

# FIRST PASSAGE PERCOLATION ON HYPERBOLIC GROUPS

RIDDHIPRATIM BASU AND MAHAN MJ

ABSTRACT. We study first passage percolation (FPP) on a Gromov-hyperbolic group  $G$  with boundary  $\partial G$  equipped with the Patterson-Sullivan measure  $\nu$ . We associate an i.i.d. collection of random passage times to each edge of a Cayley graph of  $G$ , and investigate classical questions about asymptotics of first passage time as well as the geometry of geodesics in the FPP metric. Under suitable conditions on the passage time distribution, we show that the ‘velocity’ exists in  $\nu$ -almost every direction  $\xi \in \partial G$ , and is almost surely constant by ergodicity of the  $G$ -action on  $\partial G$ . For every  $\xi \in \partial G$ , we also show almost sure coalescence of any two geodesic rays directed towards  $\xi$ . Finally, we show that the variance of the first passage time grows linearly with word distance along word geodesic rays in every fixed boundary direction. This provides an affirmative answer to a conjecture in [BZ12, BT17].

## CONTENTS

1. Introduction	1
2. Preliminaries	4
3. Automatic structure, Patterson-Sullivan measures and frequency	9
4. Approximating FPP geodesics	15
5. Velocity	18
6. Direction of $\omega$ -geodesic rays	26
7. Coalescence	30
8. Linear growth of variance along word geodesics	39
9. Discussion and Future Directions	42
Appendix A.	43
Acknowledgments	45
References	46

## 1. INTRODUCTION

First passage percolation (FPP) is a well-known probabilistic model for fluid flow through random media. It assigns i.i.d. weights to edges of a graph and analyses the first passage time (i.e., the weight of the minimum weight path) as well as the geodesic (the optimal path) between any two points. For the Cayley graph of  $\mathbb{Z}^2$  with respect to standard generators, this was introduced by Hammersley and Welsh [HW65] more than fifty years back. While  $\mathbb{Z}^2$  and more generally  $\mathbb{Z}^d$  have been

---

*Date:* September 8, 2019.

*2010 Mathematics Subject Classification.* 60K35, 82B43, 20F67 (20F65, 51F99, 60J50) .

*Key words and phrases.* first passage percolation, hyperbolic group, Patterson-Sullivan measure, ergodicity.

investigated thoroughly, the literature on other background geometries is sparse. In the special case of (Gromov) hyperbolic geometry [Gro85], some results have been established by Benjamini, Tessler and Zeitouni [BZ12, BT17] and a number of test questions have been raised there. In particular, [BZ12] established the tightness of fluctuations of the passage time from the center to the boundary of a large ball, and [BT17] established the almost sure existence of bi-geodesics in hyperbolic spaces. The aim of this paper is to undertake a more detailed study into FPP on Cayley graphs of hyperbolic groups and address the fundamental questions that have been thoroughly investigated and, in many cases, answered for FPP on the Euclidean lattice. It turns out that some of the basic questions, (e.g., the existence of the limiting constant for the scaled expected passage time along a direction) becomes harder in our setting, whereas some of the other questions (e.g., fluctuations of passage times and coalescence of geodesics) which are unresolved (or only resolved under strong unproven assumptions) in the Euclidean settings can be addressed in the hyperbolic setting primarily due to the nice behavior forced upon geodesics by the underlying geometry.

Let us first briefly describe our main results informally, while postponing the formal statements and precise assumptions to later in the paper. For the rest of this paper,  $G$  will be a hyperbolic group,  $\Gamma = \Gamma(G, S)$  a Cayley graph of  $G$  with respect to a finite (symmetric) generating set  $S$ . Let  $\nu$  denote the Patterson-Sullivan measure on the boundary  $\partial G$  of  $G$ . We shall consider first passage percolation with i.i.d. positive weights on the edges of  $\Gamma$  coming from a sufficiently nice distribution. The main results of this paper are the following:

- (1) **Velocity exists and is constant:** One of the first results in Euclidean FPP is that, under minimal conditions, the expected first passage time (from origin) in any direction grows linearly in the Euclidean distance and there exists a limiting constant (often referred to as time constant or velocity, we shall use the latter). This is a straightforward consequence of sub-additivity for rational directions. One can also show that the velocity varies continuously with the direction and upgrade this to a shape theorem [CD81]. In the hyperbolic setting, the situation is quite different in flavor. We parametrize directions in  $G$  by points in the boundary  $\partial G$ . We show that for  $\nu$ -almost every  $\xi \in \partial G$ , the velocity (i.e., limit of the linearly scaled expected passage time)  $v(\xi)$  exists along a word geodesic ray from the identity element in the direction  $\xi$  (Theorem 5.1). Further,  $v(\xi)$  is constant almost everywhere. We provide examples (Section 5.3) to show that we cannot replace "almost everywhere" by "everywhere".
- (2) **Linear Variance:** Once the first order behavior has been established, the next natural question is to understand the fluctuation of passage times along the word geodesic in a fixed direction. This question remains unresolved in Euclidean FPP, though it is widely believed that the fluctuations are subdiffusive, exhibiting a power law behavior with exponent  $\chi = \chi(d) < \frac{1}{2}$  for all dimensions  $d$ . In particular, it is predicted that  $\chi(2) = \frac{1}{3}$ . However the best rigorous upper bound on the exponent still remains  $\frac{1}{2}$  for FPP on  $\mathbb{Z}^d$  ([Kes93, BKS03]). In contrast, in the hyperbolic setting one expects the variance to grow linearly in the word distance, and this was conjectured in [BZ12, Question 5], [BT17, Section 4]. We confirm this conjecture (Theorem 8.1).

It is well understood from the study of FPP in the Euclidean case that questions of fluctuations of the first passage times are intimately connected with the geometry of finite and semi-infinite geodesics. Indeed our proof of Theorem 8.1 requires understanding the geometry of geodesics as well as semi-infinite geodesic rays, and yield some results that are interesting in their own right.

- (3) **Direction of Geodesic Rays:** Almost surely FPP geodesic rays have a well-defined direction (Theorems 6.6 and 6.7). As in the case of word-geodesics, these are parametrized by points  $\xi \in \partial G$ . A similar result is known in the Euclidean setting only under the unproven assumption of uniform curvature of the limit shape [New95].
- (4) **Coalescence of Geodesic Rays:** We show that for each direction  $\xi \in \partial G$ ,  $o_1, o_2 \in G$ , the pair of FPP geodesic rays from  $o_1$  and  $o_2$  in the direction  $\xi$  almost surely coalesce, i.e., the set of edges in the symmetric difference of the two geodesic rays outside a sufficiently large ball centered at the identity element is empty (Theorem 7.9). This is of course not true in general for word geodesic rays in hyperbolic groups (see Section 5.3.1): geodesic rays in  $\Gamma$  converging to the same  $\xi \in \partial G$  eventually lie in a uniformly bounded neighborhood of each other. Coalescence in hyperbolic groups forces FPP geodesics to actually coincide beyond a point—a strictly stronger phenomenon. It follows from coalescence that for each  $\xi \in \partial G$ , almost surely there exists a unique geodesic ray from the identity in direction  $\xi$ . In the planar Euclidean setting, coalescence and uniqueness of geodesic rays is known either in almost every direction or under some additional unproven assumptions such as the differentiability of the boundary of the limit shape [LN96, DH17].

**1.1. Outline of the paper.** The rest of this paper is organized as follows. In Section 2, we make formal definitions and set up basic notations for FPP on  $\Gamma$  (Section 2.1), recall preliminaries on hyperbolic groups (Section 2.2) and collect some standard probabilistic tools (Section 2.3) that we shall need in the rest of the paper.

The next three sections aim at establishing the existence of velocity (Theorem 5.1). Section 3 first recalls Cannon’s theorem on the existence of automatic structures, and its connections with the Patterson-Sullivan measure [CF10, CM15]. Next, we use these results to establish the main technical lemma of this section (Lemma 3.18) that proves the existence of frequency of occurrence of geodesic words along a geodesic ray  $[1, \xi)$  from the identity in direction  $\xi \in \partial G$ . The main aim of Section 4 is to establish an approximation result, Theorem 4.1, for FPP geodesics. We look at large cylindrical neighborhoods  $N_B([1, \xi))$  and look at passage times  $T_B(x, y)$  between  $x, y \in [o, \xi)$  when restricted to  $N_B([o, \xi))$ . We describe precisely in Theorem 4.1 how  $T_B(x, y)$  approximates the passage time  $T(x, y)$  in  $\Gamma$  between  $x, y$ . Finally, in Section 5, we prove that velocity exists (Theorem 5.1). Counterexamples are also provided to show that this is a statement about a full measure subset of  $\partial G$  with respect to the Patterson-Sullivan measure; and cannot be upgraded to a statement about all of  $\partial G$ .

Section 6 establishes the fact that FPP geodesic rays almost surely have a well-defined direction in  $\partial G$  (Theorems 6.6 and 6.7). The main technical tool here is Proposition 6.3, which is an adaptation to our context of the main theorem of [BT17].

Section 7 proves coalescence of geodesics (Theorem 7.9). The main geometric tool for this is the construction of hyperplanes (Section 7.1). We combine this geometric tool with probabilistic estimates and concentration inequalities in Section 7.2 to establish coalescence.

Section 8 uses the technology developed in Section 7 to prove a conjecture of Benjamini, Tessler and Zeitouni [BT17, BZ12] which asserts that variance in passage time increases linearly with distance along word geodesic rays.

## 2. PRELIMINARIES

In this section, we formally define the first passage percolation model and collect together some basic notions from the classical theory of FPP. We also very briefly recall the preliminaries of the theory of hyperbolic groups and collect together some useful probabilistic estimates that we shall use throughout the paper.

**2.1. Preliminaries on first passage percolation.** Let  $\Gamma$  be a graph and let  $V, E$  denote the vertex and edge set of  $\Gamma$ . Consider  $\Gamma$  as a metric space with the graph distance metric  $d$  (where each edge in  $\Gamma$  is assigned unit length):  $d(x, y)$  is the minimum length of an edge-path joining  $x, y$ . A minimum length path connecting  $x, y \in \Gamma$  will simply be referred to as a geodesic and denoted  $[x, y]$ .

Fix a probability measure  $\rho$  on  $[0, \infty)$  and equip the Borel  $\sigma$ -algebra on the product space  $\Omega = [0, \infty)^E$  with the product measure  $\mathbb{P} = \rho^{\otimes E}$ . A typical element of  $(\Omega, \mathbb{P})$  will be denoted by  $\omega = \{\omega(e)\}_{e \in E}$  and the random variables  $X_e : \Omega \rightarrow [0, \infty)$  given by  $X_e(\omega) = \omega(e)$  will be independent and identically distributed with law  $\rho$ . Setting the edge length of the edge  $e$  equal to  $X_e$  defines a random metric on  $\Gamma$  (a priori, this is only a pseudo-metric; however, assuming henceforth that  $\rho$  does not put any mass on 0, it is indeed a metric), the **first passage percolation** (FPP) metric. More precisely we have the following definitions.

**Definition 2.1.** Let  $\gamma = \{e_1, \dots, e_k\}$  be an edge path. For  $\omega \in (\Omega, \mathbb{P})$ , the  $\omega$ -length of  $\gamma$  is defined to be

$$\ell_\omega(\gamma) := \sum_{e \in \gamma} \omega(e).$$

Define

$$d_\omega(x, y) := \inf_\gamma \ell_\omega(\gamma),$$

where  $\gamma$  ranges over edge paths connecting  $x, y \in V$ . The random variable  $T(x, y)$  defined on  $(\Omega, \mathbb{P})$  by

$$T(x, y)(\omega) = d_\omega(x, y)$$

will be called the **first passage time** between  $x$  and  $y$ .

We shall assume throughout that  $\rho$  is continuous, i.e., it does not have any atoms; in particular it does not put any mass on 0. Under such hypotheses it is easy to see that paths attaining the first passage time exist and are unique almost surely.

**Definition 2.2.** A path that realizes  $d_\omega(x, y)$  will be called an  $\omega$ -geodesic, denoted  $[x, y]_\omega$ . Observe that under our hypothesis on  $\rho$ , for  $\mathbb{P}$ -a.e.  $\omega \in \Omega$ , there is a unique  $\omega$ -geodesic between each pair of points in  $\Gamma$ . For fixed vertices  $x, y \in \Gamma$ , this ( $\mathbb{P}$ -a.e. well-defined) random path  $\Upsilon(x, y)$  (i.e.,  $\Upsilon(x, y)(\omega)$  denotes the  $\omega$ -geodesic between  $x$  and  $y$ ) will be called the **FPP-geodesic** between  $x$  and  $y$ .

The study of first passage percolation on a graph usually focuses on understanding asymptotic properties of  $T(x, y)$  and  $\Upsilon(x, y)$  for two points far away in the underlying metric of the graph.

**Assumptions on  $\rho$ :** Throughout we shall assume that the passage time distribution  $\rho$  satisfies the following conditions:

- i. The support of  $\rho$  is contained in  $[0, \infty)$ .
- ii. There are no atoms in  $\rho$ .
- iii.  $\rho$  has sub-Gaussian tails, i.e.,

$$(1) \quad \exists a > 0 \text{ such that } \int e^{ax^2} d\rho(x) < \infty.$$

Observe that our conditions are somewhat stricter than the ones usually assumed in the study of Euclidean FPP. Indeed, for the study of shape theorems or fluctuations, it is customary to only assume that the mass of the atom at 0 is smaller than the critical probability of Bernoulli percolation and some appropriate moment conditions. Existence of geodesics can also sometimes be ascertained under weaker hypotheses than above. The above conditions are not even optimal for our proofs, but in the interests of transparency we have chosen to go with the simplest set of assumptions which still covers a wide class of distributions. Some of the proofs become easier if one assumes a stronger condition that the support of the passage time distribution is bounded away from 0 and infinity, but our hypotheses already deals with the essential difficulties of working with passage times which are unbounded and can take values arbitrarily close to 0. One obvious area of improvement is (1). In fact, this hypothesis is only invoked in the proof of Theorem 4.1, at every other place, the proof only requires an exponential tail decay of  $\rho$ . Even Theorem 4.1 (and hence all results in this paper) can be proved if  $\rho$  is an exponential distribution (i.e.,  $\rho$  has density  $e^{-\lambda x}$  on  $\mathbb{R}_+$  for some  $\lambda > 0$ ), however in the interest of brevity and clarity of exposition, we shall refrain from trying to get optimal hypotheses in our results.

**2.2. Preliminaries on hyperbolic groups.** We collect here some of the basic notions and tools from hyperbolic metric spaces that we shall need in this paper and refer the reader to [Gro85, GdlH90, CDA90, BH99] for more details on Gromov-hyperbolicity. For us,  $\Gamma$  will denote the Cayley graph of a group  $G$ , typically hyperbolic, with respect to a finite symmetric generating set  $S$ . Thus, the vertex set  $V = V(\Gamma)$  consists of  $\{g | g \in G\}$  and the edge set  $E = E(\Gamma)$  consists of pairs  $\{(g, h) | g, h \in G; g^{-1}h \in S\}$ . Thus  $G$  acts on the left by isometries (graph-isomorphisms) on  $\Gamma$ . A geodesic in a geodesic metric space  $(X, d_X)$  joining  $x, y$  will be denoted as  $[x, y]$ . The  $c$ -neighborhood of a set  $A$  in a metric space  $(X, d)$  will be denoted as  $N_c(A)$ .

**Definition 2.3.** [Gro85] *A geodesic metric space  $(X, d)$  is said to be  $\delta$ -hyperbolic if for all  $x, y, z \in X$ ,  $[x, y] \subset N_\delta([x, z] \cup [y, z])$ . A geodesic metric space  $(X, d)$  is said to be hyperbolic if it is  $\delta$ -hyperbolic for some  $\delta \geq 0$ .*

*A finitely generated group  $G$  is said to be hyperbolic with respect to some finite symmetric generating set  $S$  if the Cayley graph  $\Gamma = \Gamma(G, S)$  is hyperbolic.*

**Definition 2.4.** [Gro85] A map  $f : (X, d_X) \rightarrow (Y, d_Y)$  between metric spaces is said to be a  $(K, \epsilon)$ -**quasi-isometric embedding** if for all  $x_1, x_2 \in X$ ,

$$\frac{1}{K}d_X(x_1, x_2) - \epsilon \leq d_Y(f(x_1), f(x_2)) \leq Kd_X(x_1, x_2) + \epsilon.$$

A  $(K, \epsilon)$ -**quasi-isometric embedding**  $f : (X, d_X) \rightarrow (Y, d_Y)$  is said to be a  $(K, \epsilon)$ -**quasi-isometry**, if further,  $Y \subset N_K(f(X))$ .

A  $(K, \epsilon)$ -**quasi-isometric embedding**  $f : I \rightarrow (Y, d_Y)$  is said to be a  $(K, \epsilon)$ -**quasi-geodesic**, if  $I$  is an interval (finite, semi-infinite or bi-infinite) in  $\mathbb{R}$ .

A subset  $A$  of a geodesic metric space  $(X, d_X)$  is said to be  $\kappa$ -**quasiconvex**, if for all  $x_1, x_2 \in A$ , and any geodesic  $[x_1, x_2] \subset X$ ,  $[x_1, x_2] \subset N_\kappa(A)$ .

It was shown by Gromov [Gro85] that if  $G$  is hyperbolic with respect to some finite symmetric generating set  $S$ , it is hyperbolic with respect to any other finite symmetric generating set  $S'$ . Thus, hyperbolicity is a property of finitely generated groups, not their generating sets. This follows from the following theorem that says that hyperbolicity is invariant under quasi-isometry:

**Theorem 2.5** (Gromov). [Gro85, GdlH90][BH99, p. 401] Given  $\delta, \epsilon \geq 0$  and  $K \geq 1$ , there exists  $\delta' \geq 0$  such that the following holds.

Let  $(X, d)$  be  $\delta$ -hyperbolic and  $f : (X, d) \rightarrow (Y, d')$  be a  $(K, \epsilon)$ -quasi-isometry. Then  $(Y, d')$  is  $\delta'$ -hyperbolic.

Theorem 2.5 allows us the freedom to choose any Cayley graph of a hyperbolic group  $G$ . The qualitative results we prove in this paper will thus be independent of the generating set  $S$ .

**Definition 2.6.** [Gro85][BH99, p. 410] For any  $x, y, o$  in a metric space  $(X, d)$ , the Gromov inner product of  $x, y$  is given by

$$\langle x, y \rangle_o = \frac{1}{2}(d(x, o) + d(y, o) - d(x, y)).$$

The Gromov inner product above can be used to define the **Gromov boundary**  $\partial X$  of a hyperbolic  $(X, d)$  as follows [Gro85, GdlH90, BH99]. Fix a base-point  $o \in X$ . We consider sequences  $\{x_n\}$  in  $X$  satisfying the condition that  $\langle x_n, x_m \rangle_o \rightarrow \infty$  as  $m, n \rightarrow \infty$ . Two such sequences  $\{x_n\}$  and  $\{x'_n\}$  are defined to be equivalent if  $\langle x_n, x'_n \rangle_o \rightarrow \infty$  as  $n \rightarrow \infty$ .

**Definition 2.7.** [BH99, p. 431] The boundary  $\partial X$  of  $X$  is defined (as a set) to be the set of equivalence classes of sequences  $\{x_n\}$  as above. We write  $x_n \rightarrow \xi$ , if  $\xi \in \partial X$  is the equivalence class of  $\{x_n\}$ .

The Gromov inner product extends to  $\partial X$  as follows [BH99, p. 401]: Let  $\xi, \xi' \in \partial X$ . Then

$$\langle \xi, \xi' \rangle_o := \sup \liminf_{m, n \rightarrow \infty} \langle x_n, x'_m \rangle_o,$$

where  $\{x_n\}$  (resp.  $\{x'_m\}$ ) range over sequences in the equivalence class defining  $\xi$  (resp.  $\xi'$ ). The Gromov inner product  $\langle \xi, \xi' \rangle_o$  can be used to define a metric on  $\partial X$ .

**Definition 2.8.** A metric  $d_v$  on  $\partial X$  is said to be a **visual metric with parameter**  $a > 1$  with respect to the base-point  $o$  if there exist  $k_1, k_2 > 0$  such that

$$k_1 a^{-\langle \xi, \xi' \rangle_o} \leq d_v(\xi_1, \xi_2) \leq k_2 a^{-\langle \xi, \xi' \rangle_o}.$$

**Proposition 2.9.** [BH99, p. 435] *Given  $\delta \geq 0$ , there exists  $a > 1$  such that if  $(X, d)$  is  $\delta$ -hyperbolic, then for any base-point  $o$ , a visual metric  $d_v$  with parameter  $a > 1$  exists on  $\partial X$  with respect to the base-point  $o$ . Further, for any  $o' \in X$  a visual metric with parameter  $a > 1$  and with respect to the base-point  $o'$  is equivalent (as a metric) to  $d_v$ .*

*There exists a natural topology on  $\hat{X} = X \cup \partial X$  such that*

- (1)  $X$  is open and dense in  $\hat{X}$ ,
- (2)  $\partial X$  and  $\hat{X}$  are compact if  $X$  is proper,
- (3) the subspace topology on  $\partial X$  agrees with that given by  $d_v$ .

We call  $\hat{X}$  the **Gromov compactification** of  $X$ . For  $\xi \in \partial X$  and  $o \in X$ , a geodesic ray from  $o$  and converging to  $\xi \in \partial X$  will be denoted by  $[o, \xi)$ . For  $\xi_1 \neq \xi_2 \in \partial X$ , a bi-infinite geodesic  $f : \mathbb{R} \rightarrow X$  converging to  $\xi_1, \xi_2$  as  $s \in \mathbb{R}$  tends to  $\pm\infty$  will be denoted by  $(\xi_1, \xi_2)$ .

**Lemma 2.10** (Morse Lemma). [BH99, p. 401] *Given  $\delta, \epsilon \geq 0$  and  $K \geq 1$ , there exists  $\kappa \geq 0$  such that the following holds:*

*Let  $(X, d)$  be a  $\delta$ -hyperbolic space. Let  $f : I \rightarrow X$  be a  $(K, \epsilon)$ -quasi-geodesic with  $I = [a, b]$  finite. Then  $f(I) \in N_\kappa([f(a), f(b)])$  for any geodesic  $[f(a), f(b)]$  in  $X$  joining  $f(a), f(b)$ . In particular, any two  $(K, \epsilon)$ -quasi-geodesics joining  $x, y \in X$  lie in a  $\kappa$ -neighborhood of each other.*

*When  $I = [0, \infty)$  is semi-infinite,  $[f(a), f(b)]$  is replaced by  $[f(a), \xi)$  where  $\xi$  is a unique point in  $\partial X$ . Finally, when  $I = \mathbb{R}$ ,  $[f(a), f(b)]$  is replaced by  $(\xi_1, \xi_2)$  where  $\xi_1, \xi_2$  are unique points in  $\partial X$ . Further, any two  $(K, \epsilon)$ -quasi-geodesic rays joining  $x \in X$  (or  $\xi' \in \partial X$ ) to  $\xi \in \partial X$  lie in a  $\kappa$ -neighborhood of each other.*

The Gromov boundary (Definition 2.7) can be defined also in terms of asymptote-classes of geodesic rays: Define (semi-infinite) geodesic rays  $\gamma, \gamma' : [0, \infty) \rightarrow X$  to be asymptotic if there exists  $C_0 \geq 0$  such that for all  $t \geq 0$ ,  $(\gamma(t), \gamma'(t)) \leq C_0$ . The next Lemma says that geodesic rays  $\gamma, \gamma'$  are asymptotic if and only if they converge to the same  $\xi \in \partial X$ :

**Lemma 2.11.** [BH99, p. 427] *Let  $X$  be  $\delta$ -hyperbolic. Let  $\gamma, \gamma' : [0, \infty) \rightarrow X$  be asymptotic geodesic rays. Then*

- (1) *There exist  $m, m' \in [0, \infty)$  such that for all  $t \in [0, \infty)$ ,  $d(\gamma(t+m), \gamma'(t+m')) \leq 4\delta$ .*
- (2) *There exists  $\xi \in \partial X$  such that for any pair of sequences  $\{t_n\}, \{s_n\}$  in  $[0, \infty)$  diverging to infinity, the sequences  $\{\gamma(t_n)\}, \{\gamma'(s_n)\}$  lie in the equivalence class of  $\xi$ .*

The Patterson-Sullivan measure (Definition 3.6) will be crucially used in this paper. For now, it suffices to say that it is a Borel measure  $\nu$  supported on  $\partial G$  (with respect to the topology defined by the visual metric), and is quasi-invariant under the natural action of  $G$  on  $\partial G$ .

**2.3. Probabilistic Tools.** Here we record the basic probabilistic tools of concentration bounds and the FKG inequality that we shall use throughout. Note that these are mostly standard, but we shall provide appropriate references (or proofs) to make the exposition self-contained.

**Concentration Inequalities:** We shall have occasion to use a number of concentration inequalities for sums of i.i.d. variables. The first one we need is the Chernoff Inequality (see e.g. [Ver18, Theorem 2.3.1]).

