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 $\mathcal{T}(S)$ equipped with the Teichmüller metric $d_{\mathcal{T}}$, the Teichmüller space equipped with the Weil-Petersson metric d_{WP} and the mapping class group $\mathrm{Mod}(S)$ 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. One major goal in understanding the large scale geometry of a metric space is prove quasi-isometric rigidity of the space; that any quasi-isometry is bounded distance from an isometry. The usual starting point 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.

Date: October 16, 2014.

The first author is partially supported by NSF grants DMS 0905912 and DMS 1201422. The second author is partially supported by NSF grant DMS-0905907. The third author is partially supported by by NSF grant DMS-1007811.

Theorem A (Image of a box is locally standard)**.** *Let $\mathcal{X}$ be either of Teichmüller space $\mathcal{T}(S)$ equipped with the Teichmüller metric $d_{\mathcal{T}}$, Teichmüller space with the Weil-Petersson metric d_{WP}, (or pants graph) or the mapping class group* $\mathrm{Mod}(S)$ *equipped with the word metric d_W. For all K, C and for all R_0, ϵ_0 there exists R_1 such that if B is a box of size at least R_1 and $f\colon B \to \mathcal{X}$ 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 $\mathcal{X}$.*

As a corollary, we determine the large scale rank of the space $\mathcal{X}$. Define the *topological rank* of $\mathcal{X}$, $\mathrm{rank}_{top}(\mathcal{X})$, 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 $\mathcal{X}$ is quasi-isometric to $(\mathcal{T}(S), d_{Wp})$, annuli are also considered inessential.) The dimension of a standard flat in $\mathcal{X}$ is at most $\mathrm{rank}_{top}(\mathcal{X})$.

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 $f\colon B \to \mathcal{X}$ is a (K,C)–quasi-isometric embedding, then $n \leq \mathrm{rank}_{top}(\mathcal{X})$. Furthermore, for $n = \mathrm{rank}_{top}(\mathcal{X})$, there is a quasi-isometric embedding of a Euclidean n dimensional half space union into $\mathcal{X}$.*

Define the *geometric rank* of $\mathcal{X}$, $\mathrm{rank}_{geo}(\mathcal{X})$, to be largest integer n so that for any R there is a quasi-isometric embedding f of a ball $B \subset \mathbb{R}^n$ of radius R into $\mathcal{X}$. 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 $\mathcal{X}$ are equal. Namely, if $\mathcal{X}$ is either* $(\mathrm{Mod}(S), d_W)$ *or* $\mathcal{T}(S), d_{\mathcal{T}})$ *then*

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

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

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

Remark 1.1. In the case of $\mathrm{Mod}(S)$ there are quasi-isometric embeddings of all of Euclidean space of dimension, the geometric rank into $\mathrm{Mod}(S)$. In Theorem 1.3 of [Bo] B.Bowditch proves that there is a quasi-isometric embedding of a Euclidean n-dimensional half-space into $(\mathcal{T}(S), d_{\mathcal{T}})$ if and only if $n \leq \mathrm{rank}_{top}(\mathcal{T}(S), d_{\mathcal{T}})$. In Theorem 1.4 he shows that there are quasi-isometric embeddings of the *entire* Euclidean space of dimension $\mathrm{rank}_{top}(\mathcal{T}(S), d_{\mathcal{T}})$ into Teichmüller space if and only if the genus of S is at most 1 or S is a closed surface of genus 2. We are grateful to him for pointing out an error in a previous version of this paper.

Remark 1.2. 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 (8) (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 is 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 $\mathrm{Mod}(S)$ 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 $\mathrm{Mod}(S)$ 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 develop further 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 ϵ_0 and R_0 there is R_1 so that for $R \geq R_1$ the following holds. Let $f\colon B \to \mathcal{Y}$ 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 $R' \overset{*}{\succ} R_0$ so that f restricted to B' is ϵ_0–efficient on scale R'.*

Here, efficient means that the image of every line in the box satisfies a reverse triangle inequality 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 that if $f : \mathbb{R}^n \to \mathbb{R}^m$ is lipschitz, then f is differentiable almost everywhere. In Theorem D the sub-box B' is an analogue of a point of differentiability.

The importance of efficiency lies first of all in the fact that in a product space the projection of an efficient path to a factor is still efficient. The corresponding statement for quasi-geodesics is false. Furthermore efficient maps into hyperbolic spaces are easy to control (Proposition 2.11). We also use the construction of Bestvina-Bromberg-Fujiwara [BBF10]. They embed the mapping class group into a product of finitely many hyperbolic spaces. Their construction is axiomatic and can be adopted easily to embed any of our spaces $\mathcal{X}$ 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 conclusion of this discussion, Theorem 4.9 will be that efficient paths fellow travel paths with the same endpoints that have nice properties. These latter paths which we call preferred paths will play the role of geodesics.

The power then of Theorem D lies in the fact that one can add the assumption of efficiency for free, just by replacing B with a sub-box B'. Altogether this will mean on large boxes the image of every line fellow travels a preferred path.

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 $\mathcal{X}$ 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 $\mathcal{Y}$ and $\mathcal{Z}$ are geodesic metric spaces. We say a map $f\colon \mathcal{Y} \to \mathcal{Z}$ is quasi-Lipschitz if there are constants K and C so that

$$d_{\mathcal{Z}}\big(f(x_1), f(x_2)\big) \leq K d_{\mathcal{Y}}(x_1, x_2) + C.$$

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 $\mathfrak{a}$ and $\mathfrak{b}$ represent various quantities and let M and C' be uniform constants. We say $\mathfrak{a}$ *is less than* $\mathfrak{b}$ *up to a multiplicative error*, $\mathfrak{a} \stackrel{*}{\prec} \mathfrak{b}$, if $\mathfrak{a} \leq M\,\mathfrak{b}$. We say $\mathfrak{a}$ and $\mathfrak{b}$ are *comparable*, $\mathfrak{a} \stackrel{*}{\asymp} \mathfrak{b}$, if we have both $\mathfrak{a} \stackrel{*}{\prec} \mathfrak{b}$ and $\mathfrak{b} \stackrel{*}{\prec} \mathfrak{a}$.

Using the similar notation when the error is additive or both additive and multiplicative, we say $\mathfrak{a} \stackrel{+}{\prec} \mathfrak{b}$ if $\mathfrak{a} \leq \mathfrak{b} + C'$ and $\mathfrak{a} \prec \mathfrak{b}$ if $\mathfrak{a} \leq M\mathfrak{a} + C'$. Again, $\mathfrak{a} \stackrel{+}{\asymp} \mathfrak{b}$ if we have both $\mathfrak{a} \stackrel{+}{\prec} \mathfrak{b}$ and $\mathfrak{b} \stackrel{+}{\prec} \mathfrak{a}$ and $\mathfrak{a} \asymp \mathfrak{b}$ if we have both $\mathfrak{a} \prec \mathfrak{b}$

and $\mathfrak{b} \prec \mathfrak{a}$. Also, we often use the notation $\mathfrak{a} = O(\mathfrak{b})$ to mean $\mathfrak{a} \overset{*}{\prec} \mathfrak{b}$. For example

$$\mathfrak{a} \overset{+}{\prec} \mathfrak{b} + O(\mathfrak{c}) \quad \Longleftrightarrow \quad \mathfrak{a} \leq \mathfrak{b} + M\mathfrak{c} + C',$$

for uniform constants M and C'.

Using this notation we may write

$$\mathfrak{a} \overset{+}{\succ} \mathfrak{b} \quad \text{and} \quad \mathfrak{b} \overset{+}{\succ} \mathfrak{c} \quad \Longrightarrow \quad \mathfrak{a} \overset{+}{\succ} \mathfrak{c}.$$

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 $\overset{+}{\succ}$ 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 and points along a line satisfy the reserves triangle inequality. We emulate these concepts by introducing the notion of an ϵ–efficient paths where the points along this path satisfy the reverse triangle inequality up to a small multiplicative error.

Definition 2.1. Let $\mathcal{Z}$ be a metric space, $\gamma\colon [a,b] \to \mathcal{Z}$ be quasi-Lipschitz and $R > 0$ be a scale so that $|b-a| \overset{*}{\prec} R$. 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* γ *on the scale* r to be

$$\Delta(\gamma, r) = \min_{r\text{–partitions}} \sum_{i=0}^{m-1} d_{\mathcal{Z}}(z_i, z_{i+1}). \tag{1}$$

We say γ is ϵ–efficient on the scale R if

$$\Delta(\gamma, \epsilon R) \leq d_{\mathcal{Z}}\big(\gamma(a), \gamma(b)\big) + O(\epsilon R). \tag{2}$$

We establish some elementary properties of efficient paths.

Lemma 2.2. *Consider a map* $\gamma\colon [a,b] \to \mathcal{Z}$.

(1) *Suppose* γ *is* ϵ*–efficient on scale* R *and* k *is a uniformly bounded integer. Then for* k *points* $a \leq s_i \leq b$

$$\sum_{i=0}^{k-1} d_{\mathcal{Z}}\big(\gamma(s_i), \gamma(s_{i+1})\big) = d_{\mathcal{Z}}\big(\gamma(a), \gamma(b)\big) + O(\epsilon R).$$

(2) *For* $[c,d] \subset [a,b]$, *if* γ *is* ϵ*–efficient at scale* R *so is* $\gamma' = \gamma|_{[c,d]}$.

(3) *Assume* $\mathcal{Z} = \mathcal{Z}_1 \times \ldots \times \mathcal{Z}_l$ *equipped with the* L^1*–metric and let* γ_i *be the projection of* γ *to* $\mathcal{Z}_i$. *Then, if* γ *is* ϵ*–efficient at scale* R *so is every* γ_i.

