Hierarchically hyperbolic spaces
I: Curve complexes for cubical groups

JASON BEHRSTOCK

MARK F HAGEN

ALESSANDRO SISTO

In the context of CAT(0) cubical groups, we develop an analogue of the theory of curve complexes and subsurface projections. The role of the subsurfaces is played by a collection of convex subcomplexes called a *factor system*, and the role of the curve graph is played by the *contact graph*. There are a number of close parallels between the contact graph and the curve graph, including hyperbolicity, acylindricity of the action, the existence of hierarchy paths, and a Masur–Minsky-style distance formula.

We then define a *hierarchically hyperbolic space*; the class of such spaces includes a wide class of cubical groups (including all virtually compact special groups) as well as mapping class groups and Teichmüller space with any of the standard metrics. We deduce a number of results about these spaces, all of which are new for cubical or mapping class groups, and most of which are new for both. We show that the quasi-Lipschitz image from a ball in a nilpotent Lie group into a hierarchically hyperbolic space lies close to a product of hierarchy geodesics. We also prove a rank theorem for hierarchically hyperbolic spaces; this generalizes results of Behrstock and Minsky, of Eskin, Masur and Rafi, of Hamenstädt, and of Kleiner. We finally prove that each hierarchically hyperbolic group admits an acylindrical action on a hyperbolic space. This acylindricity result is new for cubical groups, in which case the hyperbolic space admitting the action is the contact graph; in the case of the mapping class group, this provides a new proof of a theorem of Bowditch.

1 Introduction

Cube complexes and groups that act on them are fundamental objects in geometric group theory. Examples of cubical groups — groups acting geometrically on CAT(0) cube complexes — are right-angled (and many other) Coxeter groups (see Davis [27] and Niblo and Reeves [54]), right-angled Artin groups (see Charney and Davis [23]), and, more generally, graph products of abelian groups (see Ruane and Witzel [60]).

Other examples of cubical groups include: groups satisfying sufficiently strong small-cancellation conditions (see Wise [70]); many 3–manifold groups, including all hyperbolic ones (see Bergeron and Wise [9], Kahn and Markovic [42] and Wise [72]) and some graph manifold groups (see Hagen and Przytycki [34]); hyperbolic free-by-$\mathbb{Z}$ groups (see Hagen and Wise [35]); etc. Despite the attention cubical groups have attracted in recent years, their large-scale geometry has remained rather opaque, with a few salient exceptions, notably the resolution of the rank rigidity conjecture by Caprace and Sageev [21], characterizations of linear divergence, relative hyperbolicity, and thickness in terms of combinatorial data (see Behrstock and Charney [4], Behrstock and Hagen [5] and Hagen [32]), analysis of quasiflats in the 2–dimensional case (see Bestvina, Kleiner and Sageev [14]) and top-dimensional quasiflats in general (see Huang [39]).

Recently, there has been enormous progress in understanding the mapping class group and related spaces. Highlights have included resolutions of the ending lamination conjecture by Brock, Canary and Minsky [17], the rank conjecture by Behrstock and Minsky [8], quasi-isometric rigidity (see Behrstock, Kleiner, Minsky and Mosher [7]), finite asymptotic dimension (see Bestvina, Bromberg and Fujiwara [11]), and a number of others. Proofs of each of these results have featured the curve complex in a central position.

Motivated by the vital role the curve complex has played in unveiling the geometry of the mapping class group, in this work we develop analogues of those tools in the context of cubical groups. In particular, for cubical groups we develop versions of the machinery of curve complex projections and hierarchies initiated by Masur and Minsky [49; 50] as well as subsequent tools including the consistency and realization theorems of Behrstock [3] and Behrstock, Kleiner, Minsky and Mosher [7]. We note that right-angled Artin groups are a particularly interesting class of groups to which the tools we develop can be applied. Finally, we define *hierarchically hyperbolic spaces*, which provide a framework that encompasses mapping class groups, Teichmüller space, and a large class of cubical groups including the class of *compact special* groups of Haglund and Wise [37]. This allows us to prove new results in both the mapping class group and cubical contexts.

1.1 Geometry of contact graphs

In Part I we develop a number of basic aspects of the geometry of *contact graphs*, extending a study which was initiated by Hagen [33]. The contact graph $\mathcal{CX}$ of the CAT(0) cube complex $\mathcal{X}$ is the intersection graph of the *hyperplane carriers*; in other words, there is a vertex for each hyperplane of $\mathcal{X}$, and two vertices are adjacent if

the corresponding hyperplanes are not separated by a third. Since the contact graph is always hyperbolic (in fact, a quasi-tree) [33], it is a natural candidate for a "curve complex" in the context of cubical groups. The main results of Part I are summarized below. Recall that the WPD property, as defined by Bestvina and Fujiwara [13], is a form of properness of an action "in the direction" of particular elements; see Section 5 for the precise definition. It has important applications to bounded cohomology (see eg Bestvina, Bromberg and Fujiwara [10]), is closely related to the Bestvina–Bromberg–Fujiwara construction [11], and provides an equivalent characterization of acylindrical hyperbolicity (see Osin [55]) which is used in a number of applications.

Theorem A *Let $\mathcal{X}$ be a* CAT(0) *cube complex and $\mathcal{CX}$ its contact graph. Then:*

(1) **WPD property** *If $\mathcal{X}$ is uniformly locally finite and $g \in \mathrm{Aut}(\mathcal{X})$ is rank-one with the property that no positive power of g stabilizes a hyperplane, then g acts as a loxodromic WPD element on $\mathcal{CX}$.*

(2) **Hierarchy paths** *Let $x, y \in \mathcal{X}$ be 0–cubes. Then there exist hyperplanes $H_0, \ldots, H_k$ with $x \in \mathcal{N}(H_0)$, $y \in \mathcal{N}(H_k)$ and combinatorial geodesics $\gamma_i \to \mathcal{N}(H_i)$ such that $H_0, H_1, \ldots, H_k$ is a geodesic of $\mathcal{CX}$ and $\gamma_0 \gamma_1 \cdots \gamma_k$ is a geodesic joining x and y.*

(3) **Contractibility** *If the set of 1–cubes of $\mathcal{X}$ is countable and nonempty, then the flag complex spanned by $\mathcal{CX}$ is contractible.*

In the case of mapping class groups, analogues of Theorem A(1) and Theorem A(2) were proved by Bestvina and Fujiwara [13] and by Masur and Minsky [50]. The curve complex is not contractible, so Theorem A(3) provides a way in which the contact graph is simpler than the curve complex.

Theorem A has applications to random walks. In particular, from parts of (1)–(2) Theorem A and the main result of Mathieu and Sisto [51], when the nonelementary group $G < \mathrm{Aut}(\mathcal{X})$ contains a rank-one element, random paths in G stay close to geodesics with high probability. Further, this property has applications to various parameters associated with the random walk, including rate of escape and entropy.

1.2 Factor systems

The mapping class group, $\mathcal{MCG}(S)$, of a surface S is associated with the curve complex of S, together with the collection of curve complexes of subsurfaces of S; this association underlies the hierarchy machinery of Masur and Minsky [50]. Analogously, a CAT(0) cube complex $\mathcal{X}$ contains a profusion of convex subcomplexes, each of which is itself a CAT(0) cube complex, and the judicious choice of a family of convex

subcomplexes enables the creation of hierarchy machinery. The role of the collection of subsurfaces is played by a *factor system* $\mathfrak{F}$ in $\mathcal{X}$, which is a particular uniformly locally finite collection of convex subcomplexes. Table 1 summarizes the analogy between the mapping class group and a CAT(0) cube complex with a geometric group action and a factor system. We emphasize that although the tools and the results we obtain have parallels for mapping class groups the techniques that we employ are very different.

mapping class group $\mathcal{MCG}(S)$	cube complex $\mathcal{X}$ with factor system $\mathfrak{F}$ and G–action
curve complex $\mathcal{CS}$ is hyperbolic [49]	contact graph $\mathcal{CX}$ and factored contact graph $\widehat{\mathcal{CX}}$ are hyperbolic, indeed, are quasi-trees (Theorem 2.8, Proposition 8.25)
$\mathcal{MCG}(S)$ acts on $\mathcal{CS}$ acylindrically [15]	G acts on $\mathcal{CX}, \widehat{\mathcal{CX}}$ acylindrically (Corollary 14.5)
Nielsen–Thurston classification [67]	loxodromic/reducible/generalized reducible (Theorem 2.9)
$\exists$ quasi-geodesics in $\mathcal{MCG}(S)$ shadowing geodesics in $\mathcal{CS}$ [50]	$\exists$ geodesics in $\mathcal{X}$ shadowing geodesics in $\mathcal{CX}, \widehat{\mathcal{CX}}$ (Proposition 3.1, Proposition 8.23)
subsurfaces	subcomplexes in $\mathfrak{F}$ (Definition 8.1)
projections to subsurfaces [50]	projection to $\widehat{\mathcal{C}}F$ for $F \in \mathfrak{F}$ (Section 8.3)
formula computing distance in $\mathcal{MCG}(S)$ in terms of curve complex distances [50]	formula computing distance in $\mathcal{X}$ in terms of factored contact graph distances (Theorem 9.1)
bounded geodesic image [50]	bounded geodesic image (Proposition 8.20)
nested, disjoint, overlapping subsurfaces	parallel into, orthogonal, transverse elements of $\mathfrak{F}$
large link lemma [50]	large link lemma (Proposition 9.4)
consistency and realization [3; 7]	consistency and realization (Theorem 12.4)

Table 1

The collection of subcomplexes which constitute a factor system $\mathfrak{F}$ in $\mathcal{X}$ includes $\mathcal{X}$, as well as all combinatorial hyperplanes of $\mathcal{X}$; further, this collection is closed under the following operation: if $F, F' \in \mathfrak{F}$ and F has (combinatorial) projection onto F' of diameter more than some specified threshold, then the projection of F onto F' lies in $\mathfrak{F}$. This implies that sufficiently large hyperplanes of any codimension belong to any factor system, and indeed each factor system $\mathfrak{F}$ contains a minimal factor system consisting of $\mathcal{X}$, all combinatorial hyperplanes, and the closure of this family under the above

projection. This minimal factor system is $\mathrm{Aut}(\mathcal{X})$–invariant since each automorphism preserves the set of hyperplanes.

The reader should have in mind the following example, which already shows that the class of groups G acting geometrically on cube complexes with G–invariant factor systems is very large. Let Γ be a finite simplicial graph and let $\widetilde{S}_\Gamma$ be the universal cover of the Salvetti complex of the corresponding right-angled Artin group A_Γ, so that $\widetilde{S}_\Gamma$ is a CAT(0) cube complex on which A_Γ acts properly with a single orbit of 0–cubes; see Charney and Davis [23]. Each induced subgraph Λ of Γ yields a monomorphism $A_\Lambda \to A_\Gamma$ and an A_Λ–equivariant embedding $\widetilde{S}_\Lambda \hookrightarrow \widetilde{S}_\Gamma$. The set of all such subcomplexes of $\widetilde{S}_\Gamma$, and all of their A_Γ–translates, forms a factor system for $\widetilde{S}_\Gamma$, described in detail in Section 8.2. This, and the fact that the existence of a factor system is inherited by convex subcomplexes (Lemma 8.5), enables the study of groups that are *virtually special* in the sense of Haglund and Wise [37] using factor systems:

Proposition B *Let $\overline{\mathcal{X}}$ be a special cube complex with finitely many hyperplanes. Then the universal cover $\mathcal{X}$ of $\overline{\mathcal{X}}$ contains a factor system, and hence contains a factor system that is invariant under the action of $\pi_1 \overline{\mathcal{X}}$.*

In Corollary 8.8, for special cube complexes, we describe the factor system explicitly in terms of the hyperplanes of $\overline{\mathcal{X}}$. Proposition B also enables one to study many cubical groups which are far from being special: in Section 11, using Proposition 8.3 together with Burger and Mozes [19] and Wise [71] we show there exist many non-virtually special groups G which act geometrically on a CAT(0) cube complex $\mathcal{X}$ with a factor system. Moreover, we produce many examples which, unlike those of Burger and Mozes and of Wise, do not admit equivariant embeddings into products of trees; these will be used in Section 11.

Each $F \in \mathfrak{F}$ is a convex subcomplex, and is thus a CAT(0) cube complex whose hyperplanes have the form $H \cap F$, where H is a hyperplane of $\mathcal{X}$. This gives a natural injective graph homomorphism $\mathcal{C}F \to \mathcal{C}\mathcal{X}$, whose image is an induced subgraph [33]. Just as the elements of the factor system stand in analogy to the subsurfaces, the graphs $\mathcal{C}F$, where $F \in \mathfrak{F}$, essentially play the role of the curve complexes of the subsurfaces. In order to obtain Theorem 9.1 — our analogue of the Masur–Minsky distance formula — we must modify each $\mathcal{C}F$ slightly, by coning off each subgraph which is the contact graph of some smaller element of $\mathfrak{F}$. It is the resulting *factored contact graphs* $\widehat{\mathcal{C}}F$ that actually play the role of curve complexes. In Section 8.5, we show that factored contact graphs are all quasi-trees. Moreover, when $\mathfrak{F}$ is the minimal factor system described above, $\widehat{\mathcal{C}}\mathcal{X}$ and $\mathcal{C}\mathcal{X}$ are quasi-isometric.

In Section 9, we prove the following analogue for cubical groups of the celebrated Masur–Minsky distance formula [50, Theorem 6.12]. Their formula has become an

essential tool in studying the geometry of the mapping class group. Later, we will take the existence of such a formula as one of the characteristic features of a hierarchically hyperbolic space.

Theorem C (distance formula) *Let $\mathcal{X}$ be a* CAT(0) *cube complex with a factor system $\mathfrak{F}$. Let $\overline{\mathfrak{F}}$ contain exactly one representative of each parallelism class in $\mathfrak{F}$. Then there exists $s_0 \geq 0$ such that for all $s \geq s_0$, there are constants $K \geq 1$, $C \geq 0$ such that for all $x, y \in \mathcal{X}^{(0)}$,*

$$\mathsf{d}_\mathcal{X}(x, y) \asymp_{K,C} \sum_{F \in \overline{\mathfrak{F}}} \{\!\{\mathsf{d}_{\widehat{\mathcal{C}}F}(\pi_F(x), \pi_F(y))\}\!\}_s.$$

(Here, $\{\!\{A\}\!\}_s = A$ if $A \geq s$ and 0 otherwise. The notation $\asymp_{K,C}$ means "up to bounded multiplicative and additive error".)

In Theorem C, we use the notion of *parallelism*: two convex subcomplexes F, F' of $\mathcal{X}$ are *parallel* if, for all hyperplanes H, we have $H \cap F \neq \varnothing$ if and only if $H \cap F' \neq \varnothing$. Equivalently, F, F' are parallel if and only if $\mathcal{C}F, \mathcal{C}F'$ are the same subgraph of $\mathcal{C}\mathcal{X}$; parallel subcomplexes are isomorphic. Just as the Masur–Minsky distance formula involves summing over all curve complexes of subsurfaces, by identifying parallel elements of the factor system, our sum is over all factored contact graphs without repetition.

Another important property of the curve complex $\mathcal{C}S$ is that the action of $\mathcal{MCG}(S)$ on $\mathcal{C}S$ is *acylindrical*, by a result of Bowditch [15]. We obtain an analogous result for actions on (factored) contact graphs arising from actions on cube complexes, and in Section 14 we will show this holds in considerably greater generality. The statement in the cubical case is:

Theorem D (acylindrical hyperbolicity from factor systems) *Let the group G act properly and cocompactly on the* CAT(0) *cube complex $\mathcal{X}$ and suppose that $\mathcal{X}$ contains a factor system. Then the induced action of G on the contact graph $\mathcal{C}\mathcal{X}$ of $\mathcal{X}$ is acylindrical.*

Theorem D and the results of Maher and Tiozzo [46] combine to yield the following, which is related to work of Nevo and Sageev [53] on Poisson boundaries of cube complexes:

Corollary E (Poisson boundary) *Let the group G act properly and cocompactly on the* CAT(0) *cube complex $\mathcal{X}$ and suppose that $\mathcal{X}$ contains a factor system. Let μ be a probability distribution on G with finite entropy whose support generates a nonelementary group acting on $\mathcal{C}\mathcal{X}$ and let ν be the hitting measure on $\partial\mathcal{C}\mathcal{X}$. Then $(\partial\mathcal{C}\mathcal{X}, \nu)$ is isomorphic to the Poisson boundary of (G, μ).*

Theorem D also allows one to produce free subgroups of G freely generated by finite collections of high powers of elements, each of which acts loxodromically on $\mathcal{CX}$ (Corollary 14.6).

Using Theorem C, together with the tools in Bestvina, Bromberg and Fujiwara [11], the fact that factored contact graphs are quasi-trees, and the machinery we develop in Section 10, we prove:

Theorem F *Let G act properly and cocompactly on the CAT(0) cube complex $\mathcal{X}$ and suppose that $\mathcal{X}$ contains a G–invariant factor system. If the action of G is **hereditarily flip-free**, then G quasi-isometrically embeds in the product of finitely many quasi-trees.*

Moreover, there exist such $G, \mathcal{X}$ such that, for all finite-index subgroups $G' \leq G$, there is no G'–equivariant isometric embedding of $\mathcal{X}$ in the product of finitely many simplicial trees.

Hereditary flip-freeness is a mild technical condition on the factor system $\mathfrak{F}$ which holds for most of the examples we have discussed. For the second assertion of the theorem, we exploit the existence of cocompactly cubulated groups with no finite quotients [19; 71]. These groups are lattices in products of trees stabilizing the factors; this property gives rise to a factor system. The space for this example is assembled from these pieces in such a way that the existence of a factor system persists, but there is no longer a finite equivariant coloring of the hyperplanes with intersecting hyperplanes colored differently. This lack of a coloring precludes the existence of an isometric embedding in a product of finitely many trees.

1.3 Comparison to the theory of the extension graph of a right-angled Artin group

In the special case where $\mathcal{X} = \widetilde{S}_\Gamma$ is the universal cover of the Salvetti complex S_Γ of a right-angled Artin group A_Γ, the machinery of factor systems and contact graphs is not the first attempt to define an analogue of the curve complex and the attendant techniques. Kim and Koberda [43] introduced the *extension graph* Γ^e associated to the finite simplicial graph Γ (and thus to A_Γ). This graph has a vertex for each conjugate of each standard generator of A_Γ (ie vertex of Γ), with adjacency recording commutation. In the same paper it is shown that, like $\mathcal{C}\widetilde{S}_\Gamma$, the extension graph is always quasi-isometric to a tree, and in Kim and Koberda [44], the analogy between $\mathcal{MCG}(S)$, with its action on $\mathcal{CS}$, and A_Γ, with its action on Γ^e, is extensively developed: it is shown, for instance, that this action is acylindrical and obeys a loxodromic-elliptic dichotomy. It is observed in [44] that, except in exceptional cases, there is a surjective graph homomorphism $\mathcal{C}\widetilde{S}_\Gamma \to \Gamma^e$, where $\widetilde{S}_\Gamma$ is the universal cover of the Salvetti

complex of A_Γ, which is also a quasi-isometry, so many such *geometric* results about the action of A_Γ on Γ^e can be deduced from the results of the present paper about the action of A_Γ on $\mathcal{C}\widetilde{S}_\Gamma$. It should be strongly emphasized that the papers [44; 43] also explore interesting and less purely geometric issues, particular to right-angled Artin groups, that cannot be treated with factor system tools.

The authors of [44] also set up some version of hierarchy machinery, with the role of subsurfaces being played by subgroups of A_Γ of the form $A_{\mathrm{Lk}(v)}$, where v is a vertex of Γ, and their conjugates. When Γ has girth at least 5, they obtain a distance formula [44, Proposition 65], but, as they note, the formula they give has significant differences with the Masur–Minsky distance formula for the mapping class group. For example, the sum is taken over specified projections, which depend on the points whose distance is being estimated, rather than over all projections. Another significant distinction is that their distance formula does not measure distance in the right-angled Artin group A_Γ, but rather it measures the *syllable length* in that space (although not a perfect analogy: their metric is more similar to the Weil–Petersson metric on Teichmüller space than to the word metric on the mapping class group). The extension graph seems unable to capture distance in the right-angled Artin group via a hierarchical construction, since the extension graph is bounded when $A_\Gamma = \mathbb{Z}$. In the present paper, the geometric viewpoint afforded by factor systems, and in particular the existence of hierarchy paths (Proposition 8.23) and a large link lemma (Proposition 8.20) allows us to overcome these issues.

1.4 Hierarchically hyperbolic spaces

Our aim in the last part of the paper is to develop a unified framework to study mapping class group and CAT(0) cube complexes from a common perspective. To this end, we axiomatize the machinery of factored contact graphs/curve complexes, distance formula, etc, to obtain the definition of a *hierarchically hyperbolic space*, which is formally stated in Definition 13.1. This notion includes the two classes of groups just mentioned and allows one to prove new results for both of these classes simultaneously. Hierarchically hyperbolic spaces come with a notion of *complexity*: complexity 0 corresponds to bounded spaces, infinite-diameter δ–hyperbolic spaces have complexity 1, and higher-complexity hierarchically hyperbolic spaces coarsely contain direct products.

Roughly, a space $\mathcal{X}$ is hierarchically hyperbolic if $\mathcal{X}$ can be equipped with a set $\mathfrak{S}$ of uniformly Gromov-hyperbolic spaces, and projections $\mathcal{X} \to W$, with $W \in \mathfrak{S}$. These projections are required to satisfy various properties reminiscent of those satisfied by subsurface projections in the mapping class group case and projections to factored contact graphs in the case of CAT(0) cube complexes with factor systems. Hence

a space $\mathcal{X}$ may be hierarchically hyperbolic in multiple ways, ie with respect to projections to distinct families of hyperbolic spaces.

Remark (HHS is a QI-invariant property)   It is easily seen from Definition 13.1 that, if $\mathcal{X}$ is hierarchically hyperbolic by virtue of its projections to a set $\mathfrak{S}$ of hyperbolic spaces, and $\mathcal{Y} \to \mathcal{X}$ is a quasi-isometry, then we can compose each projection with the quasi-isometry and conclude that $\mathcal{Y}$ is hierarchically hyperbolic with respect to the same set $\mathfrak{S}$.

The motivating examples of hierarchically hyperbolic spaces are as follows:

Theorem G   (hierarchically hyperbolic spaces)

(1) A CAT(0) *cube complex with a factor system $\mathfrak{F}$ is hierarchically hyperbolic with respect to the set of factored contact graphs $\widehat{\mathcal{C}}W$, with $W \in \mathfrak{F}$. (This is summarized in* Remark 13.2.)

(2) *Let S be a connected, oriented hyperbolic surface of finite type. Then $\mathcal{MCG}(S)$ is hierarchically hyperbolic with respect to the collection of curve complexes of subsurfaces of S* (Behrstock [3], Behrstock, Kleiner, Minsky and Mosher [7] and Masur and Minsky [49; 50]).

(3) *Teichmüller space $\mathcal{T}(S)$ with the Weil–Petersson metric is hierarchically hyperbolic with respect to curve complexes of nonannular subsurfaces of S* (Behrstock [3], Behrstock, Kleiner, Minsky and Mosher [7], Brock [16] and Masur and Minsky [49; 50]).

(4) $\mathcal{T}(S)$ *with the Teichmüller metric is hierarchically hyperbolic with respect to curve complexes of nonannular subsurfaces and combinatorial horoballs associated to annuli* (Durham [29], Eskin, Masur and Rafi [30], Masur and Minsky [49] and Rafi [59]).

In a forthcoming paper we will show that fundamental groups of nongeometric 3–manifolds are also hierarchically hyperbolic, and that a metric space that is hyperbolic relative to a collection of hierarchically hyperbolic subspaces is hierarchically hyperbolic; see Behrstock, Hagen and Sisto [6]. We note that it is already known that relatively hyperbolic groups admit a distance formula; see Sisto [65]. Another interesting question is whether a right-angled Artin group endowed with the syllable length metric is a hierarchically hyperbolic space.

In Section 13, after defining hierarchically hyperbolic spaces, we study quasi-Lipschitz maps from balls in $\mathbb{R}^n$, and more general nilpotent Lie groups, into these spaces. The next three results will all follow directly from our Theorem 13.11, which provides a single unifying statement in terms of asymptotic cones. Our first result is a generalization

of a result of Eskin, Masur and Rafi [30, Theorem A] which is about mapping class groups and Teichmüller spaces.

Theorem H (quasi-boxes in hierarchically hyperbolic spaces) *Let $\mathcal{X}$ be a hierarchically hyperbolic space. Then for every $n \in \mathbb{N}$ and every K, C, R_0, ϵ_0 the following holds. There exists R_1 so that for any ball $B \subseteq \mathbb{R}^n$ of radius at least R_1 and $f: B \to \mathcal{X}$ a (K, C)–quasi-Lipschitz map, there is a ball $B' \subseteq B$ of radius $R' \geq R_0$ such that $f(B')$ lies inside the $\epsilon_0 R'$–neighborhood of a standard box.*

In Theorem H, we do not require B, B' to be centered at the same point in $\mathbb{R}^n$.

Our proof uses methods different from those used in [30]. Our approach is much shorter and does not rely on partitions of the set of subsurfaces (or an analogue thereof), which plays an important role in their proof. In particular, we do not rely on the results from [11]; this is one reason why our results can be applied in the case of CAT(0) cube complexes where the techniques of [30] would fail. However, our approach and theirs share some commonalities; for instance, we use Rademacher's theorem (applied to maps that arise at the level of asymptotic cones), while in [30], the authors use a coarse differentiation result.

Using a generalization of Rademacher's theorem due to Pansu, we consider the case of quasi-Lipschitz maps from more general nilpotent Lie groups.

Theorem I (restriction on nilpotent groups in hierarchically hyperbolic spaces) *Let $\mathcal{X}$ be a hierarchically hyperbolic space. Then for every simply connected nilpotent Lie group $\mathcal{N}$ with a left-invariant Riemannian metric and every K, C, there exists R with the following property. For every (K, C)–quasi-Lipschitz map $f: B \to \mathcal{X}$ from a ball in $\mathcal{N}$ into $\mathcal{X}$ and for every $n \in \mathcal{N}$ we have $\operatorname{diam}(f(B \cap n[\mathcal{N}, \mathcal{N}])) \leq R$. In particular, if a finitely generated nilpotent group admits a quasi-isometric embedding into $\mathcal{X}$ then it is virtually abelian.*

The final conclusion of Theorem I is known in the case where $\mathcal{X}$ is a CAT(0) space; see Pauls [58]. Although it does not appear to be in the literature, the conclusion of Theorem I for $\mathcal{MCG}$ can be alternatively proved using results of Hume [40] and Pauls [58].

The following theorem generalizes the rank theorems from Behrstock and Minsky [8], Eskin, Masur and Rafi [30], Hamenstädt [38] and Kleiner [45]:

Theorem J (rank) *Let $\mathcal{X}$ be a hierarchically hyperbolic space with respect to a set $\mathfrak{S}$. If there exists a quasi-isometric embedding $\mathbb{R}^n \to \mathcal{X}$ then n is at most the maximal cardinality of a set of pairwise-orthogonal elements of $\mathfrak{S}$ and, in particular, at most the complexity of $\mathcal{X}$.*

When $\mathcal{X}$ is a CAT(0) cube complex with a factor system and $\mathrm{Aut}(\mathcal{X})$ acts cocompactly, then such a space naturally has two hierarchically hyperbolic structures. One of these structures has hierarchy paths that are combinatorial geodesics and one with CAT(0) geodesics; the first is obtained explicitly in Section 8.3 and the existence of the latter follows from the first via a simple argument about projections of CAT(0) geodesics and convex hulls of ℓ_1 geodesics to factored contact graphs. By Theorem H and cocompactness, the existence of a quasi-isometric embedding $\mathbb{R}^n \to \mathcal{X}$ then implies that $\mathcal{X}$ contains both an ℓ_1–isometrically embedded copy of $\mathbb{R}^n$ with the standard tiling and, in the CAT(0) metric, an isometrically embedded flat of dimension n. We thus recover the cubical version of a theorem of Kleiner; see [45, Theorem C]. We also note that, in the special case of top-dimensional quasiflats in CAT(0) cube complexes, Huang has very recently proved a stronger statement [39, Theorem 1.1].

Finally, we relate hierarchically hyperbolic groups to acylindrically hyperbolic groups, as studied by Osin [55]. Although natural, the definition of an *automorphism* of a hierarchically hyperbolic space is technical, but includes, in the relevant cases, all elements of $\mathcal{MCG}$ and all isometries of a cube complex with a factor system. The "maximal element" S and $\widehat{C}S$ referred to below are $\mathcal{X}$ and its factored contact graph when $\mathcal{X}$ is a CAT(0) cube complex, while they are the surface S and its curve complex when G is the mapping class group of S.

Theorem K *Let $\mathcal{X}$ be hierarchically hyperbolic with respect to the set $\mathfrak{S}$ of hyperbolic spaces and let $G \leq \mathrm{Aut}(\mathfrak{S})$ act properly and cocompactly on $\mathcal{X}$. Let S be the maximal element of $\mathfrak{S}$ and denote by $\widehat{C}S$ the corresponding hyperbolic space. Then G acts acylindrically on $\widehat{C}S$.*

In the case of the $\mathcal{MCG}$, which acts acylindrically on the curve complex, our argument provides a new proof of a result of Bowditch [15]. We note that our proof is substantially different and is also rather short (although it relies on some amount of machinery).

Throughout the section on hierarchically hyperbolic spaces, we make use of a notion of "gate", which, in addition to the uses in this paper, we believe will be very useful for other applications as well. This notion simultaneously generalizes gates/projections on (certain) convex subspaces in the cubical context and coarsely Lipschitz retractions of the marking complex on its natural subspaces associated with subsurfaces. Nonetheless, the definition we give exploits a different point of view, which turns out to be very convenient and allows us to give different (and more concise) proofs than previously existing ones for a number of results about mapping class groups and CAT(0) cube complexes.

Acknowledgements

2 Background

We assume that the reader is familiar with basic properties of CAT(0) cube complexes, hyperplanes, median graphs, etc, and refer the reader to, for example [2; 24; 25; 33; 32; 36; 72; 73] for background on these concepts as they are used in this paper. Nonetheless, in Section 2.1 and Section 2.2, we will review cubical convexity and make explicit the notions of *gate*, *projection* and *parallelism*, since they play a fundamental role, and recall several useful facts about the contact graph. In Section 2.3, we briefly discuss background on right-angled Artin groups. Since, in Section 13, we will use asymptotic cones, we refer the reader to [28] for the relevant background.

2.1 Convex subcomplexes, combinatorial hyperplanes, gates and parallelism

Throughout, $\mathcal{X}$ is a CAT(0) cube complex and $\mathcal{H}$ is the set of hyperplanes. Unless stated otherwise, we work in the 1–skeleton of $\mathcal{X}$, and we denote by $d_\mathcal{X}$ the graph metric on $\mathcal{X}^{(1)}$. The *contact graph* $\mathcal{CX}$, defined in [33], is the graph whose vertex set is $\mathcal{H}$, with two vertices adjacent if the corresponding hyperplanes are not separated by a third.

The subcomplex $K \subset \mathcal{X}$ is *full* if K contains each cube of $\mathcal{X}$ whose 1–skeleton appears in K. A subcomplex $K \subset \mathcal{X}$ is *isometrically embedded* if K is full and $K \cap \bigcap_i H_i$ is connected for all $\{H_i\} \subset \mathcal{H}$. In this case, we say that $H \in \mathcal{H}$ *crosses* K when $K \cap H \neq \varnothing$. The term "isometrically embedded" is justified by the well-known fact (see eg [33]) that, if $K \subset \mathcal{X}$ is isometrically embedded in this sense, then $K^{(1)} \hookrightarrow \mathcal{X}^{(1)}$ is an isometric embedding with respect to graph metrics. The isometrically embedded subcomplex $K \subset \mathcal{X}$ is *convex* if any of the following equivalent conditions is met:

(1) K coincides with the intersection of all combinatorial halfspaces (see below) containing K.

(2) Let $x, y, z \in \mathcal{X}$ be 0–cubes with $x, y \in K$. Then the median of x, y, z lies in K.

