Branching interlacements and tree-indexed random walks in tori

Qingsan Zhu

School of Mathematical Sciences, Tel Aviv University, Tel Aviv, 69978 ISRAEL
qingsanz@mail.tau.ac.il

Abstract. We introduce a model of branching interlacements for general critical offspring distributions. It consists of a countable collection of infinite tree-indexed random walk trajectories on $\mathbb{Z}^d$, $d \geq 5$. We show that this model turns out to be the local limit of the tree-indexed random walk in a discrete torus, conditioned on the size proportional to the volume of the torus. This generalizes the previous results of Angel, Ráth and the author, for the critical geometric offspring distribution. Our model also includes the model of random interlacements introduced by Sznitman as a degenerate case.

To obtain the local convergence, we establish results on decomposing large random trees into small trees, local limits of random trees around a prefixed vertex, and asymptotics of the visiting probability of a set by a tree-indexed random walk with a given size in a torus. These auxiliary results are interesting in themselves. As another application, we show that when $d \geq 5$ the cover time of a $d$-dimensional torus of side-length $N$ by tree-indexed random walks is concentrated at $N^d \log N^d / \text{BCap}\{0\}$.

Keywords: Branching interlacements · tree-indexed random walk · branching random walk · local limit · Galton-Watson tree · cover time · branching capacity.

1 Introduction

This article introduces a model of branching interlacements which consists of a countable collection of infinite tree-indexed random walk trajectories on $\mathbb{Z}^d$, $d \geq 5$. A non-negative parameter $u$ measures the amount of trajectories that enter the picture. The union of the range of these tree-indexed random walk trajectories defines the branching interlacement at level $u$. It is an infinite translation invariant random subset of $\mathbb{Z}^d$. This model is of special interest since it offers a microscopic description of the structure left by tree-indexed random walks. In fact, we show that the branching interlacement at level $u$ does appear as the local limiting distribution of the trace of the tree-indexed random walk in a large discrete torus, with a size proportional to the volume of the torus.

Let us first recall a similar model, the random interlacements model, which was introduced by Sznitman in [24]. This model describes the local picture left

* The author is supported by ISF grant 1207/15 and ERC starting grant 676970
by the trace of a random walk in a discrete torus when it runs up to times proportional to the volume of the torus. Roughly speaking, the model of random interlacements can be constructed via a Poisson point process with intensity measure \( \nu_0 \). The measure \( \nu_0 \) is supported on the space of doubly infinite nearest-neighbour trajectories modulo time shifts on \( \mathbb{Z}^d, d \geq 3 \). The parameter \( u \geq 0 \), referred to as the level, measures the amount of trajectories that enter the picture. The union of the range of trajectories contained in the support of this Poisson point process defines the random subset \( I^u \) of \( \mathbb{Z}^d \), called the random interlacement at level \( u \). Its law can be characterized as the unique distribution on \( \{0, 1\}^{\mathbb{Z}^d} \) such that

\[
P[I^u \cap K = \emptyset] = \exp(-u \cdot \text{Cap}(K)), \quad \text{for every finite } K \subseteq \mathbb{Z}^d,
\]

where \( \text{Cap}(K) \) is the (discrete) capacity of \( K \). In addition, there is an equivalent way to construct a set with the same law as \( I^u \cap K \).

Property 1. Let \( N_K \) be a Poisson random variable with parameter \( u \cdot \text{Cap}(K) \), and \( (X^j)_{j \geq 1} \) be i.i.d. random walks with the harmonic measure from infinity of \( K \) as the initial measure. Then \( K \cap \left( \bigcup_{j \geq 1} \text{Range}(X^j) \right) \) has the same distribution as \( I^u \cap K \).

On the other hand, branching interlacements (only) for the critical geometric offspring distribution were constructed in [4] (or see [32]). The construction is essentially based on the fact that the so-called contour walk of the Galton-Watson tree with the critical offspring distribution is just the simple random walk which is Markovian. This article extends this construction to general critical offspring distributions, by introducing two shift transformations and the corresponding invariant measures on a certain set of infinite trees.

Before entering the formal construction of our model, let us emphasize that either of the two characterizations above for random interlacements can be used to characterize the branching interlacement at level \( u \), except that we need to use the corresponding subjects for branching random walks. Branching capacity \( \text{BCap}(K) \) and the corresponding harmonic (entering) measure from infinity \( m_K \) for any finite subset \( K \subset \mathbb{Z}^d \) were introduced by the author in [28] (or see [32]). They are analogues to capacity and harmonic measure of random walk, in the setting of critical branching random walks. We write \( K \subset \mathbb{Z}^d \) to indicate that \( K \) is a finite subset of \( \mathbb{Z}^d \).

Theorem 1. For any \( u > 0 \), we can define a random subset of \( \mathbb{Z}^d \), called the branching interlacement at level \( u \) and denoted by \( \mathcal{I}^u \). Its law is characterized by

\[
P[\mathcal{I}^u \cap K = \emptyset] = \exp(-u \cdot \text{BCap}(K)), \quad \text{for any } K \subset \mathbb{Z}^d.
\]

Similarly, we have the following.

Proposition 2 Let \( N_K \) be a Poisson random variable with parameter \( u \cdot \text{BCap}(K) \), and \( (X^j)_{j \geq 1} \) be i.i.d. branching random walks with \( m_K \) as the initial measure. Then \( K \cap \left( \bigcup_{j \geq 1} \text{Range}(X^j) \right) \) has the same distribution as \( \mathcal{I}^u \cap K \).
We now describe our model. The model of branching interlacements is constructed via a Poisson point process on the space of infinite tree-indexed random walks modulo an equivalence relation. Hence, the key step for the construction is to produce a measure on that space as the intensity measure.

We consider tree-indexed random walks, also called random spatial trees or branching random walks. Fix a probability measure \( \mu \) on \( \mathbb{N} \) which serves as the offspring distribution (for the random tree mechanism) and a probability measure \( \theta \) on \( \mathbb{Z}^d \) which serves as the jump distribution (for the random walk mechanism). We assume that \( \mu \) is nondegenerate (i.e. not the Dirac mass at 1) with mean one and finite variance, and that \( \theta \) is with mean zero and finite range, not supported on any strict subgroup of \( \mathbb{Z}^d \). First, we introduce a probability measure \( \Pi_c \) on the space of infinite rooted ordered trees. In fact, this measure is supported on the subset, denoted by \( T_{\infty} \), of those trees which have a unique infinite ray from the root. Such infinite rays are called spines. The measure \( \Pi_c \) can be uniquely determined by the following property. For any \( t \in T_{\infty} \) and a non-root vertex \( v_0 \) of \( t \) in the spine, the probability under \( \Pi_c \) of being equal to \( t \) up to vertex \( v_0 \) is

\[
\frac{1}{2} \prod_v \mu(d(v) - 1),
\]

The product is over all vertices that are in the same component of the root, when deleting \( v_0 \); \( d(v) \) is the degree of \( v \); 1/2 is used for a normalization, i.e. to make \( \Pi_c \) a probability measure. We still need to define an equivalence relation, denoted by \( \cong_c \). Informally, \( t_1 \cong_c t_2 \) if and only if \( t_2 \) is \( t_1 \) re-rooted at another corner. Figure 1 explains how to re-root a tree at different corners. A key feature of \( \Pi_c \) is that this measure is invariant under the corner shift, which can be easily verified using (1.3).

**Fig. 1.** Re-root a tree at different corners
We now introduce a measure on the set of spatial trees. Let $\mathcal{T}_\infty^*$ be the set of all pairs $(t, S_t)$, where $t \in \mathbb{T}_\infty$ and $S_t : t \to \mathbb{Z}^d$ is a map from the vertex set $t$ into $\mathbb{Z}^d$. For simplicity, we also write $t$ for the vertex set of tree $t$. Similarly, we can define a measure $\Pi^*_c$ as follows. For any $(t, S_t) \in \mathcal{T}_\infty^*$ and one non-root vertex $v_0$ of $t$ in the spine, the total measure, under $\Pi^*_c$, of the set of all those spatial trees that are equal to $(t, S_t)$ up to $v_0$ is

$$\frac{1}{2} \prod_v \mu(d(v) - 1) \prod_{(v, v')} \theta(S_t(v') - S_t(v)). \tag{1.4}$$

The first product is, as before, over all vertices that are in the same component as the root, when deleting $v_0$; the second product is over all edges at this component, oriented in such a way that $v$ is closer to $v_0$ than $v'$. Note that $\Pi^*_c$ is not a probability measure, but a $\sigma$-finite measure. The mass of the set of all spatial trees that send the root to a prefixed point is one. Similarly, we can define an equivalence relation between spatial trees, also denoted by $\equiv$, $(t_1, S_{t_1}) \equiv (t_2, S_{t_2})$ if and only if $t_1 \equiv_c t_2$ and as maps, $S_{t_1} = S_{t_2}$. Note that when $t_1 \equiv_c t_2$, $t_2$ has the same vertex set as $t_1$ via a graph isomorphism. Therefore, $S_{t_1}$ and $S_{t_2}$ will have the same domain under this isomorphism. Similarly to the tree case, since both products are independent of the choice of the corner rooted, one can see that $\Pi^*_c$ is invariant under the corner shift.

Due to this invariance property, we can obtain a measure $\nu_\omega$ on $\mathcal{W} = \mathcal{T}_\infty^*/\equiv$, induced by $\Pi^*_c$. The branching interlacement at level $u$ is governed by a Poisson point process $\omega = \sum_{i \geq 0} \delta_{\omega_i}$ on $\mathcal{W}$, with intensity measure $2\nu_\omega$. Denote by $\mathbb{P}^u$ the law which turns $\omega$ into a Poisson point process with intensity $2\nu_\omega$. Then the branching interlacement at level $u$ is defined by

$$\mathcal{I}^u(\omega) = \bigcup_{i \geq 0} \text{Range}(\omega_i), \text{ if } \omega = \sum_{i \geq 0} \delta_{\omega_i},$$

where $\text{Range}(\omega_i)$ is the range of any element in the equivalence class of $\omega_i$. The vacant set at level $u$ is $\mathcal{V}^u = \mathbb{Z}^d \setminus \mathcal{I}^u$.

**Remark 1.** The reason why we put the constant factor “2” in the intensity $2\nu_\omega$ is to make (1.2) consistent with (1.1). In fact, for any $A \subset \subset \mathbb{Z}^d$, $\nu_\omega(\mathcal{W}_A) = \text{BCap}(A)/2$, where $\mathcal{W}_A$ is the set of all equivalence classes of those spatial trees that intersect $A$, cf. [2.3]. Note that $\mathbb{P}^u[\mathcal{I}^u \cap A = \emptyset] = \exp(-2\nu_\omega(\mathcal{W}_A))$ then (1.2) follows. Moreover, we will define another invariant measure $\Pi^*_c$ on (a proper subset of) $\mathcal{T}_\infty^*$, and thus obtain $\nu_\omega$ on the space of equivalence classes of spatial trees. It turns out that $\nu_\omega(\mathcal{W}_A) = 2\nu_c(\mathcal{W}_A)$ for any $A \subset \subset \mathbb{Z}^d$. Hence, using $\nu_\omega$ instead of $2\nu_c$, we also obtain the branching interlacement set with the same distribution, cf. Section 2. The factor “2” comes from the fact that on average, a vertex in a large tree has two corners.

Until now, we have not used the assumption that $d \geq 5$. In fact, we need this assumption to guarantee that when $u > 0$ $\mathcal{I}^u$ is nontrivial, i.e. a nonempty proper subset of $\mathbb{Z}^d$. It is showed in [29] that the spatial tree with law $\Pi^*_c$ (at
least when $\mu$ has finite third moment and $\theta$ is symmetric) almost surely visits every vertex on $\mathbb{Z}^d$ for $d \leq 4$. On the other hand, on $\mathbb{Z}^d, d \geq 5$, every finite subset is visited finitely many times almost surely. It is worth pointing out that most of our construction can be applied to a transient tree-indexed random walk attached to an infinite locally finite connected graph, see Remark 5.10.

Similarly to random interlacements, the branching interlacement set $I^u$ presents a long range dependence:

$$\text{Cov}(1_{x \in I^u}, 1_{y \in I^u}) \sim \frac{cu}{\|x-y\|^{d-4}}, \text{ as } \|x-y\| \to \infty.$$  

We now turn to our main result that the branching interlacement gives the local limit of the tree-indexed random walk in a large torus, conditioned on the size proportional to the volume of the torus. By a tree-indexed random walk with size $n$ and starting point $x \in \mathbb{Z}^d$, we mean a random spatial tree $(t, S_t)$ as follows. First, $t$ is a Galton-Watson tree with offspring distribution $\mu$, conditioned on the total number of vertices being $n$. Second, conditionally on $t$, the conditional probability weight of $S_t$ is determined by

$$1_{S_t(o) = x} \cdot \prod_{(v,v')} \theta(S_t(v') - S_t(v)),$$

where $o$ is the root and the product is over all edges, oriented in such a way that $v$ is closer to the root than $v'$. In fact, we will assume that $\theta$ is symmetric for the local limit result, and then the orientation is not needed. Denote by $\phi = \phi_N : \mathbb{Z}^d \to T_N = (\mathbb{Z}/N\mathbb{Z})^d$ the canonical projection map induced by mod $N$. Via this $\phi$, one can define the tree-indexed random walk in $T_N$ with size $n$ and a uniform starting point in an obvious way. In Theorem 14 we show that assuming further that $\mu$ has finite exponential moments and that $\theta$ is symmetric and aperiodic, for any $K \subset \subset \mathbb{Z}^d$, $u \in (0, \infty)$ and $n = n(N)$, an integer-valued function satisfying $\lim_{N \to \infty} n(N)/N^d = u$,

$$\lim_{N \to \infty} P[\text{Range}(S^{n,N}) \cap \phi(K) = \emptyset] = \exp(-uB\text{Cap}(K)),$$

where $S^{n,N}$ is a tree-indexed random walk in $T_N$ with size $n$ and a uniform starting point. By the inclusion-exclusion principle, this result implies the local convergence of the trace left by a tree-indexed random walk.

**Theorem 3.** Under the same assumptions above, for any $A \subset K \subset \subset \mathbb{Z}^d$, we have

$$\lim_{N \to \infty} P[\text{Range}(S^{n,N}) \cap \phi(K) = \phi(A)] = P[I^u \cap K = A].$$

We now briefly describe our idea to show the local convergence. The method is similar to the one for “the law of rare events”. Similarly to the case of the geometric distribution in [4], two main ingredients are needed. One is to find exact asymptotics for the visiting probability of a set in a torus by a tree-indexed random walk with a small size (cf. Theorem 15); the other is to decompose a large
random tree into small random trees (cf. Theorem 13). As mentioned earlier, in the geometric case, the contour walk of the random tree is just the simple random walk. Note that a (simple) random walk is Markovian, and that exact counting of paths is usually achievable for simple random walk. Moreover, in the geometric case, the corresponding random tree with a given size is invariant under the corner shift. In our case, we do not have such nice properties. This creates serious difficulties when one tries to establish the corresponding results. To overcome such difficulties, we switch to the so-called Lukasiewisz walks which still have the Markov property. One might compare the proofs in this article with the ones in [4].

It is worth pointing out that we also establish two intermediate results on random trees, which are interesting in their own right. On the one hand, as we mentioned earlier, we show (cf. Theorem 13) that, with high probability, a large random tree, precisely, a Galton-Watson tree conditioned on the total size, can be decomposed into small random trees with the same distributions as the Galton-Watson trees conditioned on size, without losing too many vertices. We can even require that those small subtrees are relatively far from each other.

On the other hand, since we aim to establish the local limit of tree-indexed random walks, we first establish a result (cf. Theorem 10) for the local limit of large random trees, i.e. the local picture around a vertex which is not too close to the root. It is well-known that the local limit around the root is given by the so-called Galton-Watson tree conditioned on survival. Local limits of large random trees have been studied in [1, 23]. However, the local limits constructed there are around a uniformly selected vertex in the random tree, while we consider the local limit around any prefixed vertex as long as it is not too close to the root, say, the \( m \)-th vertex in the Depth-first search order when \( m \) is not too small or too big. We also give bounds for the error terms.

Finally, as another application of our results on random trees and tree-indexed random walks in tori, we obtain a result (Theorem 16) on the cover time of a torus by tree-indexed random walks. Cover times of finite graphs by a simple random walk have been studied extensively (see e.g. [2, 5, 7]). One important example is the cover time of the discrete torus. The cover time \( \text{Cov}_N \) of the torus \( T_N = (\mathbb{Z}/N\mathbb{Z})^d \) is the time taken for a simple random walk to visit every vertex of \( T_N \). It is well-known that when \( d \geq 3 \), \( E[\text{Cov}_N] \sim g(0)N^d \log N^d \) where \( g \) is the discrete Green function, and \( \text{Cov}_N/(g(0)N^d \log N^d) \to 1 \) in probability. Recently, by constructing a strong coupling between random walks in tori and random interlacements, it is proved that the fluctuations of \( \text{Cov}_N \) are governed by the so-called Gumbel distribution [4]. The above is for cover times by random walks. However, we do not find any reference for cover times by tree-indexed random walks. It is even vague to define the cover time by tree-indexed random walks, since unlike random walk, there does not exist a simple way to couple tree-indexed random walks (random trees) with different sizes such that the smaller one is embedded into the bigger one. It is even not well-understood whether there exists such a coupling or not (see e.g. Problem 1.15 in [10] and [9, 16]). Nevertheless, we show that for \( d \geq 5 \), the cover time of \( T_N \)
is concentrated at $N^d \log N^d / B\text{Cap}(\{0\})$ in the following sense. For any $\epsilon > 0$, the probability that every vertex in $\mathbb{T}_N$ is visited by a tree-indexed random walk with size $n = n(N)$ goes to one if
\[ n(N) \geq (1 + \epsilon) N^d \log N^d / B\text{Cap}(\{0\}), \]
and zero if $n(N) \leq (1 - \epsilon) N^d \log N^d / B\text{Cap}(\{0\})$, as $N \to \infty$. Note that $g(0)N^d \log N^d / B\text{Cap}(\{0\}) = N^d \log N^d / B\text{Cap}(\{0\})$, and notice the obvious analogy between the results for a random walk and a tree-index random walk.

There are many natural questions about branching interlacements that are untouched in this article. For example, does the vacant set percolate or not? The answer is similar to the random interlacement case: there exists a critical value $u^* \in (0, \infty)$, such that when $u < u^*$ it does and when $u > u^*$, it does not, see the forthcoming paper [31]. Moreover, it is probable that by constructing suitable couplings between tree-indexed random walks in tori and branching interlacements, we could obtain new results on cover times as well as the largest component in the vacant set of the torus. Since the model of random interlacements has been constructed, a great effort has been made and many beautiful results have been established in that model itself and its relation with random walk (see e.g. [5, 18, 20–22, 25, 27] and the references therein). It seems that many results for random interlacements can be generalized to our model.

The article is organized as follows. In Section 2, we construct the model of branching interlacements. The main task is to construct the measures $\nu_c$ and $\nu_v$ entering the intensity of the Poisson point process that we are after. In Section 3, we consider random trees and establish our results of the local limit and of the decomposition of large random trees. Section 4 deals with asymptotics of the visiting probability by a tree-indexed random walk in a torus and the local convergence of tree-indexed random walks. Section 5 is devoted to the cover time result.

