# THE GEOMETRY OF THE DISK COMPLEX 

HOWARD MASUR AND SAUL SCHLEIMER


#### Abstract

This is a work-in-progress and should not be distributed. We prove that the disk complex is Gromov hyperbolic. Remarkably, a necessary step in the proof is a similar analysis of the geometry of the arc complex.

As an application, we find an algorithm which computes the Hempel distance of a Heegaard splitting, up to an error bounded by a function of the genus.


## Contents

1. Introduction ..... 2
1.1. The curve complex ..... 3
1.2. The disk complex ..... 3
1.3. Heegaard distance ..... 5
Acknowledgements ..... 5
2. Background on coarse geometry ..... 5
2.1. Quasi-isometry ..... 5
2.2. Geodesics ..... 6
2.3. Hyperbolicity ..... 6
2.4. A hyperbolicity criterion ..... 7
3. Background on subsurface projection ..... 7
3.1. More complexes ..... 7
3.2. Surfaces and subsurfaces ..... 8
3.3. Natural maps ..... 10
4. Holes in general and the lower bound on distance ..... 11
4.1. Holes, defined ..... 12
4.2. Projection to holes is Lipschitz ..... 13
4.3. Infinite diameter holes ..... 13
4.4. A lower bound on distance ..... 14
5. Holes for the nonorientable surface ..... 14
6. Holes for the arc complex ..... 15
7. Background on three-manifolds ..... 16

[^0]7.1. Compressions 16
7.2. Band sums 18
7.3. Handlebodies and I-bundles 18
8. Holes for the disk complex 19
9. Holes for the disk complex - annuli 20
10. Holes for the disk complex - compressible 23
10.1. Compression sequences of essential disks 23
10.2. Proving the theorem 26
11. Holes for the disk complex - incompressible 26
11.1. Diagonals of polygons 26
11.2. Improving disks 27
11.3. Building the I-bundle 30
11.4. A vertical annulus parallel into the boundary 31
11.5. Finding a pseudo-Anosov map 32
11.6. Corollaries 32
12. Paths in general and the upper bound on distance 34
12.1. The requirements 34
12.2. Basic, inductive, and shortcut intervals 35
12.3. The partition 36
12.4. The upper bound 38
13. Background on Teichmüller space 41
14. Paths for the nonorientable surface 43
15. Paths for the arc complex 44
16. Background on train tracks 44
17. Paths for the disk complex 47
18. The distance estimate 47
19. Hyperbolicity 47
19.1. Index in a hole 48
19.2. Back and sidetracking 49
19.3. Projection control 49
19.4. Finding the midpoint of a side 50
19.5. Another side of the triangle 52
20. An algorithm to coarsely compute Hempel distance 53

Appendix A. Teichmüller geodesics do not backtrack 55
References 59

## 1. Introduction

We use $S_{g, b, c}$ to denote the compact connected surface of genus $g$ with $b$ boundary components and $c$ cross-caps. If the surface is orientable
we omit the subscript $c$ and write $S_{g, b}$. If the surface is closed and orientable we simply write $S_{g}$.

A simple closed curve $\alpha \subset S$ is essential if $\alpha$ does not bound a disk in $S$. The curve $\alpha$ is non-peripheral if $\alpha$ is not isotopic to a component of $\partial S$.

### 1.1. The curve complex.

Definition 1.1 (Harvey [Har81]). The curve complex $\mathcal{C}(S)$ is the simplical complex with vertex set being isotopy classes of essential, nonperipheral curves in $S$. There is a $k$-simplex for every collection of $k+1$ distinct isotopy classes having pairwise disjoint representatives.
Definition 1.2. Fix $\alpha$ and $\beta$, vertices of $\mathcal{C}(S)$. The distance $d_{S}(\alpha, \beta)$ is the minimum possible number of edges of a path in the one-skeleton $\mathcal{C}^{1}(S)$ which starts at $\alpha$ and ends at $\beta$.

For example, if $d_{S}(\alpha, \beta) \geq 3$ then $\alpha$ and $\beta$ fill $S$ : every essential non-peripheral curve in $S$ intersects one of $\alpha$ or $\beta$. An important tool for this paper is the following theorem of the first author and Yair Minsky [MM99]:
Theorem 1.3. The curve complex of an orientable surface is Gromov hyperbolic.
We will also need this result for nonorientable surfaces; as an application of our techniques we give a proof in Corollary 5.4.
1.2. The disk complex. Let $V_{g}$ denote the handlebody of genus $g$ : the three-manifold obtained by taking a closed regular neighborhood of a polygonal, finite, connected graph in $\mathbb{R}^{3}$. The genus of the boundary is the genus of the handlebody. A properly embedded disk $D \subset V$ is essential if $\partial D \subset \partial V$ is essential.
Definition 1.4 (McCullough [McC91]). The disk complex $\mathcal{D}(V)$ is the simplicial complex with vertex set being proper isotopy classes of essential disks in $V$. We have a $k$-simplex for every collection of $k+1$ distinct isotopy classes having pairwise disjoint representatives.

As with the curve complex, define $d_{\mathcal{D}}(D, E)$ to be the distance in the one-skeleton of $\mathcal{D}(V)$ between the disks $D$ and $E$. Our driving interest is in the intrinsic metric on $\mathcal{D}(V)$ and how it differs from that of $\mathcal{C}(\partial V)$. In fact, we will necessarily treat much more general complexes. In this introduction we restrict to statements about the disk complex. An early motivation for our investigation was:
Theorem 8.3. The natural simplicial inclusion $\nu: \mathcal{D}(V) \rightarrow \mathcal{C}(\partial V)$ is not a quasi-isometric embedding.

We develop a theory of the metric obstructions, or holes, arising in Theorem 8.3. The complete classification of holes for $\mathcal{D}(V)$ is given by Theorems 9.1, 10.7, and 11.1. These theorems require an delicate analysis of JSJ decompositions of $V$ caused by widely separated disks in $\mathcal{D}(V)$.

Next, inspired by the distance estimates for the the mapping class group and the pants complex given in [MM00], we obtain an similar estimate for the disk complex:

Theorem 18.1. Distance in the disk complex $\mathcal{D}(V)$ between two disks $D$ and $E$ is quasi-equal to the sum of distances between $D$ and $E$, as projected to holes $X \subset \partial V$ for $\mathcal{D}(V)$. To be precise, for any handlebody $V$ there is a constant $c_{0}=c_{0}(V)$ so that, for any $c \geq c_{0}$ there are constants $a \geq 1$ and $b \geq 0$ so that

$$
d_{\mathcal{D}}(D, E) \stackrel{a, b}{=} \sum\left[d_{X}(D, E)\right]_{c}
$$

independent of the choice of $D$ and $E$. Here the sum ranges over the set of holes $X \subset \partial V$ for the disk complex.

The paper [MM00] verifies its distance estimates by constructing and then using the hierarchy machinary. However, hierarchies do not appear to be flexible enough to deal with the disk complex. Instead we turn to surgery sequences of essential disks, as developed in [MM], and to Lee Mosher's version of train track splitting sequences (Section 16). In fact, surgery and splitting sequences are too flexible; an inductive straightening procedure is developed is Section 12.

With the distance estimate in hand we prove:
Theorem 19.3. The intrinsic metric $d_{\mathcal{D}}$ on $\mathcal{D}(V)$ is Gromov hyperbolic.

The closest precursor to Theorem 19.3 is the paper by Brock and Farb [BF] proving the hyperbolicity of the pants complex of the fiveholed sphere and twice-holed torus. Another proof may be found in Behrstock's thesis [Beh04]. The former uses an area criterion (see Bowditch [Bow91]) while the latter, following [MM00], proves that certain projection maps are coarsely contracting. We instead use a version of quasi-geodesic stability, see Theorem 2.8.

We end this overview with a remark: to understand the geometry of the disk complex, one is unavoidably forced to consider the corresponding questions in the arc complex of a bounded surface as well as the curve complex of a nonorientable surface. Thus our preliminary material is unavoidably somewhat general.
1.3. Heegaard distance. We now turn to an application. Fix a Heegaard splitting: a triple $(S, V, W)$ consisting of a surface and two handlebodies where $V \cap W=\partial V=\partial W=S$. Hempel [Hem01] defines the quantity $d_{S}(V, W)=d_{S}(\mathcal{D}(V), \mathcal{D}(W))$ to be the minimal distance in $\mathcal{C}(S)$ between a essential disk of $V$ and of $W$. Note that a splitting can be completely determined by giving a pair of pants decompositions (maximal simplices), say $\mathbb{D}$ in $\mathcal{D}(V)$ and $\mathbb{E}$ in $\mathcal{D}(W)$. The resulting triple $(S, \mathbb{D}, \mathbb{E})$ is called a Heegaard diagram.

We prove:
Theorem 20.8. There is a constant $R_{2}=R_{2}(S)$ and an algorithm which, given a Heegaard diagram $(S, \mathbb{D}, \mathbb{E})$, computes a number $N$ so that

$$
\left|d_{S}(V, W)-N\right| \leq R_{2}
$$

Here is the context for Theorem 20.8. A long outstanding problem is to find an algorithm which, given a Heegaard diagram, decides whether or not the underlying splitting surface is reducible: has distance zero. Several early ascents of the Poincaré Conjecture fell at essentially this point.

Acknowledgements. We thank Jason Behrstock, Yair Minsky, Lee Mosher, Hossein Namazi, and Kasra Rafi for many enlightening conversations.

## 2. Background on coarse geometry

Here we review a few ideas from coarse geometry. See [BH99], [CDP90], or even [Gro87] for a fuller discussion.
2.1. Quasi-isometry. As a bit of notation, suppose $r, s, a, b$ are nonnegative real numbers, with $a \geq 1$. If $s \leq a \cdot r+b$ then we write $s \leq_{a, b} r$. If $s \leq_{a, b} r$ and $r \leq_{a, b} s$ then we write $s \stackrel{a, b}{=} r$ and call $r$ and $s$ quasi-equal with constants $(a, b)$. We also define the cut-off function $[r]_{c}$ where $[r]_{c}=0$ if $r<c$ and $[r]_{c}=r$ if $r \geq c$.

Suppose that $\left(\mathcal{X}, d_{\mathcal{X}}\right)$ and $\left(\mathcal{Y}, d_{\mathcal{Y}}\right)$ are metric spaces. $\mathrm{A} \operatorname{map} f: \mathcal{X} \rightarrow$ $\mathcal{Y}$ is an $(a, b)$ quasi-isometric embedding for $a \geq 1, b \geq 0$ if, for every $x, y \in \mathcal{X}$,

$$
d_{\mathcal{X}}(x, y) \stackrel{a, b}{=} d_{\mathcal{Y}}(f(x), f(y))
$$

The map $f$ is a quasi-isometry, and $\mathcal{X}$ is quasi-isometric to $\mathcal{Y}$, if $f$ is an $(a, b)$ quasi-isometric embedding and the image of $f$ is $b$-dense: the $b$-neighborhood of the image equals all of $\mathcal{Y}$.

We also allow set-valued maps: $f: \mathcal{X} \rightarrow \mathcal{P}(\mathcal{Y})$ where $\mathcal{P}(\mathcal{Y})$ is the power set of $\mathcal{Y}$. If $Z \subset \mathcal{X}$ we adopt the convention $f(Z)=\bigcup_{z \in Z} f(z)$.

The $\operatorname{map} f$ is an $(a, b)$ quasi-isometric embedding if $d_{\mathcal{X}}(x, y)^{a, b} d_{\mathcal{Y}}\left(x^{\prime}, y^{\prime}\right)$ for all $x^{\prime} \in f(x)$ and all $y^{\prime} \in f(y)$. It necessarily follows that the diameter of $f(x)$ is at most $b$ for any $x \in \mathcal{X}$. Again, $f$ is a quasiisometry if $f(\mathcal{X})$ is $b$-dense in $\mathcal{Y}$.
2.2. Geodesics. Fix an interval $[u, v] \subset \mathbb{R}$. A geodesic, connecting $x$ to $y$ in $\mathcal{X}$, is a $(1,0)$ quasi-isometric embedding $f:[u, v] \rightarrow \mathcal{X}$ with $f(u)=x$ and $f(v)=y$. Often the exact choice of $f$ is unimportant and all that matters are the endpoints $x$ and $y$. We then denote the image of $f$ by $[x, y] \subset \mathcal{X}$.

Fix now intervals $[m, n],[p, q] \subset \mathbb{Z}$. An $(a, b)$ quasi-isometric embed$\operatorname{ding} g:[m, n] \rightarrow \mathcal{X}$ is called an $(a, b)$ quasi-geodesic in $\mathcal{X}$. A function $g:[m, n] \rightarrow \mathcal{X}$ is an ( $a, b, c$ ) unparametrized quasi-geodesic in $\mathcal{X}$ if

- there is an increasing function $h:[p, q] \rightarrow[m, n]$ so that $g \circ$ $h:[p, q] \rightarrow \mathcal{X}$ is an (a,b) quasi-geodesic in $\mathcal{X}$ and
- for all $i \in[p, q-1]$, $\operatorname{diam}_{\mathcal{X}}(g[h(i), h(i+1)]) \leq c$.
(Compare to the definition of ( $K, \delta, s$ ) quasi-geodesics found in [MM99].)
2.3. Hyperbolicity. Suppose that $\mathcal{X}$ is a geodesic metric space. For convenience we now assume that $\mathcal{X}$ is a graph with metric induced by giving all edges length one.
Definition 2.1. The space $\mathcal{X}$ is $\delta$-hyperbolic if, for any three points $x, y, z$ in $\mathcal{X}$ and for any geodesics $k=[x, y], g=[y, z], h=[z, x]$, the triangle $g h k$ is $\delta$-slim: the $\delta$-neighborhood of $h \cup k$ contains $g$.

For the remainder of this section we assume that $\mathcal{X}$ is hyperbolic, $x, y, z \in \mathcal{X}$ are points, and $k, g, h$ are geodesics as in Definition 2.1.
Definition 2.2. We take $\rho_{k}(z)$ to be the closest points projection of $z$ to $k$ :

$$
\rho_{k}(z)=\left\{w \in k \mid d_{\mathcal{X}}(z, w)=d_{\mathcal{X}}(z, k)\right\} .
$$

We now list several lemmas useful in the sequel.
Lemma 2.3. The closest points projection $\rho_{k}(z)$ has diameter at most $4 \delta$.
Lemma 2.4. There is a point on $g$ within distance $2 \delta$ of $\rho_{k}(z)$. The same holds for $h$.
Lemma 2.5. The diameter of $\rho_{g}(x) \cup \rho_{h}(y) \cup \rho_{k}(z)$ is at most $6 \delta$.
Lemma 2.6. Suppose that $z^{\prime}$ is another point in $\mathcal{X}$ so that $d_{\mathcal{X}}\left(z, z^{\prime}\right) \leq$ $R$. Then $d_{\mathcal{X}}\left(\rho_{k}(z), \rho_{k}\left(z^{\prime}\right)\right) \leq 2 R+4 \delta$.
Lemma 2.7. Suppose that $k^{\prime}$ is another geodesic in $X$ so that the endpoints of $k^{\prime}$ are within distance $R$ of the points $x$ and $y$. Then $d_{X}\left(\rho_{k}(z), \rho_{k^{\prime}}(z)\right) \leq R+12 \delta$.
2.4. A hyperbolicity criterion. Here we give a criterion for a metric space to be Gromov hyperbolic. Again, [BH99] discusses in depth related issues and definitions.
Theorem 2.8. The graph $\mathcal{X}$ is Gromov hyperbolic if and only if there are constants $M \geq 0, \delta \geq 0$ and, for all $x, y \in \mathcal{X}^{0}$, a collection $G(x, y)$ of edge paths from $x$ to $y$ with the following properties:

- For all $x, y \in \mathcal{X}^{0}$ and for all $k \in G(x, y)$ if $d_{X}(x, y) \leq 1$ then the path $k$ has length at most $M$.
- For all $x, y, z \in \mathcal{X}^{0}$ and for all $k \in G(x, y), g \in G(y, z), h \in$ $G(z, x)$ the triangle ghk is $\delta$-slim.

Proof. It is enough to prove stability of geodesics: There is a constant $R=R(M, \delta)$ so that any geodesic $c$ in $\mathcal{X}$, connecting $x$ to $y$, and any path $k \in G(x, y)$ satisfy $d_{\text {Haus }}(c, k) \leq R$. It then follows that geodesic triangles are $R+\delta$ slim.

To prove stability, simply follow the plan of Proposition III.H.1.6 and Theorem III.H.1.7 in [BH99] replacing geodesic triangles by triangles of paths from $G$, as necessary. The details are left as an exercise for the interested reader.

## 3. Background on subsurface projection

This section recalls some of the theory of subsurface projections following [MM00].
3.1. More complexes. Suppose that $S$ is a compact connected surface with boundary. A properly embedded $\operatorname{arc} \beta \subset S$ is essential if $\beta$ is not properly isotopic into $\partial S$.
Definition 3.1. The arc complex $\mathcal{A}(S)$ is the simplical complex with vertex set being proper isotopy classes of essential arcs in $S$. We have a $k$-simplex for every collection of $k+1$ distinct classes which have pairwise disjoint representatives.

In identical fashion define $\mathcal{A C}(S)$, the arc and curve complex of $S$, which contains all essential arcs and all essential non-peripheral curves.

Given essential curves or $\operatorname{arcs} \alpha, \beta$ in a surface $X$ the geometric intersection number, $\iota(\alpha, \beta)$, is the minimum intersection possible between $\alpha$ and any $\beta^{\prime}$ properly isotopic to $\beta$. If $\alpha$ and $\beta$ realize their geometric intersection number then $\alpha$ is tight with respect to $\beta$. If they do not realize their geometric intersection then we may tighten (properly isotope) $\beta$ until they do.

Note that the curve complexes $\mathcal{C}\left(S_{1,1}\right)$ and $\mathcal{C}\left(S_{0,4}\right)$, with the current definition, are nonempty but have no edges. It is useful to alter the definition in these cases as follows:

Definition 3.2. Suppose that $X$ is one of $S_{1,1}$ or $S_{0,4}$. Add edges to $\mathcal{C}(X)$ between all vertices with geometric intersection exactly one if $X$ $S_{1,1}$ or two if $X=S_{0,4}$. In both cases the result is the Farey graph.

Notice also that, with current definitions, the curve complex of an annulus $S_{0,2}$ is empty while the arc complex is finite. Following [MM00]:
Definition 3.3. Fix $X \cong S_{0,2}$. We redefine $\mathcal{C}(X)$ to be the complex with vertices being proper isotopy classes of essential arcs where all isotopies fix the boundary of $X$ pointwise. Also, place an edge between two distinct classes if they have representatives with disjoint interiors.

As above, given $\alpha$ and $\beta$ vertices in $\mathcal{C}(S)$ define $d_{S}(\alpha, \beta)$ to be the minimal number of edges required for any path from $\alpha$ to $\beta$ in the one-skeleton of $\mathcal{C}(S)$. Recall that the geometric intersection of a pair of curves gives an upper bound for their distance, as follows:
Lemma 3.4. Suppose that $S$ is a compact connected surface which is not an annulus. If $\alpha, \beta \subset S$ are essential and non-peripheral curves with $\iota(\alpha, \beta)>0$ then $d_{S}(\alpha, \beta) \leq 2 \log _{2}(\iota(\alpha, \beta))+2$.
This form of the inequality, stated for closed orientable surfaces, may be found in [Hem01]. A careful proof in the bounded orientable case is given in [Sch]. The nonorientable case is then an exercise. In the case of an annulus, no bound is required: If $X \cong S_{0,2}$ then an easy induction proves that

$$
\begin{equation*}
d_{X}(\alpha, \beta)=1+\iota(\alpha, \beta) \tag{3.5}
\end{equation*}
$$

(see [MM00, Equation 2.3]) for distinct vertices $\alpha, \beta \in \mathcal{C}(X)$.
3.2. Surfaces and subsurfaces. If $X \subset S$ is a connected compact subsurface we call $X$ essential exactly when all boundary components of $X$ are essential in $Z$.
Definition 3.6. An essential subsurface $X \subset S$ is cleanly embedded when a component $\delta \subset \partial X$ is isotopic into $\partial S$ if and only if $\delta$ is equal to a component of $\partial S$.
As a particular case: if $X \subset S$ is a cleanly embedded annulus and $S$ is not an annulus then $\partial X$ is not parallel into $\partial S$.
Definition 3.7. Suppose $X, Y \subset S$ are essential subsurfaces. If $X$ is cleanly embedded in $Y$ then we say that $X$ is nested in $Y$. If $\partial X$ cuts $Y$ and reversely then we say that $X$ and $Y$ overlap.
Definition 3.8. A compact connected surface $S$ is simple if $\mathcal{A C}(S)$ has finite diameter.

Here we address the question of exactly which surfaces are simple.

Lemma 3.9. Suppose $S$ is a connected compact surface. The following are equivalent:

- $S$ is not simple.
- The diameter of $\mathcal{A C}(S)$ is at least five.
- $S$ admits a filling lamination.
- $S$ admits a pseudo-Anosov map or $S$ is an annulus.

Lemma 4.6 of [MM99] shows that pseudo-Anosov maps have quasigeodesic orbits, when acting on the associated curve complex. A Dehn twist acting on $\mathcal{C}\left(S_{0,2}\right)$ has geodesic orbits.

Note that Lemma 3.9 is only used in this paper when $\partial S$ is nonempty. The closed case is included for completeness. We find Figure 1 to be a highly convenient reference.


Figure 1. The notation $g b c$ stands for $S_{g, b, c}$. Surfaces in boxes are simple. Euler characteristic is constant along lines of slope $1 / 2$.

Proof of Lemma 3.9. We only sketch the proof. If $S$ admits a pseudoAnosov map then the stable lamination is filling. If $S$ admits a filling lamination then, by an argument of Kobayashi [Kob88], $\mathcal{A C}(S)$ has infinite diameter. (This argument is also sketched in [MM99], page 124, after the statement of Proposition 4.6.)

Clearly, if the diameter of $\mathcal{A C}$ is infinite then the diameter is at least equal to five. To finish, one may check directly that all the boxed surfaces in Figure 1 have $\mathcal{A C}(S)$ with diameter at most four. (The difficult cases, 012 and $3 \mathbb{P}$, are discussed by Scharlemann [Sch82].) All surfaces not in boxes, other than $\mathbb{A}$, admit pseudo-Anosov maps. The
orientable cases follow from Thurston's construction [Thu88]. Penner's generalization [Pen88] covers the nonorientable cases.
3.3. Natural maps. We now define several maps between our complexes. Remember that our main concern is with coarse geometric properties. Thus our definitions are allowed to take certain liberties; in particular we are not interested in the exact location of image values but only in their location up to bounded error. To avoid making choices we often use set-valued maps.

Begin by fixing attention on a non-annular, non-simple surface $S$. Choose a hyperbolic metric on $S$ so that all components of $\partial S$ are totally geodesic. Fix also a cleanly embedded $X \subset S$. Let $\widehat{S}$ be the cover of $S$ where $X$ lifts homeomorphically and where $\widehat{S} \cong \operatorname{interior}(X)$. Compactify $\widehat{S}$ via points in its Gromov boundary. In particular, any simple closed curve in $S$ now lifts to a simple closed curve or to a collection of properly embedded arcs. Identify $\mathcal{A C}(X)$ with $\mathcal{A C}(\widehat{S})$.
Definition 3.10. Fix $\alpha \in \mathcal{A C}(S)$. We define the cutting map $\kappa_{X}: \mathcal{A C}(S) \rightarrow$ $\mathcal{P}(\mathcal{A C}(X))$ as follows: the set $\kappa_{X}(\alpha)$ contains all isotopy classes of lifts of $\alpha$ to $\widehat{S}$ which are essential and non-peripheral.

