ACYLINDRICALLY HYPERBOLIC GROUPS AND COUNTING PROBLEMS

INHYEOK CHOI

Abstract. We show that Morse elements are generic in acylindrically hyperbolic groups. As an application, we observe that fully irreducible outer automorphisms are generic in the outer automorphism group of a finite-rank free group.

Keywords. Acylindrical action, Morse, weak proper discontinuity, counting problem, genericity
MSC classes: 20F67, 30F60, 57K20, 57M60, 60G50

1. Introduction

In non-positively curved manifolds and groups, certain geodesics or group elements exhibit hyperbolicity. A quasi-geodesic γ is said to be *Morse* if every quasi-geodesic of uniform quality connecting points on γ lies in a common neighborhood of γ. A group element g is called a *Morse element* if its orbit $\{g^i\}_{i\in\mathbb{Z}}$ is an unbounded Morse quasi-geodesic in the group.

In globally hyperbolic spaces such as CAT(-1) spaces and Gromov hyperbolic spaces, every geodesic is Morse (of uniform quality). This corresponds to the fact that every infinite-order element in a word hyperbolic group is loxodromic and is Morse. Furthermore, "most" elements in a word hyperbolic group are Morse. To formulate this, given a group G and its generating set S, let $B_S(n)$ be the collection of group elements whose S-word length is at most n. We can ask if the proportion of Morse elements in $B_S(n)$ tends to 1 as n tends to infinity. This is indeed the case when G is an infinite word hyperbolic group [Dan], [GTT18], [Yan20].

Morse elements are found in many other groups with flat parts. One classic example is the mapping class group $\mathrm{Mod}(\Sigma)$ of a finite-type hyperbolic surface Σ, whose Morse elements are precisely pseudo-Anosov mapping classes. In [Cho24a], the author proved that the asymptotic density of pseudo-Anosovs in the mapping class group is 1.

In this paper we focus on *acylindrically hyperbolic groups*, which are vast generalizations of hyperbolic groups that include $\mathrm{Mod}(\Sigma)$, $\mathrm{Out}(F_N)$, rank-1 CAT(0) groups and many 3-manifold groups, as well as many Coxeter groups , one-relator groups and free-by-cyclic groups ([BF02], [BF09], [CF10], [BF10], [MO15], [Sis18]). Despite its generality, acylindrically hyperbolic groups have shown to share many dynamical properties with word hyperbolic groups. We refer the readers to the first few pioneering papers in this direction: [BF02], [Ham08], [DGO17], [Sis18].

We add one more property to this list: genericity of Morse elements. Our main theorem is:

Theorem A. *Let G be an acylindrically hyperbolic group. Then for every finite generating set S of G, we have*
$$\lim_{n\to+\infty} \frac{\#\{g \in B_S(n) : g \text{ is Morse}\}}{\#B_S(n)} = 1.$$

This generalizes W. Yang's result on groups with strongly contracting element [Yan20]. This can be also compared with A. Sisto's theorem that simple random walks on acylindrically hyperbolic groups favor Morse elements ([Sis18, Theorem 1.6]). In fact, non-elementary random walks on any Gromov hyperbolic space favor loxodromics [CM15], [MT18], but one cannot hope such a result for counting problems (see the following subsections).

Date: May 4, 2025.

In view of the equivalent definitions of acylindrically hyperbolic groups in [Osi16] (especially in relation to [BF02]), Theorem A is a restatement of the following more explicit theorem.

Theorem 1.1. *Let G be a group generated by a finite set $S \subseteq G$. Suppose that G acts on a Gromov hyperbolic space X and that there exists $g \in G$ that is a loxodromic isometry of X with the WPD property (cf. Definition 2.5). Then for any $M > 0$, we have*

$$\lim_{n \to +\infty} \frac{\#\big\{g \in B_S(n) : g \text{ is a WPD loxodromic element with } \tau_X(g) > M\big\}}{\#B_S(n)} = 1.$$

Indeed, if $g \in G$ serves as a WPD loxodromic on a Gromov hyperbolic space, then g is a Morse element in G ([Sis16, Theorem 1], [Osi16, Theorem 1.4]).

We note a theorem by B. Wiest [Wie17] that was applied to the mapping class group by M. Cumplido and B. Wiest [CW18]: for any finitely generated group G having a non-elementary action on a Gromov hyperbolic space, the density of loxodromics is bounded away from 0. Hence, the main point of Theorem A and 1.1 is that the density has limit 1. Such a claim does not hold for general non-elementary actions.

Two important examples of acylindrically hyperbolic groups beyond hierarchically hyperbolic groups (HHGs) are $\mathrm{Out}(F_N)$ and $\mathrm{Aut}(F_N)$, the outer automorphism group and the automorphism group of the free group of rank $N \geq 3$. Theorem A tells us that most elements are Morse in large word metric balls in these groups.

We can say more by focusing on a specific $\mathrm{Out}(F_N)$-action, namely, the one on the free factor complex $\mathcal{FF}_N$ studied by M. Bestvina and M. Feighn [BF14]. Bestvina and Feighn proved that:

(1) $\mathcal{FF}_N$ is Gromov hyperbolic,
(2) the elements of $\mathrm{Out}(F_N)$ that are loxodromic isometries of $\mathcal{FF}_N$ are precisely the fully irreducible outer automorphisms, and
(3) every fully irreducible outer automorphism has the WPD property.

Using this action and building upon the stability result in [KMPT22], we prove:

Theorem 1.2. *Let $G = \mathrm{Out}(F_N)$ be the outer automorphism group of the free group of rank N for some $N \geq 2$. Then for any finite generating set S of G, we have*

$$\lim_{n \to +\infty} \frac{\#\{g \in B_S(n) : g \text{ is an ageometric triangular fully irreducible element}\}}{\#B_S(n)} = 1.$$

This is a counting version of Kapovich–Maher–Pfaff–Taylor's result that random walks on $\mathrm{Out}(F_N)$ favor ageometric triangular fully irreducibles [KMPT22, Theorem A]. There are also versions of random walk theory on $\mathrm{Aut}(F_N)$ and $\mathrm{Out}(F_N)$ using "non-backtracking" paths ([KKS07], [KP15]). In particular, I. Kapovich and C. Pfaff proved that non-backtracking random walks favor geometric triangular fully irreducibles as well.

We record a cute application to the mapping class group $\mathrm{Mod}(\Sigma)$. It seems hard to apply the method of [Cho24a] to general non-elementary subgroups of $\mathrm{Mod}(\Sigma)$, as there is no distance formula for them. However, since they still act on the curve complex $\mathcal{C}(\Sigma)$ with a WPD loxodromic element, we observe that:

Corollary 1.3. *Let $G \leq \mathrm{Mod}(\Sigma)$ be a non-elementary subgroup of the mapping class group and let S be a finite generating set of G. Then for any $M > 0$, we have*

$$\lim_{n \to +\infty} \frac{\#\big\{g \in B_S(n) : g \text{ is pseudo-Anosov with stretch factor} \geq M\big\}}{\#B_S(n)} = 1.$$

In particular, pseudo-Anosovs are generic in the Torelli group. This generalizes the result of I. Gekhtman, S. Taylor, and G. Tiozzo regarding word hyperbolic groups acting on a Gromov hyperbolic space [GTT18, Theorem 1.12].

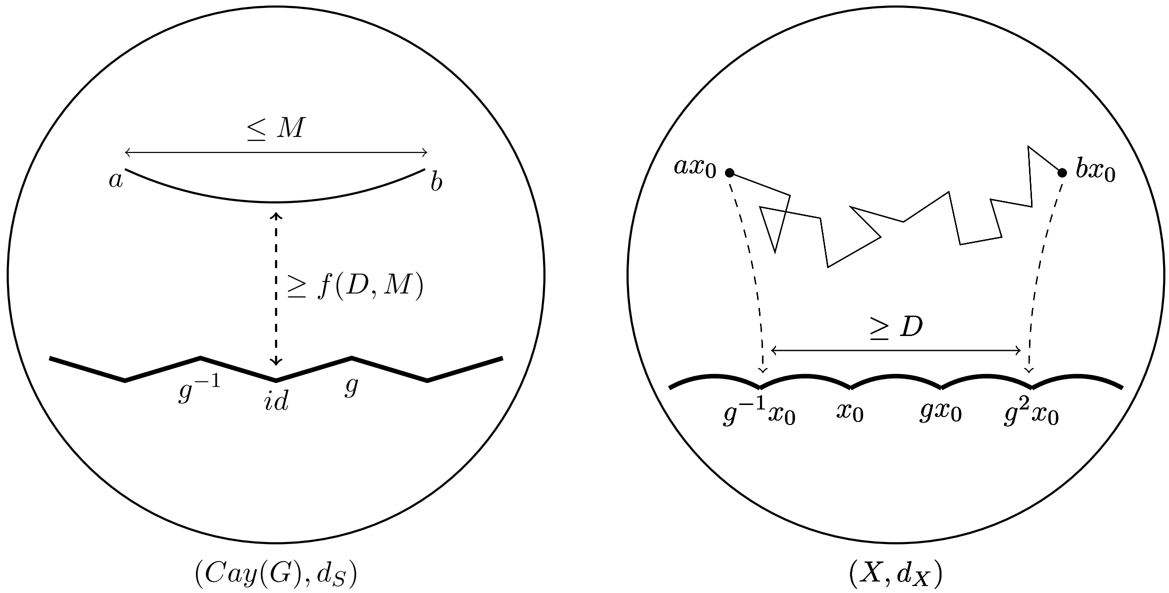

FIGURE 1. Schematics for $f(D, M)$ in Subsection 1.1

1.1. **Comparison with other groups.** To better illustrate Theorem 1.1, let us compare four groups that act on a Gromov hyperbolic space: the free group F_2 of rank 2, the mapping class group $\mathrm{Mod}(\Sigma)$, the outer automorphism group $\mathrm{Out}(F_N)$ of the free group of rank $n \geq 2$, and the direct product $F_2 \times F_3$ of two free groups. All of these act on some Gromov hyperbolic space.

Let G be a group acting on a hyperbolic space X and let S be a finite generating set of G. Given a group element $g \in G$ and a basepoint $x_0 \in X$, let us define a function $f : \mathbb{R}^2 \to \mathbb{R} \cup \{+\infty\}$ as follows. For any M-short word metric geodesic $[a, b] \subseteq G$, if a and b are D-apart along $\{g^i x_0\}_{i\in\mathbb{Z}}$, then $[a, b]$ must pass through an $f(D, M)$-neighborhood of $\{g^i\}_{i\in\mathbb{Z}}$ *in* G.

First, F_2 has proper action on its own Cayley graph $Cay(F_2)$. This implies that any coarse stabilizer of $v \in F_2$ is finite. Furthermore, each $g \in F_2 \setminus \{id\}$ has the so-called *strong contracting property*: if a geodesic $[a, b] \subseteq F_2$ makes nontrivial progress along $\{g^i\}_{i\in\mathbb{Z}}$, then $[a, b]$ passes through a bounded neighborhood of $\{g^i\}_{i\in\mathbb{Z}}$. In other words, $f(D, M)$ is constant in M for large enough D.

Second, $\mathrm{Mod}(\Sigma)$ acts on the ambient curve complex $\mathcal{C}(\Sigma)$ and tuples of subsurface curve complexes $\mathcal{C}(U), U \subsetneq \Sigma$. Fixing a simple closed curve $x_0 \in \mathcal{C}(\Sigma)$, each $g \in \mathrm{Mod}(\Sigma)$ gives rise to shadows $d_U(x_0, gx_0)$ on various $\{\mathcal{C}(U) : U \subseteq \Sigma\}$, using which the word metric on $\mathrm{Mod}(\Sigma)$ can be (coarsely) estimated via distance formula [MM00]. One consequence of the distance formula and is the *weakly contracting property* of pseudo-Anosov orbits [Beh06], [DR09]. Explicitly, for each pseudo-Anosov mapping class g, there exists $\epsilon > 0$ such that if an M-short geodesic $[a, b] \subseteq \mathrm{Mod}(\Sigma)$ makes progress D along $\{g^i\}_{i\in\mathbb{Z}}$, then $[a, b]$ passes through a $f(D, M) := (M \cdot e^{-\epsilon D})$-neighborhood of $\{g^i\}_{i\in\mathbb{Z}}$.

There is no direct analogue of the distance formula for $\mathrm{Out}(F_N)$. As a result, we do not know whether fully irreducible outer automorphisms (which are analogues of pseudo-Anosovs) have the weakly contracting property on the Cayley graph of $\mathrm{Out}(F_N)$. However, every fully irreducible outer automorphism g has the WPD property (for various hyperbolic actions, cf. [BF10], [Man14], [BF14]), i.e., the joint coarse stabilizer of g^i and g^j is finite when $|i - j|$ is large. This leads to an implicit contracting property, i.e., for every D and M the value of $f(D, M)$ is finite.

Finally, consider a trivial projection of $F_2 \times F_3$ onto the first factor F_2. This gives rise to a natural action of $F_2 \times F_3$ on $Cay(F_2)$. This action has not only a large point stabilizer, but also a large global stabilizer. Namely, $\{id\} \times F_3$ acts trivially on $Cay(F_2)$. In addition, there is no contraction along loxodromics on F_2, i.e., $f(D, M) = +\infty$. In general, if $X \times Y$ is a product space,

	δ-hyperbolic space	$f(D,M)$ for a fixed D	Density of non-loxodromics
F_2	$Cay(F_2)$	constant in M	$\lesssim \lambda^{-n}$ for some $\lambda > 1$
$\mathrm{Mod}(\Sigma)$	$\mathcal{C}(\Sigma)$	linear in M	$\lesssim n^{-k}$ $(\forall k)$
$\mathrm{Out}(F_N)$	$\mathcal{FF}_N$	finite	tends to 0
$F_2 \times F_3$	$Cay(F_2)$	$+\infty$	can be bounded away from 0

FIGURE 2. Properties of the four actions and the density estimates

a D-long geodesic γ can have D-large projection onto $X \times \{id\}$, regardless of the distance of γ from $X \times \{id\}$. Figure 2 summarizes the discussion so far.

