PERCOLATION IN ACYLINDRICALLY HYPERBOLIC GROUPS

INHYEOK CHOI AND DONGGYUN SEO

Abstract. Let G be an acylindrically hyperbolic group. We prove that Bernoulli bond percolation on every Cayley graph of G has a nonuniqueness phase, in which there are infinitely many infinite clusters. This generalizes Hutchcroft's result for Gromov hyperbolic graphs to relatively hyperbolic groups, mapping class groups and rank-1 CAT(0) groups for example.
Keywords. relative hyperbolicity, CAT(0) cube complex, percolation
MSC classes: ???

1. Introduction

In geometric group theory, groups are often studied as a geometric object. Some groups resemble $\mathbb{Z}^d$ while some others resemble free groups and surface groups. In between them there is a varying degree of hyperbolicity. Here, the notion of hyperbolicity can either be phrased internally using the Dehn function, small cancellation or the prevalence of Morse elements, or by using the group action on hyperbolic spaces, e.g., word hyperbolicity, relative hyperbolicity, hierarchical hyperbolicity and acylindrical hyperbolicity. For an overview in this aspect, we refer to M. Bestvina's survey [Bes23]. There have been efforts to study these hyperbolicity of groups by means of stochastic processes such as random walks and Markov chains ([Sis18], [MT18], [MS20], [GS21]).

On the other hand, groups naturally arise in probability theory as sources of many homogeneous graphs with vertex-transitive automorphism group. In this paper, we study percolation in groups. It was classically studied for Euclidean lattices $\mathbb{Z}^d$ in relation to physical situations where liquid passes through a porous medium. I. Benjamini and O. Schramm considered its generalization to Cayley graphs of groups and sketched the general landscape of the expected phenomena [BS96]. See also papers by I. Benjamini, R. Lyons, Y. Peres and O. Schramm ([BLPS99b], [BLPS99a]).

Given a connected, locally finite (simplicial) graph Γ, *Bernoulli bond percolation* on Γ is defined by endowing independent Bernoulli random variables with expectation p to the edges. Edges whose Bernoulli RV takes value 0 are deleted, and those with Bernoulli RV taking value 1 are retained. We can then ask how the connected components, i.e., clusters, of the resulting

Date: August 5, 2025.

random subgraph $\Gamma[p]$ are shaped. To this end, we define two parameters, called the *critical parmeter*

$$p_c = p_c(\Gamma) := \inf \{p \in [0,1] : \Gamma[p] \text{ contains an infinite cluster almost surely}\}.$$

and the *uniqueness threshold*

$$p_u = p_u(\Gamma) := \inf \{p \in [0,1] : \Gamma[p] \text{ contains a unique infinite cluster almost surely}\}.$$

See Subsection 2.1 for further basics of the percolation theory.

Benjamini and Schramm posed several conjectures regarding percolation in the Cayley graphs of groups beyond $\mathbb{Z}^d$. Among them is the following:

Conjecture 1.1 ([BS96, Conjecture 6]). *A connected, locally finite, quasi-transitive graph Γ is non-amenable if and only if $p_c(\Gamma) < p_u(\Gamma)$.*

See [HJ06] for an overview of this conjecture.

Let us list some facts for the context. Given a connected graph Γ, we have $0 \le p_c \le p_u \le 1$ by definition. It is a fact that for each $p > p_c$, the random graph $\Gamma[p]$ almost surely has an infinite cluster. When Γ is quasi-transitive in addition, for each $p > p_u$ the random graph $\Gamma[p]$ almost surely has a unique infinite cluster. This is due to O. Häggström and Y. Peres [HP99] for unimodular cases, and due to R. H. Schonmann [Sch99] in general (see also [HPS99]). Lastly, C. M. Newman and L. S. Schulman proved in [NS81] for quasi-transitive Γ that, for each $p \in (0,1)$ there exists $N_\infty(p) \in \{0, 1, +\infty\}$ such that the number of infinite clusters in $\Gamma[p]$ is almost surely $N_\infty(p)$.

Hence for quasi-transitive graphs, $N_\infty(p) = 0$ almost surely for $p < p_c$, $N_\infty(p) = +\infty$ almost surely for $p_c < p < p_u$ and $N_\infty(p) = 1$ almost surely for $p > p_u$. In particular, if $p_c < p_u$ then there exists (uncountably many) p such that $\Gamma[p]$ has infinitely many infinite clusters almost surely.

Now, for *non-amenable* quasi-transitive graphs, it is known that $N_\infty(p_c) = 0$ almost surely. This is due to I. Benjmaini, R. Lyons, Y. Peres and O. Schramm [BLPS99b, Theorem 1.1] and is generalized by T. Hutchcroft to graphs with exponential growth [Hut16, Theorem 1.2]. Hence, for a non-amenable quasi-transitive graph, $p_c < p_u$ if and only if $N_\infty(p) = +\infty$ for some (countably many) p's.

Let us go back to the conjecture. The equality $p_c(\mathbb{Z}^d) = p_u(\mathbb{Z}^d)$ was observed by M. Aizenman, H. Kesten and C. M. Newman [AKN87], and by R. M. Burton and M. Keane [BK89]. A. Gandolfi, M. S. Keane and C. M. Newman observed in [GKN92] that Burton and Keane's method generalizes to amenable graphs. Hence, the only nontrivial direction is the "only if" direction. A significant breakthrough was made by T. Hutchcroft, who showed the conjecture for non-amenable quasi-transitive graphs that admit an action by a non-unimodular group [Hut20b]. We note that the first example of a quasi-transitive graph Γ for which all the three cases—$N_\infty = 0$, $N_\infty = \infty$ and $N_\infty = 1$—take place was the direct product of trees and $\mathbb{Z}$, which admits a non-unimodular automorphism group [GN90].

Hence, the remaining case is non-amenable graphs with unimodular automorphism groups. The most natural examples in this category are Cayley graphs of non-amenable groups.

We thus focus on Cayley graphs of non-amenable groups. Let G be a group with a finite generating set S. The Cayley graph $\Gamma(G, S)$ consists of the vertex set G and the edge set $\{\overline{vw} : \exists s \in S \cup S^{-1} : v = ws\}$. The group G naturally acts as graph automorphisms and the action is vertex transitive, i.e., for each $v, w \in V(\Gamma) = G$ there exists $g \in G$ such that $gv = w$. The action is also vertex-faithful, i.e., each vertex has trivial stabilizer.

A strong evidence for Conjecture 1.1 is given by I. Pak and T. Smirnova-Nagnibeda [PSN00], who proved that every non-amenable group has a Cayley graph for which $p_c < p_u$. One can then ask if there are groups all of whose *every* Cayley graph satisfy Conjecture 1.1.

In this regard, Hutchcroft showed that the non-uniqueness phase exists for every Cayley graph of non-amenable word hyperbolic groups [Hut19]. In fact, Hutchcroft proved the result for more general quasi-transitive Gromov hyperbolic graphs, by using the Bonk-Schramm embedding of such graphs into a real hyperbolic space $\mathbb{H}^d$.

The main point of this paper is to generalize Hutchcroft's result to other non-amenable groups. Namely, we have:

Theorem A. *Let G be an acylindrically hyperbolic group and let Γ be its Cayley grpah. Then we have $p_c(\Gamma) < p_u(\Gamma)$; in particular, there exist uncountably many $p \in (0, 1)$ such that $\Gamma[p]$ has infinitely many infinite clusters.*

Acylindrically hyperbolic groups encompass word hyperbolic groups, relatively hyperbolic groups and many other groups that act on a Gromov hyperbolic space in a nontrivial way. These groups have shown to exhibit interesting dynamical and group-theoretical properties, ([BF02], [Ham08], [DGO17]), as well as probabilistic behaviour ([Sis18], [MS20], [Cho25]), that are shared with word hyperbolic groups. We list some examples of acylindrically hyperbolic groups:

- (non-elementary) relatively hyperbolic groups;
- non-elementary Kleinian groups (possibly with $\mathbb{Z}^d$ subgroups);
- free products of nontrivial groups;
- the mapping class group of a finite-type hyperbolic surface;
- the outer automorphism group $\mathrm{Out}(F_n)$ of the free group F_n;
- the automorphism group $\mathrm{Aut}(G)$ of a hyperbolic group G;
- rank-1 CAT(0) groups such as irreducible CAT(0) cubical groups,
- many Artin groups and 3-manifold groups.

We refer readers to D. Osin's survey [Osi16] for more details.

Fix a vertex $v \in G$. We denote by $\chi_p(v)$ the expected size (the number of vertices) of the cluster of v. This does not depend on the choice of v in the case of Cayley graphs, so we will drop v and write χ_p. Note that $\chi_p < +\infty$ for $p < p_c$ and $\chi_p \gtrsim (p_c - p)^{-1}$ are by M. Aizenman and C. M.

Newman [AN84, Proposition 3.1]. In the course of the proof, we show that $\chi_p \lesssim (p_c - p)^{-1}$, following Hutchcroft's criterion.

Theorem B. *Let G and Γ be as in Theorem A. Then at $p_c = p_c(\Gamma)$, we have $\Delta_{p_c}(\Gamma) < +\infty$.*

This triangle diagram is intimately related to the *mean-field critical behavior*. We rely on Hutchcroft's L^2-boundedness criterion, which also has some other implications on the mean-field critical behavior. We refer to [Hut19] and [Hut20a] for these.

1.1. **The hitchhicker's guide to the nonuniqueness.** This paper is concerned with probabilistic phenomena on geometric objects that entail hyperbolicity. Luckily, probabilistic ingredients were already given by Hutchcroft [Hut19]. Namely, the gap $p_c < p_u$ and the triangle diagram bound $\Delta_{p_c} < +\infty$ follow from Equation 2.1 and 2.2. Theorem 2.8 and 2.16 provide a way to guarantee these equations. The rest of the paper will focus on the proof that acylindrically hyperbolic groups satisfy the assumptions of Theorem 2.8 and 2.16.

Thus, for readers familiar with the theory of acylindrically hyperbolic groups, the quickest way to read this paper is as follows:

(1) read Definition 2.10 and 2.12,
(2) read Theorem 2.8 and Theorem 2.16,
(3) read Section 5 and study Proposition 5.1,
(4) read Section 6 and study Proposition 6.1, and
(5) read Section 7 and combine Proposition 7.7, 7.10 and 7.12.

Notwithstanding, we recommend the readers to read from Section 2 to Section 7. This is because:

- Section 2 contains the basics of percolation theory and overview of Hutchcroft's theory. This helps understand Theorem 2.8 and Theorem 2.16.
- Subsection 2.3 describes in detail the intuition behind our strategy.
- Section 3 provides necessary hyperbolic geometry.
- Section 4 gives a different proof of Hutchcroft's hyperbolic magic lemma without using Benjamini-Schramm's Euclidean magic lemma nor Bonk-Schramm's embedding theorem. This is not needed for general acylindrically hyperbolic group but is an important prototype containing essential ideas.

  Moreover, when restricted to *groups acting properly on a Gromov hyperbolic space*, Proposition 4.6 and 6.1 are sufficient for the main theorem. Hence, readers who are mainly interested in relatively hyperbolic groups (such as non-elementary Kleinian groups and free products of groups) may read Section 4 and Section 6 only.

This paper is mostly self-contained but hides two secret ingredients. First, Hutchcroft's approach is eventually based on the analysis of the transfer operator. We invite readers to [Hut19, Section 2] for details of operator analysis. Second, the argument for barriers (Proposition 6.1 and 7.12) can be explained by the fact that acylindrically hyperbolic groups act on a quasi-tree that enjoys the bottleneck property, which arises from Bestvina-Bromberg-Fujiwara's construction ([BBF15], [BBFS19]).

Acknowledgement. The authors thank Tom Hutchcroft and Alessandro Sisto for their comments on percolation theory and hyperbolic geometry.

2. Preliminaries I: Probability theory

We assume that readers are familiar with finitely generated groups and simplicial graphs. When a group G is given a generating set S, we define the word metric d_S as

$$\begin{aligned} d_S(g,h) &:= \inf\left\{n \geq 0 : \exists s_1, \ldots, s_n \in S \cup S^{-1}[h = g s_1 \cdots s_n]\right\}, \\ \|g\|_S &:= d_S(id, g). \end{aligned}$$

2.1. **Basics of percolation theory.**

This subsection is intended as a quick introduction to percolation theory. We refer to [Gri89] for further details. Readers who are familiar with percolation theory or want to keep it as a blackbox can skip this subsection.

Let $\Gamma = (\mathcal{V}, \mathcal{E})$ be a connected simplicial graph. We will focus on the case that $\mathcal{V}$ is countable and Γ has uniformly bounded valence. Let $0 \leq p \leq 1$. On the product space $\Omega = \{0,1\}^{\mathcal{E}}$ indexed by edges of Γ, we can endow the product of Bernoulli measures with expectation p. That means, $\mathbb{P}(\omega \in \Omega : \omega(e) = 1) = p$ for each e, and $\omega(e)$ and $\omega(e')$ are independent for $e \neq e'$. Each $\omega \in \Omega$ gives rise to a graph $\Gamma(\omega)$, which is the subgraph of Γ after removing those edges e with $\omega(e) = 0$.

When p is small, we remove many edges in probability. Hence the clusters, the connected components of $\Gamma(\omega)$, are likely to be bounded. One can imagine that the expected It is convenient to use the notation:

$$v \leftrightarrow_\omega w \;\Leftrightarrow\; \text{“}v \text{ and } w \text{ are connected in } \Gamma(\omega)\text{”}$$

for $v, w \in \mathcal{V}$. Given $v \in \mathcal{V}$ we define the cluster

$$C_\omega(v) := \{w \in \mathcal{V} : w \leftrightarrow_\omega v\}.$$

Then by Fubini's theorem, we have

$$\mathbb{E}_p \,\#C_\omega(v) = \sum_{w \in \mathcal{V}} \mathbb{P}_p(v \leftrightarrow_\omega w).$$

By convention, we will write

$$v \leftrightarrow_\omega \infty \;\Leftrightarrow \#C_\omega(v) = +\infty.$$

(Recall that we focus on locally finite graphs.)

Note that the space $\Omega = \{0,1\}^{\mathcal{E}}$ and the random graph $\Gamma(\omega) \subseteq \Gamma$ for $\omega \in \Omega$ are defined without reference to p. The parameter p affects the underlying probability measure only. To express the role of p more explicitly, we denote the random graph by $\Gamma[p]$ and the underlying measure by $\mathbb{P}_p$.

As we mentioned just before, one can ask if the expected size of clusters increase as p increases. Furthermore, one can ask if the expected size of clusters depends on the choice of the root vertex. These questions can be answered using the following tools.

The space Ω is given a natural order: for $\omega, \omega' \in \Omega$ we write $\omega \leq \omega'$ if $\omega(e) \leq \omega(e')$ for each $e \in \mathcal{E}$. We say that an event $A \subseteq \Omega$ is *increasing* if

$$\forall \omega, \omega' \in \Omega \big[[\omega \in A \wedge \omega \leq \omega'] \Rightarrow \omega' \in A \big].$$

Standard examples of increasing events include $\{\omega : v \leftrightarrow_\omega w\}$ for given $v, w \in \mathcal{V}$, or $\{\omega : v \leftrightarrow_\omega +\infty\}$ for a given $v \in \mathcal{V}$.

Fact 2.1. *Let $A \subseteq \Omega$ be an increasing event. Then $\mathbb{P}_p(A) \leq \mathbb{P}_{p'}(A)$ holds for each $0 \leq p \leq p' \leq 1$.*

We now state the Harris-FKG inequality, which was first described by T. E. Harris [Har60] and later generalized by C. M. Fortuin, P. W. Kasteleyn and J. Ginibre [FKG71]:

Proposition 2.2. *[Harris-FKG] Let $A, B \subseteq \Omega$ be increasing events. Then*

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

for each p.

The Harris-FKG inequality can be used, for example, to show that the average size of clusters does not depend on the choice of the root vertex.

The Harris-FKG inequality looks like a generalization of the strict equality for independent events. One can ask if the reverse inequality also holds in certain circumstances. The BK inequality partially explains this.

Given an increasing event A and $\omega \in A$, there can be a set $W \subseteq \{e \in \mathcal{E} : \omega(e) = 1\} \subseteq \mathcal{E}$ such that $1_W \subseteq A$ holds, i.e.,

$$\forall \omega' \in \Omega \big[\forall e \in W [\omega'(e) = 1] \Rightarrow \omega' \in A \big].$$

In this situation, we call W a *witness* for A in ω. If $W' \subseteq W$ are both witnesses for E in ω, we say that W' is a *sub-witness* of W.

For example, let $v, w \in \mathcal{V}$, let $\omega \in \Omega$ and suppose that there exists a path $(e_1, \ldots, e_n)$ in $\Gamma(\omega)$ connecting v to w. Then this path becomes a witness for $A := \{v \leftrightarrow w\}$ in ω: $\{e_1, \ldots, e_n\}$ are all given value 1 by ω, and for $\omega' \in \Omega$, $\omega'(e_i) = 1$ for each i implies $\omega' \in A$.

Now, for two increasing events $A, B \subseteq \Omega$, we define another measurable set $A \circ B \subseteq \Omega$, called the *disjoint occurrence* of A and B, as follows:

$$A \circ B := \left\{ \omega \in \Omega : \begin{array}{c} \omega \in A \cap B, \ \exists \text{witness } W \subseteq \mathcal{E} \text{ for } A \text{ in } \omega \text{ and} \\ \exists \text{witness } W' \subseteq \mathcal{E} \text{ for } B \text{ in } \omega \text{ such that } W \cap W' = \emptyset \end{array} \right\}.$$

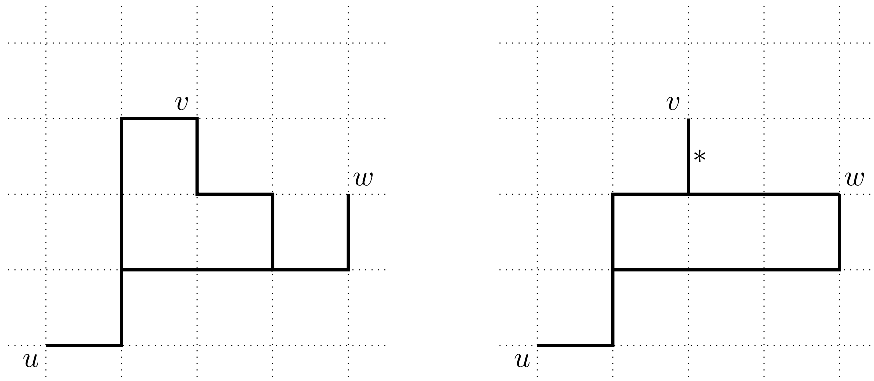

FIGURE 1. Two configurations from percolation in $\mathbb{Z}^2$. In the left configuration, the path $u \to\uparrow\uparrow\uparrow\to v$ and $v \downarrow\to\downarrow\to\uparrow w$ are disjoint witnesses for $A := \{u \leftrightarrow v\}$ and $B := \{v \leftrightarrow w\}$, respectively. In the right configuration, there are several witnesses for the events A and B in ω, but none of them are disjoint; they all contain Edge $(*)$.

For example, in Figure 1 shows two different configurations in $A \cap B$, where $A := \{u \leftrightarrow v\}$ and $B := \{v \leftrightarrow w\}$. The left configuration lies in $A \circ B$, whereas the right one does not.

We can now state the BK inequality, which is due to J. van den Berg and H. Kesten [vdBK85].

Proposition 2.3. *[BK inequality] Let $A, B \subseteq \Omega$ be increasing events for which every witness has a finite sub-witness. Then for each p we have*

$$\mathbb{P}_p(A \circ B) \leq \mathbb{P}_p(A) \cdot \mathbb{P}_p(B).$$

Often, we want a more precise information about the growth of $\mathbb{P}_p(A)$ for a given event A. For this it is beneficial to have a formula for derivatives of $\mathbb{P}_p(A)$. Russo's formula serves this purpose.

Given an increasing event $A \subseteq \Omega$ and $w \in \Omega$, we say that $e \in \mathcal{A}$ is *pivotal* for the event A if ω enters A after turning e on, and ω is excluded from A by turning e off. More formally, e is pivotal for A (given ω) if $\omega^e \in A$ and $\omega_e \notin A$ for

$$\omega^e(f) := \begin{cases} 1 & f = e \\ \omega(f) & f \in \mathcal{E} \setminus \{e\} \end{cases}, \quad \omega_e(f) := \begin{cases} 0 & f = e \\ \omega(f) & f \in \mathcal{E} \setminus \{e\} \end{cases}.$$

We now record Russo's formula [Rus81]:

Proposition 2.4. *Let $A \subseteq \Omega$ be an increasing event. Then*

$$\left(\frac{d}{dp}\right)_+ \mathbb{P}_p(A) \geq \sum_{e \in \mathcal{E}} \mathbb{P}_p(e \textit{ is pivotal for } A) = \frac{1}{1-p} \sum_{e \in \mathcal{E}} \mathbb{P}_p(e \textit{ is closed and pivotal for } A).$$

Here, $(d/dp)_+$ denotes the lower right Dini derivative.

We now focus on the percolation in
$\mathbb{P}_p$ is G-ergodic: every G-invariant event occurs with probability 0 or 1.
We now define the *critical parameter* for Γ:

$$p_c = p_c(\Gamma) := \inf\{p \in [0,1] : \mathbb{P}_p(C(id) \text{ is infinite}) > 0\}.$$

Then all clusters are almost surely finite for $p < p_c$. By ergodicity of $\mathbb{P}_p$ under the G-action, there is almost surely an infinite cluster for $p > p_c$.

We then define the *uniqueness threshold* for Γ:

$$p_u = p_u(\Gamma) := \inf\{p \in [0,1] : \mathbb{P}_p(\text{there is a unique infinite cluster}) > 0\}.$$

For every value of p, the number of infinite clusters is almost surely constant and is among $\{0, 1, \infty\}$. Furthermore, the number is infinite for $p_c < p < p_u$. It is a theorem by Häggström, Peres and Schonmann that there is almost surely a unique infinite cluster for every $p > p_u$.

2.2. **Overview of Hutchcroft's strategy.** We now explain Hutchcroft's theory in [Hut19]. Throughout, Γ will be a Cayley graph of a finitely generated group G. This graph is connected, has a uniformly bounded valency and is vertex-transitive.

Recall that for a given parameter $0 \le p \le 1$ we defined the random graph $\Gamma[p]$ by randomly deleting edges from Γ. We define the two-point function

$$\tau_p(g,h) := \mathbb{P}_p(g \leftrightarrow h) = \mathbb{P}_p(\exists \,\text{path connecting } g \text{ and } h \text{ in } \Gamma[p]).$$

We abbreviate $\tau_p(id, g)$ by $\tau_p(g)$. Then $\tau_p(g,h) = \tau_p(g^{-1}h)$ for each $g, h \in G$. We now introduce the triangle diagram

$$\Delta_p := \sup_{g \in G} \sum_{h,k \in G} \tau_p(g,h)\tau_p(h,k)\tau_p(k,g).$$

We call the expected size of the identity cluster the *susceptibility*:

$$\chi_p := \mathbb{E}_p[C(id)] = \sum_{g \in G} \tau_p(g).$$

It is a fact that $\chi_p < +\infty$ for $0 \le p < p_c$ and $\lim_{p \nearrow p_c} \chi_p = +\infty$.

Let us now define

$$\iota_p := 1 - \sup\left\{ \frac{\sum_{g,h \in K} \tau_p(g,h)}{\chi_p \cdot \# A} : A \subseteq G \text{ finite} \right\}.$$

A naive counting shows that $\iota_p \ge 0$ always holds. It is however nontrivial to show that ι_p gets closer to 1 as $p \nearrow p_c$, which is one of our main goals.

We can now state:

Theorem 2.5 ([Hut19, Proposition 2.7])**.** *Let Γ be a Cayley graph of a finitely generated group. If*

$$\liminf_{p \nearrow p_c} \frac{p_c - p}{1-p} \chi_p \sqrt{1 - \iota_p^2} = 0,$$

then $p_c(\Gamma) < p_u(\Gamma)$ and $\Delta_{p_c}(\Gamma) < +\infty$.

Hence, we can conclude $p_c(\Gamma) < p_u(\Gamma)$ once we establish

$$\limsup_{p \nearrow p_c} (p_c - p)\chi_p < +\infty, \tag{2.1}$$

$$\lim_{p \nearrow p_c} \sup \left\{ \frac{\sum_{g,h \in A} \tau_p(g,h)}{\chi_p \cdot \#A} : A \subseteq G \text{ finite} \right\} = 0. \tag{2.2}$$

In order to show Equation 2.1, Hutchcroft proved the following for Gromov hyperbolic graph that admits a vertex-transitive action by a unimodular group. We restrict ourselves to the case of Cayley graphs.

Proposition 2.6 (Supporting Hyperplane Theorem, [Hut19, Corollary 4.3])**.** *Let G be a non-elementary word hyperbolic group with a finite generating set S. Then there exists $r > 0$ such that the following holds.*

For each finite set $A \subseteq G$ there exists $A' \subseteq A$ with $\#A' \geq \#A/2$ such that for each $u \in A'$, there exists $v \in G$ with $d_S(u,v) \leq r$ such that $H_G(u,v)$ is a proper discrete halfspace with $A \subseteq H_G(u,v)$.

This is accompanied by:

Proposition 2.7. *Let G be a non-elementary word hyperbolic group with a finite generating set S. Then there exists $R > 0$ such that, for each $u, v \in G$ that gives rise to a proper discrete halfspace $H_G(u,v)$, there exists $g \in G$ with $\|g\|_S \leq 2d_S(u,v) + R$ such that $H_G(u,v)$ and $gH_G(u,v)$ are disjoint.*

The precise shape of $H_G(u,v)$ is not important. We only need that $H_G(u,v)$ is large enough to contain A, but also small enough such that some reasonably close translates of $H_G(u,v)$ do not overlap. Let us put them in a more abstract language:

Theorem 2.8 ([Hut19, Section 5.1])**.** *Let $\Gamma = Cay(G,S)$ be the Cayley graph of a finitely generated group G. Let $\mathscr{H} = \{H(g) : g \in G\}$ be a collection of subsets of G. Suppose that there exists $R > 0$ such that the following holds:*

> *For each finite set $A \subseteq G$ there exists $A' \subseteq A$ with $\#A' \geq \#A/2$ such that for each $a \in A'$, there exists $g, h \in G$ such that $\|g\|_S, \|h\|_S \leq R$, $A \subseteq aH(g)$ and $H(g) \cap hH(g) = \emptyset$.*

Then Equation 2.1 holds for Γ.

This is proven in [Hut19, Subsection 5.1] for proper discrete halfspaces in G. We present Hutchcroft's proof in Appendix A for completeness.

The proof of Equation 2.2 is more involved. Using Benjamini and Schramm's magic lemma for Euclidean spaces, Hutchcroft proved a magic lemma for real hyperbolic space $\mathbb{H}^d$:

Proposition 2.9 (hyperbolic magic lemma, [Hut19, Proposition 4.1])**.** *Let X be a closed convex set of $\mathbb{H}^d$ and let Y be a coarsely dense and uniformly locally finite subset of X. Then for every $\epsilon > 0$ there exists a constant $N(\epsilon)$ such that for every finite set $A \subset Y$ there exists a subset $A' \subseteq A$ with the following properties:*

(1) $\#A' \geq (1-\epsilon)\#A$.
(2) *For each* $v \in A'$, *there exists a pair of halfspaces* $H_1, H_2 \subseteq \mathbb{H}^d$ *such that* $d_{\mathbb{H}^2}(v, H_1 \cup H_2) \geq \epsilon^{-1}$ *and* $\#\big(A \setminus (H_1 \cup H_2)\big) \leq N(\epsilon)$.