**Theorem 2.12** (Chernoff Inequality). *Let  $X_i$  be independent Bernoulli variables with  $\theta := \sum \mathbb{E}X_i$ . Then for  $\alpha > 0$*

$$\mathbb{P}\left(\sum X_i > (1 + \alpha)\theta\right) \leq e^{\theta(\alpha - (1+\alpha)\log(1+\alpha))}.$$

We next need concentration results for sums of i.i.d. random variables with sub-exponential tails (see e.g. [Ver18, Theorem 2.8.1]).

**Theorem 2.13** (Concentration for sums of i.i.d. sub-exponential random variables). *Let  $X_i$  be i.i.d. non-negative random variables with distribution  $\nu$  such that for some  $a > 0$  we have  $\int_0^\infty e^{ax} d\nu(x) < \infty$  and  $\mathbb{E}[X_i] = \mu$ . Then for each  $\delta > 0$ , we have*

$$\mathbb{P}\left(\sum_{i=1}^n X_i \geq (1 + \delta)n\mu\right) \leq e^{-cn}$$

for some  $c = c(\delta, \nu) > 0$ . Further,  $c \geq c_1\delta$  for some  $c_1 > 0$  if  $\delta$  is sufficiently large.

As any sub-exponential random variable is also sub-Gaussian, clearly the above result will also hold for  $X_i \sim \rho$  where  $\rho$  satisfies our hypotheses on the passage time distribution.

The next result we shall need shows that for a sum of i.i.d. sub-Gaussian random variables the total contribution coming from terms that are sufficiently large is only a small fraction with high probability. This is a less standard result, even though it follows from essentially the same arguments as classical concentration inequalities. We provide a short proof in the appendix for completeness.

**Theorem 2.14.** *Let  $X_i$  be i.i.d. sub-Gaussian non-negative random variables (i.e.,  $\mathbb{P}(X_i \geq t) \leq C_1 e^{-c_1 t^2}$  for some  $C_1, c_1 > 0$  and  $t > 0$ ). Let  $\epsilon > 0$  be fixed. Then for  $M > 0$  sufficiently large, there exists  $c(M) > 0$  such that we have for all  $n \geq n_0(c_1, C_1)$ .*

$$\mathbb{P}\left(\sum_{i=1}^n (X_i - M)_+ \geq \epsilon n\right) \leq e^{-cn}.$$

Further as  $M \rightarrow \infty$  the constant  $c = c(M)$  above also goes to  $\infty$ .

We postpone the proof to Appendix A.

**FKG Inequality:** Finally we recall the standard FKG correlation inequality, which is widely used in the study of FPP and related percolation models. We call a Borel subset of  $\Omega$  increasing (resp. decreasing) if  $\omega \in A$  implies  $\omega' \in A$  if  $\omega'(e) \geq \omega(e)$  for all  $e \in E$  (resp.  $\omega'(e) \leq \omega(e)$  for all  $e \in E$ ). The following is a variant of the standard FKG inequality on product spaces (see e.g. [Kes03, Lemma 2.1])

**Theorem 2.15.** *For any two increasing (or decreasing) Borel subsets  $A$  and  $B$  of  $\Omega$ , we have*

$$\mathbb{P}(A \cap B) \geq \mathbb{P}(A) \cdot \mathbb{P}(B).$$

In particular, Theorem 2.15 shows that conditioning on an increasing (resp. decreasing) event makes another increasing (resp. decreasing) event more likely.



**Basic Set-up:** As mentioned above,  $G$  will denote a Gromov-hyperbolic group. A Cayley graph of  $G$  with respect to a fixed finite symmetric generating set  $S$  will be denoted by  $\Gamma = \Gamma(G, S)$ . The word-metric on  $\Gamma$  with respect to  $S$  will be denoted by  $d$ . Thus  $d(x, y)$  equals the minimum number of edges in an edge-path joining  $x, y$ . We shall often use  $|y|$  as a shorthand for  $d(1, y)$ . We shall consider FPP on  $\Gamma$  with edge weights distributed according to a measure  $\rho$  satisfying the hypothesis in Section 2.1.

We shall now prepare the ground for one of our main results. Let  $1$  denote the vertex of  $\Gamma$  corresponding to the identity element of  $G$ . Our objective is to study the asymptotics of the first passage time  $T(x, y)$  for  $x, y \in \Gamma$  as  $d(x, y) \rightarrow \infty$ . By group invariance it suffices to set  $x = 1$ , and study  $T(1, y)$  for large  $|y|$ . In analogy with the Euclidean case, it is natural to study  $T(1, x_n)$  as  $x_n$  moves along some fixed direction parametrized by an element in  $\partial G$ . Let  $\xi \in \partial G$  be fixed and let  $[1, \xi)$  denote a fixed word geodesic ray from  $1$  in the direction  $\xi$ , i.e.,  $\{1 = x_0, x_1, \dots, x_n, \dots\}$  such that  $x_n \rightarrow \xi$ , and each finite subpath of  $[1, \xi)$  is a word geodesic between the corresponding endpoints. Clearly, this would imply  $d(1, x_n) = n$ . It is not too difficult to show that  $\mathbb{E}T(1, x_n)$  grows linearly in  $n$ , and we shall show that a limiting velocity  $v(\xi) := \lim_{n \rightarrow \infty} \frac{\mathbb{E}T(1, x_n)}{n}$  exists (Theorem 5.1) for almost every direction  $\xi \in \partial G$ . The next three sections are devoted to the proof of this theorem.

### 3. AUTOMATIC STRUCTURE, PATTERSON-SULLIVAN MEASURES AND FREQUENCY

This section is devoted to recalling and developing some of the technical tools from hyperbolic geometry that will go into the proof of Theorem 5.1. We shall introduce an appropriate measure  $\nu$  on  $\partial G$  and show that for  $\nu$ -almost every  $\xi \in \partial G$ , there exists a limiting frequency of occurrence of fixed length geodesic words along  $[1, \xi)$  (Lemma 3.18), provided the length lies in  $d\mathbb{N}$  for a suitable  $d$ .

In Sections 3.1 and 3.2, we recall some facts about symbolic dynamics and hyperbolic groups [Gro85, CP93]. We refer to the excellent set of notes [Cal13] where the necessary basics are summarized. Most of the relatively recent material here is due to Calegari, Fujiwara and Maher [CF10, CM15]. In Section 3.3, we introduce the notion of ordered frequency—a modification of the counting function due to Rhemtulla [Rhe68], rediscovered by Brooks [Bro81]. We use a vector-valued Markov chain argument (Proposition 3.14) along with the tools recalled from [CF10, CM15] to prove that ordered frequencies exist (Lemma 3.18).

**3.1. Automatic structures on hyperbolic groups.** Our starting point is a theorem of Cannon [Can84, Can91], saying that a hyperbolic group admits an automatic structure. We say briefly what this means, referring the reader to [Can84, Can91, Cal13] for details. Since the generating set  $S$  of  $G$  is chosen to be symmetric,  $S$  generates  $G$  as a semigroup. Then [Can84] there exists a finite state automaton  $\mathcal{G}$  (equivalently, a finite digraph with directed edges labeled by  $S$  and a distinguished initial state  $1$ ) such that

- (1) any word obtained by starting at  $1$  and reading letters successively on  $\mathcal{G}$  gives a geodesic in the Cayley graph  $\Gamma = \Gamma(G, S)$  (the empty word being sent to the identity element in  $G$ ). The set of all such words is denoted by  $\mathbb{L}$  and is called the formal language accepted by  $\mathcal{G}$ .

- (2) Let  $e : \mathbb{L} \rightarrow G$  denote the evaluation map, sending  $w \in \mathbb{L}$  to the element  $g \in G$  it represents. For all  $g \in G$ , there exists a unique  $w \in \mathbb{L}$  such that  $e(w) = g$ .

A language  $\mathbb{L}$  generated as above by reading words on a finite state automaton  $\mathcal{G}$  (without reference to a group) is called a **prefix-closed regular language**. We also refer to  $\mathbb{L}$  as the language accepted by  $\mathcal{G}$ . If, as above,  $\mathbb{L}$  encodes geodesics in  $\Gamma$ , it is called a **geodesic language**. The collection of geodesics in  $\Gamma$  obtained in the process is called a **geodesic combing**, or simply, a combing of  $\Gamma$ . The finite directed and labeled graph  $\mathcal{G}$  is said to **parametrize** the combing. Cannon's theorem thus proves the following:

**Theorem 3.1.** [Can84] *For  $G$  hyperbolic and any symmetric generating set  $S$ , there exists a finite state automaton  $\mathcal{G}$  that parametrizes a combing of  $\Gamma = \Gamma(G, S)$  corresponding to the prefix-closed regular geodesic language  $\mathbb{L}$  accepted by  $\mathcal{G}$ .*

Directed paths in  $\mathcal{G}$  starting at the initial vertex 1 are in one-to-one correspondence with words in  $\mathbb{L}$ . We use the suggestive notation 1 for the initial state as the evaluation map  $e$  is assumed to send  $1 \in \mathcal{G}$  to the identity element  $1 \in G$ . Let  $\mathcal{P}^0$  denote this collection of directed paths and let  $\mathcal{P}_n^0$  denote the subset of  $\mathcal{P}^0$  consisting of directed paths of length  $n$  (starting at 1 by definition). The evaluation map  $e : \mathbb{L} \rightarrow G$  then naturally gives an evaluation map  $e : \mathcal{P}^0 \rightarrow G$  by sending a word in  $\mathbb{L}$  or equivalently a labeled path in  $\mathcal{P}^0$  to the element in  $G$  it represents (we use the same letter  $e$  for both). The set of *all* directed paths (without restriction on the base-point) in  $\mathcal{G}$  will be denoted as  $\mathcal{P}$  and the set of all directed paths of length  $n$  in  $\mathcal{P}$  will be denoted by  $\mathcal{P}_n$ . An important ingredient in the proof of Theorem 3.1 is the following:

**Definition 3.2.** *For  $g \in G$ , let  $y \in \mathcal{P}^0$  be the unique geodesic word such that  $e(y) = g$ . The cone  $\text{cone}(g)$  consists of the image under  $e$  of all paths extending  $y$ .*

Note that  $\text{cone}(g)$  is the image under  $e$  of the cylinder set in  $\mathcal{P}^0$  determined by  $y$ . The underlying directed graph of  $\mathcal{G}$  is also called a **topological Markov chain** and its vertices are called **states**. Let  $\mathcal{V}(\mathcal{G})$  denote the set of states. We define an equivalence relation on  $\mathcal{V}(\mathcal{G})$  by declaring  $v_1, v_2$  to be equivalent if there are directed paths from  $v_1$  to  $v_2$  and vice versa. Each equivalence class is called a **component** and the resulting quotient directed graph is denoted  $C(\mathcal{G})$ . Then  $C(\mathcal{G})$  has no directed loops.

Let  $V$  denote the complex vector space of complex functions on  $\mathcal{V}(\mathcal{G})$ . Let  $M$  denote the transition matrix of  $\mathcal{G}$ : thus  $M_{kl}$  equals one if there is a directed edge from the vertex labeled  $k$  to the vertex labeled  $l$  and is zero otherwise. Let  $\lambda$  denote its maximal eigenvalue. Similarly, for each component  $C$ , let  $M_C$  denote the transition matrix of  $C$  and let  $\lambda_C$  denote its maximal eigenvalue. Note that  $\lambda_C$  is real as the transition matrix is non-negative (by Perron-Frobenius). For all  $C$ ,  $\lambda_C \leq \lambda$ . A component  $C$  is said to be **maximal** if  $\lambda_C = \lambda$ . Recall that  $B_n(1)$  denotes the  $n$ -ball about  $1 \in G$  (recall that 1 denotes the identity element of  $G$ ). Let

$$G_n := \{g \in G \mid d(1, g) = n\}$$

denote its 'boundary', the  $n$ -sphere.

**Theorem 3.3.** [Coo93] [CF10, Lemma 4.15] [CM15, Theorem 3.7] *Let  $G, \mathcal{G}, C(\mathcal{G}), M, \lambda$  be as above. Then each directed path in  $C(\mathcal{G})$  is contained in at most one maximal*

component. There exist  $K$  such that

$$\frac{1}{K}\lambda^n \leq |G_n| \leq K\lambda^n$$

Theorem 3.3 shows in particular, that the algebraic and geometric multiplicities of the maximal eigenvalue  $\lambda$  are equal. Calegari and Fujiwara [CF10, Lemmas 4.5, 4.6] show that the following limits exist for all  $v \in V$ :

$$(2) \quad \begin{aligned} r(v) &:= \lim_{n \rightarrow \infty} n^{-1} \sum_{i=0}^n \lambda^{-i} M^i v. \\ l(v) &:= \lim_{n \rightarrow \infty} n^{-1} \sum_{i=0}^n \lambda^{-i} (M^T)^i v. \end{aligned}$$

Further  $r(v)$  (resp.  $l(v)$ ) equals the projection of  $v$  onto the right (resp. left)  $\lambda$ -eigenspace of  $M$ .

**Definition 3.4.** Let  $v_i$  denote the **initial vector**, taking the value 1 on the initial state 1 and zero elsewhere. Let  $v_u$  denote the **uniform vector**, taking the constant value 1 on all  $x \in \mathcal{V}(\mathcal{G})$ . Let  $N$  be the matrix given by

$$(1) \quad N_{pq} = \frac{M_{pq} r(v_u)_q}{\lambda r(v_u)_p} \text{ if } r(v_u)_p \neq 0, \\ (2) \quad N_{pp} = 1 \text{ and } N_{pq} = 0 \text{ for } p \neq q, \text{ if } r(v_u)_p = 0.$$

Define a probability measure  $\mu$  on  $\mathcal{V}(\mathcal{G})$  by  $\mu_p = \frac{r(v_u)_p l(v_i)_p}{\sum_p r(v_u)_p l(v_i)_p}$ .

Lemma 4.9 of [CF10] shows that  $N$  is a stochastic matrix preserving  $\mu$ . The measure  $\mu$  and the matrix  $N$  define measures on  $\mathcal{P}_n$  by the usual strategy of defining measures on cylinders in path spaces. Let  $\sigma = v_0 v_1 \cdots v_n \in \mathcal{P}_n$ . Then define

$$\mu(\sigma) := \mu(v_0) N_{v_0 v_1} N_{v_1 v_2} \cdots N_{v_{n-1} v_n}.$$

**3.2. Patterson-Sullivan measures.** We shall opt for two different normalizations for the Patterson-Sullivan measures (Definition 3.6 below) following [Coo93] and [CF10].

**Theorem 3.5.** [Coo93, Cal13, CF10] Define two sequences of finite measures on  $\Gamma \cup \partial G$  by

$$\nu_n = \frac{\sum_{|g| \leq n} \lambda^{-|g|} \delta_g}{\sum_{|g| \leq n} \lambda^{-|g|}},$$

and

$$\widehat{\nu}_n = \frac{1}{n} \sum_{|g| \leq n} \lambda^{-|g|} \delta_g.$$

Let  $\nu$  and  $\widehat{\nu}$  be respectively the weak limits of  $\nu_n$  and  $\widehat{\nu}_n$  (up to subsequential limits). Then,  $\nu$  and  $\widehat{\nu}$  are supported on  $\partial G$ . Further, any two subsequential limits are absolutely continuous with respect to each other with uniformly bounded Radon-Nikodym derivatives.

**Definition 3.6.** The measures  $\nu$  and  $\widehat{\nu}$  (and sometimes their normalized versions) will be called the **Patterson-Sullivan measures** on  $\partial G$ .

It turns out [CF10] that  $\widehat{\nu}_n, \widehat{\nu}$  are finite (but not necessarily probability) measures in the same measure classes as  $\nu_n$  and  $\nu$  [Coo93] respectively. Further, the Radon-Nikodym derivatives of  $\widehat{\nu}$  with respect to  $\nu$  on  $\partial G$  are uniformly bounded away from zero and infinity.

**Remark 3.7.** We shall refer to  $\nu_n$  and  $\widehat{\nu}_n$  as **approximants** of  $\nu$  and  $\widehat{\nu}$  respectively (their existence and basic properties are proven in [Coo93]). Both  $\nu$  and  $\widehat{\nu}$  will be used in what follows as some properties are easier to state for one than the other. However, owing to the fact that they are absolutely continuous with respect to each other with uniformly bounded Radon-Nikodym derivatives, statements about one hold for the other up to uniformly bounded constants.

Patterson-Sullivan measures of cones  $\text{cone}(g)$  are given by

$$\nu(\text{cone}(g)) = \lim_{n \rightarrow \infty} \nu_n(\text{cone}(g)),$$

and

$$\widehat{\nu}(\text{cone}(g)) = \lim_{n \rightarrow \infty} \widehat{\nu}_n(\text{cone}(g)).$$

The relation between  $\mu$ ,  $N$  (Definition 3.4) and  $\nu$  is given as follows:

**Lemma 3.8.** [CM15, Sections 3.3, 3.4] *Given  $g \in G$ , let  $n \in \mathbb{N}$  and  $\sigma = 1v_1v_2 \cdots v_n \in \mathcal{P}_n^0$  be the unique directed path such that  $e(\sigma) = g$ . Then*

$$\nu(\text{cone}(g)) = N_{1v_1} N_{v_1v_2} \cdots N_{v_{n-1}v_n}.$$

**Definition 3.9.** *For  $\sigma = 1v_1v_2 \cdots v_n \in \mathcal{P}_n^0$  (as in Lemma 3.8 above), we define*

$$\nu(\sigma) = \nu(\text{cone}(g)) = N_{1v_1} N_{v_1v_2} \cdots N_{v_{n-1}v_n}.$$

*The Patterson-Sullivan measure  $\nu$  on  $\text{cone}(g)$  also gives a measure on  $G_n$  for all  $n$ , simply by defining  $\nu(g) := \nu(\text{cone}(g))$ .*

Definition 3.9 thus gives us a well-defined way of lifting the Patterson-Sullivan measure  $\nu$  to the path space  $\mathcal{P}^0$ , by identifying the cylinder set corresponding to  $g$  with the Borel subset of  $\partial G$  given by the boundary of  $\text{cone}(g)$  (see [Coo93] for details). Let  $\mathcal{S}$  denote the left shift taking a sequence of vertices to the sequence omitting the first vertex. In particular,  $\mathcal{S}(\mathcal{P}^0) = \mathcal{P}$ . Then [CF10, Lemma 4.19], there exists a constant  $c > 0$  such that

$$(3) \quad c\mu = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathcal{S}_*^i \widehat{\nu},$$

where  $\mathcal{S}_*$  denotes the push-forward. We caution the reader that in [CF10], the constant  $c$  is not explicit.

The following Proposition shows that the Patterson-Sullivan measure  $\nu$  and the uniform measure on  $G_n$  are equivalent to each other with uniform constants.

**Proposition 3.10.** [CF10, Section 4][CM15, Proposition 3.11] *There exist  $K, C \geq 0$  such that the following holds. For any  $n \in \mathbb{N}$  and  $g \in G_n$ , let  $B(g, C) = (N_C(g) \cap G_n)$  and let  $B_0(g, C) = e^{-1}(B(g, C)) \subset \mathcal{P}_n^0$  denote the pre-image of  $B(g, C)$  under the evaluation map. Then*

$$\frac{1}{K} \nu(B_0(g, C)) / \nu(\mathcal{P}_n^0) \leq |B(g, C)| / |G_n| \leq K \nu(B_0(g, C)) / \nu(\mathcal{P}_n^0).$$

The following Proposition gives a quantitative estimate on the behavior of typical geodesics (recall that Definition 3.9 gives a well-defined way to lift  $\nu$  to  $\mathcal{P}_\infty^0$ ).

**Proposition 3.11.** [CF10, Proposition 4.10][CM15, Proposition 3.12] *There exist  $c_1, c > 0$  such that the following holds. Let  $\sigma$  be a path in  $(\mathcal{P}_n^0, \nu)$ . Then, apart from a prefix of size at most  $c_1 \log(n)$ ,  $\sigma$  is entirely contained in a single maximal component of  $\mathcal{G}$  with probability  $1 - O(e^{-cn})$ .*

Also, if  $C$  is a maximal component of  $\mathcal{G}$ , then, as  $n \rightarrow \infty$ , a path  $\sigma$  enters and stays in  $C$  with probability  $\mu(C)$ , where  $\mu(C)$  is computed from (3).

Let  $\mathcal{P}_\infty^0$  (resp.  $\mathcal{P}_\infty$ ) denote the collection of infinite paths in  $\mathcal{P}^0$  (resp.  $\mathcal{P}$ ).

**Lemma 3.12.** [Cal13, Lemma 3.5.1] *The evaluation map  $e : \mathcal{P}_n^0 \rightarrow G$  extends continuously to  $\bar{e} : \mathcal{P}_\infty^0 \rightarrow \partial G$  such that the cardinality  $|\bar{e}^{-1}(\xi)|$  is uniformly bounded independent of  $\xi \in \partial G$ .*

Let  $\mathcal{P}(C) \subset \mathcal{P}$  denote the collection of paths contained in a maximal component  $C$  and let  $\mathcal{P}_\infty(C)$  denote the collection of infinite paths contained in  $C$ . From (3), Remark 3.7 and Proposition 3.11, we have the following immediate corollary:

**Corollary 3.13.** *For  $\nu$ -a.e.  $\sigma \in \mathcal{P}_\infty^0$ , there is a unique maximal component  $C$  so that  $\mathcal{S}^i(\sigma) \in \mathcal{P}_\infty(C)$  for all  $i$  sufficiently large.*

We shall define a  $\nu$ -full subset  $\mathcal{P}'_\infty \subset \mathcal{P}_\infty^0$  of paths starting at the identity element such that certain mixing conditions are satisfied along all trajectories  $\sigma \in \mathcal{P}'_\infty$  (recall that the evaluation map identifies  $\sigma \in \mathcal{P}_\infty^0$  with semi-infinite geodesic words in  $\Gamma$ ). This leads us to the notion of ordered frequencies.

**3.3. Frequency.** In this subsection we shall introduce the notion of ordered frequencies. This is a refinement of the counting function of Rhemtulla [Rhe68], rediscovered by Brooks [Bro81]. We shall use the technology recalled in Sections 3.1 and 3.2 to prove below the main technical lemma of this section: ordered frequencies exist along almost every word geodesic ray (Lemma 3.18). We shall need some basic facts from the theory of ergodic Markov chains.

**3.3.1. Markov Chain Trajectories.** We refer to [LP17] for details on mixing in Markov chains. Let  $P$  denote the transition matrix of an irreducible (but not necessarily aperiodic) Markov chain on a finite state space  $\Sigma$ . Note that we are not assuming reversibility of the Markov chain. Let  $d$  denote the period of  $P$ . Let  $k, n \in \mathbb{N}$  with  $k$  a multiple of  $d$ ,  $n \gg k$  be fixed and let us also fix  $\mathbf{x} = (x_1, x_2, \dots, x_k) \in \Sigma^k$ . Let  $\{X_n^x\}_{n \geq 1}$  denote a realization of the chain starting with  $x \in \Sigma$ . Let  $N_n(\mathbf{x}, x)$  denote the number of positive integers  $i \leq \frac{n}{k} - 1$  such that  $(X_{ik+1}^x, \dots, X_{(i+1)k}^x) = \mathbf{x}$ . We have the following result.

**Proposition 3.14.** *For each  $x \in \Sigma$ , the following holds almost surely. For all  $k \in d\mathbb{N}$ ,  $\mathbf{x} \in \Sigma^k$ , and  $\epsilon > 0$ , there exists  $f(\mathbf{x}, x) \geq 0$  (non-random) and  $N_0$  (random but finite, depending on  $k, \mathbf{x}, \epsilon$ ) such that we have for all  $n \geq N_0$*

$$\frac{N_n(\mathbf{x}, x)}{n} \in (f(\mathbf{x}, x) - \epsilon, f(\mathbf{x}, x) + \epsilon).$$

*In particular, this implies*

$$\frac{N_n(\mathbf{x}, x)}{n} \rightarrow f(\mathbf{x}, x)$$

*almost surely.*

The proof is standard and uses the fact that associated vector-valued Markov chain, restricted to each of its recurrent component (which is determined by the initial state) is aperiodic and hence mixing. We provide the argument in Appendix A for completeness.

**3.3.2. Ordered Frequency.** We now turn to a refinement of the counting function of [Rhe68, Bro81]. Let  $d$  denote the L.C.M. of the periods of (the topological Markov chain underlying) the finite state automaton  $\mathcal{G}$  parametrizing a geodesic combing  $\mathbb{L}$  of  $\Gamma$ . (In the proofs below, it will suffice to take  $d$  to be the L.C.M. of the periods of the maximal components.) By Corollary 3.13, for  $\nu$ -a.e.  $\sigma$ , there exists a unique maximal component  $C$  such that  $\sigma$  eventually lies in  $\mathcal{P}_\infty(C)$ . Let  $\mathcal{P}_\infty^{\max} \subset \mathcal{P}_\infty^0$  denote the collection of such semi-infinite geodesic words.

Let  $n$  be a multiple of  $d$ . For a geodesic word  $w = g_1 g_2 \cdots g_n$  of length  $n$ , and  $\sigma \in \mathcal{P}_\infty^{\max}$ , we shall define a notion of frequency of occurrence of  $w$  in  $\sigma$ . Let  $\{y_i\}$ ,  $i = 0, 1, \dots$  be the sequence of (ordered equispaced) points on  $\sigma$ , such that  $e(y_0) = 1$  and  $d(e(y_i), 1) = in$ ; so that  $d(e(y_i), e(y_{i+1})) = n$ .

