

# Residual finiteness, QCERF, and fillings of hyperbolic groups

IAN AGOL  
DANIEL GROVES  
JASON FOX MANNING

We prove that if every hyperbolic group is residually finite, then every quasi-convex subgroup of every hyperbolic group is separable. The main tool is relatively hyperbolic Dehn filling.

[20E26](#), [20F67](#)

A group  $G$  is *residually finite* (or *RF*) if for every  $g \in G \setminus \{1\}$ , there is some finite group  $F$  and an epimorphism  $\phi: G \rightarrow F$  so that  $\phi(g) \neq 1$ . In more sophisticated language  $G$  is RF if and only if the trivial subgroup is closed in the profinite topology on  $G$ .

If  $H < G$ , then  $H$  is *separable* if for every  $g \in G \setminus H$ , there is some finite group  $F$  and an epimorphism  $\phi: G \rightarrow F$  so that  $\phi(g) \notin \phi(H)$ . Equivalently, the subgroup  $H$  is separable in  $G$  if it is closed in the profinite topology on  $G$ .

If every finitely generated subgroup of  $G$  is separable,  $G$  is said to be *LERF*, or *subgroup separable*. If  $G$  is hyperbolic, and every quasi-convex subgroup of  $G$  is separable, we say that  $G$  is *QCERF*.

In this paper, we show that if every hyperbolic group is RF, then every hyperbolic group is QCERF.

**Theorem 0.1** *If all hyperbolic groups are residually finite, then every quasi-convex subgroup of a hyperbolic group is separable.*

**Remark 0.2** Theorem 0.1 states that the existence of a non-residually finite hyperbolic group is equivalent to the existence of a non-separable quasi-convex subgroup of some hyperbolic group. This equivalence was guessed by Dani Wise in [20]. Wise (*op. cit.*) and Minasyan [17] noticed independently that an argument of Long and Niblo [12] can be used to show that residual finiteness for all hyperbolic groups implies separability of all *almost malnormal* quasi-convex subgroups.

In the other direction, Kapovich and Wise show in [10] that if every hyperbolic group has a finite index subgroup, then every hyperbolic group is residually finite. Together

with our result, this gives the statement: If every hyperbolic group has a finite index subgroup, then every hyperbolic group is QCERF.

To prove Theorem 0.1, for a hyperbolic group  $G$  with quasi-convex subgroup  $H < G$  and  $g \in G - H$  an element to separate, we would like to find a hyperbolic quotient  $\varphi : G \rightarrow K$ , such that  $\varphi(H) < K$  is finite, and  $\varphi(g) \notin \varphi(H)$ . Then since  $K$  is assumed to be hyperbolic and therefore residually finite, we may separate  $g$  from  $H$ . One natural way to attempt to find such a quotient  $\varphi$  would be to “kill” a large finite-index normal subgroup  $H' \triangleleft H$ , and hope that the quotient of  $G$  is still hyperbolic and that  $H$  projects to  $H/H'$ . This actually works if  $H$  is malnormal in  $G$ . The difficulty with this procedure if  $H$  is not malnormal is that one must make sure that for any  $k \in G - H$ , if  $U = H \cap H^k \neq \emptyset$ , then  $H' \cap U = H' \cap U^{k^{-1}} \subset H$ , otherwise killing  $H'$  would force a larger subset of  $H$  to be killed. Thus, we need to take into account intersections between  $H$  and its conjugates, which motivates considering the following definitions.

Let  $H^g = gHg^{-1}$ . The following was defined in [6].

**Definition 0.3** Let  $H$  be a subgroup of a group  $G$ . The elements  $\{g_i \mid 1 \leq i \leq n\}$  of  $G$  are said to be *essentially distinct* if  $g_i H \neq g_j H$  for  $i \neq j$ . Conjugates  $\{H^{g_i} \mid 1 \leq i \leq n\}$  of  $H$  by essentially distinct elements are called *essentially distinct conjugates*.

It should be remarked that essentially distinct conjugates may coincide if  $H$  is not equal to its own normalizer.

**Definition 0.4** The *height* of an infinite subgroup  $H < G$  is  $n$  if there exists a collection of  $n$  essentially distinct conjugates of  $H$  such that the intersection of all the elements of the collection is infinite and  $n$  is maximal possible. The height of a finite subgroup is 0.

For example, an infinite, malnormal subgroup has height 1, whereas an infinite normal subgroup has height equal to its index. The most relevant result about height for our purposes is the following theorem of Gitik, Mitra, Rips, and Sageev.

**Theorem 0.5** [6] *A quasi-convex subgroup of a hyperbolic group has finite height.*

The proof of Theorem 0.1 will be by induction on height, using the following theorem, which is the main technical result of this paper.

**Theorem 0.6** *Let  $G$  be a torsion-free residually finite hyperbolic group, let  $H$  be a quasi-convex subgroup of  $G$  of height  $k$ , and let  $g \in G \setminus H$ . There is a quotient  $\eta : G \rightarrow \bar{G}$  so that;*

- (1)  $\bar{G}$  is hyperbolic;
- (2)  $\eta(H)$  is quasi-convex in  $\bar{G}$ ;
- (3)  $\eta(g) \notin \eta(H)$ ; and
- (4) the height of  $\eta(H)$  in  $\bar{G}$  is at most  $k - 1$ .

**Proof of Theorem 0.1 from Theorem 0.6** Let  $G$  be a hyperbolic group and  $H$  a quasi-convex subgroup of  $G$  of height  $k$ . We prove that  $H$  is separable by induction on height.

The base case is when  $H$  has height zero, which means  $H$  is finite. Since  $G$  is residually finite it is straightforward to separate any  $g \in G \setminus H$  from the finite set  $H$ .

Assume that  $k \geq 1$ . We claim that it suffices to prove  $H$  is separable in the special case that  $G$  is torsion-free. Indeed, let  $G_0 \leq G$  be a torsion-free subgroup of finite-index. Such a  $G_0$  exists because  $G$  is residually-finite and  $G$  has only finitely many conjugacy classes of torsion elements (see, e.g. [4]). Further, let  $H_0 = G_0 \cap H$ . An elementary argument shows that the height of  $H_0$  is at most  $k$ . Equally, if  $H_0$  is separable in  $G_0$  then  $H$  is separable in  $G$ . To see this, note that since  $G_0$  is of finite-index in  $G$ , the profinite topology on  $G_0$  coincides with the subspace topology induced by the profinite topology on  $G$ . Thus, if  $H_0$  is closed in the profinite topology on  $G_0$  then it is closed in the profinite topology on  $G$ . The subgroup  $H$  is a finite union of cosets of  $H_0$ , and is therefore closed in the profinite topology on  $G$ .

We have now reduced to the case that  $G$  is torsion-free. Let  $g \in G \setminus H$ . By Theorem 0.6 there is a hyperbolic quotient  $\bar{G}$  of  $G$  which separates  $g$  from  $H$ , and the image of  $H$  in  $\bar{G}$  is quasi-convex and has height at most  $k - 1$ . Theorem 0.1 follows by induction.  $\square$

The remainder of the paper is devoted to the proof of Theorem 0.6.

## 0.1 Outline

In Section 1 we recall some definitions from [8], on the geometry of relative hyperbolicity. In Section 2, we collect some results about relatively hyperbolic Dehn filling. In Section 3, we study the peripheral structure on a hyperbolic group induced by a quasi-convex subgroup. The proof of Theorem 0.6 occupies Section 4; we first prove a pair of technical lemmas, and then prove the various conclusions of 0.6 hold for appropriately chosen Dehn fillings of the peripheral structure induced by a quasi-convex subgroup. The pieces are put together in Subsection 4.5. Finally we discuss some generalizations in Section 5.

## 0.2 Acknowledgments

Agol thanks Kevin Whyte for a useful conversation. We would also like to thank Igor Belegradek for corrections to an earlier version of the conclusion, and Eduardo Martínez-Pedroza, who pointed out a serious error in an earlier version of the proof of Lemma 4.2.

Agol was partially supported by NSF grant DMS-0504975, and the Guggenheim foundation. Groves was supported by NSF grant DMS 0504251.

# 1 The cusped space of a relatively hyperbolic group

In this section we briefly recall the main constructions of [8]. Briefly, given a finitely generated group  $G = \langle S \rangle$  and a finite collection of finitely generated subgroups  $\mathcal{P}$ , we build a “cusped space”  $X(G, \mathcal{P}, S)$  by first forming the Cayley graph of  $G$  and then gluing a “horoball” onto each translate of an element of  $\mathcal{P}$ .

**Definition 1.1** Let  $\Gamma$  be any 1–complex. The *combinatorial horoball based on  $\Gamma$* , denoted  $\mathcal{H}(\Gamma)$ , is the 2–complex formed as follows:

- $\mathcal{H}^{(0)} = \Gamma^{(0)} \times (\{0\} \cup \mathbb{N})$
- $\mathcal{H}^{(1)}$  contains the following three types of edges. The first two types are called *horizontal*, and the last type is called *vertical*.
  - (1) If  $e$  is an edge of  $\Gamma$  joining  $v$  to  $w$  then there is a corresponding edge  $\bar{e}$  connecting  $(v, 0)$  to  $(w, 0)$ .
  - (2) If  $k > 0$  and  $0 < d_\Gamma(v, w) \leq 2^k$ , then there is a single edge connecting  $(v, k)$  to  $(w, k)$ .
  - (3) If  $k \geq 0$  and  $v \in \Gamma^{(0)}$ , there is an edge joining  $(v, k)$  to  $(v, k + 1)$ .
- $\mathcal{H}^{(2)}$  contains 2–cells (described explicitly in [8, Definition 3.1]) which ensure that  $\mathcal{H}$  satisfies a linear isoperimetric inequality, with constant independent of  $\Gamma$ .

**Remark 1.2** As the full subgraph of  $\mathcal{H}(\Gamma)$  on the vertices  $\Gamma^{(0)} \times \{0\}$  is isomorphic to  $\Gamma$ , we may think of  $\Gamma$  as a subset of  $\mathcal{H}(\Gamma)$ .

**Definition 1.3** Let  $\Gamma$  be a graph and  $\mathcal{H}(\Gamma)$  the associated combinatorial horoball. Define a *depth* function

$$D : \mathcal{H}(\Gamma) \rightarrow [0, \infty)$$

which satisfies:

- (1)  $D(x) = 0$  if  $x \in \Gamma$ ,
- (2)  $D(x) = k$  if  $x$  is a vertex  $(v, k)$ , and
- (3)  $D$  restricts to an affine function on each 1–cell and on each 2–cell.

**Definition 1.4** [Osin] Suppose that  $G$  is generated by  $S$  with respect to  $\{H_\lambda\}_{\lambda \in \Lambda}$ . This means  $G$  is a quotient of

$$F = F(S) * (*_{\lambda \in \Lambda} H_\lambda),$$

where  $F(S)$  is the free group on the alphabet  $S$ . Suppose that  $N$  is the kernel of the canonical quotient map from  $F$  to  $G$ . If  $N$  is the normal closure of the set  $\mathcal{R}$  then we say that

$$\langle S, \{H_\lambda\}_{\lambda \in \Lambda} \mid \mathcal{R} \rangle,$$

is a *relative presentation* for  $G$  with respect to  $\{H_\lambda\}_{\lambda \in \Lambda}$ .

We say that  $G$  is *finitely presented relative to*  $\{H_\lambda\}_{\lambda \in \Lambda}$  if we can choose  $\mathcal{R}$  and  $S$  to be finite.

**Definition 1.5** *The cusped space*  $X(G, \mathcal{P}, S)$ . Let  $G$  be a finitely generated group which is finitely presented relative to  $\mathcal{P} = \{P_1, \dots, P_m\}$ , a family of finitely generated subgroups of  $G$ . Let  $S$  be a generating set for  $G$  so that  $P_i \cap S$  generates  $P_i$  for each  $i \in \{1, \dots, m\}$ . For each  $i \in \{1, \dots, m\}$ , let  $T_i$  be a left transversal for  $P_i$  (i.e. a collection of representatives for left cosets of  $P_i$  in  $G$  which contains exactly one element of each left coset). Let  $\Gamma = \Gamma(G, S)$  be the Cayley graph of  $G$ . To  $\Gamma$ , equivariantly attach 2–cells coming from the finite relative presentation to obtain a 2–complex  $\bar{\Gamma}$ .