The precise shape of halfspaces H_i in Proposition 2.9 is again not important, but we will have to impose some "smallness" of H_i in terms of ϵ. More precisely, we need that $\mathbb{E}_p \#\{g \in C(id) : gx_0 \in H_i\} \lesssim \epsilon\chi_p$. Let us introduce some terminology.

Definition 2.10. *Let* G *be a group with a finite generating set* S. *For subsets* $A, B, C \subseteq G$, *we say that* B *is a* d_S-barrier between A and C *if every* d_S*-path* $(g_0, g_1, \ldots, g_n) \subseteq G$ *starting at* A *(i.e.,* $g_0 \in A$*) and ending at* C *(i.e.,* $g_n \in C$*) intersects* B *(i.e.,* $\exists i[g_i \in B]$*).*

We record Hutchcroft's observation about barriers:

Lemma 2.11 ([Hut19, Proof of Lemma 5.4])**.** *Let* G *be a group with a finite generating set* S. *Let* $A, B \subseteq G$ *be such that* B *is a* d_S*-barrier between* id *and* A. *Then*

$$\mathbb{E}_p \#\big(C(id) \cap A\big) \leq \mathbb{E}_p \#\big(C(id) \cap B\big) \cdot \chi_p$$

for each $0 \leq p < p_c$.

Proof. For each $a \in A$ and $b \in B$ we define $E_b := \{\omega : id \leftrightarrow b\}$ and $F_{b,a} := \{\omega : b \leftrightarrow a\}$. Then

$$\begin{aligned}\sum_{a\in A, b\in B} \mathbb{P}_p(E_b \circ F_{b,a}) &\leq \sum_{a\in A, b\in B} \mathbb{P}_p(E_b)\,\mathbb{P}_p(F_{b,a}) = \sum_{b\in B}\mathbb{P}_p(E_b)\cdot\sum_{a\in A}\mathbb{P}_p(F_{b,a}) \\ &\leq \sum_{b\in B}\mathbb{P}_p(E_b)\cdot \mathbb{E}_p\#C(b) = \sum_{b\in B}\mathbb{P}_p(E_b)\cdot\chi_p = \mathbb{E}_p\#\big(C(id)\cap B\big)\cdot\chi_p.\end{aligned}$$

Meanwhile, for each $a \in A$ we claim that

$$\cup_{b\in B}(E_b \circ F_{b,a}) = \{\omega : id \leftrightarrow a\}.$$

The inclusion "$\subseteq$" is clear. Now for "$\supseteq$", let $\omega \in \Omega$ be a configuration such that $id \leftrightarrow a$. Take a shortest path P in $\Gamma(\omega)$ connecting id and a, which does not revisit a vertex twice. Since B is a d_S-barrier between id and A, P visits a vertex $b \in B$. Then the subpaths of P between id and b, and between b and a, are disjoint (finite) witnesses for E_b and $F_{b,a}$, respectively. Hence, $\omega \in E_b \circ F_{b,a}$ as desired.

In conclusion, we have

$$\begin{aligned}\mathbb{E}_p\#\big(C(id)\cap A\big) &= \sum_{a\in A}\mathbb{P}_p\{id \leftrightarrow a\} \\ &\leq \sum_{a\in A, b\in B}\mathbb{P}_p(E_b\circ F_{b,a}) \leq \mathbb{E}_p\#\big(C(id)\cap B\big)\cdot\chi_p. \quad \square\end{aligned}$$

Having Lemma 2.11 in hand, it is desirable to construct a barrier B between the origin and a halfspace whose "capacity" $\mathbb{E}_p \#(C(id) \cap B)$ is uniformly small for all $0 \leq p < p_c$. For example, B should not be the entire

G; the susceptibility $\chi_p = \mathbb{E}_p \#C(id)$ tends to infinity as $p \nearrow p_c$. Likewise, B should not contain an arbitrarily large d_S-metric ball. A geometric intuition is that if B is a *codimension 1* subset of the ambient set, then the portion of the cluster in B is finite because the cluster tends to escape B before growing large in it. The following notion captures this phenomenon.

Definition 2.12. *Let G be a group with a finite generating set S. We say that a set $B \subseteq G$ is* r-roughly branching *if there exists a subset $B' \subseteq G$ such that:*

(1) B is contained in the r-neighborhood of B' in the word metric d_S.
(2) For every $k \geq 1$, if $g_1, \ldots, g_k$ and $h_1, \ldots, h_k$ are distinct sequences of elements of B', then $g_1 \cdots g_k \neq h_1 \cdots h_k$.

Lemma 2.13 ([Hut19, Lemma 5.5])**.** *Let G be a group with a finite generating set S. Then for each $r > 0$ there exists M such that for every r-roughly branching subset $B \subseteq G$ and for every $0 \leq p \leq p_c$ we have $\mathbb{E}_p \#\big(C(id) \cap B\big) \leq M$.*

We sketch the proof for a 0-roughly branching set B; see [Hut19] for a full proof. By the Harris-FKG inequality, $\tau_p(gh) \geq \tau_p(g) \cdot \tau_p(h)$ for each $g, h \in G$. Hence, for $g_1, \ldots, g_k \in B$, we have $\tau_p(g_1 \cdots g_k) \geq \tau_p(g_1) \cdots \tau_p(g_k)$. Meanwhile, the k-th convolution map from B^k to b: $(g_1, \ldots, g_k) \mapsto g_1 \cdots g_k$ is injective by the assumption. This implies

$$\chi_p \geq \sum_{g \in B^k} \tau_p(g) = \sum_{g_1, \ldots, g_k \in B} \tau_p(g_1 \cdots g_k) \geq \prod_{i=1}^{k} \left(\sum_{g_i \in B} \tau_p(g_i) \right).$$

For a given $0 \leq p < p_c$, this is true regardless of k. Note that $\chi_p < +\infty$. This forces that $\sum_{g \in B} \tau_p(g) \leq 1$ for each $0 \leq p < p_c$. Since $p \mapsto \sum_{g \in B} \tau_p(g)$ is a lower semicontinous function on $[0, 1]$ (cf. [Hut16, Lemma 5]), the same bound holds for $p = p_c$ as well.

We finally state the "smallness" of halfspaces in $\mathbb{H}^d$ in terms of nested barriers.

Proposition 2.14 ([Hut19, Lemma 5.6])**.** *Let $X \ni x_0$ be a closed convex set of $\mathbb{H}^d$ and suppose that $G \leq \operatorname{Isom}(X)$ properly and coboundedly embeds into X by the orbit map. Let S be a finite generating set of G. Then there exist $r, R > 0$ such that for each halfspace $H \subseteq \mathbb{H}^d$, there exists an r-roughly branching subset*

$$B = B_1 \sqcup B_2 \sqcup \ldots \sqcup B_{\lfloor d(x_0, H)/R \rfloor} \subseteq G$$

such that B_i is a d_S-barrier between id and $\{g : gx_0 \in H\}$ for each $i = 1, \ldots, \lfloor d(x_0, H)/R \rfloor$.

In the above, the capacity of B is uniformly bounded in p and H; hence, there exists B_i such that $\mathbb{E}_p(\#C(id) \cap B_i) \lesssim 1/d(x_0, H)$. By Lemma 2.11 we conclude

Corollary 2.15 ([Hut19, Lemma 5.4]). *Let $X \subseteq \mathbb{H}^d$ and $G \le \mathrm{Isom}(X)$ be as in Proposition 2.14. Then there exists $K > 0$ such that for each halfspace $H \subseteq \mathbb{H}^d$ we have*

$$\mathbb{E}_p\left(\#C(id) \cap \{g \in G : gx_0 \in H\}\right) \le \frac{K}{d(x_0, H)}\chi_p$$

for each $0 \le p < p_c$.

Proposition 2.9 and Corollary 2.15 describe all we need for halfspaces. Let us now state an abstract version:

Theorem 2.16 ([Hut19, Proof of Proposition 5.2]). *Let $\Gamma = Cay(G, S)$ be the Cayley graph of a finitely generated group G. Suppose that there exists $r > 0$, and for each $D, E > 0$ there exist*

$$S_D = \sqcup_{i=1}^{\infty} S_{D;i} \subseteq G, \quad \mathcal{G}_{D,E} \subseteq G$$

and a collection $\mathscr{H}_D$ of subsets of G such that

1. *S_D is r'-roughly branching for some $r' = r'_D$,*
2. *for each $\mathcal{H} \in \mathscr{H}_D$ there exists an r-roughly branching subset $B = B_1 \sqcup \ldots \sqcup B_D \subseteq G$ such that B_i is a d_S-barrier between id and $\mathcal{H}$ for $i = 1, \ldots, D$;*
3. *for each $D, E > 0$, $\sqcup_{i \ge E} S_{D;i}$ is a d_S-barrier between id and $\mathcal{G}_{D,E}$.*

Suppose that for each $\epsilon > 0$ and $D, E > 0$, there exists a constant $N = N(\epsilon, D, E)$ such that for every finite set $A \subseteq G$ there exists $A' \subseteq A$ satisfying:

1. *$\#A' \ge (1 - \epsilon)\#A$;*
2. *For each $a \in A'$ there exist $\mathcal{H}_1, \mathcal{H}_2 \in \mathscr{H}_D$ such that*

$$\#\big(A \setminus a \cdot (\mathcal{H}_1 \cup \mathcal{H}_2 \cup \mathcal{G}_{D,E})\big) \le N.$$

Then Equation 2.2 holds for Γ.

In Hutchcroft's original formulation for Gromov hyperbolic graphs, the set S_D and $\mathcal{G}_{D,E}$ are not needed. It is not hard to adapt Hutchcroft's proof to the current version; we include it for completeness.

Proof. By Lemma 2.13, for each D we have $\sum_{g \in S_D} \tau_{p_c}(g) < +\infty$. Furthermore, note that $\sqcup_{i=1}^{E} S_{D;i}$ exhausts S_D as E increases. Hence, for each $D > 0$ and $\eta > 0$ there exists $E = E(D, \epsilon) > 0$ such that $\sum_{g \in \sqcup_{i \ge E} S_{D;i}} \tau_{p_c}(g) \le \epsilon$. Then by Lemma 2.11, we have $\sum_{g \in \mathcal{G}_{D,E}} \tau_p(g) \le \epsilon \cdot \chi_p$ for each $0 < p < p_c$.

Now, let $M = M(r)$ for r as in Lemma 2.13. Then by Assumption (2) and Lemma 2.11. we have $\sum_{g \in \mathcal{H}} \tau_p(g) \le M\chi_p/D$ for each $0 < p < p_c$ and for each $\mathcal{H} \in \mathscr{H}_D$.

Let us now fix $\epsilon > 0$. We take $D > M/\epsilon$, and then $E = E(D, \epsilon)$. Now let $N = N(\epsilon, D, E)$. Lastly, recall that $\lim_{p \nearrow p_c} \chi_p = +\infty$; there exists p_0 such that $\chi_p \ge N/\epsilon$ for $p_0 < p < p_c$.

We now claim that

$$\frac{\sum_{g,h \in A} \tau_p(g, h)}{\#A} \le 5\epsilon\chi_p \quad (\forall \text{ finite } A \subseteq G, \forall p_0 < p < p_c).$$

To observe this, let $A \subseteq G$ be a finite set and let $p_0 < p < p_c$. Let $A' \subseteq A$ be as in the proposition for ϵ, D, E. Then we have

$$\begin{aligned}\sum_{g,h\in A} \tau_p(g,h) \leq & \sum_{g\in A\setminus A', h\in G} \tau_p(g,h) + \sum_{g\in A', h\in A\setminus g\cdot(\mathcal{H}_1(g)\cup\mathcal{H}_2(g)\cup\mathcal{G}_{D,E})} \tau_p(g,h) \\ & + \sum_{g\in A', k\in \mathcal{H}_1(g)\cup\mathcal{H}_2(g)\cup\mathcal{G}_{D,E}} \tau_p(id,k) \\ & \leq \epsilon(\#A)\cdot\chi_p + (\#A')\cdot N + (\#A')\cdot\left(\frac{M}{D}\chi_p + \frac{M}{D}\chi_p + \epsilon\chi_p\right) \leq 5\epsilon(\#A)\chi_p.\end{aligned}$$

Since ϵ is arbitrary, we conclude that Equation 2.2 holds. □

Combining the aforementioned facts about word hyperbolic groups and convex subsets of $\mathbb{H}^d$, together with the Bonk-Schramm embedding theorem, Hutchcroft showed that word hyperbolic groups satisfy Equation 2.1 and 2.2 in Theorem 2.5.

2.3. **Intuition and examples.** We now explain our strategy in detail.

Our primary example will be the free group $F_2 \simeq \langle a, b\rangle$ with the generating set $S = \{a, b, c\}$. Its Cayley graph $\Gamma = Cay(F_2, S)$ is a regular 6-valent tree whose each edge is labeled with a or b. Now, if we quotient out all the edges labeled with a, then the resulting graph Γ' becomes a regular ∞-valent tree. The identity vertex id is now connected with countably infinitely many vertices $\{a^i b^{\pm 1} : i \in \mathbb{Z}\}$. One can instead consider the Cayley graph with respect to an infinite generating set $S' = S \cup \{a^i : i \in \mathbb{Z}\}$; this Cayley graph and Γ' are quasi-isometric.

At first it seems confusing to consider this ∞-valent tree instead of the original 6-valent tree. But this construction is natural for acylindrically hyperbolic groups. Acylindricallly hyperbolic groups may have non-Gromov hyperbolic Cayley graphs, but they act on a Gromov hyperbolic space that comes from this construction.

Let us first discuss the strategy for Equation 2.1.

The classical halfspaces in $\mathbb{H}^d$ or Gromov hyperbolic spaces work for Proposition 2.7. To be precise, given a δ-hyperbolic space $X \ni x_0$ and $x, y \in X$, we define

$$\mathcal{H}_{half}(x,y) := \{g \in G : d_X(gx_0, x) \leq d_X(gx_0, y)\}.$$

Then for every non-elementary isometry group $G \leq \mathrm{Isom}(X)$, there exist independent loxodromics $\{f_1, f_2, f_3\} \subseteq G$ and $R > 0$ such that, for every pair of elements $u \in G$ such that $d_X(x_0, ux_0) \geq R$, there exists $i \in \{1, 2, 3\}$ such that $\mathcal{H}_{half}(x_0, ux_0)$ and $uf_iu^{-1}\mathcal{H}_{half}(x_0, ux_0)$ are disjoint. Hence, it is straightforward to generalize Proposition 2.7 to non-elementary isometry groups of Gromov hyperbolic spaces.

Meanwhile, it is harder to generalize Proposition 2.6 in terms of $\mathcal{H}_{half}(x, y)$ for non-elementary actions on a Gromov hyperbolic space. To illustrate this, consider $G = F_2 \times \mathbb{Z}$, a group acting on the Cayley graph $\Gamma = Cay(F_2, S)$

by left multiplication of the first factor: $(a, b) \cdot x := ax$. Let $x_0 = id \in \Gamma$. Now given $D > 0$, consider a set $A \subseteq G$ whose $> 99\%$ is concentrated on $id \in \Gamma$ and the remaining $< 1\%$ covers $\{x \in \Gamma : d_S(id, x) < D\}$. In other words, we consider

$$A = \{(id, k) : 0 \le k \le 5^{D+10}\} \cup \{(g, 0) : g \in F_2, \|g\|_S \le D\}.$$

Let us give a word metric on G, say, by the generating set $S' := \{(a, 0), (b, 0), (0, 1)\}$. Then for each $g = (id, k)$ for some k, there is no $h \in G$ such that $d_{S'}(g, h) \le D$ and $\mathcal{H}_D(gx_0, hx_0)$ contains A. So $> 99\%$ of elements of A cannot satisfy the condition in Proposition 2.6.

Roughly speaking, this is because of the distortion between the geometry of G and Γ. It is possible to charge a single vertex u in Γ with arbitrarily many elements of G. For each $vx_0 \in \Gamma$ with $d_S(vx_0, ux_0) \le D$, it is also easy to make $\mathcal{H}_{half}(ux_0, vx_0)$ fail the condition in Proposition 2.6; we just charge vx_0 with one element of G. This does not cost too much, as the number of D-neighbors of ux_0 in Γ is bounded.

This pathology is remedied when we impose the so-called *weak proper discontinuity (WPD)*. Let us go back to the example F_3 acting on $\Gamma' \ni x_0 = id$. It is possible that a single vertex $id \in \Gamma'$ can be charged by many elements of G, namely, $\{a^i : i \in \mathbb{Z}\}$. But for these elements, $\{\Gamma' \setminus \mathcal{H}_{half}(a^i x_0, a^i b^D x_0) : i \in \mathbb{Z}\}$ are all *disjoint*, as x_0 has valency ∞ and the edges $\overrightarrow{a^i x_0 \, a^i b^D x_0}$ are distinct. Hence, it costs a lot to charge $\Gamma' \setminus \mathcal{H}_{half}(a^i x_0, a^i b^D x_0)$ for each i: it cannot be done with 1% of A.

Indeed, for F_2 acting on Γ', and more for generally WPD actions, Proposition 2.6 does hold. We will prove this in Section 5.

Let us now discuss Equation 2.2. In Section 4 we will prove an analogue of Proposition 2.9 for proper actions on a Gromov hyperbolic space. We sketch the idea for $F_2 = \langle a, b, \rangle$ acting on the Cayley graph $\Gamma = Cay(F_3, \{a, b\})$. Let $x_0 = id \in \Gamma$.

Suppose that $A \subseteq F_3$ is the sphere $\{g \in F_2 : \|g\|_S = R\}$. Then from the viewpoint of each $a \in A$, most elements of A are in the direction of id. It is hence sensible to pick

$$H_1(g) = \mathcal{H}_D(g, id) := \{u \in F_2 : g^{-1}u \text{ and } g^{-1} \cdot id \text{ share the initial } D\text{-long subword}\}$$

and remove it from A. Then we have $\#(A \setminus H_1(g)) \le \#\{u \in G : d_S(u, g) \le 2D\} \le (2\#S)^{2D}$ for each $a \in A$. This bound is independent of R.

Let us now consider the ball $A = \{g \in F_2 : \|g\|_S \le R\}$. The same bound holds for elements in the outmost sphere. But the bound gets worse as we go into deeper inner sphere. Nonetheless, it suffices to consider only the 10 outmost spheres $\{g : R - 10 \le \|g\|_S \le R\}$, as they account for $> 99\%$ of A. In summary, we have

$$\#(A \setminus H_1(g)) \le \#\{u \in G : d_S(u, g) \le 2(D + L)\} \le (2\#S)^{2(D+L)}$$

for $g \in \{u : R - L \le \|u\|_S \le R\}$, whose number is at least $(1 - 5^{-L})\#A$. This bound depends on the choice of D and L but not on R.

From this example we can try the following. In an arbitrary finite set $A \subseteq F_3$, for each $g \in A$ we take $H_1(g) = \mathcal{H}_D(g, id)$ and see if $A \setminus H_1(g)$ has uniformly bounded cardinality. If it does not, we regard g as an element "deep inside" and remove it. This removal is not critical as long as there are exponentially fewer "inner" elements than "outer" elements.

This strategy unfortunately does not work for an arbitrary finite set $A \subseteq F_3$. As a counterexample, consider $A = \{a^{10i} : i = 0, \ldots, R\}$, a sequence of points along a geodesic from id. Then the cardinality of $A \setminus \mathcal{H}_5(g, id)$ is bounded by N for only N many g's at the end of A, which compose a negligible portion of A. Geometrically, this subset has linear growth instead of exponential growth; the outmost spheres are negligible compared to the inner part. Indeed, for any $N > 0$ there exists R such that

$$\#\{a \in A : \exists \,\text{halfspace } H \text{ such that } d(a, H) = 5 \text{ and } \#(A \backslash H) \leq N\} \leq 0.1\#A$$

for $A = \{a^i : i = 0, \ldots, R\}$.

This is the reason we need to exclude two halfspaces for each $g \in A$ instead of one. In the example $A = \{a^{10i} : i = 0, \ldots, R\}$, a^{5R} is considered a "pre-inner" point, as $A \setminus \mathcal{H}_5(a^{5R}, id)$ contains $R/2$-many elements of A. But there is only one direct "child" of a^{5R} when viewed from id, namely, a^{5R+10}. Having only one child is not desirable for the exponential growth. Thus, we will regard a^{5R} as not genuinely inner. Then how do we cope with the largeness of $A \setminus \mathcal{H}_5(a^{5R}, id)$? We simply erase $H_2(a) := \mathcal{H}_5(a^{5R}, a^{5R+10})$. Then the number of elements of $A \setminus (\mathcal{H}_5(a^{5R}, id) \cup \mathcal{H}_5(a^{5R}, a^{5R+10})$ will be bounded.

Let us refine this strategy. We fix a bound N not depending on the size of A. Let us collect problematic points

$$\mathcal{A} := \{g \in A : \#\big(A \backslash (H_1(g) \cup H_2(g))\big) \geq N \text{ for all halfspaces } H_1, H_2 \text{ that are } D\text{-far from } id\}.$$

Then each $g \in \mathcal{A}$ is either an"outmost" element in $\mathcal{A}$ or might have some "children" in $\mathcal{A}$. In the former case, $A \setminus \mathcal{H}_D(g, id)$ is supposed to be small, and there should be only few "problematic" such elements. That means, we wish that the number of elements of $\mathcal{A}$ without "children" in $\mathcal{A}$ will be bounded.

In the latter case, if g has a lone child $h \in \mathcal{A}$, then g is considered not "deeply inner", and $A \setminus (\mathcal{H}_D(g, id) \cup \mathcal{H}_D(g, h))$ is morally small. Thus, we wish that the number of "inner but not deeply inner" elements in $\mathcal{A}$ is also bounded. If $g \in \mathcal{A}$ has more than two children in $\mathcal{A}$, then we declare that g is "deeply inner". We give up such g, but this will not be a huge loss.

After this procedure, we are left with some non-deeply-inner points $\mathcal{G} = \{g_1, g_2, \ldots\} \subseteq A$. For those elements $g \in \mathcal{G}$, $A \setminus (H_1(g) \cup H_2(g))$ has at least N elements. Now, if $A \setminus (H_1(g) \cup H_2(g))$ are disjoint for distinct $g \in \mathcal{G}$, then we can bound $\#\mathcal{G}$ in terms of $\#A$. Moreover, if deeply inner points are much fewer than not deeply inner points, then we can bound the cardinality of $\mathcal{A}$ in terms of A, This strategy indeed works for locally finite subsets of Gromov hyperbolic spaces, which is the content of Proposition 2.9. There are some concerns. What if different "inner" points share a direct child? The hyperbolicity prevents this from happening. In a Gromov hyperbolic space, every "lineage" is "linearly ordered", and no bypass is allowed. What if different non-deeply-inner points g_1, g_2 have non-disjoint $A \setminus (H_1(g_i) \cup H_2(g_i))$? Again, hyperbolicity is at play. If these sets overlap, then g_1 and g_2 are aligned when viewed from x_0. This means that one of g_1, g_2 is the descendent of the other one. With more care, it can be shown that one of g_1, g_2 is "deeply inner" and should have been removed from $\mathcal{G}$. These technical points will be studied in depth in Section 4.

Let us now talk about "smallness" of halfspaces H in terms of the capacity $\mathbb{E}_p[\#C(id) \cap H]$. In the real hyperbolic space $\mathbb{H}^d$, a halfspace H that is D-far from the origin x_0 is barred by roughly branching disjoint union of $\sim D$ barriers. This intuitively makes sense because "codimension-1" submanifolds disconnect $\mathbb{H}^d$ into two parts. There is a notion of codimension 1 subgroups for certain class of hyperbolic groups (such as cubical groups), but we will employ more general and abstract machinery.

Consider $F_3 = \langle a, b \rangle$ once again. In between id and $H = H_{100}(id, a^{100}) := \{a^{100} \cdot w,\ w \text{ does not start with } a\}$, which are spaced horizontally, we can place nine disjoint sets $B_i' := \{a^{10i} w,\ w \text{ does not start with } a^{\pm}\}$. Equivalently, B_i' is the collection of points p whose projection $\pi_{[id, a^{100}]}(p)$ onto $[id, a^{100}]$ is precisely a^{10i}. Then each of $B_1', \ldots, B_9'$ is indeed a d_S-barrier between id and H. The issue is that B_i''s are too large and are not "codimension-1". In fact, B_i''s contain an arbitrarily large d_S-metric balls, and indeed $\mathbb{E}_p[\#C(id) \cap B_i']$ is not uniformly bounded in $p \in (0, p_c)$.

We can instead consider "vertical" barriers $B_i := \{a^{10i} b^k : k \in \mathbb{Z}\}$ that are "thin" and are branching. Let us explain why B_1 is indeed a barrier. Suppose to the contrary that a d_S-path $(id = g_0, g_1, \ldots, g_N \in H)$ avoids B_1. In this path "the initial power of a appearing in g_i" grows from id to a^{100} along P. Hence, there is a moment $i(1)$ where $g_{i(1)} = a^{10} w$ for some w not starting with $a^{\pm 1}$. Since P avoids B_1, w contains some a. That means, $g_{i(1)} = a^{10} b^k a \cdot v$ in its reduced form for some v.

We claim that the letter a after $a^{10} b^k$ cannot be erased along the path, i.e., g_i starts with $a^{10} b^k a$ for every $i \geq i(1)$. If a is to be erased at some step $i, i \geq i(1)$, the only possibility is $g_i = a^{10} b^k a$ and $g_{i+1} = a^{10} b^k \in B_1$, a contradiction to the assumption. But if every g_i starts with $a^{10} b^k a$, including at $i = N$, then g_N cannot land in $H = H_{100}(id, a^{100})$. This is a contradiction.

It might look like this strategy hinges on the fact that F_2 is a (quasi-)tree and is not applicable to, say, a surface group. In fact, this strategy can be applied to WPD actions on a Gromov hyperbolic space. This is secretly related to the fact that acylindrically hyperbolic groups act on a quasi-tree thanks to Bestvina-Bromberg-Fujiwara's construction [BBF15]. We explain this in Section 6.

Proposition 2.9 is about locally finite sets of Gromov hyperbolic spaces. Thus, it can handle proper group actions on a Gromov hyperbolic space. For non-proper actions, Proposition 2.9 does not give an effective bound for element counting (as opposed to orbit counting). We hence need to exclude elements that contributes to non-properness from $A \setminus (\mathcal{H}_1 \cup \mathcal{H}_2)$. This is the reason we introduce S_D and $\mathcal{G}_{D,E}$ in Theorem 2.16.

How do we define S_D and $\mathcal{G}_{D,E}$? Recall that G contains a loxodromic isometry f of $X \ni x_0$, whose orbit $\{f^i x_0\}_{i\in\mathbb{Z}}$ is quasi-isometrically embedded in X. Hence, the powers of f are witnesses of "properness". In contrast, there can be elements $g \in G$ such that $[x_0, gx_0]$ are not fellow traveling with a translate of $[x_0, f^i x_0]$ for a long time. Such elements are manifestation of non-properness, and it is best to remove them from consideration.

