Groups Geom. Dyn. 17 (2023), 805–838
DOI 10.4171/GGD/722

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 infinite order 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. In this setting, we establish a linear bound on the length of the shortest conjugator for any pair of conjugate Morse elements. For a subclass of these groups, including, in particular, all virtually compact special groups, we prove a sharper result by obtaining a linear bound on the length of the shortest conjugator between a suitable power of any pair of conjugate infinite order elements.

The conjugacy length function is the minimal function which bounds the length of a shortest conjugator between any two conjugate elements of a given group, in terms of the sum of the word lengths of the elements. When a set of elements in a group has a linear conjugacy length function, we say that set has the linear conjugator property. For any subset of a group satisfying the linear conjugator property, and given two elements of that subset, there is an exponential-time algorithm which determines whether or not the given elements are conjugate. One of Dehn's classic decision problems is the conjugacy problem, which asks if there is an algorithm to decide conjugacy given any pair of elements in a given group. Even in groups where the conjugacy problem is unsolvable for arbitrary pairs of elements, establishing the linear conjugator property for a particular subset allows one to solve the conjugacy problem for that subset.

An early established result about hyperbolic groups is they have the linear conjugacy property [24], thereby providing a quantitative certification of how complicated a conjugator needs to be. Exploiting the parallels between pseudo-Anosovs in the mapping class group and loxodromic elements in a hyperbolic group, Masur and Minsky proved the analogous result that the set of pseudo-Anosov elements satisfies the linear conjugator property [25]. These results beg the question of whether shortest conjugators of "hyperbolic-like" elements should be linear in the length of the elements being conjugated (see Conjecture B for a precise formulation).

In the presence of non-positive curvature, the linear conjugator property is surprisingly common, as we show in this paper, extending an already interesting class of known

2020 Mathematics Subject Classification. Primary 20F65; Secondary 20F10, 20E45, 20F67.
Keywords. Conjugators, conjugacy problem, hierarchically hyperbolic group, hyperbolic space.

examples. Previously established cases of the linear conjugator property include the following: mapping class groups (established for pseudo-Anosovs in [25], generalized to all elements in [34]; see also [2] for a later, unified proof); hyperbolic elements in semisimple Lie groups [30]; arbitrary elements in lamplighter groups [31]; non-peripheral elements in a relatively hyperbolic group [12]; Morse elements in groups acting on CAT(0) spaces [2]; and Morse elements in a prime 3-manifold [2]. 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 [32] (and we give a new proof below).

In light of this, we will work in the general setting of hierarchically hyperbolic groups, introduced by Behrstock, Hagen and Sisto [5]. This class of groups is quite large, encompassing many groups of interest, including: mapping class groups [3]; right-angled Artin groups, and more generally fundamental groups of compact special cube complexes [5] and other CAT(0) cube complexes [21]; 3-manifold groups with no Nil or Sol components [3]; and lattices in products of trees, i.e., as constructed by Burger and Mozes, Wise, and others, see [5, 13–15, 23, 26, 35]. There are a number of other examples, as well, for instance, groups obtained from combination theorems, including taking graphs of hierarchically hyperbolic groups and graph products of hierarchically hyperbolic groups [3, 7, 33], or by taking certain quotients of a hierarchically hyperbolic group [4].

The first theorem is new for most hierarchically hyperbolic groups; it also provides a unified proof for the previously known cases. An element in a finitely generated group is called *Morse* if its orbit in the group is a quasigeodesic with the property that any (λ, c)-quasigeodesic beginning and ending on this orbit is completely contained within a uniformly bounded neighborhood of this orbit. We note that Morse elements in a group are ones whose geometry in the Cayley graph is similar to that of the axis of a loxodromic isometry of a hyperbolic space (via the Morse lemma); in a hierarchically hyperbolic group, the Morse elements can be characterized in several equivalent ways, see [1, Theorem B].

Theorem A. *Let $(G, \mathfrak{S})$ be a hierarchically hyperbolic group. There exist constants K, C such that if $a, b \in G$ are Morse elements which are conjugate in G, then there exists $g \in G$ with $ga = bg$ and*

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

One special case of the above theorem is a new proof that conjugate pseudo-Anosov elements in the mapping class group have a linear bound on the length of their shortest conjugator; this case was the main theorem of [25].

A natural conjecture arising from Theorem A is the following generalization:

Conjecture B. *In a finitely generated group, the set of Morse elements satisfy the linear conjugator property.*

Understanding exactly how the linear conjugator property is related to hyperbolic properties in a group remains a rich question, and with Theorem A, hierarchically hyperbolic groups provide a good place to study this. For instance, we conjecture that there exist hierarchically hyperbolic groups where the conjugacy length function is exponential. Accordingly, we do not believe the linear conjugator property holds for all elements in all hierarchically hyperbolic groups, but it does in a number of important examples, which leads us to ask:

Question C. *Under what conditions does a hierarchically hyperbolic group satisfy the linear conjugator property for all elements?*

In Section 3, we introduce a family of hierarchically hyperbolic groups in which the notion of orthogonality carries with it not just geometric implications, but also a useful algebraic structure. The way in which the algebraic structure is related to orthogonality in these groups generalizes the relationship between commutativity and orthogonality in mapping class groups and compact special groups. This family is defined through a series of conditions called the $\mathbf{F}_U$ stabilizers, orthogonal decomposition, and commutative properties (see Section 3 for the precise definitions).

After showing in Proposition 3.10 that many groups satisfy the properties we introduce and that being in this family is preserved by various combination theorems, we then study conjugators in these groups. The following generalizes Theorem A by removing the hypothesis that the elements are Morse:

Theorem D. *Let $(G, \mathfrak{S})$ be a hierarchically hyperbolic group satisfying the $\mathbf{F}_U$ stabilizers, orthogonal decomposition, and commutative properties. 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*

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

In particular, compact special groups (i.e., fundamental groups of compact cube complexes which are special in the sense of Haglund–Wise [22]) satisfy the $\mathbf{F}_U$ stabilizers, orthogonal decomposition, and commutative properties. Therefore, Theorem D holds for all virtually compact special groups. We note that [17] established a linear time solution to the conjugacy problem for fundamental groups of compact special cube complexes. This result does not a priori establish the linear conjugator property of Theorem D, although we believe that their approach could be used to do so.

We believe that the linear conjugator property will in general fail for cubulated groups without the hypothesis that the cube complex is special. Our proof relies heavily on the close relationship between orthogonality and commutation, something which can fail for CAT(0) cubical groups which are not special, even though they may be hierarchically hyperbolic groups. The Burger–Mozes groups [13, 14], for instance, are plausibly a counterexample; see [5, Section 8.2.2], or Wise's construction [35].

1. Background

1.1. Hyperbolic geometry

We begin by gathering several facts about δ-hyperbolic metric spaces and refer the reader to [11] for further details.

A map of metric spaces $f\colon (X, \mathsf{d}_X) \to (Y, \mathsf{d}_Y)$ is a (λ, c)-*quasi-isometric embedding* if for all $x, y \in X$

$$\frac{1}{\lambda}\mathsf{d}_X(x, y) - c \leq \mathsf{d}_Y(f(x), f(y)) \leq \lambda \mathsf{d}_X(x, y) + c.$$

A (λ, c)-*quasigeodesic* is a (λ, c)-*quasi-isometric embedding* of an interval $I \subseteq \mathbb{R}$ into X, and a *geodesic* is an isometric embedding of I into X. In both cases, we allow f to be a *coarse map*, that is, a map which sends points in I to uniformly bounded diameter sets in X. A (coarse) map $f\colon [0, T] \to X$ is an *unparametrized* (λ, c)-*quasigeodesic* if there exists a strictly increasing function $g\colon [0, T'] \to [0, T]$ such that the following hold:

- $g(0) = f(0)$,
- $g(T') = f(T)$,
- $f \circ g\colon [0, T'] \to X$ is a (λ, c)-quasigeodesic, and
- for each $j \in [0, T'] \cap \mathbb{N}$, we have the diameter of $f(g(j)) \cup f(g(j+1))$ is at most c.

If $Y \subseteq X$ is a subspace, then for any constant $K \geq 0$, we denote the closed K-neighborhood of Y in X by

$$\mathcal{N}_K(Y) = \{x \in X \mid \mathsf{d}_X(x, Y) \leq K\}.$$

We may write $\mathcal{N}_K^X(Y)$ to emphasize that the neighborhood is being taken in X.

A subspace $Y \subseteq X$ is σ-*quasiconvex* if any geodesic in X with endpoints in Y is contained in $\mathcal{N}_\sigma(Y)$. The subspace Y is called *quasiconvex* if it is σ-quasiconvex for some σ.

If X is a geodesic metric space and $x, y \in X$, we let $[x, y]$ denote a geodesic from x to y. If we want to emphasize the metric space X, we write $[x, y]_X$.

Definition 1.1 (δ-hyperbolic space). Fix $\delta \geq 0$. A metric space X is δ-*hyperbolic* if given any $x, y, z \in X$ and any geodesics α, β, γ between them, we have $\alpha \cup \beta \subseteq \mathcal{N}_\delta(\gamma)$. If the particular choice of δ is not important, we simply say that X is *hyperbolic*.

Quasi-geodesics in a hyperbolic spaces satisfy two important properties: a local-to-global property and the Morse lemma. A path p is an L-*local* (λ, c)-*quasigeodesic* if every subpath p of length at most L is a (λ, c)-quasigeodesic.

Lemma 1.2 (Local-to-global property). *Let X be a δ-hyperbolic metric space and fix $\ell_0 \geq 0$. There exists $L = L(\ell_0, \delta)$ depending only on δ and ℓ_0 such that if $\ell \in [0, \ell_0]$ and $\gamma\colon I \to X$ is an L-local $(1, \ell)$-quasigeodesic, then γ is a global $(2, \ell)$-quasigeodesic.*

Lemma 1.3 (Morse lemma). *Let X be a δ-hyperbolic metric space, and fix $\lambda \geq 1$ and $c \geq 0$. There exists a constant σ depending only on δ, λ, and c such that if γ_1 and γ_2 are (λ, c)-quasigeodesics in X with the same endpoints, then $\gamma_1 \subseteq \mathcal{N}_\sigma(\gamma_2)$.*

We say σ is the *Morse constant* associated to (λ, c)-quasigeodesics in a δ-hyperbolic space.

Let G act by isometries on a δ-hyperbolic metric space X. Then $h \in G$ is

- *elliptic* if it has bounded orbits;
- *loxodromic* if the map $\mathbb{Z} \to X$ defined by $n \mapsto h^n x$ is a quasi-isometric embedding for some (equivalently, any) $x \in X$;
- *parabolic* otherwise.

Isometries of a hyperbolic space can also be characterized by their limit sets in the Gromov boundary ∂X of X. An element $h \in G$ is elliptic, parabolic, or loxodromic if the limit set of h has cardinality 0, 1, or 2, respectively. If the limit set of h has cardinality 2, we call these limit points $h^{\pm\infty}$.

Loxodromic isometries will play a particularly important role in this paper, and we discuss them in more depth. For the rest of the subsection, fix a group G acting by isometries on a δ-hyperbolic space X, and fix an element $h \in G$ that is loxodromic with respect to this action. The *translation length* of h is $[h]_X := \inf_{x \in X} \mathsf{d}_X(x, hx)$, or simply $[h]$ if the space X is clear. The *stable translation length* of h in X is

$$\tau_X(h) := \lim_{n \to \infty} \frac{\mathsf{d}_X(x_0, h^n x_0)}{n}$$

for some (equivalently, any) $x_0 \in X$. These two quantities are related by $\tau_X(h) \leq [h]_X \leq \tau_X(h) + 16\delta$.

The element h acts on X as translation along a quasigeodesic axis which connects the two limit points $h^{\pm\infty}$ of h in ∂X. Up to passing to powers, such an axis can be chosen to be a *uniform quality* quasigeodesic, that is, with quasigeodesic constants which depending only on δ and not on the choice of loxodromic isometry. As this will be important in this paper, we now describe the construction of such an axis. The following lemma summarizes results from [16, Section 3].

Lemma 1.4 (Construction of an ℓ-nerve). *Let G act on a δ-hyperbolic space X. Suppose $h \in G$ is loxodromic and $\tau_X(h) \geq L_S\delta - 16\delta$, where L_S depends only on δ (and is more explicitly described in* [16, Definition 2.8]*). Then for any $\ell \in [0, \delta]$, there exists a $(2, \ell)$-quasigeodesic γ_h^X in X which connects the limit points $h^{\pm\infty}$ of h in ∂X, called the ℓ-*nerve *of h. The ℓ-nerve of h is $(\ell + 8\delta)$-quasiconvex and preserved by h.*

We briefly recall the construction of the ℓ-nerve and refer the reader to [16, Definition 3.3 and subsequent remark] for further details. Fix $\ell \in [0, \delta]$. By the definition of the translation length $[h]$, there exists $x \in X$ such that $\mathsf{d}_X(x, hx) \leq [h] + \ell/2$. Thus we can find a $(1, \ell/2)$-quasigeodesic γ from x to hx. If γ has length T, then $[h] \leq T < [h] + \ell/2$.

Extend γ to a bi-infinite path γ_h^X using the action of $\langle h\rangle$; that is, γ_h^X is the concatenation of the segments $h^i\gamma$ for $i \in \mathbb{Z}$. In particular, for any h for which $\tau_X(h) \geq L_S\delta - 16\delta$, we have that γ is an $L_S\delta$-local $(1,\ell)$-quasigeodesic. By Lemma 1.2, γ_h^X is therefore a global $(2,\ell)$-quasigeodesic.

Definition 1.5 (Quasi-geodesic axis). Let G act on a δ-hyperbolic space X, and fix the constant L_S from Lemma 1.4 and a constant $\ell \in [0,\delta]$. Suppose $h \in G$ is a loxodromic isometry of X, and let $k \in \mathbb{N}$ be such that $\tau_X(h^k) \geq L_S\delta - 16\delta$. A *$(2,\ell)$-quasigeodesic axis* of h^k in X is an ℓ-nerve $\gamma_{h^k}^X$. If the quasigeodesic constants are not important, we simply call $\gamma_{h^k}^X$ an *axis* of h^k.

Suppose $h, g \in G$ and h is loxodromic with respect to the action of G on a hyperbolic metric space X with $\tau_X(h) \geq L_S\delta - 16\delta$ and a $(2,\ell)$-quasigeodesic axis γ_h^X. Then ghg^{-1} is also loxodromic with respect to the action of G on X with translation length $\tau_X(ghg^{-1}) = \tau_X(h)$, and it follows from the construction of the ℓ-nerve that $g\gamma_h^X$ is a $(2,\ell)$-quasigeodesic axis of ghg^{-1}.

1.2. Hierarchically hyperbolic spaces

We recall the definition of a hierarchically hyperbolic space as given in [3]. The definition is in the setting of a *quasigeodesic metric space*, that is, a metric space in which any two points can be connected by a uniform quality quasigeodesic.

Definition 1.6 (Hierarchically hyperbolic space). The quasigeodesic space $(\mathcal{X}, \mathsf{d}_{\mathcal{X}})$ is a *hierarchically hyperbolic space* (HHS) if there exist $\delta \geq 0$, an index set $\mathfrak{S}$, and a set $\{\mathcal{C}W : W \in \mathfrak{S}\}$ of δ-hyperbolic spaces $(\mathcal{C}W, \mathsf{d}_W)$, satisfying the following conditions:

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

(2) *Nesting*. $\mathfrak{S}$ is equipped with a partial order $\sqsubseteq$, and either $\mathfrak{S} = \emptyset$ 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 $\rho_W^V \subset \mathcal{C}W$ with $\mathsf{diam}_{\mathcal{C}W}(\rho_W^V) \leq \xi$. There is also a *projection* $\rho_V^W : \mathcal{C}W \to 2^{\mathcal{C}V}$. We call the elements of the index set $\mathfrak{S}$ *domains*.