An **ordered occurrence** of  $w$  in  $\sigma$  is a pair  $y_i, y_{i+1}$  such that  $y_i^{-1} y_{i+1}$  equals  $w$  as an ordered word. Let  $n_w([y_0, y_i])$  be the number of distinct ordered occurrences of  $w$  in  $[y_0, y_i] \subset \sigma$ .

**Definition 3.15.** *Let  $d, \sigma, y_j$  be as above. The **ordered frequency** of occurrence of  $w$  along  $\sigma$  is defined to be*

$$f_w(\sigma) := \lim_{i \rightarrow \infty} \frac{n_w([y_0, y_i])}{i},$$

*provided the limit exists.*

**Remark 3.16.** *Let  $C$  be the unique maximal component  $\sigma$  eventually lies in. Let  $j \in \mathbb{N}$  be the least integer such that  $y_j$  onward,  $\sigma$  lies in  $C$ . Then, observe that the existence of  $f_w(\sigma)$  is equivalent to the existence of  $\lim_{i \rightarrow \infty} \frac{n_w([y_j, y_i])}{i-j}$ . This observation will be useful in the proof of Lemma 3.18 below.*

**Remark 3.17.** We have made Definition 3.15 only for  $w$  with length a multiple of  $d$ , though the definition per se works for arbitrary  $w$ . This is to address the fact that maximal components  $C$ , while irreducible, need not be mixing. The existence of ordered frequencies (Lemma 3.18 below) will be important for a coarse-graining argument in Section 5.2 to establish the existence of velocity.

Irreducibility of topological Markov chains corresponding to maximal components  $C$  give us the following.

**Lemma 3.18** (Ordered frequencies exist). *For  $d$  and  $\mathcal{P}_\infty^{\max}$  as above and for almost all  $\sigma = \{x_0, x_1, \dots\} \in \mathcal{P}_\infty^{\max}$  the following holds: For each geodesic word  $w$  of length a multiple of  $d$  and each  $\epsilon > 0$  there exists  $f_w(\sigma)$  and  $N_0 = N_0(\sigma, \epsilon, w)$  such that for all  $i \geq N_0$  we have*

$$\frac{n_w([y_0, y_i])}{i} \in (f_w(\sigma) - \epsilon, f_w(\sigma) + \epsilon);$$

*In particular, the ordered frequency  $f_w(\sigma)$  exists for all such  $w$ .*

*Proof.* By Theorem 3.3, the Markov chain  $N$  (Definition 3.4) restricted to  $C$  is irreducible as  $C$  is maximal. Note also that maximal components corresponding to  $N$  are maximal for the topological Markov chain  $M$ . Further, Lemma 3.8 shows that the law of  $\sigma \in (\mathcal{P}_\infty^0, \nu)$  is the same as the law of trajectories given by the Markov chain  $N$ .

Let  $|w| = k$ , where  $k$  is a multiple of  $d$ . Hence, the associated vector-valued Markov chain of  $k$ -tuples is mixing. Let  $w = g_1 \cdots g_k \in \mathcal{P}_k$ , where  $g_i$ 's are generators of  $G$ . The ordered frequency  $f_w(\sigma)$  equals the frequency of occurrence of

the  $k$ -tuple  $(g_1, \dots, g_k)$  from the state space  $\Sigma^k$  by Proposition 3.14. Since, the vector-valued Markov chain of  $k$ -tuples is mixing, it follows from Proposition 3.14 and Remark 3.16 that there exists a full measure subset of  $\mathcal{P}_\infty(C)$  for which  $f_w(\sigma)$  exists.  $\square$

Henceforth we fix  $\mathcal{P}'_\infty \subset \mathcal{P}_\infty^0$  to satisfy the conclusion of Lemma 3.18. Also, let  $\partial G' = \bar{e}(\mathcal{P}'_\infty)$  be the  $\nu$ -full subset of  $\partial G$  obtained as the image of  $\mathcal{P}'_\infty$ . Since  $e(\mathcal{P}'_\infty) \subset \partial G$  has full measure with respect to the Patterson-Sullivan measure  $\nu$  (Corollary 3.13), we have the following:

**Corollary 3.19.** *For  $\nu$ -a.e.  $\xi \in \partial G$ , there exists a geodesic ray  $\sigma = [1, \xi) \in \mathcal{P}'_\infty$ .*

#### 4. APPROXIMATING FPP GEODESICS

The aim in this section is to develop another technical tool needed for the proof of Theorem 5.1. Recall that  $x_n$  is a point at distance  $n$  from the identity element 1 along some fixed word geodesic along some boundary direction, and our objective is to understand  $\mathbb{E}T(1, x_n)$ . The basic idea is to prove that  $T(1, x_n)$  can be approximated by a sum of many independent random variables. To this end we shall consider FPP geodesics between  $x, y \in [1, \xi)$  constrained to lie in  $N_B([x, y])$ . We shall show (see Theorem 4.1 below for a precise statement) that FPP geodesic lengths between  $x, y \in \Gamma$  can be well approximated by the weight of the smallest weight path joining  $x, y$  in  $N_B([x, y])$  for large  $B$ . By the Morse Lemma 2.10 it is irrelevant which geodesic one chooses. Notice that if the support of  $\rho$  is bounded away from 0 and  $\infty$ , this is almost trivial (see Lemma 4.2), but for more general  $\rho$  one needs to work a bit more.

Fix a point  $x_n$  with  $d(1, x_n) = n$ , and a geodesic  $[1, x_n]$ . For  $B \in \mathbb{N}$ , let  $T_B(1, x_n)$  denote the weight of the smallest weight path in  $\Gamma$  joining 1 and  $x_n$  that does not exit  $N_B([1, x_n])$ . The main result in this section is the following.

**Theorem 4.1.** *Given  $\epsilon > 0$ , there exists  $B = B(\epsilon) \in \mathbb{N}$  and  $c = c(\epsilon), C = C(\epsilon) > 0$  such that for all  $n \in \mathbb{N}$ , all  $x_n$  with  $d(1, x_n) = n$ , and all choices of  $[1, x_n]$  we have*

$$\mathbb{P}\left(T_B(1, x_n) \geq T(1, x_n) + \epsilon n\right) \leq Ce^{-cn}.$$

Observe that it suffices to prove Theorem 4.1 only for  $n$  sufficiently large, and we shall take  $n$  to be sufficiently large in the proof without explicitly mentioning it every time. Before proceeding with the proof of Theorem 4.1, we start with the following simple, deterministic, test case:

**Lemma 4.2.** *Given  $\delta, K$ , there exists  $B$  such that the following holds. Suppose that  $(X, d)$  is a  $\delta$ -hyperbolic graph and that  $\rho$  is supported in  $[\frac{1}{K}, K]$ . Then for all  $\omega \in (\Omega, \mathbb{P})$  and  $x, y \in X$*

$$[x, y]_\omega \subset N_B([x, y]).$$

*Proof.* Let  $d_\omega$  denote the metric on  $X$  induced by  $\omega \in (\Omega, \mathbb{P})$ . Then the identity map from  $X$  to itself gives a  $K$ -bi-Lipschitz map from  $(X, d)$  to  $(X, d_\omega)$ . The Lemma is now an immediate consequence of the Morse Lemma 2.10.  $\square$

The remainder of the proof of Theorem 4.1 is a truncation argument which has little to do with the hyperbolicity assumption and works for FPP on any bounded degree connected graph. The first lemma we need shows that the (word) length of the FPP-geodesic between two points at distance  $n$  is  $O(n)$  with exponentially

small failure probability. This is a rather standard result; analogous results have been proved in the Euclidean case in [Kes86] and in the hyperbolic graph context in [BT17, Section 2].

**Lemma 4.3.** *There exists  $n_0 \in \mathbb{N}$ ,  $R < \infty$  and  $c > 0$  (depending both on  $\rho$  and  $|S|$ ) such that for all pairs of vertices  $(u, v)$  in  $\Gamma$  with  $d(u, v) \geq n_0$  we have*

$$\mathbb{P}\left(\ell(\Upsilon(u, v)) \geq Rd(u, v)\right) \leq e^{-cd(u, v)}.$$

Further,  $c \geq c_1\sqrt{R}$  for some  $c_1 > 0$  for  $R$  sufficiently large.

*Proof.* Fix  $\delta > 0$  (to be chosen arbitrarily small later) and choose  $\eta \in (0, 1/2)$  such that  $\rho([0, \eta]) \leq \delta$  (this can be done as there is no atom at 0). Call an edge  $e$  **good** if the weight of  $e$  is at least  $\eta$  and **bad** otherwise. Observe that the weight of a path is at least  $\eta$  times the number of good edges in the path. Observe that by our assumption on  $\rho$  and Theorem 2.13, it follows, by choosing  $R \gg \int_0^\infty x d\rho$  that

$$\mathbb{P}\left(T(u, v) \geq \sqrt{R}d(u, v)\right) \leq e^{-cd(u, v)}$$

for some  $c > 0$ . Also,  $c \geq c_1\sqrt{R}$  for some  $c_1 > 0$  for  $R$  sufficiently large.

Hence, to prove the lemma, it suffices to prove that all paths between  $u$  and  $v$  with word length larger than  $Rd(u, v)$  has  $\omega$ -length larger than  $\sqrt{R}d(u, v)$  with exponentially small (in  $d(u, v)$ ) failure probability. Noticing that the number of (self-avoiding) paths of length  $j$  starting from  $u$  is bounded above by  $|S|^j$ , the above probability is bounded above by

$$\sum_{j=Rd(u, v)+1}^{\infty} |S|^j \mathbb{P}(A_j)$$

where  $A_j$  denotes the event that a self avoiding path  $\gamma$  of length  $j$  has weight smaller than  $\sqrt{R}d(u, v)$ . Now observe that  $\mathbb{P}(A_j)$  is further bounded above by the probability that the number of bad edges on  $\gamma$  is at least  $j - \frac{\sqrt{R}d(u, v)}{\eta}$ . Now choose  $R$  sufficiently large so that for all  $j > Rd(u, v)$ , we have  $j - \frac{\sqrt{R}d(u, v)}{\eta} > \frac{j}{2}$ . Observe that number of bad edges on  $\gamma$  is a sum of i.i.d. Bernoulli variables with expectation  $\delta j$ . By choosing  $\delta$  sufficiently small and using Chernoff inequality 2.12, it follows that

$$\mathbb{P}(A_j) \leq e^{-cj \log(1/2\delta)}$$

where the constant  $c$  is absolute (i.e., does not depend on  $j$  or  $\delta$ ). Now by choosing  $\delta$  sufficiently small this term decays sufficiently fast to kill the entropy term  $|S|^j$  and hence we get

$$\sum_{j=Rd(u, v)+1}^{\infty} |S|^j \mathbb{P}(A_j) \leq e^{-cRd(u, v)}.$$

The exponent here also can be made arbitrarily large by making  $R$  large, thus completing the proof.  $\square$

Lemma 4.3 has the following immediate corollary that we shall use in Section 8.

**Corollary 4.4.** *There exists  $C' > 0$  such that for each  $n$  and for each  $x_n$  with  $d(1, x_n) = n$  we have  $\mathbb{E}[\ell(\Upsilon(1, x_n))] \leq C'n$ .*



We now move towards the truncation argument and the proof of Theorem 4.1. Let  $\epsilon > 0$  be fixed and let  $R$  be as in Lemma 4.3. Let us fix  $0 < \epsilon' < \frac{\epsilon}{2R}$  and  $M = M(R, \epsilon) > 0$  to be chosen sufficiently large later. For  $\omega \in (\Omega, \mathbb{P})$  let us define  $\omega' \in \Omega$  by setting, for all edges  $e \in \Gamma$ ,

- (1)  $\omega'(e) = \omega(e)$  if  $\omega(e) \in [\epsilon', M]$ ;
- (2)  $\omega'(e) = \epsilon'$  if  $\omega(e) \leq \epsilon'$ ;
- (3)  $\omega'(e) = M$  if  $\omega(e) \geq M$ .

The main idea is to use the fact that the geodesic  $[1, x_n]_{\omega'}$ , in the environment  $\omega'$ , must lie within a bounded neighborhood of  $[1, x_n]$ . We then show that for the right choice of  $M$  and  $\epsilon'$ , the geodesics  $[1, x_n]_{\omega}$  and  $[1, x_n]_{\omega'}$  are close in length except for a very small measure set of  $\omega$ 's. We now make this heuristic precise.

*Proof of Theorem 4.1.* Let  $\epsilon'$  be fixed as above and let  $M$  be fixed sufficiently large to be specified later. Let  $B = B(\epsilon', M)$ , given by Lemma 4.2 be such that for all  $\omega \in \Omega$ , we have that  $[1, x_n]_{\omega'}$  is contained in  $N_B([1, x_n])$ . Note that  $[1, x_n]_{\omega'}$  is not necessarily unique but the above conclusion holds for all choices of  $[1, x_n]_{\omega'}$  by Lemma 4.2.

Let  $X'(e)$  denote the random variable define by  $X'(e)(\omega) = \omega'(e)$ . Observe that, trivially,  $\mathbb{E}X'(e) \leq \mathbb{E}X(e) + \epsilon'$ . Set  $R' = \frac{2(\mathbb{E}X(e) + \epsilon')}{\epsilon'}$ , and let  $\gamma^B(\omega)$  (resp.  $\gamma^B(\omega')$ ) denote the path that attains weight  $T_B(1, x_n)(\omega)$  (resp.  $T_B(1, x_n)(\omega')$ ). Notice that by our choice of  $B$ , we have  $\gamma^B(\omega') = [1, x_n]_{\omega'}$ . Let us define the following events:

$$\mathcal{A} = \{\omega \in \Omega : \ell([1, x_n]_{\omega}) \leq Rn\};$$

$$\mathcal{A}' = \{\omega \in \Omega : \ell([1, x_n]_{\omega'}) \leq R'n\}.$$

$$\mathcal{B} = \left\{ \omega \in \Omega : \sum_{e \in [1, x_n]_{\omega'}} (\omega(e) - \omega'(e))_+ \leq \frac{\epsilon n}{2} \right\}.$$

We shall first show that, on  $\mathcal{A} \cap \mathcal{A}' \cap \mathcal{B}$ , we have  $T_B(1, x_n) \leq T(1, x_n) + \epsilon n$ . Observe that for any  $\omega \in \mathcal{A}$ , we have that

$$(4) \quad d_{\omega}(1, x_n) \geq \ell_{\omega'}([1, x_n]_{\omega}) - Rn\epsilon' \geq d_{\omega'}(1, x_n) - \frac{\epsilon n}{2}$$

by our choice of  $\epsilon'$ . Further, we have, for all  $\omega \in \mathcal{A}' \cap \mathcal{B}$

$$(5) \quad \ell_{\omega}(\gamma^B(\omega)) \leq \ell_{\omega}([1, x_n]_{\omega'}) \leq d_{\omega'}(1, x_n) + \frac{\epsilon n}{2}.$$

Comparing (4) and (5), we get that for all  $\omega \in \mathcal{A} \cap \mathcal{A}' \cap \mathcal{B}$

$$\ell_{\omega}(\gamma^B(\omega)) \leq d_{\omega}(1, x_n) + \epsilon n,$$

as desired.

It remains now to provide an appropriate lower bound for  $\mathbb{P}(\mathcal{A} \cap \mathcal{A}' \cap \mathcal{B})$ . Note first that Lemma 4.3 gives  $\mathbb{P}(\mathcal{A}^c) \leq e^{-cn}$  for some  $c > 0$ . Next observe that in the event that  $\ell_{\omega'}([1, x_n]) \leq 2n\mathbb{E}[X'(e)]$ , one obtains that  $\mathcal{A}'$  holds using the definition of  $R'$ . Using Theorem 2.13 we get  $\mathbb{P}((\mathcal{A}')^c) \leq e^{-cn}$  for some  $c > 0$ .

To complete the proof of the theorem it now suffices to show that

$$(6) \quad \mathbb{P}(\mathcal{A}' \cap \mathcal{B}^c) \leq e^{-cn}$$

for some  $c > 0$ .

For this, simply note that  $\mathcal{A}' \cap \mathcal{B}^c \subseteq \mathcal{B}''$  where  $\mathcal{B}''$  is the event

$$\mathcal{B}'' = \left\{ \omega \in \Omega : \sup_{\gamma: \ell(\gamma) \leq R'n} \sum_{e \in \gamma} (\omega(e) - \omega'(e))_+ \geq \frac{\epsilon n}{2} \right\}$$

where the supremum is taken over all self-avoiding paths  $\gamma$  from 1 to  $x_n$  that have length  $\leq R'n$ . Notice now that  $(\omega(e) - \omega'(e))_+ = (\omega(e) - M)_+$ . Further, for each fixed self-avoiding path  $\gamma$  from 1 to  $x_n$  of length  $\leq R'n$  we have, using Theorem 2.14, for  $M$  sufficiently large,

$$\mathbb{P} \left( \sum_{e \in \gamma} (X(e) - M)_+ \geq \frac{\epsilon n}{2} \right) \leq e^{-cn},$$

where  $c = c(M)$  can be made arbitrarily large by making  $M$  arbitrarily large. There are at most  $|S|^{R'n}$  self avoiding paths of length bounded by  $R'n$ . By choosing  $M$  appropriately large and taking a union bound over such paths we conclude that  $\mathbb{P}(\mathcal{B}'') \leq e^{-cn}$  for some  $c > 0$  which completes the proof of (6) and hence the theorem.  $\square$

**Remark 4.5.** Notice that the proof of Theorem 4.1 is above is the only place in this paper where we have used the sub-Gaussian tail hypothesis (1). Clearly the above proof is not optimal and the conditional can easily be relaxed. For example, it is not too hard to see that (6) could be deduced easily if instead of (1), we assumed that  $\rho$  has sub-exponential tails together with the **bounded mean residual life property**, i.e., if  $\rho$  has unbounded support then

$$(7) \quad \sup_{M>0} (\rho([M, \infty)))^{-1} \int_M^\infty x d\rho(x) < \infty.$$

However, as already mentioned, we are not trying to get optimal hypotheses in this paper so we shall not discuss this in more detail.

## 5. VELOCITY

Recall that the Patterson-Sullivan measure on  $\partial G$  is denoted as  $\nu$ . The aim of this Section is to prove the following (see Definition 5.2 below):

**Theorem 5.1.** *For  $\nu$ -a.e.  $\xi \in \partial G$ , the velocity  $v(\xi)$  in the direction of  $\xi$  exists. Further,  $v(\xi)$  is constant  $\nu$ -almost everywhere in  $\partial G$ .*

Recall from Definition 2.1 that the random variable  $T(x, y)$  is defined as

$$T(x, y)(\omega) = d_\omega(x, y)$$

for  $\omega \in (\Omega, \mathbb{P})$ , where

$$d_\omega(x, y) := \inf_\gamma l_\omega(\gamma)$$

is the length of an  $\omega$ -geodesic between  $x, y$ . We first show a law of large numbers result for approximate passage times defined below.

**5.1. Approximate velocity.** We recall from Section 4, the  $B$ -neighborhood versions of the random variable  $T(x, y)$ . For  $x, y \in G$ , define  $d_{\omega, B}(x, y) := \inf_{\gamma} l_{\omega}(\gamma)$ , where  $\gamma$  ranges over all paths from  $x$  to  $y$  contained in  $N_B([x, y])$ . Recall that the  $B$ -passage time  $T_B(x, y)$  from  $x$  to  $y$  is a random variable defined on  $(\Omega, \mathbb{P})$  by the following:

$$T_B(x, y)(\omega) = d_{\omega, B}(x, y).$$

Similarly, for subsets  $U, V$  of  $N_B([x, y])$  the  $B$ -passage time between  $U, V$  will be denoted as  $T_B(U, V)$  when  $x, y$  are understood.

The expected  $B$ -passage time from  $x$  to  $y$  is then given by

$$\mathbb{E}T_B(x, y) := \int_{\Omega} T_B(x, y)(\omega) d\mathbb{P}.$$

For the following definition, let us fix  $\xi \in \partial G, B > 4\delta$  and a word geodesic  $[1, \xi) = \{1 = x_0, x_1, \dots, x_n, \dots\}$  such that  $d(1, x_n) = n$  and  $x_n \rightarrow \xi$ .

**Definition 5.2.** We define the upper (resp. lower)  $B$ -velocity in the direction of  $\xi$  to be

$$\overline{v}_B(\xi) := \limsup_{n \rightarrow \infty} \frac{\mathbb{E}T_B(1, x_n)}{n},$$

(resp.)

$$\underline{v}_B(\xi) := \liminf_{n \rightarrow \infty} \frac{\mathbb{E}T_B(1, x_n)}{n}.$$

If  $\overline{v}_B(\xi) = \underline{v}_B(\xi)$ , we say that the  $B$ -velocity  $v_B(\xi)$  in the direction of  $\xi$  exists and equals  $\overline{v}_B(\xi) = \underline{v}_B(\xi)$ .

For  $\xi \in \partial G, B > 0, \epsilon > 0$ , we say that the  $(B, \epsilon)$ -velocity in the direction of  $\xi$  exists if

$$\overline{v}_B(\xi) - \underline{v}_B(\xi) \leq \epsilon.$$

For  $\xi \in \partial G$ , we define the upper (resp. lower) velocity in the direction of  $\xi$  to be

$$\overline{v}(\xi) := \limsup_{n \rightarrow \infty} \frac{\mathbb{E}(T(1, x_n))}{n},$$

(resp.)

$$\underline{v}(\xi) := \liminf_{n \rightarrow \infty} \frac{\mathbb{E}(T(1, x_n))}{n}.$$

If  $\overline{v}(\xi) = \underline{v}(\xi)$ , we say that the velocity  $v(\xi)$  in the direction of  $\xi$  exists and equals  $\overline{v}(\xi) = \underline{v}(\xi)$ .

**Remark 5.3.** Observe that a priori the quantities  $v_B(\xi), v(\xi)$  etc. defined above need not be well defined as they might depend on the choice of the word geodesic ray  $[1, \xi)$ . However, by considering two choices  $\{1 = x_0, x_1, \dots, x_n, \dots\}$  and  $\{1 = y_0, y_1, \dots, y_n, \dots\}$  of word geodesic ray from 1 in the direction  $\xi$  it follows that  $|T(1, x_n) - T(1, y_n)| \leq T(x_n, y_n)$ . Since  $d(x_n, y_n)$  is uniformly bounded by  $2\delta$  (see the proof of [BH99, Theorem 1.13] for instance), it follows that  $n^{-1}\mathbb{E}T(x_n, y_n) \rightarrow 0$  and hence  $v$  is indeed well-defined. The same argument works for  $v_B$  since  $B > 4\delta$ .

**Remark 5.4.** Velocity usually refers to the inverse of the quantity in Definition 5.2. We are following the convention from [HM95] where it is termed speed. In [Kes93] the same quantity is called the ‘‘time constant’’.

We collect some basic properties of  $B$ -velocity.

**Proposition 5.5.** *For a.e.  $\xi \in \partial G$ ,*

- (1)  $\lim_{B \rightarrow \infty} \overline{v_B}(\xi) = \overline{v}(\xi)$ .
- (2)  $\lim_{B \rightarrow \infty} \underline{v_B}(\xi) = \underline{v}(\xi)$ .

*Proof.* It follows from Theorem 4.1 (and the obvious fact  $\sup_n \frac{1}{n} \mathbb{E}(T_B(1, x_n)) < \infty$ ) that given  $\epsilon > 0$ , there exists  $B_0$  such that for all  $B \geq B_0$ ,

$$\frac{1}{n} (\mathbb{E}(T_B(1, x_n)) - \mathbb{E}(T(1, x_n))) \leq \epsilon.$$

The given Proposition follows immediately.  $\square$

We shall need a basic theorem from Patterson-Sullivan theory.

**Theorem 5.6.** [Coo93] *Let  $G$  be a hyperbolic group and let  $(\partial G, \nu)$  denote its boundary equipped with a Patterson-Sullivan measure. Then the  $G$ -action on  $(\partial G, \nu)$  is ergodic.*

**Lemma 5.7.** *For a.e.  $\xi \in \partial G$ , and every  $B > 0$ ,*

- (1)  $\overline{v_B}(\xi)$  is constant.
- (2)  $\underline{v_B}(\xi)$  is constant.
- (3)  $\overline{v}(\xi)$  is constant.
- (4)  $\underline{v}(\xi)$  is constant.

*Proof.* Each of the functions  $\overline{v_B}(\xi)$ ,  $\underline{v_B}(\xi)$ ,  $\overline{v}(\xi)$  and  $\underline{v}(\xi)$  are invariant under the  $G$ -action. This is because the formulae for  $[o, g] \cup g.[o, \xi]$  versus  $[o, g.\xi]$  differ by only finitely many terms. Ergodicity of the  $G$ -action on  $\partial G$  (Theorem 5.6) furnishes the conclusion.  $\square$

The next lemma allows us to approximate  $B$ -passage times between  $x, y$  by the  $B$ -passage times between the balls  $N_B(x), N_B(y)$  provided  $d(x, y) \gg 1$ .

