A Central Limit Theorem for Almost Local Additive Tree Functionals

Dimbinaina Ralaivaosaona · Matas Šileikis · Stephan Wagner

Received: 30 September 2018 / Accepted: 24 August 2019 / Published online: 16 September 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

An additive functional of a rooted tree is a functional that can be calculated recursively as the sum of the values of the functional over the branches, plus a certain toll function. Svante Janson recently proved a central limit theorem for additive functionals of conditioned Galton–Watson trees under the assumption that the toll function is local, i.e. only depends on a fixed neighbourhood of the root. We extend his result to functionals that are “almost local” in a certain sense, thus covering a wider range of functionals. The notion of almost local functional intuitively means that the toll function can be approximated well by considering only a neighbourhood of the root. Our main result is illustrated by several explicit examples including natural graph-theoretic parameters such as the number of independent sets, the number of matchings, and the number of dominating sets. We also cover a functional stemming from a tree reduction procedure that was studied by Hackl, Heuberger, Kropf, and Prodinger.

The first author was partially supported by the Division for Research Development (DRD) of Stellenbosch University. The second author was supported by the Czech Science Foundation, Grant Number GJ16-07822Y, with institutional support RVO:67985807. The third author was supported by the National Research Foundation of South Africa, Grant 96236. An extended abstract of this paper appeared in the Proceedings of the 29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms, AofA 2018, see [15].
Keywords  Galton–Watson trees · Additive functional · Almost local · Central limit theorem

1 Introduction

A functional $F$ that assigns a real value $F(T)$ to every finite rooted ordered tree is said to be additive if it satisfies a recursion of the form

$$F(T) = \sum_{i=1}^{k} F(T_i) + f(T),$$

where $T_1, T_2, \ldots, T_k$ are the branches (i.e. subtrees consisting of a child of the root and all its descendants) of $T$ and $f$ is a so-called toll function that assigns a real value to every rooted tree. If $T$ only consists of the root (so that $k = 0$), the sum equals 0 and hence $F(T) = f(T)$. Of course, every functional $F$ is additive in this sense (for a suitable choice of $f$), so the usefulness of the concept depends on what is known about the toll function $f$.

An important special case of an additive functional is the number of occurrences of a prescribed “fringe subtree”. A fringe subtree is an induced subtree of a rooted tree that consists of one of the nodes and all its descendants. Now fix a rooted tree $S$. We say that $S$ occurs on the fringe of $T$ if there is a fringe subtree of $T$ that is isomorphic to $S$ (when we consider ordered trees, where the order of branches matters, “isomorphic” is to be understood in the ordered sense as well). The number of occurrences of $S$ as a fringe subtree in $T$ (i.e. the number of nodes $v$ of $T$ for which the fringe subtree rooted at $v$ is isomorphic to $S$) is an additive functional, which we shall denote by $F_S(T)$. Indeed, one has

$$F_S(T) = \sum_{i=1}^{k} F_S(T_i) + f_S(T),$$

where

$$f_S(T) = \begin{cases} 1 & S \text{ is isomorphic to } T, \\ 0 & \text{otherwise}. \end{cases}$$

This is because an occurrence of $S$ in $T$ is either an occurrence in one of the branches, or comprises the entire tree $T$. Every additive functional can be expressed as a (possibly infinite) linear combination of these elementary functionals: it is easy to see (for example by induction) that a functional satisfying (1) can be expressed as

$$F(T) = \sum_S f(S) F_S(T).$$
Note that even though the sum is formally infinite, for any $T$ only a finite number of $F_S(T)$ are nonzero. Functionals of the form $F_S$ are known to be asymptotically normally distributed in various classes of trees, notably simply generated trees/Galton–Watson trees [9,20], which are the topic of this paper, and several other models of random trees [5,6,16]. In view of this and several other important examples of additive functionals that satisfy a central limit theorem, general schemes have been devised that yield a central limit theorem under different technical assumptions. This includes work on simply generated trees/Galton–Watson trees [9,20] (labelled trees, plane trees and $d$-ary trees are well-known special cases) as well as Pólya trees [20] and increasing trees [5,16,20] (specifically recursive trees, $d$-ary increasing trees and generalised plane-oriented recursive trees). It is worth mentioning, however, that there are many instances of additive functionals that are not normally distributed in the limit, since the toll functions can be quite arbitrary. A well-known example is the case of the path length, i.e. the sum of the distances of all nodes to the root in conditioned Galton–Watson trees. It satisfies (1) with toll function

$$f(T) = \lvert T \rvert - 1,$$

and, when suitably normalised, its limiting distribution for simply generated trees is the Airy distribution (see [17]).

The aforementioned results, while giving rather general conditions on the toll function that imply normality, are unfortunately still insufficient to cover all possible examples one might be interested in. This paper is essentially an extension of Janson’s work [9] on local functionals in conditional Galton–Watson trees. By weakening the conditions on the toll functions made in [9], we arrive at a new general central limit theorem that can be applied to a variety of examples that were not previously covered. Several such examples are presented in detail in this paper including natural graph-theoretic statistics and an open problem from a paper of Hackl, Heuberger, Kropf, and Prodinger [4] on tree reductions.

A local functional (as considered in Janson’s paper [9]) is a functional for which the value of the toll function can be determined from the knowledge of a fixed neighbourhood of the root. A typical example is the number of nodes with a given outdegree $r$, where the corresponding toll function is completely determined by the root degree: its value is 1 if the root degree is $r$, and 0 otherwise. We relax this condition somewhat (to what we call “almost local functionals”) in our main theorem. Intuitively speaking, functionals that satisfy our conditions have toll functions that can be approximated well from knowledge of a neighbourhood of the root, with the approximation getting better the wider the neighbourhood is chosen.

The model of random trees that we consider here are conditioned Galton–Watson trees: these are determined by an offspring distribution $\xi$ (we abuse the notation slightly and also denote by $\xi$ a random variable with distribution $\xi$). The Galton–Watson process starts from a single node, the root. At time $t$, all nodes at level/depth $t$ (i.e., at distance $t$ from the root) generate a number of children according to the offspring distribution $\xi$, the numbers of children of different nodes on the same level being mutually independent. The outcome of this process, which ends when all nodes at level $t$ generate 0 children, is a random tree $T$ (almost surely finite if we assume
$E\xi = 1$ and $\text{Var}\xi > 0$). By conditioning the process to “die out” when the total number of nodes is $n$ (of course, we only consider $n$ for which such an event occurs with nonzero probability) we obtain a conditioned Galton–Watson tree, which will be denoted by $T_n$.

We assume that the offspring distribution satisfies $E\xi = 1$ and $\text{Var}\xi \in (0, \infty)$. The condition $E\xi = 1$ is less restrictive than it might seem: it can be shown (see [8, Section 4]) that if $E\xi > 1$ then there is a (tilted) offspring distribution with mean 1 that gives the same distribution of $T_n$ for every $n$; when $E\xi < 1$, the same is true in most (but not all) cases, for example, whenever $E\xi^t < \infty$ for every $t$.

Conditioned Galton–Watson trees are known to be essentially equivalent to so-called simply generated trees [3, Section 3.1.4]. Classical examples include rooted labelled trees (corresponding to a Poisson distribution for $\xi$), plane trees (corresponding to a geometric distribution for $\xi$) and binary trees (with a distribution whose support is $\{0, 2\}$).

The main theorem of this paper, including the precise conditions we impose on the toll function $f$, will be stated in the following section along with some general remarks. Its proof requires several steps: we first provide some auxiliary results in Sect. 3, followed by a discussion of mean and variance of additive functionals whose toll functions satisfy our technical conditions (see Sect. 4). The final step, namely the proof of the central limit theorem, is performed in Sect. 5. At the end of the paper, we present some examples of natural tree functionals that satisfy the conditions of our main theorem but are not covered by previous results. Specifically, we will investigate three graph invariants of a somewhat similar nature (the number of independent sets, the number of matchings, and the number of dominating sets) for which log-normal laws are established. Finally, we prove central limit theorems for some functionals based on iterative tree reductions. All these examples will be presented in Sect. 6.

We conclude the introduction with some more notation: for a tree $T$, we let $T^{(M)}$ be its restriction to the first $M$ levels, i.e. all nodes whose distance to the root is at most $M$. A local functional as defined above is thus a functional for which the value of $f(T)$ is determined by $T^{(M)}$ for some fixed $M$ (the “cut-off”). The conditioned Galton–Watson tree $T_n$ is known to converge in the local topology induced by these restrictions to the (infinite) size-biased Galton–Watson tree $\hat{T}$ as defined by Kesten [11], see also [8, Section 5]: the tree is constructed from an infinite path starting with the root by attaching to each node of the path a random number of additional branches, each of which is an independent copy of $T$. The numbers of additional branches are independent and distributed as the random variable $\hat{\xi} - 1$, where $\hat{\xi}$ satisfies $P(\hat{\xi} = k) = kP(\xi = k)$ for $k = 0, 1, \ldots$. Given the degree of a node on the infinite path, the number of branches to the left (right) is uniformly distributed. Moreover, for any rooted tree $T$ one has (see [8, (5.11)], where $\mu = E\xi = 1$)

$$P(\hat{T}^{(M)} = T) = w_M(T)P(T^{(M)} = T),$$

where $w_M(T)$ is the number of nodes of depth $M$ in $T$.

For a rooted tree $T$ (possibly infinite), we let $\text{deg}(T)$ denote the degree of the root of $T$. Finally, it will be convenient for us to use the Vinogradov notation $\ll$ interchange-
ably with the $O$-notation. The notations $f(n) \ll g(n)$ and $f(n) = O(g(n))$ will both mean that $|f(n)| \leq K g(n)$ for a fixed positive constant $K$ and all $n$ for which $g(n)$ is positive. The same convention will apply when the expression in the $O$-term depends on other variables or trees. These expressions will always be positive, except possibly for a very small finite number of cases.

2 The General Theorem

Let us now formulate our main result, which is a central limit theorem for additive functionals under suitable technical conditions on the toll function $f$.

**Theorem 1** Let $T_n$ be a conditioned Galton–Watson tree of order $n$ with offspring distribution $\xi$, where $\xi$ satisfies $E\xi = 1$ and $0 < \sigma^2 := \text{Var}\xi < \infty$. Assume further that $E\xi^{2\alpha+1} < \infty$ for some integer $\alpha \geq 0$. Consider a functional $f$ of finite rooted ordered trees with the property that

$$f(T) = O(\text{deg}(T)^\alpha). \quad (3)$$

Furthermore, assume that there exists a sequence $(p_M)_{M \geq 1}$ of positive numbers with $p_M \to 0$ as $M \to \infty$ such that

- for every $M$, $N \in \{1, 2, \ldots\}$, such that $N \geq M$

$$E|f(\hat{T}^{(M)}) - E\left(f(\hat{T}^{(N)}) | \hat{T}^{(M)}\right)| \leq p_M, \quad (4)$$

and

- there is a sequence of positive integers $(M_n)_{n \geq 1}$ such that for large enough $n$

$$E\left|f(T_n) - f\left(T_n^{(M_n)}\right)\right| \leq p_{M_n}. \quad (5)$$

If $a_n := n^{-1/2}(n^{\max\{\alpha, 1\}} p_{M_n} + M_n^2)$ satisfies

$$\lim_{n \to \infty} a_n = 0, \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{a_n}{n} < \infty, \quad (6)$$

then

$$\frac{F(T_n) - n\mu}{\sqrt{n}} \overset{d}{\to} \mathcal{N}(0, \gamma^2) \quad (7)$$

where $\mu = E f(T)$, for some $0 \leq \gamma < \infty$.

In (7), the numerator $F(T_n) - n\mu$ can be replaced by $F(T_n) - E F(T_n)$ as we will see in Proposition 7 that

$$E F(T_n) = \mu n + o(\sqrt{n}), \quad \text{as} \ n \to \infty. \quad (8)$$
The existence and finiteness of the expectation \( \mu = \mathbb{E} f(T) \) is guaranteed by (3) and \( \mathbb{E} \xi^{2\alpha+1} < \infty \). We will also show that the variance satisfies the following asymptotic formula:

\[
\text{Var } F(T_n) = \gamma^2 n + o(n), \quad \text{as } n \to \infty.
\]

(9)

This is proved together with Theorem 1 in Sect. 5. Although we prove that the constant \( \gamma \) exists, we are unable to provide a general argument that can tell whether or not \( \gamma \neq 0 \). However, in the applications in Sect. 6, we show that \( \gamma \neq 0 \) for the logarithm of the number of independent sets. The argument there can also be applied to the logarithm of the number of matchings and the logarithm of the number of dominating sets.

**Remark 2** The proof of Theorem 1 is a generalisation of Janson’s proof of his theorem for bounded and local functionals, see [9, Theorem 1.13]. The boundedness condition is now replaced by (3) assuming finiteness of higher moments of the offspring distribution \( \xi \). However, the main difficulty to overcome is the fact that our toll function is no longer local. To give a simple example, an essential part of the proof is to give a meaning to the “expectation” \( \mathbb{E} f(\hat{T}) \). The functional \( f \) does not need to be defined on infinite trees. When \( f \) is local with a cut-off \( M \), then \( f(\hat{T}) = f(\hat{T}(M)) \) by definition. So, \( \mathbb{E} f(\hat{T}) \) is simply defined to be \( \mathbb{E} f(\hat{T}(M)) \). In our case, where \( f \) is not necessarily local, we define \( \mathbb{E} f(\hat{T}) \) in the most natural way as

\[
\mathbb{E} f(\hat{T}) := \lim_{M \to \infty} \mathbb{E} f(\hat{T}(M)),
\]

(10)

which may not exist in general. However, if \( f \) satisfies (4), then we can show that \( \mathbb{E} f(\hat{T}) \) exists and is finite. Indeed,

\[
|\mathbb{E} f(\hat{T}(M)) - \mathbb{E} f(\hat{T}(N))| = \left| \mathbb{E} \left( f(\hat{T}(M)) - \mathbb{E} \left( f(\hat{T}(N)) \big| \hat{T}(M) \right) \right) \right| \leq \mathbb{E} \left| f(\hat{T}(M)) - \mathbb{E} \left( f(\hat{T}(N)) \big| \hat{T}(M) \right) \right| \leq p_M,
\]

(11)

which tends to zero as \( M \to \infty \), uniformly for \( N \geq M \). In other words, \( (\mathbb{E} f(\hat{T}(M)))_{M \geq 1} \) is a Cauchy sequence, so the limit (10) exists.