We finish this section with a remark on constants and notations. Throughout the text, we use $C, c, C_1, c_2$ etc. to denote positive constants depending only on the dimension $d$, the offspring distribution $\mu$ and the jump distribution $\theta$. These constants may change from place to place. Dependence of constants on additional parameters will be made or stated explicitly. For example, $C(\lambda)$ stands for a positive constant depending on $d, \mu, \theta$ and $\lambda$. For functions $f(x)$ and $g(x)$, we write $f \sim g$ if $\lim_{x \to \infty} (f(x)/g(x)) = 1$. We write $f \preceq g$ and $f \succeq g$, respectively, if there exists a constant $C$ such that, $f \leq Cg$ and $f \geq Cg$. We use $f \asymp g$ to express that $f \preceq g \preceq f \succeq Cg$. We write $f \ll g$ if $\lim_{x \to \infty} (f(x)/g(x)) = 0$. We use the notation $[a, b] = [a, b] \cap \mathbb{Z}$ for $a \leq b$. For any finite or infinite set $A$, we write either $|A|$ or $\#A$ for the cardinality of $A$. We write $a \wedge b$ for $\min\{a, b\}$.

2 Branching interlacements

2.1 Finite and infinite trees

We consider rooted ordered trees (also called plane trees). Recall that a tree is rooted if one vertex is distinguished as the root $o$. If $v \neq o$ is a non-root vertex, then the parent of $v$ is just the unique neighbour of $v$ closer to $o$ than $v$ (hence $o$
has no parent). Conversely, for any vertex \( v \), the neighbours of \( v \) that are further away from \( o \) than \( v \) are the children of \( v \). The number of children of \( v \) is the out-degree \( d_\circ(v) \) of \( v \). We write \( d(v) \) for the regular degree of \( v \). Recall further that a rooted tree is ordered if the children of each vertex are ordered.

We are interested in Galton-Watson trees (GW-trees) and their companions. Given a probability measure \( \mu \) on \( \mathbb{N} \), the GW-tree (with offspring distribution \( \mu \)) can be defined recursively, starting with the root and then giving each vertex, independently, a random number of children according to \( \mu \). Throughout this work, we fix \( \mu \) and always assume that \( \mu \) is critical (i.e. \( \sum_{i \in \mathbb{N}} i \mu(i) = 1 \)) and with finite variance \( \sigma^2 > 0 \).

We also need to consider the so-called GW-tree conditioned on survival, denoted by \( T^\infty \), defined as follows.

– Each vertex is normal or special.
– The root is special.
– A normal vertex produces only normal individuals, independently, according to \( \mu \).
– A special vertex produces individuals independently, according to the so-called size-biased distribution \( \tilde{\mu} \) (i.e. \( \tilde{\mu}(k) = k \mu(k) \)). One of them, chosen uniformly at random, is special; the others (if any) are normal.

For convenience, we denote the corresponding sample space by

\[
T^\infty = \{ t : t \text{ has a unique infinite path starting from the root; there are infinite vertices on each side of the infinite path} \}.
\]

For \( t \in T^\infty \), the unique infinite path starting from the root is called the spine of \( t \). Note that for any \( t \in T^\infty \), the depth-first search (from the root) explores only those vertices on the left hand side of \( t \), i.e. the vertices that are either in the spine or the descendants of those vertices which are elder siblings of the spine vertices (any vertex is regarded as a descendant and an ancestor of itself). For any \( t \in T^\infty \), we do have a full order, called the depth-first search order from infinity, on the vertex set \( t \) as follows. First divide \( t \) into two subsets \( A_1 \) and \( A_2 \). Let \( A_1 \) be the set of those vertices that are strictly on the left hand side of \( t \), i.e. the descendants of those vertices which are elder siblings of the spine vertices; let \( A_2 = t \setminus A_1 \). Any vertex in \( A_1 \) goes after any vertex in \( A_2 \). The order inside \( A_1 \) is consistent with the depth-first search order from the root. For any \( v_1 \neq v_2 \in A_2 \), let \( v'_1 \) (\( v'_2 \) be the furthest vertex in the spine, from the root, that is an ancestor of \( v_1 \) (\( v_2 \)). Then \( v_1 \) is before \( v_2 \) if and only if either \( v'_1 \) is a strict descendant of \( v'_2 \), or \( v'_1 = v'_2 \) and \( v_1 \) goes before \( v_2 \) in the depth-first search order from \( v'_1 \) in the subtree grafted to the right hand side of \( v'_1 \), i.e. the subtree generated by those vertices on the right hand side of the spine that are descendants of \( v'_1 \) but not descendants of the next spine vertex after \( v'_1 \). Intuitively, the depth-first search order from infinity is just the order of traversals by a particle from infinity travelling the tree along the edges from the right to the left. See Figure 3 for an illustration of this order.
2.2 Invariant measures on infinite trees

Let $T^c$ be the random tree defined as follows.

- Each vertex, except the root, is normal or special.
- The root produces $i$ individuals with probability $i\mu(i - 1)/2$. One of them, chosen uniformly at random, is special; the others (if any) are normal.
- A normal vertex produces only normal individuals, independently, according to $\mu$.
- A special vertex produces individuals independently, according to the size-biased distribution $\tilde{\mu}$. One of them, chosen uniformly at random, is special; the others (if any) are normal.

Denote by $\Pi_c$ the law of $T^c$. For any $t \in T_\infty$ and any non-root vertex $v$ of $t$ in the spine, $v$ divides $t$ into a finite number of components (subtrees). Write $\text{Comp}(t,v)$ for the component containing the root. For any $t_1, t_2 \in T_\infty$ and one non-root vertex $v_2$ in the spine of $t_2$, we say $t_1$ is equal to $t_2$ up to $v_2$, if there exists a vertex $v_1$ in the spine of $t_1$, such that there exists a plane tree isomorphism (keeping the graph, the root and the order structures) between $\text{Comp}(t_1,v_1)$ and $\text{Comp}(t_2,v_2)$, sending $v_1$ to $v_2$. One can easily check that for any $t \in T_\infty$ and $v_0$, a non-root vertex in the spine of $t$,

$$P[T^c \text{ is equal to } t \text{ up to } v_0] = \frac{1}{2} \prod_v \mu(d(v) - 1), \quad (2.1)$$

where the product is over all vertices in $\text{Comp}(t,v_0)$ excluding $v_0$.

We now define the corner shift $\tau_c$ on $T_\infty$. Intuitively, $\tau_c(t)$ is $t$ re-rooted at the next corner. The definition of corners is obvious when we draw a plane tree ‘standardly’ in a plane (see Figure 1 or Figure 2). Note that a vertex with degree $k$ has $k$ corners (around it). Formally, for $t \in T_\infty$, its image $\tau_c(t)$ can be obtained as follows. Let $v$ be the first child of the root $o$ in $t$. Re-root $t$ at $v$ and regard $o$ as the last child of $v$. Then this new tree (with root $v$) is just $\tau_c(t)$. Note that $\tau_c$ is a bijection on $T_\infty$ and that $\tau_c^{-1}(t)$ is $t$ re-rooted at the previous corner. Moreover, all corners of $t$ can be ordered in a two-sided sequence $(c_n)_{n \in \mathbb{Z}}$. Informally, we image a particle that starts from the root, and then explores the left hand side of the tree via the edges. The sequence of traversals gives $(c_n)_{n \geq 0}$. Imaging a particle exploring the right hand side of the tree gives $(c_{-n})_{n \geq 0}$. Then for any corner $c = c_i$, we define $\tau_c(t,c)$, called $t$ re-rooted at $c$, to be $\tau_c^{(i)}(t)$. One could draw $\tau_c(t,c)$ easily from the drawing of $t$ and the location of $c$ (see Figure 1 and Figure 2 for illustrations).

From (2.1), we obtain

**Proposition 4** $\Pi_c$ is invariant under $\tau_c$.

We now turn to another probability measure $\Pi_v$ and the corresponding shift $\tau_v$. Let $T^v$ ($\Pi_v$) be (the law of) the random tree defined as follows.

- Each vertex, except the root, is normal or special.
The root produces $i$ individuals with probability $\mu(i-1)$. The last individual is special; the others (if any) are normal.

A normal vertex produces only normal individuals, independently, according to $\mu$.

A special vertex produces individuals independently, according to the size-biased distribution $\tilde{\mu}$. One of them, chosen uniformly at random, is special; the others (if any) are normal.

Note that $\Pi_v$ is supported on a proper subset of $T_\infty$, denoted by $T'_\infty$, consisting of those trees in which the last child of the root is the one in the spine. Note that for any $t \in T_\infty$, there is a unique corner around $o$, such that the tree re-rooted at this corner belongs to $T'_\infty$. Similarly to (2.1), for any $t \in T'_\infty$ and $v_0$, a non-root vertex in the spine of $t$,

$$P[T^v \text{ is equal to } t \text{ up to } v_0] = \prod_v \mu(d(v) - 1), \quad (2.2)$$

where the product is over all vertices in $\text{Comp}(t, v_0)$ excluding $v_0$.

For any $t \in T'_\infty$ and a vertex $v \in t$, it is elementary to see that there is a unique corner around $v$, such that the tree $t$ re-rooted at this corner belongs to $T'_\infty$. We call this new tree $t$ re-rooted at $v$. For any $t \in T'_\infty$, we define $\tau_v(t)$ to be the tree $t$ re-rooted at the next vertex due to the Depth-first search order from infinity (this vertex is also the first vertex that is not in the spine, due to the Depth-first search order from the root). Note that $\Pi_v$ is a bijection on $T'_\infty$ and $\Pi^{-1}_v(t)$ is just $t$ re-rooted at the previous vertex due to the Depth-first search order from infinity. Therefore we can define $\tau_v^{(n)}$ for every $n \in \mathbb{Z}$. Similarly, since the right hand side of (2.2) is independent of the choice of the root, we have

**Proposition 5** \(\Pi_v\) is invariant under $\tau_v$. 

---

**Fig. 2.** An illustration of the corner shift.
Remark 2. The vertex shift $\tau_v$ and the corresponding measure $\Pi_v$ are related to the ones in [14] with the difference that only subtrees on the left side of the spine are considered there.

2.3 Random walk indexed by an infinite tree

Let $\theta$ be a probability measure on $\mathbb{Z}^d$, $d \geq 5$. Throughout this work, we fix $\theta$ and always assume that $\theta$ is centered (i.e. with mean zero), with finite range and not supported in a strict subgroup of $\mathbb{Z}^d$. Let $t \in T_\infty$. Since we have not assumed that $\theta$ is symmetric, we first need to specify the orientation of edges. Intuitively, the edge is oriented in such a way that the starting point is always closer to infinity than the end point. Formally, for the edges in the spine, the orientation is from the child to the parent; for all other edges, the orientation is from the parent to the child. We write $E(t)$ for the set of all oriented edges of $t$. The random walk indexed by $t$, starting at $x \in \mathbb{Z}^d$ (with jump distribution $\theta$) is a random function $S : t \to \mathbb{Z}^d$, such that the root is mapped to $x$ and the random variables $(S(v_2) - S(v_1))_{(v_1, v_2) \in E(t)}$ are independent and distributed according to $\theta$. Write $P^x_t$ for the law of the random walk indexed by $t$ starting at $x$.

Let $T'_\infty (T'_\infty)$ be the set of all pairs $(t, S)$ where $t \in T_\infty (t \in T'_\infty)$ and $S : t \to \mathbb{Z}^d$. We define $P^x_c (P^x_v)$ by declaring that $P^x_c (P^x_v)$ is the law of $(T, S)$, where $T$ is distributed according to $\Pi_c (\Pi_v)$ and conditionally on $T = t$, $S$ is distributed according to $P^x_t$.

For any $(t_1, S_1), (t_2, S_2) \in T_\infty (T'_\infty)$ and one non-root vertex $v_2$ of $t_2$ in the spine, we say $(t_1, S_1)$ is equal to $(t_2, S_2)$ up to $v_2$, if there exist a vertex $v_1$ in the spine of $t_1$ and a plane tree isomorphism $f$ between $\text{Comp}(t_2, v_2)$ and $\text{Comp}(t_1, v_1)$, such that $f(v_2) = v_1$ and $S_2(v) = S_1(f(v))$ for any $v$ in $\text{Comp}(t_2, v_2)$. From the definition of $P^x_c (P^x_v)$, one can see that, for any $(t, S_0) \in
\( \mathbb{T}_\infty \) or \( \mathbb{T}_\infty^* \) and one non-root vertex \( v_0 \) of \( t \) in the spine,

\[
P^x_C(T,S) \text{ is equal to } (t, S_0) \text{ up to } v_0 = \frac{1}{2} \prod_v \mu(d(v) - 1) \prod_{(v,v')} \theta(S_0(v') - S_0(v)),
\]

\[
P^x_C(T,S) \text{ is equal to } (t, S_0) \text{ up to } v_0 = \prod_v \mu(d(v) - 1) \prod_{(v,v')} \theta(S_0(v') - S_0(v)).
\]

In each formula, the first product is over all vertices in \( \text{Comp}(t, v_0) \) excluding \( v_0 \), and the second is over all oriented edges in \( \text{Comp}(t, v_0) \).

We now define the corresponding shift transformation \( \tau^*_C (\tau^*_C) \) on \( \mathbb{T}_\infty \) (\( \mathbb{T}_\infty^* \)). For \( t^* = (t, S_0) \in \mathbb{T}_\infty^* (\mathbb{T}_\infty^*), \) set \( \tau^*_C (t^*) = (t_1, S_1) \) \( (t_2, S_2) \), where \( t_1 = \tau_C(t) \) \( (t_2 = \tau_C(t)) \) and keep the spatial locations of \( t_1 \) \( (t_2) \). In fact, \( \tau^*_C (t^*) \) is \( t^* \) re-rooted at the next corner and \( \tau^*_C (t^*) \) is \( t^* \) re-rooted at the next vertex (under the Depth-first search order from infinity). Note that the right hand sides of the formulas in the last paragraph are independent of the choice of the corner or the vertex rooted. Hence we have (when both the old root and the new root are in \( \text{Comp}(t, v) \) excluding \( v \))

\[
P^x_C(t, S_0) \text{ is equal to } (t, S_0) \text{ up to } v
\]

\[
= P^x_C(t, S_0) \text{ is equal to } \tau^*_C((t, S_0)) \text{ up to } v, \]

\[
P^x_C(t, S_0) \text{ is equal to } (t, S_0) \text{ up to } v
\]

\[
= P^x_C(t, S_0) \text{ is equal to } \tau^*_C((t, S_0)) \text{ up to } v, \]

where \( o, o_1, o_2 \) are respectively, the roots of \( t, \tau_C(t), \tau_C(t). \) Hence we have

**Proposition 6** \( (P^x_C)_{x \in \mathbb{Z}^d} ((P^x_V)_{x \in \mathbb{Z}^d}) \) is invariant under \( \tau^*_C (\tau^*_V) \). Precisely, for any measurable subset \( A \) of \( \mathbb{T}_\infty \) or \( \mathbb{T}_\infty^* \),

\[
\sum_{x \in \mathbb{Z}^d} P^x_C(A) = \sum_{x \in \mathbb{Z}^d} P^x_C(\tau^*_C(A)), \text{ or } \sum_{x \in \mathbb{Z}^d} P^x_C(A) = \sum_{x \in \mathbb{Z}^d} P^x_V(\tau^*_V(A)).
\]

### 2.4 Construction of branching interlacements

We can construct the model of branching interlacements using either \( (\mathbb{T}_\infty^*, \tau^*_C, (P^x_C)_{x \in \mathbb{Z}^d}) \) or \( (\mathbb{T}_\infty^*, \tau^*_V, (P^x_C)_{x \in \mathbb{Z}^d}) \). There is no significant difference between them. We use \( (\mathbb{T}_\infty^*, \tau^*_C, (P^x_C)_{x \in \mathbb{Z}^d}) \) here and will mention the main difference later. Simply write

\( W_0 = \mathbb{T}_\infty^*, \tau = \tau^*_C, \text{ and } \tau_k = \tau(k), \text{ for } k \in \mathbb{Z}. \)

For any \( w \in W_0 \), we can view \( w \) as a function: \( \mathbb{Z} \to \mathbb{Z}^d \) such that

\( w(k) = \text{the image of the root of } \tau_k(w). \)

**Remark 3.** We now see the advantage of introducing \( \tau; \tau \) plays as a time-translation and \( w \) is like a two-sided random walk path (though not Markovian). Then we can follow the construction of random interlacements or the branching interlacements for the geometric case. In fact, if we let \( \mu \) be the degenerate probability measure or the critical geometric probability measure, we will recover random interlacements introduced in [24] (modifications needed for the sample space) or the branching interlacements for the geometric distribution in [4].
It is easy to see that (e.g. see Proposition 3.4 in [29]), for any $K \subset\subset \mathbb{Z}^d$ (when $d \geq 5$ as assumed), the set $\{n \in \mathbb{Z} : w(n) \in K\}$ is almost surely finite (under $P^v_x$ or $P^c_x$). Hence, we can concentrate only on those ‘transient’ trajectories and let

$$W = \{w \in W_0 : \lim_{|n| \to \infty} w(n) = \infty\}.$$

Define the set of spatial trees modulo time-shift by $W = W / \sim$, where $\sim$ is the equivalence relation $w_1 \sim w_2$, if $\tau_k(w_1) = w_2$ for some $k \in \mathbb{Z}$.

Denote the canonical projection by $\pi : W \to \overline{W}$ which sends each element in $W$ to its equivalence class in $\overline{W}$.

For any $K \subset\subset \mathbb{Z}^d$, let

$$W_K = \{w \in W : w(n) \in K, \text{ for some } n \in \mathbb{Z}\}, \quad \overline{W}_K = \pi(W_K). \quad (2.3)$$

For any $w \in W_K$, define the ‘entrance time’ by

$$H_K(w) = \inf\{n : w(n) \in K\}.$$

Since $w$ is ‘transient’, $H_K(w) \in (-\infty, \infty)$ when $w \in W_K$. We can partition $W_K$ according to the entrance time.

$$W_K = \bigcup_{n \in \mathbb{Z}} W^n_K, \quad \text{where } W^n_K = \{w \in W_K : H_K(w) = n\}.$$

Define $t_K : W_K \to W^n_0$ ($\overline{t}_K : \overline{W}_K \to \overline{W}^0_K$) by $t_K(w) = w_0$ ($\overline{t}_K(w) = w_0$), where $w_0$ is the unique element in $W^n_0$ with $w_0 \sim w (\pi(w_0) = \overline{w})$.

For every $K \subset\subset \mathbb{Z}^d$, we now define a measure $Q_K$ on $W$ by

$$Q_K(\bullet) = \sum_{x \in K} P^x_0(\bullet \cap W^n_K). \quad (2.4)$$

It turns out that for different $K \subset K' \subset \subset \mathbb{Z}^d$, $Q_K$ and $Q_{K'}$ are consistent in the following sense.

**Theorem 7.** For any measurable event $A \subseteq \overline{W}$, and $K \subseteq K' \subset\subset \mathbb{Z}^d$, we have

$$Q_K(\pi^{-1}(A) \cap W_K) = Q_{K'}(\pi^{-1}(A) \cap W_K). \quad (2.5)$$

Therefore we can define a measure on $\overline{W}$.

