University of Bath

Alternative formats
If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

And/Or trees: A local limit point of view

Nicolas Broutin* and Cécile Mailler†

June 8, 2017

Abstract

We present here a new and universal approach for the study of random and/or trees, unifying in one framework many different models, including some novel ones not yet understood in the literature. An and/or tree is a Boolean expression represented in (one of) its tree shapes. Fix an integer k, take a sequence of random (rooted) trees of increasing size, say $(t_n)_{n \geqslant 1}$, and label each of these random trees uniformly at random in order to get a random Boolean expression on k variables.

We prove that, under rather weak local conditions on the sequence of random trees $(t_n)_{n \geqslant 1}$, the distribution induced on Boolean functions by this procedure converges as n tends to infinity. In particular, we characterise two different behaviours of this limit distribution depending on the shape of the local limit of $(t_n)_{n \geqslant 1}$: a *degenerate* case when the local limit has no leaves; and a non-degenerate case, which we are able to describe in more details under stronger conditions. In this latter case, we provide a relationship between the probability of a given Boolean function and its *complexity*.

The examples covered by this unified framework include trees that interpolate between models with logarithmic typical distances (such as random binary search trees) and other ones with square root typical distances (such as conditioned Galton–Watson trees).

Keywords: random trees, local limit, and/or trees, random Boolean functions.

1 Introduction

The problem of generating a *complex* random Boolean function and understanding its typical properties can be traced back to the pioneering work of Riordan and Shannon [38], in which the authors studied uniformly random k-variable Boolean functions (for large integer k). However, the uniform distribution is only natural if one represents the functions by a truth table (assigning a uniformly random value to every possible entry vector). Another way of representing a Boolean function is by Boolean *expressions*, and significant efforts have been made towards defining probability distributions on Boolean functions via of their representation by Boolean expressions. For a general introduction to such questions and related problems, we refer the reader to the survey article by Gardy [19] and the references therein.

The concept of and/or trees arises as a natural representation of a Boolean expression, and the many different standard distributions of random trees can be used to sample random Boolean expressions. More precisely, an expression is equivalent to a tree whose internal nodes are labelled by the Boolean connectors 'and' ($\wedge$) or 'or' ($\vee$), and whose leaves are labelled by literals among $x_1, \bar{x}_1, \ldots, x_k, \bar{x}_k$, for some integer $k \geqslant 1$, where $\bar{x}$ denotes the negation of x (see Figure 1). We call such trees *and/or trees*. The origins of this line of research may be traced back to the works of Woods [39]. Amongst other things, he proved the existence of a limit probability distribution for Boolean function represented by sequences of trees of increasing sizes, and conjectured that there might be a relationship between the *complexity* of a function and the probability that it is sampled.

Woods [39] considered functions represented by uniformly random Cayley trees (general rooted trees on $[n] = \{1, 2, \ldots, n\}$). The case of Boolean functions encoded by uniformly random binary and/or trees, called the Catalan tree model, was first studied by Lefmann and Savický [27]. Both papers proved the existence of a natural probability distribution on k-variable Boolean functions which is the weak limit of the probability induced by the uniform random trees on $[n]$ and by the uniform binary and/or trees of size n, respectively. Since the results are similar, for the sake of presentation and only in this introduction, we focus on the case of binary trees in [27]: for a function f in k Boolean variables, the corresponding limit probability is denoted by $\mathbb{P}_k^{\mathrm{uni}}(f)$. Lefmann and Savický [27] also obtain bounds on

---

*Inria Paris-Rocquencourt, Domaine de Voluceau, 78153 Le Chesnay, France and NYU Shanghai. `nicolas.broutin@inria.fr`.

†Department of Mathematical Sciences, University of Bath, Claverton Down, BA2 7AY Bath, UK. `c.mailler@bath.ac.uk`. This author would like to thank EPSRC for support through the grant EP/K016075/1.

$\mathbb{P}_k^{\text{uni}}(f)$ in terms of the complexity $L(f)$ of the function f (being the minimal number of *leaves* of a tree representing f) of the kind suggested by Woods [39]. More precisely, they prove that, for all $k \geqslant 2$,

$$\exp(-c_1 L(f) \log k) \leqslant \mathbb{P}_k^{\text{uni}}(f) \leqslant \exp(-c_2 L(f) k^{-3}), \tag{1}$$

for two constants $c_1, c_2 > 0$. The results in [27] were then reproduced and slightly improved by Chauvin et al. [7] (who replace the k^{-3} in the right-hand side of (1) by k^{-2}). These bounds are the only results in the literature that hold for fixed k; as a side remark, we will show in this article how to improve further the upper bound by a very simple symmetry argument. This improvement is significant for two reasons: the new upper bound constrains the probability of functions with complexity $o(k^2)$, and we will show using a class of functions called *read-once* that it is sharp.

More precise bounds seem hard to obtain without considering the limit as k tends to infinity. The first result in this direction was by Kozik [26] who showed that for all integer k_0, for all k_0-variable Boolean function f,

$$\mathbb{P}_k^{\text{uni}}(f) = \Theta\left(\frac{1}{k^{L(f)+1}}\right) \text{ when } k \to +\infty, \tag{2}$$

where the constants involved in the Θ-term depend on f.

More recently, it was proved by Chauvin et al. [8] that if one replaces the uniformly random binary trees underlying the distribution defined in Lefmann and Savický [27] by random binary search trees (see, e.g., Knuth [25]), then the behaviour of the family of distributions induced on Boolean function is radically different: indeed, writing $\mathbb{P}_k^{\text{bst},n}$ for the probability induced on k-variable Boolean functions by trees of size n, one has

$$\mathbb{P}_k^{\text{bst},n}(\{\text{True}, \text{False}\}) \to 1, \text{ as } n \to +\infty,$$

for all integer k, where True and False are the two constant functions: $((x_1, \ldots, x_k) \mapsto \text{True})$ and $((x_1, \ldots, x_k) \mapsto \text{False})$; we say that this distribution is *degenerate*.

Although slight generalisations were considered in the literature (see for example Genitrini et al. [20]), essentially only the two models of random and/or trees described above (uniform binary tree and random binary search tree) have been studied. However, there is no a priori reason why one should choose the underlying tree to be binary (as already mentioned in Genitrini et al. [20], the conjunction and disjunction connectors are associative), nor any reason that justifies the uniform or the random binary search tree distributions, apart of course from the fact that one can do some explicit computations in these cases. This is why we initiate in this paper a more general approach that is independent of the underlying family of trees.

Our aim is to place the previous studies in a common framework by introducing a family of distributions on Boolean functions defined as weak limits of distributions coming from tree representations. This family will include most of the and/or tree models studied in the literature, as well as some models which behaviour was unknown up to now. This paper is the first attempt at describing this family of probability distributions in greater generality. In particular, by taking a local limit point of view, we generalize and greatly simplify the proofs of previously known results. We extract essential properties that are needed to prove weak convergence of the probability distributions induced by a sequence of random trees, and to obtain bounds on the probability of a given Boolean function in terms of its complexity.

The main insight given by this common framework is that the key property of the underlying sequence of random trees of increasing sizes is its local limit as opposed to its scaling limit: on the one hand, a local limit with no leaf *near the root* will induce a distribution on Boolean functions that is degenerate or concentrated on the two constant functions True and False (the case of random binary search tree of [8] is a typical example); on the other hand, a sequence locally converging to an infinite spine with reasonably small sub-trees hanging from it will verify Equation (2). As corollaries, we obtain new proofs of the results of Lefmann and Savický [27], Chauvin et al. [7] and Kozik [26] discussed above.

Furthermore, we enlarge the family of Galton–Watson trees for which one can obtain such results and also consider Ford's alpha model [16], a model that can be seen as interpolating between the Galton–Watson family and the binary search tree, which has not up till now been considered in the context of and/or trees.

2 Description of the framework and main results

2.1 Boolean trees and Boolean functions

In this paper, for all integer k, a k-variable **Boolean tree** (or an and/or tree), is a rooted finite tree having no node of arity one (i.e. having exactly one child), whose internal nodes are labelled by the disjunction connective $\vee$ or by the conjunction connective $\wedge$, and whose leaves are labelled by literals taken in a finite fixed set $\{x_1, \bar{x}_1, \ldots, x_k, \bar{x}_k\}$, where $\bar{x}_i$ stands for the negation of the variable x_i.
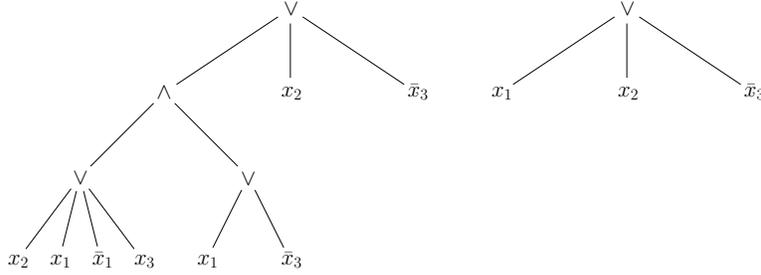
Figure 1: *Both this and/or trees represent the Boolean function $(x_1, x_2, x_3) \mapsto x_1 \vee x_2 \vee \bar{x}_3$. The left tree has size 8. The right tree has size 3 and is a minimal tree of the Boolean function $(x_1, x_2, x_3) \mapsto x_1 \vee x_2 \vee \bar{x}_3$, which thus has complexity 3.*

The set of and/or Boolean expressions can be defined recursively as follows: a Boolean expression is either a literal or its negation, or the conjunction of at least two Boolean expressions, or the disjunction of at least two Boolean expressions. Any Boolean tree is equivalent to a Boolean expression.

Recall that for all integer k, a k-variable **Boolean function** is a mapping from $\{\text{False}, \text{True}\}^k$ onto $\{\text{False}, \text{True}\}$. For every assignment of the variables $(x_1, \ldots, x_k) \in \{\text{False}, \text{True}\}^k$, one can apply Boolean algebra rules from the leaves of a Boolean tree up to its root, and the value obtained at the root can be defined as the image of $(x_1, \ldots, x_k)$; thus, every Boolean tree represents a Boolean function (but several trees can represent the same Boolean function as seen in Figure 1). We denote by $f[\tau]$ the Boolean function represented by the Boolean tree τ.

The **size** of a tree t, denoted by $\|t\|$, is the number of its leaves. The **complexity** of a non-constant Boolean function f, denoted by $L(f)$, is the size of the smallest trees that represent it, which we call **minimal trees** of f (see Figure 1 for an example). The complexity of the two constant Boolean functions, denoted respectively by True and False, is 0.

For any integer k, given a rooted tree t having no node of arity one, we define its **randomly k-labelled version $\hat{t}$** as follows (note that in order to keep notations simple the dependence in k is not explicit in this notation): each internal node of t chooses a label in $\{\wedge, \vee\}$ uniformly at random, each leaf of t chooses a label in $\{x_1, \bar{x}_1, \ldots, x_k, \bar{x}_k\}$ uniformly at random, independently from each other. For a Boolean function g, we denote by $P_k[t](g)$ the probability that the randomly k-labelled version of t represents g (in other words, $P_k[t](g)$ is the probability that $f[\hat{t}]$ equals g).

Except if mentioned otherwise, all trees considered in this article are assumed to contain no node of arity one. This assumption is natural in the context of and/or trees since the two logical connectives $\wedge$ and $\vee$ are binary operators. Note that thanks to associativity, they can also be considered as r-ary operators for all $r \geqslant 2$, which is why we do not restrict ourselves to binary trees, as sometimes done in the literature (see for example [7, 26, 27]).

2.2 The local topology, infinite trees, and continuity results

One of the main goals of the paper is to properly extend the above definition of a distribution $P_k[t]$ on the set of Boolean functions to (a certain class of) infinite trees. Our approach relies on approximations of the infinite trees by sequences of growing trees in the local topology around the root, which we now introduce.

For a (rooted) tree t we define the truncation at height $h \geqslant 0$ as the subtree induced on the nodes at distance at most h from the root, and denote it by t^h. A sequence of rooted trees $(t_n)_{n \geqslant 1}$ is said to **converge locally** to a tree t if for every integer $h \geqslant 1$ there exists n_h large enough that for all $n \geqslant n_h$, the truncations t_n^h are isomorphic to t^h. (Note that the limit tree t does not have to be infinite.) For a tree t, we define its number of **ends** as the number of disjoint paths to infinity (more precisely, the limit of the number of connected components of the forest induced by t on the set of nodes at distance at least $h \geqslant 1$, as $h \to \infty$; an end may also be defined as an equivalence class of infinite paths, where two paths are equivalent if their symmetric difference is finite). When an infinite tree t has a single end, the unique infinite simple path is called the **spine**.

The idea is to identify the distribution on Boolean functions encoded by an infinite tree as the limit distribution of the functions encoded by approximating sequences of growing trees. So given a sequence of trees $(t_n)_{n \geqslant 1}$ locally converging to a tree t, one is led to showing that the Boolean function $f[\hat{t}_n]$ converges in distribution, as n tends to infinity, to a limit that only depends on t. The two following continuity theorems prove that this is the case when the limiting tree has no leaf or when it has finitely many ends. Examples of such sequences of trees will be given later.

3

TREES WITHOUT LEAVES. Our first theorem deals with trees without leaves, that yield degenerate distributions on Boolean functions:

Theorem 2.1. *(a) Suppose that $(t_{k,n})_{n \geqslant 1}$ is a sequence of trees converging locally to an infinite tree without leaves. Then, as $n \to \infty$,*

$$P_k[t_{k,n}](\text{True}) = P_k[t_{k,n}](\text{False}) \to \frac{1}{2}.$$

(b) Conversely, if $(t_{k,n})_{n \geqslant 1}$ converges locally to a tree with at least one leaf, then there exists a function $f \notin \{\text{True}, \text{False}\}$ such that

$$\liminf_{n \to \infty} P_k[t_{k,n}](f) > 0.$$

Theorem 2.1 implies the following straightforward corollary, which settles a conjecture of Chauvin et al. [9]. For $\sigma \in \mathbb{N}$, let $\mathbb{T}_\sigma$ denote the collection of rooted trees such that the leaf closest to the root lies at distance at least σ.

Corollary 2.2. *Let $(T_n)_{n \geqslant 1}$ be a sequence of random trees. Then*

$$\mathbb{P}(\text{f}[\hat{T}_n] \in \{\text{True}, \text{False}\}) \to 1 \quad \textit{if and only if} \quad \inf_{\sigma \geqslant 1} \lim_{n \to \infty} \mathbb{P}(T_n \in \mathbb{T}_\sigma) = 1$$

Some special cases of Corollary 2.2 have already been proved in the literature. If, for all $n \geqslant 0$, T_n is almost surely equal to the balanced binary tree of height n (i.e. the unique binary tree which has 2^n leaves, all lying at height n), then, it is proven by Fournier et al. [17] that $\mathbb{P}(\text{f}[\hat{T}_n] = \{\text{True}, \text{False}\}) \to 1$.

The case when T_n is a random binary search tree of size n, is treated by Chauvin et al. [8]. Recall that a binary search tree of size n is the rooted binary tree constructed as follows: Given a list of distinct real numbers $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$, the root is labelled with π_1, and the tree is recursively obtained by repeating the construction with the lists contaning the π_i that are smaller and larger than π_1 for the left and right subtrees of the root, respectively. The random binary search tree is then the binary search tree obtained when π is a sequence of i.i.d. random variables, for example uniformly distributed in $(0, 1)$ (see e.g. [25]). Chauvin et al. [8]'s result is a direct consequence of Corollary 2.2 since Devroye [13] proved that the *fill-up* or *saturation level* (the height of the leaf closest to the root) of the random binary search tree of size n is asymptotically equivalent to $c \log n$ in probability, where $c = .3733\ldots$ is the unique solution smaller than one of $c \log((2e)/c) = 1$ (see also [36]). The same holds when the underlying tree shape is any of the classical random search trees based on the divide-and-conquer paradigm, for instance quad-trees or k-d trees built from uniformly random point sets, tries [29], or more generally any example that fits in the framework of [6].

TREES WITH FINITELY MANY ENDS. Our second continuity result concerns trees with finitely many ends for which the distribution on Boolean functions is non-degenerate. The first theorem below ensures convergence to a limiting probability distribution when n goes to infinity. The fact that this distribution is non-degenerate is the next main result, discussed in Section 2.3.

Theorem 2.3. *Fix $k \geqslant 1$. Suppose that $(t_{k,n})_{n \geqslant 1}$ is a sequence of trees which converges locally to t_k, and that t_k has only finitely many ends. Then, there exists a probability distribution $P_k[t_k]$ such that for every Boolean function f one has, as $n \to \infty$,*

$$P_k[t_{k,n}](f) \to P_k[t_k](f).$$

This convergence result is relatively easy to apply to a whole range of examples: It is known since Grimmett [21] that the family trees of critical Galton–Watson trees conditioned on the total progeny of increasing sizes converge locally to an infinite tree with a single end (see also Aldous and Steele [3]), so that the convergence results of Lefmann and Savický [27] and Chauvin et al. [7] are straightforward consequences of Theorem 2.3. We also provide in Section 3.3 novel examples of applications to random unordered trees (see Marckert and Miermont [31]), and other random trees arising from fragmentation processes (see Haas and Miermont [23]).

2.3 Trees with a unique end: properties of the limit distribution.

In the non-degenerate case, we describe further the behaviour of the limit distribution. For technical reasons, we restrict ourselves to random trees whose local limit has a unique end, and under further (but reasonable) assumptions on the limit tree of the family $(T_n)_{n \geqslant 1}$, we are then able to prove the equivalent of Kozik [26]'s result (see Equation (2)). As in previous approaches by analytic combinatorics, the proof will take two steps: first estimate the probability of the two constant functions (True and False) and then deduce from it the probability of a general Boolean function. More precisely, we are able to derive the asymptotic leading term of $P_k[T_n](f)$ when k tends to infinity, for any k_0-variable Boolean function f (for any fixed integer k_0), in terms of its complexity.

Our main result in this direction needs further definitions before being properly stated, and we thus postpone its statement to Section 5 (cf. Theorem 5.4). However, it reads informally as follows: if the family $(T_n)_{n \geqslant 1}$ of random trees converges locally to an infinite spine on which are hung some i.i.d. forests, and if we can reasonably control the size of these forests, then, for any integer k_0, for any k_0-variable Boolean function f, asymptotically as $k \to +\infty$,

$$\lim_{n \to +\infty} P_k[T_n](f) = \Theta\left(\frac{1}{k^{L(f)+1}}\right).$$

We apply this result to two main examples: critical Galton–Watson trees and Ford's alpha tree: the following two theorems are proved in Sections 5.4.1 and 5.4.2.

Theorem 2.4. *Let* $(T_n)_{n \geqslant 1}$ *be a critical Galton–Watson process of offspring distribution ξ for which there exists a constant $a > 0$ such that $\mathbb{E}[e^{a\xi}] < \infty$, and such that $\mathbb{P}(\xi = 1) = 0$. Then, for any integer k_0, for any k_0-variable Boolean function f, asymptotically as $k \to +\infty$,*

$$\lim_{n \to +\infty} P_k[T_n](f) = \Theta\left(\frac{1}{k^{L(f)+1}}\right).$$

Note that the case of Catalan trees studied by Lefmann and Savický [27], Chauvin et al. [7] and Kozik [26], is a particular case of Theorem 2.4.

Ford's alpha model (see [16]) is defined as follows (cf. Figure 2): T_1^α has a unique node to which is linked the root-edge. To build T_{n+1}^α from T_n^α, weight each internal edge (the root-edge and any edge that is not linked to a leaf is an internal edge) by α and each external edge by $1 - \alpha$. Pick at random an edge with probability proportional to these weights, add an internal node in the middle of this edge and link a new leaf to this new internal node. We call T_n^α the alpha tree of size n. The model interpolates between the Catalan tree ($\alpha = 1/2$) and the random binary search tree ($\alpha = 0$). It has been known since [22] that the expected height of the n-leaf alpha-tree behaves as n^α, except for the extremal case $\alpha = 0$ since the random binary search tree's height behaves as $\ln n$ (see [13]). This model thus permits to explore a whole range of *tree shapes*. The following theorem shows that only the case $\alpha = 0$ corresponding to the random binary search tree induces a degenerate distribution on Boolean functions. Any other alpha tree verifies Equation (2):

Theorem 2.5. *Let* $(T_n^\alpha)_{n \geqslant 1}$ *be a sequence of alpha trees. Then, for all $\alpha \in (0, 1]$, for any integer k_0, for all k_0-variable Boolean function f, as $k \to +\infty$,*

$$\lim_{n \to +\infty} P_k[T_n^\alpha](f) = \Theta\left(\frac{1}{k^{L(f)+1}}\right).$$

The behaviour of the distribution on Boolean functions that is induced by the alpha model was unknown up to now. Being able to prove Theorem 2.5 as a corollary of our main result is thus a significant step forward, when one compares it to the estimates that were available until now, that only concern very specific subcases.

Remark: In Theorem 2.4 and 2.5, as well as in Equation (2), the constants involved in the Θ-term depend on the function f.