### 3 Auxiliary Results

In this section, we give some useful results that we will need in the proof of our main theorem. Throughout the rest of the paper, the offspring distribution \( \xi \) is assumed to satisfy the following conditions (as in Theorem 1):

- \( \mathbb{E} \xi = 1 \),
- \( 0 < \sigma^2 := \text{Var } \xi < \infty \), and
- \( \mathbb{E} \xi^{2\alpha+1} < \infty \) for some integer \( \alpha \geq 0 \).

The distribution of the number of nodes at level \( k \), \( w_k \), for the three random trees \( T \), \( \hat{T} \), and \( T_n \) will play an important role in our proof. This parameter has been studied in [7], and in particular, the following results were proved there, see [7, Theorem 1.13, Lemma 2.2, and Lemma 2.3] (note that \( T_\infty \) is used there for \( \hat{T} \)).
Lemma 3 Assume that $\xi$ satisfies $E\xi = 1$, $0 < \sigma^2 := \text{Var } \xi < \infty$ and $E\xi^{2\alpha + 1} < \infty$. Then for every positive integer $r \leq \max\{2\alpha, 1\}$, we have

$$E(w_k(T)^r) = O(k^{\alpha - 1}), \ E(w_k(\hat{T})^r) = O(k^r), \ \text{and } E(w_k(T_n)^r) = O(k^r), \quad (12)$$

where the implicit constant factors in the $O$-bounds depend on the offspring distribution $\xi$ only.

For a rooted tree $T$, we know that $|T^{(M)}| = \sum_{k=0}^{M} w_k(T)$. Hence, we can immediately deduce from this lemma, with $r = 1$, that

$$E|T^{(M)}| = O(M), \ E|\hat{T}^{(M)}| = O(M^2), \ \text{and } E|T_n^{(M)}| = O(M^2). \quad (13)$$

In fact, it can be shown that $Ew_k(T) = (E\xi)^k$ and hence $E|T^{(M)}| = M + 1$. We are also going to make extensive use of the higher moments of the root degree. By definition, the distribution of $\text{deg}(T)$ is $\xi$, so we know the higher moments of $\text{deg}(T)$. On the other hand, note that $\text{deg}(T) = w_1(T)$. So, as particular cases of the estimates in (12), we have

$$E(\text{deg}(\hat{T})^r) < \infty \ \text{and } E(\text{deg}(T_n)^r) = O(1), \quad (14)$$

for every positive integer $r \leq \max\{2\alpha, 1\}$, where the implied constant in the second estimate depends only on $\xi$.

It is well known that $T_n$ converges locally to the infinite random tree $\hat{T}$ in the sense that

$$P\left(T_n^{(M)} = T\right) \to P\left(\hat{T}^{(M)} = T\right), \ \text{as } n \to \infty,$$

for any fixed integer $M \geq 0$ and a fixed tree $T$. Janson obtained a bound on the rate of convergence of this estimate, in the proof of [9, Lemma 5.9] (see (5.42) there). His result can be formulated as follows: For any tree $T$ with $|T| \leq n/2$, and any $M \geq 0$ we have

$$P\left(T_n^{(M)} = T\right) = P\left(\hat{T}^{(M)} = T\right) \left(1 + O\left(\frac{|T|}{n^{1/2}}\right)\right), \quad (15)$$

where the constant in the $O$ notation is independent of $T$ and $M$. Here $T$ and $M$ may depend on $n$, and this is crucial for our purposes. As a consequence of this result we can bound the difference between the two expectations $E f(T_n^{(M)})$ and $E f(\hat{T}^{(M)})$ explicitly in terms of $M$, where $M$ is also allowed to depend on $n$.

Lemma 4 If $\xi$ satisfies $E\xi = 1$, $0 < \sigma^2 := \text{Var } \xi < \infty$ and $E\xi^{2\alpha + 1} < \infty$, and $f$ satisfies $f(T) = O(\text{deg}(T)^\alpha)$, as in (3), then we have

$$|E f(T_n^{(M)}) - E f(\hat{T}^{(M)})| \leq O\left(n^{-1/2} M^2 E(\text{deg}(\hat{T})^{\alpha + 1}) + n^{-1} M^2 E(\text{deg}(T_n)^{\alpha + 1})\right), \quad (16)$$

where the constant in the $O$ notation is independent of $n$ and $M$. 

\(\square\) Springer
Proof We have

\[
|\mathbb{E} f(T_n^{(M)}) - \mathbb{E} f(\hat{T}^{(M)})| \\
= \left| \sum_T f(T) \mathbb{P}\left( T_n^{(M)} = T \right) - \sum_T f(T) \mathbb{P}\left( \hat{T}^{(M)} = T \right) \right| \\
\leq \sum_{|T| \leq n/2} |f(T)| \left[ \mathbb{P}\left( T_n^{(M)} = T \right) - \mathbb{P}\left( \hat{T}^{(M)} = T \right) \right] \\
+ \sum_{|T| > n/2} |f(T)| \left[ \mathbb{P}\left( T_n^{(M)} = T \right) + \mathbb{P}\left( \hat{T}^{(M)} = T \right) \right].
\]

We begin by estimating the sum over \(|T| \leq n/2\). Using (15) and the bound (3) on \(f(T)\), we obtain

\[
\sum_{|T| \leq n/2} |f(T)| \left[ \mathbb{P}\left( T_n^{(M)} = T \right) - \mathbb{P}\left( \hat{T}^{(M)} = T \right) \right] \\
\ll \sum_T \mathbb{P}\left( \hat{T}^{(M)} = T \right) \frac{\deg(T)^\alpha |T|}{n^{1/2}}.
\]

Now, the right-hand side can be bounded as follows:

\[
\sum_T \mathbb{P}\left( \hat{T}^{(M)} = T \right) \frac{\deg(T)^\alpha |T|}{n^{1/2}} \\
= n^{-1/2} \mathbb{E}(\deg(\hat{T}^{(M)})^\alpha |\hat{T}^{(M)}|) \\
= n^{-1/2} \mathbb{E}\left( \deg(\hat{T}^{(M)})^\alpha \mathbb{E}\left( |\hat{T}^{(M)}| \mid \deg(\hat{T}^{(M)}) \right) \right).
\]

Conditioning on \(\deg(\hat{T}^{(M)})\) (which is the same as \(\deg(\hat{T})\) for \(M \geq 1\), \(\hat{T}\) consists of a root, a copy of \(\hat{T}\) and \(\deg(\hat{T}) - 1\) independent copies of \(T\). Thus, by the estimates in (13), we have

\[
\mathbb{E}\left( |\hat{T}^{(M)}| \mid \deg(\hat{T}^{(M)}) \right) \ll M^2 + M \deg(\hat{T}^{(M)}) \ll M^2 \deg(\hat{T}^{(M)}).
\]  

(17)

Therefore,

\[
\mathbb{E}\left( \deg(\hat{T}^{(M)})^\alpha \mathbb{E}\left( |\hat{T}^{(M)}| \mid \deg(\hat{T}^{(M)}) \right) \right) \ll M^2 \mathbb{E}(\deg(\hat{T}^{(M)})^{\alpha+1}),
\]

which yields

\[
\sum_T \mathbb{P}\left( \hat{T}^{(M)} = T \right) \frac{\deg(T)^\alpha |T|}{n^{1/2}} \ll n^{-1/2} M^2 \mathbb{E}(\deg(\hat{T}^{(M)})^{\alpha+1}).
\]  

(18)
Next we estimate the sum over $|T| > n/2$, which we split further as follows:

$$\sum_{|T| > n/2} |f(T)| \left( \mathbb{P}(T^{(M)} = T) + \mathbb{P}(\hat{T}^{(M)} = T) \right)$$

$$\ll \sum_{|T| > n/2} \mathbb{P}(\hat{T}^{(M)} = T) \deg(T)^\alpha + \sum_{|T| > n/2} \mathbb{P}(\hat{T}^{(M)} = T) \deg(T)^\alpha.$$ 

We have

$$\sum_{|T| > n/2} \mathbb{P}(\hat{T}^{(M)} = T) \deg(T)^\alpha$$

$$= \sum_{k \geq 1} k^\alpha \mathbb{P}\left(|\hat{T}^{(M)}| > n/2 \text{ and } \deg(\hat{T}) = k\right)$$

$$= \sum_{k \geq 1} k^\alpha \mathbb{P}\left(\deg(\hat{T}) = k\right) \mathbb{P}\left(|\hat{T}^{(M)}| > n/2 \bigg| \deg(\hat{T}) = k\right).$$

Markov’s inequality yields

$$\mathbb{P}\left(|\hat{T}^{(M)}| > n/2 \bigg| \deg(\hat{T}) = k\right) \leq \frac{2\mathbb{E}(|\hat{T}^{(M)}| \big| \deg(\hat{T}) = k)}{n} \ll \frac{kM^2}{n},$$

where the last estimate follows from (17). Thus,

$$\sum_{|T| > n/2} \mathbb{P}(\hat{T}^{(M)} = T) \deg(T)^\alpha \ll n^{-1} M^2 \mathbb{E}(\deg(\hat{T})^{\alpha+1}). \quad (19)$$

Finally, for the last term, we proceed in a similar fashion:

$$\sum_{|T| > n/2} \mathbb{P}(T^{(M)} = T) \deg(T)^\alpha$$

$$= \sum_{k \geq 1} k^\alpha \mathbb{P}\left(|T^{(M)}| > n/2 \text{ and } \deg(T^{(M)}) = k\right)$$

$$= \sum_{k \geq 1} k^\alpha \mathbb{P}\left(\deg(T^{(M)}) = k\right) \mathbb{P}\left(|T^{(M)}| > n/2 \bigg| \deg(T^{(M)}) = k\right). \quad (20)$$

If $T_{n,1}, T_{n,2}, \ldots, T_{n,k}$ are the branches of $T_n$, given that $\deg(T_n) = k$, then, conditioning on their sizes $|T_{n,i}| = n_i$ for every $i$, they are $k$ independent conditioned Galton–Watson trees $T_{n,1}, T_{n,2}, \ldots, T_{n,k}$. On the other hand, if $\deg(T_n) = k$, then we have

$$|T_n^{(M)}| = 1 + \sum_{i=1}^{k} |T_{n,i}^{(M-1)}|.$$
Thus,
\[
E \left( |T_n^{(M)}| \mid \text{deg}(T_n) = k, |T_{n,1}| = n_1, \ldots, |T_{n,k}| = n_k \right)
= 1 + \sum_{i=1}^{k} E \left( |T_{n_i}^{(M-1)}| \right) \ll kM^2
\]
which again follows from the last estimate in (13). Multiplying each conditional probability by \( P \left( |T_{n,1}| = n_1, \ldots, |T_{n,k}| = n_k \mid \text{deg}(T_n) = k \right) \) and summing over all choices of \( n_1, \ldots, n_k \), we obtain \( E \left( |T_n^{(M)}| \mid \text{deg}(T_n) = k \right) \ll kM^2 \), which, via Markov’s inequality, yields
\[
P \left( |T_n^{(M)}| > n/2 \mid \text{deg}(T_n) = k \right) \ll n^{-1}kM^2.
\]
Therefore, making use of (20), the latter estimate, and (14) once again, we have
\[
\sum_{|T| > n/2} P \left( T_n^{(M)} = T \right) \text{deg}(T)^\alpha
\ll n^{-1}M^2 \sum_{k \geq 1} k^{\alpha+1} P \left( \text{deg}(T_n) = k \right) = n^{-1}M^2 E(\text{deg}(T_n)^{\alpha+1}). \tag{21}
\]
Combining the estimates (18), (19), and (21), we finally arrive at the estimate
\[
|E f(T_n^{(M)}) - E f(\hat{T}^{(M)})| \ll n^{-1/2}M^2 E(\text{deg}(\hat{T})^{\alpha+1}) + n^{-1}M^2 E(\text{deg}(T_n)^{\alpha+1}), \tag{22}
\]
which completes the proof of the lemma. \( \square \)

The following lemma will be useful in the estimate of the variance in the next section. First, we start with some operators on functionals. For any toll function \( f \) of an additive functional \( F \), we denote by \( f^{(0)} \) the centred toll function which is defined by
\[
f^{(0)}(T) := f(T) - E f(T|T|),
\]
and let \( F^{(0)} \) be the additive functional associated with \( f^{(0)} \). Furthermore, for any subset \( I \) of \( \mathbb{N} \), let \( f_I \) be the functional defined by
\[
f_I(T) = \begin{cases} f(T) & \text{if } |T| \in I, \\ 0 & \text{otherwise}, \end{cases}
\]
and we denote by \( F_I \) the additive functional whose toll function is \( f_I \). To clarify, by \( f^{(0)}_I \), we mean \( (f_I)^{(0)} \) (and likewise for \( f^{(1)}_I \), which will be defined later).
Lemma 5 Assume that \( \xi \) and \( f \) satisfy the conditions of Theorem 1 and let \((p_M)_{M \geq 1}\) and \((M_n)_{n \geq 1}\) be the corresponding sequences. Then, uniformly for any subset \( I \) of \( \mathbb{N} \), we have

\[
\mathbb{E}(f_I^{(0)}(T_n) F_I^{(0)}(T_n)) \ll \mathbb{E}(\deg(T_n)^{2\alpha}) + M_n^2 \mathbb{E}(\deg(T_n)^{\alpha+1}) + n^{\max\{\alpha, 1\}} p_M n, \tag{23}
\]

where the constants implicit in \( \ll \) do not depend on \( I \).