Note that $\kappa_{S}$ is the identity map. If $\alpha$ can be isotoped out of $X$ then $\kappa_{X}(\alpha)=\emptyset$.
Definition 3.11. Suppose that $X$ is not an annulus. Fix $\alpha \in \mathcal{A C}(S)$. We define the surgery map $\sigma_{X}: \mathcal{A C}(X) \rightarrow \mathcal{P}(\mathcal{C}(X))$ as follows: Let $U$ be a closed regular neighborhood of $\alpha \cup \partial X$. Let $\sigma_{X}(\alpha)$ be the set of isotopy classes of components of $\partial U$ which are essential and nonperipheral in $X$.

Recall that $f(A)=\bigcup_{a \in A} f(a)$.
Definition 3.12. If $X$ is not an annulus define the subsurface projection $\pi_{X}: \mathcal{A C}(S) \rightarrow \mathcal{P}(\mathcal{C}(X))$ to be the composition $\pi_{X}=\sigma_{X} \circ \kappa_{X}$. If $X$ is an annulus then set $\pi_{X}=\kappa_{X}$

Fix $\alpha \in \mathcal{A C}(S)$. If $\pi_{X}(\alpha)=\emptyset$ then $\alpha$ misses the surface $X$. If $\pi_{X}(\alpha)$ is nonempty then $\alpha$ cuts the surface $X$. If $\alpha, \beta \in \mathcal{A C}(S)$ both cut $X$ we write $d_{X}(\alpha, \beta)=\operatorname{diam}_{X}\left(\pi_{X}(\alpha) \cup \pi_{X}(\beta)\right)$. This is the projection distance between $\alpha$ and $\beta$ in $X$. As a simple observation we have:
Lemma 3.13. Suppose $\alpha, \beta \in \mathcal{A C}(S)$ are disjoint and cut $X$. Then $\operatorname{diam}_{X}\left(\pi_{X}(\alpha)\right)$ and $d_{X}(\alpha, \beta)$ are at most two.

This follows because for any $\delta, \epsilon \in \pi_{X}(\alpha) \cup \pi_{X}(\beta)$ we find $\iota(\delta, \epsilon) \leq 4$. See Lemma 2.3 of [MM00] for a very similar statement. As a consequence we have:

Corollary 3.14. Fix $X \subset S$. Suppose that $\left\{\beta_{i}\right\}_{i=0}^{n}$ is a path in $\mathcal{A C}(S)$. Suppose that $\beta_{i}$ cuts $X$ for all $i$. Then $d_{X}\left(\beta_{0}, \beta_{n}\right) \leq 2 n$.

If some vertex of $\left\{\beta_{i}\right\}$ does not cut $X$ then the conclusion need not hold: the projection distance $d_{X}\left(\beta_{0}, \beta_{n}\right)$ may be arbitarily large compared to $n$.

A kind of converse to Lemma 3.13 is:
Lemma 3.15. For every $a \in \mathbb{N}$ there is a number $b \in \mathbb{N}$ with the following property: for any $\alpha, \beta \in \mathcal{A C}(S)$ if $d_{X}(\alpha, \beta) \leq a$ for all $X \subset S$ then $\iota(\alpha, \beta) \leq b$.

Corollary D of [CR05] gives a more precise relation between projection distances and intersection number.

Proof of Lemma 3.15. We only sketch the contrapositive: Suppose we are given a sequence of curves $\alpha_{n}, \beta_{n}$ so that $\iota\left(\alpha_{n}, \beta_{n}\right)$ tends to infinity. Passing to subsequences and applying elements of the mapping class group we may assume that $\alpha_{n}=\alpha_{0}$ for all $n$. Setting $c_{n}=\iota\left(\alpha_{0}, \beta_{n}\right)$ and passing to subsequences again we may assume that $\beta_{n} / c_{n}$ converges to $\lambda \in \mathcal{P} \mathcal{M} \mathcal{L}(S)$. Let $Y$ be any connected component of the subsurface filled by $\lambda$, choosen so that $\alpha_{0}$ cuts $Y$. Note that $\pi_{Y}\left(\beta_{n}\right)$ converges to $\lambda \mid Y$. Again applying Kobayashi's argument [Kob88], the distance $d_{Y}\left(\alpha_{0}, \beta_{n}\right)$ tends to infinity.

## 4. Holes in general and the lower bound on distance

Suppose that $S$ is a compact connected surface. A multi-curve is a collection of disjoint non-parallel essential non-peripheral curves or arcs in $S$.

In this paper a combinatorial complex $\mathcal{G}(S)$ will always have isotopy classes of certain multi-curves is $S$ as vertices. Vertices will be connected by edges only if there are representatives which are disjoint. We also always assume that $\mathcal{G}$ is connected. There is a natural map $\nu: \mathcal{G} \rightarrow \mathcal{A C}(S)$ taking a vertex of $\mathcal{G}$ to the isotopy classes of the components.

One may also entertain complexes where edges are placed between multi-curves with bounded intersection, perhaps of a specified kind. Examples in the literature include the pants complex [Bro03] [BDM], the Hatcher-Thurston complex [HT80], and the complex of separating curves [BM04].

It almost always will suffice to study subcomplexes $\mathcal{G} \subset \mathcal{A C}(S)$. However, the more general case arises when dealing with the curve complex or the arc complex of a nonorientable surface. This allows
us to avoid dealing with the Teichmüller space of the nonorientable surface. It also illustrates the generality of our techniques.

For any combinatorial complex $\mathcal{G}$ defined in this paper other than the curve complex we will denote distance in the one-skeleton of $\mathcal{G}$ by $d_{\mathcal{G}}(\cdot, \cdot)$. Distance in $\mathcal{C}(S)$ will always be denoted by $d_{S}(\cdot, \cdot)$.
4.1. Holes, defined. We now develop a technique to measure the failure of the natural map $\nu: \mathcal{G} \rightarrow \mathcal{A C}(S)$ to be a quasi-isometric embedding. Suppose that $S$ is non-simple. Suppose that $\mathcal{G}(S)$ is connected combinatorial complex of multi-curves in $S$. Suppose that $X \subset S$ is an cleanly embedded subsurface. A vertex $\alpha \in \mathcal{G}$ cuts $X$ if some component of $\alpha$ cuts $X$.
Definition 4.1. We say $X \subset S$ is a hole for $\mathcal{G}$ if every vertex of $\mathcal{G}$ cuts $X$.

Almost equivalently, if $X$ is a hole then the subsurface projection $\pi_{X}: \mathcal{G}(S) \rightarrow \mathcal{C}(X)$ never takes the empty set as a value. Note that the entire surface $S$ is always a hole, regardless of our choice of $\mathcal{G}$. Boundary parallel annuli cannot be cleanly embedded (unless $S$ is also an annulus) and so cannot be holes. A hole $X \subset S$ is a strict hole if $X$ is not homeomorphic to $S$.
Example 4.2. Suppose that $S=S_{g, b}$ with $b>0$ and consider the arc complex $\mathcal{A}(S)$. The holes, up to isotopy, are exactly the cleanly embedded surfaces which contain all of $\partial S$. So, for example, if $S$ is planar then only $S$ is a hole for $\mathcal{A}(S)$. The same holds if $S=S_{1,1}$. In all other cases the arc complex admits infinitely many holes.
Definition 4.3. If $X$ is a hole and if $\pi_{X}(\mathcal{G}) \subset \mathcal{C}(X)$ has diameter at least $R$ we say that the hole $X$ has diameter at least $R$.
Example 4.4. Continuing the example above: Since the mapping class group acts on the arc complex, all non-simple holes for $\mathcal{A}(S)$ have infinite diameter.

Suppose now that $X, Y \subset S$ are disjoint holes for $\mathcal{G}$. In the presence of symmetry there can be a relationship between $\pi_{X} \mid \mathcal{G}$ and $\pi_{Y} \mid \mathcal{G}$ as follows:

Definition 4.5. Suppose that $X, Y$ are holes for $\mathcal{G}$ of infinite diameter. Then $X$ and $Y$ are paired if there is a homeomophism $\tau: X \rightarrow Y$ and a constant $K$ so that

$$
d_{Y}\left(\pi_{Y}(\gamma), \tau\left(\pi_{X}(\gamma)\right) \leq K\right.
$$

for every $\gamma \in \mathcal{G}$.
4.2. Projection to holes is Lipschitz. The following lemma is used repeatedly throughout the paper:
Lemma 4.6. Suppose that $\mathcal{G}(S)$ is a combinatorial complex. Suppose that $X$ is a hole for $\mathcal{G}$. Then for any $\alpha, \beta \in \mathcal{G}$ we have

$$
d_{X}(\alpha, \beta) \leq 2+2 \cdot d_{\mathcal{G}}(\alpha, \beta)
$$

The additive error is required only when $\alpha=\beta$.
Proof. This follows directly from Corollary 3.14 and our assumption that vertices of $\mathcal{G}$ connected by an edge represent disjoint multi-curves.

We deduce the well-known:
Lemma 4.7. The inclusion $\nu: \mathcal{C}(S) \rightarrow \mathcal{A C}(S)$ is a quasi-isometry. The surgery map $\sigma_{S}: \mathcal{A C}(S) \rightarrow \mathcal{P}(\mathcal{C}(S))$ is a quasi-inverse for $\nu$.

Proof. Fix a pair of essential, non-peripheral curves $\alpha$ and $\beta$ inside of $S$. Since $\nu$ is an inclusion we have $d_{\mathcal{A C}}(\alpha, \beta) \leq d_{S}(\alpha, \beta)$. As $S$ is a hole for $\mathcal{A C}(S)$ by Lemma 4.6 we have $d_{S}(\alpha, \beta) \leq 2+2 \cdot d_{\mathcal{A C}}(\alpha, \beta)$.

Note that the composition $\sigma_{S} \circ \nu=\operatorname{Id} \mid \mathcal{C}(S)$. Also, for any arc $\alpha \in \mathcal{A}(S)$ we have $d_{\mathcal{A C}}\left(\alpha, \nu\left(\sigma_{S}(\alpha)\right)\right)=1$. Finally, $\mathcal{C}(S)$ is 1-dense in $\mathcal{A C}(S)$, as any $\operatorname{arc} \gamma \subset S$ is disjoint from the one or two curves of $\sigma_{S}(\gamma)$.
4.3. Infinite diameter holes. Brian Bowditch posed several questions (Newton Institute, August 2003) regarding the geometry of $\mathcal{A}(S)$ : in particular, is the inclusion $\nu: \mathcal{A}(S) \rightarrow \mathcal{A C}(S)$ a quasi-isometry? The absence of a natural quasi-inverse is one hint; the presence of holes for $\mathcal{A}$ is another. In fact, we have:
Lemma 4.8. Suppose that $\mathcal{G}(S)$ is a combinatorial complex, and $X \subset$ $S$ is a strict hole of infinite diameter. Then $\nu: \mathcal{G} \rightarrow \mathcal{A C}(S)$ is not a quasi-isometric embedding.

This lemma and Example 4.2 completely determines when the inclusion of $\mathcal{A}(S)$ into $\mathcal{A C}(S)$ is a quasi-isometric embedding and so answers Bowditch's initial question. It quickly becomes clear that the set of holes tightly constrains the intrinsic geometry of a combinatorial complex.
Lemma 4.9. Suppose that $\mathcal{G}(S)$ is a combinatorial complex invariant under the natural action of $\mathcal{M C G}(S)$. Then every non-simple hole for $\mathcal{G}$ has infinite diameter. Furthermore, if $X, Y \subset S$ are disjoint nonsimple holes for $\mathcal{G}$ then there is a quasi-isometric embedding of $\mathbb{Z}^{2}$ into $\mathcal{G}$.

We will not use Lemmas 4.8 or 4.9 in an essential way and accordingly omit proofs. Instead our interest lies in proving the far more powerful distance estimate for $\mathcal{G}(S)$.
4.4. A lower bound on distance. Here we see that the sum of projection distances in holes gives a lower bound for distance.
Theorem 4.10. Fix S, a compact connected non-simple surface. Suppose that $\mathcal{G}(S)$ is a combinatorial complex. Then there is a constant $c_{0}$ so that for all $c \geq c_{0}$ there are constants $a \geq 1, b \geq 0$ satisfying

$$
\sum\left[d_{X}(\alpha, \beta)\right]_{c} \leq_{a, b} d_{\mathcal{G}}(\alpha, \beta)
$$

Here $\alpha, \beta \in \mathcal{G}$ and the sum is taken over all holes $X$ for the complex $\mathcal{G}$.

The proof follows the proof of Theorems 6.10 and 6.12 of [MM00], practially word for word. The only changes necessary are to

- replace the sum over all subsurfaces by the sum over all holes,
- replace Lemma 2.5 of [MM00], which records how markings differing by an elementary move project to an essential subsurface, by Lemma 4.6 of this paper, which records how $\mathcal{G}$ projects to a hole.
One major goal of this paper is to give criteria sufficent obtain the reverse inequality; Theorem 12.1.


## 5. Holes for the nonorientable surface

Fix $F$ a compact, connected, and nonorientable surface. Let $S$ be the orientation double cover with covering map $\rho_{F}: S \rightarrow F$. Let $\tau: S \rightarrow S$ be the associated involution; so for all $x \in S, \rho_{F}(x)=\rho_{F}(\tau(x))$.
Definition 5.1. A multi-curve or multi-arc $\gamma \subset S$ is symmetric if $\tau(\gamma) \cap \gamma=\emptyset$ or $\tau(\gamma)=\gamma$. A multi-curve or multi-arc $\gamma$ is invariant if there is a curve or arc $\gamma^{\prime} \subset F$ so that $\gamma=\rho_{F}^{-1}\left(\gamma^{\prime}\right)$. The same definitions holds for subsurfaces $X \subset S$.
Definition 5.2. The invariant complex $\mathcal{C}^{\tau}(S)$ is the simplical complex with vertex set being isotopy classes of invariant multi-curves. There is a $k$-simplex for every collection of $k+1$ distinct isotopy classes having pairwise disjoint representatives.

Notice that $\mathcal{C}^{\tau}(S)$ is simplicially isomorphic to $\mathcal{C}(F)$. There is also a natural map $\nu: \mathcal{C}^{\tau}(S) \rightarrow \mathcal{C}(S)$. We will prove:
Lemma 5.3. $\nu: \mathcal{C}^{\tau}(S) \rightarrow \mathcal{C}(S)$ is a quasi-isometric embedding.
It thus follows from the hyperbolicity of $\mathcal{C}(S)$ that:
Corollary 5.4. $\mathcal{C}(F)$ is Gromov hyperbolic.

One half of the proof of Lemma 5.3 is straight-forward: since $\nu$ sends adjacent vertices to adjacent edges we have

$$
\begin{equation*}
d_{S}(\alpha, \beta) \leq d_{\mathcal{C}^{\tau}}(\alpha, \beta) \tag{5.5}
\end{equation*}
$$

as long as $\alpha$ and $\beta$ are distinct in $\mathcal{C}^{\tau}(S)$. In fact, since the surface $S$ itself is a hole for $\mathcal{C}^{\tau}(S)$ we may deduce a slightly weaker lower bound from Lemma 4.6 or indeed from Theorem 4.10.

The other half of the proof consists of showing that $S$ is the only hole for $\mathcal{C}^{\tau}(S)$ with large diameter. After a discussion of Teichmüller geodesics we will prove:
Lemma 14.1. There is a constant $K$ with the following property: Suppose that $\alpha, \beta$ are invariant multi-curves or arcs in $S$ and $X \subset S$ so that $d_{X}(\alpha, \beta)>K$. Then $X$ is symmetric.

From this it follows that:
Corollary 5.6. With $K$ as in Lemma 14.1: If $X \subset S$ is a hole for $\mathcal{C}^{\tau}(S)$ with diameter greater than $K$ then $X=S$.

Proof. Suppose that $X \subset S$ is a strict subsurface, cleanly embedded. Suppose that $\operatorname{diam}_{X}\left(\mathcal{C}^{\tau}(S)\right)>K$. Thus $X$ is symmetric. It follows that $\partial X \backslash \partial S$ is also symmetric. Since $\partial X$ does not cut $X$ deduce that $X$ is not a hole for $\mathcal{C}^{\tau}(S)$.

This, together with the upper bound (Theorem 12.1), proves Lemma 5.3.

## 6. Holes for the arc complex

Here we generalize the definition of the arc complex and classify its holes.

Definition 6.1. Suppose that $S$ is a non-simple surface with boundary. Let $\Delta$ be a nonempty collection of components of $\partial S$. The arc complex $\mathcal{A}(S, \Delta)$ is the full subcomplex of $\mathcal{A}(S)$ containing all essential arcs $\alpha \subset S$ with $\partial \alpha \subset \Delta$.

Note that $\mathcal{A}(S, \partial S)$ and $\mathcal{A}(S)$ are identical. If $X$ is non-simple, cleanly embedded in $S$, and $\Delta \subset X$ then the meaning of $\mathcal{A}(X, \Delta)$ is clear.
Lemma 6.2. Suppose $X \subset S$ is cleanly embedded. If $\Delta \subset \partial X$ then $X$ is a hole for $\mathcal{A}(S, \Delta)$.

This follows directly from the definition of a hole. We now have an straight-forward observation:
Lemma 6.3. If $X, Y \subset S$ are holes for $\mathcal{A}(S, \Delta)$ then $X \cap Y \neq \emptyset$.

The proof follows immediately from Lemma 6.2. Lemma 4.9 indicates that Lemma 6.3 is essential to proving that $\mathcal{A}(S, \Delta)$ is Gromov hyperbolic.

In order to prove the upper bound theorem for $\mathcal{A}$ we will use pants decompositions of the surface $S$. In an attempt to avoid complications in the nonorientable case we must carefully lift to the orientation cover. We require the following definitions.

Suppose that $F$ is non-simple, nonorientable, and has nonempty boundary. Let $\rho_{F}: S \rightarrow F$ be the orientation double cover and let $\tau: S \rightarrow S$ be the induced involution. Fix $\Delta^{\prime} \subset \partial F$ and let $\Delta=\rho_{F}^{-1}\left(\Delta^{\prime}\right)$.
Definition 6.4. We define $\mathcal{A}^{\tau}(S, \Delta)$ to be the invariant arc complex: vertices are invariant multi-arcs and simplices arise from disjointness.

Again, $\mathcal{A}^{\tau}(S, \Delta)$ is simplically isomorphic to $\mathcal{A}\left(F, \Delta^{\prime}\right)$. Now, it follows from Lemma 14.1 that all holes for $\mathcal{A}^{\tau}(S, \Delta)$, with sufficiently large diameter, are symmetric subsurfaces $X \subset S$ so that $\Delta \subset X \cup \tau(X)$. If $X \cap \tau(X)=\emptyset$ then the subsurfaces $X$ and $\tau(X)$ are paired holes, as in Definition 4.5. Notice as well that all non-simple symmetric holes $X \subset S$ for $\mathcal{A}^{\tau}(S, \Delta)$ have infinite diameter.

Unlike $\mathcal{A}\left(F, \Delta^{\prime}\right)$ the complex $\mathcal{A}^{\tau}(S, \Delta)$ may have disjoint holes. Nonetheless, all infinite diameter holes for the invariant arc complex interfere: any such hole $Z$ intersects any other such hole $X$ or intersects a hole paired with $X$.
Lemma 6.5. Suppose that $X, Z \subset S$ are holes for $\mathcal{A}^{\tau}(S, \Delta)$, both of infinite diameter. Then $Z$ intersects $X$ or $\tau(X)$.
Proof. Notice that $\Delta \subset \partial X \cup \partial Y$. Since $\Delta$ is also contained in $Z \cup \tau(Z)$, the subsurface $Z$ meets either $X$ or $Y$.

## 7. Background on three-manifolds

Before discussing the holes in the disk complex, we record a few facts about handlebodies and $I$-bundles needed in the sequel.

Fix $M$ a compact connected irreducible three-manifold. Recall that $M$ is irreducible if every embedded two-sphere in $M$ bounds a threeball. Recall that if $N$ is a closed submanifold of $M$ then $\operatorname{fr}(N)$ is the closure of $\partial N \backslash \partial M$.
7.1. Compressions. Suppose that $F$ is a surface embedded in $M$. Then $F$ is compressible if there is a disk $B$ embedded in $M$ with $\operatorname{interior}(B) \cap \partial M=\emptyset, B \cap F=\partial B$, and $\partial B$ essential in $F$. Any such disk $B$ is called a compression of $F$.

In this situation form a new surface $F^{\prime}$ as follows: Let $N$ be a closed regular neighborhood of $B$. First remove from $F$ the annulus $N \cap F$.

Now form $F^{\prime}$ by gluing on both disk components of $\partial N \backslash F$. We say that $F^{\prime}$ is obtained by compressing $F$ along $B$. If no such disk exists we say $F$ is incompressible.
Definition 7.1. An embedded surface $F$ is boundary compressible if there is a disk $B$ embedded in $M$ with

- interior $(B) \cap \partial M=\emptyset$,
- $\partial B$ is a union of connected $\operatorname{arcs} \alpha$ and $\beta$,
- $\alpha \cap \beta=\partial \alpha=\partial \beta$,
- $B \cap F=\alpha$ and $\alpha$ is properly embedded in $F$,
- $B \cap \partial M=\beta$, and
- $\beta$ is essential in $\partial M \backslash \partial F$.

A disk, like $B$, with boundary partitioned into two arcs is called a bigon. Note that this definition of boundary compression is slightly weaker than others found in the literature; the $\operatorname{arc} \alpha$ is often required to be essential in $F$. We do not require this additional property because, for us, $F$ will usually be a properly embedded disk in a handlebody.

Just as for compressing disks we may boundary compress $F$ along $B$ to obtain a new surface $F^{\prime}$ : Let $N$ be a closed regular neighborhood of $B$. First remove from $F$ the rectangle $N \cap F$. Now form $F^{\prime}$ by gluing on both bigon components of $\operatorname{fr} N \backslash F$. Again, $F^{\prime}$ is obtained by boundary compressing $F$ along $B$. If no such bigon exists then $F$ is boundary incompressible.
Remark 7.2. Recall that any surface $F$ properly embedded in a handlebody $V_{g}, g \geq 2$, is either compressible or boundary compressible.

Suppose now that $F$ is properly embedded in $M$ and $\Gamma$ is a multicurve in $\partial M$.

Remark 7.3. Suppose that $F^{\prime}$ is obtained by a boundary compression of $F$ performed in the complement of $\Gamma$. Suppose that $F^{\prime}=F_{1} \cap F_{2}$ is disconnected and, after tightening, each $F_{i}$ meets $\Gamma$. Then $\iota\left(\partial F_{i}, \Gamma\right)<$ $\iota(\partial F, \Gamma)$ for $i=1,2$.