For each  $i$ , and each  $t \in T_i$ , let  $\Gamma_{i,t}$  be the full subgraph of the Cayley graph  $\Gamma(G, S)$  which contains  $tP_i$ . Each  $\Gamma_{i,t}$  is isomorphic to the Cayley graph of  $P_i$  with respect to the generators  $P_i \cap S$ . Then we define

$$X(G, \mathcal{P}, S) = \bar{\Gamma} \cup (\cup \{\mathcal{H}(\Gamma_{i,t}) \mid 1 \leq i \leq m, t \in T_i\}),$$

where the graphs  $\Gamma_{i,t} \subset \Gamma(G, S)$  and  $\Gamma_{i,t} \subset \mathcal{H}(\Gamma_{i,t})$  are identified as suggested in Remark 1.2.

**Definition 1.6** A horoball of  $X(G, \mathcal{P}, S)$  is the subgraph  $\mathcal{H}(\Gamma_{i,t})$  for some  $i$  and  $t$ . For  $l \in \mathbb{N}$ , an  $l$ -horoball is the subgraph of  $\mathcal{H}(\Gamma_{i,t})$  whose vertices are all of distance at least  $l$  from the Cayley graph  $\Gamma$ .

**Remark 1.7** Once a horoball is specified, the vertex of  $X(G, \mathcal{P}, S)$  connected by a vertical geodesic of length  $n$  to the group element  $g$  can be conveniently referred to by the ordered pair  $(g, n)$ , and we will often do so.

**Remark 1.8** Whenever  $X(G, \mathcal{P}, S)$  is to be thought of as a metric space, we will always implicitly ignore the 2-cells, and regard  $\mathcal{H}(\Gamma)^{(1)}$  as a metric graph with all edges of length one.

Relative hyperbolicity was first defined by Gromov in [7]. We use the following characterization (See [8, Section 3] for this characterization and others):

**Proposition 1.9** Let  $G$  be a finitely generated group, and let  $\mathcal{P}$  be a finite collection of finitely generated subgroups. The following are equivalent:

- (1)  $G$  is hyperbolic relative to  $\mathcal{P}$  in the sense of Gromov.
- (2) The space  $X(G, \mathcal{P}, S)$  is Gromov hyperbolic for some finite generating set  $S$ .
- (3) The space  $X(G, \mathcal{P}, S)$  satisfies a linear isoperimetric inequality.

Most of our geometric arguments therefore take place in some cusped space  $X = X(G, \mathcal{P}, S)$ . For most of the paper, we will work either with arbitrary geodesics in this space, or with *regular* geodesics, i.e. geodesics whose intersection with any horoball is vertical except possibly for a single horizontal sub-segment. In Subsection 4.4, we will need to use paths between points in  $X$  (and sometimes  $\partial X$ ) whose behavior is even more controlled. These are the *preferred paths* of [8], and we refer to that paper for a detailed discussion.

## 2 Filling hyperbolic and relatively hyperbolic groups

Let  $G$  be hyperbolic relative to a finite collection  $\mathcal{P} = \{P_1, \dots, P_m\}$ , as in the previous section. A *filling* of  $G$  is determined by a choice of subgroups  $N_j \triangleleft P_j$ , called *filling kernels*; we write the quotient after filling as  $G(N_1, \dots, N_m)$ . If  $S$  is a generating set for  $G$  which contains generating sets for each  $P_i$ , then for each  $i$  we define the *algebraic slope length*, denoted  $|N_i|_{P_i}$ , to be the length of the shortest nontrivial element of  $N_i$ , measured in the generators  $S \cap P_i$ .

We collect here some results about filling from [8] (see also [19]):

**Theorem 2.1** *Let  $G$  be a torsion-free group, which is hyperbolic relative to a collection  $\mathcal{P} = \{P_1, \dots, P_m\}$  of finitely generated subgroups. Suppose that  $S$  is a generating set for  $G$  so that for each  $1 \leq i \leq m$  we have  $P_i = \langle P_i \cap S \rangle$ . Let  $F \subseteq G$  be a finite set.*

*There exists a constant  $B$  depending only on  $G, \mathcal{P}, S,$  and  $F$  so that for any collection  $\{N_i\}_{i=1}^m$  of subgroups satisfying*

- $N_i \trianglelefteq P_i$  and
- $|N_i|_{P_i} \geq B,$

*then the following hold, where  $K$  is the normal closure in  $G$  of  $N_1 \cup \dots \cup N_m,$  and  $G(N_1, \dots, N_m) = G/K:$*

- (1) [8, Theorem 9.1] *The map  $P_i/N_i \xrightarrow{\iota_i} G(N_1, \dots, N_m)$  given by  $pN_i \mapsto pK$  is injective for each  $i.$*
- (2) [8, Theorem 7.2]  *$G(N_1, \dots, N_m)$  is hyperbolic relative to the collection  $\mathcal{Q} = \{\iota_i(P_i/N_i) \mid 1 \leq i \leq m\}.$*
- (3) [8, Corollary 9.7] *The projection from  $G$  to  $G(N_1, \dots, N_m)$  is injective on  $F.$*

The following lemma is needed in the proof of Proposition 2.3. Its statement involves both ‘relative Dehn functions’ and the ‘coned-off Cayley complex’ of a relatively hyperbolic group. We refer the reader to [18, Section 2.3] and [8, Definition 2.47] for the definitions.

**Lemma 2.2** *Suppose that  $G$  is hyperbolic relative to  $\mathcal{P}$  and that*

$$\langle X, \{P_\lambda\}_{\lambda \in \Lambda} \mid \mathcal{R} \rangle,$$

*is a finite relative presentation for  $G$  with respect to  $\mathcal{P}.$  Let  $M = \max_{r \in \mathcal{R}} |r|.$*

*Suppose further that that  $G$  has a linear relative Dehn function with constant  $K.$  Then the coned-off Cayley complex of  $G$  with respect to  $\mathcal{P}$  has a linear isoperimetric function with constant at most  $(M + 1)K + 1.$*

**Proof** Let  $\hat{C}$  be the coned-off Cayley complex. Start with a loop  $c$  in the 1–skeleton of  $\hat{C}.$  We may clearly assume that  $c$  is embedded.

Let  $\Gamma$  be the Cayley graph of  $G$  with respect to the generating set  $X \cup (\cup_\lambda P_\lambda \setminus \{1\}).$  Any loop in  $\Gamma$  can be filled with a disk whose 2–cells have boundary labelled either by elements of  $\mathcal{R}$  (‘ $\mathcal{R}$ –cells’) or by a relation in one of the  $P_\lambda$  (‘ $\mathcal{P}$ –cells’).

We replace the loop  $c$  by a loop  $c'$  in  $\Gamma$  by taking each sub-segment of  $c$  of length 2 which has a cone point as its midpoint and replacing it with the corresponding edge of  $\Gamma$ .

Clearly  $|c'| \leq |c|$ . There is therefore some filling of  $c'$  in  $\Gamma$  with at most  $K|c|$   $\mathcal{R}$ -cells (and we do not need any information about the number of  $\mathcal{P}$ -cells).

These  $\mathcal{R}$ -cells lift to a partial filling  $\omega$  of  $c$  in  $\hat{C}$ . There is a collection  $\nu_1, \dots, \nu_k$  of embedded loops, each of which is in the closed star of some cone vertex  $x_i \in \hat{C}$  so that as (oriented) 1-cycles, the boundary of  $\omega$  is  $c - \sum_i \nu_i$ .

The edges in the  $\nu_i$  are of three types:

- (1) edges on the boundary of some  $\mathcal{R}$ -cell
- (2) edges in the Cayley graph whose interior do not intersect any  $\mathcal{R}$ -cell; and
- (3) the cone edges removed when constructing  $c'$  from  $c$ .

There are at most  $MK|c|$  edges of the first type, and at most  $|c|$  total edges of the second and third types. Since the cone on any graph has isoperimetric constant 1, each loop  $\nu_i$  can be filled with a disk  $\Delta_i$  of area at most  $|\nu_i|$ . The required filling of  $c$  is given by the 2-chain  $\omega + \sum_i \Delta_i$ . It is straightforward to see that this can be realized by a disk. Since the area of  $\omega$  is at most  $K|c|$ , and the sum of the areas of the  $\Delta_i$  is equal to

$$\sum_i |\nu_i| \leq (MK + 1)|c|,$$

we get the required isoperimetric constant for  $\hat{C}$ . □

**Proposition 2.3** *Suppose that  $G$ ,  $\mathcal{P}$  and  $S$  are as in the hypothesis of Theorem 2.1, and let  $F = \emptyset$ . There is some  $\delta$  and  $B$  so that for any hyperbolic filling  $G' = G(N_1, \dots, N_m)$  and  $\mathcal{Q}$  as in Theorem 2.1 with  $|N_i|_{P_i} \geq B$  for all  $i$  the space  $X(G', \mathcal{Q}, S)$  is  $\delta$ -hyperbolic.*

**Proof** By [8, Theorem 3.25],  $G$  is hyperbolic relative to  $\mathcal{P}$ . By the Appendix of [18] this means that the relative Dehn function of  $G$  with respect to  $\mathcal{P}$  is linear. Let  $C$  be the constant of this linear function. By [19, Lemma 5.3], there is a finite set  $\mathcal{A} \in G$  so that if each  $N_i \cap \mathcal{A} = \emptyset$  then the relative Dehn function for  $G'$  with respect to  $\mathcal{Q}$  is linear with constant at most  $3C$ . Let  $B$  be so large that the ball of radius  $B$  about 1 in  $G$  contains  $\mathcal{A}$ .

Given a finite relative presentation for  $G$ , there is an obvious finite relative presentation for  $G'$ , and the maximum length of a relator does not increase. Let  $M$  be the maximum length of a relator in the given finite relative presentation for  $G$  (which is used to

calculate the constant  $C$  above). By Lemma 2.2, the coned-off Cayley complex for  $G'$  has a linear isoperimetric function with constant at most  $3(M + 1)C + 1$ . Let  $C' = 3(M + 1)C + 1$ . By [8, Theorem 3.24] this implies that the cusped space for  $G'$  has a linear isoperimetric function with constant  $3C'(2C' + 1)$ . Now, by [5, Theorem III.H.2.9], the constant of hyperbolicity for the cusped space for  $G'$  can be calculated explicitly in terms of this isoperimetric constant, and  $\max\{M, 5\}$ , the maximum length of an attaching map of a 2-cell for the cusped space. Putting all of these estimates together shows that this constant of hyperbolicity is uniform over all sufficiently long fillings.  $\square$

**Remark 2.4** There is a direct proof of the above result using the results of [8] rather than [19]. However, the output of the main theorem of [8] is a bound on the constant for a linear *homological* isoperimetric inequality for the space  $X(G', \mathcal{Q}, S)$ . In order to apply this, one needs to make the constant of hyperbolicity in the conclusion of [8, Theorem 2.29] explicit in terms of the homological isoperimetric constant. This would involve rewriting [5, Theorem III.H.2.9] in a homological setting. Feeling that this would be too much of a diversion, we chose the shorter but more circuitous proof above.

### 3 Quasi-convexity

Suppose that  $H$  and  $G$  are both relatively hyperbolic groups. Let  $\mathcal{P} = \{P_1, \dots, P_n\}$  be the peripheral subgroups of  $G$ , and let  $\mathcal{D} = \{D_1, \dots, D_m\}$  be the peripheral subgroups of  $H$ . Let  $\phi: H \rightarrow G$  be a homomorphism. If every  $\phi(D_i) \in \mathcal{D}$  is conjugate in  $G$  into some  $P_j \in \mathcal{P}$ , we say that the map  $\phi$  respects the peripheral structure on  $H$ . Let  $S$  and  $T$  be finite relative generating sets for  $G$  and  $H$  respectively.

**Lemma 3.1** *If  $\phi: H \rightarrow G$  is a homomorphism which respects the peripheral structure on  $H$ , then  $\phi$  extends to an  $H$ -equivariant lipschitz map  $\check{\phi}$  from (the zero-skeleton of)  $X(H, \mathcal{D}, T)$  to  $X(G, \mathcal{P}, S)$ . If  $\phi$  is injective, then  $\check{\phi}$  is proper.*

**Proof** We first associate with each  $D_i \in \mathcal{D}$  an element  $c_i \in G$ . Since  $\phi$  respects the peripheral structure, there is some  $P_{j_i} \in \mathcal{P}$  and some  $c \in G$  so that  $\phi(D_i) \subseteq cP_{j_i}c^{-1}$ . We let  $c_i$  be some shortest such  $c$ , with respect to the generators  $S$ .