Proof Since the left side of (23) is zero for \( n \notin I \), we may assume without loss of generality that \( n \in I \). We decompose \( F_I^{(0)}(T_n) \) according to the depth \( d(v) \) of the nodes:

\[
F_I^{(0)}(T_n) = \sum_{v \in T_n} f_I^{(0)}(T_n, v) = \sum_{d(v) < M} f_I^{(0)}(T_n, v) + \sum_{d(v) \geq M} f_I^{(0)}(T_n, v) =: S_1 + S_2, \tag{24}
\]

where \( T_n, v \) denotes the fringe subtree of \( T_n \) rooted at \( v \). Notice that \( f_I^{(0)} \) might not necessarily satisfy all conditions of Theorem 1. However, (3) is satisfied by \( f_I^{(0)} \). Hence, we have

\[
\mathbb{E}|f_I^{(0)}(T_n) S_1| \ll \mathbb{E}\left( \deg(T_n)^\alpha \sum_{d(v) < M} \deg(T_n, v)^\alpha \right) = \mathbb{E} \left( \deg(T_n)^\alpha \mathbb{E} \left( \sum_{d(v) < M} \deg(T_n, v)^\alpha \bigg| \deg(T_n) \right) \right).
\]

Next, for any positive integer \( m \leq M \), we have

\[
\mathbb{E} \left( \sum_{d(v) < m} \deg(T_n, v)^\alpha \bigg| T_n^{(m-1)} \right) = \sum_{d(v) < m-1} \deg(T_n, v)^\alpha + O(w_{m-1}(T_n)). \tag{25}
\]

This is because the \( w_{m-1}(T_n) \) fringe subtrees with roots at level \( m - 1 \), conditioned on their sizes, are conditioned Galton–Watson trees and thus by (14) the moments of the root degrees are \( O(1) \). Similarly, by conditioning on the sizes of the \( \deg(T_n) \) subtrees and applying (12) we obtain \( \mathbb{E} \left( w_{m-1}(T_n) \bigg| \deg(T_n) \right) = O(m \deg(T_n)) \). Hence taking the conditional expectation of (25), we have

\[
\mathbb{E} \left( \sum_{d(v) < m} \deg(T_n, v)^\alpha \bigg| \deg(T_n) \right) = \mathbb{E} \left( \sum_{d(v) < m-1} \deg(T_n, v)^\alpha \bigg| \deg(T_n) \right) + O(m \deg(T_n)).
\]

Thus, iterating from \( M \), we obtain

\[
\mathbb{E} \left( \sum_{d(v) < M} \deg(T_n, v)^\alpha \bigg| \deg(T_n) \right) \ll \deg(T_n)^\alpha + M^2 \deg(T_n).
\]
Therefore,
\[ |\mathbb{E} f_I^{(0)} (T_n) S_1| \ll \mathbb{E}(\deg(T_n)^{2\alpha}) + M^2 \mathbb{E}(\deg(T_n)\alpha + 1). \] (26)

We now turn to a bound for \( |\mathbb{E} f_I^{(0)} (T_n) S_2| \). Note first that
\[ S_2 = \sum_{d(v) = M} F_I^{(0)} (T_{n,v}). \] (27)

We condition on \( T_n(M) \) and the sizes of the fringe subtrees \( T_{n,v_i}, i = 1, \ldots, w_M(T_n) \), induced by nodes at level \( M \). Conditionally, each \( T_{n,v_i} \) is distributed as \( T_{n_i} \), where \( n_i = |T_{n,v_i}|. \) From the definition of \( f_I^{(0)} \), we know that \( \mathbb{E} f_I^{(0)} (T_m) = 0 \) for every \( m \geq 1 \). Note that
\[ F_I^{(0)} (T_m) = \sum_{k=1}^m F_{[k]}^{(0)} (T_m) \mathbb{I}_{[k \in I]}, \]
where \( \mathbb{I}_{[k \in I]} \) denotes the indicator function of the set \( I \). The sum is only over \( k \leq m \) since, trivially, \( F_{[k]}^{(0)} (T_m) = 0 \) for \( k > m \). It follows (see [9, (6.25)]) that \( \mathbb{E} F_I^{(0)} (T_m) = 0 \) for every \( m \geq 1 \) and therefore, by (27) and the law of total expectation, we have
\[ \mathbb{E} \left( S_2 | T_n(M) \right) = 0. \]

Writing \( X := \mathbb{E} (f_I (T_n) | T_n(M)) \) and \( X^{(0)} := \mathbb{E} (f_I^{(0)} (T_n) | T_n(M)) \), we have
\[ \mathbb{E}(X^{(0)} S_2) = \mathbb{E} \left( \mathbb{E} \left( X^{(0)} S_2 | T_n(M) \right) \right) = \mathbb{E} \left( X^{(0)} \mathbb{E} \left( S_2 | T_n(M) \right) \right) = 0. \]

Hence,
\[ |\mathbb{E} f_I^{(0)} (T_n) S_2| = |\mathbb{E} (S_2 (f_I^{(0)} (T_n) - X^{(0)}))| \leq ||S_2||_{\infty} \mathbb{E} |f_I^{(0)} (T_n) - X^{(0)}|, \] (28)
where \( ||S_2||_{\infty} \) stands for the essential supremum of the random variable \( |S_2| \). Recalling that \( f_I \) satisfies condition (3), we have that
\[ |S_2| \leq \sum_{d(v) \geq M} |f_I^{(0)} (T_{n,v})| \ll \sum_{v \in T_n} \deg(T_{n,v})^\alpha. \]

Since \( \alpha \) is a nonnegative integer, the last term is bounded above by \( (\sum_{v \in T_n} \deg(T_{n,v}))^\alpha \) (which is equal to \( (n - 1)^\alpha \)) when \( \alpha \geq 1 \) and by \( n \), when \( \alpha = 0 \). Hence, we get
\[ ||S_2||_{\infty} \ll n^{\max\{\alpha,1\}}. \] (29)
By definition, we know that \( f_I(T_n) = f_I^{(0)}(T_n) + \mathbb{E} f_I(T_n) \) and we can also verify that \( X = X^{(0)} + \mathbb{E} f_I(T_n) \). Therefore, since we assumed that \( n \in I \), we have
\[
\mathbb{E} | f_I^{(0)}(T_n) - X^{(0)} | = \mathbb{E} | f_I(T_n) - X | = \mathbb{E} | f(T_n) - \mathbb{E}(f(T_n) | T_n^{(M)}) |.
\]

(30)

We can use (5) to estimate the right-hand side of the above equation. For the rest of the proof, we choose \( M = M_n \) (as defined in Theorem 1). We have
\[
| f(T_n) - \mathbb{E}(f(T_n) | T_n^{(M)}) | \\
\leq | f(T_n) - f(T_n^{(M)}) | + | f(T_n^{(M)}) - \mathbb{E}(f(T_n) | T_n^{(M)}) | \\
= | f(T_n) - f(T_n^{(M)}) | + \mathbb{E} \left( f(T_n^{(M)}) - f(T_n) | T_n^{(M)} \right) | \\
\leq | f(T_n) - f(T_n^{(M)}) | + \mathbb{E} \left( | f(T_n^{(M)}) - f(T_n) | | T_n^{(M)} \right) .
\]

Taking the expectations in the latter inequality, and using the condition (5) (with \( M = M_n \)), we obtain
\[
\mathbb{E} | f(T_n) - \mathbb{E}(f(T_n) | T_n^{(M)}) | \leq 2pM .
\]

This, together with (28), (29), and (30) implies
\[
\mathbb{E} | f_I^{(0)}(T_n) S_2 | \ll n^{\max[\alpha, 1]} pM .
\]
Combining this with (24) and (26), we obtain (23). \( \square \)

4 Mean and Variance

We first look at the expectation \( \mathbb{E} f(T_n) \). As is also the case in [9], one of the key observations in the proof of Theorem 1 is the fact that \( \mathbb{E} f(T_n) \) is asymptotically equal to \( \mathbb{E} f(\hat{T}) \) (which is finite, cf. Remark 2) with an explicit bound on the error term. This is made precise in the next lemma.

Lemma 6 Assume that \( \xi \) and \( f \) satisfy the conditions of Theorem 1 and let \( (p_M)_{M \geq 1} \) and \( (M_n)_{n \geq 1} \) be the corresponding sequences, then
\[
\mathbb{E} f(T_n) = \mathbb{E} f(\hat{T}) + O(p_{M_n} + n^{-1/2} M_n^2) .
\]

(31)

Proof We let \( M_n \) be defined as in Theorem 1, but write \( M = M_n \) for easier reading. Notice first that
\[
| \mathbb{E} f(T_n) - \mathbb{E} f(\hat{T}) | \\
\leq | \mathbb{E} f(T_n) - \mathbb{E} f(T_n^{(M)}) | + | \mathbb{E} f(\hat{T}^{(M)}) - \mathbb{E} f(\hat{T}) | \\
+ | \mathbb{E} f(T_n^{(M)}) - \mathbb{E} f(\hat{T}^{(M)}) | .
\]

(32)
The first term on the right side is at most $p_M$ by assumption (5). The second term is also bounded above by $p_M$, since (11) implies

$$|\mathbb{E} f(\hat{T}) - \mathbb{E} f(\hat{T}(M))| = \lim_{N \to \infty} |\mathbb{E} f(\hat{T}(N)) - \mathbb{E} f(\hat{T}(M))| \leq p_M.$$  

By Lemma 4 we have

$$|\mathbb{E} f(T_n(M)) - \mathbb{E} f(\hat{T}(M))| \ll n^{-1/2}M^2 \mathbb{E}(\text{deg}(\hat{T})^{\alpha+1}) + n^{-1}M^2 \mathbb{E}(\text{deg}(T_n)^{\alpha+1}).$$

In view of (14), the moment $\mathbb{E}(\text{deg}(\hat{T})^{\alpha+1})$ is finite and $\mathbb{E}(\text{deg}(T_n)^{\alpha+1})$ is $O(1)$ as $n \to \infty$. Therefore, we conclude that

$$|\mathbb{E} f(T_n) - \mathbb{E} f(\hat{T})| \ll p_M + n^{-1/2}M^2 = p_Mn + n^{-1/2}Mn^2,$$

which is equivalent to the statement in the lemma. \qed

Lemma 6 is already enough to prove the estimate for the mean $\mathbb{E} F(T_n)$ as it is stated in (8). The following proposition is a consequence of [9, Theorem 1.5].

**Proposition 7** If $\xi$ and $f$ satisfy the conditions of Theorem 1, then

$$\mathbb{E} F(T_n) = n\mu + o(\sqrt{n}),$$

where $\mu = \mathbb{E} f(T)$.

**Proof** Consider the shifted functional $F'(T)$ whose toll function is defined by $f'(T) = f(T) - \mathbb{E} f(\hat{T})$. Since $\mathbb{E} f(\hat{T})$ does not depend on $T$, we have

$$\mathbb{E} F(T_n) = \mathbb{E} F'(T_n) + n\mathbb{E} f(\hat{T}).$$  

(33)

Lemma 6 and the assumptions of Theorem 1 imply that $\mathbb{E} f'(T_n) = \mathbb{E} f(T_n) - \mathbb{E} f(\hat{T}) \ll p_Mn + n^{-1/2}Mn^2 \to 0$, as $n \to \infty$. Therefore applying [9, Theorem 1.5(i)] to $F'$ we get that

$$\mathbb{E} F'(T_n) = n\mathbb{E} f'(T) + o(\sqrt{n}) = n(\mathbb{E} f(T) - \mathbb{E} f(\hat{T})) + o(\sqrt{n}).$$  

(34)

Substituting (34) into (33) completes the proof. \qed

Using the same notation as in Lemma 5 we obtain the following estimate of the variance.

**Lemma 8** Assume that $\xi$ and $f$ satisfy the conditions of Theorem 1 and let $(p_M)_{M \geq 1}$ and $(M_n)_{n \geq 1}$ be the corresponding sequences. Moreover, set $a_k = k^{-1/2}(k^{\max\{\alpha,1\}}p_{M_k} + M_k^2)$ (as in Theorem 1) and $\mu_k = \mathbb{E} f(T_k)$. Then, for any subset $I$ of $\mathbb{N}$, we have

$$n^{-1/2} \text{Var} (F_I(T_n))^{1/2} \ll \left( \sup_{k \in I} a_k + \sum_{k \in I} \frac{a_k}{k} \right)^{1/2} + \sup_{k \in I} |\mu_k| + \sum_{k \in I} \frac{|\mu_k|}{k}.$$

(35)
where the constants implicit in \( \ll \) do not depend on \( I \).

**Proof** We follow the proof of \([9, \text{Theorem 6.12}]\). We start with a decomposition \( f_1(T) = f_1^{(0)}(T) + f_1^{(1)}(T) \), where \( f_1^{(0)}(T) = f_1(T) - \mu_{|T|} \) and \( f_1^{(1)}(T) = \mu_{|T|} \) if \( |T| \in I \) and both are zero otherwise. This induces a decomposition \( F_1 = F_1^{(0)} + F_1^{(1)} \) of the functional \( F_1 \), where \( F_1^{(0)} \) and \( F_1^{(1)} \) are the additive functionals defined by the toll functions \( f_1^{(0)} \) and \( f_1^{(1)} \) respectively. In view of Minkowski’s inequality \((\text{Var}(X + Y))^{1/2} \leq (\text{Var}X)^{1/2} + (\text{Var}Y)^{1/2} \), we can estimate the variances \( \text{Var} F_1^{(0)}(T_n) \) and \( \text{Var} F_1^{(1)}(T_n) \) separately.

By applying \([9, \text{Theorem 6.7}]\) to \( F_1^{(1)} \) and noting that \( \sqrt{\mathbb{E}(f_1^{(1)}(T))} = |\mu_k|_{k \in I} \), we obtain

\[
    n^{-1/2} \text{Var} \left( F_1^{(1)}(T) \right)^{1/2} \ll \sup_{k \in I} |\mu_k| + \sum_{k \in I} \frac{|\mu_k|}{k}, \tag{36}
\]

where the implicit constants are independent of \( I \). This is because the constants in the bound provided by the aforementioned \([9, \text{Theorem 6.7}]\) are independent of the functional.