(3) $K^{(1)}$ contains every geodesic of $\mathcal{X}^{(1)}$ whose endpoints lie in $K^{(1)}$.

(4) Let c be an n–cube of $\mathcal{X}$, with $n \geq 2$. Suppose that c has n codimension-1 faces that lie in K. Then $c \subset K$.

(5) The inclusion $K \to \mathcal{X}$ is a local isometry.

(Recall that a combinatorial map $\phi\colon K \to \mathcal{X}$ of cube complexes, with $\mathcal{X}$ non-positively-curved, is a *local isometry* if ϕ is locally injective and, for each $x \in K^{(0)}$, the map induced by ϕ on the link of x is injective and has image a full subcomplex of the link of $\phi(x)$; see [37; 72; 73] for more on local isometries and local convexity.)

A convex subcomplex $K \subseteq \mathcal{X}$ is itself a CAT(0) cube complex, whose hyperplanes are the subspaces of the form $H \cap K$, where $H \in \mathcal{H}$. A useful mantra, following from the definition of convexity, is: "if K is convex, then any two hyperplanes that cross K and cross each other must cross each other inside of K". The *convex hull* of $Y \subseteq \mathcal{X}$ is the intersection of all convex subcomplexes containing Y.

It follows immediately from the definition that if $K \subseteq \mathcal{X}$ is convex, then for all $x \in \mathcal{X}^{(0)}$ there exists a unique closest 0–cube $\mathfrak{g}_K(x)$, called the *gate* of x in K. The gate is characterized by the property that $H \in \mathcal{H}$ separates $\mathfrak{g}_K(x)$ from x if and only if H separates x from K.

As discussed in [33], it follows from the definition of convexity that the inclusion $K \hookrightarrow \mathcal{X}$ induces an injective graph homomorphism $\mathcal{C}K \to \mathcal{C}\mathcal{X}$ whose image is a full subgraph: just send each $H \cap K$ to H. This allows us to define a projection $\mathfrak{g}_K\colon \mathcal{X} \to K$. The map $x \mapsto \mathfrak{g}_K(x)$ extends to a cubical map $\mathfrak{g}_K\colon \mathcal{X} \to K$ as follows. Let c be a cube of $\mathcal{X}$ and let $H_1, \ldots, H_d$ be the collection of pairwise-crossing hyperplanes crossing c. Suppose that these are labeled so that $H_1, \ldots, H_s$ cross K, for some $0 \leq s \leq d$, and suppose that $H_{s+1}, \ldots, H_d$ do not cross K. Then the 0–cubes of c map by $\mathfrak{g}_K$ to the 0–cubes of a uniquely determined cube $\mathfrak{g}_K(c)$ of K in which the hyperplanes $H_1, \ldots, H_s$ intersect, and there is a cubical collapsing map $c \cong [-1, 1]^d \to [-1, 1]^s \cong \mathfrak{g}_K(c)$ extending the gate map on the 0–skeleton. This map is easily seen to be compatible with the face relation for cubes, yielding the desired cubical map $\mathfrak{g}_K\colon \mathcal{X} \to K$, whose salient property is that for all $x \in \mathcal{X}$, a hyperplane H separates x from K if and only if H separates x from $\mathfrak{g}_K(x)$. The next lemma follows easily from the definitions and is used freely throughout this paper:

Lemma 2.1 Let $\mathcal{X}$ be a CAT(0) cube complex, let $A \subseteq B \subseteq \mathcal{X}$ be convex subcomplexes, and let $x, y \in \mathcal{X}$ be 0–cubes. Then any hyperplane separating $\mathfrak{g}_A(x), \mathfrak{g}_A(y)$ separates $\mathfrak{g}_B(x), \mathfrak{g}_B(y)$ (and hence separates x, y).

Definition 2.2 (parallel) Convex subcomplexes $F, F' \subseteq \mathcal{X}$ are *parallel* if for each hyperplane H, we have $H \cap F \neq \varnothing$ if and only if $H \cap F' \neq \varnothing$. Parallelism is clearly an equivalence relation.

Hyperplanes lead to important examples of parallel subcomplexes. For each hyperplane $H \in \mathcal{H}$, let $\mathcal{N}(H)$ denote its *carrier*, ie the union of all closed cubes intersecting H. Then there is a cubical isometric embedding $H \times [-1, 1] \cong \mathcal{N}(H) \hookrightarrow \mathcal{X}$, and we denote by $H^{\pm}$ the images of $H \times \{\pm 1\}$. These convex subcomplexes are *combinatorial hyperplanes*. Observe that H^+ and H^- are parallel: a hyperplane crosses $H^{\pm}$ if and only if it crosses H. (Very occasionally, it will be convenient to refer to *combinatorial halfspaces*; a combinatorial halfspace associated to $H \in \mathcal{H}$ is a component of $\mathcal{X} - H \times (-1, 1)$. A component of the boundary of such a combinatorial halfspace is one of H^+ or H^-.)

Remark 2.3 (parallelism and dual cube complexes) The reader accustomed to thinking of cube complexes using Sageev's construction of the cube complex dual to a space with walls [62; 24] might appreciate the following characterization of parallelism: the convex subcomplex F can be viewed as the dual to a wallspace whose underlying set is $\mathcal{X}^{(0)}$ and whose walls are the hyperplanes in the vertex set of $\mathcal{C}F$; one must check that in this case, the restriction quotient $\mathcal{X} \to F$ is split by a cubical isometric embedding $F \to \mathcal{X}$. In general, the splitting need not be unique, depending on a choice of basepoint, and the images of the various embeddings are exactly the representatives of the parallelism class of F.

Observe that F, F' are parallel if and only if their contact graphs are the same subgraph of $\mathcal{C}\mathcal{X}$. Parallel subcomplexes are isomorphic and in fact, the following stronger statement holds, and we shall use it throughout the paper.

Lemma 2.4 Let $F, F' \subseteq \mathcal{X}$ be convex subcomplexes. The following are equivalent:

(1) F and F' are parallel;

(2) there is a cubical isometric embedding $F \times [0, a] \to \mathcal{X}$ whose restrictions to $F \times \{0\}$ and $F \times \{a\}$ factor as $F \times \{0\} \cong F \hookrightarrow \mathcal{X}$ and $F \times \{a\} \cong F' \hookrightarrow \mathcal{X}$, respectively, and $[0, a]$ is a combinatorial geodesic segment crossing exactly those hyperplanes that separate F from F'.

Hence there exists a convex subcomplex E_F such that there is a cubical embedding $F \times E_F$ with convex image such that for each F' in the parallelism class of F, there exists a 0–cube $e \in E_F$ such that $F \times \{e\} \to \mathcal{X}$ factors as $F \times \{e\} \xrightarrow{\mathrm{id}} F' \hookrightarrow F$.

Proof Let F, F' be parallel, let $\mathcal{H}(F)$ be the set of hyperplanes crossing F (hence $\mathcal{H}(F') = \mathcal{H}(F)$) and let $\mathcal{S}$ be the finite set of hyperplanes separating F from F'. Let E be the cubical convex hull of $F \cup F'$. Then every element of $\mathcal{H}(F)$ crosses E, and the same is true of $\mathcal{S}$, since any geodesic starting on F and ending on F' must contain a 1–cube dual to each element of $\mathcal{S}$. Conversely, let $H \in \mathcal{H}$ cross the hull of $F \cup F'$ and suppose that $H \notin \mathcal{H}(F) \cup \mathcal{S}$. Then either H crosses F but not F', which is impossible, or F and F' lie in the same halfspace $\overleftarrow{H}$ associated to H. But then $\overleftarrow{H}$ contains a combinatorial halfspace $\overleftarrow{H}^*$ that contains F, F', so $E \cap \overleftarrow{H}^*$ is a convex subcomplex containing $F \cup F'$ and properly contained in E (since H does not cross $E \cap \overleftarrow{H}^*$). This contradicts that E is the convex hull. Hence the set of hyperplanes crossing E is precisely $\mathcal{H}(F) \cup \mathcal{S}$. Each hyperplane in $\mathcal{H}(F)$ crosses each hyperplane in $\mathcal{S}$, and it follows from [21, Proposition 2.5] that $E \cong F \times I$ for some convex subcomplex I such that the set of hyperplanes crossing I is precisely $\mathcal{S}$. It is easily verified that I is the convex hull of a geodesic segment. This proves (1) $\implies$ (2); the other direction is obvious.

The "hence" assertion follows from an identical argument, once $\mathcal{S}$ is taken to be the set of hyperplanes H with the property that for some F', F'' in the parallelism class of F, we have that H separates F from F'. $\qquad \square$

Lemma 2.5 *For each convex subcomplex $F \subseteq \mathcal{X}$, either F is unique in its parallelism class, or F is contained in a combinatorial hyperplane.*

Proof If F is parallel to some $F' \neq F$, then as discussed above, there is a cubical isometric embedding $F \times [0, a] \to \mathcal{X}$, with $a \geq 1$, whose restriction to $F \times \{0\}$ is the inclusion $F \hookrightarrow \mathcal{X}$. Cubical isometric embeddings take combinatorial hyperplanes to subcomplexes of combinatorial hyperplanes, and $F \times \{0\}$ is a combinatorial hyperplane of $F \times [0, a]$ since $a \geq 1$, so the claim follows. $\qquad \square$

The following will be used often starting in Section 8.1.

Lemma 2.6 *If F, F' are convex subcomplexes, then $\mathfrak{g}_F(F')$ and $\mathfrak{g}_{F'}(F)$ are parallel subcomplexes. Moreover, if $F \cap F' \neq \varnothing$, then $\mathfrak{g}_F(F') = \mathfrak{g}_{F'}(F) = F \cap F'$.*

Proof Suppose that H is a hyperplane crossing $\mathfrak{g}_F(F')$. Then H separates $\mathfrak{g}_F(x)$ and $\mathfrak{g}_F(y)$ for some $x, y \in F'$. By Lemma 2.1, H separates x, y and hence crosses F'. Thus H crosses both F and F'. Conversely, suppose that a hyperplane H crosses F and F', separating $x, y \in F'$. Then H cannot separate x from $\mathfrak{g}_F(x)$ or y from $\mathfrak{g}_F(y)$, so H separates $\mathfrak{g}_F(x)$ and $\mathfrak{g}_F(y)$, and in particular crosses $\mathfrak{g}_F(F')$. Thus the set of hyperplanes crossing $\mathfrak{g}_F(F')$ is precisely the set of hyperplanes H that

cross both F and F'. Similarly, the set of hyperplanes crossing $\mathfrak{g}_{F'}(F)$ is the set of hyperplanes H that cross both F and F'. Hence $\mathfrak{g}_F(F')$, $\mathfrak{g}_{F'}(F)$ are parallel.

Suppose that $F \cap F' \neq \varnothing$ and let $x \in F \cap F'$. Then $\mathfrak{g}_F(x) = x$, by definition, so $\mathfrak{g}_F(F') \supseteq F \cap F'$. On the other hand, let $y \in \mathfrak{g}_F(F')$, so $y = \mathfrak{g}_F(y')$ for some $y' \in F'$. Let m be the median of y, y', x for some $x \in F \cap F'$. By convexity of F and of F', we have $m \in F \cap F'$ and $d_{\mathcal{X}}(y, y') \geq d_{\mathcal{X}}(y', m)$. Hence $y = m$, so $y \in F \cap F'$. Thus $\mathfrak{g}_F(F') \subseteq F \cap F'$, as required. □

The following lemmas concern the projection of geodesics onto hyperplanes:

Lemma 2.7 *Let $\alpha \subseteq \mathcal{X}$ be a combinatorial geodesic and let $K \subseteq \mathcal{X}$ be a convex subcomplex. Then $\mathfrak{g}_K(\alpha)$ is a geodesic in K. Moreover, suppose that there exists R such that $d_{\mathcal{X}}(a, K) \leq R$ and $d_{\mathcal{X}}(a', K) \leq R$, where a, a' are the initial and terminal 0–cubes of α. Then $d_{\mathcal{X}}(t, K) \leq 2R$ for all 0–cubes t of α.*

Proof Let $\alpha' = \mathfrak{g}_K(\alpha)$. Any hyperplane separating $t, t' \in \alpha'$ separates $s, s' \in \alpha$ with $\mathfrak{g}_K(s) = t$ and $\mathfrak{g}_K(s') = t'$, by the definition of gates. Hence the set of hyperplanes crossing α' has the following properties: if H, H' cross α' and are separated by H'', then H'' crosses α'; if H, H', H'' are pairwise-disjoint hyperplanes crossing α', then one of them separates the other two. Hence, if $a_0, \ldots, a_m$ is an ordered sequence of 0–cubes of α, then $\mathfrak{g}_K(a_0), \ldots, \mathfrak{g}_K(a_m)$ has the property that $\mathfrak{g}_K(a_i), \mathfrak{g}_K(a_{i+1})$ are either equal or adjacent for all i. If a hyperplane H separates $\mathfrak{g}_K(a_i)$ from $\mathfrak{g}_K(a_{i+1})$ and also separates $\mathfrak{g}_K(a_j)$ from $\mathfrak{g}_K(a_{j+1})$, then H must separate a_i from a_{i+1} and a_j from a_{j+1}, a contradiction. This proves the first assertion.

To prove the second assertion, let $\alpha' = \mathfrak{g}_K(\alpha)$. Let $\mathcal{H}$ be the set of hyperplanes separating α from $\mathfrak{g}_K(\alpha)$. For all $t \in \alpha$, any hyperplane V separating t from K must not cross K. Hence V separates a or a' from K. The total number of such V is at most $2R - |\mathcal{H}|$. □

2.2 Contact and crossing graphs and complexes

As before, $\mathcal{X}$ is an arbitrary CAT(0) cube complex and $\mathcal{CX}$ its contact graph. Recall that hyperplanes H, H' represent adjacent vertices of $\mathcal{CX}$ if $\mathcal{N}(H) \cap \mathcal{N}(H') \neq \varnothing$, which occurs if either $H \cap H' \neq \varnothing$ (in which case H, H' *cross*, denoted $H \perp H'$), or if there are 1–cubes, dual to H, H' respectively, with a common 0–cube that does not form the corner of a 2–cube, in which case H, H' *osculate*. If H, H' are adjacent in $\mathcal{CX}$, we say that they *contact*, denoted $H \perp\!\!\!\perp H'$. As we just did, we often abuse notation by saying two hyperplanes are adjacent, when we really mean that the vertices

represented by these hyperplanes are adjacent; similarly, we will talk about a sequence of hyperplanes forming a geodesic, etc. We sometimes refer to the *crossing graph* $\mathcal{C}_\sharp \mathcal{X}$ of $\mathcal{X}$, the spanning subgraph of $\mathcal{C}\mathcal{X}$ formed by removing open edges that record osculations.

The following statements are the beginning of the analogy between, on one hand, $\mathcal{X}$ and its contact graph and, on the other hand, $\mathcal{MCG}(S)$ and the curve complex of the surface S.

Theorem 2.8 (hyperbolicity of contact graphs [33]) *For any* CAT(0) *cube complex* $\mathcal{X}$, *there is a simplicial tree* T *and a* $(10, 10)$*–quasi-isometry* $\mathcal{C}\mathcal{X} \to T$.

Under fairly mild geometric hypotheses, whether or not the quasi-tree $\mathcal{C}\mathcal{X}$ is actually a quasi-point can be detected by examining the *simplicial boundary*; moreover, if $\mathcal{X}$ admits an essential, proper, cocompact group action, then $\mathcal{C}\mathcal{X}$ is either a join of infinite subgraphs, or is unbounded [32]. This is closely related to the action of rank-one elements on $\mathcal{C}\mathcal{X}$ (recall that $g \in \mathrm{Aut}(\mathcal{X})$ is *rank-one* if it is hyperbolic and no axis bounds a half-flat in $\mathcal{X}$). As for the extension graph of a right-angled Artin group, there is a Nielsen–Thurston-style classification of automorphisms describing how elements of $\mathrm{Aut}(\mathcal{X})$ act on the contact graph.

Theorem 2.9 (classification of cubical automorphisms [32]) *Let* $\mathcal{X}$ *be a* CAT(0) *cube complex such that every clique in* $\mathcal{C}\mathcal{X}$ *is of uniformly bounded cardinality. Let* $g \in \mathrm{Aut}(\mathcal{X})$. *Then one of the following holds:*

(1) **Invariant clique** *There exists* $n > 0$ *and a hyperplane* H *such that* $g^n H = H$.

(2) **Invariant join** *There exists a subgraph* Λ *of* $\mathcal{C}\mathcal{X}$ *such that* $g\Lambda \subset \Lambda$ *and the following holds:* $\Lambda = \Lambda_0 \sqcup \Lambda_1$, *and each vertex of* Λ_0 *is adjacent to all but finitely many vertices of* Λ_1. *In this case,* g *is not a rank-one element.*

(3) **Loxodromic** g *has a quasigeodesic axis in* $\mathcal{C}\mathcal{X}$.

In the last case, g *is contracting and no positive power of* g *stabilizes a hyperplane.*

We will not use the subgraph Λ in the rest of this paper; that description follows from Theorem D of [32]. We will use the following: if $g \in \mathrm{Aut}(\mathcal{X})$ is not rank-one, then it either stabilizes a cube in $\mathcal{X}$ and thus has a bounded orbit in $\mathcal{C}\mathcal{X}$, or its combinatorial axis lies uniformly close to a half-flat F in $\mathcal{X}$. As explained in [32, Proposition 5.1], the set of hyperplanes crossing F has uniformly bounded diameter in $\mathcal{C}\mathcal{X}$, and hence the same is true of the $\langle g \rangle$ orbit in $\mathcal{C}\mathcal{X}$.

2.3 Special cube complexes, right-angled Artin groups, and the extension graph

The *right-angled Artin group* A_Γ presented by the graph Γ is the group with presentation

$$\langle \text{Vertices}(\Gamma) \mid \{[v, w] : v, w \in \text{Vertices}(\Gamma), \{v, w\} \in \text{Edges}(\Gamma)\} \rangle,$$

ie the generators are the vertices of Γ and two generators commute if and only if the corresponding vertices span an edge of Γ. The *Salvetti complex* S_Γ is the non-positively-curved cube complex with a single 0–cube and an n–cube for each set of n pairwise-commuting generators, for $n \geq 1$. Note that S_Γ is compact, and A_Γ finitely generated, if and only if Γ is finite. See the survey [22] for more details.

The universal cover $\widetilde{S}_\Gamma$ is a CAT(0) cube complex with special features. For each induced subgraph Λ of Γ, the inclusion $\Lambda \hookrightarrow \Gamma$ induces an injective local isometry $S_\Lambda \to S_\Gamma$, lifting to a A_Λ–equivariant convex embedding $\widetilde{S}_\Lambda \hookrightarrow \widetilde{S}_\Gamma$. In particular, when Λ is the link of a vertex v, then $\widetilde{S}_\Lambda$ is a combinatorial hyperplane in $\widetilde{S}_\Gamma$ which is a copy of a hyperplane whose dual 1–cubes are labeled by the generator v. Moreover, when Λ is the star of v, the subcomplex $\widetilde{S}_\Lambda$ is isomorphic to $\mathbf{R} \times \widetilde{S}_{\text{Lk}(v)}$, where $\mathbf{R}$ is a convex subcomplex isometric to $\mathbb{R}$ and stabilized by $\langle v \rangle$.

The *extension graph* Γ^e associated to Γ, introduced by Kim and Koberda [43], is the graph with a vertex for each A_Γ conjugate of each generator v with $v \in \Gamma^{(0)}$ and an edge joining the distinct vertices corresponding to v^g, w^h if and only if $[v^g, w^h] = 1$. Like the contact graph of a cube complex, Γ^e is always a quasi-tree [43], and in fact, in many situations, Γ^e is quasi-isometric to $\mathcal{C}\widetilde{S}_\Gamma$, as explained in [44].

Part I   Geometry of the contact graph

We now discuss projection to the contact graph and identify "hierarchy paths" in a cube complex. These results will be reworked in Part II once we have introduced factor systems.

Disc diagram techniques, originating in unpublished notes of Casson and developed in [25; 62; 72], have proven to be a useful tool in studying the geometry of $\mathcal{C}\mathcal{X}$ (see [26; 33; 32]), and we will continue to make use of these here.

3   Hierarchy paths

The next proposition establishes Theorem A(2).

Proposition 3.1 (hierarchy paths) *Let $\mathcal{X}$ be a CAT(0) cube complex with contact graph $\mathcal{CX}$ and let $x, y \in \mathcal{X}$ be 0–cubes. Then there exist hyperplanes $H_0, \ldots, H_k$ with $x \in \mathcal{N}(H_0)$, $y \in \mathcal{N}(H_k)$, and combinatorial geodesics $\gamma_i \to \mathcal{N}(H_i)$, such that $H_0 \perp\!\!\!\perp H_1 \perp\!\!\!\perp \cdots \perp\!\!\!\perp H_k$ is a geodesic of $\mathcal{CX}$ and $\gamma_0 \gamma_1 \cdots \gamma_k$ is a geodesic joining x, y.*

Proof Since the set of hyperplane carriers covers $\mathcal{X}$, there exist hyperplanes H, H' with $x \in \mathcal{N}(H)$, $y \in \mathcal{N}(H')$. Let $k = d_{\mathcal{CX}}(H, H')$, and let $H = H_0, H_1, \ldots, H_k = H'$ be a geodesic sequence of hyperplanes in $\mathcal{CX}$, so that $\mathcal{N}(H_i) \cap \mathcal{N}(H_{i+1}) \neq \varnothing$ for $0 \leq i \leq k-1$ and $\mathcal{N}(H_i) \cap \mathcal{N}(H_j) = \varnothing$ for $|i - j| > 1$. Let γ_0 be a combinatorial geodesic of $\mathcal{N}(H_0)$ joining x to a 0–cube in $\mathcal{N}(H_0) \cap \mathcal{N}(H_1)$. For $1 \leq i \leq k-1$, let γ_i be a geodesic of $\mathcal{N}(H_i)$ joining the terminal 0–cube of γ_{i-1} to a 0–cube of $\mathcal{N}(H_i) \cap \mathcal{N}(H_{i+1})$. Finally, let γ_k be a geodesic of $\mathcal{N}(H_k)$ joining the terminal 0–cube of γ_{k-1} to y. Since hyperplane carriers are convex subcomplexes of $\mathcal{X}$, each γ_i is a geodesic of $\mathcal{X}$ and in particular contains at most one 1–cube dual to each hyperplane. Let $\gamma = \gamma_0 \cdots \gamma_k$.

We now show that the above choices can be made in such a way that the path γ is immersed in $\mathcal{X}$, ie γ has no self-intersections. Since $\mathcal{N}(H_i) \cap \mathcal{N}(H_j) = \varnothing$ for $|i - j| > 1$ we can restrict our attention to intersections between γ_i and γ_{i+1}. Any such point of intersection must lie in $\mathcal{N}(H_i) \cap \mathcal{N}(H_{i+1})$ and we can then replace γ_i and γ_{i+1} by γ_i' and γ_{i+1}', where $\gamma_i' \subset \gamma_i$ and $\gamma_{i+1}' \subset \gamma_{i+1}$ are the geodesic subpaths obtained by restricting to the subpaths before and after the intersection point, respectively. Applying this procedure for each i to the first intersection point ensures that γ is immersed.

Let τ be a combinatorial geodesic of $\mathcal{X}$ joining y to x, so that $\gamma\tau$ is a closed path in $\mathcal{X}$, and let $D \to \mathcal{X}$ be a disc diagram with boundary path $\gamma\tau$. Suppose that D has minimal area among all disc diagrams with this boundary path, and that the choice of geodesic of $\mathcal{CX}$, and the subsequent choice of paths γ_i, were made in such a way as to minimize the area of D. See Figure 1.
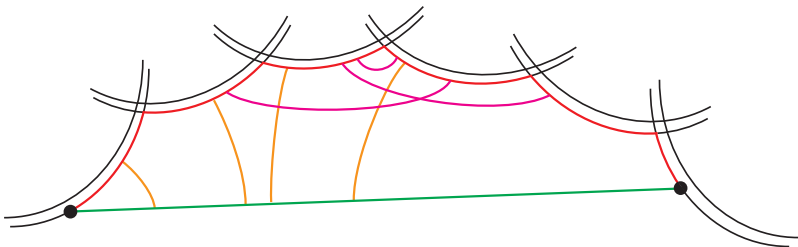

Figure 1: The diagram D, showing some possible (in orange) and impossible (in pink) dual curves

Note that for all i, if K, L are dual curves emanating from γ_i, then K, L do not intersect. Otherwise, convexity of $\mathcal{N}(H_i)$ would enable a new choice of γ_i, lowering the area of D.

Let K be a dual curve in D emanating from γ. It suffices to show that K must end on τ. Indeed, it will then follow that $|\gamma| \leq |\tau|$, whence γ is a geodesic of the desired type. To this end, suppose that K emanates from a 1–cube of γ_i. Then K cannot end on γ_i, since γ_i contains exactly one 1–cube dual to the hyperplane to which K maps. K cannot end on γ_j with $|i - j| > 2$, for then $\mathcal{CX}$ would contain a path (H_i, V, H_j), where V is the hyperplane to which K maps, contradicting the fact that $d_{\mathcal{CX}}(H_i, H_j) = |i - j|$. If K ends on $\gamma_{i\pm 2}$, then $\gamma_{i\pm 1}$ could be replaced by a path in $\mathcal{N}(V)$ that is the image of a path in the interior of D, contradicting our minimal-area choices. Indeed, recall that we chose the $\mathcal{CX}$–geodesic and the associated geodesics γ_i so that the resulting diagram D (which can be constructed given any such choices) is of minimal area among all diagrams constructible by any choice of such $\mathcal{CX}$ and $\mathcal{X}$ geodesics.

Finally, suppose K ends on γ_{i+1}. Then no dual curve emanating from the part of $\gamma_i \gamma_{i+1}$ subtended by the 1–cubes dual to K can cross K. Hence the 1–cube of γ_i dual to K is equal to the 1–cube of γ_{i+1} dual to K, contradicting the fact that γ is immersed. $\square$

Definition 3.2 A geodesic $\gamma = \gamma_1 \cdots \gamma_n$ such that $\gamma_i \to \mathcal{N}(H_i)$ for $1 \leq i \leq n$ and $H_1 \perp\!\!\!\perp \cdots \perp\!\!\!\perp H_n$ is a geodesic of $\mathcal{CX}$ is a *hierarchy path*. The geodesic $H_1 \perp\!\!\!\perp \cdots \perp\!\!\!\perp H_n$ *carries* γ, and each γ_i is a *syllable* of γ.

The following is immediate from the product structure of hyperplane-carriers:

Proposition 3.3 *Let* $\gamma = \gamma_1 \cdots \gamma_n$ *be a hierarchy path carried by* $H_1 \perp\!\!\!\perp \cdots \perp\!\!\!\perp H_n$. *Then there exists a hierarchy path* $\gamma' = \gamma'_1 \cdots \gamma'_n$, *joining the endpoints of* γ *and carried by* $H_1 \perp\!\!\!\perp \cdots \perp\!\!\!\perp H_n$, *such that for each* i, *we have* $\gamma'_i = A_i e$, *where the combinatorial geodesic* A_i *lies in* $H_i \times \{\pm 1\} \subset H_i \times [-1, 1] \cong \mathcal{N}(H_i)$ *and* e *is either a 0–cube or a 1–cube dual to* H_i.

A hierarchy path satisfying the conclusion of [Proposition 3.3] is *reduced*. A reduced hierarchy path $\gamma_1 \cdots \gamma_n$ is a *hereditary hierarchy path* if it is trivial or if for each i, the subpath path $A_i \to H_i \times \{\pm 1\}$ of γ_i is a hereditary reduced hierarchy path in $H_i \times \{\pm 1\}$. From [Proposition 3.3] and the definition of a hereditary reduced hierarchy path, we obtain:

Corollary 3.4 *Let* $\mathcal{X}$ *be a finite-dimensional* CAT(0) *cube complex and let* $x, y \in \mathcal{X}^{(0)}$. *Then* x *and* y *are joined by a hereditary reduced hierarchy path.*

4 Projection to the contact graph

The notion of projecting geodesics in $\mathcal{X}$ to $\mathcal{CX}$ was discussed in [32, Section 2], motivating the following definition:

Definition 4.1 (projection to the contact graph)   Let K be a convex subcomplex of $\mathcal{X}$. For each $k \in K$, let $\{H_i \cap K\}_{i \in \mathcal{I}}$ be the maximal collection of hyperplanes of K with $k \in \bigcap_{i \in \mathcal{I}} \mathcal{N}(H_i)$. Define $\rho_K \colon K \to 2^{\mathcal{CK}}$ by setting $\rho_K(k) = \{H_i\}_{i \in \mathcal{I}}$. The *projection of $\mathcal{X}$ on $\mathcal{CK}$* is the map $\pi_K = \rho_K \circ \mathfrak{g}_K \colon \mathcal{X} \to 2^{\mathcal{CK}}$. Note that π_K assigns to each point of $\mathcal{X}$ a clique in the contact graph of K and hence a clique in $\mathcal{CX}$.

If H is a hyperplane, let $H^{\pm}$ be $H \times \{\pm 1\} \subset \mathcal{N}(H) \cong H \times [-1, 1]$. There is a bijection between hyperplanes which intersect H^+ and ones which intersect H^-; moreover, associated hyperplanes contact (respectively, cross) in H^+ if and only if they contact (respectively, cross) in H^-, whence $\mathcal{C}H^+$ and $\mathcal{C}H^-$ are both the same subset of $\mathcal{CX}$. Abusing notation slightly, we let π_H denote the projection from $\mathcal{X}$ to $\mathcal{C}H^+ = \mathcal{C}H^- \subset \mathcal{CX}$. This map is defined as in Definition 4.1 since $H^{\pm}$ is a convex subcomplex, and it is not hard to see that it is independent of whether we took gates in H^+ or H^- (another option is to pass to the first cubical subdivision and just project to H and then to its contact graph, since subdividing makes H into a subcomplex).

Let $x, y \in \mathcal{X}$ be 0–cubes, and let H be a hyperplane that does not separate x from y. Let H^+ be the copy of H bounding $\mathcal{N}(H)$ that lies in the component of $\mathcal{X} - H$ containing $\{x, y\}$. It is easily checked that any hyperplane that separates $\mathfrak{g}_{H^+}(x)$ from $\mathfrak{g}_{H^+}(y)$ must separate x from y. In particular, if H does not separate x from y and it does not cross any hyperplane separating x from y, then $\pi_H(x) = \pi_H(y)$. In other words, if $\gamma \subset \mathcal{X}$ is a geodesic with $\pi_{\mathcal{X}}(\gamma) \cap B_1^{\mathcal{C}_\sharp \mathcal{X}}(H) = \varnothing$, then $\pi_H(x) = \pi_H(y)$, where x, y are the endpoints of γ. Let $x' \in \gamma$ be a 0–cube and suppose that the hyperplane U crosses H and separates the pair x, x', hence separating $\mathfrak{g}_{H^+}(x)$ from $\mathfrak{g}_{H^+}(x')$. Then U separates x, y, and hence belongs to $\pi_{\mathcal{X}}(\gamma) \cap B_1^{\mathcal{C}_\sharp \mathcal{X}}(H)$, contradicting our assumption. Hence $\pi_H(\gamma) = \pi_H(x) = \pi_H(y)$. Thus:

Proposition 4.2 (bounded geodesic image)   *For any $H \in \mathcal{CX}$ and any geodesic $\gamma \subseteq \mathcal{X}$ whose image in $\mathcal{CX}$ is disjoint from $B_1^{\mathcal{C}_\sharp \mathcal{X}}(H)$, the projection $\pi_H(\gamma)$ is a clique.*

5 Weak proper discontinuity of the action on the contact graph

We now consider the elements which act *weakly properly discontinuously (WPD)*, in the sense of [13]. In particular, we study the *WPD elements* of a group G acting on $\mathcal{X}$,

and prove Theorem A(1). By definition, as an isometry of $\mathcal{CX}$, some $h \in G$ is WPD if for all $\epsilon \geq 1$ and hyperplanes H, there exists $N > 0$ such that

$$\left| \{ g \in G : d_{\mathcal{CX}}(H, gH) < \epsilon, \, d_{\mathcal{CX}}(h^N H, gh^N H) < \epsilon \} \right| < \infty.$$

In the presence of a factor system and a cocompact group action, we achieve a stronger conclusion in Section 14, namely that the action on the contact graph is acylindrical.

