CONJUGATOR LENGTHS IN HIERARCHICALLY HYPERBOLIC GROUPS

CAROLYN ABBOTT AND JASON BEHRSTOCK

Abstract. In this paper we establish upper bounds on the length of the shortest conjugator between pairs of elements in a wide class of groups. We obtain a general result which applies to all hierarchically hyperbolic groups, a class which includes mapping class groups, right angled Artin groups, Burger–Mozes-type groups, most 3–manifold groups, and many others. One case of our result in this setting is a linear bound on the length of the shortest conjugator for any pair of conjugate Morse elements. In a special case, namely, for virtually compact special cubical groups, we can prove a sharper result by obtaining a linear bound on the length of the shortest conjugator between any pair of infinite order elements. In a more general case, that of acylindrically hyperbolic groups, we establish an upper bound on the length of shortest conjugators, but in this generality the bound may not be linear.

An early result proven about hyperbolic groups is that for any pair of conjugate elements the shortest conjugator has linear length [Lys89]. Exploiting the parallels between pseudo-Anosovs in the mapping class group and loxodromic elements in a hyperbolic group, Masur–Minsky proved the analogous result that conjugate pseudo-Anosov elements always have a short conjugator [MM00]. A linear bound was later established between arbitrary conjugate elements in the mapping class group by Tao [Tao13] (see [BD14] for a later, unified proof).

The linear conjugator property is surprisingly common, as we show in this paper, extending an already interesting class of known examples. Known examples of the linear conjugator property include: hyperbolic elements in semi-simple Lie groups [Sal14]; arbitrary elements in lamplighter groups [Sal16]; Morse elements in groups acting on CAT(0) spaces, [BD14]; and, Morse elements in a prime 3–manifold [BD14]. Additionally, right-angled Artin groups enjoy the linear conjugator property; this result is not explicitly stated in the literature, but it follows from work in [Ser89] (and we give a new proof below).

We will work in the context of hierarchically hyperbolic groups, a general class of groups introduced by Behrstock–Hagen–Sisto [BHS17b] (see Section 1). This class of groups includes: mapping class groups [BHS15]; right angled Artin groups, and more generally fundamental groups of compact special CAT(0) cube complexes [BHS17b] and other CAT(0) cube complexes [HS16]; 3–manifold groups with no Nil or Solv components [BHS15]; lattices in products of trees, i.e., as constructed by Burger–Mozes, Wise, and others, see [BHS17b, BM97, BM00, Cap17, JW09, Rat07, Wis07]; as well as a number of other examples, for instance groups obtained from combination theorems [BHS15] [Spr18] or by taking certain quotients [BHS17a].

The first theorem generalizes several of the known cases mentioned above, while applying to a number of new examples as well. Relevant definitions will be provided in the background section; for now, we remark that an important case of this theorem is when a and b are infinite order Morse elements of G (in the notation of the theorem this occurs when $\mathrm{Big}(a)$ is the nest-maximal domain). We note that Morse elements in hierarchically hyperbolic groups can be characterized in several equivalent ways, see [ABD, Theorem B].

Abbott was supported by NSF grant DMS-1803368.
Behrstock was supported by NSF grant DMS-1710890.

Theorem A. *Let $(G, \mathfrak{S})$ be a hierarchically hyperbolic group. There exist constants K, C such that if $a, b \in G$ are infinite order elements which are conjugate in G and b has the property that* $\mathrm{Big}(b)$ *is a maximal collection of pairwise orthogonal domains, then there exists $g \in G$ with $ga = bg$ and*

$$|g| \leqslant K(|a| + |b|) + C.$$

In Theorem A, as well as below, we need only include a hypothesis about the big set of one of the elements, since a conjugation between two elements conjugates the set of domains on which they have large projections. An easy-to-visualize non-Morse example to which the above theorem applies is any mapping class group element obtained by applying powers of a Dehn twists along every curve in a pants decomposition (here the collection of annuli around those curves constitute the maximal collection of pairwise orthogonal domains needed to apply the theorem).

In the case of a group acting geometrically on a compact CAT(0) cube complex which is special in the sense of Haglund–Wise [HW08], the following result provides a strengthening of Theorem A by bounding the lengths of conjugators for *any* pair of conjugate infinite order elements:

Theorem B. *Let G be a virtually compact special CAT(0) cubical group. There exist constants K, C and N such that if $a, b \in G$ are infinite order elements which are conjugate in G, then there exists $g \in G$ with $ga^N = b^N g$ and*

$$|g| \leqslant K(|a| + |b|) + C.$$

In Remark 3.3 below, we give some indication of why this result might fail without the hypothesis of virtually special for the cube complex.

We note that [CGW09] establish a linear time solution to the conjugacy problem for fundamental groups of compact special CAT(0) cube complexes. Their result doesn't a priori establish the linear conjugator property of Theorem B, although we believe that their approach could be used to do so. Nonetheless, we include a proof here, since this result is not explicitly in the literature and our proof is a very short application of our general approach.

Our techniques for bounding conjugator length can be applied in the general acylindrically hyperbolic setting to provide the following result, which is stated in a more explicit form as Theorem 4.1. We note that this result generalizes [BD14, Theorem 7.4] which considers the case of acylindrical actions on simplicial trees.

Theorem C. *Let G be an acylindrically hyperbolic group. There exist an associated quasi-tree T, a function $f \colon \mathbb{Z} \to \mathbb{Z}$, and a constant R such that: for any two conjugate elements $a, b \in G$ which are loxodromic with respect to the action of G on T, there exists $g \in G$ such that $ga = bg$ and*

$$|g| \leqslant f(R|a| + R|b|) + |a| + |b|.$$

For a given group G, both the constant R and the function f are fixed for any given quasi-tree, T. This provides uniformity to the bound on the length of shortest conjugators for any element acting loxodromically on T and, more generally, for those elements that act loxodromically on any one of a finite set of quasi-trees. Hence, if a particular acylindrically hyperbolic group embeds into a finite product of quasi-trees, then one obtains a uniform bound on the lengths of shortest conjugators. It is an interesting question whether or not a given group embeds in a finite product of quasi-trees and Theorem C provides an additional motivation to study this question.

Acknowlegments. We thank Mark Hagen for many interesting discussions about hierarchically hyperbolic spaces, and, in particular, for the discussion which led to Remark 3.3. We thank Jacob Russell for feedback on an early draft.

1. Background

We recall the definition of a hierarchically hyperbolic space as given in [BHS15].

Definition 1.1 (Hierarchically hyperbolic space)**.** The quasigeodesic space $(\mathcal{X}, \mathsf{d}_{\mathcal{X}})$ is a *hierarchically hyperbolic space* if there exists $\delta \geqslant 0$, an index set $\mathfrak{S}$, and a set $\{\mathcal{C}W : W \in \mathfrak{S}\}$ of δ–hyperbolic spaces $(\mathcal{C}U, \mathsf{d}_U)$, such that the following conditions are satisfied:

(1) **(Projections.)** There is a set $\{\pi_W \colon \mathcal{X} \to 2^{\mathcal{C}W} \mid W \in \mathfrak{S}\}$ of *projections* sending points in $\mathcal{X}$ to sets of diameter bounded by some $\xi \geqslant 0$ in the various $\mathcal{C}W \in \mathfrak{S}$. Moreover, there exists K so that each π_W is (K, K)–coarsely Lipschitz.

(2) **(Nesting.)** $\mathfrak{S}$ is equipped with a partial order $\sqsubseteq$, and either $\mathfrak{S} = \varnothing$ or $\mathfrak{S}$ contains a unique $\sqsubseteq$–maximal element; when $V \sqsubseteq W$, we say V is *nested* in W. (We emphasize that $W \sqsubseteq W$ for all $W \in \mathfrak{S}$.) For each $W \in \mathfrak{S}$, we denote by $\mathfrak{S}_W$ the set of $V \in \mathfrak{S}$ such that $V \sqsubseteq W$. Moreover, for all $V, W \in \mathfrak{S}$ with $V \sqsubsetneq W$ there is a specified subset $\rho^V_W \subset \mathcal{C}W$ with $\mathsf{diam}_{\mathcal{C}W}(\rho^V_W) \leqslant \xi$. There is also a *projection* $\rho^W_V \colon \mathcal{C}W \to 2^{\mathcal{C}V}$.

(3) **(Orthogonality.)** $\mathfrak{S}$ has a symmetric and anti-reflexive relation called *orthogonality*: we write $V \perp W$ when V, W are orthogonal. Also, whenever $V \sqsubseteq W$ and $W \perp U$, we require that $V \perp U$. We require that for each $T \in \mathfrak{S}$ and each $U \in \mathfrak{S}_T$ for which $\{V \in \mathfrak{S}_T \mid V \perp U\} \neq \varnothing$, there exists $W \in \mathfrak{S}_T - \{T\}$, so that whenever $V \perp U$ and $V \sqsubseteq T$, we have $V \sqsubseteq W$. Finally, if $V \perp W$, then V, W are not $\sqsubseteq$–comparable.

(4) **(Transversality and consistency.)** If $V, W \in \mathfrak{S}$ are not orthogonal and neither is nested in the other, then we say V, W are *transverse*, denoted $V \pitchfork W$. There exists $\kappa_0 \geqslant 0$ such that if $V \pitchfork W$, then there are sets $\rho^V_W \subseteq \mathcal{C}W$ and $\rho^W_V \subseteq \mathcal{C}V$ each of diameter at most ξ and satisfying:

$$\min\left\{\mathsf{d}_W(\pi_W(x), \rho^V_W), \mathsf{d}_V(\pi_V(x), \rho^W_V)\right\} \leqslant \kappa_0$$

for all $x \in \mathcal{X}$.

For $V, W \in \mathfrak{S}$ satisfying $V \sqsubseteq W$ and for all $x \in \mathcal{X}$, we have:

$$\min\left\{\mathsf{d}_W(\pi_W(x), \rho^V_W), \mathsf{diam}_{\mathcal{C}V}(\pi_V(x) \cup \rho^W_V(\pi_W(x)))\right\} \leqslant \kappa_0.$$

The preceding two inequalities are the *consistency inequalities* for points in $\mathcal{X}$.

Finally, if $U \sqsubseteq V$, then $\mathsf{d}_W(\rho^U_W, \rho^V_W) \leqslant \kappa_0$ whenever $W \in \mathfrak{S}$ satisfies either $V \sqsubsetneq W$ or $V \pitchfork W$ and $W \not\perp U$.

(5) **(Finite complexity.)** There exists $n \geqslant 0$, the *complexity* of $\mathcal{X}$ (with respect to $\mathfrak{S}$), so that any set of pairwise–$\sqsubseteq$–comparable elements has cardinality at most n.