Next, we consider \( \text{Var}(F_1^{(0)}(T_n)) \). Since \( \mathbb{E}f_1^{(0)}(T_k) = 0 \) for every \( k \), inequality (6.28) from \([9]\) holds (where \( f_k \) is what in our notation is \( f_{[k]} \)). We have

\[
    \frac{1}{n} \text{Var} \left( F_1^{(0)}(T_n) \right) \leq 2 \sum_{k=1}^{n} \frac{\mathbb{P}(S_{n-k} = n-k)}{\mathbb{P}(S_n = n-1)} \pi_k \mathbb{E}(f_1^{(0)}(T_k)F_1^{(0)}(T_k)), \tag{37}
\]

where \( \pi_k = \mathbb{P}(|T| = k) \), and \( S_k \) is the sum of \( k \) independent copies of \( \xi \). From \([9, \text{Lemma 5.2}]\), we know that

\[
    \frac{\mathbb{P}(S_{n-k} = n-k)}{\mathbb{P}(S_n = n-1)} \ll \frac{n^{1/2}}{(n-k+1)^{1/2}},
\]

uniformly for \( 1 \leq k \leq n \). Recalling that \( \pi_k = O(k^{-3/2}) \), which is a well-known fact but can also be found in \([9, (4.13)]\), we obtain

\[
    \frac{1}{n} \text{Var} F_1^{(0)}(T_n) \ll \sum_{k=1}^{n} \frac{n^{1/2}}{(n-k+1)^{1/2}k^{3/2}} |\mathbb{E}(f_1^{(0)}(T_k)F_1^{(0)}(T_k))|, \tag{38}
\]

For \( k \notin I \), we have \( \mathbb{E}(f_1^{(0)}(T_k)F_1^{(0)}(T_k)) = 0 \), and for \( k \in I \) we use Lemma 5 to estimate \( \mathbb{E}(f_1^{(0)}(T_k)F_1^{(0)}(T_k)) \). Once again, by means of the second estimate in (14), both \( \mathbb{E}(|\text{deg}(T_k)^{2\alpha}) \) and \( \mathbb{E}(|\text{deg}(T_k)^{\alpha+1}) \) are bounded above by constants. Thus, for \( k \in I \), we deduce that

\[
    \mathbb{E}(f_1^{(0)}(T_k)F_1^{(0)}(T_k)) \ll k^{\max[\alpha, 1]} \rho_{M_k} + M_k^2 = k^{1/2}a_k, \tag{39}
\]

\( \blacksquare \) Springer
where the implicit constants are independent of \( I \). Applying (39) to (38), we get
\[
\frac{1}{n} \text{Var} F_I^{(0)} (T_n) \ll \sum_{k=1}^{n} \frac{n^{1/2} a_k \|_{\{k \in I\}}}{(n - k + 1)^{1/2} k} \quad \ll \sum_{k=1}^{n/2} \frac{a_k \|_{\{k \in I\}}}{k} + \left( \sup_{k \in I} a_k \right) \sum_{n/2 \leq k \leq n} \frac{1}{(n - k + 1)^{1/2} n^{1/2}}.
\]

Noting that the last sum on the right side is bounded by a constant, we obtain
\[
\frac{1}{n} \text{Var} F_I^{(0)} (T_n) \ll \sup_{k \in I} a_k + \sum_{k \in I} \frac{a_k}{k}.
\]

The proof is complete by applying Minkowski’s inequality to combine (36) and (40).

\( \square \)

5 Central Limit Theorem

We use a truncation argument as in the proof of [9, Theorem 1.5]. This is formulated in the following lemma:

**Lemma 9** Let \((X_n)_{n \geq 1}\) and \((W_{N,n})_{N,n \geq 1}\) be sequences of centred random variables. If for some random variables \(W_N, N = 1, 2, \ldots, \) and \(W\) we have

- \(W_{N,n} \overset{d}{\to} W_N \) for every \( N \geq 1 \), and \( W_N \overset{d}{\to} W \),
- \( \text{Var}(X_n - W_{N,n}) = O(\sigma_N^2) \) uniformly in \( n \), and \( \sigma_N^2 \to 0 \),

then \( X_n \overset{d}{\to} W \).

If we assume further that \( X_n, W_{N,n}, W_N, \) and \( W \) have finite second moment for every \( n, N \geq 1 \), and

- \( \text{Var} W_{N,n} \overset{d}{\to} \text{Var} W_N \) and \( \text{Var} W_N \overset{d}{\to} \text{Var} W \),

then we also have \( \text{Var} X_n \overset{d}{\to} \text{Var} W \).

**Proof** The convergence in distribution is a consequence of [10, Theorem 4.28] or [1, Theorem 4.2]. So it suffices to prove convergence of variance. By the triangle inequality and Minkowski’s inequality we have
\[
\left| \sqrt{\text{Var}(X_n)} - \sqrt{\text{Var}(W)} \right| \leq \left| \sqrt{\text{Var}(X_n)} - \sqrt{\text{Var}(W_{N,n})} \right| + \left| \sqrt{\text{Var}(W_{N,n})} - \sqrt{\text{Var}(W_N)} \right| + \left| \sqrt{\text{Var}(W_N)} - \sqrt{\text{Var}(W)} \right| \leq \sqrt{\text{Var}(X_n - W_{N,n})} + \left| \sqrt{\text{Var}(W_{N,n})} - \sqrt{\text{Var}(W_N)} \right| + \left| \sqrt{\text{Var}(W_N)} - \sqrt{\text{Var}(W)} \right|.
\]
Since $\text{Var}(W_{N,n}) \rightarrow_{n} \text{Var}(W_N)$, the middle term in the last line tends to zero as $n \rightarrow \infty$ for any fixed $N$. Hence, we deduce that

$$\limsup_n \left| \sqrt{\text{Var}(X_n)} - \sqrt{\text{Var}(W)} \right| \ll \sigma + \left| \sqrt{\text{Var}(W_N)} - \sqrt{\text{Var}(W)} \right|.$$ 

Since the right-hand side can be made arbitrarily small by increasing $N$, we deduce that

$$\text{Var}(X_n) \rightarrow_{n} \text{Var}(W),$$

which completes the proof. $\square$

**Proof of Theorem 1 and Equation (9)** We may assume, without loss of generality, that $\mathbb{E} f(\hat{T}) = 0$, by subtracting $\mathbb{E} f(\hat{T})$ from $f$ if it is not zero, because shifting $f$ by a constant will only add a deterministic term in $F(T_n)$ and $f$ still satisfies the conditions of Theorem 1 where the sequences $(p_M)_{M \geq 1}$ and $(M_n)_{n \geq 1}$ remain the same. For each $k$, let $\mu_k$ denote the expectation $\mathbb{E} f(T_k)$ as before. By Lemma 6, we have

$$|\mu_k| = |\mathbb{E} f(T_k)| \ll p_{M_k} + k^{-1/2} M_k^2 \leq a_k. \quad (41)$$

For a positive integer $N$, let $f^{(N)}$ be the truncated functional defined by $f^{(N)}(T) = f(T) \mathbb{I}_{[|T| < N]}$ and let $F^{(N)}$ be the additive functional associated with the toll function $f^{(N)}$. It is important to notice that $f^{(N)}$ is local, for any fixed $N$. Note further that $\mathbb{E} f^{(N)}(T_k) = \mu_k$ if $k < N$, and $\mathbb{E} f^{(N)}(T_k) = 0$ otherwise. Hence, we have $|\mathbb{E} f^{(N)}(T_k)| \leq |\mu_k|$ for all positive integers $N$ and $k$. Let

$$W_{N,n} := \frac{F^{(N)}(T_n) - \mathbb{E} F^{(N)}(T_n)}{\sqrt{n}}, \quad \text{and} \quad X_n := \frac{F(T_n) - \mathbb{E} F(T_n)}{\sqrt{n}}.$$ 

Since $f^{(N)}$ has finite support, by [9, Theorem 1.5], we have

$$W_{N,n} \overset{d}{\rightarrow}_{n} \mathcal{N}(0, \gamma_N^2),$$

where

$$\gamma_N^2 = \lim_{n \rightarrow \infty} n^{-1} \text{Var} F^{(N)}(T_n)$$

$$= 2 \mathbb{E} \left( f^{(N)}(T) (F^{(N)}(T) - |T| \mu^{(N)}) \right) - \text{Var} f^{(N)}(T) - \frac{(\mu^{(N)})^2}{\sigma^2},$$

and $\mu^{(N)} = \mathbb{E} f^{(N)}(T)$. We remark that Janson actually has $\mu^{(N)}_N$ in the numerator of $W_{N,n}$ instead of $\mathbb{E} F^{(N)}(T_n)$ in [9], but it is also shown there that (8) is satisfied whenever the toll function has finite support, so this does not make a difference.
Next we need to show that \( \lim_{N \to \infty} \gamma_N \) exists. To that end, we take an arbitrary integer \( M \geq N \). We have

\[
\gamma_M - \gamma_N = \lim_{n \to \infty} n^{-1/2} \left( (\text{Var} F^{(M)}(T_n))^{1/2} - (\text{Var} F^{(N)}(T_n))^{1/2} \right).
\]

If we apply Minkowski’s inequality to the random variables \( F^{(M)}(T_n) - F^{(N)}(T_n) \) and \( F^{(N)}(T_n) \), we obtain

\[
(\text{Var} F^{(M)}(T_n))^{1/2} \leq \left[ \text{Var} \left( F^{(M)}(T_n) - F^{(N)}(T_n) \right) \right]^{1/2} + (\text{Var} F^{(N)}(T_n))^{1/2},
\]

and the same with \( M \) and \( N \) interchanged. Consequently,

\[
|\gamma_M - \gamma_N| = \lim_{n \to \infty} n^{-1/2} |(\text{Var} F^{(M)}(T_n))^{1/2} - (\text{Var} F^{(N)}(T_n))^{1/2}| \leq \limsup_{n \to \infty} n^{-1/2} \left( \text{Var} \left[ F^{(M)}(T_n) - F^{(N)}(T_n) \right] \right)^{1/2}.
\]

The toll function associated with the functional \( F^{(M)} - F^{(N)} \) is \( f^{(M)} - f^{(N)} \), which is zero for all trees of order smaller than \( N \). We apply Lemma 8 with \( I = \mathbb{N} \cap [N, M] \) to estimate the standard deviation \( (\text{Var}[F^{(M)}(T_n) - F^{(N)}(T_n)])^{1/2} \). We obtain

\[
|\gamma_M - \gamma_N| \ll \left( \sup_{k \geq N} a_k + \sum_{k = N}^{\infty} \frac{a_k}{k} \right)^{1/2} + \sup_{k \geq N} |\mu_k| + \sum_{k = N}^{\infty} \frac{|\mu_k|}{k}
\]

\[
\ll \left( \sup_{k \geq N} a_k + \sum_{k = N}^{\infty} \frac{a_k}{k} \right)^{1/2} + \sup_{k \geq N} a_k + \sum_{k = N}^{\infty} \frac{a_k}{k}.
\]

The last line follows from (41). By condition (6) of Theorem 1, we deduce that \( |\gamma_M - \gamma_N| \to_{N} 0 \) uniformly for \( M \geq N \). Hence, the sequence \( (\gamma_N)_{N \geq 1} \) is a Cauchy sequence, which implies that \( \gamma := \lim_{N \to \infty} \gamma_N \) exists and is finite.

Similarly, we have

\[
[\text{Var}(X_n - W_{N,n})]^{1/2} = n^{-1/2} \left( \text{Var} \left[ F(T_n) - F^{(N)}(T_n) \right] \right)^{1/2}.
\]

Once again, Lemma 8 applies here, where the set \( I \) is \( \mathbb{N} \cap [N, \infty) \). We obtain

\[
n^{-1/2} \left( \text{Var}[F(T_n) - F^{(N)}(T_n)] \right)^{1/2} \ll \left( \sup_{k \geq N} a_k + \sum_{k = N}^{\infty} \frac{a_k}{k} \right)^{1/2} + \sup_{k \geq N} a_k + \sum_{k = N}^{\infty} \frac{a_k}{k}.
\]

Therefore, we conclude that \( [\text{Var}(X_n - W_{N,n})]^{1/2} \) tends to zero as \( N \to \infty \) uniformly in \( n \), so Lemma 9 applies, where \( W_N \sim \mathcal{N}(0, \gamma_N^2) \) and \( W \sim \mathcal{N}(0, \gamma^2) \). With the estimate of \( \mathbb{E} F(T_n) \) in (8), which is proved in Proposition 7, we obtain (7) and the
The proof of Theorem 1 is complete. The variance estimate in Eq. (9) follows from the second part of Lemma 9.

6 Examples

In this section, we give several applications of our main theorem, focusing on functionals that are not covered by any previous results. The first example is treated in detail in Subsect. 6.1, where we prove a non-degenerate central limit theorem for the logarithm of the number of independent sets in \( T_n \). In other examples, we only verify that the conditions of Theorem 1 are satisfied by the corresponding functionals, without proving the non-degeneracy in each case (however, the approach of our first example can be applied to the others as well).

6.1 The Number of Independent Sets

An independent set is a set of nodes which does not contain two nodes that are adjacent. The number of independent sets (this number is also known as the Fibonacci number of \( T \); see [12]) was studied in the random plane tree by Kirschenhofer, Prodinger and Tichy [12] who determined a formula for its expectation (see also [13,18]). However, in order to obtain a limiting distribution, we study the logarithm of the number of independent sets rather than the number itself.

Let us formulate our result here as a theorem:

**Theorem 10** Let \( I(T) \) be the total number of independent sets of \( T \) and let \( T_n \) be a conditioned Galton–Watson tree of order \( n \) with offspring distribution \( \xi \), where \( \xi \) satisfies \( \mathbb{E}\xi = 1 \) and \( 0 < \sigma^2 := \text{Var} \xi < \infty \). There exist constants \( \mu > 0 \) and \( \gamma > 0 \) (both depending on \( \xi \)) such that

\[
\frac{\log I(T_n) - n\mu}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, \gamma^2)
\]

as \( n \to \infty \).

**Proof** Let \( I_0(T) \) be the number of independent sets of \( T \) that do not contain the root. The quantities \( I \) and \( I_0 \) satisfy the following recursive formulas, where \( T_1, \ldots, T_{\text{deg}(T)} \) stand for the branches:

\[
I_0(T) = \prod_i I(T_i), \quad (42)
\]

\[
I(T) = I_0(T) + \prod_i I_0(T_i). \quad (43)
\]

The first identity holds since every independent set of \( T \) that does not contain the root uniquely decomposes into independent sets in the branches. The second identity holds for essentially the same reason, taking into account those independent sets of \( T \) that contain the root, which can therefore not contain any of the roots of the branches.
Note that (42) and (43) are also satisfied by \( T = \bullet \), a tree consisting of a single node. Anticipating a log-normal limit distribution, we define an additive functional \( F(T) := \log I(T) \). From (42) and (43) it follows that the associated toll function is

\[
f(T) = F(T) - \sum_i F(T_i) = \log \left( \frac{I(T)}{\prod_i I(T_i)} \right) = \log \left( \frac{I(T)}{I_0(T)} \right)
\]

since \( I_0(T_i) \leq I(T_i) \), it follows immediately that \( 0 \leq f(T) \leq \log 2 \). Hence, the condition (3) of Theorem 1 is satisfied with \( \alpha = 0 \).