Proposition 5.1 *Let $\mathcal{X}$ be a CAT(0) cube complex on which the group G acts metrically properly by isometries. Suppose that $h \in G$ is loxodromic on $\mathcal{CX}$. Then h is WPD.*

Proof Fix $\epsilon \geq 1$ and let H be a hyperplane. By hypothesis, $\langle h \rangle H$ is a quasigeodesic in $\mathcal{CX}$. It follows (see [32, Theorem 2.3]) that there exists $M = M(h, H, \mathcal{X})$ such that for all $N \geq 0$, and any hierarchy path $\gamma = \gamma_1 \cdots \gamma_p$ joining some $x \in H$ to $h^N x \in h^N H$ and carried on a geodesic $H = H_0 \pitchfork H_1 \pitchfork \cdots \pitchfork H_p = h^N H$ of $\mathcal{CX}$, we have $|\gamma_i| \leq M$ for $0 \leq i \leq p$. (Recall that γ is a geodesic, being a hierarchy path; hence γ contains at most one 1–cube dual to each hyperplane.)

Suppose that $g \in G$ has the property that

$$d_{\mathcal{CX}}(H, gH) < \epsilon \quad \text{and} \quad d_{\mathcal{CX}}(h^N H, gh^N H) < \epsilon.$$

Suppose, moreover, that N has been chosen so that $p > 2\epsilon + 12$, and choose i so that $|p/2 - i| \leq 2$. Choose a point $y \in \gamma_i$ and let $\mathcal{H}(y)$ be the set of hyperplanes separating y from $gy \in g\gamma_i$.

We claim that each $W \in \mathcal{H}(y)$ intersects γ and $g\gamma$. Indeed, suppose that W intersects γ but not $g\gamma$. Suppose that W intersects γ_j with $j \geq i$. Then W separates $x, h^N x$ (since γ is a geodesic) but not $gx, gh^N x$, so W separates x, gx or $h^N x, gh^N x$. Since W also separates y, gy, the former must hold. Hence

$$p = d_{\mathcal{CX}}(H, h^N H) \leq \epsilon + |p - i| \leq \tfrac{1}{2}p + \epsilon,$$

contradicting our choice of N. The case where W intersects γ_j, $j < i$, is similar.

Hence, if W crosses γ_s with $s \geq i$, then W crosses $g\gamma_t$, with $t \leq i$. (We have the reverse conclusion if $s \leq i$.) The fact that $H_0 \pitchfork H_1 \pitchfork \cdots \pitchfork H_p$ and its g–translate are $\mathcal{CX}$–geodesics implies that W intersects γ_s with $|i - s| \leq \epsilon + 6$. Hence $d_{\mathcal{X}}(y, gy) \leq M(\epsilon + 6)$, whence the number of such g is finite since the action of G on $\mathcal{X}$ is proper. $\qquad \square$

We say that a group is nonelementary if it does not contain a cyclic subgroup of finite index, and that an action of the group G is WPD if it admits a WPD element and G is nonelementary.

Corollary 5.2 (characterization of WPD elements) *Let $\mathcal{X}$ be a uniformly locally finite CAT(0) cube complex. Then for any nonelementary group G acting properly on $\mathcal{X}$, one of the following holds:*

(1) *The induced action of G on $\mathcal{CX}$ is WPD, and the WPD elements are precisely those $h \in G$ that are rank-one and do not have positive powers that stabilize hyperplanes.*

(2) *Every $h \in G$ is either not rank-one, or has a positive power stabilizing a hyperplane.*

Proof Apply Proposition 5.1 and Theorem 2.9. □

The following is an application to determining acylindrical hyperbolicity. Following [21], we say that the action of the group G on the CAT(0) cube complex $\mathcal{X}$ is *essential* if every halfspace contains points in a fixed G–orbit arbitrarily far away from the associated hyperplane.

Corollary 5.3 *Let G be a nonelementary group acting properly and essentially on a uniformly locally finite CAT(0) cube complex $\mathcal{X}$. Suppose that $\mathcal{X}$ is not a product of unbounded subcomplexes and at least one of the following two holds:*

(1) *G acts cocompactly on $\mathcal{X}$.*

(2) *G acts with no fixed point in the simplicial boundary $\partial_\triangle \mathcal{X}$ in the sense of [32].*

Then G is acylindrically hyperbolic.

Proof By [55, Theorem 1.2] it suffices to find a hyperbolic space on which G acts with a loxodromic WPD element. By [32, Theorem 5.4] there exists a rank-one element $g \in G$ no positive power of which stabilizes a hyperplane. The conclusion follows from Corollary 5.2. □

A major motivation for studying WPD actions arises from a result of Bestvina and Fujiwara relating WPD actions to bounded cohomology. Recall that the space $\widetilde{\mathrm{QH}}(G)$, which is the quotient of the space of quasimorphisms of G by the subspace generated by bounded functions and homomorphisms $G \to \mathbb{R}$, coincides with the kernel of the map $\mathrm{H}^2_b(G, \mathbb{R}) \to \mathrm{H}^2(G, \mathbb{R})$. Theorem 7 of [13] asserts that if G admits a WPD action, then $\widetilde{\mathrm{QH}}(G)$ is infinite-dimensional. This yields an alternative proof of the dichotomy obtained by Caprace and Sageev [21, Theorem H], as a consequence of rank-rigidity: a group G admitting a sufficiently nice action on a CAT(0) cube complex $\mathcal{X}$ that is not a product has infinite-dimensional $\widetilde{\mathrm{QH}}(G)$; instead of using rank-rigidity and results of [18; 20; 52], one can deduce their result from rank-rigidity, Corollary 5.2, and [13].

6 Contractibility

We now prove Theorem A(3). The *contact complex* $C_\bullet \mathcal{X}$ of the CAT(0) cube complex $\mathcal{X}$ is the flag complex with 1–skeleton $C\mathcal{X}$.

Theorem 6.1 *Let $\mathcal{X}$ be a CAT(0) cube complex with countable 0–skeleton and at least one 1–cube. Then $C_\bullet \mathcal{X}$ is contractible.*

Proof For each $n \geq 0$, choose a convex, compact subcomplex $\mathcal{B}_n$ in such a way that $\mathcal{B}_m \subseteq \mathcal{B}_n$ for $m \leq n$ and $\bigcup_{n \geq 0} \mathcal{B}_n = \mathcal{X}$. This choice is possible since $\mathcal{X}^{(0)}$ may be written as an increasing union of finite sets; $\mathcal{B}_n$ can then be taken to be the cubical convex hull of the nth one.

Since each $\mathcal{B}_n$ is convex in $\mathcal{X}$, it is itself a CAT(0) cube complex whose hyperplanes have the form $H \cap \mathcal{B}_n$, where H is a hyperplane of $\mathcal{X}$. Moreover, the map $H \cap \mathcal{B}_n \to H$ induces an embedding $C_\bullet \mathcal{B}_n \to C_\bullet \mathcal{X}$ whose image is a full subcomplex (ie any $k + 1$ 0–simplices of $C_\bullet \mathcal{B}_n$ span a k–simplex of $C_\bullet \mathcal{B}_n$ if and only if their images in $C_\bullet \mathcal{X}$ span a k–simplex of $C_\bullet \mathcal{X}$). Thus the set $\{C_\bullet \mathcal{B}_n\}_{n \geq 0}$ provides a filtration of $C_\bullet X$ by full subcomplexes, each of which is the contact complex of a compact CAT(0) cube complex. Indeed, every hyperplane intersects all but finitely many of the $\mathcal{B}_n$, and hence appears as a 0–simplex in all but finitely many of the subcomplexes $C_\bullet \mathcal{B}_n$.

For any $m \geq 0$ and any continuous map $f \colon \mathbb{S}^m \to C_\bullet \mathcal{X}$, compactness of im f implies that there exists $n \geq 0$ such that im $f \subseteq C_\bullet \mathcal{B}_n$. By Lemma 6.2, $C_\bullet \mathcal{B}_{n+1}$ is contractible, since it is the contact complex of a compact CAT(0) cube complex. Hence the inclusion $C_\bullet \mathcal{B}_n \hookrightarrow C_\bullet \mathcal{B}_{n+1}$ is null-homotopic, whence $\mathbb{S}^m \xrightarrow{f} C_\bullet \mathcal{B}_n \hookrightarrow C_\bullet \mathcal{B}_{n+1} \hookrightarrow C_\bullet \mathcal{X}$ is null-homotopic. It then follows from Whitehead's theorem [68; 69] that $\mathcal{X}$ is contractible. $\square$

Lemma 6.2 *Let $\mathcal{X}$ be a compact CAT(0) cube complex with at least one 1–cube. Then $C_\bullet \mathcal{X}$ is contractible.*

Proof We will argue by induction on $n = |C_\bullet \mathcal{X}^{(0)}|$, ie the number of hyperplanes in $\mathcal{X}$. When $n = 1$, the cube complex $\mathcal{X}$ is necessarily a single 1–cube, so $C_\bullet \mathcal{X}$ is a 0–simplex. Let $n \geq 1$. Since $\mathcal{X}$ has finitely many hyperplanes, there exists a hyperplane $H \subset \mathcal{X}$ such that one of the components of $\mathcal{X} - H$ has closure $H \times [0, 1]$, where $\mathcal{N}(H)$ is identified with $H \times [-1, 1]$ and H with $H \times \{0\}$. This generalizes the case in which $\mathcal{X}$ is a tree and H is the midpoint of an edge containing a degree-1 vertex; accordingly, such a hyperplane H will be called a *leaf hyperplane*. When H is a leaf hyperplane, we always denote by H^+ the halfspace $H \times [0, 1]$ and by H^- the other closed halfspace.

By Lemma 6.3, $\mathcal{A} = \mathrm{Cl}(\mathcal{X} - \mathcal{N}(H))$ is a convex proper subcomplex of $\mathcal{X}$, so that $\mathcal{C}_\bullet \mathcal{A}$ embeds in $\mathcal{C}_\bullet \mathcal{X}$ as a full subcomplex. Moreover, by Lemma 6.4, there is a convex proper subcomplex $H' \subset \mathcal{X}$ such that $\mathcal{C}_\bullet \mathcal{X} \cong \mathcal{C}_\bullet \mathcal{A} \cup_L (L \star \{H\})$, where L is a subcomplex of $\mathcal{C}_\bullet \mathrm{Cl}(\mathcal{X} - \mathcal{N}(H))$ isomorphic to $\mathcal{C}_\bullet H'$. Each of $\mathcal{A}$ and H' is a CAT(0) cube complex whose set of hyperplanes corresponds bijectively to a subset of $\mathcal{X}^{(0)} - \{H\}$, so by induction, $\mathcal{C}_\bullet \mathcal{A}$ and L are contractible. Since L is contractible, there is a homotopy equivalence $f \colon L \star \{H\} \to L$ that is the identity on L, whence the pasting lemma yields a homotopy equivalence $\mathcal{C}_\bullet \mathcal{X} \to \mathcal{C}_\bullet \mathcal{A}$. Since $\mathcal{A}$ is contractible, the same is therefore true of $\mathcal{C}_\bullet \mathcal{X}$. $\qquad\square$

Lemma 6.3 *Let H be a leaf hyperplane of the CAT(0) cube complex $\mathcal{X}$. Then $\mathrm{Cl}(\mathcal{X} - \mathcal{N}(H))$ is a convex subcomplex of $\mathcal{X}$.*

Proof Since $H^+ \subset \mathcal{N}(H)$, the subcomplex $\mathrm{Cl}(\mathcal{X} - \mathcal{N}(H))$ is exactly the convex hull of the halfspace $H^- \cap \mathcal{X}^{(0)}$ of the 0–skeleton induced by H. $\qquad\square$

Lemma 6.4 *Let H be a leaf hyperplane of the compact CAT(0) cube complex $\mathcal{X}$. Then there is convex subcomplex $H' \subsetneq \mathcal{X}$ such that there is an isomorphism*

$$\mathcal{C}_\bullet \mathcal{X} \cong \mathcal{C}_\bullet \mathrm{Cl}(\mathcal{X} - \mathcal{N}(H)) \cup_L (L \star \{H\}),$$

where L is a subcomplex of $\mathcal{C}_\bullet \mathrm{Cl}(\mathcal{X} - \mathcal{N}(H))$ isomorphic to $\mathcal{C}_\bullet H'$.

Proof Let $\mathcal{A} = \mathrm{Cl}(\mathcal{X} - \mathcal{N}(H))$. Let $\{V_1, \ldots, V_s\}$ be the hyperplanes of $\mathcal{X}$ that contact H, and let L be the full subcomplex of $\mathcal{C}_\bullet \mathcal{X}$ generated by $\{V_1, \ldots, V_s\}$. Lemma 7.11 of [33] implies that there is an isometrically embedded subcomplex $H' \subset \mathcal{X}$ such that a hyperplane intersects H' if and only if that hyperplane belongs to $\{V_1, \ldots, V_k\}$. By replacing H' if necessary by its convex hull, we may assume that H' is convex, and hence $\mathcal{C}_\bullet H' \cong L$. The decomposition is obvious from the definitions. $\qquad\square$

7 Automorphisms of the contact graph

We now provide a short example which shows that the analogue of Ivanov's theorem, that the mapping class group is the automorphism group of the curve graph [41], fails to hold for the contact graph. Indeed, the example below shows that this failure holds even if one considers an edge-colored version of the contact graph differentiating edges associated to crossing hyperplanes from those associated to osculating ones. We note that in the case of RAAGs, the fact that the automorphism group of the contact graph may be much larger than the RAAG itself is familiar from the case of the extension

graph, where Kim and Koberda prove that the automorphism group of the extension graph is often uncountable [44, Theorem 66].

Example 7.1 In the cube complex pictured in Figure 2 one sees that the automorphism group of the cube complex is virtually cyclic, and in particular, pairs of valence-zero vertices, such as those labeled a, b, cannot be permuted willy-nilly. Whereas in the contact graph, the pair of simplices (labeled a', b'), one for each hyperplane separating off such a vertex, can be swapped arbitrarily; thus the contact graph contains a $(\mathbb{Z}/2\mathbb{Z})^\infty$ subgroup.
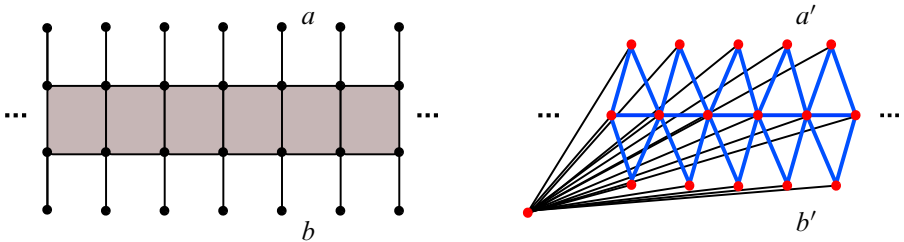

Figure 2: Cube complex $\mathcal{X}$ (left) and its associated contact graph $\widehat{\mathcal{C}}\mathcal{X}$ (right)

Part II Factor systems, projections, and the distance formula

Throughout, $\mathcal{X}$ is a CAT(0) cube complex. The setting in which we will work is established in Definition 8.1 and forces $\mathcal{X}$ to be uniformly locally finite.

8 Factored contact graphs

8.1 Factor systems

Definition 8.1 (factor system) A set $\mathfrak{F}$ of subcomplexes which satisfies the following is called a *factor system* in $\mathcal{X}$:

(1) $\mathcal{X} \in \mathfrak{F}$.

(2) Each $F \in \mathfrak{F}$ is a nonempty convex subcomplex of $\mathcal{X}$.

(3) There exists $\Delta \geq 1$ such that for all $x \in \mathcal{X}^{(0)}$, at most Δ elements of $\mathfrak{F}$ contain x.

(4) Every nontrivial convex subcomplex parallel to a combinatorial hyperplane of $\mathcal{X}$ is in $\mathfrak{F}$.

(5) There exists $\xi \geq 0$ such that for any pair of subcomplexes $F, F' \in \mathfrak{F}$, either $\mathfrak{g}_F(F') \in \mathfrak{F}$ or $\operatorname{diam}(\mathfrak{g}_F(F')) < \xi$.

Convexity of each $F \in \mathfrak{F}$ implies that each F is a CAT(0) cube complex whose hyperplanes have the form $H \cap F$, where H is a hyperplane of $\mathcal{X}$, and the map $H \cap F \mapsto H$ induces an injective graph homomorphism $\mathcal{C}F \hookrightarrow \mathcal{C}\mathcal{X}$ whose image is an induced subgraph, which, by an abuse of notation, we also denote $\mathcal{C}F$.

If $\mathfrak{F}$ is a factor system for $\mathcal{X}$, then for each $F \in \mathfrak{F}$ the set $\mathfrak{F}_F = \{H \in \mathfrak{F} : H \subseteq F\}$ is a factor system in F such that each point in F lies in at most $\Delta - 1$ elements of $\mathfrak{F}_F$. (The distinction between $\mathfrak{F}_F$ and $\{H \cap F : H \in \mathfrak{F}\}$ is small: the latter consists of the former, together with some subcomplexes of diameter at most ξ. It is mainly for convenience that we choose to work, everywhere, with $\mathfrak{F}_F$ rather than $\{H \cap F : H \in \mathfrak{F}\}$.)

8.2 Examples of factor systems

8.2.1 Special groups Universal covers of special cube complexes, defined in [37], contain factor systems provided the co-special action has finitely many orbits of hyperplanes, as we shall show below. This provides a very large class of groups for which the distance formula from Section 9 holds. Let us start by studying universal covers of Salvetti complexes.

The following definition is tailored to the proof of Proposition 8.3.

Definition 8.2 Let Γ be a simplicial graph. A collection $\mathcal{R}$ of subgraphs of Γ is *rich* if

(1) $\Gamma \in \mathcal{R}$,

(2) all links of vertices of Γ are in $\mathcal{R}$, and

(3) if $A, B \in \mathcal{R}$ then $A \cap B \in \mathcal{R}$.

The collection of all subgraphs of Γ is a rich family. Also, any graph admits a minimal rich family, consisting of Γ together with all nonempty intersections of links of vertices.

Proposition 8.3 *Let Γ be a finite simplicial graph, $A(\Gamma)$ a right-angled Artin group, and $\widetilde{S}_\Gamma$ the universal cover of its Salvetti complex. Let $\mathcal{R}$ be a rich family of subgraphs of Γ.*

Let $\mathfrak{F}$ be the ($A(\Gamma)$–invariant) collection of convex subcomplexes of $\widetilde{S}_\Gamma$ containing all lifts of the sub-Salvetti complexes S_Λ of S_Γ, for all $\Lambda \in \mathcal{R}$. Then $\mathfrak{F}$ is a factor system.

In other words, $\mathfrak{F}$ contains a subcomplex stabilized by each conjugate of $A(\Lambda)$, for each subgraph Λ of Γ.

Proof of Proposition 8.3 In the definition of a factor system, item (1) holds because $\Gamma \in \mathcal{R}$. Items (2) and (3) are clear. Item (4) holds because combinatorial hyperplanes in $\widetilde{S}_\Gamma$ are exactly lifts of sub-Salvetti complexes of S_Γ corresponding to links of vertices. More precisely, if a hyperplane H is dual to a 1–cube corresponding to the generator $v \in \Gamma^{(0)}$, then the combinatorial hyperplanes on the two sides of H are lifts of $S_{\mathrm{Lk}(v)} \subseteq S_\Gamma$.

The content of the proposition is hence that item (5) holds (for $\xi = 0$). To simplify the notation, we will identify the 0–skeleton of $\widetilde{S}_\Gamma$ with $A(\Gamma)$. Under this identification, the 0–skeleton of each $F \in \mathfrak{F}$ corresponds to a coset of $A(\Lambda) < A(\Gamma)$ for some subgraph $\Lambda \in \mathcal{R}$, and each such coset is the 0–skeleton of some $F \in \mathfrak{F}$.

Let $F, F' \in \mathfrak{F}$, whose 0–skeleta are (possibly after applying an element of $A(\Gamma)$) $A(\Lambda_0)$ and $gA(\Lambda_1)$, for some subgraphs Λ_0, Λ_1 of Γ and $g \in A(\Gamma)$. We can assume, using the action of some $g \in A(\Gamma_0)$, that $1 \in \mathfrak{g}_F(F')$. Recall from Lemma 2.4 and Lemma 2.6 that $\mathfrak{g}_F(F')$ is in a natural way one of the factors in a product region R of $\widetilde{S}_\Gamma$, the other factor being (naturally identified with) a possibly trivial geodesic γ from 1 to g, up to changing the choice of g within the same coset of $A(\Lambda_1)$. Also, $\mathfrak{g}_F(F') \times \{1\}$ is contained in F and $\mathfrak{g}_F(F') \times \{g\}$ is contained in F'. Let Λ_2 be the link in Λ of the set of vertices of Λ that label some 1–cube along γ. The following claim concludes the proof.

Claim *The 0–skeleton of $\mathfrak{g}_F(F')$ is $A(\Lambda)$, where $\Lambda = \bigcap_{i=0}^{2} \Lambda_i \in \mathcal{R}$.*

Let us first show $\mathfrak{g}_F(F')^{(0)} \subseteq A(\Lambda)$. Consider a geodesic δ joining 1 to $h \in \mathfrak{g}_F(F')^{(0)}$. Any 1–cube e of δ is contained in F as well as parallel to a 1–cube of F', which implies that the label v of e belongs to $\Lambda_0 \cap \Lambda_1$. Let us now show that e also belongs to Λ_2. Once we have done that, it is clear that h can be written as a product of generators each belonging to $A(\Lambda)$.

Fix any 1–cube e' of γ. The product region R contains a square with two 1–cubes parallel to e and two 1–cubes parallel to e'. This means that the labels of e, e' commute and are distinct. As this holds for any e', the label of e belongs to Λ_2, as required.

We are left to show $A(\Lambda) \subseteq \mathfrak{g}_F(F')^{(0)}$. If $h \in A(\Lambda)$, there exists a geodesic δ from 1 to h whose 1–cubes are labeled by elements of γ. It is then easy to see that $\widetilde{S}_\Gamma$ (and in fact R) contains a product region naturally identified with $\delta \times \gamma$ with the property that $\delta \times \{1\}$ is contained in F and $\delta \times \{g\}$ is contained in F'. In particular, δ, and thus h, is contained in $\mathfrak{g}_F(F')$, as required. $\square$

Definition 8.4 (induced factor system) Let $\mathcal{X}$ be a CAT(0) cube complex with a factor system $\mathfrak{F}$ and let $\mathcal{Y} \subseteq \mathcal{X}$ be a convex subcomplex. The *induced factor system* $\mathfrak{F}_\mathcal{Y}$ is the set of nonempty subcomplexes of $\mathcal{Y}$ of the form $F \cap \mathcal{Y}$, for $F \in \mathfrak{F}$.

Lemma 8.5 Let $\mathcal{X}$ be a CAT(0) cube complex with a factor system $\mathfrak{F}$ and let $\mathcal{Y} \subseteq \mathcal{X}$ be a convex subcomplex. Then $\mathfrak{F}_\mathcal{Y}$ is a factor system in $\mathcal{Y}$.

Proof Item (1) of Definition 8.1 holds since $\mathcal{X} \cap \mathcal{Y} = \mathcal{Y}$. Item (2) follows since intersections of convex subcomplexes are convex. Item (3) follows since $\mathfrak{F}$ is a uniformly locally finite collection. To verify item (4), let H^+ be a combinatorial hyperplane of $\mathcal{Y}$ and let H be the (genuine) hyperplane whose carrier contains H^+ as one of its bounding copies. By convexity of $\mathcal{Y}$, the hyperplane H has the form $W \cap \mathcal{Y}$ for some hyperplane W of $\mathcal{X}$, and hence $H^+ = W^+ \cap \mathcal{Y}$, where W^+ is one of the combinatorial hyperplanes bounding the carrier of W. But $W^+ \in \mathfrak{F}$, by item (4), so $W^+ \cap \mathcal{Y} = H^+ \in \mathfrak{F}_\mathcal{Y}$.

It suffices to verify item (5), namely that $\mathfrak{F}_\mathcal{Y}$ is closed under (large) projection. To that end, let $F, F' \in \mathfrak{F}$ and suppose that $\mathrm{diam}(\mathfrak{g}_{F \cap \mathcal{Y}}(F' \cap \mathcal{Y})) \geq \xi$, where ξ is the constant associated to $\mathfrak{F}$ by Definition 8.1. Then, by item (5), applied to $\mathfrak{F}$, we have $\mathfrak{g}_F(F') \in \mathfrak{F}$. It thus suffices to show that $\mathfrak{g}_F(F') \cap \mathcal{Y} = \mathfrak{g}_{F \cap \mathcal{Y}}(F' \cap \mathcal{Y})$. For convenience, let $A = F \cap \mathcal{Y}$, $A' = F' \cap \mathcal{Y}$. Since $A \subseteq \mathcal{Y}$, we have $\mathfrak{g}_A(A') \subseteq \mathcal{Y}$. Since $A \subseteq F$ and $A' \subseteq F'$, we have $\mathfrak{g}_A(A') \subseteq \mathfrak{g}_F(F')$. Hence $\mathfrak{g}_A(A') \subseteq \mathfrak{g}_F(F') \cap \mathcal{Y}$.

Conversely, suppose that $x \in \mathfrak{g}_F(F') \cap \mathcal{Y}$. Then $x \in A$, since $x \in F \cap \mathcal{Y}$. Since $x \in \mathfrak{g}_F(F')$, there exists $x' \in F'$ such that a hyperplane H separates x' from F if and only if H separates x' from x. Let $z = \mathfrak{g}_\mathcal{Y}(x')$. Note that convexity of $\mathcal{Y}$ is used here to make z well-defined. Let V be a hyperplane separating z from x. Then either V separates x' from x, and hence from F, whence V separates z from A, or V separates x, x' from z. Suppose the latter. Since V separates x' from z, it must separate x' from Y since $z = \mathfrak{g}_\mathcal{Y}(x')$. But then V cannot cross $\mathcal{Y}$, and hence cannot separate $z \in \mathcal{Y}$ from $x \in \mathcal{Y}$, a contradiction. Thus V separates z from x if and only if V separates z from A, so $x = \mathfrak{g}_A(z)$. It remains to show that $z \in A'$, but this follows from Lemma 8.6. $\square$

Lemma 8.6 Let $\mathcal{X}$ be a CAT(0) cube complex and let $\mathcal{Y}, \mathcal{Z}$ be convex subcomplexes, with $A = \mathcal{Y} \cap \mathcal{Z}$. Then for all $x \in \mathcal{Z}$, we have $\mathfrak{g}_\mathcal{Y}(x) \in A$.

Proof Let $y = \mathfrak{g}_\mathcal{Y}(x)$, let $a = \mathfrak{g}_A(x)$, and let m be the median of x, a, and z. Then $m \in \mathcal{Z}$ since it lies on a geodesic from x to a, and $x, a \in \mathcal{Z}$, and $\mathcal{Z}$ is convex. Similarly, m lies on a geodesic from y to a, and thus $m \in \mathcal{Y}$. It follows that $m \in A$, whence $m = a$ since $\mathrm{d}_\mathcal{X}(x, m) \leq \mathrm{d}_\mathcal{X}(x, a)$. But then $\mathrm{d}_\mathcal{X}(x, a) \leq \mathrm{d}_\mathcal{X}(x, y)$, so $y = a$. $\square$

If C is a special cube complex then C admits a local isometry into some Salvetti complex by [37, Theorem 1.1], and this is the Salvetti complex of a finitely generated right-angled Artin group when C has finitely many immersed hyperplanes (eg when

C is compact special). Such a local isometry lifts to a convex embedding at the level of universal covers, whence Proposition 8.3 and Lemma 8.5 immediately imply that the universal cover of C admits a factor system. (Note that finite generation of the right-angled Artin group is important, since otherwise the universal cover of the Salvetti complex does not have a factor system because each 0–cube is contained in infinitely many combinatorial hyperplanes in this case.) We will now describe such factor systems.

Definition 8.7 Let C be a special cube complex whose set $\mathcal{H}$ of immersed hyperplanes is finite. For every $\mathcal{A} \subseteq \mathcal{H}$ and 1–cubes e, e' of C, write $e \sim_{\mathcal{A}} e'$ if there is a path $e = e_0 \cdots e_n = e'$ in the 1–skeleton of C such that each e_i is dual to some immersed hyperplane from $\mathcal{A}$. Let $C_{\mathcal{A}}$ be the collection of full subcomplexes of C whose 1–skeleton is an equivalence class of 1–cubes with respect to the equivalence relation $\sim_{\mathcal{A}}$.

Notice that each $D \in C_{\mathcal{A}}$ is locally convex.

Corollary 8.8 (factor systems for special groups) *Let $\mathcal{X}$ be the universal cover of a special cube complex C with finitely many immersed hyperplanes, the set of which we denote by $\mathcal{H}$. Then $\mathcal{X}$ admits a factor system: the collection of all lifts of subcomplexes in $\bigcup_{\mathcal{A} \subseteq \mathcal{H}} C_{\mathcal{A}}$ is a factor system for $\mathcal{X}$.*

Proof Let Γ be the *crossing graph* of C, which has a vertex for each immersed hyperplane, with two vertices adjacent if the corresponding immersed hyperplanes have nonempty intersection. Then for each immersed hyperplane H of C there is a corresponding 1–cube e_H in S_{Γ}. In [37, Theorem 1.1] it is shown that there is a local isometry $\phi \colon C \to S_{\Gamma}$ such that if the 1–cube e is dual to the immersed hyperplane H then it gets mapped isometrically to e_H.

The local isometry ϕ lifts to a convex embedding $\widetilde{\phi} \colon \mathcal{X} \to \widetilde{S}_{\Gamma}$ (see eg [73, Lemma 3.12]), so in view of Proposition 8.3 and Lemma 8.5 there is a factor system on $\mathcal{X}$ consisting of all preimages of elements of the factor system $\mathfrak{F}$ for $\widetilde{S}_{\Gamma}$, where $\mathfrak{F}$ is the factor system associated to the collection of all subgraphs of Γ described in Proposition 8.3.

From now on we identify the 0–skeleton of $\widetilde{S}_{\Gamma}$ with $A(\Gamma)$, and regard the 1–cubes of $\widetilde{S}_{\Gamma}$ as labeled by an immersed hyperplane of C (they are naturally labeled by vertices of Γ, which are immersed hyperplanes of C).

Let us consider an element of $F \in \mathfrak{F}$ that intersects $\widetilde{\phi}(\mathcal{X})$ in, say, $g \in A(\Gamma)$. We want to show that $F' = F \cap \widetilde{\phi}(\mathcal{X})$ is the image via $\widetilde{\phi}$ of the lift $\widetilde{D}$ of some $D \in \bigcup_{\mathcal{A} \subseteq \mathcal{H}} C_{\mathcal{A}}$.

The 0–skeleton of F is $gA(\Lambda)$ for some subgraph Λ of Γ whose set of vertices will be denoted $\mathcal{A}$. By convexity of F' we know that we can connect g to any element

of the 0–skeleton of F' (that is to say $gA(\Lambda) \cap \widetilde{\phi}(\mathcal{X})$) by a path in the 1–skeleton of F' all whose 1–cubes are labeled by elements of $\mathcal{A}$. On the other hand, if we can connect g to, say, h by a path in the 1–skeleton of F' all whose 1–cubes are labeled by elements of $\mathcal{A}$, then $h \in gA(\Lambda)$. These facts easily imply that the 0–skeleton of F' coincides with the 0–skeleton of $\widetilde{\phi}(\widetilde{D})$ for some $D \in C_{\mathcal{A}}$, which in turn implies $\widetilde{\phi}(\widetilde{D}) = F \cap \widetilde{\phi}(\mathcal{X})$.

