LARGE SCALE RANK OF TEICHMÜLLER SPACE

ALEX ESKIN, HOWARD MASUR, AND KASRA RAFI

Abstract. Let $\mathcal{X}$ be quasi-isometric to either the mapping class group equipped with the word metric, or to Teichmüller space equipped with either the Teichmüller metric or the Weil-Petersson metric. We introduce a unified approach to study the coarse geometry of these spaces. We show that the quasi-Lipschitz image in $\mathcal{X}$ of a box in $\mathbb{R}^n$ is locally near a standard model of a flat in $\mathcal{X}$. As a consequence, we show that, for all these spaces, the geometric rank and the topological rank are equal. The methods are axiomatic and apply to a larger class of metric spaces.

1. Introduction

In this paper we study the large scale geometry of several metric spaces: the Teichmüller space $T(S)$ equipped with the Teichmüller metric $d_T$, the Teichmüller space equipped with the Weil-Petersson metric $d_{WP}$ and the mapping class group Mod$(S)$ equipped with the word metric $d_W$. (Brock [Bro03] showed that the Weil-Petersson metric is quasi-isometric to the pants complex). Even though the definitions of distance in these spaces are very different, they share a key feature, namely, an inductive structure. That is, they are a union of product regions associated to lower complexity surfaces with the gluing pattern given by the curve complex.

Let $S$ be a possibly disconnected surface of finite hyperbolic type and let $\mathcal{X} = \mathcal{X}(S)$ be a metric space that is quasi-isometric to one of the metric spaces mentioned above. The usual starting point of understanding the large scale geometry of a metric space is to understand its flats. By a flat here we mean a quasi-isometric image of Euclidean space. We analyze quasi-Lipschitz maps from a large box $B \subset \mathbb{R}^n$ into $\mathcal{X}$. Our goal is to give a description of the image of such a map on a large sub-box $B' \subset B$. We show that the image of $B'$ looks like a standard flat up to a small linear error. A standard flat is a product of preferred paths associated to disjoint subsurfaces of $S$ (see Definition 4.4 and Definition 7.1).

Our main theorem is the following.
**Theorem A** (Image of a box is locally standard). Let $X$ be either of Teichmüller space $\mathcal{T}(S)$ equipped with the Teichmüller metric $d_T$, Teichmüller space with the Weil-Petersson metric $d_{WP}$, (or pants graph) or the mapping class group $\text{Mod}(S)$ equipped with the word metric $d_W$. For all $K, C$ and for all $R_0, \epsilon_0$ there exists $R_1$ such that if $B$ is a box of size at least $R_1$ and $f: B \to X$ is a $(K, C)$–quasi-Lipschitz map, then there is a sub-box $B' \subset B$ of size $R' \geq R_0$ such that $f(B')$ lies inside an $O(\epsilon_0 R')$–neighborhood of a standard flat in $X$.

As a corollary, we determine the large scale rank of the space $X$. Define the *topological rank* of $X$, $\text{rank}_{\text{top}}(X)$, to be largest integer $r$ so that there are pairwise disjoint *essential* subsurfaces $W_1, \ldots, W_r$ in $S$. (In all cases, a thrice-punctured sphere $W$ is considered inessential. Also, as we shall see, when $X$ is quasi-isometric to $(\mathcal{T}(S), d_{WP})$, annuli are also considered inessential.) The dimension of a standard flat in $X$ is at most $\text{rank}_{\text{top}}(X)$.

**Theorem B** (Geometric rank). For every $K$ and $C$, there is a constant $R_2$ so that if $B$ is a box of size at least $R_2$ in $\mathbb{R}^n$ and $f: B \to X$ is a $(K, C)$–quasi-isometric embedding, then $n \leq \text{rank}_{\text{top}}(X)$. Furthermore, for $n = \text{rank}_{\text{top}}(X)$, there is a quasi-isometric embedding of $\mathbb{R}^n$ into $X$.

Define the *geometric rank* of $X$, $\text{rank}_{\text{geo}}(X)$, to be largest integer $n$ so that there is a quasi-isometric embedding $f: \mathbb{R}^n \to X$. Also, let $g$ be the genus of $S$, $p$ be the number of punctures of $S$ and $c$ be the number of component of $S$.

**Corollary C.** The topological and the geometric rank of $X$ are equal. Namely, if $X$ is either $(\text{Mod}(S), d_W)$ or $\mathcal{T}(S), d_T$ then

$$\text{rank}_{\text{geo}}(X) = \text{rank}_{\text{top}}(X) = 3g + p - 3c,$$

and if $X$ is $(\mathcal{T}(S), d_{WP})$, then

$$\text{rank}_{\text{geo}}(X) = \text{rank}_{\text{top}}(X) = \left\lfloor \frac{3g + p - 2c}{2} \right\rfloor.$$

**Remark** 1.1. Theorem A and Theorem B hold for a larger class of metric spaces than are discussed above. Essentially, one needs a mapping class group action and a distance formula similar to Equation (7) (see Masur-Schleimer [MS13] for examples of such distance formulas). The definition of an essential surface has to be modified to mean any type of surface that appears in the associated distance formula.

For many such spaces, e.g, the arc complex and the disk complex are known to be Gromov hyperbolic and therefore have geometric rank one [MS13], hence the corollary is already known. Others, such as the Hatcher-Thurston complex and the separating curve complex, are not Gromov hyperbolic and our discussion applies to prove the geometric rank and topological rank are equal. These complexes have been omitted to simplify the exposition.
History. The idea of studying the rank of these objects was introduced by Brock-Farb [BF06]. In the case when $\mathcal{X}$ is the pants graph Corollary [C] was first proven in that paper when the surface is the twice punctured torus. They also showed that the topological rank is always at most as large as the geometric rank and conjectured Corollary [C] for all genera. Corollary [C] was then proven for all genera in the case when $\mathcal{X}$ is quasi-isometric to the mapping class group with the word metric or Teichmüller space with the Weil-Petersson metric by Behrstock-Minsky and Hamenstädt [BM08, Ham07].

The rank statement is used to prove the quasi-isometric rigidity of $\text{Mod}(S)$ by Behrstock-Kleiner-Minsky-Mosher in [BKMM08] and by Hamenstädt in [Ham07]. The case of Teichmüller space with the Teichmüller metric had not been studied previously. Note that the map that sends $\text{Mod}(S)$ to the orbit of a point in $\mathcal{T}(S)$ is not a quasi-isometry or even a quasi-isometric embedding because of the thin regions in $\mathcal{T}(S)$ which locally look like products of horoballs. Unlike [BM08] which uses asymptotic cones, our approach as outlined below is to study the local behavior of a quasi-Lipschitz maps.

Main tools. To prove our theorems we need to adopt/modify some tools that already exist in the literature. The first one is the idea of coarse differentiation. This was introduced in the context of geometric group theory by Eskin-Fisher-Whyte [EFW06, EFW07] (see references in that paper for its use in other contexts) and used to prove quasi-isometric rigidity of lattices in Sol and in the quasi-isometry classification of lamplighter groups. The statement they used is similar to Theorem 2.7 below which holds for quasi-Lipschitz maps between more general metric spaces. However, since we are mostly concerned with maps where the domain is a subset of $\mathbb{R}^n$, we prove the following statement which is cleaner and easier to use.

**Theorem D** (Coarse Differentiation). For every $\epsilon_0$ and $R_0$ there is $R_1$ so that for $R \geq R_1$ the following holds. Let $f : B \to Y$ be a quasi-Lipschitz map where $B$ is a box of size $R$ in $\mathbb{R}^n$. Then, there is a box $B' \subset B$ of size $R' > R_0$ so that $f$ restricted to $B'$ is $\epsilon_0$-efficient on scale $R'$.

Here, efficient means that the reverse triangle inequality is satisfied up to a small multiplicative error. That is; lines are mapped to lines up to a sub-linear error. One should think of the above theorem as a coarse version of the Rademacher’s theorem:

**Theorem E** (Rademacher). Suppose $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ is lipschitz, then $\varphi$ is differentiable almost everywhere.

In Theorem D the sub-box $B'$ is an analogue of a point of differentiability. The power of this theorem lies in the fact that one can add the assumption of efficiency for free, just by replacing $B$ with a sub-box $B'$.

We also use the construction of Bestvina-Bromberg-Fujiwara [BBF10]. They embed the mapping class group into a product of finitely many hyperbolic spaces. However, their construction is axiomatic and can be adopted
easily to embed any of our spaces into a product of finitely many hyperbolic spaces. The $L^1$-metric on this product induces a metric on the space $\mathcal{X}$ and this is the metric with respect to which we apply the coarse differentiation theorem. Note that, the notion of efficiency is not preserved under quasi-isometry and the choice of metric here is essential. The key is that the projection to each factor of an efficient map into a product space is still efficient. And efficient maps into hyperbolic spaces are easier to control (Proposition 2.11).

Finally, we use the realization theorem of Behrstock-Minsky-Mosher-Kleiner [BKMM08]. They provide a description of the image of the mapping class group in the product of curve complexes. We adopt it to provide the description of the image of $\mathcal{X}$. This is necessary to translate back the information obtained in each hyperbolic factor to information in $\mathcal{X}$.

**Outline of the paper.** Section 2 is devoted to the development of coarse differentiation theory and to the discussion of efficient maps. The main result is Theorem 2.4 as discussed above. We also establish the basic properties of efficient maps and prove that efficient paths in Gromov hyperbolic spaces stay close to geodesics; an analogue of the Morse Lemma.

In §3 we discuss the combinatorial model for each of the spaces considered in the paper. The three seemingly different metric spaces above have very similar models. Namely, Teichmüller space equipped with the Weil-Petersson metric is quasi-isometric to the pants graph [Bro03]. The mapping class group is quasi-isometric to the marking graph [MM00] by work of Masur-Minsky and Teichmüller space equipped with the Teichmüller metric is quasi-isometric to the space of augmented markings by work of Rafi and Durham. [Raf10, Dur13].

The advantage of this approach is that we can measure relative complexity of two points $x, y \in \mathcal{X}$ from the point of view of a subsurface $W$. This is the distance in the curve complex of $W$ between the projections of $x$ and $y$ to $W$. The curve complex of every surface $W$ is known to be Gromov hyperbolic. We then define a coarse metric on each of these combinatorial models using a distance formula which is the sum over relative complexity from the point of view of different subsurfaces. Since we work in the category of spaces up to quasi-isometries, the distance needs to be defined only up a multiplicative error.

In Section 4 we introduce the notion of preferred paths. These are paths whose projections to every curve complex is a quasi-geodesic and they replace the notion of geodesics in our spaces. The main statement in the section is Theorem 4.9 which shows that an efficient path stays near a preferred path joining its endpoints. Hence, the outcome of the Coarse Differentiation Theorem is indeed a box where straight lines are mapped to straight lines up to the first order. This is the key tool for the rest of the paper. The proof uses the construction in [BBF10] which allows one to embed $\mathcal{X}$ into a product of hyperbolic spaces. The projection of the efficient paths into each
factor stays near a geodesic in that factor. We then use this and consistency theorem (Theorem 3.2) to build the preferred path in $X$ tracing the given efficient path.

Section 5 establishes some properties of preferred paths. The main ones are fellow traveling properties that say that under certain conditions, preferred paths that begin and end near the same point fellow travel in the middle (Proposition 5.8 and Proposition 5.9). These statements are used in the succeeding sections to build big boxes with the required properties. In §6 the main inductive step is proven (Theorem 6.1) and in §7 we assemble the proofs of the main theorems.

**Treatment of constants.** Suppose that $Y$ and $Z$ are geodesic metric spaces. We say a map $\varphi: Y \to Z$ is quasi-Lipschitz if there are constants $K$ and $C$ so that

$$d_Z(\varphi(x_1), \varphi(x_2)) \leq K d_Y(x_1, x_2) + C.$$  

We fix constants $K$ and $C$ once and for all. We also fix an upper-bound for the complexity of the surface $S$ and the dimension $n$. When we say a constant is uniform, we mean its value depends only on $K$, $C$, the topology of $S$ and the value of $n$ only. Similarly, we will use terms like quasi-isometric embedding or quasi-isometry to mean that the associated constants are the same as $K$ and $C$ fixed above.

To simplify presentation, we try to avoid naming uniform constants whenever possible. Instead, we adopt the following notations. Let $a$ and $b$ represent various quantities and let $M$ and $C'$ be uniform constants. We say $a$ is less than $b$ up to a multiplicative error, $a \prec b$, if $a \leq M b$. We say $a$ and $b$ are comparable, $a \preceq b$, if we have both $a \prec b$ and $b \prec a$.

Using the similar notation when the error is additive or both additive and multiplicative, we say $a \succ b$ if $a \leq b + C'$ and $a \succ b$ if $a \leq Ma + C'$. Again, $a \succ b$ if we have both $a \prec b$ and $b \prec a$ and $a \asymp b$ if we have both $a \prec b$ and $b \prec a$. Also, we often use the notation $a = O(b)$ to mean $a \preceq b$. For example

$$a \prec b + O(c) \iff a \leq b + Mc + C',$$

for uniform constants $M$ and $C'$.

Using this notation we may write

$$a \succ b \quad \text{and} \quad b \succ c \implies a \succ c.$$  

Here, the additive error in the last inequality is the sum of the additive errors in the first two inequalities and hence is still a uniform constant. That is, different occurrences of $\succ$ have different implied constants. But as long as we use statements of this type a uniformly bounded number of times, all the implied constant are still uniform.
2. Coarse Differentiation

Being differentiable means that to first order, lines are mapped to lines. We emulate this concept by introducing the notion of an $\epsilon$-efficient paths which is a path that satisfies the reverse triangle inequality up to a small multiplicative error.

**Definition 2.1.** Let $Z$ be a metric space, let $\gamma: [a, b] \to Z$ be quasi-Lipschitz and let $R > 0$ be a scale so that $|b - a| < R$. An $r$-partition of $[a, b]$ is a set of times $a = t_0 < t_1 < \ldots < t_m = b$ so that $(t_{i+1} - t_i) \leq r$. Let $z_i = \gamma(t_i)$. We define the coarse length of $\gamma$ on the scale $r$ to be

\[
\Delta(\gamma, r) = \min_{r\text{-partitions}} \sum_{i=0}^{m-1} d_Z(z_i, z_{i+1}).
\]

We say $\gamma$ is $\epsilon$-efficient on the scale $R$ if

\[
\Delta(\gamma, \epsilon R) \leq d_Z(\gamma(a), \gamma(b)) + O(\epsilon R).
\]

We establish some elementary properties of efficient paths.

**Lemma 2.2.** Consider a map $\gamma: [a, b] \to Z$.

1. For times $a = s_0, \ldots, s_k = b$, where $k$ is a uniformly bounded integer, we have, if $\gamma$ is $\epsilon$-efficient then

\[
\sum_{i=0}^{k-1} d_Z(\gamma(s_i), \gamma(s_{i+1})) = O(\epsilon R).
\]

2. For $[c, d] \subset [a, b]$, if $\gamma$ is $\epsilon$-efficient at scale $R$ so is $\gamma' = \gamma|_{[c, d]}$.

3. Assume $Z = Z_1 \times \ldots \times Z_l$ equipped with the $L^1$-metric and let $\gamma_i$ be the projection of $\gamma$ to $Y_i$. Then, if $\gamma$ is $\epsilon$-efficient at scale $R$ so is every $\gamma_i$.

4. If $Z'$ is a subset of $Z$ and $\gamma$ is an $\epsilon^2$-efficient path that is contained in an $O(\epsilon^2 R)$-neighborhood of $Z'$ then the closest point projection of $\gamma$ to $Z'$ is an $\epsilon$-efficient path.