(3) *Orthogonality*. $\mathfrak{S}$ has a symmetric and anti-reflexive relation called *orthogonality*: we write $V \perp W$ when V and 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 $\{V \in \mathfrak{S}_T \mid V \perp U\} \neq \emptyset$, there exists $W \in \mathfrak{S}_T - \{T\}$, so that whenever $V \perp U$ and $V \sqsubseteq T$, we have $V \sqsubseteq W$. The domain W is called the *container* associated to U in T. Finally, if $V \perp W$, then V and 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 and W are *transverse*, denoted $V \pitchfork W$. There exists $\kappa_0 \geq 0$ 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

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

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

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

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

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

(5) *Finite complexity*. There exists $n \geq 0$, 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 $\lambda \geq 1$ and $E \geq \max\{\xi, \kappa_0\}$ 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, $\mathsf{d}_W(\pi_W(x), \rho^{T_i}_W) \leq N$ 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 $\mathsf{diam}_{\mathcal{C}V}(\rho^W_V(\gamma)) \leq E$ 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

- $\mathsf{d}_{V_j}(\pi_{V_j}(x), p_j) \leq \alpha$ for all j,
- for each j and each $V \in \mathfrak{S}$ with $V_j \sqsubseteq V$, we have $\mathsf{d}_V(\pi_V(x), \rho^{V_j}_V) \leq \alpha$, and
- if $W \pitchfork V_j$ for some j, then $\mathsf{d}_W(\pi_W(x), \rho^{V_j}_W) \leq \alpha$.

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

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 $\mathsf{d}_U(\pi_U(x), p)$. When necessary for clarity, we may also write $\mathcal{C}(U)$ instead of $\mathcal{C}U$.

An important consequence of being a hierarchically hyperbolic space is the following distance formula, which relates distances in $\mathcal{X}$ to distances in the hyperbolic spaces $\mathcal{C}U$ for $U \in \mathfrak{S}$. Give $a, b \in \mathbb{R}$, the notation $\{\!\{a\}\!\}_b$ denotes the quantity which is a if $a \geq b$ and is 0 otherwise. Given C, D, we say $a \asymp_{C,D} b$ if $C^{-1}a - D \leq b \leq Ca + D$. We use $a \asymp_D b$ if $|a - b| \leq D$, and we use $a \preceq_{C,D} b$ if $a \leq Cb + D$.

Theorem 1.7 (Distance formula for HHS, [3]). *Let $(\mathcal{X}, \mathfrak{S})$ be a hierarchically hyperbolic space. Then there exists s_0 such that for all $s \geq s_0$, there exist C, D so that for all $x, y \in \mathcal{X}$,*

$$\mathsf{d}_{\mathcal{X}}(x, y) \asymp_{C,D} \sum_{U \in \mathfrak{S}} \{\!\{\mathsf{d}_U(x, y)\}\!\}_s.$$

The distance formula says that the distance between two points in $\mathcal{X}$ can be approximated by measuring the distances between their projections to the hyperbolic spaces, and, moreover, that we only need to consider hyperbolic spaces for which that projection is sufficiently large.

Definition 1.8 (Relevant domains). For any constant $R \geq s_0$ and any two points $x, y \in \mathcal{X}$, we say $U \in \mathfrak{S}$ is *relevant* (with respect to x, y, R) if $\mathsf{d}_U(x, y) \geq R$; if we want to emphasize the constant R, we say that U is *R-relevant* (with respect to x, y). We denote the set of R-relevant domains by $\mathbf{Rel}(x, y; R)$.

In other words, the set of R-relevant domains for a pair of points $x, y \in \mathcal{X}$ are the domains which appear in the distance formula for x and y with the threshold $s = R$.

Notation 1.9. Given a hierarchically hyperbolic space $(\mathcal{X}, \mathfrak{S})$, we let E denote a constant greater than any of the constants occurring in Definition 1.6 and greater than the constant s_0 from Theorem 1.7.

Definition 1.10 (Hierarchy path). Given a hierarchically hyperbolic space $(\mathcal{X}, \mathfrak{S})$ and a constant $\lambda \geq 1$, a (λ, λ)-*hierarchy path* $\gamma \subset \mathcal{X}$ is a (λ, λ)-quasigeodesic in $\mathcal{X}$ with the property that for each $U \in \mathfrak{S}$ the path $\pi_U(\gamma)$ is an unparametrized (λ, λ)-quasigeodesic in $\mathcal{C}U$.

By [3, Theorem 4.4], for any sufficiently large λ, any two points $x, y \in \mathcal{X}$ are connected by a (λ, λ)-hierarchy path. We fix such a constant $\lambda > E$, and let $\mu(x, y) \subseteq \mathcal{X}$ denote a (λ, λ)-hierarchy path from x to y.

Definition 1.11 (Hierarchically hyperbolic group). A finitely generated group G is a *hierarchically hyperbolic group* if some (hence any) Cayley graph of G is a hierarchically hyperbolic space, and the hierarchically hyperbolic structure is G-invariant. In particular, a hierarchically hyperbolic group is a finitely generated group G, equipped with a specific choice of finite generating set, such that there is a hierarchically hyperbolic space $(G, \mathfrak{S})$ satisfying the following properties:

- G acts cofinitely on $\mathfrak{S}$, preserving the relations $\sqsubseteq$, $\pitchfork$ and $\perp$.
- For each $U \in \mathfrak{S}$ and $g \in G$, there is an isometry $g\colon \mathcal{C}U \to \mathcal{C}(gU)$, and if $h \in G$, then the isometry $gh\colon \mathcal{C}U \to \mathcal{C}(ghU)$ is equal to the composition $\mathcal{C}U \xrightarrow{h} \mathcal{C}(hU) \xrightarrow{g} \mathcal{C}(ghU)$.
- For each $U \in \mathfrak{S}$ and $g, x \in G$, we have $g\pi_U(x) = \pi_{gU}(gx)$.
- For each $U, V \in \mathfrak{S}$ such that $U \pitchfork V$ or $U \sqsubsetneq V$ and each $g \in G$, we have $\rho^{gU}_{gV} = g\rho^U_V$.

Given a hierarchically hyperbolic group $(G, \mathfrak{S})$, we use d_G to denote the distance in the group G with respect to some (fixed) finite generating set.

1.3. Gate maps and standard product regions

In analogy with quasiconvex subspaces of hyperbolic spaces, there is a notion of a hierarchically quasiconvex subspace of a hierarchically hyperbolic space $\mathcal{X}$.

Definition 1.12 (Hierarchically quasiconvex). Let $(\mathcal{X}, \mathfrak{S})$ be a hierarchically hyperbolic space. A subspace $\mathcal{Y}$ of $\mathcal{X}$ is *k-hierarchically quasiconvex* for some $k\colon [0,\infty) \to [0,\infty)$ if the following hold:

(1) For all $U \in \mathfrak{S}$, the projection $\pi_U(\mathcal{Y})$ is a $k(0)$-quasiconvex subspace of $\mathcal{C}U$.

(2) For every $\kappa > 0$ and every point $x \in \mathcal{X}$ satisfying $\mathsf{d}_U(\pi_U(x), \pi_U(\mathcal{Y})) \le \kappa$ for all $U \in \mathfrak{S}$, we have $\mathsf{d}_{\mathcal{X}}(x, \mathcal{Y}) \le k(\kappa)$.

The first condition says that the subspace $\mathcal{Y}$ projects to a (uniformly) quasiconvex subspace in every hyperbolic space, while the second condition ensures that all points in $\mathcal{X}$ which project near $\mathcal{Y}$ in every hyperbolic spaces are near $\mathcal{Y}$ in $\mathcal{X}$.

As is the case for quasiconvex subspaces of hyperbolic spaces, if $\mathcal{Y}$ is a hierarchically quasiconvex subspace of a hierarchically hyperbolic space $\mathcal{X}$, then there is a well-defined "nearest point projection" from $\mathcal{X}$ to $\mathcal{Y}$, called a gate map.

Definition 1.13 (Gate maps). If $(\mathcal{X}, \mathfrak{S})$ is a hierarchically hyperbolic group and $\mathcal{Y}$ is a hierarchically quasiconvex subspace of $\mathcal{X}$, then the *gate map* is a coarsely-Lipschitz map $\mathfrak{g}_{\mathcal{Y}}\colon \mathcal{X} \to 2^{\mathcal{Y}}$, so that for each $x \in \mathcal{X}$, the image $\mathfrak{g}_{\mathcal{Y}}(x)$ is a subset of the points in $\mathcal{Y}$ with the property that for each $U \in \mathfrak{S}$ the set $\pi_U(\mathfrak{g}_{\mathcal{Y}}(x))$ uniformly coarsely coincides with the closest point projection in $\mathcal{C}U$ of $\pi_U(x)$ to $\pi_U(\mathcal{Y})$.

The following lemma shows that gate maps are uniformly coarsely equivariant.

Lemma 1.14 ([29, Lemma 4.16]). *Let $(G, \mathfrak{S})$ be a hierarchically hyperbolic group, and let $\mathcal{Y}$ be a k-hierarchically quasiconvex subspace of G. Then there exists a constant A depending on $(G, \mathfrak{S})$ and k such that for every $g, x \in G$, we have*

$$\mathsf{d}_G(g\mathfrak{g}_{\mathcal{Y}}(x), \mathfrak{g}_{g\mathcal{Y}}(gx)) \le A.$$

We now recall an important family of hierarchically quasiconvex subspaces in a hierarchically hyperbolic space called *standard product regions* introduced in [5, Section 13] and studied further in [3]. The definition we give can be found in [28, Definition 2.20] and is also discussed in [4, Section 1.2.1] .

Definition 1.15 (Standard product region). Let $(\mathcal{X}, \mathfrak{S})$ be a hierarchically hyperbolic space, and let $U \in \mathfrak{S}$. The *standard product region for U* is the set

$$\mathbf{P}_U = \{x \in \mathcal{X} \mid \mathsf{d}_V(x, \rho_V^U) \le E \text{ for all } V \in \mathfrak{S} \text{ with } V \pitchfork U \text{ or } V \sqsupsetneq U\}.$$

Note that if $S \in \mathfrak{S}$ is $\sqsubseteq$-maximal, then $\mathbf{P}_S = \mathcal{X}$.

In other words, given $U \in \mathfrak{S}$ and $V \in \mathfrak{S}$ satisfying $V \pitchfork U$ or $V \sqsupsetneq U$, the product region $\mathbf{P}_U$ is precisely the set of points which project near ρ^U_V in $\mathcal{C}V$. It thus follows from this definition that for such U, V, we have $\rho^U_V \asymp_E \pi_V(\mathbf{P}_U)$; that is, the projection ρ^U_V is coarsely equal to the projection of the product region $\mathbf{P}_U \subseteq \mathcal{X}$ into $\mathcal{C}V$.

Though it is not obvious from this definition, the product region $\mathbf{P}_U$ is quasi-isometric to a space with decomposes as a direct product of two factors, $\mathbf{F}_U$ and $\mathbf{E}_U$. As these factors will be important in this paper, we describe them in detail. See [3, Section 5.2] for additional details. We first define $\mathbf{F}_U$ and $\mathbf{E}_U$ as abstract spaces. In the paragraphs following the definitions, we explain that these spaces admit embeddings into $\mathcal{X}$. Unless otherwise noted, we will always think of these embeddings, rather than the abstract spaces themselves.

Definition 1.16 (Nested partial tuple ($\mathbf{F}_U$)). Let $\mathfrak{S}_U = \{V \in \mathfrak{S} \mid V \sqsubseteq U\}$. Fix $\kappa \geq E$ 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 consistency inequalities of Definition 1.6 (4)).

Definition 1.17 (Orthogonal partial tuple ($\mathbf{E}_U$)). Let $\mathfrak{S}^{\perp}_U = \{V \in \mathfrak{S} \mid V \perp U\} \cup \{W\}$, where W is a $\sqsubseteq$-minimal element such that $V \sqsubseteq W$ for all $V \perp U$. Fix $\kappa \geq E$, and let $\mathbf{E}_U$ be the set of κ-consistent tuples in $\prod_{V \in \mathfrak{S}^{\perp}_U - \{A\}} 2^{\mathcal{C}V}$.

Remark 1.18. The particular choice of constant κ will not be important in this paper. For simplicity, given a hierarchically hyperbolic group, we fix $\kappa = E$, and for each domain U we consider only spaces $\mathbf{F}_U$ and $\mathbf{E}_U$ defined using E-consistent tuples.

Given $\mathcal{X}$ and $U \in \mathfrak{S}$, there is a well-defined map $\phi_U \colon \mathbf{F}_U \times \mathbf{E}_U \to \mathcal{X}$. The precise definition of this map is not necessary for this paper; we refer the interested reader to [3, Construction 5.10]. The product region $\mathbf{P}_U$ defined in Definition 1.15 is coarsely equal to the image $\phi_U(\mathbf{F}_U \times \mathbf{E}_U)$ in $\mathcal{X}$. In this paper, we will only work with $\mathbf{P}_U$ and $\mathbf{F}_U$. For all results that we state for $\mathbf{F}_U$, analogous statements also hold for $\mathbf{E}_U$.

Fixing any $e \in \mathbf{E}_U$ restricts ϕ_U to a map $\phi^{\sqsubseteq} \colon \mathbf{F}_U \times \{e\} \to \mathcal{X}$. In general, this map $\phi^{\sqsubseteq}$ depends on the choice of $e \in \mathbf{E}_U$. When the basepoint is immaterial (or understood), we abuse notation and consider $\mathbf{F}_U$ to be a subspace of $\mathcal{X}$, that is, $\mathbf{F}_U = \operatorname{im} \phi^{\sqsubseteq}$.

It is proven in [3, Lemma 5.5] that standard product regions $\mathbf{P}_U$ and their factors $\mathbf{F}_U \times \{e\}$ for each $e \in \mathbf{E}_U$ (considered as subspaces of $\mathcal{X}$) are uniformly hierarchically quasiconvex. Therefore, there are well-defined gate maps $\mathfrak{g}_{\mathbf{P}_U} \colon \mathcal{X} \to \mathbf{P}_U$ and $\mathfrak{g}_{\mathbf{F}_U \times \{e\}} \colon \mathcal{X} \to \mathbf{F}_U \times \{e\}$ for each $U \in \mathfrak{S}$ and each $e \in \mathbf{E}_U$.

Remark 1.19. We note that the gate map $\mathfrak{g}_{\mathbf{F}_U \times \{e\}}$ depends on the choice of $e \in \mathbf{E}_U$. However, the image of the gate map in $\mathcal{C}V$ for any $V \sqsubseteq U$ is independent of this choice (see [4, Remark 1.16]). That is, if $e, e' \in \mathbf{E}_U$, then for any $x \in \mathcal{X}$, we have

$$\pi_V(\mathfrak{g}_{\mathbf{F}_U \times \{e\}}(x)) = \pi_V(\mathfrak{g}_{\mathbf{F}_U \times \{e'\}}(x)).$$

In statements where we only consider the image of the gate map in the hyperbolic spaces, we simplify notation and write $\mathfrak{g}_{\mathbf{F}_U}$.

The following lemma provides a formula for computing the distance between a point and a product region. It is an immediate consequence of [6, Corollary 1.28]; we give a sketch of the proof here for completeness.