It can be similarly shown that for any $D \in \bigcup_{\mathcal{A} \subseteq \mathcal{H}} C_{\mathcal{A}}$, we have that $\phi(\widetilde{D})$ is of the form $F \cap \widetilde{\phi}(\mathcal{X})$ for some $F \in \mathfrak{F}$, and this concludes the proof. $\qquad\square$

The same proof goes through to show the following more general version of the corollary:

Corollary 8.9 *Let $\mathcal{X}$ be the universal cover of the special cube complex C, whose set of immersed hyperplanes is finite, and let Γ be the crossing graph of the immersed hyperplanes of C. If $\mathcal{R}$ is a rich collection of subgraphs of Γ, then the collection of all lifts of subcomplexes in $\bigcup_{\Lambda \in \mathcal{R}} C_{\Lambda^{(0)}}$ is a factor system for $\mathcal{X}$.*

8.2.2 A nonspecial example

Example 8.10 Let S, T be wedges of finitely many circles and let D be a compact, 2–dimensional, non-positively-curved cube complex with the following properties:

(1) $D^{(1)} = S \cup T$.

(2) The universal cover $\widetilde{D}$ of D is $\widetilde{S} \times \widetilde{T}$.

(3) $H = \pi_1 D$ has no nontrivial finite quotient.

Such D were constructed by Burger and Mozes [19], whose complex actually has simple fundamental group, and by Wise [71]. Let $\widetilde{\alpha} \to \widetilde{S}$ and $\widetilde{\beta} \to \widetilde{T}$ be nontrivial combinatorial geodesics mapping to immersed closed combinatorial paths $\alpha, \beta \to D$. Since each of $\widetilde{S}, \widetilde{T}$ is convex in $\widetilde{S} \times \widetilde{T}$, each of the maps $\alpha \to D$, $\beta \to D$ is a local isometry of cube complexes. Hence $[\alpha], [\beta] \in H - \{1\}$ respectively lie in the stabilizers of the hyperplanes $\widetilde{S} \times \{o\}$ and $\{o'\} \times \widetilde{T}$ of $\widetilde{S} \times \widetilde{T}$, where (o', o) is a basepoint chosen so that o', o are arbitrary midpoints of 1–cubes of $\widetilde{S}, \widetilde{T}$. Suppose moreover that $|\alpha| = |\beta| = r \geq 1$.

Let $R = [0, r] \times [0, 2]$. Regard R as a 2–dimensional non-positively-curved cube complex in the obvious way, and form a complex Z_0 from $R \sqcup D$ by identifying $\{0\} \times [0, 2]$ with $\{r\} \times [0, 2]$ and identifying the images of $[0, r] \times \{0\}$ and $[0, r] \times \{2\}$ with α and β, respectively. Since α and β are locally convex in both R and D, the complex Z_0 is non-positively-curved, and its fundamental group is $H *_{\langle [\alpha] \rangle^t = \langle [\beta] \rangle}$.

Let $A = [0,1] \times [0,r]$ with the obvious cubical structure, and attach A to Z_0 by identifying $\{0\} \times [0,r]$ with the "meridian" in R, ie the image of $[0,r] \times \{1\}$. Finally, attach a square by gluing a length-2 subpath of its boundary path to the image of $[0,1] \times \{0\} \cup [0,1] \times \{r\} \subset A$ in $A \cup Z_0$. The resulting complex is Z. Since C is locally convex in Z_0 and $\{0\} \times [0,r]$ is locally convex in A, the complex Z is a non-positively-curved cube complex. See Figure 3.
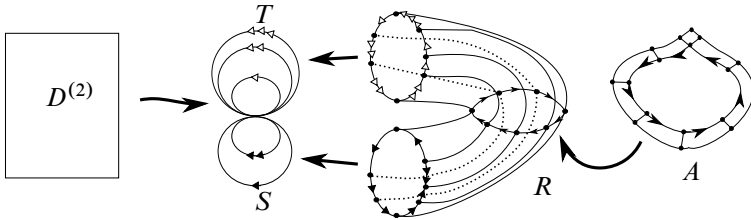

Figure 3: A complex Z with $r = 7$. The square at left represents the 2–skeleton of D. Note that A is the carrier of a self-intersecting hyperplane, A is attached to R by identifying the "inner" boundary component with the waist of R, and R is attached to $S \cup T$ by a local isometry of its boundary.

The following properties of Z will be needed in Section 11.

Proposition 8.11 *The universal cover $\widetilde{Z}$ of Z has a factor system.*

Proof Let Z' be constructed exactly as in the construction of Z in Example 8.10, except beginning with $S \times T$ instead of with the complex D. The 1–skeleta of D and $S \times T$ are identical, so the paths α, β exist in $S \times T$, rendering this construction possible. It is easily verified that the universal cover of Z' is isomorphic to $\widetilde{Z}$ and that Z' has a finite-sheeted special cover. The claim follows from Corollary 8.8. □

Proposition 8.12 *Let $G \leq_{\mathrm{fi}} \pi_1 Z$. Then there exists a hyperplane H of $\widetilde{D}$ and $g \in G$ such that gH crosses H.*

Proof Let $\overline{W} \to Z$ be the immersed hyperplane dual to the image of $[0,1] \times \{0\} \subset A$ in Z. Then $\overline{W}$ self-crosses. Moreover, if $\widehat{Z} \to Z$ is a finite-sheeted cover such that some component $\widehat{C}$ of the preimage of the meridian C has the property that $\widehat{C} \to C$ is bijective, then some lift of W to $\widehat{Z}$ is again a self-crossing immersed hyperplane. But C is homotopic into D, whose fundamental group has no finite quotients. Hence *any* finite cover of Z has a self-crossing immersed hyperplane. □

Having shown that there are many interesting groups with factor systems, below we show the utility of these systems. First, though, we propose the following question, to which we expect a positive answer; we believe that a proof of this will require significant work.

Question 8.13 Does every CAT(0) cube complex which admits a proper, cocompact group action contain a factor system?

8.3 Factored contact graphs

Definition 8.14 (factored contact graphs, projection to the factored contact graph) Let $\mathfrak{F}$ be a factor system for $\mathcal{X}$. For each convex subcomplex $F \subseteq \mathcal{X}$, the *factored contact graph* $\widehat{\mathcal{C}}F$ is obtained from $\mathcal{C}F$ as follows.

Let $F' \in \mathfrak{F}_F - \{F\}$ and suppose that F' is either parallel to a combinatorial hyperplane or has diameter $\geq \xi$, and let $[F']$ be its parallelism class. We add a vertex $v_{F'}$ to $\mathcal{C}F$, *corresponding to this parallelism class* $[F']$, and join $v_{F'}$ by an edge to each vertex of $\mathcal{C}F' \subset \mathcal{C}F$, ie each newly added vertex $v_{F'}$ gets connected to each vertex of $\mathcal{C}F$ which corresponds to a hyperplane of F that intersect F'. We repeat this for each such $F' \in \mathfrak{F}_F - \{F\}$. We emphasize that $\widehat{\mathcal{C}}F$ consists of $\mathcal{C}F$ together with a cone-vertex for each *parallelism class* of subcomplexes in $\mathfrak{F}_F - \{F\}$ that are parallel to combinatorial hyperplanes or have diameter $\geq \xi$ (or both). Also, observe that $\mathcal{C}F$ is an induced subgraph of $\widehat{\mathcal{C}}F$.

The *projection of $\mathcal{X}$ to $\widehat{\mathcal{C}}F$* is the map $\pi_F \colon \mathcal{X} \to 2^{\widehat{\mathcal{C}}F^{(0)}}$ obtained by composing the projection to $\mathcal{C}F$ given in Definition 4.1 with the inclusion $\mathcal{C}F \hookrightarrow \widehat{\mathcal{C}}F$. Recall that this map assigns to each point of $\mathcal{X}$ a clique in $\widehat{\mathcal{C}}F$ consisting of the vertices corresponding to the hyperplanes whose carriers contain the given point.

Remark 8.15 When $F \in \mathfrak{F}$ is a single 0–cube, $\mathcal{C}F = \varnothing$, so $\widehat{\mathcal{C}}F = \varnothing$. Hence $|2^{\widehat{\mathcal{C}}F^{(0)}}| = 1$ and π_F is defined in the obvious way. When the constant ξ from Definition 8.1 is at least 1, no $F \in \mathfrak{F}$ is a single point.

Lemma 8.16 *Let F, F' be parallel. Then there is a bijection $f \colon \mathfrak{F}_F \to \mathfrak{F}_{F'}$ such that $f(H)$ is parallel to H for all $H \in \mathfrak{F}_F$.*

Proof If $F'' \in \mathfrak{F}$ is contained in F, then there is a parallel copy F''' of F'' in F'. Provided either F'' is a combinatorial hyperplane or has diameter at least ξ, we have $F''' \in \mathfrak{F}$. The existence of f now follows easily. $\qquad\square$

The next lemma is immediate.

Lemma 8.17 *Let $F \in \mathfrak{F}$ and let F' be a convex subcomplex parallel to F. Then $\mathcal{C}F = \mathcal{C}F'$ and the bijection $f \colon \mathfrak{F}_F \to \mathfrak{F}_{F'}$ from Lemma 8.16 descends to a bijection of parallelism classes inducing an isomorphism $\widehat{\mathcal{C}}F \to \widehat{\mathcal{C}}F'$ that extends the identity $\mathcal{C}F \to \mathcal{C}F'$. Moreover, $\pi_F = \pi_{F'}$.*

Remark 8.18 (crossing, osculation and coning)   In the case where $\mathfrak{F}$ is the factor system obtained by closing the set of subcomplexes parallel to hyperplanes under projections with diameter $\geq \xi$, then $\widehat{\mathcal{C}}\mathcal{X}$ is quasi-isometric to $\mathcal{C}\mathcal{X}$. This is because we are coning off contact graphs of intersections of hyperplane-carriers, and these subgraphs are already contained in vertex-links in $\mathcal{C}\mathcal{X}$. More generally, let $H_I = \bigcap_{i \in I} H_i^+$ be an intersection of combinatorial hyperplanes. The hyperplanes of H_I have the form $W \cap H_I$, where W is a hyperplane that *crosses* each of the hyperplanes H_i. Coning off $\mathcal{C}(W^\pm \cap H_I)$ in $\mathcal{C}H_I$ does not affect the quasi-isometry type, since $\mathcal{C}(W^\pm \cap H_I)$ is the link of a vertex of $\mathcal{C}H_I$. It is when W is disjoint from H_I but exactly one of $W^\pm$ is not — ie precisely when W *osculates with each* H_i — that coning off $\mathcal{C}(W^+ \cap H_I) \subset \mathcal{C}H_I$ affects the geometry.

The crucial property of projections to factored contact graphs is:

Lemma 8.19 (bounded projections)   *Let $\mathfrak{F}$ be a factor system for $\mathcal{X}$ and let ξ be the constant from Definition 8.1. Let $F, F' \in \mathfrak{F}$. Then one of the following holds:*

(1) $\mathrm{diam}_{\widehat{\mathcal{C}}F}(\pi_F(F')) < \xi + 2$.

(2)   *F is parallel to a subcomplex of F'.*

Proof   If $\mathrm{diam}(\mathfrak{g}_F(F')) \geq \xi$, then either $\pi_F(F')$ is coned off in $\widehat{\mathcal{C}}F$ or $\mathfrak{g}_F(F') = F$. $\square$

We also note the following version of Proposition 4.2 for factored contact graphs:

Proposition 8.20 (bounded geodesic image II: the factoring)   *Let $F \in \mathfrak{F}$ and let $\gamma \to F$ be a combinatorial geodesic. Suppose $U \in \mathfrak{F}_F$ and that $\mathrm{d}_{\widehat{\mathcal{C}}F}(\pi_F(\gamma), \widehat{\mathcal{C}}U) \geq 1$. Then $\mathfrak{g}_U(\gamma)$ consists of a single 0–cube. Hence $\pi_U(\gamma)$ is a clique.*

Proof   Suppose that $\mathfrak{g}_U(\gamma)$ contains distinct 0–cubes x, y and let H be a hyperplane separating them. Then H crosses U and thus corresponds to a vertex of $\mathcal{C}U \subset \widehat{\mathcal{C}}U \subseteq \widehat{\mathcal{C}}F$. On the other hand, H cannot separate any 0–cube of γ from U and thus separates the endpoints of γ. Hence H is a vertex of $\pi_F(\gamma)$. $\square$

8.4   Hierarchy paths revisited

Let $\mathfrak{F}$ be a factor system for $\mathcal{X}$ and let $F \in \mathfrak{F}$. The vertices of $\widehat{\mathcal{C}}F$ are naturally associated to subcomplexes of $\mathcal{X}$: to each vertex V of $\mathcal{C}F$ ($\subset \widehat{\mathcal{C}}F$), we associate one of the two combinatorial hyperplanes $V^\pm$ bounding $\mathcal{N}(V)$. The remaining vertices are cone points corresponding to parallelism classes of subcomplexes in $\mathfrak{F}_F$. A path of $\widehat{\mathcal{C}}F$ is a sequence $(v_0, \ldots, v_n)$ of vertices with consecutive vertices adjacent.

A combinatorial path $\gamma \to F$ is *carried* by the path $(v_0, \ldots, v_n)$ of $\widehat{\mathcal{C}}F$ if $\gamma = \gamma_0 e_0 \gamma_1 e_1 \cdots \gamma_n e_n$, where for $0 \leq i \leq n$, we have $\gamma_i \to T_i$ with T_i associated to v_i and $|e_i| \leq 1$. In this situation, we say that the sequence $T_0, \ldots, T_n$ *carries* γ and *represents* $(v_0, \ldots, v_n)$.

Proposition 8.21 *Let* $v_0, v_1, \ldots, v_r$ *be a path in* $\widehat{\mathcal{C}}\mathcal{X}$ *and let* $x \in T_0$, $y \in T_r$ *be* 0–*cubes, where* T_0, T_r *are subcomplexes associated to* v_0, v_r, *respectively. Then for* $1 \leq i \leq r - 1$, *we can choose for each vertex* v_i *an associated subcomplex* T_i *so that* $T_0, \ldots, T_r$ *carries a path in* $\mathcal{X}$ *joining* x *to* y.

Proof We argue by induction on r. If $r = 0$, then the claim follows by path-connectedness of T_0. Now let $r \geq 1$. Suppose first that v_r is a cone-vertex, so that v_{r-1} is a hyperplane-vertex, and let V be the corresponding hyperplane. Since V crosses T_r, either combinatorial hyperplane T_{r-1} bounding the carrier of V intersects T_r. There exists a 0–cube $y' \in T_{r-1} \cap T_r$. On the other hand, if v_r is a hyperplane-vertex, then T_r is a specified combinatorial hyperplane parallel to a hyperplane crossing each subcomplex T_{r-1} associated to v_{r-1}, if v_{r-1} is a cone-vertex. If v_{r-1} is also a hyperplane-vertex, then at least one of the possible choices of T_{r-1} is a combinatorial hyperplane intersecting the combinatorial hyperplane T_r. Again, we have a 0–cube $y' \in T_r \cap T_{r-1}$. In either case, by induction, we can choose $T_1, \ldots, T_{r-2}$ to carry a path joining x to y', which we concatenate with a path in T_r joining y' to y. The path e_{r-2} is nontrivial if v_{r-1} corresponds to a hyperplane separating T_r from every possible choice of T_{r-2}. $\qquad\square$

Remark 8.22 (explicit description of paths carried by geodesics) In the proof of Proposition 8.21, we used the fact that a path $v_0, v_1, \ldots, v_r$ in $\widehat{\mathcal{C}}F$ has the property that, if v_i is a cone-vertex, associated to some $U \in \mathfrak{F}_F$, then $v_{i\pm 1}$ are hyperplane vertices, associated to hyperplanes $H_{i\pm 1}$ that cross U, so that all four combinatorial hyperplanes $H_{i\pm 1}^{\pm}$ intersect U. Hence we can and shall *always* assume that $|e_{i\pm 1}| = 0$ when v_i is a cone-vertex; see Figure 4. This enables us to use the proof of Proposition 3.1 verbatim in the proof of Proposition 8.23.

In view of the previous proposition, we will use the same notation for a vertex of $\widehat{\mathcal{C}}F$ as for some representative subcomplex. A *hierarchy path* is a geodesic of F that is carried by a geodesic of $\widehat{\mathcal{C}}F$.

Proposition 8.23 *Let* $F \in \mathfrak{F}$. *Then any two* 0–*cubes* $x, y \in F$ *are joined by a hierarchy path.*

Proof Let τ be a combinatorial geodesic of $\mathcal{X}$ joining x to y; since F is convex, $\tau \subseteq F$. Let $T \in \mathfrak{F}_F$ contain x (such a T exists because every 0–cube of F is
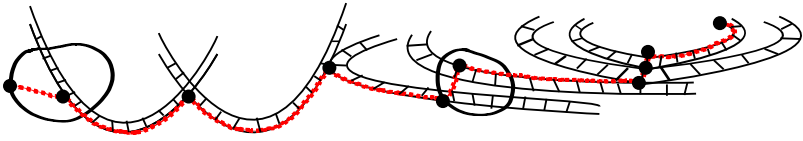
Figure 4: The two round regions are elements of $\mathfrak{F}$; the ladders represent carriers of hyperplanes. The dotted path is carried by a geodesic in the factored contact graph. Transitions between various γ_i and e_i are indicated by bold vertices. Notice that there are numerous such paths, but we choose paths with e_i trivial whenever possible.

contained in a combinatorial hyperplane of F) and let $T' \in \mathfrak{F}_F$ contain y. Let $T = T_0, T_1, \ldots, T_n = T'$ be a geodesic of $\widehat{\mathcal{C}}F$. Then we can choose the T_i in their parallelism classes so that $\mathcal{N}_1(T_i) \cap \mathcal{N}_1(T_{i+1}) \neq \varnothing$ for $0 \leq i \leq n-1$. Using Proposition 8.21, for each i, let $\gamma_i \to T_i$ be a geodesic segment, chosen so that $\gamma_1 e_1 \cdots \gamma_n e_n$ is a piecewise-geodesic path joining x to y, with each $|e_i| \leq 1$ and $|e_i| = 0$ except possibly if γ_i, γ_{i+1} lie in disjoint combinatorial hyperplanes representing vertices of $\mathcal{C}F$ (ie non-cone-vertices). The disc diagram argument from the proof of Proposition 3.1 can now be repeated verbatim to complete the proof, because $\mathfrak{F}_F$ satisfies property (4) from Definition 8.1. $\qquad\square$

8.5 Factored contact graphs are quasi-trees

In this section, we will use the following criterion of Manning, which determines if a geodesic metric space is quasi-isometric to a tree:

Proposition 8.24 (bottleneck criterion [48]) *Let (Z, d) be a geodesic metric space. Suppose that there exists $\delta \geq 0$ such that for all $x, y \in Z$, there exists $m = m(x, y)$ such that $d(x, m) = d(y, m) = \frac{1}{2}d(x, y)$ and every path joining x to y intersects the closed δ–ball about m. Then (Z, d) is $(26\delta, 16\delta)$–quasi-isometric to a tree.*

Proposition 8.25 *Let $F \in \mathfrak{F}$. Then $\widehat{\mathcal{C}}F$ is $(78, 48)$–quasi-isometric to a simplicial tree.*

Proof Denote by $\mathcal{H}$ the set of vertices of $\widehat{\mathcal{C}}F$ corresponding to hyperplanes of F and by $\mathcal{V}$ the set of cone-vertices. Let T, T' be complexes associated to vertices of $\widehat{\mathcal{C}}F$ (regarded as representatives, to be chosen, of parallelism classes of subcomplexes in $\mathfrak{F}_F$). Let $T = T_0, T_1, \ldots, T_n = T'$ be a geodesic of $\widehat{\mathcal{C}}F$ joining T to T' and let m be the midpoint of this geodesic. Suppose, moreover, that $T_0, \ldots, T_n$ carries a hierarchy path $\gamma = \gamma_0 \cdots \gamma_n$ such that each hyperplane intersecting γ separates T from T'; indeed, we can choose γ to connect a point $p \in \mathfrak{g}_T(T')$ to its gate $\mathfrak{g}_{T'}(p)$. We may

assume that $n \geq 4$, for otherwise Proposition 8.24 is satisfied by m with $\delta = \frac{3}{2}$. Hence, since every vertex adjacent to an element of $\mathcal{V}$ is in $\mathcal{H}$, there exists i so that T_i is a combinatorial hyperplane and $d_{\widehat{C}F}(T_i, m) \leq 1$. Let T_i' be the hyperplane such that T_i is in the image of $T_i' \times \{\pm 1\}$ under $T_i' \times [-1, 1] \cong \mathcal{N}(T_i') \hookrightarrow F$. Since every hyperplane crossing γ_i separates T from T', we have that either T_i' separates T from T', or T_i' crosses a hyperplane that separates T from T'. Hence, there is a hyperplane H such that H separates T, T' and $d_{\widehat{C}F}(H, m) \leq 2$. Let $T = S_0, \ldots, S_k = T'$ be another path in $\widehat{C}F$ joining T to T'. Then there exists j such that either S_j is a combinatorial copy of H, or H crosses S_j. Thus $d_{\widehat{C}F}(m, S_j) \leq 3$, and the claim follows from Proposition 8.24. $\qquad\square$

9 The distance formula

The main theorem of this section is:

Theorem 9.1 (distance formula) *Let $\mathcal{X}$ be a CAT(0) cube complex and let $\mathfrak{F}$ be a factor system. Let $\overline{\mathfrak{F}}$ contain exactly one representative of each parallelism class in $\mathfrak{F}$. Then there exists $s_0 \geq 0$ such that for all $s \geq s_0$, there are constants $K \geq 1$, $C \geq 0$ such that, for all $x, y \in \mathcal{X}^{(0)}$,*

$$d_{\mathcal{X}}(x, y) \asymp_{K,C} \sum_{F \in \overline{\mathfrak{F}}} \{\!\!\{ d_{\widehat{C}F}(\pi_F(x), \pi_F(y)) \}\!\!\}_s.$$

The rest of this section is devoted to proving Theorem 9.1. Throughout, $\mathcal{X}$ and $\mathfrak{F}$ are as in the statement of the theorem. The constants $\xi \geq 1$, $\Delta \geq 1$ are those from Definition 8.1. For convenience, if $F \in \mathfrak{F}$ and $x, y \in \mathcal{X}$, we write $d_{\widehat{C}F}(x, y) := d_{\widehat{C}F}(\pi_F(x), \pi_F(y))$.

9.1 Projection lemmas

Lemma 9.2 *Let $\mathcal{X}$ be a CAT(0) cube complex, let $A \subseteq B \subseteq \mathcal{X}$ be convex subcomplexes, and let $x, y \in \mathcal{X}$ be 0–cubes. Then*

$$d_{\widehat{C}A}(\pi_A(x), \pi_A(y)) \leq d_{\mathcal{X}}(\mathfrak{g}_A(x), \mathfrak{g}_A(y)) \leq d_{\mathcal{X}}(\mathfrak{g}_B(x), \mathfrak{g}_B(y)).$$

Proof The first inequality follows since projection to the factored contact graph is distance nonincreasing. The second inequality is Lemma 2.1. $\qquad\square$

Lemma 9.3 *Let $x, y \in \mathcal{X}^{(0)}$ and let $T_0, T_1, \ldots, T_r$ represent a geodesic in $\widehat{C}\mathcal{X}$ with $x \in T_0$, $y \in T_r$. For $0 \leq i \leq r$, let $x_i = \mathfrak{g}_{T_i}(x)$, $y_i = \mathfrak{g}_{T_i}(y)$. Then*

$$d_{\mathcal{X}}(x, y) - d_{\widehat{C}\mathcal{X}}(x, y) \leq \sum_{i=0}^{r} d_{T_i}(x_i, y_i) \leq 3 d_{\mathcal{X}}(x, y).$$

Proof For $0 \leq i \leq r$, let $\mathcal{S}_i$ be the set of hyperplanes separating x_i, y_i, so that $|\mathcal{S}_i| = d_{\mathcal{X}}(x_i, y_i) = d_{T_i}(x_i, y_i)$. For each i, observe that if the hyperplane U separates x from y and crosses T_i, then U separates x_i from y_i. Otherwise, U would separate x_i from x (say), but the only such hyperplanes are those separating x from T_i, so this situation would contradict the fact that U crosses T_i. Hence the set of hyperplanes separating x from y and crossing some T_i is $\bigcup_{i=0}^r \mathcal{S}_i$. Each hyperplane H separating x, y either crosses some T_i, or there exists i so that $\mathcal{N}(H)$ contains T_i in one of the bounding copies of H. Hence

$$d_{\mathcal{X}}(x, y) \leq \left| \bigcup_{i=0}^r \mathcal{S}_i \right| + r \leq \sum_{i=0}^r |\mathcal{S}_i| + r,$$

which establishes the first inequality.

For some $i \neq j$, let $U \in \mathcal{S}_i \cap \mathcal{S}_j$. Then $|i - j| \leq 2$, since otherwise T_i, U, T_j would provide a shortcut contradicting the fact that $d_{\widehat{\mathcal{C}}(\mathcal{X}, \mathcal{P})}(T_0, T_r) = r$. Hence at most three elements of $\{\mathcal{S}_i\}_{i=0}^r$ contain U. Thus $\sum_{i=0}^r |\mathcal{S}_i| \leq 3 |\bigcup_{i=0}^r \mathcal{S}_i|$, and the second inequality follows. $\qquad \square$

Proposition 9.4 (large link lemma) *Let $T_0, T_1, \ldots, T_r$ be a geodesic in $\widehat{\mathcal{C}}\mathcal{X}$ between 0–cubes $x \in T_0$ and $y \in T_r$. Let $F \in \mathfrak{F} - \{\mathcal{X}\}$ have the property that $d_{\widehat{\mathcal{C}}F}(x, y) \geq 4\xi + 10$. Then there exists $i \in \{0, 1, \ldots, r\}$ such that F is parallel to some $F' \in \mathfrak{F}_{T_i}$. Moreover, any geodesic of $\mathcal{X}$ contained in $\bigcup_i T_i$ that joins x, y passes through a subcomplex of T_i parallel to F'.*

Proof Recall that we can choose T_i within its parallelism class for $1 \leq i \leq r - 1$ so that $\mathcal{N}_1(T_i) \cap T_{i+1} \neq \varnothing$. We first exhibit $J \leq 3$ and $i \leq r$ such that for some j satisfying $0 \leq j \leq J$ we have $\mathrm{diam}_{\widehat{\mathcal{C}}F}(\pi_F(T_{i+j})) \geq \xi$ and for all i' satisfying $i' < i$ or $i' > i + J$ we have $\mathrm{diam}_{\widehat{\mathcal{C}}F}(\pi_F(T_{i'})) = 0$.

Let V be a hyperplane separating $\mathfrak{g}_F(x)$ from $\mathfrak{g}_F(y)$. Then V intersects F and hence V cannot separate x or y from their gates in F. Hence V must separate x from y; and thus V intersects some $\mathcal{N}_1(T_i)$. Either V separates $\mathfrak{g}_{T_i}(x)$ from $\mathfrak{g}_{T_i}(y)$, or V is the unique hyperplane separating T_i from $T_{i\pm 1}$. Let $\mathcal{G}$ be the set of hyperplanes V of the former type.

Let $V \in \mathcal{G}$. Then V intersects T_i for at least one and at most three values of i. If some other hyperplane $V' \in \mathcal{G}$ separates $\mathfrak{g}_F(x)$ from $\mathfrak{g}_F(y)$, then V' separates $\mathfrak{g}_{T_j}(x)$ from $\mathfrak{g}_{T_j}(y)$ for at least one and at most three values of j, and for any such i, j, we have $|i - j| \leq 4$; otherwise T_i, V, F, V', T_j would be a shortcut from T_i to T_j, since $F \neq \mathcal{X}$ represents a vertex in $\widehat{\mathcal{C}}\mathcal{X}$. Let $i, i+1, \ldots, i + J$ be the indices for which T_{i+j}

is crossed by a hyperplane separating $\mathfrak{g}_F(x)$ from $\mathfrak{g}_F(y)$; equivalently, the indices such that $\mathfrak{g}_{T_k}(F)$ is trivial if $k < i$ or $k > i + J$.

The graph $\pi_F\left(\bigcup_{j=0}^{J} \mathfrak{g}_F(T_{i+j})\right) = \pi_F\left(\bigcup_{j=0}^{r} \mathfrak{g}_F(T_i)\right)$ has diameter at least $4\xi + 10$, since it contains $\pi_F(x), \pi_F(y)$, whence $\mathrm{diam}_{\widehat{\mathcal{C}}F}(\mathfrak{g}_F(T_{i+j})) \geq \xi$ for some j. Indeed, there is at most one hyperplane separating T_{i+j} from T_{i+j+1}, whence at most one hyperplane separates their gates in F. We then have that F is parallel to a subcomplex of T_{i+j} by Lemma 8.19.

Let $\gamma \subset \cup_k T_k$ be a geodesic joining x, y and let $\gamma_{i+j} = \mathfrak{g}_{T_{i+j}}(\gamma)$. Then γ_{i+j} has a subpath τ joining $\mathfrak{g}_{T_{i+j}}(\mathfrak{g}_F(x))$ to $\mathfrak{g}_{T_{i+j}}(\mathfrak{g}_F(y))$. The path τ is parallel to a path in F and hence belongs to some parallel copy of F in T_{i+1}. $\qquad\square$

Lemma 9.5 *There exists $s_0 \geq 0$ such that for all $x, y \in \mathcal{X}^{(0)}$ and any hierarchy path $\gamma_0 e_0 \cdots \gamma_r e_r$ joining x to y and carried by a geodesic $T_0, \ldots, T_r$ of $\widehat{\mathcal{C}}\mathcal{X}$, one of the following holds for all $F \in \mathfrak{F}$:*

(1) *there exist $i \leq r$ and $F' \in \mathfrak{F}_{T_i}$ such that F is parallel to F';*

(2) $\mathrm{d}_{\widehat{\mathcal{C}}F}(x, y) < s_0$;

(3) $F = \mathcal{X}$.

Proof Let $s_0 = 4\xi + 10$ and let $F \neq \mathcal{X}$ be an element of $\mathfrak{F}$ with $\mathrm{d}_{\widehat{\mathcal{C}}F}(x, y) \geq s_0$. By Proposition 9.4, there exists i so that F is parallel to some $F' \in \mathfrak{F}_{T_i}$, as required. $\quad\square$

9.2 Proof of the distance formula

We have now assembled all ingredients needed for:

Proof of Theorem 9.1 For each $F \in \overline{\mathfrak{F}}$, let $\overline{\overline{\mathfrak{F}}}_F$ consist of exactly one element from each parallelism class in $\mathfrak{F}_F$. We will argue by induction on Δ, which we recall is the maximal number of elements in the factor system which can contain any given vertex.

Base case When $\Delta = 1$, the fact that $\mathcal{X}$ and each combinatorial hyperplane belongs to $\mathfrak{F}$ ensures that $\mathcal{X}$ consists of a single 0–cube, so we are summing over the empty set of projections, and we are done.

Induction hypothesis Assume $\Delta \geq 2$. For each $F \in \overline{\mathfrak{F}} - \{\mathcal{X}\}$, the set $\mathfrak{F}_F$ is a factor system in F. Every 0–cube of F is contained in at most $\Delta - 1$ elements of $\mathfrak{F}_F$, since each 0–cube of F lies in at most Δ elements of $\mathfrak{F}$, one of which is $\mathcal{X}$ (which is not contained in F). We can therefore assume, by induction on Δ, that for all $s \geq s_0$, where $s_0 \geq 2$ is the constant from Lemma 9.5, there exist $K' \geq 1$, $C' \geq 0$ so that for all $F \in \overline{\mathfrak{F}} - \{\mathcal{X}\}$ and all $x, y \in \mathcal{X}$ we have

$$\mathrm{d}_F(\mathfrak{g}_F(x), \mathfrak{g}_F(y)) \asymp_{K',C'} \sum_{T \in \overline{\overline{\mathfrak{F}}}_F} \{\!\{\mathrm{d}_{\widehat{\mathcal{C}}T}(\mathfrak{g}_F(x), \mathfrak{g}_F(y))\}\!\}_s.$$

Choosing a hierarchy path By Proposition 8.23, there exists a hierarchy path $\gamma = \gamma_0 e_0 \cdots \gamma_r e_r$ that joins x to y and is carried on a geodesic $T_0, \ldots, T_r$ in $\widehat{\mathcal{C}}\mathcal{X}$ from T_0 to T_r. For each i, let $x_i = \mathfrak{g}_{T_i}(x)$ and $y_i = \mathfrak{g}_{T_i}(y)$. Recall that each γ_i lies in $\mathcal{N}_1(T_i)$. At the cost of adding 1 to our eventual multiplicative constant K, we can assume that each e_i is trivial.