**Proof.** Let $t_0, \ldots, t_m$ be the $\epsilon R$-partition of $[a, b]$ achieving the minimum in the definition of $\Delta(\gamma, \epsilon R)$. Adding a point $s_j$ to the partition will increase the sum in Equation (1) by at most $O(\epsilon R)$. In fact, if $t_i \leq s_j \leq t_{i+1}$ then $|t_i - t_{i+1}| \leq \epsilon R$ and, since $\gamma$ is quasi-Lipschitz,

\[
d_Z(\gamma(t_i), \gamma(s_j)) + d_Z(\gamma(s_j), \gamma(t_{i+1})) = O(\epsilon R).
\]

Since the number of points $s_j$ is uniformly bounded, adding all times $s_j$ to the partition will increase the sum by at most $O(\epsilon R)$. Now, removing all $t_i$, will only decrease the sum and hence part one of the lemma holds.

To see the second part, as above, let $t_0, \ldots, t_m$ be a set of times where the sum in Equation (1) is nearly minimal and so that the times $c$ and $d$ are
included in the set \( \{ t_i \} \). Letting \( c = t_j \) and \( d = t_k \) we have

\[
\Delta(\gamma, \epsilon R) \geq \sum_{i=1}^{m} d_Z(z_i, z_{i+1}) - O(\epsilon R)
\]

\[
\geq d_Z(\gamma(a), \gamma(c)) + \sum_{i=j}^{k-1} d_Z(z_i, z_{i+1}) + d_Z(\gamma(d), \gamma(b)) - O(\epsilon R)
\]

\[
\geq d_Z(\gamma(a), \gamma(c)) + \Delta(\gamma', \epsilon R) + d_Z(\gamma(d), \gamma(b)) - O(\epsilon R).
\]

Also, by definition,

\[
\Delta(\gamma, \epsilon R) \leq d_Z(\gamma(a), \gamma(b)) + O(\epsilon R).
\]

Hence,

\[
\Delta(\gamma', \epsilon R) \leq d_Z(\gamma(a), \gamma(b)) - d_Z(\gamma(a), \gamma(c)) - d_Z(\gamma(d), \gamma(b)) + O(\epsilon R)
\]

\[
\leq d_Z(\gamma(c), \gamma(d)) + O(\epsilon R).
\]

This finishes the proof of part 2.

We prove the third part for \( l = 2 \). The general case is similar. Consider the partition \( t_0, \ldots, t_m \) that achieves the minimum for \( \Delta(\gamma, \epsilon R) \). Since \( Z \) is equipped with the \( L^1 \)-metric, we have

\[
\sum_{i=1}^{m} d_{Z_1}(z_i, z_{i+1}) + d_{Z_2}(z_i, z_{i+1}) \leq d_{Z_1}(\gamma(a), \gamma(b)) + d_{Z_2}(\gamma(a), \gamma(b)) + O(\epsilon R).
\]

But, by triangle inequality, we have

\[
\sum_{i=1}^{m} d_{Z_2}(z_i, z_{i+1}) \geq d_{Z_2}(\gamma(a), \gamma(b)).
\]

Subtracting Equation (4) from Equation (3) we obtain

\[
\Delta(\gamma_1, \epsilon R) \leq \sum_{i=1}^{m} d_{Z_1}(z_i, z_{i+1}) \leq d_{Z_1}(\gamma(a), \gamma(b)) + O(\epsilon R).
\]

To see the last part, again let \( t_1, \ldots, t_m \) be the optimal subdivision (note that, in this case, \( \gamma \) is \( \epsilon^2 \)-efficient). Choose a sub-partition \( s_1, \ldots, s_l \) so that

\[
|s_{i+1} - s_i| \lessapprox \epsilon R,
\]
Then, $l \lesssim \frac{1}{R}$. Also, let $\gamma'$ be the path obtained from composing $\gamma$ with the closest point projection to $Z'$. Let $z_i = \gamma(s_i)$ and $z'_i = \gamma'(s_i)$. Then

$$\Delta(\gamma', \epsilon R) \leq \sum_{i=1}^{l} d_{Z'}(z'_i, z'_{i+1})$$

$$\leq \sum_{i=1}^{l} d_{Z}(z'_i, z_i) + d_{Z}(z_i, z_{i+1}) + d_{Z}(z_{i+1}, z'_{i+1})$$

$$\leq \Delta(\gamma, \epsilon R) + l \cdot O(\epsilon^2 R)$$

$$\leq \Delta(\gamma, \epsilon^2 R) + O(\epsilon R) \leq d_{Z'}(z'_1, z'_l) + O(\epsilon R).$$

In the last inequality, we used the fact that the pairs $z_1, z'_1$ and $z_l, z'_l$ are $\epsilon^2 R$-close. This finishes the proof. \qed

**Definition 2.3.** For a geodesic metric space $Y$, we say a set $B \subset Y$ is a **box** of size $R$ if there is a point $y \in Y$ and radii $R_l$ and $R_u$ so that $R_l \asymp R_u$ and

$$N_{R_l}(y) \subset B \subset N_{R_u}(y).$$

The **size** of the box $B$ is any number $R$ with $R_l \asymp R_u \asymp R_u$.

A map $f : B \to Z$ from a box of size $R$ in $Y$ to a metric space $Z$ is **$\epsilon$-efficient** if, for any geodesic $\gamma : [a, b] \to B$, the path $f \circ \gamma$ is $\epsilon$-efficient at scales $R$.

**Theorem 2.4** (Coarse Differentiation). For every $\epsilon_0$ and $R_0$ there is $R_1$ so that for $R \geq R_1$ the following holds. Let $f : B \to Z$ be a quasi-Lipschitz map where $B$ is a box of size $R$ in $\mathbb{R}^n$. Then, there is a box $B' \subset B$ of size $R' \asymp R_0$ so that $f$ restricted to $B'$ is $\epsilon_0$-efficient.

**Remark 2.5.** Note that the size of the error, $\epsilon_0 R'$, depends on the size of the box. An $\epsilon_0$-efficient map from a much larger box is allowed to have a much larger error. What we control is the size of the error at a proportion of the size of the box.

**2.1. Choosing the correct scales.** We first prove a much coarser version of Theorem 2.4 for quasi-Lipschitz maps between more general metric spaces (Theorem 3 below). The statement of Theorem 3 is more complicated but the proof is straightforward and is a direct analogue of proof of Theorem 1. Then, we clean up the statement when the domain in $\mathbb{R}^n$ to obtain Theorem 2.4.

**Definition 2.6.** A family of geodesics $\mathcal{F}$ in a metric space $Y$ is a **locally finite family** if, for any compact subset $B$ of $Y$, only finitely many geodesics in $\mathcal{F}$ intersect $B$.
Let $B$ be a box of size $L$. Define $\mathcal{F}_B$ to be the collection of restrictions of paths in $\mathcal{F}$ to $B$ that are long. More precisely, let

$$\mathcal{F}_B = \{ \gamma \mid \gamma \text{ is a connected component of } \gamma' \cap B, \gamma' \in \mathcal{F}, \ |\gamma| \gtrsim L \}. $$

For any scale $R$, let $\mathcal{F}_{B,R}$ be the collection of geodesic segments obtained by subdividing $\gamma \in \mathcal{F}_B$ into segments of length $R$. Any segment of length less than $R$ in the end is ignored.

**Theorem 2.7.** Let $\mathcal{F}$ be a locally finite family of geodesics in $Y$. For any $\epsilon > 0$ and $\theta > 0$, there exist a constant $L_0$ such that the following holds. Let $L > L_0$, $B \subset Y$ be a box of size $L$ and $\varphi: B \to \mathbb{Z}$ be a quasi-Lipschitz map. Then, there exist a scale $R \leq L_0$ such that, for at least $(1 - \theta)$ fraction of segments $\gamma \in \mathcal{F}_{B,R}$, the path $\varphi \circ \gamma$ is $\epsilon$-efficient on scale $R$.

Morally, the lemma can be considered as stating that $\varphi|_B$ is nearly affine on scales $R$, up to an error of $O(\epsilon R)$.

**Remark 2.8.** In the lemma $R$ depends on $\epsilon, \theta, K, C$, and also on $B$. However the proof will find $R$ as one of finitely many values as long as $\epsilon, \theta, K$ and $C$ are fixed.

**Proof of Theorem 2.7.** Pick $r_0 \geq C$ (the additive error in the the definition of a quasi-Lipshitz map) and inductively let

$$r_m = \frac{r_{m-1}}{\epsilon}.$$ 

Let $M$ be a large positive integer (to be determined below) and let $L_0 = r_M$. For $m = 1, \ldots , M$, and $\gamma \in \mathcal{F}_B$, subdivide $\gamma$ into segments of length $r_m$. Let $\theta_m(\gamma)$ be the fraction of these segments for which the image under $\varphi$ is not $\epsilon$-efficient on the scale $r_m$.

Suppose $\tilde{\gamma}$ is one of these sub-segments of $\gamma$ with length $r_m$. Notice that by definition $\Delta(\tilde{\gamma}, r_m) = d_Y(\phi(x_m), \phi(y_m))$ where $x_m, y_m$ are the endpoints of $\tilde{\gamma}$. If $\varphi(\tilde{\gamma})$ is not $\epsilon$-efficient on the scale $r_m$, then

$$\Delta(\tilde{\gamma}, r_{m-1}) - \Delta(\tilde{\gamma}, r_m) \gtrsim \epsilon r_m.$$ 

Adding over all sub-segments of $\gamma$ we obtain

$$\Delta(\gamma, r_{m-1}) - \Delta(\gamma, r_m) \gtrsim \epsilon L \theta_m(\gamma).$$ 

After iterating this over $m$ as $m$ goes from $M$ down to 1 we get

$$\Delta(\gamma, r_0) - \Delta(\gamma, r_M) \gtrsim \epsilon L \sum_{m=1}^{M} \theta_m(\gamma).$$

Using the fact that $\phi$ is quasi-Lipschitz and $r_0 > C$ we have

$$\Delta(\gamma, r_0) \leq \frac{L}{r_0}(Kr_0 + C) \lesssim KL.$$
Hence,
\[ KL \geq \epsilon L \sum_{m=1}^{M} \theta_m(\gamma). \]
And thus
\[ \sum_{m=1}^{M} \theta_m(\gamma) \lesssim \frac{K}{\epsilon}. \]
Average over all geodesics \( \gamma \in \mathcal{F}_B \) to get
\[ \sum_{m=1}^{M} \left( \frac{1}{|\mathcal{F}_B|} \sum_{\gamma \in \mathcal{F}_B} \theta_m(\gamma) \right) \lesssim \frac{2K}{\epsilon}. \]
Choosing \( M \) large enough, we can ensure that for some \( m \), the quantity in the parenthesis is less than \( \theta \). But that is exactly the proportion of segments in \( \mathcal{F}_{B,r_m} \) whose image is not \( \epsilon \)-efficient at scale \( r_m \).

**Proof of Theorem 2.4.** Let \( R_0 \) and \( \epsilon_0 < 1 \) be given. Choose a family \( \mathcal{F} \) of geodesics in \( \mathbb{R}^n \) as follows: pick a finite set of directions \( \mathcal{D} \) in the unit sphere \( S^{n-1} \subset \mathbb{R}^n \) that is \( (\epsilon_0)^2 \) dense in \( S^{n-1} \) with the size \( |\mathcal{D}| \approx \frac{1}{(\epsilon_0)^2} \). For a direction \( \vec{v} \in \mathcal{D} \), let \( \mathcal{F}_{\vec{v}} \) be a family of parallel lines in the direction \( \vec{v} \) where the distance between nearby lines is comparable to 1. Then
\[ \mathcal{F} = \bigcup_{\vec{v} \in \mathcal{D}} \mathcal{F}_{\vec{v}} \]
is a locally finite family of geodesics in \( \mathbb{R}^n \). Let
\[ \epsilon = \epsilon_0^2, \quad \text{and} \quad \theta = \epsilon_0^5. \]
Apply Theorem 2.7 to obtain the constant \( L_0 \). We claim that the theorem holds for \( R_1 = L_0 \).

Now assume a box \( B \) of size \( L \geq L_0 \) and a quasi-Lipschitz map \( f: B \to \mathbb{Z} \) are given and let \( R \) be the scale obtained from Theorem 2.7. Let \( \mathcal{B} \) be a collection of disjoint sub-boxes of \( B \) so that

1. each box in \( \mathcal{B} \) is of size \( R \),
2. there is \( L' \approx L \) so that, for \( B' \in \mathcal{B} \), a neighborhood of radius \( L' \) around \( B' \) is contained in \( B \), and
3. the sum of the volumes of boxes in \( \mathcal{B} \) is a definite proportion of the volume of \( B \). This is equivalent to saying that a definite proportion of geodesics in \( \mathcal{F}_{B,R} \) intersect a box in \( \mathcal{B} \).

We will show that, for some box \( B' \in \mathcal{B} \), the map \( f|_{B'} \) is \( \epsilon_0 \)-efficient.

Let \( B' \in \mathcal{B} \) be a box that contains a geodesic \( \gamma \) that is not \( \epsilon_0 \)-efficient at scale \( R \). Let \( \vec{v} \) be the direction closest to the direction of \( \gamma \). Let
\[ \mathcal{N}(B', \vec{v}) \subset \mathcal{F}_{B,R} \]
be the set of geodesic segments in \( \mathcal{F}_{B,R} \) that are in the direction of \( \vec{v} \), intersect \( B' \) and are not \( \epsilon \)-efficient on scale \( R \).
Claim: Every geodesic in $F_{\vec{v}}$ that intersects an $\epsilon R$–neighborhood of $\gamma$ contains a segment in $\mathcal{NE}(B',\vec{v})$.

![Diagram](image)

**Proof of claim:** Assume $\beta \in F_{\vec{v}}$ intersects an $\epsilon R$–neighborhood of $\gamma$. The condition (2) of $B$ implies that $|\beta \cap B| \sim L$ and its subsegments of length $R$ are included in $F_{B,R}$. Since, the difference between the direction of $\gamma$ and $\beta$ is at most $\epsilon = \epsilon_2^0$, in fact $\gamma$ is contained in an $\epsilon R$–neighborhood of $\beta$. The segment $\gamma$ may have length longer than $R$ but by at most a uniform multiple. Hence, there is a segment $\beta_0 \subset \beta$ where $\gamma$ is included in an $\epsilon R$–neighborhood of $\beta_0$ and $\beta_0$ is a union of segments $\beta_1, \ldots, \beta_k \in F_{B,R}$ (refer to Fig. 1). We show that if $\gamma$ is not $\epsilon_0$–efficient on scale $R$, then $\beta_0$ will not be $\epsilon$–efficient on scale $R$. This would imply (Lemma 2.2) that at least one of the segments $\beta_j$ is not $\epsilon$–efficient on scale $R$ which is what we claimed.

Assume, for contradiction, that $\beta_0$ is $\epsilon$–efficient on scale $R$. Then every sub-segment of $\beta_0$ is also efficient on scale $R$ (Lemma 2.2). Hence, after taking a subsegment, we can assume that the end points of $\beta_0$ and $\gamma$ are $\epsilon R$–close. Consider the $r$–subdivision point $s_i, i = 1, \ldots, l$, along $\beta_0$ and find associated points $t_i$ along $\gamma$ so that $|t_i - s_i| \leq \epsilon R$. Note that, $l \approx \frac{R}{r} \approx \frac{1}{\epsilon_0}$. We have:

1. (Triangle inequality) $\Delta(\gamma, r) \leq \Delta(\beta_0, r) + \epsilon R \cdot l$
2. ($\beta_0$ is $\epsilon$–efficient) $\leq d_Y(f(s_0), f(s_l)) + O(\epsilon R) + \epsilon R \frac{1}{\epsilon_0}$
3. (the end points are close.) $\leq d_Y(f(t_0), f(t_l)) + O(\epsilon R) + \epsilon_0 R$.