2.4 Discussion and remarks

It is important to note that Theorem 2.3 is about deterministic sequences of trees. So, in the frameworks of the previous results where the trees are random, the limit distribution we construct is conditional on the limit tree. This type of results makes it hard to resort to arguments based on analytic combinatorics [15] and the Drmota–Lalley–Woods theorem giving asymptotics for generating functions satisfying a certain type of system of equations, since these techniques are very intimately related to counting problems. The techniques we use here are probabilistic, as in part of Chauvin et al. [7]. One of the drawbacks is that we cannot guarantee that the limit probability $P_k[T_\infty]$ charges every function in k variables, though it should certainly be the case.

We only consider the case of and/or trees, but one could of course think of other models of Boolean expressions. For instance, the case of Boolean expressions encoded by trees labelled by implications connectors have been treated by Fournier et al. [18].

Put together, Theorems 2.1 and 2.3 already give a pretty good idea of the properties of random Boolean expressions obtained by labelling large trees uniformly and independently of the tree. It would be interesting to study what happens when one deviates from this setting. One can probably relax the condition of uniformity without much harm, but the dependence of the labelling and the tree seems to be a more challenging obstacle. For example, our setting does not include non plane and/or trees as defined in Genitrini et al. [20], a model that takes into account the commutativity of the
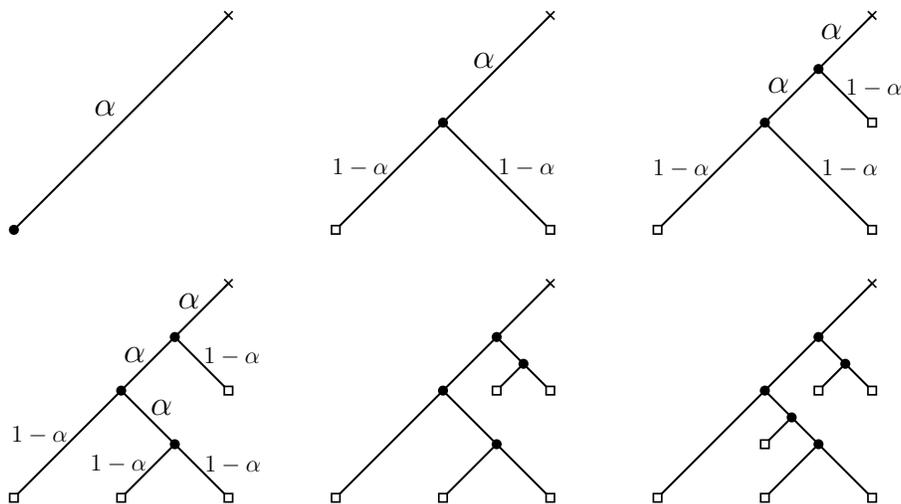
Figure 2: *A possible realisation of the first few steps of the construction of alpha trees. The alpha tree itself is obtained when removing the edge attached to the crossed node (called the root-edge) and the cross node itself. The internal nodes are represented by disks and leaves by squares.*

conjunction and disjunction connectives; this model is not covered by our framework since it cannot be described as the random uniform labelling of a random tree.

In the introduction, we evoked two different representations of a Boolean function: the truth table and Boolean expressions seen as trees. One could also think about Boolean functions that are not represented by trees but by circuits modelled by directed acyclic graphs with a single sink.

Finally, it would be interesting to look more precisely at what happens when the underlying limit tree has infinitely many ends as well as leaves. Note for example that Theorem 2.1 does not rule out the possibility that the limit Boolean function be a measurable function of the labelled limit tree (even when it has no leaves). We strongly believe that if one considers the growth of the number of ends which intersect the ball of radius d around the root (in the graph distance) as fixed, then for every small enough number of variables one should be able to define the limit Boolean function. Can one make such a claim more precise using for instance the branching number or the malthusian parameter in the case of Galton–Watson trees (as in Balogh et al. [4])?

Plan of the paper. Section 3 is devoted to the convergence to a limit probability distribution as n tends to infinity: it contains the proofs of Theorems 2.1 and 2.3 and provides some examples of families of trees for which these theorems apply. Section 4 is focusing on the Catalan tree case: we present simple arguments to tighten Inequality (1). In Section 5 we prove the analog of Kozik [26]'s result (Equation (2)) in our general setting. This stronger result only holds under further moment assumptions which we discuss by providing examples.

3 Continuity in the local topology

3.1 The degenerate case: proof of Theorem 2.1

Note that local convergence to an infinite tree with no leaves is equivalent to the divergence of the saturation level (being the height of the closest leaf to the root), and in this section we phrase Theorem 2.1 in this framework. For an integer $\sigma \geqslant 0$, we denote by $\mathbb{T}_\sigma$ the set of all rooted trees with a saturation level at least σ, so in particular $\mathbb{T}_0$ is the set of all rooted trees. Our proof of Theorem 2.1 consists in estimating the probability that the random Boolean functions assigns two different values to two (distinct) points $a, b \in \{0, 1\}^k$. (This is already the approach in [17] and [8].) Note that we use the canonical notation $0 = \text{False}$ and $1 = \text{True}$.

Let a and b be two distinct elements of $\{0, 1\}^k$, and let $\alpha, \beta \in \{0, 1\}$. Let us define the following probability (where the probability $\mathbb{P}$ refers to the uniform random k-labelling): for all trees t,

$$\mathbb{P}_t^{\alpha\beta}(a, b) := \mathbb{P}(\mathrm{f}[\hat{t}](a) = \alpha \text{ and } \mathrm{f}[\hat{t}](b) = \beta),$$

and the following supremum:
$$S_\sigma^{\alpha\beta}(a,b) = \sup\{\mathbb{P}_t^{\alpha,\beta}(a,b) : t \in \mathbb{T}_\sigma\}.$$

Lemma 3.1. *For every $k \geqslant 1$, there exists a constant $\sigma_0 \geqslant 1$ such that for all $\sigma \geqslant \sigma_0$, one has*

$$\sup_{a \neq b \in \{0,1\}^k} S_\sigma^{10}(a,b) = \sup_{a \neq b \in \{0,1\}^k} S_\sigma^{01}(a,b) \leqslant \frac{2}{\sigma}.$$

Proof. Fix any two distinct points $a, b \in \{0,1\}^k$, and, if there is no possible confusion, write $\mathbb{P}_t^{\alpha\beta}$ instead of $\mathbb{P}_t^{\alpha\beta}(a,b)$ and $S_\sigma^{\alpha\beta}$ instead of $S_\sigma^{\alpha\beta}(a,b)$. First of all, the symmetries of the labelling imply immediately that $\mathbb{P}_t^{10} = \mathbb{P}_t^{01}$ and $\mathbb{P}_t^{11} = \mathbb{P}_t^{00}$. Indeed, since the probabilities of $\wedge$ and $\vee$ are equal, and the probabilities of a variable x_i and its negation $\bar{x}_i$ are also equal, we can conclude that for any finite tree t and for any Boolean function f and its negation $\neg f$ we have

$$P_k[t](f) = P_k[t](\neg f).$$

Moreover, for every tree t, we have $\mathbb{P}_t^{10} + \mathbb{P}_t^{11} = \frac{1}{2}$, and thus, $S_\sigma^{\alpha\beta} \leqslant \frac{1}{2}$, for all $\alpha, \beta \in \{0,1\}$ and for all $\sigma \in \mathbb{N}$.

Let $t \in \mathbb{T}_\sigma$. Let $r \geqslant 2$ be the degree of the root of t (recall that, by assumption, t contains no node of arity one, see Section 2.1) and denote by $t_1, \ldots, t_r$ the subtrees rooted at the root's children. For all $i \in \{1, \ldots, r\}$, $t_i \in \mathbb{T}_{\sigma-1}$. Moreover, $f[\hat{t}](a) = 1$ and $f[\hat{t}](b) = 0$ imply that

- either the root is labelled by $\wedge$ and $f[\hat{t}_i](a) = 1$ for all $1 \leqslant i \leqslant r$, and at least some $f[\hat{t}_i](b) = 0$,

- or the root is labelled by $\vee$ and $f[\hat{t}_i](b) = 0$ for all $1 \leqslant i \leqslant r$, and at least some $f[\hat{t}_i](a) = 1$.

It follows that

$$\mathbb{P}_t^{10} = \frac{1}{2}\left(\prod_{i=1}^r (\mathbb{P}_{t_i}^{10} + \mathbb{P}_{t_i}^{11}) - \prod_{i=1}^r \mathbb{P}_{t_i}^{11}\right) + \frac{1}{2}\left(\prod_{i=1}^r (\mathbb{P}_{t_i}^{10} + \mathbb{P}_{t_i}^{00}) - \prod_{i=1}^r \mathbb{P}_{t_i}^{00}\right)$$

$$= \frac{1}{2^r} - \prod_{i=1}^r \mathbb{P}_{t_i}^{11} = \frac{1}{2^r} - \prod_{i=1}^r \left(\frac{1}{2} - \mathbb{P}_{t_i}^{10}\right).$$

Now, for all $i \in \{1, \ldots, r\}$, we have $\mathbb{P}_{t_i}^{10} \leqslant S_{\sigma-1}^{10}$, which implies that, for every tree $t \in \mathbb{T}_\sigma$ with a root of degree r,

$$\mathbb{P}_t^{10} \leqslant \frac{1}{2^r} - \left(\frac{1}{2} - S_{\sigma-1}^{10}\right)^r.$$

Thus,

$$S_\sigma^{10} = \sup_{t \in \mathbb{T}_\sigma} \mathbb{P}_t^{10} \leqslant \sup_{r \geqslant 2}\left\{\frac{1}{2^r} - \left(\frac{1}{2} - S_{\sigma-1}^{10}\right)^r\right\}.$$

One easily verifies that, for all $x \in [0, 1/2]$, one has

$$\sup_{r \geqslant 2}\left\{\frac{1}{2^r} - \left(\frac{1}{2} - x\right)^r\right\} = \frac{1}{4} - \left(\frac{1}{2} - x\right)^2,$$

which implies, since $S_\sigma^{10} \leqslant \frac{1}{2}$ for all integer σ,

$$S_\sigma^{10} \leqslant \frac{1}{4} - \left(\frac{1}{2} - S_{\sigma-1}^{10}\right)^2.$$

Let us define the sequence $(u_\sigma)_{\sigma \geqslant 1}$ as follows: $u_1 = S_1^{10}$ and $u_{\sigma+1} = g(u_\sigma)$, where $g(x) = 1/4 - (1/2 - x)^2 = x - x^2$. Since $S_\sigma^{10} \leqslant 1/2$ for every $\sigma \geqslant 1$, a straightforward induction on σ shows that, for all $\sigma \geqslant 1$, one then has $u_\sigma \geqslant S_\sigma^{10}$.

Note that 0 is the unique fixed point of g and that $[0,1]$ is stable by g. Moreover, $g'(0) = 1$ and $g''(0) = -2$. By a standard result about inductive sequences: as $\sigma \to +\infty$,

$$u_\sigma \sim -\frac{2}{\sigma g''(0)} = \frac{1}{\sigma}.$$

It follows that, for all $a \neq b$, there exists $\sigma_{a,b}$ such that, for all $\sigma \geqslant \sigma_{a,b}$, $S_\sigma^{10}(a,b) \leqslant 2/\sigma$. Let $\sigma_0 = \max_{a \neq b \in \{0,1\}^k} \sigma_{a,b}$ to conclude the proof. $\square$

With Lemma 3.1 under our belt, the proof of Theorem 2.1 (a) appears now as a straightforward application of the union bound.

Proof of Theorem 2.1 (a). Fix an arbitrary integer $\sigma \geqslant 1$. Then, by assumption, there exists $n_0 \geqslant 1$ large enough such that for all $n \geqslant n_0$, one has $t_n \in \mathbb{T}_\sigma$. Now, by definition of S_σ^{10}, for all $n \geqslant n_0 = n_0(\sigma)$ we have

$$
\begin{aligned}
\mathbb{P}(\mathrm{f}[\hat{t}_n] \notin \{\mathrm{True}, \mathrm{False}\}) &= \mathbb{P}(\exists a, b \in \{0,1\}^k : \mathrm{f}[\hat{t}_n](a) \neq \mathrm{f}[\hat{t}_n](b)) \\
&\leqslant \sum_{a \neq b} \mathbb{P}(\mathrm{f}[\hat{t}_n](a) \neq \mathrm{f}[\hat{t}_n](b)) \\
&\leqslant 2^{2k} \sup_{a \neq b} 2 S_\sigma^{10}(a,b).
\end{aligned}
$$

Now for any $\sigma \geqslant \sigma_0$ in Lemma 3.1, and all $n \geqslant n_0(\sigma)$, we have

$$
\mathbb{P}(\mathrm{f}[\hat{t}_n] \notin \{\mathrm{True}, \mathrm{False}\}) \leqslant \frac{2^{2k+2}}{\sigma}.
$$

Letting $\sigma \to \infty$ completes the proof. $\qquad\square$

We now move on to the lower bound in Theorem 2.1.

Lemma 3.2. *Let $\mathcal{A}_\sigma$ be the set of trees with saturation level equal to σ. Then, for every integer $\sigma \geqslant 1$,*

$$
I_\sigma := \inf_{t \in \mathcal{A}_\sigma} \mathbb{P}(\mathrm{f}[\hat{t}] \notin \{\mathrm{True}, \mathrm{False}\}) \geqslant \frac{1}{2^\sigma}.
$$

Proof. Let t be a tree in $\mathcal{A}_\sigma$ and consider the associated randomly labelled tree $\hat{t}$. Let us denote by x the label of ℓ, one of its leaves at height σ; by $\diamond_0, \ldots, \diamond_{\sigma-1}$ the connectives of the nodes between the root and the leaf ℓ ($\diamond_0$ being the label of the root and $\diamond_{\sigma-1}$ the label of the parent of ℓ), and by $g_0, \ldots, g_{\sigma-1}$ the random boolean functions calculated by the forests hanging along the path from the root to ℓ (from top to bottom). Therefore, the function calculated by $\hat{t}$ is given by

$$
\mathrm{f}[\hat{t}] = (g_0 \diamond_1 (g_2 \diamond_2 (\ldots (g_{\sigma-1} \diamond_{\sigma-1} x)))).
$$

Let j be the minimum $m \in \{0, \ldots, \sigma-1\}$ such that g_m depends on x; if such m does not exist, we let $j = \sigma$. Let $\rho = g_{j+1} \diamond_{j+1} (\ldots (g_{\sigma-1} \diamond_{\sigma-1} x))$. We can choose $\diamond_j$ in $\{\wedge, \vee\}$ such that $g_j \diamond_j \rho$ does depend on x: if $g_j \neq \mathrm{False}$, then $x \wedge g_j$ does depend on x and if $g_j = \mathrm{False}$, then $x \vee g_j = x$ does depend on x. By induction, we then can choose $\diamond_j, \ldots, \diamond_0$ such that $\mathrm{f}[\hat{t}]$ does depend on x, and the probability, conditionally on all the rest of the tree (and for any such conditioning), that $\diamond_0, \ldots, \diamond_j$ are actually equal to this choice is equal to $2^{-(j+1)}$.

To conclude, if we denote by $\mathrm{f}[\hat{t}](x=1)$ (resp. $\mathrm{f}[\hat{t}](x=0)$) the restriction of $\mathrm{f}[\hat{t}]$ to the subset of $\{0,1\}^n$ where $x = 1$ (resp. $x = 0$),

$$
\mathbb{P}(\mathrm{f}[\hat{t}](x=1) \neq \mathrm{f}[\hat{t}](x=0)) \geqslant \frac{1}{2^\sigma},
$$

and thus

$$
\mathbb{P}(\mathrm{f}[\hat{t}] \neq \mathrm{True} \text{ and } \mathrm{f}[\hat{t}] \neq \mathrm{False}) \geqslant \frac{1}{2^\sigma}.
$$

The last inequality holds for every $t \in \mathcal{A}_\sigma$: taking the infimum proves Lemma 3.2. $\qquad\square$

Proof of Theorem 2.1 (b). By assumption, the saturation level of $(t_n)_{n \geqslant 1}$ does not diverge, and there exists $\sigma \geqslant 1$ such that we can find an infinite subsequence $J \subseteq \mathbb{N}$ such that for all $n \in J$, $t_n \notin \mathbb{T}_\sigma$. Without loss of generality, we suppose now that $J = \mathbb{N}$. Then, for all $n \in \mathbb{N}$,

$$
\begin{aligned}
\mathbb{P}(\mathrm{f}[\hat{t}_n] \notin \{\mathrm{True}, \mathrm{False}\}) &= \sum_{m=0}^{\sigma-1} \sum_{t \in \mathcal{A}_m} \mathbb{P}(\mathrm{f}[\hat{t}_n] \notin \{\mathrm{True}, \mathrm{False}\}) \mathbf{1}_{\{t_n = t\}} \\
&\geqslant \sum_{m=0}^{\sigma-1} \sum_{t \in \mathcal{A}_m} I_m \mathbf{1}_{\{t_n = t\}},
\end{aligned}
$$

since $I_m = \inf_{t \in \mathcal{A}_m} \mathbb{P}(\mathrm{f}[\hat{t}] \notin \{\mathrm{True}, \mathrm{False}\})$. Therefore, since $t_n \in \mathcal{A}_m$ for some $m < \sigma$, we have by Lemma 3.2

$$
\mathbb{P}(\mathrm{f}[\hat{t}_n] \notin \{\mathrm{True}, \mathrm{False}\}) \geqslant \sum_{m=0}^{\sigma-1} \frac{1}{2^m} \sum_{t \in \mathcal{A}_m} \mathbf{1}_{\{t_n = t\}} \geqslant \frac{1}{2^{\sigma-1}},
$$

which proves the claim. $\qquad\square$

3.2 Finitely many ends: proof of Theorem 2.3

In this section, we prove Theorem 2.3. So far, the Boolean function associated to a labelled tree has only been defined for finite trees. One of the main ingredients of the proof of Theorem 2.3 is the following lemma, which proves that the Boolean function $f[\hat{t}]$ is also well-defined when t has a unique infinite path, which we refer to as the *spine*.

Given a tree t and an integer h, we denote by t^h the tree obtained from t by removing all nodes having height greater than h.

Lemma 3.3. *Let t be a locally finite tree with at most one end. Then there exists an (random) integer $h \geqslant 1$ such that for all $i \geqslant h$, $f[\hat{t}^i] = f[\hat{t}^h]$ almost surely. The Boolean function $f[\hat{t}]$ is then defined as $f[\hat{t}^h]$.*

Proof. Let us denote by $(u_i)_{i \geqslant 0}$ the sequence of nodes along the spine of t (starting from the root), and write $\diamond_i$ the label of u_i in $\hat{t}$. For convenience, we introduce another truncation of the infinite tree t: we let $t^{[h]}$ denote the subtree of t containing the root when the spine is cut between the nodes u_h and u_{h+1}; then $t^{[h]}$ is finite for every $h \geqslant 0$. Let $(t_{i,j})_{j \geqslant 1}$ denote the sequence of finite subtrees rooted at the children of u_i in an arbitrary order (that is, we except the tree rooted at u_{i+1}). Note that by assumption, for every $i \geqslant 1$, $(t_{i,j})_{j \geqslant 1}$ is actually a non-empty (because t contains no unary nodes by assumption) finite sequence of finite trees. Let $f_{i,j} := f[\hat{t}_{i,j}]$ (the Boolean function represented by the random labelling of the tree $t_{i,j}$), and note that $(f_{i,j})_{i,j \geqslant 1}$ is independent of the sequence of labels along the spine $(\diamond_i)_{i \geqslant 1}$. Say that two sequences $\mathbf{f}_i := (f_{i,j})_{j \geqslant 1}$ and $\mathbf{f}_m := (f_{m,j})_{j \geqslant 1}$ are equivalent if the collections of Boolean functions they are made of are identical (the multiplicities and ordering may be different); we then write $\mathbf{f}_i \sim \mathbf{f}_m$.

The proof consists in finding an integer $h \geqslant 1$ such that replacing the subtree of t rooted at u_h by *any* other tree, finite or not, does not affect the Boolean function computed by the truncations $\hat{t}^{[i]}$, for all $i \geqslant h$. It is then possible to safely define $f[\hat{t}]$ as $f[\hat{t}^{[h]}]$. Let $h(\hat{t})$ be the minimal height $h \geqslant 0$ such that for all $i \geqslant h$ one has $f[\hat{t}^{[i]}] = f[\hat{t}^{[h]}]$. In order to complete the proof, it suffices to show that $h(\hat{t})$ is almost surely finite.

