

Weights of uniform spanning forests on nonunimodular transitive graphs*

Pengfei Tang[†]

Abstract

Considering the wired uniform spanning forest on a nonunimodular transitive graph, we show that almost surely each tree of the wired uniform spanning forest is light. More generally we study the tilted volumes for the trees in the wired uniform spanning forest.

Regarding the free uniform spanning forest, we consider several families of nonunimodular transitive graphs. We show that the free uniform spanning forest is the same as the wired one on Diestel–Leader graphs. For grandparent graphs, we show that the free uniform spanning forest is connected and has branching number bigger than one. We also show that each tree of the free uniform spanning forest is heavy and has branching number bigger than one on a free product of a nonunimodular transitive graph with one edge when the free uniform spanning forest is not the same as the wired.

Keywords: uniform spanning forests; nonunimodular transitive graphs; mass-transport principle.

MSC2020 subject classifications: 60K35.

Submitted to EJP on August 26, 2019, final version accepted on September 25, 2021.

1 Introduction

Let $G = (V(G), E(G))$ be a locally finite, connected infinite graph. The wired uniform spanning forest (WUSF) and the free uniform spanning forest (FUSF) are weak limits of the uniform spanning tree measures on an exhaustion of the graph G , with wired and free boundary conditions respectively. The WUSF and FUSF can be disconnected but each component is an infinite tree. Pemantle [21] proved for \mathbb{Z}^d , the WUSF is the same as FUSF and he also showed that the WUSF(FUSF) on \mathbb{Z}^d is connected iff $d \leq 4$. Since then WUSF and FUSF have been extensively studied. Hutchcroft studied many geometric properties of WUSF in [11]. For example Hutchcroft gave the volume growth

*Pengfei Tang was partially supported by the Natural Science Foundation under grant DMS-1612363.

[†]Department of Mathematics, Indiana University, USA. Current affiliation: Department of Mathematical Sciences, Tel Aviv University, Israel. E-mail: pengfeitang@mail.tau.ac.il

dimension, the spectral dimension and the walk dimension for the trees in WUSF. His results are about a class of high-dimensional graphs and this class of graphs contains nonunimodular transitive graphs. In the present paper, we continue the study of the geometries of trees in WUSF and FUSF on a special family of graphs—nonunimodular transitive graphs. We are particularly interested in the geometry of trees in WUSF and FUSF with respect to the “level structure” of the underlying nonunimodular transitive graphs. For more background on uniform spanning forests see [3] or Chapter 4 and 10 of [18].

Let $\text{Aut}(G)$ denote the automorphism group of G . Suppose $\text{Aut}(G)$ has a nonunimodular closed subgroup $\Gamma \subset \text{Aut}(G)$ that acts transitively on G . In particular such G must be nonamenable [18, Proposition 8.14] and hence there are infinitely many components for WUSF on G [18, Corollary 10.27]. There is a unique left Haar measure $|\cdot|$ on Γ (up to a multiplicative constant). For each $x \in V(G)$, let $\Gamma_x := \{\gamma \in \Gamma : \gamma x = x\}$ denote the stabilizer of x and $m(x) := |\Gamma_x|$. We call this function $m : V \rightarrow (0, \infty)$ the **weight function** for (G, Γ) . For a cluster C , we define its weight $m(C) := \sum_{x \in C} m(x)$ and call C a Γ -**light** cluster or a Γ -**heavy** cluster according to $m(C) < \infty$ or $m(C) = \infty$ respectively. For simplicity we will just say C is a light (or heavy) cluster if Γ is well-understood from the context.

Although each component of the WUSF or FUSF on G is an infinite tree, it may happen that some tree is light and has branching number bigger than one. We prove that if there is nonunimodular subgroup $\Gamma \subset \text{Aut}(G)$ that acts transitively on G , then each tree of the WUSF on G is Γ -light a.s.

Theorem 1.1. *Let $\text{Aut}(G)$ denote the automorphism group of G . Suppose $\text{Aut}(G)$ has a nonunimodular closed subgroup $\Gamma \subset \text{Aut}(G)$ that acts transitively on G . Then each tree of the WUSF is Γ -light almost surely.*

From now on if G is a transitive graph and $\Gamma \subset \text{Aut}(G)$ is a closed nonunimodular subgroup that acts transitively on G , then we call (G, Γ) a **nonunimodular transitive pair**.

Benjamini, Lyons, Peres and Schramm gave several equivalent conditions for WUSF = FUSF; see [3, Theorem 7.3]. In particular, WUSF = FUSF for amenable transitive graphs. For the case WUSF \neq FUSF, we make the following conjecture.

Conjecture 1.2. *Suppose $\text{Aut}(G)$ has a nonunimodular closed subgroup $\Gamma \subset \text{Aut}(G)$ that acts transitively on G and WUSF \neq FUSF on G . Then each tree in the FUSF is Γ -heavy and has branching number bigger than one almost surely.*

Theorem 1.3. *Let G be a free product of a nonunimodular transitive graph with one edge such that WUSF \neq FUSF on G or a grandparent graph. Then each connected tree of the FUSF on G is heavy and has branching number bigger than one almost surely. In fact, when G is a grandparent graph, the FUSF on G is connected almost surely.*

Remark 1.4. Recently Pete and Timár [22] disproved the heaviness part of Conjecture 1.2. So it is now natural to ask the question—“which transitive nonunimodular graphs have the property that every tree in the FUSF is heavy?” Theorem 1.3, Remark 5.12 and Pete and Timár’s Theorem 1.1 and 1.3 in [22] suggest this question is a quantitative issue (see Section 6 of [22] for some related problems).

A more fundamental problem than the branching number part of Conjecture 1.2 is to show that every tree in FUSF has infinitely many ends on nonunimodular transitive graphs when FUSF \neq WUSF. This is the remaining case of Question 15.8 of [3]; the unimodular case has been answered positively in [12, 26].

The paper is organized as follows. In Section 2 we give some preliminaries—we review the definition of WUSF and FUSF, Wilson’s algorithm, nonunimodular transitive graphs and the tilted mass-transport principle. In particular, we review the “level structure”

for a nonunimodular transitive pair (G, Γ) . The level structure is induced by the weight function $m : V \rightarrow (0, \infty)$. Let $u \sim v$ denote that u and v are neighboring vertices in G . Let $L_n(x)$ be the set of vertices y such that $\frac{m(y)}{m(x)} \approx e^{nt_0}$, where $e^{t_0} := \max\{\frac{m(v)}{m(u)} : u \sim v\}$. See Subsection 2.3 for the precise definition of $L_n(x)$. We also review the so-called **tilted volumes** introduced by Hutchcroft [10]. For a set $C \subset V(G)$ and $x \in C$, the so-called tilted volumes of C are defined as: $|C|_{x,\lambda} := \sum_{y \in C} \frac{m(y)^\lambda}{m(x)^\lambda}$, where $\lambda \in \mathbb{R}$. The tilted volume can be viewed as a generalization of the size $|C|$ and weight $m(C)$: when $\lambda = 0$, $|C|_{x,\lambda}$ is just the size of C ; when $\lambda = 1$, $|C|_{x,\lambda}$ is just a normalization of $m(C)$.

In Section 3 we consider WUSF on a toy model $(\mathbb{T}_{b+1}, \Gamma_\xi)$, namely regular tree together with a subgroup of automorphisms that fixes an end. The WUSF on the toy model has close relations to critical percolation and Galton–Watson trees (Lemma 3.1). These relations and the tree structure of the underlying graph \mathbb{T}_{b+1} make the study of WUSF on the toy model relatively easy. Moreover the results for WUSF on the toy model shed light on further studies for WUSF on general nonunimodular transitive graphs (Proposition 3.8, Question 3.10, Proposition C.1 and Proposition C.9).

In Section 4 we study the geometry of the trees in WUSF with respect to the level structure of the nonunimodular transitive pair. We will consider quantities such as $\mathbb{E}[|T_x \cap L_n(x)|]$ and $\mathbb{P}[T_x \cap L_n(x) \neq \emptyset]$. Similar quantities have been studied for Bernoulli percolation in [10]. These quantities are not only interesting in themselves but also pave the way for proving Theorem 1.1. We will study when the tilted volumes of the trees in WUSF are finite (Proposition 4.11) and what the tail behaviors of the tilted volumes are (Proposition 4.24). In particular, Theorem 1.1 is a special case of Proposition 4.11. Some other objects like the past and the future of a vertex will also be studied in Section 4.

In Section 5 we consider FUSF on Diestel–Leader graphs and grandparent graphs. In Proposition 5.3 we show that FUSF on a Diestel–Leader graph is the same as WUSF. For the FUSF on a grandparent graph, we show it is just one tree (Proposition 5.8) and has branching number bigger than one (Proposition 5.9).

In Section 6 we consider FUSF on free products of nonunimodular transitive graphs with \mathbb{Z}_2 and prove Proposition 6.3. Theorem 1.3 then follows from Proposition 5.8, 5.9 and 6.3. A key ingredient of proving Proposition 6.3 is to compare the weight of a tree in the FUSF to a branching random walk and then use Biggins’ theorem [16].

In Section A of the appendix we give the details of the proof the lower bounds of (4.29) and (4.30) (The lower bounds are not used in the main part). In Section B of the appendix we give the proof of Proposition 4.15. We also provide more quantitative results for WUSF on the toy model in Section C of the appendix.

2 Preliminaries

2.1 Uniform spanning forests

If G is a finite connected graph, then it has finitely many spanning trees. The **uniform spanning tree** (UST) on G is the uniform measure on the set of spanning trees of G and denoted by $\text{UST}(G)$. The Aldous–Broder algorithm and Wilson’s algorithm are well-known methods to generate the UST on a finite graph G .

Suppose $G = (V, E)$ is a locally finite, connected infinite graph. An exhaustion of G is a sequence of finite connected subgraphs $G_n = (V_n, E_n)$ of G such that $V_n \subset V_{n+1}$ and $G = \bigcup G_n$. The weak limit of $\text{UST}(G_n)$ exists and is independent of the choice of the exhaustion; for example see [18, Section 10.1]. We call the weak limit of $\text{UST}(G_n)$ **free uniform spanning forest** (FUSF). From the graph G , first identify the vertices outside G_n to a single vertex, say z_n , and then remove loop-edges at z_n but keep multiple edges. The graph obtained in this way is denoted by G_n^W . We now also assume that G_n is the graph **induced** in G by V_n . Then the weak limit of $\text{UST}(G_n^W)$ exists and is independent

of the choice of such exhaustion; for example see [18, Section 10.1]. We call the weak limit of $\text{UST}(G_n^W)$ **wired uniform spanning forest** (WUSF).

Járai and Redig [13] also introduced a v -WUSF on G , which can be roughly understood as a wired spanning forest with v wired to ∞ . Suppose $v \in V(G_n)$ for every G_n in the above exhaustion and let \hat{G}_n be the graph obtained from G_n^W by identifying v and z_n . Then the v -WUSF on G is the weak limit of $\text{UST}(\hat{G}_n)$.

The connected component of v in the v -WUSF is finite almost surely if G is a transient transitive graph [17, Proposition 3.1 and Theorem 7.4].

2.2 Wilson’s algorithm

Next we review Wilson’s algorithm for generating WUSF on a transient graph G . This is called Wilson’s method rooted at infinity [3, Theorem 5.1]. For finite graphs or recurrent graphs see section 3 of [3] for more details.

For a path w on G that visits every vertex finitely many times, we can define a loop erasure of w , namely let $\text{LE}(w)$ be the self-avoiding path obtained by erasing the loops chronologically as they are created.

Suppose G is a transient graph. Fix an arbitrary ordering (v_1, v_2, \dots) of the vertices of G . Set $F_0 = \emptyset$. Given F_{n-1} for $n \geq 1$, we construct F_n as follows. If $v_n \in F_{n-1}$, then let $F_n = F_{n-1}$. Otherwise start a simple random walk from v_n and let τ_n denote the first hitting time of F_{n-1} . In particular, $\tau_n = \infty$ if the simple random walk never hits F_{n-1} . Let P_n denote the random walk path stopped at τ_n . Loop erase this random walk path and denote it by $\text{LE}(P_n)$. Set $F_n = F_{n-1} \cup \text{LE}(P_n)$. Finally set $F := \bigcup_n F_n$. Then F has the law of WUSF on G , and in particular, its law does not depend on the ordering one chose.

Let G be a transient network and let \mathfrak{F} be a sample of WUSF on G generated by using Wilson’s algorithm. Then for every edge e of \mathfrak{F} , there is a unique orientation such that e is crossed by the loop-erased random walk in that direction. The resulting oriented graph is called **oriented wired uniform spanning forest** and denoted by OWUSF. The OWUSF also does not depend on the ordering of the vertices. From OWUSF one can get WUSF by forgetting the orientation. It is also easy to see that in the OWUSF, every vertex has exactly one edge emanating from it.

Suppose G is a transient transitive graph. Then every tree in the WUSF on G will have only one end almost surely [17, Theorem 7.4]. Let $\tilde{\mathfrak{F}}$ be a sample of OWUSF on G and let \mathfrak{F} be the spanning forest obtained from $\tilde{\mathfrak{F}}$ by forgetting its orientation. Then \mathfrak{F} has the law of WUSF. For each vertex $x \in V(G)$, let T_x denote the connected component of x in \mathfrak{F} . Since T_x is one-ended almost surely, there is a unique infinite ray $\eta = (v_0, v_1, \dots)$ starting from $v_0 = x$ representing the unique end of T_x . Then the edge (v_0, v_1) is also the unique edge emanating from x in \mathfrak{F} .

Given a sample \mathfrak{F} of WUSF on G and $u \in V(G)$, we define **the future** of u to be the unique oriented ray starting from u and denote it by $\mathfrak{F}(u, \infty)$. We take the convention that $u \in \mathfrak{F}(u, \infty)$. We also define **the past** of u to be the subgraph of \mathfrak{F} spanned by those vertices $v \in \mathfrak{F}$ such that $u \in \mathfrak{F}(v, \infty)$ and denote it by $\mathfrak{P}(u)$.

Let \mathfrak{F}_v be a sample of v -WUSF on G . Then one can generate \mathfrak{F}_v by running Wilson’s algorithm rooted at infinity but starting with $F_0 = \{v\}$, i.e. the forest with a single vertex v and no edge. We can also orient $e \in \mathfrak{F}_v$ as the way it is crossed by the loop-erased random walk in the Wilson’s algorithm. Then each vertex but v has exactly one edge emanating from it. We can define the past and future for every vertex u in \mathfrak{F}_v according to this orientation and denote them by $\mathfrak{P}_v(u)$ and $\mathfrak{F}_v(u, \infty)$ respectively. In particular, the future of v in \mathfrak{F}_v is the single vertex v itself and the past of v is the connected component of v in \mathfrak{F}_v . Let \mathcal{T}_v denote the tree containing v in \mathfrak{F}_v . Most notation here coincides with the ones listed on page 15 of [11] for the reader’s convenience.

Given a general oriented forest F of G , if there is an oriented path from u to v in F ,

then u is said to be in the past of v and v is said to be in the future of u . Let $\text{past}_F(v)$ denote the past of v in the oriented forest F , namely, the set of vertices which lie in the past of v in F . One lemma we shall need is the following stochastic domination result.

Lemma 2.1 (Lemma 2.1 of [11]). *Let G be an infinite network and \mathfrak{F} be a sample of OWUSF on G . For each $v \in V(G)$, let \mathfrak{F}_v be an oriented v -WUSF of G . Suppose K is a finite set of vertices of G and define $F(K) := \bigcup_{u \in K} \mathfrak{F}(u, \infty)$ and $F_v(K) := \bigcup_{u \in K} \mathfrak{F}_v(u, \infty)$. Then for every $u \in K$ and every increasing event $\mathcal{A} \subset \{0, 1\}^E$ we have*

$$\mathbb{P}(\text{past}_{\mathfrak{F} \setminus F(K)}(u) \in \mathcal{A} \mid F(K)) \leq \mathbb{P}(\mathfrak{T}_u \in \mathcal{A}) \tag{2.1}$$

and similarly

$$\mathbb{P}(\text{past}_{\mathfrak{F}_v \setminus F_v(K)}(u) \in \mathcal{A} \mid F(K)) \leq \mathbb{P}(\mathfrak{T}_u \in \mathcal{A}). \tag{2.2}$$

This lemma is stated for OWUSF and we will often use it for WUSF because we can first sample an OWUSF and then get WUSF by forgetting the orientation of the OWUSF.

2.3 Nonunimodular transitive graphs and the tilted mass-transport principle

Next we recall the tilted mass-transport principle. We will restrict to transitive graphs. Suppose G is an infinite, locally finite connected graph and $\Gamma \subset \text{Aut}(G)$ is a subgroup of automorphisms that acts transitively on G . For $x, y \in V(G)$, let $|\Gamma_{xy}|$ denote the number of vertices in the set $\{\gamma x : \gamma \in \Gamma_x\}$, where Γ_x is the stabilizer of x . There is a unique (up to a multiplicative constant) nonzero left Haar measure $|\cdot|$ on Γ and we denote by $m(x) = |\Gamma_x|$ the Haar measure of the stabilizer Γ_x . Then a simple criterion for unimodularity of Γ is given as follows.

Proposition 2.2 (Trofimov [27]). *Suppose $\Gamma \subset \text{Aut}(G)$ acts transitively on G . Then Γ is unimodular if and only if for all $x, y \in V(G)$,*

$$|\Gamma_{xy}| = |\Gamma_{yx}|.$$

The grandparent graph and the Diestel–Leader graph $DL(q, r)$ with $q \neq r$ are typical examples of nonunimodular transitive graphs. For more examples see section 3 of [25].

The following lemma is well known (for a proof, see for example formula (1.28) and Lemma 1.29 in [29]):

Lemma 2.3. *Suppose $\Gamma \subset \text{Aut}(G)$ acts on $G = (V, E)$ transitively, then for all $x, y \in V$*

$$\frac{m(x)}{m(y)} = \frac{|\Gamma_{xy}|}{|\Gamma_{yx}|} = \frac{m(\gamma x)}{m(\gamma y)}, \forall \gamma \in \Gamma.$$

Given the group Γ , define the **modular function** $\Delta : V \times V \in [0, \infty]$ as follows:

$$\Delta(x, y) := \frac{m(y)}{m(x)}.$$

Then from the above lemma 2.3 we know Δ is a Γ diagonally invariant function, i.e. $\Delta(x, y) = \Delta(\gamma x, \gamma y), \forall \gamma \in \Gamma$. Another important property for the modular function is the **cocycle identity**:

$$\Delta(u, v)\Delta(v, w) = \Delta(u, w), \forall u, v, w \in V(G).$$

For more background on modular functions defined here see Section 2.1 of [10].

For a cluster K of G , a vertex $v \in K$ and a parameter $\lambda \in \mathbb{R}$, Hutchcroft [10] introduced the **tilted volume** as follows:

$$|K|_{v, \lambda} := \sum_{y \in K} \Delta(v, y)^\lambda.$$

The tilted mass-transport principle is a useful technique when dealing with nonunimodular transitive graphs. Actually the word ‘tilted’ can be omitted, and the mass-transport principle was defined without the word [2]. The ‘tilted mass-transport principle’ (TMTP) was first used in [10] for a different way of writing the mass-transport principle.

Proposition 2.4 (Proposition 2.2 of [10]). *With the same notations as in the above definition of modular function, suppose $F : V^2 \rightarrow [0, \infty]$ is invariant under the diagonal action of Γ . Then*

$$\sum_{x \in V} F(\rho, x) = \sum_{x \in V} F(x, \rho) \Delta(\rho, x).$$

We will often use the following form of the above tilted mass-transport principle. Suppose ω is a Γ -invariant bond percolation process on $G = (V, E)$. Suppose $f : V^2 \times \{0, 1\}^E \rightarrow [0, \infty]$ is Γ -invariant, that is,

$$f(x, y, \omega) = f(\gamma x, \gamma y, \gamma \omega), \forall x, y \in V, \gamma \in \Gamma.$$

Then applying the tilted mass-transport principle with $F(x, y) = \mathbb{E}[f(x, y, \omega)]$ one has

$$\mathbb{E} \left[\sum_{x \in V} f(\rho, x, \omega) \right] = \mathbb{E} \left[\sum_{x \in V} f(x, \rho, \omega) \Delta(\rho, x) \right]. \tag{2.3}$$

When the percolation ω is clear from the context, we will often write $f(x, y)$ instead of $f(x, y, \omega)$ in the above equation.

Next we recall some terminology from percolation theory. Suppose $G = (V, E)$ is a locally finite, connected graph. Let $2^E = \{0, 1\}^E$ be the collection of all subsets $\eta \subset E$ and let \mathcal{F}_E be the σ -field generated by sets of the form $\{\eta : e \in \eta\}$ where e runs over all edges in E . A bond percolation on G is a pair (\mathbf{P}, ω) , where ω is a random element in 2^E and \mathbf{P} is the law of ω . For simplicity sometimes we will just say ω is a bond percolation. The interested reader can refer to [18, Chapter 7 and 8] for more background on percolation theory. If the law \mathbf{P} is invariant under a subgroup Γ of automorphisms, then we call (\mathbf{P}, ω) a Γ -invariant percolation on G . In particular, WUSF and FUSF can be viewed as $\text{Aut}(G)$ -invariant percolation processes on an infinite graph G .

Suppose (\mathbf{P}, ω) is an automorphism-invariant percolation on G and \mathbf{E} is the corresponding expectation operator. An edge e is called open if $\omega(e) = 1$; otherwise it is called closed. For $x \in V$, the **cluster** of x is the connected component of x in the subgraph formed by open edges and denoted by C_x . In case ω has the law of WUSF or FUSF, we will also denote the connected component of x by T_x since each connected component is a tree.

Note that the modular function satisfies $\Delta(y, x) = \Delta(x, y)^{-1}$. An immediate application of the tilted mass-transport principle is the following observation at the beginning of section 3 of [10]:

Lemma 2.5. *Suppose G is an infinite, locally finite connected graph and Γ is a subgroup of automorphisms that acts transitively on G . Suppose (\mathbf{P}, ω) is a Γ -invariant percolation on G and x is an arbitrary vertex in G . Then the function $f(\lambda) := \mathbf{E}[|C_x|_{x, \lambda}]$ is symmetric about $\lambda = \frac{1}{2}$:*

$$\mathbf{E}[|C_x|_{x, \lambda}] = \mathbf{E}[|C_x|_{x, 1-\lambda}].$$

Note that f is also convex, whence f is decreasing on $(-\infty, \frac{1}{2}]$ and increasing on $[\frac{1}{2}, \infty)$.

Finally we take a look at the level structure of nonunimodular transitive graphs; see also Section 2.2 of [10].

Suppose G is an infinite, locally finite connected graph and Γ is a nonunimodular subgroup of automorphisms that acts transitively on G . For $x, y \in V(G)$, if x, y are neighbors then we write $x \sim y$.

Define

$$t_0 = t_0(G) := \max\{\log \Delta(x, y) : x, y \in V(G), x \sim y\}$$

For each $s \leq t$ and $v \in V$ we define the **slab**

$$S_{s,t}(v) := \{x \in V : s \leq \log \Delta(v, x) \leq t\}.$$

If $t \geq s + t_0$ and a simple path $\pi := (v_0, \dots, v_n)$ in G crosses the slab $S_{s,t}(v)$ in the sense that $\log \Delta(v, v_0) \leq s$ and $\log \Delta(v, v_n) \geq t$ then it must contain some vertex from the slab $S_{s,t}(v)$ by the definition of t_0 .

Now we introduce a construction which will help us when applying the tilted mass-transport principle. This construction comes from [25] and the setup here is borrowed from Section 2.2 of [10].

Fix an arbitrary vertex v_0 of G , let U_{v_0} be a uniform $[0, 1]$ random variable independent of the WUSF and v -WUSF we shall consider. For every other $v \in V(G)$, let

$$U_v := U_{v_0} - \frac{1}{t_0} \log \Delta(v_0, v) \pmod{1}.$$

Notice the law of the collection of random variables $U := \{U_v : v \in V(G)\}$ does not depend on the choice of v_0 .

Given the collection of random variables U , the **separating layers** are defined to be

$$L_n(v) := \{x \in V : (n + U_v - 1)t_0 \leq \log \Delta(v, x) \leq (n + U_v)t_0\}, n \in \mathbb{Z}.$$

We also call $L_n(v)$ the n -th slab relative to v . We will use \mathbb{E}, \mathbb{P} to denote the expectation operator and probability measure for the joint law of U and WUSF or v -WUSF on G . Note that the cocycle identity ($\Delta(x, v) \cdot \Delta(v, x) = 1$) implies that

$$x \in L_n(v) \Leftrightarrow -n + (1 - U_v) - 1 \leq \frac{1}{t_0} \log \Delta(x, v) \leq -n + 1 - U_v. \tag{2.4}$$

In particular since $1 - U_v$ has the same law as U_x (uniform on $[0, 1]$) we have

$$\mathbb{P}[x \in L_n(v)] = \mathbb{P}[v \in L_{-n}(x)]. \tag{2.5}$$

Suppose $f : V^2 \times \{0, 1\}^E \times [0, 1]^V \rightarrow [0, \infty]$ is invariant under the diagonal action of Γ , then like (2.3) we have the following form of tilted mass-transport principle (equation (2.2) in [10])

$$\mathbb{E} \left[\sum_{x \in V} f(\rho, x, \mathfrak{F}, U) \right] = \mathbb{E} \left[\sum_{x \in V} f(x, \rho, \mathfrak{F}, U) \Delta(\rho, x) \right], \tag{2.6}$$

where \mathfrak{F} is a sample of WUSF and $U := \{U_v : v \in V(G)\}$ is defined as above.

If we let $\{\rho \leftrightarrow x\}$ denote the event that ρ and x are in the same connected component in \mathfrak{F} and set $f(\rho, x, \mathfrak{F}, U) = 1_{\{\rho \leftrightarrow x, x \in L_n(\rho)\}}$, then

$$\begin{aligned} \mathbb{E}[|T_\rho \cap L_n(\rho)|] &\stackrel{(2.6)}{=} \sum_{x \in V} \mathbb{E}[1_{\{\rho \leftrightarrow \rho, \rho \in L_n(x)\}} \Delta(\rho, x)] = \sum_{x \in V} \mathbb{E}[1_{\{\rho \leftrightarrow \rho\}}] \mathbb{E}[1_{\{\rho \in L_n(x)\}}] \Delta(\rho, x) \\ &\stackrel{(2.5)}{=} \sum_{x \in V} \mathbb{E}[1_{\{\rho \leftrightarrow \rho\}}] \mathbb{E}[1_{\{x \in L_{-n}(\rho)\}}] \Delta(\rho, x) \\ &= \mathbb{E} \left[\sum_{x \in V} 1_{\{\rho \leftrightarrow \rho, x \in L_{-n}(\rho)\}} \Delta(\rho, x) \right] \asymp \exp(-t_0 n) \mathbb{E}[|T_\rho \cap L_{-n}(\rho)|], \tag{2.7} \end{aligned}$$

where the last equality holds up to a factor of $e^{\pm t_0}$.

The constants c_i in the present papers are always positive constants. Their values only depend on the underlying graph G and the subgroup Γ . The constants c_i may carry different values at different appearances. The notations we use in this paper are similar to the ones listed on page 15 of [11]. However since we use Γ to denote subgroup of automorphisms, we will use slightly different notations for the future of a vertex and the paths connecting two vertices in the WUSF and v -WUSF. We summarize the notations here for the reader's convenience.

- $\mathfrak{F}, \mathfrak{F}_v$ A sample of the WUSF and v -WUSF respectively.
- \mathfrak{F}^f A sample of the FUSF.
- $x \leftrightarrow y$ The events that x and y are in the same connected component of \mathfrak{F} .
- T_v, \mathfrak{T}_v The connected components of v in the WUSF and v -WUSF respectively.
- $\mathfrak{F}(x, \infty), \mathfrak{F}_v(x, \infty)$ The future of x in \mathfrak{F} and \mathfrak{F}_v respectively.
- $\mathfrak{P}(x), \mathfrak{P}_v(x)$ The past of x in \mathfrak{F} and \mathfrak{F}_v respectively.
- $d_G(x, y)$ The graph distance from x to y . When the underlying graph G is clear from the context, we often write it as $d(x, y)$.
- $\{X_k^x\}_{k \geq 0}$ A simple random walk on G starting from x . For $y \neq z$, we take X^y, X^z to be independent.
- σ_y^x The first visit time of y by a simple random walk starting from x .
- τ_{-n}^x The last visit time of $L_{-n}(x)$ by a simple random walk starting from x .
- \asymp This denotes an equality that holds up to positive multiplicative constants. More precisely, for two positive functions f, g on $(0, \infty)$, $f(R) \asymp g(R)$ means that there exists $R_0 > 0$ and $c_1, c_2 > 0$ such that $c_1 f(R) \leq g(R) \leq c_2 f(R)$ for all $R \geq R_0$, and the implicit constants c_1, c_2, R_0 only depend on the graph. Sometimes we also use $f(n) = \Theta(g(n))$ to denote that $f(n) \asymp g(n)$.
- \asymp_λ Similar to the above, but the implicit constants also depend on λ .
- \preceq, \preceq_λ and \succeq, \succeq_λ are defined similarly to \asymp, \asymp_λ .

3 A toy model

Let G be a regular tree \mathbb{T}_{b+1} of degree $b + 1$, where $b \geq 2$. Given an end ξ of G , let Γ_ξ be the subgroup of automorphisms that fixes this end ξ . One can check that Γ_ξ acts transitively on G and Γ_ξ is nonunimodular by Proposition 2.2. The **toy model** $(\mathbb{T}_{b+1}, \Gamma_\xi)$ is just this regular tree \mathbb{T}_{b+1} together with the subgroup Γ_ξ .

The FUSF on \mathbb{T}_{b+1} is trivial, namely it equals the tree itself almost surely. The WUSF on \mathbb{T}_{b+1} can be generated using Wilson's algorithm. The description of T_x in the following lemma is due to Häggström (1998) [8].

Lemma 3.1. Consider WUSF on the regular tree \mathbb{T}_{b+1} . Let x be a fixed vertex and T_x be the tree containing x in the WUSF.

Pick a ray (x_0, x_1, x_2, \dots) starting from $x_0 = x$ “uniformly”, i.e., $\mathbb{P}[(x_0, \dots, x_n) = \gamma] = \frac{1}{b^n - 1(b+1)}$ for any self-avoiding path γ started from x_0 with length n . Then the tree T_x has the law of the connected component of x in the union of this ray with an independent Bernoulli bond percolation with parameter $\frac{1}{b}$.

Proof. The proof given here comes from the second paragraph of Section 11 on [3, page 42] and we include it for readers’ convenience. The WUSF on \mathbb{T}_{b+1} can be generated using Wilson’s algorithm: For a vertex $x \in \mathbb{T}_{b+1}$, start a simple random walk from x on the tree \mathbb{T}_{b+1} , loop erase this random walk path chronologically to get a ray η starting from x . Let y_1, \dots, y_b denote the neighbors of x not on the ray η . For $i = 1, \dots, b$ start independent simple random walks from the vertex y_i and let $A_i := \{ \text{the simple random walk starting from } y_i \text{ hits } x \}$. Obviously given η , the events A_i are independent. If A_i occurs then put the edge connecting y_i and x to the WUSF; otherwise y_i, x will be in different components in the WUSF. On the event A_i , we add only the edge (y_i, x) to the WUSF and repeat the process for y_i . Obviously $\mathbb{P}(A_i | \eta) = \frac{1}{b}$. Thus the tree of x is the union of the ray η , and independent random trees attached to the vertices of η . For the root x , the random tree attached is a critical Galton–Watson tree with binomial progeny distribution $\text{Bin}(b, 1/b)$. For each other vertex on the ray η , the first generation of the random tree has binomial progeny distribution $\text{Bin}(b - 1, 1/b)$ while the subsequent generations have progeny distribution $\text{Bin}(b, 1/b)$. This analysis can be extended to give the whole WUSF easily.

From the above description, it is easy to see the random trees attached also have the law of clusters in an independent Bernoulli bond percolation with $p = \frac{1}{b}$. \square

Similarly one has the following description of \mathfrak{T}_x for the toy model.

Lemma 3.2. Let x be a fixed vertex of \mathbb{T}_{b+1} . Consider x -WUSF on the regular tree \mathbb{T}_{b+1} and the tree \mathfrak{T}_x that contains x . Then \mathfrak{T}_x has the law of the cluster C_x in an independent Bernoulli bond percolation with $p = \frac{1}{b}$.

3.1 Two point function and first moment for the toy model

The following is a simple application of the symmetry of \mathbb{T}_{b+1} and Lemma 3.1.

Proposition 3.3. Let \mathbb{P} denote the law of WUSF on the regular tree \mathbb{T}_{b+1} , where $b \geq 2$. Let $x, y \in V(G)$ be two arbitrary vertices and $\{x \leftrightarrow y\}$ denote the event that x, y are in the same tree of WUSF. Set $n = \text{dist}(x, y)$ to be the graph distance of x, y . Then

$$\mathbb{P}[x \leftrightarrow y] = \frac{1}{b^n} \left[1 + n \cdot \frac{b-1}{b+1} \right]. \tag{3.1}$$

In particular, the probability that the tree T_x intersects with a high slab $L_n(x)$ decays exponentially:

$$\mathbb{P}[T_x \cap L_n(x) \neq \emptyset] = \frac{1}{b^n} \left[1 + n \cdot \frac{b-1}{b+1} \right] \asymp n e^{-t_0 n}, n \geq 1.$$

Proof. Let $\eta_x = (x_0, x_1, x_2, \dots)$ be the ray starting from x representing the end ξ . Since the WUSF is invariant under the whole automorphism group of \mathbb{T}_{b+1} , and for $k \geq 1$ there are $(b+1)b^{k-1}$ vertices on \mathbb{T}_{b+1} with graph distance k to x , by symmetry $\mathbb{P}(x_k \in \mathfrak{F}(x, \infty)) = \frac{1}{(b+1)b^{k-1}}$.