(6) **(Large links.)** There exist $\lambda \geqslant 1$ and $E \geqslant \max\{\xi, \kappa_0\}$ such that the following holds. Let $W \in \mathfrak{S}$ and let $x, x' \in \mathcal{X}$. Let $N = \lambda \mathsf{d}_W(\pi_W(x), \pi_W(x')) + \lambda$. Then there exists $\{T_i\}_{i=1,\dots,\lfloor N \rfloor} \subseteq \mathfrak{S}_W - \{W\}$ such that for all $T \in \mathfrak{S}_W - \{W\}$, either $T \in \mathfrak{S}_{T_i}$ for some i, or $\mathsf{d}_T(\pi_T(x), \pi_T(x')) < E$. Also, $\mathsf{d}_W(\pi_W(x), \rho^{T_i}_W) \leqslant N$ for each i.

(7) **(Bounded geodesic image.)** There exists $E > 0$ such that for all $W \in \mathfrak{S}$, all $V \in \mathfrak{S}_W - \{W\}$, and all geodesics γ of $\mathcal{C}W$, either $\mathsf{diam}_{\mathcal{C}V}(\rho^W_V(\gamma)) \leqslant E$ or $\gamma \cap \mathcal{N}_E(\rho^V_W) \neq \varnothing$.

(8) **(Partial Realization.)** There exists a constant α with the following property. Let $\{V_j\}$ be a family of pairwise orthogonal elements of $\mathfrak{S}$, and let $p_j \in \pi_{V_j}(\mathcal{X}) \subseteq \mathcal{C}V_j$. Then there exists $x \in \mathcal{X}$ so that:
   - $\mathsf{d}_{V_j}(x, p_j) \leqslant \alpha$ for all j,
   - for each j and each $V \in \mathfrak{S}$ with $V_j \sqsubseteq V$, we have $\mathsf{d}_V(x, \rho^{V_j}_V) \leqslant \alpha$, and
   - if $W \pitchfork V_j$ for some j, then $\mathsf{d}_W(x, \rho^{V_j}_W) \leqslant \alpha$.

(9) **(Uniqueness.)** For each $\kappa \geqslant 0$, there exists $\theta_u = \theta_u(\kappa)$ such that if $x, y \in \mathcal{X}$ and $\mathsf{d}_{\mathcal{X}}(x, y) \geqslant \theta_u$, then there exists $V \in \mathfrak{S}$ such that $\mathsf{d}_V(x, y) \geqslant \kappa$.

For ease of readability, given $U \in \mathfrak{S}$, we typically suppress the projection map π_U when writing distances in $\mathcal{C}U$, i.e., given $x, y \in \mathcal{X}$ and $p \in \mathcal{C}U$ we write $\mathsf{d}_U(x,y)$ for $\mathsf{d}_U(\pi_U(x), \pi_U(y))$ and $\mathsf{d}_U(x,p)$ for $\mathsf{d}_U(\pi_U(x), p)$.

Notation 1.2. Given a hierarchically hyperbolic space $(\mathcal{X}, \mathfrak{S})$ we let E denote a constant larger than any of the constants occurring in the definition above.

A group G is said to be *hierarchically hyperbolic* if there exists a hierarchically hyperbolic space $(\mathcal{X}, \mathfrak{S})$ with the property that G acts geometrically on $\mathcal{X}$ and this action permutes elements of $\mathfrak{S}$ with finitely many G-orbits, while preserves nesting, orthogonality, and transversality and, for each $U \in \mathfrak{S}$ inducing a uniform quasi-isometry between $\mathcal{C}U$ and its image. Often we just denote a HHG as $(G, \mathfrak{S})$.

Given a hierarchically hyperbolic space $(\mathcal{X}, \mathfrak{S})$ and a constant $M \geqslant 1$, a (M,M)*–hierarchy path*, $\gamma \subset \mathcal{X}$ is an (M,M)–quasigeodesic in $\mathcal{X}$ with the property that for each $U \in \mathfrak{S}$ the path $\pi_U(\gamma)$ is an unparametrized quasigeodesic in $\mathcal{C}U$. By [BHS15, Theorem 4.4], for any sufficiently large M, any two points $x, y \in \mathcal{X}$ are connected by an (M,M)–hierarchy path. We fix such a constant $M > E$, and let $[x,y]$ denote an (M,M)–hierarchy path from x to y.

1.1. **Product Regions and Gates.**

We now recall an important construction of subspaces in a hierarchically hyperbolic space called *standard product regions* introduced in [BHS17b, Section 13] and studied further in [BHS15], which one can consult for further details. We begin by defining the two factors in the product space.

Definition 1.3 (Nested partial tuple $(\mathbf{F}_U)$)**.** Let $\mathfrak{S}_U = \{V \in \mathfrak{S} \mid V \sqsubseteq U\}$. Fix $\kappa \geqslant E$ and let $\mathbf{F}_U$ be the set of κ–consistent tuples in $\prod_{V \in \mathfrak{S}_U} 2^{\mathcal{C}V}$ (i.e., tuples satisfying the conditions of Definition 1.1.(4)).

Definition 1.4 (Orthogonal partial tuple $(\mathbf{E}_U)$)**.** Let $\mathfrak{S}_U^{\perp} = \{V \in \mathfrak{S} \mid V \perp U\} \cup \{A\}$, where A is a $\sqsubseteq$–minimal element W such that $V \sqsubseteq W$ for all $V \perp U$. Fix $\kappa \geqslant E$, let $\mathbf{E}_U$ be the set of κ–consistent tuples in $\prod_{V \in \mathfrak{S}_U^{\perp} - \{A\}} 2^{\mathcal{C}V}$.

Definition 1.5 (Product regions in $\mathcal{X}$)**.** Given $\mathcal{X}$ and $U \in \mathfrak{S}$, there are coarsely well-defined maps $\phi^{\sqsubseteq}, \phi^{\perp} \colon \mathbf{F}_U, \mathbf{E}_U \to \mathcal{X}$ which extend to a coarsely well-defined map $\phi_U \colon \mathbf{F}_U \times \mathbf{E}_U \to \mathcal{X}$. We refer to $\mathbf{F}_U \times \mathbf{E}_U$ as a *product region*, which we denote $\mathbf{P}_U$.

We often abuse notation and use the notation $\mathbf{E}_U, \mathbf{F}_U$, and $\mathbf{P}_U$ to refer to the image in $\mathcal{X}$ of the associated set. In [BHS15, Lemma 5.9] it is proven that these standard product regions have the property that they are "hierarchically quasiconvex subsets" of $\mathcal{X}$. Here we leave out the definition of hierarchically quasiconvexity, because its only use here is that, as proven in [BHS15, Lemma 5.5], product regions have "gates," which we now define.

If $(G, \mathfrak{S})$ is a hierarchically hyperbolic group and $\mathcal{Y}$ is a hierarchically quasiconvex subspace of G, then the *gate map* is a coarsely-Lipschitz map $\mathfrak{g}_{\mathcal{Y}} \colon G \to 2^{\mathcal{Y}}$, so that for each $g \in G$, the image $\mathfrak{g}_{\mathcal{Y}}(g)$ is a subset of the points in $\mathcal{Y}$ with the property that for each $U \in \mathfrak{S}$ the set $\pi_U(\mathfrak{g}_{\mathcal{Y}}(g))$ uniformly coarsely coinsides with the closest point projection in $\mathcal{C}U$ of $\pi_U(g)$ to $\pi_U(\mathcal{Y})$. In the case that G is a CAT(0) cubical group, then $\mathfrak{g}_{\mathcal{Y}}(g)$ is a unique point in $\mathcal{Y}$, whence the gate map is a well-defined map $\mathfrak{g}_{\mathcal{Y}} \colon G \to \mathcal{Y}$.

It is shown in [BHS15, Proposition 5.11] that for any $U \in \mathfrak{S}$, $(\mathbf{F}_U, \mathfrak{S}_U)$ is a hierarchically hyperbolic space, where the projections and hyperbolic spaces are those inherited from $(\mathcal{X}, \mathfrak{S})$.

The following lemma which provides a formula for computing the distance between a point and a product region, is an immediate consequence of [BHS17c, Lemma 1.19].

Lemma 1.6 ([BHS17c]). *Let $(\mathcal{X}, \mathfrak{S})$ be a hierarchically hyperbolic space. Fix $U \in \mathfrak{S}$ and let $\mathcal{Y} = \{Y \in \mathfrak{S} \mid Y \pitchfork U$ or $Y \sqsupseteq U\}$. Then given any $x \in \mathcal{X}$,*

$$\mathsf{d}_{\mathcal{X}}(x, \mathbf{P}_U) \asymp \sum_{Y \in \mathcal{Y}} \left\{\!\!\left\{ \mathsf{d}_Y(x, \rho_Y^U) \right\}\!\!\right\}. \tag{1}$$

1.2. **Big sets.** Let $(G, \mathfrak{S})$ be a hierarchically hyperbolic group and consider $g \in G$. As in [DHS17], we define

$$\operatorname{Big}(g) = \{U \in \mathfrak{S} \mid \pi_U(\langle g \rangle) \text{ is unbounded}\}.$$

Note that all $U, V \in \operatorname{Big}(g)$ satisfy $U \perp V$, so ρ_S^U is within uniformly bounded distance of ρ_S^V in $\mathcal{C}S$. We also observe that since the elements of $\operatorname{Big}(g)$ must all be pairwise orthogonal, it follows immediately that $|\operatorname{Big}(g)|$ is uniformly bounded by the HHS constants.

Lemma 1.7. *Let $(G, \mathfrak{S})$ be a hierarchically hyperbolic group. An element $g \in G$ is finite order if and only if $\operatorname{Big}(g) = \varnothing$.*

Proof. In [DHS17, Proposition 6.4] it was proven that an automorphism of G is elliptic if and only if $\operatorname{Big}(g) = \varnothing$. The result follows from this, since a group element acts elliptically on its Cayley graph if and only if the element is of finite order. □

Following [DHS17, Remark 1.14], for any $U \in \mathfrak{S}$, we consider the group of automorphisms $\mathcal{A}_U = \{g \in \operatorname{Aut}(G, \mathfrak{S}) \mid g \cdot U = U\}$. For any $U \in \mathfrak{S}$ there is a restriction homomorphism $\theta_U \colon \mathcal{A}_U \to \operatorname{Aut}(\mathfrak{S}_U)$, defined by letting $\theta_U(g)$ act as g on $\mathfrak{S}_U$. By a slight abuse of notation, we say that $\mathcal{A}_U$ (or a subgroup of $\mathcal{A}_U$) acts on $\mathcal{C}U$, when actually it is $\theta_U(\mathcal{A}_U)$ (or the image of the subgroup under θ_U) which acts.

Given an infinite order element $g \in G$ and a $U \in \mathfrak{S}$ for which g is loxodromic with respect to the action of $\mathcal{A}_U$ on $\mathcal{C}U$, we let $\tau_U(g)$ denote the translation length of g in this action.