(4) *If $\mathcal{Z}'$ is a subset of $\mathcal{Z}$ and γ is an ϵ^2–efficient path that is contained in an $O(\epsilon^2 R)$–neighborhood of $\mathcal{Z}'$ then the closest point projection of γ to $\mathcal{Z}'$ is an ϵ–efficient path.*

Proof. Let $t_0, \dots, t_m$ be an ϵR–partition of $[a,b]$ achieving the minimum in the definition of $\Delta(\gamma, \epsilon R)$. Add the points s_j to the partition. This 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 γ is quasi-Lipschitz,

$$d_{\mathcal{Z}}\big(\gamma(t_i), \gamma(s_j)\big) + d_{\mathcal{Z}}\big(\gamma(s_j), \gamma(t_{i+1})\big) = O(\epsilon R).$$

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 $t_0, \dots, t_m$ 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 $\{t_i\}$. Let $z_i = \gamma(t_i)$. Letting $c = t_j$ and $d = t_k$ we have

$$\begin{aligned}
\Delta(\gamma, \epsilon R) &\geq \sum_{i=1}^{m} d_{\mathcal{Z}}(z_i, z_{i+1}) - O(\epsilon R) \\
&\geq d_{\mathcal{Z}}\big(\gamma(a), \gamma(c)\big) + \sum_{i=j}^{k-1} d_{\mathcal{Z}}(z_i, z_{i+1}) + d_{\mathcal{Z}}\big(\gamma(d), \gamma(b)\big) - O(\epsilon R) \\
&\geq d_{\mathcal{Z}}\big(\gamma(a), \gamma(c)\big) + \Delta(\gamma', \epsilon R) + d_{\mathcal{Z}}\big(\gamma(d), \gamma(b)\big) - O(\epsilon R).
\end{aligned}$$

Also, by definition,

$$\Delta(\gamma, \epsilon R) \leq d_{\mathcal{Z}}\big(\gamma(a), \gamma(b)\big) + O(\epsilon R).$$

Hence,

$$\begin{aligned}
\Delta(\gamma', \epsilon R) &\leq d_{\mathcal{Z}}\big(\gamma(a), \gamma(b)\big) - d_{\mathcal{Z}}\big(\gamma(a), \gamma(c)\big) - d_{\mathcal{Z}}\big(\gamma(d), \gamma(b)\big) + O(\epsilon R) \\
&\leq d_{\mathcal{Z}}\big(\gamma(c), \gamma(d)\big) + O(\epsilon R).
\end{aligned}$$

This finishes the proof of part 2.

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

$$\sum_{i=1}^{m} d_{\mathcal{Z}_1}(z_i, z_{i+1}) + d_{\mathcal{Z}_2}(z_i, z_{i+1}) \leq d_{\mathcal{Z}_1}\big(\gamma(a), \gamma(b)\big) + d_{\mathcal{Z}_2}\big(\gamma(a), \gamma(b)\big) + O(\epsilon R). \tag{3}$$

But, by triangle inequality, we have

$$\sum_{i=1}^{m} d_{\mathcal{Z}_2}(z_i, z_{i+1}) \geq d_{\mathcal{Z}_2}\big(\gamma(a), \gamma(b)\big). \tag{4}$$

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

$$\Delta(\gamma_1, \epsilon R) \leq \sum_{i=1}^{m} d_{\mathcal{Z}_1}(z_i, z_{i+1}) \leq d_{\mathcal{Z}_1}\big(\gamma(a), \gamma(b)\big) + O(\epsilon R).$$

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

$$|s_{i+1} - s_i| \stackrel{*}{\asymp} \epsilon R,$$

Then, $l \stackrel{*}{\prec} \frac{1}{\epsilon}$. Also, let γ' be the path obtain from composing γ with the closest point projection to $\mathcal{Z}'$. Let $z_i = \gamma(s_i)$ and $z'_i = \gamma'(s_i)$. Then

$$\begin{aligned}
\Delta(\gamma', \epsilon R) &\leq \sum_{i=1}^{l} d_{\mathcal{Z}'}(z'_i, z'_{i+1}) \\
&\leq \sum_{i=1}^{l} d_{\mathcal{Z}}(z'_i, z_i) + d_{\mathcal{Z}}(z_i, z_{i+1}) + d_{\mathcal{Z}}(z_{i+1}, z'_{i+1}) \\
&\leq \Delta(\gamma, \epsilon R) + l \cdot O(\epsilon^2 R) \\
&\leq \Delta(\gamma, \epsilon^2 R) + O(\epsilon R) \leq d_{\mathcal{Z}'}(z'_1, z'_l) + O(\epsilon R).
\end{aligned}$$

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. $\square$

Definition 2.3. A *box* in $\mathbb{R}^n$ is a product of intervals, namely $B = \prod_{i=1}^{n} I_i$, where I_i is an interval in $\mathbb{R}$. We say a box B is *of size* R if for every i, $|I_i| \stackrel{*}{\succ} R$ and if the diameter of B is less than R. Note that if B is of size R and of size R', then $R \stackrel{*}{\asymp} R'$.

A map $f\colon B \to \mathcal{Z}$ from a box of size R in $\mathbb{R}^n$ to a metric space $\mathcal{Z}$ is called *ϵ–efficient* if, for any geodesic $\gamma\colon [a,b] \to B$, the path $f \circ \gamma$ is ϵ–efficient at scales R.

Let B be a box of size L in $\mathbb{R}^n$ and let $\underline{B}$ be a central sub box of B with comparable diameter (say a half). For any constant $0 < R < L$, let $\mathcal{B}_R$ be a subdivision $\underline{B}$ to boxes of size R. That is,

1. boxes in $\mathcal{B}_R$ are of size R,
2. they are contained in $\underline{B}$ and hence their distance to the boundary of B is comparable to L,
3. they have disjoint interiors and
4. their union is $\underline{B}$.

For any metric space $\mathcal{Z}$, we prove that any quasi-Lipschitz maps from B to $\mathcal{Z}$ is coarsely differentiable almost everywhere in a central box of comparable size:

Theorem 2.4 (Coarse Differentiation)**.** *For every ϵ_0, θ_0 and R_0 there is L_0 so that the following holds. For $L \geq L_0$, let $f\colon B \to \mathcal{Z}$ be a quasi-Lipschitz map where B is a box of size L in $\mathbb{R}^n$. Then there is a scale $R \geq R_0$ so that the proportion of boxes $B' \in \mathcal{B}_R$ where $f|_{B'}$ is ϵ_0–efficient is at least* $(1-\theta_0)$.

Remark 2.5. Note that the size of the error, $\epsilon_0 R$, depends on the size of the boxes. An ϵ_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 as a proportion of the size of the box.

Theorem 2.4 is stronger than what we need as we will need only one efficient box. However, this more general statement may be useful for other applications of coarse differentiation.

2.1. **Choosing the correct scales.** We first prove a much coarser differentiation statement. In a sense, the statement of Theorem 2.4 is a direct analogue of Rademacher's theorem, but the proof of Rademacher's theorem a direct analogue of proof of Theorem 2.7 below.

Definition 2.6. A family $\mathcal{F}$ of geodesics in $\mathbb{R}^n$ is called *locally finite* if, for any compact subset B of $\mathbb{R}^n$, 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

$$\mathcal{F}_B = \left\{ \gamma \;\middle|\; \gamma = \gamma' \cap B, \quad \gamma' \in \mathcal{F}, \quad |\gamma| \overset{*}{\succ} L \right\}.$$

For $\gamma \in \mathcal{F}_B$ we say a set of points $\mathrm{G}(\gamma, r)$ along γ is an *r–grid for γ* if they subdivide γ to segments of size exactly r except perhaps for the two end segments which may have a size less than r. An r–grid $\mathrm{G}(r)$ is a collection of r–grids for every segment in $\mathcal{F}_B$. When an r–grid $\mathrm{G}(\gamma, r) = \{p_1, \dots, p_k\}$ is fixed, we define

$$\overline{\Delta}(\gamma, r) = \sum_{i=1}^{k-1} d_{\mathcal{Z}}(f(p_i), f(p_{i+1})).$$

This is essentially the same as the definition of Δ except the sum is over a given r–grid instead of minimum over all r–partitions. Given a scale R, a segment $\gamma \in \mathcal{F}_B$ with an ϵR–grid $\mathrm{G}(\epsilon R)$, we define $\mathcal{F}(\gamma, R)$ to be the set of all subsegments of γ of length R that start and end at points in $\mathrm{G}(\gamma, \epsilon R)$. We also define

$$\mathcal{F}_B(R) = \bigcup_{\gamma \in \mathcal{F}_B} \mathcal{F}(\gamma, R).$$

Theorem 2.7. *Let $\mathcal{F}$ be a locally finite family of geodesics in $\mathbb{R}^n$. For any $\epsilon > 0$, $\theta > 0$ and R_0, there exist a constant L_0 such that the following holds. Let $L > L_0$, $B \subset \mathbb{R}^n$ be a box of size L and $f \colon B \to \mathcal{Z}$ be a quasi-Lipschitz map. Then, there exist a scale $R \geq R_0$ and an ϵR–grid $G(\epsilon R)$ such that, for at least $(1-\theta)$ fraction of segments $\gamma' \in \mathcal{F}_B(R)$,*

$$\overline{\Delta}(\gamma', \epsilon R) \overset{*}{\prec} d_{\mathcal{Z}}(f(a), f(b)) + \epsilon R. \tag{5}$$

where a, b are the endpoints of γ'.

Note that, Equation (5) implies that γ' is ϵ–efficient at scale R. But this statement is more suitable for our proof. Morally, the lemma states that $f|_B$ is nearly affine on scales R, up to an error of $O(\epsilon R)$.