The more information we have about the growth of $f(D,M)$, the better asymptotics of the density of non-loxodromics we can prove. In F_2, the proportion of non-loxodromic elements in $B_S(n)$ decays exponentially fast in n. This is proved by W. Yang [Yan20] in groups with strongly contracting elements, including relatively hyperbolic groups and small cancellation groups.

For the mapping class group, the function $f(D,M)$ grows at most linearly in M. Using this property, it is shown in [Cho24a] that the density of non-pseudo-Anosovs in $B_S(n)$ decays faster than n^{-k} for any $k > 0$. Similar growth behaviour of $f(D,M)$ is observed in HHGs with Morse elements, because loxodromics on the top curve space have the weakly contracting property. Rank-1 CAT(0) groups also fall into this category, as the strongly contracting property of a rank-1 element on the CAT(0) space implies its weak contracting property in the group.

Without control of $f(D,M)$, loxodromics can either be generic or non-generic depending on the generating set S. Indeed, there exist two finite generating sets S and S' of $F_2 \times F_3$, such that loxodromics (for the action on $Cay(F_2)$) are generic in S but not in S'. We refer readers to [GTT18, Example 1]. This simple example also tells us that genericity of Morse elements of a group may not be preserved through a quasi-isometry.

This paper deals with $\mathrm{Out}(F_N)$ and others of its ilk. There is no *a priori* control on the growth of $f(D,M)$ for acylindrically hyperbolic groups. Our main point is that, nonetheless, the finiteness of $f(D,M)$ is sufficient to conclude the genericity of loxodromics.

1.2. **Another side of the story: random walks.** There are two popular models to sample a random element in a group G. One is the counting method as in Theorem A. Namely, we consider a large word metric ball and choose an element with respect to the uniform measure. The other one is the random walk model: we put a probability measure μ on a generating set S of G (e.g., the uniform measure when S is finite) and investigate its n-fold convolution μ^{*n}.

For example, given a G-action on a Gromov hyperbolic space X, one can ask if $\mathbb{P}_{\mu^{*n}}(g \text{ is loxodromic})$ converges to 1 as n tends to infinity. This is closely related to a description of a typical sample path drawn on X, called *ray approximation* or *geodesic tracking*, that was pursued for word hyperbolic groups by V. Kaimanovich [Kai94]; see [Kai00] also. It was J. Maher's observation that neither the properness of X nor the properness of the action is necessary. As a result, Maher proved in [Mah11] that $\mathbb{P}_{\mu^{*n}}(g \text{ is pseudo-Anosov})$ converges to 1 in the mapping class group (I. Rivin independently proved this result using different method in [Riv08]).

Maher's observation was later generalized by D. Calegari and J. Maher [CM15], and once again by J. Maher and G. Tiozzo in [MT18]: they proved that $\mathbb{P}_{\mu^{*n}}(g \text{ is loxodromic})$ converges to 1 as long as the G-action on X is non-elementary (i.e., S generates two independent loxodromics). In particular, random walks do not care if the group has a large subgroup with trivial action, given that they hit non-elementary loxodromic elements for a positive probability. Maher-Tiozzo's result indeed applies to all 4 group actions in Subsection 1.1.

Consequently, for the uniform measure μ_S on a finite generating set S of G, the genericity of loxodromics with respect to μ_S^{*n} does not imply the genericity with respect to (uniform measure on $B_S(n)$). This is anticipated by the fact that the two measures differ by an exponential factor in n.

If one is allowed to pick their favorite generating set S for G, then one can bring the estimates from random walks to the counting problem. This was indeed the strategy of [Cho24b], where the author proved that every finitely generated weakly hyperbolic group has a finite generating set S for which loxodromics are generic. Since the asymptotic density may depend on the choice of S (as shown in [GTT18, Example 1]), this strategy does not establish Theorem A.

1.3. **Beyond hyperbolic spaces.** The method for Theorem 1.1 does not require global hyperbolicity of the space X. It only uses the *strongly contracting property* and the WPD property of $g \in G$ in X. For simplicity, however, we will not pursue this generality. It should be noted that the previous assumption does not imply that g is strongly contracting *in* G, i.e., with respect to the word metric. For example, the author does not know whether fully irreducibles are weakly contracting with respect to the word metrics (cf. [BD14, Question 6.8]).

For example, the method for Theorem 1.1 applies to finitely generated groups acting on a CAT(0) space (not necessarily cocompactly) that involves a rank-1 isometry with the WPD property. The study of strongly contracting isometries and their dynamics is growing rapidly. We refer the readers to the references in [ACT15], [Yan19], [Yan20], [Cou22], [SZ23], [DMGZ25].

In fact, the very notion of acylindrically hyperbolic group was already formulated in terms of contracting elements by A. Sisto [Sis18], who generalized Maher-Tiozzo's random walk theory in [MT18] to non-hyperbolic spaces. We also note a recent construction by H. Petyt and A. Zalloum [PZS24, Theorem B] that justifies why it suffices to consider WPD action on hyperbolic spaces.

1.4. **Open questions.** The methods in [Cho24a] and this paper still do not answer:

Question 1.4. *Are pseudo-Anosovs exponentially generic in every word metric on* $\mathrm{Mod}(\Sigma)$*?*

There are two types of word metrics for which exponential genericity of pseudo-Anosovs is known. One comes from generating sets mostly consisting of independent pseudo-Anosovs [Cho24b]. The other recent one is due to L. Ding, D. Martínez-Granado and A. Zalloum [DMGZ25], where the authors consider the $\mathrm{Mod}(\Sigma)$-action on an injective metric space (Y, d_Y) and collect orbit points in a large d_Y-ball. It seems hard to push either method to handle arbitrary word metric.

For non-HHGs, we can ask:

Question 1.5. *Are fully irreducibles exponentially generic in* $\mathrm{Out}(F_N)$ *with respect to every word metric? Or, is it at least true for some* $\alpha > 0$ *that*

$$\frac{\#B_S(n) \cap \{\textit{fully irreducibles}\}}{\#B_S(n)} \lesssim n^{-\alpha}?$$

This question might be answered for a given group G whenever we know the growth of the function $f(D, M)$ in Subsection 1.1.

One can ask more details about generic elements. In the random walk side we have strong law of large numbers (SLLN): for any non-elementary random walk $(Z_n)_{n>0}$ on a Gromov hyperbolic space, there exists $\lambda \in (0, +\infty]$ such that $\lim_n \frac{\|Z_n\|_X}{n} = \lim_n \frac{\tau_X(Z_n)}{n} = \lambda$ almost surely ([CM15], [MT18], [BCK21]). Here the key point is the linear growth of displacement and translation length. We pose:

Question 1.6. *Do generic fully irreducibles have linearly growing translation length? Namely, given a finite generating set* S *of* $\mathrm{Out}(F_N)$*, does there exist a linear function* $f : \mathbb{R} \to \mathbb{R}$ *such that*

$$\lim_{n \to +\infty} \frac{\#B_S(n) \cap \{g \in \mathrm{Out}(F_N) : \tau_{\mathcal{FF}}(g) \geq f(n)\}}{\#B_S(n)} = 1? \tag{1.1}$$

Our method does provide a diverging function f for which Equation 1.1 holds, but we have no control on the growth of f. For the mapping class group, the author anticipates that the method in [Cho24a] guarantees $f(n) \gtrsim \sqrt{n}$. The results of [Cho24b] and [DMGZ25] imply that $f(n) \gtrsim n$ works for *certain* finite generating set S.

Finally, we state a question related to Question 1.4.

Question 1.7. *Does $G = \mathrm{Mod}(\Sigma)$ or $G = \mathrm{Out}(F_N)$ have purely exponential growth? That means, for (some or every) finite generating set S of G, does there exist $K, \lambda > 1$ such that*

$$\frac{1}{K}\lambda^n \le \#B_S(n) \le K\lambda^n? \quad (\forall n > 0)$$

This question is answered by W. Yang for groups with strongly contracting elements [Yan19, Theorem B]. Meanwhile, we do not know the answer for $\mathrm{Mod}(\Sigma)$ for any finite generating set.

1.5. **Plans.** After reviewing preliminaries in Section 2, we observe a variant of A. Sisto's geometric separation lemma [Sis16, Lemma 3.3] in Section 3. We then prove the main theorem in Section 4.

Acknowledgments. The author thanks Jason Behrstock for discussion about weakly contracting property in the mapping class group, Jason Manning for discussion about acylindrical actions, and Alessandro Sisto for discussion about Lemma 3.2.

This work was first conceived during the program "Randomness and Geometry" at the Fields Institute. The author thanks the Fields Institute and the organizers of the program for their hospitality. Some progress was made while the author was visiting June E Huh Center for Mathematical Challenges at KIAS in January 2025.

2. Preliminaries

In this section, we collect some notions and facts about acylindrically hyperbolic groups. We refer to Gromov's seminal paper [Gro87] and standard textbooks [CDP90], [GdlH90].

A metric space is said to be *geodesic* if every pair of points can be connected by a geodesic. For two points x and y in this space, we denote by $[x, y]$ an arbitrary geodesic connecting x to y. Given $\delta > 0$, we say that a geodesic metric space is δ-hyperbolic if every geodesic is δ-slim.

Given a geodesic $\gamma : I \to X$, we will sometimes denote the image $Im(\gamma) \subseteq X$ by γ. Based on this convention, we define the *closest point projection* $\pi_\gamma : X \to 2^\gamma$ by

$$y \in \pi_\gamma(x) \;\Leftrightarrow\; d_X(x, y) = \inf\big\{d_X(x, p) : p \in \gamma\big\}.$$

We say that two geodesics $\gamma : [0, L] \to X$ and $\eta : [0, L'] \to X$ are *ϵ-fellow traveling* if

$$d_X(\gamma(0), \eta(0)) < \epsilon,\; d_X(\gamma(L), \eta(L')) < \epsilon \text{ and } d_{Haus}(\gamma, \eta) < \epsilon.$$

The fellow traveling property is transitive: if γ_1 and γ_2 are ϵ-fellow traveling; γ_2 and γ_3 are ϵ'-fellow traveling, then γ_1 and γ_3 are $(\epsilon + \epsilon')$-fellow traveling. Furthermore, we have:

Fact 2.1. *Let X be a δ-hyperbolic space and let $x, y, z, w \in X$ be such that $d_X(x, y) < \epsilon$ and $d_X(z, w) < \epsilon'$. Then $[x, z]$ and $[y, z]$ are $(\epsilon + \delta)$-fellow traveling. Moreover, $[x, z]$ and $[y, w]$ are $(\epsilon + \epsilon' + 2\delta)$-fellow traveling.*

For each $x \in X$, $\pi_\gamma(x)$ may not be a singleton. Nevertheless, its diameter is bounded and $\pi_\gamma(\cdot)$ is coarsely Lipschitz. The following is a consequence of [CDP90, Proposition 10.2.1], which follows from the tree approximation lemma [CDP90, Théorème 8.1], [GdlH90, Théoréme 2.12].

Fact 2.2. *Let X be a δ-hyperbolic space.*

(1) Let $x, y \in X$ and let γ be a geodesic in X. Then $\pi_\gamma(x) \cup \pi_\gamma(y)$ has diameter at most $d_X(x, y) + 12\delta$.

(2) *Let $x, y \in X$, let γ be a geodesic in X and let $p \in \pi_\gamma(x)$ and $q \in \pi_\gamma(y)$. Suppose that p appears earlier than q on γ and that $d_X(p, q) > 20\delta$. Then any geodesic $[x, y]$ between x and y contains a subsegment that is 20δ-fellow traveling with $[p, q]$.*

Corollary 2.3 ([Sis18, Lemma 4.1])**.** *Let X be a δ-hyperbolic space, let γ be a geodesic in X, let $x, y \in X$ and let η be a subsegment of γ that contains $\pi_\gamma(x) \cup \pi_\gamma(y)$. Then $\pi_\gamma([x, y])$ is contained in the 60δ-neighborhood of η.*

Proof. Suppose to the contrary that there exist $z \in [x, y]$, $p \in \pi_\gamma(x)$, $q \in \pi_\gamma(y)$, $r \in \pi_\gamma(z)$ such that $d_X(p, r), d_X(q, r) \geq 60\delta$ and such that p, q are to the right of r. Let p_0 be the point on γ to the right of r such that $d_X(r, p_0) = 60\delta$. Then Fact 2.2(2) implies that there exist a subsegment $[r', p']$ of $[z, x]$ and a subsegment $[r'', p'']$ of $[z, y]$ such that $d_X(r', r), d_X(r'', r) < 20\delta$ and $d_X(p', p_0), d_X(p'', p_0) < 20\delta$. We then observe that

$$\begin{aligned} 40\delta &> d_X(p', p_0) + d_X(p_0, p'') \geq d_X(p', p'') \geq d_X(p', r') + d_X(r'', p'') \\ &\geq [d_X(p_0, r) - d_X(p_0, p') - d_X(r, r')] + [d_X(p_0, r) - d_X(p_0, p'') - d_X(r, r'')] > 20\delta + 20\delta, \end{aligned}$$

a contradiction. Similar contradiction happens when p, q are both to the left of r. ☐

For $x, y, z \in X$, we define the *Gromov product* of y and z based at x by

$$(y, z)_x := \frac{1}{2}\big[d_X(y, x) + d_X(x, z) - d_X(y, z)\big].$$

Gromov hyperbolicity has the following consequence.