Lemma 1.8. *Let $(G, \mathfrak{S})$ be a hierarchically hyperbolic group. There exists a constant $T > 0$ such that for every infinite order element $g \in G$ and every $U \in \operatorname{Big}(g)$, we have $\tau_U(g) \geqslant T$.*

Proof. Let $g \in G$ and $U \in \operatorname{Big}(g)$. Thus $\langle g \rangle \subseteq \mathcal{A}_U$, and the induced action of $\langle g \rangle$ on $\mathcal{C}U$ is loxodromic. Since this action has trivial point stabilizers, $\langle g \rangle$ acts acylindrically on $\mathcal{C}U$ by [DHS18, Prop. 2.2]. Moreover, the constants of acylindricity for this action as obtained in [DHS18, Prop. 2.2] are uniform over all elements $g \in G$ for which $U \in \operatorname{Big}(g)$. Given an acylindrical action on a hyperbolic space, it is proven in [Bow08, Lemma 2.2] that there is a uniform lower bound on the translation length of loxodromic elements which depends only on the hyperbolicity constant of the space and the constants of acylindricity. Therefore there is a uniform lower bound on the translation length of all elements $g \in G$ for which $U \in \operatorname{Big}(g)$. Since the action of G on $\mathfrak{S}$ is cofinite, the result follows. □

An element g of an HHG is said to be *irreducible* if $\operatorname{Big}(g)$ consists of only the nest-maximal element of the hierarchically hyperbolic structure. Otherwise, g is said to be *reducible*.

Let $a \in G$ be an infinite order irreducible element. It follows from [BHS17b, Theorem K] that G acts acylindrically on $\mathcal{C}S$. Since a is an infinite order irreducible element, it is not elliptic with respect to this action, and thus it follows by [Bow08, Lemma 2.2] a acts loxodromically on $\mathcal{C}S$. It then follows from the distance formula that $\langle a \rangle$ is a quasi-isometrically embedded copy of $\mathbb{Z}$, and thus yield s a quasi-axis for a which we denote γ_a.

The next lemma follows immediately from thinness of quadrilaterals in the hyperbolic space $\mathcal{C}S$:

Lemma 1.9. *Let $(\mathcal{X}, \mathfrak{S})$ be a hierarchically hyperbolic space and let G be a group acting geometrically on $\mathcal{X}$. Let $a \in G$ be an irreducible element with the property that one of its orbits on $\mathcal{X}$ is quasi-isometrically embedded, and let γ_a be an associated quasi-geodesic axis* *in $\mathcal{X}$. Given any two points $x, y \in \mathcal{X}$, let $x' = \mathfrak{g}_{\gamma_a}(x), y' = \mathfrak{g}_{\gamma_a}(y)$. Then there exist points $\xi, \xi' \in \pi_S([x, y])$ such that the following hold:*

(1) $\mathsf{d}_S(\xi, [x, x']) \sim O(1)$;
(2) $\mathsf{d}_S(\xi', [y, y']) \sim O(1)$;
(3) $[\pi_S(\xi), \pi_S(\xi')] \subset \mathcal{N}_\delta(\pi_S(\gamma_a))$.

For any constant $K \geqslant E$ and any two points $x, y \in \mathcal{X}$, we say $U \in \mathfrak{S}$ is *relevant* (with respect to x, y, K) if $\mathsf{d}_U(x, y) \geqslant K$; if we want to emphasize the constant K, we say that U is *K–relevant* (with respect to x, y). We denote by $\mathbf{Rel}(x, y; K)$ the set of relevant domains.

Lemma 1.10. *Let $(G, \mathfrak{S})$ be a hierarchically hyperbolic group, and suppose $b \in G$ satisfies $\mathrm{Big}(b) \neq \varnothing$. Then there is a uniform constant N such that the following holds: for any $n \geqslant N$, any constant $K \geqslant E$, and any two points $x_1, x_2 \in \mathcal{X}$, if there exists a domain $A \in \mathfrak{S}$ with $A \in \mathbf{Rel}(x_1, x_2; K) \cap \mathbf{Rel}(b^n x_1, b^n x_2; K)$, then $A \not\sqsubseteq U$ for any $U \in \mathrm{Big}(b)$.*

Proof. Let $U \in \mathrm{Big}(b)$. Then b is loxodromic with respect to the action of $\mathcal{A}_U$ on $\mathcal{C}U$, and one of its orbits on $\mathcal{C}U$ is quasi-isometrically embedded. Moreover, by Lemma 1.8, $\tau_U(b) \geqslant T$. Therefore, we can choose a constant N depending only on T and E such that for all $n \geqslant N$, we have $\mathcal{N}_E([x_1, x_2]) \cap \mathcal{N}_E([b^n x_1, b^n x_2]) = \varnothing$.

Now, suppose by way of a contradiction that $A \subsetneq U$ for some $U \in \mathrm{Big}(b)$. Then, by the bounded geodesic image axiom and the fact that $A \in \mathbf{Rel}(x_1, x_2; K)$ we have that $\rho^A_U \subset \mathcal{N}_E(\pi_U([x_1, x_2]))$; similarly, $A \in \mathbf{Rel}(b^n x_1, b^n x_2; K)$ yields that $\rho^A_U \subset \mathcal{N}_E([b^n x_1, b^n x_2])$. This contradicts the fact we showed above that $\mathcal{N}_E([x_1, x_2]) \cap \mathcal{N}_E([b^n x_1, b^n x_2]) = \varnothing$. □

2. Proof of Theorem A

We break the proof of the Theorem into two parts depending on whether or not the elements are irreducible. In the first case, we assume a and b are irreducible elements, i.e., $\mathrm{Big}(a) = \{S\}$. This result is then used explicitly in one of the cases in the second part; additionally, the analysis in the latter case follows a similar framework.

In the second case we assume a is reducible. Note that, as in the statement of the theorem, we also have an additional hypothesis that $\mathrm{Big}(a)$ is a maximal collection of pairwise orthogonal domains.

2.1. Irreducible case.

Suppose $a, b \in G$ are two infinite order irreducible elements and there exists an element $g \in G$ such that $ga = bg$. (Note, since these two elements are conjugate, irreducibility of b follows from just assuming irreducibility of a.) Let γ_a and γ_b denote their respective quasi-axes in $\mathcal{C}S$. We note that a and b preserve their respective quasi-axes in $\mathcal{C}S$. We also fix a constant R large enough so that for all $U \neq S$, we have $\mathsf{d}_U(1, a) < R$ and $\mathsf{d}_U(1, b) < R$, which exists by [DHS17, Lemma 6.6], since $\mathrm{Big}(a) = \mathrm{Big}(b) = \{S\}$.

Applying Lemma 1.9 to the points $1, b$, and the quasi-axis γ_b (respectively, $1, a$, and γ_a) we obtain $\xi, \xi' \in [1, b]$ (respectively, $\nu, \nu' \in [1, a]$) as provided by the lemma.

We break up our analysis for irreducible elements into two cases. In the first, the distance term in S makes up a definite percentage of $\mathsf{d}_G(1, g)$. If the first case doesn't hold, then a definite percentage is made up by all domains except S, which is the second case. We look at the subsurfaces distance only above the threshold R; if all the projections are less than R, then the uniqueness axiom provides a uniform bound on $\mathsf{d}_G(1, g)$.

Case 1. $\mathsf{d}_G(1, g) \asymp \mathsf{d}_S(1, g)$.

Since $g\gamma_a = \gamma_b$, it follows that after premultiplying g by a power of b we have $\mathsf{d}_S(\xi, g\nu) \leqslant \tau(b) + \delta + K$, where $\tau(b) = \tau_S(b)$ is the translation length of b and K is a uniform constant

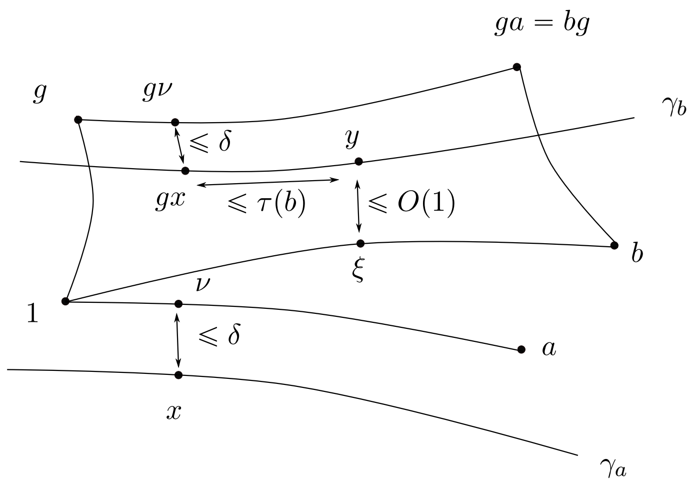

FIGURE 1. Set-up for irreducible case.

depending on the HHS constant on projection bounds (see Figure 1). Then,

$$\begin{aligned}\mathsf{d}_S(1,g) &\leqslant \mathsf{d}_S(1,\xi) + \mathsf{d}_S(\xi, g\nu) + \mathsf{d}_S(g\nu, g)\\ &\leqslant \mathsf{d}_S(1,\xi) + \tau(b) + \delta + K + \mathsf{d}_S(g\nu, g)\\ &\leqslant \mathsf{d}_S(1,a) + 2\mathsf{d}_S(1,b) + K,\end{aligned}$$

where the last inequality follows from the fact that $\mathsf{d}_S(1,b) \geqslant \tau(b)$.

Case 2. $\mathsf{d}_G(1,g) \asymp \sum_{Y \subsetneq S} \{\!\{\mathsf{d}_Y(1,g)\}\!\}_R$.

Since $\mathrm{Big}(b) = \{S\}$, while all domains in $\mathbf{Rel}(1,g;R)$ are proper subsets of S, it follows from Lemma 1.10 that there exists a uniform choice of an integer N such that, by replacing b with b^N, we may assume $\mathbf{Rel}(1,g;R) \cap \mathbf{Rel}(b,bg;R) = \varnothing$. Therefore, if $Y \in \mathbf{Rel}(1,g;R)$, then the triangle inequality yields:

$$\mathsf{d}_Y(1,g) \leqslant \mathsf{d}_Y(1,b) + R + \mathsf{d}_Y(g,bg).$$

By raising the threshold to $3R$, the above ensures that every domain in $\mathbf{Rel}(1,g;3R)$ is in $\mathbf{Rel}(1,b;R)$ or $\mathbf{Rel}(g,bg;R)$. Hence, if $Y \in \mathbf{Rel}(1,g;3R)$:

$$\mathsf{d}_Y(1,g) \leqslant 2(\mathsf{d}_Y(1,b) + \mathsf{d}_Y(g,bg)).$$

This equation applied term-wise, together with the distance formula applied to each of $\sum_{Y \subsetneq S} \{\!\{\mathsf{d}_Y(1,g)\}\!\}_{3R}$, $\sum_{Y \subsetneq S} \{\!\{\mathsf{d}_Y(1,b)\}\!\}_R$, and $\sum_{Y \subsetneq S} \{\!\{\mathsf{d}_Y(g,bg)\}\!\}_R$, yields a constant K so that:

$$\mathsf{d}_G(1,g) \leqslant K(\mathsf{d}_G(1,b) + \mathsf{d}_G(g,bg)) + K.$$

Since $bg = ga$, we have $\mathsf{d}_G(g,bg) = \mathsf{d}_G(1,a)$ and, thus, the desired bound on $\mathsf{d}_G(1,g)$.