Let u denote the highest vertex in the future of x , i.e., the vertex u is the unique one in $\mathfrak{F}(x, \infty)$ such that $\Delta(x, u) = \sup\{\Delta(x, v) : v \in \mathfrak{F}(x, \infty)\}$. Let A_k denote the event that

$\Delta(x, u) = b^k$ for $k \geq 0$. Let y denote the highest vertex in the past of u and $B_{k,k'}$ be the event that $A_k \cap \{\Delta(u, y) = b^{k'}\}$ for $k' \geq 0$. See Figure 7 for a typical $B_{k,k'}$.

For $k \geq 0$, the event A_k defined above is just the event $\{x_k \in \mathfrak{F}(x, \infty)\} \setminus \{x_{k+1} \in \mathfrak{F}(x, \infty)\}$. Hence for $k > 0$, $\mathbb{P}(A_k) = \mathbb{P}(x_k \in \mathfrak{F}(x, \infty)) - \mathbb{P}(x_{k+1} \in \mathfrak{F}(x, \infty)) = \frac{b-1}{b+1} \cdot \frac{1}{b^k}$. Combining with the fact that $\mathbb{P}(x_0 \in \mathfrak{F}(x, \infty)) = 1$, one has

$$\mathbb{P}[A_k] = \begin{cases} \frac{b}{b+1}, & k = 0 \\ \frac{b-1}{b+1} \cdot \frac{1}{b^k}, & k > 0. \end{cases}$$

In particular, we see that the future of x , $\mathfrak{F}(x, \infty) = (x_0, \dots, x_n, \dots)$ will eventually go down to the lower slabs in the sense that $\log \Delta(x_0, x_n) \rightarrow -\infty$ as n tends to infinity. In particular, this implies that for $n \geq 0$,

$$\mathbb{P}[T_x \cap L_{-n}(x) \neq \emptyset] = \mathbb{P}[\mathfrak{F}(x, \infty) \cap L_{-n}(x) \neq \emptyset] = 1.$$

By Lemma 3.1 one has

$$\mathbb{P}[B_{k,k'}] = \mathbb{P}[A_k] \cdot \frac{b-1}{b} \cdot \frac{1}{b^{k'}}, \quad \forall k, k' \geq 0.$$

The $n = 0$ case of (3.1) is trivial. We assume $n > 0$ in the following. By symmetry we can assume y is the unique vertex such that $\Delta(x, y) = b^n$ and $\text{dist}(x, y) = n$. Hence

$$\mathbb{P}[x \leftrightarrow y] = \sum_{k=n}^{\infty} \mathbb{P}[A_k] + \sum_{k=0}^{n-1} \sum_{k'=n-k}^{\infty} \mathbb{P}[B_{k,k'}].$$

Then simple calculation shows that $\mathbb{P}[x \leftrightarrow y] = \frac{1}{b^n} [1 + n \cdot \frac{b-1}{b+1}]$. □

Proposition 3.4. Consider WUSF on the toy model $(\mathbb{T}_{b+1}, \Gamma_\xi)$. Fix an arbitrary vertex x of \mathbb{T}_{b+1} . Then

$$\mathbb{E}[|T_x \cap L_n(x)|] \asymp \begin{cases} (n \vee 1)e^{-t_0 n} = \frac{n \vee 1}{b^n} & \text{if } n \geq 0 \\ |n| & \text{if } n < 0 \end{cases}. \tag{3.2}$$

Proof. Let $d(u, v)$ be the graph distance of u, v in \mathbb{T}_{b+1} . Fix $n \geq 0$. Note that for an integer $k \geq 0$,

$$|\{y: d(x, y) = n + 2k, \Delta(x, y) = b^n\}| = \begin{cases} 1 & \text{if } k = 0 \\ (b-1)b^{k-1} & \text{if } k \geq 1 \end{cases}.$$

Then by Proposition 3.3 one has for $n \geq 0$,

$$\begin{aligned} \mathbb{E}[|T_x \cap L_n(x)|] &= \sum_{y \in L_n(x)} \mathbb{P}[x \leftrightarrow y] \\ &= 1 \cdot \frac{1}{b^n} [1 + n \cdot \frac{b-1}{b+1}] + \sum_{k=1}^{\infty} (b-1)b^{k-1} \cdot \frac{1}{b^{n+2k}} [1 + (n+2k) \cdot \frac{b-1}{b+1}] \asymp \frac{n \vee 1}{b^n}. \end{aligned}$$

The $n < 0$ case is similar, just using the following observation instead

$$|\{y: d(x, y) = |n| + 2k, \Delta(x, y) = b^n\}| = \begin{cases} b^{|n|} & \text{if } k = 0 \\ (b-1)b^{|n|+k-1} & \text{if } k \geq 1 \end{cases}. \quad \square$$

Corollary 3.5. For the toy model $(\mathbb{T}_{b+1}, \Gamma_\xi)$, $\mathbb{E}[|T_x|_{x,\lambda}] < \infty$ if and only if $\lambda \in (0, 1)$.

Proof. If we decompose T_x according to its intersection with different slabs, we get

$$\mathbb{E}[|T_x|_{x,\lambda}] \asymp \sum_{n \in \mathbb{Z}} \mathbb{E}[|T_x \cap L_n(x)|] \cdot \exp(t_0 \lambda n). \tag{3.3}$$

Then (3.3) and (3.2) yields the conclusion. Also for $\lambda \in (0, 1)$, by (3.3) one has

$$\mathbb{E}[|T_x|_{x,\lambda}] \asymp 1 + \sum_{n=1}^{\infty} n(e^{-t_0 \lambda n} + e^{-t_0(1-\lambda)n}) \asymp \frac{1}{\lambda^2(1-\lambda)^2}. \tag{3.4} \quad \square$$

Proposition 3.6. *For the toy model $(\mathbb{T}_{b+1}, \Gamma_\xi)$, if $\lambda \leq 0$, then $|T_x|_{x,\lambda} = \infty$ a.s. If $\lambda > 0$, then $|T_x|_{x,\lambda} < \infty$ a.s. In particular, the $\lambda = 1$ case tells us that T_x is light almost surely.*

Proof. Let $\{X_n\}_{n \geq 0}$ be a simple random walk started at x . If we sample \mathfrak{F} starting with $\{X_n\}_{n \geq 0}$ and let $\mathfrak{F}(x, \infty) = (x_0, x_1, x_2, \dots)$ be future of x . Then from the proof of 3.3 one has that $\Delta(x_0, x_n) \rightarrow 0$ almost surely ($\mathbb{P}(\bigcup_{k \geq 0} A_k) = 1$). Thus for $\lambda \leq 0$, $|T_x|_{x,\lambda} \geq |\mathfrak{F}(x, \infty)|_{x,\lambda} = \sum_{n=0}^{\infty} \Delta(x_0, x_n)^\lambda = \infty$.

For $\lambda \in (0, 1)$, Corollary 3.5 then implies $|T_x|_{x,\lambda} < \infty$ almost surely.

For $\lambda > 0$, we can write

$$|T_x|_{x,\lambda} = \sum_{y \in T_x, \Delta(x,y) \geq 1} \Delta(x, y)^\lambda + \sum_{y \in T_x, \Delta(x,y) < 1} \Delta(x, y)^\lambda.$$

The first summation is finite almost surely since $|T_x|_{x, \frac{1}{2}} < \infty$ and then it is a summation over finitely many vertex. For $\lambda \geq 1$, the second summation is bounded above by $\sum_{y \in T_x, \Delta(x,y) < 1} \Delta(x, y)^{\frac{1}{2}} \leq |T_x|_{x, \frac{1}{2}} < \infty$. Thus for $\lambda \geq 1$, one also has $|T_x|_{x,\lambda} < \infty$ almost surely. \square

3.2 On the asymptotic behavior of $T_x \cap L_{-n}(x)$.

We have seen that $\mathbb{E}[|T_x \cap L_{-n}(x)|] \asymp n$ for the toy model for $n > 0$. A natural question is what can we say about the almost sure behavior of $|T_x \cap L_{-n}(x)|$. Since T_x is light almost surely, $T_x \cap L_n(x) = \emptyset$ a.s. if n is large. So when talking about almost sure behavior, we are only interested in the asymptotic behavior of $|T_x \cap L_{-n}(x)|$ when n tends to infinity.

Proposition 3.7. *Consider WUSF on the toy model $(\mathbb{T}_{b+1}, \Gamma_\xi)$. Let $x \in \mathbb{T}_{b+1}$ be an arbitrary vertex and T_x be its connected component in the WUSF. Then $\frac{|T_x \cap L_{-n}(x)|}{n}$ converges in distribution to a random variable Z and Z has Gamma distribution with density function $f(z) = Cz \exp(-\frac{2z}{1-1/b}) \mathbf{1}_{\{z > 0\}}$, where C is a normalizing constant. Moreover*

$$\frac{\log |T_x \cap L_{-n}(x)|}{\log n} \rightarrow 1 \quad \text{a.s.}$$

Proof. Let v be the last vertex on the future of x such that $\Delta(v, x) = 1$. In particular, if x is the highest point on its future, then $v = x$. Denote the path on the future of x from x to v by $\pi(x, v)$. Let N be the largest number k such that there is some vertex x_i in $\pi(x, v) \setminus \{v\}$ such that the bush at x_i intersects $L_{-k}(x)$. Since all the bushes are finite trees almost surely, N is finite almost surely.

For $n > N$, $|T_x \cap L_{-n}(x)| = 1 + Z_n$, where Z_n is the size of the n -th generation of a critical branching process with immigration. Indeed the 1 on the right hand side is the contribution of the future and Z_n is given as follows:

$$Z_0 = 0, Z_n = \sum_{j=1}^{Z_{n-1}} Y_{n,j} + I_n, n \geq 1,$$

where $Y_{i,j}$'s are i.i.d. random variables with the progeny distribution $\text{Bin}(b, \frac{1}{b})$ and the immigration I_i 's are i.i.d. random variables with distribution $\text{Bin}(b - 1, \frac{1}{b})$. In fact the immigration I_i is the number of children of the vertex on the future in the previous generation.

Since we are only interested in the normalized asymptotic behavior, it suffices to show $\frac{Z_n}{n}$ converges in distribution to a random variable with Gamma distribution. This is a classical result regarding critical branching process with immigration; for example see Theorem 3 in [19].

For the almost sure result, it suffices to show that $\frac{\log Z_n}{\log n} \rightarrow 1$ a.s., which is due to Theorem 1.1 of [28]. □

In view of the above proposition, one might ask whether $\frac{|T_x \cap L_{-n}(x)|}{n}$ converges almost surely to a random variable with Gamma distribution. However Proposition 3.7 together with the following proposition imply that the limit does not exist.

Proposition 3.8. *For the toy model, using the same notation as Proposition 3.7, one has*

$$\limsup_{n \rightarrow \infty} \frac{|T_x \cap L_{-n}(x)|}{n \log \log n} \asymp 1 \text{ a.s.}$$

Proof. Similar to the proof of Proposition 3.7, it suffices to show that

$$\limsup_{n \rightarrow \infty} \frac{Z_n}{n \log \log n} \leq 1 \text{ a.s.} \tag{3.5}$$

and

$$\limsup_{n \rightarrow \infty} \frac{Z_n}{n \log \log n} \geq 1 \text{ a.s.} \tag{3.6}$$

The inequality (3.5) is a direct consequence of Remark 2.2 of [28].

Next we show (3.6).

If $I_i = t > 0$, we write $GW_{i,1}, \dots, GW_{i,t}$ to be the descendant trees of these t people immigrated in generation i . We use $|GW_{i,j} \cap L_{-2^k}(x)|$ to denote the contribution of $GW_{i,j}$ to Z_{2^k} .

For a constant $c > 0$, define the events $A_k(c)$ for $k \geq 2$ as follows:

$$A_k(c) := \bigcup_{i=2^{k-1}}^{2^{k-1}+2^{k-2}} \{I_i \neq 0, |GW_{i,j} \cap L_{-2^k}(x)| \geq c \cdot 2^k \log k \text{ for some } 1 \leq j \leq I_i\}.$$

We claim that for some small enough constant c , $\sum_{k=2}^{\infty} \mathbb{P}[A_k(c)] = \infty$. Since the A_k 's are independent events, by the Borel–Cantelli lemma one has $\mathbb{P}[A_k \text{ i.o.}] = 1$. Notice on the event $A_k(c)$, $Z_{2^k} \geq c \cdot 2^k \log k$. Therefore (3.6) holds.

Now it remains to prove the claim that $\sum_{k=2}^{\infty} \mathbb{P}[A_k(c)] = \infty$ for some small constant $c > 0$.

For simplicity, write $B_{i,k}(c) = \{I_i \neq 0, |GW_{i,j} \cap L_{-2^k}(x)| \geq c \cdot 2^k \log k \text{ for some } 1 \leq j \leq I_i\}$.

For $i \in [2^{k-1}, 2^{k-1} + 2^{k-2}]$, $n = 2^k - i \in [2^{k-2}, 2^{k-1}]$. By Kolmogorov's estimate (see for example Theorem 12.7 of [18]) one has

$$\mathbb{P}[|GW_{i,j} \cap L_{-2^k}(x)| > 0] \geq \frac{1}{n} \geq \frac{1}{2^k}. \tag{3.7}$$

By the inequality (2.2) on page 588 and Theorem 3.3 in [20] and $\text{Var}(Y_{i,j}) = 1 - \frac{1}{b}$, there exists a constant c_1 that the conditional probability $P(c) := \mathbb{P}[|GW_{i,j} \cap L_{-2^k}(x)| > c \cdot 2^k \log k \mid |GW_{i,j} \cap L_{-2^k}(x)| > 0]$ satisfies

$$P(c) \geq \exp\left(-\frac{2}{1-1/b} \frac{c \cdot 2^k \log k}{n}\right) - 1.74c_1 \sqrt{\frac{\log n}{n}} \geq \frac{1}{k^{c \cdot c_2}}, \tag{3.8}$$

where $c_2 = \frac{8}{1-1/b}$ (since $\frac{2^k}{n} \leq 4$).

Therefore for $i \in [2^{k-1}, 2^{k-1} + 2^{k-2}]$

$$\mathbb{P}[B_{i,k}(c)] \geq \mathbb{P}[I_i = 1] \times \mathbb{P}[|GW_{i,j} \cap L_{-2^k}(x)| > c \cdot 2^k \log k] \stackrel{(3.7), (3.8)}{\succeq} \frac{1}{2^k \cdot k^{c \cdot c_2}}.$$

Using independence of $B_{i,k}(c)$ for different i 's, one has

$$\mathbb{P}[A_k(c)] = 1 - \prod_{i=2^{k-1}}^{2^{k-1}+2^{k-2}} [1 - \mathbb{P}[B_{i,k}(c)]] \succeq \frac{1}{k^{c \cdot c_2}},$$

whence $\sum_{k=2}^{\infty} \mathbb{P}[A_k(c)] = \infty$ for small enough constant $c > 0$. □

Remark 3.9. From the proof of Proposition 3.8 one actually obtains that

$$\limsup_{n \rightarrow \infty} \frac{Z_n}{n \log \log n} \asymp 1$$

for general critical branching process with immigration Z_n if the conditions in Theorem 1.1 of [28] are satisfied with $\delta = 1$ (to use Theorem 3.3 of [20]).

Proposition 3.7 and 3.8 implies that almost surely $\lim_{n \rightarrow \infty} \frac{|T_x \cap L_{-n}(x)|}{n}$ does not exist. Proposition 3.7 also implies that for the toy model, for every $\varepsilon > 0$, a.s. $|T_x \cap L_{-n}(x)| \succeq n^{1-\varepsilon}$. Can one improve this lower bound for the toy model?

Question 3.10. What is the asymptotic behavior of $|T_x \cap L_{-n}(x)|$ as $n \rightarrow \infty$ for WUSF on general nonunimodular transitive graphs?

4 The geometry of trees in WUSF with respect to the level structure

Throughout the section we will assume (G, Γ) is a nonunimodular transitive pair. We are interested in understanding the geometry of the tree T_x with respect to the level structure induced by the modular function. For example, we wonder whether the tree T_x is light. Moreover we are also interested in detailed information such as $\mathbb{P}[T_x \cap L_n(x) \neq \emptyset]$, $\mathbb{E}[|T_x \cap L_n(x)|^k]$, $\mathbb{E}[|T_x|_{x,\lambda}]$ and $\mathbb{P}[|T_x|_{x,\lambda} > R]$. Along the way similar quantities for \mathfrak{T}_x , $\mathfrak{P}(x)$ and $\mathfrak{F}(x, \infty)$ will also be studied.

The dependencies of the lemmas and propositions in the section can be summarized in Figure 1.

4.1 Simple random walk on nonunimodular transitive graphs

Wilson's algorithm is a very useful tool for studying WUSF. Here we begin with a simple lemma about simple random walk on a nonunimodular transitive graph.

Lemma 4.1. *Suppose $\{X_n\}_{n \geq 0}$ is a simple random walk on G and $\lambda \in \mathbb{R}$. Then one has*

$$\mathbb{E}[\Delta(X_0, X_1)^\lambda] < 1 \text{ if and only if } \lambda \in (0, 1).$$

In particular, if $\lambda \in (0, 1)$, then

$$\sum_{k=0}^{\infty} \mathbb{E}[\Delta(X_0, X_k)^\lambda] < \infty. \tag{4.1}$$

Moreover, one also has

$$\mathbb{E}[\log \Delta(X_0, X_1)] < 0.$$

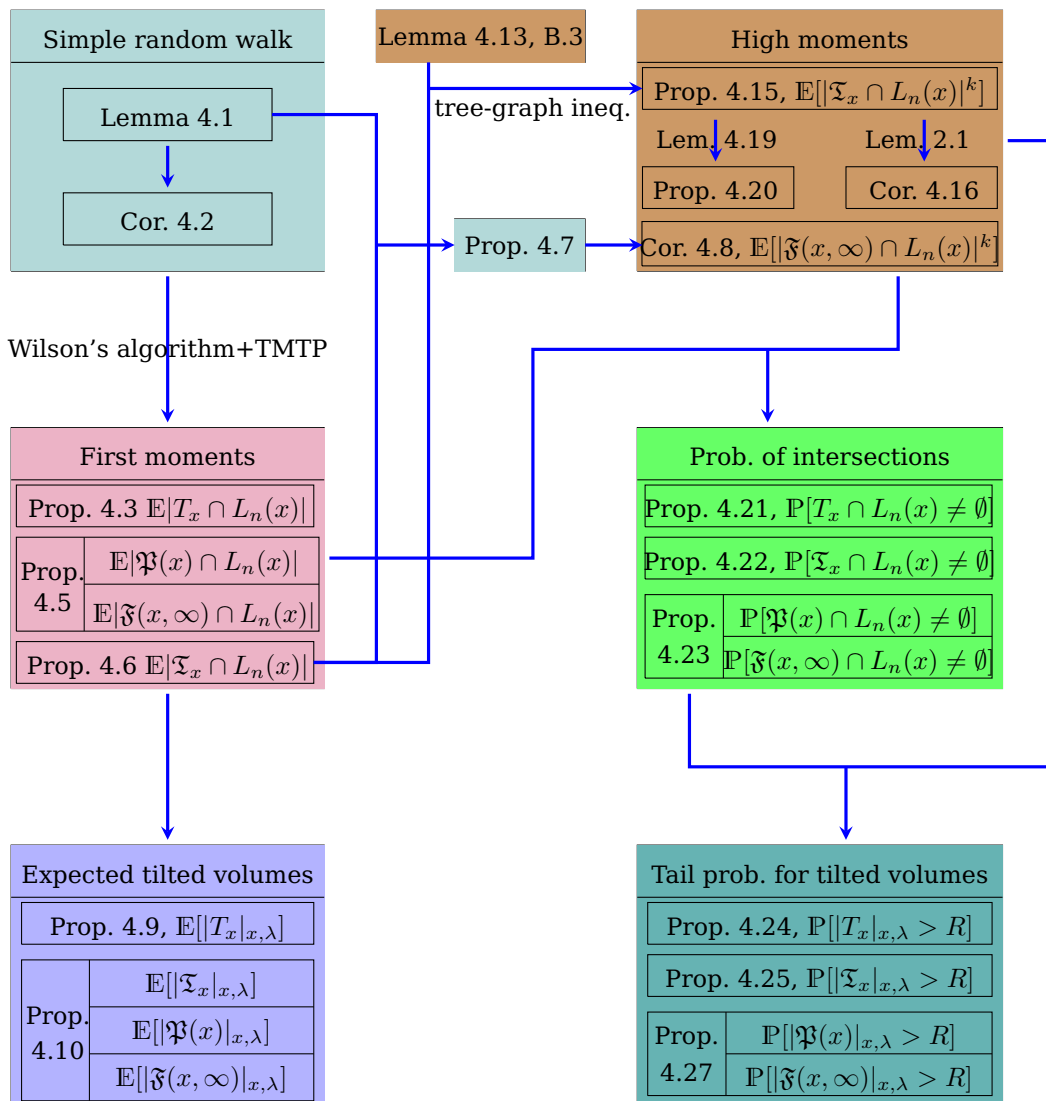


Figure 1: The dependencies among the lemmas and propositions in Section 4.

Proof. Let D denote the degree of the transitive graph G . For each $x \in V$, let $\{X_n^x\}_{n \geq 0}$ be a simple random walk starting at $X_0^x = x$.

Define a random function $f : V^2 \rightarrow [0, \infty]$ to be $f(x, y) = \mathbf{1}_{\{y=X_1^x\}} \Delta(x, y)^\lambda$, then the tilted mass-transport principle and the cocycle identity yield that

$$\begin{aligned} \mathbb{E}[\Delta(X_0, X_1)^\lambda \mid X_0 = x] &= \mathbb{E}\left[\sum_{y \in V} f(x, y)\right] \\ &= \mathbb{E}\left[\sum_{y \in V} f(y, x) \Delta(x, y)\right] = \mathbb{E}[\Delta(X_0, X_1)^{1-\lambda} \mid X_0 = x]. \end{aligned}$$

In particular, taking $\lambda = 0$ one has that

$$1 = \mathbb{E}\left[\sum_{y \in V} f(x, y)\right] = \mathbb{E}\left[\sum_{y \in V} f(y, x) \Delta(x, y)\right] = \sum_{y \sim x} \frac{1}{D} \Delta(x, y) = \mathbb{E}[\Delta(X_0, X_1) \mid X_0 = x].$$

Since the above two equations are true for all $x \in V(G)$, one has

$$\mathbb{E}[\Delta(X_0, X_1)^\lambda] = \mathbb{E}[\Delta(X_0, X_1)^{1-\lambda}]$$

and

$$1 = \mathbb{E}[\Delta(X_0, X_1)].$$

For $\lambda \in (0, 1)$, the function $x \mapsto x^\lambda$ is a strictly concave function on $[0, \infty]$ and $\Delta(X_0, X_1)$ takes more than one value with positive probability by nonunimodularity. By Jensen's inequality one has that

$$\mathbb{E}[\Delta(X_0, X_1)^\lambda] < \mathbb{E}[\Delta(X_0, X_1)]^\lambda = 1.$$

Using the cocycle identity and independence, one has $\mathbb{E}[\Delta(X_0, X_k)^\lambda] = (\mathbb{E}[\Delta(X_0, X_1)^\lambda])^k$, whence (4.1) follows from the fact that $\mathbb{E}[\Delta(X_0, X_1)^\lambda] < 1$.

For $\lambda > 1$ or $\lambda < 0$, the function $x \mapsto x^\lambda$ is a strictly convex function on $[0, \infty]$. By Jensen's inequality one has that

$$\mathbb{E}[\Delta(X_0, X_1)^\lambda] > \mathbb{E}[\Delta(X_0, X_1)]^\lambda = 1.$$

Since $x \mapsto \log x$ is strictly concave on $(0, \infty)$, $\mathbb{E}[\log \Delta(X_0, X_1)] < \log \mathbb{E}[\Delta(X_0, X_1)] = 0$. \square

Lemma 4.1 implies that simple random walk on nonunimodular transitive graphs has a drift towards the lower slabs.

Corollary 4.2. Fix $x \in V(G)$ and suppose $\{X_n^x\}_{n \geq 0}$ is a simple random walk on G starting at x . Let $n \vee 1 = \max\{1, n\}$ denote the maximum of n and 1. Then

$$\mathbb{E} \left[\sum_{k=0}^{\infty} \mathbf{1}_{\{X_k^x \in L_{-n}(x)\}} \right] \asymp 1 \quad \text{and} \quad \mathbb{E} \left[\sum_{k=0}^{\infty} k \cdot \mathbf{1}_{\{X_k^x \in L_{-n}(x)\}} \right] \asymp n \vee 1. \quad (4.2)$$

For an integer $n \geq 0$, let $\tau_{-n}^x := \sup\{k : X_k \in L_{-n}(x)\}$ be the last visit time of the slab $L_{-n}(x)$. Then there exist constants $c_1, c_2, c_3 > 0$ such that

$$\mathbb{P}[\tau_{-n}^x \geq R] \leq c_1 \exp(-c_2 R), \quad \forall R \geq c_3(n \vee 1). \quad (4.3)$$

Proof. By the cocycle identity $\log \Delta(X_0^x, X_k^x) = \sum_{j=1}^k \log \Delta(X_{j-1}^x, X_j^x)$. By transitivity of the underlying graph G , $\log \Delta(X_{j-1}^x, X_j^x)$ are independent, identically distributed random variables. The law of large numbers and Lemma 4.1 then imply that

$$\frac{1}{k} \log \Delta(X_0^x, X_k^x) = \frac{1}{k} \sum_{j=1}^k \log \Delta(X_{j-1}^x, X_j^x) \rightarrow \mathbb{E}[\log \Delta(X_0^x, X_1^x)] < 0 \text{ a.s.} \quad (4.4)$$

Also there exists a constant $c \in (0, 1)$ such that

$$\mathbb{P}[\log \Delta(X_0^x, X_1^x) = -t_0, \log \Delta(X_0^x, X_k^x) \leq -2t_0, \forall k \geq 2] \geq c. \quad (4.5)$$

Write $N_{-n} = \sum_{k=0}^{\infty} \mathbf{1}_{\{X_k^x \in L_{-n}(x)\}}$. Each time a simple random walk visits $L_n(x)$, with probability bounded below by some positive constant the simple random walk will leave $L_n(x)$ and never visit it again. Then the strong Markov property implies that

$$\mathbb{P}[N_{-n} \geq k] \leq (1 - c)^{k-1}, \quad \forall k \geq 1. \quad (4.6)$$

Hence the upper bound in the first part of (4.2) holds.

Equation (4.4) also implies that $\tau_{-n}^x < \infty$ almost surely. The lower bound for (4.2) are straightforward since $\sum_{k=0}^{\infty} \mathbf{1}_{\{X_k^x \in L_{-n}(x)\}} \geq 1$ and $\sum_{k=0}^{\infty} k \cdot \mathbf{1}_{\{X_k^x \in L_{-n}(x)\}} \geq n - 1$ almost surely.

For $k \geq R \geq \frac{2(n+1)t_0}{-\mathbb{E}[\log \Delta(X_0, X_1)]}$, one has $-\frac{(n+1)t_0}{k} \geq -\frac{(n+1)t_0}{R} \geq \frac{1}{2}\mathbb{E}[\log \Delta(X_0, X_1)] > \mathbb{E}[\log \Delta(X_0, X_1)]$. Therefore for such R

$$\begin{aligned} \mathbb{P}[\tau_{-n}^x \geq R] &= \mathbb{P}[\cup_{k \geq R} \{X_k^x \in L_{-n}(x)\}] \leq \sum_{k \geq R} \mathbb{P}[X_k^x \in L_{-n}(x)] \\ &\leq \sum_{k \geq R} \mathbb{P}[\log \Delta(X_0^x, X_k^x) \geq (-n-1)t_0] \\ &\leq \sum_{k \geq R} \mathbb{P}\left[\frac{1}{k} \log \Delta(X_0^x, X_k^x) \geq -\frac{(n+1)t_0}{k}\right] \\ &\leq \sum_{k \geq R} \mathbb{P}\left[\frac{1}{k} \log \Delta(X_0^x, X_k^x) \geq \frac{1}{2}\mathbb{E}[\log \Delta(X_0, X_1)]\right]. \end{aligned}$$

By large deviation principle (e.g. [5, Theorem 2.7.7]) we know there exists constants $c_4, c_5 > 0$ such that

$$\mathbb{P}\left[\frac{1}{k} \log \Delta(X_0^x, X_k^x) \geq \frac{1}{2}\mathbb{E}[\log \Delta(X_0, X_1)]\right] \leq c_4 \exp(-c_5 k), \quad \forall k \geq 1. \quad (4.7)$$

Hence taking $c_3 = \frac{4t_0}{-\mathbb{E}[\log \Delta(X_0, X_1)]}$, then for $n \geq 0$ and $R \geq c_3(n \vee 1) \geq \frac{2(n+1)t_0}{-\mathbb{E}[\log \Delta(X_0, X_1)]}$ one has

$$\mathbb{P}[\tau_{-n}^x \geq R] \leq \sum_{k=c_3(n \vee 1)}^{\infty} c_4 \exp(-c_5 k) = \frac{c_4}{1 - \exp(-c_5)} \exp(-c_5 c_3(n \vee 1)). \quad (4.8)$$

Then (4.3) holds if one takes $c_3 = \frac{4t_0}{-\mathbb{E}[\log \Delta(X_0, X_1)]}$, $c_1 = \frac{c_4}{1 - \exp(-c_5)}$ and $c_2 = c_5 c_3$.

Notice that for $k \geq 1, j \geq c_3(n \vee 1)$, there exists constants $c_6, c_7 > 0$ such that

$$\mathbb{P}[N_{-n} = k, \tau_{-n}^x = j] \leq \min\{(1-c)^{k-1}, c_1 \exp(-c_2 j)\} \leq c_6 \exp(-c_7(k+j)). \quad (4.9)$$

Hence

$$\begin{aligned} \sum_{k=1}^{\infty} \sum_{j=c_3(n \vee 1)}^{\infty} k \cdot j \cdot \mathbb{P}[N_{-n} = k, \tau_{-n}^x = j] &\leq c_6 \sum_{k=1}^{\infty} k \exp(-c_7 k) \sum_{j=1}^{\infty} j \exp(-c_7 j) \\ &= \frac{c_6 \exp(-2c_7)}{(1 - \exp(-c_7))^4} < \infty. \end{aligned} \quad (4.10)$$

Note

$$\begin{aligned} \sum_{k=1}^{\infty} \sum_{j=1}^{c_3(n \vee 1)} k \cdot j \cdot \mathbb{P}[N_{-n} = k, \tau_{-n}^x = j] &\leq \sum_{k=1}^{\infty} \sum_{j=1}^{c_3(n \vee 1)} k \cdot c_3(n \vee 1) \cdot \mathbb{P}[N_{-n} = k, \tau_{-n}^x = j] \\ &\leq c_3(n \vee 1) \sum_{k=1}^{\infty} k \mathbb{P}[N_{-n} = k] \\ &\leq c_3(n \vee 1) \sum_{k=1}^{\infty} k(1-c)^{k-1} = \frac{c_3(n \vee 1)}{c^2}. \end{aligned} \quad (4.11)$$

Combining (4.10) and (4.11) one has that there exists a constant $c_8 > 0$ such that

$$\mathbb{E}[N_{-n} \tau_{-n}^x] = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} k \cdot j \cdot \mathbb{P}[N_{-n} = k, \tau_{-n}^x = j] \leq c_8(n \vee 1).$$

Note that $\sum_{k=0}^{\infty} k \cdot \mathbf{1}_{\{X_k^x \in L_{-n}(x)\}} \leq N_{-n} \tau_{-n}^x$, whence the upper bound in the second part of (4.2) holds. \square

4.2 First moment of intersections with a slab and the expected tilted volumes

We summarize the expectations of $\mathbb{E}[|\star \cap L_n(x)|]$ for $\star \in \{T_x, \mathfrak{T}_x, \mathbb{P}(x), \mathfrak{F}(x, \infty)\}$ in Table 1. Based on these first moments we can also derive estimates on the tilted volumes.

Table 1: First moments for the intersections with a slab

Quantities ($n \in \mathbb{Z}$)	Results	Position
$\mathbb{E}[T_x \cap L_n(x)]$	$\begin{cases} \asymp (n \vee 1)e^{-t_0 n} & \text{if } n \geq 0 \\ \asymp n & \text{if } n < 0 \end{cases}$	Proposition 4.3
$\mathbb{E}[\mathfrak{P}(x) \cap L_n(x)]$	$\begin{cases} \asymp e^{-t_0 n} & \text{if } n \geq 0 \\ \asymp 1 & \text{if } n < 0 \end{cases}$	Proposition 4.5
$\mathbb{E}[\mathfrak{F}(x, \infty) \cap L_n(x)]$	$\begin{cases} \asymp e^{-t_0 n} & \text{if } n \geq 0 \\ \asymp 1 & \text{if } n < 0 \end{cases}$	Proposition 4.5
$\mathbb{E}[\mathfrak{T}_x \cap L_n(x)]$	$\begin{cases} \asymp e^{-t_0 n} & \text{if } n \geq 0 \\ \asymp 1 & \text{if } n < 0 \end{cases}$	Proposition 4.6

Proposition 4.3. *There exist constants $c_1 > 0, c_2 > 1$ such that for every $n \geq 1$*

$$1 \leq \mathbb{E}[|T_x \cap L_0(x)|] \leq c_2 \text{ and } c_1 n \leq \mathbb{E}[|T_x \cap L_{-n}(x)|] \leq c_2 n \tag{4.12}$$

and

$$c_1 n \exp(-t_0(n + 1)) \leq \mathbb{E}[|T_x \cap L_n(x)|] \leq c_2 n \exp(-t_0(n - 1)), \forall n \geq 1 \tag{4.13}$$

The following lemma is a simple application of Wilson’s algorithm. The inequality (4.14) comes from the proof of Theorem 13.1 in [3] where expected quadratic growth with respect to extrinsic graph metric was proved. Also a version of (4.15) is used in the proof of Lemma 6.6 in [11] and the Lemma 6.6 is closely related to the quadratic growth with respect to intrinsic graph metric. See Theorem 1.3, Corollary 6.4 and 6.12 in [11] for the quadratic growth with respect to intrinsic graph metric.