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

Proof of Theorem 2.7. Pick $r_0 \geq \max\{R_0, C\}$ (C is 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 $L_1 = r_M$. Choose an arbitrary r_0–grid $\mathrm{G}(r_0)$ for $\mathcal{F}_B$ and let θ_1 be the fraction of segments in $\mathcal{F}_B(r_1)$ for which Equation (5) does not hold. If $\theta_1 \leq \theta$ then we are done. Thus assume $\theta_1 > \theta$.

For $\gamma \in \mathcal{F}_B$, we choose an r_1–grid $\mathrm{G}(\gamma, r_1) \subset \mathrm{G}(\gamma, r_0)$ as follow: Note that $\mathrm{G}(\gamma, r_1)$ is essentially a *decomposition* of γ into segments of length r_1 (except for the subsegments in the ends). That is, we are choosing a non-overlapping subset of $\mathcal{F}(\gamma, r_1)$ so that the next segment starts where the previous segment ended. We choose the decomposition $\mathcal{D}(\gamma, r_1)$ so that the proportion $\theta_1(\gamma)$ of segments that do not satisfy Equation (5) is maximum. Hence, the average of these proportions is larger than θ_1.

For $m = 1, \dots, M$, we proceed the same way. If $\theta_m \leq \theta$ we are done. Otherwise, for every γ, we choose the decomposition $\mathcal{D}(\gamma, r_m)$ where the proportion $\theta_m(\gamma)$ of segments that do not satisfy Equation (5) is maximum and use it to define the r_m–grid $\mathrm{G}(\gamma, r_m)$. Again we have

$$\frac{\sum_\gamma \theta_m(\gamma)}{|\mathcal{F}_B|} \geq \theta_m > \theta.$$

We show that, if M is large enough this contradict the assumption that f is quasi-Lipschitz. First, note that:

$$\overline{\Delta}(\gamma, r_{m-1}) - \overline{\Delta}(\gamma, r_m) \overset{*}{\succ} (\epsilon\, r_m)\, \theta_m(\gamma) |\mathcal{D}(\gamma, r_m)| \overset{*}{\asymp} \epsilon\, L\, \theta_m(\gamma).$$

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

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

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

$$\Delta(\gamma, r_0) \leq \frac{L}{r_0}(K r_0 + C) \overset{*}{\prec} KL.$$

Hence,

$$KL \overset{*}{\succ} \epsilon\, L \sum_{m=1}^{M} \theta_m(\gamma),$$

and thus
$$\frac{K}{\epsilon} \stackrel{*}{\succ} \sum_{m=1}^{M} \theta_m(\gamma).$$
Average over all geodesics $\gamma \in \mathcal{F}_B$ to get
$$\frac{K}{\epsilon} \stackrel{*}{\succ} \sum_{m=1}^{M} \left(\frac{1}{|\mathcal{F}_B|} \sum_{\gamma \in \mathcal{F}_B} \theta_m(\gamma) \right) \stackrel{*}{\succ} \sum_{m=1}^{M} \theta_m \geq M\theta.$$
Choosing M large enough we obtain a contradiction. Hence, for some m, $\theta_m \leq \theta$ and we are done. $\square$

Proof of Theorem 2.4. Let L_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 vectors $\mathcal{V}$ in the unit sphere $S^{n-1} \subset \mathbb{R}^n$ that is $(\epsilon_0)^2$ dense in S^{n-1} with the size $|\mathcal{V}| \stackrel{*}{\asymp} \frac{1}{(\epsilon_0)^2}$. For a direction $\vec{v} \in \mathcal{V}$, 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
$$\mathcal{F} = \bigcup_{\vec{v} \in \mathcal{V}} \mathcal{F}_{\vec{v}}$$
is a locally finite family of geodesic in $\mathbb{R}^n$. Let
$$\epsilon = \epsilon_0^2, \quad \text{and} \quad \theta \ll \theta_0 \, \epsilon^{n+2}.$$
Apply Theorem 2.7 to obtain the constants L_0. Assume a box B of size $L \geq L_0$ and a quasi-Lipschitz map $f\colon B \to \mathcal{Z}$ are given and let R be the scale obtained from Theorem 2.7.

Let $\mathcal{B}_R$ be a collection of disjoint sub-boxes of B giving a decomposition of a central box in B as in the statement of Theorem 2.4. Let $B' \in \mathcal{B}$ be a box that contains a geodesic β that is not ϵ_0–efficient at scale R. Let $\vec{v}$ be the direction closest to the direction of β. Let
$$\mathcal{NE}(B', \vec{v}) \subset \mathcal{F}_B(R)$$
be the set of geodesic segments in $\mathcal{F}_{B,R}$ that are in the direction of $\vec{v}$, intersect B' and are not ϵ_0–efficient on scale R.

Claim: Every geodesic in $\mathcal{F}_{\vec{v}}$ that intersects an ϵR–neighborhood of β contains a segment in $\mathcal{NE}(B', \vec{v})$.

Proof of claim: Assume $\gamma \in \mathcal{F}_{\vec{v}}$ intersects an ϵR–neighborhood of β. The condition (2) of description of $\mathcal{B}_R$ implies that $|\gamma \cap B| \stackrel{*}{\succ} L$ and its subsegments of length R that start and end in $\mathrm{G}(\epsilon R)$ are included $\mathcal{F}_{B,R}$. Since, the difference between the direction of γ and β is at most $\epsilon = \epsilon_0^2$, in fact β is contained in an ϵR–neighborhood of γ. Also, the length of β is less than the diameter of B' which is less than R. Hence, there is a segment $\gamma_0 \in \mathcal{F}_B(\gamma, R)$ of length R where β is included in an ϵR–neighborhood of γ_0 (refer to Fig. 1). We show that if β is not ϵ_0–efficient on scale R, then γ_0 will not be ϵ–efficient on scale R which is what we claimed.

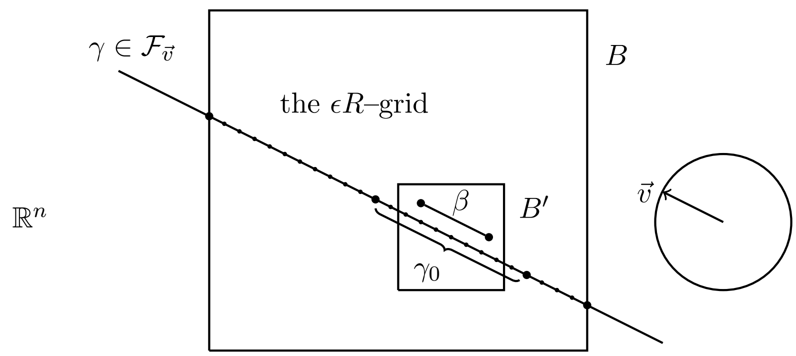

FIGURE 1. The arc γ is in a ϵR–neighborhood of β_0.

Assume, for contradiction, that γ_0 is ϵ–efficient on scale R. Then every sub-segment of γ_0 is also efficient on scale R (Lemma 2.2). Choose a sub-segment γ_1 of γ_0 so that the end points of γ_1 and β are ϵR–close. We now apply the last conclusion of Lemma 2.2 with $\mathcal{Z}' = f(\gamma_0)$ to conclude that β is ϵ_0–efficient, a contradiction. □

Let $\mathcal{F}_{B'}(R)$ be the subset of $\mathcal{F}_B(R)$ consisting of segments that intersect B'. In every direction $\vec{v} \in \mathcal{V}$ there are at most $\frac{R^{n-1}}{\epsilon}$ segments in $\mathcal{F}_{B'}(R)$. This is because a cross section of B' perpendicular to $\vec{v}$ has an area at most R^{n-1} and the grid has size ϵR. Assuming B' contains an non-efficient segment β, the number of geodesics in $\mathcal{F}_{\vec{v}}$ that intersect an ϵR–neighborhood of β is of order of $(\epsilon R)^{n-1}$ (which is the area of a cross section of an ϵR–neighborhood of β perpendicular to $\vec{v}$). That is

$$\frac{\big|\mathcal{NE}(B', \vec{v})\big|}{|\mathcal{F}_{B'}(R)|} \overset{*}{\succ} \frac{1}{|\mathcal{V}|} \cdot \frac{(\epsilon R)^{n-1}}{\frac{R^{n-1}}{\epsilon}} \overset{*}{\asymp} \epsilon^{n+2}.$$

Note that a definite proportion of segments in $\mathcal{F}_B(R)$ intersect some box $B' \in \mathcal{B}$ and each segment in $\mathcal{F}_B(R)$ intersects at most a uniform umber of boxes. Hence

$$|\mathcal{F}_B(R)| \overset{*}{\asymp} \sum_{B' \in \mathcal{B}} |\mathcal{F}_{B'}(R)|.$$

Define

$$\mathcal{NE}(B') = \bigcup_{\vec{v} \in \mathcal{V}} \mathcal{NE}(B', \vec{v}) \qquad \text{and} \qquad \mathcal{NE} = \bigcup_{B' \in \mathcal{B}} \mathcal{NE}(B')$$

Assume the proportion of boxes B' that contain a non-efficient segment is larger than θ_0. Since the sizes of $\mathcal{F}_{B'}(R)$ are comparable for every B', we have

$$\frac{|\mathcal{NE}|}{|\mathcal{F}_B(R)|} \geq \frac{1}{|\mathcal{B}|} \sum_{B' \in \mathcal{B}} \frac{\big|\mathcal{NE}(B')\big|}{|\mathcal{F}_{B'}(R)|} \overset{*}{\succ} \frac{\theta_0 |\mathcal{B}|}{|\mathcal{B}|} \epsilon^{n+2} \gg \theta.$$

The contradiction finishes the proof. □

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 $\mathcal{X}$ is a Gromov hyperbolic space and $\gamma\colon [a,b] \to \mathcal{X}$ is ϵ–efficient on scale R. Then γ stays in an $O(\epsilon R)$–neighborhood of a geodesic ℓ joining $\gamma(a)$ and $\gamma(b)$.*

Proof. Suppose, for a large M, that the path γ leaves an $M\epsilon R$–neighborhood of ℓ. Fix a constant $D_0 \geq C$. We can find times $c, d \in [a,b]$ so that at times c, d, γ is distance D_0 from ℓ; for $t \in [c,d]$, $\gamma(t)$ is at least D_0 away from ℓ, and so that in between c and d the path γ travels to a point at of a distance $M\epsilon R$ from ℓ.