With this in mind, we informally define

$$S_D := \left\{ g \in G : \begin{array}{c} [x_0, gx_0] \text{ does not fellow travel with} \\ \text{a translate of } Ax(f) \text{ for more than length } D \end{array} \right\}.$$

Then S_D becomes a roughly branching set (Proposition 7.10). This set corresponds to the collection $\mathcal{N}_D \subseteq F_2$ of words that do not have a^D as a subword. In the Cayley graph of F_2, words in $\mathcal{N}_D$ are reached from id by moving in an "almost vertical" direction. This resembles "vertical hyperplanes" in $Cay(F_2)$, and it is easy to escape these hyperplanes by adjoining a long enough horizontal step a^{2D}.

Next, we informally define

$$\mathcal{G}_{D,E} := \left\{ g \in G : \begin{array}{c} [x_0, gx_0] \text{ does not fellow travel with} \\ hAx(f) \text{ for more than length } D \\ \text{for some } h \in B_S(id, E) \end{array} \right\}.$$

This is the collection of words that do not fellow travel with $Ax(f)$ "in the beginning". In the example of F_2, this corresponds to the halfspace H that is E-far from id: in order to reach H, one has to move in the "vertical" direction for length E at first, but is allowed to move freely afterwards. Intuitively, $\mathcal{G}_{D,E}$ is barriered by $\mathcal{N}_D \cap \{g : \|g\|_S \geq E/2\}$. We formally prove this in Proposition 7.12.

3. Preliminaries II: Hyperbolic geometry

Given three real numbers A, B, C, we write $A =_C B$ if $|A - B| < C$.

A *geodesic* on a metric space (X, d) is an isometric embedding $\gamma : I \to X$ of a closed connected subset $I \subseteq \mathbb{R}$ into X. We will frequently refer to the image of γ as γ. Throughout, every metric space (X, d) is assumed to be *geodesic*, i.e., every pair of points are connected by a geodesic segment. We will however not assume that (X, d) is locally compact or complete.

3.1. Gromov hyperbolicity.

Let (X, d_X) be a geodesic metric space. Given a set $A \subseteq X$, we define its R-neighborhood

$$\mathcal{N}_R(A) := \{x \in X : \exists a \in A\, [d_X(a, x) \leq R]\}$$

for each $R > 0$. We define the Hausdorff distance between two sets

$$d_X(A, B) := \inf\{R \geq 0 : A \subseteq \mathcal{N}_R(B) \wedge B \subseteq \mathcal{N}_R(A)\}.$$

We will say that two sets $A, B \subseteq X$ are *R-equivalent* if they are within Hausdorff distance R.

For $x, y \in X$, we denote by $[x, y]$ an arbitrary geodesic between x and y. Note that such a geodesic may not be unique.

We now recall the notion of Gromov hyperbolicity due to M. Gromov [Gro87]. The version we present here is E. Rips' one. Comprehensive expositions can be found in [CDP90] and [BH99].

Definition 3.1. *Let (X, d) be a metric space. For a given $\delta > 0$, we say that (X, d) is* δ-hyperbolic *if every geodesic triangle is δ-thin, that means,*

$$\forall x, y, z \in X \left[[x, z] \subseteq \mathcal{N}_\delta([x, y]) \cup \mathcal{N}_\delta([y, z])\right].$$

We say that (X, d) is Gromov hyperbolic *if it is δ-hyperbolic for some $\delta > 0$.*

The following is immediate:

Lemma 3.2. *Let x, y, x', y' be points on a δ-hyperbolic space X such that $d_X(x, x'), d_X(y, y') < D$. Then $[x, y]$ and $[x', y']$ are $(2\delta + D)$-equivalent.*

Model examples of Gromov hyperbolic spaces are simplicial/$\mathbb{R}$-trees and real hyperbolic space $\mathbb{H}^n$. In these spaces, the following phenomenon happens: if you walk forward for some distance, and walk into another direction *without huge backtracking*, and walk into yet another direction without huge backtracking, and so on, then you will never come back to the original place. In order to formulate the property rigorously, let us define:

Definition 3.3. *Let (X, d) be a geodesic metric space. For $x, y, z \in X$, we define the* Gromov product of y and z based at x *by*

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

For example, in the standard Cayley graph of $F_2 = \langle a, b\rangle$, we have $(aaba|aab^{-1}ab)_{id} = 2$ since $aaba$ and $aab^{-1}ab$ share the first two letters.

We now formulate the local-to-global phenomenon mentioned above:

Lemma 3.4. *Let $x_0, x_1, \ldots, x_n$ be points on a δ-hyperbolic space where*

$$(x_{i-1}|x_{i+1})_{x_i} + (x_i|x_{i+2})_{x_{i+1}} \leq d_X(x_i, x_{i+1}) - 24\delta \quad (i = 1, \ldots, n-2).$$

Then there are points $y_1, \ldots, y_{n-1}$ on $[x_0, x_n]$, in the order

$$d_X(x_0, y_1) \leq d_X(x_0, y_2) \leq \ldots \leq d_X(x_0, y_{n-1}),$$

such that

$$d_X(x_i, y_i) =_{12\delta} (x_{i-1}|x_{i+1})_{x_i} \quad (i = 1, \ldots, n-1).$$

In particular, we have

$$d_X(x_0, x_n) \geq \sum_{i=1}^{n} d_X(x_{i-1}, x_i) - 2 \cdot \sum_{i=1}^{n-1} \Big((x_{i-1}|x_{i+1})_{x_i} + 12\delta\Big).$$

Another useful lemma is:

Lemma 3.5 ([Bon96, Lemma 1.3], [BH99, Prop III.H.1.17])**.** *Let x, y, z be points on a δ-hyperbolic space. Then the initial $(y|z)_x$-long subsegments of $[x, y]$ and $[x, z]$ are 4δ-fellow traveling in a synchronized manner.*

That means, if $\gamma : [0, d_X(x, y)] \to X$ represents $[x, y]$ and $\eta : [0, d_X(x, z)] \to X$ represents $[x, z]$, then $d_X(\gamma(t), \eta(t)) < 4\delta$ for $0 \le t \le (y|z)_x$.

From this property, it follows that:

Lemma 3.6 ([BH99, Prop III.H.1.22])**.** *Let x, y, z, w be points on a δ-hyperbolic space. Then we have*

$$(x|y)_w \ge \min\big((x|z)_w, (z|y)_w\big) - 4\delta.$$

In fact, Lemma 3.4 can be deduced from Lemma 3.5 and Lemma 3.6. Indeed, an induction implies that $(x_{i-1}|x_{n+1})_{x_i} =_{4\delta} (x_{i-1}|x_{i+1})_{x_i}$ for each $1 \le i < n$. Another induction implies that $(x_i|x_k)_{x_j} =_{8\delta} (x_{j-1}|x_{j+1})_{x_j}$ for each $i \le j \le k$. Yet another induction implies that $(x_i|x_n)_{x_0}$ increases in i, and the points y_i on $[x_0, x_n]$ whose distance from x_0 is $(x_i|x_n)_{x_0}$ realize the desired property.

Let us now turn to isometries. Let (X, d) be a Gromov hyperbolic space and let g be its isometry. We say that g is *loxodromic* if there exists $\tau > 0$ such that $d_X(x_0, g^n x_0) \ge \tau n$ for each n.

Prototypes of loxodromic isometries are the loxodromic isometries of $\mathbb{H}^n$. They act as a translation along an infinite geodesic. An isometry g of a Gromov hyperbolic space X is called an *axial loxodromic* if there exists $\tau > 0$ and a geodesic $\gamma : \mathbb{R} \to X$ such that $g(\gamma(t)) = \gamma(t + \tau)$ for each $t \in \mathbb{R}$. In this case, we call γ an *axis* of g and denote it by $Ax(g)$. By rescaling the metric d_X globally, it is not hard to render g *unital*, i.e., $\tau = 1$.

In general, given a group G acting on a Gromov hyperbolic space X and a loxodromic isometry $g \in G$, it is not hard to put another metric on X that is G-equivariantly quasi-isometric to the original one, so that g becomes a unital axial loxodromic isometry. See e.g. [BF02, Proposition 6.(2)].

We now introduce the nearest point projection.

Definition 3.7. *Let (X, d) be a metric space and let $A \subseteq X$ be a locally compact subset. We define the* nearest point projection $\pi_A(\cdot) : X \to 2^A$ *as*

$$\pi_A(x) := \{a \in A : d_X(x, a) = \min_{y \in A} d_X(x, y)\}.$$

For $B, C \subseteq X$, we use the notation $\operatorname{diam}_A(B) := \operatorname{diam}_X(\pi_A(B))$ *and* $d_A(B, C) := \operatorname{diam}_X(\pi_A(B \cup C))$.

Lemma 3.8. *Let x, y, z be points on a δ-hyperbolic space X. Let $p \in [x, y]$ be such that $d_X(x, p) = (y|z)_x$. Then $\pi_{[x,y]}(z)$ is contained in $\mathcal{N}_{8\delta}(p)$.*

In Gromov hyperbolic spaces, geodesics exhibit the so-called *contracting property*. The following is one formulation of the contracting property.

Lemma 3.9 ([CDP90, Proposition 10.2.1]). *Let (X, d_X) be a δ-hyperbolic space, let $x, y \in X$, let γ be a geodesic and let $p \in \pi_\gamma(x)$, $q \in \pi_\gamma(y)$. Then we have*

$$d_X(p, q) \leq \max\big(12\delta, 12\delta + d_X(x, y) - d_X(x, p) - d_X(q, y)\big).$$

This lemma has the following corollary.

Corollary 3.10. *Let (X, d_X) be a δ-hyperbolic space, let $x, y \in X$ and let γ be a geodesic in X.*

1. *(coarse Lipschitzness) We have $\mathrm{diam}(\pi_\gamma(x)) \leq 12\delta$ and $d_\gamma(x, y) \leq d_X(x, y) + 12\delta$.*
2. *(constriction) Let $p \in \pi_\gamma(x)$, $q \in \pi_\gamma(y)$. Suppose that $d_X(p, q) > 12\delta$. Then $[p, q]$ is within Hausdorff distance 12δ from some subsegment $[x', y']$ of $[x, y]$, where $d_X(x', p), d_X(y', q) < 10\delta$.*
3. *(no backtracking) Let $z \in [x, y]$. Then $\pi_\gamma(z)$ is contained in the 12δ-neighborhood of $[\pi_\gamma(x), \pi_\gamma(y)]$.*
4. *(equivalent geodesics) Let γ' be a geodesic whose endpoints are pairwise D-near with the ones of γ. Then $\pi_\gamma(x)$ and $\pi_{\gamma'}(x)$ are $(2D + 28\delta)$-equivalent.*
5. *Let γ' be a subgeodesic of γ and suppose that $d_{\gamma'}(x, y) > 12\delta$. Then $d_\gamma(x, y) \geq d_{\gamma'}(x, y) - 64\delta$.*

We include the proof in Appendix B for completeness.

We now review the notion of weak proper discontinuity introduced in [BF02].

Definition 3.11. *Let (X, d_X) be a Gromov hyperbolic space and let $G \leq \mathrm{Isom}(X)$. We say that the action of G on X is* proper *if*

$$\forall R > 0 \Big[\# \Big\{g \in G : d_X(x_0, gx_0) < R\Big\} < +\infty\Big].$$

Let $f \in G$ be a loxodromic isometry. We say that G has weakly properly discontinuous (WPD) *along f, or that f has the WPD property, if*

$$\forall R > 0\, \exists L > 0 \Big[\# \big\{g \in G : d_X(x_0, gx_0) < R \text{ and } d_X(f^L x_0, gf^L x_0) < R\big\} < +\infty\Big].$$

If f is axial in addition, we call it an axial WPD loxodromic*.*

If G has WPD action on a Gromov hyperbolic space and is not virtually cyclic, then we call it an acylindrically hyperbolic group*.*

If G acts properly on a Gromov hyperbolic space, then every loxodromic element has the WPD property automatically. We record a theorem by M. Bestvina, K. Bromberg and K. Fujiwara.

Definition 3.12. *Let G be a group and let $f \in G$. We define the* elementary closure *of f by*

$$EC(f) := \big\{g \in G : \exists N > 0 [gf^N g^{-1} = f^N] \vee [gf^N g^{-1} = f^{-N}]\big\}.$$

Theorem 3.13 ([BBF15, Theorem H])**.** *Let G be an acylindrically hyperbolic group. Then G contains an element f and admits an isometric action on a Gromov hyperbolic space X, and there exists a constant $K > 0$ such that the following holds:*

(1) f is a unital, axial WPD loxodromic isometry of X;
(2) for each $g \in G$, either
  (a) (bounded projection) the nearest point projection of $Ax(f)$ onto $gAx(f)$ has diameter $\le K$, or
  (b) $g \in EC(f)$ and $gAx(f) = Ax(f)$.

Moreover, the cyclic subgroup $\langle f \rangle$ is a finite-index subgroup of $EC(f)$.

Note that in Theorem 3.13, every finite generating set S of G contains an element $g \in G$ that falls into Case (2-a), as S generates a non-virtually cyclic group.

The following is a well-known fact about acylindrically hyperbolic groups and is a basic ingredient of the quasi-tree construction in [BBF15]. We sketch the proof for reader's convenience.

Lemma 3.14. *Let X be a δ-hyperbolic space and let $\gamma_1, \ldots, \gamma_n$ be geodesics with mutually K_0-bounded projections. Let $z \in X$ be a point such that $d_{\gamma_i}(z, \gamma_{i+1}) \ge 5K_0 + 100\delta$ for each $i = 1, \ldots, n-1$. Then $d_{\gamma_i}(z, q) \ge d_{\gamma_i}(z, \gamma_{i+1}) - (2K_0 + 112\delta)$ for each i and for each $q \in \gamma_n$.*

Before proving it, let us observe a simple fact:

Lemma 3.15. *Let X be a geodesic metric space, let $x, y, z \in X$ and let $w \in [x, y]$. Then $(w|z)_y \le (x|z)_y$ holds.*

Let $u \in \mathcal{N}_K([x, y])$. Then $(u|z)_y \le (x|z)_y + K$ holds.

Proof. The first statement follows from $d_X(x, y) = d_X(x, w) + d_X(w, y)$ and $d_X(x, z) \le d_X(x, w) + d_X(w, z)$. The second statement follows from the first one and the triangle inequality. □

Proof of Lemma 3.14. Let $p_i \in \pi_{\gamma_i}(z)$ and $q_i \in \pi_{\gamma_i}(\gamma_{i+1})$ be the ones such that $d_X(p_i, q_i) = d_{\gamma_i}(z, \gamma_{i+1})$; for $i = n$, we pick arbitrary $q_n \in \gamma_n$.

Then for $i = 1, \ldots, n-1$, p_i and q_i are $(5K_0 + 100\delta)$-far. Furthermore, $q_i \in \pi_{\gamma_i}(\gamma_{i+1})$ and $\pi_{\gamma_i}(q_{i+1})) \in \pi_{\gamma_i}(\gamma_{i+1})$ are K_0-close because of the bounded projection assumption. Corollary 3.10(2) tells us that $[z, q_{i+1}]$ passes through the 10δ-neighborhood of p_i and $(10\delta + K_0)$-neighborhood of q_i, in order. This implies that $(z|q_{i+1})_{q_i} < K_0 + 10\delta$ for $i = 1, \ldots, n-1$.

For the same reason, q_{i-1} is $(K_0 + 10\delta)$-close to $[z, q_i]$ for each $i \ge 2$. Lemma 3.15 tells us that

$$(q_{i-1}|q_{i+1})_{q_i} \le (z|q_{i+1})_{q_i} + (K_0 + 10\delta) \le 2K_0 + 20\delta.$$

Moreover, q_{i-1} is also $(K_0 + 10\delta)$-close to $[z, p_i]$ for each $i \ge 2$, and p_i is 10δ-close to $[z, q_i]$ by Corollary 3.10(2). This implies that

$$(q_{i-1}|q_i)_{p_i} \le (z|q_i)_{p_i} + (K_0 + 10\delta) \le K_0 + 20\delta.$$

Since $d_X(p_i, q_i) \geq 5K_0 + 100\delta$, we conclude $d_X(q_{i-1}, q_i) \geq 4K_0 + 80\delta$ for $i = 2, \ldots, n-1$.

We can now apply Lemma 3.4 to the points

$$(z, q_1, \ldots, q_n).$$

It follows that $(z|q_n)_{q_i} < 2K_0 + 32\delta$. Meanwhile, since $d_X(p_i, q_i) > 10K_0 + 130\delta$ and $d_X(p_i, [z, q_i]) < 10\delta$, we have $(p_i|z)_{q_i} \geq 10K_0 + 120\delta$. Lemma 3.6 tells us that $(p_i|q_n)_{q_i} \leq 2K_0 + 36\delta$. By Lemma 3.8, q_i is $(2K_0 + 44\delta)$-close to $\pi_{[p_i, q_i]}(q_n)$ and $d_{[p_i, q_i]}(z, q_n) \geq d_{\gamma_i}(z, \gamma_{i+1}) - (2K_0 + 44\delta)$. By Corollary 3.10(5), we have $d_{\gamma_i}(z, q_n) \geq d_{\gamma_i}(z, \gamma_{i+1}) - (2K_0 + 110\delta)$ as desired. □

We now record another consequence of the WPD property.

Lemma 3.16 ([Sis16, Lemma 3.3], [Cho25, 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 Gromov yperbolic space $X \ni x_0$ with a WPD loxodromic element $f \in G$. 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 $\|h\|_S \leq M$. Then $\pi_{[x_0, f^k x_0]}(\{gx_0, ghx_0\})$ has diameter at most D_0.

4. Hyperbolic magic lemma

The following is called a hyperbolic magic lemma [Hut19, Proposition 4.1]. Hutchcroft proved it under the assumption that X is the real hyperbolic space $\mathbb{H}^n$ and A lies in a quasi-convex set.

A subset $Y \subseteq X$ of a metric space is *uniformly locally finite* if

$$\sup_{y \in Y} \#\big(\mathcal{N}_R(y) \cap Y\big) < +\infty \quad (\forall R > 0).$$

The vertex set of a Cayley graph of a finitely generated group is uniformly locally finite. More generally, if a group G acts properly on a metric space $X \ni x_0$, then the G-orbit $G \cdot x_0$ is uniformly locally finite.

Given $x, y \in X$ and $D > 0$, we define

$$\mathcal{H}_D(x, y) := \{z \in X : (z|y)_x \geq D\}.$$

(Note that this set can well be empty). Sets of this sort are called *halfspaces* rooted at x with radius parameter D.

Proposition 4.1. *Let X be a δ-hyperbolic space and let Y be a uniformly locally finite subset of X. Then for each $\epsilon, D > 0$ there exists a constant $N = N(\epsilon, D, Y)$ such that for every finite set $A \subseteq Y$ there exists a subset $A' \subseteq A$ satisfying:*

(1) $\#A' \geq (1-\epsilon)\#A$;
(2) For each $a \in A'$ there exist halfspaces $\mathcal{H}_1, \mathcal{H}_2 \in X$ rooted at a with radius parameter D such that $\#\big(A \setminus (\mathcal{H}_1 \cup \mathcal{H}_2)\big) \leq N$.

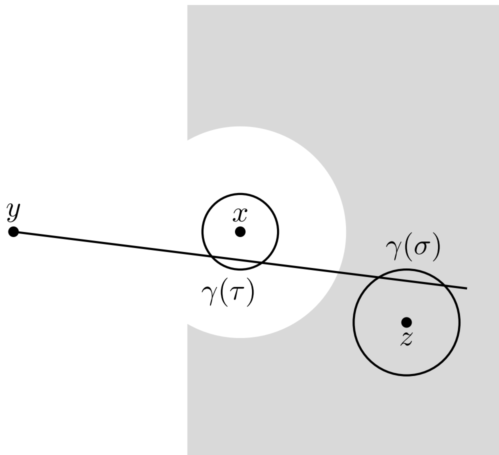

FIGURE 2. Definition of anti-halfspace $\mathfrak{A}_D(x, y)$.

Proof. For simplicity, we assume $1 \le \delta \le 0.0001D$. Because Y is locally uniformly finite, we have

$$\sup_{y \in Y} \#\big(\mathcal{N}_{100D}(y) \cap Y\big) =: M < +\infty.$$

We set $N = 2M/\epsilon$. Note that N depends on ϵ, D, Y but not on the choice of $A \subseteq Y$.

Now let $A \subseteq Y$ be a finite set. Let A' be the collection of the elements of A that satisfy the condition (2) in the statement. Our goal is to show $\#A' \ge (1 - \epsilon)\#A$.

For a technical reason we introduce a variation of the notion of halfspace. Given $x, y \in X$ and $r > 0$, let us define the *anti-halfspace*

$$\mathfrak{A}_D(x, y) := \left\{ z \in X \setminus \mathcal{N}_{6D}(x) : \begin{array}{c} \exists \text{ a geodesic } \gamma : [0, \tau] \to X \text{ and } 0 \le \tau_1 \le \tau \text{ such that} \\ \gamma(0) = y,\ d_X(\gamma(\tau_1), x) < D + 100\delta \text{ and } d_X(\gamma(\tau), z) < D + 200\delta \end{array} \right\}.$$

This is morally the complement of $\mathcal{H}_D(x, y)$ but not quite exactly.

We record two elementary observations.

Observation 4.2. *For every $x, y \in X$ we have $x \notin \mathfrak{A}_D(x, y)$. If $d_X(x, y) \ge 100D$ moreover, then $y \notin \mathfrak{A}_D(x, y)$.*

Observation 4.3. *Let $x, y \in X$ and let $z \in \mathfrak{A}_D(x, y)$. Then $d_X(y, z) \ge d_X(x, y) + D$.*

Let us now collect problematic elements, i.e.,

$$\mathcal{A}_1 := A \backslash A' = \left\{ a \in A : \begin{array}{c} \#\big(A \setminus (\mathcal{H}_1 \cup \mathcal{H}_2)\big) \ge N \text{ for every halfspaces} \\ \mathcal{H}_1, \mathcal{H}_2 \text{ rooted at } a \text{ with distance parameter } D \end{array} \right\}.$$

We now pick a maximally $100D$-separated subset $\mathcal{A}_2$ of $\mathcal{A}_1$, i.e., we have

(1) $d_X(a, a') \ge 100D$ for each pair of distinct elements $a, a' \in \mathcal{A}_2$;
(2) $\mathcal{A}_2$ is a maximal subset of $\mathcal{A}_1$ satisfying this property.

Then $\bigcup_{a \in \mathcal{A}_2}(\mathcal{N}_{100D}(a) \cap Y)$ covers entire $\mathcal{A}_1$ (if not, a missed element can be added to $\mathcal{A}_2$ and break the maximality). Since $\mathcal{N}_{100D}(a) \cap Y$ has at most M elements for each $a \in \mathcal{A}_2$, we have

$$\#\mathcal{A}_2 \geq \frac{1}{M} \cdot \#\mathcal{A}_1.$$

The proof will be done once we show that $\#\mathcal{A}_2 \leq \frac{\epsilon}{M}\#A$. For this one might wish to create disjoint complements of halfspaces rooted at each element of $\mathcal{A}_2$. However, the complement of halfspaces rooted at distinct elements of $\mathcal{A}_2$ might intersect. Our next goal is to extract some portion of $\mathcal{A}_2$ for which we can create disjoint complements of halfspaces.

Let us first prepare empty collections $\mathcal{B} = \mathcal{U} = \mathcal{G} = \emptyset$. They are meant to be collections of bad, undecided and good elements. Fix a basepoint $x_0 \in X$. Enumerate $\mathcal{A}_2$ by the distance from x_0, i.e., let $\mathcal{A}_2 = \{a_1, a_2, \ldots, a_{\#\mathcal{A}_2}\}$ be such that $d_X(x_0, a_i) \leq d_X(x_0, a_{i+1})$ for each i. At each step $i = 1, \ldots, \#\mathcal{A}_2$, we will define a point $b_i \in X$ and put a_i in either $\mathcal{B}$ or $\mathcal{G}$; this decision is final and shall not be modified further. We may put some other elements of $\mathcal{A}_2$ in $\mathcal{U}$, whose their classification will change later. We will keep the balance $\#\mathcal{B} \leq \#\mathcal{U} + \#\mathcal{G}$ throughout. After the last step there will be no element of $\mathcal{U}$, so we will have $\#\mathcal{B} \leq \#\mathcal{G}$.

We now describe the procedure. At step i, we first declare $\mathfrak{A}_i := \mathfrak{A}_D(a_i, x_0)$.

(1) If $\mathcal{A}_2 \cap \mathfrak{A}_i$ has no element, then we declare that $a_i \in \mathcal{G}$ and $b_i := x_0$.
(2) If not, pick $b_i \in \mathcal{A}_2 \cap \mathfrak{A}_i$ that is the *closest* to x_0. We then declare $\mathfrak{A}'_i := \mathfrak{A}_D(a_i, b_i)$.
   (a) If $\mathcal{A}_2 \cap \mathfrak{A}_i \cap \mathfrak{A}'_i$ has no element, then we declare that $a_i \in \mathcal{G}$.
   (b) If not, we pick $c_i \in \mathcal{A}_2 \cap \mathfrak{A}_i \cap \mathfrak{A}'_i$ that is the *closest* to x_0. We then declare $a_i \in \mathcal{B}$ and $b_i, c_i \in \mathcal{U}$.

(If an element in $\mathcal{U}$ is declared good or bad, it is not undecided anymore; we remove it from $\mathcal{U}$.)

Till step i, $\mathcal{G} \cup \mathcal{B}$ comprises of elements from $\{a_1, \ldots, a_i\}$; they do not contain any of $a_{i+1}, a_{i+2}, \ldots$. $(*)$ Let us observe what happens at step i.

In case (1), $\mathcal{G}$ gains one more element that might be from $\mathcal{U}$ or not. $\mathcal{B}$ does not change. Overall, $\#\mathcal{B}$ stays the same and $\#\mathcal{U} + \#\mathcal{G}$ does not decrease. Similar situation happens in Case (2-a).

In case (2-b), $\mathcal{B}$ gains one element a_i, which might be from $\mathcal{U}$. In exchange, $\mathcal{U}$ gains elements b_i and c_i. Observation 4.3 guarantees that $d_X(x_0, b_i), d_X(x_0, c_i) \geq d_X(x_0, a_i) + D$. Since $\mathcal{A}_2$ was labelled with respect to the distance from x_0, we conclude that $b_i, c_i \in \{a_{i+1}, a_{i+2}, \ldots\}$; in other words, neither b_i nor c_i come from $\mathcal{G} \cup \mathcal{B}$. We thus confirm that elements are never re-classified once they are put in $\mathcal{G} \cup \mathcal{B}$.