For  $h \in H$ , we define  $\check{\phi}(h) = \phi(h)$ . A vertex in a horoball of  $X(H, \mathcal{D}, T)$  is determined by a triple  $(sD_i, h, n)$ , where  $s \in H$ ,  $D_i \in \mathcal{D}$ ,  $h \in sD_i$ , and  $n \in \mathbb{N}$ . We define

$$\check{\phi}(sD_i, h, n) = (\phi(s)c_iP_{j_i}, \phi(h)c_i, n).$$

Let  $a = \max\{|\phi(t)|_S \mid t \in T\}$ , and let  $b = \max\{|c_i|_S\}$ ; the map  $\check{\phi}$  is  $\alpha$ -lipschitz for  $\alpha = \max\{a, b + 1\}$ .

Properness is left to the reader.  $\square$

Recall that a *filling* of  $G$  is determined by a choice of subgroups  $N_j \triangleleft P_j$ , called *filling kernels*; we write the quotient after filling as  $G(N_1, \dots, N_m)$ .

**Definition 3.2** If  $\phi$  is a homomorphism which respects the peripheral structure on  $H$ , then any filling of  $G$  induces a filling of  $H$  as follows. For each  $i$ , there is some  $c_i = c(D_i)$  in  $G$  and some  $P_{j_i}$  in  $\mathcal{P}$  so that  $c_i P_{j_i} c_i^{-1}$  contains  $\phi(D_i)$ . The induced filling kernels  $K_i \triangleleft D_i$  are given by

$$K_i = \phi^{-1}(c_i N_{j_i} c_i^{-1}) \cap D_i.$$

The induced filling is  $H(K_1, \dots, K_n)$ . The map  $\phi$  induces a homomorphism

$$\check{\phi}: H(K_1, \dots, K_n) \rightarrow G(N_1, \dots, N_m).$$

**Definition 3.3** Suppose  $G$  is a relatively hyperbolic group, relative to  $\mathcal{P}$ , and that  $H < G$  is hyperbolic relative to  $\mathcal{D}$  and that the inclusion of  $H$  into  $G$  respects the peripheral structure. A filling  $G \rightarrow G(N_1, \dots, N_m)$  is an  $H$ -filling if whenever  $H \cap P_i^g$  is nontrivial,  $N_i^g \subseteq s D_j s^{-1} \subseteq H$  for some  $s \in H$ , and  $D_j \in \mathcal{D}$ .

### 3.1 Induced peripheral structures

Let  $G$  be a hyperbolic group, and let  $H < G$  be a quasi-convex subgroup. Recall that according to Theorem 0.5,  $H$  has finite *height* (see Definition 0.4). We will construct a peripheral structure for  $G$  using the infinite intersections of maximal collections of essentially distinct conjugates of  $H$ .

**Lemma 3.4** *There are only finitely many  $H$ -conjugacy classes of subgroups  $H \cap H^g$  such that  $|H \cap H^g| = \infty$ .*

**Proof** If  $g_1$  and  $g_2$  lie in the same double coset  $HgH$ , then the subgroups  $H \cap H^{g_1}$  and  $H \cap H^{g_2}$  are conjugate in  $H$ . By [6, Lemma 1.2], there is an upper bound on the minimal length of a coset representative of  $HgH$  such that  $|H \cap H^g| = \infty$ .  $\square$

Using induction on the height we obtain the following.

**Corollary 3.5** *There are only finitely many  $H$ -conjugacy classes of intersections  $H \cap H^{g_2} \cap \cdots \cap H^{g_j}$ , where  $j \leq n$ , with  $n$  the height of  $H$  in  $G$  and  $\{1, g_2, \dots, g_j\}$  are essentially distinct.*

Choosing one subgroup of this form per  $H$ -conjugacy class and taking its commensurator in  $H$ , we obtain a system  $\mathcal{D}$  of (quasi-convex) subgroups of  $H$  which we will call the *malnormal core of  $H$* . The collection  $\mathcal{D}$  gives rise to a collection of peripheral subgroups  $\mathcal{P}$  for  $G$  in two steps:

- (1) Change  $\mathcal{D}$  to  $\mathcal{D}'$  by replacing each element of  $\mathcal{D}$  by its commensurator in  $G$ .
- (2) Eliminate redundant entries of  $\mathcal{D}'$  to obtain  $\mathcal{P} \subseteq \mathcal{D}'$  which contains no two elements which are conjugate in  $G$ .

Call  $\mathcal{P}$  the peripheral structure on  $G$  induced by  $H$ . This peripheral structure is only well-defined up to replacement of some elements of  $\mathcal{P}$  by conjugates. On the other hand, replacing  $H$  by a commensurable subgroup of  $G$  does not affect the induced peripheral structure. We consider two peripheral structures on a group to be the same if the same group elements are parabolic in the two structures.

**Observation 3.6** *Let  $H_1$  and  $H_2$  be quasi-convex subgroups of the hyperbolic group  $G$  with the same limit sets in  $\partial G$ . The peripheral structures induced by  $H_1$  and  $H_2$  are the same.*

In the next two observations and lemma, we consider a hyperbolic group  $G$  and a quasi-convex subgroup  $H$ . We let  $\mathcal{D}$  be the malnormal core of  $H$ , and let  $\mathcal{P}$  be the peripheral structure on  $G$  induced by  $H$ . Finally,

$$\check{\iota}: X(H, \mathcal{D}, T) \rightarrow X(G, \mathcal{P}, S)$$

is the extension of the inclusion map given by Lemma 3.1.

**Observation 3.7** *If  $\mathcal{P}$  is the peripheral structure on  $G$  induced by  $H$ , and  $P \in \mathcal{P}$ , then  $H \cap P$  is finite index in  $P$ .*

**Observation 3.8** *If  $\mathcal{P}$  is the peripheral structure on  $G$  induced by  $H$ , and  $h \in H$  is parabolic with respect to that structure, then  $h$  is conjugate in  $H$  to an element of  $D$  for some  $D$  in the malnormal core  $\mathcal{D}$  of  $H$ .*

**Lemma 3.9** *There is a constant  $\beta$  satisfying the following: Let  $A$  be a horoball of  $X(G, \mathcal{P}, S)$ , and suppose that  $H \cap \text{Stab}_G(A)$  contains an element of infinite order. A  $\beta$ -neighborhood of the image of  $\check{\iota}$  contains  $A$ .*

**Proof** For each  $D \in \mathcal{D}$ , there is some (unique)  $P = P(D) \in \mathcal{P}$  and some  $c = c(D) \in G$  (chosen as in the proof of Lemma 3.1) so that  $D < cPc^{-1}$ . By Observation 3.7,  $D$  is finite index in  $cPc^{-1}$ . Since  $\mathcal{D}$  is finite, there is some constant  $\beta_1$ , independent of  $D$ , so that  $cP$  is contained in a  $\beta_1$ -neighborhood of  $D$  in  $G$ .

Let  $h$  be the infinite order element of  $H \cap \text{Stab}_G(A)$ . Observation 3.8 implies that  $h$  is already parabolic in  $H$ , so  $h \in sDs^{-1}$  for some  $s \in H$  and  $D \in \mathcal{D}$ .

The horoball  $A$  is attached to some coset  $tP_i$  for  $t \in G$  and  $P_i \in \mathcal{P}$ . Since parabolics cannot have infinite intersection without coinciding, it follows that  $P_i = P(D)$ ; we may take  $t = sc$ .

It follows from the first paragraph that  $tP_i = scP_i$  is contained in a  $\beta_1$ -neighborhood of  $sD$ . Moreover, elements of  $tP_i$  are uniformly close to elements of  $sDc$ , and elements of the horoball  $A$  are uniformly close to elements of the form

$$\check{y}(sD, h, n) = (tP_i, hc, n).$$

In other words, the vertices of  $A$  which do not lie in  $G$  are all contained in some  $\beta_2$ -neighborhood of the image of  $\check{y}$ .

We may therefore take  $\beta = \max\{\beta_1, \beta_2\}$ . □

**Example 3.10** Let  $G = \langle a, b \rangle$  be a free group of rank 2 and let  $H = \langle a^2, ba^3b^{-1} \rangle$ . In this case one must take commensurators twice, once to get the malnormal core and a second time to get the induced peripheral structure. Indeed, the minimal infinite intersections of  $H$  with its conjugates are conjugate in  $H$  either to

$$\langle a^6 \rangle = H \cap H^a \cap H^b \cap H^{ab^{-1}} \cap H^{a^2b^{-1}}$$

or

$$\langle ba^6b^{-1} \rangle = H \cap H^{bab^{-1}} \cap H^{ba^2b^{-1}} \cap H^b \cap H^{ba},$$

so  $H$  has height 5. The malnormal core of  $H$  is

$$\mathcal{D} = \{\langle a^2 \rangle, \langle ba^3b^{-1} \rangle\}$$

and the induced peripheral structure on  $G$  is  $\mathcal{P} = \{\langle a \rangle\}$ .

**Definition 3.11** Let  $\phi: H \rightarrow G$  be a homomorphism which respects the peripheral structure. We say that  $\phi(H)$  is  $C$ -relatively quasi-convex in  $G$  if  $\check{\phi}$  has  $C$ -quasi-convex image. If  $H < G$  and  $\phi$  is the inclusion map, we say that  $H$  is a relatively quasi-convex subgroup of  $G$ .

The relative quasi-convexity of  $\phi(H)$  does not depend on the choice of relative generating sets  $S$  and  $T$ , though the constant  $C$  does depend on  $S$  and  $T$ .

**Proposition 3.12** *Let  $H$  be a quasi-convex subgroup of the torsion-free hyperbolic group  $G$ , and let  $\mathcal{D}$  be the malnormal core of  $H$ . Let  $\mathcal{P}$  be the peripheral structure on  $G$  induced by  $H$ .*

- (1)  $H$  is hyperbolic relative to  $\mathcal{D}$ .
- (2)  $G$  is hyperbolic relative to  $\mathcal{P}$ .
- (3) With respect to the above peripheral structures,  $H$  is a relatively quasi-convex subgroup of  $G$ .

**Proof** The first two assertions are essentially contained in [3, Proposition 7.11]. By construction, the elements of  $\mathcal{D}$  are quasi-convex, non-conjugate, and any pair of conjugates of elements of  $\mathcal{D}$  are either equal or intersect in a finite set. They are also equal to their commensurator, are hence to their normalizer. These are the hypotheses of [3, Proposition 7.11]. The same properties hold for  $\mathcal{P}$  in  $G$ .

We now consider the third property. Let  $X = X(G, \mathcal{P}, S)$  be the cusped space of  $G$ . Let  $X_H$  be the zero-skeleton of the cusped space of  $H$ , and let  $Y$  be the image of the proper map  $\check{\phi}: X_H \rightarrow X$  from Lemma 3.1. Let  $x$  and  $y$  be vertices of  $Y$ .

We need to prove that there is a constant  $C$  (independent of  $x$  and  $y$ ) so that a geodesic in  $X$  between  $x$  and  $y$  lies within  $C$  of  $Y$ .

**Case 1** *The points  $x$  and  $y$  lie deep (deeper than  $50\delta$ ) in the same horoball.*

In this case the geodesic between  $x$  and  $y$  lies entirely in the horoball (see [8, Lemma 3.26]). Any geodesic between  $x$  and  $y$  is Hausdorff distance at most 4 from a regular geodesic, which is vertical except for a horizontal segment of length at most three (see [8, Lemma 3.10]). The vertical sub-segments start at points in  $Y$ , so by construction of cusped spaces and the map  $\check{\phi}$ , the vertical sub-segments lie in  $Y$  also. Therefore in this case we can take  $C = 6$ .

**Case 2**  *$x$  and  $y$  lie at depth no more than  $50\delta$  in  $X$ .*

In this case, consider the space  $X'$  which consists of all vertices in  $X$  at depth at most  $100\delta$ . This space is quasi-isometric to the group  $G$ , and  $H$  is a quasi-convex subset of  $X'$ , with quasi-convexity constant  $\lambda$ , say. Choose a geodesic  $\gamma$  in  $X'$  between  $x$  and  $y$ . We may assume that  $\gamma$  is “regular” in each horoball, in the following sense: If  $\gamma$

contains vertices at depth  $90\delta$  in the horoball, then that part of  $\gamma$  between depth 0 and depth  $90\delta$  consists of two vertical segments.

