

COUNTING LOXODROMICS FOR HYPERBOLIC ACTIONS

ILYA GEKHTMAN, SAMUEL J. TAYLOR, AND GIULIO TIOZZO

ABSTRACT. Let $G \curvearrowright X$ be a nonelementary action by isometries of a hyperbolic group G on a (not necessarily proper) hyperbolic metric space X . We show that the set of elements of G which act as loxodromic isometries of X has density one in the word metric on G . That is, for any finite generating set of G , the proportion of elements in G of word length at most n which are X -loxodromics approaches 1 as $n \rightarrow \infty$. We also establish several results about the behavior in X of the images of typical geodesic rays in G ; for example, we prove that they make linear progress in X and converge to the Gromov boundary ∂X . We discuss various applications, in particular to mapping class groups, $\text{Out}(F_N)$, and right-angled Artin groups.

1. INTRODUCTION

Let $G \curvearrowright X$ be a nonelementary action of a finitely generated group G on a hyperbolic metric space X . At this level of generality, a complete understanding of the dynamical properties of each element of G may be out of reach, and so it is natural to ask: *How does a typical element of G act on X ?*

The meaning of *typical* depends crucially how we choose to model a “random” element. In the literature, there are mainly two models.

- (1) The *counting model*. Let B_n be the ball of radius n with respect to some G -invariant metric on the group. Then we call a property P *generic* if

$$\frac{\#\{g \in B_n : g \text{ has } P\}}{\#B_n} \rightarrow 1 \quad \text{as } n \rightarrow \infty.$$

- (2) The *random walk model*. Let μ be a probability measure on G , and consider a sequence (g_n) of random, i.i.d. elements of G with distribution μ . Then the n^{th} step of the random walk is defined as the product

$$w_n := g_1 g_2 \dots g_n$$

Then we call a property *typical for a random walk* if the probability

$$\mathbb{P}(w_n \text{ has } P) \rightarrow 1 \quad \text{as } n \rightarrow \infty,$$

where $\mathbb{P} = \mu^{\mathbb{N}}$ is the product measure on the space of increments.

Both types of counting are related to sequences of measures on the group: namely, in the counting model, for each n one considers the uniform measure on the ball of radius n :

$$\nu_n := \frac{1}{\#B_n} \sum_{g \in B_n} \delta_g$$

and its limit in a suitable compactification (when it exists) is usually called a *Patterson-Sullivan (PS) measure*. In the second case, for each n one has the convolution measure $\mu_n := \underbrace{\mu \star \dots \star \mu}_{n \text{ times}}$,

which is precisely the distribution of the n^{th} step of the random walk. In this case, the limit measure in a suitable compactification is the *hitting measure* for the random walk.

The two types of counting, however, need not yield the same notion of genericity; indeed, an underlying theme of much research has been the question: *Are Patterson-Sullivan measures also hitting measures for a certain random walk?*

This is for instance the main question of [Ver00]. As Furstenberg showed [Fur71], for a semisimple Lie group G , the appropriate boundary is $B = G/P$, where P is a minimal parabolic subgroup. Furthermore, the analogue of the Patterson-Sullivan measure is the unique measure ν on B which is invariant by a maximal compact subgroup. In fact, he proved that for any lattice Λ in G there exists a measure on Λ whose hitting measure is ν . Moreover, for groups acting properly and cocompactly on δ -hyperbolic manifolds, Connell and Muchnik [CM07b] showed that there exists a certain random walk which produces the Patterson-Sullivan measure on the boundary. They also extended this result to arbitrary Gibbs measures in the case the space is CAT(-1) [CM07a].

However, for actions of hyperbolic groups G on spaces X , if the harmonic measure and PS measure are in the same measure class, then the metrics on G and on X are quasi-isometric [BHM11]. Interestingly, for hyperbolic groups Gouëzel, Mathéus, and Maucourant [GMM15] recently proved that, when considering a word metric on G , harmonic measures for symmetric random walks of finite support are never in the same measure class as Patterson-Sullivan measures, unless the group is virtually free. One should compare this to Ledrappier's result that for compact hyperbolic surfaces the hitting measure for Brownian motion coincides with the PS measure if and only if the surface has constant curvature [Led95]. There are very many examples of harmonic measures which are singular with respect to the PS-type measure, see e.g. [KLP11].

In this paper, we will use random walk methods to prove results about genericity with respect to counting in balls in the *word metric* on G . In particular, we consider the case of G a word hyperbolic group, acting on a δ -hyperbolic metric space X and describe the behavior of typical geodesics in G with respect to the corresponding Patterson-Sullivan measure. Note that we do not assume that the space X is proper, hence the results hold for actions on infinite-dimensional manifolds or locally infinite graphs.

Just as in the case of isometries of the hyperbolic plane, an isometry g of a δ -hyperbolic space can be either *loxodromic* (also called *hyperbolic*), *elliptic*, or *parabolic*. In particular, it is called loxodromic if its *translation length* $\tau(g) := \lim_{n \rightarrow \infty} \frac{d(g^n x, x)}{n}$ is positive for some $x \in X$. Our first theorem states that loxodromic isometries are generic with respect to counting in the word metric:

Theorem 1.1 (Genericity of loxodromics). *Let G be a hyperbolic group with a nonelementary action by isometries on a separable¹, δ -hyperbolic metric space X . Then X -loxodromics are generic on G , i.e.*

$$\frac{\#\{g \in B_n : g \text{ is } X\text{-loxodromic}\}}{\#B_n} \rightarrow 1,$$

as $n \rightarrow \infty$.

Recall that two loxodromic isometries f and g of X are *independent* if their fixed point sets in ∂X are disjoint, and the action $G \curvearrowright X$ is *nonelementary* if G contains two independent loxodromic isometries of X ². We actually prove the stronger result that X -loxodromics are generic with respect to counting in spheres S_n in the Cayley graph of G .

The corresponding result for the random walk model has been proven by Rivin [Riv08] and Maher [Mah11] for the action of the mapping class group on the curve complex. In that case, an element is loxodromic if and only if it is pseudo-Anosov. This generalizes earlier work [Kai94] for proper actions, and led to several generalizations including [CM15, Sis11]. In particular, Maher-Tiozzo

¹In our work, we assume that all spaces are separable in order to apply results from [MT14]. However, Gruber-Sisto-Tessera [GST17, Remark 4] observe that in the current setting one may drop this hypothesis by replacing X with a separable metric space quasi-isometric to it.

²Actions satisfying this condition are sometimes called *of general type* [CdCM⁺15].

[MT14] proved that for *any* nonelementary group of isometries of a δ -hyperbolic space, loxodromic elements are typical along random walks. The interest in random walks on mapping class groups goes back to Dunfield-Thurston [DT06], where they are used to discuss the virtual Haken conjecture for random 3-manifolds.

It is important to note that for the counting model, genericity of loxodromics is *false* for general groups of isometries of a δ -hyperbolic space. This is shown in the following example (suggested to us by Sam Kim).

Example 1 (Nongenericity in general). Denote the free group of rank r by F_r and fix a free basis as a generating set. Let $G = F_2 \times F_3$ with the product generating set and let X denote a Cayley graph for F_2 . Consider the action $G \curvearrowright X$ in which the F_2 factor acts by left multiplication and the right factor acts trivially. If we denote the set of loxodromics for the action by LOX , then its complement is the set

$$\text{LOX}^c = \{(1, y) : y \in F_3\}$$

and a simple computation³ yields

$$\lim_{n \rightarrow \infty} \frac{\#(\text{LOX} \cap B_n)}{\#B_n} = \frac{2}{3} \neq 1.$$

This is in contrast to the fact that for any simple random walk of G , the probability that the n^{th} step of the random walk is loxodromic goes to 1 as $n \rightarrow \infty$. Note that this G is (geodesically) automatic and has pure exponential growth, so these hypotheses are still not sufficient.

Returning to the situation of a hyperbolic group G , our results demonstrate that much of the typical geometry of G is preserved under the action $G \curvearrowright X$. Let us fix a finite generating set S for G and denote the corresponding Patterson–Sullivan measure on ∂G by ν . Intuitively, this measure is defined by taking weak limits of the uniform distribution on balls B_n in the word metric as $n \rightarrow \infty$ (with the previous definition, it is only defined up to a multiplicative constant, but we actually fix a normalization: see Section 5.1).

One consequence of our study is that (a particular normalization of) the Patterson–Sullivan measure ν on ∂G decomposes as a countable sum of harmonic measures associated to random walks on G , in the following sense.

Proposition 1.2. *There exists a finite collection of measures ν_1, \dots, ν_r on ∂G , which are harmonic measures for random walks on G with finite exponential moment, and such that the Patterson–Sullivan measure ν can be written as*

$$\nu = \sum_{g \in G} a_g g_* \nu_{i(g)},$$

where each a_g is a non-negative, real coefficient, and $i(g) \in \{1, \dots, r\}$.

³Letting $S_n(H)$ denote the sphere of radius n in the Cayley graph for the group H , we have for $n \geq 1$

$$\#(S_n(F_2 \times F_3) \cap \text{LOX}^c) = \#S_n(F_3) = 6(5^{n-1}).$$

Since

$$\#S_n(F_2 \times F_3) = \sum_{k=0}^n \#S_k(F_2) \#S_{n-k}(F_3) = 6(5^{n-1}) + \sum_{k=1}^{n-1} 4(3^{k-1}) \cdot 6(5^{n-k-1}) + 4(3^{n-1})$$

we get by summing over all the S_n

$$\lim_{n \rightarrow \infty} \frac{\#(\text{LOX} \cap B_n)}{\#B_n} = \frac{2}{3} \neq 1.$$

One should compare Proposition 1.2 with the main theorem of Connell–Muchnik [CM07b], who show that for certain actions of hyperbolic groups, the Patterson–Sullivan measure is actually the harmonic measure for some random walk on G . We note that their result does not apply in our setting (the Gromov boundary of a hyperbolic group) since the Patterson–Sullivan measure ν does not necessarily have the property that the Radon–Nikodym derivative of $g_*\nu$ with respect to ν is continuous. See also Remark 2.15 of [GMM15].

There is a substantial literature studying typical behavior in finitely generated groups, beginning with Gromov’s influential works [Gro87, Gro93, Gro03]. More recent developments can be found, for example, in [AO96, Arz98, BMR03, Cha95, KMSS03, KS05, KRSS07, Ol’92].

Despite this interest, few works have addressed counting problems with respect to the *word metric* on G , when G is not necessarily free. For example, Pollicott and Sharp [PS98] compare word length in a cocompact lattice of $\text{Isom}(\mathbb{H}^n)$ to distance between orbit points in \mathbb{H}^n . In a different vein, Calegari and Fujiwara [CF10] study the generic behavior of a bicomposable function on a hyperbolic group and, in particular, establish a Central Limit Theorem for such functions. Finally, Wiest [Wie14] recently showed that if a group G satisfies a weak automaticity condition and the action $G \curvearrowright X$ on a hyperbolic space X satisfies a strong “geodesic word hypothesis,” then the loxodromics make up a *definite proportion* of elements of the n ball (when counting with respect to certain normal forms). This geodesic word hypothesis essentially requires geodesics in the group G , given by the normal forms, to project to unparameterized quasigeodesics in the space X under the orbit map. In our work, on the other hand, we do not assume any nice property of the action except it being by isometries.

1.1. Typical geodesic rays in G . To establish Theorem 1.1, we study the behavior of ν -typical geodesic rays of G and the paths they determine in the space X under the orbit map $G \rightarrow X$. That the images of geodesics in G have any controlled behavior when projected to X is particularly surprising at this level of generality. A few examples to bear in mind are G acting on a locally infinite hyperbolic graph, G acting on one of its nonelementary hyperbolic quotient groups, or G acting on \mathbb{H}^n with dense orbits.

We first show that typical geodesic rays of G converge to the boundary in $X \cup \partial X$:

Theorem 1.3 (Convergence to the boundary of X). *Let G be a hyperbolic group with a nonelementary action by isometries on a separable, hyperbolic metric space X . For every $x \in X$ and ν -almost every $\eta \in \partial G$, if $(g_n)_{n \geq 0}$ is a geodesic ray in G converging to η , then the sequence $g_n x$ in X converges to a point in the boundary ∂X .*

For $x \in X$, let $\Phi = \Phi_x : G \rightarrow X$ given by $\Phi(g) = gx$ denote the orbit map of the action $G \curvearrowright X$. Theorem 1.3 implies the existence of a ν -measurable, G -equivariant map $\partial\Phi : \partial G \rightarrow \partial X$ from the Gromov boundary of G to the Gromov boundary of X .

In fact, if we define

$$\partial^X G = \{\eta \in \partial G : \Phi([1, \eta)) \subset X \text{ converges to a point in } \partial X\},$$

then Φ extends to a map $\partial\Phi : \partial^X G \rightarrow \partial X$ and the set $\partial^X G$ has full ν -measure by Theorem 1.3.

In addition, we show that for almost every $\eta \in \partial G$, the path $\Phi([1, \eta))$ makes linear progress in the space X :

Theorem 1.4 (Positive drift). *Let G be a hyperbolic group with a nonelementary action by isometries on a separable, hyperbolic metric space X . Then there exists $L > 0$ such that for every $x \in X$ and ν -almost every $\eta \in \partial G$, if $(g_n)_{n \geq 0}$ is a geodesic in G converging to η , then*

$$\lim_{n \rightarrow \infty} \frac{d_X(x, g_n x)}{n} = L > 0.$$

Finally, we prove that the projection to X of a generic geodesic ray in G is approximated by a quasigeodesic ray in X , up to sublinear error.

Theorem 1.5 (Sublinear tracking). *Let G be a hyperbolic group with a nonelementary action by isometries on a separable, hyperbolic metric space X . Then, for ν -almost every boundary point $\eta \in \partial G$, there exists a quasigeodesic ray γ in X such that, for any geodesic sequence $g_n \rightarrow \eta$, one has*

$$\lim_{n \rightarrow \infty} \frac{d_X(g_n x, \gamma)}{n} = 0.$$

We note that if X has the property that each point in ∂X is the endpoint of a geodesic ray, then the quasigeodesic γ may be taken to be a geodesic.

Theorems 1.3, 1.4, and 1.5 are proven using a Markov chain (introduced by Calegari–Fujiwara [CF10]) on the directed graph Γ which parameterizes a geodesic combing of G . That is, directed paths in Γ evaluate to geodesics in the group G . Let \mathbb{P} be the corresponding Markov measure on the space of sample paths (w_n) which begin at the “identity” vertex of Γ . (See Section 6 for details.) By combining Theorem 1.3 and Theorem 1.4 we show that along a \mathbb{P} -typical path in Γ one encounters elements of G which are loxodromic for the action on X and whose translation lengths grow linearly.

Theorem 1.6 (Linear growth of X -translation lengths). *Let G be a hyperbolic group with a nonelementary action by isometries on a separable hyperbolic metric space X . Then there is an $L_1 > 0$ such that*

$$\mathbb{P}(\tau_X(w_n) \geq L_1 n) \rightarrow 1,$$

as $n \rightarrow \infty$.

In fact, we show that the constant L_1 appearing in Theorem 1.6 can be taken to be $L - \epsilon$ for any $\epsilon > 0$, where L is the drift constant of Theorem 1.4.

1.2. Genericity in G and the action $G \curvearrowright X$. Finally, we translate results about typical geodesics in the Patterson–Sullivan measure into statements about counting in the balls $B_n \subset G$. We emphasize that our genericity results also hold for counting in spheres S_n (i.e. elements of word length exactly n) which gives a priori stronger information.

First, we show that for a fixed $x \in X$, the ratio of the displacement $d_X(x, gx)$ to the word length of g is uniformly bounded from below for a generic subset of G .

Theorem 1.7 (Genericity of positive drift). *Let G be a hyperbolic group with a nonelementary action by isometries on a separable hyperbolic metric space X . Then there exists $L_1 > 0$ such that*

$$\frac{\#\{g \in B_n : d_X(gx, x) \geq L_1 |g|\}}{\#B_n} \rightarrow 1 \quad \text{as } n \rightarrow \infty.$$

By combining Theorem 1.7 with estimates for the Gromov product between generic elements we show that, generically, the translation length of g with respect to the action $G \curvearrowright X$ grows linearly with $|g|$.

Theorem 1.8 (Genericity of linear growth). *Let G be a hyperbolic group with a nonelementary action by isometries on a separable hyperbolic metric space X . Then there is an $L_1 > 0$ so that*

$$\frac{\#\{g \in B_n : \tau_X(g) \geq L_1 |g|\}}{\#B_n} \rightarrow 1,$$

as $n \rightarrow \infty$.

Just as above, the constant L_1 appearing in Theorem 1.7 and Theorem 1.8 can be taken to be $L - \epsilon$ for any $\epsilon > 0$, where L is the drift constant of Theorem 1.4. Since $\tau_X(g) > 0$ if and only if g is loxodromic for the action $G \curvearrowright X$, Theorem 1.8 immediately implies that X -loxodromics are generic in G , proving Theorem 1.1.

We remark that it is an interesting problem to establish the rate of convergence in Theorem 1.7 and Theorem 1.8. Work of Mathieu–Sisto [MS14] and Sunderland [Sun] obtains an exponential decay on drift for the random walk, which may be useful for proving an exponential convergence rate in our setting.

1.3. Applications. We now collect some immediate applications of our main results.

1.3.1. Splittings and quotients of hyperbolic groups. Applying Theorem 1.8 directly to the action $G \curvearrowright C_S(G)$ of a hyperbolic group on its Cayley graph, we see that generic elements of G are infinite order and their translation lengths grow linearly in word length. This recovers a result of Pallavi Dani from her thesis [Dan05]. The main point of our results is that they apply to a much more general setting.

For example, recall that a splitting of a group G is an action of G on a simplicial tree T – this is equivalent to realizing G as the fundamental group of a graph of groups [Ser03]. The splitting is minimal if there is no invariant subtree. For $g \in G$, we denote by g^* the shortest representative of the conjugacy class of g in T/G .