Furthermore, note that $b_i \in \mathfrak{A}_D(a_i, x_0) \cap \mathcal{A}_2$ and $c_i \in \mathfrak{A}_D(a_i, b_i) \cap \mathcal{A}_2$. By Observation 4.2, the former membership implies that $b_i \neq a_i$ and $d_X(b_i, a_i) \geq 100D$ (as A_2 is $100D$-separated), and the latter membership implies that $c_i \notin \{a_i, b_i\}$. In particular, b_i, c_i are distinct. If b_i, c_i are not from $\mathcal{U}$ at step

$i-1$ and are genuinely new additions to $\mathcal{U}$, then we can conclude that $\#\mathcal{U}$ increases at least by 1 in Case (2-b). It remains to show

Claim 4.4. *For $i < j$ such that $a_i, a_j \in \mathcal{B}$, we have $\{b_i, c_i\} \cap \{b_j, c_j\} = \emptyset$.*

Proof of Claim 4.4. Suppose first to the contrary that $b_i \in \{b_j, c_j\}$. Then by the construction of b_i, b_j and c_j, we have

$$b_i \in \mathfrak{A}_D(a_i, x_0) \cap \mathfrak{A}_D(a_j, x_0).$$

Let $\gamma : [0, \tau] \to X$ be a geodesic starting at x_0 and $0 \le \tau_1 \le \tau$ be such that

$$d_X\big(\gamma(\tau_1), a_i\big) < D + 100\delta, \quad d_X\big(\gamma(\tau), b_i\big) < D + 200\delta.$$

Let $\gamma' : [0, \sigma] \to X$ be a geodesic starting at x_0 and $0 \le \sigma_1 \le \sigma$ be such that

$$d_X\big(\gamma'(\sigma_1), a_j\big) < D + 100\delta, \quad d_X\big(\gamma'(\sigma), b_i\big) < D + 200\delta.$$

Let $L_{min} := \big(\gamma(\tau)\big|\gamma'(\sigma)\big)_{x_0}$. Lemma 3.5 tells us that $d_X(\gamma(t), \gamma'(t)) < 4\delta$ for $0 \le t \le L_{min}$. Note that

$$\tau - L_{min} = \big(x_0\big|\gamma'(\sigma)\big)_{\gamma(\tau)} \le d_X\big(\gamma(\tau), \gamma'(\sigma)\big) \le 2D + 400\delta.$$

We conclude that $L_{min} \ge \tau - 3D$. Similarly $L_{min} \ge \sigma - 3D$.

Meanwhile, recall that a_j and $b_i \in \{b_j, c_j\}$ are distinct elements of a $100D$-separated set $\mathcal{A}_2$. It follows that $d_X(a_j, b_i) > 100D$ and $\sigma - \sigma_1 \ge 97D$. In other words, we have $\sigma_1 \le \sigma - 97D \le L_{min} - 94D \le \tau - 94D$.

We now have

$$d_X\big(\gamma(\sigma_1), a_j\big) \le d_X\big(\gamma(\sigma_1), \gamma'(\sigma_1)\big) + d_X\big(\gamma'(\sigma_1), a_j\big) \le D + 120\delta. \tag{4.1}$$

This implies that

$$100D \le d_X(a_i, a_j) \le d_X\big(a_i, \gamma(\tau_1)\big) + d_X\big(\gamma(\tau_1), \gamma(\sigma_1)\big) + d_X\big(\gamma(\sigma_1), a_j\big) \le 3D + |\tau_1 - \sigma_1|,$$

i.e., $\tau_1 \ge \sigma_1 + 97D$ or $\tau_1 \le \sigma_1 - 97D$. In the former case, we have

$$\begin{aligned} d_X(x_0, a_i) &\ge \tau_1 - d_X\big(\gamma(\tau_1), a_i\big) \\ &\ge \tau_1 - (D + 100\delta) \ge \sigma_1 + 95D \\ &\ge d_X\big(x_0, \gamma'(\sigma_1)\big) + d_X\big(\gamma'(\sigma_1), a_j\big) + 90D \ge d_X(x_0, a_j) + 90D, \end{aligned}$$

contradicting the ordering of $\{a_1, a_2, \ldots\}$. Hence, the latter case holds.

We now have timing $0 \le \tau_1 \le \sigma_1 \le L_{min} \le \tau$ for the geodesic γ. Recall also Inequality 4.1. We conclude $a_j \in \mathcal{A}_2 \cap \mathfrak{A}_D(a_i, x_0)$. Moreover,

$$\begin{aligned} d_X(x_0, a_j) &\le d_X(x_0, \gamma'(\sigma_1')) + d_X(\gamma'(\sigma_1'), a_j) \\ &\le \sigma_1 + D + 100\delta \le \sigma - 95D \\ &\le d_X(x_0, \gamma'(\sigma)) - d_X(\gamma'(\sigma), b_i) - 90D \le d_X(x_0, b_i) - 90D. \end{aligned}$$

This contradicts the minimality of b_i with respect to the distance from x_0. Hence, we have $b_i \notin \{b_j, c_j\}$.

Now suppose to the contrary that $c_i \in \{b_j, c_j\}$. This implies that

$$c_i \in \mathfrak{A}_D(a_i, b_i) \cap \mathfrak{A}_D(a_i, x_0) \cap \mathfrak{A}_D(a_j, x_0).$$

We pick a geodesic $\gamma : [0, \tau] \to X$ starting at x_0 and $0 \le \tau_1 \le \tau$ such that

$$d_X(\gamma(\tau_1), a_i) < D + 1000\delta, \quad d_X(\gamma(\tau), c_i) < D + 200\delta.$$

The previous argument tells us the following: since $c_i \in \mathfrak{A}_D(a_i, x_0) \cap \mathfrak{A}_D(a_j, x_0)$, there exists $\tau_1 + 97D \le \sigma_1 \le \tau - 90D$ such that $d_X(\gamma(\sigma_1), a_j) \le D + 120\delta$. In particular, $a_j \in \mathfrak{A}_D(a_i, x_0)$. Moreover, a_j is closer than c_i to x_0.

Now consider a geodesic $\eta : [0, L'] \to X$ starting at b_i and $0 \le \tau_1' \le \tau'$ such that

$$d_X(\eta(\tau_1'), a_i) < D + 100\delta, \quad d_X(\eta(\tau'), c_i) < D + 200\delta.$$

We first consider the geodesic triangle connecting $\gamma(\tau_1), \gamma(\tau)$ and $\eta(\tau_1')$. By the δ-slimness of the triangle, $\gamma(\sigma_1) \in \gamma|_{[\tau_1, \tau]}$ is δ-close to either $[\gamma(\tau_1), \eta(\tau_1')]$ or $[\eta(\tau_1'), \gamma(\tau)]$. Meanwhile, the former one is contained in $\mathcal{N}_{3D}(\gamma(\tau_1))$, whereas $\gamma(\sigma_1)$ is at least $97D$-far from $\gamma(\tau_1)$. Hence, $\gamma(\sigma_1)$ is δ-close to some point $p \in [\eta(\tau_1'), \gamma(\tau)]$.

Next, we observe the geodesic triangle connecting $\eta(\tau_1'), \eta(\tau')$ and $\gamma(\tau)$. This time, p is δ-close to either $[\eta(\tau_1'), \eta(\tau')]$ or $[\eta(\tau'), \gamma(\tau)]$. The latter one is contained in $\mathcal{N}_{1.5D}(c_i)$ and hence in $\mathcal{N}_{3D}(\gamma(\tau))$. Meanwhile, $\gamma(\sigma_1)$ is $90D$-far from $\gamma(\tau)$, so p is $89D$-far from $\gamma(\tau)$. Hence, p cannot be δ-close to $[\eta(\tau'), \gamma(\tau)]$, and is rather δ-close to $[\eta(\tau_1'), \eta(\tau')]$.

In conclusion, $\gamma(\sigma_1)$ is 2δ-close to some point $q \in \eta|_{[\tau_1', \tau']}$. This q is $(D + 122\delta)$-close to a_j. It follows that $a_j \in \mathfrak{A}_D(a_i, b_i)$.

In conclusion, $a_j \in \mathfrak{A}_D(a_i, b_i) \cap \mathfrak{A}_D(a_i, x_0)$ and is closer than c_i to x_0. This contradicts the minimality of c_i. □

Thanks to the claim, we conclude that $\#\mathcal{B} \le \#\mathcal{U} + \#\mathcal{G}$ at each step. Recall that $a_i \in \mathcal{A}_2$ is declared good or bad at step i and is not affected thereafter. Hence, after the last step, there is no element of $\mathcal{U}$ left. This means that $\#\mathcal{B} \le \#\mathcal{G}$, and $\mathcal{G}$ takes up at least half of $\mathcal{A}_2$.

Now, with the final $\mathcal{G}$ in hand, for each $a_i \in \mathcal{G}$ we define

$$K_i := X \setminus \big(\mathcal{H}_D(a_i, x_0) \cup \mathcal{H}_D(a_i, b_i)\big).$$

Since $a_i \in \mathcal{G} \subseteq \mathcal{A}_2$, we have $\#(K_i \cap A) \ge N$. The remaining claim is:

Claim 4.5. *For every pair of distinct elements $a_i, a_j \in \mathcal{G}$, K_i and K_j do not intersect.*

To check this claim, suppose to the contrary that K_i and K_j has a common element z for some $i < j$ such that $a_i, a_j \in \mathcal{G}$. This means that $(x_0|z)_{a_i}, (x_0|z)_{a_j} < D$. Now let $\gamma : [0, L] \to X$ be the geodesic connecting x_0 to z. Lemma 3.4 guarantees timings $\tau, \tau' \in [0, L]$ such that $d_X(\gamma(\tau), a_i), d_X(\gamma(\tau'), a_j) < D + 12\delta$.

Recall that a_i and a_j are $100D$-apart. This implies that $|\tau - \tau'| > 97D$. If $\tau' \le \tau - 97D$, then we have

$$d_X(x_0, a_j) < d_X(x_0, a_i) - 97D + 2(D + 12\delta) \le d_X(x_0, a_i) - 90D,$$

which contradicts our labelling convention of elements of $\mathcal{A}_2$. Hence, $\tau \le \tau' - 97D$ holds. In particular, $a_j \in \mathfrak{A}_D(a_i, x_0)$.

Now note that $(x_0|z)_{a_j} < D$ and that

$$(x_0|a_i)_{a_j} \ge \big(\gamma(0)\big|\gamma(\tau)\big)_{\gamma(\tau')} - d_X(\gamma(\tau), a_i) - d_X(\gamma(\tau'), a_j) \ge 97D - (2D + 24\delta) \ge 90D.$$

Gromov's 4-point condition (Lemma 3.6) tells us that $(a_i|z)_{a_j} \le D + 4\delta$.

Meanwhile, $(b_i|z)_{a_i} < D$ because $z \notin \mathcal{H}_D(a_i, b_i)$. This time, $(z|a_j)_{a_i}$ is similar to $(z|\gamma(\tau'))_{\gamma(\tau)} \ge 97D$; we have $(z|a_j)_{a_i} \ge 90D$. Another application of Gromov's 4-point condition leads to $(b_i|a_j)_{a_i} < D + 4\delta$.

Now, Lemma 3.4 applies to the sequence (b_i, a_i, a_j, z) as $d_X(a_i, a_j) \ge 90D \ge 2(D+4\delta)$. We obtain a geodesic η from b_i to z that passes through the $(D+16\delta)$-neighborhoods of a_i and a_j in order. We conclude $a_j \in \mathfrak{A}_D(a_i, b_i)$.

In summary, $a_j \in \mathfrak{A}_2 \cap \mathfrak{A}_D(a_i, x_0) \cap \mathfrak{A}_D(a_i, b_i)$. This contradicts the goodness of a_i. Hence, z cannot exist, and K_i and K_j are disjoint.

With Claim 4.5 in hand, we have

$$\#A \ge \sum_{i: a_i \in \mathcal{G}} \#(K_i \cap A) \ge N \cdot \#\mathcal{G} \ge N \cdot \frac{\#\mathcal{A}_2}{2} \ge N \cdot \frac{\#\mathcal{A}_1}{2M} \ge \frac{1}{\epsilon}(\#A - \#A').$$

This ends the proof. □

Now suppose that a group G is acting properly on $X \ni x_0$. Then the stabilizer of x_0 is finite, and the G-orbit of x_0 is uniformly locally finite. By Proposition 4.1, we conclude that:

Proposition 4.6. *Let X be a δ-hyperbolic space with a basepoint x_0 and let G be a group acting properly on X. Then for each $\epsilon, D > 0$ there exists a constant $N = N(\epsilon, D, Y)$ such that for every finite set $A \subseteq G$ there exists a subset $A' \subseteq A$ satisfying:*

(1) $\#A' \ge (1-\epsilon)\#A$;
(2) For each $a \in A'$ there exist halfspaces $\mathcal{H}_1, \mathcal{H}_2 \subseteq X$ rooted at ax_0 with radius parameter D such that $\#\big(\{g \in A : gx_0 \notin (\mathcal{H}_1 \cup \mathcal{H}_2)\}\big) \le N$.

5. Supporting hyperplane lemma and the critical exponent γ

With an additional assumption that G is non-elementary, the hyperbolic magic lemma implies the following supporting hyperplane lemma. Still, we will prove it for general acylindrical actions.

We will work with the following form of halfspaces: given $x, y \in X$, let

$$\mathcal{H}_{half}(x, y) := \{z \in X : z \text{ is closer to } x \text{ than } y\}.$$

Proposition 5.1. *Let X be a δ-hyperbolic space with a basepoint x_0 and let G be a non-virtually cyclic group acting on X with a WPD loxodromic element f. Let S be a finite generating set. Then there exists D_0 such that, for each $\epsilon > 0$ and $D > D_0$ there exists a constant $N = N(\epsilon, D)$ such that for every finite set $A \subseteq G$ there exists a subset $A' \subseteq A$ satisfying:*

(1) $\#A' \ge (1-\epsilon)\#A$;

(2) For each $a \in A'$ there exist $b \in G$ such that $d_S(a,b) \le N$ and

$$\{gx_0 : g \in A\} \subseteq \mathcal{H}_{half}(bx_0, b \cdot f^D x_0)$$

and such that $\mathcal{H}_{half}(bx_0, bf^D x_0)$ and $bf^D w f^{-D} b^{-1} \cdot \mathcal{H}_{half}(bx_0, bf^D x_0)$ are disjoint.

This will follow from a weaker statement:

Proposition 5.2. *Let X be a δ-hyperbolic space and let G be a non-virtually cyclic group acting on X with a unital, axial WPD loxodromic element f. Let $x_0 \in Ax(f)$. Let S be a finite generating set. Then there exists D_0 such that, for each $\epsilon > 0$ and $D > D_0$ there exists a constant $N = N(\epsilon, D)$ such that for every finite set $A \subseteq G$ there exists a subset $A' \subseteq A$ satisfying:*

(1) $\#A' \ge (1-\epsilon)\#A$;
(2) For each $a \in A'$ there exist $b \in G$ such that $d_S(a,b) \le N$ and

$$\#\big(\{gx_0 : g \in A\} \setminus \mathcal{H}_{half}(bx_0, bf^D x_0)\big) \le N.$$

Proof. Let K be as in Theorem 3.13 and let $K_0 \ge K + 100\delta$. By enlarging K_0 if necessary, we may assume that $d_X(x_0, sx_0) \le K_0$ for each $s \in S$.

Note that $\{g \in G : g^2 \in EC(f)\}$ contains $EC(f)$ as an index-2 subgroup, which is virtually cyclic. Since G is not virtually cyclic, we can take $w \in S \setminus \{g \in G : g^2 \in EC(f)\}$. Then Theorem 3.13 guarantees that $\mathrm{diam}_\gamma(\gamma') \le K$ for distinct axes $\gamma, \gamma' \in \{Ax(f), w^{-1}Ax(f), wAx(f)\}$. By Corollary 3.10(5), $\mathrm{diam}_\kappa(\kappa') \le K_0$ for any subgeodesics κ, κ' of γ, γ' as well.

We then have:

Observation 5.3. *For $k, l \in \mathbb{Z}$ and distinct $m, n \in \{1, 0, -1\}$, we have*

$$\big(w^m f^k x_0 \,\big|\, w^n f^l x_0\big)_{x_0} \le 6K_0 + 8\delta.$$

To see this, note that $w^m[x_0, f^k x_0]$ has K_0-small projection onto $w^n[x_0, f^l x_0]$, which is $2d_X(w^m x_0, w^n x_0)$-close to $w^n x_0$. It follows that the projection of $w^m f^k x_0$ onto $w^n[x_0, f^l x_0]$ is $(K_0 + 2K_0|m-n|)$-close to $w^n x_0$, This implies

$$\big(w^m f^k x_0 \,\big|\, w^n f^l x_0\big)_{w^n x_0} \le 5K_0 + 8\delta.$$

Now the desired inequality follows from $d_X(x_0, w^n x_0) \le K_0$.

It follows that:

Observation 5.4. *There exists $K_1 > 0$ such that for each $g \in G$ either*

(1) $(gx_0|f^i x_0)_{x_0} < K_1$ for every $i \in \mathbb{Z}$ or
(2) $(gx_0|wf^i x_0)_{x_0} < K_1$ for every $i \in \mathbb{Z}$.

We define $\mathscr{W} : G \to \{id, w\}$ using the above observation. Namely, for each $g \in G$ we pick $\mathscr{W}(g) \in \{id, w\}$ such that $(g^{-1}x_0|\mathscr{W}(g)f^i x_0)_{x_0} < K_1$ for each $i \in \mathbb{Z}$, i.e., $(x_0|g\mathscr{W}(g)f^i x_0)_{gx_0} < K_1$. Let $D_0 = 10^4(K_1 + K_0 + \delta + 1)$.

We now begin the proof. Let $D > D_0$. Recall the definition of the elementary closure $EC(f)$ of f. Since $EC(f)$ is a finite extension of $\langle f \rangle$ and since $\{f^i x_0\}_{i \in \mathbb{Z}}$ is locally finite, the set

$$\big\{g \in EC(f) : d_X(x_0, gx_0) \le 2D + 4\delta\big\}$$

is finite. Hence, they are contained in $B_S(R') \subseteq G$ for some $R' > 0$. Now let $R := R' + 3 + D \cdot \|f\|_S$.

Given $g \in G$ we define the *anti-halfspace*

$$\mathfrak{A}_{\pm}(g) := \left\{ h \in G : \begin{array}{c} \exists 0 \le \tau_1 \le \tau,\ \exists \text{ geodesic } \gamma : [0,\tau] \to X \text{ starting at } x_0 \text{ such that} \\ d_X\big(\gamma(\tau_1),\, g\mathscr{W}(g) f^D w^{\pm 1} f^D x_0\big) < 0.02D, \\ d_X\big(\gamma(\tau),\, h\mathscr{W}(h) f^D x_0\big) < 0.02D. \end{array} \right\}.$$

Observation 5.5. *For each $g \in G$, each element $h \in \mathfrak{A}_{\pm}(g)$ satisfies that $d_X(x_0, hx_0) > d_X(x_0, gx_0) + 0.5D$. Moreover, $\mathfrak{A}_+(g)$ and $\mathfrak{A}_-(g)$ are disjoint.*

Proof of Observation 5.5. Suppose first that $h \in \mathfrak{A}_+(g)$. Let $0 \le \tau_1 \le \tau$ and let $\gamma : [0,\tau] \to X$ be the geodesic realizing the membership of h in $\mathfrak{A}_+(g)$. Now for the sequence

$$(y_0, y_1, y_2, y_3, y_4) := \big(x_0,\, g\mathscr{W}(g)x_0,\, g\mathscr{W}(g)f^D x_0,\, g\mathscr{W}(g) f^D w f^D x_0,\, \gamma(\tau_1)\big),$$

we observe that

(1) $(y_0|y_2)_{y_1} \le 0.01D, (y_1|y_3)_{y_2} \le 0.01D, (y_2|y_4)_{y_3} \le 0.02D$.
(2) $d_X(y_1, y_2), d_X(y_2, y_3) \ge 0.95D$.

The first item follows from the fact that $(y_0|y_2)_{gx_0} \le K_1$, $d_X(y_1, gx_0) < K_0$, $(y_1|y_3)_{y_2} \le 6K_0 + 8\delta$, $d_X(y_3, y_4) \le 0.02D$. The second item is due to the fact $d_X(x_0, f^D x_0) = D$ and $d_X(x_0, sx_0) \le K_0$.

By Lemma 3.4, there exist $0 \le t_1 \le t_2 \le t_3 \le \tau_1$ such that

$$d_X\big(\gamma(t_1), y_1\big), d_X\big(\gamma(t_2), y_2\big) \le 0.011D,\ d_X\big(\gamma(t_3), y_3\big) \le 0.021D. \tag{5.1}$$

This implies that

$$\begin{aligned}\tau_1 = d_X(x_0, \gamma(\tau_1)) &= d_X(x_0, \gamma(t_1)) + d_X(\gamma(t_1), \gamma(t_2)) + d_X(\gamma(t_2), \gamma(t_3)) + d_X(\gamma(t_3), \gamma(\tau)) \\ &\ge d_X(x_0, y_1) + d_X(y_1, y_2) + d_X(y_2, y_3) - 2(0.011D + 0.011D + 0.021D) \\ &\ge \big(d_X(x_0, gx_0) - d_X(gx_0, g\mathscr{W}(g)x_0)\big) + 1.9D - 0.09D \ge d_X(x_0, gx_0) + 1.8D.\end{aligned}$$

Meanwhile, note that

$$\begin{aligned}\tau_1 \le \tau &\le d_X(x_0, h\mathscr{W}(h) f^D x_0) + 0.02D \\ &\le d_X(x_0, hx_0) + d_X(hx_0, h\mathscr{W}(h)f^D x_0) + 0.02D \\ &\le d_X(x_0, hx_0) + d_X(x_0, f^D x_0) + 0.021D = d_X(x_0, hx_0) + 1.021D.\end{aligned}$$

Comparing these two inequalities lead to $d_X(x_0, gx_0) + 0.5D < d_X(x_0, hx_0)$.

Note that $\big(\gamma(t_2) \,\big|\, \gamma(\tau)\big)_{\gamma(\tau_1)} = 0$ as $t_2 \le \tau_1 \le \tau$. Since y_2, y_3 and $h\mathscr{W}(h)f^D x_0$ are $0.02D$-close to $\gamma(t_2)$, $\gamma(\tau)$ and $\gamma(\tau')$, respectively, we have

$$\big(g\mathscr{W}(g)f^D x_0 \,\big|\, h\mathscr{W}(h)f^D x_0\big)_{g\mathscr{W}(g)f^D w f^D x_0} < 0.06D. \tag{5.2}$$

Now, let us suppose that $h \in \mathfrak{A}_-(g)$ in addition and deduce contradiction. Just as we had Inequality 5.2, we have

$$(g\mathscr{W}(g)f^D x_0 | h\mathscr{W}(h)f^D x_0)_{g\mathscr{W}(g)f^D w^{-1} f^D x_0} \le 0.06D. \tag{5.3}$$

Finally, Observation 5.3 tells us that

$$(g\mathscr{W}(g)f^D w f^D x_0 | g\mathscr{W}(g)f^D w^{-1} f^D x_0)_{g\mathscr{W}(g)f^D x_0} \le 0.01D.$$

Furthermore, we know that $d_X(g\mathscr{W}(g)f^D x_0, g\mathscr{W}(g)f^D w^{\pm 1} f^D x_0) \ge 0.95D$. By Lemma 3.4, we have $d_X(h\mathscr{W}(h)f^D x_0, h\mathscr{W}(h)f^D x_0) \ge 0.95D + 0.95D - 2 \cdot (0.06D + 0.01D + 0.06D + 100\delta) \ge 1.8D$, a contradiction. □

We then observe:

Observation 5.6. *Suppose that $\mathfrak{A}_+(g)$ and $\mathfrak{A}_+(h)$ has nonempty intersection, and suppose that $d_S(g,h) > R$. Then either $g \in \mathfrak{A}_+(h)$ or $h \in \mathfrak{A}_+(g)$.*

Proof of Observation 5.6. Let us pick an element $a \in \mathfrak{A}_+(g) \cap \mathfrak{A}_+(h)$.

Without loss of generality, we suppose that $d_X(x_0, gx_0) \ge d_X(x_0, hx_0)$. Let γ (γ', resp.) be the geodesic and let $\tau_1 \le \tau$ (σ_1, σ, resp.) be the timing that realize the membership of a in $\mathfrak{A}_+(g)$ ($\mathfrak{A}_+(h)$, resp.). Then τ and σ differ by at most $0.04D$, as

$$\tau =_{0.02D} d_X\big(x_0, a\mathscr{W}(a)f^D x_0\big) =_{0.02D} \sigma.$$

We now let $L_{min} := \big(\gamma(\tau) \,\big|\, \gamma'(\sigma)\big)_{x_0}$, which satisfies

$$L_{min} := d\big(x_0, \gamma(\tau)\big) - \big(x_0 \,\big|\, \gamma'(\sigma)\big)_{\gamma(\tau)} \ge \tau - 0.04D.$$

Similarly, L_{min} is greater than $\sigma - 0.04D$. By Lemma 3.5, we have

$$d_X(\gamma(t), \gamma'(t)) < 4\delta \quad (0 \le t \le L_{min}).$$

We observed earlier that Lemma 3.4 applies to the sequence

$$\big(x_0,\, g\mathscr{W}(g)x_0,\, g\mathscr{W}(g)f^D x_0,\, g\mathscr{W}(g)f^D w f^D x_0,\, \gamma(\tau_1)\big).$$

In particular, there exist $0 \le t_1 \le t_2 \le \tau_1$ such that

$$d_X\big(g\mathscr{W}(g)x_0,\, \gamma(t_1)\big), d_X\big(g\mathscr{W}(g)f^D x_0,\, \gamma(t_2)\big) \le 0.011D.$$

Note that

$$|t_1 - d_X(x_0, gx_0)| \le 0.011D + d_X(x_0, \mathscr{W}(g)x_0) \le 0.012D.$$

By Lemma 3.2, $g\mathscr{W}(g)[x_0, f^D x_0]$ and $\gamma([t_1, t_2])$ are $0.012D$-equivalent. This forces $t_2 - t_1 =_{0.022D} d_X(x_0, f^D x_0) = D$. Similarly, $g\mathscr{W}(g)f^D w[x_0, f^D x_0]$ and $\gamma([t_2, \tau_1])$ are $0.021D$-equivalent and $\tau_1 - t_2 =_{0.032D} D$.

Similarly, there exist $0 \le s_1 \le s_2 \le \sigma_1$ for γ' such that

$$d_X\big(h\mathscr{W}(h)x_0, \gamma(s_1)\big) d_X\big(h\mathscr{W}(h)f^D x_0, \gamma'(s_2)\big) \le 0.01D$$

Note here that t_2 or s_2 are much smaller than τ_1 or σ_1, respectively, so they are smaller than L_{min}. In particular, $d_X(\gamma(t_2), \gamma'(t_2)) < 4\delta$ holds. Moreover, s_1 is $0.012D$-close to $d_X(x_0, hx_0)$. Since we assumed that hx_0 is closer than gx_0 to x_0, we obtain

$$t_1 \ge s_1 - 0.024D.$$

Observe that the geodesic γ' and the two timing σ_1, t_2 satisfy

$$d_X\big(\gamma'(\sigma_1), h\mathscr{W}(h)f^D w f^D x_0\big) < 0.02D,$$
$$\begin{aligned} d_X\big(\gamma'(t_2), g\mathscr{W}(g)f^D x_0\big) &\le d_X(\gamma'(t_2), \gamma(t_2)) + d_X\big(\gamma(t_2), g\mathscr{W}(g)f^D x_0\big) \\ &\le 4\delta + 0.011D < 0.02D. \end{aligned}$$

In the remaining, we will show that $\sigma_1 \le t_2$. This will guarantee $g \in \mathfrak{A}_+(h)$ and end the proof.

