CONVEX COCOMPACTNESS IN MAPPING CLASS GROUPS VIA QUASICONVEXITY IN RIGHT-ANGLED ARTIN GROUPS

JOHANNA MANGAHAS AND SAMUEL J. TAYLOR

ABSTRACT. We characterize convex cocompact subgroups of mapping class groups that arise as subgroups of specially embedded right-angled Artin groups. That is, if the right-angled Artin group $G < \mathrm{Mod}(S)$ satisfies certain conditions that imply G is quasi-isometrically embedded in $\mathrm{Mod}(S)$, then a purely pseudo-Anosov subgroup $H < G$ is convex cocompact in $\mathrm{Mod}(S)$ if and only if it is combinatorially quasiconvex in G. We use this criterion to construct convex cocompact subgroups of $\mathrm{Mod}(S)$ whose orbit maps into the curve complex have small Lipschitz constants.

1. INTRODUCTION

Convex cocompact subgroups of mapping class groups are those finitely generated subgroups whose orbits in Teichmüller space are quasiconvex [FM02], or equivalently, whose orbit map into the curve complex defines a quasi-isometric embedding of the group [Ham05, KL08]. Such subgroups of mapping class groups are of interest because of their close connection to surface group extensions. Letting S denote the closed surface of genus $g \geq 2$ and $\mathring{S}$ denote S punctured at $p \in S$, the well-known Birman exact sequence

$$1 \longrightarrow \pi_1(S, p) \longrightarrow \mathrm{Mod}(\mathring{S}) \overset{f}{\longrightarrow} \mathrm{Mod}(S) \to 1$$

gives rise to an extension E_G of $\pi_1(S, p)$ for each subgroup $G < \mathrm{Mod}(S)$, obtained by setting $E_G = f^{-1}(G)$. Theorems in [FM02] and [Ham05] combine to say that E_G is Gromov-hyperbolic if and only if G is convex cocompact. Hence, to understand the prevalence and properties of hyperbolic surface group extensions we are left to study convex cocompact subgroups of $\mathrm{Mod}(S)$.

The results of this paper determine the conditions of convex cocompactness for mapping class subgroups contained in certain *admissible* embedded right-angled Artin groups, including the groups constructed in [CLM12]. We denote by $A(\Gamma)$ the right-angled Artin group associated to the graph Γ, and say a subgroup $H < A(\Gamma)$ is *quasiconvex in $A(\Gamma)$* if it is quasiconvex in the word metric using the generating set corresponding to vertices of Γ.

Theorem 1.1. *Suppose $A(\Gamma) < \mathrm{Mod}(S)$ is admissible. Then $H < A(\Gamma)$ is convex cocompact if and only if it is quasiconvex in $A(\Gamma)$ and all nontrivial elements of H are pseudo-Anosov.*

Suppose $H < \mathrm{Mod}(S)$ has a finite $K(H, 1)$. As observed in [KL07], if H is purely pseudo-Anosov but is *not* convex cocompact, then the extension E_H has a finite $K(E_H, 1)$ and no Baumslag-Solitar subgroups, but is *not* hyperbolic. A well-known conjecture attributed to Gromov asserts that such a group does not exist (see [KL07] and [FM02]). By Theorem 1.1 an affirmative answer to the following question would produce a counter-example to this conjecture.

Question 1.2. Does there exist an admissible embedding $A(\Gamma) \hookrightarrow \mathrm{Mod}(S)$ for which some non-quasiconvex $H < A(\Gamma)$ is all-pseudo-Anosov?

A constructive and stronger result lies behind the condition for convex cocompactness in Theorem 1.1. We are able to identify when $H < A(\Gamma)$ is purely pseudo-Anosov by checking only finitely many mapping classes. Below, $\ell(\Gamma, K)$ is a positive integer-valued function depending only on Γ and $K \geq 0$ (see Section 3.3 for its definition), and $A(\Gamma)$-*length* is word length in $A(\Gamma)$ with respect to the generating set corresponding to vertices of Γ.

Theorem 1.3. *Suppose $A(\Gamma) < \mathrm{Mod}(S)$ is admissible and H is K-quasiconvex in $A(\Gamma)$. If each nonidentity $h \in H$ whose $A(\Gamma)$-length is at most $\ell(\Gamma, K)$ is pseudo-Anosov, then H is convex cocompact.*

Theorem 1.3 provides flexible means to build explicit examples of convex cocompact subgroups of mapping class groups, distinguishing it from other constructions in the literature. See Section 7 for some examples.

We briefly survey other methods of producing convex cocompact subgroups of mapping class groups. The simplest are free groups generated by sufficiently high powers of any finite family of independent pseudo-Anosov mapping classes [FM02]. Later, Min [Min08] created virtually free examples isomorphic to $G*H$ for arbitrary finite subgroups $G, H < \mathrm{Mod}(S)$, by conjugating one of these groups by a sufficiently high power of a pseudo-Anosov. A third set of examples live in certain hyperbolic groups embedded in the mapping class group $\mathrm{Mod}(\mathring{S})$ of a once-punctured surface $\mathring{S}$; these are the surface group extensions E_G described above. Generalizing [KLS09], Dowdall, Kent, and Leininger prove that, when E_G is hyperbolic, its quasiconvex all-pseudo-Anosov subgroups are convex cocompact in $\mathrm{Mod}(\mathring{S})$ [DKRL12]. The convex cocompact subgroups considered in this paper are most similar in spirit to these last examples, with E_G replaced by $A(\Gamma)$ and cubical CAT(0) geometry playing a role similar to hyperbolicity. Here, the idea is to replace quasiconvex orbits in Teichmüller space with combinatorially quasiconvex orbits in $\widetilde{S}_\Gamma$, the CAT(0) cube complex associated to $A(\Gamma)$.

To highlight a difference between the all-pseudo-Anosov free groups constructed by our method and earlier examples, we describe a family of nonelementary convex cocompact subgroups of $\mathrm{Mod}(S_g)$ whose orbit maps into the curve complex have Lipschitz constants on the order of $1/g$. Let $\mathcal{C}(S)$ denote the curve complex of S and $\ell_S(f)$ the stable translation length of $f \in \mathrm{Mod}(S)$ in $\mathcal{C}(S)$ (see Section 2 for definitions).

Theorem 1.4. *Let S_g be a surface of genus g for some $g \geq 3$. Then for any $N \geq 1$ there exists a convex cocompact $H = \langle w_1, \ldots w_N \rangle < \mathrm{Mod}(S_g)$ with the following property: there is*

an $\alpha \in \mathcal{C}(S_g)$ so that for any $h \in H$,

$$d_{\mathcal{C}(S_g)}(\alpha, h\alpha) \leq |h|_H \cdot \frac{4}{g-1} + 2,$$

where $|.|_H$ denotes word length in H with the given generators. In particular, $\ell_S(w_i) \leq \frac{4}{g-1}$ for each $i = 1, \ldots, N$.

The pseudo-Anosovs w_i appearing in Theorem 1.4 are themselves gth powers. Since it is known that the minimum stable translation length in the curve complex is of order $1/g^2$ [GT11], it would be interesting to construct rank ≥ 2 convex cocompact subgroups whose orbit maps into the curve complex have Lipschitz constants on the order of $1/g^2$.

This paper is organized as follows: Section 2 contains background on surfaces and mapping class groups. Section 3 briefly reviews some cubical geometry and proves the main technical property of quasiconvex subgroups of right-angled Artin groups used in this paper. The definition of admissibility is then given in Section 4, where some properties of admissible homomorphisms from right-angled Artin groups into mapping class groups are established. The proof of Theorem 1.3 is given in Section 5. Section 6 completes the proof of Theorem 1.1 and Section 7 contains explicit constructions of convex cocompact mapping class subgroups. The paper concludes with Section 8, which proves Theorem 1.4.

Acknowledgments. We thank Centre de Recerca Matemàtica and the Polish Academy of Sciences for hosting both authors during parts of this research.

2. Surfaces and mapping classes

2.1. **Quasiconvexity and quasi-isometry.** Let (X, d_X) and (Y, d_Y) be metric spaces. For constants $K \geq 1$ and $L \geq 0$, a map $f : X \to Y$ is a (K, L)-*quasi-isometric embedding* if for all $x_1, x_2 \in X$

$$\frac{d_X(x_1, x_2) - L}{K} \leq d_Y(f(x_1), f(x_2)) \leq K(d_X(x_1, x_2) + L).$$

In addition, if every point in Y is within a bounded distance from the image $f(X)$, then f is a *quasi-isometry* and X and Y are said to be *quasi-isometric*. Where I is a subinterval of $\mathbb{R}$ or $\mathbb{Z}$, we call a (K, L)-quasi-isometric embedding $f : I \to Y$ a (K, L)-*quasi-geodesic*. If $K = 1$ and $L = 0$, then $f : I \to Y$ is a *geodesic*. We say Y is a *geodesic metric space* if for all $y_1, y_2 \in Y$, there is a geodesic $f : [a, b] \to Y$ with $f(a) = y_1$ and $f(b) = y_2$. For example, if G is a graph, i.e. a 1-dimensional CW complex, then giving unit length to each of its edges turns G into a geodesic metric space using the induced path metric.

For convenience, we write K-quasigeodesic or K-quasi-isometric embedding to mean a (K, K)-quasigeodesic or (K, K)-quasi-isometric embedding respectively.

We say a subset X' of a geodesic metric space X is K-*quasiconvex* if for any $x, y \in X'$ and any geodesic $[x, y]$ between x, y in X,

$$[x, y] \subset N_K(X').$$

We say X' is *quasiconvex* if it is K-quasiconvex for some K. When we speak of a quasiconvex subgroup H of a group G, we have fixed a word metric on G with respect to some finite generating set (changing generating sets can change which subgroups are quasiconvex). For CAT(0) cube complexes we additionally define *combinatorial quasiconvexity* in Section 3.2.

2.2. **Surface topology basics.** Here we recall some relevant information about surfaces. For additional details, we refer the reader to [MM00, FM12]. Fix a surface S of genus g with p punctures and b boundary components. The *complexity* of S is the quantity $\xi(S) = 3g - 3 + p + b$. In this paper we will only consider surfaces with $\xi(S) \geq 1$ and we will often not distinguish between boundary components and punctures of S. By a *subsurface X* of S, we mean a compact submanifold such that the homomorphism $\pi_1 X \to \pi_1 S$ induced by inclusion is injective. Hence, all subsurfaces are assumed to be *essential*. The subsurface $X \subset S$ is *nonannular* if X is not homeomorphic to an annulus. Denote by $\Omega(S)$ the set of isotopy classes of nonannular subsurfaces of S. Although annuli play an important part in the analysis of [MM00] and subsequent work, they are not considered in this paper. A *curve* γ in S is an essential, simple loop in S, i.e. the image of a π_1-injective embedding of the circle into S. Recall that γ is *essential* if it is homotopically nontrivial and not parallel to a boundary component or puncture. As is standard in the subject, we often do not distinguish between a curve and its isotopy class or a subsurface and its isotopy class.

The *curve complex $\mathcal{C}(S)$* of the surface S is the graph with vertex set the collection of isotopy classes of curves in S. Vertices v and w are joined by an edge in $\mathcal{C}(S)$ if v and w have disjoint representatives in S. When S has complexity 1, that is when S is a once-punctured torus or a four-times punctured sphere, this definition produces a graph without edges and so is modified as follows: the vertices of $\mathcal{C}(S)$ are unchanged, but two vertices are joined by an edge if the corresponding curves have the minimal number of intersections among pairs of curves on S. Hence, curves on the once-punctured surface intersecting once are joined by an edge, as are curves intersecting twice on the four-times punctured sphere. With this definition, if S has complexity 1 then $\mathcal{C}(S)$ is the standard Farey graph. In general, we consider $\mathcal{C}(S)$ with its standard graph metric where each edge is assigned unit length. In Section 2.3 we describe the action of $\mathrm{Mod}(S)$ on $\mathcal{C}(S)$ by isometry. The foundational result in the study of the coarse geometry of the mapping class group is the following:

Theorem 2.1. [MM99] *For S with $\xi(S) \geq 1$, $\mathcal{C}(S)$ is Gromov hyperbolic.*

Recall that a geodesic metric space is Gromov hyperbolic if there exists a $\delta \geq 0$ so that for any points x, y, z and geodesics $[x, y], [y, z]$ and $[z, x]$ between the three points,

$$[x, y] \subset N_\delta([y, z] \cup [z, x])$$

where N_δ denotes a δ-neighborhood in $\mathcal{C}(S)$.

A *pants decomposition P* of S is a maximal collection of pairwise connected vertices of $\mathcal{C}(S)$, or in terms of the surface, a maximal collection of isotopy classes of pairwise disjoint curves in S. A *marking μ* of S is a pants decomposition $P = \{\gamma_1, \ldots, \gamma_{\xi(S)}\}$ with the following additional structure: for each $\gamma_i \in P$ there is a corresponding curve β_i contained in

$S \setminus (P \setminus \gamma_i)$ that intersects γ_i the minimal number of times possible. In other words, if X is the complexity 1 component of $S \setminus (P \setminus \gamma_i)$ then β_i is any adjacent curve to γ_i in $\mathcal{C}(X)$. Then $\mu = \{(\gamma_1, \beta_1), \ldots, (\gamma_{\xi(S)}, \beta_{\xi(S)})\}$ is a marking of S. The underlying pants decomposition of the marking μ is called the *base* of μ and is denoted base(μ). In the terminology of [MM00], what we have described is a complete clean marking of S and in this reference the authors construct the marking complex $\mathcal{M}(S)$ whose vertices are markings of S and edges are determined by certain elementary moves on markings. The details, which may be found in [MM00], will not be reviewed here; however, it suffices for us to recall that $\mathcal{M}(S)$ is a locally finite, connected graph and the natural action of Mod(S) is proper and cocompact. Hence, by the Švarc-Milnor lemma, $\mathcal{M}(S)$ is quasi-isometric to Mod(S).