2.2. **Reducible case.** Let $a, b \in G$ be two infinite order elements and let $g \in G$ be such that $ga = bg$. For this case we assume that $\mathrm{Big}(a)$ is a nest-maximal collection of pairwise orthogonal domains.

As $\mathrm{Big}(a)$ and $\mathrm{Big}(b)$ are finite sets (whose size is bounded by the complexity of $(G, \mathfrak{S})$) which are stabilized by a and b, respectively, it follows that by replacing a and b by a sufficiently high (uniform) power, we can assume that $\mathrm{Big}(a)$ and $\mathrm{Big}(b)$ are fixed pointwise by a and b, respectively. As a and b are conjugate, it follows that for each $U \in \mathrm{Big}(a)$ we have $gU \in \mathrm{Big}(b)$.

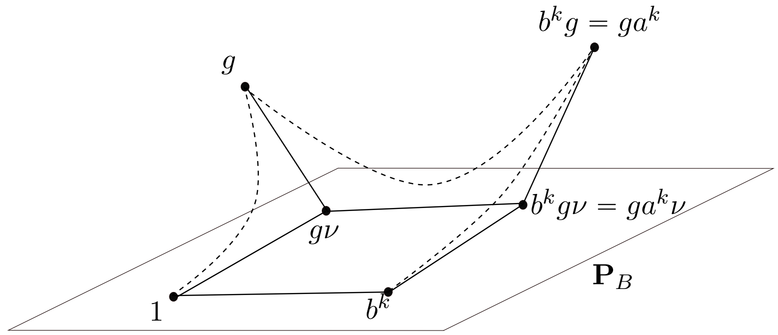

FIGURE 2. Conjugate elements a and b, with conjugator g, in G. Solid segments are hierarchy paths, while dotted segments are geodesics.

Fix some $A \in \mathrm{Big}(a)$ and let $B = gA \in \mathrm{Big}(b)$. Consider $\mathbf{P}_A, \mathbf{P}_B \subseteq G$, the standard product regions associated to A and B, respectively. Without loss of generality, we may assume $1 \in \mathbf{P}_B$. Since a and b are conjugate elements, $g \cdot \mathbf{P}_A = \mathbf{P}_B$. Moreover, since $\mathrm{Big}(b)$ is a pairwise orthogonal set, for every $Z \in \mathrm{Big}(b)$ we have $\mathbf{F}_Z \subseteq \mathbf{P}_B$.

Fix a threshold R larger than any of the HHS constants for G. Let $\nu = \mathfrak{g}_{\mathbf{P}_A}(1)$, so that $g\nu = g\mathfrak{g}_{\mathbf{P}_A}(1) = \mathfrak{g}_{\mathbf{P}_B}(g)$ (see Figure 2). We begin with the triangle inequality:

$$\mathsf{d}_G(1, g) \leqslant \mathsf{d}_G(1, g\nu) + \mathsf{d}_G(g, g\nu). \tag{2}$$

Let $\mathcal{Y} = \{Y \in \mathfrak{S} \mid \text{there exists } Z \in \mathrm{Big}(b) \text{ satisfying either } Y \pitchfork Z \text{ or } Y \sqsupseteq Z\}$. For conciseness, we abbreviate the statement $U \in \mathfrak{S} \backslash \mathcal{Y}$ by writing $U \notin \mathcal{Y}$; note that these are the domains which are nested in or equal to some $Z \in \mathrm{Big}(b)$ (they can't be orthogonal to all $Z \in \mathrm{Big}(b)$, since by construction $\mathrm{Big}(b)$ is maximal). By Lemma 1.6, it follows that for a given threshold R, every $U \in \mathfrak{S}$ appearing in the distance formula for $\mathsf{d}_G(1, g)$ is in exactly one of the follow sets: $\mathbf{Rel}(g, g\nu; R)$, which occurs exactly when $U \in \mathcal{Y}$; or $\mathbf{Rel}(1, g\nu; R)$, which occurs exactly when $U \notin \mathcal{Y}$. Therefore, combining the distance formula and this observation, up to uniformly bounded multiplicative and additive constants (which we suppress from the notation), the right-hand side of (2) can be written as:

$$\sum_{Y \in \mathcal{Y}} \{\!\{\mathsf{d}_Y(1, g)\}\!\}_R + \sum_{Y \notin \mathcal{Y}} \{\!\{\mathsf{d}_Y(1, g)\}\!\}_R. \tag{3}$$

Up to a multiplicative constant of two, we can replace the sum in (3) by the max of $\sum_{Y \in \mathcal{Y}} \{\!\{\mathsf{d}_Y(1, g)\}\!\}_R$ and $\sum_{Y \notin \mathcal{Y}} \{\!\{\mathsf{d}_Y(1, g)\}\!\}_R$. We separately analyze the cases where one or the other of the summations is the maximum.

Case 1. $\mathsf{d}_G(1, g) \asymp \sum_{Y \in \mathcal{Y}} \{\!\{\mathsf{d}_Y(1, g)\}\!\}_R$.

Consider some $U \in \mathcal{Y}$ which is R–relevant for $1, g$. Since $1 \in \mathbf{P}_B$, it follows from Lemma 1.6 that $U \in \mathbf{Rel}(g, g\nu; R)$. As g is loxodromic for the action on $\mathcal{C}Z$ for all $Z \in \mathrm{Big}(b)$, we have $\mathsf{d}_Z(g\nu, b^kg\nu) \geqslant k\tau_Z(b) \geqslant kT$, where T is the constant provided by Lemma 1.8. Thus there exists a constant depending only on the constants of the HHS structure, such that for all k larger than this constant and all $Z \in \mathrm{Big}(b)$, we have $Z \in \mathbf{Rel}(g\nu, b^kg\nu; R)$. Note that this implies Z is R–relevant for g, b^kg and $g, b^k\nu$, as well.

Each $U \in \mathcal{Y}$ satisfies the conditions for at least one of the following two subcases. In either subcase we will give a bound for $\mathsf{d}_U(1, g)$. In the first subcase we prove the following for a constant K depending only on the HHS constants: $\mathsf{d}_U(1, g) \leqslant K\mathsf{d}_U(1, b) + \mathsf{d}_U(g, b^Kg) + K$.

In the second subcase we prove: $\mathsf{d}_U(1, g) \leqslant K\mathsf{d}_U(1, b) + K\mathsf{d}_U(1, a) + K$. Once we establish these bounds for each $\mathcal{C}U$, then we apply the distance formula for a larger threshold so that the additive error drops out of the coordinate-wise summation. For the terms from the first subcase we also use the fact that $\mathsf{d}_G(g, b^K g) = \mathsf{d}_G(g, ga^K) = \mathsf{d}_G(1, a^K)$. Together, this establishes, for a larger value of K, that:

$$\mathsf{d}_G(1,g) \preceq K\mathsf{d}_G(1,b) + K\mathsf{d}_G(1,a) + K. \tag{4}$$

Note that through the proof we will increase the constant represented by K several times, but at each step the new value will still be a *uniform* constant in that it will depend only on the constants of the underlying HHS and not on the particular choice of elements. In particular, we start by assuming that K is larger than the constants in the HHS axioms, so that we can use the consistency axiom, etc, as need.

Case 1a. There exists $Z \in \mathrm{Big}(b)$ such that $U \pitchfork Z$.

Consider the domain $b^k U \in \mathfrak{S}$, which is relevant for $b^k g\nu, b^k$. Note that $\mathsf{d}_Z(\rho^U_Z, \rho^{b^k U}_Z) \asymp \mathsf{d}_Z(\rho^U_Z, b^k \rho^U_Z) \succeq k\tau_Z(b) \geqslant kT$. Since b fixes $\mathrm{Big}(b)$ pointwise, $b^k Z = Z$ and so $b^k U \pitchfork Z$. Thus, for any constant C, there exists a larger uniform choice of K such that for all $k \geqslant K$, we have $\mathsf{d}_Z(\rho^U_Z, \rho^{b^k U}_Z) \geqslant C$. Accordingly, we choose C large enough (depending only on the HHS constants), so that applying [BHS15, Lemma 2.1] yields $U \pitchfork b^k U$ for all $k \geqslant K$. For the rest of this subcase we assume $k \geqslant K$.

We will show that $U \notin \mathbf{Rel}(b^k, b^k g; R)$. Notice that by the assumption of this subcase, this is equivalent to showing that $U \notin \mathbf{Rel}(b^k g\nu, b^k g; R)$. Let $\mathbf{Rel}_{max}(g, b^k g; R)$ be any subset of $\sqsubseteq$–incomparable elements of $\mathbf{Rel}(g, b^k g; R)$ that includes U, Z, and $b^k U$. By [BHS15, Proposition 2.8], for any $x, y \in \mathcal{X}$, there is a partial order $\preceq$ on $\mathbf{Rel}_{max}(x, y; R)$ defined as follows: given $U, V \in \mathbf{Rel}_{max}(x, y; R)$, $U \preceq V$ if $U = V$ or if $U \pitchfork V$ and $\mathsf{d}_U(\rho^V_U, y) \leqslant \kappa$.

As above, we can uniformly choose C and K large enough so that $\mathsf{d}_Z(\rho^U_Z, b^k g)$ is large enough to apply the consistency axiom. Then, in the partial order on $\mathbf{Rel}_{max}(g, b^k g; R)$ it follows that $U \preceq Z \preceq b^k U$. Thus $\rho^{b^k U}_U \sim \pi_U(b^k g) \sim \rho^Z_U$.

Moreover, in $\mathcal{C}Z$, we have $\pi_Z(b^k g) \sim \pi_Z(b^k g\nu)$. Since $\mathsf{d}_Z(\rho^U_Z, b^k g\nu)$ is large enough to apply the consistency axiom, it follows that $\mathsf{d}_U(\rho^Z_U, b^k g\nu)$ is uniformly bounded above, also. Combining this with the above fact that $\pi_U(b^k g) \sim \rho^Z_U$, it follows that $\mathsf{d}_U(b^k g, b^k g\nu)$ is uniformly bounded above, and therefore (after possibly increasing R), we have $U \notin \mathbf{Rel}(b^k, b^k g; R)$. Since this means $\mathsf{d}_U(b^K, b^K g) < R$, we have:

Therefore,

$$\mathsf{d}_U(1,g) \leqslant \mathsf{d}_U(1,b^K) + \mathsf{d}_U(g, b^K g) + R \leqslant K\mathsf{d}_U(1,b) + \mathsf{d}_U(g, b^K g) + R.$$

Case 1b. There exists $Z \in \mathrm{Big}(b)$ with $U \sqsupsetneq Z$.