Proposition 1.9. *Suppose that $G \curvearrowright T$ is a minimal splitting of a hyperbolic group G such that T has at least 3 ends. Then the set of elements which are not conjugate into a vertex stabilizer is generic.*

Moreover, there is an $L > 0$ such that the set of $g \in G$ having the property that

$$\#(\text{edges crossed by } g^*) \geq L|g|$$

is generic in G .

Similar statements can be made about other useful actions of G on hyperbolic spaces. These include the quasi-trees of Bestvina–Bromberg–Fujiwara [BBF15] or hyperbolic graphs obtained by coning-off uniformly quasiconvex subsets of the hyperbolic group G [KR14]. We state one further general application, which follows directly from Theorem 1.7 and Theorem 1.3.

Theorem 1.10 (Epimorphisms are generically bilipschitz). *Let $\phi: G \rightarrow H$ be a surjective homomorphism between nonelementary hyperbolic groups. Then there is an $L > 0$ such that the set of $g \in G$ for which*

$$|\phi(g)| \geq L|g|$$

is generic in G .

Moreover, there is a subset $\partial^H G$ of ∂G with $\nu(\partial^H G) = 1$ and a boundary map $\partial\phi: \partial^H G \rightarrow \partial H$ extending the homomorphism $\phi: G \rightarrow H$.

1.3.2. Mapping class groups and $\text{Out}(F_N)$. Our main result on the genericity of loxodromics is in part motivated by a long-standing conjecture, going back to W. Thurston, about the mapping class group of an orientable surface S with $\chi(S) < -1$.

Conjecture 1.11 ([Far06, Conjecture 3.15]). *Let $\text{Mod}(S)$ be the mapping class group of an orientable surface S with $\chi(S) < -1$. Then pseudo-Anosov mapping classes are generic in $\text{Mod}(S)$.*

While Conjecture 1.11 seems at the moment out of reach, our main result does imply the corresponding statement for hyperbolic subgroups of $\text{Mod}(S)$. Recall that $\text{Mod}(S)$ is *not* itself hyperbolic and so the techniques of this paper do not directly apply. We say that a subgroup G of $\text{Mod}(S)$ is irreducible if no finite index subgroup of G fixes a multicurve on S .

Theorem 1.12 (Genericity in $\text{Mod}(S)$). *Let G be a nonelementary hyperbolic group which is an irreducible subgroup of $\text{Mod}(S)$. Then pseudo-Anosov mapping classes are generic in G with respect to any generating set of G .*

Proof. The mapping class group $\text{Mod}(S)$ acts by isometries on the curve complex $\mathcal{C}(S)$ of the surface S , which is hyperbolic by [MM99]. Since G is an irreducible subgroup of $\text{Mod}(S)$ which is not virtually cyclic, G has a nonelementary action on $\mathcal{C}(S)$; this follows from the subgroup structure theorems of [BLM83, Iva92]. Since the loxodromics of the action $\text{Mod}(S) \curvearrowright \mathcal{C}(S)$ are exactly the pseudo-Anosov mapping classes, again by [MM99], the result follows from Theorem 1.1. \square

We remark that hyperbolic, irreducible subgroups of $\text{Mod}(S)$, i.e. those subgroups to which Theorem 1.12 applies, are abundant. For example, there are several constructions of right-angled Artin subgroups of $\text{Mod}(S)$ [CLM12, Kob10] and such subgroups are well-known to contain a variety of hyperbolic subgroups (see Section 1.3.3).

The next application of Theorem 1.12 was pointed out to us by H. Masur and J. Tao. Let S be a closed surface of genus at least 2 and set $\mathring{S} = S \setminus p$ for some $p \in S$. Recall that there is a natural map $\text{Mod}(\mathring{S}) \rightarrow \text{Mod}(S)$ whose kernel is $\pi_1(S, p)$ [Bir69]. Moreover, the mapping class $f_\gamma \in \text{Mod}(\mathring{S})$ corresponding to $\gamma \in \pi_1(S, p)$ is pseudo-Anosov if and only if γ is *filling* [Kra81] – i.e. every curve homotopic to γ cuts S into disks. Hence, applying Theorem 1.12 to the induced action of $\pi_1(S)$ on the curve graph of \mathring{S} gives the following:

Corollary 1.13. *Let S be a closed surface of genus $g \geq 2$. Then filling curves are generic in $\pi_1(S)$ with respect to any generating set.*

There is also a direct analogue for hyperbolic subgroups of $\text{Out}(F_N)$. Similar to the discussion for mapping class groups, there are various techniques for constructing hyperbolic subgroups of $\text{Out}(F_N)$. See for example [Tay15]. We say that a subgroup G of $\text{Out}(F_N)$ is irreducible if no finite index subgroup of G fixes a free factor of F_N .

Theorem 1.14 (Genericity in $\text{Out}(F_N)$). *Let G be a nonelementary hyperbolic group which is an irreducible subgroup of $\text{Out}(F_N)$. Then fully irreducible automorphisms are generic in G with respect to any generating set of G .*

Moreover, if G is also not contained in a mapping class subgroup of $\text{Out}(F_N)$, then atoroidal, fully irreducible automorphisms are generic in G .

Proof. We mimic the proof of Theorem 1.12 using the action of $\text{Out}(F_N)$ on two free group analogues of the curve complex. First, since $G \leq \text{Out}(F_N)$ is irreducible and not virtually cyclic, the main result of [HM13] implies that G contains “independent” fully irreducible automorphisms. This implies that G has a nonelementary action on \mathcal{FF}_N , the free factor complex of F_N , which is hyperbolic by [BF14]. We then apply Theorem 1.1 to obtain the first part of the theorem since the loxodromic isometries of \mathcal{FF}_N are exactly the fully irreducible automorphisms [BF14].

To get the moreover statement, we use the action of $\text{Out}(F_N)$ on \mathcal{CS}_N , the co-surface graph of F_N . By [DT16], \mathcal{CS}_N is a hyperbolic graph and the loxodromic elements of the action $\text{Out}(F_N) \curvearrowright \mathcal{CS}_N$ are exactly the atoroidal fully irreducible automorphisms. By [Uya15], $G \leq \text{Out}(F_N)$ must contain an atoroidal element of $\text{Out}(F_n)$ for otherwise G is contained in a mapping class subgroup of $\text{Out}(F_N)$, contrary to our hypothesis. From this, it follows easily that the action $G \curvearrowright \mathcal{CS}_N$ is nonelementary and so another application of Theorem 1.1 completes the proof. \square

1.3.3. *Special hyperbolic groups and loxodromics in RAAGs.* As a final application, we recall the celebrated result of Agol, who building on work of Wise and his collaborators showed that every hyperbolic cubulated group G' has a finite index subgroup G which embeds into a right-angled Artin group $A(\Gamma)$ [Ago13]. Following work of Haglund and Wise, a finitely generated group which

embeds into a right-angled Artin group $A(\Gamma)$ is called *special* [HW08]. For additional information of RAAGs and special groups see, for example, [Wis12].

As $A(\Gamma)$ is itself a CAT(0) group, an important role is played by its elements which act by rank-1 isometries on its associated CAT(0) cube complex, its so-called Salvetti complex. Following [KK14], we call such elements of $A(\Gamma)$ *loxodromic*. Equivalently, loxodromic elements of $A(\Gamma)$ are characterized algebraically by having cyclic centralizer in $A(\Gamma)$ [Ser89, BF09] and geometrically as being the Morse elements of $A(\Gamma)$ [BC12]. Being of geometric interest, one might ask how frequently elements of a hyperbolic cubulated group G get mapped to loxodromics of $A(\Gamma)$ via the embeddings provided by Agol’s theorem, and our next application shows that this is indeed generically the case. Note this is in contrast to the well-known result that the only subgroups of $A(\Gamma)$ for which *each* nontrivial element is loxodromic are free subgroups [CLM12, KK14, KMT14].

For the statement of the theorem, recall that the join of two graphs Γ_1 and Γ_2 is the graph formed by connected every vertex of Γ_1 to every vertex of Γ_2 . A subgroup of $A(\Gamma)$ is called a join subgroup if it corresponds to $A(\Lambda)$, where Λ is a join subgraph of Γ . See [BC12] for additional details.

Theorem 1.15 (Genericity in hyperbolic, special groups). *Let G be a nonelementary hyperbolic group and $\phi: G \rightarrow A(\Gamma)$ be an injective homomorphism into a right-angled Artin group $A(\Gamma)$. If the image of G is not contained in a conjugate of a join subgroup, then*

$$\frac{\#\{g \in B_n : \phi(g) \text{ is loxodromic in } A(\Gamma)\}}{\#B_n} \rightarrow 1.$$

Proof. By Theorem 5.2 of [BC12], if $\phi(G)$ is not contained in a conjugate of a join subgroup, then $\phi(G)$ contains at least 2 independent loxodromic elements. This implies that the induced action of G on the extension graph Γ^e is nonelementary. Since Γ^e is hyperbolic and the loxodromic isometries of $A(\Gamma) \curvearrowright \Gamma^e$ are exactly the loxodromic elements of $A(\Gamma)$ ([KK14]), applying Theorem 1.1 to the action $G \curvearrowright \Gamma^e$ completes the proof. \square

1.4. Summary of paper. In Section 2 and Section 3 we present the theory of Markov chains on directed graphs that will be needed for our counting arguments. In Section 4 we describe a process by which we turn a Markov chain into a number of random walks, using first return probabilities. Section 5 then presents background material on hyperbolic groups and spaces, geodesic combings, and random walks.

These techniques are combined in Section 6 to show that sample paths in the geodesic combing of G converge to the boundary of X and have positive drift. This proves Theorem 1.3 and Theorem 1.4. We also show that with probability going to 1, the n th step of a sample path in the Markov chain is X -loxodromic with translation length growing linearly in n . One of the main steps in establishing these results is to show that the first return probabilities discussed above determine nonelementary measures on the group G .

Finally, in Section 7, we turn our statements about the Markov measure into statements about counting in the Cayley graph of the group G . This culminates in the proofs of Theorem 1.7 and Theorem 1.8.

Acknowledgements. The authors thank Nathan Dunfield, Camille Horbez, Ilya Kapovich, Sam Kim, Joseph Maher, Howard Masur, Kasra Rafi and Jing Tao for helpful suggestions. We also thank the referee for several corrections and suggestions that improved our paper. The first author is partially supported by NSF grant DMS-1401875 and ERC advanced grant “Moduli” of Prof. Ursula Hamenstädt, the second author is partially supported by NSF grants DMS-1400498 and DMS-1744551, and the third author is partially supported by NSERC and the Connaught fund.

2. COUNTING PATHS IN GRAPHS

We start by setting up our notation and recalling some fundamental facts about directed graphs. Note that we will first deal with general graphs, and introduce the group structure only later. We mostly follow Calegari–Fujiwara [CF10].

2.1. Almost semisimple graphs. Let Γ be a finite, directed graph with vertex set $V(\Gamma) = \{v_0, v_1, \dots, v_{r-1}\}$. The *adjacency matrix* of Γ is the $r \times r$ matrix $M = (M_{ij})$ defined so that M_{ij} is the number of edges from v_i to v_j .

Such a graph is *almost semisimple* of growth $\lambda > 1$ if the following hold:

- (1) There is an *initial vertex*, which we denote as v_0 ;
- (2) For any other vertex v , there is a (directed) path from v_0 to v ;
- (3) The largest modulus of the eigenvalues of M is λ , and for any eigenvalue of modulus λ , its geometric multiplicity and algebraic multiplicity coincide.

The (directed) *universal cover* of Γ is the countable tree $\tilde{\Gamma}$ whose vertex set is the set of finite directed paths in Γ starting from v_0 , and there is a directed edge $p_1 \rightarrow p_2$ in $\tilde{\Gamma}$ if the path p_2 is the concatenation of p_1 with an edge of Γ . By a path in Γ or $\tilde{\Gamma}$, we will always mean a directed path.

The set $\tilde{\Gamma}$ is naturally a rooted tree, where the root is the path of length 0 from v_0 to itself, hence it can be identified with v_0 . Moreover, $\tilde{\Gamma}$ is naturally endowed with a distance coming from the tree structure: there, a vertex has distance n from the initial vertex if and only if it represents a path of length n . We denote as $|g|$ the distance between $g \in \tilde{\Gamma}$ and the root. There is a natural graph map $\tilde{\Gamma} \rightarrow \Gamma$ which sends each vertex in $\tilde{\Gamma}$ representing a path in Γ to its endpoint, seen as a vertex of Γ . For any vertex g of $\tilde{\Gamma}$, we denote as $[g]$ the corresponding vertex of Γ .

Given two vertices v_1, v_2 of a directed graph, we say that v_2 is *accessible from* v_1 and write $v_1 \rightarrow v_2$ if there is a path from v_1 to v_2 . Then we say that two vertices are *mutually accessible* if $v_1 \rightarrow v_2$ and $v_2 \rightarrow v_1$. Mutually accessibility is an equivalence relation, and equivalence classes are called *irreducible components* of Γ .

The *boundary* of $\tilde{\Gamma}$ is the set $\partial\tilde{\Gamma}$ of infinite paths in Γ starting from the initial vertex. The set $\tilde{\Gamma} \cup \partial\tilde{\Gamma}$ carries a natural metric, where the distance between two (finite or infinite) paths γ_1 and γ_2 is $d(\gamma_1, \gamma_2) = 2^{-P(\gamma_1, \gamma_2)}$ where $P(\gamma_1, \gamma_2)$ is the length of the longest path from the initial vertex which is a common prefix of both γ_1 and γ_2 . This way, $\tilde{\Gamma} \cup \partial\tilde{\Gamma}$ is a compact metric space, and $\partial\tilde{\Gamma}$ is a Cantor set.

2.2. Counting measures. For each n , we denote as $S_n \subseteq \tilde{\Gamma}$ the sphere of radius n around the origin, i.e. the set of vertices at distance n from the initial vertex. This defines a sequence P^n of *counting measures* on $\tilde{\Gamma}$: namely, for each set $A \subset \tilde{\Gamma}$ we define

$$P^n(A) := \frac{\#(A \cap S_n)}{\#S_n}.$$

For any subset $A \subseteq \tilde{\Gamma}$, we define the *growth* $\lambda(A)$ of A as

$$\lambda(A) := \limsup_{n \rightarrow \infty} \sqrt[n]{\#(A \cap S_n)}.$$

By construction, the growth of $\tilde{\Gamma}$ is λ (see also Lemma 2.3).

For each vertex v of Γ which lies in a component C , let $\mathcal{P}_v(C)$ denote the set of finite paths in Γ based at v which lie entirely in C . Moreover, for any lift g of v to $\tilde{\Gamma}$, we let $\mathcal{P}_g(C)$ be the set of endpoints of paths in $\tilde{\Gamma}$ based at g whose projection to Γ lies entirely in C .

We call an irreducible component C of Γ *maximal* if for some (equivalently, any) g which projects to an element of C , the growth of $\mathcal{P}_g(C)$ equals λ .

2.3. Vertices of small and large growth.

Definition 2.1. For each vertex $g \in \tilde{\Gamma}$, the *cone* of g , denoted as $\text{cone}(g)$, is the set of (finite or infinite) paths in $\tilde{\Gamma}$ starting at v_0 and passing through g .

We now define vertices to be of large or small growth, according to the growth of their cone. More precisely, following Calegari-Fujiwara [CF10], we define the linear map $\rho : \mathbb{R}^{V(\Gamma)} \rightarrow \mathbb{R}^{V(\Gamma)}$ as

$$\rho(v) := \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^N \frac{M^n v}{\lambda^n}$$

The limit exists by the almost semisimplicity assumption (3). If we denote by $\mathbf{1}$ the vector in $\mathbb{R}^{V(\Gamma)}$ all of whose coordinates are 1, note that by construction $(M^n \mathbf{1})_i$ equals the number of paths of length n starting at v_i , i.e. for any $g \in \tilde{\Gamma}$ we have

$$(M^n \mathbf{1})_i = \#(\text{cone}(g) \cap S_{n+|g|})$$

where $[g] = v_i$. In particular, $(M^n \mathbf{1})_0 = \#S_n$.

Definition 2.2. We say a vertex v_i of Γ has *small growth* if $\rho(\mathbf{1})_i = 0$, and has *large growth* if $\rho(\mathbf{1})_i > 0$.

Note that if v_i is of small growth and there is an edge $v_i \rightarrow v_j$, then v_j is also of small growth. We denote as LG the set of vertices of large growth. Note that if a vertex belongs to a maximal irreducible component, then it has large growth, but the converse is not necessarily true; indeed, the cone of a vertex may have maximal growth, while its irreducible component may have smaller growth.

Lemma 2.3. Let Γ be an almost semisimple graph of growth $\lambda > 1$.

(1) There exists $c > 0$ such that for any vertex v of large growth and any $n \geq 0$,

$$c^{-1} \lambda^n \leq \#\{\text{paths from } v \text{ of length } n\} \leq c \lambda^n$$

(2) There exists $c > 0$ and $\lambda_1 < \lambda$ such that, for any vertex v of small growth and any $n \geq 0$,

$$\#\{\text{paths from } v \text{ of length } n\} \leq c \lambda_1^n$$

(3) Let v be a vertex which belongs to a maximal component C . Then there exists $c > 0$ such that for any $n \geq 0$,

$$c^{-1} \lambda^n \leq \#\{\text{paths in } \mathcal{P}_v(C) \text{ of length } n\} \leq c \lambda^n$$

Proof. Let us first assume that $v = v_0$ is the initial vertex, which has large growth by construction. By writing M in Jordan normal form, one has $\|M^n\| \asymp \lambda^n$ for any submultiplicative norm. Here, the symbol \asymp means that the two quantities have ratios which are uniformly bounded. Since all norms on the space of matrices are equivalent, we also have

$$\sum_{i,j} (M^n)_{ij} \asymp \lambda^n$$

which can be rewritten as

$$\sum_i (M^n \mathbf{1})_i \asymp \lambda^n$$

Now, for each vertex v_j there is a path of length l_j from v_0 to v_j , hence

$$(M^{n+l_j} \mathbf{1})_0 \geq (M^n \mathbf{1})_j$$

Moreover, since the degree of each vertex is bounded by some d ,

$$(M^{n+l_j} \mathbf{1})_0 \leq d^{l_j} (M^n \mathbf{1})_0$$

thus putting together the estimates yields

$$(M^n \mathbf{1})_0 \leq \sum_i (M^n \mathbf{1})_i \leq \sum_i d^i (M^n \mathbf{1})_0$$

which gives

$$\#S_n = (M^n \mathbf{1})_0 \asymp \lambda^n$$

proving (1) in this case.