Lemma 1.20 ([6]). *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 $Y \sqsupsetneq U\}$. Then for all $s \geq s_0$ and any $x \in \mathcal{X}$,*

$$\mathsf{d}_{\mathcal{X}}(x, \mathbf{P}_U) \asymp_{C,D} \sum_{Y \in \mathcal{Y}} \{\!\{\mathsf{d}_Y(x, \rho_Y^U)\}\!\}_s,$$

where s_0, C, and D are the constants from Theorem 1.7.

Sketch of proof. To each bounded set $\mathcal{A} \subset \mathcal{X}$, we associate a tuple $(\mathcal{A}_V)_{V \in \mathfrak{S}}$ whose components are the projections of $\mathcal{A}$ to $\mathcal{C}V$ for each $V \in \mathfrak{S}$, i.e., $\mathcal{A}_V = \pi_V(\mathcal{A})$. We will consider the case $\mathcal{A} = \mathfrak{g}_{\mathbf{P}_U}(x) \subset \mathbf{P}_U$. By [4, Remark 1.16], if $V \sqsubseteq U$ or $V \perp U$, we have $\pi_V(\mathfrak{g}_{\mathbf{P}_U}(x)) = \pi_V(x)$. Combining this with the definition of $\mathbf{P}_U$ (Definition 1.15), we have

$$(\mathfrak{g}_{\mathbf{P}_U}(x))_V = \begin{cases} \rho_V^U & \text{if } V \in \mathcal{Y}, \\ \pi_V(x) & \text{otherwise.} \end{cases}$$

There is a constant K_0 depending only on $(G, \mathfrak{S})$ such that

$$\mathsf{d}_{\mathcal{X}}(x, \mathbf{P}_U) \asymp_{K_o} \mathsf{d}_{\mathcal{X}}(x, \mathfrak{g}_{\mathbf{P}_U}(x))$$

by [6, Lemma 1.27]. From the above discussion, we see that the only components of the tuple $(x_V)_{V \in \mathfrak{S}}$ associated to x and the tuple $(\mathfrak{g}_{\mathbf{P}_U}(x))_{V \in \mathfrak{S}}$ associated to $\mathfrak{g}_{\mathbf{P}_U}(x)$ which differ in $\mathcal{C}V$ occur when $V \in \mathcal{Y}$. Thus the distance from x to $\mathfrak{g}_{\mathbf{P}_U}(x)$ in $\mathcal{X}$ can be approximated using only the domains $V \in \mathcal{Y}$. ∎

Lemma 1.20 gives the following geometric picture. Let $x, y \in \mathcal{X}$ and $U \in \mathfrak{S}$, and consider $x' = \mathfrak{g}_{\mathbf{P}_U}(x)$ and $y' = \mathfrak{g}_{\mathbf{P}_U}(y)$. Let V be a domain that is relevant for x and y. Then any distance in $\mathcal{C}V$ contributes either to the distance from x or y *to* the product region $\mathbf{P}_U$ or to the distance *within* $\mathbf{P}_U$, but not both (see Figure 1). In particular, if $V \pitchfork U$ or $V \sqsupsetneq U$, then V is relevant for either x, x' or y, y' but not for x', y'. Any other V is relevant for x', y' but not for x, x' or y, y'.

1.4. Axial elements in hierarchically hyperbolic groups

Let $(G, \mathfrak{S})$ be a hierarchically hyperbolic group, and fix the constant L_S from Lemma 1.4 and $\ell \in [0, \delta]$. (Note that the constant δ is part of the definition of $(G, \mathfrak{S})$; see Definition 1.6.) Following [18], for an element $h \in G$ we define

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

Lemma 1.21. *Let $(G, \mathfrak{S})$ be a hierarchically hyperbolic group. An element $h \in G$ is finite order if and only if* $\operatorname{Big}(h) = \varnothing$.

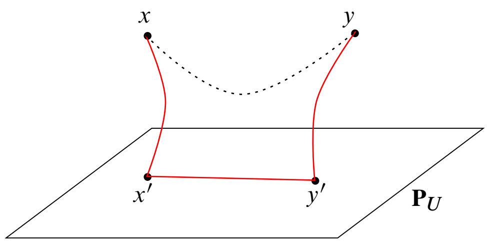

Figure 1. Geometric picture of Lemma 1.20. Domains which are relevant for x, y are relevant for either the horizontal red segment or (at least one of) the vertical red segments.

Proof. In [18, Proposition 6.4], it is proven that an element $h \in G$ is elliptic if and only if $\mathrm{Big}(h) = \varnothing$. 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. ∎

Definition 1.22 (Axial element). An element $h \in G$ with $\mathrm{Big}(h) \neq \varnothing$ is called *axial*.

Lemma 1.21 shows that every infinite order element of a hierarchically hyperbolic group is axial. By [18, Lemma 6.7], the elements of $\mathrm{Big}(h)$ are pairwise orthogonal. As the number of pairwise orthogonal domains in a hierarchically hyperbolic group is uniformly bounded by the constants in the definition of a hierarchically hyperbolic space [3, Lemma 2.1], it follows that $|\mathrm{Big}(h)|$ is uniformly bounded independently of the choice of h. As noted in [18], since $h\colon \mathcal{C}U \to \mathcal{C}(hU)$ is an isometry, we have $hU \in \mathrm{Big}(h)$ whenever $U \in \mathrm{Big}(h)$. Moreover, by [18, Lemma 6.3], there is a constant M depending only on the constants in the definition of a hierarchically hyperbolic space such that for all $h \in G$ and $U \in \mathrm{Big}(h)$, we have $h^M U = U$. In other words, by passing to a uniform power, we may assume that h fixes its big set elementwise. Moreover, by passing to this uniform power, we may assume that h is a loxodromic isometry of $\mathcal{C}U$ for any $U \in \mathrm{Big}(h)$ by [19, Theorem 3.1]. We let $\tau_U(h)$ denote the stable translation length of h in this action and let γ_h^U be a $(2,\ell)$-quasigeodesic axis of h in $\mathcal{C}U$ (see Definition 1.5).

Remark 1.23 (Acylindrical actions). The action of a group G on a metric space X is *acylindrical* if for all $\varepsilon \geq 0$, there exist constants $R(\varepsilon), N(\varepsilon) \geq 0$ such that for all $x, y \in X$ satisfying $\mathsf{d}_X(x,y) \geq R(\varepsilon)$, there are at most $N(\varepsilon)$ elements $g \in G$ for which

$$\mathsf{d}_X(x, gx) \leq \varepsilon \quad \text{and} \quad \mathsf{d}_X(y, gy) \leq \varepsilon.$$

By [5, Theorem K], G acts acylindrically on $\mathcal{C}S$, where S is the $\sqsubseteq$-maximal element of $\mathfrak{S}$. An immediate consequence of this is a lower bound on the translation length $\tau_S(h)$ that depends only on the hierarchy constants [10, Lemma 2.2].

Let $U \in \mathfrak{S}$, and let H be a subgroup of G which fixes U, so that H acts on $\mathcal{C}U$. If $U \neq S$, it is not necessarily the case that H acts acylindrically on $\mathcal{C}U$, and it remains an open question whether there is a uniform lower bound on $\tau_U(h)$ in general. We deal with this issue in the present paper by assuming such a uniform lower bound as a hypothesis.

Hierarchical acylindricity is a standard assumption requiring that the action of H on $\mathcal{C}U$ is acylindrical for all such U: this would also ensure a uniform lower bound on translation length.

The following lemma is a straightforward consequence of the hyperbolicity of the spaces $\mathcal{C}U$.

Lemma 1.24. *Let $(\mathcal{X}, \mathfrak{S})$ be a hierarchically hyperbolic space, and let G be a group acting geometrically on $\mathcal{X}$. Fix a basepoint $x_0 \in \mathcal{X}$, the constant L_S from Lemma 1.4, and $\ell \in [0, \delta]$. Then there exist constants $K_0, L \geq 0$ such that the following holds. Let $h \in G$ be an axial element so that $hU = U$ for each $U \in \operatorname{Big}(U)$ and $\tau_U(h) \geq L_S\delta$. For any $k \geq K_0$ let $x', y' \in \mathcal{C}U$ be the closest points on γ_h^U to $\pi_U(x_0)$ and $\pi_U(h^k x_0)$, respectively.*

There exists a point ξ on the subpath of γ_h^U from x' to y' such that $\mathsf{d}_U(\xi, x') \leq L$ and $\mathsf{d}_U(\xi, \pi_U(\mu(x, y))) \leq L$.

Proof. Recall that the image of any (λ, λ)-hierarchy path in $\mathcal{C}U$ is a (unparametrized) (λ, λ)-quasigeodesic. Since the axis of h in $\mathcal{C}U$ is a $(2, \ell)$-quasigeodesic, the concatenation (in the appropriate order) of $\pi_U(\mu(x_0, h^k x_0))$, $[\pi_U(x_0), x']_{\mathcal{C}U}$, $[\pi_U(h^k x_0), y']_{\mathcal{C}U}$, and a subpath of γ_h^U forms a $(2, \ell)$-quasigeodesic quadrilateral Q in $\mathcal{C}U$. Let M be the Morse constant associated to $(2, \ell)$-quasigeodesics in a δ-hyperbolic space. Fix K_0 so that $L_S\delta K_0 > 4\delta + 4M + 1$, and let $k \geq K_0$. Note that K_0 is independent of the choice of axial element h.

The quadrilateral Q is $(2M + 2\delta)$-thin, that is, given any point z on a side of Q, there is a point on one of the other three sides of Q at distance at most $2\delta + 2M$ from z. Let v and w be points on the subpath of γ_h^U between x' and y' so that $\mathsf{d}_S(x', v) = \lceil 4\delta + 4M + 1 \rceil$ and $\mathsf{d}_S(y', w) = \lceil 4\delta + 4M + 1 \rceil$. We claim that the subpath β of γ_h^U from v to w is contained in the $(2\delta + 2M)$-neighborhood of $\pi_U(\mu(x, y))$. Let z be a point on the subpath of γ_h^U from x' to y'. Then there is a point z' on one of the other three sides satisfying $\mathsf{d}_U(z, z') \leq 2M + 2\delta$. Suppose z' lies on the geodesic $[\pi_U(x), x']$. As x' is the nearest point on γ_h^U to x (hence also to z'), we must have

$$\mathsf{d}_U(z', x') \leq \mathsf{d}_U(z', z) \leq 2M + 2\delta.$$

The same holds if z' lies on the geodesic $[\pi_U(y), y']$. Thus if z lies on β, then z' must lie on $\pi_U(\mu(x, y))$, as desired.

Since the map π_U is G-equivariant and, in particular, $h^k \pi_U(x_0) = \pi_U(h^k x_0)$, we also have $y' = h^k x'$. Thus

$$\mathsf{d}_U(x', y') \geq k\tau_U(h) \geq K_0 L_S \delta \geq 4\delta + 4M + 1.$$

It follows that β is non-empty. We let ξ be the point on β closest to x', so that $\mathsf{d}_U(\xi, x') = \lceil 4\delta + 2M + 1 \rceil$. Taking $L = \lceil 4\delta + 2M + 1 \rceil$ completes the proof. ∎

2. Proof of Theorem A

Let $(G, \mathfrak{S})$ be a hierarchically hyperbolic group. The authors and Durham show in [1, Corollary 3.8] that by possibly changing the hierarchy structure on G, we may assume that $(G, \mathfrak{S})$ has *unbounded products*. In this paper, we do not directly use the definition of unbounded products, rather we only need the following consequence about Morse elements in the structure $(G, \mathfrak{S})$, which follows from [1, Theorem 4.4 and Corollary 5.5]: if $h \in G$ is an infinite order Morse element, then h is axial and $\mathrm{Big}(h) = \{S\}$, where S is the $\sqsubseteq$-maximal element of $\mathfrak{S}$.

We begin by fixing the constants that will be used throughout the proof. Definition 1.6 provides a constant δ such that $\mathcal{C}U$ is δ-hyperbolic for all $U \in \mathfrak{S}$. Let L_S be the constant from Lemma 1.4, and fix $\ell \in [0, \delta]$. Let E be as in Notation 1.9; in particular, E is larger than any of the hierarchy constants for G. Let T be the lower bound on translation length in the acylindrical action on $\mathcal{C}S$ noted in Remark 1.23. Fix $\lambda \geq \max\{2, \ell\}$ so that any two points $x, y \in G$ are connected by a (λ, λ)-hierarchy path. Let K_0, L be the constants from Lemma 1.24, and fix a constant $R > 2E$.

Finally, set

$$K = \max\left\{2\delta, R, K_0, \left\lceil \frac{4L + 3E}{T} \right\rceil + 2, 2L\right\}. \tag{2.1}$$

This constant K is uniform, in the sense that it depends only on the hierarchy constants for $(G, \mathfrak{S})$.

Let $a, b \in G$ be two infinite order Morse elements and suppose there exists $g \in G$ such that $ga = bg$. Since $(G, \mathfrak{S})$ has unbounded products, we have $\mathrm{Big}(a) = \mathrm{Big}(b) = \{S\}$. For simplicity of notation, we denote the asymptotic translation length of b in $\mathcal{C}S$ by $\tau(b)$. Note that S is fixed by the action of G on $\mathfrak{S}$. Since g conjugates a^i to b^i for any $i \in \mathbb{Z}$, we first replace a and b by sufficiently high powers so that $\tau(b) \geq L_S\delta$. By Remark 1.23, such a power can be chosen uniformly (that is, depending only on the hierarchy constants, and not the choice of elements a and b).

Let $\gamma_b = \gamma_b^S$ be a $(2, \ell)$-quasigeodesic axis of b in $\mathcal{C}S$. Then $\gamma_a = \gamma_a^S = g^{-1}\gamma_b$ is a $(2, \ell)$-quasigeodesic axis of a in $\mathcal{C}S$. We now fix a quadrilateral of (λ, λ)-hierarchy paths in G: $\mu(1, g)$, $\mu(1, b^K)$, $b^K\mu(1, g) = \mu(b^K, b^K g)$, and $g\mu(1, b^K) = \mu(g, gb^K) = \mu(g, a^K g)$.

Our first step is to replace g with a different conjugator whose length we are able to bound in G. Since $K \geq K_0$, we may apply Lemma 1.24 to each of the axes γ_a, γ_b in $\mathcal{C}S$ and the points $1, a^K \in G$ and $1, b^K \in G$, respectively. This yields a point $z' \in \gamma_a$ and a point $w' \in \gamma_b$ such that $z' \in \mathcal{N}_L(\pi_S(\mu(1, a^K)))$ and $w' \in \mathcal{N}_L(\pi_S(\mu(1, b^K)))$. Moreover, if x is a point on γ_a nearest to $\pi_S(1)$ and y is a point on γ_b nearest to $\pi_S(1)$, then $\mathsf{d}_S(z', x) \leq L$ and $\mathsf{d}_S(w', y) \leq L$. See Figure 2. Let $z \in \pi_S(\mu(1, a^K))$ and $w \in \pi_S(\mu(1, b^K))$ be points nearest to z' and w', respectively. Since g fixes S, we have $gz \in \pi_S(\mu(g, ga^K)) = \pi_S(\mu(g, b^K g)) \subseteq \mathcal{C}S$ and $gz' \in g\gamma_a = \gamma_b$. Since g is an isometry, we have $\mathsf{d}_S(gz', gx) = \mathsf{d}_S(z', x) \leq L$.