Fact 2.4 ([Cho24a, Lemma A.3])**.** *Let X be a δ-hyperbolic space. Let $x, y, z \in X$ and let $p \in [y, z]$ be such that $d_X(p, y) = (x, z)_y$. Then $\pi_{[y,z]}(x)$ is contained in the 8δ-neighborhood of p.*

Definition 2.5. *Let G be a finitely generated group acting on a δ-hyperbolic space (X, d_X) with a basepoint $x_0 \in X$. We say that a loxodromic element $\varphi \in G$ has the* WPD (weak proper discontinuity) property *if for each K there exists N, M such that*

$$\#\Big(Stab_K(x_0, \varphi^N x_0) := \big\{g \in G : d_X(x_0, g x_0) < K \text{ and } d_X(\varphi^N x_0, g\varphi^N x_0) < K\big\}\Big) < M.$$

We say that a finitely generated group G is *acylindrically hyperbolic* if it admits an isometric action on a δ-hyperbolic space with a WPD loxodromic element $\varphi \in G$. An acylindrically hyperbolic group G is said to be *non-elementary* if it is not virtually cyclic. The following fact is a consequence of [BF02, Proposition 6(1), (2)]. The proof is sketched in [Cho24a, Fact 2.2].

Fact 2.6. *Let G be a non-virtually cyclic group with a generating set S. Suppose that G acts on a δ-hyperbolic space $X \ni x_0$ with a WPD loxodromic element $\varphi \in G$. Then there exists $E_0 > 0$ such that the following hold.*

(1) *For each $g \in G$, there exist $s, t \in S \cup \{id\}$ such that $(\varphi^i x_0, sgx_0)_{x_0} \leq E_0$ for all $i > 0$ and $(\varphi^j x_0, tgx_0)_{x_0} \leq E_0$ for all $j < 0$.*
(2) *Let $n > 0$ and $g \in G$. Let $\gamma := [x_0, \varphi^n x_0]$, let $p \in \pi_\gamma(gx_0)$ and let $q \in \pi_\gamma(g\varphi^n x_0)$. Suppose that p appears earlier than q along γ and suppose that $d_X(p, q) > E_0$. Then $d_S(\varphi^i, g\varphi^j) < E_0$ for some $i, j \in \{0, 1, \ldots, n\}$. (cf. Figure 3)*

We now recall the notion of alignment.

Definition 2.7. *Let $K > 0$ and let $\gamma_1, \gamma_2, \ldots, \gamma_n$ be geodesics (which can be degenerate, i.e., points) in a metric space X. We say that $(\gamma_1, \ldots, \gamma_n)$ is K-aligned if for each $i = 1, \ldots, n-1$ we have*

$$\operatorname{diam}\big(\pi_{\gamma_i}(\gamma_{i+1}) \cup (\textit{ending point of } \gamma_i)\big) < K \textit{ and}$$
$$\operatorname{diam}\big(\pi_{\gamma_{i+1}}(\gamma_i) \cup (\textit{beginning point of } \gamma_{i+1})\big) < K.$$

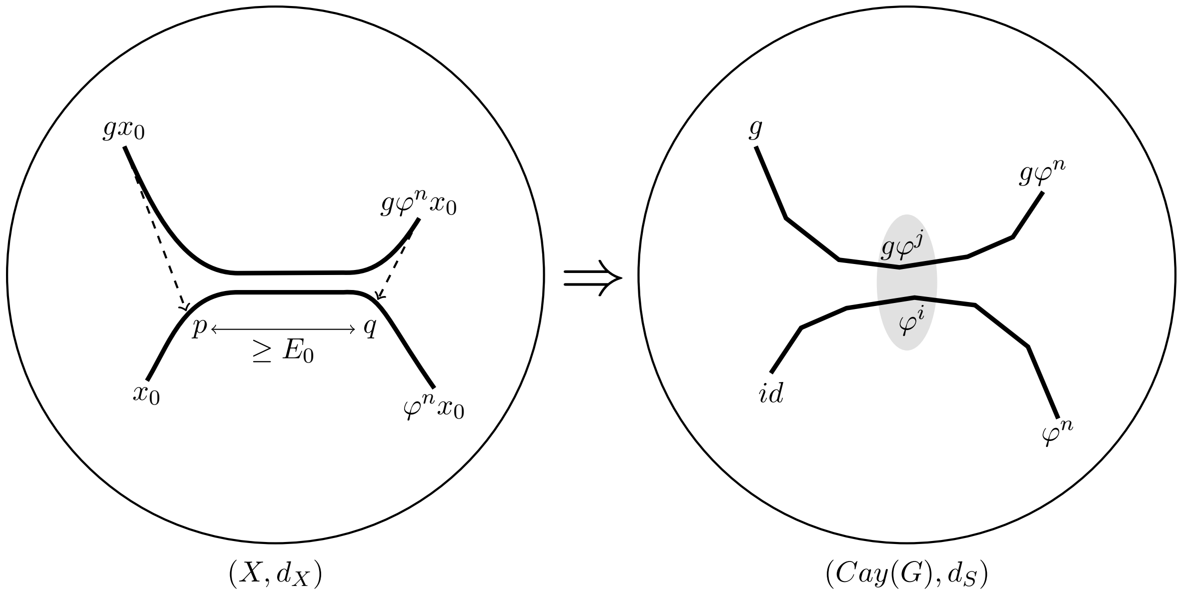

FIGURE 3. Picture for the situation in Fact 2.6(2)

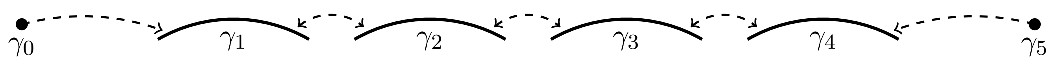

FIGURE 4. Aligned sequence of geodesics $(\gamma_0, \gamma_1, \ldots, \gamma_5)$ is aligned, where γ_0, γ_5 are degenerate geodesics, i.e., points.

The following facts are straightforward, whose proofs can be found in [Cho24a, Appendix].

Fact 2.8. *Let γ be a geodesic in a metric space. Let γ_1 and γ_2 be subsegments of γ, with γ_1 appearing earlier than γ_2. Let κ_1 and κ_2 be geodesics that are K-fellow traveling with γ_1 and γ_2, respectively. Then (κ_1, κ_2) is $6K$-aligned.*

Fact 2.9. *The following holds for each $K > 0$ and $L \geq 12K$. Let γ be a geodesic in a metric space and let γ_1 and γ_2 be subsegments of γ such that $\gamma_1 \cap \gamma_2$ has length L. Let $[x, y]$ and κ_2 be geodesics that are K-fellow traveling with γ_1 and γ_2, respectively. Then $\pi_\kappa(x)$ appears earlier than $\pi_\kappa(y)$ along κ, and $d_X(\pi_\kappa(x), \pi_\kappa(y)) > L - 10K$.*

We now record a version of Behrstock's inequality [Beh06, Theorem 4.3] (cf. [Sis18, Lemma 2.5]) and its consequences. The proofs can be found in [Cho24a, Section 3, Appendix].

Fact 2.10. *Let X be a δ-hyperbolic space. Let $x \in X$ and let (γ_1, γ_2) be a K-aligned sequence of geodesics in X. Then either (x, γ_2) is $(K + 60\delta)$-aligned or (γ_1, x) is $(K + 60\delta)$-aligned.*

Definition 2.11. *Let $K > 0$ and let $\gamma_1, \gamma_2, \ldots, \gamma_n$ be finite geodesics on $\mathcal{C}(\Sigma)$ (which includes the case of degenerate geodesics, i.e., points in $\mathcal{C}(\Sigma)$). We say that $(\gamma_1, \ldots, \gamma_n)$ is K-aligned if for each $i = 1, \ldots, n-1$ we have*

$$\mathrm{diam}_{\mathcal{C}}\left(\pi_{\gamma_i}(\gamma_{i+1}) \cup (\textit{ending point of } \gamma_i)\right) < K \textit{ and}$$
$$\mathrm{diam}_{\mathcal{C}}\left(\pi_{\gamma_{i+1}}(\gamma_i) \cup (\textit{beginning point of } \gamma_{i+1})\right) < K.$$

Fact 2.12. *Let X be a δ-hyperbolic space. Let $n \geq 3$ and let $(\gamma_1, \ldots, \gamma_n)$ be a K-aligned sequence of geodesics in X. Suppose that $\gamma_2, \ldots, \gamma_{n-1}$ are longer than $2K + 120\delta$. Then (γ_i, γ_j) is $(K + 60\delta)$-aligned for each $1 \leq i < j \leq n$.*

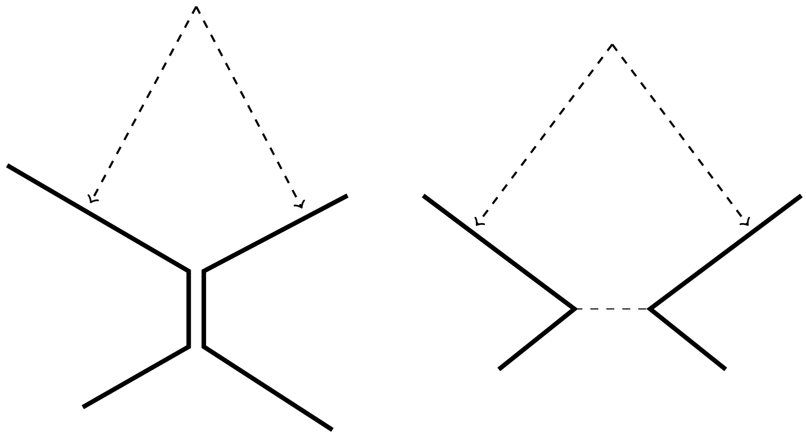

Figure 5. Prohibited configurations in a δ-hyperbolic space explaining Fact 2.10

Fact 2.13. *Let X be a δ-hyperbolic space. Let $x, y \in X$ and let $\gamma_1, \ldots, \gamma_n$ be geodesics in X, longer than $2K + 140\delta$ each, such that $(x, \gamma_1, \ldots, \gamma_n, y)$ is K-aligned.*

Then there exist disjoint subsegments $\eta_1, \ldots, \eta_n$ of $[x, y]$ such that

(1) $\eta_1, \ldots, \eta_n$ are in order from left to right along $[x, y]$, i.e., η_i appears earlier than η_{i+1} along $[x, y]$ for each $i = 1, \ldots, n-1$, and
(2) γ_i and η_i are $(K + 80\delta)$-fellow traveling for each $i = 1, \ldots, n$.

Let G be a group and let $S \subseteq G$ be its finite generating set. The word metric d_S is defined by

$$d_S(g, h) := \min \left\{ n \in \mathbb{Z}_{\geq 0} : \begin{array}{c} \exists\, a_1, a_2, \ldots, a_n \in S,\ \epsilon_1, \epsilon_2, \ldots, \epsilon_n \in \{1, -1\} \\ \text{such that } g^{-1}h = a_1^{\epsilon_1} a_2^{\epsilon_2} \cdots a_n^{\epsilon_n}. \end{array} \right\}$$

We use the notation for the word norm $\|g\|_S := d_S(id, g)$. We define

$$B_S(n) := \big\{ g \in \mathrm{Mod}(\Sigma) : d_S(id, g) \leq n \big\}.$$

We denote by $[g, h]_S$ an arbitrary d_S-geodesic between $g, h \in G$. By a d_S-path, we mean a sequence of group elements $P = (g_1, g_2, \ldots, g_n)$ such that $d_S(g_i, g_{i+1}) = 1$ for each i; we denote n by $Len(P)$.

When the group G acts on a metric space $X \ni x_0$, we often define

$$\begin{aligned} \|g\|_X &:= d_X(x_0, gx_0) \quad (g \in G), \\ K_{Lip} &:= \max_{s \in S} \|s\|_X. \end{aligned}$$

Then we have $\|g\|_X \leq K_{Lip} \|g\|_S$ for each $g \in G$.

3. WPD property and contraction

It is well-known that a loxodromic isometry φ of a δ-hyperbolic space $X \ni x_0$ has strictly positive asymptotic translation length $\tau := \lim_n d_X(x_0, \varphi^n x_0)/n$. Moreover, its orbit $\{\varphi^i x_0\}_{i \in \mathbb{Z}}$ is a quasigeodesic and hence quasi-convex. In summary,

Fact 3.1. *Let φ be a loxodromic isometry of a δ-hyperbolic space $X \ni x_0$. Then there exists $\mathcal{G} > 0$ such that the sequence $(\varphi^i x_0, \ldots, \varphi^j x_0)$ and the geodesic $[\varphi^i x_0, \varphi^j x_0]$ are $\mathcal{G}$-fellow traveling for each $i \leq j$. Furthermore, the sequence $(\varphi^i x_0)_{i \in \mathbb{Z}}$ is a $\mathcal{G}$-coarse geodesic, i.e.,*

$$d_X(\varphi^i x_0, \varphi^l x_0) \geq d_X(\varphi^i x_0, \varphi^j x_0) + d_X(\varphi^j x_0, \varphi^l x_0) - \mathcal{G} \quad (\forall i \leq j \leq l).$$

In Subsection 1.1, we claimed that $f(D, M) < +\infty$ for each $D, M > 0$ for every acylindrically hyperbolic group. We prove a variant of this fact.