By the definition of $\mathrm{Big}(b)$ and the HHS nesting axiom, there exists uniform constants $E, K \geqslant 0$ such that for all $k \geqslant K$, we have $\rho^Z_U \subseteq N_E([1, b^k])$ and $\rho^A_U \subseteq N_E([1, a^k])$. Therefore, $\mathsf{d}_U(1, \rho^Z_U) \leqslant \mathsf{d}_U(1, b^k) + E$ and $\mathsf{d}_U(1, \rho^A_U) \leqslant \mathsf{d}_U(1, a^k) + E$. Additionally, since $ga = bg$, we have that $g\rho^A_U$ is coarsely equal to ρ^Z_U (we let E denote this uniform choice of coarseness constant, as well). Thus

$$\mathsf{d}_U(1,g) \leqslant \mathsf{d}_U(1,\rho^Z_U) + \mathsf{d}_U(\rho^Z_U, g\rho^A_U) + \mathsf{d}_U(g\rho^A_U, g) \leqslant K\mathsf{d}_U(1,b) + K\mathsf{d}_U(1,a) + 3E.$$

Case 2. $\mathsf{d}_G(1,g) \asymp \sum_{Y \notin \mathcal{Y}} \{\!\{\mathsf{d}_Y(1,g)\}\!\}_R$.

In this case, $\mathsf{d}_G(1,g) \asymp \mathsf{d}_G(1, g\nu)$. Our goal will be to bound $\mathsf{d}_G(1, g\nu)$ by a linear function of $\mathsf{d}_G(1,b) + \mathsf{d}_G(g\nu, bg\nu)$. Since $\mathrm{Big}(b)$ is a maximal collection of pairwise disjoint domains, any $U \in \mathbf{Rel}(1, g; R)$ satisfying $U \notin \mathcal{Y}$ is either in $\mathrm{Big}(b)$, or nested in an element of $\mathrm{Big}(b)$. Below we obtain the appropriate bound in $\mathcal{C}U$, as in Case (1), for each of these two subcases separately. Applying the distance formula will then yield a uniform constant K' for which:

$$\mathsf{d}_G(1,g) \preceq K'\mathsf{d}_G(1,b) + K'\mathsf{d}_G(1,a) + K'. \tag{5}$$

Case 2a. $U \subsetneq Z$ for some $Z \in \mathrm{Big}(b)$.

By Lemma 1.10, there exists a uniform constant K such that for all $k \geqslant K$, the set of R–relevant domains for $1, g$ does not intersect the set of R–relevant domains for $b^k, b^k g$. Since $U \in \mathbf{Rel}(1, g; R)$, we have $U \notin \mathbf{Rel}(b^k, b^k g; R)$ and thus $\mathsf{d}_U(b^k, b^k g) < R$. Hence, the triangle inequality yields:

$$\mathsf{d}_U(1, g) \leqslant \mathsf{d}_U(1, b^K) + \mathsf{d}_U(g, b^K g) + R \leqslant K\mathsf{d}_U(1, b) + \mathsf{d}_U(g, b^K g) + R.$$

Case 2b. $U = Z$ for some $Z \in \mathrm{Big}(b)$.

In this case, the result from the irreducible case for $\theta_U(b)$ (applied in the HHS $\mathbf{F}_U$) provides a uniform constant K for which we have:

$$\mathsf{d}_U(1, g) \leqslant K\mathsf{d}_U(1, b) + K\mathsf{d}_U(g, bg) + K.$$

This completes the proof of the theorem. □

3. Special CAT(0) cubical groups

In [BHS17b] it was proven that compact CAT(0) cube complexes which are special, in the sense of Haglund–Wise [HW08], are hierarchically hyperbolic. Below we recall some facts about the particular hierarchically hyperbolic structure obtained in [BHS17b]; we then this this structure to proof Theorem B.

3.1. **Background.** A compact special CAT(0) cube complex, admits the following hierarchically hyperbolic structure, as established in [BHS17b, Theorem G]. We describe the structure in the case of a right-angled Artin group A_Γ; this implicitly describes the structure in the more general case of a special cube complex, as these are each a convex subset of some right-angled Artin group. The set $\mathfrak{S}$ consists of the collection of non-empty convex subcomplexes coming from left cosets of the form A_Δ where Δ is a subgraph of Γ, considered up to the equivalence of *parallelism*. Two convex subcomplexes F, F' of $\mathcal{X}$ are *parallel* if for all hyperplanes H, we have $H \cap F \neq \varnothing$ if and only if $H \cap F' \neq \varnothing$. Since elements of $\mathfrak{S}$ are equivalence classes of convex subsets, we often write $[U] \in \mathfrak{S}$ to denote an equivalence class and $U \in [U]$ to denote a representative.

Throughout the rest of this section we fix a compact special CAT(0) cubical group, G, and consider the hierarchically hyperbolic structure described above, which we denote $(G, \mathfrak{S})$. The following two lemmas provide information about gate maps and product regions in the CAT(0) setting.

Lemma 3.1 ([BHS17b])**.** *Let $(G, \mathfrak{S})$ be a CAT(0) cubical group with its hierarchically hyperbolic structure and let $\mathcal{Y} \subset G$ be a convex subcomplex of G. Then there is a cubical map $\mathfrak{g}_\mathcal{Y} \colon G \to \mathcal{Y}$ which, for every $g \in G$, assigns the unique 0–cube $y \in \mathcal{Y}$ which is closest to g.*

For notational simplicity, when $\mathcal{Y} = \mathbf{F}_U$ for some $[U] \in \mathfrak{S}$, we write $\mathfrak{g}_U \colon G \to \mathbf{F}_U$.

Lemma 3.2 ([BHS17b, Lemma 2.4])**.** *Let $(G, \mathfrak{S})$ be a CAT(0) cubical group with its hierarchically hyperbolic structure and let $[U] \in \mathfrak{S}$. Then there is a cubical embedding $\mathbf{F}_U \times \mathbf{E}_U \hookrightarrow G$ with convex image.*

Remark 3.3. In a right-angled Artin group, two elements commute if and only if all the generators in a factorization of one of the elements commute with all the generators in a factorization of the other element. Hence, in the Salvetti complex of a right-angled Artin group, we have that two elements span a periodic plane if and only if they commute. Similarly, if a group is special (in the sense of Haglund-Wise), then it embeds as a subgroup of a right-angled Artin group, so inherits this property as well. Further, if a group is virtually special, then, up to taking powers, two elements commute if and only if the product of their axes spans a periodic plane. Note, that non-special CAT(0) cube-complexes need not have this property, as is the case, for example, with Wise's anti-torus [Wis07].

Remark 3.4. Let $\mathfrak{S}'$ be the (finite) fundamental domain for the action of G on $\mathfrak{S}$. If $[U] \in \mathfrak{S}'$, then there is a representative $U \in [U]$ such that $1 \in \mathbf{F}_U$. Additionally, $g \in \mathrm{Stab}(\mathbf{F}_U)$ if and only if the vertex labeled by g is in $\mathbf{F}_U$. In this case, Lemma 3.2 and Remark 3.3 imply that, for a virtually special group, if the vertices labeled by g and h are in $\mathbf{F}_U$ and $\mathbf{E}_U$, respectively, then, up to taking powers, g and h commute in G.

3.2. **Proof of Theorem B.** By raising a and b to a uniform power, we may assume that they are elements of a compact special group CAT(0) cubical group. Let $\mathfrak{S}'$ be a finite fundamental domain for the action of G on $\mathfrak{S}$, and for each $[U] \in \mathfrak{S}'$, fix the representative $U \in [U]$ such that $1 \in \mathbf{F}_U$.

If $\mathrm{Big}(b)$ is a maximal collection of pairwise orthogonal domains in $\mathfrak{S}$, then the result follows by Theorem A, so we may suppose this is not the case. Extend $\mathrm{Big}(b)$ to a maximal collection $\mathcal{O}$ of pairwise orthogonal domains in $\mathfrak{S}$, and fix $B \in \mathrm{Big}(b)$. We consider the associated product region $\mathbf{P}_B$ in G.

For each $[V] \in \mathcal{O}$, fix the representative $V \in [V]$ such that $\mathfrak{g}_{\mathbf{P}_B}(1) \in \mathbf{F}_V$. For every $[Y] \in \mathfrak{S}\backslash\mathcal{O}$, fix any representative $Y \in [Y]$. We fix a transversal $\mathcal{T}$ with respect to $\{\mathrm{Stab}(\mathbf{F}_U) \mid [U] \in \mathfrak{S}'\}$ as follows. For each $[U] \in \mathfrak{S}$, choose the coset representative of $\mathrm{Stab}(\mathbf{F}_U)$ to be the label of the vertex $\mathfrak{g}_U(1)$. Notice that if $[V] \in \mathcal{O}$, then since $\mathbf{P}_B$ is a product region, $\mathfrak{g}_{\mathbf{P}_B}(1) = \mathfrak{g}_V(1)$.

By replacing b with a uniform power, we may assume that b acts trivially on $\mathcal{C}W$ for every domain $[W] \in \mathrm{Big}(b)^{\perp}$. Thus $b \in \mathrm{Stab}\left(\prod_{[V]\in\mathrm{Big}(b)} \mathbf{F}_V\right)$.

Let $\mathcal{O}\backslash \mathrm{Big}(b) = \{[V_1], \dots, [V_m]\}$. By our choice of transversal, there is some $t \in \mathcal{T}$ such that for each i, we have $V_i = tV_i'$ for some $[V_i'] \in \mathfrak{S}'$. By Lemma 3.1, for each i there is a unique 0–cell $v_i \in \mathbf{F}_{V_i}$ such that $v_i = \mathfrak{g}_{V_i}(g)$. The label of v_i is tg_i' for some $g_i' \in \mathrm{Stab}(\mathbf{F}_{V_i'})$. Define

$$g_{V_i} := tg_i't^{-1} \in \mathrm{Stab}(\mathbf{F}_{V_i}).$$

For each i and all $U \in \mathcal{O}\backslash\{\mathrm{Big}(b) \cup V_i\}$, we have $\mathsf{d}_U(t, tg_i') \leqslant K$ for some uniform constant K, since the path from t to tg_i' lies in $\mathbf{F}_{V_i}$ and $V_i \perp U$. Also, since $t = \mathfrak{g}_{V_i}(1)$, the only domains that contribute to $\mathsf{d}_G(1,t)$ are those which are transverse to V_i or in which V_i is nested, thus we also have $\mathsf{d}_U(1,t) \leqslant K$. Thus we have:

$$\mathsf{d}_U(1, g_{V_i}) \leqslant \mathsf{d}_U(1,t) + \mathsf{d}_U(t, tg_i') + \mathsf{d}_U(1, t^{-1}) \leqslant 3K.$$

Now, $b \in \mathrm{Stab}\left(\prod_{W\in\mathrm{Big}(b)} \mathbf{F}_W\right) \subseteq \mathrm{Stab}(\mathbf{E}_{V_i})$ and $g_{V_i} \in \mathrm{Stab}(\mathbf{F}_{V_i})$; thus, given our choice of transversal, there exists $b' \in \mathrm{Stab}(\mathbf{E}_{V_i'})$ such that $b = tb't^{-1}$. Moreover, by Remark 3.4, we have $[b', g_{V_i}'] = 1$, and so $[b, g_{V_i}] = 1$ for each i. Therefore, $g_i = g_{V_i}^{-1} g$ also conjugates a to b.