**Lemma 5.8.** *For all  $B > 0$ , and  $\epsilon > 0$ , there exists  $M \geq 0$  such that for all  $n \geq M$ , if  $d(x, y) = n$ , then*

$$\frac{\mathbb{E}T_B(x, y) - \mathbb{E}T_B(N_B(x), N_B(y))}{n} \leq \epsilon.$$

*Proof.* We start with the following deterministic inequality (for every  $\omega \in (\Omega, \mathbb{P})$ ):

$$T_B(N_B(x), N_B(y))(\omega) \leq T_B(x, y)(\omega) \leq T_B(N_B(x), N_B(y))(\omega) + \max_{u \in N_B(x)} T(x, u)(\omega) + \max_{v \in N_B(y)} T(y, v)(\omega).$$

The first inequality above is obvious. We turn to the second inequality. Let  $u \in N_B(x)$  and  $v \in N_B(y)$  be such that  $T_B(\{u\}, \{v\}) = T_B(N_B(x), N_B(y))$ . Consider the path from  $x$  to  $y$  obtained by concatenating the restricted geodesics between  $x$  and  $u$ ,  $u$  and  $v$ ,  $v$  and  $y$ . Then

$$T_B(x, y)(\omega) \leq T_B(\{u\}, \{v\})(\omega) + T(x, u)(\omega) + T(y, v)(\omega).$$

(We caution the reader here that by  $T_B(\{u\}, \{v\})$  we mean the passage time between  $u, v$  when restricted to  $N_B([x, y])$ .)

Therefore, to prove the lemma, it would suffice to show the existence of a uniform (not depending on  $d(x, y)$ ) upper bound on

$$\mathbb{E} \max_{u \in N_B(x)} T(x, u) + \mathbb{E} \max_{v \in N_B(y)} T(y, v).$$

We shall only bound the first term above, the second term is bounded by an identical argument.

Observe that  $\max_{u \in N_B(x)} T(x, u) \leq \sum_{u \in N_B(x)} T(x, u)$  and further that for any  $u \in N_B(x)$  we have  $\mathbb{E}T(x, u) \leq d(x, u) \int x \, d\rho(x) \leq B \int x \, d\rho(x)$ . Observe that  $|N_B(x)| \leq |S|^B$  (recall that  $S$  is the generating set of  $\Gamma$ ). Hence

$$\mathbb{E} \max_{u \in N_B(x)} T(x, u) \leq |S|^B B \int x \, d\rho(x)$$

which is independent of  $d(x, y)$  completing the proof of the lemma.  $\square$

**Remark 5.9.** Recall that we have assumed that  $G$  is  $\delta$ -hyperbolic. Also, assume that  $B \geq 2\delta$ , so that  $N_B([1, \xi])$  is  $\delta$ -quasiconvex. There exists  $\eta$  (depending only on  $\delta$ ) and  $i_0$  (depending only on  $B, \delta$ ) such that for all  $i \geq i_0$ , the  $(B + \eta)$ -neighborhood of  $x_i$  disconnects  $N_B([1, \xi])$ , i.e. paths in  $N_B([1, \xi])$  from 1 to  $\xi$  necessarily go through the  $(B + \eta)$ -balls about  $x_i$  (see [Mit97, Section 4.2] for instance). Now take an ordered sequence of points  $y_0 = 1, y_1, y_2, \dots$  on  $[1, \xi]$  such that  $d(y_i, y_{i+1}) > 4(B + \eta)$ . Then the above disconnection property of  $(B + \eta)$ -neighborhoods shows that any path  $\gamma \subset N_B([1, \xi])$  can be decomposed into connected pieces whose interiors lie either in  $N_B([1, \xi]) \setminus (\cup_i N_{(B+\eta)}(y_i))$  or entirely inside some  $N_{(B+\eta)}(y_i)$ .

Note that  $N_{B+\eta}(z)$  has cardinality bounded by  $a^{B+\eta}$  for some  $a > 0$  depending only on  $\Gamma$ , and its diameter is  $2(B + \eta)$ . By group-invariance, the expected passage time from any point in  $N_{B+\eta}(z)$  to any other is bounded in terms of  $B, \eta$  and the passage time distribution  $\rho$ . The number of connected pieces in the above decomposition that "backtrack", i.e. begin and end on the same  $N_{(B+\eta)}(y_i)$  is thus bounded by  $a^{B+\eta}$  and for any such piece, the expected passage time is bounded in terms of  $B, \eta, \rho$ .

Lemma 5.8 along with the above observation will be useful in understanding the behavior of  $\omega$ -geodesics constructed as a concatenation of segments that travel from  $N_B(y_i)$  to  $N_B(y_{i+1})$ , where  $\{y_i\}$  is a suitable coarse-graining of  $[1, \xi]$ .

## 5.2. Coarse-graining and existence of velocity.

**Definition 5.10.** For  $n \in \mathbb{N}$  and  $[1, \xi]$  a geodesic ray, let  $y_0, y_1, y_2, \dots$  be the sequence of points on  $[1, \xi]$  with  $y_0 = 1$  and  $d(y_i, 1) = ni$ . Let

$$\mathbb{E}T([1, \xi], B, n, i) := \mathbb{E}(T_B(y_i, y_{i+1}))$$

denote the expected  $B$ -passage time from  $y_i$  to  $y_{i+1}$  along  $[1, \xi]$ . The **coarse-grained  $(B, 0)$  velocity at scale  $n$  along  $[1, \xi]$**  is said to exist if the limit

$$\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{i=0}^{m-1} \mathbb{E}T([1, \xi], B, n, i)$$

of Cesaro averages of the sequence  $\{\mathbb{E}T([1, \xi], B, n, i)\}_i$  exists, and is then defined as

$$v_B([1, \xi], n) := \frac{1}{n} \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{i=0}^{m-1} \mathbb{E}T([1, \xi], B, n, i).$$

More generally, let

$$\overline{v}_B([1, \xi], n) := \frac{1}{n} \limsup_{m \rightarrow \infty} \frac{1}{m} \sum_{i=0}^{m-1} \mathbb{E}T([1, \xi], B, n, i),$$

and

$$\underline{v}_B([1, \xi], n) := \frac{1}{n} \liminf_{m \rightarrow \infty} \frac{1}{m} \sum_{i=0}^{m-1} \mathbb{E}T([1, \xi], B, n, i).$$

The **coarse-grained**  $(B, \epsilon)$  **velocity at scale**  $n$  along  $[1, \xi]$  is said to exist if  $\overline{v}_B([1, \xi], n) - \underline{v}_B([1, \xi], n) \leq \epsilon$ .

**Proposition 5.11.** *Let  $d$  and  $\mathcal{P}'_\infty$  be as obtained from Lemma 3.18. Given  $B \gg 1$  (large) and  $\epsilon > 0$  (small) there exists  $M$  such that for all  $n \in d\mathbb{N} \cap [M, \infty)$  and for all  $\sigma = [1, \xi] \in \mathcal{P}'_\infty$ , the coarse-grained  $(B, \epsilon)$  velocity at scale  $n$  in direction  $\xi$  exists.*

*Proof.* Recall that  $\partial G' \subset \partial G$  denotes the  $\nu$ -full subset of  $\partial G$  obtained as the image of  $\mathcal{P}'_\infty$  under the evaluation map (Corollary 3.19).

Given  $B, \epsilon$  as in the statement of the Proposition, there exists  $M \geq 0$  by Lemma 5.8, such that for all  $n \geq M$ , if  $d(x, y) = n$ , then

$$(8) \quad \frac{\mathbb{E}T_B(x, y) - \mathbb{E}T_B(N_B(x), N_B(y))}{n} \leq \frac{\epsilon}{2}.$$

We next construct a coarse-graining of  $[1, \xi]$  at scale  $n$ , i.e. let  $y_0, y_1, y_2, \dots$  be the sequence of points on  $[1, \xi]$  with  $y_0 = 1$  and  $d(y_i, 1) = ni$ . Then we have (see Figure 1)

$$\mathbb{E}T_B(1, y_m) \leq \sum_{i=0}^{m-1} \mathbb{E}T_B(y_i, y_{i+1}) = \sum_{i=0}^{m-1} \mathbb{E}T([1, \xi], B, n, i)$$

For a lower bound of  $\mathbb{E}T_B(1, y_m)$ , observe the following. From Remark 5.9, the  $(B + \eta)$ -balls about  $y_i$  disconnect  $N_B([1, \xi])$  for all  $i$ , provided the coarse-graining scale  $n$  is large enough. Hence there exists  $B'$  (depending on  $B, \eta$ ) such that the path attaining the weight  $T_B(1, y_m)$  has disjoint (across  $i$ ) sub-paths  $\gamma_i$  contained in  $N_B(y_i, y_{i+1})$  with endpoints  $u_i, v_i$  contained in  $N_{B+\eta}(y_i)$  and  $N_{B+\eta}(y_{i+1})$  respectively, and such that there exist paths from  $y_i$  to  $u_i$  and  $v_i$  to  $y_{i+1}$  contained in  $N_B(y_i, y_{i+1})$  with word length bounded above by  $B'$ . (This statement follows from the last part of Remark 5.9.) Notice that, by arguing as in Lemma 5.8 it follows that the expected  $\omega$ -length of the maximum over all paths of length  $\leq B'$  from  $y_i$  is at most  $B''$  for some  $B''$  independent of  $i$ . This implies that the expected length of the path between  $u_i$  and  $v_i$  described above is at least  $\mathbb{E}T_B(N_B(y_i), N_B(y_{i+1})) - 2B''$ . It therefore follows that

$$(9) \quad \mathbb{E}T_B(1, y_m) \geq \sum_{i=0}^{m-1} \mathbb{E}T_B(N_B(y_i), N_B(y_{i+1})) - 2mB''.$$

Thus,  $\frac{1}{mn} \mathbb{E}T_B(1, y_m)$  is sandwiched between the quantities  $\frac{1}{mn} \sum_{i=0}^{m-1} \mathbb{E}T_B(y_i, y_{i+1})$  and  $(\frac{1}{mn} \sum_{i=0}^{m-1} \mathbb{E}T_B(N_B(y_i), N_B(y_{i+1})) - \frac{B''}{n})$ . By (8), the difference between the latter two quantities is at most  $(\frac{\epsilon}{2} + \frac{B''}{n})$ . To prove the Proposition, it therefore suffices to show that the Cesaro averages  $\frac{1}{m} \sum_{i=0}^{m-1} \mathbb{E}T_B(y_i, y_{i+1})$  and  $\frac{1}{m} \sum_{i=0}^{m-1} \mathbb{E}T_B(N_B(y_i), N_B(y_{i+1}))$  converge as  $m \rightarrow \infty$ .

To prove these convergence statements, we invoke Lemma 3.18: for every geodesic word  $w$  of length  $n$  and  $\sigma = [1, \xi] \in \mathcal{P}'_\infty$ , the ordered frequency  $f_w(\sigma)$  exists. Hence, using group invariance,

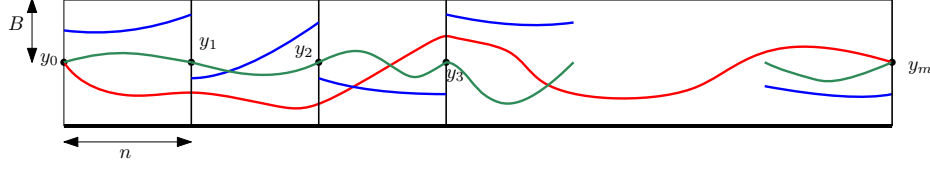


FIGURE 1. Approximating  $T_B(1, y_m)$  as in the proof of Proposition 5.11. The red path denotes the path attaining  $T_B(1, y_m)$ , i.e., the smallest FPP-length path connecting  $y_0 = 1$  and  $y_m$  that is contained in  $N_B([y_0, y_m])$ . The blue segments denote the paths attaining  $T_B(N_B(y_i), N_B(y_{i+1}))$  and the green segments denote the paths attaining  $T_B(y_i, y_{i+1})$ . Clearly, the blue segments are shorter (in FPP-length) than the green segments, and the FPP length of the red curve is sandwiched between the sum of lengths of the green segments and the blue segments.

$$(10) \quad \frac{1}{m} \sum_{i=0}^{m-1} \mathbb{E}T_B(N_B(y_i), N_B(y_{i+1})) \rightarrow \sum_{|w|=n} f_w(\sigma) \mathbb{E}T_B(N_B(1), N_B(e(w)))$$

as  $m \rightarrow \infty$ , and

$$(11) \quad \frac{1}{m} \sum_{i=0}^{m-1} \mathbb{E}T_B(y_i, y_{i+1}) \rightarrow \sum_{|w|=n} f_w(\sigma) \mathbb{E}T_B(1, e(w))$$

as  $m \rightarrow \infty$ . This completes the proof of the Proposition.  $\square$

Since  $\epsilon \rightarrow 0$  as  $n \rightarrow \infty$  in Lemma 5.8, this gives us the following immediate Corollary:

**Corollary 5.12.** *For all  $B > 0$  and all  $\sigma = [1, \xi] \in \mathcal{P}'_\infty$ ,*

$$\overline{v}_B([1, \xi], n) - \underline{v}_B([1, \xi], n) \rightarrow 0$$

as  $n \rightarrow \infty$ .

The main technical theorem of this section is the following:

**Theorem 5.13.** *For every  $B > 4\delta$  and  $\xi \in \partial G'$ , the  $B$ -velocity  $v_B(\xi)$  in the direction of  $\xi$  exists.*

*Proof.* It suffices, by Remark 5.3 to consider a geodesic ray  $\sigma \in \mathcal{P}'_\infty$  in the direction  $\xi \in \partial G'$ . Let us denote  $\sigma = \{1 = x_0, x_1, \dots, x_N, \dots\}$ . Recall that by definition of  $\mathcal{P}'_\infty$ , we have that for all  $n \in d\mathbb{N}$  as above and  $\epsilon > 0$ , there exists  $N_0 \in \mathbb{N}$  such that the following holds. For all  $N \geq N_0$  (such that  $N$  a multiple of  $n$ ), the fraction of ordered occurrences of each geodesic word  $w$  with  $|w| = n$  in the first  $N$ -length segment of  $\sigma$  is in  $[f_w(\sigma) - \epsilon, f_w(\sigma) + \epsilon]$ .

It therefore follows from the sandwiching argument of Proposition 5.11 that given  $n \in \mathbb{N}$  sufficiently large and  $\epsilon > 0$ , there exists  $N'_0$  such that for all  $N \geq N'_0$  (again  $N$  is a multiple of  $n$ ) we have

$$\frac{\mathbb{E}T_B(1, x_N)}{N} \in [v_B([1, \xi], n) - \epsilon, \overline{v_B}([1, \xi], n) + \epsilon].$$

Observe now that for  $N_1, N_2 \in \mathbb{N}$  with  $|N_1 - N_2| \leq n$  we also have that

$$|\mathbb{E}T_B(1, x_{N_1}) - \mathbb{E}T_B(1, x_{N_2})| \leq \mathbb{E}T_B(x_{N_1}, x_{N_2}) \leq cn$$

for some absolute constant  $c > 0$ . This implies that

$$\frac{\mathbb{E}T_B(1, x_N)}{N} \in [v_B([1, \xi], n) - \epsilon, \overline{v_B}([1, \xi], n) + \epsilon]$$

holds for all  $N$  larger than some  $N''_0$ , not merely the multiples of  $n$ . Using Corollary 5.12, this implies (by taking  $n$  arbitrarily large) immediately that  $\{\frac{\mathbb{E}T_B(1, x_N)}{N}\}_N$  is Cauchy, and hence

$$v_B(\xi) := \lim_{N \rightarrow \infty} \frac{\mathbb{E}T_B(1, x_N)}{N}$$

exists, thus completing the proof.  $\square$

We are now ready to prove Theorem 5.1.

*Proof of Theorem 5.1.* Theorem 5.13 shows that for all  $B > 0$ ,  $v_B(\xi)$  in the direction of  $\xi$  exists for a.e.  $\xi \in (\partial G, \nu)$ . Hence by Lemma 5.7,  $v_B(\xi)$  is constant for a.e.  $\xi \in (\partial G, \nu)$ . Letting  $B \rightarrow \infty$ , Theorem 5.1 is now an immediate consequence of Proposition 5.5.  $\square$

It is clear that  $v(\xi)$  is upper bounded by  $\int x d\rho(x)$ . One can also show, along the lines of Lemma 4.3, that  $v(\xi)$  is uniformly bounded away from 0. This is done in the following lemma.

**Lemma 5.14.** *There exists  $\epsilon = \epsilon(\rho, \Gamma) > 0$  and  $c = c(\epsilon) > 0$  such that for any  $n \in \mathbb{N}$  and any  $x_n \in \Gamma$  with  $d(1, x_n) = n$  we have*

$$\mathbb{P}(T(1, x_n) \leq \epsilon n) \leq e^{-cn}.$$

*Proof.* By Lemma 4.3, by choosing  $R$  sufficiently large, it suffices to show that with probability at least  $1 - e^{-cn}$ , each path  $\gamma$  of length  $\leq Rn$  connecting 1 and  $x_n$  satisfies  $\ell_\omega(\gamma) \geq \epsilon n$ . There are at most  $D^{Rn}$  many such paths. Hence it suffices to show that for  $\epsilon$  sufficiently small, the probability that any fixed path of length between  $n$  and  $Rn$  has  $\omega$ -length  $\leq \epsilon n$  has probability upper bounded by  $e^{-c(\epsilon)n}$  and  $c(\epsilon)$  can be made arbitrarily large by making  $\epsilon$  arbitrarily small. It suffices to prove the above statement for a fixed path of length  $n$ . Let  $a(\epsilon)$  denote the probability the the weight of a specific edges is  $\leq \sqrt{\epsilon}$ . The probability described above can be upper bounded by  $\mathbb{P}(\text{Bin}(n, a(\epsilon)) \geq (1 - \sqrt{\epsilon})n)$ , and the desired conclusion follows, as in Lemma 4.3, using Chernoff's inequality and noting that  $a(\epsilon) \rightarrow 0$  as  $\epsilon \rightarrow 0$ .  $\square$

It follows from Lemma 5.14 that  $v(\xi)$  is uniformly bounded away from 0.



**5.3. Special cases and examples.** Let  $g \in G$  be of infinite order. Then  $g$  acts by North-South dynamics on  $\partial G$  with an attracting fixed point (denoted  $g^\infty$ ) and a repelling fixed point (denoted  $g^{-\infty}$ ). Note that the repelling fixed point of  $g$  is the attracting fixed point of  $g^{-1}$  and vice versa.

**Definition 5.15.** *An attracting point of an infinite order element in  $\partial G$  is called a pole.*

Note that the set of poles is countable; hence of measure zero with respect to the Patterson-Sullivan measure.

**Lemma 5.16.** *For every pole  $\xi$ , the velocity  $v(\xi)$  exists.*

*Proof.* This is a reprise of an analogous argument in the Euclidean case (see e.g. [Kes93, Theorem A]) and we only sketch the argument. Let  $\xi = g^\infty$ . The sequence  $\{g^n\}$  defines a quasigeodesic by the classification of isometries of a hyperbolic space [Gro85]. Clearly,

$$\mathbb{E}T(1, g^{m+n}) \leq \mathbb{E}T(1, g^m) + \mathbb{E}T(g^m, g^{m+n}) = \mathbb{E}T(1, g^m) + \mathbb{E}T(1, g^n),$$

where the last equality follows by group-invariance. The Lemma is now a consequence of Kingman's sub-additive ergodic theorem.  $\square$

**5.3.1. FPP on  $\mathbb{Z}$ .** We show first that FPP on Cayley graphs of  $\mathbb{Z}$  with respect to different generating sets, but with the same passage time distribution, may exhibit different speeds along the same direction. This will be an ingredient for a counterexample in Section 5.3.2 below. Consider the following two Cayley graphs of  $\mathbb{Z}$ :  $\Gamma_1 = \Gamma_1(\mathbb{Z}, \pm 1)$  and  $\Gamma_2 = \Gamma_2(\mathbb{Z}, \pm 1, \pm 2)$ . Consider FPP on  $\Gamma_1$  and  $\Gamma_2$  with the same continuous passage time distribution  $\rho$  with mean  $\mu \in (0, \infty)$ . Let  $a_\rho \in [0, \infty)$  and  $b_\rho \in (0, \infty]$  denote the infimum and the supremum of support of  $\rho$ .

To distinguish between the two graphs, we shall denote the corresponding passage times by  $T^{\Gamma_1}(\cdot, \cdot)$  and  $T^{\Gamma_2}(\cdot, \cdot)$  respectively. Clearly  $\frac{\mathbb{E}T^{\Gamma_1}(0, n)}{n} \rightarrow \mu$  as  $n \rightarrow \infty$ . Notice that  $d^{\Gamma_2}(0, 2n) = n$  and hence the following lemma gives an example illustrating the claim above.

**Lemma 5.17.** *If  $2a_\rho < b_\rho$ , then*

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E}T^{\Gamma_2}(0, 2n)}{n} < \mu.$$

*Proof.* Observe that

$$T^{\Gamma_2}(0, 2n) \leq \sum_{i=1}^n \min(X_i, Y_{2i-1} + Y_{2i})$$

where  $X_i$  and  $Y_i$  are independent families of i.i.d. variables with distribution  $\rho$ . Therefore, it suffices to show that

$$\mathbb{E}[X_i - \min(X_i, Y_{2i-1} + Y_{2i})] > 0.$$

Clearly if  $2a_\rho < b_\rho$ , by assumption of continuity of  $\rho$ , there exists  $\delta < \frac{b_\rho - 2a_\rho}{4}$  such that  $\mathbb{P}(Y_i \in [a_\rho, a_\rho + \delta]) \geq \delta$  and  $\mathbb{P}(X_i \in [b_\rho - \delta, b_\rho]) \geq \delta$ . It therefore follows that the non-negative variable  $Z_i := X_i - \min(X_i, Y_{2i-1} + Y_{2i})$  satisfies  $\mathbb{P}(Z_i \geq \delta) \geq \delta^3$  and hence  $\mathbb{E}Z_i \geq \delta^4$ , completing the proof of the lemma.  $\square$

5.3.2. *FPP on the free group.* We next give an example to show that there exists a hyperbolic group  $G$  and a generating set  $S$  such that

- (1) there exist  $\xi_1, \xi_2 \in \partial G$  such that  $v(\xi_1), v(\xi_2)$  exist *but are unequal*.
- (2) there exists  $\xi \in \partial G$  such that  $v(\xi)$  does not exist.

Let  $G = F_2 = \langle a, b \rangle$  be the free group on 2 generators. Fix  $S = \{a^{\pm 1}, b^{\pm 1}, b^{\pm 2}\}$  to be the generating set and let  $\Gamma = \Gamma(G, S)$ . Let  $\xi_1 = a^\infty$ ,  $\xi_2 = b^\infty$ . By Lemma 5.16,  $v(\xi_1), v(\xi_2)$  exist. By Lemma 5.17,  $v(\xi_1) > v(\xi_2)$ .

We now construct a direction  $\xi \in \partial G$  such that  $v(\xi)$  does not exist. Let  $w = a^{m_1}b^{n_1}a^{m_2}b^{n_2}\dots$  be an infinite reduced word such that  $m_i, n_i$  are defined recursively (as a tower function) by

- (1)  $m_1 = 1$ ,
- (2)  $n_i = 2^{2^{m_i}}$ , for  $i \geq 1$ ,
- (3)  $m_{i+1} = 2^{2^{n_i}}$ , for  $i \geq 1$ .

Let  $\xi \in \partial G$  denote the boundary point represented by  $w$ . Let  $u_k, v_k$  denote the finite subwords of  $w$  given by

$$u_k = a^{m_1}b^{n_1}a^{m_2}b^{n_2}\dots a^{m_k},$$

and

$$v_k = a^{m_1}b^{n_1}a^{m_2}b^{n_2}\dots a^{m_k}b^{n_k}$$

Since the sequence  $\{m_1, n_1, m_2, n_2, \dots, m_k, n_k, \dots\}$  grows like the tower function, the length of  $u_k$  (resp.  $v_k$ ) is dominated by  $m_k$  (resp.  $n_k$ ). Further, since every vertex of  $\Gamma$  disconnects it, we have (by group-invariance),

$$\frac{T(1, u_k)}{T(1, a^{m_k})} \rightarrow 1,$$

and

$$\frac{T(1, v_k)}{T(1, b^{n_k})} \rightarrow 1$$

as  $k \rightarrow \infty$ . Hence,

$$v(1, u_k) \rightarrow v(a^\infty) = v(\xi_1)$$

and

$$v(1, v_k) \rightarrow v(b^\infty) = v(\xi_2)$$

as  $k \rightarrow \infty$ . Since  $v(\xi_1) \neq v(\xi_2)$ ,  $v(\xi)$  does not exist.

## 6. DIRECTION OF $\omega$ -GEODESIC RAYS

Recall again our basic set-up:  $G$  is a hyperbolic group and  $\Gamma = \Gamma(G, S)$  is a Cayley graph with respect to a finite symmetric generating set. In Definition 2.1 we have defined  $\omega$ -geodesics between  $x, y \in \Gamma$ . We would like to extend this notion to  $\omega$ -geodesic rays. The Gromov boundary (resp. compactification) of  $\Gamma$  is denoted as  $\partial G$  (resp.  $\hat{\Gamma}$ ) (since  $\partial G$  is independent of the generating set  $S$  we have not used  $\partial\Gamma$  to denote the boundary of  $\Gamma$ , using the suggestive notation  $\partial G$  instead).