Since  $H$  is  $\lambda$ -quasi-convex in  $X'$ , there is an element of  $H$  within  $\lambda$  of any point in  $\gamma$ . Now consider  $\gamma$  as a subset of  $X$ , using the natural inclusion of  $X'$  in  $X$ . We will replace  $\gamma$  with a  $10\delta$ -local geodesic  $\bar{\gamma}$  in  $X$  with endpoints  $x$  and  $y$ . The path  $\bar{\gamma}$  will be seen to lie in a uniformly bounded neighborhood of  $Y$ .

Let  $\sigma$  be a sub-segment of  $\gamma$  lying entirely below depth  $90\delta$ . Any such  $\sigma$  is contained in a unique maximal segment  $\hat{\sigma}$  lying below depth  $80\delta$ . To define  $\bar{\gamma}$ , we replace each such  $\hat{\sigma}$  with an  $X$ -geodesic consisting of two vertical and one horizontal sub-segment.

This yields a continuous path  $\bar{\gamma}$  from  $x$  to  $y$  which we claim is a  $10\delta$ -local geodesic in  $X$ . Consider a sub-segment  $I$  of  $\bar{\gamma}$  of length  $10\delta$ . We must show that  $I$  is a geodesic in  $X$ . If  $I$  lies completely beneath depth  $80\delta$  it is obvious that  $I$  is geodesic.

Suppose  $I$  lies entirely above depth  $80\delta$ . Any path joining the endpoints of  $I$  which is not entirely contained in  $X'$  must have length at least  $40\delta$ . Since  $I$  has length  $10\delta$  and is a geodesic in  $X'$ ,  $I$  is a geodesic in  $X$ .

Finally, between depths  $70\delta$  and  $90\delta$ ,  $\bar{\gamma}$  is vertical, and hence geodesic. In particular, if  $I$  crosses depth  $80\delta$ , then  $I$  is geodesic. This shows that  $\bar{\gamma}$  is a  $10\delta$ -local geodesic between  $x$  and  $y$ .

We claim that  $\bar{\gamma}$  lies in a bounded neighborhood of  $Y$ . This is clear for those parts of  $\bar{\gamma}$  which lie in  $\gamma$ , so let  $\tau$  be a maximal sub-segment of  $\bar{\gamma}$  which is not contained in  $\gamma$ . Then  $\tau$  is contained in a single horoball  $A$ . We now split into two subcases, depending on the length of  $\tau$ . Let  $A_0$  be the part of  $A$  at depth 0, and  $A_\lambda$  be the  $\lambda$ -neighborhood of  $A_0$  in  $X$ . Also let  $G_A = \text{Stab}(A)$ . Then  $G_A$  acts cocompactly on  $A_\lambda$ . Let  $K$  be the number of vertices in  $A_\lambda/G_A$ .

**Case 2.1** *The length of  $\tau$  is less than  $(3 + 40\delta + 2 \log_2(\lambda(K + 1)))$ .*

Then each point in  $\tau$  is at most  $(3 + 40\delta + 2 \log_2(\lambda(K + 1)))/2$  from an endpoint of  $\tau$ . However, the endpoints of  $\tau$  lie in  $\gamma$ , which lies in a  $\lambda$ -neighborhood of  $Y$ . Thus in this case each element of  $\tau$  lies within  $(3 + 40\delta + 2 \log_2(\lambda(K + 1)))/2 + \lambda$  of  $Y$ .

**Case 2.2** *The length of  $\tau$  is at least  $(3 + 40\delta + 2 \log_2(\lambda(K + 1)))$ .*

In this case, consider the path  $\tilde{\tau} \subset \gamma$  which joins the endpoints of  $\tau$ . The path  $\tilde{\tau}$  has length at least  $\lambda(K + 1)$ . Each point in  $\tilde{\tau}$  is within at most  $\lambda$  from a point in  $H$ , so there are at least  $K + 1$  distinct points,  $\{h_1, \dots, h_{K+1}\}$ , all within  $\lambda$  of  $\tilde{\tau}$  and each of

these points lies in  $A_\lambda$ . By the choice of  $K$ , there is  $h_i \neq h_j$  in the same  $G_A$ -orbit, so  $h_i h_j^{-1} \in G_A \setminus \{1\}$ . Since  $G$  is torsion-free,  $h_i h_j^{-1}$  has infinite order, and by Lemma 3.9, a  $\beta$ -neighborhood of  $Y$  contains  $A$ .

Let  $\eta$  be a geodesic joining  $x$  to  $y$  in  $X$ . By [5, III.H.1.13(1)],  $\eta$  lies in a  $2\delta$  neighborhood of  $\bar{\gamma}$ , which we have already shown lies in a bounded neighborhood of  $Y$ .

**Case 3** *Either  $x$  or  $y$  lies inside a  $50\delta$ -horoball, but we are not in Case 1.*

If  $x$  or  $y$  lies in a horoball, it lies directly beneath a point in  $H$  at depth 0 (i.e. in the Cayley graph of  $G$ ) in  $X$ . Either appending or deleting<sup>1</sup> the vertical paths from  $x$  to depth 0 and similarly for  $y$  we obtain a path which is a  $10\delta$ -local geodesic. Since  $10\delta$ -local geodesics are  $(\frac{7}{3}, 2\delta)$ -quasi-geodesics [5, III.H.1.13(3)], the proposition follows now from Cases 1 and 2.  $\square$

**Remark 3.13** We direct the interested reader to [16] for a much more general theorem from which part (3) of Proposition 3.12 follows.

In general, even if  $G$  is hyperbolic, a relatively quasi-convex subgroup (with respect to some relatively hyperbolic structure on  $G$ ) need not be quasi-convex in  $G$ . However, the following lemma is straightforward.

**Lemma 3.14** *Suppose that  $G$  is hyperbolic relative to a collection of finite subgroups. Then  $G$  is hyperbolic and any relatively quasi-convex subgroup of  $G$  is quasi-convex.*

## 4 Proof of Theorem 0.6

### 4.1 Projections of geodesics to cusped spaces of quotients

The key technical lemma is the following:

**Lemma 4.1** *Fix a relatively hyperbolic group  $G$  with peripheral subgroups  $\mathcal{P}$  and compatible generating set  $S$ . Choose  $\delta > 0$  so that the cusped spaces  $X = X(G, \mathcal{P}, S)$  and  $X' = X(G(N_1, \dots, N_m), \mathcal{Q}, S)$  are both  $\delta$ -hyperbolic, whenever  $G(N_1, \dots, N_m)$  is a sufficiently long filling of  $G$ .*

---

<sup>1</sup>Whether a vertical path is appended or deleted depends on whether the geodesic from  $x$  to  $y$  initially goes up or down (the case when it goes horizontal is treated as if it goes down).

Let  $L \geq 10\delta$ , and let  $D \geq 3L$ . Let  $F = \{g \in G \mid d_X(g, 1) \leq 2D\}$ , and let  $G \xrightarrow{\pi} G(N_1, \dots, N_m)$  be any hyperbolic filling of  $G$  which is injective on  $F$  and so that  $X' = X(G(N_1, \dots, N_m), \mathcal{Q}, \mathcal{S})$  is  $\delta$ -hyperbolic. (We denote the induced map from  $X$  to  $X'$  also by  $\pi$ .) Let  $\gamma$  be a regular geodesic in  $X$  joining two elements of  $G$ . One of the following occurs:

- (1) There is a  $10\delta$ -local geodesic with the same endpoints as  $\pi(\gamma)$  which is contained in a 2-neighborhood of  $\pi(\gamma)$  and coincides with  $\pi(\gamma)$  everywhere in an  $L$ -neighborhood of the Cayley graph of  $G(N_1, \dots, N_m)$ .
- (2) There is a coset  $tP_i$  whose corresponding horoball intersects  $\gamma$  in a sub-segment  $[g_1, g_2]$  of length at least  $2D - 20\delta - 4$  but there is some  $n \in N_i$  with  $d_X(g_1, g_2n) \leq 2L + 3$ .

**Proof** It is straightforward to verify that the map  $\pi: X \rightarrow X'$  induced by the filling is injective on balls of radius  $10\delta$  centered on points either in the Cayley graph or at depth at most  $D - 10\delta - 2$  in a horoball. It follows that sub-segments of  $\gamma$  straying no further than  $D - 10\delta - 2$  from the Cayley graph project to  $10\delta$ -local geodesics. Let  $B$  be a horoball which  $\gamma$  penetrates to depth greater than  $D - 10\delta - 2$ . The horoball  $B$  intersects the Cayley graph of  $G$  in some coset  $tP$  for  $t \in G$  and  $P_i \in \mathcal{P}$ . Let  $g_1$  and  $g_2$  be the initial and terminal vertices of  $\gamma \cap B$ . There are three possibilities:

- (a)  $\pi(\gamma \cap B)$  is geodesic,
- (b)  $\pi(\gamma \cap B)$  is not geodesic but  $d_{X'}(\pi(g_1), \pi(g_2)) \geq 2L + 3$ ,
- (c)  $\pi(\gamma \cap B)$  is not geodesic and  $d_{X'}(\pi(g_1), \pi(g_2)) < 2L + 3$ .

We first claim that if (a) or (b) holds for every horoball which  $\gamma$  penetrates to depth greater than  $D - 10\delta - 2$ , then conclusion (1) of the Lemma holds. We argue by constructing a new  $10\delta$ -quasi-geodesic  $\gamma'$  in  $X'$  which agrees with  $\pi(\gamma)$  everywhere in an  $L$ -neighborhood of the Cayley graph of  $G(N_1, \dots, N_m)$  and inside those horoballs of  $X'$  which  $\pi(\gamma)$  intersects in a geodesic segment. Whenever a sub-segment  $\sigma$  of  $\pi(\gamma)$  is of type (b), we can replace it by a shorter, but still  $10\delta$ -local geodesic segment as follows. The segment  $\sigma$  is composed of two vertical sub-segments and a short (length 2 or 3) horizontal sub-segment at depth  $d > D - 10\delta - 2$ . Since  $\pi(\sigma)$  is not geodesic, the images in  $X'$  of the vertical sub-segments must come within a horizontal distance of 3 of one another at some smaller depth  $d'$ . The assumption that  $d_{X'}(\pi(g_1), \pi(g_2)) \geq 2L + 3$  forces  $d' > L$ . Modifying  $\sigma$  by removing the part lying below depth  $d'$  and replacing it with a horizontal geodesic leaves a geodesic  $\sigma'$  which still goes to depth at least  $L$ . Therefore making all possible such modifications leaves a  $10\delta$ -local geodesic  $\gamma'$  satisfying conclusion (1) of the Lemma.

Now suppose that there is some horoball  $B$  so that  $\gamma$  penetrates  $B$  to depth greater than  $D - 10\delta - 2$ , but  $\pi(\gamma \cap B)$  satisfies condition (c) above. The image of  $P_i$  in  $G(N_1, \dots, N_m)$  is canonically isomorphic to  $P_i/N_i$ , so there is some  $n \in N_i$  so that  $d_X(g_1, g_2n) \leq 2L + 3$ . Since  $\gamma$  is geodesic,  $d_X(g_1, g_2) > 2(D - 10\delta - 2)$ , and so conclusion (2) holds.  $\square$

In our current applications, we will only ever apply this lemma to a geodesic with both endpoints in a quasi-convex subgroup. In this context, more can be said.

**Lemma 4.2** *Let  $G$ ,  $\mathcal{P}$ ,  $S$ , and  $L \geq 10\delta$  be as in the hypothesis of Lemma 4.1. Let  $H$  be a  $\lambda$ -relatively quasi-convex subgroup of  $G$ , and let  $\alpha$  be the Lipschitz constant for the extension of the inclusion map in Lemma 3.1.*

*Let  $D \geq 3L + 100\lambda + 4\alpha$ , and let  $F = \{g \in G \mid d_X(g, 1) \leq 2D\}$ . Suppose  $G \xrightarrow{\pi} G(N_1, \dots, N_m)$  is an  $H$ -filling which is injective on  $F$  and so that  $X' = X(G(N_1, \dots, N_m), \mathcal{Q}, S)$  is  $\delta$ -hyperbolic. Let  $K_H < \ker(\pi) \cap H$  be the kernel of the induced filling of  $H$  (as in Definition 3.2). Finally, suppose that  $\gamma$  is a geodesic joining 1 to  $h$  for some  $h \in H$ .*

*If conclusion (2) of Lemma 4.1 holds, then there is some  $k \in K_H$  satisfying  $|kh|_X < |h|_X$ .<sup>2</sup>*

**Proof** Let  $g_1, g_2 \in tP_i$  and  $n \in N_i$  be as in the conclusion to Lemma 4.1, and let  $B$  be the horoball in  $X$  which contains  $tP_i$ . We have

- (1)  $d_X(g_1, g_2) \geq 2D - 20\delta - 4 \geq 6L + 200\lambda + 8\alpha - 20\delta - 4$ , but
- (2)  $d_X(g_1, g_2n) < 2L + 3$ .