Suppose to the contrary that $\sigma_1 > t_2$. Then we have

$$\big[t_1 + 0.024D, t_2\big] \subseteq [s_1, \sigma_1] = [s_1, s_2] \cup [s_2, \sigma_1].$$

In particular, one of $[s_1, s_2]$ and $[s_2, \sigma_1]$ should overlap with $[t_1 + 0.024D, t_2]$ for length at least $\frac{1}{2}(t_2 - t_1 - 0.024D) \ge 0.46D$.

(1) $I := [t_1, t_2] \cap [s_1, s_2]$ is longer than $0.46D$. Recall that $g\mathscr{W}(g)[x_0, f^D x_0]$ and $\gamma([t_1, t_2])$ are $0.012D$-equivalent. Hence, there exist $p_1, q_1 \in g\mathscr{W}(g)[x_0, f^D x_0]$ that are $0.012D$-close to $\gamma(\min I)$ and $\gamma(\max I - 0.05D)$, respectively. We can also take $p_2, q_2 \in h\mathscr{W}(h)[x_0, f^D x_0]$ that are $0.012D$-close to $\gamma'(\min I)$ and $\gamma'(\max I - 0.05D)$, respectively.

Since $\min I \le \max I - 0.05D \le \tau - 0.05D \le L_{min}$, γ and γ' are 4δ-fellow traveling at $t = \min I, \max I - 0.05D$. We thus have

$$d_X(p_1, p_2), d_X(q_1, q_2) \le 0.025D, \quad d_X(p_2, q_2) \ge |I| - 0.05D - 2\cdot 0.012D \ge 0.4D.$$

This means that the projection of $\{p_1, q_1\} \subseteq g\mathscr{W}(g)[x_0, f^D x_0]$ onto $h\mathscr{W}(h)[x_0, f^D x_0]$ is larger than $0.2D$. By Corollary 3.10(5), we have $d_{h\mathscr{W}(h)Ax(f)}(g\mathscr{W}(g)Ax(f)) \ge 0.1D \ge K_0$. Theorem 3.13 implies that $\mathscr{W}(g)^{-1} \cdot g^{-1}h\mathscr{W}(h) \in EC(f)$. Moreover, note that

$$\begin{aligned} d_X(g\mathscr{W}(g)x_0, h\mathscr{W}(h)x_0) &\le d_X(g\mathscr{W}(g)x_0, p_1) + d_X(p_1, p_2) + d_X(p_2, h\mathscr{W}(h)x_0) \\ &\le d_X(x_0, f^D x_0) + 4\delta + d_X(x_0, f^D x_0). \end{aligned}$$

In summary, we have

$$\mathscr{W}(g)^{-1} \cdot g^{-1}h\mathscr{W}(h) \in EC(f) \cap \{u \in G : d_X(x_0, ux_0) \le 2D + 4\delta\}.$$

In other words, $\mathscr{W}(g)^{-1} \cdot g^{-1}h\mathscr{W}(h) \in B_S(R')$ and $g^{-1}h \in B_S(R)$. This is a contradiction.

(2) $I := [t_1, t_2] \cap [s_2, \sigma_1]$ is longer than $0.46D$. In this case, we can similarly take points $p_1, q_1 \in g\mathscr{W}(g)[x_0, f^D x_0]$ that are $0.012D$-close to $\gamma(\min I)$ and $\gamma(\max I - 0.05D)$, respectively. We can also pick p_2, q_2 on $h\mathscr{W}(h)f^D w[x_0, f^D x_0]$ that are $0.021D$-close to $\gamma'(\min I)$ and $\gamma'(\max I - 0.05D)$, respectively. Again, $d_X(\gamma(t), \gamma'(t)) \le 4\delta$ for $t = \min I, \max I - 0.05D$. Then we have

$$d_X(p_1, p_2), d_X(q_1, q_2) \le 0.034D, \quad d_X(p_2, q_2) \ge |I| - 0.05D - 2\cdot 0.021D \ge 0.36D.$$

Then the projection of $\{p_1, q_1\} \subseteq g\mathscr{W}(g)[x_0, f^D x_0]$ onto $h\mathscr{W}(h)f^D w[x_0, f^D x_0]$ is larger than $0.2D$. By Corollary 3.10(2), $g\mathscr{W}(g)Ax(f)$ has projection $\ge 0.1D$ onto $h\mathscr{W}(h)f^D wAx(f)$. Theorem 3.13 implies that

$\mathscr{W}(g)^{-1}\cdot g^{-1}h\mathscr{W}(h)f^D w \in EC(f)$. Moreover, note that

$$d_X(g\mathscr{W}(g)x_0, h\mathscr{W}(h)f^D wx_0) \le d_X(g\mathscr{W}(g)x_0, p_1) + d_X(p_1,p_2) + d_X(p_2, h\mathscr{W}(h)f^D wx_0) \le d_X(x_0, f^D x_0) + 4\delta + d_X(x_0, f^D x_0).$$

In summary, we have

$$\mathscr{W}(g)^{-1}\cdot g^{-1}h\mathscr{W}(h)f^D w \in EC(f)\cap \{u\in G : d_X(x_0, ux_0)\le 2D+4\delta\}.$$

In other words, $\mathscr{W}(g)^{-1}\cdot g^{-1}h\mathscr{W}(h)f^D w \in B_S(R')$ and $g^{-1}h \in B_S(R)$. This is a contradiction.

Hence, neither situation can happen and we conclude $g \in \mathfrak{A}_+(h)$. ☐

For the same reason, we have

Observation 5.7. *Let $\epsilon,\epsilon' \in \{+,-\}$. Suppose that $\mathfrak{A}_\epsilon(g)$ and $\mathfrak{A}_{\epsilon'}(h)$ has nonempty intersection, and suppose that $d_S(g,h) > R$. Then either $g \in \mathfrak{A}_{\epsilon'}(h)$ or $h \in \mathfrak{A}_\epsilon(g)$.*

We finally need:

Observation 5.8. *For each $g \in G$, we have*

$$G\setminus \mathfrak{A}_\pm(g) \subseteq \mathcal{H}_{half}\big(g\mathscr{W}(g)f^D w^{\pm 1} f^{2D}x_0, g\mathscr{W}(g)f^D w^{\pm 1}f^{3D}x_0\big).$$

Proof of Observation 5.8. For convenience, let us write $\mathfrak{g} := g\mathscr{W}(g)f^D w^{\pm 1}$.

Let us pick $u \in G\setminus \mathcal{H}_{half}(\mathfrak{g}f^{2D}x_0, \mathfrak{g}f^{3D}x_0)$. Let $k,l\in\mathbb{Z}$ be such that

$$\pi_{\mathfrak{g}Ax(f)}(ux_0)\cap[\mathfrak{g}f^k, \mathfrak{g}f^{k+1}] \neq \emptyset, \quad \pi_{\mathfrak{g}Ax(f)}(u\mathscr{W}(g)f^D x_0)\cap [\mathfrak{g}f^l, \mathfrak{g}f^{l+1}]\neq\emptyset.$$

Corollary 3.10(2) tells us that

$$d_X(\mathfrak{g}f^i x_0, ux_0) =_{30\delta+1} d_X(\mathfrak{g}f^i x_0, \mathfrak{g}f^k x_0) + d_X(\mathfrak{g}f^k x_0, ux_0) = |i-k| + d_X(\mathfrak{g}f^k x_0, ux_0). \quad (\forall i\in\mathbb{Z})$$

Since ux_0 is not closer to $\mathfrak{g}f^{2D}x_0$ than to $\mathfrak{g}f^{3D}x_0$, we conclude that $k \ge 2.45D$. Meanwhile, note that $d_X(ux_0, u\mathscr{W}(u)f^D x_0) \le 1.01D$. By Corollary 3.10(1), $l \ge 1.42D$. This means that

$$\big(\mathfrak{g}f^D x_0 \,\big|\, u\mathscr{W}(u)f^D x_0\big)_{\mathfrak{g}f^{2D}x_0} \le 0.6D. \tag{5.4}$$

Now for

$$(y_0,y_1,y_2,y_3,y_4) := \big(x_0,\, g\mathscr{W}(g)x_0, g\mathscr{W}(g)f^D x_0, \mathfrak{g}f^D x_0, \mathfrak{g}f^{2D}x_0\big)$$

we have $(y_{i-1}|y_{i+1})_{y_i} \le 0.01D$ for $i = 1,2,3$ and $d_X(y_{i-1},y_i)\ge 0.95D$ for $i = 2,3,4$. Combining this with Inequality 5.4, we can apply Lemma 3.4 and conclude that $[x_0, u\mathscr{W}(u)f^D x_0]$ is $0.011D$-close to $g\mathscr{W}(g)f^D w^{\pm 1} f^D x_0$. Hence, $x \in \mathfrak{A}^\pm(g)$. ☐

This time, we will define

$$\mathcal{A}_1 := \{a \in A : \#\big(A\setminus \mathcal{H}_{half}(vx_0, vf^D x_0)\big) \ge N \text{ for } v = a\mathscr{W}(a)f^D wf^{2D}, a\mathscr{W}(a)f^D w^{-1}f^{2D}\}.$$

We now pick a subset $\mathcal{A}_2$ of $\mathcal{A}_1$ that is maximally R-separated *in the word metric d_S*, i.e., we have

(1) $d_S(a,a') \ge R$ for each pair of distinct elements $a,a' \in \mathcal{A}_2$;

(2) $\mathcal{A}_2$ is a maximal subset of $\mathcal{A}_1$ satisfying this property.

Then $\bigcup_{a \in \mathcal{A}_2} (a \cdot B_S(R) \cap A)$ covers entire $\mathcal{A}_1$. We conclude

$$\#\mathcal{A}_2 \geq \frac{1}{\#B_S(R)} \cdot \#\mathcal{A}_1 \leq \frac{1}{(2\#S)^R} \#\mathcal{A}_1.$$

As before, we prepare empty collections $\mathcal{B} = \mathcal{U} = \mathcal{G} = \emptyset$. Enumerate $\mathcal{A}_2$ by the distance from x_0, i.e., let $\mathcal{A}_2 = \{a_1, a_2, \ldots, a_{\#\mathcal{A}_2}\}$ be such that $d_X(x_0, a_i) \leq d_X(x_0, a_{i+1})$ for each i. At each step $i = 1, \ldots, \#\mathcal{A}_2$, we will put a_i in either $\mathcal{B}$ or $\mathcal{G}$; this decision is final and shall not be modified further. We may put some other elements of $\mathcal{A}_2$ in $\mathcal{U}$, whose their classification will change later. When a_i is declared good, then we will also define its sign $\sigma(a_i) \in \{+1, -1\}$. This way, we will obtain a function $\sigma : \mathcal{G} \to \{+1, -1\}$ in the end.

We will keep the balance $\#\mathcal{B} \leq \#\mathcal{U} + \#\mathcal{G}$ throughout. Finally, after the last step there will be no $\mathcal{U}$-element. At the end we will have $\#\mathcal{B} \leq \#\mathcal{G}$.

We now describe the procedure. At step i,

(1) if $\mathcal{A}_2 \cap \mathfrak{A}_+(a_i)$ has no element, then we declare $a_i \in \mathcal{G}$ and $\sigma(a_i) = +1$;
(2) if not (1) and if $\mathcal{A}_2 \cap \mathfrak{A}_-(a_i)$ has no element, then we declare $a_i \in \mathcal{G}$ and $\sigma(a_i) = -1$;
(3) if not (1) and (2), we pick $b_i \in \mathcal{A}_2 \cap \mathfrak{A}_+(a_i)$ and $c_i \in \mathcal{A}_2 \cap \mathfrak{A}_-(a_i)$ whose orbit points are the closest to x_0. We declare $a_i \in \mathcal{B}$ and $b_i, c_i \in \mathcal{U}$.

Till step i, $\mathcal{G} \cup \mathcal{B}$ comprises of elements from $\{a_1, \ldots, a_i\}$; they do not contain any of $a_{i+1}, a_{i+2}, \ldots$. $(*)$ We now describe what happens at step i.

In case (1) or (2), $\mathcal{G}$ gains one more element that might be from $\mathcal{U}$ or not. $\mathcal{B}$ does not change. Overall, $\#\mathcal{B}$ stays the same and $\#\mathcal{U} + \#\mathcal{G}$ does not decrease. Similar situation happens in Case (2-a).

In case (2-b), $\mathcal{B}$ gains one element a_i, which might be from $\mathcal{U}$. In exchange, $\mathcal{U}$ gains elements b_i and c_i. Here Observation 5.5 guarantees that $d_X(x_0, b_i x_0), d_X(x_0, c_i x_0) > d_X(x_0, a_i x_0)$ and that b_i and c_i are distinct. Since $\mathcal{A}_2$ was labelled with respect to the distance from x_0, we conclude that $b_i, c_i \in \{a_{i+1}, a_{i+2}, \ldots\}$; in other words, neither b_i nor c_i come from $\mathcal{G} \cup \mathcal{B}$. Hence, we conclude that elements in $\mathcal{G} \cup \mathcal{B}$ are never re-classified.

As before, we need to show that for each $i < j$ such that $a_i, a_j \in \mathcal{B}$, $\{b_i, c_i\}$ and $\{b_j, c_j\}$ are disjoint. Suppose to the contrary that $b_i = b_j$. Then Observation 5.6 tells us that either (1) $a_i \in \mathfrak{A}_+(a_j)$ or (2) $a_j \in \mathfrak{A}_+(a_i)$. In the latter case, we have $d_X(x_0, a_i) > d_X(x_0, a_i)$, contradicting the labelling scheme. In the former case, we have

$$d_X(x_0, a_i x_0) < d_X(x_0, a_j x_0) < d_X(x_0, b_i x_0)$$

by Observation 5.5. This violates the minimality of b_i. Hence, $b_i = b_j$ cannot happen. Likewise, using Observation 5.7 we can exclude the cases $b_i = c_j$, $c_i = b_j$ and $c_i = c_j$. Thus, $\{b_i, c_i\}$ and $\{b_j, c_j\}$ are disjoint.

By the same logic as in the previous proof, we have $\#\mathcal{G} \geq \#\mathcal{B}$ in the end. Moreover, for distinct $a, b \in \mathcal{G}$, $\mathfrak{A}_{\sigma(a)}(a)$ and $\mathfrak{A}_{\sigma(b)}(b)$ are disjoint; if not, Observation 5.7 implies either $b \in \mathfrak{A}_{\sigma(a)}(a)$ or $a \in \mathfrak{A}_{\sigma(b)}(b)$, contradicting the goodness of a and b.

Since $X \setminus \mathcal{H}_{half}(a\mathscr{W}(a)f^D w^{\sigma(a)} f^{2D}x_0, a\mathscr{W}(a)f^{3D}w^{\sigma(a)}f^{2D}x_0) \subseteq \mathcal{A}_{\sigma(a)}(a)$ has at least N elements for $a \in \mathcal{G} \subseteq \mathcal{A}_1$, we conclude

$$\#\mathcal{A} \geq N \cdot \#\mathcal{G} \geq N \cdot \frac{\#\mathcal{A}_2}{2} \geq \frac{N}{(2\#S)^R}\#\mathcal{A}_1 \geq \frac{1}{\epsilon}\#\mathcal{A}_1.$$

This ends the proof. □

We now need:

Lemma 5.9. *Let X be a δ-hyperbolic space with a basepoint x_0 and let G be non-virtually cyclic group acting on X with an axial WPD loxodromic element f. Let $x_0 \in Ax(f)$. Then there exists D_0 such that the following holds for each $D > D_0$.*

Let $A \subseteq X$ be a finite set in $X \setminus \mathcal{H}_{half}(x_0, f^D x_0)$ and let $N = \#A$. Then there exists $a_1, \ldots, a_N \in \{f^D, wf^D\}$ such that

(1) $H := \mathcal{H}_{half}(f^D a_1 \cdots a_N x_0, f^D a_1 \cdots a_N f^D x_0)$ contains A, and
(2) H and $f^D a_1 \cdots a_N \cdot f^D \cdot w \cdot f^{-D} \cdot a_N^{-1} \cdots a_1^{-1} f^{-D} H$ are disjoint.

Proof. Let K_0, K_1, D_0 be as in the proof of Proposition 5.2. Suppose $D > D_0$. We claim that the 2^N halfspaces

(5.5)
$$\Big\{X \setminus \mathcal{H}_{half}(f^D a_1 \cdots a_N x_0, f a_1^D \cdots a_N f^D x_0) : a_1, \ldots, a_N \in \{f^D, wf^D\}\Big\}$$

are mutually disjoint subsets of $X \setminus \mathcal{H}_{half}(x_0, f^D x_0)$.

To see the disjointness, let $z \notin \mathcal{H}_{half}(f^D a_1 \cdots a_N x_0, f a_1^D \cdots a_N f^D x_0)$ and $z' \notin \mathcal{H}_{half}(f^D b_1 \cdots b_N x_0, f b_1^D \cdots b_N f^D x_0)$ for some $(a_1, \ldots, a_N) \neq (b_1, \ldots, b_N) \in \{f^D, wf^D\}^N$. Let m be the minimal one such that $a_m \neq b_m$.

Let $z_i := f^D a_1 \cdots a_i x_0$ and $z_i' := f^D b_1 \cdots b_i x_0$ for $i \geq m-1$. We observe that Lemma 3.4 applies to

$$(z, z_N, z_{N-1}, \ldots, z_m, z_{m-1} := z'_{m-1}, z'_m, \ldots, z'_N, z').$$

Indeed, we check that

$$\begin{aligned}
(z_i|z_{i-2})_{z_{i-1}}, (z_i'|z_{i-2}')_{z_{i-1}'} &\leq 6K_0 + 8\delta \leq 0.01D \quad (i = m+1, \ldots, N),\\
(z|z_{N-1})_{z_N}, (z'|z'_{N-1})_{z'_N} &\leq \frac{1}{2} d_X(x_0, f^D x_0) \qquad\qquad \leq 0.5D,\\
(z_m|z'_m)_{z_{m-1}} &\leq 0.01D,\\
d_X(z_i, z_{i-1}), d_X(z_i', z_{i-1}') &\geq 0.99D \qquad\qquad (i = m, \ldots, N).
\end{aligned}$$

Consequently, we have $d_X(z, z') \geq 2 \cdot 0.9D$ and $z \neq z'$.

For the same reason, for each $a_1, \ldots, a_N \in \{f^D, wf^D\}$, $\mathcal{H}_{half}(x_0, f^D x_0)$ and $X \setminus \mathcal{H}_{half}(f^D a_1 \cdots a_N x_0, f^D a_1 \cdots a_N f^D x_0)$ are disjoint. This implies that the latter is contained in $X \setminus \mathcal{H}_{half}(x_0, f^D x_0)$. Hence, the sets in Display 5.5 are indeed 2^N disjoint subsets of $X \setminus \mathcal{H}_{half}(x_0, f^D x_0$. One of them should avoid A by the pigeonhole principle. Item (1) of the conclusion now follows.

Moreover, a similar logic shows that $X \backslash \mathcal{H}_{half}(x_0, f^{-D}x_0)$ and $X \backslash \mathcal{H}_{half}(wx_0, wf^{-D}x_0)$ are disjoint, as $Ax(f)$ and $wAx(f)$ have K_0-bounded projections onto each other. This leads to Item (2) of the conclusion. □

Proposition 5.1 now follows from Proposition 5.2 and Lemma 5.9. Therefore acylindrically hyperbolic groups satisfy the assumption of Theorem 2.8.

6. Branching set

Recall the notions of barriers and roughly branching sets (Definition 2.10, 2.12). Recall that $\mathcal{H}_R(x_0, y) := \{z \in X : (z|y)_{x_0} > R\}$. Our aim is to show:

Proposition 6.1. *Let X be a δ-hyperbolic space and let $G \le \mathrm{Isom}(X)$ be a non-virtually cyclic group with a unital, axial WPD element $f \in G$. Let $x_0 \in Ax(f)$. Let S be a finite generating set of G. Then there exists $r > 0$ such that the following holds.*

Let $R > 0$ and $y \in X$. Then there exists an r-branching subset $B = B_1 \sqcup \ldots \sqcup B_{R/r} \subseteq G$ such that, for every $g \in G$ such that $gx_0 \in \mathcal{H}_R(x_0, y)$, every d_S-path connecting id to g passes through each of $B_1, \ldots, B_{R/r}$.

Proof. Let $K_0 = K > 1000\delta$ be as in Theorem 3.13 for G and f. Recall that $EC(f)$ is a virtually cyclic subgroup of G. Now let

$$\mathscr{A} := \{gAx(f) : g \in G\}.$$

Note that $\mathrm{diam}_\gamma(\gamma') \le K_0$ for distinct axes $\gamma, \gamma' \in \mathscr{A}$,

By enlarging K_0, we can guarantee that $d_X(x_0, sx_0) < K_0$ for each $s \in S$. Since G is not virtually cyclic, we can take $w \in S \setminus \{g \in G : g^2 \in EC(f)\}$. Then we have $d_X(x_0, wx_0) \le K_0$. By enlarging K_0 once again, we can guarantee the following:

Observation 6.2. *for every $x_1, x_2 \in X$ either*

1. *$(x_j \,|\, p)_{x_0} < K_0$ for each $j \in \{1, 2\}$ and $p \in Ax(f)$;*
2. *$(x_j \,|\, wp)_{x_0} < K_0$ for each $j \in \{1, 2\}$ and $p \in Ax(f)$, or*
3. *$(x_j \,|\, w^{-1}p)_{x_0} < K_0$ for each $j \in \{1, 2\}$ and $p \in Ax(f)$.*

We now describe the roughly branching barrier. Let

$$I_i := \big[100K_0 i - 25K_0, 100K_0 i + 25K_0\big], \quad J_i := \big[100K_0 i - K_0, 100K_0 i + K_0\big],$$

$$B_i := \left\{ g \in G : \begin{array}{c} (gx_0 \,|\, y)_{x_0} \in I_i, \\ \forall \gamma \in \mathscr{A} \big[d_\gamma(x_0, y) \ge 5K_0 \vee d_\gamma(x_0, gx_0) \le 100K_0 \big] \end{array} \right\}.$$

We claim that:

Claim 6.3. *Let $P = (g_1, g_2, \ldots, g_N)$ be a d_S-path such that $(g_1x_0|y)_{x_0} \in I_0$ and $(g_Nx_0|y)_{x_0} \in I_2$. Then there exists i such that $g_i \in B_1$.*

Proof of Claim 6.3. For this proof, let

$$\mathscr{A}' := \{\gamma \in \mathscr{A} : d_\gamma(x_0, y) < 5K_0\}.$$

Suppose to the contrary that P does not pass through B_1. Recall that for each $z, z' \in X$, $(z|y)_{x_0}$ and $(z'|y)_{x_0}$ differ by at most $d_X(z, z')$. Hence, along the d_S-path $(g_1, g_2, \ldots, g_N)$, the quantity $(g_i x_0|y)_{x_0}$ changes by at most K_0 at each step i. Since $(g_i x_0|y)_{x_0}$ changes from less than $20K_0$ to more than $180K_0$, there exists a step $i(1)$ for which $(g_{i(1)}x_0|y)_{x_0}$ lies in J_1.

Since we supposed that P does not pass through B_1, x_0 and $g_{i(1)}x_0$ are $100K_0$-separated along some $\gamma \in \mathscr{A}'$. In particular,

$$\mathcal{C}_0 := \left\{\gamma \in \mathscr{A}' : d_\gamma(x_0, g_{i(1)}x_0) \geq 80K_0\right\}$$

is non-empty. We pick $\gamma_0 \in \mathcal{C}_0$ that is the closest to x_0.

At this moment, we observe:

Observation 6.4. *Let $\gamma_1, \gamma_2, \ldots, \gamma_n \in \mathscr{A}'$ and $z \in X$ be such that:*

(1) $d_{\gamma_{i-1}}(x_0, \gamma_i) \geq 50K_0$ for $1 \leq i \leq n$;

(2) $d_{\gamma_n}(x_0, z) \geq 50K_0$.

Then $(z|y)_{x_0}$ and $(g_{i(1)}x_0|y)_{x_0}$ are $22K_0$-close. In particular, $(z|y)_{x_0}$ lies in I_1 and not in I_2.

To see this, suppose that $\gamma_1, \ldots, \gamma_n \in \mathscr{A}'$ and $z \in X$ satisfy the assumption. Then $d_{\gamma_0}(x_0, z) \geq 46K_0$ by Lemma 3.14. Let $p \in \pi_{\gamma_0}(x_0)$ and $q \in \pi_{\gamma_0}(z)$. Then $[x_0, z]$ is $0.01K_0$-close to p. Meanwhile, recall that $d_{\gamma_0}(x_0, y) < 5K_0$, which implies $d_{\gamma_0}(y, z) \geq 40K_0$. By Corollary 3.10(2), $[y, z]$ is $0.01K_0$-close to $\pi_{\gamma_0}(y)$, which is $5K_0$-close to p. In conclusion, $[x_0, z]$ and $[y, z]$ are $0.01K_0$-close and $5.01K_0$-close to p, respectively. Hence, $(z|y)_{x_0}$ and $(p|y)_{x_0}$ differ by at most $10.1K_0$.

Now recall that $d_{\gamma_0}(x_0, g_{i(1)}x_0) \geq 100K_0$. For the same reason, $(g_{i(1)}x_0|y)_{x_0}$ and $(p|y)_{x_0}$ differ by at most $10.1K_0$. Observation 6.4 now follows.

Let us now go back to the proof of the claim. If $d_{\gamma_0}(x_0, g_j x_0) \geq 80K_0$ for all $j \geq i(1)$, then Observation 6.4 tells us that the $(g_N x_0|y)_{x_0}$ lies in I_1 and not in I_2, a contradiction. Hence, we can pick the earliest $i(2) > i(1)$ such that $d_{\gamma_0}(x_0, g_{i(2)}x_0) \leq 80K_0$. By the coarse Lipschitzness of $\pi_{\gamma_0}(\cdot)$, we have $d_{\gamma_0}(x_0, g_{i(2)}x_0) \geq 78K_0$, and Observation 6.4 still tells us that $(g_{i(2)}x_0|y)_{x_0}) \in I_1$. Since $g_{i(2)} \in P$ is assumed not to be in B_1, the collection

$$\mathcal{C}_1 := \{\gamma \in \mathscr{A}' : d_\gamma(x_0, g_{i(2)}x_0) > 100K_0\}$$

is nonempty. We pick $\gamma_1 \in \mathcal{C}_1$ that is the closest to x_0. Clearly $\gamma_1 \neq \gamma_0$.

Note that $[x_0, g_{i(2)}x_0]$ has large projections onto both γ_0 and γ_1. Let η_0 and η_1 be subsegments of $[x_0, g_{i(2)}x_0]$ that are 12δ-equivalent to the two projections, respectively. Then η_0 is at least $77K_0$-long and η_1 is at least $99K_0$-long. Moreover, recall that $\mathrm{diam}_{\gamma_0}(\gamma_1) < K_0$ as distinct axes in $\mathscr{A}$ have K_0-bounded projection. This implies that η_0 and η_1 overlap for length less than $2K_0$.