For a curve γ and a nonannular subsurface X of S we define the *subsurface projection of γ to X*, denoted $\pi_X(\gamma)$, as follows: first realize γ and ∂X with minimal intersection, for example by taking geodesic representatives in some hyperbolic metric on S. If γ does not intersect ∂X then either $\gamma \subset X$, in which case we set $\pi_X(\gamma) = \{\gamma\}$, or γ does not intersect X and $\pi_X(\gamma)$ is the empty set. Otherwise, $\gamma \cap X = \{\gamma_1, \ldots, \gamma_k\}$ is an nonempty collection of essential arcs in X and $\pi_X(\gamma)$ is the subset of $\mathcal{C}(X)$ whose elements are isotopic to the boundary of a regular neighborhood of the union of ∂X and γ_i for some $1 \leq i \leq k$. In other words, the curves of $\pi_X(\gamma)$ arise from performing surgery along ∂X to the arcs of $\gamma \cap \partial X$. This gives a subset of $\mathcal{C}(X)$ with $\text{diam}_X(\pi_X(\gamma)) \leq 2$, where diam_X denotes the diameter of a collection of vertices in $\mathcal{C}(X)$. Observe that if $\pi_X(\gamma) = \emptyset$ then $\gamma \in N_1(\partial X)$. When $\pi_X(\gamma) \neq \emptyset$, then we say that γ *cuts* X; otherwise γ *misses* X.

The following result of Masur-Minsky gives control over the projection of a geodesic in the curve complex to the curve complex of a subsurface. Its main application in this paper is to force curve complex geodesics to run though prescribed regions of $\mathcal{C}(S)$, thereby guaranteeing a definite length.

Theorem 2.2 (Bounded Geodesic Image Theorem [MM00]). *Given S as above, there exists $K_{BGI} \geq 0$ so that if g is a geodesic in $\mathcal{C}(S)$ and Y is a subsurface of S such that $\text{diam}_Y(\pi_Y(g)) \geq K_{BGI}$, then there is a vertex α of g with $\pi_Y(\alpha) = \emptyset$.*

We can also project markings to the curve complex of a subsurface. For a marking μ and (nonannular) subsurface X the projection of μ to $\mathcal{C}(X)$ is defined as

$$\pi_X(\mu) = \cup_{\gamma \in \text{base}(\mu)} \pi_X(\gamma).$$

For α, β either curves or markings, we set

$$d_X(\alpha, \beta) = \text{diam}_X(\pi_X(\alpha) \cup \pi_X(\beta)),$$

when defined. With this notation, it is well known that the projection from the marking complex to the curve complex of a subsurface is coarsely 4-Lipschitz.

Given connected, non-isotopic, proper subsurfaces X and Y of S, the three possibilities for their relative position in S are captured by subsurface projection. We say X and Y are *disjoint* if $\pi_X(\partial Y) = \emptyset = \pi_Y(\partial X)$, or equivalently, if X and Y have isotopy class representatives with no points in common. Otherwise we say X and Y *intersect essentially*,

and write $X \cap Y \neq \emptyset$; there are two ways this may happen. It may be that, up to switching X and Y, we have $\pi_Y(\partial X) \neq \emptyset$ but $\pi_X(\partial Y) = \emptyset$, which implies that $X \subset Y$ up to isotopy; in this case we say X and Y are *nested*. Finally, if both $\pi_X(\partial Y)$ and $\pi_Y(\partial X)$ are nonempty, we say X and Y *overlap* and write $X \pitchfork Y$. A collection $\mathbb{X} = \{X_1, \ldots, X_n\}$ of nonannular proper subsurfaces of S is *cleanly embedded* if any two subsurfaces in the collection that intersect must overlap, i.e.

$$X_i \cap X_j \neq \emptyset \implies X_i \pitchfork X_j \quad \text{for } i \neq j.$$

In this case, we can consider the associated *coincidence graph* $\Gamma_{\mathbb{X}}$ whose vertices are labeled $v_1, \ldots, v_n$ in correspondence to the subsurfaces of $\mathbb{X}$, and in which v_i is joined by an edge to v_j if and only if $X_i \cap X_j = \emptyset$. See, for example, Figure 1. In general, a collection of subsurfaces $\{X_1, \ldots, X_n\}$ *fills* the surface S if for any curve γ, there is an i such that $\pi_{X_i}(\gamma) \neq \emptyset$. That is, there is a subsurface X_i in the collection that is cut by γ.
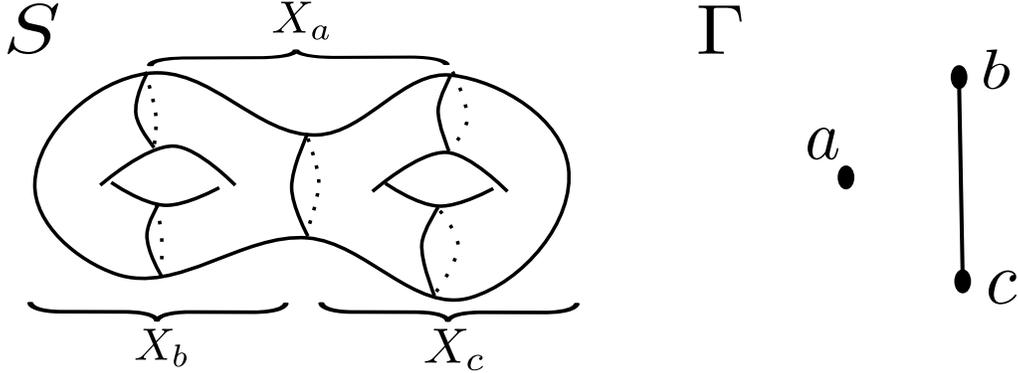

FIGURE 1. Cleanly embedded subsurfaces with coincidence graph

2.3. **Mapping class group basics.** The mapping class group of the surface S is the group of isotopy classes of orientation preserving homeomorphisms of S and is denoted $\mathrm{Mod}(S)$. For a nonannular subsurface X and $f \in \mathrm{Mod}(S)$, if f has a representative homeomorphism that is the identity outside of X then f is said to be *supported* on X. The mapping classes supported on X are precisely those mapping classes in the image of the natural map $\mathrm{Mod}(X) \to \mathrm{Mod}(S)$ (see [FM12] for details). We say f is *fully supported* on X if f is supported on X and the restriction of f to X, denoted f_X, is pseudo-Anosov. Recall that $g \in \mathrm{Mod}(X)$ is *pseudo-Anosov* if no positive power of g fixes an isotopy class of curve in X. See [FLP79] for the standard definition of pseudo-Anosov and the equivalence to the definition given here. There is a natural action of $\mathrm{Mod}(S)$ on the curve complex $\mathcal{C}(S)$ obtained by extending the action of $\mathrm{Mod}(S)$ on isotopy classes of curves.

For $f \in \mathrm{Mod}(S)$, define the *(stable) translation length* of f on $\mathcal{C}(S)$ as follows:

$$\ell_S(f) = \lim_{n \to \infty} \frac{d_S(\alpha, f^n(\alpha))}{n}.$$

where α is any vertex in $\mathcal{C}^0(S)$, the vertex set of $\mathcal{C}(S)$. It is a standard exercise to verify that this limit exists and is independent of α, and that $\ell_S(f^n) = n \cdot \ell_S(f)$. In [MM99], it is shown that if $f \in \mathrm{Mod}(S)$ is pseudo-Anosov then $\ell_S(f) \geq c > 0$, where c depends only on the topology of S. If f is fully supported on the subsurface X then we define $\ell_X(f) = \ell_X(f_X)$ which we note is clearly positive. For more discussion on translation length see Section 8.

Finally, recall that the Teichmüller space of S, denoted $\mathcal{T}(S)$, is the space of marked hyperbolic structures on the surface S. We will consider $\mathcal{T}(S)$ with its Teichmüller metric $d_{\mathcal{T}}$ and refer the reader to [FM12] for precise definitions and basic properties, including details about the action of $\mathrm{Mod}(S)$ on $\mathcal{T}(S)$ by isometries. With this background at hand, we define the subgroups of the mapping class group that are of interest in this paper.

Definition 1 (Farb–Mosher [FM02])**.** *Let S be a surface with $\xi(S) \geq 1$. Then a finitely generated subgroup $H < \mathrm{Mod}(S)$ is* convex cocompact *if for some $\mathcal{X} \in \mathcal{T}(S)$ the orbit $H \cdot \mathcal{X}$ is quasiconvex (with respect to the Teichmüller metric).*

Farb and Mosher introduced convex cocompact subgroups of mapping class groups in [FM02], confirmed that the definition does not depend on the basepoint $\mathcal{X}$, and proved its equivalence to a detailed condition requiring both hyperbolicity of the subgroup and its well-behaved, cocompact action on a "weak" convex hull in $\mathcal{T}(S)$, justifying the analogy with convex cocompact Kleinian groups. Of particular relevance to this paper is the following characterization of convex cocompact subgroups of mapping class groups:

Theorem 2.3. [KL08, Ham05] *Let S be a surface with $\xi(S) \geq 1$ and H be a finitely generated subgroup of $\mathrm{Mod}(S)$. Then H is convex cocompact if and only if for some (any) $\alpha \in \mathcal{C}^0(S)$ the orbit map*

$$H \to \mathcal{C}(S)$$

given by $h \mapsto h \cdot \alpha$ is a quasi-isometric embedding.

Using this characterization, it is immediate that if H is convex cocompact then any $h \in H$ with infinite order is pseudo-Anosov.

2.4. **A partial order on subsurfaces.** We recall the partial order on connected subsurfaces described in [CLM12]. The origins of this ordering can be found in [MM00, BKMM12]. Fix markings μ, μ' and fix $K \geq 20$. Define $\Omega(K, \mu, \mu') = \{Y \subset S : d_Y(\mu, \mu') \geq K\}$. Recall that connected subsurfaces X, Y *overlap* if both $\pi_Y(\partial X)$ and $\pi_X(\partial Y)$ are not empty, or equivalently, X and Y intersect essentially and one cannot be isotoped to be contained in the other. Define the relation $X \prec Y$ to mean X and Y overlap and

$$d_X(\mu, \partial Y) \geq 10.$$

The following properties of $\prec$ are verified in [CLM12].

Lemma 2.4. *Let $K \geq 20$ and choose $X, Y \in \Omega(K, \mu, \mu')$ that overlap. Then X and Y are ordered and the following are equivalent*

(1) $X \prec Y$
(2) $d_X(\mu, \partial Y) \geq 10$
(3) $d_X(\mu, \partial Y) \geq K - 4$
(4) $d_X(\mu', \partial Y) \leq 4$

(5) $d_Y(\mu', \partial X) \geq 10$
(6) $d_Y(\mu', \partial X) \geq K - 4$
(7) $d_Y(\mu, \partial X) \leq 4$

Moreover, $\prec$ is a strict partial order on $\Omega(K, \mu, \mu')$.

The following observation will be used in combination with the Bounded Geodesic Image Theorem (Theorem 2.2) to show when large subsurface projections combine to build up distance in the curve complex.

Lemma 2.5. *If $K \geq 20$, $X, Y, Z \in \Omega(K, \mu, \mu')$, and $X \prec Y \prec Z$, then any curve disjoint from both X and Z is also disjoint from Y.*

Proof. Suppose γ is disjoint from both X and Z but it intersects Y. Because both X and Z are ordered with Y, they overlap with Y. So we have that $d_Y(\partial X, \gamma), d_Y(\gamma, \partial Z) \leq 2$. Then

$$
\begin{aligned}
K \leq d_Y(\mu, \mu') &\leq d_Y(\mu, \partial X) + d_Y(\partial X, \gamma) + d_Y(\gamma, \partial Z) + d_Y(\partial Z, \mu') \\
&\leq 4 + 2 + 2 + 4 = 12
\end{aligned}
$$

using Lemma 2.4. This contradicts $K \geq 20$. $\qquad\square$

3. Right-angled Artin groups

3.1. **RAAGs and normal forms.** Let Γ be a simplicial graph with vertex set $V(\Gamma) = \{v_1, \ldots, v_n\}$ and edge set $E(\Gamma) \subset V(\Gamma) \times V(\Gamma)$. The right-angled Artin group, $A(\Gamma)$, associated to Γ is the group presented by

$$
\langle v_i \in V(\Gamma) : [v_i, v_j] = 1 \iff (v_i, v_j) \in E(\Gamma) \rangle.
$$

The v_i will be referred to as the *standard generators* of $A(\Gamma)$. Note that if Γ is the graph with n vertices and no edges then $A(\Gamma) = \mathbb{F}_n$. At the other extreme, if Γ is the complete graph of n vertices then $A(\Gamma) = \mathbb{Z}^n$. Because of this, $A(\Gamma)$ is often said to "interpolate" between free and free abelian groups. While simple to define, right-angled Artin groups have been at the center of major recent developments in geometric group theory and low-dimensional topology.

In this section, we briefly recall a normal form for elements of a right-angled Artin group. For details see Section 4 of [CLM12] and the references provided there. Fix a word $w = x_1^{e_1} \ldots x_k^{e_k}$ in the vertex generators of $A(\Gamma)$, with $x_i \in \{v_1, \ldots, v_n\}$ for each $i = 1, \ldots, k$. Each $x_i^{e_i}$ together with its index, which serves to distinguish between duplicate occurrences of the same generator, is a *syllable* of the word w. Let $\mathrm{syl}(w)$ denote the set of syllables for the word w. We consider the following 3 moves that can be applied to w without altering the element in $A(\Gamma)$ it represents:

(1) If $e_i = 0$, then remove the syllable $x_i^{e_i}$.
(2) If $x_i = x_{i+1}$ as vertex generators, then replace $x_i^{e_i} x_{i+1}^{e_{i+1}}$ with $x_i^{e_i + e_{i+1}}$.
(3) If the vertex generators x_i and x_{i+1} commute, then replace $x_i^{e_i} x_{i+1}^{e_{i+1}}$ with $x_{i+1}^{e_{i+1}} x_i^{e_i}$.