Lemma 4.4. *Recall that $\{x \leftrightarrow y\}$ denotes the event that x, y are in the same connected component of the WUSF sample \mathfrak{F} . Then*

$$\mathbb{P}[x \leftrightarrow y] \leq \sum_{m=0}^{\infty} (m + 1) \mathbb{P}[X_m^x = y], \tag{4.14}$$

where $\{X_m^x\}_{m \geq 0}$ is a simple random walk on G started from x .

Let $y \in \mathfrak{T}_x$ denote the event that y lies in the connected component of x in the x -WUSF sample \mathfrak{F}_x . Then

$$\mathbb{P}[y \in \mathfrak{T}_x] = \mathbb{P}[x \in \mathfrak{T}_y] \leq \sum_{m=0}^{\infty} \mathbb{P}[X_m^x = y]. \tag{4.15}$$

Proof. From the reversibility of simple random walk and the regularity of G , one has that for any vertices x, y and $k \leq m$,

$$\sum_{z \in V} \mathbb{P}[X_k^x = z] \mathbb{P}[X_{m-k}^y = z] = \sum_{z \in V} \mathbb{P}[X_k^x = z] \mathbb{P}[X_{m-k}^z = y] = \mathbb{P}[X_m^x = y].$$

By Wilson’s algorithm,

$$\begin{aligned} \mathbb{P}[x \leftrightarrow y] &\leq \mathbb{P}[X^x \text{ intersects } X^y] \\ &\leq \sum_{z \in V} \sum_{m=0}^{\infty} \sum_{k=0}^m \mathbb{P}[X_k^x = z] \mathbb{P}[X_{m-k}^y = z] \\ &= \sum_{m=0}^{\infty} \sum_{k=0}^m \mathbb{P}[X_m^x = y] = \sum_{m=0}^{\infty} (m+1) \mathbb{P}[X_m^x = y]. \end{aligned}$$

If we use Wilson’s algorithm to sample \mathfrak{F}_x starting with the simple random walk X^y , then we have that

$$\mathbb{P}[y \in \mathfrak{F}_x] = \mathbb{P}[\sigma_x^y < \infty]$$

where σ_x^y denotes the first visit time of x by the simple random walk X^y . The reversibility of simple random walk on G and the regularity of G then imply that $\mathbb{P}[y \in \mathfrak{F}_x] = \mathbb{P}[\sigma_x^y < \infty] = \mathbb{P}[\sigma_y^x < \infty] = \mathbb{P}[x \in \mathfrak{F}_y]$.

Therefore one has that

$$\mathbb{P}[y \in \mathfrak{F}_x] = \mathbb{P}[x \in \mathfrak{F}_y] = \mathbb{P}[\sigma_y^x < \infty] \leq \sum_{m=0}^{\infty} \mathbb{P}[X_m^x = y]. \quad \square$$

Proof of Proposition 4.3. We first establish the upper bounds in (4.12).

For $n \geq 0$, one has that

$$\begin{aligned} \mathbb{E}[|T_x \cap L_{-n}(x)|] &= \mathbb{E}\left[\sum_{y \in V} \mathbf{1}_{\{y \in L_{-n}(x)\}} \cdot \mathbf{1}_{\{y \leftrightarrow x\}}\right] \\ &\stackrel{(4.14)}{\leq} \sum_{y \in V} \mathbb{E}[\mathbf{1}_{\{y \in L_{-n}(x)\}}] \mathbb{E}\left[\sum_{m=0}^{\infty} (m+1) \mathbf{1}_{\{X_m^x = y\}}\right] \\ &= \mathbb{E}\left[\sum_{m=0}^{\infty} (m+1) \mathbf{1}_{\{X_m^x \in L_{-n}(x)\}}\right] \\ &\stackrel{(4.2)}{\leq} c_0 + c_0(n \vee 1). \end{aligned} \tag{4.16}$$

Taking $c_2 = 2c_0$ one has the upper bounds in (4.12). Since $x \in T_x \cap L_0(x)$, the lower bound $\mathbb{E}[|T_x \cap L_0(x)|] \geq 1$ is trivial.

Next we prove the inequality $\mathbb{E}[|T_x \cap L_{-n}(x)|] \geq c_1 n$ for $n \geq 1$ using a similar strategy as the one used in the proof of Theorem 13.1 of [3].

Denote by $g(u, v) = \sum_{m=0}^{\infty} \mathbb{P}[X_m^u = v]$ the Green function for simple random walk on G . Since Γ is a closed subgroup of automorphisms that acts transitively on G , G is nonamenable (Proposition 8.14 of [18]) and hence the spectral radius $\rho(G) < 1$. By the Cauchy–Schwarz inequality $p_n(u, v) = \mathbb{P}[X_n^u = v] = \langle P^n \mathbf{1}_u, \mathbf{1}_v \rangle \leq \rho(G)^n$ and hence that

$$g(u, v) \leq \sum_{n \geq d(u,v)} \rho(G)^n \leq \frac{1}{1 - \rho(G)} \rho(G)^{d(u,v)},$$

where $d(u, v)$ is the graph distance of u, v in G .

We use Wilson’s algorithm to generate the WUSF sample \mathfrak{F} by starting with a simple random walk. Then the future of x is a subset of X^x . By (4.4) one has that $1 \leq |\mathfrak{F}(x, \infty) \cap L_{-n}(x)| < \infty$ for every $n \geq 0$. Also

$$\mathbb{P}[y \leftrightarrow x] = \mathbb{P}[X^y \cap \mathfrak{F}(x, \infty) \neq \emptyset].$$

Let $B(x, n) := \{y \in V(G) : d(x, y) \leq n\}$ denote the ball of radius n centered at x . We shall show that there exists a constant $c_1 > 0$ such that for any set $S \subset B(x, n)$ that

contains exactly one vertex at distance k from x for $1 \leq k \leq n$, the following inequality holds.

$$\mathbb{E} \left[\sum_{y \in L_{-n}(x)} \mathbf{1}_{\{L_y(S) > 0\}} \right] \geq c_1 n, \tag{4.17}$$

where $L_y(S) := \sum_{w \in S} \sum_{m \geq 0} \mathbf{1}_{\{X_m^y = w\}}$ is the total occupation time of S by X^y .

If $S \subset \mathfrak{F}(x, \infty) \cap B(x, n)$ and S contains exactly one vertex from each sphere $\partial B(x, k)$ for $k = 1, \dots, n$, then

$$\mathbb{P}[y \leftrightarrow x] \geq \mathbb{P}[L_y(S) > 0]. \tag{4.18}$$

Since simple random walk X^w on G visits $S_{s,t}(w)$ for every $t \leq 0, t - s \geq t_0$ almost surely, one has that $\sum_{y \in L_{-n}(x)} g(w, y) \geq 1, \forall w \in S$. Thus $\sum_{y \in L_{-n}(x), w \in S} g(w, y) \geq |S| = n$. Hence for $n \geq 1$,

$$\begin{aligned} \mathbb{E} \left[\sum_{y \in L_{-n}(x)} L_y(S) \right] &= \mathbb{E} \left[\sum_{y \in L_{-n}(x)} \sum_{w \in S} g(y, w) \right] \\ &= \mathbb{E} \left[\sum_{y \in L_{-n}(x)} \sum_{w \in S} g(w, y) \right] \\ &\geq n. \end{aligned} \tag{4.19}$$

For any $w, v \in S, d(w, v) \geq |d(w, x) - d(v, x)|$, whence for any $w \in S$ one has

$$\mathbb{E}[L_w(S)] \leq \sum_{j \geq 0} 2C\rho(G)^{j-1} =: c_3.$$

The Markov property implies that $\mathbb{E}[L_y(S) | L_y(S) > 0] \leq \max_{w \in S} \mathbb{E}[L_w(S)] \leq c_3$. Hence

$$\mathbb{E}[L_y(S)] \leq c_3 \mathbb{P}[L_y(S) > 0].$$

This together with (4.19) implies (4.17) with $c_1 = \frac{1}{c_3}$.

Thus for $n \geq 1$ by conditioning on $\mathfrak{F}(x, \infty)$ one obtains the lower bounds in (4.12):

$$\begin{aligned} \mathbb{E}[|T_x \cap L_{-n}(x)|] &= \sum_{y \in V} \mathbb{P}[y \in L_{-n}(x)] \mathbb{P}[y \leftrightarrow x] \\ &\stackrel{(4.18)}{\geq} \sum_{y \in V} \mathbb{P}[y \in L_{-n}(x)] \mathbb{P}[L_y(S) > 0] \\ &= \sum_{y \in V} \mathbb{E}[\mathbf{1}_{\{y \in L_{-n}(x)\}}] \cdot \mathbb{E}[\mathbf{1}_{\{L_y(S) > 0\}}] \\ &= \sum_{y \in V} \mathbb{E}[\mathbf{1}_{\{y \in L_{-n}(x)\}} \mathbf{1}_{\{L_y(S) > 0\}}] \text{ by independence} \\ &\stackrel{(4.17)}{\geq} c_1 n. \end{aligned} \tag{4.20}$$

The inequalities (4.13) follow from (4.12) and (2.7). □

Proposition 4.5. *There exists a constant $c_1 > 0$ such that for every $n \geq 0$,*

$$1 \leq \mathbb{E}[|\mathfrak{F}(x, \infty) \cap L_{-n}(x)|] \leq c_1. \tag{4.21}$$

There exists constants $c_2, c_3 > 0$ such that for every $n \geq 1$,

$$c_2 \leq \mathbb{E}[|\mathfrak{P}(x) \cap L_{-n}(x)|] \leq c_3 e^{t_0}. \tag{4.22}$$

Moreover one also has that

$$c_2 \exp(-t_0(n+1)) \leq \mathbb{E}[|\mathfrak{F}(x, \infty) \cap L_n(x)|] \leq c_3 \exp(-t_0 n) \tag{4.23}$$

and

$$\exp(-t_0(n+1)) \leq \mathbb{E}[|\mathfrak{P}(x) \cap L_n(x)|] \leq c_1 \exp(-t_0(n-1)). \tag{4.24}$$

Proof. We will prove (4.21), the upper bound in (4.23) and the lower bound in (4.22). The rest will follow from these inequalities and the tilted mass-transport principle.

If we sample \mathfrak{F} starting with X^x , then $\mathfrak{F}(x, \infty) \subset \{X_k^x : k \geq 0\}$. Write $\mathfrak{F}(x, \infty) = (v_0 = x, v_1, v_2, \dots)$. By (4.4) one has that $\log \Delta(X_0^x, X_k^x) \rightarrow -\infty$ as $k \rightarrow \infty$, in particular $\log \Delta(v_0, v_k) \rightarrow -\infty$. Thus $1 \leq |\mathfrak{F}(x, \infty) \cap L_{-n}(x)|$ for every $n \geq 0$ almost surely, whence one has the lower bound in (4.21): $1 \leq \mathbb{E}[|\mathfrak{F}(x, \infty) \cap L_{-n}(x)|]$.

The upper bound in (4.21) also follows from the above inclusion $\mathfrak{F}(x, \infty) \subset \{X_k^x : k \geq 0\}$.

$$\mathbb{E}[|\mathfrak{F}(x, \infty) \cap L_{-n}(x)|] \leq \mathbb{E}\left[\sum_{m=0}^{\infty} \mathbf{1}_{\{X_m^x \in L_{-n}(x)\}}\right] \stackrel{(4.2)}{\leq} c_0 = c_1.$$

Next we prove the upper bound in (4.23). Recall that σ_x^y denotes the first visit time of y by a simple random walk starting from x . By the reversibility of simple random walk and the transitivity of G , one has $\mathbb{P}[\sigma_y^x < \infty] = \mathbb{P}[\sigma_x^y < \infty]$. Hence for $n \geq 1$,

$$\begin{aligned} \mathbb{E}\left[\sum_{y \in V} \mathbf{1}_{\{y \in L_n(x), \sigma_y^x < \infty\}}\right] &\stackrel{\text{TMTP}}{=} \mathbb{E}\left[\sum_{y \in V} \mathbf{1}_{\{x \in L_n(y), \sigma_x^y < \infty\}} \Delta(x, y)\right] \\ &\asymp e^{-t_0 n} \mathbb{E}\left[\sum_{y \in L_{-n}(x)} \mathbf{1}_{\{\sigma_y^x < \infty\}}\right], \end{aligned} \tag{4.25}$$

where the last equality holds up a multiplicative constant $e^{\pm t_0}$.

Since $\mathbb{P}[y \in \mathfrak{F}(x, \infty)] \leq \mathbb{P}[\sigma_y^x < \infty]$, one has that

$$\begin{aligned} \mathbb{E}[|\mathfrak{F}(x, \infty) \cap L_n(x)|] &\leq \mathbb{E}\left[\sum_{y \in V} \mathbf{1}_{\{y \in L_n(x), \sigma_y^x < \infty\}}\right] \\ &\stackrel{(4.25)}{\leq} e^{t_0 - t_0 n} \mathbb{E}\left[\sum_{m=0}^{\infty} \mathbf{1}_{\{X_m^x \in L_{-n}(x)\}}\right] \leq c_0 e^{t_0 - t_0 n}. \end{aligned}$$

Taking $c_3 = c_0 e^{t_0}$ one has the upper bound in (4.23).

Finally we prove the lower bound in (4.22).

Let $y \in V$ and let X^y be a simple random walk started at y . Sample \mathfrak{F} starting with the simple random walk X^y . Let $\mathcal{A}(y, x)$ be the event that $\sigma_x^y < \infty$ and that the sets $\{X_m^y : 0 \leq m < \sigma_x^y\}$ and $\{X_m^y : m \geq \sigma_x^y\}$ are disjoint, so that $y \in \mathfrak{F}(x)$ on the event $\mathcal{A}(y, x)$ and hence

$$\mathbb{E}[|\mathfrak{F}(x) \cap L_{-n}(x)|] \geq \sum_{y \in V} \mathbb{P}[y \in L_{-n}(x)] \cdot \mathbb{P}[\mathcal{A}(y, x)].$$

Let Y^x be an independent simple random walk also started at x . We also use X^x to denote the set of vertices $\{X_m^x : m \geq 0\}$ visited by the random walk X^x and write similarly $Y_+^x := \{Y_m^x : m \geq 1\}$. By time-reversal one has that

$$\mathbb{P}[\mathcal{A}(y, x)] \geq \mathbb{P}[\sigma_y^x < \infty, X^x \cap Y_+^x = \emptyset].$$

Since X^x hits $L_{-n}(x)$ for every $n \geq 0$ almost surely, one has that

$$\mathbb{E}[|\mathfrak{F}(x) \cap L_{-n}(x)|] \geq \sum_{y \in V} \mathbb{P}[y \in L_{-n}(x)] \cdot \mathbb{P}[\sigma_y^x < \infty, X^x \cap Y_+^x = \emptyset] \stackrel{(4.2)}{\geq} \mathbb{P}[X^x \cap Y_+^x = \emptyset].$$

Since G is a nonamenable transitive graph, there exists a positive constant c_2 such that $\mathbb{P}[X^x \cap Y_+^x = \emptyset] \geq c_2$ (see [18, Theorem 10.24]). Thus we have the lower bound in (4.22).

Just like (2.7), the tilted mass-transport principle implies that for any $n \in \mathbb{Z}$,

$$\mathbb{E}[|\mathfrak{P}(x) \cap L_n(x)|] = \mathbb{E} \left[\sum_{y \in V} \mathbf{1}_{\{y \in L_{-n}(x)\}} \mathbf{1}_{\{y \in \mathfrak{F}(x, \infty)\}} \Delta(x, y) \right] \asymp e^{-t_0 n} \mathbb{E}[|\mathfrak{F}(x, \infty) \cap L_{-n}(x)|],$$

where the last equality holds up to a factor $e^{\pm t_0}$. This together with (4.21) implies (4.24); this together with the upper bound in (4.23) implies the upper bound in (4.22) and this together with the lower bound in (4.22) implies the lower bound in (4.23) \square

Next we give estimates of $\mathbb{E}[|\mathfrak{F}_x \cap L_n(x)|]$.

Proposition 4.6. *There exists positive constants c_1, c_2 such that for all $n \geq 0$,*

$$c_1 \leq \mathbb{E}[|\mathfrak{F}_x \cap L_{-n}(x)|] \leq c_2 \tag{4.26}$$

and

$$c_1 \exp(-t_0(n + 1)) \leq \mathbb{E}[|\mathfrak{F}_x \cap L_n(x)|] \leq c_2 \exp(-t_0(n - 1)). \tag{4.27}$$

Proof. As in the proof of Lemma 4.4, by Wilson’s algorithm and the reversibility of simple random walk one has that

$$\mathbb{P}[y \in \mathfrak{F}_x] = \mathbb{P}[\sigma_x^y < \infty] = \mathbb{P}[\sigma_y^x < \infty],$$

where σ_x^y is the first visit time of x by the simple random walk X^y .

Hence

$$\begin{aligned} \mathbb{E}[|\mathfrak{F}_x \cap L_{-n}(x)|] &= \mathbb{E} \left[\sum_{y \in V} \mathbf{1}_{\{y \in L_{-n}(x)\}} \cdot \mathbf{1}_{\{y \in \mathfrak{F}_x\}} \right] \\ &= \mathbb{E} \left[\sum_{y \in V} \mathbf{1}_{\{y \in L_{-n}(x)\}} \cdot \mathbf{1}_{\{\sigma_y^x < \infty\}} \right] \end{aligned} \tag{4.28}$$

Since simple random walk X^x visits $L_{-n}(x)$ for every $n \geq 0$ almost surely, one has $\sum_{y \in V} \mathbf{1}_{\{y \in L_{-n}(x)\}} \cdot \mathbf{1}_{\{\sigma_y^x < \infty\}} \geq 1$ a.s.. Then (4.28) implies that the lower bound in (4.26) holds with $c_1 = 1$.

On the other hand,

$$\mathbb{E}[|\mathfrak{F}_x \cap L_{-n}(x)|] \stackrel{(4.28)}{\leq} \mathbb{E} \left[\sum_{m=0}^{\infty} \mathbf{1}_{\{X_m^x \in L_{-n}(x)\}} \right] \stackrel{(4.2)}{\leq} c_0.$$

Hence the upper bound in (4.26) holds with $c_2 = c_0$.

By Lemma 4.4 one has $\mathbb{P}[y \in \mathfrak{F}_x] = \mathbb{P}[x \in \mathfrak{F}_y]$. Then the tilted mass-transport principle gives the relation between $\mathbb{E}[|\mathfrak{F}_x \cap L_{-n}(x)|]$ and $\mathbb{E}[|\mathfrak{F}_x \cap L_n(x)|]$, namely (4.25). This together with (4.26) yields (4.27). \square

For the intersection of a simple random walk trajectory with a slab, we have the following results.

Proposition 4.7. *Let X^x denote a simple random walk on the transitive graph G started from x . Then for $n \geq 0$ one has that*

$$\mathbb{P}[|\{X_m^x : m \geq 0\} \cap L_n(x)| \geq k] = e^{-\Theta(k) - t_0 n} \tag{4.29}$$

and

$$\mathbb{P}[|\{X_m^x : m \geq 0\} \cap L_{-n}(x)| \geq k] = e^{-\Theta(k)} \tag{4.30}$$

In particular, one has that

$$\mathbb{P}[\{X_m^x : m \geq 0\} \cap L_n(x) \neq \emptyset] \asymp \exp(-t_0 n), n \geq 0. \tag{4.31}$$

Also a simple consequence is that there exists a constant $c > 0$ such that

$$\mathbb{E}[|\{X_m^x : m \geq 0\} \cap L_n(x)|^k] \leq \begin{cases} c^k k! \exp(-t_0 n) & \text{if } n \geq 0 \\ c^k k! & \text{if } n < 0 \end{cases}. \tag{4.32}$$

Proof. Note that $\mathbb{P}[\sigma_y^x < \infty] = \mathbb{P}[\sigma_x^y < \infty]$ and

$$|\{X_m^x : m \geq 0\} \cap L_n(x)| = \sum_{y \in L_n(x)} \mathbf{1}_{\{\sigma_y^x < \infty\}}.$$

Hence

$$\mathbb{E}[|\{X_m^x : m \geq 0\} \cap L_n(x)|] = \mathbb{E}\left[\sum_{y \in L_n(x)} \mathbf{1}_{\{\sigma_y^x < \infty\}}\right] = \mathbb{E}[|\mathfrak{T}_x \cap L_n(x)|]$$

By Proposition 4.6 one has that $\mathbb{E}[|\mathfrak{T}_x \cap L_n(x)|] \leq \exp(-t_0 n)$ and then

$$\mathbb{P}[\{X_m^x : m \geq 0\} \cap L_n(x) \neq \emptyset] \leq \mathbb{E}[|\{X_m^x : m \geq 0\} \cap L_n(x)|] \asymp \exp(-t_0 n). \tag{4.33}$$

This gives the upper bound of (4.31).

Write $N_n = \sum_{m=0}^{\infty} \mathbf{1}_{\{X_m^x \in L_n(x)\}}$ for $n \in \mathbb{Z}$ as in the proof of Corollary 4.2. Similar to (4.6), by the strong Markov property of simple random walk there exists a constant $c \in (0, 1)$ such that

$$\mathbb{P}[N_n \geq k | N_n > 0] \leq (1 - c)^{k-1}, k \geq 1. \tag{4.34}$$

In particular,

$$\mathbb{E}[N_n^k | N_n > 0] \leq \int_0^{\infty} k y^{k-1} (1 - c)^{y-1} dy = \frac{1}{1 - c} \cdot \frac{k!}{(-\log(1 - c))^k} \tag{4.35}$$

Notice that $|\{X_m^x : m \geq 0\} \cap L_n(x)| \leq N_n$ and $\mathbb{P}[N_n > 0] = \mathbb{P}[\{X_m^x : m \geq 0\} \cap L_n(x) \neq \emptyset]$. Therefore we have the upper bound:

$$\begin{aligned} \mathbb{E}[|\{X_m^x : m \geq 0\} \cap L_n(x)|^2] &\leq \mathbb{P}[N_n > 0] \mathbb{E}[N_n^2 | N_n > 0] \stackrel{(4.35)}{\leq} \mathbb{P}[N_n > 0] \\ &= \mathbb{P}[\{X_m^x : m \geq 0\} \cap L_n(x) \neq \emptyset] \stackrel{(4.33)}{\leq} \exp(-t_0 n) \end{aligned} \tag{4.36}$$

By second moment method one has the lower bound of (4.31), namely

$$\mathbb{P}[\{X_m^x : m \geq 0\} \cap L_n(x) \neq \emptyset] \geq \frac{(\mathbb{E}[|\{X_m^x : m \geq 0\} \cap L_n(x)|])^2}{\mathbb{E}[|\{X_m^x : m \geq 0\} \cap L_n(x)|^2]} \stackrel{(4.33), (4.36)}{\geq} \exp(-t_0 n).$$

Hence for $n \geq 0$, by (4.34) and (4.31) there exists a constant $c_1 > 0$ such that

$$\mathbb{P}[|\{X_m^x : m \geq 0\} \cap L_n(x)| \geq k] \leq \mathbb{P}[N_n \geq k] \leq e^{-t_0 n - c_1 k}. \tag{4.37}$$

For $n \geq 0$, the simple random walk X^x hits every $L_{-n}(x)$ almost surely. Thus $\mathbb{P}[N_{-n} > 0] = 1$. Hence by (4.34)

$$\mathbb{P}[|\{X_m^x : m \geq 0\} \cap L_{-n}(x)| \geq k] \leq \mathbb{P}[N_{-n} \geq k] \leq e^{-c_1 k}. \tag{4.38}$$

Recall for a nonnegative random variable Y and $p > 0$, $\mathbb{E}[Y^p] = \int_0^{\infty} p y^{p-1} \mathbb{P}[Y > y] dy$ (for example see Lemma 2.2.13 in [5]). This fact and estimates (4.37, 4.38) imply (4.32).

It remains to show that there exists $c_2 > 0$ such that for all $n \in \mathbb{Z}, k \geq 1$

$$\mathbb{P}[|\{X_m^x : m \geq 0\} \cap L_n(x)| \geq k] \geq \begin{cases} e^{-t_0 n - c_2 k} & \text{if } n \geq 0 \\ e^{-c_2 k} & \text{if } n < 0 \end{cases}. \tag{4.39}$$

and we defer its proof to the appendix. □

In the above proof we use the first and second moment to obtain the estimates (4.31). The estimates can also be obtained by using optimal stopping theorem since $\Delta(X_0, X_n)$ is a martingale (see the proof of Lemma 4.1).

Corollary 4.8. *For the future $\mathfrak{F}(x, \infty)$, there exists a constant $c > 0$ such that*

$$\mathbb{E}[|\mathfrak{F}(x, \infty) \cap L_n(x)|^k] \leq \begin{cases} c^k k! e^{-t_0 n} & \text{if } n \geq 0 \\ c^k k! & \text{if } n < 0 \end{cases}$$

Next we extend Corollary 3.5 to all nonunimodular transitive graphs.

Proposition 4.9. *The expected tilted volume $\mathbb{E}[|T_x|_{x,\lambda}]$ is finite if and only if $\lambda \in (0, 1)$.*

Proof. The proof is the same as Corollary 3.5. One just need to replace (3.2) with the corresponding estimates (4.12) and (4.13) in the general case. \square

Similarly Proposition 4.6 and 4.5 yield the following proposition. We omit its proof.

Proposition 4.10. *The expected tilted volume $\mathbb{E}[|\mathfrak{T}_x|_{x,\lambda}]$ is finite if and only if $\lambda \in (0, 1)$. Similarly,*

- $\mathbb{E}[|\mathfrak{F}(x, \infty)|_{x,\lambda}] < \infty$ is finite if and only if $\lambda \in (0, 1)$.
- $\mathbb{E}[|\mathfrak{P}(x)|_{x,\lambda}] < \infty$ is finite if and only if $\lambda \in (0, 1)$.

Now we are ready to prove Theorem 1.1. In fact it is just the $\lambda = 1$ case in the following proposition.

Proposition 4.11. *If $\lambda \leq 0$, then $|T_x|_{x,\lambda} = \infty$ a.s. If $\lambda > 0$, then $|T_x|_{x,\lambda} < \infty$ a.s.*

Proof. Let $\{X_n\}_{n \geq 0}$ be a simple random walk started at x . If we sample \mathfrak{F} starting with $\{X_n\}_{n \geq 0}$ and let $\mathfrak{F}(x, \infty) = (x_0, x_1, x_2, \dots)$ be future of x . Then Corollary 4.2 implies that $\Delta(x_0, x_n) \rightarrow 0$ almost surely. Thus for $\lambda \leq 0$, $|T_x|_{x,\lambda} \geq |\mathfrak{F}(x, \infty)|_{x,\lambda} = \infty$.

For $\lambda > 0$, use the same proof as Proposition 3.6 with Proposition 4.9 replacing the role of Corollary 3.5. \square

Moreover (4.76) in Proposition 4.24 gives a quantitative tail bound of the weight of the tree T_x in the WUSF.

Remark 4.12. Given Proposition 4.3 there is also another way of proving Theorem 1.1. We just sketch the idea here. One can use the update-tolerance [9] and TMTP to show that if T_x is heavy with positive probability, then T_x will intersect $L_n(x)$ with infinitely many vertices with positive probability, which contradicts the fact that $\mathbb{E}[|T_x \cap L_n(x)|] < \infty$.

4.3 High moments

We shall establish some upper bounds on high moments of $|\mathfrak{T}_x \cap L_n(x)|$ and $|T_x \cap L_n(x)|$. The high moments for $|\mathfrak{F}(x, \infty) \cap L_n(x)|$ have already been given in Corollary 4.8. The high moments for the past $|\mathfrak{P}(x) \cap L_n(x)|$ can be obtained from the corresponding result for \mathfrak{T}_x using Lemma 2.1. See Table 2 for a summary of these results.

Write $\tau(x, y) := \mathbb{P}[y \in \mathfrak{T}_x]$. By transitivity one has $\tau(x, y) = \mathbb{P}[\sigma_x^y < \infty] = \mathbb{P}[\sigma_y^x < \infty] = \tau(y, x)$, where σ_x^y is hitting time of x by a simple random walk started from y . In the following we will use the convention $\tau(x, x) = 1$. Write $\tau(x_0, x_1, \dots, x_k) := \mathbb{P}[\bigcap_{i=1}^k \{x_i \in \mathfrak{T}_{x_0}\}]$. The following lemma is an analogue of Lemma 6.89 in [7].

Lemma 4.13. *For all vertices x_0, x_1, x_2 we have*

$$\tau(x_0, x_1, x_2) \leq \sum_{u \in V} \tau(x_0, u) \tau(u, x_1) \tau(u, x_2) \tag{4.40}$$

Table 2: High moments for the intersections with a slab

Quantities ($n \in \mathbb{Z}, k \geq 2$)	Results	Position
$\mathbb{E}[\mathfrak{F}(x, \infty) \cap L_n(x) ^k]$	$\begin{cases} \leq c^k k! e^{-t_0 n} & \text{if } n \geq 0 \\ \leq c^k k! & \text{if } n < 0 \end{cases}$	Corollary 4.8
$\mathbb{E}[\mathfrak{T}_x \cap L_n(x) ^k]$	$\begin{cases} \leq c^k (k!)^2 e^{-t_0 n} & \text{if } n \geq 0 \\ \leq c^k (k!)^2 n ^{k-1} & \text{if } n < 0 \end{cases}$	Proposition 4.15
$\mathbb{E}[\mathfrak{P}(x) \cap L_n(x) ^k]$	$\begin{cases} \leq c^k (k!)^2 e^{-t_0 n} & \text{if } n \geq 0 \\ \leq c^k (k!)^2 n ^{k-1} & \text{if } n < 0 \end{cases}$	Corollary 4.16
$\mathbb{E}[T_x \cap L_n(x) ^k]$	$\begin{cases} \leq c^k (k!)^5 (n \vee 1) e^{-t_0 n} & \text{if } n \geq 0 \\ \leq c^k (k!)^5 n ^{2k} & \text{if } n < 0 \end{cases}$	Proposition 4.20

Proof. Root the tree \mathfrak{T}_{x_0} at x_0 . Let γ_i be the path on the tree from x_0 to x_i for $i = 1, 2$. Let $s(x_0, x_1, x_2)$ be the last vertex on the path $\gamma_1 \cap \gamma_2$. If we sample \mathfrak{T}_{x_0} from independent simple random walks started from u, x_1, x_2, \dots , then we know that

$$\mathbb{P}[x_1, x_2 \in \mathfrak{T}_{x_0} \text{ and } s(x_0, x_1, x_2) = u] \leq \tau(x_0, u)\tau(u, x_1)\tau(u, x_2) \tag{4.41}$$

Notice if $x_1 = x_2$, then $s(x_0, x_1, x_2) = x_1 = x_2$ the above inequality is still true. If $x_1 = x_0 \neq x_2$, then $s(x_0, x_1, x_2) = x_0 = x_1$ and the above inequality is still true. Summing this inequality over all the possible choices of $s(x_0, x_1, x_2)$ one obtains the desired inequality (4.40). \square

Observation 4.14. If $u \in L_m(x)$ and $x_1 \in L_n(x)$, then

$$x_1 \in L_{n-m}(u).$$

In other words,

$$L_n(x) = L_{n-m}(u).$$

Also by Proposition 4.6 one has

$$\mathbb{E}\left[\sum_{x_1 \in L_n(x)} \tau(u, x_1) \mid U\right] \asymp \exp(-t_0(n - m) \vee 0) \tag{4.42}$$

Proposition 4.15. There exists $c_1 > 0$ such that for all $n \geq 0, k \geq 2$, one has

$$\mathbb{E}[|\mathfrak{T}_x \cap L_n(x)|^k] \leq c_1^k (k!)^2 \exp(-t_0 n) \tag{4.43}$$

and

$$\mathbb{E}[|\mathfrak{T}_x \cap L_{-n}(x)|^k] \leq c_1^k (k!)^2 (n \vee 1)^{k-1}. \tag{4.44}$$

Proof. Here we only prove the case $k = 2$. The proof of the general case will be given in the appendix.

Note $|\mathfrak{T}_x \cap L_n(x)|^2 = \sum_{x_1, x_2 \in L_n(x)} \tau(x, x_1, x_2)$. Using Lemma 4.13 one has

$$\begin{aligned} \mathbb{E}[|\mathfrak{T}_x \cap L_n(x)|^2] &= \mathbb{E}\left[\sum_{x_1, x_2 \in L_n(x)} \tau(x, x_1, x_2)\right] \\ &\leq \mathbb{E}\left[\sum_{u \in V, x_1, x_2 \in L_n(x)} \tau(x, u)\tau(u, x_1)\tau(u, x_2)\right] \\ &\leq \sum_{j=-\infty}^{\infty} \mathbb{E}\left[\sum_{u \in L_j(x), x_1, x_2 \in L_n(x)} \tau(x, u)\tau(u, x_1)\tau(u, x_2)\right] \\ &\stackrel{(4.42)}{\asymp} \sum_{j=-\infty}^{\infty} \exp(-t_0(j \vee 0)) \cdot \exp(-2t_0(n - j) \vee 0) \end{aligned} \tag{4.45}$$

Hence for $n \geq 0$, one has

$$\mathbb{E}[|\mathfrak{F}_x \cap L_n(x)|^2] \preceq \sum_{j=-\infty}^0 e^{-2t_0 n} e^{2t_0 j} + \sum_{j=0}^n e^{-2t_0 n} e^{t_0 j} + \sum_{j=n+1}^{\infty} e^{-t_0 j} \asymp e^{-t_0 n}.$$

Similarly for $n \leq -1$, one can compute that

$$\mathbb{E}[|\mathfrak{F}_x \cap L_{-n}(x)|^2] \preceq |n|. \quad \square$$

An immediate corollary of Proposition 4.15 is the upper bound on $|\mathfrak{F}(x) \cap L_n(x)|$ using Lemma 2.1.