It is often useful to restrict our attention to boundary compressions meeting a single subsurface of $\partial M$. So suppose that $X \subset \partial M$ is an essential subsurface. Suppose that $\partial F$ is tight with respect to $\partial X$. Suppose $B$ is a boundary compression of $F$. If $B \cap \partial M \subset X$ we say that $F$ is boundary compressible into $X$.
Lemma 7.4. Suppose that $M$ is irreducible. Fix $X$ an essential subsurface of $\partial M$. Let $F \subset M$ be a properly embedded, incompressible surface. Suppose that $\partial X$ and $\partial F$ are tight and that $X$ compresses in $M$. Then either:

- $F \cap X=\emptyset$,
- $F$ is boundary compressible into $X$, or
- $F$ is a disk with $\partial F \subset X$.

Proof. Suppose that $X$ is compressible via a disk $E$. Isotope $E$ to make $\partial E$ tight with respect to $\partial F$. This can be done while maintaining $\partial E \subset X$ because $\partial F$ and $\partial X$ are tight. Since $M$ is irreducible and $F$ is incompressible we may isotope $E$, rel $\partial$, to remove all simple closed curves of $F \cap E$. If $F \cap E$ is nonempty then an outermost bigon of $E$ gives the desired boundary compression lying in $X$.

Suppose instead that $F \cap E=\emptyset$ but $F$ does meet $X$. Let $\delta \subset X$ be a simple arc meeting each of $F$ and $E$ in exactly one endpoint. Let $N$ be a closed regular neighborhood of $\delta \cup E$. Note that $\operatorname{fr}(N) \backslash F$ has three components. One is a properly embedded disk parallel to $E$ and the other two $B, B^{\prime}$ are bigons attached to $F$. At least one of these, say $B^{\prime}$ is trivial in the sense that $B^{\prime} \cap \partial M$ is a trivial arc embedded in $\partial M \backslash \partial F$. If $B$ is non-trivial then $B$ provides the desired boundary compression.

Suppose that $B$ is also trivial. It follows that $\partial E$ and one component $\gamma \subset \partial F$ cobound an annulus $A \subset X$. So $D=A \cup E$ is a disk with $(D, \partial D) \subset(M, F)$. As $\partial D=\gamma, F$ is incompressible, and $M$ is irreducible deduce that $F$ is isotopic to $E$.
7.2. Band sums. Band sum is the inverse operation to boundary compression: Fix a pair of disjoint properly embedded surfaces $F_{1}, F_{2} \subset M$. Let $F^{\prime}=F_{1} \cup F_{2}$. Fix a simple $\operatorname{arc} \delta \subset \partial M$ so that $\delta$ meets each of $F_{1}$ and $F_{2}$ in exactly one point of $\partial \delta$. Let $N \subset M$ be a closed regular neighborhood of $\delta$. Form a new surface by adding to $F^{\prime} \backslash N$ the rectangle component of $\operatorname{fr}(N) \backslash F^{\prime}$. The surface $F$ obtained is the result of band summing $F_{1}$ to $F_{2}$ along $\delta$. It is straight-forward to show that $F$ has a boundary compression dual to $\delta$ yielding $F^{\prime}$ : that is, there is a boundary compression $B$ for $F$ so that $\delta \cap B$ is a single point and compressing $F$ along $B$ gives $F^{\prime}$.
Remark 7.5. Note that, if $F$ is a band sum of $F^{\prime}$ then every component of $\partial F$ in $\partial M$ is disjoint from, or parallel to, every curve of $\partial F^{\prime}$.
7.3. Handlebodies and I-bundles. Recall that handlebodies are irreducible.

Suppose that $F$ is a compact connected surface with at least one boundary component. Let $T$ be the orientation $I$-bundle over $F$. If $F$ is orientable then $T \cong F \times I$. If $F$ is not orientable then $T$ is the unique $I$-bundle over $F$ with orientable total space. We call $T$ the $I$-bundle and $F$ the base space. Let $\rho_{F}: T \rightarrow F$ be the associated bundle map. Note that $T$ is homeomorphic to a handlebody.

If $A \subset T$ is a union of fibres of the map $\rho_{F}$ then $A$ is vertical with respect to $T$. In particular take $\partial_{v} T=\rho_{F}^{-1}(\partial F)$ to be the vertical boundary of $T$. Take $\partial_{h} T$ to be the union of the boundaries of all of the fibres: this is the horizontal boundary of $T$. Note that $\partial_{h} T$ is always incompressible in $T$ while $\partial_{v} T$ is incompressible in $T$ as long as $F$ is not homeomorphic to a disk.

Note that, as $\left|\partial_{v} T\right| \geq 1$, any vertical surface in $T$ can be boundary compressed. However no vertical surface in $T$ may be boundary compressed into $\partial_{h} T$.

We end this section with:
Lemma 7.6. Suppose that $F$ is a compact, connected surface with $\partial F \neq \emptyset$. Let $\rho_{F}: T \rightarrow F$ be the orientation I-bundle over $F$. Let $X$ be a component of $\partial_{h} T$. Let $D \subset T$ be a properly embedded disk. If

- $\partial D$ is essential in $\partial T$,
- $\partial D$ and $\partial X$ are tight, and
- D cannot be boundary compressed into $X$
then $D$ may be properly isotoped to be vertical with respect to $T$.


## 8. Holes for the disk complex

Here we begin to classify the holes for the disk complex, a more difficult analysis than that of the arc complex. To fix notation let $V$ be a handlebody. Let $S=S_{g}=\partial V$. Recall that there is a natural inclusion $\nu: \mathcal{D}(V) \rightarrow \mathcal{C}(S)$.
Remark 8.1. The notion of a hole $X \subset \partial V$ for $\mathcal{D}(V)$ may be phrased in several different ways:

- every essential disk $D \subset V$ cuts the surface $X$,
- $\overline{S \backslash X}=\overline{\partial V \backslash X}$ is incompressible in $V$, or
- $X$ is disk busting in $V$.

The classification of holes $X \subset S$ for $\mathcal{D}(V)$ breaks roughly into three cases: either $X$ is an annulus, is compressible in $V$, or is incompressible in $V$. In each case we obtain a result:
Theorem 9.1. Suppose $X$ is a hole for $\mathcal{D}(V)$ and $X$ is an annulus. Then the diameter of $X$ is at most 3 .
Theorem 10.7. Suppose $X$ is a compressible hole for $\mathcal{D}(V)$ with diameter at least 15. Then there are a pair of essential disks $D, E \subset V$ so that

- $\partial D, \partial E \subset X$ and
- $\partial D$ and $\partial E$ fill $X$.

Thus $X$ supports a pseudo-Anosov map.

Theorem 11.1. Suppose $X$ is an incompressible hole for $\mathcal{D}(V)$ with diameter at least $4 \log _{2}(g)+60$. Then there is an $I$-bundle $\rho_{F}: T \rightarrow F$ embedded in $V$ so that

- $\partial_{h} T \subset \partial V$,
- $X$ is isotopic in $S$ to a component of $\partial_{h} T$,
- some component of $\partial_{v} T$ is boundary parallel into $\partial V$,
- F supports a pseudo-Anosov map.

As a corollary of these theorems we have:
Corollary 8.2. If $X$ is hole for $\mathcal{D}(V)$ with diameter at least $4 \log _{2}(g)+$ 60 then $X$ has infinite diameter.

Proof. We only give a sketch. If $X$ is a hole with diameter at least $4 \log _{2}(g)+60$ then either Theorem 10.7 or 11.1 applies.

If $X$ is compressible then Dehn twists, in opposite directions, about the given disks $D$ and $E$ yields a automorphisms $f: V \rightarrow V$ so that $f \mid X$ is pseudo-Anosov. This follows from Thurston's construction [Thu88]. By Lemma 3.9 the hole $X$ has infinite diameter.

If $X$ is incompressible then $X \subset \partial_{h} T$ where $\rho_{F}: T \rightarrow F$ is the given $I$-bundle. Let $f: F \rightarrow F$ be the given pseudo-Anosov map. So $g$, the suspension of $f$, gives a automorphism of $V$. Again it follows that the hole $X$ has infinite diameter.

Applying Lemma 4.8 we find another corollary:
Theorem 8.3. If $\partial V$ contains a strict hole with diameter at least $4 \log _{2}(g)+60$ then the inclusion $\nu: \mathcal{D}(V) \rightarrow \mathcal{C}(\partial V)$ is not a quasiisometric embedding.

## 9. Holes for the disk complex - Annuli

The proof of Theorem 9.1 occupies the rest of this section. This proof shares many features with the proofs of Theorems 10.7 and 11.1. However, the exceptional definition of $\mathcal{C}\left(S_{0,2}\right)$ prevents a unified approach. Fix $V$, a handlebody.
Theorem 9.1. Suppose $X$ is a hole for $\mathcal{D}(V)$ and $X$ is an annulus. Then the diameter of $X$ is at most 3 .

Assume, to obtain a contradiction, that $X$ has diameter at least 4. Suppose that $D \in \mathcal{D}(V)$ is a disk choosen to minimize $D \cap X$. Among all disks $E \in \mathcal{D}(V)$ with $d_{X}(D, E) \geq 2$ choose one which minimizes $|D \cap E|$. Isotope $D$ and $E$ to make the boundaries tight and also tight with respect to $\partial X$. Tightening triples of curves is not canonical; nonetheless there is a tightening so that $S \backslash(\partial D \cup \partial E \cup X)$ contains no triangles. See Figure 2.


Figure 2. Triangles outside of $X$ (see the left side) can be moved in (see the right side). This decreases the number of points of $D \cap E \cap(S \backslash X)$.

After this tightening we have:
Claim. Every arc of $\partial D \cap X$ meets every arc of $\partial E \cap X$ at least once.
Proof. This follows from two observations. First, after the tightening we find that the arcs $D \cap X$ and $E \cap X$ lie inside of $X$ just as $\pi_{X}(D)$ and $\pi_{X}(E)$ lie inside of $\widehat{X}$, the annular cover of $\partial V$. Second, by Equation 3.5, $d_{X}(D \cap X, E \cap X)=1+\iota(D \cap X, E \cap X)$.

Claim. There is an outermost bigon $B \subset E \backslash D$ with the following properties:

- $\partial B=\alpha \cup \beta$ where $\alpha=B \cap D, \beta=\partial B \backslash \alpha \subset \partial E$,
- $\partial \alpha=\partial \beta \subset X$, and
- $|\beta \cap X|=2$.

Furthermore, $|D \cap X|=2$.
See the lower right of Figure 3 for a picture.
Proof. Consider the intersection of $D$ and $E$, thought of as a collection of arcs and curves in $E$. Any simple closed curve component of $D \cap E$ can be removed by an isotopy of $E$, fixed on the boundary. (This follows from the irreducibility of $V$ and an innermost disk argument.) Since we have assumed that $|D \cap E|$ is minimal it follows that there are no simple closed curves in $D \cap E$.

So consider any outermost bigon $B \subset E \backslash D$ : a subdisk where $B \cap D$ is a single arc lying in the boundary of $B$. Let $\alpha=B \cap D$. Let $\beta=\partial B \backslash \alpha=B \cap \partial V$. Note that $\beta$ cannot completely contain a component of $E \cap X$ as this would contradict either the fact that $B$ is outermost or the fact that every arc of $E \cap X$ meets some arc of $D \cap X$. Using this observation, Figure 3 lists the possible ways for $B$ to lie inside of $E$.


Figure 3. The arc $\alpha$ cuts $B$ off of $E$. The darker part of $\partial E$ are the arcs of $E \cap X$. Either $\beta$ is disjoint from $X$, $\beta$ is contained in $X, \beta$ meets $X$ in a single subarc, or $\beta$ meets $X$ in two subarcs.

Let $D^{\prime}$ and $D^{\prime \prime}$ be the two essential disks obtained by boundary compressing $D$ along the bigon $B$. Suppose $\alpha$ is as shown in one of the first three pictures of Figure 3. It follows that either $D^{\prime}$ or $D^{\prime \prime}$ has, after tightening, smaller intersection with $X$ than $D$ does, a contradiction. We deduce that $\alpha$ is as pictured in lower right of Figure 3.

Boundary compressing $D$ along $B$ still gives disks $D^{\prime}, D^{\prime \prime} \in \mathcal{D}(V)$. As these cannot have smaller intersection with $X$ we deduce that $|D \cap X|=$ 2 and the claim holds.

Using the same notation as in the proof above, let $B$ be an outermost bigon of $E \backslash D$. We now study how $\alpha \subset \partial B$ lies inside of $D$.
Claim. The arc $\alpha \subset D$ connects distinct components of $D \cap X$.
Proof. Suppose not. Then there is a bigon $C \subset D \backslash \alpha$ with $\partial C=\alpha \cup \gamma$ and $\gamma \subset \partial D \cap X$. The disk $C \cup B$ is essential and meets $X$ at most once after tightening. If $C \cup B$ is disjoint from $X$ then $X$ is not a hole, a contradiction. If $C \cup B$ meets $X$ exactly once then we may form a new disk $C^{\prime}$ by boundary compressing $X$ along $C \cup B$. Then this disk $C^{\prime}$ is disjoint from $X$, a contradiction.

We finish the proof of Theorem 9.1 by noting that $D \cup B$ is a tripod: $D \cup B$ is homeomorphic to $\Upsilon \times I$ where $\Upsilon$ is the simplicial graph with three edges and four vertices, three of valence one. We may choose the homeomorphism so that $(D \cup B) \cap X=\Upsilon \times \partial I$. It follows that we may properly isotope $D \cup B$ until $(D \cup B) \cap X$ is a pair of arcs. Recall that $D^{\prime}$ and $D^{\prime \prime}$ are the disks obtained by boundary compressing $D$ along $B$. It follows that one of $D^{\prime}$ or $D^{\prime \prime}$ (or both) meets $X$ in at most a
single arc. As in the previous paragraph we find that $X$ is not a hole, a contradiction.

## 10. Holes for the disk complex - Compressible

The proof of Theorem 10.7 occupies the second half of this section.
10.1. Compression sequences of essential disks. Fix a nonempty, pairwise disjoint collection of essential curves $\Gamma \subset \partial V$. Fix also a essential disk $D \subset V$. Properly isotope $D$ to make $\partial D$ tight with respect to $\Gamma$.

Assume for the moment that $D$ meets some component of $\Gamma$.
Definition 10.1. Define a compression sequence of essential disks as follows: $\left\{\left(\Delta_{k}, B_{k}\right)\right\}_{k=1}^{n}$ where $\Delta_{1}=\{D\}, B_{k}$ is a boundary compression of $\Delta_{k}, B_{k}$ is disjoint from $\Gamma$, and $\Delta_{k+1}$ is obtained by boundary compressing $\Delta_{k}$ along $B_{k}$ and tightening with respect to $\Gamma$. Note that $\Delta_{k}$ is a collection of exactly $k$ pairwise disjoint disks properly embedded in $V$. We further require, for $k \leq n$ that every disk of every $\Delta_{k}$ meets some component of $\Gamma$. We call a compression sequence maximal if either

- no disk of $\Delta_{n}$ can be boundary compressed into $S \backslash \Gamma$ or
- there is a component $Z \subset S \backslash \Gamma$ and a boundary compression of $\Delta_{n}$ into $S \backslash \Gamma$ yielding an essential disk $D^{\prime}$ with $\partial D^{\prime} \subset Z$.
We say that such maximal sequences end essentially or end in $Z$, respectively.

Note that all compression sequences must end by Remark 7.3. Given a maximal sequence we may relate the various disks in the sequence as follows:
Definition 10.2. Fix $X$, a component of $S \backslash \Gamma$. Fix $D_{k} \in \Delta_{k}$. A disjointness pair for $D_{k}$ is an ordered pair $(\alpha, \beta)$ of essential $\operatorname{arcs}$ in $X$ where

- $\alpha \subset D_{k} \cap X$,
- $\beta \subset \Delta_{n} \cap X$, and
- $d_{\mathcal{A}}(\alpha, \beta) \leq 1$.

If $\alpha \neq \alpha^{\prime}$ then the two disjointness pairs $(\alpha, \beta)$ and $\left(\alpha^{\prime}, \beta\right)$ are distinct, even if $\alpha$ is properly isotopic to $\alpha^{\prime}$. A similar remark holds for the second coordinate.

We require the following somewhat technical lemma:
Lemma 10.3. Fix $\Gamma \subset \partial V$ a nonempty collection of pairwise disjoint essential curves. Suppose that $D$ meets $\Gamma$ and choose a maximal sequence starting at $D$. Fix any component $X \subset \partial V \backslash \Gamma$ so that $\Delta_{n}$ cuts
X. Fix any disk $D_{k} \in \Delta_{k}$ which also cuts $X$. Then either $D_{k} \in \Delta_{n}$ or there are four distinct disjointness pairs $\left\{\left(\alpha_{i}, \beta_{i}\right)\right\}_{i=1}^{4}$ for $D_{k}$ where each of the arcs $\left\{\alpha_{i}\right\}$ appears as the first coordinate of at most two pairs.

Proof. We induct on $n-k$. If $D_{k}$ is contained in $\Delta_{n}$ there is nothing to prove. If $D_{k}$ is contained in $\Delta_{k+1}$ we are done by induction. If not, then $D_{k}$ is the unique disk of $\Delta_{k}$ which is boundary compressed at level $k$. Let $D_{k+1}, D_{k+1}^{\prime} \in \Delta_{k+1}$ be the two disks obtained after boundary compressing $D_{k}$ along the bigon $B_{k}$.

Let $\delta$ be the band sum arc dual to $B_{k}$. We may assume that $|\Gamma \cap \delta|$ is minimal over all arcs dual to $B_{k}$. It follows that the band sum of $D_{k+1}$ with $D_{k+1}^{\prime}$ along $\delta$ is tight, without any isotopy. (This is where we use the fact that $B_{k}$ is a boundary compression in the complement of $\Gamma$, as opposed to being a general boundary compression of $D_{k}$ in $V$.) There are now three possibilities: $|X \cap \partial \delta|$ equals zero, one, or two. We deal with each in turn.

First suppose that $X \cap \partial \delta=\emptyset$. Then $D_{k} \cap X$ contains $\left(D_{k+1} \cup D_{k+1}^{\prime}\right) \cap$ $X$. If $D_{k+1}$ and $D_{k+1}^{\prime}$ are both components of $\Delta_{n}$ then choose any arc $\beta \subset D_{k+1} \cap X$ and any arc $\beta^{\prime} \subset D_{k+1}^{\prime} \cap X$. The four disjointness pairs are then all ordered combinations of $\beta$ and $\beta^{\prime}$. Suppose instead that $D_{k+1}$, say, is not a component of $D_{n}$. Then as $D_{k+1} \cap X \subset D_{k} \cap X$ the disk $D_{k}$ inherits four disjointness pairs from $D_{k+1}$.

Second suppose that $X \cap \partial \delta$ is a single point of $D_{k+1}$, contained in $\gamma \subset D_{k+1} \cap X$. Let $\alpha, \alpha^{\prime}$ be the arcs of $D_{k} \cap X$ meeting a regular neighborhood of $\delta \cup \gamma$. Both of these are disjoint from $\gamma$ in $X$, by Remark 7.5. Let $\beta$ be any arc of $D_{k+1}^{\prime} \cap X$.

If $D_{k+1} \notin \Delta_{n}$ and $\gamma$ is not the first coordinate of one of $D_{k+1}$ 's four pairs then $D_{k}$ inherits disjointness pairs from $D_{k+1}$. If $D_{k+1}^{\prime} \notin \Delta_{n}$ then $D_{k}$ inherits disjointness pairs from $D_{k+1}^{\prime}$.

Thus we may assume that both $D_{k+1}$ and $D_{k+1}^{\prime}$ are in $\Delta_{n}$ or that only $D_{k+1}^{\prime} \in \Delta_{n}$ while $\gamma$ appears as the first arc of disjointness pair for $D_{k+1}$. In case of the former the required disjointness pairs are $(\beta, \beta)$, $(\alpha, \beta),(\alpha, \gamma)$, and $\left(\alpha^{\prime}, \gamma\right)$. In case of the latter we do not know if $\gamma$ is allowed to appear as the second coordinate of a pair. However we are given four disjointness pairs for $D_{k+1}$ and are told that $\gamma$ appears as the first coordinate of at most two of these pairs. Hence the other two pairs are inherited by $D_{k}$. The pairs $(\beta, \beta)$ and $(\alpha, \beta)$, say, give the desired conclusion.

Third suppose that $X \cap \partial \delta$ meets $\gamma \subset D_{k+1}$ and $\gamma^{\prime} \subset D_{k+1}^{\prime}$. Let $\alpha$ and $\alpha^{\prime}$ be two arcs of $D_{k} \cap X$ meeting the regular neighborhood of $\left(\delta \cup \gamma \cup \gamma^{\prime}\right) \cap X$. These are disjoint from all arcs of $\left(D_{k+1} \cup D_{k+1}^{\prime}\right) \cap X$ again by Remark 7.5. Suppose both $D_{k+1}$ or $D_{k+1}^{\prime}$ lie in $\Delta_{n}$. Then
the desired pairs are $(\alpha, \gamma),\left(\alpha^{\prime}, \gamma\right),\left(\alpha, \gamma^{\prime}\right)$, and $\left(\alpha^{\prime}, \gamma^{\prime}\right)$. If $D_{k+1}^{\prime} \in \Delta_{n}$ while $D_{k+1}$ is not then $D_{k}$ inherits two pairs from $D_{k+1}$. We add to these the pairs $\left(\alpha, \gamma^{\prime}\right)$, and $\left(\alpha^{\prime}, \gamma^{\prime}\right)$. If neither disk lies in $\Delta_{n}$ then $D_{k}$ inherits two pairs from each and the proof is complete.

Lemma 10.3 allows us to relate a disk $D$ to another disk $D^{\prime}$ obtained from $D$ by repeated boundary compressions in the complement of $\Gamma$.
Lemma 10.4. Fix $X \subset \partial V$, a hole for $\mathcal{D}(V)$. For any disk $D \in \mathcal{D}(V)$ there is a disk $D^{\prime}$ with the following properties:

- $\partial X$ and $\partial D^{\prime}$ are tight.
- If $X$ is incompressible then $D^{\prime}$ is not boundary compressible into $X$ and $d_{\mathcal{A}}\left(D, D^{\prime}\right) \leq 3$.
- If $X$ is compressible then $\partial D^{\prime} \subset X$ and $d_{\mathcal{A C}}\left(D, D^{\prime}\right) \leq 3$.

Here $\mathcal{A}=\mathcal{A}(X)$ and $\mathcal{A C}=\mathcal{A C}(X)$.
Proof. If $\partial D \subset X$ then the lemma is trivial. So assume, by Remark 8.1, that $D$ cuts $\partial X$. Choose a maximal sequence starting at $D$.

Assume first that the sequence is non-trivial $(n>1)$. By Lemma 10.3 there is a disk $E \in \Delta_{n}$ so that $D \cap X$ and $E \cap X$ contain disjoint arcs.

If the sequence ends essentially then choose $D^{\prime}=E$ and the lemma is proved. If the sequence ends in $X$ then there is a boundary compression of $\Delta_{n}$, disjoint from $\partial X$, yielding the desired disk $D^{\prime}$ with $\partial D^{\prime} \subset X$. Since $E \cap D^{\prime}=\emptyset$ we again obtain the desired bound.