For $\sigma \in A(\Gamma)$, set $\mathrm{Min}(\sigma)$ equal to the set of words in the standard generators of $A(\Gamma)$ that have the fewest syllables among words representing σ. We refers to words in $\mathrm{Min}(\sigma)$ as *normal representatives* of σ. We also refer to a word in the standard generators as *normal* if it is normal for the element of $A(\Gamma)$ that it represents. Hermiller and Meiler showed in [HM95] that any word representing σ can be brought to any word in $\mathrm{Min}(\sigma)$ by application of the three moves above. Since these moves increase neither word nor syllable length, we see that words in $\mathrm{Min}(\sigma)$ are also minimal length with respect to the standard generators, and that any two words in $\mathrm{Min}(\sigma)$ differ by repeated application of move (3) only. For normal words w, w', we will occasionally use the notation $w \sim w'$ to denote that w and w' differ by a repeated application of move (3). In other words, $w \sim w'$ if and only if there is a $\sigma \in A(\Gamma)$ with $w, w' \in \mathrm{Min}(\sigma)$. It is verified in [CLM12] that for any $\sigma \in A(\Gamma)$ and $w, w' \in \mathrm{Min}(\sigma)$ there is a natural bijection between $\mathrm{syl}(w)$ and $\mathrm{syl}(w')$, which extends the obvious bijection between normal form words differing by a single application of move (3). Thus we define, for $\sigma \in A(\Gamma)$, $\mathrm{syl}(\sigma) := \mathrm{syl}(w)$ using any $w \in \mathrm{Min}(\sigma)$. This permits us to define, for each $\sigma \in A(\Gamma)$, a strict partial order $\prec$ on the set $\mathrm{syl}(\sigma)$ by setting $x_i^{e_i} \prec x_j^{e_j}$ if and only if for every $w \in \mathrm{Min}(\sigma)$ the syllable $x_i^{e_i}$ appears to the left of $x_j^{e_j}$ in the spelling of w.

One can imagine that the generators and their inverses correspond to directions in the Cayley graph. For example, the standard Cayley graphs for either $F_2 = \langle a, b \rangle$ or $\mathbb{Z}^2 = \langle a, b \mid ab = ba \rangle$ have four directions; typically a and a^{-1} correspond to east and west respectively, while b and b^{-1} point north and south. A word in normal form represents a geodesic path that also minimizes changes in direction.

Example 1. Take Γ to be the graph with vertex set $\{a, b, c, d\}$ and edge set $\{(a, b), (b, c), (c, d)\}$ and consider the word $w = acbd$. We see that w is in normal form and is equivalent through move (3) to the words $abcd$, $bacd$, $abdc$, and $badc$. Hence, the only ordered syllables of w are $a \prec c$, $a \prec d$, and $b \prec d$.

Our first lemma states immediate properties of our partial order and normal forms. We use this lemma without mention, in particular to prove a second, more technical lemma critical to our proof of the main theorem.

Lemma 3.1. *Subwords of normal words are themselves normal words. Unordered syllables of a normal word correspond to generators that commute.*

Lemma 3.2. *Suppose w is a normal word containing distinct, unordered syllables $p, q \in \mathrm{syl}(w)$ separated by the subword M. That is, up to switching p and q,*

$$w = x_1^{e_1} \cdots x_k^{e_k} \qquad p = x_i^{e_i} \qquad q = x_j^{e_j} \qquad M = x_{i+1}^{e_{i+1}} \cdots x_{j-1}^{e_{j-1}},$$

so pMq is the smallest subword of w containing p and q. Then M has normal representative $M' = LR$ where L and R are (possibly empty) subwords commuting with p and q respectively.

Proof. We induct on the syllable length of M. The claim is vacuous if M is the empty word. If M non-empty, observe that each syllable of M is ordered with at most one of p and q, since p and q are not ordered. Find the first syllable in M from the left which is ordered with p.

If no such syllable exists then the claim is true for $M = L$ and R empty. Otherwise, call this syllable s and observe that s and q are not ordered and therefore commute. Furthermore, $pMq = pL_1sM_1q$ where L_1 commutes with p by construction. Inductively, M_1 has normal representative L_2R_2 where L_2 commutes with s and R_2 commutes with q. Thus pMq has normal representative $L_1pL_2qsR_2$ got by a repeated application of move (3). Inductively, $L_2 = L_3R_3$ where p commutes with L_3 and q commutes with R_3. We have shown that the word M has normal representatives

$$M \sim L_1sM_1 \sim L_1sL_2R_2 \sim L_1L_2sR_2 \sim L_1L_3 \cdot R_3sR_2$$

where $L = L_1L_3$ commutes with p and $R = R_3sR_2$ commutes with q. $\square$

3.2. **Cubical geometry.** In this section we briefly review some geometry of non-positively curved cube complexes, which we use throughout the paper. Good references for this material are [BH09, HW08, Hag08]. First, recall that a simplicial complex is a *flag complex* if any $n+1$ pairwise adjacent vertices span an n-simplex. A subcomplex X of a simplicial complex Y is a *full subcomplex* when any simplex in Y whose vertices are in X is contained in X.

Definition 2. *A* cube complex X *is the space formed by isometrically gluing Euclidean unit cubes along their faces.* X *is* non-positivley curved (NPC) *if the link of each vertex is a flag simplicial complex.* X *is* CAT(0) *if it is* NPC *and simply connected.*

Because of their combinatorial nature, local isometries between NPC cube complexes have a particularly simple description; see [Cha07, BH09, Hag08] for details. A map between cube complexes is said to be *cubical* if it maps open cubes homeomorphically onto open cubes. Denote by $\text{lk}(x)$ the link of the vertex x of X.

Theorem 3.3. *Let* X, Y *be cube complexes,* Y NPC, *and* $f : X \to Y$ *a cubical map such that for each* $x \in X^0$ *the induced map on the link of x is injective and $f(\text{lk}(x))$ is a full subcomplex of* $\text{lk}(f(x))$ *in* Y. *Then* X *is* NPC, *and* f *is a local isometry. Further,* $f_* : \pi_1(X) \to \pi_1(Y)$ *is injective and the induced map on the universal covers* $\widetilde{f} : \widetilde{X} \to \widetilde{Y}$ *is an isometric embedding with* $\widetilde{f}(\widetilde{X}) \subset \widetilde{Y}$ *a convex subcomplex.*

Given a simplicial graph Γ with associated right-angled Artin group $A(\Gamma)$, we recall the definition of the *Salvetti complex* S_Γ. This is the cube complex defined as follows: begin with a rose, denoted S_Γ^1, with vertex x and $|V(\Gamma)|$ petals oriented and labeled by the vertices of Γ. Now attach 2-cubes for each edge of Γ corresponding to a commutation relator in $A(\Gamma)$. Specifically, if u and w are vertices of Γ joined by an edge, then a square is attached to S_Γ^1 with boundary label $uwu^{-1}w^{-1}$. For $n > 2$, we attach an n-cube for each set of n pairwise commuting generators (for each n-clique of Γ) whose attaching map restricted to each face is the characteristic map for that face. It is easy verified that S_Γ is an NPC cube complex with fundamental group $A(\Gamma)$, see [BH09] for example. We denote its universal cover $\widetilde{S}_\Gamma$ and fix a lift $\widetilde{x}$ of the unique vertex x of S_Γ. By the construction, it follows that $(\widetilde{S}_\Gamma^1, \widetilde{x})$ is isomorphic to the Cayley graph for $A(\Gamma)$ with its standard generating set. In this paper,

our cube complexes arise as compact locally convex subcomplexes of covers of the Salvetti complex. These are known as compact *special* cube complexes in the literature.

Besides the induced path metric on a CAT(0) cube complex X, we will also be interested in the graph metric on the 1-skeleton of X. We refer to this metric as *combinatorial distance* on X^0. Although we will not need these facts, we note that combinatorial distance can be alternatively characterized as the number of hyperplanes separating two vertices or the number of hyperplanes intersected by the CAT(0) geodesic between the vertices. Geodesics in this metric on X^1 are called *combinatorial geodesics*. For more on combinatorial distance and hyperplanes, see [Hag08], or the appendix of [HW08]. In the special case of $\widetilde{S}_\Gamma$, we note that combinatorial distance agrees with distance in $A(\Gamma)$ with its standard generating set via the identification $A(\Gamma) \to \widetilde{S}_\Gamma$ given by $g \mapsto g \cdot \widetilde{x}$.

Finally, recall that a collection of vertices $Y \subset X^0$ for X a CAT(0) cube complex is called *combinatorially K-quasiconvex* in X if *every* combinatorial geodesic between points of Y stays within combinatorial distance K from Y. Again translating to the group $A(\Gamma)$, quasiconvexity of the subgroup $H < A(\Gamma)$ with respect to the standard generators is equivalent to combinatorial quasiconvexity of $H \cdot \widetilde{x}$ in $\widetilde{S}_\Gamma$.

3.3. **Quasiconvex subgroups of RAAGs.** We will make use of the following theorem of Haglund. Here, the *convex hull* of a subcomplex Y of the CAT(0) cube complex X is the intersection of all convex subcomplexes containing Y. We recall that a subcomplex $Y \subset X$ is convex (with respect to the CAT(0) metric) if and only if Y is full and combinatorially convex [Hag08, HW08], where the subcomplex $Y \subset X$ is full if for each vertex $v \in Y^0$, the link of v in Y is a full subcomplex of the link of v in X.

Theorem 3.4. [Hag08] *Let X be a uniformly locally finite CAT(0) cube complex. Then for any $K \geq 0$ there exists an $L \geq 0$ such that the convex hull of any combinatorially K-quasiconvex subcomplex Y is contained in the L-neighborhood of Y.*

For a simplicial graph Γ and $K \geq 0$, define the function $f(\Gamma, K)$ as follows: first use Theorem 3.4 to choose an L so that the L-neighborhood of any combinatorially K-quasiconvex subcomplex Y of the CAT(0) cube complex $\widetilde{S}_\Gamma$ contains the convex hull of Y. Now define $f(\Gamma, K)$ to be the number of vertices of combinatorial distance $\leq L$ from $\widetilde{x} \in \widetilde{S}_\Gamma$, where $\widetilde{x}$ is a lift of the unique vertex of S_Γ. Hence, $f(\Gamma, K)$ counts the number of elements in the L-ball about the identity in the Cayley graph of $A(\Gamma)$ with its standard generators. We now obtain the following consequence of Theorem 3.4.

Lemma 3.5. *Let $H < A(\Gamma)$ be a quasiconvex subgroup with respect to the standard generators of $A(\Gamma)$. Then there exists a pointed compact cube complex (C, x) and a cubical local isometry $\phi : (C, x) \to (S_\Gamma, x)$ with $H = \phi_*(\pi(C, x))$. Furthermore if H is K-quasiconvex, then C has at most $f(\Gamma, K)$ vertices.*

Proof. Since quasiconvexity of H in the standard generators of $A(\Gamma)$ is equivalent to combinatorial quasiconvexity of $H \cdot \widetilde{x}$, Theorem 3.4 provides a convex subcomplex $\widetilde{C}$ containing the convex hull of $H \cdot \widetilde{x}$ in $\widetilde{S}_\Gamma$ that is H-invariant and cocompact. Set $C = \widetilde{C}/H$ and let

$\phi : C \to S_\Gamma$ denote the composition $\widetilde{C}/H \to \widetilde{S}_\Gamma/H \to S_\Gamma$. As the composition of local isometries, ϕ is a local isometry and $H = \phi_*(\pi(C, x))$.

To conclude, note that if H is K-quasiconvex then the number of vertices of C is no greater than the number of vertices of $N_L(H \cdot \widetilde{x})/H$, as $\widetilde{C} \subset N_L(H \cdot \widetilde{x})$. This is in turn bounded by the number of vertices of $N_L(\widetilde{x})$ in $\widetilde{S}_\Gamma$, because the projection map $N_L(\widetilde{x}) \to N_L(H \cdot \widetilde{x})/H$ is surjective. Hence, the number of vertices of C is bounded above by $f(\Gamma, K)$. $\square$

Label the oriented edges of S_Γ by the vertex generators of $A(\Gamma)$ as in the construction of S_Γ. Using the cubical map ϕ, pull back these labels to the oriented edges of C so that oriented loops in C at x correspond to words in H written in the standard generators of $A(\Gamma)$. We note the following implication of the geometric properties of C. First, call an edge path in a nonpositively curved cube complex X a *combinatorial local geodesic* if each of its lifts to $\widetilde{X}$ is a combinatorial geodesic. Note that this property is preserved under local isometry of cube complexes.

Lemma 3.6. *With the notation above, oriented edge-loops based at $x \in C$ which are combinatorial local geodesics are in bijective correspondence with minimal length words in H with respect to the standard generators of $A(\Gamma)$.*

Proof. By Lemma 3.5, a loop based at x in S_Γ corresponding to a word in H is a combinatorial local geodesic in S_Γ if and only if it is contained in C and is a combinatorial local geodesic there. More precisely, if γ is a combinatorial local geodesic loop of (S_Γ, x) with $[\gamma] \in H$, then its lifts based at $\widetilde{x}$ in $\widetilde{S}_\Gamma$ are contained in $\widetilde{C}$, and so γ corresponds to a combinatorial local geodesic in C. Conversely, combinatorial local geodesic loops of (C, x) map through ϕ to combinatorial local geodesic loops of (S_Γ, x) representing elements of H.