Note that the transpose matrix M^T has the following interpretation. Let e_i be the basis vector of $\mathbb{R}^{V(\Gamma)}$ which corresponds to v_i , and for a set A of vertices, let V_A be the span of $(e_i)_{v_i \in A}$. Then by definition $M^T e_i = \sum_k \#\{v_i \rightarrow v_k\} e_k$. Hence we also get $M^T V_A \subseteq V_B$ where B is the set of vertices which can be reached in one step from A .

Let us now pick a vertex v , and consider the set \mathcal{V} of all vertices which are accessible from v . Then the matrix M^T leaves invariant the span V' of all basis vectors corresponding to elements of \mathcal{V} . Let M_1 be the restriction of M^T to V' ; then either M_1 is almost semisimple of growth λ , or it has spectral radius $< \lambda$. In the first case, the above proof shows that

$$\#(\text{cone}(v) \cap S_n) \asymp \lambda^n$$

which implies that v has large growth; in the second case, then $(M_1^n \mathbf{1})_v \leq \|M_1^n\| \leq c\lambda_1^n$ with $\lambda_1 < \lambda$, hence v has small growth, proving (2).

To prove (3), let us consider the set \mathcal{W} of vertices which belong to the maximal component C , let V'' be the span of its basis vectors. Now, M_1^T leaves invariant V'' , and let M_2 be the restriction of M_1^T to V'' . Again, either M_2 is almost semisimple or it has spectral radius strictly smaller than λ . However, since C is maximal the spectral radius cannot be strictly smaller than λ , hence if $v = v_i$ we have

$$\#\{\text{paths in } \mathcal{P}_v(C) \text{ of length } n\} = (M_2^n \mathbf{1})_i \asymp \lambda^n.$$

□

Remark 2.4. In order to clarify the connection between large growth vertices and maximal components, let us note that one can prove the following: if v is a vertex which belongs to a component C , then v has large growth if and only if there exists a descendent component of C (possibly C itself) which is maximal.

Let us denote as S_n the set of elements of $\tilde{\Gamma}$ at distance n from the initial vertex. If $g \in S_n$, then we denote \hat{g} the element along the path from the initial vertex to g at distance $n - \lfloor \log n \rfloor$ from the initial vertex.

Proposition 2.5. *We have*

$$\frac{\#\{g \in S_n : \hat{g} \text{ of small growth}\}}{\#S_n} \rightarrow 0$$

as $n \rightarrow \infty$.

Proof. By Lemma 2.3 (1), the number of paths from the origin of length $n - \lfloor \log n \rfloor$ is at most $c\lambda^{n - \lfloor \log n \rfloor}$, and by Lemma 2.3 (2) for each vertex of small growth the number of outgoing paths of length $\lfloor \log n \rfloor$ is at most $c\lambda_1^{\lfloor \log n \rfloor}$, so the total number of paths of length n from the origin which passes at time $n - \lfloor \log n \rfloor$ through a vertex of small growth is at most

$$c\lambda^{n - \lfloor \log n \rfloor} \cdot c\lambda_1^{\lfloor \log n \rfloor} \leq c^2 \lambda^{n+1} n^{\log \lambda_1 - \log \lambda}$$

which is negligible with respect to the cardinality of S_n , of the order of λ^n . □

2.4. The Patterson-Sullivan measure. We define the *Patterson-Sullivan measure* ν as the limit of the measures

$$\nu_N := \frac{\sum_{|g| \leq N} \lambda^{-|g|} \delta_g}{\sum_{|g| \leq N} \lambda^{-|g|}}$$

on $\tilde{\Gamma} \cup \partial\tilde{\Gamma}$. As a consequence of the almost semisimplicity one gets the following proposition. For its statement, recall that $\mathbf{1}$ denotes the vector with all entries equal to 1.

Proposition 2.6. *The limit $\nu := \lim_{N \rightarrow \infty} \nu_N$ exists as a limit of measures on $\tilde{\Gamma} \cup \partial\tilde{\Gamma}$, and it is given on cones by the formula:*

$$(1) \quad \nu(\text{cone}(g)) = \begin{cases} \frac{\rho(1)_i}{\rho(1)_0} \lambda^{-|g|} & \text{if } [g] = v_i \text{ has large growth} \\ 0 & \text{otherwise} \end{cases}$$

In particular, ν is supported on $\partial\tilde{\Gamma}$. The measure ν is called the Patterson-Sullivan measure.

For a proof, see the discussion preceding Lemma 4.19 of [CF10]. The only difference is that they consider the limit of the measures $\hat{\nu}_N := \frac{1}{N} \sum_{|g| \leq N} \lambda^{-|g|} \delta_g$, so the limit may not be a probability measure. We normalize by dividing by $\rho(1)_0$ so that ν is a probability measure.

3. MARKOV CHAINS

Let Γ be a finite directed graph, with vertex set $V(\Gamma)$ and edge set $E(\Gamma)$.

A *transition probability* μ for Γ is a function $\mu : E(\Gamma) \rightarrow \mathbb{R}^{\geq 0}$ such that for each vertex $v \in V(\Gamma)$, the total sum of the probabilities of the edges going out from v is 1:

$$\sum_{e \in \text{Out}(v)} \mu(e) = 1$$

where $\text{Out}(v)$ denotes the set of outgoing edges from vertex v . A *vertex distribution* for Γ is a function $p : V(\Gamma) \rightarrow \mathbb{R}^{\geq 0}$ such that $\sum_{v \in V(\Gamma)} p(v) = 1$.

Let Ω denote the set of all infinite paths in Γ , no matter what the starting vertex is. Given a transition probability μ and a vertex distribution p , we define a probability measure $\mathbb{P}_{p,\mu}$ on Ω by assigning to each cylinder set

$$C_{e_1, \dots, e_n} = \{\omega = (\omega_n) \in \Omega : \omega_1 = e_1, \dots, \omega_n = e_n\}$$

the probability

$$\mathbb{P}_{p,\mu}(C_{e_1, \dots, e_n}) = p(v_0) \mu(e_1) \cdots \mu(e_n)$$

where v_0 is the initial vertex of the edge e_1 . Finally, if $\gamma = e_1 \dots e_n$ is a finite path in Γ or $\tilde{\Gamma}$, we define

$$\mu(\gamma) = \mu(e_1) \dots \mu(e_n)$$

to be the product of the transition probabilities of its edges.

3.1. Recurrence. Let us now fix a transition probability μ on the graph Γ .

Then for each vertex v , we denote as $\mathbb{P}_v = \mathbb{P}_{\delta_v, \mu}$ the probability measure on Ω where the vertex distribution is the δ -measure at v . Clearly, this measure is supported on the set of infinite paths starting at v , which we denote by Ω_v . Finally, we denote simply as $\mathbb{P} = \mathbb{P}_{v_0}$ the probability measure on the set of infinite paths starting at the initial vertex v_0 . Set $\Omega_0 = \Omega_{v_0}$.

The measure \mathbb{P} defines a Markov process on $\tilde{\Gamma}$, which starts at the initial vertex v_0 and moves along the graph according to the transition probabilities. To be precise, let us define the map $g_n : \Omega_0 \rightarrow \tilde{\Gamma}$ which associates to each infinite path starting at the initial vertex its prefix of length n , seen as a vertex of $\tilde{\Gamma}$. This defines a sequence of *Markov measures* (\mathbb{P}^n) on $\tilde{\Gamma}$ by setting $\mathbb{P}^n(A)$ as the probability that the Markov chain hits A at time n , that is $\mathbb{P}^n(A) = \mathbb{P}(g_n \in A)$.

Definition 3.1. A vertex v of Γ is *recurrent* if:

- (1) there is a path from v_0 to v of positive probability; and
- (2) whenever there is a path from v to another vertex w of positive probability, there is also a path from w to v of positive probability.

Two recurrent vertices v_1 and v_2 are *equivalent* if there exists a path from v_1 to v_2 of positive probability, and a path from v_2 to v_1 of positive probability. A *recurrent component* is an equivalence class of recurrent vertices.

For each n , let us define $X_n : \Omega \rightarrow \Gamma$ so that $X_n(\omega)$ denotes the endpoint of the finite prefix of ω of length n . Once we fix a measure on Ω , the sequence (X_n) is a stochastic process with values in Γ (in our case, a Markov chain) and X_n is the location of the n th step of the chain.

Lemma 3.2. Let Γ be a directed graph with transition probability μ .

- (1) For \mathbb{P} -almost every path in Ω , there exists a recurrent component C and an index N such that X_n belongs to C for all $n \geq N$. Moreover, X_n visits every vertex of C infinitely many times.
- (2) There exists $c > 0$ such that, for any recurrent vertex v , for each $n \geq 0$ we have

$$\mathbb{P}_v(\tau_v^+ = n) \leq ce^{-cn}$$

where $\tau_v^+ = \min\{n \geq 1 : X_n = v\}$ denotes the first return time to vertex v .

Proof. First note that there must be at least one recurrent vertex. In fact, let us consider the set V' of vertices accessible from v_0 with a path of positive probability. If $v \in V'$ is not recurrent, then there exists a vertex $f(v) \in V'$ such that there is a path of positive probability from v to $f(v)$ and there is no path of positive probability from $f(v)$ to v . Thus, if none of the vertices in V' is recurrent, there is a map $f : V' \rightarrow V'$ and since V' is finite there must be a periodic cycle, i.e. v such that $f^k(v) = v$ for some k . That is, there is a path of positive probability from $f(v)$ to v , but this contradicts the definition of f . We conclude that Γ has a recurrent vertex.

For each pair of vertices v_i, v_j such that there is a path from v_i to v_j of positive probability, let us pick one such path γ_{ij} , and let \mathcal{L} be the (finite) collection of all γ_{ij} . Now, let r be the maximum length of all the elements of \mathcal{L} , and $p := \min_{\gamma \in \mathcal{L}} \mu(\gamma) > 0$. For each non-recurrent vertex v_i which can be reached from v_0 with positive probability, there is a path γ in \mathcal{L} from v_i to some recurrent v_j . Then for each non-recurrent v_i we have

$$\mathbb{P}_{v_i}(X_n \text{ not recurrent for all } n \leq r) \leq 1 - p$$

Thus, for any k we have

$$\mathbb{P}_{v_i}(X_n \text{ not recurrent for all } n \leq kr) \leq (1 - p)^k$$

yielding the first claim if $k \rightarrow \infty$. Now, if v_i and v_j lie in the same recurrent component, then there are paths in \mathcal{L} from v_i to v_j and from v_j to v_i , hence

$$\mathbb{P}_{v_i}(X_n \neq v_j \text{ for all } n \leq r) \leq 1 - p$$

and by repeated application of this argument (as you never leave a recurrent component)

$$\mathbb{P}_{v_i}(X_n \neq v_j \text{ for all } n \leq kr) \leq (1 - p)^k$$

which implies the second claim. The claim about the return time follows from the above inequality setting $v_i = v_j = v$ a recurrent vertex. \square

3.2. Markov measure. On the other hand, given an almost semisimple graph Γ of growth λ , we construct a Markov chain in the following manner. Let us consider the graph Γ' given by adding to Γ an edge f_v from v to v for every vertex v of Γ of small growth. Then we define a transition probability $\mu: E(\Gamma') \rightarrow \mathbb{R}^{\geq 0}$ as follows: If v_i has large growth, then for each directed edge e from v_i to v_j we define

$$(2) \quad \mu(e) := \frac{\rho(1)_j}{\lambda \rho(1)_i}.$$

If v_i has small growth, we define $\mu(f_{v_i}) = 1$ and $\mu(e) = 0$ for every other edge with initial vertex v_i . With this definition, if $\mu(v_i \rightarrow v_j)$ denotes the probability of going from v_i to v_j and if v_i has large growth, then

$$\mu(v_i \rightarrow v_j) = \frac{M_{ij} \rho(1)_j}{\lambda \rho(1)_i}.$$

If v_i has small growth, $\mu(v_i \rightarrow v_j) = 0$ for $i \neq j$ and $\mu(v_i \rightarrow v_i) = 1$. With this observation, one can easily check that μ defines a Markov matrix for Γ as in [CF10, Lemma 4.9]. Note that, as we will see shortly, in the corresponding Markov chain the probability of ever reaching a small growth vertex starting at a large growth vertex is 0, so the Markov measure is supported on infinite paths in Γ (rather than Γ'). Hence, how we chose to define the transition probabilities at edges out of small growth vertices was inconsequential.

For this choice of transition probability μ , it is now easy to explicitly compute the n -step distributions \mathbb{P}^n of the associated Markov chain. In fact, from eq. (2) one gets for each vertex $g \in \tilde{\Gamma}$

$$(3) \quad \mathbb{P}^n(\{g\}) = \begin{cases} \frac{\rho(1)_i}{\rho(1)_0} \lambda^{-n} & \text{if } |g| = n, [g] = v_i \text{ has large growth} \\ 0 & \text{otherwise} \end{cases}$$

In greater detail, if $|g| = n$ write the corresponding path in Γ starting at v_0 as the concatenation of edges $e_1 e_2 \cdots e_n$. If $[g] = v_i$ has large growth, then so do all the edges appearing as endpoints of the e_i (see Lemma 2.3). In this case, eq.(3) is verified by multiplying together the transition probabilities for each e_i . When $[g]$ has small growth, then we take the first edge, say e_i , whose terminal vertex is of small growth. (Recall that v_0 has large growth.) Naming the initial vertex of e_i by v_k and its terminal vertex by v_l , then $\mu(e_i) = \frac{\rho(1)_l}{\lambda \rho(1)_k} = 0$ since v_l has small growth and v_k has large growth. This shows that in this case $\mathbb{P}^n(\{g\}) = 0$, as required.

One can check from direct calculation, comparing formulas (1) and (3) that the sequence (\mathbb{P}^n) of measures on $\tilde{\Gamma} \cup \partial \tilde{\Gamma}$ converges to the Patterson-Sullivan measure ν . In particular, if we identify Ω_0 with $\partial \tilde{\Gamma}$ via path lifting to the base vertex, then $\nu = \mathbb{P}$.

Lemma 3.3. *Let Γ be an almost semisimple graph and let μ be given by eq. (2). Then an irreducible component of Γ is maximal if and only if it is recurrent.*

Proof. Let C be an irreducible component of Γ , and $v_i \in C$ a vertex. First note that if C is recurrent, then every vertex of C has large growth and \mathbb{P}_{v_i} -almost every path which starts at v_i stays in C . Hence by equation (2) one gets

$$1 = \sum_{v_j \in C} \mathbb{P}_{v_i}(X_n = v_j) = \sum_{j \in C} \frac{M_{ij}^n \rho(1)_j}{\lambda^n \rho(1)_i} \leq \left(\frac{\max_{v_j \in C} \rho(1)_j}{\rho(1)_i} \right) \frac{\sum_{v_j \in C} M_{ij}^n}{\lambda^n}$$

Note that $\sum_{j \in C} M_{ij}^n$ equals the number of paths of length n which start from v_i and lie entirely in C , hence by the above estimate such number is bounded below by $c\lambda^n$, proving that C is maximal.

Conversely, suppose that C is maximal. Then each vertex of C has large growth and so for each v in C there is a path from v_0 to v of positive probability. If there is a path from v to some vertex w of positive probability, then w also has large growth (eq. (2)). Since Γ is almost semisimple, no

two maximal components are joined by a directed path [Cal13, Lemma 3.4.2]. Hence, $w \in C$ since any vertices of large growth must have a path to a maximal component (see Remark 2.4). Now any directed path from w back to v in C has positive probability by definition of μ , showing that C is recurrent and completing the proof. \square

Finally, the next lemma compares the n -step distribution for the Markov chain to the counting measure.

Lemma 3.4. *There exists $c > 1$ such that, for each $A \subseteq \tilde{\Gamma}$, we have the inequality*

$$(4) \quad c^{-1} \mathbb{P}^n(A) \leq P^n(A \cap LG) \leq c \mathbb{P}^n(A).$$

Proof. By definition,

$$P^n(A \cap LG) = \frac{\#(A \cap LG \cap S_n)}{\#S_n}$$

while from formula (3)

$$\mathbb{P}^n(A) = \left(\frac{1}{\rho(1)_0} \right) \sum_{g \in A \cap LG \cap S_n} \rho(1)_g \lambda^{-n}$$

Now, since there are a finite number of vertices, we have $c^{-1} \leq \rho(1)_g \leq c$ for any g of large growth, hence

$$\mathbb{P}^n(A) = \left(\frac{1}{\rho(1)_0} \right) \sum_{g \in A \cap LG \cap S_n} \rho(1)_g \lambda^{-n} \asymp \#(A \cap LG \cap S_n) \lambda^{-n}$$

and since $\#S_n \asymp \lambda^n$ we get the claim. \square

4. FROM MARKOV CHAINS TO RANDOM WALKS

Let Γ be the finite graph, v_0 its initial vertex, and \mathbb{P} the Markov measure on the space Ω_0 of infinite paths starting at the initial vertex.

We will associate to each recurrent vertex of the Markov chain a random walk, and use previous results on random walks to prove statements about the asymptotic behavior of the Markov chain.