To see the last inequality, note that since the end points of $\beta_0$ and $\gamma$ are close in $\mathbb{R}^n$, their images are $O(\epsilon R)$ close in $Y$ since $f$ is a quasi-Lipschitz map. This shows that $\gamma$ was $\epsilon_0$–efficient which is a contradiction. □

Let $F_{B',R}$ be the subset of $F_{B,R}$ consisting of segment that intersect $B'$. The ratio of the number of geodesics in $F_{\vec{v}}$ that intersect an $\epsilon R$–neighborhood of $\gamma$ to the number of geodesics in $F_{B',R}$ in the direction $\vec{v}$ is $\epsilon$. Hence, as a consequence of the claim, we have that if a box $B'$ contains
a non-efficient path, then
\[ |\mathcal{N}(\mathcal{E}(B', \vec{v}))| \gtrsim \frac{1}{|\mathcal{D}|} \times \epsilon \times |\mathcal{F}, B', R| \gtrsim \epsilon_0^4 |\mathcal{F}, B', R|. \]
But the proportion of non-efficient segments in \( B \) is \( \theta = \epsilon_0^5 \). Thus, not every box in \( B \) can contain a non-efficient segment. This finishes the proof. \( \square \)

2.2. Efficient map into a hyperbolic space. The following is the first use of efficient paths when the target is Gromov hyperbolic and is similar to the familiar Morse argument.

**Lemma 2.9.** Suppose \( X \) is a Gromov hyperbolic space and \( \gamma: [a, b] \to X \) is \( \epsilon \)-efficient on scale \( R \). Then \( \gamma \) stays in an \( O(\epsilon R) \)-neighborhood of a geodesic \( \ell \) joining \( \gamma(a) \) and \( \gamma(b) \).

**Proof.** Suppose, for a large \( M \), that the path \( \gamma \) leaves an \( M \epsilon R \)-neighborhood of \( \ell \). Fix a constant \( D_0 \geq C \) and let the times \( c, d \in [a, b] \) be so that, for \( t \in [c, d] \), \( \gamma(t) \) is at least \( D_0 \) away from \( \ell \) and so that in between \( c \) and \( d \) the path \( \gamma \) travels to a point at of a distance \( M \epsilon R \) from \( \ell \).

By Lemma 2.2, \( \gamma' = \gamma|_{[c, d]} \) is still \( \epsilon \)-efficient. Let \( c = t_1 < \ldots < t_N = d \) be a partition so that, for \( z_i = \gamma(t_i) \),

\[ \Delta(\gamma', \epsilon R) \gtrsim \sum_i d_X(z_{i+1}, z_i). \]

(In fact, by definition of \( \Delta \), we can choose \( t_i \) so that the two sides are equal. However, we are about to modify the partition \( t_i \).) We can remove some of the times \( t_i \) so that

\[ d_X(z_{i+1}, z_i) \gtrsim \epsilon R. \]

Note that after removing points from the partition, Equation (5) still holds. We have \( N \gtrsim M \).

By the contraction property of hyperbolic spaces, there is a uniform constant \( D_1 \) and a projection map \( \pi: X \to \ell \) such that

\[ d_X(\pi(z_i), \pi(z_{i+1})) \leq D_1. \]

Thus, we have

\[ d_X(\gamma(c), \gamma(d)) \leq 2D_0 + N D_1. \]

On the other hand

\[ \Delta(\gamma, \epsilon R) \gtrsim \sum i d_X(z_{i+1}, z_i) \gtrsim N \epsilon R. \]

From the assumption that \( \gamma' \) is \( \epsilon \)-efficient on scale \( R \), we have

\[ 2D_0 + N D_1 + O(\epsilon R) \gtrsim N \epsilon R. \]

This implies \( N \) is uniformly bounded. Hence, \( M \) is also uniformly bounded. This finishes the proof. \( \square \)

We now consider an efficient map from a box to a hyperbolic space. First we need the following lemma.
Lemma 2.10. Given \( n \) and \( N \), there is \( \sigma = \sigma(n, N) > 0 \) such that for each \( L \), if \( \{ C_i \} \) is a collection of \( N \) convex bodies in \( \mathbb{R}^n \) that cover a ball \( B \) of radius \( L \), then some \( C_i \) contains a ball of radius \( \sigma L \).

Proof. For a convex set \( C \), let \( R = R(C) \) be the out-radius of a convex set: the radius of the smallest ball that contains it. Let \( r = r(C) \) be the in-radius: the largest ball contained in the set and \( w = w(C) \) be the width: the maximum distance between supporting hyperplanes.

From Theorem 1 in [HCSSG04] we have, for some \( \kappa = \kappa(n) > 0 \), and any convex set \( C \) that
\[
\text{Vol}(C) \leq \kappa R^n \int_0^{\arcsin \frac{w}{2R}} \cos^n \theta d\theta \leq \kappa R^n \arcsin \frac{w}{2R} \leq \kappa R^n \arcsin \frac{cr}{2R}.
\]

The last inequality follows from the Steinhagen inequality, which states that there is a constant \( c = c(n) > 0 \), such that
\[ w \leq cr. \]

Since the convex sets \( C_i \) cover the ball of radius \( L \), for some \( c' > 0 \), there is some \( C = C_i \) with
\[ \text{Vol}(C) \geq c'L^n/N. \]

This implies that \( R \geq c''L \), for some constant \( c'' = c''(N) > 0 \).

We will show \( r \geq \sigma L \) by arguing in two cases. Assume,
\[ \arcsin \frac{cr}{2R} \geq \frac{\pi}{4}. \]

Then \( \frac{cr}{2R} \geq \frac{\sqrt{2}}{2} \) and so
\[ cr \geq \sqrt{2}R \geq \sqrt{2}c''L, \]
and we are done by taking \( \sigma = \frac{\sqrt{2}c''}{c} \). Now assume
\[ \arcsin \frac{cr}{2R} \leq \frac{\pi}{4} \]
so that
\[ \arcsin \frac{cr}{2R} \leq \frac{cr}{R}. \]
But then
\[ c'L^n/N \leq \text{Vol}(C) \leq \kappa R^n \frac{cr}{R} = \kappa c R^{n-1} r \]
and so
\[ r \geq \frac{c'}{cN\kappa} \left( \frac{L}{R} \right)^{n-1} L, \]
and again we are done by taking \( \sigma = \frac{c'(c''-1)N\kappa}{c(c'')^n-1}. \)

\[ \square \]

Proposition 2.11. Suppose \( Z \) is a Gromov hyperbolic space and \( f: B \to Z \) is an \( \epsilon \)-efficient map from a box of size \( R \) in \( \mathbb{R}^n \) to \( Z \). Then, there is a sub-box \( B' \subset B \) with \( |B'| \asymp |B| \), so that the image \( f(B') \) lies in an \( O(\epsilon R) \)-neighborhood of a line \( \ell' \) in \( Z \).
Proof. By taking a sub-box we assume \( B = [0, R]^n \) and let \( \ell_i \) be the edges of the box \( B \). Given a line \( \ell \subset B \) denote by \( \ell' \) a geodesic in \( Z \) joining the \( f \) image of its endpoints.

We first prove by induction on \( n \) that, for any \( q \in B \), \( f(q) \) is within \( O(\epsilon R) \) of a point in some \( \ell_i' \). We start with \( n = 2 \) and \( \ell_1, \ell_2, \ell_3, \ell_4 \) the four edges of \( B \) arranged in counterclockwise order. We have by Lemma 2.9 that each point of \( f(\ell_i) \) is within \( O(\epsilon R) \) of \( \ell_i' \). Now let \( \ell_q \) be the line through \( q \) parallel to \( \ell_1 \). Take the rectangle with sides \( \ell_q, \ell_1 \) and subsegments \( m_2 \subset \ell_2 \) and \( m_4 \subset \ell_4 \). Lemma 2.9 implies that the end points of \( m_2' \) are within \( O(\epsilon R) \) of \( \ell_2' \) and hence \( m_2' \) are within \( O(\epsilon R) \) of \( \ell_2' \). The same holds for \( m_4' \) and \( \ell_4' \). The quadrilateral bounded by \( \ell_1', m_2', \ell_2', m_4' \) is \( O(1) \) thin, which implies that \( f(q) \) is within \( O(\epsilon R) \) of one of the other three sides and therefore within \( O(\epsilon R) \) of one of the \( \ell_i' \).

Now suppose the statement is true for boxes in \( \mathbb{R}^{n-1} \) and \( B \subset \mathbb{R}^n \). Take again the geodesics \( \ell_i \) that correspond to the edges of the box \( B \) and any point \( q \in B \). It lies on a face \( B_q^{n-1} \) parallel to the a face of \( B \). Let \( \tau_i \) be the edges of \( B_q^{n-1} \). By induction, \( f(q) \) lies within \( O(\epsilon R) \) of some \( \tau_i' \). Since each \( \tau_i \) itself lies in an \( n-1 \) dimensional face, again by induction, each point of \( f(\tau_i) \) lies within \( O(\epsilon R) \) of the union of \( \ell_i' \). Thus \( f(q) \) is within \( O(\epsilon R) \) of some \( \ell_i' \), completing the induction step.

Now, fix any \( n \) and one of the geodesics \( \ell_i' \). If \( n+1 \) points \( q_1, \ldots, q_{n+1} \) span an \( n \) simplex \( \Lambda \) and are such that each \( f(q_j) \) is within \( O(\epsilon R) \) of \( \ell_i' \), then the image under \( f \) of every point of \( \Lambda \) is within \( O(\epsilon R) \) of \( \ell_i' \). By the Caratheodory theorem, the convex hull of the set of points mapped within \( O(\epsilon R) \) of \( \ell_i' \) is the union of such simplices, and therefore the convex hull is a convex set of points mapped within \( O(\epsilon R) \) of \( \ell_i' \). We conclude by Lemma 2.10 that there is a box \( B' \) with \( |B'| \gtrsim |B| \) consisting of points mapped within \( O(\epsilon R) \) of one of the \( \ell_i' \). \( \square \)

3. Combinatorial Model

Let \( S \) be a possibly disconnected surface of finite hyperbolic type. Define the complexity of \( S \) to be

\[
\xi(S) = \sum_W (3g_W + p_W - 4),
\]

where the sum is over all connected components \( W \) of \( S \), \( g_W \) is the genus of \( W \) and \( p_W \) is the number of punctures.

Let \((\mathcal{T}(S), d_T)\) represent the Teichmüller space equipped with the Teichmüller metric, \((\mathcal{T}(S), d_{WP})\) represent Teichmüller space equipped with the Weil-Petersson metric and \((\text{Mod}(S), d_W)\) represent the mapping class group equipped with the word metric. We construct combinatorial models for these spaces.

Let \( \mathcal{P}(S) \) be the space of pants decompositions of \( S \). That is, a point \( P \) in \( \mathcal{P}(S) \) is a free homotopy class of maximum number of disjoint essential
simple closed curves. Define a marking \((P, \{\tau_\alpha\}_{\alpha \in P})\) to be a pants decomposition together with a transverse curve \(\tau_\alpha\), for each pants curve \(\alpha\). The transverse curves are assumed to be disjoint from other curves in \(P\) and to intersect \(\alpha\) minimally (see [MM00] for more details). The space of all markings is denoted by \(\mathcal{M}(S)\). An augmented marking \((P, \{\tau_\alpha\}_{\alpha \in P}, \{\ell_\alpha\}_{\alpha \in P})\) is a marking together with a positive real number \(\ell_\alpha\) (length of \(\alpha\)) associated to every pants curve \(\alpha\). The length of each curve is assumed to be less than the Bers constant for the surface \(S\). The space of augmented markings is denoted by \(\mathcal{AM}(S)\) (see [Raf10] and also [Dur13] for slightly different definition and extensive discussion of \(\mathcal{AM}(S)\)). We will use these spaces as combinatorial models for, respectively, \((\mathcal{T}(S), d_{WP})\), \((\text{Mod}(S), d_W)\) and \((\mathcal{T}(S), d_{\mathcal{T}})\). Assume \(\mathcal{X} = \mathcal{X}(S)\) is one of these model spaces. Later in this section we will equip \(\mathcal{X}\) with a coarse metric.

### 3.1. Curve complex

Let \(W\) be a subsurface of \(S\). We always assume a subsurface is connected (unless specified otherwise) and that the embedding \(W \subset S\) induces an injective map \(\pi_1(W) \to \pi_1(S)\). We also exclude the cases where \(W\) is a thrice-punctured sphere or an annulus going around a puncture.

Let \(\mathcal{C}(W)\) be the curve graph of \(W\) with metric \(d_{\mathcal{C}(W)}\). This is a graph where the vertices are free homotopy class of non-trivial non peripheral simple closed curve (henceforth, simply referred to as curves) and edges are pairs of curves intersecting minimally (see [MM00] for precise definition and discussion). We make a special definition for the case of annuli. For an annulus \(A\),

- \(\mathcal{C}(A)\) is a horoball in \(\mathbb{H}^2\) when \(\mathcal{X} = \mathcal{AM}(S)\).
- \(\mathcal{C}(A)\) is \(\mathbb{Z}\) when \(\mathcal{X} = \mathcal{M}(S)\).
- \(\mathcal{C}(A)\) is a point when \(\mathcal{X} = \mathcal{P}(S)\).

The curve complex of every subsurface is Gromov hyperbolic in all cases. This is clear when \(W\) is an annulus and is a theorem of Masur-Minsky [MM99] in other cases.

For every subsurface \(W\) of \(S\), there is a coarsely defined projection map (see [MM00] for general discussion and [Dur13] for the case of augmented markings)

\[
\pi_W : \mathcal{X} \to \mathcal{C}(W).
\]

We sketch the definition here. Assume first that \(W\) is not an annulus. Given \(x \in \mathcal{X}\) (recall that in all three cases \(x\) contains a pants decomposition which we denote by \(P_x\)) choose any pants curve \(\gamma \in P_x\) that intersects \(W\). If \(\gamma \subset W\) then choose the projection to be \(\gamma\). If \(\gamma\) is not contained in \(W\) then \(\gamma \cap W\) is a collection of arcs with endpoints on \(\partial W\). Choose one such arc and perform a surgery using this arc and a subarc of \(\partial W\) to find a point in \(\mathcal{C}(W)\). The choice of different arcs or different choices of intersecting pants curves determines a set of diameter 2 in \(\mathcal{C}(W)\); hence the projection is coarsely defined.
For annuli $A$ the definition is slightly different. When $\mathcal{X} = \mathcal{P}(S)$ the projection map is trivially defined since $C(A)$ is just a point. When $\mathcal{X} = \mathcal{M}(S)$, consider the annular cover $\tilde{A}$ of $S$ associated to $A$. Identify the space of arcs in $\tilde{A}$ (homotopy classes of arc connecting different boundaries of $\tilde{A}$ relative to their end points) with $\mathbb{Z}$ by identifying some arc $\omega_0$ with zero and sending every other arc $\omega$ to the signed intersection number between $\omega$ and $\omega_0$. Define $\pi_A(x)$ by lifting the pants deposition $P_x$ and transverse curves $\tau_\alpha$ to $\tilde{A}$. At least one of these curves lifts to an arc connecting different boundaries of $\tilde{A}$ and different ones have bounded intersection number. Hence the map is coarsely defined. We refer to this number as the twisting number of $x$ around $\alpha$ and denote it by twist$_\alpha(x)$, which coarsely defined integer.