Further, let \( \rho(T) := \frac{I_0(T)}{I(T)} \). By (42) and (43), functional \( \rho \) also satisfies a recursion, namely

\[
\rho(T) = \frac{1}{1 + \prod_i \rho(T_i)}.
\]

Observe that (44) and (45) imply

\[
f(T) = -\log \rho(T).
\]

In order to measure the difference between \( f(T) \) and \( f(T^{(M)}) \) in terms of \( M \), we define the exact bounds on \( \rho \) given the first \( M \) levels:

\[
\rho_{\text{inf}}^M(T) := \inf \{ \rho(S) : S^{(M)} = T^{(M)} \}, \quad \rho_{\text{sup}}^M(T) := \sup \{ \rho(S) : S^{(M)} = T^{(M)} \},
\]

where \( S \) ranges over all rooted trees. From (46) it follows that for any tree \( T \)

\[
|f(T) - f(T^{(M)})| \leq \log(\rho_{\text{sup}}^M(T)/\rho_{\text{inf}}^M(T)) =: \tau^M(T).
\]

In view of (45) we have the trivial bounds \( 1/2 \leq \rho_{\text{inf}}^0(T) \leq \rho_{\text{sup}}^0(T) \leq 1 \), which imply

\[
\tau^0(T) \leq \log 2.
\]

For \( M \geq 1 \) the functions \( \rho_{\text{inf}}^M(T) \) and \( \rho_{\text{sup}}^M(T) \) can be determined recursively using (45), which gives

\[
\rho_{\text{sup}}^M(T) = \frac{1}{1 + \prod_i \rho_{\text{inf}}^{M-1}(T_i)} \quad \text{and} \quad \rho_{\text{inf}}^M(T) = \frac{1}{1 + \prod_i \rho_{\text{sup}}^{M-1}(T_i)}.
\]

Using (49) and writing \( \Pi := \prod_i \rho_{\text{sup}}^{M-1}(T_i) \) and \( \Sigma := \sum_j \tau^{M-1}(T_j) \), we get

\[
\tau^M(T) = \log \left( \frac{1 + \prod_i \rho_{\text{sup}}^{M-1}(T_i)}{1 + \prod_i \rho_{\text{inf}}^{M-1}(T_i)} \right) = -\log \left( \frac{1 + e^{-\Sigma \Pi}}{1 + \Pi} \right).
\]
Applying Jensen’s inequality to the convex function \( x \mapsto -\log x \) and using \( \Pi \leq 1 \), we deduce
\[
\tau^M(T) \leq \frac{\Pi}{1 + \Pi} \sum_{i} \tau^{M-1}(T_i).
\]
(50)

Let \( v_1, v_2, \ldots, v_{w_M(T)} \) be the nodes of \( T \) at level \( M \). By applying (50) recursively \( M \) times and using (48), we obtain a bound
\[
\tau^M(T) \leq 2^{-M} \sum_{i=1}^{w_M(T)} \tau^0(T_{v_i}) \leq \frac{\log 2}{2M} w_M(T).
\]
(51)

Combining (47) and (51) we obtain
\[
|f(T) - f(T^{(M)})| \leq \frac{\log 2}{2M} w_M(T).
\]
(52)

Now we are ready to verify that the remaining conditions of Theorem 1 are satisfied by our toll function. Note that for any \( N \geq M \), we have
\[
\mathbb{E} \left| f(\hat{T}^{(M)}) - \mathbb{E}\left(f(\hat{T}^{(N)}) \mid \hat{T}^{(M)}\right) \right| \leq \mathbb{E}\left(\mathbb{E}\left(|f(\hat{T}^{(M)}) - f(\hat{T}^{(N)})| \mid \hat{T}^{(M)}\right)\right).
\]
Using (52), we deduce that for any \( N \geq M \),
\[
\mathbb{E}\left(|f(\hat{T}^{(M)}) - f(\hat{T}^{(N)})| \mid \hat{T}^{(M)}\right) \leq \frac{\log 2}{2M} \mathbb{E}\left(w_M(\hat{T}^{(N)}) \mid \hat{T}^{(M)}\right).
\]

By taking the expectations, and using \( w_M(\hat{T}^{(N)}) = w_M(\hat{T}) \) as well as the estimate \( \mathbb{E}w_M(\hat{T}) = O(M) \) (see (12)), we get
\[
\mathbb{E}\left| f(\hat{T}^{(M)}) - \mathbb{E}\left(f(\hat{T}^{(N)}) \mid \hat{T}^{(M)}\right) \right| \ll M 2^{-M}.
\]
(53)

To check the condition (5) we use (52) and \( \mathbb{E}w_M(T_n) = O(M) \) (see (12)) and get
\[
\mathbb{E}|f(T_n) - f(T_n^{(M)})| = \mathbb{E}\left(\mathbb{E}\left(|f(T_n) - f(T_n^{(M)})| \mid T_n^{(M)}\right)\right)
\leq \mathbb{E}\left(\frac{\log 2}{2M} w_M(T_n)\right) \ll M 2^{-M},
\]
(54)

where the implied constant is independent of \( n \). To sum up, (53) and (54) show that assumptions (4) and (5) of Theorem 1 hold for a suitable choice of \( p_M \) and \( M_n \) with \( p_M \ll M 2^{-M} \) and \( M_n \ll \log n \), which implies that condition (6) is satisfied.

In the following, we show that the variance constant \( \gamma \) in Theorem 1 and in Eq. (9) is strictly positive for the functional \( F(T) = \log I(T) \). The approach that we use also applies (mutatis mutandis) to our other examples in the following sections, so we will not explicitly prove positivity of \( \gamma \) in all those cases.
As a first step, choose two trees $S_1$ and $S_2$ with the same number of nodes that both have a positive probability, i.e. $\mathbb{P}(T = S_1) > 0$ and $\mathbb{P}(T = S_2) > 0$, and also satisfy $I(S_1) > I(S_2)$ and $I_0(S_1) > I_0(S_2)$. This is always possible, for example in the following way: let $d$ be a possible outdegree for the given offspring distribution, i.e. $\mathbb{P}(\xi = d) > 0$. Now let $S_1$ be a complete $d$-ary tree of height 3 (the root has $d$ children, each of which has $d$ children, each of which has again $d$ children, which are leaves), and let $S_2$ be a $d$-ary caterpillar with the same number of nodes, consisting of $d^2 + d + 1$ internal nodes that form a path, and $d^3$ leaves (each internal node has $d − 1$ leaf children, except for the last, which has $d$ leaf children). One can verify that both inequalities hold for this choice of $S_1$ and $S_2$ for all $d \geq 2$.

The key observation is that replacing a fringe subtree isomorphic to $S_2$ in a tree by $S_1$ increases the number of independent sets by at least a fixed factor greater than 1. This is always possible, for example in the following way: let $d$ be a possible outdegree for the given offspring distribution, i.e. $\mathbb{P}(\xi = d) > 0$. Now let $S_1$ be a complete $d$-ary tree of height 3 (the root has $d$ children, each of which has $d$ children, each of which has again $d$ children, which are leaves), and let $S_2$ be a $d$-ary caterpillar with the same number of nodes, consisting of $d^2 + d + 1$ internal nodes that form a path, and $d^3$ leaves (each internal node has $d − 1$ leaf children, except for the last, which has $d$ leaf children). One can verify that both inequalities hold for this choice of $S_1$ and $S_2$ for all $d \geq 2$.

The key observation is that replacing a fringe subtree isomorphic to $S_2$ in a tree by $S_1$ increases the number of independent sets by at least a fixed factor greater than 1. To see this, suppose that $S$ is a fringe subtree of a tree $T$ rooted at $r$, let $T'$ be the tree obtained by removing the entire fringe subtree $S$ from $T$, and let $v$ be the parent of $r$ in $T'$. If $A$ is the number of independent sets of $T'$ that do not contain $v$, and $B$ the number of independent sets of $T'$ that contain $v$, then we have

$$I(T) = AI(S) + BI_0(S).$$

As a consequence of this representation, we find that the values of $I(T)$ for $S = S_1$ and $S = S_2$ differ at least by a factor of $\eta = \min \left\{ \frac{I(S_1)}{I(S_2)}, \frac{I_0(S_1)}{I_0(S_2)} \right\} > 1$ (and at most by $\max \left\{ \frac{I(S_1)}{I(S_2)}, \frac{I_0(S_1)}{I_0(S_2)} \right\}$).

Now consider a large random tree $T_n$ with $n$ nodes. We replace each occurrence of $S_1$ or $S_2$ as a fringe subtree by a marked leaf. The resulting tree, which has some number of marked leaves, is denoted by $T_n^\ast$. Given that $T_n^\ast$ has marked leaves $v_1, v_2, \ldots, v_m$, the original tree $T_n$ is obtained by replacing each marked node $v_i$ by a tree $R_i \in \{S_1, S_2\}$. Conditioned on $T_n^\ast$, the fringe subtrees $R_1, \ldots, R_m$ are independent and identically distributed with probabilities $\mathbb{P}(R_i = S_1) = \mathbb{P}(|T_n^\ast_{|S_i|} = S_1 | |T_n^\ast_{|S_i|} \in \{S_1, S_2\}) =: p > 0$ and $\mathbb{P}(R_i = S_2) = 1 − p > 0$; hence the distribution of each $R_i$ only depends on the choice of $S_1$ and $S_2$.

Still conditioning on $T_n^\ast$, we would like to determine a lower bound for the variance of $F(T_n) = \log I(T_n)$. Iterated application of the law of total variance yields

$$\text{Var}(F(T_n)|T_n^\ast) = \mathbb{E} \left( \text{Var}(F(T_n)|T_n^\ast, R_1, R_2, \ldots, R_{m-1})|T_n^\ast \right)$$

$$+ \sum_{j=2}^{m-1} \mathbb{E} \left( \text{Var}(\mathbb{E}(F(T_n)|T_n^\ast, R_1, R_2, \ldots, R_j)|T_n^\ast, R_1, R_2, \ldots, R_{j-1})|T_n^\ast \right)$$

$$+ \text{Var}(\mathbb{E}(F(T_n)|T_n^\ast, R_1)|T_n^\ast).$$

As mentioned before, replacing a fringe subtree isomorphic to $S_2$ by a fringe subtree isomorphic to $S_1$ increases the number of independent sets at least by a factor $\eta$ (thus increases the logarithm by at least $\log \eta$), regardless of the remaining shape of the tree. It is easy to see, that, given a realization of $T_n^\ast, R_1, \ldots, R_{j-1}$, the random variable
Algorithmica (2020) 82:642–679

\( \mathbb{E}(F(T_n)|T_n^*, R_1, \ldots, R_j) \) takes two values (note that for \( j = m \) this variable reduces to \( F(T_n) \)), at least \( \log \eta \) apart, the larger one with probability \( p \) and the other one with probability \( 1 - p \). Therefore its conditional variance is deterministically at least a fixed constant \( c := p(1 - p) \log^2 \eta \), which shows that

\[
\text{Var}(F(T_n)|T_n^*) \geq cm,
\]

uniformly for all possible choices of \( T_n^* \). Applying the law of total variance once again and recalling that random variable \( m \) is the total number of fringe subtrees isomorphic to \( S_1 \) or \( S_2 \), we find that

\[
\text{Var} F(T_n) \gg \mathbb{E} F_{S_1}(T_n) + \mathbb{E} F_{S_2}(T_n).
\]

The two functionals \( F_{S_1} \) and \( F_{S_2} \), counting fringe subtrees isomorphic to \( S_1 \) and \( S_2 \) respectively, are additive functionals whose means are linear in \( n \) with nonzero constants by our choice of \( S_1 \) and \( S_2 \) (see for example [9, (1.10)]). Therefore, it follows that

\[
\text{Var} F(T_n) \gg n,
\]

which, by Eq. (9), shows that \( \gamma > 0 \). Thus we have a non-degenerate central limit theorem for the logarithm of the number of independent sets.

There are various ways to see why \( \mu > 0 \) holds as well: for example, recall that \( \mu = \mathbb{E} f(T) \). Since the toll function \( f \) is strictly positive for all trees in this example, we must have \( \mu > 0 \). One can also use an indirect argument: if we had \( \mu \leq 0 \), then the central limit theorem would imply that \( \log I(T_n) < 0 \) with positive limiting probability, or equivalently \( I(T_n) < 1 \) with positive limiting probability. Since this is clearly impossible, we must have \( \mu > 0 \). The latter argument actually applies to all examples in this section. \( \square \)

**Remark 11** We remark that we actually have a stronger inequality for the constant \( \mu \), namely \( \mu \geq \log((1 + \sqrt{5})/2) \). This is because the number of independent sets of a tree with \( n \) nodes is always greater than or equal to the Fibonacci number

\[
F_{n+2} = \frac{1}{\sqrt{5}} \left( \left( \frac{1 + \sqrt{5}}{2} \right)^{n+2} - \left( \frac{1 - \sqrt{5}}{2} \right)^{n+2} \right).
\]

This was first established by Prodinger and Tichy in [14]. Thus \( \log F_{n+2} \) is also a deterministic lower bound on the mean of \( \log I(T_n) \), implying \( \mu \geq \log((1 + \sqrt{5})/2) \).

### 6.2 The Number of Matchings

The number of matchings (i.e. sets of edges no two of which share a node) in random trees has been studied previously, and means and variances have been determined for different classes of trees [13,18,19]. Just as in the previous example, in order to obtain a limiting distribution, we will consider the logarithm of this quantity. The approach
here is very similar to our approach in the previous example. We obtain the following result:

**Theorem 12** Let $m(T)$ be the total number of matchings of $T$ and let $T_n$ be a conditioned Galton–Watson tree of order $n$ with offspring distribution $\xi$, where $\xi$ satisfies $E\xi = 1$ and $0 < \sigma^2 := \text{Var} \xi < \infty$ as well as $E\xi^3 < \infty$. There exist constants $\mu > 0$ and $\gamma > 0$ (both depending on $\xi$) such that

$$\frac{\log m(T_n) - n\mu}{\sqrt{n}} \overset{d}{\to} \mathcal{N}(0, \gamma^2)$$

as $n \to \infty$.