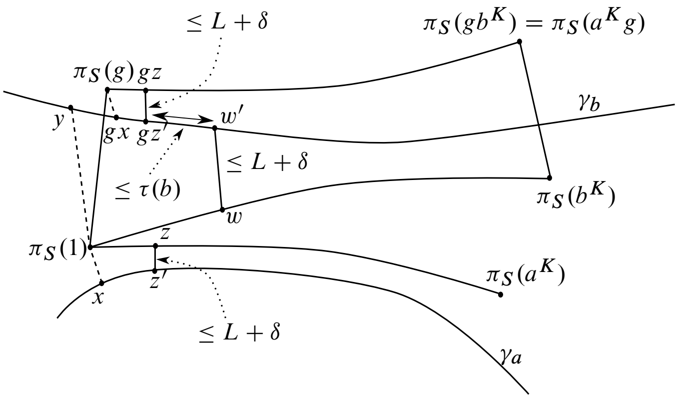

Figure 2. The geometry of the axes of a and b in $\mathcal{C}S$.

By possibly premultiplying g by a power of b, we may assume that $\mathsf{d}_S(gz', w') \leq \tau(b)$ (while still conjugating a to b). Thus we have

$$\mathsf{d}_S(y, gx) \leq \mathsf{d}_S(y, w') + \mathsf{d}_S(w', gz') + \mathsf{d}_S(gz', gx) \leq \tau(b) + 2L. \tag{2.2}$$

Our goal is to bound the length of this new conjugator, which, by abuse of notation, we will still call g.

We will show that for each $U \in \mathfrak{S}$, we have

$$\mathsf{d}_U(1, g) \leq 2K\mathsf{d}_U(1, b) + \mathsf{d}_U(g, b^K g) + K, \tag{2.3}$$

where K is as in (2.1). After establishing this bound for each $U \in \mathfrak{S}$, we apply the distance formula (Theorem 1.7) with threshold R to obtain

$$\mathsf{d}_G(1, g) \preceq_{C,D} 2K\mathsf{d}_G(1, b) + \mathsf{d}_G(g, b^K g) + K.$$

Finally, we use the fact that $\mathsf{d}_G(g, b^K g) = \mathsf{d}_G(g, ga^K) = \mathsf{d}_G(1, a^K) \leq K\mathsf{d}_G(1, a)$, which establishes that

$$\mathsf{d}_G(1, g) \preceq_{C,D} 2K\mathsf{d}_G(1, b) + K\mathsf{d}_G(1, a) + K,$$

where C, D are the constants given by the distance formula (Theorem 1.7). (Note that by assumption, R is sufficiently large to serve as a threshold in the distance formula.) This will provide the desired bound in G.

Fix $U \in \mathfrak{S}$. If $U \notin \mathbf{Rel}(1, g; R)$, then we have $\mathsf{d}_U(1, g) \leq R \leq K$, and (2.3) holds. Thus we assume for the rest of the proof that $U \in \mathbf{Rel}(1, g; R)$. There are two cases to consider: either $U = S$ or $U \subsetneq S$. We will deal with each of these possibilities individually.

Case 1: $U = S$. In this case, we have (as seen in Figure 2):

$$\begin{aligned}
\mathsf{d}_S(1, g) &\leq \mathsf{d}_S(1, w) + \mathsf{d}_S(w, w') + \mathsf{d}_S(w', gz') + \mathsf{d}_S(gz', gz) + \mathsf{d}_S(gz, g) \\
&\leq \mathsf{d}_S(1, b^K) + 2L + \tau_S(b) + \mathsf{d}_S(g, b^K g) \\
&\leq 2K\mathsf{d}_S(1, b) + \mathsf{d}_S(g, b^K g) + K,
\end{aligned}$$

where the final inequality follows from the fact that $\mathsf{d}_S(1,b) \geq \tau_S(b)$ and (2.1). Therefore, (2.3) holds in this case.

Case 2: $U \subsetneq S$. As we are assuming that U is R-relevant for $1, g$, we must have $\rho_S^U \subseteq \mathcal{C}S$ is contained in the E-neighborhood of a geodesic in $\mathcal{C}S$ from $\pi_S(1)$ to $\pi_S(g)$ by the bounded geodesic image axiom (Definition 1.6 (7)). As geodesic quadrilaterals in δ-hyperbolic spaces are 2δ-thin, it follows that ρ_S^U is contained in the $(E+2\delta)$-neighborhood of $[\pi_S(1), y] \cup [y, gx] \cup [gx, \pi_S(g)]$, where these are geodesics in $\mathcal{C}S$. Since x and gx are the nearest point projections of $\pi_S(1)$ and $\pi_S(g)$ onto γ_b, respectively, it follows from (2.2) that the projection of $[\pi_S(1), y] \cup [y, gx] \cup [gx, \pi_S(g)]$ onto γ_b has diameter at most $\tau(b) + 2L$. In particular, since nearest point projection maps in hyperbolic spaces are Lipschitz, the nearest point on γ_b to ρ_S^U is distance at most $\tau(b) + 2L + E$ from y.

By an analogous argument, if U is also R-relevant for $b^K, b^K g$, we must have that ρ_S^U is contained in the $(E+2\delta)$-neighborhood of $[\pi_S(b^K), b^k y] \cup [b^k y, b^k gx] \cup [b^k gx, \pi_S(b^k g)]$. In particular, the nearest point on γ_b to ρ_S^U is at distance at most $\tau(b) + 2L + E$ from $b^K y$.

However, our choice of K in (2.1) ensures that

$$\mathsf{d}_S(y, b^K y) \geq K\tau(b) \geq \Big(\Big\lceil \frac{4L+3E}{T} \Big\rceil + 2\Big)\tau(b) \geq 4L + 3E + 2\tau(b),$$

which is a contradiction. Thus U is not R-relevant for $b^K, b^K g$, and so

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

Therefore,

$$\begin{aligned}
\mathsf{d}_U(1,g) &\leq \mathsf{d}_U(1,b^K) + \mathsf{d}_U(b^K, b^K g) + \mathsf{d}_U(b^K g, g) \\
&\leq K\mathsf{d}_U(1,b) + \mathsf{d}_U(b^K, b^K g) + R \\
&\leq K\mathsf{d}_U(1,b) + \mathsf{d}_U(b^K, b^K g) + K,
\end{aligned}$$

where the final inequality follows from our choice of K in (2.1).

Hence (2.3) holds in this case, which completes the proof of Theorem A. ■

3. A family of hierarchically hyperbolic groups

In this section, we highlight three properties which isolate some of the nice features of compact special groups and which appear in many other contexts as well. We will show in Proposition 3.10 that many hierarchically hyperbolic groups satisfy these three properties, which we call $\mathbf{F}_U$ *stabilizers*, *orthogonal decomposition*, and *commutativity*.

Fix a hierarchically hyperbolic group $(G, \mathfrak{S})$. If $\mathcal{U} \subseteq \mathfrak{S}$ is a collection of pairwise orthogonal domains, we denote the *container* of $\mathcal{U}$ in S by $C_{\mathcal{U}}$ (Definition 1.6 (3)); by definition, each domain V which is orthogonal to every $U \in \mathcal{U}$ is nested into $C_{\mathcal{U}}$. We say G has *clean containers* if for every collection of pairwise orthogonal domains $\mathcal{U}$, the container $C_{\mathcal{U}}$ is orthogonal to every $U \in \mathcal{U}$. If $\mathcal{U} = \{U\}$, we write C_U instead of $C_{\{U\}}$.

Recall that for any domain $U \in \mathfrak{S}$, we identify $\mathbf{P}_U$ with $\mathbf{F}_U \times \mathbf{E}_U$ (see the discussion after Definition 1.17). If a subgroup $H \leq G$ fixes a domain $U \in \mathfrak{S}$ (in the action of G on $\mathfrak{S}$), then whenever $V \sqsubseteq U$ or $V \perp U$, we have $hV \sqsubseteq U$ or $hV \perp U$, respectively, for each $h \in H$. It follows that H stabilizes the product region $\mathbf{P}_U$ and each of its factors $\mathbf{F}_U$ and $\mathbf{E}_U$.

Definition 3.1. For any $U \in \mathfrak{S}$, let G_U be the subgroup of G that fixes U in the action of G on $\mathfrak{S}$ and that stabilizes $\mathbf{F}_U \times \{e\}$ for each $e \in \mathbf{E}_U$.

Equivalently, G_U is the subgroup which stabilizes each factor of $\mathbf{F}_U \times \mathbf{E}_U$ and acts as the identity on the second factor. We note that when G has clean containers, the second factor $\mathbf{E}_U$ is isometric to $\mathbf{F}_{C_U}$ by Lemma 3.6.

Example 3.2. Right-angled Artin groups and, more generally, compact special groups, provide a good example to have in mind when reading this section. With the standard hierarchically hyperbolic group structure given in [5], such groups are hierarchically acylindrical, and have clean containers [1, Proposition 7.2]. One nice property of right-angled Artin groups is that two elements commute if and only if all the generators in a cyclically reduced factorization of one of the elements commute with all the generators in a cyclically reduced factorization of the other element. Hence, in the Salvetti complex of a right-angled Artin group, G, we have that two elements span a periodic plane if and only if they commute. Similarly, if a group is compact special it embeds as a quasiconvex subgroup of a right-angled Artin group and thus inherits this property as well. Further, if a group is virtually compact special, then, up to taking powers, two elements commute if and only if they span a periodic plane. For these groups $U, V \in \mathfrak{S}$ are orthogonal if and only if they have associated subcomplexes of the cube complex which span a direct product. Hence, it follows that given $U, V \in \mathfrak{S}$ which are orthogonal, the subgroup which fixes U in the action on $\mathfrak{S}$ and which stabilizes the subset $\mathbf{F}_U \times \{e\}$ for each $e \in \mathbf{E}_U$ has the property that it commutes with the similarly defined subset for V. In other words, elements of G_U and G_V commute. In particular, if $g \in G$ fixes each $U_i \in \mathrm{Big}(G)$, then g can be written as a product of elements in G_{U_i}.

3.1. The $\mathbf{F}_U$ stabilizers, orthogonal decomposition, and commutative properties

We will now extract and formalize the properties which we described above for right-angled Artin groups.

3.1.1. The $\mathbf{F}_U$ stabilizers property. Since $(G, \mathfrak{S})$ is a hierarchically hyperbolic group, there is a finite fundamental domain $\mathfrak{S}'$ for the action of G on $\mathfrak{S}$. We may choose $\mathfrak{S}'$ to have the property that for each $U \in \mathfrak{S}'$, there exists $e \in \mathbf{E}_U$ such that $1 \in \mathbf{F}_U \times \{e\}$, where 1 is the identity element of G. We denote this copy of $\mathbf{F}_U$ by $\mathbb{F}_U$. For such domains U, we always have $G_U \subseteq \mathbb{F}_U$. To see this, consider any $f \notin \mathbb{F}_U$. Since $1 \in \mathbb{F}_U$ and $f = f \cdot 1 \notin \mathbb{F}_U$, the element f does not stabilize $\mathbb{F}_U$, so $f \notin G_U$.

The first property says that for all $U \in \mathfrak{S}'$, the sets $\mathbb{F}_U$ and G_U are coarsely equal.

Definition 3.3 ($\mathbf{F}_U$ stabilizers). A hierarchically hyperbolic group $(G, \mathfrak{S})$ satisfies the $\mathbf{F}_U$ *stabilizers property* if there exists a constant ν depending only on the hierarchy constants such that $\mathrm{d}_G(f, G_U) \le \nu$ for each $U \in \mathfrak{S}'$ and any $f \in \mathbb{F}_U$.

The $\mathbf{F}_U$ stabilizers property implies that for domains $U \in \mathfrak{S}'$, the subgroup G_U inherits many geometric properties from $\mathbb{F}_U$, including hierarchical quasiconvexity. In a hierarchically hyperbolic group, there is a function $k\colon [0,\infty) \to [0,\infty)$ so that for any $U \in \mathfrak{S}$, the subspace $\mathbf{F}_U$ is k-hierarchically quasiconvex [3, Construction 5.10]. If the group has the $\mathbf{F}_U$ stabilizers property, then since G_U and $\mathbb{F}_U$ are at uniformly bounded distance whenever $U \in \mathfrak{S}'$, there is a function $k'\colon [0,\infty) \to [0,\infty)$ depending only on k and E so that the subgroup G_U is also hierarchically quasiconvex for any $U \in \mathfrak{S}'$. It then follows from [3, Lemma 5.5] that there is a well-defined gate map $\mathfrak{g}_{G_U}\colon G \to G_U$. Moreover, for any $g \in G$, each coset gG_U of G_U is also k'-hierarchically quasiconvex in G, so we also have a well-defined gate map $\mathfrak{g}_{gG_U}\colon G \to gG_U$. These gate maps will be important for defining the two additional properties we introduce in this section.

The next lemma says that $(G_U, \mathfrak{S}_U)$ is a hierarchically hyperbolic group, where $\mathfrak{S}_U = \{V \in \mathfrak{S} \mid V \sqsubseteq U\}$.

Lemma 3.4. *Let $(G, \mathfrak{S})$ be a hierarchically hyperbolic group satisfying the $\mathbf{F}_U$ stabilizers property. For any $U \in \mathfrak{S}'$, $(G_U, \mathfrak{S}_U)$ is a hierarchically hyperbolic group.*

Proof. The $\mathbf{F}_U$ stabilizers property says that the subgroup G_U is at uniformly bounded distance from $\mathbb{F}_U$. In particular, $\mathbb{F}_U$ and G_U are quasi-isometric. Since $(\mathbb{F}_U, \mathfrak{S}_U)$ is a hierarchically hyperbolic space [3, Proposition 5.11], this immediately implies that $(G_U, \mathfrak{S}_U)$ is a hierarchically hyperbolic space, where the associated hyperbolic spaces and maps are the same as those for $(\mathbb{F}_U, \mathfrak{S}_U)$. It remains to show that $(G_U, \mathfrak{S}_U)$ is a hierarchically hyperbolic *group*. For this, note that G_U stabilizes $\mathfrak{S}_U$ by definition. Since $(G, \mathfrak{S})$ is a hierarchically hyperbolic group and $G_U \le G$, the four additional conditions from Definition 1.11 hold because they hold for the action of G on $\mathfrak{S}$. For example, since G acts cofinitely on $\mathfrak{S}$ and preserves the relations $\sqsubseteq$, $\pitchfork$, and $\perp$, so does G_U. Similar arguments show the other three conditions hold. ∎

3.1.2. The orthogonal decomposition property. The next property allows any infinite order element which fixes a collection of pairwise-orthogonal domains $\{U_1, \dots, U_k\}$ to be decomposed into a product of elements in G_{U_i}. Before defining this property, the following lemma establishes that for each $i = 1, \dots, k$, there is a preferred $\mathbf{F}_{U_i} \times \{e_i\}$ which we denote by $\mathbb{F}_{U_i}$. The careful reader will note that if U_i is already in the fundamental domain $\mathfrak{S}'$, then the choice given by the lemma is consistent with our previous choice of $\mathbb{F}_{U_i}$.