Assume now that the sequence is trivial $(n=1)$. Then take $E=D \in$ $\Delta_{n}$ and the proof is identical to that of the previous paragraph.

Remark 10.5. At first sight Lemma 10.4 appears surprising; after all, any pair of curves in $\mathcal{C}(X)$ can be connected by a sequence of band sums. However, the sequences of band sums arising in Lemma 10.4 are quite special due to the fact that $D$ has bounded genus.

When $V$ has the structure of an $I$ bundle an important special case of Lemma 10.4 appears: Let $T \cong F \times I, X \cup Y=\partial_{h} T$, and $\rho_{F}: T \rightarrow F$ be the natural projection. If $\alpha, \beta$ are $\operatorname{arcs}$ in $\partial_{h} T$ define $d_{\mathcal{A}(F)}(\alpha, \beta)=$ $d_{\mathcal{A}(F)}\left(\rho_{F}(\alpha), \rho_{F}(\beta)\right)$.
Lemma 10.6. Suppose that $D$ is an essential disk in $T \cong F \times I$. Tighten $\partial D$ with respect to $\partial \partial_{h} T$. Then $d_{\mathcal{A}(F)}(D \cap X, D \cap Y) \leq 6$.
Proof. As in Lemma 10.4 obtain a disk $D^{\prime}$ which cannot be boundary compressed into $X$ or $Y$. As $D^{\prime}$ is $\partial_{X}$ incompressible Lemma 7.6 proves $D^{\prime}$ may be isotoped to be vertical with respect to $T$ (that is, $D^{\prime}$ is a union of fibres). Thus $D^{\prime} \cap X$ is a single arc as is $D^{\prime} \cap Y$. We find that with $d_{\mathcal{A}(X)}\left(D, D^{\prime}\right) \leq 3$ and $d_{\mathcal{A}(Y)}\left(D, D^{\prime}\right) \leq 3$ and the lemma follows.
10.2. Proving the theorem. It is now trival to prove Theorem 10.7.

We first recall the statement:
Theorem 10.7. Suppose $X$ is a compressible hole for $\mathcal{D}(V)$ with diameter at least 15. Then there are a pair of essential disks $D, E \subset V$ so that

- $\partial D, \partial E \subset X$ and
- $\partial D$ and $\partial E$ fill $X$.

Thus $X$ supports a pseudo-Anosov map.
Proof. Choose disks $D^{\prime}$ and $E^{\prime}$ in $\mathcal{D}(V)$ so that $d_{X}\left(D^{\prime}, E^{\prime}\right) \geq 15$. By Lemma 10.4 there are disks $D$ and $E$ so that $\partial D, \partial E \subset X, d_{X}\left(D^{\prime}, D\right) \leq$ 6 , and $d_{X}\left(E^{\prime}, E\right) \leq 6$. It follows from the triangle inequality that $d_{X}(D, E) \geq 3$ and the theorem is proved.

## 11. Holes for the disk complex - incompressible

This section classifies incompressible holes for the disk complex.
Theorem 11.1. Suppose $X$ is an incompressible hole for $\mathcal{D}(V)$ with diameter at least $4 \log _{2}(g)+60$. Then there is an $I$-bundle $\rho_{F}: T \rightarrow F$ embedded in $V$ so that

- $\partial_{h} T \subset \partial V$,
- $X$ is a component of $\partial_{h} T$,
- some component of $\partial_{v} T$ is boundary parallel into $\partial V$,
- F supports a pseudo-Anosov map.

Here is a short plan of the proof: We are given $X$, an incompressible hole for $\mathcal{D}(V)$. Following Lemma 10.4 we may assume that $D, E$ are essential disks, without boundary compressions into $X$ or $S \backslash X$, with $d_{X}(D, E)>4 \log _{2}(g)+48$. Examine the intersection pattern of $D$ and $E$ to find rectangles $R$ and $Q$. The intersection of these rectangle in $V$ will determine the desired $I$-bundle $T$. The third and fourth conclusions of the theorem follow from standard facts about primitive annuli. The fifth requires another application of Lemma 10.4 as well as Lemma 3.9.
11.1. Diagonals of polygons. Before beginning the proof of Theorem 11.1 we must briefly discuss diagonals of polygons.

Let $D$ be a $2 n$ sided regular polygon. Label the sides of $D$ with the letters $X$ and $Y$ in an alternating fashion. Any side labeled $X$ (or $Y$ ) will be called an $X$ side (or $Y$ side).
Definition 11.2. An arc $\gamma$ properly embedded in $D$ is a diagonal if the points of $\partial \gamma$ lie in the interiors of distinct sides of $D$. If $\gamma$ and $\gamma^{\prime}$
are diagonals for $D$ which together meet three different sides then we say that $\gamma$ and $\gamma^{\prime}$ are non-parallel.
Lemma 11.3. Suppose that $\Gamma \subset D$ is a collection of pairwise disjoint non-parallel diagonals. Then there is an $X$ side of $D$ meeting at most eight diagonals of $\Gamma$.

Proof. First, an easy counting argument shows that $|\Gamma| \leq 4 n-3$. Now, if every $X$ side meets at least nine non-parallel diagonals then $|\Gamma| \geq \frac{9}{2} n>4 n-3$, a contradiction.
11.2. Improving disks. Suppose now that $X$ is an incompressible hole for $\mathcal{D}(V)$ with diameter at least $4 \log _{2}(g)+60$. (From Theorem 9.1 it follows that $X$ is not an annulus.) Let $Y=\overline{S \backslash X}$.

Choose disks $D^{\prime}$ and $E^{\prime}$ in $V$ so that $d_{X}\left(D^{\prime}, E^{\prime}\right) \geq 4 \log _{2}(g)+60$. By Lemma 10.4 there are a pair of disks $D$ and $E$ so that both are essential in $V$, cannot be boundary compressed into $X$ or into $\overline{S \backslash X}$, and so that $d_{\mathcal{A}(X)}\left(D^{\prime}, D\right) \leq 3$ and $d_{\mathcal{A}(X)}\left(E^{\prime}, E\right) \leq 3$. Thus $d_{X}\left(D^{\prime}, D\right) \leq 6$ and $d_{X}\left(E^{\prime}, E\right) \leq 6$ (Lemma 4.6). By the triangle inequality $d_{X}(D, E) \geq$ $4 \log _{2}(g)+60-12=4 \log _{2}(g)+48$.

Recall, as well, that $\partial D$ and $\partial E$ are tight with respect to $\partial X$. We may further assume that $\partial D$ and $\partial E$ are tight with respect to each other. Also, minimize the quantities $|X \cap(\partial D \cap \partial E)|$ and $|D \cap E|$ while keeping everything tight. Now consider $D$ and $E$ to be evensided polygons, with vertices being the points $\partial D \cap \partial X$ and $\partial E \cap \partial X$ respectively. Now let $\Gamma=D \cap E$. See Figure 4 for one a priori possible collection $\Gamma \subset D$.


Figure 4. Can $\Gamma \subset D$ contain simple closed curves and non-diagonal arcs?

From our assumptions and the irreducibility of $V$ it follows that $\Gamma$ contains no simple closed curves. Suppose now that there is an $\gamma \subset \Gamma$ so that, in $D$, both endpoints of $\gamma$ lie in the same side of $D$ - so $\gamma$ is not a diagonal for $D$. Then there is an outermost such arc, say
$\gamma^{\prime} \subset \Gamma$, cutting a bigon $B$ out of $D$. It follows that $B$ is a boundary compression of $E$ which is disjoint from $\partial X$. But this contradicts the construction of $E$. We deduce that all arcs of $\Gamma$ are diagonals for $D$ and, via a similar argument, for $E$.

Let $\alpha \subset D \cap X$ be an $X$ side of $D$ meeting at most eight distinct types of diagonal of $\Gamma$. Choose $\beta \subset E \cap X$ similarly. As $d_{X}(D, E) \geq$ $4 \log _{2}(g)+48$ we have that $d_{X}(\alpha, \beta) \geq 4 \log _{2}(g)+48-4=4 \log _{2}(g)+44$. As $d_{X}(\alpha, \beta)=d_{X}\left(\sigma_{X}(\alpha), \sigma_{X}(\beta)\right)$ and $\iota\left(\sigma_{X}(\alpha), \sigma_{X}(\beta)\right) \leq 4 \iota(\alpha, \beta)+4$ it follows from Lemma 3.4 that $\iota(\alpha, \beta) \geq 2^{\frac{4 \log _{2}(g)+44-6}{2}}-1>2^{13}$

Now break $\alpha$ and $\beta$ into subarcs, each subarc meeting all of the diagonals of fixed type and only meeting the diagonals of that type. There is a pair of such subarcs $\alpha^{\prime} \subset \alpha$ and $\beta^{\prime} \subset \beta$ with the following property:

$$
\left|\alpha^{\prime} \cap \beta^{\prime}\right| \geq \frac{\iota(\alpha, \beta)}{64}
$$

Note that it follows that $\left|\alpha^{\prime} \cap \beta^{\prime}\right| \geq 128$.
Claim 11.4. The graph $\Theta^{\prime}=\alpha^{\prime} \cup \beta^{\prime}$ fills $X$.
Proof. Suppose not. Let $W \subset X$ be the surface obtained by taking a regular neighborhood of $\Theta^{\prime}$ union with all disks and boundary parallel annuli (in $X$ ) of $X \backslash \Theta^{\prime}$. Thus $W$ is an essential subsurface strictly contained in $X$. We now assume, breaking symmetry, that $d_{X}(\alpha, \partial W) \geq \frac{4 \log _{2}(g)+44}{2}=2 \log _{2}(g)+22$.

As $\alpha$ and $\beta$ are tight in $X$ it follows that both $\alpha$ and $\beta$ are tight with respect to $\partial W$. Let $A=\left\{\alpha_{i}\right\}_{i=0}^{n}=\alpha \cap W$ be the collection of subarcs of $\alpha$ meeting $W$, with $\alpha^{\prime} \subset \alpha_{0}$. Partition $A$ into subsets $A_{k}$ so that

- $\alpha_{0} \in A_{0}$ and
- $\alpha_{i} \in A_{h}, \alpha_{j} \in A_{k}$ are properly isotopic if and only if $h=k$.

Let $\mathcal{A}$ be the collection of partitions $A_{k}$ so that $\left.\left|A_{k}\right| \geq 65\right\}$.
Subclaim. Cutting the surface $W$ along $\mathcal{A}$ gives a collection of disks.
Suppose, for a contradiction, that $W^{\prime} \subset W \backslash \mathcal{A}$ is not a disk. Deduce that if $W$ is not an annulus there is an essential curve $\gamma \subset W^{\prime}$ with $|\gamma \cap \alpha| \leq 64 \cdot(-3 \chi(W))$. This is because $\alpha \cap W$ yields at most $-3 \chi(W)$ different proper isotopy classes of arcs. If $W$ is an annulus (and so $\left.W^{\prime}=W\right)$ then $|\gamma \cap \alpha| \leq 64$.

In either case $|\gamma \cap \alpha| \leq 64 \cdot 6 g \leq 512 g$ where $g$ is the genus of the handlebody $V$. By Lemma 3.4 deduce that $d_{X}(\gamma, \alpha) \leq 2\left(9+\log _{2}(g)\right)+$ $2=2 \log _{2}(g)+20$. As $\gamma \subset W^{\prime} \subset W$ it follows that $d_{X}(\gamma, \partial W) \leq 1$ and so $d_{X}(\alpha, \partial W) \leq 2 \log _{2}(g)+21$. This contradicts our assumption that $d_{X}(\alpha, \partial W) \geq 2 \log _{2}(g)+22$.

On the other hand:
Subclaim. $A_{0}$ contains at most 64 arcs.
If $A_{0}$ contains 65 arcs or more then $|\alpha \cap \beta|>64 \cdot\left|\alpha^{\prime} \cap \beta^{\prime}\right| \geq \iota(\alpha, \beta)$, a contradiction.

So let $W_{0}$ be the disk component of $W \backslash \mathcal{A}$ containing $\alpha_{0}=\alpha^{\prime}$. Note that $\alpha_{0}$ separates $W_{0}$. Thus every time $\beta^{\prime}$ crosses $\alpha_{0}$, except for the last, $\beta^{\prime}$ must also cross some partition $A_{k} \in \mathcal{A}$. We deduce that:

$$
\iota(\alpha, \beta) \geq\left|\alpha \cap \beta^{\prime}\right| \geq 65\left(\left|\alpha^{\prime} \cap \beta^{\prime}\right|-1\right)>64 \cdot\left|\alpha^{\prime} \cap \beta^{\prime}\right| \geq \iota(\alpha, \beta)
$$

(The strict inequality follows from the fact that $\left|\alpha^{\prime} \cap \beta^{\prime}\right| \geq 128$.) As this is a contradiction, Claim 11.4 is proven.

Let $R \subset D$ be the rectangle with top side equal to $\alpha^{\prime}$, left and right sides equal to diagonals $D \cap E$ of $D$, and bottom side in $\partial D \backslash \partial X$ parallel (along the diagonals) to $\alpha^{\prime}$. Denote this bottom side of $R$ by $\alpha^{\prime \prime}$. See Figure 5. We define $Q \subset E$ and $\beta^{\prime \prime} \subset \partial Q$ similarly.


Figure 5. The rectangle $R \subset D$ is surrounded by the dotted line. The arc $\alpha^{\prime}$ in $\partial D \cap X$ is indicated. In general the arc $\alpha^{\prime \prime}$ may lie in $X$ or in $Y$.

As above we have $\Theta^{\prime}=\alpha^{\prime} \cup \beta^{\prime}$. Let $\Theta^{\prime \prime}=\alpha^{\prime \prime} \cup \beta^{\prime \prime}$. There are now two possibilities: either $\Theta^{\prime}$ and $\Theta^{\prime \prime}$ are disjoint or they are not. In the first case let $\Theta=\Theta^{\prime}$. In the second case let $\Theta=\Theta^{\prime} \cup \Theta^{\prime \prime}$. In either case notice that, by Claim 11.4, the graph $\Theta$ is connected and fills $X$.

The following claim will be useful in the next section:
Claim 11.5. The graph $\Theta^{\prime}$ is not contained in any disk $C$ embedded in $\partial V$. The same holds for $\Theta^{\prime \prime}$.

Proof. This is clear for $\Theta^{\prime}$ because $\Theta^{\prime}$ fills $X$, an essential subsurface of $\partial V$. Now suppose that $\Theta^{\prime \prime} \subset C$, a disk. Note that $\left|\alpha^{\prime \prime} \cap \beta^{\prime \prime}\right|=\left|\alpha^{\prime} \cap \beta^{\prime}\right|>$ 1. We are supposing that all intersections of $\alpha^{\prime \prime}$ and $\beta^{\prime \prime}$ occur in $C$. It follows that $\partial D$ and $\partial E$ are not tight on $\partial V$, a contradiction.
11.3. Building the I-bundle. As in the previous section we are given a pair of rectangles $R \subset D$ and $Q \subset E$ so that $R \cap \partial V=\alpha^{\prime} \cup \alpha^{\prime \prime}$ and $Q \cap \partial V=\beta^{\prime} \cup \beta^{\prime \prime}$. We also have a connected graph $\Theta$ which fills $X$. Note that $R \cup Q$ is an $I$-bundle and $\Theta$ is the component of its horizontal boundary meeting $X$. See Figure 6.


Figure 6. $R \cup Q$ is an $I$-bundle because all arcs of intersection are parallel.

Let $T_{0}$ be a regular neighborhood of $R \cup Q$, taken in $V$. This has the structure of an $I$-bundle. Note that $\partial_{h} T_{0} \subset \partial V, \partial_{h} T_{0} \cap X$ is a component of $\partial_{h} T_{0}$, and this component fills $X$ due to Claim 11.4. We wish to enlarge $T_{0}$ to obtain the correct $I$-bundle in $V$.

Begin by enumerating all annuli $\left\{A_{i}\right\} \subset \partial_{v} T_{0}$ with the property that some component of $\partial A_{i}$ is inessential in $\partial V$. Suppose that we have built the $I$-bundle $T_{i}$ and are now considering the annulus $A=A_{i}$. Let $\gamma \cup \gamma^{\prime}=\partial A \subset \partial V$ with $\gamma$ inessential in $\partial V$. Let $B \subset \partial V$ be the disk which $\gamma$ bounds. By induction we assume that no component of $\partial_{h} T_{i}$ is contained in a disk embedded in $\partial V$ (the base case holds by Claim 11.5). It follows that $B \cap T_{i}=\partial B=\gamma$. Thus $B \cup A$ is isotopic, rel $\gamma^{\prime}$, to be a properly embedded disk $B^{\prime} \subset V$. As $\gamma^{\prime}$ lies in $X$ or $\overline{\partial V \backslash X}$, both incompressible, $\gamma^{\prime}$ must bound a disk $C \subset \partial V$. Note that $C \cap T_{i}=\partial C=\gamma^{\prime}$, again using the induction hypothesis.

It follows that $B \cup A \cup C$ is an embedded two-sphere in $V$. As $V$ is a handlebody $V$ is irreducible. Thus $B \cup A \cup C$ bounds a three-ball $U_{i}$ in $V$. Choose a homeomorphism $U_{i} \cong B \times I$ so that $B$ is identified with $B \times\{0\}, C$ is identified with $B \times\{1\}$, and $A$ is identified with $\partial B \times I$. We form $T_{i+1}=T_{i} \cup U_{i}$ and note that $T_{i+1}$ still has the structure of an $I$-bundle. Recalling that $A=A_{i}$ we have $\partial_{v} T_{i+1}=\partial_{v} T_{i} \backslash A_{i}$. Also $\partial_{h} T_{i+1}=\partial_{h} T_{i} \cup(B \cup C) \subset \partial V$. It follows that no component of $\partial_{h} T_{i+1}$
is contained in a disk embedded in $\partial V$. Similarly, $\partial_{h} T_{i+1} \cap X$ is a component of $\partial_{h} T_{i+1}$ and this component fills $X$.

After dealing with all of the annuli $\left\{A_{i}\right\}$ in this fashion we are left with an $I$-bundle $T$. Now all components of $\partial \partial_{v} T$ [sic] are essential in $\partial V$. All of these lying in $X$ are peripheral in $X$. This is because they are disjoint from $\Theta \subset \partial_{h} T$, which fills $X$, by induction. It follows that the component of $\partial_{h} T$ containing $\Theta$ is isotopic to $X$.

This finishes the construction of the promised $I$-bundle $T$ and demonstrates the first two conclusions of Theorem 11.1. For future use we record:
Remark 11.6. Every curve of $\partial \partial_{v} T=\partial \partial_{h} T$ is essential in $S=\partial V$.
11.4. A vertical annulus parallel into the boundary. Here we obtain the third conclusion of Theorem 11.1: at least one component of $\partial_{v} T$ is boundary parallel in $\partial V$.

Fix $T$ an $I$-bundle with the incompressible hole $X$ a component of $\partial_{h} T$.
Claim 11.7. All components of $\partial_{v} T$ are incompressible in $V$.
Proof. Suppose that $A \subset \partial_{v} T$ was compressible. By Remark 11.6 we may compress $A$ to obtain a pair of essential disks $B$ and $C$. Note that $\partial B$ is isotopic into the complement of $\partial_{h} T$. So $\overline{S \backslash X}$ is compressible, contradicting Remark 8.1.

Claim 11.8. Some component of $\partial_{v} T$ is boundary parallel.
Proof. Since $\partial_{v} T$ is incompressible (Claim 11.7) by Remark 7.2, we find that $\partial_{v} T$ is boundary compressible in $V$. Let $B$ be a boundary compression for $\partial_{v} T$. Let $A$ be the component of $\partial_{v} T$ meeting $B$. Let $\alpha$ denote the arc $A \cap B$.

The arc $\alpha$ is either essential or inessential in $A$. Suppose $\alpha$ is inessential in $A$. Then $\alpha$ cuts a bigon, $C$, out of $A$. Since $B$ was a boundary compression the disk $D=B \cup C$ is essential in $V$. Since $B$ meets $\partial_{v} T$ in a single arc, either $D \subset T$ or $D \subset \overline{V \backslash T}$. The former implies that $\partial_{h} T$ is compressible and the latter that $X$ is not a hole. Either gives a contradiction.

It follows that $\alpha$ is essential in $A$. Now carefully boundary compress $A$ : Let $N$ be the closure of a regular neighborhood of $B$, taken in $V \backslash A$. Let $A^{\prime}$ be the closure of $A \backslash N$ (so $A^{\prime}$ is a rectangle). Let $B^{\prime} \cup B^{\prime \prime}$ be the closure of $\operatorname{fr}(N) \backslash A$. Both $B^{\prime}$ and $B^{\prime \prime}$ are bigons, parallel to $B$. Form $D=A^{\prime} \cup B^{\prime} \cup B^{\prime \prime}$ : a properly embedded disk in $V$. If $D$ is essential then, as above, either $D \subset T$ or $D \subset \overline{V \backslash T}$. Again, either gives a contradiction.

It follows that $D$ is inessential in $V$. Thus $D$ cuts a closed three-ball $U$ out of $V$. There are two final cases: either $N \subset U$ or $N \cap U=B^{\prime} \cup B^{\prime \prime}$. If $U$ contains $N$ then $U$ contains $A$. Thus $\partial A$ is contained in the disk $U \cap \partial V$. This contradicts Remark 11.6. Deduce instead that $W=U \cup N$ is a solid torus with meridional disk $B$. Thus $W$ gives a parallelism between $A$ and the annulus $\partial V \cap \partial W$, as desired.

Remark 11.9. Similar considerations prove that the multicurve
$\left\{\partial A \mid A\right.$ is a boundary parallel component of $\left.\partial_{v} T\right\}$
is disk busting for $V$.
11.5. Finding a pseudo-Anosov map. Here we prove that the base surface $F$ of the $I$-bundle $T$ admits a pseudo-Anosov map.

As in Section 11.2, pick essential disks $D^{\prime}$ and $E^{\prime}$ essential disks in $V$ so that $d_{X}\left(D^{\prime}, E^{\prime}\right) \geq 4 \log _{2}(g)+60$. Lemma 10.4 provides disks $D$ and $E$ which cannot be boundary compressed into $X$ or into $\overline{S \backslash X}$ - thus $D$ and $E$ cannot be boundary compressed into $\partial_{h} T$. Also, as above, $d_{X}(D, E) \geq 4 \log _{2}(g)+60-12=4 \log _{2}(g)+48$.

After isotoping $D$ to minimize intersection with $\partial_{v} T$ it must be the case that all components of $D \cap \partial_{v} T$ are essential arcs in $\partial_{v} T$. By Lemma 7.6 we conclude that $D$ may be isotoped in $V$ so that $D \cap T$ is vertical in $T$. The same holds of $E$. Choose $A$ and $B$, components of $D \cap T$ and $E \cap T$. Each are vertical rectangles. Note that we still have $d_{X}(A, B) \geq 4 \log _{2}(g)+48$.