4.1. Return times. For each sample path $\omega \in \Omega_0$, let us define $n(k, v, \omega)$ as the k^{th} time the path ω lies at the vertex v . In formulas,

$$n(k, v, \omega) := \begin{cases} 0 & \text{if } k = 0 \\ \min(h > n(k-1, v, \omega) : X_h(\omega) = v) & \text{if } k \geq 1 \end{cases}$$

To simplify notation, we will write $n(k, v)$ instead of $n(k, v, \omega)$ when the sample path ω is fixed. Moreover, we denote

$$\tau(k, v) := n(k, v) - n(k-1, v).$$

These are the return times to the vertex v .

Definition 4.1. For every vertex v of Γ , the *loop semigroup* of v is the set L_v of loops in the graph Γ which begin and end at v . It is a semigroup with respect to concatenation. A loop in L_v is *primitive* if it is not the concatenation of two (non-trivial) loops in L_v .

Let us now pick a recurrent component C , and let Ω_C be the set of all infinite paths from the initial vertex which enter C and remain inside C forever. We denote as \mathbb{P}_C the conditional probability of \mathbb{P} on Ω_C .

Let us now fix a vertex v of C . Then \mathbb{P}_C -almost every infinite path γ in Ω_C passes through v infinitely many times, hence such γ can be decomposed as a concatenation $\gamma = \gamma_0 \gamma_1 \dots \gamma_n \dots$

where γ_0 is a path from v_0 to v which visits v only once, and each γ_i for $i \geq 1$ is a primitive loop at v . This defines for each recurrent vertex $v \in C$ a measurable map

$$\begin{aligned} \varphi_v : (\Omega_C, \mathbb{P}_C) &\rightarrow (L_v)^\mathbb{N} \\ (e_n) &\rightarrow (\gamma_1, \gamma_2, \dots) \end{aligned}$$

where $\gamma_k = e_{n(k,v)+1} \dots e_{n(k+1,v)}$.

We now define the *first return measure* μ_v on the set of primitive loops by setting, for each primitive loop $\gamma = e_1 \dots e_n$ with edges e_1, \dots, e_n ,

$$\mu_v(e_1 \dots e_n) = \mu(e_1) \dots \mu(e_n).$$

Extend μ_v to the entire loop semigroup L_v by setting $\mu_v(\gamma) = 0$ if $\gamma \in L_v$ is not primitive. Since almost every path starting at v visits v infinitely many times, the measure μ_v is a probability measure.

Lemma 4.2. *For any recurrent vertex v which belongs to a component C , we have*

$$(\varphi_v)_\star \mathbb{P}_C = \mu_v^\mathbb{N}.$$

Proof. It suffices to show that the measures $\mu_v^\mathbb{N}$ and $(\varphi_v)_\star \mathbb{P}_C$ agree on cylinder sets of the form $C_{l_1, l_2, \dots, l_k} \subset (L_v)^\mathbb{N}$. We may also assume that each l_i is primitive, otherwise C_{l_1, l_2, \dots, l_k} has $\mu_v^\mathbb{N}$ -measure 0 and $\varphi_v^{-1}(C_{l_1, l_2, \dots, l_k})$ is empty. Hence, $\mu_v^\mathbb{N}(C_{l_1, l_2, \dots, l_k}) = \mu_v(l_1) \dots \mu_v(l_k)$.

Let H_v be the set of paths in Ω_C which start at v_0 and meet v only at their terminal endpoint. Then we have the disjoint union

$$\varphi_v^{-1}(C_{l_1, l_2, \dots, l_k}) = \bigcup_{\gamma \in H_v} \gamma \cdot l_1 l_2 \dots l_k \cdot \Omega_v.$$

So by the Markov property

$$\begin{aligned} \mathbb{P}_C(\varphi_v^{-1}(C_{l_1, l_2, \dots, l_k})) &= \sum_{\gamma \in H_v} \mathbb{P}_C(\gamma) \mu_v(l_1) \dots \mu_v(l_k) \\ &= \mu_v(l_1) \dots \mu_v(l_k), \end{aligned}$$

where the second equality follows from the fact that almost every path in Ω_C hits v and that each l_i is a primitive loop. This completes the proof. \square

By Lemma 3.2, for every recurrent vertex v , the first return measure μ_v has finite exponential moment, i.e. there exists a constant $\alpha > 0$ such that

$$(5) \quad \int_{L_v} e^{\alpha|\gamma|} d\mu_v(\gamma) < \infty.$$

5. GROUPS, COMBINGS AND RANDOM WALKS

The connection between hyperbolic groups and Markov chains is through the theory of geodesic combings.

Definition 5.1. A *finite state automaton* over the alphabet S is a finite, directed graph Γ , with a distinguished *initial vertex* v_0 and such that its edges are labeled by elements of S , in such a way that no vertex has two outgoing edges with the same label.

Let G be a countable group, and $S \subset G$ a finite set which generates G as a semigroup. Then we say that G has a *geodesic combing* with respect to S if there exists a finite state automaton Γ such that:

- (1) every path in Γ starting at the initial vertex is a geodesic in the word metric for S ;

- (2) for every element $g \in G$, there exists exactly one path in Γ which starts at the initial vertex and represents g .

In this paper, we will only need geodesic combings on hyperbolic groups. For detailed introduction, see [Cal13]. We begin by reviewing some basics on hyperbolic metric spaces.

5.1. Hyperbolic spaces and groups. A geodesic metric space X is δ -hyperbolic, where $\delta \geq 0$, if for every geodesic triangle in X , each side is contained within the δ -neighborhood of the other two sides. Each hyperbolic space has a well-defined *Gromov boundary* ∂X , and we refer the reader to [BH09, Section III.H.3], [GdlH90], or [KB02, Section 2] for definitions and properties. We note that when the metric space X is proper (meaning that closed metric balls are compact), $X \cup \partial X$ is a compactification of X . If a group G acts on X by isometries, then there is an induced action $G \curvearrowright \partial X$ by homeomorphisms.

Recall that we have defined a *nonelementary* action of G on X as an action $G \curvearrowright X$ for which there are $g, h \in G$ which act loxodromically on X and whose fixed point sets on the Gromov boundary of X are disjoint. We say that two such elements of G are independent loxodromics for the action $G \curvearrowright X$. This definition extends equally well to an action of a semigroup L on X . We will require the following criterion which characterizes when a semigroup action $L \curvearrowright X$ is nonelementary:

Proposition 5.2 ([DSU14, Proposition 7.3.1]). *Suppose that L is a semigroup which acts on the hyperbolic metric space X by isometries. If the limit set $\Lambda_L \subset \partial X$ of L on the boundary of X is nonempty and L does not have a finite orbit in ∂X , then the action $L \curvearrowright X$ is nonelementary⁴, that is L contains two independent loxodromics.*

If G is a finitely generated group, then G is *hyperbolic* if for some (any) finite generating set S , the associated Cayley graph $C_S(G)$ is δ -hyperbolic for some $\delta \geq 0$. Since $C_S(G)$ is locally finite, the Gromov boundary $\partial G = \partial C_S(G)$ is compact. In fact, the boundary of G can be identified with the set of asymptote classes of geodesic rays starting at the identity, where two geodesic rays $\gamma_1, \gamma_2 : [0, \infty) \rightarrow G$ are asymptotic if $\sup_t d(\gamma_1(t), \gamma_2(t)) < \infty$. Since the boundary ∂G is compact and Hausdorff, it is a Baire space – the countable union of closed nowhere dense sets has empty interior.

Hyperbolic groups have automatic structures by the well-known theorem of Cannon:

Theorem 5.3 ([Can84]). *Let G be a hyperbolic group, and S a finite generating set. Then G has a geodesic combing with respect to S .*

Let Γ be the directed graph associated to the geodesic combing of G , whose edges are labeled by generators from S . As Γ is a directed graph, we will use the notation and terminology established in Section 2. That Γ comes from a combing of G implies that there is a map $\text{ev}: \tilde{\Gamma} \rightarrow G$, the *evaluation map*, which associates to each directed path p in Γ beginning at v_0 , the element $\text{ev}(p) \in G$ represented by the word obtained by concatenating the edge labels of p . By definition, $|\text{ev}(p)| = |p|$ and $\text{ev}: \tilde{\Gamma} \rightarrow G$ is bijective. We remark that this map extends to *all* directed paths in Γ (no matter which vertex they start from), again by concatenating edge labels, and we will continue to denote this map by ev . Note that the extension still maps directed paths to geodesics, but is no longer injective since paths which start at different vertices of Γ may still read the same edge labels. Moreover, since paths in $\tilde{\Gamma}$ beginning at v_0 map to geodesics in G beginning at 1, ev induces a boundary map, denoted $\text{ev}: \partial \tilde{\Gamma} \rightarrow \partial G$, which is continuous, surjective, closed, and finite-to-one [Cal13].

⁴Das–Simmons–Urbanski call such an action *of general type*.

In fact, the graph Γ is almost semisimple [CF10, Lemma 4.15], this follows from a theorem of Coornaert. Recall that S_n is the set of elements of G of word length n with respect to a generating set S .

Theorem 5.4 ([Coo93, Theorem 7.2]). *Let G be a nonelementary hyperbolic group with generating set S . Then there are constants $\lambda > 0$, $c \geq 1$ such that for any n*

$$c^{-1}\lambda^n \leq \#S_n \leq c\lambda^n.$$

We remark that since Γ parameterizes the geodesic combing of G , the constant $\lambda > 0$ in Theorem 5.4 is the growth rate of Γ .

Finally, the map $\text{ev}: \partial\tilde{\Gamma} \rightarrow \partial G$ pushes forward the Patterson-Sullivan measure on $\partial\tilde{\Gamma}$ to the usual *Patterson-Sullivan measure* on ∂G , which we continue to denote by ν . The measure ν is a quasiconformal measure on ∂G , which can alternatively be defined, up to bounded ratios, as a weak limit of uniform measure on the balls B_n . For details, see [Coo93] or Section 4.3 of [CF10].

Lemma 5.5 ([Coo93][Cal13]). *The Patterson-Sullivan measure on ∂G has full support, i.e. every open set has positive measure. Moreover, the action of G on ∂G is ergodic, i.e. any G -invariant subset of ∂G has either zero or full measure.*

Finally, we turn to the definition and basic properties of shadows in the δ -hyperbolic space X . For $x, y \in X$, the *shadow in X around y based at x* is

$$S_x(y, R) = \{z \in X : (y, z)_x \geq d(x, y) - R\},$$

where $R > 0$ and $(y, z)_x = \frac{1}{2}(d(x, y) + d(x, z) - d(y, z))$ is the usual *Gromov product*. The *distance parameter* of $S_x(y, R)$ is by definition the number $r = d(x, y) - R$, which up to an additive constant depending only on δ , measures the distance from x to $S_x(y, R)$. Indeed, $z \in S_x(y, R)$ if and only if any geodesic $[x, z]$ 2δ -fellow travels any geodesic $[x, y]$ for distance $r + O(\delta)$.

We will use shadows in both the hyperbolic group G and the hyperbolic space X , but the location of the shadow will always be clear from context. For example, $S_1(g, R)$ will usually denote a shadow in G based at the identity element.

The following lemma is an easy exercise in metric geometry.

Lemma 5.6. *For each $D \geq 0$, and each x, y in a metric space, we have*

$$N_D(S_x(y, R)) \subseteq S_x(y, R + D).$$

Returning to the directed graph Γ parameterizing the geodesic combing for G , we say a vertex v is *continuable* if there exists at least one infinite path starting from v . Otherwise, v is said to be *non-continuable*. Elements of g are called continuable or non-continuable according to the property of their vertex. Note that g is continuable if and only if it lies on an infinite geodesic ray in the combing.

Lemma 5.7. *Let r be the number of vertices in Γ . Then for every $g \in G$, there exists $h \in G$ with $d_G(g, h) \leq r$ and such that the vertex of h is continuable.*

Proof. Note that every path of length $\geq r$ in the graph passes through some vertex at least twice, and a vertex which supports a loop is continuable. If $|g| \leq r$, just pick $h = 1$; otherwise, consider the path in the combing from 1 to g , and look at the final subpath of length r (the one ending in g). By the previous observation, such subpath has to hit a continuable vertex, proving the claim. \square

Lemma 5.8. *There exists a constant $c \geq 0$ which depends only on δ such that, for any g and any $R \geq 0$, the shadow $S_1(g, R + c)$ in G has maximal growth.*

Proof. Let us first assume that g lies on an infinite geodesic ray γ in the Cayley graph of G based at 1 and set $D = |g|$. Moreover, let g_n be a point on γ at distance $n + D + c$ from 1, where c is a constant which depends on δ . We claim that the ball of radius n around g_n is contained in $S_1(g, c) \subseteq S_1(g, R + c)$. Indeed, if $h \in B_n(g_n)$, then

$$(h, g_n)_1 \geq d_G(1, g_n) - d_G(h, g_n) \geq n + D + c - n = D + c$$

Thus, a closest point projection p of h to $[1, g_n]$ lies on the segment $[g, g_n]$. As a consequence, $d(1, h) = d(1, g) + d(g, p) + d(p, h) + O(\delta)$ and $d(g, h) = d(g, p) + d(p, h) + O(\delta)$, hence

$$(h, g)_1 = d_G(1, g) + O(\delta)$$

which proves the claim.

Let us now prove the general case. Given $g \in G$, by Lemma 5.7 there exists $h \in G$ continuable such that $d_G(g, h) \leq r$. By the previous paragraph, the shadow $S_1(h, c)$ has maximal growth, and by writing out the Gromov products and using the triangle inequality

$$S_1(h, c) \subseteq S_1(g, c + r)$$

hence the claim follows by replacing c with $c + r$. □

5.2. Random walks. If X is a separable hyperbolic space and G acts on X by isometries, then a probability measure μ on G is said to be *nonelementary* with respect to the action $G \curvearrowright X$ if the semigroup generated by the support of μ contains two independent loxodromics.

We will need the fact that a random walk on G whose increments are distributed according to a nonelementary measure μ almost surely converge to the boundary of X and has positive drift in X .

Theorem 5.9 ([MT14, Theorems 1.1, 1.2]). *Let G be a countable group which acts by isometries on a separable hyperbolic space X , and let μ be a nonelementary probability distribution on G . Fix $x \in X$. Then, almost every sample path $(w_n x)$ converges to a point in the boundary of ∂X , and the resulting hitting measure ν is nonatomic.*

Moreover, if μ has finite first moment, then there is a constant $L > 0$ such that for almost every sample path

$$\lim_{n \rightarrow \infty} \frac{d(x, w_n x)}{n} = L > 0.$$

The constant $L > 0$ in Theorem 5.9 is called the *drift* of the random walk (w_n) .

6. THE MARKOV PROCESS AND THE ACTION $G \curvearrowright X$

Let G be a hyperbolic group with a nonelementary action $G \curvearrowright X$ on a hyperbolic space X . In this section, we relate the geometry of generic geodesics in G , with respect to Patterson–Sullivan measure, to the geometry of X under the orbit map $G \rightarrow X$. These results will be used in the next section to prove genericity statements in G with respect to counting in the Cayley graph.

We begin by fixing the notation needed throughout the remainder of the paper. Let G be a hyperbolic group with finite generating set S and let Γ be the directed graph parameterizing the geodesic combing given by Theorem 5.3. The directed edges of Γ are labeled by $s \in S$ and we denote by Ω the set of all infinite paths in Γ starting at any vertex $v \in \Gamma$.

As Γ is almost semisimple, we apply the machinery of Section 3.2 to make the directed graph Γ into a Markov chain with transition probability μ (see Equation 2). This induces the measures \mathbb{P}_v on the set Ω , one for each vertex $v \in \Gamma$ of large growth. As directed paths in Γ evaluate to geodesics in G , the measure $\mathbb{P} = \mathbb{P}_{v_0}$, for v_0 the vertex associated to $1 \in G$, is thought of as a measure on the

(combing) geodesic rays in G starting at 1. The other measures \mathbb{P}_v have similar interpretations. Let

$$w_n: \Omega \rightarrow G$$

be the variable which associates to each path p in Ω the element of G spelled by concatenating the first n edges of p . That is, $\mathbb{P}_v(w_n = g)$ is the probability of hitting $g \in G$ after n steps starting at $v \in \Gamma$.

We will apply the notation and terminology developed in Sections 2, 3, and 4 to the Markov chain Γ . For example, for a subset $A \subset G$, $P^n(A)$ is the proportion of elements in the sphere S_n of radius n about $1 \in G$ which are contained in A . We also fix $\lambda > 0$ to be the growth rate of G , as in Theorem 5.4, and note that λ is exactly the growth of Γ , as defined in Section 2. Finally, for $g \in G$, we set

$$\text{cone}(g) = \text{ev}(\text{cone}(\gamma)) \subset G \cup \partial G,$$

where $\gamma \in \tilde{\Gamma}$ with $\text{ev}(\gamma) = g$. Informally, $\text{cone}(g) \subset G \cup \partial G$ is the set of points which are reachable by geodesics of G (parameterized by Γ) which pass through g .

6.1. Patterson–Sullivan measure as a combination of harmonic measures. We begin by realizing the Patterson–Sullivan measure on ∂G as a combination of harmonic measures associated to random walks.

In Section 4, to each vertex v in Γ we associated a loop semigroup L_v . Under evaluation, these semigroups map to subsemigroups of G , which we denote by Γ_v . In other words, Γ_v is the semigroup of elements of G which can be spelled by directed loops at the vertex $v \in \Gamma$. Hence, the first return measure μ_v on L_v induces a probability measure, still denoted μ_v , on $\Gamma_v \leq G$, which has finite exponential moment (see eq. (5)).

Let \mathcal{R} be the set of recurrent vertices of Γ and let ν_v be the harmonic measure on ∂G corresponding to the random walk μ_v on G . Let N_v be the set of finite directed paths in Γ beginning at v_0 and ending at v which do not meet a recurrent vertex of Γ in their interior.