Now consider the case $\mathcal{X} = \mathcal{AM}(S)$. Let $B$ be the Bers constant of the surface $S$. For an annulus $A$, we identify $C(A)$ with the subset $H \subset \mathbb{R}^2$ of all points in $\mathbb{R}^2$ where the $y$–coordinate is larger than $1/B$. Note that, for an augmented marking $x = (P, \{\tau_\alpha\}, \{l_\alpha\})$, the twisting number twist$_\alpha(x)$ can still be defined as above. If the core curve of $A$ is in $P_x$ we define $\pi_A(x) = (\text{twist}_\alpha(x), 1/l_\alpha)$, otherwise $\pi_A(x) = (\text{twist}_\alpha(x), 1/B)$.

Also, for subsurfaces $U$ and $V$ we have a projection map

$$\pi_{U,V}: C(U) \to C(V),$$

which is defined on the subset of $C(U)$ consisting of curves that intersect $V$. Here $U$ is non-annular; for an annulus $A$, elements of $C(A)$ cannot be projected to other subsurfaces. When the context is clear, we denote all these projection maps simply by $\pi$. By construction, all projection maps are quasi-Lipschitz.

### 3.2. Distance Formula.

For $x, y \in \mathcal{X}$, define the $W$–projection distance between $x$ and $y$ to be:

$$d_W(x, y) = d_{C(W)}(\pi_W(x), \pi_W(y)).$$

In fact, when $W$ is an annulus with core curve $\alpha$, we sometimes denote this distance with $d_\alpha(x, y)$. We define the distance in $\mathcal{X}$ using these projection distances. For a threshold $T > 0$ large enough, define

$$d_X(x, y) = \sum_{W \in \mathcal{W}_T(x, y)} d_W(x, y),$$

where $\mathcal{W}_T(x, y)$ is the set of subsurfaces with $d_W(x, y) \geq T$. This is not a real metric since the distance between different points may be zero and the triangle inequality does not hold. However, it is symmetric and the triangle inequality holds up to a multiplicative error. That is, for $x, y, z \in \mathcal{X}$

$$d_X(x, y) + d_X(y, z) \preceq d_X(x, z).$$
Also, changing the threshold changes the metric by only uniform additive and multiplicative constants. That is, for $T' \geq T$ we have (see [MM00, Raf07])

$$
\sum_{W \in W_T(x,y)} d_W(x,y) \preceq \sum_{W \in W_{T'}(x,y)} d_W(x,y).
$$

Even though this is not a metric, it makes sense to say $X$ is quasi-isometric to another metric space. In fact, in the category of metric spaces up to quasi-isometry, this notion of distance is completely adequate. We fix a threshold $T$ once and for all so that $d_X(x,y)$ is a well defined number for all $x, y \in X$. The threshold $T$ needs to be large enough so that statement in the rest of this subsection hold.

There is a coarsely defined map

$$(\mathcal{T}(S), d_{WP}) \to \mathcal{P}(S)$$

sending a Riemann surface $X$ to the shortest pant decomposition in $X$ which is, by [Bro03], a quasi-isometry. Hence $\mathcal{P}(S)$ with the above metric is a combinatorial model for the Weil-Petersson metric.

A point in $(\text{Mod}(S), d_W)$ can be coarsely represented as a marking [MM99]. That is, there is a coarsely defined map

$$(\text{Mod}(S), d_W) \to \mathcal{M}(S),$$

which can be defined by, for example, fixing a point $x_0 \in \mathcal{M}(S)$ and sending a mapping class $\phi \in \text{Mod}(S)$ to the marking $\phi(x_0)$. It is shown in [MM00] that this map is a quasi-isometry.

A point in $(\mathcal{T}(S), d_T)$ can be coarsely represented as an augmented marking [Raf10]. That is, there is a coarsely defined map

$$(\mathcal{T}(S), d_T) \to \mathcal{AM}(S)$$

defined as follows. A point in $X$ in Teichmüller space is mapped to the augmented marking $x = (P, \{\tau_\alpha\}, \{l_\alpha\})$ where $P$ is the shorts pants decomposition in $X$, for $\alpha \in P$, $\tau_\alpha$ is the shortest transverse curve to $\alpha$ in $X$ and $l_\alpha$ is the hyperbolic length of $\alpha$ in $X$. It follows from [Raf07] that this map is a quasi-isometry. (Again, see [Dur13] for more details in this case.)

By $(\mathcal{X}(S), d_{\mathcal{X}})$ we denote one of the model spaces above. When the context is clear, we use $\mathcal{X}$ instead of $\mathcal{X}(S)$. However, often we need to talk about $\mathcal{X}(W)$ when $W$ is a subsurface of $S$. For example, if $\mathcal{X}(S)$ is the space of pants decompositions of $S$, then $\mathcal{X}(W)$ is the space of pants decompositions of $W$.

3.3. Bounded projection, consistency and realization. In this section, we review some properties of the projection maps. We will also drive a coarse characterization of the image of the curve complex projections of points in $\mathcal{X}$ similar to [BKMM08].

We start with a Theorem from [MM00].
Theorem 3.1 (Bounded Geodesic Image Theorem). There exists a constant $M_0$ so that the following holds. Assume $V \subseteq U$ are subsurfaces of $S$ and $\beta_1, \ldots, \beta_k$ is a geodesic in $C(U)$. Then either there is some $\beta_j$ that is disjoint from $V$ or $d_V(\beta_1, \beta_k) \leq M_0$.

Masur-Minsky proved this theorem for $\mathcal{M}(S)$ but it then holds for augmented markings as well (and trivially it holds for $\mathcal{P}(S)$). The only case to check is when $V$ is an annulus with a core curve $\alpha$ and $C(V)$ is a horoball. Then, applying the $\mathcal{M}(S)$ version of Theorem 3.1, we have that either some $\beta_j$ is disjoint from $\alpha$ or

$$\text{twist}_\alpha(\beta_1) \approx \text{twist}_\alpha(\beta_k).$$

Also,

$$\pi_V(\beta_i) = (\text{twist}_\alpha(\beta_i), 1/B).$$

Hence,

$$d_\alpha(\beta_1, \beta_k) \approx \log |\text{twist}_\alpha(\beta_1) - \text{twist}_\alpha(\beta_k)| = O(1).$$

Thus, Theorem 3.1 holds for $\mathcal{AM}(S)$ as well.

The other important property of the projection maps is the consistency and realization result of Behrstock-Kleiner-Mosher-Minsky. First we recall that for subsurfaces $U, V$ the notation $U \shortmid V$ means that $U \cap V \neq \emptyset$ and neither is contained in the other. Consider the following consistency condition on a tuple

$$z \in \prod_U C(U).$$

Denote the coordinate of $z$ in $C(U)$ with $z_U$. For a constant $M$, we say $z$ is $M$-consistent if

1. Whenever $U \shortmid V$,

$$\min (d_U(z_U, \partial V), d_V(z_V, \partial U)) \leq M.$$

2. If $V \subseteq U$, then

$$\min (d_U(z_U, \partial V), d_V(z_V, z_U)) \leq M.$$

To any $z \in \mathcal{X}$, the tuple of projections of $z$ is a tuple $z$ so that $z_U = \pi_U(z)$.

The following, in case $\mathcal{X}$ is $\mathcal{M}(S)$ and $\mathcal{P}(S)$, is Theorem 4.3 in [BKMM08]. However it holds true for $\mathcal{AM}(S)$ as well.

Theorem 3.2 (Consistency and Realization). The tuples that are consistent are essentially those that are tuples of projections. More precisely, there is a constant $M_1$, so that

1. For $z \in \mathcal{X}$, the tuple of projection $z$ of $z$ is $M_1$-consistent.

2. If a tuple $z$ is $M$-consistent for some uniform $M$, then there is a realization $z \in \mathcal{X}$ so that

$$\forall U \quad d_U(z, z_U) = O(1).$$
Proof. As mentioned before, this is known for $\mathcal{P}(S)$ and $\mathcal{M}(S)$. We verify the theorem in the case of $\mathcal{AM}(S)$.

First we check part (1). For any $z \in \mathcal{X}$, let $P_z$ be the associated pants decomposition. The non-annular projections of $z$ are the same as projections of $P_z$ and by the $\mathcal{P}(S)$–case of Theorem 3.2 these projections are consistent. Let $A$ be an annulus and $U$ be any other surface intersecting $A$. As in Theorem 3.1, the consistency still holds because

- the image of $\pi_{U,A}$ is always on the boundary the horocycle $\mathcal{C}(A)$,
- the distance in $\mathcal{C}(A)$ between two points on the boundary is the log of the twisting difference, and
- the twisting distance is bounded as a consequence of the consistency theorem for $\mathcal{M}(S)$.

That is, the consistency constant for $\mathcal{AM}(S)$ is no larger than that of $\mathcal{M}(S)$.

To see part (2) we need to construct an augmented marking from a consistent tuple $z$. Use the realization part of Theorem 3.2 for $\mathcal{P}(S)$ to construct a pants decomposition $P_0$ so that, for every subsurface $U$ that is not an annulus, $d_U(P_0, z_U) = O(1)$. Still, for some curves $\alpha$, the projection of $P$ to $\mathcal{C}(\alpha)$ may not be close to $z_\alpha$. We show that the set of such curves is a multi-curve.

Let $\alpha$ and $\beta$ be two curves that intersect $P_0$ so that $d_\alpha(P_0, z_\alpha)$ and $d_\beta(P_0, z_\beta)$ are both large. We show that $\alpha$ and $\beta$ are disjoint. Assume, for contradiction, that they intersect and let $U$ be the surface they fill. Then $z_U$ intersects either $\alpha$ to $\beta$ (say $\alpha$ without loss of generality). By construction of $P_0$

$$d_U(P_0, z_U) = O(1) \xrightarrow{\pi_{U,A} \text{ is quasi-Lipschitz}} d_\alpha(P_0, z_U) = O(1),$$

and by consistency of coordinates of $z$

$$d_\alpha(z_U, z_\alpha) = O(1).$$

Now, the triangle inequality implies that $d_\alpha(P_0, z_\alpha) = O(1)$ which is a contradiction.

We also note that, for any such curve $\alpha$ (where $d_\alpha(P_0, z_\alpha)$ is large) and for every subsurface $U$ intersecting $\alpha$,

$$d_U(\alpha, P_0) \lesssim d_U(\alpha, z_U) = O(1).$$

This is because, if $z_U$ is far from $\alpha$ in $\mathcal{C}(U)$, then the projection of $z_U$ to $\mathcal{C}(\alpha)$ is defined and is (z is constant) near $z_\alpha$. But, as above, $d_U(P_0, z_U) = O(1)$ which implies $d_\alpha(P_0, z_U) = O(1)$. This is a contradiction.

Let $\alpha$ be the multi curve consisting of all the curves above. Since, $d_U(P_0, \alpha) = O(1)$ for every non-annular subsurface $U$, $\alpha$ can be extended to a pants decomposition $P$ with $d_U(P_0, P) = O(1)$ for every non-annular subsurface $U$. That is,

$$d_V(P, z_V) = O(1), \quad \text{for every subsurface } V \text{ intersecting } P.$$
We now complete $P$ into an augmented marking. For a curve $\beta \in P$, write $z_\beta \in \mathcal{C}(\beta)$ as
$$z_\beta = (t_\beta, l_\beta),$$
where $t_\beta$ is an integer and $l_\beta$ is a real number less than the Bers constant. Let $\tau_\beta$ be a curve intersecting $\beta$ minimally and disjoint from other curves in $P$ with twist $\beta(\tau_\beta) \leq t_\beta$ (this can always be achieved by applying Dehn twists around $\beta$). Now,
$$x = (P, \{\tau_\beta\}_{\beta \in P}, \{l_\beta\}_{\beta \in P}),$$
is the desired augmented marking. This is because, for $\beta \in P$, the projections to $\mathcal{C}(\beta)$ are close to $z_\beta$ by construction and, for every other subsurface $V$, the projection of $x$ to $V$ is the same as the projection of $P$ to $V$. □

The following corollary of this theorem will be useful later.

**Corollary 3.3.** Let $U$ and $V$ be two subsurfaces where $\partial V$ intersects $U$. The for any $x \in X$, if
$$d_U(x_U, \partial V) > M_1$$
then
$$d_V(x_V, x_U) = O(1).$$

**Proof.** There are two cases. If $V \subsetneq U$, then this is immediate from part two of the consistency condition.

Otherwise, $\partial U$ intersects $V$ and has a defined projection to $\mathcal{C}(V)$. Since $x_U$ is disjoint from $\partial U$, $d_V(x_U, \partial U) = O(1)$. On the other hand the assumption of the corollary and part one of the consistency condition say that $d_V(x_V, \partial U) = O(1)$. The corollary follows from a the triangle inequality. □

**3.4. Product regions.** For every subsurface $W$ of $S$, we have a projection map
$$\phi_W : \mathcal{X} \to \mathcal{X}(W),$$
defined by Theorem 3.2. Namely since the projections of a point $x \in \mathcal{X}$ to subsurfaces of $S$ are consistent, the projections to subsurfaces of $W$ are also consistent and hence can be realized by a point in $\mathcal{X}(W)$. For points $x, y \in \mathcal{X}$, we define
$$d_{\mathcal{X}(W)}(x, y) = d_{\mathcal{X}(W)}(\phi_W(x), \phi_W(y)).$$
The subsurface $W$ is allowed to be an annulus in which case, $\mathcal{X}(W) = \mathcal{C}(W)$.

For a curve system $\alpha$, let $\mathcal{X}_\alpha$ be the set of points in $x \in \mathcal{X}$ where $\alpha$ is a subset of the pants decomposition $P_x$ associated to $x$. Consider a point $x \in \mathcal{X}_\alpha$. Since every curve in $x$ is disjoint from $\alpha$, the projection of $x$ to any subsurface intersecting $\alpha$ is distance at most 2 from the the projection of $\alpha$. Therefore for a sufficiently large threshold $T$ and for any $x, y \in \mathcal{X}_\alpha$, the set $\mathcal{W}_T(x, y)$ consists only of subsurfaces disjoint from $\alpha$, each of which is contained in some component $W$ of $S \setminus \alpha$. Therefore, the map
$$\Phi : \mathcal{X}_\alpha \to \prod_W \mathcal{X}(W)$$
where
$$\Phi = \prod_W \phi_W$$
is a quasi-isometry (the image is coarsely onto). Here the product space is equipped with the $L^1$-metric. A version of this theorem for Teichmüller space was first proved by Minsky [Min96] and is known as the product regions theorem. We see that the fact that $\Phi$ is a quasi-isometry is essentially immediate from the distance formula. But the proof of the distance formula in [Raf07] used Minsky’s product regions theorem.

There is also a projection map $\phi_{\mathcal{X}_\alpha} : \mathcal{X} \rightarrow \mathcal{X}_\alpha$; choosing a point in each $\mathcal{X}(W)$ for $W$ a component of $S \setminus \alpha$ and taking a union results in a point in $\mathcal{X}_\alpha$. We define $d_{\mathcal{X}_\alpha}$ to mean the distance between projections to $\mathcal{X}_\alpha$. That is, for $x,y \in \mathcal{X}$, we define

$$d_{\mathcal{X}_\alpha}(x,y) := d_{\mathcal{X}}(\phi_{\mathcal{X}_\alpha}(x),\phi_{\mathcal{X}_\alpha}(y)).$$

Note that the projection of $\phi_{\mathcal{X}_\alpha}(x)$ to $\mathcal{X}(W)$ is close to $\phi_W(x)$ because $\phi_W$ was defined using the consistency result. Therefore,

$$d_{\mathcal{X}_\alpha}(x,y) \asymp \sum_W d_{\mathcal{X}(W)}(x,y),$$

where the sum is over components of $S \setminus \alpha$.