In this manner, we obtain a conjugator

$$g'' = g_{V_1}^{-1} \cdots g_{V_m}^{-1} g$$

such that $\mathsf{d}_U(1, g'') \leqslant 3K$ for each $U \in \mathcal{O}\backslash \mathrm{Big}(b)$. Notice that since the g_{V_i} commute, the element g'' does not depend on the ordering of the domains in $\mathcal{O}\backslash \mathrm{Big}(b)$.

We claim that $\mathsf{d}_G(1, g'') \leqslant K(|a| + |b|) + C$. To show this, we follow the argument as in the proof of Theorem A. The bound in Cases 1a, 1b, 2a, and 2b holds exactly as in the proof of Theorem A. However, there is now an additional case which we must now treat. In the notation of the proof of Theorem A, this new case is: $U \in \mathcal{Y} \cap \mathbf{Rel}(1, g; R)$.

Case 2c. $U \in \mathcal{Y} \cap \mathbf{Rel}(1, g; R)$**.**

Suppose this U is orthogonal to every element of $\mathrm{Big}(b)$. Then $U \in \mathcal{O}\backslash \mathrm{Big}(b)$, and $\mathsf{d}_U(1, g'') \leqslant 3K$. By increasing the threshold above $3K$, we ensure that these domains do not contribute to the distance formula.

This completes the proof. □

4. Acylindrically hyperbolic groups

The action of a group G on a metric space X is *acylindrical* if for all $\varepsilon \geqslant 0$, there exist constants $R(\varepsilon), N(\varepsilon) \geqslant 0$ such that for all $x, y \in X$ satisfying $\mathsf{d}_X(x, y) \geqslant R(\varepsilon)$, $|\{g \in G \mid \mathsf{d}_X(x, gx) \leqslant \varepsilon \text{ and } \mathsf{d}_X(y, gy) \leqslant \varepsilon\}| \leqslant N(\varepsilon)$. A group is *acylindrically hyperbolic* if it admits an acylindrical action on a non-elementary hyperbolic space. An element $g \in G$ is *generalized loxodromic* if there exists an acylindrical action of G on a hyperbolic space with respect to which g is loxodromic.

Let G be an acylindrically hyperbolic group, let $g \in G$ be a generalized loxodromic element, and let $E(g)$ be the largest virtually cyclic subgroup containing $\langle g\rangle$. Consider the set $\mathbf{Y} = \{aE(g) \mid a \in G\}$. Then if K is a sufficiently large constant, Bestvina, Bromberg, and Fujiwara in [BBF15] give a construction of a quasi-tree $\mathcal{P}_K(\mathbf{Y})$ on which G acts. This construction was later refined in [BBFS17]. In both [BBF15] and [BBFS17], the construction of the quasi-tree was done in a much more general setting, but we will only need this particular case.

Our goal in this section will be to establish the following bound on shortest conjugators in acylindrically hyperbolic groups.

Theorem 4.1. *Let G be an acylindrically hyperbolic group and $h \in G$ a generalized loxodromic element. Let $\mathbf{Y} = \{cE(h) \mid c \in G\}$, and let K be sufficiently large so that the projection complex $T = \mathcal{P}_K(\mathbf{Y})$ is a quasi-tree. Then there exists a function $f\colon \mathbb{Z} \to \mathbb{Z}$ and a constant R such that for any two conjugate elements a, b that are loxodromic with respect to the action of G on T, there exists $g \in G$ such that $ga = bg$ and*

$$|g| \leqslant f(R|a| + R|b|) + |a| + |b|.$$

4.1. **Background.** We briefly recall the construction of the quasi-tree we will be using and refer the reader to [BBFS17] for further details. Considering $aE(g)$ as a subset of the Cayley graph of G with respect to a fixed generating set, for any $a, b \in G$, one can define a nearest point projection of $aE(g)$ onto $bE(g)$, which will have uniformly bounded diameter. Define metrics d_a on $\mathbf{Y}$ by letting $\mathsf{d}_a(bE(g), cE(g))$ be the diameter of the union of the nearest point projections of $bE(g)$ and $cE(g)$ onto $aE(g)$. For any $K \geqslant 0$, we define $\mathcal{P}_K(\mathbf{Y})$ to be the graph whose vertices are $\mathbf{Y}$, and in which two vertices $aE(g)$ and $bE(g)$ are connected by an edge if $\mathsf{d}_Y(aE(g), bE(g)) \leqslant K$ for all $Y \in \mathbf{Y}$. For sufficiently large K, the graph $T = \mathcal{P}_K(\mathbf{Y})$ is a quasi-tree. The fundamental domain of this action is a single vertex, Y_0, whose stabilizer is the subgroup $E(g)$. Moreover, by [BBFS17, Theorem 5.6], the action of G on T is acylindrical. The existence of an acylindrical action on a quasi-tree was first proven by Balasubramanya in [Bal17], using a slightly different construction.

We fix K large enough so that $\mathcal{P}_K(\mathbf{Y})$ is a quasi-tree. For any $\varepsilon \geqslant 0$, let $R(\varepsilon), N(\varepsilon)$ be the constants of acylindricity for the action of G on $\mathcal{P}_K(\mathbf{Y})$.

Given $X, Y \in \mathbf{Y}$, let $\mathbf{Y}_K(X, Y) = \{Z \in \mathbf{Y} \mid \mathsf{d}_Z(X, Y) \geqslant K\}$. By [BBFS17, Proposition 2.3], one can define a total order on $\mathbf{Y}_K(X, Y) \cup \{X, Y\}$, which we denote by $<$. The path $Y_K(X, Y) \cup \{X, Y\} = \{X < Z_1 < Z_2 < \cdots < Z_k < Y\}$ is called a *standard path* in $\mathcal{P}_K(\mathbf{Y})$. Note that in the quasi-tree, we use the term "path," to refer to a sequence of vertices each distance one from the previous. We will need the following results about standard paths.

Lemma 4.2. [BBFS17, Lemma 3.6] *If K is sufficiently large, the following holds. Given $X, Y, Z \in \mathbf{Y}$, the union $\mathbf{Y}_K(X, Y) \cup \mathbf{Y}_K(Y, Z)$ contains all but at most two elements of $\mathbf{Y}_K(X, Z)$, and if there are two such elements they are consecutive.*

Lemma 4.3. [BBFS17, Corollary 3.8] *Standard paths are uniform quasi-geodesics.*

It was proven in [BBFS17, Theorem 3.5] that T is a quasi-tree by showing that it satisfies the following version of Manning's bottleneck property:

Lemma 4.4 ([BBFS17])**.** *Let X, Z be vertices in T. Consider the standard path $\mathbf{Y}_K(X,Z) \cup \{X, Z\} = \{X < X_1 < \cdots < X_k < Z\}$ from X to Z. Then for any path $X = Y_0, \ldots, Y_n = Z$ from X to Z and any $i \in \{1, \ldots, k\}$, there exists $j \in \{0, \ldots, n\}$ such that $\mathsf{d}_T(X_i, Y_j) \leqslant 2$.*

We also need the following generalization of [MSW11, Proposition 2.2]. Given the action of a group G on a metric space X and a point $x \in X$, we say $\mathrm{Stab}_\eta(x) = \{g \in G \mid \mathsf{d}_X(x, gx) \leqslant \eta\}$ is the *η–quasi-stabilizer* of x, or simply the *quasi-stabilizer* of x if the value of η is not important. Notice that in general, $\mathrm{Stab}_\eta(x)$ is not a subgroup of G.

Lemma 4.5. *Let G be a finitely generated group acting on a metric space X, and let $x, y \in X$. Then for any $\varepsilon \geqslant 0$ and any $r \geqslant 0$, there exists a constant $r' \geqslant 0$ such that*

$$N_r(\mathrm{Stab}_\eta(x)) \cap N_r(\mathrm{Stab}_\eta(y)) \subseteq N_{r'}(\mathrm{Stab}_\eta(x) \cap \mathrm{Stab}_\eta(y)).$$

The proof closely follows that of [MSW11, Proposition 2.2]; since it is short we include it here for completeness.

Proof. Consider the word metric on G with respect to some fixed finite generating set, S. Fix $\eta, r \geqslant 0$ and let $x, y \in X$. Given any element $t \in N_r(\mathrm{Stab}_\eta(x)) \cap N_r(\mathrm{Stab}_\eta(y)) \subset \Gamma$, it follows immediately that there exists an element $h_t \in \mathrm{Stab}_\eta(x)$ and $g_t \in \mathrm{Stab}_\eta(y)$ so that the element $\gamma_t = h_t^{-1} g_t$ satisfies $\|\gamma_t\|_S \leqslant 2r$.

Consider all elements $\gamma \in G$ such that $\|\gamma\|_S \leqslant 2r$ and such that the equation $\gamma = h^{-1}g$ has a solution $(h, g) \in \mathrm{Stab}_\eta(x) \times \mathrm{Stab}_\eta(y)$. Choose a solution (h_γ, g_γ) so that $\|h_\gamma\|_S = N_\gamma$ is minimal. Let $N = \max_\gamma N_\gamma$.

Setting $\gamma = \gamma_t$, we have

$$\gamma = h_\gamma^{-1} g_\gamma = h_t^{-1} g_t.$$

Since $h_t, h_\gamma \in \mathrm{Stab}_\eta(x)$, it follows that $h_t h_\gamma^{-1} \in \mathrm{Stab}_{2\eta}(x)$; similarly, $g_t g_\gamma^{-1} \in \mathrm{Stab}_{2\eta}(y)$. Thus

$$h_t h_\gamma^{-1} (= g_t g_\gamma^{-1}) \in \mathrm{Stab}_{2\eta}(x) \cap \mathrm{Stab}_{2\eta}(y).$$

Moreover, since $h_t = (g_t g_\gamma^{-1}) h_\gamma$ it follows that h_t is within distance N of $\mathrm{Stab}_{2\eta}(x) \cap \mathrm{Stab}_{2\eta}(y)$. Therefore, the element t is within distance $N + r$ of an element of $\mathrm{Stab}_{2\eta}(x) \cap \mathrm{Stab}_{2\eta}(y)$. Setting $r' = N + r$ completes the proof. □

Before giving a proof of the main result of this section we establish a preliminary lemma.