The lemma then reduces to the observation that since $(\widetilde{S}_\Gamma^1, \widetilde{x})$ is isomorphic to the Cayely graph of $A(\Gamma)$ with the standard vertex generators, combinatorial local geodesics in S_Γ correspond to minimal length words in $A(\Gamma)$. $\square$

The following proposition gives us control over the normal forms of elements of H.

Proposition 3.7. *Let H be a K-quasiconvex subgroup of $A(\Gamma)$. There is an $\ell = \ell(\Gamma, K) \geq 0$ so that for $h \in H$ and a minimal length representative $w = w_1 w_2 \ldots w_n$ of h in the standard generators of $A(\Gamma)$, any subword $w_{i+1} w_{i+2} \ldots w_j$ of w with length $j - i \geq \ell/3$ contains as a subword a conjugate of an element of H of $A(\Gamma)$-length $\leq \ell$.*

Proof. Since H is quasiconvex in $A(\Gamma)$ with the standard generators, $H \cdot \widetilde{x}$ is combinatorially quasiconvex in $\widetilde{S}_\Gamma$. Hence, there is a cube complex C and a cubical local isometry $\phi : (C, x) \to (S_\Gamma, x)$ with $H = \phi_*(\pi_1(C, x))$, as discussed above. Set $\ell = 3f(\Gamma, K)$.

If w is as in the statement of the proposition, then by Lemma 3.6 above w corresponds to a combinatorial local geodesic in C. Since subpaths of combinatorial geodesics in $\widetilde{C}$ are themselves combinatorial geodesics, the subword $w_{i+1} w_{i+2} \ldots w_j$ is also a combinatorial local geodesic of C. This implies, in particular, that the edge path in C spelled by $w_{i+1} w_{i+2} \ldots w_j$

does not backtrack. If $j - i \geq f(\Gamma, K)$, then this edge path must repeat a vertex, say $v \in C$, since the number of vertices of C is less than $f(\Gamma, K)$ by Lemma 3.5.

Let α be the subpath of $w_{i+1}w_{i+2}\ldots w_j$ beginning and ending at v. Again, α is a combinatorial local geodesic loop, and its length is no more than $f(\Gamma, K)$. Since the combinatorial distance from x to v is also at most $f(\Gamma, K)$, we have that α represents a word in $A(\Gamma)$ conjugate to a loop in C based at x of length $\leq 3f(\Gamma, K) = \ell$. This complete the proof. $\square$

For concrete examples of the local isometries $C \to S_\Gamma$ considered above, see Section 7.

4. Admissible $A(\Gamma)$ in Mod(S)

4.1. **The [CLM12] construction.** In order to introduce notation and motivate the condition to which our theorem applies, we begin by describing the types of the homomorphisms constructed in [CLM12]. Begin with a finite simplicial graph Γ and a collection of cleanly embedded, connected nonannular subsurfaces $\mathbb{X} = \{X_1, \ldots, X_n\}$ in S, whose coincidence graph is $\Gamma = \Gamma_{\mathbb{X}}$. We label the vertices of Γ by $v_1, \ldots, v_n$ so that these indices agree with subsurfaces they represent in $\mathbb{X}$. Since mapping classes supported on disjoint subsurfaces commute, any choice of mapping classes $\mathbb{F} = \{f_1, \ldots, f_n\}$ with $f_i \in \text{Mod}(X_i) \leq \text{Mod}(S)$ determines a homomorphism $\phi = \phi_{\mathbb{F}} : A(\Gamma) \to \text{Mod}(S)$ with $\phi(v_i) = f_i$. Suppose we have fixed such a homomorphism where the mapping classes are fully supported on their respective surfaces. Recall this implies that for each $1 \leq i \leq n$, $\ell_{X_i}(f_i) > 0$. Informally, the main theorem from [CLM12] concludes that if these translation lengths are large enough, depending only on $\mathbb{X}$, then the induced homomorphism into $A(\Gamma)$ is a quasi-isometric embedding. Here are some of the relevant details.

For $\sigma \in A(\Gamma)$ with $w = x_1^{e_1} \ldots x_k^{e_k} \in \text{Min}(\sigma)$, define $J : \{1, \ldots, k\} \to \{1, \ldots, n\}$ so that $x_i = v_{J(i)}$ as standard generators of $A(\Gamma)$. Hence, $\phi(x_i) = f_{J(i)}$ is supported on $X_{J(i)}$. Write

$$X^w(x_i^{e_i}) = \phi(x_1^{e_1} \ldots x_{i-1}^{e_{i-1}})(X_{J(i)})$$

for $i = 2, \ldots, k$ and $X^w(x_1^{e_1}) = X_{J(1)}$. This defines a map

$$X^w : \text{syl}(w) \to \Omega(S).$$

It is verified in [CLM12] that this map is well-defined for $\sigma \in A(\Gamma)$ independent of the choice of normal representative w, so we set $X^\sigma := X^w$ for $w \in \text{Min}(\sigma)$.

Define $\text{SUBS}(\sigma)$ as the image of the map $X^\sigma : \text{syl}(\sigma) \to \Omega$. Thus we may associate to every syllable s in $\text{syl}(\sigma)$ the subsurface $Y = X^\sigma(s)$ in $\text{SUBS}(\sigma)$, without reference to the particular indexing of syllables in w. The subsurface Y is not to be confused with the support of the syllable $\phi(s) = f_j^i$, denoted $\text{supp}(\phi(s))$, which is X_j.

Example 2. Consider the surface S and the collection of subsurfaces $\mathbb{X} = \{X_a, X_b, X_c\}$ given in Figure 1 with coincidence graph Γ. Let $\mathbb{F} = \{f_a, f_b, f_c\}$ be a collection of mapping classes that are fully supported on $\mathbb{X}$ and let $\phi : A(\Gamma) \to \text{Mod}(S)$ be the induced homomorphism. Consider the normal form word $w = abca$. Then $\text{SUBS}(w) = \{X_a, f_a X_b, f_a f_b X_c, f_a f_b f_c X_a\}$. A single application of move (3) yields the word $w' = acba$ with $\text{SUBS}(w') =$

$\{X_a, f_a X_c, f_a f_c X_b, f_a f_c f_b X_a\}$. As X_b and X_c are disjoint, f_b and f_c commute and so $\mathrm{SUBS}(w) = \mathrm{SUBS}(w')$, as claimed above.

4.2. **Admissibility.** Suppose that we have fixed the homomorphism $\phi_{\mathbb{F}} : A(\Gamma) \to \mathrm{Mod}(S)$ as in the previous section. The following definition collects the properties necessary for the proof of our main theorem.

Definition 3. *Fix the marking $\mu \in \mathcal{M}(S)$. Let Γ be the coincidence graph of a collection $\mathbb{X} = \{X_1, \ldots, X_n\}$ of cleanly embedded, nonannular, proper subsurfaces of S. Choose $\mathbb{F}$ a collection of mapping classes fully supported on $\mathbb{X}$. The homomorphism $\phi = \phi_{\mathbb{F}} : A(\Gamma) \to Mod(S)$ is K-admissible if the following conditions hold for each $\sigma \in A(\Gamma)$:*
- *For each syllable $x_i^{e_i}$ of σ, we have $d_{X^\sigma(x_i^{e_i})}(\mu, \phi(\sigma)\mu) \geq K|e_i|$. In particular, $X^\sigma(x_i^{e_i}) \in \Omega(K, \mu, \phi(\sigma)\mu)$.*
- *The map $X^\sigma : \mathrm{syl}(\sigma) \to \mathrm{SUBS}(\sigma)$ is injective and order preserving, where $\mathrm{SUBS}(\sigma) \subset \Omega(K, \mu, \phi(\sigma)\mu)$ is given the induced ordering.*

The main result from [CLM12] can then be rephrased as

Theorem 4.1. [CLM12] *Let $\mathbb{X} = \{X_1, \ldots, X_n\}$ be cleanly embedded in S with co-incidence graph $\Gamma_{\mathbb{X}}$. Then for any K there is a $c_0 \geq 0$ so that the following holds: if $\mathbb{F} = \{f_1, \ldots, f_n\}$ is a collection of mapping classes fully supported on $\mathbb{X}$ and for each $f_i \in \mathrm{Mod}(X_i)$ we have $\ell_{X_i}(f_i) \geq c_0$, then the induced homomorphism $\phi_{\mathbb{F}} : A(\Gamma) \to \mathrm{Mod}(S)$ is K-admissible. For K sufficiently large, any K-admissible homomorphism is a quasi-isometric embedding into $\mathrm{Mod}(S)$ and has a quasi–isometric orbit map into Teichmüller space.*

We remark that, while in the proof of Theorem 4.1 showing that large translation length implies admissibility is quite involved, once admissibility is demonstrated, showing that the homomorphisms are quasi-isometric embeddings into $\mathrm{Mod}(S)$ and Teichmüller space follows immediately from the distance formulas of Masur-Minsky [MM00] and Rafi [Raf07], respectively. In [Tay13], the second author gives a similar criterion for quasi-isometric embeddings of right-angled Artin groups into $\mathrm{Out}(F_n)$.

Definition 4. *Assuming the set-up of Definition 3, $\phi_{\mathbb{F}} : A(\Gamma) \to \mathrm{Mod}(S)$ is admissible if it is K-admissible for some $K \geq \max\{K_{BGI}, 20\}$, with K_{BGI} as in Theorem 2.2.*

4.3. **Filling words, filling blocks, and ℓ-short filling subgroups.** Fix an admissible homomorphism $\phi_{\mathbb{F}} : A(\Gamma) \to \mathrm{Mod}(S)$. In this section we define three notions of "filling," for words, subwords, and subgroups in $A(\Gamma)$ respectively. We emphasize that these conditions depend on the particular admissible embedding of $A(\Gamma)$ into $\mathrm{Mod}(S)$, which we denote by ϕ for brevity. These definitions allow us to develop the technical machinery used in the proof of Theorem 5.2, from which we derive the "if" direction of Theorem 1.1: that quasiconvex, all-pseudo-Anosov subgroups of admissible embeddings $A(\Gamma)$ are convex cocompact.

Given $\sigma \in A(\Gamma)$, recall from Section 3.1 that $\mathrm{syl}(\sigma)$ is the set of syllables of a normal representative for σ. Call σ *cyclically reduced* if it has the least number of syllables $|\mathrm{syl}(\sigma)|$ among all of its conjugates. If σ is cyclically reduced, we say that σ *fills* if the collection of

subsurfaces SUBS(σ) fills S; that is, for any curve γ there is a subsurface $Y \in$ SUBS(σ) such that $\pi_Y(\gamma)$ is not empty. For arbitrary $\sigma \in A(\Gamma)$, we say that σ fills if σ is conjugate to σ' where σ' is cyclically reduced and fills. Note that this is equivalent to requiring that all conjugates σ' of σ are such that the collection of subsurfaces SUBS(σ') fills S. We say a word w in the standard generators of $A(\Gamma)$ *fills*, or is a *filling word*, if and only if $w \in \mathrm{Min}(\sigma)$ for some $\sigma \in A(\Gamma)$ that fills.

Filling words relate to pseudo-Anosov mapping classes in an unsurprising way. The idea for the next lemma appears in [CLM12], but we include a proof for completeness. Let $\mathbb{X}_\sigma$ be the collection of subsurfaces

$$\mathbb{X}_\sigma = \{X \in \mathbb{X} : X = \mathrm{supp}(\phi(x_i)) \text{ for some } x_i^{e_i} \in \mathrm{syl}(\sigma)\}.$$

Lemma 4.2. *The set of subsurfaces* SUBS(σ) *fills S if and only if $\mathbb{X}_\sigma$ fills S.*

Proof. Fix an indexing by i of syllables $x_i^{e_i}$ in $\mathrm{syl}(\sigma)$, their supports X_i in $\mathbb{X}$, and their associated subsurfaces $Y_i = \phi(x_1^{e_1} \cdots x_{i-1}^{e_{i-1}})X_i$ in SUBS(σ). If $\mathbb{X}_\sigma$ fills S, then given a curve γ we may consider the smallest index I such that γ cuts X_I. Because $\phi(x_1^{e_1} \cdots x_{I-1}^{e_{I-1}})$ fixes γ, we know $\gamma = \phi(x_1^{e_1} \cdots x_{I-1}^{e_{I-1}})(\gamma)$ cuts $Y_I = \phi(x_1^{e_1} \cdots x_{I-1}^{e_{I-1}})X_I$. Since γ was arbitrary, we conclude SUBS(σ) fills S. On the other hand, if some curve γ misses all X_i in $\mathbb{X}_\sigma$, then because $\gamma = \phi(x_1^{e_1} \cdots x_{i-1}^{e_{i-1}})(\gamma)$ for any i, it is easy to see that γ is also disjoint from all Y_i in SUBS(σ). $\qquad\square$

Lemma 4.2 implies that, if SUBS(σ) fails to fill, $\phi(\sigma)$ fixes the isotopy class of some curve and thus cannot be pseudo-Anosov. Therefore, for any $\sigma \in A(\Gamma)$, if $\phi(\sigma)$ is pseudo-Anosov, then σ fills. The converse is shown in [CLM12]: if σ fills then $\phi(\sigma)$ is pseudo-Anosov. However, we do not directly use this fact, which can be recovered from Theorem 5.2 in the case H is a cyclic group. Indeed, rather than assume that $\phi(H)$ is a purely pseudo-Anosov subgroup of the mapping class group, Theorem 5.2 requires the (*a priori* weaker) filling assumption on a finite collection of elements of the subgroup H. We formalize this assumption with a definition: say that for $H \leq A(\Gamma)$ and $\ell > 0$, H is *ℓ-short filling* if for all $h \in H$ with $|h|_{A(\Gamma)} \leq \ell$, h fills. We remind the reader that this condition is for a fixed admissible homomorphism $\phi : A(\Gamma) \to \mathrm{Mod}(S)$.