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

$$\Delta(\gamma', \epsilon R) \overset{*}{\succ} \sum_i d_{\mathcal{X}}(z_{i+1}, z_i). \tag{6}$$

(In fact, by definition of Δ, 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

$$d_{\mathcal{X}}(z_{i+1}, z_i) \overset{*}{\asymp} \epsilon R.$$

Note that after removing points from the partition, Equation (6) still holds. Since $t_{i+1} - t_i \overset{*}{\asymp} \epsilon R$ we have $N \overset{*}{\succ} M$. By the contraction property of hyperbolic spaces, there is a uniform constant D_1 and a projection map $\pi\colon \mathcal{X} \to \ell$ such that

$$d_{\mathcal{X}}\big(\pi(z_i), \pi(z_{i+1})\big) \leq D_1.$$

By using these projected points to ℓ and since $\gamma(c)$ and $\gamma(d)$ are at distance N_0 from ℓ we have

$$d_{\mathcal{X}}\big(\gamma(c), \gamma(d)\big) \leq 2D_0 + ND_1.$$

On the other hand

$$\Delta(\gamma', \epsilon R) \overset{*}{\succ} \sum d_{\mathcal{X}}(z_{i+1}, z_i) \overset{*}{\succ} N\epsilon R.$$

From the assumption that γ' is ϵ–efficient on scale R, we have

$$2D_0 + ND_1 + O(\epsilon R) \overset{*}{\succ} N\epsilon R.$$

This implies N is uniformly bounded. Hence, M is also uniformly bounded. This finishes the proof. □

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 σL.*

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 minimum distance between supporting hyperplanes.

From Theorem 1 in [HCSSG04] we have, for some $\kappa = \kappa(n) > 0$, and any convex set C that

$$\mathrm{Vol}(C) \leq \kappa R^n \int_0^{\arcsin \frac{w}{2R}} \cos^n \theta d\theta \leq \kappa R^n \arcsin \frac{w}{2R} \leq \kappa R^n \arcsin \frac{c\,r}{2R}.$$

The last inequality follows from the Steinhagen inequality, which states that there is a constant $c = c(n) > 0$, such that

$$w \leq c\,r.$$

Since the convex sets C_i cover the ball of radius L, for some $c' > 0$, there is some $C = C_i$ with

$$\mathrm{Vol}(C) \geq c' L^n / N.$$

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,

$$\arcsin \frac{c\,r}{2R} \geq \frac{\pi}{4}.$$

Then $\frac{c\,r}{2R} \geq \frac{\sqrt{2}}{2}$ and so

$$c\,r \geq \sqrt{2} R \geq \sqrt{2} c'' L,$$

and we are done by taking $\sigma = \frac{\sqrt{2}c''}{c}$. Now assume

$$\arcsin \frac{c\,r}{2R} \leq \frac{\pi}{4}$$

so that

$$\arcsin \frac{c\,r}{2R} \leq \frac{c\,r}{R}.$$

But then

$$c' L^n / N \leq \mathrm{Vol}(C) \leq \kappa\, R^n \frac{c\,r}{R} = \kappa\, c\, R^{n-1}\, r$$

and so

$$r \geq \frac{c'}{c\,N\kappa} \left(\frac{L}{R}\right)^{n-1} L,$$

and again we are done by taking $\sigma = \frac{c'}{c\,(c'')^{n-1} N\kappa}$. □

Proposition 2.11. *Suppose $\mathcal{Z}$ is a Gromov hyperbolic space and $f\colon B \to \mathcal{Z}$ is an ϵ–efficient map from a box of size R in $\mathbb{R}^n$ to $\mathcal{Z}$. Then, there is a sub-box $B' \subset B$ with $|B'| \stackrel{*}{\asymp} |B|$, so that the image $f(B')$ lies in an $O(\epsilon R)$–neighborhood of a line ℓ' in $\mathcal{Z}$.*

Proof. By taking a sub-box we assume $B = [0, R]^n$ and let ℓ_i be the edges of the box B. Given a line $\ell \subset B$ denote by ℓ' a geodesic in $\mathcal{Z}$ 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 ℓ_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 ℓ'_i. Now let ℓ_q be the line through q parallel to ℓ_1. Take the rectangle with sides ℓ_q, ℓ_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 ℓ'_2 and hence m'_2 are within $O(\epsilon R)$ of ℓ'_2. The same holds for m'_4 and ℓ'_4. The quadrilateral bounded by $\ell'_1, m'_2, \ell'_q, m'_4$ 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 ℓ'_i.

Now suppose the statement is true for boxes in $\mathbb{R}^{n-1}$ and $B \subset \mathbb{R}^n$. Take again the geodesics ℓ_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 τ_i be the edges of B_q^{n-1}. By induction, $f(q)$ lies within $O(\epsilon R)$ of some τ'_i. Since each τ_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 ℓ'_i. Thus $f(q)$ is within $O(\epsilon R)$ of some ℓ'_i, completing the induction step.

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

3. Combinatorial Model

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

$$\xi(S) = \sum_W (3\mathsf{g}_W + \mathsf{p}_W - 4), \tag{7}$$

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 $(\mathcal{T}(S), d_{\mathcal{T}})$ represent the Teichmüller space equipped with the Teichmüller metric, $(\mathcal{T}(S), \mathrm{d}_{WP})$ represent Teichmüller space equipped with the Weil-Petersson metric and $(\mathrm{Mod}(S), d_W)$ 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 τ_α, for each pants curve α. The transverse curves are assumed to be disjoint from other curves in P and to intersect α 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 l_α (length of α) associated to to every pants curve α. 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 nom-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 $X = \mathcal{M}(S)$.
- $\mathcal{C}(A)$ 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)

$$\pi_W \colon \mathcal{X} \to \mathcal{C}(W).$$

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 γ. If γ is not contained in W then $\gamma \cap W$ is a collection of arcs with endpoints on ∂W. Choose one such arc and perform a surgery using this arc and a subarc of ∂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 ω_0 with zero and sending every other arc ω to the signed intersection number between ω and ω_0. Define $\pi_A(x)$ by lifting the pants deposition P_x and transverse curves τ_α 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 α and denote it by $\operatorname{twist}_\alpha(x)$, 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 $\mathcal{C}(A)$ with the subset $H \subset \mathbb{R}^2$ of all points in $\mathbb{R}^2$ where the y–coordinate is larger than $1/\mathsf{B}$. Note that, for an augmented marking $x = (P, \{\tau_\alpha\}, \{l_\alpha\})$, the twisting number $\operatorname{twist}_\alpha(x)$ can still be defined as above. If the core curve of A is in P_x we define

$$\pi_A(x) = \big(\operatorname{twist}_\alpha(x), 1/l_\alpha\big),$$

otherwise

$$\pi_A(x) = \big(\operatorname{twist}_\alpha(x), 1/B\big).$$

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

$$\pi_{U,V} \colon \mathcal{C}(U) \to \mathcal{C}(V).$$

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

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

In fact, when W is an annulus with core curve α, 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

$$d_{\mathcal{X}}(x, y) = \sum_{W \in \mathcal{W}_T(x,y)} d_W(x, y), \tag{8}$$

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 $x, y, z \in \mathcal{X}$

$$d_{\mathcal{X}}(x, y) + d_{\mathcal{X}}(y, z) \overset{*}{\succ} d_X(x, z).$$

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])

$$\sum_{W \in \mathcal{W}_T(x,y)} d_W(x, y) \overset{*}{\asymp} \sum_{W \in \mathcal{W}_{T'}(x,y)} d_W(x, y). \tag{9}$$

Even though this is not a metric, it makes sense to say $\mathcal{X}$ 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 $d_{\mathcal{X}}(x, y)$ is a well defined number for all $x, y \in \mathcal{X}$. 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 $(\mathrm{Mod}(S), d_W)$ can be coarsely represented as a marking [MM99]. That is, there is a coarsely defined map

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

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

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

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

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 $\alpha \in P$, τ_α is the shortest transverse curve to α in X and l_α is the hyperbolic length of α 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 $V \subsetneq U$ are subsurfaces of S and $\beta_1, \ldots, \beta_k$ is a geodesic in $\mathcal{C}(U)$. Then either there is some β_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 α and $\mathcal{C}(V)$ is a horoball.

Then, applying the $\mathcal{M}(S)$ version of Theorem 3.1, we have that either some β_j is disjoint from α or

$$\mathrm{twist}_\alpha(\beta_1) \stackrel{+}{\asymp} \mathrm{twist}_\alpha(\beta_k).$$

Also,

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

Hence,

$$d_\alpha(\beta_1, \beta_k) \stackrel{+}{\asymp} \log |\,\mathrm{twist}_\alpha(\beta_1) - \mathrm{twist}_\alpha(\beta_k)| = O(1).$$

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 $U \pitchfork V$ means that $U \cap V \neq \emptyset$ and neither is contained in the other. Consider the following *consistency condition* on a tuple

$$\mathsf{z} \in \prod_U \mathcal{C}(U).$$

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

(1) Whenever $U \pitchfork V$,

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

(2) If $V \subsetneq U$, then

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

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 α, the projection of P to $\mathcal{C}(\alpha)$ may not be close to z_α. We show that the set of such curves is a multi-curve.