**Corollary 1.** There exists a unique $\sigma$-finite measure $\nu_v$ on $\overline{W}$ which satisfies:

for all $K \subset\subset \mathbb{Z}^d$,

$$\nu_v|_{\overline{W}_K} = Q_K \circ \pi^{-1}. \quad (2.6)$$
Proof (Proof of Theorem 7). Write \( B = t_K(\pi^{-1}(A) \cap W_K) \). Since \( Q_K(Q_{K'}) \) is supported on \( W^0_K \), we have
\[
Q_K(\pi^{-1}(A) \cap W_K) = Q_K(B) \quad \text{and} \quad Q_{K'}(\pi^{-1}(A) \cap W_K) = Q_{K'}(t_{K'}(B)).
\]
It suffices to show
\[
Q_K(B) = Q_{K'}(t_{K'}(B)). \tag{2.7}
\]
We partition \( W^0_K \) according to the entrance times and the entrance points of \( K \) and \( K' \). For any \( x \in K, y \in K' \) and \( n \in \mathbb{Z}^- = \{0, -1, -2, \ldots\} \), write
\[
A_{x,n,y} = \{ w \in W : w(0) = x, H_K(w) = 0, w(n) = y, H_{K'}(w) = n \}.
\]
On \( A_{x,n,y} \), \( t_{K'} \) is injective, \( t_{K'}(w)(\bullet) = w(\bullet + n) \) and
\[
t_{K'}(A_{x,n,y}) = \{ w \in W : w(0) = y, H_K(w) = -n, w(-n) = x, H_{K'}(w) = 0 \}.
\]
Let \( B_{x,n,y} = B \cap A_{x,n,y} \). Then \( B \) has a countable partition.
\[
B = \bigcup_{x \in K, y \in K', n \in \mathbb{Z}^-} B_{x,n,y}.
\]
In order to show (2.7), it suffices to show
\[
Q_K(B_{x,n,y}) = Q_{K'}(t_{K'}(B_{x,n,y})).
\]
By (2.4), one can get
\[
Q_K(B_{x,n,y}) = P^x_v[w(n) = y, H_K(w) = 0, w(0) = x, w \in \pi^{-1}(A)]
= P^y_v[w(-n) = x, H_K(w) = -n, H_{K'}(w) = 0, w(0) = y, w \in \pi^{-1}(A)]
= Q_{K'}(t_{K'}(B_{x,n,y})),
\]
where for the second line we use the fact that \( (P^x_v)_{x \in \mathbb{Z}^d} \) is invariant under \( \tau_v \), see Proposition 6.

We have constructed the measure \( \nu_v \) which will be used as the intensity measure in the construction of the branching interlacement Poisson point process. Before moving to the Poisson point process, let us calculate \( \nu_v(W_K) \).
\[
\nu_v(W_K) = Q_K(W_K) = Q_K(W^0_K) = \sum_{x \in K} P^x_v(W^0_K).
\]
Note that \( P^x_v(W^0_K) \) is just the escape probability \( \operatorname{Es}_K(x) \) introduced by the author in [28], \( \operatorname{BCap}(K) \) is also introduced there as the sum of \( \operatorname{Es}_K(x) \) over all \( x \in K \). Hence, we have

**Proposition 8**
\[
\nu_v(W_K) = \operatorname{BCap}(K). \tag{2.8}
\]
Remark 4. If we use the ‘last visiting time’ \( D_K(w) = \sup \{ n : w(n) \in K \} \) instead of the ‘first visiting time’ \( H_K(w) \), then we get another equivalent definition of \( \text{BCap}(K) \): \( \text{BCap}(K) = \sum_{x \in K} \text{Esc}_K(x) \). See [28] for the definition of \( \text{Esc}_K(x) \). Note that in [28], \( \sum_{x \in \mathbb{W}} \text{Esc}_K(x) = \sum_{x \in \mathbb{W}} \text{Esc}_K(x) \) is proved by establishing the exact asymptotics of the visiting probability of \( K \) and for general graphs, this method may fail since the visiting probability may behave quite differently.

We now introduce the space of locally finite point measures on \( \mathbb{W} \).

\[
\Omega = \left\{ \omega = \sum_{n \in I} \delta_{\overline{w}_n}, \text{ where } \overline{w}_n \in \mathbb{W}, \ I \subseteq \mathbb{N} \right\}, \quad \text{and } \omega(\mathbb{W}) < \infty \text{ for any } K \subset \subset \mathbb{Z}^d \right\}. \tag{2.9}
\]

For any \( u \in [0, \infty) \), let \( \mathbb{P}^u \) be the law of a Poisson point process on \( \Omega \) with intensity measure \( u \nu \). We can now define the branching interlacement at level \( u \).

**Definition 1.** The branching interlacement at level \( u \) is defined to be the random subset of \( \mathbb{Z}^d \) given by

\[
I^u = I^u(\omega) = \bigcup_{n \geq 0} \text{Range}(\overline{w}_n), \quad \text{where } \omega = \sum_{n \geq 0} \delta_{\overline{w}_n} \text{ has law } \mathbb{P}^u, \tag{2.10}
\]

and for \( \overline{w} \in \mathbb{W} \), \( \text{Range}(\overline{w}) = w(\mathbb{Z}) \), such that \( w \in \mathbb{W} \) is any element satisfying \( \pi(w) = \overline{w} \). The vacant set of branching interlacement at level \( u \) is defined by

\[
V^u = V^u(\omega) = \mathbb{Z}^d \setminus I^u(\omega). \tag{2.11}
\]

**Proposition 9** For any \( u \geq 0 \) and \( K \subset \subset \mathbb{Z}^d \), we have

\[
P[K \subseteq V^u(\omega)] = \exp(-u \cdot \nu_K(\mathbb{W})) \tag{2.12}
\]

Proof.\[
P[K \subseteq V^u(\omega)] = \exp(-u \cdot \nu_K(\mathbb{W})) \tag{2.12}
\]

**Remark 5.** 1. Using the inclusion-exclusion principle, one can see that (2.12) uniquely determines the laws of \( I^u \) and \( V^u \). Precisely, for any \( A \subseteq K \subset \subset \mathbb{Z}^d \), we have,

\[
P[I^u \cup K = A] = \sum_{B \subseteq A} (-1)^{|A \setminus B|} P[I^u \cup (K \setminus B) = \emptyset]
\]

\[
= \sum_{B \subseteq A} (-1)^{|A \setminus B|} \exp(-u \text{BCap}(K \setminus B)).
\]
2. As mentioned earlier, if we let $W = \mathbb{T}_\infty$, $\tau = \tau^*$ and replace $P_x^\tau$ by $P_x^\tau$ in (2.4), we can establish another measure $\nu_c$ on $W$. It is elementary to see that $\nu_c(W_K) = B\text{Cap}(K)/2$, for any $K \subset \mathbb{Z}^d$. Therefore, we can use $2\nu_c$ for the intensity measure in the construction of the branching interlacement set at level $u$ without changing its law.

3. Obviously $\nu_c$ is invariant under spatial translation: $\nu \to \nu + x$, for every $x \in \mathbb{Z}^d$. Hence, $I^u$ is space-translation invariant (i.e. $I^u \equiv I^u + x$ for every $x \in \mathbb{Z}^d$). On the other hand, $\nu_c$ is invariant under the time inversion while $\nu_v$ is not.

4. Since $I^u$ consists of infinite tree-indexed random walk trajectories, $V^u$ possesses the following screening effect: if $K'$ separates $K$ from infinity (with respect to the connectedness induced by $\theta$), then 

$$\{K' \subseteq V^u\} \subseteq \{K \subseteq V^u\}.$$

5. In [28], it is showed that the branching capacity of the cube (or the ball) $B_0(r)$ with radius $r$ behaves like

$$\text{BCap}(B_0(r)) \propto r^{d-4}.$$ 

Hence, for some positive constants $c, C$,

$$\exp(-Cru^{d-4}) \leq P[B_0(r) \subseteq V^u] \leq \exp(-cur^{d-4}).$$

In particular, there is generally not exponential decay with $|A|$ of $P[A \subseteq V^u]$.

6. Since $\text{BCap}(K \cup K') \leq \text{BCap}(K) + \text{BCap}(K')$, one can see that the events \{\{K \subseteq V^u\}\} and \{\{K' \subseteq V^u\}\} are positively correlated:

$$P[\{K \cup K' \subseteq V^u\}] \geq P[\{K \subseteq V^u\}]P[\{K' \subseteq V^u\}].$$

On the other hand, the FKG inequality holds for $V^u$. More precisely, for every pair of increasing random variables $X, Y$ with finite second moment with respect to the law of $V^u$, denoted by $Q^u$, we have

$$\int XYdQ^u \geq \int XdQ^u \int YdQ^u.$$

This can be proved in much the same way as in [26].

7. Using the results in [28], it is not difficult to get

$$\text{BCap}(\{x\}) + \text{BCap}(\{y\}) - \text{BCap}(\{x, y\}) \sim \frac{c}{\|x - y\|^{d-4}},$$

where $\|x\| = \sqrt{x \cdot Q^{-1}x}/\sqrt{d}$, with $Q$ being the covariance matrix of $\theta$. From this, one can see

$$\text{Cov}(1_{x \in I^v}, 1_{y \in I^v}) = \text{Cov}(1_{x \in V^u}, 1_{y \in V^u}) \sim \frac{cu}{\|x - y\|^{d-4}}$$

as $\|x - y\| \to \infty$. 

8. The constant "4" in the exponent comes from the fact that tree-indexed random walks (with regularity assumptions) behave like 4-dimensional subjects (while random walks behave like 2-dimensional subjects). This fact is reflected in many results on tree-indexed random walk, see e.g. [12–14, 19, 28–30, 32].

9. The measure $\nu_v|_{W_K \circ \stackrel{\sim}{t}^{-1} K}$ on $W_K^0$ is

$$\sum_{x \in K} \text{ES}_K(x) P^\nu_{x}[\bullet|H_K = 0] = \text{BCap}(K) \sum_{x \in K} \frac{\text{ES}_K(x)}{\text{BCap}(K)} P^\nu_{x}[\bullet|H_K = 0].$$

Note that the sum on the right hand site is a probability law (on the space of tree-indexed random walks). It is not difficult to verify that the entering measure (of $K$) by a tree-indexed random walk with this law coincides with the ‘branching harmonic measure’ of $K$, $m_K$, introduced in Section 9 in [28]. From this, one can obtain Proposition 2.

10. The constructions here for the random walk in $\mathbb{Z}^d, d \geq 5$ can be straightforwardly generalized to the case of an infinite locally finite connected graph $G = (V, E)$ with vertex set $V$ and (undirected) edge set $E$, endowed with positive weights $\lambda(e) > 0, e \in E$, such that the corresponding nearest neighbor walk on $V$ with transition probability

$$p(x, y) = \frac{\lambda(\{x, y\})}{\sum_{z: \{x, z\} \in E} \lambda(\{x, z\})}$$

is ‘branching transient’ in the sense that any finite subset of $V$ is visited finitely many times by the tree-indexed random walk with law $P^\nu_x$, almost surely. Note that the corresponding random walk on $V$ is reversible and has a stationary measure $\lambda(x) = \sum_{z: \{x, z\} \in E} \lambda(\{x, z\})$. In this set-up some of our definitions need to be slightly modified. For example, one should insert multipliers $\lambda(x)$ into the sum in (2.4).

3 Local limits of random trees and decomposing large random trees into small ones

In this section, we consider random trees (GW-trees) conditioned on size. Recall that we fix the offspring distribution $\mu$ and always assume that $\mu$ is critical (with mean one) and with finite variance $\sigma^2 > 0$. For simplicity, we assume further that span($\mu$) = 1 and leave the minor modifications when span($\mu$) > 1 to the reader. To make our arguments smooth, we concentrate mainly on the case when $\mu$ has finite exponential moments i.e. $\sum_{i \in \mathbb{N}} c^i \mu(i) < \infty$, for some $c > 1$. However, for the local limit result of random trees, we do not need this assumption. We write $T$ for an unconditioned GW-tree and $T_f$ for the sample space, i.e. the set of finite rooted order trees. We define $T^n$, called the GW-tree with size $n$, as $T$ conditioned on $|T| = n$ (recall that, for a tree $t$, we also write $t$ for its vertex set). For any $t \in T_f$, we adopt the depth-first search order (from the root) on $t$. 
We write \( V_i(t) \) (even \( V_i \) when \( t \) is obvious) for the \( i \)-th vertex of \( t \) starting from \( V_0(t) = a \), the root of \( t \). The main tool we will use is the so-called Lukasiewisz walks (L-walks). For any \( t \in T \) with \( |t| = n \), its L-walk \( L_t : \{0, 1, \ldots, n\} \to \mathbb{Z}^d \) is defined by \( L_t(0) = 0 \) and \( L_t(k) = \sum_{i=0}^{k-1} (d_o(V_i(t)) - 1) \), for \( k \in [1, n] \). It is easy to check that the L-walk of the GW-tree \( T \) is distributed as a random walk on \( \mathbb{Z} \) with jump distribution \( \mu_1 \) given by \( \mu_1(j) = \mu(j+1) \) for every \( j \geq -1 \), which starts at 0 and stops at the fist visiting time of \(-1\). In particular, \(|T|\) has the same distribution as the fist visiting time of such a random walk.

### 3.1 Some estimates on L-walks

In this subsection we collect useful lemmas about L-walks.

For a sequence \( x = (x_1, \ldots, x_n) \), the walk with step \( x \) is just the partial sum sequence \( s_0 = 0, s_1, \ldots, s_n \) where \( s_j = \sum_{i=1}^j x_i \). Say the walk fist visits \( b \) at time \( k \) if \( s_k \neq b \) for \( i < k \) and \( s_k = b \). For \( i \in \{1, \ldots, n\} \), the \( i \)-th cyclic shift of \( x \), denoted by \( x^{(i)} \), is the sequence of length \( n \) whose \( j \)-th term is \( x_{i+j} \) with \( x_{i+j} = x_{i+j-n} \) for \( i + j > n \).

**Lemma 1.** Let \( x = (x_1, \ldots, x_n) \) be a sequence with values in \( \mathbb{Z} \) and sum \(-k \) for some \( 1 \leq k \leq n \). Then there are at most \( k \) distinct \( i \in \{1, \ldots, n\} \) such that the walk with step \( x^{(i)} \) fist visits \(-k \) at time \( n \). Moreover, if the values of \( x_i \) are in \(-1, 0, 1, 2, \ldots \), then there are exactly \( k \) distinct such \( i \in \{1, \ldots, n\} \).

**Proof.** When \( x_i \in \{-1, 0, 1, \ldots\} \), it follows from Kemperman’s formula (see Lemma 6.1 [17]). For the general case, one can replace each \( x_i = -j \in \{-2, -3, \ldots\} \) by \( j \) successive \(-1\). Then the new sequence has exactly \( k \) cyclic shifts which first visit \(-k \) at the end. Therefore the original sequence has at most \( k \) cyclic shifts which first visit \(-k \) at the end.

Let \( L(n) = \sum_1^n X_i \) where \( X_i \) are i.i.d. with distribution \( \mu_1 \). Set \( H = H_{-1} = \inf\{k \geq 0, L(k) = -1\} \) the fist visiting time of \(-1\). By the lemma above, we get

\[
P(|T| = n) = P(H = n) = \frac{1}{n} P(L(n) = -1).
\]

If \( (T_i)_{i \in \mathbb{N}^+} \) are independent copies of \( T \), by concatenating their L-walks, one can get

\[
P \left( \sum_{i=1}^m |T_i| = n \right) = P(H_{-m} = n) = \frac{m}{n} P(L_n = -m), \tag{3.1}
\]

where \( H_{-m} \) is the fist visiting time of \(-m\).

We also need standard estimates for random walks. For \( x \in \mathbb{Z}, n \in \mathbb{N}^+ \), write

\[
p_n(x) = \frac{1}{\sqrt{2\pi\sigma}} \frac{1}{\sqrt{n}} \exp\left(-\frac{x^2}{2n\sigma^2}\right).
\]
Lemma 2. (See Theorem 2.3.9, Theorem 2.3.11 and Corollary A.2.7 in [11] and Lemma 2.1 in [10]) There exists $c$, such that

$$P(L(n) = x) = p_n(x) + o\left(\frac{1}{\sqrt{n}}\right),$$

$$P(L(n) = x) \leq \frac{1}{\sqrt{n}} \exp\left(-\frac{c^2}{n}\right), \text{ when } x \leq 0.$$ 

Moreover, if $\mu$ has finite exponential moments, then for some $c_1, c_2$,

$$P[L(n) = x] = p_n(x) \exp\left(O\left(\frac{1}{\sqrt{n}} + \frac{|x|^3}{n^2}\right)\right), \text{ when } |x| \leq c_1n,$$

$$P\left[\max_{0 \leq j \leq n} L(j) \geq r\sqrt{n}\right] \leq \exp\left(-c_2 r^2\right), \text{ when } r \in [0, c_1\sqrt{n}].$$

By this lemma and (3.1), we get

$$P[|T| = n] \sim \frac{1}{\sqrt{2\pi \sigma n^{3/2}}}, \quad (3.2)$$

$$P[\sum_{i=1}^{m} |T_i| = n] \leq \frac{m}{n^{3/2}} \exp\left(-\frac{cm^2}{n}\right), \quad (3.3)$$

$$P[\sum_{i=1}^{m} |T_i| = n] \sim \frac{1}{\sqrt{2\pi \sigma}} \frac{m}{n^{3/2}} \exp\left(-\frac{m^2}{2n\sigma^2}\right), \text{ when } m = O(\sqrt{n}). \quad (3.4)$$

Since we are interested in GW-trees with given sizes, from the view of $L$-walks, we need to study the random walk $L = (L(i))_{i \in \mathbb{N}}$ conditioned on $H_1 = n$. Write $P^n$ for the law of $L$ conditioned on $H_1 = n$.

Lemma 3. Assume that $\mu$ has finite exponential moments. Then, there exists $c > 0$ such that, for $m \in [1, n-1]$ and $h \in [0, cA]$ with $A = m \wedge (n-m)$,

$$P^n[L(m) = h] \leq \frac{(h + 1)^2}{A^{3/2}} \exp(-c\frac{(h + 1)^2}{A}),$$

$$P^n[L(m) \leq h] \leq \frac{(h + 1)^3}{\sqrt{A}},$$

$$P^n[L(m) \geq h] \leq \exp(-c\frac{(h + 1)^2}{A}).$$

Proof.

$$P^n(L(m) = h) = \frac{P[H = n, L(m) = h]}{P[H = n]}$$

$$= \frac{P[L(m) = h; L(i) \geq 0, i \in [0, m]] P_h[H = n - m]}{P[H = n]},$$

where $P_h$ is the law of $L$ when staring at $h$ and we use the Markov property in the last line.
By the equations after Lemma 1, we have
\[ P_h[H = n - m] = P(H_{-h-1} = n - m) \]
\[ = \frac{h + 1}{n - m} P(L(n - m) = -h - 1) \leq \frac{h + 1}{(n - m)^{3/2}} \exp\left(-c\frac{(h + 1)^2}{n - m}\right), \]

\[ P[H = n] \sim \frac{1}{n^{3/2}}. \]

For the other term, by inserting one up-step and then reversing the walk, one can see (for any fixed \( j \) with \( \mu_{-1}(j) > 0 \))
\[ P[L(m) = h; L(i) \geq 0, i \in [0, m]] \]
\[ \leq P[L(m + 1) = h + j; L(i) \geq j, i \in [1, m + 1]] / \mu_{-1}(j) \]
\[ \leq P_{h+j} [L^{-}(m + 1) = 0; L^{-}(i) \geq 1, i \in [0, m]] / \mu_{-1}(j) \]
\[ \leq \frac{h + j}{m + 1} P(L^{-}(m + 1) = -h - j) \leq \frac{h + 1}{m^{3/2}} \exp\left(-c\frac{(h + 1)^2}{m}\right), \]

where \( L^{-} \) is the reversed random walk (i.e. with jump distribution \( \mu_{-1}(\bullet) \equiv \mu_{-1}(\bullet) \) and in the last line we apply Lemma 1 and Lemma 2.

From what have been proved, it follows that
\[ P^n(L(m) = h) \leq \frac{(h+1)^2}{(m(n-m))^{3/2}} \exp\left(-c\frac{(h+1)^2 + (h+1)^2}{n-m}\right) \]
\[ \leq \frac{(h+1)^2}{A^{3/2}} \exp\left(-c\frac{(h+1)^2}{A}\right). \]

This completes the proof of the first assertion. By summation and the last lemma, one can get the other assertions.