Lemma 3.5. *Let $(G, \mathfrak{S})$ be a hierarchically hyperbolic group with the $\mathbf{F}_U$ stabilizers property, and let $\mathcal{U} = \{U_1, \dots, U_k\}$ be a maximal collection of pairwise-orthogonal domains in $\mathfrak{S}$. Then there exist $t \in G$ and copies $\mathbf{F}_{U_i} \times \{e_i\}$ such that the following hold for all i:*

- *$U_i = tU_i'$ for some $U_i' \in \mathfrak{S}'$;*
- *$\mathbf{F}_{U_i} \times \{e_i\} = t\mathbb{F}_{U_i'}$;*
- *$G_{U_i} = tG_{U_i'}t^{-1}$;*
- *$\mathsf{d}_G(t, \mathfrak{g}_{\mathbf{P}_{\mathcal{U}}}(1)) \leq E\nu$, where ν is the constant from Definition 3.3.*

Proof. Consider the product region $\mathbf{P}_{\mathcal{U}}$ associated to $\mathcal{U}$, and let t' be any point in $\mathfrak{g}_{\mathbf{P}_{\mathcal{U}}}(1)$. For the first part of the proof, it will be convenient to distinguish between the abstract product region $\mathbf{P}_{\mathcal{U}} = \mathbf{F}_{U_1} \times \cdots \times \mathbf{F}_{U_k}$ and its image $\phi_{\mathcal{U}}(\mathbf{P}_{\mathcal{U}}) \subseteq G$ (see the discussion after Definition 1.17). Let $(t_1', \dots, t_k') \in \mathbf{P}_{\mathcal{U}}$ be such that $\phi_{\mathcal{U}}(t_1', \dots, t_k') = t'$. We will adjust each t_i' individually to find a new point $(t_1, \dots, t_k)$, which will determine the points e_i in the statement. At the ith stage, we change the ith coordinate of the point in $\mathbf{P}_{\mathcal{U}}$ to ensure that it lies in a coset of G_{U_i} that is completely contained in the associated copy of $\mathbf{F}_{U_i}$. In subsequent steps, we will adjust later coordinates: this may change which coset of G_{U_i} the point lies in, but it will simultaneously translate the copy of $\mathbf{F}_{U_i}$ so that this new coset is still contained in the new copy of $\mathbf{F}_{U_i}$, as desired. After changing all coordinates, the desired element t will be $\phi_{\mathcal{U}}(t_1, \dots, t_k)$.

We begin with $i = 1$. Since $\mathfrak{S}'$ is a fundamental domain, there are some $f_1' \in G$ and $U_1' \in \mathfrak{S}'$ such that $\phi_{\mathcal{U}}(\mathbf{F}_{U_1}, t_2', \dots, t_k') = f_1'\mathbb{F}_{U_1'}$. Since $G_{U_1'} \subseteq \mathbb{F}_{U_1'}$, we have $f_1'G_{U_1'} \subseteq \phi_{\mathcal{U}}(\mathbf{F}_{U_1}, t_2', \dots, t_k')$. By the $\mathbf{F}_U$ stabilizers property, there is an element $t_1 \in \mathbf{F}_{U_1}$ with $\mathsf{d}_G(t', \phi_{\mathcal{U}}(t_1, t_2', \dots, t_k')) \leq \nu$ and $\phi_{\mathcal{U}}(t_1, t_2', \dots, t_k') \in f_1'G_{U_1'}$.

We fix t_1 from the previous paragraph and consider $i = 2$. The point $\phi_{\mathcal{U}}(t_1, t_2', \dots, t_k')$ is in $\phi_{\mathcal{U}}(t_1, \mathbf{F}_{U_2}, t_3', \dots, t_k')$. Again, as above, there are some $f_2' \in G$ and $U_2' \in \mathfrak{S}'$ for which $\phi_{\mathcal{U}}(t_1, \mathbf{F}_{U_2}, t_3', \dots, t_k') = f_2'\mathbb{F}_{U_2'}$. Also, as above, we can find an element $t_2 \in \mathbf{F}_{U_2}$ with $\mathsf{d}_G(\phi_{\mathcal{U}}(t_1, t_2', \dots, t_k'), \phi_{\mathcal{U}}(t_1, t_2, t_3', \dots, t_k')) \leq \nu$ and $\phi_{\mathcal{U}}(t_1, t_2, t_3', \dots, t_k') \in f_2'G_{U_2'}$.

Continuing in this way for each i yields a point $(t_1, t_2, \dots, t_k) \in \mathbf{F}_{U_1} \times \cdots \times \mathbf{F}_{U_k}$. Letting $t = \phi_{\mathcal{U}}(t_1, \dots, t_k)$, it follows from the triangle inequality that

$$\mathsf{d}_G(t', t) \leq k\nu \leq E\nu, \tag{3.1}$$

where the final inequality holds because any collection of pairwise orthogonal domains has cardinality bounded by E.

We now return to our convention of identifying $\mathbf{P}_{\mathcal{U}}$ with its image $\phi_{\mathcal{U}}(\mathbf{P}_{\mathcal{U}}) \subseteq G$. We have shown that, for each i, we have $t \in \mathbf{F}_{U_i} \times \{e_i\}$ for some e_i. Precisely,

$$e_i = \phi_U^{\perp}(t_1, \dots, \hat{t}_i, \dots, t_k),$$

where $\hat{t}_i$ indicates that the term t_i does not appear in the tuple.

We now show that $\mathbf{F}_{U_i} \times \{e_i\}$ satisfies the conclusion of the lemma for each i. The final bullet point holds by (3.1).

There is an element $f_i \in G$ such that $\mathbf{F}_{U_i} \times \{e_i\} = f_i \mathbb{F}_{U_i'}$, where $U_i = f_i U_i'$, and $t \in f_i G_{U_i}$. Thus $f_i G_{U_i'} = t G_{U_i'}$, and $G_{U_i} = t G_{U_i} t^{-1}$, so the third bullet point holds. Also, since $t = f_i q_i$ for some $q_i \in G_{U_i'}$, we have

$$t\mathbb{F}_{U_i'} = f_i q_i \mathbb{F}_{U_i'} = f_i \mathbb{F}_{U_i'},$$

so the second bullet point holds. Finally, $tU_i' = f_i q_i U_i' = f_i U_i' = U_i$, which shows that the first bullet point holds and concludes the proof of the lemma. ■

The following lemma is presumably well known, but is not in the literature. An immediate corollary of this is that an axial element fixes the container associated to its big set.

Lemma 3.6. *Let $(G, \mathfrak{S})$ be a hierarchically hyperbolic group with clean containers, and let $\{U_1, \dots, U_k\}$ be a (non-maximal) collection of pairwise orthogonal domains. There exists a unique $C \in \mathfrak{S}$ such that if for each i, a domain $V \in \mathfrak{S}$ satisfies $V \perp U_i$, then $V \sqsubseteq C$.*

Proof. First, by Definition 1.6 (3), some C exists with the desired property, what is needed is to prove uniqueness. So suppose that both C and C' satisfy this property. Since the containers are clean, each of C and C' is orthogonal to U_i for each i. Thus, since C is a container and since C' is orthogonal to all the U_i, we must have that $C' \sqsubseteq C$. Similarly, $C \sqsubseteq C'$. Thus $C = C'$, as desired. ■

Definition 3.7 (Orthogonal decomposition). Let $(G, \mathfrak{S})$ be a hierarchically hyperbolic group with clean containers which satisfies the $\mathbf{F}_U$ stabilizers property, and let $h \in G$ be an infinite order element. Let $\{U_1, \dots, U_{k+1}\}$ be a maximal collection of pairwise orthogonal domains of $\mathfrak{S}$ so that $\mathrm{Big}(h) = \{U_1, \dots, U_k\}$ and U_{k+1} is the container associated to $\mathrm{Big}(h)$ in S. Suppose $h \in G$ fixes $\mathrm{Big}(h)$ elementwise. By Lemma 3.5, there exist $t \in G$ and, for each $i = 1, \dots, k$, a domain $U_i' \in \mathfrak{S}'$ with $U_i = tU_i'$. The label of the vertex $\mathfrak{g}_{tG_{U_i}}(h)$ is th_i' for some $h_i' \in G_{u_i'} \subseteq \mathbb{F}_{U_i'}$. Define

$$h_{U_i} := th_i't^{-1} \in tG_{U_i'}t^{-1} = G_{U_i}. \tag{3.2}$$

The group $(G, \mathfrak{S})$ satisfies the *orthogonal decomposition property* if the following two properties hold for all axial elements $h \in G$. First, there is a uniform lower bound on the translation length $\tau_{U_i}(h)$ for each $U_i \in \mathrm{Big}(h)$ (this uniformity only depends on the hierarchy constants and not the choice of h). Second, after possibly relabeling the domains of $\mathrm{Big}(h)$, we have

$$h = h_{U_1} h_{U_2} \dots h_{U_k} = th_1' \dots h_k' t^{-1}.$$

We say $h_{U_1} h_{U_2} \dots h_{U_k}$ is a *decomposition* of h.

This decomposition may depend on the order of the factors. In particular, it may be the case that h_{U_i} does not commute with h_{U_j}, because elements of G_{U_i} and G_{U_j} may not commute. However, the final property we discuss will require that such elements do commute, and so the order of the factors will not be important for the groups we consider.

3.1.3. The commutative property. The final property ensures that G_U and G_V commute whenever $U \perp V$.

Definition 3.8 (Commutative property). A hierarchically hyperbolic group $(G, \mathfrak{S})$ with the $\mathbf{F}_U$ stabilizers property satisfies the *commutative property* if $[G_U, G_V] = 1$ whenever $U \perp V$.

The following lemma is a consequence of the commutative property.

Lemma 3.9. *Let G be a hierarchically hyperbolic group satisfying the $\mathbf{F}_U$ stabilizers, orthogonal decomposition, and commutative properties. Let $h \in G$ be an axial element which fixes* $\mathrm{Big}(h) = \{H_1, \dots, H_k\}$ *elementwise, and let C be the clean container associated to* $\mathrm{Big}(h)$. *Then there exists a uniform constant K such that $(h^K)_C = 1$, where $(h^K)_C$ is the factor corresponding to C in the decomposition of h^K with respect to $\{H_1, \dots, H_k, C\}$ and* 1 *is the identity element of $G_C \leq G$.*

Proof. First, note that by Lemma 3.6, h fixes $\{H_1, \dots, H_k, C\}$ elementwise, and so the decomposition $h = h_{H_1} \dots h_{H_k} h_C$ of h with respect to this set is well defined. Recall that $h_C \in G_C$ is an element of the hierarchically hyperbolic group $(G_C, \mathfrak{S}_C)$. Since $C \notin \mathrm{Big}(h)$, h_C is not an axial element of G_C. Therefore, h_C must be finite order by Lemma 1.21. By [20, Theorem G], there are finitely many conjugacy classes of finite order elements in a hierarchically hyperbolic group, and therefore there is a uniform constant K such that h_C^K is the identity element of G_C.

By the commutative property, we have

$$h^K = (h_{H_1} \dots h_{H_k} h_C)^K = (h_{H_1})^K \dots (h_{H_k})^K (h_C)^K = (h_{H_1})^K \dots (h_{H_k})^K.$$

From this decomposition, it is clear that $(h^K)_C = 1$. ∎

3.1.4. Examples. We now give several examples of hierarchically hyperbolic groups satisfying the three properties defined above. Moreover, additional examples can be built using combination theorems, of which there are several in the literature (see, for instance, [3, 7, 8, 27]).

Proposition 3.10. *Let Ξ be the set of hierarchically hyperbolic groups with clean containers which satisfy the $\mathbf{F}_U$ stabilizers, orthogonal decomposition, and commutative properties. Then the following groups are in Ξ:*

(1) *hyperbolic groups,*

(2) *compact special groups,*

(3) *groups hyperbolic relative to a collection of groups in Ξ, and*

(4) *direct products of groups in Ξ.*

Proof. We consider each class of groups in turn.

(1) The statement is immediate for hyperbolic groups G, as they all admit hierarchically hyperbolic structures with a single domain S, and the action on $\mathcal{C}S$ is acylindrical.

For this domain, $\mathbf{F}_S$ is a Cayley graph of the group and $G_S = G$. As there is no orthogonality, the orthogonal decomposition and commutative properties vacuously hold.

(2) For compact special groups, we use the standard structure described in [3]. This structure satisfies the three properties by a completely analogous argument to the one given for right-angled Artin groups in Example 3.2.

(3) Let G be a group which is hyperbolic relative to a collection $\mathcal{P}$ of hierarchically hyperbolic groups with clean containers satisfying the $\mathbf{F}_U$ stabilizers, orthogonal decomposition, and commutative properties. Then G is a hierarchically hyperbolic group by [3, Theorem 9.1] and has clean containers by [1, Proposition 7.4]. For each $P \in \mathcal{P}$, let $(P, \mathfrak{S}_P)$ be a hierarchically hyperbolic group structure for P, and for each left coset gP, let $\mathfrak{S}_{gP}$ be a copy of $\mathfrak{S}_P$, with the associated hyperbolic spaces and projections. Let $\widehat{G}$ be the hyperbolic space formed from G by coning off each left coset of each $P \in \mathcal{P}$. Then the hierarchically hyperbolic group structure on G is given by $\mathfrak{S} = \{\widehat{G}\} \sqcup_{gP \in G\mathcal{P}} \mathfrak{S}_{gP}$. The domain $\widehat{G}$ is the unique $\sqsubseteq$-maximal domain, and if $U \in \mathfrak{S}_{gP}$ and $V \in \mathfrak{S}_{g'P'}$, where $gP \neq g'P'$, then $U \pitchfork V$. We refer the reader to [3, Section 9] for details of this structure, but note one important feature of the structure $(G, \mathfrak{S})$: any pair of orthogonal domains are contained in some $gP \in G\mathcal{P}$.

We first check that the $\mathbf{F}_U$ stabilizes property holds. A fundamental domain for the action of G on $\mathfrak{S}$ is given by $\mathfrak{S}' = \{\widehat{G}\} \sqcup_{P \in \mathcal{P}} \mathfrak{S}_P$. Let $U \in \mathfrak{S}'$. If $U = \widehat{G}$, then $\mathbf{F}_U = G$ and $G_U = G$, so the property holds for this domain. Now suppose $U \in \mathfrak{S}_P$ for some $P \in \mathcal{P}$. Then $\mathbf{F}_U \subseteq P$. If $g \notin P$, then $g\mathbf{F}_U \subseteq gP \neq P$, and so $g \notin G_U$. Therefore, G_U is a subgroup of P in this case. Since $(P, \mathfrak{S}_P)$ satisfies the $\mathbf{F}_U$ stabilizers property, it follows that $(G, \mathfrak{S})$ does, as well.

We now check the orthogonal decomposition property. Since G is hyperbolic relative to $\mathcal{P}$, every infinite order element $h \in G$ is either loxodromic with respect to the action of G on $\widehat{G}$, in which case $\mathrm{Big}(h) = \{\widehat{G}\}$ or is conjugate into some $P \in \mathcal{P}$, in which case we consider the conjugate $ghg^{-1} \in P$. In the first case, the action of G on $\widehat{G}$ is acylindrical, and so there is a uniform lower bound on the translation length of h, and we have the trivial orthogonal decomposition of h. In the second case, there is a uniform lower bound on the translation length of ghg^{-1} in each domain in $\mathrm{Big}(ghg^{-1})$ by the assumption that each P satisfies the orthogonal decomposition property. Translation length is invariant under conjugacy, and so we obtain a uniform lower bound on the translation length of h in each domain in $\mathrm{Big}(h)$. There is also an orthogonal decomposition of ghg^{-1} coming from the assumption on $(P, \mathfrak{S}_P)$. Since $\mathrm{Big}(h) = g^{-1}\,\mathrm{Big}(ghg^{-1})$, conjugating each term in the decomposition of ghg^{-1} by g^{-1} yields an orthogonal decomposition for h.

Finally, the commutative property follows immediately from the construction of the orthogonal decomposition in the previous paragraph and the fact that $(P, \mathfrak{S}_P)$ satisfies the commutative property for each $P \in \mathcal{P}$.