Lemma 3.2. *Let G be a non-virtually cyclic group with a finite generating set $S \subseteq G$. Suppose that G acts on a δ-hyperbolic space $X \ni x_0$ with a WPD loxodromic element $\varphi \in G$. Then there exists $D_0 > 0$, and for each $k, M > 0$ there exists $R = R(k, M) > 0$, such that the following holds.*

Let $g, h \in G$ be such that $\|g\|_S > R$ and $\|h\|_S \le M$. Then $\pi_{[x_0, \varphi^k x_0]}(\{gx_0, ghx_0\})$ has diameter at most D_0.

This lemma closely resembles [Sis16, Lemma 3.3] and [MS20, Lemma 8.1]. Here, the crucial point is that D_0 is uniform and is independent from k, M and R.

Proof. Let $\mathcal{G} > 0$ be the constant for φ as in Fact 3.1. For $K = 24\mathcal{G} + 130\delta$, we pick N such that $Stab_K(x_0, \varphi^N x_0)$ is finite using the WPD property of φ. We then set $D_0 := 1002\mathcal{G} + ND_\varphi + 1000\delta$, where $D_\varphi := d_X(x_0, \varphi x_0)$.

To prove the lemma, let $k, M > 0$ and denote $\gamma := [x_0, \varphi^k x_0]$. Suppose to the contrary that there does not exist R for (k, M). That means, suppose that there exist a sequence $(g_1, g_2, \ldots)$ of distinct elements of G and a sequence $(h_1, h_2, \ldots)$ in $B_S(M)$ such that

$$\operatorname{diam}\big(\pi_\gamma(\{g_i x_0, g_i h_i x_0\})\big) \ge D_0 \quad (\forall i > 0).$$

Let p_i, q_i be points in $\pi_\gamma(\{g_i x_0, g_i h_i x_0\})$ that are at least D_0-apart. Recall that the nearest point projection of a single point onto γ has diameter at most $20\delta < D_0$ (Fact 2.2(1)). Hence, up to relabelling, we can say that $p_i \in \pi_\gamma(g_i x_0)$ and $q_i \in \pi_\gamma(g_i h_i x_0)$.

Since $\gamma = [x_0, \varphi^k x_0]$ is compact and $B_S(M)$ is finite, we can take a subsequence and assert that:

$$\begin{aligned} h_1 = h_2 = \ldots &=: h, \\ d_X(p_i, p_j), d_X(q_i, q_j) &< \mathcal{G} \quad (\forall i, j > 0). \end{aligned}$$

Either p_i's appear earlier than or later than q_i's. In the latter case, we can replace h with h^{-1}, g_i with $g_i h$ and swap p_i's with q_i's. Hence, we may assume that p_i's appear earlier than q_i's.

By Fact 2.2(2), there exists $[\alpha_i, \beta_i] \subseteq [g_i x_0, g_i h x_0]$ that is 20δ-fellow traveling with $[p_i, q_i]$. Note that $0 \le d_X(g_i x_0, \alpha_i) \le d_X(x_0, hx_0)$. By taking further subsequence, we can obtain T such that

$$\big|d_X(g_i x_0, \alpha_i) - T\big| < \delta \quad (\forall i > 0).$$

By Fact 3.1, there exist $l, m \in \mathbb{Z}$ such that $d_X(\varphi^l x_0, p_1) < \mathcal{G}$ and $d_X(\varphi^m x_0, q_1) < \mathcal{G}$. Note that

$$D_\varphi \cdot |l - m| \ge d_X(\varphi^l x_0, \varphi^m x_0) > d_X(p_1, q_1) - 2\mathcal{G} > 1000\mathcal{G} + ND_\varphi. \tag{3.1}$$

Here, if $m \le l$ then

$$\begin{aligned} d_X(x_0, q_1) &\le d_X(x_0, \varphi^m x_0) + \mathcal{G} \le d_X(x_0, \varphi^l x_0) - d_X(\varphi^m x_0, \varphi^l x_0) + 2\mathcal{G} \\ &\le d_X(x_0, \varphi^l x_0) - 1000\mathcal{G} + 2\mathcal{G} < d_X(x_0, \varphi^l x_0) - 998\mathcal{G} \le d_X(x_0, p_1) - 997\mathcal{G}. \end{aligned}$$

This contradicts the fact that p_1 appears earlier than q_1 on γ. Hence, we have $l < m$.

Now Inequality 3.1 implies that $l + N$ lies between l and m. By Fact 3.1, $\varphi^{l+N} x_0$ lies in a $\mathcal{G}$-neighborhood of $[\varphi^l x_0, \varphi^m x_0]$. Note that $d_X(\varphi^l x_0, p_i) \le d_X(\varphi^l x_0, p_1) + d_X(p_1, p_i) \le 2\mathcal{G}$ for each i, and similarly $\varphi^m x_0$ and q_i are $2\mathcal{G}$-close.

By Fact 2.1 $[\varphi^l x_0, \varphi^m x_0]$ is $(4\mathcal{G} + 2\delta)$-fellow traveling with $[p_i, q_i]$, which is 20δ-fellow traveling with $[\alpha_i, \beta_i]$. Thus, there exists $c_i \in [\alpha_i, \beta_i]$ such that $d_X(c_i, \varphi^{l+N} x_0) < 5\mathcal{G} + 22\delta$. We have

$$\begin{aligned} \big|d_X(\alpha_i, c_i) - d_X(\varphi^l x_0, \varphi^{l+N} x_0)\big| &\le d_X(\alpha_i, \varphi^i x_0) + d_X(c_i, \varphi^{l+N} x_0) \\ &\le d_X(\alpha_i, p_i) + d_X(p_i, p_1) + d_X(p_1, \varphi^l x_0) + d_X(c_i, \varphi^{l+N} x_0) \\ &\le 20\delta + \mathcal{G} + \mathcal{G} + (5\mathcal{G} + 22\delta) \le 7\mathcal{G} + 42\delta. \end{aligned}$$

Now, for each i we have four points $g_i g_1^{-1} \alpha_1$, $g_i g_1^{-1} c_1$, α_i on the geodesic

$$g_i \cdot g_1^{-1}([g_1 x_0, g_1 h x_0]) = [g_i x_0, g_i h x_0].$$

Recall that $d_X(g_i x_0, g_i g_1^{-1}\alpha_1) = d_X(g_1 x_0, \alpha_1)$ and $d_X(g_i x_0, \alpha_i)$ are both δ-close to T. This implies that $g_i g_1^{-1}\alpha_1$ and α_i are 2δ-close. Hence, we have

$$\begin{aligned}
d_X(\varphi^l x_0, g_i g_1^{-1} \cdot \varphi^l x_0) &\le d_X(\varphi^l x_0, \alpha_i) + d_X(\alpha_i, g_i g_1^{-1}\alpha_1) + d_X(g_i g_1^{-1}\alpha_1, g_i g_1^{-1} \cdot \varphi^l x_0) \\
&\le d_X(\varphi^l x_0, p_i) + d_X(p_i, \alpha_i) + 2\delta + d_X(\varphi^l x_0, p_1) + d_X(p_1, \alpha_1) \\
&\le (20\delta + 2\mathcal{G}) + 2\delta + (\mathcal{G} + 20\delta) = 42\delta + 3\mathcal{G}.
\end{aligned}$$

Next, $d_X(g_i x_0, c_i) = d_X(g_i x_0, \alpha_i) + d_X(\alpha_i, c_i)$ is $(7\mathcal{G} + 43\delta)$-close to $T + d_X(x_0, \varphi^N x_0)$. So is $d_X(g_i x_0, g_i g_1^{-1} c_1) = d_X(g_1 x_0, c_1)$. Hence, c_i and $g_i g_1^{-1} c_1$ are $(14\mathcal{G} + 86\delta)$-close. We conclude

$$\begin{aligned}
d_X(\varphi^{l+N} x_0, g_i g_1^{-1} \cdot \varphi^{l+N} x_0) &\le d_X(\varphi^{l+N} x_0, c_i) + d_X(c_i, g_i g_1^{-1} c_1) + d_X(g_i g_1^{-1} c_1, g_i g_1^{-1} \cdot \varphi^{l+N} x_0) \\
&\le (5\mathcal{G} + 22\delta) + (14\mathcal{G} + 86\delta) + (5\mathcal{G} + 22\delta) < 24\mathcal{G} + 130\delta.
\end{aligned}$$

To summarize, $\varphi^{-l} g_i g_1^{-1} \varphi^l$ belongs to $Stab_K(x_0, \varphi^N x_0)$ for each i. Furthermore, we have

$$\varphi^{-l} g_i g_1^{-1} \varphi^l = \varphi^{-l} g_j g_1^{-1} \varphi^l \ \Leftrightarrow\ g_i = g_j.$$

Since $g_1, g_2, \ldots$ are distinct, it follows that $Stab_K(x_0, \varphi^N x_0)$ is infinite, a contradiction. □

Proposition 3.3. *Let G be a non-virtually cyclic group and let $S \subseteq G$ be its finite generating set. Suppose that G acts on a δ-hyperbolic space $X \ni x_0$ with a WPD loxodromic element $\varphi \in G$. Then for each $K > 0$ there exists $L_0 = L_0(K)$ such that, for each $L \ge L_0$ and for each $M > 0$ there exists $R_0 = R_0(L, M) > 0$ satisfying the following.*

Let P_l be a d_S-path connecting $a_l \in G$ to $b_l \in G$ for $l = 1, 2$. Let $g_1, \ldots, g_m \in G$ be such that

$$(a_i x_0, g_1[x_0, \varphi^L x_0], \ldots, g_m[x_0, \varphi^L x_0], b_i x_0) \text{ is } K\text{-aligned.} \quad (i = 1, 2)$$

Let $k \le m$. Then one of the following happens:

(1) For each subset $I \subseteq \{1, \ldots, m\}$ of cardinality k, there exist $i \in I$ such that

$$d_S(P_l, g_i) < R_0 \quad (l = 1, 2); \text{ or,}$$

(2) $Len(P_1) + Len(P_2) \ge M \cdot k$.

Proof. Let D_0 be as in Lemma 3.2. Let $K_{Lip} = \max_{s \in S} \|s\|_X$, let $\tau := \lim_n \frac{1}{n} d_X(x_0, \varphi^n x_0)$ and let

$$L_0 = \frac{1}{\tau}\big(2K + 2K_{Lip} + 1000\delta + D_0\big).$$

Now given $L > L_0$, we will declare $R_0 = R_0(L, M)$ following Lemma 3.2.

Consider P_1, P_2 and g_i's as in the proposition. Note that

$$\operatorname{diam}_X([x_0, \varphi^L x_0]) \ge \tau L \ge 2K + 2K_{Lip} + 1000\delta + D_0. \tag{3.2}$$

Thanks to this inequality, we can apply Fact 2.12 to the K-aligned translates of $[x_0, \varphi^L x_0]$.

We will negate the case (1) and prove that (2) holds. For this, let $I \subseteq \{1, \ldots, m\}$ be a k-element subset such that g_i is not simultaneously R_0-close to P_1 and P_2 for each $i \in I$. Let

$$I_l := \{i : d_S(P_l, g_i) \ge R_0\}. \quad (l = 1, 2)$$

Then $I_1 \cup I_2$ has cardinality at least k. For convenience, we will denote $g_i[x_0, \varphi^L x_0]$ by γ_i.

For each $i \in \{1, \ldots, m\}$, let p_i be the latest vertex of P_1 such that $(p_i x_0, \gamma_i)$ is $(K + 60\delta)$-aligned. Such a vertex exists because $(a_i x_0, \gamma_i)$ is $(K + 60\delta)$-aligned by Fact 2.12. Now let

$$V_i := \{v \in P_1 : v \text{ comes later than } p_i \text{ along } P_1 \text{ and } (\gamma_i, v x_0) \text{ is } (K + 60\delta)\text{-aligned}\};$$

V_i is nonempty because (γ_i, b_i) is $(K + 60\delta)$-aligned by Fact 2.12. Let q_i be the earliest vertex in V_i. Lastly, let p_i' be the vertex of P_1 right after p_i, and let q_i' be the one right before q_i.