We now begin to work in the base surface $F$. Recall that $\rho_{F}: T \rightarrow F$ is an $I$-bundle. Take $\alpha=\rho_{F}(A)$ and $\beta=\rho_{F}(B)$. Note that the natural $\operatorname{map} \mathcal{C}(F) \rightarrow \mathcal{C}(X)$, defined by taking a curve to its lift, is distance non-increasing (see Equation 5.5). Thus $d_{F}(\alpha, \beta) \geq 4 \log _{2}(g)+48$. By Theorem 9.1 the surface $F$ cannot be an annulus. Thus, by Lemma 3.9 the subsurface $F$ supports a pseudo-Anosov map and we are done.
11.6. Corollaries. We now deal with the possibility of disjoint holes for the disk complex.
Lemma 11.10. Suppose that $X$ is incompressible hole for $\mathcal{D}(V)$ supported by the $I$-bundle $\rho_{F}: T \rightarrow F$. Suppose that $D \subset V$ is an essential disk.

- If $F$ is orientable then set $X \cup Y=\partial_{h} T$. Then $d_{\mathcal{A}(F)}(D \cap X, D \cap$ $Y) \leq 5$.
- If $F$ is nonorientable then set $X=\partial_{h} T$. Let $\tau: X \rightarrow X$ be the associated involution. Then $d_{X}\left(D, \mathcal{C}^{\tau}(X)\right) \leq 6$.

Proof. By Lemma 10.4 there is a disk $D^{\prime} \subset V$ which is tight with respect to $\partial_{h} T$ and which cannot be boundary compressed into $\partial_{h} T$ (or into the complement). Also, for any component $X \subset \partial_{h} T$ we have $d_{\mathcal{A}(X)}\left(D, D^{\prime}\right) \leq 3$.

Properly isotope $D^{\prime}$ to minimize $D^{\prime} \cap \partial_{v} T$. Then $D^{\prime} \cap \partial_{v} T$ is properly isotopic, in $\partial_{v} T$, to a collection of vertical arcs. Let $E=D^{\prime} \cap T$. Deduce that the components of $E$ are, after a proper isotopy in $T$, vertical with respect to $T$.

Pick any component $E^{\prime}$ of $E$ and let $\alpha \subset F$ be the arc of intersection between $E^{\prime}$ and $F$. Let $\alpha^{\prime}=E^{\prime} \cap \partial_{h} T$. This is, after a proper isotopy of $D^{\prime}$, a subarc of $D^{\prime} \cap \partial_{h} T$. Note that $\alpha^{\prime}$ is symmetric under the involution $\tau: \partial_{h} T \rightarrow \partial_{h} T$. The conclusion follows.

Recall Lemma 6.3: all holes for the arc complex intersect. This cannot hold for the disk complex. For example if $\rho_{F}: T \rightarrow F$ is an $I$ bundle over an orientable surface then take $V=T$ and notice that both components of $\partial_{h} T$ are holes for $\mathcal{D}(V)$. However, by the first conclusion of Lemma 11.10, $X$ and $Y$ are paired holes, in the sense of Definition 4.5. So, as with the invariant arc complex (Lemma 6.5), all holes for the disk complex interfere:

Lemma 11.11. Suppose that $X, Z \subset \partial V$ are holes for $\mathcal{D}(V)$, both of infinite diameter. If $X \cap Z=\emptyset$ then there is an $I$-bundle $T \cong F \times I$ in $V$ so that $\partial_{h} T=X \cup Y$ and $Y \cap Z \neq \emptyset$.

Proof. Suppose that $X \cap Z=\emptyset$. It follows from Remark 8.1 that both $X$ and $Z$ are incompressible. Let $\rho_{F}: T \rightarrow F$ be the $I$-bundle in $V$ with $X \subset \partial_{h} T$, as provided by Theorem 11.1. We also have a component $A \subset \partial_{v} T$ so that $A$ is boundary parallel. Let $U$ be the solid torus component of $V \backslash A$.

Let $\alpha=\rho_{F}(A)$. Choose any essential $\operatorname{arc} \delta \subset F$ with both endpoints in $\alpha \subset \partial F$. It follows that $\rho_{F}^{-1}(\delta)$, together with two meridional disks of $U$, forms an essential disk $D$ in $V$.

Now, if $F$ is nonorientable then $\partial D \subset X \cup(\partial U \backslash A)$. Also, $Z$ cannot meet $U$ (because $Z \cap X=\emptyset$ ) and it follows that $D \cap Z=\emptyset$, a contradiction.

Deduce that $F$ is orientable. Let $X \cup Y=\partial_{h} T$. So $\partial D \subset(X \cup$ $(\partial U \backslash A) \cup Y)$. Now, if $Z$ misses $Y$ then $Z \subset \partial U \backslash A$. In this case, $Z$ is itself an annulus, contradicting Theorem 9.1. It follows that $Z$ cuts $Y$ and we are done.
12. Paths in general and The upper bound on distance

Needs work.
Suppose that $\mathcal{G}=\mathcal{G}(S)$ is a combinatorial complex. The goal of the section is to give conditions on $\mathcal{G}$, the path requirements, which imply that the distance $d_{\mathcal{G}}$ is bounded above by the sum of projections to holes. We begin by stating the target theorem.
Theorem 12.1. Fix $S$ a compact connected non-simple surface. Suppose that $\mathcal{G}(S)$ is a combinatorial complex satisfying the path requirements. Then there is a constant $c_{0}=c_{0}(S)$ so that for all $c \geq c_{0}$ there are constants $a \geq 1, b \geq 0$ satisfying

$$
d_{\mathcal{G}}(\alpha, \beta) \leq_{a, b} \sum\left[d_{X}(\alpha, \beta)\right]_{c} .
$$

Here $\alpha, \beta \in \mathcal{G}$ and the sum is taken over all holes $X$ for $\mathcal{G}$.
The proof is more difficult than that of the lower bound, Theorem 4.10. Essentially, we must build a path in $\mathcal{G}$ which is not too long. We will be given a path $\Lambda=\left\{\mu_{n}\right\}$ of markings (defined below) which is locally an unparametrized quasi-geodesic: when $X$ is a hole for $\mathcal{G}$ the path $\pi_{X}(\Lambda)$ will be an unparametrized quasi-geodesic.

We will also be given a combinatorial path $\Gamma=\left\{\gamma_{i}\right\}$ lying in $\mathcal{G}$. A reindexing function ties $\Gamma$ to $\Lambda$. It will follow that $\Gamma$ is also locally an unparametrized quasi-geodesic, with respect to the holes for $\mathcal{G}$. Unfortunately $\Gamma$ is almost surely too long to give the upper bound; it may spend time traveling through $\mathcal{C}(Y)$ where $Y$ is not a hole.

Our goal is to partition $\Gamma$ into basic intervals where the path $\Gamma$ makes definite progress through $\mathcal{C}(S)$, shortcut intervals where $\Gamma$ wastes time in $\mathcal{C}(Y)$ for some non-hole $Y \subset S$, and inductive intervals where $\Gamma$ enters and exits $\mathcal{C}(X)$ for some strict hole $X \subset S$. We do not alter $\Gamma$ inside of the basic intervals, we shorten $\Gamma$ so that it only spends a constant amount of time in non-holes, and we call on induction to cope with strict holes. A similar scheme then gives an upper bound on the length of the shortened $\Gamma$. Theorem 12.1 then follows. We begin by setting out the path requirements.
12.1. The requirements. We are given the following data: $\mathcal{G}=\mathcal{G}(S)$, a combinatorial complex, as well as vertices $\alpha, \beta \in \mathcal{G}$. We are also given $\Lambda=\left\{\mu_{n}\right\}_{n=0}^{N}$ a collection of markings. There is also a sufficiently large constant $c_{0}$, depending only on the topology of $S$.
Definition 12.2. Recall that a marking $\mu=\left\{\nu_{i}\right\}$ is a set of curves in $S$ with the following properties:

- $\iota\left(\nu_{i}, \nu_{j}\right) \leq c_{0}$ and
- the $\nu_{i}$ fill $S$.


## Requirements for the marking path.

(1) $\iota\left(\alpha, \mu_{0}\right), \iota\left(\beta, \mu_{N}\right) \leq c_{0}$.
(2) For any hole $X$ for $\mathcal{G}$, the map $n \mapsto \pi_{X}\left(\mu_{n}\right)$ is an unparametrized quasi-geodesic with uniform constants.
(3) $\Lambda$ locally satisfies the reverse triangle inequality: For any subsurface $X$, and for any $m<n<p$, we have

$$
d_{X}\left(\mu_{m}, \mu_{p}\right) \geq d_{X}\left(\mu_{m}, \mu_{n}\right)+d_{X}\left(\mu_{n}, \mu_{p}\right)-c_{0}
$$

From the marking path $\left\{\mu_{n}\right\}$ we will extract a combinatorial path $\Gamma=\left\{\gamma_{i}\right\}_{i=0}^{K}$ consisting of vertices in $\mathcal{G}$. We also will have a reindexing function $r:[0, K] \rightarrow[0, N]$. There is a uniform constant $c_{1} \in \mathbb{R}$ with the following properties:

## Requirements for the combinatorial path.

(1) $\alpha=\gamma_{0}$ and $\beta=\gamma_{K}$.
(2) The reindexing map is strictly increasing.
(3) $\iota\left(\gamma_{i}, \mu_{r(i)}\right) \leq c_{1}$, for every $i \in[0, K]$.
(4) $\iota\left(\gamma_{i}, \gamma_{i+1}\right) \leq c_{1}$, for every $i \in[0, K]$.

Finally we will be given, for every essential subsurface $X \subset S$, a (possibly empty) interval $J_{X} \subset[0, N]$. For every $n \in J_{X}$ we say that $X$ is accessible from the marking $\mu_{n}$. There are uniform constants $c_{2}, c_{3} \in \mathbb{R}$ with the following properties:

## Requirements for accessibility.

(1) If $d_{X}(\alpha, \beta)>c_{2}$ then $J_{X}$ is nonempty.
(2) If $m, n \leq \min J_{X}$ or if $m, n \geq \max J_{X}$ then $d_{X}\left(\mu_{m}, \mu_{n}\right) \leq c_{3}$.
(3) If $X$ is nested in $Y$ then for any $m, n \in J_{X}$ we have $d_{Y}\left(\mu_{m}, \mu_{n}\right) \leq$ $c_{3}$.
(4) If $X$ and $Y$ overlap then for any $m, n \in J_{X} \cap J_{Y}$ we have $d_{X}\left(\mu_{m}, \mu_{n}\right), d_{Y}\left(\mu_{m}, \mu_{n}\right) \leq c_{3}$.
12.2. Basic, inductive, and shortcut intervals. Our goal is to partition the indices of $\Gamma=\left\{\gamma_{i}\right\}_{i=0}^{K}$ into basic, inductive, and shortcut intervals. We first identify the relevant surfaces. Take $L_{0} \geq c_{3}$ and sufficiently large. We are given $\alpha, \beta \in \mathcal{G}$ and a combinatorial path $\Gamma=\left\{\gamma_{i}\right\}_{i=0}^{K}$. Set $\alpha_{S}=\alpha$ and $\beta_{S}=\beta$. Define

$$
B_{S}=\left\{Y \subsetneq S \mid d_{Y}\left(\alpha_{S}, \beta_{S}\right) \geq 3 L_{0}\right\}
$$

In general we are given a hole $X \subset S$, an interval $I_{X}=[h, k] \subset[0, K]$, and a set of subsurfaces $B_{X}$. Set $\alpha_{X}=\gamma_{h}$ and $\beta_{X}=\gamma_{k}$. We are told that:

- If $X \neq S, d_{X}\left(\alpha_{X}, \beta_{X}\right) \geq L_{0}$.
- For all $Y \in B_{X}, Y \subsetneq X$ and $d_{Y}\left(\alpha_{X}, \beta_{X}\right) \geq 3 L_{0}$.
- For all $Y \notin B_{X}$, if $Y \subsetneq X$ and $Y$ is a hole then $d_{Y}\left(\alpha_{X}, \beta_{X}\right) \leq 5 L_{0}$.

For any subinterval $[i, j] \subset I_{X}$ we define

$$
B_{X}(i, j)=\left\{Y \in B_{X} \mid d_{Y}\left(\gamma_{i}, \gamma_{j}\right) \geq 3 L_{0}\right\}
$$

We may now define basic, shortcut, and inductive subintervals of $I_{X}$. Definition 12.3. If $B_{X}(i, j)$ is empty then $[i, j]$ is a basic interval. We require:

$$
\begin{equation*}
d_{\mathcal{G}}\left(\gamma_{i}, \gamma_{j}\right) \leq_{a, b} d_{X}\left(\gamma_{i}, \gamma_{j}\right) \tag{12.4}
\end{equation*}
$$

Definition 12.5. Suppose $Y \in B_{X}$ is a non-hole. If $r([i, j]) \subset J_{Y}$ then $[i, j]$ is a shortcut interval associated to $Y$. We require a constant $L_{1}=L_{1}(X)$ so that:

$$
\begin{equation*}
d_{\mathcal{G}}\left(\gamma_{i}, \gamma_{j}\right) \leq_{a, b} L_{1} \tag{12.6}
\end{equation*}
$$

Definition 12.7. Suppose $Y \in B_{X}$ is a hole. If $r([i, j]) \subset J_{Y}$ and $d_{Y}\left(\gamma_{i}, \gamma_{j}\right) \geq L_{0}$ then $[i, j]$ is an inductive interval associated to $Y$. The induction hypothesis is:

$$
\begin{equation*}
d_{\mathcal{G}}\left(\gamma_{i}, \gamma_{j}\right) \leq_{a, b} \sum\left[d_{Z}\left(\gamma_{i}, \gamma_{j}\right)\right]_{c} \tag{12.8}
\end{equation*}
$$

where the sum is over all holes $Z$ for $\mathcal{G}(S)$.
If $[i, j]$ is inductive and associated to $Y$ then we may write $I_{Y}=[i, j]$ and set $\alpha_{Y}=\gamma_{i}, \beta_{Y}=\gamma_{j}$. In addition we may define the set of subsurfaces:

$$
B_{Y}=\left\{Z \in B_{X}(i, j) \mid Z \subsetneq Y\right\}
$$

Lemma 12.9. For $L_{0}$ sufficiently large, the following holds. Suppose $I_{Y}=[i, j]$ is inductive. If $Z \subsetneq Y, Z$ is a hole, and $Z \notin B_{Y}$ then $d_{Z}\left(\alpha_{Y}, \beta_{Y}\right) \leq 5 L_{0}$.

Proof. If $d_{Z}\left(\alpha_{Y}, \beta_{Y}\right)<3 L_{0}$ then there is nothing to prove. Supposing the contrary, since $Z \subsetneq Y \subsetneq X$ and $Z \notin B_{Y}$, we have $d_{Z}\left(\alpha_{X}, \beta_{X}\right)<3 L_{0}$. As the path $\pi_{Z}\left(\mu_{n}\right)$ is an unparametrized quasi-geodesic in $\mathcal{C}(Z)$, the same holds for the path $\pi_{Z}\left(\gamma_{k}\right)$, because $Z$ is a hole. Thus for $L_{0}$ sufficiently large, $\pi_{Z}\left(\alpha_{Y}\right)$ and $\pi_{Z}\left(\beta_{Y}\right)$ lie within $L_{0}$ of the geodesic between $\pi_{Z}\left(\alpha_{X}\right)$ and $\pi_{Z}\left(\beta_{X}\right)$. The conclusion follows from the triangle inequality.
12.3. The partition. We now partition the interval $I_{X}$. In general, we are given a subinterval $[i, j] \subset I_{X}$. (To begin with, $[i, j]=I_{X}$.)

If $B_{X}(i, j)$ is empty then $[i, j]$ is basic and we are done. Suppose then that $B_{X}(i, j)$ is nonempty. If there is a hole $Y \in B_{X}(i, j)$ then choose that hole $Y \in B_{X}(i, j)$ so that

- $Y$ is maximal with respect to the partial order (by inclusion) on $B_{X}(i, j)$
- $J_{Y} \cap r([i, j])$ is as large as possible. (Subject to the first condition.)
If there are no holes in $B_{X}(i, j)$ then instead take $Y$ to be a non-hole satisfying these same two conditions as above.

In either case let $i^{\prime}, j^{\prime} \in[i, j]$ be the first and last indices, respectively, so that $r\left(i^{\prime}\right), r\left(j^{\prime}\right) \in J_{Y}$. Partition $[i, j]=\left[i, i^{\prime}-1\right] \cup\left[i^{\prime}, j^{\prime}\right] \cup\left[j^{\prime}+1, j\right]$. (The first or third parts, or both, may be empty). Take $I_{Y}=\left[i^{\prime}, j^{\prime}\right]$ and set $\alpha_{Y}=\gamma_{i^{\prime}}, \beta_{Y}=\gamma_{j^{\prime}}$.

We will show later that if $Y$ is a non-hole then $I_{Y}$ is a shortcut interval. If $Y$ is a hole, then set

$$
B_{Y}=\left\{Z \in B_{X}\left(i^{\prime}, j^{\prime}\right) \mid Z \subsetneq Y\right\}
$$

We must check that $I_{Y}$ is an inductive interval - the only nontrivial point is that $d_{Y}\left(\gamma_{i^{\prime}}, \gamma_{j^{\prime}}\right) \geq L_{0}$. But this follows immediately from (2) of accessibility requirements and the triangle inequality.

Now partition $\left[i, i^{\prime}-1\right]$ and $\left[j^{\prime}+1, j\right]$ in similar fashion. This completes our recursive description to the partition of $I_{X}$.

We establish preliminary results about the partition.
Lemma 12.10. If $Y, Z$ are holes that either overlap or are nested, then $I_{Z} \subset J_{Y}$ implies $Z \subset Y$.

Proof. This follows immediately from the definition of the intervals $I_{Z}$, the assumption $M_{0} \geq C_{3}$, and the accessibility requirement (2) for nested subsurfaces and requirement (3) for overlapping surfaces.

Proposition 12.11. For $L_{0}$ large enough, the partition of $[m, n]$ has the properties that
(1) For any hole $Y$ there is at most 1 inductive interval $I_{Y}$ associated to $Y$.
(2) For any hole $Y$ there are at most 2 inductive intervals $I_{Z}$ such that $J_{Y} \cap I_{Z} \neq \emptyset$, such that $Y \subsetneq Z$, or $Z$ and $Y$ overlap. There is at most one interval $I_{Z}$ for $Z$ paired with $Y$.
(3) Suppose there exists $l \geq N_{1}$, intervals $I_{Y_{i}}$ that are shortcut intervals associated to a nonholes $Y_{i}$ or inductive intervals associated to holes $Y_{i}$. Then $\left\{\partial Y_{i}\right\}$ fill $X$, and $d_{X}\left(\partial Y_{1}, \partial Y_{l}\right) \geq l / N_{1}-2$. In particular, there are at most $N_{1}$ shortcut intervals $I_{Y}$ associated to any fixed $Y$.

Proof. We prove the first conclusion. If there were two intervals $I_{1}=$ $\left[p_{1}, q_{1}\right]$ and $I_{2}=\left[p_{2}, q_{2}\right]$ associated to the same hole $Y$ with $q_{1} \leq p_{2}$,
then we claim that

$$
d_{Y}\left(\mu_{q_{2}}, \mu_{p_{1}}\right) \geq L_{0}
$$

If that is the case, then we would have a contradiction, for then an interval, either $\left[p_{1}, q_{2}\right]$ or an interval containing it, would have been chosen as the inductive interval rather than either $\left[p_{i}, q_{i}\right]$. We need to prove the claim. Let $\alpha_{Y}=\pi_{Y}\left(\gamma_{p_{1}}\right)$ and $\beta_{Y}=\pi_{Y}\left(\gamma_{q_{2}}\right)$. We notice that since the path $\pi_{Y}\left(\mu_{n}\right)$ is an unparametrized quasi-geodesic, and $\mathcal{C}(Y)$ is a hyperbolic space, there is a constant $D$ depending on these constants such that for any two points on the path, any point between them is within $D$ of the geodesic joining the endpoints. If $q_{1}=p_{2}$ we apply it to the endpoints $\alpha_{Y}$ and $\beta_{Y}$ with $\pi_{Y}\left(\gamma_{q_{1}}\right)$ between them to conclude that

$$
\left.d_{Y}\left(\beta_{Y}, \alpha_{Y}\right)\right) \geq 2 L_{0}-2 D \geq L_{0}
$$

for $L_{0}$ large enough, proving the claim. If $p_{2}>q_{1}$, then dividing the interval $\left[p_{1}, q_{2}\right]$ into $\left[p_{1}, p_{2}\right]$ with interior point $q_{1}$, and $\left[p_{2}, q_{2}\right]$, we conclude that

$$
d_{Y}\left(\mu_{q_{2}}, \mu_{p_{1}}\right) \geq 2 L_{0}-3 D \geq L_{0}
$$

again for $L_{0}$ big enough, proving the claim.
We prove the second conclusion. The interval $J_{Y}$ may be contained in a single $I_{W}$ for $W$ paired with $Y$. If $J_{Y}$ intersected three or more $I_{Z}$ such that either $Y$ and $Z$ overlap, or $Y \subset Z$, then $J_{Y}$ would contain one of them. This is impossible by Lemma 12.10. Thus it intersects at most 2 such $I_{Z}$.

We prove the third statement. If there were more than $N_{1}$ such $I_{Y_{i}}$, then Lemma ?? says that some hole or nonhole $Z$ containing a collection of these $Y_{i}$ would satisfy $d_{Z}\left(\mu_{m_{1}}, \mu_{n_{k}}\right) \geq 3 M_{0}$, where $\mu_{m_{1}}$ is the furthest point to the left, and $\mu_{n_{k}}$ the furthest to the right. Suppose $Z$ is a proper subsurface of $X$. If $Z$ is a hole, then either $Z$ or some $W$ containing $Z$ would have been chosen to define an inductive interval $I$ rather than any of the $I_{Y_{i}}$. This is a contradiction. If $Z$ is a nonhole, then since it contains no holes, either it or some $W$ containing it would have been chosen rather than $I_{Y_{i}}$, and again we have a contradiction.