**Proof** For a rooted tree $T$ let $m_0(T)$ be the number of matchings of $T$ that do not cover the root (by this, we mean matchings that do not contain an edge incident to the root). Using similar arguments as for the number of independent sets, one finds that these functionals satisfy the following recursive formulas (including the case when $T$ consists of a single node):

$$m_0(T) = \prod_i m(T_i), \quad (55)$$

$$m(T) = m_0(T) + \sum_i m_0(T_i) \prod_{j \neq i} m(T_j). \quad (56)$$

Defining an additive functional $F(T) := \log m(T)$, we observe from (55) and (56) that the associated toll function is

$$f(T) = F(T) - \sum_i F(T_i) = \log m(T) - \sum_i \log m(T_i) = -\log \left( \frac{m_0(T)}{m(T)} \right). \quad (57)$$

We define $\rho(T) := \frac{m_0(T)}{m(T)}$, which, by (55) and (56), also satisfies a recursion, namely

$$\rho(T) = \frac{1}{1 + \sum_i \rho(T_i)}. \quad (58)$$

From (57) it follows that $f(T) = -\log \rho(T)$, which, in view of (58), implies that $0 \leq f(T) \leq \log(1 + \text{deg}(T))$. Hence, condition (3) of Theorem 1 is satisfied by $f$ with $\alpha = 1$.

To estimate the distance between $T$ and $T^{(M)}$, we define

$$\rho_{\text{inf}}^M(T) := \inf \{ \rho(S) : S^{(M)} = T^{(M)} \}, \quad \rho_{\text{sup}}^M(T) := \sup \{ \rho(S) : S^{(M)} = T^{(M)} \},$$

where $S$ ranges over all rooted trees, so that

$$|f(T) - f(T^{(M)})| \leq \log \frac{\rho_{\text{sup}}^M(T)}{\rho_{\text{inf}}^M(T)} =: \tau^M(T). \quad (59)$$
The functionals $\rho^M_{\text{inf}}(T)$ and $\rho^M_{\text{sup}}(T)$, $M = 0, 1, 2, \ldots$, satisfy the recursions

$$\rho^M_{\text{sup}}(T) = \frac{1}{1 + \sum_i \rho^{M-1}_{\text{inf}}(T_i)} \quad \text{and} \quad \rho^M_{\text{inf}}(T) = \frac{1}{1 + \sum_i \rho^{M-1}_{\text{sup}}(T_i)}. \tag{60}$$

Using (60), and denoting $\rho_i = \rho^{M-1}_{\text{sup}}(T_i)$, we get

$$\tau^M(T) = -\log \left( \frac{1 + \sum_i \rho^{M-1}_{\text{inf}}(T_i)}{1 + \sum_i \rho^{M-1}_{\text{sup}}(T_i)} \right) = -\log \left( \frac{1 + \sum_i \rho_i \exp(-\tau^{M-1}(T_i))}{1 + \sum_i \rho_i} \right).$$

Since the argument of the logarithm on the right side is a convex combination of expressions $\exp(-\tau^{M-1}(T_i))$, $i = 1, 2, \ldots$, applying Jensen’s inequality to the convex function $x \mapsto -\log x$ yields, for $T$ with $\deg(T) \geq 1$,

$$\tau^M(T) \leq \frac{1}{1 + \sum_i \rho_i} \sum_i \rho_i \tau^{M-1}(T_i) \leq \frac{\max_i \rho_i}{1 + \max_i \rho_i} \sum_i \tau^{M-1}(T_i) \leq \frac{1}{2} \sum_i \tau^{M-1}(T_i). \tag{61}$$

But for $M \geq 1$, the inequality $\tau^M(T) \leq \frac{1}{2} \sum_i \tau^{M-1}(T_i)$ is also satisfied for the only tree $T$ with $\deg(T) = 0$. Unlike the previous example involving the number of independent sets, $\tau^0$ is not bounded by a constant, but the situation is saved by bounding $\tau^1$ by the root degree instead. From (60) it is clear that $\rho^{1}_{\text{sup}}(T) = 1$ and $\rho^{1}_{\text{inf}}(T) = (1 + \deg(T))^{-1}$ for every $T$. Therefore

$$\tau^1(T) = \log(1 + \deg(T)) \leq \deg(T). \tag{62}$$

Let $v_1, v_2, \ldots, v_{W_{M-1}}(T)$ be the nodes at level $M - 1$ of $T$. By iterating (61) $M - 1$ times and applying (62), we obtain

$$\tau^M(T) \leq 2^{-(M-1)} \sum_{i=1}^{w_{M-1}(T)} \tau^1(T_{v_i}) \leq 2^{-(M-1)} \sum_{i=1}^{w_{M-1}(T)} \deg(T_{v_i}) \leq 2^{-(M-1)} w_{M}(T). \tag{63}$$

Combining (59) and (63), we obtain

$$|f(T) - f(T^{(M)})| \leq 2^{-(M+1)} w_{M}(T). \tag{64}$$

Since (64) differs from the corresponding inequality (52) for the number of independent sets just by a constant factor, checking of the conditions (4) and (5) works precisely in the same way and shows that the conditions of Theorem 1 are again satisfied for some choice of $p_M$, $M_n$ satisfying $p_M \ll M 2^{-M}$ and $M_n \ll \log n$. \hfill \Box
6.3 The Number of Dominating Sets

Recall that a dominating set $D \subseteq V(T)$ is a set of nodes so that every node of the tree is either in $D$ or has a neighbour in $D$. Let $d(T)$ be the number of dominating sets in $T$. Extreme values of $d$ in trees were studied by Bród and Skupień [2].

**Theorem 13** Let $T_n$ be a conditioned Galton–Watson tree of order $n$ with offspring distribution $\xi$, where $\xi$ satisfies $\mathbb{E}\xi = 1$ and $0 < \sigma^2 := \text{Var}\xi < \infty$ as well as $\mathbb{E}\xi^3 < \infty$. There exist constants $\mu > 0$ and $\gamma > 0$ (both depending on $\xi$) such that

$$\frac{\log d(T_n) - n\mu}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, \gamma^2)$$

as $n \to \infty$.

Applying Theorem 1 to the number of dominating sets $d(T)$ is more complicated than the cases of independent sets and matchings. We consider two auxiliary parameters, defining $d_0(T)$ to be the number of dominating sets not containing the root and $d_*(T)$ to be the number of sets dominating everything except for the root (in particular such a set contains neither the root nor a child of the root). A bit of consideration reveals that the following recursive formulas are satisfied for any tree $T$ with branches $T_1, \ldots, T_{\text{deg}(T)}$, the reasoning being similar to the previous two examples:

$$d_*(T) = \prod_i d_0(T_i),$$

$$d_0(T) = \prod_i d(T_i) - \prod_i d_0(T_i),$$

$$d(T) = d_0(T) + \prod i (d(T_i) + d_*(T_i)).$$

Considering the additive functional $F(T) = \log d(T)$, we get that the corresponding toll function is

$$f(T) = F(T) - \sum_i F(T_i) = \log \frac{d(T)}{\prod_i d(T_i)} = \log \frac{d(T)}{d_0(T) + \prod_i d_0(T_i)} = \log \frac{d(T)}{d_0(T) + d_*(T)}.$$

Defining

$$\rho_*(T) := \frac{d_*(T)}{d(T)} \quad \text{and} \quad \rho_0(T) := \frac{d_0(T)}{d(T)},$$

from (65) we obtain

$$f(T) = -\log \left( \rho_*(T) + \rho_0(T) \right).$$

(66)
It is easy to see that $0 \leq \rho_*(T) \leq 1$ and $0 \leq \rho_0(T) \leq \frac{1}{2}$. This implies $f(T) \geq -\log(\frac{3}{2})$. On the other hand, using the recursive formulas for $d_0$ and $d$ as well as $\rho_* \leq 1$, we get

$$f(T) = \log \frac{d(T)}{\prod_i d(T_i)} = \log \left( \frac{\prod_i d(T_i) - \prod_i d_0(T_i) + \prod_i (d(T_i) + d_*)}{\prod_i d(T_i)} \right) \leq \log \left( 1 + \prod_i (1 + \rho_*(T_i)) \right) \leq \log \left( 1 + 2^{\deg(T)} \right) \leq (\log 2) \deg(T) + 1. \quad (67)$$

Hence condition (3) is satisfied with $\alpha = 1$.

Preparing to estimate the distance between $f(T)$ and $f(T^M)$, we first note that the functionals $\rho_*$ and $\rho_0$ satisfy the recursions

$$\rho_*(T) = \frac{\prod_i \rho_0(T_i)}{1 - \prod_i \rho_0(T_i) + \prod_i (1 + \rho_*(T_i))}, \quad (68)$$

$$\rho_0(T) = \frac{1}{1 - \prod_i \rho_0(T_i) + \prod_i (1 + \rho_*(T_i))}. \quad (69)$$

We further define

$$\rho_{0,\sup}^M(T) := \inf \{ \rho_0(S) : S^M = T^M \},$$

$$\rho_{0,\inf}^M(T) := \sup \{ \rho_0(S) : S^M = T^M \}, \quad (70)$$

$$\rho_{*,\sup}^M(T) := \inf \{ \rho_*(S) : S^M = T^M \},$$

$$\rho_{*,\inf}^M(T) := \sup \{ \rho_*(S) : S^M = T^M \}, \quad (71)$$

where $S$ ranges over all rooted trees as before. Let

$$\tau_{0}^M(T) := \log \frac{\rho_{0,\sup}^M(T)}{\rho_{0,\inf}^M(T)}, \quad \tau_*^M(T) = \log \frac{\rho_{*,\sup}^M(T)}{\rho_{*,\inf}^M(T)},$$

whenever the denominator is nonzero. If it is zero, then let the corresponding $\tau$ be 0 if the numerator also equals 0 and $\infty$ if the numerator is positive.

We further obtain a bound for $\tau_0^1$ in terms of $\deg(T)$. If $\deg(T) = 0$, that is $T = \bullet$ is the tree with a single node, then it is easy to see from the definitions that $\tau_0^1(T) = 0$. If $\deg(T) \geq 1$, then recursion (69) and the inequalities $\rho_0 \leq 1/2$ and $\rho_* \leq 1$ imply

$$\rho_0(T) \geq \frac{1 - 2^{-\deg(T)}}{1 - 2^{-\deg(T)} + 2^{\deg(T)}} \geq \frac{1}{2}, \quad (72)$$

whence $\tau_0^1(T) \leq \log \left( \frac{1}{2} + 2^{\deg(T)} \right) \ll \deg(T)$. We conclude that for all trees $T$, we have

$$\tau_0^1(T) \ll \deg(T) = w_1(T). \quad (73)$$
Unfortunately, it is not possible to bound $\tau_1(T)$, since whenever $T$ has a leaf of depth one, we have $d_1(T) = 0$. On the other hand, any other tree has $d_0(T) > 0$ whence $\tau_1(T) = \infty$. We can, however, bound $\tau_2(T)$. If $T$ has a leaf at depth one, then $\rho_*,\sup(T) = \rho_*,\inf(T) = 0$ and hence $\tau_2(T) = 0$. If instead we assume that no node of depth one is a leaf, then one can bound from below the numerator of (68) using (72) and bound from above the denominator of (68) using (73) and (74) we can bound 

$$
\tau_2(T) = \log \frac{\rho_*,\sup(T)}{\rho_*,\inf(T)} \leq \log \prod_i 1/(1 + 2^{\deg(T_i)+1}) \leq \deg(T) + \sum_i \deg(T_i) = w_1(T) + w_2(T).
$$

(74)

Our goal is now to show the following.

**Lemma 14** For some constants $a > 1$ and $0 < c < 1$, the functional $\eta^M(T) := a\tau_0^M(T) + \tau_*^M(T)$ satisfies, for $M \geq 3$,

$$
\eta^M(T) \leq \begin{cases} 
  c \sum_{v \in V_1(T)} \eta^{M-1}(T_v), & \deg(T) \geq 2, \\
  c^2 \sum_{v \in V_2(T)} \eta^{M-2}(T_v), & \deg(T) \leq 1,
\end{cases}
$$

where $V_i(T)$ are the nodes of $T$ at depth $i$.

Before we prove Lemma 14, let us show how it implies that the remaining conditions of Theorem 1 are satisfied.

By applying Lemma 14 recursively to $T_v$ for nodes $v$ at depth $M - 4$ or less, we are eventually left with a linear combination over nodes at depths $M - 3$ and $M - 2$, and hence we obtain a bound

$$
\eta^M(T) \leq c^{M-3} \sum_{v \in V_{M-3}(T)} \eta^3(T_v) + c^{M-2} \sum_{v \in V_{M-2}(T)} \eta^2(T_v).
$$

(75)

Note that by definition $\tau_0^M(T)$ and $\tau_*^M(T)$ increase as $M$ decreases. By using this together with (73) and (74) we can bound $\eta^3(T_v)$ and $\eta^2(T_v)$ by the number of nodes of $T_v$ at depth 1 or 2 (possibly multiplied by a constant), so that we obtain

$$
\eta^M(T) \ll c^M (w_{M-2}(T) + w_{M-1}(T) + w_M(T))
$$

(76)

and since $\tau_0^M(T), \tau_*^M(T) \leq \eta^M(T)$, using (66) we obtain

$$
|f(T) - f(T^M)| \leq \log \frac{\rho_{0,\inf}(T)^{e\tau_0^M(T)} + \rho_{*,\inf}(T)e^{e\tau_*^M(T)}}{\rho_{0,\inf}(T) + \rho_{*,\inf}(T)} \leq \log e^{\eta^M(T)} = \eta^M(T),
$$

(77)
which together with (76) and the same bounds for the expectations of $w_k$ that have been used in the previous examples shows that conditions (4) and (5) are satisfied for a suitable choice of $p_M$ and $M_n$ with $p_M \ll M e^M$ and $M_n \ll \log n$, so that (6) is satisfied.

**Proof of Lemma 14** Before considering the two cases $\deg(T) \geq 2$ and $\deg(T) \leq 1$, let us bound the ratio of $A_{\text{sup}} := \prod_i \left(1 + \rho_{*,\text{sup}}^m(T_i)\right)$ and $A_{\text{inf}} := \prod_i \left(1 + \rho_{*,\text{inf}}^m(T_i)\right)$, $m \geq 1$.