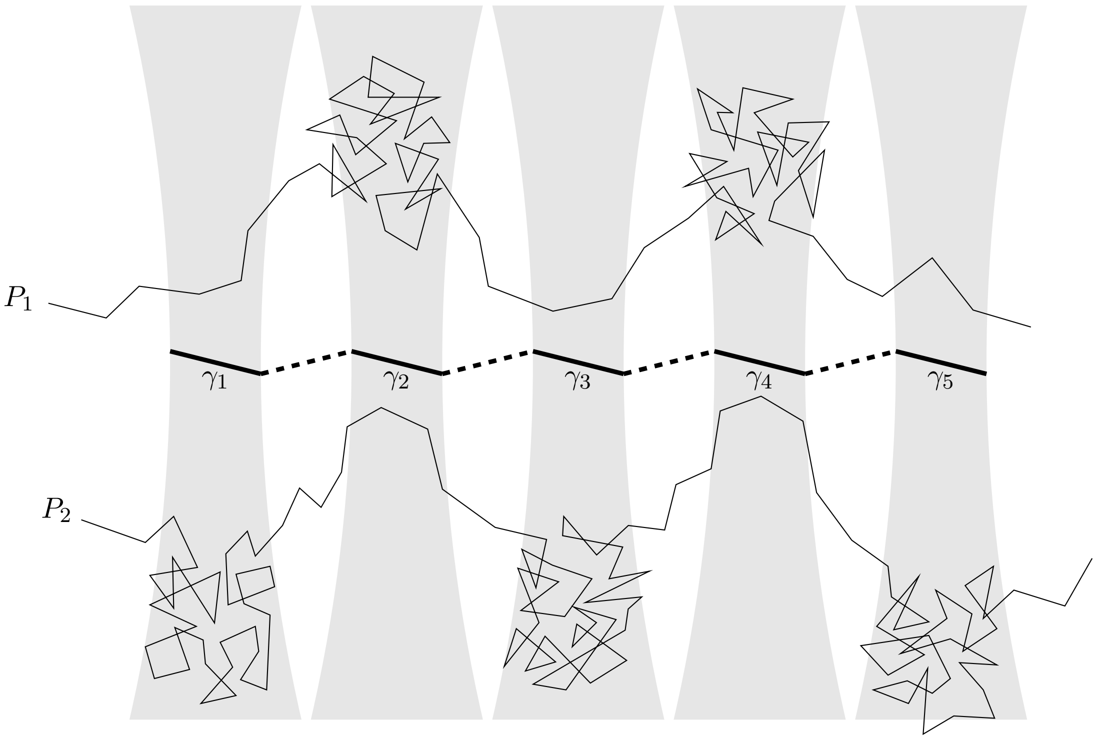

FIGURE 6. Schematics for Proposition 3.3. Thick segments represent $\gamma_i := g_i[x_0, \varphi^L x_0]$ for $i = 1, \ldots, 5$. The i-th Shadowed regions consist of points that project onto γ_i in the middle. In this example, P_1 is far away from γ_2 and γ_4, whereas P_2 is far away from $\gamma_1, \gamma_3, \gamma_5$. This makes P_1 and P_2 lengthy.

Note that $\pi_{\gamma_i}(\{p_i x_0, p_i' x_0\})$ has diameter at most $K_{Lip} + 12\delta$ by Fact 2.2(1). Hence, $\pi_{\gamma_i}(p_i' x_0)$ is contained in the beginning $(K_{Lip} + K + 80\delta)$-subsegment of γ_i and does not meet the ending $(K_{Lip} + 60\delta)$-subsegment (Inequality 3.2). This implies $p_i' \notin V_i$, and q_i comes later than p_i'. Let

$$V_i' := \{v \in P_1 : v \text{ in between } p_i \text{ and } q_i \text{ (excluding } p_i, q_i)\}.$$

We have then observed that $p_i', q_i' \in V_i'$ and V_i' is nonempty. For each $v \in V_i'$ we conclude that

$$\text{neither } (\gamma_i, v x_0) \text{ nor } (v x_0, \gamma_i) \text{ is } (K + 60\delta)\text{-aligned.}$$

Now repeated application of Fact 2.10 tells us that for $i \neq j$,

$$v \in V_i' \Rightarrow \left\{ \begin{array}{ll} (\gamma_j, v x_0) \text{ is } (K+60\delta)\text{-aligned} & \text{if } j < i \\ (v x_0, \gamma_j) \text{ is } (K+60\delta)\text{-aligned} & \text{if } j > i \end{array} \right\} \Rightarrow v \notin V_j'.$$

In conclusion, $V_1', \ldots, V_m'$'s are disjoint subpaths of P_1.

Now observe for each i that

$$\begin{aligned} \operatorname{diam}_X(\pi_{\gamma_i}(\{p_i', q_i'\})) &\geq \operatorname{diam}(\gamma_i) - 2(K + 60\delta) - \operatorname{diam}(\pi_{\gamma_i}(p_i x_0, p_i' x_0)) - \operatorname{diam}(\pi_{\gamma_i}(q_i x_0, q_i' x_0)) \\ &\geq \tau L - 2(K + 60\delta) - 2 \cdot (K_{Lip} + 20\delta) > D_0. \end{aligned}$$

If $i \in I_1$, we additionally know that $d_S(p_i', g_i) \geq R_0$. Lemma 3.2 then implies that $d_S(p_i', q_i') > M$. Hence, $Len(V_i') \geq M$ for each $i \in I_1$. Summing up, we obtain $Len(P_1) \geq M \cdot \#I_1$.

Similar logic implies $Len(P_2) \geq M \cdot \#I_2$. We thus conclude

$$Len(P_1) + Len(P_2) \geq M \cdot (\#I_1 \cup \#I_2) \geq M \cdot k. \qquad \square$$

4. Proof of Theorem 1.1

Throughout, let G be a non-virtually cyclic group with a finite generating set S. Suppose that G acts on a δ-hyperbolic space $X \ni x_0$ with a WPD loxodromic element $\varphi \in G$. When a constant L is understood, we will use the notation

$$\Upsilon_L := [x_0, \varphi^L x_0].$$

Since G contains independent loxodromics, G has exponential growth. In other words,

$$\lambda_S := \liminf_n \frac{\ln \# B_S(n)}{n} > 1.$$

This immediately implies that:

Fact 4.1. *For each sufficiently large n we have*

$$\# B_S(0.9n)/\# B_S(n) \le \lambda_S^{0.05n}.$$

Let us fix some more constants for the proof. Let E_0 be as in Fact 2.6. Let $K_{Lip} := \max_{s\in S} \|s\|_X$, let $\tau := \lim_n \|\varphi^n\|_X/n$ (so that $\|\varphi^k\|_X \ge k\tau$ for each k). Let $F_0 := \|\varphi\|_S$. Finally, let L_0 be as in Proposition 3.3 for $K = 100(E_0 + 1000\delta + K_{Lip})$, and

$$L_1 := L_0 + \frac{1}{\tau} 100(E_0 + 1000\delta + K_{Lip}).$$

This choice implies that:

Fact 4.2. *For each $L > L_1$,*

$$d_X(x_0, \varphi^L x_0) - 40(E_0 + 1000\delta + K_{Lip}) \ge \tau L - 40(E_0 + 1000\delta + K_{Lip}) \ge 0.5\tau L + 140\delta.$$

Given $L, \epsilon > 0$, we define

$$\mathcal{V}_{L,\epsilon}(n) := \left\{ g \in B_S(n) : \begin{array}{c} \text{there exist } h_1, \ldots, h_{\epsilon n} \in G \text{ such that} \\ (x_0, h_1\Upsilon_L, \ldots, h_{\epsilon n}\Upsilon_L, gx_0) \text{ is } (6E_0 + 300\delta)\text{-aligned} \end{array} \right\},$$

$$\mathcal{BAD}_{L,\epsilon}(n) := B_S(n) \setminus \big(B_S(0.9n) \cup \mathcal{V}_{L,\epsilon}(n)\big).$$

Lemma 4.3. *For each $L > L_1$ and $\epsilon > 0$, we have*

$$\limsup_n \frac{\#\mathcal{BAD}_{L,\epsilon}(n)}{\# B_S(n)} < 5 \cdot (2E_0 + 4\,L\,F_0 + 5) \cdot (\# S)^{E_0 + 3\,L\,F_0 + 4} \cdot \epsilon.$$

Proof of Lemma 4.3. Let us define a map

$$F : Dom(F) := \mathcal{BAD}_{L,\epsilon}(n) \times \{1, \ldots, 0.9n\} \to B_S(n)$$

as follows. Given $(g, i) \in \mathcal{BAD}_{L,\epsilon}(n) \times \{1, \ldots, 0.9n\}$, we first fix a d_S-geodesic representative $g = a_1 a_2 \cdots a_{\|g\|_S}$. By Fact 2.6, there exist $s = s(g,i)$ and $t = t(g,i)$ in $S \cup \{id\}$ such that

$$\big(s^{-1} \cdot (a_1 \cdots a_i)^{-1} x_0, \varphi^L x_0\big)_{x_0} < E_0, \big(\varphi^{-L} x_0, t \cdot a_{i + F_0 L + 3} \cdots a_{\|g\|} x_0\big)_{x_0} < E_0.$$

We then define

$$h(g,i) := a_1 \cdots a_i \cdot s,\ h'(g,i) := t \cdot a_{i+F_0 L + 3} \cdots a_{\|g\|},\ F(g,i) := h(g,i)\varphi^L h'(g,i).$$

Note that $F(g,i) \in B_S(n)$ because

$$\begin{aligned} \|F(g,i)\|_S &\le \|h(g,i)\|_S + \|\varphi^L\|_S + \|h'(g,i)\|_S \\ &\le (i+1) + F_0\,L + (\|g\|_S - i - F_0\,L - 2) + 1 \le \|g\|_S \le n. \end{aligned}$$

Before the proof, we first declare

$$T := (2E_0 + 4\,L\,F_0 + 5) \cdot \# B_S(E_0 + 2\,L\,F_0) \cdot \# B_S(F_0\,L + 4)$$

and $R_0 = R_0(L, 4/\epsilon)$ as in Proposition 3.3.

Claim 4.4. *Let m be the maximum number of elements $(g_1, i_1), \ldots, (g_m, i_m) \in \mathcal{BAD}_{L,\epsilon}(n) \times \{1, \ldots, 0.9n\}$ such that*

(1) $F(g_1, i_1) = \ldots = F(g_m, i_m) =: U$,

(2) $\Big(x_0,\, h(g_1, i_1) \cdot \Upsilon_L,\, \ldots,\, h(g_m, i_m) \cdot \Upsilon_L,\, Ux_0\Big)$ is $6(E_0 + 30\delta)$-aligned.

Then $\#F^{-1}(U) \le 2T \cdot m$ holds for each $U \in B_S(n)$.

Proof of Claim 4.4. Fix an arbitrary $U \in B_S(n)$. For each $(g, i) \in F^{-1}(U)$, we have:

(1) $(x_0, h(g, i)\varphi^L x_0)_{h(g,i)x_0} < E_0$; hence $\pi_{h(g,i)\Upsilon_L}(x_0)$ is $(E_0 + 8\delta)$-close to $h(g, i)x_0$ (Fact 2.4).

(2) Similarly, the projection of Ux_0 onto $h(g, i)\Upsilon_L$ is $(E_0 + 8\delta)$-close to $h(g, i)\varphi^L x_0$.

(3) $h(g, i)x_0$ and $h(g, i)\varphi^L x_0$ are at least $\tau\, L$-apart, which is much larger than 20δ.

Now Fact 2.2 guarantees a subsegment $\gamma(g, i)$ of $[x_0, Ux_0]$ and a subsegment $\eta = [p, q]$ of $h(g, i)\Upsilon_L$ that are 20δ-fellow traveling. Here, p and $h_k x_0$, and q and $h_k \varphi^L x_0$ are pairwise $(E_0 + 8\delta)$-close. Hence, $\gamma(g, i)$ and $h(g, i)\Upsilon_L$ are $(E_0 + 30\delta)$-fellow traveling. It follows that $\gamma(g, i)$'s are longer than $\tau\, L - 2(E_0 + 30\delta) \ge 25(E_0 + 30\delta)$.

We now pick a maximal subset $\mathcal{A}$ of $F^{-1}(U)$ such that

$$\text{for any } (g, i), (g', i') \in \mathcal{A},\ \gamma(g, i) \text{ and } \gamma(g', i') \text{ overlap for length at most } 12(E_0 + 30\delta).$$

We claim that $\#F^{-1}(U) \le T \cdot \#\mathcal{A}$. To show this, pick an arbitrary $(g, i) \in F^{-1}(U)$. Let $a_1 \cdots a_{\|g\|_S}$ be the geodesic representative for g that was used when defining

$$h(g, i) := a_1 \cdots a_i \cdot s(g, i), \quad h(g, i)' := t(g, i) \cdot a_{i+F_0\, L+3} + \cdots a_{\|g\|_S}.$$

By the maximality of $\mathcal{A}$, there exists $(\mathfrak{g}, \mathfrak{i}) \in \mathcal{A}$ such that $\gamma(g, i)$ and $\gamma(\mathfrak{g}, \mathfrak{i})$ overlap for length at least $12(E_0 + 30\delta)$. Recall that $h(g, i)\Upsilon_L$ and $h(\mathfrak{g}, \mathfrak{i})\Upsilon_L$ are $(E_0 + 30\delta)$-fellow traveling $\gamma(g, i)$ and $\gamma(\mathfrak{g}, \mathfrak{i})$, respectively. By Fact 2.9, $\pi_{h(g,i)\Upsilon_L}(h(\mathfrak{g}, \mathfrak{i})x_0)$ appears earlier than $\pi_{h(g,i)\Upsilon_L}(h(\mathfrak{g}, \mathfrak{i})\varphi^L x_0)$. Moreover, they are $(12(E_0 + 30\delta) - 10(E_0 + 30\delta))$-apart and hence E_0-apart. By Fact 2.6, we conclude that $\varphi^{-a} \cdot h(g, i)^{-1} h(\mathfrak{g}, \mathfrak{i})\varphi^b \in B_S(E_0)$ for some $a, b \in \{0, \ldots, L\}$. We conclude that

$$h(g, i) \in h(\mathfrak{g}, \mathfrak{i}) \cdot \{\varphi^a : a = 0, \ldots, L\} \cdot B_S(E_0) \cdot \{\varphi^{-a} : a = 0, \ldots, L\} \subseteq h(\mathfrak{g}, \mathfrak{i}) B_S(E_0 + 2\, L\, F_0).$$

This also implies that $\|h(g, i)\|_S$ and $\|h(\mathfrak{g}, \mathfrak{i})\|_S$ differ by at most $E_0 + 2\, L\, F_0$, and hence

$$i \in \big[\mathfrak{i} - (E_0 + 2\, L\, F_0 + 2), \mathfrak{i} + (E_0 + 2\, L\, F_0 + 2)\big].$$

Note also that

$$h(g, i)' = \varphi^{-\, L} h(g, i)^{-1} \cdot U$$

is determined as soon as $h(g, i)$ is determined.