Thus $Z=X$ and so the $\left\{\partial Y_{i}\right\}$ fill $X$. Again let $g_{X}$ be a geodesic in $\mathcal{C}(X)$ joining $\gamma_{i}$ and $\gamma_{j}$. For each $Y_{i}$ there is a curve $\omega_{i} \in m$ such that $d_{X}\left(\omega_{i}, \partial Y_{i}\right) \leq 1$. For any $\omega$, the set of $\partial Y_{i}$ within distance 1 of $\omega$ do not fill $X$. There are at most $N_{1}$ of these $Y_{i}$. This forces the total number of $\omega \in g_{X}$ that are within distance 1 of some $\partial Y_{i}$ to be at least $l / N_{1}$, giving the desired estimate.
12.4. The upper bound. We are now equipped to prove Theorem 12.1. Recall we are given curves $\gamma_{h}, \gamma_{k}$ and a hole $X$. We construct a path $l$
in $\mathcal{G}$ joining $\gamma_{h}$ and $\gamma_{k}$. For each basic interval $[r(i), r(j)]$ with corresponding curves $\gamma_{i}, \gamma_{j}$ we take the given subpath joining $\gamma_{i}$ and $\gamma_{j}$. For each shortcut interval $[r(i), r(j)]$ we take the shortcut path of length at most $N_{0}$ joining $\gamma_{i}, \gamma_{j}$, guaranteed to exist by assumption (??). For each inductive interval we take the path defined inductively through that interval, guaranteed to exist by (??).
Proposition 12.12. For some fixed $a, b, c$ the path $l$ satisfies

$$
|l| \leq_{a, b} \sum\left[d_{Y}(\alpha, \beta)\right]_{c}
$$

the sum over holes $Y \subset X$.
Proof. Let $N_{S}$ be the number of shortcut intervals. The path though each shortcut interval is bounded by $N_{0}$. Therefore we need to prove that $N_{S}$ is bounded in terms of $d_{X}\left(\gamma_{h}, \gamma_{k}\right)$. By Proposition 12.111 ), and 2) any hole $Y$ is associated to at most 1 inductive interval $I_{Y}=$ $[r(i), r(j)]$ and any $J_{Y}$ overlaps at most 2 other $I_{Z_{1}}=\left[r\left(p_{1}\right), r\left(q_{1}\right)\right], I_{Z_{2}}=$ $\left[r\left(p_{2}\right), r\left(q_{2}\right)\right]$ and we can assume $r\left(q_{1}\right) \leq r(i)<r(j) \leq r\left(p_{2}\right)$. The path in $\mathcal{C}(Y)$ is an unparametrized quasi-geodesic. Therefore

$$
d_{Y}\left(\gamma_{p_{1}}, \gamma_{q_{1}}\right)+d_{Y}\left(\gamma_{i}, \gamma_{j}\right)+d_{Y}\left(\gamma_{p_{2}}, \gamma_{q_{2}}\right) \leq_{a, b} d_{Y}\left(\gamma_{p_{1}}, \gamma_{q_{2}}\right)
$$

Thus the path length of $l$ through each hole $Y$ inductively is bounded by $K_{1} d_{Y}\left(\gamma_{h}, \gamma_{l}\right)+K_{2}$, for fixed $K_{1}, K_{2}$. Thus if we let $N_{I}$ denote the number of inductive intervals it follows that it is enough to prove that there are constants $a, b$ such that

$$
\begin{equation*}
\sum_{I_{B}}\left|I_{B}\right|+N_{S}+N_{I} \leq_{a, b} d_{X}\left(\gamma_{h}, \gamma_{k}\right) \tag{12.13}
\end{equation*}
$$

where the sum is over all basic intervals $I_{B}$. We prove this bound.
Let $\delta$ be the hyperbolicity constant for the space $\mathcal{C}(X)$. The fact that the sequence $\left\{\gamma_{i}\right\}$ is an unparametrized quasi-geodesic in $\mathcal{C}(X)$ says that there are constants $a, b$ and $D>\delta$ such that for any partition of the $\left\{\gamma_{i}\right\}$,

$$
\begin{equation*}
P=\gamma_{i}=\beta_{1}, \beta_{2} \ldots, \beta_{M}=\gamma_{j} \tag{12.14}
\end{equation*}
$$

with the property that for all $p, d_{X}\left(\beta_{p}, \beta_{p+1}\right) \geq D$, then

$$
\begin{equation*}
\sum_{p=1}^{M-1} d_{X}\left(\gamma_{p}, \gamma_{p+1}\right) \leq_{a, b} d_{X}\left(\gamma_{i}, \gamma_{j}\right) \tag{12.15}
\end{equation*}
$$

We partition the entire interval $[r(h), r(l)]$ as

$$
I_{1} \cup J_{1} \cup I_{2} \cup J_{2} \cup \ldots \cup I_{k} \cup J_{k}
$$

where the intervals $I_{i}$ are basic and satisfy

$$
\left|I_{i}\right| \geq 4 D
$$

Each interval $J_{i}$ is a connected union of basic, shortcut, and inductive intervals with the following properties. Each $J_{i}$ begins with either a shortcut or inductive interval associated with some proper subsurface $Y_{i}{ }^{1}$. It is followed by a possibly empty basic interval with length at most $4 D$, followed by a shortcut or inductive interval associated to some $Y_{i}^{2}$, then a basic interval again with length at most $4 D$, and so forth ending with a shortcut or inductive interval associated to some subsurface $Y_{i}^{l(i)}$. (We may have $I_{1}=\emptyset$ or $J_{k}=\emptyset$.)

We wish to form a partition $P$ of the entire interval $[r(h), r(k)]$. Suppose the shortcut and inductive intervals in $J_{i}$ are associated to $Y_{i}^{1}, \ldots Y_{i}^{n_{i}}$. Let $\omega_{i}, \omega_{i}^{\prime}$ the initial and terminal curves. Let

$$
K_{1}=K_{0}+C_{2} .
$$

By (??)

$$
d_{X}\left(\omega_{i}, \omega_{i}^{\prime}\right) \geq d_{X}\left(\partial Y_{i}^{1}, \partial Y_{i}^{n_{i}}\right)-2 K_{1}
$$

The contribution to the left side of (12.13) from $J_{i}$ is

$$
\begin{equation*}
n_{i}+\sum_{I_{B} \cap J}\left|I_{B}\right| \leq n_{i}(4 D+1) . \tag{12.16}
\end{equation*}
$$

By 3) of Proposition 12.11,

$$
\left.d_{X}\left(\partial Y_{1}, \partial Y_{i}^{n_{i}}\right)\right) \geq n_{i} / N_{1}-2
$$

and therefore

$$
\begin{equation*}
d_{X}\left(\omega_{i}, \omega_{i}^{\prime}\right) \geq n_{i} / N_{1}-2-2 K_{1} . \tag{12.17}
\end{equation*}
$$

We group the set of $J_{i}$ into two types. The first possibility is that

$$
d_{X}\left(\omega_{i}, \omega_{i}^{\prime}\right) \leq D
$$

so

$$
n_{i} / N_{1}-2-2 K_{1} \leq D
$$

Then

$$
n_{i}+\sum_{I_{B} \cap J_{i}}\left|I_{B}\right| \leq n_{i}(4 D+1) \leq(4 D+1)\left(D+2+2 K_{1}\right) N_{1} .
$$

This quantity then can by incorporated into the length through a basic interval of length at least $4 D$ bordering $J_{i}$ by multiplying by a fixed factor (depending on $D$ ) and so can be ignored in terms of (12.13). In our partition $P$ we include the midpoints $\alpha_{i}, \tau_{i}$ of the basic intervals $I, I^{\prime}$ on each side of $J_{i}$. Let $I_{1 / 2} \subset I, I_{1 / 2}^{\prime} \subset I^{\prime}$ the half intervals bordering
$J_{i}$. Then a simple thin triangle argument in the hyperbolic space $\mathcal{C}(X)$, and the fact that $D>\delta$, shows that

$$
d_{X}\left(\alpha_{i}, \tau_{i}\right) \geq D
$$

and furthermore

$$
\left|I_{1 / 2}\right|+\left|I_{1 / 2}^{\prime}\right| \leq_{a, b} d_{X}\left(\alpha_{i}, \tau_{i}\right)
$$

The second possibility for $J_{i}$ is that the endpoints satisfy

$$
d_{X}\left(\omega_{i}, \omega_{i}^{\prime}\right) \geq D
$$

In this case we include the endpoints $\omega_{i}, \omega_{i}^{\prime}$ in the partition. Combining (12.16) and (12.17) we get a bound

$$
\begin{equation*}
n_{i}+\sum_{I_{B} \cap J_{i}}\left|I_{B}\right| \leq(4 D+1) N_{1}\left(2+2 K_{1}+d_{X}\left(\omega_{i}, \omega_{i}^{\prime}\right)\right) . \tag{12.18}
\end{equation*}
$$

We now have a partition of the sequence of curves such that two consecutive curves are at least $D$ apart. The points of the partition include the endpoints of intervals $J_{i}$, if $J_{i}$ is of the second type, and the midpoints of intervals $I$ surrounding $J_{i}$ if it is of the first type.

Let $I_{p}=\left[r(p), r\left(p^{\prime}\right)\right]$ be a basic interval of length at least $4 D$ or half of a basic interval of length at least $2 D$ with endpoints $\gamma_{p}, \gamma_{p^{\prime}}$. By (??) there exists $C_{4}$ such that

$$
\begin{equation*}
d_{X}\left(\gamma_{p}, \gamma_{p^{\prime}}\right) \geq \frac{\left|I_{p}\right|}{C_{4}} \tag{12.19}
\end{equation*}
$$

Combining (12.18) and (12.19) and applying (12.15), we conclude that the left side of (12.13) which can be expressed as $\sum_{p=1}^{k}\left|I_{p}\right|+$ $\sum_{p=1}^{k} n_{p}+\sum_{I_{B} \cap J_{i}}\left|I_{B}\right|$ satisfies

$$
\begin{aligned}
\sum_{p}\left|I_{p}\right|+\sum_{p} n_{p}+\sum_{I_{B} \cap J_{p}}\left|I_{B}\right| \leq & C_{4} \sum_{p} d_{X}\left(\gamma_{p}, \gamma_{p}^{\prime}\right)+(4 D+1) N_{1}\left(2+2 K_{1}+\sum_{p} d_{X}\left(\omega_{p}, \omega_{p}^{\prime}\right)\right) \\
& \leq_{a, b} d_{X}\left(\gamma_{h}, \gamma_{k}\right)
\end{aligned}
$$

## 13. Background on Teichmüller space

Fix now a surface $S=S_{g, n}$ of genus $g$ with $n$ punctures. Let $\mathcal{M}=$ $\mathcal{M}(S)$ be the Teichmüller space of $S$. This is slightly non-standard notation; however this fits with our general framework where $\mathcal{M}$ is a marking space.

Recall that $\mathcal{M}(S)$ is the set of equivalence classes of conformal structures $\mu$ on $S$. Here $\mu_{1} \sim \mu_{2}$ if there is a conformal map $f: \mu_{1} \rightarrow \mu_{2}$
which is isotopic to the identity. We equip Teichmüller space with the Teichmüller metric. That is,

$$
d_{\mathcal{M}}\left(\mu_{1}, \mu_{2}\right)=\inf _{f}\left\{\frac{1}{2} \log K(f)\right\} .
$$

Here $K(f)$ is the maximal dilatation of $f$. The infimum ranges over all maps $f: \mu_{1} \rightarrow \mu_{2}$ isotopic to the identity and it is realized by a Teichmüller map. This, in turn, may be defined in terms of a quadratic differential.

Definition 13.1. A holomorphic quadratic differential $q(z) d z^{2}$ on a Riemann surface $\mu_{0}$, possibly with punctures, is an assignment of a holomorphic function $q_{z}(z)$ to each local holomorphic coordinate chart. If $z$ and $\zeta$ are overlapping charts then we require the equality

$$
q_{z}(z)=q_{\zeta}(\zeta)\left(\frac{d \zeta}{d z}\right)^{2}
$$

to hold in the intersection of the charts.
The sum of the orders of the zeroes and poles of $q$ totals $4 g-4$. At any point away from the zeroes and poles there is a natural coordinate $z=x+i y$ with the property that $q_{z} \equiv 1$. In this natural coordinate the foliation by lines $y=c$ is called the horizontal foliation. The foliation by lines $x=c$ is called the vertical foliation.

Now fix a quadratic differential $q$. For every $t \in \mathbb{R}$ there is a new Riemann surface $\mu_{t}$ and associated quadratic differential $g_{t}(q)$ on $\mu_{t}$. We are also given a map $f_{t}: \mu_{0} \rightarrow \mu_{t}$, called a Teichmüller map. The Teichmüller map is defined by the property that in the natural coordinates $z=x+i y$ of $q$ and $z_{t}=x_{t}+i y_{t}$ of $g_{t}(q)$ we have

$$
x_{t}=e^{t} x, \quad y_{t}=e^{-t} y .
$$

The Teichmüller map is the unique map realizing the infimum in the definition of the Teichmüller distance.

Let $\sigma_{t}$ be the unique hyperbolic metric in the conformal class of $\mu_{t}$. Let $P_{t}$ be a Bers pants decomposition of $S$;
$P_{t}$ of shortest length curves, called a Bers pants decomposition and also a shortest marking, again denoted $\mu_{t}$. For any subsurface $X$ there is a projection $\pi_{X}: \mathcal{M} \rightarrow \mathcal{C}(X)$ which first associates to $\mu \in \mathcal{M}$ its minimal length marking and then projects that marking into $X$. This

I don't think that is true. Consider an annulus.
map is Lipschitz. Let $\epsilon_{1}$ a Margulis constant so that intersecting curves cannot both be of Poincare length smaller than $\epsilon_{1}$.

## 14. Paths for the nonorientable surface

Again fix $F$ a compact, connected, and nonorientable surface. Let $S$ be the orientation double cover with covering map $\rho_{F}: S \rightarrow F$. Let $\tau: S \rightarrow S$ be the associated involution. Let $\nu: \mathcal{C}^{\tau}(S) \rightarrow \mathcal{C}(S)$ be the natural map from the complex of $\tau$-invariant curves to $\mathcal{C}(S)$. As noted above, the inclusion is Lipschitz.

Our goal here is to obtain a reverse inequality:

$$
d_{\mathcal{C}^{\tau}}(\alpha, \beta) \geq_{a, b} d_{S}(\alpha, \beta)
$$

We begin by giving the promised proof of:
Lemma 14.1. There is a constant $K$ with the following property: Suppose that $\alpha, \beta$ are invariant multi-curves or arcs in $S$ and $X \subset S$ so that $d_{X}(\alpha, \beta)>K$. Then $X$ is symmetric.

Proof. Let $Y=\tau(X)$. Since $\alpha$ and $\beta$ are invariant deduce $d_{Y}(\alpha, \beta)=$ $d_{X}(\alpha, \beta)$. We can assume $d_{\mathcal{C}}(\alpha, \beta) \geq 3$. This implies that $\alpha, \beta$ fill the surface $S$, and so there exists a quadratic differential $q$ on some Riemann surface $X_{0}$ which has as its horizontal leaves, closed curves homotopic to $\alpha$ and vertical leaves homotopic to $\beta$. If $\alpha$ has two components then we require that the moduli of the corresponding cylinders coincide. The same is true for $\beta$. Since $\tau$ preserves $\alpha$ and $\beta, \tau^{*} q$ is another quadratic differential on $X_{0}$ all of whose horizontal leaves are homotopic to $\alpha$ and vertical leaves homotopic to $\beta$. But then $\tau^{*} q=q$ and so $q$ is symmetric.

Let $g_{t}(q),-\infty \leq t \leq \infty$ the corresponding path of quadratic differentials. Each is $\tau$ symmetric. For each time $t$, let $\sigma_{t}$ be the hyperbolic metric on the surface of $g_{t}(q)$ and $\mu_{t}$ the shortest marking. The hyperbolic metric is $\tau$ invariant. Let $\epsilon_{1}$ a Margulis constant so that intersecting curves cannot both be of Poincare length smaller than $\epsilon_{1}$.

By Theorem ( ) of Rafi there is a constant $K$ such that if $d_{X}(\alpha, \beta) \geq$ $3 K$, there are $t_{1}<t_{2}$ so that

$$
\begin{equation*}
d_{X}\left(\mu_{t_{1}}, \mu_{t_{2}}\right) \geq K \tag{14.2}
\end{equation*}
$$

and such that for all $t_{1} \leq t \leq t_{2}$

$$
\sigma_{t}(\partial X)<\epsilon_{1} .
$$

Since $\sigma_{t}$ is symmetric, we have $\sigma_{t}(\tau(\partial X))=\sigma_{t}(\partial X)$. This implies $i(\partial X, \tau(\partial X)=\emptyset$. We cannot have any component of $\tau(\partial X)$ contained in $X$ for this would violate (14.2). Thus either $\tau(X)=X$ or $\tau(X) \cap X=$ $\emptyset$.

As $t \rightarrow-\infty, \rho_{t}(\alpha) \rightarrow 0$ and no other curve has hyperbolic length going to 0 and as $t$ We can let We have a sequence $\gamma_{n}$ of shortest curves on these surfaces. The function

$$
i: \mathcal{A C} \times \mathcal{M} \rightarrow \mathbb{R}^{+}
$$

is given by the length of the shortest geodesic in the homotopy class.
We need to define sequences $\alpha=\gamma_{1}, \ldots, \gamma_{I}=\beta$ of curves, a corresponding sequence $\mu_{n}$ of points in Teichmuller space that satisfy the path requirements. We also have to define the accessible intervals so as to satisfy the Accessible requirements. We also need then to check (??), (??) and (??).

We check the marking requirements on our sequences in Teichmuller space and curves. We already have Condition 1). Condition 2) is exactly Theorem ( ) of [?] for the whole surface, and Theorem 1 of [?] in the case of subsurfaces. Condition 3) holds since intersection number is always a bound for distance. shortest curves can intersect a bounded number of times.

We check the combinatorial path requirements. In all cases there was no reindexing so 2 ) holds. Condition 3 ) holds since bounded length curves can intersect only a bound number of times. A bounded length surgery path produces curves that connect them. Condition 4) follows for example from the distance formula given in Theorem 1.1 of [?]. We are given

$$
d_{T}\left(\mu_{i}, \mu_{i+1}\right)=1
$$

where $d_{T}(\cdot, \cdot)$ is Teichmuller distance. On the other hand

$$
d_{T}\left(\mu_{i}, \mu_{i+1}\right) \stackrel{a, b}{=} d_{Y}\left(\mu_{i}, \mu_{i=1} .\right.
$$

This implies there is a bound for any subsurface projection which implies there is a bound for $i\left(\gamma_{i}, \gamma_{i+1}\right)$

## 15. Paths for the arc complex

Here we build the pair of paths required to obtain the upper bound on distance for the arc complex $\mathcal{A}(S, \Delta)$.

## 16. Background on train tracks

Here we give the necessary definitions and theorems regarding train tracks. This treatment is essentially due to Lee Mosher.

Recall that a generic train track $\tau \subset S$ is a smooth $\left(C^{2}\right)$ embedded trivalent graph. As usual we call the vertices switches and the edges branches. At every switch the tangents of the three branches agree.

Also, there are exactly two incoming branches and one outgoing branch at each switch. See Figure 7 for the local model of a switch.
incoming


Figure 7. The local model of a train track
Let $N=N(\tau)$ be an closed $I$-bundle neighborhood of $\tau$. Note that $\partial N$ splits into two collections of arcs: the vertical boundary $\partial_{v} N$ and the horizontal boundary $\partial_{h} N$. (If $\tau$ has no switches then $\partial_{v} N$ is empty.) See Figure 8. Conversely, given a foliated neighborhood $N$, taking the quotient of the $I$-fibres recovers $\tau$, up to isotopy.


Figure 8. A vertically foliated neighborhood of $\tau$.
A simple closed curve $\alpha$ is carried by $\tau$, written $\alpha \prec \tau$, if $\alpha$ lies in the interior of $N(\tau)$, transverse to the vertical fibres. If $\alpha$ is an arc then we additionally require that $\partial \alpha \subset \partial_{v} N$. The weights of $\alpha$ on $\tau$ are then a non-negative integer for each branch of $\tau$ : the number of times $\alpha$ meets any fibre in the neighborhood of that branch. The set of all allowable weights on $\tau$, denoted $P(\tau)$ is a cone over $\mathbb{Z}$. A simple closed curve carried by $\tau$, with weights lying on an extremal ray of $P(\tau)$, is called a vertex of $\tau$. Denote the set of vertices of $\tau \operatorname{by} \operatorname{Vert}(\tau)$. If $\tau$ and $\sigma$ are tracks, and $Y \subset S$ is an essential surface, then define $d_{Y}(\tau, \sigma)=d_{Y}(\operatorname{Vert}(\tau), \operatorname{Vert}(\sigma))$.

If $\alpha \prec \tau$ then we may split the track $\tau$ along $\alpha$ : Form $N=N(\tau)$, let $n(\alpha)$ be an open neighborhood of $\alpha$, let $N^{\prime}=N-n(\alpha)$, and set $\tau^{\prime}=N^{\prime} / I$. If all weights of $\alpha$ on $\tau$ are two or less then we call $\tau^{\prime}$ a central splitting of $\tau$. Notice that there are only finitely many central splittings possible of any given track.

A train track $\sigma$ is a subtrack of $\tau$ if $\sigma$ is obtained by deleting branches of $\tau$. In this situation we write $\sigma \subset \tau$. We now have a crucial definition:

Definition 16.1. An essential surface $Y \subset S$ is accessible from $\tau$ if $Y$ can be realized as follows: there is a subtrack $\sigma \subset \tau$ and a central splitting $\sigma^{\prime}$ of $\sigma$ so that a regular neighborhood $N$ of $\sigma^{\prime}$, plus any disk components of $S \backslash N$ to $N$, is isotopic to $Y$. We will denote $\sigma^{\prime}$ by $\tau \mid Y$.
Actually, I'm pretty sure that $\tau \mid Y$ is only defined up to some fuzzyness...

There is a bound, depending only on $S$, on the number of surfaces accessible from any one given train track.

We say that a train track $\sigma$ is obtained from $\tau$ by sliding if $\sigma$ and $\tau$ are related as in Figure 9. We say that a train track $\sigma$ is obtained from $\tau$ by splitting if $\sigma$ and $\tau$ are related as in Figure 10.


Figure 9. All slides take place in a small regular neigborhood of the affected branch.


Figure 10. There are three kinds of splitting: right, left, and central.

If $\sigma$ is a train track obtained from $\tau$ by a sequence of slides and exactly one splitting then we say that $\sigma$ is obtained via a wide splitting. In this case we write $\sigma \prec \tau$. We remark that if there is a wide splitting from $\tau$ to $\sigma$ then there is one using an a priori bounded number of slides. We shall always use these parsimonious wide splittings. For an in-depth discussion see Section 3.13 of Mosher's monograph [Mos].
Definition 16.2. A splitting sequence is a collection $\left\{\tau_{i}\right\}$ of train tracks so that $\tau_{0} \succ \tau_{1} \succ \ldots \succ \tau_{n}$.

We may now state the first of Mosher's results:
Theorem 16.3 (Mosher). Fix a surface $S$. There are constants $(a, b, c)$ with the following property: Suppose that $\left\{\tau_{i}\right\}$ is a splitting sequence in $S$ and $Y \subset S$ is an essential surface. Then the map $i \mapsto \pi_{Y}\left(\operatorname{Vert}\left(\tau_{i}\right)\right)$ is an ( $a, b, c$ ) unparametrized quasi-geodesic.

Note that, when $Y=S$, Theorem 16.3 is essentially due to the first author and Minsky; see Theorem 1.3 of [MM]. To see why Theorem 16.3 holds when $Y$ a strict subsurface we define:
Definition 16.4. For any essential subsurface $Y \subset S$ and any splitting sequence $\left\{\tau_{i}\right\}$ define the accessible region $J_{Y}$ to be set of indices so that $Y$ is accessible from $\tau_{i}$.
Lemma 16.5 (Mosher). The accessible region $J_{Y}$ is a (possibly empty) interval.

We now have:
Theorem 16.6 (Mosher). Fix a surface $S$. There is a constant $K$ with the following property: Suppose that $\left\{\tau_{i}\right\}$ is a splitting sequence and $Y$ is an essential subsurface. There is a splitting sequence $\left\{\sigma_{k}\right\}$ in $Y$, parametrized by $J_{Y}=[i, j]$, so that

- If $k \leq i$ (respectively, $j \leq k$ ), then

$$
d_{Y}\left(\tau_{k}, \tau_{i}\right) \leq K \quad\left(\text { resp. } d_{Y}\left(\tau_{j}, \tau_{k}\right) \leq K\right) .
$$

- For $k \in J_{Y}$, the tracks $\tau_{k} \mid Y$ and $\sigma_{k}$ have splitting distance at most $K$.
Note that Theorem 16.3 follows from the cited result of [MM] and Theorem 16.6.