Denote $\rho_i = \rho_{*,\text{sup}}^m(T_i)$, $\tau_i = \tau_{*,\text{sup}}^m(T_i)$, and $\Sigma_* = \sum_i \tau_{*,\text{sup}}^m(T_i)$. Since $\rho_{*,\text{sup}}^m(T_i) = \rho_{*,\text{inf}}^m(T_i)$ when $\deg(T_i) = 0$, we can further assume that $\deg(T_i) \geq 1$ for all $i$. Hence, in view of recursion (68) and the bound $\rho_0(T) \leq 1/2$ we can assume $\rho_i \leq 1/2$. By definition, $\rho_{*,\text{inf}}^m(T_i) = \rho_i e^{-\tau_i}$, which gives

$$
\log \frac{A_{\text{sup}}}{A_{\text{inf}}} = \sum_i \left(- \log \frac{1 + \rho_i e^{-\tau_i}}{1 + \rho_i}\right) \leq \sum_i \frac{\rho_i}{1 + \rho_i} \tau_i \leq \frac{1}{3} \sum_i \tau_i = \frac{1}{3} \Sigma_*,
$$

(78)

where the first inequality follows from Jensen’s inequality applied to the convex function $x \mapsto -\log x$.

**First case:** $\deg(T) \geq 2$. Since the right-hand side of (69) is a decreasing function of each $\rho_0(T_i)$ and each $\rho_*(T_i)$, we have

$$
\tau_0^M(T) \leq \log \left(\frac{1 - \prod_i \rho_{0,\text{inf}}^{M-1}(T_i) + \prod_i \left(1 + \rho_{*,\text{sup}}^{M-1}(T_i)\right)}{1 - \prod_i \rho_{*,\text{sup}}^{M-1}(T_i) + \prod_i \left(1 + \rho_{*,\text{inf}}^{M-1}(T_i)\right)} \cdot \frac{1 - \prod_i \rho_{0,\text{inf}}^{M-1}(T_i)}{1 - \prod_i \rho_{0,\text{sup}}^{M-1}(T_i)}\right).
$$

Since $\prod_i \rho_{0,\text{inf}}^{M-1}(T_i) \leq \prod_i \rho_{0,\text{sup}}^{M-1}(T_i) \leq 1$ and $\prod_i \left(1 + \rho_{*,\text{inf}}^{M-1}(T_i)\right) \leq \prod_i \left(1 + \rho_{*,\text{sup}}^{M-1}(T_i)\right)$, the first fraction in the argument of the logarithm is at most $\prod_i \left(1 + \rho_{*,\text{sup}}^{M-1}(T_i)\right)/\prod_i \left(1 + \rho_{*,\text{inf}}^{M-1}(T_i)\right)$. Therefore, writing $\Pi = \prod_i \rho_{0,\text{sup}}^{M-1}(T_i)$ and $\Sigma_0 = \sum_i \tau_0^{M-1}(T_i)$, we get

$$
\tau_0^M(T) \leq \log \frac{\prod_i \left(1 + \rho_{*,\text{sup}}^{M-1}(T_i)\right)}{\prod_i \left(1 + \rho_{*,\text{inf}}^{M-1}(T_i)\right)} + \log \left(\frac{1 - \Pi e^{-\Sigma_0}}{1 - \Pi}\right).
$$

(79)

The first term is at most $\Sigma_*/3$ by (78) with $m = M - 1$. Turning to the second term in (79), and using $\Pi \leq 1/4$ (since $\deg(T) \geq 2$), we get

$$
\log \left(\frac{1 - \Pi e^{-\Sigma_0}}{1 - \Pi}\right) \leq \log \left(\frac{1 - \frac{1}{4} e^{-\Sigma_0}}{1 - \frac{1}{4}}\right) \leq \frac{1}{3} \Sigma_0,
$$

(80)

where the second inequality follows from Jensen’s inequality: $1 \leq \frac{3}{4} e^{\Sigma_0/3} + \frac{1}{4} e^{-\Sigma_0}$.
Combining (79), (78) with $m = M - 1$, and (80) we obtain

$$
\tau^M_0(T) \leq \frac{1}{3} \sum_i \tau^{M-1}_0(T_i) + \frac{1}{3} \sum_i \tau^*_M(T_i).
$$

(81)

We now proceed to $\tau^*_M(T)$. Since the right-hand side of (68) is increasing in each $\rho_0(T_i)$ and decreasing in each $\rho^*_0(T_i)$, we have

$$
\tau^*_M(T) \leq \log \left( \frac{\prod_i \rho^M_{0,\text{sup}}(T_i)}{\prod_i \rho^M_{0,\text{inf}}(T_i)} \cdot \frac{1 - \prod_i \rho^{M-1}_{0,\text{inf}}(T_i)}{1 - \prod_i \rho^{M-1}_{0,\text{sup}}(T_i)} \right).
$$

(82)

Keeping the notations $\Pi, A_{\text{sup}}, A_{\text{inf}}, \Sigma_0$ as above and noting that (78) implies $A_{\text{sup}} \leq A_{\text{inf}} e^{\Sigma_*/3}$, we have

$$
\tau^*_M(T) \leq \log \left( \frac{\prod_i \rho^M_{0,\text{sup}}(T_i)}{\prod_i \rho^M_{0,\text{inf}}(T_i)} \cdot \frac{1 - \prod_i \rho^{M-1}_{0,\text{inf}}(T_i)}{1 - \prod_i \rho^{M-1}_{0,\text{sup}}(T_i)} \right) + \log \left( \frac{1 - \Pi e^{-\Sigma_0} + A_{\text{sup}}}{1 - \Pi + A_{\text{inf}}} \right)
\leq \Sigma_0 + \log \left( \frac{1 - \Pi e^{-\Sigma_0} + A_{\text{inf}} e^{\Sigma_*/3}}{1 - \Pi + A_{\text{inf}}} \right)
\leq \Sigma_0 + \log \left( \frac{1 - \Pi e^{-\Sigma_0} + A_{\text{inf}} e^{\Sigma_*/3}}{1 - \Pi + A_{\text{inf}}} \right)
= \Sigma_0 + \frac{1}{3} \Sigma_* + \log \left( \frac{1 - \Pi e^{-\Sigma_0} + A_{\text{inf}}}{1 - \Pi + A_{\text{inf}}} \right).
$$

(82)

Further shortening $A = A_{\text{inf}}$, we note that the argument of the last logarithm is decreasing in $\Pi \geq 1$ and increasing in $\Pi \leq 1/4$, hence we obtain

$$
\frac{1 - \Pi e^{-\Sigma_0} + A}{1 - \Pi + A} \leq \frac{2 - \frac{1}{4} e^{-\Sigma_0}}{2 - \frac{1}{4}} = \frac{8 - e^{-\Sigma_0}}{7} \leq e^{\Sigma_0/7}
$$

where the last inequality follows from Jensen’s inequality: $\frac{7}{8}e^{\Sigma_0/7} + \frac{1}{8}e^{-\Sigma_0} \geq 1$. Putting the last estimate into (82), we conclude

$$
\tau^*_M(T) \leq \frac{8}{7} \Sigma_0 + \frac{1}{3} \Sigma_*.
$$

Now combining this inequality with (81) and choosing $a = 13/7$, say, we obtain

$$
\eta^M(T) \leq \left( \frac{13}{7} + \frac{8}{7} \right) \Sigma_0 + \left( \frac{13}{7} + \frac{1}{3} \right) \Sigma_* \leq \frac{20}{21} \Sigma_0 + \frac{20}{21} \Sigma_*

= \frac{20}{21} \sum_i \eta^{M-1}(T_i).
$$

(83)
Second case: \( \deg(T) \leq 1 \). Since the case \( \deg(T) = 0 \) is trivial, let us further assume \( \deg(T) = 1 \). Let us write \( T' \) for \( T \) with the root removed and let us denote the branches of the root of \( T' \) by \( T_1, T_2, \ldots \) (if there are no such branches, then \( T \) consists only of the root and a single leaf, and the lemma is also trivial). Using (69) and (68), after some straightforward simplifications we obtain

\[
\rho_0(T) = \frac{1 - \rho_0(T')}{2 - \rho_0(T') + \rho_0(T')} = \frac{\prod_i (1 + \rho_*(T_i))}{1 + 2 \prod_i (1 + \rho_*(T_i))}
\]

and

\[
\rho_*(T) = \frac{\rho_0(T')}{2 - \rho_0(T') + \rho_0(T')} = \frac{1 - \prod_i \rho_0(T_i)}{1 + 2 \prod_i (1 + \rho_*(T_i))}.
\]

By obvious monotonicity properties, we obtain that

\[
\tau_0^M(T) \leq \log \left( \frac{\prod_i (1 + \rho_{*,\sup}(T_i))}{\prod_i (1 + \rho_{*,\inf}(T_i))} \cdot \frac{1 + 2 \prod_i (1 + \rho_{*,\inf}(T_i))}{1 + 2 \prod_i (1 + \rho_{*,\sup}(T_i))} \right) \leq \log \left( \frac{\prod_i (1 + \rho_{*,\inf}(T_i))}{\prod_i (1 + \rho_{*,\sup}(T_i))} \right) \quad (84)
\]

Writing \( A = \prod_i (1 + \rho_{*,\inf}(T_i)) \), \( \Sigma_* = \sum_i \tau_{*,\inf}^M(T_i) \) and using (78) with \( m = M - 2 \) we get

\[
\tau_0^M(T) \leq \log \left( A e^{\Sigma_*/3} / A \right) = \frac{1}{3} \Sigma_*. \quad (85)
\]

On the other hand, writing \( \Pi = \prod_i \rho_{0,\sup}(T_i) \leq 1/2 \) and using (78) with \( m = M - 2 \), we get

\[
\tau_{*,\inf}^M(T) \leq \log \left( \frac{1 - \prod_i \rho_{0,\inf}(T_i)}{1 - \prod_i \rho_{0,\sup}(T_i)} \cdot \frac{1 + 2 \prod_i (1 + \rho_{*,\sup}(T_i))}{1 + 2 \prod_i (1 + \rho_{*,\inf}(T_i))} \right) \leq \log \left( \frac{1 - \Pi e^{-\Sigma_0}}{1 - \Pi} \right) + \log \left( \frac{\prod_i (1 + \rho_{*,\sup}(T_i))}{\prod_i (1 + \rho_{*,\inf}(T_i))} \right) \leq \log (2 - e^{-\Sigma_0}) + \frac{1}{3} \Sigma_*.
\]

(86)

The basic inequality \( e^{\Sigma_0} + e^{-\Sigma_0} \geq 2 \) implies \( \log(2 - e^{-\Sigma_0}) \leq \Sigma_0 \), which together with (86) implies

\[
\tau_{*,\inf}^M(T) \leq \Sigma_0 + \frac{1}{3} \Sigma_*.
\]

(87)

This, combined with (85), implies (recall we chose \( a = 13/7 \))

\[
\eta^M(T) = \frac{13}{7} \tau_0^M(T) + \tau_{*,\inf}^M(T) \leq \Sigma_0 + \frac{20}{21} \Sigma_* \leq \frac{20}{21} \sum_i \eta_{*,\inf}^{M-2}(T_i).
\]

(88)
Combining (88) with (83) completes the proof of Lemma 14 with $c = \sqrt{20/21}$. □

Now we know that all conditions of Theorem 1 are indeed satisfied (with $\alpha = 1$).

**Remark 15** As in the example of independent sets, we can use a deterministic lower bound on the number of dominating sets (due to Bród and Skupień, see [2]) to provide a better bound on $\mu$: in this example, we have $\mu \geq (\log 5)/3$ for all offspring distributions.

### 6.4 Tree Reductions

In a recent paper [4], Hackl, Heuberger, Kropf, and Prodinger studied various natural tree reduction procedures based on repeatedly deleting parts of a tree by certain operations until only the root remains. Before we describe these processes, we first need to define a few terms. Given a tree with root $o$, a leaf is called an *old leaf* if it is the leftmost child of its parent (with an exception that $o$ cannot be an old leaf). A *path* is a maximal (with respect to containment) fringe subtree with a single leaf, rooted at a node other than $o$. An *old path* is a maximal fringe subtree with a single leaf, rooted at a node other than $o$ which is the leftmost child of its parent. Note that although old leaves are leaves, an old path need not be a path (but is contained in a path; see Fig. 1).

In [4], the tree is reduced until only the root remains by repeating one of the following operations:

- (a) Leaf-reduction: all leaves are deleted in each round,
- (b) Old leaf-reduction: all old leaves are deleted in each round,
- (c) Path-reduction: all paths are deleted in each round,
- (d) Old path-reduction: all old paths are deleted in each round.

In any of the above operations, when a node is deleted, the incident edges are also deleted from the tree. An example is given in Fig. 1, where, for each of the four operations, the parts of the tree that are about to be deleted in the next round of reduction are dashed.

For a given positive integer $r$, and for a tree $T$, let $X_r(T)$ be the number of nodes in the reduced tree after the first $r$ steps of one of the above reductions. The authors of [4] proved asymptotic estimates for the mean and variance as well as a central limit theorem for $X_r(T_n)$ for the uniform random plane (=ordered) tree on $n$ nodes in the cases (a)–(c). For the case (d), they gave asymptotic estimates for mean and
variance of $X_r(T_n)$, but left the central limit theorem as an open problem. We show that asymptotic normality of the functional $X_r(T_n)$ can also be derived from Theorem 1 for conditioned Galton–Watson trees, of which the uniform random plane tree is a special case. We let

$$F_r(T) = |T| - X_r(T),$$

which corresponds to the number of nodes of $T$ other than the root that have been deleted after $r$ steps according to one of the reductions above. We obtain the following result:

**Theorem 16** Let $T_n$ be a conditioned Galton–Watson tree of order $n$ with offspring distribution $\xi$, where $\xi$ satisfies $E\xi = 1$ and $0 < \sigma^2 := \text{Var} \xi < \infty$. Consider one of the four reduction procedures described at the beginning of this subsection, and fix a positive integer $r$, so that $F_r(T)$ denotes the number of deleted nodes after $r$ steps of the procedure. Then, we have the following cases:

- if $F_r$ is the function in the reduction procedure (a) or (c) and $E\xi^3 < \infty$, then there exist constants $\mu > 0$ and $\gamma \geq 0$ (depending on $\xi$, the specific procedure and the value of $r$) such that

$$\frac{F_r(T_n) - n\mu}{\sqrt{n}} \xrightarrow{d} N(0, \gamma^2)$$

as $n \to \infty$, and

- if $F_r$ is the function in the reduction procedure (b) or (d), then there exist constants $\mu > 0$ and $\gamma \geq 0$ (depending on $\xi$, the specific procedure and the value of $r$) such that

$$\frac{F_r(T_n) - n\mu}{\sqrt{n}} \xrightarrow{d} N(0, \gamma^2)$$

as $n \to \infty$.