Finally, in order to reconstruct $g = a_1 \cdots a_{\|g_l\|_S}$ from $h(g, i)$ and $h(g, i)$, it suffices to pick $c := s_l^{-1} a_{i_l+1} \cdots a_{i_l + F_0\, L + 2} t^{-1} \in B_S(F_0\, L + 4)$ and multiply $h(g, i)$, c and $h(g', i')$. In summary, we have

$$F^{-1}(U) \subseteq \bigcup_{(\mathfrak{g}, \mathfrak{i}) \in \mathcal{A}} \left(\Big\{ h(\mathfrak{g}, \mathfrak{i}) f \cdot c \cdot \varphi^{-\, L} f^{-1} h(\mathfrak{g}, \mathfrak{i})^{-1} U : f \in B_S(E_0 + 2\, L\, F_0),\ c \in B_S(F_0\, L + 4) \Big\} \times I(\mathfrak{i}) \right)$$

where $I(\mathfrak{i}) := [\mathfrak{i} - (E_0 + 2\, L\, F_0 + 2), \mathfrak{i} + (E_0 + 2\, L\, F_0 + 2)]$. From this, we conclude $\#F^{-1}(U) \le T \cdot \#\mathcal{A}$.

Next, let us enumerate $\mathcal{A}$ as

$$\mathcal{A} = \{(\mathfrak{g}_1, \mathfrak{i}_1), (\mathfrak{g}_2, \mathfrak{i}_2), \ldots\}$$

so that $\gamma(\mathfrak{g}_l, \mathfrak{i}_l)$ starts earlier than $\gamma(\mathfrak{g}_{l+1}, \mathfrak{i}_{l+1})$ along $[x_0, Ux_0]$, for each l. Then the beginning point of $\gamma(\mathfrak{g}_2, \mathfrak{i}_2)$ is later than that of $\gamma(\mathfrak{g}_1, \mathfrak{i}_1)$ and earlier than that of $\gamma(\mathfrak{g}_3, \mathfrak{i}_3)$. $(*)$ Moreover, $\gamma(\mathfrak{g}_2, \mathfrak{i}_2)$ does not contain $\gamma(\mathfrak{g}_3, \mathfrak{i}_3)$, as their overlap should not be longer than $12(E_0 + 30\delta)$ while $\gamma(\mathfrak{g}_3, \mathfrak{i}_3)$ is longer than $25(E_0 + 30\delta)$. Hence, the ending point of $\gamma(\mathfrak{g}_2, \mathfrak{i}_2)$ is earlier than that of $\gamma(\mathfrak{g}_3, \mathfrak{i}_3)$. $(**)$

At this point, if $\gamma(\mathfrak{g}_1, \mathfrak{i}_1)$ and $\gamma(\mathfrak{g}_3, \mathfrak{i}_3)$ intersect, then $\gamma(\mathfrak{g}_3, \mathfrak{i}_3)$ is completely covered by $\gamma(\mathfrak{g}_1, \mathfrak{i}_1)$ and $\gamma(\mathfrak{g}_3, \mathfrak{i}_3)$ due to $(*)$ and $(**)$. We would then have

$$\mathrm{diam}_X \left(\gamma(\mathfrak{g}_1, \mathfrak{i}_1) \cap \gamma(\mathfrak{g}_2, \mathfrak{i}_2)\right) + \mathrm{diam}_X \left(\gamma(\mathfrak{g}_2, \mathfrak{i}_2) \cap \gamma(\mathfrak{g}_3, \mathfrak{i}_3)\right) \geq \mathrm{diam}_X \left(\gamma(\mathfrak{g}_2, \mathfrak{i}_2)\right) \geq 25(E_0 + 30\delta),$$

which contradicts to the bound $12(E_0 + 30\delta)$ on $\mathrm{diam}_X \left(\gamma(\mathfrak{g}_1, \mathfrak{i}_1) \cap \gamma(\mathfrak{g}_2, \mathfrak{i}_2)\right)$ and $\mathrm{diam}_X \big(\gamma(\mathfrak{g}_2, \mathfrak{i}_2) \cap \gamma(\mathfrak{g}_3, \mathfrak{i}_3)\big)$. Hence, we conclude that $\gamma(\mathfrak{g}_1, \mathfrak{i}_1)$ and $\gamma(\mathfrak{g}_3, \mathfrak{i}_3)$ do not intersect.

With the same logic, we conclude that $\gamma(\mathfrak{g}_l, \mathfrak{i}_l)$'s for odd integers l are disjoint subsegments of $[x_0, Ux_0]$, in order from left to right along $[x_0, Ux_0]$. Recall again that $\gamma(\mathfrak{g}_l, \mathfrak{i}_l)$ and $h(\mathfrak{g}_l, \mathfrak{i}_l))\Upsilon_L$ are $(E_0 + 30\delta)$-fellow traveling. Now Fact 2.8 tells us that

$$\left(x_0, h(\mathfrak{g}_1, \mathfrak{i}_1)\Upsilon_L,\, h(\mathfrak{g}_3, \mathfrak{i}_3)\Upsilon_L,\, \ldots,\, h(\mathfrak{g}_{2\lceil \#\mathcal{A}/2\rceil+1}, \mathfrak{i}_{2\lceil \#\mathcal{A}/2\rceil+1})\Upsilon_L,\, Ux_0\right)$$

is $6(E_0 + 30\delta)$-aligned. This implies that $m \geq \lceil \#\mathcal{A}/2 \rceil \geq \frac{1}{2T}\#F^{-1}(U)$ as desired. □

Now let $(g_1, i_1), \ldots, (g_m, i_m) \in \mathcal{BAD}_{L,\epsilon}(n) \times \{1, \ldots, 0.9n\}$ the elements as in Claim 4.4:

(1) $F(g_1, i_1) = \ldots = F(g_m, i_m) =: U$,
(2) $\Big(x_0,\, h(g_1, i_1) \cdot \Upsilon_L,\, \ldots,\, h(g_m, i_m) \cdot \Upsilon_L,\, Ux_0\Big)$ is $6(E_0 + 30\delta)$-aligned.

It remains to prove that $m < 2\epsilon n$ for large enough n. We will prove it for every $n \geq 32R_0K_{Lip}/\tau\, L\,\epsilon$. Suppose to the contrary that $m \geq 2\epsilon n$. Let us denote the d_S-geodesic representative used for g_1 by $a_1 \cdots a_{\|g_1\|_S}$, so that $h(g_1, i_1) = a_1 \cdots a_{i_1} s(g_1, i_1)$. We will abbreviate $h(g_l, i_l)$ by h_l, $h'(g_l, i_l)$ by h'_l, $s(g_l, i_l)$ by s_l and $t(g_l, i_l)$ by t_l.

We focus on a particular vertex on the d_S-geodesic $Ug_1^{-1} \cdot [x_0, g_1]_S$, namely

$$v := Ug_1^{-1} \cdot a_1 \cdots a_{i_1 + F_0\, L + 2} = a_1 \cdots a_{i_1} \cdot s_1 \cdot \varphi^L \cdot t_1 x_0 = h_1 \cdot \varphi^L \cdot t_1.$$

Then vx_0 is K_{Lip}-close to $h_1 \cdot \varphi^L x_0$. Since $\left(h_1\varphi^L x_0,\, h_2[x_0, \varphi^L x_0]\right)$ is $(6E_0 + 180\delta)$-aligned, Fact 2.2(1) tells us that $\left(vx_0,\, h_2\Upsilon_L\right)$ is $(6E_0 + 200\delta + K_{Lip})$-aligned.

Now, Fact 2.10 tells us that either:

(1) $(Ug_1^{-1}x_0, h_{\epsilon n}\Upsilon_L)$ is $(6E_0 + 240\delta)$-aligned, or
(2) $(h_{\epsilon n - 1}\Upsilon_L, Ug_1^{-1}x_0)$ is $(6E_0 + 240\delta)$-aligned.

In Case (1), we conclude that

$$\left(x_0,\, g_1U^{-1}h_{\epsilon n}\Upsilon_L,\, g_1U^{-1}h_{\epsilon n+1}\Upsilon_L,\, \ldots,\, g_1U^{-1}h_m\Upsilon_L,\, g_1x_0\right)$$

is $(6E_0 + 240\delta)$-aligned. This contradicts the fact that $g_1 \notin \mathcal{V}_{L,\epsilon}(n)$.

In the latter case, we have:

$$(vx_0, h_1\Upsilon_L, \ldots, h_{\epsilon n-1}\Upsilon_L, y_i)$$

is $(6E_0 + 240\delta + K_{Lip})$-aligned for $y_1 = Ug_1^{-1}x_0$ and $y_2 = Ux_0$. Here, the alignment of $(h_{\epsilon n-1}\Upsilon_L, y_2)$ is due to Fact 2.12. Let P_1 be the first half of the geodesic $Ug_1^{-1}[x_0, g_1x_0]$ connecting $Ug_1^{-1}x_0$ to vx_0, and let P_2 be the latter half connecting vx_0 to Ux_0. Then $Len(P_1) + Len(P_2) \leq \|g_1\|_S \leq n$.

Recall that $R_0 = R_0(L, 4/\epsilon)$ is chosen as in Proposition 3.3 and that $L \geq L_1$ is longer than $L_0(K)$ for $K = 6E_0 + 240\delta + K_{Lip}$. Since $Len(P_1) + Len(P_2) \leq n \leq (4/\epsilon) \cdot (\epsilon n/4)$, the paths should satisfy the first alternative in Proposition 3.3 for $k = \epsilon n/4$. In particular, there exists $i \in \{0.5\epsilon n, \ldots, 0.75\epsilon n\}$ such that $d_S(P_1, h_i), d_S(P_2, h_i) \leq R_0$. Let $u_1 \in P_1$ and $u_2 \in P_2$ be the vertices realizing the distance.

Meanwhile, note that $(vx_0, h_2\Upsilon_L, \ldots, h_{i-1}\Upsilon_L, h_ix_0)$ is $(6E_0 + 200\delta + K_{Lip})$-aligned. Fact 2.13 implies that there exist $i - 2 \geq 0.25\epsilon n$ disjoint subsegments of $[vx_0, h_ix_0]$, each longer than $\tau\, L - 2(6E_0 + 200\delta + K_{Lip}) - 160\delta \geq 0.5\tau\, L$. This implies that

$$d_S(h_i, v) \geq \frac{1}{K_{Lip}} d_X(vx_0, h_ix_0) \geq \frac{1}{K_{Lip}} \cdot 0.5\tau L \cdot 0.35\epsilon n.$$

This implies that

$$d_S(u_1, v) \geq d_S(h_i, v) - d_S(h_i, u_1) \geq \frac{\tau\, L\epsilon n}{8K_{Lip}} - R_0 \geq 3R_0.$$

Meanwhile, u_1, v and u_2 are aligned along a d_S-geodesic $Ug_1^{-1}[x_0, g_1 x_0]$. This leads to a contradiction

$$2R_0 \geq d_S(u_1, h_i) + d_S(u_2, h_i) \geq d_S(u_1, u_2) \geq d_S(u_1, v) \geq 3R_0$$

In conclusion, $m \leq 2\epsilon n$ holds for m in Claim 4.4 when $n \geq 32R_0 K_{Lip}/\tau\, L\epsilon$. This implies that

$$\#\mathcal{BAD}_{L,\epsilon}(n) \times 0.9n = \#\big(\operatorname{Dom} F\big) = \sum_{U \in B_S(n)} \big(\#F^{-1}(U)\big) \leq 4T\epsilon n \cdot \#B_S(n).$$

We conclude

$$\frac{\#\mathcal{BAD}_{L,\epsilon}(n)}{\#B_S(n)} \leq 5T\epsilon. \quad (\forall n \geq 32R_0 K_{Lip}/\tau\, L\epsilon)$$

□

Let us now define

$$\mathcal{W}_{L,\epsilon}(n) := \left\{ g \in B_S(n) : \begin{array}{c} \exists h_1, \ldots, h_{\epsilon n} \in G \text{ such that the sequences } (x_0, h_1\Upsilon_L, \ldots, h_{\epsilon n}\Upsilon_L, gx_0), \\ (g^{-1}x_0, h_1\Upsilon_L) \text{ and } (h_{\epsilon n}\Upsilon_L, g^2 x_0) \text{ are each } (6E_0 + 360\delta)\text{-aligned} \end{array} \right\}.$$

Lemma 4.5. *Let $L > L_1$. Then the following is true for $g \in \mathcal{W}_{L,\epsilon}(n)$:*

(1) g is a loxodromic isometry on X with $\tau_X(g) \geq 0.5\tau\, L\epsilon n$.
(2) g has the WPD property and hence is Morse ([Sis16, Theorem 1]*).*
(3) There exists a conjugate ψ of φ such that for each large enough i, the projections of $g^{-i}x_0$ and $g^i x_0$ onto $[\psi^{-i}x_0, \psi^i x_0]$ are at least $\tau\, L/2$-apart, with the former one coming first.