We finish with an estimate of $d_{\mathcal{X}}$ using the projection distances $d_{\mathcal{X}_\alpha}$.

**Lemma 3.4.** Suppose $x,y \in \mathcal{X}$ and $\alpha_1, \ldots, \alpha_k$ is a geodesic in $C(S)$ joining $\alpha_1 \in P_x$ to $\alpha_k \in P_y$. Then

$$d_{\mathcal{X}}(x,y) \asymp \sum_{i=1}^k d_{\mathcal{X}_{\alpha_i}}(x,y).$$

**Proof.** Let $M_0$ be the constant in Theorem 3.1. Then any subsurface $U$ with $d_{\mathcal{X}}(x,y) \geq M_0$, is disjoint from some $\alpha_j$ and so is a subset of $S \setminus \alpha_j$. Thus it appears as a term in some $d_{\mathcal{X}(W)}(x,y)$, where $W$ is a component of $S \setminus \alpha_j$. We are done by Equation (9). $\square$

4. **Efficient paths are nearly geodesics**

In this section we show that efficient paths are nearly geodesics in the space $\mathcal{X}$. To do this, we use the construction of Bestvina-Bromberg-Fujiwara [BBF10] which gives a quasi-isometric embedding of the mapping class group to a finite product of Gromov hyperbolic spaces. Their construction is completely axiomatic and works, essentially without modifications, for any of our spaces $\mathcal{X}$. We first review their construction.

4.1. **A quasi-tree of curve complexes.** We summarize some statements in [BBF10]. Fix a threshold $K$. Let $Y$ be a collection of subsurfaces of $S$ with the property that if $V,W \in Y$ then

- $V \pitchfork W$, and
- every curve in $W$ intersects $V$. Hence, the domain of $\pi_{W,V}$ is $C(W)$. 

The space $\mathcal{X}$ is a quasi-tree of curve complexes.
Define $P_K(Y)$ to be a graph whose vertices are elements of $Y$ and two vertices $V, W$ are connected with an edge if for all $U \in Y$
\[ d_U(V, W) \leq K. \]

In [BBF10] it is shown that $P_K(Y)$ is quasi-isometric to a tree.

Define $C(Y)$ to be the space obtained from $P_K(Y)$ by attaching a copy of $C(W)$ for every vertex $W \in Y$ as follows: The vertex set of $C(Y)$ is the union of vertex sets of $C(W), W \in Y$. If $V$ and $W$ are joined by an edge in $P_K(Y)$, then we join the vertex $\pi_V(\partial W)$ in $C(V)$ to $\pi_W(\partial V)$ in $C(W)$. It follows from [BBF10] Theorem E, using the fact that each space $C(W)$ is hyperbolic, that the resulting space $C(Y)$ is also hyperbolic.

Furthermore, in [BBF10], the authors show that when $S$ is connected:

- the subsurfaces of $S$ can be decomposed into finitely many disjoint subsets $Y_1, \ldots, Y_k$ each having the transversality property mentioned above.
- there is a finite index subgroup $\Gamma$ in $\text{Mod}(S)$ which fixes each $Y_i$

When $S$ is disconnected, we decompose the subsurfaces of each component as above let $Y^1, \ldots, Y^k$ be the list of all such collections.

We can assume $Y^i$ contain only essential subsurfaces, that is, the thrice punctured spheres are always excluded, and in the case $\mathcal{X}$ is the pants graph, annuli are also excluded. Let
\[ C = \prod_{j=1}^k C(Y^j) \]

equipped with the $L^1$-metric. Thus $C$ is a product of finitely many hyperbolic spaces.

We define a projection map $\Psi^j: \mathcal{X} \to C(Y^j)$ as follows: For $x \in \mathcal{X}$, choose a subsurface $W_j \in Y^j$ that minimizes
\[ \max_{\alpha \in \mathcal{F}_x} i(\alpha, \partial W_j), \]
where the minimum is over $W_j \in Y^j$. Define
\[ \Psi^j(x) = x_{W_j}. \]

Recall that $x_{W_j}$ is the projection of $x$ to the curve complex $C(W_j)$. That is, $x_{W_j}$ is a point in $C(W_j)$ and hence is a point in $C(Y^j)$.

**Remark 4.1.** Our definition of the projection is slightly different from that of [BBF10]. They use the action of the mapping class group to define the projection. However, in case $Y^j$ consists of annuli and $\mathcal{X}$ is the augmented marking space, $x_{W_j}$ is a point in the horoball $H \subset \mathbb{H}^2$ and not a curve. In particular, the action of the mapping class group is not coarsely transitive. But in the other two cases, the two definitions match.
We claim that the consistency condition (§3.3) shows that this map is coarsely well-defined; that is, up to a bounded distance in $C$, the image is independent of the choice of $W_j$.

First we note that since the finite index subgroup $\Gamma$ acts preserving each $C(Y_j)$ the above minimum is uniformly bounded by a constant independent of $x$. We need to check that distinct choices $W_j$ and $V_j$ give points at bounded distance in $C(Y_j)$. First, we check, for all $U \in Y_j$, that

$$d_U(W_j, V_j) = O(1).$$

Choose an $\alpha \in x$ that intersect $U$. Since $i(\alpha, W_j) = O(1)$, we have

$$d_U(\alpha, W_j) = O(1)$$

and similarly for $\alpha$ and $V_j$. The claim now follows from the triangle inequality. Therefore, $W_j$ and $V_j$ are connected by an edge in $P_K(Y_j)$. To show that $x_{W_j}$ and $x_{V_j}$ are close in $C(Y_j)$, we need that

$$d_{Y_j}(x_{V_j}, \partial W_j) = O(1) \quad \text{and} \quad d_{W_j}(x_{W_j}, \partial V_j) = O(1).$$

This holds since for every curve $\alpha \in x$,

$$i(\alpha, \partial W_j) = O(1) \quad \text{and} \quad i(\alpha, \partial V_j) = O(1).$$

We often denote $\Psi^j(x)$ by $x_j$. Now define a map

$$\Psi: \mathcal{X} \to C \quad \text{with} \quad \Psi = \prod_j \Psi^j.$$

The following theorem is proven in [BBF10] for the mapping class group.

**Lemma 4.2.** There is $K' > K$ so that, for every $x, y \in \mathcal{X}$, and

$$d_C(\Psi(x), \Psi(y)) \geq \frac{1}{2} \sum_{W \in \mathcal{W}_{K'}(x,y)} d_W(x, y),$$

The proof of this lemma in [BBF10] uses only the hyperbolicity of each curve complex and the consistency condition detailed in §3.3 and works verbatim in our case. Hence we omit the proof. As a consequence we have the following theorem which is also proven in [BBF10] for the mapping class group. We give a proof here because our projection maps $\psi^j$ are defined differently from [BBF10].

**Theorem 4.3.** The map $\Psi$ is a quasi-isometric embedding from $\mathcal{X}$ into $C$.

**Proof.** For $x \in \mathcal{X}$, the map $\Psi$ is defined by

$$\Psi(x) = (x_{W_1}, \ldots, x_{W_k}),$$

for some subsurface $W_j \in Y_j$, $1 \leq j \leq k$. We show $\Psi^j$ is quasi-Lipschitz which essentially the same as the proof the $\Psi^j$ is coarsely well defined. If $d_\mathcal{X}(x, x') = O(1)$, then

$$d_U(x, x') = O(1) \quad \forall U.$$
We have $i(x, W_j) = O(1)$, and so
\[ d_U(x, W_j) = O(1) \quad \text{and similarly} \quad d_U(x', W'_j) = O(1). \]
Together this gives
\[ d_U(W_j, W'_j) = O(1), \]
which implies that $W_j$ and $W'_j$ are connected by an edge in $P_K(Y^j)$. We also know that $x$ is close to $x'$ which has bounded intersection with $\partial W'_j$. Hence
\[ d_W(x_j, \partial W'_j) = O(1), \quad \text{and similarly} \quad d_W(x'_j, \partial W_j) = O(1). \]
Therefore,
\[ d_C(Y^j)(\Psi^j(x), \Psi^j(x')) = O(1). \]
This means, the maps $\Psi^j$ are quasi-Lipschitz, and so is $\Psi$.

We need to find a lower bound for the distance between $\Psi(x)$ and $\Psi(y)$.

By Lemma 4.2, there is $K' > K$ so that,
\[ d_C(\Psi(x), \Psi(y)) \geq \frac{1}{2} \sum_{W \in W_{K'}(x, y)} d_W(x, y), \]
and since the distance formula works for any threshold,
\[ d_X(x, y) \preceq \sum_{W \in W_{K'}(x, y)} d_W(x, y). \]
Hence
\[ d_C(\Psi(x), \Psi(y)) \preceq d_X(x, y). \]
This finishes the proof.}

4.2. **Preferred paths and efficient paths.** Since the space $X$ is not hyperbolic, a quasi-geodesic connecting two points in $X$ may not be well-behaved. Instead, we define a notion of preferred path connecting two points in $X$.

**Definition 4.4.** Given $x, y \in X$, we say a quasi-geodesic $\omega: [a, b] \to X$ is a preferred path connecting $x$ to $y$ if,
- $\omega(a) = x$, $\omega(b) = y$
- for every subsurface $U$, the map
  \[ \omega_U = \pi_U \circ \omega : [a, b] \to C(U) \]
  is an unparametrized quasi-geodesic.

**Lemma 4.5.** For any $x, y \in X$ there is a preferred path connecting $x$ to $y$.

**Proof.** In the case of the mapping class group and the pants complex a resolution of a hierarchy is a preferred path [MM00]. In the case that $X$ is Teichmüller space with the Teichmüller metric, such a path is constructed in [Raf07, Theorem 5.7].
Remark 4.6. It is known that the image of a Teichmüller geodesic is not a always preferred path (there may be backtracking in annuli). It is unknown if the image of Weil-Petersson geodesics or a geodesic in the mapping class group is a preferred paths.

Now let \( U \) be a subsurface of \( S \) and let \( x_U \) and \( y_U \) be the projections of \( x \) and \( y \) to \( C(U) \). Denote a geodesic segment in \( C(U) \) connecting \( x_U \) to \( y_U \) by \( [x,y]_U \). Given \( \kappa > 0 \), let

\[
G(x, y, \kappa) = \left\{ z \in X \mid \forall U, d_U(z, [x,y]_U) \leq \kappa \right\}.
\]

In a sense, this set is the union of all points in all preferred paths.

Lemma 4.7. There is a constant \( \kappa_0 \) depending only on the topology of \( S \) and the constant involved in the definition of a preferred path so that, for any preferred path \( \omega: [a,b] \to X \) and any \( a \leq t \leq b \),

\[ \omega(t) \in G(x, y, \kappa_0). \]

Proof. Since \( C(U) \) is Gromov hyperbolic and the projection of \( \omega \) is an unparametrized quasi-geodesic, it stays in a uniform neighborhood of the geodesic connecting its end points. \( \square \)

When \( \kappa_0 \) is fixed, we drop \( \kappa_0 \) and denote this set by \( G(x, y) \).

Lemma 4.8. Let \( x, y \in X \) and \( w, z \in G(x, y) \). Then

- For any subsurface \( U \),
  \[ d_U(w, z) \preceq d_U(x, y). \]
  In fact, \( [w,z]_U \) is contained in a uniform neighborhood of \( [x,y]_U \).
- (convexity) if \( w, z \in G(x, y) \) then \( G(w, z) \) is contained in a uniform neighborhood of \( G(x, y) \).

Proof. Since \( C(U) \) is Gromov hyperbolic, if both \( z_U \) and \( w_U \) are close to \( [x,y]_U \) so is \( [z,w]_U \). Hence the length of \( [z,w]_U \) is less than \( [x,y]_U \) and any point close to \( [z,w]_U \) is also close to \( [x,y]_U \). \( \square \)

The following is the main theorem of this subsection and states that efficient paths fellow travel preferred paths.

Theorem 4.9. Let \( \gamma: [0,R] \to X \) be an \( \epsilon \)-efficient path connecting \( x = \gamma(0) \) to \( y = \gamma(R) \). Then, the image of \( \gamma \) stays in an \( O(\epsilon R) \)-neighborhood of \( G(x, y) \). Moreover, it stays in an \( O(\epsilon R) \)-neighborhood of a preferred path connecting \( x \) to \( y \).

Before we prove this theorem, we need a few technical lemmas. Given \( x, y, z \in X \) and a subsurface \( W \), let \( \eta_W \) be the center of the triangle \((x_W, y_W, z_W)\) guaranteed by the hyperbolicity of \( C(W) \). That is, \( \eta_W \) is \( \delta_W \)-close to all three geodesics \([x,y]_W\), \([y,z]_W\) and \([x,z]_W\), where \( \delta_W \) is the hyperbolicity constant of \( C(W) \).

Lemma 4.10. The set \( \{\eta_W\} \) is \( O(1) \)-consistent.
Proof. Let $U, V$ arbitrary domains which are not disjoint. We can assume that $U \not\subset V$ and hence $\pi_U(\partial V)$ is defined. We know that the projections of $x, y, z$ to $U$ and $V$ are themselves $M_1$–consistent.

Consider the triangle $\Delta_U$ with vertices $x_U, y_U, z_U$ in $\mathcal{C}(U)$. Note that, if $\partial V$ is uniformly close to all three edges of $\Delta_U$, then $\partial V$ is uniformly close to $\eta_U$ and we are done. Hence, we can, without loss of generality, assume that no point in $[x, y]_U$ is near $\partial V$ in $\mathcal{C}(U)$. In fact, since $\eta_U$ is $\delta$–close to $[x, y]_U$, we can assume every point in the convex hall of $x_U, y_U, \eta_U$ is more than $M_1$ away from $\partial V$.

This implies that $x_U, y_U$ and $\eta_U$ have defined projections to $V$ and Theorem 3.1 implies that their projections are a bounded distance from one another. That is,

$$d_V(\eta_U, x_U) = O(1), \quad d_V(\eta_U, y_U) = O(1), \quad \text{and} \quad d_V(x_U, y_U) = O(1).$$

On the other hand, because $d_U(x_U, \partial V)$ and $d_U(y_U, \partial V)$ are both larger than $M_1$, by Corollary 3.3

$$d_V(x_V, x_U) \leq M_1 \quad \text{and} \quad d_V(y_V, y_U) \leq M_1.$$ 

And by triangle inequality

$$d_V(x_V, y_V) = O(1).$$

From the definition of $\eta_V$, we have

$$d_V(\eta_V, [x, y]_V) \leq \delta_V.$$ 

Again, using the triangle inequality, we get

(10) \quad $$d_V(\eta_U, \eta_V) = O(1).$$

This is the consistency condition when $V \subset U$.

Thus assume $U \cap V$. Since $\eta_U$ and $\partial U$ are disjoint, $d_V(\eta_U, \partial U) = O(1)$. This and Equation (10) imply that

$$d_V(\partial U, \eta_V) = O(1),$$

which is the required consistency condition in this case. □

Since the tuple of centers is consistent, it has a realization $\eta$. We call $\eta$ the center of the triangle $x, y$ and $z$.

Lemma 4.11. For any $x, y, z \in \mathcal{X}$, let $\eta$ be the center of the triangle $\Delta$ with vertices $(x, y, z)$. Let $x_j, y_j, z_j$ and $\eta_j$ be projections of $x, y, z$ and $\eta$ to $\mathcal{C}(Y^j)$. Then $\eta_j$ is near the center of the triangle $(x_j, y_j, z_j)$ in $\mathcal{C}(Y^j)$.