Consider a normal word $w = s_1 \cdots s_k$ where s_i are syllables in $\mathrm{syl}(w)$. Suppose w' is a subword consisting of a sequence of consecutive syllables, so $w' = s_i s_{i+1} \cdots s_j$ where $i < j$. If SUBS(w') fills S, we say w' is a *filling block* for w. Note that, while every filling word has a filling block, filling blocks can also appear in non-filling words. We are interested in filling blocks because, as Lemma 4.4 shows, these have consequences about syllable order. Theorem 4.1 relates syllable order to subsurface order, and we prove Theorem 5.2 by relating subsurface order to distances in the curve complex.

The proof of Theorem 5.2 relies on the lemmas that follow. We have already mentioned Lemma 4.4, which shows that filling blocks block commutation. Lemma 4.6 show how to find filling blocks for ℓ-short filling H. Before either of these, we need to relate SUBS(w') to SUBS(w) when w' is a subword of w. Given $u, v \in A(\Gamma)$, write $u \cdot$ SUBS(v) to denote the collection $\{\phi(u)Y : Y \in$ SUBS(v)$\}$.

Lemma 4.3. *Suppose w is a normal word with length k in the standard generators of $A(\Gamma)$, partitioned into subwords by $w = BME$, so*

$$w = x_1 \cdots x_k \qquad B = x_1 \cdots x_i \qquad M = x_{i+1} \cdots x_j \qquad 1 \leq i < j \leq k$$

where x_i are generators of $A(\Gamma)$, and E may be empty. Then $B \cdot \mathrm{SUBS}(M) \subset \mathrm{SUBS}(w)$, and any syllable $s' \in \mathrm{syl}(M)$ with associated subsurface $Y \in \mathrm{SUBS}(M)$ is a subword of the syllable $s \in \mathrm{syl}(w)$ with associated subsurface $\phi(B)Y$.

Proof. This follows by inspection, using Lemma 3.1 and the definition of SUBS. $\square$

Lemma 4.4. *Suppose the normal word $w = s_1 \cdots s_k$ contains a filling block $w' = s_i s_{i+1} \cdots s_j$, where s_n are syllables in $\mathrm{syl}(w)$ for $1 \leq n \leq k$. Then each syllable s_n is ordered with some syllable of w'.*

Proof. Let $Y \in \mathrm{SUBS}(w)$ be the subsurface associated to the syllable s_n, that is $Y = X^w(s_n)$. Write $w = Bw'E$ where B and E are the prefix and suffix respectively of the subword w' in w. Because $\mathrm{SUBS}(w')$ fills, so does its homeomorphic image $B \cdot \mathrm{SUBS}(w')$, which is a subset of $\mathrm{SUBS}(w)$ by Lemma 4.3. In particular we have $Z \in B \cdot \mathrm{SUBS}(w') \subset \mathrm{SUBS}(w)$ such that $\pi_Z(\partial Y)$ is not empty. Let s_m be the syllable of w associated to Z; observe that s_m is also a syllable of w'. Write X_n and X_m for the supports of s_n and s_m respectively.

We will show that s_n and s_m are ordered. Suppose not. Then w has normal representatives $w_1 = bs_n s_m e$ and $w_2 = bs_m s_n e$ which each represent some $\sigma \in A(\Gamma)$. Since either of X^{w_1} and X^{w_2} can be used to determine the bijection $X^\sigma : \mathrm{syl}(\sigma) \to \mathrm{SUBS}(\sigma)$, we know

$$Y = \phi(b)X_n = \phi(bs_m)X_n \qquad \text{while} \qquad Z = \phi(bs_n)X_m = \phi(b)X_m.$$

This means $\phi(s_n)X_m = X_m$, which requires that X_m and X_n are either the same or they are disjoint subsurfaces, by admissibility. Either case contradicts our finding that

$$\pi_Z(\partial Y) = \pi_{\phi(b)X_m}(\partial(\phi(b)X_n)) = \phi(b)(\pi_{X_m}(\partial X_n))$$

is not empty. $\square$

Finally, we prove two lemmas to find filling blocks whenever $H < A(\Gamma)$ is quasiconvex and ℓ-short filling for appropriate ℓ.

Lemma 4.5. *Given an admissible embedding $\phi : A(\Gamma) \to \mathrm{Mod}(S)$, suppose $H < A(\Gamma)$ is quasiconvex and ℓ-short filling, where ℓ is the constant determined in Proposition 3.7. Then for any $h \in H$, every syllable $x^e \in \mathrm{syl}(h)$ has exponent $e < \ell/3$.*

Proof. Supposing the contrary, let $w \in \mathrm{Min}(h)$ with syllable x^e. Since normal form words are length minimizing, Proposition 3.7 says that x^e contains a subword conjugate to $h' \in H$ where $|h'|_{A(\Gamma)} \leq \ell$. Then by hypothesis, h' fills. Thus all its conjugates fill, including the subword of x^e. But any such subword is merely a power of the generator x, whose image under the admissible embedding ϕ is supported on a proper subsurface of S. In particular, such a subword cannot fill, our contradiction. $\square$

Lemma 4.6. *Given an admissible embedding $\phi : A(\Gamma) \to \mathrm{Mod}(S)$, suppose $H < A(\Gamma)$ is K-quasiconvex and ℓ-short filling, where $\ell = \ell(\Gamma, K)$ is the constant determined in Proposition 3.7. Then for any $h \in H$ and any $w \in \mathrm{Min}(h)$, every subword of w of length at least ℓ contains a filling block for w.*

Proof. Let w' be a subword of w of length at least ℓ. Then w' partitions into subwords $w' = BME$ where M has length $\ell/3$ while B and E have length at least $\ell/3$. By Proposition 3.7, M has a subword M' which, as an element of $A(\Gamma)$, is conjugate to $h' \in H$ such that $|h'|_{A(\Gamma)} \leq \ell$. By the hypothesis that H is ℓ-short filling, h' is a filling word, and therefore so is its conjugate M'. This means $\mathrm{SUBS}(M')$ fills. If M' is not already a filling block for w, it must be that the first and/or last syllables of M' are proper subwords of syllables of w. By Lemma 4.5, since B and E have length greater than $\ell/3$, it is possible to complete M' to a filling block of w which is entirely contained in w'. $\qquad\square$

5. Proof of convex cocompactness and Theorem 1.3

In this section, we provide the conditions for when a subgroup of the mapping class group is convex compact, given that it comes from an admissible embedding of a right-angled Artin group. The idea is the following: fix a marking μ and begin with a normal word w of $H \leq A(\Gamma)$. If w has n syllables, admissibility implies that there are at least n subsurfaces with large subsurface projection between μ and its translate $\phi(w)\mu$. An application of the Bounded Geodesic Image Theorem (Theorem 2.2) implies that any geodesic between the marking and its translate in $\mathcal{C}(S)$ must enter a 1-neighborhood in the curve complex of the boundary of each of the subsurfaces of large projection distance. This will force curve complex distance of at least $\frac{n}{k}$ so long as we have a k-to-1 association between subsurfaces of large projection distance and, on the geodesic between the markings, vertices that are distance 1 from the boundaries of these subsurfaces in $\mathcal{C}(S)$. Hence, the difficulty is determining such a k that is independent of the word w. This is where the cubical geometry developed in Section 3.3 is employed. In particular, the existence of the compact, locally convex subcomplex C, coming from quasiconvexity of H, controls the regularity with which one encounters filling subwords in the spelling of w. This regularity, along with the order-preserving map X^w, implies that syllables sufficiently spaced out in the spelling of w correspond to filling subsurfaces. Hence, two subsurfaces coming from such syllables cannot both be disjoint from a single vertex of our geodesics, thereby forcing up the amount that w displaces the original marking. We now give the details, beginning with the assumption that, in a normal representative of a word in H, filling blocks occur with a definite frequency.

Proposition 5.1. *For a marking μ of S and an admissible homomorphism $\phi : A(\Gamma) \to \mathrm{Mod}(S)$, suppose $H \leq A(\Gamma)$ satisfies the following: there exists $\rho \geq 0$ such that for every $h \in H$, any normal form $w \in \mathrm{Min}(h)$ has the property that every subword of w of length $\geq \rho$ contains a filling block. Then for any $h \in H$,*

$$d_S(\mu, \phi(h)\mu) \geq |h|_{A(\Gamma)}/6\rho - 2.$$

Proof. Suppose that an admissible homomorphism $\phi : A(\Gamma) \to \mathrm{Mod}(S)$ is given, and let $\rho \geq 0$ be as in the statement of the proposition. Given $h \in A(\Gamma)$, choose a normal representative $w \in \mathrm{Min}(h)$. Then $w = x_1 \cdots x_N$ where x_i are standard generators (not syllables) and $N = |h|_{A(\Gamma)}$. Let m be the smallest integer strictly greater than $N/6\rho - 1$. We can partition w into m subwords of length 6ρ by setting

$$W_j = x_{I(j)+1} \cdots x_{I(j+1)} \qquad I(j) = 6\rho \cdot (j-1)$$

so that $w = W_1 \cdots W_m W_{m+1}$, where W_{m+1} is some possibly empty suffix.

For $1 \leq j \leq m$, each W_j admits a decomposition

$$W_j = P_j Q_j R_j, \quad \text{where} \quad |P_j| = 3\rho \quad |Q_j| = \rho \quad |R_j| = 2\rho,$$

and $|\cdot|$ denotes the length of each subword; recall that subwords of w are all minimal length for the elements of $A(\Gamma)$ they represent. Let p_j be the first syllable of P_j that is also a syllable of w for $1 \leq i \leq m$. There are m of these, and we claim that these syllables can be associated to distinct vertices in a fixed geodesic $[\mu, \phi(h)\mu]$ in the curve complex, as follows.

Let $Y_j \in \mathrm{SUBS}(w)$ correspond to the syllable p_j, that is, $Y_j = X^w(p_j)$. Here the index corresponds to the W_j subwords, not the syllable indexing of w. Since ϕ is admissible, we know that $d_{Y_j}(\mu, \phi(h)\mu) \geq K \geq K_{BGI}$. By Theorem 2.2, any curve complex geodesic $[\mu, \phi(h)\mu]$ contains a vertex v_j with empty projection to Y_j. It remains to show that these v_j are distinct for j between 1 and m; then we can conclude that

$$d_S(\mu, \phi(h)\mu) \geq m - 1 \geq |h|_{A(\Gamma)}/6\rho - 2,$$

completing the proof of the theorem. Towards contradiction, suppose there exists $j < k$ such that Y_j and Y_k contribute the same vertex in $[\mu, \phi(h)\mu]$. This means there exists a curve $\gamma = v_j = v_k$ disjoint from both Y_j and Y_k.

Now consider each Q_j and recall $|Q_j| = \rho$. Our hypothesis ensures Q_j contains a filling block for w. In particular, this gives a syllable $q \in \mathrm{syl}(w)$ that falls in Q_j and is associated to a subsurface $Z \in \mathrm{SUBS}(w)$ for which $\pi_Z(\gamma)$ is not empty.

Let us show that Z and Y_j are ordered; by admissibility it suffices to show that q and p_j are ordered. Recall p_j is the first syllable of P_j that is also a syllable of w, and P_j has length 3ρ. Lemma 4.5 assures us that the suffix of p_j in P_j has length at least 2ρ. Since we know q lies in Q_j, we see that the shortest subword of w containing p_j and q has the form $p_j M q$, where M has length at least 2ρ. If p_j and q are not ordered, then Lemma 3.2 says M has normal representative LR where p_j commutes L and q commutes with R. On the other hand, since M has length at least 2ρ, one of the pair L, R is longer than ρ, and thus contains a filling block (by our hypothesis) which includes syllables ordered with each of p_j and q (by Lemma 4.4). This contradicts the characterization of L and R from Lemma 3.2. So it must be p_j and q are ordered. Then so are Y_j and Z. As w is a normal word, evidently $Y_j \prec Z$.

Similarly, because R_j has length at least 2ρ, we can conclude that $Z \prec Y_k$. We have shown $Y_j \prec Z \prec Y_k$. By Lemma 2.5, since γ is disjoint from both Y_j and Y_k, it must also be disjoint from Z. This contradicts our choice of Z such that $\pi_Z(\gamma)$ is not empty. Hence, no curve is simultaneously disjoint from Y_j and Y_k for $j \neq k$. $\square$

By combining Lemma 4.6 with Proposition 5.1, we obtain:

Theorem 5.2. *Fix a marking μ of S and an admissible homomorphism $\phi : A(\Gamma) \to Mod(S)$. Let H be a K-quasiconvex subgroup of $A(\Gamma)$. For $\ell = \ell(\Gamma, K)$ as determined in Proposition 3.7, if H is ℓ-short filling, then $\phi(H)$ is a convex cocompact subgroup of the mapping class group. In particular, for any $h \in H$,*

$$d_S(\mu, \phi(h)\mu) \geq |h|_{A(\Gamma)}/6\ell - 2.$$

Proof. We show that the orbit map from H to $\mathcal{C}(S)$ sending h to $\phi(h)\mu$ is a quasi-isometric embedding. This will prove that $\phi(H) < Mod(S)$ is convex cocompact. Recall that since H is quasiconvex, it is undistorted in $A(\Gamma)$ and so we may use the word norm on $A(\Gamma)$ in place of a word norm for H. Being quasiconvex, H is finitely generated, and therefore the orbit map is coarsely Lipshitz. In detail, if $h_1, \ldots, h_s$ generates H and $C = \max\{d_S(\mu, \phi(h_i)\mu) : 1 \leq i \leq s\}$ then $d_S(\mu, \phi(h)\mu) \leq C \cdot |h|$, where $h \in H$ and $|\cdot|$ denotes the word norm in these generators. Hence, it suffices to prove the lower bound on curve complex distance given in the statement of the theorem.