Assume that there exists two integers $i < j$ such that $\diamond_i = \wedge$, $\diamond_j = \vee$, and $\mathbf{f}_i \sim \mathbf{f}_j$. Then, there exists $m \geqslant 1$ such that $f_{j,m} = f_{i,1}$. For all $x \in \{0,1\}^k$ such that $f_{i,1}(x) = 0$, we have $f[\hat{t}^{[m]}](x) = f[\hat{t}^{[i]}](x)$ for all $m \geqslant i$, and for all $x \in \{0,1\}^k$ such that $f_{i,1}(x) = f_{j,m}(x) = 1$, we have $f[\hat{t}^{[m]}](x) = f[\hat{t}^{[j]}](x)$ for all $m \geqslant j$. Thus, for all $x \in \{0,1\}^k$, for all $h \geqslant j$, $f[\hat{t}^{[m]}](x) = f[\hat{t}^{[j]}](x)$. The same result holds if $\diamond_i = \vee$ and $\diamond_j = \wedge$; we define $\bullet_i = \wedge$ (resp. $\vee$) if $\diamond_i = \vee$ (resp. $\wedge$). We have shown that if there exist two integers $i < j$ such that $\diamond_j = \bullet_i$ and $\mathbf{f}_i \sim \mathbf{f}_j$, then $f[\hat{t}^{[j]}] = f[\hat{t}^{[m]}]$ for every $m \geqslant j$. If this occurs, $h(\hat{t}) \leqslant j$, and we are now looking for such a pair (i,j) of integers, that we call *good* hereafter.

Invoking the pigeon hole principle, we know that among the $2^{2^{2^k}} + 1 =: s_k$ first sequences in $(\mathbf{f}_i)_{i \geqslant 1}$ at least two are equivalent: almost surely, there exists $i < j \leqslant s_k$ such that $\mathbf{f}_i \sim \mathbf{f}_j$. Thus, since $(\mathbf{f}_i)_{i \geqslant 1}$ is independent of $(\diamond_i)_{i \geqslant 1}$ and the latter is an i.i.d. sequence, with probability $\frac{1}{2}$ we have $\diamond_j = \bullet_i$ and (i,j) is a good pair. If $\diamond_j \neq \bullet_i$, we look at the next s_k sequences of $(\mathbf{f}_i)_{i \geqslant 1}$, find two equivalent sequences by the pigeon hole principle, and the corresponding indices happen to be a good pair with probability $\frac{1}{2}$. Continuing in this manner, we see that the number of groups of size s_k one has to look at before finding a good pair follows a geometric distribution of parameter $1/2$, so that $h(\hat{t})$ is almost surely finite and the proof is complete. $\square$

This proof contains a first cutting algorithm of an infinite randomly labelled Boolean tree $\hat{t}$, which is far from optimal, but sufficient to prove the continuity of $f[\cdot]$ in the local topology. The infinite Boolean tree can certainly be simplified further and we introduce a refined trimming algorithm later on.

If, instead of having a single end, the tree t has finitely many ends, Lemma 3.3 still holds and the proof remains the same: one has to continue the cutting algorithm as long as an end remains. Or equivalently, one just uses the previous algorithm for the portions of he tree below height d, where the ends have all been separated. This is straightforward and we omit a formal proof of the following lemma:

Lemma 3.4. *Let t be a tree with finitely many ends. Then there exists a random integer $h \geqslant 1$ such that for all $i \geqslant h$, $f[\hat{t}^i] = f[\hat{t}^h]$ almost surely. The Boolean function $f[\hat{t}]$ is then defined as $f[\hat{t}^h]$.*

Lemma 3.5. *Let $(t_n)_{n \geqslant 1}$ a sequence of unlabelled trees converging locally to a tree t_∞ having finitely many spines. Then, there exists (random) integers $h \geqslant 1$ and $n_0 \geqslant 1$ such that for all $i \geqslant h$, and all $n \geqslant n_0$,*

$$f[\hat{t}_n^i] \stackrel{(law)}{=} f[\hat{t}_\infty^i] = f[\hat{t}_\infty].$$

Proof. Lemma 3.4 tells us that there exists almost surely an integer $h \geqslant 1$ such that, for all $k \geqslant h$, $f[\hat{t}_\infty^k] = f[\hat{t}_\infty^h]$, the latter serving as a definition for $f[\hat{t}_\infty]$. On the other hand, for this value of h, since $t_n \to t$ in the local topology, there exists a integer $n_0 \geqslant 1$ such that, for all $n \geqslant n_0$, $t_n^h = t_\infty^h$, which implies that $f[\hat{t}_n^h]$ and $f[\hat{t}_\infty^h]$ have the same distribution.

In other words, there exists almost surely an integer $h \geqslant 1$ such that, for all $i \geqslant h$, there exists an integer $n_i \geqslant 1$ such that, for all $n \geqslant n_i$,

$$\mathrm{f}[\hat{t}_n^i] \stackrel{(law)}{=} \mathrm{f}[\hat{t}_\infty^i] = \mathrm{f}[\hat{t}_\infty],$$

which proves the claim. $\qquad\square$

The following result is a direct consequence of Lemma 3.5:

Theorem 3.6. *Let $(T_{k,n})_{n \geqslant 1}$ a sequence of unlabelled random trees converging in distribution to a local limit T_∞ having a finite number of infinite branches with probablity one. For any k-variable Boolean function f, let us denote by $\mathbb{P}_{n,k}(f)$ the probability that $\mathrm{f}[\hat{T}_n] = f$. Asymptotically as n tends to infinity, the sequence $(\mathbb{P}_{n,k})_{n \geqslant 1}$ converges to an asymptotic probability distribution $\mathbb{P}_k$ such that, for every Boolean function f, $\mathbb{P}_k(f) = \mathbb{P}(\mathrm{f}[\hat{T}_\infty] = f)$.*

Most natural examples have finitely many ends, but it is natural to ask whether this condition is necessary. Theorem 2.1 treats the case when the limit tree has no leaves and possibly infinitely many ends. One could ask what could be said in a less extreme case, when there are infinitely many ends as well as leaves in the limit tree. This question remains open.

3.3 A few natural examples

To prove the existence of a limit probability distribution, it is enough to prove the local weak convergence towards a limit tree that has finitely many ends. We provide in this section examples of sequences of random trees that have finitely many ends, and thus to which Theorem 2.3 can be applied. In three of these examples, namely the conditioned critical Galton–Watson trees, the non-plane binary trees, and the fragmentation trees, the sequence of trees $(t_{n,k})_{n \geqslant 0}$ and its local limit t_k do not depend on k, the number of variables. For the so-called *associative tree*, treated in Section 3.3.3, which is a natural example from the literature, the sequence of trees and its limit both depend on the number of variables.

3.3.1 Conditioned critical Galton–Watson trees

Let ξ be an integer-valued random variable. The Galton–Watson tree of progeny distribution ξ is a random rooted tree in which every node has a number of children that is an independent copy of ξ.

Assume that the progeny distribution verifies that $\mathbb{E}[\xi] = 1$, and that $p_1 = \mathbb{P}(\xi = 1) = 0$ (recall that all trees considered in this article have no unary nodes). We let GW_ξ denote the distribution of a Galton–Watson tree with reproduction distribution ξ. Let T_n be a random Galton–Watson tree with progeny distribution ξ, conditioned on the total population being n (if such a size is possible). Then, it is well-known that T_n converges locally in distribution to an infinite tree T_∞ described as follows (see, e.g., [1, 24, 28]). Let $\hat{\xi}$ be the size-biased distribution associated to ξ defined by $\mathbb{P}(\hat{\xi} = i) = i\mathbb{P}(\xi = i)$, $i \geqslant 0$. Let $(\hat{\xi}_i)_{i \geqslant 0}$ be a sequence of i.i.d. copies of $\hat{\xi}$. Then there exists a unique self-avoiding infinite path in T_∞, consisting of the nodes $(u_i)_{i \geqslant 0}$; and for every $i \geqslant 0$, u_i has $\hat{\xi}_i \geqslant 1$ children, one of each is u_{i+1} and the others are the roots of i.i.d. copies of unconditioned GW_ξ random trees. (See Figure 3.) The random trees discussed in [27] and [7] correspond to the special case where ξ is such that $\mathbb{P}(\xi = 2) = \mathbb{P}(\xi = 0) = 1/2$. The results of Woods [39] on Cayley trees (uniformly random labelled trees) also fit in the framework since the shapes of a Cayley tree of size n and of a Galton–Watson tree with Poisson(1) offspring conditioned on the total progeny to be n have the same distribution. (Note however, that the arguments in [7] could be probably be extended to critical Galton–Watson trees with an offspring distribution ξ having exponential moments; here such an assumption on moment is not necessary.)

3.3.2 Non-plane binary trees

A rooted non-plane binary unlabelled tree is either a single external node, or it consists of an unordered pair of such trees. These trees can be seen as the equivalence classes of the usual (plane) binary trees where two trees are deemed equivalent if it is possible to transform one tree into the other by swapping the left and right children of a finite collection of nodes. These trees originate in the work of Pólya [37], and have been enumerated by Otter [33]; in the following we refer to them simply as *unordered trees*. They are different from the conditioned Galton–Watson trees of the previous example in an essential way, and in particular they lack the nice probabilistic representation as a branching process [5, 23, 31]. Let y_n denote the number of rooted binary unordered trees with n labelled leaves. Then, Otter [33] proved that

$$y_n = \kappa \rho^{-n} n^{-3/2}(1 + O(1/n)), \tag{3}$$

for some constants $\kappa > 0$ and $\rho \in (1/4, 1/2)$. Note in particular that there is a constant $C > 0$ such that $y_n \leqslant C\rho^{-n}n^{-3/2}$ for all $n \geqslant 0$. We would like to prove that if T_n is a sequence of uniformly random binary unordered trees on n leaves,
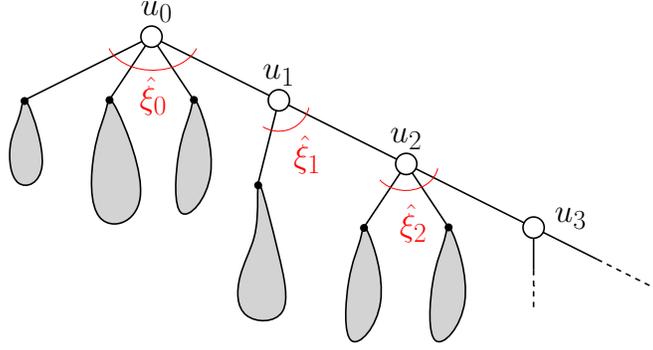
Figure 3: *The critical Galton–Watson tree of reproduction random variable ξ conditioned to be infinite: the $\hat{\xi}_i$ are independent copies of the size biased version of ξ, and the grey subtrees are independent Galton–Watson trees of reproduction random variable ξ.*

then T_n converges in distribution in the local sense to an infinite tree with a single infinite path. To prove this, we consider $\lambda_1^n \geqslant \lambda_2^n$ the sizes of the two subtrees of the root. For all $i < n/2$, $\lambda_2^n = i$ implies that $\lambda_1^n > \lambda_2^n$ and therefore, there cannot be any symmetry that involves the root, implying that

$$\mathbb{P}\left(\lambda_2^n = i\right) = \frac{y_{n-i} \cdot y_i}{y_n} = \rho^i y_i + O(i/n). \tag{4}$$

In particular, we have, for all $j < n/2$,

$$\mathbb{P}\left(\lambda_2^n \geqslant j\right) \leqslant \sum_{i \geqslant j}^{n/2} \frac{y_{n-i} \cdot y_i}{y_n} \leqslant \frac{C^2}{y_n} \sum_{i \geqslant j}^{n/2} \rho^{-n}(n-i)^{-3/2} i^{-3/2} \leqslant A j^{-1/2},$$

for a constant A and all n large enough. This implies that $(\lambda_2^n)_{n \geqslant 1}$ is tight, so that by (4), it converges in distribution to a (real) random variable, say X.

Let $(X_i)_{i \geqslant 0}$ be a sequence of i.i.d. copies of X, and conditional on that, let $(Y_i)_{i \geqslant 0}$ denote a sequence of independent random rooted binary unordered trees of respective sizes X_i. Finally, let T_∞ be the binary tree consisting of a single infinite path $(u_i)_{i \geqslant 0}$ to which one appends the trees Y_i by adding an edge between u_i and the root of Y_i. Then, T_∞ is the local weak limit of T_n.

3.3.3 The associative tree

Suppose that for all $n \geqslant 1$, $G_{n,k}$ is uniformly distributed among all trees with n nodes (instead of n leaves as in the majority of examples) labelled with 'and' and 'or' on the internal nodes, and the literals $\{x_1, \bar{x}_1, \ldots, x_k, \bar{x}_k\}$ on the leaves. Let us denote by $T_{n,k}$ the random unlabelled tree obtained by forgetting the labels of $G_{n,k}$. Since for every t with ℓ leaves and $n - \ell$ internal nodes, there are $(2k)^\ell$ different labellings of the leaves and $2^{n-\ell}$ labelling of the internal nodes, the probability that $T_{n,k}$ is equal to a given tree t with ℓ leaves and $n - \ell$ internal nodes is proportional to k^ℓ.

Recall that, given a sequence of weights $(\omega_i)_{i \geqslant 0}$, the n-node simply generated tree is defined as follows (see, e.g., [32]):

- For an n-node rooted tree t, let its weight $w(t) = \prod_{\nu \in t} \omega_{\mathrm{out}(\nu)}$ where the product is over the nodes ν of t and $\mathrm{out}(u)$ denotes the number of children of a node u;

- an n-node simply generated tree associated with the weight sequence $(\omega_i)_{i \geqslant 0}$ is then an n-node rooted tree sampled with probability proportional to its weight $w(t)$.

Thus, $T_{n,k}$ is the simply generated tree with weights $w_0 = k$, $w_1 = 0$ and $w_i = 1$ for all $i \geqslant 2$. Note that the simply generated tree with weight sequence

$$\pi_i = \frac{1}{k\left(1 + \frac{1}{1+\sqrt{k}}\right)} \left(\frac{\sqrt{k}}{1 + \sqrt{k}}\right)^i w_i \tag{5}$$

has the same law as $T_{n,k}$, and the sequence $(\pi_i)_{i \geqslant 1}$ is a probability sequence. Then $T_{n,k}$ has the same law as a (critical) Galton–Watson tree with offspring distribution $(\pi_i)_{i \geqslant 0}$ conditioned on being of size n. In view of Section 3.3.1, we know

that such a tree locally converges to an infinite tree with one infinite end. Remark that this local limit, even unlabelled, depends on k.

3.3.4 Fragmentation trees

Consider a family of probability distributions $\mathbf{q} := (q_n)_{n \geqslant 1}$ such that q_n is a distribution on the set of partitions of the integer n. A partition of n is a non-increasing integer sequence $(\lambda_1^n, \lambda_2^n, \ldots, \lambda_n^n)$ of sum n; as an example, the partitions of 3 are $(3), (2, 1)$ and $(1, 1, 1)$. For $n = 1$, we assume that q_1 charges the partition (1), but also the empty sequence $\varnothing$. We require that for all $n \geqslant 1$, $q_n((n)) < 1$, so that q_n does not only charge the partition (n). Then the family $\mathbf{q}$ induces a family of random fragmentation trees which are defined as the genealogical trees of the fragmentation of a collection of n indistinguishable items, or balls. The tree T^n on n leaves is rooted, and this root represents the collection of n first items. With probability $q_n(\lambda_1, \lambda_2, \ldots, \lambda_p)$, the collection (n) is split into p subcollections of sizes $\lambda_1 \geqslant \cdots \geqslant \lambda_p$ with $\lambda_1 + \cdots + \lambda_p = n$. The root of T^n then has p children which are independent copies of $T^{\lambda_1}, \ldots, T^{\lambda_p}$. Note that when $q_n((n)) > 0$ it is possible that the collection remains unchanged, and that the root of T^n has only one child. (There is a similar model of random trees with n nodes, see Haas and Miermont [23] for details.) The model emcompasses the ones in [2, 6, 10, 14].

As we have already seen, the relevant information for and/or trees is located around the root, and we shall investigate conditions on $\mathbf{q}$ under which a sequence of random fragmentation trees converges locally (in distribution). Write $\ell_\downarrow^1 = \{(x_1, x_2, \ldots) : x_1 \geqslant x_2 \geqslant \cdots \geqslant 0 : \sum_{i \geqslant 1} x_i < \infty\}$ equipped with the usual ℓ^1 norm.

Proposition 3.7. *Let $\mathbf{q} = (q_n)_{n \geqslant 1}$ be a family of probability distributions such that q_n is a distribution on the set of partitions of the integer n with $q_n((n)) < 1$. Let $(\lambda_i^n)_{i \geqslant 1}$ be a random variable under q_n. If $(\lambda_i^n)_{i \geqslant 2}$ converges in distribution in $\ell_\downarrow^1$ as $n \to \infty$, then T^n converges locally in distribution to a limit random tree T^∞ with a single end.*

Proof. We first describe the limit tree. Let η denote the limit distribution of $(\lambda_i^n)_{i \geqslant 2}$ under q_n, as $n \to \infty$. Let $(\mathbf{s}^i)_{i \geqslant 0}$ denote of sequence of i.i.d. random variables distributed like η, where $\mathbf{s}^i = (s_j^i)_{j \geqslant 1}$, for $i \geqslant 0$. Note that, since the convergence holds in $\ell_\downarrow^1$, and since $\mathbf{s}^i$ is a sequence of integers, then, with probability one, there exists $\xi(i) < \infty$ such that $s_j^i = 0$ for $j \geqslant \xi(i)$. Consider the tree T constructed as follows: there is a unique half-infinite path u_i, $i \geqslant 0$, and T is rooted at u_0. Aside from u_{i+1}, the node u_i has $\xi(i)$ extra children v_j^i, $j = 1, \ldots, \xi(i)$. Then, for $i \geqslant 0$ and $1 \leqslant j \leqslant \xi(i)$, the node v_j^i is the root of a tree T_j^i which is independent of everything else, and distributed like a $\mathbf{q}$-fragmentation tree on s_j^i leaves. It should be clear that the tree T we have just described is indeed the local limit of a sequence of trees $\mathbf{q}$-fragmentation trees T^n on n leaves. $\qquad\square$

We note that Proposition 3.7 applies in particular in the case of the alpha-gamma model of Chen et al. [10] provided that $\gamma > 0$. The model also encompasses the trees defined by Ford [16] and the discrete stable trees of Marchal [30].

The alpha-gamma model introduced in [10] is a random process on the space of leaf-labelled trees. There are two parameters: $\alpha \in [0, 1]$ and $\gamma \in [0, \alpha]$. For $n = 1$ and $n = 2$, there is a unique n-leaf-labelled tree (see Figure 2). Given that T_n has been constructed we assign weight $1 - \alpha$ to each of the n edges adjacent to the leaves, a weight γ to each of the other edges and weight $(k - 1)\alpha - \gamma$ to each vertex of degree $k + 1 \geqslant 3$. So there is a total weight of $n - \alpha$. Then, an element —a vertex v_n or an edge $\{a_n, c_n\}$— is picked randomly according to the weight distribution and the tree T_{n+1} is constructed as follows:

- if we picked a vertex v_n, then we add the leaf $n + 1$ and the edge $\{v_n, n + 1\}$;

- if we picked an edge $\{a_n, c_n\}$ we split it into two and attach the new leaf $n + 1$ to the midpoint. More formally, we replace $\{a_n, c_n\}$ by three edges $\{a_n, b_n\}$, $\{b_n, c_n\}$ and $\{b_n, n + 1\}$.