Corollary 4.16. *There exists a constant $c > 0$ such that for all $n \geq 0, k \geq 2$,*

$$\mathbb{E}[|\mathfrak{F}(x) \cap L_n(x)|^k] \leq \begin{cases} c^k (k!)^2 e^{-t_0 n} & \text{if } n \geq 0 \\ c^k (k!)^2 |n|^{k-1} & \text{if } n < 0 \end{cases}.$$

Corollary 4.17. *There exist constants $c_1, c_2, c_3, c_4 > 0$ such that for all $n \geq 0, k \geq 1$,*

$$\mathbb{P}[|\mathfrak{F}_x \cap L_n(x)| \geq k] \leq c_2 e^{-t_0 n} \cdot e^{-c_1 \sqrt{k}} \tag{4.46}$$

and for all $n > 0, k \geq 1$

$$\mathbb{P}[|\mathfrak{F}_x \cap L_{-n}(x)| \geq k] \leq \frac{c_4}{n} e^{-c_3 \sqrt{\frac{k}{n}}} \tag{4.47}$$

Proof. Taking $c_1 = \frac{1}{2\sqrt{c}}$, then

$$\begin{aligned} \mathbb{E}\left[e^{c_1 \sqrt{|\mathfrak{F}_x \cap L_n(x)|}}\right] &\leq \mathbb{E}\left[e^{c_1 \sqrt{|\mathfrak{F}_x \cap L_n(x)|}}\right] + \mathbb{E}\left[e^{-c_1 \sqrt{|\mathfrak{F}_x \cap L_n(x)|}}\right] \\ &= 2 \sum_{k=0}^{\infty} \frac{c_1^{2k}}{(2k)!} \mathbb{E}[|\mathfrak{F}_x \cap L_n(x)|^k] \\ &\stackrel{(4.43)}{\leq} 2 \sum_{k=0}^{\infty} \frac{c_1^{2k}}{(2k)!} c^k \cdot (k!)^2 \cdot e^{-t_0 n} \\ &\leq 2 \sum_{k=0}^{\infty} (c_1^2 c)^k \cdot e^{-t_0 n} < 3e^{-t_0 n}. \end{aligned} \tag{4.48}$$

Hence by Markov's inequality for all $n \geq 0, k \geq 1$ one has that

$$\begin{aligned} \mathbb{P}[|\mathfrak{F}_x \cap L_n(x)| \geq k] &= \mathbb{P}[\exp(c_1 \sqrt{|\mathfrak{F}_x \cap L_n(x)|}) \geq \exp(c_1 \sqrt{k})] \\ &\leq \frac{3e^{-t_0 n}}{\exp(c_1 \sqrt{k})} \leq c_2 e^{-t_0 n} \cdot e^{-c_1 \sqrt{k}}. \end{aligned}$$

The proof of (4.47) is similar. For small $c_3 > 0$, one has that $\mathbb{E}[\exp(c_3 \sqrt{\frac{|\mathfrak{F}_x \cap L_{-n}(x)|}{n}})] \leq \frac{c_4}{n}$ and then use Markov's inequality. □

Question 4.18. Is it the case that for all $n \geq 0, k \geq 1$,

$$\mathbb{P}[|\mathfrak{F}_x \cap L_n(x)| \geq k] = e^{-t_0 n - \Theta(\sqrt{k})}?$$

The above Corollary 4.17 establishes the upper bound. As we will see in the appendix, the corresponding lower bound is also true for the toy model.

Using the estimates for $\mathbb{E}[|\mathfrak{F}_x \cap L_n(x)|^k]$ we can derive upper bounds for $\mathbb{E}[|T_x \cap L_n(x)|^k]$. Recall $U = \{U_v : v \in V\}$ are the labels we used to define the slabs.

Lemma 4.19. *Suppose $\mathfrak{F}(x, \infty) = (x_0, x_1, x_2, \dots)$ is the future of x in the WUSF sample \mathfrak{F} . Let \mathfrak{T}_{x_i} be the tree of x_i in x_i -WUSF, sampled independently of \mathfrak{F} , U and each other. Then for every $n \in \mathbb{Z}, k \geq 2$, one has*

$$\mathbb{E}[|T_x \cap L_n(x)|^k \mid \mathfrak{F}(x, \infty), U] \leq k! \sum_{\substack{k_0, k_1, \dots \geq 0: \\ k_0 + k_1 + \dots = k}} \prod_{i: k_i \neq 0} \frac{1}{k_i!} \mathbb{E}[|\mathfrak{T}_{x_i} \cap L_n(x)|^{k_i} \mid U]. \quad (4.49)$$

Proof. The proof is similar to the one of Lemma 6.8 in [11]. We present the details for reader's convenience. In the following we fix $n \in \mathbb{Z}$.

Given the future $\mathfrak{F}(x, \infty) = (x_0, x_1, x_2, \dots)$ of x and $i \geq 0$, we call the connected component of x_i in $T_x \setminus \{x_0, x_1, \dots, x_{i-1}, x_{i+1}, x_{i+2}, \dots\}$ the i -th **bush** of T_x and denote it by $\text{Bush}_i(x)$. Denote by N_i the number of vertices in $L_n(x) \cap \text{Bush}_i(x)$. In particular, $\text{Bush}_0(x) = \mathfrak{P}(x)$. By the lightness of T_x , almost surely only finitely many N_i 's are nonzero.

Notice that

$$|T_x \cap L_n(x)|^k = (N_0 + N_1 + \dots)^k = \sum_{\substack{k_0, k_1, \dots \geq 0: \\ k_0 + k_1 + \dots = k}} k! \prod_{i: k_i \neq 0} \frac{N_i^{k_i}}{k_i!} \quad (4.50)$$

Since $N_i = \sum_{y \in V} \mathbf{1}_{\{y \in L_n(x) \cap \text{Bush}_i(x)\}}$, $N_i^{k_i} = \sum_{y_j: 1 \leq j \leq k_i} \mathbf{1}_{\{y_j \in L_n(x) \cap \text{Bush}_i(x) \text{ for } j=1, \dots, k_i\}}$ for $k_i > 0$ and then for $k_i > 0$

$$\mathbb{E}[N_i^{k_i} \mid \mathfrak{F}(x, \infty), U] = \sum_{y_j: 1 \leq j \leq k_i} \mathbb{P}\left[\bigcap_{1 \leq j \leq k_i} y_j \in L_n(x) \cap \text{Bush}_i(x) \mid \mathfrak{F}(x, \infty), U\right].$$

Similarly for any sequence $(k_i)_{i \geq 0}$ such that $k_i \geq 0, \sum_{i=0}^{\infty} k_i = k$ one has that

$$\begin{aligned} & \mathbb{E}\left[\prod_{i: k_i \neq 0} N_i^{k_i} \mid \mathfrak{F}(x, \infty), U\right] \\ &= \prod_{i: k_i \neq 0} \sum_{\substack{y_{i,j}: \\ 1 \leq j \leq k_i}} \mathbb{P}\left[\bigcap_{i: k_i \neq 0} \bigcap_{1 \leq j \leq k_i} y_{i,j} \in L_n(x) \cap \text{Bush}_i(x) \mid \mathfrak{F}(x, \infty), U\right] \end{aligned} \quad (4.51)$$

For each $i \geq 0$, let $Y_i = \{y_{i,1}, \dots, y_{i,k_i}\}$ be a finite (possibly with multiplicity) collection of vertices of G and $W_i = \{w_{i,1}, \dots, w_{i,m_i}\}$ the corresponding set of vertices of Y_i without multiplicity. In particular if $k_i = 0$ then $m_i = 0$ and W_i is an empty set. Let \mathcal{A}_i be the event that for every vertex $w \in W_i, w \in \text{Bush}_i(x)$.

Let $\{X^{i,j} : i \geq 0, m_i \neq 0, 1 \leq j \leq m_i\}$ be a collection of independent simple random walks, independent of $\mathfrak{F}(x, \infty)$, such that $X_0^{i,j} = w_{i,j}$ for each $i \geq 0$ with $m_i \neq 0$ and $1 \leq j \leq m_i$.

For each $i \geq 0$ such that $m_i \neq 0$, let \mathcal{B}_i be the event that, if we sample x_i -WUSF using Wilson's algorithm, starting with the random walks $X^{i,1}, \dots, X^{i,m_i}$, then for every vertex $w \in W_i, w$ is connected to x_i in x_i -WUSF.

It is easy to see that if we sample \mathfrak{F} conditional on $\mathfrak{F}(x, \infty)$ using Wilson's algorithm starting with $X^{0,1}, \dots, X^{0,m_0}$, then $X^{1,1}, \dots, X^{1,m_1}$, and so on, then we have $\mathcal{A}_i \subset \mathcal{B}_i$. Therefore

$$\mathbb{P}\left[\bigcap_{i: k_i \neq 0} \mathcal{A}_i \mid \mathfrak{F}(x, \infty), U\right] \leq \prod_{i: k_i \neq 0} \mathbb{P}[\mathcal{B}_i \mid \mathfrak{F}(x, \infty), U] \quad (4.52)$$

Summing over all the possible choices of the sets Y_i such that $Y_i \subset L_n(x), i \geq 0$ and $\sum_{i=0}^{\infty} k_i = k$, by (4.50), (4.51) and (4.52) we obtain (4.49). \square

Considering the high moments for the intersection $T_x \cap L_n(x)$, we have the following upper bounds.

Proposition 4.20. For all $n \geq 0, k \geq 2$ there exists a constant $c_1 > 0$ such that

$$\mathbb{E}[|T_x \cap L_n(x)|^k] \leq c_1^k (k!)^5 (n \vee 1) e^{-t_0 n} \tag{4.53}$$

and

$$\mathbb{E}[|T_x \cap L_{-n}(x)|^k] \leq c_1^k (k!)^5 (n \vee 1)^{2k}. \tag{4.54}$$

Proof. By Proposition 4.6 and 4.15 one has that there exists $c_2 > 0$ such that for all $k \geq 1$,

$$\mathbb{E}[|\mathfrak{F}_x \cap L_n(x)|^k] \leq c_2^k (k!)^2 (|n \wedge 0| \vee 1)^{k-1} \exp(-t_0(n \vee 0)) = \begin{cases} c_2^k (k!)^2 |n|^{k-1} & \text{if } n < 0 \\ c_2^k (k!)^2 e^{-t_0 n} & \text{if } n \geq 0 \end{cases} \tag{4.55}$$

Suppose $\mathfrak{F}(x, \infty) = (x_0, x_1, x_2, \dots)$ is the future of x . From the Observation 4.14 we know that if $u \in L_m(x)$ for some integer m , then $L_n(x) = L_{n-m}(u)$. Hence for any positive integer t by (4.55) we have

$$\begin{aligned} \mathbb{E}[|\mathfrak{F}_u \cap L_n(x)|^t \mid U] &= \mathbb{E}[|\mathfrak{F}_u \cap L_{n-m}(u)|^t] \\ &\leq c_2^t (t!)^2 (|(n-m) \wedge 0| \vee 1)^{t-1} \exp(-t_0((n-m) \vee 0)) \end{aligned}$$

Let $m(x_i)$ denote the integer m such that $x_i \in L_m(x)$ (if there are two such m 's, taking the smaller one). Then using Lemma 4.19 one has that for any $n \in \mathbb{Z}, k \geq 2$

$$\begin{aligned} \text{LHS} &:= \mathbb{E}[|T_x \cap L_n(x)|^k] \\ &\preceq c_2^k k! \sum_{\substack{k_0, k_1, \dots \geq 0: \\ k_0 + k_1 + \dots = k}} \mathbb{E} \left[\prod_{i: k_i \neq 0} \frac{(k_i!)^2}{k_i!} (|(n - m(x_i)) \wedge 0| \vee 1)^{k_i - 1} \left(\frac{e^{t_0 m(x_i)}}{e^{t_0 n}} \wedge 1 \right) \right] \\ &\preceq c_2^k (k!)^2 \sum_{\substack{k_0, k_1, \dots \geq 0: \\ k_0 + k_1 + \dots = k}} \mathbb{E} \left[\prod_{i: k_i \neq 0} (|(n - m(x_i)) \wedge 0| \vee 1)^{k_i - 1} \left(\frac{e^{t_0 m(x_i)}}{e^{t_0 n}} \wedge 1 \right) \right] \\ &\preceq c_2^k (k!)^2 \sum_{\substack{k_0, k_1, \dots \geq 0: \\ k_0 + k_1 + \dots = k}} \mathbb{E} \left[\prod_{i: k_i \neq 0} (|(n - m(x_i)) \wedge 0| \vee 1)^{k_i} \left(\frac{e^{t_0 m(x_i)}}{e^{t_0 n}} \wedge 1 \right) \right] \end{aligned} \tag{4.56}$$

where in the second inequality we use $\prod_{i: k_i \neq 0} k_i! \leq k!$.

Write $I := \{i : k_i \neq 0\}$. Then $1 \leq |I| \leq k$ and for a fixed set I of indices, the number of positive solutions $(k_i : i \in I)$ for the equation $\sum_{i \in I} k_i = k$ is $\binom{k}{|I|-1} \leq 2^k$. Set

$$f(k_i, m, n) := \begin{cases} 1 & \text{if } m \leq n \\ k_i & \text{if } m > n \end{cases} \quad \text{For this function } f, \text{ it is easy to check that}$$

$$(|(n - m(x_i)) \wedge 0| \vee 1)^{k_i} \left(\frac{e^{t_0 m(x_i)}}{e^{t_0 n}} \wedge 1 \right) = (|(n - m(x_i)) \wedge 0| \vee 1)^{f(k_i, m, n)} \left(\frac{e^{t_0 m(x_i)}}{e^{t_0 n}} \wedge 1 \right)^{f(k_i, m, n)}.$$

Then for each sequence of integers (k_0, k_1, \dots) such that $k_i \geq 0$ and $\sum_{i=0}^\infty k_i = k$, the term $\prod_{i: k_i \neq 0} (|(n - m(x_i)) \wedge 0| \vee 1)^{k_i} \left(\frac{e^{t_0 m(x_i)}}{e^{t_0 n}} \wedge 1 \right)$ appears in the expansion of

$$\left(\sum_{i=0}^\infty (|(n - m(x_i)) \wedge 0| \vee 1) \left(\frac{e^{t_0 m(x_i)}}{e^{t_0 n}} \wedge 1 \right) \right)^t,$$

where $t = \sum_{i \in I} f(k_i, m(x_i), n) \leq \sum_{i \in I} k_i = k$. Since this term appears at most $\binom{k}{|I|-1} \leq 2^k$ times in

$$\sum_{\substack{k_0, k_1, \dots \geq 0: \\ k_0 + k_1 + \dots = k}} \mathbb{E} \left[\prod_{i: k_i \neq 0} (|(n - m(x_i)) \wedge 0| \vee 1)^{k_i} \left(\frac{e^{t_0 m(x_i)}}{e^{t_0 n}} \wedge 1 \right) \right],$$

one has that

$$\begin{aligned} \mathbb{E}[|T_x \cap L_n(x)|^k] &\leq 2^k c_2^k (k!)^2 \sum_{t=1}^k \mathbb{E} \left(\sum_{i=0}^{\infty} (|(n - m(x_i)) \wedge 0| \vee 1) \left(\frac{e^{t_0 m(x_i)}}{e^{t_0 n}} \wedge 1 \right) \right)^t \\ &= 2^k c_2^k (k!)^2 \sum_{t=1}^k \mathbb{E} \left(\sum_{j=-\infty}^{\infty} (|(n - j) \wedge 0| \vee 1) \left(\frac{e^{t_0 j}}{e^{t_0 n}} \wedge 1 \right) |\mathfrak{F}(x, \infty) \cap L_j(x)| \right)^t. \end{aligned}$$

Set $S_t := \mathbb{E} \left(\sum_{j=-\infty}^{\infty} (|(n - j) \wedge 0| \vee 1) \left(\frac{e^{t_0 j}}{e^{t_0 n}} \wedge 1 \right) |\mathfrak{F}(x, \infty) \cap L_j(x)| \right)^t$ for $1 \leq t \leq k$. The above inequality becomes

$$\mathbb{E}[|T_x \cap L_n(x)|^k] \leq c_3^k (k!)^2 (S_1 + \dots + S_k) \tag{4.57}$$

By Proposition 4.5 one has that for $n \geq 0$

$$S_1 \leq \sum_{j=-\infty}^{-1} \frac{e^{t_0 j}}{e^{t_0 n}} + \sum_{j=0}^n \frac{1}{e^{t_0 n}} + \sum_{j=n+1}^{\infty} (j - n) e^{-t_0 j} \leq (n \vee 1) e^{-t_0 n}. \tag{4.58}$$

Similarly if $n < 0$, then

$$S_1 \leq \sum_{j=-\infty}^{n-1} \frac{e^{t_0 j}}{e^{t_0 n}} + \sum_{j=n}^0 ((j - n) \vee 1) + \sum_{j=1}^{\infty} (j - n) e^{-t_0 j} \leq |n|^2. \tag{4.59}$$

By Hölder’s inequality one has that for any $t \geq 2$ and nonnegative sequences $(a_j)_{j \in \mathbb{Z}}, (b_j)_{j \in \mathbb{Z}}$

$$\left(\sum_{j=-\infty}^{\infty} a_j b_j \right)^t \leq \left(\sum_{j=-\infty}^{\infty} a_j^{\frac{t}{t-1}} \right)^{t-1} \left(\sum_{j=-\infty}^{\infty} b_j^t \right),$$

with equality when $b_j = c a_j^{\frac{1}{t-1}}$.

For $2 \leq t \leq k$, applying the above Hölder’s inequality with $a_j > 0$ and $b_j = \frac{1}{a_j} \times (|(n - j) \wedge 0| \vee 1) \left(\frac{e^{t_0 j}}{e^{t_0 n}} \wedge 1 \right) |\mathfrak{F}(x, \infty) \cap L_j(x)|$ and then taking expectations one has that

$$\begin{aligned} S_t &\leq \left(\sum_{j=-\infty}^{\infty} a_j^{\frac{t}{t-1}} \right)^{t-1} \cdot \sum_{j=-\infty}^{\infty} \frac{1}{a_j^t} (|(n - j) \wedge 0| \vee 1)^t \left(\frac{e^{t_0 j}}{e^{t_0 n}} \wedge 1 \right)^t \mathbb{E} |\mathfrak{F}(x, \infty) \cap L_j(x)|^t \\ &\leq \left(\sum_{j=-\infty}^{\infty} a_j^{\frac{t}{t-1}} \right)^{t-1} \cdot \sum_{j=-\infty}^{\infty} \frac{1}{a_j^t} (|(n - j) \wedge 0| \vee 1)^t \left(\frac{e^{t_0 j}}{e^{t_0 n}} \wedge 1 \right)^t (ct)^t e^{-t_0(j \vee 0)}, \end{aligned} \tag{4.60}$$

where in the second inequality we use (4.32) and $t! \leq t^t$. Pick a_j such that

$$a_j^{\frac{1}{t-1}} = \frac{1}{a_j} (|(n - j) \wedge 0| \vee 1) \left(\frac{e^{t_0 j}}{e^{t_0 n}} \wedge 1 \right) (ct) e^{-\frac{t_0(j \vee 0)}{t}} =: \tilde{b}_j, \tag{4.61}$$

i.e., for $n \geq 0$,

$$a_j^{\frac{t}{t-1}} = \begin{cases} (j - n) \cdot 1 \cdot (ct) e^{-\frac{t_0 j}{t}} & \text{if } j > n \\ 1 \cdot \frac{e^{t_0 j}}{e^{t_0 n}} \cdot (ct) e^{-\frac{t_0 j}{t}} & \text{if } 0 \leq j \leq n \\ 1 \cdot \frac{e^{t_0 j}}{e^{t_0 n}} \cdot (ct) & \text{if } j < 0 \end{cases} \tag{4.62}$$

and for $n < 0$,

$$a_j^{\frac{t}{t-1}} = \begin{cases} (j-n) \cdot 1 \cdot (ct)e^{-\frac{t_0j}{t}} & \text{if } j > 0 \\ (j-n) \cdot 1 \cdot (ct) & \text{if } n < j \leq 0 \\ 1 \cdot \frac{e^{t_0j}}{e^{t_0n}} \cdot (ct) & \text{if } j \leq n \end{cases} \tag{4.63}$$

With a_j given above and \tilde{b}_j as defined in (4.61), $\tilde{b}_j = a_j^{\frac{1}{t-1}}$ and (4.60) becomes

$$\begin{aligned} S_t &\leq \left(\sum_{j=-\infty}^{\infty} a_j^{\frac{t}{t-1}} \right)^{t-1} \cdot \sum_{j=-\infty}^{\infty} \tilde{b}_j^t \\ &= \left(\sum_{j=-\infty}^{\infty} a_j^{\frac{t}{t-1}} \right)^{t-1} \cdot \sum_{j=-\infty}^{\infty} a_j^{\frac{t}{t-1}} = \left(\sum_{j=-\infty}^{\infty} a_j^{\frac{t}{t-1}} \right)^t \end{aligned} \tag{4.64}$$

Hence for $n \geq 0, 2 \leq t \leq k$, by (4.62) and (4.64) one has

$$\begin{aligned} S_t &\leq \left(\sum_{j=-\infty}^{-1} ct \frac{e^{t_0j}}{e^{t_0n}} + \sum_{j=1}^n ct \frac{e^{t_0j}}{e^{t_0n}} e^{-\frac{t_0j}{t}} + \sum_{j=n+1}^{\infty} (j-n)cte^{-\frac{t_0j}{t}} \right)^t \\ &\leq (ct)^t \left(\frac{c_4}{e^{t_0n}} + \frac{c_4 e^{t_0(1-\frac{1}{t})n}}{e^{t_0n}(e^{t_0(1-\frac{1}{t})} - 1)} + c_4 e^{-\frac{t_0n}{t}} \int_0^{\infty} x e^{-\frac{t_0x}{t}} dx \right)^t \\ &\leq (ct)^t \left(\frac{c_4}{e^{t_0n}} + \frac{c_5}{e^{\frac{t_0n}{t}}} + c_4 e^{-\frac{t_0n}{t}} \left(\frac{t}{t_0}\right)^2 \right)^t \\ &\leq c_6^t t^{3t} e^{-t_0n} \leq c_7^t (t!)^3 e^{-t_0n}, \end{aligned} \tag{4.65}$$

where we use $t^t \leq c_8^t t!$ for some $c_8 > 1$ (by Stirling’s formula for example) in the last inequality.

For $n < 0, 2 \leq t \leq k$, by (4.63) and (4.64) one has

$$\begin{aligned} S_t &\leq \left(\sum_{j=-\infty}^n ct \frac{e^{t_0j}}{e^{t_0n}} + \sum_{j=n+1}^0 (j-n)ct + \sum_{j=1}^{\infty} (j-n)cte^{-\frac{t_0j}{t}} \right)^t \\ &\leq (ct)^t \left(c_9 + |n|^2 + c_9|n| \int_0^{\infty} e^{-\frac{t_0x}{t}} dx + c_9 \int_0^{\infty} x e^{-\frac{t_0x}{t}} dx \right)^t \\ &= (ct)^t \left(c_9 + |n|^2 + c_9|n| \frac{t}{t_0} + c_9 \left(\frac{t}{t_0}\right)^2 \right)^t \\ &\leq (ct)^t (c_{10}(|n| + t)^2)^t \leq c_{11}^t t^t (|n|t)^{2t} \\ &\leq c_{12}^t (t!)^3 |n|^{2t}, \end{aligned} \tag{4.66}$$

Now we are ready to get the final conclusion. For $n \geq 0$, by (4.57), (4.58) and (4.65) we get

$$\mathbb{E}[|T_x \cap L_n(x)|^k] \leq c_1^k (k!)^5 (n \vee 1) e^{-t_0n}.$$

For $n < 0$, by (4.57), (4.59) and (4.66) we get

$$\mathbb{E}[|T_x \cap L_n(x)|^k] \leq c_1^k (k!)^5 |n|^{2k}. \quad \square$$

Table 3: Probabilities of nonempty intersections with a slab

Quantities ($n \in \mathbb{Z}$)	Results	Position
$\mathbb{P}[T_x \cap L_n(x) \neq \emptyset]$	$\begin{cases} \asymp ne^{-t_0n} & \text{if } n > 0 \\ \stackrel{(4.69)}{=} 1 & \text{if } n \leq 0 \end{cases}$	Proposition 4.21
$\mathbb{P}[\mathfrak{T}_x \cap L_n(x) \neq \emptyset]$	$\begin{cases} \asymp e^{-t_0n} & \text{if } n \geq 0 \\ \asymp \frac{1}{ n } & \text{if } n < 0 \end{cases}$	Proposition 4.22
$\mathbb{P}[\mathfrak{P}(x) \cap L_n(x) \neq \emptyset]$	$\begin{cases} \asymp e^{-t_0n} & \text{if } n \geq 0 \\ \asymp \frac{1}{n} & \text{if } n < 0 \end{cases}$	Proposition 4.23
$\mathbb{P}[\mathfrak{F}(x, \infty) \cap L_n(x) \neq \emptyset]$	$\begin{cases} \asymp e^{-t_0n} & \text{if } n > 0 \\ \stackrel{(4.69)}{=} 1 & \text{if } n \leq 0 \end{cases}$	Proposition 4.23

4.4 Probabilities for intersections with a slab

Another natural question to consider is the decay of the probability of a vertex connecting to a certain slab. Similar question was analyzed for Bernoulli percolation [10, Lemma 5.2].

Proposition 4.21. *Let $\{T_x \cap L_n(x) \neq \emptyset\}$ denote the event that T_x has a nonempty intersection with the slab $L_n(x)$. Then for $n \geq 1$*

$$\mathbb{P}[T_x \cap L_n(x) \neq \emptyset] \asymp n \exp(-t_0n). \tag{4.67}$$

In particular

$$\lim_{n \rightarrow \infty} -\frac{1}{t_0n} \log \mathbb{P}[T_x \cap L_n(x) \neq \emptyset] = 1. \tag{4.68}$$

Proof. By Lemma 4.1 one has that $\mathfrak{F}(x, \infty)$ intersects every $L_n(x)$ almost surely. Since $\mathfrak{F}(x, \infty) \subset T_x$, one has

$$\mathbb{P}[T_x \cap L_n(x) \neq \emptyset] = \mathbb{P}[\mathfrak{F}(x, \infty) \cap L_n(x) \neq \emptyset] = 1, \forall n \leq 0. \tag{4.69}$$

On the one hand, for $n \geq 1$ using second moment method and (4.13) and (4.53) one has

$$\mathbb{P}[T_x \cap L_n(x) \neq \emptyset] \geq \frac{(\mathbb{E}[|T_x \cap L_n(x)|])^2}{\mathbb{E}[|T_x \cap L_n(x)|^2]} \succeq ne^{-t_0n}.$$

On the other hand, note when $T_x \cap L_n(x) \neq \emptyset$, $|T_x \cap L_n(x)| \geq 1$. Thus by Markov's inequality one has that

$$\mathbb{P}[T_x \cap L_n(x) \neq \emptyset] = \mathbb{P}[|T_x \cap L_n(x)| \geq 1] \leq \mathbb{E}[|T_x \cap L_n(x)|] \stackrel{(4.13)}{\preceq} n \exp(-t_0n).$$

Combining the above two inequalities one has the conclusion. □

Since the tree \mathfrak{T}_x in the x -WUSF is almost surely finite, $\mathbb{P}[\mathfrak{T}_x \cap L_n(x) \neq \emptyset]$ and $\mathbb{P}[\mathfrak{T}_x \cap L_{-n}(x) \neq \emptyset]$ both decay to zero as $n \rightarrow \infty$.

Proposition 4.22. *For the tree \mathfrak{T}_x in the x -WUSF and $n \geq 0$, one has*

$$\mathbb{P}[\mathfrak{T}_x \cap L_n(x) \neq \emptyset] \asymp \exp(-t_0n) \tag{4.70}$$

In particular,

$$\lim_{n \rightarrow \infty} -\frac{1}{t_0n} \log \mathbb{P}[\mathfrak{T}_x \cap L_n(x) \neq \emptyset] = 1. \tag{4.71}$$

Moreover for $n \geq 1$, the decay of $\mathbb{P}[\mathfrak{T}_x \cap L_{-n}(x) \neq \emptyset]$ is much slower.

$$\mathbb{P}[\mathfrak{T}_x \cap L_{-n}(x) \neq \emptyset] \asymp \frac{1}{n}. \tag{4.72}$$

Proof. The proof of (4.70) is similar to (4.67) and thus we omit it.

By second moment method one has the lower bound in (4.72), namely

$$\mathbb{P}[\mathfrak{T}_x \cap L_{-n}(x) \neq \emptyset] \geq \frac{(\mathbb{E}[|\mathfrak{T}_x \cap L_{-n}(x)|])^2}{\mathbb{E}[|\mathfrak{T}_x \cap L_{-n}(x)|^2]} \stackrel{(4.26), (4.44)}{\succeq} \frac{1}{n}.$$

Let $\text{diam}_{\text{int}}(\mathfrak{T}_x)$ denote the intrinsic diameter of the finite tree \mathfrak{T}_x . If $\mathfrak{T}_x \cap L_{-n}(x) \neq \emptyset$, then $\text{diam}_{\text{int}}(\mathfrak{T}_x) \geq n - 1$. Hence for $n \geq 1$ one has

$$\mathbb{P}[\mathfrak{T}_x \cap L_{-n}(x) \neq \emptyset] \leq \mathbb{P}[\text{diam}_{\text{int}}(\mathfrak{T}_x) \geq n - 1] \preceq \frac{1}{n},$$

where the last inequality is due to Theorem 7.1 of [11]. □

For the future $\mathfrak{F}(x, \infty)$, it intersects every $L_{-n}(x), n \geq 0$ almost surely, and the probability $\mathbb{P}[\mathfrak{F}(x, \infty) \cap L_n(x) \neq \emptyset]$ decays exponentially. For the past $\mathfrak{P}(x)$, it is a finite tree and hence $\mathbb{P}[\mathfrak{P}(x) \cap L_n(x) \neq \emptyset]$ and $\mathbb{P}[|\mathfrak{P}(x) \cap L_{-n}(x)| \neq \emptyset]$ both decay to zero as n tending to infinity. We summarize these in the following Proposition.

Proposition 4.23. *Suppose $n \geq 0$. For the future one has the following asymptotic behavior:*

$$\mathbb{P}[\mathfrak{F}(x, \infty) \cap L_n(x) \neq \emptyset] \asymp \exp(-t_0 n) \tag{4.73}$$

For the past one has

$$\mathbb{P}[\mathfrak{P}(x) \cap L_n(x) \neq \emptyset] \asymp \exp(-t_0 n) \tag{4.74}$$

and

$$\mathbb{P}[\mathfrak{P}(x) \cap L_{-n}(x) \neq \emptyset] \asymp \frac{1}{n \vee 1}. \tag{4.75}$$

Proof. By (4.23) one has $\mathbb{E}[|\mathfrak{F}(x, \infty) \cap L_n(x)|] \asymp e^{-t_0 n}$. From Corollary 4.8 one has $\mathbb{E}[|\mathfrak{F}(x, \infty) \cap L_n(x)|^2] \preceq e^{-t_0 n}$. Using the first-moment and second-moment method as in the proof of Proposition 4.21 again one obtains (4.73).

Similarly for the past, by (4.24) one has $\mathbb{E}[|\mathfrak{P}(x) \cap L_n(x)|] \asymp e^{-t_0 n}$. From Corollary 4.16 one has $\mathbb{E}[|\mathfrak{P}(x) \cap L_n(x)|^2] \preceq e^{-t_0 n}$. Using the first-moment and second-moment method again one obtains (4.74).

By Lemma 2.1 and Proposition 4.15 one has

$$\mathbb{E}[|\mathfrak{P}(x) \cap L_n(x)|^2] \preceq n \exp(-t_0 n).$$

From Proposition 4.5 one has $\mathbb{E}[|\mathfrak{P}(x) \cap L_n(x)|] \succeq \exp(-t_0 n)$. Using Markov's inequality one then has the lower bound of (4.74).

By Lemma 2.1 and Proposition 4.22 one has the upper bound in (4.75):

$$\mathbb{P}[\mathfrak{P}(x) \cap L_{-n}(x) \neq \emptyset] \leq \mathbb{P}[\mathfrak{T}_x \cap L_{-n}(x) \neq \emptyset] \preceq \frac{1}{n \vee 1}.$$

By (4.22) one has $\mathbb{E}[|\mathfrak{P}(x) \cap L_{-n}(x)|] \asymp 1$. From Corollary 4.16 one has $\mathbb{E}[|\mathfrak{P}(x) \cap L_{-n}(x)|^2] \preceq n \vee 1$. Then the proof of the lower bound of (4.75) is an application of second-moment method again. □

4.5 Tail bounds for tilted volumes

The results for this subsection are summarized in Table 4.