Since $H \leq A(\Gamma)$ is K-quasiconvex and ℓ-short filling, Lemma 4.6 implies that for any $h \in H$ and any $w \in \text{Min}(h)$, every subword of w of length at least ℓ contains a filling block for w. Hence, we see that H and $\phi : A(\Gamma) \to Mod(S)$ satisfy the hypotheses of Proposition 5.1 with $\ell = \ell(\Gamma, K)$ playing the role of $\rho \geq 0$. We conclude that

$$d_S(\mu, \phi(h)\mu) \geq |h|_{A(\Gamma)}/6\ell - 2,$$

completing the proof of the theorem. $\qquad\square$

Since pseudo-Anosov mapping classes of H are necessarily filling, Theorem 1.3 from the introduction follows immediately from Theorem 5.2:

Proof of Theorem 1.3. Suppose that H is K-quasiconvex in $A(\Gamma)$ and let $\ell = \ell(\Gamma, K)$ be as in Theorem 5.2. By assumption, each nonidentity $h \in H$ of $A(\Gamma)$-length at most ℓ is pseudo-Anosov. Hence, H is ℓ-short filling. By Theorem 5.2, H is convex cocompact as a subgroup of $Mod(S)$. $\qquad\square$

For the same reason, we have one direction of Theorem 1.1 as a corollary:

Corollary 5.3. *Let $\phi : A(\Gamma) \to Mod(S)$ be an admissible homomorphism and suppose $H \leq A(\Gamma)$ is quasiconvex. If $\phi(H)$ is purely pseudo-Anosov, then it is convex cocompact in $Mod(S)$.*

Observe that although Corollary 5.3 is simpler to state, Theorem 5.2 provides a means of verifying that a quasiconvex subgroup $H \leq A(\Gamma)$ is actually purely pseudo-Anosov, assuming that one knows the quasiconvexity constant for H.

6. NECESSITY OF QUASICONVEXITY IN $A(\Gamma)$ AND THE PROOF OF THEOREM 1.1

Fix an admissible homomorphism $\phi : A(\Gamma) \to Mod(S)$. Having shown that for $H < A(\Gamma)$ with $\phi(H)$ purely pseudo-Anosov, quasiconvexity of H in $A(\Gamma)$ implies convex cocompactness

of $\phi(H)$ in $\text{Mod}(S)$, it is natural to ask whether quasiconvexity of H is a necessary condition for $\phi(H)$ to be convex cocompact. To answer this question, we first review some facts about the geometry of Teichmüller space.

Recall that the ϵ-thick part of Tecihmüller space is the region determined by

$$\mathcal{T}_{\geq\epsilon}(S) = \{X \in \mathcal{T}(S) : \ell_X(\gamma) \geq \epsilon \text{ for all } \gamma \in \mathcal{C}^0(S)\}$$

and a set in $\mathcal{T}(S)$ is called ϵ-*cobounded* if it resides in the ϵ-thick part. Although Teichmüller space itself is in no ordinary sense negatively curved, a driving principle in the study of the coarse geometry of $\mathcal{T}(S)$ is that the thick part $\mathcal{T}_{\geq\epsilon}$ has many hyperbolic-like features. See for example [KL07]. One manifestation of this principle is Minsky's theorem that Teichmüller geodesics which remain in the thick part of Teichmüller space are strongly contracting. More precisely,

Theorem 6.1 (Theorem 4.2, [Min96]). *Given K and ϵ, there exists a B so that if α is a K-quasigeodesic path in $\mathcal{T}(S)$ whose endpoints are connected by an ϵ-cobounded Teichmüller geodesic τ, then α is contained in the B-neighborhood of τ.*

We can now show that quasiconvexity of $H < A(\Gamma)$ is necessary for $\phi(H)$ to be convex cocompact in $\text{Mod}(S)$. The idea is that quasiconvexity in $\mathcal{T}(S)$ can be pulled back to quasiconvexity in $A(\Gamma)$. Recall that admissible homomorphisms induce quasi-isometric orbit maps into $\mathcal{T}(S)$.

Theorem 6.2. *Consider $A(\Gamma) < \text{Mod}(S)$ so that for $X \in \mathcal{T}(S)$ the orbit map*

$$\phi : A(\Gamma) \to \mathcal{T}(S)$$

$$g \mapsto g \cdot X$$

is a K-quasi-isometric embedding, where $\mathcal{T}(S)$ is given the Teichmüller metric. If $H < A(\Gamma)$ is convex cocompact as a subgroup of $\text{Mod}(S)$, then H is quasiconvex in $A(\Gamma)$.

Proof. Fix D, the quasiconvexity constant for $H \cdot X$ in $\mathcal{T}(S)$. Note that $N_D(H \cdot X)$ is ϵ-cobounded for some $\epsilon > 0$. This follows from the facts that X is ϵ'-thick and that, if $Y \in \mathcal{T}(S)$ with $d_{\mathcal{T}(S)}(X, Y) \leq D$, then for any simple closed curve γ, $l_Y(\gamma) \geq e^{-2D} l_X(\gamma) \geq e^{-2D} \epsilon'$.

Let $h, h' \in H$ and suppose α is a geodesic path in $A(\Gamma)$ from h to h'. Then $\phi(\alpha)$ is a K-quasigeodesic joining $h \cdot X$ and $h' \cdot X$. Let τ be the Teichmuller geodesic between these points. By quasiconvexity, $\tau \subset N_D(H \cdot X)$ and so τ is ϵ-cobounded.

By Theorem 6.1, there exists B depending only on ϵ and K so that any K-quasigeodesic with endpoints on τ is contained in a B-neighborhood of τ. Hence, $\phi(\alpha) \subset N_B(\tau) \subset N_{D+B}(H \cdot X)$. So if $p \in \alpha \subset A(\Gamma)$, then there is an $h \in H$ with

$$d_{\mathcal{T}(S)}(\phi(p), h \cdot X) \leq D + B$$

and so

$$\begin{aligned} d_{A(\Gamma)}(p, h) &\leq K(d_{\mathcal{T}(S)}(\phi(p), h \cdot X)) + K \\ &= K(D + B) + K \end{aligned}$$

This implies that $\alpha \subset N_{K(D+B)+K}(H)$, as required. $\square$

We can now complete the proof of Theorem 1.1.

Proof of Theorem 1.1. Let $\phi : A(\Gamma) \to \mathrm{Mod}(S)$ be an admissible homomorphism. First suppose that each element of $\phi(H)$ is pseudo-Anosov and that $H \leq G$ is quasiconvex. Then Corollary 5.3 states that $\phi(H)$ is convex cocompact. Conversely, if we assume that $\phi(H)$ is convex cocompact, then since the orbit map into $\mathcal{T}(S)$ induced by an admissible homomorphism is a quasi-isometric embedding, Theorem 6.2 implies that H is quasiconvex in $A(\Gamma)$. Since each infinite-order element of a convex cocompact subgroup in $\mathrm{Mod}(S)$ is psuedo-Anosov and $A(\Gamma)$ is torsion-free, we have that each nonidentity element of $\phi(H)$ is pseudo-Anosov. This completes the proof. $\qquad\square$

7. Some examples

In this section we produce some examples of convex cocompact subgroups of mapping class groups. We encourage the reader to observe the algorithmic nature of these examples, and that in each case quasiconvexity of the subgroup is proven by constructing the compact complex C.

Example 3. Take Γ with $E(\Gamma) = \emptyset$, so that $A(\Gamma) = \mathbb{F}(V(\Gamma))$, the free group on the vertex set of Γ. In this case, any finitely generated $H \leq A(\Gamma)$ is quasiconvex and C is the Stallings core graph for H, as explained in [Sta83]. Briefly, one begins with a rose R_H with rank(H) petals and subdivides and labels each petal with a generator of H written in terms on the standard generators of $A(\Gamma)$. This produces a map $R_H \to S_\Gamma = R_{|V(\Gamma)|}$ and iteratively folding the graph R_H produces a graph G_H and an immersion $G_H \to R_{|V(\Gamma)|}$ with image H at the fundamental group level. This is the desired C, since graph immersions are local isometries (when each edge is given unit length). Hence, when $A(\Gamma)$ is free, for any finitely generated $H < A(\Gamma)$ and any admissible homomorphism $\phi : A(\Gamma) \to \mathrm{Mod}(S)$, $\phi(H)$ is convex cocompact if and only if $\phi(H)$ is purely pseudo-Anosov.

Example 4. Take Γ as in Figure 1 with any admissible realization of $A(\Gamma)$ in $\mathrm{Mod}(S)$ fully supported on X_a, X_b, X_c. Then $A(\Gamma) = \langle a, b, c | [b, c] = 1 \rangle$. Consider $H = \langle bca, babc \rangle$. We construct the complex C described below. This will be the desired complex once we verify that the induced map $C \to S_\Gamma$ is a local isometry, as in Section 3.2.

To construct C, we begin as in the example above by building an immersion from a based graph into the 1-skeleton of S_Γ, see Figure 2. In this case, a single fold is required. This map, however, is far from a local isometry into S_Γ because of the absence of 2-cells. We add such 2-cells to the graph representing the commuting relations between b and c, first adding the squares between adjacent occurrences of b and c in the graph. The resulting complex admits an obvious cubical map into S_Γ. This, however, is not a local isometry as the image of the link at the base point is not a full subcomplex of $\mathrm{lk}(v, S_\Gamma)$. To rectify this, attach an additional square representing commuting relations between b and c, giving four 2-cells in total. Call the resulting complex C. We then verify that the induced map on the link of each vertex satisfies the conditions for $C \to S_\Gamma$ to be a local isometry. Hence, H is a

quasiconvex subgroup of $A(\Gamma)$. From this, we can readily check that each $h \in H$ is filling and so by Theorem 1.3, H is convex cocompact as a subgroup of the mapping class group.
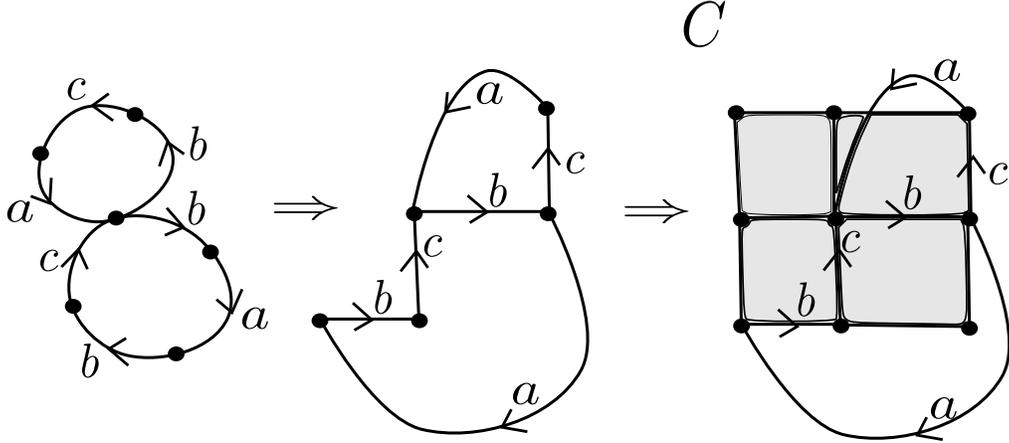

FIGURE 2. Constructing C

Example 5. We augment the example above by adding the element $b^2c^2a^2$ to the generating set of H. Note that inspection of the complex C reveals that this word is not already in the subgroup H. Call the resulting subgroup $H' = \langle abc, cba, a^2b^2c^2 \rangle$.

To build C', the complex with a local isometry to S_Γ with fundamental group H', we add to the complex C with a relative version of the construction above. This requires attaching a loop labeled $b^2c^2a^2$ to the base vertex, performing 2 folds, and adding 5 squares. Once this is done, it is again easily seen that C' has the required properties. Hence, H' is a convex cocompact subgroup of the mapping class group (through the homomorphism fixed above).

In the above examples, we used the existence of the complex C to verify that our subgroup was quasiconvex with all filling elements. The following example shows how attempting to construct C can also show that a particular group will not be purely pseudo-Anosov.

Example 6. As in the previous example, we add to the subgroup H without changing $A(\Gamma) \leq \mathrm{Mod}(S)$. Suppose we wish to add the generator bca^2 to obtain the subgroup $H'' = \langle abc, cab, a^2bc \rangle$. Attaching the necessary cells to the complex C, we see that some word in H'' is conjugate to a and hence not filling. Adding the word $b^{-1}ca^2$, however, will result in a quasiconvex subgroup all of whose elements are filling. Again, this is immediate by building the associated complex.