**Lemma 4.** Let \( k = k(n) \in [[1, n-1]] \) with \( A = A(n) = k(n) \land (n - k(n)) \to \infty. \)
Then
\[ \lim_{n \to \infty} P(V_k \text{ in } T^n \text{ is a leaf}) = \mu(0). \]

Moreover, if \( \mu \) has finite exponential moments, then
\[ \sup_{n,k \in \mathbb{N}^+; A < k < n - A} P(V_k \text{ in } T^n \text{ is a leaf}) = \mu(0) \left(1 + O\left(\frac{1}{A^{3/8}}\right)\right). \]

**Proof.** It is standard that for any \( \epsilon > 0 \), we can find \( \delta \in (0, 1/2) \), such that (when \( A \) is sufficiently large)
\[ P^n(L(k) \in [\delta \sqrt{A}, \frac{1}{\delta} \sqrt{A}]) < \epsilon. \]

Conditionally on the event \( L_{T^n}(k) = h \), we have (when \( h \in [[\delta \sqrt{A}, \delta^{-1} \sqrt{A}]] \))
\[ P[V_k \text{ in } T^n \text{ is a leaf}|L_{T^n}(k) = h] = \frac{P[H = n, L(k) = h, L(k + 1) = h - 1]}{P[H = n, L(k) = h]} \]
\[ = \frac{\mu(0) P_{h-1}[H = n - k - 1]}{P_h[H = n - k]} = \frac{\mu(0) \frac{h}{n-k-1} P(L(n - k - 1) = -h)}{P_h[H = n - k] \frac{h+1}{n-k} P(L(n-k) = -h - 1)} \to \mu(0), \]
where we use the Markov property and Lemma \[ \text{2} \]. The proof of the first assertion is complete.

When \( \mu \) has finite exponential moments. By the last lemma, we can find large \( C \), such that
\[
P^n[L(k) \in [A^{3/8}, C \sqrt{A \log A}]] \leq \frac{1}{A^{4/8}}.
\]
Similarly, we have (when \( h \in ([A^{3/8}, C \sqrt{A \log A}] )
\[
P(V_k \text{ in } T^n \text{ is a leaf}[L_T^n (k) = h]) = \frac{\mu(0) \frac{h}{n-k+1} P(L(n-k-1) = -h)}{\frac{h+1}{n-k} P(L(n-k) = -h-1)} = \frac{n-k}{h+1} \frac{\bar{p}_{n-k-1}(-h)(1 + O(1/(n-k-1)^{0.4}))}{\bar{p}_{n-k}(-h-1)(1 + O(1/(n-k)^{0.4}))} = \mu(0)(1 + O(\frac{1}{A^{3/8}})).
\]
We finish the proof of the second assertion.

### 3.2 Local limits of large Galton-Watson trees at a prefixed vertex

Let us introduce an unrooted random tree \( \mathcal{T} \), which appears in \[ \text{123} \]. Intuitively, \( \mathcal{T} \) is \( T^v \) re-rooted at infinity. Formally, \( \mathcal{T} \) is defined as follows. Start with a semi-infinite path \( u_0, u_1, \ldots \). Let \( u_0 \) be the root of an independent copy of the unconditioned GW-tree. For each \( i \geq 1 \), \( u_i \) produces, independently, a random number of children according to \( \tilde{\mu} \), the size-biased probability measure of \( \mu \). Let \( u_{i+1} \) be identified with a uniformly chosen child of \( u_i \) and all other children of \( u_i \) become the root of an independent copy of the unconditioned GW-tree.

We follow the notations used in \[ \text{23} \]. Let \( t \in \mathcal{T}_l \), \( v \in t \), and \( l \in \mathbb{N} \). When the vertex \( v \) has an \( l \)-th ancestor \( v_l \), we define the pointed plane tree \( H_l(t, v) \) as \( [t]_{v_l} \) (we write \( [t]_v \) for the subtree of \( t \) generated by all descendants of \( v \), rooted at \( v \)). Otherwise, when the vertex \( v \) has no \( l \)-th ancestor, we simply set \( H_l(t, v) = o \) for some inessential symbol \( o \). Let \( H_l(\mathcal{T}, u_0) \) be the subtree of \( \mathcal{T} \) generated by all descendants of \( u_0 \), rooted at \( u_0 \) and pointed at \( u_0 \).

We now state our theorem.

**Theorem 10.** For any \( k = k(n) \in [1, n/2] \) and \( l = l(n) \in [0, k(n)] \) satisfying
\[
\lim_{n \to \infty} k(n) = \infty \text{ and } \lim_{n \to \infty} k(n)/(l(n) + 1)^2 = \infty.
\]
Then we have
\[
\lim_{n \to \infty} \sup_{m \in [k(n), n-k(n)]} d_{TV}(H_l(T^n, V_m), H_l(\mathcal{T}, u_0)) = 0,
\]
where \( d_{TV}(\bullet, \bullet) \) is the total variance distance. Moreover, if \( \mu \) has finite exponential moments, then we have the following bounds for the error term:
\[
\sup_{m \in [k(n), n-k(n)]} d_{TV}(H_l(T^n, V_m), H_l(\mathcal{T}, u_0)) \leq \begin{cases} \frac{1}{k^{3/8}} + \left(\frac{l+1}{\sqrt{n}}\right)^{2/3}, & \text{when } ((l+1)n)^{2/3} \leq k; \\ \frac{l+1}{k^{3/8} \sqrt{n}} + \frac{k}{n}, & \text{when } ((l+1)n)^{2/3} \geq k. \end{cases}
\]
We give two lemmas before proving the theorem.

Lemma 5. \[ P(|H_l(\mathcal{X}, u_0)| \geq N) \leq \frac{l + 1}{\sqrt{N}}. \]

Proof. Note that \( |H_l(\mathcal{X}, u_0)| \) can be written as the sum of \( |H_0(\mathcal{X}, u_0)| \) and the independent differences

\[ |H_i(\mathcal{X}, u_0)| - |H_{i-1}(\mathcal{X}, u_0)| \overset{d}{=} S_{\xi - 1}, \quad i = 1, \ldots, l; \]

where \( \xi \) has the size-biased distribution \( \tilde{\mu} \) and \( S_m = \xi_1 + \cdots + \xi_m + 1 \)

with \( (\xi_j)_{j \in \mathbb{N}} \) being independent copies of \( |T| \). Hence we can bound \( |H_l(\mathcal{X}, u_0)| \) stochastically by \( S_{M_l} \) with \( M_l = \xi_1 + \cdots + \xi_{l+1} \),

the sum of \( l + 1 \) independent copies of \( \hat{\xi} \). By (3.3),

\[ P(S_m \geq n) \leq \sum_{i} P(M_l = i) P(S_i \geq n) \leq \sum_{i} P(M_l = i) \frac{i}{\sqrt{N}} = \frac{(l + 1)E[\hat{\xi}]}{\sqrt{N}}. \]

Hence, we get

\[ P(S_{M_l} \geq N) \leq \sum_{i} P(M_l = i) P(S_i \geq N) \leq \sum_{i} P(M_l = i) \frac{i}{\sqrt{N}} = \frac{(l + 1)E[\hat{\xi}]}{\sqrt{N}}. \]

For \( t \in T^l \) and \( v \in t \), we simply write \((t, v)\) for the pointed tree \( t \) pointed at \( v \). Let \( E_{k,l} \) denote the set of pointed trees \((t, v)\) of a plane tree \( t \) having at most \( k \) vertices and a vertex \( v \) with height \( l \), such that \( P(H_l(\mathcal{X}, u_0) = (t, v)) > 0 \). For such \((t, v_0)\) \( \in E_{k,l} \), it is elementary to see that

\[ P[H_l(\mathcal{X}, u_0) = (t, v_0)] = \prod_{v \in t} \mu(d_v(v)). \]

Lemma 6. For any \((t, v)\) \( \in E_{k,l} \) \( n \geq |t| \) and \( m \geq a \) with \( a \) being the number of vertices in \( t \) that are before \( v \), we have

\[ \frac{P[H_l(T^n, V_m) = (t, v)]}{P[H_l(\mathcal{X}, u_0) = (t, v)]} = \frac{P(V_{m-a} \text{ in } T^{n-|t|+1} \text{ is a leaf})}{\mu(0)} \cdot \frac{P(|T| = n - |t| + 1)}{P(|T| = n)}. \]

Proof. Note that the out-degree of the \( i \)-th vertex \( V_i \) in \( T \) is uniquely determined by the difference of the L-walk values at \( i \) and \( i + 1 \). Hence \( H_l(T^n, V_m) = (t, v) \) if and only if the translated L-walk of \( T^n \) restricted on \([m-a, m+|t|-a]\) equals to the L-walk of \( t \), i.e.

\[ L_{T^n}(i + m - a) - L_{T^n}(m - a) = L_t(i), \quad i = 0, 1, \ldots, |t|. \]
As before, write \( L(i) = X_1 + \cdots + X_i \) for the random walk with jump distribution \( \mu_{-1} \). Then

\[
P[\forall i \leq n \exists j \in [0, m-a] \in [m-a, n] : L(j) - L(i) = x]
\]

By conditioning on the value of \( L(m-a) \), we can rewrite the numerator as follows.

\[
\sum_{x} P[L(i) \geq 0, i \in [0, m-a] ; L(m-a) = x] \times P[(\ast) ; L(j) \geq 0, j \in [m-a, n] ; L(n) = -1 | L(m-a) = x]
\]

where we use the Markov property twice and note that \( \ast \) is a factor.

We now turn to the other terms. By inserting an extra down-step between them (which corresponds to a factor \( \mu(0) \)), one can connect the L-walks on \([0, m-a] \) and \([|t| + m-a, n] \). From this, we get

\[
\sum_{x} P[L(i) \geq 0, i \in [0, m-a] ; L(m-a) = x] \times P[L(j) \geq 0, j \in [m+|t|-a, n] ; L(n) = -1 | L(m-a) = x-1]
\]

where we use the Markov property twice and note that \( \ast \) together with \( L(m-a) = x \) implies \( L(|t| + m-a) = x-1 \) (and hence \( x \geq 1 \)). Also by the Markov property, we obtain that

\[
P[(\ast) ; L(m-a) = x] = \prod_{v \in \mathcal{V}} \mu(0)(v) = P[\forall i \in [0, m-a] \in [m-a, n] : L(i) = x]
\]

We now turn to the other terms. By inserting an extra down-step between them (which corresponds to a factor \( \mu(0) \)), one can connect the L-walks on \([0, m-a] \) and \([|t| + m-a, n] \). From this, we get

\[
\sum_{x} P[L(i) \geq 0, i \in [0, m-a] ; L(m-a) = x] \times P[L(j) \geq 0, j \in [m+|t|-a, n] ; L(n) = -1 | L(m-a) = x-1]
\]

where we use the Markov property twice and note that \( \ast \) together with \( L(m-a) = x \) implies \( L(|t| + m-a) = x-1 \) (and hence \( x \geq 1 \)). Also by the Markov property, we obtain that

\[
P[(\ast) ; L(m-a) = x] = \prod_{v \in \mathcal{V}} \mu(0)(v) = P[\forall i \in [0, m-a] \in [m-a, n] : L(i) = x]
\]

We now turn to the other terms. By inserting an extra down-step between them (which corresponds to a factor \( \mu(0) \)), one can connect the L-walks on \([0, m-a] \) and \([|t| + m-a, n] \). From this, we get
From Lemma 4, Lemma 6 and Lemma 2, we have (when $|s|/k \to 0$)

$$\Xi^n_{l,m}(s) = \Xi_l(s)(1 + o(1)).$$

(3.6)

Hence,

$$d_{TV}(\Xi_l, \Xi^n_{l,m}) \leq \sum_{s \in A_l: |s| > \epsilon k} \Xi_l(s) + \sum_{s \in A_l: |s| \leq \epsilon k} |\Xi_l(s) - \Xi^n_{l,m}(s)|$$

$$\leq \frac{l + 1}{\sqrt{k}} + \sum_{s \in A_l: |s| \leq \epsilon k} \Xi_1(s) |o(1)|.$$

We can choose $\epsilon$ small enough to make the second term as small as we want and then let $n \to \infty$ to make the first term also small. This completes the proof of first assertion. When $\mu$ has finite exponential moments, by Lemma 4, Lemma 6 and (3.2), one can get (when $|s| \leq x/2$ and $x \leq k$)

$$\Xi^n_{l,m}(s) = \Xi_l(s)(1 + O(\frac{1}{k^{3/8}}) + O(\frac{x}{n})).$$

Case I: $((l + 1)n)^{2/3} \leq k$. Let $x = ((l + 1)n)^{2/3}$, we have

$$d_{TV}(\Xi_l, \Xi^n_{l,m}) \leq \sum_{s \in A_l: |s| > x/2} \Xi_l(s) + \sum_{s \in A_l: |s| \leq x/2} |\Xi_l(s) - \Xi^n_{l,m}(s)|$$

$$\leq \frac{l + 1}{\sqrt{x}} + O\left(\frac{1}{k^{3/8}}\right) + O\left(\frac{x}{n}\right) = O\left(\frac{(l + 1)^{2/3}}{n^{1/3}}\right) + O\left(\frac{1}{k^{3/8}}\right).$$

Case II: $((l + 1)n)^{2/3} \geq k$. we have

$$d_{TV}(\Xi_l, \Xi^n_{l,m}) \leq \sum_{s \in A_l: |s| > k/2} \Xi_l(s) + \sum_{s \in A_l: |s| \leq k/2} |\Xi_l(s) - \Xi^n_{l,m}(s)|$$

$$\leq \frac{l + 1}{\sqrt{k}} + O\left(\frac{1}{k^{3/8}}\right) + O\left(\frac{k}{n}\right).$$

3.3 Decomposing large random trees into small ones

In this subsection we consider GW-trees conditioned on size. From now on, we always assume that $\mu$ has exponential moments (in addition to the previous assumptions). Given a $t \in T_f$ with size $|t| = n$, an interval $[a, b] \subset [0, n]$ (with $a, b \in \mathbb{N}$) and a non-root vertex $v \in t$, we say $v$ is an M-vertex in $[a, b]$ (of $t$) and $[t]_v$ is an M-subtree in $[[a, b]]$ (of $t$), if

- all vertices of $[t]_v$ are in $\{V_a, \ldots, V_b\}$;
- $v$ is not in the unique path connecting $V_a$ and $V_b$;
- the parent of $v$ is in the unique path connecting $V_a$ and $V_b$.

We now state a lemma.
Lemma 7. For any \( \alpha \in (0, 1) \), there exist \( C, c_1, c_2 \) (depending on \( \alpha \)), such that: for any \( N = N(n) \in [\lceil n^\alpha, n \rceil] \) and \( m = m(n) \in [C \sqrt{\log(n + 1)}, c_1 \sqrt{N}] \), we have

\[
P[\exists a \in [[0, n - N(n) - 1]] : \{ k \in [[a, a + N(n)]] : V_k \text{ is an M-vertex in } \mathcal{L} \} \geq m(n) \sqrt{N(n)}] \leq \exp(-c_2 m^2(n));
\]

\[
P[\exists a \in [[0, n - N(n) - 1]] : d(V_{a}(T^n), V_{a+N(n)}(T^n)) \geq m(n) \sqrt{N(n)}] \leq \exp(-c_2 m^2(n)),
\]

where \( d(\bullet, \bullet) \) is the graph distance.

Proof. Fix any \( a \in [[0, n - N - 1]] \). Write \( I = [[a, a + N(n)]] \), \( v_1 = V_a \), \( v_2 = V_{a+N(n)} \) and \( v' \) for the latest common ancestor of \( v_1, v_2 \). We begin with the first assertion. If the number of M-vertices in \( I \) is larger than \( m(n) \sqrt{N(n)} \), then either the number of M-vertices attached to the (simple) path connecting \( v_1, v' \) is larger than \( m(n) \sqrt{N(n)/2} \), or the number of M-vertices attached to the path connecting \( v_2, v' \) is larger than \( m(n) \sqrt{N(n)/2} \).

For the first case, assume that all M-subtrees attached to the path connecting \( v_1, v' \) are \( \mathbf{t}_1, \ldots, \mathbf{t}_j \) and \( |\mathbf{t}_1| + \cdots + |\mathbf{t}_j| = J \). Then we have

\[ j \geq m(n) \sqrt{N(n)/2}, J < N(n). \]  

(3.7)

From the relation between a tree and its L-walk, one can easily see that the translated L-walk of \( T^n \) restricted on \( [[a + 1, a + J + 1]] \), is just the concatenation of L-walks of \( \mathbf{t}_1, \ldots, \mathbf{t}_j \). Precisely,

\[ L_{T^n}(a + 1 + k) - L_{T^n}(a + 1) = L_{\mathbf{t}_i}(k - |\mathbf{t}_1| - \cdots - |\mathbf{t}_{i-1}|) - (i - 1), \]

when \( |\mathbf{t}_1| + \cdots + |\mathbf{t}_{i-1}| \leq k \leq |\mathbf{t}_1| + \cdots + |\mathbf{t}_i| \). Especially, we have

\[ \min_{k \in [[a + 1, a + N]]} L_{T^n}(k) - L_{T^n}(a + 1) \leq L_{T^n}(a + J + 1) - L_{T^n}(a + 1) = - j. \]  

(3.8)

The L-walk of \( T^n \) is a conditioned random walk conditioned on \( H_{-1} = n \). we know that

\[ P(H_{-1} = n) \sim cn^{-3/2}. \]  

(3.9)

On the other hand, by Lemma 2

\[ P[ \min_{k \in [[0, N]]} L(k) \leq - j] \leq N \exp(-c_2 j^2/N), \]

where as before, \( L \) is the (unconditioned) random walk starting at 0. By this and (3.7), we get

\[ P(3.8) \leq N \exp(-c m^2(n)). \]

Noting that (by choosing \( C \) large enough) \( \exp(-c m^2(n)) \) decays faster than any polynomials of \( n \), we get the desired bound for this case.
For the case when the number of M-vertices attached to the path connecting \(v_2, v'\) is larger than \(m(n)\sqrt{N(n)}/2\), we can reduce it to the first case by considering the L-walk from the right side. For a \(t \in \mathbb{T}_t\), its L-walk from the right side can be obtained as follows. First, for every vertex, reverse the order of its children. Then the L-walk of this tree (with the reversed order) is just the L-walk from the right side of \(t\). We leave the details of the reduction to the reader.

We now turn to the second assertion. When \(d(v_1, v_2) \geq m(n)\sqrt{N(n)}\), either \(d(v_2, v') \geq m(n)\sqrt{N(n)}/2\) or \(d(v_1, v') \geq m(n)\sqrt{N(n)}/2\). As in the proof of the first assertion, (the latter case can be reduced to the former one) we can assume that

\[
d(v_2, v') \geq m(n)\sqrt{N(n)}/2.
\]

Consider the translated L-walk of \(T^n\) in \(I, L' : [[0, N(n)]] \to \mathbb{Z} : L'(i) = L_{T^n}(a + i) - L_{T^n}(a).
\)

Note that \(L'\) can be regarded as the concatenation of L-walks of GW-trees (though the traverse of the last GW-tree is generally not ‘finished’) and (3.10) corresponds to the event that the height of the \(N(n)\)-th vertex in these GW-trees is at least \(m(n)\sqrt{N(n)}/2\). Since the L-walk of \(T^n\) is a conditioned random walk conditioned on the event \(H_{-1} = n\), which has probability of order \(n^{-3/2}\), it suffices to show that for an unconditioned random walk \(L\) starting at 0 with distribution \(\mu_{-1}\), (which can be regarded as the concatenation of L-walks of independent GW-trees), the probability of the event that the height of the \(N(n)\)-th vertex in the corresponding GW-tree is bigger than \(m(n)\sqrt{N(n)}/2\) has the desired bound.