Let α and β 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 α and β are disjoint. Assume, for contradiction, that they intersect and let U be the surface they fill. Then z_U intersects either α to β (say α without loss of generality). By construction of P_0

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

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 α (where $d_\alpha(P_0, z_\alpha)$ is large) and for every subsurface U intersecting α,

$$d_U(\alpha, P_0) \stackrel{+}{\asymp} d_U(\alpha, z_U) = O(1).$$

This is because, if z_U is far from α in $\mathcal{C}(U)$, then the projection of z_U to $\mathcal{C}(\alpha)$ is defined and is (z is constant) near z_α. 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 $\boldsymbol{\alpha}$ be the multi curve consisting of all the curves above. Since, $d_U(P_0, \boldsymbol{\alpha}) = O(1)$ for every non-annular subsurface U, $\boldsymbol{\alpha}$ can be extended to a pants decomposition P with $d_U(P_0, P) = O(1)$ for every non-annular subsurface U. That is,

$$d_V(P, z_V) = O(1), \qquad \text{for every subsurface } V \text{ intersecting } P.$$

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_β is an integer and l_β is a real number less than the Bers constant. Let τ_β be a curve intersecting β minimally and disjoint from other curves in P with $\operatorname{twist}_\beta(\tau_\beta) \stackrel{+}{\asymp} t_\beta$ (this can always be achieved by applying Dehn twists around β). Now,

$$x = (P, \{\tau_\beta\}_{\beta \in P}, \{l_\beta\}_{\beta_P}).$$

is the desired augmented marking. This is because, for $\beta \in P$, the projections to $\mathcal{C}(\beta)$ are close to z_β 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 ∂V intersects U. The for any $x \in X$, if*

$$d_U(x_U, \partial V) > M_1 \qquad \textit{then} \qquad d_V(x_V, x_U) = O(1).$$

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

Otherwise, ∂U intersects V and has a defined projection to $\mathcal{C}(V)$. Since x_U is disjoint from ∂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

$$\phi_W \colon \mathcal{X} \to \mathcal{X}(W),$$

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

$$d_{\mathcal{X}(W)}(x, y) = d_{\mathcal{X}(W)}\big(\phi_W(x), \phi_W(y)\big)$$

The subsurface W is allowed to be an annulus in which case, $\mathcal{X}(W) = \mathcal{C}(W)$.

For a curve system $\boldsymbol{\alpha}$, let $\mathcal{X}_{\boldsymbol{\alpha}}$ be the set of points in $x \in \mathcal{X}$ where $\boldsymbol{\alpha}$ is a subset of the pants decomposition P_x associated to x. Consider a point $x \in \mathcal{X}_{\boldsymbol{\alpha}}$. Since every curve in x is disjoint from $\boldsymbol{\alpha}$, the projection of x to any subsurface intersecting $\boldsymbol{\alpha}$ is distance at most 2 from the the projection of $\boldsymbol{\alpha}$. 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 $\boldsymbol{\alpha}$, each of which is contained in some component W of $S \setminus \boldsymbol{\alpha}$. Therefore, the map

$$\Phi \colon \mathcal{X}_{\boldsymbol{\alpha}} \to \prod_W \mathcal{X}(W) \qquad \text{where} \qquad \Phi = \prod_W \phi_W$$

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 Φ 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 $\phi_{\mathcal{X}_{\boldsymbol{\alpha}}} \colon \mathcal{X} \to \mathcal{X}_{\boldsymbol{\alpha}}$; 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

$\mathcal{X}_{\boldsymbol{\alpha}}$. We define $d_{\mathcal{X}_\alpha}$ to mean the distance between projections to $\mathcal{X}_{\boldsymbol{\alpha}}$. That is, for $x, y \in \mathcal{X}$, we define

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

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

$$d_{\mathcal{X}_{\boldsymbol{\alpha}}}(x,y) \asymp \sum_W d_{\mathcal{X}(W)}(x,y), \tag{10}$$

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

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

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

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

Proof. Let M_0 be the constant in Theorem 3.1. Then any subsurface U with $d_U(x,y) \geq M_0$, is disjoint from some α_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 (10). □

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 $\mathbf{Y}$ be a collection of subsurfaces of S with the property that if $V, W \in \mathbf{Y}$ then

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

Define $\mathbf{P}_K(\mathbf{Y})$ to be a graph whose vertices are elements of $\mathbf{Y}$ and two vertices V, W are connected with an edge if for all $U \in \mathbf{Y}$

$$d_U(V,W) \leq K.$$

In [BBF10] it is shown that $\mathbf{P}_K(\mathbf{Y})$ is quasi-isometric to a tree.

Define $\mathcal{C}(\mathbf{Y})$ to be the space obtained from $\mathbf{P}_K(\mathbf{Y})$ by attaching a copy of $\mathcal{C}(W)$ for every vertex $W \in \mathbf{Y}$ as follows: The vertex set of $\mathcal{C}(\mathbf{Y})$ is the union of vertex sets of $\mathcal{C}(W)$, $W \in \mathbf{Y}$. If V and W are joined by an edge in $\mathbf{P}_K(\mathbf{Y})$, then we join the vertex $\pi_V(\partial W)$ in $\mathcal{C}(V)$ to $\pi_W(\partial V)$ in $\mathcal{C}(W)$. It follows from [BBF10, Theorem E], using the fact that each space $\mathcal{C}(W)$ is hyperbolic, that the resulting space $\mathcal{C}(\mathbf{Y})$ 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 $\mathbf{Y}^1, \dots, \mathbf{Y}^k$ each having the transversality property mentioned above.
- there is a finite index subgroup Γ in $\mathrm{Mod}(S)$ which fixes each $\mathbf{Y}^i$

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

We can assume $\mathbf{Y}^j$ 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

$$\mathbf{C} = \prod_{j=1}^{k} \mathcal{C}(\mathbf{Y}^j)$$

equipped with the L^1–metric. Thus $\mathbf{C}$ is a product of finitely many hyperbolic spaces.

We define a projection map $\Psi^j \colon \mathcal{X} \to \mathcal{C}(\mathbf{Y}^j)$ as follows: For $x \in \mathcal{X}$, choose a subsurface $W_j \in \mathbf{Y}^j$ that minimizes

$$\max_{\alpha \in P_x} i(\alpha, \partial W_j),$$

where the minimum is over $W_j \in \mathbf{Y}^j$. Define

$$\Psi^j(x) = x_{W_j}.$$

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

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 $\mathbf{Y}^j$ 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 $\mathbf{C}$, the image is independent of the choice of W_j.

First we note that since the finite index subgroup Γ acts preserving each $\mathcal{C}(\mathbf{Y}^j)$ 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 $\mathcal{C}(\mathbf{Y}^j)$. First, we check, for all $U \in \mathbf{Y}^j$, 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 α and V_j. The claim now follows from the triangle inequality. Therefore, W_j and V_j are connected by an edge in $\mathbf{P}_K(\mathbf{Y}^j)$. To show that x_{W_j} and x_{V_j} are close in $\mathcal{C}(\mathbf{Y}^j)$, we need that

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

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

$$\Psi\colon\ \mathcal{X} \to \mathbf{C} \qquad \text{with} \qquad \Psi = \prod_j \Psi^j.$$

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*

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

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 ψ^j are defined differently from [BBF10].

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

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

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

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

$$d_U(x, x') = O(1) \qquad \forall U.$$

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 $\mathbf{P}_K(\mathbf{Y}^j)$. We also know that x is close to x' which has bounded intersection with $\partial W_j'$. Hence

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

Therefore,

$$d_{\mathcal{C}(\mathbf{Y}^j)}\big(\Psi^j(x), \Psi^j(x')\big) = O(1).$$

This means, the maps Ψ^j are quasi-Lipschitz, and so is Ψ.

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,

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

and since the distance formula works for any threshold,

$$d_{\mathcal{X}}(x, y) \overset{*}{\asymp} \sum_{W \in \mathcal{W}_{K'}(x,y)} d_W(x, y).$$

Hence

$$d_{\mathbf{C}}\big(\Psi(x), \Psi(y)\big) \overset{*}{\succ} d_{\mathcal{X}}(x, y).$$

This finishes the proof. □

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

Definition 4.4. Given $x, y \in \mathcal{X}$, we say a quasi-geodesic $\omega\colon [a, b] \to \mathcal{X}$ is a *preferred path* connecting x to y if,

- $\omega(a) = x$, $\omega(b) = y$
- for every subsurface U, the map
$$\omega_U = \pi_U \circ \omega\colon [a, b] \to \mathcal{C}(U)$$
is an unparametrized quasi-geodesic.

Lemma 4.5. *For any $x, y \in \mathcal{X}$ 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 $\mathcal{X}$ 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 $\mathcal{C}(U)$. Denote a geodesic segment in $\mathcal{C}(U)$ connecting x_U to y_U by $[x, y]_U$. Given $\kappa > 0$, let

$$\mathcal{G}(x, y, \kappa) = \Big\{ z \in \mathcal{X} \;\Big|\; \forall U,\, d_U\big(z, [x, y]_U\big) \leq \kappa \Big\}.$$

This notion was introduced in [BKMM08] where they call it the hull. In a sense, this set if the union of all points in all preferred paths.