Enumeration of the nonvanishing terms Let $s \geq s_0$ and let $F \in \mathfrak{F}$ be such that $\mathsf{d}_{\widehat{\mathcal{C}}F}(x, y) \geq s$. By Lemma 9.5, either F is parallel to some $F' \in \mathfrak{F}_{T_i}$ for some i, or $F = \mathcal{X}$.

Let $\mathfrak{F}_{\{T_i\}}$ be the set of all $F \in \mathfrak{F}$ that are parallel to a proper subcomplex of some T_i, and let $\overline{\mathfrak{F}}_{\{T_i\}}$ contain exactly one representative for each parallelism class of elements of $\mathfrak{F}_{\{T_i\}}$. Recalling that $\mathsf{d}_{\widehat{\mathcal{C}}F}(x, y) = \mathsf{d}_{\widehat{\mathcal{C}}F'}(x, y)$ when F is parallel to F', we get

$$\sum_{F \in \overline{\mathfrak{F}}} \{\!\{\mathsf{d}_{\widehat{\mathcal{C}}F}(x, y)\}\!\}_s = \mathbf{I} + \mathbf{II} + \{\!\{r\}\!\}_s,$$

where

$$\mathbf{I} = \sum_{i=0}^{r} \{\!\{\mathsf{d}_{\widehat{\mathcal{C}}T_i}(x, y)\}\!\}_s,$$

$$\mathbf{II} = \sum_{F \in \overline{\mathfrak{F}}_{\{T_i\}}} \{\!\{\mathsf{d}_{\widehat{\mathcal{C}}F}(x, y)\}\!\}_s \asymp_{3,0} \sum_{i} \sum_{F \in \overline{\mathfrak{F}}_{T_i} - \{T_i\}} \{\!\{\mathsf{d}_{\widehat{\mathcal{C}}F}(x, y)\}\!\}_s.$$

The estimate of **II** follows since no factor is parallel to both T_i and $T_{i'}$ when $|i - i'| \geq 3$.

The upper bound By the inductive hypothesis,

$$\mathbf{I} + \mathbf{II} \asymp_{K', C'} \sum_{i=0}^{r} \mathsf{d}_{T_i}(\mathfrak{g}_{T_i}(x), \mathfrak{g}_{T_i}(y)).$$

Hence, by Lemma 9.3,

$$\frac{\mathsf{d}_{\mathcal{X}}(x, y) - r}{3K'} - C' \leq \mathbf{I} + \mathbf{II},$$

and we obtain the upper bound

$$\begin{aligned}
\mathsf{d}_{\mathcal{X}}(x, y) &\leq 3K'(\mathbf{I} + \mathbf{II} + r) + 3C'K' \\
&= 3K' \sum_{F \in \overline{\mathfrak{F}} - \{\mathcal{X}\}} \{\!\{\mathsf{d}_{\widehat{\mathcal{C}}F}(x, y)\}\!\}_s + 3K'\mathsf{d}_{\widehat{\mathcal{C}}\mathcal{X}}(x, y) + 3C'K' \\
&\asymp_{1,s} \sum_{F \in \overline{\mathfrak{F}} - \{\mathcal{X}\}} \{\!\{\mathsf{d}_{\widehat{\mathcal{C}}F}(x, y)\}\!\}_s + 3K'\{\!\{\mathsf{d}_{\widehat{\mathcal{C}}\mathcal{X}}(x, y)\}\!\}_s + 3C'K'.
\end{aligned}$$

The lower bound We have that $d_{\mathcal{X}}(x, y) \geq r$ and we have the estimate $\mathbf{I} + \mathbf{II} \leq 3K'd_{\mathcal{X}}(x, y) + C'r$, by the induction hypothesis and Lemma 9.3. Hence

$$\mathbf{I} + \mathbf{II} + \{\!\{r\}\!\}_s \leq (3K' + C' + 1)d_{\mathcal{X}}(x, y),$$

and we are done. $\qquad\square$

10 Projection of parallelism classes

Let $\mathcal{X}$ be CAT(0) cube complex with a factor system $\mathfrak{F}$, and let $\overline{\mathfrak{F}}$ be the set of parallelism classes of elements of $\mathfrak{F}$. We have previously defined projections from $\mathcal{X}$ to factored contact graphs of elements of $\mathfrak{F}$, and noted that $\pi_F \colon \mathcal{X} \to \hat{\mathcal{C}}F$ is independent of the choice of representative of the parallelism class $[F]$. However, care must be taken in order to define projections of parallelism classes:

Proposition 10.1 (projections of parallelism classes are bounded or cover) *Let $F', F'' \in \mathfrak{F}$ with F' parallel to F''. Then one of the following holds for each $F \in \mathfrak{F}$:*

(1) $d_{\hat{\mathcal{C}}F}(\pi_F(F'), \pi_F(F'')) \leq \xi + 2$, *where ξ is a constant of the factor system;*

(2) *there is a cubical isometric embedding $F' \times F \to \mathcal{X}$ such that for some $f', f'' \in F$, the subcomplexes F', F'' are the images of $F' \times \{f'\}$, $F' \times \{f''\}$, respectively, and for some $f \in F'$, the subcomplex F is the image of $\{f\} \times F$. In particular, $\bigcup_{F_1 \in [F']} \pi_F(F_1) = \hat{\mathcal{C}}F$.*

Moreover, whether assertion (1) or (2) holds for F', F'' with respect to F is independent of the choice of parallelism class representative of F.

Proof By Lemma 2.4, there is a convex subcomplex B and a cubical isometric embedding $F' \times B \to \mathcal{X}$ with the following properties: B is the convex hull of a shortest geodesic $[0, b]$ joining F' to F'' and F', F'' are respectively the images of $F' \times \{0\}$ and $F' \times \{b\}$, and in fact $[0, b]$ joins $\mathfrak{g}_{F'}(F'')$ to $\mathfrak{g}_{F''}(F')$ (the fact that $[0, b]$ is shortest and joins $\mathfrak{g}_{F'}(F'')$ to $\mathfrak{g}_{F''}(F')$ follows from the fact that it only crosses the hyperplanes that separate F' from F'', which is part of Lemma 2.4). Since B is, at minimum, contained in a combinatorial hyperplane (indeed, cubical isometric embeddings take hyperplanes to hyperplanes), it is contained in some minimal $Y \in \mathfrak{F}$. If $b \leq \xi$, then the first assertion holds. Hence suppose $b \geq \xi$. Then either $\pi_F(Y)$ has diameter at most $\xi + 2$, in which case the first assertion again holds, or by Lemma 8.19, F is parallel to a subcomplex F_1 of Y containing B. By minimality of Y, we have $F_1 = Y$. Similarly, for all $f \in F'$, we have a parallel copy $\{f\} \times F_1$ of F containing $\{f\} \times B$. Item (2) follows. $\qquad\square$

Definition 10.2 (orthogonal) Let $[F_1], [F_2] \in \overline{\mathfrak{F}}$. If there are $F_1', F_1'' \in [F_1]$ such that Proposition 10.1(2) holds for F_1', F_1'' projected to F_2, then $[F_1]$ and $[F_2]$ (and F_1 and F_2) are said to be *orthogonal*.

Remark 10.3 Lemma 8.19 implies that either F is parallel to a subcomplex of F' or $\pi_F(F')$ has uniformly bounded diameter. In the latter case, Proposition 10.1 implies that either $\bigcup_{F_1 \in [F']} \pi_F(F_1)$ is uniformly bounded, or $[F]$ and $[F_1]$ are orthogonal. In fact, Lemma 8.19 says that if F is not parallel to a subcomplex of F', then $\pi_F(F'')$ has diameter bounded by $\xi + 2$ for each F'' parallel to F'. On the other hand, whenever $\pi_F(F'')$ has diameter bounded by $\xi + 2$ for each F'' parallel to F', Proposition 10.1 implies that either $\bigcup_{F_1 \in [F']} \pi_F(F_1)$ has diameter bounded by $3\xi + 6$ or $[F]$ and $[F']$ are orthogonal.

Motivated by the remark, we give the following two definitions. Definition 10.5 provides projections between parallelism classes; as explained in Remark 10.3, Proposition 10.1 gives conditions ensuring that these projections are coarsely well-defined.

Definition 10.4 (transverse) Parallelism classes $[F], [F'] \in \overline{\mathfrak{F}}$ are *transverse* if they are not orthogonal and if F is not parallel to a subcomplex of F', or vice versa, for some (hence all) $F \in [F]$, $F' \in [F']$.

Definition 10.5 (projection of parallelism classes, projection distances) Let $[F] \in \overline{\mathfrak{F}}$. The *projection to $[F]$ (really to $\widehat{C}F$, which depends only on the parallelism class)* is the map $\overline{\mathfrak{F}} - \{[F]\} \to 2^{\widehat{C}F}$ given by $\pi_{[F]}([F']) = \bigcup_{F_1 \in [F']} \pi_F(F_1)$.

For each $[Y] \in \overline{\mathfrak{F}}$, define a function $\mathsf{d}_Y^\pi \colon (\overline{\mathfrak{F}} - \{[Y]\})^2 \to [0, \infty]$ by

$$\mathsf{d}_Y^\pi([F], [F']) = \operatorname{diam}_{\widehat{C}Y}\big(\pi_{[Y]}([F]) \cup \pi_{[Y]}([F'])\big).$$

We fix, until the end of the section, any subset $\overline{\mathfrak{F}}_{\mathrm{tr}} \subset \overline{\mathfrak{F}}$ with the property that for each distinct pair $[F], [F'] \in \overline{\mathfrak{F}}_{\mathrm{tr}}$ we have that $[F]$ is transverse to $[F']$.

Observe that for each $F \in \overline{\mathfrak{F}}_{\mathrm{tr}}$, the restriction of d_F^π to $\overline{\mathfrak{F}}_{\mathrm{tr}}$ takes uniformly bounded values. From the definition, it is obvious that $\mathsf{d}_F^\pi([F'], [F'']) = \mathsf{d}_F^\pi([F''], [F'])$ and $\mathsf{d}_F^\pi([F'], [F'']) + \mathsf{d}_F^\pi([F''], [F''']) \geq \mathsf{d}_F^\pi([F'], [F'''])$ for all $[F], [F'], [F''], [F'''] \in \overline{\mathfrak{F}}_{\mathrm{tr}}$. Moreover, the fact that the right-hand side of the distance formula (Theorem 9.1) is finite shows that, for a suitable η, we have $|\{F : \mathsf{d}_F^\pi([F'], [F'']) \geq \eta\}| < \infty$ for all $[F'], [F''] \in \overline{\mathfrak{F}}_{\mathrm{tr}}$.

Proposition 10.6 *For all $[F] \neq [F']$ in $\overline{\mathfrak{F}}_{\mathrm{tr}}$, we have $\operatorname{diam}_{\widehat{C}F}(\pi_F(F')) \leq \xi + 2$.*

Proof This is just Lemma 8.19 combined with the fact that F cannot be parallel to a subcomplex of F' by our assumptions on $\overline{\mathfrak{F}}_{\mathrm{tr}}$. □

The next result is the analogue for cubical groups of the inequality that Behrstock established in [3] for the mapping class group. Versions of this inequality in other contexts have appeared in [11; 12; 61; 66]

Proposition 10.7 (Behrstock inequality) *For all $[F], [F'], [F''] \in \overline{\mathfrak{F}}_{tr}$,*

$$\min\{\mathsf{d}_F^\pi([F'], [F'']), \mathsf{d}_{F'}^\pi([F], [F''])\} \leq 3\xi + 6.$$

Proof Let F, F', F'' satisfy $\mathsf{d}_F^\pi([F'], [F'']) > 3\xi + 6$. Let $x \in F'$. We claim that there exists Y parallel to F and $y \in Y$ so that $\mathfrak{g}_{F''}(x) = \mathfrak{g}_{F''}(y)$, which will conclude the proof.

Let γ be a geodesic from x to $\mathfrak{g}_{F''}(x)$. By Lemma 10.8, there exists Y parallel to F such that γ intersects Y, say at y. Since y is on γ, we have $\mathfrak{g}_{F''}(x) = \mathfrak{g}_{F''}(y)$, as required. $\square$

Lemma 10.8 *Let $F \in \mathfrak{F}$ and let $x, y \in \mathcal{X}^{(0)}$ satisfy $\mathsf{d}_{\widehat{C}F}(\pi_F(x), \pi_F(y)) > 2\xi + 4$. Then any geodesic from x to y enters a parallel copy of F.*

Proof Let γ be a geodesic from x to y. By Lemma 8.17 — which implies

$$\mathsf{d}_{\widehat{C}F}(\pi_F(a), \pi_F(b)) = \mathsf{d}_{\widehat{C}F'}(\pi_{F'}(a), \pi_{F'}(b))$$

for all $a, b \in \mathcal{X}$ whenever F, F' are parallel — we can assume that F was chosen within its parallelism class so that $\mathsf{d}_{\mathcal{X}}(F, \gamma)$ is minimal.

Observe that no hyperplane separates γ from F. Indeed, suppose that H is such a hyperplane. Then $\mathfrak{g}_F(x), \mathfrak{g}_F(y) \subset \mathfrak{g}_F(H)$. By our assumption that $\mathsf{d}_{\widehat{C}F}(\pi_F(x), \pi_F(y)) > 2\xi + 4$, we have $\mathrm{diam}_{\widehat{C}F}(\mathfrak{g}_F(H)) > 2\xi + 4 > \xi + 2$, whence, by Lemma 8.19, F is parallel to a subcomplex F' of the combinatorial hyperplane H^- on $\mathcal{N}(H)$ which is separated from F by H. For any $z \in \gamma$, any hyperplane separating z from F' separates z from F, whence $\mathsf{d}_{\mathcal{X}}(\gamma, F') < \mathsf{d}_{\mathcal{X}}(\gamma, F)$, contradicting our choice of F.

Let $\mathfrak{L}$ be the set of hyperplanes separating x from $\{y\} \cup F$ and let $\mathfrak{R}$ be the set of hyperplanes separating y from $\{x\} \cup F$. If $\mathfrak{L} = \varnothing$, then either $x \in F$, since no hyperplane separates x and y from F (for otherwise, by convexity of halfspaces, that hyperplane would separate γ from F). Similarly, $\mathfrak{R} \neq \varnothing$.

Let $L \in \mathfrak{L}$ be closest to y and let $R \in \mathfrak{R}$ be closest to x. Suppose that $L \cap R = \varnothing$. Then there exists a 0–cube $z \in \gamma$ so that no hyperplane in $\mathfrak{L} \cup \mathfrak{R}$ separates z from F. Hence any hyperplane separating z from F separates $\{x, y\}$, and thus γ, from F, which is impossible.

Thus L and R cross. Since $\mathfrak{g}_F(x), \mathfrak{g}_F(y) \in \mathfrak{g}_F(L) \cup \mathfrak{g}_F(R)$, arguing as above yields, say, $\mathrm{diam}_{\widehat{C}F}(\mathfrak{g}_F(L)) > \xi + 2$, so F is parallel into L, violating our choice of F as above. $\square$

Observe that Propositions 10.6 and 10.7, together with the discussion preceding them, imply the following, which we record for later convenience:

Corollary 10.9 (BBF axioms) *Let $\mathfrak{F}$ be a factor system and let $\overline{\mathfrak{F}}$ be the set of parallelism classes in $\mathfrak{F}$. Let $\overline{\mathfrak{F}}_{\mathrm{tr}}$ be a subset of $\overline{\mathfrak{F}}$ such that $[F], [F']$ are transverse for all distinct $[F], [F'] \in \overline{\mathfrak{F}}_{\mathrm{tr}}$. Then $\{\mathrm{d}_Y^{\pi} : [Y] \in \overline{\mathfrak{F}}_{\mathrm{tr}}\}$ satisfies Axioms (0)–(4) of [11, Section 2.1].*

11 Efficient embeddings into products of trees

We now let $\mathcal{X}$ be a CAT(0) cube complex with a proper cocompact action by a group G, and we let $\mathfrak{F}$ be a G–invariant factor system (recall that the existence of *any* factor system ensures the existence of a G–invariant one). In this section we produce a very particular G–equivariant quasi-isometric embedding of $\mathcal{X}$ into the product of finitely many quasi-trees. We begin with the following fact, which is well-known in the study of cubical groups:

Proposition 11.1 (embeddings in products of trees) *Suppose that there exists $G' \leq_{\mathrm{fi}} G$ such that no hyperplane in $\mathcal{X}$ crosses its G'–translates. Then there exist a finite collection $\{T_1, \ldots, T_k\}$ of simplicial trees, a G–action on $\prod_{i=1}^k T_i$, and a G–equivariant isometric embedding $\mathcal{X} \to \prod_{i=1}^k T_i$.*

Conversely, suppose that G acts properly and cocompactly on a CAT(0) cube complex $\mathcal{X}$, and that there is a G–equivariant cubical isometric embedding $\mathcal{X} \to \prod_{i=1}^k T_i$ with each T_i a simplicial tree. Then there exists a finite-index subgroup $G' \leq G$ such that for all hyperplanes H of $\mathcal{X}$ and all $g \in G'$, the hyperplanes gH and H do not cross.

Proof By passing to the normal core, we can assume G' is normal. Let $H_1, \ldots, H_k$ be a complete list of representatives of G'–orbits of hyperplanes. For each i, the cube complex dual to the wallspace $(\mathcal{X}^{(0)}, G' \cdot H_i)$ is a simplicial tree T_i on which G' acts by isometries, and G acts on $\prod_{i=1}^k T_k$, permuting the factors, with G' stabilizing each factor (if $h \in G'$ and $g \in G$ and $H_j = gH_i$, then $hH_j = hgH_i = g(g^{-1}hg)H_i = gH_i$ since G' is normal). The existence of the embedding follows from [26, Corollary 1]; equivariance is easily checked.

Conversely, let $\mathcal{X} \to \prod_{i=1}^k T_i$ be a G–equivariant embedding, and let G' be the kernel of the action of G on the factors of $\prod_{i=1}^k \mathcal{X}$. Since $\mathcal{X} \to \prod_{i=1}^k T_i$ sends hyperplanes to hyperplanes, and sends crossing hyperplanes to crossing hyperplanes, the coloring of the hyperplanes of $\prod_{i=1}^n T_i$ by the tree factor that they cross pulls back to a G'–invariant

coloring of the hyperplanes of $\mathcal{X}$ by k colors, with the property that hyperplanes of like color do not cross. By definition, two hyperplanes in the same G' orbit have the same color. □

An element of $\mathfrak{F} - \{\mathcal{X}\}$ is *maximal* if it is not properly contained in any other element.

Definition 11.2 (hereditarily flip-free) Let G act properly and cocompactly on the cube complex $\mathcal{X}$. The action is *flip-free (with respect to $\mathfrak{F}$)* if for each cubical isometric embedding $A \times B \to \mathcal{X}$ such that $A, B \in \mathfrak{F} - \{\mathcal{X}\}$ and each of A, B is parallel to a maximal element of $\mathfrak{F} - \{\mathcal{X}\}$, no $g \in G$ has gA parallel to B. The action is *hereditarily flip-free* with respect to $\mathfrak{F}$ if for each $H \in \mathfrak{F}$, the action of $\mathrm{Stab}_G(H)$ on H is flip-free with respect to $\mathfrak{F}_H$.

Lemma 11.3 (coloring the factor system) *Let G act properly and cocompactly on $\mathcal{X}$, let $\mathfrak{F}$ be a G–invariant factor system, and suppose that there exists $G' \leq_{\mathrm{fi}} G$ such that the action of G' on $\mathcal{X}$ is hereditarily flip-free with respect to $\mathfrak{F}$. Then there exists $k \in \mathbb{N}$ and a coloring $\chi \colon \mathfrak{F} \to \{1, \dots, k\}$ such that:*

 (1) *if $F, F' \in \mathfrak{F}$ are parallel, then $\chi(F) = \chi(F')$;*

 (2) *$\chi(gF) = \chi(F)$ for all F and all $g \in G'$;*

 (3) *if F is parallel to a proper subcomplex of F', then $\chi(F) \neq \chi(F')$;*

 (4) *if $[F]$ and $[F']$ are orthogonal, then $\chi(F) \neq \chi(F')$.*

Hence there is a G'–invariant coloring $\chi \colon \overline{\mathfrak{F}} \to \{1, \dots, k\}$ such that for $1 \leq i \leq k$, the set $\overline{\mathfrak{F}}_i = \chi^{-1}(i)$ is a G'–invariant collection satisfying the hypotheses of [Corollary 10.9](#).

Proof For $n \geq 1$ and $\{q_i\}_{i=1}^n$ determined below, let $\{c_1^0, \dots, c_{q_n}^0\}$, $\{c_1^1, \dots, c_{q_1}^1\}, \dots,$ $\{c_1^n, \dots, c_{q_n}^n\}$ be disjoint sets of colors, and let ε be an extra color.

Define the relation $\sim$ on $\mathfrak{F}$ to be the transitive closure of the union of the following two relations: $F \sim F'$ if F, F' are parallel; $F \sim F'$ if there exists $g \in G'$ such that $F = gF$. Since the G–action preserves parallelism, if $F \sim F'$, then there exists $g \in G'$ with F, gF' parallel.

We say that $F \in \mathfrak{F} - \{\mathcal{X}\}$ is *maximal* if it is not parallel to a proper subcomplex of any other $F' \in \mathfrak{F} - \{\mathcal{X}\}$. Notice that maximal elements exist. In fact, if we had a chain $F_1, \dots, F_N$, with N larger than the local finiteness constant Δ of the factor system, so that F_i is parallel to a proper subcomplex of F_{i+1}, then there would be some 0–cube in $\bigcap_i \mathfrak{g}_{F_N}(F_i)$, violating local finiteness. Let $F_1, \dots, F_n \in \mathfrak{F}$ be maximal elements of $\mathfrak{F} - \{\mathcal{X}\}$, chosen so that any other maximal $F \in \mathfrak{F} - \{\mathcal{X}\}$ satisfies

$F \sim F_i$ for some unique i; there are finitely many $\sim$–classes since our hypotheses ensure that there are finitely many G–orbits in $\mathfrak{F}$. Let $\mathfrak{F}_0 \subseteq \mathfrak{F}$ consist of those F such that $F \sim F_i$ for some i, so that each maximal factor subcomplex is in $\mathfrak{F}_0$. Define $\chi_0 \colon \mathfrak{F}_0 \to \{c_1^0, \ldots, c_n^0\}$ by $\chi_0(F) = c_i^0$, where i is the unique index with $F \sim F_i$. The coloring χ_0 of $\mathfrak{F}_0$ satisfies items (1)–(3) by definition (in particular, maximality ensures that item (3) holds). Item (4) follows from the assumption that the action of G' is flip-free. Indeed, if $[F]$ and $[F']$ are orthogonal then, by definition, we can choose F, F' within their parallelism classes so that $F \times F' \hookrightarrow \mathcal{X}$ isometrically. But if $\chi_0(F) = \chi_0(F')$, then F is parallel to gF' for some $g \in G'$, contradicting flip-freeness since each of F, F' is parallel to a maximal element, namely a translate of some F_i.

Note that since $\mathfrak{F}_0$ is a locally finite collection and G' acts on $\mathcal{X}$ cocompactly, $G'_F :=$ $\mathrm{Stab}_{G'}(F)$ acts cocompactly on F for each $F \in \mathfrak{F}$. For each $F \in \mathfrak{F}$, let $\mathfrak{F}_F$ be the induced factor system, so that $\bigcup_F \mathfrak{F}_F = \mathfrak{F}$. By hereditary flip-freeness of the overall action of G' on $\mathcal{X}$, the action of G'_F on each F is flip-free with respect to $\mathfrak{F}_F$. Hence, by induction, for $1 \le i \le n$, we have a coloring $\chi'_{F_i} \colon \mathfrak{F}_{F_i} \to \{c_1^i, \ldots, c_{q_i}^i\}$ satisfying all four conclusions of the proposition with respect to the G'_{F_i}–action.

We define $\chi_i \colon \bigcup_{F \sim F_i} \mathfrak{F}_F \to \{c_1^i, \ldots, c_{q_i}^i\}$ as follows. If F is parallel to F_i, then each $F' \in \mathfrak{F}_F$ is parallel to some $F_i' \in \mathfrak{F}_{F_i}$, and if F' is also parallel to $F_i'' \in \mathfrak{F}_{F_i}$, then $\chi'_{F_i}(F_i') = \chi'_{F_i}(F_i'')$. Hence let $\chi_i(F') = \chi_{F_i}(F_i')$. We now extend G'–equivariantly to obtain χ_i. More precisely, for each $F \sim F_i$, choose $g \in G'$ so that gF is parallel to F_i, and define χ_i on $\mathfrak{F}_F$ so that $\chi_i(F') = \chi_i(gF')$ for all $F' \in \mathfrak{F}_F$. This is independent of the choice of gF because of how χ_i was defined on the subcomplexes parallel to F_i. Moreover, since χ'_{F_i} is G'_{F_i}–invariant, and parallel complexes have identical stabilizers, this is independent of the choice of g.

Let $F \sim F_i$ and let $F', F'' \in \mathfrak{F}_F$. Suppose that $\chi_i(F') = \chi_i(F'')$. Choose $g \in G'$ so that gF is parallel to F_i. Then

$$\chi_i(gF') = \chi_i(gF'') = \chi_i(F_i') = \chi_i(F_i''),$$

where $F_i', F_i'' \in \mathfrak{F}_{F_i}$ are parallel to gF', gF'', respectively. Hence neither of F_i'' nor F_i' is parallel to a subcomplex of the other, and they are not orthogonal, by our induction hypothesis. Hence the same is true of F', F'', so that χ_i correctly colors $\mathfrak{F}_F$.

For $i \ge 0$, we extend χ_i to the rest of $\mathfrak{F}$ by letting $\chi_i(H) = \varepsilon$ when $H \notin \bigcup_{F \sim F_i} \mathfrak{F}_F$ (or $H \notin \mathfrak{F}_0$ when $i = 0$). We finally define the coloring χ on $\mathfrak{F}$ by

$$\chi(F) = (\chi_0(F), \chi_1(F), \ldots, \chi_n(F)).$$

(The total number of colors is $(n+1)\prod_{i=1}^n (q_i + 1)$.)

Let $F, F' \in \mathfrak{F}$. If the set of i for which F is contained in $\mathfrak{F}_H$ for some $H \sim F_i$ differs from the corresponding set for F', then $\chi(F) \neq \chi(F')$ since there is at least one coordinate in $\chi(F)$ equal to ε for which the corresponding coordinate in $\chi(F')$ is not ε, or vice versa. It follows from the definition of the χ_i that F cannot be parallel to a proper subcomplex of F', and they cannot be orthogonal, if $\chi(F) = \chi(F')$. Hence χ is the desired coloring. $\qquad\square$

We can now prove Theorem F:

Proposition 11.4 (embedding in a product of quasi-trees) *Let $G, \mathcal{X}$, and $\overline{\overline{\mathfrak{F}}}$ be as in Lemma 11.3. Then there is a partition $\overline{\overline{\mathfrak{F}}} \cup \{[\mathcal{X}]\} = \bigsqcup_{i=1}^{k} \overline{\mathfrak{F}}_i$ such that each $\mathfrak{C}(\overline{\mathfrak{F}}_i)$ is a quasi-tree, G acts by isometries on $\prod_{i=1}^{k} \mathfrak{C}(\overline{\mathfrak{F}}_i)$, and there is a G–equivariant quasi-isometric embedding $\mathcal{X} \to \prod_{i=1}^{k} \mathfrak{C}(\overline{\mathfrak{F}}_i)$.*

Proof Lemma 11.3 allows us to finitely and equivariantly color the elements of $\mathfrak{F}$ so that elements of the same color are transverse. By Corollary 10.9 and [11, Theorem A], each color gives rise to a quasi-tree and a map from $\mathcal{X}$ to each quasi-tree. Comparing Theorem 9.1 to the distance estimate [11, Theorem 4.13] shows, as in [11, Section 5.3], that these maps give the desired quasi-isometric embedding in the product of the quasi-trees associated to the various colors. $\qquad\square$

Remark 11.5 Proposition 11.6 shows that Proposition 11.4 does not follow from Proposition 11.1. Any stipulation that no hyperplane crosses its G'–translates is satisfied when, for example, hyperplane-stabilizers are separable (this uses Scott's criterion for separability [64]). Such separability occurs when the action of G on $\mathcal{X}$ is virtually cospecial (eg when G is word-hyperbolic [1]); this is stronger than assuming the action of G on $\mathcal{X}$ is virtually flip-free.

Proposition 11.6 *There exists a group G acting properly and cocompactly on a CAT(0) cube complex $\mathcal{X}$ in such a way that $\mathcal{X}$ admits a G–equivariant quasi-isometric embedding into the product of finitely many quasi-trees, but $\mathcal{X}$ does not admit a G'–equivariant isometric embedding in a finite product of trees for any finite-index $G' \leq G$.*

Proof We will actually find G and $\mathcal{X}$ such that the following three properties hold:

(1) $\mathcal{X}$ admits a factor system $\mathfrak{F}$.

(2) For every finite-index subgroup $G' \leq G$, there exists a hyperplane H of $\mathcal{X}$ and some $g \in G'$ such that the hyperplanes H, gH cross.

(3) The action of G'' on $\mathcal{X}$ is hereditarily flip-free for some $G'' \leq_{\text{fi}} G$.

Let Z be the complex from Example 8.10, whose notation we will use throughout this proof, let $G = \pi_1 Z$, and let $\mathcal{X} = \widetilde{Z}$. Property (1) is Proposition 8.11 and property (2) is Proposition 8.12. Thus, by Proposition 11.1, $\mathcal{X}$ has no G–equivariant isometric embedding in a finite product of trees. The existence of the desired quasi-isometric embedding will follow from Proposition 11.4 once we have established property (3).

To that end, it will be convenient to compute an explicit factor system for $\mathcal{X}$. We take $\xi \geq r + 1$ and let $\mathfrak{F}$ be the smallest subset containing $\mathcal{X}$, all combinatorial hyperplanes of $\mathcal{X}$, and $\mathfrak{g}_F(F')$ whenever $F, F' \in \mathfrak{F}$ and $\mathrm{diam}(\mathfrak{g}_F(F')) \geq \xi$.

We first describe the combinatorial hyperplanes in $\mathcal{X}$. We can and shall assume that the paths $\alpha, \beta \colon [0, r] \to S, T$ from Example 8.10 are surjective. This is for convenience only; it reduces the number of isomorphism types of hyperplane that we need to consider. Let $\mathcal{X}_0$ be the universal cover of Z_0, which decomposes as a tree of spaces whose vertex-spaces are copies of $\widetilde{S} \times \widetilde{T}$ and whose edge-spaces are lines. Then $\mathcal{X}$ is formed from $\mathcal{X}_0$ by attaching all lifts of the carrier of the self-crossing hyperplane W (which is a segment of length $r + 2$) along a segment of length r.

$\mathcal{X}$ thus has compact hyperplanes, all of which are lifts of W. Each corresponds to a cardinality-2 parallelism class of combinatorial hyperplanes. The projection of such a combinatorial hyperplane W^+ onto some other hyperplane is either an isomorphism or has image of diameter at most r, whence this projection is excluded from $\mathfrak{F}$ by our choice of ξ. The projection of any other hyperplane onto W^+ also has diameter at most r and is similarly excluded. Every product subcomplex of the form $W^+ \times E$ with W^+ a compact combinatorial hyperplane has the property that E is contained in a 1–cube and thus is not a combinatorial hyperplane. Hence the compact hyperplanes do not affect whether the action of G is flip-free.

The remaining combinatorial hyperplanes are of two types: there are *two-ended* combinatorial hyperplanes, which are parallel to lifts of $\widetilde{C}$, and *bushy* hyperplanes, which are trees of the following type. Each is formed by beginning with a copy of $\widetilde{S}$, attaching a path of length 2 to each vertex, attaching a copy of $\widetilde{T}$ to the end of each of these paths, etc. Then, at the midpoint of each of the length-2 paths (ie the points at which the hyperplane intersects a lift of $\widetilde{C}$), attach a single 1–cube (dual to a compact hyperplane). Observe that any subcomplex parallel to a bushy hyperplane is a bushy hyperplane, but that two-ended hyperplanes are parallel to lines properly contained in bushy hyperplanes.