Proof. First we claim that, for every $W \in \mathcal{Y}^j$, $d_W(x, x_j) = O(1)$. Let $x_j$ be a curve $x_V$ in a surface $V \in \mathcal{Y}^j$.

$x_V$ is disjoint from $\partial V \quad \Rightarrow \quad d_W(\partial V, x_V) = O(1).$

and

$$i(x, \partial V) = O(1) \quad \Rightarrow \quad d_W(\partial V, x) = O(1).$$
The claim follows from the triangle inequality.

In [BBF10] Lemma 3.13 it is shown that a geodesic in $C(Y^j)$ connecting $x_j$ to $y_j$ is a bounded Hausdorff distance from a union of geodesics $[x_j, y_j]_W$, where the union is over the subsurface where $d_W(x_j, y_j)$ is large. The same holds for the geodesic connecting $x_j$ to $\eta_j$.

But, as a consequence of above claim, the geodesic $[x_j, y_j]_W$ is a bounded Hausdorff distance from $[x, y]_W$ and $[x_j, \eta_j]_W$ is a bounded Hausdorff distance from $[x, \eta]_W$. Also, by assumption, we know that $[x, \eta]_W$ is contained in a bounded neighborhood of $[x, y]_W$. Therefore, $[x_j, \eta_j]_W$ is contained in a bounded neighborhood of $[x_j, y_j]_W$. In particular, if $d_W(x_j, \eta_j)$ is large, so is $d_W(x_j, y_j)$.

That is, every subsurface that appears in the geodesic connecting $x_j$ to $\eta_j$ also appears in the geodesic connecting $x_j$ to $y_j$ and the portion of the geodesic $[x_j, \eta_j]$ that is in $W$ stays near the geodesic $[x_j, y_j]$. Thus, $\eta_j$ is itself close to the $[x_j, y_j]$.

The same hold for $x_j, z_j$ and $y_j, z_j$. Since, $C(Y^j)$ is Gromov hyperbolic, and $\eta_j$ is close to all three geodesics, it is near the center of the triangle. □

We now prove the theorem.

**Proof of Theorem 4.9.** Let $\gamma_j = \Psi_j \circ \gamma$ be the projection of the path $\gamma$ to $C(Y^j)$. By Lemma 2.2, each $\gamma_j$ is still $\epsilon$–efficient. Since $C(Y^j)$ is hyperbolic, by Lemma 2.9, $\gamma_j(t)$ is within $O(\epsilon R)$ distance of the geodesic $[x_j, y_j]$. Let $z = \gamma(t)$ and let $\eta$ be the center of $x, y, z$. From the construction, we have $\eta \in G(x, y)$. We estimate the distance between $\eta$ and $z$.

By Lemma 4.11, $\eta_j$ is the center of $x_j, y_j, z_j$ The distance from $z_j$ to $[x_j, y_j]$ is, up to an additive error, the distance from $z_j$ to the center $\eta_j$. Therefore,

$$d_{C(Y^j)}(z_j, \eta_j) \leq d_{C(Y^j)}([x_j, y_j]) = O(\epsilon R).$$

It follows, since $\Psi$ is coarsely Lipschitz and the metric in $C$ is the $L^1$–metric, that

$$d(x, \eta) \leq \sum_j d_{C(Y^j)}(z_j, \eta_j) = O(\epsilon R).$$

This finishes the proof of the first statement of the Theorem.

We prove the second statement, namely that, $\gamma$ stays in an $O(\epsilon R)$–neighborhood of a preferred path connecting $x$ to $y$. Let $\eta(t)$ be the center of $x, y$ and $\gamma(t)$. The issue is that $\eta(t)$ may not trace a preferred path since the $\epsilon$–efficient path $\gamma$ is allowed to backtrack up to $O(\epsilon R)$. We proceed therefore as follows.

For a time $t$ and a subsurface $W$ let $\eta_W(t)$ be the projection of $\eta(t)$ to $W$. Consider the geodesic $[x, y]_W$ in $C(W)$. Let $s \in [0, t]$ be a time where $d_W(x_W, \eta_W(s))$ is maximized (see Fig. 2) and define

$$\omega_W(t) = \eta_W(s).$$
Figure 2. The point \( \omega_W(t) \) is defined to be the point \( \eta_W(s) \) that is farthest along in \([x, y]_W\), for \( s \in [0, t] \).

Note that \( \omega_W(t) \) is an unparametrized quasi-geodesic since it stays close to the geodesic \([x, y]_W\) and does not backtrack.

We prove \( \{\omega_W(t)\} \) is consistent. Pick two intersecting surfaces \( U \) and \( V \). Suppose first that \( U \sqsupseteq V \) and, without loss of generality, \( d_V(\partial U, y) = O(1) \). Assuming
\[
    d_U(\omega_U(t), \partial V) \quad \text{is large,}
\]
we need to show \( d_V(\omega_V(t), \partial U) = O(1) \). Since \( \omega_U(t) = \eta_U(s) \) and \( \{\eta_W(s)\} \) is consistent, Equation (11) implies
\[
    d_V(\eta(s), \partial U) = O(1).
\]
Now \( \omega_V(t) \) comes after \( \eta_V(s) \) along \([x, y]_V\) and hence is close to the geodesic \([\eta(s), y]_V\). That is, the projections of \( \eta(s) \), \( y \), \( \partial U \) and \( \omega_V(t) \) to \( V \) are all close to each other. In particular,
\[
    d_V(\omega_V(t), \partial U) = O(1).
\]
Next assume \( V \sqsubset U \) and \( d_U(\omega_U(t), \partial V) \) is large, or equivalently \( d_U(\eta(s), \partial V) \) is large. In \( C(U) \), one of the segments \([x, \omega_U(t)]_U \) or \([\omega(t), y]_U \) is far from \( \partial V \). Assume first that \([x_U, \omega_U(t)]_U \) is far from \( \partial V \). This implies that
\[
    d_V(x_U, \omega_U(t)) = O(1).
\]
Now consider \( \omega_V(t) = \eta_V(s') \) for some \( s' \leq t \). Since \( \eta_U(s') \) comes before \( \omega_U(t) = \eta_U(s) \), we have that \( d_U(\eta_U(s'), \partial V) \) is large. By Consistency we have
\[
    d_V(\eta_U(s'), \omega_V(t)) = O(1)
\]
and since \( \eta_U(s') \) comes between \( x_U \) and \( \omega_U(t) \) we also have
\[
    d_V(\eta_U(s'), \omega_U(t)) = O(1).
\]
From the triangle inequality we get
\[
    d_V(\omega_U(t), \omega_V(t)) = O(1),
\]
which proves Consistency in this case.

The remaining case is when \([\omega(t), y]_U \) is far from \( \partial V \). Consistency implies
\[
    d_V(\omega_V(t), \eta_V(s)) = O(1),
\]
where \( \omega_U(t) = \eta_U(s) \). We also have
\[
\delta_V(y_U, \omega_U(t)) = O(1) \quad \text{and so} \quad \delta_V(y_U, \eta_V(s)) = O(1).
\]
Now \( \omega_V(t) \) is further along the geodesic \([x, y]_V\) than \( \eta_V(s) \) and so
\[
\delta_V(y_U, \omega_V(t)) = O(1),
\]
and applying the triangle inequality once more we have
\[
\delta_V(\omega_V(t), \omega_U(t)) = O(1),
\]
the desired inequality for Consistency.

Finally, we need to prove
\[
\delta_X(\gamma(t), \omega(t)) = O(\epsilon R).
\]

Let \( \omega_j(t) \) be the projection \( \omega(t) \) to \( C(Y_j) \). We observe that \( \omega_j(t) \) is near the point \( \eta_j(s), s \in [0, t] \), that is farthest along in \([x_j, y_j]\). This is because a geodesic is \( C(Y_j) \) is a union of geodesics in subsurfaces \( W_1, \ldots, W_k \) appearing in natural order, and since all these subsurfaces intersect, if a point \( z_j \) is ahead \( z'_j \) along \([x_j, y_j]\), then the projection of \( z_j \) is ahead of the projection \( z'_j \) in every subsurface. Therefore there is \( s \in [0, t] \) so that
\[
\omega_j(t) = \eta_j(s).
\]

From part one of Lemma 2.2 we have
\[
d_{C(Y_j)}(x_j, \gamma_j(s)) + d_{C(Y_j)}(\gamma_j(s), \gamma_j(t)) + d_{C(Y_j)}(\gamma_j(t), y_j) \\
\leq d_{C(Y_j)}(x_j, y_j) + O(\epsilon R).
\]

But \( C(Y_j) \) is Gromov hyperbolic, so projection to \([x_j, y_j]\) is distance decreasing. Furthermore \( d_{C(Y_j)}(\eta_j, [x_j, y_j]) = O(1) \). Hence
\[
d_{C(Y_j)}(x_j, \eta_j(s)) + d_{C(Y_j)}(\eta_j(s), \eta_j(t)) + d_{C(Y_j)}(\eta_j(t), y_j) \\
\leq d_{C(Y_j)}(x_j, y_j) + O(\epsilon R).
\]

But we know \( \eta_j(t) \) comes before \( \eta_j(s) \). Therefore,
\[
d_{C(Y_j)}(\eta_j(s), \eta_j(t)) = O(\epsilon R),
\]
and hence,
\[
d_{C(Y_j)}(\omega_j(t), \eta_j(t)) = O(\epsilon R).
\]

Since this is true for every \( j \), we also have
\[
\delta_X(\omega(t), \eta(t)) = O(\epsilon R).
\]

This finishes the proof. \(\square\)
5. Behavior of Preferred paths

In this section, we analyze preferred paths more carefully to obtain more control over their behavior. In Proposition 5.5 we show (up taking a subsurface) that if a preferred path is making progress in a subsurface, it has to stay close to the set of points in $X$ that contain the boundary of that subsurface in their pants decomposition. This is analogous to the main results in [Raf05] for Teichmüller geodesics. At the end of the section we prove two fellow traveling results Proposition 5.8 and Proposition 5.9 for preferred paths. We start by proving a few lemmas.

The following Lemma give a bound on the thickness of $G(x,y)$ in terms of projection distances $d_{X(W)}(x,y)$.

Lemma 5.1. For any $D > 0$, if $d_{X(W)}(x,y) \leq D$ for every subsurface $W$, then any point $z \in G(x,y)$ is within distance $O(D)$ of any preferred path $\gamma$ joining $x,y$.

Proof. Let $\gamma$ be a preferred path connecting $x$ to $y$ and let $z \in G(x,y)$. If $S$ is not connected then $S = W_1 \cup ... \cup W_k$ and, for any time $t$,

$$d_X(z, \gamma(t)) = \sum_i d_{X(W_i)}(z, \gamma(t)) \leq \sum_i d_{X(W_i)}(x,y) = O(D).$$

(Lemma 4.8)

Assume $S$ is connected and let $\gamma(t)$ be a point with $d_{C(S)}(z, \gamma(t)) = O(1)$. Such a point exists because $d_S(z, [x,y]) = O(1)$ by the definition $G(x,y)$ and $\gamma$ connects $x$ to $y$.

Let $\alpha_1, \ldots, \alpha_k$ be the geodesic in the $C(S)$ connecting a pants curve in $z$ to a pants curve in $\gamma(t)$ with $k = O(1)$. By Lemma 4.8 for every $\alpha_i$, we have

$$d_{X_{\alpha_i}}(z, \gamma(t)) \preceq d_{X_{\alpha_i}}(x_1, x_2) \leq D,$$

And by Lemma 3.4

$$d_X(z, \gamma(t)) \preceq \sum_i d_{X_{\alpha_i}}(z, \gamma(t)) = O(D).$$

In preparation for the next lemma we recall a result of Rafi-Schleimer. They give the following definition.

Definition 5.2. Given a pair of points $x, y \in X(S)$, thresholds $T_1 \geq T_0 > 0$, and a subsurface $W$, a collection $\Omega$ of subsurfaces $W_i \subset W$ is an antichain in $W$ for $x$ and $y$ if

- if $W_i \in \Omega$ then $d_{C(W_i)}(x,y) \geq T_0$.
- if $d_V(x,y) \geq T_1$, then $V \subset W_i$ where $W_i \in \Omega$
- if $W_i, W_j \in \Omega$ then $W_i$ is not a proper subsurface of $W_j$.

The size of $\Omega$ is a lower bound for the distance in the curve complex.
Lemma 5.3 ([RS09]). There is a constant $A = A(T_0, T_1)$ such that
$$|\Omega| \leq A d_W(x, y).$$

In the next lemma we show that if $\gamma$ is moving in some $X(W)$ but it is not close to $X_{\partial W}$ it is because it is really moving in a subsurface of $W$.

Lemma 5.4. For all sufficiently large $M, D$ and any subsurface $W$ of $S$, if $\gamma: [a, b] \to X$ is a preferred path connecting $x$ to $y$ such that for all $t$,
$$d_X(\gamma(t), X_{\partial W}) \geq M \quad \text{and} \quad d_X(W)(x, y) \geq D,$$
then there is a proper subsurface $V \subsetneq W$ such that
$$d_X(V)(x, y) \approx d_X(W)(x, y).$$

Proof. Let $M_0$ be the constant for Theorem 3.1. We argue in two cases.

Case 1. Assume $d_W(x, y) \leq M_2$, for some uniform constant $M_2 = O(M_0)$. By the distance formula
$$d_X(W)(x, y) \approx \sum_{U \in W_{M_2}(x, y)} d_U(x, y).$$

Here $W_{M_2}(x, y)$ is the collection of sub-surfaces $U \subset W$ where $d_U(x, y) \geq M_2$. Since $d_W(x, y) < M_2$, $W$ itself is not in the sum. Consider the anti-chain $\Omega$ in $W$ for $x$ and $y$. Then Lemma 5.3 applied with $T_0 = T_1 = M_2$ implies that $\Omega = \{V_1, \ldots, V_k\}$ where $k = O(M_0)$. Each subsurface in $W_{M_2}(x, y)$ is a subset of some $V_i$ and the number of sub-surfaces $V_i$ is uniformly bounded. Hence, for $V$ equal to some $V_i$, we have
$$d_X(W)(x, y) \approx \sum_{U \in V_{M_2}(x, y)} d_U(x, y),$$

where $V_{M_2}(x, y)$ is the collection of sub-surfaces $U \subset V$ where $d_U(x, y) \geq M_2$. That is
$$d_X(V)(x, y) \approx d_X(W)(x, y).$$

Case 2. Assume $d_W(x, y)$ is large compared to $M_0$. We argue this case cannot occur. Choose $z \in \gamma$ whose projection to $C(W)$ is at the midpoint of the quasi-geodesic $\pi_W \circ \gamma$. From our assumption, we know that both
$$d_W(x, z) \quad \text{and} \quad d_W(z, y)$$
are large compared to $M_0$. Let $w \in X_{\partial W}$ be the projection of $z$ to $X_{\partial W}$. Then, for all $U$ disjoint from $W$,
$$d_U(z, w) = O(1).$$

By the assumption of the Lemma $d_X(z, w)$ is large and by the distance formula, there is some $U$ where $d_U(z, w)$ is large. From the previous equation, we know that $U$ has to intersect $W$. There are two cases.

Consider first the possibility that $W \subset U$ and $d_U(z, \partial W)$ large. Since the quasi-geodesic $[x, y]_U$ in $C(U)$ passes through $z_U$, either $[x, z]_U$ or $[z, y]_U$
stays far from \( \partial W \). Then, by Theorem 3.1 either \( d_W(x, z) \leq M_0 \) or \( d_W(y, z) \leq M_0 \) which is a contradiction.