At the current stage we need a relation between the height of a vertex of a forest and its L-walk. The L-walk of a forest is just the concatenation of all L-walks of the trees in the forest. A key fact is that the height of the \(N\)-th vertex has the same distribution as the number of ‘record times’ of \(L\) in \([[1, N]]\). More precisely, set \(R_0 = 0\) and \(R_j = \inf\{k > R_{j-1} : L(k) = \max_{i \in [[0,k]]} L(i)\}\). Then, the height of the \(N\)-th vertex in the forest of independent GW-tress has the same distribution as \([\{j \in \mathbb{N}^+ : R_j \leq N\}]\). Moreover, the random variables \(L(R_j) - L(R_{j-1}), j = 1, 2, \ldots\) are i.i.d. with the distribution given by: \(P(L(R_i) = i) = \mu_{-1}([i, \infty])\). Especially, \(E[L(R_1)] = \sigma^2/2\). For more details about this, see Section 1.3 in [13]. With the help of these facts, we now get (write \(k = [m(n)\sqrt{N(n)}/2]\))

\[
P[\{j \in \mathbb{N}^+ : R_j \leq N\} \geq k] = P[R_k \leq N]
\]

\[
\leq P\left[\sum_{j=1}^k (L(R_j) - L(R_{j-1})) \leq \frac{k\sigma^2}{4}\right] + P\left[\sum_{j=1}^k (L(R_j) - L(R_{j-1})) \geq \frac{k\sigma^2}{4}, R_k \leq N\right]
\]

\[
\leq P\left[\sum_{j=1}^k (L(R_j) - L(R_{j-1})) \leq \frac{k\sigma^2}{4}\right] + P\left[\max_{i \leq N} L(i) \geq \frac{k\sigma^2}{4}\right].
\]
Now the last assertion in Lemma 2 gives the the desired bounds for both terms (note that \( \mu_{-1} \) and \( L(R_1) \) have finite exponential moments and that \( E[L(R_j) - L(R_{j-1})] = \sigma^2/2 \).

**Lemma 8.** For any \( k, j \in \mathbb{N}^+ \), let \( \xi_1, \ldots, \xi_j \) be i.i.d. with the distribution of \(|T|\), and \( \tilde{\xi}_i = \xi_i 1_{\xi_i \leq k} \). There exists \( C > 0 \), such that, when \( t \geq Cj \sqrt{k} \),

\[
P[\tilde{\xi}_1 + \cdots + \tilde{\xi}_j > t] \leq \exp\left(-\frac{t}{k}\right).
\]

**Proof.** We know that \( P(\xi_i = n) \propto n^{-3/2} \). From this we get \( E\tilde{\xi}_i \propto \sqrt{k} \) and \( E\tilde{\xi}_i^2 \propto k^{3/2} \). A simple application of the Bernstein inequality gives the desired bound.

**Proposition 11** For any \( \alpha \in (0,1) \), there exist \( C_1, C_2, c_1, c_2 \) (depending on \( \alpha \)), such that, for any \( n \) sufficiently large, \( N(n) \in [n^\alpha, n] \), \( b(n) \in [C_1, \sqrt{n}] \) and \( a(n) \in [(C_2 b(n))^2 \log n, c_1 b(n)^2 N(n)] \), we have the following. For \( I \), any interval in \([0, n]\) with length \( N(n) \), let \( t_1, \ldots, t_j \) be all \( M \)-subtrees in \( I \) of \( T^n \). Then

\[
P\left[ \sum_{|t_i| \leq N(n)/a(n)} |t_i| > \frac{N(n)}{b(n)} \right] \leq \exp\left(-c_2 \frac{a(n)}{(b(n))^2} \right).
\]

In other words, if we discard all small \( M \)-subtrees (with size at most \( N(n)/a(n) \)) in \( I \), with high probability, we would not lose too many (bigger than \( N(n)/b(n) \)) vertices.

**Proof.** As in the proof of Lemma 7, write \( v_1, v_2 \) for the first and the last vertices of \( T \) in \( I \) and \( v' \) for the latest common ancestor of \( v_1, v_2 \). Assume \( t_1, \ldots, t_j \) are all \( M \)-subtrees in \( I \). Let \( l \) be the graph distance between \( v_2 \) and \( v' \). Note that the set of all vertices in \( \{V_i : i \in I\} \) except \( v_1 \) is the disjoint union of the set of vertices in \( t_1, \ldots, t_j \) and the set of vertices in the unique path connecting \( v_2 \) and \( v' \) (including \( v_2 \) but not \( v' \)). Due to Lemma 7 by discarding an event with small probability, we can assume that

\[
j \leq c_0 \frac{\sqrt{N} a^{1/2}}{b}, \quad l \leq c_0 \frac{\sqrt{N} a^{1/2}}{b},
\]

where we will pick up \( c_0 \) very small and then let \( C_1 \) be large enough and \( c_1 \) small enough (such that we can apply Lemma 7, Lemma 8 and Lemma 2).

Note that we have

\[
\sum_{i=1}^j |t_i| = N(n) - l.
\]

Moreover, it is elementary to see that conditionally on the value of \( j \) and \( l \), the sizes \((|t_1|, \ldots, |t_j|)\) have the same distribution of \((\xi_1, \ldots, \xi_j)\) in the last lemma conditioned on their sum being \( N - l \). Hence we only need to estimate (under the same notations in the last lemma)

\[
P[\tilde{\xi}_1 + \cdots + \tilde{\xi}_j > t | \xi_1 + \cdots + \xi_j = N - l].
\]
with \( k = N/a \) and \( t = N/b \). By choosing \( c_0 \) small enough, \( t, k, j \) can be satisfied with the requirement in Lemma 8. Hence, we have

\[
P[\tilde{\xi}_1 + \cdots + \tilde{\xi}_j > \frac{N}{t}] \leq \exp(-\frac{a}{b}).
\]

On the other hand, we have (from (3.1) and the third assertion of Lemma 2)

\[
P[\xi_1 + \cdots + \xi_j = N - l] \geq \frac{1}{N^{3/2}} \exp(-\frac{a}{2b}),
\]

by choosing \( C_1 \) and \( C_2 \) large enough. Combining the two inequalities above, we get

\[
P[\tilde{\xi}_1 + \cdots + \tilde{\xi}_j > \frac{N}{t} | \xi_1 + \cdots + \xi_j = N - l] \leq \frac{N^{3/2}}{2b} \exp(-\frac{a}{2b}).
\]

By choosing \( C_1 \) large enough, we have

\[
\frac{N^{3/2}}{2b} \exp(-\frac{a}{2b}) \leq \exp(-\frac{a}{b^2}).
\]

The proof is complete.

Remark 6. In this proposition, conditionally on the value of \((j; |t_1|, \ldots, |t_j|)\), \(t_1, \ldots, t_j\) are actually independent and distributed as the GW-trees with their sizes. One can easily see this from the following. When conditioning on the values of the L-walk in \( I \) outside the intervals corresponding to \( t_1, \ldots, t_j \), the L-walk restricted on the interval corresponding to \( t_i \) is just a conditioned random walk conditioned (only) on fixed starting point and end point and achieving its minimum only at the end point.

**Proposition 12** For any positive \( A > 2, \epsilon, \alpha \) with \( \alpha > 2 \epsilon \), there exists \( C = C(\epsilon, \alpha, A) \), satisfying the following. When \( n \in \mathbb{N} \) is sufficiently large, for any \( k = k(n), L = L(n) \in [[1, n]] \) satisfying \( k^A \geq n \geq L \log L, L \geq k^{2+\alpha}, \) with probability bigger than \( 1 - Ck^{1+\epsilon}/\sqrt{L} \), we can find a vertex \( v \) in the \( k \)-th generation of \( T^n \), such that

- the subtree \([T^n]_v\) has at least \( n - L \) vertices;
- conditionally on the size (even on the subtree generated by all edges not in \([T^n]_v\)), \([T^n]_v\) is distributed as a GW-tree with the given size.

**Proof.** For any tree \( t \), write \( w_i(t) \) for the number of vertices in the \( i \)-th generation and \([t]_i\) for the subtree generated by all vertices of height \( \leq i \). In order to prove this proposition, we need two lemmas.

**Lemma 9.** For any positive \( \beta, \gamma \), there exists \( C = C(\beta, \gamma) > 0 \), such that

\[
P[w_k(T^n) \geq k^{1+\beta}] \leq \frac{C}{k^\gamma},
\]

and

\[
P[\sum_{i=1}^k w_i(T^n) \geq k^{2+\beta}] \leq \frac{C}{k^\gamma}.
\]
Proof. Lemma 2.2 in [10] states that 

\[ E(w_k(T^n))^r \leq C(r)k^{r-1}. \]

From this, one can get

\[ P[w_k(T^n) \geq x] \leq \frac{C(r)k^{r-1}}{x^r}. \]

(3.12)

Hence, we have

\[ P[w_k(T^n) \geq k^{1+\beta}] \leq \frac{C(r)k^{r-1}}{(k^{1+\beta})^r} = \frac{C(r)}{k^{1+\beta^r}}. \]

By choosing \( r \) large enough, we get the first assertion.

For the second one,

\[ P[\sum_{i=0}^k w_i(T^n) \geq k^{2+\beta}] \leq \sum_{i=1}^{k} P[w_i(T^n) \geq k^{1+\beta}/2] \]

\[ \leq \sum_{i=1}^{k} C(r)k^{i-1}(k^{1+\beta}/2)^r \leq kC(r)\frac{2^rk^{r-1}}{k(1+\beta)^r} = \frac{2^rC(r)}{k^{1+\beta^r}}. \]

By choosing \( r \) large enough, we get the second assertion.

As before, let \( \xi_1, \xi_2, \ldots, \xi_m \) be i.i.d. copies of \(|T|\).

Lemma 10. For any \( N \geq l \geq m \geq 2 \) with \( N \geq 2l \geq m^2 \),

\[ P[\max_{1 \leq i \leq m} \{\xi_i\} \leq N - l|\sum_{i=1}^m \xi_i = N] \leq \frac{m}{\sqrt{l}} + \frac{m}{\sqrt{N/\log N}}. \]

(3.13)

Proof. We divide the term on the left hand side into three pieces and estimate them separately.

\[ \sum_{j=l}^{\lfloor 0.9N \rfloor} P[\max_i \{\xi_i\} = N - j|\sum_i \xi_i = N] + \sum_{j=\lfloor N/5 \log N \rfloor}^{\lfloor 0.1N \rfloor} P[\max_i \{\xi_i\} = j|\sum_i \xi_i = N] \]

\[ + P[\max_i \{\xi_i\} \leq N/(5 \log N)|\sum_i \xi_i = N]. \]

The first one is bounded by

\[ \sum_{j=l}^{\lfloor 0.9N \rfloor} \frac{P[\max_i \{\xi_i\} = N - j, \sum_i \xi_i = N]}{P[\sum_i \xi_i = N]} \leq \sum_{j=l}^{\lfloor 0.9N \rfloor} \frac{mP[\xi_1 = N - j, \sum_{i=2}^m \xi_i = j]}{m \sum_{j=l}^{\lfloor 0.9N \rfloor} P[L(N) = -m]} \]

\[ \leq \sum_{j=l}^{\lfloor 0.9N \rfloor} \frac{mN^{-3/2}(m-1)j^{-3/2}}{mN^{-3/2}} \leq \frac{m}{\sqrt{l}}. \]
The second one can be analyzed similarly.

\[
\sum_{j=\lceil N/(5 \log N) \rceil}^{\lfloor 0.1N \rfloor} \frac{P[\max_i \xi_i = j; \sum_i \xi_i = N-j]}{P[\sum_i \xi_i = N]} \lesssim \sum_{j=\lceil N/(5 \log N) \rceil}^{\lfloor 0.1N \rfloor} \frac{mP[\xi_1 = j; \sum_2^m \xi_i = N-j]}{P[\sum_i \xi_i = N]} \\
\lesssim \sum_{j=\lceil N/(5 \log N) \rceil}^{\lfloor 0.1N \rfloor} \frac{m \gamma^{3/2}(m-1)N^{-3/2}}{mN^{-3/2}} \asymp \frac{m}{\sqrt{N/\log N}}.
\]

The last term is bounded by

\[
P[\max_i \xi_i \leq N/(5 \log N), \sum_i \xi_i = N] \leq \frac{P[\sum_i \xi_i \leq N/(5 \log N)]}{mN^{-3/2}}.
\]

Applying Lemma 8 with \( t = N, j = m, k = \lceil N/(5 \log N) \rceil \), we obtain that the numerator is less than \( N^{-5/2} \) and finish the proof.

We are ready to prove our proposition. Let \( \beta = \frac{1}{2}(\epsilon \wedge \alpha) \). By Lemma 9, we can assume that (by discarding an event with small probability)

\[
M \doteq |[T^n]_k| \leq k^{2+\beta}, \quad m = w_k([T^n]_k) \leq k^{1+\beta}.
\]

Write \( v_1, \ldots, v_m \) for all vertices in the \( k \)-generation and \( t_i \) for \( [T^n]_{v_i} \) (\( i = 1, \ldots, m \)). As before, it is elementary to see that conditionally on \( M, m \) (even on \( [T^n]_k \)), \( |t_1|, \ldots, |t_m| \) are distributed as \( \xi_1, \ldots, \xi_m \) conditioned on \( \sum_{i=1}^m \xi_i = n-M+m \). Let \( N = n-M+m \) and \( l = L/2 \). Applying Lemma 10 we have (note that \( M, m \ll L \ll n \), hence \( n-M+m-L/2 \geq n-L \))

\[
P[\max_i |t_i| \geq n-L] \geq 1 - C \frac{k^{1+\beta}}{\sqrt{L}}.
\]

Let \( [T^n]_o \) be the one that achieves the maximum. The first requirement holds and the second is true due to the same reason as in Remark 6. We finish the proof.

We now state our result for decomposing a large random tree into small ones.

**Theorem 13.** For any \( d_0, e > 0, \alpha, \beta, \delta \in (0, d_0) \) with \( \epsilon_0 = \frac{3}{2} \beta - \alpha - d_0 - \frac{3}{2} \delta > 0 \), \( e \in (0, \epsilon_0) \), and \( n(N) \), a function of \( N \) with \( n(N) \in [N^{d_0}/(\log N)^{e}, N^{d_0}(\log N)^{e}] \), there exists \( C \) (depending on all parameters except \( N \)), such that, for any sufficiently large \( N \in \mathbb{N} \), with probability at least \( 1 - C N^{-e} \), in \( T^n \), we can find rooted subtrees \( T_1, \ldots, T_m \) (note that \( m \) is also random), where \( T_i \) is rooted at \( v_i \), the unique vertex in \( T_i \) closest to the root of \( T^n \), \( o \), satisfying the following:

1. For every \( i \in \{1, \ldots, m\} \), \( N^{\beta - 2\delta}/(\log N)^{1.01} \leq |T_i| \leq N^{\beta} \) and the distance between \( v_i \) and \( o \) is at least \( |N^\alpha| \);
2. Let $\hat{T}$ be the graph generated by all edges not in any $T_i$. Then $\hat{T}$ is a tree and $|\hat{T}| \leq n/N^\delta$.
3. Let $v_i$ ($i = 1, \ldots, m$) be the unique path starting from $v_i$ towards the root of $T_i$, with length $|N^\alpha|$. Then for any $i \in \{1, \ldots, m\}$, all $T_j$ except $T_i$, are in the same component of $T \setminus v_i$.
4. Conditionally on $\{m; |T_1|, \ldots, |T_m|; \hat{T}\}$ (and even the locations of $v_i$ in $\hat{T}$), the trees $T_i$ are independent and distributed as the GW-trees with their sizes.

**Proof.** This theorem is a combination of Proposition 11 and Proposition 12. We only need to designate the parameters.

Note that since $\epsilon_0 > 0$, we have $0 < d_0 - \beta + 2\delta < \beta/2 - \alpha - 3\delta/2$ and $(\beta/2 - \alpha - 3\delta/2) - (d_0 - \beta + 2\delta) = \epsilon_0$. Let $\gamma = d_0 - \beta + 2\delta + (\epsilon + \epsilon_0)/2$. Note that $\beta/2 - \alpha - 3\delta/2 - \gamma = (\epsilon_0 - \epsilon)/2$. Set $a = N^2(\log N)^{1.01}/3, b = 2N^\delta, k = |N^\alpha|$ and $L = N^{2\alpha+2\gamma}/2$. First, divide $[0, n]$ (deterministically) into disjoint intervals with lengths in $[[N^3/2, N^3]]$. Then, for each interval, apply Proposition 11. Then, (with high enough probability, since the error term decays faster than any polynomial of $N$) without losing too many vertices ($< n/2N^3$), we obtain a number of $M$-subtrees, say $T'_1, \ldots, T'_m$, with sizes in $[[3N^\beta-2\delta/2(\log N)^{1.01}, N^3]]$. Then, for each $T'_i$, by applying Proposition 12, we can obtain, with high probability ($1 - Ck^{1.01}+0.1(\epsilon_0-\epsilon)/d_0)/\sqrt{L} \geq 1 - CN^0.1(\epsilon_0-\epsilon)/N^\gamma$), a subtree $T_i$, which is at distance $k$ from the root of $T'_i$, without losing more than $L$ vertices. We argue that $T_1, \ldots, T_m$ satisfy all conditions. Conditions (3), (4), and part of (1) and (2) are satisfied automatically from the propositions. We only need to check the quantitative parts of Conditions (1) and (2).

First, since $2\alpha + 2\gamma = (\beta - 2\delta) - \delta - (\epsilon_0 - \epsilon)$, one can see that the size of $T_i$ is at least

$$\frac{3N^\beta-2\delta}{2(\log N)^{1.01}} - L \geq \frac{3N^\beta-2\delta}{2(\log N)^{1.01}} - \frac{N^{\beta-2\delta}}{2(\log N)^{1.01}} = \frac{N^{\beta-2\delta}}{(\log N)^{1.01}}.$$

Hence, Condition (1) is satisfied.

Second, consider the total number of vertices that we lose. In the first step (when applying Proposition 11), we lose at most $n/2N^3$ vertices. On the other hand, note that (when $N$ is large enough)

$$m < \frac{n}{3N^\beta-2\delta/2(\log N)^{2.001}} \leq \frac{n}{N^\beta-2\delta+0.1(\epsilon_0-\epsilon)}.$$

Therefore, in the second step (when applying Proposition 12), the number of vertices we lose is at most

$$mL \leq \frac{n}{2N^\beta-2\delta-2\alpha-2\gamma-0.1(\epsilon_0-\epsilon)} = \frac{n}{2N^\delta+2(\beta/2-\alpha-3\delta/2-\gamma)-0.1(\epsilon_0-\epsilon)} = \frac{n}{2N^\delta+0.9(\epsilon_0-\epsilon)} < \frac{n}{2N^\delta}.$$

Combining both, one can get Condition (2).

Finally, note that

$$m \frac{N^{0.1(\epsilon_0-\epsilon)}}{N^\gamma} \leq \frac{N^{d_0}}{N^{\gamma+(\beta-2\delta)-0.3(\epsilon_0-\epsilon)}} = \frac{1}{N^{0.2\epsilon_0+0.8\epsilon}} \leq \frac{1}{N^\epsilon}.$$
It means that with probability at least $1 - CN^{-\epsilon}$, we can get all $T_i$ from $T'_i$ as desired. The proof is complete.

4 The local limit of tree-indexed random walks in tori

In this and the next sections, we consider tree-indexed random walks in the discrete torus $T_N = (\mathbb{Z}/N\mathbb{Z})^d$ of side-length $N$ ($d \geq 5$). For technical reasons, we assume that the jump distribution $\theta$ is aperiodic, symmetric, with finite range and not supported on any strict subgroup of $\mathbb{Z}^d$, and that the offspring distribution $\mu$ is critical, nondegenerate, with finite exponential moments. For simplicity, we assume further that $\mu$ is with span 1. The case when the span is bigger than 1 can be treated with minor standard modifications. We aim to show the local convergence result in this section. Write $\varphi : \mathbb{Z}^d \to T_N$ for the canonical projection map induced by mod $N$. We now state our main theorem.