Lemma 3.8. *If $(\lambda_i^n)_{i \geqslant 1}$ is distributed according to $q_n^{\alpha,\gamma}$ and if $0 < \gamma \leqslant \alpha \leqslant 1$, then $(\lambda_i^n)_{i \geqslant 2}$ converges in distribution with respect to $\ell_\downarrow^1$ as $n \to \infty$.*

Proof. To prove convergence in distribution in $\ell_\downarrow^1$, it suffices to prove that $(\sum_{i \geqslant 2} \lambda_i^n)_{n \geqslant 1}$ is tight, and that $(\lambda_2^n, \lambda_3^n, \ldots)_{n \geqslant 1}$ converges in distribution for the product topology. The split distributions $(q_n^{\alpha,\gamma})_{n \geqslant 1}$ induced by the alpha-gamma model are given in Proposition 10 of [10]: for all integers $k, n_1, \ldots, n_k$ such that $n_1 + \ldots, +n_k = n$,

$$
q_n^{\alpha,\gamma}(n_1, \ldots, n_k)
$$
$$
= \frac{\Gamma(1 - \alpha)}{\Gamma(n - \alpha)} \left(\gamma + \frac{1 - \alpha - \gamma}{n(n - 1)} \sum_{i \neq j} n_i n_j \right) \frac{\binom{n}{n_1, \ldots, n_k}}{m_1! \ldots m_n!} \alpha^{k-2} \frac{\Gamma(k - 1 - \gamma/\alpha)}{\Gamma(1 - \gamma/\alpha)} \prod_{j=1}^k \frac{\Gamma(n_j - \alpha)}{\Gamma(1 - \alpha)}, \tag{6}
$$

where $m_i = \#\{j : n_j = i\}$, for $1 \leqslant i \leqslant n$. For every fixed $k \geqslant 1$, and $(n_2, \ldots, n_k)$ note that, writing $\tilde{n} = n_2 + n_3 + \cdots + n_k$, we can rewrite $q_n^{\alpha,\gamma}(n - n_2 - \cdots - n_k, n_2, \ldots, n_k)$ as

$$q_n^{\alpha,\gamma}(n - \tilde{n}, n_2, \ldots, n_k) = c_n \left(\gamma + \frac{1 - \alpha - \gamma}{n(n-1)} \sum_{i \neq j} n_i n_j \right) \frac{\alpha^{k-2}}{m_1! \ldots m_n!} \frac{\Gamma(k - 1 - \gamma/\alpha)}{\Gamma(1 - \gamma/\alpha)} \prod_{j=2}^k \frac{\Gamma(n_j - \alpha)}{\Gamma(1 - \alpha)}$$

where, as $n \to \infty$,

$$c_n = \binom{n}{n - \tilde{n}, n_2, \ldots, n_k} \frac{\Gamma(n - \tilde{n} - \alpha)}{\Gamma(n - \alpha)} \sim \frac{1}{n_2! \cdots n_k!}.$$

It follows that

$$q_n^{\alpha,\gamma}(n - \tilde{n}, n_2, \ldots, n_k) \sim \frac{\gamma}{n_2! \cdots n_k!} \frac{\alpha^{k-2}}{m_1! \ldots m_n!} \frac{\Gamma(k - 1 - \gamma/\alpha)}{\Gamma(1 - \gamma/\alpha)} \prod_{j=2}^k \frac{\Gamma(n_j - \alpha)}{\Gamma(1 - \alpha)}. \tag{7}$$

Now to complete the proof, it suffices to check that the right-hand side above indeed defines a probability distribution on the set of non-increasing sequences of integers. To this aim, first observe that for all n large enough, one has $m_1! m_2! \cdots m_n! = m_1! m_2! \cdots m_{\tilde{n}}!$. Then, we can further rewrite (7) as

$$q_n^{\alpha,\gamma}(n - \tilde{n}, n_2, \ldots, n_k) \sim \mu(\tilde{n}) \times p_{\alpha,-\gamma}(m_1, m_2, \ldots, m_{\tilde{n}}), \tag{8}$$

where

$$\mu(\tilde{n}) = \gamma \frac{\Gamma(\tilde{n} - \gamma)}{\Gamma(1 - \gamma)\tilde{n}!}$$

and

$$p_{\alpha,\theta}(m_1, m_2, \ldots, m_{\tilde{n}}) = \frac{\tilde{n}!}{\prod_{i=1}^n (i!)^{m_i} m_i!} \alpha^{k-1} \frac{\Gamma(k + \theta/\alpha)/\Gamma(1 + \theta/\alpha)}{\Gamma(\tilde{n} + \theta)/\Gamma(1 + \theta)} \prod_{j=1}^{\tilde{n}} \left(\frac{\Gamma(j - \alpha)}{\Gamma(1 - \alpha)} \right)^{m_j}.$$

Now, $p_{\alpha,\theta}(m_1, \ldots, m_{\tilde{n}})$ is the probability distribution associated to Pitman's generalization of the Ewens sampling formula, see Proposition 9 of [35]. Finally, for any $\gamma > 0$,

$$\sum_{\tilde{n} \geqslant 1} \mu(\tilde{n}) = \sum_{\tilde{n} \geqslant 1} \gamma \frac{\Gamma(\tilde{n} - \gamma)}{\Gamma(1 - \gamma)\tilde{n}!} = \sum_{\tilde{n} \geqslant 1} \frac{\Gamma(\tilde{n} - \gamma)}{-\Gamma(-\gamma)\tilde{n}!} = 1$$

and $(\mu(i))_{i \geqslant 0}$ is also a probability distribution related to the negative binomial distribution. This shows in particular that $(\sum_{i \geqslant 2} \lambda_i^n)_{n \geqslant 1}$ converges in distribution to $\mu(\tilde{n})$, which, together with (8), completes the proof. $\qquad \square$

4 Galton–Watson trees: improved Lefmann and Savický bounds

In this section we obtain the Lefmann and Savický [27] bounds by a branching argument, improve them via a very simple symmetry argument, and extend them to all Galton–Watson trees with progenies having exponential tails. This is also the occasion to introduce the *trimming procedure* that will be crucial in the remainder of the paper (Section 5).

4.1 A symmetry argument

Let us recall Equation (1), by Chauvin et al. [7] which states that if T_n is uniformly distributed among binary trees having n leaves, then there exist two constants $c_1, c_2 > 0$, for all $k \geqslant 1$, for any k-variable Boolean function f

$$\exp(-c_1 L(f) \log k) \leqslant P_k[T_n](f) \leqslant \exp(-c_2 L(f) k^{-2}). \tag{9}$$

The bound in (9) suffers from two main problems: first it does not constrain in any way the probability of functions of complexity of order $o(k^2)$; and second, it has only been proved for the case of binary trees. In the following, we tighten the upper bound and generalize it to any critical Galton–Watson tree conditioned on being infinite, under the condition that the offspring distribution has exponential moments. A simple observation also allows us to strengthen the upper bound above.

First we need to define the notion of essential variables:

Definition 4.1. *Fix an integer k. Let f be a k-variable Boolean function: $f : (x_1, \ldots, x_k) \mapsto f(x_1, \ldots, x_k)$. For all $1 \leqslant i \leqslant k$, we say that the variable x_i is **essential** for f if and only if $f_{|x_i=0} \neq f_{|x_i=1}$ (meaning that the restriction of f to the subspace where $x_i = 1$ is not the same Boolean function as the restriction of f to the subspace where $x_i = 0$).*

Theorem 4.2. *Suppose that $\mathbb{E}[\xi] = 1$, $\mathbb{P}(\xi = 1) = 0$, and that there exists $a > 0$ such that $\mathbb{E}[e^{a\xi}] < \infty$. Let T_∞ be a Galton–Watson tree with offspring distribution ξ, conditioned on being infinite. Then, there exists constants $c_1, c_2 > 0$ such that, for every $k \geqslant 2$, for any k-variable Boolean function f, we have*

$$\exp(-c_1 L(f) \log k) \leqslant P_k[T_\infty](f) \leqslant \exp\left(-c_2 L(f) k^{-2} - \log \binom{k}{\mathtt{Ess}(f)}\right), \tag{10}$$

where $\mathtt{Ess}(f)$ is the number of essential variables of f.

This section is devoted to achieving two main goals: we first prove by a symmetry argument the improved upper bound of Theorem 4.2 relying on the looser upper bound proved by Chauvin et al. [7]; we then propose in Section 4.2 a simpler and more general proof for Chauvin et al. [7]'s result that applies to a wider class of Galton–Watson trees.

Let us first improve Chauvin et al. [7]'s upper bound. Their proof relies on the apparently blunt upper bound

$$\mathbb{P}(\mathrm{f}[\tau] = f) \leqslant \mathbb{P}(\|\tau\| \geqslant L(f)),$$

for all k-variable Boolean trees τ and all k-variable Boolean functions. This inequality is loose since $\mathbb{P}(\|\tau\| \geqslant L(f))$ actually bounds the probability of the collection of functions which may be obtained from f by permutations of the variables. More precisely, consider any Boolean function f with $\gamma := \mathtt{Ess}(f)$ essential variables. Assume without loss of generality that the essential variables of f are $x_1, \ldots, x_\gamma$. Consider a Boolean function f and a permutation π of $(x_1, \ldots, x_k)$ which does not map $\{x_1, \ldots, x_\gamma\}$ into itself. Then, by symmetry, the Boolean function $f \circ \pi$ has the same probability as f. Furthermore, for any such permutation π the functions $f \circ \pi$ and f are distinct. Thus, we actually have, for all $k \geqslant 1$

$$\binom{k}{\mathtt{Ess}(f)} \mathbb{P}(\mathrm{f}[\tau] = f) \leqslant \mathbb{P}(\|\tau\| \geqslant L(f)) \leqslant \exp(-c_2 L(f) k^{-2}).$$

Applying this inequality to $\tau = T_\infty$, and writing $\mathbb{P}(f) := \mathbb{P}_k[T_\infty](f)$, we get

$$\mathbb{P}(f) \leqslant \exp\left(-c_2 L(f) k^{-2} - \log \binom{k}{\mathtt{Ess}(f)}\right).$$

Moreover, it is interesting to note that this new upper-bound happens to be optimal in some cases, at least at the level of exponents, since it is achieved for read-once functions (see [34, page 25] for a definition of such functions): Consider a k-variable read-once function f, meaning in particular that $L(f) = \mathtt{Ess}(f)$, and suppose further that $L(f) = \mathtt{Ess}(f) = o(k)$. Then, in view of Equation (10), there exists a constant c_3 such that

$$\exp(-c_1 L(f) \log k) \leqslant \mathbb{P}(f) \leqslant \exp(-c_3 L(f) \log k),$$

so that, neither the upper nor the lower bound can be significantly improved without considering exponents that would depend on other parameters than the mere complexity $L(f)$.

4.2 Proof of Theorem 4.2: the trimming procedure

The lower bound is exactly the one from Theorem 1.1 of [27] and we shall not reproduce the argument. Our proof of the upper bound relies on a refined analysis of a certain trimming procedure, which removes the portions of the tree that do not influence the Boolean function it encodes. The cutting procedure is similar to the one used by Chauvin et al. [7], but modified in order to simplify the analysis but also to make it more powerful. (The interested reader can easily verify that our procedure removes more nodes than the one in [7].)

Consider an and/or tree τ. Let $\diamond_u$ denote the label of a node u. We also let $\Lambda(u)$ denote the collection of children of u that are leaves. Given a leaf w, we denote its label by $\mathtt{Lab}(w)$. We associate a set of **constraints** to every node, such that the sequences of sets seen when following paths away from the root are increasing (for the inclusion). Inductively define the constraints sets for all node using the following rules: for the root r, we set $C_r := \varnothing$; for a node v that is a child of u,

- if u is labelled by 'and' then $C_v = C_u \cup \bigcup_{w \in \Lambda(u)} \{\mathtt{Lab}(w) = \mathrm{True}\}$;

- if u is labelled by 'or' then $C_v = C_u \cup \bigcup_{w \in \Lambda(u)} \{\mathtt{Lab}(w) = \mathrm{False}\}$.

14

Figure 4: *The above tree is an example of an and/or tree τ to which we apply the trimming procedure* `trim`. *The dashed parts are possibly infinite and/or trees. The tree* `trim`(τ) *is highlighted in bold and in blue. The constraint sets of some nodes have been added in the shaded boxes; note that due to space constraints,* True *is represented by* 1 *and* False *by* 0. *For example, consider the node ν above. The parent of ν has constraint set $\{x_1 = 0\}$, and a sibling of ν is a leaf labelled by the literal x_2, then, since the parent of ν is labelled by $\wedge$, its constraint set is $\{x_1 = 0\} \cup \{x_2 = 1\}$.*

We say that a node is **consistent** if there exists an assignment of the variables satisfying all the constraints of its constraint set; note that all descendants of a non-consistent node are also non-consistent. We denote by `trim`(τ) the labelled tree obtained from τ, by keeping only the nodes that are consistent (an example is given in Figure 4). For a subtree τ' of τ, we let `trim`(τ') denote the portion of `trim`(τ) that is in τ' (with its labels).

In the procedure of Chauvin et al. [7] the label of the leaf v does not affect the constraint set of v itself. This is now possible in this improved trimming procedure. Also, as a consequence of the definition, if any node v is inconsistent, so are all its siblings. It follows that some internal nodes of τ end up having no progeny in `trim`(τ) (see Figure 4), and thus become leaves of `trim`(τ). For these nodes, we adopt the convention that a leaf of `trim`(τ) that is an internal node in τ labelled by $\wedge$ (resp. $\vee$) has Boolean value False (resp. True). We first verify that this trimming procedure does not modify the Boolean function that the tree represents:

Lemma 4.3. *For every and/or tree τ,* `trim`(τ) *calculates the same Boolean function as τ.*

Proof. Let ν be an inconsistent node of τ. Let us prove that the tree obtained by cutting ν and all its progeny from τ calculates the same Boolean function as τ. The fact that ν is inconsistent means that there exist two internal nodes ν_1 and ν_2 on the path between the root of τ and ν such that assigning x to False makes the tree rooted at ν_1 calculate a constant function (more precisely True if ν_1 is labelled by $\vee$ or False otherwise) and assigning x to True makes the tree rooted at ν_2 calculate a constant function. Note that the restriction on $\{x = \text{False}\}$ (resp. $\{x = \text{True}\}$) of the Boolean function calculated by τ and of the Boolean function calculated by the tree obtained from τ by cutting all progeny of ν_1 (resp. ν_2) are equal. Let us assume for example that ν_1 is an ancestor of ν_2 (note that our reasoning also holds when $\nu_1 = \nu_2$). Then, the tree obtained by cutting all progeny of ν_2 calculates the same Boolean function as τ, implying the result. $\qquad\square$

Finally, we also define the **size** of `trim`(τ) as the number of its leaves that are labelled by literals in $\{x_1, \bar{x}_1, \ldots, x_k, \bar{x}_k\}$ (or the number of leaves that were already leaves in τ). With this definition, we see that the functions True and False are computed by trees of size zero (a single internal node labelled by $\vee$ or $\wedge$), which agrees with our previous convention that they should have complexity $L(\text{True}) = L(\text{False}) = 0$.

Proposition 4.4. *Let ξ be an integer-valued random variable such that $\mathbb{E}[\xi] = 1$ and $\mathbb{P}(\xi = 1) = 0$. Suppose further that there exists an $a > 0$ such that $\mathbb{E}[e^{a\xi}] < \infty$. Let T_∞ be a Galton–Watson tree with offspring distribution ξ, conditioned on being infinite. Recall that $\hat{T}_\infty$ stands for the randomly k-variable labelled version of T_∞. Then, there exists a constant $c > 0$ such that, for any $\varpi > 0$, for any integer $k \geqslant 1$, we have*

$$\mathbb{P}\left(\|\text{trim}(\hat{T}_\infty)\| \geqslant \varpi\right) \leqslant \exp(-c\varpi/k^2).$$

15

The rest of the section is devoted to proving this proposition.

A node u may have multiple children, and the set of constraints of its children may be inconsistent even if $C(u) = \varnothing$. However, to simplify the analysis we will only search for inconsistencies at the children of nodes u for which we already have $C(u) \neq \varnothing$. In order to bound the size of the trimmed portion of a tree, we decompose the tree into a (maximal) subtree which contains only nodes with empty sets of constraints (together with the leaves that may be attached to it), to which are grafted subtrees whose internal nodes have non-empty constraint sets.

Consider T_∞, a Galton–Watson tree with critical offspring distribution ξ conditioned to be infinite. This is the random tree we have introduced in Section 3.3. Recall that a critical Galton–Watson tree conditioned on being infinite can be obtained by size-biasing the progeny distribution the nodes on a single infinite path from the root: Write $(\hat{\xi}_i)_{i \geqslant 0}$ for a sequence of i.i.d. copies of $\hat{\xi}$, the size-biased version of ξ (recall that $\mathbb{P}(\hat{\xi} = i) = i\mathbb{P}(\xi = i)$). Then T_∞ consists of an infinite backbone $(u_i)_{i \geqslant 0}$ such that the node u_i has $\hat{\xi}_i$ children, one of which is u_{i+1}. All the other offsprings of the nodes u_i, $i \geqslant 0$, are the roots of independent (unconditioned) GW_ξ random trees (see Figure 3).

Let us first consider T, an unconditioned Galton–Watson tree of progeny distribution ξ, and prove that $\|\mathtt{trim}(\hat{T})\|$ has exponential tails:

Proposition 4.5. *Let T be an unconditioned Galton–Watson tree of progeny distribution ξ, such that there exists a constant $a > 0$ verifying $\mathbb{E}[e^{a\xi}] < \infty$, and such that $\mathbb{P}(\xi = 1) = 0$. Then, $\|\mathtt{trim}(\hat{T})\|$ has exponential tails.*

To prove Proposition 4.5, let us colour the nodes of $\mathtt{trim}(\hat{T})$ as follows: the nodes having an empty constraint set in blue, and the nodes having a non empty constraint set in red. It is enough to prove that the random sizes of all these different clusters, which we denote respectively by $\|\mathtt{blue}\|$ and $\|\mathtt{red}\|$ have exponential tails. Note that the different red clusters are i.i.d. and therefore, their sizes all have the same law: $\|\mathtt{red}\|$ is the size of one of them and not the size of their union. Recall that the random labelling of T is defined for a certain integer k, and that $\|\mathtt{trim}(\hat{T})\|$, $\|\mathtt{red}\|$ and $\|\mathtt{blue}\|$ thus depend on k.

The following two lemmas concern each of the two coloured random trees.

Lemma 4.6. *Let ξ be an integer-valued random variable such that $\mathbb{E}[\xi] = 1$ and $\mathbb{P}(\xi = 1) = 0$. Suppose that there exists a constant $a > 0$ such that $\mathbb{E}[e^{a\xi}] < \infty$. Then, there exists a constant $c > 0$ such that, for all integers k and m,*

$$\mathbb{P}\left(\|\mathtt{red}\| \geqslant m\right) \leqslant \exp(-cm/k^2)$$

Proof. Note that a red cluster is a Galton–Watson tree of progeny ξ, whose root has a non-empty constraint set. Let us introduce an alternative way of sampling the k-trimmed version. We do this by introducing two types of nodes, say black and white. The root of the tree is black, meaning that its constraint set is non-empty: assume without loss of generality that this constraint set contains $\{x_1 = \mathrm{True}\}$. In the following the black nodes will be the internal nodes and the white nodes will be the leaves. Note that every black node has a constraint set that contains $\{x_1 = \mathrm{True}\}$. The white nodes are leaves and have no children, they give a chance to their siblings to become inconsistent: a leaf labelled by x_1 (resp. $\bar{x}_1$) and whose father is labelled by $\vee$ (resp. $\wedge$) makes its siblings inconsistent. The black nodes reproduce as follows
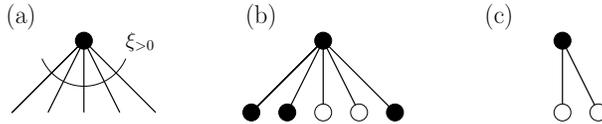

Figure 5: *Reproduction of black nodes: (a) first sample the number of children with distribution the law of $\xi_{>0}$, (b) colour the children independently at random in white with probability p_0 or black otherwise, (c) if at least one of the children is white then, with probability $1/(2k)$, delete all the black children.*

(see Figure 5): first sample $\xi_{>0}$, a copy of ξ conditioned on $\xi \geqslant 1$. Then colour each node white with probability $p_0 = \mathbb{P}(\xi = 0)$, or black with probability $1 - p_0$. Now comes the trimming part: if there is at least one white child, the black siblings are all removed with probability $1/2k$. One easily verifies that the tree obtained by this branching process is stochastically larger than the tree obtained by applying the k-trimming procedure described above (in the k-trimming procedure, the constraint-sets of the black nodes possibly contain more than one constraint, and each white node gives a chance to trim all its siblings including itself, and not only its black siblings). The matrix M of mean offspring of this branching process is

$$M = \begin{bmatrix} 0 & 0 \\ \mathbb{E}[W] & \mathbb{E}[B] \end{bmatrix},$$

16

where W and B denote respectively the number of white and black children of a black node. In particular, the largest eigenvalue M is $\mathbb{E}[B]$, and it is this value that characterises the asymptotic behaviour of the branching process.