**Definition 6.1.** For  $\omega \in (\Omega, \mathbb{P})$ , a semi-infinite (resp. bi-infinite) path  $\sigma_\omega$  is said to be an  $\omega$ -geodesic ray (resp. a bi-infinite  $\omega$ -geodesic) if every finite subpath of  $\sigma_\omega$  is an  $\omega$ -geodesic.

A path  $\sigma$  is said to **accumulate on**  $\xi \in \partial G$  if there exist  $v_n \in \sigma$  such that  $v_n \rightarrow \xi$  as  $n \rightarrow \infty$ .

**Defining directions of  $\omega$ -geodesic rays:** We briefly motivate the notion of direction of an  $\omega$ -geodesic ray. In Euclidean space  $\mathbb{R}^n$ , a direction  $u$  is identified with an element of  $T_I(0) \subset T_0(\mathbb{R}^n)$ , the unit tangent sphere at  $0 \in \mathbb{R}^n$  contained in the tangent space  $T_0(\mathbb{R}^n)$  to  $\mathbb{R}^n$  at  $0$ . Since tangent spaces are not available in our setup, we have to interpret  $T_I(0)$  appropriately for  $\Gamma$ . The exponential map from  $T_0(\mathbb{R}^n)$  to  $\mathbb{R}^n$  is a diffeomorphism sending  $tu$  (with  $t \in \mathbb{R}_+$  and  $u \in T_I(0)$ ) to the unique geodesic in  $\mathbb{R}^n$  starting at  $0$  and in the direction given by  $u$ . This allows us to identify  $T_I(0)$  to the boundary  $\partial\mathbb{R}^n$  given by asymptote classes of geodesics as in Lemma 2.11. We shall thus parametrize directions of  $\omega$ -geodesic rays in  $\Gamma$  by points  $\xi \in \partial G$ . To this end we make the following definition.

**Definition 6.2.** *An  $\omega$ -geodesic ray  $\sigma_\omega$  accumulating on  $\xi \in \partial G$  has direction  $\xi$  if its only accumulation point in  $\partial G$  is  $\xi$ .*

The main objective of this section is to establish that Definition 6.2 is indeed a natural definition, every  $\omega$ -geodesic ray has a direction (Theorem 6.6) and there exist  $\omega$ -geodesic rays in each direction  $\xi \in \partial G$ . (Theorem 6.7). The analogous statements for Euclidean lattices (where direction, as usual, is parametrized by points on the unit sphere) is due to Newman [New95], where it is proved under additional curvature assumptions on the limiting shape of random balls; recent partial progress without these assumptions has been made in [Hof08, DH14, DH17]. The hyperbolic geometry allows these results to go through without such assumptions in our case.

We shall say that a sequence  $x_n$  in  $\Gamma$  satisfies  $x_n \rightarrow \infty$  as  $n \rightarrow \infty$  if  $d(x_n, 1) \rightarrow \infty$  as  $n \rightarrow \infty$ . The main idea is to observe that if an  $\omega$ -geodesic ray starting from  $o$  has two distinct accumulation points then they must have points  $x_n, y_n \rightarrow \infty$  on them such that the word geodesic  $[x_n, y_n]$  passes through a bounded neighborhood of  $o$ . This would force  $[x_n, y_n]_\omega$  to also almost surely pass through finite neighborhoods of  $o$ ; which will lead to a contradiction. The first step of making this argument precise is the following proposition.

**Proposition 6.3.** *Let  $C \geq 0$  and  $o \in \Gamma$  be given. Then for a.e.  $\omega \in (\Omega, \mathbb{P})$ , there exists  $R_\omega > 0$  such that the following holds:*

*For every sequence  $x_n, y_n \rightarrow \infty$  such that  $[x_n, y_n]$  passes through  $N_C(o)$  (i.e., some word geodesic between  $x_n$  and  $y_n$  passes through  $N_C(o)$ ), the  $\omega$ -geodesic  $[x_n, y_n]_\omega$  passes through the  $R_\omega$ -neighborhood of  $o$ .*

The proof of this proposition adapts the proof of [BT17, Theorem 1.3], except that we do not assume  $x_n, y_n$  to lie on a fixed Morse geodesic passing through  $o$ , thus requiring some additional work.

For  $A \subset \Gamma$  and  $p \in \Gamma$ , define a **nearest-point projection** of  $p$  onto  $A$  as  $\Pi(p) = a$  if  $d(p, a) = d(p, A)$ . Nearest-point projections onto geodesics (or more generally quasiconvex sets) are coarsely well-defined in the following sense (see the proof of [BH99, Theorem 1.13, p405] and [Mj14, Lemma 2.20]):

**Lemma 6.4.** *Let  $x \in \Gamma$  and  $u, v \in [1, \xi]$  be such that  $d(x, [1, \xi]) = d(x, u) = d(x, v)$ . Then  $d(u, v) \leq 2\delta$ . More generally, given  $\kappa > 0$ , there exists  $C_0$  such that if  $A \subset \Gamma$  is  $\kappa$ -quasiconvex, then for any  $x \in \Gamma$  and  $u, v \in A$  with  $d(x, A) = d(x, u) = d(x, v)$ , we have  $d(u, v) \leq C_0$ .*

We shall need the following geometric lemma, similar in content (and proof) to Lemma 2.2 and Lemma 3.1 of [BT17].

**Lemma 6.5.** *Let  $\Gamma$  be as above. Then there exist  $A, C', R_0 > 0$  such that for all  $R \geq R_0$  the following holds. Let  $[x, y] \subset \Gamma$  be a geodesic and  $o \in [x, y]$ . Let  $\sigma$  be a path joining  $x, y$  such that  $\sigma \cap N_{100R}(o) = \emptyset$ . Then there exist  $u, v \in \sigma$  such that  $d(u, [x, y]) = R = d(v, [x, y])$ ,  $[u, v] \cap N_A(o) \neq \emptyset$ . Further,  $u, v$  satisfy the following properties in addition. Let  $\sigma_{uv}$  be the subpath of  $\sigma$  between  $u, v$  and let  $\Pi$  denote nearest-point projection onto  $[x, y]$ . Then  $\Pi(u), \Pi(v) \in [x, y]$  lie on opposite sides of  $o$  with  $d(\Pi(u), o) \geq 99R$ ,  $d(\Pi(v), o) \geq 99R$ . Also,*

- (1)  $\frac{\ell(\sigma_{uv})}{2R + d(\Pi(u), \Pi(v))} \rightarrow \infty$  as  $R \rightarrow \infty$ .
- (2)  $d(o, \sigma_{uv}) \leq C'\ell(\sigma_{uv})$ .

The proof uses basic hyperbolic geometry and is postponed to Appendix A. We are now ready to prove Proposition 6.3.

*Proof of Proposition 6.3.* For the purpose of this proof, let us define  $\mu := \int x d\rho(x)$ . Let  $\Omega_1 = \Omega_1(T)$  denote the set of all configurations such that

$$\sup_{\gamma} (\ell_{\omega}(\gamma) - T\mu\ell(\gamma)) < \infty$$

where the supremum is taken over all paths  $\gamma$  passing through  $N_C(o)$ . We claim that, if  $T$  is sufficiently large,  $\mathbb{P}(\Omega_1(T)) = 1$ . Indeed, since  $N_C(o)$  is finite, it suffices to consider separately the paths passing through each vertex in  $N_C(o)$ . Clearly, there are at most  $n|S|^n$  many paths of length  $n$  passing through a fixed  $o_1 \in N_C(o)$  and by choosing  $T$  sufficiently large and using Theorem 2.13 it follows that the probability of any such path  $\gamma$  satisfying  $\ell_{\omega}(\gamma) \geq T\mu\ell(\gamma)$  is at most  $e^{-2|S|^n}$ . Taking a union bound over all paths of length  $n$  passing through  $o_1$ , followed by a union bound over all choices of  $o_1$  and an application of Borel-Cantelli lemma completes the proof of the claim. From now on, we shall fix a full measure subset  $\Omega_1$  satisfying the above property.

Next, let  $\Omega_2 = \Omega_2(\epsilon)$  be the set of all configurations such that

$$\inf_{\gamma} (\ell_{\omega}(\gamma) - \epsilon\ell(\gamma)) > -\infty$$

where the infimum is taken over all paths  $\gamma$  such that  $d(o, \gamma) \leq C'\ell(\gamma)$  (where  $C'$  is as in Lemma 6.5). By the same argument as in the proofs of Lemma 4.3 and Lemma 5.14 (cf. [BT17, Lemma 2.5]) there exists  $\epsilon > 0$  sufficiently small such that  $\mathbb{P}(\Omega_2) = 1$ . Let us fix a full measure subset  $\Omega_2$  satisfying the above property.

We now show that the full measure subset  $\Omega' := \Omega_1 \cap \Omega_2$  satisfies the hypothesis in the statement of the proposition. We argue by contradiction. Suppose for  $\omega \in \Omega'$ , there exist  $R_n \uparrow \infty$  and two sequences  $x_n, y_n \rightarrow \infty$  such that  $[x_n, y_n]$  passes through  $N_C(o)$  but  $[x_n, y_n]_{\omega}$  avoids  $N_{(100R_n + C)}(o)$ .

Let  $\Pi_n$  denote nearest point projection onto  $[x_n, y_n]$  and  $o_n \in N_C(o) \cap [x_n, y_n]$ . Since  $N_C(o)$  is finite, we can assume after passing to a subsequence if necessary that  $o_n = o'$  for all  $n$ . Hence,  $[x_n, y_n]_{\omega}$  avoids  $N_{(100R_n)}(o')$  for all  $n$ . By Lemma 6.5, there exist  $A, C' \geq 0$  (depending only on  $\delta$ ) and  $u_n, v_n$  on  $[x_n, y_n]_{\omega}$  such that,

- (1)  $d(u_n, [x_n, y_n]) = R_n = d(v_n, [x_n, y_n])$ ,
- (2)  $\Pi_n(u_n), \Pi_n(v_n) \in [x_n, y_n]$  lie on opposite sides of  $o$  with  $d(\Pi_n(u_n), o') \geq 99R_n$ ,  $d(\Pi_n(v_n), o') \geq 99R_n$ ,
- (3)  $\frac{\ell([u_n, v_n]_{\omega})}{2R_n + d(\Pi_n(u_n), \Pi_n(v_n))} \rightarrow \infty$  as  $n \rightarrow \infty$ . In fact, for some  $\alpha > 0$  depending only on the hyperbolicity constant  $\delta$ ,  $\ell([u_n, v_n]_{\omega}) \geq d(\Pi_n(u_n), \Pi_n(v_n))e^{\alpha R_n}$
- (4)  $d(o', [u_n, v_n]_{\omega}) \leq C'\ell([u_n, v_n]_{\omega})$ .

Using the definition of  $\Omega_2$  we can therefore write

$$\ell_\omega([u_n, v_n]_\omega) \geq \epsilon \ell([u_n, v_n]_\omega) - r_1(\omega) \geq \epsilon B_n(2R_n + d(\Pi_n(u_n), \Pi_n(v_n))) - r_1(\omega)$$

where,  $B_n \rightarrow \infty$  as  $n \rightarrow \infty$ .

The path between  $u_n$  and  $v_n$  obtained by concatenating  $[u_n, \Pi_n(u_n)]$ ,  $[\Pi_n(u_n), \Pi_n(v_n)]$  and  $[\Pi_n(v_n), v_n]$  passes through  $o'$  (by Property 2 above). Using the definition of  $\Omega_1$  we also have

$$\ell_\omega([u_n, v_n]_\omega) \leq T\mu(2R_n + d(\Pi_n(u_n), \Pi_n(v_n))) + r_2(\omega).$$

Notice that  $r_1$  and  $r_2$  do not depend on  $n$ . Comparing these inequalities we get  $(\epsilon B_n - T\mu)d(\Pi_n(u_n), \Pi_n(v_n)) \leq 2T\mu R_n + r_1(\omega) + r_2(\omega)$ , which is a contradiction for large enough  $n$ , since  $d(\Pi_n(u_n), \Pi_n(v_n)) \geq R_n$  and  $B_n \rightarrow \infty$  as  $n \rightarrow \infty$ .  $\square$

We can now prove that almost surely all  $\omega$ -geodesic rays have direction.

**Theorem 6.6.** *For  $o \in \Gamma$ , for a.e.  $\omega \in (\Omega, \mathbb{P})$ , all  $\omega$ -geodesic rays starting at  $o$  has a direction  $\xi \in \partial G$ .*

*Proof.* Since an  $\omega$ -geodesic ray necessarily visits infinitely many  $x \in \Gamma$ , it has some accumulation point  $\xi \in \partial G$ .

Let  $C' \geq 0$  be fixed. We first show that almost surely there does not exist an  $\omega$ -geodesic ray starting at  $o$  such that it has two accumulation points  $\xi, \eta \in \partial G$  with  $\langle \xi, \eta \rangle_o \leq C'$ . Observe that there exists  $C \geq 0$  (depending only on  $C'$  and  $\delta$ ) with the following property: if such an  $\omega$ -geodesic ray  $\sigma$  existed then there would be points  $x_n, y_n \in \sigma$  with  $x_n, y_n \rightarrow \infty$  such that (i)  $y_n$  belongs to the restriction of  $\sigma$  between  $o$  and  $x_n$  (ii)  $[x_n, y_n]$  passes through  $N_C(o)$  (since we can choose  $C = C' + 2\delta$ , [BH99, Chapter III.H.1]). Let  $\Omega' = \Omega'(C)$  be as in Proposition 6.3. By Proposition 6.3, for all  $\omega \in \Omega'$ ,  $\sigma$  returns infinitely often to  $N_C(o)$ , implying that  $\sigma$  is self-intersecting, a contradiction. Letting  $C' \rightarrow \infty$  completes the proof of the theorem.  $\square$

Next we show that for every  $\xi \in \partial G$ , there is an  $\omega$ -geodesic ray with direction  $\xi$ .

**Theorem 6.7.** *Fix  $\xi \in \partial G$  and  $x_n \in \Gamma$  such that  $x_n \rightarrow \xi$ . For a.e.  $\omega \in (\Omega, \mathbb{P})$  the sequence of  $\omega$ -geodesics  $[o, x_n]_\omega$  from  $o$  to  $x_n$  converges (up to subsequence) to an  $\omega$ -geodesic ray  $[o, \xi]_\omega$  having direction  $\xi$ .*

*Proof.* The proof is similar to the previous one. Fixing  $C' \geq 0$  we shall show that for a.e.  $\omega$ , any sub-sequential limit  $[o, \xi]_\omega$  of  $[o, x_n]_\omega$  (sub-sequential limits exist by a compactness argument) cannot accumulate at  $\eta \in \partial G$  with  $\langle \xi, \eta \rangle_o \leq C'$ . Observe again, that there exists  $C \geq 0$  with the following property: if  $[o, \xi]_\omega$  had such an accumulation point, then (if necessary, passing to a subsequence), there will be points  $y_n \rightarrow \infty$  such that  $y_n \in [o, x_n]_\omega$  and  $[x_n, y_n]$  passes through  $N_C(o)$  (choosing  $C = C' + 2\delta$ , as in the proof of Theorem 6.6). Note that the difference from the previous case is that  $x_n$  does not necessarily lie on  $[o, \xi]_\omega$ . Nevertheless, by considering  $\omega \in \Omega' = \Omega'(C)$  given in Proposition 6.3 we see that  $[y_n, x_n]_\omega$  must pass through a point  $z_n$  in a finite  $(R_\omega)$  neighborhood of  $o$ . Since  $y_n \in [o, x_n]_\omega$  it follows that  $y_n \in [o, z_n]_\omega$  which implies that there exists infinitely many points on the finite union of  $\omega$ -geodesics  $\cup_{o' \in N_{R_\omega}(o)} [o, o']_\omega$ , a contradiction. As before we finish the proof by taking  $C' \rightarrow \infty$ .  $\square$

## 7. COALESCENCE

We shall prove in this section that for FPP on  $\Gamma$ , semi-infinite geodesics in a fixed direction almost surely coalesce. This question is of fundamental importance in FPP on  $\mathbb{Z}^d$ , with progress being made under curvature assumptions [New95, LN96] and with more recent progress using Busemann functions in [Hof08, DH14, DH17, AH]. Some of the finer questions have in recent years been addressed in the exactly solvable models of exponential directed last passage percolation on  $\mathbb{Z}^2$  [Cou11, FP05, Pim16, BSSa, BHS] but most of the fundamental questions remain open for FPP with general weights, even on  $\mathbb{Z}^d$ . For spaces with negative curvature, asymptotic coalescence is a folklore expectation due to the thin triangles condition of Gromov. For FPP on Cayley graphs of hyperbolic groups we shall establish this: semi-infinite geodesics in the same boundary direction almost surely coalesce.

**7.1. Hyperplanes and their properties.** In this subsection we deduce some of the basic estimates from hyperbolic geometry that will be needed to prove coalescence. Let  $\xi \in \partial G$  be a boundary point and let  $[1, \xi) = \{1 = x_0, x_1, \dots\}$  be a geodesic ray from 1 to  $\xi$ .

**Definition 7.1.** Fix  $\xi$ ,  $[1, \xi) = \{1 = x_0, x_1, \dots\}$  as above. For any  $x_i \in [1, \xi)$  with  $i > 2\delta$ , define the **elementary hyperplane**  $H_i^e (= H_i^e(\xi))$  through  $x_i$  as follows:

$$H_i^e := \{x \in \Gamma \mid d(x, [1, \xi)) = d(x, x_i)\}.$$

Finally define the **hyperplane**  $H_i (= H_i(\xi))$  through  $x_i$  to be the  $2\delta$ -neighborhood of  $H_i^e$ :

$$H_i = N_{2\delta}(H_i^e).$$

(See Figure 2 below.)

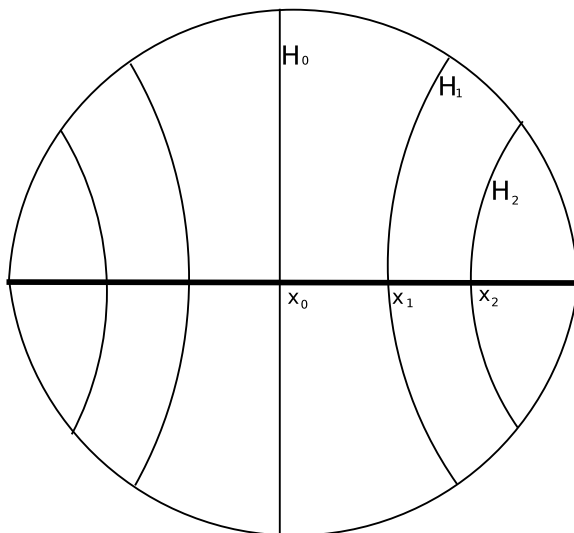


FIGURE 2. Hyperplanes in the hyperbolic plane  $\mathbf{H}^2$  and their exponential divergence

Equivalently, if  $\Pi$  denotes nearest-point projection onto  $[1, \xi)$ ,  $H_i^e = \Pi^{-1}(x_i)$  (see Lemma 6.4).

**Remark 7.2.** The reason for thickening  $H_i^e$  to  $H_i$  in Definition 7.1 above is to take care of the fact that nearest-point projections are only coarsely well-defined in the sense of Lemma 6.4.

We record some of the properties of hyperplanes. The following Lemma says that hyperplanes are quasiconvex and separate  $\Gamma$ . Thus the resulting half-spaces may be regarded as nested.

**Lemma 7.3.** *For all  $\xi$  and  $i$ , hyperplanes  $H_i$  are  $\delta$ -quasiconvex.*

*Hyperplanes  $H_i$  separate  $\Gamma$ , i.e. there exist **half-spaces**  $H_i^+, H_i^-$  such that*

- (a) *any two points of  $H_i^+$  (resp.  $H_i^-$ ) that can be connected by a path in  $\Gamma$  can be joined by a path lying in  $H_i^+ \cup H_i$  (resp.  $H_i^- \cup H_i$ ),*
- (b) *any path from  $H_i^+$  to  $H_i^-$  intersects  $H_i$ .*
- (c) *There exists  $C_1$  depending on  $\delta$  alone such that for all  $k \geq C_1$ ,  $H_{i+k} \subset H_i^+$ ,  $H_{i+k}^+ \subset H_i^+$ ,  $H_{i-k} \subset H_i^-$ ,  $H_{i-k}^- \subset H_i^-$ .*

*Proof.* Let  $\Pi : \Gamma \rightarrow [1, \xi)$  denote nearest-point projection. Let  $x \in H_i^e$ . Then  $\Pi(x) = x_i$ . Every point  $y \in [x, x_i]$  satisfies  $\Pi(y) = x_i$ , and hence  $[x, x_i] \subset H_i^e$ . For any  $u, v \in H_i$ ,  $[u, v] \subset N_\delta([u, x_i] \cup [u, x_i]) \subset N_\delta(H_i)$ , i.e.  $H_i^e$  is  $\delta$ -quasiconvex. The first assertion follows (since a  $2\delta$ -neighborhood of a  $\delta$ -quasiconvex set is  $\delta$ -quasiconvex).

Define  $H_i^{e+} = \cup_{j \geq i} H_j$ ,  $H_i^{e-} = \Gamma \setminus H_i^{e+}$ . Finally, define  $H_i^+$  (resp.  $H_i^-$ ) to be a  $4\delta$ -neighborhood of  $H_i^{e+}$  (resp.  $H_i^{e-}$ ). Essentially the same argument as in the previous paragraph establishes the quasiconvexity of  $H_i^{e+}$  and  $H_i^{e-}$ .

Now, let  $\gamma$  be any path in  $\Gamma$  joining  $u, v \in H_i^+$ . Let  $\Pi'$  denote nearest-point projection onto  $H_i$ . By Lemma 6.4, nearest point projections onto hyperplanes are coarsely well defined. Projecting every maximal subpath of  $\gamma$  contained in  $H_i^-$  (if any) onto  $H_i$  using  $\Pi'$  gives us a new path  $\gamma'$  joining  $u, v$  and contained in  $H_i^+$ , proving Item (a).

Let  $\sigma$  be a path in  $\Gamma$  joining  $u \in H_i^-$  to  $v \in H_i^+$ . Then  $\Pi(u) = x_j$  for some  $j < i$  and  $\Pi(v) = x_k$  for some  $k > i$ . Since nearest-point projections are coarsely Lipschitz [Mit98, Lemma 3.2], it follows (essentially by continuity) that there exists  $w \in \sigma$  such that  $\Pi(w) = x_i$ , i.e.  $w \in H_i$ . Item (b) follows.

Item (c) now follows from the second assertion of Lemma 7.4 below, which implies that  $[x_i, x_{i+D}]$  is coarsely the shortest path between  $H_i, H_{i+D}$  whenever  $D$  is large enough.  $\square$

The following is a general fact [BH99, Chapter 3.H.1] (see also [Mit98, Lemma 3.1] and Figure 2 above).

**Lemma 7.4** (Exponential divergence of hyperplanes). *For  $D$  sufficiently large and any  $i \in \mathbb{N}$ ,  $H_i, H_{i+D}$  diverge exponentially, i.e. there exists  $R_0 \geq 0$  and  $C, \alpha > 0$  such that for  $u \in H_i$  and  $v \in H_{i+D}$  with  $d(u, x_i) \geq R \geq R_0$  and  $d(v, x_{i+D}) \geq R \geq R_0$ , any path joining  $u, v$  lying outside the open  $R$ -neighborhood of  $[1, \xi)$  has length at least  $CDe^{\alpha R}$ .*

*Further,  $[u, x_i] \cup [x_i, x_{i+D}] \cup [x_{i+D}, v]$  is a  $(1, 4\delta)$ -quasigeodesic lying in a  $4\delta$ -neighborhood of  $[u, v]$ .*

**Definition 7.5.** For  $C_1$  as in Lemma 7.3 and  $D \geq C_1$ , the **region**  $\mathbb{B}_{i,D}$  between  $H_i$  and  $H_{i+D}$  is defined to be

$$\mathbb{B}_{i,D} = H_i \cup (H_i^+ \cap H_{i+D}^-) \cup H_{i+D}.$$

Let  $a \in H_i$  and  $b \in H_{i+D}$ . Let  $\gamma$  be a path from  $a$  to  $b$  in the region  $\mathbb{B}_{i,D}$  between hyperplanes. By Lemma 7.4,  $[a, x_i] \cup [x_i, x_{i+D}] \cup [x_{i+D}, b]$  is a  $(1, 4\delta)$ -quasigeodesic lying in a  $4\delta$ -neighborhood of the geodesic  $[a, b]$ . Then there exist  $x'_i, x'_{i+D}$  on  $[a, b]$  such that  $d(x_i, x'_i) \leq 4\delta$  and  $d(x_{i+D}, x'_{i+D}) \leq 4\delta$ . Thus nearest point projections (of  $x \in \Gamma$ ) onto  $[a, b]$  and  $[a, x_i] \cup [x_i, x_{i+D}] \cup [x_{i+D}, b]$  lie in a uniformly bounded neighborhood of each other.