(4) Assume $G = G_1 \times G_2$, and suppose $(G_1, \mathfrak{S}_1)$, $(G_2, \mathfrak{S}_2)$ are hierarchically hyperbolic groups with clean containers which satisfy the $\mathbf{F}_U$ stabilizers, orthogonal decomposition, and commutative properties. Then G is a hierarchically hyperbolic group by [3,

Proposition 8.27] and has clean containers by [1, Proposition 7.3]. The hierarchy structure on G is given by $\mathfrak{S} = \{S, U_1, U_2\} \sqcup \mathfrak{S}_1 \sqcup \mathfrak{S}_2 \sqcup \{V_U \mid U \in \mathfrak{S}_1 \cup \mathfrak{S}_2\}$, where S is the unique $\sqsubseteq$-maximal element, U_i is a domain into which all domains in $\mathfrak{S}_i$ nest, and for each $U \in \mathfrak{S}_i$, the domain V_U is a domain into which all domains in $\mathfrak{S}_j$ with $j \neq i$ and all domains in $\mathfrak{S}_i$ orthogonal to U nest. The only important relation between domains for this proof is orthogonality. In addition to any orthogonality among domains in $\mathfrak{S}_1$ or $\mathfrak{S}_2$, we have that all domains in $\mathfrak{S}_1$ are orthogonal to all domains in $\mathfrak{S}_2$, $U_1 \perp U_2$, and $V_U \perp U$ for each $U \in \mathfrak{S}_1 \cup \mathfrak{S}_2$. By construction, $(G, \mathfrak{S})$ has clean containers. See [3, Section 8] for further details on this structure.

When we refer to subsets of G_i or the $(G_i, \mathfrak{S}_i)$ structure, we append a superscript i to the notation. For example, if $U \in \mathfrak{S}_i$, then $\mathbf{F}_U^i$ is the corresponding subset of G_i.

We first check the $\mathbf{F}_U$ stabilizers property. If $U = S$, there is nothing to check, so suppose first that $U \in \mathfrak{S}_1$. Let G_U^1 denote the subgroup from the structure $(G_1, \mathfrak{S}_1)$ which stabilizers $\mathbf{F}_U^1 \times \{e\}$ for each $e \in \mathbf{E}_U^1$. In the structure $(G, \mathfrak{S})$, there are additional domains orthogonal to U; in particular, every domain in $\mathfrak{S}_2$ is orthogonal to U. We have $\mathbf{F}_U = \mathbf{F}_U^1$, but now $\mathbf{E}_U = \mathbf{E}_U^1 \times G_2$. Therefore, we have $(g_1, g_2) \in G_U$ if and only if $g_1 \in G_U^1$ and $g_2 = 1$. Thus $G_U \simeq G_U^1 \times \{1\}$. Since $(G_1, \mathfrak{S}_1)$ satisfies the $\mathbf{F}_U$ stabilizers property, G_U^1 is coarsely equal to $\mathbf{F}_U^1$. The above discussion then implies that G_U is coarsely equal to $\mathbf{F}_U$. Similarly, if $U \in \mathfrak{S}_2$, then $G_U \simeq \{1\} \times G_U^2$, and we again have that G_U is coarsely equal to $\mathbf{F}_U$.

Suppose next that $U = U_1$. Then $\mathbf{F}_U = G_1$, and $\mathbf{E}_U = G_2$. Since G is the direct product of G_1 and G_2, we have that $G_{U_1} = G_1$, and so G_{U_1} is coarsely equal to $\mathbf{F}_U$. The analogous argument holds if $U = U_2$.

Finally, fix $U \in \mathfrak{S}_1$, and consider the domain V_U. Let C_U be the container associated to U in the $\sqsubseteq$-maximal domain of $\mathfrak{S}_1$. Then $\mathbf{F}_{V_U} = \mathbf{E}_U = \mathbf{E}_U^1 \times G_2 = \mathbf{F}_{C_U}^1 \times G_2$, and $\mathbf{E}_{V_U} = \mathbf{E}_{C_U}^1$. It follows that $G_{V_U} \simeq G_C^1 \times G_2$. Since $\mathbf{F}_{C_U}^1$ is coarsely equal to $G_{C_U}^1$, we also have that $\mathbf{F}_{V_U}$ is coarsely equal to G_{V_U}, as desired. An analogous argument holds if we fix $U \in \mathfrak{S}_2$. Therefore, $(G, \mathfrak{S})$ satisfies the $\mathbf{F}_U$ stabilizers property.

The orthogonal decomposition and commutative properties both follow immediately because they hold in each $(G_i, \mathfrak{S}_i)$ and G_1 and G_2 commute. ■

The $\mathbf{F}_U$ stabilizers, orthogonal decomposition, and commutative properties all involve orthogonality and properties of product regions. Hence, intuitively, if a combination theorem does not add any additional orthogonality relations (or only in a trivial way, such as by adding domains whose associated hyperbolic space is bounded diameter), then such a combination of groups in Ξ should, in general, yield a group in Ξ. For example, we expect that trees of groups in Ξ satisfying the hypotheses of the combination theorem in [3, Theorem 8.6] are also in Ξ. In particular, combined with Proposition 3.10 (3), (4), this would show that for hierarchically hyperbolic groups $\pi_1(M)$ where M is the fundamental group of compact 3-manifolds with no Nil or Sol in its prime decomposition, then $\pi_1(M)$ is in Ξ.

3.2. A non-example: the mapping class group

We briefly explain why the standard hierarchy structure on the mapping class group fails to satisfy the $\mathbf{F}_U$ stabilizers property. Notwithstanding this fact, we believe that a modification of the properties from this section can be used to make the present approach work for the mapping class group, as well. We do not carry this out, though, because the approaches we see for doing so are all technical, and the present results are already known for mapping class groups. We record this fact for those using these properties in the future with an eye towards other applications.

The standard hierarchically hyperbolic group structure $\mathfrak{S}$ on the mapping class group of a surface S is described in [3, Theorem 11.1]. The domains $U \in \mathfrak{S}$ correspond to homotopy classes of essential, not necessarily connected, *open* subsurfaces $U \subseteq S$. Two domains are orthogonal if the corresponding subsurfaces are disjoint. In particular, the annuli about the boundary curves of a subsurface do not intersect the subsurface; thus an annulus around a boundary curve is a domain orthogonal to the subsurface. A finite fundamental domain $\mathfrak{S}'$ for the action of $\mathcal{MCG}(S)$ on $\mathfrak{S}$ is provided by taking a collection of subsurfaces, one for each homeomorphism type of subsurface. For each $U \in \mathfrak{S}'$, $\mathbf{F}_U$ is coarsely equal to the mapping class group of the subsurface associated to U, and $\mathbf{E}_U$ is coarsely equal to the mapping class group of the complementary closed subsurface $S - U$.

One subtlety in the hierarchically hyperbolic structure on mapping class groups is that while elements of $\mathcal{MCG}(S)$ supported on disjoint subsurfaces commute, elements supported on disjoint *closed* subsurfaces are distinct, while two elements supported on disjoint *open* surfaces may coincide. A simple example of this is found by taking a product of elements in a once-punctured torus which generate the Dehn twists along the boundary. Taking the genus two surface obtained by doubling along the boundary curve, we see that we can generate the same Dehn twist by a product of elements on either of the open once-punctured tori separated by that curve.

Associated to a closed subsurface V, which includes its boundary components, is an element of $\mathfrak{S}$ consisting of the disjoint union of the interior of V, which we will denote by $\mathring{V}$, with annuli around the elements $\alpha_1, \ldots, \alpha_k$ of ∂V. The Dehn twist about a boundary curve in ∂V can be represented as a product of mapping class elements supported on the interior of V, even those these are orthogonal domains. Accordingly the stabilizer of $\mathring{V}$, in the action of G on $\mathfrak{S}$, is (possibly up to finite index if V is homeomorphic to $S - V$) a central extension of $\mathcal{MCG}(\mathring{V}) \times \mathcal{MCG}(S - V)$ by $\mathbb{Z}^k$, where $\mathbb{Z}^k$ is generated by Dehn twists along the boundary curves α_i, see, e.g., [9]. The domains $\mathring{V}$, $S - V$, and the annuli around each α_i form a maximal collection of pairwise orthogonal domains. If this was a semidirect product instead of a central extension, this would yield the $\mathbf{F}_U$ stabilizers and orthogonal decomposition properties. However, the fact that $\mathcal{MCG}(\mathring{V})$ does not act cocompactly on $\mathbf{F}_U$ means that the $\mathbf{F}_U$ stabilizers property does not hold in this structure.

We note, though, that any open subsurface $\mathring{U}$ is contained in a larger subsurface U obtained by taking the union of $\mathring{U}$ and all the annuli which bound $\mathring{U}$. For this subsurface U, the subgroup G_U of the mapping class group of S which stabilizes $\mathbf{F}_U$ and fixes $\mathbf{E}_U$ pointwise *can* be identified with $\mathcal{MCG}(U)$. This is a weaker version of the $\mathbf{F}_U$ stabilizers property. We expect that this weaker version might be useful in future work.[1]

3.3. Conjugators in hierarchically hyperbolic groups

We are now ready to prove Theorem D, which we restate for the convenience of the reader.

Theorem D. *Let $(G, \mathfrak{S})$ be a hierarchically hyperbolic group satisfying the $\mathbf{F}_U$ stabilizers, orthogonal decomposition, and commutative properties. 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*

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

Proof. Fix a hierarchically hyperbolic group $(G, \mathfrak{S})$ and a finite fundamental domain $\mathfrak{S}'$ for the action of G on $\mathfrak{S}$ as at the beginning of this section. Assume that $(G, \mathfrak{S})$ satisfies the $\mathbf{F}_U$ stabilizers, orthogonal decomposition, and commutative properties. For each $U \in \mathfrak{S}$, we fix $\mathbb{F}_U = \mathbf{F}_U \times \{e\}$ as described in Lemma 3.5.

We fix the same constants as in the beginning of the proof of Theorem A, and let σ be the Morse constant for (λ, λ)-quasigeodesics in a δ-hyperbolic space. Fix the function $k'\colon [0, \infty) \to [0, \infty)$ so that G_U is k'-hierarchically quasiconvex whenever $U \in \mathfrak{S}'$, and let A be the constant from Lemma 1.14 applied to k'-hierarchically quasiconvex subspaces. We further increase R so that $R > \max\{3E + A, E + E\nu + A + \nu, s_0\}$ and K so that

$$\begin{aligned} K = \max \Big\{ & 2\delta, 3R, K_0, \left\lceil \frac{4L + 4\delta + E}{T} \right\rceil + 2, 2L + 2\delta, \\ & \frac{6E + A + \sigma + 1}{T}, 3E + 2\sigma \Big\}. \end{aligned} \tag{3.3}$$

Let $a, b \in G$ be two infinite order elements, and suppose there exists $g \in G$ such that $ga = bg$. Then $g\,\mathrm{Big}(a) = \mathrm{Big}(b)$. Let C be the container associated to $\mathrm{Big}(b)$ in S, so that $\mathrm{Big}(b) \cup \{C\} = \{B_1, \ldots, B_k, C\}$ is a maximal collection of pairwise orthogonal domains. Since g conjugates a^i to b^i for any i, we first replace a and b by sufficiently high powers so that the following conditions are satisfied:

(a) $\mathrm{Big}(a)$ and $\mathrm{Big}(b)$ are fixed pointwise by a and b, respectively;

(b) b has the decomposition $b = b_1 \cdots b_k$ with respect to $\mathrm{Big}(b) = \{B_1, \ldots, B_k, C\}$, where $b_i = b_{B_i}$ is as in (3.2);

(c) $\tau_V(b) \geq L_S \delta$ for every $V \in \mathrm{Big}(b)$.

[1] We note that a related property to this is studied in forthcoming work of Montse Casals-Ruiz, Mark Hagen, and Ilya Kazachkov.

Such powers exist and can be chosen uniformly (that is, depending only on the hierarchy constants, and not on the choice of elements a and b) by the discussion after Definition 1.22 in the first case, the orthogonal decomposition property and Lemma 3.9 in the second case, and the assumed bound on translation length in the orthogonal decomposition property in the third case. Lemma 3.5 applied to $\mathrm{Big}(b)$ provides an element $t \in G$ such that $b_i = tb_i'$, where $b_i' \in G_{U_i'} \subseteq \mathbb{F}_{U_i'}$ for $U_i' \in \mathfrak{S}'$ and $U_i = tU_i'$ for all $i = 1, \ldots, k$.

For each $Z \in \mathrm{Big}(b)$, let γ_b^Z be a $(2, \ell)$-quasigeodesic axis of b in $\mathcal{C}Z$. Then $\gamma_a^{g^{-1}Z} := g^{-1}\gamma_b^Z$ is a $(2, \ell)$-quasigeodesic axis of a in $\mathcal{C}(g^{-1}Z)$. We now fix a quadrilateral of (λ, λ)-hierarchy paths $\mu(1, g)$, $\mu(1, b^K)$, $b^K\mu(1, g) = \mu(b^K, b^K g)$, and $g\mu(1, b^K) = \mu(g, gb^K) = \mu(g, a^K g)$ in G.

Step 1: *Changing the conjugator.* Our first step is to replace g by a (possibly) different conjugator whose length we are able to bound in G. We will do this by first premultiplying g by a power of $b_i \in G_{B_i}$ for each $B_i \in \mathrm{Big}(b)$. By the commutative property, any power of b_i commutes with b, and so this new element will still conjugate a to b. This is analogous to how we changed the conjugator in the proof of Theorem A, when $\mathrm{Big}(a) = \mathrm{Big}(b) = \{S\}$. In that situation, the orthogonal decomposition of b was simply $b = b_S$, and we premultiplied the conjugator by a power of b. In the current situation, we need to be a bit more careful because not only may b have more than one term in its orthogonal decomposition, but now $\mathrm{Big}(a)$ and $\mathrm{Big}(b)$ may be *different* collections of domains. Because of this, we will need to estimate distances in multiple hyperbolic spaces. Finally, we will alter g in the clean container C associated to $\mathrm{Big}(b)$.

Fix $Z \in \mathrm{Big}(b)$ and let $b_Z = \mathfrak{g}_{G_Z}(b)$. Since $K \geq K_0$, we may apply Lemma 1.24 to each of the axes γ_b^Z in $\mathcal{C}Z$ and $\gamma_a^{g^{-1}Z}$ in $\mathcal{C}(g^{-1}Z)$, and the points $1, a^K \in G$ and $1, b^K \in G$, respectively. This yields a point $z' \in \gamma_a^{g^{-1}Z}$ and a point $w' \in \gamma_b^Z$ such that $z' \in \mathcal{N}_L(\pi_{g^{-1}Z}(\mu(1, a^K)))$ and $w' \in \mathcal{N}_L(\pi_Z(\mu(1, b^K)))$, where these neighborhoods are taken in $\mathcal{C}g^{-1}Z$ and $\mathcal{C}Z$, respectively. Moreover, if x is a nearest point on $\gamma_a^{g^{-1}Z}$ to $\pi_{g^{-1}Z}(1)$ in $\mathcal{C}(g^{-1}Z)$ and y is a nearest point on γ_b^Z to $\pi_Z(1)$ in $\mathcal{C}Z$, then

$$\mathsf{d}_{g^{-1}Z}(x, z') \leq L \qquad \text{and} \qquad \mathsf{d}_Z(y, w') \leq L.$$

Let $z \in \pi_{g^{-1}Z}(\mu(1, a^K))$ and $w \in \pi_Z(\mu(1, b^K))$ be nearest points to z' and w', respectively, so that $\mathsf{d}_{g^{-1}Z}(z, z') \leq L$ and $\mathsf{d}_Z(w, w') \leq L$. See Figure 3.