We have

$$W \sim \mathrm{Bin}(\xi_{>0}, p_0) \qquad \text{and} \qquad B = \frac{\mathbf{1}_{\{W>0\}}}{2k} \delta_0 + \left(1 - \frac{\mathbf{1}_{\{W>0\}}}{2k}\right)(\xi_{>0} - W),$$ (11)

where $\mathrm{Bin}(n, p)$ stands for the binomial distribution of parameters $n \in \mathbb{N}$ and $p \in [0, 1]$, and δ_0 is the Dirac mass at 1. Let $\bar{W} = \xi_{>0} - W$. Note that this new random variable has law $\mathrm{Bin}(\xi_{>0}, 1 - p_0)$, and that

$$B = \frac{\mathbf{1}_{\{\bar{W}<\xi_{>0}\}}}{2k} \delta_0 + \left(1 - \frac{\mathbf{1}_{\{\bar{W}<\xi_{>0}\}}}{2k}\right) \bar{W}.$$

Thus,

$$\begin{aligned}
\mathbb{E}[B] &= \mathbb{E}\bar{W} - \frac{1}{2k}\mathbb{E}[\bar{W}\mathbf{1}_{\{\bar{W}<\xi_{>0}\}}] \\
&= (1 - p_0)\mathbb{E}\xi_{>0} - \frac{1}{2k}\left(\mathbb{E}\bar{W} - \mathbb{E}[\bar{W}\mathbf{1}_{\{\bar{W}=\xi_{>0}\}}]\right) \\
&= \frac{1 - p_0}{1 - p_0} - \frac{1}{2k}\left((1 - p_0)\mathbb{E}\xi_{>0} - \mathbb{E}[\xi_{>0}(1 - p_0)^{\xi_{>0}}]\right) \\
&= 1 - \frac{1}{2k}\left(1 - \mathbb{E}[\xi_{>0}(1 - p_0)^{\xi_{>0}}]\right),
\end{aligned}$$

because $\mathbb{E}\xi_{>0} = \frac{1}{1-p_0}$ since $\mathbb{E}\xi = 1$. Note that, with $p_i = \mathbb{P}(\xi = i)$, we have

$$\mathbb{E}[\xi_{>0}(1 - p_0)^{\xi_{>0}}] = \sum_{i \geqslant 1} i p_i (1 - p_0)^{i-1} < 1,$$

since $p_1 = \mathbb{P}(\xi = 1) = 0$ by assumption. We thus have that there exists a positive constant c such that,

$$\mathbb{E}B = 1 - \frac{c}{k}.$$

It follows that the black subtree rooted at a black node is a subcritical Galton–Watson tree. Let Ξ denote the total progeny of the black subtree, and $(B_i)_{i \geqslant 1}$ be i.i.d. copies of B. Then,

$$\mathbb{P}(\Xi \geqslant n) \leqslant \mathbb{P}\left(\sum_{i=1}^{n}(B_i - 1) \geqslant 0\right) \leqslant \exp(-n\Psi^\star(1 - \mathbb{E}[B])),$$

by Cramér's theorem [12], where $\Psi^\star$ is the large deviations rate function of $B - \mathbb{E}[B]$. More precisely, write $\Psi(t) := \log \mathbb{E}[e^{t(B - \mathbb{E}[B])}]$, and observe that $\Psi(t) < +\infty$ since ξ has exponential moments, and $\Psi(\varpi) \sim \Psi''(0)\varpi^2/2$ as $\varpi \to 0$. Then we have, setting $\alpha = 1/\Psi''(0)$, as $\varpi \to 0$,

$$\Psi^\star(\varpi) = \sup_{t \in \mathbb{R}}\{t\varpi - \Psi(t)\} \geqslant \alpha\varpi^2 - \Psi(\alpha\varpi) \sim x\varpi^2\alpha/2.$$

This guarantees that there exists a constant $c > 0$ such that

$$\mathbb{P}(\Xi \geqslant n) \leqslant e^{-cn/k^2}.$$ (12)

In order to recover the size of the whole tree, it suffices to add the missing white nodes. Writing $(W_i)_{i \geqslant 1}$ for a sequence of i.i.d copies of W defined in (11), it follows from (12) that

$$\mathbb{P}(\|\mathtt{red}\| \geqslant m) \leqslant \mathbb{P}\left(\sum_{i \geqslant 1} W_i \mathbf{1}_{\{i \leqslant \Xi\}} \geqslant m\right) \leqslant \mathbb{P}\left(\sum_{1 \leqslant i \leqslant 2m/\mathbb{E}[W]} W_i \geqslant m\right) + e^{-c_2 m/(k^2 \mathbb{E}[B])} \leqslant e^{-c_1 m} + e^{-c_2 m/k^2},$$

by a second use of Cramér's theorem for large deviations. Indeed, since ξ has exponential tails, W also does and the theorem applies. $\qquad\square$

Lemma 4.7. *Let ξ be an integer-valued random variable such that $\mathbb{E}[\xi] = 1$, and $\mathbb{P}(\xi = 1) = 0$. Suppose that there exists a constant $a > 0$ such that $\mathbb{E}[e^{a\xi}] < \infty$. Then, there exists a constant $c > 0$ such that, for all integers k and m,*

$$\mathbb{P}(\|\mathtt{blue}\| \geqslant m) \leqslant \exp(-cm/k^2).$$

17

Proof. The proof is very similar to that of Lemma 4.6. We now have black, white, and in addition green nodes which are the ones having a white sibling. Black nodes are the ones having an empty constraint set, the white ones are the leaves and the green ones are nodes having a non-empty constraint set (being then the roots of independent copies of a red cluster). Clearly, the black subtree is dominated by the black subtree considered in the proof of Lemma 4.6, so its size also has exponential tails. Then the nodes to add are only the white and green nodes, and the distribution is precisely that of $\xi_{>0}$ conditioned on $W > 0$. This also has exponential tails since

$$\mathbb{P}\left(\xi_{>0} \geqslant \varpi \mid W > 0\right) \leqslant p_0^{-1}\mathbb{P}\left(\xi_{>0} \geqslant 0\right).$$

We omit the straightforward details. □

Lemmas 4.6 and 4.7 tell us that each red subtree has exponential tails, and that their number (equal to $\|\mathtt{blue}\|$) also has exponential tails, which implies Proposition 4.5.

Proof of Proposition 4.4. Let T_∞ be a Galton–Watson tree of progeny ξ conditioned on being infinite. It can be described as an infinite spine on which independent (unconditioned) Galton–Watson of progeny distribution ξ are hanging. Let us associate to every node of $\hat{T}_\infty$ its constraint set as explained in the trimming procedure, and let us first focus on the nodes of the spine $(u_i)_{i\geqslant 0}$ (see Figure 6). The first nodes of this sequence have empty constraint sets. Therefore, the trees hanging onto them fall under Proposition 4.5. The following nodes on the spine have non-empty constraint sets and therefore, the unconditioned Galton–Watson trees hanging on them have the same law as the red clusters studied in Lemma 4.6. In all cases, the trimmed versions of the subtrees hanging on the spine have exponential tails.

It thus only remains to prove that the total number of trees hanging on the trimmed spine has exponential tails. Let us denote by $\hat{W}_i$ the random number of leaves of node u_i. The sequence $(\hat{W}_i)_{i\geqslant 0}$ is i.i.d. and we denote by $\hat{W}$ a random variable having the common law of the $\hat{W}_i$'s. Recall that we have two different kinds of nodes on the spines: nodes with empty constraints sets (at the top of the tree), and nodes with non-empty constraints sets. Let us denote by N the number of nodes on the spine with empty constraint sets, and by $N + M$ the total number of nodes on the trimmed spine. Let us prove that the total number of trees hanging on nodes of the spine

$\hat{T}_\infty$

infinite spine

Figure 6: *The first N nodes of the spine have empty constraint sets, the following M nodes have non-empty constraint sets, and all nodes below that are non-consistent, and thus get cut during the trimming procedure. The white subtrees fall under Proposition 4.5, the gray ones under Lemma 4.6.*

with non-empty constraint sets has exponential tails. The proof that the total number of trees hanging on nodes of the spine having *empty* constraints set has exponential tails follows the same outline and is actually simpler: this case will be left to the reader.

So let us treat the case of the M nodes of the trimmed spine having non-empty constraint sets: let ν be such a node. Assume without loss of generality that its constraint set is $\{x_1 = \mathrm{True}\}$. Since the spine has not been cut before level $N + M$, the node ν cannot be inconsistent: thus the leaf-children of ν cannot be labelled by x_1 (resp. $\bar{x}_1$ depending on the connector labelling ν). We denote by $\hat{S}$ the indicator of the event "no leaf-child of ν is labelled by x_1", and by $\hat{W}$ the number of leaf-children of ν. We have:

$$\hat{W} = \mathrm{Bin}(\hat{\xi} - 1, p_0), \qquad \text{and} \qquad \hat{S} = \mathrm{Ber}\left((1 - 1/2k)^{\hat{W}}\right).$$

Therefore, if we denote by $m := \min\{n \geqslant 1 \mid p_n > 0\}$, where $p_n := \mathbb{P}(\xi = n)$, then

$$\mathbb{P}(\hat{S} = 1) = \mathbb{E}[(1 - 1/2k)^{\hat{W}}] \geqslant \mathbb{P}(\hat{W} = 0) \geqslant \mathbb{P}(\hat{\xi} = m)(1 - p_0)^{m-1} =: \kappa > 0,$$

and it follows that, $i \geqslant 0$ and $k \geqslant 1$,

$$\mathbb{P}(\hat{\xi} - 1 \geqslant i \mid \hat{S} = 1) \leqslant \frac{\mathbb{P}(\hat{\xi} - 1 \geqslant i)}{\mathbb{P}(\hat{S} = 1)} \leqslant \kappa^{-1}\mathbb{P}(\hat{\xi} \geqslant i + 1).$$

In particular, if $\hat{\xi}$ has exponential moments, so does the number of siblings of a node u_i of the spine conditional on not being cut. Let $(X_i)_{i\geqslant 1}$ be an i.i.d. sequence of copies of $\hat{\xi} - 1$ conditional on $\hat{S} = 1$, and M for a geometric random
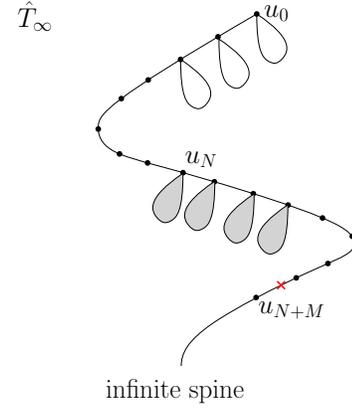
variable with success parameter $\eta := \mathbb{P}(\hat{S} = 0) < 1$ (since $\mathbb{P}(\hat{W} > 0) > 0$). Then, writing Δ for the collection of trees hanging on node of the trimmed spine having a non-empty constraint set, we have, for all integer m:

$$
\begin{aligned}
\mathbb{P}\left(|\Delta| \geqslant m\right) &\leqslant \mathbb{P}\left(\sum_{i \geqslant 1} X_i \mathbf{1}_{\{i \leqslant M\}} \geqslant m\right) \\
&\leqslant \mathbb{P}\left(\sum_{i=1}^{m} X_i \geqslant n\right) + \mathbb{P}\left(M \leqslant m\right) \\
&\leqslant \mathbb{P}\left(\sum_{i=1}^{\lceil \epsilon m \rceil} X_i \geqslant n\right) + (1 - \eta)^{\epsilon m} \\
&\leqslant e^{-cm},
\end{aligned}
\tag{13}
$$

for $\epsilon > 0$ small enough and c a constant depending on ϵ, which concludes the proof. $\qquad\square$

5 Improved relations between complexity and probability

Section 3.2 was devoted to proving that, when the sequence of random trees $(T_n)_{n \geqslant 1}$ converges locally in distribution to an infinite tree with finitely many ends T_∞, then the distribution of the random Boolean function calculated by the random k-variable labelling of T_n converges to an asymptotic distribution $\mathbb{P}_k = P_k[T_\infty]$ when $n \to +\infty$ (cf. Theorem 3.6).

We here state and prove an equivalent of the result of Kozik [26] (see Equation (2)): fix an integer k_0, and a k_0-variable Boolean function f, we are able to understand the behaviour of $\mathbb{P}_k(f)$ when k tends to infinity. Remark though that we need stronger assumptions than the one needed to get convergence to the asymptotic distribution: (A) We restrict ourselves to random trees whose local limit has a unique end, although we believe that the result holds for local limits having finitely many ends. More importantly, (B) we need assumptions that permit to control the sizes of the finite trees attached to the infinite spine.

5.1 Controlling the repetitions

First of all, let us prove the following crucial lemma concerning the probability that the k-trimmed version (according to `trim`) of a randomly labelled tree contains repetitions. The number of **repetitions** in a labelled and/or tree is defined as the difference between the number of its leaves and the number of distinct variables that appear as leaf-labels of this tree. As an example, the left tree in Figure 1 has 5 repetitions since it has size 8 and is labelled by 3 distinct variables, namely x_1, x_2 and x_3.

Lemma 5.1. *(a) There exists an integer k_0 such that, for any $k \geqslant k_0$, for any tree t_∞, for any integer q,*

$$
\mathbb{P}(\mathtt{trim}(\hat{t}_\infty) \text{ contains at least } q \text{ repetitions}) \leqslant \frac{\mathbb{E}[\|\mathtt{trim}(\hat{t}_\infty)\|^{2q}] + 2\mathrm{e}\,\mathbb{E}[\|\mathtt{trim}(\hat{t}_\infty)\|^{q}]}{k^q}.
$$

(b) For any infinite tree t_∞, and for all integers p and q, there exists a constant $K_{p,q}(t_\infty)$ such that

$$
\mathbb{P}(\mathtt{trim}(\hat{t}_\infty) \text{ has size } p \text{ and contains at least } q \text{ repetitions}) \leqslant \frac{K_{p,q}(t_\infty)}{k^{q+1}}.
$$

Proof. (a) For a subtree t of $\hat{t}_\infty$, we denote by $\mathtt{shape}(t)$ the tree t in which the nodes that are leaves of $\hat{t}_\infty$ have been unlabelled. We emphasize the fact that an element $\mathtt{shape}(t)$ may have leaves labelled by connectives ($\wedge$ or $\vee$), and that all its internal nodes are labelled by connectives.

Let us decompose the event $\{\mathtt{trim}(\hat{t}_\infty)$ contains at least q repetitions$\}$ according to the different possible realisations of $\mathtt{trim}(\hat{t}_\infty)$. We denote by $\mathtt{sub}(t_\infty)$ the support of the random variable $\mathtt{shape}(\mathtt{trim}(\hat{t}_\infty))$ (where the randomness comes from the random labelling of the tree t_∞). We let $\mathtt{sub}_m(t_\infty)$ be the subset of $\mathtt{sub}(t_\infty)$ consisting of the trees having m nodes that are leaves of t_∞. With the definition of size of a trimmed tree, the elements of $\mathtt{sub}_m(t_\infty)$ all have size m. Given a tree $t \in \mathtt{sub}(t_\infty)$, we denote by $\Delta(t)$ (resp. $\Delta^{[\geqslant q]}(t)$, resp. $\Delta^{[0]}(t)$, resp. $\Delta^{[<q]}(t)$) the set of all different leaf-labelled version of t (resp. having at least q repetitions, resp. having no repetitions, resp. having at most $q - 1$ repetitions); we see an element $\ell \in \Delta(t)$ as a function from the set of leaves of t (that are also leaves of t_∞) to $\{x_1, \bar{x}_1, \ldots, x_k, \bar{x}_k\}$. Given a leaf w of t, we denote by $\ell(w)$ its label according to $\ell \in \Delta(t)$.

Recall that $\mathtt{Lab}(w)$ denotes the random label of the leaf w in $\hat{t}_\infty$. For $t \in \mathtt{sub}(t_\infty)$ and $\ell \in \Delta(t)$, we also write $\mathtt{Lab}(t : \hat{t}_\infty) = \ell$ for the event "$\mathtt{Lab}(w) = \ell(w)$ for every leaf w of t_∞ that is also contained in t".

The addition of an index $\mathtt{notrim}$ means that we restrict ourselves to the set of labellings ℓ of t, such that conditionally on $\mathtt{Lab}(t : \hat{t}_\infty) = \ell$, the trimming procedure leaves all the nodes of t consistent, or in other words such that $\mathtt{trim}(t) = \ell$.

In the following we argue conditionally to the random labelling of the internal nodes of $\hat{t}_\infty$, and the reasoning is valid for any such labelling: we denote by $\mathbb{P}^\star$ this conditional probability. On the one hand, using Markov's inequality, we have, for any k,

$\mathbb{P}^\star(\mathtt{trim}(\hat{t}_\infty)$ contains at least q repetitions$)$

$$= \mathbb{P}^\star(\|\mathtt{trim}(\hat{t}_\infty)\| \geqslant \sqrt{k}) + \sum_{m=1}^{\lfloor \sqrt{k} \rfloor} \sum_{t \in \mathtt{sub}_m(t_\infty)} \sum_{\ell \in \Delta_{\mathtt{notrim}}^{[\geqslant q]}(t)} \mathbb{P}^\star(\mathtt{shape}(\mathtt{trim}(\hat{t}_\infty)) = t \text{ and } \mathtt{Lab}(t : \hat{t}_\infty) = \ell)$$

$$\leqslant \frac{\mathbb{E}^\star[\|\mathtt{trim}(\hat{t}_\infty)\|^{2q}]}{k^q} + \sum_{m=1}^{\lfloor \sqrt{k} \rfloor} \sum_{t \in \mathtt{sub}_m(t_\infty)} \sum_{\ell \in \Delta_{\mathtt{notrim}}^{[\geqslant q]}(t)} \mathbb{P}^\star(\mathtt{Lab}(t : \hat{t}_\infty) = \ell)\mathbb{P}^\star(\mathtt{shape}(\mathtt{trim}(\hat{t}_\infty)) = t \mid \mathtt{Lab}(t : \hat{t}_\infty) = \ell)$$

$$= \frac{\mathbb{E}[\|\mathtt{trim}(\hat{t}_\infty)\|^{2q}]}{k^q} + \sum_{m=q}^{\lfloor \sqrt{k} \rfloor} \frac{1}{(2k)^m} \sum_{t \in \mathtt{sub}_m(t_\infty)} \sum_{\ell \in \Delta_{\mathtt{notrim}}^{[\geqslant q]}(t)} \mathbb{P}^\star(\mathtt{shape}(\mathtt{trim}(\hat{t}_\infty)) = t \mid \mathtt{Lab}(t : \hat{t}_\infty) = \ell). \tag{14}$$

Indeed, for every fixed $t \in \mathtt{sub}_m(t_\infty)$, and $\ell \in \Delta(t)$ the m leaves of $\hat{t}_\infty$ that are in t have independent labels, so that, for all $\ell \in \Delta(t)$,

$$\mathbb{P}^\star(\mathtt{Lab}(t : \hat{t}_\infty) = \ell) = \mathbb{P}^\star(\mathtt{Lab}(w) = \ell(w) \text{ for all leaf } w \text{ of } t_\infty \text{ that is also a leaf of } t) = \frac{1}{(2k)^m}.$$

We also used that $\mathbb{E}^\star[\|\mathtt{trim}(\hat{t}_\infty)\|^{2q}] = \mathbb{E}[\|\mathtt{trim}(\hat{t}_\infty)\|^{2q}]$ since the distribution of $\|\mathtt{trim}(\hat{t}_\infty)\|$ does not depend on the labels of the internal nodes of $\hat{t}_\infty$.

Note that, on the other hand, we have, for all integer m,

$$\mathbb{P}^\star(\|\mathtt{trim}(\hat{t}_\infty)\| = m) = \frac{1}{(2k)^m} \sum_{t \in \mathtt{sub}_m(t_\infty)} \sum_{\ell \in \Delta_{\mathtt{notrim}}(t)} \mathbb{P}^\star(\mathtt{shape}(\mathtt{trim}(\hat{t}_\infty)) = t \mid \mathtt{Lab}(t : \hat{t}_\infty) = \ell).$$

To upper bound the number of repetitions, let us look at the following ratio for all $1 \leqslant m < \sqrt{k}$,

$$\frac{\sum_{t \in \mathtt{sub}_m(t_\infty)} \sum_{\ell \in \Delta_{\mathtt{notrim}}^{[\geqslant q]}(t)} \mathbb{P}^\star(\mathtt{shape}(\mathtt{trim}(\hat{t}_\infty)) = t \mid \mathtt{Lab}(t : \hat{t}_\infty) = \ell)}{\sum_{t \in \mathtt{sub}_m(t_\infty)} \sum_{\ell \in \Delta_{\mathtt{notrim}}(t)} \mathbb{P}^\star(\mathtt{shape}(\mathtt{trim}(\hat{t}_\infty)) = t \mid \mathtt{Lab}(t : \hat{t}_\infty) = \ell)} = \frac{1}{1 + \mathtt{rat}} \leqslant \frac{1}{\mathtt{rat}},$$

where

$$\mathtt{rat} := \frac{\sum_{t \in \mathtt{sub}_m(t_\infty)} \sum_{\ell \in \Delta_{\mathtt{notrim}}^{[<q]}(t)} \mathbb{P}^\star(\mathtt{shape}(\mathtt{trim}(\hat{t}_\infty)) = t \mid \mathtt{Lab}(t : \hat{t}_\infty) = \ell)}{\sum_{t \in \mathtt{sub}_m(t_\infty)} \sum_{\ell \in \Delta_{\mathtt{notrim}}^{[\geqslant q]}(t)} \mathbb{P}^\star(\mathtt{shape}(\mathtt{trim}(\hat{t}_\infty)) = t \mid \mathtt{Lab}(t : \hat{t}_\infty) = \ell)}$$

$$\geqslant \frac{\sum_{t \in \mathtt{sub}_m(t_\infty)} \sum_{\ell \in \Delta^{[0]}(t)} \mathbb{P}^\star(\mathtt{shape}(\mathtt{trim}(\hat{t}_\infty)) = t \mid \mathtt{Lab}(t : \hat{t}_\infty) = \ell)}{\sum_{t \in \mathtt{sub}_m(t_\infty)} \sum_{\ell \in \Delta_{\mathtt{notrim}}^{[\geqslant q]}(t)} \mathbb{P}^\star(\mathtt{shape}(\mathtt{trim}(\hat{t}_\infty)) = t \mid \mathtt{Lab}(t : \hat{t}_\infty) = \ell)},$$

because a labelled tree with no repetition cannot be trimmed, so for every t, and for every integer $q \geqslant 1$,

$$\Delta^{[0]}(t) = \Delta_{\mathtt{notrim}}^{[0]}(t) \subseteq \Delta_{\mathtt{notrim}}^{[<q]}(t)$$

Given $t \in \mathtt{sub}_m(t_\infty)$, we denote by ∂t the the set of nodes of t that have no children in t, but that are not leaves of t_∞; so ∂t is the internal vertex boundary of t inside t_∞. Each node of ∂t has all its progeny that is inconsistent, but is not itself inconsistent (see Figure 7).