For each  $z \in \gamma$ , let  $\Pi(z)$  denote a nearest point projection onto  $[a, x_i] \cup [x_i, x_{i+D}] \cup [x_{i+D}, b]$ . We assume below that  $\Pi$  is surjective (instead of being only coarsely so) to avoid cluttering the discussion. Let  $u$  be the last point on  $\gamma$  projecting to  $x_i$  and let  $v$  be the first point projecting to  $x_{i+D}$ . Suppose that  $\gamma$  lies outside the  $R$ -neighborhood of  $[x_i, x_{i+D}]$ , i.e.  $\gamma \cap N_R([1, \xi]) = \emptyset$ . Let  $u_1 \in \gamma$  be the last point on  $\gamma$  before  $u$  such that  $d(u_1, [a, b]) = R$ . Let  $a_1 = \Pi(u_1)$ . Then  $a_1 \in [a, x_i]$ . Similarly, let  $v_1 \in \gamma$  be the first point on  $\gamma$  after  $v$  such that  $d(v_1, [a, b]) = R$ . Let  $b_1 = \Pi(v_1)$ . Then  $b_1 \in [x_{i+D}, b]$ . See Figure 3 below.

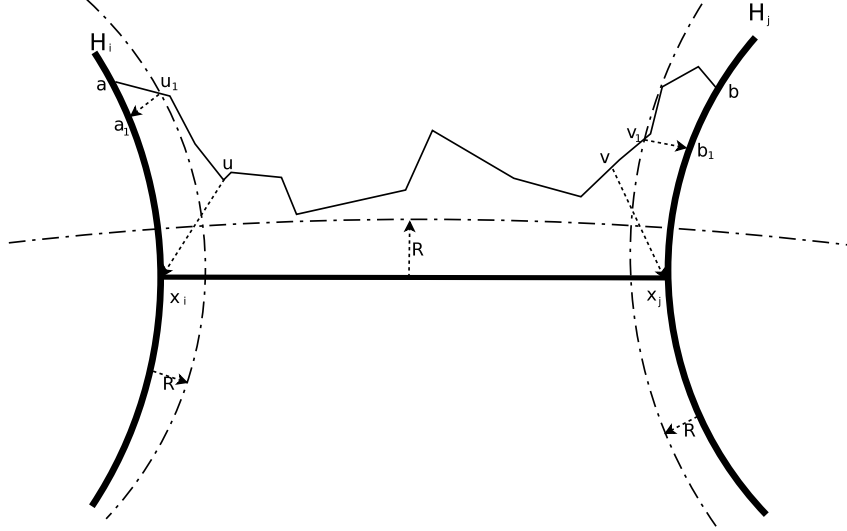


FIGURE 3. Exponential divergence of paths away from geodesics.

Let  $\gamma'$  be the subpath of  $\gamma$  between  $u_1$  and  $v_1$ . By exponential divergence of geodesics [BH99, p. 412-413] we have the following:

**Lemma 7.6.** *There exists  $R_0 \geq 0$  and  $\alpha > 0$  depending only on  $\delta$  such that for  $R \geq R_0$ ,  $D \geq R_0$ ,  $\gamma \cap N_R([1, \xi]) = \emptyset$  and  $\gamma', u_1, v_1, a_1, b_1$  as above,*

$$l(\gamma') \geq \left( d(a_1, x_i) + d(x_i, x_{i+D}) + d(x_{i+D}, b_1) \right) e^{\alpha R}.$$



**7.2. Coalescence of  $\omega$ -geodesic rays.** We now make the necessary definitions for  $\omega$ -geodesics rays.

Recall the following from Definitions 6.1 and 6.2. For  $\xi \in \partial G$  and  $\omega \in \Omega$ , a semi-infinite path  $\gamma = \{x_i\}_{i \in \mathbb{N}}$  is called a semi-infinite  $\omega$ -geodesic (or an  $\omega$ -geodesic ray) in direction  $\xi$  if every finite segment of  $\gamma$  is an  $\omega$ -geodesic between the respective endpoints,  $x_n \rightarrow \xi$ , and further for any sequence of points  $y_n$  on  $\gamma$  with  $d(y_n, 1) \rightarrow \infty$ , we have  $y_n \rightarrow \xi$ . We restate Theorem 6.7 as follows for ready reference within this section:

**Lemma 7.7.** *For each  $o \in \Gamma$  and  $\xi \in \partial G$ , for a.e.  $\omega \in \Omega$  there exists a  $\omega$ -geodesic ray started at  $o$  in direction  $\xi$ .*

The next lemma asserts that for each fixed direction and starting point, the geodesic ray is almost surely unique, proof of which is obtained together with the proof of the coalescence result (Theorem 7.9) below.

**Lemma 7.8.** *For each  $\xi \in \partial G$  and each  $o \in \Gamma$ , for a.e.  $\omega \in \Omega$ , there exists a unique  $\omega$ -geodesic ray (denoted  $[o, \xi)_\omega$ ) starting from  $o$  in direction  $\xi$ .*

We now state the main theorem of this section.

**Theorem 7.9.** *For a.e.  $\omega \in (\Omega, \mathbb{P})$ ,  $o_1, o_2 \in G$  and any direction  $\xi$ ,  $\omega$ -geodesic rays  $[o_1, \xi)_\omega$  and  $[o_2, \xi)_\omega$  a.s. coalesce.*

As mentioned before Lemma 7.8 and Theorem 7.9 will be proved together. First we record the following simple geometric fact which will be useful throughout this section. Let  $[1, \xi) = \{1 = x_0, x_1, \dots\}$  denote a geodesic ray that corresponds to the direction  $\xi$ . For  $i \geq 0$  and  $D \in \mathbb{N}$ , let  $H_{iD}$  denote the hyperplane perpendicular to  $[1, \xi)$  at  $x_{iD}$ . Then for  $D$  sufficiently large, any  $o \in \Gamma$ , and for a.e.  $\omega \in (\Omega, \mathbb{P})$ , every  $\omega$ -geodesic ray from  $o$  to  $\xi$  must cross  $H_{iD}$  for all sufficiently large  $i$  (see Lemma 7.3).

Before starting with the formal arguments let us present briefly the idea of the proof. Fix a direction  $\xi \in \partial G$ . Let us define the (almost surely well-defined) random object  $\Upsilon_1$  (resp.  $\Upsilon_2$ ) by setting  $\Upsilon_1(\omega) = [o_1, \xi)_\omega$  (resp.  $\Upsilon_2(\omega) = [o_2, \xi)_\omega$ ). Without loss of generality we shall assume that  $o_1 = 1$ . Let us fix, as above, a geodesic ray  $[1, \xi) = \{1 = x_0, x_1, \dots\}$ . For  $D \in \mathbb{N}$ , recall that  $H_{iD}$  denotes the hyperplane perpendicular to  $[1, \xi)$  at  $x_{iD}$ . Without loss of generality we shall also assume that  $o_2$  is such that the projection of  $o_2$  to  $[1, \xi)$  is 1; for if the projection to  $[1, \xi)$  is given by  $\Pi(o_2) = x_i$ , then we simply translate  $x_i$  to 1 and proceed with the argument below. It will be clear that the same proof can be easily modified to include the more general case. Consider now the FPP geodesics  $\Upsilon_1$  and  $\Upsilon_2$ . Notice that by Lemma 7.3, if  $D$  is sufficiently large, then for each  $i \in \mathbb{Z}_{\geq 0}$ , and for  $j = 1, 2$ ,  $\Upsilon_j$  must have a last exit from  $H_{iD}$ . We shall denote it as  $U_j = U_j(i, D)$ , where the dependence on  $i$  and  $D$  will be suppressed when clear from context. Further, there is a first entry point into  $H_{(i+1)D}$  following  $U_j$  which we shall denote as  $V_j$ . Let  $\Upsilon_{1,i}$  (resp.  $\Upsilon_{2,i}$ ) denote the restriction of  $\Upsilon_1$  (resp.  $\Upsilon_2$ ) between  $U_1$  and  $V_1$  (resp.  $U_2$  and  $V_2$ ). Notice that  $\Upsilon_{1,i}$  (resp.  $\Upsilon_{2,i}$ ) is the FPP geodesic between  $U_1$  and  $V_1$  (resp.  $U_2$  and  $V_2$ ) that is also further restricted to lie in the region  $\mathbb{B}_{iD, D}$ .

The basic idea behind the proof of Theorem 7.9 is to show that  $\Upsilon_{1,i}$  and  $\Upsilon_{2,i}$  intersect with probability bounded below independently of  $i$ . Theorem 7.9 (and also Lemma 7.8) will follow from the next proposition.

**Proposition 7.10.** *For  $D = D(\rho, \Gamma)$  sufficiently large, there exists  $\beta = \beta(D, \Gamma, \rho) > 0$  independent of  $i$  such that for each  $i \in \mathbb{N}$  there exists an event  $B_i$  depending only on the configuration restricted to  $N_{D/10}(\mathbb{B}_{iD, D})$  with the following properties:*

- (i)  $\mathbb{P}(B_i) \geq \beta$ .
- (ii) *For all sufficiently large  $i$ , on  $B_i$ , every pair of  $\omega$ -geodesic rays started from either  $1$  or  $o_2$  in the direction  $\xi$  intersect on  $[x_{iD}, x_{(i+1)D}]$ .*

Postponing the proof of Proposition 7.10, let us first quickly show how this lemma implies Lemma 7.8 and Theorem 7.9.

*Proof of Lemma 7.8.* Observe that if two distinct  $\omega$ -geodesic rays intersect infinitely often, it implies that there exist multiple  $\omega$ -geodesics between some pairs of vertices. By continuity of the passage time distribution  $\rho$ ,  $[u, v]_\omega$  is unique for each  $u, v \in \Gamma$  and for a.e  $\omega$ . Now, on this probability one event, if there exist distinct  $\omega$ -geodesic rays started from  $o$  in direction  $\xi$  then they cannot intersect infinitely often. Observe on the other hand the events  $\{B_{2i} : i \geq 1\}$  as in Proposition 7.10 depend on disjoint sets of edges, hence are independent, which in turn using the Borel-Cantelli Lemma implies  $\limsup_{i \rightarrow \infty} B_i$  is an almost sure event. On this almost sure event, distinct  $\omega$ -geodesic rays from  $o$  in direction  $\xi$ , if exist, must intersect infinitely often, leading to a contradiction. This completes the proof of the lemma.  $\square$

*Proof of Theorem 7.9.* As in the above proof, notice that  $\limsup_{i \rightarrow \infty} B_i$  is an almost sure event. This in particular, implies, recalling the notation from the discussion preceding Proposition 7.10, that almost surely, for infinitely many  $i$ ,  $\Upsilon_{1,i} \cap \Upsilon_{2,i} \neq \emptyset$  and in particular,  $\Upsilon_1$  and  $\Upsilon_2$  almost surely intersect. By Lemma 7.8 (actually this is merely a consequence of continuity of passage time distribution) it follows that if  $\Upsilon_1$  and  $\Upsilon_2$  intersect, they must coalesce (almost surely), completing the proof of the theorem.  $\square$

It remains to prove Proposition 7.10. The proof is divided into three cases depending on the support of the passage time distribution  $\rho$ :

- Case (i):  $\rho$  has support  $\subseteq [a, b]$  with  $0 < a < b < \infty$ ,
- Case (ii): the passage time can take arbitrarily small values, but not arbitrarily large values,
- Case (iii): the passage time can also take arbitrarily large values.

**Case (i): Support of  $\rho$  is bounded away from 0 and  $\infty$ :** This is the easiest case. Assume without loss of generality that the support is  $[a, b]$  with  $0 < a < b < \infty$ . By continuity of the passage time distribution this implies some mass around  $a$  as well as  $b$  (this is what we really need). Now, by the Morse Lemma 2.10, there exists  $R$ , depending on  $a, b$  such that  $\Upsilon_1$  and  $\Upsilon_2$  are contained in  $N_R([1, \xi])$  and  $N_R([o_2, \xi])$  respectively. Further, since  $d(o_2, 1) = d(o_2, [1, \xi])$ , it follows that  $[o_2, 1] \cup [1, \xi]$  is a quasigeodesic with constant depending only on  $\delta$  (see [Mit98, Lemma 3.1] for instance) and hence we can assume, by increasing  $R$  if necessary, that  $\Upsilon_2$  is contained in  $N_R([o_2, 1] \cup [1, \xi])$ . Hence there exists  $c$  depending only on  $a, \delta$  such that for each  $i > cd(o_2, 1) + c$ , and any  $\omega$ -geodesic ray from either  $1$  or  $o_2$  in the direction  $\xi$  will have a sub-segment contained in  $N_R([x_{iD}, x_{(i+1)D}])$  with starting and ending points contained in  $N_R(\{x_{iD}\})$  and  $N_R(\{x_{(i+1)D}\})$  respectively. Proposition 7.10 in this case will follow from Lemma 7.11 below.

**Lemma 7.11.** *Suppose  $\rho$  has support  $[a, b]$  bounded away from 0 and  $\infty$ , and let  $R$  be as above. Then there exists  $D = D(R, \Gamma)$  sufficiently large,  $\beta = \beta(D, R, \rho, \Gamma) > 0$  and an event  $B_i$  for each  $i$  depending only on the edges in  $N_R([x_{iD}, x_{(i+1)D}])$  with  $\mathbb{P}(B_i) \geq \beta$ , such that for all  $\omega \in B_i$  we have the following: for each  $u \in N_R(\{x_{iD}\})$  and  $v \in N_R(\{x_{(i+1)D}\})$ , and a path  $\gamma$  between  $u$  and  $v$  contained in  $N_R([x_{iD}, x_{(i+1)D}])$ , there exists a path  $\gamma'$  between  $u$  and  $v$  that intersects  $[x_{iD}, x_{(i+1)D}]$  in at least  $\frac{3D}{5}$  edges and  $\ell_\omega(\gamma') \leq \ell_\omega(\gamma)$ .*

*Proof.* Let  $\epsilon > 0$  be such that  $a + \epsilon < a + 10\epsilon < b - \epsilon$  and  $\rho([a, a + \epsilon]), \rho([b - \epsilon, b]) > 0$ . The geodesic subsegment of  $[1, \xi)$  between  $x_{iD}$  and  $x_{(i+1)D}$  will be denoted by  $[x_{iD}, x_{(i+1)D}]$  as usual. Let  $S = S_i$  be the union of

- (1) the set of all edges on  $[x_{iD}, x_{(i+1)D}]$ ,
- (2) the set of all edges that lie on some geodesic joining  $x_{iD}$  to some point  $u \in N_R(\{x_{iD}\})$ ,
- (3) and the set of all edges that lie on some geodesic joining  $x_{(i+1)D}$  to some point  $v \in N_R(\{x_{(i+1)D}\})$ .

Let  $T := N_R([x_{iD}, x_{(i+1)D}]) \setminus S$ . Let  $B_i$  denote the event

$$B_i := \{\omega \in \Omega : \omega(e) \leq a + \epsilon \forall e \in S, \omega(e) \geq b - \epsilon \forall e \in T\}.$$

Clearly, there exists  $\beta = \beta(D, R, \rho, \Gamma) > 0$  such that  $\mathbb{P}(B_i) \geq \beta$ . Fix  $\omega \in B_i$ . Consider  $u \in N_R(\{x_{iD}\})$  and  $v \in N_R(\{x_{(i+1)D}\})$ , and a path  $\gamma$  between  $u$  and  $v$  contained in  $N_R([x_{iD}, x_{(i+1)D}])$  that has an intersection with  $[x_{iD}, x_{(i+1)D}]$  of cardinality less than  $\frac{3D}{5}$ . For  $D$  sufficiently large ( $D \gg R$ ) it follows that  $\gamma$  must intersect  $T$  in at least  $\frac{3D}{10}$  many edges. This is because the number of edges in  $N_R(\{x_{iD}\})$  or  $N_R(\{x_{(i+1)D}\})$  can be made much smaller than  $D$  by choosing  $D \gg R$  (e.g.  $D \sim e^{aR}$  for  $a \gg 1$ ). Again by choosing  $D$  sufficiently large (depending on  $R$  and some fixed  $\epsilon/2 > \epsilon' > 0$ ) we can ensure that  $\ell(\gamma) \geq (1 - \epsilon')D$  (since the number of edges of  $\gamma$  in  $N_R(\{x_{iD}\})$  or  $N_R(\{x_{(i+1)D}\})$  can be made much smaller than  $D$  by choosing  $D \gg R$  as above). By definition of  $B_i$  it follows that,  $\ell_\omega(\gamma) \geq \frac{3D}{10}(b - \epsilon) + a(\frac{7D}{10} - \epsilon'D)$ . Consider now the path  $\gamma'$  obtained by concatenating  $[u, x_{iD}]$ ,  $[x_{iD}, x_{(i+1)D}]$  and  $[x_{(i+1)D}, v]$ . Clearly by definition of  $B_i$ ,  $\ell_\omega(\gamma') \leq (a + \epsilon)(D + 2R)$ . Again, by choosing  $D$  sufficiently large depending on  $R, \epsilon$  and  $\epsilon'$  we can ensure  $\ell_\omega(\gamma') \leq \ell_\omega(\gamma)$ , as desired.  $\square$

We shall prove statements similar to Lemma 7.11 for the cases (ii) and (iii) as well. However, in those cases we are not afforded the luxury of knowing that the paths  $\Upsilon_1$  and  $\Upsilon_2$  are contained in some  $R$  neighborhood of a fixed word geodesic. Hence we shall consider all FPP geodesics from some point in  $H_{iD}$  to some point in  $H_{(i+1)D}$  and show that on a positive probability event all those geodesics must spend sufficient time on the word geodesic segment  $[x_{iD}, x_{(i+1)D}]$ . Although the statement of this lemma is identical for both of the cases, the proofs are slightly different and hence we shall state the results separately.

**Case (ii): Support of  $\rho$  is bounded away from  $\infty$  but not from 0:** We start with the case where 0 is contained in the support of  $\rho$ . We have the following analogue of Lemma 7.11 in this case.

**Lemma 7.12.** *Suppose that the support of  $\rho$  contains 0 and is bounded away from  $\infty$ . Then there exists  $D = D(\Gamma)$  sufficiently large,  $\beta = \beta(D, \rho, \Gamma) > 0$  and an event*

$B_i$  for each  $i$  depending only on the edges in  $N_{D/10}(\mathbb{B}_{iD,D})$  with  $\mathbb{P}(B_i) \geq \beta$ , such that for all  $\omega \in B_i$  we have the following: for each  $u \in H_{iD}$  and  $v \in H_{(i+1)D}$ , and a path  $\gamma$  from  $u$  to  $v$  contained in the region  $\mathbb{B}_{iD,D}$ , there exists a path  $\gamma'$  between  $u$  and  $v$  that intersects  $[x_{iD}, x_{(i+1)D}]$  in at least  $\frac{3D}{5}$  edges and  $\ell_\omega(\gamma') \leq \ell_\omega(\gamma)$ .

*Proof.* Let us set  $j = (i + 0.01)D$ ,  $j' = (i + 0.99)D$ ,  $k = (i + 0.02)D$ ,  $k' = (i + 0.98)D$  and assume without loss of generality that these are all integers (taking  $D$  a large enough multiple of 100). Let  $S = S_{i,d}$  denote the set of all edges contained in  $N_d([x_{iD}, x_j]) \cup N_d([x_{j'}, x_{(i+1)D}]) \cup [x_{iD}, x_{(i+1)D}]$ . We call the edges in  $\mathbb{B}_{k,0.96D}$  (i.e., the edges in the region between  $H_k$  and  $H_{k'}$ ) except those on  $[1, \xi)$  **unspecified edges**. Let us now define the following events. Let  $\mathcal{C} = \mathcal{C}(\epsilon) := \{X(e) \leq \epsilon \forall e \in S\}$ . Let  $\mathcal{D} = \mathcal{D}(d)$  (resp.  $\mathcal{D}'$ ) denote the event that all  $\omega$ -geodesics between pairs of points, one each in  $H_{iD}$  and  $H_j$  (resp. one each in  $H_{j'}$  and  $H_{(i+1)D}$ ) that are also contained in  $\mathbb{B}_{iD,D}$  enter  $N_d([x_{iD}, x_j])$  (resp.  $N_d([x_{j'}, x_{(i+1)D}])$ ). Finally, let  $\mathcal{E} = \mathcal{E}(\epsilon)$  denote the event that the total weight of unspecified edges along any path from a point in  $N_d([x_{iD}, x_j])$  to a point in  $N_d([x_{j'}, x_{(i+1)D}])$  that intersects  $[x_{iD}, x_{(i+1)D}]$  in less than  $\frac{3D}{5}$  edges is at least  $2\epsilon D$ .

We claim that, for appropriate choices of  $d, \epsilon$  and for  $D$  sufficiently large,  $B_i := \mathcal{C} \cap \mathcal{D} \cap \mathcal{D}' \cap \mathcal{E}$  satisfies the condition in the statement of the lemma. Clearly, for  $D$  sufficiently large compared to  $d$ ,  $B_i$  depends only on the edges in  $N_{D/10}(\mathbb{B}_{iD,D})$ , as required. First let us fix  $\omega \in B_i$ . Consider a path  $\gamma$  as in the statement of the lemma. Clearly, without loss of generality we can replace, in  $\gamma$ , the last crossing of  $H_{iD}$  to  $H_j$  and the last crossing of  $H_{j'}$  to  $H_{(i+1)D}$  by the  $\omega$ -geodesics between the respective endpoints. With a minor abuse of notation we shall also call this path  $\gamma$ . If  $\gamma$  intersects  $[x_{iD}, x_{(i+1)D}]$  in at least  $\frac{3D}{5}$  edges then there is nothing to prove, so let us suppose the contrary. By definition of  $\mathcal{D}$  and  $\mathcal{D}'$  there exist points  $w_1 \in \gamma \cap N_d([x_{iD}, x_j])$  and  $w_2 \in \gamma \cap N_d([x_{j'}, x_{(i+1)D}])$ . Let  $w'_1$  and  $w'_2$  denote the projections of  $w_1$  and  $w_2$  onto  $[1, \xi)$  respectively. Let  $\gamma'$  be the path obtained from  $\gamma$  by replacing the segment between  $w_1$  and  $w_2$  (call it  $\gamma_*$ ) by the path obtained by concatenating  $[w_1, w'_1]$ ,  $[w'_1, w'_2]$  (i.e., the restriction of  $[1, \xi)$  between  $w'_1$  and  $w'_2$ ) and  $[w'_2, w_2]$ . Clearly it suffices to show that  $\ell_\omega([w_1, w'_1]) + \ell_\omega([w'_1, w'_2]) + \ell_\omega([w'_2, w_2]) \leq \ell_\omega(\gamma_*)$ . By the definition of  $\mathcal{C}$  the left hand side above is at most  $2\epsilon D$  (for  $D$  sufficiently large compared to  $d$ ) whereas by definition of  $\mathcal{E}$ , the right hand side is bounded below by  $2\epsilon D$ . Hence the event  $B_i$  defined above satisfies the conclusion of the lemma.

It remains to show a lower bound for  $\mathbb{P}(B_i)$ . Observe first that for  $D$  sufficiently large the events  $\mathcal{C} \cap \mathcal{D} \cap \mathcal{D}'$  and  $\mathcal{E}$  depend on disjoint sets of edges (Lemma 7.3) and hence  $\mathbb{P}(B_i) = \mathbb{P}(\mathcal{E})\mathbb{P}(\mathcal{C} \cap \mathcal{D} \cap \mathcal{D}')$ . Observe also that both  $\mathcal{C}$  and  $\mathcal{D} \cap \mathcal{D}'$  are decreasing events in the weight configuration on  $S$  and hence by the FKG inequality (Theorem 2.15), we have  $\mathbb{P}(\mathcal{C} \cap \mathcal{D} \cap \mathcal{D}') \geq \mathbb{P}(\mathcal{C})\mathbb{P}(\mathcal{D} \cap \mathcal{D}')$ . By our assumption on  $\rho$ , for each  $\epsilon > 0$ , there exists  $\beta_1 = \beta_1(D, d, \epsilon, \rho) > 0$  such that  $\mathbb{P}(\mathcal{C}) \geq \beta_1$ . Hence to complete the proof, it suffices to show that  $\mathbb{P}(\mathcal{E})$  and  $\mathbb{P}(\mathcal{D} \cap \mathcal{D}')$  are also both bounded away from 0. This is done below in Lemma 7.13 and Lemma 7.14, invoking which the proof is completed.  $\square$

**Lemma 7.13.** *In the set-up above, there exists  $\epsilon > 0$  such that  $\mathbb{P}(\mathcal{E}) \geq \frac{1}{2}$  for all  $D$  sufficiently large.*

*Proof.* The proof will use essentially the same argument as in the proof of Lemma 4.3 and Lemma 5.14. For any path  $\gamma'$  satisfying the condition in the definition

of  $\mathcal{E}$ , let  $\gamma'_*$  denote a maximal subpath contained in the region  $\mathbb{B}_{k,0.96D}$ . Clearly,  $\ell(\gamma'_*) \geq 0.95D$  and the number of unspecified edges in  $\gamma'_* \geq \ell(\gamma'_*) - 0.7D$  (since the fraction of edges of  $\gamma'$  in  $N_d([x_{iD}, x_j]) \cup N_d([x_{j'}, x_{(i+1)D}])$  can be made arbitrarily small by choosing  $D \gg d$ ). For all  $h \geq 0.95D$ , and a  $\gamma'_*$  as above with  $\ell_{\gamma'_*} = h$ ; it follows by choosing  $\epsilon$  sufficiently small and using Theorem 2.12 as in the proof of Lemma 5.14, that the probability that the total weight of unspecified edges in  $\gamma'_*$  is  $\leq 2\epsilon D$  is bounded above by  $e^{-c(\epsilon)h}$  where  $c$  can be made arbitrarily small by making  $\epsilon$  sufficiently small. Taking a union bound over all  $\gamma'_*$  with  $\ell(\gamma'_*) = h$  (at most  $|S|^h$  in number where  $S$  is the generating set for  $\Gamma$ ) and then taking a union bound over all  $h \geq 0.95D$  the result follows.  $\square$

**Lemma 7.14.** *In the set-up of proof of Lemma 7.12, there exists  $d_0$  (independent of  $D$ ) such that for all  $d \geq d_0$  and all  $D$  sufficiently large we have  $\mathbb{P}(\mathcal{D} \cap \mathcal{D}') \geq \frac{1}{2}$ .*

*Proof.* Observe first that for  $D$  sufficiently large,  $\mathcal{D}$  and  $\mathcal{D}'$  depend on disjoint sets of edges (by Lemma 7.4), and hence are independent. So it suffices to show that  $\mathbb{P}(\mathcal{D}), \mathbb{P}(\mathcal{D}') \geq \frac{3}{4}$ . We shall only show this lower bound for  $\mathbb{P}(\mathcal{D})$ , the corresponding argument for  $\mathcal{D}'$  is identical. In fact we shall show that  $\mathbb{P}(\mathcal{D}^c) \leq \frac{1}{4}$ . For this, notice that in  $\mathcal{D}^c$  there must exist a geodesic from  $H_{iD}$  to  $H_j$  that does not intersect  $N_d([x_{iD}, x_j])$  where  $d$  is sufficiently large to be appropriately chosen later.