Let T, T' be subcomplexes parallel to two-ended hyperplanes and let B, B' be bushy hyperplanes. If B, B' intersect some lift of $\widetilde{S} \times \widetilde{T}$ but do not intersect, then B and B' are parallel, so that $\mathfrak{g}_B(B') = B$. If B, B' are distinct and intersect, then $\mathfrak{g}_B(B')$ is a single point. If B, B' do not intersect a common vertex-space, then $\mathfrak{g}_B(B') = \mathfrak{g}_B(T)$

for some two-ended hyperplane separating B from B'. But $\mathfrak{g}_B(T)$ is either a single point or is a parallel copy of T for any bushy B and two-ended T. Hence $\mathfrak{F}$ consists of $\mathcal{X}$, combinatorial hyperplanes and their parallel copies. Maximal elements of $\mathfrak{F}-\{\mathcal{X}\}$ are bushy hyperplanes or lifts of $\widetilde{C}$.

Let G'' be the unique index-2 subgroup of G (this is the kernel of the map $G \to \mathbb{Z}_2$ sending the element represented by C to 1). Then the action of G'' on $\mathcal{X}$ is flip-free. Indeed, $\mathcal{X}$ does not contain the product of two bushy hyperplanes, so we need only consider the case of a product $\widetilde{\alpha} \times \widetilde{\beta}$ in some $\widetilde{S} \times \widetilde{T}$. But any G''–translate of $\widetilde{\beta}$ is either in the same product piece — and thus parallel to β — or at even distance from $\widetilde{\beta}$ in the Bass–Serre tree, and hence is again a component of the preimage of β, and hence not parallel to $\widetilde{\alpha}$. To verify hereditary flip-freeness, it suffices to note that each element of $\mathfrak{F}-\{\mathcal{X}\}$ is a tree, and thus contains no nontrivial products. $\square$

12 Consistency and realization

In this section, we prove an analogue in the cubical context of the consistency theorem in the mapping class group; see [7, Theorem 4.3]. (For an analogue of this theorem in Teichmüller space with the Teichmüller metric, see [30, Section 3].) For the purposes of this section, $\mathcal{X}$ is an arbitrary CAT(0) cube complex with a factor system $\mathfrak{F}$. Let ξ be the constant from Definition 8.1. For each $F \in \mathfrak{F}$, let

$$\mathfrak{B}F = \{S \subseteq \mathcal{C}F^{(0)} : \operatorname{diam}_{\widehat{\mathcal{C}}F}(S) \leq 1\}.$$

If $U, V \in \mathfrak{F}$, and V is parallel to a proper subcomplex of U, then there is a map $\rho_V^U \colon \mathfrak{B}U \to 2^{\widehat{\mathcal{C}}V^{(0)}}$ defined as follows: let $b \in \mathfrak{B}U$ be a clique in $\mathcal{C}U$ and let

$$\rho_V^U(b) = \bigcup_W \pi_V(W),$$

where W varies over all combinatorial hyperplanes parallel to hyperplanes of U in the clique b.

Let $\vec{b} \in \prod_{[F]\in\overline{\mathfrak{F}}} \mathfrak{B}F$ be a tuple, whose $[F]$–coordinate we denote by b_F. The tuple $\vec{b}$ is *realized* if there exists $x \in \mathcal{X}$ such that $\pi_F(x) = b_F$ for all $[F] \in \overline{\mathfrak{F}}$; note that by Lemma 8.17 b_F is independent of the choice of representative $F \in [F]$. In this section we first give a set of *consistency conditions* on coordinate projections which hold for any element of $x \in \mathcal{X}$ (for the analogue in the mapping class group, see [3]). We then show that this set of necessary consistency conditions on a vector of coordinates is, essentially, also sufficient for that vector to be realized by an element of $\mathcal{X}$.

Remark 12.1 In this section, we use the sets $\mathfrak{B}F$, simply because the projection of a point to the contact graph is always a clique. In the more general context introduced in Section 13 — specifically in Definition 13.1 — we abstract the conditions on tuples in $\prod_{[F]\in\overline{\mathfrak{F}}} \mathfrak{B}F$ provided by the next several results, in terms of tuples in $\prod_{S\in\mathfrak{S}} 2^S$, where $\mathfrak{S}$ is some set of uniformly hyperbolic spaces, and the tuple restricted to each coordinate is a uniformly bounded (though not necessarily diameter–1) set. In the cubical setting, instead of doing this using $S = \widehat{\mathcal{C}}F$, we instead use $\mathfrak{B}F$ because it yields more refined statements.

Recall that the parallelism classes $[F], [F'] \in \overline{\mathfrak{F}}$ are *transverse* if they are not orthogonal and if F is not parallel to a subcomplex of F', or vice versa, for some (hence all) $F \in [F]$, $F' \in [F']$. We begin with the consistency conditions, which take the form of the following inequalities:

Proposition 12.2 (realized tuples are consistent) *Let $\vec{b} \in \prod_{[F]\in\overline{\mathfrak{F}}} \mathfrak{B}F$ be realized by $x \in \mathcal{X}$. There exists $\kappa_0 = \kappa_0(\xi)$ such that for all $U, V \in \mathfrak{F}$, the following hold:*

(1) *If U and V are transverse, then*
$$\min\{\mathrm{d}_{\widehat{\mathcal{C}}U}(b_U, \pi_U(V)), \mathrm{d}_{\widehat{\mathcal{C}}V}(b_V, \pi_V(U))\} \leq \kappa_0;$$

(2) *If V is parallel to a proper subcomplex of U, then*
$$\min\{\mathrm{d}_{\widehat{\mathcal{C}}U}(b_U, \pi_U(V)), \mathrm{diam}_{\widehat{\mathcal{C}}V}(b_V \cup \rho_V^U(b_U))\} \leq \kappa_0.$$

A tuple $\vec{b}$ (realized or otherwise) satisfying the conclusions of the above proposition is said to be κ_0–*consistent*. More generally, a tuple $\vec{b}$ satisfying inequalities (1) and (2) from Proposition 12.2, with κ_0 replaced by some $\kappa \geq \kappa_0$, will be called κ–*consistent*.

Proof of Proposition 12.2 First suppose that neither U nor V is parallel to a subcomplex of the other. If $\mathrm{d}_{\widehat{\mathcal{C}}U}(b_U, \pi_U(V)) > 2\xi + 4$, then $\mathrm{d}_U(\mathfrak{g}_U(x), \mathfrak{g}_U(V)) > \xi$, so that the set $\mathcal{H}_U$ of hyperplanes crossing U and separating x from V has cardinality at least ξ. Suppose also that $\mathrm{d}_{\widehat{\mathcal{C}}V}(b_V, \pi_V(U)) > 2\xi + 4$, so that the set $\mathcal{H}_V$ of hyperplanes crossing V and separating x from U has cardinality at least ξ. Let $W \in \mathcal{H}_V$ and let W^+ be an associated combinatorial hyperplane not separated from x by any hyperplane crossing V. Notice that $\pi_U(W^+)$ contains $\pi_U(x) = b_U$ (because W separates x from U) and intersects $\pi_U(V)$. In particular, $\mathrm{diam}_{\widehat{\mathcal{C}}U}(\pi_U(W^+)) > \xi + 2$ and hence, by Lemma 8.19, U is parallel to a subcomplex U' of W^+ containing $\mathfrak{g}_{W^+}(x)$. We have $\mathrm{diam}_{\widehat{\mathcal{C}}V}(\pi_V(U')) \leq \xi + 2$ for otherwise V would be parallel into U', which is impossible since U' is parallel to U. Also, $\pi_V(U')$ contains $\pi_V(x) = b_V$ and hence $\mathrm{d}_{\widehat{\mathcal{C}}V}(\pi_V(U), \pi_V(U')) > \xi + 2$, ie Proposition 10.1(1) does not hold, so U and V are orthogonal by Proposition 10.1.

Let V be parallel to a proper subcomplex of U and let $\{W_i\}$ be the set of combinatorial hyperplanes of U containing $\mathfrak{g}_U(x)$. Suppose that $d_{\widehat{C}U}(b_U, \pi_U(V)) > \xi + 2$. Then for each i, we have that $\mathfrak{g}_V(W_i)$ is a single point (since otherwise there would be a hyperplane intersecting $\mathfrak{g}_V(W_i)$ and hence intersecting both V and W_i), hence $\mathfrak{g}_V(W_i) = \mathfrak{g}_V(x)$, whence $b_V = \rho_V^U(b_U)$. In either case, $\kappa_0 = \xi + 2$ suffices. $\qquad\square$

Remark 12.3 We note that Proposition 10.7 can alternatively be proven as an immediate consequence of Proposition 12.2 and Lemma 8.19. The formulation here is very close to that in the mapping class group; see [3] and [7].

Theorem 12.4 (consistency and realization) *Let $\kappa_0 \geq 1$ be the constant from Proposition 12.2. For each $\kappa \geq \kappa_0$, there exists $\theta \geq 0$ such that, if $\vec{b} \in \prod_{[F] \in \overline{\mathfrak{F}}} \mathcal{B}F$ is κ–consistent, then there exists $y \in \mathcal{X}$ such that $d_{\widehat{C}F}(b_F, \pi_F(y)) \leq \theta$ for all $[F] \in \overline{\mathfrak{F}}$.*

Proof For each $[U] \in \overline{\mathfrak{F}}$, let E_U be the smallest convex subcomplex of $\mathcal{X}$ with the property that $y \in E_U$ if $\pi_U(y)$ intersects the 2κ–neighborhood of b_U in $\widehat{C}U$. By Lemma 8.17, we have $E_U = E_{U'}$ if $[U] = [U']$.

Moreover, we claim that $\pi_U(E_U)$ has diameter at most 5κ. Indeed, let $\mathcal{E}$ be the set of all 0–cubes x such that $\pi_U(x) \cap \mathcal{N}_{2\kappa}(b_U) \neq \varnothing$. Given $x, y \in \mathcal{E}$, let $H_1, \ldots, H_k$ and $V_1, \ldots, V_\ell$ be sequences of subcomplexes, representing geodesics in $\widehat{C}U$, so that $x \in H_1$, $y \in V_\ell$, and H_k, V_1 correspond to points of b_U. Hence there is a combinatorial path from x to y and carried on $\bigcup_i H_i \cup \bigcup_j V_j$; by construction, each 0–cube on this path lies in $\mathcal{E}$. Hence $\mathcal{E}$ is 1–coarsely connected. Now let $z \in E_U - \mathcal{E}$ be a 0–cube and let H be a hyperplane whose carrier contains z and which crosses a geodesic from z to a point of $\mathcal{E}$. Since E_U is the convex hull of $\mathcal{E}$, the hyperplane H separates two 0–cubes of $\mathcal{E}$ lying in $\mathcal{N}(H)$. Hence $d_{\widehat{C}U}(\pi_U(z), b_U) \leq 2\kappa + 1$ and thus $\pi_U(E_U) \leq 5\kappa$.

Suppose $[U], [V] \in \overline{\mathfrak{F}}$ are transverse. By κ–consistency and Lemma 8.19, we have that either $V \subseteq E_U$ or $U \subseteq E_V$. The same is true if V is parallel to a proper subcomplex of U, by κ–consistency. Finally, if $U \times V$ isometrically embeds in $\mathcal{X}$, then some representative of $[V]$ lies in E_U and some representative of $[U]$ lies in E_V. Hence, in all cases, at least one representative subcomplex corresponding to one of the cliques b_U or b_V belongs to $E_U \cap E_V$, ie $E_U \cap E_V \neq \varnothing$ for all U, V. Thus, by the Helly property, for any finite collection $[F_1], \ldots, [F_n]$ of elements of $\overline{\mathfrak{F}}$, we can find an $x \in \mathcal{X}$ such that $d_{\widehat{C}F_i}(\pi_{F_i}(x), b_{F_i}) \leq 5\kappa$ for each i.

Fix any 0–cube $x_0 \in \mathcal{X}$ and let θ_1 be a (large) constant to be chosen later. Also, denote by $\overline{\mathfrak{F}}_{\max}$ the set of all $[F] \in \overline{\mathfrak{F}} - \{\mathcal{X}\}$ such that

- $d_{\widehat{\mathcal{C}}F}(\pi_F(x_0), b_F) > \theta_1$, and
- F is not parallel to a proper subcomplex of a representative of $[F'] \in \overline{\mathfrak{F}} - \{\mathcal{X}\}$ for which $d_{\widehat{\mathcal{C}}F'}(\pi_{F'}(x_0), b_{F'}) > \theta_1$.

Let $l = d_{\widehat{\mathcal{C}}\mathcal{X}}(\pi_{\mathcal{X}}(x_0), b_{\mathcal{X}})$. We claim that $\overline{\mathfrak{F}}_{\max}$ has finitely many elements. Suppose not, and let $[F_1], \ldots, [F_{p(l+5\kappa)+1}] \in \overline{\mathfrak{F}}_{\max}$, for some p to be determined below so that $d_{\widehat{\mathcal{C}}F}(\pi_{F_i}(x_0), b_{F_i}) > \theta_1$ for $1 \le i \le p(l+5\kappa)+1$. Consider $x \in \mathcal{X}$ satisfying $d_{\widehat{\mathcal{C}}F_i}(\pi_{F_i}(x), b_{F_i}) \le 5\kappa$ for each i and such that $d_{\widehat{\mathcal{C}}\mathcal{X}}(\pi_{\mathcal{X}}(x), b_{\mathcal{X}}) \le 5\kappa$. The existence of such a point x was established above.

Thus, there exists $m \le l + 5\kappa$ such that $[T_0], \ldots, [T_m] \in \mathfrak{F} - \{\mathcal{X}\}$ is a sequence, with T_0 a combinatorial hyperplane parallel to a hyperplane representing a vertex of $\pi_{\mathcal{X}}(x_0)$ and T_m enjoying the same property with x replacing x_0, representing a geodesic in $\widehat{\mathcal{C}}\mathcal{X}$ joining $\pi_{\mathcal{X}}(x_0)$ to $\pi_{\mathcal{X}}(x)$. Suppose that we have chosen constants θ_1 and κ so that $\theta_1 \ge 6\kappa - 6 \ge 5\kappa + 4\xi + 10$. Hence, for each i, we have $d_{\widehat{\mathcal{C}}F_i}(\pi_{F_i}(x_0), \pi_{F_i}(x)) \ge \theta_1 - 5\kappa \ge 4\xi + 10$. Proposition 9.4 now implies that each F_i is parallel to a subcomplex of T_j for some $j \in \{0, \ldots, m\}$. Hence $F_i \in \overline{\mathfrak{F}}_{\max} \cap \overline{\mathfrak{F}}_{T_j}$, so that it suffices to show that the latter is finite.

Note that for each j, the set $\overline{\mathfrak{F}}_{\max} \cap \overline{\mathfrak{F}}_{T_j}$ is the set of maximal elements $H \in \overline{\mathfrak{F}}_{T_j}$ with $d_{\widehat{\mathcal{C}}H}(\pi_H(\pi_{T_j}(x_0)), b_H) > \theta_1$. This follows from the fact that $\mathfrak{g}_H(\mathfrak{g}_{T_j}(x_0)) = \mathfrak{g}_H(x_0)$. Either $\overline{\mathfrak{F}}_{\max} \cap \overline{\mathfrak{F}}_{T_j} = \{T_j\}$, or $d_{\widehat{\mathcal{C}}T_j}(\pi_{T_j}(x_0), b_{T_j}) \le \theta_1$ and, by induction on Δ, there exists $p \ge 1$ with $|\overline{\mathfrak{F}}_{\max} \cap \overline{\mathfrak{F}}_{T_j}| \le p$. Hence $|\overline{\mathfrak{F}}_{\max}| \le p(l + 5\kappa)$.

The finiteness of $\overline{\mathfrak{F}}_{\max}$ holds for any choice of x_0, but it will be convenient for what follows to assume that x_0 has been chosen so that $\pi_{\mathcal{X}}(x_0)$ contains the clique $b_{\mathcal{X}}$.

As before, choose $x \in \mathcal{X}$ so that $d_{\widehat{\mathcal{C}}F}(\pi_F(x), b_F) \le 5\kappa$ for each $[F] \in \overline{\mathfrak{F}}_{\max}$. Let $\overline{\mathfrak{F}}'_{\max}$ be the set of all $[F] \in \overline{\mathfrak{F}}_{\max}$ with the property that for each $[F'] \in \overline{\mathfrak{F}}_{\max}$ transverse to $[F]$, the closest parallel copy of F to x is closer in $\mathcal{X}$ than the closest parallel copy of F' to x.

Claim 1 *Let $[F_1], [F_{-1}] \in \overline{\mathfrak{F}}_{\max}$ be transverse. Then there exists a unique $i \in \{\pm 1\}$ such that $d_{\widehat{\mathcal{C}}F_i}(\pi_{F_i}(x_0), \pi_{F_i}(F_{-i})) \le \kappa$. Moreover, $d_{\mathcal{X}}(F_i, x) < d_{\mathcal{X}}(F_{-i}, x)$, assuming $F_{\pm i}$ is the closest representative of $[F_{\pm i}]$ to x.*

Proof of Claim 1 Suppose first that $d_{\widehat{\mathcal{C}}F_1}(\pi_{F_1}(x_0), \pi_{F_1}(F_{-1})) > \kappa$. Then, by Proposition 12.2, we have

$$d_{\widehat{\mathcal{C}}F_{-1}}(\pi_{F_{-1}}(x_0), \pi_{F_{-1}}(F_1)) \le \kappa,$$

and we are done. If $d_{\widehat{\mathcal{C}}F_1}(\pi_{F_1}(x_0), \pi_{F_1}(F_{-1})) \le \kappa$, then

$$d_{\widehat{\mathcal{C}}F_1}(\pi_{F_1}(x), \pi_{F_1}(F_{-1})) > \kappa.$$

Proposition 12.2 implies that $d_{\widehat{C}F_{-1}}(\pi_{F_{-1}}(x), \pi_{F_{-1}}(F_1)) \leq \kappa$. But then

$$d_{\widehat{C}F_{-1}}(\pi_{F_{-1}}(x_0), \pi_{F_{-1}}(F_1)) \geq \theta_1 - \kappa - \xi - 2 > \kappa.$$

To prove the "moreover" clause, notice that $d_{\widehat{C}F_i}(\pi_{F_i}(x), \pi_{F_i}(F_{-i})) > 4\xi + 10$. Hence, by Proposition 9.4, for any point x' on any parallel copy of F_{-i} there is a geodesic from x' to x passing through a parallel copy of F_i. $\square$

If $\overline{\mathfrak{F}}'_{max} = \varnothing$ then we can choose $y = x_0$. Otherwise, $\overline{\mathfrak{F}}'_{max} \neq \varnothing$, and Claim 1 implies that the elements of $\overline{\mathfrak{F}}'_{max}$ are pairwise orthogonal; hence $\mathcal{X}$ contains $P = \prod_{[F]\in\overline{\mathfrak{F}}'_{max}} F$. By Lemma 2.4, there is a convex subcomplex $Q \subset \mathcal{X}$ such that the convex hull of the union of all subcomplexes parallel to P is isomorphic to $P \times Q$. Since $|P| > 1$ by our choice of θ_1, there is at least one hyperplane crossing P, whence Q is not unique in its parallelism class and hence lies in a combinatorial hyperplane U. By induction on Δ, there exists $y_U \in U$ such that $d_{\widehat{C}W}(\pi_W(y_U), b_W) \leq \theta_2$ for some fixed θ_2 and all $[W] \in \mathfrak{F}_U$. Let $y_Q = \mathfrak{g}_Q(y_U)$.

By induction on Δ, for each $[F] \in \overline{\mathfrak{F}}'_{max}$, we can choose $F \in [F]$ and a point $y_F \in F$ such that $d_{\widehat{C}H}(\pi_H(y_F), b_H) \leq \theta_2$ for all $[H] \in \mathfrak{F}_F$ and some θ_2. Importantly, θ_2 can be chosen to depend only on κ, ξ and $\Delta - 1$, ie θ_2 is independent of $\overline{\mathfrak{F}}_{max}$, and we can thus assume that $\theta_1 > \theta_2 + 2\kappa$. By pairwise orthogonality, there exists $y \in P \times Q$ such that $\mathfrak{g}_F(y) = y_F$ for all $[F] \in \overline{\mathfrak{F}}'_{max}$ and $\mathfrak{g}_Q(y) = y_Q$. We now choose θ so that $d_{\widehat{C}W}(\pi_W(y), b_W) \leq \theta$ for all $[W] \in \mathfrak{F}$.

First, if W is parallel to a proper subcomplex of some $[F] \in \overline{\mathfrak{F}}'_{max}$, then any $\theta \geq \theta_2$ suffices.

Second, suppose that some $[F] \in \overline{\mathfrak{F}}'_{max}$ is parallel to a proper subcomplex of W. Suppose $d_{\widehat{C}W}(b_W, \pi_W(F)) \leq \kappa + 2\theta_1$. Notice that $\pi_W(y) = \pi_W(y_F)$ because F is parallel to a subcomplex of W, and hence

$$d_{\widehat{C}W}(b_W, \pi_W(y_F)) = d_{\widehat{C}W}(b_W, \pi_W(y)) \leq 2\kappa + 2\theta_1.$$

Hence suppose $d_{\widehat{C}W}(b_W, \pi_W(F)) > \kappa + 2\theta_1$, so $\text{diam}_{\widehat{C}F}(b_F \cup \rho_F^W(b_W)) \leq \kappa$. Since, up to parallelism, W properly contains F and $[F] \in \mathfrak{F}_{max}$, we can conclude that if $W \neq \mathcal{X}$ then we have $d_{\widehat{C}W}(b_W, \pi_W(x_0)) \leq \theta_1$, while the same inequality also holds in the case $W = \mathcal{X}$ because of our choice of x_0. Hence $d_{\widehat{C}W}(\pi_W(x_0), \pi_W(F)) > \kappa + \theta_1$. The inequality

$$d_{\widehat{C}W}(\pi_W(x_0), \pi_W(F)) > \kappa + \theta_1 \geq d_{\widehat{C}W}(\pi_W(x_0), b_W)$$

implies that a hierarchy path joining $\mathfrak{g}_W(x_0)$ to a point in W projecting to b_W cannot pass through an element of $[F]$. This fact together with Proposition 9.4 gives the

inequality $d_{\widehat{C}F}(\pi_F(x_0), \rho_F^W(b_W)) \leq 4\xi + 10$. Hence $d_{\widehat{C}F}(\pi_F(x_0), b_F) \leq 4\xi + 10 + \kappa$, which contradicts $F \in \overline{\mathfrak{F}}_{\max}$, so that this case does not arise.

Third, suppose that $[W]$ is transverse to some $[F] \in \overline{\mathfrak{F}}'_{\max}$. If $d_{\widehat{C}F}(\pi_F(y), \pi_F(W)) > \kappa + \theta_2$, then by consistency of $\vec{b}$, Proposition 12.2, and the fact $\pi_W(y) \in \pi_W(F)$ (since y lies in a parallel copy of F), we have $d_{\widehat{C}W}(\pi_W(y), b_W) \leq \kappa$. Hence, suppose $d_{\widehat{C}F}(\pi_F(y), \pi_F(W)) \leq \kappa + \theta_2$. In this case $d_{\widehat{C}F}(\pi_F(x_0), \pi_F(W)) > \theta_1 - \theta_2 - \kappa$, whence $d_{\widehat{C}W}(\pi_W(x_0), \pi_W(F)) \leq \kappa$.

Now, since y belongs to (a parallel copy of) F, we have $d_{\widehat{C}W}(\pi_W(y), \pi_W(F)) = 0$. We now wish to argue that the inequality $d_{\widehat{C}W}(\pi_W(y), b_W) \geq \theta_1 + \xi + 2 + \kappa$ is impossible, thus concluding the proof in this case. In fact, if it holds, then

$$d_{\widehat{C}W}(\pi_W(x_0), b_W) \geq d_{\widehat{C}W}(\pi_W(y), b_W) - \operatorname{diam}(\pi_W(F)) - d_{\widehat{C}W}(\pi_W(F), \pi_W(x_0))$$
$$\geq \theta_1.$$

Hence W is contained in some W' with $[W'] \in \overline{\mathfrak{F}}_{\max}$. Notice that $[W']$ must be transverse to $[F]$, since F cannot be parallel into W nor vice versa, by maximality of F and W', and since, if $[W']$ was orthogonal to $[F]$, then $[W]$ would be as well.

Note $d_{\widehat{C}F}(\pi_F(y), \pi_F(W')) \leq \kappa$, since this holds for W. Then we have the inequality $d_{\widehat{C}W'}(\pi_{W'}(x_0), \pi_{W'}(F)) \leq \kappa$, by an argument above. But then, by Claim 1, the closest parallel copy of W' to x is closer than the closest parallel copy of F, in contradiction with $[F] \in \overline{\mathfrak{F}}'_{\max}$.

The only case left is when W is orthogonal to all $[F] \in \overline{\mathfrak{F}}'_{\max}$, which implies that W is, up to parallelism, contained in Q. In this case, $\pi_W(y) = \pi_W(\mathfrak{g}_Q(y_Q)) = \pi_W(y_Q)$, whence we are done by the choice of y_Q. $\qquad\square$

Remark 12.5 Let $\kappa \geq \kappa_0$ and let $\vec{b}$ be a κ–consistent tuple. Theorem 12.4 yields a constant θ and a 0–cube $x \in \mathcal{X}^{(0)}$ such that $d_{\widehat{C}F}(x, b_F) \leq \theta$ for all $F \in \mathfrak{F}$. By Theorem 9.1, any other 0–cube y with this property is at a uniformly bounded distance — depending on κ and ξ — from x. The definition of a factor system implies that $\mathcal{X}$ is uniformly locally finite, so the number of such 0–cubes y is bounded by a constant depending only on κ, ξ, and the growth function of $\mathcal{X}^{(0)}$.

Part III   Hierarchically hyperbolic spaces

13   Quasi-boxes in hierarchically hyperbolic spaces

In this section, we work in a level of generality which includes both mapping class groups and each CAT(0) cube complex $\mathcal{X}$ with a factor system $\mathfrak{F}$.

Definition 13.1 (hierarchically hyperbolic space) The metric space $(\mathcal{X}, d_{\mathcal{X}})$ is a *hierarchically hyperbolic space* if there exists $\delta \geq 0$, an index set $\mathfrak{S}$, and a set $\{\widehat{\mathcal{C}}W \mid W \in \mathfrak{S}\}$ of δ–hyperbolic spaces, such that the following conditions are satisfied:

(1) **Projections** There is a set $\{\pi_W : \mathcal{X} \to 2^{\widehat{\mathcal{C}}W} \mid W \in \mathfrak{S}\}$ of *projections* sending points in $\mathcal{X}$ to sets of diameter bounded by some $\xi \geq 0$ in the various $\widehat{\mathcal{C}}W \in \mathfrak{S}$.

(2) **Nesting** $\mathfrak{S}$ is equipped with a partial order $\sqsubseteq$, and either $\mathfrak{S} = \varnothing$ or $\mathfrak{S}$ contains a unique $\sqsubseteq$–maximal element; when $V \sqsubseteq W$, we say V is *nested* in W. We require that $W \sqsubseteq W$ for all $W \in \mathfrak{S}$. For each $W \in \mathfrak{S}$, we denote by $\mathfrak{S}_W$ the set of $V \in \mathfrak{S}$ such that $V \sqsubseteq W$. Moreover, for all $V, W \in \mathfrak{S}$ with V properly nested into W there is a specified subset $\rho_W^V \subset \widehat{\mathcal{C}}W$ with $\operatorname{diam}_{\widehat{\mathcal{C}}W}(\rho_W^V) \leq \xi$. There is also a *projection* $\rho_V^W : \widehat{\mathcal{C}}W \to 2^{\widehat{\mathcal{C}}V}$. (The similarity in notation is justified by viewing ρ_W^V as a coarsely constant map $\widehat{\mathcal{C}}V \to 2^{\widehat{\mathcal{C}}W}$.)

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

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

$$\min\{d_{\widehat{\mathcal{C}}W}(\pi_W(x), \rho_W^V), d_{\widehat{\mathcal{C}}V}(\pi_V(x), \rho_V^W)\} \leq \kappa_0$$

for all $x \in \mathcal{X}$; alternatively, in the case $V \sqsubseteq W$, then for all $x \in \mathcal{X}$ we have

$$\min\{d_{\widehat{\mathcal{C}}W}(\pi_W(x), \rho_W^V), \operatorname{diam}_{\widehat{\mathcal{C}}V}(\pi_V(x) \cup \rho_V^W(\pi_W(x)))\} \leq \kappa_0.$$

Suppose that either $U \subsetneq V$ or $U \pitchfork V$, and either $U \subsetneq W$ or $U \pitchfork W$. Then if $V \pitchfork W$, we have

$$\min\{d_{\widehat{\mathcal{C}}W}(\rho_W^U, \rho_W^V), d_{\widehat{\mathcal{C}}V}(\rho_V^U, \rho_V^W)\} \leq \kappa_0$$

and if $V \subsetneq W$, we have

$$\min\{d_{\widehat{\mathcal{C}}W}(\rho_W^U, \rho_W^V), \operatorname{diam}_{\widehat{\mathcal{C}}V}(\rho_V^U \cup \rho_V^W(\rho_W^U))\} \leq \kappa_0.$$

Finally, if $V \sqsubseteq U$ or $U \perp V$, then $d_{\widehat{\mathcal{C}}W}(\rho_W^U, \rho_W^V) \leq \kappa_0$ whenever $W \in \mathfrak{S} - \{U, V\}$ satisfies either $V \sqsubseteq W$ or $V \pitchfork W$ and either $U \sqsubseteq W$ or $U \pitchfork W$.

(5) **Finite complexity** There exists $n \geq 0$, the *complexity* of $\mathcal{X}$ (with respect to $\mathfrak{S}$), such that any sequence (U_i) with U_i properly nested into U_{i+1} has length at most n.

(6) **Distance formula** There exists $s_0 \geq \xi$ such that for all $s \geq s_0$ there exist constants K, C such that for all $x, x' \in \mathcal{X}$,

$$d_{\mathcal{X}}(x, x') \asymp_{(K,C)} \sum_{W \in \mathfrak{S}} \{\!\{d_{\widehat{C}W}(\pi_W(x), \pi_W(x'))\}\!\}_s.$$

We often write $\sigma_{\mathcal{X},s}(x, x')$ to denote the right-hand side of item (6); more generally, given $W \in \mathfrak{S}$, we denote by $\sigma_{W,s}(x, x')$ the corresponding sum taken over $\mathfrak{S}_W$.

(7) **Large links** There exists $\lambda \geq 1$ such that the following holds. Let $W \in \mathfrak{S}$ and let $x, x' \in \mathcal{X}$. Let $N = \lambda d_{\widehat{C}W}(\pi_W(x), \pi_W(x')) + \lambda$. Then there exists $\{T_i\}_{i=1,\ldots,\lfloor N \rfloor} \subseteq \mathfrak{S}_W - \{W\}$ such that for all $T \in \mathfrak{S}_W - \{W\}$, either $T \in \mathfrak{S}_{T_i}$ for some i, or $d_{\widehat{C}T}(\pi_T(x), \pi_T(x')) < s_0$. Also, $d_{\widehat{C}W}(\pi_W(x), \rho_W^{T_i}) \leq N$ for each i.

(8) **Bounded geodesic image** For all $W \in \mathfrak{S}$, all $V \in \mathfrak{S}_W - \{W\}$, and all geodesics γ of $\widehat{C}W$, either $\mathrm{diam}_{\widehat{C}V}(\rho_V^W(\gamma)) \leq B$ or $\gamma \cap \mathcal{N}_E(\rho_W^V) \neq \varnothing$ for some uniform B, E.

