) \right) \in B_{k-l,i} \right), \tag{3.1} $$

where, for $a_0, \ldots, a_\theta \in \mathbb{N}^m$, we defined $\Lambda := \bigotimes_{j=1}^{\theta} a_j \in \mathbb{N}^{m \times \theta}$ to be the matrix for which $\Lambda_{ij} = a_j(i)$.
We proceed by defining the conditional measure ${}^t\tilde{Q}^{(i)}_{lk}$ on $A^{(i)}_{k-l} \cap \mathcal{T}^t$. To do this, define for each $t \in \{1, \ldots, \theta\}$ and $0 \le l \le k - k_0 - 1$ on the same probability space as W^t_l, the random matrices

$$ {}^t X_{lk} = \begin{pmatrix} {}^t X^{(1,1)}_{lk} & \ldots & {}^t X^{(1,\theta)}_{lk} \\ \vdots & \ddots & \vdots \\ {}^t X^{(m,1)}_{lk} & \ldots & {}^t X^{(m,\theta)}_{lk} \end{pmatrix}, $$

such that conditional on W^t_l, all columns are independent and the distribution of the jth column satisfies

$$ ({}^t X^{(1,j)}_{lk}, \ldots, {}^t X^{(m,j)}_{lk}) \mid W^t_l \sim \mathrm{Multi}\left(W^t_{l,j}, ({}^j p^{(1)}_{l+1,k}, \ldots, {}^j p^{(m)}_{l+1,k}) \right). $$

This determines the full joint distribution of $(W_l^t, {}^tX_{lk})$. For a type t node at level l, the distribution of its children over the θ types is given by the random vector W_l^t. Furthermore, the jth column of ${}^tX_{lk}$ represents how the type j children of this type t node are distributed over the m classes. For $l = k - k_0$, we define for each $T \in A_{k_0}^{(i)} \cap \mathcal{T}^t$

$$ {}^t\tilde{Q}_{k-k_0,k}^{(i)}(T) = \frac{P_{k-k_0,k}^t(T)}{P_{k-k_0,k}^t(A_{k_0}^{(i)} \cap \mathcal{T}^t)}, $$

as a probability measure on $A_{k_0}^{(i)} \cap \mathcal{T}^t$. Next, we inductively define the probability measures ${}^t\tilde{Q}_{lk}^{(i)}$ on $A_{k-l}^{(i)} \cap \mathcal{T}^t$ for each $0 \leq l \leq k - k_0 - 1$ such that for each $T \in A_{k-l}^{(i)} \cap \mathcal{T}^t$

$$ {}^t\tilde{Q}_{lk}^{(i)}(T) = \frac{\mathbb{P}({}^tX_{lk} = C_{k-l}(T) \mid {}^tX_{lk} \in B_{k-l}^{(i)})}{D(C_{k-l}(T))} \prod_{j=1}^{m} \prod_{\tilde{T} \prec T : \tilde{T} \in A_{k-l-1}^{(j)}} {}^{r(\tilde{T})}\tilde{Q}_{l+1,k}^{(j)}(\tilde{T}), $$

where we extended the definition of D (see Section 1.1) to integer-valued matrices, and once again empty products are taken to be 1. Note that this definition is valid for all $T \in \mathcal{T}_{k-l}^t$: we simply get ${}^t\tilde{Q}_{lk}^{(i)}(T) = 0$ whenever $T \notin A_{k-l}^{(i)}$. We can now define the alternative measure $\tilde{P}_{lk}^t$ on $\mathcal{T}_{k-l}^t$:

$$ \tilde{P}_{lk}^t(T) = \sum_{i=1}^{m} {}^tp_{lk}^{(i)} {}^t\tilde{Q}_{lk}^{(i)}(T). $$

3.1 Construction according to the conditional measure

We can describe the random tree $\mathbb{T} \sim \tilde{P}_{lk}^t$ as follows. The root of the tree has type t. To construct the tree, we first toss an m-sided coin to determine in which of the m classes $\mathbb{T}$ is, giving probability ${}^tp_{lk}^{(i)}$ to the ith class $A_{k-l}^{(i)}$. If $\mathbb{T} \in A_{k-l}^{(i)}$, then we choose it according to ${}^t\tilde{Q}_{lk}^{(i)}$. This means that we choose $(\tilde{W}_{lk}^t, {}^t\tilde{X}_{lk})$, where $\tilde{W}_{lk}^t$ counts the numbers of children of $\mathbb{T}$ of each type and ${}^t\tilde{X}_{lk}$ counts for each type the numbers of children that will lie in each of the m classes, according to

$$ (\tilde{W}_{lk}^t, {}^t\tilde{X}_{lk}) \sim (W_l^t, {}^tX_{lk}) \mid {}^tX_{lk} \in B_{k-l}^{(i)}. $$