We adapt the notation from Lemma 7.6 as follows. We replace  $R$  in Lemma 7.6 by  $d$  here (as it will be determined later). For any  $x_p, x_l$ , the half-planes through them will be denoted by  $H_p, H_l$  as usual. For a path  $\gamma$  from  $H_p$  to  $H_l$ ,  $a(\gamma)$  (resp.  $b(\gamma)$ ) will denote the initial (resp. terminal) point on  $H_p$  (resp.  $H_l$ ). Further, suppose  $\gamma$  lies outside the  $d$ -neighborhood of  $[x_p, x_l]$ , i.e.  $\gamma \cap N_d([1, \xi]) = \emptyset$ . We shall use nearest-point projection  $\Pi$  onto  $[a(\gamma), x_p] \cup [x_p, x_l] \cup [x_l, b(\gamma)]$ . Let  $u = u(\gamma)$  be the last point on  $\gamma$  projecting to  $x_p$  and let  $v = v(\gamma)$  be the first point projecting to  $x_l$ . Let  $u_1 = u_1(\gamma) \in \gamma$  be the last point on  $\gamma$  before  $u$  such that  $d(u_1, [a(\gamma), b(\gamma)]) = d$ . Let  $a_1(\gamma) = \Pi(u_1)$ . Thus  $a_1(\gamma) \in [a(\gamma), x_p]$ . Similarly, let  $v_1 = v_1(\gamma) \in \gamma$  be the first point on  $\gamma$  after  $v$  such that  $d(v_1, [a(\gamma), b(\gamma)]) = d$ . Let  $b_1(\gamma) = \Pi(v_1)$ ; thus  $b_1 \in [x_l, b(\gamma)]$ .

We shall specialize now to  $k = iD$  and  $l = j (= (i + 0.01D))$  (as in Lemma 7.12). For  $j_1, j_2 \geq 0$ , let  $\mathcal{F}_{j_1, j_2}$  denote the event that a geodesic  $\gamma$  exists such that  $d(a_1(\gamma), x_{iD}) \in [j_1d, (j_1 + 1)d]$  and  $d(b_1(\gamma), x_j) \in [j_2d, (j_2 + 1)d]$ . Clearly,

$$\mathbb{P}(\mathcal{D}^c) \leq \sum_{j_1, j_2 \geq 0} \mathbb{P}(\mathcal{F}_{j_1, j_2}).$$

Let  $\mathcal{G}_{j_1, j_2}$  now denote the event that for all  $u \in (N_{(j_1+1)d}(x_{iD}) \setminus N_{j_1d}(x_{iD})) \cap N_d(H_{iD})$  and for all  $v \in (N_{(j_2+1)d}(x_j) \setminus N_{j_2d}(x_j)) \cap N_d(H_{iD})$ , we have  $\ell(\Upsilon(u, v)) \geq \frac{1}{2}e^{\alpha d}d(u, v)$  where  $\alpha$  is as in Lemma 7.6. Now for  $d$  sufficiently large and  $D$  sufficiently large depending on  $d$ , it follows that for all  $u, v$  as above we have  $d(u, v) \geq \frac{1}{2}(j_1 + j_2 + 2)R$  (by Lemma 7.4) and also  $\ell([a_1, x_{iD}] \cup [x_{iD}, x_{(i+1)D}] \cup [x_{(i+1)D}, b_1]) \geq \frac{1}{2}d(u, v)$  (again by Lemma 7.4). Using Lemma 7.6, it follows that for such choices of  $d$  and  $D$  we have  $\mathcal{F}_{j_1, j_2} \subseteq \mathcal{G}_{j_1, j_2}$  and hence  $\mathbb{P}(\mathcal{F}_{j_1, j_2}) \leq \mathbb{P}(\mathcal{G}_{j_1, j_2})$ . Observe now that the total number of pairs of  $u, v$  as above is bounded above by  $|S|^{(j_1+j_2+2)d}$  and hence using Lemma 4.3 we get  $\mathbb{P}(\mathcal{G}_{j_1, j_2}) \leq |S|^{(j_1+j_2)d} e^{-c(j_1+j_2+2)d}$ , and by choosing  $d$  (and hence  $e^{\alpha d}$ ) sufficiently large we can make the  $c$  above arbitrarily large. It

follows that for a suitable choice of  $d$  and  $D$  we get  $\mathbb{P}(\mathcal{G}_{j_1, j_2}) \leq e^{-c(j_1 + j_2 + 2)^d}$ . Summing over all  $j_1, j_2 \geq 0$  we get the desired upper bound on  $\mathbb{P}(\mathcal{D}^c)$ . This completes the proof of the lemma.  $\square$

**Case (iii): Support of  $\rho$  is noncompact:** Let us now turn to case (iii), where  $X(e)$  can take arbitrarily large values. The corresponding lemma in this case is the following.

**Lemma 7.15.** *Suppose the support of  $\rho$  is non-compact. Then there exists  $D = D(\Gamma, \rho)$  sufficiently large,  $\beta = \beta(D, R, \rho, \Gamma) > 0$  and an event  $B_i$  for each  $i$  depending only on the edges in  $N_{D/10}(\mathbb{B}_{iD, D})$  with  $\mathbb{P}(B_i) \geq \beta$  such that for all  $\omega \in B_i$  we have the following: for each  $u \in H_{iD}$  and  $v \in H_{(i+1)D}$ , and a path  $\gamma$  from  $u$  to  $v$  contained in the region  $\mathbb{B}_{iD, D}$ , there exists a path  $\gamma'$  between  $u$  and  $v$  that intersects  $[x_{iD}, x_{(i+1)D}]$  in at least  $\frac{3D}{5}$  edges and  $\ell_\omega(\gamma') \leq \ell_\omega(\gamma)$ .*

*Proof.* We shall use the same notations as in the proof of Lemma 7.12. Let  $d$  be such that the conclusion of Lemma 7.14 holds (notice that Lemma 7.14 did not use the condition that support of  $\rho$  contained 0), and let  $\mathcal{D}, \mathcal{D}'$  be as before. Let  $\mathcal{E}' = \mathcal{E}'(M)$  denote the event that for all edges in the region between  $H_k$  and  $H_{k'}$  (except those on  $[1, \xi]$ ) that are within a distance  $100d$  of  $[1, \xi]$ , we have  $X(e) \geq M$ . Let  $\mathcal{C}'$  denote the set of all  $\omega \in \Omega$  such that for all  $w_1 \in N_d([x_{iD}, x_j])$  and for all  $w_2 \in N_d([x_{j'}, x_{(i+1)D}])$  we have  $\ell_\omega([w_1, w'_1]) + \ell_\omega([w'_1, w'_2]) + \ell_\omega([w'_2, w_2]) \leq 2D\mathbb{E}X(e)$  (recall that  $w'_1$  and  $w'_2$  are projections of  $w_1$  and  $w_2$  respectively onto  $[x_{iD}, x_{(i+1)D}]$ ). Finally, let  $\mathcal{E}''$  denote the event that the total weight of unspecified edges (as defined in the proof of Lemma 7.12, although they are not all unspecified in the changed context) along any path from a point in  $N_d([x_{iD}, x_j])$  to a point in  $N_d([x_{j'}, x_{(i+1)D}])$  that intersects  $[x_{iD}, x_{(i+1)D}]$  in less than  $\frac{3D}{5}$  edges is at least  $2D\mathbb{E}X(e)$ .

Our claim now is that  $B_i := \mathcal{C}' \cap \mathcal{D} \cap \mathcal{D}' \cap \mathcal{E}' \cap \mathcal{E}''$  does our job. We first show that the appropriate conditions are satisfied for this event. Fix a path  $\gamma$  as in the statement of the lemma. As in the proof of Lemma 7.12, we shall assume without loss of generality that in  $\gamma$ , the last crossing of  $H_{iD}$  to  $H_j$  and the last crossing of  $H_{j'}$  to  $H_{(i+1)D}$  are the  $\omega$ -geodesics between the respective endpoints. As before, let  $w_1 \in \gamma \cap N_d([x_{iD}, x_j])$  and  $w_2 \in \gamma \cap N_d([x_{j'}, x_{(i+1)D}])$  and let  $w'_1$  and  $w'_2$  be their respective projections. By the definition of  $\mathcal{C}'$ , it suffices to show that the  $\omega$ -length of the segment of  $\gamma$  between  $w_1$  and  $w_2$  is at least  $2\mathbb{E}X(e)D$ , which is guaranteed by the definition of  $\mathcal{E}''$ .

Now to show the lower bound for  $\mathbb{P}(B_i)$ , observe first that  $\mathcal{C}' \cap \mathcal{D} \cap \mathcal{D}'$  and  $\mathcal{E}' \cap \mathcal{E}''$  depend on disjoint sets of edges for  $D$  sufficiently large (Lemma 7.3) and hence they are independent. So it suffices to bound from below the respective probabilities separately. For  $\mathcal{C}'$  observe that if  $D$  is sufficiently large, we have  $\ell([w_1, w'_1]) + \ell([w'_1, w'_2]) + \ell([w'_2, w_2]) \leq 1.5D$  hence using Theorem 2.13 we get that for each fixed  $w_1$  and  $w_2$  we have  $\ell_\omega([w_1, w'_1]) + \ell_\omega([w'_1, w'_2]) + \ell_\omega([w'_2, w_2]) \leq 2D\mathbb{E}X(e)$  with failure probability at most  $e^{-cD}$ . Taking a union bound over all pairs of  $(w_1, w_2)$  (these are polynomially many in  $D$  for a fixed  $d$ ) we get  $\mathbb{P}(\mathcal{C}') \geq \frac{3}{4}$ . As mentioned above, the proof of Lemma 7.14 remains valid in this set-up also and hence we get  $\mathbb{P}(\mathcal{C}' \cap \mathcal{D} \cap \mathcal{D}') \geq \frac{1}{4}$ .

As for  $\mathcal{E}' \cap \mathcal{E}''$ , observe that  $\mathbb{P}(\mathcal{E}') \geq \beta_2(M, d, D, \rho) > 0$  for each  $M$  by our hypothesis on  $\rho$ . Observe now that any path satisfying the condition in the definition of  $\mathcal{E}''$  must either have  $50d$  many edges that are set to weight at least  $M$  by  $\mathcal{E}'$  or

have at least  $\frac{1}{1000}e^{\alpha d}D$  many unspecified edges (provided  $d$  and  $D$  are sufficiently large) by Lemma 7.6. By choosing  $M \geq \frac{2D\mathbb{E}X(e)}{50d}$ , we can ensure that the condition in the definition of  $\mathcal{E}''$  is satisfied in the former case. In the latter case, we repeat the argument in the proof of Lemma 7.13. Observe that if  $d$  is chosen sufficiently large, then the proof of Lemma 7.13 implies that with probability at least  $\frac{1}{2}$ , every path  $\gamma$  satisfying the conditions in the definition of  $\mathcal{E}''$  with at least  $\frac{1}{1000}e^{\alpha d}D$  many unspecified edges will have  $\ell_\omega(\gamma) \geq 2D\mathbb{E}X(e)$ . Finally noticing that  $\mathcal{E}'$  and  $\mathcal{E}''$  are both increasing events in the weight configurations of the unspecified edges, we invoke the FKG inequality (Theorem 2.15) to conclude that  $\mathbb{P}(\mathcal{E}' \cap \mathcal{E}'') \geq \beta_3 > 0$ . This completes the proof of the lemma.  $\square$

We are now in a position to complete the proof of Proposition 7.10.

*Proof of Proposition 7.10.* Depending on the support of  $\rho$  consider the event  $B_i$  defined in Lemma 7.11, Lemma 7.12 or Lemma 7.15, and let  $D$  be sufficiently large so that the conclusion of those lemmas hold. The event  $B_i$  in each case depends on the configuration restricted to the edges in  $N_{D/10}(\mathbb{B}_{iD,D})$  and satisfies the required probability lower bound. Observe that, for all sufficiently large  $i$  and on  $B_i$ , every  $\omega$ -geodesic ray from either  $1$  or  $o_2$  (that must cross from  $H_{iD}$  to  $H_{(i+1)D}$ ) must intersect  $[x_{iD}, x_{(i+1)D}]$  in at least  $\frac{3D}{5}$  many edges. Clearly this implies that any two such geodesic ray must intersect on  $[x_{iD}, x_{(i+1)D}]$  completing the proof of the proposition.  $\square$

## 8. LINEAR GROWTH OF VARIANCE ALONG WORD GEODESICS

The aim of this section is to use the technology in the previous section to prove that under suitable conditions on the distribution  $\rho$  on edges, the variance of the length of the FPP geodesic between points on a semi-infinite geodesic ray  $[1, \xi)$  in a fixed direction  $\xi$  grows linearly with the word distance on  $\Gamma$ . Fix  $\xi \in \partial G$  and a geodesic ray  $[1, \xi) = \{x_0 = 1, x_1, x_2, \dots\}$  in  $\Gamma$ . The following is the main result of this section.

**Theorem 8.1.** *In the above set-up, there exists  $0 < C_1 < C_2 < \infty$  such that*

$$C_1 n \leq \text{Var}(T(1, x_n)) \leq C_2 n.$$

Because the FPP geodesic  $\Upsilon(1, x_n)$  is expected to remain close to the word geodesic (see Section 4), the linear growth of variance was conjectured by Benjamini, Tessler and Zeitouni [BZ12, Question 5], [BT17, Section 4]. Theorem 8.1 above proves their conjecture. As already pointed out in the introduction, in contrast, in the Euclidean setting, the variance is expected to grow sub-linearly in all dimensions with exponent strictly less than 1, but the best known result so far for a general FPP on  $\mathbb{Z}^d$  is an upper bound of  $O(\frac{n}{\log n})$  of the variance [BKS03].

Linear growth of variance (or the related behavior of diffusive fluctuations) has often been observed in constrained models of first passage percolation and its variants. Gaussian fluctuations or linear variance have been shown for first passage percolation across thin cylinders, and in certain one-dimensional graphs [CD13, Ahl15] (see also [DPJ, BB17] for similar results). Unlike these results, for FPP on hyperbolic groups, the FPP-geodesic is not restricted to a uniformly bounded neighborhood of the word geodesic (except for the special case when the support of  $\rho$  is bounded away from 0 and  $\infty$ ). So we cannot consider FPP geodesics restricted to lie in a

thin cylinder. The upper bound for the variance in Theorem 8.1 will follow from a standard Poincaré inequality. For the lower bound, [BT17] already speculated that results like Proposition 6.3 could be useful for showing that the variance grows linearly. We shall use for effective and quantitative version of such results, obtained in Section 7, and show that a linear number of vertices contributed a uniformly positive amount to the variance, thereby getting a lower bound of the matching order. This is philosophically similar to the proof of [BSSb, Proposition 7.2].

**Upper bound:** As mentioned above, the upper bound in Theorem 8.1 will use a standard argument due to Kesten [Kes93], who used it to obtain a linear upper bound on variance in Euclidean FPP. Kesten’s argument is rather general and the proof of (1.13) in [Kes93] shows the following in our set-up.

**Proposition 8.2.** *There exists  $C > 0$  such that  $\text{Var}(T(1, x_n)) \leq CE[\ell(\Upsilon(1, x_n))]$ .*

The upper bound in Theorem 8.1 follows immediately from Proposition 8.2 together with Corollary 4.4.

**Lower Bound:** Fix  $n \in \mathbb{N}$ . Consider the passage time  $T_n := T(1, x_n)$ . Let  $D = D(\Gamma, \rho)$  be such that the conclusion of Lemma 7.11, Lemma 7.12 or Lemma 7.15 holds depending on the support of  $\rho$ . Recall that  $\mathbb{B}_{i,k}$  denotes the region between the hyperplanes  $H_i$  and  $H_{i+k}$ . For  $i \geq 1$ , let  $\mathcal{G}_i$  denote the  $\sigma$ -algebra generated by the random variables  $\{X(e) : e \in \mathbb{B}_{(2i-1)D,D}\}$ . Note that when  $D$  is sufficiently large, these sets are mutually disjoint (as  $i$  varies). Let  $\mathcal{G}_*$  denote the  $\sigma$ -algebra generated by the remaining edge weights, i.e.,

$$\{X(e) : e \in \Gamma \setminus (\cup_{i=1}^{\infty} \mathbb{B}_{(2i-1)D,D})\}.$$

We define a filtration  $\{\mathcal{F}_i\}_{i \geq 0}$  by setting  $\mathcal{F}_0 = \emptyset$ ,  $\mathcal{F}_1 = \mathcal{G}_*$ , and for  $i \geq 2$ , set  $\mathcal{F}_i$  to be the  $\sigma$ -algebra generated by  $\cup_{j=2}^i \mathcal{G}_{i-1} \cup \mathcal{G}_*$ . Consider the Doob Martingale of  $T_n$  with respect to this filtration  $\{\mathcal{F}_i\}_{i \geq 0}$  given by

$$M_i = \mathbb{E}[T_n \mid \mathcal{F}_i].$$

It is a standard fact (see, e.g. [ADH17, Section 3.3]) using the orthogonality of martingale difference sequences that

$$(12) \quad \text{Var}(T_n) = \mathbb{E} \left[ \sum_{i=1}^{\infty} \text{Var}(M_i \mid \mathcal{F}_{i-1}) \right].$$

This type of decomposition of variance is a standard method of proving upper and lower bounds of the variance in models of first and last passage percolation. The lower bound in Theorem 8.1 is an immediate consequence of (12) together with the following lemma.

**Lemma 8.3.** *There exists  $c = c(\rho, \Gamma, D) > 0$  such that for each  $2 \leq i \leq \frac{n}{4D}$  we have  $\text{Var}(M_i \mid \mathcal{F}_{i-1}) \geq c$ .*

*Proof.* Let  $2 \leq i \leq \frac{n}{4D}$  be fixed. We shall divide the set of all edges in  $\Gamma$  into three parts. Let  $S = S_i$  denote the set of all edges whose weights generate the  $\sigma$ -algebra  $\mathcal{F}_{i-1}$ . Let  $V = V_i$  denote the set of all edges in  $\cup_{j \geq i} \mathbb{B}_{(2j-1)D,D}$ . Finally let  $U = U_i$  denote the set of all edges in  $\Gamma \setminus (S_i \cup V_i)$ , i.e., the set of all edges in  $\mathbb{B}_{(2i-3)D,D}$ . We shall write any  $\omega \in \Omega$  as  $\omega = (\omega_S, \omega_V, \omega_U)$  where  $\omega_S, \omega_V$ , and  $\omega_U$  are the restrictions of  $\omega$  to  $S_i, V_i$  and  $U_i$  respectively. Condition on  $\mathcal{F}_{i-1}$  (i.e., fix



$\omega_S$  for the rest of this proof). Using the standard fact that  $\text{Var} X \geq \frac{1}{4} \mathbb{E}[X - X']^2$  when  $X$  and  $X'$  have the same distribution and are defined on the same probability space we have

$$\text{Var}(M_i | \mathcal{F}_{i-1}) \geq \frac{1}{4} \int \left( \int (T_n(\omega_S, \omega_V, \omega_U) - T_n(\omega_S, \omega_V, \omega'_U) d\omega_V) \right)^2 d\omega_U d\omega'_U,$$

where  $\omega'_U$  is an copy of  $\omega_U$  (i.e., an assignment of  $\rho$ -distributed i.i.d. random weights on the edges  $U_i$ ) which need not be independent of  $\omega_U$ . In fact we shall consider an appropriate coupling of  $(\omega_U, \omega'_U)$  below. To see why the above equation is true, note that for a fixed  $\omega_S$ ,  $\int T_n(\omega_S, \omega_V, \omega_U) d\omega_V$  and  $\int T_n(\omega_S, \omega_V, \omega'_U) d\omega_V$  are two coupled copies of  $\mathbb{E}[T_n | \mathcal{F}_{i-1}]$ .

To prove the lemma, it now suffices to show that there exists a coupling of  $(\omega_U, \omega'_U)$  such that under this coupling there is a subset  $\mathcal{B}$  of configurations  $(\omega_U, \omega'_U)$  with  $\mathbb{P}(\mathcal{B}) \geq c_1 > 0$ , and on  $\mathcal{B}$

$$(13) \quad \int (T_n(\omega_S, \omega_V, \omega_U) - T_n(\omega_S, \omega_V, \omega'_U) d\omega_V \geq c_2 > 0.$$

Indeed, (13) will imply  $\text{Var}(M_i | \mathcal{F}_{i-1}) \geq \frac{c_1 c_2^2}{4}$ , and establish the desired result. We establish (13) in Lemma 8.4 below, invoking which finishes the proof of Lemma 8.3 and hence that of Theorem 8.1.  $\square$

Before establishing (13), let us briefly explain the basic idea. Roughly what we want to show is that once we have fixed  $\omega_S$  and  $\omega_V$  there is still enough randomness in the configuration  $\omega_U$  such that we can ensure that for two coupled copies  $(\omega_U, \omega'_U)$  with the same marginal we can ensure with uniformly positive probability that  $(T_n(\omega_S, \omega_V, \omega_U) - T_n(\omega_S, \omega_V, \omega'_U))$  is uniformly bounded away from 0. To show this we use the technology from Sec. 7 to ensure that with a positive probability  $\omega_U$  can be chosen in such a way that the geodesic  $\Upsilon(1, x_n)$  intersects the segment  $[x_{(2i-3)D}, x_{(2i-2)D}]$ . We then obtain  $\omega'_U$  from  $\omega_U$  by resampling the weights on  $[x_{(2i-3)D}, x_{(2i-2)D}]$  while keeping the other weights fixed. Now on the positive probability event that each changed co-ordinate in  $\omega'_U$  is strictly smaller than the corresponding co-ordinate in  $\omega$  we get the desired decrease in  $T_n(\omega_S, \omega_V, \omega'_U)$  compared to  $T_n(\omega_S, \omega_V, \omega_U)$ . We now make this argument rigorous in the following lemma.

**Lemma 8.4.** *In the set-up of proof of Lemma 8.3 we have the following: there exists a coupling  $(\omega_U, \omega'_U)$  with both marginals  $\rho^{\otimes U}$  such that under this coupling there is a subset  $\mathcal{B}$  of configurations  $(\omega_U, \omega'_U)$  with  $\mathbb{P}(\mathcal{B}) \geq c_1 > 0$ , and on  $\mathcal{B}$*

$$\int (T_n(\omega_S, \omega_V, \omega_U) - T_n(\omega_S, \omega_V, \omega'_U) d\omega_V \geq c_2 > 0.$$

*Proof.* Let  $I$  denote the set of edges on  $[x_{(2i-3)D}, x_{(2i-2)D}]$ , and let  $U^* = U_i \setminus I$ . Let  $\omega_{U^*}$  (resp.  $\omega_I$ ) denote the restriction of  $\omega_U$  to  $U^*$  (resp.  $I$ ). Let  $a \geq 0$  denote the essential infimum of  $\rho$ , i.e., the smallest point in the support of  $\rho$ . For  $\epsilon > 0$ , let  $\mathcal{A}(\epsilon)$  denote the set of all possible weight configurations  $\omega_{U^*}$  such that in the environment  $\omega = (\omega_S, \omega_V, \omega_{U^*}, \omega_I)$  the geodesic  $\Upsilon(1, x_n)$  passes through at least one edge in  $I$  for all choices of  $\omega_V$  provided  $\omega_e \leq a + \epsilon$  for all  $e \in I$ . If  $D$  is sufficiently large and  $\beta$  is as in Lemma 7.11, Lemma 7.12 or Lemma 7.15 (depending on the support of  $\rho$ ) we have

$$\mathbb{P}(\mathcal{A}(\epsilon)) + D\rho([a, a + \epsilon]) \geq \beta.$$

Indeed, notice that by Lemma 7.11, Lemma 7.12 or Lemma 7.15 we have with probability at least  $\beta$ ,  $\omega_U$  is such that every geodesic from  $H_{(2i-3)D}$  to  $H_{(2i-2)D}$  passes through at least one edge in  $I$ . Further, for each  $\omega_U$  for which the above happens we either have  $\omega_e < a + \epsilon$  for some  $e \in I$  or we have that the projection  $\omega_{U^*} \in \mathcal{A}(\epsilon)$ . Clearly, since  $\rho$  does not have an atom at  $a$ , it follows that by choosing  $\epsilon$  sufficiently small we get  $\mathbb{P}(\mathcal{A}(\epsilon)) \geq \frac{\beta}{2}$ . Since  $\rho$  is assumed to be continuous it follows that for each  $\epsilon > 0$  sufficiently small there exists  $0 < \epsilon' < \epsilon'' < \epsilon$  such that  $h(\epsilon, \epsilon', \epsilon'') = \min\{\rho([a, a + \epsilon']), \rho([a + \epsilon'', a + \epsilon])\} > 0$ .