Lemma 4.6. *Let $\varepsilon \geqslant 0$. For any $M \geqslant 0$ there exists a constant $f(M)$ satisfying the following. Consider any point $y \in T$ and the set:*

$$V_\varepsilon(y) = \{g \in G \mid g \in \mathrm{Stab}_\varepsilon(Y_0) \text{ and } \mathsf{d}_S(g, \pi^{-1}(N_\varepsilon(y))) \leqslant M\}.$$

If $\mathsf{d}_T(Y_0, y) \geqslant R(6\varepsilon)$, then either $V_\varepsilon(y) = \varnothing$ or the diameter of $V_\varepsilon(y)$ (in the word metric on G with respect to a fixed finite generating set) is at most $f(M)$.

Proof. Throughout the proof we will use d_G to denote the word metric on G with respect to a fixed finite generating set; we write d_T to denote distance in the quasi-tree T.

Suppose $1 \notin V_\varepsilon(y)$ and consider some fixed $k \in V_\varepsilon(y)$, so that $1 \in k^{-1}(V_\varepsilon(y))$. Any element $\alpha \in k^{-1}(V_\varepsilon(y))$ satisfies $\alpha \in \mathrm{Stab}_{2\varepsilon}(Y_0)$, and thus there exists $\beta \in \pi^{-1}(N_\varepsilon(k^{-1}y))$ for which $d_S(\alpha, \beta) \leqslant M$. Hence, $\alpha \in V_{2\varepsilon}(k^{-1}y)$. Thus, either $V_\varepsilon(y) = \varnothing$ and the lemma is proven, or the diameter of $V_\varepsilon(y)$ is bounded by the diameter of $V_{2\varepsilon}(k^{-1}y)$, a set which contains 1. Thus, after possibly renaming y and replacing ε by 2ε, we will assume $1 \in V_\varepsilon(y)$. For the remainder of the argument we set $V(y) = V_\varepsilon(y)$.

Since $1 \in V(y)$, there is an element $h \in \pi^{-1}(N_\varepsilon(y))$ satisfying $\mathsf{d}_G(1,h) \leqslant M$. Notice that the definition of the map π implies that $h \in \pi^{-1}(N_\varepsilon(y))$ is equivalent to $hY_0 \in N_\varepsilon(y)$, and thus

$$\mathsf{d}_T(hY_0, y) \leqslant \varepsilon. \tag{6}$$

Our goal is to use the acylindricity of the action of G on T to bound the diameter of $V(y)$. Since acylindricity will give a bound on the intersection of the quasi-stabilizers of Y_0 and y, we begin by establishing relationships between $\mathrm{Stab}_\varepsilon(y)$, $\mathrm{Stab}_\varepsilon(hY_0)$, $h\,\mathrm{Stab}_\varepsilon(Y_0)$, and their intersections. In particular, we will show

$$\mathrm{Stab}_{4\varepsilon}(hY_0) \subseteq \mathrm{Stab}_{6\varepsilon}(y), \tag{7}$$

$$\pi^{-1}(N_\varepsilon(y)) \subseteq h\,\mathrm{Stab}_{2\varepsilon}(Y_0), \tag{8}$$

and there exists a constant R' depending only on M such that

$$\mathrm{Stab}_\varepsilon(Y_0) \cap N_{2M}(\mathrm{Stab}_{2\varepsilon}(hY_0)) \subseteq N_{R'}(\mathrm{Stab}_{4\varepsilon}(Y_0) \cap \mathrm{Stab}_{4\varepsilon}(hY_0)). \tag{9}$$

We first establish (7). Let $g \in \mathrm{Stab}_{4\varepsilon}(hY_0)$, so that $\mathsf{d}_T(ghY_0, hY_0) \leqslant 4\varepsilon$. Combining this with (6) and the triangle inequality gives

$$\begin{aligned}
\mathsf{d}_T(gy, y) &\leqslant \mathsf{d}_T(gy, ghY_0) + \mathsf{d}_T(ghY_0, hY_0) + \mathsf{d}_T(hY_0, y) \\
&= \mathsf{d}_T(y, hY_0) + \mathsf{d}_T(ghY_0, hY_0) + \mathsf{d}_T(hY_0, y) \\
&\leqslant \varepsilon + 4\varepsilon + \varepsilon \\
&= 6\varepsilon.
\end{aligned}$$

Thus $g \in \mathrm{Stab}_{6\varepsilon}(y)$, proving (7).

We next establish (8). Let $g \in \pi^{-1}(N_\varepsilon(y))$, so that $gY_0 \in N_\varepsilon(y)$ and thus $\mathsf{d}_T(gY_0, y) \leqslant \varepsilon$. As before, combining this with (6) and the triangle inequality gives

$$\begin{aligned}
\mathsf{d}_T(Y_0, h^{-1}gY_0) &= \mathsf{d}_T(hY_0, gY_0) \\
&\leqslant \mathsf{d}_T(hY_0, y) + \mathsf{d}_T(y, gY_0) \\
&\leqslant 2\varepsilon.
\end{aligned}$$

Thus $h^{-1}g \in \mathrm{Stab}_{2\varepsilon}(Y_0)$, and so $g \in h\,\mathrm{Stab}_{2\varepsilon}(Y_0)$, proving (8).

We finally establish (9). Note first that

$$N_M(h\,\mathrm{Stab}_{2\varepsilon}(Y_0)) \subseteq N_{2M}(h\,\mathrm{Stab}_{2\varepsilon}(Y_0)h^{-1}) = N_{2M}(\mathrm{Stab}_{2\varepsilon}(hY_0)). \tag{10}$$

By Lemma 4.5 there is a constant R' depending only on M such that

$$\mathrm{Stab}_\varepsilon(Y_0) \cap N_{2M}(\mathrm{Stab}_{2\varepsilon}(hY_0)) \subseteq N_{R'}(\mathrm{Stab}_{4\varepsilon}(Y_0) \cap \mathrm{Stab}_{4\varepsilon}(hY_0)),$$

and so (9) follows.

We are now ready to bound the diameter of $V(y)$. Let $g \in V(y)$. By the definition of $V(y)$, $g \in \mathrm{Stab}_\varepsilon(Y_0)$ and $g \in N_M(\pi^{-1}(N_\varepsilon(y)))$. It then follows from (8) that $g \in \mathrm{Stab}_\varepsilon(Y_0) \cap N_M(h\,\mathrm{Stab}_{2\varepsilon}(Y_0))$. Thus it suffices to find a bound on the diameter of $\mathrm{Stab}_\varepsilon(Y_0) \cap N_M(h\,\mathrm{Stab}_{2\varepsilon}(Y_0))$.

We have the following inclusions:

$$\begin{aligned}
\mathrm{Stab}_\varepsilon(Y_0) \cap N_M(h\,\mathrm{Stab}_{2\varepsilon}(Y_0) &\subseteq \mathrm{Stab}_\varepsilon(Y_0) \cap N_{2M}(\mathrm{Stab}_{2\varepsilon}(hY_0) \\
&\subseteq N_{R'}(\mathrm{Stab}_{4\varepsilon}(Y_0) \cap \mathrm{Stab}_{4\varepsilon}(hY_0)) \\
&\subseteq N_{R'}(\mathrm{Stab}_{6\varepsilon}(Y_0) \cap \mathrm{Stab}_{6\varepsilon}(y)),
\end{aligned}$$

where the first inclusion follows from (10), the second from (9), and the third from (7).

Since $\mathsf{d}_T(Y_0, y) \geqslant R(6\varepsilon)$, the acylindricity of the action of G on T implies that

$$|\,\mathrm{Stab}_{6\varepsilon}(Y_0) \cap \mathrm{Stab}_{6\varepsilon}(y)| \leqslant N(6\varepsilon).$$

Thus the diameter of $N_{R'}(\mathrm{Stab}_{6\varepsilon}(Y_0) \cap \mathrm{Stab}_{6\varepsilon}(y))$ is bounded, which in turns implies that the diameter of $\mathrm{Stab}_\varepsilon(Y_0) \cap N_M(h\,\mathrm{Stab}_{2\varepsilon}(Y_0))$ is bounded. This bound depends on our choice of y, as well as on M, ε, and our choice of h. For a given h, there may be infinitely many $y \in T$ such that $h \in \pi^{-1}(N_\varepsilon(y))$ and $\mathsf{d}_G(1,h) \leqslant M$. For each h, we take the infimum over all such y of the diameter of $N_{R'}(\mathrm{Stab}_{6\varepsilon}(Y_0) \cap \mathrm{Stab}_{6\varepsilon}(y))$, which implies that the diameter of $\mathrm{Stab}_\varepsilon(Y_0) \cap N_M(h\,\mathrm{Stab}_{2\varepsilon}(Y_0))$ is bounded by some constant $D(h, M, \varepsilon)$ which depends only on M, ε, and our choice of the element h.

There are only finitely many choices for h in a ball of radius M about 1 in G with the fixed finite generating set, and thus we let

$$f(M) = \max\{D(h, M, \varepsilon) \mid \mathsf{d}_G(1,h) \leqslant M\},$$

completing the proof. □

4.2. **Proof of Theorem 4.1.** We will use d_G to denote the word metric on G with respect to a fixed finite generating set; we write d_T to denote distance in the quasi-tree T.

Let A_a and A_b be quasi-axes of a and b, respectively, in T. Let $p \in A_a$ be a closest point projection of Y_0 to A_a, let γ_a be a standard path in T from Y_0 to p, and let P_a be a standard path connecting p to $a^m p$ for some m, as determined below. Also let P_b be a standard path connecting gp to $ga^m p = b^m gp$. Consider the vertices $Y_0, aY_0, a^2Y_0, \ldots, a^mY_0$ in T, and connect consecutive points $a^{i-1}Y_0$ and a^iY_0 by a standard path $\gamma_{a,i}$ for $i = 1, \ldots, m$. (See Figure 3.)

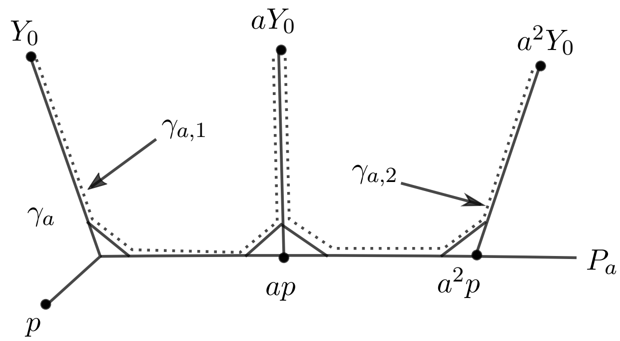

FIGURE 3. The configuration of standard paths in the quasi-tree T. The dotted paths are $\gamma_{a,i}$.

Let q be a closest point projection of Y_0 to A_b, let γ_b be a standard path in T from Y_0 to q, and let P_b be a standard path connecting q to $b^m q$. Since $A_b = gA_a$, after replacing g by $b^r g$ for some r, we may assume that $gp = q$ and that $gP_a = P_b$. (See Figure 4.)