Consider now the possibility that \( U \cap W \) so \( d_U(z, \partial W) \) is large. By the first consistency condition we have that \( d_W(z, \partial U) \) is small. The assumption that \( d_U(x, z) \) and \( d_W(y, z) \) are large (and the triangle inequality) now implies that both \( d_W(x, \partial U) \) and \( d_W(y, \partial U) \) are large. Again, the first consistency condition implies \( d_U(x, \partial W) \) and \( d_U(y, \partial W) \) are small, so \( d_U(x, y) \) is small by the triangle inequality. This in turn implies \( d_U(x, z) \) and \( d_U(y, z) \) are small, and using triangle inequality one more time, we conclude that \( d_U(z, \partial W) \) is small. This is a contradiction. 

**Proposition 5.5.** There exists constant \( D_0 \) such that given a subsurface \( W \subset S \) and a preferred path \( \gamma \colon [a, b] \to X \) connecting \( x \) to \( y \) where

\[
D = d_X(W)(x, y) \geq D_0,
\]

there is a subsurface \( V \subset W \) and a sub-interval \( [c, d] \subset [a, b] \) so that

- \( d_X(V)(\gamma(c), \gamma(d)) \approx D \)
- for \( t \in [c, d] \), \( d_X(\gamma(t), \partial W) = O(1) \).

**Proof.** We use induction on complexity of subsurfaces. If \( \gamma \) does not come within \( M \) of \( \partial W \) then Lemma 5.4 applies. Let \( W' \subset W \) be a subsurface such that

\[
d_X(W')(x, y) \approx D.
\]

Since \( W' \) has lower complexity than \( W \), Proposition 5.5 applies by induction. That is, there is a subsurface \( V \subset W' \subset W \) with the desired properties.

Thus assume \( \gamma \) does in fact come within \( M \) of \( \partial W \), and let \( z_1 \) and \( z_2 \) be points in \( \gamma \) marking the first and the last times \( \gamma \) is within \( M \) of \( \partial W \). We have, either

\[
d_X(W)(x, z_1) \approx D, \quad d_X(W)(z_1, z_2) \approx D, \quad \text{or} \quad d_X(W)(z_2, y) \approx D.
\]

If \( d_X(S_1)(z_1, z_2) \approx D \), then we are done after taking \( V = W \). In the other two cases, (say \( d_X(W)(x, z_1) \approx D \)) the path connecting \( x \) and \( z_1 \) does not come close to \( \partial W \) but travels large distance in \( X_W \). Hence, we again can apply Lemma 5.4 and induction to finish the proof. \( \square \)

5.1. **Steady Progress.** Consider a preferred path that stays near the space \( X_{\partial W} \) for some subsurface \( W \). Sometimes, it is desirable that \( \gamma \) makes steady progress in the curve complex of \( W \). We make this notion precise:

**Definition 5.6.** Suppose \( \gamma \colon [a, b] \to X_{\partial W} \) is a preferred path connecting \( x = \gamma(a) \) and \( y = \gamma(b) \). Let \( L = d_X(W)(x, y) \) and let \( a = t_0 < t_1 < t_2 < t_3 < t_4 < t_5 = b \) such that \( y_i = \gamma(t_i) \) satisfy

\[
d_X(W)(y_i, y_{i+1}) = L/5.
\]

For a constant \( C_0 \), we say \( \gamma \) makes \( C_0 \)-steady progress in \( W \) if, for \( i = 0, \ldots, 4 \)

\[
d_W(y_i, y_{i+1}) \geq C_0.
\]
Note that if $\gamma$ makes $C_0$ steady progress for some $C_0$ then it makes $C'_0$ steady progress for $C'_0 < C_0$.

The next lemma says that we can find subsurfaces where there is steady progress. The constant $D_0$ appears in Proposition 5.5, Lemma 5.7 and Proposition 5.8. This means that we choose $D_0$ large enough that all three statements hold. In the lemma below, $D_0$ seems to depend on $C_0$ which is an open variable. But, in fact, the value of $C_0$ is fixed in Proposition 5.8 and should be taught of as a fixed constant.

**Lemma 5.7.** For every $C_0$, there is $D_0$ such that for any surface $W$, if $\gamma$ is a preferred path joining $x, y \in N_{O(1)}(X_{\partial W})$ and

$$D := d_X(W)(x, y) \geq D_0$$

then, for some interval $[c, d] \subset [a, b]$ and some subsurface $V \subset W$, we have

- $d_X(V)(\gamma(c), \gamma(d)) \gtrsim D$.
- $\gamma|_{[c, d]}$ makes $C_0$-steady progress in $V$.

**Proof.** We first note that by definition, a quasi-geodesic in a lowest complexity subsurface makes steady progress; otherwise since it is a quasi-geodesic it would have to make progress in the curve complex of some proper subsurface, but there are none. The proof is now by induction on complexity. If $\gamma$ does not make steady progress in $W$, then for some $i$, $d_W(y_i, y_{i+1}) = O(1)$. If so, we use the anti-chains (Lemma 5.3) and argue as in Case 1 in the proof of Lemma 5.4 to conclude that there exists a subsurface $V \subset W$ where

$$d_X(V)(y_i, y_{i+1}) \gtrsim D.$$ 

Here the implied constant only depends on $C_0$ and not on $D$. Now, an induction on the complexity of $W$ implies the lemma (replace $\gamma$ with the preferred path connecting $y_i$ to $y_{i+1}$ and $W$ with $V$). □

We now prove a pair of fellow traveling lemmas. The first states that, if the end points of two preferred paths are close compared to their lengths, and the first one makes steady progress in $W$ then the middle part of the second one also stays near $X_{\partial W}$.

**Proposition 5.8.** There are constants $c_0, c_1, C_0, D_0$ with the following property. Suppose $\gamma$ is a preferred path joining $x, y \in X_{\partial W}$ that makes $C_0$-steady progress in $W$, let $z$ be the midpoint of $\gamma$ and assume $D := d_X(W)(x, y) \geq D_0$. Suppose $\gamma'$ is a preferred path joining $x'$ and $y'$ with

$$d_X(x, x') \leq c_0 D, \quad \text{and} \quad d_X(y, y') \leq c_0 D.$$ 

Then, there is a subsegment of $\gamma'$ with length comparable to $D$ that stays in a bounded neighborhood of $X_{\partial W}$. In fact, for $z'$ on $\gamma'$, if

$$d_X(z', z) \leq c_1 D \quad \text{then} \quad d_X(z', X_{\partial W}) = O(1).$$

**Proof.** Let $x = y_0, \ldots, y_5 = y$ be as in Definition 5.6. Consider the projection $\pi_W(\gamma)$ to $C(W)$. We have $d_W(x, y_1) \gtrsim C_0$. Even though the distance in $X$
between $y_1$ and $x$ is $D/5$, there may be points much closer to $x$ in $X$ whose projection to $\mathcal{C}(W)$ is still near $\pi_W(y_1)$. However, this can not happen if we travel a few steps towards $y$ along $[x,y]_W$. We make this precise.

**Claim.** There is $\beta_1 \in [x,y]_W$ so that

- $d_W(y_1,\beta_1) \leq 2\delta$.
- for any $z_1 \in X$,

\[ d_W(z_1,\beta_1) \leq \delta \implies d_X(x,z_1) \succ D. \]

We remind the reader that $\delta$ is the hyperbolicity constant for $\mathcal{C}(W)$.

**Proof of Claim.** We know $d_X(x,y_1) \succ D/5$ and that the distance in $X(W)$ is the sum of subsurface projections to subsurfaces in $W_T(x,y_1)$. The boundary of any such subsurface is near a curve in $[x,y]_W$.

Let $\beta_1$ be the curve along $[x,y]_W$ that is $2\delta$ away (towards $y$) from the projection of $y_1$ to $W$. For $T$ larger than $M_0$, by Theorem 3.1 the projection of $[\beta_1,y]_W$ to any subsurface $U \in W_T(x,y_1)$ has uniformly bounded diameter $M_0$ (every curve in $[\beta_1,y]_W$ intersects $U$). In fact, for $z_1 \in X$, where $d_W(z_1,\beta_1) \leq \delta$, every curve in the geodesic $[z_1,y]_W$ also intersects $U$. Hence,

\[ d_U(x,z_1) \succ d_U(x,y_1). \]

Therefore, $d_X(x,z_1) \succ d_X(x,y_1) = D/5$. \hfill \Box

\begin{figure}[h]
\centering
\includegraphics[width=\linewidth]{figure3.png}
\caption{The projections of $y_0, \ldots, y_5$ in $\mathcal{C}(W)$ are at least $C_0$ apart. For $i = 1, \ldots, 4$, the $\beta_i$ in $[x,y]_W$ is $2\delta$ away from the projection of $y_i$ in the indicated direction.}
\end{figure}

Similarly we find

- a curve $\beta_4$ near the projection of $y_4$ so that, for any $z_4$ whose projection to $W$ is $\delta$ close to $\beta_4$, we have

\[ d_X(z_4,y) \succ D. \]

- curves $\beta_2$ and $\beta_3$, near the shadows of $y_2$ and $y_3$ respectively, so that, for any $z_2, z_3 \in X$,

\[ d_W(z_2,\beta_2) \leq \delta, \quad d_W(z_3,\beta_3) \leq \delta \implies d_X(z_2,z_3) \succ D. \]
If \( c_0 \) is small enough, any point in the path \([x, x']_W\) has a distance of at least \( C_0 \) (up to an additive error) from any point in \([\beta_2, \beta_3]_W\). The same holds for curves in \([y, y']_W\). Hence, if \( C_0 \) is much larger than the hyperbolicity constant \( \delta \), it follows from the hyperbolicity of the curve complex that the \( \delta \)-neighborhood of the path \([x', y']_W\) has to contain \([\beta_2, \beta_3]_W\). In particular,

\[
d_W(x', y') \gtrsim C_0.
\]

This means the path \([x', y']_W\) passes near \( X_{\beta W} \). Let \( z'_1 \) and \( z'_4 \) be the first and the last time the path \([x', y']_W\) is near \( X_{\beta W} \). A \( \delta \)-neighborhood of the geodesic \([z'_1, z'_4]_W\) must also contain \([\beta_2, \beta_3]_W\).

This means that there are points \( z'_2 \) and \( z'_3 \) along \([z'_1, z'_4]_W\) whose projection to \( C(W) \) is \( \delta \)-close to \( \beta_2 \) and \( \beta_3 \) respectively. Thus, by Equation (12),

\[
d_X(z'_2, z'_3) \gtrsim D \quad \implies \quad d_X(z'_1, z'_4) \gtrsim D.
\]

This is the desired subsegment of \( \gamma' \). To see the last assertion of the theorem, note that if \( d_X(z', z) \leq c_1D \) for \( c_1 \) small enough, then \( z' \) is indeed in the segment \([z'_2, z'_3]\). This finishes the proof. \( \square \)

**Proposition 5.9.** Assume \( S \) is connected. There is a constant \( C_1 \) with the following property. For two pairs of points \( x, y \) and \( x', y' \), suppose

\[
d_S(x, x') = O(\delta) \quad \text{and} \quad d_S(y, y') = O(\delta).
\]

Suppose \( z' \in \mathcal{G}(x', y') \) is such that

\[
d_S(z', x) \geq C_1 \quad \text{and} \quad d_S(z', y) \geq C_1.
\]

Then

\[
z' \in \mathcal{G}(x, y).
\]

**Proof.** We need to show that, for any subsurface \( U \),

\[
d_U(z', [x, y]_U) = O(1).
\]

If \( U = S \) then we know from the hyperbolicity of \( C(S) \) that either \( z'_S \) is close to \([x, x']_S\), \([x, y]_S\) or \([y, y']_S\). Note that different paths connecting two points in \( C(S) \) are a bounded distance apart in the \( d_C(S) \). Since, \( x'_S \) is much closer to \( x_S \) than \( z'_S \), then \( z'_S \) is far from the path \([x, x']_S\) and similarly from \([y, y']_S\). Thus it has to be near \([x, y]_S\).

Now assume \( U \neq S \). If \( \partial U \) is not close to \( z'_S \) in \( C(S) \) then, without loss of generality, we can assume \( \partial U \) is far from \([z', y]_S\) (otherwise \( \partial U \) would be far from \([x, z']_S\)). Therefore, by Theorem 3.1 \( d_U(z', y) = O(1) \). That is, \( z_U \) is close to \( y_U \) and hence close to \([x, y]_U\).

If \( \partial U \) is near \( z'_S \) in \( C(S) \), then it is far from \([x, x']_S\) and \([y, y']_S\). Therefore

\[
d_U(x, x') = O(1) \quad \text{and} \quad d_U(y, y') = O(1).
\]

Hence, \([x, y]_U \) is near \([x', y']_U \). But we know that \( z'_U \) is near \([x', y']_U \) and therefore it is near \([x, y]_U \). This finishes the proof. \( \square \)
6. Local structure of efficient maps

We want to prove the following result. We assume $S$ is connected.

**Theorem 6.1.** For all $R_0, \epsilon_0$ there is $R_1 \geq R_0$ and $\epsilon_1 < \epsilon_0$ so that if $B$ is a box in $\mathbb{R}^n$ with $|B| = R \geq R_1$ and $f: B \to X$ is an $\epsilon R$-efficient map with $\epsilon \leq \epsilon_1$, then there is a sub-box $B' \subset B$ with $R' = |B'| \geq R_0$ such that one of the following holds:

- $R' \succ \sqrt[3]{\epsilon^2} R$ and, for some curve $\alpha$, $f(B')$ lies within an $O(\sqrt[3]{\epsilon} R')$-neighborhood of $X_\alpha$.
- $R' \succ \sqrt[3]{\epsilon} R$ and there exist $x, y$ so that $f(B')$ lies within an $O(\sqrt[3]{\epsilon} R')$-neighborhood of a preferred path in $X$ joining $x$ to $y$.

**Proof.** We can assume that the diameter $f(R)$ is at least $\sqrt[3]{\epsilon} R$; otherwise the first case holds by taking $B' = B$, $R' = R$ and any curve $\alpha \in \pi_S \circ f(B)$. Thus choose $x = f(p), y = f(q)$ where $p, q \in B$ such that

$$d_X(x, y) \geq \sqrt[3]{\epsilon} R.$$ 

Identify the geodesic segment $[p, q]$ with an interval in $\mathbb{R}$. By Theorem 4.9, $f([p, q])$ stays in the $O(\epsilon R)$-neighborhood of a preferred path $\gamma$ joining $x, y$.

Now suppose, for some proper subsurface $W \subseteq S$, that

$$d_X(x, y) \geq \sqrt[3]{\epsilon} R.$$ 

We now claim the first conclusion of the theorem holds. That is, there is a sub-box $B'$ of size

$$R' \succ d_X(x, y) \succ \sqrt[3]{\epsilon^2} R$$

and a curve $\alpha$ such that $f(B')$ lies within an $O(\sqrt[3]{\epsilon} R')$-neighborhood of $X_\alpha$. We prove the claim.

It follows from Proposition 5.5 and Lemma 5.7 that we can find a subinterval $[d_1, d_2] \subset [a, b]$, a subsurface $V$ and $C_0$ such that

1. $d_X(V)(\gamma(d_1), \gamma(d_2)) \succ R'$,
2. $\gamma([d_1, d_2])$ stays in a $O(1)$ neighborhood of $X_{\partial V}$, and
3. the path $\gamma|_{[d_1, d_2]}$ makes $C_0$-steady progress in $V$.