Let us now choose the coupling  $(\omega_U, \omega'_{U'})$  as follows: Let  $\omega_{U^*} = \omega'_{U^*}$  and let  $\omega'_I$  denote an independent copy of  $\omega_I$ . Let  $\mathcal{B} = \mathcal{B}(\epsilon, \epsilon', \epsilon'')$  denote the set of all configurations  $(\omega_U, \omega'_{U'})$  such that  $\omega_{U^*} = \omega'_{U^*} \in \mathcal{A}(\epsilon)$ ,  $\omega_e \in [a + \epsilon'', a + \epsilon]$  for all  $e \in I$  and  $\omega'_e \in [a, a + \epsilon']$  for all  $e \in I$ .

Clearly, for appropriate choices of the parameters we have  $\mathbb{P}(\mathcal{B}(\epsilon, \epsilon', \epsilon'')) \geq \frac{\beta h(\epsilon, \epsilon', \epsilon'')^2}{2} = c_1 > 0$ . For  $(\omega_U, \omega'_{U'}) \in \mathcal{B}$ , a.e.,  $\omega_V$  (and the fixed choice of  $\omega_S$ ) we finally need to establish:

$$(T_n(\omega_S, \omega_V, \omega_U) - T_n(\omega_S, \omega_V, \omega'_{U'})) \geq \epsilon'' - \epsilon' = c_2 > 0.$$

Indeed, observe that, changing  $\omega_U$  to  $\omega'_{U'}$  decreases the length of any path by at least  $\ell c_2$  where  $\ell$  is the number of edges of  $I$  used by the path. By definition of  $\mathcal{B}$ , the geodesic must pass through at least one edge of  $I$  and hence we get the above equation, completing the proof of the lemma.  $\square$

## 9. DISCUSSION AND FUTURE DIRECTIONS

We have, in this paper, investigated some of the fundamental questions for first passage percolation on the Cayley graph of a hyperbolic group. In contrast to  $\mathbb{Z}^d$ , some of the results, e.g., the existence of velocity along almost every direction required much more work owing to the more complicated geometry of the underlying graph, whereas the hyperbolic geometry helped us resolve some other problems that are well-known to be difficult in Euclidean FPP (e.g. coalescence of geodesics). To maintain transparency of exposition, we have often worked with sub-optimal arguments and many of our results can possibly be strengthened. We finish with a discussion of some of these and a few of the many remaining open questions.

First of all, we have worked with a rather strong assumption on the passage time distribution  $\rho$ , which one should be able to relax to a large extent. It would be interesting to know the optimal conditions on  $\rho$  under which the results like Theorem 5.1 or Theorem 8.1 hold.

Observe also that Theorem 5.1 only gives a convergence in mean for first passage times along word geodesics to almost every boundary direction. Using Theorem 8.1 one can immediately upgrade this to an in probability convergence. Developing stronger concentration inequalities as in [Kes93], possibly requiring some additional conditions, would let one further upgrade this to an almost sure convergence.

After establishing the law of large number and order of fluctuations, the next natural question is to ask for the scaling limit for  $T(1, x_n)$  for  $x_n$  along a word geodesic ray. Comparing with results from [CD13, Ahl15], one would expect a positive answer to the following question which was conjectured also in [BZ12].

**Question 9.1** (Central Limit Theorem). Does  $\frac{T(1, x_n) - \mathbb{E}T(1, x_n)}{\sqrt{\text{Var}(T(1, x_n))}}$  converge weakly to a standard Gaussian variable?

More on the geometric side of things, by Theorem 5.1, existence and continuity of velocity  $v(\xi)$  on  $\partial G$  implies that  $v(\xi)$  is constant for all  $\xi \in \partial G$ , whereas we have also given examples to show that this need not be true in general. It is clear that if  $G$  is free of rank  $m$  and the generating sets are of the form  $\{a_i^{\pm j} | i = 1 \cdots m; j = 1, \cdots k\}$  for some  $k \geq 1$ , then  $v(\xi)$  exists and is constant for all  $\xi \in \partial G$ . Are there other examples? Towards this, we propose:

**Question 9.2.** Find conditions on a hyperbolic group  $G$  such that  $v(\xi)$  exists and is continuous (and is hence constant) on  $\partial G$ ?

What we have in mind here is the following: if  $B$ -neighborhoods of  $[1, \xi)$ ,  $\xi \in \partial G$ , are all nearly isometric (in a suitable sense), then  $v(\xi)$  is constant. Can one deduce restrictions on the geometry of  $\Gamma$  from continuity of velocity  $v(\xi)$ ?

#### APPENDIX A.

We provide in this section the postponed proofs of Theorem 2.14, Proposition 3.14 and Lemma 6.5.

*Proof of Theorem 2.14.* For  $j \geq 0$ , let us denote  $Y_{i,j} = (X_i - M)1_{\{X_i \in [2^j M, 2^{j+1} M]\}}$ . Let  $\epsilon > 0$  be fixed and let us choose  $L > 0$  sufficiently large. We shall show that by choosing  $M = M(\epsilon, L)$  sufficiently large we can ensure

$$(14) \quad \mathbb{P} \left( \sum_{i=1}^n Y_{i,j} \geq \epsilon 2^{-(j+1)} n \right) \leq e^{-Ln}$$

for  $j \leq \log n$ . Observe that by the sub-Gaussian tails of  $X_i$  we have

$$(15) \quad \mathbb{P} \left( \sum_{j \geq \log n} \sum_{i=1}^n Y_{i,j} \neq 0 \right) \leq n e^{-cn^2 M^2},$$

and hence together with (15), (14) completes the proof. Notice now that for a fixed  $j \leq \log n$ ,

$$\mathbb{P} \left( \sum_{i=1}^n Y_{i,j} \geq \epsilon 2^{-(j+1)} n \right) \leq \mathbb{P} \left( \sum_{i=1}^n 1(Y_{i,j} > 0) \geq \epsilon M^{-1} 4^{-(j+1)} n \right)$$

Observe that  $\sum_{i=1}^n 1(Y_{i,j} > 0)$  is a  $\text{Bin}(n, p_{j,M})$  variable where  $p_{j,M} = \mathbb{P}(X_i \in [2^j M, 2^{j+1} M]) \leq C_1 e^{-c_1 4^j M^2}$ . It therefore follows by a Chernoff inequality that for  $M$  sufficiently large the above probability is bounded above by

$$\exp \left( -\epsilon M^{-1} 4^{-(j+1)} n (c_1 4^j M^2 - \log C_1 + \log(\epsilon M^{-1} 4^{-(j+1)})) \right).$$

Denoting  $q_j := \epsilon M^{-1} 4^{-(j+1)}$  it follows that the above probability is bounded above by

$$\exp \left( -n \left( \frac{cM\epsilon}{8} - \log C_1 + q_j \log q_j \right) \right)$$

and the proof is completed by noting that  $x \log x$  is bounded away from  $-\infty$  for  $x \in [0, 1]$  and choosing  $M$  sufficiently large.  $\square$

**Proof of Proposition 3.14:** The proof of Proposition 3.14 below is fairly standard in the theory of finite Markov chains, but we provide it for completeness. For this we need to consider the vector valued Markov chain  $\{\mathbf{X}_n\}_{n \geq 1}$  obtained from the Markov chain  $\{X_i\}_{i \geq 1}$  where  $\mathbf{X}_i = (X_{(i-1)k+1}, \dots, X_{ik})$  for  $i \geq 1$ . Let  $P^{(k)}$  denote

the transition matrix of its chain. Also let us say an element  $\mathbf{x} = (x_1, x_2, \dots, x_k) \in \Sigma^k$  is admissible if  $P(x_i, x_{i+1}) > 0$  for all  $i = 1, 2, \dots, k-1$ . Notice that if a state  $\mathbf{x}$  is not admissible then almost surely the word  $\mathbf{x}$  never occurs in the trajectory of the chain  $\{X_i\}_{i \geq 1}$ . We first need the following lemma.

**Lemma A.1.** *If  $k$  is a multiple of  $d$ , then each of the recurrent components of  $P^{(k)}$  is aperiodic. Further, every admissible state in  $\Sigma^k$  belongs to a recurrent component of  $P^{(k)}$ .*

*Proof.* Let  $\mathbf{x} \in \Sigma^k$  denote an admissible state starting with some  $x_1 \in \Sigma^{(k)}$ . As  $k$  is a multiple of  $d$  and chain  $P$  is irreducible with period  $d$ , starting from  $\mathbf{x}$  any admissible state in  $\Sigma^k$  starting with  $x_i$  can be reached in  $P^{(k)}$  for all  $x_i$  in the  $d$ -periodic orbit of  $x_1$  in  $P$  (i.e., all  $x_i$  that can be reached from  $x_1$  in multiples of  $d$  steps). Clearly, these states form a recurrent component, and hence all admissible states belong to one such recurrent component. It now remains to show that the recurrent components are all aperiodic. For this, notice that since the chain  $P$  has period  $d$ , and  $k$  is a multiple of  $d$ , it follows that for all sufficiently large  $i$ ,

$$\mathbb{P}(X_{ik+1} = x_1 \mid (X_1, X_2, \dots, X_k) = \mathbf{x}) > 0.$$

This immediately shows that  $\mathbb{P}(\mathbf{X}_i = \mathbf{x} \mid \mathbf{X}_1 = \mathbf{x}) > 0$  for all  $i$  sufficiently large, thus showing that the component of  $\mathbf{x}$  is aperiodic, as desired.  $\square$

We can now complete the proof of Proposition 3.14.

*Proof of Proposition 3.14.* As  $k$  is a multiple of  $d$ , clearly if  $\mathbf{x}$  is not admissible or does not start with a state in the  $d$ -periodic orbit of  $x$  (in  $P$ ),  $N_n(\mathbf{x}, x) = 0$  and there is nothing to prove. So let us assume the contrary. Let  $\underline{\pi}$  denote the unique invariant measure of the irreducible aperiodic recurrent component of  $P^{(k)}$  containing  $\mathbf{x}$ . We shall show that the conclusion of the proposition holds with  $(\mathbf{x}, x) = \underline{\pi}(\mathbf{x})$  (notice that the reason to retain the dependence on  $x$  is that the starting state  $x$  uniquely determines which recurrent component a particular realization of the vector valued chain  $\{\mathbf{X}_n\}_{n \geq 1}$  belongs to). Let  $x$  and  $\mathbf{x}$  be as above and let  $\epsilon > 0$  be fixed. We shall show that

$$(16) \quad \limsup_{n \rightarrow \infty} \left| \frac{N_n(x_1, x)}{n} - \underline{\pi}(\mathbf{x}) \right| \leq \epsilon$$

almost surely, which gives the desired result.

It is a standard result from the theory of (not necessarily reversible) finite Markov chains that for every irreducible aperiodic finite Markov chain with transition matrix  $Q$  and stationary distribution  $\pi$  there exists  $\alpha = \alpha(Q) < 1$  such that for all  $x$  in the state space we have

$$(17) \quad \|Q^t(x, \cdot) - \pi(\cdot)\|_{TV} \leq \alpha^t,$$

where  $\|\cdot\|_{TV}$  denotes the total variation norm. Let us denote by  $Q$  the restriction of  $P^{(k)}$  to the recurrent component of  $\mathbf{x}$  and choose  $L$  sufficiently large so that  $L\alpha^L \leq \frac{\epsilon}{2}$ . Observe that

$$N_n(\mathbf{x}, x) = \sum_{i=1}^L N_{n,i}(\mathbf{x}, x)$$

where

$$N_{n,i}(\mathbf{x}, x) := \sum_{j=0}^{\frac{n}{L}-1} 1(\mathbf{X}_{jL+i} = \mathbf{x}).$$

Next, by (17), for each fixed  $i$ , and for  $j \geq 1$ , the indicator functions  $1(\mathbf{X}_{jL+i} = \mathbf{x})$  are stochastically dominated above and below by a family of i.i.d. Bernoulli random variables with parameters  $\pi(\mathbf{x}) + \frac{\epsilon}{2L}$  and  $\pi(\mathbf{x}) - \frac{\epsilon}{2L}$  respectively. A Chernoff inequality (Theorem 2.12) now yields that

$$\mathbb{P} \left( \left| \frac{1}{n} N_{n,i}(\mathbf{x}, x) - \frac{\pi(\mathbf{x})}{L} \right| \geq \frac{\epsilon}{L} \right) \leq e^{-c(\epsilon)n/L}.$$

Summing over all  $i$  and applying the Borel-Cantelli Lemma now yields (16) which completes the proof of the proposition.  $\square$

*Proof of Lemma 6.5.* Though  $\Pi$  is only coarsely Lipschitz and coarsely surjective with uniformly bounded constants depending only on  $\delta$ , we shall assume below that  $\Pi$  is, in fact, surjective and continuous for ease of exposition. Let  $w \in \sigma$  be such that  $\Pi(w) = o$ . Consider the  $R$ -neighborhood  $N_R([x, y])$  of  $[x, y]$ . Let  $u$  denote the last point on  $\sigma$  before  $w$  where  $\sigma$  exits  $N_R([x, y])$  so that  $d(u, \Pi(u)) = R$ . Similarly, let  $v$  be the first point after  $w$  where  $\sigma$  re-enters  $N_R([x, y])$  so that  $d(v, \Pi(v)) = R$ . Since  $\sigma \cap N_{100R}(o) = \emptyset$ , it follows from the triangle inequality that  $\Pi(u), \Pi(v) \in [x, y]$  lie on opposite sides of  $o$  with  $d(\Pi(u), o) \geq 99R$ ,  $d(\Pi(v), o) \geq 99R$ .

We shall show that  $u, v$  satisfy the properties required by the Lemma. For  $R$  large enough  $[u, \Pi(u)] \cup [\Pi(u), \Pi(v)] \cup [\Pi(v) \cup v]$  (concatenated in this order) forms a  $(1, 4\delta)$ -quasigeodesic (see [BH99, Chapter III.H.1] or [Mit98, Lemma 3.1] or Lemma 7.4 below). Hence there exists  $A > 0$  (depending only on  $\delta$ ) such that  $[u, v] \cap N_A(o) \neq \emptyset$ .

Using the standard fact that paths leaving large neighborhoods of a geodesic are exponentially inefficient in hyperbolic space ([BH99, Chapter III.H.1] or Lemma 7.4) we have that there exists  $R_0 > 0, \alpha > 0$  depending only on the hyperbolicity constant  $\delta$  of  $\Gamma$ , such that  $\ell(\sigma_{uv}) \geq d(\Pi(u), \Pi(v))e^{\alpha R}$  for  $R \geq R_0$ . Item (1) of the Lemma follows immediately.

Next,  $d(o, \sigma_{uv}) \leq R + d(\Pi(u), \Pi(v))$ . Choosing  $R_0$  large enough, it follows from the previous paragraph that there exists  $C'$  depending only on  $\delta$  such that for  $R \geq R_0$ ,  $d(o, \sigma_{uv}) \leq C'\ell(\sigma_{uv})$ , proving Item (2) of the Lemma.  $\square$

#### ACKNOWLEDGMENTS

We thank Danny Calegari for useful email correspondence, particularly for explaining to us one of the key points of [CF10]. RB was partially supported by an ICTS-Simons Junior Faculty Fellowship and a Ramanujan Fellowship from Govt. of India. MM was partly supported by a DST JC Bose Fellowship, Matrics research project grant MTR/2017/000005 and CEFIPRA project No. 5801-1. MM also acknowledges partial support from the grant 346300 for IMPAN from the Simons Foundation and the matching 2015-2019 Polish MNiSW fund. This paper began at the International Centre for Theoretical Sciences (ICTS) during the program - Probabilistic Methods in Negative Curvature (Code: ICTS/PMNC2019/03) and was completed during a visit to Indian Institute of Science, Bengaluru. We gratefully acknowledge the hospitality of these institutions.

## REFERENCES

- [ADH17] Antonio Auffinger, Michael Damron, and Jack Hanson. *50 years of first-passage percolation*, volume 68 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2017.
- [AH] D. Ahlberg and C. Hoffman. Random coalescing geodesics in first-passage percolation. Preprint, arXiv:1609.02447.
- [Ahl15] Daniel Ahlberg. Asymptotics of first-passage percolation on one-dimensional graphs. *Adv. in Appl. Probab.*, 47(1):182–209, 03 2015.
- [BB17] Riddhipratim Basu and Nayantara Bhatnagar. Limit theorems for longest monotone subsequences in random mallows permutations. *Ann. Inst. H. Poincaré Probab. Statist.*, 53(4):1934–1951, 11 2017.
- [BH99] Martin R. Bridson and André Haefliger. *Metric spaces of non-positive curvature*, volume 319 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.
- [BHS] Riddhipratim Basu, Christopher Hoffman, and Allan Sly. Non-existence of bigeodesics in exactly solvable models of last passage percolation. Preprint, arXiv:1811.04908.
- [BKS03] Itai Benjamini, Gil Kalai, and Oded Schramm. First passage percolation has sublinear distance variance. *Ann. Probab.*, 31(4):1970–1978, 2003.
- [Bro81] Robert Brooks. Some remarks on bounded cohomology. In *Riemann surfaces and related topics: Proceedings of the 1978 Stony Brook Conference (State Univ. New York, Stony Brook, N.Y., 1978)*, volume 97 of *Ann. of Math. Stud.*, pages 53–63. Princeton Univ. Press, Princeton, N.J., 1981.
- [BSSa] Riddhipratim Basu, Sourav Sarkar, and Allan Sly. Coalescence of geodesics in exactly solvable models of last passage percolation. Preprint, arXiv:1704.05219.
- [BSSb] Riddhipratim Basu, Sourav Sarkar, and Allan Sly. Invariant measures for TASEP with a slow bond. Preprint arXiv 1704.07799.
- [BT17] Itai Benjamini and Romain Tessera. First passage percolation on a hyperbolic graph admits bi-infinite geodesics. *Electron. Commun. Probab.*, 22:Paper No. 14, 8, 2017.
- [BZ12] Itai Benjamini and Ofer Zeitouni. Tightness of fluctuations of first passage percolation on some large graphs. In *Geometric aspects of functional analysis*, volume 2050 of *Lecture Notes in Math.*, pages 127–132. Springer, Heidelberg, 2012.
- [Cal13] Danny Calegari. The ergodic theory of hyperbolic groups. In *Geometry and topology down under*, volume 597 of *Contemp. Math.*, pages 15–52. Amer. Math. Soc., Providence, RI, 2013.
- [Can84] James W. Cannon. The combinatorial structure of cocompact discrete hyperbolic groups. *Geom. Dedicata*, 16(2):123–148, 1984.
- [Can91] James W. Cannon. The theory of negatively curved spaces and groups. In *Ergodic theory, symbolic dynamics, and hyperbolic spaces (Trieste, 1989)*, Oxford Sci. Publ., pages 315–369. Oxford Univ. Press, New York, 1991.
- [CD81] J. Theodore Cox and Richard Durrett. Some limit theorems for percolation processes with necessary and sufficient conditions. *Ann. Probab.*, 9(4):583–603, 08 1981.
- [CD13] Sourav Chatterjee and Partha S. Dey. Central limit theorem for first-passage percolation time across thin cylinders. *Probability Theory and Related Fields*, 156(3):613–663, 2013.
- [CDA90] M. Coornaert, T. Delzant, and A. Papadopoulos. Geometrie et theorie des groupes. *Lecture Notes in Math., vol. 1441, Springer Verlag*, 1990.
- [CF10] Danny Calegari and Koji Fujiwara. Combable functions, quasimorphisms, and the central limit theorem. *Ergodic Theory Dynam. Systems*, 30(5):1343–1369, 2010.
- [CM15] Danny Calegari and Joseph Maher. Statistics and compression of scl. *Ergodic Theory Dynam. Systems*, 35(1):64–110, 2015.
- [Coo93] M. Coornaert. Mesures de Patterson-Sullivan sur le bord d'un espace hyperbolique au sens de Gromov. *Pacific J. Math.* 159, no. 2, pages 241–270, 1993.
- [Cou11] David Coupier. Multiple geodesics with the same direction. *Electronic Communications in Probability*, 16(46):517–527, 2011.
- [CP93] Michel Coornaert and Athanase Papadopoulos. *Symbolic dynamics and hyperbolic groups*, volume 1539 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1993.

- [DH14] Michael Damron and Jack Hanson. Busemann functions and infinite geodesics in two-dimensional first-passage percolation. *Communications in Mathematical Physics*, 325(3):917–963, 2014.
- [DH17] Michael Damron and Jack Hanson. Bigeodesics in first-passage percolation. *Communications in Mathematical Physics*, 349(2):753–776, 2017.
- [DPJ] Partha S. Dey, Ron Peled, and Matthew Joseph. Longest increasing path within the critical strip. Preprint, arXiv: 1808.08407.
- [FP05] Pablo A Ferrari and Leandro PR Pimentel. Competition interfaces and second class particles. *Annals of probability*, pages 1235–1254, 2005.
- [GdlH90] E. Ghys and P. de la Harpe(eds.). Sur les groupes hyperboliques d’après Mikhael Gromov. *Progress in Math. vol 83, Birkhauser, Boston Ma.*, 1990.
- [Gro85] M. Gromov. Hyperbolic Groups. in *Essays in Group Theory, ed. Gersten, MSRI Publ.,vol.8, Springer Verlag*, pages 75–263, 1985.
- [HM95] Olle Häggström and Ronald Meester. Asymptotic shapes for stationary first passage percolation. *Ann. Probab.*, 23(4):1511–1522, 1995.
- [Hof08] Christopher Hoffman. Geodesics in first passage percolation. *The Annals of Applied Probability*, 18(5):1944–1969, 2008.
- [HW65] J. M. Hammersley and D. J. A. Welsh. First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory. In *Proc. Internat. Res. Semin., Statist. Lab., Univ. California, Berkeley, Calif*, pages 61–110. Springer-Verlag, New York, 1965.
- [Kes86] Harry Kesten. Aspects of first passage percolation. In *École d’été de probabilités de Saint-Flour, XIV—1984*, volume 1180 of *Lecture Notes in Math.*, pages 125–264. Springer, Berlin, 1986.
- [Kes93] Harry Kesten. On the speed of convergence in first-passage percolation. *Ann. Appl. Probab.*, 3(2):296–338, 1993.
- [Kes03] Harry Kesten. *From Classical to Modern Probability: CIMPA Summer School 2001*, chapter First-Passage Percolation, pages 93–143. Birkhäuser Basel, 2003.
- [LN96] Cristina Licea and Charles M. Newman. Geodesics in two-dimensional first-passage percolation. *Ann. Probab.*, 24(1):399–410, 01 1996.
- [LP17] David A. Levin and Yuval Peres. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2017. Second edition of [ MR2466937], With contributions by Elizabeth L. Wilmer, With a chapter on “Coupling from the past” by James G. Propp and David B. Wilson.
- [Mit97] M. Mitra. Ending Laminations for Hyperbolic Group Extensions. *Geom. Funct. Anal. vol.7 No. 2*, pages 379–402, 1997.
- [Mit98] M. Mitra. Cannon-Thurston Maps for Trees of Hyperbolic Metric Spaces. *Jour. Diff. Geom.* 48, pages 135–164, 1998.
- [Mj14] M. Mj. Cannon-Thurston Maps for Surface Groups. *Ann. of Math.*, 179(1), pages 1–80, 2014.
- [New95] Charles M Newman. A surface view of first-passage percolation. In *Proceedings of the International Congress of Mathematicians*, pages 1017–1023. Springer, 1995.
- [Pim16] Leandro P. R. Pimentel. Duality between coalescence times and exit points in last-passage percolation models. *Ann. Probab.*, 44(5):3187–3206, 2016.
- [Rhe68] A. H. Rhemtulla. A problem of bounded expressibility in free products. *Proc. Cambridge Philos. Soc.*, 64:573–584, 1968.
- [Ver18] Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2018.

RIDDHIPRATIM BASU, INTERNATIONAL CENTRE FOR THEORETICAL SCIENCES, TATA INSTITUTE OF FUNDAMENTAL RESEARCH, BENGALURU, INDIA  
*E-mail address:* `rbasu@icts.res.in`

MAHAN MJ, SCHOOL OF MATHEMATICS, TATA INSTITUTE OF FUNDAMENTAL RESEARCH. 1, HOMI BHABHA ROAD, MUMBAI-400005, INDIA  
*E-mail address:* `mahan@math.tifr.res.in`