Proposition 4.24. *For $\lambda > 0$, the tilted volume $|T_x|_{x, \lambda}$ has the following property.*

$$\mathbb{P}[|T_x|_{x, \lambda} \geq R] \asymp_{\lambda} R^{-\frac{1}{\lambda}} \log R, \forall R \geq 2. \tag{4.76}$$

Table 4: Probabilities of nonempty intersections with a slab

Quantities ($R \geq 2$)	Results	Position
$\mathbb{P}[T_x _{x,\lambda} \geq R]$	$\begin{cases} \asymp_\lambda R^{-\frac{1}{\lambda}} \log R & \text{if } \lambda > 0 \\ \stackrel{(4.77)}{=} 1 & \text{if } \lambda \leq 0 \end{cases}$	Proposition 4.24
$\mathbb{P}[\mathfrak{T}_x _{x,\lambda} \geq R]$	$\begin{cases} \asymp_\lambda R^{-\frac{1}{\lambda}} & \text{if } \lambda > 0 \\ \asymp \frac{1}{\sqrt{R}} & \text{if } \lambda = 0 \\ \asymp_\lambda \frac{1}{\log R} & \text{if } \lambda < 0 \end{cases}$	Proposition 4.25
$\mathbb{P}[\mathfrak{P}(x) _{x,\lambda} \geq R]$	$\begin{cases} \asymp_\lambda R^{-\frac{1}{\lambda}} & \text{if } \lambda > 0 \\ \asymp \frac{1}{\sqrt{R}} & \text{if } \lambda = 0 \\ \asymp_\lambda \frac{1}{\log R} & \text{if } \lambda < 0 \end{cases}$	Proposition 4.27
$\mathbb{P}[\mathfrak{F}(x, \infty) _{x,\lambda} \geq R]$	$\begin{cases} \asymp_\lambda R^{-\frac{1}{\lambda}} \log R & \text{if } \lambda > 0 \\ \stackrel{(4.77)}{=} 1 & \text{if } \lambda \leq 0 \end{cases}$	Proposition 4.27

Proof. By Lemma 4.1 the future $\mathfrak{F}(x, \infty)$ intersects every $L_n(x), n \leq 0$ almost surely. In particular this implies that for every $\lambda \leq 0$,

$$\mathbb{P}[|T_x|_{x,\lambda} \geq R] = \mathbb{P}[|\mathfrak{F}(x, \infty)|_{x,\lambda} \geq R] = 1, R \geq 2. \tag{4.77}$$

Taking $n_0 = \lceil 1 + \frac{\log R}{t_0 \lambda} \rceil$, then $1 + \frac{\log R}{t_0 \lambda} \leq n_0 \leq 2 + \frac{\log R}{t_0 \lambda}$ and

$$e^{-2t_0} R^{-\frac{1}{\lambda}} \leq \exp(-t_0 n_0) \leq e^{-t_0} R^{-\frac{1}{\lambda}} \tag{4.78}$$

On the event that $\{T_x \cap L_{n_0}(x) \neq \emptyset\}$, one has $|T_x|_{x,\lambda} \geq \exp(t_0 \lambda (n_0 - 1)) \geq R$. Hence by Proposition 4.21

$$\mathbb{P}[|T_x|_{x,\lambda} \geq R] \geq \mathbb{P}[T_x \cap L_{n_0}(x) \neq \emptyset] \geq n_0 \exp(-t_0 n_0) \geq \frac{1}{\lambda} R^{-\frac{1}{\lambda}} \log R.$$

Now we derive the upper bound. If $T_x \cap L_{n_0}(x) = \emptyset$, then $T_x \cap L_j(x) = \emptyset$ for all $j \geq n_0$. If $y \in L_j(x)$, then $\Delta(x, y)^\lambda \leq e^{t_0(j+1)\lambda}$. Thus on the event $T_x \cap L_{n_0}(x) = \emptyset$, $|T_x|_{x,\lambda} \leq \sum_{j=-\infty}^{n_0-1} e^{t_0(j+1)\lambda} |T_x \cap L_j(x)|$. Using the union bounds we have

$$\begin{aligned} \mathbb{P}[|T_x|_{x,\lambda} \geq R] &\leq \mathbb{P}[T_x \cap L_{n_0}(x) \neq \emptyset] + \mathbb{P}\left[\sum_{j=-\infty}^{n_0} e^{t_0(j+1)\lambda} |T_x \cap L_j(x)| \geq R\right] \\ &=: J_1 + J_2 \end{aligned} \tag{4.79}$$

By Proposition 4.21 we have

$$J_1 \leq n_0 \exp(-t_0 n_0) \stackrel{(4.78)}{\leq} \frac{2t_0 \lambda + \log R}{\lambda} R^{-\frac{1}{\lambda}} \leq_\lambda R^{-\frac{1}{\lambda}} \log R. \tag{4.80}$$

Taking a small constant $c > 0$ and $h_j = \frac{c}{(|j-n_0| \vee 1)^2}$ such that $\sum_{j=-\infty}^{n_0} h_j < \sum_{j=-\infty}^{\infty} h_j \leq 1$, using the union bound we get

$$J_2 \leq \sum_{j=-\infty}^{n_0} \mathbb{P}[e^{t_0(j+1)\lambda} |T_x \cap L_j(x)| \geq R h_j]. \tag{4.81}$$

Taking a positive integer $k = \lceil \frac{1}{\lambda} \rceil + 1 > \frac{1}{\lambda}$, by Markov's inequality one has

$$J_2 \leq \sum_{j=-\infty}^{n_0} \frac{1}{R^k h_j^k} \mathbb{E}[e^{t_0(j+1)\lambda k} |T_x \cap L_j(x)|^k] \tag{4.82}$$

By Proposition 4.20 we have

$$\begin{aligned}
 J_2 &\preceq_\lambda \sum_{j=-\infty}^0 \frac{e^{t_0(j+1)\lambda k}}{R^k h_j^k} (|j \vee 1|)^{2k} + \sum_{j=1}^{n_0} \frac{e^{t_0(j+1)\lambda k}}{R^k h_j^k} j \exp(-t_0 j) \\
 &\preceq_\lambda \frac{1}{R^k} \sum_{j=-\infty}^0 e^{t_0 \lambda k j} (|j \vee 1|)^{2k} (|j - n_0| \vee 1)^{2k} + \frac{1}{R^k} \sum_{j=1}^{n_0} j (|j - n_0| \vee 1)^{2k} e^{t_0(\lambda k - 1)j} \\
 &\preceq_\lambda \frac{1}{R^k} \sum_{j=-\infty}^0 e^{t_0 \lambda k j} (|j \vee 1|)^{4k} n_0^{2k} + \frac{n_0 e^{t_0(\lambda k - 1)n_0}}{R^k} + \frac{n_0}{R^k} \sum_{j=1}^{n_0 - 1} (n_0 - j)^{2k} e^{t_0(\lambda k - 1)j} \\
 &\preceq_\lambda \frac{n_0^{2k}}{R^k} + n_0 e^{-t_0 n_0} + \frac{n_0 e^{t_0(\lambda k - 1)n_0}}{R^k} \int_0^{n_0} (n_0 - x)^{2k} e^{t_0(\lambda k - 1)(x - n_0)} dx \\
 &\preceq_\lambda \frac{(\log R)^{2+2\lceil \frac{1}{\lambda} \rceil}}{R^{\frac{1}{\lambda}+1}} + n_0 e^{-t_0 n_0} + n_0 e^{-t_0 n_0} \cdot \int_0^\infty y^{2k} e^{-t_0(\lambda k - 1)y} dy \\
 &\preceq_\lambda \frac{(\log R)^{2+2\lceil \frac{1}{\lambda} \rceil}}{R^{\frac{1}{\lambda}+1}} + n_0 e^{-t_0 n_0} + n_0 e^{-t_0 n_0} \\
 &\preceq_\lambda R^{-\frac{1}{\lambda}} \log R,
 \end{aligned} \tag{4.83}$$

where in the fourth to the sixth inequalities we use $R^k \asymp_\lambda e^{t_0 \lambda k n_0}$ by the choice of n_0 .

Therefore

$$\mathbb{P}[|T_x|_{x,\lambda} \geq R] \leq J_1 + J_2 \preceq_\lambda R^{-\frac{1}{\lambda}} \log R. \quad \square$$

Since the tree \mathfrak{T}_x in the x -WUSF is almost surely finite, $|\mathfrak{T}_x|_{x,\lambda} < \infty$ for all $\lambda \in \mathbb{R}$.

Proposition 4.25. *The tail probability $\mathbb{P}[|\mathfrak{T}_x|_{x,\lambda} \geq R]$ satisfies the following inequalities.*

- If $\lambda < 0$, then

$$\mathbb{P}[|\mathfrak{T}_x|_{x,\lambda} \geq R] \asymp_\lambda \frac{1}{\log R}, \quad R \geq 2. \tag{4.84}$$

- If $\lambda = 0$, then $|\mathfrak{T}_x|_{x,\lambda} = |\mathfrak{T}_x|$ is the size of the tree \mathfrak{T}_x and

$$\mathbb{P}[|\mathfrak{T}_x| \geq R] \asymp \frac{1}{\sqrt{R}}, \quad R \geq 2. \tag{4.85}$$

- If $\lambda > 0$, then

$$\mathbb{P}[|\mathfrak{T}_x|_{x,\lambda} \geq R] \asymp_\lambda R^{-\frac{1}{\lambda}}, \quad R \geq 2. \tag{4.86}$$

Proof. We only prove (4.84) since (4.85) was proved to hold for more general networks in Theorem 7.1 of [11] and the proof of (4.86) is similar to the ones in Proposition 4.24.

Now we assume $\lambda < 0$ and $R \geq 2$. Taking $n_0 = \lceil \frac{\log R}{-t_0 \lambda} + 1 \rceil$, then $\exp(-t_0(n_0 - 1)\lambda) \geq R$. Hence if $\mathfrak{T}_x \cap L_{-n_0}(x) \neq \emptyset$, then $|\mathfrak{T}_x|_{x,\lambda} \geq \exp(-t_0(n_0 - 1)\lambda) \geq R$. Therefore by Proposition 4.22 one has that

$$\mathbb{P}[|\mathfrak{T}_x|_{x,\lambda} \geq R] \geq \mathbb{P}[\mathfrak{T}_x \cap L_{-n_0}(x) \neq \emptyset] \succeq \frac{1}{n_0} \succeq_\lambda \frac{1}{\log R}.$$

On the other hand, there exists a constant $C > 0$ small enough such that for all $R \geq 2D$

$$\sum_{y \in B(x, C \log R)} \Delta(x, y)^\lambda < D^{C \log R + 1} \exp(-t_0 \lambda C \log R) \leq R.$$

Therefore, for $R \geq 2D$

$$\mathbb{P}[|\mathfrak{T}_x|_{x,\lambda} \geq R] \leq \mathbb{P}[\text{diam}_{\text{int}}(\mathfrak{T}_x) \geq C \log R] \preceq \frac{1}{\log R},$$

where the last inequality is again due to Theorem 7.1 of [11]. □

Remark 4.26. As $\lambda \rightarrow 0$, the tilted volume $|\mathfrak{T}_x|_{x,\lambda} \rightarrow |\mathfrak{T}_x|$. So for properly related λ and R , the probabilities $\mathbb{P}[|\mathfrak{T}_x|_{x,\lambda} \geq R]$ and $\mathbb{P}[|\mathfrak{T}_x| \geq R]$ should be close. This is due to the dependence of λ implicitly in (4.84) and (4.86). Indeed if $\lambda = -\frac{\log R}{\sqrt{R}}$, then $\mathbb{P}[|\mathfrak{T}_x|_{x,\lambda} \geq R] \geq \frac{1}{n_0} \geq \frac{1}{\sqrt{R}}$, where $n_0 = \lceil \frac{\log R}{-t_0 \lambda} + 1 \rceil$. Moreover for $\lambda = -\frac{\log R}{\sqrt{R}}$, let $j_0 = \frac{\log 2}{t_0 \lambda}$, if $j \geq j_0$, then $\exp(t_0 j \lambda) \leq 2$. Thus $\mathbb{P}[|\mathfrak{T}_x|_{x,\lambda} \geq R] \leq \mathbb{P}[\mathfrak{T}_x \cap L_{j_0}(x) \neq \emptyset] + \mathbb{P}[2|\mathfrak{T}_x| \geq R] \leq \frac{1}{|j_0|} + \frac{1}{\sqrt{R}} \leq \frac{\log R}{\sqrt{R}}$.

For the future and past of x in the WUSF sample \mathfrak{F} , one can also consider the tail probability for the titled volumes. We summarize the results in the following proposition and omit the proofs since they are similar to the ones for T_x and \mathfrak{T}_x .

Proposition 4.27. *If $\lambda > 0$, then*

$$\mathbb{P}[|\mathfrak{F}(x, \infty)|_{x,\lambda} \geq R] \asymp_\lambda R^{-\frac{1}{\lambda}}$$

and

$$\mathbb{P}[|\mathfrak{P}(x)|_{x,\lambda} \geq R] \asymp_\lambda R^{-\frac{1}{\lambda}}.$$

If $\lambda = 0$, then

$$\mathbb{P}[|\mathfrak{P}(x)| \geq R] \asymp \frac{1}{\sqrt{R}}.$$

If $\lambda < 0$, then

$$\mathbb{P}[|\mathfrak{P}(x)|_{x,\lambda} \geq R] \asymp_\lambda \frac{1}{\log R}.$$

5 FUSF on Diestel-Leader graphs and grandparent graphs

5.1 FUSF=WUSF on Diestel-Leader graphs

Diestel-Leader graphs are a family of transitive graphs constructed by Diestel and Leader in [4] as possible examples of transitive graphs that are not roughly isometric to any Cayley graph. Later Eskin, Fisher, and Whyte [6] showed that the Diestel-Leader graph $DL(q, r)$ with $q \neq r$ is indeed not roughly isometric to any Cayley graph.

Next we give a precise definition of the Diestel-Leader graph $DL(q, r)$. For more details see [30]. Given a regular tree \mathbb{T}_{b+1} with $b \geq 2$, fix a root $o \in \mathbb{T}_{b+1}$ and an end ξ . For each vertex $v \in \mathbb{T}_{b+1}$, there is a unique ray η_v starting from v and representing ξ . We call the unique neighbor of v on the ray η_v the **parent** of v with respect to ξ , and denote it by v^- . For $x, y \in \mathbb{T}_{b+1}$, we define $x \wedge y$ (w.r.t. ξ) to be the first intersection vertex of η_x and η_y . We define the **Busemann function** $\mathfrak{h} : \mathbb{T}_{b+1} \rightarrow \mathbb{Z}$ with respect to o, ξ as follows:

$$\mathfrak{h}(x) = d(o, x \wedge o) - d(x, x \wedge o),$$

where $d(u, v)$ is the graph distance between u and v on \mathbb{T}_{b+1} . We also define the **horocycles** H_k with respect to o, ξ as $H_k = H_k(o, \xi) := \{x \in \mathbb{T}_{b+1} : \mathfrak{h}(x) = k\}$. Note that changing the base o will only change the Busemann function by adding a constant.

Suppose $q \geq r \geq 2$ are two positive integers and $\mathbb{T}_{q+1}, \mathbb{T}_{r+1}$ are two regular trees with degree $q + 1, r + 1$ respectively. Fix roots o_1, o_2 and reference ends ω_1, ω_2 for $\mathbb{T}_{q+1}, \mathbb{T}_{r+1}$ respectively.

Definition 5.1. *The set of vertices of Diestel-Leader graph $DL(q, r)$ is given by*

$$DL(q, r) = \{x_1 x_2 \in \mathbb{T}_{q+1} \times \mathbb{T}_{r+1} : \mathfrak{h}(x_1) + \mathfrak{h}(x_2) = 0\},$$

where we use \mathfrak{h} as the Busemann functions on $\mathbb{T}_{q+1}, \mathbb{T}_{r+1}$ w.r.t. to o_1, ω_1 and o_2, ω_2 .

The neighborhood relation is given by

$$x_1 x_2 \sim y_1 y_2 \text{ if and only if } x_1 \sim y_1 \text{ and } y_1 \sim y_2.$$

A way of visualizing of $DL(q, r)$ is described on page 418 of [30]; see also Figure 2 there for an example $DL(2, 2)$.

Let P denote the transition operator corresponding to simple random walk on $DL(q, r)$. Namely for $x_1x_2 \in DL(q, r)$, $p(x_1x_2, y_1y_2) = \frac{\mathbf{1}_{\{x_1x_2 \sim y_1y_2\}}}{q+r}$ and for functions $h : DL(q, r) \rightarrow \mathbb{R}$, $Ph(x_1x_2) = \sum_{y_1y_2} p(x_1x_2, y_1y_2)h(y_1y_2)$.

A (P -)harmonic function h is the one that satisfies $Ph = h$.

Let P_1, P_2 denote the projection of P on \mathbb{T}_{q+1} and \mathbb{T}_{r+1} respectively:

$$p_1(x_1, y_1) = \begin{cases} \frac{1}{q+r} & \text{if } y_1^- = x_1 \\ \frac{r}{q+r} & \text{if } y_1 = x_1^- \\ 0 & \text{otherwise,} \end{cases} \quad , \quad p_2(x_2, y_2) = \begin{cases} \frac{1}{q+r} & \text{if } y_2^- = x_2 \\ \frac{q}{q+r} & \text{if } y_2 = x_2^- \\ 0 & \text{otherwise.} \end{cases}$$

Woess proved the following decomposition theorem about positive harmonic functions.

Theorem 5.2 (Theorem 2.3 of [30]). *If h is a non-negative P -harmonic function on $DL(q, r)$, then there are non-negative P_i -harmonic functions $h_i, i = 1, 2$ on \mathbb{T}_{q+1} and \mathbb{T}_{r+1} respectively, such that*

$$h(x_1x_2) = h_1(x_1) + h_2(x_2), \quad \forall x_1x_2 \in DL(q, r)$$

Proposition 5.3. *FUSF is the same as WUSF on Diestel–Leader graphs.*

Proof. By Theorem 7.3 of [3], it suffices to show that every harmonic Dirichlet function on $DL(q, r)$ is a constant function and this is an easy consequence of Theorem 5.2.

In fact, suppose there are non-constant harmonic Dirichlet functions on $DL(q, r)$. Then there are non-constant bounded harmonic Dirichlet functions on $DL(q, r)$ (Theorem 3.73 of [23]), whence there are non-constant non-negative harmonic Dirichlet functions on $DL(q, r)$.

Let h be a non-constant non-negative harmonic Dirichlet function on $DL(q, r)$. By Theorem 5.2, there exist non-negative functions h_1 and h_2 on \mathbb{T}_{q+1} and \mathbb{T}_{r+1} such that $h(x_1x_2) = h_1(x_1) + h_2(x_2), \forall x_1x_2 \in DL(q, r)$.

Since h is not a constant function, at least one of h_1, h_2 is also not a constant. Without loss of generality, we assume that h_1 is not a constant. Suppose $x_1, y_1 \in \mathbb{T}_{q+1}$ are two neighboring vertices such that $y_1^- = x_1$ and $h_1(x_1) \neq h_1(y_1)$. We first show that for any $z_1 \in \mathbb{T}_{q+1}$ such that $z_1^- = x_1$, one has $h_1(z_1) = h_1(y_1)$. Suppose $z_1 \neq y_1$.

Since h is a harmonic Dirichlet function,

$$\begin{aligned} \infty &> \sum_{x_2 \in \mathbb{T}_{r+1}: \mathfrak{h}(x_2) = -\mathfrak{h}(x_1)} |h(x_1x_2) - h(y_1x_2^-)|^2 + |h(x_1x_2) - h(z_1x_2^-)|^2 \\ &\geq \sum_{x_2 \in \mathbb{T}_{r+1}: \mathfrak{h}(x_2) = -\mathfrak{h}(x_1)} \frac{1}{2} |h(y_1x_2^-) - h(z_1x_2^-)|^2 \\ &= \sum_{x_2 \in \mathbb{T}_{r+1}: \mathfrak{h}(x_2) = -\mathfrak{h}(x_1)} \frac{1}{2} |h_1(y_1) - h_1(z_1)|^2. \end{aligned} \tag{5.1}$$

Since there are infinitely many $x_2 \in \mathbb{T}_{r+1}$ such that $\mathfrak{h}(x_2) = -\mathfrak{h}(x_1)$, we have $h_1(y_1) = h_1(z_1)$. From this and the fact that h_1 is P_1 harmonic, one has that h_1 is constant on each horocycle of \mathbb{T}_{q+1} . Similarly h_2 is also a constant on each horocycle of \mathbb{T}_{r+1} , whence $h(x_1x_2)$ only depends on which horocycle x_1 lies in. Then h must be a constant function on $DL(q, r)$ to have finite Dirichlet energy. This contradicts with the choice that h is a non-constant function. \square

Now by Proposition 5.3, the study of FUSF on $DL(qr)$ reduces to WUSF. In particular we know that each component of FUSF on $DL(q, r)$ with $q > r \geq 2$ is one-ended and light almost surely.

5.2 FUSF on grandparent graphs

We first recall the definition of grandparent graphs. For more details see Section 7.1 of [18].

Consider a regular tree \mathbb{T}_{b+1} with degree $b + 1 \geq 3$. Let ξ be a fixed end of \mathbb{T}_{b+1} . As in the previous subsection, for each v in \mathbb{T}_{b+1} , there is a unique ray $\eta_v = (v_0, v_1, v_2, \dots)$ that represents the end ξ starting at $v_0 = v$. We call v_2 the ξ -**grandparent** of v . Throughout this subsection we let G be the graph obtained from \mathbb{T}_{b+1} by adding the edges (v, v_2) between v and its ξ -grandparent for all $v \in \mathbb{T}_{b+1}$. It is well known that G is a nonunimodular transitive graph. For two vertices x, y in G , we denote by $d_G(x, y), d_{\mathbb{T}}(x, y)$ the graph distance of x, y in G and \mathbb{T}_{b+1} respectively.

Fix a base point v and let $\eta_v = (v_0, v_1, v_2, \dots)$ be the unique ray that represents the end ξ starting at $v_0 = v$. We consider the following exhaustion of G . For $n \geq 1$, let G_n be the subgraph of G induced by vertices $\{x : d_{\mathbb{T}}(x, v) \leq n\}$.

For $k, n \geq 1$, let $\mathcal{P}_{k,n}$ denote the set of self-avoiding paths that connect v_0, v_1 in G_n with length k .

We start with an estimate on the size of $\mathcal{P}_{k,n}$.

Lemma 5.4. *If $1 \leq k \leq n + 1$, then $|\mathcal{P}_{k,n}| \asymp b^k$. If $k \geq n + 2$, then $\mathcal{P}_{k,n} = \emptyset$.*

Proof. For a self-avoiding path $\pi = (w_0, w_1, \dots, w_k)$ in G , if $\Delta(w_i, w_{i+1}) = b^{-2}$, then we say the step from w_i to w_{i+1} is **downward** 2 levels, and denote it by $w_i \xrightarrow{-2} w_{i+1}$. Similarly we define downward 1 level, upward 1 level and upward 2 levels. If $w_i \xrightarrow{-1} w_{i+1}$ or $w_i \xrightarrow{+1} w_{i+1}$, then the edge $e = (w_i, w_{i+1})$ is an edge in \mathbb{T}_{b+1} , we call it a **tree** edge; if $w_i \xrightarrow{-2} w_{i+1}$ or $w_i \xrightarrow{+2} w_{i+1}$, then the edge $e = (w_i, w_{i+1})$ is an edge connecting a vertex to its grandparent, we call it a **grandparent** edge.

The following relations are obvious:

- if $w_{i-1} \xrightarrow{+1} w_i$, then w_i is the parent of w_{i-1} ;
- if $w_{i-1} \xrightarrow{+2} w_i$, then w_i is the grandparent of w_{i-1} ;
- if $w_{i-1} \xrightarrow{-1} w_i$, then w_{i-1} is the parent of w_i ;
- if $w_{i-1} \xrightarrow{-2} w_i$, then w_{i-1} is the grandparent of w_i .

Observation 5.5. *If the path $\pi = (w_0, w_1, \dots, w_k) \in \mathcal{P}_{k,n}$ from v_0 to v_1 has visited a vertex x and its parent x^- by time i , then the remaining part of the path (w_{i+1}, \dots, w_k) must lie in the connected component of v_1 in the induced subgraph of $G \setminus \{x, x^-\}$. This is because $\{x, x^-\}$ is a cutset for G and π is self-avoiding.*

Claim 5.6. *A self-avoiding path $\pi = (w_0, w_1, \dots, w_k)$ in G connecting v_0 and v_1 has one of the following forms. If k is even, then exactly one of the following statements holds. See Figure 2 and 3 for an illustration of typical paths of each form.*

- (a) π has the form $w_0 \xrightarrow{-2} \dots \xrightarrow{-2} w_{t-1} \xrightarrow{-1} w_t \xrightarrow{+2} w_{t+1} \xrightarrow{+2} \dots \xrightarrow{+2} w_{2t+2}$. Here $w_0 = v_0$ and $w_{2t+2} = v_1$.
- (b) π has the form $w_0 \xrightarrow{+2} \dots \xrightarrow{+2} w_t \xrightarrow{-1} w_{t+1} \xrightarrow{-2} \dots \xrightarrow{-2} w_{2t}$.
- (b') π has the form $w_0 \xrightarrow{+2} \dots \xrightarrow{+2} w_t \xrightarrow{-1} w_{t+1} \xrightarrow{+2} w_{t+2} \xrightarrow{-2} \dots \xrightarrow{-2} w_{2t+2}$, where w_{t+1} and w_{t+3} are different children of w_{t+2} .
- (c) For some positive integers α, β such that $k = 2 + 2\alpha + 2\beta$, π has the form $w_0 \xrightarrow{+2} \dots \xrightarrow{+2} w_\alpha \xrightarrow{-2} w_{\alpha+1} \xrightarrow{-2} \dots \xrightarrow{-2} w_{\alpha+\beta} \xrightarrow{-1} w_{\alpha+\beta+1} \xrightarrow{+2} \dots \xrightarrow{+2} w_{\alpha+2\beta+2} \xrightarrow{-2} \dots \xrightarrow{-2} w_{2+2\alpha+2\beta}$, where $w_{\alpha-1}$ and $w_{\alpha+1}$ have different parents.

(c') For some positive integers α, β such that $k = 2\alpha + 2\beta$, π has the form $w_0 \xrightarrow{+2} \dots \xrightarrow{+2} w_\alpha \xrightarrow{-2} w_{\alpha+1} \xrightarrow{-2} \dots \xrightarrow{-2} w_{\alpha+\beta} \xrightarrow{-1} w_{\alpha+\beta+1} \xrightarrow{+2} \dots \xrightarrow{+2} w_{\alpha+2\beta+1} \xrightarrow{-2} \dots \xrightarrow{-2} w_{2\alpha+2\beta}$, where $w_{\alpha-1}$ and $w_{\alpha+1}$ have the same parent (e.g. $w_{\alpha+2\beta+1}$).

If k is odd, then exactly one of the following statements holds.

(d) π has the form $w_0 \xrightarrow{-2} \dots \xrightarrow{-2} w_t \xrightarrow{+1} w_{t+1} \xrightarrow{+2} w_{t+2} \xrightarrow{+2} \dots \xrightarrow{+2} w_{2t+1}$.

(e) π has the form $w_0 \xrightarrow{+2} \dots \xrightarrow{+2} w_t \xrightarrow{+1} w_{t+1} \xrightarrow{-2} \dots \xrightarrow{-2} w_{2t+1}$.

(f) For some positive integers α, β such that $k = 1 + 2\alpha + 2\beta$, π has the form $w_0 \xrightarrow{+2} \dots \xrightarrow{+2} w_\alpha \xrightarrow{-2} w_{\alpha+1} \xrightarrow{-2} \dots \xrightarrow{-2} w_{\alpha+\beta} \xrightarrow{+1} w_{\alpha+\beta+1} \xrightarrow{+2} \dots \xrightarrow{+2} w_{\alpha+2\beta+1} \xrightarrow{-2} \dots \xrightarrow{-2} w_{1+2\alpha+2\beta}$, where $w_{\alpha-1}$ and $w_{\alpha+1}$ have different parents.

(f') For some positive integers α, β such that $k = 2\alpha + 2\beta - 1$, π has the form $w_0 \xrightarrow{+2} \dots \xrightarrow{+2} w_\alpha \xrightarrow{-2} w_{\alpha+1} \xrightarrow{-2} \dots \xrightarrow{-2} w_{\alpha+\beta} \xrightarrow{+1} w_{\alpha+\beta+1} \xrightarrow{+2} \dots \xrightarrow{+2} w_{\alpha+2\beta} \xrightarrow{-2} \dots \xrightarrow{-2} w_{2\alpha+2\beta-1}$, where $w_{\alpha-1}$ and $w_{\alpha+1}$ have the same parent (e.g. $w_{\alpha+2\beta}$).

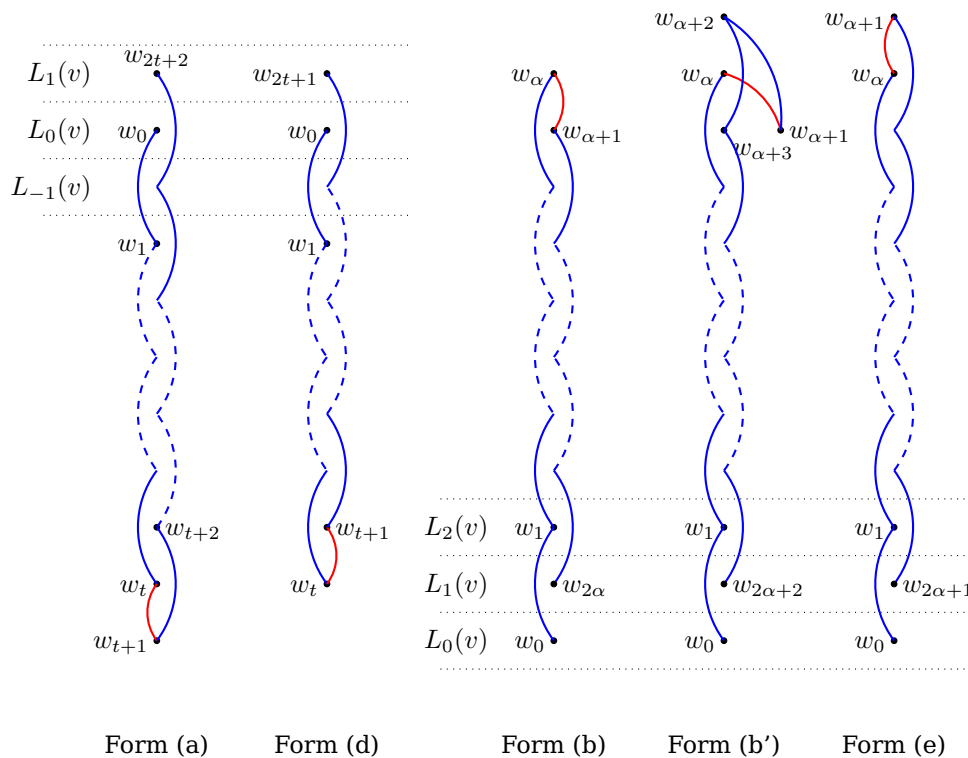


Figure 2: A systematic drawing of the five classes (a),(d),(b),(b'),(e) of paths for $\mathcal{P}_{k,n}$, where $v = w_0 = v_0$.

Proof of Claim 5.6. The proof of this claim is a case-by-case study and repeated applications of Observation 5.5 with appropriate choices of x, x^- .

Suppose $\pi = (w_0, \dots, w_k)$ is a self-avoiding path connecting v_0 and v_1 with length k . By parity π must use at least one tree edge.

Case one: (First step is downward 2 levels)

If the first step of π is downward 2 levels, i.e. $w_0 \xrightarrow{-2} w_1$, then the next step cannot be upward 2 levels, otherwise it will not be self-avoiding. By the same reason the path can only go downward 2 levels each step before encountering the tree edge. Let t be the number of steps before encountering a tree edge, i.e., $w_0 \xrightarrow{-2} \dots \xrightarrow{-2} w_t$.

Now look at the next step $w_t \rightarrow w_{t+1}$, it must be a tree edge by the choice of t . If $w_t \xrightarrow{-1} w_{t+1}$, then apply Observation 5.5 with $x = w_{t+1}$ and $x^- = w_t$ since $w_t \xrightarrow{-1} w_{t+1}$. Note that v_1 and the descendants of w_{t+1} are in different connected component of $G \setminus \{w_t, w_{t+1}\}$. Thus the next step must be $w_{t+1} \xrightarrow{+2} w_{t+2}$. Apply Observation 5.5 with $x = w_{t+2}$ and $x^- = w_{t+1}$, the next step must be $w_{t+2} \xrightarrow{+2} w_{t+3}$. Continue in this way and then π must have form (a).

Using similar reasoning, if $w_t \xrightarrow{+1} w_{t+1}$, to avoid cycle on π the remaining steps after time $t + 1$ can only be upward 2 levels until reaching v_1 .

In sum if the first step of π is downward 2 levels, then π must be of the form (a) or (d).

Case two: (First step is upward 2 levels)

Suppose the first step of π is upward 2 levels, i.e. $w_0 \xrightarrow{+2} w_1$. Let α be the number of steps of π which is upward 2 levels until another type of step is encountering, i.e. $w_0 \xrightarrow{+2} \dots \xrightarrow{+2} w_\alpha$ but $w_\alpha \rightarrow w_{\alpha+1}$ is not upward 2 levels.

Sub-case 2(a) Suppose the step $w_\alpha \rightarrow w_{\alpha+1}$ is a tree edge.

If $w_\alpha \xrightarrow{+1} w_{\alpha+1}$, then applying Observation 5.5 with $x = w_\alpha, x^- = w_{\alpha+1}$ it must be the case that $w_{\alpha+1} \xrightarrow{-2} w_{\alpha+2}$ and $w_{\alpha+2}$ is an ancestor of v_1 . So on and so forth, the remaining steps can only be downward 2 levels until hitting v_1 . Thus in this sub-case the path π must be of the form (e).

If $w_\alpha \xrightarrow{-1} w_{\alpha+1}$ and $w_{\alpha+1}$ is an ancestor of v_1 (i.e. $w_{\alpha+1} = w_{\alpha-1}^-$), then applying Observation 5.5 with $x = w_{\alpha-1}, x^- = w_{\alpha+1}$, it must be the case that $w_{\alpha+1} \xrightarrow{-2} w_{\alpha+2}$ and $w_{\alpha+2}$ is an ancestor of v_1 . So on and so forth, the remaining steps can only be downward 2 levels until hitting v_1 . Thus in this sub-case the path π must be of the form (b).

If $w_\alpha \xrightarrow{-1} w_{\alpha+1}$ and $w_{\alpha+1}$ is not an ancestor of v_1 (i.e. $w_{\alpha+1} \neq w_{\alpha-1}^-$), then applying Observation 5.5 with $x = w_{\alpha+1}, x^- = w_\alpha$, it must be the case that $w_{\alpha+1} \xrightarrow{+2} w_{\alpha+2}$. Next applying Observation 5.5 with $x = w_\alpha, x^- = w_{\alpha+2}$, it must be the case that $w_{\alpha+2} \xrightarrow{-2} w_{\alpha+3}$ and $w_{\alpha+3}$ is an ancestor of v_1 (i.e. $w_{\alpha-1}^- = w_{\alpha+3}$). Applying Observation 5.5 with $x = w_{\alpha-1}, x^- = w_{\alpha+3}$, it must be the case that $w_{\alpha+3} \xrightarrow{-2} w_{\alpha+4}$ and $w_{\alpha+4}$ is an ancestor of v_1 . So on and so forth, the remaining steps can only be downward 2 levels until hitting v_1 . Thus in this sub-case the path π must be of the form (b').