Proof. For the first item, we claim that

$$\big(\ldots, g^{-1}\gamma_1, \ldots, g^{-1}\gamma_{\epsilon n}, \gamma_1, \ldots, \gamma_{\epsilon n}, g\gamma_1, \ldots, g\gamma_{\epsilon n}, \ldots\big) \tag{4.1}$$

is $(12E_0 + 900\delta)$-aligned, where $\gamma_i := h_i \Upsilon_L$. The only nontrivial part is the $(12E_0 + 900\delta)$-alignment of $(\gamma_{\epsilon n}, g\gamma_1)$. First, observe that $(\gamma_{\epsilon n}, gx_0)$ and $(\gamma_{\epsilon n}, g^2 x_0)$ are each $(6E_0 + 360\delta)$-aligned. By Corollary 2.3, $(\gamma_{\epsilon n}, z)$ is $(6E_0 + 420\delta)$-aligned for each $z \in [gx_0, g^2 x_0]$.

Meanwhile, Fact 2.13 implies that $g\gamma_1$ is contained in the $(6E_0 + 440\delta)$-neighborhood of $[x_0, gx_0]$. Fact 2.2(1) implies that $\pi_{\gamma_{\epsilon n}}(g\gamma_1)$ is contained in the $(12E_0 + 900\delta)$-long ending subsegment of $\gamma_{\epsilon n}$.

By a symmetric argument, we can similarly observe that $\pi_{g\gamma_1}(\gamma_{\epsilon n})$ is contained in the $(12E_0 + 900\delta)$-long ending subsegment of $g\gamma_1$. This concludes the desired alignment.

Now Fact 2.13 applies to the $(12E_0 + 900\delta)$-aligned sequence $(x_0, \gamma_1, \ldots, \gamma_{\epsilon n}, g\gamma_1, \ldots, g\gamma_{\epsilon n}, \ldots, g^k x_0)$ and concludes that

$$\begin{aligned} d_X(x_0, g^k x_0) &\geq \sum_{i=0}^{k-1} \sum_{j=1}^{\epsilon n} \big(\operatorname{diam}_X(g^i \gamma_j) - (2(12E_0 + 900\delta) + 160\delta)\big) \\ &\geq \epsilon nk \cdot \big(\tau\, L - (2(12E_0 + 900\delta) + 160\delta)\big) \geq \epsilon nk \cdot \frac{1}{2}\tau\, L\,. \end{aligned}$$

This implies that $\tau_X(g) \geq 0.5\tau\, L\epsilon n$.

In order to discuss WPD property, let $K > 0$. Because g is loxodromic, there exists N such that $d_X(g^{\pm N}x_0, h_1\Upsilon_L) \geq K + 1000\delta$. We then claim that $Stab_K(x_0, g^{2N}x_0)$ is finite. Suppose to the contrary that $Stab_K(x_0, g^{2N}x_0)$ is not contained in any finite d_S-metric ball. Then we can take infinitely many distinct elements $g_1, g_2, \ldots \in Stab_K(x_0, g^{2N}x_0)$.

Combining the alignment of the sequence in Display 4.1 and Fact 2.12, we observe that

$$(x_0, g^N h_1 \Upsilon_L, g^{2N} x_0)$$

is $(12E_0+960\delta)$-aligned. Since $d_X(x_0, g_i x_0) \leq K \leq d_X(x_0, g^N h_1 \Upsilon_L) - 1000\delta$, the contraposition of Fact 2.2(2) tells us that $\pi_{g^N h_1 \Upsilon_L}(\{x_0, g_i x_0\})$ has diameter at most 20δ. Similarly, $\pi_{g^N h_1 \Upsilon_L}(\{g_0^{2N}, g_i g^{2N} x_0\})$ is also 20δ-small. Hence, $(g_i x_0, g^N h_1 \Upsilon_L, g_i g^{2N} x_0)$ is also $(12E_0+980\delta)$-aligned. Hence, $[g_i x_0, g_i g^{2N} x_0]$ contains a subsegment η_i that is $(12E_0 + 1060\delta)$-fellow traveling with $g^N h_1 \Upsilon_L$.

Now, $g_i^{-1}\eta_i$'s are subsegments of $[x_0, g^{2N} x_0]$ that is longer than $\tau L - 2(12E_0 + 1060\delta) \geq 0.5\tau L$. Since $[x_0, g^{2N} x_0]$ is compact, by passing to subsequence, we may assume that $g_i^{-1}\eta_i$'s converge to a subsegment of $[x_0, g^{2N} x_0]$ of length at least $0.5\tau L$. Also, these subsegments are $(12E_0 + 1060\delta)$-fellow traveling with $g_i^{-1} g^N h_1 \Upsilon_L$ and $g_j^{-1} g^N h_1 \Upsilon_L$, respectively. Since $0.5\tau L > 12(12E_0 + 1060\delta) + E_0$, Fact 2.9 implies for large i, j that $\pi_{g_i^{-1} g^N h_1 \Upsilon_L}(g_j^{-1} g^N h_1 \Upsilon_L)$ is E_0-large and is orientation-matching. Now Fact 2.6(2) implies that

$$h_1^{-1} g^{-N} g_i g_j^{-1} g^N h_1 \subseteq B_S(E_0 + 2\, L\, F_0).$$

In particular, $g_i g_j^{-1}$ is uniformly bounded for every pair of g_i, g_j. This contradicts the infinitude of $Stab_K(x_0, g^{2N} x_0)$. The WPD property of g is now proven.

The third item holds for $\psi = h_1 \varphi h_1^{-1}$. Indeed, when N is sufficiently large, $[g^{-N} x_0, g^N x_0]$ and $[h_1 \varphi^{-N} h_1^{-1} x_0, h_1 \varphi^N h_1 x_0]$ both contain subsegments that are $0.01\tau L$-fellow traveling with a τL-long geodesic $h_1 \Upsilon_L$. We omit the detail. $\square$

We now claim that $\mathcal{V}_{L,3\epsilon}(n) \setminus \mathcal{W}_{L,\epsilon}(n)$ is non-generic.

Lemma 4.6. *For each $L > L_1$ and $\epsilon > 0$, there exists $\lambda > 1$ such that*

$$\lim_{n \to +\infty} \frac{\# \mathcal{V}_{L,3\epsilon}(n) \setminus \mathcal{W}_{L,\epsilon}(n)}{\# B_S(n)} \leq \lambda^{-n}$$

for all large enough n.

Proof. Before the proof, let $R_1 = R_1(L, 12/\epsilon)$ be as in Proposition 3.3. Let us first define

$$\begin{aligned} \mathcal{K}_1 &:= \bigcup_{r=\epsilon n}^{n/2} \big\{ abca^{-1} : a \in B_S(r), b \in B_S(n - 2r + 2R_1), c \in B_S(2R_1) \big\}, \\ \mathcal{K}_2 &:= \bigcup_{r, r' \geq 0, r + r' \leq (1-\epsilon)n} \big\{ acba^{-1} : a \in B_S(r + 2R_1), b \in B_S(r' + 2R_1), c \in B_S(2R_1) \big\}. \end{aligned}$$

We also define $\mathcal{K}_i^{-1} := \{g^{-1} : g \in \mathcal{K}_i\}$ for $i = 1, 2$. Then we have

$$\#\big(\mathcal{K}_1 \cup \mathcal{K}_2 \cup \mathcal{K}_1^{-1} \cup \mathcal{K}_2^{-1}\big) \lesssim 2 \cdot n^2 \lambda_S^{(1-0.5\epsilon)n}.$$

This is exponentially smaller than $\# B_S(n)$.

It remains to prove that $\mathcal{V}_{L,3\epsilon}(n) \setminus \big(\mathcal{K}_1 \cup \mathcal{K}_2 \cup \mathcal{K}_1^{-1} \cup \mathcal{K}_2^{-1}\big)$ is contained in $\mathcal{W}_{L,\epsilon}(n)$. To show this, let $g \in \mathcal{V}_{L,3\epsilon}(n) \setminus \big(\mathcal{K}_1 \cup \mathcal{K}_2 \cup \mathcal{K}_1^{-1} \cup \mathcal{K}_2^{-1}\big)$. Then there exists $h_1, \ldots, h_{3\epsilon n} \in G$ such that

$$(x_0, h_1 \Upsilon_L, \ldots, h_{3\epsilon n} \Upsilon_L, g x_0) \tag{4.2}$$

is $(6E_0 + 300\delta)$-aligned. Then $(x_0, h_i \Upsilon_L, g x_0)$ is $(6E_0 + 360\delta)$-aligned for each i by Fact 2.12.

Fact 2.10 guarantees that the following dichotomy holds: either

(1) $(g^{-1} x_0, h_{\epsilon n + 1} \Upsilon_L)$ is $(6E_0 + 360\delta)$-aligned, or
(2) $(h_{\epsilon n} \Upsilon_L, g^{-1} x_0)$ is $(6E_0 + 360\delta)$-aligned.

We claim that Case (1) holds. Suppose to the contrary that Case (2) holds. That means,

$$(x_0, h_1 \Upsilon_L, \ldots, h_{\epsilon n} \Upsilon_L, g^{\pm 1} x_0) \text{ is } (6E_0 + 360\delta)\text{-aligned.}$$

Pick a d_S-geodesic path P connecting id to g. Then $g^{-1} P$ is a path connecting g^{-1} to id.

We now apply Proposition 3.3. Since $Len(P) + Len(g^{-1}P) \le 2n \le (12/\epsilon)\cdot(\epsilon n/6)$, the first alternative in Proposition 3.3 should hold for $k = \epsilon n/6$. In particular, there exists $i \in \{\epsilon n/2, \ldots, \epsilon n\}$ such that $d_S(h_i, P), d_S(h_i, g^{-1}P) \le R_1$. Let $v \in P$ and $g^{-1}u \in g^{-1}P$ be the vertices realizing the distance. Here, as before, the alignment of the sequence in Display 4.2 implies that $[x_0, h_i x_0]$ contains i disjoint subsegments longer than $\mathrm{diam}_X(\Upsilon_L) - 2(6E_0 + 520\delta) \ge 0.5\tau L$. Hence,

$$d_S(id, h_i) \ge \frac{1}{K_{Lip}} d_X(x_0, h_i x_0) \ge \frac{1}{4K_{Lip}} \tau\, L\, \epsilon n.$$

This implies

$$\|v\|_S \ge \|h_i\|_S - d_S(h_i, v) \ge \frac{1}{4K_{Lip}} \tau\, L\, \epsilon n - R_1 \ge \epsilon n. \quad (\text{when } n \ge R_1/\epsilon)$$

Meanwhile, since $(h_i x_0, h_{i+1}\Upsilon_L, \ldots, h_{3\epsilon n}\Upsilon_L, g x_0)$ is also aligned, we have

$$d_S(h_i, g) \ge \frac{1}{K_{Lip}} d_X(h_i x_0, g x_0) \ge \frac{1}{K_{Lip}} \tau\, L\, \epsilon.$$

This implies $d_S(v, g) \ge \epsilon n$. Note that $\|g^{-1}u\|_S = \|g\|_S - \|u\|_S$ and $\|v\|_S$ differ by at most $2R_0$. $(*)$

We now divide the cases:

(1) $\epsilon n \le \|v\|_S \le \|g\|_S/2$. Recall that id, u, v, g are on the same d_S-geodesic P. This means

$$\|v^{-1}u\|_S = d_S(v, u) = \big|\|v\|_S - \|u\|_S\big| = \Big|\|v\|_S + \big(\|g^{-1}u\|_S - \|g\|_S\big)\Big|.$$

Thanks to $(*)$, we have

$$\Big|\|v\|_S + \big(\|g^{-1}u\|_S - \|g\|_S\big)\Big| \le \big|2\|v\|_S - \|g\|_S\big| + 2R_0 = \|g\|_S - 2\|v\|_S + 2R_0.$$

Finally, $g^{-1}u$ and v are $2R_0$-close so $u^{-1}g \cdot v \in B_S(2R_0)$. This implies the contradiction

$$g = v \cdot (v^{-1}u) \cdot (u^{-1}gv) \cdot v^{-1} \in \mathcal{K}_1.$$

(2) $\|g\|_S/2 \le \|v\|_S \le \|g\|_S - \epsilon n$. In this case, $(*)$ implies that

$$\|u\|_S \le \|g\|_S - \|v\|_S + 2R_0, \quad \|u^{-1}v\|_S \le \big|2\|v\|_S - \|g\|_S\big| + 2R_0 = 2\|v\|_S - \|g\|_S + 2R_0.$$

We also have $u^{-1}g \cdot v^{-1} \in B_S(2R_0)$. Note that $\|g\|_S - \|v\|_S$, $2\|v\|_S - \|g\|_S$ are positive integers whose sum is at most $\|v\|_S \le \|g\|_S - \epsilon n \le n - \epsilon n$. These facts lead to a contradiction

$$g = u \cdot (u^{-1}gv) \cdot (v^{-1}u) \cdot u^{-1} \in \mathcal{K}_2.$$

We can thus conclude that Case (1) holds. Meanwhile, Fact 2.10 asserts that either

(a) $(h_{2\epsilon n}\Upsilon_L, g^2 x_0)$ is $(6E_0 + 360\delta)$-aligned, or
(b) $(g^2 x_0, h_{2\epsilon n+1}\Upsilon_L)$ is $(6E_0 + 360\delta)$-aligned.

In Case (b), we are led to the alignment that

$$(g^{\pm 1}x_0, g^{-1}h_{2\epsilon n+1}\Upsilon_L, \ldots, g^{-1}h_{3\epsilon n}\Upsilon_L, x_0) \text{ is } (6E_0 + 360\delta)\text{-aligned}.$$

A similar argument as before implies $g \in \mathcal{K}_1^{-1} \cup \mathcal{K}_2^{-1}$, a contradiction. Hence, Case (a) must hold.