Lemma 4.7. *There is a constant κ_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 $\omega\colon [a,b] \to \mathcal{X}$ and any $a \leq t \leq b$,*

$$\omega(t) \in \mathcal{G}(x, y, \kappa_0).$$

Proof. Since $\mathcal{C}(U)$ is Gromov hyperbolic and the projection of ω is an unparametrized quasi-geodesic, it stays in a uniform neighborhood of the geodesic connecting its end points. □

When κ_0 is fixed, we drop κ_0 and denote this set by $\mathcal{G}(x, y)$.

Lemma 4.8. *Let $x, y \in \mathcal{X}$ and $w, z \in \mathcal{G}(x, y)$. Then*

- *For any subsurface U,*

$$d_U(w, z) \overset{+}{\prec} d_U(x, y).$$

  In fact, $[w, z]_U$ is contained in a uniform neighborhood of $[x, y]_U$.
- *(convexity) if $w, z \in \mathcal{G}(x, y)$ then $\mathcal{G}(w, z)$ is contained in a uniform neighborhood of $\mathcal{G}(x, y)$.*

Proof. Since $\mathcal{C}(U)$ is Gromox 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 $[x, y]_U$ □

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

Theorem 4.9. *Let $\gamma\colon [0, R] \to \mathcal{X}$ be an ϵ–efficient path connecting $x = \gamma(0)$ to $y = \gamma(R)$. Then, the image of γ stays in an $O(\epsilon R)$–neighborhood of $\mathcal{G}(x, y)$. 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 $x, y, z \in \mathcal{X}$ and a subsurface W, let η_W be the center of the triangle (x_W, y_W, z_W) guaranteed by the hyperbolicity of $\mathcal{C}(W)$. That is, η_W is δ_W–close to all three geodesics $[x, y]_W$, $[y, z]_W$ and $[x, z]_W$, where δ_W is the hyperbolicity constant of $\mathcal{C}(W)$.

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

This implies that x_U, y_U and η_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

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

And by triangle inequality

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

From the definition of η_V, we have

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

Again, using the triangle inequality, we get

$$d_V(\eta_U, \eta_V) = O(1). \tag{11}$$

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

Thus assume $U \pitchfork V$. Since η_U and ∂U are disjoint, $d_V(\eta_U, \partial U) = O(1)$. This and Equation (11) 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 η. We call η the center of the triangle x, y and z.

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

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

$$x_V \text{ is disjoint from } \partial V \quad \Longrightarrow \quad d_W(\partial V, x_V) = O(1).$$

and

$$\mathrm{i}(x, \partial V) = O(1) \quad \Longrightarrow \quad d_W(\partial V, x) = O(1).$$

The claim follows from the triangle inequality.

In [BBF10, Lemma 3.13] it is shown that a geodesic in $\mathcal{C}(\mathbf{Y}^j)$ 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 η_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 η_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, η_j is itself close to the $[x_j, y_j]$.

The same hold for x_j, z_j and y_j, z_j. Since, $\mathcal{C}(\mathbf{Y}^j)$ is Gromov hyperbolic, and η_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 γ to $\mathcal{C}(\mathbf{Y}^j)$. By Lemma 2.2, each γ_j is still ϵ–efficient. Since $\mathcal{C}(\mathbf{Y}^j)$ 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 η be the center of x, y, z. From the construction, we have $\eta \in \mathcal{G}(x, y)$. We estimate the distance between η and z.

By Lemma 4.11, η_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 η_j. Therefore,

$$d_{\mathcal{C}(\mathbf{Y}^j)}(z_j, \eta_j) \overset{+}{\prec} d_{\mathcal{C}(\mathbf{Y}^j)}\big(z_j, [x_j, y_j]\big) = O(\epsilon R).$$

It follows, since Ψ is coarsely Lipschitz and the metric in $\mathbf{C}$ is the L^1–metric, that

$$d_{\mathcal{X}}(z, \eta) \overset{*}{\prec} \sum_j d_{\mathcal{C}(\mathbf{Y}^j)}(z_j, \eta_j) = O(\epsilon R).$$

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

We prove the second statement, namely that, γ 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 ϵ–efficient path γ is allowed to backtrack up to $O(\epsilon R)$. We proceed therefore as follows.

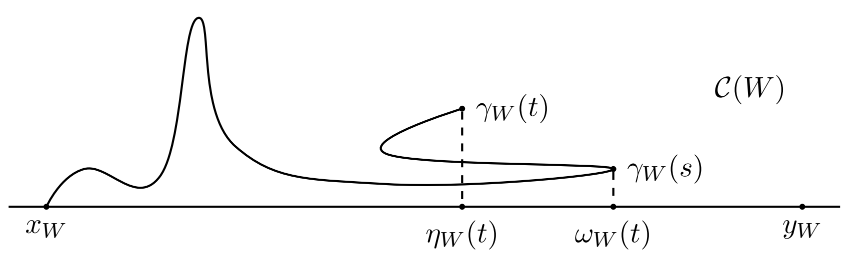

FIGURE 2. The point $\omega_W(t)$ is defined to be the point $\eta_W(s)$ that is farthest along in $[x, y]_W$, for $s \in [0, t]$.

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 $\mathcal{C}(W)$. Let $s \in [0, t]$ be a time where $d_W\big(x_W, \eta_W(s)\big)$ is maximized (see Fig. 2) and define

$$\omega_W(t) = \eta_W(s).$$

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 $U \pitchfork V$ and, without loss of generality, $d_V(\partial U, y) = O(1)$. Assuming

$$d_U\big(\omega_U(t), \partial V\big) \qquad \text{is large,} \tag{12}$$

we need to show $d_V\big(\omega_V(t), \partial U\big) = O(1)$. Since $\omega_U(t) = \eta_U(s)$ and $\{\eta_W(s)\}$ is consistent, Equation (12) implies

$$d_V\big(\eta(s), \partial U\big) = O(1).$$

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, ∂U and $\omega_V(t)$ to V are all close to each other. In particular,

$$d_V\big(\omega_V(t), \partial U\big) = O(1).$$

Next assume $V \subsetneq U$ and $d_U(\omega_U(t), \partial V)$ is large, or equivalently $d_U(\eta(s), \partial V)$ is large. In $\mathcal{C}(U)$, one of the segments $[x, \omega(t)]_U$ or $[\omega(t), y]_U$ is far from ∂V. Assume first that $[x_U, \omega_U(t)]_U$ is far from ∂V. This implies that

$$d_V\big(x_U, \omega_U(t)\big) = O(1).$$

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

$$d_V\big(\eta_U(s'), \omega_V(t)\big) = O(1)$$

and since $\eta_U(s')$ comes between x_U and $\omega_U(t)$ we also have

$$d_V\big(\eta_U(s'), \omega_U(t)\big) = O(1).$$

From the triangle inequality we get

$$d_V\big(\omega_U(t), \omega_V(t)\big) = O(1),$$

which proves Consistency in this case.

The remaining case is when $[\omega(t), y]_U$ is far from ∂V. Consistency implies

$$d_V\big(\omega_U(t), \eta_V(s)\big) = O(1),$$

where $\omega_U(t) = \eta_U(s)$. We also have

$$d_V\big(y_U, \omega_U(t)\big) = O(1) \quad \text{and so} \quad d_V\big(y_U, \eta_V(s)\big) = O(1).$$

Now $\omega_V(t)$ is further along the geodesic $[x,y]_V$ than $\eta_V(s)$ and so

$$d_V\big(y_U, \omega_V(t)\big) = O(1),$$

and applying the triangle inequality once more we have

$$d_V\big(\omega_V(t), \omega_U(t)\big) = O(1),$$

the desired inequality for Consistency.

Finally, we need to prove

$$d_{\mathcal{X}}\big(\gamma(t), \omega(t)\big) = O(\epsilon R).$$

Let $\omega_j(t)$ be the projection $\omega(t)$ to $\mathcal{C}(\mathbf{Y}^j)$. 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 $\mathcal{C}(\mathbf{Y}^j)$ 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

$$d_{\mathcal{C}(\mathbf{Y}^j)}\big(x_j, \gamma_j(s)\big) + d_{\mathcal{C}(\mathbf{Y}^j)}\big(\gamma_j(s), \gamma_j(t)\big) + d_{\mathcal{C}(\mathbf{Y}^j)}\big(\gamma_j(t), y_j\big) \\ \leq d_{\mathcal{C}(\mathbf{Y}^j)}(x_j, y_j) + O(\epsilon R).$$

But $\mathcal{C}(\mathbf{Y}^j)$ is Gromov hyperbolic, so projection to $[x_j, y_j]$ is distance decreasing. Furthermore $d_{\mathcal{C}(\mathbf{Y}^j)}\big(\eta_j, [x_j, y_j]\big) = O(1)$. Hence

$$d_{\mathcal{C}(\mathbf{Y}^j)}\big(x_j, \eta_j(s)\big) + d_{\mathcal{C}(\mathbf{Y}^j)}\big(\eta_j(s), \eta_j(t)\big) + d_{\mathcal{C}(\mathbf{Y}^j)}\big(\eta_j(t), y_j\big) \\ \leq d_{\mathcal{C}(\mathbf{Y}^j)}(x_j, y_j) + O(\epsilon R).$$

But we know $\eta_j(t)$ comes before $\eta_j(s)$. Therefore,

$$d_{\mathcal{C}(\mathbf{Y}^j)}\big(\eta_j(s), \eta_j(t)\big) = O(\epsilon R),$$

and hence,

$$d_{\mathcal{C}(\mathbf{Y}^j)}\big(\omega_j(t), \eta_j(t)\big) = O(\epsilon R).$$

Since this is true for every j, we also have

$$d_{\mathcal{X}}\big(\omega(t), \eta(t)\big) = O(\epsilon R).$$

This finishes the proof. □

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 $\mathcal{X}$ 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 $\mathcal{G}(x,y)$ in terms of projection distances $d_{\mathcal{X}(W)}(x,y)$.