We use quasi-convexity to approximate  $g_1$  and  $g_2$  by elements of  $Hc \cap tP_i$  for some small  $c \in G$ . If  $g \in tP_i$  and  $m \in \mathbb{N}$ , we will write  $(g, m)$  for the unique vertex of  $B$  connected to  $g$  by a vertical geodesic of length  $m$ , as in Remark 1.7.

By (1), the geodesic  $\gamma$  penetrates the horoball  $B$  to depth at least  $D - 10\delta - 4 > 2\lambda$ ; in particular,  $\gamma$  passes through  $(g_1, \lambda + 1)$  and  $(g_2, \lambda + 1)$ . Let  $\iota: H \rightarrow G$  be the inclusion map, and  $\check{\iota}$  the extension from Lemma 3.1. Since  $H$  is  $\lambda$ -relatively quasi-convex, there are points  $z_1 = \check{\iota}((s_1 D_{j_1}, h_1, n_1))$  and  $z_2 = \check{\iota}((s_2 D_{j_2}, h_2, n_2))$  in  $B$  within  $\lambda$  of  $(g_1, \lambda + 1)$  and  $(g_2, \lambda + 1)$ , respectively. Note that  $d_X(z_j, h_j) \leq 2\lambda + \alpha$  and thus  $d_X(h_j, g_j) \leq 4\lambda + \alpha$  for  $j = 1, 2$ . Thus

$$(3) \quad d_X(h_1, h_2) \geq d_X(g_1, g_2) - (8\lambda + 2\alpha) > 0.$$

---

<sup>2</sup>(writing  $|\cdot|_X$  for  $d_X(\cdot, 1)$ )

In particular  $h_1 \neq h_2$ . Moreover, since  $\check{\iota}$  is  $H$ -equivariant,  $h_2 h_1^{-1} B$  intersects  $B$  in its interior. Hence  $h_2 h_1^{-1}$  fixes the horoball  $B$  and the coset  $tP_i$ . Because the filling kernels  $\{N_1, \dots, N_m\}$  are assumed to determine an  $H$ -filling, we have

$$tN_i t^{-1} \subseteq sD_l s^{-1} \subseteq H$$

for some  $s \in H$  and some  $D_l \in \mathcal{D}$ .

For  $j \in \{1, 2\}$  we have  $g_j = tp_j$  for  $p_j \in P_i$ . As  $N_i \triangleleft P_i$ ,  $g_2 n = tp_2 n = tn' p_2$  for some  $n' \in N_i$ . Let  $k = tn' t^{-1}$ , so that  $g_2 n = k g_2$ .

We claim that  $k \in K_H$ . Indeed,  $K_l = cN_i c^{-1}$  is the induced filling kernel in  $D_l$ , if  $c = c(D_l)$ . Moreover  $tN_i t^{-1} = scN_i c^{-1} s^{-1} = sK_l s^{-1}$  lies in  $K_H$ . In particular  $k \in sK_l s^{-1} \subset K_H$ .

Let  $h' = kh$ . It remains to show that  $|h'|_X < |h|_X$ .

Note first that  $d_X(h'_2, g_2 n) = d_X(h_2, g_2)$ . It follows that

$$(4) \quad |h_1^{-1} h'_2|_X < 8\lambda + 2\alpha + 2L + 3.$$

Clearly  $h = h_1(h_1^{-1} h_2)(h_2^{-1} h)$ . Furthermore, each of  $h_1$  and  $h_2$  lies no more than  $4\lambda + \alpha$  from a geodesic connecting 1 to  $h$ . Thus

$$(5) \quad |h_1|_X + |h_1^{-1} h_2|_X + |h_2^{-1} h|_X \leq |h|_X + 16\lambda + 4\alpha.$$

We can factorize  $h'$  as  $h' = h_1(h_1^{-1} h'_2)(h_2^{-1} h)$ . It follows that

$$(6) \quad \begin{aligned} |h'|_X &\leq |h_1|_X + |h_1^{-1} h'_2|_X + |h_2^{-1} h|_X \\ &< |h|_X + 16\lambda + 4\alpha - (|h_1^{-1} h_2|_X - |h_1^{-1} h'_2|_X), \end{aligned}$$

by (5).

Inequalities (3), (4), and (1) imply that

$$\begin{aligned} |h_1^{-1} h_2|_X - |h_1^{-1} h'_2|_X &> |g_1^{-1} g_2|_X - (8\lambda + 2\alpha) - (8\lambda + 2\alpha + 2L + 3) \\ &= |g_1^{-1} g_2|_X - (16\lambda + 4\alpha + 2L + 3) \\ &\geq 200\lambda + 2L + 20\delta - 4 - (16\lambda + 4\alpha + 2L + 3) \\ &> 16\lambda. \end{aligned}$$

Applied to (6), this yields  $|h'|_X < |h|_X$ , as required.  $\square$

## 4.2 The image of $H$ is quasi-convex

**Proposition 4.3** *Let  $G$  be a relatively hyperbolic group and  $H$  a subgroup as in Definition 3.2, and suppose that  $H$  is  $\lambda$ -relatively quasi-convex in  $G$ . For all sufficiently long  $H$ -fillings  $G(N_1, \dots, N_m)$  of  $G$ , the image in  $G(N_1, \dots, N_m)$  of the induced filling  $H(K_1, \dots, K_n)$  is  $\lambda'$ -relatively quasi-convex, for some  $\lambda'$  independent of the filling.*

**Proof** We fix a (relative, compatible) generating set  $S$  for  $G$ , and let  $\delta$  be the uniform constant of hyperbolicity for cusped spaces  $X = X(G, \mathcal{P}, S)$  and  $X' = X(G(N_1, \dots, N_m), \mathcal{Q}, S)$  provided by Proposition 2.3. It is useful to assume that both  $\delta$  and  $\lambda$  are integers.

We will apply Lemmas 4.1 and 4.2 with  $L = 10\delta$  and  $D = 100\lambda + 100\delta$ . “Sufficiently long” then means that the filling is injective on  $F = \{g \in G \mid d_X(g, 1) \leq 2D\}$  (and that  $X'$  is  $\delta$ -hyperbolic).

By [5, III.H.1.13], any  $10\delta$ -local geodesic in  $X'$  is a  $(7/3, 2\delta)$ -quasi-geodesic. Let  $R$  be the constant of quasi-geodesic stability for  $(7/3, 2\delta)$ -quasi-geodesics in a  $\delta$ -hyperbolic space (see [5, III.H.1.7]). We show that it is sufficient to take  $\lambda' = \lambda + R + 2\delta + 2$ .

Let  $\iota: H \rightarrow G$  be inclusion and  $\phi: H \rightarrow G(N_1, \dots, N_m)$  be  $\pi \circ \iota$  where  $\pi: G \rightarrow G(N_1, \dots, N_m)$  is the filling map. Recall from Lemma 3.1 that we have induced maps  $\check{\iota}$  and  $\check{\phi}$  from the cusped space for  $H$  to  $X$  and  $X'$ , respectively.

**Claim 4.3.1** *Let  $\bar{h} \in \pi(H)$ . Any geodesic in  $X'$  joining 1 to  $\bar{h}$  stays in an  $(\lambda + R + 2)$ -neighborhood of the image of  $\check{\phi}$ .*

**Proof** (Claim 4.3.1) We choose  $h \in H$  of minimal  $X$ -length projecting to  $\bar{h}$ , and let  $\gamma$  be a regular geodesic joining 1 to  $h$  in  $X$ .

By Lemmas 4.1 and 4.2, and the minimality of  $h$ , there is a  $10\delta$ -local geodesic with endpoints  $1 = \pi(1)$  and  $\bar{h} = \pi(h)$  which is contained in a 2-neighborhood of  $\pi(\gamma)$ .

Any geodesic from 1 to  $\pi(h)$  therefore lies in a  $(R + 2)$ -neighborhood of  $\pi(\gamma)$ , by quasi-geodesic stability. Since  $\gamma$  lies in a  $\lambda$ -neighborhood of the image of  $\check{\iota}$ , any geodesic from 1 to  $\pi(h)$  lies in a  $(\lambda + R + 2)$ -neighborhood of the image of  $\check{\phi}$ .  $\square$

Claim 4.3.1 suffices to prove the Proposition, as follows: Let  $X_H$  be the zero-skeleton of the cusped space of  $H$ , and let  $Y = \check{\phi}(X_H)$ . Let  $x_1, x_2$  be elements of  $Y$ .

If  $x_1$  and  $x_2$  lie in the same horoball one may use the convexity of  $\delta$ -horoballs ([8, Lemma 3.26]) to see that any geodesic joining them stays in a  $(2\delta + 2)$ -neighborhood of  $Y$ .

Suppose therefore that  $x_1$  and  $x_2$  lie in different horoballs. Each  $x_i$  is connected by a vertical geodesic to some  $h_i c$  for  $h_i \in \phi(H)$  and  $|c|_X < \alpha$ , where  $\alpha$  is the lipschitz constant from Lemma 3.1. Except for  $h_i c$  itself, this vertical geodesic contains only vertices of  $Y$ . The geodesic between  $h_1$  and  $h_2$  is a  $\phi(H)$ -translate of one between 1 and  $h_1^{-1}h_2$ , and so this geodesic stays in a  $(\lambda + R + 2)$ -neighborhood of  $Y$  by Claim 4.3.1. The two vertical segments, the geodesic between  $h_1$  and  $h_2$ , and the geodesics from  $h_1$  to  $h_1 c$  and from  $h_2$  to  $h_2 c$  form five sides of a geodesic hexagon, the sixth side of which can be taken to be any geodesic joining  $x_1$  to  $x_2$ . This sixth side stays within a  $4\delta$ -neighborhood of the other five, and therefore within  $\lambda + R + 2 + 4\delta + \alpha$  of  $Y$ .

The Proposition is proved, for  $\lambda' = \lambda + R + 2 + 4\delta + \alpha$ .  $\square$

The following result is not required for the proof of Theorem 0.6, but may be of independent interest.

**Proposition 4.4** *Let  $H < G$  be a relatively quasi-convex subgroup. For any sufficiently large  $H$ -filling  $G(N_1, \dots, N_m)$  of  $G$ , the induced map from the induced filling  $H(K_1, \dots, K_n)$  into  $G(N_1, \dots, N_m)$  is injective.*

**Proof** As above, choose a compatible generating set  $S$  for  $G$  with peripheral structure  $\mathcal{P} = \{P_1, \dots, P_m\}$  and let  $\delta$  be a constant of hyperbolicity which suffices both for  $X(G, \mathcal{P}, S)$  and for the cusped space  $X'$  of any sufficiently long hyperbolic filling of  $G$ . Let  $\lambda$  be the constant of (relative) quasi-convexity for  $H$ .

Once again, we will apply Lemmas 4.1 and 4.2 with  $L = 10\delta$  and  $D = 100\lambda + 100\delta$ . ‘‘Sufficiently long’’ again means that the filling is injective on  $F = \{g \in G \mid d_X(g, 1) \leq 2D\}$  (and that  $X'$  is  $\delta$ -hyperbolic).

Let  $\pi: G \rightarrow G(N_1, \dots, N_m)$  be such a filling. Let  $h \in \ker(\pi) \cap H$  be nontrivial. We must show that  $h \in K_H$ , the kernel of the induced filling of  $H$ . Let  $\gamma$  be a geodesic in  $X$  from 1 to  $h$ . Note that  $\pi(\gamma)$  is a loop. Suppose that conclusion (1) of Lemma 4.1 holds. Then there is a  $10\delta$ -local geodesic loop based at 1 in  $X'$ , which coincides with  $\pi(\gamma)$  on an initial segment of length  $L \geq 10\delta$ . This is impossible since there are no nontrivial  $10\delta$ -local geodesic loops in a  $\delta$ -hyperbolic space.