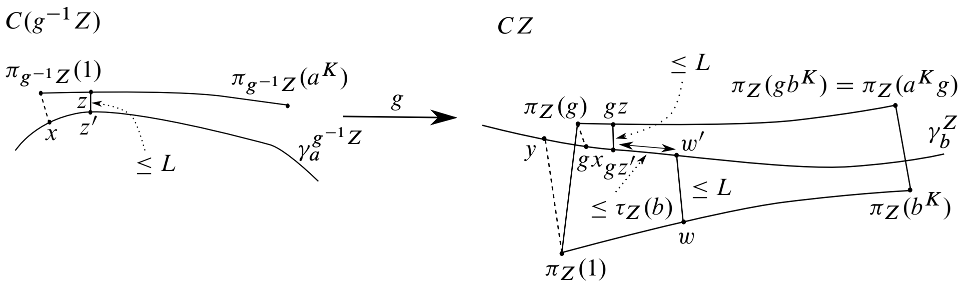

Figure 3. The geometry of the axes of a and b in $\mathcal{C}g^{-1}Z$ and $\mathcal{C}Z$, respectively.

Since the isometry g maps $\mathcal{C}(g^{-1}Z)$ to $\mathcal{C}Z$, we have $gz' \in g\gamma_a^{g^{-1}Z} = \gamma_b^Z$ and $gz \in g\pi_{g^{-1}Z}(\mu(1, a^K)) = \pi_Z(\mu(g, ga^K)) = \pi_Z(\mu(g, b^K g))$. Moreover,

$$\mathsf{d}_Z(gz, gz') = \mathsf{d}_{g^{-1}Z}(z, z') \le L \quad \text{and} \quad \mathsf{d}_Z(gz', gx) = \mathsf{d}_{g^{-1}Z}(z', x) \le L.$$

By possibly premultiplying g by a power of b_Z, we may assume that $\mathsf{d}_Z(gz', w') \le \tau_Z(b)$. Moreover, this new element also conjugates a to b, because b_Z commutes with b by the commutative property.

We perform the above procedure for each $Z \in \mathrm{Big}(b)$ and possibly premultiply g by a (possibly different) power m_Z of each b_Z.

We now alter g in the clean container C associated to $\mathrm{Big}(b)$. Let $t \in G$ be as in Lemma 3.5 applied to $\{B_1, \ldots, B_k, C\}$, so that $C = tC'$ for some $C' \in \mathfrak{S}'$. The label of the vertex $\mathfrak{g}_{tG_C}(g)$ is $tg_{C'}$, where $g_{C'} \in G_{C'} \subseteq \mathbb{F}_{C'}$. Let $g_C := tg_{C'}t^{-1} \in G_C$.

We claim that $g_C^{-1}g$ conjugates a to b and $\mathsf{d}_V(1, g_C^{-1}g) \le A$ for each $V \sqsubseteq C$. The commutative property and condition (b) ensure that g_C^{-1} commutes with b, hence $g_C^{-1}g$ conjugates a to b.

We have

$$\mathfrak{g}_{\mathbb{F}_C}(g_C^{-1}g) \asymp_A g_C^{-1}\mathfrak{g}_{g_C\mathbb{F}_C}(g) = tg_{C'}^{-1}t^{-1}\mathfrak{g}_{\mathbb{F}_C}(g),$$

where the first estimate follows from Lemma 1.14 and the second from the definition of g_C and the fact that $g_C\mathbb{F}_C = \mathbb{F}_C$.

By the $\mathbf{F}_U$ stabilizers property, $\mathsf{d}_G(\mathfrak{g}_{\mathbb{F}_C}(g), \mathfrak{g}_{G_C}(g)) \le \nu$. We also have

$$tg_{C'}^{-1}t^{-1}\mathfrak{g}_{G_C}(g) = tg_{C'}^{-1}t^{-1}(tg_{C'}) = t.$$

Thus

$$\begin{aligned}\mathsf{d}_G(t, \mathfrak{g}_{\mathbb{F}_C}(g_C^{-1}g)) &\le \mathsf{d}_G(t, tg_{C'}^{-1}t^{-1}\mathfrak{g}_{\mathbb{F}_C}(g)) + \mathsf{d}_G(tg_{C'}^{-1}t^{-1}\mathfrak{g}_{\mathbb{F}_C}(g), \mathfrak{g}_{\mathbb{F}_C}(g_C^{-1}g)) \\ &\le \nu + A. \end{aligned} \tag{3.4}$$

By [4, Remark 1.16] and Remark 1.19, we have $\pi_V(\mathfrak{g}_{\mathbb{F}_C}(g_C^{-1}g)) = \pi_V(g_C^{-1}g)$. Since the projection maps π are Lipschitz, it thus follows from (3.4) that $\mathsf{d}_V(t, g_C^{-1}g) \le A + \nu$ for all $V \sqsubseteq C$.

By Lemma 3.5, we have $\mathsf{d}_G(t, \mathfrak{g}_{\mathbf{P}_\mathcal{U}}(1)) \le E\nu$, where $\mathcal{U} = \{B_1, \ldots, B_k, C\}$. The only domains which are E-relevant for $1, \mathfrak{g}_{\mathbf{P}_\mathcal{U}}(1)$ are those which are transverse to some element of $\mathcal{U}$ or into which some element of $\mathcal{U}$ properly nests by Lemma 1.20. In particular, $\mathsf{d}_V(1, \mathfrak{g}_{\mathbf{P}_\mathcal{U}}(1)) \le E$ for all $V \sqsubseteq C$. By the triangle inequality and the fact that the maps π_U are Lipschitz, we have for all $V \sqsubseteq C$

$$\mathsf{d}_V(1, g_C^{-1}g) \le \mathsf{d}_V(1, \mathfrak{g}_{\mathbf{P}_\mathcal{U}}(1)) + \mathsf{d}_V(\mathfrak{g}_{\mathbf{P}_\mathcal{U}}(1), t) + \mathsf{d}_V(t, g_C^{-1}g) \le E + E\nu + A + \nu < R.$$

This yields a new element $(\prod_{Z \in \mathrm{Big}(b)} b_Z^{m_Z})g_C^{-1}g$, which also conjugates a to b. We have shown that this new conjugator, which by an abuse of notation we still call g, satisfies the following properties:

$$\mathsf{d}_Z(y, gx) \le \mathsf{d}_Z(y, w') + \mathsf{d}_Z(w', gz') + \mathsf{d}_Z(gz', gz) \le \tau_Z(b) + 2L \tag{3.5}$$

for each $Z \in \mathrm{Big}(b)$, and

$$\mathsf{d}_V(1, g) < R \tag{3.6}$$

whenever $V \sqsubseteq C$.

Step 2: *Bounding the length of* g. Our goal is to bound the length of g in G. As in the proof of Theorem A, we will show that for each $U \in \mathfrak{S}$, we have

$$\mathsf{d}_U(1, g) \leq 2K\mathsf{d}_U(1, b) + \mathsf{d}_U(g, b^K g) + K, \tag{3.7}$$

where K is as in (3.3). After establishing this bound for each $U \in \mathfrak{S}$, we then apply the distance formula with threshold R and the fact that $\mathsf{d}_G(g, b^K g) = \mathsf{d}_G(g, ga^K) = \mathsf{d}_G(1, a^K)$, which, as in the proof of Theorem A, establishes that

$$\mathsf{d}_G(1, g) \preceq_{C,D} 2K\mathsf{d}_G(1, b) + K\mathsf{d}_G(1, a) + K,$$

where C, D are the constants given by the distance formula (Theorem 1.7). (Note that by assumption, R is sufficiently large to serve as a threshold in the distance formula.) This will provide the desired bound in G.

Fix $U \in \mathfrak{S}$. If $U \notin \mathbf{Rel}(1, g; R)$, then we have $\mathsf{d}_U(1, g) \leq R \leq K$, and (3.7) holds. Thus we assume for the rest of the proof that $U \in \mathbf{Rel}(1, g; R)$. There are five cases to consider: there is some $Z \in \mathrm{Big}(b)$ such that $U = Z$; there is some $Z \in \mathrm{Big}(b)$ such that $U \sqsubsetneq Z$; there is some $Z \in \mathrm{Big}(b)$ such that $U \sqsupsetneq Z$; there is some $Z \in \mathrm{Big}(b)$ such that $U \pitchfork Z$; and $U \perp Z$ for all $Z \in \mathrm{Big}(b)$.

Cases 1 *and* 2. There is some $Z \in \mathrm{Big}(b)$ such that $U = Z$ or $U \sqsubsetneq Z$.

These two cases follow almost exactly as in the proof of Theorem A, the distinction being that Z plays the role of S and we measure distances in both $\mathcal{C}(g^{-1}Z)$ and $\mathcal{C}Z$. In case 2, one must also use (3.5) in place of (2.2).

Case 3. There is some $Z \in \mathrm{Big}(b)$ such that $U \sqsupsetneq Z$.

By our choice of K, we have $KT \geq E$, and thus $\mathsf{d}_Z(1, b^K) \geq E$. Applying the bounded geodesic image axiom (Definition 1.6 (7)) to $\pi_U(\mu(1, b^K))$ in $\mathcal{C}U$, we obtain $\rho^Z_U \subseteq \mathcal{N}_{E+\sigma}(\pi_U(\mu(1, b^K)))$ in $\mathcal{C}U$, and hence

$$\mathsf{d}_U(1, \rho^Z_U) \leq \mathsf{d}_U(1, b^K) + E + \sigma. \tag{3.8}$$

Additionally, $g^{-1}Z \in \mathrm{Big}(a)$ and $g^{-1}U \sqsupsetneq g^{-1}Z$. The choice of K ensures that

$$\mathsf{d}_{g^{-1}Z}(1, a^K) \geq E,$$

so applying the bounded geodesic image axiom to $\pi_U(\mu(1, a^K))$ in $\mathcal{C}(g^{-1}U)$ yields

$$\rho^{g^{-1}Z}_{g^{-1}U} \subseteq \mathcal{N}_{E+\sigma}(\pi_U(\mu(1, a^K)))$$

in $\mathcal{C}(g^{-1}U)$. Applying the isometry g, we obtain

$$g\rho^{g^{-1}Z}_{g^{-1}U} \subseteq \mathcal{N}_{E+\sigma}(g\pi_{g^{-1}U}(\mu(1, a^K))) = \mathcal{N}_{E+\sigma}(\pi_U(\mu(g, b^K g)))$$

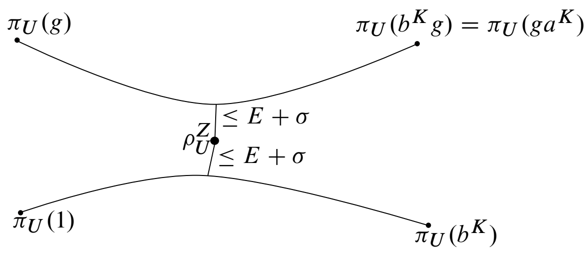

Figure 4. Case 3.

in $\mathcal{C}U$. See Figure 4. Moreover, projection maps in a hierarchically hyperbolic group are G-equivariant, and so $g\rho^{g^{-1}Z}_{g^{-1}U} = \rho^Z_U$. Thus

$$\mathsf{d}_U(g, \rho^Z_U) = \mathsf{d}_U(g, g\rho^{g^{-1}Z}_{g^{-1}U}) \leq \mathsf{d}_U(g, b^K g) + E + \sigma. \tag{3.9}$$

Therefore, by the triangle inequality, (3.8), and (3.9), we have

$$\begin{aligned}\mathsf{d}_U(1, g) &\leq \mathsf{d}_U(1, \rho^Z_U) + \mathsf{diam}_{\mathcal{C}U}(\rho^Z_U) + \mathsf{d}(\rho^Z_U, g)\\ &\leq \mathsf{d}_U(1, b^K) + 3E + 2\sigma + \mathsf{d}_U(g, b^K g)\\ &\leq K\mathsf{d}_U(1, b) + \mathsf{d}_U(g, b^K g) + K,\end{aligned}$$

where the final inequality follows because $K \geq 3E + 2\sigma$.

Case 4. There is some $Z \in \mathrm{Big}(b)$ such that $U \pitchfork Z$.

Consider the product region $\mathbf{P}_Z$, and let $\xi = \mathfrak{g}_{\mathbf{P}_Z}(g)$ and $\nu = \mathfrak{g}_{\mathbf{P}_Z}(g)$. See Figure 5.

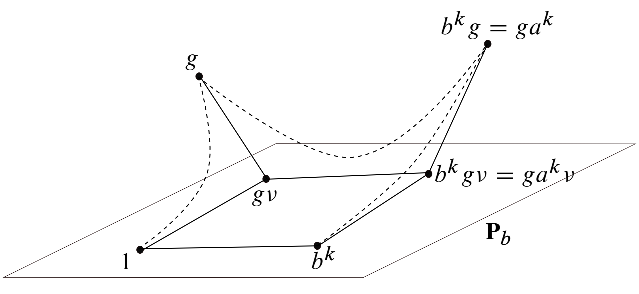

Figure 5. Conjugate elements a and b, with conjugator g, in G. Solid segments are hierarchy paths, while dotted segments are geodesics.

Since we are assuming that U is relevant for 1, g and $U \pitchfork Z$, it follows from Lemma 1.20 that $U \in \mathbf{Rel}(1, \xi; R) \cup \mathbf{Rel}(g, \nu; R)$. As b is loxodromic with respect to the action on $\mathcal{C}Z$ for all $Z \in \mathrm{Big}(b)$, we have $\mathsf{d}_Z(\nu, b^K\nu) \geq K\tau_Z(b) \geq KT \geq R$. Thus $Z \in \mathbf{Rel}(\nu, b^K\nu; R)$. Similarly, $Z \in \mathbf{Rel}(\xi, b^K\xi; R)$. Note that this implies $Z \in \mathbf{Rel}(1, b^K; R) \cap \mathbf{Rel}(1, b^K\xi; R) \cap \mathbf{Rel}(g, b^K g; R) \cap \mathbf{Rel}(g, b^K\nu; R)$, as well.

Claim 3.11. *If U is R-relevant for g, $g\nu$, then U is not R-relevant for $b^K\nu$, $b^K g$. If U is R-relevant for 1, ξ, then U is not R-relevant for $b^K\xi$, b^K.*

Proof. We will prove the first statement. The proof of the second statement is completely analogous.