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

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

$$d_{\mathcal{X}}\big(z, \gamma(t)\big) = \sum_i d_{\mathcal{X}(W_i)}\big(z, \gamma(t)\big)$$

$$\text{(Lemma 4.8)} \qquad \leq \sum_i d_{\mathcal{X}(W_i)}(x,y) = O(D).$$

Assume S is connected and let $\gamma(t)$ be a point with

$$d_{\mathcal{C}(S)}\big(z, \gamma(t)\big) = O(1).$$

Such a point exists because $d_S(z, [x,y]_S) = O(1)$ by the definition $\mathcal{G}(x,y)$ and γ connects x to y.

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

$$d_{\mathcal{X}_{\alpha_i}}\big(z, \gamma(t)\big) \overset{*}{\prec} d_{\mathcal{X}_{\alpha_i}}(x_1, x_2) \leq D,$$

And by Lemma 3.4

$$d_{\mathcal{X}}\big(z, \gamma(t)\big) \overset{*}{\prec} \sum_i d_{\mathcal{X}_{\alpha_i}}\big(z, \gamma(t)\big) = O(D). \qquad \square$$

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 $x, y \in \mathcal{X}(S)$, thresholds $T_1 \geq T_0 > 0$, and a subsurface W, a collection Ω of subsurfaces $W_i \subsetneq W$ is an antichain in W for x and y if

- if $W_i \in \Omega$ then $d_{\mathcal{C}(W_i)}(x,y) \geq T_0$.
- 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 Ω 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 γ is moving in some $\mathcal{X}(W)$ but it is not close to $\mathcal{X}_{\partial W}$ 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 $\gamma\colon [a,b] \to \mathcal{X}$ is a preferred path connecting x to y such that for all t,*

$$d_{\mathcal{X}}\big(\gamma(t), \mathcal{X}_{\partial W}\big) \geq M \quad \text{and} \quad d_{\mathcal{X}(W)}(x,y) \geq D,$$

then there is a proper subsurface $V \subsetneq W$ such that

$$d_{\mathcal{X}(V)}(x,y) \overset{*}{\asymp} d_{\mathcal{X}(W)}(x,y).$$

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

Case 1. Assume $d_W(x,y) \le M_2$, for some uniform constant $M_2 = O(M_0)$. By the distance formula

$$d_{\mathcal{X}(W)}(x,y) \stackrel{*}{\asymp} \sum_{U \in \mathcal{W}_{M_2}(x,y)} d_U(x,y).$$

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

$$d_{\mathcal{X}(W)}(x,y) \stackrel{*}{\asymp} \sum_{U \in \mathcal{V}_{M_2}(x,y)} d_U(x,y),$$

where $\mathcal{V}_{M_2}(x,y)$ is the collection of subsurfaces $U \subset V$ where $d_U(x,y) \ge M_2$. That is

$$d_{\mathcal{X}(V)}(x,y) \stackrel{*}{\asymp} d_{\mathcal{X}(W)}(x,y).$$

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 $\mathcal{C}(W)$ 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 $d_{\mathcal{X}}(z,w)$ 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 $W \subsetneq U$ and $d_U(z, \partial W)$ large. Since the quasi-geodesic $[x,y]_U$ in $\mathcal{C}(U)$ passes through z_U, either $[x,z]_U$ or $[z,y]_U$ stays far from ∂W. Then, by Theorem 3.1, either $d_W(x,z) \le M_0$ or $d_W(y,z) \le M_0$ which is a contradiction.

Consider now the possibility that $U \pitchfork W$ 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 $d_W(x,z)$ 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 $\gamma \colon [a,b] \to \mathcal{X}$ connecting x to y where*

$$D = d_{\mathcal{X}(W)}(x,y) \ge D_0,$$

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

- $d_{\mathcal{X}(V)}\big(\gamma(c), \gamma(d)\big) \stackrel{*}{\asymp} D$
- *for $t \in [c, d]$, $d_{\mathcal{X}}\big(\gamma(t), \mathcal{X}_{\partial V}\big) = O(1)$.*

Proof. We use induction on complexity of subsurfaces. If γ does not come within M of $\mathcal{X}_{\partial W}$ then Lemma 5.4 applies. Let $W' \subsetneq W$ be a subsurface such that

$$d_{\mathcal{X}(W')}(x, y) \stackrel{*}{\asymp} D.$$

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 γ does in fact come within M of $\mathcal{X}_{\partial W}$, and let z_1 and z_2 be points in γ marking the first and the last times γ is within M of $\mathcal{X}_{\partial W}$. We have, either

$$d_{\mathcal{X}(W)}(x, z_1) \stackrel{*}{\asymp} D, \quad d_{\mathcal{X}(W)}(z_1, z_2) \stackrel{*}{\asymp} D, \quad \text{or} \quad d_{\mathcal{X}(W)}(z_2, y) \stackrel{*}{\asymp} D.$$

If $d_{\mathcal{X}(S_1)}(z_1, z_2) \stackrel{*}{\asymp} D$, then we are done after taking $V = W$. In the other two cases, (say $d_{\mathcal{X}_{\partial W}}(x, z_1) \stackrel{*}{\asymp} D$) the path connecting x and z_1 does not come close to $\mathcal{X}_{\partial W}$ but travels large distance in $\mathcal{X}_W$. Hence, we again can apply Lemma 5.4 and induction to finish the proof. □

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

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

$$d_{\mathcal{X}(W)}(y_i, y_{i+1}) = L/5.$$

For a constant C_0, we say γ 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 γ 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 γ is a preferred path joining $x, y \in \mathcal{N}_{O(1)}(\mathcal{X}_{\partial W})$ and*

$$D := d_{\mathcal{X}(W)}(x, y) \geq D_0$$

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

- $d_{\mathcal{X}(V)}\big(\gamma(c), \gamma(d)\big) \overset{*}{\succ} D$.
- $\gamma|_{[c,d]}$ *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 γ 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 $V \subsetneq W$ where

$$d_{\mathcal{X}(V)}(y_i, y_{i+1}) \overset{*}{\succ} D.$$

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 γ 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 $\mathcal{X}_{\partial W}$.

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

$$d_{\mathcal{X}}(x, x') \leq c_0 D, \quad \textit{and} \quad d_{\mathcal{X}}(y, y') \leq c_0 D.$$

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

$$d_{\mathcal{X}}(z', z) \leq c_1 D \qquad \textit{then} \qquad d_{\mathcal{X}}(z', \mathcal{X}_{\partial W}) = O(1).$$

Proof. Let $x = y_0, \ldots, y_5 = y$ be as in Definition 5.6. Consider the projection $\pi_W(\gamma)$ to $\mathcal{C}(W)$. We have $d_W(x, y_1) \overset{*}{\succ} C_0$. Even though the distance in $\mathcal{X}$ between y_1 and x is $D/5$, there may be points much closer to x in $\mathcal{X}$ 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 $z_1 \in \mathcal{X}$,

$$d_W(z_1, \beta_1) \leq \delta \quad \Longrightarrow \quad d_{\mathcal{X}}(x, z_1) \overset{*}{\succ} D.$$

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

Proof of Claim. We know $d_{\mathcal{X}(W)}(x, y_1) \overset{+}{\succ} D/5$ and that the distance in $\mathcal{X}(W)$ is the sum of subsurface projections to subsurfaces in $\mathcal{W}_T(x, y_1)$. The boundary of any such subsurface is near a curve in $[x, y_1]_W$.

Let β_1 be the curve along $[x, y]_W$ that is 2δ 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 $U \in \mathcal{W}_T(x, y_1)$ has uniformly bounded diameter M_0 (every curve in $[\beta_1, y]_W$ intersects U). In fact, for $z_1 \in \mathcal{X}$, where $d_W(z_1, \beta_1) \le \delta$, every curve in the geodesic $[z_1, y]_W$ also intersects U. Hence,

$$d_U(x, z_1) \overset{+}{\succ} d_U(x, y_1).$$

Therefore, $d_{\mathcal{X}(W)}(x, z_1) \overset{*}{\succ} d_{\mathcal{X}(W)}(x, y_1) = D/5$. $\square$

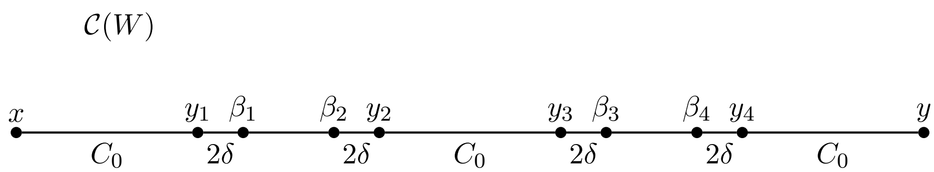

FIGURE 3. The projections of $y_0, \dots, y_5$ in $\mathcal{C}(W)$ are at least C_0 apart. For $i = 1, \dots 4$, the β_i in $[x, y]_W$ is 2δ away from the projection of y_i in the indicated direction.