8. SHORT-TRANSLATION CONVEX COCOMPACT SUBGROUPS OF $\mathrm{Mod}(S)$

As an application of our main theorem, we construct convex cocompact subgroups of mapping class groups $\mathrm{Mod}(S_g)$ whose orbit maps into the curve complex have Lipschitz

constant on the order of $1/g$, where g denotes the genus of the surface. This is in contrast to other examples of convex cocompact subgroups occurring in the literature, where one does not have such control over the group constructed. This is clear in the first two constructions of convex cocompact subgroups of mapping class groups that are discussed in the introduction. It is also the case that, in the mapping class group of a once-punctured surface, a pseudo-Anosov in the image of the point-pushing map (the map $\pi_1(S_g) \to \mathrm{Mod}(\mathring{S}_g)$ discussed in Section 1) has minimum possible translation length equal to 1, independent of genus [Zha13].
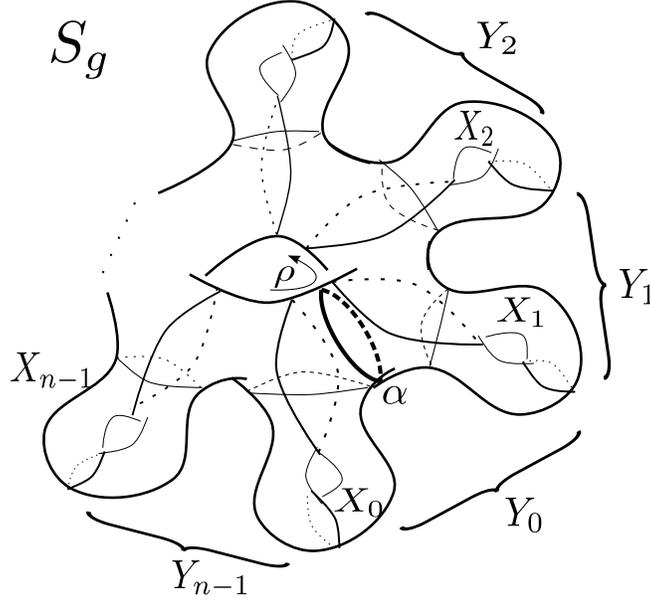

FIGURE 3. S_g with subsurfaces

Fix $n \geq 2$ and let S_g be a surface of genus $g = n + 1$. Referring to Figure 3, let f_0 be a mapping class fully supported on X_0, let g_0 be fully supported on Y_0, and let ρ be the order-n counterclockwise rotation shown (see Section 2.3 for the definition of fully supported). Here X_0 is a once-punctured torus and Y_0 is a four-times punctured sphere. We will require these maps to be pseudo-Anosov with sufficiently large translation length in the curve complex. Define
$$f_i = \rho^i f_0 \rho^{-i} \qquad \text{and} \qquad g_i = \rho^i g_0 \rho^{-i}.$$
By construction, f_i is fully supported on X_i and g_i on Y_i, where $X_i = \rho^i X_0$ and $Y_i = \rho^i Y_0$. Note that we have only defined finitely many distinct maps, and may consider the index i as an integer mod n.

Observe by disjointness of the subsurfaces involved, we have the following:

- f_i commutes with f_j for all $1 \leq i, j \leq n$,
- g_i commutes with g_j for all $1 \leq i, j \leq n$, and
- f_i commutes with g_j for $j \neq i - 1, i$.

Hence, if Γ is the coincidence graph for these subsurfaces, Γ^c is a cycle of length $2n$, alternating f_i and g_i in cyclic order. If the initial f_0, g_0 are chosen with large enough translation length in the curve complexes of X_0 and Y_0 respectively, where "large enough" is independent of n, then by Theorem 4.1 we have an admissible embedding

$$A(\Gamma) \to \mathrm{Mod}(S),$$

and so we identify $A(\Gamma)$ with its image under this homomorphism.

We now demonstrate a few calculations. Set $\phi = \rho g_0 f_0$ and $\psi = \phi^n$, so

$$
\begin{aligned}
\psi = \phi^n &= \rho g_0 f_0 \rho g_0 f_0 \ldots \rho g_0 f_0 \\
&= \rho g_0 f_0 \rho^{-1} \rho^2 g_0 f_0 \rho^{-2} \ldots \rho^n g_0 f_0 \\
&= g_1 f_1 g_2 f_2 \ldots g_n f_n \\
&= (g_1 g_2 \ldots g_{n-1})(f_1 g_n)(f_2 f_3 \ldots f_n).
\end{aligned}
$$

Let $w_i = (\rho g_0^i f_0^i)^n$, so that $w_1 = \psi$ and w_i replaces each g_j and f_j in ψ with its ith power:

$$w_i = (g_1^i g_2^i \ldots g_{n-1}^i)(f_1^i g_n^i)(f_2^i f_3^i \ldots f_n^i)$$

Let α be the separating curve in Y_0 disjoint from both X_0 and X_1, as shown in Figure 3.

We can now prove the following theorem, which gives examples of convex cocompact subgroups of the mapping class group whose generators have small translation length in the curve complex. We note that Theorem 8.1 immediately implies Theorem 1.4 from the introduction by taking H to be the subgroup of $\mathrm{Mod}(S_g)$ given in the statement of Theorem 8.1.

Theorem 8.1. *Let S be as above with genus $g = n+1$ for $n \geq 2$, and let $H = \langle w_1, \ldots w_N \rangle < A(\Gamma) < \mathrm{Mod}(S_g)$ as above, for any $N \geq 1$. Then H is a convex cocompact subgroup of $\mathrm{Mod}(S)$ and with α as above,*

$$d_{\mathcal{C}(S)}(\alpha, w_i^p \alpha) \leq 2,$$

for all $i = 1, \ldots, N$ and $p \leq \frac{n}{2}$. In particular, $\ell_S(w_i) \leq \frac{4}{g-1}$.

Proof. The statement about translation length follows from Theorem 8.5 below.

Proving convex cocompactness of H require additional notation as well as a few lemmas. Set for $1 \leq i \leq N$

$$
\begin{aligned}
B_i &= g_1^i \ldots g_{n-1}^i, \\
M_i &= f_1^i g_n^i, \\
E_i &= f_2^i \ldots f_n^i.
\end{aligned}
$$

Observe that $B_i^{-1} = B_{-i}$, $E_i^{-1} = E_{-i}$, but $M_i^{-1} \neq M_{-i}$. In what follows we do not allow negative indices for the M_i. Symbols without subscripts will be used when the subscript is not important (e.g. $B = B^{-1}$ but $M \neq M^{-1}$). In any case, the occurrence of such a symbol will aways denote a non-zero subscript (and power). The first lemma gives a procedure for assembling a *particular* normal form for any $h \in H$.

Lemma 8.2. *With notation as above, we have the following facts:*

(1) $w_i = B_i M_i E_i$
(2) $w_i w_j^{-1} = B_i M_i E_{i-j} M_j^{-1} B_j^{-1}$
   $w_i^{-1} w_j = E_i^{-1} M_i^{-1} B_{j-i} M_j E_j$
(3) $B_i M_j^{\pm 1} E_k$ *and its inverse are filling words where $j \geq 1$.*
(4) *Suppose that h is a word of $A(\Gamma)$ that is a concatenation of Bs, Ms, and Es so that adjacent pairs appear as BM, ME, EB or their inverses ($M^{-1}B$, EM^{-1}, or BE). Then h is in normal form with respect to the standard generators $f_1, \ldots, f_n, g_1, \ldots, g_n$ of $A(\Gamma)$.*

Proof. Fact 1 is immediate from the definitions, as is Fact 2 once it is recalled that all letters used in the spelling of B (and E) pairwise commute. Fact 3 follows since these words have the least number of syllables in their conjugacy classes and the supports of the generators used in their spelling fill S. Fact 4 requires some verification. First observe that any word with at most one occurrence of B, M and E is in normal form, since in this case every generator occurs in a syllable at most once. Suppose toward contradiction that h as in Fact 4 is not in normal form. Then there is a syllable that can be shuffled to the right using move (3) and combined with another syllable elsewhere in the word. This syllable must commute with every letter in between. By Fact 3 and the fact that each generator occurs only within a B, M or E such shuffling can only occur between consecutive occurrence of the same symbol B, M or E; otherwise it would have to commute with a filling word, an impossibility. Hence, the remaining cases are when the syllables that are to be combined occur between consecutive B, M or E expressions. For the B case, such a syllable is a power of g_j and must be commuted passed an entire expression E to combine with another power of g_j. This, however, is a contradiction since g_j fails to commute with both f_j and f_{j+1}, one of which occurs in E. The case where the syllable to be combined is contained in E is similar. The final case is when the syllable is contained in M or M^{-1} and is therefore either a power of f_1 or g_n. However, in M, f_1 cannot be commuted past g_n, and g_n cannot be commuted past f_n which occurs in the expression E. In either case, neither syllable can be shuffled right to combine with another syllable. In M^{-1}, g_n^{-1} cannot commute past f_1^{-1}, and f_1^{-1} cannot commute past g_1 which occurs in the expression B, so we have a similar contradiction. This completes the proof that no syllable of h can be combined with another syllable and shows that h is in normal form. $\square$

By Proposition 5.1, to prove that H is a convex cocompact subgroup of $\mathrm{Mod}(S_g)$ it suffices to show that there is an $\ell \geq 0$ so that for $h \in H$, any $w \in \mathrm{Min}(h)$ has the property that any subword of w of $A(\Gamma)$-length greater than ℓ contains a filling block.

First suppose that $h \in H$ is written in terms of the generators $w_1, \ldots, w_N$ with no consecutive occurrences of a generator and its inverse (i.e. h is written as a reduced word in the free group generated by the symbols $w_1, \ldots, w_N$). Now use Facts 1 and 2 above to transform h into it its BME spelling. We note that since h is reduced in the free group generated by the w_i, within the BME spelling no subscript is zero (see Fact 2) and so expressions E and B are always separated by a M or M^{-1}. To verify that this rewriting of h is in normal form, we use Fact 4 above. This is immediate from Fact 2 and the fact that no zero subscript

occurs in the rewriting. Denote the normal form representative of h so obtained by w^* and refer to it as the *BME normal form*. The next lemma shows that this normal form has the desired property.

Lemma 8.3. *For any $h \in H$, its BME normal form w^* has the property that every subword of w^* with $A(\Gamma)$-length at least $b = 3Nn + 4N$ has a filling block.*

Proof. Assume that r is a subword of w^* of $A(\Gamma)$-length at least $3Nn + 4N$. In this spelling, no more than the first Nn letters occur in r before the first full occurrence of B, M or E. Whichever symbol occurs first, we see that there is a terminal subword of r of length at least $2Nn + 4N$ that begins with either B, M or E. Checking the possible cases, we see that within r we have a subword w' which is a consecutive occurrence of BME up to cyclic permutations and inverses. By Fact 3 of the previous lemma, w' is a filling block. Thus, r contains a filling block as required. $\square$

Our final lemma needed for the proof gives a strong ordering property for the BME normal forms. Note that our choice of $A(\Gamma) \leq \mathrm{Mod}(S_g)$ is such that the support of every generator is required to fill the surface S_g.

Lemma 8.4. *There is an $L \geq 0$ so that if $h \in H$ has BME normal form w^* and s is a syllable of w^*, then any syllable separated from s in w^* by L or more syllables is ordered with s. Hence, no more than $2L$ syllables in h are unordered with $s \in \mathrm{syl}(h)$ in any normal form for h.*

Proof. Let Γ^c denote the complement graph of Γ. As previously observed, Γ^c is connected; let d denote its diameter. Take $L = d \cdot b$.

Now let t be a syllable of w^* so that at least L syllables occur between s and t in w^*. Let v and u be standard generators of $A(\Gamma)$ so that s is a power of v and t is a power of u. Then, for some $k \leq d$, there is a path $v = v_0, v_1, \ldots, v_k = u$ in Γ^c; by construction, v_i and v_{i+1} do not commute. Since $k \cdot b \leq d \cdot b$, we may select $w_1, w_2, \ldots, w_{k-1}$ a sequence of disjoint filling blocks between s and t occurring in the order as read from s to t. Being filling blocks, each w_i must include every generator, so for $1 \leq i < k$ we choose s_i to be a syllable in w_i that is a power of v_i. Then by construction

$$s \prec s_1 \prec \ldots \prec s_{k-1} \prec t.$$

Hence, by transitivity $s \prec t$. For the conclusion of the lemma, recall that the number of syllables unordered with a given s in $\mathrm{syl}(h)$ is independent of choice of normal form. $\square$

We can now conclude the proof of the theorem. Set $\ell' = b + 4LN + 1$ and let w be any normal form for $h \in H$. Towards contradiction, suppose that there is a subword r of w of length at least ℓ' that does not contain an occurrence of some generator; call this generator v. Enlarge this subword, continuing to denote it r, so that either the first or last syllable of r is a power of v and the other terminal syllable is either a terminal syllable of w or another syllable that is a power of v. This is possible since w itself is filling and so must contain at least one occurrence of a power of v.

Let w^* be the BME normal form for h; note that, by Lemma 8.3, there is no subword of w^* longer than b that fails to contain each generator of $A(\Gamma)$. Since $w, w^* \in \text{Min}(h)$, these words are related by a repeated application of the type (3) moves described in Section 3.1. There are two cases, depending on whether r contains one or two occurrences of v. If the first and last syllables of r are powers of v then denote these by s_1 and s_2. Since between these syllables in w there is no occurrence of v, the type (3) moves from w to w^* must commute $\ell' - b$ letters past either s_1 or s_2. Hence, at least $(\ell' - b)/2$ letters must be commuted past, say, s_1. Since no syllable contains more than N letters, at least $(\ell' - b)/(2N)$ syllables are commuted past s_1 in the process of applying type (3) moves. All of these syllables must be unordered with s_1. However, by Lemma 8.4 the number of syllables unordered with s_1 is bounded by $2L$. Hence,

$$\frac{\ell' - b}{2N} \leq 2L,$$

contradicting our choice of ℓ'. The case where r contains only one occurrence of the generator v, and therefore contains the first or last syllable of w, is similar and left to the reader.

Let $\ell = \ell' + 2N$. Given w' a subword of w of length at least ℓ, let r be the subword consisting of the middle ℓ' letters of w'. We have seen that r contains an occurrence of each generator. By Lemma 4.2, r is a filling block for w unless one or both of its terminal syllables are not complete syllables in w. Appending up to N letters to the left and the right of r gives a filling block for w contained in w'. $\qquad\square$

Let S_g be as above with genus $g = n + 1$ for $n \geq 2$, and let $H = \langle w_1, \ldots w_N \rangle \leq A(\Gamma) \leq \text{Mod}(S_g)$ as above, for any $N \geq 1$. For any element $h \in H$, let $|h|_H$ denote its word length with respect to the generating set $\{w_1, \ldots w_N\}$. Let α be the curve in Figure 3.