In conclusion, the following sequence is $(6E_0 + 360\delta)$-aligned:

$$(g^{-1}x_0, h_{\epsilon n+1}\Upsilon_L, \ldots, h_{2\epsilon n}\Upsilon_L, g^2 x_0)$$

Also, $(x_0, h_{\epsilon n+1}\Upsilon_L)$ and $(h_{2\epsilon n}\Upsilon_L, g x_0)$ are $(6E_0 + 360\delta)$-aligned. Hence $g \in \mathcal{W}_{L,\epsilon}(n)$. □

We can now finish the proof of Theorem 1.2.

Proof of Theorem 1.2. We again start by fixing the constants E_0, τ, K_{Lip}, F_0, L_1. Take $L \geq L_1$ large enough such that $\tau L \geq M$.

By Lemma 4.5, it suffices to show that for each $\eta > 0$ there exists $\epsilon > 0$ such that

$$\limsup_{n\to+\infty} \frac{\#\big(B_S(n) \setminus \mathcal{W}_{L,\epsilon}(n)\big)}{\#B_S(n)} \leq \eta. \tag{4.3}$$

To this end, we take

$$\epsilon := \frac{1}{30(2E_0 + 4\,L\,F_0 + 5)(\#S)^{E_0+3\,L\,F_0+4}} \cdot \eta.$$

Then by Fact 4.1, Lemma 4.3 and Lemma 4.6,

$$\begin{aligned}
\lim_{n\to+\infty} \frac{\#B_S(0.9n)}{\#B_S(n)} &= \lim_{n\to+\infty} \frac{\#\big(\mathcal{V}_{L,3\epsilon}(n) \setminus \mathcal{W}_{L,\epsilon}(n)\big)}{\#B_S(n)} = 0,\\
\limsup_{n\to+\infty} \frac{\#\mathcal{BAD}_{L,3\epsilon}(n)}{\#B_S(n)} &< \eta/2.
\end{aligned}$$

Moreover, we have

$$\begin{aligned}
B_S(n) \setminus \mathcal{W}_{L,\epsilon}(n) &\subseteq B_S(0.9n) \cup \Big(\big(B_S(n) \setminus B_S(0.9n)\big) \setminus \mathcal{W}_{L,\epsilon}(n)\Big)\\
&\subseteq B_S(0.9n) \cup \Big(B_S(n) \setminus \big(B_S(0.9n) \cup \mathcal{V}_{L,3\epsilon}(n)\big)\Big) \cup \big(\mathcal{V}_{L,3\epsilon}(n) \setminus \mathcal{W}_{L,\epsilon}(n)\big)\\
&= B_S(0.9n) \cup \mathcal{BAD}_{L,3\epsilon}(n) \cup \big(\mathcal{V}_{L,3\epsilon}(n) \setminus \mathcal{W}_{L,\epsilon}(n)\big).
\end{aligned}$$

Hence, Equation 4.3 holds. □

Proof of Theorem 1.2. We only list additional observations needed for Theorem 1.2. For detailed explanations about the notion of principal/triangular/ageometric fully irreducible outer automorphism in $\mathrm{Out}(F_n)$, refer to [AKKP19] and [KMPT22].

By [AKKP19, Example 6.1], there exists a principal fully irreducible $\varphi \in \mathrm{Out}(F_N)$. Now [KMPT22, Remark 5.4] provides a *lone axis* γ for φ, which is necessarily a periodic greedy folding line. Further, every fully irreducibe $g \in \mathrm{Out}(F_N)$ has a simple (periodic) folding axis.

Pick a basepoint $x_0 \in \mathcal{FF}_N$. For now, let us denote the projection map from the Outer space CV_N to $\mathcal{FF}$ by Π. Then [KMPT22, Proposition 8.1] guarantees that:

Fact 4.7. *There exists $M_0 > 0$ such that the following holds. If g is a fully irreducible and if the $d_{\mathcal{FF}}$-nearest point projections of $g^{-i}x_0$ and $g^i x_0$ onto $[\varphi^{-i}x_0, \varphi^i x_0]_{\mathcal{FF}}$ is at least M_0-apart, the first projection coming first, then g is ageometric and triangular.*

The original [KMPT22, Proposition 8.1] is formulated in terms of Pr_γ, but this can be replaced with the $d_{\mathcal{FF}}$-nearest point projection onto $\Pi(\gamma)$ by [DT17, Lemma 4.2]. Furthermore, $[\varphi^{-i}x_0, \varphi^i x_0]_{\mathcal{FF}}$ uniformly fellow travels with subsegments $\Pi(\gamma_i)$ of $\Pi(\gamma)$, where γ_i exhausts γ as i tends to infinity. This justifies the reformulation.

Given Fact 4.7, we take $M > M_0$ and run the proof of Theorem 1.1: for each $\eta > 0$ there exists $\epsilon > 0$ such that $\mathcal{W}_{L,\epsilon}$ has asymptotic density $\geq 1 - \eta$. For this ϵ, elements of $\mathcal{W}_{L,\epsilon}(n)$ for large enough n satisfy the assumption of Fact 4.7 by Lemma 4.5. Hence, $\mathcal{W}_{L,\epsilon}(n)$ consists of ageometric triangular fully irreducibles for large enough n. By shrinking η, we conclude Theorem 1.2. □

References

[ACT15] Goulnara N. Arzhantseva, Christopher H. Cashen, and Jing Tao. Growth tight actions. *Pacific J. Math.*, 278(1):1–49, 2015.

[AKKP19] Yael Algom-Kfir, Ilya Kapovich, and Catherine Pfaff. Stable strata of geodesics in outer space. *Int. Math. Res. Not. IMRN*, (14):4549–4578, 2019.

[BCK21] Hyungryul Baik, Inhyeok Choi, and Dongryul Kim. Linear growth of translation lengths of random isometries on Gromov hyperbolic spaces and Teichmüller spaces. *arXiv preprint arXiv:2103.13616*, 2021.

[BD14] Jason Behrstock and Cornelia Druţu. Divergence, thick groups, and short conjugators. *Illinois J. Math.*, 58(4):939–980, 2014.

[Beh06] Jason A. Behrstock. Asymptotic geometry of the mapping class group and Teichmüller space. *Geom. Topol.*, 10:1523–1578, 2006.

[BF02] Mladen Bestvina and Koji Fujiwara. Bounded cohomology of subgroups of mapping class groups. *Geom. Topol.*, 6:69–89, 2002.

[BF09] Mladen Bestvina and Koji Fujiwara. A characterization of higher rank symmetric spaces via bounded cohomology. *Geom. Funct. Anal.*, 19(1):11–40, 2009.

[BF10] Mladen Bestvina and Mark Feighn. A hyperbolic $\mathrm{Out}(F_n)$-complex. *Groups Geom. Dyn.*, 4(1):31–58, 2010.

[BF14] Mladen Bestvina and Mark Feighn. Hyperbolicity of the complex of free factors. *Adv. Math.*, 256:104–155, 2014.

[CDP90] M. Coornaert, T. Delzant, and A. Papadopoulos. *Géométrie et théorie des groupes*, volume 1441 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1990. Les groupes hyperboliques de Gromov. [Gromov hyperbolic groups], With an English summary.

[CF10] Pierre-Emmanuel Caprace and Koji Fujiwara. Rank-one isometries of buildings and quasi-morphisms of Kac-Moody groups. *Geom. Funct. Anal.*, 19(5):1296–1319, 2010.

[Cho24a] Inhyeok Choi. Counting pseudo-anosovs as weakly contracting isometries. *arXiv preprint arXiv:2408.00603*, 2024.

[Cho24b] Inhyeok Choi. Pseudo-Anosovs are exponentially generic in mapping class groups. *Geom. Top.*, 28:1923–1955, 2024.

[CM15] Danny Calegari and Joseph Maher. Statistics and compression of scl. *Ergodic Theory Dynam. Systems*, 35(1):64–110, 2015.

[Cou22] Rémi Coulon. Patterson-Sullivan theory for groups with a strongly contracting element. *arXiv preprint arXiv:2206.07361*, 2022.

[CW18] María Cumplido and Bert Wiest. A positive proportion of elements of mapping class groups is pseudo-Anosov. *Bull. Lond. Math. Soc.*, 50(3):390–394, 2018.

[Dan] Pallavi Dani. Genericity of infinite-order elements in hyperbolic groups. *https://www.math.lsu.edu/ pdani/research/hyp.pdf*.

[DGO17] F. Dahmani, V. Guirardel, and D. Osin. Hyperbolically embedded subgroups and rotating families in groups acting on hyperbolic spaces. *Mem. Amer. Math. Soc.*, 245(1156):v+152, 2017.

[DMGZ25] Lihuang Ding, Dídac Martínez-Granado, and Abdul Zalloum. Growth tightness and genericity for word metrics from injective spaces. *arXiv preprint arXiv:2407.02378*, 2025.

[DR09] Moon Duchin and Kasra Rafi. Divergence of geodesics in Teichmüller space and the mapping class group. *Geom. Funct. Anal.*, 19(3):722–742, 2009.

[DT17] Spencer Dowdall and Samuel J. Taylor. The co-surface graph and the geometry of hyperbolic free group extensions. *J. Topol.*, 10(2):447–482, 2017.

[GdlH90] Étienne Ghys and Pierre de la Harpe. Le bord d'un espace hyperbolique. In *Sur les groupes hyperboliques d'après Mikhael Gromov (Bern, 1988)*, volume 83 of *Progr. Math.*, pages 117–134. Birkhäuser Boston, Boston, MA, 1990.

[Gro87] M. Gromov. Hyperbolic groups. In *Essays in group theory*, volume 8 of *Math. Sci. Res. Inst. Publ.*, pages 75–263. Springer, New York, 1987.

[GTT18] Ilya Gekhtman, Samuel J. Taylor, and Giulio Tiozzo. Counting loxodromics for hyperbolic actions. *J. Topol.*, 11(2):379–419, 2018.

[Ham08] Ursula Hamenstädt. Bounded cohomology and isometry groups of hyperbolic spaces. *J. Eur. Math. Soc. (JEMS)*, 10(2):315–349, 2008.

[Kai94] Vadim A. Kaimanovich. The Poisson boundary of hyperbolic groups. *C. R. Acad. Sci. Paris Sér. I Math.*, 318(1):59–64, 1994.

[Kai00] Vadim A. Kaimanovich. The Poisson formula for groups with hyperbolic properties. *Ann. of Math. (2)*, 152(3):659–692, 2000.

[KKS07] Vadim Kaimanovich, Ilya Kapovich, and Paul Schupp. The subadditive ergodic theorem and generic stretching factors for free group automorphisms. *Israel J. Math.*, 157:1–46, 2007.

[KMPT22] Ilya Kapovich, Joseph Maher, Catherine Pfaff, and Samuel J. Taylor. Random outer automorphisms of free groups: attracting trees and their singularity structures. *Trans. Amer. Math. Soc.*, 375(1):525–557, 2022.

[KP15] Ilya Kapovich and Catherine Pfaff. A train track directed random walk on $\mathrm{Out}(F_r)$. *Internat. J. Algebra Comput.*, 25(5):745–798, 2015.

[Mah11] Joseph Maher. Random walks on the mapping class group. *Duke Math. J.*, 156(3):429–468, 2011.

[Man14] Brian Mann. *Some hyperbolic Out(FN)-graphs and nonunique ergodicity of very small FN-trees.* ProQuest LLC, Ann Arbor, MI, 2014. Thesis (Ph.D.)–The University of Utah.

[MM00] H. A. Masur and Y. N. Minsky. Geometry of the complex of curves. II. Hierarchical structure. *Geom. Funct. Anal.*, 10(4):902–974, 2000.

[MO15] Ashot Minasyan and Denis Osin. Acylindrical hyperbolicity of groups acting on trees. *Math. Ann.*, 362(3-4):1055–1105, 2015.

[MS20] Pierre Mathieu and Alessandro Sisto. Deviation inequalities for random walks. *Duke Math. J.*, 169(5):961–1036, 2020.

[MT18] Joseph Maher and Giulio Tiozzo. Random walks on weakly hyperbolic groups. *J. Reine Angew. Math.*, 742:187–239, 2018.

[Osi16] D. Osin. Acylindrically hyperbolic groups. *Trans. Amer. Math. Soc.*, 368(2):851–888, 2016.

[PZS24] Harry Petyt, Abdul Zalloum, and Davide Spriano. Constructing metric spaces from systems of walls, 2024.

[Riv08] Igor Rivin. Walks on groups, counting reducible matrices, polynomials, and surface and free group automorphisms. *Duke Math. J.*, 142(2):353–379, 2008.

[Sis16] Alessandro Sisto. Quasi-convexity of hyperbolically embedded subgroups. *Math. Z.*, 283(3-4):649–658, 2016.

[Sis18] Alessandro Sisto. Contracting elements and random walks. *J. Reine Angew. Math.*, 742:79–114, 2018.

[SZ23] Alessandro Sisto and Abdul Zalloum. Morse subsets of injective spaces are strongly contracting, 2023.

[Wie17] Bert Wiest. On the genericity of loxodromic actions. *Israel J. Math.*, 220(2):559–582, 2017.

[Yan19] Wen-yuan Yang. Statistically convex-cocompact actions of groups with contracting elements. *Int. Math. Res. Not. IMRN*, (23):7259–7323, 2019.

[Yan20] Wen-yuan Yang. Genericity of contracting elements in groups. *Math. Ann.*, 376(3-4):823–861, 2020.

Cornell University, 310 Malott Hall, Ithaca, NY, USA,
June E Huh Center for Mathematical Challenges, KIAS, 85 Hoegi-ro, Dongdaemun-gu, Seoul 02455, South Korea
Email address: `inhyeokchoi48@gmail.com`