Consider first the case of $\ell \in \Delta^{[0]}(t)$. Given a node $v \in \partial t$, let us denote by $D(v)$ the number of leaves in t whose father is an ancestor of v. Conditional on $\{\mathtt{Lab}(t) = \ell\}$, these $D(v)$ leaves are labelled according to ℓ, and since ℓ contains no repetition, they are labelled by $D(v)$ distinct variables. These $D(v)$ leaves define a set of $D(v)$ literals such that: if at least one of the leaves whose parent is v is labelled by one of these literals (or its negation, depending on the labelling of the internal nodes which we conditioned on), then all the children of v are inconsistent (and thus cut by the trimming
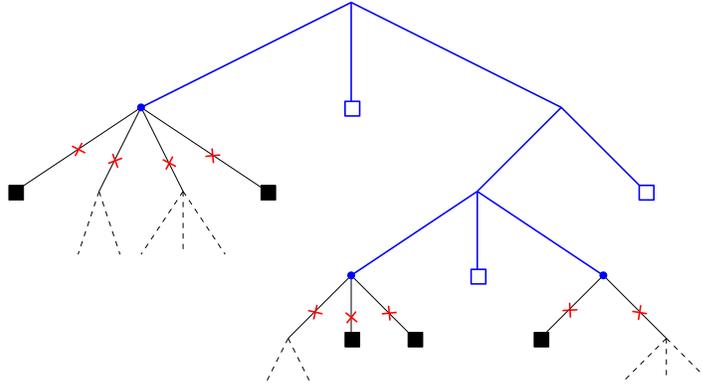
Figure 7: *This figure represents and infinite but locally finite tree t_∞ (the dashed parts are possibly infinite trees that we don't represent in this picture) and one of its subtrees t in bold blue. To have $\mathrm{trim}(\hat{t}_\infty) = t$ the nine cuts in red must have been made during the trimming procedure. The three internal nodes marked by a black dot are the elements of ∂t, and if we denote by v the leftmost one, $D(v) = 1$ and $F(v) = 2$.*

procedure). Let us denote by $F(v)$ the number of leaves in t_∞ whose parent is v; not that for $v \in \partial t$, none of these leaves can be in t. Let also $p(d, f)$ be the probability that at least one among f leaves is labelled by $\alpha_1, \alpha_2, \ldots,$ or α_d (where $\{\alpha_1, \ldots \alpha_d\}$ is any fixed subset of d elements of $\{x_1, \bar{x}_1, \ldots, x_k, \bar{x}_k\}$), or two among those f leaves are labelled by a literal and its negation. (This is clearly independent of the set $\{\alpha_1, \alpha_2, \ldots, \alpha_d\}$, provided it has cardinality d.) Then, the arguments above and the independence of the leaf labels in $\hat{t}_\infty$ imply that given a subtree t and a labelling of this subtree $\ell \in \Delta^{[0]}(t)$, conditioned on the labels of the internal nodes of $\hat{t}_\infty$,

$$\mathbb{P}^\star(\mathtt{shape}(\mathrm{trim}(\hat{t}_\infty)) = t \mid \mathtt{Lab}(t : \hat{t}_\infty) = \ell) = \prod_{v \in \partial t} p(D(v), F(v)).$$

Note that this probability does not depend on the labels of the internal nodes of $\hat{t}_\infty$. This is due to the fact that in the trimming procedure, the labels $\wedge$ and $\vee$, literals and their negations behave symmetrically: the constraint $\alpha = \mathrm{True}$ is generated by a leaf labelled by α whose parent is labelled by $\wedge$ as well as by a leaf labelled by $\bar{\alpha}$ whose parent is labelled by $\vee$.

The same arguments are valid for every subtree t of t_∞, and for every $\ell \in \Delta^{[\geqslant q]}(t)$, except that since there may be some repetitions in ℓ, for every leaf v in ∂t, the number of labellings permitting to trim is at most $D(v)$ (and not exactly $D(v)$ as above). Therefore, since the function $p(d, f)$ is increasing in d, for every $t \in t_\infty$, and $\ell \in \Delta^{[\geqslant q]}(t)$:

$$\mathbb{P}^\star(\mathtt{shape}(\mathrm{trim}(\hat{t}_\infty)) = t \mid \mathtt{Lab}(t : \hat{t}_\infty) = \ell) \leqslant \prod_{v \in \partial t} p(D(v), F(v)). \tag{15}$$

Thus, since the cardinalities $|\Delta^{[\geqslant q]}(t)|$ and $|\Delta^{[0]}(t)|$ depend on the size m of t but not on t itself, we have

$$\mathtt{rat} \geqslant \frac{\sum_{t \in \mathtt{sub}_m(t_\infty)} \prod_{v \in \partial t} p(D(v), F(v)) \, |\Delta^{[0]}(t)|}{\sum_{t \in \mathtt{sub}_m(t_\infty)} \prod_{v \in \partial t} p(D(v), F(v)) \, |\Delta^{[\geqslant q]}_{\mathtt{notrim}}(t)|}$$

$$\geqslant \frac{\sum_{t \in \mathtt{sub}_m(t_\infty)} \prod_{v \in \partial t} p(D(v), F(v)) \, |\Delta^{[0]}(t)|}{\sum_{t \in \mathtt{sub}_m(t_\infty)} \prod_{v \in \partial t} p(D(v), F(v)) \, |\Delta^{[\geqslant q]}(t)|} \geqslant \frac{|\Delta^{[0]}_m|}{|\Delta^{[\geqslant q]}_m|},$$

since, for any integer q, for any subtree t of t_∞, $|\Delta^{[\geqslant q]}_{\mathtt{notrim}}(t)| \leqslant |\Delta^{[\geqslant q]}(t)|$, where $|\Delta^{[\geqslant q]}_m|$ (resp. $|\Delta^{[0]}_m|$) is the number of ways to label m leaves with at least q repetitions (resp. no repetition). We have,

$$|\Delta^{[0]}_m| = k(k-1) \cdots (k - m + 1) = \frac{k!}{(k-m)!},$$

and,

$$|\Delta^{[\geqslant q]}_m| = \sum_{v=1}^{m-q} \left\{ {m \atop v} \right\} k(k-1) \cdots (k - v + 1) \leqslant m^q \sum_{v=1}^{m-q} \left\{ {m - q \atop v} \right\} \frac{k!}{(k-v)!} \leqslant m^q k^{m-q},$$

21

where for all integers $n, p \geqslant 1$, $\left\{{n \atop p}\right\}$ is the Stirling number of second kind, i.e. the number of ways to partition a set of n elements into p non-empty parts. We have used the standard equality (see, e.g., [11, p. 207])

$$\sum_{p=1}^{n} \left\{{n \atop p}\right\} x(x-1)\cdots(x-p+1) = x^n,$$

for all integers n and x. Thus,

$$\mathtt{rat} \geqslant \frac{k(k-1)\cdots(k-m+1)}{m^q k^{m-q}} \geqslant \frac{k^q}{m^q}\left(1 - \frac{m}{k}\right)^m \geqslant \frac{k^q}{2em^q},$$

for all $m < \sqrt{k}$, for all $k \geqslant k_0$, where k_0 is defined as $k_0 = \max\{k \geqslant 1 : (1 - 1/\sqrt{k})^{\sqrt{k}} < 1/2e\}$. We thus get

$$\mathbb{P}(\mathtt{trim}(\hat{t}_\infty) \text{ contains at least } q \text{ repetitions}) \leqslant \frac{\mathbb{E}\|\mathtt{trim}(\hat{T}_\infty)\|^{2q}}{k^q} + 2e \sum_{m=1}^{\lfloor\sqrt{k}\rfloor} \frac{m^q}{k^q} \mathbb{P}(\|\mathtt{trim}(\hat{t}_\infty)\| = m)$$

$$= \frac{\mathbb{E}\|\mathtt{trim}(\hat{t}_\infty)\|^{2q} + 2e\, \mathbb{E}\|\mathtt{trim}(\hat{t}_\infty)\|^q}{k^q},$$

for all $k \geqslant k_0$, as desired.

(b) We follow the line of argument and use the same notations as in the proof of statement (a). We have (see Equation (14))

$$\mathbb{P}^\star(\mathtt{trim}(\hat{t}_\infty) \text{ has size } p \text{ and at least } q \text{ repetitions})$$

$$= \frac{1}{(2k)^p} \sum_{t \in \mathtt{sub}_p(t_\infty)} \sum_{\ell \in \Delta_{\mathtt{notrim}}^{[\geqslant q]}(t)} \mathbb{P}^\star(\mathtt{shape}(\mathtt{trim}(\hat{t}_\infty)) = t \mid \mathtt{Lab}(t : \hat{t}_\infty) = \ell),$$

and (see Equation (15)),

$$\mathbb{P}^\star(\mathtt{shape}(\mathtt{trim}(\hat{t}_\infty)) = t \mid \mathtt{Lab}(t : \hat{t}_\infty) = \ell) \leqslant \prod_{v \in \partial t} p(D(v), F(v)).$$

Recall that $p(d, f)$ is, by definition, the probability that either at least one among f leaves is labelled by one of $\alpha_1, \alpha_2, \ldots, \alpha_d$, or two among those f leaves are labelled by a literal and its negation. Using the bound $(1-x)^n \geqslant 1 - nx$ for all integer $n \geqslant 1$ and for all $x \in (0, 1)$, we obtain

$$1 - p(d, f) = \frac{(2k-d)!}{(2k-d-f)!(2k)^f} \geqslant \left(1 - \frac{d+f}{2k}\right)^f \geqslant 1 - \frac{(d+f)f}{2k},$$

so that

$$p(d, f) \leqslant \frac{(d+f)f}{2k}.$$

Note that since $\hat{t}_\infty$ is an infinite tree and $t \in \mathtt{sub}_p(t_\infty)$ has size p, the set ∂t is non empty. Let v be a node of ∂t: this node has height at most p because the tree t_∞ and thus t have no unary node (by definition of an and/or tree, see Section 2.1). Since t_∞ is locally finite, the constant

$$F_{\mathtt{max}} := \max\{F(v) : v \in t_\infty \text{ having height at most } p\} \tag{16}$$

is finite, and a similar argument implies that the constant

$$D_{\mathtt{max}} := \max\{D(v) : v \in t_\infty \text{ having height at most } p\} \tag{17}$$

is also finite, so that that for all node $v \in \partial t$

$$p(D(v), F(v)) \leqslant \frac{(D_{\mathtt{max}} + F_{\mathtt{max}})F_{\mathtt{max}}}{2k},$$

and thus,

$$\mathbb{P}^\star(\mathtt{trim}(\hat{t}_\infty) = \ell \mid \mathtt{Lab}(t) = \ell) \leqslant \frac{(D_{\mathtt{max}} + F_{\mathtt{max}})F_{\mathtt{max}}}{2k}.$$

Therefore, we get

$$\mathbb{P}^\star(\mathtt{trim}(\hat{t}_\infty) \text{ has size } p \text{ and at least } q \text{ repetitions}) \leqslant \frac{(D_{\max} + F_{\max})F_{\max}}{(2k)^{p+1}} \, |\mathtt{sub}_p(t_\infty)| \, |\Delta_p^{[\geqslant q]}|.$$

Recall that

$$|\Delta_p^{[\geqslant q]}| \leqslant p^q \frac{k!}{(k-p+q)!} \leqslant p^q k^{p-q},$$

which gives

$$\mathbb{P}^\star(\mathtt{trim}(\hat{t}_\infty) \text{ has size } p \text{ and at least } q \text{ repetitions}) \leqslant (D_{\max} + F_{\max})F_{\max} \, |\mathtt{sub}_p(t_\infty)| \, \frac{p^q}{2^{p+1}k^{q+1}}.$$

Noting that $|\mathtt{sub}_p(t_\infty)|$ is finite concludes the proof with the following choice of $K_{p,q}(t_\infty)$

$$K_{p,q}(t_\infty) := (D_{\max} + F_{\max})F_{\max} \, |\mathtt{sub}_p(t_\infty)| \, \frac{p^q}{2^{p+1}}, \tag{18}$$

which completes the proof of (b). $\qquad\square$

5.2 Probability of the two constant functions

Let T_∞ be the limit tree of the random sequence $(T_n)_{n \geqslant 0}$, and suppose that it has a unique end. Let $u_0, u_1, \ldots$ be the nodes of the spine, starting from the root. Recall that, as already mentioned, the distribution $(T_n)_{n \geqslant 0}$ and thus T_∞ may depend on k (we will provide such an example at the end of the section). For all $i \geqslant 0$, we denote by A_i the random forest of finite subtrees hanging from the node u_i.

Let us denote by L_i the number of leaves of $\mathtt{trim}(\hat{A}_i)$. In the following, we will assume that

(H) the sequence $(A_i)_{i \geqslant 0}$ is i.i.d.

When (H) holds, we will denote by A a random variable having the common law of the A_i, and by $\mathtt{trim}(\hat{A})$ the trimmed version of its random labelling. Observe that (H) implies that $(L_i)_{i \geqslant 0}$ is a sequence of i.i.d. random variables as well, and we denote by L a random variable having their common law.

Given a forest a, and two of its leaves ℓ_1 and ℓ_2 we denote by C_{ℓ_1,ℓ_2} the number of nodes (of out-degree at least two) of the union of the path from ℓ_1 to one root of a and the path from ℓ_2 to one root (possibly the same) of a. We denote by C_a the minimum of those C_{ℓ_1,ℓ_2} taken on all couple of leaves (ℓ_1, ℓ_2) and by N_a the number of such couples that realize this minimum.

Lemma 5.2. *Suppose that the hypothesis of Theorem 3.6 holds, that T_∞ has almost surely a unique end and satisfies (H). Then, asymptotically as k tends to infinity,*

$$\frac{1}{k} \, \mathbb{E}\left[\frac{N_A}{2^{C_A}}\right] \leqslant \mathbb{P}_k(\text{True}) = \mathbb{P}_k(\text{False}) \leqslant \frac{(2\mathsf{e} + 1)\mathbb{E}[L] + \mathbb{E}[L^2]}{k}.$$

Example. Before proceeding to the proof, let us present an example that shows that the moment conditions in Lemma 5.2 cannot be removed altogether. Consider a sequence $(X_i)_{i \geqslant 0}$ of i.i.d. copies of an integer-valued random variable X. Let T_∞ consists of an infinite spine $(u_i)_{i \geqslant 0}$, and such that, for each $i \geqslant 0$, the node u_i has X_i leaf-children aside from u_{i+1}. Then $L = X, C = 0$ and $N = X(X-1)/2$. Then, Lemma 5.2 implies that, if $\mathbb{E}[X^2] = o(k)$ as $k \to \infty$, we have

$$\frac{\mathbb{E}[X^2] - \mathbb{E}[X]}{2k} \leqslant \mathbb{P}_k(\text{True}) = \mathbb{P}_k(\text{False}) \leqslant \frac{\mathbb{E}[X^2] + (2\mathsf{e} + 1)\mathbb{E}[X]}{k},$$

implying that $\mathbb{P}_k(\text{True}) = \mathbb{P}_k(\text{False}) = o(1)$; while if $\mathbb{E}[X^2] = \Theta(k)$ then $\mathbb{P}_k(\text{True}) = \mathbb{P}_k(\text{False}) = \Theta(1)$.

Corollary 5.3. *Suppose that the assumptions of Lemma 5.2 are satisfied. If additionally, we assume that $\limsup_{k \to +\infty} \mathbb{E}L^2 < +\infty$, then asymptotically as k tends to infinity,*

$$\mathbb{P}_k(\text{True}) = \mathbb{P}_k(\text{False}) = \Theta(1/k).$$

Proof. The upper bound is straightforward from Lemma 5.2 and from the additional hypothesis. For the lower bound observe that $\limsup_{k\to+\infty}\mathbb{E}L^2 < +\infty$ implies that there exists an integer $K > 0$ such that the probability that $L \leqslant K$ is greater than a positive constant c (note that these two constants do not depend on k). Remark as well that, since $\mathtt{trim}(\hat{A})$ is smaller than A, $N_A \geqslant N_{\mathtt{trim}(\hat{A})}$ and $C_A \leqslant C_{\mathtt{trim}(\hat{A})}$, which implies

$$\frac{1}{k}\,\mathbb{E}\left[\frac{N_A}{2^{C_A}}\right] \geqslant \frac{1}{k}\,\mathbb{E}\left[\frac{N_{\mathtt{trim}(\hat{A})}}{2^{C_{\mathtt{trim}(\hat{A})}}}\right] \geqslant \frac{c}{k2^K} = \Omega(1/k). \qquad \square$$

Proof of Lemma 5.2. The constant function True is represented by any tree that has two leaves labelled by some variable and its negation both connected to the root by a path of $\vee$ connectives. We thus have

$$\mathbb{P}_k(\text{True}) \geqslant \frac{1}{k}\sum_{a\in\mathrm{Supp}A}\mathbb{P}(A = a)\cdot\frac{N_A}{2^{C_A}} = \frac{1}{k}\,\mathbb{E}\left[\frac{N_A}{2^{C_A}}\right].$$

Let us now prove the upper bound. Assuming that $\hat{T}_\infty$ calculates True implies that

- if the root is $\wedge$-labelled, then the conjunction of the subtrees af $\hat{A}_0$ calculates True;

- if the root is $\vee$-labelled,

  - either the disjunction of the subtrees of $\hat{A}_0$ calculates True;
  - or the infinite subtree of the root of $\hat{T}_\infty$ calculates True;
  - or the disjunction of the subtrees of $\hat{A}_0$ calculates a non-constant function η, the infinite subtree of the root calculates a non-constant function f, and $\eta \vee f$ is the constant function True: in this case, we say that there is *compensation*.

Assumption (H) tells us that the infinite subtree of the root of $\hat{T}_\infty$ is distributed as $\hat{T}_\infty$. Thus, the above implication tells us that

$$\mathbb{P}(\mathrm{f}[\hat{T}_\infty] = \text{True}) \leqslant \frac{1}{2}\mathbb{P}(\mathrm{f}[\hat{A}_\vee] = \text{True}) + \frac{1}{2}\mathbb{P}(\mathrm{f}[\hat{T}_\infty] = \text{True})$$
$$+ \frac{1}{2}\mathbb{P}(\mathrm{f}[\hat{A}_\wedge] = \text{True}) + \frac{1}{2}\mathbb{P}(\text{compensation}),$$

where $\mathrm{f}[\hat{A}_\vee]$ (resp. $\mathrm{f}[\hat{A}_\wedge]$) stands for the Boolean function calculated by the disjunction (resp. conjunction) of the subtrees of the forest $\hat{A}$. Using the fact that $\mathbb{P}(\mathrm{f}[\hat{A}_\wedge] = \text{True}) \leqslant \mathbb{P}(\mathrm{f}[\hat{A}_\vee] = \text{True})$, we obtain

$$\mathbb{P}(\mathrm{f}[\hat{T}_\infty] = \text{True}) \leqslant 2\mathbb{P}(\mathrm{f}[\hat{A}_\vee] = \text{True}) + \mathbb{P}(\text{compensation}). \tag{19}$$

Note that the event $\{\mathrm{f}[\hat{A}_\vee] = \text{True}\} = \{\mathrm{f}[\mathtt{trim}(\hat{A})_\vee] = \text{True}\}$ is contained in the event "at least two leaves of $\mathtt{trim}(\hat{A})$ are labelled by the same variable":

$$\mathbb{P}(\mathrm{f}[\hat{A}_\vee] = \text{True}) \leqslant \mathbb{P}(\mathtt{trim}(\hat{A})\text{ contains at least 1 repetition}),$$

which implies, in view of Lemma 5.1:

$$\mathbb{P}\left(\mathrm{f}[\hat{A}_\vee] = \text{True}\right) \leqslant \frac{2\mathrm{e}\mathbb{E}[\|\mathtt{trim}(\hat{A})\|] + \mathbb{E}[\|\mathtt{trim}(\hat{A})\|^2]}{k} \leqslant \frac{2\mathrm{e}\mathbb{E}L + \mathbb{E}L^2}{k}. \tag{20}$$

Let us now study the probability of the event $\{\text{compensation}\}$. On this event, the disjunction of the subtrees of $\hat{A}_1$ calculates a non-constant function, which we denote by η. Note that η has then at least one essential variable. Let us denote by f the random Boolean function calculated by the infinite subtree (rooted at ν_2). Let us prove that the event $\{\eta \vee f \equiv \text{True}\}$ is contained into the event "at least one essential variable of η is an essential variable for f". Assume for a contradiction that no essential variable of η is essential for f: then there exists an assignation w of the essential variables of η such that $\eta_{|w} = \text{False}$ and $f_{|w} \neq \text{True}$ thus, $f \vee \eta_{|w} \neq \text{True}$, which is impossible since $f \vee \eta_{|w} = \text{True}$ by assumption. Thus there exists at least one essential variable of η which is also essential for f. Let us denote by Γ the (random) number of essential variables of η. Note that, with probability one, $\Gamma \leqslant L_1$, and that L_1 is (unconditioned and) distributed as L. Then,

$$\mathbb{P}(f \vee \eta = \text{True and compensation}) \leqslant \sum_{\gamma=1}^{k}\mathbb{P}(\Gamma = \gamma)\sum_{j=1}^{\gamma}\mathbb{P}(x_j\text{ is an essential variable of }\eta)$$

24

$$\leqslant \sum_{\gamma=1}^{k} \mathbb{P}(\Gamma = \gamma)\frac{\gamma}{k}$$

$$\leqslant \frac{\mathbb{E}L}{k}. \tag{21}$$

Combining Equations (20), (21) and (19) completes the proof. □

Example: the associative trees. We already introduced this example in Section 3.3.3: Let $G_{n,k}$ be a random tree uniformly distributed among trees having n nodes (leaves and internal nodes), labelled with k variables and such that no internal node has a unique child. Forget the labels of $G_{n,k}$, it gives a random non-labelled tree $T_{n,k}$ such that $\hat{T}_{n,k}$ and $G_{n,k}$ have the same distribution. In this case, $T_{n,k}$ really depends on k as so does the random variable $A = A_k$ (being the common law of the random forests of trees hanging on the infinite spine). We have shown in Subsection 3.3 that $T_{n,k}$ is a critical Galton–Watson tree. Using Equation (5), one can check that, with high probability when $k \to +\infty$, $C_A = 0$, and that $\mathbb{E}[N_A 2^{-C_A}] \sim k$ as k tends to infinity. Lemma 5.2 applies and gives that $\mathbb{P}_k(\text{True}) = \Theta(1)$ (a result already proved in [20]).