Sub-case 2(b) Suppose the step $w_\alpha \rightarrow w_{\alpha+1}$ is downward 2 levels. To avoid cycle, $w_{\alpha+1}$ must be a different grandchildren of w_α other than $w_{\alpha-1}$. Similar as in the first case, from the time α the path can only go downward 2 levels each step before encountering a tree edge, say there are β such steps. Now the path looks like $w_0 \xrightarrow{+2} \dots \xrightarrow{+2} w_\alpha \xrightarrow{-2} w_{\alpha+1} \xrightarrow{-2} \dots \xrightarrow{-2} w_{\alpha+\beta} \rightarrow \dots$. By the choice of β , the next step $w_{\alpha+\beta} \rightarrow w_{\alpha+\beta+1}$ is a tree edge. Using the same reasoning as in the first case, after time $\alpha + \beta + 1$ the path can only be upward 2 levels each step until it hits the ray η_v . Then from that hitting vertex, to avoid cycles the path can only go downward 2 levels each step until hitting v_1 . Thus in this sub-case the path π must be of the form (c) or (c') or (f) or (f') depending on the parity of the length of π and whether $w_{\alpha-1}$ and $w_{\alpha+1}$ has the same parent. We leave the details of this **Sub-case 2(b)** to the reader.

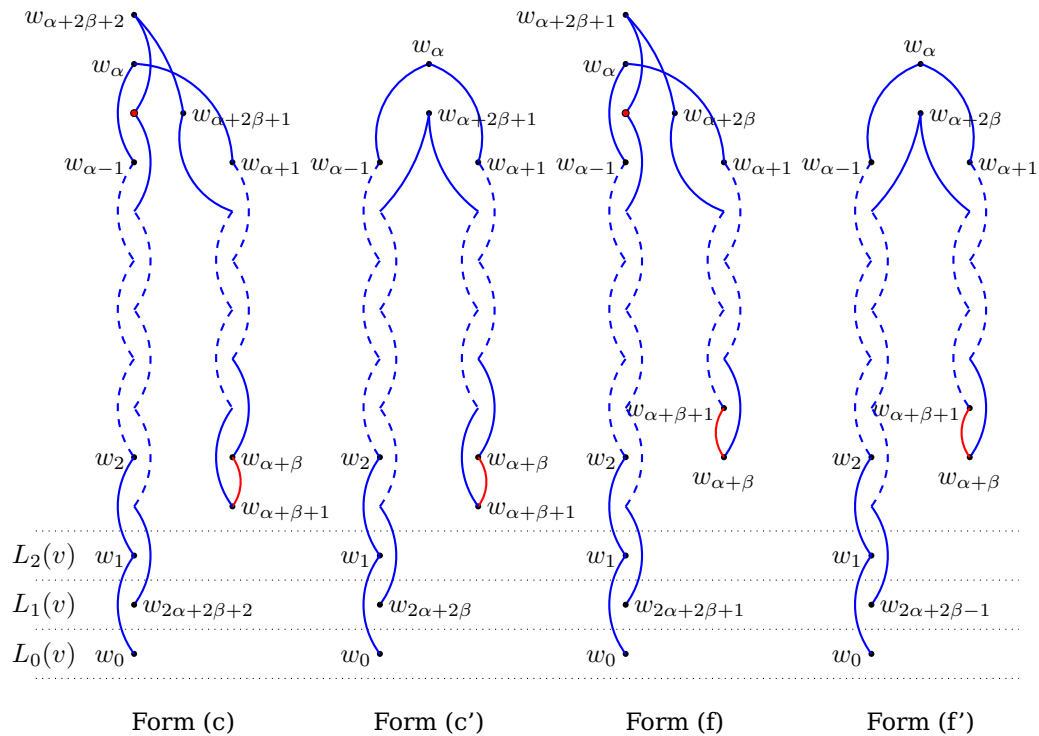


Figure 3: A systematic drawing of the four classes (c),(c'),(f),(f') of paths for $\mathcal{P}_{k,n}$, where $v = w_0 = v_0$. In form (c) and (f), the parent of $w_{\alpha-1}$ is denoted by relatively large and red node.

In particular from the above analysis, we see that each self-avoiding path π that connects v_0, v_1 uses exactly one tree edge. \square

Also from the forms of the path in Claim 5.6, we see that if a self-avoiding path $\pi = (w_0, \dots, w_k)$ connecting v_0 and v_1 with length k is one of the forms (a), (b'),(c), (d) and (f), then

$$\max\{d_{\mathbb{T}}(w_0, w_i) : 1 \leq i \leq k\} = (k - 1) \vee 1;$$

If π is of the form (c') or (f'), then

$$\max\{d_{\mathbb{T}}(w_0, w_i) : 1 \leq i \leq k\} = k + 1.$$

If π is of the form (b) or (e), then

$$\max\{d_{\mathbb{T}}(w_0, w_i) : 1 \leq i \leq k\} = k.$$

Therefore, if $k \geq n + 2$ then $\mathcal{P}_{k,n} = \emptyset$.

Suppose k is even and $k = 2t + 2$. The number of paths of the form (a) is $b^{2t+1} = b^{k-1}$ because for the first t steps one has b^2 choices for the grandchildren and b choices for w_{t+1} . Once w_0, \dots, w_{t+1} are fixed, the remaining vertices w_{t+2}, \dots, w_{2t+2} are fixed. The number of paths of the form (b) is just 1. The number of paths of the form (b') is $b - 1$ since one has $b - 1$ choices for $w_{\alpha+1}$. The number of paths of the form (c) is $\sum_{\beta=1}^{t-1} (b - 1)b^{2\beta}$ (Given $\alpha, \beta, w_0, \dots, w_{\alpha}$ are fixed, one has $b(b - 1)$ choices for $w_{\alpha+1}$, and b^2

choices for each of $w_{\alpha+2}, \dots, w_{\alpha+\beta}$ and b choices for $w_{\alpha+\beta+1}$. Once $w_0, \dots, w_{\alpha+\beta+1}$ are fixed, the remaining path is fixed). Using similar counting argument, we have that the number of paths of the form (c') is $\sum_{\beta=1}^t (b-1)b^{2\beta-1}$.

In sum if $k \in [1, n)$ is even, then $|\mathcal{P}_{k,n}| = b^{k-1} + 1 + \sum_{\beta=1}^{(k-4)/2} (b-1)b^{2\beta} + \sum_{\beta=1}^{(k-2)/2} (b-1)b^{2\beta-1} \asymp b^k$.

Similarly if $k \in [1, n)$ is odd, one also has $|\mathcal{P}_{k,n}| \asymp b^k$.

If $k = n$ or $n + 1$, the estimate $|\mathcal{P}_{k,n}| \asymp b^k$ is also true because we only need to deduct the contributions of cases (b), (c'), (e) and (f') from the previous expression for $|\mathcal{P}_{k,n}|$. \square

Next we recall some notation and results regarding loop-erased random walk on a graph from [14].

Let $S(t)$ be a discrete time Markov chain on a countable state space X with transition probabilities $p(x, y)$. For a subset A of X , define the hitting time

$$\tau_A := \inf\{t \geq 0 : S(t) \in A\}$$

and the Green function

$$G(x, A) := \sum_{j=0}^{\infty} \mathbf{P}^x[S(j) = x; S(t) \notin A, t = 0, \dots, j],$$

where \mathbf{P}^x denote the measure of the Markov chain $S(t)$ started from x .

Fix a base point $o \in X$ and we assume that $S(t)$ is irreducible. Let $B \subset X$ be a subset with $\mathbf{P}^o[\tau_B < \infty] = 1$. Suppose $o \notin B$, start the Markov chain at o , let it run until it hits B , and then erase the loops chronologically. Then we get a probability measure μ on the self-avoiding paths from o to B . In particular, if $w = [w_0, \dots, w_k]$ is such a self-avoiding path from o to B with $w_0, \dots, w_{k-1} \notin B$, then

$$\mu(w) = \left[\prod_{j=1}^k p(w_{j-1}, w_j) \right] \cdot \left[\prod_{j=0}^{k-1} G(w_j, B \cup A_{j-1}) \right], \tag{5.2}$$

where $A_{-1} = \emptyset, A_j = \{w_0, \dots, w_j\}$ for $j \geq 1$ (see Proposition 3.2 of [15]).

Proposition 5.7 (Proposition 3.3 of [15]). *Suppose S is irreducible, $B \neq \emptyset$ and that $w_0, \dots, w_{k-1} \notin B$, define*

$$f(w_0, \dots, w_{k-1}; B) = \prod_{j=0}^{k-1} G(w_j, B \cup A_{j-1}),$$

where $A_{-1} = \emptyset$ and $A_j = \{w_0, \dots, w_j\}$. Then $f(w_0, \dots, w_{k-1}; B)$ is a symmetric function of w_0, \dots, w_{k-1} .

Proposition 5.8. *The FUSF on the grandparent graph G is connected almost surely.*

Proof. Recall G_n is the sub-graph of G induced by vertices $\{x : d_{\mathbb{T}}(x, v_0) \leq n\}$. Let T_n be a uniform spanning tree on G_n .

We will show that there exists positive constants c_{23}, c_{24} such that

$$\mathbb{P}[d_{T_n}(v_0, v_1) \geq k] \leq c_{24} \exp(-c_{23}k). \tag{5.3}$$

Start a simple random walk on G_n from v_0 and stop at the first hit of v_1 , then loop-erase this random walk path. Then the self-avoiding path we get has the law of the path

from v_0 to v_1 in T_n . Let $S(t)$ be the simple random walk on G_n , let o be v_0 and $B = \{v_1\}$. By (5.2)

$$\mathbb{P}[d_{T_n}(v_0, v_1) \geq k] = \sum_{j=k}^{\infty} \sum_{w \in \mathcal{P}_{j,n}} \mu(w). \tag{5.4}$$

By Lemma 5.4, (5.3) holds trivially for $k \geq n + 2$. In the following we assume $k \leq n + 1$.

Notice that in our case the first product in (5.2) equals $\prod_{j=0}^{k-1} \frac{1}{\deg_{G_n}(w_j)}$, which is again symmetric for w_0, \dots, w_{k-1} . Hence by Proposition 5.7, we have $\mu(w)$ is a symmetric function of w_0, \dots, w_{k-1} . Using the connection between effective resistance and Green function (see Proposition 2.1 of [18]) one has that for any reordering w'_0, \dots, w'_{k-1} of w_0, \dots, w_{k-1}

$$\mu(w) = \prod_{j=0}^{k-1} \mathcal{R}(w'_j \leftrightarrow B \cup A'_{j-1}), \tag{5.5}$$

where $A'_{-1} = \emptyset$, $A'_j = \{w'_0, \dots, w'_j\}$ and $\mathcal{R}(w'_j \leftrightarrow B \cup A'_{j-1})$ is the effective resistance from w'_j to $B \cup A'_{j-1}$ in G_n .

For $w = (w_0, \dots, w_k) \in \mathcal{P}_{k,n}$, since w has one of the forms listed in the proof of Lemma 5.4, using a case by case analysis it is easy to see that we can reorder w_0, \dots, w_{k-1} as w'_0, \dots, w'_{k-1} such that except at most 8 j 's in $\{0, \dots, k - 1\}$, one has w'_{j-2}, w'_{j-1} are the grandparent and parent of w'_j and the grandchildren of w'_j are also in G_n . For such an ordering and an index j such that w'_{j-2}, w'_{j-1} are the grandparent and parent of w'_j and the grandchildren of w'_j are in G_n , one has that

$$\mathcal{R}(w'_j \leftrightarrow B \cup A'_{j-1}) \leq \frac{b + 4}{b^2 + 4b + 8} < \frac{1}{b}.$$

Indeed, from the local structure one has that

$$\mathcal{R}(w'_j \leftrightarrow B \cup A'_{j-1}) \leq \mathcal{R}(w'_j \leftrightarrow \{w'_{j-2}, w'_{j-1}\}) \stackrel{\text{Fig.4}}{\leq} \frac{1}{1 + 1 + b \cdot \frac{1}{1 + \frac{1}{1+b/2}}} = \frac{b + 4}{b^2 + 4b + 8} < \frac{1}{b}.$$

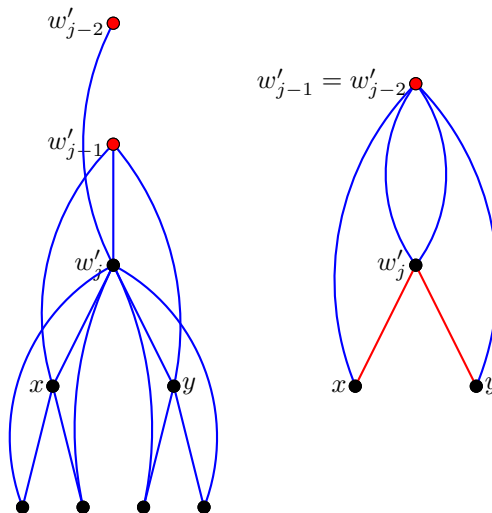


Figure 4: Local structure for estimating $\mathcal{R}(w'_j \leftrightarrow \{w'_{j-2}, w'_{j-1}\})$, the right half is a network reduction with red edges with conductance $1 + b/2$. All blue edges has conductance 1.

For the other j 's we use trivial estimates

$$\mathcal{R}(w'_j \leftrightarrow B \cup A'_{j-1}) \leq k + 1.$$

Now by (5.5) there exists a constant $c_{25} > 0$ such that

$$\max_{w \in \mathcal{P}_{k,n}} \mu(w) \leq (k + 1)^8 \left(\frac{b + 4}{b^2 + 4b + 8} \right)^{k-8} \leq c_{25} \left(\frac{b + 4}{b^2 + 4b + 7} \right)^k.$$

This combined with (5.4) and Lemma 5.4 yields (5.3).

Since G_n is an exhaustion of G , T_n converges weakly to the FUSF on G . Thus (5.3) implies that in the FUSF sample \mathfrak{F}^f on G , the probability that $d_{\mathfrak{F}^f}(v_0, v_1) \geq k$ decays exponentially in k . In particular, v_0 and v_1 are connected almost surely in \mathfrak{F}^f . By transitivity, almost surely any vertex of G is in the same connected component as its parent in \mathfrak{F}^f . Therefore the FUSF \mathfrak{F}^f is connected almost surely. \square

Now we know that the FUSF sample \mathfrak{F}^f on a grandparent graph G is just a tree. Next we consider the branching number of \mathfrak{F}^f .

Proposition 5.9. *The FUSF on a grandparent graph G has branching number strictly larger than one.*

Proof. From Proposition 5.8 we know that there is at least one tree edge in \mathfrak{F}^f , otherwise there would be at least two trees in \mathfrak{F}^f .

For $x \in V(G)$, let y be the parent of x and z be the grandparent of x . Let x_1, \dots, x_b be the children of x . Note that $\{x, y\}$ is a cutset for G , and $G \setminus \{x, y\}$ has $b + 1$ connected components, one containing z and other b ones each containing a unique child of x . We denote the connected component containing z by $K_0(x)$ and the connected components containing x_i by $K_i(x)$ for $i = 1, \dots, b$. Let $\widehat{K}_i(x)$ be the subgraph of G induced by $K_i(x) \cup \{x, y\}$ for $i = 0, 1, \dots, b$.

Conditioned on the event that the tree edge $e = (x, y) \in \mathfrak{F}^f$, one has the following observation:

$$\text{The conditional distributions of } \mathfrak{F}^f \cap \widehat{K}_i(x), i = 0, 1, \dots, b \text{ are independent.} \quad (5.6)$$

To see (5.6), consider an exhaustion of G by the balls $B(x, n)$. Sample UST on $B(x, n)$ using Wilson's algorithm with $F_0 = \{x\}$ and the first simple random walk X^y starting from y . Then the edge $e = (y, x)$ belongs to the uniform spanning tree $\text{UST}(B(x, n))$ if and only if the first time X^y hits x via the edge (y, x) . Suppose the event $e = (y, x) \in \text{UST}(B(x, n))$ occurs and continue Wilson's algorithm. Suppose we already obtain F_{k-1} and the next simple random walk starts from w . Since $\{x, y\}$ is a cutset for $B(x, n)$, the simple random walk X^w will hit F_{k-1} before visiting other connected component of $B(x, n) \setminus \{x, y\}$. Therefore the $b + 1$ trees $\text{UST}(B(x, n)) \cap \widehat{K}_i(x), i = 0, 1, \dots, b$ are independent conditioned on $e = (y, x) \in \text{UST}(B(x, n))$. By the definition of FUSF, one has (5.6).

Next we show that conditioned on the event $e = (x, y) \in \mathfrak{F}^f$, for each $i = 1, \dots, b$, almost surely there is another tree edge in $\mathfrak{F}^f \cap \widehat{K}_i(x)$. If not, we define a mass transport as follows:

$$f(u, v) := \mathbf{1}_{\{v \text{ is the nearest ancestor of } u \text{ such that } (v, v^-) \in \mathfrak{F}^f\}},$$

where v^- denotes the parent of v .

Then the mass sent out from a vertex is at most one. But if conditioned on the event $e = (x, y) \in \mathfrak{F}^f$, with positive probability there is no other tree edge in $\mathfrak{F}^f \cap \widehat{K}_i(x)$, then x will receive infinite mass with positive probability. This contradicts with the tilted mass-transport principle.

Thus there is a large constant $M > 0$ such that the following inequality holds

$$\mathbb{P}[\exists e' = (v, v^-) \in \mathfrak{F}^f \cap \widehat{K}_i(x) \text{ s.t. } d_{\mathfrak{F}^f}(v, x) \in [1, M] \mid e = (x, y) \in \mathfrak{F}^f] > \frac{1}{b}. \quad (5.7)$$

Also observe that conditioned on $e = (x, y) \in \mathfrak{F}^f$ and $e' = (v, v^-) \in \mathfrak{F}^f \cap \widehat{K}_i(x)$, the conditional distributions of $\mathfrak{F}^f \cap \widehat{K}_i(v)$ are independent and are the same as the distribution $\mathfrak{F}^f \cap \widehat{K}_i(v)$ conditioned only on $e' = (v, v^-) \in \mathfrak{F}^f$.

Conditioned on the tree edge $e = (x, y) \in \mathfrak{F}^f$, we call a component $\widehat{K}_i(x)$ **good** if there exists some edge $e' = (v, v^-) \in \mathfrak{F}^f \cap \widehat{K}_i(x)$ s.t. $d_{\mathfrak{F}^f}(v, x) \in [1, M]$ and we also call this edge $e' = (v, v^-)$ a **good** edge for $e = (x, y)$. In a good component we pick an arbitrary good edge. Note that the size of good edges form a supercritical Galton–Watson tree in light of (5.7). With positive probability the supercritical Galton–Watson tree has branching number bigger than one [18, Cor. 5.10]. Since in \mathfrak{F}^f two good edges in neighboring generations of the Galton–Watson tree have distance at most M , with positive probability \mathfrak{F}^f also has branching number strictly larger than one. Since the branching number of \mathfrak{F}^f is a constant almost surely (Theorem 10.18 of [18]), it is strictly larger than one almost surely. \square

Remark 5.10. Proposition 5.9 is non-trivial in the sense that there exist spanning trees of G with branching number equals to one. It is also of interest to find the exact value of the branching number of the FUSF.

One natural further question is the following:

Question 5.11. Are there invariant spanning trees of the grandparent graphs with branching number equals to 1 or even being recurrent? Section 6 of [25] provides an example with branching number equals to 1 in a similar setting.

Remark 5.12. The conclusion of Proposition 5.8 and 5.9 also hold for the Cartesian product $\mathbb{T}_{b+1} \square \mathbb{Z}_2$, where the \mathbb{Z}_2 -edges in $\mathbb{T}_{b+1} \square \mathbb{Z}_2$ will play the role of the tree edges in the above proofs. We conjecture that for any finite connected graph H , almost surely the FUSF on the Cartesian product $\mathbb{T}_{b+1} \square H$ is connected. Note that every tree in the FUSF on the Cartesian product $\mathbb{T}_{b+1} \square H$ has branching number larger than one since $\mathbb{T}_{b+1} \square H$ is unimodular [12, 26].

Quite recently Pete and Timár [22] disproved the conjecture in Remark 5.12.

6 FUSF on free products of nonunimodular transitive graphs with \mathbb{Z}_2

The free product of two Cayley graphs is well known. More generally one can define the free product of two transitive graphs G_1, G_2 . For more details, see the description on page 2349 of [25].

Suppose $G_0 = (V_0, E_0)$ is a nonunimodular transitive graph and \mathbb{Z}_2 is the graph of two vertices connected by one edge. We now give the detailed definition of the free product $G_0 * \mathbb{Z}_2$ just like [25]. First take a copy of G_0 and countably many copies of \mathbb{Z}_2 . Fix a bijection from the vertices of this copy of G_0 to the copies of \mathbb{Z}_2 . Identify each vertex of this copy of G_0 with an arbitrary vertex in its image under the bijection. Call the resulting graph H_1 , it is formed by attaching an edge to each vertex of the copy of G_0 . Let I_1 denote the set of vertices on the \mathbb{Z}_2 edges that are not identified with a vertex of G_0 . Fix a bijection between I_1 and countably many new copies of G_0 . Identify every vertex of I_1 with an arbitrary vertex in its image by the bijection to obtain a graph H_2 . So H_2 is formed by attaching a G_0 copy to each vertex in I_1 . Continue this process similarly, given H_i , and if I_i is the set of vertices in H_i that were not born by identification in some previous steps, then fix a bijection between I_i to a set of infinitely

many copies of G_0 if i is odd (or infinitely many copies of \mathbb{Z}_2 if i is even). Identify every vertex in I_i with an arbitrary vertex in its image by the bijection to obtain H_{i+1} . If we view H_i as a subgraph of H_{i+1} , then finally the free product $G_0 * \mathbb{Z}_2 := \bigcup H_i$. It is easy to see G is still a transitive graph. Also we call an edge in the free product $G_0 * \mathbb{Z}_2$ a G_0 -edge if its two endpoints lies in the same copy of G_0 in the above construction. Similarly we call the other edges \mathbb{Z}_2 -edges. The free product $G = G_0 * \mathbb{Z}_2$ can be viewed as countably many disjoint copies of G_0 , say $G_0^{(i)}$ ($i = 1, 2, \dots$), connected by countably many \mathbb{Z}_2 edges in a certain way. So we can fix a labeling of the vertices of G : $V(G) = \{(v, i) : v \in V_0, i \in \mathbb{Z}^+\}$, where (v, i) indicates the vertex v in the copy $G_0^{(i)}$. Let $\Phi_i : G_0^{(i)} \rightarrow G_0$ denote the projection map, namely for each vertex $(u, i) \in V(G_0^{(i)})$ and edge $e = ((v, i), (w, i)) \in E(G_0^{(i)})$, $\Phi_i((u, i)) = u \in V_0$ and $\Phi_i(e) = (v, w) \in E_0$.

Definition 6.1. Suppose G_0 is a transitive graph with a closed subgroup Γ of automorphisms that acts transitively on G_0 . Let $G = G_0 * \mathbb{Z}_2$. Suppose ω_0 is a Γ -invariant percolation process on G_0 . We view ω_0 as a random subgraph of G_0 . For each G_0 copy $G_0^{(i)}$ in G , we take an independent percolation ω_i on $G_0^{(i)}$ such that $\Phi_i(\omega_i)$ has the same law as ω_0 . Let ω be the union of these independent percolation subgraphs and all the \mathbb{Z}_2 -edges. We call ω the **free product percolation** of ω_0 on $G = G_0 * \mathbb{Z}_2$.

Suppose (G_0, Γ) is a nonunimodular transitive pair. For the free product $G = G_0 * \mathbb{Z}_2$ defined above, let $\tilde{\Gamma}$ be the set of automorphisms γ_G of G with the following two properties:

- γ_G maps \mathbb{Z}_2 edges to \mathbb{Z}_2 edges and G_0 copies to G_0 copies;
- for each $i \in \mathbb{Z}^+$, the bijection $\varphi_i = \varphi_i(\gamma_G)$ is an element of Γ , where φ_i is given as follows. The automorphism γ_G maps the copy $G_0^{(i)}$ to another G_0 copy, say $G_0^{(j)}$. Then restricted to $G_0^{(i)}$ and its image $G_0^{(j)}$, γ_G is a bijection from $G_0^{(i)}$ to $G_0^{(j)}$. Projecting $\gamma_G|_{G_0^{(i)}}$ to the first coordinate one gets the bijection $\varphi_i : G_0 \rightarrow G_0$, namely $\varphi_i = \Phi_j \circ \gamma_G|_{G_0^{(i)}} \circ \Phi_i^{-1}$.

Observation 6.2. Suppose (G_0, Γ) is a nonunimodular transitive pair. Let $\tilde{\Gamma}$ be the set of automorphisms of $G = G_0 * \mathbb{Z}_2$ as defined above. Then $\tilde{\Gamma}$ has the following properties.

1. The set $\tilde{\Gamma}$ is a closed subgroup of $\text{Aut}(G)$ and $\tilde{\Gamma}$ acts on G transitively.
2. For each copy $G_0^{(i)}$ and two vertices $(u, i), (v, i)$ in that copy, one has $|\tilde{\Gamma}_{(u,i)}(v, i)| = |\Gamma_{uv}|$. In particular by Lemma 2.3 this implies that $\tilde{\Gamma}$ is also nonunimodular.
3. If (u, i) and (v, j) are connected in G by a \mathbb{Z}_2 -edge, then $|\tilde{\Gamma}_{(u,i)}(v, j)| = 1$.

Proposition 6.3. Suppose (G_0, Γ) is a nonunimodular transitive pair. Suppose ω_0 is a Γ -invariant percolation process on G_0 and ω is the free product percolation of ω_0 on $G = G_0 * \mathbb{Z}_2$. If almost surely every connected component of ω_0 is infinite, then each connected component of ω is $\tilde{\Gamma}$ -heavy and has branching number bigger than one.

Given Proposition 6.3 it is easy to prove Theorem 1.3.

Proof of Theorem 1.3. Notice that the FUSF on $G = G_0 * \mathbb{Z}_2$ is an example of free product percolation of $\mathfrak{F}^f(G_0)$ on G , where $\mathfrak{F}^f(G_0)$ is a FUSF sample on G_0 . Theorem 1.3 follows from the combination of Proposition 5.8, 5.9 and Proposition 6.3. \square

Proof of Proposition 6.3. The part that each connected component of ω has branching number bigger than one is obvious and we omit the details.

The part that each connected component of ω is heavy can be proved using comparison to branching random walks.

For a fix vertex $u \in V(G_0)$, since each connected component of ω_0 is infinite, writing $K_{G_0}(u)$ of the connected component of u in ω_0 , one has

$$\mathbb{E} \left[\sum_{v \in V(G_0)} \mathbf{1}_{\{v \in K_{G_0}(u)\}} \right] = \infty.$$

Using TMTF and noting $v \in K_{G_0}(u)$ if and only if $u \in K_{G_0}(v)$, one has

$$\mathbb{E} \left[\sum_{v \in V(G_0)} \mathbf{1}_{\{v \in K_{G_0}(u)\}} \frac{m(v)}{m(u)} \right] = \infty. \tag{6.1}$$

By monotone convergence theorem there exists a large constant M such that

$$\mathbb{E} \left[\sum_{v \in V(G_0)} \mathbf{1}_{\{v \in K_{B_{G_0}(u,M)}(u)\}} \frac{m(v)}{m(u)} \right] > e, \tag{6.2}$$

where $B_{G_0}(u, M)$ denotes the ball in G_0 with center u and radius M and $K_{B_{G_0}(u,M)}(u)$ denote the connected component of u when one consider the percolation of ω_0 restricted in $B_{G_0}(u, M)$.

For $x = (u, i) \in V(G)$, there is a unique \mathbb{Z}_2 -edge incident to x and we suppose the other vertex of the \mathbb{Z}_2 -edge is $x' = (u', j)$. Let $K(x)$ denote the connected component of x in the free product percolation ω . Let $K_h(x)$ denote the connected component of x if we delete the \mathbb{Z}_2 -edge $e = (x, x')$ from $K(x)$.

For a fixed constant $M > 0$, we truncate $K_h(x)$ as follows. First we truncate all the edges $e = ((a, i), (b, i))$ in the same copy $G_0^{(i)}$ as $x = (u, i)$ if $\max\{d_{G_0}(a, u), d_{G_0}(b, u)\} \geq M$. For all the vertex y in the same copy $G_0^{(i)}$ as x that can be connected to x by an ω -open path staying in $B_{G_0^{(i)}}(x, M)$ (the ball in the copy $G_0^{(i)}$ with center x and radius M), we keep the \mathbb{Z}_2 -edge (y, y') and do the same truncation procedure for y' as previous for x . Keep doing this procedure and in the end we get an infinite random graph $K_h^M(x)$. This random graph $K_h^M(x)$ is just the sub-graph of $K_h(x)$ induced by those vertices that have an ω -open path to x such that between any two consecutive \mathbb{Z}_2 -edges on this path there is at most M other (G_0) -edges.

Now we show that $K_h^M(x)$ is heavy with positive probability.

We first briefly recall the definition of a branching random walk and a related result; see [16] for more details. Let $\mathcal{L} := \{X_i\}_{i=1}^N$ be a random N -tuple of real numbers, where N is also random. We can view the branching random walk as an ordered point process on the real line. An initial point x is located at the origin. It gives birth to N children x_1, \dots, x_N with random displacements X_1, \dots, X_N . Then each x_i gives birth to a random number of particles with random displacement relative to the position of x_i according to the same law as \mathcal{L} and independently of one another and of the initial displacements. This procedure continues forever or until no more particles are born.

For a particle u , let $|u|$ be the generation of u and $X(u)$ its displacement from its parent, and $S(u)$ its position (relative to the origin). Denote the initial particle at origin by 0. If u is an ancestor of v , write $u < v$. Then $S(v) = \sum_{0 < u \leq v} X(u)$.

For $\alpha \in \mathbb{R}$, write $\langle \alpha, \mathcal{L} \rangle := \sum_{i=1}^N e^{-\alpha X_i}$ and $\lambda(\alpha) := \mathbb{E}[\langle \alpha, \mathcal{L} \rangle] \in (0, \infty]$. Assume $\lambda(0) > 1$ so that the extinction probability $q < 1$. If $\lambda(\alpha) < \infty$ for some α , then $W_n(\alpha) := \frac{\sum_{|u|=n} e^{-\alpha S(u)}}{\lambda(\alpha)^n}$ is a martingale with a.s. limit $W(\alpha)$.

Set

$$\lambda'(\alpha) := \mathbb{E} \left[\sum_{i=1}^N X_i e^{-\alpha X_i} \right]$$

when the integral exists as a Lebesgue integral. Biggins' theorem tells us that if there exists a real number α such that the following three conditions holds, then the limit $W(\alpha)$ satisfies $\mathbb{P}[W(\alpha) = 0] = q < 1$.

- (1) $\lambda(\alpha) < \infty$ and $\lambda'(\alpha)$ exists and is finite;
- (2) $\mathbb{E}[\langle \alpha, \mathcal{L} \rangle \log^+ \langle \alpha, \mathcal{L} \rangle] < \infty$ and
- (3) $\alpha \lambda'(\alpha) / \lambda(\alpha) < \log \lambda(\alpha)$.

Now given $K_h^M(x)$, we construct a corresponding branching random walk in the following manner. Let $N + 1$ be the number of vertices in the connected component of $x = (u, i)$ in $K_h^M(x)$ intersecting with the copy $G_0^{(i)}$, which we denote by $K_h^M(x, B_{G_0^{(i)}})$. Note that $N \in [M, |B_{G_0}(x, M)|]$ is a finite random number. Let x_1, \dots, x_N be an arbitrary ordering of the vertices in $K_h^M(x, B_{G_0^{(i)}}) \setminus \{x\}$. Now for the corresponding branching random walk, we let the initial particle give birth to N children, each with displacement $X_i = \log \frac{m(x_i)}{m(x)}$. For each x_i , let $x'_i = (u_i, j_i)$ be the vertex incident to x_i by a \mathbb{Z}_2 -edge. Let N_{x_i} denote the number of vertices in $K_h^M(x'_i, B_{G_0^{(j_i)}}) \setminus \{x'_i\}$. In the corresponding branching random walk, we let the particle corresponding to x_i give birth to N_{x_i} new particles each with a relative displacement $\log \frac{m(y)}{m(x'_i)}$, $\forall y \in K_h^M(x'_i, B_{G_0^{(j_i)}}) \setminus \{x'_i\}$. By the third item in Observation 6.2 and Lemma 2.3 one has $m(x_i) = m(x'_i)$, thus $\log \frac{m(y)}{m(x'_i)} = \log \frac{m(y)}{m(x_i)}$.