The $\sum_{i,j} {}^t\tilde{X}_{lk}^{(i,j)}$ children are distributed over the $\sum_j \tilde{W}_{lk,j}^t$ positions uniformly at random. Then for each child of type j in $A_{k-l-1}^{(i)}$ we draw a tree according to ${}^j\tilde{Q}_{l+1,k}^{(i)}$. In this way we have described the random tree as a Galton-Watson tree with $m \times \theta$ 'types' of children and type- and level-dependent offspring distribution. Note that conditioning $\tilde{P}_{lk}^t$ on $A_{k-l}^{(i)}$ is trivial: we simply have to draw $\mathbb{T}$ according to ${}^t\tilde{Q}_{lk}^{(i)}$.

4 The two random trees are equally distributed

The following theorem shows that the construction procedure of Section 3 in fact generates trees with the same probabilities as under the original Galton-Watson measure. Fix k and k_0 and define all measures as before.

Theorem 1 *For all $0 \leq l \leq k-k_0$, $t \in \Theta$ and $T \in \mathcal{T}_{k-l}^t$,*

$$P_{lk}^t(T) = \tilde{P}_{lk}^t(T).$$

Proof: The theorem is true by construction for $l = k-k_0$. Now suppose that we have already shown that $P_{l+1,k}^t = \tilde{P}_{l+1,k}^t$ for all t. Choose $T \in \mathcal{T}_{k-l}$ and suppose $T \in A_{k-l}^{(i)}$. Before we show that $\tilde{P}_{lk}^t(T) = P_{lk}^t(T)$, we collect some useful observations. First of all, note that the number of ways to distribute the individuals over the positions in C_{k-l} can be written as a product by first assigning a type to each individual and then distributing all individuals of a given type over the classes (writing C_{k-l} for $C_{k-l}(T)$ and N for $N(T)$):

$$D(C_{k-l}) = D(N) \prod_{j=1}^{\theta} D\left(\left(C_{k-l}^{(i,j)}\right)_{i=1}^{m}\right).$$

Secondly, note that $\mathbb{P}({}^tX_{lk} \in B_{k-l}^{(i)})$ is equal to

$$\sum_{(n_1,\ldots,n_\theta)\in\mathbb{N}^\theta} \mathbb{P}\left(W_l^t = (n_1,\ldots,n_\theta)\right) \mathbb{P}\left(\bigotimes_{j=1}^{\theta} \text{Multi}\left(n_j, ({}^jp_{l+1,k}^{(1)},\ldots,{}^jp_{l+1,k}^{(m)})\right) \in B_{k-l,i}\right)$$
$$= \mathbb{P}\left(\bigotimes_{j=1}^{\theta} \text{Multi}\left(W_{l,j}^t, ({}^jp_{l+1,k}^{(1)},\ldots,{}^jp_{l+1,k}^{(m)})\right) \in B_{k-l,i}\right)$$

and by (3.1) this is exactly ${}^tp_{lk}^{(i)}$. Next, since $C_{k-l}(T)$ determines $N(T)$, we have:

$$\begin{aligned}
\mathbb{P}({}^tX_{lk} = C_{k-l}) &= \mathbb{P}\left({}^tX_{lk} = C_{k-l}, W_l^t = N\right) \\
&= \mathbb{P}\left(W_l^t = N\right) \mathbb{P}\left({}^tX_{lk} = C_{k-l} \mid W_l^t = N\right) \\
&= \mathbb{P}\left(W_l^t = N\right) \prod_{j=1}^{\theta} \mathbb{P}\left(\left({}^tX_{lk}^{(i,j)}\right)_{i=1}^{m} = \left(C_{k-l}^{(i,j)}\right)_{i=1}^{m} \mid W_{l,j}^t = N_j\right) \\
&= \mathbb{P}\left(W_l^t = N\right) \prod_{j=1}^{\theta} D\left(\left(C_{k-l}^{(i,j)}\right)_{i=1}^{m}\right) \prod_{i=1}^{m} \left({}^jp_{l+1,k}^{(i)}\right)^{C_{k-l}^{(i,j)}}
\end{aligned}$$

$$= \frac{\mathbb{P}\left(W_l^t = N\right) D(C_{k-l})}{D(N)} \prod_{j=1}^{\theta} \prod_{i=1}^{m} \left({}^{j}p_{l+1,k}^{(i)} \right)^{C_{k-l}^{(i,j)}}.$$