Let $x_i = \gamma(d_i)$. Let $s_i \in [p, q]$ so that

$$d_X(f(s_i), x_i) = O(\epsilon R).$$

For small $c$ consider any two points $t_1, t_2$ at distance $cR'$ from $s_1, s_2$ respectively and set $y_i = f(t_i)$. For $c$ small enough we have

$$d_X(x_i, y_i) \leq d_X(x_i, f(s_i)) + d_X(f(s_i), y_i) \leq O(\epsilon R) + KcR' \leq c_0 R'$$

where $c_0$ is the constant given by Proposition 5.8. Let $d \in [d_1, d_2]$ be such that $x = \gamma(d)$ is the midpoint of a $\gamma|_{[d_1, d_2]}$. Let $p$ be any point such that

$$d_X(f(p), x) = O(\epsilon R).$$
By Proposition 5.8 there is $c_1$ so that all $y \in \mathcal{G}(y_1, y_2)$ that satisfy
\[ d_X(y, x) \leq c_1 R' \]
also satisfy
\[ d_X(y, \partial V) = O(1). \]
For $c$ small, a box of size $R'$ centered at $p$ is mapped under $f$ within distance $O(cR) = O(\sqrt[3]{\epsilon} R')$ of such $y$ and so the image of the box lies within $O(\sqrt[3]{\epsilon} R')$ of $\mathcal{X}_\alpha$, for $\alpha = \partial V$. This proves the claim.

We continue the proof of the theorem. By the first part of the argument we can assume that for all $W \subset S$ that
\[ d_X(W, x, y) = O(\epsilon R^2) \]
This and Lemma 3.4 imply that $\gamma$ makes $C_0\sqrt[3]{\epsilon}$ steady progress in the entire surface $S$, for some $C_0 > 1/3\sqrt[3]{\epsilon}$. For $\epsilon$ small enough, this implies that it makes $C_0$ steady progress where $C_0$ is the fixed constant of Proposition 5.8.

Let $C_1$ be the constant of Proposition 5.9. For a small but fixed $c > 0$, take a $cR'$–neighborhood of $p$ and a $cR'$–neighborhood of $q$ where now
\[ R' = d_X(x, y) \geq 3\sqrt[3]{\epsilon} R. \]

Let $p', q'$ be any points in these neighborhoods and let $x' = f(p')$ and $y' = f(q')$. By Theorem 4.9 we can find a preferred path $\gamma'$ joining $x', y'$ within $O(\epsilon R)$ of $f([p', q'])$. Since the map $f$ is quasi-Lipschitz it follows, for $c$ sufficiently small, that
\[ d_X(x', y') \geq R'. \]

Choose any point $\hat{p}$ in the middle third of $[p', q']$. There is $z' \in \mathcal{G}(x', y')$ whose projection to $\mathcal{C}(S)$ is at least $C_1$–far from $x'_S$ and $y'_S$ and
\[ d_X(f(\hat{p}), z') = O(\epsilon R). \]

By Proposition 5.9 we know that $z' \in \mathcal{G}(x, y)$ and so $f(\hat{p})$ is within $O(\epsilon R)$ of $\mathcal{G}(x, y)$ and by Lemma 5.1 any point of $\mathcal{G}(x, y)$ is within distance $O(\sqrt[3]{\epsilon^2} R)$ of $\gamma$.

We have shown that any point in the middle third of any segment starting near $p$ and ending near $q$ is mapped to a point that is in a $O(\epsilon R)$–neighborhood of $\gamma$. But such a path covers a box of size $R'$. Thus, there is box of size $R'$ which maps within $O(\sqrt[3]{\epsilon^2} R) = O(\sqrt[3]{\epsilon} R')$ of a preferred path. We are done.

7. Proof of main theorems

We are ready to prove Theorem A and Theorem B. We first prove a version of Theorem A for efficient maps. Then, we use coarse differentiation to finish the proof.
**Definition 7.1.** Let $\alpha$ be a (possibly empty) curve system. For every connected component $W$ of $S \setminus \alpha$ (including annuli if $X$ is not the $\mathcal{P}(S)$), let $\omega_W : I_W \to \mathcal{X}(W)$ be a preferred path. Consider the box $B = \prod_W I_W \subset \mathbb{R}^n$, where $n$ is the number of components of $S \setminus \alpha$. Then

$$F : B \to \mathcal{X}_\alpha = \prod_W \mathcal{X}(W)$$

where $F = \prod_W \omega_W$, is a quasi-isometric embedding because each $\gamma_W$ is a quasi-geodesic and the product space is equipped with the $L^1$-metric. We call this map a **standard flat** in $\mathcal{X}$.

**Theorem 7.2.** Let $S$ be a surface with complexity $\xi = \xi(S)$. For given $\epsilon_0$ and $R_0$, let

$$\epsilon_\xi = \epsilon_0^{(6\xi)} \quad \text{and} \quad R_\xi = \frac{R_0}{\epsilon_\xi}.$$ 

Assume $f : B \to \mathcal{X}$ is an $\epsilon_\xi$–efficient map where $B$ is a box of size $R_\xi$ in $\mathbb{R}^n$. Then, there is a box $B' \subset B$ of size $R' \geq R_0$ so that the image $f(B')$ lies inside the $O(\epsilon_0 R')$–neighborhood of a standard flat in $\mathcal{X}$.

**Proof.** We prove the theorem by induction on the complexity $\xi = \xi(S)$ of the surface $S$ (see Equation (6)). If $\xi = 0$, then

$$S = \prod_{i=1}^m S_i,$$

where each $S_i$ is either a once-punctured torus or a four-times-punctured spheres. When $\mathcal{X}$ is the pants complex, $\mathcal{X}(S_i)$ is quasi-isometric to the Farey graph; when $\mathcal{X}$ is the augmented marking space, $\mathcal{X}(S_i)$ is isometric to a copy of the hyperbolic plane; and when $\mathcal{X}$ is the marking complex $\mathcal{X}(S_i)$, is a graph whose vertices are the edges of the Farey graph and two vertices are connected if the associated edges have a common vertex. The latter is known to be quasi-isometric to a tree. Hence, in all cases, $\mathcal{X}(S_i)$ is a Gromov hyperbolic space. That is, $\mathcal{X}$ is a product of hyperbolic spaces.

In this case $R_\xi = R_0$ and $\epsilon_\xi = \epsilon_0$. Let $f_i : B \to \mathcal{X}(S_i)$ be the projection of $f$ to $\mathcal{X}(S_i)$, by Lemma 2.2, $f_i$ is still $\epsilon_0$–efficient. Applying, Proposition 2.11 to $f_1 : B \to \mathcal{X}(S_1)$, we obtain a sub-box $B_1$ where $f_1(B_1)$ lies in an $O(\epsilon_0 R_0)$–neighborhood of a quasi-geodesic in $\mathcal{X}(S_1)$. But in this case, quasi-geodesics are also preferred paths. Now we apply Proposition 2.11 to $f_2 : B_1 \to \mathcal{X}(S_2)$ to obtain a box $B_2$ so that $f_2(B_2)$ lies in an $O(\epsilon_0 R_0)$–neighborhood of a preferred path in $\mathcal{X}(S_2)$. Continuing this way, we find a box $B_m$ where the image of every $f_i$ lies in an $O(\epsilon_0 R_0)$–neighborhood of a preferred path in $\mathcal{X}(S_i)$. This means $f(B_m)$ lies in a $O(\epsilon_0 R_0)$–neighborhood of a standard flat in $\mathcal{X}$. Note that $B_m$ has the same size as $B$ (within uniform multiplicative error). This proves the base case of the induction.
Assume now that \( \xi \) is non-zero. Apply Theorem [6.1]. If the second condition holds, we are done for
\[
R' \gtrsim \sqrt[\xi]{R} = \frac{\sqrt[\xi]{R}}{\epsilon}\geq R_0,
\]
since a preferred path is itself a standard flat and \( \sqrt[\xi]{R} = \epsilon^{2}_{\xi-1} \leq \epsilon_0 \).

Otherwise, we have a box \( \bar{B} \) of size
\[
\bar{R} \gtrsim \sqrt[\xi]{\epsilon^2} R \geq R_{\xi-1}.
\]
that maps to a \( O(\sqrt{\xi} \bar{R}) \)–neighborhood of \( X_\alpha \) for some curve \( \alpha \). The map \( f \) is \( \epsilon^{2}_{\xi-1} \)–efficient because, reducing the size of the box by some factor (in this case \( \sqrt[\xi]{\epsilon^2} \)) only makes the efficiency constant increase by the same factor and
\[
\frac{\epsilon_\xi}{\sqrt[\xi]{\epsilon^2}} = \sqrt[\xi]{R} = \epsilon^{2}_{\xi-1}.
\]

Composing \( f \) with the closest point projection map to \( X_\alpha \) and using the fact that \( X_\alpha \) is quasi-isometric to \( C(\alpha) \times X(S \setminus \alpha) \), we obtain a map
\[
\bar{f}: \bar{B} \to C(\alpha) \times X(S \setminus \alpha).
\]
By part (4) of Lemma [2.2] \( \bar{f} \) is \( \epsilon_{\xi-1} \)–efficient.

Projecting to the second factor, we have a \( \epsilon_{\xi-1} \)–efficient map from a box of size \( R_{\xi-1} \) to \( X(S \setminus \alpha) \) which by the inductive assumption has a sub-box \( B_0 \) of size at least \( R_0 \) that stays in \( O(\epsilon_0 R_0) \)–neighborhood of a standard flat in \( X(S \setminus \alpha) \). Now projecting to the first factor and applying Proposition [2.11] \((C(\alpha) \text{ is hyperbolic})\), we find a sub-box \( B' \) of size at least \( R_0 \) that stays in a \( O(\epsilon_0 R_0) \)–neighborhood of a line in \( C(\alpha) \). That is, \( f(B') \) stays in a \( O(\epsilon_0 R_0) \)–neighborhood of a standard flat in \( X \). This finishes the proof. \( \square \)

**Proof of Theorem [4]** Let \( R_1 \) be large enough that the box \( B \) is guaranteed to have a sub-box \( B_\xi \) of size at least \( R_\xi \) where the restriction of \( f \) to \( B_\xi \) is \( \epsilon_\xi \)–efficient. Apply Theorem [7.2] to \( f: B_\xi \to X \) to obtain the theorem. \( \square \)

**Proof of Theorem [8]** Pick \( 1/\epsilon_0 \gg K \) and \( R_0 \gg C/\epsilon_0 \). Apply Theorem [A] to obtain a constant \( R_1 \) and let \( R_2 \) be any constant greater than \( R_1 \). Then, the image of a sub-box \( B' \) of \( B \) of size \( R' > R_0 \) is in a \( O(\epsilon_0 R') \)–neighborhood of a flat \( F: \mathbb{R}^m \to X \). Taking a composition of \( f \), the closest point projection to the image of \( F \) and then \( F^{-1} \), we obtain a map \( G: B' \to \mathbb{R}^m \) with the property that, for \( p, q \in B' \),
\[
d_{\mathbb{R}^m}(G(p), G(q)) \asymp d_{\mathbb{R}^m}(p, q) \pm O(\epsilon R').
\]
We show that there is no such a map if \( n \) is bigger than \( \operatorname{rank}_{top}(X) \geq m \). The proof is similar to the proof of nonexistence of quasi-isometries between \( R^n \) and \( R^m \). Consider a net of \( O(\epsilon_0 R^n) \) points in \( B' \) that are pairwise \( K_1 \epsilon_0 R' \) apart, where \( K_1 \) is much larger than constants involved in Equation [14]. Then, by the choice of \( K_1 \), the image points are at least distance \( \epsilon_0 R' \)–apart
and are contained in a ball of radius $O(R')$ in $\mathbb{R}^m$. The number points in a such net is of order of $O(R'\epsilon_0)^m$. Choosing $R'$ large enough (which can be done by choosing $R_0$ large) we obtain a contradiction.

To see the second assertion, we note that if, for every subsurface $W_i$, $\mathcal{C}(W_i)$ contains an infinite geodesic, the the product of these geodesics is a quasi-isometric image of $R^n$. This fails when $W_i$ is an annulus and $X$ is the augmented marking space (a horoball does not contain a bi-infinite geodesic). In this case, we choose a $\epsilon_0$–thick point $X$ in $T(S)$ and take $B > 0$ large enough so that the set of curves of length less than $B$ fill $X$. This set contains many (but uniformly bounded number of) pants decompositions. For any such pants decomposition $P$ (where the length in $X$ of every curve in $P$ is less than $B$), the point $x \in X$ associated to $P$ (see §3.4). Consider an infinite ray for every horoball associated to a curve in $P$. The product of these rays is a quasi-isometric image of an orthant in $R^n$. The union of these is quasi-isometric to $R^n$ because the total number of orthant glued at $x$ is uniformly bounded. □

References

[BBF10] M. Bestvina, K. Bromberg, and K. Fujiwara. Constructing group actions on quasi-trees and applications to mapping class groups. preprint, arXiv:1006.1939, 2010.
[BF06] J. Brock and B. Farb. Curvature and rank of Teichmüller space. Amer. J. Math., 128(1):1–22, 2006.
[BKMM08] J. Behrstock, B. Kleiner, Y.N. Minsky, and L. Mosher. Geometry and rigidity of mapping class groups. arXiv:0801.2006, 2008.
[BM08] J. A. Behrstock and Y. N. Minsky. Dimension and rank for mapping class groups. Ann. of Math. (2), 167(3):1055–1077, 2008.
[Bro03] J. F. Brock. The Weil-Petersson metric and volumes of 3-dimensional hyperbolic convex cores. J. Amer. Math. Soc., 16(3):495–535 (electronic), 2003.
[Dur13] M. Durham. The augmented marking complex of a surface. preprint, 2013.
[EFW06] A. Eskin, D. Fisher, and K. Whyte. Coarse differentiation of quasi-isometries I: spaces not quasi-isometric to cayley graphs. arXiv:math/0607207, 2006.
[EFW07] A. Eskin, D. Fisher, and K. Whyte. Coarse differentiation of quasi-isometries II: rigidity for Sol and Lamplighter groups. arXiv:0706.0940, 2007.
[Ham07] U. Hamenstädt. Geometry of the mapping class group III: Quasi-isometric rigidity. preprint, arXiv:math.GT/051242, 2007.
[HCSSG04] M.A. Hernández Cifre, G. Salinas, and S. Segura Gomis. Two optimization problems for convex bodies in the n-dimensional space. Beiträge Algebra Geom., 45(2):549–555, 2004.
[Min96] Y.N. Minsky. Extremal length estimates and product regions in Teichmüller space. Duke Math. J., 83(2):249–286, 1996.
[MM99] H.A. Masur and Y.N. Minsky. Geometry of the complex of curves. I. Hyperbolicity. Invent. Math., 138(1):103–149, 1999.
[MM00] H.A. Masur and Y. N. Minsky. Geometry of the complex of curves. II. Hierarchical structure. Geom. Funct. Anal., 10(4):902–974, 2000.
[MS13] H. Masur and S. Schleimer. The geometry of the disk complex. J. Amer. Math. Soc., 26(1):1–62, 2013.
[Raf05] K. Rafi. A characterization of short curves of a Teichmüller geodesic. Geometry and Topology, 9:179–202, 2005.

[Raf07] K. Rafi. A combinatorial model for the Teichmüller metric. Geom. Funct. Anal., 17(3):936–959, 2007.

[Raf10] K. Rafi. Hyperbolicity in Teichmüller space. preprint [arXiv:1011.6004] 2010.

[RS09] K. Rafi and S. Schleimer. Covers and the curve complex. Geom. Topol., 13(4):2141–2162, 2009.