Suppose to the contrary that $d_X(x_0, \gamma_1) < d_X(x_0, \gamma_0)$. This implies that η_1 appears earlier than η_0 along $[x_0, g_{i(2)}x_0]$. Since they do not overlap much and since η_0 is long enough, we can take $p \in \eta_0$ such that

$$\begin{aligned} d_X(g_{i(2)}x_0, p) &\le d_X(g_{i(2)}x_0, \eta_1) - 75K_0 \\ &\le d_X\big(g_{i(2)}x_0, \pi_{\gamma_1}([x_0, g_{i(2)}x_0])\big) - 74K_0 = d_X(g_{i(2)}x_0, \gamma_1) - 74K_0. \end{aligned}$$

By Lemma 3.9 we have $d_{\gamma_1}(p, g_{i(2)}x_0) \le 12\delta$. Since p is 12δ-close to γ_0 and since γ_0 has bounded projection onto γ_1 (as they are distinct!), we conclude that $d_{\gamma_1}(\gamma_0, g_{i(2)}x_0) \le 3K_0$. As a result, we have

$$d_{\gamma_1}(x_0, \gamma_0) \ge 97K_0.$$

Let us observe $\mathcal{C}_0$ for the moment. Since $d_{\gamma_0}(x, g_{i(1)}x_0) \ge 100K_0$, there exists a point $p \in [x, g_{i(1)}x_0]$ that is 12δ-close to some $q \in \gamma_0$. Since $d_{\gamma_1}(x_0, \gamma_0) \ge 97K_0$ and $\mathrm{diam}_{\gamma_1}(\gamma_0) \le K_0$, we have $d_{\gamma_1}(x_0, q) \ge 96K_0$ and $d_{\gamma_1}(x_0, p) \ge 95K_0$. In particular, $d_{\gamma_1}([x_0, g_{i(1)}x_0]) \ge 95K_0$, which implies $d_{\gamma_1}(x_0, g_{i(1)}x_0) \ge 94K_0$ by Corollary 3.10(3). Thus, γ_1 belongs to $\mathcal{C}_0$. Since $d_X(x_0, \gamma_1) < d_X(x_0, \gamma_0)$, this contradicts the minimality of γ_0.

We therefore conclude that $d_X(x_0, \gamma_0) \le d_X(x_0, \gamma_1)$, and η_0 appears earlier than η_1. Then $d_{\eta_0}(\eta_1, g_{i_2}x_0) \le 2K_0$ and $d_{\gamma_0}(g_{i(2)}x_0, \gamma_1) \le 3K_0$. Hence, $d_{\gamma_0}(x_0, \gamma_1) \ge 87K_0$.

We keep this manner. If $d_{\gamma_1}(x_0, g_j x_0) \ge 80K_0$ for all $j > i(2)$, then $(g_N x_0|y)_{x_0}$ lies in I_1 and not in I_2 by Observation 6.4, a contradiction. Hence, there is the first moment $i(3) > i(2)$ at which $d_{\gamma_1}(x_0, g_j x_0) \le 80K_0$. Then $d_{\gamma_1}(x_0, g_j x_0) =_{2K_0} 80K_0$ and Observation 6.4 again tells us that $(g_{i(3)}x_0, y)_{x_0} \in I_1$. Since $g_{i_3} \in P \notin B_1$, the collection

$$\mathcal{C}_2 := \{\gamma \in \mathscr{A}' : d_\gamma(x_0, g_{i(3)}x_0) > 100K_0\}$$

is nonempty. We pick $\gamma_2 \in \mathcal{C}_2$ that is the closest to x_0. If γ_2 is closer than γ_1 to x_0, then we can argue the same as before that $\gamma_2 \in \mathcal{C}_1$, violating the minimality of γ_1. It follows that γ_1 is closer to γ_2, and $d_{\gamma_1}(x_0, \gamma_2) \ge 90K_0$.

If this process does not halt, we obtain infinite sequence of step numbers $i(1) < i(2) < \ldots$ for the finite path P, a contradiction. Hence, the process must halt and the path P should intersect B_1. ☐

Similarly, for $g \in G$ such that $gx_0 \in \mathcal{H}_R(x_0, y)$, every d_S-path from id to g must pass through each of $B_1, B_2, \ldots, B_{R/100K_0}$.

It remains to show that $\sqcup_{i \ge 1} B_i$ is roughly branching. Since $EC(f)$ is a finite extension of a quasi-isometrically embedded subgroup $\langle f \rangle$, the set

$$EC(f) \cap \{g \in G : d_X(x_0, gx_0) < 200K_0\}$$

is finite. Hence, it is contained in $\{g : \|g\|_S \le R'\}$ for some R'. We claim that $\sqcup_{i \ge 1} B_i$ is $(R' + 4 + 200K_0\|f\|_S)$-roughly branching.

Let us take a subset B' of $\sqcup_{i \ge 1} B_i$ that is maximally $(R'+2)$-separated (in terms of the word metric d_S). We will construct a map $F : B' \to F(B') \subseteq G$.

Given $a \in B'$, Observation 6.2 guarantees $\mathscr{W}(a) \in \{w^{-1}, id, w\}$ such that

$$\big(x_0 \,\big|\, a\mathscr{W}(a)p\big)_{ax_0} < K_0, \quad \big(y \,\big|\, a\mathscr{W}(a)p\big)_{ax_0} < K_0 \quad \big(\forall p \in Ax(f)\big).$$

Furthermore, there exists $\mathscr{W} \in \{w^{-1}, id, w\}$ such that

$$\big(y \,\big|\, \mathscr{W}^{-1}p\big)_{x_0} < K_0. \quad \big(\forall p \in Ax(f)\big)$$

Then we define

$$F(a) := a\mathscr{W}(a)f^{200K_0}\mathscr{W}.$$

We now claim that:

Claim 6.5. *If $a_1, a_2, \ldots, a_k, b_1, \ldots, b_k \in B'$ are such that*

$$F(a_1)F(a_2)\cdots F(a_k) = F(b_1)F(b_2)\cdots F(b_k),$$

then $a_1 = b_1$.

Proof of Claim 6.5. Let $U = F(a_1)\cdots F(a_k)$. Note that

$$\big(\ [x_0, a_1x_0],\ a_1[x_0, \mathscr{W}(a_1)f^{200K_0}\mathscr{W}x_0],\ F(a_1)[x_0, a_2x_0],\ F(a_1)a_2[x_0, \mathscr{W}(a_2)f^{200K_0}\mathscr{W}x_0],\ \ldots\ \big)$$

is a sequence of consecutive geodesics, each longer than $50K_0$. (Recall that $(a_ix_0|y)_{x_0} \in I_1 \cup I_2 \cup \ldots$ is at least $75K_0$.) Next, between each pair of consecutive geodesics the Gromov product is bounded by $2.1K_0$. This is because

- $(x_0|F(a_i)\mathscr{W}^{-1}x_0)_{a_ix_0} < K_0$ and $d_X(F(a_i)\mathscr{W}^{-1}x_0, F(a_i)x_0) < K_0$.
- $\big(a_i\mathscr{W}(a_i)x_0 \,\big|\, F(a_i)y\big)_{F(a_i)x_0} = \big(\mathscr{W}^{-1}f^{-200K_0}x_0 \,\big|\, y\big)_{x_0} < K_0$ and $(y|a_{i+1}x_0)_{x_0} \geq 75K_0$, which imply $\big(a_i\mathscr{W}(a_i)x_0 \,\big|\, F(a_i)a_{i+1}x_0\big)_{F(a_i)x_0} < K_0 + 4\delta$.

  Moreover, $d_X(a_i\mathscr{W}(a_i)x_0, a_ix_0) < K_0$.

By the stability lemma, there exist points $p_1, q_1, \ldots, p_k, q_k$ on $[x_0, Ux_0]$, in order from closest to farthest from x_0, such that $d_X(p_1, a_1x_0), d_X(q_1, F(a_1)x_0), \ldots$ are all smaller than $2.2K_0$. Similarly, there exist points $p_1', q_1', \ldots, p_k', q_k'$ on $[x_0, Ux_0]$, in order, such that $d_X(p_1', b_1x_0), d_X(q_1', F(b_1)x_0), \ldots \leq 2.2K_0$.

Suppose to the contrary that $d_X(x_0, p_1) > d_X(x_0, p_1') + 130K_0$. Note that

$$d_X\big(p_1', b_1\mathscr{W}(b_1)x_0\big) \leq d_X(p_1', b_1x_0) + d_X\big(x_0, \mathscr{W}(b_1)x_0\big) \leq 3.2K_0,$$

and similarly q_1' and $b_1\mathscr{W}(b_1) \cdot f^{200K_0}x_0$ are $3.2K_0$-close. By Lemma 3.2, $[p_1', q_1']$ is $3.3K_0$-equivalent to $b_1\mathscr{W}(b_1)[x_0, f^{200K_0}x_0]$.

Let q be q_1' or p_1, whichever coming earlier along $[x_0, Ux_0]$. Then,

- p_1' and q are both closer to x_0 than p_1 is. Hence, $p_1', q \in [x_0, p_1]$.
- $q \in [p_1', q_1']$ is $3.3K_0$-close to $b_1\mathscr{W}(b_1)Ax(f)$, as well as p_1'. Meanwhile, we have $d_X(p_1', q_1') =_{6.6K_0} d_X(x_0, f^{200K_0}) = 200K_0$ and $d_X(p_1', p_1) > 130K_0$. Hence, p_1' and q are $130K_0$-distant points on $[x_0, p_1]$ that are $3.3K_0$-close to $b_1\mathscr{W}(b_1)Ax(f)$.

Now observe that $[x_0, a_1x_0]$ and $[x_0, p_1]$ are $2.3K_0$-equivalent, as a_1x_0 and p_1 are $2.2K_0$-close. Hence, $b_1\mathscr{W}(b_1)Ax(f)$ is $5.6K_0$-close to points on $[x_0, a_1x_0]$ that are at least $127K_0$-distant. This implies that $\mathrm{diam}_{b_1\mathscr{W}(b_1)Ax(f)}([x_0, a_1x_0]) >$

$115K_0$. Corollary 3.10(5) then tells us that $d_{b_1\mathscr{W}(b_1)Ax(f)}(x_0, a_1x_0) > 114K_0$. Recall our definition of B_i's. We are led to $d_{b_1\mathscr{W}(b_1)Ax(f)}(x_0, y) \ge 5K_0$.

Meanwhile, our definition of $\mathscr{W}(b_1)$ tells us that $(x_0|p)_{b_1\mathscr{W}(b_1)x_0} < 2K_0$ for every $p \in b_1\mathscr{W}(b_1)Ax(f)$. Lemma 3.8 implies that the projection of x_0 onto $b_1\mathscr{W}(b_1)Ax(f)$ is $(2K_0 + 8\delta)$-close to $b_1\mathscr{W}(b_1)x_0$. Similarly, because $(y|p)_{b_1\mathscr{W}(b_1)x_0} < 2K_0$ for every $p \in b_1\mathscr{W}(b_1)Ax(f)$, the projection $\pi_{b_1\mathscr{W}(b_1)Ax(f)}(y)$ should be $(2K_0 + 8\delta)$-close to $b_1\mathscr{W}(b_1)x_0$. In conclusion, $d_{b_1\mathscr{W}(b_1)Ax(f)}(x_0, y) < 4.5K_0$. This is a contradiction.

A similar contradiction happens if $d_X(x_0, p_1') > d_X(x_0, p_1) + 130K_0$. Hence, p_1 and p_1' are $130K_0$-close. Since q_1 (q_1', resp.) appears later than p_1 (p_1'. resp.) by at least $195K_0$, we conclude that $[p_1, q_1]$ and $[p_1', q_1']$ overlap for length at least $65K_0$.

Recall that $[p_1, q_1]$ ($[p_1', q_1']$, resp.) and $a_1\mathscr{W}(a_1)[x_0, f^{200K_0}x_0]$ ($b_1\mathscr{W}(b_1)[x_0, f^{200K_0}x_0]$, resp.) are $3.3K_0$-equivalent. By Corollary 3.10(1), (4), (5), we conclude

$$\mathrm{diam}_{a_1\mathscr{W}(a_1)Ax(f)}\big(b_1\mathscr{W}(b_1)Ax(f)\big) \ge 20K_0.$$

This implies that $\mathscr{W}(a_1)^{-1}a_1^{-1}b_1\mathscr{W}(b_1)$ lies in $EC(f)$. Meanwhile, note that $d_X(a_1\mathscr{W}(a_1)x_0, b_1\mathscr{W}(b_1)x_0) \le 140K_0$. Hence, $\mathscr{W}(a_1)^{-1}a_1^{-1}b_1\mathscr{W}(b_1)$ lies in $B_S(R')$, and $d_S(a_1, b_1) \le R' + 2$. Since a_1, b_1 are chosen from an $(R' + 2)$-separated set B', this forces $a_1 = b_1$. ☐

Now an inductive argument leads to:

Claim 6.6. *If $a_1, a_2, \ldots, a_k, b_1, \ldots, b_k \in B'$ are such that*

$$F(a_1)F(a_2)\cdots F(a_k) = F(b_1)F(b_2)\cdots F(b_k),$$

then $a_i = b_i$ for each $i = 1, \ldots, k$.

It remains to check that $\sqcup_i B_i$ is contained in a bounded neighborhood of $F(B')$. Given any $a \in \sqcup_i B_i$, it is $(R' + 2)$-close to some $a' \in B'$, as B' is a maximal $(R' + 2)$-separated subset of B. Now, $F(a')$ and a' are $(2 + 200K_0\|f\|_S)$-close. In summary, a is $(R + 4 + 200K_0\|f\|_S)$-close to $F(B')$ as desired. ☐

Combining Proposition 4.6 and Proposition 6.1, we conclude that relatively hyperbolic groups satisfy the assumption of Theorem 2.16 with

$$\mathscr{H}_D := \big\{\{g \in G : gx_0 \in \mathcal{H}_{100K_0D}(x_0, y)\} : y \in X\big\}$$

and $\mathcal{G}_{D,E} := \emptyset$ for each D, E. Therefore, Cayley graphs of relatively hyperbolic groups satisfy the assumption of Theorem 2.5, and $p_c < p_u$ and $\Delta_{p_c} < +\infty$ hold for such graphs.

7. Barriers in acylindrically hyperbolic group

Let G be an acylindrical hyperbolic group with a finite generating set S. Then G acts on a suitable δ-hyperbolic space (X, d_X) with a unital, axial WPD loxodromic element $f \in G$. Let $x_0 \in Ax(f)$. We fix these choices throughout the section.

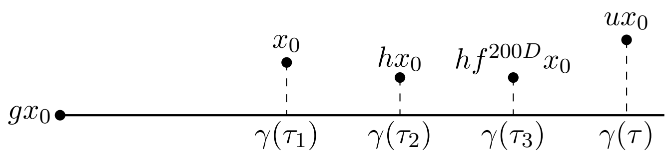

FIGURE 3. Schematics for $\mathfrak{A}^+_{D,E,f}(g)$.

The following is immediate from the δ-hyperbolicity.

Fact 7.1. *Let $i < j < k < l$ be integers, let $x \in \mathcal{N}_{j-i-2\delta}(f^i x_0)$ and let $y \in \mathcal{N}_{l-k-2\delta}(f^l x_0)$. Then there exists a subsegment $[x', y']$ of $[x, y]$ such that $x' \in \mathcal{N}_{2\delta}(f^j x_0)$ and $y' \in \mathcal{N}_{2\delta}(f^k x_0)$.*

For $D, E \geq 0$ and $u \in G$, we consider two versions of anti-halfspaces:

$$\mathfrak{A}^{\pm}_{D,E,f}(g) := \left\{ u \in G : \begin{array}{c} \exists \text{ geodesic } \gamma : [0,\tau] \to X, \, \exists 0 \leq \tau_1 \leq \tau_2 \leq \tau_3 \leq \tau, \, \exists h \in G \\ \text{such that } \|h\|_S \leq E, \gamma(0) = gx_0, \gamma(\tau_1) \in \mathcal{N}_{1.1D}(x_0), \\ \gamma(\tau_2) \in \mathcal{N}_{20\delta}(hx_0), \gamma(\tau_3) \in \mathcal{N}_{20\delta}(hf^{\pm 200D}x_0), \gamma(\tau) \in \mathcal{N}_{5D}(ux_0) \end{array} \right\},$$

$$\mathfrak{A}_{D,E,f}(g) := \mathfrak{A}^+_{D,E,f}(g) \cup \mathfrak{A}^-_{D,E,f}(g),$$

$$\mathfrak{B}^{\pm}_{D,E,f}(g) := \left\{ u \in G : \begin{array}{c} \exists \text{ geodesic } \gamma : [0,\tau] \to X, \, \exists 0 \leq \tau_1 \leq \tau_2 \leq \tau_3 \leq \tau, \, \exists h \in G \\ \text{such that } \|h\|_S \leq 2E, \gamma(0) = gx_0, \gamma(\tau_1) \in \mathcal{N}_{3D}(x_0), \\ \gamma(\tau_2) \in \mathcal{N}_{20\delta}(hx_0), \gamma(\tau_3) \in \mathcal{N}_{20\delta}(hf^{\pm 180D}x_0), \gamma(\tau) \in \mathcal{N}_{5D}(ux_0) \end{array} \right\},$$

$$\mathfrak{B}_{D,E,f}(y) := \mathfrak{B}^+_{D,E,f}(g) \cup \mathfrak{B}^-_{D,E,f}(g).$$

Some observations are in order.

Observation 7.2. *Let $D \geq 1000(\delta + 1)$ and $E > 10D\|f\|_S$. Then for each $g \in G$ we have*

$$\mathfrak{A}^{\pm}_{D,E,f}(g) \subseteq \mathfrak{B}^{\pm}_{D,E,f}(g). \quad (\forall g \in G)$$

Proof. Let $u \in \mathfrak{A}^{\pm}_{D,E,f}(g)$. Let $\gamma : [0, \tau] \to X$, $0 \leq \tau_1 \leq \tau_2 \leq \tau_3 \leq \tau$ and $h \in G$ be the ingredients for the membership. In particular, $\gamma(\tau_2)$ is 20δ-close to hx_0 and $\gamma(\tau_3)$ is 20δ-close to $hf^{\pm 200D}x_0$. By Fact 7.1, there exist $\tau_2 \leq \tau_2' \leq \tau_3' \leq \tau_3$ such that

$$d_X\big(hf^{\pm 10D}x_0, \gamma(\tau_2')\big) \leq 2\delta, \; d_X\big(hf^{\pm 190D}x_0, \gamma(\tau_3')\big) \leq 2\delta.$$

Furthermore, $\|hf^{\pm 10D}\|_S \leq \|h\|_S + 10D\|f\|_S \leq 2E$ holds. It is now clear that $u \in \mathfrak{B}^{\pm}_{D,E,f}(g)$. □

In the definition of $\mathfrak{B}^{\pm}_{D,E,f}(g)$ we have $d_X(\gamma(\tau_2), \gamma(\tau_3)) \geq 180D - 2 \cdot 20\delta$, whereas $d(x_0, \gamma(\tau_1)), d_X(ux_0, \gamma(\tau)) \leq 5D$. This leads to:

Observation 7.3. *Let $D \geq 1000(\delta + 1)$ and let $u \in \mathfrak{B}_{D,E,f}(g)$. Then $d_X(x_0, ux_0)$ and $d_X(gx_0, ux_0)$ are at least $100D$. In particular, $id, g \notin \mathfrak{B}_{D,E,f}(g)$. Moreover, we have $d_X(gx_0, ux_0) \geq d_X(gx_0, x_0) + 100D$.*

Recall that $EC(f)$ is a finite extension of a cyclic subgroup $\langle f \rangle$, which is quasi-isometrically embedded in X. Hence, the set

$$EC(f) \cap \{g : d_X(x_0, gx_0) \le 500D + 2E + 20\delta\}$$

is finite. The following observation tells us that "lineage is linear".

Observation 7.4. *For each large enough $D > 0$ and for each $E > 0$ there exists $R > 0$ such that the following holds.*

Let $u, v \in G$ and suppose that there exists $w \in u\mathfrak{A}_{D,E,f}(u^{-1}) \cap v\mathfrak{B}_{D,E,f}(v^{-1})$. Then one of the following holds.

(1) $v \in u\mathfrak{A}_{D,E,f}(u^{-1})$ and $d_X(wx_0, vx_0) < d_X(wx_0, ux_0)$;
(2) $u \in v\mathfrak{B}_{D,E,f}(v^{-1})$ and $d_X(wx_0, ux_0) < d_X(wx_0, vx_0)$, or
(3) $d_S(u, v) \ge R$.

Proof. Let $K_0 = K$ be the constant as in Theorem 3.13. Furthermore, let $D_0 = D_0$ be as in Lemma 3.16. We assume that $D > 1000(\delta + K_0 + D_0)$.

Furthermore, let $R' = R(200D, 300D\|f\|_S + E)$ be as in Lemma 3.16, and let $R = R' + E$.

By the assumption, there exist a geodesic $\gamma : [0, \tau] \to X$, $0 \le \tau_1 \le \tau_2 \le \tau_3 \le \tau$, a sign $\epsilon \in \{+1, -1\}$ and an element $h \in G$ with $\|h\|_S \le E$ such that

$$\gamma(0) = x_0,\ d_X\big(\gamma(\tau_1), ux_0\big) < 1.1D,\ d_X\big(\gamma(\tau_2), uhx_0\big) < 20\delta,$$
$$d_X\big(\gamma(\tau_3), uhf^{200\epsilon D}x_0\big) < 20\delta,\ d_X\big(\gamma(\tau), wx_0\big) < 5D.$$

Similarly, there exist a geodesic $\gamma' : [0, \sigma] \to X$, $0 \le \sigma_1 \le \sigma_2 \le \sigma_3 \le \sigma$, $\epsilon' \in \{+1, -1\}$ and $h' \in G$ with $\|h'\|_S \le 2E$ such that

$$\gamma'(0) = x_0,\ d_X\big(\gamma'(\sigma_1), vx_0\big) < 3D,\ d_X\big(\gamma'(\sigma_2), vh'x_0\big) < 20\delta,$$
$$d_X\big(\gamma'(\sigma_3), vh'f^{180\epsilon' D}x_0\big) < 20\delta,\ d_X\big(\gamma'(\sigma), wx_0\big) < 5D.$$

We define $L_{min} := (\gamma(\tau)|\gamma'(\sigma))_{x_0}$. Then we have

$$L_{min} = \tau - (x_0|\gamma'(\sigma))_{\gamma(\tau)} \ge \tau - d_X(\gamma(\tau), \gamma'(\sigma)) \ge \tau - 10D.$$

Similarly, $L_{min} \ge \tau - 10D$. Similarly $L_{min} \ge \sigma - 10D$. By Lemma 3.5, $\gamma(t)$ and $\gamma'(t)$ are 4δ-close for $0 \le t \le L_{min}$.

We claim that if $\tau_2 \le \sigma_2$ then either (1) or (3) holds, and if $\tau_2 \ge \sigma_2$ then either (2) or (3) holds. Since the latter case follows from a similar argument, we only explain the former one.

If $\tau_3 \le \sigma_1 + 1.5D$ in addition, then we have

$$\begin{aligned} d_X\big(\gamma(\sigma_1 + 1.5D), vx_0\big) &\le d_X\big(\gamma(\sigma_1 + 1.5D), \gamma(\sigma_1)\big) + d_X(\gamma(\sigma_1), \gamma'(\sigma_1)) + d_X\big(\gamma'(\sigma_1), vx_0\big) \\ &\le 1.5D + 4\delta + 3D < 4.6D. \end{aligned}$$

Here, we can feed the parameter $\sigma_1 + 1.5D$ in $\gamma(\cdot)$ because $\sigma_1 \le \sigma_1 + 1.5D \le \sigma - 150D \le L_{min} \le \tau$. Since we have the geodesic γ with timing $\tau_1 \le \tau_2 \le \tau_3 \le \sigma_1 + 1.5D$, we conclude that $vx_0 \in u\mathfrak{A}_{D,E,f}(u^{-1})$. Furthermore, since

$\tau_3 \geq \tau_2 + 199D$ we have

$$d_X(wx_0, vx_0) =_{5D} \tau - (\sigma_1 + 1.5D) \leq \tau - \tau_3$$
$$\leq \tau - \tau_2 + 199D \leq \tau - \tau_1 + 199D =_D d_X(wx_0, ux_0) + 199D.$$

If $\tau_3 \geq \sigma_1 + 1.5D$, then we claim that the intersection I of $[\sigma_1, \sigma_3]$ and $[\tau_2, \tau_3]$ is large. Indeed, there are three cases:

- First note that $[\sigma_1, \sigma_3]$ and $[\tau_2, \tau_3]$ are both $100D$-long. Hence, if one includes the other one, the intersection I must be $100D$-long.
- If $\sigma_1 \leq \tau_2$ and $\sigma_3 \leq \tau_3$, then I is $100D$-long as
$$\sigma_3 \geq \sigma_2 + 100D \geq \tau_2 + 100D.$$
- If $\sigma_1 \geq \tau_2$ and $\sigma_3 \geq \tau_3$, then I is at least $1.5D$-long by the assumption $\sigma_1 \leq \tau_3 - 1.5D$.

All in all, we have $\mathrm{diam}(I) \geq 1.5D$. In other words, the projections of $\gamma(\sigma_1)$ and $\gamma(\sigma_3)$ onto $\gamma([\tau_2, \tau_3])$ is at least $1.5D$-distant. Note that $uh[x_0, f^{200\epsilon D} x_0]$ and $\gamma([\tau_2, \tau_3])$ are 22δ-equivalent by Lemma 3.2. Hence, Corollary 3.10(4) tells us that $d_{uh[x_0, f^{200\epsilon D}x_0]}(\gamma(\sigma_1), \gamma(\sigma_3)) \geq 1.4D$. Moreover, note that

$$d_X(\gamma(\sigma_1), vx_0) \leq d_X(\gamma(\sigma_1), \gamma'(\sigma_1)) + d_X(\gamma'(\sigma_1), vx_0) \leq 4\delta + 1.1D.$$

Similarly, $\gamma(\sigma_3)$ and $vh' f^{180\epsilon' D} x_0$ are 24δ-close. By Corollary 3.10(1) we conclude

$$d_{uh[x_0, f^{200\epsilon D}x_0]}(vx_0, vh' f^{180\epsilon' D} x_0) > 0.1D > K_0.$$

Note that $\|h' f^{180\epsilon' D}\|_S \leq E + 180D\|f\|_S$. Our choice of constant R' based on Lemma 3.16 guarantees $d_S(uh, v) \leq R'$. Hence, $d_S(u, v) \leq R' + E = R$. $\square$

We now need

Observation 7.5. *Let D be large enough and let $E > 10\|f\|_S D$. Let $u, v, w \in G$ such that $w \in \mathfrak{A}^{\pm}_{D,E,f}(u)$ and $(vx_0|wx_0)_{x_0} < 2D$. Then $w \in \mathfrak{B}^{\pm}_{D,E,f}(v)$.*

Proof. Let $\gamma : [0, \tau] \to X$ be a geodesic starting at ux_0 and let $h \in G$ be the ones that realize the membership $w \in \mathfrak{A}^{\pm}_{D,E,f}(u)$. In particular, there are timing $\tau_1 \leq \tau_2 \leq \tau_3 \leq \tau$ such that

$$d_X(\gamma(\tau_1), x_0) < 1.1D, \ d_X(\gamma(\tau_2), hx_0) < 20\delta,$$
$$d_X\big(\gamma(\tau_3), hf^{\pm 200D} x_0\big) < 20\delta, \ d_X(\gamma(\tau), wx_0) < 5D.$$

Let us draw a geodesic $\eta : [0, L'] \to X$ that connects vx_0 to wx_0. Since we are assuming $(vx_0|wx_0)_{x_0} < 2D$, there exists τ_1^* such that $\eta(\tau_1^*)$ and x_0 are $(2D + 20\delta)$-close. Now $\eta([\tau_1^*, L'])$ and $\gamma([\tau_1, \tau_3])$ are $5.1D$-equivalent by Lemma 3.2. Namely, there exist $\tau_1^* \leq t_2 \leq t_3 \leq L'$ such that $\eta(t_2)$ and $\eta(t_3)$ are $5.5D$-close to hx_0 and $hf^{200\epsilon D} x_0$, respectively.