5.3 Probability of a given Boolean function

In this section, we aim at proving the analog of the result of Kozik [26], namely estimate $\mathbb{P}_k(f)$ for any given Boolean function f. To do so, we will need additional assumptions: we still assume (H) and additionally require that A and L do *not* depend on k (recall that the local limit may depend on k).

Given a Boolean function f, and a random forest Υ, we denote by $L_\Upsilon(f)$ the effective complexity of f according to Υ as follows. Let $\mathcal{M}_\Upsilon(f)$ be the set of forests in the support of Υ that can be labelled so that the disjunction of their subtrees calculates the function f; $L_\Upsilon(f)$ is the the size of the smallest forests in $\mathcal{M}_\Upsilon(f)$. In the following theorem, the quantity $L_A(f)$ is used; we recall that, under (H), A is a random forest whose distribution is the joint distribution of the A_i's, the forests hanging on the spine of the local limit of $(T_{k,n})_{n \geqslant 1}$.

For any tree t and any integer m, we denote by $N_m(t)$ the number of nodes that are at distance at most m of the root in t.

Theorem 5.4. *Suppose that for $k \geqslant 1$, $(T_{k,n})_{n \geqslant 0}$ is a sequence of unlabelled random trees converging in distribution to a local limit T_∞ with a unique end, satisfying* (H)*, and such that A and L do not depend on k. Suppose further that for all integer $m \geqslant 1$,*

(i) $\mathbb{E}[N_m^{m+2}(T_\infty)] < +\infty$, and

(ii) $\mathbb{E}[\|\mathtt{trim}(\hat{T}_\infty)\|^m] < \infty$.

Let k_0 be an integer independent of k and let f be a k_0-variable Boolean function, then there exists constants $c, C \in (0, \infty)$ such that for all k large enough

$$c \cdot \frac{\mathbb{P}_k(\text{True})}{k^{L_A(f)}} \leqslant \mathbb{P}_k(f) \leqslant C \cdot \frac{\mathbb{P}_k(\text{True}) + 1/k}{k^{L(f)}}.$$

Note that Assumption (ii) implies in particular that L has finite moments of all orders, and thus, under the assumptions of Theorem 5.4, $\mathbb{P}_k(\text{True}) = \Theta(1/k)$, in view of Corollary 5.3. Also note that in all examples considered in this article, the support of the random variable A will be such that $L_A \equiv L$ on the set of Boolean function, and the following Corollary applies to these examples:

Corollary 5.5. *Under the assumptions of Theorem 5.4 and assuming that $L_A(f) = L(f)$, we have*

$$\mathbb{P}_k(f) = \Theta\left(\frac{1}{k^{L(f)+1}}\right).$$

Proof of Theorem 5.4. If $\hat{T}_\infty$ is $\vee$-rooted, if the infinite subtree rooted at u_1 calculates False (recall that u_1 is the second node on the infinite spine, u_0 being the root of $\hat{T}_\infty$), and if the disjunction of the subtrees of A_0 calculates f, then $\hat{T}_\infty$ calculates f. Let t be a forest in $\mathcal{M}_A(f)$. Since the number of internal nodes is maximized when t is a binary tree, $\frac{1}{k^{\|t\|}} \frac{1}{2^{\|t\|-1}}$ is a lower bound of the probability that t is labelled such that $\hat{t}$ calculates f, and the probability that $\hat{A}_0$ calculates f is at least $\mathbb{P}(A = t)\frac{1}{k^{\|t\|}} \frac{1}{2^{\|t\|-1}}$. As the labels of disjoints sets of nodes are independent, the lower bound is thus proved.

We now focus on the upper bound. If $\hat{T}_\infty$ represents f, then $\mathtt{trim}(\hat{T}_\infty)$ also computes the function f, which implies that its size is at least $L(f)$. Thus in order to prove the upper bound, it suffices to show that

$$\mathbb{P}(\mathrm{f}[\hat{T}_\infty] = f \text{ and } \|\mathtt{trim}(\hat{T}_\infty)\| \geqslant L(f)) = \mathcal{O}\left(\frac{1}{k^{L(f)+1}}\right).$$

Let us first prove that

$$\mathbb{P}(\mathrm{f}[\hat{T}_\infty] = f \text{ and } \|\mathtt{trim}(\hat{T}_\infty)\| = L(f)) = \mathcal{O}\left(\frac{1}{k^{L(f)+1}}\right). \tag{22}$$

Recall that the number of repetitions is formally defined as the difference between the number of leaves and the number of pairwise different variables appearing as labels of these leaves (in their positive or negated form). Assume that $\mathtt{trim}(\hat{T}_\infty)$ has size $L(f)$ and that $d \geqslant 1$ leaves of $\mathtt{trim}(\hat{T}_\infty)$ are labelled by a non-essential variable (or its negation). Assign all the non-essential variables appearing in $\mathtt{trim}(\hat{T}_\infty)$ to True, and simplify the tree according to Boolean logic, i.e. using the following four simplifying rules: for every Boolean function h

$$\text{True} \wedge h = h \qquad \text{True} \vee h = \text{True} \qquad \text{False} \wedge h = \text{False} \qquad \text{False} \vee h = h.$$

Since we have only assigned values to non-essential variables, the tree obtained still calculates f, but since we have removed d leaves, it has size at most $L(f) - d \leqslant L(f) - 1$, which is impossible since $L(f)$ is minimum size of a tree computing f. Therefore, if $\mathtt{trim}(\hat{T}_\infty)$ has size $L(f)$, then all its leaves are labelled by essential variables, which implies that it contains exactly $L(f) - \gamma$ repetitions, where γ denotes the number of essential variables of f.

In other words,

$$\{\mathrm{f}[\hat{T}_\infty] = f \text{ and } \|\mathtt{trim}(\hat{T}_\infty)\| = L(f)\}$$
$$\subseteq \quad \{\mathtt{trim}(\hat{T}_\infty) \text{ has size } L(f) \text{ and contains at least } L(f) - \gamma \text{ repetitions}\}.$$

Note that by symmetry, the above inclusion is true for any Boolean function f having γ essential variables. It thus implies that,

$$\binom{k}{\gamma} \mathbb{P}(\mathrm{f}[\hat{T}_\infty] = f \text{ and } \|\mathtt{trim}(\hat{T}_\infty)\| = L(f))$$
$$\leqslant \mathbb{P}(\mathtt{trim}(\hat{T}_\infty) \text{ has size } L(f) \text{ and contains exactly } L(f) - \gamma \text{ repetitions})$$
$$\leqslant \frac{\mathbb{E}[K_{L(f),L(f)-\gamma}(T_\infty)]}{k^{L(f)+1-\gamma}},$$

in view of Lemma 5.1(b), where $K_{p,q}(t_\infty)$ is defined in Equation (18) for any infinite locally finite tree t_∞ and any integers p and q. Since $\binom{k}{\gamma} \geqslant (k-\gamma)^\gamma$, it is enough to prove that $\mathbb{E}[K_{L(f),L(f)-\gamma}(T_\infty)]$ is a finite constant (independent of k). Note that $|\mathtt{sub}_p(T_\infty)|$ is bounded from above by $\binom{N_p(T_\infty)}{p}$ for all integer p where $N_p(T_\infty)$ is the (random) number of nodes at height at most p in T_∞. Moreover, in view of the definitions of $A_{\mathtt{max}}$ and $F_{\mathtt{max}}$ (see Equations (16) and (17)), we have that $F_{\mathtt{max}}$ is bounded from above by the maximal out-degree of all nodes at height at most $L(f)$ in T_∞, and thus $F_{\mathtt{max}} \leqslant N_{L(f)+1}(T_\infty)$. Note also that $A_{\mathtt{max}} \leqslant L(f)$. It follows that

$$K_{L(f),L(f)-\gamma}(T_\infty) \leqslant (L(f) + N_{L(f)+1}(T_\infty))N_{L(f)+1}(T_\infty)\binom{N_{L(f)}(T_\infty)}{L(f)}\frac{L(f)^{L(f)-\gamma}}{2^{L(f)+1}},$$

and thus has finite expectation in view of Assumption (i). We thus have proved (22)

The second case, when $\|\mathtt{trim}(\hat{T}_\infty)\| > L(f)$ is similar, but simpler. Let us now prove that

$$\mathbb{P}(\mathrm{f}[\hat{T}_\infty] = f \text{ and } \|\mathtt{trim}(\hat{T}_\infty)\| \geqslant L(f) + 1) = \mathcal{O}\left(\frac{1}{k^{L(f)+1}}\right) = \mathcal{O}\left(\frac{\mathbb{P}_k(\text{True}) + 1/k}{k^{L(f)}}\right). \tag{23}$$

Assume that $\mathtt{trim}(\hat{T}_\infty)$ has $d \geqslant p - L(f)$ leaves labelled by variables that are non-essential for f. Take ℓ the left-most one, and denote by ν its closer ancestor having arity at least 2. Assign ℓ to True if its ν is labelled by $\vee$ or to False if ν is labelled by $\wedge$. This permit to assign ν to True (resp. False) and to cut all its children (among which there is at least one other leaf since ν has arity at least 2). Assign all other non-essential variables appearing in $\mathtt{trim}(\hat{T}_\infty)$ to True (this assigns some of the leaves to True, and others to False depending on the polarity of the literal labelling them). This operation does not change the function calculated by the tree and after simplification, the obtained tree has size at most $p - (d + 1) \leqslant L(f) - 1$, which is impossible.

Thus, $\texttt{trim}(\hat{T}_\infty)$ contains at least $L(f) + 1$ leaves labelled by essential variables of f, and since f has γ essential variables, it contains at least $L(f) + 1 - \gamma$ repetitions. It follows that

$$\{\|\texttt{trim}(\hat{T}_\infty)\| \geqslant L(f) + 1 \text{ and } \texttt{f}[\hat{T}_\infty] = f\}$$
$$\subseteq \{\texttt{trim}(\hat{T}_\infty) \text{ contains at least } L(f) + 1 - \gamma \text{ repetitions}\}.$$

The above inclusion is true for any Boolean function having γ essential variables, which implies that

$$\binom{k}{\gamma} \mathbb{P}\left(\texttt{trim}(\hat{T}_\infty) \text{ contains at least } L(f) + 1 - \gamma \text{ repetitions}\right)$$
$$\leqslant \frac{\mathbb{E}[\|\texttt{trim}(\hat{T}_\infty)\|^{2(L(f)+1-\gamma)}] + 2\mathrm{e}\,\mathbb{E}[\|\texttt{trim}(\hat{T}_\infty)\|^{L(f)+1-\gamma}]}{k^{L(f)+1-\gamma}},$$

in view of Lemma 5.1, which concludes the proof since $\binom{k}{\gamma} \geqslant (k - \gamma)^\gamma$ and all moments of $\|\texttt{trim}(\hat{T}_\infty)\|$ are finite. $\qquad\square$

5.4 Examples

We show here how to apply Theorem 5.4 to different random trees: we first consider critical Galton–Watson trees, and then the Ford's alpha tree.

5.4.1 Application to Galton–Watson trees (Proof of Theorem 2.4)

For all $n \geqslant 1$, let T_n be a critical Galton–Watson tree conditioned to have size n. Let us denote by ξ its reproduction random variable: in particular, we have $\mathbb{E}\xi = 1$. We also assume that there exists a positive constant a such that $\mathbb{E}\mathrm{e}^{a\xi} < +\infty$ and that $\mathbb{P}(\xi = 1) = 0$.

It is known from the literature, and mentioned earlier in Section 3.3.1 that $(T_n)_{n\geqslant 1}$ converges locally to an infinite random tree T_∞ having a unique end, on which are hanging some independent copies of the critical, unconditioned (and thus almost surely finite) Galton–Watson trees of reproduction ξ. Thus, assumption (H) is satisfied.

Moreover, the proof of Proposition 4.4 tells us that the k-trimmed subtree $\tau = \texttt{trim}(\hat{T}_\infty)$ of $\hat{T}_\infty$ is such that there exists a constant $c > 0$ such that, for all $x > 0$,

$$\mathbb{P}(\|\tau\| \geqslant x) \leqslant \exp\left(-cx/k^2\right),$$

as long as there exists $a > 0$ such that $\mathbb{E}\mathrm{e}^{a\xi} < \infty$, which is assumed here.

Thus, for all $m \geqslant 1$, for all $x \geqslant 0$,

$$\mathbb{P}(\|\texttt{trim}(\hat{T}_\infty)\|^m \geqslant x) \leqslant \exp\left(\frac{-cx^{1/m}}{k^2}\right),$$

which implies that

$$\mathbb{E}\|\texttt{trim}(\hat{T}_\infty)\|^m < +\infty,$$

which proves that these Galton–Watson trees verify assumption (ii) of Theorem 2.3.

Let us now check that Assumption (i) is also verified: By assumption, ξ has exponential moments, implying that $\hat{\xi}$ also has exponential moments. Note that, for all integers m

$$\Delta_m := N_{m+1}(T_\infty) - N_m(T_\infty) = \sum_{i=1}^{\Delta_{m-1}-1} \xi_i^{(m)} + \hat{\xi}^{(m)} + 1,$$

where the $\xi_i^{(m)}$'s are independent copies of ξ and $\hat{\xi}^{(m)}$ is a copy of ξ. Therefore, we can prove by induction on m that, for all integers m, $N_m(T_\infty)$ has exponential moments, which implies that Assumption (i) holds.

Therefore, Theorem 5.4 applies to this class of Galton–Watson trees, including the Catalan tree model as a particular case, which proves Theorem 2.4.

5.4.2 Application to the alpha model (proof of Theorem 2.5)

We consider in this section the alpha–gamma model. For technical reasons, we restrict ourselves to the case $\alpha = \gamma$, although Theorem 2.5 might also hold for other values of γ. When $\alpha = \gamma$, the probability that the random tree of size n – called the alpha tree of size n – has a left subtree of size $k \in \{1, \ldots, n-1\}$ is given by (see Equation (6) in which we take $\gamma = \alpha$)

$$q_n^\alpha(k) = \frac{\Gamma(k-\alpha)\Gamma(n-k-\alpha)}{\Gamma(n-\alpha)\Gamma(1-\alpha)} \left[\frac{\alpha}{2} \binom{n}{k} + (1-2\alpha) \binom{n-2}{k-1} \right].$$

The following lemmas aim at proving that the alpha model verifies Assumption (ii) of Theorem 5.4.

Lemma 5.6. *For all $n \geq 1$, $q_n^\alpha(1) \geq \alpha/2$.*

Proof. We have:

$$q_n^\alpha(1) = \frac{\Gamma(n-1-\alpha)}{\Gamma(n)} \left(\frac{\alpha n}{2} + (1-2\alpha) \right) = \frac{\alpha}{2} \frac{n + \frac{2(1-2\alpha)}{\alpha}}{n-1-\alpha} \geq \frac{\alpha}{2},$$

for all integers n. $\qquad\square$

As a preliminary, let us also prove that

Lemma 5.7. *For all $\delta > 0$, there exists $n_0 \in \mathbb{N}$ (independent of δ) such that, for all $n \geq n_0$,*

$$\sum_{\delta n \leq k \leq (1-\delta)n} q_n^\alpha(k) \leq \frac{C}{2}\, \delta^{-2(\alpha+1)} n^{-\alpha}, \qquad \text{where} \qquad C = 2\, \frac{\max(\alpha/2, 1-3\alpha/2)}{\Gamma(1-\alpha)}.$$

Proof. For all $\delta > 0$, asymptotically when $n \to +\infty$,

$$\sum_{\delta n \leq k \leq (1-\delta)n} q_n^\alpha(k) = \sum_{\delta n \leq k \leq (1-\delta)n} \frac{\Gamma(k-\alpha)\Gamma(n-k-\alpha)}{\Gamma(n-\alpha)\Gamma(1-\alpha)} \left[\frac{\alpha}{2}\binom{n}{k} + (1-2\alpha)\binom{n-2}{k-1} \right]$$

$$= \frac{e^{\alpha+1} + o(1)}{\Gamma(1-\alpha)} \sum_{\delta n \leq k \leq (1-\delta)n} \frac{(k-\alpha-1)^{k-\alpha-1/2}(n-k-\alpha-1)^{n-k-\alpha-1/2}}{(n-\alpha-1)^{n-\alpha-1/2}}$$

$$\left[\frac{\alpha}{2} \frac{n^{n+1/2}}{k^{k+1/2}(n-k)^{n-k+1/2}} + (1-2\alpha) \frac{(n-2)^{n-3/2}}{(k-1)^{k-1/2}(n-k-1)^{n-k-1/2}} \right]$$

$$= \frac{\max(\alpha/2, 1-3\alpha/2) + o(1)}{\Gamma(1-\alpha)} \sum_{\delta n \leq k \leq (1-\delta)n} \left(\frac{n}{k(n-k)} \right)^{\alpha+1}$$

$$\leq \frac{C}{2}(1 + o(1))\, n \left(\frac{n}{(\delta n)^2} \right)^{\alpha+1}$$

where $C = 2\,\frac{\max(\alpha/2, 1-3\alpha/2)}{\Gamma(1-\alpha)}$ does not depend on δ. $\qquad\square$

Let us define the following cutting procedure: a node having at least one child of size 1 is cut with probability $\varepsilon > 0$. Let us denote by $\tau(n)$ the size of the alpha tree of size n after having applied this cut procedure. We are going to prove that all the moments of $\tau(n)$ are finite.

Let us start with its expectation:

Lemma 5.8. *There exists a constant M_1 such that, for all $n \geq 1$, $\mathbb{E}\tau(n) \leq \frac{M_1 n^\alpha}{1 + \ln^2 n}$.*

Proof. We prove this result by induction.