Since U is relevant for g, ν, the domain $b^K U \in \mathfrak{S}$ is relevant for $b^K\nu$, $b^K g$. Moreover, since b fixes $\mathrm{Big}(b)$ pointwise, we have $b^K Z = Z$. As $U \pitchfork Z$, we must also have $b^K U \pitchfork Z$. From this, we apply the G-equivariance of the projections maps in a hierarchically hyperbolic group to conclude that $b^K \rho_Z^U = \rho_Z^{b^K U}$ in $\mathcal{C}Z$. Thus we have

$$\mathsf{d}_Z(\rho_Z^U, \rho_Z^{b^K U}) = \mathsf{d}_Z(\rho_Z^U, b^K \rho_Z^U) \succeq_E K\tau_Z(b) \geq KT.$$

Since $Z \in \mathrm{Big}(b)$, Lemma 1.20 implies that Z is not s-relevant for g, ν for any $s \geq s_0$. In particular, since $E \geq s_0$, the distance between $\pi_Z(g)$ and $\pi_Z(\nu)$ in $\mathcal{C}Z$ is bounded by E. On the other hand, since U is R-relevant for g and ν, it follows from [3, Proposition 5.17] that any hierarchy path $\mu(g,\nu)$ in G has a subpath which is contained in the E-neighborhood of $\mathbf{P}_U$. Since the projection maps π are Lipschitz, we have

$$\mathsf{d}_Z(\pi_Z(\mathbf{P}_U), \pi_Z(\mu(g,\nu))) \leq E. \tag{3.10}$$

Recall that $\rho_Z^U \asymp_E \pi_Z(\mathbf{P}_U)$ (see comments after Definition 1.15). Thus (3.10) implies

$$\mathsf{d}_Z(\rho_Z^U, \pi_Z(\mu(g,\nu))) \leq 2E.$$

Since $\pi_Z(\mu(g,\nu))$ is an unparametrized (λ,λ)-quasigeodesic, it is contained in the σ-neighborhood of a geodesic in $\mathcal{C}Z$ from $\pi_Z(g)$ to $\pi_Z(\nu)$. By the above discussion, such a geodesic necessarily has length at most E. Therefore,

$$\mathsf{d}_Z(\rho_Z^U, g) \leq 3E + \sigma.$$

By the triangle inequality, we have

$$\begin{aligned}\mathsf{d}_Z(\rho_Z^U, b^K g) &\geq \mathsf{d}_Z(g, b^K g) - \mathsf{d}_Z(\rho_Z^U, g) \geq K\tau_Z(b) - \mathsf{d}_Z(\rho_Z^U, g)\\ &\geq KT - (3E+\sigma) > 3E + A,\end{aligned}$$

where the final inequality follows from our choice of $K \geq \frac{6E+A+\sigma+1}{T}$. See Figure 6. Therefore, $\mathsf{d}_Z(\rho_Z^U, b^K g)$ is large enough to apply the consistency inequalities (Definition 1.6 (4)), yielding

$$\mathsf{d}_U(\rho_U^Z, b^K g) \leq E. \tag{3.11}$$

The same argument bounding the distance in $\mathcal{C}Z$ between g and ν applies to show that Z is not E-relevant for $b^k g$, $\mathfrak{g}_{\mathbf{P}_Z}(b^K g)$. By Lemma 1.14, we have $b^K\nu \asymp_A \mathfrak{g}_{\mathbf{P}_Z}(b^K g)$. Therefore, Z is not $(E+A)$-relevant for ν, $b^k\nu$, and so $\pi_Z(b^K g) \asymp_{E+A} \pi_Z(b^K\nu)$ in $\mathcal{C}Z$. It follows that

$$\mathsf{d}_Z(\rho_Z^U, b^K\nu) \asymp_{2E+A} \mathsf{d}_Z(\rho_Z^U, b^K g) > 3E + A,$$

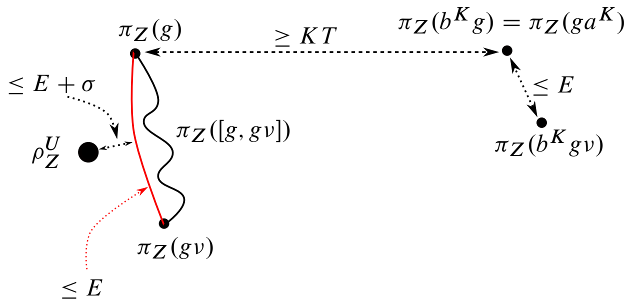

Figure 6. The arrangement of points in $\mathcal{C}Z$ in the proof of Claim 3.11 in case 4.

from which we conclude $\mathsf{d}_Z(\rho_Z^U, b^K\nu) > E$. Thus we may again apply the consistency inequalities, yielding

$$\mathsf{d}_U(\rho_U^Z, b^K\nu) \leq E.$$

Combining this with (3.11) and applying the triangle inequality yields

$$\mathsf{d}_U(b^K g, b^K \nu) \leq \mathsf{d}_U(b^K g, \rho_U^Z) + \mathsf{diam}_{\mathcal{C}U}(\rho_U^Z) + \mathsf{d}_U(\rho_U^Z, b^K\nu) \leq E + E + E = 3E.$$

Since $R > 3E$, we have $U \notin \mathbf{Rel}(b^K, b^K\nu; R)$. This completes the proof of the claim. ∎

Suppose first that $U \in \mathbf{Rel}(1, \xi; R) \cap \mathbf{Rel}(g, \nu; R)$. Then by the claim, we have that $U \notin \mathbf{Rel}(b^K, b^K\xi; R) \cup \mathbf{Rel}(b^K g, b^K\nu; R)$. By Lemma 1.20 and the fact that $U \pitchfork Z$, this is equivalent to $U \notin \mathbf{Rel}(b^K, b^K g; R)$. Thus $\mathsf{d}_U(b^K, b^K g) < R$, and so we have,

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

Now suppose that $U \notin \mathbf{Rel}(1, \xi; R) \cap \mathbf{Rel}(g, \nu; R)$. Since, as previously noted, it follows from Lemma 1.20 that $U \in \mathbf{Rel}(1, \xi; R) \cup \mathbf{Rel}(g, \nu; R)$, we must have either $U \in \mathbf{Rel}(1, \xi; R)$ or $U \in \mathbf{Rel}(g, \nu; R)$. Suppose without loss of generality that $U \in \mathbf{Rel}(1, \xi; R)$ but $U \notin \mathbf{Rel}(g, \nu; R)$. It follows from the claim that $U \notin \mathbf{Rel}(b^K, b^K\xi; R)$. Moreover, by Lemma 1.20, we have $\mathsf{d}_U(b^K\xi, \nu) \leq R$. Therefore,

$$\begin{aligned}
\mathsf{d}_U(1, g) &\leq \mathsf{d}_U(1, b^K) + \mathsf{d}_U(b^K, b^K\xi) + \mathsf{d}_U(b^K\xi, \nu) + \mathsf{d}_U(\nu, g) \\
&\leq K\mathsf{d}_U(1, b) + R + R + R \\
&\leq K\mathsf{d}_U(1, b) + K,
\end{aligned}$$

where the final inequality follows because $K \geq 3R$. Thus (3.7) holds regardless of whether $U \in \mathbf{Rel}(1, \xi; R) \cap \mathbf{Rel}(g, \nu; R)$.

Case 5. $U \perp Z$ for all $Z \in \mathrm{Big}(b)$.

Note that $U \perp Z$ for all $Z \in \mathrm{Big}(b)$ if and only if $U \sqsubseteq C$, where C is the container associated to $\mathrm{Big}(b) = \{B_1, \ldots, B_k\}$. Thus the bound $\mathsf{d}_U(1, g) \leq R$ follows immediately from (3.6).

This completes the proof of Theorem D. ∎

Remark 3.12. Theorem D establishes the linear conjugator property for suitable *powers* of pairs of conjugate infinite order elements. In particular, the conjugator whose length we bound in these theorems may not conjugate a to b. There are two additional steps necessary to extend the ideas in these proofs to show the linear conjugator property holds for *all* pairs of conjugate infinite order elements. First, one would have to deal with the fact that an element may permute the elements in its big set, an issue we avoid by passing to a power to assume that the big set is fixed elementwise. This is likely not a serious problem. Second, one would need to understand the conjugator length function for finite order elements. Recall that in the decomposition of b in the proof of Theorem D, the factor b_C corresponding to the container associated to the big set of b was a finite order element of the corresponding sub-hierarchically hyperbolic group $(G_C, \mathfrak{S}_C)$. We passed to a power so that we could assume this factor was trivial. If we do not pass to a power, we need a different way to modify the conjugator in that sub-hierarchically hyperbolic group G_C. To do this, we need to understand conjugators of finite order elements. The conjugator length function for finite order elements of hierarchically hyperbolic groups is unknown, hence this second step is currently out of reach.

Acknowledgments. We thank Mark Hagen for many interesting discussions about hierarchically hyperbolic spaces. We thank Jacob Russell for feedback on an early draft, and the anonymous referees for useful comments.

Funding. C. Abbott was supported by NSF grants DMS-1803368 and DMS-2106906. J. Behrstock was supported by NSF grant DMS-1710890.

References

[1] C. Abbott, J. Behrstock, and M. G. Durham, Largest acylindrical actions and stability in hierarchically hyperbolic groups. *Trans. Amer. Math. Soc. Ser. B* **8** (2021), 66–104 Zbl 1498.20109 MR 4215647

[2] J. Behrstock and C. Druţu, Divergence, thick groups, and short conjugators. *Illinois J. Math.* **58** (2014), no. 4, 939–980 Zbl 1353.20024 MR 3421592

[3] J. Behrstock, M. Hagen, and A. Sisto, Hierarchically hyperbolic spaces II: Combination theorems and the distance formula. *Pacific J. Math.* **299** (2019), no. 2, 257–338 Zbl 07062864 MR 3956144

[4] J. Behrstock, M. F. Hagen, and A. Sisto, Asymptotic dimension and small-cancellation for hierarchically hyperbolic spaces and groups. *Proc. Lond. Math. Soc. (3)* **114** (2017), no. 5, 890–926 Zbl 1431.20028 MR 3653249

[5] J. Behrstock, M. F. Hagen, and A. Sisto, Hierarchically hyperbolic spaces, I: Curve complexes for cubical groups. *Geom. Topol.* **21** (2017), no. 3, 1731–1804 Zbl 1439.20043 MR 3650081

[6] J. Behrstock, M. F. Hagen, and A. Sisto, Quasiflats in hierarchically hyperbolic spaces. *Duke Math. J.* **170** (2021), no. 5, 909–996 Zbl 07369844 MR 4255047

[7] F. Berlai and B. Robbio, A refined combination theorem for hierarchically hyperbolic groups. *Groups Geom. Dyn.* **14** (2020), no. 4, 1127–1203 Zbl 07362555 MR 4186470

[8] D. Berlyne and J. Russell, Hierarchical hyperbolicity of graph products. *Groups Geom. Dyn.* **16** (2022), no. 2, 523–580 Zbl 07624152 MR 4502615

[9] J. S. Birman, A. Lubotzky, and J. McCarthy, Abelian and solvable subgroups of the mapping class groups. *Duke Math. J.* **50** (1983), no. 4, 1107–1120 Zbl 0551.57004 MR 726319

[10] B. H. Bowditch, Tight geodesics in the curve complex. *Invent. Math.* **171** (2008), no. 2, 281–300 Zbl 1185.57011 MR 2367021

[11] M. R. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*. Grundlehren Math. Wiss. 319, Springer, Berlin, 1999 Zbl 0988.53001 MR 1744486

[12] I. Bumagin, Time complexity of the conjugacy problem in relatively hyperbolic groups. *Internat. J. Algebra Comput.* **25** (2015), no. 5, 689–723 Zbl 1327.20037 MR 3384078

[13] M. Burger and S. Mozes, Finitely presented simple groups and products of trees. *C. R. Acad. Sci. Paris Sér. I Math.* **324** (1997), no. 7, 747–752 Zbl 0966.20013 MR 1446574

[14] M. Burger and S. Mozes, Lattices in product of trees. *Inst. Hautes Études Sci. Publ. Math.* **92** (2000), 151–194 Zbl 1007.22013 MR 1839489

[15] P.-E. Caprace, Finite and infinite quotients of discrete and indiscrete groups. In *Groups St Andrews 2017 in Birmingham*, pp. 16–69, London Math. Soc. Lecture Note Ser. 455, Cambridge University Press, Cambridge, 2019 MR 3931408

[16] R. B. Coulon, Partial periodic quotients of groups acting on a hyperbolic space. *Ann. Inst. Fourier (Grenoble)* **66** (2016), no. 5, 1773–1857 Zbl 1397.20054 MR 3533270

[17] J. Crisp, E. Godelle, and B. Wiest, The conjugacy problem in subgroups of right-angled Artin groups. *J. Topol.* **2** (2009), no. 3, 442–460 Zbl 1181.20030 MR 2546582

[18] M. G. Durham, M. F. Hagen, and A. Sisto, Boundaries and automorphisms of hierarchically hyperbolic spaces. *Geom. Topol.* **21** (2017), no. 6, 3659–3758 Zbl 1439.20044 MR 3693574

[19] M. G. Durham, M. F. Hagen, and A. Sisto, Correction to the article Boundaries and automorphisms of hierarchically hyperbolic spaces. *Geom. Topol.* **24** (2020), no. 2, 1051–1073 Zbl 1443.20071 MR 4153656

[20] T. Haettel, N. Hoda, and H. Petyt, Coarse injectivity, hierarchical hyperbolicity, and semihyperbolicity. 2022, arXiv:2009.14053, To appear in *Geom. Topol.*

[21] M. F. Hagen and T. Susse, On hierarchical hyperbolicity of cubical groups. *Israel J. Math.* **236** (2020), no. 1, 45–89 Zbl 1453.20056 MR 4093881

[22] F. Haglund and D. T. Wise, Special cube complexes. *Geom. Funct. Anal.* **17** (2008), no. 5, 1551–1620 Zbl 1155.53025 MR 2377497

[23] D. Janzen and D. T. Wise, A smallest irreducible lattice in the product of trees. *Algebr. Geom. Topol.* **9** (2009), no. 4, 2191–2201 Zbl 1220.20039 MR 2558308

[24] I. G. Lysënok, On some algorithmic properties of hyperbolic groups. *Math. USSR Izv.* **35** (1990), no. 1, 145–163 Zbl 0697.20020

[25] H. A. Masur and Y. N. Minsky, Geometry of the complex of curves. II. Hierarchical structure. *Geom. Funct. Anal.* **10** (2000), no. 4, 902–974 Zbl 0972.32011 MR 1791145

[26] D. Rattaggi, A finitely presented torsion-free simple group. *J. Group Theory* **10** (2007), no. 3, 363–371 Zbl 1136.20026 MR 2320973

[27] B. Robbio and S. Spriano, Hierarchical hyperbolicity of hyperbolic-2-decomposable groups. 2020, arXiv:2007.13383

[28] J. Russell, From hierarchical to relative hyperbolicity. *Int. Math. Res. Not. IMRN* **2022** (2022), no. 1, 575–624 Zbl 1483.30085 MR 4366027

[29] J. Russell, D. Spriano, and H. C. Tran, Convexity in hierarchically hyperbolic spaces. 2021, arXiv:1809.09303, To appear in *Algebr. Geom. Topol.*

[30] A. Sale, Bounded conjugators for real hyperbolic and unipotent elements in semisimple Lie groups. *J. Lie Theory* **24** (2014), no. 1, 259–305 Zbl 1311.22017 MR 3186337

[31] A. Sale, Geometry of the conjugacy problem in lamplighter groups. In *Algebra and computer science*, pp. 171–183, Contemp. Math. 677, American Mathematical Society, Providence, RI, 2016 Zbl 1388.20050 MR 3589810

[32] H. Servatius, Automorphisms of graph groups. *J. Algebra* **126** (1989), no. 1, 34–60 Zbl 0682.20022 MR 1023285

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

[34] J. Tao, Linearly bounded conjugator property for mapping class groups. *Geom. Funct. Anal.* **23** (2013), no. 1, 415–466 Zbl 1286.57020 MR 3037904

[35] D. T. Wise, Complete square complexes. *Comment. Math. Helv.* **82** (2007), no. 4, 683–724 Zbl 1142.20025 MR 2341837

Received 17 February 2020.

Carolyn Abbott
Department of Mathematics, Brandeis University, 415 South St, Waltham, MA 02453, USA; carolynabbott@brandeis.edu

Jason Behrstock
Lehman College and The Graduate Center, City University of New York, 365 5th Ave, New York, NY 10016, USA; jason.behrstock@lehman.cuny.edu