Now Fact 7.1 gives timing $t_2 \leq \tau_2^* \leq \tau_3^* \leq t_3$ such that $\eta(\tau_2^*)$ and $\eta(\tau_3^*)$ are 2δ-close to $hf^{10\epsilon D} x_0$ and $hf^{190\epsilon D} x_0$, respectively. The geodesic η together with $\tau_1^* \leq \tau_2^* \leq \tau_3^* \leq L$ show that $v \in u\mathfrak{B}_{D,E,f}(g)$, as $\|hf^{10\epsilon D}\|_S \leq E + 10\|f\|_S D \leq 2E$. $\square$

Now for each $u \in G$ we define

$$\mathcal{H}_{D,E,f}(u) := \Big\{ g \in G : \big(gx_0 \big| ux_0\big)_{x_0} > D \text{ or } \nexists h \in G\Big[[\|h\|_S \le E] \wedge [d_{hAx(f)}(x_0, gx_0) \ge 250D] \Big] \Big\}.$$

We call it an f-halfspace radius parameters (D, E). This is related to anti-halfspaces $\mathfrak{A}_{D,E,f}$ because:

Lemma 7.6. *For each $D > 1000(\delta+1)$ and $E \ge 0$ there exists $F > E$ such that*

$$\big[g \notin \mathcal{H}_{D,E,f}(u)\big] \Rightarrow \big[g \in \mathfrak{A}_{D,F,f}(u)\big] \quad (\forall g, u \in G).$$

Proof. Let $K_0 := \max_{s\in S} d_X(x_0, sx_0) + \|f\|_S$. We claim that $F := E + (2K_0E + 2D + 1)K_0$ works. To see this, let $g \notin \mathcal{H}_{D,E,f}(u)$. Then there exists $h \in G$ such that $\|h\|_S \le E$ and $d_{hAx(f)}(x_0, gx_0) \ge 250D$. Now let $\gamma : [0, L] \to X$ be the geodesic connecting ux_0 to gx_0. Since $(gx_0|ux_0)_{x_0} \le D$, there exists $\tau_1 \in [0, L]$ such that $\gamma(\tau_1)$ is $1.1D$-close to x_0. Then by the coarse Lipschitzness of $\pi_{hAx(f)}(\cdot)$, we have

$$d_{hAx(f)}\big(\gamma(\tau_1), \gamma(L)\big) > d_{hAx(f)}(x_0, gx_0) - (1.1D + 12\delta) > 248D. \tag{7.1}$$

Let $i, j \in \mathbb{Z}$ be such that $\pi_{hAx(f)}(\gamma(\tau_1))$ intersects $[hf^i x_0, hf^{i+1}x_0]$ and $\pi_{hAx(f)}(\gamma(L))$ intersects $[hf^j x_0, hf^{j+1}x_0]$. Then either $j > i + 247D$ or $j < i - 247D$ due to Inequality 7.1. We will focus on the former case; the latter case can be handled in a similar way. In this case, Corollary 3.10 tells us that there exist $\tau_1 \le \tau_2 \le \tau_3 \le L$ such that $\gamma(\tau_2)$ is 12δ-close to $hf^{i+1}x_0$ and $\gamma(\tau_3)$ is 12δ-close to $hf^{i+200D+1}x_0$.

Recall that $d_X(\gamma(\tau_1), hAx(f)) \le d_X(\gamma(\tau_1), x_0) + d_X(x_0, hx_0)$. Using this, we observe that

$$\begin{aligned} d_X(hx_0, hf^{i+1}x_0) &\le d_X(hx_0, \pi_{hAx(f)}(\gamma(\tau_1)) + 1 \\ &\le d_X\big(hx_0, \gamma(\tau_1)\big) + d_X\big(\gamma(\tau_1), hAx(f)\big) \\ &\le 2d_X(x_0, hx_0) + 2D + 1 \le 2K_0E + 2D + 1. \end{aligned}$$

This means $|i+1| < 2K_0E + 2D + 1$ and $\|hf^{i+1}\|_S \le \|h\|_S + (2K_0E + 2D + 1)\|f\|_S \le F$.

All in all, our choice of timing $\tau_1 \le \tau_2 \le \tau_3 \le L$, together with $hf^{i+1} \in G$ with $\|hf^{i+1}\|_S \le F$, guarantees that $gx_0 \in \mathfrak{A}_{D,F,f}(u)$ as desired. □

We can now state:

Proposition 7.7. *Let X be a δ-hyperbolic space and let G be a non-virtually cyclic group acting on X with an axial, unital WPD loxodromic element f. Let S be a finite generating set of G. Let $x_0 \in Ax(f)$.*

Then for each $\epsilon > 0$ and for each large $D > 0$ and $E, E' \ge 0$ there exists a constant $N = N(\epsilon, D, E, E')$ such that for every finite set $A \subseteq G$ there exists a subset $A' \subseteq A$ satisfying:

(1) $\#A' \ge (1-\epsilon)\#A$;

(2) *For each $a \in A'$ there exist f-halfspaces $\mathcal{H}_1, \mathcal{H}_2 \subseteq G$ with radius parameters (D, E) such that*

$$\#\Big(A \setminus a \cdot \big(\mathcal{H}_1 \cup \mathcal{H}_2 \setminus \{g \in G : \|g\|_S \le E'\}\big)\Big) \le N.$$

Proof. Note that $\{g \in G : \|g\|_S \le E'\}$ have elements at most $(2\#S)^{E'}$. Hence, the statement for general E' will follow once we prove it for $E' = 0$. For this reason we set $E' = 0$. Let D be large enough that Observation 7.3, 7.4 and 7.5 apply, and let $E \ge 10\|f\|_S D$.

Let $F = F(D, E)$ be as in Lemma 7.6 and let $R_0 = R$ be as in Observation 7.4 for (D, F). We claim that

$$N = N(\epsilon, D, E) := \frac{2 \cdot (2\#S)^R}{\epsilon}$$

works.

Let us begin the proof by collecting problematic elements, i.e.,

$$\mathcal{A}_1 := A \backslash A' = \left\{ a \in A : \begin{array}{c} \#\big(A \setminus a(\mathcal{H}_1 \cup \mathcal{H}_2)\big) \ge N \text{ for every } f\text{-halfspaces} \\ \mathcal{H}_1, \mathcal{H}_2 \text{ with radius parameters } (D, E) \end{array} \right\}.$$

Let $\mathcal{A}_2$ be a maximally R-separated subset $\mathcal{A}_2$ of $\mathcal{A}_1$, i.e., we have

(1) $d_S(a, a') \ge R$ for each pair of distinct elements $a, a' \in \mathcal{A}_2$;
(2) $\mathcal{A}_2$ is a maximal subset of $\mathcal{A}_1$ satisfying this property.

Then $\bigcup_{a \in \mathcal{A}_2} a \cdot \{g \in G : \|g\|_S \le R\}$ covers entire $\mathcal{A}_1$. Hence, we have

$$\#\mathcal{A}_2 \ge \frac{1}{(2\#S)^R} \cdot \#\mathcal{A}_1.$$

As before, we first prepare empty collections $\mathcal{B} = \mathcal{U} = \mathcal{G} = \emptyset$. Enumerate $\mathcal{A}_2$ by the distance from x_0, i.e., let $\mathcal{A}_2 = \{a_1, a_2, \ldots, a_{\#\mathcal{A}_2}\}$ be such that $d_X(x_0, a_i) \le d_X(x_0, a_{i+1})$ for each i.

We now describe a procedure that takes place throughout $\#\mathcal{A}_2$ steps. At step i, we first declare $\mathfrak{A}_i := a_i \mathfrak{A}_{D,F,f}(a_i^{-1})$.

(1) If $\mathcal{A}_2 \cap \mathfrak{A}_i$ has no element, then we declare that $a_i \in \mathcal{G}$ and $b_i := id$.
(2) If not, pick $b_i \in \mathcal{A}_2 \cap \mathfrak{A}_i$ whose orbit point $b_i x_0$ is the *closest* to x_0. We then declare $\mathfrak{A}'_i := a_i \mathfrak{A}_{D,F,f}(a_i^{-1} b_i)$.
   (a) If $\mathcal{A}_2 \cap \mathfrak{A}_i \cap \mathfrak{A}'_i$ has no element, then we declare that $a_i \in \mathcal{G}$.
   (b) If not, we pick $c_i \in \mathcal{A}_2 \cap \mathfrak{A}_i \cap \mathfrak{A}'_i$ whose orbit point $c_i x_0$ is the *closest* to x_0. We then declare $a_i \in \mathcal{B}$ and $b_i, c_i \in \mathcal{U}$.

(If an element in $\mathcal{U}$ is declared good or bad, it is not undecided anymore; we remove it from $\mathcal{U}$.)

Till step i, $\mathcal{G} \cup \mathcal{B}$ comprises of elements from $\{a_1, \ldots, a_i\}$; they do not contain any of $a_{i+1}, a_{i+2}, \ldots$. $(*)$ Let us now observe what happens at step i.

In case (1), $\mathcal{G}$ gains one more element that might be from $\mathcal{U}$ or not. $\mathcal{B}$ does not change. Overall, $\#\mathcal{B}$ stays the same and $\#\mathcal{U} + \#\mathcal{G}$ does not decrease. Similar situation happens in Case (2-a).

In case (2-b), $\mathcal{B}$ gains one element a_i, which might be from $\mathcal{U}$. In exchange, $\mathcal{U}$ gains elements b_i and c_i. Here Observation 7.3 guarantees that $d_X(x_0, b_i x_0), d_X(x_0, c_i x_0) > d_X(x_0, a_i x_0)$. Since $\mathcal{A}_2$ was labelled with respect to the distance from x_0, we conclude that $b_i, c_i \in \{a_{i+1}, a_{i+2}, \ldots\}$; in other words, neither b_i nor c_i come from $\mathcal{G} \cup \mathcal{B}$. We thus confirm that elements are never re-classified once they are put in $\mathcal{G} \cup \mathcal{B}$.

Furthermore, Observation 7.3 guarantees that $d_X(b_i x_0, c_i x_0) > 100D$. Hence b_i and c_i are distinct elements. If b_i, c_i are genuinely new addition to $\mathcal{U}$ and are not re-used from $\mathcal{U}$ at step $i-1$, then we can conclude that $\#\mathcal{U}$ increases at least by 1 in Case (2-b). It remains to show

Claim 7.8. *For $i < j$ such that $a_i, a_j \in \mathcal{B}$, we have $\{b_i, c_i\} \cap \{b_j, c_j\} = \emptyset$.*

Proof of Claim 7.8. Suppose to the contrary that $b_i \in \{b_j, c_j\}$. That means

$$b_i \in a_i \mathfrak{A}_{D,F,f}(a_i^{-1}) \cap a_j \mathfrak{A}_{D,F,f}(a_j^{-1}).$$

Here, recall that $d_X(x_0, a_i x_0) \le d_X(x_0, a_j x_0)$ and $d_S(a_i, a_j) > R$. Observation 7.4 tells us that $a_j \in a_i \mathfrak{A}_{D,F,f}(a_i^{-1})$. (In Observation 7.4, Case 2 cannot happen because of Observation 7.3, and Case 3 cannot happen for $d_S(a_i, a_j) > R$.) Here, note that $d_X(x_0, a_j x_0) < d_X(x_0, b_j x_0)$ because of Observation 7.3. This contradicts the minimality of b_i.

Next, suppose to the contrary that $c_i \in \{b_j, c_j\}$. That means

$$c_i \in a_i \mathfrak{A}_{D,F,f}(a_i^{-1}) \cap a_i \mathfrak{A}_{D,F,f}(a_i^{-1} b_i) \cap a_j \mathfrak{A}_{D,F,f}(a_j^{-1}).$$

For the same reason as above, we have $a_j \in a_i \mathfrak{A}_{D,F,f}(a_i^{-1})$ and $d_X(c_i x_0, a_j x_0) < d_X(c_i x_0, a_i x_0)$.

We then have

$$\begin{aligned}
d_X(x_0, a_j x_0) &\ge d_X(x_0, a_i x_0) + 100D, && (\because \text{Observation 7.3}) \\
d_X(b_i x_0, a_j x_0) &\ge d_X(b_i x_0, c_i x_0) - d_X(a_j x_0, c_i x_0) \\
&\ge d_X(b_i x_0, c_i x_0) - d_X(c_i x_0, a_i x_0) \\
&=_{2D} d_X(b_i x_0, a_i x_0) && \big(\because c_i \in a_i \mathfrak{A}_{D,F,f}(a_i^{-1} b_i)\big).
\end{aligned}$$

This implies that $(x_0 | b_i x_0)_{a_j x_0} \ge 98D$. Meanwhile, $(x_0 | c_i x_0)_{a_j x_0} \le D$ as $c_i \in a_j \mathfrak{A}_{D,F,f}(a_j^{-1})$. By Lemma 3.6, we have $(b_i x_0 | c_i x_0)_{a_j x_0} < 2D$. Combining this with $c_i \in a_j \mathfrak{A}_{D,F,f}(a_j^{-1})$, we can apply Observation 7.5 to conclude that $c_i \in a_j \mathfrak{B}_{D,F,f}(a_j^{-1} b_i)$.

We thus have $c_i \in a_i \mathfrak{A}_{D,E,f}(a_i^{-1}, b_i) \cap a_j \mathfrak{A}_{D,E,f}(a_j^{-1} b_i)$. By Observation 7.4, either:

(1) $a_j \in a_i \mathfrak{A}_{D,F,f}(a_i^{-1} b_i)$ and $d_X(c_i x_0, a_j x_0) < d_X(c_i x_0, a_i x_0)$, or
(2) $a_i \in a_j \mathfrak{B}_{D,F,f}(a_j^{-1} b_i)$ and $d_X(c_i x_0, a_j x_0) > d_X(c_i x_0, a_i x_0)$.

(Again, $d_S(a_i, a_j) < R$ is ruled out.) Since we already know $d_X(c_i x_0, a_j x_0) < d_X(c_i x_0, a_i x_0)$, the former case happens.

Hence $a_j \in a_i \mathfrak{A}_{D,F,f}(a_i^{-1}) \cap a_i \mathfrak{A}_{D,F,f}(a_i^{-1} b_i)$ with $d_X(x_0, a_j x_0) < d_X(x_0, c_i x_0) - 100D$, as $c_i \in a_j \mathfrak{A}_{D,F,f}(a_j^{-1})$. This contradicts the minimality of c_i. □

Thanks to the claim, we conclude that $\#\mathcal{B} \le \#\mathcal{U} + \#\mathcal{G}$ at each step. But recall also that $a_i \in \mathcal{A}_2$ is declared good or bad at step i and is not affected thereafter. Hence, after the last step, there is no element of $\mathcal{U}$ left. Hence, we have $\#\mathcal{B} \le \#\mathcal{G}$, and $\mathcal{G}$ takes up at least half of $\mathcal{A}_2$.

Now, with the final $\mathcal{G}$ in hand, for each $i \in \{1, \ldots, \#\mathcal{A}_2\}$ such that $a_i \in \mathcal{G}$, we define

$$K_i := A \setminus a_i\big(\mathcal{H}_{D,E,f}(a_i^{-1}) \cup \mathcal{H}_{D,E,f}(a_i^{-1}b_i)\big).$$

Since $a_i \in \mathcal{G} \subseteq \mathcal{A}_2$, we have $\#K_i \ge N$. The remaining claim is:

Claim 7.9. *For every pair of distinct elements $a_i, a_j \in \mathcal{G}$, K_i and K_j do not intersect.*

To check this claim, suppose to the contrary that K_i and K_j have a common element w for some $i < j$ such that $a_i, a_j \in \mathcal{G}$. By Lemma 7.6, we have

$$w \in a_i\mathfrak{A}_{D,F,f}(a_i^{-1}) \cap a_i\mathfrak{A}_{D,F,f}(a_i^{-1}b_i) \cap a_j\mathfrak{A}_{D,F,f}(a_j^{-1}) \cap a_j\mathfrak{A}_{D,F,f}(a_j^{-1}b_j).$$

Depending on whether $d_X(wx_0, a_jx_0) \le d_X(wx_0, a_ix_0)$ or not, we have $a_j \in a_i\mathfrak{A}_{D,F,f}(a_i^{-1}) \cap a_i\mathfrak{A}_{D,F,f}(a_i^{-1}b_i)$ or $a_i \in a_j(\mathfrak{A}_{D,F,f}(a_j^{-1}) \cap a_j\mathfrak{A}_{D,F,f}(a_j^{-1}b_j)$ by Observation 7.3. This contradicts the goodness of a_i or a_j. Hence, such a common element w cannot exist and K_i and K_j are disjoint.

With Claim 4.5 in hand, we have

$$\begin{aligned} \#A &\ge \sum_{i:a_i \in \mathcal{G}} \#(K_i \cap A) \ge N \cdot \#\mathcal{G} \ge N \cdot \frac{\#\mathcal{A}_2}{2} \\ &\ge N \cdot \frac{\#\mathcal{A}_1}{2 \cdot (2\#S)^R} \ge \frac{1}{\epsilon}(\#A - \#A'). \quad \square \end{aligned}$$

We now have to check the branching property. Recall that $\mathscr{A}$ is the collection of all translates of $Ax(f)$. Let

$$\begin{aligned} \mathcal{NF}_D &:= \big\{g \in G : \forall \gamma \in \mathscr{A}[d_\gamma(x_0, gx_0) < D]\big\}, \\ \mathcal{NF}_D^{\ge i} &:= \mathcal{NF}_D \cap \{g : d_S(id, g) \ge i\}. \end{aligned}$$

We then observe that:

Proposition 7.10. *For each D, $\mathcal{NF}_D$ is r-roughly branching for some r.*

Proof. Since G is non-virtually cyclic, there exists $w \in S$ such that $Ax(f)$ and $wAx(f)$ have K_0-bounded projection onto each other. This guarantees a constant K_1 such that the following holds. For each $g \in \mathcal{NF}_D$, there exists $\mathscr{W}(g) \in \{id, w\}$ such that

$$\operatorname{diam}_{g^{-1}Ax(f)} \big(\mathscr{W}(g) \cdot Ax(f)\big) \le K_1.$$

By increasing K_1 if necessary, we can also guarantee that $d_X(x_0, wx_0) < K_1$.

We will prove the proposition for $D > 10^4(\delta + K_1 + 1)$. For each $g \in G$ we define $F(g) := g\mathscr{W}(g)f^{50D}$.

Recall that the set

$$EC(f) \cap \{g \in G : d_X(x_0, gx_0) < 100D\}$$

is a finite set. Namely, it is contained in $\{g \in G : d_S(id, g) < R'\}$ for some R'. Let $R = R' + 2$.

We now consider a subset A of $\mathcal{NF}_D$ that is maximally R-separated in the word metric d_S. Let $a_1, \ldots, a_k, b_1, \ldots, b_k \in A$ be such that

$$F(a_1)F(a_2)\cdots F(a_k) = F(b_1)F(b_2)\cdots F(b_k).$$

We then claim $a_1 = b_1$. To see this, let us define

$$\begin{aligned} p_i &:= F(a_1)\cdots F(a_i)x_0 & (i = 0, \ldots, k), \\ q_i &:= F(a_1)\cdots F(a_{i-1}) \cdot a_i \mathscr{W}(a_i)x_0 & (i = 1, \ldots, k). \end{aligned}$$

We claim that:

(1) $d_X(q_{i-1}, p_i) = 50D$, q_{i-1} is $1.1D$-close to $[p_{i-1}, p_i]$ and $d_X(p_{i-1}, p_i) > 48D$ for $i = 1, \ldots, k$;
(2) $(q_{i-1}|p_{i+1})_{p_i}, (p_{i-1}|p_{i+1})_{p_i} \le 2.2D$ for $i = 1, \ldots, k-1$.

Recall that $a_i \in \mathcal{NF}_D$. Hence, we have

$$d_{a_i \mathscr{W}(a_i) Ax(f)}(x_0, a_i x_0) \le D.$$

By the coarse Lipschitzness of the projection (Corollary 3.10(1)), we also have

$$d_{a_i \mathscr{W}(a_i) Ax(f)}\big(a_i x_0, a_i \mathscr{W}(a_i)x_0\big) \le 0.001D.$$

In summary, we have $d_{a_i \mathscr{W}(a_i) Ax(f)}(x_0, a_i \mathscr{W}(a_i)x_0) \le 1.001D$. By Corollary 3.10(5), we then have

$$d_{a_i \mathscr{W}(a_i)[x_0, f^{50D}x_0]}(x_0, a_i \mathscr{W}(a_i)x_0) \le 1.01D.$$

By Lemma 3.8, we conclude $(x_0 | a_i \mathscr{W}(a_i) f^{50D} x_0)_{a_i \mathscr{W}(a_i)x_0} \le 1.015D$. Now Lemma 3.5 tells us that $a_i \mathscr{W}(a_i)x_0$ is $1.1D$-close to $[x_0, a_i \mathscr{W}(a_i) f^{50D} x_0]$. This also implies

$$d_X\big(x_0, a_i \mathscr{W}(a_i) f^{50D} x_0\big) \ge d_X\big(a_i \mathscr{W}(a_i)x_0, a_i \mathscr{W}(a_i) f^{50D} x_0\big) - 1.1D \ge 50D - 1.1D \ge 48D.$$

Hence, we conclude Item (1).

We now observe that

$$\begin{aligned} d_{Ax(f)}(x_0, a_{i+1}x_0) &\le D, \\ d_{Ax(f)}(a_{i+1}x_0, a_{i+1}\mathscr{W}(a_{i+1})x_0) &\le 0.001D, \\ d_{Ax(f)}\big(a_{i+1}\mathscr{W}(a_{i+1})x_0, a_{i+1}\mathscr{W}(a_{i+1})f^{50D}x_0\big) &\le 0.001D. \end{aligned}$$

The first inequality is due to the membership $a_{i+1} \in \mathcal{NF}_D$. The second inequality is by Corollary 3.10(1). The third inequality is the requirement for $\mathscr{W}(a_{i+1})$. Combined with Corollary 3.10(5), these imply

$$d_{Ax(f)}\big(x_0, a_{i+1}\mathscr{W}(a_{i+1})f^{50D}x_0\big) \le 1.002D,\ d_{[f^{-50D}x_0, x_0]}\big(x_0, a_{i+1}\mathscr{W}(a_{i+1})f^{50D}x_0\big) \le 1.003D.$$

All in all, we have

$$\big(f^{-50D}x_0 \,\big|\, a_{i+1}\mathscr{W}(a_{i+1})f^{50D}x_0\big)_{x_0} \le 1.01D,$$

i.e. $(q_{i-1}|p_{i+1})_{q_i} \le 2D$. Meanwhile, by Item (1) we have

$$\begin{aligned}\big(f^{-50D}x_0 \,\big|\, f^{-50D}\mathscr{W}(a_i)^{-1}a_i^{-1}x_0\big)_{x_0} &= \big(a_i\mathscr{W}(a_i)x_0 \,\big|\, x_0\big)_{F(a_i)x_0} \\ &=_{4\delta} d_X(x_0, f^{-50D}x_0) - d_X(a_i\mathscr{W}\big(a_i)x_0, [x_0, F(a_i)x_0]\big) \\ &\ge 50D - 1.1D \ge 48D.\end{aligned}$$

Now Gromov's 4-point inequality implies that

$$\big(f^{-50D}\mathscr{W}(a_i)^{-1}a_i^{-1}x_0 \,\big|\, a_{i+1}\mathscr{W}(a_{i+1})f^{50D}x_0\big)_{x_0} \le 1.1D,$$

i.e. $(p_{i-1}|p_{i+1})_{p_i} \le 2D$. This leads to Item (2).

We can now apply Lemma 3.4 to the sequence

$$\big(x_0,\, q_1,\, p_1,\, p_2,\, \ldots,\, p_k\big).$$

Let $\gamma : [0, L] \to X$ be the geodesic connecting x_0 to $F(a_1)\cdots F(a_k)x_0$. By Lemma 3.4 there exist $\tau \le \tau'$ such that $d_X(\gamma(\tau), q_1), d_X(\gamma(\tau'), p_1) \le 2.25D$. Note that $\tau' \ge \tau + 40D$.

For the exactly same reason, there exist $\sigma \le \sigma'$ such that $\gamma(\sigma)$, $\gamma(\sigma')$ are $2.25D$-close to $b_1\mathscr{W}(b_1)x_0$ and $F(b_1)x_0$, respectively.

Suppose without loss of generality that $\tau \le \sigma$. There are two cases.

(1) If $\sigma \ge \tau + 25D$, then we have

$$d_{\gamma([\tau,\tau'])}(x_0, \gamma(\sigma)) \ge 25D.$$

Recall that $\gamma([\tau, \tau'])$ and $[q_1, p_1] = a_1\mathscr{W}(a_1)[x_0, f^{50D}x_0]$ are $2.3D$-equivalent by Lemma 3.2. Moreover, $\gamma(\sigma)$ and b_1x_0 are $2.3D$-close. These facts and Corollary 3.10(5) imply

$$d_{a_1\mathscr{W}(a_1)[x_0,f^{50D}x_0]}(x_0, b_1x_0) \ge 12D, \quad d_{a_1\mathscr{W}(a_1)Ax(f)}(x_0, b_1x_0) \ge 10D.$$

This contradicts the requirement that $b_1 \in \mathcal{NF}_D$.

(2) If $\sigma \in [\tau, \tau + 25D]$, then $\gamma([\tau, \tau'])$ and $\gamma([\sigma, \sigma'])$ overlap for length at least $15D$. Since $\gamma([\tau, \tau'])$ ($\gamma([\sigma, \sigma'])$, resp.) and $a_1\mathscr{W}(a_1)[x_0, f^{50D}x_0]$ ($b_1\mathscr{W}(b_1)[x_0, f^{50D}x_0]$, resp.) are $2.3D$-equivalent. By Corollary 3.10(1), (4) and (5), we have

$$\operatorname{diam}_{a_1\mathscr{W}(a_1)[x_0,f^{50D}x_0]}\big(b_1\mathscr{W}(b_1)Ax(f)\big) \ge 8D,\ \operatorname{diam}_{a_1\mathscr{W}(a_1)Ax(f)}\big(b_1\mathscr{W}(b_1)Ax(f)\big) \ge 7D.$$

This implies that $\mathscr{W}(a_1)^{-1}a_1^{-1}b_1\mathscr{W}(b_1) \in EC(f)$. Meanwhile, note that $a_1\mathscr{W}(a_1)x_0$ and $b_1\mathscr{W}(b_1)x_0$ are $30D$-close. Hence, we have

$$\big\|\mathscr{W}(a_1)^{-1}a_1^{-1}b_1\mathscr{W}(b_1)\big\|_S \le R',\ \|a_1^{-1}b_1\|_S \le R.$$

Since a_1, b_1 are chosen from an R-separated set A, we conclude $a_1 = b_1$ as desired.