Theorem 8.5. *For any $h \in H$,*

$$d_{\mathcal{C}(S)}(\alpha, h\alpha) \leq |h|_H \cdot \frac{4}{g - 1} + 2$$

Proof. It suffices to prove the following lemma:

Lemma 8.6. *If $|h|_H \leq n/2$ then $d_{\mathcal{C}(S)}(\alpha, h\alpha) \leq 2$.*

This suffices because, where m is the largest integer less than $|h|_H \cdot (2/n) + 1$, we may write $h = h_1 \cdots h_m$ for some h_i such that $|h_i|_H \leq \frac{n}{2}$. Using the triangle inequality, we have

$$\begin{aligned} d_{\mathcal{C}(S)}(\alpha, h\alpha) &\leq d_{\mathcal{C}(S)}(\alpha, h_1\alpha) + d_{\mathcal{C}(S)}(h_1\alpha, h_1 h_2\alpha) + \cdots + d_{\mathcal{C}(S)}(h_1 h_2 \cdots h_{m-1}\alpha, h\alpha) \\ &= \sum_{i=1}^{m} d_{\mathcal{C}(S)}(\alpha, h_i\alpha) \leq 2m \leq |h|_H \cdot (4/n) + 2. \end{aligned}$$

To obtain the theorem, recall $g = n + 1$.

So we are tasked with proving the lemma above. We rely on an obvious principle: if S is connected and has a proper subsurface S' containing the two curves α and β, then $d_{\mathcal{C}(S)}(\alpha, \beta) \leq 2$, because α and β are each disjoint from the curves of $\partial S'$.

For a list of subsurfaces $Z_1, \ldots, Z_k$ of S, we will write $\mathsf{span}\{Z_1, \ldots, Z_k\}$ for the essential subsurface of S of minimal complexity that contains each of $Z_1, \ldots, Z_k$ (up to isotopy). In other words, $\mathsf{span}\{Z_1, \ldots, Z_k\}$ is the subsurface Z of S that contains each subsurface in the collection and has the property that any curve that projects nontrivially to Z also has nontrivial projection to some Z_i. If each Z_i in the collection equals either some X_s or Y_t, observe that $\mathsf{span}\{Z_1, \ldots, Z_k\}$ is a proper subsurface of S unless *all* of the subsurfaces $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ occur in the collection. Hence, the set of all X_s and Y_t fill S, but no proper subset of this collection of subsurfaces fills S.
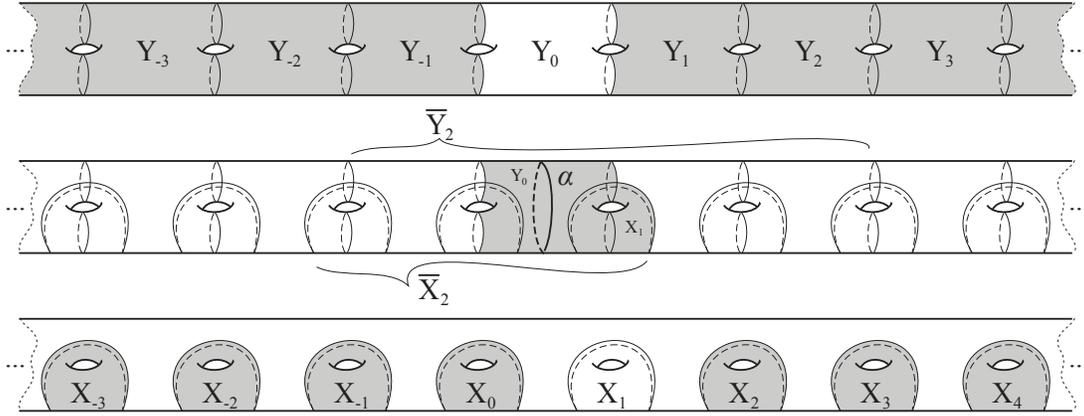

FIGURE 4. Three views of S_g. The supports of B, M, and E are shaded in the top, middle, and bottom copies respectively.

Below, we write w to represent any w_i and v to represent any w_i^{-1}. That is, any appearance of w or v satisfies $w = BME$ or $v = EM^{-1}B$ for some suitabe B, M, and E. It turns out H is a free group, although this fact is not necessary in order to obtain upper bounds on displacement. We think of a given h as a product of generators of type w or v. It is tedious but straightforward to verify by inspection the following statements:

- $\alpha, v\alpha \subset \mathsf{span}\{X_0, Y_0\}$ which is a proper subsurface of S (written "$\subsetneq S$") if $n \geq 2$
- $\alpha, w\alpha \subset \mathsf{span}\{Y_0, X_1, Y_1\} \subsetneq S$ if $n \geq 2$
- $\alpha, wv\alpha \subset \mathsf{span}\{Y_1, X_0, Y_0, X_1, Y_1\} \subsetneq S$ if $n \geq 3$
- $\alpha, vw\alpha \subset \mathsf{span}\{X_0, Y_0, X_1, Y_1, X_2\} \subsetneq S$ if $n \geq 3$
- $\alpha, w^2\alpha \subset \mathsf{span}\{Y_{-1}, X_0, Y_0, X_1, Y_1, X_2, Y_2\} \subsetneq S$ if $n \geq 4$
- $\alpha, v^2\alpha \subset \mathsf{span}\{X_{-1}, Y_{-1}, X_0, Y_0, X_1\} \subsetneq S$ if $n \geq 5$

These statements are sufficient to verify Lemma 8.6 if $n \leq 5$. They also set forth enough initial cases to complete the proof of Lemma 8.6 via induction.

For $k \geq 2$, define the subsurfaces:

$$\begin{aligned}
\overline{X}_k &= \operatorname{span}\{X_i, Y_j : -k < i < k \text{ and } -k < j < k-1\} \\
\overline{Y}_k &= \operatorname{span}\{Y_i, X_j : -k < i \leq k \text{ and } -k+1 < j \leq k\}
\end{aligned}$$

Using our comment above about how many subsurfaces X_i and Y_j are required to fill S, one can check that, for $k \leq n/2$, $\overline{X}_k$ and $\overline{Y}_k$ are proper subsurfaces of S. Thus we are done if we can establish the claim:

($\star$) If $|h| \leq k$ then $h\alpha$ is contained in either $\overline{X}_k$ or $\overline{Y}_k$.

The claim for $k = 2$ is an outcome of the definitions and the bulleted statements above. So it remains to assume the truth of ($\star$) for $k \geq 2$, and prove ($\star$) for $k + 1$.

Given h such that $|h| = k + 1$, either $h = wh'$ or $h = vh'$ where $|h'| = k$ satisfies ($\star$). We also know $h'\alpha$ is contained in either $\overline{X}_k$ or $\overline{Y}_k$. We analyze all possibilities below. Recall that $w = BME$ and $v = EM^{-1}B$ where B is supported entirely on $\cup Y_i$ and E is supported entirely on $\cup X_i$. Since $k \geq 2$ we are also assured that M is supported within $\overline{X}_k$ as well as within $\overline{Y}_k$. For the next set of statements—apparent by inspection—observe that, depending on the case, applying B or E may increase the support required to contain the image curve, while M has no effect. See Figure 4.

If $h'\alpha \subset \overline{X}_k = \operatorname{span}\{X_{-k+1}, Y_{-k+1}, \cdots, Y_{k-2}, X_{k-1}\}$,

$$\begin{aligned}
Bh'\alpha &\subset \operatorname{span}\{Y_{-k}, X_{-k+1}, Y_{-k+1}, \cdots, Y_{k-2}, X_{k-1}, Y_{k-1}\} \\
M^{-1}Bh'\alpha &\subset \operatorname{span}\{Y_{-k}, X_{-k+1}, Y_{-k+1}, \cdots, Y_{k-2}, X_{k-1}, Y_{k-1}\} \\
vh' = EM^{-1}Bh'\alpha &\subset \operatorname{span}\{X_{-k}, Y_{-k}, X_{-k+1}, Y_{-k+1}, \cdots, Y_{k-2}, X_{k-1}, Y_{k-1}, X_k\} = \overline{X}_{k+1} \\
Eh'\alpha &\subset \overline{X}_k \\
MEh'\alpha &\subset \overline{X}_k \\
wh' = BMEh'\alpha &\subset \operatorname{span}\{Y_{-k}, X_{-k+1}, Y_{-k+1}, \cdots, Y_{k-2}, X_{k-1}, Y_{k-1}\} \subset \overline{X}_{k+1}.
\end{aligned}$$

If $h'\alpha \subset \overline{Y}_k = \operatorname{span}\{Y_{-k+1}, X_{-k+2}, \cdots, X_k, Y_k\}$,

$$\begin{aligned}
Eh'\alpha &\subset \operatorname{span}\{X_{-k+1}, Y_{-k+1}, X_{-k+2}, \cdots, X_k, Y_k, X_{k+1}\} \\
MEh'\alpha &\subset \operatorname{span}\{X_{-k+1}, Y_{-k+1}, X_{-k+2}, \cdots, X_k, Y_k, X_{k+1}\} \\
wh' = BMEh'\alpha &\subset \operatorname{span}\{Y_{-k}, X_{-k+1}, Y_{-k+1}, X_{-k+2}, \cdots, X_k, Y_k, X_{k+1}, Y_{k+1}\} = \overline{Y}_{k+1} \\
Bh'\alpha &\subset \overline{Y}_k \\
M^{-1}Bh'\alpha &\subset \overline{Y}_k \\
vh' = EM^{-1}Bh'\alpha &\subset \operatorname{span}\{X_{-k+1}, Y_{-k+1}, X_{-k+2}, \cdots, X_k, Y_k, X_{k+1}\} \subset \overline{Y}_{k+1}.
\end{aligned}$$

$\square$

References

[BH09]   M.R. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*, vol. 319, Springer, 2009.

[BKMM12] Jason Behrstock, Bruce Kleiner, Yair Minsky, and Lee Mosher, *Geometry and rigidity of mapping class groups*, Geometry & Topology **16** (2012), 781–888.

[Cha07]   R. Charney, *An introduction to right-angled artin groups*, Geometriae Dedicata **125** (2007), no. 1, 141–158.

[CLM12]   Matt T. Clay, Christopher J. Leininger, and Johanna Mangahas, *The geometry of right-angled Artin subgroups of mapping class groups*, Groups Geom. Dyn. **6** (2012), no. 2, 249–278.

[DKRL12] S. Dowdall, IV Kent, P. Richard, and C.J. Leininger, *Pseudo-anosov subgroups of fibered 3-manifold groups*, arXiv preprint arXiv:1208.2495 (2012).

[FLP79]   A. Fathi, F. Laudenbach, and V. Poenaru, *Travaux de Thurston sur les surfaces*, Astérisque, vol. 66, Société Mathématique de France, Paris, 1979, Séminaire Orsay, With an English summary.

[FM02]   Benson Farb and Lee Mosher, *Convex cocompact subgroups of mapping class groups*, Geom. Topol. **6** (2002), 91–152 (electronic).

[FM12]   Benson Farb and Dan Margalit, *A primer on mapping class groups*, Princeton Mathematical Series, vol. 49, Princeton University Press, Princeton, NJ, 2012.

[GT11]   Vaibhav Gadre and Chia-Yen Tsai, *Minimal pseudo-Anosov translation lengths on the complex of curves*, Geom. Topol. **15** (2011), no. 3, 1297–1312.

[Hag08]   F. Haglund, *Finite index subgroups of graph products*, Geometriae Dedicata **135** (2008), no. 1, 167–209.

[Ham05]   Ursula Hamenstädt, *Word hyperbolic extensions of surface groups*, preprint (2005), arXiv:0807.4891v2.

[HM95]   S. Hermiller and J. Meier, *Algorithms and geometry for graph products of groups*, Journal of Algebra **171** (1995), no. 1, 230–257.

[HW08]   F. Haglund and D.T. Wise, *Special cube complexes*, Geometric and Functional Analysis **17** (2008), no. 5, 1551–1620.

[KL07]   Richard P. Kent, IV and Christopher J. Leininger, *Subgroups of mapping class groups from the geometrical viewpoint*, In the tradition of Ahlfors-Bers. IV, Contemp. Math., vol. 432, Amer. Math. Soc., Providence, RI, 2007, pp. 119–141.

[KL08]   ———, *Shadows of mapping class groups: capturing convex cocompactness*, Geom. Funct. Anal. **18** (2008), no. 4, 1270–1325.

[KLS09]   R.P. Kent, C.J. Leininger, and S. Schleimer, *Trees and mapping class groups*, Journal für die reine und angewandte Mathematik (Crelles Journal) **2009** (2009), no. 637, 1–21.

[Min96]   Y.N. Minsky, *Quasi-projections in teichmüller space*, Journal fur die Reine und Angewandte Mathematik (1996), 121–136.

[Min08]   H. Min, *Hyperbolic graphs of surface groups*, ProQuest, 2008.

[MM99]   Howard A. Masur and Yair N. Minsky, *Geometry of the complex of curves. I. Hyperbolicity*, Invent. Math. **138** (1999), no. 1, 103–149.

[MM00]   ———, *Geometry of the complex of curves. II. Hierarchical structure*, Geom. Funct. Anal. **10** (2000), no. 4, 902–974.

[Raf07]   Kasra Rafi, *A combinatorial model for the teichmüller metric*, GAFA Geometric And Functional Analysis **17** (2007), no. 3, 936–959.

[Sta83]   J.R. Stallings, *Topology of finite graphs*, Inventiones mathematicae **71** (1983), no. 3, 551–565.

[Tay13]   Samuel J Taylor, *Right-angled artin groups and out (f_n) i: quasi-isometric embeddings*, arXiv preprint arXiv:1303.6889 (2013).

[Zha13]   Chaohui Zhang, *On the minimum of asymptotic translation lengths of point-pushing pseudo-anosov maps on punctured riemann surfaces*, arXiv preprint arXiv:1301.4911 (2013).

Mathematics Department, Brown University, 151 Thayer Street, Providence, RI 02912, U.S.A.

E-mail address: mangahas@math.brown.edu

Department of Mathematics, University of Texas at Austin, 1 University Station C1200, Austin, TX 78712, U.S.A.

E-mail address: staylor@math.utexas.edu