**Theorem 14.** For any $K \subset \subset \mathbb{Z}^d$, $u \in (0, \infty)$ and $n = n(N)$, an integer-valued function satisfying $\lim_{N \to \infty} n(N)/N^d = u$, 
\[
\lim_{N \to \infty} P[\text{Range}(S_{n,N}) \cap \varphi(K) = \emptyset] = \exp(-u B\text{Cap}(K)),
\]
where $S_{n,N}$ is the tree-indexed random walk with size $n$ and a uniform starting point in $T_N$.

Through the inclusion-exclusion principle, this theorem implies the local convergence of the configuration, Theorem 10.

4.1 The visiting probability of a set by a small tree-indexed random walk

Theorem 14 gives an asymptotic formula for the probability that a tree-indexed random walk visits a subset in $T_N$, with a size proportional to the volume of the torus, $N^d$. The main result of this subsection is to give an asymptotic formula for the probability that a set is visited by a much smaller tree-indexed random walk.

**Theorem 15.** For any $\alpha_1 < \alpha_2 \in (0, d)$ fixed, $n = n(N)$ any integer-valued function of $N$ satisfying $n(N) \in [N^{\alpha_1}, N^{\alpha_2}]$, and $K \subset \subset \mathbb{Z}^d$, we have
\[
\lim_{N \to \infty} \frac{N^d}{n} P[\text{Range}(S_{n,N}) \cap \varphi(K) = \emptyset] = B\text{Cap}(K).
\]

We start with our notations. For any $m \in \mathbb{N}^+$ and $t \in T_I \cup T_\infty$ with $|t| > m$, denote by $\rho_m(t)$ and $\rho_{m}^+(t)$, respectively, the subtree generated by the first $m$ vertices of $t$ (i.e. $V_0, \ldots, V_{m-1}$) and the one generated by the first $m$ vertices together with all children of the first $m$ vertices. Note that $\rho_m^+(t)$ is uniquely determined by $\rho_m(t)$ and the degrees of the vertices that are ancestors of $V_{m-1}$. In addition, since the parent of a vertex comes before that vertex, we can see
that $\rho_m^+(t)$ is uniquely determined by $(L_k(i))_{i \in [0,m]}$. Note that in $\rho_m^+(t)$, there are exactly $L_k(m) + 1$ vertices, which are not in $\rho_m(t)$. We denote by $\partial \rho_m(t)$ the set of these vertices.

As before, write $T^n$ and $T^\infty$ for the GW-tree with $n$ vertices and the one conditioned on survival respectively. We now compare $P[\rho_m^+(T^n) = \rho_m^+(t)]$ to $P[\rho_m^+(T^\infty) = \rho_m^+(t)]$. By considering the L-walks, it is elementary to see that

$$P[\rho_m^+(T^n) = \rho_m^+(t)] = \frac{P[L(i) = L_k(i), i = 0, 1, \ldots, m; H - 1 = n]}{P[|T| = n]},$$

where as before, $L = (L(i))_{i \in \mathbb{N}}$ is a random walk starting at 0 with jump distribution $\mu$, $H - 1$ is the fist visiting time of $-1$ and $T$ is an unconditioned GW-tree. Using the Markov property at time $m$ and (3.1), one can get

$$P[\rho_m^+(T^n) = \rho_m^+(t)] = \frac{P[L(i) = L_k(i), i = 0, 1, \ldots, m]L_k(m + 1)P[L(n - m) = -(L_k(m) + 1)]}{P[|T| = n]}.$$  \hfill (4.3)

We turn to $T^\infty$. Note that when $\rho_m^+(T^\infty) = \rho_m^+(t)$, there is a unique vertex in $\partial \rho_m(t)$, which is also in the spine of $T^\infty$. On the other hand, for any $u \in \partial \rho_m(t)$, we have

$$P[\rho_m^+(T^\infty) = \rho_m^+(t), u \text{ is a vertex in the spine of } T^\infty] = P[L(i) = L_k(i), i = 0, 1, \ldots, m].$$

Hence, we get

$$P[\rho_m^+(T^\infty) = \rho_m^+(t)] = P[L(i) = L_k(i), i = 0, 1, \ldots, m](L_k(m) + 1).$$  \hfill (4.4)

Combining (4.3) and (4.4) leads to

$$\frac{P[\rho_m^+(T^n) = \rho_m^+(t)]}{P[\rho_m^+(T^\infty) = \rho_m^+(t)]} = \frac{P[L(n - m) = -(L_k(m) + 1)]}{(n - m)P[|T| = n]}.$$  \hfill (4.5)

When $n \to \infty$ and $m/n \to a (0, 1)$, by the first assertion of Lemma 2, we get

$$\frac{P[\rho_m^+(T^n) = \rho_m^+(t)]}{P[\rho_m^+(T^\infty) = \rho_m^+(t)]} = \Gamma_a(L_k(m) + 1 \sqrt{n}) + o(1),$$

where $\Gamma_a(x) = \exp(-x^2/2\alpha^2(1-a))/((1-a)^{3/2})$. Note that $\Gamma_a$ is bounded above, when $a \in (0, 1)$ is fixed. Hence, we have, when $m = [an]$,

$$P[\rho_m^+(T^n) = \rho_m^+(t)] \leq C(a)P[\rho_m^+(T^\infty) = \rho_m^+(t)].$$  \hfill (4.5)

We now give a similar relation between $T^\infty$ and $T^c$, the random tree with law $\Pi_c$. Note that $T^\infty$ and $T^c$ differ only in the finite bushes attached to the
roots. Write \( v \) for the vertex in the spine next to the root. Then, for any \( i \in \mathbb{N} \),
\[
P[v \text{ has exact } i \text{ elder siblings in } T^\infty] = \sum_{j=i+1}^{\infty} \mu(j)
\leq \sum_{j=i}^{\infty} \mu(j) = 2P[v \text{ has exact } i \text{ elder siblings in } T^c].
\]
Hence, we have
\[
P[\rho_m(T^\infty) = \rho_m(t)] \leq 2P[\rho_m(T^c) = \rho_m(t)].
\]
Combining this with (4.5), we get
\[
\text{Lemma 11.}
\]
\[
P[\rho^+_{m}(T^n) = \rho^+_{m}(t)] \leq C(n)P[\rho^+_{m}(T^\infty) = \rho^+_{m}(t)],
\]
\[
P[\rho_{m}(T^\infty) = \rho_{m}(t)] \leq P[\rho_{m}(T^c) = \rho_{m}(t)].
\]

Before the formal proof, let us set up our notations. For any (finite or infinite) spatial tree \((t, S)\), write \(X_{i} = X_{i}(t, S) = S(V_{i}(t))\). We let \( P_{n}^{m} \) be the law when \((t, S)\) is as follows. First \( t \) is a GW-tree with size \( n \) and conditionally on \( t \), \( S \) is a \( t \)-indexed random walk sending \( V_{k} \) to \( x \), i.e. the law of \( S : t \to \mathbb{Z}^d \) is determined by: \( V_{k} \) is mapped to \( x \) and the random variables \((S(v_1) - S(v_2))_{(v_1, v_2)}\) are independent and distributed according to \( \theta \), where \((v_1, v_2)\) are over all (undirected) edges of \( t \). Similarly, one can define \( P_{n}^{\infty} \) and \( P_{k,x}^{\infty} \). We write \( P_{N}^{m} \) for the law when \((t, S)\) is the tree-indexed random walk with size \( n \), in \( T_{N} \), with a uniform starting point.

\textbf{Proof (Proof of Theorem 15).} Write \( K' = \varphi(K) \). We have (when \( N \) is large enough)
\[
P_{n}^{m,N}[\{X_{0}, \ldots, X_{n-1}\} \cap K' \neq \emptyset] \geq \frac{N^d}{n}
\leq \sum_{k=1}^{n-1} P_{n}^{m,N}[X_{0}, \ldots, X_{k-1} \notin K', X_{k} = x] \frac{N^d}{n} + \sum_{x \in K'} P_{n}^{m,N}[X_{0} = x] \frac{N^d}{n}
\leq \sum_{x \in K'} \left( \frac{1}{n} \sum_{k=1}^{n-1} N^d P_{n}^{m,N}[X_{0}, \ldots, X_{k-1} \notin K', X_{k} = x] \right) + \frac{|K'|}{n}.
\]

It suffices to show that for any \( x \in K \) and \( \epsilon \in (0, 0.1) \),
\[
\lim_{N \to \infty} \frac{1}{n} \sum_{x \in K, \|x\| < (1-\epsilon)n} |N^d P_{n}^{m,N}[X_{0}, \ldots, X_{k-1} \notin K', X_{k} = \varphi(x)] - E_{S_{K}}(x)| = 0.
\]

The equality above follows if we can show the following three lemmas.
Lemma 12.

\[
\lim_{N \to \infty} \frac{1}{n} \sum_{\epsilon n < k < (1-\epsilon)n} |N^d P_n [X_0, \ldots, X_{k-1} \notin K', X_k = \varphi(x)] - P_{k,x}^n [X_0, X_1, \ldots, X_{k-1} \notin K]| = 0. \tag{4.8}
\]

Lemma 13. For any \(l = l(n) \in [1, cn]\) with \(l(n) \to \infty\),

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{\epsilon n < k < (1-\epsilon)n} |P_{n,x}^k [X_0, X_1, \ldots, X_{k-1} \notin K] - P_{k,x}^n [X_{k-l}, X_{k-l+1}, \ldots, X_{k-1} \notin K]| = 0. \tag{4.9}
\]

Lemma 14. For any \(l = l(n) \in [1, cn]\) with \(1 \ll l(n) \ll \sqrt{n}\),

\[
\lim_{n \to \infty} \sup_{\epsilon n < k < (1-\epsilon)n} |P_{n,x}^k [X_{k-l}, X_{k-l+1}, \ldots, X_{k-1} \notin K] - \mathbb{E}S_K(x)| = 0. \tag{4.10}
\]

We need the following two lemmas in the proof of Lemma 12.

Lemma 15. There exist \(C, c_1, c_2\), such that for any \(n, l \in \mathbb{N}^+\) with \(l \in [C_1 n^{1/4} (\log n)^{3/4}, c_1 \sqrt{n}]\),

\[
\sup_{0 \leq i \leq n} P_{0,x}^n [\left| X_i - x \right| > l] \leq \exp(-c_2 \frac{l^{4/3}}{n^{1/3}}). \tag{4.11}
\]

Lemma 16. For any \(x \neq y \in \mathbb{Z}^d\) and \(n \in \mathbb{N}^+\),

\[
\frac{1}{n} \sum_{0 \leq i < j < n} P_{j,x}^n (X_i = y) \leq |x - y|^{4-d} \tag{4.12}
\]

**Proof (Proof of Lemma 15).** Let \(h\) be the height of \(V_i\) in \(T^n\). By Lemma 7, we have

\[
P[h \geq l^{2/3} n^{1/3}] \leq \exp(-c \frac{l^{4/3}}{n^{1/3}}).
\]

On the other hand, by standard estimates for random walks (e.g. the last assertion in Lemma 2) we have

\[
P_{0,x}^n [\left| X_i - x \right| > l] \leq \exp(-c \frac{l^{4/3}}{n^{1/3}}).
\]

Combining both displays finishes the proof.

**Proof (Proof of Lemma 16).** By conditioning on the distance between \(V_i\) and \(V_j\), one can get

\[
\frac{1}{n} \sum_{0 \leq i < j < n} P_{j,x}^n (X_i = y) = \frac{1}{n} \sum_{0 \leq i < j < n} \sum_{k \in \mathbb{N}^+} P_{j,x}^n [d(V_i, V_j) = k] P_{j,x}^n (X_i = y | d(V_i, V_j) = k) = \frac{1}{n} \sum_{k \in \mathbb{N}^+} P_{x}^{RW} (Z(k) = y) \sum_{0 \leq i < j < n} P_{j,x}^n [d(V_i, V_j) = k],
\]
where $P_{x}^{\text{RW}}$ indicates that $Z = (Z(i))$ is a random walk starting at $x$ (with jump distribution $\theta$).

At the current stage, we need the following result, Theorem 1.1 in [8], which will also be useful later.

**Lemma 17.** There exists a constant $C$, such that for all $n, k \in \mathbb{N}^+$, the expected number of pairs of vertices with distance $k$, in $T^n$, is at most $Ckn$.

Hence

$$\frac{1}{n} \sum_{0 \leq i < j < n} P_{j,x}^n[d(V_i, V_j) = k] \preceq k.$$ 

Therefore,

$$\frac{1}{n} \sum_{0 \leq i < j < n} P_{j,x}^n(X_i = y) \preceq \sum_{k \in \mathbb{N}^+} kp_{x}^{\text{RW}}(Z(k) = y) \asymp |x-y|^{4-d},$$

where the last step is standard.

**Proof (Proof of Lemma 12).** Let $\varphi^{-1}(K') = \bigcup_{i=0}^{\infty} K_i$, such that $x \in K_0 = K$ and $K_i$ is a translated copy of $K_0$. Fix some $\lambda \in \left(\frac{1}{4}, \frac{d}{4\alpha^2}\right)$ and let $b = \lfloor \frac{n\lambda}{b} \rfloor + 1$.

$$\frac{1}{n} \sum_{cn < k < (1-\epsilon)n} |N^d P^{n, N}[X_0, \ldots, X_{k-1} \notin K; X_k = \varphi(x)] - P_{k,x}^n[X_0, \ldots, X_{k-1} \notin K]|$$

$$= \frac{1}{n} \sum_{cn < k < (1-\epsilon)n} P_{k,x}^n[X_0, X_1, \ldots, X_{k-1} \notin K] - P_{k,x}^n[X_0, X_1, \ldots, X_{k-1} \notin \varphi^{-1}(K)]$$

$$\leq \frac{1}{n} \sum_{cn < k < (1-\epsilon)n} P_{k,x}^n[\{X_0, \ldots, X_{k-1}\} \cap (\bigcup_{i \geq 1} K_i) \neq \emptyset]$$

$$\leq \frac{1}{n} \sum_{cn < k < (1-\epsilon)n} P_{k,x}^n[\sup_{0 \leq i \leq n-1} |X_i - x| > bN] +$$

$$\frac{1}{n} \sum_{cn < k < (1-\epsilon)n} P_{k,x}^n[\sup_{0 \leq i \leq n-1} |X_i - x| \leq bN, \{X_0, \ldots, X_{k-1}\} \cap (\bigcup_{i \geq 1} K_i) \neq \emptyset].$$

The first term above goes to 0, due to Lemma 15 (since $bN \geq n^\lambda$, $\lambda > 1/4$).
For the other term, we have when $N$ is large, (write $B_x(r)$ and $S_x(r)$ for the box of radius $r$ centered at $x$ and the boundary of $B_x(r)$ respectively)

\[
\frac{1}{n} \sum_{en<k<(1-\epsilon)n} P_{k,x}^n \left( \sup_{0\leq i\leq n-1} |X_i - x| \leq bN, \{X_0, \ldots, X_{k-1}\} \cap (\cup_{i\geq 1} K_i) \neq \emptyset \right) \leq \frac{1}{n} \sum_{en<k<(1-\epsilon)n} \sum_{i: K_0 \neq K_i \leq B_{i(\rho+1)N})} \sum_{y \in K_i} P_{k,x}^n[y \in \{X_0, \ldots, X_{k-1}\}] \leq \sum_{i=1}^{b+1} \sum_{j: K_j \cap S_{j}(iN) \neq \emptyset} \sum_{y \in K_j} \frac{1}{iN)^{d-4} \leq \sum_{i=1}^{b+1} \frac{1}{iN)^{d-4} \leq |K| \frac{4^4/\alpha_2 + 1}{N^{d-4}} \to 0,
\]

where the last convergence follows from $\lambda < \frac{4}{\alpha_2}$, $\alpha_2 < d$ and $n \leq N^{\alpha_2}$.

Proof (Proof of Lemma 13). Note that

\[
0 \leq P_{k,x}^n[X_{k-1}, X_{k-1+1}, \ldots, X_{k-1} \notin K] - P_{k,x}^n[X_0, \ldots, X_{k-1} \notin K] \leq P_{k,x}^n[\{X_0, \ldots, X_{k-1}\} \cap K \neq \emptyset] \leq \sum_{y \in K} P_{k,x}^n[y \in \{X_0, \ldots, X_{k-1}\}] \leq C(\epsilon) \sum_{y \in K} P_{k,x}^n[y \in \{X_0, \ldots, X_{k-1}\}],
\]

where for the last step, we use (4.6).

It would be ideal if we can establish a formula similar to (4.7), for $T^\infty$ and $T^\vee$, the random tree with law $\Pi_T$. But in this case, the following inequality is not generally true.

$P[v$ has exact $i$ elder siblings in $T^\infty] \leq P[v$ has exact $i$ elder siblings in $T^\vee]$. The following inequality is enough for our purpose.

Lemma 18.

\[
P_{k,x}^\infty[A] \leq P_{k,x}^\infty[A] + P_{k,x}^\infty[X_0 = y],
\]

where $A$ is the event $y \in \{X_0, \ldots, X_{k-1}\}$.

Proof. Write $A_0$ for the event $y \in \{X_1, \ldots, X_{k-1}\}$. We have $P_{k,x}^\infty[A] \leq P_{k,x}^\infty[A_0] + P_{k,x}^\infty[X_0 = y]$. For any $t \in T_\infty$ and $m \in \mathbb{N}^+$, write $\rho_m(t)$ for the marked tree $\rho_m(t)$ with those vertices that are in the spine of $t$, marked. Write $M$ for the set of all marked trees $t$ such that $P[\rho_{k+1}'(T^\infty) = t] > 0$. Note that

\[
P_{k,x}^\infty[A] = \sum_{t \in M} P[\rho_{k+1}'(T^\infty) = t] P_{k,x}^\infty[A | \rho_{k+1}'(T^\infty) = t],
\]

\[
P_{k,x}^\infty[A] \geq \sum_{t \in M} P[\rho_{k+1}'(T^\vee) = t] P_{k,x}^\infty[A | \rho_{k+1}'(T^\vee) = t].
\]
By a slight abuse of notation, we also write $T^\infty$ $(T^\nu)$ for the tree in the tree-indexed random walk corresponding to $P_{k,x}^\infty (P_{k,x}^\nu)$.

We divide $M$ into three disjoint subsets:

$M_1 = \{ t \in M : o$ has exactly one child and this child is marked$\}$,

$M_2 = \{ t \in M : o$ has exactly one child and this child is not marked$\}$,

$M_3 = \{ t \in M : o$ has $\geq 2$ children$\}$.