## 17. Paths for the disk complex

18. The distance estimate

Needs work.
Needs work.

Theorem 18.1. For any handlebody $V$ there is a constant $c_{0}=c_{0}(V)$ so that, for any $c \geq c_{0}$ there are constants $a \geq 1$ and $b \geq 0$ satisfying

$$
d_{\mathcal{D}}(D, E) \stackrel{a, b}{=} \sum\left[d_{X}(D, E)\right]_{c}
$$

independent of the choice of $D$ and $E$. The sum ranges over subsurfaces $X \subset \partial V$ which are holes for $\mathcal{D}(V)$.

## 19. Hyperbolicity

We remark that some of the ideas in this section are similar to the heirarchy machine of [MM00] (see also Chapters 4 and 5 of Behrstock's thesis [Beh04]). Our discussion has a different focus; we avoid the
heirarchy machine as it appears to be too rigid to deal with the disk complex. As another difference from [MM00] and [Beh04] we rely on Theorem 2.8 instead of proving that projection maps are coarsely contracting.

We wish to prove:
Theorem 19.1. Suppose that $\mathcal{G}(S)$ satisfies the distance estimate and the path requirements. Suppose also that all holes for $\mathcal{G}$ interfere. Then $\mathcal{G}$ is Gromov hyperbolic.

As corollaries we have
Theorem 19.2. The arc complex is Gromov hyperbolic.
Theorem 19.3. The disk complex is Gromov hyperbolic.
In fact, Theorem 19.1 follows quickly from:
Theorem 19.4. Fix $\mathcal{G}$, a combinatorial complex. Suppose that $\mathcal{G}$ satisfies the distance estimate. Suppose that all holes for $\mathcal{G}$ interfere. For any constants $(a, b, c)$ there is a $\delta$ so that for any triangle of paths $T \subset \mathcal{G}$ we have: if the projection of any side of $T$ into into any hole is an ( $a, b, c$ ) unparametrized quasi-geodesic, then $T$ is $\delta$-slim.

Proof of Theorem 19.1. Suppose that $\mathcal{G}(S)$ satisfies the distance estimate and the path requirements. Suppose also that all holes for $\mathcal{G}$ interfere. Then, by Theorem ???, there are uniform constants $(a, b, c)$ so that for any pair $\alpha, \beta \in \mathcal{G}$ there is a path $\mathcal{P}=\left\{\gamma_{i}\right\} \subset \mathcal{G}$ with

- for any hole $X$ for $\mathcal{G}$, the projection $\pi_{X}(\mathcal{P})$ is an $(a, b, c)$ unparametrized quasi-geodesic and
- $|\mathcal{P}| \stackrel{a, b}{=} d_{\mathcal{G}}(\alpha, \beta)$.

So if $\alpha \cap \beta=\emptyset$ then $|\mathcal{P}|$ is uniformly short. Also, by Theorem 19.4, triangles made of such paths are uniformly slim. Thus, by Theorem 2.8, $\mathcal{G}$ is Gromov hyperbolic.

The rest of this section is devoted to proving Theorem 19.4.
19.1. Index in a hole. For the following definitions, we assume that $\alpha$ and $\beta$ are fixed vertices of $\mathcal{G}$.

For any hole $X$ and for any geodesic $h \in \mathcal{C}(X)$ connecting a point of $\pi_{X}(\alpha)$ to a point of $\pi_{X}(\beta)$ we also define $\rho_{h}: \mathcal{G} \rightarrow h$ to be the map $\pi_{X} \mid \mathcal{G}: \mathcal{G} \rightarrow \mathcal{C}(X)$ followed by closest points projection to $h$. Define index $_{X}: \mathcal{G} \rightarrow \mathbb{N}$ to be the index in $X$ :

$$
\operatorname{index}_{X}(\sigma)=d_{X}\left(\alpha, \rho_{h}(\sigma)\right)
$$

Remark 19.5. Suppose that $h^{\prime}$ is a different geodesic connecting $\pi_{X}(\alpha)$ to $\pi_{X}(\beta)$ and index ${ }_{X}^{\prime}$ is defined with respect to $h^{\prime}$. Then

$$
\left|\operatorname{index}_{X}(\sigma)-\operatorname{index}_{X}^{\prime}(\sigma)\right| \leq 12 \delta+2
$$

by Lemma 2.7. Thus, if we are willing to accept a small additive error, the choice of geodesic $h$ is irrelevant. Accordingly we will supress the superscript whenever possible.
19.2. Back and sidetracking. Fix $\sigma, \tau \in \mathcal{G}$. We say $\sigma$ precedes $\tau$ by at least $K$ in $X$ if

$$
\operatorname{index}_{X}(\sigma)+K \leq \operatorname{index}_{X}(\tau)
$$

We say $\sigma$ precedes $\tau$ by at most $K$ if the inequality is reversed. If $\sigma$ precedes $\tau$ then we say $\tau$ succeeds $\sigma$.

Now take $\mathcal{P}=\sigma_{i}$ to be a path in $\mathcal{G}$ connecting $\alpha$ to $\beta$. We assume that $\sigma_{i}$ and $\sigma_{i+1}$ are disjoint.

We now formalize a pair of properties enjoyed by unparametrized quasi-geodesics to the situation at hand. The path $\mathcal{P}$ backtracks at most $K$ if for every hole $X$ and all indices $i<j$ we find that $\sigma_{j}$ precedes $\sigma_{i}$ by at most $K$. The path $\mathcal{P}$ sidetracks at most $K$ if for every hole $X$ and every index $i$ we find that

$$
d_{X}\left(\sigma_{i}, \rho_{h}\left(\sigma_{i}\right)\right) \leq K,
$$

for some geodesic $h$ connecting a point of $\pi_{X}(\alpha)$ to a point of $\pi_{X}(\beta)$.
Remark 19.6. As in Remark 19.5, allowing a small additive error makes irrelevant the choice of geodesic in the definition of sidetracking. We note that, if $\mathcal{P}$ has bounded sidetracking, one may freely use in calculation whichever of $\sigma_{i}$ or $\rho_{h}\left(\sigma_{i}\right)$ is more convenient.
19.3. Projection control. We say domains $X, Y \subset S$ overlap if $X$ and $Y$ intersect but are not nested. The following theorem (see Theorem 4.2.1 of Behrstock's thesis [Beh04]) follows from Masur and Minsky's idea (see [MM00]) of time ordered domains in $S$ :
Theorem 19.7. There is a constant $M_{1}=M_{1}(S)$ with the following property. Suppose that $X, Y$ are overlapping non-simple domains. If $\gamma \in \mathcal{A C}(S)$ cuts both $X$ and $Y$ then either $d_{X}(\gamma, \partial Y)<M_{1}$ or $d_{Y}(\partial X, \gamma)<M_{1}$.

We also require a more specialized version of Theorem 19.7 for the case where $X$ and $Y$ are nested.
Lemma 19.8. There is a constant $M_{2}=M_{2}(S)$ with the following property. Suppose that $X \subset Y$ are nested non-simple domains. Fix $\alpha, \beta, \gamma \in \mathcal{A C}(S)$ which cut both $X$ and $Y$. Fix $k=\left[\alpha^{\prime}, \beta^{\prime}\right] \subset \mathcal{C}(Y)$, a geodesic connecting a point of $\pi_{Y}(\alpha)$ to a point of $\pi_{Y}(\beta)$. Assume
that $d_{X}(\alpha, \beta) \geq M_{0}$, the constant given by the Bounded Geodesic Image Theorem [MM00]. If $d_{X}(\alpha, \gamma) \geq M_{2}$ then

$$
\operatorname{index}_{Y}(\partial X)-4 \leq \operatorname{index}_{Y}(\gamma)
$$

Symmetrically, we have

$$
\operatorname{index}_{Y}(\gamma) \leq \operatorname{index}_{Y}(\partial X)+4
$$

if $d_{X}(\gamma, \beta) \geq M_{2}$.
19.4. Finding the midpoint of a side. Let $(a, b, c)$ be arbitrary. Let $\mathcal{P}, \mathcal{Q}, \mathcal{R}$ be the sides of a triangle in $\mathcal{G}$ with vertices at $\alpha, \beta, \gamma$. We assume that each of $\mathcal{P}, \mathcal{Q}$, and $\mathcal{R}$ are ( $a, b, c$ ) unparametrized quasigeodesics when projected to any hole.

Recall that $M_{0}=M_{0}(S), M_{1}=M_{1}(S)$, and $M_{2}=M_{2}(S)$ are functions depending only on the topology of $S$. We may assume that if $T \subset S$ is an essential subsurface, then $M_{0}(S)>M_{0}(T)$.

Now choose $K_{1} \geq \max \left\{M_{0}, 4 M_{1}, M_{2}, 8\right\}+6 \delta$ sufficently large so that any ( $a, b, c$ ) unparametrized quasi-geodesic in any hole back and side tracks at most $K_{1}$.
Claim 19.9. If $\sigma_{i}$ precedes $\gamma$ in $X$ and $\sigma_{j}$ succeeds $\gamma$ in $Y$, both by at least $2 K_{1}$, then $i<j$.

Proof. To begin, as $X$ and $Y$ are holes and all holes interfere, we need not consider the possibility that $X \cap Y=\emptyset$. If $X=Y$ we immediately deduce that

$$
\operatorname{index}_{X}\left(\sigma_{i}\right)+2 K_{1} \leq \operatorname{index}_{X}(\gamma) \leq \operatorname{index}_{X}\left(\sigma_{j}\right)-2 K_{1}
$$

Thus index ${ }_{X}\left(\sigma_{i}\right)+4 K_{1} \leq \operatorname{index}_{X}\left(\sigma_{j}\right)$. Since $\mathcal{P}$ backtracks at most $K_{1}$ we have $i<j$, as desired.

Suppose instead that $X \subset Y$. Since $\sigma_{i}$ precedes $\gamma$ in $X$ we immediately find $d_{X}(\alpha, \beta) \geq 2 K_{1} \geq M_{0}$ and $d_{X}(\alpha, \gamma) \geq 2 K_{1}-2 \delta \geq M_{2}$. Apply Lemma 19.8 to deduce index ${ }_{Y}(\partial X)-4 \leq \operatorname{index}_{Y}(\gamma)$. Since $\sigma_{j}$ succeeds $\gamma$ in $Y$ it follows that $\operatorname{index}_{Y}(\partial X)-4+2 K_{1} \leq \operatorname{index}_{Y}\left(\sigma_{j}\right)$. Again using the fact that $\sigma_{i}$ precedes $\gamma$ in $X$ we have that $d_{X}\left(\sigma_{i}, \beta\right) \geq M_{2}$. We deduce from Lemma 19.8 that $\operatorname{index}_{Y}\left(\sigma_{i}\right) \leq \operatorname{index}_{Y}(\partial X)+4$. Thus

$$
\operatorname{index}_{Y}\left(\sigma_{i}\right)-8+2 K_{1} \leq \operatorname{index}_{Y}\left(\sigma_{j}\right)
$$

Since $\mathcal{P}$ backtracks at most $K_{1}$ in $Y$ we again deduce that $i<j$. The case where $Y \subset X$ is handled in symmetric fashion.

Suppose now that $X$ and $Y$ overlap. Applying Theorem 19.7 and breaking symmetry, we may assume that $d_{X}(\gamma, \partial Y)<M_{1}$. Since $\sigma_{i}$
precedes $\gamma$ we have index ${ }_{X}(\gamma) \geq 2 K_{1}$. Thus, it follows that index ${ }_{X}(\partial Y) \geq$ $2 K_{1}-2 M_{1}-4 \delta$ and so

$$
d_{X}(\alpha, \partial Y) \geq 2 K_{1}-2 M_{1}-6 \delta \geq M_{1}
$$

Applying Theorem 19.7 again, we find that $d_{Y}(\alpha, \partial X)<M_{1}$. Now, since $\sigma_{j}$ succeeds $\gamma$ in $Y$, we deduce that $\operatorname{index}_{Y}\left(\sigma_{j}\right) \geq 2 K_{1}$. Similar considerations to the above show that

$$
d_{Y}\left(\partial X, \sigma_{j}\right) \geq 2 K_{1}-M_{1}-2 \delta \geq M_{1} .
$$

Applying Theorem 19.7 one last time, we find that $d_{X}\left(\partial Y, \sigma_{j}\right)<M_{1}$. Thus $d_{X}\left(\gamma, \sigma_{j}\right) \leq 2 M_{1}$. Finally, we deduce that the difference in index (in $X$ ) between $\sigma_{i}$ and $\sigma_{j}$ is at least $2 K_{1}-4 M_{1}-4 \delta$. Since this is again greater than $K_{1}$, it follows that $i<j$.

Let $\sigma_{\alpha} \in \mathcal{P}$ be the last vertex of $\mathcal{P}$ preceding $\gamma$ by at least $2 K_{1}$ in some hole. If no such vertex of $\mathcal{P}$ exists then take $\sigma_{\alpha}=\alpha$.
Claim 19.10. There is a constant $N_{1}=N_{1}(S)$ with the following property. For every hole $X$ and geodesic $h$ connecting $\pi_{X}(\alpha)$ to $\pi_{X}(\beta)$ :

$$
d_{X}\left(\sigma_{\alpha}, \rho_{h}(\gamma)\right) \leq N_{1}
$$

Proof. Since $\sigma_{i}$ and $\sigma_{i+1}$ are disjoint we have

$$
\left|\operatorname{index}_{X}\left(\sigma_{i+1}\right)-\operatorname{index}_{X}\left(\sigma_{i}\right)\right| \leq 4 \delta+2
$$

Since $\mathcal{P}$ is a path connecting $\alpha$ to $\beta$ the image $\rho_{h}(\mathcal{P})$ is $4 \delta+2-$ dense in $h$. Thus, if index ${ }_{X}\left(\sigma_{\alpha}\right)+2 K_{1}+4 \delta+2<\operatorname{index}_{X}(\gamma)$ then we have a contradiction to the definition of $\sigma_{\alpha}$.

On the other hand, if index ${ }_{X}\left(\sigma_{\alpha}\right) \geq \operatorname{index}_{X}(\gamma)+K_{1}$ then $\sigma_{\alpha}$ succeeds $\gamma$. This directly contradicts Claim 19.9.

We deduce that the difference in index between $\sigma_{\alpha}$ and $\gamma$ in $X$ is at most $2 K_{1}+4 \delta+2$. Finally, as $\mathcal{P}$ sidetracks by at most $K_{1}$ we have

$$
d_{X}\left(\sigma_{\alpha}, \rho_{h}(\gamma)\right) \leq 3 K_{1}+4 \delta+2
$$

as desired.
We define $\sigma_{\beta}$ to be the first $\sigma_{i}$ to succeed $\gamma$ by at least $2 K_{1}$ — if no such vertex of $\mathcal{P}$ exists take $\sigma_{\beta}=\beta$. If $\alpha=\beta$ then $\sigma_{\alpha}=\sigma_{\beta}$. Otherwise, from Claim 19.9, we immediately deduce that $\sigma_{\alpha}$ comes before $\sigma_{\beta}$ in $\mathcal{P}$. A symmetric version of Claim 19.10 applies to $\sigma_{\beta}$ : for every hole $X$

$$
d_{X}\left(\rho_{h}(\gamma), \sigma_{\beta}\right) \leq N_{1}
$$

19.5. Another side of the triangle. Recall now that we are also given a path $\mathcal{R}=\left\{\tau_{i}\right\}$ connecting $\alpha$ to $\gamma$ in $\mathcal{G}$. As before, $\mathcal{R}$ has bounded back and sidetracking. Thus we again find vertices $\tau_{\alpha}$ and $\tau_{\gamma}$ the last/first to precede/succeed $\beta$ by at least $2 K_{1}$. Again, this is defined in terms the closest points projection of $\beta$ to geodesics of the form $l=\left[\pi_{X}(\alpha), \pi_{X}(\gamma)\right]$. By Claim 19.10, for every hole $X, \tau_{\alpha}$ and $\tau_{\gamma}$ are close to $\rho_{l}(\beta)$.

By Lemma 2.5, if $h=\left[\pi_{X}(\alpha), \pi_{X}(\beta)\right]$, then $d_{X}\left(\rho_{h}(\gamma), \rho_{l}(\beta)\right) \leq 6 \delta$. We deduce:
Claim 19.11. $d_{X}\left(\sigma_{\alpha}, \tau_{\alpha}\right) \leq 2 N_{1}+6 \delta$.
We now prove:
Claim 19.12. There is a constant $N_{2}=N_{2}(S)$ with the following property. For every $\sigma_{i} \leq \sigma_{\alpha}$ in $\mathcal{P}$ there is a $\tau_{j} \leq \tau_{\alpha}$ in $\mathcal{R}$ so that

$$
d_{X}\left(\sigma_{i}, \tau_{j}\right) \leq N_{2}
$$

for every hole $X$.
Proof. We only sketch the proof, as the details are similar to the discussion above. Fix $\sigma_{i} \leq \sigma_{\alpha}$.

Suppose first that no vertex of $\mathcal{R}$ precedes $\sigma_{i}$ by more than $2 K_{1}$. Fix a hole $X$ and geodesics $h=\left[\pi_{X}(\alpha), \pi_{X}(\beta)\right]$ and $l=\left[\pi_{X}(\alpha), \pi_{X}(\gamma)\right]$. Then $\rho_{l}\left(\sigma_{i}\right)$ is within distance $2 K_{1}$ of $\pi_{X}(\alpha)$. Appealing to Claim 19.11, bounded sidetracking, and hyperbolicity of $\mathcal{C}(X)$ we find that the initial segments

$$
\left[\pi_{X}(\alpha), \rho_{h}\left(\sigma_{\alpha}\right)\right], \quad\left[\pi_{X}(\alpha), \rho_{l}\left(\tau_{\alpha}\right)\right]
$$

of $h$ and $l$ respectively must fellow travel. Because of bounded backtracking along $\mathcal{P}, \rho_{h}\left(\sigma_{i}\right)$ lies on, or at least near, this initial segment of $h$. Thus by Lemma $2.7 \rho_{l}\left(\sigma_{i}\right)$ is close to $\rho_{h}\left(\sigma_{i}\right)$ which in turn is close to $\pi_{X}\left(\sigma_{i}\right)$, because $\mathcal{P}$ has bounded sidetracking. In short, $d_{X}\left(\alpha, \sigma_{i}\right)$ is bounded for all holes $X$. Thus we may take $\tau_{j}=\tau_{0}=\alpha$ and we are done.

Now suppose that some vertex of $\mathcal{R}$ precedes $\sigma_{i}$ by at least $2 K_{1}$ in some hole $X$. Take $\tau_{j}$ to be the last such vertex in $\mathcal{R}$. Following the proof of Claim 19.9 shows that $\tau_{j}$ comes before $\tau_{\alpha}$ in $\mathcal{R}$. The argument now required to bound $d_{X}\left(\sigma_{i}, \tau_{j}\right)$ is essentially identical to the proof of Claim 19.10.

By the distance estimate, we find that there is a uniform neighborhood of $\left[\sigma_{0}, \sigma_{\alpha}\right] \subset \mathcal{P}$, taken in $\mathcal{G}$, which contains $\left[\tau_{0}, \tau_{\alpha}\right] \subset \mathcal{P}$. The slimness of $\mathcal{P Q R}$ follows directly. This completes the proof of Theorem 19.4.
20. An algorithm to coarsely compute Hempel distance

Clean up.
Definition 20.1. Let $\rho_{V}: \mathcal{C}(S) \rightarrow \mathcal{P}(\mathcal{D}(V))$ be the closest points projection map: for every $\gamma \in \mathcal{C}(S)$ take $\rho_{V}(\gamma)$ equal to the set of $D \in \mathcal{D}(V)$ so that $d_{S}(\gamma, D)=\min \left\{d_{S}(\gamma, E) \mid E \in \mathcal{D}(V)\right\}$.

Recall that a maximal simplex $\mathbb{D}$ in $\mathcal{D}(V)$ is a pants decomposition of $V$. The first application of our ideas is the "projection to handlebodies" theorem:
Theorem 20.6. There is a constant $R_{1}=R_{1}(V)$ and an algorithm which, given an essential curve $\gamma \subset S$ and a pants decomposition $\mathbb{D} \subset$ $\mathcal{D}(V)$, finds a disk $E \in \mathcal{D}(V)$ so that

$$
d_{S}\left(E, \rho_{V}(\gamma)\right) \leq R_{1}
$$

Recall the definition of the closest points projection map $\rho_{V}: \mathcal{C}(S) \rightarrow$ $\mathcal{P}(\mathcal{D}(V))$. For every $\gamma \in \mathcal{C}(S)$ we take $\rho_{V}(\gamma)$ equal to the set of $D \in$ $\mathcal{D}(V)$ so that $d_{S}(\gamma, D)=\min \left\{d_{S}(\gamma, E) \mid E \in \mathcal{D}(V)\right\}$.

Recall that a maximal simplex $\mathbb{D}$ in $\mathcal{D}(V)$ is a pants decomposition of $V$. The goal of this section is to prove Theorem 20.6: there is an algorithm which finds a point of $\rho_{V}(\gamma)$, up to bounded error.

Fix a handlebody $V$ with $\partial V=S$. Fix $\gamma$ an essential curve in $S$ and $\mathbb{D}$ a pants decomposition of $V$. Let $D$ be any element of $\rho_{V}(\gamma)$.

Let $H$ be a hierarchy connecting $\mathbb{D}$ to $\gamma$. Let $\sigma$ be the last slice of $H$ preceding $D$ on $H$. Note that $d_{S}(\sigma, D)$ is bounded by $2 \delta+R$, the latter being the quasi-convexity constant of $\mathcal{D}(V)$.
Claim 20.2. There is a multicurve $\mu$ so that $d_{X}(\sigma, \mu)$ is uniformly bounded for all essential subsurfaces $X$ and either

- $\mu$ is the boundary of a disk,
- $\mu$ is the boundary of a non-hole, or
- $\mu$ is the boundary of a hole of large diameter.

Proof. Connect $\sigma$ to $D$ by a hierarchy $L$. Connect $D$ to $\mathbb{D}$ by a hierarchy $K$. If $|L|$ is small then we may take $\mu=\partial D$. Otherwise, as $d_{S}(\sigma, D)$ is bounded, we may assume that the hierarchy $L$ contains at least one large projection in a strict subsurface. Let $X$ be the first such, under time-order. We may assume that $d_{X}(\sigma, D)$ is large.

We now claim that $\mu=\partial X$ satisfies the conclusion of the claim. This is clear if $X$ is not a hole. If $X$ is a hole then, as $\sigma$ is the last slice preceding $D$ along $H$, we deduce that $d_{X}(\mathbb{D}, D)$ is also large. We are done.

We immediately deduce:

Claim 20.3. There is a disk $E \in \mathcal{D}(V)$ so that $E$ and $\mu$ meet at most twice and so that $d_{S}\left(E, \rho_{V}(\gamma)\right) \leq R_{1}$.

Also, by Lemma 3.15 we have:
Claim 20.4. The multi-curve $\mu$ and the slice $\sigma$ have bounded intersection.