(9) **Realization** For each κ there exist θ_e, θ_u such that the following holds. Let $\vec{b} \in \prod_{W \in \mathfrak{S}} 2^{\widehat{C}W}$ have each coordinate correspond to a subset of $\widehat{C}W$ of diameter at most κ; for each W, let b_W denote the $\widehat{C}W$–coordinate of $\vec{b}$. Suppose that whenever $V \pitchfork W$ we have

$$\min\{d_{\widehat{C}W}(b_W, \rho_W^V), d_{\widehat{C}V}(b_V, \rho_V^W)\} \leq \kappa$$

and whenever $V \sqsubseteq W$ we have

$$\min\{d_{\widehat{C}W}(b_W, \rho_W^V), \mathrm{diam}_{\widehat{C}V}(b_V \cup \rho_V^W(b_W))\} \leq \kappa.$$

Then the set of all $x \in \mathcal{X}$ such that $d_{\widehat{C}W}(b_W, \pi_W(x)) \leq \theta_e$ for all $\widehat{C}W \in \mathfrak{S}$ is nonempty and has diameter at most θ_u. (A tuple $\vec{b}$ satisfying the inequalities above is called κ–*consistent*.)

(10) **Hierarchy paths** There exists $D \geq 0$ such that any pair of points in $\mathcal{X}$ can be joined by a (D, D)–quasi-geodesic γ with the property that, for each $W \in \mathfrak{S}$, the projection $\pi_W(\gamma)$ is at Hausdorff distance at most D from any geodesic connecting $\pi_W(x)$ to $\pi_W(y)$. We call such quasi-geodesics *hierarchy paths*.

We say that the metric spaces $\{\mathcal{X}_i\}$ are *uniformly hierarchically hyperbolic* if each $\mathcal{X}_i$ satisfies the axioms above and all constants can be chosen uniformly.

Observe that a space $\mathcal{X}$ is hierarchically hyperbolic with respect to $\mathfrak{S} = \varnothing$, ie hierarchically hyperbolic of complexity 0, if and only if $\mathcal{X}$ is bounded. Similarly, $\mathcal{X}$ is hierarchically hyperbolic of complexity 1, with respect to $\mathfrak{S} = \{\mathcal{X}\}$, if and only if $\mathcal{X}$ is hyperbolic.

Remark 13.2 (cube complexes with factor systems are hierarchically hyperbolic spaces) In the case where $\mathcal{X}$ is a CAT(0) cube complex with a factor system, we take $\mathfrak{S}$ to be a subset of the factor system containing exactly one element from each parallelism class. We take $\{\hat{\mathcal{C}}W \mid W \in \mathfrak{S}\}$ to be the set of factored contact graphs of elements of $\mathfrak{S}$. All of the properties required by Definition 13.1 are satisfied: given $U, V \in \mathfrak{S}$, we have $U \sqsubseteq V$ if and only if U is parallel into V; orthogonality is defined as in Section 10. The last condition on orthogonality is satisfied since each element of a factor system is either unique in its parallelism class or is contained in a hyperplane. The first two inequalities in item (4) follow from Proposition 10.7; the second two follow from the first two since each ρ_V^U is the image of the convex subcomplex $U \in \mathfrak{S}$ under the projection π_V. Finiteness of complexity follows from the definition of a factor system (specifically, uniform local finiteness of the family of factors). Item (6) is provided by Theorem 9.1, item (7) by Proposition 9.4, item (8) is Proposition 8.20, item (9) is Theorem 12.4, and item (10) is Proposition 8.23. In order to ensure that nesting and orthogonality are mutually exclusive, we require that elements of the factor system $\mathfrak{S}$ are not single points.

Remark 13.3 The bounded geodesic image constant E from Definition 13.1 can be taken to be 1 for cube complexes with factor systems, as well as for the mapping class group. However, we allow an arbitrary (fixed) constant in the general definition for greater flexibility.

Remark 13.4 In the case of a factor system, the constant in item (1) of Definition 13.1 is 1, since points of $\mathcal{X}$ project to cliques in each factored contact graph. The constant in item (2) is $\xi + 2$, where ξ is the constant from Definition 8.1. For simplicity, in Definition 13.1, we use a single constant to fulfill both of these roles.

13.1 Product regions and standard boxes

Let $\mathcal{X}$ be a hierarchically hyperbolic space and let $U \in \mathfrak{S}$. With κ_0 as in Definition 13.1, we say that a tuple $\vec{a}_\sqsubseteq \in \prod_{W \sqsubseteq U} 2^{\hat{\mathcal{C}}W}$ or $\vec{a}_\perp \in \prod_{W \perp U} 2^{\hat{\mathcal{C}}W}$ is κ_0–*consistent* if the inequalities from item (4) are satisfied for each pair of coordinates a_W and a_V in the tuple with the relevant $\pi_W(x)$ term replaced by a_W (and similarly for V).

Let $\vec{a}_\sqsubseteq$ and $\vec{a}_\perp$ be κ_0–consistent. Let $\vec{a} \in \prod_{\hat{\mathcal{C}}W \in \mathfrak{S}} 2^{\hat{\mathcal{C}}W}$ be the tuple with $a_W = \rho_W^U$ for each $W \pitchfork U$ or $U \sqsubseteq W$, and such that a_W otherwise agrees with the relevant

coordinate of $\vec{a}_{\sqsubseteq}$ or $\vec{a}_{\perp}$, depending on whether $W \sqsubseteq U$ or $W \perp U$. We claim that $\vec{a}$ is κ_0–consistent. Indeed, all inequalities involving some $W \pitchfork U$ or $U \sqsubseteq W$ are satisfied because of the last three inequalities in Definition 13.1(4). Otherwise, if $W \sqsubseteq U$ or $W \perp U$, then, if $V \sqsubseteq W$, we have that V, W are either both nested into or both orthogonal to U, and in both cases the consistency inequality holds by assumption.

Item (9) of Definition 13.1 (realization) combines with the above discussion to yield a natural coarsely defined map $\phi_U \colon E_U \times F_U \to \mathcal{X}$, where E_U and F_U are the sets of consistent elements of $\prod_{W \perp U} 2^{\widehat{C}W}$ and $\prod_{W \sqsubseteq U} 2^{\widehat{C}W}$, respectively.

Remark 13.5 When $\mathcal{X}$ is a cube complex with a factor system $\mathfrak{S}$, the space F_U can be taken to be the subcomplex $U \in \mathfrak{S}$, we can take E_U to be the complex provided by Lemma 2.4, and ϕ_U to be the inclusion, which is a cubical isometric embedding.

We will denote $P_U = \phi_U(E_U \times F_U)$. In this context we define a *gate*, $\mathfrak{g}_{P_U} \colon \mathcal{X} \to P_U$, in the following way. Given $x \in \mathcal{X}$, we let $\mathfrak{g}_{P_U}(x)$ be a point that coarsely realizes the coordinates $a_W = \pi_W(x)$ for W nested into or orthogonal to U and $\pi_W(x) = \rho_W^U$ for $W \pitchfork U$ or $U \sqsubseteq W$.

By Definition 13.1(6) (distance formula), for a fixed U the subspaces $\phi_U(E_U \times \{b_{\sqsubseteq}\})$ of $\mathcal{X}$ are pairwise uniformly quasi-isometric, and similarly for F_U. Abusing notation slightly, we sometimes regard E_U as a metric space by identifying it with some $\phi_U(E_U \times \{b_{\sqsubseteq}\})$, and similarly for F_U. Moreover, E_U is a hierarchically hyperbolic space, with respect to $\{V \in \mathfrak{S} : V \perp U\} \cup \{W\}$, where W is some nonmaximal element of $\mathfrak{S}$ into which each element orthogonal to U is nested; and, similarly, F_U is a hierarchically hyperbolic space with respect to $\mathfrak{S}_U$. It is straightforward to check that $\{E_U\}_{U \in \mathfrak{S}}$ are uniformly hierarchically hyperbolic with respect to the index set described above, with constants depending only on the constants for $\mathcal{X}$.

Remark 13.6 (how to induct on complexity) Observe that if U is not $\sqsubseteq$–maximal, then any $\sqsubseteq$–chain in $\mathfrak{S}_U$ is strictly contained in a $\sqsubseteq$–chain of $\mathfrak{S}$, so that the complexity of F_U is strictly less than that of $\mathcal{X}$. Moreover, by definition, there exists $W \in \mathfrak{S}$, not $\sqsubseteq$–maximal, such that each V with $V \perp U$ satisfies $V \sqsubseteq W$. Hence E_U has complexity strictly less than that of $\mathcal{X}$. We will make use of these observations when inducting on complexity in the proof of Theorem 13.11.

A *standard 1–box* is a hierarchy path $I \to F_U$, where $U \in \mathfrak{S}$. We then inductively define a *standard n–box* to be a map of the form $\iota \colon (B \times I \subseteq \mathbb{R}^{n-1} \times \mathbb{R}) \to \mathcal{X}$, with $\iota(b, t) = \phi_U(\iota_1(b), \gamma(t))$, where $\iota_1 \colon B \to E_U$ is a standard box and $\gamma \colon I \to F_U$ is a hierarchy path. It is straightforward to show inductively that a standard box is a quasi-isometric embedding.

13.2 Results

The goal of this section is to establish the following three theorems:

Theorem 13.7 *Let $\mathcal{X}$ be a hierarchically hyperbolic space. Then for every $n \in \mathbb{N}$ and every K, C, R_0, ϵ_0 the following holds. There exists R_1 such that for any ball $B \subseteq \mathbb{R}^n$ of radius at least R_1 and $f: B \to \mathcal{X}$ a (K, C)–quasi-Lipschitz map, there is a ball $B' \subseteq B$ of radius $R' \geq R_0$ such that $f(B')$ lies inside the $\epsilon_0 R'$–neighborhood of a standard box.*

Remark In Theorem 13.7, we do not require B, B' to be centered at the same point in $\mathbb{R}^n$.

Theorem 13.8 *Let $\mathcal{X}$ be a hierarchically hyperbolic space. Then for every simply connected nilpotent Lie group $\mathcal{N}$, with a left-invariant Riemannian metric, and every K, C, there exists R with the following property. For every (K, C)–quasi-Lipschitz map $f: B \to \mathcal{X}$ from a ball in $\mathcal{N}$ into $\mathcal{X}$ and for every $h \in \mathcal{N}$ we have*

$$\mathrm{diam}(f(B \cap h[\mathcal{N}, \mathcal{N}])) \leq R.$$

In particular, any finitely generated nilpotent group which quasi-isometrically embeds into $\mathcal{X}$ is virtually abelian.

Theorem 13.9 *Let $\mathcal{X}$ be a hierarchically hyperbolic space with respect to a set $\mathfrak{S}$. If there exists a quasi-isometric embedding $f: \mathbb{R}^n \to \mathcal{X}$ then n is bounded by the maximal cardinality of a set of pairwise orthogonal elements of $\mathfrak{S}$, and in particular by the complexity of $\mathcal{X}$.*

Notation 13.10 Let ω be a nonprincipal ultrafilter, which will be arbitrary, but fixed for the rest of the paper. For a sequence of spaces (X_m), we denote its ultralimit $\lim_\omega X_m$ by X; similarly, for sequences of maps, $\lim_\omega f_m = f$, etc. For scaling constants, we will have sequences of positive real numbers (r_m); when comparing two such sequences we write $(r_m) < (r'_m)$ if $r_m < r'_m$ for ω–ae m. Also, we write $(r_m) \ll (r'_m)$ if $\lim_\omega r'_m / r_m = \infty$.

Each of the above theorems follows from:

Theorem 13.11 *Let $\{\mathcal{X}_m\}$ be uniformly hierarchically hyperbolic spaces and let $\mathcal{N}$ be a simply connected nilpotent Lie group endowed with a left-invariant Riemannian metric. Let $(f_m: B_m \to \mathcal{X}_m)$ be a sequence of quasi-Lipschitz maps with uniformly bounded constants, where $B_m \subseteq \mathcal{N}$ is a ball of radius r_m^1. Then:*

(1) *For every ultralimit $\mathcal{X}$ of $(\mathcal{X}_m)$ and corresponding ultralimit map $f\colon \mathbf{B} \to \mathcal{X}$, the map f is constant along ultralimits of cosets of $[\mathcal{N}, \mathcal{N}]$.*

(2) *Suppose that $\mathcal{N}$ is abelian and let (r_m^0) satisfy $(r_m^0) \ll (r_m^1)$. Then there exist sequences (p_m) and (r_m) with $(r_m^0) \ll (r_m) \leq (r_m^1)$ and $B(p_m, r_m) \subseteq B_m$ so that the following holds. Let $\mathcal{X}$ be the ultralimit of $(\mathcal{X}_m)$ with scaling factor (r_m) and observation point $(f(p_m))$ and let $\mathbf{B}$ be the ultralimit of the sequence $B(p_m, r_m)$ with observation point (p_m) and scaling factors (r_m). Then there exists an ultralimit $\mathbf{F}$ of standard boxes such that the ultralimit map $f\colon \mathbf{B} \to \mathcal{X}$ satisfies $f(\mathbf{B}) \subseteq \mathbf{F}$.*

We first deduce the other theorems from Theorem 13.11.

Proof of Theorem 13.7 Fix $n, K, C, \epsilon_0, R_0 > 0$. If the statement was false then there would exist a sequence of maps $f_m\colon B_m \to \mathcal{X}$ where

- each B_m is a ball of radius r_m^1 about $0 \in \mathbb{R}^n$, each f_m is (K, C)–quasi-Lipschitz and $r_m^1 \to \infty$,
- no ball $B' \subseteq B_m$ of radius $R' \geq R_0$ is such that $f_m(B')$ is contained in the $\epsilon_0 R'$–neighborhood of a standard box.

In the notation of Theorem 13.11(2), if (p_m) is the observation point of $\mathbf{B}$, then it is readily checked that, ω–ae, we have $f_m(B(p_m, r_m)) \subseteq \mathcal{N}_{\epsilon_0 r_m}(F_m)$, where $\mathbf{r} = (r_m)$, $\mathbf{F} = (F_m)$. By the proof of Theorem 13.11(2), we have that $d_{\mathbb{R}^n}(p_m, 0) \leq \frac{1}{2} r_m^1$. Letting $s_m = \min\{r_m, \frac{1}{2} r_m^1\}$, and recalling that $\lim_\omega r_m^1 = \lim_\omega r^m = +\infty$, we have $s_m \geq R_0$ ω–ae; thus for sufficiently large m, taking $B' = B(p_m, s_m)$ contradicts the second item above. $\square$

Proof of Theorem 13.8 This follows directly from Theorem 13.11(1). The statement about nilpotent groups follows from the well-known fact that every finitely generated nilpotent group is virtually a lattice in a simply connected nilpotent Lie group [47]. $\square$

Proof of Theorem 13.9 Using Theorem 13.11(2), we see that some rescaled ultralimit f of f maps an asymptotic cone of $\mathbb{R}^n$, which is itself a copy of $\mathbb{R}^n$, into a rescaled ultralimit of standard boxes, whence the claim follows. $\square$

13.3 Proof of Theorem 13.11

We now complete the proofs of Theorems 13.7, 13.8, and 13.9 by proving Theorem 13.11. Throughout, $\mathcal{X}$ is a hierarchically hyperbolic space; below we use the notation from Definition 13.1.

Fix $W \in \mathfrak{S}$ and let $\gamma \colon I \to F_W$ be a hierarchy path: we now define the "projection" from $\mathcal{X}$ to γ. Roughly speaking, we map $x \in \mathcal{X}$ to any point in γ that minimizes the $\widehat{\mathcal{C}}W$–distance from x. Let $\beta = \pi_W(\gamma(I))$ and let $c_\beta \colon \widehat{\mathcal{C}}W \to \beta$ be closest-point projection. Define a map $q_\gamma \colon \mathcal{X} \to \gamma$ as follows: for each $z \in \beta$, let $t(z)$ be an arbitrary point of γ such that $\pi_W(t(z)) = z$, and then, for each $x \in \mathcal{X}$, let $q_\gamma(x) = t(c_\beta(\pi_W(x)))$.

Let P_W denote the image in $\mathcal{X}$ of the restriction of ϕ_W to $E_W \times (\operatorname{im} \gamma)$.

Lemmas 13.12 and 13.13 accomplish in the hierarchically hyperbolic setting an analogue of a result in the mapping class group proven in [8, Theorems 3.4 and 3.5]; note that in [8] the results are formulated in the asymptotic cone, rather than in the original space.

Lemma 13.12 *Let $(\mathcal{X}, \mathrm{d})$ be a hierarchically hyperbolic space. There exists a constant $\ell \geq 1$ depending only on the constants from Definition 13.1 such that the following holds. Suppose that $\gamma \colon I \to F_W$ is a hierarchy path connecting x to y and let $M = \sup_{U \in \mathfrak{S}_W - \{W\}} \sigma_{U, s_0}(x, y)$. Let $z \in \mathcal{X}$ and let $R = \mathrm{d}_{\mathcal{X}}(z, P_W)$. Then*

$$\operatorname{diam}_{\mathcal{X}}(q_\gamma(B(z, R/\ell))) \leq \ell M + \ell.$$

Proof For a suitable constant ℓ, we will prove the following: given $v, w \in \mathcal{X}$, if $\mathrm{d}(q_\gamma(v), q_\gamma(w)) > \ell M + \ell$, then any hierarchy path α from v to w intersects the ℓM–neighborhood of P_W; then the lemma will follow by increasing ℓ.

First, fix a hierarchy path α from v to w. For ℓ large enough, the δ–hyperbolicity of $\widehat{\mathcal{C}}W$ and the fact that α is a hierarchy path guarantee the existence of points t_1, t, t_2, appearing in this order along α, such that:

(1) $\mathrm{d}_{\widehat{\mathcal{C}}W}(t, \gamma) \leq 2\delta + 2D$,

(2) $\mathrm{d}_{\widehat{\mathcal{C}}W}(t, c_\beta(\pi_W(v))), d_{\widehat{\mathcal{C}}W}(t, c_\beta(\pi_W(w))) \geq 100\delta + 100D$,

(3) $\mathrm{d}_{\widehat{\mathcal{C}}W}(t, t_i) > 10\kappa_0$ for $i \in \{1, 2\}$.

Using Definition 13.1(9) (realization) we pick any $t' \in P_W$ whose $\widehat{\mathcal{C}}U$–coordinate is κ–close to that of $c_\beta(t)$ whenever $U \sqsubseteq W$ and whose $\widehat{\mathcal{C}}U$–coordinate is κ–close to that of t whenever $U \perp W$.

By choice of t' and taking the threshold in the distance formula larger than κ we have that any term in the distance formula contributing to $\mathrm{d}_{\mathcal{X}}(t, t')$ comes from some $U \in \mathfrak{S}$ which is either nested in W or transverse to it. We will first argue that the latter terms are uniformly bounded. This is because if $U \pitchfork W$ contributes to the distance formula with any threshold $\geq \kappa_0 + \kappa$, we have $\mathrm{d}_{\widehat{\mathcal{C}}U}(\rho_U^W, \pi_U(t)) > \kappa_0$ and hence $\mathrm{d}_{\widehat{\mathcal{C}}W}(\rho_W^U, \pi_W(t)) \leq \kappa_0$ by the first consistency inequality (see Definition 13.1(4)).
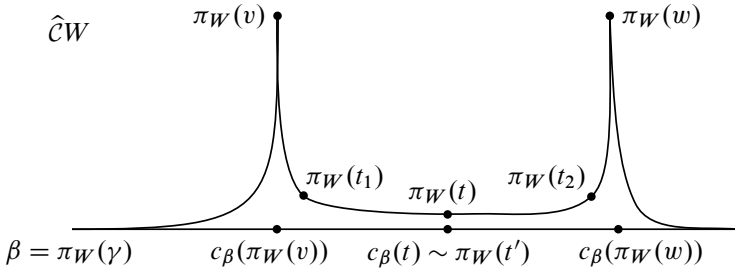
Figure 5: Sketch of Lemma 13.12

But then we also have $d_{\widehat{C}W}(\rho_W^U, \pi_W(t_i)) > \kappa_0$ and hence again by the first consistency condition we have $d_{\widehat{C}U}(\rho_U^U, \pi_U(t_i)) \leq \kappa_0$. However, since $\pi_W \circ \alpha$ is an unparametrized (D, D)–quasi-geodesic and t lies between t_1 and t_2, there is a bound on $d_{\widehat{C}U}(\rho_U^W, \pi_U(t))$ in terms of κ_0, D, δ only; this is the uniform bound we wanted. From now on, we assume that the thresholds used in the distance formula exceed this bound. We now know that $\sigma_{W,s}(t, t')$ coarsely bounds $d_{\mathcal{X}}(t, t')$ for any sufficiently large threshold s.

By Definition 13.1(7) (large link lemma), any term that contributes to $\sigma_{W,s_0}(t, t')$ is either W or an element of $\mathfrak{S}_{T_i}$, for at most $\lambda(2\delta + 2D + 1)$ elements $T_i \in \mathfrak{S}_W - \{W\}$, each within distance $\lambda(2\delta + 2D + 1)$ from $\pi_W(t)$, where $\lambda \geq 1$ is a uniform constant.

By Definition 13.1(8) (bounded geodesic image), applied to geodesics joining $\pi_W(v)$ to $\pi_W(x)$ and from $\pi_W(w)$ to $\pi_W(y)$ (or vice versa), we have $d_{\widehat{C}F}(v, x), d_{\widehat{C}F}(w, y) \leq B$ for any $F \in \mathfrak{S}_{T_i}$. Since $\pi_F(t)$ is D–close to a geodesic from $\pi_F(v)$ to $\pi_F(w)$ for $F \in \mathfrak{S}_{T_i}$ and similarly for $\pi_F(t')$ with respect to $\pi_F(v)$ and $\pi_F(w)$, we thus have

$$\sum_{F \in \mathfrak{S}_{T_i}} \{\!\{d_{\widehat{C}F}(t, t')\}\!\}_{s_0 + 2D + 2B} \leq R \sum_{F \in \mathfrak{S}_{T_i}} \{\!\{d_{\widehat{C}F}(x, y)\}\!\}_{s_0},$$

where $R = (s_0 + 2D + 2B)/s_0$. Hence

$$\begin{aligned}
\sigma_{W,s_0 + 2D + 2B}(t, t') &\leq \lambda(2\delta + 2D + 1) \\
&\quad + \lambda(2\delta + 2D + 1)R \max_i \left\{ \sum_{F \in \mathfrak{S}_{T_i}} \{\!\{d_{\widehat{C}F}(x, y)\}\!\}_{s_0} \right\} \\
&\leq \ell'M + \ell'.
\end{aligned}$$

We hence get $d_{\mathcal{X}}(t, t') \leq \ell M$, as required. □

Lemma 13.13 *Let $\{\mathcal{X}_m\}$ be uniformly hierarchically hyperbolic spaces, with respect to $\{\mathfrak{S}_m\}$, and let $\mathcal{X}$ be an ultralimit of $(\mathcal{X}_m)$. Also, let x, y be distinct points of $\mathcal{X}$. Then there exists an ultralimit F_U of the sequence (F_{U_m}), an ultralimit γ of hierarchy paths γ_m, contained in F_U and connecting distinct points x', y', and a Lipschitz map $q_\gamma \colon \mathcal{X} \to \gamma$ such that*

(1) $q_\gamma(x) = x'$, $q_\gamma(y) = y'$;

(2) q_γ restricted to P_U is the projection on the first factor, where $P_U \cong E_U \times \gamma$ is an ultralimit of (P_{U_m});

(3) q_γ is locally constant outside P_U.

Proof We first claim that there exist $U_m \in \mathfrak{S}_m$ such that

(1) $\lim_\omega (1/s_m)\sigma_{U_m, s_0}(x_m, y_m) > 0$,

(2) for any $U'_m \in \mathfrak{S}_{U_m} - \{U_m\}$, we have $\lim_\omega (1/s_m)\sigma_{U'_m, s_0}(x_m, y_m) = 0$.

In fact, since $x \neq y$, there exists some $(\widehat{C}U_m)$ satisfying the first property, by the distance formula and the existence of a $\sqsubseteq$–maximal element. Also, if $(\widehat{C}U_m)$ satisfies the first property but not the second one, then by definition we have a sequence $(\widehat{C}V_m)$ satisfying the first property and so that each V_m is ω–ae properly nested into U_m. By Definition 13.1(5), in finitely many steps we find (U_m) with the desired property.

Now let γ_m be a hierarchy path in F_{U_m} connecting $x'_m = \mathfrak{g}_{P_{U_m}}(x_m)$ to $y'_m = \mathfrak{g}_{P_{U_m}}(y_m)$. We can define q_γ to be the ultralimit of the maps q_{γ_m} as in Lemma 13.12, and all properties are easily verified. □

Proof of Theorem 13.11 The main task is to prove Theorem 13.11(2). Along the way we will point out how to adapt the first part of the argument to obtain Theorem 13.11(1).

We will prove the proposition by induction on complexity, the base case being that of complexity 0, where the distance formula implies that every ultralimit of $(\mathcal{X}_m)$ is a point.

Consider the ultralimit $\mathcal{X}^1$ of $(\mathcal{X}_m)$ with scaling factor (r_m^1) and an ultralimit map $f^1 \colon B^1 \to \mathcal{X}^1$. (For Theorem 13.11(1), we don't restrict the choice of scaling factors used in forming $\mathcal{X}^1$.)

If $f^1(B^1)$ is a single point, then the conclusion is immediate. Hence, consider $x \neq y$ in $f^1(B^1)$. Let γ, U, x', y' and $q_\gamma \colon \mathcal{X}^1 \to \gamma$ be as in Lemma 13.13. In the situation of Theorem 13.11(1), towards a contradiction, we pick $x \neq y$ in the same ultralimit of cosets of $[\mathcal{N}, \mathcal{N}]$, which we can do if $\mathcal{N}$ is not abelian. It follows by Pansu's differentiability theorem [57, Theorem 2], which applies because the asymptotic cones of $\mathcal{N}$ are Carnot groups [56], that every Lipschitz map from B^1 to $\mathbb{R}$ is constant along cosets of $[\mathcal{N}, \mathcal{N}]$; this contradicts the properties of q_γ established in Lemma 13.13 when $\mathcal{N}$ is not abelian. Theorem 13.11(1) is hence proven, and from now on we focus exclusively on Theorem 13.11(2). By Rademacher's theorem there exists $(p_m) \in (B_m^1)$ such that $\mathfrak{g}_\gamma = q_\gamma \circ f^1$ is differentiable at p and the differential is nonzero and $d_{\mathbb{R}^n}(p_m, 0) \leq \frac{1}{2}r_m^1$. In particular $B(p_m, \frac{1}{2}r_m^1) \subseteq B_m$. Abusing notation slightly, we are regarding γ as an interval in $\mathbb{R}$.

Claim 1 *For every $\epsilon > 0$ there exists $l_\epsilon > 0$ such that for any $l \leq l_\epsilon$ we have that $f^1(B(p, l))$ is contained in the ϵl–neighborhood of P_U.*

Proof of Claim 1 We identify a neighborhood of p with a neighborhood of 0 in $\mathbb{R}^n$.

We know that there exists a linear function A such that for any v_1 and v_2 in this neighborhood we have

$$|g_\gamma(v_1) - g_\gamma(v_2)| = |A(v_2 - v_1)| + o(\max\{\|v_1\|, \|v_2\|\}).$$

Moreover, consider v of norm 1 and set $Av = \theta > 0$. Given any v_1 with

$$d(f^1(v_1), E_U \times \gamma) \geq \epsilon\|v_1\|,$$

then, since f^1 is K–Lipschitz and q_γ is constant outside $E_U \times \gamma$, we have $g_\gamma(v_1) = g_\gamma(v_2)$, where $v_2 = v_1 + (\epsilon\|v_1\|/K)v$. Hence

$$0 = |g_\gamma(v_1) - g_\gamma(v_2)| = \frac{\epsilon\theta\|v_1\|}{K} + o((1 + \epsilon/K)\|v_1\|),$$

which cannot happen for $\|v_1\|$ small enough depending on ϵ. □

Claim 2 *There exists (r_m^2) with $(r_m^0) \leq (r_m^2) \ll (r_m^1)$ such that for all (r_m) with $(r_m^2) \ll (r_m) \ll (r_m^1)$ we have $f(B) \subseteq E_{U'} \times \gamma'$, where B is the ultralimit of (B_m) with observation point (p_m) and scaling factor (r_m).*

Claim 2 follows from Claim 1 by an application of the principle from nonstandard analysis called *underspill*; nonetheless, we provide a proof in the interest of self-containment.

Proof of Claim 2 In view of Claim 1, there exists a function $\epsilon \colon \mathbb{N} \to \mathbb{R}^+$ such that $\epsilon(k) \to 0$ as $k \to \infty$ and such that for every k the set $A_k \subseteq \mathbb{N}$, defined by

$$A_k = \{m \mid f_m(B(p_m, r_m^1/2^k)) \subseteq \mathcal{N}_{\epsilon(k)r_m^1/2^k}(E_{U_m} \times \gamma_m)\},$$

satisfies $\omega(A_k) = 1$. Let $k(m) = \max\{k \leq m : m \in \bigcap_{i \leq k} A_i\}$, which by the above is well-defined ω–ae and satisfies $\lim_\omega k(m) = \infty$. By definition of $k(m)$, for ω–ae m we have $f_m(B(p_m, r_m^1/2^k)) \subseteq \mathcal{N}_{\epsilon(k)r_m^1/2^k}(E_{U_m} \times \gamma_m)$ for every $k \leq k(m)$.

Let $(r_m^2) = \max\{(r_m^0), (r_m^1/2^{k(m)})\}$. We claim that if $(r_m^2) \ll (r_m) \ll (r_m^1)$, then the result holds for the scaling factor (r_m). To do this, we show for each $j \in \mathbb{N}$ that $\omega(A_j') = 1$, where

$$A_j' = \{m \mid f_m(B(p_m, j\,r_m)) \subseteq \mathcal{N}_{r_m/j}(E_{U_m} \times \gamma_m)\}.$$

Fix j. For a given m, let $k'(m)$ satisfy $r_m^1/2^{k'(m)+1} < j r_m \le r_m^1/2^{k'(m)}$. For ω–ae m we have $k'(m) \le k(m)$ since $(r_m) \gg (r_m^1/2^{k(m)})$. In particular, for ω–ae m we have

$$
\begin{aligned}
f_m(B(p_m, j r_m)) &\subseteq f_m(B(p_m, r_m^1/2^{k'(m)})) \\
&\subseteq \mathcal{N}_{\epsilon(k'(m)) r_m^1/2^{k'(m)}}(E_{U_m} \times \gamma_m) \\
&\subseteq \mathcal{N}_{2\epsilon(k'(m)) j r_m}(E_{U_m} \times \gamma_m).
\end{aligned}
$$

Since $(r_m) \ll (r_m^1)$, we have $\lim_\omega k'(m) = \infty$, and hence $2\epsilon(k'(m)) \le 1/j^2$ for ω–ae m; the claim follows. $\qquad\square$

By induction on complexity, enabled by Remark 13.6, we know that for some $(r_m) \gg (r_m^2)$ the ultralimit $\boldsymbol{g}$ of $g_m = \mathfrak{g}_{E_{U_m}} \circ f_m$ maps $\boldsymbol{B}$ into some ultralimit $\boldsymbol{F}'$ of standard boxes, whence it is readily deduced that $\boldsymbol{f}(\boldsymbol{B})$ is contained in the ultralimit $\boldsymbol{F}' \times \boldsymbol{\gamma}'$ of standard boxes. Note that, provided $r_m \le \frac{1}{2} r_m^1$, the ultralimit $\boldsymbol{B}$ contains the required ultralimit of $(B(p_m, r_m))$. $\qquad\square$

14 Acylindricity

Definition 14.1 (automorphism of a hierarchically hyperbolic space) Let $\mathcal{X}$ be a hierarchically hyperbolic space, with respect to $\mathfrak{S}$. An *automorphism* g of $(\mathcal{X}, \mathfrak{S})$ is a bijection $g\colon \mathfrak{S} \to \mathfrak{S}$ and a collection $\{g_U\colon \hat{\mathcal{C}}U \to \hat{\mathcal{C}}(gU) : U \in \mathfrak{S}\}$ of isometries such that for all $U, V \in \mathfrak{S}$ we have $g_V(\rho_V^U) = \rho_{gV}^{gU}$ and such that g preserves $\bot$, $\sqsubseteq$ and $\pitchfork$.