Assume that, for all $k \leq n-1$, $\mathbb{E}\tau(k) \leq \frac{M_1 k^\alpha}{1+\ln^2 k}$. Using the induction hypothesis,

$$\mathbb{E}\tau(n) = \sum_{k=2}^{n-2} q_n^\alpha(k) \left(\mathbb{E}\tau(k) + \mathbb{E}\tau(n-k) \right) + (1-\varepsilon)2q_n^\alpha(1)\left(1 + \mathbb{E}\tau(n-1) \right)$$

$$\leq M_1 \sum_{k=2}^{n-2} q_n^\alpha(k) \left(\frac{k^\alpha}{1+\ln^2 k} + \frac{(n-k)^\alpha}{1+\ln^2(n-k)} \right) + (1-\varepsilon)2q_n^\alpha(1)\left(1 + \frac{M_1(n-1)^\alpha}{1+\ln^2(n-1)} \right)$$

28

$$= M_1 \sum_{k=1}^{n-1} q_n^\alpha(k) \left(\frac{k^\alpha}{1 + \ln^2 k} + \frac{(n-k)^\alpha}{1 + \ln^2(n-k)} \right) - 2M_1 \varepsilon q_n^\alpha(1) \left(1 + \frac{(n-1)^\alpha}{1 + \ln^2(n-1)} \right)$$
$$+ 2(1 - \varepsilon) q_n^\alpha(1)(1 - M_1).$$

First note that
$$\left(1 + \frac{(n-1)^\alpha}{1 + \ln^2(n-1)} \right) \geqslant \frac{1 + (n-1)^\alpha}{1 + \ln^2 n} \geqslant \frac{n^\alpha}{1 + \ln^2 n}.$$

Moreover, for all $\delta > 0$, for C the constant of Lemma 5.7 and all n large enough

$$\sum_{k=1}^{n-1} q_n^\alpha(k) \left(\frac{k^\alpha}{1 + \ln^2 k} + \frac{(n-k)^\alpha}{1 + \ln^2(n-k)} \right)$$

$$= \sum_{\delta n \leqslant k \leqslant (1-\delta)n} q_n^\alpha(k) \left(\frac{k^\alpha}{1 + \ln^2 k} + \frac{(n-k)^\alpha}{1 + \ln^2(n-k)} \right) + \sum_{k < \delta n \text{ or } k > (1-\delta)n} q_n^\alpha(k) \left(\frac{k^\alpha}{1 + \ln^2 k} + \frac{(n-k)^\alpha}{1 + \ln^2(n-k)} \right)$$

$$\leqslant \frac{2[(1-\delta)n]^\alpha}{1 + \ln^2(\delta n)} \sum_{\delta n \leqslant k \leqslant (1-\delta)n} q_n^\alpha(k) + \left(\frac{(\delta n)^\alpha}{1 + \ln^2[(1-\delta)n]} + \frac{n^\alpha}{1 + \ln^2[(1-\delta)n]} \right) \sum_{k < \delta n \text{ or } k > (1-\delta)n} q_n^\alpha(k)$$

$$\leqslant \frac{2C[(1-\delta)n]^\alpha}{1 + \ln^2(\delta n)} \delta^{-2(1+\alpha)} n^{-\alpha} + [\delta^\alpha + 1] \frac{1 + \ln^2 n}{1 + \ln^2[(1-\delta)n]} \frac{n^\alpha}{1 + \ln^2 n}$$

$$\leqslant 2C(1-\delta)^\alpha \delta^{-2(1+\alpha)} + (\delta^\alpha + 1) \frac{1 + \ln^2 n}{1 + \ln^2[(1-\delta)n]} \frac{n^\alpha}{1 + \ln^2 n},$$

where we have applied Lemma 5.7. In total, we thus get:

$$\mathbb{E}\tau(n) \leqslant \frac{M_1 n^\alpha}{1 + \ln^2 n} \left[(\delta^\alpha + 1) \frac{1 + \ln^2 n}{1 + \ln^2[(1-\delta)n]} - 2\varepsilon q_n^\alpha(1) \right] + M_1 \left(2C\delta^{-2(\alpha+1)} - 2(1-\varepsilon) q_n^\alpha(1)) \right)$$
$$+ 2(1 - \varepsilon) q_n^\alpha(1).$$

Let us first fix δ such that $-2\eta := \delta^\alpha - \varepsilon\alpha < 0$. It implies that, in view of Lemma 5.6,

$$(\delta^\alpha + 1) \frac{1 + \ln^2 n}{1 + \ln^2[(1-\delta)n]} - 2\varepsilon q_n^\alpha(1) \leqslant (\delta^\alpha + 1) \frac{1 + \ln^2 n}{1 + \ln^2[(1-\delta)n]} - 2\varepsilon \frac{\alpha}{2} \to \delta^\alpha + 1 - 2\varepsilon q_n^\alpha(1) = 1 - 2\eta,$$

when $n \to +\infty$. Thus, there exists n_δ such that, for all $n \geqslant n_\delta$,

$$(\delta^\alpha + 1) \frac{1 + \ln^2 n}{1 + \ln^2[(1-\delta)n]} - 2\varepsilon q_n^\alpha(1) < 1 - \eta.$$

Thus, for all $n \geqslant n_\delta$,

$$\mathbb{E}\tau(n) \leqslant \frac{M_1 n^\alpha}{1 + \ln^2 n} + M_1 \left(2C\delta^{-2(\alpha+1)} - 2(1-\varepsilon) q_n^\alpha(1) - \frac{\eta n^\alpha}{1 + \ln^2 n} \right) + 2(1 - \varepsilon) q_n^\alpha(1).$$

There exists $\tilde{n}_\delta > n_\delta$ such that, for all $n \geqslant \tilde{n}_\delta$,

$$2C\delta^{-2(\alpha+1)} - 2(1-\varepsilon) q_n^\alpha(1) - \frac{\eta n^\alpha}{1 + \ln^2 n} < -1,$$

thus, choosing

$$M_1 := 2 + \max_{n \leqslant \tilde{n}_\delta} \left\{ \frac{1 + \ln^2 n}{n^\alpha} \mathbb{E}\tau(n) \right\}$$

we have that, for all $n \geqslant 1$,

$$\mathbb{E}\tau(n) \leqslant \frac{M_1 n^\alpha}{1 + \ln^2 n} - M_1 + 2(1 - \varepsilon) \leqslant \frac{M_1 n^\alpha}{1 + \ln^2 n},$$

which concludes the proof. $\qquad \square$

Lemma 5.9. *There exists a constant $K > 0$ such that, for all $n \geqslant 2$, $\mathbb{E}\tau(n) \leqslant K$.*

Proof. We have proved in Section 3.3 that the α tree converges locally to an infinite spine on which are attached alpha trees of random sizes $(N_i)_{i \geq 0}$, from top to bottom in the tree. The sequence $(N_i)_{i \geq 0}$ is a sequence of i.i.d. random variables having law

$$\mathbb{P}(N_i = k) = \frac{\alpha \Gamma(k - \alpha)}{k! \Gamma(1 - \alpha)} \text{ for all } k \geq 1.$$

Let us denote by $\tau(\infty)$ the size of this infinite tree after applying it the cutting procedure explained above. First note that for all $i \geq 0$ $\mathbb{P}(N_i = 1) = \alpha$. Thus, if we explore the spine from top to bottom, we cut the spine between u_i and u_{i+1} with probability at least $\alpha \varepsilon$ independently for all $i \geq 0$. Thus, the spine is cut at height at most H where H is a geometric random variable with parameter $\varepsilon \alpha$. Thus,

$$\mathbb{E}\tau(\infty) = \mathbb{E}H \mathbb{E}\tau(N_1) = \frac{1}{\alpha \varepsilon} \sum_{k \geq 1} \frac{\alpha \Gamma(k - \alpha)}{k! \Gamma(1 - \alpha)} \mathbb{E}\tau(k) \leq \frac{1}{\alpha \varepsilon} \sum_{k \geq 1} \frac{\alpha \Gamma(k - \alpha)}{k! \Gamma(1 - \alpha)} \frac{k^\alpha}{1 + \ln^2 k}.$$

Note that

$$\frac{\alpha \Gamma(k - \alpha)}{k! \Gamma(1 - \alpha)} \frac{k^\alpha}{1 + \ln^2 k} \sim \frac{\alpha}{\Gamma(1 - \alpha)} \frac{1}{k \ln^2 k},$$

when $k \to +\infty$. Thus, it is summable, implying that $\mathbb{E}\tau(\infty) < +\infty$. Finally remark that, for all $n \geq 2$,

$$\mathbb{E}\tau(n) \leq \mathbb{E}\tau(\infty) =: K,$$

which concludes the proof. $\qquad \square$

Lemma 5.10. *For all $p \geq 1$, there exists a constant K_p such that, for all $n \geq 2$, $\mathbb{E}\tau(n)^p \leq K_p$.*

Proof. Lemma 5.9 is the case $p = 1$. For all $p \geq 1$, $\mathbb{E}\tau(1)^p = 1$.
Let us fix p, and n, and assume that,

- for all $q \leq p$, for all $n \geq 2$, $\mathbb{E}\tau(n)^q \leq K_q$; and that,

- Let

$$M_p := \frac{\sum_{i=1}^{p-1} \binom{p}{i} K_i K_{p-i}}{2\varepsilon q_n^\alpha(1)},$$

and assume that, for all $k \leq n$, $\mathbb{E}\tau(k)^p \leq \frac{M_p k^\alpha}{1 + \ln^2 k}$.

Applying the induction hypothesis, we have:

$$\mathbb{E}\tau(n)^p = \sum_{k=2}^{n-2} q_n^\alpha(k) \mathbb{E}[(\tau(k) + \tau(n - k))^p] + (1 - \varepsilon)2q_n^\alpha(1) \mathbb{E}[(1 + \tau(n - 1))^p]$$

$$= \sum_{k=2}^{n-2} q_n^\alpha(k) \sum_{i=0}^{p} \binom{p}{i} \mathbb{E}\tau(k)^i \mathbb{E}\tau(n - k)^{(p - i)} + (1 - \varepsilon)2q_n^\alpha(1) \sum_{i=0}^{p} \binom{p}{i} \mathbb{E}[\tau(n - 1)^i]$$

$$\leq \sum_{k=2}^{n-2} q_n^\alpha(k) \sum_{i=1}^{p-1} \binom{p}{i} K_i K_{p-i} + M_p \sum_{k=2}^{n-2} q_n^\alpha(k) \left(\frac{k^\alpha}{1 + \ln^2 k} + \frac{(n - k)^\alpha}{1 + \ln^2(n - k)} \right)$$

$$+ (1 - \varepsilon)2q_n^\alpha(1) \sum_{i=1}^{p-1} \binom{p}{i} K_i + M_p(1 - \varepsilon)2q_n^\alpha(1) \left(1 + \frac{(n - 1)^\alpha}{1 + \ln^2(n - 1)} \right)$$

$$= \sum_{i=1}^{p-1} \binom{p}{i} K_i K_{p-i} - 2\varepsilon q_n^\alpha(1) \sum_{i=1}^{p-1} \binom{p}{i} K_i$$

$$+ M_p \left[\sum_{k=1}^{n-1} q_n^\alpha(k) \left(\frac{k^\alpha}{1 + \ln^2 k} + \frac{(n - k)^\alpha}{1 + \ln^2(n - k)} \right) - 2\varepsilon q_n^\alpha(1) \left(1 + \frac{(n - 1)^\alpha}{1 + \ln^2(n - 1)} \right) \right]$$

Reusing the calculations made in the proof of Lemma 5.9, we get that:

$$\mathbb{E}\tau(n)^p \leq \sum_{i=1}^{p-1} \binom{p}{i} K_i K_{p-i} + (1 - 2\varepsilon q_n^\alpha(1)) \frac{M_p n^\alpha}{1 + \ln^2 n}$$

30

$$\leqslant \frac{M_p n^\alpha}{1 + \ln^2 n} + \sum_{i=1}^{p-1} \binom{p}{i} K_i K_{p-i} - 2\varepsilon q_n^\alpha(1) M_p$$

$$\leqslant \frac{M_p n^\alpha}{1 + \ln^2 n},$$

in view of the value of M_p.

We have thus proved that there exists a constant M_p such that, for all $n \geqslant 2$, $\mathbb{E}\tau(n)^p \leqslant \frac{M_p n^\alpha}{1+\ln^2 n}$. To prove that this implies that $\mathbb{E}\tau(n)^p < +\infty$, let us recall that the local limit of the alpha-tree is an infinite spine on which are hung some independent alpha-trees of respective sizes $(N_i)_{i\geqslant 1}$. We have already noticed that if we apply the cutting procedure to this infinite tree, the spine is cut at level H, where H is a geometric random variable of parameter $\alpha\varepsilon$. Note that

$$\mathbb{E}\tau(N_1)^p = \sum_{k\geqslant 1} \frac{\alpha\Gamma(k-\alpha)}{k!\Gamma(1-\alpha)}\mathbb{E}\tau(k)^p \leqslant \frac{M_p \alpha}{\Gamma(1-\alpha)} \sum_{k\geqslant 1} \frac{\Gamma(k-\alpha)}{k!} \frac{k^\alpha}{1+\ln^2 k} < +\infty.$$

If we denote by $\tau(\infty)$ the size of the obtained tree, then

$$\mathbb{E}\tau(\infty)^p = \mathbb{E}\left(\sum_{i=1}^H \tau(N_i)\right)^p \leqslant \mathbb{E}\left[H^{p-1}\sum_{i=1}^H \tau(N_i)^p\right] = \mathbb{E}H^p \mathbb{E}\tau(N_1)^p =: K_p < +\infty,$$

which concludes the proof because $\mathbb{E}\tau(n)^p \leqslant \mathbb{E}\tau(\infty)^p$ for all $n \geqslant 2$. $\qquad\square$

We are now ready to prove that the alpha-tree verifies the hypothesis of Theorem 5.4, by relating the idealized cutting procedure described above to the trimming procedure. We have already proved (see Section 3.3.4) that the alpha tree locally converges to an infinite tree T_∞ which consists of an infinite spine to which are attached independent alpha trees of random i.i.d. sizes $(N_i)_{i\geqslant 0}$ where

$$\mathbb{P}(N_1 = k) = \frac{\alpha\Gamma(k-\alpha)}{k!\Gamma(1-\alpha)}.$$

Hypothesis (H) is thus verified.

The alpha tree is by construction binary (see Figure 2), and thus, Assumption (i) of Theorem 5.4 is trivially verified: for all integer $m \geqslant 1$, $N_m^{m+2}(T_\infty) \leqslant 2^{m(m+2)} < +\infty$.

Let us couple the trimming procedure applied on $\hat{T}_\infty$ to the following procedure. Let $\varepsilon = 1/2k$. At first, all nodes of the local limit of the alpha tree T_∞ are black: starting from the root,

- if a black node has a child of size 1, then colour it and all its descendent in red; otherwise, colour it in blue;

- when all nodes are coloured either in blue or in red, apply the cutting procedure above for all red parts of the tree: a red node having a child of size one is cut with probability ε.

One can couple the above procedure with `trim` so that the size of the obtained tree by the procedure above is larger than the size of $\mathtt{trim}(\hat{T}_\infty)$. The black nodes correspond to the nodes in $\hat{T}_\infty$ having empty constraints sets, the red nodes correspond to nodes having a non-empty constraint set: thus each leaf can make all its siblings inconsistent with probability at least $1/2k$. Since the internal nodes that were parent of a set of inconsistent siblings do not contribute to the size of $\mathtt{trim}(\hat{T}_\infty)$, cutting them together with their children does not affect the size of $\mathtt{trim}(\hat{T}_\infty)$ (but it might change the Boolean function calculated by the tree). Thus, the above procedure gives a random tree τ whose size stochastically dominates $\mathtt{trim}(\hat{T}_\infty)$.

Moreover this tree τ consists in a blue tree to whose leaves are attached some red trees. Let us denote by $\|\mathtt{blue}\|$ the number of the blue nodes in the above tree: Lemma 5.10 applied in the special case with $\varepsilon = 1$ gives us that all the moments of $\|\mathtt{blue}\|$ are finite. Then, the red trees attached to the leaves of the blue tree are i.i.d. There is at most $\|\mathtt{blue}\|$ of them and by Lemma 5.10 applied to $\varepsilon = \frac{1}{2k}$, their sizes all have finite moments.

Therefore, the size of the whole tree (blue nodes and red nodes) has finite moments, which permits to conclude that $\|\mathtt{trim}(\hat{T}_\infty)\|$ has finite moments. We can then apply our main result Theorem 5.4 to the alpha model and deduce Theorem 2.5.

References

[1] R. Abraham and J.-F. Delmas. Local limits of conditioned Galton–Watson trees: the infinite spine case. *Electronic Journal of Probability*, 19(2):1–19, 2014.

[2] D. Aldous. Probability distributions on cladograms. In D. Aldous and R. Pemantle, editors, *Random Discrete Structures*, pages 1–18. Springer, 1996.

[3] D. Aldous and J. M. Steele. The objective method: probabilistic combinatorial optimization and local weak convergence. In H. Kesten, editor, *Discrete and Combinatorial Probability*, pages 1–72. Springer Verlag, 2003.

[4] J. Balogh, Y. Peres, and G. Pete. Bootstrap percolation on infinite trees and non-amenable groups. *Combinatorics, Probability and Computing*, 15:715–730, 2006.

[5] N. Broutin and P. Flajolet. The distribution of the height and diameter in random non-plane binary trees. *Random Structures & Algorithms*, 41:215–252, 2012.

[6] N. Broutin, L. Devroye, E. McLeish, and M. de la Salle. The height of increasing trees. *Random Structures and Algorithms*, 32:494–518, 2008.

[7] B. Chauvin, P. Flajolet, D. Gardy, and B. Gittenberger. And/or trees revisited. *Combinatorics, Probability and Computing*, 13:475–497, 2004.

[8] B. Chauvin, D. Gardy, and C. Mailler. The growing tree distribution on Boolean functions. In *Proceedings of the 8th SIAM Workshop on Analytic and Combinatorics (ANALCO)*, pages 45–56, 2011.

[9] B. Chauvin, D. Gardy, and C. Mailler. A sprouting tree model for random Boolean functions. *Random Structures and Algorithms*, 2015. To appear.

[10] B. Chen, D. Ford, and M. Winkel. A new family of Markov branching trees: the alpha-gamma model. *Electronic Journal of Probability*, 14:400–430, 2009.

[11] L. Comtet. *Advanced Combinatorics*. Reidel Publishing Company, Dordrecht, 1974.

[12] A. Dembo and O. Zeitouni. *Large Deviation Techniques and Applications*. Springer, second edition, 1998.

[13] L. Devroye. A note on the height of binary search trees. *Journal of the ACM*, 33:489–498, 1986.

[14] L. Devroye. Universal limit laws for depth in random trees. *SIAM Journal on Computing*, 28(2):409–432, 1998.

[15] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, Cambridge, UK, 2009.

[16] D.J. Ford. Probabilities on cladograms: introduction to the alpha model. ArXiv:0511246 [math.PR], 2005.

[17] H. Fournier, D. Gardy, and A. Genitrini. Balanced And/Or trees and linear threshold functions. In *5th SIAM Workshop on Analytic and Combinatorics (ANALCO)*, pages 51–57, 2009.

[18] H. Fournier, D. Gardy, A. Genitrini, and B. Gittenberger. The fraction of large random trees representing a given Boolean function in implicational logic. *Random Structures & Algorithms*, 40(3):317–349, 2012.

[19] D. Gardy. Random Boolean functions. In *Computational Logic and Applications (CLA '05)*, volume AF of *DMTCS Proc.*, pages 1–36, 2006.

[20] A. Genitrini, B. Gittenberger, V. Kraus, and C. Mailler. Associative and commutative tree representations for Boolean functions. *Theoretical Computer Science*, 570:70–101, 2015.

[21] G.R. Grimmett. Random labelled trees and their branching networks. *Journal of the Australian Mathematical Society, Series A*, 30:229–237, 1980.

[22] B. Haas and G. Miermont. The genealogy of self-similar fragmentations with negative index as a continuum random tree. *Electronic Journal of Probability*, 9:57–97, 2004.

[23] B. Haas and G. Miermont. Scaling limits of Markov branching trees with applications to Galton–Watson and random unordered trees. *The Annals of Probability*, 40:2589–2666, 2012.

[24] H. Kesten. Subdiffusive behavior of random walk on a random cluster. *Annales de l'Institut Henri Poincaré: Probabilités et Statistiques*, 22:425–487, 1986.

[25] D. E. Knuth. *The Art of Computer Programming: Sorting and Searching*, volume 3. Addison-Wesley, Reading, MA, 1973.

[26] J. Kozik. Subcritical pattern languages for And/Or trees. In *Proceedings of the Fifth Colloquium on Mathematics and Computer Science*, volume AI, pages 437–448. DMTCS Proceedings, 2008.

[27] H. Lefmann and P. Savický. Some typical properties of large and/or Boolean formulas. *Random Structures and Algorithms*, 10(3):337–351, 1997.

[28] R. Lyons, R. Pemantle, and Y. Peres. Conceptual proofs of the $L \log L$ criteria for mean behavior of branching processes. *The Annals of Probability*, (23):1125–1138, 1995.

[29] H. Mahmoud. *Evolution of Random Search Trees*. Wiley, New York, 1992.

[30] P. Marchal. A note on the fragmentation of a stable tree. In *Fifth Colloquium on Mathematics and Computer Science*, volume AI of *DMTCS Proceedings*, pages 489–500, 2008.

[31] J.-F. Marckert and G. Miermont. The CRT is the scaling limit of unordered binary trees. *Random Structures & Algorithms*, 38:467–501, 2011.

[32] A. Meir and J. W. Moon. On the altitude of nodes in random trees. *Canadian journal of Mathematics*, 30(1978): 997–1015, 1978.

[33] R. Otter. On the number of trees. *Annals of Mathematics*, 49:583–599, 1948.

[34] M. S. Paterson, editor. *Boolean Function Complexity*. Cambridge University Press, 1992. ISBN 9780511526633. Cambridge Books Online.

[35] J. Pitman. Exchangeable and partially exchangeable random partitions. *Probability Theory and Related Fields*, 102: 145–158, 1995.

[36] B. Pittel. On growing random binary trees. *Journal of Mathematical Analysis and its applications*, 103:461–480, 1984.

[37] G. Pólya. Kombinatorishe Anzahlbestimmungen fúr Gruppen, Graphen und chemishe Verbindungen. *Acta Mathematica*, 68:145–254, 1937.

[38] J. Riordan and C.E. Shannon. The number of two-terminal series-parallel networks. *Journal of Mathematics and Physics*, 21:83–93, 1942.

[39] A.R. Woods. Coloring rules for finite trees, and probabilities of monadic second order sequences. *Random Structures and Algorithms*, 10:453–485, 1997.