Let $R_{3}=R_{3}(S)$ be the given bound on intersection number. We may now describe the algorithm:
Algorithm 20.5. We are given $\gamma \in \mathcal{C}(S)$ and $\mathbb{D} \subset \mathcal{D}(V)$. Build $H$ a hierarchy between $\gamma$ and $\mathbb{D}$; see Leasure [Lea]. Find a resolution into slices $\left\{\sigma^{i}\right\}$ of $H$. For each slice $\sigma^{i}$ list all multicurves $\mu$ so that $\iota\left(\mu, \sigma^{i}\right) \leq R_{3}$.

For every multicurve $\mu$ so produced determine whether

- a component of $\mu$ bounds a disk in $V$ (using the solution of the word problem in the free group $\pi_{1}(V)$ ),
- a component of $S \backslash \mu$ compresses in $V$ (using Haken's algorithm), or
- a component of $S \backslash \mu$ cobounds an $I$-bundle in $V$ (using the Jaco-Tollefson algorithm for detecting the JSJ decomposition).
For every $\mu$ falling into one of the three categories above the algorithm also produces a disk $E_{\mu}$ with $\iota\left(E_{\mu}, \mu\right) \leq 2$. For every disk $E_{\mu}$ so produced compute $d_{S}\left(\gamma, E_{\mu}\right)$ (again, see [Lea]). Finally take $E$ equal to any of the $E_{\mu}$ which minimized $d_{S}\left(\gamma, E_{\mu}\right)$, among all disks considered.

By Claim 20.3 the disk $E$ satisfies $d_{S}\left(E, \rho_{V}(\gamma)\right) \leq R_{1}$. This proves:
Theorem 20.6. There is a constant $R_{1}=R_{1}(V)$ and an algorithm which, given an essential curve $\gamma \subset S$ and a pants decomposition $\mathbb{D} \subset$ $\mathcal{D}(V)$, finds a disk $E \in \mathcal{D}(V)$ so that

$$
d_{S}\left(E, \rho_{V}(\gamma)\right) \leq R_{1}
$$

Remark 20.7. We obviously make no claims as to the efficiency of this algorithm.
Theorem 20.8. There is a constant $R_{2}=R_{2}(S)$ and an algorithm which, given a Heegaard diagram $(S, \mathbb{D}, \mathbb{E})$, computes a number $N$ so that

$$
\left|d_{S}(V, W)-N\right| \leq R_{2}
$$

Proof. Using Theorem 20.6 we find a disk $D$ close to $\pi_{V}(\mathbb{E})$. Similarly, we find a disk $E$ close to $\pi_{W}(D)$.

Recall that computing distance between fixed vertices in the curve complex is algorithmic [Lea]. So, we may compute the distance $d_{S}(D, E)$. By the quasi-convexity of $\mathcal{D}(V)$ and of $\mathcal{D}(W)$ (see [MM]) this is the desired number.

## Appendix A. Teichmüller geodesics do not backtrack

KASRA RAFI
Let $G: \mathbb{R} \rightarrow \mathcal{T}(S)$ be a Teichmüller geodesic, and let $\nu_{ \pm}$be the corresponding vertical and horizontal foliations. Let $Y$ be a subsurface of $S$ that is not an annulus or a pair of pants. To be more precise, for $t \in \mathbb{R}$, let $\mu_{t}$ be a shortest marking of $S$ in $G(t)$. We call the map $\rho_{Y}: \mathbb{R} \rightarrow \mathcal{C}(Y)$ a shadow of $G$ to $\mathcal{C}(Y)$ if $\rho_{Y}(t)$ is a curve in $\pi_{Y}\left(\mu_{t}\right)$. In these notes, we prove the following theorem.
Theorem A.1. Every shadow of a Teichmüller geodesic $G$ to $\mathcal{C}(Y)$ is an un-parametrized quasi-geodesic in $\mathcal{C}(Y)$.

We remark that this contrasts with the way geodesics behave in the Lipschitz metric on $\mathcal{T}(S)$, studied by Thurston in [Thu98], where the projection of a geodesic to a subsurface can be move back and forth for arbitrarily long periods. (Examples can easily be produced using Thurston's construction of minimal stretch maps [Thu98] and the results in [CR05]).

To begin, note that if $Y=S$, the above is a theorem of Masur and Minsky [MM00, Theorem 3.3], that is, we already know that the shadow of $G$ to $\mathcal{C}(S)$ does not backtrack. To prove Theorem A. 1 in general, in addition to the Masur-Minsky result we require the following theorem: Let $\mathrm{FN}_{Y}: \mathcal{T}(S) \rightarrow \mathcal{T}(Y)$ be the map defined by restriction of FenchelNielsen coordinates to $Y$; that is, choose a pants decomposition of $S$ that contains the boundary of $Y$. The forgetful map sends FenchelNielsen coordinates on $\mathcal{T}(S)$, with respect to this pants decomposition, to Fenchel-Nielsen coordinates on $\mathcal{T}(Y)$ (see [Min96]).
Theorem A.2. For every Teichmüller geodesic $G$ and every subsurface $Y$, there exists an interval $J_{Y}=[a, b]$ and a geodesic $G_{Y}: J_{Y} \rightarrow \mathcal{T}(Y)$ such that

- If $t<a$ (resp. $b<t$ ), then

$$
d_{Y}\left(\mu_{t}, \mu_{a}\right)=O(1) \quad\left(\text { resp. } d_{Y}\left(\mu_{b}, \mu_{t}\right)=O(1)\right)
$$

- For $t \in J_{Y}$,

$$
d_{\mathcal{T}}\left(\left(\mathrm{FN}_{Y} \circ G\right)(t), G_{Y}(t)\right)=O(1)
$$

From the result of Masur and Minsky, we know that any shadow of $G_{Y}$ is an un-parametrized quasi-geodesic. The first conclusion of the above theorem implies that the image a shadow of $G \mid(-\infty, a]$ and $G \mid[b, \infty)$ is a bounded set in $\mathcal{C}(Y)$ and its second conclusion implies that any shadow of $G \mid[a, b]$ fellow travels any shadow of $G_{Y}$. That is, Theorem A. 1 follows from Theorem A.2.


Figure 11. $\left(\mathrm{FN}_{Y} \circ G\right)$ and $G_{Y}$ are bounded away in $\mathcal{T}(Y)$

Theorem A. 2 is essentially follows from results contained in [Min96], [Raf05], [Raf06] and [Raf07]. We sketch the arguments below.

Let $q_{t}$ be the quadratic differential given by the geodesic $G$ at time $t$. We call the flat metric $\left|q_{t}\right|$ the $q_{t}$-metric. For a curve $\alpha$ in $S$, let $q_{t}(\alpha)$ be the $q_{t}$-length of $\alpha$ : the length of a geodesic representative of $\alpha$ in the $q_{t}$-metric. Let $\alpha$ be a boundary component of $Y$, and let $\omega$ be an arc in $Y$ with both endpoints in $\alpha$. By the $q_{t}$-length of $\omega$, we mean the length of the shortest arc representing $\omega$ that starts and ends on a geodesic representative of $\alpha$ - all considered in the $q_{t}-$ metric. Again, denote this length by $q_{t}(\omega)$. Define

$$
M_{t}(\alpha, Y)=\min _{\omega} \frac{q_{t}(\omega)}{q_{t}(\alpha)},
$$

where $\omega$ ranges over all arcs in $Y$ with both endpoints on $\alpha$, as above. Let $t_{\alpha}$ be the time when $\alpha$ is balanced (its intersections with the horizontal and the vertical foliations are equal). The following lemma is contained in the proof of Theorem 3.1 of [Raf06].
Lemma A.3. There is a uniform constant $c \geq 0$ so that

$$
M_{s}(\alpha, Y) \leq M_{t}(\alpha, Y)+c
$$

for all $s \leq t \leq t_{\alpha}$ and for all $t_{\alpha} \leq t \leq s$.
Choose a large enough $M_{0}$ (see Theorem A. 4 below). Define $J_{\alpha, Y} \subset \mathbb{R}$ to be empty when $M_{t_{\alpha}}(\alpha, Y)<M_{0}$. On the other hand, if $M_{t_{\alpha}}(\alpha, Y) \geq$ $M_{0}$, then take $J_{\alpha, Y}$ to be the largest interval containing $t_{\alpha}$ so that $M_{t}(\alpha, Y) \geq M_{0}$ for all $t \in J_{\alpha, Y}$. We now define

$$
J_{Y}=\bigcap_{\alpha \subset \partial Y} J_{\alpha, Y}
$$

Note that, by Lemma A.3, for every $t \notin J_{Y}$, there is a boundary component $\alpha$ such that $M_{t}(\alpha, Y) \leq M_{0}+c$. Immediately from the proof of Theorem 5.5 of [Raf05] we now have:

Theorem A.4. There are constants $M_{0}$ and $C$ so that, if $M_{t}(\alpha, Y) \leq$ $M_{0}+c$ for some boundary component $\alpha$, then either

$$
d_{Y}\left(\mu_{t}, \nu_{-}\right) \leq C \quad \text { or } \quad d_{Y}\left(\mu_{t}, \nu_{+}\right) \leq C
$$

This proves the first conclusion of Theorem A.2. In fact, for $t<a$,

$$
d_{Y}\left(\mu_{t}, \nu_{-}\right) \leq C \quad \text { and } \quad d_{Y}\left(\mu_{a}, \nu_{-}\right) \leq C \Longrightarrow d_{Y}\left(\mu_{t}, \mu_{a}\right) \leq 2 C
$$

and, for $t>b$,

$$
d_{Y}\left(\mu_{t}, \nu_{+}\right) \leq C \quad \text { and } \quad d_{Y}\left(\mu_{b}, \nu_{+}\right) \leq C \Longrightarrow d_{Y}\left(\mu_{t}, \mu_{b}\right) \leq 2 C
$$

To obtain the second conclusion of Theorem A.2, we must first construct the candidate geodesic arc $G_{Y}$ in $\mathcal{T}(Y)$. Our plan, as in [Raf07], is to fill all components of $\partial Y$ with locally flat punctured disks. The quadratic differential $q_{t} \mid Y$ will extend over these disks, giving an extension $\bar{q}_{t}$. The map $G_{Y}: J_{Y} \rightarrow \mathcal{T}(Y)$, where $G_{Y}(t)$ is the conformal structure of $\bar{q}_{t}$, will be a Teichmüller geodesic. We will then use a theorem of Minsky [Min96] to show that the distance in $\mathcal{T}(Y)$ between $\mathrm{FN}_{Y}(t)$ and $G_{Y}(t)$ is uniformly bounded, depending only on $M_{0}$. We begin by examining the boundary of $Y$.

For every curve $\alpha$, the $q_{t}$-geodesic representatives of $\alpha$ form a (possibly degenerate) flat annulus $F_{t}(\alpha)$. Let $\mathrm{Y}_{t}$ be a representative of the isotopy class of $Y$ that has $q_{t}$-geodesic boundaries and that is disjoint from the interior of $F_{t}(\alpha)$ for every curve $\alpha \subset \partial Y$. For $t \in J_{Y}$ and $r=M_{t}(\alpha, Y) / 2$, the $r$-regular neighborhood of a boundary component of $Y$ inside $\mathrm{Y}_{t}$ is an annulus $A=A_{t}(\alpha)$. We call the boundary component of $A$ that is a $q_{t}$-geodesic representative of $\alpha$ the inner boundary. Along the inner boundary we attach a foliated locally flat punctured disk as follows (the construction is identical to the one given in [Raf07] and is repeated here for completeness):

Let $A^{\prime}$ be the double cover of $A$, and let $q_{t}^{\prime}$ be the lift of $q_{t}$. Let $x_{1}, \ldots, x_{n}$ be the points on the inner boundary of $A$ which have angle $\theta_{i}>\pi$ in $\mathrm{Y}_{t}$. Note that this set is nonempty: if not, $\mathrm{Y}_{t}$ meets the interior of the flat annulus $F_{t}(\alpha)$, a contradiction. Denote the lifts of $x_{i}$ by $y_{i}$ and $z_{i}$. We now fill the inner boundary of $A^{\prime}$ by symmetrically adding $2(n-1)$ Euclidean triangles to obtain a disk, $D^{\prime}$, equipped with a singular flat structure such that the total angle at each point is a multiple of $\pi$ and is at least $2 \pi$.

We start by attaching a Euclidean triangle to vertices $y_{1}, y_{2}, y_{3}$, which we denote by $\triangle\left(y_{1}, y_{2}, y_{3}\right)$ (see Figure 12). We choose the angle at vertex $y_{2}, \angle y_{2}$, so that the total angle at $y_{2}, \theta_{2}+\angle y_{2}$, is a multiple of $\pi$. Assuming $0 \leq \angle y_{2}<\pi$, there is a unique such triangle. Attach an isometric triangle to $z_{1}, z_{2}, z_{3}$. Now consider points $y_{1}, y_{3}, y_{4}$. Again, there


Figure 12. The filling of the annulus $A^{\prime}$
exists a Euclidean triangle with one edge equal to the newly introduced segment $\left[y_{1}, y_{3}\right]$, another edge equal to the segment $\left[y_{3}, y_{4}\right]$ and an angle at $y_{3}$ that makes the total angle at $y_{3}$, including the contribution from the triangle $\triangle\left(y_{1}, y_{2}, y_{3}\right)$, a multiple of $\pi$. Attach this triangle to vertices $y_{1}, y_{3}, y_{4}$ and an identical triangle to vertices $z_{1}, z_{3}, z_{4}$. Continue in this fashion until finally adding triangles $\triangle\left(y_{1}, y_{n}, z_{1}\right)$ and $\triangle\left(z_{1}, z_{n}, y_{1}\right)$. Because of the symmetry, the two edges connecting $y_{1}$ and $z_{1}$ have equal length, and we can glue these together. We call the union of the added triangles $D^{\prime}$. Notice that the involution on $A^{\prime}$ extends to $D^{\prime}$. Let $D$ be the quotient of $D^{\prime}$, and note that $D$ is a punctured disk attached to the inner boundary of $A$.

For $i \neq 1$, the total angles at $y_{i}$ and at $z_{i}$ are multiples of $\pi$ and are larger than $\theta_{i}>\pi$; therefore, they are at least $2 \pi$. We have added $2(n-1)$ triangles. Hence, the sum of the total angles of all vertices is $2 \sum_{i} \theta_{i}+2(n-1) \pi$, which is a multiple of $2 \pi$. Therefore, the sum of the angles at $y_{1}$ and $z_{1}$ is also a multiple of $2 \pi$. But they are equal to each other, and each one is larger than $\pi$. This implies that they are both at least $2 \pi$.

It follows that the quadratic differential $q_{t}^{\prime}$ extends over $D^{\prime}$ symmetrically with quotient $\bar{q}_{t}$, an extension of $q_{t}$ to $D$. Then $\bar{q}_{t}$ is a quadratic differential on $Y$. Denote the Riemann surface obtained after capping off all boundaries of $\mathrm{Y}_{t}$ by $\bar{Y}_{t}$. This completes the construction.

It remains to show that the distance in $\mathcal{T}(Y)$ between $\mathrm{FN}_{Y}(G((t))$ and $G_{Y}(t)$ is uniformly bounded. For this, we examine the extremal lengths of curves in two metrics. We say two quantities $P$ and $Q$ are comparable if they are the same up to up a multiplicative error depending on the topology of $S$ only and we write $P \asymp Q$.

We know from [Min96] that, for any essential curve $\gamma$ in $Y$, the extremal lengths of $\gamma$ in $G(t)$ and in $\mathrm{FN}_{Y}(G(t))$ are comparable:

$$
\begin{equation*}
\operatorname{Ext}(\gamma, G(t)) \asymp \operatorname{Ext}\left(\gamma, \operatorname{FN}_{Y}(G(t))\right) \tag{A.5}
\end{equation*}
$$

(See the proof of Theorem 6.1 in [Min96], page 283, line 19.) We need to show that the extremal lengths of $\gamma$ in $G(t)$ and in $G_{Y}(t)$ are comparable as well. For this, we need the following lemma (see [Min96, Lemma 4.2]).
Lemma A. 6 (Minsky). Let $\Sigma$ be any Riemann surface. There exists a constant $m$, depending only on the topological type of $\Sigma$, such that if $Y \subset \Sigma$ is a subsurface of $S$ with negative Euler characteristic for which each component $\alpha$ of $\partial Y$ bounds an annulus $A_{\alpha}$ in $Y$, with modulus $\operatorname{Mod}\left(A_{\alpha}\right) \geq m$, then for any essential $\gamma$ in $Y$,

$$
\operatorname{Ext}(\gamma, Y) \doteqdot \operatorname{Ext}(\gamma, \Sigma)
$$

After choosing $M_{0}$ large enough, $\mathrm{Y}_{t}$ satisfies the condition of the above lemma, for every $t \in J_{Y}$. In fact, for $\alpha \subset \partial Y$ and $r=$ $q_{t}(\alpha) M_{t}(\alpha, Y)$, the $r$-neighborhood of $\alpha$ in $\mathrm{Y}_{t}$ is an annulus and has a modulus $\log M_{t}(\alpha, Y)$ [Raf05, Lemma 3.6]. Choosing $M_{0}$ large enough, we can assume that, for $t \in J_{Y}, \log M_{t}(\alpha, Y) \geq m_{0}$.

Now, considering $\mathrm{Y}_{t}$ as a subsurface of $G(t)$ and as a subsurface of $\overline{\mathrm{Y}}_{t}=G_{Y}(t)$, Lemma A. 6 implies that

$$
\operatorname{Ext}(\gamma, G(t)) \doteqdot \operatorname{Ext}\left(\gamma, \mathrm{Y}_{t}\right) \doteqdot \operatorname{Ext}\left(\gamma, G_{Y}(t)\right)
$$

This and (A.5) finishes the proof of Theorem A.2.

## References

[BDM] Jason Behrstock, Cornelia Drutu, and Lee Mosher. Thick metric spaces, relative hyperbolicity, and quasi-isometric rigidity. arXiv:math.GT/0512592.
[Beh04] Jason Behrstock. Asymptotic geometry of the mapping class group and Teichmüller space. PhD thesis, SUNY Stony Brook, 2004. http://www.math.columbia.edu/~jason/thesis.pdf.
[BF] Jeffrey Brock and Benson Farb. Curvature and rank of Teichmüller space. arXiv:math.GT/0109045.
[BH99] Martin R. Bridson and André Haefliger. Metric spaces of non-positive curvature. Springer-Verlag, Berlin, 1999.
[BM04] Tara E. Brendle and Dan Margalit. Commensurations of the Johnson kernel. Geom. Topol., 8:1361-1384 (electronic), 2004. arXiv:math.GT/0404445.
[Bow91] B. H. Bowditch. Notes on Gromov's hyperbolicity criterion for pathmetric spaces. In Group theory from a geometrical viewpoint (Trieste, 1990), pages 64-167. World Sci. Publ., River Edge, NJ, 1991.
[Bro03] Jeffrey F. Brock. The Weil-Petersson metric and volumes of 3-dimensional hyperbolic convex cores. J. Amer. Math. Soc., 16(3):495-535 (electronic), 2003. arXiv:math.GT/0109048.
[CDP90] M. Coornaert, T. Delzant, and A. Papadopoulos. Géométrie et théorie des groupes. Springer-Verlag, Berlin, 1990. Les groupes hyperboliques de Gromov.
[CR05] Young-Eun Choi and Kasra Rafi. Comparison Between Teichmuller and Lipschitz Metrics, 2005. math.GT/0510136.
[Gro87] Mikhael Gromov. Hyperbolic groups. In Essays in group theory, pages 75-263. Springer, New York, 1987.
[Har81] Willam J. Harvey. Boundary structure of the modular group. In Riemann surfaces and related topics: Proceedings of the 1978 Stony Brook Conference (State Univ. New York, Stony Brook, N.Y., 1978), pages 245-251, Princeton, N.J., 1981. Princeton Univ. Press.
[Hem01] John Hempel. 3-manifolds as viewed from the curve complex. Topology, 40(3):631-657, 2001. arXiv:math.GT/9712220.
[HT80] A. Hatcher and W. Thurston. A presentation for the mapping class group of a closed orientable surface. Topology, 19(3):221-237, 1980.
[Kob88] Tsuyoshi Kobayashi. Heights of simple loops and pseudo-Anosov homeomorphisms. In Braids (Santa Cruz, CA, 1986), pages 327-338. Amer. Math. Soc., Providence, RI, 1988.
[Lea] Jason Leasure. Geodesics in the complex of curves of a surface. Ph.D. thesis.
[McC91] Darryl McCullough. Virtually geometrically finite mapping class groups of 3-manifolds. J. Differential Geom., 33(1):1-65, 1991.
[Min96] Yair N. Minsky. Extremal length estimates and product regions in Teichmüller space. Duke Math. J., 83(2):249-286, 1996. arXiv:math.GT/9407222.
[MM] Howard A. Masur and Yair N. Minsky. Quasiconvexity in the curve complex. arXiv:math.GT/0307083.
[MM99] Howard A. Masur and Yair N. Minsky. Geometry of the complex of curves. I. Hyperbolicity. Invent. Math., 138(1):103-149, 1999. arXiv:math.GT/9804098.
[MM00] Howard A. Masur and Yair N. Minsky. Geometry of the complex of curves. II. Hierarchical structure. Geom. Funct. Anal., 10(4):902-974, 2000. arXiv:math.GT/9807150.
[Mos] Lee Mosher. Train track expansions of measured foliations. http://newark.rutgers.edu/~mosher/.
[Pen88] Robert C. Penner. A construction of pseudo-Anosov homeomorphisms. Trans. Amer. Math. Soc., 310(1):179-197, 1988.
[Raf05] Kasra Rafi. A characterization of short curves of a Teichmüller geodesic. Geom. Topol., 9:179-202 (electronic), 2005. arXiv:math.GT/0404227.
[Raf06] Kasra Rafi. A combinatorial model for the Teichmüller metric. To appear in Geometric and Functional Analysis, 2006. arXiv:math.GT/0509584.
[Raf07] Kasra Rafi. Thick-thin decomposition for quadratic differentials. To appear in Mathematics Research Letters, 2007. http://www.math.uconn.edu/~rafi/research/research.html/TT.pdf.
[Sch] Saul Schleimer. Notes on the curve complex. http://math.rutgers.edu/~saulsch/Maths/notes.pdf.
[Sch82] Martin Scharlemann. The complex of curves on nonorientable surfaces. J. London Math. Soc. (2), 25(1):171-184, 1982.
[Thu88] William P. Thurston. On the geometry and dynamics of diffeomorphisms of surfaces. Bull. Amer. Math. Soc. (N.S.), 19(2):417-431, 1988.
[Thu98] William P. Thurston. Minimal stretch maps between hyperbolic surfaces, 1998. arXiv:math.GT/9801039.

Department of Mathematics, University of Illinois, Chicago, Illinois 60607

E-mail address: masur@math.uic.edu
URL: http://www.math. uic.edu/~masur
Department of Mathematics, Rutgers University, Piscataway, New
Jersey 08854
E-mail address: saulsch@math.rutgers.edu
URL: http://www.math.rutgers.edu/~saulsch


[^0]:    Date: January 23, 2007.
    This work is in the public domain.