Proposition 6.1. *The Patterson–Sullivan measure ν on ∂G is a combination of the harmonic measures ν_v :*

$$\nu = \sum_{v \in \mathcal{R}} \sum_{\gamma \in N_v} \mu(\gamma) \gamma_* \nu_v.$$

In the statement of Proposition 6.1, we have identified $\gamma \in N_v$ with its evaluation $\text{ev}(\gamma) \in G$.

Proof. Let $\Pi \subset \Omega_0$ be the set of infinite paths which enter some maximal component and visit each vertex of that component infinitely often. Let $\gamma \in \Pi$, and v be the first recurrent vertex which is met along γ . Then γ can be decomposed uniquely as $\gamma = \gamma_0 \gamma_1 \dots$ where γ_0 is a path from v_0 to v which does not meet any recurrent vertex in its interior, and each γ_i for $i \geq 1$ is a primitive loop at v . Thus one can define the map $\varphi: \Pi \rightarrow \tilde{\Gamma} \times \bigcup_v (\Gamma_v)^\mathbb{N}$ as $\varphi(\gamma) = (\gamma_0, \gamma_1, \dots, \gamma_n, \dots)$. Since the set Π has full \mathbb{P} -measure (Lemma 3.2), and the Patterson–Sullivan measure ν on $\partial \tilde{\Gamma}$ is the same as \mathbb{P} under the identification $\partial \tilde{\Gamma} = \Omega_0$, the map φ can be defined as a measurable, ν -almost surely defined map $\varphi: \partial \tilde{\Gamma} \rightarrow \tilde{\Gamma} \times \bigcup_v (\Gamma_v)^\mathbb{N}$.

Moreover, one has the boundary map $\text{bnd}: G \times G^\mathbb{N} \rightarrow \partial G$ defined as

$$\text{bnd}(g_0, (g_1, \dots, g_n, \dots)) = \lim_{n \rightarrow \infty} g_0 g_1 \dots g_n$$

(whenever the limit exists).

Combining these maps with the evaluation map, we get the following commutative diagram of measurable maps:

$$\begin{array}{ccc} \partial\tilde{\Gamma} & \xrightarrow{\text{ev}} & \partial G \\ \downarrow \varphi & & \uparrow \text{bnd} \\ \tilde{\Gamma} \times \bigcup_v (\Gamma_v)^\mathbb{N} & \xrightarrow{\text{ev}} & G \times G^\mathbb{N} \end{array}$$

The claim follows by pushing forward the Patterson-Sullivan measure ν on $\partial\tilde{\Gamma}$ along the diagram. First, by disintegration we get

$$\nu = \sum_{v \in \mathcal{R}} \sum_{\gamma \in N_v} \mu(\gamma) \nu_\gamma$$

where ν_γ is the conditional probability on the set of infinite paths which start with γ . By Lemma 4.2, $\varphi_*(\nu_\gamma) = \delta_\gamma \times (\mu_v)^\mathbb{N}$, so

$$\varphi_*(\nu) = \sum_{v \in \mathcal{R}} \sum_{\gamma \in N_v} \mu(\gamma) \left(\delta_\gamma \times (\mu_v)^\mathbb{N} \right)$$

Now, by definition the pushforward of $\delta_1 \times (\mu_v)^\mathbb{N}$ by the boundary map bnd is the harmonic measure ν_v , hence by G -equivariance $\text{bnd}_*(\delta_\gamma \times (\mu_v)^\mathbb{N}) = \gamma_* \nu_v$, yielding the claim. \square

We can also obtain an analogous statement for any vertex of large growth. Indeed, for each vertex v of large growth, there exists a measure ν_v on ∂G which is the hitting measure of the Markov chain on ∂G , starting from the identity element on G .

Then we have

$$(6) \quad \nu_v = \sum_{w \in \mathcal{R}} \sum_{\gamma: v \rightarrow w} \mu(\gamma) \gamma_* \nu_w$$

Here, the sum is over all finite paths from v to w which only meet a recurrent vertex at their terminal endpoint. Note that if v is recurrent, then ν_v is the harmonic measure for the random walk on G generated by the measure μ_v , as discussed above.

6.2. The loop semigroup is nonelementary. Now suppose that X is a hyperbolic space and that $G \curvearrowright X$ is a nonelementary action. In this section, we show:

Proposition 6.2. *There is a recurrent vertex v of Γ such that the corresponding loop semigroup Γ_v is nonelementary.*

We define the *limit set* of the loop semigroup Γ_v on the boundary of G to be $\Lambda_{\Gamma_v} = \overline{\Gamma}_v \cap \partial G$, while the limit set on the boundary of X is $\Lambda_{\Gamma_v}^X = \overline{\Gamma}_v x \cap \partial X$.

For an irreducible component C of Γ , let $\mathcal{P}_v(C)$ be the set of finite paths in Γ which are based at v and lie entirely in C , and $P_v(C) = \text{ev}(\mathcal{P}_v(C))$. Similarly, $\partial\mathcal{P}_v(C)$ is the set of infinite paths in Γ which are based at v and lie entirely in C , and $\partial P_v(C) = \text{ev}(\partial\mathcal{P}_v(C))$. We remind the reader that ev is injective on directed paths in Γ which start at a fixed vertex v .

Lemma 6.3. *There exists a maximal component C of Γ and a vertex v in C such that $\partial P_v(C)$ contains an open set on ∂G .*

Proof. Every infinite path starting at the origin eventually stays forever in some component, hence we can write $\partial\tilde{\Gamma}$ as the countable disjoint union

$$\partial\tilde{\Gamma} = \bigcup_{i=1}^{\infty} \gamma_i \cdot \partial\mathcal{P}_{v_i}(C_i)$$

where γ_i is a path from the initial vertex to v_i , and C_i is the component of v_i , hence by applying the evaluation map we get

$$\partial G = \bigcup_{i=1}^{\infty} g_i \cdot \partial P_{v_i}(C_i)$$

with $g_i = \text{ev}(\gamma_i)$. Since ∂G is a Baire space and all $\partial P_{v_i}(C_i)$ are closed, then there exists some i such that $\partial P_{v_i}(C_i)$ contains an open set on ∂G .

For a shadow S in G , let $\partial S \subseteq \partial G$ be the set of equivalence classes of geodesic rays based at 1 which have a representative (g_n) which eventually lies in S (i.e. for which there exists n_0 such that $g_n \in S$ for all $n \geq n_0$).

If $\partial P_v(C)$ contains a (non-empty) open set, then by definition of the topology on ∂G there exists a nested pair of shadows $S = S_1(g, R)$ and $S' = S_1(g, R+r)$ such that $\partial P_v(C) \supseteq \partial S' \supseteq \partial S$, and r is given by Lemma 5.7. Note that by Lemma 5.8 S has maximal growth. We claim that this implies that for a constant D which depends only on δ and r we have

$$S \subseteq N_D(P_v(C)).$$

Then, since S has maximal growth, so does $N_D(P_v(C))$, hence $P_v(C)$ also has maximal growth, proving the lemma. To prove the claim, let $h \in S$. Then by Lemma 5.7, there exists h' with $d_G(h, h') \leq r$ and such that h' is continuable. By Lemma 5.6, h' belongs to S' , and since it is continuable h' lies on a geodesic ray γ based at 1 with endpoint $\zeta \in \partial G$. As $h' \in S'$, so must each vertex of γ after h' , and so we see that $\zeta \in \partial S'$. Since $\partial S' \subseteq \partial P_v(C)$, there also exists an infinite geodesic ray γ' based at 1 which lies entirely in $P_v(C)$ and terminates at $\zeta \in \partial G$. Since γ and γ' are both geodesic rays in G from 1 to $\zeta \in \partial G$, they 2δ -fellow travel. Thus, h belongs to a $(2\delta + r)$ -neighborhood of $P_v(C)$. Setting $D = 2\delta + r$ establishes the claim. \square

Lemma 6.4. *Let v be a vertex in the irreducible component C . Then*

$$\Lambda_{\Gamma_v} \supseteq \partial P_v(C).$$

Proof. Let $\xi \in \partial P_v(C)$, and (g_n) a sequence of elements of $P_v(C)$ so that $g_n \rightarrow \xi$. Then there exists for each n a path s_n from the endpoint of g_n to v , so $g_n s_n \in \Gamma_v$ and $|s_n| \leq D$, where D is the diameter of the component C . Then $d(g_n s_n, g_n) \leq D$, hence $g_n s_n$ also converges to ξ , proving the claim. \square

Lemma 6.5. *If C is a maximal irreducible component and v belongs to C , then the growth of Γ_v is maximal.*

Proof. Let D be the diameter of C . Since every path from v to any vertex v_i inside C of length n can be extended to a path of length $\leq n + D$ from v to itself, we have for each n

$$\#(P_v(C) \cap B_n) \leq \#(\Gamma_v \cap B_{n+D})$$

which, together with Lemma 2.3, proves the claim. \square

We can now give the proof of Proposition 6.2:

Proof of Proposition 6.2. Let v be the vertex given by Lemma 6.3. We first claim that the semi-group Γ_v does not have a finite orbit in ∂X . Otherwise, the subgroup $H = \langle \Gamma_v \rangle$ also has a finite orbit in ∂X . Since C is maximal, by Lemma 6.5 the semigroup Γ_v has maximal growth, hence $|B(r) \cap \Gamma_v| \asymp \lambda^r$, where λ is the growth rate of G ; then [GMM15, Theorem 4.3] implies that the subgroup H must have finite index in G . This implies that the action $G \curvearrowright \partial X$ has a finite orbit, contradicting that $G \curvearrowright X$ is nonelementary.

Using Proposition 5.2, it remains to show that $\Lambda_{\Gamma_v}^X$ contains at least 3 points. To this end, let U be an open set of ∂G contained in $\Lambda_{\Gamma_v} \subset \partial G$, which exists by Lemma 6.4. Now, the set of fixed points of loxodromic elements is a G -invariant subset of ∂G , and since the action of G on its boundary

is minimal, this set is dense. Hence, we can find $f, g \in G$, which are independent loxodromics with respect to the action $G \curvearrowright X$ and with $f^{\pm\infty}, g^{\pm\infty} \in U \subset \partial G$. Indeed, for any independent X -loxodromic elements f, g choose $u \in G$ with $u^{+\infty} \in U$ and $u^{-\infty}$ distinct from $f^{\pm\infty}, g^{\pm\infty}$ in ∂G . (If some chosen $u^{+\infty} \in U$ has $u^{-\infty}$ equal to one of $f^{\pm\infty}, g^{\pm\infty}$, then replace U with some open set $U_1 \subset U$ not containing $u^{+\infty}$ and choose $u_1 \in G$ with $u_1^{+\infty} \in U_1$. Since two infinite order elements of G cannot share exactly one limit point on ∂G , applying this procedure at most 4 times gives the required element.) Then after replacing f, g by $u^k f u^{-k}, u^k g u^{-k}$ for sufficiently large k , we obtain our desired pair of independent X -loxodromics, all of whose limit points on ∂G lie in U .

We claim that each of $f^{\pm\infty}, g^{\pm\infty} \in U \subset \partial G$ gives points of $\Lambda_{\Gamma_v}^X$. Since f, g are independent and loxodromic with respect to the action $G \curvearrowright X$, these 4 points must be distinct points of $\Lambda_{\Gamma_v}^X$, completing the proof.

Since $f^\infty \in U \subset \Lambda_{\Gamma_v}$, there is a sequence $l_i \in \Gamma_v$ with $l_i \rightarrow f^\infty$ in $G \cup \partial G$ as $i \rightarrow \infty$. As $\tilde{\Gamma} \cup \partial\tilde{\Gamma}$ is compact, after passing to a subsequence, we may assume that $l_i \rightarrow l_\infty$ in $\tilde{\Gamma} \cup \partial\tilde{\Gamma}$ and by continuity, $\text{ev}(l_\infty) = f^\infty$. Let g_i be element of G represented by the i th term of l^∞ (realizing l^∞ as an infinite path in $\tilde{\Gamma}$). Then $(g_i)_{i \geq 0}$ is a geodesic in G converging to $f^\infty \in \partial G$ and so there is a K , depending only on f , such that

$$d_G(g_i, f^{j_i}) \leq K,$$

for some sequence j_i which goes to ∞ as $i \rightarrow \infty$. Since the (fixed) orbit map $G \rightarrow X$ is L -coarsely Lipschitz,

$$d_X(g_i x, f^{j_i} x) \leq KL.$$

As $f \in G$ is loxodromic for the X -action, $f^{j_i} x$ converges to some point $f_X^\infty \in \partial X$, hence so does $g_i x$. We conclude $f_X^\infty \in \Lambda_{\Gamma_v}^X$, as required. \square

Remark 6.6. If it is a priori known that no element $g \in G$ acts parabolically on X (i.e. g is either loxodromic or has a bounded orbit), then the use of [GMM15, Theorem 4.3] in Proposition 6.2 (and in Corollary 6.11) can be removed.

6.3. Convergence to the boundary of X . In this section we show that almost every sample path for the Markov chain converges to the boundary of X . Since sample paths in the Markov chain evaluate to geodesic rays in G , this will show that the orbit of almost every geodesic ray in G converges to the boundary of X .

Let $\partial\tilde{\Gamma}^X$ be the set of infinite paths beginning at v_0 which converge to a point in ∂X when projected to X and set $\partial^X G = \text{ev}(\partial\tilde{\Gamma}^X)$. Since the orbit map $G \rightarrow X$ is Lipschitz, $\partial^X G \subset \partial G$ consists of those ξ such that the projection of any geodesic ray $[1, \xi) \subset G$ to X converges in $X \cup \partial X$ to a point of ∂X .

Lemma 6.7. *The set $\partial^X G$ is G -invariant.*

Proof. If $\eta \in \partial^X G$, then $[1, \eta)$ is a geodesic in G which projects under the orbit map to a path γ in X which converges to a point ξ in ∂X . Hence, $g\gamma$ converges to $g\xi$ in ∂X . Since $g[1, \eta)$ is a geodesic in G with the same endpoint as $[1, g\eta)$, they R -fellow-travel in G for some $R \geq 0$. Hence, the projection γ' of $[1, g\eta)$ to X must RL -fellow-travel the path $g\gamma$ in X , where L is the Lipschitz constant of the fixed orbit map $G \rightarrow X$. Hence, γ' converges to the point $g\xi$ in $X \cup \partial X$ showing that $g\eta \in \partial^X G$ as required. \square

Here we show that $\partial^X G$ has full measure in ∂G , so that the map $\Phi: \partial^X G \rightarrow \partial X$ gives a measurable map $\partial G \rightarrow \partial X$.

Theorem 6.8. *For \mathbb{P} -almost every path (w_n) in the Markov chain, the projection $(w_n x)$ to the space converges to a point in the boundary ∂X .*

Proof. We know by Lemma 3.2 that \mathbb{P} -almost every sample path enters some recurrent component C , and visits each vertex there infinitely often. Let v be a vertex in C . Then for each v , one can split the infinite path into a prefix γ_0 , and a sequence $\gamma_1, \gamma_2, \dots$ of primitive loops at v . Then the pushforward of \mathbb{P}_C via the map φ_v equals the measure $\mu_v^{\mathbb{N}}$ on the set $\Gamma_v^{\mathbb{N}}$ (as in Theorem 6.1). Now, if the measure μ_v is non-elementary, then by Theorem 5.9 we know that almost surely the sequence $(\gamma_1 \gamma_2 \dots \gamma_n x)$ converges to a point in ∂X . Since the same is true for every vertex of C , then the sequence (w_n) is partitioned in a finite number of subsequences, and for each such subsequence (w_{n_k}) the sample path $(w_{n_k} x)$ converges to a point in ∂X . Thus, we define the equivalence relation on the set of vertices of C by defining $v_i \sim v_j$ if the sequences $(w_{n(k,v_i)} x)$ and $(w_{n(k,v_j)} x)$ converge to the same point. Now, we know that if $w_{n_k} x \rightarrow \xi \in \partial X$, then also $w_{n_k+1} x \rightarrow \xi$, so this equivalence relation satisfies the hypothesis of Lemma 6.9, hence as a consequence there is only one equivalence class, thus the whole sequence $(w_n x)$ converges to the same $\xi \in \partial X$.

Since there exists at least one (recurrent) vertex v such that the semigroup Γ_v is nonelementary by Proposition 6.2, the previous argument shows that $\nu(\partial \tilde{\Gamma}^X) > 0$, thus $\nu(\partial^X G) > 0$. Since $G \curvearrowright (\partial G, \nu)$ is ergodic (Lemma 5.5) and $\partial^X G$ is invariant, then $\partial^X G$ has full measure, which completes the proof of the theorem. \square

Lemma 6.9. *Let $r \geq 1$, and \sim be an equivalence relation on the set $\{1, \dots, r\}$, and $\mathbb{N} = A_1 \sqcup \dots \sqcup A_r$ a partition of the set of natural numbers into finitely many, disjoint infinite sets. Assume moreover that if the intersection $(A_i + 1) \cap A_j$ is infinite, then $i \sim j$. Then there is only one equivalence class.*

Proof. Suppose that there are at least two equivalence classes, and pick one equivalence class. Let X be the union of all sets A_i which belong to such a class, and Y the union of all other equivalence classes. Then $\mathbb{N} = X \sqcup Y$, with both X and Y infinite. Then $(X + 1) \cap Y$ is infinite, which implies that there exists a set $A_i \subseteq X$ and a set $A_j \subseteq Y$ such that $|(A_i + 1) \cap A_j| = \infty$, contradicting the fact that i and j are in different equivalence classes. \square

Theorem 1.3 now follows immediately as a corollary to Theorem 6.8.