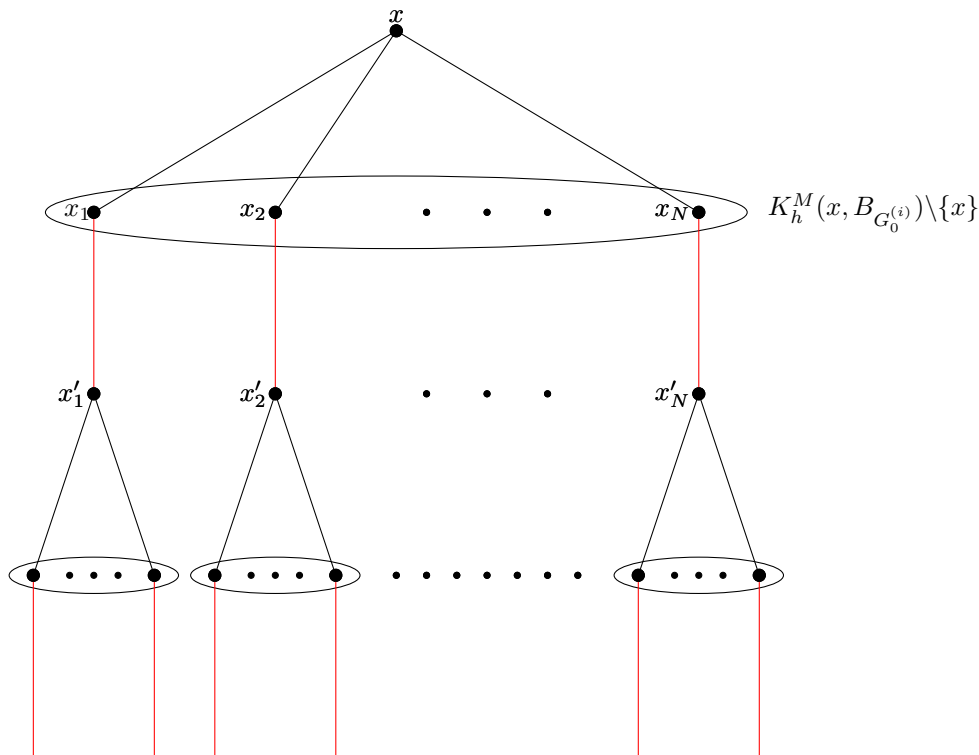


Figure 5: A systematic drawing of the family tree of the corresponding branching random walk for $K_h^M(x)$ if one contracts all the red edges

We use y to denote a vertex in $K_h^M(x)$ or its corresponding vertex in the family tree of the corresponding branching random walk. Now we take $\alpha = -1$, then $e^{-\alpha X_i} = \frac{m(x_i)}{m(x)}$,

$e^{-\alpha S(y)} = \frac{m(y)}{m(x)}$. Notice that as a subgraph of $G_0^{(i)}$, $K_h^M(x, B_{G_0^{(i)}})$ has the same law as $K_{B_{G_0}(u, M)}(u)$ as a subgraph of G_0 . By Lemma 2.3 and the second item in Observation 6.2

$$\begin{aligned} \lambda(-1) &= \mathbb{E} \left[\sum_{y \in K_h^M(x, B_{G_0^{(i)}}) \setminus \{x\}} \frac{m(y)}{m(x)} \right] \\ &= \mathbb{E} \left[\sum_{y \in V(G_0), y \neq x} \mathbf{1}_{\{y \in K_{B_{G_0}(x, M)}(x)\}} \frac{m(y)}{m(x)} \right] \stackrel{(6.2)}{>} e. \end{aligned} \tag{6.3}$$

Now we verify that the conditions listed above to use Biggins' theorem hold. Conditions (1) and (2) are trivial since in our case X_i and N are both bounded. Since $\alpha = -1$ and $\lambda(-1) > 0$, condition (3) is just $\lambda'(-1) + \lambda(-1) \log \lambda(-1) > 0$. Since $\lambda(-1) > e$, it suffices to show that

$$\lambda'(-1) + \lambda(-1) \geq 0. \tag{6.4}$$

By definition and the equivalence of $y \in K_{B_{G_0}(x, M)}(x)$ and $x \in K_{B_{G_0}(y, M)}(y)$ one has

$$\begin{aligned} \lambda'(-1) &= \mathbb{E} \left[\sum_{y \in V(G_0), y \neq x} \mathbf{1}_{\{y \in K_{B_{G_0}(x, M)}(x)\}} \frac{m(y)}{m(x)} \log \frac{m(y)}{m(x)} \right] \\ &\stackrel{\text{TMTP}}{=} \mathbb{E} \left[\sum_{y \in V(G_0), y \neq x} \mathbf{1}_{\{y \in K_{B_{G_0}(x, M)}(x)\}} \log \frac{m(x)}{m(y)} \right] \end{aligned} \tag{6.5}$$

Note that $f(t) = t + \log \frac{1}{t} \geq f(1) = 1$ on $(0, \infty)$. Hence by (6.3) and (6.5) one has (6.4):

$$\lambda'(-1) + \lambda(-1) = \mathbb{E} \left[\sum_{y \in V(G_0), y \neq x} \mathbf{1}_{\{y \in K_{B_{G_0}(x, M)}(x)\}} f\left(\frac{m(y)}{m(x)}\right) \right] \geq \mathbb{E}[N] \geq M > 0.$$

Therefore Biggins' theorem yields that with positive probability $W(\alpha) > 0$. On the event that $W(\alpha) > 0$, the corresponding $m(K_h^M(x)) := \sum_{y \in K_h^M(x)} m(y) = \infty$ since $\sum_{|u|=n} e^{-\alpha S(u)} = \sum_{|u|=n} \frac{m(u)}{m(x)}$ tends to infinity because $\lambda(-1) > 1$ and $W(\alpha) > 0$. Using the standard trick (see e.g. Proposition 5.6 in [18]) one has $m(K_h^M(x)) = \infty$ almost surely.

Therefore $K(x)$ is heavy almost surely. □

Remark 6.4. The free minimal spanning forests on $G = G_0 * \mathbb{Z}_2$ is also a free product percolation that satisfies the condition of Proposition 6.3. Hence each tree of the free minimal spanning forests on $G = G_0 * \mathbb{Z}_2$ is also heavy and has branching number bigger than one. Interested readers can refer to Chapter 11 of [18] for more background on the free minimal spanning forests.

We conclude with two further open questions on FUSF on nonunimodular transitive graphs. The first question is about the number of trees in the FUSF. Benjamini et al asked whether the number of trees of the FUSF is 1 or ∞ almost surely (Question 15.6 in [3]). Hutchcroft and Nachmias [12] answered this question positively for unimodular transitive graphs and the nonunimodular case remains open. Another question one can consider is about the indistinguishability of the trees in the FUSF on a nonunimodular transitive graph. Since the trees in the WUSF on a nonunimodular transitive graph is light almost surely, they are distinguishable by automorphism-invariant properties, e.g. the sum of degrees of the highest points in the components. So the interesting case is for the trees in the FUSF on nonunimodular transitive graphs with the property that FUSF \neq WUSF. The techniques from [12, 26, 24] might be useful for this question.

A Proof of (4.39) in Proposition 4.7

We have already seen the following property of simple random walk from Lemma 4.1 and (4.31):

$$\mathbb{P}[\{X_m^x : m \geq 0\} \cap L_n(x) \neq \emptyset] \begin{cases} \asymp e^{-t_0 n} & \text{if } n > 0 \\ = 1 & \text{if } n \leq 0 \end{cases} \tag{A.1}$$

Using (A.1) and the strong Markov property for simple random walk, to show (4.39) it suffices to show the case $n = 0$, i.e., there exists $c > 0$ such that

$$\mathbb{P}[\{X_m^x : m \geq 0\} \cap L_0(x) \geq k] \geq e^{-ck}, \forall k \geq 1 \tag{A.2}$$

The basic idea is to show that there is a large constant $c > 0$ such that for every $k \geq 1$, one can construct a simple random walk trajectory $\{X_m^x : m \geq 0\}$ starting from x such that $|\{X_m^x : m = 0, \dots, ck\} \cap L_0(x)| \geq k$.

Recall $t_0 = \max\{\log \Delta(x, y) : x \sim y\} > 0$. For $k \in \mathbb{Z}$, let $N_k(x) := \{y : d_G(x, y) = |k|, \text{ and } \log \Delta(x, y) = kt_0\}$. In particular, $N_0(x) = \{x\}$. For a set $A \subset V(G)$, write $N_k(A) := \bigcup_{x \in A} N_k(x)$.

Obviously for $k \geq 1$, $N_k(x) \subset N_1(N_{k-1}(x))$. On the other hand, if $y \in N_1(N_{k-1}(x))$, then there exists $z \in N_{k-1}(x)$ such that $y \in N_1(z)$. Hence $d_G(x, z) = k - 1$, $d_G(z, y) = 1$ and $\Delta(x, z) = e^{(k-1)t_0}$, $\Delta(z, y) = e^{t_0}$. Thus $d_G(x, y) \leq d_G(x, z) + d_G(z, y) = k$ and $\Delta(x, y) = \Delta(x, z) \cdot \Delta(z, y) = e^{kt_0}$. By the choice of t_0 , one must have $d_G(x, y) = k$ and thus $y \in N_k(x)$. Therefore one has the reverse direction $N_k(x) \supset N_1(N_{k-1}(x))$. Hence for all $k \geq 1$, $N_k(x) = N_1(N_{k-1}(x))$.

Using similar idea, one has $N_{k+m}(x) = N_k(N_m(x))$ and $N_{-k-m}(x) = N_{-k}(N_{-m}(x))$ for all $k, m \geq 0$. In particular, $|N_{k+m}(x)| \leq |N_k(x)| \cdot |N_m(x)|$ and $N_{-k-m}(x) = |N_{-k}(x)| \cdot |N_{-m}(x)|$ for all $k, m \geq 0$. Also the fact $N_{k+m}(x) = N_k(N_m(x))$, $k, m \geq 0$ implies that

$$|N_k(x)| \geq |N_j(x)|, \forall k \geq j \geq 0. \tag{A.3}$$

Consider a mass-transport function as $F_k(x, y) = \frac{\mathbf{1}_{\{y \in N_k(x)\}}}{|N_k(x)|}$. Notice that $x \in N_k(y)$ if and only if $y \in N_{-k}(x)$. By the tilted mass-transport principle we have

$$1 = \sum_{y \in V} F_k(x, y) = \sum_{y \in V} F_k(y, x) \Delta(x, y) = \sum_{y \in N_{-k}(x)} \frac{1}{|N_k(x)|} e^{-kt_0}.$$

Thus

$$|N_{-k}(x)| = e^{kt_0} |N_k(x)|, \forall k \geq 1. \tag{A.4}$$

By (A.3) and (A.4) one has that there exists a positive constant c_1 such that for all $k \geq 1$,

$$\sum_{j=0}^k |N_{-j}(x)| = |N_{-k}(x)| \sum_{j=0}^k e^{(j-k)t_0} \frac{|N_j(x)|}{|N_k(x)|} \leq |N_{-k}(x)| \sum_{j=0}^k e^{(j-k)t_0} \leq c_1 |N_{-k}(x)|. \tag{A.5}$$

Consider the following subgraph $T(x)$ of G . Let $V(T(x))$ be the union $\bigcup_{k \geq 0} N_{-k}(x)$. For each $k \geq 0$ and $y \in N_{-k-1}(x)$, there is at least one edge linking from $N_{-k}(x)$ to y . If there are more one such edges, just pick one arbitrarily. Let $E(T(x))$ be the collection of edges picked. Then obviously $T(x)$ is a tree. Write $T_{\leq n}(x) := \{y \in V(T) : d_{T(x)}(y, x) \leq n\}$.

By (A.4), $|N_{-j}(x)|$ grows exponentially fast in j . Now for a fixed integer $k \geq 1$, let j be the smallest nonnegative integer such that $|N_{-j}(x)| > k$. Then

$$|N_{-j+1}(x)| \leq k < |N_{-j}(x)| \leq |N_{-1}(x)| \cdot |N_{-j+1}(x)| \leq |N_{-1}(x)| \cdot k.$$

Pick $v \in N_j(x)$. Let the simple random walk $\{X_m^x : m \geq 0\}$ travels according to the depth-first search of the tree $T_{\leq j}(v)$ until it visits all vertices in $N_{-j}(v)$. Under

this event one has that $|\{X_m^x : m \geq 0\} \cap L_0(x)| \geq |N_{-j}(v)| > k$. Notice that there are $\sum_{i=1}^j |N_{-i}(v)| \stackrel{(A.5)}{\leq} c_1 |N_{-j}(v)| \leq c_1 |N_{-1}(x)| k$ edges in $T_{\leq j}(v)$. Also notice each edge in the tree $T_{\leq j}(v)$ has been crossed by the depth-first search at most twice. Denoting by D the degree of G , one obtains (A.2):

$$\mathbb{P}[|\{X_m^x : m \geq 0\} \cap L_0(x)| > k] \geq \frac{1}{D^{2c_1 |N_{-1}(x)| k}} \geq \exp(-c_2 k).$$

B Proof of Proposition 4.15 for general k

The proof of Proposition 4.15 uses a **tree-graph inequality**. For similar ideas and notations, see the proof of [7, Theorem 6.75], where high moments of a cluster were estimated in the Bernoulli percolation case.

Here we recall some definitions from [7, Page 135].

Definition B.1. A tree is called a **skeleton** if each vertex has degree 1 or 3. Suppose S is a skeleton, let $I(S)$ denote the set of vertices with degree 3 and call these vertices **interior vertices**. Call the other vertices of S **exterior vertices** (they are just leaves of the tree S). It is easy to see that if a skeleton has k exterior vertices, then it must have $k - 2$ interior vertices. Write $V(S), E(S)$ as the vertex set and edge set of S respectively as usual (in particular $E(S)$ is not the set of exterior vertices).

A skeleton with k exterior vertices is called **labelled** if there is an assignment of the numbers $0, 1, \dots, k - 1$ to the exterior vertices. Two labelled skeletons S_1 and S_2 are called **isomorphic** if there is a graph isomorphism from S_1 to S_2 which also preserves the labels of the exterior vertices.

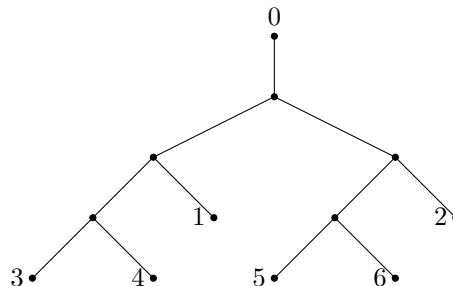


Figure 6: A labelled skeleton with $6 + 1$ exterior vertices and $6 - 1$ interior vertices.

The following lemma is an adaption of the claim from [7, Page 135-136, Section 6.3].

Lemma B.2. Suppose x_0, x_1, \dots, x_k are (possibly non-distinct) vertices of G such that $x_i \in \mathfrak{T}_{x_0}$ for $i = 1, \dots, k$. Write $\mathbf{x} = (x_0, x_1, \dots, x_k)$. Then there is a labelled skeleton S with $k + 1$ exterior vertices together with a map $\varphi_{\mathbf{x}} : V(S) \rightarrow V(G)$ such that

- the exterior vertex with label i maps to x_i by $\varphi_{\mathbf{x}}$ for $i = 0, 1, \dots, k$ and
- the edges of S correspond to edge-disjoint paths in \mathfrak{T}_{x_0} .

Here if (u, v) is an edge of S such that $\varphi_{\mathbf{x}}(u) = \varphi_{\mathbf{x}}(v)$, then its corresponding path is an empty path and we use the convention that the empty path is edge disjoint with any other path. Such a map is called an **admissible map**.

Lemma B.3. Suppose x_0, x_1, \dots, x_k are (possibly non-distinct) vertices of G . Then

$$\tau(x_0, x_1, \dots, x_k) \leq \sum_S \sum_{\varphi_{\mathbf{x}}} \prod_{(u,v) \in E(S)} \tau(\varphi_{\mathbf{x}}(u), \varphi_{\mathbf{x}}(v)) \tag{B.1}$$

where the first sum is over all labelled skeleton with $k + 1$ exterior vertices and the second sum is over all admissible maps φ_x from $V(S)$ to $V(G)$.

Proof. This is the generalization of Lemma 4.13. By Lemma B.2, one has

$$\tau(x_0, x_1, \dots, x_k) \leq \sum_S \sum_{\varphi_x} \mathbb{P}[\varphi_x(u) \in \mathfrak{T}_{x_0} \text{ for all } u \in V(S) \text{ and the paths in } \mathfrak{T}_{x_0} \text{ connecting } \varphi_x(u), \varphi_x(v) \text{ are edge-disjoint for all } (u, v) \in E(S)].$$

Place the skeleton S in the lower half plane in some arbitrary way. Order the vertices of S first according to the graph distance to the exterior vertex with label 0 (for simplicity we will denote this vertex 0 hereafter) and if two vertices are at the same distance to 0, then order them according left to right. Call this order Or . Sample \mathfrak{T}_{x_0} using independent simple random walks started from the order $\varphi_x(Or)$. Then one can see

$$\begin{aligned} \mathbb{P}[\varphi_x(u) \in \mathfrak{T}_{x_0} \text{ for all } u \in V(S) \text{ and the paths in } \mathfrak{T}_{x_0} \text{ connecting } \\ \varphi_x(u), \varphi_x(v) \text{ are edge-disjoint for all } (u, v) \in E(S)] \\ \leq \prod_{(u,v) \in E(S)} \tau(\varphi_x(u), \varphi_x(v)). \end{aligned}$$

Hence that (B.1) holds. □

Definition B.4. Suppose x_0, x_1, \dots, x_k are (possibly non-distinct) vertices of G . Suppose that a labelled skeleton S and an admissible map φ_x from $V(S)$ to $V(G)$ is given. Name the vertices of S the same as the image under φ_x . In particular, write $I(S) = \{u_1, \dots, u_{k-1}\}$. Let l_v be the integer such that $\varphi_x(v) \in L_{l_v}(x_0)$. Write $\vec{l} = (l_v : v \in V(S))$. Oriented the edges of S such that they all lead away from the root x_0 and write the set of oriented edges as $E(\vec{S})$. Define $f(S, \vec{l})$ to be

$$f(S, \vec{l}) := \sum_{\langle u,v \rangle \in E(\vec{S})} (l_v - l_u) \vee 0.$$

Lemma B.5. Use the same notation as in Definition B.4. Fix an arbitrary integer n . Let $m(S, n, j)$ be the number of $k - 1$ tuples (l_1, \dots, l_{k-1}) such that the corresponding \vec{l} with $l_{u_i} = l_i, i = 1, \dots, k - 1, l_{x_0} = 0$ and $l_{x_i} = n, i = 1, \dots, k$ satisfies

$$f(S, \vec{l}) = j,$$

where $j \geq n \vee 0$. For $n \geq 0$, one has $m(S, n, n) = 1$. Also there exists a constant $c > 0$ such that

$$m(S, n, j) \leq c^k (j - n)^{k-1} \text{ for all } n \geq 0, j > n \tag{B.2}$$

and

$$m(S, n, j) \leq c^k |n|^{k-1} (j \vee 1)^{k-1} \text{ for all } n < 0, j \geq 0 \tag{B.3}$$

Proof. For each oriented path $\gamma = \langle v_1, \dots, v_t \rangle$ in S , one has that

$$\sum_{\langle u,v \rangle \in E(\gamma)} (l_v - l_u) \vee 0 \geq \sum_{\langle u,v \rangle \in E(\gamma)} (l_v - l_u) = l_{v_t} - l_{v_1}.$$

First consider the case $n \geq 0, j \geq n$. For each inner vertex u_i , considering the oriented path η_{x_0, u_i} from x_0 to u_i in S , one has that

$$l_i = l_{u_i} - l_{x_0} \leq \sum_{\langle u,v \rangle \in E(\eta(x_0, u_i))} (l_v - l_u) \vee 0 \leq f(S, \vec{l}) = j.$$

On the other hand, there are at least two disjoint oriented paths starting from u_i to the leaves in S , say η_1, η_2 . These two paths are also disjoint with the oriented path from x_0 to u_i . Hence

$$2(n - l_i) + l_i \leq \sum_{\langle u,v \rangle \in E(\eta_1) \cup E(\eta_2)} (l_v - l_u) \vee 0 + \sum_{\langle u,v \rangle \in E(\eta(x_0, u_i))} (l_v - l_u) \vee 0 \leq f(S, \vec{l}) = j$$

Hence the above two inequalities imply that $l_i \in [2n - j, j], i = 1, \dots, k - 1$. In particular, for $j = n \geq 0$, the only tuple (l_1, \dots, l_{k-1}) such that $f(S, \vec{l}) = j$ is (n, \dots, n) and thus $m(S, n, n) = 1$. For $j > n \geq 0$, the number of tuples such that $f(S, \vec{l}) = j$ is at most $(2j - 2n + 1)^{k-1} \leq c^k (j - n)^{k-1}$.

Next for the case $n < 0, j \geq 0$, one just need to observe that there exists a large constant $c > 0$ such that $2j - 2n + 1 \leq c|n|(j \vee 1)$. □

Proof of Proposition 4.15. Since for $k \geq 2, |\mathfrak{F}_x \cap L_n(x)|^k = \sum_{x_1, \dots, x_k \in L_n(x)} \tau(x, x_1, \dots, x_k)$, using Lemma B.3 one has

$$\begin{aligned} \mathbb{E}[|\mathfrak{F}_x \cap L_n(x)|^k] &\leq \mathbb{E}\left[\sum_{x_1, \dots, x_k \in L_n(x)} \sum_S \sum_{\varphi_x} \prod_{(u,v) \in E(S)} \tau(\varphi_x(u), \varphi_x(v))\right] \\ &= \mathbb{E}\left[\sum_S \sum_{l_1, \dots, l_{k-1}} \sum_{x_1, \dots, x_k \in L_n(x)} \sum_{\substack{\varphi_x: \\ \varphi_x(u_i) \in L_{l_i}(x)}} \prod_{(u,v) \in E(S)} \tau(\varphi_x(u), \varphi_x(v))\right] \\ &\preceq c^k \sum_S \sum_{l_1, \dots, l_{k-1}} \exp(-t_0 f(S, \vec{l})), \end{aligned} \tag{B.4}$$

where in the last step is obtained as follows: for a fixed skeleton S and a sequence (l_1, \dots, l_{k-1}) of levels of the interior vertices,

- Pick a leaf of the tree S distinct from the root, say x_1 . Suppose x_1 is neighboring an interior vertex v , say $l_v = l_i$. We first sum over $x_1 \in L_n(x)$. Since x_1 is a leaf, it only appears once in the product, the term corresponding to the edge (v, x_1) . By Observation 4.14 this is the same as summing over vertices in $L_{n-l_i}(v)$ and the term can be bounded by

$$c_1 \exp(-t_0(n - l_i) \vee 0).$$

- We delete the leaf we picked in the previous step from the tree S and get a new tree. Then we pick another leaf (distinct from the root again) in the new tree and then do the same procedure as in the previous step.
- In the end we stop after $2k - 1$ steps (there are $2k - 1$ edges in $E(S)$) and by the definition of $f(S, \vec{l})$ the sum can be bounded

$$c_1^{2k-1} \exp(-t_0 f(S, \vec{l})).$$

From Formula (6.96) ([7, Page 138, Section 6.3]) we know that the number of labelled skeletons with $k + 1$ exterior vertices is $\frac{(2k-2)!}{2^{k-1}(k-1)!} \preceq 2^k k!$. Hence by (B.4) one has that

$$\mathbb{E}[|\mathfrak{F}_x \cap L_n(x)|^k] \leq (2c)^k k! \sum_{j \geq n \vee 0} m(S, n, j) \exp(-t_0 j). \tag{B.5}$$

For $n \geq 0$, by Lemma B.5 there exists a constant $c_1 > 0$ such that

$$\begin{aligned} \mathbb{E}[|\mathfrak{T}_x \cap L_n(x)|^k] &\leq c_1^k k! \sum_{j=n}^{\infty} ((j-n) \vee 1)^{k-1} \exp(-t_0 j) \\ &\leq c_1^k k! e^{-t_0 n} \int_0^{\infty} x^{k-1} e^{-t_0 x} dx \\ &= \left(\frac{c_1}{t_0}\right)^k (k!)^2 e^{-t_0 n}. \end{aligned}$$

For $n < 0$, by Lemma B.5 there exists a constant $c_2 > 0$ such that

$$\begin{aligned} \mathbb{E}[|\mathfrak{T}_x \cap L_n(x)|^k] &\leq c_2^k k! |n|^{k-1} \sum_{j=0}^{\infty} (j \vee 1)^{k-1} \exp(-t_0 j) \\ &\leq c_2^k k! |n|^{k-1} \int_0^{\infty} x^{k-1} e^{-t_0 x} dx \\ &= \left(\frac{c_2}{t_0}\right)^k (k!)^2 |n|^{k-1}. \quad \square \end{aligned}$$

C More quantitative results for the toy model

C.1 The intersection $\mathfrak{T}_x \cap L_n(x)$ when $n \geq 0$

By Lemma 3.2 we know the tree \mathfrak{T}_x in the x -WUSF has same law as C_x , where C_x denotes the cluster of x in an independent critical Bernoulli bond percolation on the regular tree \mathbb{T}_{b+1} . Many things are known about the cluster of C_x at criticality; for example see [10, Theorem 1.6].

Here we are particularly interested in the tail probability of $|C_x \cap L_n(x)|$ when $n \geq 0$.

Proposition C.1. Consider x -WUSF on the toy model $(\mathbb{T}_{b+1}, \Gamma_\xi)$. Then for $n \geq 0, k \geq 1$,

$$\mathbb{P}[|\mathfrak{T}_x \cap L_n(x)| \geq k] = e^{-t_0 n - \Theta(\sqrt{k})} \tag{C.1}$$

Question C.2. Can one generalize Proposition C.1 to WUSF on all nonunimodular transitive graphs?

In light of Lemma 3.2, the following question is also natural to ask.

Question C.3. Consider critical Bernoulli percolation on a nonunimodular transitive graph. Do we have the following estimate:

$$\mathbb{P}[|C_x \cap L_n(x)| \geq k] = e^{-t_0 n - \Theta(\sqrt{k})}, n \geq 0, k \geq 1?$$

We start with some lemmas on critical Galton–Watson trees.

Lemma C.4. Consider a critical Galton–Watson tree with binomial progeny distribution $\text{Bin}(b, 1/b)$, where $b \geq 2$. Let Z_n be the size of the n -th generation of this critical Galton–Watson tree and $Z_0 = 1$. Then there exists a constant $C > 1$ such that for all $k \geq 1, n \geq 1$,

$$\mathbb{E}[Z_n^k] \leq C^k n^{k-1} k!$$

The intuition behind this lemma is to think about the limit of Z_n . By classical results of branching process (for example Theorem 1 and 2 in [1, page 19-20]),

$$\lim_{n \rightarrow \infty} \frac{\mathbb{P}[Z_n > 0]}{\frac{2}{(1-\frac{1}{b})^n}} = 1$$

and

$$\lim_{n \rightarrow \infty} \mathbb{P}\left[\frac{Z_n}{n} > x \mid Z_n > 0\right] = \exp\left(-\frac{2x}{(1-\frac{1}{b})}\right), x \geq 0.$$

If a nonnegative random variable Z has the following distribution

- $\mathbb{P}[Z > 0] = 1 - \mathbb{P}[Z = 0] = \frac{2}{(1-\frac{1}{b})n}$ and
- $\mathbb{P}\left[\frac{Z}{n} > x \mid Z > 0\right] = \exp\left(-\frac{2x}{(1-\frac{1}{b})}\right), x \geq 0..$

then it's easy to compute that $\mathbb{E}[Z^k] = \frac{n^{k-1} \cdot k!}{\lambda^{k-1}}, \forall k \geq 1$, where $\lambda = \frac{2}{1-\frac{1}{b}}$.

Proof. For a random variable W with binomial distribution $\text{Bin}(m, p)$, its moment generating function is

$$M_W(t) := \mathbb{E}[e^{tW}] = [1 - p + pe^t]^m.$$

Notice that conditioned on Z_{m-1} , Z_m has binomial distribution $\text{Bin}(bZ_{m-1}, \frac{1}{b})$. Hence for $m \geq 2$,

$$\begin{aligned} \mathbb{E}[e^{tZ_m}] &= \mathbb{E}[\mathbb{E}[e^{tZ_m} \mid Z_{m-1}]] \\ &= \mathbb{E}\left[\left(1 - \frac{1}{b} + \frac{1}{b}e^t\right)^{bZ_{m-1}}\right] \\ &= \mathbb{E}\left[e^{(b \log(1 - \frac{1}{b} + \frac{1}{b}e^t)) \cdot Z_{m-1}}\right]. \end{aligned} \tag{C.2}$$

Using Taylor expansion at 0 one has

$$b \log\left(1 - \frac{1}{b} + \frac{1}{b}e^t\right) = t + \left(\frac{1}{2} - \frac{1}{2b}\right)t^2 + o(t^2).$$

Hence there exists a large constant $C_0 > 0$ such that for all $t \in [0, \frac{1}{C_0}]$,

$$t \leq b \log\left(1 - \frac{1}{b} + \frac{1}{b}e^t\right) \leq t + \frac{1}{2}t^2.$$

Hence for $t \in [0, \frac{1}{C_0}]$, one has

$$\mathbb{E}[e^{tZ_m}] = \mathbb{E}\left[e^{(b \log(1 - \frac{1}{b} + \frac{1}{b}e^t)) \cdot Z_{m-1}}\right] \leq \mathbb{E}\left[e^{t(1 + \frac{t}{2})Z_{m-1}}\right]. \tag{C.3}$$

If t is such that $t \cdot e < \frac{2}{n} \wedge \frac{1}{C_0}$, then one has

$$t\left(1 + \frac{t}{2}\right) < t\left(1 + \frac{1}{n}\right).$$

Take $C \geq C_0$ such that $t = \frac{1}{nC}$ satisfying $t \cdot e < \frac{2}{n} \wedge \frac{1}{C_0}$ for all $n \geq 1$, then by (C.3) one has

$$\mathbb{E}[e^{tZ_n}] \leq \mathbb{E}\left[e^{t\left(1 + \frac{1}{n}\right)Z_{n-1}}\right].$$

By our choice of C and t , one has $t\left(1 + \frac{1}{n}\right) < t \cdot e < \frac{2}{n} \wedge \frac{1}{C_0}$. Hence by (C.3) again one has

$$\mathbb{E}[e^{tZ_n}] \leq \mathbb{E}\left[e^{t\left(1 + \frac{1}{n}\right)Z_{n-1}}\right] \leq \mathbb{E}\left[e^{t\left(1 + \frac{1}{n}\right)^2 Z_{n-2}}\right].$$

Repeating this one obtains for $t = \frac{1}{nC}$

$$\mathbb{E}[e^{tZ_n}] \leq \mathbb{E}\left[e^{t\left(1 + \frac{1}{n}\right)^{n-1} Z_1}\right] \leq \mathbb{E}\left[e^{teZ_1}\right] \leq \mathbb{E}\left[e^{\frac{2Z_1}{n}}\right].$$

Since Z_1 has binomial distribution $\text{Bin}(b, \frac{1}{b})$, one has for $t = \frac{1}{nC}$,

$$\mathbb{E}[e^{tZ_n}] \leq \left[1 - \frac{1}{b} + \frac{1}{b}e^{\frac{2}{n}}\right]^b \leq 1 + \frac{C_1}{n}. \tag{C.4}$$

Hence for $k \geq 1$,

$$\frac{t^k \mathbb{E}[Z_n^k]}{k!} \leq \mathbb{E}[e^{tZ_n}] - 1 \leq \frac{C_1}{n}.$$

Rewriting this one obtains

$$\mathbb{E}[Z_n^k] \leq C_1 \cdot C^k n^{k-1} \cdot k!. \quad \square$$

Remark C.5. Notice that $Z_n \leq b^n$ almost surely. So for fixed n , $\mathbb{E}[Z_n^k]$ grows at most exponentially in k . Hence there is no constant $C > 0$ such that for all $n \geq 1, k \geq 1$

$$\mathbb{E}[Z_n^k] \geq \frac{n^{k-1} \cdot k!}{C^k}.$$

Lemma C.6. Consider a critical Galton–Watson tree with binomial progeny distribution $\text{Bin}(b, 1/b)$. Let Z_n be the size of the n -th generation of this critical Galton–Watson tree. Then there exist constants $c_1, c_2, c_3, c_4 > 0$ such that for all $n \geq 5$,

$$c_1 \exp(-c_2 n) \leq \mathbb{P}[Z_n \geq n^2] \leq c_3 \exp(-c_4 n).$$

Here the requirement of $n \geq 5$ is just to ensure $\mathbb{P}[Z_n \geq n^2] > 0$.

Proof. By (C.4) we have seen that for a large enough constant C ,

$$\mathbb{E}[e^{\frac{Z_n}{n^C}}] \leq 1 + \frac{C_1}{n}.$$

Using Markov’s inequality one has that

$$\mathbb{P}[Z_n \geq n^2] = \mathbb{P}[e^{\frac{Z_n}{n^C}} \geq e^{\frac{n}{C}}] \leq \frac{1 + \frac{C_1}{n}}{e^{\frac{n}{C}}} \leq (1 + C_1)e^{-\frac{n}{C}}.$$

It is easy to compute that

$$\mathbb{E}[Z_n^2] = 1 + \frac{(b-1)n}{b}.$$

and

$$\text{Var}(Z_n) = \frac{(b-1)n}{b}.$$

To finish the proof, we will use the following claim, to be proved soon.

Claim C.7. There exist constants $m_0 \geq 5, c_5, c_6 > 0$ such that for all $m \geq m_0$,

$$\mathbb{P}[Z_{2m} \geq 4m^2 \mid Z_m \geq m^2] \geq c_5 e^{-c_6 m}. \tag{C.5}$$

and

$$\mathbb{P}[Z_{2m+1} \geq (2m+1)^2 \mid Z_m \geq m^2] \geq c_5 e^{-c_6 m}. \tag{C.6}$$

Now for $n = 2^k$, multiplying (C.5) for $m = 2^3, \dots, 2^{k-1}$ one obtains that there exists $c_7, c_8 > 0$ such that

$$\mathbb{P}[Z_n \geq n^2] \geq \mathbb{P}[Z_8 \geq 8^2] \cdot c_5^k \exp(-c_6(2^3 + \dots + 2^k)) \geq c_7 \exp(-c_8 n).$$

For general $n \geq 5$, multiplying (C.5) or (C.6) for $m = f(n) := \lfloor \frac{n}{2} \rfloor, f^{\circ 2}(n), \dots$ one obtains

$$\mathbb{P}[Z_n \geq n^2] \geq c \cdot c_5^{\log n} \exp(-c_6(\lfloor \frac{n}{2} \rfloor + \dots)) \geq c_7 \exp(-c_8 n). \quad \square$$

Now we proceed to show Claim C.7. We will just show (C.5), the case (C.6) being similar.