Therefore we may apply Lemma 4.2 to conclude that there is a  $k \in K_H$  so that  $|kh|_X < |h|_X$ . Induction on the length of  $h$  shows that  $h \in K_H$ , as required.  $\square$

### 4.3 Keeping $g$ out of $H$

**Proposition 4.5** *Let  $H < G$  be a relatively quasi-convex subgroup, and let  $I > 0$ . There is some  $F = F(I)$  so that if  $G(N_1, \dots, N_m)$  is an  $H$ -filling of  $G$  which is injective on  $F$ , and  $g \in G \setminus H$  satisfies  $|g|_X < I$ , then  $\pi(g) \notin \pi(H)$ .*

**Proof** As above, choose a compatible generating set  $S$  for  $G$  with peripheral structure  $\mathcal{P} = \{P_1, \dots, P_m\}$  and let  $\delta$  be a constant of hyperbolicity which suffices both for  $X(G, \mathcal{P}, S)$  and for the cusped space of any sufficiently long hyperbolic filling of  $G$ . Let  $\lambda$  be the constant of (relative) quasi-convexity for  $H$ .

As usual, we will apply Lemmas 4.1 and 4.2. This time, we will choose  $L = 2I + 10\delta$  and  $D = 100\lambda + 100\delta + 6I$ . Let

$$F = \{g \in G \mid |g|_X \leq 2D\},$$

and let  $G(N_1, \dots, N_m)$  be an  $H$ -filling of  $G$  which is injective on  $F$  and so that the associated cusped space  $X'$  is  $\delta$ -hyperbolic.

If the proposition does not hold, then there is some  $g \in G \setminus H$  so  $|g|_X < I$ , and some  $h \in H$  so that  $\pi(g) = \pi(h)$ . Without loss of generality, we may pick some such  $h$  so  $|h|_X$  is minimal. Note that  $|h|_X \geq 200\lambda + 200\delta + 12I$ , by the injectivity hypothesis. We let  $\gamma$  be a geodesic joining 1 to  $h$ . By Lemmas 4.1 and 4.2, and the minimality of  $h$ , conclusion (1) of Lemma 4.1 holds.

In this case, there is a  $10\delta$ -local geodesic  $\gamma'$  in  $X'$  connecting 1 to  $\pi(g)$  which lies in a  $2$ -neighborhood of  $\pi(\gamma)$  and coincides with  $\pi(\gamma)$  in a  $(2I + 10\delta)$ -neighborhood of both 1 and  $\pi(g)$ . It follows that  $\gamma'$  has length at least  $4I + 20\delta$ . But since  $\gamma'$  is a  $10\delta$ -local geodesic, it must be a  $(7/3, 2\delta)$ -quasi-geodesic, and so the distance in  $X'$  between 1 and  $\pi(g)$  is at least

$$\frac{3}{7}(4I + 20\delta) - 2\delta > I.$$

It follows that  $|g|_X > I$ , a contradiction.  $\square$

### 4.4 Height decreases under filling

This subsection is devoted to proving that given a relatively quasi-convex subgroup  $H$ , its height decreases after any sufficiently long  $H$ -filling. Our method is the same as the one used by the second and third authors for the results in [8, Part 2]; as such it is inspired by certain hyperbolic 3-manifold arguments by Lackenby [11] and by the first author [1]. Briefly, we choose some minimal counterexample to the theorem, and derive a contradiction by using “area” estimates coming from “pleated surfaces”.

**Theorem 4.6** *Let  $G$  be a torsion-free hyperbolic group,  $H < G$  a quasi-convex subgroup of height  $k$ , and let  $\mathcal{P} = \{P_1, \dots, P_m\}$  be the peripheral structure on  $G$  induced by  $H$ . There is a finite set  $F \subset G$  so that if  $G \xrightarrow{\pi} G(N_1, \dots, N_m)$  is a hyperbolic  $H$ -filling satisfying:*

- (1)  $N_i \triangleleft P_i$  is finite index for all  $i$ , and
- (2)  $N_i \cap F = \emptyset$  for all  $i$ ,

*then  $\pi(H)$  has height strictly less than  $k$  in  $G(N_1, \dots, N_m)$ .*

**Proof** Suppose  $H < G$  is the height  $k$  quasi-convex subgroup and that  $H' < G'$  is the image after filling along finite index subgroups of the malnormal core of  $H$ . The filling map from  $G$  to  $G'$  will be called  $\pi$ . The kernel  $N$  of  $\pi$  will be normally generated by some collection of filling kernels  $N_1, \dots, N_m$ , each normal in some element of the peripheral structure on  $G$  induced by  $H$ , and each contained in  $H$ .

We must show that if some intersection of conjugates of  $H'$  is infinite, then this intersection can be lifted back up to  $G$ . In other words, we will show that infinite order elements of the intersection of essentially distinct conjugates of  $H'$  are always images of infinite order elements of the intersection of essentially distinct conjugates of  $H$ .

We choose some (compatible) generating set  $S$  for  $G$  so that  $X = X(G, \mathcal{P}, S)$  is  $\delta$ -hyperbolic. With respect to this choice,  $H$  is  $\lambda$ -relatively quasi-convex for some  $\lambda$ . By rechoosing  $\delta$ , we can assume that  $\lambda < \delta$ . We also assume, as in [8] that  $\delta$  is an integer greater than or equal to 100. All constants from [8] will be in terms of this re-chosen  $\delta$  for  $X(G, \mathcal{P}, S)$ . In particular, the auxiliary constants  $K = 10\delta$ ,  $L_1 = 1000\delta$  and  $L_2 = 3000\delta$  will be used in the argument below.

Let  $C_h = K + 12\delta + 9$  be the upper bound on the Hausdorff distance between a geodesic and the preferred path with the same endpoints, from [8, Corollary 5.12]. For each  $i$  let  $F_i$  be the ball of  $P_i$ -radius  $2^{L_2}(24 \cdot 2^{4C_h+3} + 24)$  about 1 in  $P_i$ , and let  $F = (\cup_i F_i) \setminus \{1\}$ .

Now fix a hyperbolic  $H$ -filling  $G \xrightarrow{\pi} \bar{G}$  of  $G$  so that for each  $i$ ,  $N_i$  is finite index in  $P_i$  and so that  $N_i \cap F = \emptyset$  for each  $i$ . In other words, for every nontrivial  $n$  in any  $N_i$ , the length in  $P_i$  satisfies

$$(7) \quad |n|_{P_i} > 2^{L_2}(24 \cdot 2^{4C_h+3} + 24).$$

Let  $N = \ker \pi$ . Any element  $g \in N$  is a product in  $G$  of conjugates of elements of  $\cup_i N_i$ . We define the  $N$ -area as the smallest number of such conjugates needed to write  $g$ .

In order to derive a contradiction, assume that  $\pi(H)$  has height at least  $k$ . Thus  $k$  essentially distinct conjugates of  $\pi(H)$  intersect in an infinite (and quasi-convex) subgroup of  $\tilde{G}$ . This subgroup is infinite and hyperbolic, and so it must contain an element of infinite order. It follows that  $\pi(H)$  contains infinite order elements  $a$  and  $b_i$  for  $i \in \{2, \dots, k\}$  and essentially distinct  $\{1, g_2, \dots, g_k\}$  so that

$$(8) \quad a = g_i b_i g_i^{-1}$$

for each  $i$ .

Fix a lift  $\tilde{a}$  of  $a$  closest to 1 in  $X$ , subject to the condition that  $\tilde{a} \in H$ . Now for each  $i \in \{2, \dots, k\}$  choose some  $\tilde{g}_i$  and  $\tilde{b}_i$  subject to the conditions:

$$(C1) \quad \tilde{g}_i \in \pi^{-1}(g_i),$$

$$(C2) \quad \tilde{b}_i \in \pi^{-1}(b_i) \cap H, \text{ and}$$

$$(C3) \quad \tilde{a}^{-1} \tilde{g}_i \tilde{b}_i \tilde{g}_i^{-1} \text{ has minimal } N\text{-area for all choices of } \tilde{g}_i \text{ and } \tilde{b}_i \text{ satisfying conditions (C1) and (C2).}$$

**Claim 4.6.1** For each  $i$ ,  $\tilde{a}^{-1} \tilde{g}_i \tilde{b}_i \tilde{g}_i^{-1}$  has  $N$ -area zero.

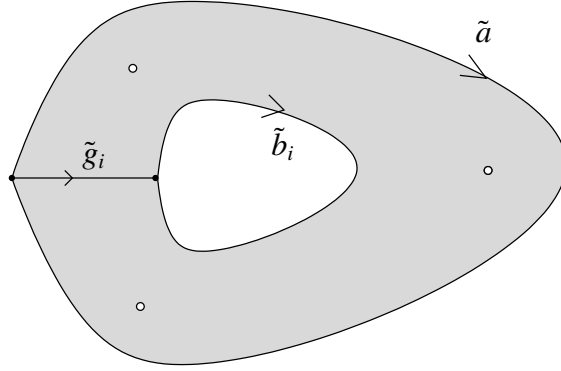
If the word  $\tilde{a}^{-1} \tilde{g}_i \tilde{b}_i \tilde{g}_i^{-1}$  has  $N$ -area zero, then it is equal in  $G$  to 1. Therefore, the claim implies that  $\tilde{a} = \tilde{g}_i \tilde{b}_i \tilde{g}_i^{-1}$  in  $G$  for each  $i$ .

The claim implies the theorem: Certainly, the elements  $\{1, \tilde{g}_2, \dots, \tilde{g}_k\}$  are essentially distinct in  $G$ . Because each  $\tilde{b}_i$  is in  $H$ , the conjugates  $H, H^{\tilde{g}_2}, \dots, H^{\tilde{g}_k}$  all contain the element  $\tilde{a}$ , and so  $H \cap H^{\tilde{g}_2} \cap \dots \cap H^{\tilde{g}_k}$  is infinite. Since  $H$  has height  $k$ , the subgroup  $H \cap H^{\tilde{g}_2} \cap \dots \cap H^{\tilde{g}_k}$  is a conjugate of  $P_j$  for some  $j$ . Since  $N_j$  was chosen to be finite index in  $P_j$ , this implies that  $a = \pi(\tilde{a})$  has finite order in  $\tilde{G}$ , contradicting the original choice of  $a$ .

**Proof of Claim 4.6.1:** If the equation  $\tilde{a}^{-1} \tilde{g}_i \tilde{b}_i \tilde{g}_i^{-1}$  has  $N$ -area  $p > 0$ , then there is some equation

$$(9) \quad \tilde{a}^{-1} \tilde{g}_i \tilde{b}_i \tilde{g}_i^{-1} = \prod_{j=1}^p \alpha_j n_j \alpha_j^{-1}$$

with each  $n_j \in N_{k_j}$  for some  $k_j$ . This equality can be represented by a punctured disk (as in [8, Part 2]) with boundary labelled  $\tilde{a}^{-1} \tilde{g}_i \tilde{b}_i \tilde{g}_i^{-1}$ . There are two sub-segments of the boundary of this disk labelled  $\tilde{g}_i$  and  $\tilde{g}_i^{-1}$ . Gluing these together yields an annulus, again with  $p$  punctures, as in Figure 1. Again as in [8, Part 2], there is a proper map from this punctured annulus into  $X/G$ , so that labelled sub-segments of the

Figure 1: The punctured annulus  $\check{\Sigma}$ .

boundary go to loops representing those elements of  $G$  described by their labels. The distinguished arc labelled  $\tilde{g}_i$  is sent by  $\phi$  to a loop representing  $\tilde{g}_i$ . To be consistent with the notation of Part 2 of [8], we refer to the punctured annulus as  $\check{\Sigma}$ , and the proper map to  $X/G$  as  $\check{\phi}$ .

We define a *reducing arc* for  $\check{\phi}$  to be a proper, essential embedding  $\sigma: \mathbb{R} \rightarrow \check{\Sigma}$ , so that  $\check{\phi} \circ \sigma$  can be properly homotoped to miss any given compact subset of  $X/G$ .

**Subclaim 4.6.1.1** *There are no reducing arcs for  $\check{\phi}$ .*

**Proof** We argue by contradiction. Suppose that  $\sigma: \mathbb{R} \rightarrow \check{\Sigma}$  is a reducing arc.

**Case 1:** We suppose first that the closure of the image of  $\sigma$  in the unpunctured annulus is a homotopically nontrivial loop. In this case, the surface  $\check{\Sigma}$  may be cut along the image of  $\sigma$  to yield a pair of surfaces. The half of the surface with boundary labelled  $\tilde{a}$  represents a proof that  $a = \pi(\tilde{a})$  is parabolic in  $\bar{G}$ . Since every parabolic element of  $\bar{G}$  has finite order, this is a contradiction.