Similarly we find

- a curve β_4 near the projection of y_4 so that, for any z_4 whose projection to W is δ close to β_4, we have
$$d_{\mathcal{X}(W)}(z_4, y) \overset{*}{\succ} D.$$
- curves β_2 and β_3, near the shadows of y_2 and y_3 respectively, so that, for any $z_2, z_3 \in X$,

$$d_W(z_2, \beta_2) \le \delta, \quad d_W(z_3, \beta_3) \le \delta \quad \Longrightarrow \quad d_{\mathcal{X}}(z_2, z_3) \overset{*}{\succ} D. \tag{13}$$

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 δ, it follows from the hyperbolicity of the curve complex that the δ–neighborhood of the path $[x', y']_W$ has to contain $[\beta_2, \beta_3]_W$. In particular,

$$d_W(x', y') \overset{+}{\succ} C_0.$$

This means the path $[x', y']$ passes near $X_{\partial W}$. Let z_1' and z_4' be the first and the last time the path $[x', y']$ is near $\mathcal{X}_{\partial W}$. A δ–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 $[z_1', z_4']$ whose projection to $\mathcal{C}(W)$ is δ–close to β_2 and β_3 respectively. Thus, by Equation (13),

$$d_{\mathcal{X}}(z_2', z_3') \overset{*}{\succ} D \quad \Longrightarrow \quad d_{\mathcal{X}}(z_1', z_4') \overset{*}{\succ} D.$$

This is the desired subsegment of γ'. To see the last assertion of the theorem, note that if $d_{\mathcal{X}}(z', z) \le c_1 D$ 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*

$$d_S(x, x') = O(\delta) \quad \textit{and} \quad d_S(y, y') = O(\delta).$$

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

$$d_S(z', x) \geq C_1 \quad \textit{and} \quad d_S(z', y) \geq C_1.$$

Then

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

Proof. We need to show that, for any subsurface U,

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

If $U = S$ then we know from the hyperbolicity of $\mathcal{C}(S)$ 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 $\mathcal{C}(S)$ are a bounded distance apart in the $d_{\mathcal{C}(S)}$. 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 ∂U is not close to z'_S in $\mathcal{C}(S)$ then, without loss of generality, we can assume ∂U is far from $[z', y]_S$ (otherwise ∂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 ∂U is near z'_S in $\mathcal{C}(S)$, 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, ϵ_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 $f\colon B \to \mathcal{X}$ is an ϵR–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:*

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

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

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

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

$$d_{\mathcal{X}(W)}(x,y) \geq \sqrt[3]{\epsilon^2}\, R.$$

We now claim the first conclusion of the Theorem holds. That is, there is a subbox B' of size

$$R' \stackrel{*}{\asymp} d_{\mathcal{X}(W)}(x,y) \stackrel{*}{\succ} \sqrt[3]{\epsilon^2} R$$

and a curve α such that $f(B')$ lies within an $O(\sqrt[3]{\epsilon} R')$–neighborhood of $\mathcal{X}_\alpha$. 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) $d_{\mathcal{X}(V)}\big(\gamma(d_1), \gamma(d_2)\big) \stackrel{*}{\succ} R'$,
(2) $\gamma\big([d_1, d_2]\big)$ stays in a $O(1)$ neighborhood of $\mathcal{X}_{\partial V}$, 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

$$d_{\mathcal{X}}(f(s_i), x_i) = O(\epsilon R).$$

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

$$\begin{aligned} d_{\mathcal{X}}(x_i, y_i) &\leq d_{\mathcal{X}}\big(x_i, f(s_i)\big) + d_{\mathcal{X}}\big(f(s_i), y_i\big) \\ &\leq O(\epsilon R) + KcR' \leq c_0 R' \end{aligned}$$

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

$$d_{\mathcal{X}}(f(p), x) = O(\epsilon R).$$

By Proposition 5.8, there is c_1 so that all $y \in \mathcal{G}(y_1, y_2)$ that satisfy

$$d_{\mathcal{X}}(y, x) \leq c_1 R'$$

also satisfy

$$d_{\mathcal{X}}(y, \mathcal{X}_{\partial V}) = O(1).$$

For c small, a box of size R' centered at p is mapped under f within distance $O(\epsilon R) = O(\sqrt[3]{\epsilon}\, R')$ 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 $W \subsetneq S$ that

$$d_{\mathcal{X}(W)}(x,y) \stackrel{*}{\prec} \sqrt[3]{\epsilon^2}\, R \tag{14}$$

This and Lemma 3.4 imply that γ makes C_0'– steady progress in the entire surface S, for some $C_0' \stackrel{*}{\succ} 1/\sqrt[3]{\epsilon}$. For ϵ 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

$$R' = d_{\mathcal{X}}(x,y) \stackrel{*}{\succ} \sqrt[3]{\epsilon}\, R.$$

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 γ' joining x', y' within $O(\epsilon R)$ of $f\big([p', q']\big)$. Since the map f is quasi-Lipschitz it follows, for c sufficiently small, that

$$d_{\mathcal{X}}(x', y') \overset{*}{\succ} R'.$$

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

$$d_{\mathcal{X}}(f(\hat{p}), z') = O(\epsilon R).$$

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 γ.

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 γ. 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. $\square$

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 $\boldsymbol{\alpha}$ be a (possibly empty) curve system. For every connected component W of $S \setminus \boldsymbol{\alpha}$ (including annuli if $\mathcal{X}$ is not the $\mathcal{P}(S)$), let $\omega_W \colon I_W \to \mathcal{X}(W)$ be a preferred path. Consider the box $B = \prod_W I_W \subset \mathbb{R}^n$, where n is the number of components of $S \setminus \boldsymbol{\alpha}$. Then

$$F \colon B \to \mathcal{X}_{\boldsymbol{\alpha}} = \prod_W \mathcal{X}(W) \qquad \text{where} \qquad F = \prod_W \omega_W,$$

is a quasi-isometric embedding because each γ_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 ϵ_0 and R_0, let*

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

Assume $f \colon B \to \mathcal{X}$ is an ϵ_ξ–efficient map where B is a box of size R_ξ 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 (7)). If $\xi = 0$, then

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

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 $f_i \colon B \to \mathcal{X}(S_i)$ be the projection of f to $\mathcal{X}(S_i)$. by Lemma 2.2, f_i is still ϵ_0–efficient. Applying, Proposition 2.11 to $f_1 \colon B \to \mathcal{X}(S_1)$, 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 $f_2 \colon B_1 \to \mathcal{X}(S_2)$ 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 ξ is non-zero. Apply Theorem 6.1. If the second condition holds, we are done for

$$R' \overset{*}{\succ} \sqrt[3]{\epsilon_\xi}\, R_\xi = \frac{\sqrt[3]{\epsilon_\xi}}{\epsilon_\xi}\, R_0 \geq R_0,$$

since a preferred path is itself is a standard flat and $\sqrt[3]{\epsilon_\xi} = \epsilon_{\xi-1}^2 \leq \epsilon_0$.

Otherwise, we have a box $\bar{B}$ of size

$$\bar{R} \overset{*}{\succ} \sqrt[3]{\epsilon_\xi^2}\, R_\xi \geq R_{\xi-1}.$$

that maps to a $O(\sqrt[3]{\epsilon_\xi}\, \bar{R})$–neighborhood of $\mathcal{X}_\alpha$ for some curve α. The map f is $\epsilon_{\xi-1}^2$–efficient because, reducing the size of the box by some factor (in this case $\sqrt[3]{\epsilon_\xi^{-2}}$) only makes the efficiency constant increase by the same factor and

$$\frac{\epsilon_\xi}{\sqrt[3]{\epsilon_\xi^2}} = \sqrt[3]{\epsilon_\xi} = \epsilon_{\xi-1}^2.$$

Composing f with the closest point projection map to $\mathcal{X}_\alpha$ and using the fact that $\mathcal{X}_\alpha$ is quasi-isometric to $\mathcal{C}(\alpha) \times \mathcal{X}(S \setminus \alpha)$, we obtain a map

$$\bar{f} \colon \bar{B} \to \mathcal{C}(\alpha) \times \mathcal{X}(S \setminus \alpha).$$

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 $\mathcal{X}(S \setminus \alpha)$ 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 $\mathcal{X}(S \setminus \alpha)$. Now projecting to the first factor and applying Proposition 2.11 ($\mathcal{C}(\alpha)$ 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 $\mathcal{C}(\alpha)$. That is, $f(B')$ stays in a $O(\epsilon_0 R_0)$–neighborhood of a standard flat in $\mathcal{X}$. This finishes the proof. □

Proof of Theorem A. Let R_1 be large enough that the box B is guaranteed (by Theorem 2.4) to have a sub-box B_ξ of size at least R_ξ where the restriction of f to B_ξ is ϵ_ξ–efficient. Note that we are not using the full force of the theorem; we need only one efficient sub-box. Apply Theorem 7.2 to $f\colon B_\xi \to \mathcal{X}$ to obtain the theorem. □

Proof of Theorem B. 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 a 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 $F\colon \mathbb{R}^m \to \mathcal{X}$. Taking a composition of f, the closest point projection to the image of F and then F^{-1}, we obtain a map $G\colon B' \to \mathbb{R}^m$ with the property that, for $p, q \in B'$,

$$d_{\mathbb{R}^m}\big(G(p), G(q)\big) \overset{*}{\asymp} d_{\mathbb{R}^n}(p,q) \pm O(\epsilon R'). \tag{15}$$

We show that there is no such a map if n is bigger than $\mathrm{rank}_{top}(\mathcal{X}) \geq m$. 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 (15). 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, then the product of these geodesics is a quasi-isometric image of R^n. This fails when W_i is an annulus and $\mathcal{X}$ is the augmented marking space (a horoball does not contain a bi-infinite geodesic). In this case, we choose a ϵ_0–thick point X in $\mathcal{T}(S)$ and a pants decomposition of curves of length at most some fixed B. The point $x \in \mathcal{X}$ associated to X is uniformly close to the product region 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 $\mathbb{R}^n$. □

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.

[Bo] B. Bowditch The Coarse geometry of the Teichmüller metric, preprint

[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.