Combining these observations gives

$$\begin{aligned}
\tilde{P}_{lk}^{t}(T) &= {}^{t}p_{lk}^{(i)}\, {}^{t}\tilde{Q}_{lk}^{(i)}(T) \\
&= {}^{t}p_{lk}^{(i)} \frac{\mathbb{P}({}^{t}X_{lk} = C_{k-l}(T) \mid {}^{t}X_{lk} \in B_{k-l}^{(i)})}{D(C_{k-l}(T))} \prod_{j=1}^{m} \prod_{\tilde{T} \prec T : \tilde{T} \in A_{k-l-1}^{(j)}} {}^{r(\tilde{T})}\tilde{Q}_{l+1,k}^{(j)}(\tilde{T}) \\
&= {}^{t}p_{lk}^{(i)} \frac{\mathbb{P}({}^{t}X_{lk} = C_{k-l}(T))}{D(C_{k-l}(T)) \mathbb{P}({}^{t}X_{lk} \in B_{k-l}^{(i)})} \prod_{j=1}^{m} \prod_{\tilde{t} \in \Theta} \prod_{\tilde{T} \prec T : \tilde{T} \in A_{k-l-1}^{(j)} \cap \mathcal{T}^{\tilde{t}}} {}^{\tilde{t}}\tilde{Q}_{l+1,k}^{(j)}(\tilde{T}) \\
&= \left(\frac{\mathbb{P}\left(W_l^t = N\right)}{D(N)} \prod_{\tilde{t}=1}^{\theta} \prod_{j=1}^{m} \left({}^{\tilde{t}}p_{l+1,k}^{(j)} \right)^{C_{k-l}^{(j,\tilde{t})}} \right) \left(\prod_{j=1}^{m} \prod_{\tilde{t} \in \Theta} \prod_{\tilde{T} \prec T : \tilde{T} \in A_{k-l-1}^{(j)} \cap \mathcal{T}^{\tilde{t}}} {}^{\tilde{t}}\tilde{Q}_{l+1,k}^{(j)}(\tilde{T}) \right) \\
&= \frac{\mathbb{P}\left(W_l^t = N\right)}{D(N)} \prod_{j=1}^{m} \prod_{\tilde{t} \in \Theta} \prod_{\tilde{T} \prec T : \tilde{T} \in A_{k-l-1}^{(j)} \cap \mathcal{T}^{\tilde{t}}} {}^{\tilde{t}}p_{l+1,k}^{(j)}\, {}^{\tilde{t}}\tilde{Q}_{l+1,k}^{(j)}(\tilde{T}) \\
&= \frac{\mathbb{P}\left(W_l^t = N\right)}{D(N)} \prod_{j=1}^{m} \prod_{\tilde{t} \in \Theta} \prod_{\tilde{T} \prec T : \tilde{T} \in A_{k-l-1}^{(j)} \cap \mathcal{T}^{\tilde{t}}} \tilde{P}_{l+1,k}^{\tilde{t}}(\tilde{T}) \\
&= \frac{\mathbb{P}\left(W_l^t = N\right)}{D(N)} \prod_{\tilde{T} \prec T} P_{l+1,k}^{r(\tilde{T})}(\tilde{T}) \\
&= P_{lk}^{t}(T).
\end{aligned}$$

□

5 Conclusion

We have demonstrated how to condition multi-type Galton-Watson trees on events having some recursive nature. More specifically, we looked at partitions of the set of trees in which each partition set is defined by some tree property. A crucial aspect of these properties is that they are determined completely by the types of children of the tree and the partition sets to which they belong. As our examples show, there is a wide variety of events fitting into this framework.

We have shown that such a conditioned tree itself is again a multi-type Galton-Watson tree, and we derived equations for the type- and level-dependent offspring distribution. These results can for example be used to compute probabilities on properties related to the events we condition on. Also, using our explicit construction procedure we can directly generate a tree that is conditioned to satisfy some property that has very low probability, which should also be useful for simulation purposes.

References

[1] E.A. Cator, H. Don – *Constructing and searching conditioned Galton-Watson trees.* arXiv:1412.5890.

[2] J. Geiger, G. Kersting – *The Galton-Watson tree conditioned on its height.* In: Probability Theory and Mathematical Statistics (Vilnius, 1998). (TEV, Vilnius, 1999), 277-286.

[3] J. Geiger – *Elementary new proofs of classical limit theorems for Galton-Watson processes.* J. Appl. Probab. 36 (1999), 301-309.

[4] S. Janson – *Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation.* Probability Surveys, Vol. 9 (2012), 103-252.

[5] Lyons, R. with Peres, Y. – *Probability on Trees and Networks.* Cambridge University Press. In preparation (2015). Current version available at `http://pages.iu.edu/~rdlyons/`.