Corollary 6.10 (Convergence to the boundary of X). *For every $x \in X$ and ν -almost every $\eta \in \partial G$, if $(g_n)_{n \geq 0}$ is a geodesic in G converging to η , then the sequence $g_n x$ in X converges to a point in the boundary ∂X .*

Theorem 6.8 was proven using only the knowledge that for *some* recurrent vertex v , Γ_v was nonelementary. However, using Theorem 6.8 we can now show that for *every* recurrent vertex, Γ_v is nonelementary. This will be crucial for what follows.

Corollary 6.11. *Suppose that $G \curvearrowright X$ is as above. Then for each recurrent vertex v of Γ , the semigroup Γ_v is nonelementary with respect to the action on X .*

Proof. Since Γ_v has maximal growth, $\nu(\partial \mathcal{P}_v(C)) > 0$. Then, identifying $\partial \mathcal{P}_v(C)$ with the corresponding endpoints in $\partial \tilde{\Gamma}$, we have that $\partial \mathcal{P}_v(C) \cap \partial \tilde{\Gamma}^X$ is infinite since $\nu(\partial \tilde{\Gamma}^X) = 1$. Hence, so is $\text{ev}(\partial \mathcal{P}_v(C) \cap \partial \tilde{\Gamma}^X) \subset \Lambda_{\Gamma_v} \cap \partial^X G$ and we conclude that $\Lambda_{\Gamma_v}^X$ is infinite.

Just as in the proof of Proposition 6.2, the subgroup $H = \langle \Gamma_v \rangle$ has finite index in G and so Γ_v cannot have a finite orbit on ∂X . Since $\Lambda_{\Gamma_v}^X \neq \emptyset$, Theorem 5.2, implies that Γ_v is nonelementary. \square

Similar to the discussion in Section 6.1, we denote as ν_v^X the corresponding hitting measure on ∂X (which is well-defined since we already proved the Markov process converges to the boundary). By equation (6) and applying convergence to the boundary we get:

$$(7) \quad \nu_v^X = \sum_{w \in \mathcal{R}} \sum_{\gamma: v \rightarrow w} \mu(\gamma) \gamma_* \nu_w^X$$

Lemma 6.12. *For any v of large growth, the measure ν_v^X is non-atomic.*

Proof. Since the random walk measures ν_w^X are non-atomic, so are the measures $\gamma_*\nu_w^X$ for each γ , hence by equation (7) the measure ν_v is also non-atomic as it is a linear combination of non-atomic measures. \square

6.4. Positive drift along geodesics. In this section, we show that almost every sample path has positive drift in X . This will imply (Corollary 6.17) the same for almost every geodesic ray in G with respect to Patterson–Sullivan measure.

We first show that there is a well-defined average return time for a recurrent vertex v of Γ :

Lemma 6.13. *For each vertex v in a recurrent component C , there exists $T_v > 0$ such that for almost every $\omega \in \Omega_C$*

$$\lim_{k \rightarrow \infty} \frac{n(k, v, \omega)}{k} = T_v.$$

Proof. We know by Lemma 3.2 that the return time to v has an exponential tail:

$$\mathbb{P}_v(\tau(1, v, \omega) \geq n) \leq ce^{-cn}$$

Thus, the average return time is finite:

$$\int \tau(1, v, \omega) d\mathbb{P}_v(\omega) = T_v < \infty$$

which implies that μ_v has finite first moment in the word metric:

$$\int |g| d\mu_v(g) = T_v.$$

Now, by definition

$$n(k, v, \omega) = \tau(0, v, \omega) + \tau(1, v, \omega) + \cdots + \tau(k, v, \omega)$$

then since the variables $\tau(1, v, \omega), \dots, \tau(k, v, \omega)$ are independent and identically distributed, we get almost surely

$$\lim_{k \rightarrow \infty} \frac{n(k, v, \omega)}{k} = T_v.$$

\square

Using these average return times, we show that sample paths in the Markov chain have a well-defined drift and hence make linear progress in X .

Theorem 6.14. *There exists $L > 0$ such that for \mathbb{P} -almost every sample path (w_n) we have*

$$\lim_{n \rightarrow \infty} \frac{d(w_n x, x)}{n} = L.$$

Proof. Let v be a recurrent vertex. Then by Proposition 6.11 the loop semigroup Γ_v is nonelementary, hence the random walk given by the return times to v has positive drift. More precisely, from Theorem 5.9, there exists a constant $\ell_v > 0$ such that for almost every sample path which enters v ,

$$\lim_{k \rightarrow \infty} \frac{d(w_{n(k,v)} x, x)}{k} = \ell_v.$$

Combined with Lemma 6.13, this implies

$$\lim_{k \rightarrow \infty} \frac{d(w_{n(k,v)} x, x)}{n(k, v)} = \frac{\ell_v}{T_v}.$$

Now, almost every infinite path visits every vertex of some recurrent component infinitely often. Thus, for each recurrent vertex v_i which belongs to a component C , there exists a constant $L_i > 0$ such that for \mathbb{P}_C -almost every path (w_n) , there is a limit

$$L_i = \lim_{k \rightarrow \infty} \frac{d(w_{n(k,v_i)}x, x)}{n(k, v_i)}.$$

Let C be a maximal component, and v_1, \dots, v_k its vertices. Our goal now is to prove that $L_1 = L_2 = \dots = L_k$. Let us pick a path $\omega \in \Omega_0$ such that the limit L_i above exists for each $i = 1, \dots, k$, and define $A_i = \{n(k, v_i), k \in \mathbb{N}\}$, and the equivalence relation $i \sim j$ if $L_i = L_j$. Since $w_{n(k,v_i)}$ and $w_{n(k,v_i)+1}$ differ by one generator, $d(w_{n(k,v_i)}x, w_{n(k,v_i)+1}x)$ is uniformly bounded, hence

$$\lim_{k \rightarrow \infty} \frac{d(w_{n(k,v_i)+1}x, x)}{n(k, v_i) + 1} = \lim_{k \rightarrow \infty} \frac{d(w_{n(k,v_i)}x, x)}{n(k, v_i)} = L_i$$

so the equivalence relation satisfies the hypothesis of Lemma 6.9, hence there is a unique limit $L = L_i$ so that

$$\lim_{n \rightarrow \infty} \frac{d(w_n x, x)}{n} = L.$$

In order to prove that the drift is the same for all maximal components, let us now pick $L = L_i$ for some recurrent vertex v_i , and consider the set ∂G_L of points ξ in ∂G such that there exists a geodesic $(g_n) \subseteq G$ with $g_n \rightarrow \xi$ and such that $\lim_{n \rightarrow \infty} \frac{d(g_n x, x)}{n} = L$. The set ∂G_L is G -invariant, and by the above statement it has positive probability, hence by ergodicity it has full measure. \square

Theorem 6.15. *For every vertex v of large growth, and for \mathbb{P}_v -almost every sample path (w_n) we have*

$$\lim_{n \rightarrow \infty} \frac{d(w_n x, x)}{n} = L.$$

Proof. By Theorem 6.14, for \mathbb{P} -almost every path which passes through v , the drift equals L . Each such path $\gamma = (w_n)$ can be decomposed as $\gamma = \gamma_0 \star \tilde{\gamma}$, where γ_0 is a finite path from v_0 to v and $\tilde{\gamma} = (\tilde{w}_n)$ is an infinite path starting from v . Thus for each n , $w_n = g_0 \tilde{w}_n$, where g_0 is the group element which represents γ_0 . Hence,

$$d(w_n x, x) - d(x, g_0 x) \leq d(\tilde{w}_n x, x) = d(w_n x, g_0 x) \leq d(w_n x, x) + d(x, g_0 x)$$

and so

$$\lim_{n \rightarrow \infty} \frac{d(\tilde{w}_n x, x)}{n} = \lim_{n \rightarrow \infty} \frac{d(w_n x, x)}{n} = L$$

as required. \square

For application in Section 7, we will need the following convergence in measure statement.

Corollary 6.16. *For any $\epsilon > 0$, and for any v of large growth,*

$$\mathbb{P}_v \left(\left| \frac{d(x, w_n x)}{n} - L \right| \geq \epsilon \right) \rightarrow 0.$$

Proof. By the theorem, the sequence of random variables $X_n = \frac{d(w_n x, x)}{n}$ converges almost surely to L . Moreover, for every n the variable X_n is bounded above by the Lipschitz constant of the orbit map $G \rightarrow X$. Thus, X_n converges to L in L^1 , yielding the claim. \square

Finally, Theorem 6.14 implies the following linear progress statement for almost every geodesic ray in G . This proves Theorem 1.4 from the introduction.

Corollary 6.17 (Positive drift). *For every $x \in X$ and ν -almost every $\eta \in \partial G$, if $(g_n)_{n \geq 0}$ is a geodesic in G converging to η , then*

$$\lim_{n \rightarrow \infty} \frac{d(x, g_n x)}{n} = L > 0.$$

Proof. Note that the measure \mathbb{P} on the set of infinite paths starting at 1 pushes forward to ν on ∂G . Thus, by Theorem 6.14, for ν -almost every $\xi \in \partial G$ there exists a geodesic (g_n) in the combing with $g_n \rightarrow \xi$ for which the statement is true; by hyperbolicity, any other geodesic (g'_n) which converges to ξ fellow travels (g_n) , hence the drift in X is the same. \square

6.5. Sublinear tracking. We are now ready to prove Theorem 1.5 from the introduction.

Proof of Theorem 1.5. It follows essentially from the argument in the proof of Theorem 6.14, together with the sublinear tracking result for random walks [MT14, Theorem 1.3]. Indeed, \mathbb{P} -almost every sample path visits all vertices in a certain recurrent component infinitely often, hence the sequence $(w_n)_{n \in \mathbb{N}}$ can be partitioned into finitely many subsequences $(w_{n(k,v)})_{k \in \mathbb{N}}$, one for each such vertex v , and by [MT14, Theorem 1.3] each subsequence tracks some quasigeodesic ray in X with sublinear error; more precisely, for each v there exists a quasigeodesic ray γ_v in X such that

$$\lim_{k \rightarrow \infty} \frac{d(w_{n(k,v)} x, \gamma_v)}{k} = 0,$$

and since $\lim_{k \rightarrow \infty} \frac{n(k,v)}{k}$ exists and is finite almost surely, one also has

$$\lim_{k \rightarrow \infty} \frac{d(w_{n(k,v)} x, \gamma_v)}{n(k,v)} = 0.$$

Note that in the proof of [MT14, Theorem 1.3] one chooses γ_v so that $\lim_{k \rightarrow \infty} w_{n(k,v)} x = \lim_{t \rightarrow \infty} \gamma_v(t)$. Now, we already proved that the whole sequence (w_n) converges, hence all limits $\lim_{t \rightarrow \infty} \gamma_v(t)$ must be independent of v . By definition of Gromov boundary, this implies that the quasigeodesic rays γ_v lie within bounded distance of each other, hence one can pick any such quasigeodesic ray $\gamma = \gamma_v$ and get

$$\lim_{n \rightarrow \infty} \frac{d(w_n x, \gamma)}{n} = 0$$

as required. \square

6.6. Decay of shadows for \mathbb{P} . Let us denote as $Sh(x, r)$ the set of shadows of the form $S_x(gx, R)$ where $g \in G$ and the distance parameter satisfies $d(x, gx) - R \geq r$. For any shadow S , we denote its closure in $X \cup \partial X$ by \bar{S} .

Proposition 6.18 (Decay of shadows). *There exists a function $f : \mathbb{R} \rightarrow \mathbb{R}$ such that $f(r) \rightarrow 0$ as $r \rightarrow \infty$, and such that for each v of large growth,*

$$\nu_v^X(\bar{S}_x(gx, R)) \leq f(r),$$

where $r = d(x, gx) - R$ is the distance parameter of the shadow.

Proof. By Lemma 6.12, the measure ν_v^X on ∂X is nonatomic. The proposition now follows from a standard measure theory argument, as in Proposition 5.1 of [MT14]. \square

Let Ω be the set of all infinite paths in Γ , starting from any vertex. We define for each n the map $w_n : \Omega \rightarrow G$ which maps each infinite path to the product of the first n edges of the path:

$$w_n((e_1, \dots, e_n, \dots)) = e_1 \dots e_n$$

and

$$w_\infty(x) := \lim_{n \rightarrow \infty} w_n x \in \partial X$$

(if it exists). Recall that by definition $\nu_v^X(A) = \mathbb{P}_v(w_\infty(x) \in A)$.

Proposition 6.19. *There exists a function $p : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ with $p(r) \rightarrow 0$ as $r \rightarrow \infty$, such that for each vertex v and any shadow $S_x(gx, R)$ we have*

$$\mathbb{P}_v\left(\exists n \geq 0 : w_n x \in S_x(gx, R)\right) \leq p(r),$$

where $r = d(x, gx) - R$ is the distance parameter of the shadow.

Proof. Let $S = S_x(gx, R)$ and $S_1 = S_x(gx, R + c_1)$, for c_1 to be determined in a moment. The idea is that for c_1 sufficiently large, the probability that the sample path $(w_n x)$ ever reaches the shadow S is dominated by the probability that it converges to a point of $\overline{S_1} \cap \partial X$.

For any $h \in G$ such that $hx \in S$, hyperbolicity of X implies that there is a constant $c_2 \geq 0$ with $c_2 = c_1 + O(\delta)$ such that the complement of S_1 is contained in the shadow

$$S^h = S_{hx}(x, d(x, hx) - c_2).$$

Note that the distance parameter of S^h is $c_2 = c_1 + O(\delta)$. Then since $h^{-1}S^h$ is a shadow based at x , Proposition 6.18 implies that we can choose c_1 large enough so that

$$\nu_w^X(h^{-1}S^h) \leq f(c_1 + O(\delta)) \leq 1/2,$$

for any large growth vertex w . This is our fixed $c_1 \geq 0$.

Now we have

$$\mathbb{P}_v(w_\infty(x) \in \overline{S_1}) \geq \mathbb{P}_v(w_\infty(x) \in \overline{S_1} \text{ and } \exists n : w_n x \in S) =$$

If there exists n such that $w_n x$ belongs to S , let n_1 be the first hitting time. Then

$$= \sum_{hx \in S} \mathbb{P}_v(w_{n_1} = h) \mathbb{P}_v(w_\infty(x) \in \overline{S_1} \mid w_{n_1} = h) =$$

hence from the Markov property of the chain

$$= \sum_{hx \in S} \mathbb{P}_v(w_{n_1} = h) \mathbb{P}_{v^h}(w_\infty(x) \in h^{-1}\overline{S_1}),$$

where v^h is terminal vertex of the path starting at v which reads the words h .

Since the complement of S_1 is contained in S^h , $(\partial X - h^{-1}\overline{S_1}) \subset h^{-1}S^h$ and so

$$\nu_w^X(h^{-1}\overline{S_1}) \geq 1 - \nu_w^X(h^{-1}S^h) \geq 1 - f(c_1 + O(\delta)) \geq 1/2,$$

by our fixed choice of c_1 . Hence, combining this with the previous estimate,

$$\mathbb{P}_v(w_\infty(x) \in \overline{S_1}) \geq \sum_{hx \in S} \mathbb{P}_v(w_{n_1} = h) \cdot \nu_{v^h}^X(h^{-1}\overline{S_1}) \geq \frac{1}{2} \mathbb{P}_v(\exists n : w_n x \in S)$$

hence

$$\mathbb{P}_v(\exists n : w_n x \in S) \leq 2 \cdot \mathbb{P}_v(w_\infty(x) \in \overline{S_1}) \leq 2 \cdot f(d(x, gx) - R - c_1)$$

so the claim is proven by taking $p(r) := 2 \cdot f(r - c_1)$. □

We conclude this section with a simple lower bound on the Gromov product:

Lemma 6.20. *Let x, y, z be points in a metric space X . Then*

$$(y, z)_x \geq d(x, z) - d(y, z).$$

Proof. By definition

$$(y, z)_x = \frac{1}{2} (d(x, y) + d(x, z) - d(y, z))$$

and by the triangle inequality $d(x, y) \geq d(x, z) - d(y, z)$, yielding the claim. □

6.7. Almost independence and Gromov products. Let \mathbb{P} denote the probability measure (on the space of infinite paths) for the Markov chain starting at the identity. Given a vertex v , we will denote as $\mathbb{P}_v(w_n = g)$ the probability that the Markov chain starting at the vertex v reads g after n steps. Let $m = \lfloor n/2 \rfloor$. Finally, we denote as u_m the random variable $u_m = w_m^{-1}w_n$, so that $w_n = w_m u_m$. Recall that by the Markov property of the chain, for any $g, h \in G$ and each m ,

$$\mathbb{P}(w_m = g \text{ and } u_m^{-1} = h) = \mathbb{P}(w_m = g) \cdot \mathbb{P}_v(w_m = h^{-1})$$

where $v = [g]$. Let V denote the vertex set of Γ .

The following three lemmas will be essential in controlling the size of $(gx, g^{-1}x)_x$ over the Markov chain (Lemma 6.26), which is needed to produce loxodromic elements of the action $G \curvearrowright X$.

Lemma 6.21. *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be any function such that $f(m) \rightarrow +\infty$ as $m \rightarrow +\infty$. Then*

$$\mathbb{P}((w_m x, u_m^{-1} x)_x \geq f(m)) \rightarrow 0$$

as $m \rightarrow \infty$.

Proof. Let us denote as $S(m) := \{(g, h) \in G \times G : (gx, hx)_x \geq f(m)\}$. Then by definition we have

$$\mathbb{P}((w_m x, u_m^{-1} x)_x \geq f(m)) = \sum_{g, h \in S(m)} \mathbb{P}(w_m = g \text{ and } u_m = h^{-1}) \leq$$

hence by the Markov property

$$\leq \sum_{v \in V} \sum_{h \in G} \mathbb{P}_v(w_m = h^{-1}) \sum_{\{g: [g]=v, (g, h) \in S(m)\}} \mathbb{P}(w_m = g) \leq$$

and forgetting the requirement that $[g] = v$

$$\leq \sum_{v \in V} \sum_{h \in G} \mathbb{P}_v(w_m = h^{-1}) \mathbb{P}(w_m x \in S_x(hx, R)) \leq$$

where $R = d(x, hx) - f(m)$, while using the estimate on shadows (Proposition 6.19) we have

$$\leq \sum_{v \in V} \sum_{h \in G} \mathbb{P}_v(w_m = h^{-1}) p(f(m)) \leq \#V \cdot p(f(m))$$

and the claim follows by decay of shadows. □

Lemma 6.22. *For each $\eta > 0$, the probability*

$$\mathbb{P}((w_n^{-1} x, u_m^{-1} x)_x \leq n(L - \eta)/2)$$

tends to 0 as $n \rightarrow \infty$.