Obviously we have

$$P_{k,x}^\infty [A|\rho'_{k+1}(T^\infty) = t] = P_{k,x}^\nu [A|\rho'_{k+1}(T^\nu) = t], \text{ for } t \in M_1 \cup M_2.$$  

Moreover, it is easy to see

$$P[\rho'_{k+1}(T^\infty) = t] = \frac{1 - \mu(0)}{\mu(0)} P[\rho'_{k+1}(T^\nu) = t], \text{ for } t \in M_1,$$

$$P[\rho'_{k+1}(T^\infty) = t] = \frac{\sum_{i \geq 1} (i - 1) \mu(i)}{\sum_{i \geq 1} \mu(i)} P[\rho'_{k+1}(T^\nu) = t], \text{ for } t \in M_2.$$  

Hence we obtain

$$\sum_{t \in M_1 \cup M_2} P[\rho'_{k+1}(T^\infty) = t] P_{k,x}^\infty [A|\rho'_{k+1}(T^\infty) = t] \leq \sum_{t \in M_1 \cup M_2} P[\rho'_{k+1}(T^\nu) = t] P_{k,x}^\nu [A|\rho'_{k+1}(T^\nu) = t] \leq P_{k,x}^\nu [A].$$

It suffices to show that

$$\sum_{t \in M_3} P[\rho'_{k+1}(T^\infty) = t] P_{k,x}^\infty [A_0|\rho'_{k+1}(T^\infty) = t] \leq P_{k,x}^\nu [A_0].$$

For any $t \in M_3$, write $f(t)$ for $t$ re-rooted at the first child of $o$ with the new root added to the set of marked vertices (and $o$ being regarded as the last child of the new root). It is easy to check that

$$P_{k,x}^\infty [A_0|\rho'_{k+1}(T^\infty) = t] \leq P_{k,x}^\nu [A|\rho'_{k+1}(T^\nu) = f(t)], \text{ for } t \in M_3.$$  

On the other hand, one can verify that

$$P[\rho'_{k+1}(T^\infty) = t] \leq P[\rho'_{k+1}(T^\nu) = f(t)], \text{ for } t \in M_3.$$  

Hence, we have

$$\sum_{t \in M_3} P[\rho'_{k+1}(T^\infty) = t] P_{k,x}^\infty [A_0|\rho'_{k+1}(T^\infty) = t] \leq \sum_{t \in M_3} P[\rho'_{k+1}(T^\nu) = f(t)] P_{k,x}^\nu [A|\rho'_{k+1}(T^\nu) = f(t)] \leq P_{k,x}^\nu [A],$$

Note that we use the fact that $f$ is injective for the last step. The proof of (4.13) is complete.
Thanks to the last lemma, (note that $P_{k,x}^\infty[X_0 = y] = P_{0,y}^\infty[X_k = x]_{k \to \infty} 0$,) it suffices to show, for any $y \in K$,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{c_n < k < (1-\epsilon)n} P_{k,x}^\infty[y \in \{X_0, \ldots, X_{k-1}\}] = 0. \quad (4.14)$$

Write $B = B(k)$ and $S(B)$ for the set of non-root vertices in the spine, that are not after $V_k$, and the range of $B$, respectively. It is standard that $P_{k,x}^\infty[V_k \in B] = o(1)$ (when $k \to \infty$). Hence we have

$$P_{k,x}^\infty[y \in \{X_0, \ldots, X_{k-1}\}] \leq P_{k,x}^\infty[y \in \{X_0, \ldots, X_{k-1}\}, V_k \notin B] + o(1). \quad (4.15)$$

By changing the spatial tree a bit, around the fist visiting point of $y$ by $S(B)$, one can show that for any $i_0 \in \mathbb{N}^+$ fixed, with $\mu(i_0) > 0$,

$$P_{k,x}^\infty[y \in S(B) \setminus (S(\{V_0, \ldots, V_{k-1}\} \setminus B)), V_k \notin B] \leq \sum_{i = 0, i_0} P_{k+i,x}^\infty[y \in S(\{V_0, \ldots, V_{k-1}\} \setminus B), V_k \notin B].$$

Here is a naive argument. Assume that $(t, S)$ is a spatial tree in the event on the left hand side. We need to argue that we can make $(t, S)$ visit $y$ for some vertex before $V_k$, not in the spine, by changing $(t, S)$ a bit, e.g. by changing one variable in some edge of $t$ or inserting $i_0$ new edges and the corresponding variables in the new edges. Let $v \in B$ be the first vertex such that $S(v) = y$ and $v_1$ be the parent of $v$. If $v$ has elder siblings, then just pick up the eldest one and change the variable in the edge connecting $v_1$ and that vertex such that the sum of those two variables in the edges connecting $v$ and that vertex is zero (we can do so since $\theta$ is symmetric). If $v$ has no elder siblings, we can achieve our goal by inserting $i_0$ elder siblings for $v$ and sending the eldest one to $y$.

From the last inequality, one can obtain

$$P_{k,x}^\infty[y \in \{X_0, \ldots, X_{k-1}\}, V_k \notin B] \leq \sum_{i = 0, i_0} P_{k+i,x}^\infty[y \in S(\{V_0, \ldots, V_{k-1}\} \setminus B), V_k \notin B]. \quad (4.16)$$

Write $B(a, b) = B(a, b; t)$ for the number of spine vertices in $\{V_i(t) : i \in [a, b]\}$. The following estimate of $B(a, b; T^w)$ is standard and easy:

$$\lim_{n \to \infty} \sup_{m \in \mathbb{N}^+} P[B(m, m + n; T^w) \geq n/2] = 0. \quad (4.17)$$

Joining (4.14) with (4.15), (4.16) and (4.17), one can see that it suffices to show that (by making $\epsilon$ smaller)

$$\lim_{n \to \infty} \frac{1}{n} \sum_{c_n < k < (1-\epsilon)n} P_{k,x}^\infty[y \in S(\{V_0, \ldots, V_{k-1}\} \setminus B(k)), V_k \notin B(k), B(k-l, k) \leq l/2] = 0. \quad (4.18)$$
Let us introduce our notations. Recall that for \( t \in T_\infty \), we have defined the depth-first search from infinity. Write \((V_i^\infty(t))_{i \in \mathbb{Z}}\) for all vertices arranged with that order such that \( V_0^\infty(t) \) is the root (we will drop \( t \) in the notation when \( t \) is obvious, as before). For any infinite spatial tree \((t, S)\), write \( Y_t = S(V_i^\infty(t)) \).

We write \( P_{k,x}^V \) for the law when \( t \) has the law of \( T^V \) and conditionally on \( t \), \( S \) is a \( t \)-indexed random walk sending \( V_k^\infty \) to \( x \).

Let us continue our proof. Note that

\[
P_{k,x}^V[y \in S\{\{V_0, \ldots, V_{k-1}\} \setminus B(k)\}, V_k \notin B(k), B(k-l, k) \leq l/2] = \sum_{m=0}^{k-1} P_{k,x}^V[y \in S\{\{V_0, \ldots, V_{k-1}\} \setminus B(k)\}, V_k \notin B(k), |B(k)| = m, B(k-l, k) \leq l/2] = \sum_{m=0}^{k-1} P_{k-m,x}^V[y \in S\{\{V_0, \ldots, V_{k-1}\} \setminus B(k)\}, V_k \notin B(k), |B(k)| = m, B(k-l, k) \leq l/2] \leq \sum_{m=0}^{k-1} P_{k-m,x}^V[y \in \{V_0, \ldots, V_{k-m-l/2}\}, V_k \notin B(k), |B(k)| = m, B(k-l, k) \leq l/2] \leq \sum_{m \leq \lfloor k^{0.6} \rfloor} P_{k-m,x}^V[y \in \{V_0, \ldots, V_{k-m-l/2}\}, |B(k)| = m] + o(1).
\]

Note that for the third line, we use the fact that conditionally on \(|B(k)| = m\) and \( V_k \notin B(k)\), \( V_k = V_{k-m}^\infty\); for the second last line, that conditionally on \( B(k-l, k) \leq l/2\),

\[
S\{\{V_0, \ldots, V_{k-1}\} \setminus B(k)\} \subset \{V_0, \ldots, V_{k-m-l/2}\};
\]

and for the last line, that \( P_{k-m,x}^\infty[|B(k)| \geq k^{0.6}] = P_{0,x}^\infty[|B(k)| \geq k^{0.6}] = o(1)\).

Hence, \( (4.18) \) can be reduced again to

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{\epsilon n < k < (1-\epsilon)n} \sum_{m \leq \lfloor k^{0.6} \rfloor} P_{k-m,x}^\infty[y \in \{V_0, \ldots, V_{k-m-l/2}\}, |B(k)| = m] = 0.
\]

We have (when \( n \) is large enough)

\[
\sum_{\epsilon n < k < (1-\epsilon)n} \sum_{m \leq \lfloor k^{0.6} \rfloor} P_{k-m,x}^\infty[y \in \{V_0, \ldots, V_{k-m-l/2}\}, |B(k)| = m] \leq \sum_{\epsilon n/2 < j < (1-\epsilon/2)n} \sum_{m \leq \lfloor j^{0.6} \rfloor} P_{j,x}^\infty[y \in \{V_0, \ldots, V_{j-l/2}\}, |B(k)| = m] \leq \sum_{\epsilon n/2 < j < (1-\epsilon/2)n} P_{j,x}^\infty[y \in \{V_0, \ldots, V_{j-l/2}\}] = \sum_{\epsilon n/2 < j < (1-\epsilon/2)n} P_{0,x}^\infty[y \in \{V_{-j}, V_{-j+1}, \ldots, V_{-l/2}\}],
\]

(4.19)
where for the last step we use Proposition 6.

Note that \( P_{0,x}^{∞}[y \in \{Y_{-j}, Y_{-j+1}, \ldots, Y_{-l/2}\}] \leq P_{0,x}^{∞}[y \in \cup_{i \leq -l/2}\{Y_i\}] \). Since on \( \mathbb{Z}^d (d \geq 5) \), every finite subset is branching transient (see Section 2 in [20]) we get \( \lim_{l \to \infty} P_{0,x}^{∞}[y \in \cup_{i \leq -l/2}\{Y_i\}] = 0 \) for any \( x, y \in \mathbb{Z}^d \). Therefore (4.19) follows and we finish the proof.

Proof (Proof of Lemma [14]). First note that the subtree in \( T^n \) generated by the vertices corresponding to \( X_{k-1}, \ldots, X_k \) is contained in \( H_l(T^n, V_k) \) (recall the definition of \( H_l \) in Section 4.2), at least when the height of \( V_k \) is bigger than \( l \). By Theorem 10, we have \( H_l(T^n, V_k) \xrightarrow{\text{unifromly for } k} H_l(\Sigma, u_0) \) (uniformly for \( k \in [[\epsilon n, (1-\epsilon)n]] \)). Note that, as an unrooted tree, \( H_l(\Sigma, u_0) \) has the same distribution as \( T^v \) up to generation \( l \). Hence, we have

\[
\lim_{n \to \infty} \max_{k \in [[\epsilon n, (1-\epsilon)n]]} |P_{n,x}^v[X_{k-1}, \ldots, X_{k-1} \notin K] - P_{0,x}^{∞}[Y_{-l}, \ldots, Y_{-1} \notin K]| = 0.
\]

Note that the second probability converges to \( \operatorname{Es}_K(x) \) when \( l \to \infty \), (recall that \( \operatorname{Es}_K(x) = P_{x}^{v}(W_{K}^{0}) \)). We finish the proof.

4.2 Proof of Theorem 14

We have constructed the ingredients we need and are ready to show our main result. We follow the arguments in [4].

Let \( T \) be the corresponding family tree in \( S^{n,N} \). Apply Theorem 13 to \( T \), with \( (d_0, e, \alpha, \beta, \delta, \epsilon) = (d, 2, 2.01, d - 0.01, 0.01, 0.01) \) (note that \( e_0 = 3\beta/2 - \alpha - d - 7\delta/2 = 0.5d - 2.06 - 0.44 \)). In fact, the values of \( \alpha, \beta, \epsilon, \delta \) are not important. What we essentially need is \( \alpha > 2 \). With high probability \( (1 - C/N^*) \), we can find subtrees \( T_1, \ldots, T_m \) of \( T \) as in the theorem. We denote by \( A \) this event. We write \( P[\bullet|(m; k_1, \ldots, k_m; t)] \) (respectively \( p(m; k_1, \ldots, k_m; t) \)) for the conditional probability conditioned (respectively the probability) that \( A \) is true, the number of \( T_i \) is \( m \), the size of \( T_i \) is \( k_i \) \( (i = 1, \ldots, m) \) and the subtree \( \hat{T} \) with \( m \) (ordered) marked vertices indicating the locations of \( v \) is \( t \). Under \( P[\bullet|(m; k_1, \ldots, k_m; t)] \), the trees \( T_1, \ldots, T_m \) are independent and distributed as the GW-trees with the given sizes. We have

\[
P[\Gamma(S^{n,N}) \cap \varphi(K) = \emptyset] = \sum p(m; k_1, \ldots, k_m; t) P[\Gamma(T) \cap \varphi(K) = \emptyset|(m; k_1, \ldots, k_m; t)] + o(1),
\]

where the sum runs over all possible values of \( T = (m; k_1, \ldots, k_m; t) \) such that \( p(T) > 0 \) (depending on \( N \)). For notational ease, we simply write \( S \) for \( S^{n,N} \). It suffices to show

\[
\lim_{N \to \infty} \max_{T} |P[\Gamma(T) \cap \varphi(K) = \emptyset|T] - \exp(-\mu\text{Cap}(K))| = 0. \tag{4.20}
\]

The equality above can be reduced to:

\[
\lim_{N \to \infty} \max_{T} |P[\Gamma(T) \cap \varphi(K) = \emptyset|T] - P[\bigcup_{i=1}^{m} \Gamma(T_i) \cap \varphi(K) = \emptyset|T]| = 0. \tag{4.21}
\]
There exist positive Lemma 19.

for all $v \in \mathcal{T}$.

For (4.23), by Condition (1) and (4) in Theorem 13 we know that $|T_i| \in [N^{\alpha_1}, N^{\alpha_2}]$ (for some $\alpha_1, \alpha_2 \in (0, d)$) and that conditionally on the size, $T_i$ is a GW-tree conditioned on that size. Hence we can apply Theorem 15. Then together with Condition (2), one can get (4.23).

We now turn to (4.22). We need the following lemma.

**Lemma 19.** There exist positive $c$ and $C$, such that, for any $N \in \mathbb{N}^+$ and $\mathcal{T} = (m; k_1, \ldots, k_m; t)$ with $p(\mathcal{T}) > 0, k \in [1, m - 1]$, then

\[
|P[(\bigcup_{i=1}^m S(T_i)) \cap \varphi(K) = \emptyset | \mathcal{T}] - P[(\bigcup_{i=1}^m S(T_i)) \cap \varphi(K) = \emptyset | \mathcal{T}]| \leq C \exp(-cN^{\alpha - 2}).
\]  

(4.24)

With this lemma one can use induction to show

\[
|P[(\bigcup_{i=1}^m S(T_i)) \cap \varphi(K) = \emptyset | \mathcal{T}] - \prod_{i=1}^m P[S(T_i) \cap \varphi(K) = \emptyset | \mathcal{T}]| \
\leq (m - 1)C \exp(-cN^{\alpha - 2}).
\]  

(4.25)

Since $m$ is bounded by a polynomial of $N$, the right hand side tends to 0, which implies (4.22).

**Proof (Proof of Lemma 19).** Let $o_1$ and $o_2$ be the endpoints of $t_k$ in Theorem 13 (say $o_1 \in T_k$). For any $x, y \in \mathbb{T}_N$, define

\[
f(x) = P[(S(T_k)) \cap \varphi(K) = \emptyset | S(o_1) = x, \mathcal{T}],
\]  

(4.26)

\[
h(y) = P[(\bigcup_{i=k+1}^m S(T_i)) \cap \varphi(K) = \emptyset | S(o_2) = y, \mathcal{T}].
\]  

(4.27)

By Condition (3) in Theorem 13 $t_k$ separates $T_k$ and $\bigcup_{i=k+1}^m T_i$, so we have

\[
P[(\bigcup_{i=1}^m S(T_i)) \cap \varphi(K) = \emptyset | S(o_1) = x, S(o_2) = y, \mathcal{T}] = f(x) \times h(y).
\]  

(4.28)
Therefore,
\[
P \left[ \bigcup_{i=k}^m S(T_i) \cap \varphi(K) = \emptyset \mid T \right] = \sum_{x,y \in T_N} f(x)h(y)P[S(o_1) = x, S(o_2) = y \mid T]
\]
\[
= N^{-d} \sum_{x,y \in T_N} f(x)h(y)P^\text{RW}_x[Z(\lfloor N^\alpha \rfloor) = y],
\]
where \(P^\text{RW}_x\) is the law of \(Z = (Z(n))_{n \in \mathbb{N}}\) which is a random walk starting at \(x\) with distribution \(\theta\) in \(T_N\). Note that
\[
P \left[ \bigcup_{i=k}^m S(T_i) \cap \varphi(K) = \emptyset \mid T \right] = N^{-d} \sum_{x \in T_N} f(x)h(y)P^\text{RW}_x[Z(\lfloor N^\alpha \rfloor) = y - N^{-d}],
\]
(4.29)
\[
P \left[ \bigcup_{i=k+1}^m S(T_i) \cap \varphi(K) = \emptyset \mid T \right] = N^{-d} \sum_{y \in T_N} h(y).
\]
(4.30)
Hence the left hand side of (4.24) is
\[
\left| N^{-d} \sum_{x \in T_N} f(x)h(y)(P^\text{RW}_x[Z(\lfloor N^\alpha \rfloor) = y - N^{-d}]) \right|
\]
\[
\leq \max_{x \in T_N} \sum_{y \in T_N} |P^\text{RW}_x[Z(\lfloor N^\alpha \rfloor) = y - N^{-d}]|.
\]
Now (4.24) follows from the following standard result in the theory of mixing time (e.g. see [15]).

Lemma 20. There exist positive numbers \(c\) and \(C\) such that for any \(N \in \mathbb{N}^+\) and \(a > 2\),
\[
\max_{x,y \in T_N} |P^\text{RW}_x[Z(\lfloor N^\alpha \rfloor) = y - N^{-d}]| \leq C \exp(-cN^{a-2}).
\]
(4.31)

5 Cover times of tori by tree-indexed random walks

The main goal of this section is to construct the result of cover times of \(d\)-dimensional tori by tree-indexed random walks conditioned on sizes. Recall that we adopt the same assumptions as in the last section on the offspring distribution \(\mu\) and the jump distribution \(\theta\) (and \(d \geq 5\)). As before, write \(S^{n,N}\) for the tree-indexed random walk in torus \(T_N\), with size \(n\) and a uniform starting point.

Theorem 16. Let \(n = n(N)\) be an integer-valued function of \(N\). For any \(\epsilon \in (0,1)\), we have if \(n(N) > (1 + \epsilon)N^d \log N^d / B\text{Cap}(\{0\})\), then
\[
\lim_{N \to \infty} P[\text{Range}(S^{n,N}) = T_N] = 1;
\]
if \(n(N) < (1 - \epsilon)N^d \log N^d / B\text{Cap}(\{0\})\), then
\[
\lim_{N \to \infty} P[\text{Range}(S^{n,N}) = T_N] = 0.
\]
We need two lemmas in our proof of the theorem above. Theorem 15 gives asymptotics for the probability that a set is visited by a tree-indexed random walk with a uniform starting point. The following proposition gives asymptotics for a tree-indexed random walk with a fixed starting point, which is not too close to the set.

**Proposition 17** For any \( \beta_1, \beta_2, \gamma \) fixed with \((2d + 4)/3 < \beta_1 < \beta_2 < d\), \((3d - 3\beta_1 + 4)/d < \gamma < 1\), \(n = n(N)\), any integer-valued function of \(N\) satisfying \(n(N) \in [N^{\beta_1}, N^{\beta_2}]\), and \(K \subset \mathbb{Z}^d\), we have

\[
\lim_{N \to \infty} \sup_{\rho(x, \varphi(K)) > N^\gamma} \frac{|N^d P[\text{Range}(\mathcal{S}) \cap \varphi(K) \neq \emptyset] - \text{BCap}(K)|}{n} = 0, \tag{5.1}
\]

where \(\mathcal{S} = \mathcal{S}_x^{n,N}\) is a tree-indexed random walk in \(T_N\) with size \(n\) and starting point \(x\) and \(\rho\) is the graph distance of the torus.