By items (6) and (9) of Definition 13.1, every automorphism g of $\mathfrak{S}$ induces a quasi-isometry $\phi_g\colon \mathcal{X} \to \mathcal{X}$ with uniformly bounded constants. Such a quasi-isometry ϕ_g can be described as follows. For $x \in \mathcal{X}$ and $U \in \mathfrak{S}$, $\pi_{gU}(\phi_g(x))$ coarsely coincides with $g_U(\pi_U(x))$. When it will not introduce confusion, we will use the notation "g" for the quasi-isometry ϕ_g as well as for the element $g \in \mathrm{Aut}(\mathfrak{S})$. In the remainder of this section, we are interested in the situation where the action of G on $\mathcal{X}$ by uniform quasi-isometries is proper and cocompact.

Let $S = S(\mathcal{X}, \mathfrak{S})$ be the unique $\sqsubseteq$–maximal element and consider $G \le \mathrm{Aut}(\mathfrak{S})$. Since each $g \in G$ preserves the nesting relation, G fixes S, and hence g_S is an isometry of $\hat{\mathcal{C}}S$ for each $g \in G$. Note that this action coarsely preserves the union of the sets in $\pi_S(\mathcal{X})$.

Remark 14.2 When $\mathcal{X}$ is a cube complex with a factor system, $S = \mathcal{X}$ and $\hat{\mathcal{C}}S$ is the factored contact graph $\hat{\mathcal{C}}\mathcal{X}$. The union of sets in the image of $\pi_{\mathcal{X}}\colon \mathcal{X} \to 2^{\hat{\mathcal{C}}\mathcal{X}^{(0)}}$ is coarsely equal to $\hat{\mathcal{C}}\mathcal{X}$, by the definition of $\pi_{\mathcal{X}}$.

Theorem 14.3 *Let the group* $G \leq \mathrm{Aut}(\mathfrak{S})$ *act properly and cocompactly on the hierarchically hyperbolic space* $\mathcal{X}$ *and let* S *be the unique* $\sqsubseteq$*–maximal element of* $\mathfrak{S}$. *Then for all* $\epsilon > 0$, *there exist constants* R, N *such that there are at most* N *elements* $g \in G$ *such that* $\mathrm{d}_{\widehat{\mathcal{C}}S}(\pi_S(x), \pi_S(gx)) < \epsilon$ *and* $\mathrm{d}_{\widehat{\mathcal{C}}S}(\pi_S(y), \pi_S(gy)) < \epsilon$ *whenever* $x, y \in \mathcal{X}$ *satisfy* $\mathrm{d}_{\widehat{\mathcal{C}}S}(\pi_S(x), \pi_S(y)) \geq R$.

Before proving Theorem 14.3, we note two corollaries. Recall that the action of the group G on the metric space (M, d) is *acylindrical* if for all $\epsilon > 0$ there exist constants R, N such that there are at most N elements $g \in G$ such that $\mathrm{d}(x, gx) < \epsilon$ and $\mathrm{d}(y, gy) < \epsilon$ whenever $x, y \in M$ satisfy $\mathrm{d}(x, y) \geq R$. (Note that if M is bounded, then any action on M is automatically acylindrical. In particular, if $\mathcal{X}$ is a cube complex decomposing as the product of unbounded subcomplexes and G acts geometrically on $\mathcal{X}$, then the action of G on $\mathcal{C}\mathcal{X}$ is trivially acylindrical.)

Corollary 14.4 (acylindrical hyperbolicity) *Let* $G \leq \mathrm{Aut}(\mathfrak{S})$ *act properly and cocompactly on the hierarchically hyperbolic space* $\mathcal{X}$. *Then* G *acts acylindrically on a hyperbolic space quasi-isometric to* $\mathcal{U} = \bigcup_{x \in \mathcal{X}} \pi_S(x)$. *In particular, if* $\mathcal{U}$ *is unbounded and* G *is not virtually cyclic, then* G *is acylindrically hyperbolic.*

Proof Let S be as in Theorem 14.3. Let $\mathcal{T}_0 \subseteq \widehat{\mathcal{C}}S$ be the union of all sets of the form $\pi_S(x)$ with $x \in \mathcal{X}$, together with all of their G–orbits, so that $\mathcal{T}_0$ is G–invariant and coarsely equal to the union of the elements of $\pi_S(\mathcal{X})$. Definition 13.1(10) thus ensures that $\mathcal{T}_0$ is quasiconvex in the hyperbolic space $\widehat{\mathcal{C}}S$, so that we can add geodesics of $\widehat{\mathcal{C}}S$ to $\mathcal{T}_0$ to form a G–hyperbolic space $\mathcal{T}$ that is G–equivariantly quasi-isometric to the union of the elements of $\pi_S(\mathcal{X})$. The action of G on $\mathcal{T}$ is thus acylindrical by Theorem 14.3. $\qquad \square$

In the cubical case, acylindricity can be witnessed by the contact graph instead of the factored contact graph.

Corollary 14.5 *Let* G *act properly and cocompactly on the* $\mathrm{CAT}(0)$ *cube complex* $\mathcal{X}$, *and suppose that* $\mathcal{X}$ *contains a* G*–invariant factor system. Then* G *acts acylindrically on* $\widehat{\mathcal{C}}\mathcal{X}$, *and hence on* $\mathcal{C}\mathcal{X}$.

Proof For any G–invariant factor system $\mathfrak{F}$, the action of G on $\widehat{\mathcal{C}}\mathcal{X}$ is acylindrical by Theorem 14.3 and Remarks 13.2 and 14.2. Let $\mathfrak{F}$ be a factor system in $\mathcal{X}$ and let ξ be the constant for $\mathfrak{F}$ from Definition 8.1. Let $\mathfrak{F}_0$ be the smallest set of convex subcomplexes of $\mathcal{X}$ that contains $\mathcal{X}$ and each subcomplex parallel to a combinatorial hyperplane, and has the property that $\mathfrak{g}_F(F') \in \mathfrak{F}_0$ whenever $F, F' \in \mathfrak{F}_0$ and $\mathrm{diam}(\mathfrak{g}_F(F')) \geq \xi$.

By definition, $\mathfrak{F}_0 \subseteq \mathfrak{F}$, so $\mathfrak{F}_0$ has bounded multiplicity and thus $\mathfrak{F}_0$ is a factor system. The associated factored contact graph of $\mathcal{X}$, on which G acts acylindrically, is G–equivariantly quasi-isometric to $\mathcal{CX}$, and the result follows. $\qquad\qquad\square$

Proof of Theorem 14.3 Fix $\epsilon > 0$, and for convenience assume that it is 100 times larger than all the constants in Definition 13.1. Let $R_0 \geq 1000\epsilon$ and consider $x, y \in \mathcal{X}$ such that $R = d_{\widehat{\mathcal{C}}S}(\pi_S(x), \pi_S(y)) \geq R_0$, and let $\mathfrak{H}$ be the set of all $g \in G$ such that $d_{\widehat{\mathcal{C}}S}(\pi_S(gx), \pi_S(x)) < \epsilon$ and $d_{\widehat{\mathcal{C}}S}(\pi_S(gy), \pi_S(y)) < \epsilon$.

We will consider, roughly speaking, the set of all $U \in \mathfrak{S}$ such that x, y project far away in U, the corresponding ρ_S^U is near the middle of a geodesic from $\pi_S(x)$ to $\pi_S(y)$, and U is $\sqsubseteq$–maximal with these properties. We do so because these U correspond to product regions "in the middle" between x and y. Formally, let $\mathfrak{L}_1$ be the set of all $U \in \mathfrak{S} - \{S\}$ with the following properties:

(1) $d_{\widehat{\mathcal{C}}U}(\pi_U(x), \pi_U(y)) > \epsilon$;

(2) $\left| d_{\widehat{\mathcal{C}}S}(\pi_S(x), \rho_S^U) - \frac{1}{2}R \right| \leq 10\epsilon$;

(3) U is not properly nested into any $U' \in \mathfrak{S} - \{S\}$ with $d_{\widehat{\mathcal{C}}U'}(\pi_{U'}(x), \pi_{U'}(y)) > \epsilon$.

When applying an element g that moves x, y a bounded amount, any $U \in \mathfrak{L}_1$ gets moved to some gU with similar properties but slightly worse constants. To capture this, we let $\mathfrak{L}_2$ be the set of all $U \in \mathfrak{S} - \{S\}$ such that

(1) $d_{\widehat{\mathcal{C}}U}(\pi_U(x), \pi_U(y)) > \frac{1}{2}\epsilon$;

(2) $\left| d_{\widehat{\mathcal{C}}S}(\pi_S(x), \rho_S^U) - \frac{1}{2}R \right| \leq 11\epsilon$;

(3) U is not properly nested into any $U' \in \mathfrak{S} - \{S\}$ with $d_{\widehat{\mathcal{C}}U'}(\pi_{U'}(x), \pi_{U'}(y)) > 2\epsilon$.

Bounding $|\mathfrak{L}_2|$ Consider a hierarchy path γ from x to y. Then there are x', y' on γ such that

- $d_{\widehat{\mathcal{C}}S}(\pi_S(x'), \pi_S(y')) \leq 23\epsilon$, and

- whenever $U \in \mathfrak{S} - \{S\}$ is such that

$$d_{\widehat{\mathcal{C}}U}(\pi_U(x), \pi_U(y)) > \tfrac{1}{2}\epsilon \quad \text{and} \quad \left| d_{\widehat{\mathcal{C}}S}(\pi_S(x), \rho_S^U) - \tfrac{1}{2}R \right| \leq 11\epsilon,$$

  then $d_{\widehat{\mathcal{C}}U}(\pi_U(x'), \pi_U(y')) > s_0$ (recall that s_0 is the minimal threshold of the distance formula).

The existence of x', y' follows since we can choose x' and y' projecting close to points on a $\widehat{\mathcal{C}}S$–geodesic from $\pi_S(x)$ to $\pi_S(y)$ that lie on opposite sides of the midpoint, at distance slightly larger than 11ϵ from the midpoint. Bounded geodesic image

(Definition 13.1(8)) guarantees that the second condition holds because it ensures that $\pi_U(x), \pi_U(y)$ coarsely coincide with $\pi_U(x'), \pi_U(y')$ (recall that ϵ is much larger than s_0 and all other constants in Definition 13.1).

By Definition 13.1(7) (large link lemma), with $W = S$ and x, x' replaced by x', y', each $U \in \mathfrak{L}_2$ is nested into one of at most $23\epsilon\lambda + \lambda$ elements T of $\mathfrak{S}$. For p as in the claim below, the number of $U \in \mathfrak{L}_2$ nested into the same such T is bounded by p, for otherwise some U (the U_i in the conclusion of the claim) would fail to satisfy the third property in the definition of $\mathfrak{L}_2$. Hence $|\mathfrak{L}_2| \leq p\lambda(23\epsilon + 1)$.

Claim *There exists p with the following property. Let $T \in \mathfrak{S}$, let $x, y \in \mathcal{X}$, and let $\{U_i\}_{i=1}^p \subseteq \mathfrak{S}_T$ be distinct and satisfy $d_{\widehat{C}U_i}(\pi_{U_i}(x), \pi_{U_i}(y)) \geq \epsilon$. Then there exist $U' \in \mathfrak{S}_T$ and i such that $U_i \subsetneq U'$ and $d_{\widehat{C}U'}(\pi_{U'}(x), \pi_{U'}(y)) > 2\epsilon$.*

Proof of claim Define the level of $Y \in \mathfrak{S}$ to be the maximal k such that there exists a $\sqsubseteq$–chain of length k in $\mathfrak{S}_Y$. The proof is by induction on the level k of a $\sqsubseteq$–minimal $T' \in \mathfrak{S}_T$ into which each S_i is nested. For the base case $k = 1$ it suffices to take $p = 2$ since in this case there is no pair of distinct $U_1, U_2 \in \mathfrak{S}_{T'}$.

Suppose that the statement holds for a given $p(k)$ when the level of T' is at most k. Suppose further that $|\{U_i\}| \geq p(k+1)$ (where $p(k+1)$ is a constant, much larger than $p(k)$, that will be determined shortly) and there exists a $\sqsubseteq$–minimal $T' \in \mathfrak{S}_T$ of level $k + 1$ into which each U_i is nested. There are two cases.

If $d_{\widehat{C}T'}(\pi_{T'}(x), \pi_{T'}(y)) > 2\epsilon$, then we are done (for $p \geq 2$). If not, then item (7) of Definition 13.1 (large link lemma) yields K and $T_1, \ldots, T_K$, each properly nested into T' (and hence of level $\leq k$), such that any given U_i is nested into some T_j. In particular, if $p(k+1) \geq Kp(k)$, there exists j such that at least $p(k)$ elements of $\{U_i\}$ are nested into T_j. By the induction hypothesis and Definition 13.1(5) (finite complexity), we are done. $\qquad \square$

Case $\mathfrak{L}_1 \neq \varnothing$ Suppose that there exists $U \in \mathfrak{L}_1$. In this case the idea is that there actually are product regions in the middle between x, y, and the action of an element moving x, y not too much permutes the gates into such product regions, so that there cannot be too many such elements.

First of all, our choice of R_0 and Definition 13.1(8) (bounded geodesic image) ensure that for all $g \in \mathfrak{H}$, we have $gU \in \mathfrak{L}_2$. Indeed, if $U \in \mathfrak{L}_1$ then

$$d_{\widehat{C}gU}(\pi_{gU}(x), \pi_{gU}(y)) \geq d_{\widehat{C}gU}(\pi_{gU}(gx), \pi_{gU}(gy)) - 2B \geq \epsilon - 2B > \tfrac{1}{2}\epsilon,$$

ensuring the first property in the definition of $\mathfrak{L}_2$. The second property follows from the fact that $\pi_S(x)$ gets moved distance $\leq \epsilon$ by g, and the third property holds for gU

because otherwise, using bounded geodesic image, we would find a contradiction with the third property of U from the definition of $\mathfrak{L}_1$.

Fix any $g \in \mathfrak{H}$ and let $P_1 = E_U \times F_U$ and $P_2 = E_{gU} \times F_{gU}$ be the spaces provided by Section 13.1. We claim that $d_{\mathcal{X}}(g\mathfrak{g}_{P_1}(x), \mathfrak{g}_{P_2}(x))$ is uniformly bounded. From this, properness of the action and the bound on $|\mathfrak{L}_2|$ yield a bound on $|\mathfrak{H}|$.

Clearly, $g\mathfrak{g}_{P_1}(x)$ coarsely coincides with $\mathfrak{g}_{P_2}(gx)$, since coordinates of gates are defined equivariantly. Hence we must show that $w = \mathfrak{g}_{P_2}(gx)$ coarsely coincides with $z = \mathfrak{g}_{P_2}(x)$.

By Definition 13.1(6) (distance formula), it suffices to show that the projections of w and z coarsely coincide in every $\widehat{\mathcal{C}}Y$ for $Y \in \mathfrak{S}$. By definition of $\mathfrak{g}_{P_2}$, it suffices to consider $Y \in \mathfrak{S}$ which is either nested into or orthogonal to gU. For such Y, ρ_S^Y coarsely coincides with ρ_S^{gU} by the final part of Definition 13.1(4) (transversality and consistency). Moreover, any geodesic from x to gx stays far from ρ_S^Y, so that Definition 13.1(8) (bounded geodesic image) gives a uniform bound on $d_{\widehat{\mathcal{C}}Y}(\pi_Y(x), \pi_Y(gx))$, as required.

Case $\mathfrak{L}_1 = \varnothing$ Suppose now that $\mathfrak{L}_1 = \varnothing$. In this case, the idea is that a hierarchy path from x to y does not spend much time in any product region near the middle, and hence it behaves like a geodesic in a hyperbolic space. Fix a hierarchy path γ in $\mathcal{X}$ joining x, y and let $p \in \gamma$ satisfy $\left| d_{\widehat{\mathcal{C}}S}(\pi_S(p), \pi_S(x)) - \frac{1}{2}R \right| \le \epsilon$. We will produce a constant M_3, depending on the constants from Definition 13.1 and on ϵ such that $d_{\mathcal{X}}(p, gp) \le M_3$. It will then follow that $|\mathfrak{H}|$ is bounded in view of properness of the action of G on $\mathcal{X}$.

We now bound $d_{\mathcal{X}}(p, gp)$ using the distance formula. First, note that

$$d_{\widehat{\mathcal{C}}S}(\pi_S(p), \pi_S(gp)) \le 10\delta + \epsilon$$

(where δ is the hyperbolicity constant of $\widehat{\mathcal{C}}S$).

If $U \in \mathfrak{S}$ contributes to the sum $\sigma_{\mathcal{X}, s_0}(p, gp)$ with threshold s_0, then, given the bound on $d_{\widehat{\mathcal{C}}S}(p, gp)$ and bounded geodesic image, $d_{\widehat{\mathcal{C}}S}(\pi_S(p), \rho_S^U) \le 2\epsilon$. Fix now any U satisfying $d_{\widehat{\mathcal{C}}S}(\pi_S(p), \rho_S^U) \le 2\epsilon$. Our goal is now to bound $d_{\widehat{\mathcal{C}}U}(\pi_U(p), \pi_U(gp))$ uniformly. It follows from the assumption that $\mathfrak{L}_1 = \varnothing$ that $d_{\widehat{\mathcal{C}}U}(\pi_U(x), \pi_U(y)) \le 3\epsilon$, and also

$$d_{\widehat{\mathcal{C}}U}(\pi_U(gx), \pi_U(gy)) = d_{\widehat{\mathcal{C}}g^{-1}U}(\pi_{g^{-1}U}(x), \pi_{g^{-1}U}(y)) \le 3\epsilon.$$

Indeed, if, say, U satisfied $d_{\widehat{\mathcal{C}}U}(\pi_U(x), \pi_U(y)) > 3\epsilon$, then either U or some other U' with $d_{\widehat{\mathcal{C}}U'}(\pi_{U'}(x), \pi_{U'}(y)) > \epsilon$ and $U \sqsubseteq U'$ would belong to $\mathfrak{L}_1$, since for $U \sqsubseteq U'$ and $U' \ne S$, the sets ρ_S^U and $\rho_S^{U'}$ coarsely coincide.

Since γ is a hierarchy path, $\mathsf{d}_{\widehat{C}U}(\pi_U(x), \pi_U(p))$ and $d_{\widehat{C}U}(\pi_U(gx), \pi_U(gp))$ are bounded by $\mathsf{d}_{\widehat{C}U}(\pi_U(x), \pi_U(y)) + 2D$ and $d_{\widehat{C}U}(\pi_U(gx), \pi_U(gy)) + 2D$, respectively, which are both bounded by, say, 4ϵ. Hence,

$$\mathsf{d}_{\widehat{C}U}(\pi_U(p), \pi_U(gp)) \leq 9\epsilon + \mathsf{d}_{\widehat{C}U}(\pi_U(x), \pi_U(gx))$$

(where the diameters of the projection sets are taken care of by the extra ϵ), and the last term is uniformly bounded by bounded geodesic image. We then get the desired bound, concluding the proof. $\square$

Corollary 14.6 *Let $G \leq \mathrm{Aut}(\mathfrak{S})$ act properly and cocompactly on the hierarchically hyperbolic space $\mathcal{X}$. Let $g, h \in G$ be hyperbolic on the maximal $S \in \mathfrak{S}$ and satisfy $g^n h \neq h g^n$, $h^n g \neq g h^n$ for all $n \neq 0$. Then there exists $N > 0$, depending on g and h, such that $\langle g^N, h^N \rangle$ is freely generated by g^N, h^N.*

Proof By Corollary 14.4, G acts acylindrically on a hyperbolic space, whence the claim follows from [31, Proposition 2.4]. $\square$

In particular, when G acts properly and cocompactly on a CAT(0) cube complex with a factor system, eg when G is compact special, the conclusion of Corollary 14.6 is satisfied. Thus acylindricity can be used to find free subgroups of groups acting on cube complexes by different means than are used in the discussion of the Tits alternative [21; 63]. The above corollary recovers Theorem 47 of [44] about subgroups of right-angled Artin groups generated by powers of elements acting loxodromically on the extension graph once we observe, as in [44], that there is a quasi-isometry from the extension graph to the contact graph (whenever the right-angled Artin group does not have a free $\mathbb{Z}$ factor).

In the case $\mathcal{X}$ is a uniformly locally finite cube complex, not necessarily equipped with a factor system, one can obtain the same conclusion, provided $g, h \in \mathrm{Aut}(\mathcal{X})$ act loxodromically on $C\mathcal{X}$, by a ping-pong argument.

References

[1] **I Agol**, *The virtual Haken conjecture*, Doc. Math. 18 (2013) 1045–1087 MR

[2] **H-J Bandelt**, **V Chepoi**, *Metric graph theory and geometry: a survey*, from "Surveys on discrete and computational geometry" (J E Goodman, J Pach, R Pollack, editors), Contemp. Math. 453, Amer. Math. Soc. (2008) 49–86 MR

[3] **J A Behrstock**, *Asymptotic geometry of the mapping class group and Teichmüller space*, Geom. Topol. 10 (2006) 1523–1578 MR

[4] **J Behrstock**, **R Charney**, *Divergence and quasimorphisms of right-angled Artin groups*, Math. Ann. 352 (2012) 339–356 MR

[5] **J Behrstock**, **M F Hagen**, *Cubulated groups: thickness, relative hyperbolicity, and simplicial boundaries*, Groups Geom. Dyn. 10 (2016) 649–707 MR

[6] **J Behrstock**, **M F Hagen**, **A Sisto**, *Hierarchically hyperbolic spaces II: Combination theorems and the distance formula*, preprint (2015) arXiv

[7] **J Behrstock**, **B Kleiner**, **Y Minsky**, **L Mosher**, *Geometry and rigidity of mapping class groups*, Geom. Topol. 16 (2012) 781–888 MR

[8] **J A Behrstock**, **Y N Minsky**, *Dimension and rank for mapping class groups*, Ann. of Math. 167 (2008) 1055–1077 MR

[9] **N Bergeron**, **D T Wise**, *A boundary criterion for cubulation*, Amer. J. Math. 134 (2012) 843–859 MR

[10] **M Bestvina**, **K Bromberg**, **K Fujiwara**, *Bounded cohomology with coefficients in uniformly convex Banach spaces*, preprint (2013) arXiv

[11] **M Bestvina**, **K Bromberg**, **K Fujiwara**, *Constructing group actions on quasi-trees and applications to mapping class groups*, Publ. Math. Inst. Hautes Études Sci. 122 (2015) 1–64 MR

[12] **M Bestvina**, **M Feighn**, *Subfactor projections*, J. Topol. 7 (2014) 771–804 MR

[13] **M Bestvina**, **K Fujiwara**, *Bounded cohomology of subgroups of mapping class groups*, Geom. Topol. 6 (2002) 69–89 MR

[14] **M Bestvina**, **B Kleiner**, **M Sageev**, *Quasiflats in* CAT(0) *2–complexes*, Algebr. Geom. Topol. 16 (2016) 2663–2676 MR

[15] **B H Bowditch**, *Tight geodesics in the curve complex*, Invent. Math. 171 (2008) 281–300 MR

[16] **J F Brock**, *The Weil–Petersson metric and volumes of* 3*–dimensional hyperbolic convex cores*, J. Amer. Math. Soc. 16 (2003) 495–535 MR

[17] **J F Brock**, **R D Canary**, **Y N Minsky**, *The classification of Kleinian surface groups, II: The ending lamination conjecture*, Ann. of Math. 176 (2012) 1–149 MR

[18] **M Burger**, **N Monod**, *Continuous bounded cohomology and applications to rigidity theory*, Geom. Funct. Anal. 12 (2002) 219–280 MR

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

[20] **P-E Caprace**, **N Monod**, *Isometry groups of non-positively curved spaces: structure theory*, J. Topol. 2 (2009) 661–700 MR

[21] **P-E Caprace**, **M Sageev**, *Rank rigidity for* CAT(0) *cube complexes*, Geom. Funct. Anal. 21 (2011) 851–891 MR

[22] **R Charney**, *An introduction to right-angled Artin groups*, Geom. Dedicata 125 (2007) 141–158 MR

[23] **R Charney**, **M W Davis**, *Finite $K(\pi, 1)s$ for Artin groups*, from "Prospects in topology" (F Quinn, editor), Ann. of Math. Stud. 138, Princeton Univ. Press (1995) 110–124 MR

[24] **I Chatterji**, **G Niblo**, *From wall spaces to* CAT(0) *cube complexes*, Internat. J. Algebra Comput. 15 (2005) 875–885 MR

[25] **V Chepoi**, *Graphs of some* CAT(0) *complexes*, Adv. in Appl. Math. 24 (2000) 125–179 MR

[26] **V Chepoi**, **M F Hagen**, *On embeddings of* CAT(0) *cube complexes into products of trees via colouring their hyperplanes*, J. Combin. Theory Ser. B 103 (2013) 428–467 MR

[27] **M W Davis**, *The geometry and topology of Coxeter groups*, London Mathematical Society Monographs Series 32, Princeton Univ. Press (2008) MR

[28] **C Druțu**, *Quasi-isometry invariants and asymptotic cones*, Internat. J. Algebra Comput. 12 (2002) 99–135 MR

[29] **M G Durham**, *The augmented marking complex of a surface*, J. Lond. Math. Soc. 94 (2016) 933–969 MR

[30] **A Eskin**, **H Masur**, **K Rafi**, *Large scale rank of Teichmüller space*, preprint (2013) arXiv

[31] **K Fujiwara**, *Subgroups generated by two pseudo-Anosov elements in a mapping class group, I: Uniform exponential growth*, from "Groups of diffeomorphisms" (R Penner, D Kotschick, T Tsuboi, N Kawazumi, T Kitano, Y Mitsumatsu, editors), Adv. Stud. Pure Math. 52, Math. Soc. Japan, Tokyo (2008) 283–296 MR

[32] **M F Hagen**, *The simplicial boundary of a* CAT(0) *cube complex*, Algebr. Geom. Topol. 13 (2013) 1299–1367 MR

[33] **M F Hagen**, *Weak hyperbolicity of cube complexes and quasi-arboreal groups*, J. Topol. 7 (2014) 385–418 MR

[34] **M F Hagen**, **P Przytycki**, *Cocompactly cubulated graph manifolds*, Israel J. Math. 207 (2015) 377–394 MR

[35] **M F Hagen**, **D T Wise**, *Cubulating hyperbolic free-by-cyclic groups: the general case*, Geom. Funct. Anal. 25 (2015) 134–179 MR

[36] **F Haglund**, *Finite index subgroups of graph products*, Geom. Dedicata 135 (2008) 167–209 MR

[37] **F Haglund**, **D T Wise**, *Special cube complexes*, Geom. Funct. Anal. 17 (2008) 1551–1620 MR

[38] **U Hamenstädt**, *Geometry of the mapping class groups, III: Quasi-isometric rigidity*, preprint (2005) arXiv

[39] **J Huang**, *Top dimensional quasiflats in* CAT(0) *cube complexes*, preprint (2014) arXiv

[40] **D Hume**, *Embedding mapping class groups into finite products of trees*, preprint (2012) arXiv

[41] **N V Ivanov**, *Automorphism of complexes of curves and of Teichmüller spaces*, Internat. Math. Res. Notices (1997) 651–666 MR

[42] **J Kahn**, **V Markovic**, *Immersing almost geodesic surfaces in a closed hyperbolic three manifold*, Ann. of Math. 175 (2012) 1127–1190 MR

[43] **S-h Kim**, **T Koberda**, *Embedability between right-angled Artin groups*, Geom. Topol. 17 (2013) 493–530 MR

[44] **S-H Kim**, **T Koberda**, *The geometry of the curve graph of a right-angled Artin group*, Internat. J. Algebra Comput. 24 (2014) 121–169 MR

[45] **B Kleiner**, *The local structure of length spaces with curvature bounded above*, Math. Z. 231 (1999) 409–456 MR

[46] **J Maher**, **G Tiozzo**, *Random walks on weakly hyperbolic groups*, preprint (2014) arXiv

[47] **A I Mal'cev**, *On a class of homogeneous spaces*, Izvestiya Akad. Nauk. SSSR. Ser. Mat. 13 (1949) 9–32 MR In Russian

[48] **J F Manning**, *Geometry of pseudocharacters*, Geom. Topol. 9 (2005) 1147–1185 MR

[49] **H A Masur**, **Y N Minsky**, *Geometry of the complex of curves, I: Hyperbolicity*, Invent. Math. 138 (1999) 103–149 MR

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

[51] **P Mathieu**, **A Sisto**, *Deviation inequalities and CLT for random walks on acylindrically hyperbolic groups*, preprint (2014) arXiv

[52] **N Monod**, *Continuous bounded cohomology of locally compact groups*, Lecture Notes in Math. 1758, Springer (2001) MR

[53] **A Nevo**, **M Sageev**, *The Poisson boundary of* CAT(0) *cube complex groups*, Groups Geom. Dyn. 7 (2013) 653–695 MR

[54] **G A Niblo**, **L D Reeves**, *Coxeter groups act on* CAT(0) *cube complexes*, J. Group Theory 6 (2003) 399–413 MR

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

[56] **P Pansu**, *Croissance des boules et des géodésiques fermées dans les nilvariétés*, Ergodic Theory Dynam. Systems 3 (1983) 415–445 MR

[57] **P Pansu**, *Métriques de Carnot–Carathéodory et quasiisométries des espaces symétriques de rang un*, Ann. of Math. 129 (1989) 1–60 MR

[58] **S D Pauls**, *The large scale geometry of nilpotent Lie groups*, Comm. Anal. Geom. 9 (2001) 951–982 MR

[59] **K Rafi**, *A combinatorial model for the Teichmüller metric*, Geom. Funct. Anal. 17 (2007) 936–959 MR

[60] **K Ruane**, **S Witzel**, CAT(0) *cubical complexes for graph products of finitely generated abelian groups*, New York J. Math. 22 (2016) 637–651 MR

[61] **L Sabalka**, **D Savchuk**, *Submanifold projection*, preprint (2012) arXiv

[62] **M Sageev**, *Ends of group pairs and non-positively curved cube complexes*, Proc. London Math. Soc. 71 (1995) 585–617 MR

[63] **M Sageev**, **D T Wise**, *The Tits alternative for* CAT(0) *cubical complexes*, Bull. London Math. Soc. 37 (2005) 706–710 MR

[64] **P Scott**, *Subgroups of surface groups are almost geometric*, J. London Math. Soc. 17 (1978) 555–565 MR

[65] **A Sisto**, *Projections and relative hyperbolicity*, Enseign. Math. 59 (2013) 165–181 MR

[66] **S J Taylor**, *Right-angled Artin groups and* Out($\mathbb{F}_n$)*, I: Quasi-isometric embeddings*, Groups Geom. Dyn. 9 (2015) 275–316 MR

[67] **W P Thurston**, *On the geometry and dynamics of diffeomorphisms of surfaces*, Bull. Amer. Math. Soc. 19 (1988) 417–431 MR

[68] **J H C Whitehead**, *Combinatorial homotopy, I*, Bull. Amer. Math. Soc. 55 (1949) 213–245 MR

[69] **J H C Whitehead**, *Combinatorial homotopy, II*, Bull. Amer. Math. Soc. 55 (1949) 453–496 MR

[70] **D T Wise**, *Cubulating small cancellation groups*, Geom. Funct. Anal. 14 (2004) 150–214 MR

[71] **D T Wise**, *Complete square complexes*, Comment. Math. Helv. 82 (2007) 683–724 MR

[72] **D T Wise**, *The structure of groups with a quasiconvex hierarchy*, preprint (2011) Available at https://tinyurl.com/Wise-quasiconvex

[73] **D T Wise**, *From riches to raags:* 3*–manifolds, right-angled Artin groups, and cubical geometry*, CBMS Regional Conference Ser. Math. 117, Amer. Math. Soc., Providence, RI (2012) MR

The Graduate Center and Lehman College, CUNY
365 5th Avenue, New York, NY 10036, United States

Current address: *Barnard College, Columbia University*
New York, NY, USA

Department of Pure Mathematics and Mathematical Statistics, University of Cambridge
Wilberforce Road, Cambridge, CB3 0WB, United Kingdom

Departement Mathematik HG G 28, ETH
Rämistrasse 101, CH-8092 Zürich, Switzerland

jason@math.columbia.edu,   mfh26@cam.ac.uk,   sisto@math.ethz.ch