**Case 2:** Now suppose that the image of  $\sigma$  does not intersect the distinguished arc in  $\check{\Sigma}$  labelled  $\tilde{g}_i$ . We may argue very much as in the proof of Claim 9.2 in the proof of Theorem 9.1 of [8]: The case that  $\sigma$  connects two distinct punctures is the same as Case 1 of that argument, and the case that  $\sigma$  connects a puncture to itself is the same as Case 4 of that argument, except that if the reducing arc  $\sigma$  represents a peripheral element which is not in  $N_i$  then the contradiction is to the conclusion of Theorem 9.1 of [8], rather than to the minimality of the diagram. Since

$$2^{L_2}(24 \cdot 2^{4C_h+3} + 24) > 12 \cdot 2^{L_2},$$

the conclusion of Theorem 9.1 of [8] *does* hold. Cases 2 and 3 of the proof from [8, Claim 9.2] do not occur. The upshot here is that if a reducing arc appeared, we would be able to choose a new expression of the form (9) with smaller  $N$ -area by performing a “boundary reduction” of some kind.

**Case 3:** Finally, we suppose that the image of  $\sigma$  intersects the arc labelled  $\tilde{g}_i$ , but we are not in Case 1.

There is a natural basepoint  $\bar{1}$  in  $X/G$ , which is the image of  $1 \in G \subset X$  under the quotient map (and also, of course, the image of any other group element). There is a canonical identification between  $\pi_1(X/G, \bar{1})$  with  $G$ . Let  $\gamma : I \rightarrow \check{\Sigma}$  be any arc with the same endpoints as the distinguished arc labelled by  $\tilde{g}_i$ . Then  $\check{\phi} \circ \gamma$  is a loop in  $X/G$  based at  $\bar{1}$  and so determines a unique element  $g_\gamma$  of  $G$ . The identity

$$a^{-1}\pi(g_\gamma)b_i\pi(g_\gamma) = 1,$$

always holds in  $\bar{G}$ , and has  $N$ -area at most  $p$ .

Since we are not in Case 1, there is an arc  $\gamma$  in  $\check{\Sigma}$ , with the same endpoints as the distinguished arc, which does not intersect  $\sigma$ . We are now in Case 2.

This completes the proof of Subclaim 4.6.1.1.  $\square$

Choose a (partially ideal) triangulation  $\mathcal{T}$  of the punctured annulus  $\check{\Sigma}$  whose vertices are the endpoints of the distinguished arc and the punctures. There are  $2p + 2$  triangles in such a triangulation.

Since there are no reducing arcs, the hypotheses of [8, Lemma 8.6] are satisfied. Let

$$\check{\phi}_{\mathcal{T}} : \check{\Sigma} \rightarrow X/K$$

be the map from [8, Lemma 8.6], which sends each edge of  $\mathcal{T}$  to a preferred path, and let

$$\check{\phi}_{\mathcal{T}} : \text{Skel}(\check{\phi}_{\mathcal{T}}) \rightarrow X/G \cup (\partial_{\mathcal{H}}X)/G,$$

be the map from [8, Remark 8.11], where it is called  $\check{\phi}_{\mathcal{T}}$ . (Elements of  $(\partial_{\mathcal{H}}X)/G$  are  $G$ -orbits of horoball centers, and are in one to one correspondence with  $\mathcal{P}$ .) We note some facts about the skeletal filling  $\text{Skel}(\check{\phi}_{\mathcal{T}})$  and the map  $\check{\phi}_{\mathcal{T}}$ :

- (1) If  $\bar{\Sigma}$  is  $\check{\Sigma}$  with punctures filled in, then the 1-complex  $\text{Skel}(\check{\phi}_{\mathcal{T}})$  embeds naturally in  $\bar{\Sigma}$  so that every edge is either
  - (a) part of one of the edges from  $\mathcal{T}$ ,
  - (b) a *rib* (with image under  $\check{\phi}_{\mathcal{T}}$  a horizontal edge at depth  $L_2$ ), or

(c) a *ligament* (with image under  $\check{\phi}_{\mathcal{T}}$  a vertex at depth  $L_2$ ),

and every vertex is either

(a) coincident with a filled-in puncture of  $\bar{\Sigma}$ ,

(b) a vertex of  $\mathcal{T}$ , or

(c) the endpoint of one or two ribs or ligaments.

The last two types of vertices will be called *ordinary* vertices.

- (2) Every vertex  $v$  coming from a puncture has a *link*, which is a circle in  $\text{Skel}(\check{\phi}_{\mathcal{T}})$  composed of ribs, ligaments, and possibly some sub-segments of edges of  $\mathcal{T}$  in the boundary of  $\check{\Sigma}$ . (See [8, Definition 8.12] for the precise definition.) The map  $\check{\phi}_{\mathcal{T}}$  sends the entire link to that part of  $X/G$  at depth  $L_2$  or more. A puncture is called *interior* if its link is composed entirely of ribs and ligaments. Otherwise it is called *exterior*.
- (3) Every path between ordinary vertices is sent by  $\check{\phi}_{\mathcal{T}}$  to a path in  $X/G$  which is either a based loop at  $\bar{1}$ , or can be made into one in a canonical way by adding vertical segments. Thus any path between ordinary vertices in  $\text{Skel}(\check{\phi}_{\mathcal{T}})$  determines an element of  $G$ .
- (4) The group element determined by the concatenation of paths between ordinary vertices in  $\text{Skel}(\check{\phi}_{\mathcal{T}})$  is the product of those determined by the paths.
- (5) The group element determined by a loop around the link of a vertex is always an element of  $N_i$  for some  $i$ . Different choices of starting point for the loop give rise to elements of  $N_i$  which are conjugate in  $P_i$ .
- (6) Two paths homotopic in  $\bar{\Sigma}$  rel their endpoints determine the same element of  $\bar{G}$ . (This follows immediately from (4) and (5).)

Figure 2 shows an example of what  $\text{Skel}(\check{\phi}_{\mathcal{T}})$  might look like, if there were three exterior and no interior punctures.

If  $T$  is a 2-simplex of  $\mathcal{T}$ , then  $\phi|_{\partial T}$  lifts to a preferred triangle  $\widetilde{\phi|_{\partial T}}: \partial T \rightarrow X$ . Let  $R(T)$  be the number of ribs in  $\text{Skel}(\widetilde{\phi|_{\partial T}})$ , and note that this number does not depend on the lift chosen. [8, Corollary 5.38] implies that  $R(T) \leq 6$ .

Let

$$A(\phi) = \sum_{T \in \mathcal{T}} R(T).$$

[8, Corollary 5.38] immediately implies

$$(10) \quad A(\phi) \leq 6(2p + 2) \leq 24p.$$

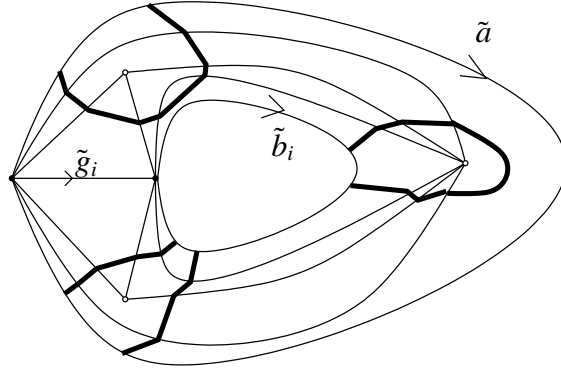


Figure 2: Dark edges are either ribs or ligaments. Hollow circles are vertices of  $\text{Skel}(\check{\phi}_{\mathcal{T}})$  coming from punctures. The (exterior) puncture at the right has a link composed of 7 ribs or ligaments, and a single edge which is part of the edge of  $\mathcal{T}$  labelled by  $\tilde{b}_i$ .

Let  $x$  be a puncture. As remarked above, the link  $\text{Lk}(x)$  is an embedded loop in  $\check{\Sigma}$  whose image in  $X/G$  lies entirely at or below depth  $L_2$ . By joining an arbitrarily chosen basepoint of  $\text{Lk}(x)$  to  $\bar{1}$  in the canonical way, we obtain an element of the filling kernel  $N_i$ , contained in the peripheral subgroup  $P_i$  for some  $i$ .

**Subclaim 4.6.1.2** *Not all punctures of  $\check{\Sigma}$  are interior.*

**Proof** Suppose that all punctures of  $\check{\Sigma}$  were interior. In this case, the image of each link  $\text{Lk}(x)$  lies entirely at depth  $L_2$  in  $X/G$ , and represents a conjugacy class of element of  $N_i$  for some  $i$ . The length of  $\check{\phi}(\text{Lk}(x))$  is the number of ribs in  $\text{Lk}(x)$ . Therefore there are at least  $24 \cdot 2^{4C_h+3} + 24$  ribs in each link. Summing over all the links, there are more than  $24p$  ribs, which contradicts (10) above.

This proves that not all punctures are interior.  $\square$

**Subclaim 4.6.1.3** *No link of a puncture hits the boundary component of  $\check{\Sigma}$  labelled  $\tilde{b}_i$ .*

**Proof** By way of contradiction, we assume that there is some puncture  $x$  so that  $\text{Lk}(x)$  intersects the boundary component of  $\check{\Sigma}$  labelled  $\tilde{b}_i$  in a sub-segment  $I$ . The idea here is that if it did, we would be able to replace  $\tilde{b}_i$  by some  $\tilde{b}'_i$  so that  $\tilde{a}^{-1}\tilde{g}_i\tilde{b}'_i\tilde{g}_i^{-1}$  had smaller  $N$ -area. Our assumption that  $\lambda$  is small with respect to  $\delta$  (and therefore with respect to  $L_2$ ) ensures that this new lift  $\tilde{b}'_i$  still lies in  $H$ .

We will choose  $\tilde{b}'_i$  to be the element determined by the path in  $\text{Skel}(\check{\phi}_{\mathcal{T}})$  which is obtained from the path labelled  $\tilde{b}_i$  by replacing  $I$  with its complement in  $\text{Lk}(x)$ . Let  $\beta$  be the group element represented by the part of the edge labelled by  $\tilde{b}_i$  which precedes  $I$ , and let  $n \in N_j$  be the element represented by the loop around  $\text{Lk}(x)$  starting at the beginning of  $I$  and going around once, clockwise. It is clear that

$$\tilde{b}'_i = \beta n \beta^{-1} \tilde{b}_i$$

maps to  $b_i$ , and has  $N$ -area less than that of  $\tilde{b}_i$ . To establish the subclaim, it remains only to establish that  $\tilde{b}'_i \in H$ .

The restriction of the map  $\check{\phi}_{\mathcal{T}}: \check{\Sigma} \rightarrow X/G$  to the edge  $e$  of  $\check{\Sigma}$  labelled  $\tilde{b}_i$  can be lifted to a preferred path  $\gamma: e \rightarrow X$  joining 1 to  $\tilde{b}_i$ . The map  $\gamma$  sends the sub-segment  $I$  into the  $L_2$ -horoball inside some 0-horoball  $A$ , corresponding to some coset  $tP_j$  of one of the peripheral subgroups.

As in the proof of Lemma 4.2, we first argue that some nontrivial element of  $H$  stabilizes the horoball  $A$ . The preferred path from 1 to  $\tilde{b}_i$  must penetrate at least  $L_2$  into the horoball  $A$ . It follows that a regular geodesic  $\gamma$  penetrates  $A$  at least to depth  $L_2 - C_h > 2500\delta$ . Let  $g_1$  and  $g_2$  be the group elements in  $tP_j$  through which this geodesic passes. By the foregoing, we must have  $d_X(g_1, g_2) > 5000\delta$ . Moreover, the geodesic  $\gamma$  passes through the vertices  $(g_1, \lambda + 1)$  and  $(g_2, \lambda + 1)$  in  $A$ . The  $\lambda$ -relative quasi-convexity of  $H$  implies that there are vertices  $z_1 = \check{\iota}((s_1 D_{j_1}, h_1, n_1))$  and  $z_2 = \check{\iota}((s_2 D_{j_2}, h_2, n_2))$  within  $\lambda$  of  $(g_1, \lambda + 1)$  and  $(g_2, \lambda + 1)$ , respectively. Since  $g_1$  and  $g_2$  are so far apart (recall we have made the assumption  $\lambda \leq \delta$ ), the element  $h_2 h_1^{-1}$  is non-trivial. By the  $H$ -equivariance of  $\check{\iota}$ , this element stabilizes  $A$ , i.e.  $H \cap tP_j t^{-1}$  is non-trivial.