By induction, we conclude that $a_i = b_i$ for each i.

It remains to show that $\mathcal{NF}_D$ is contained in a bounded neighborhood of $F(A)$ in the word metric. It is clear that $\mathcal{NF}_D$ is contained in the R-neighborhood of A, and A is contained in the $(50D\|f\|_S + 1)$-neighborhood of $F(A)$. This ends the proof. □

We now observe that $\mathcal{NF}_D^{\geq i}$ indeed serves as a "barrier". First, we record the following corollary of Lemma 3.14.

Lemma 7.11. *There exists $K > 0$ such that the following holds. Let $\gamma_0, \gamma_1, \ldots, \gamma_N \in \mathscr{A} := \{gAx(f) : g \in G\}$ and let $z \in X$ be such that:*
(1) $d_{\gamma_{i-1}}(x_0, \gamma_i) \geq K$ for $1 \leq i \leq N$, and
(2) $d_{\gamma_N}(x_0, z) \geq K$.
Then $d_{\gamma_1}(x_0, z) \geq d_{\gamma_0}(x_0, \gamma_1) - K$.

Proposition 7.12. *For each large enough $D, E > 0$, there exists $E' > 0$ such that the following holds. Let $g \in G$ be such that*
(1) there does not exist $h \in G$ such that
$$d_S(id, h) \leq E' \text{ and } d_{hAx(f)}(x_0, gx_0) \geq 250D.$$
(2) $d_S(id, g) \geq E$.
Let $(id = g_0, g_1, \ldots, g_N := g)$ be a d_S-path between id and g. Then there exists t such that $g_t \in \mathcal{NF}_{300D}^{\geq E}$.

Proof. Let D_0 be as in Lemma 3.16 and let $K_0 = K$ be as in Lemma 7.11. We assume that $D > 1000(\delta + \max_{s \in S} d_X(x_0, sx_0) + D_0 + K_0)$. Furthermore, let $E_{amp} = R(600D, E)$ be as in Lemma 3.16 and let $E' = E_{amp} + 600\|f\|_S$.

Suppose to the contrary that a d_S-path $P = (g_1, \ldots, g_N)$ never intersects $\mathcal{NF}_{300D}^{\geq E}$. We will deduce contradiction. Let
$$i(0) := \max\{i : d_S(id, g_i) \leq E\}.$$
If $g_{i(0)} \in \mathcal{NF}_{300D}$, then it is in $\mathcal{NF}_{300D}^{\geq E}$. Due to our standing assumption, this is not the case.

Thus, there exists $h \in G$ such that $d_{hAx(f)}(x_0, g_{i(0)}x_0) \geq 300D$. By replacing h with an element of $\{hf^i\}_{i \in \mathbb{Z}}$, we may suppose that $\pi_{hAx(f)}(x_0)$ intersects $[hx_0, hfx_0]$. Now consider a d_S-geodesic $Q = (id, g'_1, \ldots, g'_E =: g_{i(0)})$ connecting id and $g_{i(0)}$. By the coarse Lipschitzness of $\pi_{hAx(f)}(\cdot)$, there exist $1 \leq j \leq E$ such that $d_X(hx_0, \pi_{hAx(f)}(g'_j x_0)) =_D 30D$.

Then $\pi_{hAx(f)}(x_0)$ and $\pi_{hAx(f)}(g'_j x_0)$ are contained in $[hf^{-300D}x_0, hf^{300D}x_0]$, so we have
$$d_{[hf^{-300D}x_0, hf^{300D}x_0]}(x_0, g'_j x_0) = d_{hAx(f)}(x_0, g'_j x_0) \geq 25D.$$
Note here that $\|g'_j\|_S \leq \|g_{i(0)}\|_S = E$. Lemma 3.16 then tells us that $\|h\|_S \leq E_{amp}$ and $\|h\|_S \leq E'$.

In summary, the collection
$$\mathcal{C}_0 := \{\gamma \in \mathscr{A} : d_\gamma(x_0, g_{i(1)}x_0) \geq 300D\}$$
is a nonempty collection, and each element of $\mathcal{C}_0$ is realized as $hAx(f)$ for some $\|h\|_S \leq E'$. Let us take $\gamma_0 = h_0Ax(f) \in \mathcal{C}_0$ whose axis is the closest to x_0, i.e., the one with the smallest $d_X(x_0, h_0Ax(f))$.

If $d_{h_0Ax(f)}(x_0, g_ix_0) \geq 300D$ for all $i \geq i(0)$, including $i = N$, then it contradicts the condition on g. Hence, $d_{h_0Ax(f)}(x_0, g_ix_0) < 290D$ for some i.

Let us take the smallest such i and name it $i(1)$. By the coarse Lipschitzness of $\pi_{h_0 Ax(f)}(\cdot)$, we have $d_{h_0 Ax(f)}(x_0, g_{i(1)}x_0) \geq 289D$.

Meanwhile, by our standing assumption, the collection

$$\mathcal{C}_1 := \{\gamma \in \mathscr{A} : d_\gamma(x_0, g_{i(1)}x_0) > 300D\}$$

is nonempty. We pick $\gamma_1 \in \mathcal{C}_1$ that is the closest to x_0. Clearly $\gamma_1 \neq \gamma_0$.

Now, as in the proof of Claim 6.3, we can prove that γ_1 appears later than γ_0 along $[x_0, g_{i(2)}x_0]$; otherwise it will contradict the minimality of γ_0 in $\mathcal{C}_0$. We deduce that $d_{\gamma_0}(x_0, \gamma_1) > 289D$.

The proof goes on. If $d_{\gamma_1}(x_0, g_i x_0) \geq 300D$ for all $i \geq i(1)$, including $i = N$, then Lemma 7.11 implies that $d_{\gamma_0}(x_0, g_N x_0) > 280D$. This contradicts the condition on g.

Hence, $d_{\gamma_1}(x_0, g_i x_0) < 290D$ for some i, and we take the smallest such i as $i(2)$. We have $d_{\gamma_1}(x_0, g_{t_2}x_0) \geq 289D$. By the standing assumption,

$$\mathcal{C}_2 := \{\gamma \in \mathscr{A} : d_\gamma(x_0, g_{u(2)}x_0) > 300D\}$$

is nonempty. We pick $\gamma_2 \in \mathcal{C}_2$ that is the closest to x_0. We then observe that γ_2 appears later than γ_1; otherwise it violates the minimality of γ_1 in $\mathcal{C}_1$. We deduce that $d_{\gamma_1}(x_0, \gamma_2) > 289D$.

If $d_{\gamma_2}(x_0, g_i x_0) \geq 300D$ for all $i \geq i(2)$, then Lemma 7.11 again implies that $d_{\gamma_0}(x_0, g_N x_0) > 280D$, a contradiction. Thus, $d_{\gamma_2}(x_0, g_i x_0) < 290D$ for some i, and we take the smallest such i as $i(3)$. We have $d_{\gamma_2}(x_0, g_{i(3)}x_0) \geq 289D$.

If this process persists, it means we get an infinite sequence $i(1) < i(2) < \ldots$ in a finite sequence $0 \leq 1 \leq \ldots \leq N$. This is a contradiction. $\square$

Corollary 7.13. *Let Γ be the Cayley graph of an acylindrically hyperbolic group G. Then Equation 2.2 holds.*

Proof. Without loss of generality, we can fix an action of G on a δ-hyperbolic space X with a unital, axial WPD element $f \in G$. Let S be a finite generating set for G that gives rise to $\Gamma = Cay(G, S)$. Let $x_0 \in Ax(f)$.

Let $r > 0$ be as in Proposition 6.1. Given $D, E > 0$, we define

$$\mathscr{H}_D := \big\{\{g \in G : gx_0 \in \mathcal{H}_{r \cdot D}(x_0, ux_0)\} : u \in G\big\},$$
$$S_D := \mathcal{NF}_{300D} = \sqcup_{i=1}^{\infty} \mathcal{NF}_{300D}^i.$$

Also, let $E' = E'(D, E)$ be as in Proposition 7.12 for D and E. We then define

$$\mathcal{G}_{D,E} := \big\{g \in G : \|g\|_S \geq E,\ \nexists h \in G\big[\|h\|_S \leq E' \text{ and } d_{hAx(f)}(x_0, gx_0) \geq 250D\big]\big\}.$$

Now given $\epsilon > 0$ as well, we let $N = N(\epsilon, rD, E'(rD, E), E)$ be as in Proposition 7.7.

Then S_D is roughly branching for each D. Moreover, for each $\mathcal{H} \in \mathscr{H}_D$ there exists an r-branching subset $B = B_1 \sqcup \ldots \sqcup B_D$ that is a disjoint union of D d_S-barriers $B_1, \ldots, B_D$ between id and $\mathcal{H}$, by Proposition 6.1.

Finally, for each $D, E > 0$ we observe that $\mathcal{NF}_{300D}^{\geq E}$ is a d_S-barrier for id and $\mathcal{G}_{D,E}$ by Proposition 7.12. Lastly, Proposition 7.7 guarantees that for each finite $A \subseteq G$ there exist $A' \subseteq A$ with $\#A' \geq (1-\epsilon)\#A$ such that for each $a \in A'$, there exist $\mathcal{H}_1, \mathcal{H}_2 \in \mathscr{H}_D$ such that

$$\#\big(A \setminus a(\mathcal{H}_1 \cup \mathcal{H}_2 \cup \mathcal{G}_{D,E})\big) \leq N.$$

We can thus apply Theorem 2.16 and conclude Equation 2.2. □

Appendix A. Proof of Theorem 2.8

We summarize Hutchcroft's proof of Theorem 2.8. Let G, S, R and $\mathscr{H} = \{H(g) : g \in G\}$ be as in the assumption. Our goal is to show that there exists $\epsilon > 0$ such that $\left(\frac{d}{dp}\right)_+ \chi_p \geq \epsilon \chi_p^2$ holds for all $p_c/2 \leq p \leq p_c$. Then by integration we will have $\chi_p \leq \epsilon^{-1}(p_c - p)^{-1}$ for each $p_c/2 < p < p_c$, which leads to Equation 2.1.

We first define a $\{0,1\}$-valued function $I = I(g, A) \subseteq \Omega$ for inputs $a \in G$ and $A \subseteq G$:

$$I(a, A) := 1_{\{\exists g,h \in G[\|g\|_S, \|h\|_S \leq R \text{ and } A \subseteq aH(g) \text{ and } H(g) \cap hH(g) = \emptyset]\}}.$$

By our assumption, we have

$$\sum_{a \in A} I(a, A) \geq \frac{1}{2}\#A$$

for each $A \subseteq G$. We now define $F : \Omega \times G \to \mathbb{R}$:

$$F(\omega, a) := I\big(a, C_\omega(id)\big) 1_{a \in C_\omega(id)}.$$

We will now fix $p_c/2 < p < p_c$. We have

$$\sum_{a \in G} \mathbb{E}_p\, F(\omega, a) = \mathbb{E}_p \sum_{a \in C_\omega(id)} I\big(a, C_\omega(id)\big) \geq \mathbb{E}_p \left(\frac{1}{2}\#C_\omega(id)\right) = \frac{1}{2}\chi_p.$$

Meanwhile, we have

$$\begin{aligned}
\sum_{a \in G} \mathbb{E}_p\, F(\omega, a) &= \sum_{a \in G} \mathbb{E}_p\, F(\omega, a^{-1}) = \mathbb{E}_p \sum_{a : a^{-1} \in C_\omega(id)} I\big(a^{-1}, C_\omega(id)\big) \\
&= \mathbb{E}_p \sum_{a : id \in C_\omega(a)} I\big(id, C_\omega(a)\big) \\
&= \mathbb{E}_p \sum_{a \in C_\omega(id)} I\big(id, C_\omega(id)\big) = \mathbb{E}_p \left[\#C_{id} \cdot I\big(id, C_\omega(id)\big)\right].
\end{aligned}$$

Recall also that there are at most N choices $g, h \in G$ such that $\|g\|_S, \|h\|_S < R$, where $N = (2\#S)^{2R}$. Among those finitely many candidates, there exist a concrete, non-random g, h such that $H(g)$ and $hH(g)$ are disjoint and such that

$$\mathbb{E}_p \left[\#C(id) \cdot 1_{\{C(id) \subseteq H(g)\}}\right] \geq \frac{1}{2N}\chi_p.$$

By applying the action of h, we also have

$$\mathbb{E}_p \left[\#C(h) \cdot 1_{\{C(h) \subseteq hH(g)\}}\right] \geq \frac{1}{2N}\chi_p.$$

Now, note that the event $\{C(id) \subseteq H(g)\}$ is determined solely by edges in $H(g)$, and $\{C(h) \subseteq hH(g)\}$ is determined solely by edges in $hH(g)$. Since the two sets are disjoint, the two events are independent. We conclude that

$$\mathbb{E}_p\left[\big(\#C(id)\big) \cdot \big(\#C(h)\big) \cdot 1_{\{C(id) \subseteq H(g) \text{ and } C(h) \subseteq hH(g)\}}\right] \geq \frac{1}{4N^2}\chi_p^2.$$

We now pick a d_S-geodesic $\gamma = (g_1, g_2, \ldots, g_{\|h\|_S})$ connecting id to h. Let $e = \overrightarrow{g_n g_{n+1}}$ be the first (oriented) edge of γ that connects $H(g)$ to $\Gamma \setminus H(g)$. As described in Hutchcroft's proof, a standard conversion of finitely many states guarantees a constant c_p, which is a linear combination of finitely many products and ratios of p and $1-p$, hence bounded on compact subsets of $(0,1)$, such that

$$\mathbb{E}_p\left[\big(\#C(g_n)\big)\cdot\big(\#C(g_{n+1})\big)\cdot 1_{\{g_n \not\leftrightarrow g_{n+1}\}}\right] \geq c_p\, \mathbb{E}_p\left[\big(\#C(id)\big)\cdot\big(\#C(h)\big)\cdot 1_{\{C(id)\subseteq H(g), C(h) \subseteq hH(g)\}}\right].$$

By applying the action by g_n^{-1}, we conclude that

$$\mathbb{E}_p\left[\big(\#C(id)\big) \cdot \big(\#C(s)\big) \cdot 1_{id \not\leftrightarrow s}\right] \geq c\chi_p^2$$

for some c uniform on $0.5p_c < p < p_c$ and for some $s = s(p)$ in the generating set S.

We now recall Russo's formula. For each $g \in G$, a closed edge $e = \overrightarrow{e^- e^+} \in \mathcal{E}^{\rightarrow}$ is pivotal for the event $\{id \leftrightarrow g\}$ if and only if $id \leftrightarrow e^-$, $e^+ \leftrightarrow g$ and $e^- \not\leftrightarrow e^+$. Hence, we have

$$\left(\frac{d}{dp}\right)_+ \tau_p(g) \geq \frac{1}{1-p} \sum_{e \in \mathcal{E}^{\rightarrow}} \mathbb{P}_p(\{id \leftrightarrow e^-\} \cap \{e^- \not\leftrightarrow e^+\} \cap \{e^+ \leftrightarrow g\}).$$

Since $\tau_p(g)$ are monotonic for each $g \in G$, summing over finitely many g's and taking limits imply

$$\begin{aligned}
\left(\frac{d}{dp}\right)_+ \chi_p &\geq \frac{1}{1-p} \sum_{g \in G} \sum_{e \in \mathcal{E}^{\rightarrow}} \mathbb{P}_p(\{id \leftrightarrow e^-\} \cap \{e^- \not\leftrightarrow e^+\} \cap \{e^+ \leftrightarrow g\}) \\
&= \frac{1}{1-p} \sum_{s \in S} \sum_{h \in G} \sum_{g \in G} \mathbb{P}_p(\{id \leftrightarrow h\} \cap \{h \not\leftrightarrow hs\} \cap \{hs \leftrightarrow g\}) \\
&= \frac{1}{1-p} \sum_{s \in S} \sum_{h \in G} \sum_{g \in G} \mathbb{P}_p(\{id \leftrightarrow h\} \cap \{h \not\leftrightarrow hs\} \cap \{hs \leftrightarrow hsg\}) \\
&= \frac{1}{1-p} \sum_{g,h \in G, s \in S} \mathbb{P}_p(\{h^{-1} \leftrightarrow id\} \cap \{id \not\leftrightarrow s\} \cap \{s \leftrightarrow g\}).
\end{aligned}$$

The last summation is bounded from below by $\frac{1}{1-p}\,\mathbb{E}_p\left[\big(\#C(id)\big) \cdot \big(\#C(s)\big) \cdot 1_{id \not\leftrightarrow s}\right]$ for our choice $s = s(p)$. Hence, we conclude that $(d/dp)_+ \chi_p$ is uniformly coarsely bounded from below by χ_p^2 for $p \in (p_c/2, p_c)$. This ends the proof.

Appendix B. Proof of Corollary 3.10

We sketch the proof of Corollary 3.10.

(1) It is a direct consequence of Lemma 3.9.
(2) We have

$$d_X(x,p)+d_X(p,y)-12\delta \le d_X(x,p)+d_X(p,q)+d_X(q,y)-12\delta \le d_X(x,y).$$

Hence, $(x|y)_p \le 6\delta$. Pick $x' \in [x,y]$ such that $d_X(x,x') = (y|p)_x$. Then by Lemma 3.5, x' is 10δ-close to p. Similarly, the point $y' \in [x,y]$ such that $d_X(x,y') = (y|q)_x$ is 10δ-close to p. Note that

$$d_X(x,p) \le d_X(x,y)-d_X(q,y)-d_X(p,q)+12\delta < d_X(x,q)-12\delta+12\delta = d_X(x,q).$$

For a similar reason, we have $d_X(y,q) < d_X(y,p)$. This implies that

$$d_X(x,x') = \frac{1}{2}[d_X(x,y)+d_X(x,p)-d_X(y,p)] \le \frac{1}{2}[d_X(x,y)+d_X(x,q)-d_X(y,q)] = d_X(x,y').$$

Hence, x' comes earlier than y' along $[x,y]$. By Lemma 3.2, $[x',y']$ and $[p,q]$ are 12δ-equivalent.

(3) Suppose to the contrary that there exist $p \in \pi_\gamma(x)$, $q \in \pi_\gamma(y)$ and $r \in \pi_\gamma(z)$ such that $r \notin \mathcal{N}_{12\delta}([p,q])$. This means that $d_X(p,r) + d_X(r,q) > 24\delta + d_X(p,q)$.

Now Lemma 3.9 tells us that

$$d_X(x,z) \ge d_X(x,p)+d_X(r,z)+d_X(p,r)-12\delta,$$
$$d_X(z,y) \ge d_X(y,q)+d_X(r,z)+d_X(q,r)-12\delta.$$

This implies that

$$d_X(x,y) \ge d_X(x,p)+d_X(y,q)+\big(d_X(p,r)+d_X(r,q)\big)-24\delta > d_X(x,p)+d_X(y,q)+d_X(p,q).$$

This is a contradiction.

(4) Let $\gamma = [y,z]$ and $\gamma' = [y',z']$. Let $a \in [y,z]$ be such that $d_X(y,a) = (x|z)_y$, let $b \in [y,x]$ be such that $d_X(y,b) = (x|z)_y$, let $c \in [y,x]$ be such that $d_X(y,c) = (x|z')_y$, let $d \in [x,z']$ be such that $d_X(z',d) = (x|y)_{z'}$, let $e \in [x,z']$ be such that $d_X(z',e) = (x|y')_{z'}$, and let $f \in [y',z']$ be such that $d_X(z',f) = (x|y')_{z'}$.

Then by Lemma 3.8, a and e are 8δ-equivalent to $\pi_\gamma(x)$ and $\pi_{\gamma'}(x)$, respectively. Moreover, Lemma 3.5 tells us that $d_X(a,b), d_X(c,d), d_X(e,f) \le 4\delta$, and the triangle inequality tells us that $d_X(b,c) \le d_X(z,z') \le D$, $d_X(d,e) \le d_X(y,y') \le D$. In conclusion, $\pi_\gamma(x)$ and $\pi_{\gamma'}(x)$ are $(2D+28\delta)$-equivalent as desired.

(5) Suppose that there exist $p \in \pi_\gamma(x)$ and $q \in \pi_\gamma(y)$ that realizes $d_X(p,q) = d_{\gamma'}(x,y) > 12\delta$. By Corollary 3.10(2), there exists $x', y' \in [x,y]$ that are 10δ-close to p and q, respectively. In particular, x' and y' are 16δ-close to $\gamma' \subseteq \gamma$. This implies

$$\pi_\gamma(x') \subseteq \mathcal{N}_{10\delta}(x') \subseteq \mathcal{N}_{20\delta}(p), \quad \pi_\gamma(y') \subseteq \mathcal{N}_{10\delta}(y') \subseteq \mathcal{N}_{20\delta}(q).$$

This implies that $d_\gamma(x', y') \geq d_X(p, q) - 40\delta$. Meanwhile, Corollary 3.10(3) tells us that $d_\gamma(x, y) \geq d_\gamma(x', y') - 24\delta$. Combining these two, we conclude that

$$\mathrm{diam}_\gamma(x, y) \geq d_\gamma(x', y') - 24\delta \geq d_X(p, q) - 64\delta.$$

References

[AKN87] M. Aizenman, H. Kesten, and C. M. Newman. Uniqueness of the infinite cluster and continuity of connectivity functions for short and long range percolation. *Comm. Math. Phys.*, 111(4):505–531, 1987.

[AN84] Michael Aizenman and Charles M. Newman. Tree graph inequalities and critical behavior in percolation models. *J. Statist. Phys.*, 36(1-2):107–143, 1984.

[BBF15] Mladen Bestvina, Ken Bromberg, and Koji Fujiwara. Constructing group actions on quasi-trees and applications to mapping class groups. *Publ. Math. Inst. Hautes Études Sci.*, 122:1–64, 2015.

[BBFS19] Mladen Bestvina, Ken Bromberg, Koji Fujiwara, and Alessandro Sisto. Acylindrical actions on projection complexes. *Enseign. Math.*, 65(1-2):1–32, 2019.

[Bes23] Mladen Bestvina. Groups acting on hyperbolic spaces—a survey. In *ICM—International Congress of Mathematicians. Vol. 2. Plenary lectures*, pages 678–711. EMS Press, Berlin, [2023] ©2023.

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

[BH99] Martin R. Bridson and André Haefliger. *Metric spaces of non-positive curvature*, volume 319 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.

[BK89] R. M. Burton and M. Keane. Density and uniqueness in percolation. *Comm. Math. Phys.*, 121(3):501–505, 1989.

[BLPS99a] I. Benjamini, R. Lyons, Y. Peres, and O. Schramm. Group-invariant percolation on graphs. *Geom. Funct. Anal.*, 9(1):29–66, 1999.

[BLPS99b] Itai Benjamini, Russell Lyons, Yuval Peres, and Oded Schramm. Critical percolation on any nonamenable group has no infinite clusters. *Ann. Probab.*, 27(3):1347–1356, 1999.

[Bon96] Mario Bonk. Quasi-geodesic segments and Gromov hyperbolic spaces. *Geom. Dedicata*, 62(3):281–298, 1996.

[BS96] Itai Benjamini and Oded Schramm. Percolation beyond $\mathbf{Z}^d$, many questions and a few answers. *Electron. Comm. Probab.*, 1:no. 8, 71–82, 1996.

[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.

[Cho25] Inhyeok Choi. Acylindrically hyperbolic groups and counting problems. *arXiv preprint arXiv:2504.20985*, 2025.

[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.

[FKG71] C. M. Fortuin, P. W. Kasteleyn, and J. Ginibre. Correlation inequalities on some partially ordered sets. *Comm. Math. Phys.*, 22:89–103, 1971.

[GKN92] A. Gandolfi, M. S. Keane, and C. M. Newman. Uniqueness of the infinite component in a random graph with applications to percolation and spin glasses. *Probab. Theory Related Fields*, 92(4):511–527, 1992.

[GN90] G. R. Grimmett and C. M. Newman. Percolation in $\infty + 1$ dimensions. In *Disorder in physical systems*, Oxford Sci. Publ., pages 167–190. Oxford Univ. Press, New York, 1990.

[Gri89] Geoffrey Grimmett. *Percolation.* Springer-Verlag, New York, 1989.

[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.

[GS21] Antoine Goldsborough and Alessandro Sisto. Markov chains on hyperbolic-like groups and quasi-isometries. *arXiv preprint arXiv:2111.09837*, 2021.

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

[Har60] T. E. Harris. A lower bound for the critical probability in a certain percolation process. *Proc. Cambridge Philos. Soc.*, 56:13–20, 1960.

[HJ06] Olle Häggström and Johan Jonasson. Uniqueness and non-uniqueness in percolation theory. *Probab. Surv.*, 3:289–344, 2006.

[HP99] Olle Häggström and Yuval Peres. Monotonicity of uniqueness for percolation on Cayley graphs: all infinite clusters are born simultaneously. *Probab. Theory Related Fields*, 113(2):273–285, 1999.

[HPS99] Olle Häggström, Yuval Peres, and Roberto H. Schonmann. Percolation on transitive graphs as a coalescent process: relentless merging followed by simultaneous uniqueness. In *Perplexing problems in probability*, volume 44 of *Progr. Probab.*, pages 69–90. Birkhäuser Boston, Boston, MA, 1999.

[Hut16] Tom Hutchcroft. Critical percolation on any quasi-transitive graph of exponential growth has no infinite clusters. *C. R. Math. Acad. Sci. Paris*, 354(9):944–947, 2016.

[Hut19] Tom Hutchcroft. Percolation on hyperbolic graphs. *Geom. Funct. Anal.*, 29(3):766–810, 2019.

[Hut20a] Tom Hutchcroft. The L^2 boundedness condition in nonamenable percolation. *Electron. J. Probab.*, 25:Paper No. 127, 27, 2020.

[Hut20b] Tom Hutchcroft. Nonuniqueness and mean-field criticality for percolation on nonunimodular transitive graphs. *J. Amer. Math. Soc.*, 33(4):1101–1165, 2020.

[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.

[NS81] C. M. Newman and L. S. Schulman. Infinite clusters in percolation models. *J. Statist. Phys.*, 26(3):613–628, 1981.

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

[PSN00] Igor Pak and Tatiana Smirnova-Nagnibeda. On non-uniqueness of percolation on nonamenable Cayley graphs. *C. R. Acad. Sci. Paris Sér. I Math.*, 330(6):495–500, 2000.

[Rus81] Lucio Russo. On the critical percolation probabilities. *Z. Wahrsch. Verw. Gebiete*, 56(2):229–237, 1981.

[Sch99] Roberto H. Schonmann. Stability of infinite clusters in supercritical percolation. *Probab. Theory Related Fields*, 113(2):287–300, 1999.

[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.

[vdBK85] J. van den Berg and H. Kesten. Inequalities with applications to percolation and reliability. *J. Appl. Probab.*, 22(3):556–569, 1985.

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

Research Institute of Mathematics, Seoul National University, Seoul, Korea
Email address: seodonggyun@snu.ac.kr, seodonggyun7@gmail.com