Proof of Claim C.7. For convenience, we refer to the lexicographic orderings of labelled vertices of the Galton–Watson tree as increasing from left to right (see page 136 of [18] for the labeling). On the event $\{Z_m \geq m^2\}$, let $u_i, i = 1, \dots, m^2$ be the first m^2 individuals (from left to right) in generation m and denote by Y_i the number of children of u_i in generation $2m$. The random variables Y_i ’s are independent and have the same law as

Z_m . So by Theorem 1 on page 19 of [1] and formula (4) on page 4 of [1] one has that $\mathbb{P}[Y_i > 0] = \frac{2}{(1-1/b)m}(1 + o(1))$ and $\text{Var}(Y_i) = \frac{(b-1)m}{b} \asymp m$. Obviously we have

$$\mathbb{P}[Z_{2m} \geq 4m^2 \mid Z_m \geq m^2] \geq \mathbb{P}[Y_1 + \dots + Y_{m^2} \geq 4m^2]. \tag{C.7}$$

Write $J := \{j \in [1, m^2]: Y_j > 0\}$. Then there exists a constant $c_9 > 0$ such that $\mathbb{P}[|J| > c_9 m] \geq c_9$. In fact $\mathbb{E}[|J|] = \sum_{j=1}^{m^2} \mathbb{P}[Y_j > 0] \asymp m$. By independence of Y_j 's, $\mathbb{E}[|J|^2] = \mathbb{E}[|J|]^2 + \text{Var}(|J|) \asymp m^2 + m \cdot \frac{\lambda}{m}(1 - \frac{\lambda}{m}) \asymp m^2$, where $\lambda = \frac{2}{1-1/b}$. By Paley-Zygmund inequality one has $\mathbb{P}[|J| > c_9 m] \geq c_9$ for some small $c_9 > 0$.

Suppose $W_j = W_j(m)$'s are i.i.d. random variables with the law of $\frac{Y_j}{m}$ conditioned on $Y_j > 0$. Then by Theorem 1 and 2 in [1, page 19-20] W_j converges in distribution to an exponential distribution $\exp(\lambda)$ with $\lambda = \frac{2}{1-1/b}$ as $m \rightarrow \infty$. Hence there exists $m_0 \geq 5$ and $c_{10} > 0$ such that $\mathbb{P}(W_j \geq \frac{4}{c_9}) \geq e^{-c_{10}}$ for all $m \geq m_0$.

Therefore for $m \geq m_0$,

$$\begin{aligned} \mathbb{P}[Y_1 + \dots + Y_{m^2} \geq 4m^2] &= \mathbb{P}\left[\sum_{j \in J} Y_j \geq 4m^2\right] \\ &\geq \mathbb{P}\left[|J| > c_9 m, \sum_{j \in J} \frac{Y_j}{m} \geq 4m\right] \\ &\geq c_9 \cdot \mathbb{P}\left[\sum_{j=1}^{c_9 m} W_j \geq 4m\right] \\ &\geq c_9 \cdot \left[\mathbb{P}(W_j \geq \frac{4}{c_9})\right]^{c_9 m} \geq c_9 \exp(-c_9 c_{10} m). \end{aligned} \tag{C.8}$$

By (C.7) and (C.8) one obtains (C.5). □

Proof of Proposition C.1. The upper bound has already been proved for general nonunimodular transitive graphs in Corollary 4.17. Next we focus on the lower bound.

Let $\eta_x := (x, x_1, x_2, \dots)$ be the ray starting from x that represents ξ . Let $E_m := \{x \leftrightarrow x_m, x \not\leftrightarrow x_{m+1}\}$ denote the event that the path from x to x_m is open but the edge (x_m, x_{m+1}) is not open in Bernoulli($\frac{1}{b}$) percolation on \mathbb{T}_{b+1} . Let C_x denote the cluster of x in this Bernoulli($\frac{1}{b}$) percolation. Then

$$\mathbb{P}[E_m] = \frac{1}{b^m} \left(1 - \frac{1}{b}\right).$$

Conditioned on E_m , one has $|C_x \cap L_0(x)| = X'_0 + X'_1 + \dots + X'_m$ where $X'_i \equiv 1, X'_1, \dots, X'_m$ are independent and X'_i has the law of the size of the i -th generation of a "critical" Galton-Watson tree. The first generation of this Galton-Watson tree has binomial distribution $\text{Bin}(b - 1, \frac{1}{b})$ and the other generations have binomial progeny distribution $\text{Bin}(b, \frac{1}{b})$.

Similar to the proof of Lemma C.6, one can show that there exist constants $m_0 > 0, c > 0$ such that

$$\mathbb{P}[X'_m \geq m^2] \geq e^{-cm}, \forall m \geq m_0.$$

Now for $k \geq m_0^2$ and taking $m = \lceil \sqrt{k} \rceil$,

$$\mathbb{P}[|C_x \cap L_0(x)| \geq k] \geq \mathbb{P}[E_m] \cdot \mathbb{P}[X'_m \geq m^2] \geq c_1 \exp(-c_2 \sqrt{k}).$$

Taking sufficiently small c_1, c_2 the above inequality would also hold for $k \in [1, m_0^2]$. Since for the toy model the distribution of $|C_x \cap L_n(x)|$ given $|C_x \cap L_n(x)| > 0$ is the same as $|C_x \cap L_0(x)|$ and $\mathbb{P}[|C_x \cap L_n(x)| > 0] = e^{-t_0 n} = \frac{1}{b^n}$, one has the desired lower bound for all $n \geq 0, k \geq 1$:

$$\mathbb{P}[|C_x \cap L_n(x)| \geq k] \geq c_1 \exp(-t_0 n - c_2 \sqrt{k}).$$

By Lemma 3.2 we are done. □

Corollary C.8. For x -WUSF on the toy model $(\mathbb{T}_{b+1}, \Gamma_\xi)$, there exists a constant $C > 1$ such that for all $n \geq 0, k \geq 1$,

$$\frac{(k!)^2 e^{-t_0 n}}{C^k} \leq \mathbb{E}[|\mathfrak{F}_x \cap L_n(x)|^k] \leq C^k (k!)^2 e^{-t_0 n} \tag{C.9}$$

Proof. The upper bound has already been proved in (4.43) for general nonunimodular transitive graphs. In fact we use these k th moments to prove the upper bound in Proposition C.1.

Now we look at the lower bound.

By Proposition C.1, there exists a constant $c > 0$ such that for all $n \geq 0, k \geq 1$,

$$\mathbb{P}[|\mathfrak{F}_x \cap L_n(x)| \geq y] \geq e^{-t_0 n - c\sqrt{y}}.$$

Hence

$$\begin{aligned} \mathbb{E}[|\mathfrak{F}_x \cap L_n(x)|^k] &= \int_0^\infty k y^{k-1} \mathbb{P}[|\mathfrak{F}_x \cap L_n(x)| \geq y] dy \\ &\geq \int_0^\infty k y^{k-1} e^{-t_0 n - c\sqrt{y}} dy \\ &\geq \frac{(k!)^2 e^{-t_0 n}}{C^k} \end{aligned} \tag{C.10}$$

for some large constant $C > 0$. □

C.2 The intersection $T_x \cap L_n(x)$ when $n \geq 0$

Proposition C.9. For WUSF on the toy model $(\mathbb{T}_{b+1}, \Gamma_\xi)$, there exists a constant $C > 1$ such that for all $n \geq 0, k \geq 1$,

$$\frac{(k!)^2 (n \vee 1) e^{-t_0 n}}{C^k} \leq \mathbb{E}[|T_x \cap L_n(x)|^k] \leq C^k (k!)^2 (n \vee 1) e^{-t_0 n} \tag{C.11}$$

and

$$\mathbb{P}[|T_x \cap L_n(x)| \geq k] = (n \vee 1) e^{-t_0 n - \Theta(\sqrt{k})} \tag{C.12}$$

Lemma C.10. Let Z_n denote size of the n -th generation of a critical Galton–Watson tree with progeny distribution $\text{Bin}(b, \frac{1}{b})$. Suppose X_n has the same law as Z_n and the X_i 's are independent. Then there is a constant $c > 0$ such that for all $k \geq 1, n \geq 1, m \geq 0$,

$$\mathbb{E}[(X_n + \dots + X_{n+m})^k] \leq c^k (k!) (m+1)(n+m)^{k-1}. \tag{C.13}$$

Similarly there is a constant $c > 0$ such that for all $k \geq 1, m \geq 1$,

$$\mathbb{E}[(X_0 + \dots + X_m)^k] \leq c^k (k!) m^k. \tag{C.14}$$

Proof. We will only show (C.13). Since $X_0 \equiv 1$, the inequality (C.14) can be obtained from (C.13) easily. Observe that

$$\begin{aligned} \mathbb{E}[(X_n + \dots + X_{n+m})^k] &= \sum_{\substack{k_0, \dots, k_m \geq 0 \\ k_0 + \dots + k_m = k}} \frac{k!}{k_0! \dots k_m!} \prod_{i=0}^m \mathbb{E}[X_{n+i}^{k_i}] \\ &\stackrel{\text{Lem.C.4}}{\leq} \sum_{\substack{k_0, \dots, k_m \geq 0 \\ k_0 + \dots + k_m = k}} \frac{k!}{k_0! \dots k_m!} \prod_{i=0}^m (C^{k_i} k_i! (n+i)^{k_i-1} \vee 1) \\ &= C^k k! \sum_{\substack{k_0, \dots, k_m \geq 0 \\ k_0 + \dots + k_m = k}} \prod_{i=0}^m ((n+i)^{k_i-1} \vee 1). \end{aligned} \tag{C.15}$$

Claim C.11. For all $n \geq 1, m \geq 0, k \geq 1$,

$$\sum_{\substack{k_0, \dots, k_m \geq 0 \\ k_0 + \dots + k_m = k}} \prod_{i=0}^m ((n+i)^{k_i-1} \vee 1) \leq 3^k (m+1)(n+m)^{k-1} \tag{C.16}$$

We prove Claim C.11 by induction on m . When $m = 0$, the left hand side of (C.16) equals $n^{k-1} \vee 1 = n^{k-1}$ and the right hand side equals $3^k n^{k-1}$, so (C.16) holds for $m = 0$.

Now suppose $m \geq 1$ and (C.16) holds for $m-1$. Using the convention that $\sum_{k_m=1}^{k-1} \cdot = 0$ when $k = 1$ one has the left hand side \mathcal{L} of (C.16) satisfies

$$\begin{aligned} \mathcal{L} &= \sum_{k_m=0}^k ((n+m)^{k_m-1} \vee 1) \cdot \sum_{\substack{k_0, \dots, k_{m-1} \geq 0 \\ k_0 + \dots + k_{m-1} = k - k_m}} \prod_{i=0}^{m-1} ((n+i)^{k_i-1} \vee 1) \\ &= (n+m)^{k-1} + \sum_{\substack{k_0, \dots, k_{m-1} \geq 0 \\ k_0 + \dots + k_{m-1} = k}} \prod_{i=0}^{m-1} ((n+i)^{k_i-1} \vee 1) \\ &\quad + \sum_{k_m=1}^{k-1} ((n+m)^{k_m-1} \vee 1) \cdot \sum_{\substack{k_0, \dots, k_{m-1} \geq 0 \\ k_0 + \dots + k_{m-1} = k - k_m}} \prod_{i=0}^{m-1} ((n+i)^{k_i-1} \vee 1) \\ &\leq (n+m)^{k-1} + 3^k m (n+m-1)^{k-1} + \sum_{k_m=1}^{k-1} (n+m)^{k_m-1} 3^{k-k_m} m (n+m-1)^{k-k_m-1} \\ &\leq (n+m)^{k-1} + 3^k m (n+m-1)^{k-1} + \sum_{k_m=1}^{k-1} 3^{k-k_m} m (n+m)^{k-2} \\ &= 3^k (m+1)(n+m)^{k-1} \left[\frac{1}{3^k (m+1)} + \frac{m}{m+1} \left(1 - \frac{1}{n+m}\right)^{k-1} \right. \\ &\quad \left. + \frac{m}{(m+1)(n+m)} \sum_{k_m=1}^{k-1} 3^{-k_m} \right] \\ &\leq 3^k (m+1)(n+m)^{k-1} \left[\frac{1}{3(m+1)} + \frac{m}{m+1} + \frac{m}{(m+1)(n+m)} \cdot \frac{1}{2} \right] \\ &< 3^k (m+1)(n+m)^{k-1}. \end{aligned} \tag{C.17}$$

Hence by induction the claim holds.

Now the conclusion (C.13) follows from (C.15) and (C.16). □

Proof of Proposition C.9. Given the upper bound in (C.11), the proof of the upper bound of (C.12) is the same as Corollary 4.17 and thus we omit the details.

Given the lower bound of (C.12), the lower bound of (C.11) can be obtained in the same way as (C.10).

So it suffices to show the upper bound of (C.11) and the lower bound of (C.12).

Let $\eta_x = (x_0, x_1, x_2, \dots)$ be the ray starting from x representing the end ξ . Recall the events A_k and $B_{k,k'}$ we defined in the proof of Proposition 3.3; see Figure 7 for an example of $B_{k,k'}$.

Also recall that

$$\mathbb{P}[A_k] = \begin{cases} \frac{b}{b+1}, & k = 0 \\ \frac{b-1}{b+1} \cdot \frac{1}{b^k}, & k > 0. \end{cases}$$

and

$$\mathbb{P}[B_{k,k'}] = \mathbb{P}[A_k] \cdot \frac{b-1}{b} \cdot \frac{1}{b^{k'}} \asymp \frac{1}{b^{k+k'}}, \quad \forall k, k' \geq 0.$$

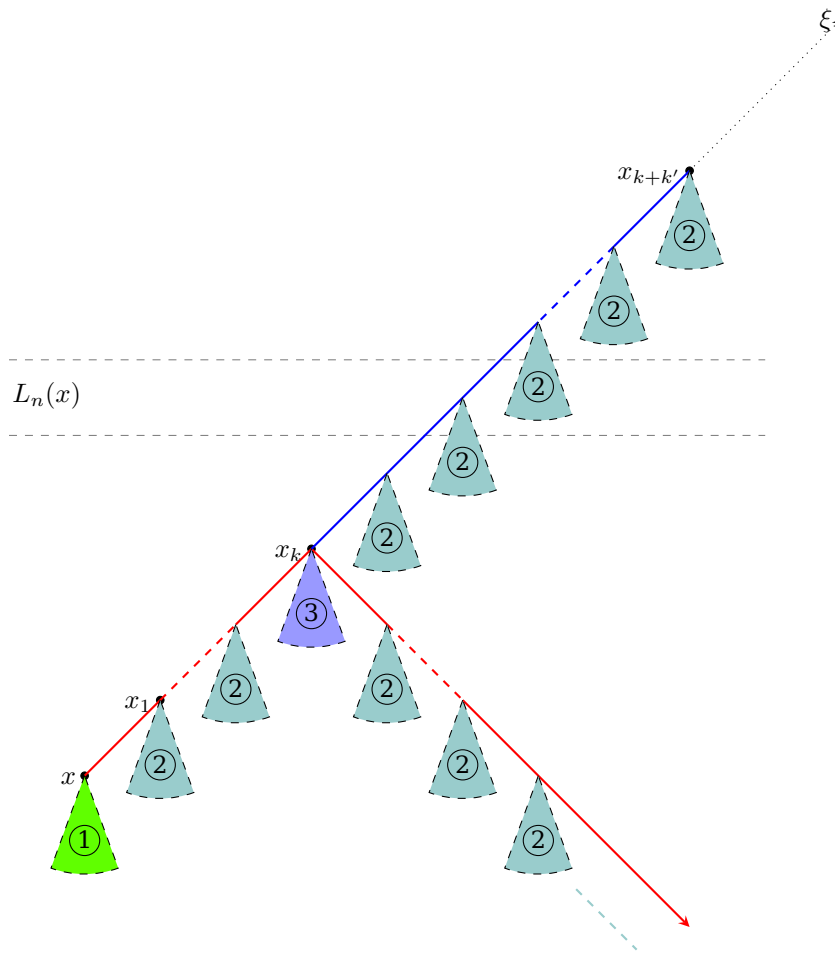


Figure 7: A sketch of the tree T_x conditioned on a typical $B_{k,k'}$. The red line denotes the future of x . There are three types of independent “critical” Galton–Watson trees attached to the future of x and vertices $x_{k+1}, \dots, x_{k+k'}$, and the only difference among them is the progeny distribution of the first generation. The distribution of the first generation distribution for the three types are $\text{Bin}(b, \frac{1}{b})$, $\text{Bin}(b - 1, \frac{1}{b})$ and $\text{Bin}(b - 2, \frac{1}{b})$ respectively.

Observation C.12. Suppose X_n, X'_n, X''_n have the law of the size of the n -th generation of “critical” Galton–Watson trees with first generation progeny distributions $\text{Bin}(b, \frac{1}{b})$, $\text{Bin}(b - 1, \frac{1}{b})$ and $\text{Bin}(b - 2, \frac{1}{b})$ respectively (for generation at least 2, the progeny distribution is always $\text{Bin}(b, \frac{1}{b})$.) In particular, when $b = 2$, $X''_0 = 1$ and $X''_n = 0$ for $n \geq 1$. Moreover assume all the random variables $\{X_i, X'_i, X''_i : i = 0, 1, 2, \dots\}$ are independent.

It is easy to see from Figure 7 that conditioned on the event $B_{k,k'}$, one has

1. if $n > k + k'$, then $|T_x \cap L_n(x)| = 0$;
2. if $0 \leq k < n \leq k + k'$, then $|T_x \cap L_n(x)|$ has the same distribution as $\sum_{j=n}^{k+k'} X'_{j-n}$;
3. if $0 < n \leq k \leq k + k'$, then $|T_x \cap L_n(x)|$ has the same distribution as $\sum_{j=n}^{k-1} X'_{j-n} + X''_{k-n} + \sum_{j=k+1}^{k+k'} X'_{j-n} + \sum_{j=n}^{k-1} \tilde{X}'_{j-n}$, where all these random variables are independent and \tilde{X}'_j has the same distribution as X'_j ;
4. if $n \leq 0$, then $|T_x \cap L_n(x)|$ has the same distribution as $X_{-n} + \sum_{j=1}^{k-1} X'_{j-n} +$

$\sum_{j=k+1}^{k+k'} X'_{j-n} + X''_{k-n} + \sum_{j=n}^{k-1} \tilde{X}'_{j-n}$, where all these random variables are independent and \tilde{X}'_j has the same distribution as X'_j .

Note that the random variables X'_j, X''_j and \tilde{X}'_j can be stochastically dominated by X_j , which has the same law as Z_j in Lemma C.4. Also note that $B_{k,k'}$ are disjoint and $\mathbb{P}[\bigcup_{k,k' \geq 0} B_{k,k'}] = 1$.

Therefore for $n > 0, t \geq 2$, one has that

$$\begin{aligned}
 \text{LHS} &:= \mathbb{E}[|T_x \cap L_n(x)|^t] \\
 &= \sum_{k=0}^{n-1} \sum_{k'=n-k}^{\infty} \mathbb{P}[B_{k,k'}] \cdot \mathbb{E}\left[\left(\sum_{j=n}^{k+k'} X'_{j-n}\right)^t\right] \\
 &\quad + \sum_{k=n}^{\infty} \sum_{k'=0}^{\infty} \mathbb{P}[B_{k,k'}] \cdot \mathbb{E}\left[\left(\sum_{j=n}^{k-1} X'_{j-n} + \sum_{j=k+1}^{k+k'} X'_{j-n} + X''_{k-n} + \sum_{j=n}^{k-1} \tilde{X}'_{j-n}\right)^t\right] \\
 &\leq C^t \sum_{k=0}^{n-1} \sum_{k'=n-k}^{\infty} \frac{1}{b^{k+k'}} \cdot \mathbb{E}\left[\left(\sum_{j=n}^{k+k'} X_{j-n}\right)^t\right] \\
 &\quad + C^t \sum_{k=n}^{\infty} \sum_{k'=0}^{\infty} \frac{1}{b^{k+k'}} \cdot \mathbb{E}\left[\left(\sum_{j=n}^{k+k'} X_{j-n}\right)^t\right] \\
 &\stackrel{(C.14)}{\leq} C^t \sum_{k=0}^{n-1} \sum_{k'=n-k}^{\infty} \frac{1}{b^{k+k'}} \cdot c^t(t!) [(k+k'-n) \vee 1]^t \\
 &\quad + C^t \sum_{k=n}^{\infty} \sum_{k'=0}^{\infty} \frac{1}{b^{k+k'}} \cdot c^t(t!) [(k+k'-n) \vee 1]^t \\
 &\leq c_2^t(t!) \sum_{j=0}^{\infty} (n+1) \cdot \frac{1}{b^{n+j}} \cdot (j \vee 1)^t \\
 &\leq c_3^t(t!)^2 \frac{n}{b^n}. \tag{C.18}
 \end{aligned}$$

This proves the upper bound of (C.11) for $n \geq 1$. The case of $n = 0$ can be dealt in a similar way and we omit the details.

Next we turn to the proof of the lower bound of (C.12).

As mentioned earlier using a similar proof to Lemma C.6, there exist constants $m_0 > 0, c > 0$ such that

$$\mathbb{P}[X'_m \geq m^2] \geq e^{-cm}, \forall m \geq m_0.$$

Notice for different pairs of (k, k') , the event $B_{k,k'}$ are disjoint. Conditioned on $B_{k,k'}$ with $k+k' > n$, one has that $|T_x \cap L_n(x)|$ stochastically dominates $X'_{k+k'-n}$. Hence for $t \geq m_0^2$, one has the desired lower bound of (C.12):

$$\begin{aligned}
 \mathbb{E}[|T_x \cap L_n(x)| \geq t] &\geq \sum_{k+k'=n+\lceil\sqrt{t}\rceil} \mathbb{P}[B_{k,k'}] \cdot \mathbb{P}[X'_{\lceil\sqrt{t}\rceil} \geq \lceil\sqrt{t}\rceil^2] \\
 &\geq \sum_{k+k'=n+\lceil\sqrt{t}\rceil} \frac{1}{b^{n+\lceil\sqrt{t}\rceil}} \cdot e^{-c\lceil\sqrt{t}\rceil} \\
 &\geq \frac{(n \vee 1)}{b^n} \cdot e^{-c_1\sqrt{t}} \tag{C.19}
 \end{aligned}$$

For $t \in [1, m_0^2]$, taking a small constant c_1 would be enough. □

Remark C.13. One can use (C.13) and a similar calculation to (C.18) in the above proof to give another proof of Proposition 4.15 for the toy model. Just use the events E_m in the proof of Proposition C.1 instead of $B_{k,k'}$.

C.3 Intersections with $L_{-n}(x)$ when $n > 0$

Proposition C.14. For WUSF on the toy model $(\mathbb{T}_{b+1}, \Gamma_\xi)$, there exist constants c_1, c_2, c_3 such that for all $n > 0, k \geq 1$,

$$\mathbb{E}[|T_x \cap L_{-n}(x)|^k] \leq c_1^k (k!)^2 n^k \tag{C.20}$$

and

$$\mathbb{P}[|T_x \cap L_{-n}(x)| \geq k] \leq c_2 e^{-c_3 \sqrt{\frac{k}{n}}} \tag{C.21}$$

Remark C.15. Comparing the upper bound of (C.11) for the toy model with (4.53) in Proposition 4.20, the power of $k!$ in (4.53) for general nonunimodular transitive graphs in Proposition 4.20 might be not optimal. Similarly the power of $k!$ on the right hand of (C.20) suggests that the power of $k!$ in (4.54) might be not optimal. Besides, the power of n in (C.20) suggests that the power of n in (4.54) might also be not optimal.

Remark C.16. The corresponding upper bounds for $\mathfrak{T}_x \cap L_{-n}(x)$ have been proved for general nonunimodular graphs; see (4.44) in Proposition 4.15 and (4.47) in Corollary 4.17.

Remark C.17. For the toy model, one can get a lower bound with the form of the upper bound in (4.47) when $1 \leq k \leq c_4 n$. To see this, note that for $1 \leq k \leq c_4 n$, $\mathbb{P}[|\mathfrak{T}_x \cap L_{-n}(x)| \geq k] \geq \mathbb{P}[X_n \geq c_4 n] \geq \frac{c_5}{n} \geq \frac{c_5}{n} e^{-c_6 \sqrt{\frac{k}{n}}}$ by Theorem 1 on page 19 of [1], where c_5 depends on c_4 .

Similarly for the toy model, when $1 \leq k \leq c_4 n$ we also have a lower bound with the form of the upper bound in (C.21) (in this case it is just $\mathbb{P}[|T_x \cap L_{-n}(x)| \geq k] \geq c_5$) and we omit the proof.

Question C.18. Are there lower bounds for $\mathbb{P}[|\mathfrak{T}_x \cap L_{-n}(x)| \geq k]$ and $\mathbb{P}[|T_x \cap L_{-n}(x)| \geq k]$ with the same form as the upper bound in (4.47) and (C.21) respectively for all $n, k \geq 1$?

Proof of Proposition C.14. The proof of (C.20) is quite similar to (C.18). By Observation C.12, for $t \geq 2$ one has

$$\begin{aligned} \text{LHS} &:= \mathbb{E}[|T_x \cap L_{-n}(x)|^t] \\ &= \sum_{k,k'=0}^{\infty} \mathbb{P}[B_{k,k'}] \mathbb{E}\left[\left(X_n + \sum_{j=1}^{k-1} X'_{j+n} + \sum_{j=k+1}^{k+k'} X'_{j+n} + X''_{k+n} + \sum_{j=-n}^{k-1} \tilde{X}'_{j+n} \right)^t \right] \\ &\leq \sum_{k,k'=0}^{\infty} \frac{c_1}{b^{k+k'}} c_2^t \mathbb{E}\left[\left(\sum_{j=0}^{k+k'+n} X_j \right)^t \right] \\ &\stackrel{(C.14)}{\leq} \sum_{k,k'=0}^{\infty} \frac{c_3^t}{b^{k+k'}} t! (k+k'+n)^t = \sum_{j=0}^{\infty} c_3^t t! \frac{j+1}{b^j} (j+n)^t \\ &\leq c_3^t t! n^t + \sum_{j=1}^{\infty} c_3^t t! \frac{2j}{b^j} (j+n)^t \leq c_3^t t! n^t + 2bc_3^t t! \int_0^{\infty} \frac{x(x+n)^t}{b^x} dx \\ &= c_3^t t! n^t + 2bc_3^t t! \sum_{j=0}^t \binom{t}{j} \int_0^{\infty} x^{j+1} n^{t-j} b^{-x} dx \end{aligned}$$

$$\begin{aligned}
 &= c_3^t t! n^t + 2bc_3^t t! \sum_{j=0}^t \binom{t}{j} n^{t-j} \frac{(j+1)!}{(\log b)^{j+2}} \\
 &\leq c_3^t t! n^t + 2bc_4^t t! \sum_{j=0}^t t!(j+1)n^t \leq c_5^t (t!)^2 n^t
 \end{aligned} \tag{C.22}$$

Similar to the proof of Corollary 4.17, one has for $c_6 < \frac{1}{2\sqrt{c_5}}$,

$$\begin{aligned}
 \mathbb{E} \left[\exp \left(c_6 \sqrt{\frac{|T_x \cap L_{-n}(x)|}{n}} \right) \right] &\leq 2 \sum_{k=0}^{\infty} \frac{c_6^{2k}}{(2k)!} \frac{1}{n^k} \mathbb{E}[|T_x \cap L_{-n}(x)|^k] \\
 &\stackrel{(C.22)}{\leq} 2 \sum_{k=0}^{\infty} \frac{c_6^{2k}}{(2k)!} c_5^k (k!)^2 < 3
 \end{aligned} \tag{C.23}$$

Hence by Markov's inequality one obtains (C.21). \square

For the toy model, the results in Proposition C.1 and Corollary C.8 also hold for the past $\mathfrak{P}(x)$ and we omit the details. For the future, it is not interesting since $|\mathfrak{F}(x, \infty) \cap L_n(x)| \in \{0, 1, 2\}$ almost surely.

References

- [1] Krishna B. Athreya and Peter E. Ney, *Branching processes*, Springer-Verlag, New York-Heidelberg, 1972, Die Grundlehren der mathematischen Wissenschaften, Band 196. MR0373040
- [2] Itai Benjamini, Russell Lyons, Yuval Peres, and Oded Schramm, *Group-invariant percolation on graphs*, *Geom. Funct. Anal.* **9** (1999), no. 1, 29–66. MR1675890
- [3] Itai Benjamini, Russell Lyons, Yuval Peres, and Oded Schramm, *Uniform spanning forests*, *Ann. Probab.* **29** (2001), no. 1, 1–65. MR1825141
- [4] Reinhard Diestel and Imre Leader, *A conjecture concerning a limit of non-Cayley graphs*, *J. Algebraic Combin.* **14** (2001), no. 1, 17–25. MR1856226
- [5] Rick Durrett, *Probability—theory and examples*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 49, Cambridge University Press, Cambridge, 2019, Fifth edition of [MR1068527]. MR3930614
- [6] Alex Eskin, David Fisher, and Kevin Whyte, *Coarse differentiation of quasi-isometries I: Spaces not quasi-isometric to Cayley graphs*, *Ann. of Math. (2)* **176** (2012), no. 1, 221–260. MR2925383
- [7] Geoffrey Grimmett, *Percolation*, second ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 321, Springer-Verlag, Berlin, 1999. MR1707339
- [8] Olle Häggström, *Uniform and minimal essential spanning forests on trees*, *Random Structures Algorithms* **12** (1998), no. 1, 27–50. MR1637387
- [9] Tom Hutchcroft, *Wired cycle-breaking dynamics for uniform spanning forests*, *Ann. Probab.* **44** (2016), no. 6, 3879–3892. MR3572326
- [10] Tom Hutchcroft, *Nonuniqueness and mean-field criticality for percolation on nonunimodular transitive graphs*, *J. Amer. Math. Soc.* **33** (2020), no. 4, 1101–1165. MR4155221
- [11] Tom Hutchcroft, *Universality of high-dimensional spanning forests and sandpiles*, *Probab. Theory Related Fields* **176** (2020), no. 1-2, 533–597. MR4055195
- [12] Tom Hutchcroft and Asaf Nachmias, *Indistinguishability of trees in uniform spanning forests*, *Probab. Theory Related Fields* **168** (2017), no. 1-2, 113–152. MR3651050
- [13] Antal A. Járai and Frank Redig, *Infinite volume limit of the abelian sandpile model in dimensions $d \geq 3$* , *Probab. Theory Related Fields* **141** (2008), no. 1-2, 181–212. MR2372969

- [14] Gregory F. Lawler, *Loop-erased random walk*, Perplexing problems in probability, Progr. Probab., vol. 44, Birkhäuser Boston, Boston, MA, 1999, pp. 197–217. MR1703133
- [15] Gregory F. Lawler, *Topics in loop measures and the loop-erased walk*, Probab. Surv. **15** (2018), 28–101. MR3770886
- [16] Russell Lyons, *A simple path to Biggins’ martingale convergence for branching random walk*, Classical and modern branching processes (Minneapolis, MN, 1994), IMA Vol. Math. Appl., vol. 84, Springer, New York, 1997, pp. 217–221. MR1601749
- [17] Russell Lyons, Benjamin J. Morris, and Oded Schramm, *Ends in uniform spanning forests*, Electron. J. Probab. **13** (2008), no. 58, 1702–1725. MR2448128
- [18] Russell Lyons and Yuval Peres, *Probability on trees and networks*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 42, Cambridge University Press, New York, 2016, Available at <http://pages.iu.edu/~rdlyons/>. MR3616205
- [19] A. G. Pakes, *On the critical Galton-Watson process with immigration*, J. Austral. Math. Soc. **12** (1971), 476–482. MR0307370
- [20] Erol A. Peköz and Adrian Röllin, *New rates for exponential approximation and the theorems of Rényi and Yaglom*, Ann. Probab. **39** (2011), no. 2, 587–608. MR2789507
- [21] Robin Pemantle, *Choosing a spanning tree for the integer lattice uniformly*, Ann. Probab. **19** (1991), no. 4, 1559–1574. MR1127715
- [22] Gábor Pete and Ádám Timár, *The free uniform spanning forest is disconnected in some virtually free groups, depending on the generator set*, arXiv preprint arXiv:2006.06387v1 (2020).
- [23] Paolo M. Soardi, *Potential theory on infinite networks*, Lecture Notes in Mathematics, vol. 1590, Springer-Verlag, Berlin, 1994. MR1324344
- [24] Pengfei Tang, *Heavy Bernoulli-percolation clusters are indistinguishable*, Ann. Probab. **47** (2019), no. 6, 4077–4115. MR4038049
- [25] Ádám Timár, *Percolation on nonunimodular transitive graphs*, Ann. Probab. **34** (2006), no. 6, 2344–2364. MR2294985
- [26] Ádám Timár, *Indistinguishability of the components of random spanning forests*, Ann. Probab. **46** (2018), no. 4, 2221–2242. MR3813990
- [27] V. I. Trofimov, *Groups of automorphisms of graphs as topological groups*, Mat. Zametki **38** (1985), no. 3, 378–385, 476. MR0811571
- [28] C. Z. Wei, *Convergence rates for the critical branching process with immigration*, Statist. Sinica **1** (1991), no. 1, 175–184. MR1101321
- [29] Wolfgang Woess, *Random walks on infinite graphs and groups*, Cambridge Tracts in Mathematics, vol. 138, Cambridge University Press, Cambridge, 2000. MR1743100
- [30] Wolfgang Woess, *Lamplighters, Diestel–Leader graphs, random walks, and harmonic functions*, Combin. Probab. Comput. **14** (2005), no. 3, 415–433. MR2138121

Acknowledgments. We thank Russell Lyons for many helpful discussions and suggestions. We thank Ádám Timár and Gábor Pete for sharing their recent work [22] and pointing out the branching number part of Remark 5.12. We thank the referees for their helpful comments and suggestions. Especially one referee pointed out a serious mistake in the previous version and proposed a useful rescue, namely the tree-graph inequality method. This also leads to sharper bounds on certain estimates in Section 4. The study of $\mathbb{P}[|\mathfrak{T}_x \cap L_n(x)| > k]$ in Proposition C.1 was also suggested by the referee. Question 5.11 was suggested by another referee.