Proof. By Lemma 6.20, and since the group acts by isometries,

$$(w_n^{-1} x, u_m^{-1} x)_x \geq d(x, w_n^{-1} x) - d(w_n^{-1} x, u_m^{-1} x) = d(w_n x, x) - d(x, w_m x)$$

Now, by Theorem 6.14,

$$\mathbb{P}(d(w_n x, x) \geq n(L - \epsilon)) \rightarrow 1 \quad \text{and} \quad \mathbb{P}(d(w_m x, x) \leq n/2(L + \epsilon)) \rightarrow 1$$

hence with probability which tends to 1

$$(w_n^{-1} x, u_m^{-1} x)_x \geq n(L - \epsilon) - \frac{n}{2}(L + \epsilon) = \frac{n}{2}(L - 3\epsilon)$$

which yields the claim setting $\eta = 3\epsilon/2$. □

Finally, we show

Lemma 6.23. *For each $\eta > 0$, the probability*

$$\mathbb{P}((w_mx, w_nx)_x \leq n(L - \eta)/2)$$

tends to 0 as $n \rightarrow \infty$.

Proof. By Lemma 6.20,

$$(w_nx, w_mx)_x \geq d(x, w_nx) - d(w_mx, w_nx) = d(x, w_nx) - d(x, u_mx)$$

so if we have $d(x, w_nx) \geq n(L - \epsilon)$ and $d(x, u_mx) \leq n(L + \epsilon)/2$ then

$$(w_nx, w_mx)_x \geq n(L - 3\epsilon)/2$$

Now, the first statement holds with probability which tends to 1 by Theorem 6.14; for the second statement, note that for each g and each m ,

$$\mathbb{P}(u_m = g) = \sum_{v \in V} \mathbb{P}([w_m] = v) \mathbb{P}_v(w_m = g)$$

and by Theorem 6.15

$$\mathbb{P}_v(d(x, w_mx) \leq n(L + \epsilon)/2) \rightarrow 1$$

for each recurrent v . Hence

$$\mathbb{P}(d(x, u_mx) \leq n(L + \epsilon)/2) \rightarrow 1$$

completing the proof. \square

6.8. Linear translation length along geodesics. For $g \in G$, let $\tau_X(g)$ be the (stable) translation length of g with respect to its action on X . Recall that $\tau_X(g) > 0$ if and only if g is loxodromic for the action $G \curvearrowright X$.

The following lemma for estimating the translation length of an isometry g of X is well-known; see for example [MT14, Proposition 5.8].

Lemma 6.24. *There exists a constant c , which depends only on δ , such that for any isometry g of a δ -hyperbolic space X and any $x \in X$ with $d(x, gx) \geq 2(gx, g^{-1}x)_x + c$, the translation length of g is given by*

$$\tau_X(g) = d(x, gx) - 2(gx, g^{-1}x)_x + O(\delta).$$

Since we already proved that the term $d(x, gx)$ grows linearly \mathbb{P} -almost surely (Theorem 6.16), to complete the proof we need to show that the Gromov product $(gx, g^{-1}x)_x$ does not grow too fast. To do so, we will use the following trick. For a proof, see [TT15].

Lemma 6.25 (Fellow traveling is contagious). *Suppose that X is a δ -hyperbolic space with basepoint x and suppose that $A \geq 0$. If $a, b, c, d \in X$ are points of X with $(a \cdot b)_x \geq A$, $(c \cdot d)_x \geq A$, and $(a \cdot c)_x \leq A - 3\delta$. Then $(b \cdot d)_x - 2\delta \leq (a \cdot c)_x \leq (b \cdot d)_x + 2\delta$.*

We apply this to prove the following:

Lemma 6.26. *Let $f : \mathbb{N} \rightarrow \mathbb{R}$ be any function such that $f(n) \rightarrow +\infty$ as $n \rightarrow +\infty$. Then*

$$\mathbb{P}\left((w_nx, w_n^{-1}x)_x \geq f(n)\right) \rightarrow 0$$

as $n \rightarrow \infty$.

Proof. Define

$$f_1(n) = \min \left\{ f(n) - 2\delta, \frac{n(L - \eta)}{2} - 3\delta \right\}$$

It is easy to see that $f_1(n) \rightarrow \infty$ as $n \rightarrow \infty$. By Lemma 6.25, if we know that:

- (1) $(w_mx, w_nx)_x \geq n(L - \eta)/2$,
- (2) $(u_m^{-1}x, w_n^{-1}x)_x \geq n(L - \eta)/2$, and

$$(3) (w_mx, u_m^{-1}x)_x \leq f_1(n) \leq n(L - \eta)/2 - 3\delta,$$

then

$$(w_nx, w_n^{-1}x)_x \leq (w_mx, u_m^{-1}x)_x + 2\delta \leq f_1(n) + 2\delta.$$

Using Lemmas 6.21, 6.22, and 6.23, the probability that conditions (1),(2), (3) hold tends to 1, hence we have

$$\mathbb{P}^n \left((w_nx, w_n^{-1}x)_x \leq f(n) \right) \rightarrow 1$$

as $n \rightarrow \infty$. □

We can now prove Theorem 1.6 from the introduction.

Theorem 6.27. *For every $\epsilon > 0$, one has*

$$\mathbb{P} \left(\tau_X(w_n) \geq n(L - \epsilon) \right) \rightarrow 1,$$

as $n \rightarrow \infty$.

Proof. If we set $f(n) = \eta n$ with $\eta > 0$, then by Lemma 6.26 and Corollary 6.16 the events $(w_nx, w_n^{-1}x)_x \leq \eta n$ and $d(x, w_nx) \geq n(L - \eta)$ occur with probability (\mathbb{P}^n) which tends to 1, hence by Lemma 6.24

$$\mathbb{P}^n \left(\tau(w_n) \geq n(L - 3\eta) \right) \geq \mathbb{P}^n \left(d(x, w_nx) - 2(w_nx, w_n^{-1}x)_x + O(\delta) \geq n(L - 3\eta) \right)$$

which approaches 1 as $n \rightarrow \infty$. This implies the statement if we choose $\epsilon > 3\eta$. □

7. GENERIC ELEMENTS AND THE ACTION $G \curvearrowright X$

Recall that for any $A \subseteq G$,

$$P^n(A) = \frac{\#(A \cap S_n)}{\#S_n},$$

where S_n is the n -sphere in G . Hence, P^n is uniform measure on the n -sphere of G .

We remind the reader that there is a constant $c > 1$ such that $c^{-1}\lambda^n \leq \#S_n \leq c\lambda^n$ (Theorem 5.4) where λ is the growth of the directed graph Γ . Also, since $\text{ev}: \tilde{\Gamma} \rightarrow G$ is bijective, we may unambiguously use graph terminology when referring to elements of G . For example, $[g] = v$ means that v is the vertex of Γ which is the terminal endpoint of the path starting at v_0 and spelling g . Further, for $g \in G$, \hat{g} denote the element of G along the unique combing geodesic from 1 to g which has distance $[\log |g|]$ from g , where $[x]$ denotes the integer part of x .

Lemma 7.1. *For any $R \geq 0$ and any $h \in G$, the group shadow $S_1(h, R) \subseteq G$ is contained in the union of finitely many cones:*

$$S_1(h, R) \subseteq \bigcup_{g \in B_{R+c}(h)} \text{cone}(g)$$

where c depends only on δ .

Proof. For each x, y, z in a metric space, by definition of shadow and Gromov product,

$$z \in S_x(y, R) \Leftrightarrow (y, z)_x \geq d(x, y) - R \Leftrightarrow (x, z)_y \leq R$$

Thus, if $g_1 \in S_1(h, R)$, then $(g_1, 1)_h \leq R$, which implies that the distance between any geodesic segment from 1 to g_1 and h is at most $R + c$, where c only depends on δ . In particular, there exists a group element h_1 on the combing geodesic from 1 to g_1 which lies at distance $\leq R + c$ from h . Thus, $g_1 \in \text{cone}(h_1)$ with $h_1 \in B_{R+c}(h)$. □

Lemma 7.2. *There is $K \geq 0$ such that for every $h \in G$ there is a $g \in G$ such that $v = [g]$ has large growth and $d_G(h, g) \leq K$.*

Proof. Recall that by Lemma 5.8 there exists an $R \geq 0$ such that for any h the group shadow $S_1(h, R)$ has maximal growth, i.e. $\#(S_1(h, R) \cap S_n) \asymp c_1 \lambda^n$ for some c_1 depending only on h . Moreover, by Lemma 7.1 the shadow $S_1(h, R)$ is contained in a finite union of cones:

$$S_1(h, R) \subseteq \bigcup_{g \in B_{R+c}(h)} \text{cone}(g)$$

where C depends only on δ . We conclude that one such cone $\text{cone}(g)$ also has large growth. This completes the proof. \square

7.1. Genericity of positive drift. The following result establishes Theorem 1.7 from the introduction using what we have shown about the Markov measures in Section 6.

Theorem 7.3. *For every $\epsilon > 0$ one has*

$$\frac{\#\{g \in S_n : d(gx, x) \geq (L - \epsilon)|g|\}}{\#S_n} \rightarrow 1 \quad \text{as } n \rightarrow \infty.$$

Proof. Let A_L denote the set of group elements

$$A_L := \{g \in G : d(gx, x) \leq L|g|\}.$$

We know by Corollary 6.16 that for any $L' < L$ one has

$$\mathbb{P}^n(A_{L'}) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Then we have

$$\mathbb{P}^n(A_{L-\epsilon}) \leq \frac{\#\{g \in S_n : \hat{g} \notin LG\}}{\#S_n} + \frac{\#\{g \in S_n \cap A_{L-\epsilon} : \hat{g} \in LG\}}{\#S_n}$$

and we know by Proposition 2.5 that the first term tends to 0. Now, by writing $g = \hat{g}h$ with $|h| = \lfloor \log |g| \rfloor$ we have that $d(gx, x) \leq (L - \epsilon)|g|$ implies

$$d(\hat{g}x, x) \leq d(gx, x) + d(\hat{g}x, gx) \leq (L - \epsilon)|g| + d(x, hx) \leq$$

hence, there exists C such that it is less than

$$\leq (L - \epsilon)|g| + C \lfloor \log |g| \rfloor \leq L'|\hat{g}|$$

for any $L - \epsilon < L' < L$ whenever $|g|$ is sufficiently large. This proves the inclusion

$$\{g \in S_n \cap A_{L-\epsilon} : \hat{g} \in LG\} \subseteq \{g \in S_n : \hat{g} \in A_{L'} \cap LG\}$$

and by Lemma 2.3 (1)

$$\#\{g \in S_n : \hat{g} \in A_{L'} \cap LG\} \leq c \lambda^{\lfloor \log n \rfloor} \#(S_{n-\lfloor \log n \rfloor} \cap A_{L'} \cap LG) \leq$$

hence by Lemma 3.4 (equation (4)) and considering the size of $S_{n-\lfloor \log n \rfloor}$

$$\leq c_1 \lambda^{\lfloor \log n \rfloor} \mathbb{P}^{n-\lfloor \log n \rfloor}(A_{L'}) \#S_{n-\lfloor \log n \rfloor} \leq c_2 \lambda^n \mathbb{P}^{n-\lfloor \log n \rfloor}(A_{L'}).$$

Finally, using that $\mathbb{P}^{n-\lfloor \log n \rfloor}(A_{L'}) \rightarrow 0$ we get

$$\limsup_{n \rightarrow \infty} \frac{\#\{g \in S_n \cap A_{L-\epsilon} : \hat{g} \in LG\}}{\#S_n} \leq \limsup_{n \rightarrow \infty} c_3 \mathbb{P}^{n-\lfloor \log n \rfloor}(A_{L'}) = 0$$

which proves the claim. \square

7.2. **Counting and decay of shadows.** For $g \in G$, we set

$$S_x^G(gx, R) = \{h \in G : hx \in S_x(gx, R)\},$$

where as usual, $S_x(gx, R)$ is the shadow in X around gx centered at the basepoint $x \in X$. We will need the following decay property for $S_x^G(gx, R) \subset G$.

Proposition 7.4. *There is a function $\rho : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ with $\rho(r) \rightarrow 0$ as $r \rightarrow \infty$ such that for every $n \geq 0$*

$$P^n(S_x^G(gx, R)) \leq \rho(d(x, gx) - R).$$

Proof. By Lemma 7.2, every $h \in S_x^G(gx, R)$ lies at distance $\leq D$ from a group element h_1 of large growth, and moreover by Lemma 5.6 such an element h_1 is contained in the shadow $S_x^G(gx, R + cD)$, where c is the Lipschitz constant of the orbit map $G \rightarrow X$. Thus,

$$S_x^G(gx, R) \subseteq N_D(S_x^G(gx, R + cD) \cap LG).$$

Now,

$$P^n(S_x^G(gx, R)) \leq P^n(N_D(S_x^G(gx, R + cD) \cap LG)) \leq$$

hence, if $\#B_D$ is the size of a ball of radius D in the Cayley graph,

$$\leq \#B_D \cdot P^n(S_x^G(gx, R + cD) \cap LG) \leq$$

and by Lemma 3.4

$$\leq c_1 \#B_D \cdot P^n(S_x^G(gx, R + cD)) \leq$$

hence by decay of shadows for \mathbb{P} (Proposition 6.19)

$$\leq c_1 \#B_D \cdot p(d(x, gx) - R - cD)$$

thus the claim is proven if we set $\rho(r) := c_1 \#B_D \cdot p(r - cD)$. \square

7.3. **Genericity of loxodromics.** For each n , let $n_1 = \lfloor \frac{n}{2} \rfloor$, $n_2 = n - n_1$. For each $g \in S_n$, let us pick its representative path from the initial vertex, and let a, b be the group elements associated respectively to the first n_1 edges, and the last n_2 edges of this path. Thus we can canonically write $g = ab$, with $a \in S_{n_1}$, $b \in S_{n_2}$, and b in the cone of a . We now show that a and b are almost independent:

Lemma 7.5. *There exists a constant $c > 0$ such that for any $n \geq 2$ the inequality*

$$P^n(a \in A, b \in B) \leq cP^{n_1}(A)P^{n_2}(B)$$

is satisfied for any subsets $A, B \subseteq G$.

Proof.

$$P^n(a \in A, b \in B) = \frac{\#\{(a, b) \in S_{n_1} \times S_{n_2} : b \in \text{cone}(a), a \in A, b \in B\}}{\#S_n} \leq$$

and forgetting the requirement that $b \in \text{cone}(a)$,

$$\leq \frac{\#\{(a, b) \in S_{n_1} \times S_{n_2} : a \in A, b \in B\}}{\#S_n} \leq \frac{P^{n_1}(A)P^{n_2}(B)\#S_{n_1}\#S_{n_2}}{\#S_n}.$$

Since $\#S_n = e^{\lambda n}$ by Theorem 5.4, we have $\#S_{n_1}\#S_{n_2} \leq c\#S_n$ for some $c \geq 0$ not depending on n . This completes the proof of the claim. \square

Lemma 7.6. For any $\epsilon > 0$,

$$P^n \left(d(x, ax) \leq \frac{n(L + \epsilon)}{2} \right) \rightarrow 1$$

and

$$P^n \left(d(x, bx) \leq \frac{n(L + \epsilon)}{2} \right) \rightarrow 1.$$

Proof. We prove the complementary statement that

$$P^n \left(d(x, ax) \geq \frac{n(L + \epsilon)}{2} \right) \rightarrow 0.$$

Indeed, by Lemma 7.5

$$P^n \left(d(x, ax) \geq \frac{n(L + \epsilon)}{2} \right) \leq cP^{n_1} \left(d(x, gx) \geq \frac{n(L + \epsilon)}{2} \right)$$

which tends to zero by Theorem 7.3, recalling that $n_1 \sim \frac{n}{2}$. The proof of the second statement is completely analogous. \square

Our goal is to prove that the translation length of a generic element in the n -sphere grows linearly in n . In order to apply Lemma 6.25, we need to check that the first half of g (which is a) and the first half of g^{-1} (which is b^{-1}) generically do not fellow travel:

Lemma 7.7. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be any function such that $f(n) \rightarrow +\infty$ as $n \rightarrow +\infty$. Then

$$P^n \left((ax, b^{-1}x)_x \geq f(n) \right) \rightarrow 0$$

as $n \rightarrow \infty$.

Proof. We compute

$$P^n \left((ax, b^{-1}x)_x \geq f(n) \right) = \frac{\#\{(g, h) \in S_{n_1} \times S_{n_2} : h \in \text{cone}(g), (gx, h^{-1}x)_x \geq f(n)\}}{\#S_n} \leq$$

and removing the requirement that $h \in \text{cone}(g)$ we have

$$\begin{aligned} &\leq \frac{\#\{(g, h) \in S_{n_1} \times S_{n_2} : (gx, h^{-1}x)_x \geq f(n)\}}{\#S_n} \leq \\ &\leq \frac{1}{\#S_n} \sum_{h \in S_{n_2}} \#\{g \in S_{n_1} : gx \in S_x(h^{-1}x, d(x, h^{-1}x) - f(n))\} \leq \end{aligned}$$

and from decay of shadows (Proposition 7.4) it follows that

$$\leq \frac{1}{\#S_n} \sum_{h \in S_{n_2}} \rho(f(n)) \#S_{n_1} \leq \frac{\#S_{n_1} \#S_{n_2} \rho(f(n))}{\#S_n} \leq c\rho(f(n)) \rightarrow 0.$$

\square

The following two lemmas are the counting analogues of Lemma 6.22 and Lemma 6.23.