**Remark 17** There are some (trivial) degenerate cases for this example: consider for instance the leaf reduction procedure applied to binary trees. The number of nodes removed in the first step is deterministic since the number of leaves in a binary tree is. As for $\mu > 0$, consider a fixed star $S$ such that $P(T = S) > 0$, and let $N_S(T)$ denote the number of occurrences of $S$ as a fringe subtree of $T$. Thus, we have

$$F_r(T) \geq N_S(T),$$

since each occurrence of $S$ will lose at least one node in the first $r$ reductions in any of the four procedures. Furthermore, note that $N_S$ is an additive local functional (it coincides with $F_S$ defined in the introduction) with a toll function $f_S = 1_{\{S\}}$. Therefore, we have

$$\mathbb{E}F_r(T_n) \geq \mathbb{E}N_S(T_n) = n \mathbb{E}f_S(T) = n P(T = S) \gg n.$$
Let us now proceed to the proof of this result. Assume first that $F_r(T)$ is obtained from the reduction procedure (a) or (c) and that $\mathbb{E}\xi^3 < \infty$. The functional $F_r$ is additive with toll function $f_r$, where

$$f_r(T) = \sum_j \eta_T(T_j)$$

and the sum is over all branches $T_j$, with

$$\eta_T(T_j) = \begin{cases} 1 & \text{if the root of } T_j \text{ is deleted within the first } r \text{ steps,} \\ 0 & \text{otherwise.} \end{cases}$$

We can immediately see that

$$0 \leq f_r(T) \leq \deg(T).$$

The upper bound is clearly sharp for the leaf- and path-reductions. Hence, the condition $\mathbb{E}\xi^3 < \infty$ is required in this case.

Next, we show that $f_r$ satisfies the remaining conditions (i.e., (4), (5), (6)) of Theorem 1 in all four cases. For a rooted tree $T$, we denote by $T^*$ the planted tree where the root of $T$ is connected to a new node, which becomes the root of $T^*$. Let $\kappa = \min\{k \geq 2 : \mathbb{P}(\xi = k) > 0\}$ (this value must exist under our assumptions on $\xi$), and let $T_0$ be the complete $\kappa$-ary tree of height $r$. It is easy to verify that $X_r(T_0^*) \neq 1$, i.e. $T_0^*$ is not reduced to the root in $r$ steps. Moreover, by definition of $T_0$, it is easy to see that $\mathbb{P}(T = T_0) > 0$ which yields (using (2))

\begin{equation}
\mathbb{P}(T^{(r)} = T_0) > 0 \quad \text{and} \quad \mathbb{P}(\hat{T}^{(r)} = T_0) > 0.
\end{equation}

For each positive integer $M$, let $B_M$ be the set of all trees $T$ (not necessarily finite) of height at least $M - 1$ such that $X_r((T^{(M-1)})^*) = 1$ (i.e. the tree $T^{(M-1)}$ vanishes after the first $r$ steps of the reduction). It is important to notice here that a rooted tree $T$ is not reduced to a single node after the first $r$ steps of the reduction if the fixed tree $T_0$ appears as a subtree of $T$ (here, by subtree, we mean a subtree of the form $T_v^{(r)}$ for some node $v$ of $T$). This observation is key in the proof of the next lemma.

**Lemma 18** Assume that $\xi$ satisfies $\mathbb{E}\xi = 1$ and $0 < \text{Var} \xi < \infty$. Then there is a positive constant $c < 1$ that depends only on $\xi$ and $r$, such that (as $M \to \infty$)

$$\mathbb{P}(T \in B_M) \ll c^M \quad \text{and} \quad \mathbb{P}(\hat{T} \in B_M) \ll c^M.$$  

**Proof** Without loss of generality we assume $M \geq r + 1$. We start with the first estimate. We notice that for $T$ to satisfy the condition of event $B_M$, $T^{(r)}$ must not be equal to $T_0$. Moreover, since trees in $B_M$ have height at least $M - 1 \geq r$, $T \in B_M$ implies $w_r(T) \geq 1$. So
\[ P(T \in B_M) = \sum_{T \neq T_0, w_r(T) \geq 1} P(T^{(r)} = T) P(T \in B_M | T^{(r)} = T). \]

Conditioning on the event \( \{T^{(r)} = T\} \), the rest of \( T \) is a forest consisting of \( w_r(T) \) independent copies of \( T \). If \( T \in B_M \), then all of them must belong to \( B_{M-r} \), hence we obtain

\[ P(T \in B_M) \leq \sum_{T \neq T_0, w_r(T) \geq 1} P(T^{(r)} = T) P(T \in B_{M-r}) \]
\[ \leq \sum_{T \neq T_0} P(T^{(r)} = T) P(T \in B_{M-r}) \]
\[ = P(T \in B_{M-r}) P(T^{(r)} \neq T_0) =: q P(T \in B_{M-r}), \]

where \( q < 1 \) by (89). Iterating this inequality gives

\[ P(T \in B_M) \leq q^{\lfloor M/r \rfloor} \ll c_1^M, \tag{90} \]

where \( c_1 := q^{1/r} < 1 \), proving the first estimate.

For the second estimate, we begin in a similar fashion, i.e. we have

\[ P(\hat{T} \in B_M) = \sum_{T \neq T_0, w_r(T) \geq 1} P(\hat{T}^{(r)} = T) P(\hat{T} \in B_M | \hat{T}^{(r)} = T). \]

Here, when conditioning on the event \( \{\hat{T}^{(r)} = T\} \), the rest of \( \hat{T} \) is a forest consisting of \( w_r(T) - 1 \) independent copies of \( T \) and an independent copy of \( \hat{T} \). Thus,

\[ P(\hat{T} \in B_M) \leq \sum_{T \neq T_0, w_r(T) \geq 1} P(\hat{T}^{(r)} = T) P(T \in B_{M-r}) \]
\[ \leq \sum_{T \neq T_0} P(\hat{T}^{(r)} = T) P(\hat{T} \in B_{M-r}) \]
\[ = P(\hat{T}^{(r)} \neq T_0) P(\hat{T} \in B_{M-r}) \]

and by iterating as in the previous case and using the second inequality in (89) we obtain

\[ P(\hat{T} \in B_M) \ll c_2^M, \]

where \( c_2 := (P(\hat{T}^{(r)} \neq T_0))^{1/r} < 1 \). The proof is complete by choosing \( c = \max\{c_1, c_2\}. \quad \Box \)

For a finite tree \( T \), the only possibility for which \( f_r(T^{(M)}) \neq f_r(T) \) is when there is a branch \( T_j \) of \( T \) such that \( T_j^{(M-1)} \) vanishes after the first \( r \) steps of the reduction of
$T^{(M)}$, but $T_j$ does not vanish after the first $r$ steps of the reduction of $T$. This means that if $f_r(T^{(M)}) \neq f_r(T)$, then $T$ must have a branch in $B_M$. Since the branches of $T$ are copies of $T$, by conditioning on $\deg(T)$ and using the union bound we obtain

$$
P \left( f_r(T^{(M)}) \neq f_r(T) \right) \leq \sum_{k=1}^{\infty} P(\xi = k) \cdot kP(T \in B_M) \ll c^M,
$$

where the second inequality follows from Lemma 18 and $E\xi = 1$. As an immediate consequence of this, we have

$$
P \left( f_r(T^{(M)}) \neq f_r(T_n) \right) \leq \frac{P \left( f_r(T^{(M)}) \neq f_r(T) \right)}{P(|T| = n)} \ll n^{3/2} c^M,
$$

using the asymptotics $P(|T| = n) = \Theta(n^{-3/2})$, see [9, (4.13)]. Hence,

$$
E |f_r(T_n^{(M)}) - f_r(T_n)| \ll n^{3/2} c^M \max_{|T| = n} |f_r(T^{(M)}) - f_r(T)| \ll n^{5/2} c^M. \tag{91}
$$

Let us denote by $E_M$ the event $\bigcup_{N > M} \{f_r(\hat{T}^{(M)}) \neq \mathbb{E}(f_r(\hat{T}^{(N)}) | \hat{T}^{(M)})\}$. Then, for any $N \geq M$, we have

$$
|f_r(\hat{T}^{(M)}) - \mathbb{E}(f_r(\hat{T}^{(N)}) | \hat{T}^{(M)})| \ll \deg(\hat{T}^{(M)})I_{E_M}.
$$

For $\hat{T}$ to be in $E_M$, $\hat{T}$ must have a branch in $B_M$. Therefore,

$$
\mathbb{E} \left| f_r(\hat{T}^{(M)}) - \mathbb{E}(f_r(\hat{T}^{(N)}) | \hat{T}^{(M)}) \right| \\
\ll \sum_{k=1}^{\infty} kP \left( \deg(\hat{T}) = k \right) \left( (k - 1)P(T \in B_M) + P(\hat{T} \in B_M) \right). \tag{92}
$$

In view of Lemma 18, we have

$$
\mathbb{E} \left| f_r(\hat{T}^{(M)}) - \mathbb{E}(f_r(\hat{T}^{(N)}) | \hat{T}^{(M)}) \right| \ll c^M \sum_{k=1}^{\infty} k^2P \left( \deg(\hat{T}) = k \right) \ll c^M, \tag{93}
$$

since $\mathbb{E}(\deg(\hat{T})^2) < \infty$ if $\mathbb{E}\xi^3 < \infty$, see (14). The estimates (91) and (93) confirm that $f_r$ satisfies all conditions of Theorem 1 for a suitable choice of $p_M$ and $M_n \to \infty$ such that $p_{M_n} \ll c_M^3$ for some $c_3 \in (c, 1)$ and $M_n \ll \log n$.

In the case of procedures (b) or (d) we have $f_r(T) \leq r$, since in each step at most one child of the root is deleted. So (91) becomes

$$
\mathbb{E} |f_r(T_n^{(M)}) - f_r(T_n)| \ll n^{3/2} c^M \max_{|T| = n} |f_r(T^{(M)}) - f_r(T)| \ll n^{3/2} c^M,
$$
and the first inequality of (93) becomes

$$\mathbb{E} \left| f_r(\hat{T}^{(M)}) - \mathbb{E}(f_r(\hat{T}^{(N)} | \hat{T}^{(M)})) \right| \ll c^M \sum_{k=1}^{\infty} k \mathbb{P}\left(\text{deg}(\hat{T}) = k\right).$$

The sum on the right-hand side is equal to $\mathbb{E} \text{deg}(\hat{T})$, which is finite whenever $\text{Var} \xi < \infty$. The rest of proof is the same as in the previous case.

Acknowledgements We thank Jim Fill for useful remarks and the anonymous referees for very detailed reports and helpful comments which helped to improve the exposition.

References

1. Billingsley, P.: Convergence of Probability Measures. Wiley, New York (1968)
2. Bród, D., Skupień, Z.: Trees with extremal numbers of dominating sets. Australas. J. Comb. 35, 273–290 (2006)
3. Drmota, M.: Random Trees: An Interplay Between Combinatorics and Probability. Springer, Berlin (2009)
4. Hackl, B., Heuberger, C., Kropf, S., Prodinger, H.: Fringe analysis of plane trees related to cutting and pruning. Aequ. Math. 92(2), 311–353 (2018)
5. Holmgren, C., Janson, S.: Limit laws for functions of fringe trees for binary search trees and random recursive trees. Electron. J. Probab. 20(4), 1–51 (2015)
6. Holmgren, C., Janson, S., Šileikis, M.: Multivariate normal limit laws for the numbers of fringe subtrees in m-ary search trees and preferential attachment trees. Electron. J. Comb. 24(2), 49 (2017)
7. Janson, S.: Random cutting and records in deterministic and random trees. Random Struct. Algorithms 29(2), 139–179 (2006)
8. Janson, S.: Simply generated trees, conditioned Galton–Watson trees, random allocations and condensation. Probab. Surv. 9, 103–252 (2012)
9. Janson, S.: Asymptotic normality of fringe subtrees and additive functionals in conditioned Galton–Watson trees. Random Struct. Algorithms 48(1), 57–101 (2016)
10. Kallenberg, O.: Foundations of Modern Probability. Probability and its Applications (New York), 2nd edn. Springer, New York (2002)
11. Kesten, H.: Subdiffusive behavior of random walk on a random cluster. Ann. Inst. Henri Poincaré Probab. Stat. 22(4), 425–487 (1986)
12. Kirschenhofer, P., Prodinger, H., Tichy, R.F.: Fibonacci numbers of graphs. III. Planted plane trees. Fibonacci Numbers and Their Applications (Patras 1984) of Applied mathematics, vol. 28, pp. 105–120. Reidel, Dordrecht (1986)
13. Klazar, M.: Twelve countings with rooted plane trees. Eur. J. Comb. 18(2), 195–210 (1997)
14. Prodinger, H., Tichy, R.F.: Fibonacci numbers of graphs. Fibonacci Q. 20(1), 16–21 (1982)
15. Ralaivaosaona, D., Šileikis, M., Wagner, S.: Asymptotic normality of almost local functionals in conditioned Galton–Watson trees. In: 29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms, vol. 110 of LIPIcs. Leibniz International Proceedings in Informatics, Art. No. 33, pp. 17. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, (2018)
16. Ralaivaosaona, D., Wagner, S.: A central limit theorem for additive functionals of increasing trees. Comb. Probab. Comput. 28(4), 618–637 (2019)
17. Takács, L.: Conditional limit theorems for branching processes. J. Appl. Math. Stoch. Anal. 4(4), 263–292 (1991)
18. Wagner, S.: Correlation of graph-theoretical indices. SIAM J. Discrete Math. 21(1), 33–46 (2007)
19. Wagner, S.: On the number of matchings of a tree. Eur. J. Comb. 28(4), 1322–1330 (2007)
20. Wagner, S.: Central limit theorems for additive tree parameters with small toll functions. Comb. Probab. Comput. 24(1), 329–353 (2015)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.