Let α be a geodesic in T from Y_0 to q. Since γ_b is a standard path, $\gamma_b \subset N_2(\alpha)$, by Lemma 4.4; similarly $P_b \subset N_2(A_b)$. Let x be a vertex on $P_b \cap \gamma_b$. Then there are points $z \in \alpha$ and $a' \in A_b$ such that $\mathsf{d}_T(x,z) \leqslant 2$ and $\mathsf{d}_T(x,z') \leqslant 2$. Thus $\mathsf{d}_T(z, A_b) \leqslant 4$. Since p is a closest point projection of Y onto A_b and z lies on a geodesic from Y to A_b, it follows that p is also a closest point projection of z onto A_b. Thus $\mathsf{d}_T(z,p) \leqslant 4$, and $\mathsf{d}_T(p,x) \leqslant \mathsf{d}_T(p,z) + \mathsf{d}_T(z,x) \leqslant 4 + 2 = 6$. Let $p' \in \gamma_a \cap P_a$ be the vertex farthest from p and let $q' \in \gamma_b \cap P_b$ be the vertex farthest from q. After applying a similar argument to vertices on $\gamma_a \cap P_a$, we have $\mathsf{d}_T(p,p') \leqslant 6$ and $\mathsf{d}_T(q,q') \leqslant 6$.

Consider the quadrilateral of standard paths whose sides are γ_a, $\gamma_{a,1}$, $a\gamma_a$, and the subpath of P_a from p to ap. Then by Lemma 4.2, there is a vertex $v \in \gamma_{a,1}$ such that $\mathsf{d}_T(v,p') \leqslant 3$ and a vertex $w \in \gamma_{b,1}$ such that $\mathsf{d}_T(w,q') \leqslant 3$. Since standard paths are quasi-geodesics (Lemma 4.3), there is a uniform constant D and vertices v', w' on the geodesics connecting Y_0 to aY_0

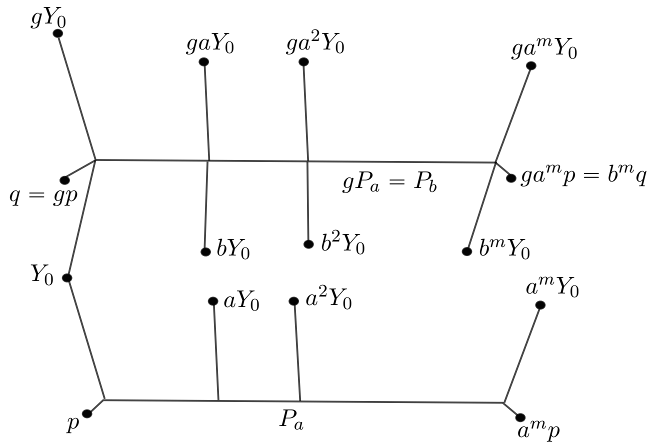

FIGURE 4. A schematic of the relevant portion of the quasi-tree T.

and Y_0 to bY_0, respectively, such that $\mathsf{d}_T(v', v) \leqslant D$ and $\mathsf{d}_T(w', w) \leqslant D$. It follows that $\mathsf{d}_T(v', p) \leqslant D + 9$ and $\mathsf{d}_T(w', q) \leqslant D + 9$. (See Figure 5.)

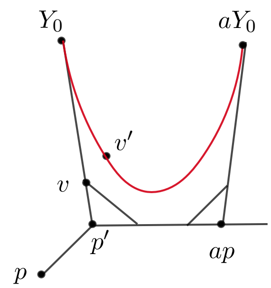

FIGURE 5. The red path is a quasi-geodesic in T from Y_0 to aY_0.

Let $L = \max_{s \in S} \mathsf{d}_T(Y_0, sY_0)$. Then there are vertices y_a, y_b on $[1, a]$ and $[1, b]$, respectively, in G such that $\mathsf{d}_T(y_aY_0, v') \leqslant L$ and $d_T(y_bY_0, w') \leqslant L$. Thus $\mathsf{d}_T(gy_aY_0, q) = \mathsf{d}_T(y_aY_0, p) \leqslant D + L + 9$ and $\mathsf{d}_T(y_bY_0, q) \leqslant D + L + 9$.

By a similar argument (letting a^mp and b^mq play the roles of p and q, respectively) there are vertices y'_a, y'_b on $[a^{m-1}, a^m]$ and $[b^{m-1}, b^m]$, respectively, such that $\mathsf{d}_T(gy'_aY_0, b^mq) \leqslant D + L + 9$ and $\mathsf{d}_T(y'_bY_0, b^mq) \leqslant D + L + 9$.

Therefore, $gy_a, y_b \in \pi^{-1}(N_{D+L+9}(q))$ and are each at distance at most $m(|a| + |b|)$ from $gy'_a, y'_b \in \pi^{-1}(N_{D+L+9}(b^mq))$, respectively. By choosing $m \geqslant R(4(D + L + 9))$, Lemma 4.6 implies that $\mathsf{d}_G(y_b, gy_a) \leqslant f(m|a| + m|b|)$ (see Figure 6).

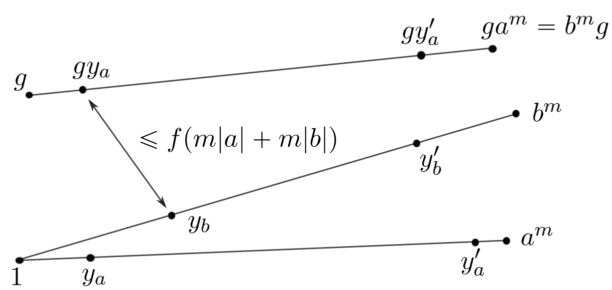

FIGURE 6. Estimating distances in G.

Finally, to complete the proof, we note that:

$$\begin{aligned} d_G(1, g) &\leqslant d_G(1, y_b) + d_G(y_b, gy_a) + d_G(gy_a, g) \\ &\leqslant |b| + f(m|a| + m|b|) + |a|. \end{aligned}$$ □

References

[ABD] Carolyn Abbott, Jason Behrstock, and Matthew G Durham. Largest acylindrical actions and stability in hierarchically hyperbolic groups. *arXiv math.GT/1705.06219*.

[Bal17] Sahana H. Balasubramanya. Acylindrical group actions on quasi-trees. *Algebr. Geom. Topol.*, 17(4):2145–2176, 2017.

[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.

[BBFS17] Mladen Bestvina, Ken Bromberg, Koji Fujiwara, and Alessandro Sisto. Acylindrical actions on projection complexes. *arXiv 1711.08722*, 2017.

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

[BHS15] Jason Behrstock, Mark F Hagen, and Alessandro Sisto. Hierarchically hyperbolic spaces II: combination theorems and the distance formula. *arXiv:1509.00632v2*, pages 1–53, 2015.

[BHS17a] J. Behrstock, M.F. Hagen, and A. Sisto. Asymptotic dimension and small-cancellation for hierarchically hyperbolic spaces and groups. *Proc. London Math. Soc.*, 114(5):890–926, 2017.

[BHS17b] Jason Behrstock, Mark F Hagen, and Alessandro Sisto. Hierarchically hyperbolic spaces I: curve complexes for cubical groups. *Geometry & Topology*, 21:1731–1804, 2017.

[BHS17c] Jason Behrstock, Mark F Hagen, and Alessandro Sisto. Quasiflats in hierarchically hyperbolic spaces. *arXiv:1704.04271*, 2017.

[BM97] M. Burger and S. Mozes. Finitely presented simple groups and products of trees. *C. R. Acad. Sci. Paris Sér. I Math.*, 324(7):747–752, 1997.

[BM00] Marc Burger and Shahar Mozes. Lattices in product of trees. *Inst. Hautes Études Sci. Publ. Math.*, (92):151–194 (2001), 2000.

[Bow08] Brian H. Bowditch. Tight geodesics in the curve complex. *Inventiones mathematicae*, 171(2):281–300, 2008.

[Cap17] Pierre-Emmanuel Caprace. Finite and infinite quotients of discrete and indiscrete groups. Preprint, arXiv:1709.05949, 2017.

[CGW09] John Crisp, Eddy Godelle, and Bert Wiest. The conjugacy problem in subgroups of right-angled Artin groups. *J. Topol.*, 2(3):442–460, 2009.

[DHS17] Matthew Durham, Mark Hagen, and Alessandro Sisto. Boundaries and automorphisms of hierarchically hyperbolic spaces. *Geom. Topol.*, 21(6):3659–3758, 2017.

[DHS18] Matthew Durham, Mark Hagen, and Alessandro Sisto. Corrigendum to boundaries and automorphisms of hierarchically hyperbolic spaces, 2018.

[HS16] M.F. Hagen and T. Susse. On hierarchical hyperbolicity of cubical groups. arXiv:1609.01313, 2016.

[HW08] Frédéric Haglund and Daniel T. Wise. Special cube complexes. *Geom. Funct. Anal.*, 17(5):1 551–1620, 2008.

[JW09] David Janzen and Daniel T. Wise. A smallest irreducible lattice in the product of trees. *Algebr. Geom. Topol.*, 9(4):2191–2201, 2009.

[Lys89] I. G. Lysënok. Some algorithmic properties of hyperbolic groups. *Izv. Akad. Nauk SSSR Ser. Mat.*, 53(4):814–832, 912, 1989.

[MM00] Howard A Masur and Yair N Minsky. Geometry of the complex of curves II: Hierarchical structure. *Geometric and Functional Analysis*, 10(4):902–974, 2000.

[MSW11] L. Mosher, M. Sageev, and K. Whyte. Quasi-actions on trees II: Finite depth Bass-Serre trees. *Mem. AMS*, 214(1008), 2011.

[Rat07] Diego Rattaggi. A finitely presented torsion-free simple group. *J. Group Theory*, 10(3):363–371, 2007.

[Sal14] Andrew Sale. Bounded conjugators for real hyperbolic and unipotent elements in semisimple Lie groups. *J. Lie Theory*, 24(1):259–305, 2014.

[Sal16] Andrew Sale. Geometry of the conjugacy problem in lamplighter groups. In *Algebra and computer science*, volume 677 of *Contemp. Math.*, pages 171–183. Amer. Math. Soc., Providence, RI, 2016.

[Ser89] Herman Servatius. Automorphisms of graph groups. *J. Algebra*, 126(1):34–60, 1989.

[Spr18] D. Spriano. Hyperbolic HHS II: Graphs of hierarchically hyperbolic groups. arXiv:1801.01850, 2018.

[Tao13] Jing Tao. Linearly bounded conjugator property for mapping class groups. *Geom. Funct. Anal.*, 23(1):415–466, 2013.

[Wis07] Daniel T Wise. Complete square complexes. *Commentarii Mathematici Helvetici*, 82(4):683–724, 2007.

University of California, Berkeley
E-mail address: c_abbott@math.berkeley.edu

Lehman College and The Graduate Center, CUNY
E-mail address: jason.behrstock@lehman.cuny.edu