The element  $\beta = tp$  for some  $p \in P_j$ , and so

$$\begin{aligned} \tilde{b}'_i &= \beta n \beta^{-1} \tilde{b}_i \\ &= t(p n p^{-1}) t^{-1} \tilde{b}_i. \end{aligned}$$

Since  $N_j$  is normal in  $P_j$ , the element  $p n p^{-1} \in N_j$ . Because of the assumption that  $\bar{G}$  is an  $H$ -filling of  $G$ , the subgroup  $tN_j t^{-1}$  lies in  $H$ . It follows that  $\tilde{b}'_i \in H$ , completing the proof of the subclaim.  $\square$

By Subclaim 4.6.1.2, some puncture or punctures are exterior; by Subclaim 4.6.1.3, the links of the exterior punctures all intersect the edge of  $\mathcal{T}$  labelled  $\tilde{a}$ , and miss the edge of  $\mathcal{T}$  labelled  $\tilde{b}_i$ . We will show that if any link of a puncture hits the side labelled  $\tilde{a}$ , we can find another lift of  $a$  whose  $X$ -length is smaller, contradicting our initial choice of a shortest lift.

Let  $r_j$  be the number of ribs in the link of the  $j$ 'th puncture. We have

$$\sum_j r_j \leq 24p.$$

Also associated to the puncture is a sub-segment of the preferred path from 1 to  $\tilde{a}$  passing through an  $L_2$ -horoball. Specifically, it passes through some points  $(x_j, L_2)$  and  $(y_j, L_2)$  in a horoball based on  $P_{k_j}$ . Let  $q_j$  be the distance in the  $L_2$ -horosphere between  $v_j = (x_j, L_2)$  and  $w_j = (y_j, L_2)$ . The ribs in the link of the  $j$ 'th puncture give an edge-path in this horosphere from  $(x_j, L_2)$  to  $w'_j = (y_j z_j, L_2)$  for some  $z_j \in N_{k_j}$ . We therefore have

$$r_j + q_j \geq 2^{-L_2} C_N$$

for each  $j$ . (If we are looking at an interior puncture, we have  $q_j = 0$ .) Thus

$$\sum_j r_j + q_j \geq 2^{-L_2} C_N p$$

and

$$\sum_j q_j \geq (2^{-L_2} C_N - 24)p.$$

We now claim that for some  $j$ ,

$$(11) \quad q_j / r_j > 2^R.$$

Indeed, if this is never the case, then

$$\sum_j q_j \leq 2^R \sum_j r_j \leq 24 \cdot 2^R p,$$

which implies that

$$2^{-L_2} C_N - 24 \leq 24 \cdot 2^R,$$

and so

$$C_N \leq 2^{L_2} (24 \cdot 2^R + 24),$$

contradicting (7).

Choose some such  $j$ , and consider the lift  $\tilde{a}'$  of  $a$  obtained by excising the word representing  $x_j^{-1} y_j$  and replacing it by a word representing  $x_j^{-1} y_j z_j$ . The distance from 1 to  $\tilde{a}'$  is at most

$$d(1, v_j) + d(v_j, w'_j) + d(w_j, \tilde{a}'),$$

which is at least  $(\log_2(q_j) - \log_2(r_j)) - 4C_h$  less than the distance from 1 to  $\tilde{a}$  (for this probably a picture should be drawn). Equation (11) implies that  $\tilde{a}'$  is actually shorter than  $\tilde{a}$ , contradicting our initial choice of  $\tilde{a}$ .

This completes the proof of Claim 4.6.1.  $\square$

By Claim 4.6.1, there are lifts  $\tilde{a}$ ,  $\tilde{g}_i$ ,  $\tilde{b}_i$  of  $a$ ,  $g_i$  and  $b_i$  respectively satisfying the conditions (C1–C3) and with

$$\tilde{a} = \tilde{g}_i \tilde{b}_i \tilde{g}_i^{-1}$$

for each  $i \in \{2, \dots, k\}$ . Since  $\{1, g_2, \dots, g_k\}$  are essentially distinct, so are the lifts  $\{1, \tilde{g}_2, \dots, \tilde{g}_k\}$ , and so  $\tilde{a}$  lies in the intersection of  $k$  essentially distinct conjugates of  $H$ . Since  $\tilde{a}$  has infinite order, and  $H$  has height  $k$ , it follows that  $\tilde{a} \in P_j$  for some  $j$ . But since  $N_j$  has finite index in  $P_j$ , the image  $a$  of  $\tilde{a}$  in  $\bar{G}$  must have finite order. This contradiction completes the proof of Theorem 4.6.  $\square$

## 4.5 The proof of Theorem 0.6

Let  $G$  be a torsion-free residually finite hyperbolic group, let  $H$  be a quasi-convex subgroup of  $G$  of height  $k$  and let  $g \in G \setminus H$ .

Let  $\mathcal{D}$  be the malnormal core of  $H$ , and let  $\mathcal{P}$  be the peripheral structure on  $G$  induced by  $H$ .

By Proposition 3.12,  $G$  is hyperbolic relative to  $\mathcal{P}$ ,  $H$  is hyperbolic relative to  $\mathcal{D}$  and  $H$  is  $\lambda$ -relatively quasi-convex in  $G$  for some  $\lambda$ .

Since all the elements of  $\mathcal{P}$  are subgroups of  $G$ , they are residually finite. Thus they contain finite-index normal subgroups  $\{N_i\}$  which induce an  $H$ -filling of  $G$  satisfying the hypotheses of Propositions 4.3, 4.5 and Theorem 4.6.

We claim that the group  $\bar{G} = G(N_1, \dots, N_m)$  satisfies the conclusion of Theorem 0.6. Let  $\eta : G \rightarrow \bar{G}$  be the canonical quotient map. By Proposition 4.3,  $\eta(H)$  is relatively quasi-convex. Since the peripheral subgroups of  $\bar{G}$  are finite, Lemma 3.14 implies that  $\eta(H)$  is actually quasi-convex in the hyperbolic group  $\bar{G}$ . Further,  $\eta(g) \notin \eta(H)$  by Proposition 4.5 and the height of  $\eta(H)$  is at most  $k - 1$  by Theorem 4.6.

## 5 Conclusion

It would be nice to extend the main result of this paper to groups which are residually hyperbolic, that is to groups  $G$  which for any element  $1 \neq g \in G$ , there is a homomorphism  $\varphi : G \rightarrow H$  onto a hyperbolic group such that  $\varphi(g) \neq 1$ . A natural class of such groups are groups  $G$  which are relatively hyperbolic, relative to a finite family  $\mathcal{P} = \{P_1, \dots, P_n\}$  of finitely generated residually finite subgroups of  $G$ . If hyperbolic groups are residually finite, then these groups are also residually finite,

by performing finite fillings on the peripheral subgroups  $\mathcal{P}$  using Theorem 2.1. To generalize Theorem 0.1 to this class of groups, we would need to identify the analogue of quasi-convex subgroups to separate. It would be natural to consider relatively quasi-convex subgroups of  $G$  (as in Definition 3.11). In order to prove that  $G$  is QCERF, it is certainly necessary that the peripheral groups are LERF. Moreover, to make sure that quasi-convex subgroups of  $G$  intersect each conjugate of a peripheral subgroup nicely, it is convenient to assume that subgroups of  $P_i$  are finitely generated. A group for which every subgroup is finitely generated will be called *slender*. Examples of slender LERF groups are polycyclic groups and, in particular, finitely generated nilpotent groups [13]. The following theorem was conjectured (without the hypothesis of torsion-freeness) in an earlier version of this paper, and proved in [14].

**Theorem 5.1** [14] *Suppose that hyperbolic groups are residually finite. Let  $G$  be a torsion-free group which is hyperbolic relative to the peripheral system  $\mathcal{P} = \{P_1, \dots, P_n\}$ . If  $P_i$  is slender and LERF for each  $i$ , then relatively quasi-convex subgroups of  $G$  are separable.*

In particular, this theorem extends Theorem 0.1 to non-uniform lattices in rank one symmetric spaces, since these are hyperbolic relative to nilpotent cusp subgroups.

**Corollary 5.2** [14] *If hyperbolic groups are residually finite, then (3–dimensional) Kleinian groups are LERF.*

Theorem 0.1 also makes it an interesting question to know whether hyperbolic groups are residually finite. See [2, Problem 4.4] for an example of hyperbolic groups which are conjectured to not be residually finite. Also, a question of Olshanskii about exponents of  $m$ -generated finite groups would imply the existence of a hyperbolic group which is not residually finite (see [15] for a discussion of this question). On the other hand, there are many examples known of residually finite hyperbolic groups (see [9] and references therein for recent progress).

## References

- [1] **I Agol**, *Bounds on exceptional Dehn filling*, *Geom. Topol.* 4 (2000) 431–449 (electronic)
- [2] **A Borisov, M Sapir**, *Polynomial maps over finite fields and residual finiteness of mapping tori of group endomorphisms*, *Invent. Math.* 160 (2005) 341–356
- [3] **B Bowditch**, *Relatively hyperbolic groups*, Preprint, available at <http://www.warwick.ac.uk/~masgak/preprints.html> (1999)

- [4] **N Brady**, *Finite subgroups of hyperbolic groups*, Internat. J. Algebra Comput. 10 (2000) 399–405
- [5] **M R Bridson, A Haefliger**, *Metric Spaces of Non-Positive Curvature*, volume 319 of *Grundlehren der mathematischen Wissenschaften*, Springer-Verlag, Berlin (1999)
- [6] **R Gitik, M Mitra, E Rips, M Sageev**, *Widths of subgroups*, Trans. Amer. Math. Soc. 350 (1998) 321–329
- [7] **M Gromov**, *Word hyperbolic groups*, from: “Essays in Group Theory”, (S M Gersten, editor), Mathematical Sciences Research Institute Publications 8, Springer-Verlag, New York (1987) 75–264
- [8] **D Groves, J F Manning**, *Dehn filling in relatively hyperbolic groups*, Israel Journal of Mathematics 168 (2008) 317–429
- [9] **F Haglund, D T Wise**, *Special cube complexes*, Geom. Funct. Anal. 17 (2008) 1551–1620
- [10] **I Kapovich, D T Wise**, *The equivalence of some residual properties of word-hyperbolic groups*, J. Algebra 223 (2000) 562–583
- [11] **M Lackenby**, *Word hyperbolic Dehn surgery*, Invent. Math. 140 (2000) 243–282
- [12] **D D Long, G A Niblo**, *Subgroup separability and 3-manifold groups*, Math. Z. 207 (1991) 209–215
- [13] **A I Mal'cev**, *On homomorphisms onto finite groups.*, American Mathematical Society Translations (2) 119 (1983) 67–79
- [14] **J F Manning, E Martínez-Pedroza**, *Separation of relatively quasiconvex subgroups* Preprint, [arXiv:0811.4001](https://arxiv.org/abs/0811.4001)
- [15] **C Martínez**, *On power subgroups of profinite groups*, Trans. Amer. Math. Soc. 345 (1994) 865–869
- [16] **E Martínez-Pedroza**, *A note on quasiconvexity and relative hyperbolic structures* Preprint, [arXiv:0811.2384](https://arxiv.org/abs/0811.2384)
- [17] **A Minasyan**, *Separable subsets of GFERF negatively curved groups*, J. Algebra 304 (2006) 1090–1100
- [18] **D V Osin**, *Relatively hyperbolic groups: intrinsic geometry, algebraic properties, and algorithmic problems*, Mem. Amer. Math. Soc. 179 (2006) vi+100
- [19] **D V Osin**, *Peripheral fillings of relatively hyperbolic groups*, Invent. Math. 167 (2007) 295–326
- [20] **D T Wise**, *The residual finiteness of negatively curved polygons of finite groups*, Invent. Math. 149 (2002) 579–617

University of California, Berkeley, 970 Evans Hall #3840, Berkeley, CA 94720-3840

Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, 322 Science and Engineering Offices (M/C 249), 851 S. Morgan St., Chicago, IL 60607-7045

Department of Mathematics, SUNY at Buffalo, Buffalo, NY 14260-2900

[ianagol@math.berkeley.edu](mailto:ianagol@math.berkeley.edu), [groves@math.uic.edu](mailto:groves@math.uic.edu), [j399m@buffalo.edu](mailto:j399m@buffalo.edu)