**Proof.** Fix any \(x\) with \(\rho(x, \varphi(K)) > N^\gamma\). By our assumptions on \(\beta_1, \gamma\), we can find \(\alpha > 0\) satisfying

\[
\beta_1 + d\gamma > 4 + 2\alpha + d, \quad \alpha + \beta_1 > d. \tag{5.2}
\]

Write \(T\) for the corresponding tree in \(\mathcal{S}\). Applying Proposition 12 with \(k = \lfloor N^{2+\alpha} \rfloor, L = \lfloor k^2 N^{2\alpha} \rfloor\) to \(T\), we obtain that, with probability at least \(1 - C N^{-\alpha - c}\) (\(C\) may depend on \(\beta_1, \beta_2, \gamma, c\)), there exists a subtree \(T_0\) which is at distance \(k\) away from the root of \(T\) without losing more than \(L\) vertices. Assume that this event has occurred and condition on the value of \(T \setminus T_0\). Write \(o_0\) for the root of \(T_0\). Similarly to the situation in the proof of Theorem 15, one can see that the law of \(\mathcal{S}(o_0)\) is very close to the uniform measure. Therefore, by Theorem 15, we have

\[
P[\mathcal{S}(T_0) \cap \varphi(K) \neq \emptyset] \sim \frac{|T_0|}{N^d} \text{BCap}(K) \sim \frac{n}{N^d} \text{BCap}(K).
\]

We point out that

\[
P[\mathcal{S}(T \setminus T_0) \cap \varphi(K) \neq \emptyset] \leq \frac{L|K|}{N^{d\gamma}}. \tag{5.3}
\]

If so, then we finish the proof since by (5.2), we have (when choosing \(c\) small enough)

\[
N^{-(\alpha - c)} \leq \frac{n}{N^{d+o(1)}}, \quad \frac{L|K|}{N^{d\gamma}} \leq \frac{n}{N^{d+o(1)}}.
\]

We still need to show (5.3). It can be obtained from the following. For any \(y \in T_N\) with \(\rho(x, y) > N^\gamma\) and \(v \in T \setminus T_0\), we have

\[
P[\mathcal{S}(v) = y] \leq \sup_{n \in \mathbb{N}} P[Z_x(n) = y] \leq N^{-d\gamma},
\]

where \(Z_x = (Z_x(n))_{n \in \mathbb{N}}\) is a random walk starting at \(x\) in \(T_N\), and the last step follows from the Local Central Limit Theorem.
Lemma 21. For any $\beta_1 < \beta_2 \in (0, d)$, $\epsilon > 0$ and $n = n(N)$, any integer-valued function of $N$ such that $n(N) \in [N^{\beta_1}, N^{\beta_2}]$, there exists $C = C(\epsilon)$, such that for any $x \neq y \in \mathbb{T}_N$,

$$P\{x, y \in \mbox{Range}(S^{n,N})\} \leq C \frac{n}{N^d} \left( \frac{1}{\rho(x, y)^{d-4}} + \frac{(\log N)^{3+\epsilon}}{N^{d-\beta_2}} \right).$$

Proof. Note that the left hand side is no greater than

$$\sum_{0 \leq i < j \leq n-1} (P^{n,N}_{i}[X_i = x, X_j = y] + P^{n,N}_{i}[X_i = y, X_j = x]) = 2 \sum_{0 \leq i < j \leq n-1} P^{n,N}_{i}[X_i = x, X_j = y].$$

Hence, we have

$$\frac{N^d}{n} P\{x, y \in \mbox{Range}(S^{n,N})\} \leq \frac{N^d}{n} \sum_{0 \leq i < j \leq n-1} P^{n,N}_{i}[X_i = x, X_j = y]$$

$$= \frac{1}{n} \sum_{0 \leq i < j \leq n-1} P^{n,N}_{i}[X_j = y] = \frac{1}{n} \sum_{0 \leq i < j \leq n-1} \sum_{k} P^{n,N}_{i}[X_j = y, d(V_i, V_j) = k]$$

$$= \frac{1}{n} \sum_{k} P[Z_x(k) = y] \sum_{0 \leq i < j \leq n-1} P^{n,N}_{i}[d(V_i, V_j) = k],$$

where $P^{n,N}, P^{n,N}_{i,x}$ are similar to the ones in Section 4, $d(V_i, V_j)$ is the graph distance of the $i$-th vertex $V_i$ and the $j$-vertex $V_j$ in the corresponding tree $T^n$, and $Z_x = (Z_x(i))_{i \in \mathbb{N}}$ is a random walk in $\mathbb{T}_N$ with starting point $x$. From Lemma 17 we have

$$\frac{1}{n} \sum_{0 \leq i < j \leq n-1} P^{n,N}_{i,x}[d(V_i, V_j) = k] \leq k.$$ 

and therefore

$$\frac{N^d}{n} P\{x, y \in \mbox{Range}(S)\} \leq \sum_{k} k P[Z_x(k) = y].$$

Note that when $x \neq y \in \mathbb{Z}^d$,

$$\sum_{k} P[Z_x(k) = y] \cdot k \leq |x - y|^{d-4}. \quad (5.4)$$

This is false for $x \neq y \in \mathbb{T}_N$ since when $x, y \in \mathbb{Z}^d$ with $\varphi(x) \neq \varphi(y)$,

$$P[Z_{\varphi(x)}(k) = \varphi(y)] = \sum_{z \in \mathbb{Z}^d} P[Z_x(k) = y + N z].$$
However, similarly to the argument in the proof of Lemma 15, one can use Lemma 15 to rule out those \( z \) with norm bigger than \( b = N^{\beta_2/4 - 1}(\log N)^{0.75 + o(1)} \). The sum over the remaining \( z \) has the desired bound:

(\text{without loss of generality, assume that } y \text{ achieves } \min_{y' : \varphi(y') = \varphi(y)} \{|x - y'|\})

\[
\sum_{z \in \mathbb{Z}^d : |z| \leq b} kP[Z_x(k) = y + Nz] \geq \sum_{z \in \mathbb{Z}^d : |z| \leq b} |x - y - Nz|^{4-d}
\]

\[
\approx |x - y|^{4-d} + \sum_{i=1}^{b} i^{d-1} \left( \frac{1}{(1+0N^d)^{d-1}} \right) \leq C(\epsilon)(\log N)^{3+\epsilon} N^{d-\beta}.
\]

The proof is complete.

**Proof (Proof of Theorem 16).** Write \( a_0 = \text{BCap}(\{0\}) \) and \( S = S^{N,N} \). We start with the upper bound. Assume that \( n(N) > (1+\epsilon)N^d \log N^d/a_0 \). Let \( T \) be the corresponding family tree in \( S \). Apply Theorem 13 with \( (d_0, c, \alpha, \beta, \delta, \epsilon) = (d, 2, 2.01, d-0.01, 0.01) \). Then, by discarding an event with small probability, we can find subtrees \( T_1, \ldots, T_m \) as in the theorem. Write \( l_1, \ldots, l_m \) for their sizes. By Theorem 15, we have (when \( N \) is large enough)

\[
P[x \in S(T_i)] \geq (1 - 0.1\epsilon) \frac{l_i}{N^d} a_0,
\]

for any \( x \in \mathbb{T}_N \). Similarly to the argument in the last section, by Lemma 19, one can get

\[
P[x \notin \bigcup_{i=1}^m S(T_i)] \leq \prod_{i=1}^m P[x \notin S(T_i)] + C \exp(-cN^{\alpha-2}).
\]

Note that (when \( N \) is large)

\[
\prod_{i=1}^m P[x \notin S(T_i)] \leq \prod_{i=1}^m (1 - (1 - 0.1\epsilon) \frac{l_i}{N^d} a_0) \leq \exp((-1 - 0.2\epsilon) \frac{\sum_{i=1}^m l_i}{N^d} a_0)
\]

\[
\leq \exp((-1 - 0.3\epsilon)(1 + \epsilon) \log(N^d)) \leq \exp((-1 + 0.6\epsilon) \log(N^d)).
\]

Hence we have

\[
P[x \notin \bigcup_{i=1}^m S(T_i)] \leq \frac{1}{N^{d(1+0.5\epsilon)}}.
\]

Since there are only \( N^d \) vertices in \( \mathbb{T}_N \), one can get \( P[\text{Range}(S) \neq \mathbb{T}_N] \leq N^{-0.50d} \). This finishes the proof of the upper bound.

We now turn to the lower bound. Fix some \( \eta < \epsilon \wedge 0.01 \). Apply Theorem 13 with

\[
(d_0, c, \alpha, \beta, \delta, \epsilon) = (d, 2, 2 + s, d - s, s, s),
\]

where \( s \) is a small number (in fact \( 9s < 0.5 \) is enough). By discarding an event with small probability, we can assume that there exist subtrees \( T_1, \ldots, T_m \) as in the theorem. Write \( l_1, \ldots, l_m \) for their sizes and \( o_1, \ldots, o_m \) for their roots. We
condition on the values of the subtree \( \hat{T} \) (recall this notation in the theorem) and \( S_{\hat{T}} \) (and the values of \( l_1, \ldots, l_m \)). Note that conditionally on these values, \((T_1, S_{T_1}), \ldots, (T_m, S_{T_m})\) are independent tree-indexed random walks with given sizes and starting points. By Proposition \ref{prop:independence} with \( \beta_1 = \beta - 2s - o(1) \) and \( \beta_2 = \beta \), we see that as long as \( \gamma > (9s + 4)/d \), \ref{eq:estimate} holds for \( S(T_i) \). Let \( \gamma = 4.5/d \). Then we have \ref{eq:estimate}.

We claim that we can pick up \( \lfloor N^{(1-\eta)d} \rfloor \) vertices (write \( F \) for the set of these vertices) in \( T_N \), such that the distance between any two vertices in \( F \) is at least \( N^\eta/8 \). Here is a way to do so. First we can find at least \( 4N^{(1-\eta)d} \) boxes with radius \( N^\eta/8 \) in \( T_N \) such that the distance between any two in them are at least \( N^\eta/8 \). Write \( B_0 \) for the set of these boxes. Note that for each \( y \in A_1 = S(\{a_i\}) \), there are at most \( (32N^\gamma /N^\eta)^d \times N^{4.5-2d_\eta} \) boxes in \( B_0 \) whose distance from \( y \) is at most \( N^\gamma \). Since there are only \( m \) (\( m \leq N^{4.5} \)) vertices in \( A_1 \), there are at least \( 3N^{(1-\eta)d} \) boxes in \( B_0 \) which are at least \( N^\gamma \) away from \( A \). On the other hand, note that there are at most \( n/N^\gamma \ll N^d \) vertices in \( S(\hat{T}) \) and that a strictly positive proportion of vertices in \( T_N \) are covered by the remaining \( 3N^{(1-\eta)d} \) boxes. Hence, most of the \( 3N^{(1-\eta)d} \) boxes are not fully covered by \( S(T) \) and in each such box we can pick up a point not in \( S(\hat{T}) \). In this way, we obtain \( F \) with the desired properties.

We argue that with high probability, \( F \) is not covered by \( R = S(\cup_i^m T_i) \), and therefore \( F \) is not covered by \( S(T) \). Then we finish the proof.

Without loss of generality, assume that \( \sum_{i=1}^m l_i \sim (1-\epsilon)N^d \log N/d/a_0 \). Write \( U = \sum_{x \in F} 1_{x \notin R} \). We need to show that it is positive with high probability. We manage to do so by using the Chebyshev inequality and hence need to estimate the expectation and the variance of \( U \). First, note that (since \( (S(T_i))_{i=1,\ldots,m} \) are independent)

\[
E[1_{x \notin R}] = P[x \notin S(\cup_i^m T_i)] = \prod_i^m P[x \notin S(T_i)] = \prod_i^m (1 - P[x \in S(T_i)]).
\]

Using \ref{eq:estimate}, we have

\[
\prod_i^m (1 - P[x \in S(T_i)]) = \exp(-\frac{(1 + o(1))a_0 \sum_{i=1}^m l_i}{N^d}) = \frac{1}{N^{(1-\epsilon)d + o(1)}}.
\]

Hence, we obtain \( E[U] = |F|/N^{(1-\epsilon)d + o(1)} = N^{d(\epsilon - \eta) + o(1)} \).

We now turn to the variance. Note that \( \text{Var}[U] = \sum_{x,y \in F} q_{x,y} \leq E[U] + \sum_{x \notin F} q_{x,y} \) where \( q_{x,y} = P[x, y \notin R] - P[x \notin R]P[y \notin R] \). Note that, as before, by independence, we have

\[
P[x, y \notin R] = \prod_i^m P[x, y \notin S(T_i)],
\]

\[
P[x \notin R]P[y \notin R] = \prod_i^m P[x \notin S(T_i)]P[y \notin S(T_i)].
\]
On the other hand,

\[ P[x, y \notin S(T_i)] = 1 - P[x \in S(T_i)] - P[y \in S(T_i)] + P[x, y \in S(T_i)]. \]

We point out an inequality and show it later: for any \( x \neq y \in F \)

\[ \frac{N^d}{l_i} P[x, y \in S(T_i)] \leq C \frac{1}{N^c}, \tag{5.5} \]

where \( C, c \) are constants which may depend on \( s, \eta, \epsilon \). Therefore, we have

\[
\frac{P[x, y \notin R]}{P[x \notin R] P[y \notin R]} \leq \prod_i \left( 1 - \frac{1 - P[x \in S(T_i)] - P[y \in S(T_i)] + P[x, y \in S(T_i)]}{(1 - P[x \in S(T_i)])(1 - P[y \in S(T_i)])} \right)
\leq \prod_i \left( 1 + C \frac{l_i}{N^d+1} \right) \leq 1 + C \frac{N}{N^c}.
\]

From this, we get \( q_{x,y} \leq \frac{P[x \notin R] P[y \notin R]}{N^c} \) and hence

\[
\sum_{x \neq y \in F} q_{x,y} \leq \left( \sum_x P[x \notin R] \right)^2 / N^c = (E[U])^2 N^{-c}.
\]

By the Chebyshev inequality, we have

\[
P[|U - E[U]| > \frac{E[U]}{2}] \leq \frac{4 \text{Var}[U]}{(E[U])^2} \leq \frac{1}{N^{c/2}}.
\]

Hence, at least \( E[U]/2 \) vertices in \( F \) are not covered by \( S(\cup_i^m T_i) \) with high probability.

We still need to show (5.5). The argument is almost the same as the one in the proof of Proposition 17 except that we need to use Lemma 21 instead of Theorem 15. As in the proof of Proposition 17 (recall that we have verified that our \( \gamma \) and the size \( l_i \) satisfy the requirements in that proposition, see the paragraph after the proof of the upper bound) by discarding an event with probability less than \( C l_i / N^{d+c} \), we can find a subtree \( T_0 \) of \( T_i \) as in the proof of proposition. Similarly, we have

\[ P[S(T_i \setminus T_0) \cap \{x, y\} \neq \emptyset] \leq \frac{l_i}{N^{d+c}}. \]

For \( S(T_0) \), its starting point is very close to the uniform measure hence we can apply Lemma 21 and obtain

\[ P[x, y \in S(T_0)] \leq \frac{l_i}{N^{d+c}}. \]

Now (5.5) follows and the proof is complete.

Acknowledgements

This work was partially done when the author was a PhD student in the University of British Columbia. The author would like to thank Omer Angel and Balázs Ráth for supervising his research and useful discussions.
References

1. David Aldous. Asymptotic fringe distributions for general families of random trees. Ann. Appl. Probab., 1(2):228-266, 1991.
2. David Aldous. On the time taken by random walks on finite groups to visit every state. Z. Wahrsch. Verw. Gebiete, 62(3):361-374, 1983.
3. David Aldous. Threshold limits for cover times. J. Theoret. Probab., 4(1):197-211, 1991.
4. Omer Angel, Balázs Ráth and Qingsan Zhu. Local limit of branching random walk on torus and branching interlacements. Preprint.
5. David Belius. Gumbel fluctuations for cover times in the discrete torus. Probab. Theory Related Fields, 157(3-4):635-689, 2013.
6. M. J. A. M. Brummelhuis and H. J. Hilhorst. Covering of a finite lattice by a random walk. Phys. A, 176(3):387-408, 1991.
7. Amir Dembo, Yuval Peres, Jay Rosen and Ofer Zeitouni. Cover times for Brownian motion and random walks in two dimensions. Ann. of Math. (2), 160(2):433-464, 2004.
8. Luc Devroye and Svante Janson. Distances between pairs of vertices and vertical profile in conditioned Galton-Watson trees. Random Structures Algorithms, 38(4):381-395, 2011.
9. Svante Janson. Conditioned Galton-Watson trees do not grow. Preprint, arXiv:math/0604141.
10. Svante Janson. Random cutting and records in deterministic and random trees. Random Structures Algorithms, 29(2):139-179, 2006.
11. Gregory F. Lawler and Vlada Limic. Random walk: a modern introduction. Cambridge University Press, 2010.
12. Jean-François Le Gall. Spatial branching processes, random snakes and partial differential equations. Birkhäuser Verlag, Basel, 1999.
13. Jean-François Le Gall. Random trees and applications. Probab. Surv., 2:245-311, 2005.
14. Jean-François Le Gall and Shen Lin. The range of tree-indexed random walk. Journal of the Institute of Mathematics of Jussieu, 15(2):271-317, 2016.
15. David A. Levin and Yuval Peres. Markov chains and mixing times. American Mathematical Society, Providence, RI, 2017. With contributions by Elizabeth L. Wilmer, With a chapter on “Coupling from the past” by James G. Propp and David B. Wilson.
16. Malwina Luczak and Peter Winkler. Building uniformly random subtrees. Random Structures Algorithms, 24(4):420-443, 2004.
17. Jim Pitman. Combinatorial stochastic processes. Springer-Verlag, Berlin, 2006. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002. With a foreword by Jean Picard.
18. Sergeï Popov and Augusto Teixeira. Soft local times and decoupling of random interlacements. J. Eur. Math. Soc. (JEMS), 17(10):2545-2593, 2015.
19. Eviatar B. Procaccia and Yuan Zhang. Connectivity properties of Branching Interlacements. Preprint, arXiv:1612.00406.
20. Balázs Ráth and Artём Sapozhnikov. On the transience of random interlacements. Electron. Commun. Probab., 16:379-391, 2011.
21. Vladas Sidoravicius and Alain-Sol Sznitman. Percolation for the vacant set of random interlacements. Comm. Pure Appl. Math., 62(6):831-858, 2009.
22. Vladas Sidoravicius and Alain-Sol Sznitman. Connectivity bounds for the vacant set of random interlacements. Ann. Inst. Henri Poincaré Probab. Stat., 46(4):976-990, 2010.

23. Benedikt Stufler. Local limits of large Galton-Watson trees rerooted at a random vertex. Ann. Inst. Henri Poincaré Probab. Stat., 55(1):155-183, 2019.

24. Alain-Sol Sznitman. Vacant set of random interlacements and percolation. Ann. of Math. (2), 171(3):2039-2087, 2010.

25. Alain-Sol Sznitman. Decoupling inequalities and interlacement percolation on $G \times \mathbb{Z}$. Invent. Math., 187(3):645-706, 2012.

26. Augusto Teixeira. Interlacement percolation on transient weighted graphs. Electron. J. Probab., 14:no.54, 1604-1628, 2009.

27. Augusto Teixeira and David Windisch. On the fragmentation of a torus by random walk. Comm. Pure Appl. Math., 64(12):1599-1656, 2011.

28. Qingsan Zhu. On the critical branching random walk I: Branching capacity and visiting probability. Preprint, arXiv:1611.10324

29. Qingsan Zhu. On the critical branching random walk II: Branching capacity and branching recurrence. Preprint, arXiv:1612.00161

30. Qingsan Zhu. On the critical branching random walk III: The critical dimension. Preprint, arXiv:1701.08917

31. Qingsan Zhu. Percolation for the vacant set of branching interlacements. In preparation.

32. Qingsan Zhu. Critical branching random walks, branching capacity and branching interlacements. PhD thesis, 2017.