Lemma 7.8. For each $\eta > 0$, the probability

$$P^n \left((b^{-1}x, g^{-1}x)_x \leq \frac{n(L - \eta)}{2} \right)$$

tends to 0 as $n \rightarrow \infty$.

Proof. By Lemma 6.20, and since the action is isometric

$$(b^{-1}x, g^{-1}x)_x \geq d(x, g^{-1}x) - d(b^{-1}x, g^{-1}x) = d(x, gx) - d(x, ax)$$

Now, for any $\epsilon > 0$, by genericity of positive drift (Theorem 7.3)

$$P^n(d(x, gx) \geq n(L - \epsilon)) \rightarrow 1$$

and by Lemma 7.6

$$P^n\left(d(x, ax) \leq \frac{n}{2}(L + \epsilon)\right) \rightarrow 1$$

so

$$P^n\left(\left((b^{-1}x, g^{-1}x)_x \geq n\left(\frac{L}{2} - \frac{3\epsilon}{2}\right)\right)\right) \rightarrow 1$$

which proves the claim setting $\eta = 3\epsilon/2$. □

Lemma 7.9. *For each $\eta > 0$, the probability*

$$P^n\left(\left((ax, gx)_x \leq \frac{n(L - \eta)}{2}\right)\right)$$

tends to 0 as $n \rightarrow \infty$.

Proof. By Lemma 6.20,

$$(ax, gx)_x \geq d(x, gx) - d(ax, gx) = d(x, gx) - d(x, bx)$$

hence, since for any $\epsilon > 0$ we have by genericity of positive drift (Theorem 7.3)

$$P^n(d(x, gx) \geq n(L - \epsilon)) \rightarrow 1$$

and by Lemma 7.6

$$P^n\left(d(x, bx) \leq \frac{n}{2}(L + \epsilon)\right) \rightarrow 1$$

we get

$$P^n\left(\left((ax, gx)_x \geq n\left(\frac{L}{2} - \frac{3\epsilon}{2}\right)\right)\right) \rightarrow 1$$

which proves the claim setting $\eta = 3\epsilon/2$. □

The following proposition establishes control of the Gromov products $(gx, g^{-1}x)_x$ with respect to our counting measures:

Proposition 7.10. *Let $f : \mathbb{N} \rightarrow \mathbb{R}$ be a function such that $f(n) \rightarrow +\infty$ as $n \rightarrow \infty$. Then*

$$P^n\left(\left((gx, g^{-1}x)_x \leq f(n)\right)\right) \rightarrow 1$$

as $n \rightarrow \infty$.

Proof. Define

$$f_1(n) = \min\left\{f(n) - 2\delta, \frac{n(L - \eta)}{2} - 3\delta\right\}$$

It is easy to see that $f_1(n) \rightarrow \infty$ as $n \rightarrow \infty$. By Lemma 6.25, if we know that:

- (1) $(ax, gx)_x \geq n(L - \eta)/2$,
- (2) $(b^{-1}x, g^{-1}x)_x \geq n(L - \eta)/2$, and
- (3) $(ax, b^{-1}x)_x \leq f_1(n) \leq n(L - \eta)/2 - 3\delta$,

then

$$(gx, g^{-1}x)_x \leq (ax, b^{-1}x)_x + 2\delta \leq f_1(n) + 2\delta.$$

Using Lemmas 7.7, 7.8, and 7.9, the probability that conditions (1),(2), (3) hold tends to 1, hence we have

$$P^n((gx, g^{-1}x)_x \leq f(n)) \rightarrow 1$$

as $n \rightarrow \infty$. □

Finally, we prove Theorem 1.8 from the introduction.

Theorem 7.11 (Linear growth of translation length). *For any $\epsilon > 0$ we have*

$$\frac{\#\{g \in S_n : \tau_X(g) \geq n(L - \epsilon)\}}{\#S_n} \rightarrow 1,$$

as $n \rightarrow \infty$.

Proof. If we set $f(n) = \eta n$ with $\eta > 0$, then by Proposition 7.10 and Theorem 7.3 the events $(gx, g^{-1}x)_x \leq \eta n$ and $d(x, gx) \geq n(L - \eta)$ occur with probability (P^n) which tends to 1, hence by Lemma 6.24

$$P^n\left(\tau(g) \geq n(L - 3\eta)\right) \geq P^n\left(d(x, gx) - 2(gx, g^{-1}x)_x + O(\delta) \geq n(L - 3\eta)\right)$$

which approaches 1 as $n \rightarrow \infty$. This implies the statement if we choose $\epsilon > 3\eta$. □

Since elements with positive translation length are loxodromic, we finally get

Corollary 7.12 (Genericity of loxodromics).

$$\frac{\#\{g \in S_n : g \text{ is } X - \text{loxodromic}\}}{\#S_n} \rightarrow 1,$$

as $n \rightarrow \infty$.

REFERENCES

- [Ago13] Ian Agol, *The virtual Haken conjecture*, with appendix by Agol, Daniel Groves, and Jason Manning. Doc. Math. **18** (2013), 1045–1087.
- [AO96] G. Arzhantseva and A. Olshanskii, *Generality of the class of groups in which subgroups with a lesser number of generators are free*, Mat. Zametki **59** (1996), 489–496.
- [Arz98] GN Arzhantseva, *Generic properties of finitely presented groups and Howson’s theorem*, Communications in Algebra **26** (1998), no. 11, 3783–3792.
- [BBF15] Mladen Bestvina, Ken Bromberg, and Koji Fujiwara, *Constructing group actions on quasi-trees and applications to mapping class groups*, Publications mathématiques de l’IHÉS **122** (2015), no. 1, 1–64.
- [BC12] Jason Behrstock and Ruth Charney, *Divergence and quasimorphisms of right-angled Artin groups*, Mathematische Annalen **352** (2012), no. 2, 339–356.
- [BF09] Mladen Bestvina and Koji Fujiwara, *A characterization of higher rank symmetric spaces via bounded cohomology*, Geom. Funct. Anal. **19** (2009), no. 1, 11–40.
- [BF14] Mladen Bestvina and Mark Feighn, *Hyperbolicity of the complex of free factors*, Adv. Math. **256** (2014), 104–155. MR 3177291
- [BH09] Martin R. Bridson and Aandré Haefliger, *Metric spaces of non-positive curvature*, vol. 319, Springer, 2009.
- [BHM11] Sébastien Blachère, Peter Haïssinsky, and Pierre Mathieu, *Harmonic measures versus quasiconformal measures for hyperbolic groups*, Ann. Sci. E.N.S. **44** (2011), no. 4, 683–721.
- [Bir69] Joan S Birman, *Mapping class groups and their relationship to braid groups*, Communications on Pure and Applied Mathematics **22** (1969), no. 2, 213–238.
- [BLM83] Joan S. Birman, Alex Lubotzky, and John McCarthy, *Abelian and solvable subgroups of the mapping class groups*, Duke Math. J. **50** (1983), no. 4, 1107–1120.
- [BMR03] Alexandre V Borovik, Alexei G Myasnikov, and Vladimir N Remeslennikov, *Multiplicative measures on free groups*, Internat. J. Algebra Comput. **13** (2003), no. 06, 705–731.

- [Cal13] Danny Calegari, *The ergodic theory of hyperbolic groups*, Geometry and topology down under, Contemp. Math **597** (2013), 15–52.
- [Can84] James W Cannon, *The combinatorial structure of cocompact discrete hyperbolic groups*, Geometriae Dedicata **16** (1984), no. 2, 123–148.
- [CdCM⁺15] Pierre-Emmanuel Caprace, Y de Cornulier, N Monod, R Tessera, et al., *Amenable hyperbolic groups*, Journal of the European Mathematical Society **17** (2015), 2903–2947.
- [CF10] Danny Calegari and Koji Fujiwara, *Combable functions, quasimorphisms, and the central limit theorem*, Ergodic Theory and Dynamical Systems **30** (2010), no. 05, 1343–1369.
- [Cha95] Christophe Champetier, *Propriétés statistiques des groupes de présentation finie*, Adv. Math. **116** (1995), no. 2, 197–262.
- [CLM12] Matt T. Clay, Christopher J. Leininger, and Johanna Mangahas, *The geometry of right-angled Artin subgroups of mapping class groups*, Groups Geom. Dyn. **6** (2012), no. 2, 249–278.
- [CM07a] Chris Connell and Roman Muchnik, *Harmonicity of Gibbs measures*, Duke Math. J. **137** (2007), no. 3, 461–509.
- [CM07b] ———, *Harmonicity of quasiconformal measures and Poisson boundaries of hyperbolic spaces*, Geom. Funct. Anal. **17** (2007), no. 3, 707–769.
- [CM15] Danny Calegari and Joseph Maher, *Statistics and compression of scl*, Ergodic Theory and Dynamical Systems **35** (2015), no. 01, 64–110.
- [Coo93] Michel Coornaert, *Mesures de Patterson-Sullivan sur le bord d’un espace hyperbolique au sens de Gromov*, Pacific Journal of Mathematics **159** (1993), no. 2, 241–270.
- [Dan05] Pallavi Dani, *Statistical properties of elements in infinite groups*, Ph.D. thesis, University of Chicago, 2005.
- [DSU14] Tushar Das, David Simmons, and Mariusz Urbański, *Geometry and dynamics in Gromov hyperbolic metric spaces: With an emphasis on non-proper settings*, arXiv preprint arXiv:1409.2155 (2014).
- [DT06] Nathan Dunfield and William P Thurston, *Finite covers of random 3-manifolds*, Invent. Math. **166** (2006), no. 3, 457–521.
- [DT16] Spencer Dowdall and Samuel J Taylor, *The co-surface graph and the geometry of hyperbolic free group extensions*, arXiv preprint arXiv:1601.00101 (2016).
- [Far06] Benson Farb, *Some problems on mapping class groups and moduli space*, Proceedings of symposia in pure mathematics, vol. 74, Providence, RI; American Mathematical Society; 1998, 2006, p. 11.
- [Fur71] Harry Furstenberg, *Random walks and discrete subgroups of Lie groups*, 1–63.
- [GdlH90] É. Ghys and P. de la Harpe (eds.), *Sur les groupes hyperboliques d’après Mikhael Gromov*, Progress in Mathematics, vol. 83, Birkhäuser Boston, Inc., Boston, MA, 1990, Papers from the Swiss Seminar on Hyperbolic Groups held in Bern, 1988. MR 1086648 (92f:53050)
- [GMM15] Sébastien Gouëzel, Frédéric Mathéus, and François Maucourant, *Entropy and drift in word hyperbolic groups*, arXiv preprint arXiv:1501.05082 (2015).
- [Gro87] Mikhael Gromov, *Hyperbolic groups*, Springer, 1987.
- [Gro93] M Gromov, *Asymptotic invariants of infinite groups*, Geometric group theory, vol. 2, vol. 182, Cambridge University Press, 1993.
- [Gro03] Mikhail Gromov, *Random walk in random groups*, Geometric and Functional Analysis **13** (2003), no. 1, 73–146.
- [GST17] Dominik Gruber, Alessandro Sisto, and Romain Tessera, *Gromov’s random monsters do not act non-elementarily on hyperbolic spaces*, arXiv preprint arXiv:1705.10258 (2017).
- [HM13] Michael Handel and Lee Mosher, *Subgroup decomposition in Out(F_n): Introduction and research announcement*, arXiv preprint arXiv:1302.2681 (2013).
- [HW08] Frédéric Haglund and Daniel T Wise, *Special cube complexes*, Geom. Funct. Anal. **17** (2008), no. 5, 1551–1620.
- [Iva92] Nikolai V. Ivanov, *Subgroups of Teichmüller modular groups*, Translations of Mathematical Monographs, vol. 115, American Mathematical Society, Providence, RI, 1992, Translated from the Russian by E. J. F. Primrose and revised by the author.
- [Kai94] Vadim A Kaimanovich, *The Poisson boundary of hyperbolic groups*, Comptes rendus de l’Académie des sciences. Série 1, Mathématique **318** (1994), no. 1, 59–64.
- [KB02] Ilya Kapovich and Nadia Benakli, *Boundaries of hyperbolic groups*, Combinatorial and geometric group theory (New York, 2000/Hoboken, NJ, 2001) **296** (2002), 39–93.
- [KK14] Sang-Hyun Kim and Thomas Koberda, *The geometry of the curve graph of a right-angled Artin group*, International Journal of Algebra and Computation **24** (2014), no. 02, 121–169.
- [KLP11] Vadim A Kaimanovich and Vincent Le Prince, *Matrix random products with singular harmonic measure*, Geom. Dedicata **150** (2011), no. 1, 257–279.

- [KMSS03] Ilya Kapovich, Alexei Myasnikov, Paul Schupp, and Vladimir Shpilrain, *Generic-case complexity, decision problems in group theory, and random walks*, J. Algebra **264** (2003), no. 2, 665–694.
- [KMT14] Thomas Koberda, Johanna Mangahas, and Samuel J. Taylor, *The geometry of purely loxodromic subgroups of right-angled Artin groups*, Trans. Amer. Math. Soc, to appear .arXiv preprint arXiv:1412.3663 (2014).
- [Kob10] Thomas Koberda, *Right-angled Artin groups and a generalized isomorphism problem for finitely generated subgroups of mapping class groups*, Geom. Funct. Anal. (2010), 1–50.
- [KR14] Ilya Kapovich and Kasra Rafi, *On the hyperbolicity of the free splitting and free factor complexes*, Groups Geom. Dyn. **8** (2014), no. 2, 391–414.
- [Kra81] Irwin Kra, *On the Nielsen-Thurston-Bers type of some self-maps of Riemann surfaces*, Acta mathematica **146** (1981), no. 1, 231–270.
- [KRSS07] Ilya Kapovich, Igor Rivin, Paul Schupp, and Vladimir Shpilrain, *Densities in free groups and \mathbb{Z}^n , visible points and test elements*, Math. Res. Lett **14** (2007), no. 2.
- [KS05] Ilya Kapovich and Paul Schupp, *Genericity, the Arzhantseva-Ol’shanskii method and the isomorphism problem for one-relator groups*, Math. Ann. **331** (2005), no. 1, 1–19.
- [Led95] Francois Ledrappier, *Applications of dynamics to compact manifolds of negative curvature*, Proceedings of the International Congress of Mathematicians (Zurich 1994), 1995, pp. 1195–1202.
- [Mah11] Joseph Maher, *Random walks on the mapping class group*, Duke Math. J. **156** (2011), no. 3, 429–468.
- [MM99] Howard A. Masur and Yair N. Minsky, *Geometry of the complex of curves. I. Hyperbolicity*, Invent. Math. **138** (1999), no. 1, 103–149.
- [MS14] P Mathieu and A Sisto, *Deviation inequalities for random walks*, arXiv preprint arXiv:1411.7865 (2014).
- [MT14] Joseph Maher and Giulio Tiozzo, *Random walks on weakly hyperbolic groups*, To appear in J. Reine Angew. Math. (2014).
- [Ol’92] A Yu Ol’shanskii, *Almost every group is hyperbolic*, Internat. J. Algebra Comput. **2** (1992), no. 01, 1–17.
- [PS98] Mark Pollicott and Richard Sharp, *Comparison theorems and orbit counting in hyperbolic geometry*, Trans. Amer. Math. Soc. **350** (1998), no. 2, 473–499.
- [Riv08] Igor Rivin, *Walks on groups, counting reducible matrices, polynomials, and surface and free group automorphisms*, Duke Math. J. **142** (2008), no. 2, 353–379.
- [Ser89] Herman Servatius, *Automorphisms of graph groups*, J. Algebra **126** (1989), no. 1, 34–60.
- [Ser03] Jean-Pierre Serre, *Trees*, Springer Monographs in Mathematics. Springer-Verlag, Berlin (2003).
- [Sis11] Alessandro Sisto, *Contracting elements and random walks*, J. Reine Angew. Math. (2011).
- [Sun] Matt Sunderland, *Linear progress with exponential decay in weakly hyperbolic groups*, arXiv:1710.05107 [math.GT].
- [Tay15] Samuel J. Taylor, *Right-angled Artin groups and $Out(F_n)$ I: quasi-isometric embeddings*, Groups Geom. Dyn. **9** (2015), no. 1, 275–316.
- [TT15] Samuel J. Taylor and Giulio Tiozzo, *Random extensions of free groups and surface groups are hyperbolic*, Int. Math. Res. Not. (2015).
- [Uya15] Caglar Uyanik, *Generalized North–South dynamics on the space of geodesic currents*, Geom. Dedicata **177** (2015), no. 1, 129–148.
- [Ver00] Anatoly Vershik, *Dynamic theory of growth in groups: Entropy, boundaries, examples*, Russ. Math. Surv. **55** (2000), no. 4, 667–733.
- [Wie14] Bert Wiest, *On the genericity of loxodromic actions*, arXiv preprint arXiv:1406.7041 (2014).
- [Wis12] Daniel T Wise, *From riches to RAAGs: 3-manifolds, right-angled Artin groups, and cubical geometry*, vol. 117, American Mathematical Soc., 2012.

DEPARTMENT OF MATHEMATICS, YALE UNIVERSITY, 10 HILLHOUSE AVE, NEW HAVEN, CT 06520, U.S.A,
E-mail address: ilya.gekhtman@yale.edu

DEPARTMENT OF MATHEMATICS, TEMPLE UNIVERSITY, 1805 NORTH BROAD STREET, PHILADELPHIA, PA 19122, U.S.A,
E-mail address: samuel.taylor@temple.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TORONTO, 40 ST GEORGE ST, TORONTO, ON M5S 2E4, CANADA,
E-mail address: tiozzo@math.utoronto.ca